



دانشگاه گیلان
۱۳۵۳ - ۱۳۷۴

جانب اول

آنالیز عددی پیشرفته

آنالیز عددی پیشرفته

مرکز نشر دانشگاه گیلان

University of Guilan Press

تدوین (به شیوه گردآوری)
دکتر داود خجسته سالکویه
استاد دانشکده علوم ریاضی دانشگاه گیلان

تدوین: دکتر داود خجسته سالکویه

BY:

Davod Khojasteh Salkuyeh, Ph.D



9786001532931



آنالیز عددی پیشرفته

تدوین (به شیوه گردآوری):

دکتر داود خجسته سالکویه

استاد دانشکده علوم ریاضی دانشگاه گیلان

مرکز نشر دانشگاه گیلان

۱۴۰۱



دانشگاه گیلان
۱۳۵۳-۱۹۷۴

شابک: ۱-۲۹۳-۱۵۳-۶۰۰-۹۷۸

سرشناسه	: خجسته سالکویه، داود، ۱۳۴۹ -
عنوان و نام پدیدآور	: آنالیز عددی پیشرفته / تدوین (به شیوه گردآوری) داود خجسته سالکویه؛ ویراستار ادبی فرشته رحمانی نژاد؛ ویراستار علمی حسین امینی خواه.
مشخصات نشر	: رشت: دانشگاه گیلان، انتشارات، ۱۴۰۱.
مشخصات ظاهری	: ۲۴۴ ص.
شابک	: 978-600-153-293-1
وضعیت فهرست نویسی	: فیبا
یادداشت	: پشت جلد به انگلیسی: Davod Khojasteh Salkuyeh, Advanced Numerical Analysis
یادداشت	: کتابنامه.
موضوع	: آنالیز عددی -- راهنمای آموزشی (عالی) Numerical analysis -- Study and teaching آنالیز عددی -- مسائل، تمرین‌ها و غیره (عالی) Numerical analysis -- Problems, exercises, etc (Higher)
شناسه افزوده	: امینی خواه، حسین، ۱۳۵۷ -، ویراستار
شناسه افزوده	: دانشگاه گیلان
رده بندی کنگره	: QA۲۹۷
رده بندی دیویی	: ۵۱۹/۴
شماره کتابشناسی ملی	: ۸۹۳۱۲۸۲
اطلاعات رکورد	: فیبا
کتابشناسی	

مرکز نشر دانشگاه گیلان

نام کتاب	: آنالیز عددی پیشرفته
گردآوری	: دکتر داود خجسته سالکویه
ویراستار ادبی	: فرشته رحمانی نژاد
ویراستار علمی	: دکتر حسین امینی خواه
نوبت چاپ	: اول، ۱۴۰۱
ناشر	: مرکز نشر دانشگاه گیلان
شمارگان	: ۱۰۰۰ نسخه

* هر گونه چاپ و تکثیر صرفاً در اختیار انتشارات دانشگاه گیلان است.*

تقدیم به

پدر و مادر عزیزم

پیشگفتار

بسیاری از پدیده‌ها در علوم پایه و مهندسی، توسط معادلات مختلف ریاضی مثل دستگاه معادلات خطی و غیرخطی، معادلات دیفرانسیل معمولی و جزئی، معادلات انتگرال و ... مدل‌بندی می‌شوند. در بیشتر موارد، حل دقیق این گونه از معادلات بسیار سخت و حتی گاهی ناممکن است. از این رو، برای حل آنها از روش‌های عددی استفاده می‌شود. با توجه به اینکه ابعاد مسائلی که در این فرایند ظاهر می‌شوند، بزرگ و حجم محاسبات زیاد است، استفاده از رایانه ضروری است. به دلیل محدودیت رایانه‌ها در نمایش رقم‌های یک عدد، اعداد در رایانه به صورت تقریبی ذخیره (گرد) می‌شوند؛ بنابراین هر محاسبه ساده با یک خطا همراه است. از این رو، محاسبات تحت تأثیر عمل گرد کردن قرار می‌گیرند و گاهی نتایج به دست آمده غیرقابل پیش‌بینی و نادرست هستند. بنابراین، مهم‌ترین پرسشی که مطرح می‌شود، این است که چه زمانی می‌توان به یک جواب به دست آمده توسط رایانه اعتماد کرد. برای یافتن پاسخ این پرسش، لازم است با سیستم محاسبات رایانه‌ای آشنا شویم تا بتوانیم خطاهای ناشی از گرد کردن اعداد در رایانه را تحلیل کنیم. پس از ارائه پیش‌نیازهایی در فصل نخست، در فصل دوم به موضوع آنالیز خطا می‌پردازیم.

