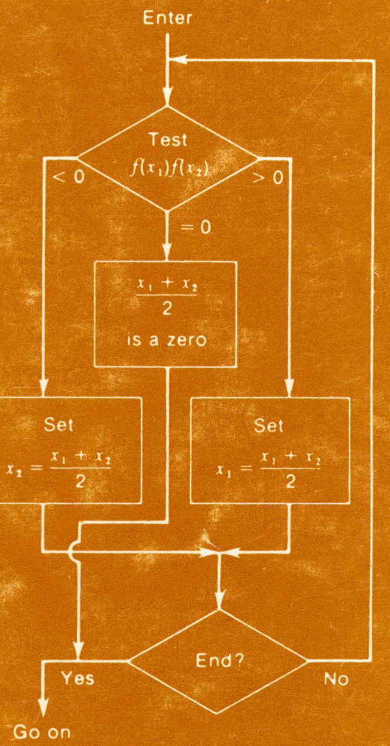
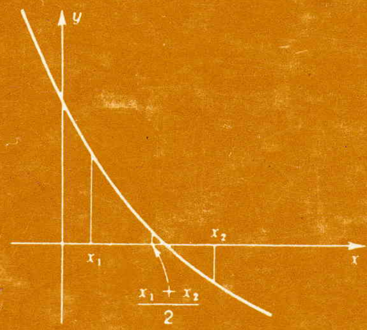
محاسبات عددی





ترجمه و تألیف:

دکتر اسماعیل بابلیان

استادیار دانشگاه تربیت معلم

دکتر خسرو مالک نژاد

استادیار دانشگاه علم و صنعت ایران

بسم الله الرحمن الرحیم

نام کتاب : محاسبات عددی
ترجمه و تالیف : دکتر اسماعیل بابلیان ــ دکتر خسرو مالک‌نژاد
ناشر: دکتر اسماعیل بابلیان ــ دکتر خسرو مالک‌نژاد
تیراژ: ۳۷۵۰ جلد
نوبت چاپ : اول
تاریخ چاپ : تیرماه ۱۳۶۶
امور لیتوگرافی : کاوه نو تلفن ۳۱۵۸۷۲
چاپ : هما
همکار فنی چاپ : خدمات چاپ پچواک تلفن ۶۷۴۲۸۴
قیمت : ۸۰۰ ریال

محاسبات عددی

ترجمه و تألیف:

دکتر اسماعیل بابلیان

استادیار دانشگاه تربیت معلم

دکتر خسرو مالک نژاد

استادیار دانشگاه علم و صنعت ایران

بنام خدا

پیشگفتار

امروزه آنالیز عددی ، یا محاسبات عددی ، در نزد متخصصین آن هم علم است و هم هنر. اما این تعبیر برای غیرمتخصصین اغلب بدفهمی‌هائی دربردارد . بعضی فکر می‌کنند که علم و هنر نامیدن آنالیز عددی تنها بخاطر پنهان کردن این حقیقت است که آنالیز عددی رشته‌ای دقیق نیست که سزاوار علم نامیدن باشد . بعضی دیگر ایراد می‌کنند که چون در ریاضیات آنالیز ، بمعنی کلاسیک آن ، قابل کاربرد در کارهای عددی نیست بنابراین "آنالیز عددی "اسمی بی‌مسمی است .

اما هیچیک از این تعبیرها صحیح نیست . پهلوی هم قرار دادن علم وهنر بخاطر اصلی موسوم به" اصل نامعلومی " است که غالبا "در حل مسائل رخ می دهد ، یعنی اینکه:

"تعیین بهترین راه جهت حل یک مسئله ممکن است محتاج جواب این مسئله یا متکی به دانستن خواصی از عوامل تشکیل‌دهندهٔ آن باشد ، که به طریق نظری و عملی قابل بدست آوردن نیستند . "

به عنوان یک هنر متخصص آنالیز عددی با انتخاب بهترین روش، واجرای مناسب آن ، برای حل یک مسئله خاص مواجه است که مستلزم توسعه هر چه بیشتر تجربه و بینش می‌باشد .

مثال ساده‌ای مطالب بالا را روشن می‌کند : دو روش معمول برای تخمین

$$\int_a^b f(x)\,dx$$

عبارتند از "قاعدهٔ ذوزنقه‌ای "و "قاعدهٔ سیمپسون " .

خطای قاعدهٔ اول $-\dfrac{(b-a)^3 f''(\eta_1)}{12n^2}$ و خطای قاعدهٔ دوم $-\dfrac{(b-a)^5 f^{(4)}(\eta_2)}{180n^4}$

است ، که در آنها n تعداد زیر فاصله‌های $[a,b]$ است و η_1 و η_2 دو نقطه از $[a,b]$ هستند .

از آنجا که آنالیز عددی یک علم است ، خطاهای فوق را ، که هنگام استفاده از این قاعده‌ها متحمل می‌شویم ، ارائه داده است . اما پاسخ به این سوءال که " از کدام قاعده استفاده کنیم " به هنر بودن این رشته مربوط است . متخصصین این رشته با استفاده از بینش، تجربه و آگاهی خود نسبت به پارامترهای مسئله مناسبترین روش را برای حل آن انتخاب می‌کنند .

به عنوان یک علم ، متخصص آنالیز عددی با فرایندهائی که مسائل ریاضی را بوسیلهٔ اعمال حسابی حل می‌کنند سر و کار دارد . بعضی اوقات این مستلزم وضع الگوریتمهائی برای حل مسائل جدید یا بهتر کردن الگوریتمهای موجود است .

به‌عنوان یک هنر، متخصص آنالیز عددی با انتخاب بهترین روش، و اجرای مناسب آن ، برای حل یک مسئله خاص مواجه است که مستلزم توسعه هرچه بیشتر تجربه و بینش می‌باشد .

پیشنیاز لازم برای این کتاب ریاضیات لازم برای دوره‌های لیسانس فنی و مهندسی و آگاهی از یک زبان کامپیوتری سطح بالا می‌باشد . مواد این کتاب برای بیشتر دانشجویان دوره‌های لیسانس فنی و علوم مناسب بوده و از آنجا که آنالیز عددی هنوز جزء بعضی از دوره‌های لیسانس نیست لذا ، این کتاب برای بعضی از دوره‌های بالاتر از لیسانس نیز مناسب است . قضایای ریاضی ، در مواردی که نتیجه آنها مهم بوده ولی اثباتشان دشوار است ، بدون برهان ذکر شده‌اند . برهانها ، تنها در مواردی که خاصیت یا اصل مهمی را روشن کنند آورده شده‌اند .

نظر به اینکه دانشجویان ضمن انجام پروژه‌های مربوط به آنالیز عددی نیاز به برنامه‌های کامپیوتری ، جهت اجرای روشهای محاسبات عددی ، دارند لذا ، در پایان مسائل حل شدهٔ هر فصل برنامه‌های کامپیوتر ضروری آورده شده است . زبان این برنامه‌ها ، باتوجه به امکانات مراکز کامپیوتری دانشگاهها ،"فرترن چهار" درنظر گرفته شده‌است .

این کتاب مجموعه وسیعی از روشهای عددی در سطحی مناسب با پیشنیاز آنها می‌باشد . بهمین دلیل در موارد لازم برای اطلاع بیشتر از روشها ، در هر زمینه به کتب مناسبی ارجاع داده شده است .

فصل مربوط به مقادیر ویژه از نقطه نظر پیشنیاز از مشکل‌ترین فصول کتاب است و بعضی از دانشجویان لازم خواهند دید که مقداری از نظریه را بدون اثبات بپذیرند .

به منظور حفظ پیوستگی ، بیشتر مثالهای حل شده همراه با مسائلی که دانشجویان باید حل کنند در پایان هر فصل گردآوری شده است . مثالها شامل مطالب اضافی ارزنده است که بدون مطالعه آنها از کتاب بهره کاملی نمی‌توان گرفت .

در پایان از خانم شهلا کهوائی برای صفحه‌آرائی و آقایان قدرت‌الله خوانساری، مسعود سیفی زارعی و محمد رضا نامدار فارغ‌التحصیلان دانشگاه علم و صنعت

ایران که در مراحل مختلف امور فنی چاپ این کتاب همکاری نموده‌اند ، کمال تشکر و قدردانی می‌شود .

اسماعیل بابلیان ــ خسرو مالک‌نژاد
خرداد ۱۳۶۶

فهرست

فهرست راهنما

فصل اول
محاسبات کامپیوتری

۱.۱ ■ مقدمه

قبل از شروع مطالعهٔ روشهای محاسباتی کامپیوتر، شاید بررسی این نکته‌که چرا در طول بیست سال گذشته کاربرد کامپیوتر در محاسبات علمی به طور چشمگیری افزایش یافته است، خالی از فایده نباشد. در گذشته، مشخص شده که مغز بشر قابلیت درک تکامل چشمگیر ریاضیات را دارد، بنابراین مزیت کامپیوتر در چیست؟

گرچه مغز انسان قادر است حوزهٔ وسیعی از افکار و معلومات را بهم ربط دهد، لکن هنگام انجام محاسبهٔ ساده‌ای نظیر $2654{\cdot}3 \times 19{\cdot}3281$ دستگاه ذهنی هم آهسته و هم ناقص عمل می‌کند. انجام محاسبهٔ فوق، برای فرد معمولی احتمالا" ۲ دقیقه طول خواهد کشید و پاسخ بدست آمده به احتمال زیاد اشتباه خواهد بود! در همان مدت زمان، جدیدترین کامپیوترها قادر به انجام 500 000 000 محاسبه نظیر محاسبه فوق می‌باشند، و تنها اشکال فنی دستگاه الکترونیکی می‌تواند خطائی معادل خطای بشری به بار آورد. شاید این امر که جواب حاصل تقریبا" همیشه دقیق نبوده تاسف‌آور باشد، اما از آنجائیکه عدم دقت محدود بوده و در حدود 10^{-12} می‌باشد، می‌توان جواب را قابل قبول دانست.

اهمیت این عدم دقت اجتناب‌ناپذیر، یکی از مباحثی است که در طول کتاب مستمرا" مطرح خواهد شد.

بنابراین یکی از مزیتهای کامپیوتر سرعت و قابل اطمینان بودن آن در انجام عملیات ساده است. مزیت عمده دیگر، مربوط به اثر روشهای محاسبه‌ای جدید می‌گردد. اگرچه دانش ریاضیات مستمرا" درحال‌توسعه است، معهذا مسائل زیادی وجود دارد که بکمک آنالیز ریاضی قابل حل نیستند. برای مثال، اگرموضوع‌انتگرال‌گیری

را درنظر بگیریم توابع زیادی هستند که هیچ فرمولی برای انتگرال آنها موجود نیست. تنها راه حل چنین مسائلی به کاربردن تقریب عددی است که متضمن استفاده از همان نوع حساب تکراری است که برای کامپیوتر ایده‌آل است. در حقیقت، سرعت کامپیوترهای جدید کاربرد روشهای تقریب عددی را در مسائل فیزیکی ممکن ساخته و در غیر اینصورت حل چنین مسائلی، به لحاظ حجم محاسبات عددی مربوطه، کاملا" غیرممکن می‌شد. این نوع مسائل، مثلا"، زمانی که مادهٔ متصلی نظیر یک مایع یا میله صلب، به منظور امکان تحلیل ریاضی، به واحدهای کوچکتری تقسیم می‌شود، پیش می‌آیند. هم‌اکنون محاسبهٔ بعضی مسائل با سریع‌ترین ماشینهای جدید چندین ساعت بطول می‌انجامد و هنوز نیاز به تقریبی باز هم بهتر که زمان محاسبه را بطور عمده کاهش دهد وجود دارد.

بدست آوردن ایده‌هایی در مورد نحوهٔ انجام محاسبات در کامپیوتر خالی از فایده نبوده و در نتیجه مسائل و محدودیتهای محاسبهٔ عددی که در آینده مورد بحث واقع می‌شوند به سهولت قابل ارزیابی هستند. درک اینکه چرا کامپیوترهای نوین براساس الکتریسته کار می‌کنند، آسان است. سرعت انتقال الکتریسته، گرچه آنی نیست اما مسلما" نزدیک به آن حدی است که به لحاظ فیزیکی امکان‌پذیر است. خاصیت دو حالتی وصل یا قطع علائم الکتریکی برای نمایش اعداد در مبنای دو کافی است و تکامل ترانزیستورها نیز، دستگاههای اتصال، ذخیره‌ساز و زمان‌سنج را به‌وجود آورده است، که همگی با سرعتهای بالا کار می‌کنند. در حقیقت سرعت این اجزاء امروزه به حدی است که زمان لازم برای عبور الکتریسیته از سیمها عاملی جدی در تاخیر به حساب می‌آید و کوششهای اخیر طراحی، به حذف سیم‌ها از طریق کوچک نمودن اجزاء و یکپارچه ساختن اجزاء از قبل ساخته شده معطوف شده است. بدیهی است که مطالعه مدارهای الکترونیکی، که امکان ساخت کامپیوترهای نوین را فراهم می‌سازد، خارج از حیطهٔ این کتاب بوده ولی در اینجا بعضی از مراحل را به منظور نشان دادن چگونگی محاسبات در داخل کامپیوتر، مورد ملاحظه قرار می‌دهیم. ابتدا، باید روش محاسبه به طور دقیق تعریف شود و دنباله‌ای از دستورالعملها که برای کامپیوتر ضروری است ارائه گردد. در روزهای اولیهٔ کامپیوتر، این کار بسیار خسته کننده بود زیرا لازم بود که مراحل ریاضی را به مراحل بسیط منطقی که برای کامپیوتر قابل اجرا است، تقسیم و یا ترجمه نمود. برنامه‌نویسی کامپیوتر به کمک دستورات رمزی ماشین صورت می‌گرفت که برای دستورالعملهای اساسی کامپیوتر قابل اجرا است.

این کار اضافی ترجمه ازنظر اهل علم و مهندسین، مانعی در استفاده موثر از ماشین بود و بنابراین کار ترجمه به خود ماشین واگذار شد. زبانهای پیشرفته نظیر الگول یا فرترن، حدود سال ۱۹۶۰ وضع شدند، این زبانها، به نحوی طراحی شده

بودند که به طریق معمول دانشمندان برای مشخص نمودن یک روش، نزدیک بوده و در عین حال هر دستورالعمل قابل ترجمه به یک دسته از رمزهای اساسی ماشینی باشد. بدین‌ترتیب زمانی که یک برنامه به زبان سطح ـ بالا به داخل کامپیوتر خوانده می‌شود، برنامه تفسیری موسوم به همگردان دستورالعملهای زبان پیشرفته را به صورت رمز ماشین تبدیل می‌کند. همچنین مسائل معینی، تا زمانیکه عددی را به‌عنوان معرف یک کامپیوتر در نظر بگیریم وجود دارد، شبیه نمادهای دستگاه دهدهی مشتمل بر 10 رقم مختلف، که مناسبترین شکل برای این منظور نبوده، چون دستگاههای الکترونیکی نیازمند به ایجاد و تشخیص 10 شکل مختلف خواهند بود. در اینصورت پیچیدگی دستگاهها احتمالاً آنها را گران و قابل اشتباه خواهد کرد. بهرحال، ساختن دستگاههای خیلی ساده که مجهز به دو علامت روشن و خاموش باشد ممکن می‌باشد. با این اسبابها ذخیره کردن یک دنباله دوتائی از اطلاعات و ترکیب اینها از راههای مختلف به منظور ارائه عملیاتی نظیر جمع و تفریق و غیره امکان‌پذیر خواهد بود. خوشبختانه، یک دستگاه عددی که تنها نیازمند به دو نشان اساسی 0 و 1 باشد وجود دارد که در آن اعداد به صورت یک رشته متشکل از صفر و یک ارائه‌می‌شوند، بنابراین دستگاه عددی (دستگاه دوتائی) برای معرفی اعداد داخل کامپیوتر به کار می‌رود. این بدان معنی است که روند انتقالی زمانیکه اعدادبه داخل کامپیوتر خوانده یا خارج می‌شوند ضروری است و این روند نسبتاً ساده می‌باشد.

روابط بین اعداد دوتائی و اعشاری در بخش ۱.۲ بحث می‌گردد. مسئله دیگری که با آن مواجه می‌شویم زمانی است که سعی در ذخیره هر دو دستگاه اعداد دوتائی و دهدهی در یک کامپیوتر داشته باشیم. اگراعداد در دستگاه اعشاری مانند $\frac{1}{3}$ یا e در نظر گرفته شوند، معرفی این اعداد با تعداد متناهی رقم اعشاری ممکن نخواهد بود، به روشی مشابه اعداد زیادی وجود دارند که نمی‌توان به وسیله تعدادی متناهی رقم دوتائی محدود نشان داد.

از آنجائیکه یک کامپیوتر تنها دارای تعدادی محدود از ارقام دوتائی برای معرفی یک عدد می‌باشد (معمولاً بین ۲۴ تا ۶۰ رقم) آنگاه هر عدد که به ارقام بیشتری نیاز داشته باشد اجباراً با تقریب نشان داده می‌شود. حال به محاسباتیکه با دستورات ساده ماشین امکان‌پذیر است می‌پردازیم. بدون هیچ مشکلی مسئله جمع و تفریق مانند ترکیبات علائم برقی انجام می‌گیرد. بهر حال، اگر کسی ضرب دو عدد را درنظر بگیرد، یک تجزیه و تحلیل مختصر برای درک اینکه چطور چنین محاسبه‌ای انجام می‌گیرد ضروری است. محققاً این عمل مشتمل بر سه مرحله است. ضرب یک رقم تنها با به کار بردن جدول ضرب، تغییر ستونها برای یکان، دهگان، صدگان و غیره و

جمع‌بندی مجموعه‌های جزئی مختلف. اولی در دستگاه دوتائی ساده است چون فقط دارای سه جزء در جدول‌ضرب می‌باشد، دومی یک عمل معمولی‌کد ماشین بوده بنابراین روند تبدیل به یک سری انتقال مکانها و جمعها منجر می‌گردد. عمل ضرب به‌طور فراوان انجام می‌گیرد و این عمل به کمک قسمت مخصوصی از دستگاه الکترونیکی انجام گرفته تا اینکه یک سری دستورات کدی ماشین را در هر لحظه ایجاد نماید. بحثهای مشابه برای تقسیم موجود بوده که سری انتقالها و تفریقها انجام می‌گیرد.

هر یک از سایر عملیات نظیر گرفتن لگاریتمها، توانها و یا تشکیل توابع مثلثاتی باید به طریقی تبدیل به یکی از طرق بالا گردد، آنگاه یک مجموعه از دستورات کدی ماشین برای ایجاد چنین توابع بکار می‌رود.

روش محاسبه برای چنین توابع، شامل روشهای تقریب بوده بنابراین خطاهای بیشتری با کاربرد توابعی از نوع فوق تولید می‌گردد. برای کامل کردن این مطلب، در کامپیوتر قسمتهای کنترل وجود داشته که ورودی، خروجی و حرکات بین ذخیره و محاسبه دستگاه را نظارت می‌کند. کار یکی از این دستگاههای کنترل عبارتست‌از به‌کار بردن دستورات ذخیره شده در کامپیوتر به‌منظور تبدیل به دستورات کدی ماشین که در محاسبات موردنظر، موثر می‌باشند.

با توجه به بحث فوق روشن می‌گردد که آشنائی با دستگاه دوتائی و خطاهای کامپیوتر پیش نیاز اساسی برای محاسبه عددی بوده و این مطلب در بقیهٔ این فصل بررسی می‌شود.

۱.۲ ■ حساب دوتائی

۱.۲.۱ ■ دستگاه دوتائی

بهترین مقدمه برای دستگاه دوتائی، که اساس عمل کامپیوتر است، توجه دقیق به دستگاه آشناتر دهدهی می‌باشد. این دستگاه مبتنی براساس ستون‌هائی که نمایشگر توان‌های 10 است می‌باشد، طوریکه 2304 به‌معنی

$$2 \times 10^3 + 3 \times 10^2 + 0 \times 10^1 + 4 \times 10^0$$

یا در حالت کلی، عدد دهدهی

$$a_n a_{n-1} \dots a_1 a_0 \qquad 0 \leqslant a_i \leqslant 9 \qquad (۱.۱\ a)$$

نمایشگر عدد دهدهی

$$a_n \times 10^n + a_{n-1} \times 10^{n-1} + \cdots + a_1 \times 10 + a_0 \qquad (۱.۱\ b)$$

می‌باشد، البته داشتن پایه‌ای به‌جز 10 امکان‌پذیر می‌باشد. بنابراین، اگر دستگاهی عددی در پایه r داشته باشیم، عدد

$$a_n a_{n-1} \ldots a_1 a_0 \qquad 0 \leqslant a_i \leqslant r-1 \tag{1.2a}$$

نمایشگر عدد دهدهی

$$a_n \times r^n + a_{n-1} \times r^{n-1} + \cdots + a_1 \times r + a_0 \tag{1.2b}$$

می‌باشد.

با انتخاب پایه 2 ، ملاحظه می‌شود که 10110 در دستگاه دهدهی به‌صورت

$$1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 22$$

خواهد بود. برجسته‌ترین فرق میان دو دستگاه تعداد ارقامی است که برای نوشتن یک عدد مورد نیاز است. در دستگاه دهدهی ارقام $0,\ldots,9$ را به کار برده و هر عدد بزرگتر از 9 به واحدهائی در پایه 10 شکسته می‌شود، در دستگاه دوتائی ارقام کمتر از پایه 2 ارقام 0 و 1 بوده و هر عدد بزرگتر به‌طور طبیعی شامل مضاربی از 2 درستونهای دیگر است.

بدین‌ترتیب، چند عدد اولیه به‌صورت زیر نوشته می‌شوند.

$$\begin{array}{ll} 0_{10} = 0_2 & 3_{10} = 11_2 \\ 1_{10} = 1_2 & 4_{10} = 100_2 \\ 2_{10} = 10_2 & 5_{10} = 101_2 \end{array}$$

و غیره.

علامت اندیس پائین برای مشخص کردن پایه هنگامی به‌کار می‌رود که امکان شبهه وجود داشته باشد. خواننده‌ایکه نسبت به دستگاه دوتائی ناآشناست باید به مثالهای ساده در پایان این فصل مراجعه نماید (مثالهای ۱۰۲ ، ۱۰۳ ، ۱۰۴ ، ۱۰۵) جزئیات عملیات حسابی در دستگاه دوتائی را می‌توان در کتب متعدد، برای مثال:

(Conte (1965)) یافت، لیکن در اینجا توجه به تبدیل اعداد از دستگاه دوتائی به دهدهی و بالعکس محدود خواهد بود.

۱.۲.۲ ■ تبدیل دوتائی به دهدهی

برای مثال عدد 101100 را در دستگاه دوتائی درنظر می‌گیریم، یک راه برای تبدیل این عدد به عدد دهدهی از دادن مقدار عددی معادل به هر ستون و جمع کردن جزءهاست، یعنی $44 = 4 + 8 + 32$ معهذا رویهٔ ساده‌تری وجود دارد که خیلی تکراری بوده، و بنابراین برای کاربرد کامپیوتر مناسب می‌باشد. فرض می‌کنیم ارقام $a_r\ (r = 0, 1, \ldots)$ بوده و از سمت راست به چپ شماره‌گذاری شده باشند از a_k ، اولین رقم سمت چپ شروع و به صورت زیر محاسبه می‌کنیم

$$b_k = a_k \tag{1.3a}$$

$$b_r = 2b_{r+1} + a_r, \qquad r = k-1, \qquad k-2, \ldots, 0 \tag{1.3b}$$

برای مثال فوق خواهیم داشت :

$$\begin{aligned} b_5 &= 1 \\ b_4 &= 1.2+0=2 \\ b_3 &= 2.2+1=5 \\ b_2 &= 2.5+1=11 \\ b_1 &= 2.11+0=22 \\ b_0 &= 2.22+0=44 \end{aligned} \tag{1.4}$$

(مثال ۱۰۱ را نیز ببینید)

۱.۲.۳ ■ تبدیل دهدهی به دوتائی

در اینجا روش مستقیم عبارتست‌از کاهش بزرگترین توان صحیح ممکن ۲ از عدد دهدهی و ثبت آن در ستون مناسب ، و ادامه این روند تا حصول باقیمانده صفر . برای مثال ، 100_{10} دارای 64 و 32 و 4 بوده که از آن کم می‌گردد ، بنابراین عـــدد دوتائـــی مناسب 1100100 خواهد گردید .

روش دیگر ، عدد دهدهی b_0 را اختیار کرده و به 2 تقسیم نموده و باقیمانده‌هارا به‌صورت a_1, a_0 ، و الاآخر ، ذخیره می‌کند .

$$b_r = \frac{b_{r-1} - a_{r-1}}{2} \tag{1.5}$$

که درآن a_{r-1} برابر 0 است اگر b_{r-1} زوج باشد و برابر 1 است اگر b_{r-1} فرد باشد

	b_r	a_r
2	44	rem.0
2	22	rem.0
2	11	rem.1
2	5	rem.1
2	2	rem.0
2	1	rem.1
	0	

(1.6)

توجه : اولین رقم دوتائی از سمت چپ در انتهای‌ستون باقیمانده‌ها ظاهر می‌گردد به‌طوری که عدد دوتائی عبارتست‌از 101100 (مثال۱۰۲ را نیز ببینید) .

۱.۳ ■ نمایش اعداد در کامپیوتر

در یک کامپیوتر اساسا" دو نوع نمایش از اعداد وجود دارد . یکی نمایش اعداد صحیح و دیگری اعداد حقیقی ، که اعمال حسابی نظیر هرکدام متفاوت است . عملیات

اعداد صحیح ساده‌تر و سریعتر بوده ، بنابراین تا حد امکان بایستی به‌کار روند . باوجود این ، در اغلب محاسبات علمی نه‌تنها کسرها مورد احتیاج قرار می‌گیرند بلکه محاسبات از اعداد بینهایت کوچک گرفته تا اعداد بینهایت بزرگ تغییر خواهد کرد . در این حالت حساب ممیز سیار به‌کار می‌رود ، که هر دو مسئله فوق را دربر می‌گیرد .

نمایش عدد دوتائی به‌شکل ممیز سیار خیلی شبیه شکل توانی عدد دهدهی است ، به‌طور مثال $0{\cdot}2342 \times 10^{-3}$ دارای مانتیس $0{\cdot}2342$ و توان -3 می‌باشد . صورت ممیز سیار عدد ماشینی دارای شکل

$$\bar{x} = a \times 2^{b} \tag{1.7}$$

است که در آن

$$a = \pm \sum_{r=1}^{t} d_r 2^{-r} \quad \text{و} \quad \tfrac{1}{2} \leqslant |a| < 1$$

و معمولا "

$$|b| \leqslant M$$

d_r ارقام دوتائی 0 یا 1 بوده که به هر توانی از 2 متناظر می‌شوند . مقادیر t و M متغیراند ، لیکن مقادیر عمومی آنها $t = 37$ و $M = 2^{8} = 256$ است . این نمایش تقریبا" ۱۱ رقم اعشار با معنی را در دامنه بین اعداد 10^{-77} و 10^{+77} بدست‌می‌دهد .

علامت ــ در x زمانی به‌کار می‌رود که لازم به تاکید باشد چون ممکن است عدد ماشینی مقدار واقعی آن را ندهد . عدد مورد نظر ممکن است ، حداکثر به اندازه یک واحد که در اولین ستون مانتیس a ظاهر نمی‌گردد خطا داشته باشد ، و مقدار خطا ، برابر 2^{-t-1} است . بنابراین خطای ممکن در $\bar{x}$ ، مساوی 2^{b-t-1} بوده که مربوط به گرد کردن ماشین در منظور کردن توان b می‌باشد . عموما" ، حساب اعداد صحیح ، به‌اعمالی نظیر شمارش اختصاص داده می‌شود ، چون اگر اعداد صحیح به‌طور متوالی ضرب شوند تعداد مکانهای دوتائی برای نمایش حاصلضرب به طور نامحدود افزوده می‌گردد . در حالت حساب ممیز سیار ، اگر عدد بزرگی تولید گردد ، برای معرفی اندازهٔ درست توان میزان می‌شود ، و اگر تعداد ارقام بزرگتر از t باشد ، آنگاه ارقام اضافی حذف شده و کمترین رقم با معنی در پایان باقی می‌ماند . حد دستگاه ، زمانی رخ می‌دهد که اندازهٔ توان از حد مجاز بیشتر شود ، این حالتی است که به سرریز مشهور می‌باشد . اگر ارقام اضافی حذف شوند ، خطای ناشی از گرد کردن اضافی تولید می‌گردد .

بنابراین ، ملاحظه می‌گردد که خطاها در کامپیوتر هم در نمایش اعداد به صورت ممیز سیار و هم در سایر محاسبات بعدی تولید می‌گردد . از آنجائیکه حساب ممیز سیار برای محاسبات علمی واقعا" اساسی است ، لازم است طرقی برای می‌نیمم کردن این

خطاها داشته باشیم .

راه بدیهی برای انجام این منظور عبارتست از افزایش تعداد ارقام با معنی ، که معادل افزایش اندازهٔ t می‌باشد . نظیر چنین طرحی برای تمام محاسبات به کار نمی‌رود زیرا این عمل نیازمند زمان بیش از حد مجاز برای کامپیوتر است ، و برای بعضی از کامپیوترها برنامه‌های مخصوصی بایدنوشته شود . معهذا ، در دستگاههای محاسبه مدرن کار کردن با طول دوبرابر خیلی معمول است . در این حالت ، با افزایش دقت ، خطاهای تولید شده از 2^{b-2t-1} بیشتر نخواهد بود . بیشتر مسائل علمی منجر به محاسبات مشتمل بر اعداد مختلط می‌گردد و راه معرفی مستقیم برای اینها در کامپیوتر وجود ندارد .

امکانات مربوط به عدد مختلط در بعضی از زبانهای پیشرفته نظیر فرترن درنظر گرفته شده است ، اما باید درنظر داشت که این عمل با نوشتن قسمتهایی از برنامه انجام می‌گیرد که وقت‌گیرمی‌باشد .

۱.۴ ■ خطاها دراعمال حسابی

در این قسمت بحث محدود به حساب ممیز سیار بوده و عمدتا" خطاهای مربوط به قوانین چهارگانه حساب‌درنظر گرفته خواهند شد . باید خطاهای ناشی از نمایش متناهی در ماشین موسوم به خطاهای گرد کردن ، و همچنین طریقی که اعمال حسابی مختلف این خطاها را ایجاد می‌کند بررسی نمائیم .

ابتدا ، باید به‌خاطر آورد که نمایش عدد x در کامپیوتر عدد $\bar{x}$ با ماکزیمم خطای ممکن $\pm 2^{b-t-1}$ خواهد بود . حال جمع دو عدد ماشینی را درنظر می‌گیریم .

$$\bar{x}_1 = a_1 . 2^{b_1} \qquad \bar{x}_2 = a_2 . 2^{b_2}, \qquad b_1 > b_2 \tag{1.8}$$

که در آن جواب $\bar{x}_3 = a_3 . 2^{b_3}$ می‌باشد . عمل جمع ابتدا با انتقال ارقام a_2 به‌راست‌به اندازه $b_1 - b_2$ انجام می‌گیرد ، به طوری که ستونهای متناظر زیر هم قرار می‌گیرند این عمل معمولا" با به کار بردن ثباتهائی با طول دو برابر انجام می‌گیرد ، بنابراین هیچ رقمی در روند انتقال ازبین نمی‌رود ، سپس دو عدد به هم افزوده می‌شوند . اگر مجموع نتیجه شدهٔ a_3 ، دارای مقداری خارج از حدود $\frac{1}{2} \leqslant |a_3| < 1$ باشد ، آنگاه‌ارقام a_3 به‌راست یا چپ حرکت داده شده و اندازهٔ توان b_3 برحسب‌آن میزان میشود . سپس این عدد با صرفنظر کردن کمترین ارقام با معنی به طول ساده کاهش می‌یابد . بنابراین ماکزیمم خطای تولید شده در روند گردکردن 2^{b_3-t-1} خواهد شد . در محاسبه خطا ، اغلب راحت‌تر خواهد بود که خطای نسبی را ، که بصورت نسبت خطای مطلق به‌نتیجه محاسبه تعریف می‌شود ، درنظر گرفت . یک تعریف اساسی از خطای نسبی عبارتست از

حاصل تقسیم خطای مطلق به‌مقدار واقعی . چون ، نتیجه محاسبه شده تنها عددیست که به آن دست‌رسی داریم لذا تعریف تقریبا" متفاوتی از خطای نسبی که در اینجا به‌کار رفت خیلی مفید می‌باشد . بدین ترتیب ، در اینحالت خطای نسبی عبارتست‌از :
$2^{b_3-t-1}/(a_3 . 2^{b_3})$ از آنجائی‌که کمترین‌مقدار ممکن a_3 عبارتست‌از $\frac{1}{2}$ ، ماکزیمم مقدار این کمیت $\pm 2^{-t}$ خواهد شد . بحثهای مشابهی برای سایر سه روندقابل‌اجرا است ، بنابراین پاسخ ماشینی برای یک محاسبه ، شامل اعداد ماشینی درست ، عبارتست‌از :

$$\bar{A} = A(1+e) \tag{1.9a}$$

که در آن

$$|e| \leqslant 2^{-t} \tag{1.9b}$$

و A پاسخ درست است که معمولا" درطول ساده به‌طور دقیق نمی‌تواند معرفی گردد . علامت ε برای خطای مطلق و e برای خطای نسبی به کار خواهد رفت . عموما" ، محاسبات در اثر خطاهای قبلی ، محاسبه‌ای یا مربوط به نمایش اعداد در ماشین ، درست نمی‌باشد . بنابراین درنظر گرفتن اینکه چطور خطاها در روندهای حسابی‌منتشر می‌گردند امری ضروری است .

فرض می‌کنیم

$$x = \bar{x} + \varepsilon_x \qquad y = \bar{y} + \varepsilon_y \tag{1.10}$$

آنگاه

$$x + y = \bar{x} + \bar{y} + \varepsilon_x + \varepsilon_y \tag{1.11}$$

بنابراین خطای جدید مساوی مجموع تک تک خطاها بوده و مقدار ماکزیمم قدر مطلق آن عبارتست از $|\varepsilon_x| + |\varepsilon_y|$
خطای نسبی دارای مقدار

$$\text{خطای نسبی} = \frac{\bar{x}}{\bar{x}+\bar{y}} \cdot \frac{\varepsilon_x}{\bar{x}} + \frac{\bar{y}}{\bar{x}+\bar{y}} \cdot \frac{\varepsilon_y}{\bar{y}} \tag{1.12}$$

بوده که یک مجموع ساده از مضارب محاسبه شدهٔ خطاهای نسبی اعداد اصلی می‌باشد .

بدست آوردن تفاضل خیلی شبیه جمع بوده با علائم منفی در سمت راست . این نکته حائز اهمیت است که در تفریق زمانیکه $\bar{x}$ و $\bar{y}$ خیلی به هم نزدیک باشند ، دو مضرب از خطاهای نسبی می‌توانند خیلی بزرگ گردند و این عمل به سادگی ، خطاهای قابل توجهی تولید می‌کند . (به مثال ۱.۵ مراجعه شود)

برای ضرب داریم

$$\begin{aligned} xy &= (\bar{x} + \varepsilon_x)(\bar{y} + \varepsilon_y) \\ &= \bar{x}\bar{y} + \varepsilon_x . \bar{y} + \varepsilon_y . \bar{x} + \varepsilon_x . \varepsilon_y \end{aligned} \tag{1.13}$$

معمولا" ، خطاهای ε_x و ε_y کوچک‌اند، بنابراین حاصلضرب $\varepsilon_x.\varepsilon_y$ می‌تواند قابل اغماض باشد، لذا خطای مطلق برابر

$$\varepsilon_x\bar{y}+\varepsilon_y\bar{x} \tag{1.14}$$

شده و خطای نسبی برابر

$$\frac{\varepsilon_x}{\bar{x}}+\frac{\varepsilon_y}{\bar{y}} \tag{1.15}$$

می‌گردد.

خطای نسبی در ضرب تقریبا" مساوی جمع خطاهای نسبی اعداد اصلی می‌گردد. نتیجه مشابهی برای تقسیم پیدا خواهیم کرد، اما در این حالت تفاضل دو خطای نسبی اصلی راداریم. برهان آن برای نتیجه تقسیم به کمک بسط نسبت، به‌کار بردن سری دو جمله‌ای و حذف جملاتی که شامل حاصلضرب خطاها باشد، بدست می‌آید.

$$\begin{aligned}\frac{x}{y}&=\frac{\bar{x}+\varepsilon_x}{\bar{y}+\varepsilon_y}\\&=\frac{\bar{x}}{\bar{y}}\left(1+\frac{\varepsilon_x}{\bar{x}}\right)\left(1+\frac{\varepsilon_y}{\bar{y}}\right)^{-1}\\&=\frac{\bar{x}}{\bar{y}}\left(1+\frac{\varepsilon_x}{\bar{x}}\right)\left(1-\frac{\varepsilon_y}{\bar{y}}+\left(\frac{\varepsilon_y}{\bar{y}}\right)^2\cdots\right)\\&\approx\frac{\bar{x}}{\bar{y}}+\frac{\varepsilon_x}{\bar{y}}-\varepsilon_y\cdot\frac{\bar{x}}{\bar{y}^2}\end{aligned} \tag{1.16}$$

بدین ترتیب، خطای مطلق برابر

$$\frac{\varepsilon_x}{\bar{y}}-\frac{\varepsilon_y.\bar{x}}{\bar{y}^2} \tag{1.17}$$

و خطای نسبی مساوی

$$\frac{\varepsilon_x}{\bar{x}}-\frac{\varepsilon_y}{\bar{y}} \tag{1.18}$$

می‌شود. بالاخره، خطای حاصل از به کار بردن توابعی از x مانند ریشه دوم، سینوس و غیره را درنظر می‌گیریم. اگر $x=\bar{x}+\varepsilon_x$ آنگاه با به کار بردن بسط سری تیلر داریم:

$$f(x)=f(\bar{x})+\varepsilon_x f'(\bar{x})+\frac{\varepsilon_x^2}{2!}f''(\bar{x}+\theta\varepsilon_x),\qquad (0\leqslant\theta\leqslant 1) \tag{1.19}$$

بدین‌ترتیب، خطای مطلق تقریبا" برابر $\varepsilon_x f'(\bar{x})$ بوده، و خطای نسبی، مشروط بر آنکه مشتق مرتبه دوم تابع بزرگ نباشد، برابر $\varepsilon_x f'(\bar{x})/f(\bar{x})$ می‌گردد. سپس خطاهای منتشره از عملیات مختلف با خطاهای ناشی از عمل گرد کردن که قبلا" بحث شده می‌توانند ترکیب گردند. محاسبهٔ دقیق کل خطای انباشته در هر گام به سه دلیل خیلی مشکل می‌باشد. اول اینکه بعضی از تخمین‌های خطا مبتنی بر حذف

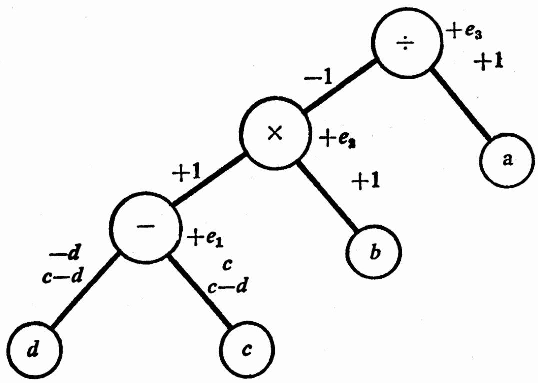
حاصلضربهای خطاها است، که در صورت افزایش خطاها قابل اغماض نخواهد بود. دوم، اندازهٔ خطا ممکن است بستگی‌به ترتیب محاسبات انجام شده داشته باشد. (به مثال ۱۰۶ مراجعه شود) سوم آنکه، حالتهایی وجود دارند که در آنها اندازهٔ خطا بستگی به اندازهٔ نتایج داشته که در این حالات بدبینانه‌ترین نتیجه پذیرفته می‌شود.

بدین‌ترتیب، کران بالای محاسبه شده برای خطاها ممکن است، تقریبا" در تمام حالات، خیلی بیشتر از خطاهای واقعی باشد. همچنین تخمین‌های آماری امید خطا امکان‌پذیر است، لیکن این عمل شامل روندهای مفصلتری است که در اینجا قابل پیگیری نیست. خواننده‌ایکه علاقمند به مطالب بیشتر در زمینه آنالیز خطا باشد می‌تواند به کتابهای:

McCracken Dorn (1964), Hamming (1962), Ralston (1965)

مراجعه کند. دو کتاب آخری شامل مطالبی پیرامون تجزیه و تحلیل آماری خطای گرد کردن می‌باشد.

بعنوان مثال برای محاسبهٔ خطا، مثال $a/[b(c-d)]$ را درنظر گرفته و خواننده را به روش نموداری با ارائه این روند آشنا خواهیم نمود، مثلا" در



شکل (۱۰۱) روند ترسیمی $a/[b(c-d)]$

McCracken (1964) و Dorn نمودار فوق، خطای نسبی را در صورت محاسبه بامرتبه $r_1=c-d$ و $r_2=b\times r_1$ و $r_3=a/r_2$ نشان می‌دهد. این نمودار با شروع از پایین پس از ضرب خطای مجتمع هر دایره، در عامل کنار خط منتهی بر دایره بعدی، به‌کار گرفته شده، سپس به خطای ناشی از گرد کردن افزوده می‌گردد. این جمع آنگاه به نام خطای نسبی انباشتگی به کار رفته و با عامل ضربی مناسب در گام بعدی به صورت ذیل ترکیب می‌گردد:

$$
\begin{aligned}
E_1 &= \frac{\bar{c}}{\bar{c}-\bar{d}} e_c - \frac{d}{\bar{c}-\bar{d}} e_d + e_1 \\
E_2 &= E_1 + e_b + e_2 \\
E_3 &= -E_2 + e_a + e_3 \\
&= \frac{-\bar{c}}{\bar{c}-\bar{d}} e_c + \frac{d}{\bar{c}-\bar{d}} e_d - e_b + e_a \\
&\quad - e_1 - e_2 + e_3
\end{aligned}
\tag{1.20}
$$

که در آن

$$|e_3| \leqslant 2^{-t} \quad ، \quad |e_2| \leqslant 2^{-t} \quad ، \quad |e_1| \leqslant 2^{-t} \tag{1.21}$$

توجه: علامت ــ مربوط به عدد ماشینی بوده که گرد شده‌ای از جواب واقعی است .

۱.۵ ■ خطاهای روشهای محاسبه ای

همانطوریکه خطاهای مربوط به روندهای حسابی را در کامپیوتر بررسی می‌کنیم ، باید خطاهای ایجاد شده در اثر کاربرد روشهای تقریبی را نیز درنظر بگیریم . یک مثال ساده عبارتست از اینکه، بخواهیم توابعی نظیر sine ، log ،ریشه دوم و غیره را بکار بریم ، که معمولا" به صورتهای جدول‌بندی شده وجود دارند . به سادگی می‌توان دریافت که ذخیرهٔ این چنین مجموعه از اعداد در کامپیوتر امکان‌پذیر نبوده بنابراین به‌شکل دیگری باید درنظر گرفته شوند .

خوشبختانه روشهای متعددی برای نمایش این توابع به کمک یک‌سری همگرا موجود است ، که به وسیلهٔ چند قانون ساده تعریف‌پذیر بوده و به سادگی قابل برنامه‌نویسی نیز هستند . تعدادی از جملات سری به هم دیگر اضافه می‌شوند تا اینکه خطای صرفنظر شده بطور قابل قبولی کوچک گردد . چنین خطائی که ناشی از نمایش متناهی از یک روند نامتناهی است ، خطای برشی نامیده می‌شود .

سری تیلر توابع برای نشان دادن کاربرد سری همگرا به کار می‌رود ، هرچند که این سری ویژه برای تولید مقادیر تابع در کامپیوتر مناسب نمی‌باشد . شکل کلی سری تیلر با باقیمانده عبارتست‌از :

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \cdots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x) + R_n \tag{1.22}$$

که در آن

$$R_n = \frac{h^n}{n!} f^{(n)}(x+\theta h), \qquad 0 \leqslant \theta \leqslant 1 \tag{1.23}$$

این رابطه در فاصله $(x, x+h)$ مشروط بر آنکه تابع پیوسته و دارای مشتقات پیوسته تا

مرتبهٔ n ام باشد برقرار می‌گردد . درصورتی که بتوان نشان داد $\lim_{n\to\infty} R_n=0$. سری فوق همگرا می‌گردد .

در محاسبهٔ کامپیوتری جملات یک به یک محاسبه شده تا آنکه R_n خیلی کوچک گردد ، بدین‌معنی که ازنظر اندازه با میزان گرد شدهٔ کامپیوتر قابل مقایسه گردد . عمل اضافی دیگری که منجر به پیشرفتی گردد نمی‌توان انجام داد. سری زیر را درنظر می‌گیریم .

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots$$

با قراردادن

$$A_0=1 \quad , \quad A_r=\frac{x\,.\,A_{r-1}}{r} \qquad r=1,2,\ldots \tag{1.24}$$

آنگاه مجموعهای

$$S_0=A_0 \qquad S_r=S_{r-1}+A_r, \qquad r=1,2,\ldots \tag{1.25}$$

مقادیر تقریبی e^x را می‌دهند .

بدین‌ترتیب مجموعه‌ای ساده از دستورات می‌تواند در ماشین ذخیره شود تا e^x مطلوب محاسبه گردد . در زبانهای سطح بالا این قسمتهای برنامه در دستگاه کامپیوتری درنظر گرفته شده، اما از آنجائیکه اینها دستورات برنامه‌نویسی شده هستند آهسته‌تر از امکانات مخابره‌ای نظیر جمع و ضرب انجام می‌گیرند . مطالعهٔ همگرائی سری خارج از بحث این‌کتاب بوده، اما خواننده باید تشخیص دهد که سری e^x به‌ازای تمام مقادیر متناهی x همگرا بوده ، و همچنین سریهائی وجود دارند که مجموع قسمت‌ها زمانیکه n افزایش یابد بدون کران افزایش پیدا می‌کنند . سری $(1-x)^{-1}$ مثالی از این نوع می‌باشد . برای مقادیر معین x داریم :

$$(1-x)^{-1}=1+x+x^2+x^3+\ldots \tag{1.26}$$

اگر مقدار $x=1\cdot1$ را در بسط بکار بریم داریم .

$$1+1\cdot1+1\cdot21+1\cdot331+\ldots \tag{1.27}$$

و می‌توان مشاهده نمود که هر یک از جملات بزرگتر از ۱ بوده و بنابراین مجموع به‌طور نامعین افزایش می‌یابد .

تکنیک تقریبی دیگری که برای مسائل مختلف محاسباتی به کار می‌رود عبارتست‌از روش تکراری . این تکنیک روندی پی در پی است که یک تقریب به جواب را وارد محاسبه می‌کند ، و از این امر استفاده نموده و تقریب بهتری را تولید می‌کند . زمانیکه دنبالهٔ نتایج درحال همگرائی است ، مرحله‌ای می‌رسد که تفاضل این مقادیر قابل توجه نبوده ،

و در این مرحله، محاسبه متوقف می‌گردد. البته، این نکته شایان توجه است که در یک شیوهٔ تکراری زمانیکه نتایج به یک مقدار ثابت همگرا نمی‌گردد، بایستی با آزمونی این روند متوقف گردد. تعریف دقیق چنین آزمونی فوق‌العاده مشکل می‌باشد. روش تکراری را می‌توان با ساده‌ترین شکل خود، برای حل یک معادلهٔ یک مجهولی به کار گرفت، مشروط بر آنکه معادله را بتوان به صورت $x=g(x)$ نوشت، که در آن نامساوی $|g'(x)|<1$ در ناحیهٔ تکرار برقرار باشد آنگاه این تکرار با انتخاب مقدار اولیه x_0 پیش رفته و محاسبه می‌گردد.

$$x_1=g(x_0)$$

این روند سپس تکرار می‌شود.

$$x_{r+1}=g(x_r),\qquad r=1,2,\ldots \tag{1.28}$$

تا اینکه همگرا گردد.

اهمیت شرط مشتق به سادگی با مثال $x+4\sqrt{(x)}-8=0$ که دارای جوابهای $x=29{\cdot}8552$ و $2{\cdot}1432$ است نشان داده می‌شود. ابتدا معادلهٔ تکراری را درنظر

$$x=8-4\sqrt{x} \tag{1.29}$$

گرفته و با این تقریب $x_0=1$ شروع می‌کنیم. جدول مقادیر تکراری به صورت زیر خواهد شد.

$$\begin{aligned} x_0&=1\\ x_1&=8-4\times 1=4\\ x_2&=8-4\times 2=0\\ x_3&=8-4\times 0=8\\ x_4&=8-4\times 2{\cdot}828=-3{\cdot}312 \end{aligned} \tag{1.30}$$

و همگرائی امکان‌پذیر نیست زیرا،

$$g'(x)=\frac{-4}{2\sqrt{x}}$$

دارای مقادیر زیر است:

(1.31)

x	0	1	4
$g'(x)$	$-\infty$	-2	-1

به هرحال، اگر طرح تکرار به شکل زیر نوشته شود،

$$\sqrt{x}=\tfrac{1}{4}(8-x) \quad \text{یا} \quad x=\tfrac{1}{16}(8-x)^2 \tag{1.32}$$

خواهیم داشت $g'(x)=\frac{1}{16}(-16+2x)$ ، که در حوزهٔ موردنظر $|g'(x)|<1$ یعنی

x	0	1	2	4
$g'(x)$	-1	$-\frac{7}{8}$	$-\frac{3}{4}$	$-\frac{1}{2}$

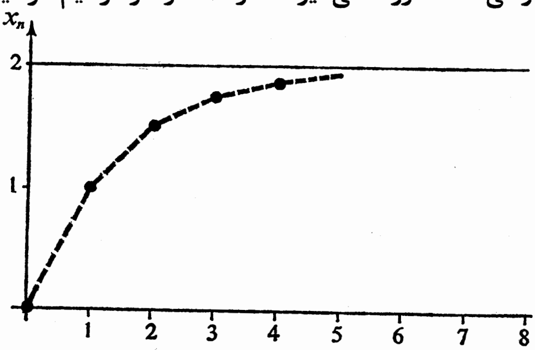
(1.33) جدول مقادیر عبارتست از

$$\begin{array}{ll} x_0=1 & x_6=1{\cdot}9625 \\ x_1=3{\cdot}0625 & x_7=2{\cdot}2782 \\ x_2=1{\cdot}5236 & x_8=2{\cdot}0461 \\ x_3=2{\cdot}6214 & x_9=2{\cdot}2155 \\ x_4=1{\cdot}8080 & x_{10}=2{\cdot}0912 \\ x_5=2{\cdot}3963 & x_{11}=2{\cdot}1821 \end{array} \tag{1.34}$$

مشاهده می‌شود، اگرچه سرعت همگرائی کند است اما این مقادیر همگرا می‌باشند. مقدار $g'(x)$ دراین نقطه $0{\cdot}7321-$ بوده و از آنجائیکه این مقدار نسبتا " به $1-$ نزدیک است، لذا سرعت همگرائی کند می‌باشد.

۱.۶ ■ شتاب همگرائی

در روندهای گام به گام، مانند روند ذکر شدهٔ فوق، ذکر روشهائی که‌به روند سرعت بخشند دارای اهمیت است. از آنجائیکه‌این ایده‌ها در انواع مسائل مختلف به کار رفته، لذا در این جا اصول اساسی سه روش شتاب دهنده مورد بحث قرار گرفته و دربخشهای بعدی برای مثالهای خاص به کار خواهند رفت. مقدمتا ً، یک طرح تکراری ساده را به صورت $x_{n+1}=Kx_n+B$ در نظر گرفته، و برای مثال فرض می‌کنیم $K=\frac{1}{2}$ و $B=1$. مقادیر متوالی تکرارها به منظور نشان دادنِ همگرائی روند به مقدار $x=2$ و اینکه همگرائی کند صورت می‌گیرد، توسط نمودار ترسیم گردیده است.



شکل ۱.۲ تکرار برای $x_{n+1}=Kx_n+B$

۱.۶.۱ ■ روش فوق - تخفیف

از نمودار فوق می‌توان دید که در هرگام فاصلهٔ جواب واقعی از مقدار تکراری نصف می‌گردد. اگر در هر مرحله گامی بزرگتر به سمت ۲ برداشته شود آنگاه روند را سریعتر خواهد نمود. این موضوع، انگیزه‌ای است که در پس روش فوق ــ تخفیف قرار داشته، و مطابق معادلات زیر عمل می‌کند. یک مقدار میانی به صورت زیرتشکیل شده

$$\tilde{x}_{n+1} = Kx_n + B \tag{1.35}$$

و سپس یک مقدار تصحیح شده با اختیار نمودن مقدار قدیمی و با افزودن مضربی از نمو محاسبه شده تشکیل می‌یابد، یعنی

$$x_{n+1} = x_n + \omega(\tilde{x}_{n+1} - x_n) \tag{1.36}$$

که معادله زیر را می‌دهد.

$$\begin{aligned} x_{n+1} &= x_n + \omega(K-1)x_n + \omega B \\ &= [\omega K + 1 - \omega]x_n + \omega B \end{aligned} \tag{1.37}$$

مقدار ω بمنظور دست یابی به همگرائی سریعتر، تغییر می‌کند و در حالات معینی، مانند تکرار ماتریسی، محاسبه بهترین مقدار ω بطور تئوری ممکن می‌باشد. اگر $\omega > 1$ روش را فوق ــ تخفیف و درصورتیکه $0 < \omega < 1$ ، روش را تحت ــ تخفیف می‌گوئیم. روش تحت ــ تخفیف بطور گسترده به کار نمی‌رود اما، درحل دستگاه معادلات خطی روش فوق ــ تخفیف به ازای $1 < \omega < 2$ روش بسیار مفیدی می‌باشد. نتیجه را به سادگی در مثال عددی پیشین به ازای $\omega = \frac{3}{2}$ می‌توان دید. حال مقادیر تکراری عبارتنداز: $0, \frac{3}{2}, 1\frac{7}{8}, 1\frac{31}{32}$ بنابراین همگرائی در این حالت دوبار سریعتر از روش تکرار مستقیم می‌باشد. در حقیقت، در این مثال اگر $\omega = 2$ باشد همگرائی در یک‌مرحله‌رخ‌می‌دهد، اما این نتایج با یک مثال ساختگی بدست آمده‌اند.

۱.۶.۲ ■ روند ایتکن

رفتار فوق (که در آن تفاضل مقدار واقعی x^* و تکراری به یک نسبت ثابت کاهش یابد) به ندرت در عمل رخ می‌دهد، اما بعد از چند تکرار اولیه، در بعضی مسائل یک تقریب خیلی نزدیک باین نسبت را می‌دهد. با فرض اینکه خاصیت نسبت ثابت درست باشد، دستور ساده‌ای که به روند Δ^2 ایتکن شهرت دارد، امکان می‌دهد که جواب نهایی محاسبه گردد.

فرض می‌کنیم
$$x^* - x_{n+1} = A(x^* - x_n) \tag{1.38}$$

سپس

$$x^* - x_{n+2} = A(x^* - x_{n+1})$$

و با تقسیم طرفین رابطه زیر نتیجه می‌شود .

$$\frac{x^* - x_{n+1}}{x^* - x_{n+2}} = \frac{x^* - x_n}{x^* - x_{n+1}} \tag{1.39}$$

$$(x^*)^2 - 2x_{n+1} . x^* + x_{n+1}^2 = (x^*)^2 - (x_n + x_{n+2})x^* + x_n x_{n+2}$$

پس از حذف اولین جمله و مرتب کردن داریم

$$x^* = \frac{-x_{n+1}^2 + x_n x_{n+2}}{x_n - 2x_{n+1} + x_{n+2}} = x_n - \frac{(x_{n+1} - x_n)^2}{x_n - 2x_{n+1} + x_{n+2}}$$

$$= x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} \tag{1.40}$$

(آخرین جمله نماد تفاضل متناهی را به کار می‌گیرد که بیانگر ریشهٔ روند می‌باشد) . البته درصورت به کار بردن این روند برای مثال قبل جواب درست خواهد داد ، چون از قانون کسر دقیقا" تبعیت می‌کند .

فرض می‌کنیم $x_n = 0$ ، $x_{n+1} = 1$ و $x_{n+2} = 1\frac{1}{2}$ باشد .

آنگاه $\Delta x_n = 1$ ، $\Delta^2 x_n = -\frac{1}{2}$ و $x^* = 2$. یک مثال عددی در مثال ۱.۷ داده شده است .

۱.۶.۳ ■ روش برونیابی ریچاردسُن

این روش برای روندهائی قابل اجراست که به طول گام h بستگی داشته ، و این مقدار را بتوان دلخواه اختیار نمود . بطور مثال ، در محاسبهٔ تقریبی انتگرال که فاصله به نوارهائی تقسیم شده و همچنین در انتگرال‌گیری گام به گام معادلات دیفرانسیل رخ می‌دهد . این روش را برای روندی که بتوان خطا را با یک سری توانی همگرانسبت به h بیان کرد می‌توان به‌کار برد ، بدین معنی که اگر مقدار واقعی و تقریبی را به ترتیب Y و Y_1 فرض کنیم داریم

$$Y = Y_1 + \sum_{j=k}^{\infty} a_j h_1^j \tag{1.41}$$

فرض می‌کنیم که دو محاسبه با مقادیر متفاوتی از h انجام گرفته و نتیجهٔ دو محاسبه بعدا" به منظور حذف جملهٔ پیشرو خطای ایجاد شده ترکیب شوند . معادلهٔ زیر (1.42)

$$Y = Y_2 + \sum_{j=k}^{\infty} a_j h_2^j$$

و (1.41) برای این منظور به کار می‌روند . پس از ضرب معادله (1.41) در h_2^k و

(1.42) در h_1^k و محاسبهٔ تفاضل آنها، معادله‌ای بدست می‌آید که ضریب $(h_1^k\ h_2)^k$ صفر می‌گردد.

$$Y[h_2^k - h_1^k] = h_2^k Y_1 - h_1^k Y_2 + \sum_{j=k+1}^{\infty} a_j [h_2^k h_1^j - h_1^k h_2^j]$$

برای سادگی فرض می‌کنیم h_2 مضربی از h_1 باشد، یعنی $h_2 = r.h_1$ ، آنگاه

$$Y = \frac{r^k Y_1 - Y_2}{r^k - 1} + \sum_{j=k+1}^{\infty} \frac{a_j.(r^k - r^j).h_1^j}{r^k - 1} \qquad (1.43)$$

تخمین جدید، اولین جملهٔ سمت راست رابطه فوق بوده و خطا با توانی‌بالاتراز h ظاهر می‌گردد. این روند را می‌توان با تکرار مقادیر بیشتر h ادامه داده جملات بیشتری از خطا را حذف نمود، اما در چنین حالت باید بطور دقیق همگرائی روند تحت بررسی قرار گرفته‌شود. این عمل برای انتگرال‌گیری رامبرگ انجام شده، که بطور متوالی فواصل نصف می‌گردند، یعنی $r = \frac{1}{2}$ و این روش برای دستهٔ بزرگی از توابع همگرا می‌باشد.

این روش در فصل هشت در مورد انتگرال‌گیری تقریبی، و در مثال ۱۰۸ که جدولی از اعداد را برای این روند بدست می‌دهد، بررسی شده است.

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ اعداد ذیل را از دستگاه دوتائی به دهدهی تبدیل نمائید

با به‌کار بردن ضرب داریم 10111 و 1010001

$$10111 = 16 + 4 + 2 + 1 = 23$$

$$\begin{array}{l}(((1 \times 2 + 0) \times 2 + 1) \times 2 \ \ + 1) \ \ \times 2 \ \ + 1 \ \ = 23 \\ \quad 2, \qquad 4, \ \ 5, \quad 10, \ \ 11, \quad 22, \ \ 23\end{array}$$

$$1010001 = 64 + 16 + 1 = 81$$

با به کار بردن ضرب داریم

$$\begin{array}{l}(((((1 \times 2 + 0) \times 2 + 1) \times 2 \ \ + 0) \ \ \times 2 \ \ + 0) \ \ \times 2 \ \ + 0) \ \ \times 2 \ \ + 1 \ \ = 81 \\ \quad 2, \ \ 2, \ \ 4, \ \ 5, \quad 10, \ \ 10, \quad 20, \ \ 20, \quad 40, \ \ 40, \quad 80, \ \ 81\end{array}$$

۲ــ اعداد زیر را از دستگاه دهدهی به دوتائی تبدیل نمائید: 28 و 53. باتفریق بزرگترین توان از ۲، در هر مرحله، داریم

$$28 = 16 + 8 + 4 = 11100$$

پس از تقسیم بر ۲ به طور متوالی داریم:

$$
\begin{array}{r|l}
2 & 28 \\ \hline
2 & 14 \quad r\ 0 \\ \hline
2 & 7 \quad r\ 0 \\ \hline
2 & 3 \quad r\ 1 \\ \hline
2 & 1 \quad r\ 1 \\ \hline
2 & 0 \quad r\ 1 \\ \hline
\end{array}
$$

11100 پاسخ

با تفریق بزرگترین توان ۲ ، در هر مرحله ، داریم :

$$53 = 32 + 16 + 4 + 1 = 110101$$

· پس از تقسیم بر ۲ متوالیا " داریم

$$
\begin{array}{r|l}
2 & 53 \\ \hline
2 & 26 \quad r\ 1 \\ \hline
2 & 13 \quad r\ 0 \\ \hline
2 & 6 \quad r\ 1 \\ \hline
2 & 3 \quad r\ 0 \\ \hline
2 & 1 \quad r\ 1 \\ \hline
2 & 0 \quad r\ 1 \\ \hline
\end{array}
$$

110101 پاسخ

۳ــ اعمال زیر را در حساب دوتائی انجام دهید

$13 + 9 = 22,\qquad 25 - 12 = 13,\qquad 5 \times 7 = 35,\qquad 28 \div 4 = 7$

$$
\begin{array}{r}
1101 \\
+\,1001 \\ \hline
10110
\end{array}
\qquad
\begin{array}{r}
11001 \\
-\,1100 \\ \hline
1101
\end{array}
\qquad
\begin{array}{r}
101 \\
\times\,111 \\ \hline
101 \\
101 \\
101 \\ \hline
100011
\end{array}
\qquad
\begin{array}{r}
111 \\
100\overline{)\,11100} \\
\underline{100} \\
110 \\
\underline{100} \\
100
\end{array}
$$

۴ ــ روش تفریق نمودن به وسیلهٔ متمم‌گیری بسیار مفید است و بطور گسترده‌ای به کار می‌رود. می‌توان آنرا با حساب دهدهی نشان داد. کاهش 273 معادل کم کردن $1000 - 727$ می‌باشد ، یعنی

$$
278 - 273 = \begin{array}{r}
278 \\
727 \\ \hline
1005
\end{array}
$$

5 = جواب

برای بدست آوردن متمم ابتدا عدد را از ۹۹۹ کاسته و سپس عدد ۱ را به یکان حاصل اضافه می‌کنیم . برای بدست آوردن تفاضل ابتدا عدد را با مکمل جمع کرده سپس از ستون هزارگان ۱ واحد کم می‌کنیم .

عمل مشابه‌در دستگاه دوتائی قابل اجرا است و متناظر با مدار الکترونیکی خیلی ساده‌ای می‌باشد . در دستگاه دوتائی عدد 3 را از 7 کم‌کنید .متمم عدد 11 عدد 101 خواهد بود . متمم را به 111 اضافه نمائید .

$$\begin{array}{r} 111 \\ \underline{101} \\ 1100 \end{array}$$

$100_2 = 4$ جواب

۵ ــ در صورتیکه محاسبه‌ای شامل تفاضل دو عدد تقریبا" مساوی باشد ، خطای قابل ملاحظه‌ای نتیجه می‌گردد . محاسبه $1/(0{\cdot}003\,146 - 0{\cdot}003\,130)$ را برای سه رقم اعشاری با معنی درنظر می‌گیریم .

$$\frac{1}{0{\cdot}315 \times 10^{-2} - 0{\cdot}313 \times 10^{-2}} = \frac{1}{0{\cdot}2 \times 10^{-4}} = 5{\cdot}00 \times 10^{4}$$

جواب درست عبارتست‌از:

$$\frac{1}{0{\cdot}16 \times 10^{-4}} = 6{\cdot}67 \times 10^{4}$$

وبنابراین خطائی درحدود 25% تولید می‌شود . این مسئله در مورد حل معادلهٔ درجه دوم هنگامی رخ می‌دهد که b^2 در مقایسه با $4ac$ خیلی بزرگ باشد . دومقدار b و $\sqrt{(b^2-4ac)}$ تقریبا" مساوی یکدیگر شده و در اینصورت یکی از ریشه‌ها شامل خطاهای بزرگ می‌گردد . برای مثال ، اگر b مثبت باشد $x = (-b - \sqrt{(b^2-4ac)})/2a$ یک ریشه خواهد بود . مسئله یافتن ریشه دیگر $x = (-b + \sqrt{(b^2-4ac)})/2a$ بوده، این ریشه دوم را می‌توان به‌کمک معادلهٔ $x_1 . x_2 = c/a$ نیز یافت . معادلهٔ $x^2 - 1000{\cdot}01x + 10 = 0$ دارای‌ریشه‌های $x_1 = 1000$ و $x_2 = 0{\cdot}01$ می‌باشد . با به کاربردن فرمول برای ۶ رقم‌اعشار $x_2 = 0{\cdot}015$ ، که دارای خطائی‌معادل 50% است . اگر بزرگترین ریشه‌ایکه ابتدا پیدا می‌کنیم $x_1 = 999{\cdot}995$ باشد ، با به‌کاربردن معادلهٔ $x_1 . x_2 = 10$ مقدار $x_2 = 0{\cdot}009\,999\,95$ و خطاها کاملا" کوچک می‌گردند .

۶ــ همانطوریکه در مثال زیر نشان داده شده است یک تغییر ساده در ترتیب محاسبات ، بعضی مواقع می‌تواند اثر قابل توجهی داشته باشد . در مثالها ، عملیات حسابی تا ۴ رقم با معنی درنظر گرفته شده‌اند . از نقطه‌نظر کامپیوتری بدین معنی است که ۸ محل برای هر محاسبه وجود دارد ، اما پاسخ بایستی برای ۴ رقم با معنی گرد شود . بیان ساده زیر این موضوع را نشان می‌دهد .

$$225{\cdot}1 - (224{\cdot}8 + 0{\cdot}1572) = 225{\cdot}1 - (225{\cdot}0) = 0{\cdot}1000$$

بصورت دیگر داریم

$$225{\cdot}1 - 224{\cdot}8 - 0{\cdot}1572 = 0{\cdot}3000 - 0{\cdot}1572 = 0{\cdot}1428$$

این دو مثال جمع نشان می‌دهد که چطور ترتیب جمعها در جواب اثر می‌کند .

991·1
327·6
1318·7
1319·
225·0
1544·
85·67
1629·67
1630·
75·61
1705·61
1706·
25·54
1731·54
1732·
3·712
1735·712
1736·
1·543
1737·543
1738·
0·1001
1738·
0·5112
1738·5112
1739·

0·5112
0·1001
0·6113
1·543
2·1543
3·712
5·866
25·54
31·406
31·41
75·61
107·02
107·0
85·67
192·67
192·7
225·0
417·7
327·6
745·3
991·1
1736·4
1736

جواب واقعی عبارتست‌از :

0·5112
0·1001
1·543
3·712
25·54
75·61
85·67
225·0
327·6
991·1
1736·3863

۷ ــ به عنوان مثال در فصل بعدی نشان داده شده است که چطور معادلهٔ درجه دوم $x_{n+1}=x_n^2+0{\cdot}16$ قابل حل می‌باشد. با انتخاب $x_0=0{\cdot}15$ دنبالهٔ تکرارها در زیر آورده شده است. اولین و دومین تفاضل محاسبه شده، و روند Δ^2 ایتکن به کار می‌رود. مقادیر خام قبلی همگرا می‌گردد. روند Δ^2 با معادلهٔ زیر تعریف می‌شود.

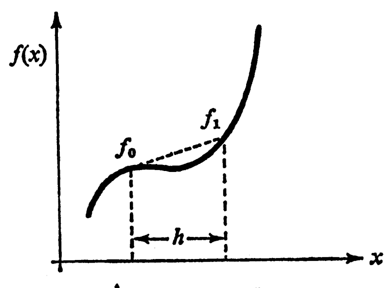
$$x_{n+1}^*=x_n-\frac{(\Delta x_n)^2}{\Delta^2 x_n}$$

اعداد تا ۶ رقم اعشار در جدول (۱۰۱) داده شده است.

جدول (۱۰۱)

x	Δx	$\Delta^2 x$	x^*
0·150000	0·032500	−0·021650	0·198686
0·182500	0·010805	−0·006744	0·199811
0·193306	0·004061	−0·002475	0·199969
0·197367	0·001586	−0·000957	0·199995
0·198953	0·000629	−0·000379	
0·199582	0·000250	−0·000150	
0·199832	0·000100	−0·000060	
0·199932	0·000040	−0·000024	
0·199972	0·000016	−0·000009	
0·199988	0·000007		
0·199995			

۸ ــ تقریب ساده‌ای برای محاسبهٔ سطح زیر یک منحنی به کمک رسم خط مستقیم بین نقاط متوالی و به کار بردن قاعدهٔ ذوزنقه مطابق شکل ۱۰۳ بدست می‌آید.



سطح $=\frac{1}{2}h[f_0+f_1]$

شکل ۱۰۳ قاعدهٔ ذوزنقه

معمولاً"، فاصله به چندین زیر فاصله تقسیم می‌شود، بنابراین فرمول کلی به صورت زیر به کار می‌رود.

$$\text{سطح}=\frac{h}{2}[f_0+2f_1+2f_2+\cdots+f_n]$$

که در آن h ، فاصله بین مقادیر متوالی تابع می‌باشد. نتایج زیر در ستون اول جدول ۱۰۲ برای انتــگرال‌گیــری از x^4 در فاصله $x=0$ و $x=2$ به ترتیب برای تعداد فواصل 4 ,8 ,16 , 32 بدست آمده‌اند.

جدول (۱۰۲)

n	Integral	R_1	R_2	R_3
4	7·06250			
		6·40104		
8	6·56641		6·39999	
		6·40006		6·40000
16	6·44165		6·40000	
		6·40001		
32	6·41042			

فرمول خطا برای قاعدهٔ انتگرال‌گیری به طریق ذوزنقه دارای شکل مناسبی برای کاربرد روش برونیابی ریچاردسن می‌باشد، یعنی

$$I=I_n+\sum_{r=1}^{\infty} a_r h^{2r}$$

بدین‌ترتیب

$$I=I_{2n}+\sum_{r=1}^{\infty} a_r\left(\frac{h}{2}\right)^{2r}$$

$$4I=4I_{2n}+a_1h^2+\sum_{r=2}^{\infty} 4a_r\left(\frac{h}{2}\right)^{2r}$$

بنابراین، جمله h^2 با تفریق این دو معادله حذف می‌گردد.

$$3I=4I_{2n}-I_n+\sum_{r=2}^{\infty} b_r(h)^{2r}$$

مقادیر در ستون R_1 بکمک فرمول

$$R_1=\frac{4I_{2n}-I_n}{3}$$

محاسبه می‌شوند. مقادیر در ستــونهــای بعــدی به‌کمک $R_m=(2^{2m}I_{2n}-I_n)/(2^{2m}-1)$ محاسبه می‌گردند. لذا آخرین ستون دارای خطائی از مرتبه h^8 می‌باشد. می‌توان دید که مقادیر روی هر قطر اریب (از چپ به راست) به مقدار واقعی همگرا می‌گردند.

$$6.4=\int_0^2 x^4\,dx=\left[\frac{x^5}{5}\right]_0^2$$

۹ــ در صورتیکه روش فوق تخفیف را روی روش نیوتن برای مسئله $x^{20}-1$ اثردهیم نتیجه جالب‌توجهی را خواهیم داشت. روش نیوتن در قسمت بعد بحث شده است. یک قسمت

از نتایج روش نیوتن ، با مقدار تقریب اولیه $x_0 = 10$ ، در جدول ۱.۳ داده شده است . ستونهای بعدی جدول نتایج روش فوق تخفیف را به ازای $\omega = 2.0$ نمایش می‌دهد .

جدول (۱.۳)روش نیوتن

Iteration	Result	Iteration	Result
1	9·500 000	38	1·423 993
2	9·025 000	39	1·352 854
3	8·573 750	40	1·285 372
4	8·145 063	41	1·221 527
5	7·737 809	42	1·161 567
6	7·350 919	43	1·106 394
7	6·983 373	44	1·058 397
8	6·634 204	45	1·022 485
9	6·302 494	46	1·004 132
10	5·987 369	47	1·000 158
11	5·688 001	48	1·000 000
12	5·403 601	49	1·000 000

Iteration	Result	Iteration	Result
1	9·000 000	13	2·541 866
2	8·100 000	14	2·287 679
3	7·290 000	15	2·058 911
4	6·561 000	16	1·853 020
5	5·904 900	17	1·667 719
6	5·314 410	18	1·500 953
7	4·782 969	19	1·350 902
8	4·304 672	20	1·216 142
9	3·874 205	21	1·096 956
10	3·486 784	22	1·004 495
11	3·138 106	23	0·995 877
12	2·824 295	24	1·004 456

مشاهده می‌شود که روندهای فوق تخفیف در آغاز میزان همگرائی سریعتری را می‌دهند، طوریکه شش تکرار اولیه پیشرفت بهتری نسبت به ۱۲ تکرار روش نیوتن دارد . لیکن ، زمانیکه جواب تکراری نزدیک به ریشه باشد ، روش نیوتن سریعا" همگرا گردیده و روش فوق تخفیف خیلی آهسته در همگرائی نوسان می‌کند . از تکرار ۴۶ روش نیوتن دارای سه گام بیشتر است . برای یک مقدار مشابه در تکرار ۲۲ ، روش فوق تخفیف ۱۵۰ گام بیشتر نیاز دارد .

اگر یک مقدار میانی برای عامل تخفیف فرض شود ، $K = 1{\cdot}5$ ، روند تکراری در سی‌امین تکرار به مقدار 1·001 752 رسیده و سپس نوسان می‌کند و در تکرار چهل و

سوم در ۶ رقم اعشار همگرا می‌گردد. بطور وضوح، روش فوق تخفیف بدون تجزیه و تحلیل دقیق نباید به‌کار رود.

۱۰ ـ مثال زیر همچنین دقت اینکه کدام روش بایستی در تعبیر نتایج کامپیوتری به‌کار رود را نشان می‌دهد. مسئله عبارتست از پیدا کردن ریشهٔ معادله

$$5x-3x^5=0$$

با به‌کار بردن معادلهٔ تکراری

$$x_{n+1}=\tfrac{3}{5}x_n^5$$

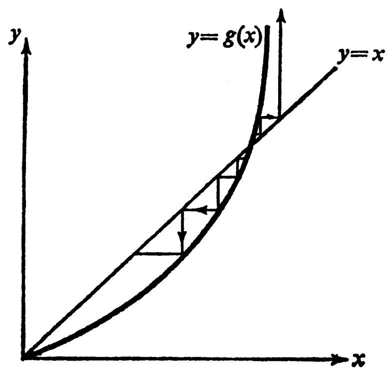
با اولین تقریب، تا ۶ رقم اعشار، $x_0=0{\cdot}880\,112$ فرض می‌شود. نتایج کامپیوتری جدول ۱۰۴ از یک برنامه کامپیوتری برای هر دو حالت بدست آمده است.

Results 1	*Results* 2
0·880 112	0·880 112
0·880 112	0·880 112
0·880 112	0·880 112
0·880 112	0·880 112
0·880 112	0·880 112
0·880 112	0·880 112
0·880 112	0·880 112
0·880 111	0·880 112
0·880 108	0·880 112
0·880 094	0·880 114
0·880 022	0·880 124
0·879 662	0·880 171
0·877 867	0·880 407
0·868 945	0·881 587
0·825 675	0·887 514
0·639 579	0·917 753
0·178 369	1·085 118
0·000 301	2·507 456
0·000 000	165·202 101
	$>10^6$

جدول ۱۰۴

ملاحظه می‌شود که مقادیر تقریب برای ۵ تکرار آغازی مساوی بوده بنابراین بنظر نمی‌رسد که روند همگرا گردد. چند مقدار اولی به علت اینکه تقریب آغازی خیلی به مقدار واقعی ریشه‌نزدیک است واگرا نمی‌گردند. بهرحال خطای گرد شده تدریجا" سیمای واقعی واگرائی روند را آشکار می‌کند. در حالت اول روند به یکی از ریشه‌ها، $x=0$ ، همگرا شده و در حالت دوم مقادیر بدون کران بزرگ می‌شوند. نمودار این روند تکراری در شکل (۱۰۴) نشان داده‌شده‌است.

دلیل متفاوت بودن دو مجموعه نتایج، عبارتست از اینکه اولین نتایج در سیستم ICL 1904 با ۳۷ محفظه در مانتیس محاسبه شده است. دومین مجموعه نتایج در



شکل ۱۰۴ تکرار واگرا

ICL 1906A ، با طول چندگانه و ۷۴ مکان دوتائی برای محاسبه‌های حسابی ، محاسبه و گرد شده‌اند . آشکارا ، روند انتقال یک برنامه از یک ماشین به دیگری ، حتی زمانیکه بطور تئوری سازگار باشند ، ممکن است به آن سادگی که تصور می‌شود نباشد .

۱۱ ــ با استفاده از برنامهٔ کامپیوتری و به روش تکرار ساده ریشهٔ مثبت معادلهٔ $f(x) = 0.5 - x + 0.2 \sin x = 0$ را با $x_0 = 0.5$ تا ۷ رقم اعشار درست محاسبه کنید . برنامهٔ کامپیوتری شماره ۱۰۱ و نتایج آن وسیله IBM 7094 در ذیل آمده است .

ALGORITM 1.1

FORTRAN PROGRAM FOR EXAMPLE *11*

```
C        TO SOLVE F(X) = 0 OR G(X) = X
         F(X) = X - .2*SIN(X) - .5
         G(X) = .5 + .2*SIN(X)
         X = .5
         Y = F(X)
         WRITE(6,4)X,Y
         DO 1 I = 1,20
         X1 = G(X)
         Y = F(X1)
         WRITE(6,5)I,X1,Y
         IF(ABS (X1 - X) .LT. 1.E-7)GO TO 2
       1 X = X1
         WRITE(6,6)
       2 STOP
       4 FORMAT(1H03X1HI8X 4HX(I) 12X 8HF(X(I)) //4X1H0I2E17.8)
       5 FORMAT(1H0 I4,2E17.8)
       6 FORMAT(//36H0FAILED TO CONVERGE IN 20 ITERATIONS)
         END
```

COMPUTER RESULTS FOR EXAMPLE *11*

I	X(I)	F(X(I))
0	0.50000000E 00	−0.95885109E−01
1	0.59588511E 00	−0.16363200E−01
2	0.61224830E 00	−0.26934631E−02
3	0.61494176E 00	−0.44042990E−03
4	0.61538219E 00	−0.71939081E−04
5	0.61545412E 00	−0.11753291E−04
6	0.61546587E 00	−0.19222498E−05
7	0.61546779E 00	−0.32037497E−06
8	0.61546810E 00	−0.55879354E−07
9	0.61546815E 00	−0.14900161E−07

■ مسائل

۱ ـ اعداد زیر را به دستگاه دوتائی تبدیل کنید .

37, 61, 211, 107, 57, 127

۲ ـ اعداد دوتائی زیر را به مقادیر دهدهی تبدیل نمائید .

1101, 11110, 1010110, 11011011, 1010101, 11001100

۳ ـ اعداد زیر را به دوتائی تبدیل کرده و محاسبات را ادامه دهید . جوابهای حاصل را پس از تبدیل به اعداد دهدهی با هم مقایسه کنید .

$23 + 17,\ 18 - 12,\ 15 \times 14,\ 63 \div 9$

۴ ـ فرض کنید

$$a = \alpha_m 10^m + \alpha_{m-1} 10^{m-1} + \alpha_{m-2} 10^{m-2} + \dots + \alpha_{m-n+1} 10^{m-n+1} + \dots$$

و a تقریبی از A باشد گوئیم a دارای n رقم با معنی درست است در صورتی که n بزرگترین عدد صحیح نامنفی باشد که $\Delta = |A - a| \leqslant 5 \times 10^{m-n}$ در مورد اعداد زیر تعداد ارقام بامعنی درست را بدست آورید .

A = 1.5789 a = 1.579

A = 25.39 a = 25.7

A = -78.39 a = -78.3

$A = \sqrt{2}$ a = 1.414

A = 10.001 a = 9.98

۵ ـ گوئیم عدد a دارای n رقم اعشار درست است در صورتی که n بزرگترین عدد صحیح نامنفی باشد که $\Delta = |A - a| \leqslant 5 \times 10^{-n-1}$ ، تعداد ارقام اعشار درست اعداد زیر را حساب کنید .

A = 1.789634 a = 1.79

A = 20.70184 a = 20.702

A = 0.005673 a = 0.006

۶ ـ مطلوبست محاسبهٔ :

الف : عدد $e^{\frac{1}{7}}$ تا سه رقم اعشار درست

ب : عدد $e^{\frac{2}{3}}$ تا چهار رقم با معنای درست

ج : عدد $\sin\frac{\pi}{12}$ تا چهار رقم اعشار درست

د : عدد $\cos\frac{\pi}{7}$ تا پنج رقم اعشار درست

۷ ـ مطلوبست محاسبهٔ :

الف : $\cos 17° 24'$ با دقت 10^{-5} .

ب : $e^{\frac{1}{2}}$ با دقت 10^{-5} .

ج : Sh 1.4 با دقت 10^{-6}.

۸ ــ گردشدهٔ هر عدد را تا سه رقم با معنا در زیر آن عدد بنویسید .

1.708321, 90.071, 0.08895, 7.445, 0.04376002

۹ ــ اعداد ذیل مفروضند . ابتدا این اعداد را از کوچک به بزرگ مرتب کنید ، سپس حاصل‌جمع آنها را به سه طریق حساب کنید .

0.5112, 0.1001, 1.543, 3.712, 25.54, 75.61, 225.0

الف : اعداد را از کوچک به بزرگ جمع کنید ، به این ترتیب که ابتدا دو عدد اول را جمع کرده و حاصل‌جمع را تا چهار رقم با معنا گرد کنید ، نتیجهٔ به‌دست‌آمده را با عدد سوم جمع کنید و حاصل‌جمع را تا چهار رقم با معنا گرد کنید و الی آخر .

ب : اعداد را مانند قسمت الف از بزرگ به کوچک جمع کنید .

ج : حاصل‌جمع اعداد را بدون گرد کردن به‌دست آورید . آیا جوابها متفاوتند؟ چرا؟

۱۰ ــ در صورتی که $\bar{a}$ و $\bar{b}$ و $\bar{c}$ و $\bar{d}$ مقادیر تقریبی از a و b و c و d باشند ، مطلوبست محاسبهٔ خطای نسبی کسر زیر :

$$\frac{abcd}{(c-d)(a-b)(c-a)}$$

الف : بدون در نظر گرفتن خطای ناشی از گرد شدن اعداد وسیلهٔ ماشین .

ب : با در نظر گرفتن خطای ناشی از گرد شدن اعداد وسیلهٔ ماشین .

۱۱ ــ عدد 0.625 را در مبنای ۲ بنویسید .

۱۲ ــ اعداد زیر را به مبنای دهدهی تبدیل کنید .

$(0.110011)_2$, $(0.111111)_2$

۱۳ ــ با استفاده از شکل مناسب معادلهٔ تکراری ریشه‌های حقیقی معادلات زیر را با دقت 10^{-3} پیدا کنید .

(a) $x = \frac{1}{(x+1)3}$

(b) $\sin X = 10(X-1)$

(c) $x = (5-x)^{\frac{1}{3}}$

(d) $e^{-X} = 10x$

(e) $x - \cos x = 0$

۱۴ـ کلیه ریشه‌های معادلهٔ $x^2 + 10 \cos x = 0$ را با استفاده از تکرار ساده با دقت 10^{-3} پیدا کنید.

۱۵ـ ثابت کنید دنباله زیر، به ازای هر $x_0 > 0$،

$$x_n = \frac{1}{2}\left(x_{n-1} + \frac{2}{x_{n-1}}\right) \qquad n > 1$$

به $\sqrt{2}$ همگرا خواهد شد.

۱۶ـ معادله $x^2 - x - 2 = 0$ مفروض است.

الف: با انتخاب شکل مناسبی از معادلهٔ تکراری مقدار تقریبی ریشه نزدیک به 2 را با $x_0 = 0$ پیدا کنید.

ب: مقدار تقریبی این ریشه را با روش Δ^2 ایتکن پیدا کرده و سرعت همگرائی هر دو روش را با هم مقایسه کنید. (دقت محاسبه 10^{-3})

۱۷ـ کامپیوتری را در نظر می‌گیریم که برای نشان دادن اعداد ممیز سیار دارای مانتیس ۸ رقمی با معنا و دو رقم برای توان می‌باشد.

فرض می‌کنیم دو عدد به صورت زیر داده باشند:

$$4735.821653 = 0.47352165 \times 10^4$$
$$329.4175251 = 0.32941752 \times 10^3$$

الف: دو عدد را با هم جمع کرده سپس نتیجه را به شکل ممیز سیار بنویسید.

ب: اگر هر دو عدد به صورت درست داده شده باشند، مقدار خطا در جمع دو عدد چه خواهد بود؟

ج: در صورتی که مانتیس‌ها به صورت بالا گرد شده باشند، مطلوبست محاسبهٔ کران خطا برای مجموع محاسبه‌شده در قسمت الف.

۱۸ـ معادلهٔ $1 - 5x - \mathrm{tg}\, x = 0.1$ مفروض است.

کوچکترین ریشهٔ حقیقی معادله را با مقدار اولیه $x_0 = 0$ به روش تکرار ساده با تقریب 10^{-4} یافته سپس با استفاده از روش Δ^2 ایتکن به کمک مقادیر تکرار ساده این ریشه را با دقت 10^{-4} بیابید.

۱۹ـ نشان دهید معادلهٔ $x = 2^{-x}$ در فاصلهٔ $[\frac{1}{3}, 1]$ یک ریشه دارد.

۲۰ـ نشان دهید $\sin x = x$ $\pi + 0.5$ در فاصلهٔ $[0, 2\pi]$ یک ریشه دارد.

۲۱ـ با استفاده از تکرار ساده یک ریشهٔ تقریبی از معادلهٔ $x - 4^{-x} = 0$ را با دقت 10^{-2} پیدا کنید.

۲۲ـ کوچکترین ریشهٔ هریک از معادلات زیر را با دقت 10^{-5} پیدا کنید.

(a) $x = 4^{-x}$

(b) $x = 5^{-x}$

(c) $x = 6^{-x}$

(d) $x = \frac{2 - e^{x} + x^{2}}{3}$

۲۳ـ اعداد زیر در یک ماشین با ۴ رقم مانتیس استاندارد شده داده شده است :

$a = 0.4523 \times 10^{4}$

$b = 0.2115 \times 10^{-3}$

$c = 0.2583 \times 10^{1}$

اعمال زیر را انجام داده و خطاها را در هریک محاسبه نمائید .

الف : $a + b + c$

ب : $a - b - c$

ج : a / c

د : ab / c

۲۴ـ مطلوبست محاسبه ریشه‌های حقیقی معادلهٔ $2x - \cos x - 3 = 0$ با استفاده از روش Δ^2 ایتکن با دقت 10^{-2} .

۲۵ـ در صورتی که برای بدست آوردن مقدار تقریبی برای $y = 3$ به کمک دو بسط به ترتیب با تقریبهای $O(h_1^5)$ و $O(h_2^5)$ که درآن $h_2 = \frac{1}{2} h_1$ مقادیر $y_1 = 2.67$ و $y_2 = 2.90$ را بدست آوریم ، با استفاده از روش برونیابی ریچاردسن تقریب جدیدی برای y یافته و نشان دهید این مقدار نسبت به دو مقدار y_1 و y_2 برای y دارای تقریب بهتری است و دلیل آن را ذکر نمائید .

۲۵ـ مطلوبست محاسبهٔ انتگرال تابع x^4 از فاصلهٔ ۰ تا ۱ به روش ذوزنقه‌ای با استفاده از $h = 1$ و $h = 0.5$ ، همچنین با استفاده از دو مقدار تقریبی که بدست می‌آورید ، به کمک روش رامبرگ ، تقریب بهتری از مقدار انتگرال بدست آورید و با مقدار واقعی انتگرال مقایسه نموده و نتیجه را توجیه کنید .

فصل دوم
حل معادلات غیرخطی

معادلاتی هستند که از نوع غیر خطی بوده و به علت وجود روشهای تحلیلی که منجر به فرمولهایی برای جواب آنها می‌شود برای خواننده شناخته شده می‌باشند و حل آنها به صورت یک دستور در می‌آید.

حل معادلات درجه دوم ، یا معادلات مثلثاتی نظیر $\sin 3x + 2\cos 3x = 2$ مثالهای ساده‌ای را تشکیل می‌دهند. به هرحال، معادلات غیر خطی زیادی وجود دارند که مستقیما " قابل حل با روشهای تحلیلی نبوده و برای حل آنها بایستی روشهای مبتنی بر تقریب به‌کار برد.

۲.۱ ■ مثالی ازتکرارساده

روش تکرار ساده برای به کار بردن آسان بوده و می‌تواند در حوزهٔ وسیعی از مسائل متعدد به کار رود. ساده‌ترین شکل معادلهٔ تکراری با تغییر نظم در معادلهٔ اصلی بطوری که x دریک طرف قرار گیرد بدست می‌آید. تقریبی برای x در طرف دوم معادله قرار داده می‌شود و مقدار جدید x محاسبه می‌گردد. این مقدار جدید x در محاسبه به کار رفته و مقدار بعدی x را داده و روند تکرار می‌گردد.

مثال ساده‌ای از چنین معادله به وسیلهٔ معادله (2.3) داده شده است.

اگر روش موفقیت‌آمیز باشد این مقادیر باید رفته رفته به مقدار واقعی جواب نزدیکتر گردد. در چنین حالتی روش را همگرا می‌گوئیم. شرایطی که همگرائی را تضمین کند قسمت مهمی از بحث این فصل را تشکیل می‌دهد. مادامی که تجزیه و تحلیل دقیق صورت نگرفته، بسیار ساده است که روشی انتخاب شود که مقادیر دنبالهٔ بدست آمده از ریشه دور گردد.

بعنوان مثال، معادلهٔ درجه دوم زیر را مطالعه‌می‌کنیم .

$$x^2 - x + 0{\cdot}16 = 0 \qquad (2.1)$$

این مثال به دلیل اینکه جوابهای $x = 0{\cdot}2$ و $x = 0{\cdot}8$ به سادگی پیدا می‌شوندانتخاب شده است . روش تکراری تنها به حل معادله درجه دوم خلاصه نمی‌گردد !
این معادله را می‌توان به صورت زیر نوشت .

$$x = x^2 + 0{\cdot}16 \qquad (2.2)$$

فرض کنیم که رسم تقریبی تابع ، دلالت کند که کوچکترین ریشه تقریبا " $0{\cdot}15$ می‌باشد . پس از جایگذاری مقدار $x_0 = 0{\cdot}15$ در سمت راست خواهیم داشت . $x_1 = 0{\cdot}1825$ این مقدار سپس می‌تواند در سمت راست گذاشته شود ، و روندمطابق دستور زیر تکرار گردد .

$$x_{n+1} = x_n^2 + 0{\cdot}16, \qquad n = 0, 1, \ldots \qquad (2.3)$$

مقادیر بدست آمده در جدول ۲۰۱ داده شده است .

تکرار ساده ، $x_{n+1} = x_n^2 + 0{\cdot}16$

x_0	0·150 000	x_6	0·199 832
x_1	0·182 500	x_7	0·199 932
x_2	0·193 306	x_8	0·199 972
x_3	0·197 367	x_9	0·199 988
x_4	0·198 953	x_{10}	0·199 995
x_5	0·199 582	x_{11}	0·199 999

جدول ۲۰۱

متذکر می‌شویم که این روند به مقدار واقعی ریشه ، $x = 0{\cdot}2$ ، همگرا بوده اما سرعت همگرائی کند می‌باشد .

بسادگی می‌توان نشان داد که این روند همیشه همگرا نیست . معادلهٔ

$$2x^2 - x = 0 \qquad (2.4)$$

با ریشه‌های $x = \frac{1}{2}$ و $x = 0$ و مقدار تقریب اولیه $x_0 = 1$ و معادلهٔ تکراری

$$x_{n+1} = 2x_n^2, \qquad n = 0, 1, \ldots \qquad (2.5)$$

را درنظر می‌گیریم .

دنبالهٔ تقریبها عبارتستاز : $x_0 = 1$ ، $x_1 = 2$ ، $x_2 = 8$ و $x_3 = 128$ که آشکارا همگرا نمی‌باشد . در این فصل یافتن تقریب آغازی خوب و انتخاب شکل تکراری‌با خواص همگرائی خوب مورد توجه است .

۲.۲ ■ یافتن تقریبهای آغازی

چند خاصیت مفید از چند جمله‌ایها که یافتن تقریبهای مربوط به ریشه‌ها را

مقدور می‌سازد، در فصل ۳ معرفی شده‌اند. در مورد یک معادلهٔ غیر خطی در حالت کلی روشهای معدودی وجود دارد. روشهای زیر حتما" مفید خواهند بود.

۱ ــ نتایج فیزیکی مسئله ممکن است منجر به تعیین تقریب خوبی برای ریشه گردد.

۲ ــ دیگر اینکه، غالبا" رسم نمودار تقریبی تابع ممکن بوده و نشانگر محل تقریبی ریشه‌ها خواهد بود.

۳ ــ گاهی معادله را می‌توان به دو تابع تجزیه نمود، محل برخورد نمودارهای دوتابع به مراتب روشنتر از نمودار تابع اصلی نشان می‌دهد که ریشه‌ها در کجا قرار دارند.

۴ ــ اگر یک برنامهٔ کامپیوتری مورد نظر باشد، محاسبهٔ مقادیر تابع می‌تواند با نظم مشخصی ادامه پیدا کند، تا آنکه دو مقدار با علامت مخالف حاصل شود. برای یک تابع پیوسته نتیجه می‌شود که این مقادیر ریشه‌ای را دربر می‌گیرند.

۵ ــ گاهی، وضع طوری است که در ناحیه‌های مشخص بعضی از قسمتهای معادله قابل اغماض می‌باشد. آنگاه ممکن است که یک ریشه تقریبی معادله، از حل قسمت باقیماندهٔ معادله بدست آید.

۲.۳ ■ تکرارساده

روش ساده‌ای که در مقدمهٔ فصل به کار رفت می‌تواند، در صورتیکه فرمول همگرا باشد، برای تکرار به کار رود. شکل عمومی روش با تغییر نظم فرمول $f(x)=0$ بدست می‌آید، که رابطه

$$x=g(x) \tag{2.6}$$

را داده و منجر به شکل تکراری

$$x_{n+1}=g(x_n), \qquad n=0,1,\ldots \tag{2.7}$$

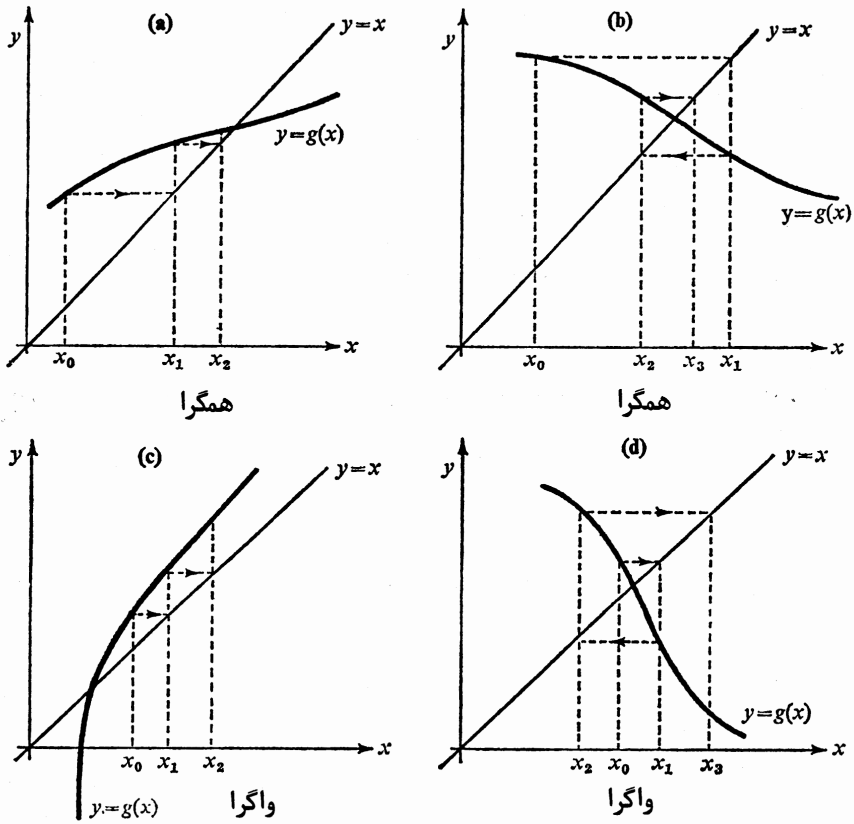
می‌گردد.

اولین تقریب x_0 را در سمت راست قرار داده و یک مقدار تقریبی جدید x_1 در سمت چپ بدست می‌آید. این مقدار را در سمت راست گذاشته و روند تکرار می‌گردد تا اینکه مقادیر همگرا گردند.