سه فصل بعدی، به ارائه روش‌های عددی برای حل مسائل مختلف اختصاص داده شده است. در فصل سوم، موضوع درونیابی مورد مطالعه قرار گرفته است. در این فصل، انواع درونیابی از قبیل درونیابی توسط چندجمله‌ای‌ها، درونیابی توسط توابع گویا، توابع مثلثاتی و توابع اسپلاین بررسی شده است. فصل چهارم به روش‌های محاسبه تقریبی انتگرال‌های معین اختصاص یافته است. در این فصل، روش‌های انتگرال‌گیری نیوتن-کاتس و فرمول‌های انتگرال‌گیری گاوس مورد بررسی قرار می‌گیرند. در فصل پایانی، به مطالعه روش‌های مختلف برای حل معادلات غیرخطی پرداخته شده است. در تمام فصل‌ها تلاش شده است تا آنجا که ممکن است، آنالیز خطای هر روش ارائه شود.

برای روش‌شدن مطالب ارائه‌شده، تلاش شده است برای هر روشی که ارائه می‌شود، یک مثال یا کاربردی از آن روش بیان شود. در پایان هر فصل تمرین‌هایی قرار داده شده است که حل آنها

به درک بهتر مطالب کمک می‌کند.

کتاب پیش رو، چندین سال به صورت جزوهٔ درسی به دانشجویان کارشناسی ارشد در درس آنالیز عددی پیشرفته در دانشگاه گیلان و دانشگاه محقق اردبیلی تدریس شده است و افراد زیادی از جمله دانشجویان این درس، اشکالات و کاستی‌های آن را به نگارنده یادآوری کرده‌اند که شایسته است از همهٔ آنها سپاسگزاری کنم. از ویراستار علمی کتاب، دکتر حسین امینی خواه، که با دقت و نکته‌سنجی خود به بررسی علمی کتاب پرداختند، صمیمانه سپاسگزاری می‌کنم. در پایان، از ویراستار ادبی کتاب، خانم فرشته رحمانی‌نژاد، که با دقت به مطالعهٔ متن کتاب پرداختند و نکات نگارشی بسیار ارزشمندی را یادآور شدند و همچنین، از واحد پژوهشی دانشگاه گیلان که زمینهٔ چاپ اثر را فراهم کردند، قدردانی می‌کنم.

داود خجسته سالکویه

دانشکدهٔ علوم ریاضی دانشگاه گیلان

تیر ۱۴۰۱

فهرست مندرجات

۱	پیش‌نیازها	۱
۱	۱.۱ پیش‌نیازهایی از جبر خطی	۱.۱
۸	۲.۱ حل دستگاه معادلات خطی	۲.۱
۸	۱.۲.۱ روش حذفی گاوس	۱.۲.۱
۱۲	۲.۲.۱ روش حذفی گاوس با محورگیری جزئی	۲.۲.۱
۱۳	۳.۱ پیش‌نیازهایی از حساب و دیفرانسیل	۳.۱
۲۱	۲ خطاها	۲
۲۱	۱.۲ اعداد ممیز شناور در دستگاه IEEE	۱.۲
۲۵	۲.۲ حساب IEEE	۲.۲
۲۷	۳.۲ آنالیز خطا	۳.۲
۳۷	۴.۲ جلوگیری از حذف فاجعه‌انگیز	۴.۲
۳۸	۵.۲ خوش‌وضع، پایداری و حساسیت	۵.۲
۵۳	۳ درونیابی	۳
۵۳	۱.۳ درونیابی با استفاده از چند جمله‌ای‌ها	۱.۳
۶۱	۲.۳ تفاضلات تقسیم‌شده و چند جمله‌ای درونیاب نیوتن	۲.۳
۶۵	۳.۳ خطای درونیابی	۳.۳

۶۷	مسئلهٔ اقل اکثر	۴.۳
۷۴	درونیابی هرمیت	۵.۳
۸۳	درونیابی توسط توابع گویا	۶.۳
۹۰	تفاضلات معکوس و کسرهای مسلسل تیل	۷.۳
۹۶	درونیابی مثلثاتی	۸.۳
۱۰۲	درونیابی توابع چندمتغیره	۹.۳
۱۰۴	درونیابی با کمک توابع اسپلاین	۱۰.۳
۱۱۰	محاسبهٔ اسپلاین مکعبی درونیاب	۱۱.۳
۱۲۱	یکی از کاربردهای مهم توابع اسپلاین	۱۲.۳
۱۲۲	همگرایی تابع اسپلاین مکعبی درونیاب	۱۳.۳
۱۲۹	انتگرال گیری عددی	۴
۱۲۹	مقدمه	۱.۴
۱۲۹	فرمول های نیوتن- کاتس	۲.۴
۱۳۵	دستورهای مرکب انتگرال گیری	۳.۴
۱۳۹	نمایش خطای پئانو	۴.۴
۱۴۴	روش رامبرگ	۵.۴
۱۴۸	فرمول های انتگرال گیری گاوس	۶.۴
۱۵۸	فرمول های انتگرال گیری گاوس- چپیشف و گاوس- لژاندر	۷.۴
۱۷۳	حل عددی معادلات غیرخطی	۵
۱۷۳	روش نیوتن- رافسون	۱.۵
۱۸۱	قضایای همگرایی در حالت کلی	۲.۵

فهرست مندرجات

ج

۱۸۸	همگرایی روش نیوتن-رافسون	۳.۵
۱۹۲	روش بیرستو	۴.۵
۱۹۸	روش نابجایی	۵.۵
۲۰۱	روش وتری	۶.۵
۲۰۴	روش Δ^2 -ایتکن	۷.۵