بوسیله نمودار می‌توان به سادگی نشان داد که نتایج تحت کدام روند همگرا خواهند شد. شکل ۲.۱(a) نشانگر نمودار نوعی از روندهای همگرا است. نمودار $y=g(x)$ رسم گردیده، و تعیین مقدار $y_0=g(x_0)$ از روی نمودار رسم شده، امکان‌پذیر می‌باشد. آنگاه مقدار $x_1=y_0$ مورد نیاز است که با رسم خط $y=x$ پیدا می‌شود.

با مشاهدهٔ خطوط بریدهٔ جهت‌دار (مطابق شکل)، می‌توان دید که این مقادیر به نقطه x^* همگرا می‌گردند، که در آن $x^*=g(x^*)$ ریشه معادله می‌باشد. در این نوع

همگرائی تفاضل بین ریشه تقریبی و واقعی همیشه دارای علامت مشابه بوده، بنابراین مقادیر بطور یکنواخت به مقدار نهائی میل می‌کند. این نوع رفتار معروف به همگرائی یکنوا می‌باشد.



شکل ۲۰۱ تکرار ساده

شکل ۲۰۱(b) نوع دیگر از همگرائی را نشان می‌دهد که مقادیر در دو طرف جواب واقعی نوسان می‌کند. این همگرائی نوسانی خیلی ساده است، زیرا دو مقدار نهائی، کرانهایی را بدست می‌دهند که جواب واقعی باید بین آنها قرار گیرد. بهر حال، بعید نیست که این روش واگرا گردد و شکلهای ۲۰۱(c) و ۲۰۱(d) دو مثال ساده از این واگرائی را نشان می‌دهد.

از مطالعهٔ نمودار روشن می‌گردد که زمانیکه گرادیان تابع $g(x)$ کمتر از گرادیان خط $y = x$ است روند همگرا خواهد شد، یعنی، در ناحیه پوشیده شده به وسیله تکرار به

$$|g'(x)| < 1 \qquad (2.8)$$

نیازمند می‌باشیم .

همچنین این مطلب به طور نظری قابل بیان می‌باشد . داریم

$$x_{n+1}=g(x_n), \qquad n=0,1,\ldots \tag{2.9}$$

و ریشه واقعی x^* در رابطه

$$x^*=g(x^*) \tag{2.10}$$

صدق می‌کند .

تفاضل این برابریها رابطهٔ زیر را می‌دهد :

$$x^*-x_{n+1}=g(x^*)-g(x_n) \tag{2.11}$$

پس از به کار بردن قضیهٔ مقدار میانگین ، سمت راست رابطهٔ فوق بصورت زیر درمی‌آید :

$$g(x^*)-g(x_n)=(x^*-x_n)g'(\zeta) \tag{2.12}$$

که درآن ζ بین x^* و x_n قرار دارد . درصورتیکه تعریف کنیم $\varepsilon_n=x^*-x_n$ معادلهٔ فوق بصورت زیر در می‌آید :

$$\varepsilon_{n+1}=g'(\zeta)\,.\,\varepsilon_n \tag{2.13}$$

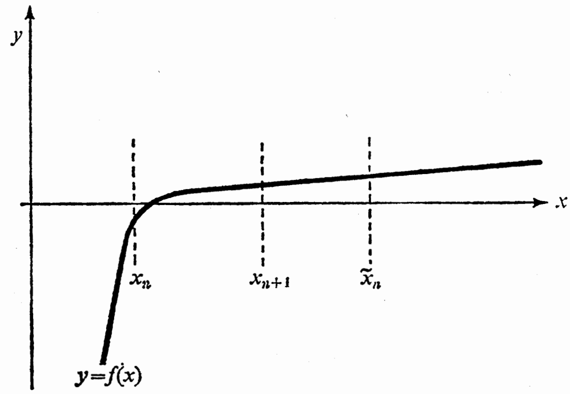
و اگر $|g'(\zeta)|<1$ ، آنگاه خطاها همواره گام به گام کوچکتر می‌گردند . برای $|g'(\zeta)|>1$ خطاها رشد خواهند کرد .

۲.۴ ■ روش دوبخشی

مسئله این است که با روش فوق امکان دارد نتوانیم دستور تکراری را بیابیم که در شرایط همگرائی (2.8) صدق کند . روش دوبخشی دارای این مزیت‌ها است که همواره همگرا بوده و خیلی ساده می‌باشد . لیکن ، سرعت همگرائی آهسته است و زمانیکه تکرارها به ریشه نزدیک شده‌اند ، معمولا" به کار بردن روش دیگر با سرعت همگرائی سریعتر مطلوب خواهد بود . این امر بوسیلهٔ مثال ۲.۱ نشان داده شده است .

فرض کنید یافتن مقادیر x بطوریکه $f(x)=0$ ضروری است . برای شروع روش ، مقادیر تابع برای یک رشته نقاط تا جائی محاسبه می‌شود که دو نقطهٔ x_0 و $\tilde{x}_0$ ، که مقادیر تابع متناظر با آنها مختلف‌العلامه باشند ، بدست آید . با فرض آنکه تابع پیوسته باشد باید یک ریشه بین x_0 و $\tilde{x}_0$ قرار گیرد . آنگاه تابع در نقطه $x_1=\frac{1}{2}(x_0+\tilde{x}_0)$ محاسبه می‌گردد . نقطهٔ $\tilde{x}_1$ ازبین x_0 و $\tilde{x}_0$ طوری انتخاب می‌شودکه علامت مقدار تابع متناظر با آن مخالف با علامت $f(x_1)$ باشد . فاصلهٔ بین x_1 و $\tilde{x}_1$ طوری بدست آمده که همچنان شامل یک ریشه بوده و اندازهٔ فاصله مذکور نصف اندازه فاصله اولی است .آنگاه روند مطابق رابطه $x_{n+1}=\frac{1}{2}(x_n+\tilde{x}_n)$ تا جائی ادامه می‌یابد که کران بالا و پائین ریشه به اندازه کافی به یکدیگر نزدیک گردند . شکل ۲.۲ نمونه‌ای از یک گام این روش را نشان

می‌دهد. علاوه بر کندی همگرائی محدودیت‌های دیگری نیز برای این روش موجود است. اگر یک ریشهٔ مضاعف، یا ریشهٔ چندگانه از مرتبه زوج، وجود داشته باشد، آنگاه تابع در همسایگی ریشه تغییر علامت نداده و روش دوبخشی را نمی‌توان به‌کار برد. همچنین، روش دوبخشی روش مناسبی برای پیدا کردن ریشه‌های مختلط نمی‌باشد.



شکل ۲۰۲ روش دو بخشی

۲.۵ ■ روش نیوتن

برای بدست آوردن روشی با سرعت همگرائی رضایت‌بخش‌تر، می‌توان روشهای معین محاسباتی به‌کار برد. اگر تکرار به نقطهٔ x_n رسیده باشد، آنگاه نموی به اندازهٔ Δx_n لازم است تا به جواب x^* برسیم. هرگاه $f(x^*)$ را به کمک سری تیلر بسط دهیم خواهیم داشت:

$$0=f(x^*)\equiv f(x_n+\Delta x_n)=f(x_n)+\Delta x_n f'(x_n)+\frac{(\Delta x_n)^2}{2!}f''(x_n)+\cdots \quad (2.14)$$

اگر Δx_n ، فاصلهٔ بین نقطهٔ تکراری و جواب واقعی، به اندازهٔ کافی کوچک باشد، آنگاه از دو جملهٔ اول طرف راست رابطه (2.14) داریم:

$$0\approx f(x_n)+\Delta x_n f'(x_n)$$

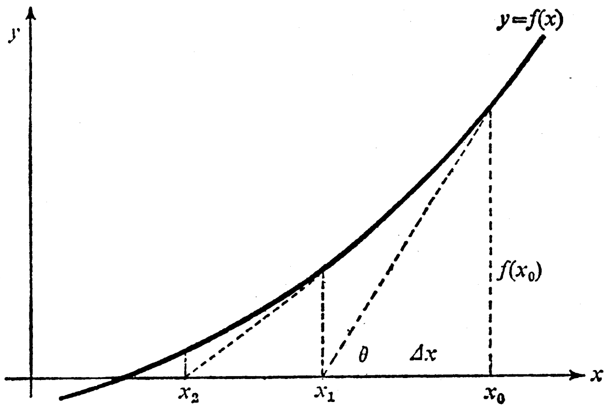
$$\Delta x_n \approx \frac{-f(x_n)}{f'(x_n)} \quad (2.15)$$

این رابطه به تشکیل فرمولی موسوم به روش نیوتن منجر می‌گردد،

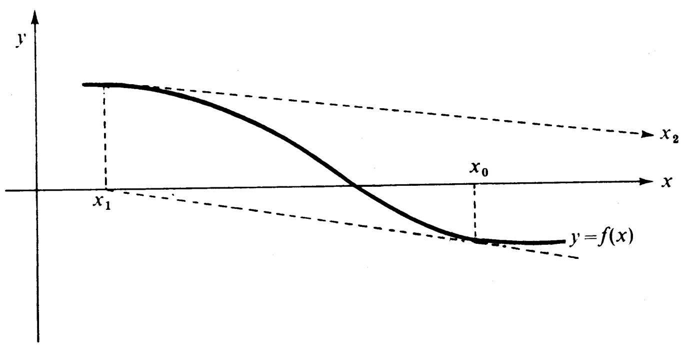
$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \quad (2.16)$$

این روش وسیله نمودار در شکل ۲۰۳ نشان داده شده است. درصورتیکه مماس بر منحنی در نقطهٔ x_0 رسم شود آنگاه داریم.

$$\tan\theta = \frac{f(x_0)}{-\Delta x} = f'(x_0) \tag{2.17}$$



شکل (۲۰۳) روش نیوتن



شکل ۲۰۴ واگرایی روش نیوتن

بنابراین گام Δx از طرق ترسیم با رسم مماسی بر منحنی در نقطه تکراری‌جاری، وتعیین نقطهٔ تلاقی آن با محور x ها پیدا می‌شود. این نقطه برای تکرار بعدی به کار گرفته می‌شود.

گرایش ترسیمی برای توضیح خواص همگرائی روش نیوتن سودمند می‌باشد. راههای گوناگونی وجود دارد که در آنها روش همگرا نمی‌باشد، و چند نمودار نوع رفتار موجود را نشان خواهد داد. شکل ۲۰۴ نشانگر مثالی از یک منحنی که به طرف محور متمایل شده می‌باشد، واضح است که روش نیوتن در اینجا واگرا می‌گردد. برای مقایسه، واضح است که برای منحنی نظیر شکل ۲۰۳، روش نیوتن همگرائی یکنوا را می‌دهد.

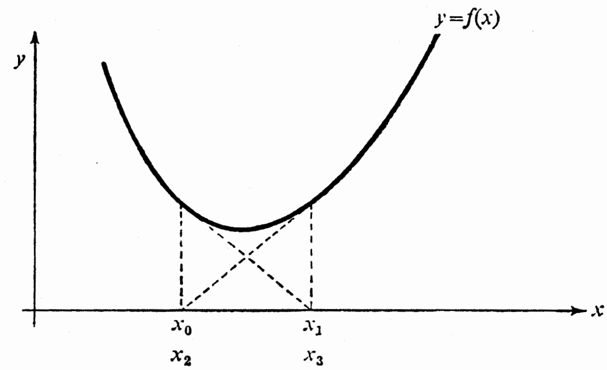
حالت دیگر به وسیلهٔ نوسان در شکل ۲.۵ نشان داده شده است. در اینجا اولین گرادیان در نقطه x_0 ، نقطه‌ای مانند x_1 را می‌دهد که در آن نقطه دومین گرادیان بدست می‌آید. آنگاه دومین گرادیان نقطه‌ای را ایجاد می‌کند که بسیار نزدیک به x_0 است، بنابراین روند بالا پیرامون نقطه‌ای نوسان می‌کند که ریشهٔ معادله نیست. (به مثال ۲.۲ نیز مراجعه شود) البته، در روند تکراری کامپیوتر، داشتن وسیله‌ای که در صورت همگرا نبودن روند متوقف گردد ضروری است. شمارنده‌ای که روند را بعد از تعداد معینی تکرار متوقف کند روش سهلی برای رسیدن به این امر است. همچنین اشکالاتی هستند که به هنگام نزدیکی $f'(x)$ به صفر پدیدار می‌شوند، مانند حالتیکه ریشه‌ها چندگانه باشد. در مثال ۲.۳ این مطلب نشان داده شده است.

زمانیکه روش نیوتن به کار می‌رود ازنظر اشکالات احتمالی، کنترل همگرائی روند و رضایت بخش بودن آن نیازمند دقت است. همگرائی هنگامی تضمین می‌شود که تابع دارای مشتق مرتبه دومی است که در ناحیه‌تکرار تغییر علامت نداده، و در شرط زیر نیز صادق باشد.

$$f(x)f''(x) > 0 \qquad (2.18)$$

اگر مشتق مرتبه دوم به سادگی محاسبه گردد آنگاه این شرط می‌تواند برای کنترل به‌کار رود. هرگاه همگرائی از طریق ریاضی قابل محاسبه نباشد، آنگاه می‌توان به وسیلهٔ برنامه‌نویسی دقیق، پیشرفت تکرارها را برای ملاحظهٔ اینکه تکرارها همگرا می‌گردند یا نه، زیرنظر گرفت.

اگر هیچ نشانه‌ای از همگرائی موجود نباشد در این صورت برنامه‌ای لازم است که به‌روش تکراری دیگری که همگرائی آن تضمین شده است بازگشت نماید.



شکل ۲.۵ روش نوسانی نیوتن

هرگاه نشان داده شود که روش همگرا است، آنگاه اگر سرعت همگرائی کند باشد

ممکن است مشکلاتی پدید آید . این امر با کار برد روش نیوتن در مورد معادلهٔ $x^{20}-1=0$ مشخص می‌گردد . اگر مقدار تقریبی اولیه $x_0=\frac{1}{2}$ فرض شود ، دنباله مقادیر زیر تولید می‌گردد .

$$x_1=26\,214{\cdot}9, \qquad x_2=24\,904{\cdot}1, \qquad x_3=23\,658{\cdot}9, \qquad x_4=22\,476{\cdot}0$$

تعداد تکرارها پیش از اینکه این دنباله یکنوا به مقدار واقعی $x=1$ همگرا گردد خیلی بزرگ خواهد بود . دلیل این امر هنگامی آشکار می‌شود که دستور نیوتن به‌صورت زیر نوشته شود .

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^{20}-1}{20x_n^{19}} \\ &= \frac{19}{20}x_n + \frac{1}{20x_n^{19}} \end{aligned} \tag{2.19}$$

جمله دوم به ازای $x>2$ بسیار کوچک بوده و تقریب $x_{n+1}=\frac{19}{20}x_n$ نشان می‌دهد که مقادیر بسیار آهسته کاهش می‌یابد .

جذابیت روش نیوتن عبارتست از اینکه ، زمانیکه خطاها کوچک هستند هر خطا با مربع خطای قبلی متناسب است که نسبت به رابطه خطی موجود برای تکرار ساده همگرائی سریعتری را بدست می‌دهد .

رابطهٔ خطا برای روش نیوتن وسیلهٔ بسط $f(x)$ حول نقطهٔ x_n اثبات می‌شود .

$$0=f(x^*)=f(x_n)+(x^*-x_n)f'(x_n)+\left(\frac{x^*-x_n}{2}\right)^2 f''(\zeta) \tag{2.20}$$

که‌در آن ζ مقداری است وابسته به x و بین x_n و x^* واقع‌می‌باشد جمله $f(x_n)/f'(x_n)$ در فرمول نیوتن از معادلهٔ (2.16) تعیین گردیده‌و پس از جایگذاری در معادله (2.20) ، رابطهٔ زیر بدست می‌آید .

$$x_{n+1}=x_n+(x^*-x_n)+\frac{(x^*-x_n)^2}{2}\frac{f''(\zeta)}{f'(x_n)} \tag{2.21}$$

بدین‌ترتیب ، به‌شرط آنکه $f'(x_n)$ و $f''(\zeta)$ مخالف صفر باشند ، خطا مضربی از مربع خطای قبلی می‌گردد .

$$x^*-x_{n+1}=-\left(\frac{x^*-x_n}{2}\right)^2\frac{f''(\zeta)}{f'(x_n)} \tag{2.22}$$

۲.۶ ■ روش وتری

مشکل بدیهی که با روش نیوتن ممکن است پیش آید محاسبهٔ $f'(x_n)$ است . برای تابعی نظیر چند جمله‌ای این امر مشکلی پیش نمی‌آورد چون الگوریتمهای ضرب

تو در تو و تقسیم ترکیبی ، شیوه‌های کارآی خودکار برای محاسبه بدست می‌دهند . برای محاسبهٔ مشتق بیشتر توابع بایستی برنامه‌مخصوصی درنظر گرفته شود و در صورت پیچیده بودن تابع مشتق ، ممکن است زمان کامپیوتری قابل توجهی صرف محاسبه آن گردد . اگر برای مشتق تقریبی به صورت :

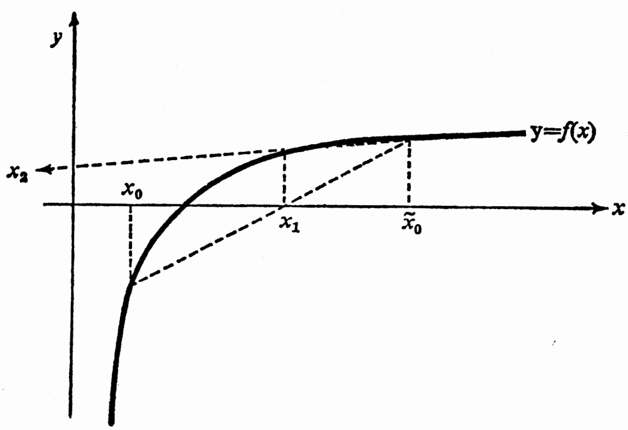
$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \tag{2.23}$$

به کار رود از بروز این مشکل جلوگیری می‌شود .

این تقریب به فرمولی موسوم به روش وتری منتهی می‌شود .

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} \\ &= \frac{x_{n-1}f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})} \end{aligned} \tag{2.24}$$

روش وتری باید با دو مقدار اولیه x_0 و $\tilde{x}_0$ آغاز شود . مقادیر $f(x_0)$ و $f(\tilde{x}_0)$ محاسبه گردد که نمایشگر دو نقطه واقع روی منحنی است . از روی دستور مشاهده می‌شود که نقطهٔ جدید x_1 به وسیله درونیابی خطی بدست می‌آید . این مطلب بوسیلهٔ ترسیم در شکل(۲۰۶)نشان داده شده‌است . نقطهٔ جدید مربوط به رشتهٔ تکرار ، نقطه‌ای است که از محل برخورد وتر حاصل از اتصال دو نقطه قبلی با محور x ها بدست می‌آید .



شکل (۲۰۶) واگرائی روش وتری

در روش وتری نقاط دنباله به طور اکید به کار می‌روند . چنانکه هر نقطه جدید پیدا شود نقطه با کمترین اندیس کنار گذاشته می‌شود . در این روش برای دنباله همانطوریکه در شکل(۲۰۶)نشان داده شده است واگرائی کاملا" امکان‌پذیر می‌باشد ، که در آن x_2

آشکارا نسبت به x_1 دورتر از ریشه است . سرعت همگرائی این روش زمانیکه مقدار تقریب به اندازه کافی به ریشه نزدیک باشد ، بیش از روش تکرار ساده بوده اما از روش نیوتن کمتر است . هر خطای متوالی بستگی به توانی از خطای قبلی بصورت زیر دارد .

$$e_{n+1} = K_n(e_n)^{(1+\sqrt{5})/2} \tag{2.25}$$

کتاب Balfour (1967) و McTernan بحث بیشتری پیرامون سرعت همگرائی ذکر کرده‌اند .

۲.۷ ■ روش نابجائی

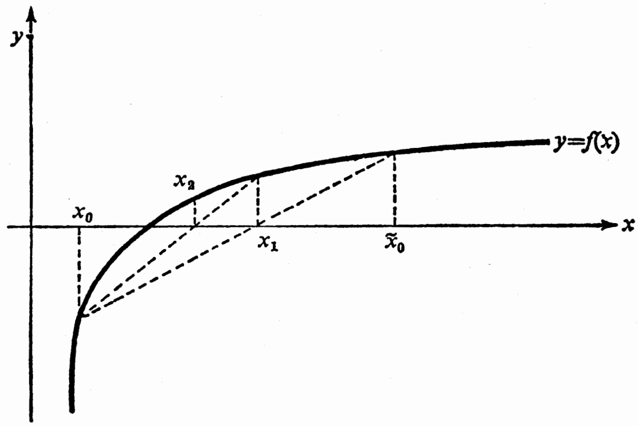
با پیرایش ساده‌ای ، از روش وتری‌روشی‌حاصل می‌گردد که همواره همگرا می‌باشد . در صورتیکه دو تقریب اولیه طوری انتخاب گردد که مقدار تابع در این دو نقطه دارای علامات مخالف باشند ، در اینصورت تولید یک دنباله از مقادیر که همیشه در این خاصیت صدق نماید امکان‌پذیر می‌باشد . مقدار x_1 وسیله فرمول درونیابی خطی پیدا می‌شود و $f(x_1)$ محاسبه می‌گردد . $\bar{x}_1$ یکی از x_0 یا $\bar{x}_0$ انتخاب می‌گردد که مقدار تابع در آن مخالف علامت $f(x_1)$ گردد .

حال مقادیر x_1 و $\bar{x}_1$ مشخص کننده فاصله کوچکتری هستند ده ریشه باید در آن قرار گیرد . این روند در هر مرحله با انتخاب دو مقدار در طرفین ریشه ادامه پیدا می‌کند . همگرائی روش نابجائی برای مسئله‌ای که در شکل ۲.۶ نشان داده شده در شکل ۲.۷ آورده شده است . نتیجهٔ پیرایش روش پیشین این است که دیگر سرعت همگرائی روش وتری قابل اجرا نمی‌باشد .

۲.۸ ■ مقایسه روشها

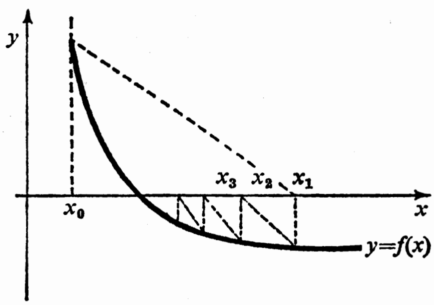
ساده‌ترین روش ، روش دوبخشی است که همواره همگرا می‌باشد . از نقطه‌نظر ضعف دارای سرعت همگرائی کند بوده‌و برای ریشه‌های چندگانه از مرتبه زوج قابل اجرا نمی‌باشد . این روش به عنوان یک روش مقدماتی جهت پیدا کردن تقریبهای خام یک ریشه کاملا" مناسب است . روش دیگری که همواره همگرا است روش نابجائی می‌باشد . این روش تحت شرایط معینی هم‌ارز روش وتری است و بنابراین دارای سرعت همگرائی خوبی است . بهر حال ، شکل ۲.۸ نشانگر وضعیتی است که نتیجه با همگرائی کاملا" کند بدست آمده است . زمانیکه شرایط برای همگرائی برقرار باشد آنگاه روشی نظیر روش وتری یا روش نیوتن به کار خواهد رفت .

در یک برنامه کامپیوتری خودکار ، ترکیبی از یک‌روش همیشه همگرا مانند روش نابجائی با روشی دارای همگرائی سریع ، نظیر روش نیوتن ، مطلوب خواهد بود .



شکل ۲۰۷ روش نابجایی

روش نابجائی فاصله‌ای را تعریف می‌کند که ریشه در آن قرار می‌گیرد، و مقدار بدست آمده بوسیله روش نیوتن‌برای‌نقطه تکراری بعدی در صورتی مورد قبول واقع می‌شود که در این فاصله قرار گیرد. اگر در این فاصله قرار نگرفت مقدار بدست آمده بوسیله روش نابجائی برای نقطه تکراری بعدی منظور می‌گردد. بدین‌ترتیب دو مقدار در دو طرف ریشه باقی می‌ماند و روش نیوتن فقط زمانی به‌کار خواهد رفت که مقداری در این ناحیه بدهد.



شکل ۲۰۸ همگرائی کند روش نابجائی

این روند تنها برای ریشه‌های مکرر از مرتبه فرد مناسب است لیکن زمانی که معادله شامل ریشه‌های چندگانه زوج باشد، ناحیه‌ای وجود دارد که درآن $f(x)f''(x) > 0$ بنابراین اگر یک تقریب به اندازه کافی نزدیک بتوان پیدا کرد، روش نیوتن با همگرائی تضمین شده می‌تواند به‌کار رود. همچنین روشهای دیگری وجود دارد که پیچیده‌تر از روش‌نیوتن بوده، اما سرعت همگرائی آنها بیشتر است. بهرحال، همان مشکل بررسی همگرائی موجود است که با پیچیده‌تر شدن شکل و بیشتر شدن مرتبه همگرائی مشکل‌تر

می‌گردد . بحث مفصل درباره سایر روشها در Traub (1964) داده شده است .

۲.۹ ■ دستگاه معادلات غیرخطی

مطالعهٔ کامل حل دستگاه معادلات موضوعی است که نیازمند تحلیل قابل ملاحظه‌ای بوده و برای مطالعهٔ غیر متخصصین مناسب نمی‌باشد . بهرحال ، بعضی از روشهای بخش پیشین می‌تواند به دستگاه معادلات تعمیم داده شود ، اگرچه تحلیل خواص همگرائی ساده نمی‌باشد . بطور مثال ، روش نیوتن برای دستگاه غیر خطی دو معادله با دو مجهول توسعه داده شده است .

$$\begin{aligned} f(x,y) &= 0 \\ g(x,y) &= 0 \end{aligned} \tag{2.26}$$

فرض می‌کنیم x_0 و y_0 دو تقریب اولیه برای جواب بوده و Δx و Δy دو نمو برای رسیدن به جواب واقعی باشند . آنگاه پس از بسط تابع دو متغیره به وسیلهٔ سری تیلر داریم :

$$\begin{aligned} 0 &= f(x_0+\Delta x, y_0+\Delta y) \\ &= f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\,\Delta x + \frac{\partial f}{\partial y}(x_0,y_0)\,\Delta y + \cdots \\ 0 &= g(x_0+\Delta x, y_0+\Delta y) \\ &= g(x_0,y_0) + \frac{\partial g}{\partial x}(x_0,y_0)\,\Delta x + \frac{\partial g}{\partial y}(x_0,y_0)\,\Delta y + \cdots \end{aligned} \tag{2.27}$$

اگر جملات مرتبه دوم صرفنظر شوند و ضرائب Δx ، Δy با :

$$\partial g/\partial y = a_{22}\ ,\ \partial g/\partial x = a_{21},\ \partial f/\partial y = a_{12},\ \partial f/\partial x = a_{11},$$

مشخص گردند دو معادلهٔ زیر را خواهیم داشت :

$$\begin{aligned} a_{11}\,\Delta x_0 + a_{12}\,\Delta y_0 &= -f(x_0,y_0) \\ a_{21}\,\Delta x_0 + a_{22}\,\Delta y_0 &= -g(x_0,y_0) \end{aligned} \tag{2.28}$$

این معادلات نسبت به نموهای Δx_0 و Δy_0 برای بدست آوردن تقریبهای جدید $x_1 = x_0 + \Delta x_0$ و $y_1 = y_0 + \Delta y_0$ قابل حل می‌باشند . روند با به کار بردن مقادیر x_1 و y_1 در سمت راست و مشتقات جزئی برای پیدا کردن مقادیر جدید تکرارمی‌گردد .

$$\begin{aligned} x_{r+1} &= x_r + \Delta x_r, \\ & \qquad\qquad r = 0,1,2,\ldots \\ y_{r+1} &= y_r + \Delta y_r, \end{aligned} \tag{2.29}$$

در حالت عمومی‌تر مجموعه‌ای از n تابع n متغیره داریم :

$$f^{(r)}=f_1[x_1^{(r)},x_2^{(r)},\ldots,x_n^{(r)}]$$
$$f_2[x_1^{(r)},x_2^{(r)},\ldots,x_n^{(r)}] \qquad (2.30)$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$f_n[x_1^{(r)},x_2^{(r)},\ldots,x_n^{(r)}]$$

ماتریس A با عناصر

$$a_{ij}=\frac{\partial f_i}{\partial x_i} \qquad (2.31)$$

تشکیل یافته و نموهای $\Delta \mathbf{x}^{(r)}$ با حل متوالی مجموعه معادلات زیر بدست می‌آید.

$$\mathbf{A}\,\Delta\mathbf{x}^{(r)}=-\mathbf{f}^{(r)}, \qquad r=0,1,\ldots \qquad (2.32)$$

مثال عددی ساده‌ای از این نوع در مثال ۲.۴ داده شده است.

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ مثال زیر کندی همگرائی را برای روش دوبخشی نشان می‌دهد. مطلوبست حل معادله $1/x+1=0$ با روشی دوبخشی.

مقادیر اولیه عبارتند از:

x	$-0{\cdot}5$	$-4{\cdot}0$
$f(x)$	$-1{\cdot}0$	$0{\cdot}75$

دنبالهٔ تکرارها عبارتند از:

x	$-2{\cdot}25$	$-1{\cdot}375$	$-0{\cdot}9375$	$-1{\cdot}115\,625$	$-1{\cdot}046\,875$
$f(x)$	$0{\cdot}555\,556$	$0{\cdot}272\,728$	$-0{\cdot}066\,666$	$0{\cdot}135\,136$	$0{\cdot}044\,770$
x	$-0{\cdot}992\,187$	$-1{\cdot}019\,531$	$-1{\cdot}005\,859$	$-0{\cdot}999\,023$	$-1{\cdot}002\,929$
$f(x)$	$-0{\cdot}007\,874$	$0{\cdot}019\,157$	$0{\cdot}005\,825$	$-0{\cdot}000\,977$	$0{\cdot}002\,921$
x	$-0{\cdot}999\,515$	$-1{\cdot}001\,222$	$-1{\cdot}000\,368$		
$f(x)$	$-0{\cdot}000\,485$	$0{\cdot}001\,221$	$0{\cdot}000\,368$		

اعداد بدست آمده از روش نیوتن عبارتند از:

x	$-0{\cdot}75$	$-0{\cdot}937\,500$	$-0{\cdot}996\,094$	$-0{\cdot}999\,850$	$-1{\cdot}000\,000$
$f(x)$	$-0{\cdot}333\,333$	$-0{\cdot}066\,666$	$-0{\cdot}003\,921$	$-0{\cdot}000\,150$	$0{\cdot}000\,000$

سرعت همگرائی در روش نیوتن افزایش بسیاری‌می‌یابد.

ــ مسئله نوسان که با روش نیوتن می‌تواند رخ دهد بطور وضوح با درنظر گرفتن کاربرد این روش در مورد معادلهٔ زیر نشان داده می‌شود.

$$x^4-8x^3+23x^2-28x+187/16=0$$

در حقیقت، معادله دارای ریشه‌هایی نزدیک به $x=1$ و $x=3$ بوده اما با شروع تقریب آغازی $x_0=1\frac{1}{2}$ داریم:

$$x_1 = 1\tfrac{1}{2} - f(1\tfrac{1}{2})/f'(1\tfrac{1}{2}) = 2\tfrac{1}{2}$$
$$x_2 = 2\tfrac{1}{2} - f(2\tfrac{1}{2})/f'(2\tfrac{1}{2}) = 1\tfrac{1}{2}$$

و روند بین مقادیر $1\tfrac{1}{2}$ و $2\tfrac{1}{2}$ نوسان می‌کند.

۳ ــ زمانیکه ریشه‌ای چندگانه موجود باشد در تعیین مقدار تقریبی نزدیک به این ریشه، به کمک روش نیوتن، مشکلاتی وجود دارد. این امر به واسطهٔ محاسبه جملات $f(x)$ نزدیک به ریشه است که اغلب حاوی تفاضل دو عدد تقریبا" مساوی می‌باشد. خطای حاصل از این عمل، با تقسیم به مقدار $f'(x)$ که به علت وجود ریشه تکراری نزدیک به صفر است، بزرگ می‌گردد.

جدول زیر نتیجهٔ محاسبات، تا ۶ رقم اعشار، برای یافتن یک ریشهٔ چندجمله‌ای زیر است.

$$f(x) = (x-1)^2(x-2)(x+3) = x^4 - x^3 - 7x^2 + 13x - 6$$
$$f'(x) = 4x^3 - 3x^2 - 14x + 13$$

x	$f(x)$	$f'(x)$	Δx
0·5	1·3125	5·75	0·228 260
0·728 260	−0·350 116	2·758 241	0·126 934
0·855 194	−0·092 544	1·335 022	0·069 320
0·924 514	−0·024 050	0·653 451	0·036 810
0·961 324	−0·006 155	0·322 639	0·019 077
0·980 401	−0·001 558	0·160 219	0·009 724
0·990 125	−0·000 393	0·079 874	0·004 920
0·995 045	−0·000 098	0·039 861	0·002 458
0·997 503	−0·000 024	0·020 032	0·001 198
0·998 701	−0·000 007	0·010 407	0·000 672
0·999 373	−0·000 001	0·005 019	0·000 199
0·999 572	−0·000 000	0·003 425	

در اینجا خطاهای ممکن از مرتبه 5×10^{-7} است که به مقدار تقریبی 5×10^{-4} تقسیم گردیده، این محاسبات خطاهائی را ارائه می‌کند که حداکثر 10^{-4} هستند. بنابراین، حتما" ممکن است سومین رقم اعشار مناسب نباشد، که به طور یقین رقم چهارم نیز قابل اعتماد نیست. این مطلب با توجه به مقدار واقعی $x = 1{\cdot}0$ تائید می‌گردد.

زمانیکه تشخیص داده شود که مقدار تابع در دامنه $0{\cdot}999\,500 \leqslant x \leqslant 1{\cdot}000\,501$، تا ۶ رقم اعشار، صفر می‌شود نتیجه غیر منتظره نخواهد بود.

۴ ــ روش حل دستگاه معادلات غیر خطی وسیلهٔ مثال ساده‌ای با دو معادلهٔ دومجهولی نشان داده می‌شود. این مثال پیگیری روند را برای خواننده آسان خواهد کرد.

معادلات زیر به وسیله روش نیوتن با تقریب آغازی $(x_0, y_0) = (3, 4)$ حل می‌شوند.

$$f(x, y) = x^2 + y^2 - 25 = 0$$
$$g(x, y) = x^2 - y^2 - 7 = 0$$

معادلات مربوط به روش نیوتن برای این دستگاه عبارتنداز:

$$2x\,\Delta x_r + 2y\,\Delta y_r = -f(x_r, y_r)$$
$$2x\,\Delta x_r - 2y\,\Delta y_r = -g(x_r, y_r)$$

بدین‌ترتیب

$$6\,\Delta x_0 + 8\,\Delta y_0 = 0$$
$$6\,\Delta x_0 - 8\,\Delta y_0 = +14$$
$$\Delta x_0 = +\tfrac{14}{12} = 1{\cdot}166\,667$$
$$\Delta y_0 = -\tfrac{14}{12} = -1{\cdot}166\,667$$
$$x_1 = 4{\cdot}166\,667, \qquad y_1 = 2{\cdot}833\,335$$

و

$$8{\cdot}333\,334\,\Delta x_1 + 5{\cdot}666\,670\,\Delta y_1 = -0{\cdot}388\,887$$
$$8{\cdot}333\,334\,\Delta x_1 - 5{\cdot}666\,670\,\Delta y_1 = -2{\cdot}333\,335$$
$$\Delta x_1 = -0{\cdot}163\,333, \qquad \Delta y_1 = 0{\cdot}171\,569$$
$$x_2 = 4{\cdot}003\,334, \qquad y_2 = 3{\cdot}004\,904$$

و

$$8{\cdot}006\,668\,\Delta x_2 + 6{\cdot}009\,808\,\Delta y_2 = -0{\cdot}056\,131$$
$$8{\cdot}006\,668\,\Delta x_2 - 6{\cdot}009\,808\,\Delta y_2 = 0{\cdot}002\,765$$
$$\Delta x_3 = -0{\cdot}003\,160, \qquad \Delta y_3 = -0{\cdot}004\,900$$

بنابراین

$$x_4 = 4{\cdot}000\,174, \qquad y_4 = 3{\cdot}000\,004$$

و می‌توان دید که روند به یکی از جوابهای درست همگرا می‌گردد.

۵ ــ ریشهٔ مثبت معادلهٔ $x^3 - x - 1 = 0$ را با الگوریتم‌های دوبخشی، نابجائی پیراسته، وتری و نیوتن با دقت 10^{-7} به کمک برنامه کامپیوتری پیدا کنید.

* الگوریتم (۲۰۱) روش دوبخشی

برای تابع $f(x)$ که در بازهٔ $[a_0,b_0]$ پیوسته باشد به طوری که $f(a_0)f(b_0) \le 0$ داریم.

برای $n = 0, 1, 2, \ldots$

قرار می‌دهیم $m = (a_n + b_n)/2$

اگر $f(a_n)f(m) \le 0$ قرار می‌دهیم $a_{n+1} = a_n,\ b_{n+1} = m$

در غیر اینصورت $a_{n+1} = m,\ b_{n+1} = b_n$

آنگاه $f(x)$ دارای یک ریشه در فاصله $[a_{n+1},b_{n+1}]$ است.

الگوریتم فوق را ذیلا به صورت زیر برنامه کامپیوتری ارائه می‌کنیم.

FORTRAN SUBROUTINE FOR THE BISECTION ALGORITHM 2.1

```
      SUBROUTINE BISECT(F,A,B,XTOL,IFLAG)
      IFLAG = 0
      N = -1
      FA = F(A)
CHECK FOR SIGN CHANGE
      IF (FA*F(B) .LE. 0.)                     GO TO 5
      IFLAG = 2
      WRITE (6,601) A,B
  601 FORMAT (43H F(X) IS OF SAME SIGN AT THE TWO ENDPOINTS
     *           2E15.7)
                                               RETURN
    5 ERROR = ABS(B - A)
    6 ERROR = ERROR/2.
CHECK FOR SUFFICIENTLY SMALL INTERVAL
      IF (ERROR .LE. XTOL)                     RETURN
      XM = (A + B)/2.
CHECK FOR UNREASONABLE ERROR REQUIREMENT
      IF (XM + ERROR .EQ. XM)                  GO TO 20
      FM = F(XM)
C  TEMPORARY PRINTOUT
      N = N + 1
      WRITE (6,606) N,A,XM,B,FA,FM
  606 FORMAT(I3,9H X-VALUES3E15.7/3X9H F-VALUES2E15.7/)
CHANGE TO NEW INTERVAL
      IF (FA*FM .LE. 0.)                       GO TO 9
      A = XM
      FA = FM
                                               GO TO 6
    9 B = XM
                                               GO TO 6
   20 IFLAG = 1
                                               RETURN
      END
```

سپس مسئله مورد نظر را با برنامهٔ زیر، و استفاده از زیربرنامه فوق، حل می‌کنیم.

```
    EXTERNAL FF
    A = 1.
    B = 2.
    CALL BISECT(FF,A,B,1.E-6,IFLAG)
    IF (IFLAG .GT. 1)                    STOP
    XI = (A + B)/2.
    ERROR = ABS(A - B)/2.
    WRITE (6,600) XI,ERROR
600 FORMAT(13H THE ROOT IS E15.7,12H PLUS/MINUS E15.7)
                                         STOP
    END
    FUNCTION FF(X)
    FF = -1. - X*(1. - X*X)
                                         RETURN
    END
```

* الگوریتم (۲۰۲) روش نابجائی

برای تابع $f(x)$ که در بازهٔ $[a_0,b_0]$ پیوسته است به طوری که $f(a_0)f(b_0) < 0$

داریم .

$n = 0, 1, 2, \ldots$

$w = [f(b_n)a_n - f(a_n)b_n]/[f(b_n) - f(a_n)]$

اگر $f(a_n)f(w) \leq 0$ قرار می‌دهیم $a_{n+1} = a_n, b_{n+1} = w$

در غیر اینصورت قرار می‌دهیم $a_{n+1} = w, b_{n+1} = b_n$

* الگوریتم (۲۰۳) روش نابجائی پیراسته

برای تابع $f(x)$ که در بازهٔ $[a_0,b_0]$ پیوسته به طوری که $f(a_0)f(b_0) < 0$ باشد داریم

قرار می‌دهیم $F = f(a_0), G = f(b_0), w_0 = a_0$

برای $n = 0, 1, 2, \ldots$

$w_{n+1} = (Ga_n - Fb_n)/(G - F)$

اگر $f(a_n)f(w_{n+1}) \leq 0$ قرار می‌دهیم

$a_{n+1} = a_n, b_{n+1} = w_{n+1}, G = f(w_{n+1})$

اگر $f(w_n)f(w_{n+1}) > 0$ قرار می‌دهیم $F = F/2$

در غیر اینصورت $a_{n+1} = w_{n+1}, F = f(w_{n+1}), b_{n+1} = b_n$

در صورتیکه $f(w_n)f(w_{n+1}) > 0$ قرار می‌دهیم $G = G/2$

آنگاه $f(x)$ دارای یک ریشه در فاصلهٔ $[a_{n+1},b_{n+1}]$ خواهد بود .

الگوریتم های (۲۰۲) و (۲۰۳) در شکل‌های (۲۰۹) و (۲۰۱۰) و برنامه کامپیوتری الگوریتم فوق در برنامه (۲۰۳) آمده است .

* الگوریتم (۲۰۴) روش وتری

برای تابع $f(x)$ و دو نقطهٔ x_{-1}, x_0 داریم :

برای $n = 0, 1, 2, \ldots$

$x_{n+1} = [f(x_n)x_{n-1} - f(x_{n-1})x_n]/[f(x_n) - f(x_{n-1})]$

برنامه کامپیوتری الگوریتم فوق در برنامه (۲۰۴) آمده است .

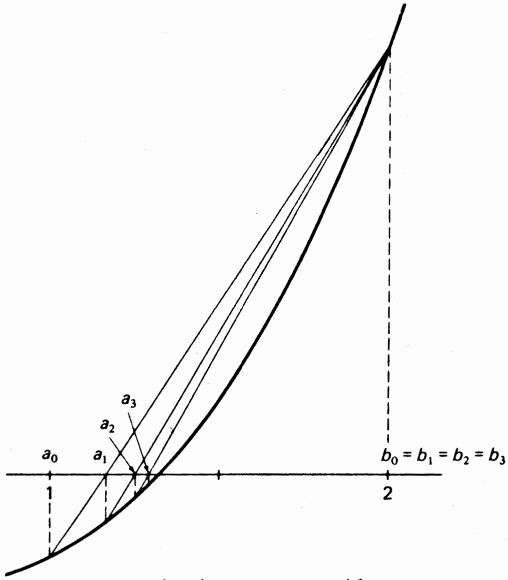
* الگوریتم (۲۰۵) روش نیوتن

برای تابع $f(x)$ که در نقطهٔ x_0 پیوسته و مشتق‌پذیر باشد الگوریتم زیررا داریم :

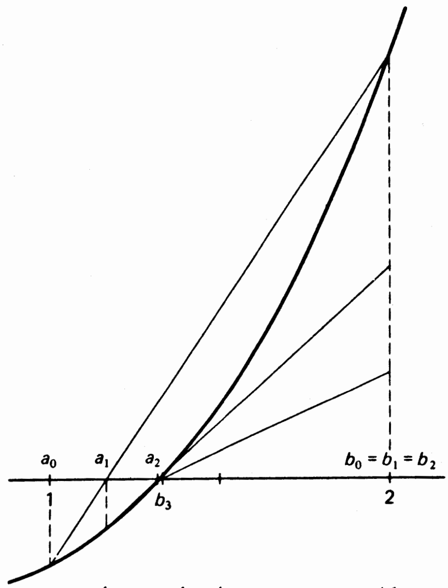
برای $n = 0, 1, 2, \ldots$

$x_{n+1} = x_n - f(x_n)/f'(x_n)$

و برنامه کامپیوتری الگوریتم فوق در برنامه ۲.۵ آمده است .



شکل ۲.۹ روش نابجائی



شکل ۲.۱۰ روش نابجائی پیراسته

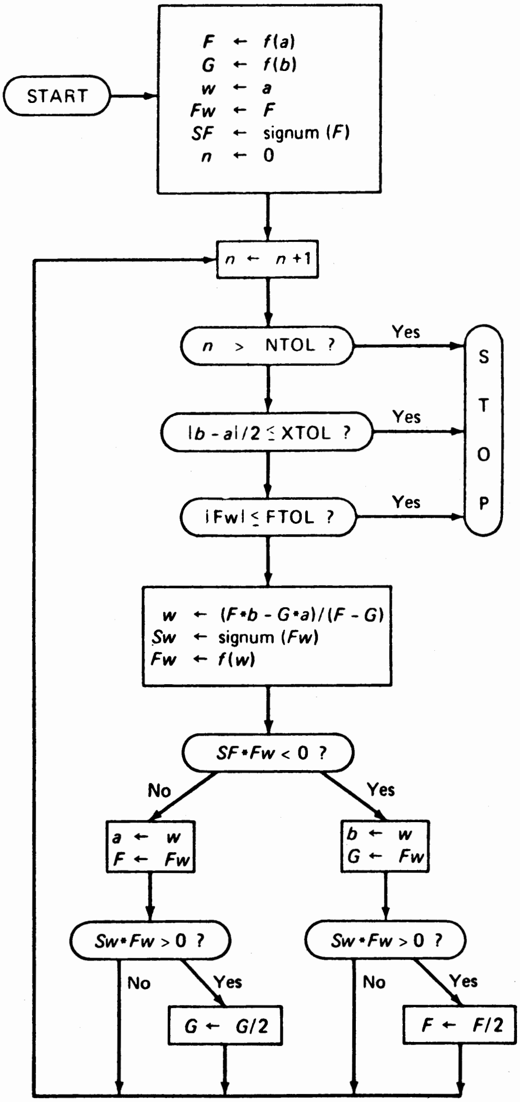
FORTRAN PROGRAM USING THE MODIFIED REGULA FALSI ALGORITHM 2.3

```
EXTERNAL FG
A = 1.
B = 2.
```

```
      CALL MRGFLS(FG,A,B,1.E-7,1.E-10,30,IFLAG)
      IF (IFLAG .GT. 2)                         STOP
      XI = (A + B)/2.
      ERROR = ABS(B - A)/2.
      FXI = FG(XI)
      WRITE (6,600) XI, ERROR, FX1
  600 FORMAT (13H THE ROOT IS E15.7,12H PLUS/MINUS E15.7/
     *         11H FG(ROOT) = E15.7)
                                                STOP
      END
      FUNCTION FG(X)
      FG = -1. - X*(1. - X*X)
                                                RETURN
      END
      SUBROUTINE MRGFLS(F,A,B,XTOL,FTOL,NTOL,IFLAG)
      IFLAG = 0
      FA = F(A)
      SIGNFA = SIGN(1.,FA)
      FB = F(B)
CHECK FOR SIGN CHANGE
      IF (SIGNFA*FB .LE. 0.)                    GO TO 5
      IFLAG = 3
      WRITE (6,601) A,B
  601 FORMAT(43H F(X) IS OF SAME SIGN AT THE TWO ENDPOINTS
     *         2E15.7)
                                                RETURN
    5 W = A
      FW = FA
      DO 20 N = 1,NTOL
CHECK FOR SUFFICIENTLY SMALL INTERVAL
      IF (ABS(B - A)/2. .LE. XTOL)              RETURN
CHECK FOR SUFFICIENTLY SMALL FUNCTION VALUES
      IF (ABS(FW) .GT. FTOL)                    GO TO 9
      A = W
      B = W
      IFLAG = 1
                                                RETURN
    9 W = (FA*B - FB*A)/(FA - FB)
      PREVFW = SIGN(1.,FW)
      FW = F(W)
C  TEMPORARY PRINT OUT
      NM1 = N - 1
      WRITE (6,609) NM1,A,W,B,FA,FW,FB
  609 FORMAT(I3, 9H X-VALUES3E16.8/3X9H F-VALUES3E16.8/)
CHANGE TO NEW INTERVAL
      IF (SIGNFA*FW .LT. 0.)                    GO TO 10
      A = W
      FA = FW
      IF (FW*PREVFW .GT. 0.) FB = FB/2.
                                                GO TO 20
   10 B = W

   FB = FW
   IF (FW*PREVFW .GT. 0.) FA = FA/2.
20 CONTINUE
   IFLAG = 2
   WRITE (6,620) NTOL
```

```
620 FORMAT(18H NO CONVERGENCE IN I5,11H ITERATIONS)
                                             RETURN
    END
```



FORTRAN PROGRAM USING THE SECANT ALGORITHM 2.4

```
      EXTERNAL FF
      X0 = 1.
      X1 = 2.
      CALL SECANT(FF,X0,X1,1.E-7,1.E-7,30,IFLAG)
      IF (IFLAG .GT. 1)                STOP
      FX1 = FF(X1)
      WRITE (6,600) X1,FX1
```

```
   600 FORMAT(13H THE ROOT IS E15.7,11H F(ROOT) = E15.7)
                                             STOP
       END
       SUBROUTINE SECANT(F,X0,X1,XTOL,FTOL,NTOL,IFLAG)
       IFLAG = 0
       F0 = F(X0)
       DELTAX = X1 - X0
       DO 20 N = 1,NTOL
       F1 = F(X1)
C   TEMPORARY PRINT OUT
       WRITE (6,600) N,X0,X1,F0,F1
   600 FORMAT (I5,9H X-VALUES 2E16.8/5X9H F-VALUES 2E16.8/)
       IF (ABS(F1) .LE. FTOL)                GO TO 30
       DELTAF = F0 - F1
       IF (DELTAF .EQ. 0.)                   GO TO 999
       DELTAX = F1/DELTAF*DELTAX
       X0 = X1
       X1 = X1 + DELTAX
       IF (ABS(DELTAX) .LE. XTOL)            RETURN
    20 F0 = F1
   999 IFLAG = 2
                                             RETURN
    30 IFLAG = 1
                                             RETURN
       END
       FUNCTION FF(X)
       FF = -1. - X*(1. - X*X)
                                             RETURN
       END
```

FORTRAN PROGRAM USING NEWTON'S METHOD, ALGORITHM 2.5

```
       EXTERNAL FF,FFDER
       X0 = 1.
       CALL NEWTON(FF,FFDER,X0,1.E-7,1.E-7,10,IFLAG)
       IF (IFLAG .GT. 1)                     STOP
       FX0 = FF(X0)
       WRITE (6,600) X0,FX0
   600 FORMAT(13H THE ROOT IS E15.7,11H F(ROOT) = E15.7)
                                             STOP
       END
       SUBROUTINE NEWTON(F,FDERIV,X0,XTOL,FTOL,NTOL,IFLAG)
       IFLAG = 0
       DO 20 N = 1,NTOL
       FX0 = F(X0)
       IF (ABS(FX0) .LT. FTOL)               RETURN
       DERIV = FDERIV(X0)
C   TEMPORARY PRINT OUT
       WRITE (6,600) N,X0,FX0,DERIV
   600 FORMAT(I5,16H X0, FX0, DERIV 3E15.7)
       IF (DERIV .EQ. 0.)                    GO TO 999
       DELTAX = FX0/DERIV
       X0 = X0 - DELTAX
    20 IF (ABS(DELTAX) .LT. XTOL)            RETURN
```

```
999 IFLAG = 2
                                            RETURN
    END
    FUNCTION FF(X)
    FF = -1. - X*(1. - X*X)
    RETURN
    END
    FUNCTION FFDER(X)
    FFDER = 3.*X*X - 1.
    RETURN
    END
```

نتایج کامپیوتری مسئله ۵ در جدول (۲۰۲) آمده است .

جدول ۲۰۲

n	Bisection x_n	ε_n	Modified regula falsi x_n	ε_n	Secant x_n	Newton x_n
					1.0	
0	1.5	$5\cdot10^{-1}$	1.5	$5\cdot10^{-1}$	2.0	1.0
1	1.25	$3\cdot10^{-1}$	1.5833333	$4\cdot10^{-1}$	1.1666667	1.5
2	1.375	$1\cdot10^{-1}$	1.6616541	$3\cdot10^{-1}$	1.2531120	1.3478261
3	1.3125	$6\cdot10^{-2}$	1.3249256	$2\cdot10^{-3}$	1.3372064	1.3252004
4	1.34375	$3\cdot10^{-2}$	1.3256293	$9\cdot10^{-4}$	1.3238501	1.3247182
5	1.328125	$2\cdot10^{-2}$	1.3256305	$9\cdot10^{-4}$	1.3247079	1.3247180
6	1.3203125	$8\cdot10^{-3}$	1.3247180	$4\cdot10^{-9}$	1.3247180	
...	....	......				
10	1.3247070	$5\cdot10^{-4}$				
...	....	......				
20	1.3247180	$5\cdot10^{-7}$				

۶ ــ با استفاده از برنامه کامپیوتری و به کمک روش نیوتن دستگاه زیر را با دقت 10^{-8} با مقادیر اولیه $x_0 = 3.4$ $y_0 = 2.2$ و بار دیگر با مقادیر اولیه $y_0 = -2$ و $x_0 = 1$ بدست آورید .

$$x + 3\log x - y^2 = 0$$
$$2x^2 - xy - 5x + 1 = 0$$

برنامه کامپیوتری در الگوریتم ۲۰۶ و نتایج در جدول ۲۰۳ آمده است .

ALGORITHM2.6

FORTRAN PROGRAM FOR EXAMPLE 2.6

```
C       EXAMPLE 2.17 BY NEWTON METHOD
C       TO SOLVE F(X,Y) = 0 AND G(X,Y) = 0
        F(X,Y) = X + 3.*ALOG10(X) - Y**2
        G(X,Y) = 2.*X**2 - X*Y - 5.*X + 1.
        FX(X,Y) = 1. + 3./(ALOG10.)*X)
```

```
      FY(X,Y) = -2.*Y
      GX(X,Y) = 4.*X - Y - 5.
      GY(X,Y) = -X
      WRITE(6,6)
   10 READ(5,2) X,Y
      A = F(X,Y)
      B = G(X,Y)
      WRITE(6,3)X,Y,A,B
      DO 1 I = 1, 20
      DELTAX = (-F(X,Y)*GY(X,Y)
     *+ G(X,Y)*FY(X,Y))/(FX(X,Y)*GY(X,Y) - FY(X,Y)*GX(X,Y))
      DELTAY = (-G(X,Y)*FX(X,Y)
     *+ F(X,Y)*GX(X,Y))/(FX(X,Y)*GY(X,Y) - FY(X,Y)*GX(X,Y))
      X = X + DELTAX
      Y = Y + DELTAY
      A = F(X,Y)
      B = G(X,Y)
      WRITE(6,4) I,X,Y,A,B
      IF(AMAX1(ABS(DELTAX),ABS(DELTAY),ABS(A),ABS(B)).LT.1.E-7)
     *                          GO TO 10
    1 CONTINUE
      WRITE(6,5)
      GO TO 10
    2 FORMAT(2E20.7)
    3 FORMAT(//4X1HI8X 1HX 16X 1HY 15X 6HF(X,Y) 12X
     16HG(X,Y) /4X1H04(1PE17.7))
    4 FORMAT(I5, 4(1PE17.7))
    5 FORMAT(36H0FAILED TO CONVERGE IN 20 ITERATIONS)
    6 FORMAT(28H EXAMPLE 2.17, NEWTON METHOD)
      END
```

جدول ۲۰۳

COMPUTER RESULTS FOR EXAMPLE 2.6

I	X	Y	F(X,Y)	G(X,Y)
0	3.4000000E 00	2.2000000E 00	1.5443677E−01	−3.6000001E−01
1	3.4899123E 00	2.2633644E 00	−4.4626594E−03	1.0471106E−02
2	3.4874446E 00	2.2616299E 00	−3.2186508E−06	7.8678131E−06
3	3.4874428E 00	2.2616286E 00	5.9604645E−08	−0.
4	3.4874428E 00	2.2616286E 00	5.9604645E−08	−0.

I	X	Y	F(X,Y)	G(X,Y)
0	1.0000000E 00	−2.0000000E 00	−3.0000000E 00	−0.
1	1.4759726E 00	−1.5240274E 00	−3.3945185E−01	2.2654986E−01
2	1.4573549E 00	−1.4011608E 00	−1.5200734E−02	2.9807091E−03
3	1.4588864E 00	−1.3967715E 00	−1.9967556E−05	−2.0265579E−06
4	1.4588902E 00	−1.3967670E 00	−1.4901161E−08	−1.1920929E−07
5	1.4588902E 00	−1.3967670E 00	0.	−0.

■ مسائل

۱ــ معادله $2x^2 - x = 0$ وسیله روش تکرار ساده حل می‌شود .

$$x_{n+1} = 2x_n^2$$

بنابراین مشتق $g(x)$ مساوی $4x$ می‌شود . مقدار تقریب اولیه ، به اشتباه

$$x_0 = \tfrac{1}{2}$$

اختیار شده ، که در آن مشتق مساوی

$$4.\tfrac{1}{2} = 2$$

است . بنابراین روند نباید همگرا باشد . اما مقادیر متوالی عبارتنداز $\tfrac{1}{2}, \tfrac{1}{2}, \ldots$.
بکمک شکل (c) ۲۰۱ چگونگی را بیان نمائید .

۲ ــ یک ریشهٔ معادلهٔ

$$x = -\log x$$

را به کمک تکرار ساده با مقدار تقریب آغازی

$$x_0 = 1{\cdot}0$$

تعیین نمائید . به منظور رسیدن به شرایط همگرائی ، تکرار چه صورتی باید داشته باشد ؟

۳ ــ مطلوبست تعیین کوچکترین ریشهٔ مثبت

$$\tan y = e^y$$

۴ ــ یک ریشهٔ معادلهٔ

$$2x^3 + x^2 - 20x + 12 = 0$$

بین $x = 0$ و $x = 1$ قرار دارد . با به کار بردن روش نابجائی اولین تقریب را بدست آورید و سپس با روش نیوتن تقریب را بهتر کنید .

۵ــ روش نیوتن را با تقریب اولیه $x = 0{\cdot}0$ در معادله $x^3 - 6{\cdot}5x^2 + 14{\cdot}0x - 10{\cdot}0 = 0$ به کار برید . فرض می‌کنیم تکرار زمانیکه تفاضل مقادیر متوالی x_r کمتر از 10^{-10} باشد متوقف گردد . جواب درست $x = 2{\cdot}0$ است . اگر پاسخ شما بطور قابل توجهی با این مقدار فرق داشته باشد ، دلائلی برای این اختلاف ذکر کنید . آیا روش نابجائی جواب بهتری می‌دهد ؟ (کامپیوتر به کار رفته باید دارای مانتیسی باشد که لااقل از ۳۷ رقم دوتایی تشکیل شده باشد .)

۶ ــ معادلهٔ $x^3 - x - 1 = 0$ مفروض است .

در صورتیکه معادلهٔ فوق دارای یک ریشهٔ حقیقی بین ۱ و ۲ باشد ، مطلوبست محاسبهٔ مقدار تقریبی این ریشه به کمک معادله تکراری مناسب تا سه رقم اعشار درست .

۷ ــ مطلوبست تعیین کوچکترین ریشهٔ مثبت معادلهٔ $x^3 + x = 1000$ تا ۴ رقم اعشار درست .

۸ ــ معادلهٔ $x^3 - 2x - 1 = 0$ مفروض است .

درصورتیکه یک ریشهٔ حقیقی معادله بین ۱ و ۲ باشد مقادیر تقریبی این ریشه را به کمک معادله تکراری مناسب با تقریب 10^{-4} پیدا کنید .

۹ ــ معادله $x^3 + 1.9x^2 - 1.3x - 2.2 = 0$ دارای یک ریشهٔ نزدیک به ۱ است . معادلهٔ فوق را به شکل $x = g(x)$ نوشته و با انتخاب $x_0 = 1$ مقدار تقریبی این ریشه را پیدا کنید (معیار دقت 10^{-3}) .

۱۰ ــ محل تقریبی ریشه‌های حقیقی $\sin x - x + 0.5 = 0$ را با رسم نمودار تعیین کنید . سپس با انتخاب مقدار تقریبی مناسبی این ریشه‌ها را تعیین کنید . (معیار دقت 10^{-3}) .

۱۱ ــ محل تقریبی ریشه‌های حقیقی $(x-3)(x-2)x = e^{-0.2x}$ را با رسم نمودار تعیین کرده و سپس به کمک روش نیوتن مقدار تقریبی کوچکترین ریشهٔ حقیقی را تا ۴ رقم اعشار درست پیدا کنید .

۱۲ ــ به کمک روش نیوتن ریشهٔ پنجم و هفتم عدد ۸ را با دقت 10^{-8} محاسبه کنید .

۱۳ ــ در صورتیکه معادلهٔ $f(x) = 0$ دارای یک ریشهٔ تکراری از مرتبه K باشد فرمول پیراستهٔ روش نیوتن در این حالت چیست ؟ و علت اینکه سرعت همگرائی تندتر می‌شود را با ذکر دلیل بیان کنید .

۱۴ ــ نشان دهید رابطهٔ خطاها در روش وتری به صورت زیر خواهد بود .

$$e_{n+1} \approx e_n e_{n-1} \frac{-\frac{1}{2}f''(\xi)}{f'(\xi)}$$

که در آن $e_n = x_n - S$ و S مقدار واقعی ریشه و ξ مقداری است در همسایگی ریشه .

۱۵ ــ ریشهٔ مثبت معادلهٔ $x^2 = \sin x$ را تا سه رقم اعشار درست به کمک روش نیوتن محاسبه نمائید .

۱۶ ــ در هریک از معادلات زیر با استفاده از رسم نمودار مقدار اولیه مناسبی یافته و کوچکترین ریشهٔ مثبت هریک را به کمک روش نیوتن با دقت 10^{-3} محاسبه نمائید .

$$4\cos x - e^x = 0$$

$$2\cos x - \cosh x = 0$$

$$e^{-x^2} - \cos x = 0$$

۱۷ــ معادلهٔ $x^3 - 2x^2 + 2x - 1 = 0$ مفروض است .
با استفاده از برنامه کامپیوتری ، به کمک روش نیوتن ، با انتخاب مقادیر اولیه $x_0 = 0$ و $x_0 = 10$ ریشهٔ معادلهٔ فوق را با معیار توقف $|f(x_n)| \leqslant 10^{-6}$ پیدا کرده و رابطهٔ خطا را در این حالت تحقیق نمائید .

۱۸ــ با استفاده از روش وتری کوچکترین ریشهٔ مثبت معادلهٔ زیر را با دقت 10^{-7} محاسبه نمائید .

$$e^{-x^2} - \log_{10} x = 0$$

۱۹ــ معادلهٔ $x - 0.2 \sin x - 0.5 = 0$ دارای یک ریشهٔ مثبت بین 0.5 و 1 میباشد .
با به‌کار بردن روشهای ذیل مقدار تقریبی این ریشه را با دقت 10^{-7} با استفاده از برنامه کامپیوتری محاسبه نمائید .

الف : دوبخشی

ب : نابجائی

ج : وتری

د : نیوتن .

۲۰ــ معادلهٔ $3x^3 + 14x^2 + 13x - 6 = 0$ مفروض است .
در صورتیکه این معادله دارای سه ریشهٔ حقیقی در نزدیکی 3- ، 2- و $\frac{1}{3}$ باشد .
به کمک برنامه کامپیوتری مقدار تقریبی هریک از سه ریشهٔ فوق را با روشهای زیر با دقت 10^{-6} محاسبه نمائید .

الف تکرار ساده

ب : وتری

ج : نابجائی

د : نیوتن

۲۱ــ مطلوبست محاسبهٔ ریشهٔ مضاعف معادلهٔ $x^3 - x^2 - x + 1 = 0$ به کمک روش پیراسته نیوتن با دقت 10^{-5}.

۲۲ــ دستگاههای غیرخطی زیر را ، با استفاده از روش نیوتن به کمک برنامه کامپیوتری ، با مقادیر اولیه داده‌شده و معیار دقت 10^{-6} حل کنید .

$$(a) \quad \begin{aligned} f(x,y) &= x^2 + y^2 - 1 = 0 \\ g(x,y) &= x^2 - y^2 + \tfrac{1}{2} = 0 \end{aligned} \qquad x_0 = y_0 = 0.5$$

(b) $$e^x + xy - 1 = 0 \qquad x_0 = 0.1$$
$$\sin xy + x + y - 1 = 0 \qquad y_0 = 0.5$$

(c) $$x^2 + y^2 = 1, \quad xy = 0 \qquad x_0 = 0.5,\ y_0 = 0.1$$

(d) $$x^2 + xy^3 = 9, \quad 3x^2y - y^3 = 4 \qquad \{x_0,y_0\} = \{1.2,2.5\}, \{-2,2.5\}, \{-1.2,-2.5\}, \{2,-2.5\}$$

(e) $$x - \sinh y = 0, \quad 2y - \cosh x = 0 \qquad x_0 = 0.6,\ y_0 = 0.6$$

۲۳ــ دستگاه غیرخطی $\begin{matrix} f(x,y) = 0 \\ g(x,y) = 0 \end{matrix}$ را به صورت

$$x_{i+1} = x_i - \frac{f(x_i,y_i)}{f_x(x_i,y_i)}$$

$$y_{i+1} = y_i - \frac{g(x_{i+1},y_i)}{g_y(x_{i+1},y_i)}$$

درآورید . سپس دستگاههای (e ۲۹ ــ a ۲۹) را با روش فوق حل کرده و نتایج دو روش را با هم مقایسه و توجیه نمائید .

۲۴ــ با استفاده از روش نیوتن تقریبی از ریشه‌های دستگاه زیر را با مقادیر اولیه داده‌شده بدست آورید (تعداد تکرارها : ۳) .

$$2x^2 + y^2 - 4z = 0$$
$$x^2 + y^2 + z^2 = 1 \qquad x_0 = y_0 = z_0 = 0.5$$
$$3x^2 - 4y + z^2 = 0$$

۲۵ــ با استفاده از روش تکرار ساده دستگاههای زیر را حل کنید (به کمک برنامهٔ کامپیوتری با انتخاب دقت مناسب) .

$$1. \quad \left.\begin{aligned} x &= \log \frac{y}{z} + 1, \\ y &= 0.4 + z^2 - 2x^2, \\ z &= 2 + \frac{xy}{20}, \end{aligned}\right\} x_0 = 1, \quad y_0 = 2.2, \quad z_0 = 2.$$

✓ $$2. \quad \left.\begin{aligned} x + x^2 - 2yz &= 0.1, \\ y - y^2 + 3xz &= -0.2, \\ z + z^2 + 2xy &= 0.3, \end{aligned}\right\} x_0 = y_0 = z_0 = 0.$$

۲۶ــ مطلوبست حل دستگاههای زیر با دقت 10^{-5}.

a) $$\begin{aligned} x&=\sin(x+y)\\ y&=\cos(x-y)\end{aligned}$$

b) $$\begin{aligned} \sin x.\,\mathrm{sh}\,y &= 0.2\\ \cos x.\,\mathrm{ch}\,y &= 1.2\end{aligned}$$

۲۷ ـ دستگاه زیر را با استفاده از روش تکرار ساده حل کنید (با استفاده از برنامهٔ کامپیوتری).

$$\begin{aligned} &3x_1 - \cos(x_2 x_3) - \tfrac{1}{2} = 0\\ &x_1^2 - 81(x_2 + .1)^2 + \sin x_3 + 1.06 = 0\\ &e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0\end{aligned}$$

$x_0 = y_0 = 0.1$
$z_0 = -0.1$

۲۸ ـ دستگاههای زیر را با روش تکرار ساده و نیوتن با دقت 10^{-4} حل کرده و نتایج را با هم مقایسه کنید. (با استفاده از کامپیوتر).

a) $$\begin{aligned} x_1^2 - 10x_1 + x_2^2 + 8 &= 0\\ x_1x_2^2 + x_1 - 10x_2 + 8 &= 0\end{aligned}$$

b) $$\begin{aligned} x_1^2 + x_2^2 - x_1 &= 0\\ x_1^2 - x_2^2 - x_2 &= 0\end{aligned}$$

c) $$\begin{aligned} 3x_1^2 - x_2^2 &= 0\\ 3x_1x_2^2 - x_1^3 - 1 &= 0\end{aligned}$$

d) $$\begin{aligned} x_1^2 + x_2 - 37 &= 0\\ x_1 - x_2^2 - 5 &= 0\\ x_1 + x_2 + x_3 - 3 &= 0\end{aligned}$$

e) $$\begin{aligned} &x_1 + \cos(x_1x_2x_3) - 1 = 0\\ &(1 - x_1)^{1/4} + x_2 + .05x_3^2 - .15x_3 - 1 = 0\\ &-x_1^2 - .1x_2^2 + .01x_2 + x_3 - 1 = 0\end{aligned}$$

f) $$\begin{aligned} 12x_1 - 3x_2^2 - 4x_3 &= 7.17\\ x_1^2 + 10x_2 - x_3 &= 11.54\\ x_2^3 + 7x_3 &= 7.631\end{aligned}$$

فصل سوم
حل معادلات چند جمله ای

۳.۱ ■ مقدمه

در این فصل یافتن ریشه‌های چند جمله‌ای

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n, \qquad a_n \neq 0$$
$$\equiv a_n(z - z_1)(z - z_2)\cdots(z - z_n) \tag{3.1}$$

موردنظر می‌باشد، که در آن a_r ها ضرائب حقیقی بوده و $z_r\ (r = 1, 2, \ldots, n)$ در حالت کلی ریشه‌های مختلط هستند. عدد α درصورتی ریشهٔ معادله گفته می‌شود که

$$f(\alpha) = 0$$

در اینجا تنها روشهای ساده‌تر یافتن ریشه‌های حقیقی موردنظر خواهد بود، اما برای حل کامل، باید تمام ریشه‌ها تعیین شوند. کتب جبرنشان می‌دهد که n ریشه‌برای یک چند جمله‌ای درجهٔ n وجود دارد. بایستی خاطرنشان ساخت ریشه‌های مختلط به صورت جفتهای مزدوج مختلط ظاهر می‌گردند، بنابراین برای هر ریشهٔ مختلط $a + ib$ مزدوج متناظرش $a - ib$ یک ریشه خواهد بود.

در نخستین نگاه، دلیلی به‌نظر نمی‌رسد که تصور شود مشکل خاصی دریافتن ریشه‌های یک چند جمله‌ای وجود داشته باشد. بطور تحلیلی یک چند جمله‌ای، تابع سادهٔ خاصی است، زیرا تمام مشتقاتش در هر ناحیه کران‌دار پیوسته بوده و می‌تواند به سهولت محاسبه گردد. بهر حال حتی برای چند جمله‌ایهایی از درجهٔ کمتر از ۲۰، مسئله محاسبه تقریبهای دقیق برای تمام صفرهای چند جمله‌ای می‌تواند فوق‌العاده مشکل باشد، خصوصا" زمانیکه ریشه‌ها مکرر باشند مشکلات حاد می‌گردد.

مثالی وسیلهٔ Wilkinson (1963) داده شده که نشان می‌دهد چطور یک تغییر خیلی کوچک در ضرائب یک چند جمله‌ای می‌تواند باعث تغییر بزرگی در تعدادی از صفرها گردد. چند جمله‌ای زیر، با صفرهای $z_k = k$ ،مستقیما" در نگاه اول این موضوع

را نشان می‌دهد .

$$f(z)=(z-1)(z-2)\cdots(z-20)$$
$$=z^{20}-210z^{19}+20615z^{18}+\cdots+20! \quad (3.2)$$

اگرچه یک تغییر به اندازه 2^{-23} در ضریب -210 تغییر کوچکی در دهمین رقم اعشار ایجاد می‌کند ، ریشه‌های $z_{10},\ldots,z_{19}$ از حقیقی به ریشه‌های مختلط بدل خواهند شد . تغییرات قابل ملاحظه نیز در مقادیر ایجاد می‌گردد ، برای مثال :

	مقدار قدیم	مقدار جدید
z_{14}	14	$13{\cdot}992+2{\cdot}519i$
z_{15}	15	$13{\cdot}992-2{\cdot}519i$

(3.3)

یافتن کلیه ریشه‌های یک چند جمله‌ای مشکل بزرگی می‌تواند باشد و توصیه می‌گردد که در حالات مشکل ، برای می‌نیمم کردن خطاهای هر برنامه کامل بایستی محاسبات با طول چندگانه به‌کار رود .

۳.۲ حساب چند جمله ای

یکی از فوائد محاسبه با چند جمله‌ای‌ها عبارتست از اینکه الگوریتمهای ساده‌ای برای محاسباتیکه به چند جمله‌ای نیازمند است ، وجود دارد . ذیلا" مسائل مربوط به یافتن مقدار چند جمله‌ای ، مشتق آن و باقیمانده آن در تقسیم بر $z-\alpha$ را در هرنقطه داده شده α بررسی می‌کنیم .

۳.۲.۱ ضرب تودرتو

در بحث حساب کامپیوتری اشاره کردیم که جمع و ضرب معمولا اعمال پایه کامپیوتر بوده ، اما سایر توابع به کمک قسمتهائی از برنامه کامپیوتر که کندتر می‌باشد انجام می‌گیرد ، همچنین این مطلب برای اعمالی نظیر z^r نیز صحیح بوده و در صورتیکه این توان و یا هر توانی به کار نرود مطلوب خواهد بود . حساب یک چند جمله‌ای بطور مستقیم نیازمند به محاسبه $n-1$ توان رسانی ، n ضرب و n جمع می‌باشد . روش ضرب تودرتو که در اینجا نشان داده شده تنها به n ضرب و n جمع نیازمنداست . روش برای یک چند جمله‌ای درجه سوم به شکل جبری

$$((a_3z+a_2)z+a_1)z+a_0 \quad (3.4)$$

به کار برده و مشاهده می‌شود که اثر این عبارتست از ضرب a_3 در z^3 و غیره ، و به صورت الگوریتم به شکل زیر نوشته می‌شود .

$$b_n=a_n$$
$$b_r=zb_{r+1}+a_r, \qquad r=n-1,n-2,\ldots,1,0 \quad (3.5)$$

که بوسیلهٔ برنامه کامپیوتری ساده‌ای محاسبه می‌گردد. b_0 مقدار چند جمله‌ای به‌ازای مقدار داده شدهٔ z می‌باشد. بطور مثال مقدار چند جمله‌ای درجه سوم $2z^3-z^2+6$ را در نقطهٔ $z=1{\cdot}1$ درنظر می‌گیریم

$$\begin{aligned} b_3 &= 2{\cdot}0 \\ b_2 &= 1{\cdot}1\times 2{\cdot}0-1{\cdot}0=1{\cdot}2 \\ b_1 &= 1{\cdot}1\times 1{\cdot}2+0=1{\cdot}32 \\ b_0 &= 1{\cdot}1\times 1{\cdot}32+6{\cdot}0=7{\cdot}452 \end{aligned} \tag{3.6}$$

این روش همیشه برای کامپیوتر یا ماشین حساب رومیزی قابل اجرا است.

۳.۲.۲ ■ تقسیم ترکیبی

روند تقسیم چند جمله‌ای به عامل $z-\alpha$ به دو دلیل مهم می‌باشد. اولا اینکه مولفه‌ای از طرح یافتن تمام ریشه‌های یک چند جمله‌ای را تشکیل داده و همچنین، قادر است‌قضیهٔ‌باقیمانده جبر رابه طور موثر در محاسبات‌کامپیوتر به کار گیرد. قضیهٔ‌باقیمانده با نوشتن یک چند جمله‌ای بصورت زیر بدست می‌آید.

$$f_n(z)=(z-\alpha)f_{n-1}(z)+R \tag{3.7}$$

اندیس دلالت بر درجهٔ چند جمله‌ای داشته و مشاهده می‌شود که تقسیم، خارج قسمتی از درجه $n-1$ و باقیماندهٔ ثابت R را می‌دهد. اگر $z=\alpha$ فرض شود، ملاحظه می‌کنیم که باقیمانده $R=f_n(\alpha)$ مقدار تابع در $z=\alpha$ را می‌دهد. ابتدا یک تقسیم طولانی بـــه $z-\alpha$ و آنگاه جدول خلاصه شده که معروف به روند تقسیم ترکیبی است را نشان می‌دهیم.

$$\begin{array}{l} z-\alpha\overline{)a_3z^3+a_2z^2+a_1z+a_0}(a_3z^2+(a_2+a_3\alpha)z+a_1+(a_2+a_3\alpha)\alpha \\ \quad\underline{a_3z^3-a_3\alpha z^2} \\ \qquad (a_2+a_3\alpha)z^2+a_1z \\ \qquad \underline{(a_2+a_3\alpha)z^2-(a_2+a_3\alpha)\alpha z} \\ \qquad\qquad [a_1+(a_2+a_3\alpha)\alpha]z+a_0 \\ \qquad\qquad \underline{[a_1+(a_2+a_3\alpha)\alpha]z-[a_1+(a_2+a_3\alpha)\alpha]\alpha} \\ \qquad\qquad\qquad a_0+[a_1+(a_2+a_3\alpha)\alpha]\alpha \end{array} \tag{3.8}$$

می‌توان دید نوشتن توانهای z غیر ضروریست، مشروط برآنکه ضرائب در ستونهای صحیح نگهداشته شوند و ستونهای سمت چپ صفرها را ایجاد می‌نماید که در تفاضل می‌توان نادیده گرفت. اگر تفاضل ایجاد شده به وسیله تغییر علامت α و افزودن به دو مقدار حاصل گردد نتایج ساده‌تر خواهند شد. آنگاه جدول زیر ایجاد می‌گردد.

$$\begin{array}{c|ccccc} & & a_3 & a_2 & a_1 & a_0 \\ +\alpha & + & 0 & p_3\alpha & p_2\alpha & p_1\alpha \\ \hline & & p_3=a_3 & p_2=a_2+p_3\alpha & p_1=a_1+p_2\alpha & p_0=a_0+p_1\alpha \end{array} \quad (3.9)$$

پائین‌ترین عنصر آخرین ستون روند، باقیمانده را می‌دهد که مقدار چند جمله‌ای در $z=a$ خواهد بود. ضرائب چند جمله‌ای خارج قسمت‌به کمک p_1 ، p_2 و p_3 بدست می‌آیند.

روند قبلی دارای همان گام‌های محاسبه‌ای است که ضرب تودرتو دارا است اما این طرح می‌تواند به تقسیم بر عامل درجه دوم توسعه داده شود، روند زمانی به کار می‌رود که زوج ریشهٔ مزدوج ـ مختلط از ریشه‌ها در معادلهٔ چند جمله‌ای رخ دهد. تقسیم ترکیبی به کمک عامل درجه دوم $z^2+\alpha z+\beta$ دارای شکل

$$\begin{array}{c|ccccc} & a_4 & a_3 & a_2 & a_1 & a_0 \\ -\alpha & & -p_4\alpha & -p_3\alpha & -p_2\alpha & \\ -\beta & & & -p_4\beta & -p_3\beta & -p_2\beta \\ \hline & p_4=a_4 & p_3=a_3-p_4\alpha & p_2=a_2-p_3\alpha-p_4\beta & p_1=a_1-p_2\alpha-p_3\beta & p_0=a_0-p_2\beta \end{array} \quad (3.10)$$

است. در این‌حالت، جملهٔ باقیمانده عبارتست از p_1z+p_0 و ضرائب چند جمله‌ای خارج قسمت p_4 ، p_3 و p_2 می‌باشد. خواننده بایستی با روند تقسیم ـ طولانی این روند را کنترل نماید. یک مثال عددی در مثال ۳.۱ داده شده است.

۳.۲.۳ ■ محاسبه مشتق

فرمول‌بندی معادلهٔ (3.7) را برای نشان دادن تقسیم به $z-\alpha$ درنظر می‌گیریم، یعنی:

$$f_n(z)=(z-\alpha)f_{n-1}(z)+R \quad (3.11)$$

و از طرفین مشتق گرفته باتوجه به اینکه R مقداری است ثابت، داریم:

$$f'_n(z)=f_{n-1}(z)+(z-\alpha)f'_{n-1}(z) \quad (3.12)$$

با جایگذاری مقدار $z=\alpha$ جمله دوم طرف راست صفر خواهد شد و مشتق وسیلهٔ مقدار $f_{n-1}(\alpha)$ بدست می‌آید. این چند جمله‌ای به کمک روند تتسیم ـ ترکیبی تشکیل شده و در نقطه $z=\alpha$ وسیله تقسیم ـ ترکیبی بیشتری یا با ضرب تودرتو می‌تواند محاسبه گردد، چون روند می‌تواند برای محاسبهٔ $f(x)$ و $f'(x)$ به‌کاررود. روندهای توضیح‌داده شده در این فصل کاربرد روش نیوتن برای چند جمله‌ای را به‌طور ویژه‌ای خیلی ساده می‌کند. مثال عددی از این روند در مثال (۳.۱) آورده شده است.

۳.۳ ■ یافتن تقریبهای آغازی

در صورت امکان تعدادی عملیات مقدماتی در یک معادله، قبل از قرار دادنِ در برنامه کامپیوتری، انجام می‌گیرد آنگاه روشی که شامل مهارت و بصیرت حل کننده است می‌تواند به کار رود.

روش بدیهی عبارتست از رسم نمودار چند جمله‌ای و در صورت امکان، تخمین وضعیت ریشه‌ها (این مطلب در فصل ۲ بحث شده است)که کمک زیادی در تخمین مقدار تقریبی ریشه‌ها می‌کند و بعضی از اینها در اینجا معرفی شده‌اند.

۳.۳.۱ ■ خواص مفید یک چندجمله ای

چند جمله‌ای به یکی از دو صورت زیر معرفی می‌شود.

$$f(z)=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_0 \tag{3.13}$$

یا

$$f(z)=a_n(z-z_1)(z-z_2)\cdots(z-z_n) \tag{3.14}$$

که در آن $z_r\ (r=1, 2, \ldots, n)$ ریشه‌های معادله می‌باشند. اگر $a_r\ (r=0, 1, 2, \ldots, n)$ حقیقی باشند دراینصورت تمام ریشه‌ها حقیقی یا بصورت زوجهای مزدوج ــ مختلط خواهند بود. بنابراین، اگر یک ریشه $z=a+ib$ وجود داشته باشد همچنین ریشه دیگری به صورت مزدوج $z=a-ib$ وجود خواهد داشت.

از مقایسه معادلات (3.13) و(3.14) روابطی بین ضرائب a_r و حاصلضربهای ریشه‌ها بدست می‌آید. معادلاتی که بیشتر مورد توجه می‌باشند عبارتند از:

$$\sum_{r=1}^{n} z_r=\frac{-a_{n-1}}{a_n} \tag{3.15}$$

$$\sum_{r=1}^{n}\sum_{s=r+1}^{n} z_rz_s=\frac{a_{n-2}}{a_n} \tag{3.16}$$

و

$$\prod_{r=1}^{n} z_r=(-1)^n\frac{a_0}{a_n} \tag{3.17}$$

درصورتیکه تمام ریشه‌ها حقیقی باشند، معادلهٔ اول می‌تواند کران بالای ریشه‌ها را بدهد. توجه داریم که ماکزیمم ریشه $x_{\max}$ در رابطهٔ زیر صدق می‌کند.

$$\begin{aligned}x_{\max}^2 &\leqslant x_1^2+x_2^2+\cdots+x_n^2\\ &=(x_1+x_2+\cdots+x_n)^2-2(x_1x_2+x_1x_3+\cdots+x_1x_n+x_2x_3+\cdots)\\ &=\left(\frac{a_{n-1}}{a_n}\right)^2-2\frac{a_{n-2}}{a_n}\end{aligned}$$

بدین معنی که

$$|x_{max}| \leqslant \sqrt{\left[\left(\frac{a_{n-1}}{a_n}\right)^2 - 2\frac{a_{n-2}}{a_n}\right]} \tag{3.18}$$

در حالتیکه ریشه تکراری است کنترل مفید است چون ممکن است استفاده از روش ویژه‌ای را پیشنهاد کند. این امر بسهولت باتوجه به اینکه ریشهٔ مورد نظر، ریشه مکرر مشتقات متوالی چند جمله‌ای است می‌تواند انجام گیرد. فرض می‌کنیم α یک ریشه از مرتبهٔ k باشد، بدین معنی که

$$f(x) = (x-\alpha)^k g(x) \tag{3.19}$$

آنگاه پس‌ازمشتق‌گیری‌از $f(x)$ ملاحظه‌می‌کنیم‌که α نیزیک‌ریشهٔ $f'(x)$ خواهدبود.

$$f'(x) = k(x-\alpha)^{k-1}g(x) + (x-\alpha)^k g'(x) \tag{3.20}$$

بوسیله مشتق‌گیری از مراتب بالاتر ملاحظه می‌شود که تمام مشتقات تا $f^{k-1}(x)$ نیز دارای ریشه $x=\alpha$ بوده و این عمل قادر است که مرتبه تکرار یک ریشه را پیدا کند.

قانون بسیار ساده‌ای موسوم به *قانون علائم دکارت* موجود است، که بعضی مواقع اطلاعات مفیدی برای تعداد ریشه‌ها می‌دهد. روش دارای اشکالی است که ممکنست بعضی مواقع هیچ اطلاعی ندهد. اگر تمام ضرائب معادله (3.13) حقیقی باشند آنگاه تعداد ریشه‌های مثبت $f(z)$ مساوی تعداد تغییرات علامت در دنباله $a_0, a_1, \ldots, a_n$ یا کمتر از این با تعداد زوج خواهد بود.

تعدادریشه‌های منفی، با روش مشابه، به چند جمله‌ای مرتبط می‌شودکه از قرار دادن $-x$ در چند جمله‌ای $f(x)$ محاسبه می‌گردد. تعداد ریشه‌های منفی $f(x)$ مساوی تعداد تغییرات علامت در چند جمله‌ای $f(-x)$ یا کمتر از این با تعداد زوج می‌باشد. بدین‌ترتیب، می‌توان دید که اگر تغییرات علامت در $f(x)$ وجود نداشته باشد در اینصورت ریشه مثبت وجود ندارد. به هر حال زمانیکه ضرائب چندین‌بار تغییر علامت دهند، اطلاع بسیار کمی حاصل می‌گردد.

۳.۳.۲ ■ رشته های ستورم

جزئیات ریاضی تئوری رشتهٔ ستورم در کتاب Ralston (1965) آورده شده است. دراینجا ارائه جزئیات محاسبه‌ای کاربرد رشته ستورم به منظور یافتن ریشه‌های چندجمله‌ای کفایت خواهد شد.

در حالت ساده یک رشته از چند جمله‌ایها به‌صورت زیر ایجاد می‌گردد:

$f_0(x)$ چند جمله‌ای اصلی و $f_1(x)$ مشتق این چند جمله‌ای می‌باشد. چند جمله‌ایهای بعدی عبارتند از باقیمانده حاصل تقسیم $f_r(x)$ به $f_{r+1}(x)$ با علامت منفی.

$$
\begin{aligned}
&f_0(x)=f(x)\\
&f_1(x)=f'(x)\\
&f_0(x)=f_1(x)q_1(x)-f_2(x)\\
&f_1(x)=f_2(x)q_2(x)-f_3(x)\\
&\cdots\cdots\cdots\cdots
\end{aligned}
\tag{3.21}
$$

در صورتیکه ریشه‌های تکراری وجود نداشته باشد دنباله با یک چند جمله‌ای $f_k(x)\ (k=n)$ که‌مقداریست ثابت متوقف می‌گردد. زمانیکه ریشه‌ها تکراری باشند روند با یک چند جمله‌ای $f_k(x)\ (k<n)$ متوقف شده و $n-k+1$ معرف تعداد ریشه‌های تکراری خواهد بود. فرض می‌کنیم، که ریشه‌های تکراری وجود نداشته باشد و دو مقدار a و b $(a<b)$ انتخاب شده و مقدار توابع رشتهٔ ستورم در این دو نقطه محاسبه گردد.

$N(a)$ را تعداد تغییرات علامت در رشتهٔ $f_0(a),f_1(a),\ldots,f_k(a)$ و $N(b)$ را تعداد تغییرات علامت در رشته $f_0(b),f_1(b),\ldots,f_k(b)$ می‌گیریم. تعداد ریشه‌های بین a و b با $N(a)-N(b)$ داده می‌شود. فرض شده است ریشه‌ای در a یا b وجود ندارد. اگر نقطه‌ای که انتخاب شده است ریشه‌ای از تابع باشد، این مقدار می‌تواند بلافاصله ثبت شده و رشتهٔ ستورم در نقطهٔ دیگری محاسبه گردد. درصورتیکه باقیمانده در نقطه‌ای از دنباله توابع وجود نداشته باشد پیچیدگی بیشتری می‌تواند ظاهر گردد. در چنین حالتی یک ریشه تکراری وجود داشته و این به وسیله آخرین عنصر رشته $f_k(x)\ (k<n)$ داده می‌شود و تئوری رشته ستورم هنوز می‌تواند قابل اجرا باشد. می‌توان رشته جدیدی با تقسیم به عامل مشترک $f_k(x)$ تشکیل داد.

$$
\tilde{f}_r(x)=\frac{f_r(x)}{f_k(x)},\qquad r=0,1,\ldots,k \tag{3.22}
$$

آنگاه نمونه تغییرات علامتی، قبل از پیداکردن موقعیت سایر ریشه‌ها می‌تواند به‌کاررود. مثالهای عددی رشته‌های ستورم در پایان فصل در مثالهای (۳۰۳) و (۳۰۴) داده شده‌اند.

۳.۴ ■ کلیه ریشه های یک چند جمله ای

۳.۴.۱ ■ شیوه تقلیلی

روش حل ساده بوده اما برای رسیدن به دقت محاسبه ممکن است هر مشکلی پیش آید که ناچیز خواهد بود. می‌دانیم که چند جمله‌ای دارای n ریشه است، تعدادی حقیقی و سایرین اعداد مزدوج ــ مختلط می‌باشند. بنابراین، ریشه‌های مختلط می‌توانند

با یک عامل درجه دوم با ضرائب حقیقی معرفی گردد. یعنی:

$$(z-\alpha-i\beta)(z-\alpha+i\beta)=(z^2-2\alpha z+\alpha^2+\beta^2) \tag{2.23}$$

نخست، انجام مطالعات مقدماتی به منظور پیدا کردن تقریب برای ریشه‌ها به کمک روشهائی که قبلا" در فصل ۲ بحث شده ضروری است. روش تکراری مناسبی برای پیدا کردن مقدار دقیقی برای ریشه به کار می‌رود. مشکلی بلافاصله ظاهر می‌شود، زیرا اتفاق می‌افتد که تکرار، به ریشهٔ مطلوب همگرا نمی‌گردد. همچنین، در بعضی حالات ریشه‌ای به روند تسلط داشته بنابراین روند تکراری در یک ناحیه وسیعی از مقادیر آغازی به این ریشه همگرا می‌گردد. در اینحالت، حذف این ریشه قبل از اینکه روش تکرار بتواند برای یافتن ریشه‌های بعدی به کار رود امریست ضروری. این عمل با روند تقلیلی انجام می‌گیرد که تقسیم ترکیبی را برای تقسیم به عامل دقیقی که پیدا شده به کار برده، و یک چند جمله‌ای جدید با یک ریشه کمتر به عنوان خارج قسمت می‌دهد. اگر ریشه‌ها مختلط باشند معمولا" به کار بردن روش تکراری برستو برای یک عامل‌درجه دوم ساده‌تر بوده چون از به کار بردن اعداد مختلط جلوگیری می‌کند. زمانیکه یک عامل درجه دوم بطور دقیق پیدا شود، روند تقسیم ترکیبی دوباره می‌تواند به کار رود و این عمل درجهٔ چند جمله‌ای را به‌اندازهٔ دو درجه تقلیل خواهد داد. این روند با یافتن یک ریشه به کمک تقلیل ادامه پیدا کرده تا آنکه تمام ریشه‌ها بدست آیند. مثالهای ۳۰۱ و ۳۰۲ روند تقلیلی، به‌ترتیب، برای یک عامل درجه اول و درجه دوم می‌باشد.

این رشته تقلیل و تکرارها متاسفانه می‌تواند خطاهای بزرگی در ریشه‌هائیکه بعدا" در روند پیدا می‌شوند ایجاد نمایند. این خطاها به دو طریق می‌توانند می‌نیمم گردند. اولا" ریشه‌ها باید، در صورت امکان، به ترتیب افزایش اندازه آنها بدست آیند. بحث این قسمت همراه با رفتار مفید یافتن ریشه‌های چند جمله‌ای در کتاب Bareiss (1967) داده شده است. ثانیا"، معکوس کردن چند جمله‌ای و آنگاه یافتن ریشه‌های جدید آن به ترتیب افزایش‌مرتبه بزرگی که متناظر با کاهش مرتبه بزرگی مجموعه ریشه‌های اولیه است. امکان‌پذیر می‌باشد.

$$f(z)=a_0+a_1z+a_2z^2+\cdots+a_nz^n \tag{2.24}$$

با تغییر متغیر $z=1/z$ تابع جدید تشکیل می‌شود.

$$\begin{aligned} f(z)&=a_0+\frac{a_1}{z}+\frac{a_2}{z^2}+\cdots+\frac{a_n}{z^n} \\ &=\frac{1}{z^n}[a_0z^n+a_1z^{n-1}+\cdots+a_nz+a_n] \end{aligned} \tag{2.25}$$

چند جمله‌ای داخل کروشه دارای ریشه‌هائی عکس ریشه‌های اصلی و باندازه معکوس مرتب شده است.

در اولین روند کوچکترین ریشه‌ها و در دومی معکوس بزرگترین ریشه‌ها خیلی دقیق خواهد بود .

جدول (۳۰۱) نتایجی از این دو نوع را نشان می‌دهد . درصورتیکه ریشه‌ها کاملا" دقیق باشد اعداد داخل جدول مقادیر $f(z)$ برای هر دو جمله‌ای معمولی S و چند جمله‌ای معکوس I رانشان‌داده ،البته ،مقادیربایستی ،صفر باشندنماد z_r $(r=1,2,\dots 7)$ برای مقدار خیلی دقیق هر ریشه در جدول می‌باشد .

جدول ۳۰۱ خطاها در حل چند جمله‌ای اصلی و معکوس

	z_1	z_2	z_3
S	$5{\cdot}12 \times 10^{-12}$	$1{\cdot}28 \times 10^{-11}$	$8{\cdot}31 \times 10^{-11}$
I	$2{\cdot}81 \times 10^{-8}$	$8{\cdot}93 \times 10^{-9}$	$6{\cdot}15 \times 10^{-9}$

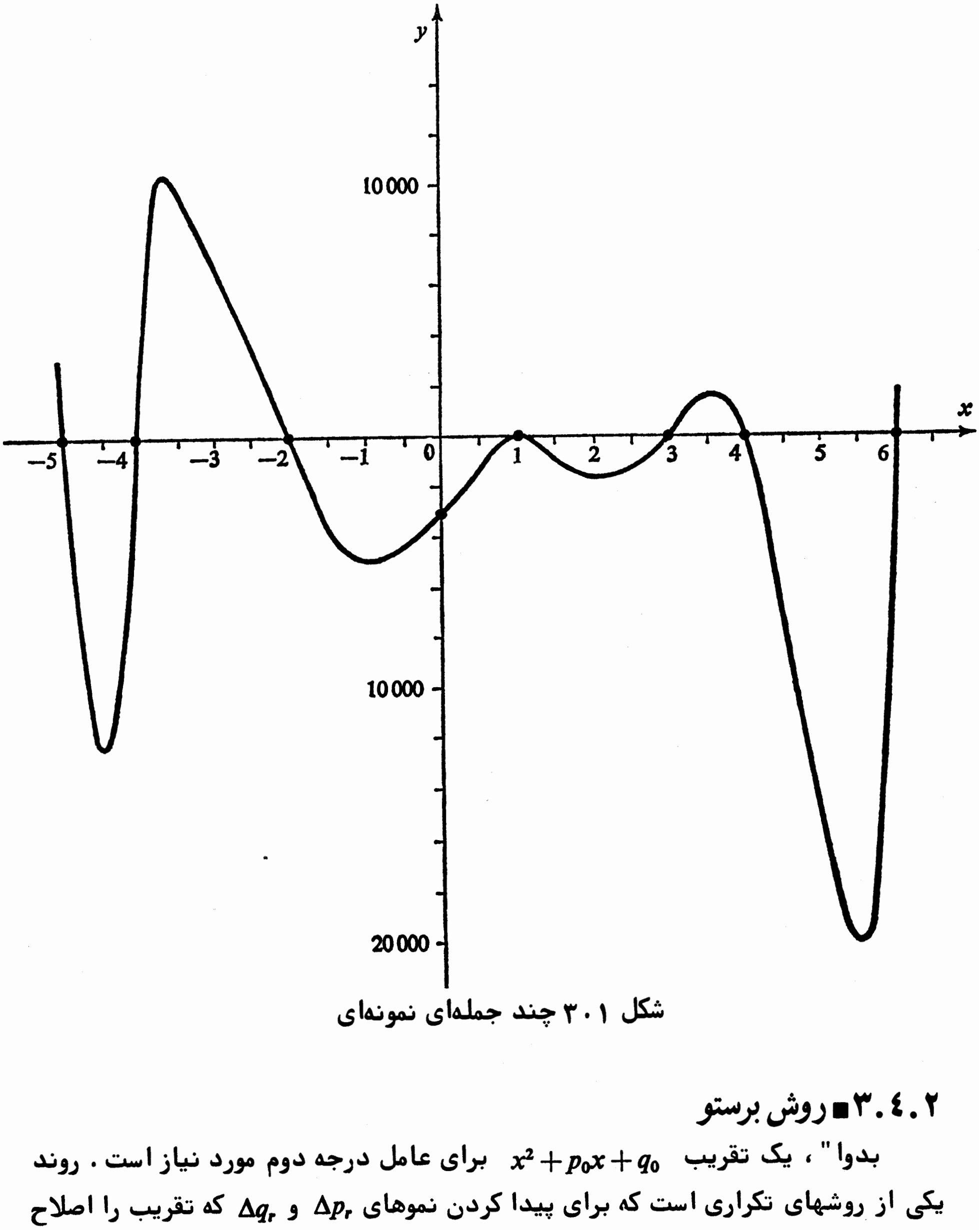
S	$9{\cdot}23 \times 10^{-10}$	$1{\cdot}83 \times 10^{-9}$	$4{\cdot}81 \times 10^{-9}$	$2{\cdot}13 \times 10^{-8}$
I	$4{\cdot}51 \times 10^{-10}$	$3{\cdot}56 \times 10^{-10}$	$9{\cdot}05 \times 10^{-11}$	$8{\cdot}54 \times 10^{-11}$
	z_4	z_5	z_6	z_7

حال انتخاب روشهای تکراری مناسب برای چند جمله‌ای را درنظر می‌گیریم . از آنجائیکه، مشتق یک چند جمله‌ای از درجه n دارای $n-1$ ریشه می‌باشد . می‌توان دید ، که یک چند جمله‌ای دارای تعداد قابل توجهی نقاط بازگشتی است که در شکل ۳۰۱ نشان داده شده است .

$$x^8-4x^7-46x^6+168x^5+553x^4-1988x^3-988x^2+5184x-2880$$
$$=(x+5)(x+4)(x+2)(x-1)^2(x-3)(x-4)(x-6)$$

روش نظیر روش نیوتن سریعا" در همسایگی ریشه همگرا می‌گردد اما در تابع نوسانی تند نظیر چند جمله‌ای با مرتبه بالا به مسائل همگرائی بستگی دارد. زمانیکه یک چند جمله‌ای را درنظر می‌گیریم روشی را به کار می‌بریم که اطمینان در همگرائی آن باشد ، حتی اگر این سبب کاهش سرعت در گام نهائی تکرار شود. روش نابجائی روشی است ساده و مناسب برای یافتن یک ریشه حقیقی و روش برستو اغلب برای یافتن یک زوج مزدوج-مختلط به‌کارمی‌رود . توجیه‌تئوری روش‌برستودر McTernan (1967) و Balfour ، بـه آسانی قابل درک برای آندسته که آشنائی با مشتقات جزئی دارند ، آورده شده است . تنها گامهای محاسبه‌ای روند در اینجا داده شده است .

نمودار فرمول بالا



شکل ۳۰۱ چند جمله‌ای نمونه‌ای

۳.۴.۲ ■ روش برستو

بدوا" ، یک تقریب $x^2+p_0x+q_0$ برای عامل درجه دوم مورد نیاز است . روند یکی از روشهای تکراری است که برای پیدا کردن نموهای Δp_r و Δq_r که تقریب را اصلاح می‌کند به کار می‌رود . رشته مقادیر

$$p_{r+1}=p_r+\Delta p_r, \qquad q_{r+1}=q_r+\Delta q_r, \qquad r=0,1,2,\ldots, \tag{3.26}$$

به کمک حل دستگاه معادلات دو مجهولی زیر بدست می‌آیند .

$$\begin{aligned} a_{11}\Delta p_r+a_{12}\Delta q_r&=b_1\\ a_{21}\Delta p_r+a_{22}\Delta q_r&=b_2 \end{aligned} \tag{3.27}$$

گام‌های روند عبارتنداز:

۱ــ چند جمله‌ای $f(x)$ به تقریب جاری، عامل درجه دوم $x^2+p_rx+q_r$، تقسیم شده و خارج قسمت $Q(x)$ و باقیماندهٔ R_1x+S_1 را می‌دهد. زمانیکه‌روند همگرا باشد هر دو R_1 و S_1 صفر خواهند شد.

۲ــ آنگاه‌تابع $xQ(x)$ به $x^2+p_rx+q_r$ تقسیم شده و باقیماندهٔ R_2x+S_2 رامی‌دهد.

۳ــ تابع $Q(x)$ به $x^2+p_rx+q_r$ تقسیم‌شده باقیمانده R_3x+S_3 رامی‌دهد. باید توجه داشت که این عمل شامل عملیات حسابی بیشتری نمی‌باشد چون اعداد مانند گام ۲ اما با ستون‌های مختلف می‌باشد.

۴ــ ضرائب در دستگاه معادلات دو مجهولی به وسیله

$$a_{11}=-R_2, a_{12}=-R_3, a_{21}=-S_2, a_{22}=-S_3, b_1=-R_1, b_2=-S_1.$$

داده شده‌اند.

۵ــ مقادیر جدید p_{r+1} و q_{r+1} بدست آمده و روند تا اینکه مقادیر p و q همگـــرا گردند تکرار می‌شود. سیمائی از این طرح نادیده گرفته شده، چون راه ساده‌ای برای آن نیست، مسئلهٔ نسبتا "مشکل یافتن تقریب آغازی عامل درجه دوم است. ممکن است عامل متناظر با سه جمله‌ای از کمترین درجه در معادله را به کار برد، اگرچه تضمینی برای همگرائی جوابها نخواهد بود. مثال ۳۰۵ جدول عددی می‌دهد که نشانگر آن است که چگونه روش برستو به کار می‌رود. روش دیگری که دارای جذابیتی قابل توجه و ایده‌آل می‌باشد روش لاگور است که با فرمول زیر تعریف می‌گردد.

$$\begin{aligned} x_{n+1} &= x_n-\frac{nf(x_n)}{f'(x_n)\pm[H(x_n)]^{\frac{1}{2}}} \\ H(x_n) &= (n-1)^2[f'(x_n)]^2-n(n-1)f(x_n)f''(x_n) \end{aligned} \tag{3.28}$$

این روش دو مجموعه از تکرارها را می‌دهد که به نزدیکترین صفر بزرگتر از a و نزدیکترین صفر کمتر از a همگرا می‌گردد که در آن a عبارتست از اولین تقریب، اگر علامت ریشه دوم موافق $f'(x_n)$ گرفته شود روش همگرا از مرتبه ۳ خواهد بود، اما این عمل مستلزم صرف محاسبه فوق‌العاده‌ای خواهد بود. جذابیت روش به علت آن است که تعدادی از صاحبنظران‌پیشنهاد می‌کنند که این روش خیلی شبیه روش نیوتن همگرا می‌گردد و این زمانیست که تقریب آغازی در همسایگی ریشه نباشد. روش برای یک چند جمله‌ای خیلی عملی‌تر است تا برای یک $f(x)$ کلی، چون مشتق دوم $f''(x_n)$ به وسیله سه بار استفاده از روند تقسیم ــ ترکیبی می‌تواند محاسبه شود، که یک الگوریتم کامپیوتری ساده می‌باشد.

۳.۵ روشهای لازم برای یافتن تمام ریشه ها

در قسمت گذشته اشاره شد که مشکلات خطاهای ایجاد شده زمانی رخ می‌دهند که روند تکرار حاصل از تقلیل، به طور تکراری به کار رود. بنابراین شایان توجه خواهد بود که تمام ریشه‌های یک چند جمله‌ای را بطور همزمان پیدا کرد. متاسفانه این نوع روشها دارای نقاط ضعفی است. زمانیکه نتایج در همسایگی مقادیر واقعی هستند، سرعت همگرائی معمولا" خیلی کندتر از سرعت همگرائی روش برای یک تک ریشه است. استراتژی خوب روشی است‌که تمام ریشه‌ها را پیدا کرده و تقریبهای آغازی خوبی داده، و با یک روش سریعا " همگرا، برای بهتر کردن دقت تک تک ریشه‌ها، تلفیق شود.

۳.۵.۱ الگوریتم تفاضل - خارج قسمت

این روش برای اولین بار بوسیله Rutishauser (1956) و بطور ساده در کتابهای Henrici (1964), ، McTernan (1967) و Balfour بیان شده است. پیش‌نیاز لازم دانشی از توابع مختلط بوده بنابراین، خارج از حوصلهٔ این کتاب می‌باشد. بهر حال، به منظور اهمیت روش و شکل سادهٔ محاسبات، فقط به ارائه گامهای کامپیوتری اکتفا می‌شود. برای چند جمله‌ای درجهٔ n ، یک جدول با $2n+1$ ستون رسم می‌گردد. دو ستون اولی و آخری جدول صفر هستند، باقیماندهٔ ستونها در دو مجموعه ستونهای متناوب قرار دارند که به کمک دو قاعده محاسبه‌ای جداگانه، قاعدهٔ خارج‌قسمت و تفاضل، بدست می‌آیند. روش تولید جدول به بهترین وجه با ساختن از ستون صفر منتهاالیه سمت‌چپ بوسیله طرحی مبتنی بر روش برنولی برای یافتن ریشه‌ها، بسادگی بیان می‌گردد.

(این قسمت در این‌کتاب بحث نشده است). بهرحال، به منظور می‌نیمم کردن خطاهای ایجاد شده، ترجیح داده می‌شود که جدول سطر به سطر تولید گردد، این نیازمند به روشی از ایجاد دو سطر اول که نسبتا" برای خواننده اختیاری است می‌باشد. درصورت لزوم در Henrici (1964) پیش نیاز قابل توجهی برای این قوانین موجـــود می‌باشد. جدول ۳.۲ برای یک معادلهٔ درجه سوم داده شده است.

جدول ۳.۲

جدول تفاضل ــ خارج قسمت

	$q_0^{(1)}$		$q_{-1}^{(2)}$		$q_{-2}^{(3)}$		
0		$\varepsilon_0^{(1)}$		$\varepsilon_{-1}^{(2)}$		0	
	$q_1^{(1)}$		$q_0^{(2)}$		$q_{-1}^{(3)}$		
0		$\varepsilon_1^{(1)}$		$\varepsilon_0^{(2)}$		0	
	$q_2^{(1)}$		$q_1^{(2)}$		$q_0^{(3)}$		
0		$\varepsilon_2^{(1)}$		$\varepsilon_1^{(2)}$		0	
	$q_3^{(1)}$		$q_2^{(2)}$		$q_1^{(3)}$		
	$\vdots$		$\vdots$		$\vdots$		

ضمناً " اندیس‌های هر ستون و قطر پیشرو ثابت می‌باشند . برای یک طرح کلی ازدرجه n روند با محاسبهٔ عناصر دو سطر اول مطابق ذیل شروع می‌شود .

$$\begin{aligned} &q_0^{(1)} = -a_{n-1}/a_n \\ &q_{1-r}^{(r)} = 0, \qquad r = 2, 3, \ldots, n \\ &\varepsilon_{1-r}^{(r)} = a_{r-1}/a_r, \qquad r = n-1, n-2, \ldots, 1 \end{aligned} \tag{3.29}$$

عناصر آنگاه به سمت راست روی سطرها و پائین روی ستون‌ها محاسبه می‌گردند و چهار مقدار به شکل رئوس لوزی مطابق ذیل بدست می‌آید .

$$\begin{array}{ccc} & \alpha_j & \\ \beta_{j+1} & & \beta_j \\ & \alpha_{j+1} & \end{array}$$

در صورتیکه α_j عناصر واقع در ستون q باشد آنگاه $\alpha_j + \beta_j = \beta_{j+1} + \alpha_{j+1}$ اگر α_j عناصر واقع در ستون ε باشد آنگاه $\alpha_j \times \beta_j = \beta_{j+1} \times \alpha_{j+1}$ را داریم . باحل سطر به سطر، تمام مقادیر α_j ، β_j و β_{j+1} معلوم شده و مجهول α_{j+1} وسیله یکی از معادلات زیر بدست می‌آید .

$$q_{j+1}^{(r)} = q_j^{(r)} + \varepsilon_j^{(r)} - \varepsilon_{j+1}^{(r-1)} \tag{3.30}$$

$$\varepsilon_{j+1}^{(r)} = \varepsilon_j^{(r)} \cdot \frac{q_j^{(r+1)}}{q_{j+1}^{(r)}} \tag{3.31}$$

بایستی توجه داشت در صورتیکه هرمقدار $q_j^{(r)}$ مساوی صفر گردد ممکن است طرح به‌طور کامل بهم بخورد ،چون در اینصورت تقسیم نمی‌تواند ادامه پیدا کند . الگوریتم فوق زمانیکه ریشه‌های چندجمله‌ای z_r $(r = 1, 2, \ldots, n)$ دررابطه $|z_1| > |z_2| > \cdots > |z_n|$ صدق کند خیلی مفید است . در این حالت می‌توان ثابت کرد که

$$\lim_{j \to \infty} q_j^{(r)} = z_r, \qquad r = 1, 2, \ldots, n \tag{3.32a}$$

$$\lim_{j \to \infty} \varepsilon_j^{(r)} = 0, \qquad r = 1, 2, \ldots, n \tag{3.32b}$$

بدین ترتیب ، اگر تمام ریشه‌ها از حیث قدر مطلق نامساوی باشند طرح به کمک مطالعه رفتار ستون‌های ε آشکار خواهد شد . اگر تعدادی از ریشه‌ها دارای قدرمطلقهای مساوی باشند یعنی $|z_1| \geqslant |z_2| \geqslant \cdots \geqslant |z_n|$ آنگاه شرایط زیر برقرار می‌گردد .

۱ ــ برای هر r طوریکه : $|z_{r-1}| > |z_r| > |z_{r+1}|$

داریم

$$\lim_{j \to \infty} q_j^{(r)} = z_r \tag{3.33a}$$

۲ـ برای هر r طوریکه : $|z_r| > |z_{r+1}|$
داریم

$$\lim_{j\to\infty} \varepsilon_j^{(r)} = 0 \qquad (3.33b)$$

بدین‌ترتیب ، جدول به گروههائی از ستون‌های q متناظر با ریشه‌هائی از حیث قدرمطلق مساوی و ستون‌های ε که به صفر میل می‌کنند تقسیم می‌گردد .

بیشترین حالت عمومی عبارتست از ریشه‌هائیکه از حیث قدرمطلق مساوی‌اند ، و این زمانی رخ می‌دهد که یک معادله با ضرائب حقیقی دارای یک زوج ریشه‌مزدوج‌ـمختلط باشد . دو ستون q ، جدا شده به کمک ستون ε که به سمت صفر میل می‌کند ، وجود خواهد داشت و دو ریشه مورد نیاز z_{r+1} و z_{r+2} جوابهای معادلهٔ درجه دوم

$$z^2 - A_r z + B_r = 0$$

خواهد بود که در آن

$$\begin{aligned} \lim_{j\to\infty} (q_{j+1}^{(r+1)} + q_j^{(r+2)}) &= A_r \\ \lim_{j\to\infty} (q_j^{(r+1)} q_j^{(r+2)}) &= B_r \end{aligned} \qquad (3.34)$$

برای مثالهای پیچیده‌تر می‌توان به کتابهای ذکر شده مراجعه کرد . نتایج عددی در مثال(۳.۶) داده شده است .

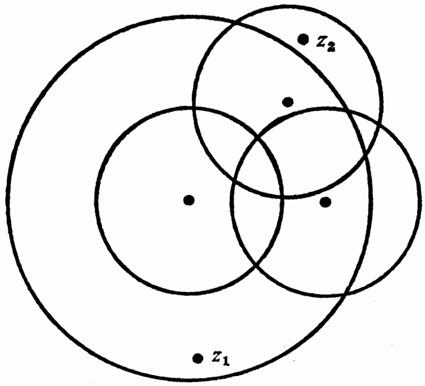
۳.۵.۲ ■ روش لمرـ شور

اگرچه روش همیشه همگرا است ، اما فوق‌العاده پر زحمت و خیلی کند همگرا می‌گردد . البته برای پیدا کردن تقریبهای آغازی مناسب می‌باشد . روش مبتنی بر رشته‌ای از محاسبات می‌باشد که بررسی می‌کند آیا یک ریشه از معادله در دایرهٔ داده شده قرار می‌گیرد یا نه . یک رشته دوایر آزمون که بتواندبرای تحقیق تمام صفحه مختلط به کار رود . اولا" ، آزمون برای بیان اینکه آیا یک ریشه در داخل دایره به مرکز مبداٴ مختصات و شعاع واحد قرار می‌گیرد یا نه به کار می‌رود . اگر این آزمون موفقیت‌آمیز نباشد یک انتقال از معادله اصلی قادر خواهد بود یک بررسی در یک دایره با شعاع دو برابر شعاع دایره اصلی انجام گیرد . در صورتیکه هنوز موفقیت‌آمیز نباشد ، انتقال بعدی بررسی در داخل دایره‌ای به شعاع چهار برابر شعاع دایره‌اصلی را می‌دهد . این روند تا اینکه‌یک ریشه در حلقهٔ $R < |z| < 2R$ قرار گیرد ادامه پیدا می‌کند . سطحی که در آن ریشه قرار می‌گیرد ۸ دایره متقاطع به شعاع $4R/5$ با مراکز

$$z = \frac{3R e^{2\pi i k/8}}{2\cos(\pi/8)}, \qquad k = 0, 1, \ldots, 7 \qquad (3.35)$$

میباشد .

یکی از این دوایر بایستی شامل یک ریشه باشد . همچنین تشخیص اینکه بیشتر از یک دایره ممکن است شامل یک ریشه باشد حائز اهمیت بوده ، چون دوایر متقاطع سطح وسیعتری از دوایر متحدالمرکز را می‌پوشاند و هم‌چنین بعضی سطوح دوباره پوشیده



شکل ۳۰۲ روش لمر ــ شور

می‌شوند . حال دایره‌ای که شامل یک ریشه می‌باشد وجود دارد ، بنابراین یک بررسی در دوایر متوالی که هر کدام دارای شعاع نصف اندازه شعاع قبلی است که حلقه ایجاد می‌کند حاوی یک ریشه بوده و آنگاه روند می‌تواند یک بار دیگر تکرار گردد .

تفصیلات ریاضی و الگوریتم کامل این روش در Ralston (1965) یا مقاله اصلی Lehmer (1961) وجود دارد ، اما توصیه می‌شود اگر برنامه معمولی مشخصی در کامپیوتر موجود است بهتر است که به کار گرفته شود . مشکلات اساسی روش زمانی است که چندین ریشه در یک سطح تحقیق قرار گیرد یا زمانیکه یک ریشه جدا از حلقهٔ بیان شده بررسی گردد . این مطلب به سادگی به کمک شکل (۳۰۲) دیده می‌شود که نشان می‌دهد روش یک ریشهٔ z_1 که تدریجا" به ریشهٔ دیگر z_2 همگرا می‌گردد را پیدا می‌کند .

۳.۵.۳ ■ روش مربع سازی ریشه گریف

این روش در ساده‌ترین شکلش به صورت معادله‌ایست که ریشه‌هایش از حیث قدر مطلق نامساوی بوده و از این معادله ، معادله دومی نتیجه می‌شود که ریشه‌هایش مربع ریشه‌های معادله اصلی می‌گردد . روند تا زمانیکه ریشه‌های جدید ، 2^m برابر ریشه‌های اصلی شود ادامه پیدا می‌کند . برای m به اندازه کافی بزرگ اندازه ریشه‌های اصلی با یک محاسبه ساده می‌تواند بیان گردد .

روش به علت روند مربع کردن که جداسازی ریشه‌ها را افزایش می‌دهد و منجر به سادگی روند دریافتن ریشه شده، جذاب می‌گردد. آنالیز خطا نشان می‌دهد که بدست آوردن مقادیر دقیق برای ریشه‌ها زمانی که خیلی بهم نزدیک هستندمشکل است. اگرچه در ساده‌ترین شکلش برای حالتی که ریشه‌ها از حیث اندازه مساوی هستند قادر نمی‌باشد اما تغییر و تبدیلی برای این روش وجود دارد که پس ازآن، روش قادر به اعمال خواهد بود. فرض می‌کنیم ضرائب چند جمله‌ایهای متعدد با $a_r^{(J)}$ که در آن $a_r^{(0)}=a_r$ مطابق مثال (۳.۱) تعریف شده است. ضرائب چند جمله‌ای جدید به صورت زیربدست‌می‌آیند.

$$\begin{aligned} a_n^{(1)} &= [a_n^{(0)}]^2 \\ a_{n-1}^{(1)} &= -[a_{n-1}^{(0)}]^2 + 2a_n^{(0)}a_{n-2}^{(0)} \\ a_{n-2}^{(1)} &= [a_{n-2}^{(0)}]^2 - 2a_{n-1}^{(0)}a_{n-3}^{(0)} + 2a_n^{(0)}a_{n-4}^{(0)} \end{aligned} \tag{3.36}$$

و با یک قانون مشابه با اندیسهای مختلف عناصر بیشتری از دنباله ایجاد خواهد شد. یک رابطهٔ معمولی بین ضرائب چند جمله‌ای و مجموع حاصلضربهای ریشه‌های Z_r موجود است. برای مثال:

$$\begin{aligned} \frac{a_{n-1}}{a_n} &= -\sum_{r=1}^{n} Z_r \\ \frac{a_{n-2}}{a_n} &= \sum_{r=1}^{n}\sum_{s=r+1}^{n} Z_r Z_s \\ \frac{a_0}{a_n} &= (-1)^n Z_0 . Z_1 \cdots Z_n \end{aligned} \tag{3.37}$$

که در آن جمع روی تمام ترکیبات مقادیر Z_r و Z_s بین 1 و n انجام گرفته است. در حالت مورد بحث، بعلت عملیات توان روی ریشه‌ها، بیشتر جمله‌ها قابل اغماض می‌گردند. بنابراین، تقریبهای زیر برای چند جمله‌ایها با مرتبه، به اندازه کافی بالا در رشته گریف می‌توانند به کار رود. برای سهولت فرض می‌کنیم:

$$|Z_1| > |Z_2| > \cdots > |Z_n|$$

$$\begin{aligned} &Z_1^{2J} \approx -a_{n-1}/a_n \\ &Z_1^{2J}Z_2^{2J} \approx a_{n-2}/a_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &Z_1^{2J}Z_2^{2J} \cdots Z_n^{2J} \approx (-1)^n(a_0/a_n) \end{aligned} \tag{3.38}$$

و با فرض ریشه مناسب، مقادیر ریشه‌های متعدد Z_r می‌توانند پیدا گردند. علامت Z_r با جایگذاری در دنبالهاصلی پیدا می‌شود، نباید فریب بیان ساده روند گریف ذکر شده در فوق را خورد. مشکلات با ریشه‌های مساوی از حیث قدر مطلق ایجاد می‌گردد و همگرائی ضعیف در روند را قبل از اینکه یک برنامه کامپیوتری مناسب به کار رود بایستی درنظر گرفت. برای مطالعه قویا " فصل نوشته شده توسط Bareiss در Ralston و Wilf جلد

دو سال (1967) توصیه می‌شود که مقدار زیادی از بحث ارزشمند در مسئله حل چند جمله‌ای را داده و همچنین تفصیل نمودار گردشی برای روند مربع‌سازی ریشه، شیوه‌ای پیش رفته را می‌دهد.

برنامه‌نویسی این روش توصیه نمی‌گردد مگر درصورتیکه نیاز مبرم به حل مشکل کامپیوتری که باید برطرف گردد، باشد.

۳.۶ ■ مقایسه روشها

رضایت بخش‌ترین برنامه با هدف معمولی وسیله Bareiss (1967) بیان‌شده‌است. این روش همگرائی خیلی سریعتر نسبت به روش لمر ــ شور را نشان می‌دهد، هم‌چنین شامل عملیات خیلی پیچیده‌تری می‌باشد. ساده‌ترین گرایش عبارتست‌از نوشتن برنامه خصوصی و به کاربردن آلگوریتم تفاضل ــ خارج قسمت به منظور بدست آوردن تقریبهایی برای ریشه‌ها، و کاربرد روشهایی با خواص همگرائی خوب به منظور اصلاح ریشه نظیر روش نیوتن برای ریشه‌های متمایز و برستو برای ریشه‌های مختلط ــ مزدوج. اگر برنامهٔ کامپیوتری استانداردی برای روش لمر ــ شور موجود باشد به جای الگوریتم تفاضل ــ خارج قسمت به کار می‌رود. یافتن کلیه ریشه‌های یک چند جمله‌ای ساده نبوده و جائیکه‌یک برنامه کامپیوتری معتبر موجود باشد به‌کار خواهد رفت. بهر حال، باید تشخیص داد که هنوز چند جمله‌ایهایی موجودهستند که با روشهای عددی معمولی قابل حل نبوده و برای بدست آوردن جوابهای با معنی محتاج به تجزیه و تحلیلی دقیق ریاضی می‌باشد.

مثالهای حل شده وبرنامه های کامپیوتری

۱ ــ مطلوبست تعیین مقدار چند جمله‌ای $5x^4-2x^3+x-1$ و مشتق آن در نقطهٔ $x=2$ با به کار بردن قضیهٔ باقیمانده و روند تقسیم ــ ترکیبی.

باقیماندهٔ تقسیم چند جمله‌ای بر $x-2$ مقدار چند جمله‌ای در نقطهٔ $x=2$ را می‌دهد. باقیماندهٔ حاصل تقسیم خارج قسمت به $x-2$، مقدار مشتق را می‌دهد.

$$
\begin{array}{r|rrrr|l}
 & 5 & -2 & 0 & 1 & -1 \\
2 & 0 & 10 & 16 & 32 & 66 \\
\hline
 & 5 & 8 & 16 & 33 & 65=f(2) \\
2 & 0 & 10 & 36 & 104 & \\
\hline
 & 5 & 18 & 52 & 137=f'(2) &
\end{array}
$$

۲ــ دو ریشه از معادلهٔ $x^4+2x^3-6x^2-22x-15$ عبارتنداز زوج مزدوج ــ مختلط $-2+i$ و $-2-i$ که متناظر با ریشه‌های عامل درجه دوم x^2+4x+5 هستند.

تقسیم ترکیبی به عامل درجه دوم ، چند جمله‌ای را آماده برای‌یک حل تکراری بعدی می‌کند .

$$\begin{array}{r|rrrrr}
 & 1 & 2 & -6 & -22 & -15 \\
-4 & 0 & -4 & 8 & 12 & 0 \\
-5 & 0 & 0 & -5 & 10 & +15 \\
\hline
 & 1 & -2 & -3 & \multicolumn{1}{|r}{0} & 0
\end{array}$$

باقیمانده تائید می‌کند که x^2+4x+5 یک عامل از چند جمله‌ای است ، و عامل دیگر x^2-2x-3. است .

۳ ــ مطلوبست تعیین محل تقریبی یکی از ریشه‌های x^3-3x^2+2x-1 با به کار بردن روش رشتهٔ ستورم .

$$f_0(x)=x^3-3x^2+2x-1$$
$$f_1(x)=3x^2-6x+2$$
$$f_2(x)=2x/3+\tfrac{1}{3}$$
$$f_3(x)=-\tfrac{23}{4}$$

دو تابع آخری به کمک دو جدول تقسیم ترکیبی زیر داده می‌شوند .

$$\begin{array}{r|rrrr}
 & 1 & -3 & 2 & -1 \\
2 & 0 & 2 & -2 & 0 \\
-\frac{2}{3} & 0 & 0 & -\frac{2}{3} & +\frac{2}{3} \\
\hline
 & 1 & -1 & \multicolumn{1}{|r}{-\frac{2}{3}} & -\frac{1}{3}
\end{array}$$

$$\begin{array}{r|rrr}
 & 3 & -6 & 2 \\
-\frac{1}{2} & 0 & -\frac{3}{2} & \frac{15}{4} \\
\hline
 & 3 & -\frac{15}{2} & \multicolumn{1}{|r}{\frac{23}{4}}
\end{array}$$

علائم این توابع در نقاط مختلف ، در جدول زیر آورده شده است .

x	$-\infty$	0	∞	1	2	3
$f_0(x)$	$-$	$-$	$+$	$-$	$-$	$+$
$f_1(x)$	$+$	$+$	$+$	$-$	$+$	$+$
$f_2(x)$	$-$	$+$	$+$	$+$	$+$	$+$
$f_3(x)$	$-$	$-$	$-$	$-$	$-$	$-$
	2	2	1	2	2	1

بدین ترتیب ، ریشه بین $x=2$ و $x=3$ قرار می‌گیرد .

۴ ــ مطلوبست تعیین محل تقریبی ریشه‌های چند جمله‌ای $x^4-9x^2+4x+12$ با به کار بردن روش رشته ستورم.

جدول ستورم به صورت زیر بسط داده‌می‌شود :

$$f_0(x)=x^4-9x^2+4x+12$$
$$f_1(x)=4x^3-18x+4$$
$$f_2(x)=18x^2/4-3x-12$$
$$f_3(x)=50x/9-\tfrac{100}{9}$$

زمانیکه این عامل به $f_2(x)$ تقسیم گردد ، باقیمانده‌ای که دلالت به ریشه تکراری به

صورت عامل $f_3(x)$ یعنی $x-2$ باشد نخواهد داشت. بسط رشته ستورم به وسیله تقسیم به این عامل بدست می‌آید.

$$f_0^*(x)=x^3+2x^2-5x-6$$
$$f_1^*(x)=4x^2+8x-2$$
$$f_2^*(x)=18x/4+6$$
$$f_3^*(x)=\frac{50}{9}$$

جدول زیر موقعیت ریشه‌ها را نشان می‌دهد.

x	$-\infty$	-4	-2	0	∞
$f_0^*(x)$	$-$	$-$	$+$	$-$	$+$
$f_1^*(x)$	$+$	$+$	$-$	$-$	$+$
$f_2^*(x)$	$-$	$-$	$-$	$+$	$+$
$f_3^*(x)$	$+$	$+$	$+$	$+$	$+$
	3	3	2	1	0

یک ریشه بین -4 و -2، یک ریشه بین -2 و 0، و یک ریشه بین 0 و ∞ وجود دارد. ریشهٔ آخری یک ریشهٔ مضاعف در $x=2$ بوده که قبلاً" پیدا شده است. بدین‌ترتیب، موقعیت تقریبی تمام ریشه‌ها معلوم می‌گردد.

۵ ــ مطلوبست تعیین یک عامل‌درجه دوم از چند جمله‌ای زیر به روش برستو.

$$P(x)=2x^6+15x^5+43x^4+51x^3+41x^2+108x+160$$

سه ضریب آخری برای پیشنهاد اولین تقریب x^2+3x+4 به کار می‌رود. در روش برستو، تقسیم ترکیبی را برای تقسیم $P(x)$، $Q(x)$ و $xQ(x)$ به عامل آزمایشی به کار می‌بریم.

$$P(x)=Q(x)(x^2+3x+4)+R_1x+S_1$$
$$Q(x)=B(x)(x^2+3x+4)+R_3x+S_3$$
$$xQ(x)=C(x)(x^2+3x+4)+R_2x+S_2$$

نموها برای ضرائب از معادلات

$$-R_2\Delta p_0-R_3\Delta q_0=-R_1$$
$$-S_2\Delta p_0-S_3\Delta q_0=-S_1$$

پیدا شده و تقریبهای بعدی عبارتنداز $3+\Delta p_1$ و $4+\Delta q_1$. برای سهولت تمام علائم منفی می‌توانند حذف گردند. روند فوق آنگاه با عامل درجه دوم تکرار می‌گردد.

	2	15	43	51	41	108	160
-3	0	-6	-27	-24	27	-108	0
-4	0	0	-8	-36	-32	36	-144
	2	9	8	-9	36	36	16

بنابراین

$$Q(x) = 2x^4 + 9x^3 + 8x^2 - 9x + 36, \qquad R_1 = 36, \qquad S_1 = 16$$

	2	9	8	−9	36	0
−3	0	−6	−9	27	−18	0
−4	0	0	−8	−12	36	−24
$Q(x)$	2	3	−9	6	72	
$xQ(x)$				6	54	−24

$$R_2 = 54, \qquad S_2 = -24, \qquad R_3 = 6, \qquad S_3 = 72$$

لهذا ، دو دستگاه معادله بصورت :

$$54\Delta p_0 + 6\Delta q_0 = 36$$
$$-24\Delta p_0 + 72\Delta q_0 = 16$$
$$\Delta p_0 = 0{\cdot}6, \qquad \Delta q_0 = 0{\cdot}4$$

در می‌آید .

باید توجه داشت ، که به کاربردن دقت بالا زمانیکه مقادیر هنوز دقیق نباشند در قدمهای اولیه تکرار غیر ضروری است . تعداد ارقام به کار رفته در محاسبه افزایش پیدا کرده‌بطوریکه تقریبهای نزدیک‌تر بدست می‌آید .

تقریب جدید برای عامل درجه دوم عبارتست‌از $x^2 + 3{\cdot}6x + 4{\cdot}4$ و دنباله محاسبات فوق تکرار می‌گردد .

	2	15	43	51	41	108	160
−3·6	0	−7·2	−28·08	−22·03	19·26	−120·02	0
−4·4	0	0	−8·80	−34·32	−26·92	23·53	−146·69
	2	7·8	6·12	−5·35	33·34	11·51	13·31

$$Q(x) = 2x^4 + 7{\cdot}8x^3 + 6{\cdot}12x^2 - 5{\cdot}35x + 33{\cdot}34, \qquad R_1 = 11{\cdot}51, \qquad S_1 = 13{\cdot}31$$

	2	7·80	6·12	−5·35	33·34	0·00
−3·6	0·00	−7·20	−2·16	17·42	−33·94	0·00
−4·4	0·00	0·00	−8·80	−2·64	21·29	−41·49
	2	0·60	−4·84	9·43	54·63	
				9·43	20·69	−41·49

$$R_2 = 20{\cdot}69, \qquad S_2 = -41{\cdot}49, \qquad R_3 = 9{\cdot}43, \qquad S_3 = 54{\cdot}63$$

آنگاه دستگاه دو معادله زیر حل می‌گردد .

$$20{\cdot}69\Delta p_1 + 9{\cdot}43\Delta q_1 = 11{\cdot}51$$
$$-41{\cdot}49\Delta p_1 + 54{\cdot}63\Delta q_1 = 13{\cdot}31$$
$$\Delta p_1 = 0{\cdot}33, \qquad \Delta q_1 = 0{\cdot}49$$

عامل درجه دوم جدید عبارتست‌از $x^2 + 3{\cdot}93x + 4{\cdot}89$ که بدستگاه معادلات‌زیر

$$15{\cdot}981\Delta p_2 + 6{\cdot}366\Delta q_2 = 1{\cdot}782$$
$$-31{\cdot}129\Delta p_2 + 40{\cdot}999\Delta q_2 = 2{\cdot}337$$

با جوابهای $\Delta p_2 = 0{\cdot}068$ و $\Delta q_2 = 0{\cdot}108$ منجر می‌گردد .

عامل درجه دوم جدید $x^2 + 3{\cdot}998x + 4{\cdot}998$ به‌دستگاه معادلات

$$16{\cdot}8943\Delta p_3 + 5{\cdot}0654\Delta q_3 = 0{\cdot}0438$$
$$-25{\cdot}3168\Delta p_3 + 37{\cdot}1457\Delta q_3 = -0{\cdot}0226$$

با جوابهای $\Delta p_3 = 0{\cdot}0020$ و $\Delta q_3 = 0{\cdot}0019$ ، و عامل جدید

$$x^2 + 4{\cdot}0000x + 4{\cdot}9999$$

منجر می‌گردد . دستگاه معادلات :

$$\begin{aligned}17{\cdot}002\,81\,\Delta p_4 + 4{\cdot}999\,00\,\Delta q_4 &= 0{\cdot}000\,45\\ -24{\cdot}994\,50\,\Delta p_4 + 36{\cdot}998\,81\,\Delta q_4 &= 0{\cdot}003\,65\end{aligned}$$

جوابهای $\Delta p_4 = 0{\cdot}000\,00$ و $\Delta q_4 = 0{\cdot}000\,09$ و عامل ، $x^2 + 4{\cdot}000\,00x + 4{\cdot}999\,99$ را می‌دهد .

تقسیم به این عامل درجه دوم باقیمانده $0{\cdot}000\,00x + 0{\cdot}000\,32$ را داده و بدین‌ترتیب روند به طور واقعی به عامل صحیح همگرا شده است . بعد از تقسیم به این عامل یک معادله درجه دوم نتیجه می‌شود و کاربرد بیشتر روش برستو دو عامل باقیمانده را ایجاد می‌نماید .

۶ ــ جدول صفحهٔ ۸۴ نتایج بدست آمده از کاربرد الگوریتم تفاضل ــ خارج قسمت برای چند جمله‌ای زیر را نشان می‌دهد .

$$P(x) = (x+2)(x+1)(x-1)(x-2)(x-4) = x^5 - 4x^4 - 5x^3 + 20x^2 + 4x - 16$$

ریشه $x = 4{\cdot}0$ بطور وضوح بین دو ستون صفرنشان داده شده است . دو زوج از ریشه‌های مساوی از حیث قدرمطلق ایجاد پنج ستون می‌کند که دو تا از ستونهای طرفین صفر بوده و ستون‌های ردیف ۲ و ۴ مقادیر مورد نیاز برای محاسبه ریشه‌ها را می‌دهد .

داریم .

$$\begin{aligned}A &= \lim_{j\to\infty}(q_{j+1}^{(r+1)} + q_j^{(r+2)}) = 0\\ B &= \lim_{j\to\infty}(q_j^{(r+1)}q_j^{(r+2)}) = -4\end{aligned}$$

ریشه‌ها عبارتنداز جوابهای معادله

$$x^2 - 4 = 0$$

که ریشه‌های $x = \pm 2$ را می‌دهد .

متشابها" با زوج دوم دارای

$$A = 0, \qquad B = -1$$

ریشه‌های $x = \pm 1$ را می‌دهد .

۷ ــ با استفاده از برنامه کامپیوتری و به کمک روش نیوتن و ضرب تودرتو ریشه مثبت معادله زیر را با دقت 10^{-7} پیدا کنید .

$$x^5 - 3.7x^4 + 7.4x^3 - 10.8x^2 + 10.8x - 6.8 = 0$$

زمان رانش برنامه ۰/۲۷ ثانیه

0·00000	4·00000	1·25000	0·00000	−4·00000	0·00000	0·20000	0·00000	−4·00000	0·00000	0·0000
0·00000	5·25000	−1·25000	−5·25000	3·20000	4·20000	−0·20000	−4·20000	3·80952	4·00000	0·0000
0·00000	4·00000	0·25000	−0·80000	−3·20000	0·80000	0·04762	−0·19048	−3·80952	0·19048	0·0000
0·00000	4·25000	−0·25000	−4·25000	3·04762	4·04762	−0·04762	−4·04762	3·76471	4·00000	0·0000
0·00000	4·00000	0·05952	−0·95238	−3·04762	0·95238	0·01176	−0·23529	−3·76471	0·23529	0·0000
0·00000	4·05952	−0·05952	−4·05952	3·01176	4·01176	−0·01176	−4·01176	3·75367	4·00000	0·0000
0·00000	4·00000	0·01471	−0·98824	−3·01176	0·98824	0·00293	−0·24633	−3·75367	0·24633	0·0000
0·00000	4·01471	−0·01471	−4·01471	3·00293	4·00293	−0·00293	−4·00293	3·75092	4·00000	0·0000
0·00000	4·00000	0·00367	−0·99707	−3·00293	0·99707	0·00073	−0·24908	−3·75092	0·24908	0·0000
0·00000	4·00367	−0·00367	−4·00367	3·00073	4·00073	−0·00073	−4·00073	3·75023	4·00000	0·0000
0·00000	4·00000	0·00092	−0·99927	−3·00073	0·99927	0·00018	−0·24977	−3·75023	0·24977	0·0000
0·00000	4·00092	−0·00092	−4·00092	3·00018	4·00018	−0·00018	−4·00018	3·75006	4·00000	0·0000
0·00000	4·00000	0·00023	−0·99982	−3·00018	0·99982	0·00005	−0·24994	−3·75006	0·24994	0·0000
0·00000	4·00023	−0·00023	−4·00023	3·00005	4·00005	−0·00005	−4·00005	3·75001	4·00000	0·0000
0·00000	4·00000	0·00006	−0·99995	−3·00005	0·99995	0·00001	−0·24999	−3·75001	0·24999	0·0000
0·00000	4·00006	−0·00006	−4·00006	3·00001	4·00001	−0·00001	−4·00001	3·75000	4·00000	0·0000
0·00000	4·00000	0·00001	−0·99999	−3·00001	0·99999	0·00000	−0·25000	−3·75000	0·25000	0·0000
0·00000	4·00001	−0·00001	−4·00001	3·00000	4·00000	−0·00000	−4·00000	3·75000	4·00000	0·0000
0·00000	4·00000	0·00000	−1·00000	−3·00000	1·00000	0·00000	−0·25000	−3·75000	0·25000	0·0000
0·00000	4·00000	−0·00000	−4·00000	3·00000	4·00000	−0·00000	−4·00000	3·75000	4·00000	0·0000
0·00000	4·00000	0·00000	−1·00000	−3·00000	1·00000	0·00000	−0·25000	−3·75000	0·25000	0·0000

ابتدا الگوریتم نیوتن و ضرب تودرتو را در نظر گرفته آنگاه برنامه کامپیوتری را می‌آوریم .

* الگوریتم (۳۰۱) نیوتن و ضرب تودرتو

برای $m = 0, 1, \ldots$

قرار می‌دهیم $z = x_m,\ b_n = a_n,\ c_n = b_n$

برای $k = n-1, \ldots, 1$

قرار می‌دهیم

$$\begin{aligned} b_k &= a_k + zb_{k+1} \\ c_k &= b_k + zc_{k+1} \\ b_0 &= a_0 + zb_1 \\ x_{m+1} &= x_m - b_0/c_1 \end{aligned}$$

برنامه کامپیوتری فوق در الگوریتم (۳۰۱) و نتایج در جدول (۳۰۳) آمده است .

ALGORITHM 3.1

FORTRAN PROGRAM FOR EXAMPLE 7

```
C     NEWTONS METHOD FOR FINDING A REAL ROOT OF A POLY-
C     NOMIAL P(X) = 0
      DIMENSION A(10)
      DATA (A(I), I = 1,6)/ -6.8, 10.8, -10.8,7.4, -3.7,1./
      READ (5,3) X
      WRITE (6,4)
      DO 2 J = 1,20
      B = A(6)
      C = A(6)
      DO 1 I = 1,4
      K = 6 - I
      B = A(K) + X*B
    1 C = B + X*C
      B = A(1) + X*B
      WRITE (6,5) J,X,B , C
      DELTAX = B/C
      IF(ABS(DELTAX) .LT. 1.E-7 .AND. ABS(B) .LT.
     11.E-7) STOP
    2 X = X - DELTAX
      WRITE (6,6)
      STOP
    3 FORMAT(E20.8)
    4 FORMAT(34H1NEWTONS METHOD FOR FINDING A REAL
     *21H ROOT OF A POLYNOMIAL//4X1HI10X1HX14X4HB(0)
     *13X4HC(1)/)
    5 FORMAT(I5, 3(1PE17.7))
    6 FORMAT(36H0FAILED TO CONVERGE IN 20 ITERATIONS)
      END
```

جدول (۳-۳)

COMPUTER RESULTS FOR EXAMPLE 7

I	X	B(0)	C(1)
1	1.5000000E 00	−1.0625001E−00	3.7124998E 00
2	1.7861953E 00	7.2393334E−01	9.6004875E 00
3	1.7107894E 00	8.0013633E−02	7.5470622E 00
4	1.7001875E 00	1.3663173E−03	7.2905675E 00
5	1.7000000E 00	4.7683716E−07	7.2861013E 00
6	1.7000000E 00	−1.1920929E−07	7.2860994E 00
7	1.7000000E 00	−5.9604645E−08	7.2860998E 00

■ مسائل

۱ــ مطلوبست محاسبه چند جمله‌ای $4x^4-2x^3+5x-3$ در نقطه $x=-2$ به وسیله روش ضرب تو در تو.

۲ــ محل تقریبی یک ریشه از x^3-3x^2+x-1 را با به‌کاربردن رشتهٔ ستورم تعیین کنید.

۳ ــ با به‌کار بردن روش برستو دو عامل درجه دوم از $2x^4+7x^3+5x^2-4x+32$ را تعیین کنید.

برای اولین تقریب سه جملهٔ آخر معادله را به کار برید.

۴ ــ الگوریتمی بنویسید که به کمک فرمول نیوتن و با استفاده از ضرب تودرتو بتوان ریشه‌های حقیقی یک چندجمله‌ای را یافت، آنگاه به کمک الگوریتم اخیر کلیه ریشه‌های حقیقی معادلهٔ $x^3-x-1=0$ را با دقت 10^{-4} محاسبه نمائید.

۵ ــ با استفاده از الگوریتم مسئلهٔ ۴ ریشه‌های حقیقی مثبت معادلات زیر را پیدا کنید (با دقت 10^{-4}).

(a) $x^4+6x^2-1=0$

(b) $3x^5-2x^3+x-2=0$

(c) $x^{12}-11x^{11}+8x^7-2=0$

(d) $x^4+2.8x^3-0.38x^2-6.3x-4.2=0$

۶ ــ با استفاده از قضیه ستورم تعداد و محل تقریبی ریشه‌های معادلهٔ زیر را پیدا کنید.

$$x^6+5x^5+4x^4+3x^3+2x^2+x+1=0$$

ثانیا به کمک برنامهٔ کامپیوتری کلیه ریشه‌های معادله فوق را با دقت 10^{-6} محاسبه نمائید.

۷ ــ با استفاده از روش برستو کلیه ریشه‌های معادلهٔ زیر را پیدا کنید. (با مقادیر اولیه متفاوت).

$$x^3-6x^2+11x-6=0$$

$(p_0, q_0)=(4,-3)$

$(p_0, q_0)=(5,-6)$

$(p_0, q_0)=(3,-2)$

۸ ــ با استفاده از روش برستو یک ریشه موهومی معادله زیر را پیدا کنید. (دقت محاسبه 10^{-7}).

$$x^5+0.597x^4-0.062701x^3+0.0834x^2-0.127x-0.367=0$$

۹ ــ مطلوبست محاسبه ریشه‌های موهومی معادلات زیر به کمک روش برستو.

(a) $x^4 + 3x^2 + 1 = 0$

(b) $x^4 + 2x^3 + 6x^2 - 13x + 48 = 0$

(c) $x^4 + 2x^2 + x + 1 = 0$

۱۰ــ برای نشان دادن خطرات در یافتن ریشه‌های چندجمله‌ای با رتبهٔ بالا معادلهٔ زیر را در نظر می‌گیریم .

$$x^7 - 28x^6 + 322x^5 - 1960x^4 + 6763x^3 - 13132x^2 + 13068x - 5040 = 0$$

ریشه‌های واقعی عبارتند از: $1,2,3,4,5,6,7$

با استفاده از برنامهٔ کامپیوتری ، به کمک روش نیوتن و ضرب تودرتو ، ریشه‌های معادلهٔ فوق را به ترتیب صعودی ، با انتخاب مقادیر اولیه 6.9، 5.9، 4.9 ، 3.9 ، 2.9، 1.9 ، 0.9 و معیار دقت 10^{-7} و آنگاه ریشه‌ها را به ترتیب نزولی یافته و در دو ستون چاپ نمائید .

ثانیا : با تغییر ضریب -13132 به -13133 نتایج قسمت فوق را یافته و پس از چاپ جوابها نتایج را توجیه نمائید .

۱۱ــ چندجمله‌ای $x^8 - 170x^6 + 739x^4 - 39712x^2 + 51200 = 0$ مفروض است .
معادلهٔ فوق دارای ریشه‌های $\pm\sqrt{2}$ ، ± 2 ، ± 8 و ± 10 می‌باشد .
با استفاده از الگوریتم نیوتن و ضرب تودرتو و با انتخاب مقادیر اولیه برای هریک مساوی ٪ 10 ریشه‌های واقعی ، ریشه‌های فوق را به ترتیب صعودی و آنگاه به ترتیب نزولی یافته و در ستون سوم پس از تغییر ضریب x^2 به -39710 ریشه‌ها را به ترتیب محاسبه و نتایج را توجیه نمائید .

۱۲ــ الگوریتم لمر ــ شور را برای $4x^3 - 8x^2 + 9x - 18 = 0$ به کار ببرید .

۱۳ــ قضیه ستورم را برای یافتن تعداد ریشه‌های حقیقی معادلهٔ زیر به کار ببرید .

$$x^6 + 4x^5 + 4x^4 - x^2 - 4x - 4 = 0$$

۱۴ــ با استفاده از روش مربع سازی ، ریشه‌های معادلهٔ زیر را پیدا کنید .

$$x^4 - 3x^3 - 54x^2 - 150x - 100 = 0$$

۱۵ــ با استفاده از الگوریتم تفاضل ــ خارج‌قسمت ریشه‌های معادلات زیر را پیدا کنید .

(a) $x^3 - 9x^2 + 26x - 24 = 0$

(b) $x^3 - 7x^2 + 10x - 2 = 0$

فصل چهارم

حل دستگاه معادلات خطی

۴.۱ ■ مقدمه

مسائل فیزیکی و عددی زیادی وجود دارد که جواب آنها به وسیله حل دستگاه معادلات خطی بدست می‌آید. زمانیکه تعداد مجهولات کم باشد مسئله می‌تواند نسبتا" ساده باشد، و اغلب در سطح مقدماتی در ریاضیات مطالعه می‌گردد. اگرچه، زمانیکه تعداد معادلات مستقل با تعداد مجهولات مساوی نباشد مشکلاتی وجود دارد، زمانیکه n معادله مستقل و n مجهول وجود داشته باشد مسئله دارای یک جواب یکتا می‌باشد. هنگامیکه تعداد چنین معادلاتی کمتر از n باشد در اینصورت جواب یکتا نمی‌تواند بدست آید.

اگر تعداد معادلات از n بیشتر باشد در اینصورت دو امکان وجود خواهد داشت. به علت وجود وابستگی خطی تعدادی از معادلات با سایرین ممکن است معادلات اضافی ایجاد کند. شکل دوم عبارتست از اینکه تعدادی از معادلات با هریک از سایر معادلات سازگار نباشند. اغلب این حالت‌زمانی رخ می‌دهد که سعی در برازاندن‌یک شکل‌ریاضی منتخب به یک مجموعه نتایج تجربی باشد.

در جستجو برای دست یافتن به خطای تجربی تعداد خیلی بزرگتری از مشاهدات نسبت به تعداد لازم برای بیان پارامترها در معادله گرفته می‌شود. روشی که به این موقعیت پرداخته است عبارتست از روش کمترین مربعات که در فصل ۷ آمده است. در این فصل به‌حالتیکه به n معادله خطی با n مجهول محدود شده و به صورت زیر نوشته می‌شود پرداخته و به صورت زیر خواهیم نوشت.

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned} \tag{4.1}$$

علامت اندیس مضاعف کاملا" ساده بوده، اگر درنظر بگیریم اندیس اولی مربوط به سطری است که عنصر در آن ظاهر می‌گردد و اندیس دوم مربوط به شماره ستون می‌باشد.

اکثرا" با شکل ماتریسی که مجموعه‌ای از ضرائب $a_{ij}\ (i,j=1,2,\ldots,n)$ را با ماتریس $\mathbf{A}$ و عناصر $x_j\ (j=1,2,\ldots,n)$ را به صورت بردار $\mathbf{X}$ نمایش می‌دهدآشنائی دارند. با به کار بردن قوانین جبر ماتریسی، مجموعه معادلات را می‌توان به صورت

$$\mathbf{AX}=\mathbf{B} \tag{4.2}$$

نمایش داد.

ازنظر ریاضی، در صورتیکه $\det(\mathbf{A})$ مخالف صفر باشد این معادلات دارای جوابی یکتا خواهد بود. از نقطه‌نظر عددی، از آنجائیکه مفهوم صفر در کامپیوتر نسبتا" صریح نمی‌باشد، در این مورد مشکلی وجود دارد. نتیجهٔ یک محاسبه که پاسخ آن باید صفر گرفته‌شود ممکن است مقدار 10^{-50} بدست آید، برای مثال نتیجه‌ای از داده، یا خطای گرد شده را می‌توان درنظر گرفت. بطور وضوح مقدار خیلی کوچکی برای دترمینان می‌تواند به اندازهٔ یک مقدار صفر مشکل آفرین باشد چون ممکن است این دو مقدار غیر قابل تشخیص باشد.

چنین مجموعه‌هائی از معادلات بد وضع نامیده می‌شوند. در حقیقت با تغییر جزئی در مقادیر ثابت آنها تغییرات خیلی بزرگ در جواب ایجاد می‌گردد. مثلا"، در مجموعه‌های زیر از معادلات به ازای 1% تغییر در ضرائب تقریبا" 200% تغییر در جواب مشاهده می‌شود.

$$\begin{array}{lll}
x+100y=3 & x+101y=3 & x+100y=3\\
\underline{x+100y=6} & \underline{x+100y=6} & \underline{x+101y=6}\\
\text{No solution} & \qquad y=-3 & \qquad y=3\\
 & \qquad x=306 & \qquad x=-297
\end{array} \tag{4.3}$$

به کمک نمودار، می‌توان دید که خطوط تقریبا" موازی بوده بنابراین تغییر خیلی کوچکی در شیب موجب تفاوت قابل توجهی در نقطه برخورد آنها خواهد بود.

دو نوع روش برای حل معادلات (4.1) مورد بحث قرار خواهد گرفت. روشهای مستقیم مبتنی بر روشهای حذف که خیلی شبیه به روشهائی است که در دبیرستان برای حل‌دستگاه معادله بررسی می‌شود. روشهای غیر مستقیم یا تکراری، که مبتنی بر محاسبه رشته‌ای از تقریبها بوده که احتمالا" به یک مقدار، به اندازه کافی نزدیک به مقدار واقعی همگرا می‌گردد. برای استفاده کننده حائز اهمیت است که خواص این نوع روشها را درک کرده بطوریکه با انتخاب مناسبی از روش، در موقعیت ویژه‌ای بتواند به کار برد. فایده روشهای مستقیم عبارتست از اینکه مقدار محاسبات ثابت بوده و قبلا" قابل بیان

می‌باشد، اما در روشهای تکراری، محاسبه باید به طور نامحدود ادامه‌یابد تا اینکه جوابها به مقداری با دقت کافی همگرا گردند.

پارامتر بعدی که باید درنظر گرفت عبارتست از تعداد ضرایبی که دارای مقدار صفر می‌باشند. برای مثال، در بیشتر مسائلی که در حل معادلات دیفرانسیل و معادلات با مشتقات جزئی پیش می‌آید، مجموعه‌ای بزرگ از دستگاه معادلات تولید می‌گردد که هر یک دارای فقط چند ضریب غیر صفر می‌باشند. چنین ماتریسی، ماتریس تنک نامیده می‌شود. این مطلب بعدا" نیز دیده خواهد شد که به طور قابل توجهی باعث افزایش جذابیت روشهای تکراری می‌شود، چون کاری که انجام می‌شود متناسب با تعداد عناصر غیر صفر است.

بهرحال در حالتیکه تعدادی از معادلات تنک دارای ساختمانهای خیلی ساده و معینی باشند، باید یک کلمه احتیاط را نیز به‌کار برد. در این حالت طرح روش مستقیمی که مقدار محاسبه به کار رفته مستقیما" متناسب با تعداد عناصر غیر صفر باشد ضروری است. در این صورت مزیت روشهای تکراری از دست رفته و طبیعت متناهی روش مستقیم این مطلب را خیلی مناسبتر می‌سازد.

۴.۲ ■ روشهای مستقیم

۴.۲.۱ ■ حذف به طریق گاوس

اساس روش مستقیم مبتنی بر حذف ساده مجهولات بوده، در مثالهای دبیرستان برای خیلی‌ها آشنا می‌باشد. در محاسبهٔ دستی پیدا کردن ساده‌ترین ترتیب حذف مجهولات با توجه به مهارت شخص عمل کننده تعیین می‌گردد. در اینجا دو نقطه ضعف وجود دارد، برای مثال، ترتیب اتفاقی اعمال برای محاسبه در یک کامپیوتر مناسب نبوده، وترتیب اصولی محاسبات ضروری بوده، بنابراین اگر در محاسبات وقفه‌ای ایجاد گردد شناسائی دلیل ایجاد وقفه امکان‌پذیر خواهد بود. جایگذاری که در آن محاسبه دستی با پاسخ $0=0$ مشخص می‌گردد وآیا در اینجا مسئله اصلی، نادرست فرموله شده، یا اینکه گامهای محاسبه بطور نادرست انجام گرفته‌اند روشن می‌باشد.

ابتدا روند سادهٔ حذف به طریق گاوس توضیح داده خواهد شد، اما باید به خاطر داشت برای روندی که ازنظر محاسباتی دقیق باشد، باید شامل روالهایی برای مقیاس ماتریس و محورگیری جزئی در اثناء حذف باشد. این موضوع بعدا" در این قسمت بحث می‌گردد. در این رفتار ساده فرض خواهیم کرد که تمام مخرجها مخالف صفر هستند. معادلاتیکه حل می‌گردند عبارتنداز:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{4.4}$$

اولین معادله برای بعد ذخیره شده و متغیر x_1 از $n-1$ معادله باقیمانده وسیله تفریق مضرب مناسبی از اولین معادله از هریک از سایر معادلات حذف می‌گردد . فرض می‌کنیم ضرائب اصلی باعلائم‌زیر نشان داده شده باشند .

$$a_{ij}^{(1)} = a_{ij}, \qquad i,j = 1,2,\ldots,n \tag{4.5}$$

$$b_i^{(1)} = b_i, \qquad i = 1,2,\ldots,n \tag{4.6}$$

ضرائب جدید پس از به کار بردن مضارب بدست آمده ،

$$m_{i1} = a_{i1}^{(1)}/a_{11}^{(1)}, \qquad i = 2,3,\ldots,n \tag{4.7}$$

و تشکیل عناصر جدیدی را می‌دهند .

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1}a_{1j}^{(1)}, \qquad \begin{matrix} i = 2,3,\ldots,n \\ j = 1,2,\ldots,n \end{matrix} \tag{4.8}$$

$$b_i^{(2)} = b_i^{(1)} - m_{i1}b^{(1)}, \qquad i = 2,3,\ldots,n \tag{4.9}$$

می‌توان دید که عناصر در ستون اول ، $j=1$ ، دارای مقادیر

$$a_{i1}^{(2)} = a_{i1}^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \cdot a_{11}^{(1)} = 0 \tag{4.10}$$

بوده بنابراین اولین متغیر x_1 در $n-1$ معادله بعدی حذف شده است . در صورتیکه از اولین سطر صرفنظر گردد معادلات دارای همان شکل معادلات (4.4) بوده اما با یک سطر و ستون کمتر .

روند قبلی ، $n-1$دفعه تواما " تکرار می‌گردد تا اینکه معادلهٔ باقیمانده فقط دارای یک مجهول گردد که می‌تواند به سادگی حل شود . درهر مرحله از روند زمانیکه متغیر x_k حذف می‌گردد ضرائب بصورت زیر تشکیل می‌شوند .

$$m_{ik} = a_{ik}^{(k)}/a_{kk}^{(k)}, \qquad i = k+1, k+2, \ldots, n \tag{4.11}$$

و عناصر جدید به صورت زیر تشکیل می‌شوند .

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)} \qquad \begin{matrix} i = k+1, k+2, \ldots, n \\ j = k, k+1, \ldots, n \end{matrix} \tag{4.12}$$

$$b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)} \qquad i = k+1, k+2, \ldots, n \tag{4.13}$$

نتیجه این روند حذف ، یک دستگاه بالامثلثی از معادلات داده شده وسیله معادلات زیر است .

$$
\begin{array}{rl}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} & \\
a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n = b_2^{(2)} & \\
0 \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{nn}^{(n)}x_n = b_n^{(n)} &
\end{array}
\tag{4.14}
$$

که در آن عناصر زیر قطر صفر می‌باشند . حل این معادلات به وسیله جایگذاری پسرو ساده می‌باشد . آخرین معادله دارای جواب

$$x_n = b_n^{(n)}/a_{nn}^{(n)}$$

بوده و این مقدار می‌تواند در پائین‌ترین معادله پس از جایگذاری، مقدار x_{n-1} را بدهد و الی آخر.

با به‌کار بردن معادلات قبلی مقادیر تمام متغیرها قابل محاسبه می‌باشند . جدولی از نتایج حذف به طریق گاوس در مثال (۴۰۱)داده شده است .

۴.۲.۲ ■ طرفهای سمت راست متعدد

بسیار اتفاق می‌افتد که چندین معادله با مجموعه ضرائب یکسان در ماتریس A با طرفهای سمت راست مختلف بایستی حل شوند . در این حالت برگزاری صحیح روش، از زمان لازم کامپیوتر به میزان قابل توجهی کم می‌کند . اگر معادلات(4.7) ، (4.8) ، (4.11) و (4.12) که گام‌های محاسبه‌ای برای تقلیل به شکل بالا مثلثی ، می‌باشد مورد آزمایش قرار گیرند می‌توان دید که این محاسبات مستقل از جملات طرف دوم می‌باشند . بعد از آنکه این روند تقلیلی یک مرتبه انجام گرفت ، مشروط بر آنکه ضرائب m_{ij} ذخیره گردند نیازی به تکرار نخواهد بود . اگر تمام طرفهای سمت راست از ابتدا معلوم باشند امکان اینکه تمام اینها به‌مجرد اینکه‌تبدیل به بالا مثلثی انجام شد ، عملیات لازم انجام گیرد موجود است به عبارت دیگر ، ذخیره ضرائب منجر به حل دستگاه با طرفهای سمت راست دیگر با حداقل زحمت می‌گردد .

معادلات باقیمانده (4.9) و (4.13) فقط به جملات b_i و ضرائب m_{ij} که هنگام تبدیل مثلثی در اولین رفتار ذخیره شده است بستگی دارد . بنابراین، فقط محاسبات در معادلات (4.9) و (4.13) برای طرفهای سمت راست بعدی ادامه پیدا می‌کند جالب است که تعداد عناصر صفر تولید شده درست مساوی تعداد ضرائب m_{ij} است ، بنابراین به عوض ذخیره کردن مقدار صفر در فضای ذخیره‌ای کامپیوترمقدار m_{ij} ذخیره‌می‌گردد .

برای مثال، در اولین ستون ضریبی که در سطر i ذخیره می‌گردد به صورت m_{i1} و معمولا" محل i,j شامل m_{ij} برای $j < i$ خواهد بود .

۴.۲.۳ ■ یافتن معکوس

حالت ویژه‌ای از یک مجموعه معادلات با چندین طرف راست زمانی رخ می‌دهد که عکس ماتریس غیر منفرد A مورد نیاز است. این وسیله جواب X از معادله $AX = I$ بدست می‌آید که در آن A دارای ضرائب طرف سمت چپ معادله (4.4) و طرف سمت راست عبارتست‌از:

$$I = \begin{bmatrix} 1 & 0 & 0 & & & 0 \\ 0 & 1 & 0 & & & \cdot \\ \cdot & 0 & 1 & & & \cdot \\ \cdot & \cdot & 0 & & & \cdot \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & & & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.15}$$

از حل به نوبت دستگاهها بااین طرفهای سمت راست، ستون‌های ماتریس معکوس A^{-1} بدست می‌آید. باید توجه داشت که فقط در حالات نادر، یافتن ماتریس A^{-1} ضروری است چون برای حل یک دستگاه معادلات نیازی به آن نیست، بهترین روش کم‌زحمت حذف به طریق گاوس می‌باشد. شاید جای آن باشد که اشاره شود روش دترمینان کرامر برای حل معادلات و فرمول $A^{-1} = \mathrm{Adj}\,A/\det(A)$ برای کاربرد کامپیوتر کاملا" نامناسب می‌باشد. دترمینان یک ماتریس کاملا"، به سادگی‌انجام یک ضرب در روند حذف گاوس محاسبه می‌گردد. اگر روند مطابق توضیح فوق ادامه یابد، آنگاه دترمینان عبارتست‌از حاصلضرب عناصر قطری ماتریس مثلثی که در روند جایگذاری پسرو به کار می‌رود.

۴.۲.۴ ■ محورگیری جزئی

برای ماتریسهای بزرگ می‌توان دید که حذف گاوس مشتمل بر تعداد اعمال حسابی قابل توجهی است، و در هر مرحله از روند، محاسبات، اعداد محاسبه شده در مرحله قبلی را به کار می‌برد. این موقعیتی است کلاسیک که خطای بزرگی می‌تواند ایجاد گردد و بنابراین حائز اهمیت است که تمام سعی خود را به منظور می‌نیمم کردن خطاهای ساخته شده به‌کار گیریم.

از معادلات (4.12) و (4.13) می‌توان دید عملی که بیشتر اوقات رخ می‌دهد عبارتست از ضرب بوسیله m_{ij}. در ضرب عدد هر خطایی که موجود باشد، در m_{ij} ضرب خواهد شد، بنابراین، بایستی این ضرائب به هراندازه که ممکن است‌کوچک‌گردد، و محققا" تا مقدار کمتر از واحد، بطوری که خطاها به وسیله ضرب بزرگ نشوند.

این عمل بسادگی انجام می‌گیرد. درصورتیکه عنصر، محوری، $a_{kk}^{(k)}$ بزرگترین عنصر $a_{ik}^{(k)}$ درهمان ستون برای $i \geqslant k$ باشد از اینرو داریم.

$$|m_{ij}| \leqslant 1, \qquad \begin{matrix} i = 1, 2, \dots, n \\ j < i \end{matrix} \tag{4.16}$$

برای ایفاء این پیشنهاد ، به گام اضافی که شامل محاسبه بسیار کمی است نیازمند می‌باشیم . در این مرحله از روند که در آن متغیر بعدی x_k موردنظر حذف می‌باشد در ستون موردنظر تحقیقی روی عناصر زیرقطر انجام گرفته ، و عنصریکه از لحاظ قدر مطلق بزرگترین است مشخص می‌گردد . آنگاه سطر شامل این عنصر با سطر محوری عوض شده بنابراین اکنون بزرگترین عنصر در محل محوری قرار می‌گیرد .

این عمل مضارب کوچکتر از واحد از لحاظ قدر مطلق را می‌دهد . عمل فوق در هر مرحله از روند انجام گرفته و هر کجا که ضرورت داشته باشد تعویض انجام می‌گیرد همچنین روند تعویض در صورتیکه یک عنصر محوری $a_{kk}^{(k)}$ صفر باشد نیز انجام می‌گیـــرد جز زمانیکه ماتریس منفرد ، یا تقریبا" منفرد ، باشد روند تحقیق حداقل یک عنصر غیر صفر در هر ستون را پیدا خواهد کرد ، بنابراین تقسیم به صفر رخ نخواهد داد . بایستی توجه داشت که روند محورگیری جزئی برای تامین دقت درصورتیکه مسئله بطور کلی بد وضع باشد اساسی است .

همچنین توسعه این ایده وبه‌کار بردن محورگیری کامل برای ماتریس امکان‌پذیر است .

این روند شامل تحقیق تمام زیرماتریسهای باقیمانده است که در آن بزرگترین قدرمطلق عناصر در محل محورگیری قرار گیرد . این عمل ممکن است نه تنها منجر به تغییر ترتیب سطرها بلکه به تغییر ترتیب متغیرها در معادلات گردد ، و نسبت به محورگیری جزئی دارای برنامه‌نویسی پیچیده‌تری می‌باشد . از آنجائیکه بدست آوردن دقت به وسیله محورگیری کامل خیلی قابل توجه نیست لذا استراتژی محورگیری کامل اغلب به کار نمی‌رود .

۴.۲.۵ ■ مقیاس کردن

با توجه مختصر می‌توان دید که استراتژی محورگیری جزئی به تنهائی کافی نمی‌باشد . اگر دو معادلهٔ زیر را درنظر بگیریم .

$$\begin{aligned} 4x + 3y &= 10 \\ 3x - 2y &= 12 \end{aligned} \tag{4.17}$$

آنگاه عنصر محوری قبلا" بزرگترین عنصر از لحاظ قدرمطلق می‌باشد . معهذا عمل ساده ضرب معادله دوم در عدد ۲ بدین معنی است که دو معادله محتاج به تعویض و قرار دادن بزرگترین عنصر از لحاظ قدرمطلق در محل محور است . بنابراین انتخاب

سطر محور اختیاری است .

مشخص شده که خاصیت اختیاری انتخاب سطر محور مانعی برای توسعه بیشتر روندهای دقیق حذفی می‌باشد . برای این مسئله، راه حل عبارتست از مقیاس کردن ماتریس طوریکه سطرها بیشتر به طریقی که تعریف می‌شود قابل مقایسه باشند . این عمل معمولا " با نرمال کردن به یکی از دو طریق زیر انجام می‌گیرد . سطرها می‌توانند به عنصریکه از لحاظ قدرمطلق بزرگترین است تقسیم شده،بنابراین بزرگترین عنصر جدید واحد خواهد بود ، یا اینکه عناصر هر سطر به

$$d_i = \sqrt{\left(\sum_{j=1}^{n} a_{ij}^2\right)}$$

تقسیم گردد . هم‌چنین مقیاس کردن می‌تواند تفاوتی در دقت جوابها ایجاد کند، روش متعارفی از مقیاس کردن که جامع و قابل قبول باشد وجود ندارد . آندسته که علاقمند باشند می‌توانند به کار (1963) Bauer و (1961) Wilkinson که در این موضوع می‌باشد مراجعه کنند .

۴.۲.۶ ■ بدوضعی Ill-conditioning

البته ، امکان اینکه حتی درصورتیکه محورگیری جزئی به‌کار گرفته شود روش متوقف گردد وجوددارد اما روش حذف گاوس نشانی از یک نوع بدوضعی که مسئله از آن ناشی شده است را خواهد داد . این مطلب می‌تواند بوسیلهٔ معادلات ساده دو متغیرهٔ زیر نشان داده شود .

$$\begin{array}{ll} 2x+3y=10 & 2x+3y=10 \\ \underline{4x+6y=20} & \underline{4x+6y=22} \\ \quad 0 = 0 & \quad 0 = 2 \end{array} \tag{4.18}$$

در اولین مثال ، روند حذفی عنصر خیلی کوچکی را نشان خواهد داد که نه فقط در محل محورگیری است بلکه همچنین در طرف سمت‌راست نیز قرار می‌گیرد . این مطلب دلالت می‌کند که دستگاه n معادله به‌طور خطی مستقل ازهم هستند . درحالت دوم طرف سمت راست نزدیک به صفر نمی‌باشد . این مطلب دلالت می‌کند که معادلات ناسازگار بوده و باید اشتباهی در فرمول مسئله رخ داده باشد . بدیهی است که در هر مرحله ، عنصر محوری باید بررسی شده و درصورتیکه اندازه عنصر محوری از مقدار کوچک انتخابی کمتر گردد تصمیم مناسبی اتخاذ شود .

۴.۲.۷ ■ اصلاح جواب

مانند یک پیش‌بینی مقدماتی ، زمانیکه جواب مجموعه معادلات پیدا شود برای

اینکه بدانیم آنها تاچه اندازه در دستگاه صدق می‌کنند باید مقادیر در معادلات اصلی گذاشته شوند. درصورتیکه خطاهای بزرگ وجود داشته باشند آنگاه بطور وضوح جوابها رضایت‌بخش نمی‌باشند. متاسفانه، در معادلات بد وضع حتی اگر جوابهای محاسبه شده از جوابهای واقعی اختلاف قابل‌توجهی داشته باشد ممکن است پس از جایگذاری خطاهای کاملا" کوچک پیدا شود.

فرض می‌کنیم جوابهای محاسبه شده به صورت

$$\mathbf{X}^{(0)} \equiv [x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}]^T \tag{4.19}$$

باشند.

باقیمانده‌ها را با دقت مضاعف با جایگذاری در معادله (4.4) تشکیل می‌دهیم.

$$\mathbf{R}^{(0)} \equiv [r_1^{(0)}, r_2^{(0)}, \ldots, r_n^{(0)}]^T \tag{4.20}$$

به‌صورت ماتریسی داریم

$$\mathbf{AX}^{(0)} - \mathbf{B} = \mathbf{R}^{(0)} \tag{4.21}$$

باقیمانده‌های $\mathbf{R}^{(0)}$ زمانیست که جوابها غیر قابل قبول باشند، اما آنها منجر به روشی برای اصلاح جواب که کاملا" موثر می‌باشد خواهد شد. ازآنجائیکه، جواب واقعی در

$$\mathbf{AX} - \mathbf{B} = \mathbf{0} \tag{4.22}$$

صدق می‌کند، آنگاه تفاضل معادله (4.21) از معادله (4.22) معادله

$$\mathbf{A}(\mathbf{X} - \mathbf{X}^{(0)}) = -\mathbf{R}^{(0)} \tag{4.23}$$

را می‌دهد.

مقدار $\mathbf{X} - \mathbf{X}^{(0)}$ مقداریست‌که بایستی‌به‌اولین جواب محاسبه شده به‌منظور بدست‌آوردن جواب درست اضافه گردد. درصورتیکه‌این $\mathbf{X}^{(0)}$ درست محاسبه‌گردد مسئله به طور کامل حل می‌شود. در حقیقت، $\mathbf{X} - \mathbf{X}^{(0)}$ مانند قبل به وسیله حل دستگاه معادلات خطی پیدا می‌گردد. باید توجه داشت که به هر حال زمان لازم به‌طور قابل توجهی کاهش یافته و این به دلیل تقلیل به صورت مثلثی و ذخیره ضرائب است. حل معادله (4.23) و به‌کار بردن این جواب برای اصلاح $\mathbf{X}^{(0)}$ این مطلب را مشخص می‌کند که اغلب دقت را بطور قابل توجهی افزایش داده و حداقل یک اصلاح از این نوع در حالاتی که در آن دقت موردتوجه‌باشد توصیه می‌گردد. $\mathbf{E}^{(0)}$ را جواب $\mathbf{AE} = -\mathbf{R}^{(0)}$ می‌گیریم‌وتقریب‌جدید برای جواب وسیله‌ی $\mathbf{X}^{(1)} = \mathbf{X}^{(0)} + \mathbf{E}^{(0)}$ داده می‌شود.

اصلاحات بیشتری با تشکیل باقیمانده‌های

$$\mathbf{AX}^{(p)} - \mathbf{B} = \mathbf{R}^{(p)} \tag{4.24}$$

به‌صورت دقت مضاعف و آنگاه حل مکرر معادلات

$$\mathbf{AE} = -\mathbf{R}^{(p)}, \qquad p = 1, 2, \ldots \tag{4.25}$$

می‌تواند ساخته شود .
جوابهای $\mathbf{E}^{(p)}$ این معادلات تقریب‌های جدیدی برای جوابها را می‌دهند .

$$\mathbf{X}^{(p+1)} = \mathbf{X}^{(p)} + \mathbf{E}^{(p)} \tag{4.26}$$

این روند معمولا" بیش از یک یا دوبار انجام نمی‌گیرد ، چون دفعات بیشتر اصلاحات مختصری ایجاد می‌نماید .

۴.۲.۸ ■ سایر روشهای مستقیم

در کتابهای متعدد چندین نوع از حذف به طریق گاوس توضیح داده شده ، برای مثال طرحی که برای کار با ماشین رومیزی معمول است عبارتستاز حذف جردن . در این نوع شکل ماتریس آخری ، بعد از حذف ، قطری است . هر معادله دارای فقط یک متغیر بوده ، بنابراین ، روند جایگذاری پسرو به‌کار نمی‌رود و مقادیر متغیرها می‌توانند مستقیما" بدست آیند . روش خیلی شبیه به روش حذف گاوس ادامه پیدا کرده اما در هر مرحله متغیر x_k نه تنها از معادلات بعدی بلکه همچنین از تمام معادلات قبلی حذف می‌گردد .معادلات (4.12) و (4.13) برای تمام‌مقادیر $i \neq k$ به‌کار می‌روند . حذف‌جردن تقریبا" تعداد $n^3/2$ عمل در مقایسه با $n^3/3$ عمل در حذف گاوس احتیاج دارد ،بنابراین برای کاربرد عمومی توصیه نمی‌گردد .

گروه دیگری از روشها می‌توانند تحت عنوان تجزیه مثلثی توضیح داده شوند ، اینها شامل روشهای مختلف Crout ، Doolittle و Choleski می‌باشند . روش محاسبه به گونه‌ای متفاوت با حذف گاوس ارائه می‌شود اگرچه روشها خیلی شبیه‌اند . این روشها مبتنی بر یک سری از ضرائب است که به وسیله روند جایگذاری پسرو ماتریس را به صورت مثلثی تبدیل می‌کند . این روش نسبت به حذف گاوس برای آن دسته از کامپیوترها که دارای امکان جمع‌کردن حاصلضربهایی به صورت $\sum_{j=1}^{n} a_j b_j$ درطول‌مضاعف است دارای مزیتی می‌باشد . جذابیت این امکان عبارتست از اینکه جریمهٔ زمانی کار با طول مضاعف معمولی را ندارد . ضرب یک عدد در کامپیوتر همیشه مشتمل بر تولید عددی با طول ـ مضاعف ، قبل از قطع کردن ، و ادامهٔ عملیات می‌باشد .
ماشینهائیکه موردنظرما هستند برش به طول معمولی بعد از اینکه جمع حاصلضربها انجام گرفت صورت می‌گیرد ازآنجائیکه حذف گاوس مجموع حاصلضربها را معرفی نمی‌کند روش تجزیه مثلثی جائیکه امکان جمع با طول ـ مضاعف باشد ترجیح داده می‌شود .
فرض شده است که تجزیه مثلثی از نظر تئوری امکان‌پذیر بوده ، یعنی ماتریس می‌تواند به صورت حاصلضرب دوماتریس نوشته شود .

$$\mathbf{A} = \mathbf{L}.\mathbf{U} \tag{4.27}$$

که در آن U ماتریس بالا مثلثی L ماتریس پائین‌مثلثی می‌باشد. به‌محض اینکه این ماتریس‌ها پیدا شوند مجموعه معادلات در دو مرحله حل می‌شوند. هریک از اینها شامل حل یک مجموعه معادلات با یک ماتریس مثلثی بوده، این عمل باعث سادگی می‌شود و به وسیله جایگذاری پیشرو و پسرو انجام می‌گیرد. معادله به صورت

$$\mathbf{L.U.X=B} \tag{4.28}$$

درآمده و بردار Y را طوری می‌یابیم که $\mathbf{L.Y=B}$ بوده و آنگاه معادلات $\mathbf{U.X=Y}$ را حل می‌کنیم. معرفی محورگیری مانند حذف گاوس امکان‌پذیر بوده، و این عمل معمولا" انجام می‌گیرد. اگرچه وسیله Wilkinson (1961) برای یک ماتریس مثبت متقارن نشان داده شده است که خطاها وسیله حذف محورگیری جزئی به‌طور قابل‌توجهی اضافه نمی‌گردد.

این روش را وسیله مجموعه معادلات سه متغیره‌ای تشریح می‌کنیم. ضرائب را باید طوری پیدا کنیم که

$$\begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{32} \end{pmatrix}\begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{4.28}$$

یک عنصر آزادی در انتخاب مقادیر L و U وجود دارد و در این حالت، برای سهولت عناصر قطری ماتریس U مساوی واحد درنظر گرفته شده‌اند.

ضرائب L و U با مساوی هم قراردادن ضرائب طرفهای راست و چپ معادله (4.28) قابل محاسبه می‌باشند. معادلات مربوط به ستون اول A مقادیر l_{11}, l_{21}, l_{31} را مستقیما" می‌دهد.

$$l_{11}=a_{11}, \qquad l_{21}=a_{21}, \qquad l_{31}=a_{31} \tag{4.30}$$

معادلات مربوط به ستون دوم می‌توانند برای مقادیر u_{12}, l_{22}, l_{32} حل گردند.

$$l_{11}u_{12}=a_{12}, \qquad l_{21}u_{12}+l_{22}=a_{22}, \qquad l_{31}u_{12}+l_{32}=a_{32} \tag{4.31}$$

و معادلات مربوط به ستون سوم می‌توانند مقادیر u_{13}, u_{23}, l_{33} را بدهد.

$$\begin{aligned} l_{11}u_{13} &= a_{13} \\ l_{21}u_{13}+l_{22}u_{23} &= a_{23} \\ l_{31}u_{13}+l_{32}u_{23}+l_{33} &= a_{33} \end{aligned} \tag{4.32}$$

شکل کلی این الگوریتم در Keller (1966) و Isaacson توضیح داده شده است. تعداد عملیات مانند حذف گاوس بوده و از نقطه نظر امکان جمع حاصلضربها به صورت طول مضاعف، روشی است که برای بیشتر کاربردهای عمومی که روش مستقیم مورد احتیاج باشد توصیه می‌گردد.

درحالتیکه ماتریس متقارن باشد، کاهش مقدار عملیات به خاطر مزیت متقارن بودن امکان‌پذیر می‌باشد. درصورتیکه عناصر قطری L و U مساوی گردند. آنگاه $U = L^T$ و تنها عناصر L نیاز به محاسبه یا ذخیره می‌باشد. این روش موسوم به تجزیه به عوامل Choleski می‌باشد.

روش تجزیه مثلثی در فصل مربوط به یافتن مقادیر ویژه دوباره اشاره می‌گردد که در آن اساس روش، یافتن تمام مقادیر ویژه یک ماتریس می‌باشد. یک مثال عددی در (۴.۴) داده شده است.

۴.۲.۹ ■ تعدادی شکلهای ویژه ماتریس

شکلی از ماتریس که بطور فراوان ظاهر می‌گردد عبارتست از ماتریس سه قطری که دارای شکل

$$\begin{bmatrix} a_1 & c_1 & & & & \\ b_2 & a_2 & c_2 & & \mathbf{0} & \\ & \ddots & \ddots & \ddots & & \\ & & & b_{n-1} & a_{n-1} & c_{n-1} \\ & \mathbf{0} & & & b_n & a_n \end{bmatrix}$$

می‌باشد. تمام عناصر آن به استثنای عناصر روی قطر اصلی، بالا و پائین قطر اصلی صفر می‌باشند. چنین ماتریسی می‌تواند در حل مسائل مقدار ــ کرانه‌ای در معادلات دیفرانسیل معمولی در حالت $|a_i| \geqslant |b_i| + |c_i|$ رخ دهد.

بعدا" ملاحظه خواهد شد که این شرط، کاربرد روشهای تکراری که در فصل بعد به کار می‌رود را مطمئن می‌سازد. همچنین، از آنجائیکه این روشهای تکراری به ویژه برای ماتریسها با تعداد بزرگی از عناصر صفر مناسب می‌باشند، ممکن است روشهای تکراری را برای این نوع معادله به کار برد. در حقیقت، به علت اینکه عناصر خیلی نزدیک به قطر متمرکز شده‌اند، امکان بکارگیری نوعی از حذف گاوس که فقط عناصر غیر صفر را محاسبه کرده و از موقعیت تنکی به طور کامل استفاده شود وجود دارد. از آنجائیکه کار لازم در روش تکراری بستگی به تعداد تکرارها دارد روش مستقیم در این حالت ترجیح داده می‌شود.

اگر معادله ماتریسی به‌صورت $AX = V$ باشد آنگاه معادلاتی برای کاهش به‌صورت مثلثی در این الگوریتم که به الگوریتم توماس Thomas موسوم است تشکیل می‌گردد.

$$\alpha_1 = a_1 \qquad \gamma_1 = c_1/\alpha_1 \qquad u_1 = v_1/\alpha_1 \tag{4.34}$$

$$\begin{aligned} &\alpha_i = a_i - b_i\gamma_{i-1}, && i = 2, 3, \ldots, n \\ &u_i = (v_i - b_i u_{i-1})/\alpha_i, && i = 2, 3, \ldots, n \\ &\gamma_i = c_i/\alpha_i && i = 2, 3, \ldots, n-1 \end{aligned} \tag{4.35}$$

و جواب جایگذاری پسرو با

$$\begin{aligned} &x_n = u_n, \\ &x_i = u_i - \gamma_i x_{i+1}, \qquad i = n-1, n-2, \ldots, 1 \end{aligned} \tag{4.36}$$

داده می‌شود .

نتایج عددی برای این روش در مثال(۴۰۵)داده شده است .

برای قوت فرموله کردن محاسبه در این روش می‌توان نشان داد که برای مجموعه‌های معینی از a_i ، b_i و c_i که عموما" رخ می‌دهند ، ضرائب γ_i دارای قدرمطلق کوچکتر از واحد با همان تاثیر در حذف گاوس معمولی وسیله محورگیری می‌باشند . این مطلب در Isaacson و Keller (1966) مورد بحث قرار گرفته است . حقیقت قابل‌توجه درباره الگوریتم فوق عبارتستاز اینکه تعداد عملیات موردنیاز تقریبا" مساوی $5n$ در مقام مقایسه با $n^3/3$ حذف گاوس برای n بزرگ در یک ماتریس کامل می‌باشد . همچنین بایستی توجه داشت که می‌توان روش مشابهی برای ماتریسهایی که عناصر آنها پیرامون قطر متمرکز شده‌اند و در آن عرض نوار بزرگتر از ۳ باشد طراحی کرد .

صورت دیگری از ماتریس که عموما" رخ می‌دهد عبارتست ازماتریس بلوکی سه قطری که در حل تفاضل ــ متناهی معادلات با مشتقات جزئی رخ می‌دهد . این صورت شبیه معادله (4.33) بوده اما در این حالت عناصر به‌صورت ماتریس می‌باشند . معادلات (4.34) ، (4.35) و (4.36) مشروط برآنکه کمیت ماتریسی α_i^{-1} جایگزین خارج قسمت α_i گردد . برای تبدیل به صورت مثلثی می‌توانند به کار روند . روند شامل محاسبه معکوس یک ماتریس بوده اما این عمل وسیله حذف گاوس می‌تواند ارزانتر انجام گیرد . این روش به طور قابل توجهی خیلی موثرتر از تکرار ساده یا حذف گاوس کامل می‌باشد . توضیحات بیشتر در Smith (1965) موجود می‌باشد .

۴.۳ ■ روشهای تکراری

۴.۳.۱ ■ روش ژاکوبی Jacobi

در فصل ۲ دیدیم که به کار بردن یک روش تکراری که جواب را تا حصول همگرائی بهتر کند امکان‌پذیر است . این ایده به‌سهولت می‌تواند برای دستگاه معادلات خطی ، همانطوریکه وسیله مثال زیر نشان داده شده است ، به‌کار رود .

معادلات زیر را درنظر می‌گیریم :

$$\begin{aligned} 10x_1 + \quad x_2 + \quad x_3 &= 24 \\ -x_1 + 20x_2 + \quad x_3 &= 21 \\ x_1 - \quad 2x_2 + 100x_3 &= 300 \end{aligned} \tag{4.37}$$

به‌صورت زیر داریم :

$$\begin{aligned} x_1 &= \tfrac{1}{10}(24 - x_2 - x_3) \\ x_2 &= \tfrac{1}{20}(21 + x_1 - x_3) \\ x_3 &= \tfrac{1}{100}(300 - x_1 + 2x_2) \end{aligned} \tag{4.38}$$

درصورتیکه تقریب اولیه به‌صورت $(0,0,0)$ باشد آنگاه تکرار دیگر $(2{\cdot}4, 1{\cdot}05, 3{\cdot}00)$ بوده و تکرارهای متوالی مقادیر $(1{\cdot}995, 1{\cdot}02, 2{\cdot}997)$ و $(1{\cdot}9983, 0{\cdot}9999, 2{\cdot}999\,55)$ می‌باشند . جایگذاری مقادیر $(2, 1, 3)$ نشان می‌دهد که روند سریعا" به جواب واقعی همگرا می‌گردد .

محققا" روشن است که تکرار می‌تواند روش خیلی مفیدی برای حل باشد . حال نشان می‌دهیم که این روش تحت شرایط معینی می‌تواند روش خیلی ناپایداری باشد . معادلات (4.37) در ترتیب دیگری می‌تواند معادلات زیر را بدهد .

$$\begin{aligned} x_1 &= 300 + \ 2x_2 - 100x_3 \\ x_2 &= \ 24 - 10x_1 - \quad x_3 \\ x_3 &= \ 21 + \quad x_1 - \ 20x_2 \end{aligned} \tag{4.39}$$

با به‌کار بردن مقادیر $(0,0,0)$ به‌عنوان‌اولین‌تقریب ،تقریبهای‌متوالی $(300, 24, 21)$ و $(-3879, -2997, -1752-)$ را داده و امکان اینکه این رشته به مقادیر $(2, 1, 3)$ همگرا گردد وجود نخواهد داشت .

بدست آوردن شرایطی که همگرائی را مطمئن سازد امریست ضروری. این موضوع بعدا" در این فصل توضیح داده‌می‌شود .

روش فوق که به روش ژاکوبی موسوم است خیلی به‌ندرت به کار می‌رود زیرا نیازمند به اصلاحات متعددی است که بتواند سرعت همگرائی را بالا ببرد . اگرچه، تحت شرایط معینی سایر روش‌ها ممکن است همگرا نباشند و احتمالا" روش ژاکوبی ممکن است مناسب باشد .

۴.۳.۲ ■ روش گاوس-سایدل

در مطالعهٔ مثال ساده فوق خواننده ممکن است تعجب کند که چرا یک دسته از مقادیر پیدا شده و همگی در معادلات جایگذاری می‌گردند . منطقی به‌نظر می‌رسد که به

محض اینکه مقادیر جدید پیدا شوند فورا" در معادلات بعدی جایگذاری شوند. این ایده اساس روش گاوس ــ سایدل بوده که می‌تواند اثری قابل توجه در بهتر شدن سرعت همگرائی معادلات معینی داشته باشد.

به منظور دست یافتن به طریقه‌ای که در آن روشهای مختلف کارمی‌کند مناسب است که ضرائب به سه گروه تقسیم شوند، یعنی، مجموعه عناصر قطری، عناصر بالا و پائین قطر اصلی. ساده‌ترین راه بیان این روشها به کار بردن صورت ماتریسی است. ماتریس A به سه قسمت،متناظر با سه مجموعه از ضرائب بحث شده فوق، تجزیه می‌شود. معادلهٔ

$$\mathbf{AX} = \mathbf{B}$$

به

$$(\mathbf{L} + \mathbf{D} + \mathbf{U})\mathbf{X} = \mathbf{B} \tag{4.40}$$

تبدیل می‌شود.

برای سهولت معادلات را با تقسیم به عناصر قطری مقیاس کرده بنابراین D تبدیل به ماتریس واحد I خواهد شد.

روش ژاکوبی از انتقال تمام جمل به استثنای جملات قطری، با به‌کار بردن عمل تکرار به طریق زیر نتیجه می‌گردد.

$$\mathbf{X}^{(r+1)} = (-\mathbf{L} - \mathbf{U})\mathbf{X}^{(r)} + \mathbf{B}, \qquad r = 0, 1, \ldots \tag{4.41}$$

به هر حال روش گاوس ــ سایدل عناصر $x_1^{(r+1)}$, $x_2^{(r+1)}$ و غیره را بمحض محاسبه شدن در سمت راست قرار می‌دهد. بدین‌ترتیب معادلات تکراری به‌صورت

$$\mathbf{X}^{(r+1)} = -\mathbf{L}.\mathbf{X}^{(r+1)} - \mathbf{U}\mathbf{X}^{(r)} + \mathbf{B}$$

یا

$$(\mathbf{I} + \mathbf{L})\mathbf{X}^{(r+1)} = -\mathbf{U}\mathbf{X}^{(r)} + \mathbf{B} \tag{4.42}$$

درمی‌آید.

با نوشتن کامل این معادلات داریم.

$$\begin{aligned}
x_1^{(r+1)} &= -(a_{12}x_2^{(r)} + a_{13}x_3^{(r)} + \cdots + a_{1n}x_n^{(r)}) + b_1 \\
x_2^{(r+1)} &= -(a_{21}x_1^{(r+1)} + a_{23}x_3^{(r)} + \cdots + a_{2n}x_n^{(r)}) + b_2 \\
x_3^{(r+1)} &= -(a_{31}x_1^{(r+1)} + a_{32}x_2^{(r+1)} + \cdots + a_{3n}x_n^{(r)}) + b_3 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_n^{(r+1)} &= -(a_{n1}x_1^{(r+1)} + a_{n2}x_2^{(r+1)} + \cdots + a_{n,n-1}x_{n-1}^{(r+1)}) + b_n
\end{aligned} \tag{4.43}$$

زمانیکه هر دو روش ژاکوبی و گاوس ــ سایدل همگرا شوند می‌توان نشان داد که روش گاوس ــ سایدل نسبت به روش ژاکوبی سریعتر همگرا است. مثال(۴۰۳)نتایج روش ژاکوبی و گاوس ــ سایدل را نشان می‌دهد، روش فوق تخفیف در زیر مورد بحث قرار می‌گیرد. اختلاف در سرعت همگرائی کاملا" تمیز داده می‌شود.

۴.۳.۳ ■ فوق تخفیف

در فصل ۱ نشان داده شد که اگر یک روند تکراری به کندی همگرا گردد ، بعضی مواقع انتخاب گامهای بزرگتر از مقدار محاسبه شده امکان‌پذیر بوده ، بدین‌ترتیب همگرائی تسریع می‌گردد . این شیوه تقریبا" همیشه برای دستگاه معادلات خطی که در حل معادلات دیفرانسیل بیضوی ظاهر می‌گردد به کار می‌رود . این معادلات در شرایطی که همگرائی را تضمین کند صدق کرده اما سرعت همگرائی ، به ویژه برای یک دستگاه بزرگ ، خیلی کند است . بنابراین مقادیر محاسبه شده از روند گاوس ــ سایدل در حکم مقادیر میانی هستند و معادلهٔ زیر برای یافتن مقادیر پیراسته به‌کار می‌رود .

$$\mathbf{X}^{(r+1)} = \mathbf{X}^{(r)} + \omega(\tilde{\mathbf{X}}^{(r+1)} - \mathbf{X}^{(r)}) \qquad (4.44)$$

$\tilde{\mathbf{X}}^{(r+1)}$ مقدار محاسبه شده به وسیله روند گاوس ــ سایدل می‌باشد . از معادله می‌توان دید که مقدار جدید با ضرب نمو حاصل از تقریب آخری در عامل ω بدست می‌آید . زمانیکه $\omega > 1$ روش فوق ــ تخفیف را داریم و در صورتیکه $\omega < 1$ گفته می‌شود که دستگاه معادلات تحت تخفیف قرار گرفته است برای معادلات بیضوی مقدار ω معمولا" در نامساوی $1 < \omega < 2$ صدق می‌کند به بیان ماتریسی معادله به‌صورت زیر درمی‌آید .

$$\begin{aligned}\mathbf{X}^{(r+1)} &= \mathbf{X}^{(r)} + \omega(-\mathbf{L}\mathbf{X}^{(r+1)} - \mathbf{U}\mathbf{X}^{(r)} + \mathbf{B} - \mathbf{X}^{(r)}) \\ &= [-\omega\mathbf{U} + (1-\omega)\mathbf{I}]\mathbf{X}^{(r)} - \omega\mathbf{L}\mathbf{X}^{(r+1)} + \omega\mathbf{B} \qquad (4.45)\end{aligned}$$

آشکارا ، در به کار بردن روش فوق ــ تخفیف انتخاب ω مشکل اصلی می‌باشد . البته ، ماتریسهایی که در حل معادلات با مشتقات جزئی ظاهر می‌شوند اغلب دارای شکلهای سادهٔ ویژه‌ای بوده و برای بعضی از اینها محاسبه موثرترین ω امکان‌پذیر می‌باشد . به‌خصوص بیشتر تخمین زدن به کمتر تخمین زدن آن ترجیح دارد . برای یک ماتریس خیلی کلی ، به‌کار بردن فوق ــ تخفیف ، در صورتیکه یک سری تجربیات با پارامترهای مختلف تخفیف برای بیان بیشتر مقادیر مناسب امکان‌پذیر باشد ، میسر خواهد بود .

۴.۳.۴ ■ همگرائی روشهای تکراری

به منظور بررسی روشهای متعدد ، معادلات فوق را به صورت کلی زیر می‌نویسم .

$$\mathbf{X}^{(r+1)} = \mathbf{P}\mathbf{X}^{(r)} + \mathbf{C} \qquad (4.46)$$

جواب آخری $\mathbf{X}$ به‌وسیله معادله

$$\mathbf{X} = \mathbf{P}\mathbf{X} + \mathbf{C} \qquad (4.47)$$

مشخص می‌شود ، بنابراین خطای $\mathbf{E}^{(r)} = \mathbf{X} - \mathbf{X}^{(r)}$ به‌وسیلهٔ

$$\begin{aligned}\mathbf{X} - \mathbf{X}^{(r+1)} &= \mathbf{P}(\mathbf{X} - \mathbf{X}^{(r)}) \\ \mathbf{E}^{(r+1)} &= \mathbf{P}\mathbf{E}^{(r)} \qquad (4.48) \\ &= \mathbf{P}^{r+1}\mathbf{E}^{(0)}\end{aligned}$$

داده می‌شود .

بدیهی است که همگرائی تنها زمانی انجام می‌گیرد که ماتریس $\mathbf{P}$ متوالیا " منجر به تقلیل خطای $\mathbf{E}^{(r)}$ شود ، شرطی که وضوحا " ضروری است عبارتست از اینکه تمام مقادیر ویژه $\mathbf{P}$ دارای قدرمطلق کوچکتر از واحد باشند . به‌عبارت دیگر اگر $|\lambda_i| > 1$ و $\mathbf{v}_i$ یک بردار ویژهٔ متناظر آن باشد آنگاه با قراردادن $\mathbf{E}^{(0)} = \mathbf{v}_i$

$$\begin{aligned} \mathbf{E}^{(r+1)} &= \mathbf{P}^{r+1}\mathbf{v}_i \\ &= \lambda^{r+1}\mathbf{v}_i \end{aligned} \tag{4.49}$$

که با افزایش r بدون کران افزایش می‌یابد . شرط $|\lambda_i| < 1$ همچنین‌یک‌شرط کافی‌بوده و بنابراین ، اطلاع از مقادیر ویژه بیان خواهد کرد که آیا تکرار همگرا خواهد شد یا نه . تعریف مقدار ویژه و بردار ویژه در فصل ۹ داده شده است .