۲۱۳

منابع

۲۱۵

واژه‌نامه فارسی به انگلیسی

۲۲۳

واژه‌نامه انگلیسی به فارسی

فصل ۱

پیش نیازها

در این فصل برخی از مفاهیم، تعاریف، روش‌ها و قضایای موردنیاز در فصل‌های بعدی را مرور می‌کنیم. فرض ما بر این است که خواننده با مفاهیمی مثل میدان، فضای برداری، دترمینان، و رتبه آشنایی کامل دارد. در سراسر این کتاب، مجموعه اعداد حقیقی (میدان اعداد حقیقی) و مختلط (میدان اعداد مختلط) را به ترتیب با \mathbb{R} و \mathbb{C} نشان می‌دهیم. همچنین، مجموعه بردارهای n تایی با مؤلفه‌های حقیقی را با \mathbb{R}^n و مجموعه ماتریس‌های $m \times n$ با درایه‌های حقیقی را با $\mathbb{R}^{m \times n}$ نشان می‌دهیم. به همین ترتیب، مجموعه \mathbb{C}^n و $\mathbb{C}^{m \times n}$ تعریف می‌شود.

۱.۱ پیش‌نیازهایی از جبر خطی

ابتدا تعریف زوج ویژه یک ماتریس و چند نکته در مورد آن را ارائه می‌کنیم.

تعریف ۱.۱ فرض کنید $A \in \mathbb{C}^{n \times n}$. اگر به ازای یک بردار ناصفر $x \in \mathbb{C}^n$ و عدد $\lambda \in \mathbb{C}$ داشته باشیم $Ax = \lambda x$ ، آنگاه زوج مرتب (λ, x) را یک زوج ویژه A می‌گویند. λ را مقدار ویژه و x را بردار ویژه A می‌گویند. مجموعه مقادیر ویژه A را طیف A می‌گویند و با $\sigma(A)$ نشان می‌دهند. همچنین، شعاع طیفی ماتریس A به صورت زیر تعریف می‌شود

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

اگر (λ, x) یک زوج ویژه ماتریس A باشد، آنگاه از رابطه $Ax = \lambda x$ خواهیم داشت

$$(\lambda I - A)x = 0.$$

از اینکه بردار ویژه x ناصفر است، باید داشته باشیم $\det(\lambda I - A) = 0$. تعریف می‌کنیم

$$p(\lambda) = \det(\lambda I - A).$$

می‌توان دید که $p(\lambda)$ یک چندجمله‌ای درجه n است که ریشه‌های آن مقادیر ویژه A هستند. $p(\lambda)$ را چندجمله‌ای مشخصه A می‌گویند.

مثال ۱.۱ زوج‌های ویژه ماتریس

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix},$$

را به دست آورید.

حل: با صفر قراردادن چندجمله‌ای مشخصه ماتریس A ، داریم

$$p(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & -1 \\ -2 & \lambda - 3 \end{vmatrix} = (\lambda - 1)(\lambda - 4) = 0.$$

بنابراین، مقادیر ویژه A عبارتند از $\lambda_1 = 1$ و $\lambda_2 = 4$. از این رو، $\sigma(A) = \{1, 4\}$ و $\rho(A) = 4$. برای محاسبه بردار ویژه متناظر با $\lambda_1 = 1$ ، از رابطه $(\lambda_1 I - A)x = 0$ ، داریم

$$\begin{pmatrix} -1 & -1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

حل این دستگاه، با حل معادله $x_1 + x_2 = 0$ ، یا $x_2 = -x_1$ ، معادل است. بنابراین، هر بردار به صورت $x = (x_1, -x_1)^T$ ، یک بردار ویژه A متناظر با $\lambda_1 = 1$ است، به شرطی که x_1 ناصفر باشد. برای مثال، بردار $x = (1, -1)^T$ ، یک بردار ویژه متناظر با $\lambda_1 = 1$ است. به طور مشابه می‌توان دید که بردار $x = (1, 2)^T$ یک بردار ویژه متناظر با $\lambda_2 = 4$ است.

تعریف ۲.۱ فرض کنید V یک فضای برداری روی میدان \mathbb{F} باشد. یک نرم روی V ، تابعی است

مثل $\|\cdot\|: V \rightarrow \mathbb{R}$ ، به طوری که در شرایط زیر صدق کند

(الف) به ازای هر $x \in V$ ، $\|x\| \geq 0$. به علاوه $\|x\| = 0$ اگر و تنها اگر $x = 0$ ؛

(ب) به ازای هر $\alpha \in \mathbb{F}$ و $x \in V$ ، $\|\alpha x\| = |\alpha| \|x\|$ ؛

(ج) به ازای هر $x, y \in V$ ، $\|x + y\| \leq \|x\| + \|y\|$ (نامساوی مثلث).

اگر $\|\cdot\|$ یک نرم روی V باشد، آنگاه به‌ازای هر $x, y \in V$ داریم

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|.$$

از اینجا به بعد نرم‌های تعریف‌شده روی \mathbb{C}^n و $\mathbb{C}^{m \times n}$ را به‌ترتیب، نرم برداری و نرم ماتریسی می‌نامیم.

برای هر $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$ نرم‌های برداری متداول عبارتند از

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}},$$

که به‌ترتیب نرم-۱، نرم بینهایت و نرم-۲ (نرم اقلیدسی) هستند. همچنین، مهم‌ترین نرم‌های ماتریسی روی $\mathbb{C}^{n \times n}$ عبارتند از

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|,$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

$$\|A\|_2 = \rho(A^H A)^{\frac{1}{2}},$$

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}},$$

که به‌ترتیب به نرم-۱، نرم بینهایت، نرم-۲ و نرم فروبنیوس معروفند.

مثال ۲.۱ اگر $x = (1, 2, -2)^T$ آنگاه

$$\|x\|_1 = |1| + |2| + |-2| = 5,$$

$$\|x\|_\infty = \max\{|1|, |2|, |-2|\} = 2$$

$$\|x\|_2 = \sqrt{|1|^2 + |2|^2 + |-2|^2} = \sqrt{9} = 3.$$

همچنین، فرض کنید

$$A = \begin{pmatrix} 2 & -2 \\ 5 & -1 \end{pmatrix}.$$

در این صورت، داریم

$$\|A\|_1 = \max\{|2| + |5|, |-2| + |-1|\} = 7,$$

$$\|A\|_\infty = \max\{|2| + |-2|, |5| + |-1|\} = 6,$$

$$\|A\|_F = \sqrt{|2|^2 + |-2|^2 + |5|^2 + |-1|^2} = \sqrt{34}.$$

برای محاسبه $\|A\|_2$ ، داریم

$$A^T A = \begin{pmatrix} 29 & -9 \\ -9 & 5 \end{pmatrix}.$$

به‌سادگی می‌توان دید که $\sigma(A^T A) = \{2, 32\}$. از این رو، $\|A\|_2 = \sqrt{32}$.

تعریف ۳.۱ (نرم طبیعی) فرض کنید $\|\cdot\|$ یک نرم برداری روی \mathbb{C}^n باشد و $A \in \mathbb{C}^{n \times n}$. در این صورت، نرم طبیعی متناظر با این نرم برداری به صورت

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \max_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|}, \quad (1.1)$$

تعریف می‌شود.

می‌توان دید که نرم‌های طبیعی تولیدشده توسط نرم‌های برداری نرم-۱، نرم بینهایت و نرم-۲ به ترتیب نرم‌های ماتریسی طبیعی نرم-۱، نرم بینهایت و نرم-۲ هستند. همچنین، اگر $\|\cdot\|_v$ یک نرم برداری و $\|\cdot\|_M$ نرم طبیعی متناظر باشد، آنگاه به‌ازای هر $A, B \in \mathbb{C}^{n \times n}$ و $x \in \mathbb{C}^n$ داریم

$$\|AB\|_M \leq \|A\|_M \|B\|_M, \quad \|Ax\|_v \leq \|A\|_M \|x\|_v.$$

خاصیت اول را خاصیت ضربی نرم و خاصیت دوم را خاصیت سازگاری نرم ماتریسی نسبت به نرم برداری مولد آن می‌گویند.

تعریف ۴.۱ فرض کنید V یک فضای برداری روی میدان \mathbb{F} است. یک ضرب داخلی روی V ، تابعی است که به هر زوج از بردارهای x و y در V یک اسکالر مثل $\langle x, y \rangle$ در \mathbb{F} را نسبت می‌دهد به طوری که در شرایط زیر صدق کند

(الف) $\langle x, x \rangle$ حقیقی باشد و $\langle x, x \rangle \geq 0$. به‌علاوه، $\langle x, x \rangle = 0$ اگر و تنها اگر $x = 0$ ؛

(ب) برای هر $\alpha \in \mathbb{F}$ ، $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$ ؛

(ج) برای هر $z \in V$ ، $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$ ؛

(د) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

هر فضای برداری را که یک ضرب داخلی در آن تعریف شده باشد، یک فضای ضرب داخلی می‌نامند.

فرض کنید V یک فضای ضرب داخلی روی میدان \mathbb{F} و $\langle \cdot, \cdot \rangle$ یک ضرب داخلی در V باشد. در این صورت، اگر به‌ازای هر $x \in V$ قرار دهیم

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad (2.1)$$

آنگاه $\|\cdot\|$ ، یک نرم روی V تعریف می‌کند. اگر V یک فضای ضرب داخلی روی \mathbb{F} باشد و $x, y \in V$ آنگاه

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

که در آن $\|\cdot\|$ نرم تعریف‌شده در رابطه (۲.۱) است. این نامساوی به نامساوی کوشی-شوارتز معروف است.