تعیین مقادیر ویژه به نوبهٔ خود مشکل بوده و اما شرایط متعددی وجود دارد که می‌توان کنترل کرد که کدامیک ، کران فوقانی برای قدر مطلق مقادیر ویژه خواهد بود . در صورتیکه یک‌چنین کران بالایی کمتر از واحد باشد روند همگرا خواهد شد . اگر این کرانهای بالا بزرگتر از واحد باشند آنگاه هنوز امکان اینکه مقادیر ویژه دارای قدر مطلقهای کوچکتر از واحد باشد وجود دارد ، بنابراین شرط کران بالایی برای همگرائی شرط لازم نیست .

در اولین دو روند تکراری موردنظر فوق ، ماتریس $\mathbf{P}$ دارای شکلهای

$$-\mathbf{L}-\mathbf{U} \quad \text{ژاکوبی} \tag{4.50}$$

$$(\mathbf{I}+\mathbf{L})^{-1}(-\mathbf{U}) \quad \text{گاوس‐سایدل} \tag{4.51}$$

می‌باشد .

می‌توان نشان داد که شرط غالب‌قطری ماتریس $\mathbf{A}$ یک شرط کافی برای اطمینان همگرائی روند در ماتریسهای فوق می‌باشد . (برای مثال به Keller (1966) و Isaacson مراجعه شود) . ماتریسی اکیدا " غالب‌قطری گفته می‌شود که داشته باشیم :

$$d_r < 1 \qquad \text{برای} \quad r = 1, 2, \ldots, n \tag{4.52}$$

که در آن :

$$d_r = \frac{\sum_{j=1}^{n}{}' |a_{rj}|}{|a_{rr}|} \tag{4.53}$$

علامت ′ دلالت می‌کند که مقدار a_{rr} از مجموع حذف گردیده است . اگر $d_r \leqslant 1$ برای $r = 1, 2, .., n$ ، و برای حداقل یک مقدار از r داشته باشیم $d_r < 1$ آنگاه‌ماتریس به‌طور ضعیف‌غالب قطری گفته می‌شود . این شرط برای همگرائی روند تکراری‌یک‌شرط کافی می‌باشد . شرط دیگری که همگرائی‌را تضمین خواهد کرد زمانی است که ماتریس $\mathbf{A}$ معین

مثبت باشد . چنین ماتریسهایی اغلب در مسائل فیزیکی و در حل مسائل کمترین مربعات رخ می‌دهند . ماتریس درصورتی معین مثبت گفته می‌شود که برای هر بردار غیر صفر داشته باشیم $X^TAX > 0$. از آنجائیکه بررسی این خاصیت خیلی مشکل است ، خاصیت غالب قطری بیشتر برای بررسی اینکه آیا همگرائی تضمین می‌شود یا نه به‌کار می‌رود . برای همگرائی روش فوق ــ تخفیف در روشهای معین باشکلهای ویژه نتایج معلومی را می‌توان گرفت . خوانندهٔ مشتاق را به کتاب Varga (1962) که در آن این مسائل به گونه‌ای دقیق و مفصل بحث شده ارجاع می‌دهیم .

۴.۴ ■ ماتریسهای تُنک

کلمه تنک برای بیان ماتریسی که‌دارای تعداد زیادی از عناصر صفر است به‌کار می‌رود . برای این نوع ماتریس، مقدار زمان محاسبه‌ای برای روشهای گوناگون یا طرحهای تکراری از الگوی معمولی تبعیت نمی‌کند . یک روش تکراری معمولا" mn^2 عمل برای حل احتیاج داشته که در آن m تعداد تکرارها و n تعداد معادلات می‌باشد . به هر حال ، اگر روشی برای محاسباتیکه فقط دارای عناصر غیر صفر می‌باشد برنامه‌نویسی گردد ، آنگاه مقدار زمان محاسبه‌ای متناسب با تعداد عناصر غیر صفر می‌باشد . این مطلب می‌تواند زمان محاسبه‌ای را به سطحی که روشهای تکراری را از روشهای مستقیم بیشتر اقتصادی نماید تبدیل کند . بنابراین ، برای یک ماتریس بطور تصادفی تنک با ساختمان غیرویژه احتمالا" یک روش تکراری به کار خواهد رفت . دو موقعیت وجود دارد که در آن این حالت نخواهد بود ، مثلا" اگر ماتریس غالب قطری نباشد آنگاه بایستی یک روش مستقیم برای جلوگیری از مشکلات همگرائی به‌کار رود ، یا اگر ماتریس دارای یک ترکیب ویژه‌ای باشد آنگاه اغلب الگوریتمهای ویژه می‌تواند برای بهره‌برداری این طرح به‌کار رود .

در حالت تنک تصادفی که در آن تعداد عناصر غیر صفر ماتریس در ابتدا کوچک بوده ، مانند 5% و یا 10% ، بعضی دگرگونیهای جالب در روش‌حذفی گاوس وجود دارد . مسئله‌ای که در روش متعارف ظاهر می‌گردد عبارتست از اینکه هرتفاضلی از سطرها ، عناصر غیر صفر بیشتری تولید می‌نماید ، و برای ماتریس بدوا" خیلی تنک امکان دارد که سریعا" پر گردد . در برنامه استراتژیهای متعددی برای تقلیل ظاهر شدن این اعضای غیرصفر درنظر گرفته شده که فقط برای محاسبه عناصر غیر صفر طرح‌ریزی می‌گردد .

یک استراتژی ساده که نتایج تجربی خوبی داده است عبارتست از به کار بردن نوع جدیدی از استراتژی محورگیری . در هر مرحله از روند ، سطر محور ، مانند سطری با کوچکترین عددی از عناصر غیرصفر که دارای عنصر رضایت‌بخش ، در محل محور ، نیز هست انتخاب می‌گردد . البته ، کوچک نبودن عنصر در محل محورگیری مهم بوده ، چون

با استراتژی خلاصه شده در اینجا عنصر محوری معمولا" مانند آنچه که قبلا" در روش گاوس اشاره شد بزرگترین نخواهد بود.

استثنای دیگر زمانی است که ماتریس دارای یک شکل ساده بوده، و اجازه می‌دهد که شکل موثرتری از روش گاوس اتخاذ گردد، بنابراین زمان محاسبه‌ای متناسب با $n^3/3$ باقی نمی‌ماند. این مطلب بامثالی به وسیلهٔ الگوریتم توماس نشان داده شده‌است که یک فرمول‌بندی از طرح حذفی گاوس بوده که در یک ماتریس سه قطری به‌کار رفته است، یعنی،

$$\begin{bmatrix} a_1 & c_1 & & & \\ b_2 & a_2 & c_2 & 0 & \\ & & & & \\ & 0 & b_{n-1} & a_{n-1} & c_{n-1} \\ & & & b_n & a_n \end{bmatrix} \tag{4.54}$$

این مطلب در قسمت(۴.۲)توضیح داده شده است.

۴.۵ ■ مقایسه روشها

در هر روش مقدار زمان محاسبه‌ای لازم مسئله مهمی است و تفصیلات شمارشهای عملی برای روشهای متعدد در جدول(۴.۱)داده شده‌اند. در بیشتر کامپیوترها زمان لازم برای ضرب و تقسیم خیلی طولانی‌تر از جمع و تفریق می‌باشد، بنابراین، شمارش عملی فقط برای ضرب و تقسیم‌ها درنظر گرفته می‌شود.

تعداد عملیات در روش‌های متعارف

حذفی گاوس تجزیه مثلثی	$\dfrac{n^3}{3}+n^2-\dfrac{n}{3}$
معکوس و ضرب ماتریس	n^3+n^2
حذف جردن	$\dfrac{n^3}{2}+n^2-\dfrac{n}{2}$
با طرف سمت راست اضافی برای تمام روشهای فوق	n^2
حل سه قطری به وسیله الگوریتم توماس	$5n-4$
با سمت راست اضافی	$3n-2$
روشهای تکراری	$r.n^2$
باطرف سمت راست اضافی	rn^2

جدول ۴.۱

n عبارتست از تعداد معادلات

r عبارتست از تعداد تکرارهای لازم برای معیارهای مشخص همگرائی .

بدیهی است در ردهٔ روشهای مستقیم روشهائی که ترکیب ویژه‌ای را در ماتریس به انجام میرساند نسبت به روشهای متعارف اقتصادی‌تر است . برای یک ماتریس عمومی که وسیله روشهای مستقیم حل می‌شود حذف گاوس یا تجزیهٔ مثلثی به کار خواهد رفت . در صورتیکه کامپیوتری با امکانات جمع‌حاصلضربها با طول ــ مضاعف وجود داشته باشد آنگاه الگوریتم تجزیهٔ مثلثی نتایج خیلی دقیق خواهد داد . زمانیکه ماتریس متقارن باشد روش تجزیه مثلثی مفید خواهد بود ، چون در اینصورت کاهش شماره عملیات و ذخیره به 50% امکان‌پذیر می‌گردد .

روشهای تکراری برای ماتریسهای تنک تصادفی به کار می‌روند که در آن سرعت همگرائی خوب می‌باشد . اگر روش برای مسئله‌ای بکندی همگرا گردد ممکن است ترجیح داده شود که روشهای مستقیم را به‌کار برد ، اگرچه ماتریس تنک باشد . روش محورگیری متناظر به سطر با کوچکترین تعداد عناصر غیر صفر دارای نتایج تجربی مشوقی است ، اگرچه خطری وجود دارد ، که خطاها ممکن است در اثر پیراستن روند محورگیری نرمال جمع شوند .

در غیاب هرکدام از سایر روندها ، روند تکراری که معمولا" به کار می‌رود عبارتست از روش گاوس ــ سایدل ، چون این‌روش معمولا" دارای سرعت همگرائی سریعتر از روش ژاکوبی می‌باشد . اگر ماتریس دارای شکل ویژه‌ای باشد ، مانند آنهائیکه در بعضی معادلات با مشتقات نسبی ظاهر می‌شوند ، آنگاه ممکن است امکان محاسبه یک عامل مناسب فوق ــ تخفیف ω وجود داشته باشد که در این صورت به منظور سرعت بخشیدن به همگرائی به‌کار خواهد رفت . برای یک ماتریس عمومی انتخاب عامل فوق ــ تخفیف مشکل بوده ، اما ممکن است درصورتیکه سرعت همگرائی آهسته باشد و چندین محاسبه شبیه هم وجود داشته باشد ارزنده گردد . یک‌سری از محاسبات با مقادیر آزمایشی ω انجام گرفته و آنگاه یک مقدار مناسب ω انتخاب می‌گردد .

درصورتیکه روش گاوس ــ سایدل برای همگرائی شکست بخورد آنگاه روش دیگری مانند ژاکوبی امکان‌پذیر می‌باشد .

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ جدول اعداد صفحهٔ ۱۱۰ طرح حذف گاوس را برای ۴ معادله با چهار مجهول نشان می‌دهد ، ضرائب ، یک ماتریس پایین مثلثی تشکیل داده که می‌تواند در محل‌هائی که عناصر بوسیله روند حذف گاوس تولید می‌گردند ذخیره شوند . در ماتریس بالا

مثلثی ، جایگذاری پسرو به‌کار رفته و بادرنظر گرفتن بالاترین سطر هر یک از مجموعه معادلات تشکیل می‌گردد . آخرین سطر اعداد ، مقادیر x_4, x_3, x_2, x_1 که بوسیله روند جایگذاری پسرو بدست می‌آیند را می‌دهد .
هر مجموعه از معادلات همانطوریکه در قسمت ۴۰۲۰۱ اشاره شد بوسیله کاهش مضرب مناسبی از معادلهٔ اولی از بقیه معادلات‌بدست می‌آید . این عمل مجموعه جدیدی از معادلات با یک بعد کمتر از مجموعه قبلی ایجاد می‌کند .

۲ــ مثالی از یک ماتریس بدوضع ، قطعه‌ای از ماتریس هیلبرت‌می‌باشد .

$$\begin{matrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{matrix}$$

نتایج معکوس این ماتریس با 6 رقم با معنی و با 3 رقم بامعنی درمثال (۷۰۶) داده شده است .

۳ــ جدول نتایج مندرج در صفحات ۱۱۱ و ۱۱۲ یک مقایسه جالب بین روشهای تکراری مختلف برای حل دستگاه معادلات خطی را می‌دهد . معادله ماتریسی که باید حل گردد عبارتست‌از $\mathbf{AX} = \mathbf{B}$ که در آن

$$\mathbf{A} = \begin{bmatrix} 1{\cdot}0 & -0{\cdot}7 & 0{\cdot}0 \\ -0{\cdot}7 & 1{\cdot}0 & -0{\cdot}7 \\ 0{\cdot}0 & -0{\cdot}7 & 1{\cdot}0 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} -4 \\ 34 \\ -44 \end{bmatrix}$$

مقدار بزرگترین قدر مطلق مقادیر ویژه ، سرعت همگرائی را بیان‌می‌کند . به‌طور ایده‌آل این مقدار باید اساسا" کمتر از واحد باشد ، اگرچه شرط تئوری به سادگی نیازمند یک مقدار کمتر از واحد هست .

مقدار این کمیت شعاع‌طیفی (the spectral radius) برای‌روش‌ژاکوبی 0·9898 برای روش گاوس ــ سایدل 0·98 ، برای روش فوق تخفیف 0·755 با می‌نیمم عامل فوق تحفیف 1·755 می‌باشد . می‌توان دید که روش ژاکوبی بعد از 500 تکرار هنوز نسبتا" دور از جواب واقعی بوده ، درصورتیکه روش فوق ــ تخفیف در 70 تکرار همگرائی رضایت‌ــ بخشی دارا می‌باشد .

مثال حل شده شماره ۱

					طرف سمت راست
Multipliers	1·400 000 000 0	0·350 000 000 0	0·980 000 000 0	0·029 224 000 0	2·639 600 000 1
0·250 000 000 0	0·350 000 000 0	0·980 000 000 0	0·029 224 000 0	0·006 633 000 0	0·669 855 000 0
0·700 000 000 0	0·980 000 000 0	0·029 224 000 0	0·006 633 000 0	0·002 887 000 0	0·189 451 000 0
0·020 874 285 7	0·029 224 000 0	0·006 633 000 0	0·002 887 000 0	0·000 935 000 0	0·056 897 000 0
	Multipliers	0·892 500 000 0	−0·215 776 000 0	−0·000 673 000 0	0·009 955 000 0
	−0·241 765 826 3	−0·215 776 000 0	−0·679 367 000 0	−0·017 569 800 0	−1·658 269 000 1
	−0·000 754 061 6	−0·000 673 000 0	−0·017 569 800 0	0·000 324 969 9	0·001 797 235 4
		Multiplier	−0·731 534 263 0	−0·017 732 508 4	−1·655 862 221 3
		0·024 240 161 1	−0·017 732 508 4	0·000 324 462 4	0·001 804 742 1
				0·000 754 301 3	0·041 943 109 2
	x_1	x_2	x_3	x_4	
	55·605 249 316 4	0·915 666 845 0	0·274 460 797 7	0·015 126 719 1	

Jacobi's method

Iteration number	x_1	x_2	x_3	Iteration number	x_1	x_2	x_3
0	0·0000	0·0000	0·0000	300	10·4830	19·0341	−29·5170
1	−4·0000	34·0000	−44·0000	301	9·3239	20·6761	−30·6761
2	19·8000	0·4000	−20·2000	302	10·4733	19·0534	−29·5267
3	−3·7200	33·7200	−43·7200	303	9·3374	20·6626	−30·6626
4	19·6040	0·7920	−20·3960	304	10·4638	19·0723	−29·5362
5	−3·4456	33·4456	−43·4456	305	9·3506	20·6494	−30·6494
6	19·4119	1·1762	−20·5881	306	10·4546	19·0909	−29·5454
7	−3·1767	33·1767	−43·1767	307	9·3636	20·6364	−30·6364
8	19·2237	1·5526	−20·7763	308	10·4455	19·1091	−29·5545
9	−2·9132	32·9132	−42·9132	309	9·3763	20·6237	−30·6237
100	13·6417	12·7166	−26·3583	400	10·1759	19·6482	−29·8241
101	4·9016	25·0984	−35·0984	401	9·7538	20·2462	−30·2462
102	13·5689	12·8623	−26·4311	402	10·1724	19·6553	−29·8276
103	5·0036	24·9964	−34·9964	403	9·7587	20·2413	−30·2413
104	13·4975	13·0050	−26·5025	404	10·1689	19·6622	−29·8311
105	5·1035	24·8965	−34·8965	405	9·7635	20·2365	−30·2365
106	13·4275	13·1449	−26·5725	406	10·1655	19·6689	−29·8345
107	5·2014	24·7986	−34·7986	407	9·7682	20·2318	−30·2318
108	13·3590	13·2820	−26·6410	408	10·1622	19·6755	−29·8378
109	5·2974	24·7026	−34·7026	409	9·7729	20·2271	−30·2271
200	11·3262	17·3476	−28·6738	490	10·0709	19·8583	−29·9291
201	8·1433	21·8567	−31·8567	491	9·9008	20·0992	−30·0992
202	11·2997	17·4007	−28·7003	492	10·0694	19·8611	−29·9306
203	8·1805	21·8195	−31·8195	493	9·9028	20·0972	−30·0972
204	11·2737	17·4526	−28·7263	494	10·0681	19·8639	−29·9319
205	8·2169	21·7831	−31·7831	495	9·9047	20·0953	−30·0953
206	11·2482	17·5036	−28·7518	496	10·0667	19·8666	−29·9333
207	8·2525	21·7475	−31·7475	497	9·9066	20·0934	−30·0934
208	11·2232	17·5535	−28·7768	498	10·0654	19·8693	−29·9346
209	8·2875	21·7125	−31·7125	499	9·9085	20·0915	−30·0915

Gauss–Seidel method

Iteration number	x_1	x_2	x_3	Iteration number	x_1	x_2	x_3
0	0·0000	0·0000	0·0000	200	10·1436	20·2010	−29·8593
1	−4·0000	31·2000	−22·1600	201	10·1407	20·1970	−29·8621
2	17·8400	30·9760	−22·3168	202	10·1379	20·1930	−29·8649
3	17·6832	30·7565	−22·4705	203	10·1351	20·1892	−29·8676
4	17·5295	30·5414	−22·6211	204	10·1324	20·1854	−29·8702
5	17·3789	30·3305	−22·7686	205	10·1298	20·1817	−29·8728
6	17·2314	30·1239	−22·9133	206	10·1272	20·1781	−29·8754
7	17·0867	29·9214	−23·0550	207	10·1246	20·1745	−29·8779
8	16·9450	29·7230	−23·1939	208	10·1221	20·1710	−29·8803
9	16·8061	29·5285	−23·3300	209	10·1197	20·1676	−29·8827
100	11·0826	21·5157	−28·9390	290	10·0233	20·0326	−29·9772
101	11·0610	21·4853	−28·9603	291	10·0223	20·0320	−29·9776
102	11·0397	21·4556	−28·9811	292	10·0224	20·0313	−29·9781
103	11·0189	21·4265	−29·0014	293	10·0219	20·0307	−29·9785
104	10·9986	21·3980	−29·0214	294	10·0215	20·0301	−29·9789
105	10·9786	21·3700	−29·0410	295	10·0211	20·0295	−29·9794
106	10·9590	21·3426	−29·0602	296	10·0206	20·0289	−29·9798
107	10·9398	21·3158	−29·0790	297	10·0202	20·0283	−29·9802
108	10·9210	21·2895	−29·0974	298	10·0198	20·0278	−29·9806
109	10·9026	21·2637	−29·1154	299	10·0194	20·0272	−29·9810

Successive over-relaxation
$\omega = 1{\cdot}75$

Iteration number	x_1	x_2	x_3	Iteration number	x_1	x_2	x_3
0	0·0000	0·0000	0·0000	35	10·0969	20·1232	−29·9238
1	−7·0000	50·9250	−14·6169	36	10·0782	20·0967	−29·9387
2	60·6331	77·6762	29·1159	37	10·0598	20·0759	−29·9530
3	42·6784	89·1910	10·4220	38	10·0481	20·0595	−29·9623
4	70·2502	91·4302	27·1854	39	10·0369	20·0467	−29·9711
5	52·3143	88·3146	10·7963	40	10·0295	20·0366	−29·9769
6	61·9496	82·3778	15·8156	41	10·0227	20·0287	−29·9822
7	47·4506	75·2177	3·2800	42	10·0181	20·0225	−29·9858
8	49·5538	67·8081	3·6049	43	10·0139	20·0176	−29·9891
9	38·8996	60·7120	−5·3315	44	10·0111	20·0138	−29·9913
10	38·1974	54·2268	−6·5736	45	10·0086	20·0108	−29·9933
11	30·7797	48·4824	−12·6788	46	10·0068	20·0085	−29·9947
12	29·3062	43·5067	−14·1952	47	10·0052	20·0066	−29·9959
13	24·3161	39·2681	−18·2502	48	10·0042	20·0052	−29·9967
14	22·8664	35·7037	−19·5753	49	10·0032	20·0041	−29·9975
15	19·5872	32·7369	−22·2159	50	10·0026	20·0032	−29·9980
16	18·4122	30·2879	−23·2354	51	10·0020	20·0025	−29·9985
17	16·2935	28·2802	−24·9302	52	10·0016	20·0019	−29·9988
18	15·4231	26·6437	−25·6638	53	10·0012	20·0015	−29·9991
19	14·0712	25·3163	−26·7397	54	10·0010	20·0012	−29·9993
20	13·4590	24·2440	−27·2464	55	10·0007	20·0009	−29·9994
21	12·6046	23·3808	−27·9237	56	10·0006	20·0007	−29·9995
22	12·1881	22·6883	−28·2641	57	10·0005	20·0006	−29·9996
23	11·6521	22·1340	−28·6877	58	10·0004	20·0004	−29·9997
24	11·3751	21·6916	−28·9120	59	10·0003	20·0003	−29·9998
25	11·0408	21·3391	−29·1756	60	10·0002	20·0003	−29·9998
26	10·8598	21·0588	−29·3213	61	10·0002	20·0002	−29·9999
27	10·6522	20·8363	−29·4846	62	10·0001	20·0002	−29·9999
28	10·5353	20·6599	−29·5782	63	10·0001	20·0001	−29·9999
29	10·4069	20·5202	−29·6791	64	10·0001	20·0001	−29·9999
30	10·3321	20·4098	−29·7387	65	10·0001	20·0001	−30·0000
31	10·2529	20·3226	−29·8008	66	10·0000	20·0001	−30·0000
32	10·2055	20·2538	−29·8385	67	10·0000	20·0000	−30·0000
33	10·1568	20·1996	−29·8767	68	10·0000	20·0000	−30·0000
34	10·1269	20·1568	−29·9004	69	10·0000	20·0000	−30·0000

۴ــ ماتریس زیر را به‌صورت حاصلضرب $\mathbf{LU}$ یک ماتریس مثلثی پایین و بالا تجزیه کنید.

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & -1 & 0 \\ 1 & -2 & 3 & 1 \\ 2 & -3 & 5 & 1 \\ -1 & 2 & -1 & 6 \end{bmatrix}$$

نشان دهید این تجزیه چگونه قادر به حل معادله $\mathbf{AX} = \mathbf{B}$ خواهد شد که در آن

$$\mathbf{B}^T = [0, -5, -5, 3]$$

چون $\mathbf{LU} = \mathbf{A}$ داریم.

$$\begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}\begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 4 & 2 & -1 & 0 \\ 1 & -2 & 3 & 1 \\ 2 & -3 & 5 & 1 \\ -1 & 2 & -1 & 6 \end{bmatrix}$$

$$\begin{array}{ll} l_{11}=4 & l_{11}=4 \\ l_{11}u_{12}=2 & u_{12}=0\cdot5 \\ l_{11}u_{13}=-1 & u_{13}=-0\cdot25 \\ l_{11}u_{14}=0 & u_{14}=0 \\ & \\ l_{21}=1 & l_{21}=1 \\ l_{21}u_{12}+l_{22}=-2 & l_{22}=-2\cdot5 \\ l_{21}u_{13}+l_{22}u_{23}=3 & u_{23}=-1\cdot3 \\ l_{21}u_{14}=l_{22}u_{24}=1 & u_{24}=-0\cdot4 \\ & \\ l_{31}=2 & l_{31}=2 \\ l_{31}u_{12}+l_{32}=-3 & l_{32}=-4 \\ l_{31}u_{13}+l_{32}u_{23}+l_{33}=5 & l_{33}=0\cdot3 \\ l_{31}u_{14}+l_{32}u_{24}+l_{33}u_{34}=0 & u_{34}=-2\cdot0 \\ & \\ l_{41}=-1 & l_{41}=-1 \\ l_{41}u_{12}+l_{42}=2 & l_{42}=2\cdot5 \\ l_{41}u_{13}+l_{42}u_{23}+l_{43}=-1 & l_{43}=2\cdot0 \\ l_{41}u_{14}+l_{42}u_{24}+l_{43}u_{34}+l_{44}=6 & l_{44}=11\cdot0 \end{array}$$

لذا ، معادله زیر را حل می‌نمائیم .

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 1 & -2\cdot5 & 0 & 0 \\ 2 & -4 & 0\cdot3 & 0 \\ -1 & 2\cdot5 & 2\cdot0 & 11\cdot0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}\begin{bmatrix} 0 \\ -5 \\ -5 \\ 3 \end{bmatrix}$$

$y_1=0, \quad y_2=2, \quad y_3=10, \quad y_4=-2$

آنگاه شکل بالامثلثی حل می‌گردد .

$$\begin{bmatrix} 1 & 0\cdot5 & -0\cdot25 & 0 \\ 0 & 1 & -1\cdot3 & -0\cdot4 \\ 0 & 0 & 1 & -2\cdot0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \\ 10 \\ -2 \end{bmatrix}$$

$x_4=-2, \quad x_3=6, \quad x_2=9, \quad x_1=-3$

با قرار دادن این نتایج در معادله اصلی می‌توان دید که جوابهای واقعی بدست آمده‌اند .

۵ـ مطلوبست حل مجموعه معادلات سه قطری زیر :

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 3 & -2 & -3 & 0 & 0 \\ 0 & 4 & -4 & 2 & 0 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}=\begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

الگوریتم توماس که بوسیله معادلات (4.34), (4.35) و (4.36) داده شـده به کـار می‌رود .

$$\alpha_1 = 2 \qquad \gamma_1 = \tfrac{1}{2} = 0{\cdot}5$$
$$\alpha_2 = -2 - 3 \times 0{\cdot}5 = -3{\cdot}5 \qquad \gamma_2 = -3/(-3{\cdot}5) = 0{\cdot}857\,142$$
$$\alpha_3 = -4 - 4 \times 0{\cdot}857\,143 = -7{\cdot}428\,752 \qquad \gamma_3 = 2/(-7{\cdot}428\,572) = -0{\cdot}269\,231$$
$$\alpha_4 = 3 + 1 \times (-0{\cdot}269\,231) = 2{\cdot}730\,769 \qquad \gamma_4 = 2/2{\cdot}730\,769 = -0{\cdot}732\,394$$
$$\alpha_5 = 7{\cdot}661\,970$$

$$u_1 = \tfrac{3}{2} = 1{\cdot}5$$
$$u_2 = (2 - 3 \times 1{\cdot}5)/(-3{\cdot}5) = 0{\cdot}714\,286$$
$$u_3 = (1 - 4 \times 0{\cdot}714\,826)/(-7{\cdot}428\,572) = 0{\cdot}25$$
$$u_4 = (0 + 1 \times 0{\cdot}25)/2{\cdot}730\,769 = 0{\cdot}091\,549$$
$$u_5 = (-1 - 5 \times 0{\cdot}091\,549)/7{\cdot}661\,970 = -0{\cdot}190\,257$$

جواب با جایگذاری پسرو عبارتست‌از :

$$x_5 = -0{\cdot}190\,257$$
$$x_4 = 0{\cdot}091\,549 + 0{\cdot}732\,394 \times (-0{\cdot}190\,257) = -0{\cdot}047\,794$$
$$x_3 = 0{\cdot}25 + 0{\cdot}269\,231 \times (-0{\cdot}047\,794) = 0{\cdot}237\,132$$
$$x_2 = 0{\cdot}714\,286 - 0{\cdot}857\,143 \times 0{\cdot}237\,132 = 0{\cdot}511\,030$$
$$x_1 = 1{\cdot}5 - 0{\cdot}5 \times 0{\cdot}511\,030 = 1{\cdot}244\,485$$

این جوابها را می‌توان در معادلات اصلی قرار داده صحت محاسبات را کنترل کرد .

※ برنامه‌های کامپیوتری

الگوریتم‌ های روش‌ های حذفی

برای حل دستگاه $AX = B$ ماتریس W از رتبه $n\times(n+1)$ که شامل عناصر ماتریس A و بردار B می‌باشد را در نظر می‌گیریم .

$$W = (w_{ij}): \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} & b_n \end{bmatrix}$$

و P بردار محورگیری می‌باشد .

$$\mathbf{p} = (p_i)$$

$$p_i = i,\ i = 1, \ldots, n.$$

* الگوریتم (۴.۱) جایگذاری پَسرو (SUBST)

با داشتن n ستون آخری ماتریس W و بردار n بعدی P الگوریتم جایگذاری پسرو به صورت زیر خواهد شد :

برای $k = 1, \ldots, n$

قرار می‌دهیم

$$\tilde{b}_k = b_{p_k} - \sum_{j=1}^{k-1} w_{p_k j}\tilde{b}_j$$

برای $k = n, n-1, \ldots, 1$

قرار می‌دهیم

$$x_k = \frac{\tilde{b}_k - \sum_{j=k+1}^{n} w_{p_k j}x_j}{w_{p_k k}}$$

زیر برنامه الگوریتم فوق در برنامه کامپیوتری (۴.۱) آورده شده است .

ALGORITHM 4.1

```
      SUBROUTINE SUBST(W,B,X,IPIVOT,N)
      DIMENSION W(N,N),B(N),X(N),IPIVOT(N)
      IF (N .GT. 1)                             GO TO 10
      X(1) = B(1)/W(1,1)
                                                RETURN
   10 IP = IPIVOT(1)
      X(1) = B(IP)
      DO 15 K = 2,N
      IP = IPIVOT(K)
      KM1 = K - 1
      SUM = 0.
      DO 14 J = 1,KM1
   14 SUM = W(IP,J)*X(J) + SUM
   15 X(K) = B(IP) - SUM
C
      X(N) = X(N)/W(IP,N)
      K = N

      DO 20 NP1MK = 2,N
      KP1 = K
      K = K - 1
      IP = IPIVOT(K)
      SUM = 0.
      DO 19 J = KP1,N
   19 SUM = W(IP,J)*X(J) + SUM
   20 X(K) = (X(K) - SUM)/W(IP,K)
                                                RETURN
      END
```

* الگوریتم (۴.۲) حذف گاوس با محورگیری جزئی مقیاس شده (FACTOR)

این الگوریتم برای ماتریس A از مرتبه N×N یک ماتریس W از مرتبه N×N و محورگیری

را به صورت بردار IPIVOT با بعد N ذخیره کرده و آماده برای استفاده از زیربرنامه SUBST می‌گردد . همچنین باید یک بردار اضافی N بعدی مانند D برای ذخیره کردن اندازه سطرهای ماتریس A به کار برد . زیربرنامه الگوریتم فوق تحت SUBST به‌صورت برنامه کامپیوتری (۴.۲) آورده شده است .

ALGORITHM 4.2

```
      SUBROUTINE FACTOR(A,W,IPIVOT,D,N,IFLAG)
      DIMENSION A(N,N),W(N,N),IPIVOT(N),D(N)
      IFLAG = 1
C     INITIALIZE W, IPIVOT, D
      DO 10 I = 1,N
      IPIVOT(I) = I
      ROWMAX = 0.
      DO 9 J = 1,N
    9 ROWMAX = AMAX1(ROWMAX,ABS(W(I,J)))
      IF (ROWMAX .EQ. 0.)                  GO TO 999
   10 D(I) = ROWMAX
C     GAUSS ELIMINATION WITH SCALED PARTIAL PIVOTING.
      NM1 = N - 1
      IF (NM1 .EQ. 0)                      RETURN
      DO 20 K = 1,NM1
      J = K
      KP1 = K + 1
      IP = IPIVOT(K)
      COLMAX = ABS(W(IP,K))/D(IP)
      DO 11 I = KP1,N
      IP = IPIVOT(I)
      AWIKOV = ABS(W(IP,K))/D(IP)
      IF (AWIKOV .LE. COLMAX)              GO TO 11
      COLMAX = AWIKOV
      J = I
   11 CONTINUE
      IF (COLMAX .EQ. 0.)                  GO TO 999
C
      IPK = IPIVOT(J)

      IPIVOT(J) = IPIVOT(K)
      IPIVOT(K) = IPK
      DO 20 I = KP1,N
      IP = IPIVOT(I)
      W(IP,K) = W(IP,K)/W(IPK,K)
      RATIO = -W(IP,K)
      DO 20 J = KP1,N
   20 W(IP,J) = RATIO*W(IPK,J) + W(IP,J)
      IF (W(IP,N) .EQ. 0.)                 GO TO 999
                                           RETURN
  999 IFLAG = 2
                                           RETURN
      END
```

۶ ــ با استفاده از برنامه کامپیوتری و به کمک حذف گاوسی و عمل محورگیری جزئی مقیاس‌شده معکوس ماتریس زیر را پیدا کنید .

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & -3 \\ -2 & 3 & -1 \end{bmatrix}$$

معکوس A به کمک زیربرنامه‌های FACTOR و SUBST و همچنین نمونه داده‌ها و ماتریس به کمک برنامه کامپیوتری ۴۰۳ بدست می‌آید .

مرتبه N ماتریس A باید جزئی از داده‌ها باشد بنابراین امکان اینکه در حین اجرای برنامه بعد درست ماتریس A مشخص گردد وجود ندارد . به عبارت دیگر زیربرنامه‌های FACTOR و SUBST ماتریس‌های A یا W از مرتبه N×N را می‌گیرد . بنابراین برای سهولت در برنامه فرترن زیرماتریس A به صورت یک‌بعدی ذخیره می‌گردد . این به این منظور است که عنصر (I,J) از دوبعدی (N,M) معادل $(I+(J-1)*N)$ از یک‌بعدی می‌باشد .

همچنین برای ذخیره کردن ستون Jام ماتریس A^{-1} از بردار یک‌بعدی AINV استفاده می‌گردد . زیربرنامه SUBST عنصر $(1+(J-1)*N)$ از AINV را می‌دهد که مساوی اولین عنصر بردار X در SUBST می‌شود . جواب دستگاه به صورت $AX = I_j$ ذخیره می‌گردد .

ALGORITHM 4.3

```
      IBEG = 1
      DO 30 J = 1,N
      B(J) = 1.
      CALL SUBST(A,B,AINV(IBEG),IPIVOT,N)
      B(J) = 0.
   30 IBEG = IBEG + N
      WRITE (6,630)
  630 FORMAT(24H1THE COMPUTED INVERSE IS //)
      DO 31 I = 1,N
   31 WRITE  (6,631) I,(AINV(J),J = I,NSQ,N)
  631 FORMAT(5H0ROW I2,8E15.7/(7X8E15.7))
                                            GO TO 1
      END
```

SAMPLE INPUT

```
 3
2.     3.    -1.
4.     4.    -3.
-2.    3.    -1.
```

RESULTING OUTPUT

THE COMPUTED INVERSE IS

```
ROW  1  0.2500000E 00    0.0              -0.2499999E  00
ROW  2  0.5000000E 00   -0.1999998E 00     0.9999996E-01
ROW  3  0.1000000E 01   -0.6000000E 00    -0.2000000E  00
```

FORTRAN PROGRAM FOR CALCULATING THE INVERSE OF A GIVEN MATRIX

```
      DIMENSION A(900),AINV(900),B(30),IPIVOT(30)
    1 READ (5,501) N
  501 FORMAT(I2)
C     READ IN MATRIX ROW BY ROW.
      NSQ = N*N
      DO 10 I = 1,N
   10 READ  (5,510) (A(J),J = I,NSQ,N)
  510 FORMAT(5E15.7)
C
      CALL FACTOR(A,A,IPIVOT,B,N,IFLAG)
                                     GO TO (20,11),IFLAG
   11 WRITE (6,611)
  611 FORMAT(19H1MATRIX IS SINGULAR)
                                     GO TO 1
C
   20 DO 21 I = 1,N
   21 B(I) = 0.
```

۷ ــ با استفاده از برنامه کامپیوتری به طریق حذفی دستگاه سه‌قطری زیر را حل کنید .

$$\begin{aligned} 2x_1 \; - \; x_2 \qquad\qquad &= 1 \\ -x_{i-1} + 2x_i \quad - x_{i+1} &= 0 \qquad i = 2, \ldots, n-1 \\ - \; x_{n-1} + 2x_n \quad &= 0 \end{aligned}$$

برای $n = 10.$

ابتدا الگوریتم زیر را در نظر می‌گیریم آنگاه زیربرنامه تحت TRID را برای حل مسئلهٔ فوق به کار می‌بریم . برنامهٔ کامپیوتری با جواب دستگاه به صورت الگوریتم ۴.۴ آورده شده است .

الگوریتم (۴۰۴) حل دستگاه سه‌قطری (TRID)

برای دستگاه سه‌قطری داده‌شده به صورت :

$$a_i x_{i-1} + d_i x_i + c_i x_{i+1} = b_i \qquad i = 1, \ldots, n \text{ (with } a_1 = c_n = 0)$$

داریم .

برای $k = 2, \ldots, n$

اگر $d_{k-1} = 0$ عمل حذف با شکست مواجه شده باید برنامه را متوقف کرد .

در غیر اینصورت قرار می‌دهیم

$$m = \frac{a_k}{d_{k-1}}$$

و ادامه عملیات

$$d_k = d_k - m * c_{k-1}$$
$$b_k = b_k - m * b_{k-1}$$

اگر $d_n = 0$ عمل حذف با شکست مواجه شده باید برنامه را متوقف کرد .
در غیر اینصورت

$$x_n = \frac{b'_n}{d'_n}$$

و ادامه عملیات

$$k = n - 1, \ldots, 1,$$

$$x_k = \frac{b'_k - c_k x_{k+1}}{d'_k}$$

برنامه فرترن زیر به صورت TRID(SUB, DIAG, SUP, B, N)
که در آن SUB ، DIAG ، SUP و B بردارهای N بعدی هستند که دستگاه فوق را به صورت زیر می‌توان نشان داد .

$$\text{SUB}(i)x_{i-1} + \text{DIAG}(i)x_i + \text{SUP}(i)x_{i+1} = \text{B}(i) \qquad i = 1, \ldots, \text{N}$$

که SUB(1) و SUP(N) صرفنظر شده‌اند .
زیربرنامه محتویات بردار DIAG را پس از تغییر دادن به صورت بردار B جواب دستگاه برمی‌گرداند .
جواب این برنامه از کامپیوتر IBM 360 بدست آمده است .

ALGORITHM 4.4

FORTRAN PROGRAM FOR EXAMPLE 7

```
      DIMENSION A(10),D(10),C(10),B(10)
      N = 10
      DO 10 I = 1,N
      A(I) = -1.
      D(I) = 2.
      C(I) = -1.
   10 B(I) = 0.
      B(1) = 1.
      CALL TRID(A,D,C,B,N)
      WRITE (6,610) (I,B(I),I = 1,N)
  610 FORMAT(16H1THE SOLUTION IS /(I5,E15.7))
      STOP
      END
      SUBROUTINE TRID(SUB,DIAG,SUP,B,N)
      DIMENSION SUB(30),DIAG(30),SUP(30),B(30)
      IF (N .GT. 1)                      GO TO 10
      B(1) = B(1)/DIAG(1)
                                         RETURN
   10 DO 11 K = 2,N
      RATIO = -SUB(K)/DIAG(K - 1)
```

```
      DIAG(K) = DIAG(K) + RATIO*SUP(K - 1)
   11 B(K) = B(K) + RATIO*B(K - 1)
      B(N) = B(N)/DIAG(N)
      K = N
      DO 12 NP1MK = 2,N
      K = K - 1
   12 B(K) = (B(K) - SUP(K)*B(K + 1))/DIAG(K)
                                         RETURN
      END
```

OUTPUT

```
THE SOLUTION IS
  1  0.9090915E  00
  2  0.8181832E  00
  3  0.7272751E  00
  4  0.6363666E  00
  5  0.5454577E  00
  6  0.4545485E  00
  7  0.3636391E  00
  8  0.2727295E  00
  9  0.1818197E  00
 10  0.9090990E-01
```

۸ ــ با استفاده از برنامه کامپیوتری به روش گاوس ــ سایدل دستگاه خطی زیر را برای $n = 20$ با معیار دقت داده‌شده و مقدار اولیه $x^{(0)} = 0$ بدست آورید.
معیار دقت 10^{-6}

$$\frac{\|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\|}{\|\mathbf{x}^{(m)}\|} < 10^{-6}$$

$$2x_1 - x_2 = 1$$
$$-x_{i-1} + 2x_i - x_{i+1} = 0 \qquad i = 2, \dots, n-1$$
$$-x_{n-1} + 2x_n = 1$$

برای $n = 10$

ماتریس ضرایب دستگاه فوق تنک و سه‌قطری می‌باشد. برنامه زیر تحت نام GSTRI برای حل دستگاه فوق به صورت برنامه کامپیوتری ۴.۵ و نتایج در جدول ۴.۵ آورده شده است.

ALGORITHM 4.5

FORTRAN PROGRAM FOR EXAMPLE 8

```
      DIMENSION AL(20),AD(20),AR(20),B(20),X(20)
      DO 10 J = 1,20
      AL(J) = -1.
      AD(J) = 2.
```

```
      AR(J) = -1.
      B(J) = 0.
   10 X(J) = 0.
      B(1) = 1.
      B(20) = 1.
C
      CALL GSTRI(20,AL,AD,AR,B,X,1.E - 6)
      PRINT 600, (J,X(J),J = 1,20)
  600 FORMAT (16H1THE SOLUTION IS/(I5,E20.8))
      STOP
      END
C
      SUBROUTINE GSTRI(N,AL,AD,AR,B,X,EPS)
C
C     TO SOLVE A TRIDIAGONAL LINEAR SYSTEM OF THE FORM
C       AD(1)*X(1) + AR(1)*X(2)                     = B(1)
C      AL(I)*X(I - 1) + AD(I)*X(I) + AR(I)*X(I + 1) = B(I), I = 2, . . . ,
C                 N - 1, AL(N)*X(N - 1) + AD(N)*X(N) = B(N)
C     BY GAUSS-SEIDEL ITERATION. STARTING WITH THE GIVEN X,
C     THE ITERATION IS CARRIED OUT UNTIL EITHER MORE THAN
C     1000 STEPS HAVE BEEN TAKEN OR ELSE THE RELATIVE CHANGE
C     IN X DURING AN ITERATION STEP IS LESS THAN THE GIVEN
C     EPS.
C
      DIMENSION AL(1),AD(1),AR(1),B(1),X(1)
      DO 10 J = 1,N
      IF (AD(J) .EQ. 0.)                         GO TO 999
      AL(J) = AL(J)/AD(J)
      AR(J) = AR(J)/AD(J)
   10 B(J) = B(J)/AD(J)
C
      DO 30 ITER = 1,1000
      I = 1
      DIFMAX = 0.
      XMAX = 0.
C
      XNEW = B(1) - AR(1)*X(2)
                                                 GO TO 13
   11 XNEW = B(N) - AL(N)*X(N - 1)
                                                 GO TO 13
   12 XNEW = B(I) - AL(I)*X(I - 1) - AR(I)*X(I + 1)
C
   13 DIFF = XNEW - X(I)
      X(I) = XNEW
      XMAX = AMAX1(XMAX,ABS(XNEW))
      DIFMAX = AMAX1(DIFMAX,ABS(DIFF))
C
      I = I + 1
      IF (I - N)                                 12,11,19
C
   19 TEST = DIFMAX/XMAX
      IF (TEST .LT. EPS)                         RETURN
   30 CONTINUE
  999 PRINT 600
  600 FORMAT (23H ITERATION UNSUCCESSFUL)
                                                 RETURN
      END
```

جدول ۴.۵

m	$x_1^{(m)}$	$x_{10}^{(m)}$	$\dfrac{\lVert \mathbf{x}^{(m)} - \mathbf{x}^{(m-1)} \rVert_\infty}{\lVert \mathbf{x}^{(m)} \rVert_\infty}$	$r_{\min}^{(m)}$	$r_{\max}^{(m)}$
30	0.89844144	0.34722738	1.63 (−2)	0.963 · · ·	0.988 · · ·
100	0.97746982	0.86366559	3.15 (−3)	0.977969 · · ·	0.977976 · · ·
140	0.99082111	0.94448888	1.27 (−2)	0.97778708 · · ·	0.97779869 · · ·
···	....	....	....	....	....
200	0.99761522	0.98557816	3.28 (−4)	0.97778641 · · ·	0.97778660 · · ·
···	....	....	....	....	....
300	0.99974774	0.99847447	3.47 (−5)	0.97778640 · · ·	0.97778640 · · ·
400	0.99997332	0.99983863	3.67 (−6)	0.97778640 · · ·	0.97778640 · · ·
455	0.99999224	0.99995309	1.07 (−6)	0.97778640 · · ·	0.97778640 · · ·
456	0.99999275	0.99995615	9.60 (−7)	0.97778640 · · ·	0.97778640 · · ·

■ مسائل

۱ــ مطلوبست محاسبه A^{-1} به وسیله تجزیه مثلثی

$$A=\begin{bmatrix} 4 & 6 & 8 \\ 6 & 10 & 17 \\ 8 & 17 & 25 \end{bmatrix}$$

۲ــ مطلوبست حل مجموعه معادلات $AX=B$ که در آن

$$A=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 12 & 2 \\ 5 & 10 & 1 \end{pmatrix} \qquad B=\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

۳ــ با استفاده از خط کش مهندسی مطلوبست حل دستگاه زیر بکمک روش گاوســ سایدل .

$$\begin{aligned} 9x_1-2x_2&=4{\cdot}1 \\ 18x_2-2x_3&=1{\cdot}3 \\ 2{\cdot}1x_1-15x_3&=3{\cdot}2 \end{aligned}$$

۴ــ ماتریس A را به‌صورت $L.L^T$ در آورید .

$$A=\begin{bmatrix} 9 & -18 & 3 \\ -18 & 61 & 34 \\ 3 & 34 & 81 \end{bmatrix}$$

۵ــ مطلوبست تجزیه ماتریس زیر به ماتریسهای بالا مثلثی و پایین مثلثی .

$$\begin{bmatrix} 1 & 7 & 5 \\ 4 & 29 & 24 \\ 10 & 64 & 29 \end{bmatrix}$$

۶ــ مطلوبست حل دستگاه معادلات زیر با به‌کار بردن روش گاوســ سایدل با مقادیر اولیه تقریب

$x_3=2{\cdot}0$ و $x_2=0$ ، $x_1=2{\cdot}0$

$$\begin{aligned} 3x_1+2x_2-\tfrac{1}{2}x_3&=6 \\ x_1-4x_2+x_3&=4 \\ -x_1\qquad +2x_3&=2 \end{aligned}$$

همچنین ، دستگاه را با روش ژاکوبی حل کرده و سرعت همگرائی دو روش را مقایسه کنید .

۷ ــ دستگاه زیر را به روش حذفی گاوس و با به کار بردن محورگیری جزئی حل کنید و کلیه محاسبات را تا ۲ رقم اعشار گرد کنید و دقت جوابها را پیدا کنید .

$$\begin{aligned} 0.20x_1+0.32x_2+0.12x_3+0.30x_4&=0.94 \\ 0.10x_1+0.15x_2+0.24x_3+0.32x_4&=0.81 \\ 0.20x_1+0.24x_2+0.46x_3+0.36x_4&=1.26 \\ 0.60x_1+0.40x_2+0.32x_3+0.20x_4&=1.52 \end{aligned}$$

۸ ــ با استفاده از زیربرنامه TRID دستگاه زیر را حل کنید .

$$\begin{aligned} -2(1+h^2)x_1 + x_2 &= 1 \\ x_{i-1} - 2(1+h^2)x_i + x_{i+1} &= 0 \qquad i = 2, 3, \ldots, n-1 \\ x_{n-1} - 2(1+h^2)x_n &= 1 \end{aligned}$$

برای $n = 30 \quad h = 0.1$

۹ ــ دستگاه زیر داده شده است .

$$\begin{aligned} 0.1410\cdot 10^{-2}x_1 + 0.4004\cdot 10^{-1}x_2 &= 0.1142\cdot 10^{-1} \\ 0.2000\cdot 10^{0}x_1 + 0.4912\cdot 10^{1}x_2 &= 0.1428\cdot 10^{1} \end{aligned}$$

در صورتیکه کلیه اعمال محاسباتی را با حساب مانتیس ۴ رقمی انجام دهیم دستگاه فوق را :

الف : بدون محورگیری جزئی

ب : با محورگیری جزئی

حل کرده و آنگاه نتایج را با مقدار واقعی دستگاه $x_1 = 1.000$, $x_2 = 0.2500$ مقایسه کنید .

۱۰ ــ دستگاه زیر داده شده است :

$$\begin{aligned} 0.0003x_1 + 1.566x_2 &= 1.569 \\ 0.3454x_1 - 2.436x_2 &= 1.018 \end{aligned}$$

در صورتیکه کلیه اعمال محاسباتی را با حساب مانتیس ۴ رقمی انجام دهیم دستگاه فوق را در حالات زیر حل کنید .

الف : با محورگیری جزئی

ب : بدون محورگیری جزئی

ج : معادله اول دستگاه را در 10^m ضرب سپس حل کنید .

د : نتایج الف تا ب را توجیه نمائید .

(m عددی است صحیح)

۱۱ ــ دستگاه ذیل را با محورگیری جزئی حل کرده و سپس در صورت نیاز یک بار عمل تصحیح را به روش حذفی گاوس انجام دهید . کلیه اعمال محاسباتی را تا ۴ رقم اعشار انجام دهید .

$$\begin{aligned} 4.01x_1 + 1.23x_2 + 1.43x_3 - .73x_4 &= 5.94 \\ 1.23x_1 + 7.41x_2 + 2.41x_3 + 3.02x_4 &= 14.07 \\ 1.43x_1 + 2.41x_2 + 5.79x_3 - 1.11x_4 &= 8.52 \\ -.73x_1 + 3.02x_2 - 1.11x_3 + 6.41x_4 &= 7.59 \end{aligned}$$

۱۲ ــ با استفاده از برنامه کامپیوتری دستگاه زیر را به روش گاوس حل کنید .

$$
\begin{aligned}
x_1 + 2x_2 + x_3 - x_4 - x_5 + 2x_6 + x_7 - x_8 &= 2\\
2x_1 + 4x_2 + 3x_3 - x_4 - x_5 + 5x_6 + 3x_7 - x_8 &= 1\\
-x_1 - x_2 + 2x_4 + 2x_5 - x_6 + 2x_8 &= -3\\
x_2 + 2x_3 + 3x_4 - 4x_5 - 3x_6 - 2x_7 - x_8 &= 0\\
x_1 + x_2 + 3x_3 - 6x_4 + 2x_5 + x_6 + 5x_7 + 2x_8 &= 3\\
-2x_1 - 2x_2 + 2x_3 + 5x_5 - 2x_6 - 3x_7 &= 0\\
5x_1 + 7x_2 + 5x_3 - 3x_4 + x_5 + 4x_6 + 11x_7 + 2x_8 &= 12\\
4x_1 + 6x_2 - 4x_4 + 10x_6 + 5x_7 - 4x_8 &= 11
\end{aligned}
$$

۱۳ ــ دستگاه زیر را با استفاده از برنامه کامپیوتری از روش گاوس ــ سایدل حل کنید .

$$
\begin{aligned}
x + \cos y &= 0.90710678\\
x^2 + \tan y - e^z &= 0.67212056\\
y + 5z &= -4.21460184
\end{aligned}
$$

۱۴ ــ دستگاه زیر داده شده است .

$$
\begin{aligned}
1.012x_1 - 2.132x_2 + 3.104x_3 &= 1.984\\
-2.132x_1 + 4.096x_2 - 7.013x_3 &= -5.049\\
3.104x_1 - 7.013x_2 + .014x_3 &= -3.895
\end{aligned}
$$

این دستگاه را ، به روش گاوس، در حالات زیر حل کنید .

الف : بدون محورگیری

ب : با محورگیری جزئی

ج : محورگیری کامل

د : جوابها را با جواب واقعی ($\frac{1}{4}$ و $\frac{1}{3}$ و $\frac{1}{2}$) مقایسه و توجیه نمائید .

کلیه اعمال محاسباتی را بعد از هر عمل تا ۴ رقم با معنی انجام دهید .

۱۵ ــ ثابت کنید ، برای دستگاه:

$$
\begin{aligned}
2x_1 - x_2 + x_3 &= 1\\
x_1 + x_2 + x_3 &= 6\\
x_1 + x_2 - 2x_3 &= 0
\end{aligned}
$$

روش گاوس ــ سایدل همگرا بوده و ژاکوبی واگرا .

۱۶ ــ دستگاه معادلات خطی زیر را به روش حذفی گاوس حل کنید .
(با حساب مانتیس ۳ رقمی)

a)
$$
\begin{aligned}
x_1 + \tfrac{1}{2}x_2 + \tfrac{1}{3}x_3 &= \tfrac{11}{6}\\
5x_1 + \tfrac{10}{3}x_2 + \tfrac{5}{2}x_3 &= \tfrac{65}{6}\\
\tfrac{100}{3}x_1 + 25x_2 + 20x_3 &= \tfrac{235}{3}
\end{aligned}
$$

(با حساب مانتیس ۴ رقمی)

b) $1.003x_1 + 58.09x_2 = 68.12$
$5.550x_1 + 321.8x_2 = 377.8$

(با حساب مانتیس ۲ رقمی)

c) $3.9x_1 + 1.6x_2 = 5.5$
$6.8x_1 + 2.9x_2 = 9.7$

(با دقت ساده)

d) $\frac{1}{4}x_1 + \frac{1}{5}x_2 + \frac{1}{6}x_3 = 9$
$\frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = 8$
$\frac{1}{2}x_1 + x_2 + 2x_3 = 8$

(با حساب مانتیس ۲ رقمی)

e) $4.56x_1 + 2.18x_2 = 6.74$
$2.79x_1 + 1.38x_2 = 4.13$

۱۷ــ نشان دهید دستگاههای زیر بدوضع می‌باشند.

(با حساب مانتیس ۵ رقمی)

$$\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}$$

(با حساب مانتیس ۶ رقمی)

$$\begin{bmatrix} 1 & 2 \\ 1.00001 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.00003 \end{bmatrix}$$

۱۸ــ دستگاههای زیر را به روش ژاکوبی و گاوس ـ سایدل با دقت 10^{-6} حل کنید.

a)
$$\begin{aligned} 10x_1 - x_2 + 2x_3 \phantom{{}+3x_4} &= 6, \\ -x_1 + 11x_2 - x_3 + 3x_4 &= 25, \\ 2x_1 - x_2 + 10x_3 - x_4 &= -11, \\ 3x_2 - x_3 + 8x_4 &= 15, \end{aligned}$$

b)
$$\begin{aligned} 2x_1 - x_2 + 10x_3 \phantom{{}+8x_4} &= -11, \\ 3x_2 - x_3 + 8x_4 &= -11, \\ 10x_1 - x_2 + 2x_3 \phantom{{}+8x_4} &= 6, \\ -x_1 + 11x_2 - x_3 + 3x_4 &= 25. \end{aligned}$$

c) $$\begin{bmatrix} 4 & -1 & 0 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & -1 & 4 & 0 & 0 & -1 \\ -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 0 \\ 6 \\ -2 \\ 6 \end{bmatrix}$$

فصل پنجم
حل معادلات دیفرانسیل معمولی

۵.۱ ■ مقدمه

سرعت تغییر یک متغیر کرارا" در مسائل فیزیکی و معادلات ریاضی مربوط به یک مسئله، که اغلب برحسب مشتقات متغیر فرمول‌بندی می‌شوند، ظاهر می‌شود. به‌عنوان مثالی ساده، بیشتر خوانندگان با معادلات حرکت مرتبط با فاصله x، سرعت dx/dt و شتاب d^2x/dt^2 آشنائی دارند. برای تعدادی‌ازاین معادلات‌بدست‌آوردن‌جواب‌بوسیله آنالیز ریاضی‌امکان‌پذیر می‌باشد، اما دو مسئله ممکن‌است ازاین‌عمل‌ممانعت نماید. اولا" اگرچه بعضی مسائل به وسیلهٔ آنالیز نسبتا" ساده‌ای قابل حل هستند اما معادلات دیفرانسیل زیادی وجود دارند که حل آنها نیازمند به مهارت بالائی از دانش ریاضی می‌باشد. ثانیا"، آشنائی گسترده‌ای با مسائل فیزیکی فورا" نشان می‌دهد که معادلات دیفرانسیل زیادی وجود دارند که جواب آنها را نمی‌توان با یک صورت سادهٔ ریاضی ارائه کرد. روشهای تقریبی حل آنها تنها راه دسترسی به جواب می‌باشد. اما، مشکلات زیادی درکاربرد روشهای عددی وجود دارد، درصورتیکه آنالیز ریاضی فرمولی صریح برای حل آنها بیان کند نباید با عجله متوسل به یک روش عددی شد. حتی اگر مسئله دارای فرمولی صریح نباشد ممکن است بهتر باشد که وسیلهٔ آنالیز ریاضی مسئله را به‌صورت مناسبتری برای محاسبه عددی آماده کرد.

روند انتگرال‌گیری، اعداد ثابت دلخواه معرفی می‌کند که وسیلهٔ شرایط اضافی معلوم در تابع یا مشتق آن قابل تعیین می‌باشند، برای مثال، معادله دیفرانسیل مرتبه سوم به سه شرط اضافی برای تعیین سه ثابت دلخواه که ظاهر می‌شود نیاز دارد. بعنوان مثالی ساده، درصورتیکه از $dy/dx=2$ انتگرال بگیریم، جواب $y=2x+A$ بدست می‌آید، اگر شرط $y=4$ به‌ازای $x=0$ رادرنظر بگیریم آنگاه $A=4$.

مطابق با روشی که این شرایط مشخص میگردند دو نوع‌مسئله وجود دارد. اگر تمام

شرایط مورد نیاز تنها در یک نقطه داده شوند در اینصورت یک مسئله مقدار ــ آغازی داریم ، و روش حل از نقطه معلوم شروع شده و گام به‌گام درطول مجموعه‌مقادیرانتگرال‌گیری حرکت می‌نماید . به هر حال ، اگر شرایط داده شده در بیش از یک نقطه باشد برای شروع محاسبه، تنها اطلاعات در هر نقطه کافی نبوده و روش محاسبه شامل حل یک دستگاه معادلات ، یا به‌کارگیری مقادیر تخمینی در هر نقطه میباشد . سپس این مقادیر تخمینی ، همچنانکه محاسبات پیش می‌رود ، بوسیلهٔ تکرار تصحیح می‌شوند . این مسئله نوع دوم به مسئله مقدارـ مرزی موسوم بوده و آندسته که به طور جدی این نوع از مسئله را درنظر میگیرند محتاج به مطالعه پیش نیاز ریاضی قابل توجهی میباشند . قسمتهای بعدی فقط به مسئله مقدار ــ آغازی اختصاص خواهد داشت .

۵.۲ ■ مسئله‌مقدار- آغازی

شکل عمومی مسئله مقدار ــ آغازی وسیلهٔ مشخص کردن معادله‌ای برای مشتق و شرطی روی تابع در تنها یک نقطه داده می‌شود ، یعنی

$$\frac{dy}{dx}=f(x,y) \tag{5.1a}$$

$$y(a)=s \tag{5.1b}$$

تذکر : $f(x,y)$ نمایشگر رابطه‌ایست که شامل y و x میباشد ، مانند :

$\sin x+\sin y$ ، x^2+y^2 و غیره .

در اولین نظر ممکن است این معادلات برای کاربرد کلی خیلی‌مقدماتی به‌نظر آید اما با توسیع علامت ، توسعهٔ قابل توجه حیطهٔ عمل آنها امکان‌پذیر می‌باشد .

$$\frac{d^2y}{dx^2}+\sin x.\frac{dy}{dx}+\cos x.y=3 \tag{5.2a}$$

$$y(a)=1, \qquad \left(\frac{dy}{dx}\right)_{x=a}=-1 \tag{5.2b}$$

با قرار دادن $z=dy/dx$ می‌توان معادلات را به‌صورت زیر نوشت .

$$\frac{dz}{dx}=-\sin x.z-\cos x.y+3$$

$$\frac{dy}{dx}=z \tag{5.3a}$$

و

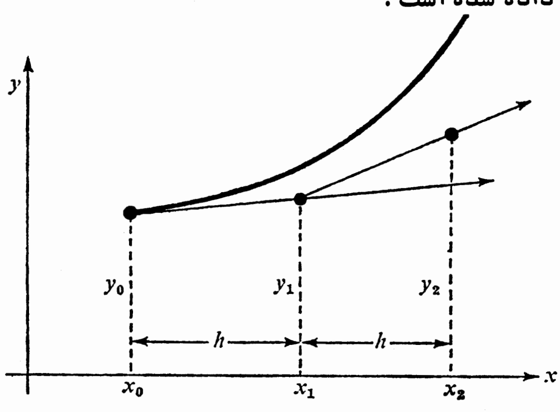
$$z(a)=-1$$

$$y(a)=+1 \tag{5.3b}$$

به‌مثال (۵۰۱)نیز مراجعه کنید. معادلات خیلی شبیه معادله(5.1) بوده اما دو معادله برای مشتق و دو شرط آغازی داریم. آن دسته که با نماد برداری آشنائی دارند در صورتی که به جای f, s, y مقادیر برداری جایگزین شود خواهیم دید که همان صورت معادلات می‌توانند به‌کار روند. بنابراین، معادلاتی از هر مرتبه که نسبت به مشتقات متعدد خطی باشند می‌توانند به‌همان صورت متعارف معادله (5.1) معرفی شده و روشهای حل توضیح داده شده در زیر به‌کار روند.

روشهای حل عددی علاوه بر طرق گوناگون می‌تواند از فرمول تفاضل ــ متناهی و سری تیلر بریده شده ناشی گردند. نتایج نشان می‌دهد که تقریب ساخته شده در هر گام خطائی تولید می‌کند. نشان داده خواهد شد که مفید بودن این روشها نه فقط به اندازهٔ این خطاها، بلکه به‌طریق زیاد شدن آنها، همچنانکه‌درطول ناحیه‌انتگرال‌گیری پیش میرویم، بستگی دارد. مفهوم سازگاری را که به خطای ناشی شده در هر نقطهٔ ویژه مرتبط می‌گردد، و مفهوم پایداری را که به خطایی که با پیشرفت محاسبه مربوط است بررسی می‌کنیم. بهرحال، قبل از توجه به ریاضیاتی که در انتخاب فرمول مناسبی وجود دارد، درنظر گرفتن حل یک معادله دیفرانسیل مرتبه اول ازنقطه‌نظر نمودار آموزنده خواهد بود.

معادله (5.1b) مقدار آغازی جواب و معادله (5.1a) مقدار مشتق در هر نقطه را می‌دهد. درصورتیکه بخواهیم منحنی را دنبال کنیم منطقی به‌نظر میرسد که ازنقطهٔ معلوم آغازی (x_0, y_0) $x_0 = a, y_0 = s,$ شروع کرده و از این نقطه در جهت مماس حرکت نمائیم. چنانچه این خط مستقیم تعقیب شود از منحنی جواب دور می‌شویم، بنابراین بعد از طی‌فاصله کوچک h در جهت x مختصات جدید $(x_0 + h, y_1)$ محاسبه می‌شودواین نقطه مانندیک‌نقطه پایه‌به‌کار می‌رود جهت مماس این نقطه جدید محاسبه شده وگام‌بعدی در این جهت، نقطه $(x_0 + 2h, y_2)$ می‌شود. این روند به روش اویلر موسوم بوده و در شکل (۵۰۱) نشان داده شده است.



شکل (۵۰۱)

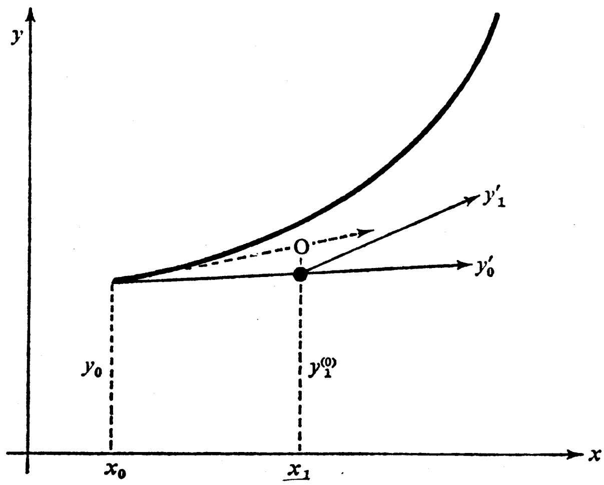
روش اویلر

این شیوه گام به گام را می‌توان تا اینکه تمام ناحیه انتگرال‌گیری را دربر گیرد تکرار کرد. باید توجه داشت که برای سهولت، طول گامها مساوی انتخاب شده‌اند. برای این روش همچنین طول گامهای مختلف h در گامهای مختلف امکان‌پذیر می‌باشد. از نمودار می‌توان دید که این روش بعید به‌نظر میرسد که بتواند جواب خیلی دقیقی را بدهد، حتی اگر طول گام کوچک باشد، چون جواب تقریبی همیشه در زیر و دور از منحنی از نوع فوق میباشد. بهر حال، وسیلهٔ یک پیرایش ساده جواب بسیار بهتری می‌تواند بدست آید.

بعد از اولین گام که جواب تقریبی y_1 در نقطه x_0+h بدست آمد آنگاه محاسبه مقدار مشتق در این نقطه امکان‌پذیر می‌گردد. حال مقادیر مشتق را در هر دو انتهای فاصله‌داشته‌و منطقی به‌نظرمیرسد که‌برای انتخاب‌جهت بهتر از نقطه آغازی، از مقدار متوسط این دو مقدار شروع نمائیم. بنابراین یک‌شیوه تصحیح باشروع از نقطه آغازی (x_0, y_0) به‌کار رفته و با به کارگیری مقدار متوسط جهت مماس تقریب جدیدی برای جواب در نقطه $x=x_0+h$ که در شکل (۵.۲) نشان داده شده است بدست می‌آید. این‌روند پیش‌بینی، آنگاه وسیله تصحیحی تا پایان فاصله می‌تواند ادامه داشته باشد. محاسبات در ذیل خلاصه می‌گردد.

۱ــ مطابق روش قبل $y_1^{(0)}$ را پیدا کنید.

$$y_1^{(0)} = y_0 + hf(x_0, y_0) \tag{5.4}$$



شکل (۵.۲) روش اویلر ــ ذوزنقه‌ای o معرف نقطه $(x_0+h, y_1^{(1)})$ است.

۲ــ مشتق را در نقطه $(x_0+h, y_1^{(0)})$ محاسبه نمائید.

۳ــ مقدار متوسط مشتقات را در انتهای فواصل برای ادامه از y_0 تا تقریب جدید $y_1^{(1)}$ به‌کار برید.

$$y_1^{(1)} = y_0 + \frac{h}{2}[f(x_0, y_0) + f(x_0 + h, y_1^{(0)})] \qquad (5.5)$$

۴ــ تصحیحات بیشتری را در این مرحله با به‌کار بردن آخرین مقدار به‌منظور دست یافتن تصحیح بعدی می‌توان به‌کار برد.

$$y_1^{(r+1)} = y_0 + \frac{h}{2}[f(x_0, y_0) + f(x_0 + h, y_1^{(r)})], \qquad r = 1, 2, \ldots \qquad (5.6)$$

۵ــ بعد از تعداد کافی از تصحیحات مقدار جدید y به عنوان نقطه شروع فاصله بعدی می‌تواند درنظر گرفته شود و گامهای (۴ ــ ۱) با اندیسهایی که نمو آنها واحد است تکرار می‌گردد. این شیوه تا رسیدن به نقطه انتهائی فاصله ادامه پیدا می‌کند.

روش فوق، که به روش اویلر ــ ذوزنقه‌ای موسوم است، مثال ساده‌ای از فرمول پیشگو ــ اصلاحگر بوده که پیش‌بینی آغازی وسیله روند تکراری تصحیح می‌گردد. روند تکرار فوق روندی بسیار ابتدایی است و باید توجه داشت که هر روش تکراری متعارف می‌تواند به‌جای آن به‌کار رود.

شکل کلی معادله‌ای که حل می‌شود وسیلهٔ رابطه زیر بیان می‌گردد.

$$y_{n+1} = \phi_n + h\beta_{n+1} f(x_n + h, y_{n+1}) \qquad (5.7)$$

مقدار ϕ_n معرف قسمتهایی از فرمول است که به مقادیر قبلی بستگی داشته بنابراین در تکرار عوض نمی‌گردد. بجز وقتیکه تابع $f(x, y)$ طوری است که معادلهٔ (5.7) دارای جوابی صریح می‌باشد، حل تکراری ضروری است و این جواب با نوشتن معادله به‌صورت

$$y_{n+1} - \phi_n - h\beta_{n+1} f(x_n + h, y_{n+1}) = 0 \qquad (5.8)$$

وبه‌کارگیری یکی از روشهای فصل ۲ بدست می‌آید. شیوه تکرار مستقیم خیلی زیاد به‌کار می‌رود، اما روش گیر Gear که بعدا" دراین فصل بحث خواهد شد، روند تکراری خیلی دقیقی است که به‌کار می‌رود.

روشهایی که در عمل به‌کار می‌رود مبتنی برهمان اساس روش فوق است، اما اطلاعات چندین مقدار قبلی x به‌منظور دست یافتن به دقت بالاتر به‌کار می‌رود. ردهٔ دیگری از روشها که به طور گسترده‌ای به‌کار می‌رود از نوع رونگه ــ کوتا می‌باشد. این روشها مبتنی بر محاسبهٔ تقریبهای مشتقات نه تنها در نقاط انتهائی فاصله بلکه در نقاط میانی نیز می‌باشد. مهمترین امتیاز این روش عبارتست از اینکه محاسبه، تنها مقادیر در نقطه آغازی را به‌کار میگیرد و اطلاعات قبلی استفاده و یا ذخیره نمی‌شود.

۵.۳ ■ روشهای پیشگو ــ اصلاحگر

۵.۳.۱ ■ صورت کلی معادلات

روشهای پیشگو ــ اصلاحگر در کاربــرد معمولی مبتنی بر اطلاعات چندین نقطه

قبلی است ذخیره تقریبهای جواب و مشتقات آن طوریکه در محاسبات بعدی به‌کار رود ضروری است . بنابراین بایستی توجه خاص به‌محاسبهٔ چندین مقدار آغازی داشت زیرا مقادیر گذشته در آن مرحله موجود نمی‌باشد . به هرحال ، ابتدا هم خود را در مسئلهٔ اصلی که انتخاب فرمول مناسب است متمرکز خواهیم کرد .

ناحیه انتگرال‌گیری به فواصل مساوی h تقسیم می‌شود بنابراین ، برای جواب در فاصله $a \leqslant x \leqslant b$ ، داریم .

$$h = (b-a)/N \tag{5.9a}$$
$$x_n = a + nh, \qquad n = 0, 1, \ldots, N \tag{5.9b}$$

تشخیص بین جواب واقعی ، که با $y(x_n)$ و جواب تقریبی که با y_n نشان داده می‌شود حائزاهمیت می‌باشد . درحالت اول یک تابع تعریف شده برای تمام‌مقادیر $x\ (a \leqslant x \leqslant b)$ و در حالت دوم یک مجموعه متناهی از مقادیر $y_n\ (n = 0,\ 1, \ldots, N)$ داریم مفهوم f_n برای بیان $f(x_n, y_n)$ به‌کار خواهد رفت .

گرایش معمول برای بدست آوردن فرمول عبارتست از توجه به مسئله ، مانند یکی از روشهای انتگرال‌گیری تقریبی و به‌کار بردن روشهای تفاضل ــ متناهی برای ایجاد فرمول . اگر از معادلهٔ دیفرانسیل (5.1a) در فاصلهٔ x_{n+k} و x_n انتگرال بگیریم‌داریم

$$y(x_{n+k}) - y(x_n) = \int_{x_n}^{x_{n+k}} f(x, y)\, dx \tag{5.10}$$

این انتگرال مستقیما" قابل حل نمی‌باشد زیرا y در زیر علامت انتگرال آمده است فرمولهای متعددی وسیلهٔ محاسبات تفاضل ــ متناهی درکتابهائی نظیر Redish (1961) یا *Modern Computing Methods* (1961) آورده شده است که در اینجا دنبال نخواهد شد .

فرمولی از این نوع که غالبا" به‌کار می‌رود عبارتست از فرمول باز یا بستهٔ نیوتن‌ــ کاتس که به ترتیب متناظر با فرمول پیشگو و اصلاحگر می‌باشد . تعدادی از مثالهای ساده این‌فرمول در جدول(۵۰۱)داده شده است فرمول دیگری که اغلب به‌آن مراجعه خواهیم کرد عبارتست‌از فرمول Adams-Bashforth و Adams-Moulton ، مثالهای ساده‌ای در جدول (۵۰۲) نیز آورده شده است .

یافتن فرمول از نقطه‌نظر کلی‌تر حائز اهمیت است زیرا ، اساسی را مشخص می‌کند که برای مسائل متعدد به‌کار می‌رود . روش عمومی عبارتست‌از انتخاب فرمولی شامل تعدادی از پارامترها که به‌طور اختیاری انتخاب می‌شوند طوریکه در خاصیت ریاضی یا محاسبه‌ای مطلوبی صدق نماید آنگاه می‌توان مجموعه‌ای از شرایط که باید برقرار باشند درنظر گرفت ، و این شرایط مقادیر قابل قبول پارامترها را مشخص می‌کند . دراین حالت

اطلاعات موجود مشتمل بر مقادیر محاسبه شدهٔ جواب و مشتقات آن در رشته‌ای از نقاط می‌باشد. شکل انتخابی برای فرمول انتگرال‌گیری ترکیب خطی از این مقادیر خواهد بود.

$$\sum_{r=0}^{k} \alpha_r y_{n+r} - h \sum_{r=0}^{k} \beta_r f_{n+r} = 0 \tag{5.11}$$

که در آن α_r و β_r ضرائب ثابت بوده که برای دست یافتن به بهترین تقریب انتخاب می‌گردند. از آنجائیکه اندازهٔ ضرائب بوسیلهٔ ضرب این معادله درعدد ثابت اختیاری می‌تواند تغییر کند برای سهولت در تجزیه و تحلیل بعدی ضرائب را بوسیله تقسیم به α_k استاندارد می‌نمائیم. لذا، ضریب α_k همیشه واحد گرفته می‌شود. توجه کنید تمام مقادیر $y_{n+r}\ (r=0,1,\ldots,k-1)$ مقادیر محاسبه شده قبلی هستند که ذخیره نیز شده‌اند و بلافاصله برای محاسبه موجود می‌باشند. بنابراین، هر فرمولی که در آن $\beta_k = 0$ قادر به محاسبهٔ y_{n+k} بطور صریح می‌گردد. این فرمولها به نامهای متعدد از قبیل فرمول صریح، باز یا پیشگو نامیده می‌شود. درحالتیکه $\beta_k \neq 0$ مقدار y_{n+k} نیز در $f_{n+k} = f(x_{n+k}, y_{n+k})$ رخ می‌نماید که به صورت غیر خطی بوده‌و بنابراین، نیازمند عمل تکرار خواهد بود. این فرمول ضمنی، بسته، یا اصلاحگر نامیده می‌گردد. ضرائب α_r و β_r طوری انتخاب می‌شود که دقت و پایداری برقرار گردد.

۵.۳.۲ ■ دقت

یکی از خواص مهم فرمول عبارتست‌از اینکه خطا، وسیلهٔ کاربرد تقریب در یک نقطه ویژه ارائه می‌گردد. این خطا نه فقط بمختصات نقطه موردنظر، بلکه به مشتق آن نیز بستگی دارد. بنابراین، بیان اینکه کدام فرمول نسبت به دیگری دقیق‌تر است نسبتا" مشکل می‌باشد. راهی که اغلب برای تعریف دقت به‌کار می‌رود عبارتست‌از جانشین کردن مقادیر متناظر با جواب واقعی در فرمول چندگامی (5.11) و بسط سری تیلر هر جمله. آنگاه مرتبه دقت به وسیلهٔ کمترین توان h که دارای ضریب غیر صفر می‌باشد مشخص می‌گردد. درصورتی‌که اولین جمله‌ای که حذف نشده است شامل توان h^{p+1} باشد، آنگاه فرمول از مرتبه p گفته می‌شود ضروری است که مرتبهٔ دقت حداقل یک باشد. و فرمولی‌که در این شرایط صدق کند سازگار گفته می‌شود. مطابق این تعریف از دقت، ضرائب طوری باید انتخاب گردند که حتی‌الامکان p بزرگ باشد، یعنی توانهای بیشتری از h باید ضریب صفر داشته باشند.

به‌عنوان مثال، فرمول صریحی که فقط برای سه نقطه به‌کار می‌رود درنظر گرفته و ضرائب طوری انتخاب‌خواهند شد که فرمولی با بیشترین دقت بدست آید. گرچه، نشان

داده خواهد شد که فرمول ایجاد شده به علت اینکه قویا" ناپایدار است ازنظر محاسبه‌ای کاربرد ندارد، خطاها گام به گام، با پیشرفت محاسبات، با الگوی غیر قابل کنترلی افزایش پیدا می‌کنند. فرمولی که درنظر گرفته می‌شود عبارتست‌از:

$$y_{n+2}+\alpha_1 y_{n+1}+\alpha_0 y_n-h[\beta_1 f_{n+1}+\beta_0 f_n]=0 \tag{5.12}$$

بطوریکه با جاگذاری مقادیر جواب واقعی، فرمول خطا بدست می‌آید.

$$y(x_{n+2})+\alpha_1 y(x_{n+1})+\alpha_0 y(x_n)-h[\beta_1 f(x_{n+1},y(x_{n+1}))+\beta_0 f(x_n,y(x_n))]=E_{n+2} \tag{5.13}$$

بسطهای سری تیلر برای جملات متعدد عبارتنداز:

$$y(x_{n+2})=y(x_n)+2hy'(x_n)+\frac{4h^2}{2}y''(x_n)+\frac{8h^3}{6}y'''(x_n)+\frac{16h^4}{24}y^{(\mathrm{IV})}(x_n)+\cdots \tag{5.14a}$$

$$y(x_{n+1})=y(x_n)+hy'(x_n)+\frac{h^2}{2}y''(x_n)+\frac{h^3}{6}y'''(x_n)+\frac{h^4}{24}y^{(\mathrm{IV})}(x_n)+\cdots \tag{5.14b}$$

$$y(x_n)=y(x_n) \tag{5.14c}$$

$$f[x_{n+1},y(x_{n+1})]=y'(x_{n+1})=y'(x_n)+hy''(x_n)+\frac{h^2}{2}y'''(x_n)+\frac{h^3}{6}y^{(\mathrm{IV})}(x_n)+\cdots \tag{5.14d}$$

$$f[x_n,y(x_n)]=y'(x_n) \tag{5.14e}$$

حال، جملات برحسب α_r و β_r طوری انتخاب می‌گردند که جملات با مرتبه پائین‌ترنسبت به h دارای ضریب صفر باشد. بدین‌ترتیب ضریب $y(x_n)$ درصورتی صفر خواهد بود که $1+\alpha_1+\alpha_0=0$. دراین راه چهار معادله سازگار بدست می‌آید که قابل حل برای چهار مجهول می‌باشد.

$$\begin{aligned}
&1+\alpha_1+\alpha_0=0 \qquad &&2+\frac{\alpha_1}{2}-\beta_1=0\\
&2+\alpha_1-\beta_1-\beta_0=0 \qquad &&\frac{4}{3}+\frac{\alpha_1}{6}-\frac{\beta_1}{2}=0
\end{aligned} \tag{5.15}$$

این معادلات دارای جواب‌های

$$\begin{aligned}
&\alpha_1=4 \qquad \alpha_0=-5\\
&\beta_1=4 \qquad \beta_0=2
\end{aligned} \tag{5.16}$$

می‌باشد و این ضرائب خطائی از مرتبه h^4 می‌دهد (به‌مثال ۵.۵ نیز مراجعه کنید).

۵.۳.۳ ■ پایداری

موضوع پایداری در قسمت(5.5) با تفصیلات بیشتر مورد بحث قرار گرفته است اما آشکارا می‌توان دید که اصرار به خواستن دقت زیاد، همانند مثال فوق ازنقطه‌نظر پایداری می‌تواند به‌نتایج خیلی بدی منتهی گردد. در حقیقت این مطلب تاییدی خواهد بود که در روندهای عددی بایستی همواره راهی که خطاها سرتاسر روند بزرگ

می‌شوند مانند دقت هر واحد محاسبه‌ای درنظر گرفته شوند . جهت توضیح این مطلب

$$\frac{dy}{dx}=0 \tag{5.17a}$$

$$y(0)=1 \tag{5.17b}$$

که به ازای تمام مقادیر x دارای جواب $y=1$ است را درنظر می‌گیریم . در صورتیکه برای معادله ساده‌ای از این نوع روش عددی به‌کار نمی‌رود ، انتظار داریم که این روش برای حل این معادله مناسب باشد .

معادله تفاضل ــ متناهی (5.12) دارای شکل

$$y_{n+2}+4y_{n+1}-5y_n=0 \tag{5.18}$$

بوده از آنجائیکه یک معادله خطی است به سادگی قابل حل می‌باشد . بحث خوبی از حل‌معادلات‌تفاضل‌ـ متناهی‌در Goldberg (1958) موجود می‌باشد ، امابرای‌هدف‌ما چند نتیجه ساده کفایت خواهد کرد . حل معادله (5.18) به‌وسیله جانشین کردن یک جواب آزمایشی $y_n=A(z)^n$ که به معادله چند جمله‌ای کمکی منتهی می‌گردد قابل محاسبه می‌باشد . اگر ریشه‌های چند جمله‌ای کمکی متمایز باشند ، مقادیر z را برای y_n که در معادله تفاضل ــ متناهی صدق می‌کند می‌دهد . دراین حالت معادله کمکی با جایگذاری به‌صورت

$$Az^n[z^2+4z-5]=0 \tag{5.19}$$

بدست می‌آید ، که دارای دو جواب با معنی $z_1=1$ و $z_2=-5$ می‌باشد . کلیه جوابهای معادله تفاضلی عبارتستاز :

$$\begin{aligned} y_n &= A_1(z_1)^n+A_2(z_2)^n \\ &= A_1(+1)^n+A_2(-5)^n \end{aligned} \tag{5.20}$$

مقادیر A_1 و A_2 به‌کمک شرایط آغازی پیدا می‌شوند و در این حالت ، در حدود دقت کامپیوتر تقریبا" A_1 مساوی واحد و A_2 مساوی صفر خواهد شد . بدین‌ترتیب ، اولین جمله متناظر با جواب بوده اما دومین جمله ، جمله‌ای است کاذب که دارای بیشترین خواص غیرقابل قبول می‌باشد . چنین جملات کاذبی همواره زمانیکه روشهای گام‌به‌گام به‌منظور افزایش دقت‌طرح به کار می‌رود رخ خواهد داد . دراین حالت بزرگی جمله کاذب با مضربی‌از 5 در هر گام افزایش پیدا می‌کند ، بنابراین حتی اگر A_2 کوچک باشد این جمله تدریجا" به جواب واقعی غالب می‌گردد . برای‌مثال ، بعدازفقط 10 گام جمله $(z_2)^n$ تقریبا" برابر 10^7 شده درصورتیکه درآغاز اگر A_2 کمتر از 10^{-7} باشد جمله اضافی به بزرگی جمله واقعی میرسد و ازاین نقطه حائز اهمیت می‌باشد .

بدیهی است جائیکه فرمول جملات کاذب معرفی کند باید بلافاصله این جملات

ازبین بروند و شرایطی که این خاصیت را مطمئن سازد در قسمت (۵۰۵) مورد بحث قرار گرفته است. عدم پایداری قوی در معادله فوق نشان داده شده و بوسیله جمله مشتق در نقطه x_{n+2} بسادگی قابل حذف می‌باشد اما دقت افزایش پیدا نمی‌کند. این عمل چهار معادله با پنج مجهول را داده که در آن یک مجهول به‌صورت پارامتر آزاد درنظر گرفته شده که به‌منظور دست یافتن به‌خواص پایداری خوب مطابق ذیل انتخاب می‌گردد.

مجموعه معادلات به‌صورت زیر خواهد بود.

$$\begin{aligned} &1+\alpha_1+\alpha_0 = 0 \qquad && 2+\frac{\alpha_1}{2}-2\beta_2-\beta_1=0 \\ &2+\alpha_1-\beta_2-\beta_1-\beta_0=0 \qquad && \frac{4}{3}+\frac{\alpha_1}{6}-2\beta_2-\frac{\beta_1}{2}=0 \end{aligned} \tag{5.21}$$

تمام ضرائب به‌صورت جملاتی از یک ضریب تنها، مانند α_0، بیان می‌گردد و آنگاه این ضریب به‌منظور دست‌یافتن به پایداری انتخاب خواهد شد.

$$\begin{aligned} &\alpha_1=-1-\alpha_0 \qquad && \beta_2=\frac{5}{12}+\frac{\alpha_0}{12} \\ &\alpha_0=\alpha_0 && \beta_1=\frac{2}{3}-\frac{2}{3}\alpha_0 \\ & && \beta_0=\frac{-1}{12}-\frac{5}{12}\alpha_0 \end{aligned} \tag{5.22}$$

با به‌کارگیری جواب آزمایشی $A(z)^n$ مانند قبل معادله کمکی زیر بدست می‌آید.

$$z^2-(1+\alpha_0)z+\alpha_0=0 \tag{5.23}$$

که دارای جوابهای $z=1$ یا $z=\alpha_0$ است.

به‌منظور اطمینان از اینکه با افزایش n جواب کاذب افزایش پیدا نمی‌کند به ناحیه $-1\leqslant\alpha_0<+1$ نیازمند هستیم و امکان اینکه چندین فرمول پایدار رابدهدوجوددارد. تعدادی مثال در ذیل داده شده است.

$$\begin{array}{ccc} \alpha_0 & -1 & 0 \\ \alpha_1 & 0 & -1 \\ \beta_0 & \frac{1}{3} & -\frac{1}{12} \\ \beta_1 & \frac{4}{3} & \frac{8}{12} \\ \beta_2 & \frac{1}{3} & \frac{5}{12} \end{array} \tag{5.24}$$

اولین ستون قانون سیمپسون Simpson و دومین ستون یکی از فرمولهای ضمنی Adams–Moulton است که سیمائی از روش Nordsieck درشکلی پیراسته می‌باشد.

باید توجه داشت که این گرایش با مطالعه مسئله‌ساده تعریف شده به‌وسیله (5.17) توسعه یافته است. هرطرح تفاضل ــ متناهی که ناپایدار باشد درکاربرد معمولی، برای این مثال مناسب نمی‌باشد. طرحی را قویا " ناپایدار گوئیم، که درآن هر ریشهٔ معادله کمکی دارای قدرمطلق بزرگتر از واحد بوده، یا هر ریشه مکرر دارای قدرمطلق مساوی

واحد باشد. بهرحال، حتی اگر توجه خود را محدود به طرحهائی نمائیم که قویا" ناپایدار نباشند طرحهای دیگری موجود هستند که فقط بهطور ضعیف پایدار میباشند. برای مسائل معینی نتایج رضایتبخشی میتواند بدست آید اما برای سایر مسائل روشها نباید بهکار رود این قسمتها با تفصیلات و بحث بیشتر در قسمت ۵۰۵۰۳ ارائه شدهاند (همچنین بهمثال ۵۰۵ مراجعه شود).

۵.۳.٤ ■ روشهای محاسبه

اگرچه بعضی مواقع درنشان دادن عامل موثر در انتخاب ضرائب وسیلهاستفاده کننده وقت صرف میشود، درحالت کلی، مهارت شخصی در بهکارگیری روش، موثر خواهد بود. در نمایش یک روش چندگامی مانند روش پیشگوی Adams–Bashforth و روش اصلاحگر Adams–Moulton گامهای متعدد موجودمیباشد.

۱ـ در شروع روند روشی برای یافتن چند مقدار آغازی لازم است، چون یک روش چند گامی نوعی، مقادیر را در چندین نقطه متوالی بهکار میبرد و در ابتدا فقط یک مقدار موجود است.

۲ـ فرمول پیشگو برای بدست آوردن اولین تخمین مقدار جدید y بهکار میرود.

$$y^{(0)}_{n+4} = y_{n+3} + \frac{h}{24}[55f_{n+3} - 59f_{n+2} + 37f_{n+1} - 9f_n] \qquad (5.25)$$

۳ـ مقدار پیشبینی شده جهت تعیین ضریب زاویه محاسبه میگردد.

$$f^{(0)}_{n+4} = f(x_{n+4}, y^{(0)}_{n+4})$$

۴ـ یکفرمول اصلاحگر بهطور تکراری بهکار میرود.

$$y^{(r+1)}_{n+4} = y_{n+3} + \frac{h}{24}[9f^{(r)}_{n+4} + 19f_{n+3} - 5f_{n+2} + f_{n+1}] \qquad r = 0, 1, \ldots \qquad (5.26)$$

۵ـ بعد از کامل شدن عمل تکرار، گامهای ۴ ـ ۲، با نمواندیسها باندازهٔ یک واحد، جهت یافتن مقدار دیگر y تکرار شده و این روند تا رسیدن به پایان فاصله انتگرال گیری ادامه مییابد.

مسئله یافتن مقادیر آغازی برای چندگامی درطرحهای مختلف به راههای متعدد حل شده است. روش بسط سری تیلر برای چندین گام میتواند بهکار رود، اما این بیشتر برای محاسبه دستی مناسب است تا کاربرد در کامپیوتر. (مثال ۵۰۲ راببینید) روش دیگر عبارتستاز کاربرد طرح گام ـ ساده اویلر ذوزنقهای با اندازه گام کوچکتر از طرح انتگرالگیری اصلی طوریکه دقت قابل مقایسه تامین گردد. (مثال ۵۰۴ راببینید) یا روشی از رونگه ـ کوتا که در قسمت ۵۰۴ توضیح داده شده که برای طرح تک گامی مناسب

می‌باشد به ویژه طرح رونگه ــ کوتا ــ مرسن Runge-Kutta-Merson معیاری راخواهد داد که وسیله آن می‌توان خطائی که در گامهای انتگرال‌گیری آغازی به‌اندازه کافی کوچک باشد کنترل نمود. روشهای خودکار Nordsieck و Gear مذکور در قسمت (۵۰۶)، یک سری از روشهای چندگامی با مراتب مختلف را به‌کار می‌برد، طوریکه روند از یک طرح گام ــ ساده شروع شده و برای رسیدن به مرتبه مورد نیاز زمانیکه نقاط بعدی پیدا می‌شود یک سری فرمولهای مختلف را می‌سازد.

توجه در طرح روند شروع، به‌منظور اطمینان از اینکه دقت حداقل معادل دقت طرح انتگرال‌گیری اصلی باشد، امریست ضروری. از آنجائیکه طرح اصلی معمولاً شامل توان خیلی بالائی از h در جمله خطا است. این بدان معنی است که روشهای شروع بایستی مبتنی‌بر کاهش بیشتر اندازه گام‌باشد. حال اینکه چطور اصلاحگر بایستی به‌کار رود، باید بررسی شود. ساده‌ترین گرایش عبارتست‌از ادامه تکرار برای هرچند دفعه به منظور اطمینان از دقت کافی. عملاً، این بدین معنی است که تفاضل بین تکرارهای متوالی از مقدار کوچک مشخص کوچکتر گردد. البته، داشتن روند همگرا و شرایطی که همگرائی را مطمئن سازد ضروری است و بسهولت می‌توان آن را بدست آورد. شکل کلی معادله اصلاحگر عبارتست‌از:

$$y_{n+k} = h\beta_k f(x_{n+k}, y_{n+k}) + \phi_{n+k-1} \qquad (5.27)$$

که در آن ϕ_{n+k-1} شامل تمام جملاتی است که قبلاً مقادیر معلومی را دارند و بنابراین در اثناء روند تکرار تغییر نمی‌کنند معادله تکرار به‌صورت

$$y_{n+k}^{(r+1)} = h\beta_k f(x_{n+k}, y_{n+k}^{(r)}) + \phi_{n+k-1}, \qquad r = 0, 1, \ldots \qquad (5.28)$$

خواهد بود و ما به‌خطائی که به‌صورت

$$\begin{aligned} e_{n+k}^{(r+1)} &= y_{n+k} - y_{n+k}^{(r+1)} \\ &= h\beta_k[f(x_{n+k}, y_{n+k}) - f(x_{n+k}, y_{n+k}^{(r)})] \end{aligned} \qquad (5.29)$$

داده شده باشد توجه داریم.

اگر چه برای همگرائی شرایط خیلی‌پیچیده‌ای می‌توان تعریف کرد اما در صورتیکه $f(x,y)$ به‌ویژه نسبت به y مشتق‌پذیر باشد امکان یافتن یک شرط ساده که برای همگرائی کافی باشد موجود است. با به‌کار بردن قضیه مقدار میانگین حساب دیفرانسیل داریم.

$$f(x, y_1) - f(x, y_2) = (y_1 - y_2)\frac{\partial f}{\partial y}(x, \zeta) \qquad (5.30)$$

که در آن ζ در فاصله (y_1, y_2) قرار گرفته است. در ناحیه روند انتگرال‌گیری K را مساوی $K = \max|\partial f/\partial y|$ قرار میدهیم. آنگاه

$$
\begin{aligned}
|e_{n+k}^{(r+1)}| &= h\,|\beta_k|\left|\frac{\partial f}{\partial y}\right|\,|y_{n+k}-y_{n+k}^{(r)}| \\
&\leqslant hK\,|\beta_k|\,|e_{n+k}^{(r)}| \\
&\leqslant [(hK)\,|\beta_k|]^{r+1}\,|e_{n+k}^{(0)}|
\end{aligned}
\tag{5.31}
$$

در صورتیکه $hK|\beta_k|<1$ خطا در هر تکرار نقصان پیدا کرده و زمانی که r به سمت بینهایت میل کند دارای حد صفر می‌گردد. بنابراین شرط مطلوب برای همگرائی ،به‌ازای مقادیر x و y که در رشته تکرار رخ می‌دهد ، عبارتست‌از :

$$h\,|\beta_k|\left|\frac{\partial f}{\partial y}\right|<1 \tag{5.32}$$

بدین‌ترتیب ، اگر ماکزیمم مقدار $|\partial f/\partial y|$ را بدانیم یک اندازه مناسب برای h می‌توان انتخاب کرد .

بایستی توجه داشت که ما نه‌تنها بامقدار $\partial f/\partial y$ در حوزهٔ جواب بلکه با مقدار آن درهرناحیه مقادیری که روند تکرار بایستی به آن برسد سروکار داریم . البته ، این ناحیه بعدی می‌تواند خیلی بزرگتر از **ناحیه جوابها باشد** .

همچنین باید درنظر بگیریم که زمان لازم برای انجام چندین تکرار چه میزان می‌باشد . زمانیکه محاسبه مشتق موردنظر باشد (اغلب این حالت است) زمان لازم برای دوبرابر تعداد تکرارها با زمان لازم برای دو برابر تعداد گامها با نصف اندازه گام h مساوی است . خطای برشی یک فرمول به‌توان h بستگی دارد ، بنابراین کاهش اندازه گام ، خطا را در یک الگوی قوی‌تر خواهد کاست .

درصورتیکه فرمول به‌کار رفته در معادله (5.25) را که دارای جمله خطائی به‌صورت $(251/720)h^5f^{\mathrm{IV}}(\zeta_1)$ است درنظر بگیریم ، تنصیف فاصله ، خطا را با عاملی درحدود 32 در هر مرحله کاهش خواهد داد . افزایش دو خطا این عامل را به 16 تقلیل می‌دهد . به هر حال ، فرمول اصلاحگر دارای خطائی است به‌صورت $(-19/720)h^5f^{\mathrm{IV}}(\zeta_2)$ طوریکه اگر اصلاحگر برای همگرائی به‌کار می‌رفت کاهش درخطا با مقایسه فرمول پیشگو در یک عامل از مرتبه $13 \approx \frac{251}{720}\cdot\frac{720}{19}$ خواهد بود . بدیهی است ، که تنصیف گام با عامل 16 مرجح است . برای بیشتر فرمولها نتایج مشابهی نتیجه می‌گردد . بنابراین روش تکرار برای همگرائی در حالت کلی اقتصادی نمی‌باشد .

تزکیه‌ای برای طرح محاسبه‌ای با به‌کار بردن پیشگو و اصلاحگر ، که فایده آن نیز به اثبات رسیده است ، وجود دارد . یک پیشگو و اصلاحگر با مرتبه جمله خطای یکسان انتخاب می‌گردد و بنابراین یک ترکیب خطی از مقدار پیش‌بینی شده و مقدار اصلاح شده می‌تواند برای کاهش خطا به‌کار رود . دلیلی برای این روند در زیر آورده شده است ، اما باید توجه داشت که مطالعه دقیق این مسائل خارج از حوصله این کتاب می‌باشد .