مثال ۳.۱ به‌ازای هر دو بردار $x, y \in \mathbb{C}^n$ تعریف می‌کنیم

$$\langle x, y \rangle = x^H y = \sum_{i=1}^k \bar{x}_i y_i,$$

که در آن x_i ها و y_i ها به‌ترتیب مؤلفه‌های x و y هستند. به‌سادگی می‌توان نشان داد که $\langle \cdot, \cdot \rangle$ یک ضرب داخلی روی \mathbb{C}^n است. این ضرب داخلی، ضرب داخلی استاندارد روی \mathbb{C}^n نامیده می‌شود. نرم تولیدشده توسط این ضرب داخلی، نرم-۲ است؛ زیرا برای هر $x \in \mathbb{C}^n$ داریم

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{i=1}^n \bar{x}_i x_i \right)^{\frac{1}{2}} = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} = \|x\|_2.$$

همچنین، نامساوی کوشی-شوارتز به‌صورت

$$|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2,$$

یا

$$\left| \sum_{i=1}^n \bar{x}_i y_i \right| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |y_i|^2 \right)^{\frac{1}{2}},$$

است.

تعریف ۵.۱ ماتریس $A \in \mathbb{R}^{n \times n}$ را معین مثبت متقارن می‌گوییم، هرگاه

$$(الف) \quad A^T = A \text{ (متقارن باشد)};$$

$$(ب) \quad \text{برای هر بردار ناصفر } x \in \mathbb{R}^n, \quad x^T A x > 0.$$

ماتریس A را نیمه‌معین مثبت متقارن می‌گویند، هرگاه (الف) برقرار باشد و برای هر $x \in \mathbb{R}^n$ ، $x^T A x \geq 0$.

مثال ۴.۱ ماتریس

$$A = \begin{pmatrix} 5 & -4 \\ -4 & 6 \end{pmatrix},$$

معین مثبت متقارن است؛ زیرا هم متقارن است و هم برای هر بردار ناصفر $x = (x_1, x_2)^T$ داریم

$$\begin{aligned} x^T A x &= (x_1, x_2) \begin{pmatrix} 5 & -4 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= x_1^2 + 4(x_1 - x_2)^2 + 2x_2^2 > 0. \end{aligned}$$

تعریف ۶.۱ فرض کنید $A \in \mathbb{R}^{n \times n}$ و $1 \leq i_1 < i_2 < \dots < i_k \leq n$. در این صورت، ماتریس

$$A_k = \begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \dots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \dots & a_{i_2 i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \dots & a_{i_k i_k} \end{pmatrix},$$

یک زیرماتریس اصلی A نامیده می‌شود. در صورتی که $i_j = j$ ، $j = 1, 2, \dots, k$ ، ماتریس A_k یک زیرماتریس اصلی پیشرو ماتریس A نامیده می‌شود.

مثال ۵.۱ فرض کنید

$$A = \begin{pmatrix} 8 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 1 \end{pmatrix}.$$

در این صورت، زیرماتریس اصلی که توسط اندیس‌های $i_1 = 1$ و $i_3 = 3$ به دست می‌آید به صورت

$$A = \begin{pmatrix} 8 & 1 \\ 1 & 1 \end{pmatrix},$$

است. همچنین، ماتریس

$$A = \begin{pmatrix} 8 & 2 \\ 2 & 5 \end{pmatrix},$$

یک زیرماتریس اصلی پیشروی ماتریس A است که توسط اندیس‌های $i_1 = 1$ و $i_2 = 2$ تولید می‌شود.

قضیه ۱.۱ فرض کنید ماتریس $A \in \mathbb{R}^{n \times n}$ متقارن است.

(الف) A معین مثبت است اگر و تنها اگر درمیان همه زیرماتریس‌های اصلی پیشرو آن مثبت باشند.

(ب) A معین مثبت است اگر و تنها اگر مقادیر ویژه A مثبت باشند.

مثال ۶.۱ ماتریس

$$A = \begin{pmatrix} 9 & -3 & 2 \\ -3 & 5 & -2 \\ 2 & -2 & 1 \end{pmatrix},$$

را در نظر بگیرید. داریم

$$\begin{aligned} \det(A_1) &= 9 > 0, \\ \det(A_2) &= \begin{vmatrix} 9 & -3 \\ -3 & 5 \end{vmatrix} = 36 > 0, \\ \det(A_3) &= \det(A) = 4 > 0. \end{aligned}$$

از این رو، با توجه به قضیه ۱.۱، ماتریس A معین مثبت متقارن است؛ زیرا هم متقارن است و هم دترمینان تمام زیرماتریس‌های اصلی پیشرو A مثبت هستند.

مثال ۷.۱ ماتریس

$$A = \begin{pmatrix} 4 & -2 \\ -2 & 7 \end{pmatrix},$$

را در نظر بگیرید. مجموعه مقادیر ویژه ماتریس A برابر با $\{3, 8\}$ است. بنابراین، با کمک قضیه ۱.۱ می‌بینیم که ماتریس A معین مثبت متقارن است؛ زیرا هم متقارن است و هم مقادیر ویژه آن مثبت هستند.

فرض کنید $x_i, i = 1, \dots, n$ عدد حقیقی متمایز باشند. در این صورت، ماتریس

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix},$$

یک ماتریس واندرموند نامیده می‌شود. می‌توان دید که ماتریس واندرموند نامنفرد است و

$$\det(V) = \prod_{i < j} (x_i - x_j). \quad (3.1)$$

برای مثال، اگر $n = 4$ ، آنگاه

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{pmatrix},$$

$$\det(V) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4) \text{ و}$$

۲.۱ حل دستگاه معادلات خطی

حل بسیاری از مسائل در علوم و مهندسی، به حل دستگاه معادلات خطی با ابعاد بزرگ منجر می‌شود. در این بخش، روش حذفی گاوس برای حل این گونه از دستگاه‌ها را مرور می‌کنیم.

۱.۲.۱ روش حذفی گاوس

دستگاه معادلات خطی

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n = b_3, \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = b_n, \end{cases} \quad (4.1)$$

را در نظر بگیرید، که در آن a_{ij} ها و b_i ها اعداد حقیقی و x_i ها مجهولات دستگاه هستند. روشن است که با سه عمل

(۱) تعویض دو معادله،

(۲) افزودن مضربی از طرفین یک معادله به طرفین معادله دیگر،

(۳) ضرب طرفین یک معادله در یک عدد ناصفر،

دستگاه (۴.۱) به دستگاه دیگری تبدیل می‌شود، که جواب آن با جواب دستگاه (۴.۱) یکسان است. این اعمال را اعمال سطری مقدماتی می‌گویند.

در روش حذفی گاوس، با کمک این سه عمل، دستگاه را به دستگاهی تبدیل می‌کنیم که ماتریس ضرایب آن یک ماتریس بالامثلثی باشد که به راحتی قابل حل است. در ادامه، الگوریتم روش حذفی گاوس را شرح می‌دهیم. نخست، ماتریس افزوده

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} & b_n \end{array} \right),$$

را تشکیل می‌دهیم. عملیات بالامثلثی‌سازی ماتریس ضرایب در روش حذفی گاوس در $(n - 1)$ گام انجام می‌پذیرد.

در گام نخست، اگر $a_{11} = 0$ ، آنگاه الگوریتم شکست می‌خورد. در غیر این صورت، قرار می‌دهیم

$$\ell_{i1} = \frac{a_{i1}}{a_{11}}, \quad i = 2, 3, \dots, n.$$

اکنون به‌ازای $i = 2, 3, \dots, n$ ، $-\ell_{i1}$ برابر سطر اول ماتریس افزوده را به سطر i م این ماتریس اضافه می‌کنیم. با انجام این اعمال، درایه‌های مکان $(i, 1)$ ، $i = 2, 3, \dots, n$ برابر با صفر خواهند شد. در واقع، ماتریس افزوده به صورت

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & \hat{a}_{22} & \hat{a}_{23} & \cdots & \hat{a}_{2n} & \hat{b}_2 \\ 0 & \hat{a}_{32} & \hat{a}_{33} & \cdots & \hat{a}_{3n} & \hat{b}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \hat{a}_{n2} & \hat{a}_{n3} & \cdots & \hat{a}_{nn} & \hat{b}_n \end{array} \right),$$

خواهد بود.

در گام دوم، درایه‌های مکان $(i, 2)$ ، $i = 3, \dots, n$ را صفر می‌کنیم. در صورتی که $\hat{a}_{22} = 0$ ، الگوریتم شکست می‌خورد. در غیر این صورت، قرار می‌دهیم

$$\ell_{i2} = \frac{\hat{a}_{i2}}{\hat{a}_{22}}, \quad i = 3, 4, \dots, n.$$

به‌ازای $i = 3, 4, \dots, n$ ، $-\ell_{i2}$ برابر سطر دوم ماتریس افزوده را به سطر i م این ماتریس اضافه می‌کنیم. با انجام این اعمال درایه‌های مکان $(i, 2)$ ، $i = 3, 4, \dots, n$ برابر با صفر خواهند شد. در واقع ماتریس افزوده به صورت زیر به دست می‌آید

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & \hat{a}_{22} & \hat{a}_{23} & \cdots & \hat{a}_{2n} & \hat{b}_2 \\ 0 & 0 & \tilde{a}_{33} & \cdots & \tilde{a}_{3n} & \tilde{b}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \tilde{a}_{n3} & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right).$$

با ادامه این روند، پس از $(n-1)$ گام، ماتریس ضرایب دستگاه بالامثلثی خواهد شد که دستگاه حاصل را به سادگی می‌توان حل کرد. در واقع، از آخرین معادله متغیر x_n محاسبه می‌شود. با جای‌گذاری x_n در معادله $(n-1)$ م، متغیر x_{n-1} به دست می‌آید. با ادامه این روند، همه متغیرها قابل محاسبه خواهند بود. این شیوه از جای‌گذاری را جای‌گذاری پسرو می‌گویند. مقادیری که روی قطر ماتریس بالامثلثی حاصل قرار دارند، درایه‌های محوری نامیده می‌شوند.

مثال ۸.۱ دستگاه معادلات خطی

$$\begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 6 & 2 & 1 \\ 1 & 4 & 4 & 2 \\ 4 & 2 & 8 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \\ 8 \\ 5 \end{pmatrix},$$

را با روش حذفی گاوس حل کنید.