به‌عنوان مثالی از این روش، معادلهٔ پیشگو (5.25) همراه با معادله اصلاحگر (5.26) به‌کار خواهد رفت. اگر این معادلات تفاضل ــ متناهی برای جواب واقعی به‌کار رود خواهیم داشت.

$$y(x_{n+4}) = y(x_{n+3}) + \frac{h}{24}[55f(x_{n+3}, y(x_{n+3})) - 59f(x_{n+2}, y(x_{n+2})) + 37f(x_{n+1}, y(x_{n+1})) - 9f(x_n, y(x_n))] + \frac{251}{720}h^5 f^{\mathrm{IV}}(\zeta_1, y(\zeta_1)) \quad (5.33)$$

و

$$y(x_{n+4}) = y(x_{n+3}) + \frac{h}{24}[9f(x_{n+4}, y(x_{n+4})) + 19f(x_{n+3}, y(x_{n+3})) - 5f(x_{n+2}, y(x_{n+2})) + f(x_{n+1}, y(x_{n+1}))] - \frac{19}{720}h^5 f^{\mathrm{IV}}(\zeta_2, y(\zeta_2)) \quad (5.34)$$

که در آن ζ_1 و ζ_2 مقادیری از x در فاصله $[x_n, x_{n+4}]$ می‌باشند. روشی که به کار می‌رود به دو فرض بستگی دارد و با انحراف از این دو ممکن است روش رضایت‌بخش نگردد. بدواً فرض شده است که $f^{\mathrm{IV}}(\zeta, y(\zeta))$ در فاصله داده شده ثابت است بنابراین جملات خطا فقط با مضارب ثابتی با هم فرق دارند. خطای موجود دراین فرض برای یک تابع خوش رفتار درفاصله کوتاه کوچک خواهد شد. همچنین، اولین پنج جمله در سمت راست وسیله مقادیر محاسبه شده در نقاط متناظر زمانیکه y_{n+4} محاسبه می‌شود جانشین می‌گردد. بدین ترتیب، جمله خطای دو معادله فوق تنها درصورتی خطای واقعی را ارائه خواهد کرد که مقادیر به‌کار رفته در معادله تفاضل ــ متناهی به‌طور تصادفی درست باشند.

اساس روش عبارتست‌از به‌کارگیری دو معادله به‌منظور حذف خطای معرفی شده به صورت جمله h^5، اما خواننده بایستی فرض نماید که هر دو در غیاب خطای انباشتگی و طبیعت تقریبی مفروضات می‌باشند. بعنوان مثال، اگر مقادیر پیش‌بینی و اصلاح شده به‌ترتیب $y_{n+4} = 3{\cdot}195$ و $y_{n+4} = 3{\cdot}141$ باشند آنگاه

$$\begin{aligned} 3{\cdot}195 &= y_{n+3} + \frac{h}{24}[55f_{n+3} - 59f_{n+2} + 37f_{n+1} - 9f_n] \\ 3{\cdot}141 &= y_{n+3} + \frac{h}{24}[9f_{n+4} + 19f_{n+3} - 5f_{n+2} + f_{n+1}] \end{aligned} \quad (5.35)$$

حال معادلات (5.33) و (5.34) به‌صورت زیر خواهد بود.

$$\begin{aligned} y(x_{n+4}) &\approx 3{\cdot}195 + \tfrac{251}{720}E \\ y(x_{n+4}) &\approx 3{\cdot}141 - \tfrac{19}{720}E \end{aligned} \quad (5.36)$$

حال حذف E امکان‌پذیر می‌گردد.

$$(19+251)y(x_{n+4}) = 19 \times 3{\cdot}195 + 251 \times 3{\cdot}141 \qquad (5.37)$$
$$y(x_{n+4}) = 3{\cdot}145$$

آنگاه در هرگام روند سه مرحله پیش‌بینی تصحیح و پیشرفت قبل از حرکت به‌گام دیگر انجام می‌گردد.

۵.۴ ■ روشهای رونگه - کوتا

۵.۴.۱ ■ تعدادی فرمول نمونه ای رونگه - کوتا

روشهای قبلی مقادیر تابع و مشتقات آن را در چندین نقطه قبلی به‌کار گرفت. از آنجائیکه در این حالات مقادیر قبلی موجودنمی‌باشند، این عمل در شروع روش و کاهش فاصله ایجاد مشکلاتی می‌نماید. روشهای رونگه ــ کوتا، تمام‌مقادیر تابع و مشتق آن را که در شروع از یک نقطه آغازی در هر گام محاسبه‌ای احتیاج است ایجاد می‌نماید. بدین‌ترتیب این روشها برای شروع حل یا تغییر اندازه گام ایده‌آل می‌باشند. اما نقاط ضعفی نیز دارند که نیازمند به محاسبه چندین مشتق در هر گام می‌باشند. نحوه یافتن فرمول رونگه ــ کوتا به‌علت طولانی بودن نشان داده نمی‌شود، اما برای نشان دادن دلیل طرح و شیوه محاسبه‌ای یک مثال نمونه به‌کار خواهد رفت.

طرح معمول مرتبه ــ چهارم به‌وسیله معادلات زیر داده شده است که برای سادگی اندازه گام ثابت فرض می‌گردد (به‌مثال ۵۰۶ مراجعه شود). در یک طرح با تفصیل بیشتر، اندازه گام با پیشروی انتگرال‌گیری به‌منظور رسیدن‌به معیار دقت تغییر می‌یابد.

$$\begin{aligned} k_1 &= hf(x_n, y_n) \\ k_2 &= hf(x_n + h/2, y_n + k_1/2) \\ k_3 &= hf(x_n + h/2, y_n + k_2/2) \\ k_4 &= hf(x_n + h, y_n + k_3) \end{aligned} \qquad (5.38)$$

$$y_{n+1} = y_n + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \qquad (5.39)$$

از آنجائیکه $dy/dx = f(x,y) \approx \Delta y/h$ می‌توان دید که k_r تخمینهای نمو y را در نقطه پایانی سمت چپ، دوبار در نقطهٔ میانی و یک بار در نقطه پایانی سمت راست معرفی می‌کند. در هر محاسبه تمام تخمینهای قبلی نمو، k_r موجود بوده و می‌تواند برای محاسبه مقدار y در $f(x,y)$ به‌کاررود. یک مقدار وزین از y افزایش حاصل پیدا کرده آنگاه به‌منظور یافتن جواب بعدی به مقدار قبلی اضافه می‌گردد.

در فرمول فوق در هر طرف راست تنها یک k_r ظاهر می‌شود که آنرا فرمولی برای کاربرد در ماشین ــ رومیزی ممکن می‌سازد. اما از آنجائیکه k_r در ترکیبات مختلف در سمت راست می‌تواند به‌کار رود چندین فرمول به‌صورت کلی

$$k_r = h_n f(x_n + \alpha_r h_n, y_n + \sum_{s=1}^{r-1} \beta_{rs} k_s) \tag{5.40}$$

موجود می باشد .

توجه : اندیس n برای مشخص ساختن تغییرپذیری اندازه گام ، درصورت لزوم ، در هر گام می باشد .

یکی از مشکلات ناشی از کاربرد فرمول رونگه ــ کوتا عبارتستاز اینکه جمله خطا خیلی پیچیده است و برای تخمین خطای برشی یا بیان اندازه گام به سهولت نمی‌توان آنرا به‌کار برد . بهرحال ، روشی که برای به‌کار بردن تقریب ساده به‌منظور رسیدن به تقریبی از خطا که بتواند برای تغییر دادن اندازه گام به‌کار رود طراحی شده است . این روش، که به روش رونگه ــ کوتا ــ مرسن Runge–Kutta–Merson معروف است ، یک محاسبه مشتق اضافی به‌کار برده و زمانیکه این تابع مشتق نسبت به x و y خطی باشد یک تخمین خطای دقیق تولید می‌کند . اگر اندازهٔ گام به‌اندازه کافی کوچک باشد خطی بودن مشتق تقریب منطقی خواهد بود .

معادلات روش رونگه ــ کوتا ــ مرسن عبارتستاز :

$$\begin{aligned}
k_1 &= hf(x_n, y_n)\\
k_2 &= hf(x_n + h/3, y_n + k_1/3)\\
k_3 &= hf(x_n + h/3, y_n + k_1/6 + k_2/6)\\
k_4 &= hf(x_n + h/2, y_n + k_1/8 + 3k_3/8)\\
k_5 &= hf(x_n + h, y_n + k_1/2 - 3k_3/2 + 2k_4)
\end{aligned} \tag{5.41}$$

$$y_{n+1} = y_n + \tfrac{1}{6}[k_1 + 4k_4 + k_5] \tag{5.42}$$

بایستی توجه داشت که هردوی این فرمولها و معادله قبلی (5.38) دارای یک جمله خطا از مرتبه h^5 بوده ، اما دراین حالت یک مشتق اضافی باید محاسبه شود . تخمین خطا وسیله

$$E \approx \tfrac{1}{30}[2k_1 - 9k_3 + 8k_4 - k_5] \tag{5.43}$$

داده می‌شود . درصورتیکه $f(x, y) = Ax + By + C$ این تخمین برابر مقدار دقت خواهد بود .

دقت یک فرمول معمولا" به‌صورت جملاتی از بسط سری تیلر تعریف می‌شود ، مانند فرمول پیشگو ــ اصلاحگر ، اما در این حالت به بسطی از جملات دو متغیر x و y نیاز می‌باشد . مرتبه دقت دوباره وسیله اولین توان h غیر صفر در بسط تعیین می‌گردد . تفصیلات این محاسبات در Conte (1965) و Bull (1966) موجود می‌باشد .

۵.۴.۲ ■ پایداری

خواص پایداری فرمول رونگه ــ کوتا با تعریف قبلی نسبتا" متفاوت است ، زیرا

روشهای رونگه‌ـ کوتا جوابهای کاذب تولید نمی‌کند . در روشهای چندگامی این جوابهای کاذب باید به‌صورت یک مقدار قابل اغماض برای جواب باشد ، اما این مسئله در روشهای رونگه ــ کوتا به‌سادگی دقت تفاضل ــ متناهی است که جواب تقریبی از جواب واقعی را پیدا می‌کند . ناپایداری به‌صورت ناپایداری جزئی نامیده می‌شود . زیرا کاهش در اندازه h ناپایداری را حذف خواهد کرد ، اگر چه ممکن است کاهش h برای اندازه مورد نیاز اقتصادی نباشد . زمانی این ناپایداری خیلی واضح می‌گردد که جواب واقعی‌با تابع نمائی کاهش یابد . این‌عمل منجر به تبدیل یک جواب تفاضل ــ متناهی افزایشی با انتخاب اندازه خیلی بزرگ از h می‌گردد . معمولا" محدودیت h شکل زیر را دارد ،

$$h\left|\frac{\partial f}{\partial y}\right| < K$$

که درآن K مقداری است بین $2{\cdot}0$ و $2{\cdot}8$. این ناپایداری زمانی یک مسئله خواهد شد که یک جزء غالب را ایجاب کند و این موجب تغییر یک جزء توانی نزولی به یک جواب تفاضل ــ متناهی افزایشی شود . همچنین زمانیکه حل مجموعه‌ای از معادلات موردنظر باشدناپایداری مورد توجه قرار می‌گیرد زیرا باید معیار پایداری جزئی برای هر جزء صادق گردد ، که می‌تواند یک ضرورت برای یک اندازه گام خیلی کوچک ایجاد نماید . باید توجه داشت که روش رونگه ــ کوتا ــ مرسن یک اندازه گام حتی کوچکتر از معیار پایداری ــ جزئی تعریف شده در فوق را خواهد داد .این مطلب درقسمت(۵۰۵۰۴) با تفصیل بیشتر بحث شده است .

۵.٤.۳ ■ توسعه به دستگاه معادلات دیفرانسیل

توسعهٔ روشهای رونگه ــ کوتا به دستگاه معادلات کاملا" ساده بوده و معادلات مورد بحث در اینجا نشان داده شده است . درصورتیکه مسئله

$$y' = f(x, y, z) \qquad z' = g(x, y, z)$$
$$y_0 = s_1 \qquad z_0 = s_2 \tag{5.44}$$

را درنظر بگیریم ، آنگاه طرح محاسبه‌ای زیر می‌تواند به‌کار رود .

$$\begin{aligned} k_1 &= hf(x_n, y_n, z_n) \qquad m_1 = hg(x_n, y_n, z_n) \\ k_2 &= hf(x_n + h/2, y_n + k_1/2, z_n + m_1/2) \\ &\qquad\qquad m_2 = hg(x_n + h/2, y_n + k_1/2, z_n + m_1/2) \end{aligned} \tag{5.45}$$

و الی آخر

$$y_{n+1} = y_n + \tfrac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4] \tag{5.46a}$$
$$z_{n+1} = z_n + \tfrac{1}{6}[m_1 + 2m_2 + 2m_3 + m_4] \tag{5.46b}$$

فقط نکته‌ای که باید توجه داشت آن‌است که باید k_1 و m_1 هر دو قبل از k_2 و m_2 محاسبه شوند. توسعه طرح فوق برای چندین متغیر ساده بوده و بسهولت قابل برنامه‌نویسی در کامپیوتر می‌باشد.

۵.۵ ■ پایداری

۵.۵.۱ ■ ناپایداری ذاتی والقائی

زمانیکه محاسبات کامپیوتری به‌صورت گام به گام باشند، بررسی نه‌تنها اندازه خطاهای ایجاد شده در هرگام، بلکه طریقی که این خطاها در سراسر روند بزرگ می‌شوند نیز مهم می‌باشد. مطالعه پایداری با خاصیت اخیر مورد نظر خواهد بود. عبارت پایداری، وسیله راههای نسبتا" مختلفی، توسط مولفه‌های متعددی تعریف شده است. بنابراین در درک معنی پایداری باید دقت شود.

دو نوع ناپایداری موجود است که باید از یکدیگر تمیز داده شود، که عبارتنداز: ناپایداری ذاتی و ناپایداری القائی. در ناپایداری ذاتی تغییرات کوچک در شرایط مسئله باعث تغییرات بزرگ در جواب واقعی شده و این ناپایداری باید به‌طور اجتناب‌ناپذیر در طرح تفاضل ــ متناهی اثر کند. ناپایداری ذاتی به خطای بزرگی وابسته می‌گردد که وسیله روش حل ایجاد شده و اگر رشد خطا بدتر از رشد جواب واقعی نباشد ممکن است جدی‌نباشد.

تشخیص بین ناپایداری روشهای چندگامی زمانیکه یکی از جوابهای کاذب بسرعت افزایش پیدا کرده ونسبت به جواب واقعی غالب گردد و ناپایداری جزئی روشهای رونگه ــ کوتا که جواب تفاضل ــ متناهی جواب واقعی را به‌طور نامناسب معرفی می‌کند (زیرا گام خیلی بزرگ به‌کار رفته است) مهم می‌باشد.

زمانیکه جواب واقعی صعودی باشد باید مفاهیم پایداری خیلی دقیق فرمول‌بندی شود. در این حالت امکان افزایش این چنین خطاها نیز هست و خاصیتی که در اینجا لازم است عبارتست از اینکه خطاها از جواب واقعی سریعتر افزایش پیدا نکند. طرحی نسبتا" پایدار نامیده می‌شود اگر خاصیت اخیر برقرار باشد، بنابراین اگر خطای مطلق بزرگ شود خطای نسبی افزایش نمی‌یابد. تاکنون، روندهای متعارفی که قادر به آزمون پایداری باشد در دسترس نیست لذا این مطلب از نقطه‌نظر تئوری نسبت به عملی با اهمیت‌تر می‌باشد.

۵.۵.۲ ■ ناپایداری ذاتی

این نوع ناپایداری زمانی رخ می‌دهد که تغییرات‌کوچک درشرایط مسئله‌دیفرانسیلی

تعریف شده منجر به تغییر بزرگی در جواب معادله دیفرانسیل گردد و ناپایداری عامل اساسی در مسئله فرمول‌بندی باشد. ازآنجائیکه طرح تفاضل ــ متناهی تا حدامکان از خواص معادله دیفرانسیل منتج می‌گردد، همچنین یک تغییر کوچک منجر به افزایش خطای برشی یا خطای ناشی‌از گرد شدن می‌شود، این مشگل به‌طور ذاتی در فرمول‌بندی مسئله موجود می‌باشد و تنها راه حذف ناپایداری عبارتست‌از سعی در دوباره‌فرمول‌بندی کردن مسئله طوریکه حساسیت قوی نسبت به تغییرات مختصر ازبین برود.

مثالی از این ناپایداری را می‌توان زمانیکه جواب معادله دیفرانسیل شامل یک جمله غالب است که وسیله مجموعه ویژه‌ای از شرایط آغازی قابل حذف می‌باشد بررسی کرد. معادله زیر را درنظر می‌گیریم:

$$y' = y - x \tag{5.47}$$

که دارای جواب درست

$$y = Ae^x + x + 1 \tag{5.48}$$

می‌باشد.

درصورتیکه شرایط آغازی $y=1$ وقتی $x=0$ باشد آنگاه داریم $A=0$ وجزء غالب جواب که e^x است به‌طورتئوری ازبین می‌رود. بهرحال، خطاهای کوچک درطرح تفاضل متناهی این مولفه را معرفی خواهد کرد و به‌دلیل رشد سریع خود بزودی در جواب واقعی ادغام خواهد شد.

۵.۵.۳ ■ ناپایداری قوی وپایداری ضعیف

مسائل مخصوصی با روشهای چند گامی وجود دارد از آنجائیکه‌به‌کار بردن روشهای بامرتبه بالاتر و دقت بیشتربه‌طور خودکار جوابهای اضافی معرفی می‌کند و تشخیص این جوابهای کاذب بدخیم می‌گردد. به‌طور ایده‌آل این جوابها کاذب بایستی سریعا" ازبین بروند یا در کمترین مقدار خود باشند. یا اینکه افزایش پیدا نکنند.

بدترین حالت وقتی است که در شیوه‌ای با کاهش اندازه h نتوان تصحیح را انجام داد و یکی از جوابهای کاذب افزایش پیدا کند، در حقیقت جواب در یک مثال قویا" ناپایدار زمانیکه h بسمت صفر میل کند بدون کران افزایش پیدا می‌کند. مثالی از یک طرح قویا" ناپایدار در بخش ۵.۳.۲ آورده شده است. چنین طرحهائی غیرقابل قبول میباشند اما ممکن است طرحهایی که قویا" ناپایدار هستند برای دسته‌ای مسائل نامناسب باشند. روشهای به‌طور ضعیف پایدار برای بعضی مسائل دارای جوابهای کاذب افزایشی می‌باشند و زمانی که روشهای به‌طور ضعیف پایدار به‌کار می‌روند بایستی توجه بیشتری شود. این خواص ازنظر ریاضی به‌شرح زیر تعریف می‌گردد.

معادله دیفرانسیل

$$
\begin{aligned}
y' &= f(x,y) \\
y(a) &= s
\end{aligned}
\tag{5.49}
$$

را درنظر گرفته و معادله تفاضل ــ متناهی زیر را به‌کار می‌بریم .

$$\sum_{r=0}^{k} \alpha_r y_{n+r} - h \sum_{r=0}^{k} \beta_r f_{n+r} = 0 \tag{5.50}$$

اگر $k > 1$ آنگاه طرح تفاضلی دارای جوابهای بیشتری از طرح دیفرانسیلی می‌گردد و ناپایداری با روشهای چندگامی که شرکت دارند ممکن است ظاهر گردد . تجزیه و تحلیل این روشها با دقت کامل ریاضی به‌وسیله Dahlquist (1956) و Henrici (1962) آورده شده است . در این جا معرفی نتایج و بررسی آن بدون دلیل کفایت خواهد شد . معادله تفاضلی (5.50) که برای y_{n+k} حل می‌گردد یک معادله‌غیرخطی با ضرائب ثابت می‌باشد . برای سهولت مفهوم معادلات چند جمله‌ای

$$\rho(z) = \sum_{r=0}^{k} \alpha_r z^r \tag{5.51}$$

و

$$\sigma(z) = \sum_{r=0}^{k} \beta_r z^r \tag{5.52}$$

معرفی شده‌اند . درصورتیکه به تمام طرحهای تفاضلی که برای y ثابت پایدار می‌گردند تکیه کنیم یعنی $f(x,y) \equiv 0$. آنگاه تنها جزء خطی طرح تفاضلی باقی می‌ماند ، و جوابها به‌وسیله ریشه‌های معادله کمکی $\rho(z) = 0$ داده می‌شود . درصورتیکه ریشه‌های این معادله z_r $(r = 1, 2, \ldots, k)$ باشند آنگاه اگر تمام ریشه‌هامتمایز باشند کلیه جوابهاعبارتنداز :

$$y_n = A_1(z_1)^n + \cdots + A_k(z_k)^n \tag{5.53}$$

همیشه یک ریشه اصلی $z_1 = 1$ که متناظر با جواب واقعی است وجود دارد و زمانی که هر ریشه $z_r (r \neq 1)$ دارای قدر مطلق بزرگتر از واحد باشد (ناپایداری قوی) ظاهر می‌گردد . می‌توان دید که زمانی که n افزایش یافته و طرح را بی‌فایده سازد چنین ریشه‌ای بدون کران افزایش حاصل می‌کند . همچنین نشان داده می‌شود که درصورتی که ریشه‌های تکراری z_r با $|z_r| = 1$ باشد این رشد بدون کران اتفاق می‌افتد . اگر تمام ریشه‌ها در شرط $|z_r| \leqslant 1$ و ریشه‌های چندگانه در $|z_r| < 1$ صدق نمایند ، آنگاه ناپایداری قوی نمی‌تواند پیش آید ، اما ممکن است روش از خاصیتی موسوم به پایداری ضعیف تبعیت نماید . این مشکل در طرحهای تفاضل ــ متناهی برای کلاسهای مخصوصی از مسائل رخ می‌دهد . یک طرح به‌طور ضعیف پایدار ممکن است برای مشتق معین توابع کاملا" صادق

باشد اما در سایر مسائل نتایج نامرغوب داشته باشد، زیرا معادله تفاضل ـ متناهی کامل دارای ریشه‌هائی است که با ریشه‌هائی که بوسیله معادله کمکی داده می‌شود فرق دارد، چون جملات $h\beta_r f_{n+r}$ نیز باید درنظر گرفته شود. ممکن است یک یا بیشتر ریشه‌های این معادله کاملا" خارج دایره به‌شعاع واحد قرار گیرد و منجر به یک افزایش جواب کاذب شده و سبب پایداری ضعیف گردد.

این رفتار به‌ویژه زمانی که جواب واقعی قویا" نزولی باشد آشکار می‌گردد و نظیر چنین جواب، معادله زیر

$$\begin{aligned} y' &= -y \\ y(0) &= 1 \end{aligned} \tag{5.54}$$

است که دارای جواب واقعی $y = e^{-x}$ می‌باشد. برای مثال، اصلاحگر سیمپسون

$$y_{n+2} = y_n + \frac{h}{3}[f_{n+2} + 4f_{n+1} + f_n] \tag{5.55}$$

دارای جواب

$$y_n = A_1[z_1]^n + A_2[z_2]^n \tag{5.56}$$

بوده و z_r تقریبهایی برای e^{-h} و $-e^{h/3}$ می‌باشد. درصورتیکه مقادیر آغازی دقیق باشند مطمئن می‌سازد که A_1 تقریبا" برابر واحد و A_2 تقریبا" برابر صفر می‌باشد، بنابراین بدوا مولفه دوم کوچک خواهد شد، هرچند که این مولفه به‌طور توانی افزایش پیدا می‌کند و تدریجا" به جواب واقعی غالب می‌گردد. اندازه مولفهٔ دوم به اندازه A_2 و حاصلضرب hn بستگی دارد. بدین‌ترتیب اثـر پایداری ضعیف بدو طریق می‌توانـد کاهش یابد. اگر شرایط آغازی خیلی دقیق درنظر گرفته شود، برای مثال با کاهش طول گام h، آنگاه A_2 خیلی کوچک خواهد شد. درصورتی‌که طول فاصله انتگرال‌گیری کوچک باشد آنگاه hn نیز کوچک شده بنابراین $e^{hn/3}$ کوچک باقی می‌ماند. بدیهی است که طرحهای به‌طور ضعیف پایدار برای انتگرال‌گیری روی طول فواصل مناسب نمی‌باشد. Dahlquist چندین خاصیت مهم از طرح‌های چند گامی را به شرح زیر به ثبوت رسانیده است.

۱ــ درصورتی که طرح پایدار باشد رتبه دقت یک روش چندگامی بایستی محدود شود. از آنجائیکه تعداد $2k+2$ ضریب در معادله (5.11) موجود است به‌طور تئوری امکان رسیدن به دقتی از رتبه $2k+1$ موجود می‌باشد. بهرحال، رتبه یک عملگر پایدار که در آن k فرد است نمی‌تواند از $k+1$ بیشتر گردد. اگر k زوج باشد، رتبه عملگر پایدار نمی‌تواند از $k+2$ بیشتر شود.

۲ــ شرایط لازم و کافی برای رسیدن به ماکزیمم رتبه دقت $k+2$، که پایداری را تامین

کند ، عبارتست از اینکه k زوج بوده و تمام ریشه‌های $\rho(z)$ دارای قدرمطلق واحد باشد . و روابط زیر بین ضرائب $\rho(z)$ و $\sigma(z)$ برقرار گردد .

$$\alpha_r = -\alpha_{k-r}, \qquad \beta_r = \beta_{k-r} \tag{5.57}$$

۳ــ فرمول پایدار با ماکزیمم رتبه دقت همواره دارای خاصیت پایداری ضعیف خواهدبود .

۵.۵.۴ ■ پایداری جزئی

پایداری جزئی عبارتی برای بیان پدیده‌ای است که درآن اندازهٔ گام h بسیار بزرگ سبب معرفی یک جواب نزولی وسیله جواب تفاضل ــ متناهی صعودی می‌گردد . این عبارت کرارا" در ارتباط با روشهای رونگه ــ کوتا به‌کار می‌رود و ناپایداری به ویژه زمانی آشکار می‌گردد که این روشها برای جواب معادلات سخت stiff به‌کار روند ، از آنجائیکه آن زمانی رخ می‌دهد که بیشترین شرط مبرم در h به‌وسیله کمترین مولفه معنی‌دار جواب باشد ، جمله معادله سخت زمانی که مولفه‌های یک جواب در میزانهای با تفاوت فاحش کاهش می‌یابد به‌کار می‌رود ، برای مثال ،درمعادلات (5.63)-(5.66) .این نوع ناپایداری واقعی نمی‌باشد چون از یکی از مولفه‌های جوابهای غیر دقیق بدست می‌آید نه از بزرگ‌سازی خطاها .

مثالهای زیر نشان می‌دهد که چطور این نوع ناپایداری ظاهر می‌گردد . روش رونگه ــ کوتا از رتبه دوم را درنظر گرفته

$$\begin{aligned} k_1 &= hf(x_n, y_n) \\ k_2 &= hf(x_n + h, y_n + k_1) \end{aligned} \tag{5.58}$$

$$y_{n+1} = y_n + \tfrac{1}{2}[k_1 + k_2] \tag{5.59}$$

و برای معادله دیفرانسیل رتبه ــ اول

$$\begin{aligned} y' &= \lambda y \\ y(0) &= A \end{aligned} \tag{5.60}$$

که دارای جواب $y = Ae^{\lambda x}$ است به‌کار می‌بریم . فرمول رونگه ــ کوتارابطه زیررامی‌دهد .

$$y_{n+1} = y_n\left[1 + h\lambda + \frac{h^2\lambda^2}{2}\right] \equiv y_n\phi_n \tag{5.61}$$

ونمودار شکل ۵.۳ نشان می‌دهد که چطور این عامل برای مقادیر $\lambda = \pm 1$ تغییرمی‌کند .

برای $\lambda = 1$ ملاحظه می‌شود که طرح تفاضل ــ متناهی جواب را در هر گام به‌منظور تعقیب جواب واقعی افزایش داده و بهرحال بدین معنی است که خطا همچنان بزرگ می‌گردد . این ناپایداری نمی‌تواند جزء ردهٔ ناپایداری واقعی قرار گیرد زیرا این خاصیت جزء ذات معادله دیفرانسیل می‌باشد . درچنین حالتی لازم است طرح نسبتا"

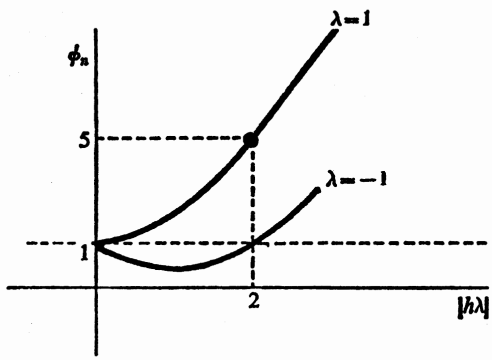
پایدار باشد . بدین‌معنی ، که خطاها در طرح تفاضلی از جواب واقعی سریعتر رشد پیدا نکنند اما این موضوع مشکلی است که در اینجا دنبال نخواهد شد .

برای $\lambda=-1$ به طرحی که نسبتا" پایدار باشد برای کاهش به همان میزان e^{-x} نیازمند می‌باشیم ، ملاحظه می‌شود که این فقط در ناحیه کوچکی از $h\lambda$ روی خواهد داد . در حقیقت ، برای $h\lambda>2$ خطاها بزرگ خواهند شد و به این پدیده ، ناپایداری جزئی گفته می‌شود . بهرحال ، جائیکه مولفه نمائی قسمت با ارزش جواب است ، مانند $h\lambda<2$ شرطی به اندازه کافی دقیق برای رسیدن به دقت مناسب نخواهد بود .

برای مجموعه معادلات با جوابهای نمائی کاهشی با مقادیر متفاوتی از λ مسئله فوق توجه ویژه‌ای را به‌خود اختصاص داده است . این نوع معادلات به معادلات سخت stiff موسوم می‌باشند . در این حالت برای شرط

$$|h\lambda_i|<2 \tag{5.62}$$

لازم است که تمام مقادیر λ_i را به‌کار برد و درغیر این‌صورت‌یکی از جوابهای کاهشی به کمک یک جواب تفاضل ــ متناهی افزایشی تقریب زده شده است که فورا" به سایر جوابها غالب خواهد گردید .



شکل (۵.۳) ــ نمودار رونگه ــ کوتا عامل ϕ_n

روش‌های رونگه ــ کوتا برای این نوع مسئله مناسب نیستند و روشی مانند گیر Gear بایستی به‌کار رود . این روش در قسمت ۵.۶ بحث شده است .

یک معادله سخت ساده این مسئله را نشان خواهد داد. معادله دیفرانسیل

$$y''+101y'+100y=0 \tag{5.63}$$

به‌وسیله مجموعه معادلات

$$\begin{aligned} y' &= z \\ z' &= -(100y+101z) \end{aligned} \tag{5.64}$$

قابل بیان بوده، که‌درآن y' مشتق نسبت به x می‌باشد. درصورتیکه شرایط آغازی به صورت

$$\begin{aligned} y(0) &= 2 \\ z(0) &= -101 \end{aligned} \tag{5.65}$$

باشد، آنگاه جواب واقعی بوسیله

$$y = e^{-x} + e^{-100x} \tag{5.66}$$

بدست می‌آید. زمانی که $x = 1$ است اولین قسمت جراب تقریبا" معادل 0·368 و دومین قسمت کمتر از 10^{-8} می‌باشد. آزمون بیانگر این است که دومین قسمت جواب قابل اغماض است، بنابراین صرفنظر می‌گردد. اگرچه، درصورتی‌که از شرط پایداری ــ جزئی صرفنظر شود آنگاه دومین قسمت جواب شروع به افزایش کرده و بزودی از اولین قسمت بزرگترخواهد گردید. بنابراین، اساسا" شرط (5.62) برای تمام مقادیر λ_i برقرار بوده و این محدودیت جدی در اندازه h را می‌تواند ایجاد نماید.

۵۰۶ ■ روشهای پیشرفته

بدلیل محاسبه اضافی که از به‌کار بردن روشهای رونگه ــ کوتا نسبت به روشهای پیشگو ــ اصلاحگر نتیجه می‌گردد. ضروری است‌نوعی روش پیشگو ــ اصلاحگر که برای ساختن روندهائی که بستگی به دو مسئله شروع جواب و تغییر اندازه گام داشته موجود باشد. Nordsieck (1962) یک برنامه کامپیوتری توصیف‌کرده که دارای این دو خاصیت می‌باشد و برای تعدیل اندازه گام دو آزمون به‌کار می‌برد، یکی برای کنترل خطای برشی و دیگری برای اطمینان از پایداری روش، روش اساسا" یک فرمول‌بندی مجدد از فرمول تفاضل ــ متناهی Adams بوده که برای $h\lambda$ کوچک دارای خاصیت پایداری خوبی می‌باشد. جملاتیکه ذخیره می‌گردند عبارتنداز بسط تیلر. قوت این روش عبارتست‌از اینکه جائیکه مشتقات بزرگ وجود دارند عضوی از مقیاس‌گیری جایگزین توانی از h ذخیره شده می‌گردد. خواننده علاقمند برای تفصیل بیشتر بایستی به مقاله مراجعه کند.

اگرچه توضیحات روشن ممکن است برای بعضی‌ها جالب نباشد اما برنامه کامپیوتری نوشته شده‌ای از روش موجود بوده و دراین صورت برای مقصود کلی نسبت به روش رونگه ــ کوتا ــ مرسن که وقت کمتری را میگیرد توصیه می‌شود. اگرچه مسائلی وجود دارند که شامل معادلات سخت بوده، به‌دلیل به‌کارگیری اندازه‌گام کوچک که باید به‌طور خودکار در روند انتخاب گردد، روش فوق‌ایده‌آل نمی‌باشد. روشی منسوب به Gear (1968) با شرکت یک روند آغازی خودکار موجود بوده که مقدار قابل‌توجهی از زمان مورد نیاز برای جواب معادلات سخت را کم میکند. همچنین این روش جواب

تکراری خیلی پیچیده‌ای از معادله اصلاحگر همراه دارد اما این جواب تاثیر چندانی در اساس روش برای معادلات سخت نخواهد داشت.

مطالعه روش گیر Gear جالب توجه بوده، چون، اگرچه چنین بخشی برای کسی که روشهای عددی را به‌کار می‌برد اساسی نمی‌باشد، یک دید مفید از مسائل مربوط به انتخاب روش می‌نیممی برای معادلات ناسخت. non-stiff را ارائه می‌دهد.

$$\sum_{r=0}^{k} \alpha_r y_{n+r} - h \sum_{r=0}^{k} \beta_r f_{n+r} = 0 \tag{5.67}$$

ومعادله‌متناظربه‌وسیله‌جایگزینی‌مقادیرجواب $y(x_{n+r})$ برای y_{n+r} و $f(x_{n+r}, y(x_{n+r}))$ برای f_{n+r} به‌دست می‌آید. یعنی

$$\sum_{r=0}^{k} \alpha_r y(x_{n+r}) - h \sum_{r=0}^{k} \beta_r f(x_{n+r}, y(x_{n+r})) = E_{n+k} \tag{5.68}$$

پس از تفریق اولین معادله از دومین و به‌کار بردن قضیه میانگین داریم.

$$f(x_{n+r}, y(x_{n+r})) - f(x_{n+r}, y_{n+r}) = [y(x_{n+r}) - y_{n+r}] \frac{\partial f}{\partial y}(x_{n+r}, \zeta) \tag{5.69}$$

که در آن ζ در فاصله داده شده وسیله نقاط انتهائی y_{n+r}, و $y(x_{n+r})$ قرار دارد. معادله‌اخیر معادله

$$\sum_{r=0}^{k} \alpha_r e_{n+r} - h \sum_{r=0}^{k} \beta_r \lambda(x_{n+r}, \zeta) e_{n+r} = E \tag{5.70}$$

را می‌دهد، که درآن

$$e_{n+r} = y(x_{n+r}) - y_{n+r} \qquad \text{و} \qquad \lambda \equiv \frac{\partial f}{\partial y}$$

مطالعه پایداری، با مفهوم Dahlquist ، برای ریشه‌های چند جمله‌ای کمکی تشکیل شده با معادله تفاضلی در طرف سمت چپ معادله (5.70) مورد توجه می‌باشد، مانند

$$\sum_{r=0}^{k} \alpha_r z^r - h \sum_{r=0}^{k} \beta_r \lambda(x_{n+r}, \zeta) z^r = 0 \tag{5.71}$$

زمانی که $h=0$ یک ریشه این معادله $z=1$ می‌گردد اما، بجز این ریشه، برای‌پایداری لازم است که قدر مطلق تمام ریشه‌ها از واحد کمتر باشد همچنین امید می‌رود که این شرط برای مقدار بزرگ h تا حدامکان برقرار گردد.

زمانی که مقدار $h\lambda$ خیلی کوچک و قابل اغماض یا خیلی بزرگ باشد دو مشکل می‌تواند پیش آید. اگر $h\lambda$ خیلی کوچک باشد آنگاه تمام ریشه‌ها از تغییرات کوچک ریشه‌های $\rho(z) \equiv \sum_{r=0}^{k} \alpha_r z^r = 0$ بدست می‌آید و دراین صورت منطقی به‌نظر می‌رسد که تمام ریشه‌ها را مساوی صفر گرفت تا مساوی ریشه اصلی $z=1$. درصورتیکه $h\lambda$ بزرگ شود امکان اینکه ریشه‌ها به سرعت از واحد بزرگتر نگردند وجود دارد، زیرا این ریشه‌ها

زمانی که $h=0$ ازصفر شروع می‌شوند . این معادله $\rho(z)=z^k-z^{k-1}$ رامی‌دهدکه‌مربوط به روش Nordsieck است . حال ، چون h معمولا" بدلیل وابسته بودن به پایداری و دقت ، تاحدامکان بزرگ انتخاب می‌گردد ، از آنجا نتیجه می‌گردد ، زمانی که $h\lambda$ خیلی کوچک باشد λ نیز خیلی کوچک گردد ، و زمانی که چندین متغیر موردنظر باشد بایستی تمام λ_i ها کوچک باشند .

اگرچه ، برای معادلات‌سخت معمولا" حداقل یکی از λ_i ها بزرگ است و توجه به معادله (5.67) زمانی که $h\lambda_i$ بزرگ باشد حائز اهمیت می‌باشد . درصورتیکه λ_i ثابت باشد ریشه‌ها تغییرات کوچکی از

$$\sigma(z)\equiv\sum_{r=0}^{k}\beta_r z^r=0 \tag{5.72}$$

خواهد بود و منطقی به‌نظر می‌رسد که معادله انتخاب شده z^k باشد که دراین صورت تمام ریشه‌ها مساوی صفر می‌گردد . همچنین بایستی ریشه اساسی $z_1=1$ از چند جمله‌ای $\rho(z)$ وجود داشته باشد ، این شرایط به‌کار رفته وسیله Gear را تعریف می‌نماید . روش Gear مبتنی براین معادله کمکی است و بنابراین ، احتمالا" نتایج خوبی برای معادلات سخت خواهد داد ، و این محققا" با نتایج عملی تحقق پیدا می‌کند .

بایستی به‌خاطر داشت که دلیل منطقی ، نظیر روش فوق یک بینش مفید در پیش نیاز روشها را می‌دهد اما نتایج مستدل تنها می‌تواند مبتنی بر آنالیز ریاضی قوی باشد . بویژه بایستی توجه داشت که برای مقادیر کلی $h\lambda$ نه $\rho(z)$ نه $\sigma(z)$ غالب می‌گردند و انتخاب یک روش با بیشترین پایداری روشن نمی‌شود .

۵.۷ ■ مقایسه روش‌ها

غالبا" روشهای رونگه ــ کوتا برای جواب مسائل مهندسی در کامپیوتر به‌کار می‌روند زیرا آنها به روش آغازی ویژه‌ٔی نیاز ندارند و بنابراین برنامه‌نویسی آنها خیلی ساده می‌باشد . بهرحال ، سه نقطه ضعف روشهای رونگه ــ کوتا قابل اغماض نمی‌باشد ، نظریه اینکه دو سیمای اشاره شده روش درفوق . با برنامه‌نویسی دقیقی به‌صورت روشهای چندگامی می‌تواند معرفی گردد ، مانند روشهای ، Gear و Nordsieck ، از این شرایط نتیجه می‌گردد که درحالت کلی روشهای چندگامی ترجیح داده می‌شود ، تنها باستثناء زمانی که سهولت روش خیلی مهم باشد .

تقریبا" تنها نقطه ضعف مهم روشهای رونگه ــ کوتا عبارتست از اینکه شکل جمله خطا فوق‌العاده پیچیده است ، و این مشکل کاملا" با روش مرسن برطرف نمی‌گردد ، در حالیکه فرمول ساده‌ای برای خطا در روشهای چندگامی موجود است ، و این درصورتی است که مشتق مناسب قابل محاسبه باشد . سیمای نامطلوب دیگر روشهای رونگه ــ کوتا

عبارتست‌ازاینکه تعداد بسیاری محاسبه مشتق برای هرگام موردنیاز می‌باشد . روش رونگه ــ کوتا ــ مرسن نیازمند پنج محاسبه از نوع فوق بوده ، درحالیکه در روش پیشگوــ اصلاحگر اغلب دو یا سه اصلاح کافی‌می‌باشد . این مطلب زمان لازم برای کامپیوتر را در روشهای چندگامی کم کرده و در حالیکه تابع مشتق شامل محاسبه توابع ویژه یا محاسبه پیچیده‌ای باشد این زمان می‌تواند قابل توجه باشد .

سومین مسئله‌ای که در کاربرد روشهای رونگه ــ کوتا شرکت دارد در قسمت مربوط به پایداری بحث گردید که در آن نشان داده شده است که چطور روشهای رونگه ــ کوتا برای معادلات سخت کافی نمی‌باشد . با مقایسه روش Gear مشخص می‌گردد که زمان قابل توجهی نسبت به روش قبلی از زمان کامپیوتر کاسته می‌شود . برای انتخاب یکی از روشهای Gear و Nordsieck از اطلاعات مربوط به مقادیر λ_i ومقادیر ویژه ژاکوبین $\partial f/\partial y$ یک دستگاه معادلات استفاده می‌شود . درحالتیکه مقدار λ_i بزرگ باشد روش Gear توصیه می‌گردد ، و زمانیکه تمام λ_i کوچک باشد روش Nordsieck ترجیح‌داده می‌شود . توصیه‌های فوق تنها مبتنی بر تجزیه هستند چون قبل از رسیدن به نتایج مستدل ، تجزیه و تحلیل کامل مسئله موردنظر ضروری است .

جدول ۵۰۱

روشهای نیوتن ــ کاتس

باز

$y_{n+1}=y_{n-1}+2hy'_n$ Midpoint

$y_{n+1}=y_{n-2}+\frac{3h}{2}[y'_n+y'_{n-1}]$

$y_{n+1}=y_{n-3}+\frac{4h}{3}[2y'_n-y'_{n-1}+2y'_{n-2}]$

بسته

$y_{n+1}=y_n+\frac{h}{2}[y'_{n+1}+y'_n]$ trapezoidal

$y_{n+1}=y_{n-1}+\frac{h}{3}[y'_{n+1}+4y'_n+y'_{n-1}]$ Simpson

$y_{n+1}=y_{n-2}+\frac{3h}{8}[y'_{n+1}+3y'_n+3y'_{n-1}+y'_{n-2}]$

روشهای نوع ــ پیشگو باز مبتنی بر انتگرال‌گیری فرمول تفاضل پسرو نیوتن است . معادلات پیشگو از نوع بسته با همان روش با مقدار متفاوت p نتیجه می‌گردد .

جدول ۵۰۲

روشهای آدامز ــ بشفورث

Adams–Bashforth methods

باز

$$\nabla y_{n+1} = h(1 + \tfrac{1}{2}\nabla + \tfrac{5}{12}\nabla^2 + \cdots)y'_n$$

$$y_{n+1} = y_n + hy'_n \quad \text{Euler}$$

$$y_{n+1} = y_n + \frac{h}{2}[3y'_n - y'_{n-1}]$$

$$y_{n+1} = y_n + \frac{h}{12}[23y'_n - 16y'_{n-1} + 5y'_{n-2}]$$

بسته

$$\nabla y_{n+1} = h(1 - \tfrac{1}{2}\nabla - \tfrac{1}{12}\nabla^2 - \cdots)y'_{n+1}$$

$$y_{n+1} = y_n + hy'_{n+1}$$

$$y_{n+1} = y_n + \frac{h}{2}[y'_{n+1} + y'_n) \quad \text{trapezoidal}$$

$$y_{n+1} = y_n + \frac{h}{12}[5y'_{n+1} + 8y'_n - y'_{n-1}]$$

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ معادله زیر را به‌شکل متعارف معادله (5.1) تبدیل کنید .

$$a(x)y'' + b(x)y' + c(x)y = 0, \qquad y(0) = s_1, \qquad y'(0) = s_2$$

باقراردادن $y' = z$ و $y'' = z'$ و فرض $a(x) > 0$

$$z' = -\frac{b(x)}{a(x)}z - \frac{c(x)}{a(x)}y$$

بدین‌ترتیب

$$\begin{pmatrix} y \\ z \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -\frac{c(x)}{a(x)} & -\frac{b(x)}{a(x)} \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

$$\begin{pmatrix} y \\ z \end{pmatrix}_{x=0} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$$

حال مانند دستگاه دو معادله با دو مجهول مرتبه اول حل می‌شود .

هر معادله با مرتبه بالاتر که نسبت به مشتقات متعدد خطی باشد می‌تواند در این شکل‌برداری معرفی گردد .

۲ــ مطلوبست تعیین مقدار y در نقاط $x = 1{\cdot}0, 2{\cdot}0, 0{\cdot}5$ وسیله روش سری‌تیلربرای‌معادله

معادله $dy/dx = e^x$

$$y(0) = 1$$

ابتدا برای مشتقات متعدد صورتهای زیر را پیدا کنید .

$$\frac{d^2y}{dx^2}=e^x, \qquad \frac{d^3y}{dx^3}=e^x, \quad \text{etc.}$$

با شروع $x=0$ خواهیم داشت $y(0)=1$ و $y'(0)=1$ و $y''(0)=1$ و غیره با به کار بردن سری تیلر

$$y(x+h)=y(x)+hy'(x)+\frac{h^2y''}{2!}(x)+\cdots+\frac{+h^{n-1}}{(n-1)!}y^{(n-1)}(x)+\frac{h^n}{n!}y^{(n)}(x+\theta h),$$
$$0\leqslant\theta\leqslant 1$$

داریم

$$y(1\cdot0)=1+1+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\cdots\approx 2\cdot7$$
$$y(2\cdot0)=1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}+\cdots\approx 7\cdot0$$
$$y(0\cdot5)=1+\frac{1}{2}+\frac{1}{2!}\left(\frac{1}{2}\right)^2+\frac{1}{3}\left(\frac{1}{2}\right)^3+\frac{1}{4!}\left(\frac{1}{2}\right)^4+\cdots\approx 1\cdot65$$

جمله خطا قابل محاسبه می‌باشد و در این حالت ملاحظه کران فوقانی برای خطا ساده‌تر است زیرا e^x تابعی است افزایشی و بنابراین مقدار ماکزیمم بازاء $\theta=1$ رخ می‌دهد . حالت $y(2\cdot0)$ که در بالا محاسبه شده است را درنظر می‌گیریم . جمله خطا دارای یک مقدار ماکزیمم

$$\frac{h^n}{n!}e^h=\frac{2^5}{5!}e^2\approx 2\cdot0$$

و جملات خیلی بیشتری برای رسیدن به دقت منطقی موردنیاز است . همانطوریکه در کتاب پیشنهاد شده است راه کاهش تعداد جملات موردنیاز عبارتست‌از یافتن مقدار بازاء $x=2$ به وسیله کاربرد سری تیلر در گامهائی که برای مثال نقاط $0\cdot0$, $0\cdot5$, $1\cdot0$, $1\cdot5$ در تکرارهای متوالی به‌صورت نقاط اساسی برای بسط به‌کار روند .

۳ــ مطلوبست حل معادله زیر به روش اویلر در ناحیه $0\leqslant x\leqslant 2\cdot0$ با طول گام $h=0\cdot5$

$$\frac{dy}{dx}=y$$
$$y(0)=1$$

جواب واقعی عبارتست‌از : $y=e^x$

$$y_{n+1}=y_n+hy_n=(1+h)y_n$$
$$y_1=(1+h)\times 1=1\cdot5$$
$$y_2=(1+0\cdot5)\times 1\cdot5=2\cdot25$$
$$y_3=1\cdot5\times 2\cdot25=3\cdot375$$
$$y_4=1\cdot5\times 3\cdot375=5\cdot062$$

توجه خواهد شد که هر مقدار فقط به آخرین مقدار نیاز دارد . بنابر این روش آغازی ویژه‌ای مورد نیاز نمی‌باشد .

۴ـ مطلوبست حل معادله فوق به روش اویلر ـ ذوزنقه‌ای تنها با یک اصلاح. روند در شکل (۵۰۲) به کمک نمودار نشان داده شده است.

$$y_{n+1}^p = y_n + hf(x_n, y_n)$$
$$y_{n+1}^c = y_n + \frac{h}{2}[f(x_n, y_n) + f(x_{n+1}, y_{n+1}^p)]$$
$$y_1^p = (1+0{\cdot}5)1 = 1{\cdot}5 \qquad y_1^c = 1 + 0{\cdot}25[1 + 1{\cdot}5] = 1{\cdot}625$$
$$y_2^p = (1{\cdot}5)1{\cdot}625 \approx 2{\cdot}438 \qquad y_2^c = 1{\cdot}625 + 0{\cdot}25[1{\cdot}625 + 2{\cdot}438] = 2{\cdot}641$$
$$y_3^p = (1{\cdot}5)2{\cdot}641 \approx 3{\cdot}962 \qquad y_3^c = 2{\cdot}641 + 0{\cdot}25[2{\cdot}641 + 3{\cdot}962] = 4{\cdot}292$$
$$y_4^p = (1{\cdot}5)4{\cdot}292 \approx 6{\cdot}438 \qquad y_4^c = 4{\cdot}292 + 0{\cdot}25[4{\cdot}292 + 6{\cdot}438] = 6{\cdot}974$$

توجه، در مثالی که به روش اویلر حل شد خطا در y_4 تقریبا " $31{\cdot}5\%$ بوده درحالیکه در دومین حالت تقریبا " $5{\cdot}6\%$ می‌گردد.

۵ـ پایداری و دقت فرمول مبتنی بر قاعدهٔ سیمپسون را بررسی کنید.

$$y_{n+2} = y_n + \frac{h}{3}[f(x_{n+2}, y_{n+2}) + 4f(x_{n+1}, y_{n+1}) + f(x_n, y_n)]$$

پایداری قوی را درنظر بگیرید. معادله کمکی عبارتست از $z^2 - 1 = 0$ که‌دارای‌ریشه‌های $z = \pm 1$ می‌باشد بنابراین فرمول ناپایدار نیست. پایداری ضعیف تنها برای معادله ویژه‌ای بررسی میگردد.

معادله

$$y' = -\lambda y, \qquad \lambda > 0$$
$$y(0) = 1$$

را درنظر گرفته که دارای جواب $y = e^{-\lambda x}$ می‌باشد. حال معادله کمکی به‌صورت

$$z^2 - 1 - \frac{h}{3}[-\lambda z^2 - 4\lambda z - \lambda] = 0$$
$$z^2\left[1 + \frac{h\lambda}{3}\right] + z\left[\frac{4h\lambda}{3}\right] - 1 + \frac{h\lambda}{3} = 0$$

درآمده‌ریشه‌های این معادله عبارتست‌از:

$$z = \frac{\dfrac{-4h\lambda}{3} \pm \sqrt{\left[\dfrac{16h^2\lambda^2}{9} - 4\left(-1 + \dfrac{h^2\lambda^2}{9}\right)\right]}}{2\left(1 + \dfrac{h\lambda}{3}\right)}$$

$$z_1 = \left[1 - \frac{2h\lambda}{3} + \frac{h^2\lambda^2}{6} - \frac{h^4\lambda^4}{72}\cdots\right]\left[1 - \frac{h\lambda}{3} + \frac{h^2\lambda^2}{9} - \frac{h^3\lambda^3}{27} + \frac{h^4\lambda^4}{81}\cdots\right]$$
$$= 1 - h\lambda + \frac{h^2\lambda^2}{2} - \frac{h^3\lambda^3}{6} + \frac{h^4\lambda^4}{24}\cdots$$

که متناظر با بسط سری تیلر $e^{-h\lambda}$ تا جمله h^4 می‌باشد.

$$z_2 = \left[-1 - \frac{2h\lambda}{3} - \frac{h^2\lambda^2}{6} + \frac{h^4\lambda^4}{72}\cdots\right]\left[- \frac{h\lambda}{3} + \frac{h^2\lambda^2}{9} - \frac{h^3\lambda^3}{27} + \frac{h^4\lambda^4}{81}\cdots\right]$$
$$= -1 - \frac{h\lambda}{3} - \frac{h^2\lambda^2}{18} + \frac{h^3\lambda^3}{54} + \frac{5}{648}h^4\lambda^4\cdots$$

متناظربا بسط سری تیلر $e^{h\lambda/3}$ – تا جمله h^2 می‌گردد، بنابراین، برای مقادیر کوچک جمله اول جواب $e^{-\lambda x}$ را به‌طور تقریبی تعیین کرده و دومین جمله، جمله‌ایست که سریعاً" افزایشی‌است، بدین معنی فرمول پایداری ضعیف را ارائه‌می‌دهد. در مواقعی که جواب واقعی تابع نمائی افزایشی است، جواب کاذب کاهش یافته‌وبنابراین ازنظر قابل اغماض بودن با اهمیت خواهد بود. بدین‌ترتیب، می‌توان دید که پایداری ضعیف تنها یک مشکل برای معادلات معینی می‌باشد.

$$y(x_{n+2}) = y + 2hy' + \frac{4h^2}{2}y'' + \frac{8h^3}{6}3y''' + \frac{16h^4}{24}y'''' + \frac{32h^5}{120}y''''' + \cdots$$
$$f(x_{n+2}, y(x_{n+2})) = y' + 2hy'' + \frac{4h^2}{2}y''' + \frac{8h^3}{6}y'''' + \frac{16h^4}{24}y''''' + \cdots$$
$$f(x_{n+1}, y(x_{n+1})) = y' + hy'' + \frac{h^2}{2}y''' + \frac{h^3}{6}y'''' + \frac{h^4}{24}y''''' + \cdots$$
$$f(x_n, y(x_n)) = y'$$

جملات سمت راست بدون اندیس نمایش مقادیر واقعی بازاء x_n بوده، یعنی $y \equiv y(x_n)$ جمله خطا به‌وسیله

$$y(x_{n+2}) - y(x_n) - \frac{h}{3}[f(x_{n+2}, y(x_{n+2})) + 4f(x_{n+1}, y(x_{n+1})) + f(x_n, y(x_n))] = E_{n+2}$$

داده می‌شود.

در صورتیکه جملات متعدد تماما" بهم‌افزوده شوند جملات تا h^4 دارای ضریب صفر بوده و E_{n+2} دارای جمله‌ای به‌صورت $-h^5/90y''''' = -h^5/90f''''(x_n, y(x_n))$ می‌باشد.

۶ – فرمول رونگه – کوتا مرتبه‌چهارم متعارف را برای حل معادله زیر درناحیه $0 \leqslant x \leqslant 0{\cdot}2$ با طول گام $h = 0{\cdot}1$ به‌کار برید.

$$y' = -2y$$
$$y(0) = 1$$

فرمول رونگه – کوتا عبارتست‌از:

$$k_1 = hf(x_n, y_n)$$
$$k_2 = hf(x_n + h/2, y_n + k_1/2)$$
$$k_3 = hf(x_n + h/2, y_n + k_2/2)$$
$$k_4 = hf(x_n + h, y_n + k_3)$$
$$y_{n+1} = y_n + \tfrac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4]$$

$k_1 = 0{\cdot}1(-2{\cdot}0 \times 1{\cdot}0) = -0{\cdot}2$ $\quad$ $k_2 = 0{\cdot}1(-2{\cdot}0 \times 0{\cdot}9) = -0{\cdot}18$

$k_3 = 0{\cdot}1(-2{\cdot}0 \times 0{\cdot}91) = -0{\cdot}182$ $\quad$ $k_4 = 0{\cdot}1(-2{\cdot}0 \times 0{\cdot}818) = -0{\cdot}1636$

بنابراین

$$y_{0\cdot1} = 1 + \tfrac{1}{6}[-0\cdot2 - 0\cdot36 - 0\cdot364 - 0\cdot1636] \approx 0\cdot8187$$

$$k_1 = 0\cdot1(-2\cdot0 \times 0\cdot8187) = -0\cdot1627 \qquad k_2 = 0\cdot1(-2\cdot0 \times 0\cdot7368) = -0\cdot1474$$
$$k_3 = 0\cdot1(-2\cdot0 \times 0\cdot7450) = -0\cdot1490 \qquad k_4 = 0\cdot1(-2\cdot0 \times 0\cdot6697) = -0\cdot1339$$

$$y = 0\cdot8187 + \tfrac{1}{6}[-0\cdot1627 - 0\cdot2948 - 0\cdot2980 - 0\cdot1339] = 0\cdot6705$$

جواب واقعی عبارتست‌از $y = e^{-2x}$ که می‌تواند جهت کنترل خطا به‌کار رود . خطا در $x = 0\cdot2$ تقریبا" رقم ۳ در چهارمین محل اعشاری است . بهرحال ، درصورتیکه مقداری برای h مانند $1\cdot0$ که در شرط پایداری جزئی $h|\lambda| < 2\cdot7$ صادق است به کاربریم نتایج به‌طور ناامیدوارانه غیر دقیق بدست می‌آیند .

$$k_1 = -2\cdot0 \times 1\cdot0 = -2\cdot0 \qquad k_2 = -2\cdot0 \times 0\cdot0 = 0\cdot0$$
$$k_3 = -2\cdot0 \times 1\cdot0 = -2\cdot0 \qquad k_4 = -2\cdot0 \times 1\cdot0 = +2\cdot0$$
$$y_1 = 1 + \tfrac{1}{6}[-2\cdot0 - 4\cdot0 + 2\cdot0]$$
$$= 0\cdot3333$$

نتیجه واقعی عبارتست‌از $y = 0\cdot1353$ که خطائی معادل 146% را می‌دهد . بدین‌ترتیب محدودیت دقت در اندازه h از شرط پایداری جزئی موردنیاز خیلی فاصله می‌گیرد .

۷ ــ مسئله مقدار اولیه زیر را با نگهداری شش رقم اعشار ، به روش اویلر حل کنید .

$$y' = y \qquad y(0) = 1$$

حل ــ با استفاده از فرمول

$$\boxed{y_{n+1} = y_n + hf(x_n, y_n)}$$

و $h = 0.01$ بدست می‌آوریم .

$$y(0.01) \approx y_1 = 1 + 0.01 = 1.01$$
$$y(0.02) \approx y_2 = 1.01 + 0.01(1.01) = 1.0201$$
$$y(0.03) \approx y_3 = 1.0201 + 0.01(1.0201) = 1.030301$$
$$y(0.04) \approx y_4 = 1.030301 + 0.01(1.030301) = 1.040606$$

چون جواب دقیق معادله $y = e^x$ است مقدار دقیق در $x = 0.04$ عبارتست از 1.0408 . واضح است که برای بدست آوردن دقت بیشتر ، باید مقدار بسیار کوچکتری برای h بگیریم . اگر قرار دهیم $h = 0.005$ مقادیر زیر را بدست می‌آوریم .

$$y(0.005) \approx y_1 = 1.0050$$
$$y(0.010) \approx y_2 = 1.0100$$
$$y(0.015) \approx y_3 = 1.0151$$
$$y(0.020) \approx y_4 = 1.0202$$
$$y(0.025) \approx y_5 = 1.0253$$
$$y(0.030) \approx y_6 = 1.0304$$
$$y(0.035) \approx y_7 = 1.0356$$
$$y(0.040) \approx y_8 = 1.0408$$

این نتایج تا چهار رقم اعشار دقیق‌اند .

۸ــ با به کار بردن روش تیلر جواب معادله دیفرانسیل زیر را در $x = 2.1$ تا پنج رقم اعشار درست بدست آورید .

$$xy' = x - y \qquad y(2) = 2$$

حل ــ چندین مشتق اول و مقادیرشان در $x = 2$ و $y = 2$ عبارتند از :

$$y' = 1 - \frac{y}{x} \qquad y'_0 = 0$$

$$y'' = \frac{-y'}{x} + \frac{y}{x^2} \qquad y''_0 = \tfrac{1}{2}$$

$$y''' = \frac{-y''}{x} + \frac{2y'}{x^2} - \frac{2y}{x^3} \qquad y'''_0 = -\tfrac{3}{4}$$

$$y^{iv} = \frac{-y'''}{x} + \frac{3y''}{x^2} - \frac{6y'}{x^3} + \frac{6y}{x^4} \qquad y^{iv}_0 = \tfrac{3}{2}$$

بسط تیلر در اطراف $x_0 = 2$ عبارتست از :

$$\begin{aligned} y(x) &= y_0 + (x-2)y'_0 + \tfrac{1}{2}(x-2)^2 y''_0 + \tfrac{1}{6}(x-2)^3 y'''_0 + \tfrac{1}{24}(x-2)^4 y^{iv}_0 + \cdots \\ &= 2 + (x-2)0 + \tfrac{1}{4}(x-2)^2 - \tfrac{1}{8}(x-2)^3 + \tfrac{1}{16}(x-2)^4 + \cdots \end{aligned}$$

در $x = 2.1$ بدست می‌آوریم .

$$\begin{aligned} y(2.1) &= 2 + 0.0025 - 0.000125 + 0.0000062 - \cdots \\ &\approx 2.00238 \end{aligned}$$

چون جملات سری تیلر نزولی و مختلف‌العلامه‌اند این نتیجه تا پنج رقم اعشار درست است .

۹ ــ برای استفاده از فرمول چندگامی (روش آدامز ــ بشفورث)

$$\boxed{y_{n+1} = y_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})}$$

باید چهار مقدار آغازی داشته باشیم . این مقادیر باید از منبع مستقلی بدست آیند و برای نشان دادن اینکه چگونه از فرمول فوق استفاده می‌کنند ، چندگام از انتگرال‌گیری از معادله

$$\begin{aligned} y' &= -y^2 \\ y(1) &= 1 \end{aligned}$$

را با $h = 0.1$ انجام می‌دهیم . جواب دقیق این مسئله $y = \frac{1}{x}$ است . در جدول ۵.۳ زیر چهار مقدار آغازی از جواب دقیق بدست آمده‌اند و بقیه از فرمول .

جدول ۵۰۳

x_n	y_n	$f_n = -y_n^2$	$y(x_n) = 1/x_n$
1.0	1.00000000	−1.00000000	
1.1	0.90909091	−0.82644628	0.90909091
1.2	0.83333333	−0.69444444	0.83333333
1.3	0.76923077	−0.59171598	0.76923077
1.4	0.71443632	−0.51041926	0.71428571
1.5	0.66686030	−0.44470266	0.66666667
1.6	0.62524613	−0.39093272	0.62500000

۱۰ــ معادلهٔ

$$y' = x + y \qquad y(0) = 0$$

را از $x = 0$ تا $x = 1$ به روش آدامز ــ بشفورث حل کنید .

یک برنامه فرترن و نتایج آن برای مسئلهٔ فوق در زیر آمده است . جواب دقیق این مسئله عبارتست از $y = e^x - 1 - x$. چهار مقدار آغازی از این جواب بدست آمده‌اند . اولین ستون نتایج ، مقادیر x_n را با $h = \frac{1}{32}$ ، ستون دوم مقادیر y_n را که از فرمول آدامز ــ بشفورث حساب شده‌اند ، ستون سوم مقادیر $y(x_n)$ را که از جواب دقیق بدست آمده‌اند و ستون چهارم خطای مطلق را نشان می‌دهند . نتایج تا ۶ رقم بامعنی درست هستند .

FORTRAN PROGRAM FOR EXAMPLE 10

```
C        ADAMS-BASHFORTH METHOD
         DIMENSION F(4)
C
         SOLN(X) = EXP(X) - 1.0 - X
C
C        ** INITIALIZE
         WRITE (6,500)
         N = 4
         H = 1./32.
         NSTEP = 32
         Y0 = 0.0
         X0 = 0.0
         EN = 0.0
         NZ = 0
```

```
C
C        ** COMPUTE FIRST FOUR POINTS USING EXACT SOLUTION
         F(1) = X0 + Y0
         WRITE (6,510) NZ,X0,Y0,Y0,EN
         XN = X0
         DO 20 I = 2,4
         K = I - 1
         XN = XN + H
         YN = SOLN(XN)
         F(I) = XN + YN
         WRITE (6,510) K,XN,YN,YN,EN
      20 CONTINUE
C
C        ** BEGIN ITERATION
         DO 50 K = 4,NSTEP
         YN = YN + (H/24.)*(55.*F(N) - 59.*F(N - 1) + 37.*F(N - 2) -
        X9.*F(N - 3))
         XN = XN + H
         F(N - 3) = F(N - 2)
         F(N - 2) = F(N - 1)
         F(N - 1) = F(N)
         F(N) = XN + YN
         YXN = SOLN(XN)
         EN = YN - YXN
         WRITE (6,510) K,XN,YN,YXN,EN
      50 CONTINUE
         STOP
     500 FORMAT (23H1ADAMS-BASHFORTH METHOD/
        X1H04X1HN6X2HXN18X2HYN18X5HY(XN)15X15HEN = YN - Y(XN))
     510 FORMAT (1H 2XI3,4X 4(E16.8,4X))
         END
```

COMPUTER RESULTS FOR EXAMPLE *10*

N	XN	YN	Y(XN)	EN = YN − Y(XN)
0	0.	0.	0.	0.
1	0.31250000E-01	0.49340725E-03	0.49340725E-03	0.
2	0.62500000E-01	0.19944459E-02	0.19944459E-02	0.
3	0.93750000E-01	0.45351386E-02	0.45351386E-02	0.
4	0.12500000E-00	0.81484411E-02	0.81484467E-02	−0.55879354E-08
5	0.15625000E-00	0.12868421E-01	0.12868434E-01	−0.12922101E-07
6	0.18750000E-00	0.18730211E-01	0.18730238E-01	−0.26309863E-07
7	0.21875000E-00	0.25770056E-01	0.25770098E-01	−0.41676685E-07
8	0.25000000E-00	0.34025350E-01	0.34025416E-01	−0.65192580E-07
9	0.28125000E-00	0.43534677E-01	0.43534756E-01	−0.78696758E-07
10	0.31250000E-00	0.54337843E-01	0.54337934E-01	−0.90803951E-07
11	0.34375000E-00	0.66475919E-01	0.66476032E-01	−0.11269003E-06
12	0.37500000E-00	0.79991280E-01	0.79991400E-01	−0.12014061E-06
13	0.40625000E-00	0.94927646E-01	0.94927788E-01	−0.14156103E-06
14	0.43750000E-00	0.11133012E-00	0.11133029E-00	−0.16111881E-06
15	0.46875000E-00	0.12924525E-00	0.12924545E-00	−0.19185245E-06
16	0.50000000E 00	0.14872105E-00	0.14872126E-00	−0.21234155E-06
17	0.53125000E 00	0.16980705E-00	0.16980730E-00	−0.24400651E-06
18	0.56250000E 00	0.19255438E-00	0.19255465E-00	−0.26822090E-06

N	XN	YN	Y(XN)	EN = YN − Y(XN)
19	0.59375000E 00	0.21701577E-00	0.21701607E-00	−0.29988587E-06
20	0.62500000E 00	0.24324562E-00	0.24324594E-00	−0.31664968E-06
21	0.65625000E 00	0.27130008E-00	0.27130044E-00	−0.34645200E-06
22	0.68750000E 00	0.30123707E-00	0.30123746E-00	−0.39115548E-06
23	0.71875000E 00	0.33311634E-00	0.33311677E-00	−0.42840838E-06
24	0.75000000E 00	0.36699954E-00	0.36700001E-00	−0.46566129E-06
25	0.78125000E 00	0.40295030E-00	0.40295079E-00	−0.49173832E-06
26	0.81250000E 00	0.44103424E-00	0.44103476E-00	−0.52526593E-06
27	0.84375000E 00	0.48131907E-00	0.48131964E-00	−0.56624413E-06
28	0.87500000E 00	0.52387466E 00	0.52387527E 00	−0.61094761E-06
29	0.90625000E 00	0.56877308E 00	0.56877375E 00	−0.66310167E-06
30	0.93750000E 00	0.61608872E 00	0.61608934E 00	−0.71525574E-06
31	0.96875000E 00	0.66589829E 00	0.66589907E 00	−0.77486038E-06
32	0.09999999E 01	0.71828098E 00	0.71828181E 00	−0.82701445E-06

■ مسائل

۱ــ پایداری و دقت معادله تفاضل ــ متناهی معادله

$$y_{n+3}=y_{n+2}+\frac{h}{12}[23f_{n+2}-16f_{n+1}+5f_n]$$

را بررسی کنید.

۲ــ کدام روش یا روشها را برای نتایج زیر به‌کار می‌برید؟

(a) تابع مشتق پیچیده بوده و نیازمند برنامه‌های ویژه‌ای برای محاسبه خود می‌باشد.

(b) اندازه تابع مشتق در نقاط مختلف ناحیه به‌طور قابل توجهی تغییر می‌کند.

(c) حل دستگاه دومعادله دو مجهولی زیر مورد نظر می‌باشد.

$$\begin{aligned} y' &= -10y+6z, \qquad y(0)=4 \\ z' &= 13{\cdot}5y-10z, \qquad z(0)=0 \end{aligned}$$

۳ــ مثال حل شده شماره ۴ را تا حصول به همگرائی تا شش رقم اعشار حل کرده و دقت را با جوابی که از یک تکرار بدست می‌آید مقایسه نمائید.

۴ــ برنامه‌ای برای حل معادله

$$\begin{aligned} y' &= \sqrt{(1-y^2)} \\ y(0) &= -1 \end{aligned}$$

نوشته و روش پیشگو ــ اصلاحگر و Adams وسیله معادلات (5.25) و (5.26) با دو اصلاح در ناحیه 0 تا $0{\cdot}9$ ، $h=0{\cdot}1$ را به‌کار برید.

۵ــ معادله فوق را با کاربردی از اصلاحگر که از یک گام نتیجه‌می‌گردد حل کرده و جمله خطا را همانطوریکه در قسمت ۵۰۳۰۳ اشاره‌شده‌ونتایج‌را باروش‌فوق مقایسه‌کنید.

۶ــ فرمول‌زیر را درنظر بگیرید:

$$y_{n+2}=\alpha_1 y_{n+1}+\alpha_0 y_n+h\beta_1 f_{n+1}$$

مطلوبست تعیین a_0 ، a_1 و β_1 به‌منظور رسیدن به ماکزیمم مرتبه تقریب. جمله خطا را تعیین کرده و دستورات خواص پایداری فرمول را مشخص نمائید.

۷ ــ برای معادلهٔ

$$y' = xy+1 \qquad y(0)=1$$

مقدار $y(0.1)$ را به روش سری تیلر تا ۶ رقم اعشار پیدا کنید.

۸ ــ برنامه‌ای بنویسید که معادلهٔ

$$y' = \frac{x}{y}+\frac{y}{x} \qquad y(1)=0.5$$

را از $x=1$ تا $x=2$ به روش آدامز ــ بشفورث حل کنید.

برای $h = \frac{1}{64}$ و $h = \frac{1}{128}$ بگیرید و مقدار h را طوری تخمین بزنید که ۷ رقم اعشار درست بدهد.

۹ ــ نشان دهید که رابطهٔ زیر مشروط بر آنکه

$$|(9h/24)(\partial f/\partial y)| < 1$$

$$y_{n+1}^{(k)} = y_n + \frac{h}{24}[9f(x_{n+1}, y_{n+1}^{(k-1)}) + 19f_n - 5f_{n-1} + f_{n-2}]$$

همگرا خواهد شد.

۱۰ ــ با استفاده از فرمول روش پیشگو ــ اصلاحگر آدامز مولتن معادلهٔ

$$y' = y + x^2, \; y(0) = 1$$

را از $x = 0$ تا $x = 2$ با $h = 0.1$ حل کنید. مقادیر آغازی را چنین انتخاب کنید.

$$\begin{aligned} y(0) &= 1.000000 \\ y(0.1) &= 1.105513 \\ y(0.2) &= 1.224208 \\ y(0.3) &= 1.359576 \end{aligned}$$

۱۱ ــ معادلهٔ

$$y' = x - \frac{1}{y} \qquad y(0) = 1$$

را از $x = 0$ تا $x = 0.2$ با استفاده از فرمول روش پیشگو ــ اصلاحگر رتبه دوم

$$y_{n+1}^{(0)} = y_n + hf(x_n, y_n)$$

$$y_{n+1}^{(k)} = y_n + \frac{h}{2}[f(x_n, y_n) + f(x_{n+1}, y_{n+1}^{(k-1)})] \qquad k = 1, 2, \ldots$$

حل کنید.

* گام ۱

به روش اویلر

$$\begin{aligned} y_1^{(0)} &= 0.9 \\ y_1^{(1)} &= 0.8994 \\ y_1^{(2)} &= 0.8994 \end{aligned}$$

با فرمول فوق

چون $y_1^{(1)}$ و $y_1^{(2)}$ تا چهار رقم اعشار یکسانند این جواب را می‌پذیریم و قرار می‌دهیم:

$$f(x_1, y_1) = -1.0118$$

* گام ۲

به روش اویلر

$$y_2^{(0)} = 0.8994 + 0.1(-1.0118) = 0.7982$$

با فرمول فوق

$$y_2^{(1)} = 0.8994 + 0.05\left[-1.0118 + \left(0.2 - \frac{1}{0.7982}\right)\right] = 0.7962$$

$$y_2^{(2)} = 0.8994 + 0.05\left[-1.0118 + \left(0.2 - \frac{1}{0.7962}\right)\right] = 0.7960$$

$$y_2^{(3)} = 0.7960$$

مجددا $y_2 = 0.7960$ را می‌پذیریم و به گام بعدی می‌رویم .

فصل ششم
تفاضل های متناهی

۶.۱ ■ رفتار تفاضلات متناهی

سی‌سال قبل کارهای مشتمل بر استفاده از روشهای عددی افزایش یافت زیرا محاسبات یا بوسیلهٔ خط‌کش محاسبه و یا به کمک ماشین حسابهای رومیزی انجام‌می‌شد. در این محیط تاکیدی برآن روشها که ازنظر محاسباتی ساده و برای محاسبات دستی مناسب بود وجود داشت. روشهای مبتنی بر تفاضلات، بسیار مناسب این وضعیت بودند و بطور وسیعی به‌کار می‌رفتند.

با پیدایش‌کامپیوترهای رقمی الکترونیکی با سرعتهای بسیار زیاد عملیاتی، اما با ظرفیت حافظه‌ای‌کم، ذخیرهٔ جداول بزرگ تفاضلات متناهی کمتر امکان‌پذیر بود و روشها به طرق دیگر فرمولبندی شد. باز هم وضعیت تغییر پیدا کرده و محدودیت حافظه در اکثر وسایل محاسباتی علمی مسئله اصلی نمی‌باشد. این بدان معنی است که فرمولبندی تفاضلات متناهی دو مرتبه عملی است و ممکن است مجددا" جایگاه‌ارزشمندی در روشهای عددی پیدا کند.

روشهای تفاضلات متناهی روی توابعی اعمال می‌شود که مقادیرش در نقاط متساوی‌الفاصله‌موجودباشد، یک سری نقطه مانند $x_n = x_0 + n.h \ (n=0,1,\dots,N)$ که برایشان مقادیر متناظر $f_n \ (n=0,1,\dots,N)$ وجود دارند. با تفریق مقادیر تابعی متوالی‌ستون تفاضلات مرتبهٔ اول را می‌توان ساخت.

$$\Delta f_r = f_{r+1} - f_r \tag{6.1}$$

به‌طریقی مشابه ستون تفاضلات مرتبهٔ دوم را می‌توان از فرمول زیر حساب کرد.

$$\Delta^2 f_r = \Delta f_{r+1} - \Delta f_r \tag{6.2}$$

تفاضلات مرتبهٔ بالاتر را می‌توان به‌طریقی مشابه پیدا کرد. جدول تفاضل متناهی روش مفیدی برای یافتن تقریب یک تابع، به هنگامیکه آن را بتوان تقریبا" با یک چند

جمله‌ای نمایش داد ، بدست می‌دهد . در چنین حالتی ستونهای تفاضلات بعد از نقطهٔ معینی قابل اغماض می‌شوند و تنها به ستونهای قبل از آن نقطه از تفاضلات نیاز است ، که در محاسبات شرکت داده شوند .

مثلا " ، جدول (۶.۱) که تفاضلات مربوط به یک چند جمله‌ای درجه سوم را نشان می‌دهد تفاضلات ستون چهارم و تمام ستونهای بعد از آن ، در حالت چند جمله‌ای درجهٔ n ، ستون $n+1$ ام تفاضلات و تمام ستونهای بعد از آن صفرند . عکس آن نیز صادق است ، یعنی اگر معلوم شد ستونی از تفاضلات صفر است مقادیر تابع روی یک‌چند‌جمله‌ای از درجهٔ مساوی با بزرگترین تفاضلات که ناصفرند قرار دارد . اما در حالت یک مثال عملی ستونهای تفاضلات ممکن‌است به علت خطای ناشی از گرد کردن هرگز دقیقا " صفر نشوند . اکثرا " توابع مورد بررسی یک چند جمله‌ای نیستند اما باید بوسیلهٔ چند جمله‌ای مناسبی تقریب زده شوند . در چنین حالتی یک جدول تفاضلی ساخته می‌شود و ستونهای مختلف موردآزمایش قرار می‌گیرند . اگر ستون k از جدول تفاضلی آنقدر کوچک باشد که قابل اغماض‌به‌حساب‌آید جدول تفاضلی در این نقطه قطع می‌شود . این با تقریب بوسیلهٔ یک چند جمله‌ای درجهٔ $k-1$ معادل است .

x_n	f_n	Δ	Δ^2	Δ^3	Δ^4
1	1				
		7			
2	8		12		
		19		6	
3	27		18		0
		37		6	
4	64		24		0
		61		6	
5	125		30		0
		91		6	
6	216		36		
		127			
7	343				

جدول ۶.۱
جدول تفاضلی برای یک چند جمله‌ای درجه سوم

اما مثال دیگری نشان خواهد داد که این رفتار برای تمام توابع اتفاق نمی‌افتد . تابع $f(x)=2^x$ را که در جدول (۶.۲) نشان داده شده درنظر بگیرید .

x_n	f_n	Δ	Δ^2	Δ^3
1	2			
		2		
2	4		2	...
		4		
3	8		4	...
		8		
4	16		8	...
		16		
5	32		16	...
		32		
6	64			

جدول ۶.۲
جدول تفاضلی برای 2^x

درحالت فوق تمام ستونهای تفاضلات قابل توجه هستند و نمی‌توان از هیچیک از آنها صرفنظر نمود. جدول (۶.۳) که تفاضلات $\sqrt{x}$ را نشان می‌دهد نوع دیگری ازرفتاراست که در آن ستونهای تفاضلات کوچکتر می‌شوند اما الگوی سیستماتیکی باقی می‌ماند.

جدول ۶.۳

جدول تفاضلی برای $x^{\frac{1}{2}}$

x_n	f_n	Δ	Δ^2	Δ^3	Δ^4
1	1·000				
		0·414			
2	1·414		−0·096		
		0·318		0·046	
3	1·732		−0·050		−0·028
		0·268		0·018	
4	2·000		−0·032		−0·009
		0·236		0·009	
5	2·236		−0·023		−0·002
		0·213		0·007	
6	2·449		−0·016		
		0·197			
7	2·646				

در بحث بعدی این فصل فرض می‌شود که توابع مورد بررسی را می‌توان بحد کافی بوسیلهٔ یک چند جمله‌ای تقریب کرد.

۶.۲ ■ خطاها درجدول تفاضل متناهی

اکثر مقادیر تابع را نمی‌توان دقیقا" با تعداد متناهی رقم نمایش داد به‌طوری که جدول مقادیر تابع به علت خطای گردکردن شامل خطاهایی هستند. خطاها همچنین ممکن است بخاطر اشتباهات تحریری، نظیر جابجا نوشتن ارقام یا تکرارارقام به‌غلط، رخ دهد، بعنوان مثال 977 به جای 997 . دیدن اینکه چگونه این خطا وقتی ستونهای مختلف تفاضلات حساب می‌شوند انباشته می‌گردند، آگاهی دهنده است. این را می‌توان بوسیلهٔ بررسی یک جدول تفاضلی که در آن هر مؤلفه صفر است و وارد نمودن خطای e در یک مقدار تابع، نشان داد. جدول (۶.۴) رشد خطا را نشان می‌دهد.

جدول ۶.۴

انباشتگی خطا در جدول تفاضلی

f_n	Δ	Δ^2	Δ^3	Δ^4
0				
	0			
0		0		
	0		0	
0		0		$+e$
	0		$+e$	
0		$+e$		$-4e$
	$+e$		$-3e$	
e		$-2e$		$6e$
	$-e$		$+3e$	
0		$+e$		$-4e$
	0		$-e$	
0		0		$+e$
	0		0	
0		0		

ضرائب ، در هر سطر ، ضرائب دو جمله‌ای $_nC_r = n!/(n-r)!r!$ هستند که از نظراندازه با افزایش n رشد می‌کنند .

یکی از مسائل تفاضلات متناهی طریقی است که خطا بوسیلهٔ فرایند تفاضل کردن بزرگ می‌شود ، اما دانستن الگوی تولید شده بوسیلهٔ خطاها می‌تواند به شناسایی و حذف خطاها کمک کند . جدول (۶۰۵) را درنظر بگیرید .

x_n	f_n	Δ	Δ^2	Δ^3	Δ^4	Δ^5
$x_0 = 0$	100					
		1				
0·5	101		14			
		15		36		
1·0	116		50		24	
		65		60		0
1·5	181		110		24	
		175		84		−27
2·0	356		194		−3	
		369		81		135
2·5	725		275		132	
		644		213		−270
3·0	1369		488		−138	
		1132		75		+270
3·5	2501		563		132	
		1695		207		−135
4·0	4196		770		−3	
		2465		204		
4·5	6661		974			
		3439				
5·0	10100					

جدول (۶۰۵)
یک جدول تفاضلی
با یک خطا

نسبت جملات آخرین ستون عبارتست‌از 1:5:10:10:5 که نتیجه می‌دهد خطا در پنج ستون قبل رخ داده است ، زیرا این اعداد ضرائب دو جمله‌ای $_5C_r$ هستند . اندازهٔ خطا 27— است و محل خطا با تعقیب خط شیب‌دار قطری که بطرف پایین و بسوی ستون اول کشیده می‌شود دقیقا" مشخص می‌شود . خطایی به‌اندازهٔ 27— در مؤلفهٔ 1369 وجود دارد . این مطلب را می‌توان به آسانی با به‌کار بردن فرمول $f(x) = 100 + 16x^4$ ، که اعداد جدول (6.5) را می‌دهد ، آزمایش کرد . دو مثال دیگر از انتشار خطا در یک جدول تفاضل متناهی در مثال‌های (۶۰۱) و (۶۰۲) داده شده است .

پس ،هر جا که معلوم باشد که یک تابع بوسیلهٔ یک چند جمله‌ای تقریب می‌شود پیدا کردن خطاها به این طریق امکان دارد . این مسئله وقتی اعداد گرد شده به‌کار روند مشکلتر می‌شود ، زیرادر این صورت خطاها مضارب درستی از ضرائب چند جمله‌ای نخواهند بود .

۶.۳ ■ عملگرهای تفاضل متناهی

ما قبلا" عملگر تفاضل پسرو را که بوسیلهٔ رابطهٔ

$$\Delta f_r = f_{r+1} - f_r \tag{6.3}$$

تعریف‌می‌شود به‌کار برده‌ایم . این عملگر معروف است زیرا این عمل با استفاده از نقطهٔ بعدی دنباله و نقطه حاضر، جهت تعیین تفاضل، تعریف می‌شود. به طریقی مشابه عملگر تفاضل پسرو تعریف می‌شود .

$$\nabla f_r = f_r - f_{r-1} \tag{6.4}$$

عملگرهای دیگر که به‌کار می‌روند به استفاده از اندیسهای نقاط میانی احتیاج دارند . عملگــر تفاضــل مرکزی با رابطه زیر تعریف می‌شود .

$$\delta f_{r+1/2} = f_{r+1} - f_r \tag{6.5}$$

عملگر میانگین با رابطه زیر تعریف می‌شود .

$$\mu f_{r+1/2} = \tfrac{1}{2}[f_{r+1} + f_r] \tag{6.6}$$

برای تعیین ارتباط بیـــن این عملگرها ، عملگر دیگری لازم است ، عملگر انتقال ، که به طریق زیر تعریف می‌شود .

$$Ef_r = f_{r+1} \tag{6.7}$$

هنگام استفاده از عملگرها مناسب است که جبر عملگرها را صرفنظر از مقادیر تابعی که روی آن اثر می‌کنند بررسی کرد . در چنین حالاتی لازم است که شرائطی را که تحت آن این جبر قابل اعمال است بطور دقیق تجزیه و تحلیل کنیم . این تجزیه و تحلیل در کتابهای درسی استاندارد دربارهٔ تفاضلات متناهی نظیر Hildebrand (1956) یا Redish (1961) ارائه شده است .

برخی نتایج با فرض اینکه تحت شرایط مناسب که برای اطمینان برقراری آنها لازم است ، موجود است ، ذکر می‌شوند .

1 $\Delta f_r = f_{r+1} - f_r$

ازاین‌رو

$$\begin{aligned} \Delta &\equiv E - 1 \\ E &\equiv 1 + \Delta \end{aligned} \tag{6.8}$$

2 $\nabla f_r = f_r - f_{r-1}$

پس

$$\begin{aligned} \nabla &\equiv 1 - E^{-1} \\ E &\equiv (1 - \nabla)^{-1} \end{aligned} \tag{6.9}$$

3 $\delta f_{r+1/2} = f_{r+1} - f_r$

پس

$$\delta \equiv E^{1/2} - E^{-1/2} \qquad (6.10)$$

از معادلات (6.8) و (6.9)

$$4 \quad 1+\Delta \equiv \frac{1}{1-\nabla}$$

بنابراین

$$\Delta \equiv \frac{1}{1-\nabla} - 1 \equiv \frac{\nabla}{1-\nabla} \qquad (6.11)$$

$$\nabla \equiv 1 - \frac{1}{1+\Delta} \equiv \frac{\Delta}{1+\Delta} \qquad (6.12)$$

روابط بسیار بیشتری از این نوع ، که خواننده می‌تواند برای خود بدست آورد ، وجود دارد . برخی از روابط فوق جهت بدست آوردن روابط تفاضل متناهی استاندارد ، در بخشهای بعد ، به‌کار می‌روند .

همچنین داشتن فرمولهای مربوط به تفاضلات مراتب بالا و مقادیر تابع مفیداست .

$$\begin{aligned} \Delta^2 f_r &= \Delta f_{r+1} - \Delta f_r \\ &= f_{r+2} - f_{r+1} - (f_{r+1} - f_r) \\ &= f_{r+2} - 2f_{r+1} + f_r \\ &= (E-1)^2 f_r \end{aligned} \qquad (6.13)$$

و بطور عکس

$$\Delta^r = (E-1)^r \qquad (6.14)$$

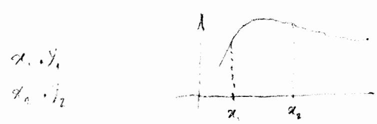
اگر به جدول (۶۰۵) رجوع کنیم می‌توانیم عناصر متناظر با تفاضلات مختلف را مشاهده کنیم . دنباله‌ای از تفاضلات پیشرو با اندیس یکسان در امتداد خط شیبدار قطری ، به طرف پایین ، رخ می‌دهد .

$$f_2 = 116, \quad \Delta f_2 = 65, \quad \Delta^2 f_2 = 110, \quad \Delta^3 f_2 = 84, \quad \text{etc.}$$

توجه کنید که عنصر 110 هم $\Delta^2 f_2$ وهم $\nabla^2 f_4$ است . تفاضلات مرکزی با اندیس ثابت را روی یک خط افقی می‌توان یافت .

$$f_3 = 181, \quad \delta^2 f_3 = 110, \quad \delta^4 f_3 = 24$$

مثالهای(۶۰۳) و (۶ ... ای بیشتری در رابطه با عملگرهای مختلف و مقادیر تابع نشان می‌دهند .



۶.۴ ■ درونیابی

با توسـ ... علامتگذاری فوق (که در کتابهای درسی ریاضی درستی آن ثابت می‌شود) بدسـ آوردن فرمولهایی که مقدار تابع را در یک نقطه غیر جدولی تخمین

می‌زند امکان‌پذیر است . در حالت جدولی نظیر جدول (۶.۱) یا تصحیح شدهٔ جدول (۶.۵)تفاضلات سرانجام صفر می‌شوندو لذا ، فرمولهایی می‌توان به‌کار برد که ، صرفنظر از خطاهای گرد کردن ، دقیق هستند .

اما ، با جدولی نظیرجدول(۶.۳)تفاضلات متناهی مرتبهٔ بالا کوچک خواهند بود ولی‌نه صفر. در چنین حالتی تعدادی متناهی از ستونهای تفاضلی به‌کار می‌روند و این خطائی ، به خاطر قطع دنبالهٔ تفاضلات متناهی ، وارد می‌کند. مثلا" ، اگر جدول (۶.۳) بعد از ستون سوم تفاضلات قطع شود تقریب معادل خواهد بود با تقریبی بوسیلهٔ یک چند جمله‌ای درجهٔ سوم بر مجموعهٔ خاصی از چهار نقطه.

توسیع این فرمول به طریق زیر بدست می‌آید . اگر p و k اعداد صحیحی باشند

$$\begin{aligned} f_{p+k} &= Ef_{p+k-1} \\ &= E^2 f_{p+k-2} \\ &= E^p f_k \end{aligned} \tag{6.15}$$

حال ، فرض کنیم این فرمول وقتی f_k یک نقطه جدول باشد که درآن $x_k = x_0 + k.h$ بازهم برقرار باشد اما p یک مقدارکسری بطوری‌که f_{p+k} نقطه جدولی نباشد .بااستفاده از معادلهٔ (6.8) ، $E \equiv 1 + \Delta$ ، داریم

$$\begin{aligned} f_{p+k} &= E^p f_k \\ &\equiv (1+\Delta)^p f_k \\ &= \left(1 + p\Delta + \frac{p.(p-1)}{2!}\Delta^2 \cdots\right) f_k \\ &= f_k + p\,\Delta f_k + \left(\frac{p^2-p}{2}\right)\Delta^2 f_k \cdots \end{aligned} \tag{6.16}$$

این فرمول ، اگر مقادیر تابع یک چندجمله‌ای را نمایش دهند ، مختوم خواهد بود .

این فرمول‌به فرمول تفاضل پسرو نیوتن معروف است و مقادیر درامتداد قطری بطرف پایین جدول تفاضلی‌رابه‌کار می‌برد . بنابراین ، بخصوص در شروع جدول مفید است ، زیرا مقادیر تفاضل مرکزی کافی روی یک خط قطری (یا تفاضلات پسرو در امتداد خط شیبداری به طرف بالا) جهت استفاده در فرمول وجود ندارد .

فرمولی مشابه که تفاضلات پسرو را به‌کار می‌برد بسادگی بادلیل مشابه بدست می‌آید .

$$\begin{aligned} f_{p+k} &= E^p f_k \\ &= (1-\nabla)^{-p} f_k \\ &= (1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 \cdots) f_k \end{aligned}$$

$$= 1 + p\nabla f_k + \frac{(p^2+p)}{2}\nabla^2 f_k \cdots \qquad (6.17)$$

این فرمول مخصوصا" برای درونیابی در انتهای جدول، جائیکه فقط تفاضلات پسرو موجودند، مناسب است.

در کتابهای درسی که شامل تحلیل ریاضی روشهای تفاضل متناهی هستند (ر.ک. Hildebrand (1956) یا Hartree (1958)) نشان داده شده است که این فرمولهای تفاضل پیشرو و یا پسرو به دقیقی فرمولهای مبتنی بر تفاضلات مرکزی نیستند. فرمولهای تفاضل مرکزی شناخته شده توسط بسل Bessel، استرلینگ Stirling و اورت Everett وجود دارد و یکی از آنها بطور نرمال، جائی که تفاضلات مناسب موجود باشند، به‌کار می‌رود.

استفاده از فرمولی نظیر فرمول تفاضل پیشرو نیوتن به بهترین وجه بوسیلهٔ یک مثال نشان داده شده است. اگر بخواهیم مقدار $f(0\cdot7)$ را با به‌کار بردن فرمول تفاضل پیشرو نیوتن، که بر نقطهٔ $x=0\cdot5$ استوار است، پیدا کنیم، با استفاده از داده‌های جدول (۶.۵) داریم.

$$x_{p+k} = x_k + p \times h$$
$$0\cdot7 = 0\cdot5 + p \times 0\cdot5$$

بنابراین

$$p = 0\cdot4 \qquad (6.18)$$

پس، تقریبات متوالی عبارتنداز:

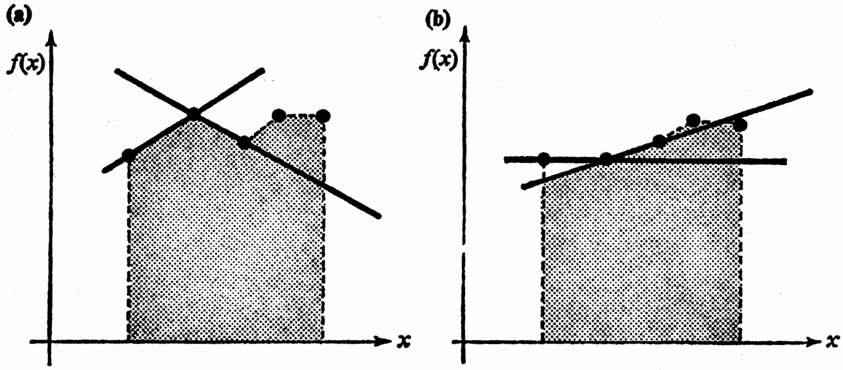
$$\begin{aligned} 101 + 0\cdot4 \times 15 &= 107 \\ 101 + 0\cdot4 \times 15 \frac{+(0\cdot4)\times(-0\cdot6)\times 50}{2!} &= 101 \\ 101 + 6\cdot0 - 6\cdot0 \frac{+(0\cdot4)(-0\cdot6)(-1\cdot6)\times 60}{6} &= 104\cdot84 \\ 101 + 6\cdot0 - 6\cdot0 + 3\cdot84 \frac{+(0\cdot4)(-0\cdot6)(-1\cdot6)(-2\cdot6)\times 24}{4\times 6} &= 103\cdot8416 \end{aligned} \qquad (6.19)$$

مقدار اخیر مقدار دقیق است زیرا تابع یک چند جمله‌ای درجه چهار است، مانند $16x^4+100$، و تمام تفاضلات تا تفاضلات چهارم منظور شده‌اند.

۶.۵ ■ فرمولهایی برای مشتق گیری و انتگرال گیری

قبل از ارائه مثالهایی از کاربرد فرمولهای تفاضلی برای مشتق‌گیری و انتگرال‌گیری شاید با ارزش باشد که به مسئله از طریق نمودار توجه کنیم تا مشکلات موجود را بهتر

احساس کنیم . دو مثال مربوط به شکلهای (b) ، (a) ۶۰۱ را درنظر بگیرید .



شکل (۶۰۱) ـــ خطا در انتگرال و مشتق

اختلاف بین دو شکل آن است که نقطه دوم دارای مقدار تابعی با جزئی اختلاف است . اختلاف اغراق شده است طوری که بسادگی در شکل دیده می‌شود اما ، مقادیر ، بطور نرمال ،حاوی خطای گرد کردن یا آزمایش هستند . اثر این خطا روی مشتق تقریبی در دو نقطه مجاور فاجعه‌آمیز است ، همانطور که در شکل نشان داده شده است . اما ، اگر انتگرال بوسیلهٔ سطح محدود شده تقریب شود ، مشاهده می‌شود که خطا در انتگرال نسبتا " کوچک است .

این بینشهای خام را می‌توان با تحلیل ریاضی از روشهای تفاضل متناهی تقویت کرد . این نشان می‌دهد که مشتق‌گیری از فرمولهای تفاضل متناهی می‌تواند به خطاهای بزرگ منجر شود ، درحالیکه فرایند انتگرال‌گیری اصولا " خوش طرح است . بنابراین ، فرمولهای مخصوص مشتق‌گیری پیگیری نمی‌شوند . هر جا ممکن باشد ، مسائل باید به‌گونه‌ای فرموله شوند که مشتق‌گیری از فرمولهای تفاضل متناهی ، یا توابع تقریب ، لازم نباشد .

فرمولهای مفیدی برای انتگرال می‌توان بدست آورد ، مثلا " ، با انتگرال‌گیری از فرمول تفاضل پیشرو نیوتن . چون $x_{p+k}=x_k+p.h$ که در آن p یک متغیر و x_k یک نقطه جدولی است ، داریم .

$$\frac{dx}{dp}=h$$

$$\int_{x_k}^{x_{k+1}} f(x)\,dx = h\int_0^1 f_{p+k}\,dp \tag{6.20}$$

$$= h\int_0^1 \left[1+p\Delta+\frac{p(p-1)}{2}\Delta^2\cdots\right] f_k\,dp$$

$$= h\left[pf_k+\frac{p^2}{2}\Delta f_k+\left(\frac{p^3}{6}-\frac{p^2}{4}\right)\Delta^2 f_k\cdots\right]_0^1$$

$$= h[f_k+\tfrac{1}{2}\Delta f_k-\tfrac{1}{12}\Delta^2 f_k\cdots] \tag{6.21}$$

دو جملهٔ اول این سری نتیجه می‌دهد :

$$I \approx \frac{h}{2}[f_k + f_{k+1}] \tag{6.22}$$

که فرمول ذوزنقه‌ای برای انتگرال‌گیری تقریبی است . با انتگرال‌گیری از x_k تا x_{k+2} و گرفتن سه جملهٔ اول ، فرمول

$$\begin{aligned} I &\approx h[2f_k + 2\Delta f_k + \tfrac{1}{3}\Delta^2 f_k] \\ &= \frac{h}{3}[f_k + 4f_{k+1} + f_{k+2}] \end{aligned} \tag{6.23}$$

بدست می‌آید که فرمول آشنای قاعدهٔ سیمپسون برای انتگرال‌گیری است .