حل: ماتریس افزودهٔ متناظر برای این دستگاه به صورت

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 5 \\ 2 & 6 & 2 & 1 & 11 \\ 1 & 4 & 4 & 2 & 8 \\ 4 & 2 & 8 & 1 & 5 \end{array} \right),$$

است.

گام ۱. قرار می‌دهیم

$$l_{21} = \frac{2}{1} = 2, \quad l_{31} = \frac{1}{1} = 1, \quad l_{41} = \frac{4}{1} = 4.$$

۲- برابر سطر اول را به سطر دوم، ۱- برابر سطر اول را به سطر سوم و ۴- سطر اول را به سطر چهارم اضافه می‌کنیم. در این صورت، خواهیم داشت

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 5 \\ 0 & 2 & -2 & -1 & 1 \\ 0 & 2 & 2 & 1 & 3 \\ 0 & -6 & 0 & -3 & -15 \end{array} \right).$$

گام ۲. قرار می‌دهیم

$$l_{32} = \frac{2}{2} = 1, \quad l_{42} = \frac{-6}{2} = -3.$$

۱- سطر دوم را به سطر سوم و ۳ برابر سطر دوم را به سطر چهارم اضافه می‌کنیم. از این رو، خواهیم داشت

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 5 \\ 0 & 2 & -2 & -1 & 1 \\ 0 & 0 & 4 & 2 & 2 \\ 0 & 0 & -6 & -6 & -12 \end{array} \right).$$

گام ۳. قرار می‌دهیم

$$l_{43} = \frac{-6}{4} = -\frac{3}{2}.$$

سپس، $\frac{3}{4}$ برابر سطر سوم را به سطر چهارم اضافه می‌کنیم. از این رو، خواهیم داشت

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 5 \\ 0 & 2 & -2 & -1 & 1 \\ 0 & 0 & 4 & 2 & 2 \\ 0 & 0 & 0 & -3 & -9 \end{array} \right).$$

می‌بینیم که ماتریس ضرایب دستگاه، بالامثلثی شده است. اکنون این دستگاه بالامثلثی را حل می‌کنیم. از معادله آخر داریم

$$-3x_4 = -9,$$

که مقدار $x_4 = 3$ را نتیجه می‌دهد. از معادله سوم داریم

$$4x_3 + 2x_2 = 2,$$

که با جای‌گذاری $x_4 = 3$ خواهیم داشت

$$4x_3 + 2 \times 3 = 2.$$

بنابراین، $x_3 = -1$. از معادله سوم داریم

$$2x_2 - 2x_3 - x_4 = 1,$$

که با جای‌گذاری $x_3 = -1$ و $x_4 = 3$ ، می‌بینیم که

$$2x_2 - 2 \times (-1) - 3 = 1.$$

از این رو، $x_2 = 1$. سرانجام از معادله اول داریم

$$x_1 + 2x_2 + 2x_3 + x_4 = 5,$$

که با جای‌گذاری $x_2 = 1$ ، $x_3 = -1$ و $x_4 = 3$ ، خواهیم داشت

$$x_1 + 2 \times 1 + 2 \times (-1) + 3 = 5.$$

بنابراین، $x_1 = 2$.

۲.۲.۱ روش حذفی گاوس با محورگیری جزئی

همان‌گونه که مشاهده شد، اگر در الگوریتم روش حذفی گاوس یکی از درایه‌های محوری صفر باشد، الگوریتم شکست می‌خورد. برای رفع این مشکل، از راهکار محورگیری استفاده می‌شود. به‌طورکلی، دو نوع محورگیری وجود دارد: محورگیری جزئی و محورگیری کلی [۱۵، ۳]. در اینجا به معرفی محورگیری جزئی می‌پردازیم. در روش محورگیری جزئی، در هرگام از روش حذفی گاوس از بین درایه‌های محوری و درایه‌های زیر آن و در همان ستون، بزرگ‌ترین درایه از لحاظ قدرمطلق را یافته و سطر مربوطه را با سطر درایه‌های محوری تعویض می‌کنیم. توجه کنید که اگر هیچ درایه‌ی ناصفیری پیدا نشود، ماتریس ضرایب منفرد است.

مثال ۹.۱ دستگاه معادلات خطی

$$\begin{cases} 2x_1 + x_2 = 3, \\ 3x_1 + 2x_2 + 2x_3 = 9, \\ x_1 + x_2 + 5x_3 = 8, \end{cases}$$

را با روش حذفی گاوس با محورگیری جزئی حل کنید.