فرمولهای مختلف بسیاری را می‌توان ، با استفاده از فرمولهای تفاضلی گوناگون و انتگرال‌گیری روی تعداد خاصی زیر فاصله ، بدست آورد . مثالهای جامعی از این فرمولها در کتابهای درسی دربارهٔ تفاضلات متناهی موجودند اما روش کلی‌تری برای تولید فرمولها ، در فصل ۸ دربارهٔ انتگرالها ، استفاده خواهد شد .

۶.۶ حل معادلات تفاضلی خطی

در مطالعهٔ معادلات دیفرانسیل ، حل یک معادلهٔ تفاضلی خطی پایداری یک روش را تعیین می‌کند . لذا ، حل یک معادلهٔ تفاضلی به‌شکل

$$\alpha_k y_{n+k} + \cdots + \alpha_1 y_{n+1} + \alpha_0 y_n = 0 \tag{6.24}$$

را بررسی می‌کنیم . این یک معادله تفاضلی خطی از درجهٔ k است که دارای k جواب مستقل خطی است . جواب کامل ترکیبی خطی از این جوابها است و بنابراین شامل k ثابت دلخواه است . این ثابتها بوسیلهٔ k شرط زیر پیدا می‌شوند .

$$y_r = s_r, \qquad r = 0, 1, \ldots, k-1 \tag{6.25}$$

جوابهای معادله (6.24) را می‌توان با جایگذاری جواب آزمایشی $y_n = A(z)^n$ و پیدا کردن مقادیری از z بطوری که معادله برقرار باشد ، بدست آورد . جایگذاری جواب آزمایشی می‌دهد :

$$Az^n[\alpha_k z^k + \alpha_{k-1} z^{k-1} + \cdots + \alpha_0] = 0 \tag{6.26}$$

این چند جمله‌ای به‌عنوان یک چند جمله‌ای کمکی شناخته می‌شود و هر ریشه z_r از این چند جمله‌ای منجر به یک جواب $y_n = A(z_r)^n$ از معادله تفاضلی می‌شود . اگر ریشه‌های مکرری از معادلهٔ تفاضلی وجود داشته باشد مشکلی بوجود می‌آید زیرا در این صورت تمام ریشه‌ها نمی‌توانند این شکل ساده را داشته باشند . درحالت یک ریشه مضاعف z_r این به دو ریشه z_r و $n.z_r$ منجر می‌شود . برای یک ریشه سه‌گانه سه جواب مستقل عبارتنداز z_r ، nz_r و $n^2 z_r$ وغیره .

پس، وقتی n ریشه متمایز وجود دارد جواب کامل عبارتست‌از :

$$y_n = A_1(z_1)^n + A_2(z_2)^n + \cdots + A_k(z_k)^n \tag{6.27}$$

اگر z_1 ریشه مضاعف باشد این‌جواب به‌صورت زیر درمی‌آید .

$$y_n = (A_1 + A_2 n)(z_1)^n + A_3(z_2)^n + \cdots + A_k(z_k)^n \tag{6.28}$$

دراین‌صورت k شرط‌اولیه‌معادله (6.25) برای‌بدست‌آوردن‌ثابتهای $A_r\ (r=1,\ldots,k)$ بوسیلهٔ حل یک دستگاه معادلات خطی ، کافی است . در حالت ریشه‌های متمایز این دستگاه عبارتست‌از :

$$\begin{aligned} &A_1 + A_2 + \cdots + A_k = s_0 \\ &A_1 z_1 + A_2 z_2 + \cdots + A_k z_k = s_1 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &A_1 z_1^{k-1} + A_2 z_2^{k-1} + \cdots + A_k z_k^{k-1} = s_{k-1} \end{aligned} \tag{6.29}$$

یک مثال حل شده در مثال (۶۰۵) داده شده است .

مثالهای حل شده وبرنامه های کامپیوتری

۱ ــ جدول زیر طریقهٔ دیگری که خطاها می‌توانند انباشته شوند را ، دریک جدول تفاضلی که خطاها متناوبا" مثبت و منفی هستند ، نشان می‌دهد . این الگو که ماکزیمم انباشتگی خطاها را می‌دهد خطاهای حداکثر دریک جدول را با مؤلفه‌های صفر نشان می‌دهد .

$+\varepsilon$					
	-2ε				
$-\varepsilon$		$+4\varepsilon$			
	$+2\varepsilon$		-8ε		
$+\varepsilon$		-4ε		$+16\varepsilon$	
	-2ε		$+8\varepsilon$		-32ε
$-\varepsilon$		$+4\varepsilon$		-16ε	
	$+2\varepsilon$		-8ε		
$+\varepsilon$		-4ε			
	-2ε				
$-\varepsilon$					

۲ــ یک جدول که در آن دو خطا انجام گرفته است برای تحلیل مشکل‌تر است،زیرا ضرائب دو جمله‌ای روی هم می‌افتند الگوی زیر مثال ممکنی را نشان می‌دهد . باید توجه شود که دریک مثال حقیقی مقادیر تابع در چند ستون اول قابل ملاحظه خواهند بود و خطاها باید از ستونهای دست راست ، جائیکه ماکزیمم تفاضل اتفاق افتاده است ، ردیابی شوند . الگوی خطا را در تفاضلات مرتبه ۳ می‌توان مشاهده کرد اما بهم‌ریختگی در ستون تفاضلی چهارم احتمالا" خیلی زیاد خواهند بود که فرصتی جهت آگاهی از خطاها بدهند .

$$
\begin{array}{ccccc}
f & \Delta & \Delta^2 & \Delta^3 & \Delta^4 \\
0 & & & & \\
 & 0 & & & \\
0 & & 0 & & \\
 & 0 & & \varepsilon_1 & \\
0 & & \varepsilon_1 & & -4\varepsilon_1 \\
 & \varepsilon_1 & & -3\varepsilon_1 & \\
\varepsilon_1 & & -2\varepsilon_1 & & 6\varepsilon_1 \\
 & -\varepsilon_1 & & 3\varepsilon_1 & \\
0 & & \varepsilon_1 & & \varepsilon_2-4\varepsilon_1 \\
 & 0 & & \varepsilon_2-\varepsilon_1 & \\
0 & & \varepsilon_2 & & +4\varepsilon_2+\varepsilon_1 \\
 & \varepsilon_2 & & -3\varepsilon_2 & \\
\varepsilon_2 & & -2\varepsilon_2 & & 6\varepsilon_2 \\
 & -\varepsilon_2 & & 3\varepsilon_2 & \\
 & & \varepsilon_2 & & \\
0 & & & & \\
 & 0 & & & \\
0 & & & &
\end{array}
$$

۳ــ رابطه بین یک فرمول بیان شده برحسب تفاضلات و همان فرمول که به‌صورت مجموعی از مقادیر یک تابع بیان شده است را بسادگی می‌توان، با استفاده از فرمولهای متعددی که عملگرهای مختلف را بهم مربوط می‌کند، تعیین کرد. مثلاً "

$$\nabla = 1 - E^{-1},\quad \nabla^2 = E^{-2} - 2E^{-1} + 1,\quad \text{etc.}$$

فرمول آدامزــ بشفورث Adams–Bashforth که برای انتگرال‌گیری از معادلات دیفرانسیل معمولی به‌کار می‌رود عبارتست‌از:

$$\nabla y_{n+1} = h[1 + \tfrac{1}{2}\nabla + \tfrac{5}{12}\nabla^2]y_n'$$

تا تفاضلات مرتبه دوم این را می‌توان به‌طریق زیر بیان‌کرد.

$$
\begin{aligned}
y_{n+1} - y_n &= hy_n' + \frac{h}{2}[y_n' - y_{n-1}'] + \frac{5h}{12}[y_n' - 2y_{n-1}' + y_{n-2}'] \\
&= \frac{23h}{12}y_n' - \frac{16h}{12}y_{n-1}' + \frac{5h}{12}y_{n-2}'
\end{aligned}
$$

۴ــ رابطه بین عملگرهای E و δ را پیدا کنید.

$$\delta f_r = f_{r+1/2} - f_{r-1/2}$$

پس

$$
\begin{aligned}
&\delta = E^{1/2} - E^{-1/2} \\
&\delta^2 = E - 2 + E^{-1} \\
&E - 2 - \delta^2 + E^{-1} = 0 \\
&E = 2 + \frac{\delta^2 + \sqrt{[(2+\delta^2)^2 - 4]}}{2}
\end{aligned}
$$

۵ ــ معادله تفاضلی خطی زیر را

$$y_{r+4} - y_{r+3} - 3y_{r+2} + y_{r+1} + 2y_r = 0$$

با شرایط اولیه

$$y_0 = 1,\ y_1 = -1,\ y_2 = 2,\ y_3 = 7$$

حل کنید . معادلهٔ کمکی زیر را تشکیل می‌دهیم .

$$z^4 - z^3 - 3z^2 + z + 2 = 0$$

با تحقیق معلوم می‌شود یک ریشه $z = 1$ موجود است . بنابراین ، با تقسیم بر $z - 1$ ریشه‌های معادله زیر را می‌یابیم .

$$z^3 - 3z - 2 = 0$$

معادلهٔ فوق ریشهٔ $z = -1$ دارد که ریشه مضاعف است و $z = 2$. لذاجواب‌کلی عبارتست‌از :

$$y_n = A(+1)^n + B(-1)^n + Cn(-1)^n + D(2)^n$$

و با استفاده از شرایط اولیه نتیجه می‌شود :

$$\begin{aligned} A+B+D&=1\\ A-B-C+2D&=-1\\ A+B+2C+4D&=2\\ A-B-3C+8D&=7 \end{aligned}$$

جمع کردن معادلات اول و سوم و تفریق مجموع معادلات سوم و چهارم از آن نتیجه می‌دهد

$$-9D = -9$$

بنابراین

$$D = 1, \qquad A + B = 0$$

و از معادلات دوم و سوم $3A - B = -8$ و لذا ،

$$A = -2, \qquad B = 2 \quad C = -4 + 4 + 1 = 1$$

۶ ــ با استفاده از جدول داده‌شده مطلوبست محاسبهٔ $\log 2.15$ با ماکزیمم دقت .

حل : از آنجائی که 2.15 بین 2.0 و 2.2 قرار دارد ، فرض می‌کنیم $x_0 = 2.00$ و فرمول پیشرو نیوتن را با $n = 4$ به کار می‌بریم .

x	$\log x$	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
2.0	0.30103				
		4139			
2.2	0.34242		−360		
		3779		57	
2.4	0.38021		−303		−11
		3476		46	
2.6	0.41497		−257		−12
		3219		34	
2.8	0.44716		−223		
		2996			
3.0	0.47712				

$$p_4(x_0+sh)=f_0+\binom{s}{1}\Delta f_0+\binom{s}{2}\Delta^2 f_0+\binom{s}{3}\Delta^3 f_0+\binom{s}{4}\Delta^4 f_0$$

با مقادیر $x_0=2.0$ و $h=0.2$ و $s=(x-x_0)/h$ مقدار $\bar{x}=2.15$ که متناظر با $s=(2.15-2.0)/0.2=0.75$

و

$$\binom{s}{1}=0.75$$

$$\binom{s}{2}=\binom{s}{1}\frac{s-1}{2}=\frac{0.75(-0.25)}{2}=-0.09375$$

$$\binom{s}{3}=\binom{s}{2}\frac{s-2}{3}=-0.09375\frac{-1.25}{3}=0.03906$$

$$\binom{s}{4}=\binom{s}{3}\frac{s-3}{4}=0.03906\frac{-2.25}{4}=-0.02197$$

بنابراین

$$\begin{aligned}\log 2.15\approx p_4(0.75)&=0.30103+(0.75)(0.04139)+(-0.09375)(-0.00360)\\&\quad+(0.03906)(0.00057)+(-0.02197)(-0.00011)\\&=0.332435\end{aligned}$$

و تخمین خطا از جمله دوم نیوتن بدست می‌آید. بدین ترتیب، از آنجائی که

$$\binom{s}{5}\Delta^5 f_0=-0.0000014$$

نتیجه تا پنج رقم اعشار با معنی درست خواهد بود. مقدار گردشده $\log 2.15$ مقدار 0.33244 خواهد بود که با نتیجه بدست‌آمده مطابقت دارد.

۷ ــ با استفاده از جدول مسئلهٔ ۶ مقدار $\log 2.9$ را با ماکزیمم دقت محاسبه نمائید.

حل: از آنجائی که 2.9 نزدیک پایان جدول می‌باشد، $x_0=3.0$ انتخاب کرده و فرمول نیوتن ــ پسرو را در نظر می‌گیریم.

برای $h=0.2$، داریم: $s=(2.9-3.0)/0.2=-\frac{1}{2}$

فرمول نیوتن ــ پسرو با $n=4$

$$p_4(x_0+sh)=f_0-\binom{-s}{1}\Delta f_{-1}+\binom{-s}{2}\Delta^2 f_{-2}-\binom{-s}{3}\Delta^3 f_{-3}+\binom{-s}{4}\Delta^4 f_{-4}$$

خواهد بود. آنگاه

$$\binom{-s}{1}=-s=\frac{1}{2}$$

$$\binom{-s}{2}=\binom{-s}{1}\frac{-s-1}{2}=\frac{1}{2}\cdot\frac{-1/2}{2}=-\frac{1}{8}$$

$$\binom{-s}{3}=\binom{-s}{2}\frac{-s-2}{3}=-\frac{1}{8}\cdot\frac{-3/2}{3}=\frac{1}{16}$$

$$\binom{-s}{4}=\binom{-s}{3}\frac{-s-3}{4}=\frac{1}{16}\cdot\frac{-5/2}{4}=-\frac{5}{128}$$

بنابراین

$$\begin{aligned}\log 2.9 \approx p_4\left(-\tfrac{1}{2}\right) &= 0.47712 - \tfrac{1}{2}(0.02996) - \tfrac{1}{8}(-0.00223)\\ &\quad - \tfrac{1}{16}(0.00034) - \tfrac{5}{128}(-0.00012)\\ &= 0.46240\end{aligned}$$

■ مسائل

۱ــ خطای جدول زیر را پیدا و آنرا تصحیح نمائید .

x	0·0	0·1	0·2	0·3	0·4	0·5
$f(x)$	4·000	4·641	5·368	6·187	7·104	8·125

x	0·6	0·7	0·8	0·9	1·0
$f(x)$	9·265	10·503	11·872	13·369	15·000

۲ ــ رابطهٔ بین عملگرهای μ و δ را پیدا کنید .

۳ ــ فرمول آدامز ــ بشفورث بستهٔ زیر تا تفاضلات دوم را برحسب مقادیر تابع بیان‌کنید .

$$\nabla y_{n+1} = h[1 - \tfrac{1}{2}\nabla - \tfrac{1}{12}\nabla^2]y'_{n+1}$$

۴ ــ معادله تفاضلی خطی زیر را

$$2y_{r+3} + y_{r+2} - 2y_{r+1} - y_r = 0$$

با شرایط اولیه $y_2 = \frac{1}{2}$ ، $y_1 = 2$ ، $y_0 = -1$ حل نمائید .

۵ ــ از فرمول تفاضل پسرو نیوتن تا تفاضلات مرتبهٔ سوم انتگرال بگیرید و فرمولهای گوناگونی برحسب مقادیر تابع ، با قطع کردن فرمول بعد از تفاضل اول و غیره ، بدست آورید .

۶ ــ با استفاده از جدول زیر مطلوبست محاسبه :

(*a*) $f(1252.5)$ (*b*) $f(1332.5)$

و در هر حالت خطا را محاسبه نمائید .

t	$x = f(t)$	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$	$\Delta^5 f$	$\Delta^6 f$
1,250.5	1.39140						
		−1444					
1,260.5	1.37696		−1469				
		−2913		55			
1,270.5	1.34783		−1414		17		
		−4327		72		−3	
1,280.5	1.30456		−1342		14		−2
		−5669		86		−5	
1,290.5	1.24787		−1256		9		8
		−6925		95		3	
1,300.5	1.17862		−1161		12		−11
		−8086		107		−8	
1,310.5	1.09776		−1054		4		8
		−9140		111		0	
1,320.5	1.00636		−943		4		
		−10083		115			
1,330.5	0.90553		−828				
		−10911					
1,340.5	0.79642						

۷ ــ مطلوبست محاسبهٔ جواب عمومی معادلات تفاضلی زیر :

$(a)\quad y_{n+1} - 3y_n = 5$
$(b)\quad y_{n+2} - 4y_{n+1} + 4y_n = n$
$(c)\quad y_{n+2} + 2y_{n+1} + 2y_n = 0$
$(d)\quad y_{n+3} - y_{n+2} + 2y_{n+1} - 2y_n = 0$
$(e)\quad y_{n+2} - y_{n+1} - y_n = 0$

۸ ــ جواب معادلات تفاضلی را تعیین کنید .

$(a)\quad y_{n+2} - 4y_{n+1} + 3y_n = 2^n \qquad y_0 = 0 \qquad y_1 = 1$
$(b)\quad y_{n+2} - y_{n+1} - y_n = 0 \qquad y_0 = 0 \qquad y_1 = 1$

۹ ــ نشان دهید که جواب عمومی معادلهٔ تفاضلی :

$$y_{n+2} + 4hy_{n+1} - y_n = 2h$$

که در آن h مقداری است مثبت و ثابت ، به صورت زیر بیان می‌گردد :

$$y_n = c_1[1 - 2h + O(h^2)]^n + c_2(-1)^n[1 + 2h + O(h^2)]^n + \tfrac{1}{2}$$

۱۰ ــ نشان دهید که جواب عمومی معادلهٔ تفاضلی :

$$y_{n+2} - (2 + h^2)y_{n+1} + y_n = h^2 \qquad h > 0$$

به صورت زیر بیان می‌گردد .

$$y_n = c_1\left[1 + h + \frac{h^2}{2} + 0(h^3)\right]^n + c_2\left[1 - h + \frac{h^2}{2} + 0(h^3)\right]^n - 1$$

فصل هفتم
برازش منحنی

۷.۱ ■ درونیابی و تقریب

دو نوع مسئله وجود دارد که تحت‌این عنوان بررسی می‌شود، ابتدا مسئله درونیابی است که شامل پیدا کردن مقادیر میانی می‌باشد، هنگامی که مقادیر در مجموعه‌ای متناهی از نقاط داده شده‌اند و دوم، مسئله تقریب به یک تابع برحوزه‌ای کامل از مقادیر بوسیلهٔ تابعی ساده می‌باشد، که برای محاسبه بیشتر مناسب است. آشکارا، هدف از تقریب باید کوچک کردن خطا تا حد امکان باشد و روشهای متفاوتی متناظر با طرق مختلف تعریف خطا ظهور می‌نماید.

در حالت درونیابی بوسیلهٔ روشهای تفاضل متناهی، تابع تقریب یک‌چند‌جمله‌ای است که مقادیر آن با مقادیر مفروضی در مجموعه‌ای از نقاط برابر است. فرض می‌کنیم f_i مقادیر مفروض در نقاط $x_i\ (i=0,1,\dots n)$ و $\phi_n(x)$ تابع تقریب باشد. اگر خطا را بوسیلهٔ رابطهٔ

$$E_n=\sum_{i=0}^{n}|\phi_n(x_i)-f_i| \tag{7.1}$$

تعریف کنیم برازش تابع بطور دقیق در $n+1$ نقطه خطای E_n را به صفر تقلیل می‌دهد. خطای تعریف شده در رابطهٔ (7.1) یقینا" می‌نیمم شده است، اما این سؤال باقی می‌ماند که آیا مقادیر در نقاطی که $x\neq x_i$ تقریبهای خوبی هستند؟

در مسئله تقریب با خطا در تمام نقاط حوزه سروکار داریم. مسائل موجود در به‌کار بردن تعریف خطایی که تنها بستگی به مقادیر در مجموعه‌ای متناهی از نقاط دارد بوضوح بوسیلهٔ مثالی که توسط رونگه Runge در 1901 کشف شد نشان داده شده است. این مثال در لنکزوس Lanczos بررسی شده است. مجموعه‌ای از نقاط متساوی‌الفاصله $x_i\ (i=0,1,\dots,n)$ جهت تقریب تابع $1/(1+25x^2)$ برحوزهٔ $[-1,+1]$

انتخاب شده‌اند. به‌ازای هر نقطه $x \neq x_i$ که $|x| > 0{\cdot}726$ ثابت شده است که خطای $|\phi_n(x) - f(x)|$ با افزایش n بدون کران زیاد می‌گردد. و این امر، با توجه به اینکه $\phi_n(x_i) = f(x_i)$, $(i = 0, 1, \ldots, n)$ که بمعنی $E_n = 0$ است، اتفاق می‌افتد.

هنگام بررسی خطا برتمام حوزه، هدف رضایت‌بخش‌تر آن است که ماکزیمم خطا را حتی‌الامکان کوچک کنیم این تقریب از نوع مینیماکس است که در آن خطا بوسیلهٔ رابطه

$$E_{\max} = \max_{a \leqslant x \leqslant b} |\phi(x) - f(x)| \tag{7.2}$$

تعریف می‌شود و تابع $\phi(x)$ چنان انتخاب می‌شود که $E_{\max}$ می‌نیمم شود. به این دلیل است که چند جمله‌ایهای چبیشف Chebyshev کاربرد وسیعی پیدا کرده‌اند.

حالت سوم حالت‌مفیدی است که‌تعداد نقاطی که مقادیر در آنها داده شده است بطور قابل ملاحظه‌ای بزرگتر از درجهٔ تقریب مطلوب است. مثلا، ممکن است استفاده از چند جمله‌ای درجه پایینی، مثلا" درجهٔ ۳، برای تقریب بر حوزه‌ای خواسته شود که دوازده مقدار تابع معلوم است. چهار نقطه برای تعیین یک چند جمله‌ای درجه سوم به‌طور یکتا کافی است و خطاها احتمالا" در نقاط باقیمانده، رخ می‌دهند.

در این وضعیت بجای صفر کردن خطا در نقاطی خاص، پایین نگاه‌داشتن خطای کلی را، تاحد امکان طلب می‌کنیم. انتخاب مناسبی از تعریف خطا بوسیلهٔ رابطهٔ

$$S_m = \sum_{i=0}^{m} [\phi_n(x_i) - f_i]^2, \qquad m \geqslant n \tag{7.3}$$

داده می‌شود، برازش کمترین مربعات بوسیله پیدا کردن تابع $\phi_n(x)$ که کمیت S_m را می‌نیمم می‌کند بدست می‌آید. اندیس n معرف آن است که تابع $\phi_n(x)$ به تعدادی پارامتر بستگی دارد که جهت بدست آوردن خاصیت کمترین مربعات می‌توان به آنها مقادیر مناسب داد. در حالت‌چند جمله‌ای این پارامترها ضرائب $a_0, a_1, \ldots, a_n$ و تابع $\phi_n(x)$ دارای $n+1$ پارامتر خواهد بود.

چند جمله‌ایها بطور وسیعی برای تقریب به‌کار می‌روند لذا، بررسی اینکه آنها مناسبترین توابع برای این منظور هستند یا نه، با ارزش است. یک مزیت بزرگ آن است که با استفاده از عملیات حسابی مهیا در کامپیوترهای رقمی محاسبهٔ مستقیم یک چند جمله‌ای یا خارج قسمت دو چند جمله‌ای امکان دارد. اما، محاسبهٔ توابع دیگر، مثل توابع نمائی یا مثلثاتی، بوسیلهٔ روشهای تقریب است. همچنین، تعیین انتگرال و مشتق چند جمله‌ایها با محاسبهٔ مستقیم آسان است. دانستن اینکه یک تقریب با به‌کار بردن چند جمله‌ای به چه نزدیکی حاصل می‌شود نیز مهم است. خوشبختانه، قضیه‌ای از وایراشتراس Weierstrass وجود دارد که نشان می‌دهد، به ازای هر تابع پیوسته بر یک فاصلهٔ محدود خطا را می‌توان با انتخاب یک چند جمله‌ای از درجهٔ باندازه کافی بالا

بدلخواه کوچک کرد .

نوع دیگر تقریب که باز هم ارزشمند است تقریب بوسیلهٔ سری فوریه Fourier است . در این حالت نشان داده می‌شود که می‌توان برای دستهٔ کاملا " وسیعتری از توابع ، مثلا " آنهایی که در شرایط دیریکله Dirichlet صدق می‌کنند ، تقریبهای بدلخواه نزدیک بدست آورد . (ر.ک. لنکزوس Lanczos (1957)) .

اکنون این فرمهای مختلف تقریب تشریح می‌شوند . بحثی دربارهٔ خواص چند جمله‌ایهای متعامد نیز ، بخاطر اهمیت آنها در تئوری تقریب ، گنجانده شده است .

۷.۲ درونیابی

۷.۲.۱ درونیابی لاگرانژ

در فصل پیش نشان داده شد که جدول تفاضل متناهی را می‌توان برای درونیابی به‌کار برد ، اما این به حالت مقادیر تابع در فاصله‌های متساوی محدود شده بود . فرم لاگرانژ چند جمله‌ای درونیاب ، یک چند جمله‌ای بدست می‌دهدکه یک تابع مفروض را در تعدادی از نقاط با محل دلخواه برازش می‌کند ، یعنی

$$P_n(x_i)=f_i, \qquad i=0,1,\ldots,n \tag{7.4}$$

یا در شکل بسط یافته

$$\begin{aligned} a_0+a_1x_0+\cdots+a_nx_0^n&=f_0\\ a_0+a_1x_1+\cdots+a_nx_1^n&=f_1\\ \cdots\cdots\cdots&\cdots\cdots\\ a_0+a_1x_n+\cdots+a_nx_n^n&=f_n \end{aligned} \tag{7.5}$$

اگر تمام x_i ها متمایز باشند ماتریس وابسته به این معادلات نامنفرد است و لذا معادلات را می‌توان بطور نظری جهت یافتن ضرائب a_i حل نمود . اما ، این معادلات می‌توانند بدوضع باشند و از آنجا ترجیح داده می‌شود که ضرائب به روش دیگری پیدا شوند .

از معادلات می‌توان دید که جوابهای a_i ترکیبی‌خطی از f_i ها می‌باشند . اگر این نوع عبارات برای a_i ها در چند جمله‌ای درج شوند و جملات مربوط به هر f_i با هم جمع شوند ، داریم .

$$P_n(x)=\sum_{i=0}^{n} l_{n,i}(x)f_i \tag{7.6}$$

که در آن $l_{n,i}(x)$ چند جمله‌ایهایی حداکثر از درجهٔ n هستند .

بدست آوردن فرم چند جمله‌ای $l_{n,i}(x)$ با به‌کار بردن معادلهٔ (7.4) که‌درونیابی لاگرانژرا تعریف می‌کندآسان می‌باشد . درنقطه x_0 تمام مقادیر $l_{n,i}(x_0)$ $(i=1,2,\ldots,n)$

باید صفر باشند تا جملات f_i متناظر حذف شوند . این یک ریشه از تمام‌چندجمله‌ایهای $l_{n,i}(x_0)\,(i\neq 0)$. را در $x=x_0$ می‌دهد . همچنین $l_{n,0}(x_0)$ بایدمساوی واحد باشد تا ضریب درست f_0 را بدهد . دلیل مشابهی برای هر نقطه می‌توان اعمال کردونشان‌داد که هر چند جمله‌ای ، ریشه‌ای در n نقطه دیگر از مجموعهٔ نقاط دارد . برای مثال ،

$$l_{n,0}(x_0)=A_0(x-x_1)(x-x_2)\cdots(x-x_n) \tag{7.7}$$

وبطور کلی

$$l_{n,i}(x)=A_i(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n) \tag{7.8}$$

شرط دوم ، که مقدار لازم را در نقطهٔ x_i می‌دهد ، پیدا کردن A_i را میسرمی‌کند ، یعنی چون $l_{n,i}(x_i)=1$ آنگاه

$$A_i=\frac{1}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)} \tag{7.9}$$

لذا چند جمله‌ایها را می‌توان بطور دقیق نوشت :

$$l_{n,i}(x)=\prod_{r=0}^{n}{}'\frac{x-x_r}{x_i-x_r} \tag{7.10}$$

پریم برای نشان دادن اینکه جمله x_i از حاصلضرب حذف شده ، به کار رفته است . اگر بخواهیم در هر نقطهٔ x درونیابی کنیم مقادیر این چند جمله‌ایها محاسبه شده و در مقادیر متناظر f_i ضرب می‌شوند . مثلا" ، مقدار درونیابی شدهٔ $f(5)$ را می‌توان به طریق زیر از جدول بدست آورد .

	x_0	x_1	x_2	x_3
x_i	1	4	7	9
f_i	2	13	122	504

$$l_{3,0}(x)=\frac{(x-4)(x-7)(x-9)}{(1-4)(1-7)(1-9)} \tag{7.11}$$

بنابراین

$$l_{3,0}(5)=\frac{1.(-2).(-4)}{-3.(-6).(-8)}=-\frac{1}{18} \tag{7.12a}$$

و متشابها"

$$l_{3,1}(5)=\frac{(5-1)(5-7)(5-9)}{(4-1)(4-7)(4-9)}=\frac{32}{45} \tag{7.12b}$$

$$l_{3,2}(5)=\frac{(5-1)(5-4)(5-9)}{(7-1)(7-4)(7-9)}=\frac{4}{9} \tag{7.12c}$$

$$l_{3,3}(5)=\frac{(5-1)(5-4)(5-7)}{(9-1)(9-4)(9-7)}=-0{\cdot}1 \tag{7.12d}$$

و از آنجا

$$P_n(5) = -\tfrac{1}{18}.2 + \tfrac{32}{45}.13 + \tfrac{4}{9}.122 - 0\cdot1.504 = 12\cdot95 \tag{7.13}$$

محاسبات بالا نسبتا" پیچیده می‌باشند و اگر مقادیر درونیابی شده در تعداد زیادی نقطه لازم باشند به محاسبات قابل ملاحظه‌ای منجر می‌شود. همچنین، می‌توان مشاهده کرد که وارد نمودن نقاط بیشتر به روش، جهت افزایش دقت، تمام ضرائب را عوض خواهد کرد و در نتیجه تمام عملیات قبلی ازبین خواهد رفت. به این دلایل روش لاگرانژ مناسبترین فرم درونیابی برای کامپیوتر نمی‌باشد، هرچند شهرت زیادی به‌عنوان روشی برای ماشینهای رومیزی دارد.

۷.۲.۲ ■ تفاضلات تقسیم شده

در اکثر کتب درسی ریاضی دربارهٔ آنالیز عددی ثابت شده است که اگر یک چند جمله‌ای در شرایط

$$P_n(x_i) = f_i, \qquad i = 0, 1, \ldots, n \tag{7.14}$$

صدق کند منحصر بفرد است، به‌شرط آنکه نقاط x_i متمایز باشند (مثلا"، ر.ک. ایساکسون و کلر Isaacson and Keller (1966)). بنابراین، مسئله عبارت ازتغییر نظم چند جمله‌ای لاگرانژ به فرم مناسبتر برای کامپیوتراست. بخصوص اضافه کردن مقادیر بیشتر به این روش باید بگونه‌ای ساده امکان‌پذیر باشد، بدون اینکه محاسبات قبلی مردود شناخته شوند.

فرض کنید چند جمله‌ای درونیابی چون $P_k(x)$ حداکثر از درجهٔ k، داریم که داده‌های واقع در نقاط $x_i\ (i = 0, 1, \ldots, k)$ را برازش می‌کند و مایلیم عنصر بعدی $P_{k+1}(x)$ را با افزودن نقطه‌درونیاب x_{k+1} تشکیل دهیم. چون خواسته شده که محاسبات قبلی احتیاج به تغییر نداشته باشند، شکلی جستجو می‌کنیم که در آن

$$P_{k+1}(x) = P_k(x) + p_{k+1}(x), \qquad k = 0, 1, \ldots, n-1 \tag{7.15}$$

حداکثر از درجهٔ $k+1$ باشد.

چون $P_k(x)$ و $p_{k+1}(x)$ در نقاط $x_i\ (i = 0, 1, \ldots, k)$ درونیابی می‌کنند داریم

$$P_{k+1}(x_i) = P_k(x_i), \qquad i = 0, 1, \ldots, k \tag{7.16}$$

وبنابراین با استفاده از معادلهٔ (7.15)، $p_{k+1}(x_i) = 0$ چون با این تحلیل تمام $k+1$ عامل آشکار شده‌اند فرم $P_{k+1}(x)$ به این صورت می‌باشد

$$p_{k+1}(x) = a_{k+1}(x - x_0)(x - x_1)\cdots(x - x_k) \tag{7.17}$$

چند جمله‌ای درونیاب که حاصل می‌شود به چند جمله‌ای درونیاب تفاضل تقسیم شدهٔ نیوتن معروف است و به فرم زیر است.

$$p_n(x)=a_0+a_1(x-x_0)+\cdots+a_n(x-x_0)(x-x_1)\cdots(x-x_{n-1}) \quad (7.18)$$

این چند جمله‌ای را می‌توان به‌فرم تو در تو نوشت که برای محاسبه مناسب است. مثلاً،

$$P_3(x)=[[a_3(x-x_2)+a_2](x-x_1)+a_1](x-x_0)+a_0 \quad (7.19)$$

خوشبختانه، ضرائب a_k رامی‌توان بسادگی کامل بوسیلهٔ ساختن جدولی مشابه جدول تفاضل متناهی تولید کرد. با جایگذاری مقادیر $x_i\ (i=0,1,\ldots,n-1)$ در معادلهٔ (7.18) و به‌کار بردن $P(x_i)=f_i$ داریم:

$$\begin{aligned}
f_0&=a_0\\
f_1&=a_0+(x_1-x_0)a_1\\
f_2&=a_0+(x_2-x_0)a_1+(x_2-x_0)(x_2-x_1)a_2\\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \quad (7.20)$$

ضرائب a_i بوسیلهٔ روابط

$$a_0=f_0 \quad (7.21a)$$

$$a_1=\frac{f_1-f_0}{x_1-x_0} \quad (7.21b)$$

داده می‌شوند. این عبارت تفاضل تقسیم شدهٔ مرتبهٔ اول نامیده می‌شود و با $f[x_1,x_0]$ یا f_{01} نمایش داده می‌شود. تفاضل مرتبهٔ دوم برای نقاط x_0, x_1, x_2 ازمعادله (7.20) بدست می‌آید.

$$a_2=\frac{(f_2-f_0)/(x_2-x_0)-f[x_1,x_0]}{x_2-x_1}=\frac{f[x_2,x_0]-f[x_1,x_0]}{x_2-x_1} \quad (7.21c)$$

وبا $f[x_2,x_1,x_0]$ یا f_{012} نمایش داده می‌شوند. ضرائب دیگر با تفاضلات تقسیم شدهٔ مراتب بالای زیر داده می‌شود.

$$a_n=f[x_n,x_{n-1},\ldots,x_0]\equiv f_{01,\ldots,n} \quad (7.22)$$

خواص گوناگون تفاضلات تقسیم شده و قضایای مربوطه در هیلدبراند (1956) وایساکسون و کلر (1966) ارائه شده‌اند. کافی است در اینجا تذکر داده شود که تغییر ترتیب جملات مقدار تفاضل تقسیم شده را عوض نمی‌کند، و عبارت خطا برای درونیابی با استفاده از $P_n(x)$ در نقطهٔ دلخواه x عبارتستاز:

$$E(x)=(x-x_0)(x-x_1)\cdots(x-x_n)f[x_0,x_1,\ldots,x_n,x] \quad (7.23)$$

ومی‌توان نشان داد

$$f[x_0,x_1,\ldots,x_n,x]=\frac{f^{(n+1)}(\zeta)}{(n+1)!} \quad (7.24)$$

که در آن ζ نقطه‌ای در بازهٔ شامل x_i ها است.

ضرائب با ساختن یک جدول متناهی تفاضل تقسیم شده، همانطور که نشان داده شده،

محاسبه می‌شوند .

$$
\begin{array}{llll}
x_0 & f_0 & & \\
 & & \dfrac{f_1-f_0}{x_1-x_0}\equiv f_{01} & \\
x_1 & f_1 & & \dfrac{f_{12}-f_{01}}{x_2-x_0}\equiv f_{012} \\
 & & \dfrac{f_2-f_1}{x_2-x_1}\equiv f_{12} & \\
x_2 & f_2 & & \dfrac{f_{23}-f_{12}}{x_3-x_1}\equiv f_{123} \\
 & & \dfrac{f_3-f_2}{x_3-x_2}\equiv f_{23} & \\
x_3 & f_3 & & \dfrac{f_{34}-f_{23}}{x_4-x_2}\equiv f_{234} \\
 & & \dfrac{f_4-f_3}{x_4-x_3}\equiv f_{34} & \\
x_4 & f_4 & &
\end{array}
\tag{7.25}
$$

پس فرمی بدست آورده‌ایم که در آن ضرائب بسادگی حساب می‌شوند و فرم نهایی محاسبهٔ سریعی برای مقادیر زیادی از x ها توسط الگوریتم ضرب تو در تو بدست می‌دهد ، و نقاط درونیابی بیشتر را می‌توان جهت بهتر کردن دقت به طرح اضافه نمود . جدول تفاضل تقسیم شده نشان می‌دهد که افزودن یک نقطهٔ دیگر فقط خط قطری دیگری در زیر جدول وارد می‌کند که ضریب اضافی a_{k+1} را بدون تغییر مجموعهٔ قبلی می‌دهد . این مطلب در مثال (۷۰۱) نشان داده شده است .

۷.۲.۳ ■ درونیابی تکراری

طرح بالا در دو مرحله عمل می‌کند ، اول ، جدول تفاضل تقسیم شده برای محاسبهٔ ضرائب کشیده می‌شود ، و دوم ، ضرب تو در تو برای پیدا کردن مقادیر درونیابی شده به‌کار می‌رود . این روند هنگامی‌که مقادیر درونیابی شده در تعداد زیادی از نقاط خواسته شوند مناسبترین است ، اما درحالتی که فقط چند مقدارخواسته شود حجم محاسبات را می‌توان با به‌کار بردن روشهای درونیابی تکراری ، مانند روشهای ایتکن و نویل Aitken and Neville تقلیل داد .

مبنای این روشها ساختن جدولی از چند جمله‌ایهاست ، به فرم مشابه جدول تفاضل تقسیم شده ، که در آن ستونهای متوالی شامل چند جمله‌ایهای مراتب بالاتر می‌باشند که با حرکت در عرض جدول داده‌های مفروض را در تعداد بیشتری از نقاط برازش می‌کند . در محاسبهٔ مقادیر درونیابی شده ، خود چند جمله‌ایها به‌کار نمی‌روند ، یک مقدار خاص از x درج می‌شود و جدولی از مقادیر متناظر با چند جمله‌ایهای متعدد تولید می‌شود . درجهٔ چند جمله‌ایها در عرض جدول افزایش می‌یابد و خوشبختانه ، دقت مقادیر نیز زیاد می‌شود . بنابراین،اگر فرایند همگرا باشد نتایج سمت راست جدول

سرانجام در حد دقت مطلوب مساوی می‌شوند. در این نقطه محاسبه متوقف می‌شود. جـدول درونیـابی خطی چند جمله‌ایهای قبلی برحسب فرمول زیر ساخته می‌شوند.

$$P_{r_1,r_2,\ldots,r_k,i,j}(x)=\frac{(x-x_i)P_{r_1,r_2,\ldots,r_k,j}-(x-x_j)P_{r_1,r_2,\ldots,r_k,i}}{(x_j-x_i)} \qquad (7.26)$$

در آغاز $P_0(x)=f_0$ و $P_1(x)=f_1$ وچند آرایه اول جدول اینکن عبارتند از

$$P_{01}(x)=\frac{(x-x_1)P_0(x)-(x-x_0)P_1(x)}{x_0-x_1} \qquad (7.27a)$$

$$P_{02}(x)=\frac{(x-x_2)P_0(x)-(x-x_0)P_2(x)}{x_0-x_2} \qquad (7.27b)$$

$$P_{012}(x)=\frac{(x-x_2)P_{01}(x)-(x-x_1)P_{02}(x)}{(x_1-x_2)} \qquad (7.27c)$$

جدول کامل فرم زیر را دارد:

$$\begin{array}{lllll}
x_0 & f_0 & & & \\
 & & P_{01}(x) & & \\
x_1 & f_1 & & P_{012}(x) & \\
 & & P_{02}(x) & & \\
x_2 & f_2 & & \ddots & \\
 & & P_{03}(x) & & \\
x_3 & f_3 & & & \\
\cdots & \cdots & \cdots & \cdots & \\
x_k & f_k & P_{0k}(x) & P_{01k}(x)\cdots P_{01\ldots k}(x) &
\end{array} \qquad (7.28)$$

اگر مقدار تابع دیگری براین جدول اضافه شود فقط با به‌کار بردن مقادیر قطری که از f_0 شروع می‌شود سطر جدید محاسبه می‌شود. این نتیجه می‌دهد که فقط مقادیر زیر این قطر احتیاج به ذخیره شدن دارند، زیرا اعداد دیگر تنها مقادیر میانی هستند.

جدول نویل Neville برهمـان فرمولهـا استـوار است اما ترکیبهای مختلفـی از نقاط جهت تشکیل ستونها به‌کار می‌رود. این در جدول زیر نشان داده شده است.

$$\begin{array}{lllll}
x_0 & f_0 & & & \\
 & & P_{01}(x) & & \\
x_1 & f_1 & & P_{012}(x) & \\
 & & P_{12}(x) & \cdots & \\
x_2 & f_2 & & & \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
x_k & f_k & P_{k-1,k}(x) & P_{k-2,k-1,k}(x) & P_{012\ldots k}(x)
\end{array} \qquad (7.29)$$

وقتی نقطهٔ جدیدی به این جدول وارد می‌شود محاسبات تنها شامل مقادیری از سطر قبلی است، به‌طوری که احتیاج به ذخیرهٔ بقیه جدول نیست. یک محاسبهٔ عددی با استفاده از جدول نویل در مثال (۷.۲) داده شده است.

روش درونیابی تکراری یا روش تفاضل تقسیم شدهٔ نیوتن هر دو برای درونیابی

با استفاده از کامپیوتر مناسب هستند. در روش تکراری راهنمایی از دقت روش از درجهٔ نزدیکی اعضای متوالی جدول بسادگی بدست می‌آید، اما هر درونیابی محاسبهٔ تمام جدول تفاضل تقسیم شده را دربر دارد. اگر نقاط زیادی باید درونیابی شوند فرم نیوتن مناسبتر است زیرا فقط یک محاسبه از جدول انجام می‌شود. لذا، هر درونیابی تنها یک ضرب تو در تو نیاز دارد که محاسبهٔ کمی را دربر می‌گیرد. معهذا، باید بررسیهای مقدماتی جهت تعیین تعداد نقاط درونیاب لازم که دقت کافی بدست دهند، انجام شود.

۷.۳ ■ برازش بوسیلهٔ روش کمترین مربّعات

خاصیت اساسی این روش این است که مجموع مربعات خطا را حتی‌الامکان کوچک می‌کند. برحسب‌اینکه یک مجموعهٔ متناهی از مقادیر را تقریب می‌کنیم یا یک تابع تعریف شده بریک حوزه را، دو حالت اتفاق می‌افتد. در حالت اول خطا به‌صورت مجموع مربعات خطاها در هر نقطه تعریف می‌شود، و در حالت دوم یک فرمول‌بندی انتگرالی لازم می‌باشد. فرمول‌بندی اخیر جهت تشریح مبنای نظری روش به‌کار خواهد رفت چرا که اغلب خوانندگان با حساب انتگرال بیشتر آشنایی دارند تا خواص جمعبندی.

روش کمترین مربعات، درحالت کلی آن، برتابع تقریبی استوار است‌که بطور خطی بر مجموعه‌ای از پارامترهای $a_0, a_1, \ldots, a_n$ وابسته است. انتگرال جمع مربعات خطا با رابطهٔ

$$S=\int_a^b [f(x)-\phi(a_0, a_1, \ldots, a_n, x)]^2\,dx \tag{7.30}$$

داده می‌شود. چون می‌خواهیم S می‌نیمم باشد مشتقات اول نسبت به ضرائب مختلف صفر خواهد بود. یعنی،

$$\frac{\partial S}{\partial a_i}=0 \tag{7.31}$$

اگر شرایط مناسب جهت دیفرانسیل‌گیری زیر علامت انتگرال برقرار باشد، این معادلات $n+1$ معادله برای ضرائب a_i می‌دهد یعنی،

$$-2\int_a^b \frac{\partial \phi}{\partial a_i}[f(x)-\phi(a_0, a_1, \ldots, a_n, x)]\,dx=0, \qquad i=0, 1, \ldots, n \tag{7.32}$$

چون ϕ تابعی خطی از ضرائب است اولین جملهٔ این معادلات ثابت هستند طوری که معادلات را می‌توان نوشت.

$$\int_a^b \frac{\partial \phi}{\partial a_i}\phi(a_0, a_1, \ldots, a_n, x)\,dx=\int_a^b \frac{\partial \phi}{\partial a_i}f(x)\,dx, \qquad i=0, 1, \ldots, n \tag{7.33}$$

اینها معادلات نرمال نامیده می‌شوند .

به‌عنوان توضیح ساده‌ای از روش، حالتی را بررسی می‌کنیم که در آن‌تقریب ،چند جمله‌ای انتخاب شده است .

$$\phi(a_0,\dots,a_n,x)=a_0+a_1x+\cdots+a_nx^n \tag{7.34}$$

در این صورت معادلات نرمال عبارتنداز :

$$\int_a^b x^i[a_0+a_1x+\cdots+a_nx^n]\,dx=\int_a^b x^i f(x)\,dx,\quad i=0,1,\dots,n \tag{7.35}$$

اگر این معادلات به فرم ماتریسی نوشته شوند

$$\begin{aligned} &u_{00}a_0+u_{01}a_1+\cdots+u_{0n}a_n=b_0\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &u_{n0}a_0+u_{n1}a_1+\cdots+u_{nn}a_n=b_n \end{aligned} \tag{7.36}$$

آنگاه ضرائب u_{ij} در طرف چپ ماتریس $\mathbf{U}$ با رابطهٔ

$$u_{ij}=\int_a^b x^i\,.\,x^j\,dx,\qquad i,j=0,1,\dots,n \tag{7.37}$$

داده می‌شوند .

متأسفانه ، این معادلات می‌توانند بسیار بد وضع باشند و حل مستقیم مسئله کمترین مربعات بوسیلهٔ این روشها توصیه نمی‌شود . مثال بالا را با $a=0$ ، $b=+1.$ درنظر بگیرید . ضرائب با رابطهٔ

$$u_{ij}=1/(i+j+1) \tag{7.38}$$

داده می‌شوند .

و

$$\mathbf{U}=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \vdots\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \vdots\\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \vdots\\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \vdots\\ \cdots & \cdots & \cdots & \cdots & \end{bmatrix} \tag{7.39}$$

این ماتریس، ماتریس‌هیلبرت است که مشهور به منجر شدن به معادلات بد وضع است . این مطلب با تعیین معکوس عددی یک نوع از این ماتریسها در مثال(۷.۶) نشان داده شده است .

به‌طور ایده‌آل، مایلیم معادلات نرمال شکل ساده‌ای داشته و جواب مؤثری به مسئله بدهد . ساده‌ترین شکل ممکن شکل قطری است ، که تعیین ضرائب $a_0,\dots,a_n$ را بطور مستقیم با تقسیم کردن بر $u_{ii}\ (i=0,1,\dots,n)$ میسر می‌کند . این شکل را می‌توان با استفاده از خواص ویژه توابع متعامد تولید کرد .

خواص اساسی توابع متعامد این است که به ازای دو عضو مختلف از دنبالهٔ متعامد $Q_k(x)\ (k=0,1,2,\dots)$

$$\int_a^b w(x)\,.\,Q_i(x)Q_j(x)\,dx=0, \qquad i\neq j \tag{7.40}$$

به شرط آنکه حدود a و b وتابع وزن $w(x)$ بطور مناسبی انتخاب شده‌باشد .بالاخص، اگر انتخاب کنیم $w(x)=1$ ، $a=-1$ ، $b=+1$ آنگاه چنـــد جمله‌ایهـای لژاندر Legendre رابه‌کار می‌بریم . یعنی

$$\phi(a_0,a_1,\dots,a_n,x)=a_0\bar{P}_0(x)+a_1\bar{P}_1(x)+\cdots+a_n\bar{P}_n(x). \tag{7.41}$$

چند جملهٔ اول این نوع چند جمله‌ایها عبارتنداز:

$$\bar{P}_0(x)=1, \qquad \bar{P}_1(x)=x, \qquad \bar{P}_2(x)=\tfrac{1}{2}(3x^2-1) \tag{7.42}$$

حال متناظر با معادلات (7.36) مجموعه‌ای تنها با جملات قطری داریم زیرا ،

$$u_{ij}=\int_{-1}^{+1}\bar{P}_i(x)\bar{P}_j(x)\,dx=0, \qquad \begin{array}{l} i=0,1,\dots,n \\ j\neq i \end{array} \tag{7.43}$$

عناصر قطری نیزشکل سادهٔ ویژه‌ای دارند .

$$u_{ii}=\int_{-1}^{+1}\bar{P}_i^2(x)\,dx=\frac{2}{2i+1} \tag{7.44}$$

بطوری که a_i ها مستقیما" بدست می‌آیند .

$$a_i=\frac{2i+1}{2}\int_{-1}^{+1}\bar{P}_i(x)f(x)\,dx, \qquad i=0,1,\dots,n \tag{7.45}$$

این دو روش متفاوت از محاسبه در مثال(۷۰۳) نشان داده شده‌است . اگر نمایش بصورت یک سری توانی از x لازم باشد می‌توان عبارات مربوط به $\bar{P}_i(x)$ را ، وقتی مقادیر a_i بدست آمدند ، در معادلات جایگذاری کرد و با تغییر نظم ، سری توانی از x را بدست آورد .

۷.٤ ■ چند جمله ایهای متعامد

بخاطر اهمیت چند جمله‌ایهای متعامد در مسائل تقریبی در اینجا بعضی از خواص مهمتر این چند جمله‌ایها را خلاصه می‌کنیم . ما دنباله‌ای از چند جمله‌ایهای $Q_k(x)$ از درجهٔ k بشرط اینکه $k\leqslant K$ به ازای مقدار ثابتی از K ،داریم .

۷.٤.۱ ■ رابطهٔ تعامد

چند جمله‌ایهای $Q_k(x)$ نسبت به تابع وزن $w(x)$ در بازهٔ $[a,b]$ متعامد گفتـــه

می‌شوند اگر

$$\int_a^b w(x)Q_i(x)Q_j(x)\,dx=0, \qquad i\neq j$$
$$w(x)\geqslant 0, \qquad a\leqslant x\leqslant b \tag{7.46}$$

بعلاوه، اگر شرایط زیر برقرار باشد چند جمله‌ایها یک مجموعهٔ متعامد نرمال تشکیل می‌دهند.

$$\int_a^b w(x)Q_i^2(x)\,dx=1 \qquad (\text{به ازای تمام مقادیر } i) \tag{7.47}$$

باید متذکر شد که این تعریف نتیجه می‌دهد که $Q_i(x)$ **بر هر چند جمله‌ای از درجه کوچکتر** از i، عمود است. این بدلیل آنست که هر چند جمله‌ای را می‌توان برحسب مجموعهٔ $Q_k(x)\ (k=0,1,\dots,i)$ بیان کرد چرا که آنها مستقل خطی هستند، یعنی،

$$P_j(x)=\sum_{k=0}^{j}a_kQ_k(x), \qquad j<i \tag{7.48}$$

ملاحظه می‌شود که:

$$\int_a^b w(x)Q_i(x)P_j(x)\,dx=\sum_{k=0}^{j}a_k\int_a^b w(x)Q_i(x)Q_k(x)\,dx=0 \tag{7.49}$$

تمام جملات بنابر خاصیت تعامد صفر هستند به‌طوری که $P_j(x)$ بر $Q_i(x)$ عمود است $(j<i)$. این خاصیت تعامد وسیله‌ای برای تولید عناصر متوالی دنباله بدست می‌دهد، اگرچه خاصیت بازگشتی که ذیلا" بحث می‌شود روش مناسبتری برای اجرای عملی می‌دهد.

بعنوان یک مثال، ضرائب چند جمله‌ای لژاندار درجهٔ دوم را می‌توان به‌طریق زیر بدست آورد. رابطهٔ تعامد برای چند جمله‌ای ax^2+bx+c می‌دهد

$$\int_{-1}^{+1}(ax^2+bx+c).1.dx=0, \qquad \frac{2a}{3}+2c=0 \tag{7.50}$$

و

$$\int_{-1}^{+1}(ax^2+bx+c).x\,dx=0, \qquad \frac{2b}{3}=0 \tag{7.51}$$

بنابراین، چند جمله‌ای لژاندر درجهٔ دوم $c(-3x^2+1)$ است. رابطهٔ تعامد به ازای تمام مقادیر c برقرار است و انتخاب این ثابت به اطلاعات عمیقتری از نظریهٔ توابع متعامد بستگی دارد. شکل معمولی این چند جمله‌ای لژاندر $\frac{1}{2}(3x^2-1)$ است. اما اگر مجموعه‌ای متعامد نرمال خواسته شود این چند جمله‌ای شکل $(\sqrt{5}/2\sqrt{2})(3x^2-1)$ را

می‌گیرد .

۷.۴.۲ ■ رابطهٔ بازگشتی

می‌توان ملاحظه کرد که روش بالا جهت پیدا کردن ضرائب برای چند جمله‌ایهای مراتب بالا غیر عملی است ،چون شامل حل دستگاهی بزرگ از معادلات می‌شود . خوشبختانه، خاصیت دیگری از توابع متعامد روش ساده‌تری، که برای استفاده از کامپیوتر مناسب است ، برای تولید آنها بدست می‌دهد . تمام مجموعه‌های متعامد دریک رابطهٔ بازگشتی به‌شکل زیر صدق می‌کنند .

$$Q_{n+1}(x) = (A_n + B_n x) Q_n(x) + C_n Q_{n-1}(x), \qquad n = 1, 2, \dots \tag{7.52}$$

بنابراین ، اگر دو چند جمله‌ای اول مجموعه معلوم باشند ، سومی را می‌توان از معادلهٔ بالا پیدا کرد ، و بعد دومی و سومی برای تولید عضو چهارم به‌کار می‌روند وبه‌همین‌ترتیب. برای چند جمله‌ایهای لژاندر رابطهٔ بازگشتی عبارتست‌از :

$$\bar{P}_{n+1}(x) = \frac{2n+1}{n+1} \, . \, x\bar{P}_n(x) - \frac{n}{n+1} \bar{P}_{n-1}(x) \tag{7.53}$$

دو عضو اول دنباله $\bar{P}_0(x) = 1$ $\bar{P}_1(x) = x$ هستند و بنابراین ،

$$\bar{P}_2(x) = \tfrac{3}{2} \, . \, x(x) - \tfrac{1}{2} \, . \, 1 = \tfrac{1}{2}(3x^2 - 1) \tag{7.54a}$$

$$\bar{P}_3(x) = \tfrac{5}{3} \, . \, x(\tfrac{3}{2}x^2 - \tfrac{1}{2}) - \tfrac{2}{3}(x) = \tfrac{1}{2}(5x^3 - 3x) \tag{7.54b}$$

۷.۴.۳ ■ تعامد گسسته

تااینجانظرخودرابه حالتی که در آن یک تابع پیوسته باید تقریب شود معطوف کرده‌ایم ، بطوری که خاصیت تعامد انتگرالی مناسب است . اما ، اگر مجموعه‌ای گسسته‌از مقادیر تابع موجود باشد مناسبتر است که خطا را فقط برحسب این نقاط تعریف کنیم ، یعنی ،

$$S_m = \sum_{i=0}^{m} [\phi_n(x_i) - f_i]^2 \tag{7.55}$$

باید بخاطر آوریم که اگر $Q_n(x)$ یک چند جمله‌ای از درجهٔ n باشد شامل $n+1$ ضریب به‌عنوان پارامتر متغیر است ، لذا به‌ازای $m \geqslant n+1$ جواب منحصر بفردی‌برای مسئله کمترین مربعات وجود دارد . تقریب کمترین مربعات معمولاً " دارای نقاط بسیـــار بیشتر m ، از تعداد ضرائب چند جمله‌ای، $n+1$ ، است .

اگر دوباره عبارتی برحسب چند جمله‌ایهای متعامد به‌کار بریم داریم :

$$\phi_n(x) = \sum_{k=0}^{n} a_k Q_k(x) \tag{7.56}$$

و با به‌کار بردن شرایط زیر برای یک می‌نیمم

$$\frac{\partial S_m}{\partial a_r}=0 \tag{7.57}$$

معادلات نرمال زیر تولید می‌شود .

$$\sum_{k=0}^{n} a_k \sum_{i=0}^{m} Q_k(x_i)Q_r(x_i)=\sum_{i=0}^{m} Q_r(x_i)f_i, \qquad r=0,1,\ldots,n \tag{7.58}$$

اگر توابع در خاصیت تعامد گسسته زیر صدق کنند

$$\sum_{i=0}^{m} Q_k(x_i)Q_r(x_i)=0, \qquad k\neq r \tag{7.59}$$

آنگاه مجموعه‌ای از معادلات قطری حاصل می‌شود که جوابهای زیر را دارند

$$a_r=\frac{\sum_{i=0}^{m} Q_r(x_i)f_i}{\sum_{i=0}^{m} Q_r^2(x_i)} \tag{7.60}$$

تمام چند جمله‌ایهای متعامد استاندارد مجموعه‌ای از نقاط $x_i\ (i=0,\ldots,m)$ دارند که برآن خاصیت تعامد گسسته برقرار است . چند جمله‌ایهایی که برای نشان دادن این خواص به‌کار می‌روند چند جمله‌ایهای چبیشف هستند ، که در رابطه با تقریب مینیماکس نیز مفید می‌باشند وبا

$$T_n(x)=\cos[n\cos^{-1}(x)], \qquad -1\leqslant x\leqslant +1 \tag{7.61}$$

تعریف می‌شوند و صفرهای این چند جمله‌ای با

$$x_j=\cos\left[\frac{(2j+1)\pi}{2n}\right], \qquad j=0,1,\ldots,n-1 \tag{7.62}$$

داده می‌شوند . به‌ازای یک‌مقدارثابت N تمام‌چندجمله‌ایهای‌چبیشف $T_n(x)\ (n<N)$ دارای خاصیت تعامد متناهی نسبت به مجموعه نقاط زیر هستند

$$x_j=\cos\left(\frac{\pi j}{N}\right), \qquad j=0,1,\ldots,N \tag{7.63}$$

مسئله اصلی مربوط به استفاده از تقریب کمترین مربعات گسسته آنست که باید قادر باشیم مقادیر تابع را در هر نقطهٔ x که فرمول لازم دارد بدست آوریم ، و این ممکن است امکان نداشته باشد . همچنین‌اگر لازم باشد درجهٔ تابع تقریب را افزایش دهیم مشکلاتی وجود دارد . درحالت پیوسته محاسبات تنها شامل پیدا کردن کمیتهای

$$a_k=\frac{\int_a^b Q_k(x)f(x)\,dx}{\int_a^b Q_k^2(x)\,dx} \tag{7.64}$$

برای مقادیر بعدی k است و این مزاحم مقادیر قبلی نمی‌شود .
در حالت گسسته جمعبندی شامل محاسبهٔ f_i دریک مجموعهٔ جدید از مقادیر x است که ، عموما " به ازای مقادیر مختلف m متفاوت می‌باشند . یکی از مزایای فرمولبندی چبیشف آن است که x_j های مرتبهٔ m در فرمولهای مراتب $2m$ و $4m$ و غیره نیز موجودند بقسمی که بعضی f_i ها از مقادیر محاسبه شدهٔ قبلی معلومند .

نکتهٔ دیگری که محتاج فکر است انتخاب مقادیر n و m است . بطور کلی ، اگر تعداد نقاط m افزایش یابد ، به قیمت محاسبات بیشتری ، تقریب نزدیکتری بدست می‌آید ، ضمنا ، باید بین دقت و زمان محاسبه تناسبی باشد . انتخاب مناسب‌مقادیر بسته به مسئلهٔ خاص دارد و قاعدهٔ کلی نمی‌توان داد .

تحقیق در انتخاب یک مقدار مناسب برای n بینش فوق‌العاده‌ای از طبیعت‌تقریبات کمترین مربعات می‌دهد . وقتی $n=m$ چند جمله‌ای درونیاب ، که از $n+1$ نقطه به‌کار رفته می‌گذرد ، حاصل می‌شود . به‌ازای مقادیر $n<m$ برازش‌کنندهٔ‌کمترین‌مربعات بطور طبیعی از نقاط عبور نمی‌کند ، و منحنی مقید به فرایند هموارسازی است . این به هنگامی که روش بر نتایج آزمایشی که مقادیر تابعی را با خطای آزمایش می دهند با ارزش است . انحرافات جزئی در اثر خطاها می‌تواند یک چند جمله‌ای کاملا " نوسانی نتیجه دهد که اصولا " اختلافات ناشی از خطاها را گسترش می‌دهد . باز هم باید سازشی بین کاربرد یک مقدار کوچک n ، که حتی اختلافات بارز توابع را هموار می‌کند ، و استفاده از یک مقدار خیلی بزرگ n به هنگامی که خطای آزمایشی قابل هموار شدن نیست ، انجام گیرد . با منحنی‌هایی که انتظار می‌رود نسبتا " هموار باشند تقریب را می‌توان با تجزیه حوزه به زیر حوزه‌هایی بدست آورد . سپس در هر زیر حوزه تقریب کمترین‌مربعاتی استوار بر یک چند جمله‌ای با درجهٔ پایین ،مثلا " درجهٔ ۳ ، به‌کار برد .

۷.۴.۴ ■ ریشه های چند جمله ایهای چبیشف

اگر مجموعه‌ای از چند جمله‌ایهای متعامد $Q_k(x)\ (k=0,1,\ldots)$ بربازه $[a,b]$ داشته باشیم ریشه‌های $x_j\ (j=1,2,\ldots k)$ از $Q_k(x)=0$ همه حقیقی و متمایزند و دربازهٔ $[a,b]$ قرار دارند .
این خاصیت به هنگام استفاده از صفرهای یک چند جمله‌ای متعامد به‌عنوان نقاط پایه برای یک تقریب کمترین مربعات گسسته ، همچنانکه در بخش قبل توصیه شد ، لازم است . اما ، اگر ریشه‌های مکرر یا مختلط وجود داشتند ممکن نبود که اینها را در روش‌کمترین مربعات به‌کار بریم.

۷.۴.۵ ■ چند جمله ایهای متعامد مهم

قبلا" دیده‌ایم که چگونه چند جمله‌ایهای لژاندر بطور طبیعی در نظریهٔ کمترین مربعات، اگر خاصیت‌تعامد شامل یک تابع وزن واحد باشد، ظاهر می‌شوند. محدودیت کاربردی آنها را به حوزه $[-1,+1]$ بسادگی می‌توان با یک تغییر متغیر مرتفع ساخت. اگر $x\epsilon[-1,+1]$ و $X\epsilon[a,b]$ آنگاه معادلهٔ $X=(x+1)(b-a)/2+a$ باعث تبدیل به‌حوزهٔ استاندارد $[-1,+1]$ خواهد شد. این تغییر متغیردر نظریهٔ کوادراتور گاوسی gaussian quadrature مهم است زیرا به‌این معناست که هر حوزهٔ متناهی را می‌توان بوسیلهٔ بحث در حوزهٔ استاندارد $[-1,+1]$ بررسی کرد. معهذا، تبدیلاتی مانند بالا نمی‌توانند بازه نیم‌محدود$[0,\infty[$ یا بازه نامحدود$]-\infty,+\infty[$ را دربرگیرند. چند جمله‌ایهای لاگور بر بازه $0,\infty$ متعامد هستند و می‌توانند برای‌انتگرال‌گیری گاوسی براین بازه به‌کار روند. چند جمله‌ایهای هرمیت‌بر$]-\infty,+\infty[$متعامد هستند. مجموعهٔ دیگری از چند جمله‌ایهای متعامد که زیاد به‌کار می‌روند. چند جمله‌ایهای چبیشف هستند که بر حوزهٔ $[-1,+1]$ با تابع وزن $(1-x^2)^{-1/2}$ متعامد هستند. با ارزش‌ترین خاصیت، نوسان مساوی آنهاست. تمام مقادیر ماکزیمم و می‌نیمم تابع در $[-1,+1]$ اتفاق می‌افتد و همه دارای یک اندازهٔ مطلق هستند. دربخش بعد نشان داده خواهد شد که این مطلب به‌خاصیت مینیماکس منجر خواهد شد.

۷.۵ ■ چند جمله ایهای چبیشف وتقریب مینیماکس

۷.۵.۱ ■ تعریف چند جمله ایها

چند جمله‌ایهای چبیشف حالت خاصی از چند جمله‌ایهای ژاکوبی Jacobi با $p=q=\frac{1}{2}$ هستند. آنها بربازهٔ $[-1,+1]$ با تابع وزن $(1-x^2)^{-1/2}$ متعامد هستند. ساده‌ترین معرفی از این چند جمله ایها ازطریق تابع کسینوس، $\cos n\theta \; (0\leqslant\theta\leqslant\pi)$ می‌باشد. تابع $\cos n\theta$ رامی‌توان با استفاده از فرمولهای مثلثاتی بسط داد، یعنی

$$\cos n\theta=\cos[(n-1)\theta]\cos\theta-\sin[(n-1)\theta]\sin\theta \qquad (7.65)$$

و سپس جملات مرتبهٔ $n-1$ را می‌توان دوباره بسط داده و بهمین‌ترتیب، تا تمام عبارت تنها شامل θ باشد و تمام مضارب θ بسط داده شده باشند. قوای زوج $\sin\theta$ که ظاهر می‌شوند را می‌توان با $1-\cos^2\theta$ جایگزین نمود بطوری که شامل قوای $\cos\theta$ باشد.

$$\cos 2\theta=2\cos^2\theta-1 \qquad (7.66a)$$

$$\begin{aligned}\cos 3\theta&=\cos\theta[2\cos^2\theta-1]-\sin\theta\,.\sin 2\theta\\&=2\cos^3\theta-\cos\theta-2\cos\theta[1-\cos^2\theta]\\&=4\cos^3\theta-3\cos\theta \qquad (7.66b)\end{aligned}$$

اگر جایگذاری $x=\cos\theta$ انجام شود توابع $\cos n\theta \equiv \cos(n\cos^{-1}x)$ معادل چند جمله‌ایهای درجهٔ n برای $-1 \leqslant x \leqslant +1$ می‌باشند. بنابراین، یک مجموعه از چند جمله‌ایها، معروف به چند جمله‌ایهای چبیشف را می‌توان با عبارت زیر تعریف نمود.

$$T_n(x)=\cos(n\cos^{-1}x), \qquad -1 \leqslant x \leqslant 1 \tag{7.67}$$

این ارتباط با توابع مثلثاتی وسیلهٔ مفیدی برای اثبات برخی از خواص چند جمله‌ایهای چبیشف است که با دوباره فرمولبندی کردن مسئله برحسب θ غالبا" به اثبات ساده‌ای منجر می‌شود. مثلا"، تعامد انتگرالی حکم می‌کند که،

$$\int_{-1}^{+1}\frac{T_r(x)T_s(x)\,dx}{(1-x^2)^{1/2}}=0, \qquad r\neq s \tag{7.68}$$

انجام تبدیل $x=\cos\theta$ می‌دهد $dx=-\sin\theta\, d\theta$ و انتگرال می‌شود:

$$\int_{-\pi}^{0}\cos r\theta.\cos s\theta\, d\theta=\tfrac{1}{2}\int_{-\pi}^{0}\cos[(r+s)\theta]+\cos[(r-s)\theta]\,d\theta \tag{7.69}$$

انتگرال $\cos n\theta$، که در آن n صحیح است، برحوزهٔ $-\pi \leqslant \theta \leqslant 0$ صفر است مگر اینکه $n=0$ بطوری‌که هر دو قسمت انتگرال صفر خواهد شد بجز وقتی که $r=s$ بنابراین، خاصیت تعامد ثابت می‌شود.

رابطه بازگشتی برای چند جمله‌ایهای چبیشف را می‌توان بهمین طریق بدست آورد.

$$\begin{aligned} T_{n+1}(x) &\equiv \cos(n+1)\theta=\cos n\theta\cos\theta-\sin n\theta\sin\theta \\ T_{n-1}(x) &\equiv \cos(n-1)\theta=\cos n\theta\cos\theta+\sin n\theta\sin\theta \end{aligned} \tag{7.70}$$

$$T_{n+1}(x)+T_{n-1}(x)=2x\cos n\theta \quad \text{یا} \quad T_{n+1}(x)=2xT_n(x)-T_{n-1}(x) \tag{7.71}$$

۷.۵.۲ ■ خاصیت مینیماکس

مهمترین خاصیت چند جمله‌ایهای چبیشف خاصیت مینیماکس می‌باشد که نتیجه‌ای از خاصیت نوسان مساوی است. چند جمله‌ای چبیشف $T_n(x)/2^{n-1}$ دربارهٔ $[-1,+1]$ دارای کمترین مقدار مطلق، نسبت به تمام چند جمله‌ایهای درجهٔ n که‌ضریب بزرگترین توان x آنها واحد است، می‌باشد. این بدین معنی است که کلیه چند جمله‌ایهای دیگر بیشتر از صفر دور می‌شوند، در نقطه‌ای در حوزه‌اش، تا چند جمله‌ای چبیشف نرمال شده که بین $\pm(\frac{1}{2})^{n-1}$ نوسان می‌کند. این خاصیت بسادگی از خواص تابع کسینوس ثابت می‌شود. $\cos n\theta$ در نقاط

$$\theta_j=\frac{j\pi}{n}, \qquad j=0,1,\ldots,n \tag{7.72}$$

(جدول (۷۰۱))

	$w(x)$	a	b	*First term*	*Second term*	*Third term*	(رابطهٔ بازگشتی)
Jacobi	$(1-x)^{-p}(1+x)^{-q}$ $p<1, q<1$	-1	$+1$				
Legendre	1 $p=0, q=0$	-1	$+1$	1	x	$\frac{1}{2}(3x^2-1)$	$P_{n+1}(x)=\frac{2n+1}{n+1}.xP_n(x)-\frac{n}{n+1}P_{n-1}(x)$
Chebyshev	$(1-x^2)^{-1/2}$ $p=\frac{1}{2}, q=\frac{1}{2}$	-1	$+1$	1	x	$2x^2-1$	$T_{n+1}(x)=2xT_n(x)-T_{n-1}(x)$
Laguerre	$e^{-\alpha x}$	0	∞	1	$1-x$	$2-4x+x^2$	$L_{r+1}(x)=(1+2r-x)L_r(x)-r^2L_{r-1}(x)$
Hermite	$e^{-\alpha^2x^2}$	$-\infty$	$+\infty$	1	$2x$	$4x^2-2$	$H_{r+1}(x)=2xH_r(x)-2rH_{r-1}(x)$

که مقادیر ± 1 را می‌گیرد ماکزیمم و می‌نیمم دارد، بطوری که چند جمله‌ای چبیشف (نرمال شده) متناوبا" مقادیر $\pm(\frac{1}{2})^{n-1}$ را می‌گیرد فرض کنید یک چندجمله‌ای چون $p_n(x)$ وجود دارد، با ضریب بزرگترین توان x واحد، که مقادیرش در تمام نقاط حوزهٔ $-1 \leqslant x \leqslant +1$ ازنظر قدرمطلق کوچکتر از $(\frac{1}{2})^{n-1}$ است. نشان‌خواهیم‌داد که این فرض منجر به تناقض خواهد شد. تابع زیر را تشکیل دهید.

$$\phi_{n-1}(x) = p_n(x) - \frac{T_n(x)}{2^{n-1}} \tag{7.73}$$

اندیس $n-1$ جهت تاکید اینکه دو جملهٔ x^n حذف می‌شوند، چون هر دو چند جمله‌ای نرمال شده‌اند طوری که ضریب پیشرو آنها واحد باشد، به‌کار رفته است.

در هر نقطهٔ $\theta_j = j\pi/n$ که j زوج باشد، می‌دانیم که $\phi_{n-1}(x)$ منفی است زیرا مقدار $-T_n(x)/2^{n-1}$ برابر $-1/2^{n-1}$ است که بنا بفرض ازنظر قدرمطلق بیشتر از مقدار $p_n(x)$ است. بهمین‌ترتیب، در نقاط θ_j که j فرد است مقدار $\phi_{n-1}(x)$ باید مثبت باشد چون $-T_n(x)/2^{n-1}$ ماکزیمم مثبت خود یعنی $(\frac{1}{2})^{n-1}$ را می‌گیرد.

بنابراین، چند جمله‌ای $\phi_{n-1}(x)$ در $n+1$ نقطه θ_j بین این مقادیر مثبت و منفی نوسان می‌کند. چون این تابع پیوسته است باید n ریشهٔ حقیقی و متمایز بین این نقاط داشته باشد. این بدین معنی است که $\phi_{n-1}(x)$، یک چند جمله‌ای درجهٔ $n-1$ دارای n صفر است. این ممکن نیست مگر اینکه $p_n(x) \equiv \frac{T_n(x)}{2^{n-1}}$. لذا، فرض اولیه باید نادرست باشد.

باید اشاره شود که نتیجهٔ بدست آمده براین حقیقت استوار بود که چندجمله‌ای‌متناوبا" $n+1$ بار به ماکزیمم و می‌نیمم فاصلهٔ خود تا صفر می‌رسد. مسئلهٔ کلی‌تر مینیماکس یا تقریب چبیشف پیدا کردن یک تابع است که دارای خاصیت نوسان مساوی نسبت به تابع دیگری تا صفرباشد. نشان دادیم که چند جمله‌ای چبیشف بهترین تقریب به صفر به مفهوم مینیماکس است. اما ممکن است بهترین تقریب به توابع دیگری نظیر تابع سینوس یا لگاریتمی خواسته شود. این بحث خارج از حیطهٔ این کتاب است اما، نشان داده خواهد شد که تقریبهای مفیدی می‌توان برمبنای چند جمله‌ایهای چبیشف، هر چند نه بهترین تقریبها به مفهوم مینیماکس، بدست آورد.

۷.۵.۳ ■ اقتصادی کردن چند جمله ایها

خاصیت بالا از چند جمله‌ایهای چبیشف را می‌توان برای بدست آوردن بهترین تقریب چند جمله‌ای از درجهٔ $n-1$ به چند جمله‌ای مفروضی از درجهٔ n به‌کار برد. اگر

چند جمله‌ای مفروض

$$p_n(x) = a_0 + a_1x + \cdots + a_nx^n, \qquad -1 \leqslant x \leqslant +1 \tag{7.74}$$

باشد چند جمله‌ای

$$q_{n-1}(x) = p_n(x) - \frac{a_n}{2^{n-1}} T_n(x) \tag{7.75}$$

را تشکیل می‌دهیم . می‌توان مشاهده کرد که $q_{n-1}(x)$ یک چند جمله‌ای با ماکزیمم درجه $n-1$ است ، زیرا ضریب x^n طوری انتخاب شده که صفر باشد . اختلاف بین دو چند جمله‌ای $q_{n-1}(x)$ و $p_n(x)$ مضربی از چند جمله‌ای چبیشف است و بنابراین ، دربارهٔ $-1 \leqslant x \leqslant +1$ با کمترین انحراف از صفر می‌باشد .

اقتصادی کردن بر حذف جملهٔ با بزرگترین درجه بوسیلهٔ تفریق مضربی از چند جمله‌ای چبیشف مناسب متکی است .

وقتی چند جمله‌ای جدید $q_{n-1}(x)$ پیدا می‌شود تشکیل بهترین تقریب به آن به همان روش امکان‌پذیر است .

$$q_{n-2}(x) = q_{n-1}(x) - \frac{b_{n-1}}{2^{n-2}} T_{n-1}(x) \tag{7.76}$$

هر عضو دنباله بهترین تقریب ، عضو قبلی خواهد بود اما باید درنظر داشت که تقریب $p_n(x)$ ، به مفهوم مینیماکس، فقط به ازای چند جمله‌ای $q_{n-1}(x)$ ، بهترین تقریب خواهد بود . مثلا " ، خطای تقریب دوم با رابطهٔ

$$q_{n-2}(x) - p_n(x) = -\frac{b_{n-1}}{2^{n-2}} T_{n-1}(x) - \frac{a_n}{2^{n-1}} T_n(x) \tag{7.77}$$

داده می‌شود و مجموع دو چند جمله‌ای چبیشف دارای خاصیت مینیماکس نخواهند بود . معهذا ، در عمل معلوم شده که این فرایند اقتصادی کردن تقریب خوبی از درجهٔ پایین می‌دهد .

ماکزیمم خطای ممکن در این تقریب بسادگی پیدا می‌شود ، چون $T_r(x)$ ماکزیمم ومی‌نیمم ± 1 دارد . پس، برای $q_{n-2}(x)$ خطایی داریم که ماکزیمم مطلقی کمتر از مقدار زیر دارد .

$$\left|\frac{b_{n-1}}{2^{n-2}}\right| + \left|\frac{a_n}{2^{n-1}}\right|$$

بسط زیر را درنظر بگیرید

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} \cdots, \qquad -1 < x \leqslant +1 \tag{7.78}$$

این بسط خیلی بکندی همگراست ، در حقیقت ، این اغلب در مورد بسط سری تیلر یک مسئله است . بنابراین، حذف دو جملهٔ آخر از تقریب فوق خطایی وارد می‌کند که

ماکزیمم مطلقی کمتر از $\frac{1}{4}+\frac{1}{5}=\frac{9}{20}$ دارد.
تقریب بالا را می‌توان برحسب چند جمله‌ایهای چبیشف، با به‌کار بردن عبارات زیر که بسادگی از فرمولهای چند جمله‌ای چبیشف بدست می‌آیند، بیان کرد.

$$\begin{aligned} 1 &= T_0(x) & x &= T_1(x) \\ x^2 &= \tfrac{1}{2}[T_0(x)+T_2(x)] & x^3 &= \tfrac{1}{4}[3T_1(x)+T_3(x)] \\ x^4 &= \tfrac{1}{8}[3T_0(x)+4T_2(x)+T_4(x)] & x^5 &= \tfrac{1}{16}[10T_1(x)+5T_3(x)+T_5(x)] \end{aligned} \tag{7.79}$$

$$\log(1+x) = -\tfrac{11}{32}T_0(x)+\tfrac{11}{8}T_1(x)-\tfrac{3}{8}T_2(x)+\tfrac{7}{48}T_3(x)-\tfrac{1}{32}T_4(x)+\tfrac{1}{80}T_5(x)-\cdots \tag{7.80}$$

حذف دو جملهٔ آخر از این تقریب خطایی با ماکزیمم مطلق کمتر از $\frac{1}{32}+\frac{1}{80}=\frac{7}{160}$ وارد می‌کند که به مقدار قابل ملاحظه‌ای بهتر از نتیجهٔ قبلی است. خواننده در خواهد یافت که این مثال برای سادگی انتخاب شده بود و اعداد مربوطه غیر واقعی بودند، اما اصل اقتصادی کردن را می‌توان با بهترین‌کارآیی جهت بدست آوردن تقریب چند جمله‌ای با درجهٔ پایین و ازدست دادن دقت نسبتا" کم، به‌کار برد (مثال ۷۰۷ نیز مشاهده شود).

۷.۵.۴ ■ بسط سری چبیشف

به‌عنوان جانشینی برای بسط بوسیلهٔ سری تیلر بهمراه اقتصادی کردن، ممکن است مستقیما" بصورت بسط چبیشف بسط دهیم. هر تابع $f(x)$ که در بازه $[-1,+1]$ پیوسته و دارای تعداد متناهی ماکزیمم و می‌نیمم باشد را می‌توان به این صورت بسط داد، یعنی

$$f(x) = \sum_{k=0}^{\infty}{}' a_k T_k(x) \tag{7.81}$$

علامت پریم مشخص کنندهٔ آن است که جملهٔ اول $a_0/2$ می‌باشد، که تعریف ضرائب a_r را به‌گونه‌ای مشابه با ضرائب فوریه Fourier ساده می‌کند.
این ضرائب می‌توانندبوسیلهٔ خاصیت تعامد استاندارد محاسبه شوند. معادلهٔ (7.81) را در $(1-x^2)^{-1/2}.T_s(x)$ ضرب نمائید و سپس بین -1 و $+1$ انتگرال بگیرید. تمام جملات طرف راست، بجز یکی، صفر می‌شوند که می‌دهد:

$$a_s = \frac{2}{\pi}\int_{-1}^{+1} \frac{T_s(x).f(x)\,dx}{(1-x^2)^{1/2}} \tag{7.82}$$

البته لازم خواهد بود که این انتگرال را، درصورتی که از روشهای کامپیوتری استفاده می‌کنیم، به طرق تقریبی حساب کنیم. لذا، ترجیح داده می‌شود که طرحی مبتنی بریک

سری متناهی داشته باشیم که از خواص تعامد متناهی استفاده کند . فرض کنید مقادیر $f(x)$ در نقاط $x_j=\cos(\pi j/n)\ (j=0,1,\ldots,n)$ معلوم باشد . تابع

$$\phi(x)=\sum_{k=0}^{n}{}'' b_k T_k(x) \tag{7.83}$$

را طوری انتخاب کنید که مقادیرش در نقاط $x_j\ (j=0,1,\ldots,n)$ همان مقادیر معلوم تابع باشد یعنی :

$$f(x_j)=\phi(x_j)=\sum_{k=0}^{n}{}'' b_k T_k(x_j), \qquad i=0,1,\ldots,n \tag{7.84}$$

علامت Σ'' مشخص می‌کند که اولین و آخرین جمله ، به‌ترتیب ، ضریب $b_n/2$ و $b_0/2$ دارند .

خاصیت تعامد متناهی با رابطهٔ

$$\sum_{j=0}^{n}{}'' T_r(x_j)T_s(x_j)=\begin{cases} n & r=s=0,n \\ n/2 & r=s\neq 0 \\ 0 & r\neq s \end{cases} \tag{7.85}$$

تعریف می‌شود .

ضرائب b_r را می‌توان با ضرب هریک از معادلات (7.84) در مقدار مناسب $T_s(x_j)\ (j=0,1,\ldots,n)$ و جمع کردن این معادلات بدست آورد . خاصیت تعامد بدین معنی است که تمام جملات بجز یکی در طرف راست صفر می‌شود ،که می‌دهد :

$$b_s=\frac{2}{n}\sum_{j=0}^{n}{}'' f(x_j)\,.\,T_s(x_j) \tag{7.86}$$

واضح است که این بسط را با کامپیوتر می‌توان مستقیما" محاسبه کرد . یک بسط بسری چبیشف ساده در مثال (۷۰۴) داده شده است .

بطور طبیعی ، دانستن اینکه ارتباط بین ضرائب تعریف شده در بسط چبیشف متناهی ، داده شده با $\phi(x)$ و ضرائب سری نامتناهی چقدر نزدیک است ، جالب می‌باشد . اگر بسط سری نامتناهی $f(x)$ را در معادلهٔ (7.86) درج کنیم داریم :

$$b_s=\frac{2}{n}\sum_{j=0}^{n}{}''\sum_{k=0}^{\infty}{}' a_k T_k(x_j)T_s(x_j) \tag{7.87}$$

با استفاده از تعریف چند جمله‌ایهای چبیشف $T_k(x)=\cos k\theta$ و $\cos\theta=x$ می‌توانیم خواص تابع کسینوس را برای نشان دادن اینکه تمام جملات $k=2Mn\pm s$ درخاصیت تعامد متناهی (7.85) صدق می‌کنند ، اگر $M\geqslant 1$ عدد صحیح مثبتی باشد ، به‌کار بریم .
این نتیجه را داریم زیرا $\cos\frac{k\,.\,\pi j}{n}=\cos\frac{k\pi s}{n}$ به ازای مقادیری از k که

$$\frac{k\pi j}{n}=2\pi M'\pm\frac{k\pi s}{n} \quad \text{یا} \quad k=\frac{2M'n}{j}\pm s \tag{7.88}$$

چون k باید صحیح باشد به آن مقادیر M' مقید هستیم که مضربی از j هستند: که می‌دهند $k=2n.M\pm s$ لذا، تنها این مقادیر k مقادیر غیر صفر درطرف راست را ناشی می‌شوند و این نتیجه می‌دهد:

$$b_s=a_s+a_{2n-s}+a_{2n+s}+a_{4n-s}+a_{4n+s}+\cdots \tag{7.89}$$

اگر این سری خواص همگرائی خوبی داشته باشد جملات مراتب بالاتر بزودی قابل اغماض شده، و b_s ها تقریب خوبی به ضرائب سری نامتناهی خواهند بود.

۷.۵.۵ ■ محاسبهٔ سری چبیشف

در فصل ۳ اشاره شد که محاسبهٔ چند جمله‌ایها باید با استفاده ازالگوریتم ضرب تو در تو انجام گیرد. لذا، یک طریقه ممکن جهت محاسبهٔ یک سری چبیشف متناهی (7.81) محاسبهٔ ضرائب قوای متفاوت x و استفاده از این الگوریتم است. اما، الگوریتمی وجود دارد که می‌تواند مستقیما" روی سری چبیشف به‌کار رود و خیلی شبیه به الگوریتم ضرب تو در تو است. این روش ترجیح داده می‌شود. زیرا، از تغییر نظم برای مرتب کردن برحسب قوای x اجتناب می‌کند. مثالی عددی بوسیلهٔ مثال (۷۰۵) داده شده است. اگر

$$c_{n+1}=c_{n+2}=0$$

و رابطه بازگشتی

$$c_r=2xc_{r+1}-c_{r+2}+b_r, \qquad r=n,n-1,\ldots,0 \tag{7.90}$$

را به‌کار برید مقدار سری چبیشف (7.81) در نقطهٔ x با رابطهٔ

$$\phi(x)=\tfrac{1}{2}[c_0-c_2] \tag{7.91}$$

داده می‌شود.

۷.۵.۶ ■ خواص دیگر

خاصیت نوسان مساوی چبیشف، هنگامی که این نقاط بعنوان مبنائی برای یک فرمول درونیابی انتخاب شوند، کاربرد جالبی دارد. در بخش ۷۰۲۰۲ نشان داده شده است که خطای فرمول درونیابی با به‌کار بردن نقاط $x_0,\ldots,x_n$ با رابطهٔ

$$(x-x_0)(x-x_1)\ldots(x-x_n)\frac{f^{(n+1)}(\zeta)}{(n+1)!} \tag{7.92}$$

داده می‌شود که در آن ζ نقطه‌ای است که در بازه‌ای که $n+1$ نقطه در آن می‌باشندقرار

دارد. طبیعتا" ، مایلیم که این عبارت خطا حتی‌الامکان کوچک باشد. باوجوداینکه کار کمی می‌توان برای می‌نیمم ساختن عبارت مشتق انجام داد یقینا" ممکن است که با انتخاب نقاط x_i $(i=0,1,\dots,n)$ به‌عنوان صفرهای چند جمله‌ای چبیشف $T_{n+1}(x)$ نوسان می‌نیمم را برای حاصلضرب جملات بدست آورد.

همچنین جالب است که چند جمله‌ایهای چبیشف در فرمولهای انتگرال‌گیری گاوس به‌کار می‌روند. آنها به فرمولی منجر می‌شوند که ضرائب مساوی دارد، که کمی از محاسبات لازم می‌کاهد. همچنین، حضور فاکتور وزنی $(1-x^2)^{-1/2}$ بدین معنی است که انتگرالهای مشخصی با قطبهایی در انتگراند آنها را نیز می‌توان بطور عددی محاسبه کرد. تفصیلات بیشتر در بخش ۸.۳ داده شده است.

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ ذیلا" یک جدول تفاضل تقسیم شده برای مقادیر تابعی زیر داده شده است.

x	$f(x)$			
1·6	0·6250			
2·9	0·3448	−0·2155		
3·7	0·2703	−0·0931	+0·0582	
4·8	0·2083	−0·0563	+0·0193	−0·0121

فرمول درونیابی مشتق شده از یک جدول تفاضل متناهی شکل زیر را دارد.

$$P_3(x)=a_0+a_1(x-x_0)+a_2(x-x_0)(x-x_1)+a_3(x-x_0)(x-x_1)(x-x_2)$$
$$=((a_3(x-x_2)+a_2)(x-x_1)+a_1)(x-x_0)+a_0$$

که در آن a_i $(i=0,1,2,3)$ به‌ترتیب تفاضلات تقسیم شده هستند. یعنی، $a_0=0{\cdot}6250$ ، $a_1=-0{\cdot}2155$ ، $a_2=0{\cdot}0582,$ ، $a_3=-0{\cdot}0121$. شکل دوم چندجمله‌ای‌درونیاب به شکل ضرب تودرتوبوده که برای محاسبات عددی مناسب است. مقادیر درونیابی شده در $x=2{\cdot}0$ و $x=4{\cdot}5$ بوسیلهٔ ضرب تودرتو به ترتیب 0·5105 و 0·2254 را بدست می‌دهند.

اگر دقت بالاتر لازم باشد یک نقطهٔ بیشتر می‌توان اضافه کرد که یک‌سطربه‌جدول تفاضل تقسیم شده اضافه می‌کند. فرض کنید نقطهٔ جدید $x=4{\cdot}0$ و $f(x)=0{\cdot}25$ باشد سطر جدید عبارتست‌از:

4·0 0·25 −0·0521 0·0140 −0·0048 0·0030

پس، ضریب اضافی $a_4=0{\cdot}0030$ و ضرب‌تودرتومقادیردرونیابی‌شده‌درنقاط $x=2{\cdot}0$ و $x=4{\cdot}5$ را به‌ترتیب 0·5058 و 0·2221 می‌دهد. مقادیردرست‌تابع 0·5 و 0·2222 هستند.

۲ــ جدول نویل برای مقادیر تابعی فوق در نقطه $x=2{\cdot}0$ در زیر نشان داده شده است.

سطر اضافی از مقادیر برای $x=4\cdot0$ رامی‌توان بعد از محاسبهٔ جدول اصلی بسادگی اضافه کرد.

x	$f(x)$			
1·6	0·6250			
2·9	0·3448	0·5387		
3·7	0·2703	0·5574	0·5176	
4·8	0·2083	0·5729	0·5224	0·5101

مؤلفهٔ اول جدول از فرمول

$$\frac{(x-x_0)f_1-(x-x_1)f_0}{x_1-x_0}=\frac{0\cdot4\times0\cdot3448+0\cdot9\times0\cdot6250}{2\cdot9-1\cdot6}$$
$$=0\cdot5387$$

بدست می‌آید. مولفهٔ اول از ستون سوم از فرمول

$$\frac{(x-x_0)f_{12}-(x-x_2)f_{01}}{x_2-x_0}=\frac{0\cdot4\times0\cdot5574+1\cdot7\times0\cdot5387}{3\cdot7-1\cdot6}$$
$$=0\cdot5176$$

محاسبه می‌شود. اضافه کردن‌نقطه $x=4\cdot0$ و $f(x)=0\cdot25$ سطر اضافی زیر را می‌دهد.

4·0 0·25 0·5625 0·5192 0·5086 0·5047

لذا دقیقترین مقدار در جدول اول 0·5101 است‌وسطراضافی‌مقدار 0·5047 رامی‌دهد.

۳ـ جدول زیر از مقادیر بایستی بوسیلهٔ یک چند جمله‌ای درجهٔ سوم و با استفاده از روش کمترین مربعات تقریب شود.

x	-3	-2	-1	0	1	2	3
$f(x)$	-27	-16	-9	1	8	18	26

عبارت خطای S با رابطهٔ

$$S=\sum_{i=0}^{6}[a_0+a_1x_i+a_2x_i^2+a_3x_i^3-f_i]^2$$

داده شده است. اعضای ماتریسی که برای پیدا کردن‌ضرائب $a_i\ (i=0,1,\ldots,3)$ به‌کار می‌رود عبارتست‌از:

$$\begin{matrix} u_0 & u_1 & u_2 & u_3\\ u_1 & u_2 & u_3 & u_4\\ u_2 & u_3 & u_4 & u_5\\ u_3 & u_4 & u_5 & u_6\end{matrix}$$

که در آن $u_r=\sum_{i=0}^{6}x_i^r$. جملات طرف راست $b_r=\sum_{i=0}^{6}x_i^rf_i$ می‌باشند. محاسبه اعضای ماتریس معادلات زیر را می‌دهد.

$$\begin{bmatrix}7&0&28&0\\0&28&0&196\\28&0&196&0\\0&196&0&1588\end{bmatrix}\begin{bmatrix}a_0\\a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}1\\244\\-2\\1720\end{bmatrix}$$

تفریق چهار برابر معادلهٔ اول از معادلهٔ سوم نتیجه می‌دهد

$$(196-112)a_2=-6$$

بنابراین :

$$a_2=-\tfrac{1}{14},\qquad a_0=\tfrac{3}{7}$$

تفریق هفت برابر معادلهٔ اول از معادلهٔ سوم نتیجه می‌دهد

$$(1588-1372)a_3=1720-1708$$

لذا ،

$$a_3=\tfrac{1}{18},\qquad a_1=\tfrac{1}{28}(244-\tfrac{196}{18})=\tfrac{1049}{126}$$

بنابراین ، چند جمله‌ای کمترین مربعات خواسته شده عبارتست از :

$$P_3(x)=\frac{3}{7}+\frac{1049}{126}x-\frac{x^2}{14}+\frac{x^3}{18}$$

این‌مسئله بوسیلهٔ شکل ماتریسی بسیار ساده شده بود . درحالت کلی جواب باید بوسیلهٔ حذف گاوسی کامل بدست آید . این مسئله را می‌توان با به‌کار بردن چند جمله‌ایهایی که براین مجموعهٔ خاص از نقاط متعامدند نیز حل نمود ، یعنی

$$Q_0(x)=1,\qquad Q_1(x)=x,\qquad Q_2(x)=x^2-4,\qquad Q_3(x)=x^3-7x$$

بنابراین ، با قرار دادن

$$P_3(x)=a_0Q_0(x)+a_1Q_1(x)+a_2Q_2(x)+a_3Q_3(x)$$

معادلات این‌چنین می‌شوند :

$$\begin{aligned}
7a_0 &= -27-16-9+1+8+18+26=+1\\
28a_1 &= 81+32+9+8+36+78=244\\
84a_2 &= -6\\
216a_3 &= 12
\end{aligned}$$

$$a_0=\tfrac{1}{7},\qquad a_1=\tfrac{61}{7},\qquad a_2=-\tfrac{1}{14},\qquad a_3=\tfrac{1}{18}$$

لذا ، چندجمله‌ای‌خواسته‌شده‌عبارتست‌از :

$$P_3(x)=\tfrac{1}{7}+\tfrac{61}{7}x-\frac{(x^2-4)}{14}+\tfrac{1}{18}(x^3-7x)$$

بعهدهٔ خواننده است که تحقیق کند این دو جواب یکسان هستند .

معادلات قطری را می‌توان مستقیما " حل کرد . با تعداد زیادی معادله ، اگر حذف گاوسی کامل برای روش چند جمله‌ای معمولی لازم باشد برای روش چند جمله‌ای متعامد حل ساده‌تری نتیجه می‌شود .

۴ـ یک سری چبیشف متناهی با چهار جمله پیدا کنید که تابع e^x را در حوزهٔ $[-1,+1]$ تقریب نماید . مایلیم ضرائب عبارت

$$\phi(x)=\frac{b_0}{2}T_0(x)+b_1T_1(x)+b_2T_2(x)+\frac{b_3}{2}T_3(x)$$

را با مفروض بودن مقادیر در نقاط مناسب برای تعامد متناهی ، یعنی

$x_j = \cos(\pi j/n), j = 0, 1, 2, 3$

پیداکنیم . مقادیر این تابع و چند جمله‌ایهای‌چبیشف دراین نقاط عبارتنداز :

x	$f(x)$	$T_0(x)$	$T_1(x)$	$T_2(x)$	$T_3(x)$
1	2·718 282	1	1	1	+1
$\frac{1}{2}$	1·648 721	1	$\frac{1}{2}$	$-\frac{1}{2}$	−1
$-\frac{1}{2}$	0·606 530	1	$-\frac{1}{2}$	$-\frac{1}{2}$	+1
−1	0·367 879	1	−1	1	−1

رابطهٔ تعامد متناهی عبارتست‌از :

$$\sum_{i=0}^{3}{}'' T_r(x_j)T_s(x_j) = \begin{cases} 0 & r \neq s \\ \dfrac{2}{n} & r = s = 1, 2 \\ n & r = s = 0, 3 \end{cases}$$

که در آن پریم مضاعف مشخص کننــدهٔ آن است که جملهٔ اول و آخــر در $\frac{1}{2}$ ضــرب شده‌اند . پس، فرمول ضرائب عبارتست‌از :

$$b_s = \frac{2}{n}\sum_{j=0}^{n}{}'' f(x_j)T_s(x_j)$$

اولین و آخرین ضریب b_0 و b_3 فاکتور وزنی $\frac{1}{2}$ دارند بقسمی که تنها یک فــرمــول می‌توان برای تعریف تمام ضرائب به‌کار برد .

$$b_0 = \frac{2}{3}\left[\frac{e}{2} + e^{1/2} + e^{-1/2} + \frac{e^{-1}}{2}\right] = 2{\cdot}532\,221$$

$$b_1 = \frac{2}{3}\left[\frac{e}{2} + \frac{1}{2}e^{-1/2} - \frac{1}{2}e^{-1/2} - \frac{e^{-1}}{2}\right] = 1{\cdot}130\,865$$

$$b_2 = \frac{2}{3}\left[\frac{e}{2} - \frac{1}{2}e^{1/2} - \frac{1}{2}e^{-1/2} + \frac{e^{-1}}{2}\right] = 0{\cdot}276\,970$$

$$b_3 = \frac{2}{3}\left[\frac{e}{2} - e^{-1/2} + e^{-1/2} - \frac{e^{-1}}{2}\right] = 0{\cdot}088\,674$$

پس

$$\phi(x) = 1{\cdot}266\,110 + 1{\cdot}130\,865T_1(x) + 0{\cdot}276\,970T_2(x) + 0{\cdot}044\,337T_3(x)$$

روش ضرب تو در تو را جهت محاسبهٔ سری چبیشف می‌توان به‌کار برد . فرمولهای داده شده در معادلات (7.90) و (7.91) به بسط‌چبیشف با $b_0/2$ به عنوان جملهٔ اول مربوط می‌شود ولی جمله آخر b_n است نه $b_n/2$. پس مقدار b_3 که باید به‌کار رود 0·044 337 است . اما مقدار b_0 در هر دو حالت 2·532 222 1 است .

بعنوان یک مثال در نقطهٔ $x = 1$

$$\begin{aligned} c_3 &= 0 - 0 + 0{\cdot}044\,337 \\ c_2 &= 2 \times 0{\cdot}044\,337 - 0 + 0{\cdot}276\,970 = 0{\cdot}365\,644 \\ c_1 &= 2 \times 0{\cdot}365\,644 - 0{\cdot}044\,337 + 1{\cdot}130\,865 = 1{\cdot}817\,816 \\ c_0 &= 2 \times 1{\cdot}817\,816 - 0{\cdot}365\,644 + 2{\cdot}532\,221 = 5{\cdot}802\,209 \end{aligned}$$

$$\begin{aligned} \text{Answer} &= \tfrac{1}{2}[5{\cdot}802\,209 - 0{\cdot}365\,644] \\ &= 2{\cdot}718\,282 \end{aligned}$$

۵ـ سری چبیشف فوق را در نقاط $x = 0, 0{\cdot}25, -0{\cdot}25$ حساب کنید و جوابها را با مقادیر واقعی، یعنی $0{\cdot}778\,801$ و $1{\cdot}284\,025$ و 1 مقایسه کنید.

$x = 0$

$$c_3 = 0 - 0 + 0{\cdot}044\,337$$
$$c_2 = 0 - 0 + 0{\cdot}276\,970$$
$$c_1 = 0 - 0{\cdot}044\,337 + 1{\cdot}130\,865 = 1{\cdot}086\,528$$
$$c_0 = 0 - 0{\cdot}276\,970 + 2{\cdot}532\,221 = 2{\cdot}255\,251$$

$$\text{Answer} = \tfrac{1}{2}[c_0 - c_2] = \frac{1{\cdot}978\,281}{2} = 0{\cdot}989\,140$$

$x = 0{\cdot}25$

$$c_3 = 0 - 0 + 0{\cdot}044\,337$$
$$c_2 = 0{\cdot}044\,337 \times \tfrac{1}{2} - 0 + 0{\cdot}276\,970 = 0{\cdot}299\,138$$
$$c_1 = 0{\cdot}299\,138 \times \tfrac{1}{2} - 0{\cdot}044\,337 + 1{\cdot}130\,865 = 1{\cdot}236\,097$$
$$c_0 = 1{\cdot}236\,097 \times \tfrac{1}{2} - 0{\cdot}299\,138 + 2{\cdot}532\,221 = 2{\cdot}851\,131$$

$$\text{Answer} = \tfrac{1}{2}[2{\cdot}851\,131 - 0{\cdot}299\,138] = 1{\cdot}275\,996$$

$x = -0{\cdot}25$

$$c_3 = 0 - 0 + 0{\cdot}044\,337$$
$$c_2 = -0{\cdot}044\,337 \times \tfrac{1}{2} - 0 + 0{\cdot}276\,970 = 0{\cdot}254\,802$$
$$c_1 = -0{\cdot}254\,802 \times \tfrac{1}{2} - 0{\cdot}044\,337 + 1{\cdot}130\,865 = 0{\cdot}959\,127$$
$$c_0 = -0{\cdot}959\,127 \times \tfrac{1}{2} - 0{\cdot}254\,802 + 2{\cdot}532\,221 = 1{\cdot}797\,855$$

$$\text{Answer} = \tfrac{1}{2}[1{\cdot}797\,855 - 0{\cdot}254\,802] = 0{\cdot}771\,527$$

۶ـ ماتریسهای هیلبرت برای حل دستگاه معادلات خطی، بدوضع هستند. این مطلب با محاسبه معکوس یک ماتریس هیلبرت سه‌درسه‌و به‌کار بردن شش رقم با معنی، که درجهٔ دقت خوبی می‌دهد، به‌کاربردن سه رقم با معنی، که نتیجه بسیار نادقیقی می‌دهد، نشان داده شده است. دترمینان ماتریس تقریبا" $0{\cdot}000\,463$ است که‌طبیعت بدوضع این ماتریس را نشان می‌دهد. جدول محاسبات با به‌کار بردن حذف گاوسی جهت پیدا کردن معکوس تا شش رقم با معنی عبارتست‌از:

1·000 000	0·500 000	0·333 333	1·000 00	0·000 00	0·000 00
0·500 000	0·333 333	0·250 000	0·000 00	1·000 00	0·000 00
0·333 333	0·250 000	0·200 000	0·000 00	0·000 00	1·000 00
	0·083 333 3	0·083 333 3	−0·500 000	1·000 00	0·000 00
	0·083 333 3	0·088 888 9	−0·333 333	0·000 00	1·000 00
		0·005 555 6	0·166 667	−1·000 00	1·000 00

با جایگذاری پسرو

$$R_3 = \quad 29{\cdot}9977 \quad -179{\cdot}986 \quad 179{\cdot}986$$
$$R_2 = -35{\cdot}9977 \quad 191{\cdot}986 \quad -179{\cdot}986$$
$$R_1 = \quad 8{\cdot}999\,62 \quad -35{\cdot}9977 \quad 29{\cdot}9977$$

نتایج را می‌توان با ضرب کردن آزمایش نمود.

$$\begin{bmatrix} 8{\cdot}999\,62 & -35{\cdot}9977 & 29{\cdot}9977 \\ -35{\cdot}9977 & 191{\cdot}986 & -179{\cdot}986 \\ 29{\cdot}9977 & -179{\cdot}986 & 179{\cdot}986 \end{bmatrix} \begin{bmatrix} 1{\cdot}000\,00 & 0{\cdot}500\,000 & 0{\cdot}333\,333 \\ 0{\cdot}500\,000 & 0{\cdot}333\,333 & 0{\cdot}250\,000 \\ 0{\cdot}333\,333 & 0{\cdot}250\,000 & 0{\cdot}200\,000 \end{bmatrix}$$

$$=\begin{bmatrix} 0.999\,993 & 0.000\,14 & -0.000\,015 \\ 0.000\,027 & 0.999\,919 & 0.000\,079 \\ -0.000\,027 & 0.000\,077 & 0.999\,923 \end{bmatrix}$$

جدول مشابهی تا سه‌رقم با معنی نتایجی می‌دهد که دقت بسیار کمی دارد.

1.00	0.500	0.333	1.00	0.00	0.00
0.500	0.333	0.250	0.00	1.00	0.00
0.333	0.250	0.200	0.00	0.00	1.00
	0.0833	0.0833	−0.500	1.00	0.00
	0.0833	0.0889	−0.333	0.00	1.00
		0.0056	0.167	−1.00	1.00

با جایگذاری پسرو

$$\begin{aligned} R_3 &= \quad 29.8 \quad -179 \quad +179 \\ R_2 &= -35.8 \quad 191 \quad -179 \\ R_1 &= \quad 8.98 \quad -35.9 \quad 29.9 \end{aligned}$$

نتایج با ضرب کردن آزمایش شده‌اند.

$$\begin{bmatrix} 8.98 & -35.9 & 29.9 \\ -35.8 & 191 & -179 \\ 29.8 & -179 & 179 \end{bmatrix}\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix}$$

$$=\begin{bmatrix} 0.987 & 0.0103 & -0.004\,66 \\ 0.0930 & 0.953 & 0.0286 \\ -0.0930 & 0.0430 & 0.973 \end{bmatrix}$$

اگر محاسبات فوق بجای سه رقم با معنی با سه رقم اعشار انجام شوند معکوس، و حاصلضرب معکوس در ماتریس اصلی به‌ترتیب چنین داده می‌شوند.

8.661	−33.858	27.834	1.001	0.015	−0.014
−33.857	178.715	−166.667	0.001	0.916	0.072
27.833	−166.667	166.667	−0.001	0.084	0.934

۷ــ سری‌تیلربرای e^x تا توان‌پنجم‌دربازه $-1\leqslant x\leqslant 1$ دارای‌خطای‌ماکزیمم $0.003\,775$ است. با بیان قوای x برحسب چند جمله‌ایهای چبیشف نشان دهید که اگر دو جملهٔ آخر بسط چبیشف حذف شوند خطای اضافی وارد شده از 0.006 کمتر است.

سری تیلر بریده truncated عبارتست‌از:

$$e^x=1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\frac{x^5}{120}$$

با خطای ماکزیمم $e^{1.0}/720=0.003\,775.$ رابطهٔ زیر برای توانهای مختلف x به کار می‌روند.

$$\begin{aligned} &1=T_0(x), \qquad x=T_1(x), \qquad x^2=\tfrac{1}{2}[T_2(x)+T_0(x)] \\ &x^3=\tfrac{1}{4}[T_3(x)+3T_1(x)], \qquad x^4=\tfrac{1}{8}[T_4(x)+4T(x)+3T_0(x)] \\ &x^5=\tfrac{1}{16}[T_5(x)+5T_3(x)+10T_1(x)] \end{aligned}$$

بنابراین،

$$\begin{aligned}e^x &= T_0(x) + T_1(x) + \tfrac{1}{4}[T_2(x) + T_0(x)] + \tfrac{1}{24}[T_3(x) + 3T_1(x)]\\ &\quad + \tfrac{1}{192}[T_4(x) + 4T_2(x) + 3T_0(x)] + \tfrac{1}{1920}[T_5(x) + 5T_3(x) + 10T_1(x)]\\ &= \tfrac{243}{192}T_0(x) + \tfrac{217}{182}T_1(x) + \tfrac{13}{48}T_2(x) + \tfrac{17}{384}T_3(x) + \tfrac{1}{192}T_4(x) + \tfrac{1}{1920}T_5(x)\end{aligned}$$

چون چند جمله‌ایهای چبیشف بین ± 1 نوسان می‌کنند ماکزیمم مطلق مجموع دو جملهٔ آخر عبارتست‌از :

$$0{\cdot}005\,208 + 0{\cdot}000\,520\,8 \approx 0{\cdot}005\,728$$

۸ ــ یک انتگرال مربوط به انتگرال بیضوی کامل با :

$$K(k) = \int_0^{\pi/2} \frac{dx}{[1 - (\sin k)^2 \sin^2 x]^{1/2}}$$

تعریف می‌شود . از یک جدول مقادیر این انتگرالها درمی‌یابیم که :

$$\begin{aligned}K(1) &= 1.5709\\ K(4) &= 1.5727\\ K(6) &= 1.5751\end{aligned}$$

با استفاده از درونیابی بوسیلهٔ یک چندجمله‌ای درجهٔ دوم تقریبی از $K(3.5)$ را بدست آورید .

داریم :

$$l_0(3.5) = \frac{(3.5-4)(3.5-6)}{(1-4)(1-6)} = \frac{1.25}{15} = 0.08333$$

$$l_1(3.5) = \frac{(3.5-1)(3.5-6)}{(4-1)(4-6)} = \frac{-6.25}{-6} = 1.04167$$

$$l_2(3.5) = \frac{(3.5-1)(3.5-4)}{(6-1)(6-4)} = \frac{-1.25}{10} = -0.12500$$

پس

$$\begin{aligned}K(3.5) &\approx (1.5709)(0.08333) + (1.5727)(1.04167) + (1.5751)(-0.12500)\\ &= 1.57225\end{aligned}$$

این تقریب درآخرین رقم خطا دارد .

۹ ــ مثال ۸ را با استفاده از فرمول نیوتن ، یعنی :

$$\boxed{p_n(x) = \sum_{i=0}^{n} f[x_0, \ldots, x_i] \prod_{j=0}^{i-1} (x - x_j)}$$

حل کنید .

در این مثال ، مجبوریم چندجمله‌ای از درجه نابیشتر از دو پیدا کنیم که در معادلات زیر صدق کند .

$$p_2(1) = 1.5709 \qquad p_2(4) = 1.5727 \qquad p_2(6) = 1.5751$$

داریم

$$K[1,4] = \frac{1.5709 - 1.5727}{1 - 4} = 0.0006$$

$$K[4,6] = \frac{1.5727 - 1.5751}{4 - 6} = 0.0012$$

بنابراین

$$K[1,4,6] = \frac{0.0006 - 0.0012}{1 - 6} = 0.00012$$

و در نتیجه

$$p_2(x) = 1.5709 + 0.0006(x - 1) + 0.00012(x - 1)(x - 4)$$

با جایگذاری $x = 3.5$ در معادلهٔ فوق بدست می‌آوریم :

$$p_2(3.5) = 1.5709 + (0.0006)(2.5) + (0.00012)(2.5)(-0.5) = 1.57225$$

که بر نتیجهٔ حاصل از مثال ۸ منطبق است .

۱۰ــ فرض کنید $f(x) = (1 + x^2)^{-1}$ ، به ازای $n = 2, 4, \ldots, 16$ چندجمله‌ای با درجهٔ نابیشتر از n را که $f(x)$ را در $n + 1$ نقطهٔ متساوی‌الفاصلهٔ

$$x_i = i\frac{10}{n} - 5 \qquad i = 0, \ldots, n$$

درونیابی می‌کند حساب کنید .

سپس تخمینی از ماکزیمم خطای درونیابی ، یعنی ،

$$E_n = \max_{-5 \leq x \leq 5} |f(x) - p_n(x)|$$

را با محاسبهٔ

$$E_n \approx \max_i |f(y_i) - p_n(y_i)|$$

که در آن

$$y_i = \frac{i}{10} - 5 \qquad i = 0, \ldots, 100$$

بدست آورید .

برنامهٔ فرترن زیر برای حل این مثال ارائه می‌شود . همانطور که خروجی نشان می‌دهد E_n ابتدا کمی نزول می‌کند ولی بعد ، همچنانکه n افزایش می‌یابد بسرعت صعود می‌کند . لذا ، برای این مثال ، درونیابی با یک چندجمله‌ای از درجهٔ پائین نتایج بهتری از درونیابی با یک چندجمله‌ای با درجهٔ بالا بدست می‌دهد .

FORTRAN PROGRAM FOR EXAMPLE 10

```
      DIMENSION X(17),D(17)
      F(X) = 1./(1. + X*X)
      WRITE (6,600)
  600 FORMAT(1H1,3X1HN5X13HMAXIMUM ERROR)
      DO 40 N = 2,16,2
```

```
      NP1 = N + 1
      H = 10./FLOAT(N)
      DO 10 I = 1,NP1
      X(I) = FLOAT(I - 1)*H - 5.
   10 D(I) = F(X(I))
CALCULATE THE COEFFICIENTS OF INTERP. POL. BY ALGORITHM 4.3
      DO 20 K = 1,N
      NP1MK = NP1 - K
      DO 20 I = 1,NP1MK
      IPK = I + K
   20 D(I) = (D(I + 1) - D(I))/(X(IPK) - X(I))
C     ESTIMATE MAXIMUM INTERPOLATION ERROR ON (-5,5).
      Y = -5.
      ERRMX = 0.
      DO 30 J = 1,101
CALCULATE PN(Y) BY ALGORITHM 4.1
      PNOFY = D(1)
      DO 29 K = 2,NP1
   29 PNOFY = D(K) + (Y - X(K))*PNOFY
      ERROR = ABS(F(Y) - PNOFY)
      IF (ERROR .GT. ERRMX) ERRMX = ERROR
   30 Y = Y + .1
   40 WRITE (6,640) N,ERRMX
  640 FORMAT(I5,E18.7)
                                          STOP
      END
```

COMPUTER OUTPUT FOR EXAMPLE 10

N	MAXIMUM ERROR
2	6.4615385E−01
4	4.3813387E−01
6	6.1666759E−01
8	1.0451739E+00
10	1.9156431E+00
12	3.6052745E+00
14	7.1920080E+00
16	1.4051542E+01

۱۱ ــ اگر خطای درونیابی $f(x)$ در نقاط $x_0, \ldots, x_n$ بوسیلهٔ $P_n(x)$ در نقطهٔ $\bar{x}$ با

$$e_n(\bar{x}) = f(\bar{x}) - p_n(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_{j=0}^{n}(\bar{x} - x_j)$$

داده شده باشد.

طول فاصلهٔ h را در یک جدول با نقاط متساوی‌الفاصله برای تابع $f(x) = \sqrt{x}$ بین 1 و 2 چنان تعیین کنید که درونیابی با چندجمله‌ای درجهٔ دوم خطائی کمتر از 5×10^{-8} داشته باشد.

بنا بفرض، جدول مورد نظر شامل $f(x_i)$ $x_i = 1 + ih,\ i = 0, \ldots, N,$ است که $N = (2-1)/h$. اگر $\bar{x} \in [x_{i-1}, x_{i+1}]$ مقدار $f(\bar{x})$ را با $p_2(\bar{x})$ تقریب می‌زنیم که در آن $p_2(x)$ چندجمله‌ای درجه‌دومی است که $f(x)$ را در x_{i-1}, x_i, x_{i+1}. درونیابی می‌کند . لذا ، خطا عبارتست از :

$$f(\bar{x}) - p_2(\bar{x}) = (\bar{x} - x_{i-1})(\bar{x} - x_i)(\bar{x} - x_{i+1}) \frac{f'''(\xi)}{3!}$$

به ازای ξ ای در (x_{i-1}, x_{i+1}) . چون ξ را نمی‌شناسیم می‌توانیم صرفا $f'''(\xi)$ را برآورد کنیم .

$$|f'''(\xi)| \leq \max_{1 \leq x \leq 2} |f'''(x)|$$

با محاسبه: $f'''(x) = \frac{3}{8} x^{-5/2}$ داریم $|f'''(\xi)| \leq \frac{3}{8}$. بعلاوه، با تغییر متغیر خطی $y = x - x_i$ ،

$$\max_{x \in [x_{i-1}, x_{i+1}]} |(x - x_{i-1})(x - x_i)(x - x_{i+1})| = \max_{y \in [-h, h]} |(y+h)y(y-h)|$$
$$= \max_{y \in [-h, h]} |y(y^2 - h^2)|$$

چون تابع $\psi(y) = y(y^2 - h^2)$ در نقاط $y = -h$ و $y = h$ صفر می‌شود ماکزیمم $|\psi(y)|$ بر $[-h, h]$ باید در یکی از نقاط اکسترمم $y = \pm h/\sqrt{3}$ اتفاق افتد . این نقاط با حل‌معادلهٔ $\psi'(y) = 3y^2 - h^2 = 0$ بدست می‌آیند که می‌دهد $y = \pm h/\sqrt{3}$ بنابراین

$$\max_{x \in [x_{i-1}, x_{i+1}]} |(x - x_{i-1})(x - x_i)(x - x_{i+1})| = \frac{2h^3}{3\sqrt{3}}$$

اکنون مطمئن شده‌ایم که به ازای هر $\bar{x} \in [1,2]$، ، در صورتیکه $p_2(x)$ چنان انتخاب شده باشد که $f(x)$ را در سه نقطه جدولی نزدیک $\bar{x}$ درونیابی کند ،

$$|f(\bar{x}) - p_2(\bar{x})| \leq \frac{(2h^3/3\sqrt{3})(3/8)}{6} = \frac{h^3}{24\sqrt{3}}$$

لذا ، قرار می‌دهیم

$$\frac{h^3}{24\sqrt{3}} < 5 \cdot 10^{-8}$$

که می‌دهد $h \approx 0.0128$ یا $N \approx 79$

■ مسائل

۱ـ درونیابی لاگرانژی را برای بدست آوردن مقادیر تابع در نقاط $2\cdot1, 1\cdot9, 1\cdot2$ پیدا نمائید ، جدول مقادیر زیر مفروض است .

x	1·0	1·5	2·6	2·8	3·0
$f(x)$	2·7183	4·4817	13·464	16·445	20·086

۲ـ درونیابی لاگرانژی را برای پیدا کردن مقادیر تابع در نقاط $x=2\cdot0$ ، $x=4\cdot5$ به‌کار برید ،جدول مقادیر زیر مفروض است .

x	1·6	2·9	3·7	4·8
$f(x)$	0·6250	0·3448	0·2703	0·2083

۳ـ چند جمله‌ایهای چبیشف $T_0(x)=1$ و $T_1(x)=x$ و رابطهٔ بازگشتی :

$T_{r+1}(x)+T_{r-1}(x)=2xT_r(x),\qquad r=1,2,\ldots,$ داده‌شده‌اند $T_2(x)$ ، $T_3(x)$ ، $T_4(x)$ را بیابید . این نتایج را جهت پیدا کردن عبارات‌مربوط به قوای x تا x^4 برحسب $T_r(x)\ (r=0,1,\ldots,4)$ به‌کار برید .

۴ـ با استفاده از تقریب به $\log(1+x)$ دربازهٔ $[0,1]$ که با رابطهٔ

$$x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}$$

داده شده است بوسیلهٔ روش اقتصادی چبیشف تقریبی از درجهٔ پایین که با این حداکثر $0\cdot12$ اختلاف داشته باشد پیدا نمائید .

۵ـ چند جمله‌ایهای لاگور در جدول ۷۰۱ داده شده‌اند ، چند جمله‌ایهای لاگور از درجهٔ ۳ و ۴ و ۵ را پیدا کنید . خاصیت تعامد را برای چند جمله‌ایهای درجه ۳ و ۴ تحقیق نمائید .

۶ـ روابط زیر را برای انتگرالهای چند جمله‌ایهای چبیشف ثابت کنید :

$$\int T_0(x)\,dx=T_1(x),\qquad \int T_1(x)\,dx=\tfrac{1}{4}T_2(x)$$

$$\int T_r(x)\,dx=\frac{1}{2}\left[\frac{T_{r+1}(x)}{r+1}-\frac{T_{r-1}(x)}{r-1}\right]$$

۷ ـ در مثال ۸ با به کار بردن یک چندجمله‌ای درونیاب درجهٔ دوم تقریبی از $K(3.0)$ را بدست آورید .

۸ ـ بعضی از مقادیر K ، از مثال ۸ ، که از یک جدول بدست آمده‌اند عبارتند از :

$K(2)=1.5713$
$K(3)=1.5719$
$K(5)=1.5738$
$K(6)=1.5751$

با به کار بردن چند جمله‌ای درونیاب درجهٔ دوم تقریبهایی از $k(4.0)$ بدست آورید. نتایج نهائی را تا چهار رقم اعشار گرد کنید. آیا در این حالت استفاده از چندجمله‌ای درجهٔ سوم مزیتی دارد؟

۹ ــ ثابت کنید اگر $f(x)$ یک چندجمله‌ای از درجهٔ نابیشتر از n باشد آنگاه چندجمله‌ای از درجهٔ نابیشتر از n که $f(x)$ را در نقاط $x_0, \ldots, x_n$ درونیابی می‌کند خود $f(x)$ است.

۱۰ ــ با استفاده از تمرین قبل ثابت کنید

$$\sum_{i=0}^{n} l_i(x) \equiv 1$$

که در آن $l_n(x)$ چندجمله‌ای‌های لاگرانژ از درجهٔ n هستند که در (۷۰۱۰) تعریف شدند.

۱۱ ــ برای مثال ۸، تقریبی با بکار بردن چندجمله‌ای درونیاب درجهٔ دوم به روش نیوتن بدست آورید.

۱۲ ــ تحقیق کنید که به ازای هر سه نقطهٔ متمایز x_0 ، x_1 ، x_2

$$f[x_0,x_1,x_2] = f[x_2,x_0,x_1] = f[x_1,x_2,x_0]$$

۱۳ ــ یک جدول از مقادیر $\cos x$ لازم است، بطوری که به ازای هر x در $[0,\pi]$ درونیابی خطی دقتی تا ۶ رقم اعشار بدست دهد. با فرض اینکه مقادیر جدولی متساوی‌الفاصله باشند، ماکزیمم تعداد عناصر جدول چیست؟

۱۴ ــ تابع تعریف‌شده با:

$$f(x) = \int_0^x \sin s^2\, ds$$

برای مقادیر متساوی‌الفاصلهٔ x، با گام $h = 0.1$، جدول‌بندی شده است. در صورتیکه درونیابی درجهٔ سوم برای محاسبهٔ $f(x)$ به ازای هر $\bar{x}$ متعلق به $[0,\frac{\pi}{2}]$ به کار برده شود ماگزیمم خطا چیست؟

۱۵ ــ وقتی $f(x) = \ln x$ ، $n = 3$، $x_0 = 1$، $x_1 = \frac{4}{3}$، $x_2 = \frac{5}{3}$، $x_3 = 2$، $\bar{x} = \frac{3}{2}$

کران پائینی برای خطای درونیابی، یعنی $|f(\bar{x}) - p_n(\bar{x})|$ چه خواهد بود؟

فصل هشتم
انتگرال گیری عددی

۸.۱ ■ مقدمه

این بخش به روشهای عددی موجود برای محاسبهٔ یک انتگرال معین اختصاص دارد، یعنی،

$$I=\int_a^b f(x)\,dx \tag{8.1}$$

دلایل گوناگونی برای اینکه چرا انجام این محاسبه بوسیلهٔ تقریب عددی بجای آنالیز ریاضی لازم یا مطلوب است، وجود دارد. مثلا"، ممکن است پیدا کردن یک فرمول ریاضی برای انتگرال مشکل یا غیر ممکن باشد یا اگر بتوان مسئله را بطور تحلیلی حل کرد ممکن است تابع موردنظر برای محاسبهٔ موثر خیلی پیچیده باشد. همچنین، ممکن است یک برنامه انتگرال‌گیری برای کتابخانهٔ کامپیوتری لازم باشد که بتواند برای یک تابع کلی، و بدون تحلیلی‌ریاضی در هر حالت خاص، به‌کاررود.

روشهایی که در اینجا بررسی خواهند شد با گرفتن تابع ساده $Q_n(x)$ که در تعدادی نقطهٔ انتخاب شده چون $x_i\ (i=0,1,\ldots,n)$ همان مقادیر $f(x)$ را دارد، و استفاده از انتگرال $Q_n(x)$ بعنوان تقریبی‌از انتگرال $f(x)$ بدست می‌آیند. بنابراین باید توابع $Q_n(x)$ برای انتگرال‌گیری ساده باشند و روشهای موردنظر در این بخش بر تقریب بوسیلهٔ توابع چند جمله‌ای، که دارای این خاصیت هستند، استوارند. نقاط x_i که در آنها قرار می‌دهیم $Q_n(x_i)=f(x_i)$ به گره‌های فرمول انتگرال گیری معروفند، مقادیر $f(x_i)$ را با نماد f_i نشان می‌دهیم.

اگر گره‌های انتخاب شده متساوی الفاصله باشند می‌توان یک‌سری فرمول بدست آورد که به فرمولهای نیوتن ــ کاتس Newton–Cotes معروفند، ساده‌ترین دو فرمول این دسته قاعدهٔ ذوزنقه‌ای و قاعدهٔ سیمپسون هستند و بخاطر سادگی آنها

فرمولهای مبتنی براین دو روش غالبا" به‌کار می‌روند، اما، محدودیت این فرمولها به نقاط متساوی‌الفاصله کمی از دقتی که می‌تواند بدست آید می‌کاهد. اگر گره‌ها بطور خاصی انتخاب شوند که ماکزیمم دقت ممکن را بدهند، می‌توان فرمولهایی که دو برابر فرمولهای نیوتن ــ کاتس دقت دارند بدست آورد. این فرمولها به فرمولهای انتگرال‌گیری گاوس معروفند و گره‌های این فرمولها صفرهای چند جمله‌ایهای متعامد مشخصی هستند. باید توجه داشت که فرمولها برحسب تعریف ریاضی خاصی از دقت، که ذیلا" مورد بحث قرار می‌گیرد، دقیقتر خواهند بود. بسادگی اتفاق می‌افتد که یک فرمول نیوتن ــ کاتس مقدار عددی نزدیکتری، از یک فرمول گاوسی بطور نظری دقیقتر، بدهد.

روش دیگری که ذیلا" بحث می‌شود براستفاده از قاعدهٔ ذوزنقه‌ای با چندین اندازهٔ مختلف از بازه‌ها و به کار بردن روش برونیابــی ریچاردســن Richardson جهت کاهش تدریجی خطا استوار است. این روش، که به انتگرال‌گیری رامبرگ معروف است، برای استفاده دریک برنامه کامپیوتری بسیار مناسب است و می‌توان نشان داد که برای هر تابع پیوسته $f(x)$ همگراست. (ر.ک. ایساکسون‌وکلر صفحهٔ 341)

۸.۲■ فرمولهای نیوتن ــ کاتس

۸.۲.۱■ مقدمه

این فرمولها انتگرال

$$I=\int_a^b f(x)\,dx$$

را بوسیلهٔ انتگرال‌گیری از چند جمله‌ای $Q_n(x)$ از درجهٔ n تقریب می‌کنند که ضرائب آن طوری انتخاب می‌شود که

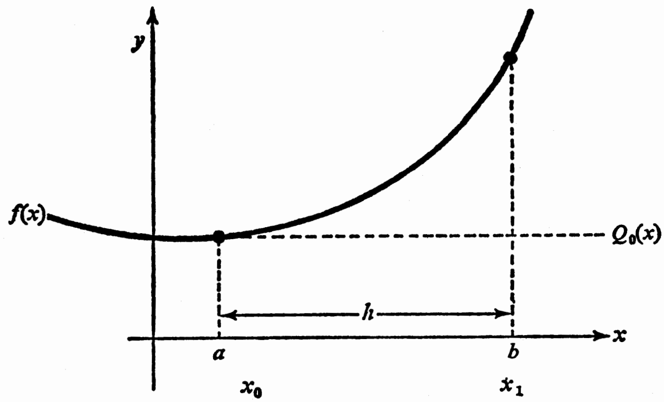
$$Q_n(x_i)=f(x_i),\qquad i=0,1,\ldots,n \tag{8.2}$$

که در آن x_i ها نقاط متساوی‌الفاصله‌اند. به‌منظور روشن بودن مطلب فرمولهای "بسته" را بررسی می‌نماییم که در این حالت

$$x_i=a+i.h,\qquad i=0,1,\ldots,n\qquad h=(b-a)/n \tag{8.3}$$

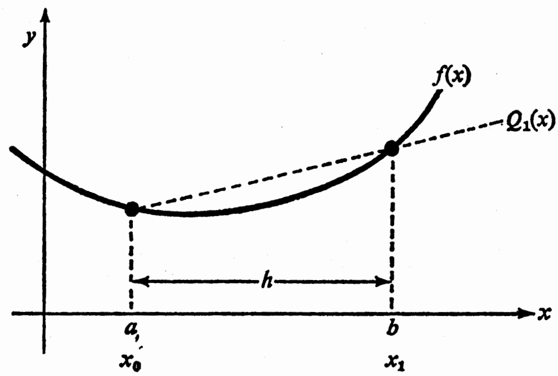
جملهٔ "بسته" نتیجه می‌دهد که هر دو نقطهٔ انتهایی گره‌های فرمول انتگرال‌گیری هستند. چند مثال ساده انگیزهٔ استفاده از این فرمولها را نشان می‌دهد. قاعدهٔ مستطیلی برتقریب بوسیله یک چندجمله‌ای ازدرجهٔ صفراستواراست. یعنی، $Q_0(x)=c_0$ مساحت این مستطیل تقریبی از انتگرال را می‌دهد.

$$\int_a^b f(x)\,dx\approx\int_a^b Q_0(x)\,dx=hf_0 \tag{8.4}$$



شکل ۸.۱ قاعدهٔ مستطیلی

بوضوح این‌تقریب خاصی است و یک راه سادهٔ بهتر کردن آن، همانطور که در شکل ۸.۲ نشان داده شده، استفاده از یک خط مستقیم بین دو نقطهٔ انتهائی منحنی است.

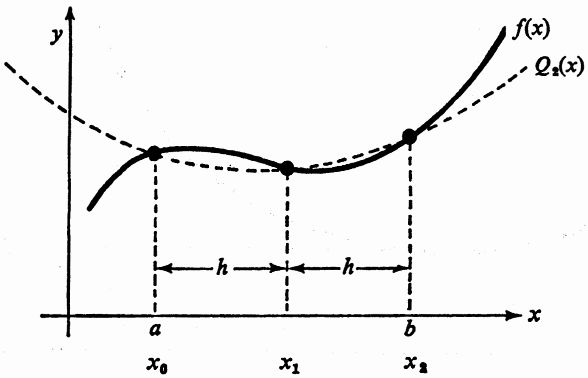


شکل ۸.۲ قاعدهٔ ذوزنقه‌ای

این‌روش‌برتقریب‌بوسیلهٔ یک‌چندجمله‌ای‌درجهٔ ۱استوار‌است. یعنی، $Q_1(x) = c_1x + c_0$ انتگرال از مساحت ذوزنقه حاصل بدست می‌آید.

$$I \approx \frac{h}{2}[f_0 + f_1] \tag{8.5}$$

با وجود اینکه انتظار می‌رود روش بالا نتایج بهتری از روش اولی بدهد بنظر می‌رسد که یک منحنی نتایج باز هم بهتر خواهد داد. در این حالت $Q_2(x) = c_2x^2 + c_1x + c_0$ برای‌تقریب‌تابع به‌کار می‌رود. مساحت‌بوسیلهٔ‌انتگرال‌گیری‌از $Q_2(x)$ وجایگذاری مقادیر f_i برای فرمول $Q_2(x_i)$ بدست می‌آید.



شکل ۸.۳ قاعدهٔ سیمپسون

تفصیلات حسابی در اینجا نشان داده نمی‌شوند اما نتیجه عبارتست‌از:

$$I \approx \frac{h}{3}[f_0 + 4f_1 + f_2] \tag{8.6}$$

یک مثال عددی که قاعدهٔ سیمپسون را به‌کار می‌برد در مثال ۸.۱ داده شده‌است.

پس می‌توان ملاحظه کرد که با گرفتن یک سری چند جمله‌ای با مرتبهٔ صعودی یک خانواده از فرمولهای انتگرال‌گیری تولید خواهد شد که شکل کلی زیر را دارند.

$$\int_a^b f(x)\,dx = \sum_{i=0}^{n} a_i f_i + E \tag{8.7}$$

که در آن مایلیم جملهٔ خطای E حتی‌الامکان کوچک باشد.

۸.۲.۲ ■ تعیین ضرائب

اکنون یک شکل کلی برای تقریب انتگرال‌داریم که دارای $2n+2$ ضریب متغیر است یعنی a_i ها و x_i ها که باید چنان انتخاب شوند که فرمول مورد نظر را حتی‌الامکان دقیق سازند. در حالت فرمولهای نیوتن ــ کاتس درجه‌ای از انتخاب با متساوی‌الفاصله گرفتن x_i ها از دست رفته است، بطوری که تنها $n+1$ پارامتر برای انتخاب وجود دارد. این ضرائب را می‌توان از نتایج تفاضل ــ متناهی بدست آورد اما بررسی این مسئله بعنوان یک فرمول کلی، با ضرائب متعدد که باید انتخاب شوند، آموزنده‌تر است. لذا ضرائب با اعمال برقراری شرایط معینی در فرمول، که منجر به خواص ریاضی یا عددی مطلوب می‌شود، بدست می‌آیند. معیار دقت که در اینجا به‌کار می‌رود صفر قرار دادن جملهٔ خطا بهنگامی که $f(x)$ هر چند جمله‌ای از درجهٔ کوچکتر یا مساوی مقدار مشخص شدهٔ n است، می‌باشد. چون $n+1$ ضریب a_i وجوددارد $n+1$

شرط لازم است تا این ضرائب بدست آیند. این شرایط با دقیق ساختن فرمول برای چند جمله‌ایهای $1, x, x^2, \ldots, x^n$ مهیا می‌شوند. باید توجه داشت که اگر این فرمول برای این چند جمله‌ایها دقیق باشد برای هر ترکیب خطی از اینها نیز دقیق است، یعنی، هر چند جمله‌ای از درجهٔ کمتر یا مساوی با n.

بعنوان یک مثال، جهت نشان دادن چگونگی یافتن ضرائب، فرمول بسته نیوتن–کاتس از مرتبهٔ 3 را بربازهٔ 0 تا $3h$ درنظر می‌گیریم:

$$\int_0^{3h} f(x)\,dx = \sum_{i=0}^{3} a_i f_i + E, \qquad x_i = 0 + i.h \tag{8.8}$$

چون این فرمول برای $1, x, x^2, x^3$ دقیق‌است‌این توابع‌را، به‌نوبت، درمعادلهٔ (8.8) جایگذاری می‌نماییم تا چهار معادله با چهار مجهول بدست آید که بسادگی‌حل‌می‌شوند.

$$f(x)=1 \qquad 3h = a_0 + a_1 + a_2 + a_3 \tag{8.9a}$$

$$f(x)=x \qquad \frac{9h^2}{2} = h.a_1 + 2h.a_2 + 3h.a_3 \tag{8.9b}$$

$$f(x)=x^2 \qquad \frac{27h^3}{3} = h^2.a_1 + 4h^2.a_2 + 9h^2.a_3 \tag{8.9c}$$

$$f(x)=x^3 \qquad \frac{81h^4}{4} = h^3.a_1 + 8h^3.a_2 + 27h^3.a_3 \tag{8.9d}$$

این معادلات جواب زیر را دارند.

$$a_0 = \frac{3h}{8}, \qquad a_1 = \frac{9h}{8}, \qquad a_2 = \frac{9h}{8}, \qquad a_3 = \frac{3h}{8} \tag{8.10}$$

۸.۲.۳ ■ بحث خطاها

بطور مشهودی ممکن است احساس شود که دقت این فرمولهای انتگرال‌گیری را می‌توان با گرفتن فرمولهای مراتب بالاتر، تا رسیدن به دقت کافی، افزایش داد. اما، دو دلیل برای شکست احتمالی این خط مشی وجود دارد، یعنی، اینکه تابع ممکن است بحد کافی با یک چند جمله‌ای تقریب نشود، که در این حالت خطای برشی بزرگ می‌شود، یا اینکه فرمولها ممکن است دارای خطای گرد شدهٔ اضافی باشند. حال این فرمولها و جملهٔ خطای آنها را مورد بررسی قرار می‌دهیم ملاحظه خواهد شد که هر دو مسئله ذکر شده اتفاق می‌افتند. بحث ریاضی مفصلی در رالستون Ralston (1965) داده شده است که در آن نشان می‌دهد شکل کلی برای فرمولهای انتگرال‌گیری نیوتن–کاتس، شامل جملهٔ خطا، عبارتست‌از:

$$\int_a^b f(x)\,dx = A_0.h.\sum_{i=0}^{n} w_i f_i + A_1 h^{k+1} f^{(k)}(\zeta) \tag{8.11}$$

که در آن n تعداد بازه‌ها ، $h=(b-a)/n$ ، و ζ مقداری در بازهٔ $[a,b]$. اســت . جدول زیر تعدادی از ضرائب را برای فرمولهای بسته ارائه می‌دهد . توجه کنیدکه‌ضرائب متقارن هستند و تنها نیاز به نشان دادن نیمی‌از جدول داریم .

جدول ۸.۱ ضرائب فرمولهای بستهٔ نیوتن ــ کاتس

n	A_0	w_0	w_1	w_2	w_3	w_4	A_1
1	$\frac{1}{2}$	1	1				$-\frac{1}{12}$
2	$\frac{1}{3}$	1	4	1			$-\frac{1}{90}$
3	$\frac{3}{8}$	1	3	3	1		$-\frac{3}{80}$
4	$\frac{2}{45}$	7	32	12	32	7	$-\frac{8}{945}$
5	$\frac{5}{288}$	19	75	50	50	75	$-\frac{275}{12096}$
6	$\frac{1}{140}$	41	216	27	272	27	$-\frac{9}{1400}$
7	$\frac{7}{17280}$	751	3577	1323	2989	2989	$-\frac{8183}{518400}$
8	$\frac{4}{14175}$	989	5888	-928	10946	-4540	$-\frac{2368}{467775}$

نتیجهٔ جالبی که از آنالیز بدست می‌آید آنست که فرمولهای با تعداد زوج نوار نه‌تنها برای چند جمله‌ایهای تا درجهٔ n بلکه‌برای‌چند جمله‌ایهای‌درجهٔ $n+1$ نیزخطای صفر می‌دهند . این بدان معنی است که در معادلهٔ (8.11) وقتی n زوج است $k=n+2$ و وقتی n فرد است $k=n+1$. در پرتو این دقت اضافی طبیعتا " یک فرمول مرتبهٔ زوج به‌کار می‌رود . استثناء در این مورد قاعدهٔ ذوزنقه‌ای است که‌بخاطرسادگی‌آن‌باارزش است . نقطه قوت دیگر آن است که‌خطابه‌مشتق‌دوم‌تابع‌مورد انتگرال‌گیری بستگی دارد . توابع سادهٔ بسیاری موجودند که مشتقات مراتب بالای آنها می‌توانند مقادیر بسیار بزرگ داشته باشند و در این حالات استفاده‌از فرمولهای مرتبهٔ بالا ، خطای برشی بزرگی می‌دهد .

علاوه بر مسئله خطای برشی از جدول ضرائب دیده می‌شود که مشکلاتی بخاطر رشد خطای گرد کردن انتظار می‌رود . اگر محاسبهٔ ماکزیممی برای خطای گرد کردن موردنظر باشد آنگاه اندازهٔ کمیت زیر مهم است .

$$B_n = A_0 . h \sum_{i=0}^{n} |w_i| \tag{8.12}$$

حال از معادلهٔ (8.9a) ملاحظه می‌شود که

$$A_0 . h \sum_{i=0}^{n} w_i = b-a$$

بطوری که ، در حالتی که ضرائب همه مثبت‌هستند $B_n=(b-a)$. اما ،برای‌فرمولهایی که در آن $n=8$ و $n \geqslant 10$ بعضی از w_i ها منفی هستند بطوری که B_n بزرگتـــر از

$b-a$ است و با n افزایش می‌یابد. کمیت V_n ، که برای تخمین آماری اندازهٔ خطا به‌کار می‌رود، نیز مهم است. یعنی، برای فرمولهای با مرتبهٔ مختلف با فاصلهٔ h یکسان

$$V_n = A_0^2 \sum_{i=0}^{n} w_i^2 \tag{8.13}$$

بخاطر تغییر علامت در ضرائب و تغییرات بزرگ در اندازهٔ آنها، که در فرمولهای مرتبهٔ بالا گسترش می‌یابد، این کمیت نیز با n افزایش می‌یابد.

۸.۲.۴ ■ فرمولهای مرکب

بخاطر مسائلی که با فرمولهای مرتبهٔ بالا پیش می‌آید باید روشهایی پیدا کرد که خطا را به سطح قابل قبولی تقلیل دهد. یک روش واضح تقسیم بازهٔ $[a,b]$ به‌زیر بازه‌های متعدد است که در هر یک از آنها یک فرمول مرتبهٔ پایین به‌کار رفته باشد. این عمل اندازهٔ h را کاهش می‌دهد و با وجود اینکه جملات خطای متعددی باید جمع شوند اثر نهائی، همچنانکه در زیر نشان داده شده، تقلیل خطای برشی است. چون اکنون جملهٔ خطا به مشتق مرتبهٔ پایینی بستگی دارد مسائل مربوط به مشتقات مراتب بالا نیز رفع شده است. ضرائب فرمولهای مرکب از مرتبهٔ پایین همگی مثبت هستند و بطورفاحش تغییرنمی‌کنند بطوری که کمیتهای B_n و V_n مناسب هستند. بحثی از خطای گردکردن در قاعدهٔ ذوزنقه‌ای مرکب، با کمی تفصیل در McCracken and Dorn (1964) ارائه شده است. فرمولهایی که اکثر اوقات به‌کار می‌روند قاعدهٔ ذوزنقه‌ای مرکب با $h=(b-a)/m$ است.

$$\begin{aligned} I_m &= \frac{h}{2}[f_0+f_1+f_1+f_2+\cdots+f_m] \\ &= \frac{h}{2}[f_0+2f_1+2f_2+\cdots+f_m] \end{aligned} \tag{8.14}$$

و قاعدهٔ سیمپسون مرکب با $h=(b-a)/m$ است که در آن m زوج است.

$$\begin{aligned} I_m &= \frac{h}{3}[f_0+4f_1+f_2+f_2+4f_3+f_4+\cdots+f_m] \\ &= \frac{h}{3}[f_0+4f_1+2f_2+4f_3+\cdots+f_m] \end{aligned} \tag{8.15}$$

اگر قرار دهیم.

$$M_2 = \max_{a\leqslant x\leqslant b} f''(x) \quad \text{and} \quad M_4 = \max_{a\leqslant x\leqslant b} f^{\mathrm{IV}}(x)$$

آنگاه خطای قاعدهٔ ذوزنقه‌ای مرکب با رابطهٔ

$$|E| \leqslant \frac{mh^3}{12}M_2 = \frac{(b-a)^3}{12m^2}M_2 \tag{8.16}$$

داده می‌شود و خطای قاعدهٔ سیمپسون مرکب با

$$|E| \leqslant \frac{mh^5}{180}M_4 = \frac{(b-a)^5}{180m^4}M_4 \tag{8.17}$$

لذا ، برای همگرائی این فرمولهای مرکب ، وقتی m زیاد می‌شود ، به‌ترتیب کراندار بودن مشتق دوم و چهارم ، که غالبا" قابل اثبات است ، لازم می‌باشد . مقایسه‌ای از قاعدهٔ ذوزنقه‌ای مرکب با فرمولهای نیوتن ــ کاتس مرتبهٔ بالا در مثال ۸.۵ داده شده است .

۸.۲.۵ فرمولهای بازنیوتن - کاتس

چون در حالت کلی فرمولهای نیوتن‌ــ کاتس بسته ، و بخصوص به فرم مرکب ، ازنظر محاسباتی کارآیی بیشتری دارند به آنها بیشتر توجه می‌شود . اما بعضی اوقات مناسب است که فرمولهای باز به‌کار برد ، که در آنها از نقاط انتهایی a و b استفاده نمی‌شود طوری که $x_i = a + i.h\ (i = 1, \ldots, n-1)$ و $h = (b-a)/n$. این فرمولها درانتگرال‌گیری ازمعادلات دیفرانسیل بعنوان فرمولهای‌اصلاحگر کاربرد پیدا می‌کنند . جدول زیر ضرائب را برای چند فرمول مرتبهٔ پایین بدست می‌دهد . جملهٔ خطا دارای شکل

$$A_1 h^{k+1} f^{(k)}(\zeta) \tag{8.18}$$

است که در آن اگر n زوج باشد $k = n$ و اگر n فرد باشد $k = n-1$.

جدول ۸.۲ ضرایب فرمولهای باز نیوتن ــ کاتس

n	A_0	w_1	w_2	w_3	w_4	A_1
2	2	1				$\frac{1}{3}$
3	$\frac{3}{2}$	1	1			$\frac{3}{4}$
4	$\frac{4}{3}$	2	-1	2		$\frac{14}{45}$
5	$\frac{5}{24}$	11	1	1	11	$\frac{95}{144}$

۸.۳ انتگرال‌گیری گاوس

۸.۳.۱ مقدمه

اگر فرمول

$$\int_a^b f(x)\,dx = \sum_{i=0}^{n} a_i f(x_i) + E \tag{8.19}$$

را درنظر بگیریم ملاحظه می‌کنیم که $2n+2$ پارامتر ، a_i ها و x_i ها ، وجود دارد که می‌توانند تغییر کنند . چون‌یک چند جمله‌ای از درجهٔ $2n+1$ به $2n+2$ شرط برای تعیین ضرائب آن نیاز دارد بنظر قابل قبول می‌رسد که $2n+2$ پارامتر چنان

انتخاب شوند که تمام چند جمله‌ایهای از درجهٔ کمتر یا مساوی $2n+1$ دارای انتگرالی با خطای صفر باشند. ذیلاً نشان می‌دهیم که در حقیقت چنین است و روش پیداکردن مقادیر a_i و x_i را بررسی می‌نماییم.

اگر روش قبلاً استفاده شده برای پیدا کردن ضرائب a_i را درنظر بگیریم دیده می‌شود که بسادگی قابل گسترش برای پیدا کردن گره‌های x_i نیست. حالتی را که در آن $n=1$ درنظر بگیرید که بااستفاده از روش بخش ۸.۲.۲ به معادلات زیر منجر می‌شود. اگر حدود انتگرال‌گیری $a=0$, $b=h$ باشند:

$$
\begin{aligned}
f(x)&=1 & h&=a_0+a_1\\
f(x)&=x & \frac{h^2}{2}&=a_0x_0+a_1x_1\\
f(x)&=x^2 & \frac{h^3}{3}&=a_0x_0^2+a_1x_1^2\\
f(x)&=x^3 & \frac{h^4}{4}&=a_0x_0^3+a_1x_1^3
\end{aligned}
\tag{8.20}
$$

با وجود اینکه این معادلات نسبت به a_i ها خطی هستند نسبت به x_i ها خطی نبوده و لذا بسادگی حل نمی‌شوند. بوضوح برای n های بزرگتر وضعیت خیلی بدتر می‌شود. خوشبختانه نظریه ریاضی مربوط به چند جمله‌ایها وسیله‌ای برای پیدا کردن گره‌ها مهیا می‌کند و ضرائب a_i متعاقباً به روش قبلی بدست می‌آیند. بحث مختصری از پیشنیاز انتگرال‌گیری گاوس در اینجا داده می‌شود، تفصیلات بیشتر در Isaacson and Keller (1966) و Ralston (1965) موجود است.

۸.۳.۲ ■ انتگرال‌گیری گاوسی وزندار

بعضی از خواص چند جمله‌ایهای متعامد در فصل ۷ بحث شدند و در بحث زیر خاصیت اساسی چند جمله‌ایهای متعامد، که مهم می‌باشد، بررسی می‌شود، یعنی، اگر مجموعه‌ای از چند جمله‌ایهای $(k=0,1,\ldots,n+1)$ $Q_k(x)$ موجود باشد که بر $[a,b]$ متعامدند آنگاه

$$
\int_a^b Q_{n+1}(x)S_k(x)\,dx=0, \qquad k=0,1,2,\ldots,n \tag{8.21}
$$

که در آن $S_k(x)$ یک چند جمله‌ای از درجهٔ k است. این خاصیت برای $S_k(x)$ دلخواه برقرار است زیرا هر چند جمله‌ای قابل بیان بصورت یک ترکیب خطی از چند جمله‌ایهای متعامد $Q_k(x)$ است.

اینک چندجمله‌ای‌دلخواه $P_{2n+1}(x)$ از درجهٔ $2n+1$ راانتخاب وبر $Q_{n+1}(x)$ تقسیم می‌کنیم تا بدست آوریم

$$P_{2n+1}(x) = Q_{n+1}(x) l_n(x) + R(x) \tag{8.22}$$

که در آن $R(x)$ حداکثر از درجهٔ n است. حال انتگرال بدو قسمت تقسیم می‌شود.

$$\int_a^b P_{2n+1}(x)\,dx = \int_a^b Q_{n+1}(x) l_n(x)\,dx + \int_a^b R(x)\,dx \tag{8.23}$$

و بنابر خاصیت تعامد انتگرال اول صفر است. پس،

$$\int_a^b P_{2n+1}(x)\,dx = \int_a^b R(x)\,dx \approx \sum_{i=0}^{n} a_i R(x_i)$$

حال اگر نقاط $x_i\ (i=0,1,\ldots,n)$ را صفرهای $Q_{n+1}(x)$ انتخاب کنیم آنگاه نتیجه می‌شود که

$$P_{2n+1}(x_i) = R(x_i), \qquad i=0,1,\ldots,n$$

زیرا اولین جمله در معادلهٔ 8.22 صفر می‌شود. بنابراین داریم

$$\int_a^b P_{2n+1}(x)\,dx = \sum_{i=0}^{n} a_i P_{2n+1}(x_i) \tag{8.24}$$

که در آن گره‌های x_i بعنوان صفرهای $Q_{n+1}(x)$ معلومند و ضرائب a_i را می‌توان مانند بخش ۸.۲.۲ پیدا کرد.

۸.۳.۳ ■ انتگرالهای ناسره ونامتناهی

اگر فرمولهای بالا را بربازه $[-1,+1]$ به‌کار بریم چند جمله‌ایهای مناسب، چند جمله‌ایهای لژاندر هستند امااستفاده از دیگر چند جمله‌ایهای متعامد با تغییرات مختصر در شرایط مسئله امکان‌پذیر است. اولاً"، توجه کنید که یک انتگرال نسبت به x بین حدود متناهی a و b توسط تبدیل زیر:

$$x = \frac{(b-a)X + a + b}{2} \tag{8.25}$$

قابل تبدیل به یک انتگرال با حدود $[-1,+1]$ می‌باشد بقسمی که چند جمله ایهای لژاندر برای هر حوزه متناهی‌با تغییر متغیر مربوطه مناسب است. نتایج عددی برای یک مثال ساده با به‌کار بردن چند جمله‌ایهای لژاندر در مثال ۸.۲ داده شده است.

اگر یک تابع وزن $w(x) > 0$ در انتگرال وارد کنیم، و چند جمله‌ایهایی که بر حسب این تابع وزن متعامد هستند به‌کار بریم، از چند جمله‌ایهای متعامد مختلف متعددی می‌توان استفاده کرد.

$$\int_{-1}^{+1} f(x)\,dx = \int_{-1}^{+1} w(x).\frac{f(x)}{w(x)}\,dx$$

بنابراین ، اگر $g(x)=f(x)/w(x)$ داریم

$$\int_{-1}^{+1} w(x)g(x)\,dx=\sum_{i=0}^{n} a_i g(x_i) \tag{8.26}$$

بعنوان یک‌مثال چند جمله‌ایهای متعامدی‌را درنظرمیگیریم که‌با $w(x)=(1-x^2)^{-1/2}$ پدید می‌آیند . این چند جمله‌ایها چند جمله‌ایهای چبیشف هستند که در فصل ۷ بررسی شدند و دو خاصیت دارند که در اینجا توجه خاصی به آنها داریم . اولاً ، ضرائب a_i از معادلهٔ (8.26) همه یک مقدار $\pi/(n+1)$ را دارند که کمی محاسبات لازم را کاهش می‌دهد و خطای گرد کردن را کم می‌کند . همچنین ، گره‌های فرمولهای انتگرال‌گیری چبیشف با فرمول سادهٔ زیر داده می‌شوند .

$$x_j=\cos\left[\frac{(2j+1)\pi}{2(n+1)}\right], \qquad j=0,1,\ldots,n \tag{8.27}$$

ملاحظه می‌شود که اگر مقداری از انتگرال به‌ازای یک مقدار از n حساب شده باشد و لازم باشد که مقدار n را جهت دقت بیشتر افزایش دهیم منطبق گرفتن بعضی ازمقادیر گره‌ها در دو فرمول امکان‌پذیر است و این محاسبات اضافی را کاهش می‌دهد . مثلاً ، اگر از گره‌های $\cos(\pi/6)$, $\cos(3\pi/6)$, $\cos(5\pi/6)$ استفاده شود درجهٔ فرمول رامی‌توان در ۳ ضرب کرد که نقاط $\cos(\theta_i)$ را بدهد که در آن θ_i با $\pi/18$ $3\pi/18$, $5\pi/18$ داده می‌شوند . یک سوم این مقادیر با مجموعهٔ قبلی یکسان است و $f(x_i)$ های مربوطه لازم به محاسبه نیست . این خاصیت را چند جمله‌ایهای متعامد دیگر دارا نیستند .

۸.۳.٤ ■ انتگرالهای ناسره و حدود نامتناهی

دلایل دیگری برای اینکه کاربرد تابع وزن بتواند مزیت داشته باشد وجود دارد . مسائلی که تاکنون بررسی شده‌اند در مورد انتگرالهای‌روی یک بازهٔ متناهی بوده‌اند که در آن مقادیر $f(x_i)$ را هر جا مورد نیاز بوده‌اند می‌توانسته‌ایم تعیین نماییم . در حالتی که $f(x_i)$ نامعین می‌شود ، حتی اگر انتگرال موجود باشد ، فرمولهای‌بالا را نمی‌توان به‌کار برد . همچنین مشکلاتی به هنگام نامعین شدن مشتق $f(x)$ موجود است زیرا این می‌تواند باعث جملهٔ خطای بی‌کران شود .

مثال ساده‌ای نشان می‌دهد که چگونه استفاده از تابع وزنی می‌تواند این مشکل را حل‌کند .انتگرال‌گیری تابع $(1-x^2)^{-1/4}$ را رویبازهٔ $[-1,+1]$ درنظربگیرید . چون این تابع در دو انتهای این حوزه نامتناهی است استفاده از یک فرمول انتگرال‌گیری بسته ممکن نخواهد بود . اما ، با به‌کار بردن انتگرال‌گیری گاوس ـ چبیشف با تابع وزنی $(1-x^2)^{-1/2}$ داریم ، و با به کار بردن معادلهٔ (8.26) ،

$$\int_{-1}^{+1}\frac{(1-x^2)^{1/4}}{(1-x^2)^{1/2}}dx=\frac{\pi}{n+1}\sum_{i=0}^{n}(1-x_i^2)^{1/4} \tag{8.28}$$

و محاسبهٔ تابع سمت راست که ساده است نتیجه می‌گردد.

استفاده از تابع وزن، انتگرال‌گیری از توابع روی یک حوزهٔ نیم نامتناهی یا نامتناهی را بوسیلهٔ انتخاب چند جمله‌ایهای متعامد مناسب نیز نتیجه می‌دهد. روی بازهٔ $[0, \infty]$ چند جمله‌ایهای لاکور Laguerre را می‌توان به‌کار برد، زیرا آنها نسبت به تابع وزن e^{-x} روی این حوزهٔ متعامدند. (ر.ک. مثال ۸۰۳) روی‌حوزهٔ $[-\infty, +\infty]$ چند جمله‌ایهای هرمیت در رالستون Ralston (1965) جدول‌بندی شده‌اند و در آن بحث بیشتری روی فرمولهای با تابع وزنی ارائه شده است.

۸.۴ ■ انتگرال گیری رامبرگ

قبلاً مشاهده شده است که قاعدهٔ ذوزنقه‌ای مرکب برای استفاده ساده است و اکنون تغییری‌در این قاعدهٔ ساده را توضیح می‌دهیم که می‌تواند به دقت زیاد منجر شود و بسیار مناسب استفاده در کامپیوتر باشد. برای دستهٔ وسیعی از توابع فرمول انتگرال‌گیری ذوزنقه‌ای و جملهٔ خطای آن را می‌توان به‌صورت زیر نوشت:

$$\frac{h}{2}(f_0+2f_1+\cdots+f_m)=I+\sum_{j=1}^{\infty}a_jh^{2J} \tag{8.29}$$

بطوری که امکان استفاده از برونیابی ریچاردسن Richardson، جهت بالا بردن دقت نتیجه، امکان دارد (ر.ک. فصل ۱). فرض کنید، بخاطر سادگی علامتگذاری، که مقدار اولیه m قوه‌ای از 2 مثلاً 2^k، باشد و تقریب داده شده بوسیلهٔ معادلهٔ (8.29) با $T_{0,k}$ نشان داده شود، یعنی،

$$T_{0,k}=I+\sum_{j=1}^{\infty}a_jh^{2J} \tag{8.30}$$

اکنون تقریب به انتگرال موردنظر را با بازهٔ نصف شده حساب می‌کنیم تا بدست آوریم

$$T_{0,k+1}=I+\sum_{j=1}^{\infty}a_j\left(\frac{h}{2}\right)^{2J} \tag{8.31}$$

حال می‌توان جملهٔ اول سریهای خطا را، با درنظر گرفتن ترکیب مناسبی از این دو معادله، حذف کرد.

$$4I-I=4T_{0,k+1}-T_{0,k}-\sum_{i=2}^{\infty}a_j\left[\frac{4\,.\,h^{2J}}{2^{2J}}-h^{2J}\right]$$

یا

$$I=\frac{4T_{0,k+1}-T_{0,k}}{3}-\sum_{j=2}^{\infty}\frac{a_j h^{2J}}{3}\left[\frac{4}{2^{2J}}-1\right] \tag{8.32}$$

اولین جملهٔ دست راست این معادله با $T_{1,k}$ مشخص می‌شود و می‌توان دید کـه اکنون جملهٔ خطا با h^4 شروع می‌شود. نصف کردن متوالی بازه دنباله‌ای از مقادیر $T_{0,k}$ می‌دهد و هرجفت متوالی می‌تواند ترکیب شود تا مقادیر $T_{1,k}$ بدست آید. سپس دنبالهٔ مقادیر $T_{1,k}$ را می‌توان به روش مشابهی ترکیب کرد تا جملهٔ خطای مربوط به h^4، بوسیله برونیابی ریچاردسن ـ حذف شود.

با استفاده از فرمول

$$T_{p,k}=\frac{1}{4^p-1}(4^pT_{p-1,k+1}-T_{p-1,k}),\qquad \begin{matrix}k=0,1,2,\dots,\\ p=1,2,\dots,\end{matrix} \tag{8.33}$$

این فرایند می‌تواند ادامه پیدا کند تا دنباله‌ای از ستونها با جملات خطایی با مرتبهٔ صعودی تشکیل شود.

$$\begin{array}{llll} h^2 & h^4 & h^6 & h^8\\ T_{0,0} & & & \\ T_{0,1} & T_{1,0} & & \\ T_{0,2} & T_{1,1} & T_{2,0} & \\ T_{0,3} & T_{1,2} & T_{2,1} & T_{3,0} \end{array} \tag{8.34}$$

با استفاده از این حقیقت که $h=(b-a)/2^k$ ملاحظه می‌شود که جملهٔ خطا برای تقریب $T_{p,k}$ از مرتبهٔ $[(b-a)/2^k]^{2p+2}$ است، و مقادیر هر ستون سریعتـــر بـه مقدار واقعی انتگرال همگرا هستند. این‌روش مخصوصا " برای استفاده مناسب است زیرا مقادیر متوالی را می‌توان جهت همگرائی فرایند مقایسه کرد. جدول مقادیر زیر باید این مطلب را روشن سازد. نتایج عددی بیشتر در مثال ۸.۴ داده شده‌است.

جدول ۸.۳ جدول انتگرال‌گیری رامبرگ برای $\sec x$

Number of intervals	$T_{0,k}$	$T_{1,k}$	$T_{2,k}$	$T_{3,k}$
1	0·948 059			
2	0·899 084	0·882 759		
4	0·885 886	0·881 487	0·881 402	
8	0·882 507	0·881 381	0·881 374	0·881 372

۸.۵ ■ مقایسه روشها

درحالتی از توابع که به صورت جدول‌بندی شده دردسترس می‌باشند باید فرمولی انتخاب گردد که مبتنی بر گرههای متساوی‌الفاصله باشد زیرا انتگرال‌گیری گاوسی شامل

درونیابی ، جهت پیدا کردن مقادیر تابع ، خواهد بود . همچنین مناسب است جائیکه داده‌ها از مشاهدات تجربی بدست می‌آیند از بازه‌های متساوی استفاده شود . انتخاب بین یک فرمول نیوتن ــ کاتس مرتبهٔ بالا و یک فرمول مرکب مرتبهٔ پایین بستگی به مقدار اطلاعاتی دارد که از مشتقات تابع داریم . در حالتی که مشتقات مرتبهٔ بالا سریعا" صفر می‌شوند یک فرمول مرتبهٔ بالا کارآتر خواهد بود ، با این فرض که دقت کامپیوتر چنان است که خطاهای گرد کردن غالب نمی‌شوند . برای مسائلی که مشتقات مرتبهٔ بالا می‌توانند کاملا" بزرگ باشند ، قواعد مرکب مرتبهٔ پایین ترجیح داده می‌شوند . قاعدهٔ سیمپسون از قاعدهٔ ذوزنقه‌ای دقیقتر است مگر آنکه‌مشتق چهارم به مراتب بزرگتر از مشتق دوم باشد . مشکلترین مسئله ، انتخاب بازهایی‌است که دقت کافی را تضمین کند . در وضعیتی که انتگرال‌گیریهای مشابه متعددی باید انجام گیرد ارزشمند است که چندین نتیجه با بازه‌های مختلف حساب شود تا طول بازه‌ای بدست آید که دقت کافی را می‌دهد . اگر تنها یک انتگرال‌گیری باید انجام شود محاسبات اضافی لازم برای انتگرال‌گیری رامبرگ قابل قبول خواهد بود ، زیرا دنبالهٔ مقادیر نشان می‌دهند که چه موقع دقت کافی بدست آمده است و می‌توان نشان داد که انتگرال‌گیری رامبرگ برای هر تابع پیوسته‌ای همگرا است .

اگر مقادیر تابع در هر مقدار از گره موجود است ، مثل حالتی که مقادیر تابع بوسیلهٔ کامپیوتر حساب می‌شوند ، آنگاه دقت اضافی موجود در انتگرال‌گیری گاوسی مفید خواهد بود . این فرمولها بخصوص وقتی محاسبهٔ مکرر انتگرالها لازم است‌با‌ارزش هستند زیرا در این صورت کارآیی محاسباتی مهم است . در چنین حالتی باید محاسبات مقدماتی جهت انتخاب مرتبهٔ فرمول انجام پذیرد اما،زمان صرف شده برای این کار قابل قبول است . برای محاسبات بادقت ساده اطمینان به‌دقت بدست آمده مشکل است‌مگر آنکه ، به‌منظور آزمایش ، بیشتر از یک مقدار برای انتگرال بدست آید . اگر لازم باشد یک فرمول مرتبه بالاتر به‌کار رود تمام گره‌ها مقادیر جدید خواهند داشت و تمام مقادیر تابع احتیاج به‌محاسبهٔ مجدد دارند ، که در حالت فرمولهای با بازهٔ متساوی اینطور نبود . لذا ، فرمولهای انتگرال‌گیری گاوسی خیلی مناسب محاسبهٔ یک انتگرال نیستند .

مثالهای حل شده وبرنامه های کامپیوتری

۱ــ در این مثال قاعدهٔ سیمپسون جهت پیدا کردن $\int_0^{\pi/4} \sec(x)\,dx$ به‌کار خواهد رفت . مقدار درست انتگرال تا شش رقم اعشار $0.881\,374$ است .

این انتگرال با دو، چهار و هشت بازه محاسبه می‌شود. مقادیر تابع $\sec x$ درنقاط مختلف عبارتنداز:

x	0°	5·625°	11·25°	16·875°	22·5°
$\sec x$	1·000000	1·004838	1·019591	1·044997	1·082392

x	28·125°	33·75°	39·375°	45°
$\sec x$	1·133888	2·202690	1·293643	1·414214

مقادیر بدست‌آمده با استفاده از قاعده سیمپسون مرکب عبارتنداز:

تعداد بازه‌ها	انتگرال	خطا	مقدار ریچاردسن
2	0·882761	0·001387	
4	0·881489	0·000115	0·881404
8	0·881383	0·000009	0·881376

جمله خطا برای قاعدهٔ سیمپسون عبارتستاز:

$$\frac{-h^4(b-a)}{180}\cdot f^{\mathrm{IV}}(\zeta_1)$$

و این را می‌توان در کوششی جهت بهتر کردن مقادیر سیمپسون، بوسیلهٔ روش برونیابی ریچارد سن، به‌کار برد. وقتی اندازهٔ گام نصف می‌شود خطا برابر می‌شود با

$$\frac{-h^4(b-a)}{16.180}\cdot f^{\mathrm{IV}}(\zeta_2)$$

تقریب جدیدی بوسیله $(16I_{2n}-I_n)/15$ بدست می‌آید و دو مقدار بدست آمده بوسیله ترکیب دو جفت از نتایج نشان داده شده‌اند.

طریقه ایده‌آل انتخاب اندازه گام، محاسبهٔ مقدار ماکزیمم مشتق مرتبهٔ چهار در فرمول خطا و به‌کار بردن آن جهت محاسبهٔ مقدار h مناسب است که حاصل‌جمع جملات خطا را کمتر از مقدار انتخاب شده‌ای کند. در اکثر مثالها، مانند این مثال، مشتق چهارم فرمول پیچیده‌ای دارد و پیدا کردن کران کار ساده‌ای نیست.

۲ــ فرمول انتگرال‌گیری گاوس ــ لژاندر با سه گره را بدست می‌آوریم و نشان خواهیم داد که این فرمول برای یک چند جمله‌ای از درجهٔ ۵ دقیق است. چند جمله‌ای لژاندر درجه ۳ عبارتستاز $\frac{1}{2}(5x^3-3x)$ که دارای ریشه‌های

$$0, \pm\sqrt{(0\cdot6)} = 0, \pm0\cdot774597$$

است. اگر چند جمله‌ایهای $1, x, \ldots,$ را جایگذاری کنیم معادلاتی برای ضریب فرمول گاوسی بدست می‌آوریم

$$\int_{-1}^{+1} f(x)\,dx = \sum_{i=0}^{2} a_i f(x_i)$$

چون

$$x_0 = -0{\cdot}774\,597,\ x_1 = 0,\ \ x_2 = 0{\cdot}775\,497$$

$$\begin{array}{lll}
f(x)=1 & 2 = a_0 + a_1 + a_2 & \\
f(x)=x & 0 = a_0x_0 + a_2x_2 & a_0 = a_2 \\
f(x)=x^2 & \frac{2}{3} = a_0x_0^2 + a_2x_2^2 & a_0 + a_2 = \frac{10}{9} \\
f(x)=x^3 & 0 = a_0x_0^3 + a_2x_2^3 & a_0 = a_2 = \frac{5}{9} \\
f(x)=x^4 & \frac{2}{5} = a_0x_0^4 + a_2x_2^4 & \\
f(x)=x^5 & 0 = a_0x_0^5 + a_2x_2^5 & \\
a_0 = a_2 = \frac{5}{9}, & a_1 = \frac{8}{9} &
\end{array}$$

ملاحظه می‌کنیم‌که

$$\int_{-1}^{+1} x^5\,dx = 0 = -\tfrac{5}{9}(0{\cdot}774\,597)^5 + \tfrac{5}{9}(0{\cdot}774\,597)^5$$
$$= 0$$

۳ــ فرمولهای گاوس ــ لاگور برای انتگرالهایی به‌شکل

$$\int_0^\infty e^{-x} f(x)\,dx = \sum_{i=0}^{n} a_i f(x_i)$$

به‌کار می‌روند . ضرائب و گره‌ها برای فرمول مرتبهٔ ۴ ذیلا" داده شده است .

گره‌ها	ضرائب	حاصلضرب
0·322 548	0·603 154	0·194 546
1·745 761	0·357 419	0·623 968
4·536 620	0·038 888	0·176 420
9·395 071	0·000 539	0·005 063
	1·000 000	0·999 997

اگر $f(x)=1$ مقدار انتگرال 1 است و مقدار تقریب عبارتست‌از :

$$0{\cdot}603\,154 + 0{\cdot}357\,419 + 0{\cdot}038\,888 + 0{\cdot}000\,539 = 1{\cdot}000\,000$$

همانگونه که تئوری پیش‌بینی می‌کند دقیق است . اگر $f(x)=x$ مقدار انتگرال عبارتست‌از :

$$\left[-e^{-x}x\right]_0^\infty + \left[\frac{e^{-x}}{-1}\right]_0^\infty = 1{\cdot}0$$

مقدار تقریب 0·999 998 است که تا حدود اعمال شده درکاربرد شش رقم اعشاردقیق است .

انتگرالی را درنظر بگیرید که شامل چند جمله‌ای $f(x)$ نیست .

$$\int_0^\infty e^{-x}\sin x\,dx = \sum_{i=0}^{3} a_i \sin(x_i)$$

مقدار واقعی انتگرال 0·5 است . تقریب مقدار زیر را می‌دهد .

$$0{\cdot}603\,154 \times 0{\cdot}316\,984 + 0{\cdot}357\,419 \times 0{\cdot}984\,732 + 0{\cdot}038\,888 \times (-0{\cdot}984\,593)$$
$$+ 0{\cdot}000\,539 \times 0{\cdot}029\,703 = 0{\cdot}504\,879$$

پس، برای فرمولی از مرتبه کمی چون 4 خطای تقریبا " 1% داریم و این مطلب در پرتو این حقیقت که $\sin x$ بطور ضعیفی بوسیلهٔ یک چند جمله‌ای روی حوزهٔ مورد سوال تقریب می‌شود، فوق‌العاده است.

۴ـ جدول انتگرال‌گیری رامبرگ زیر همگرائی این روش را نشان می‌دهد. این روش بوضوح برای این مسئله مناسب است زیرا مقدار واقعی عبارتست‌از:

$$1{\cdot}000000 = \int_0^{\pi/2} \sin x\, dx$$

تعداد بازه‌ها	$T_{0,k}$	$T_{1,k}$	$T_{2,k}$	$T_{3,k}$
1	0·785 398			
2	0·948 059	1·002 279		
4	0·987 116	1·000 135	0·999 992	
8	0·996 785	1·000 008	1·000 000	1·000 000
16	0·999 197	1·000 001	1·000 000	1·000 000

استفاده از فرایند Δ^2 ایتکن نیز نشان داده می‌شود. اگر سه مقدار آخری جدول به‌کار روند داریم:

$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$
0·987 116		
0·991 762	0·004 646	
0·994 282	0·002 520	$-0{\cdot}002\,126$

پس مقدار بهتر چنین بدست می‌آید:

$$f = f_n - \frac{(\Delta f_n)^2}{\Delta^2 f_n} = 0{\cdot}987\,116 - \frac{(0{\cdot}004\,646)^2}{-0{\cdot}002\,126}$$
$$= 0{\cdot}997\,271$$

۵ ـ به منظور مقایسه، مجددا می‌خواهیم تقریبی از $I = \int_0^1 e^{-x^2}\,dx$ را، با به کار بردن فرمول پنج‌نقطه‌ای گاوس $(k = 4)$، حساب کنیم.

ابتدا تبدیل $x = (t+1)/2$ را انجام می‌دهیم تا بدست آوریم:

$$I = \int_0^1 e^{-x^2}\,dx = \tfrac{1}{2}\int_{-1}^{1} e^{-(t+1)^2/4}\,dt = \tfrac{1}{2}\int_{-1}^{1} g(t)\,dt \approx \tfrac{1}{2}\sum_{k=0}^{4} A_k g(\xi_k)$$

با به کار بردن نقاط و وزن‌ها از جدول زیر:

k	ξ_i	A_i
1	$\xi_1 = -\xi_0 = 0.57735027$	$A_1 = A_0 = 1.0000000$
2	$\xi_1 = 0$	$A_1 = 0.88888889$
	$\xi_2 = -\xi_0 = 0.77459667$	$A_2 = A_0 = 0.55555556$
3	$\xi_2 = -\xi_1 = 0.33998104$	$A_2 = A_1 = 0.65214515$
	$\xi_3 = -\xi_0 = 0.86113631$	$A_3 = A_0 = 0.34785485$
4	$\xi_2 = 0$	$A_2 = 0.56888889$
	$\xi_4 = -\xi_0 = 0.90617985$	$A_4 = A_0 = 0.23692689$
	$\xi_3 = -\xi_1 = 0.53846931$	$A_3 = A_1 = 0.47862867$

بدست می‌آوریم:

$$I \approx \tfrac{1}{2}\{0.23692689g(-0.90617985) + 0.47862867g(-0.53846931) + 0.56888889g(0) + 0.47862867g(0.53846931) + 0.23692689g(0.90617985)\}$$
$$= 0.74682413$$

برای رسیدن به دقت قابل‌مقایسه ، برای قاعدهٔ ذوزنقه‌ای حدود 2800 زیرتقسیم ، و حال آنکه برای قاعدهٔ سیمپسون تقریبا 20 زیرتقسیم لازم است .

۶ ــ با به کار بردن کوادراتور گاوسی و $k = 3$ تقریبی از

$$I = \int_1^3 \frac{\sin^2 x}{x}\, dx$$

بدست آورید (مقدار درست ... I = 0.79482518 است .)

باز هم با تغییر متغیر x = t + 2 بازهٔ انتگرال‌گیری را به [-1,1] تبدیل می‌کنیم . در نتیجه

$$I = \int_{-1}^1 \frac{\sin^2 (t+2)}{t+2}\, dt = \int_{-1}^1 g(t)\, dt$$

با استفاده از نقاط و وزن‌ها از جدول مندرج در مثال ۵ بدست می‌آوریم :

$$I \approx 0.34735485[g(\xi_0) + g(\xi_3)] + 0.65214515[g(\xi_1) + g(\xi_2)]$$

برنامهٔ فرترن و نتیجه به قرار زیر است :

FORTRAN PROGRAM FOR EXAMPLE 6

```
C         EXAMPLE 7 GAUSSIAN INTEGRATION
          REAL I
          F(T) = SIN (T + 2.)**2/(T + 2.)
          T0 = -.86113631
          T2 =  .33998104
          T3 = -T0
          T1 = -T2
          I  = .34785485*(F(T0) + F(T3)) + .65214515*(F(T1) + F(T2))
          WRITE (6,1) I
          STOP
        1 FORMAT (34H EXAMPLE 5.6 GAUSSIAN INTEGRATION /
         15H0I = 1PE14.7)
          END

          I = 7.9482833-01
```

۷ ــ تقریبی از انتگرال

$$I(f) = \int_0^1 e^{-x^2}\, dx$$

به روش رامبرگ بدست آورید .

برنامهٔ کامپیوتری ذیل جدول زیر را برای مقادیر T_{ij} بدست می‌دهد .

جدول رامبرگ

0.7313700E 00				
0.7429838E 00	0.7468551E 00			
0.7458653E 00	0.7468258E 00	0.7468238E 00		
0.7465842E 00	0.7468238E 00	0.7468237E 00	0.7468237E 00	
0.7467639E 00	0.7468237E 00	0.7468237E 00	0.7468237E 00	0.7468237E 00
0.7468069E 00	0.7468212E 00	0.7468210E 00	0.7468210E 00	0.7468209E 00

از این جدول بهترین مقدار حاصل 0.7468237 است . مقدار واقعی عبارتست از:

$$I(f) = 0.7468241$$

اختلاف این دو بخاطر آنست که ما در نقاط لازم تقریبی از e^{-x^2} را بکار برده‌ایم .

FORTRAN PROGRAM FOR EXAMPLE 7

```
      DIMENSION T(10,10)
      M = 2
      KMAX = 6
      A = 0.
      B = 1.
      H = (B - A)/FLOAT(M)
      SUM = (F(A) + F(B))/2.
      MM1 = M - 1
      IF (MM1)                                40,10,8
    8 DO 9 I = 1,MM1
    9 SUM = SUM + F(A + FLOAT(I)*H)
   10 T(1,1) = SUM*H
      WRITE (6,600)
  600 FORMAT(1H110X15HROMBERG T-TABLE//)
      WRITE (6,601) T(1,1)
  601 FORMAT(7E15.7)
C
      DO 20 K = 2,KMAX
      H = H/2.
      M = M*2
      SUM = 0.
      DO 11 I = 1,M,2
   11 SUM = SUM + F(A + FLOAT(I)*H)
      T(K,1) = T(K - 1,1)/2. + SUM*H
      FOURJ = 1.
      DO 12 J = 2,K
      FOURJ = FOURJ*4.
C   SAVE DIFFERENCES FOR LATER CALC. OF RATIOS
      T(K - 1,J - 1) = T(K,J - 1) - T(K - 1,J - 1)
   12 T(K,J) = T(K,J - 1) + T(K - 1,J - 1)/(FOURJ - 1.)
   20 WRITE (6,601) (T(K,J),J = 1,K)
      KMAXM2 = KMAX - 2
      IF (KMAXM2)                             40,40,29
CALCULATE RATIOS
   29 WRITE (6,602)
  602 FORMAT(///11X15HTABLE OF RATIOS//)
      DO 35 K = 1,KMAXM2
      DO 30 J = 1,K
      RATIO = 0.
      IF (ABS(T(K + 1,J)) .GT. 0.) RATIO = T(K,J)/T(K + 1,J)
   30 T(K,J) = RATIO
   35 WRITE (6,603) (T(K,J),J = 1,K)
  603 FORMAT(8F10.2)
   40                                         STOP
      END
      FUNCTION F(X)
      F = EXP(-X*X)
      RETURN
      END
```

۸ ــ با توجه به فرمول قاعده ذوزنقه‌ای مرکب و خطای آن ، یعنی ،

$$\boxed{I(f) \approx T_N = h\sum_{i=1}^{N-1} f_i + \frac{h}{2}(f_0 + f_N)}$$

$$\boxed{E_N{}^T = -\frac{f''(\eta)h^2(b-a)}{12}}$$

عدد N را چنان تعیین کنید که تقریبی از $\int_0^1 e^{-x^2}\,dx$ را تا شش رقم اعشار درست بدست دهد ، فرض کنید که e^{-x^2} را می‌توان به دقت حساب کرده و این تقریب را حساب کنید .

حل :

در این مثال

$$f(x) = e^{-x^2},\ a = 0,\ b = 1$$

بنابراین خطای قاعدهٔ ذوزنقه‌ای مرکب $-f''(\eta)N^{-2}/12$ به ازای مقداری از $\eta \in (a,b)$ است . چون h معلوم نیست خطا بیشتر از

$$\max_{0\le\eta\le1} \frac{|f''(\eta)|N^{-2}}{12}$$

نیست . اما $f''(x) = e^{-x^2}(4x^2-2)$ و بعلاوه $f'''(x) = e^{-x^2}4x(3-2x^2)$ که در $x = 0$ و $x = \pm\sqrt{1.5}$ صفر می‌شود . بنابراین $\max|f''(x)|$ بر $[0,1]$ در $x = 0$ اتفاق می‌افتد ، پس

$$\max_{0\le\eta\le1} |f''(\eta)| = \max\{|f''(0)|,|f''(1)|\} = \max\{2,2e^{-1}\} = 2$$

لذا ، باید N را طوری انتخاب کنیم که :

$$\frac{2N^{-2}}{12} < 5\cdot10^{-7}$$

$$N^2 > \frac{10^6}{3} = \frac{10^7}{6\cdot5}$$

$$N > \frac{10^3}{\sqrt{3}} \approx 578$$

نتیجهٔ کامپیوتری زیر مقدار تقریبی را به ازای بعضی از مقادیر N بدست می‌دهد . منظور از I(SP) مقدار محاسبه‌شده با دقت معمولی و I(DP) مقدار محاسبه‌شده با

دقت مضاعف است .

N	$I(SP)$	$I(DP)$
50	7.4679947E-01	7.4679961D-01
100	7.4681776E-01	7.4681800D-01
200	7.4682212E-01	7.4682260D-01
400	7.4682275E-01	7.4682375D-01
800	7.4682207E-01	7.4682404D-01

برنامه این روش ذیلا آمده است .

FORTRAN PROGRAM FOR EXAMPLE 8 (SINGLE PRECISION)

```
C       EXAMPLE 5.2, TRAPEZOIDAL INTEGRATION
        F(X) = EXP(-X**2)
      1 WRITE(6,3)
        READ(5,4)A,B,N
        T = F(A)/2.
        N1 = N - 1
        H = (B - A)/FLOAT(N)
        DO 2 I = 1, N1
      2 T = F(A + FLOAT(I)*H) + T
        T = (F(B)/2. + T)*H
        WRITE(6,5)A,B,N,T
        GO TO 1
      3 FORMAT(36H EXAMPLE 5.2 TRAPEZOIDAL INTEGRATION)
      4 FORMAT(2E20.0,I5)
      5 FORMAT(18H0 INTEGRAL FROM A = 1PE14.7,7H TO B = 1PE14.7,
       19H FOR N = I5,4H IS 1PE14.7)
        END
```

۹ ــ برنامه‌ای برای قاعدهٔ ذوزنقه‌ای تصحیح‌شده بنویسید و مسئلهٔ قبل را با به کار بردن آن حل کنید .

FORTRAN PROGRAM

```
C       EXAMPLE 9   CORRECTED TRAPEZOID RULE
        F(X) = EXP(-X*X)
        FPRIME(X) = -2.*X*F(X)
        A = 0.
        B = 1.
        WRITE (6,600)
    600 FORMAT(9X1HN7X13HTRAPEZOID SUM7X13HCORR.TRAP.SUM)
        DO 10 N = 10,15
        H = (B - A)/FLOAT(N)
        NM1 = N - 1
        TRAP = (F(A) + F(B))/2.
        DO 1 I = 1,NM1
      1 TRAP = TRAP + F(A + FLOAT(I)*H)
        TRAP = H*TRAP
        CORTRP = TRAP + H*H*(FPRIME(A) - FPRIME(B))/12.
     10 WRITE (6,610) N,TRAP,CORTRP
    610 FORMAT(I10,2E20.7)
        STOP
        END
```

دقت معمولی

Single precision

N	TRAPEZOID SUM	CORR.TRAP.SUM
10	0.7462108E 00	0.7468239E 00
11	0.7463173E 00	0.7468240E 00
12	0.7463983E 00	0.7468240E 00
13	0.7464612E 00	0.7468240E 00
14	0.7465112E 00	0.7468240E 00
15	0.7465516E 00	0.7468241E 00

دقت مضاعف

Double precision

N	TRAPEZOID SUM	CORR.TRAP.SUM
10	7.4621080E-01	7.4682393E-01
11	7.4631727E-01	7.4682399E-01
12	7.4639825E-01	7.4682403E-01
13	7.4646126E-01	7.4682406E-01
14	7.4651126E-01	7.4682408E-01
15	7.4655159E-01	7.4682409E-01

۱۰ــ برنامه‌ای برای قاعدهٔ سیمپسون بنویسید و مثال ۹ را با دقت معمولی و مضاعف حل کنید . نتایج و برنامه به قرار زیر است .

* نتایج کامپیوتری مثال ۱۰

N	$I(SP)$	$I(DP)$
25	7.4682406E-01	7.4682413D-01
50	7.4682400E-01	7.4682413D-01
100	7.4682392E-01	7.4682413D-01

FORTRAN PROGRAM FOR EXAMPLE 10 (SINGLE PRECISION)

```
C       EXAMPLE 5.4 SIMPSON'S RULE
        F(X) = EXP(-X**2)
        WRITE (6,3)
      1 READ(5,4)A,B,N
        H = (B - A)/FLOAT(N)
        HOV2 = H/2.
        S = 0.
        HALF = F(A + HOV2)
        NM1 = N - 1
        DO 2 I = 1,NM1
        X = A + FLOAT(I)*H
        S = S + F(X)
      2 HALF = HALF + F(X + HOV2)
        S = (H/6.)*(F(A) + 4.*HALF + 2.*S + F(B))
        WRITE(6,5)A,B,N,S
        GO TO 1
      3 FORMAT(28H EXAMPLE 5.4 SIMPSON'S RULE)
      4 FORMAT(2E20.0,I5)
      5 FORMAT(18H0 INTEGRAL FROM A = 1PE14.7,7H TO B = 1PE14.7,
       19H FOR N = I5,4H IS 1PE14.7)
        END
```

■ مسائل

۱ـ فرمول زیر باید برای تقریب عددی یک انتگرال به‌کار رود .

$$\int_a^b f(x)\,dx \approx \sum_{i=0}^{n} a_i f(x_i)$$

معیار دقت این فرمول آن است که فرمول برای تمام چند جمله‌ایهای از درجهٔ کوچکتر یا مساوی n دقیق باشد . اگر $a=0$ ، $b=6$ و مقادیر $f(x)$ درنقاط $x_0=0$ $x_1=1,\ x_2=5,\ x_3=6,$ معلوم باشد مقدار ضرائب a_i را که در شرط دقت فوق صادق باشد پیدا کنید .

۲ـ فرمولی با شکل کلی

$$I = a_0 f(x_0) + a_1 f(x_1) + a_2 f(x_2)$$

باید برای تقریبی از انتگرال $f(x)$ بین x_0 و x_2 که درآن $x_n = x_0 + n.h$ به‌کار رود . معیار دقت آن است که فرمول برای چند جمله‌ایهای حتی‌الامکان از درجهٔ بالا دقیق باشد ، ضرائب a_2 , a_1, a_0 را پیدا کنید . آیا این فرمول به هنگام انتگرال‌گیری از یک چند جمله‌ای درجهٔ سوم دقیق است ؟

۳ـ چند جمله‌ای درجه دومی را که بربازه $[-2,2]$ برچند جمله‌ای 1 و x عمود باشد پیدا کنید . سپس گره‌های x_i و وزنهای a_i از فرمول انتگرال‌گیری گاوسی زیر را پیدا کنید .

$$I = \sum_{i=0}^{1} a_i f(x_i)$$

۴ـ نتایج بالا را با به‌کار بردن این معیار که انتگرال‌گیری گاوسی نشان داده شده برای تمام چند جمله‌ایهای تا درجه ۳ دقیق است ، آزمایش کنید و بعنوان مبنائی جهت بدست آوردن گره‌ها و وزنهای فرمول به‌کار برید .

۵ ـ با استفاده از قاعده ذوزنقه‌ای مرکب با بازه‌های $h=0.4$. انتگرال $\int_0^{0.8} \sin x\,dx$ را حساب کنید .

این را با نتیجه به‌کار بردن قاعدهٔ سیمپسون با همان بازه مقایسه کنید . آیا قاعدهٔ ذوزنقه‌ای ، با دوبرابر تعداد بازه‌ها از قاعدهٔ سیمپسون دقیقتر است ؟ اثر استفاده از برونیابی ریچارد سن روی مقادیر انتگرال ،که از دو کاربرد قاعدهٔ ذوزنقه‌ای در بالا بدست می‌آید ، چگونه است ؟

۶ ـ تقریبی از انتگرال

$$\int_0^1 e^{-x^2}\,dx$$

تا ۶ رقم اعشار درست با دقت ساده و مضاعف به کمک برنامه کامپیوتری به روش ذوزنقه‌ای محاسبه و نتایج را توجیه نمائید .

۷ ــ تابع $f(x)$ بر بازهٔ $[0,1]$ بصورت زیر تعریف شده است :

$$f(x) = \begin{cases} x & 0 \le x \le \frac{1}{2} \\ 1 - x & \frac{1}{2} \le x \le 1 \end{cases}$$

مقدار $\int_0^1 f(x)\,dx$ را با به کاربردن قاعده‌های زیر حساب کنید .

الف : قاعدهٔ ذوزنقه‌ای بر $[0,1]$

ب : قاعدهٔ ذوزنقه‌ای ابتدا بر $[0,\frac{1}{2}]$ و بعد بر $[\frac{1}{2},1]$

ج : قاعدهٔ سیمپسون بر $[0,1]$

د : قاعدهٔ تصحیح‌شده ذوزنقه‌ای بر $[0,1]$

۸ ــ هریک از پنج قاعدهٔ ذکرشده را برای پیدا کردن تقریبهائی از $I = \int_0^1 e^{-x}\,dx$ به کار برید . نتایج را با مقدار دقیق $I = -e^{-x}|_0^1 = -e^{-1} + 1 = 0.63212.$ مقایسه کنید . (در محاسبه فوق مقادیر $e^0 = 1, e^{-1} = 0.36788, e^{-1/2} = 0.60653$ را نیاز خواهید داشت)

۹ ــ با استفاده از برنامهٔ کامپیوتری با به کار بردن قاعده‌های زیر انتگرال‌های داده شده را محاسبه نمائید .

الف : قاعدهٔ ذوزنقه‌ای

ب : قاعدهٔ سیمپسون

ج : قاعدهٔ تصحیح‌شدهٔ ذوزنقه‌ای

(a) $I = \int_0^1 xe^{-x}\,dx$

(b) $I = \int_0^1 x \sin x\,dx$

(c) $I = \int_0^1 (1 + x^2)^{3/2}\,dx$

و دقت نتایج را در هر حالت محاسبه نمائید .

۱۰ ــ تخمینی از $I(f) = \int_a^b f(x)\,dx$ با تقریب 10^{-6} و با استفاده از انتگرال‌گیری رامبرگ برای هریک از توابع زیر حساب کنید .

(a) $f(x) = x^2$ — $a = 0, b = 1, M$ دلخواه

(b) $f(x) = \sin 101\pi x$ — $a = 0, b = 1, M = 1$

(c) $f(x) = 1 + \sin 10\pi x$ — $a = 0, b = 1, M = 1$

(d) $f(x) = |x - \frac{1}{2}|$ — $a = 0, b = 1, M = 1$ and $M = 3$

(e) $f(x) = \sqrt{x}$ — $a = 0, b = 1, M$ دلخواه

(f) $f(x) = \begin{cases} \dfrac{\sin x}{x} & x \ne 0 \\ 1 & x = 0 \end{cases}$ — $a = 0, b = 1$

فصل نهم
مقادیر و بردارهای ویژه

۹.۱ ■ معادلات اساسی

مسائل فیزیکی بسیاری وجود دارند که برای حل آنها مقادیری از λ که در معادلهٔ زیر صدق کند لازم است.

$$\mathbf{AX} = \lambda\mathbf{X} \tag{9.1}$$

که در آن $\mathbf{A}$ یک ماتریس $n \times n$ ، λ اسکالری موسوم به مقدار ویژه و $\mathbf{X}$ بردار ویژهٔ متناظر با آن است. چنین مسائلی در تحلیل پایداری و در معادلات ارتعاشات ساختمانها و مدارهای الکتریکی اتفاق می‌افتد. مثالی بوسیلهٔ ارتعاشات جسمی که از حالت تعادل خارج شده‌ارائه می‌شود. معادلات شکل نمونهٔ زیر را دارند.

$$\begin{aligned} a_{11}\ddot{x} + b_{11}x + a_{12}\ddot{y} + b_{12}y = 0 \\ a_{21}\ddot{x} + b_{21}x + a_{22}\ddot{y} + b_{22}y = 0 \end{aligned} \tag{9.2}$$

جوابها با $x = K_1 \cos(pt + \varepsilon_1)$ و $y = K_2 \cos(pt + \varepsilon_2)$ داده می‌شوند و جایگذاری این مقادیر در معادلات نتیجه می‌دهد

$$-p^2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \tag{9.3}$$

با قرار دادن معادلات به شکل ماتریسی، مقادیر $\lambda \equiv p^2$ لازم است که در معادلهٔ

$$\lambda\mathbf{AX} = \mathbf{BX}$$

یا

$$\mathbf{A}^{-1}\mathbf{BX} = \lambda\mathbf{X} \tag{9.4}$$

صدق کند. چون معادله (9.1) را می‌توان به‌شکل $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$ نوشت ملاحظه می‌کنیم که معادلهٔ (9.1) جوابی دارد که در آن مولفه‌ها ناصفرند اگر و فقط دترمینان $\mathbf{A} - \lambda\mathbf{I}$ صفر باشد یعنی،

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & \cdots & a_{nn}-\lambda \end{vmatrix} = 0 \tag{9.5}$$

می‌توان مشاهده کرد که بسط این دترمینان یک چند جمله‌ای درجهٔ n از λ می‌دهد و ریشه‌های این معادله مقادیر ویژه مطلوب هستند. این چند جمله‌ای به معادلهٔ مشخصه ماتریس $\mathbf{A}$ معروف است. سپس حل معادلات همگن (9.1) بردار ویژهٔ $\mathbf{X}$ را می‌دهد. در حالتی که مقادیر ویژه جملگی متمایزند به هر مقدار ویژه یک بردار ویژهٔ منحصر بفرد متناظر است. در حالت وجود مقدار ویژه‌ای با تکرار m مکن است m بردار ویژه یا کمتر وجود داشته باشد. درحالت اخیر کمتر از n بردار ویژه برای $\mathbf{A}$ موجود است‌وبردارهای ویژه نمی‌توانند یک مبنا برای فضای مورد بحث تشکیل دهند. این خصوصیت، بهنگام کوششی برای استفاده از روشهای تکراری بخش ۹.۳ سبب مشکلاتی می‌شود.

بحث‌فصل ۳ دربارهٔ حل معادلات چند جمله‌ای بر مشکل بدست آوردن جوابهای این مسئله تاکید داشت. در مسئلهٔ حاضر ضرائب چند جمله‌ای بوسیلهٔ فرایند حذف یا محاسبه دترمینان بدست می‌آیند و این بدین معنا است که ضرائب معادلهٔ مشخصه شامل خطاهای قابل ملاحظه‌ای هستند. در پرتو این مشکلات روش بسط به‌صورت یک چند جمله‌ای و بعد تعیین جواب، رضایتبخش نیست مگر برای‌ماتریسهای با مرتبهٔ خیلی پایین.

دو نوع روش وجود دارد که می‌توان به‌کار برد. روشهای تکراری برای استفاده بسیار ساده‌اند و در موقعیتهای معین یک ریشه را می‌توان بطور موثر پیدا کرد. این روشها را می‌توان تغییر داده و بیش از یک ریشه را بدست آورد امادرحالت کلی برای پیدا کردن تمام مقادیر ویژهٔ یک ماتریس به‌کار نمی‌روند. روشهایی که بهنگام لزوم تمام مقادیر ویژه به‌کار می‌روند برتبدیلات استوارند که ماتریس را به‌شکل ساده‌ای، که به آسانی برای پیدا کردن مقادیر ویژه حل می‌شود، تبدیل می‌کنند. اگر تبدیلات مشابهی انجام شود ماتریسهای جدید مقادیر ویژهٔ یکسان و برابر با ماتریس اولیه دارند و رابطهٔ ساده‌ای بین بردارهای ویژهٔ قدیم و جدید وجود دارد.

۹.۲ ■ چند نتیجه ماتریسی مفید

ارائه بعضی قسمتهای نظریهٔ ماتریسها که‌پیشنیاز کار در این فصل است مفید خواهد بود. معهذا، فرض می‌شود که خواننده قدری با حساب ماتریسها منجمله نظریه مقدار ویژه و بردار ویژه آشنائی داشته باشد. دراین فصل فرض می‌شود که ماتریس عناصر حقیقی دارد.

۱ــ یک ماتریس را متقارن گویند اگر$(i,j=1,2,\dots,n)$ $a_{ij}=a_{ji}$. مثالهای زیر این نکته را روشن می‌کنند .

$$\begin{bmatrix} 1 & -2 & 1 & -3 \\ -2 & 2 & -3 & 0 \\ 1 & -3 & -4 & 1 \\ -3 & 0 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 3 & -5 & 0 & 1 \\ 3 & 1 & -3 & 0 \\ -2 & 2 & 1 & -5 \\ -1 & -2 & 3 & 3 \end{bmatrix}$$

متقارن نامتقارن

۲ــ یک ماتریس را متعامد گویند اگر

$$\begin{aligned} \mathbf{a}_i^T\mathbf{a}_j &= 0, \qquad i\neq j \\ \mathbf{a}_i^T\mathbf{a}_i &= 1 \end{aligned} \tag{9.6}$$

که در آن بردار $\mathbf{a}_i$ ستون i ام $\mathbf{A}$ است . باید توجه کرد که معکوس یک ماتریس متعامد را بدون محاسبه می‌توان بدست آورد زیرابنابر دستگاه (9.6) $\mathbf{A}^T\mathbf{A}=\mathbf{I}$ ، لذا ، $\mathbf{A}^{-1}=\mathbf{A}^T$

۳ــ مقادیر ویژهٔ یک ماتریس متقارن حقیقی ، حقیقی‌هستند . بردارهای ویژهٔ متناظر با مقادیر ویژهٔ متمایز متعامدند .

۴ــ اگر یک ماتریس مرتبهٔ n دارای n مقدار ویژهٔ متمایز باشد آنگاه n بردار ویژه مستقل خطی وجود دارد که می‌توانند یک مبنا برای فضای بردارها تشکیل دهند . لذا . یک بردار دلخواه را می‌توان برحسب بردارهای ویژه بیان کرد ، یعنی

$$\mathbf{y}=\sum_{r=1}^{n} a_r\mathbf{X}^{(r)} \tag{9.7}$$

که در آن $\mathbf{X}^{(r)}$ $(r=1,2,\dots,n)$ بردارهای ویژهٔ مستقل خطی هستند .

۵ــ اگر $\mathbf{X}^{(i)}$ بردار ویژه‌ای متناظر با مقدار ویژهٔ λ_i باشد آنگاه $\mathbf{A}\mathbf{X}^{(i)}=\lambda_i\mathbf{X}^{(i)}$ و $\mathbf{A}^k\mathbf{X}^{(i)}=\lambda_i^k\mathbf{X}^{(i)}$. پس، اثر ضرب متوالی $\mathbf{A}$ در یک بردار ویژه عبارتست‌از ضرب متوالی آن بردار در اسکالر λ_i

۶ ــ دو ماتریس $\mathbf{A}$ و $\mathbf{B}$ را متشابه گویند درصورتی‌که یک ماتریس نامنفرد چون $\mathbf{P}$ وجود داشته باشد بقسمی که $\mathbf{B}=\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. مشاهدهٔ اینکه ماتریسهای متشابه مقادیر ویژه یکسان دارند ساده است.زیرا،اگر

$$\mathbf{A}\mathbf{X}=\lambda\mathbf{X}$$

آنگاه

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{X}=\lambda\mathbf{P}^{-1}\mathbf{X}$$

واگر

$$\mathbf{X}=\mathbf{P}\mathbf{Y} \tag{9.8}$$

آنگاه

$$P^{-1}APY = \lambda Y \tag{9.9}$$

بـــردارهای ویژهٔ A را می‌توان از بـــردارهای ویـــژهٔ B ، بوسیله رابطهٔ $X = PY$ بدست آورد .

۷ــ اگر $X^{(r)}$ بردار ویژه‌ای از یک ماتریس باشـد آنگاه هر مضـرب آن نیز یک بردار ویژه است . بعضی اوقات مناسب است که بردار ویژه نرمال شود و این را می‌توان بدو طریق انجام داد . یک روش نرمال سازی تقسیم تمام عناصر یک بردار به‌بزرگترین‌عنصر آن ، از نظر قدر مطلق،است بقسمی که بردارها ماکزیمم عنصرشان واحد باشد . در روش دیگر . هر عنصر را می‌توان برجذر مجموع مربعات عناصر بردار تقسیم نمود که در این حالت بردارها طول واحد دارند .

۹.۳ ▪ روشهای تکراری

۹.۳.۱ ▪ پیدا کردن مقدار ویژه حقیقی با بزرگترین اندازهٔ مطلق

روشهای تکراری اغلب برای مسائلی که در آنها فقط یک یا دو مقدار ویژه باید محاسبه شوند مناسبند ، اگرچه طرقی برای تعمیم این روشها جهت پیدا کردن تمام ریشه‌ها وجود دارد . در این بخش فرض می‌شود که تمام مقادیرویژه‌هامتمایزند . بنابراین ، بردارهای ویژه مستقل خطی هستند و یک بردار دلخواه را می‌توان به‌فرم زیر بیان کرد .

$$y = \sum_{r=1}^{n} a_r X^{(r)} \tag{9.10}$$

برای پیدا کردن بزرگترین مقدار ویژه و بردار ویژهٔ متناظر آن بردار دلخواه $y^{(0)}$ متوالیا " در ماتریس A ضرب می‌شود . درحالت کلی ، دنبالهٔ بردارهای $y^{(k)}$ به بردار ویژهٔ موردنظر و نسبت عناصر بردارهای ستوالی به‌مقدار ویژهٔ موردنظر همگرا خواهند بود . مثلا " ، فرض کنید λ_1 بزرگترین مقدار ویژه و $X^{(1)}$ بردار ویـــژهٔ متناظرش باشد . آنگاه ،

$$y^{(k)} = A^k y^{(0)} \tag{9.11}$$

$$= A^k \sum_{r=1}^{n} a_r X^{(r)}$$

$$= \sum_{r=1}^{n} a_r \lambda_r^k X^{(r)}$$

$$= \lambda_1^k \left[a_1 X^{(1)} + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k X^{(2)} + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k X^{(n)} \right] \tag{9.12}$$

مقادیر $(\lambda_r/\lambda_1)^k$ $(r \neq 1)$ وقتی k به‌سمت ∞ میل می‌کنـد به صفـر میل می‌کننـد و در نتیجه تمام جملات بجز جمله اول قابل اغماض می‌شوند . پس، وقتی k افزایش می‌یابد

$\mathbf{y}^{(k)}$ به مضربی از $\mathbf{X}^{(1)}$ ونسبت عنصر $y_i^{(k+1)}$ به عنصر متناظرش $y_i^{(k)}$ به λ_1 میل می‌کند.

می‌توان مشاهده کرد که اگر بردار اولیه شامل مولفهٔ بزرگی از $\mathbf{X}^{(1)}$ باشد همگرائی سریعتر خواهد بود. اگر مولفهٔ $\mathbf{X}^{(1)}$ اصلا" نباشد، یعنی، $a_1 = 0$ آنگاه، دیده‌می‌شود که دنباله نمی‌تواند به بردار ویژهٔ غالب همگرا باشد. معهذا، خطای گرد کردن بطور عادی مولفۀای از $\mathbf{X}^{(1)}$ تولید می‌کند که از این پس بزرگ شده و سرانجام غالب می‌شود. اگر همگرائی با یک انتخاب خاص بردار اولیه، کند باشد بعضی اوقات پیشرفت رضایتبخشی می‌توان با انتخاب بردار اولیهٔ دیگری بدست آورد. سرعت همگرائی بوضوح به نسبت قدر مطلق دو مقدار ویژه بزرگتر از بقیه بستگی دارد. وقتی این نسبت تقریبا" واحد باشد همگرائی بسیار کندی نتیجه خواهد شد.

روند محاسباتی کمی با فرایند توضیح داده شده در بالا فرق می‌کند بطوری‌که بتوان از رشد بدون کران $\mathbf{y}^{(k)}$ اجتناب کرد.

۱ــ بردار $\mathbf{y}^{(0)}$ را نرمال ساخته تا بزرگترین عنصر واحد داشته باشد.

۲ــ این بردار در $\mathbf{A}$ ضرب می‌شود.

۳ــ بردار جدید را با تقسیم هر عنصرش بر بزرگترین عنصر آن، که آن را با q_k نشان می‌دهیم، نرمال می‌سازیم

۴ــ ماتریس $\mathbf{A}$ را مرتبا" در بردار $\mathbf{y}^{(k)}$ ضرب می‌کنیم و بردار حاصل را برعامل q_k تقسیم می‌کنیم تا اینکه اختلاف q_k و q_{k+1} مقدار کوچک مشخص شود. مقدار q_k بزرگترین مقدار ویژه را می‌دهد و بردار $\mathbf{y}^{(k)}$ بردار ویژه متناظر آن است. نتایج کامپیوتری برای این روش در مثال ۹.۱ داده شده است.

۹.۳.۲ ■ ریشه های مختلط مزدوج

اگر بزرگترین ریشه مطلق مختلط باشد آنگاه، چون $\mathbf{A}$ حقیقی است، ریشهٔ دیگری با قدرمطلق مساوی وجود دارد که مزدوج آن می‌باشد. تحلیل قبلی دیگر مناسب نیست و ضرب کردن بوسیلهٔ $\mathbf{A}$ دنباله‌ای از بردارها که همگرا باشد تولید نمی‌کند. اگر بردار اولیه دلخواه $\mathbf{y}^{(0)}$ مثل قبل برحسب بردارهای ویژه بسط داده شود بعد از چندین بار ضرب بوسیله $\mathbf{A}$ تمام جملات بجز دو تا قابل اغماض خواهند بود، بنابراین،

$$\mathbf{y}^{(k)} \approx a_1\lambda_1^k\mathbf{X}^{(1)} + a_2\bar{\lambda}_1^k\bar{\mathbf{X}}^{(1)} \qquad (9.13)$$

تحلیل زیر نشان می‌دهد که چگونه می‌توان مقادیر ویژه را پیدا کرد. فرض کنید λ_1 و $\bar{\lambda}_1$ جوابهای معادلهٔ درجه دوم زیر باشند.

$$\lambda^2 + a\lambda + b = 0 \qquad (9.14)$$

که در آن a و b در این مرحله مقادیر نامعلومند . سه بردار متوالی از دنباله ، یعنی ، $\mathbf{y}^{(k)}$, $\mathbf{y}^{(k+1)}$, $\mathbf{y}^{(k+2)}$ را گرفته و ترکیب خطی زیر را تشکیل می‌دهیم .

$$\begin{aligned}
&\mathbf{y}^{(k+2)} + a\mathbf{y}^{(k+1)} + b\mathbf{y}^{(k)} \\
&\approx a_1\lambda_1^{k+2}\mathbf{X}^{(1)} + a_2\bar{\lambda}_1^{k+2}\bar{\mathbf{X}}^{(1)} + a[a_1\lambda_1^{k+1}\mathbf{X}^{(1)} + a_2\bar{\lambda}_1^{k+1}\bar{\mathbf{X}}^{(1)}] \\
&\qquad + b[a_1\lambda_1^{k}\mathbf{X}^{(1)} + a_2\bar{\lambda}_1^{k}\bar{\mathbf{X}}^{(1)}] \\
&= a_1\lambda_1^{k}\mathbf{X}^{(1)}[\lambda_1^2 + a\lambda_1 + b] + a_2\bar{\lambda}_1^{k}\bar{\mathbf{X}}^{(1)}[\bar{\lambda}_1^2 + a\bar{\lambda}_1 + b] = 0
\end{aligned} \tag{9.15}$$

پس با دانستن a و b می‌توان یک ترکیب خطی از سه بردار متوالی را تشکیل داد که تقریبا" مساوی بردار صفر شود . در مسئله‌ای که بررسی می‌کنیم محاسبات را معکوس انجام می‌دهیم ، یعنی ، سعی می‌کنیم مقادیر a و b را چنان تعیین کنیم که یک ترکیب خطی از سه‌بردار متوالی بردار صفر را بدهند . اگر هر دو مولفهٔ این بردارها ، در نظر گرفته شوند و ترکیب خطی موردنظر مساوی صفر قرار داده شود دو معادله برای دو مجهول a و b تولید می‌شود و این معادلات را می‌توان بطور عادی حل کرد .

$$\begin{aligned}
y_r^{(k+2)} + ay_r^{(k+1)} + by_r^{(k)} &= 0 \\
y_s^{(k+2)} + ay_s^{(k+1)} + by_s^{(k)} &= 0
\end{aligned} \tag{9.16}$$

وقتی مقادیر a و b معلوم باشند . معادلهٔ (9.14) را می‌توان حل کرد تا مقادیر λ_1 و $\bar{\lambda}_1$ بدست آیند . بعد بردار ویژه $\mathbf{X}^{(1)}$ از دو بردار متوالی پیدا می‌شود .

$$\mathbf{y}^{(k+1)} - \bar{\lambda}_1\mathbf{y}^{(k)} = a_1\lambda_1^{k}[\lambda_1 - \bar{\lambda}_1]\mathbf{X}^{(1)} \tag{9.17}$$

و $\mathbf{X}^{(1)}$ به روش مشابهی بدست می‌آید .

$$\mathbf{y}^{(k+1)} - \lambda_1\mathbf{y}^{(k)} = a_2\bar{\lambda}_1^{k}[\bar{\lambda}_1 - \lambda_1]\bar{\mathbf{X}}^{(1)} \tag{9.18}$$

عملا" فرآیند تکرار تا زمانی که مقادیر a و b ، صرفنظر از اینکه‌کدام مولفه‌های بردارهای $\mathbf{y}^{(k)}$ به‌کار روند و مقادیر از یک تکرار به تکرار بعدی تغییر نکند ، کاملا" ثابت شوند ادامه می‌یابد . نتایج کامپیوتری برای این روش در مثال ۹.۳ داده شده است .

۹.۳.۳ ▪ تکرار معکوس برای کوچکترین ریشهٔ حقیقی

با تغییر ساده‌ای در روش بالا می‌توان تکرار را برای بدست آوردن کوچکترین مقدار ویژه به‌کار برد . جریمه آن‌است که‌فرآیند جدید شامل حل مکرر دستگاههای معادلات خطی است . این فرآیند براین خاصیت استوار است که مقادیر ویژه $\mathbf{A}^{-1}$ معکوس مقادیر ویژهٔ $\mathbf{A}$ هستند . بنابراین ، کوچکترین مقدار ویژهٔ $\mathbf{A}$ بزرگترین مقدار ویژهٔ $\mathbf{A}^{-1}$ است . بردارهای ویژهٔ $\mathbf{A}$ و $\mathbf{A}^{-1}$ یکسان هستند . لذا ، با $\mathbf{A}^{-1}$ تکرار انجام می‌دهیم . یعنی ،

$$\mathbf{y}^{(k)} = \mathbf{A}^{-1}\mathbf{y}^{(k-1)} \tag{9.19}$$

اما در فصل ۴ نشان داده شد که‌فرایند تعیین $\mathbf{A}^{-1}$ بطور صریح ، موثر نبوده و باید با

حذف گاوس جایگزین شود. در این مسئله این عمل را می‌توان با بدست آوردن دنبالهٔ بردارهای $\mathbf{y}^{(k)}$ بوسیلهٔ حل متوالی معادلات زیر انجام داد.

$$\mathbf{A}\mathbf{y}^{(k)} = \mathbf{y}^{(k-1)}, \qquad k = 1, 2, \ldots \tag{9.20}$$

بعد از اولین تبدیل به فرم مثلثی، عملیات بعدی شامل ضرب بوسیلهٔ یک ماتریس مثلثی و جایگذاری پسرو است بقسمی که زمان محاسبه کاهش می‌یابد. بعد از تعیین هر جواب بردار $\mathbf{y}^{(k)}$ بوسیله تقسیم به عنصر q_k، با بزرگترین اندازه، نرمال می‌شود. فرایند هنگامی متوقف می‌شود که اختلاف بین مقادیر متوالی q_k کمتر از مقدار مشخص شده‌ای باشد. مقدار کوچکترین مقدار ویژهٔ $\mathbf{A}$ بوسیلهٔ q_k^{-1} و بردار ویژه با $\mathbf{y}^{(k)}$ داده می‌شود.

۹.۳.۴ ■ پیدا کردن نزدیکترین مقدار ویژه به یک مقدار مفروض

فرض کنید $\mathbf{B} = \mathbf{A} - p\mathbf{I}$ که در آن نزدیکترین مقدارویژه‌به p خواسته‌شده‌است. بردارهای ویژهٔ $\mathbf{A}$ در رابطهٔ $\mathbf{AX} = \lambda\mathbf{X}$ صدق می‌کنند و از آنجا، آنها در معادلهٔ

$$(\mathbf{A} - p\mathbf{I})\mathbf{X} = (\lambda - p)\mathbf{X} \tag{9.21}$$

نیز صدق می‌کنند. نتیجه اینکه بردارهای ویژهٔ $\mathbf{B}$ و $\mathbf{A}$ یکسان هستند و مقادیر ویژه جدید $\lambda - p$ هستند پس، مقدار ویژهٔ λ_t که نزدیکترین به p است با کوچکترین مقدار ویژه $\mathbf{B}$ متناظراست و روش بخش قبلی را می‌توان برای‌یافتن $(\lambda_t - p)^{-1}$ به‌کار برد. اگر مقسوم‌علیه مانند قبل q_k باشد آنگاه $\lambda_t = p + 1/q_k$ و بردار ویژه‌برابر $\mathbf{y}^{(k)}$ است نتایج کامپیوتری برای این روش در مثال ۹.۲ داده شده است.

۹.۳.۵ ■ توسیع روش بالا

روش تکراری در حالاتی که بزرگترین و کوچکترین مقدار ویژه، یا مقدار ویژهٔ خاصی خواسته شده باشد مناسبترین است و مسائل فیزیکی متعددی که از این نوع هستند وجود دارد. اکنون فرایندی برای توسیع طرح بالا، جهت یافتن تمام مقادیر ویژه یک ماتریس، بررسی خواهد شد. اما، فرایندی که تشریح می‌شود برای ماتریسهای بزرگ توصیه نمی‌شود زیرا، انباشتگی خطاها می‌تواند بسیار زیاد باشد. مبنای‌این روش حذف کردن بزرگترین‌مقدار ویژه و بردار ویژه به‌گونه‌ای است که طرح تکراری بخش ۹.۲/۱ بزرگترین مقدار ویژه بعدی را بدست دهد این فرایند به تقلیل معروف است و دو روش تشریح خواهد شد، یکی قابل اعمال بر هر ماتریسی است، و دومی محدود به ماتریسهای متقارن است.

فرض کنید λ_1 بزرگترین مقدار ویژه و $\mathbf{X}^{(1)}$ بردار ویژهٔ نظیر آن باشد. ماتریس

A برحسب اینکه کدام عنصر $X^{(1)}$ ماکزیمم باشد افراز می‌شود . فرض کنید که اولین عنصر بزرگترین باشد که در آن حالت ماتریس A با رابطهٔ

$$A = \begin{pmatrix} a_1 \\ B \end{pmatrix} \tag{9.22}$$

نمایش داده می‌شود که در آن a_1 اولین سطر A و B ماتریس $n-1 \times n$ از بقیه‌سطرها است . چون بزرگترین مولفه بردار $X^{(1)}$ اولین مولفه است ، تمام بردارهای مورد بحث نرمال خواهند شد بگونه‌ای که اولین مولفه آنها واحد باشد .

سپس ماتریس

$$A_1 = A - X^{(1)}a_1 \tag{9.23}$$

تشکیل می‌شود . نشان داده خواهد شد که A_1 مقادیر ویژهٔ $\lambda_i \ (i=2,3,\ldots,n)$ را داراست که همانند A هستند و مقدار ویژهٔ باقیمانده صفر است .

هر مقدار ویژه دیگر λ_i از A را با بردار ویژهٔ نظیرش $X^{(i)}$ درنظر بگیرید .

$$A_1(X^{(1)} - X^{(i)}) = A(X^{(1)} - X^{(i)}) - X^{(1)}a_1(X^{(1)} - X^{(i)}) \tag{9.24}$$

حال چون a_1 اولین سطر A است ، حاصلضرب $a_1X^{(i)}$ اولین عنصر $\lambda_iX^{(i)}$ است . این برابر λ_i است زیرا ، بردارهانرمال شده‌اند . پس، طرف راست معادلهٔ (9.24) برابر است با :

$$\begin{aligned} &\lambda_1X^{(1)} - \lambda_iX^{(i)} - X^{(1)}(\lambda_1 - \lambda_i) \\ &= \lambda_i(X^{(1)} - X^{(i)}) \end{aligned} \tag{9.25}$$

این نشان می‌دهد که مقادیر ویژهٔ $\lambda_i \ (i=2,3,\ldots,n)$ مقادیر ویژهٔ ماتریس‌جدید هستند و بردارهای ویژه متناظر $X^{(1)} - X^{(i)}$ می‌باشند . همچنین

$$\begin{aligned} A_1X^{(1)} &= AX^{(1)} - X^{(1)}a_1X^{(1)} \\ &= AX^{(1)} - \lambda_1X^{(1)} = 0 \end{aligned} \tag{9.26}$$

بطوری که مقدار ویژه باقیمانده صفر است . اکنون می‌توان ماتریس A_1 را برای تکرار به‌کار برده بزرگترین مقدار ویژه بعدی را بدست آورد و بهمین ترتیب برای بقیه .

باید تذکر داده شود که می‌توان یک ساده کننده‌ای به‌کار گرفت که محاسبات لازم کاهش یابد . اولین مولفهٔ تمام بردارهای ویژهٔ ماتریس جدید صفر است و لذا ، تنها نیاز به یافتن $n-1$ مولفه از بردارهای ویژه می‌باشد . همچنین اعضای سطر اول ماتریس باید صفر باشد زیرا ، اولین سطر A و $X^{(1)}a_1$ هر دو a_1 است . بنابر این ، مسئله‌را می‌توان بعنوان یک مسئلهٔ $n-1$ بعدی ، با حذف سطر اول و ستون اول ماتریس A_1 حل کرد .

وقتی که بردار ویژهٔ $r^{(2)}$ متناظر با بزرگترین مقدار ویژهٔ ماتریس A_1 ، بدست

آمده باشد می‌توان آن را برای یافتن بردار ویژهٔ $\mathbf{A}$ ، متناظر با مقدار ویژهٔ λ_2 ، به‌کار برد. فرض می‌شود که λ_2 بزرگترین مقدار ویژهٔ بعدی است. از بحث قبلی داریم که $\mathbf{r}^{(2)}$ مضربی از $\mathbf{X}^{(1)} - \mathbf{X}^{(2)}$ است. یعنی

$$c\mathbf{r}^{(2)} = \mathbf{X}^{(1)} - \mathbf{X}^{(2)}$$

پس

$$\mathbf{X}^{(2)} = \mathbf{X}^{(1)} - c\mathbf{r}^{(2)} \tag{9.27}$$

که در آن c ثابتی است که باید بدست آید. همچنین

$$\mathbf{a}_1\mathbf{X}^{(2)} = \mathbf{a}_1\mathbf{X}^{(1)} - c\mathbf{a}_1\mathbf{r}^{(2)}$$
$$\lambda_2 = \lambda_1 - c\mathbf{a}_1\mathbf{r}^{(2)}$$

و این مقدار c را می‌دهد

$$c = (\lambda_1 - \lambda_2)/\mathbf{a}_1\mathbf{r}^{(2)} \tag{9.28}$$

که می‌تواند بعداً" در معادله (9.27) قرار گرفته $\mathbf{X}^{(2)}$ را بدست دهد.

در حالت متقارن بودن، روش دیگری وجود دارد که در مقابل خطاهای گرد کردن کمتر حساس است. در این تحلیل بردارها با تقسیم به مجموع مربعات عناصر بردار نرمال شده‌اند. این عمل طول بردارها را واحد می‌کند. ماتریس

$$\mathbf{A}_1 = \mathbf{A} - \lambda_1\mathbf{X}^{(1)}[\mathbf{X}^{(1)}]^T \tag{9.29}$$

را تشکیل دهید. $\mathbf{A}_1$ مقادیر ویژه همانند $\mathbf{A}$ دارد. یعنی، $\lambda_i\ (i=2,3,\ldots,n)$ وریشه باقیمانده $\mathbf{A}_1$ صفر است. این بخاطر

$$\begin{aligned}\mathbf{A}_1\mathbf{X}^{(i)} &= \mathbf{A}\mathbf{X}^{(i)} - \lambda_1\mathbf{X}^{(1)}[\mathbf{X}^{(1)}]^T\mathbf{X}^{(i)}\\ &= \lambda_i\mathbf{X}^{(i)}, \qquad i\neq 1\end{aligned} \tag{9.30}$$

نتیجه می‌شود، زیرا بردارهای ویژه یک ماتریس متقارن متعامد هستند. همچنین

$$\begin{aligned}\mathbf{A}_1\mathbf{X}^{(1)} &= \mathbf{A}\mathbf{X}^{(1)} - \lambda_1\mathbf{X}^{(1)}[\mathbf{X}^{(1)}]^T\mathbf{X}^{(1)}\\ &= \mathbf{A}\mathbf{X}^{(1)} - \lambda_1\mathbf{X}^{(1)}\\ &= 0\end{aligned} \tag{9.31}$$

تحلیل بالا همچنین نشان می‌دهد که بردارهای ویژهٔ نظیر مقادیر ویژهٔ $\lambda_i\ (i=2,3,\ldots,n)$ برای ماتریسهای $\mathbf{A}$ و $\mathbf{A}_1$ یکسان هستند. این روش در مثال عددی (۹۰۱) نشان داده شده است.

۹.۴ ■ روشهای تبدیلی

۹.۴.۱ ■ روش ژاکوبی

روشهای این بخش از تبدیلات متشابه استفاده کرده تبدیلی از ماتریس با همان

مقادیر ویژه، اما با فرم ساده‌تر، بدست می‌آورند. ماتریسهای تبدیلی که به‌کار می‌روند ماتریسهای متعامد هستند، زیرا می‌توان نشان داد که چنین ماتریسهایی برای می‌نیمم کردن خطاهای فرایند مناسبترین هستند.

ساده‌ترین شکلی که می‌توان بدست آورد شکل قطری است زیرا، در این صورت مقادیر ویژه اعضای قطری بوده مستقیما" در دسترس خواهند بود. روش ژاکوبی جهت تولید یک شکل قطری بوسیله حذف سیستماتیک عناصر غیر قطری، طراحی شده است. اما، این روش یک فرایند تکراری است و به تعدادی نامتناهی مرحله نیاز دارد. این باعث دو نقیصه می‌شود، اولا"، فرایند ممکن است با کندی زیاد همگرا باشد یا اصلا" همگرا نباشد، و ثانیا"، نیاز به قطع فرایند ممکن است خطاهایی وارد کند که یافتن جوابهای درست را جدا" به مخاطره اندازد.

طرح محاسباتی سر راست می‌باشد. ماتریس جدید $A_1 = P_1^{-1}AP_1$ با استفاده‌از ماتریس P_1 که یک صفر غیر قطری در A تولید می‌کند، تشکیل می‌شود. متاسفانه، در فرایند ژاکوبی، با بدست آوردن هر صفر جدید معمولا" عنصر جدیدی در محل‌صفرقبلی ایجاد می‌شود فرایند یا بوسیلهٔ کار سیستماتیک در امتداد یک سطر و سپس سطر بعدی، با تولید عناصر صفر، یا بوسیله حذف بزرگترین صفر غیر قطری در هر مرحله، ادامه می‌یابد. وقتی تمام عناصر غیر قطری ازنظر قدرمطلق کمتر از عدد کوچک مشخصی باشند فرایند خاتمه می‌یابد. دراین صورت، مقادیر ویژه عناصر قطری گرفته می‌شوند.

چون محاسبات با ماتریسهای متعامد انجام می‌شود. احتیاجی به محاسبهٔ عکس این ماتریسها، که با رابطه $P_r^{-1} = P_r^T$ داده می‌شود، نیست. ماتریس نهایی عبارتست‌از:

$$A_r = P_r^T P_{r-1}^T \cdots P_1^T A P_1 \cdots P_{r-1} P_r \tag{9.32}$$

و اگر $Y^{(r)}$ بردار ویژه A_1 باشد بردار ویژهٔ ماتریس اولیه A عبارتست‌از:

$$P_1 \cdots P_{r-1} P_r Y^{(r)} \tag{9.33}$$

ماتریسهای متعامدی که در روشهای ژاکوبی‌و گیونز Givens به‌کار می‌روند توسیعهایی از ماتریسهای دوران در دستگاه دوبعدی هستند. ماتریس دوران $n \times n$ جهت‌دوران در صفحه (r,s) بوسیله ماتریس واحد $n \times n$ با چهار تغییر زیر، داده می‌شود.

$$a_{rr} = \cos\theta, \qquad a_{rs} = -\sin\theta$$
$$a_{sr} = \sin\theta, \qquad a_{ss} = \cos\theta$$

مثلا"، ماتریس زیر با دوران در صفحهٔ $2,3$ متناظر است.

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c & -s & 0 \\ 0 & s & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \begin{array}{l} c = \cos\theta \\ s = \sin\theta \end{array} \tag{9.34}$$

تبدیل $P_1^TAP_1$ ماتریس زیر را می‌دهد

$$\begin{bmatrix} a_{11} & ca_{12}+sa_{13} & -sa_{12}+ca_{13} & a_{14} \\ ca_{21}+sa_{31} & c^2a_{22}+csa_{23}+csa_{32}+s^2a_{33} & -csa_{22}+c^2a_{23}-s^2a_{32}+csa_{33} & ca_{24}+sa_{34} \\ -sa_{21}+ca_{31} & -csa_{22}-s^2a_{23}+c^2a_{32}+csa_{33} & s^2a_{22}-csa_{23}-csa_{33}+c^2a_{32} & ca_{34}-sa_{24} \\ a_{41} & ca_{42}+sa_{43} & -sa_{42}+ca_{43} & a_{44} \end{bmatrix} \tag{9.35}$$

روش ژاکوبی یک عنصر را ، با انتخاب مقدار θ بقسمی که عنصر واقع در محل $(2,3)$ صفر شود ، صفر می‌کند. این روش بطور نرمال برای ماتریسهای متقارن به‌کار می‌رود بقسمی که لازم است

$$(c^2-s^2)a_{23}+cs(a_{33}-a_{22})=0 \quad \text{or} \quad \tan 2\theta=\frac{2a_{23}}{a_{22}-a_{33}} \tag{9.36}$$

نتایج کامپیوتری برای این روش در مثال ۹.۴ داده شده است .

۹.٤.۲ ■ روش گیونز

روش گیونز بر تبدیلات ماتریسی از همان نوع روش ژاکوبی استوار است . اما چنان طراحی شده که هر صفری که ایجاد می‌شود در تبدیلات بعدی حفظ می‌شود . وقتی دورانی در صفحه (r,s) انجام می‌شود عنصر $(r-1,s)$ ، به‌ازای $r=1,2,\dots,n-1$ و $s=r+2,r+3,\dots,n$ حذف می‌شود پس، برای دوران فوق در صفحهٔ $2,3$ مقدار θ باید انتخاب شود تا در روابط زیر صدق کند .

$$-sa_{12}+ca_{13}=0$$

$$\tan\theta=\frac{a_{13}}{a_{12}}$$

یا بطور کلی‌تر

$$\tan\theta=\frac{a_{r-1,s}}{a_{r-1,r}} \tag{9.37}$$

می‌توان مشاهده کرد که عناصر قطر اصلی ، و عناصر بلافاصله بالا و پایین عناصر قطر اصلی غیر صفر باقی می‌مانند پس، نتیجه نهایی این فرایند فرم قطری ساده نیست بلکه فرم معروف سه قطری است.

$$\begin{bmatrix} x & x & & & 0 & & \\ x & x & x & & & & \\ & x & x & x & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & x & x & x \\ & & 0 & & & x & x & x \\ & & & & & & x & x \end{bmatrix} \tag{9.38}$$

مقادیر ویژه فرم سه قطری را ، مثل حالت قطری ، فورا" نمی‌توان بدست آورد . اما ، روش حل به اندازه‌ای ساده است که تبدیل به فرم سه قطری‌را گام محاسباتی با ارزشی می‌سازد . این روش در بخش ۹.۵ توضیح داده می‌شود .

بسادگی می‌توان ملاحظه کرد که اگر محل حذف بطور سیستماتیک در امتداد سطر اول و شروع از عنصر 1 و 3 و بعد در امتداد سطر دوم و شروع از عنصر2,4 وغیره انجام شود صفرها حفظ خواهند شد . نتایج کامپیوتری برای‌این روش در مثال ۹.۵ داده شده‌است.

اگر ماتریس A متقارن نباشد روش گیونز را باز هم می‌توان اعمال کرد . اما ،فرم نهایی به‌شکل سه قطری متقارن نخواهد بود بلکه به‌شکل هسنبرگ Hessenberg است .

$$\begin{bmatrix} x & x & & & \\ x & x & x & 0 & \\ \cdots & \cdots & \cdots & & \\ \cdots & \cdots & \cdots & \cdots & \\ \cdots & \cdots & x & x & x \\ \cdots & \cdots & \cdots & x & x \end{bmatrix} \qquad (9.39)$$

۹.۴.۳ ■ روش هاوس هُلدر

با وجود اینکه روش گیونز مزیت قابل توجهی بر روش ژاکوبی داشت بوسیلهٔ روش هاوس‌هلدر جانشین شده است . این روش نیز تبدیلات متعامد را برای تغییر یک ماتریس متقارن به فرم سه قطری یا غیر متقارن را به فرم هسنبرگ به‌کار می‌برد.مزیت این روش آن است که تمام صفرهای ممکن در یک سطر تنها بوسیلهٔ یک تبدیل تولید می‌شود . بنابراین روش هاوس هلدر فقط $n-2$ تبدیل متشابه لازم دارد ، در مقایسه‌با روش گیونز که به $(n^2-3n+2)/2$ تبدیل نیاز دارد . روش هاوس هلدر ازنظر محاسبه‌ای بیشتر پیچیده است اما ، صرفه‌جویی قابل ملاحظه‌ای در وقت کامپیوتر وجود دارد . همچنین ، تقلیل عملیات کامپیوتری انتشار خطا را کاهش می‌دهد .

این روش با معادلات

$$\begin{aligned} \mathbf{A}_0 &= \mathbf{A} \\ \mathbf{A}_r &= \mathbf{P}_r^T \mathbf{A}_{r-1} \mathbf{P}_r, \qquad r = 1, 2, \ldots \end{aligned} \qquad (9.40)$$

تعریف می‌شود که در آن

$$\mathbf{P}_r = \mathbf{I} - 2\omega^{(r)}[\omega^{(r)}]^T, \qquad [\omega^{(r)}]^T\omega^{(r)} = 1 \qquad (9.41)$$

بعنوان یک مثال ، ماتریس $\mathbf{P}_1$ که صفرهایی در محلهای (1,3) و (1,4) ایجـــاد می‌کنداز

$$(\omega^{(1)})^T = (0, \omega_2^{(1)}, \omega_3^{(1)}, \omega_4^{(1)})$$

$$\mathbf{P}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-2[\omega_2^{(1)}]^2 & -2[\omega_2^{(1)}][\omega_3^{(1)}] & -2[\omega_2^{(1)}][\omega_4^{(1)}] \\ 0 & -2[\omega_2^{(1)}][\omega_3^{(1)}] & 1-2[\omega_3^{(1)}]^2 & -2[\omega_3^{(1)}][\omega_4^{(1)}] \\ 0 & -2[\omega_2^{(1)}][\omega_4^{(1)}] & [\omega_3^{(1)}][\omega_4^{(1)}] & 1-2[\omega_4^{(1)}]^2 \end{bmatrix} \tag{9.42}$$

استفاده می‌کند مقادیر $\omega_2^{(1)}$ $\omega_3^{(1)}$, و $\omega_4^{(1)}$ را می‌توان از $\mathbf{A}_0$ محاسبه کرد . تفصیل این محاسبات را می‌توان در ویلکنسن Wilkinson (1960) و رالستون Ralston (1965) یافت.

۹.۵ ■ مقادیر ویژه یک ماتریس سه‌قطری

هرچند روشهای فوق‌الذکر اثر قابل توجهی در ساده کردن فرم ماتریسها داشتند ، اگر ماتریس سه قطری نتیجه شده برای حل ساده‌ای مناسب‌نبود ارزش کمی می‌داشتند . در واقع ، فرم سه قطری به یک دنبالهٔ ستورم Sturm منجر می‌شود که برای محاسبه ساده است . سپس، تقریبی از ریشه‌ها را می‌توان به سادگی یافت و بعد ، مثلا" ، با روش دو بخشی بهتر نمود .

دنبالهٔ ستورم بوسیلهٔ یک دنبالهٔ بازگشتی به‌شرح زیر تولید می‌شود . فرض کنید $f_r(\lambda)$ مقدار دترمینان

$$\begin{vmatrix} a_1-\lambda & b_2 & & & \\ b_2 & a_2-\lambda & b_3 & & 0 \\ & b_3 & a_3-\lambda & b_4 & \\ 0 & & \cdots\cdots & \cdots\cdots & b_r \\ & & & b_r & a_r-\lambda \end{vmatrix} \tag{9.43}$$

باشد که چند جمله‌ای مشخصهٔ یک ماتریس سه قطری است . بسط این دترمینان برحسب ستون آخر می‌دهد

$$f_r(\lambda) = (a_r-\lambda)f_{r-1}(\lambda) - b_r^2 f_{r-2}(\lambda) \tag{9.44}$$

که برای $r = n, n-1, \ldots 2$ برقرار است . اگر تعریف کنیم

$$\begin{aligned} f_0(\lambda) &= 1 \\ f_1(\lambda) &= a_1 - \lambda \end{aligned} \tag{9.45}$$

آنگاه معادلات (9.44) به ازای $r = 2, 3, \ldots, n$ و (9.45) یک دنباله از توابع تولید می‌کندکه‌یک دنبالهٔ ستورم است . تعداد تغییرعلامتهای این دنباله‌رامی‌توان برای‌مقادیر

گوناگون λ حساب کرد و محل تقریبی ریشه‌ها را ، همانگونه‌که در فصل ۳ شرح داده‌شد ، تعیین نمود . نتایج عددی برای این روش در مثال ۹.۶ داده شده‌است .

۹.۶ ■ روشهای دیگر

دو روش وجود دارد که می‌توان برای یافتن تمام مقادیر ویژهٔ یک ماتریس حقیقی یا مختلط به‌کار برد . این روشها روشهای موجود به هنگام یافتن تمام مقادیر ویژه یک ماتریس هستند . چون برنامه‌نویسیهای نسبتا "مفصلی برای این روشها لازم است . توصیه نمی‌شود که استفاده کننده از این روشها برای اجرای آنها اقدام کند . برای این روشها روتینهای کامل توسط ویلکنسن و دیگران تولید ، آزمایش و چاپ شده است و هرجا ممکن باشد باید از این برنامه‌های استاندارد استفاده کرد .

ساده‌ترین‌این‌روشهاالگوریتم L–R است‌که‌اولین‌بار توسط روتیشوزر Rutishauser تشریح شد . نام L–R از روند محاسباتی که شامل تجزیهٔ تکراری‌یک دنباله‌ازماتریسها ، به فرم مثلثی چپ و مثلثی‌راست،گرفته شده است . برای مطابقت با علامتگذاری قبلی برای ماتریسهای پایین و بالا مثلثی به‌ترتیب حروف L و U به‌کار خواهند رفت .

ما دنباله‌ای از ماتریسها ، بوسیلهٔ تجزیهٔ مثلثی هر عضو دنباله ،تشکیل‌می‌دهیم . به‌منظور بدست آوردن روش فرض می‌شود که تمام ماتریسها به‌گونه‌ای هستند که تجزیه مثلثی ممکن است .

قرار دهید

$$\mathbf{A}=\mathbf{A}_1=\mathbf{L}_1\mathbf{U}_1 \tag{9.46}$$

و تشکیل دهید .

$$\begin{aligned}\mathbf{A}_2&=\mathbf{U}_1\mathbf{L}_1=\mathbf{L}_2\mathbf{U}_2\\ \mathbf{A}_r&=\mathbf{U}_{r-1}\mathbf{L}_{r-1}=\mathbf{L}_r\mathbf{U}_r, \qquad r=1,2,\ldots\end{aligned} \tag{9.47}$$

می‌توان ملاحظه کرد که این ماتریسها با $\mathbf{A}_1$ متشابه هستند و لذا مقادیر ویژهٔ یکسان دارند ، زیرا :

$$\begin{aligned}\mathbf{A}_2&=(\mathbf{U}_1)\mathbf{L}_1=\mathbf{L}_1^{-1}\mathbf{A}_1\mathbf{L}_1\\ \mathbf{A}_3&=\mathbf{L}_2^{-1}\mathbf{A}_2\mathbf{L}_2=\mathbf{L}_2^{-1}\mathbf{L}_1^{-1}\mathbf{A}_1\mathbf{L}_1\mathbf{L}_2\end{aligned} \tag{9.48}$$

این دنباله از ماتریسها غالبا "به یک ماتریس به فرم بالا مثلثی‌بلوکی ، که در آن هر بلوک با مقادیر ویژهٔ یکسان ازنظر قدر مطلق متناظر است ، همگرا می‌باشد . در حالت یک ماتریس حقیقی با مقادیر ویژهٔ متمایز ، این مقادیر ویژه ، وقتی همگرائی داشته باشیم ، به‌ترتیب‌نزولی از چپ به راست روی قطر اصلی ظاهر می‌شوند . لذا ، در صورتی‌که شرایط برای تجزیهٔ مثلثی برقرار باشد ، روش فوق روند تکراری ساده‌ای که

برای استفاده کامپیوتر مناسب است ارائه می‌کند. چند نتیجهٔ عددی در مثال ۹.۰۷ داده شده است. هنگامیکه مقادیر ویژه بدست آمدند بردارهای ویژه از ماتریس اولیه بدست می‌آیند. بحث شرایط همگرائی این‌روند از حیطهٔ عمل این کتاب خارج است. تفصیلات در سطحی مناسب آن افرادی‌که زمینهٔ ریاضی دارند در رالستون و پارلت Parlett (1967) آمده است. ارائه دقیقتر آن در ویلکنسن آمده است.

روش دوم تبدیلات متعامد را، به‌خاطر خواص پایداری بسیار خوبی که دارد، در روند وارد می‌کند. ماتریسها به‌حاصلضرب Q_rU_r که در آن U_r بالامثلثی است تجزیه می‌شوند. پس،

$$A = A_1 = Q_1U_1 \tag{9.49}$$

$$A_r = U_{r-1}Q_{r-1} = Q_rU_r, \qquad r = 2, 3, \ldots \tag{9.50}$$

متذکر می‌شویم که مانند قبل این ماتریسها متشابهند زیرا:

$$A_r = U_{r-1}Q_{r-1} = Q_{r-1}^{-1}A_{r-1}Q_{r-1} \tag{9.51}$$

این روش پیچیده‌تر است و وقت‌گیرتر از روش L–R می‌باشد. اما، دارای منفعت‌بیشتر پایداری است. همانند روش L–R ماتریسها به فرم بالا مثلثی همگرایند، بامقادیرویژهٔ واقع بر قطر اصلی. تجزیهٔ اساسی $A_r = Q_rU_r$ را می‌توان برای هر ماتریسی بدست آورد، که برای حالت تجزیه L–U چنین نیست. وقتی تمام مقادیر ویژه یک ماتریس لازم است، اگر روش Q–R از منبع قابل اعتمادی در یک کامپیوتر در دسترس‌است این روش باید استفاده شود.

بیشتر بحث قبلی بر ماتریسهای متقارن متمرکز شده بود که به فرمهای ساده‌تری برای حل و طرحهای محاسباتی پایدارتر منجر می‌شوند. برای ماتریسهای کلی‌تر ممکن است بدوضعی کاملا" جدی رخ دهد. روش پیشنهادی برای یک ماتریس‌کلی تبدیل به فرم هسنبرگ (ر.ک. معادلات (9.39)) بوسیلهٔ تبدیلات هاوس هلدر و متعاقبا" تبدیل به فرم مثلثی بلوکی بوسیلهٔ الگوریتم Q–R است.

مثالهای حل شده

۱ــ دو تا از بزرگترین مقادیر ویژه و بردارهای ویژهٔ متناظر آنها را از ماتریس زیر، به روش توانی پیدا کنید.

$$\begin{pmatrix} 2{\cdot}05 & 1{\cdot}30 & 4{\cdot}00 \\ 1{\cdot}30 & 2{\cdot}15 & 3{\cdot}70 \\ 4{\cdot}00 & 3{\cdot}70 & 8{\cdot}40 \end{pmatrix}$$

با فرض اینکه λ_1 بزرگترین مقدار ویژه باشد تقلیل $A_1 = A - \lambda_1 X^{(1)}X^{(1)T}$ را به‌کار برید، بردارآغازی [1, 1, 1] است.

محاسبات برای مقدار ویژه

اول عبارتست از :

x_1		x_2		x_3		λ	
4·565217*E*	−1	4·440994*E*	−1	1·000000*E*	0	1·61000000*E*	1
4·644941*E*	−1	4·421754*E*	−1	1·000000*E*	0	1·18692547*E*	1
4·646905*E*	−1	4·417781*E*	−1	1·000000*E*	0	1·18940252*E*	1
4·647077*E*	−1	4·417531*E*	−1	1·000000*E*	0	1·18933409*E*	1
4·647089*E*	−1	4·417514*E*	−1	1·000000*E*	0	1·18933173*E*	1
4·647089*E*	−1	4·417512*E*	−1	1·000000*E*	0	1·18933155*E*	1
4·647089*E*	−1	4·417512*E*	−1	1·000000*E*	0	1·18933153*E*	1
4·647089*E*	−1	4·417512*E*	−1	1·000000*E*	0	1·18933153*E*	1
1·189332*E*	1						

پس $\lambda_1 = 11{\cdot}89332$ and $\mathbf{X}_1^T = [0{\cdot}4647089, 0{\cdot}4417512, 1{\cdot}0000000]$

روند تقلیل ، ماتریس جدید $\mathbf{A}_1 = \mathbf{A} - \lambda_1 \mathbf{X}^{(1)} \mathbf{X}^{(1)T}$ را می‌دهد .

2·298481*E*	−1	−4·302322*E*	−1	8·324315*E*	−2
−4·302322*E*	−1	5·052453*E*	−1	−2·325999*E*	−2
8·324315*E*	−2	−2·325999*E*	−2	−2·840871*E*	−2

با بردار آغازی $[1,1,1]$ جدول تکرارها عبارتست از :

x_1		x_2		x_3		λ	
−1·000000*E*	0	4·418020*E*	−1	2·695424*E*	−1	1·17140964*E*	−1
−6·141836*E*	−1	1·000000*E*	0	−1·563346*E*	−1	6·47181000*E*	−1
−7·559143*E*	−1	1·000000*E*	0	−9·047109*E*	−2	7·73123187*E*	−1
−7·344848*E*	−1	1·000000*E*	0	−1·004296*E*	−1	8·32568328*E*	−1
−7·375264*E*	−1	1·000000*E*	0	−9·901613*E*	−2	8·23580271*E*	−1
−7·370906*E*	−1	1·000000*E*	0	−9·921863*E*	−2	8·24855999*E*	−1
−7·371530*E*	−1	1·000000*E*	0	−9·918966*E*	−2	8·23673234*E*	−1
−7·371441*E*	−1	1·000000*E*	0	−9·919381*E*	−2	8·24699383*E*	−1
−7·371453*E*	−1	1·000000*E*	0	−9·919321*E*	−2	8·24695641*E*	−1
−7·371452*E*	−1	1·000000*E*	0	−9·919330*E*	−2	8·24696176*E*	−1
−7·371452*E*	−1	1·000000*E*	0	−9·919329*E*	−2	8·24696100*E*	−1
8·246961*E*	−1						

بنابراین $\lambda_2 = 0{\cdot}8246961$ and $[\mathbf{X}^{(2)}]^T = [-0{\cdot}7371452, 1{\cdot}000000, -0{\cdot}09919329]$.

۲ــ نزدیکترین مقدار ویژه به $5{\cdot}0$ و بردار ویژهٔ متناظر آن را برای ماتریس زیر به روش توانی پیداکنید .

$$\begin{pmatrix} 3 & 6 & 8 \\ 4 & 11 & 16 \\ 6 & 15 & 40 \end{pmatrix}$$

ماتریس $\mathbf{A} - 5\mathbf{I}$ تشکیل شده و تکرار معکوس به‌کار می‌رود . ماتریس جدید عبارتست از :

$$\begin{matrix} -2 & 6 & 8 \\ 4 & 6 & 16 \\ 6 & 15 & 35 \end{matrix}$$

جوابهای حذف گاوسی عبارتنداز :

λ	x_3		x_2		x_1	
15·465986	−5·510204E	−1	1·000000E	0	7·346939E	−1
15·542812	−5·575104E	−1	1·000000E	0	7·462063E	−1
15·542087	−5·574888E	−1	1·000000E	0	7·460399E	−1
15·542098	−5·574893E	−1	1·000000E	0	7·460424E	−1

۳ــ جدول اعداد زیر نتایج محاسبات‌را برای پیدا کردن بزرگترین مقدار ویژه از نظر قدرمطلق یک ماتریس، بوسیلهٔ تکرار، و حالتی که در آن دو ریشه مختلط مزدوج با قدر مطلق یکسان وجود دارد، نشان می‌دهد. تنها نوزده تکرار اول جدولبندی شده‌اند. مقادیر از تکرارهای 20 و 21 و 22 برای تشکیل دو معادله جهت تعیین ضرائب a و b به‌کار رفته‌اند. (ر.ک معادلات (9.16)) وقتی این معادلات حل شوند مقادیر a و b در معادلهٔ درجهٔ دوم درج شده که برای بدست آوردن مقادیر ویژهٔ مطلوب حل می‌شود. هر سه بردار ویژه متوالی بهمان‌روش، به‌کار رفته و می‌توان مشاهده کرد که مقادیر ویژه همگرا هستند. در هر مرحله مقادیر مختلفی برای بردارهای ویژه وجود دارد زیرا بردار ویژه را می‌توان در اسکالر دلخواهی ضرب کرد. بردارهای ویژه عبارتند از:

$$[-i, -1/\sqrt{2}, -1/\sqrt{2}, 0] \quad \text{و} \quad [i, -1/\sqrt{2}, -1/\sqrt{2}, 0]$$

هر یک از بردارهای ویژه در دو ستون‌جدولبندی شده‌اند، اول قسمتهای حقیقی و دوم قسمتهای موهومی بردارها هستند.

محاسبات با کامپیوتر ICL 1904A و عملیات 2·60 ثانیه طول کشیده است.

A

3·000000000 0	−1·414213562 4	1·414213562 4	0·000000000 0
1·414213562 4	2·250000000 0	−0·750000000 0	−0·353553390 6
−1·414213562 4	−0·750000000 0	2·250000000 0	−0·353553390 6
0·000000000 0	−0·353553390 6	−0·353553390 6	1·060660171 8

$\mathbf{Y}^{(0)}$

1·000000000 0
1·000000000 0
1·000000000 0
1·000000000 0

$\mathbf{Y}^{(1)}$

1·000000000 0
0·853553390 6
−0·089255651 0
0·117851130 2

$\mathbf{Y}^{(2)}$

0·496034143 1
1·000000000 0
−0·683595710 2
−0·043220464 0

$Y^{(3)}$
−0·256 608 042 4
1·000 000 000 0
−0·854 814 933 0
−0·045 325 228 9

$Y^{(4)}$
−1·000 000 000 0
0·749 864 746 1
−0·676 232 963 8
−0·029 297 798 6

$Y^{(20)}$	$Y^{(21)}$	$Y^{(22)}$
1·000 000 000 0	5·085 482 313 8	17·512 893 882 6
−0·737 328 667 9	−0·797 773 233 7	4·798 636 911 2
0·737 330 018 2	0·797 775 699 9	−4·798 632 406 8
−0·000 000 623 4	−0·000 001 138 6	−0·000 002 079 6

A,B = −9·6114, 16·9074

مقادیر ویژهٔ مزدوج = 7·293 075 +i* 0·000 000, 2·318 283 +i* 0·000 000

بردار ویژه		بردار ویژه	
−0·556 244 160 4	0·000 000 000 0	0·443 755 839 6	0·000 000 000 0
−0·183 236 503 8	0·000 000 000 0	−0·920 565 171 7	0·000 000 000 0
0·183 236 637 2	0·000 000 000 0	0·920 566 655 4	0·000 000 000 0
−0·000 000 061 6	0·000 000 000 0	−0·000 000 685 0	0·000 000 000 0

$Y^{(23)}$	$Y^{(24)}$	$Y^{(25)}$
0·994 975 167 3	0·156 498 674 4	−11·995 685 128 9
1·000 000 000 0	4·407 107 252 6	13·442 644 080 4
−0·999 999 789 9	−4·407 106 868 9	−13·442 643 379 6
−0·000 000 097 0	−0·000 000 177 1	−0·000 000 323 6

A,B = − 6·0000, 13·0000

مقادیر ویژهٔ مزدوج = 3·000 000 +i* 2·000 000, 3·000 000 +i* −2·000 00

بردار ویژه		بردار ویژه	
−0·497 487 583 7	−0·707 106 756 8	0·497 487 583 7	−0·707 106 756 8
−0·500 000 000 0	0·351 776 813 2	0·500 000 000 0	0·351 776 813 2
0·499 999 895 0	-0·351 776 874 8	−0·499 999 895 0	−0·351 776 874 8
0·000 000 048 5	0·000 000 028 5	−0·000 000 048 5	0·000 000 028 5

$Y^{(26)}$	$Y^{(27)}$	$Y^{(28)}$
−1·000 000 000 0	−3·892 894 605 7	−10·357 367 634 2
0·315 685 923 9	−0·467 155 800 7	−6·906 851 769 0
−0·315 685 906 6	0·467 155 832 3	6·906 851 826 8
−0·000 000 008 0	−0·000 000 014 6	−0·000 000 026 6

A,B = −6·0000, 13·0000

مقادیر ویژهٔ مزدوج = 3·000 000 +i* 2·000 000, 3·000 000 +i* −2·000 000

بردار ویژه		بردار ویژه	
0·500 000 000 0	−0·223 223 655 5	−0·500 000 000 0	−0·223 223 655 5
−0·157 842 962 0	−0·353 553 393 1	0·157 842 962 0	−0·353 553 393 1
0·157 842 953 3	0·353 553 388 1	−0·157 842 953 3	0·353 553 388 1
0·000 000 004 0	0·000 000 002 3	−0·000 000 004 0	0·000 000 002 3

$Y^{(29)}$	$Y^{(30)}$	$Y^{(31)}$
−0·326 186 047 6	1·849 868 977 7	15·339 632 485 1
−0·999 999 997 0	−3·461 296 725 1	−7·767 780 381 8
1·000 000 000 0	3·461 296 730 7	7·767 780 391 9
−0·000 000 001 4	−0·000 000 002 5	−0·000 000 004 6

A,B = −6·0000, 13·0000

مقادیر ویژهٔ مزدوج = 3·000000 +i* 2·000000, 3·000000 +i* −2·000000

بردار ویژه		بردار ویژه	
0·163 093 023 8	0·707 106 779 4	−0·163 093 023 8	0·707 106 779 4
0·499 999 998 5	−0·115 324 183 5	−0·499 999 998 5	−0·115 324 183 5
−0·500 000 000 0	0·115 324 182 7	0·500 000 000 0	0·115 324 182 7
0·000 000 000 7	0·000 000 000 4	−0·000 000 000 7	0·000 000 000 4

$Y^{(32)}$	$Y^{(33)}$	$Y^{(34)}$
1·000 000 000 0	3·066 970 229 5	5·401 821 376 8
−0·023 677 551 6	1·343 180 907 6	8·366 893 616 3
0·023 677 551 8	−1·343 180 907 1	−8·366 893 615 3
−0·000 000 000 1	−0·000 000 000 2	−0·000 000 000 4

A, B = −6·0000, 13·0000

مقادیر ویژهٔ مزدوج = 3·000000 +i* 2·000000, 3·000000 +i* −2·000000

بردار ویژه		بردار ویژه	
−0·500 000 000 0	0·016 742 557 3	0·500 000 000 0	0·016 742 557 3
0·011 838 775 8	0·353 553 390 5	−0·011 838 775 8	0·353 553 390 5
−0·011 838 775 9	−0·353 553 390 6	0·011 838 775 9	−0·353 553 390 6
0·000 000 000 1	0·000 000 000 0	−0·000 000 000 1	0·000 000 000 0

۴ــ روش ژاکوبی برای تبدیل عناصر غیر قطری به‌صفر، از تبدیلات متشابه استفاده می‌کند. در حالت یک ماتریس متقارن عناصر متقارن نیز به صفر تبدیل می‌شوند. عناصر غیر قطری بطور سیستماتیک حذف شده تا اینکه عناصر غیر قطری کاملا" صفر شوند. دراین‌صورت، عناصر قطری مقادیر تقریبی مقادیر ویژه را بدست می‌دهند. موثرترین روش برای حذف عناصر غیر قطری، درنظر گرفتن یک مقدار آغازی و مرور سیستماتیک سطرها و حذف تمام عناصری که قدرمطلقی بزرگتر از مقدار آغازی دارند می‌باشد. سپس مقدار آغازی کم‌کم تقلیل یافته تا اینکه به سطح قابل قبول نهائی برسد. در نتایج کامپیوتری ذیل، ماتریس A معرف ماتریس اولیه و تبدیل یافته‌های متوالی آن است، ماتریس C معرف ماتریسی است که در تبدیل متشابه $C_r^T A_{r-1} C_r$ به‌کار می‌رود، و بردارهای ویژهٔ نهایی ستونهای ماتریس CE که از $C_1C_2\cdots C_r$ تشکیل می‌شود، هستند. نتیجهٔ اخیر به این علت حاصل می‌شود که‌بردارهای ویژه ماتریس قطری نهائی ستونهای ماتریس واحد است، و بااستفاده از نظریه‌ای که قبلا" ارائه‌شد، بردارهای ویژهٔ ماتریس اولیه ستونهای $C_1C_2\cdots C_rI$ هستند. درمحاسبات ارائه شده 16 مرحله موجود است و مقدار آغازی تا 10^{-5} تقلیل یافته است. برای تقلیل حجم اعداد ارائه شده تنها چند تبدیل اولی و آخری ارائه‌شده‌اند. محاسبات

با کامپیوتر ICL 1904A و عملیات 2·66 ثانیه طول کشیده‌است .

A

10·000 000 000 0	7·000 000 000 1	4·000 000 000 0	1·000 000 000 0
7·000 000 000 1	11·000 000 000 0	1·000 000 000 0	2·000 000 000 0
4·000 000 000 0	1·000 000 000 0	5·000 000 000 0	3·000 000 000 0
1·000 000 000 0	2·000 000 000 0	3·000 000 000 0	4·000 000 000 0

THRES = 0·10E 01

1 THETA = −0·749 744 $I=1$ $J=2$

C

0·731 863 050 7	0·681 451 740 8	0·000 000 000 0	0·000 000 000 0
−0·681 451 740 8	0·731 863 050 7	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

3·482 165 576 3	0·000 000 000 0	2·246 000 462 1	−0·631 040 430 9
−0·000 000 000 1	17·517 834 424 5	3·457 670 013 9	2·145 177 842 2
2·246 000 462 1	3·457 670 013 9	5·000 000 000 0	3·000 000 000 0
−0·631 040 430 9	2·145 177 842 2	3·000 000 000 0	4·000 000 000 0

CE

0·731 863 050 7	0·681 451 740 8	0·000 000 000 0	0·000 000 000 0
−0·681 451 740 8	0·731 863 050 7	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

2 THETA = − 0·622 472 $I=1$ $J=3$

C

0·812 439 666 0	0·000 000 000 0	0·583 045 271 9	0·000 000 000 0
0·000 000 000 0	1·000 000 000 0	0·000 000 000 0	0·000 000 000 0
−0·583 045 271 9	0·000 000 000 0	0·812 439 666 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

1·870 329 024 1	−2·015 978 153 5	−0·000 000 000 0	−2·261 818 092 7
−2·015 978 153 6	17·517 834 424 5	2·809 148 271 2	2·145 177 842 2
−0·000 000 000 0	2·809 148 271 3	6·611 836 552 6	2·069 393 858 4
−2·261 818 092 7	2·145 177 842 2	2·069 393 858 4	4·000 000 000 0

CE

0·594 594 572 5	0·681 451 740 8	0·426 709 291 4	0·000 000 000 0
−0·553 638 424 7	0·731 863 050 7	−0·397 317 215 5	0·000 000 000 0
−0·583 045 271 9	0·000 000 000 0	0·812 439 666 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

3 THETA = 0·565 395 $I=1$ $J=4$

C

0·844 376 860 3	0·000 000 000 0	0·000 000 000 0	−0·535 749 678 3
0·000 000 000 0	1·000 000 000 0	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·535 749 678 3	0·000 000 000 0	0·000 000 000 0	0·844 376 860 3

A

0·435 225 372 6	−0·552 966 964 9	1·108 677 093 8	0·000 000 000 1
−0·552 966 964 9	17·517 834 424 5	2·809 148 271 2	2·891 398 178 4
1·108 677 093 9	2·809 148 271 3	6·611 836 552 6	1·747 348 288 9
0·000 000 000 0	2·891 398 178 4	1·747 348 288 9	5·435 103 651 5

CE

0·502 061 898 3	0·681 451 740 8	0·426 709 291 4	−0·318 553 850 9
−0·467 479 474 8	0·731 863 050 7	−0·397 317 215 5	0·296 611 607 9
−0·492 309 936 1	0·000 000 000 0	0·812 439 666 0	0·312 366 316 9
0·535 749 678 3	0·000 000 000 0	0·000 000 000 0	0·844 376 860 3

4 THETA = 0·237 850 $I=2$ $J=3$

C

1·000 000 000 0	0·000 000 000 0	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·971 846 882 0	−0·235 613 322 9	0·000 000 000 0
0·000 000 000 0	0·235 613 322 9	0·971 846 882 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

0·435 225 372 6	−0·276 180 126 6	1·207 750 760 8	0·000 000 000 1
−0·276 180 126 6	18·198 880 761 1	0·000 000 000 0	3·221 694 840 9
1·207 750 760 9	−0·000 000 000 1	5·930 790 216 0	1·016 903 053 7
0·000 000 000 0	3·221 694 840 9	1·016 903 053 7	5·435 103 651 5

CE

0·502 061 898 3	0·762 805 143 6	0·254 136 985 4	−0·318 553 850 9
−0·467 479 474 8	0·617 645 594 5	−0·558 568 182 4	0·296 611 607 9
−0·492 309 936 1	0·191 421 609 4	0·789 566 956 2	0·312 366 316 9
0·535 749 678 3	0·000 000 000 0	0·000 000 000 0	0·844 376 860 3

5 THETA = 0·233 747 $I=2$ $J=4$

C

1·000 000 000 0	0·000 000 000 0	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·972 805 225 4	0·000 000 000 0	−0·231 624 682 2
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·231 624 682 2	0·000 000 000 0	0·972 805 225 4

A

0·435 225 372 6	−0·268 669 470 3	1·207 750 760 8	0·063 970 134 1
−0·268 669 470 3	18·965 965 501 3	0·235 539 846 7	−0·000 000 000 0
1·207 750 760 9	0·235 539 846 6	5·930 790 216 0	0·989 248 604 4
0·063 970 134 1	−0·000 000 000 1	0·989 248 604 4	4·668 018 911 2

CE

0·502 061 898 3	0·668 275 895 2	0·254 136 985 4	−0·486 575 349 7
−0·467 479 474 8	0·669 551 431 2	−0·558 568 182 4	0·145 483 357 5
−0·492 309 936 1	0·258 567 690 7	0·789 566 956 2	0·259 533 615 9
0·535 749 678 3	0·195 578 521 9	0·000 000 000 0	0·821 414 221 9

6 THETA = 0·207 059 $I=3$ $J=1$

C

0·978 639 711 7	0·000 000 000 0	0·205 582 865 7	0·000 000 000 0
0·000 000 000 0	1·000 000 000 0	0·000 000 000 0	0·000 000 000 0
−0·205 582 865 7	0·000 000 000 0	0·978 639 711 7	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

0·181 513 143 7	−0·311 353 569 6	−0·000 000 000 1	−0·140 768 849 4
−0·311 353 569 7	18·965 965 501 3	0·175 274 808 0	−0·000 000 000 0
−0·000 000 000 0	0·175 274 807 9	6·184 502 444 8	0·981 269 132 5
−0·140 768 849 4	−0·000 000 000 1	0·981 269 132 5	4·668 018 911 2

CE

0·439 091 501 6	0·668 275 895 2	0·351 923 869 9	−0·486 575 349 7
−0·342 661 930 8	0·669 551 431 2	−0·642 742 775 1	0·145 483 357 5
−0·644 115 491 5	0·258 567 690 7	0·671 491 090 9	0·259 533 615 9
0·524 305 910 7	0·195 578 521 9	0·110 140 954 2	0·821 414 221 9

THRES = 0·10E 00
7 THETA = 0·016 569 $I=1$ $J=2$

C

0·999 862 737 3	−0·016 568 242 0	0·000 000 000 0	0·000 000 000 0
0·016 568 242 0	0·999 862 737 3	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

0·176 353 854 3	0·000 000 000 0	0·002 903 995 4	−0·140 749 527 1
−0·000 000 000 1	18·971 124 790 9	0·175 250 749 3	0·002 332 292 3
0·002 903 995 4	0·175 250 749 2	6·184 502 444 8	0·981 269 132 5
−0·140 749 527 1	0·002 332 292 3	0·981 269 132 5	4·668 018 911 2

CE

0·450 103 387 5	0·660 909 191 6	0·351 923 869 9	−0·486 575 349 7
−0·331 521 605 9	0·675 136 832 5	−0·642 742 775 1	0·145 483 357 5
−0·639 743 066 4	0·269 204 060 3	0·671 491 090 9	0·259 533 615 9
0·527 474 335 3	0·186 864 849 1	0·110 140 954 2	0·821 414 221 9

THRES = 0·10E −03
15 THETA = −0·000 029 $I=1$ $J=2$

C

0·999 999 999 6	0·000 029 430 4	0·000 000 000 0	0·000 000 000 0
−0·000 029 430 4	0·999 999 999 6	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	1·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·000 000 000 0	1·000 000 000 0

A

0·171 752 891 4	0·000 000 000 0	−0·000 000 567 6	−0·000 000 529 5
−0·000 000 000 1	18·973 543 783 6	−0·000 000 000 0	−0·000 000 042 8
−0·000 000 567 6	−0·000 000 000 1	6·664 858 254 3	−0·000 075 562 4
−0·000 000 529 5	−0·000 000 042 8	−0·000 075 562 4	4·189 845 072 8

CE

0·432 011 697 6	0·665 155 644 9	0·088 255 154 5	−0·602 615 042 1
−0·322 845 225 7	0·666 374 137 9	−0·518 019 542 0	0·428 219 829 9
−0·634 823 972 1	0·278 727 841 8	0·719 401 054 4	−0·042 088 450 4
0·553 295 018 9	0·189 272 994 8	0·454 234 416 0	0·672 094 822 8

THRES = 0·10E −04
16 THETA = −0·000 031 $I=3$ $J=4$

C

1·000 000 000 0	0·000 000 000 0	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	1·000 000 000 0	0·000 000 000 0	0·000 000 000 0
0·000 000 000 0	0·000 000 000 0	0·999 999 999 5	0·000 030 530 1
0·000 000 000 0	0·000 000 000 0	−0·000 030 530 1	0·999 999 999 5

A

0·171 752 891 4	0·000 000 000 0	−0·000 000 567 6	−0·000 000 529 5
−0·000 000 000 1	18·973 543 783 6	−0·000 000 000 0	−0·000 000 042 8
−0·000 000 567 5	−0·000 000 000 1	6·664 858 256 7	0·000 000 000 0
−0·000 000 529 5	−0·000 000 042 8	0·000 000 000 0	4·189 845 070 4

CE

0·432 011 697 6	0·665 155 644 9	0·088 273 552 3	−0·602 612 347 3
−0·322 845 225 7	0·666 374 137 9	−0·518 032 615 4	0·428 204 014 5
−0·634 823 972 1	0·278 727 841 8	0·719 402 339 1	−0·042 066 487 0
0·553 295 018 9	0·189 272 994 8	0·454 213 896 7	0·672 108 690 3

THRES = 0·10E −05

۵ ــ ارقام زیر نتایج محاسبه‌ای برای تبدیل ماتریس $\mathbf{A}_0$ به یک ماتریس سه قطری بوسیلهٔ روش گیونز است . تبدیلات به شکل $\mathbf{A}_r = \mathbf{P}_r^T \mathbf{A}_{r-1} \mathbf{P}_r$ هستند که در آن ماتریس $\mathbf{P}_r$ یک ماتریس دوران ، همانگونه که در بخش ۹.۴.۲ بحث شده است ، می‌باشد مثلاً " ، اولین عنصری که باید صفر شود عنصر $1,3$ است که با استفاده از ماتریس دورانی که در صفحهٔ $(2,3)$ دوران می‌دهد حاصل می‌شود ، زاویهٔ θ از فرمول زیرپیدامی‌شود .

$$\tan\theta = \frac{a_{13}}{a_{12}} = 3$$

بنابراین

$$\sin\theta = \frac{3}{\sqrt{10}} = 0{\cdot}948\,683, \qquad \cos\theta = \frac{1}{\sqrt{10}} = 0{\cdot}316\,228$$

مقادیر ویژهٔ ماتریس سه‌قطری حاصل بوسیلهٔ روش‌دنبالهٔ ستورم (ر.ک. مثال ۹.۶) بدست خواهند آمد .

$$\mathbf{A}_0 = \begin{bmatrix} 2{\cdot}000\,000 & 1{\cdot}000\,000 & 3{\cdot}000\,000 & 2{\cdot}000\,000 \\ 1{\cdot}000\,000 & 4{\cdot}000\,000 & 2{\cdot}000\,000 & 1{\cdot}000\,000 \\ 3{\cdot}000\,000 & 2{\cdot}000\,000 & 3{\cdot}000\,000 & -3{\cdot}000\,000 \\ 2{\cdot}000\,000 & 1{\cdot}000\,000 & -3{\cdot}000\,000 & 1{\cdot}000\,000 \end{bmatrix}$$

$$\mathbf{P}_1 = \begin{bmatrix} 1{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}316\,228 & -0{\cdot}948\,683 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}948\,683 & 0{\cdot}316\,228 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 & 1{\cdot}000\,000 \end{bmatrix}$$

$$\mathbf{A}_1 = \begin{bmatrix} 2{\cdot}000\,000 & 3{\cdot}162\,278 & -0{\cdot}000\,000 & 2{\cdot}000\,000 \\ 3{\cdot}162\,278 & 4{\cdot}300\,000 & -1{\cdot}900\,000 & -2{\cdot}529\,822 \\ -0{\cdot}000\,000 & -1{\cdot}900\,000 & 2{\cdot}700\,000 & -1{\cdot}897\,367 \\ 2{\cdot}000\,000 & -2{\cdot}529\,822 & -1{\cdot}897\,367 & 1{\cdot}000\,000 \end{bmatrix}$$

$$P_2=\begin{bmatrix} 1{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}845\,154 & 0{\cdot}000\,000 & -0{\cdot}534\,522 \\ 0{\cdot}000\,000 & 0{\cdot}000\,000 & 1{\cdot}000\,000 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}534\,522 & 0{\cdot}000\,000 & 0{\cdot}845\,154 \end{bmatrix}$$

$$A_2=\begin{bmatrix} 2{\cdot}000\,000 & 3{\cdot}741\,657 & -0{\cdot}000\,000 & -0{\cdot}000\,000 \\ 3{\cdot}741\,657 & 1{\cdot}071\,429 & -2{\cdot}619\,978 & -2{\cdot}574\,998 \\ -0{\cdot}000\,000 & -2{\cdot}619\,978 & 2{\cdot}700\,000 & -0{\cdot}587\,975 \\ -0{\cdot}000\,000 & -2{\cdot}574\,998 & -0{\cdot}587\,975 & 4{\cdot}228\,571 \end{bmatrix}$$

$$P_3=\begin{bmatrix} 1{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 1{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}000\,000 \\ 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}713\,203 & -0{\cdot}700\,958 \\ 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}700\,958 & 0{\cdot}713\,203 \end{bmatrix}$$

$$A_3=\begin{bmatrix} 2{\cdot}000\,000 & 3{\cdot}741\,657 & -0{\cdot}000\,000 & 0{\cdot}000\,000 \\ 3{\cdot}741\,657 & 1{\cdot}071\,429 & -3{\cdot}673\,540 & 0{\cdot}000\,000 \\ -0{\cdot}000\,000 & -3{\cdot}673\,540 & 2{\cdot}863\,165 & 0{\cdot}753\,990 \\ 0{\cdot}000\,000 & 0{\cdot}000\,000 & 0{\cdot}753\,990 & 4{\cdot}065\,406 \end{bmatrix}$$

۶ــ جدول زیر محاسبات مربوط به دنبالهٔ ستورم را در یک سری نقطه، به‌هنگام پیدا کردن تقریبهایی از مقادیر ویژه نشان می‌دهد.

$$\begin{bmatrix} 1 & -2 & 0 & 0 \\ -2 & 1 & -2 & 0 \\ 0 & -2 & 1 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}$$

λ	-10	-2	0	2	10
f_0	1	1	1	1	1
f_1	11	3	1	-1	-9
f_2	117	5	-3	-3	77
f_3	1243	3	-7	7	-657
f_4	11962	-11	5	-2	6262
Number of sign changes	0	1	2	3	4

پس، ریشه‌ها، بین 10 و 2، 2 و 0، 0 و 2-، 2- و 10- قرار دارند.لذا، برای تمام ریشه‌ها محلهای تقریبی بدست آمده است.

۷ــ جدولی از محاسبات روش L-R ذیلا" داده شده است. جوابهای واقعی برای سه مقدار ویژه عبارتند از:

3·550510 و 6·000000 و 8·449490 . ماتریسهای واقع در ستونهای سمت چپ از مولفه‌های کامل ماتریسهای بالا مثلثی و پایین مثلثی ،بدون اعضای قطری آن که واحدند، تشکیل شده‌اند. ستونهای سمت راست ماتریسهای جدید را که بوسیلهٔ حاصلضرب U_rL_r تشکیل شده‌اند، می‌دهند. بنابراین، درحالت اول ماتریس به صورت زیر تجزیه می‌شود

$$\begin{bmatrix}8 & 1 & 0\\ 1 & 6 & 1\\ 0 & 1 & 4\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0{\cdot}125 & 1 & 0\\ 0 & 0{\cdot}170\,213 & 1\end{bmatrix}\begin{bmatrix}8 & 1 & 0\\ 0 & 5{\cdot}875 & 1\\ 0 & 0 & 3{\cdot}829\,787\end{bmatrix}$$

ماتریسی جدید بصورت‌زیر تشکیل می‌شود

$$\begin{bmatrix}8 & 1 & 0\\ 0 & 5{\cdot}875 & 1\\ 0 & 0 & 3{\cdot}829\,787\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0{\cdot}125 & 1 & 0\\ 0 & 0{\cdot}170\,213 & 1\end{bmatrix}$$
$$=\begin{bmatrix}8{\cdot}125 & 1 & 1\\ 0{\cdot}734\,375 & 6{\cdot}045\,213 & 1\\ 0 & 0{\cdot}651\,880 & 3{\cdot}829\,787\end{bmatrix}$$

می‌توان ملاحظه کرد که همگرائی کند است ولی عناصر ماتریس پایین مثلثی درحال کوچک شدن هستند و عناصر قطری به مقادیر واقعی مقادیر ویژه میل می‌کنند .

8·0	1·000 000	0	8·125	1·000 000	0
0·125	5·875	1·000 000	0·734 375	6·045 230	1·000 000
0	0·170 213	3·829 787	0	0·651 880	3·829 787
8·125	1·000 000	0	8·215 385	1·000 000	0
0·090 385	5·954 845	1·000 000	0·538 229	6·062 679	1·000 000
0	0·107 834	3·721 953	0	0·401 353	3·721 953
8·215 385	1·000 000	0	8·280 900	1·000 000	0
0·065 515	5·997 164	1·000 000	0·392 904	6·064 088	1·000 000
0	0·066 924	3·655 029	0	0·244 609	3·655 029
8·280 900	1·000 000	0	8·328 347	1·000 000	0
0·047 447	6·016 641	1·000 000	0·285 472	6·057 296	1·000 000
0	0·040 655	3·614 374	0	0·146 942	3·614 374
8·328 347	1·000 000	0	8·362 624	1·000 000	0
0·034 277	6·023 019	1·000 000	0·206 451	6·047 416	1·000 000
0	0·024 397	3·589 977	0	0·087 585	3·589 977
8·362 624	1·000 000	0	8·387 311	1·000 000	0
0·024 687	6·022 729	1·000 000	0·148 683	6·037 271	1·000 000
0	0·014 542	3·575 435	0	0·051 994	3·575 435
8·387 311	1·000 000	0	8·405 038	1·000 000	0
0·017 727	6 019 544	1 000 000	0·106 708	6·028 182	1·000 000
0	0·008 638	3·566 797	0	0·030 810	3·566 797
8·405 038	1·000 000	0	8·417 734	1·000 000	0
0·012 696	6·015 486	1·000 000	0·076 373	6·020 608	1·000 000
0	0·005 122	3·561 675	0	0·018 243	3·561 675
8·417 734	1·000 000	0	8·426 807	1·000 000	0
0·009 073	6·011 535	1·000 000	0·054 543	6·014 570	1·000 000
0	0·003 035	3·558 640	0	0·010 800	3·558 640
8·426 807	1·000 000	0	8·433 280	1·000 000	0
0·006 473	6·008 097	1·000 000	0·038 890	6·009 895	1·000 000
0	0·001 798	3·556 842	0	0·006 395	3·556 842

8·433 280	1·000 000	0	8·437 891	1·000 000	0
0·004 611	6·005 284	1·000 000	0·027 690	6·006 349	1·000 000
0	0·001 065	3·555 777	0	0·003 787	3·555 777
8·437 891	1·000 000	0	8·441 173	1·000 000	0
0·003 282	6·003 067	1·000 000	0·019 702	6·003 698	1·000 000
0	0·000 631	3·555 146	0	0·002 243	3·555 146
8·441 173	1·000 000	0	8·443 507	1·000 000	0
0·002 334	6·001 364	1·000 000	0·014 007	6·001 738	1·000 000
0	0·000 374	3·554 772	0	0·001 329	3·554 772
8·443 507	1·000 000	0	8·445 166	1·000 000	0
0·001 659	6·000 079	1·000 000	0·009 954	6·000 301	1·000 000
0	0·000 222	3·554 550	0	0·000 789	3·554 550

■ مسائل

۱ـ ماتریس تبدیل ژاکوبی‌را که صفری در محل (1,3) از تبدیل یافتهٔ ماتریس زیر تولید می‌کند پیدا کنید. تبدیل را کامل کرده به فرم قطری درآورید و تمام مقادیر ویژه و بردارهای ویژه را بدست آورید.

$$\begin{bmatrix} 7 & 0 & 6 \\ 0 & 5 & 0 \\ 6 & 0 & 2 \end{bmatrix}$$

۲ـ مقدار ویژه با بزرگترین قدر مطلق و بردار ویژه متناظر نرمال شدهٔ آن را از ماتریس زیر بدست آورید.

$$\begin{bmatrix} 3 & 6 & 8 \\ 4 & 11 & 16 \\ 6 & 15 & 40 \end{bmatrix}$$

۳ـ نزدیکترین مقدار ویژهٔ ماتریس زیر را به 4 بدست آورید.

$$\begin{bmatrix} 1 & 2 & 0 \\ 4 & 3 & 0 \\ 5 & 6 & 7 \end{bmatrix}$$

۴ـ مقدار ویژه با بزرگترین مقدار مطلق را از ماتریس زیر پیدا کنید، و بردار ویژهٔ متناظر را نیز بدست آورید.

$$\begin{bmatrix} 0 & \sqrt{2} & 1 \\ -\sqrt{2} & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

۵ ــ مقادیر و بردارهای ویژهٔ ماتریس زیر را بدست آورید.

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

۶ـ ماتریس زیر را با استفاده از تبدیل گیونز به سه قطری تبدیل کنید. روش دنبالهٔ ستورم را بهمراه روش دو بخشی جهت یافتن مقادیر ویژه ماتریس سه قطری حاصل به‌کار برید. همچنین بردارهای ویژه متناظر با تمام مقادیر ویژه را بدست آورید.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$$

۷ ــ اگر A و B دو ماتریس مربع و هم‌مرتبه باشند ثابت کنید مقادیر ویژهٔ AB و BA یکسان هستند.

۸ ــ اگر ماتریس $I-AB$ معکوس داشته باشد ثابت کنید $I-BA$ نیز معکوس دارد.

۹ ــ اگر X_1 ، X_2 ، . . . و X_k بردارهای ویژهٔ A نظیر مقدار ویژهٔ λ باشند

ثابت کنید هر ترکیب خطی از این بردارها نیز یک بردار ویژهٔ A نظیر λ است .

۱۰ ــ اگر λ مختلط و مقدار ویژهٔ A باشد ثابت کنید $\bar{\lambda}$ نیز مقدار ویژهٔ A است . ضمنا نشان دهید که بردارهای ویژهٔ نظیر $\bar{\lambda}$ مزدوج بردارهای ویژهٔ نظیر λ است .

۱۱ ــ اگر A حقیقی و λ مقدار ویژه‌ای مختلط از A باشد ثابت کنید تمام بردارهای ویژهٔ نظیر λ مختلط هستند . (یک بردار را مختلط گوئیم در صورتیکه حداقل یک مولفهٔ آن مختلط باشد .)

۱۲ ــ مقادیر ویژه و بردارهای ویژهٔ ماتریس‌های زیر را تعیین کنید .

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \qquad C = \begin{bmatrix} a & 1 \\ 0 & b \end{bmatrix}$$

در ماتریس C وقتی b به a میل کند بردارهای ویژه چگونه تغییر می‌کنند ؟

۱۳ ــ ماتریس حقیقی A را متقارن چپ ، یا اوریب ، نامند اگر $A = -A^t$. اگر A حقیقی و متقارن چپ باشد ثابت کنید مقادیر ویژهٔ A صفر یا موهومی محض‌اند .

واژه‌نامه فارسی به انگلیسی

الف

اصلاحگر	corrector
الگوریتم	algorithm
انتگرال	integral
ـ معین	definite —
ـ ناسره	improper —
ـ نامعین	indefinite —
انتگرال گیری	integration
اندازه	size, magnitud
اولیه	initial

ب

بازگشتی	recurrence
باقیمانده	remainder
بالا مثلثی	upper triangular
بد طرح	illposed
بد وضع	ill-conditioned
بد وضعی	ill-conditioning
بردار	vector
برونیابی	extrapolation
برهان	proof

پ

پایدار	stable
پایداری	stability
پایین مثلثی	lower triangular
پیوسته	continuous

ت

تابع	function
ـ متناوب	periodic —
ـ وزن	weight —
تبدیل	transformation
ـ خطی	linear —
تجزیه	factorization
ترانهاد	transpose
تعویض	interchange
ـ ستون	column —
ـ سطر	row —
تفاضلات	differences
ـ تقسیم شده	divided —
ـ متناهی	finite —
تقریب	approximation
ـ کمترین مربعات	least square —
ـ مینیماکس	minimax —
تکرار	iteration
تکراری	iterative
تناقض	contradiction
توان	power

ج

جایگذاری	substitution
ـ پسرو	backward —
ـ پیشرو	forward —
جواب	solution

واژه‌نامه فارسی به انگلیسی

چ

چند جمله‌ای	polynomial
ــ درونیاب	interpolating —
ــ تقریب	approximating —

ح

حذف	elimination
ــ گاوسی	Gaussian —

خ

خطا	error
ــ ی برش	truncated —
ــ ی جامع	global —
ــ ی ذاتی	inherent —
ــ ی مجتمع	accumulated —
ــ ی مطلق	absolute —
ــ ی نسبی	relative —
خطی	linear
خوش طرح	wellposed
خوش وضع	well-conditioned

د

داده‌ها	data
دترمینان	determinant
درونیابی	interpolation
ــ خطی	linear —
دنباله	sequence
دوتائی	binary

ر

رابطه	relation
ـ بازگشتی	recurrence —
رتبه	rank
رقم	digit
ـ با معنی	significant —
روش	method
ـ تکراری	iterative —
ـ تک گامی	one step —
ـ دو بخشی	bisection —
ـ نابجائی	false position —
ـ نیوتن	Newton —
ـ وتری	secant —
روند	process

ز

زیر ماتریس	submatrix

س

ستون	column
سرریز	overflow
سری	series
سطر	row

ش

شتاب	acceleration
شرط	condition
ـ اولیه	initial —
ـ مرزی	boundary —

ص

صفر	zero
ـ چندگانه	multiple

ع

عضو	element, member
عملگر	operator
ـ تفاضل پسرو	backward difference —
ـ تفاضل پیشرو	forward difference —
ـ تفاضل مرکزی	central difference —
عنصر	element
ـ محوری	pivotal —

غ

غالب	dominant
ـ قطری	diagonally —

ق

قاعده	rule
ـ ذوزنقه‌ای	trapezoidal —
ـ سیمپسون	Simpson —
ـ نقطه میانی	mid-point —
قضیه	theorem
ـ مقدار میانگین	mean value —

ک

کامل	complete
کران	bound
کل	tolal
کمترین مربعات	least squares
کوادراتور	quadrature

گ

گام	step
ـ به ـ	— by —
گرد کردن	rounding

م

ماتریس	matrix
ـ افزوده	augmented —
ـ سه قطری	tridiagonal —
ـ قطر غالب	diagonally dominant —
ـ متعامد	orthogonal —
ـ متقارن	symmetric —
ـ مثلثی	triangular —
ـ مربع	square —
ـ معکوس	inverse —
ـ معین مثبت	positive definite —
ـ واحد	unit —
ـ یکانی	unitary —
مانتیس	mantissa
مانده	residual
مبنا	base
متعامد	orthogonal
متقارن	symmetric
متناهی	finite
مثلثی	triangular
مجموعه	set
محور گیری	pivoting
ـ جزئی	partial —
ـ کلی	complete —
مرکزی	central
مستقل	independent
ـ خطی	linear —
مشتق	derivative
ـ جزئی	partial —
معادلات	equations
ـ تفاضلی	difference —

ــ سخت	stiff —
ــ مثلثی	triangular —
ــ نرمال	normal —
معکوس	inverse
معین	definite
مقدار	value
ــ آغازی	starting —
ــ اولیه	initial —
ــ مرزی	boundary —
مقدار ویژه	eigenvalue
ممیز سیار	floating point
منفرد	singular

ن

نابجایی	false posilion
ناپایدار	unstable
ناپایداری	unstability
ــ ذاتی	inherent —
ــ ضعیف	weack —
ناسره	improper
نامتناهی	infinite
نا معین	indefinite
نرم	norm
ــ برداری	vector —
ــ ماتریسی	matrix —
نرمال	normal
نسبی	relative
نقطه میانی	mid-point
نما	exponent
نمایش	representation
ــ دوتائی	binary —

و

وابسته	dependent
ــ خطی	linear —
واگرا	divergent

واگرایی	divergence
وزن	weight

ی

یکنوا	monotonic
یکنواخت	uniform

ه

همگرا	convergent
همگرایی	convergence
همگردان	compiler
هموار	smooth

واژه‌نامه انگلیسی به فارسی

A

absolute	مطلق
acceleration	شتاب
algorithm	الگوریتم
approximation	تقریب

B

base	مبنا
binary	دوتائی
— system	دستگاه ـ
bound	کران
boundary	مرز
— condition	شرط ـ ی
— value	مقدار ـ ی

C

central	مرکزی
— operator	عملگر ـ
chopping	قطع کردن
column	ستون
— interchange	تعویض ـ
compiler	همگردان

complete	کامل
—pivoting	محور گیری —
condition	شرط
consistent	سازگار
— equations	معادلات —
continuous	پیوسته
contradiction	تناقض
convergence	همگرائی
convergent	همگرا
correct	درست
correction	تصحیح
corrector	اصلاحگر

D

data	داده‌ها
definite	معین
— integral	انتگرال —
deflation	تقلیل
dependent	وابسته
derivative	مشتق
determinant	دترمینان
diagonal	قطر
differences	تفاضلات
digit	رقم
dimension	بعد
divergent	واگرا
divided differences	تفاضلات تقسیم شده
domain	قلمرو ــ حوزه ــ ناحیه
dominant	غالب

E

economization	اقتصادی کردن
element	عنصر ــ عضو
elimination	حذف
equal	مساوی
equation	معادله

واژه‌نامه انگلیسی به فارسی

error	خطا
estimate	تخمین
exponent	نما
exponential	نمائی
extrapolation	برونیابی

F

factor	عامل
factorization	تجزیه
false position	نابجایی
finite	متناهی
fixed point	نقطه ثابت
floating point	ممیز سیار
formula	فرمول
Fourier series	سری فوریه
fraction	کسر
function	تابع

I

ill-conditioned	بدوضع ــ بدخیم
illposed	بد طرح
improper	ناسره
— integral	انتگرال ــ
independent	مستقل
integration	انتگرال گیری
interchange	تعویض
interpolation	درونیابی
iteration	تکرار
itrative	تکراری
— method	روش ــ

L

least square	کمترین مربعات

— method	روش —
linear	خطی
— interpolation	درونیابی —

M

mantissa	مانتیس
matrix	ماتریس
maximum	ماکزیمم
method	روش
mid-point	نقطه میانی
— rule	قاعدهٔ —
minimum	مینیمم
monotonic	یکنوا
multiple	چندگانه
— root	ریشه —

N

neibourhood	همسایگی
nonsingular	نامنفرد
— matrix	ماتریس —

O

operator	عملگر
orthogonal	متعامد
— functions	توابع —
— series	سری —
overflow	سرریز
over-relaxation	فوق تخفیف

P

pivot	محور
— element	عضو — ی

pivoting	محور گیری
Polynomial	چند جمله‌ای
power	توان
predictor	پیشگو
process	روند
product	ضرب

R

rank	رتبه
recurrence	بازگشت
— relation	رابطهٔ ـ ی
relation	رابطه
relative	نسبی
—error	خطای ـ
remainder	باقیمانده
representation	نمایش
residual	مانده
root	ریشه
rounding	گرد کردن
row	سطر
—interchange	تعویض ـ
rule	قاعده

S

secant method	روش وتری
sequence	دنباله
series	سری
set	مجموعه
singular	منفرد
— matrix	ماتریس ـ
size	اندازه
smooth	هموار
solution	حل ـ جواب
space	فضا
sparse	تنک
stability	پایداری

stable	پایدار
submatrix	زیر ماتریس
substitution	جایگذاری
symmetric	متقارن
symmetry	تقارن

T

theorem	قضیه
transformation	تبدیل
transpose	ترانهاد
triangular	مثلثی
— matrix	ماتریس —
tridiagonal matrix	ماتریس سه قطری

U

unstability	ناپایداری
unstable	ناپایدار

V

vector	بردار

W

weak	ضعیف
weight	وزن
— function	تابع —
well-conditioned	خوش وضع
wellposed	خوش طرح

Z

zero	صفر

منابع

1. Balfour, A., and McTernan, A. J., 1967, *The Numerical Solution of Equations,* Heinemann, London.
2. Bareiss, E. H., 1967, The numerical solution of polynomial equations and the resultant procedures. In Ralston and Wilf, 1967.
3. Bauer, F. L., 1963, Optimally scaled matrices, *Num. Math.,* **5**, 73-87.
4. Buckingham, R. A., 1957, *Numerical Mehtods,* Pitman, London.
5. Bull, G., 1966, *Computational Methods and Algol,* Harrap, London.
6. Burden, R. L., Faires, T. D. and Reynolds, A. C., Numerical Analysis, Prindle, Weber and, *Schmidt, 1978*, Boston, Massachusetts.
7. Butler, R. and Kerr, E., 1962, *An Introduction to Numerical Mehtods,* Pitman, London.
8. Conte, S. D., 1965, *Elementary Numerical Analysis,* McGraw-Hill, New York.
9. Dahlquist, G., 1956, Convergence and stability in the numerical integration of ordinary differential equations, *Math. Scand.*, **4**, 33-53.
10. Davis, P. J., 1964, *Interpolation and Approximation,* Blaisdell, New York.
11. Davis, P. J. and Rabinowitz, P., 1967, *Numerical Integration,* Blaisdell, Waltham, Mass.
12. De, Boor, C. D., Elementary Numerical Analysis International student Edition, 1972, London.
13. Forsythe, G. E. and Moler, C. B., 1967, *Computer Solution of Linear Algebraic Systems,* Prentice-Hall, Englewood Cliffs, N. J.
14. Fox, L. (Ed.), 1962, *Numerical Solution of Ordinary and Partial Differential Equations,* Pergamon Press, Oxford.
15. Fox, L., 1964, *An Introduction to Numerical Linear Algebra,* Clarendon Press, Oxford.
16. Fox, L. and Mayers, D. F., 1968, *Computing Methods for Scientists and Engineers,* Clarendon Press, Oxford.

17. Francis, J. G. F., 1961, The Q–R transformation–a unitary analogue to the L–R transformation, *Comp. J.*, **4**, 265-271, 332-345.
18. Gear, C. W., 1968, The automatic integration of stiff ordinary differential equations, *Proc. IFIPS Conf.*, Edinburgh.
19. Goldberg, S., 1958, *Introduction to Difference Equations*, John Wiley, New York.
20. Hamming, R. W., 1962, *Numerical Methods for Scientists and Engineers*, McGraw-Hill, New York.
21. Hartree, D. R., 1958, *Numerical Analysis*, Oxford University Press, London.
22. Henrici, P., 1962, *Discrete Variable Mehtods in Ordinary Differential Equations*, John Wiley, New York.
23. Henrici, P., 1963, *Error Propagation for Difference Mehtods*, John Wiley, New York.
24. Henrici, P., 1964, *Elements of Numerical Analysis*, John Wiley, New York.
25. Hildebrand, F. B., 1956, *Introduction to Numerical Analysis*, McGraw-Hill, New York.
26. Householder, A. S., 1953, *Principles of Numerical Analysis*, McGraw-Hill, New York.
27. Isaacson, E. and Keller, H. B., 1966, *Analysis of Numerical Methods*, John Wiley, New York.
28. Jonson, L. W. and Dean Ricess, R. Numerical Analysis 2Ed, 1982, Addison -Wesley Publishing Company, London.
29. Keller, H. B., 1968, *The Numerical Solution of Boundary Value Problems*, Blaisdell New York.
30. Kopchenova, N. V. and Maron, I. A., Computational Mathematics, Mir Publishes, Moscow,
31. Krylov, V. I., 1962, *Approximate Calculation of Integrals*, trans. Stroud, A. H., Macmillan, New York.
32. Lafara, R., Computer Method for Science and Engineering, International Textbook Company Ltd, London.
33. Lanczos, C., 1957, *Applied Analysis*, Pitman, London.
34. Lehmer, D. H., 1961, A machine method for solving polynomial equations, *J. A. C. M.*, **8**, 151-162.
35. McCracken, D. D., and Dorn, W. S., 1964, *Numerical Methods and* FORTRAN *Programming*, John Wiley, New York.
36. *Modern Computing Methods*, 1961, HMSO, London.
37. Nordsieck, A., 1962, Numerical integration of ordinary differential equations, *Maths Comp.*, **16**, 22-49.
38. Ostrowski, A. M., 1960, *Solution of Equations and Systems of Equations*, Academic Press, New York.
39. Parlett, B. N., 1967, The L–U and Q–R algorithms. In Ralston and Wilf, 1967.

40. Ralston, A., 1965, *A First Course in Numerical Analysis,* McGraw-Hill, New York.
41. Ralston, A. and Wilf, H. S., 1960, *Mathematical Mehtods for Digital Computers,* Vol. 1, John Wiley, New York.
42. Ralston, A. and Wilf, H. S., 1967, *Mathematical Mehtods for Digital Computers,* Vol. 2, John Wiley, New York.
43. Redish, K. A., 1961, *An Introduction to Computational Methods,* English Universities Press, London.
44. Rice, J. R., 1964, *The Approximation of Functions,* Vol. 1, Addison-Wesley, New York.
45. Rutishauser, H. 1956, Der Quotienten-Differenzen-Algorithmus. *Mitteilungen aus dem Institut fur angew. Math. No. 7,* Birkhauser, Basel and Stuttgart.
46. Rutishauser, H., 1958, Solution of eigenvalue problems with the L–R transformation, *App. Math. Ser. Nat. Bur. Stand.,* **49**, 47-81.
47. Scarborough, J. B., 1958, *Numerical Mathematical Analysis,* Johns Hopkins Press, Baltimore.
48. Smith, G. D., 1964, *The Numerical Solution of Partial Differential Equations,* Oxford University Press, London.
49. Stroud, A. H., and Secrest, D., 1966, *Gaussian Quadrature Formulae,* Prentice-Hall, Englewood Cliffs, N. J.
50. Todd, J., 1962, *Survey of Numerical Analysis,* McGraw-Hill, New York.
51. Traub, J. F., 1964, *Iterative Methods for the Solution of Equations,* Prentice-Hall, Englewood Cliffs, N. J.
52. Varga, R. S., 1962, *Matrix Iterative Analysis,* Prentice-Hall, Englewood Cliffs, N. J.
53. Wilkinson, J. H., 1960, Householder's method for the solution of the algebraic eigenproblem, *Comp. J.,* **3**, 23-27.
54. Wilkinson, J. H., 1961, Error analysis of direct methods of matrix inversion, *J. A. C. M.,* **8**, 281-330.
55. Wilkinson, J. H., 1963, *Rounding Errors In Algebraic Processes,* HMSO, London.
56. Wilkinson, J. H., 1965, *The Algebraic Eigenvalue Problem,* Oxford University Press, London.
57. Williams, P. W., 1979, Numerical Computation, Thomas Nelson and Sons Ltd, London.

Numerical Computation

By:
K. Maleknejad, Ph. D
E. Babolian, Ph. D