حل: ماتریس افزوده

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & 3 \\ 3 & 2 & 2 & 9 \\ 1 & 1 & 5 & 8 \end{array} \right),$$

را تشکیل می‌دهیم. بزرگ‌ترین درایه از لحاظ قدرمطلق در ستون اول ماتریس، درایه (۲, ۱) است. بنابراین، سطر دوم و اول ماتریس افزوده را جایگزین یکدیگر می‌کنیم. در نتیجه، ماتریس افزوده به صورت زیر تبدیل می‌شود

$$\left(\begin{array}{ccc|c} 3 & 2 & 2 & 9 \\ 2 & 0 & 1 & 3 \\ 1 & 1 & 5 & 8 \end{array} \right).$$

با حذف درایه‌های زیر درایه (۱, ۱)، داریم

$$\left(\begin{array}{ccc|c} 3 & 2 & 2 & 9 \\ 0 & -\frac{4}{3} & -\frac{1}{3} & -3 \\ 0 & \frac{1}{3} & \frac{13}{3} & 5 \end{array} \right).$$

اکنون، از میان درایه‌های $(۲, ۲)$ و $(۳, ۲)$ ، درایه $(۲, ۲)$ از لحاظ قدرمطلق بزرگ‌تراست و بنابراین، نیازی به تعویض سطر نیست. با صفرکردن درایه $(۳, ۲)$ ، خواهیم داشت

$$\left(\begin{array}{ccc|c} ۳ & ۲ & ۲ & ۹ \\ ۰ & -\frac{۴}{۳} & -\frac{۱}{۳} & -۳ \\ ۰ & ۰ & \frac{۱۷}{۴} & \frac{۱۷}{۴} \end{array} \right).$$

اینک، از جای‌گذاری پس‌رواستفاده می‌کنیم. معادله سوم به صورت

$$\frac{۱۷}{۴}x_3 = \frac{۱۷}{۴},$$

است که با حل آن داریم $x_3 = ۱$. با جای‌گذاری $x_3 = ۱$ در معادله دوم، یعنی

$$-\frac{۴}{۳}x_2 - \frac{۱}{۳}x_3 = -۳,$$

مقدار $x_2 = ۲$ به دست می‌آید. سرانجام با جای‌گذاری x_2 و x_3 در معادله اول، یعنی

$$۳x_1 + ۲x_2 + ۲x_3 = ۹,$$

خواهیم داشت $x_1 = ۱$.

۳.۱ پیش‌نیازهایی از حساب و دیفرانسیل

در این بخش، برخی از مفاهیم و قضایای حساب و دیفرانسیل را مرور می‌کنیم.

قضیه ۲.۱ (قضیه مقدار میانی) فرض کنید تابع f در بازه $[a, b]$ پیوسته است و

$$\min_{x \in [a, b]} f(x) \leq r \leq \max_{x \in [a, b]} f(x).$$

در این صورت، یک $\xi \in [a, b]$ وجود دارد به طوری که $f(\xi) = r$.

قضیه ۳.۱ (قضیه مقدار میانگین) فرض کنید f و g دو تابع حقیقی مقدار پیوسته روی بازه $[a, b]$ هستند، به طوری که $g \geq ۰$. در این صورت، یک نقطه مثل ξ در بازه $[a, b]$ وجود دارد به طوری که

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

در حالت خاص، اگر $g(x) = 1$ ، آنگاه قضیه مقدار میانگین به صورت

$$\int_a^b f(x) dx = f(\xi)(b-a).$$

خواهد بود.

قضیه ۴.۱ (قضیه اساسی حساب و دیفرانسیل) فرض کنید تابع f در بازه $[a, b]$ پیوسته است. در این صورت

$$F(x) = \int_a^x f(t) dt, \quad x \in [a, x],$$

یک تابع مشتق‌پذیر است و به ازای هر $x \in [a, b]$ ، $F'(x) = f(x)$. تابع F را تابع اولیه f می‌گویند.

قضیه ۵.۱ (سری تیلر) فرض کنید f یک تابع حقیقی روی بازه $[a, b]$ ، n یک عدد طبیعی، $f^{(n)}$ روی $[a, b]$ پیوسته و به ازای هر $t \in (a, b)$ مقدار $f^{(n+1)}(t)$ موجود است. همچنین، فرض کنید α و x دو عدد متمایز در بازه $[a, b]$ هستند. در این صورت، یک ξ بین α و x وجود دارد به طوری که

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(\alpha)}{k!} (x-\alpha)^k + E_n(x),$$

که در آن

$$E_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-\alpha)^{n+1}.$$

با استفاده از قضیه ۵.۱، اگر بازه $[a, b]$ شامل نقاط x و $x+h$ باشد، آنگاه خواهیم داشت

$$f(x+h) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} h^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}. \quad (5.1)$$

که در آن ξ عددی بین x و $x+h$ است. قضیه بعد، صورت انتگرالی مانده در بسط تیلر را ارائه می‌کند [۴].

قضیه ۶.۱ اگر مشتقات تابع f تا مرتبه $n+1$ تابع f پیوسته باشند، آنگاه مانده $E_n(x)$ در قضیه ۵.۱ را می‌توان به صورت

$$E_n(x) = \frac{1}{n!} \int_{\alpha}^x f^{(n+1)}(t) (x-t)^n dt,$$

نمایش داد.

برهان: با فرض

$$u = \frac{1}{n!}(x-t)^n, \quad dv = f^{(n+1)}(t)dt,$$

با کمک فرمول انتگرال‌گیری جزء‌به‌جزء، خواهیم داشت

$$\begin{aligned} E_n(x) &= \int_{\alpha}^x u dv = uv|_{\alpha}^x - \int_{\alpha}^x v du \\ &= \frac{1}{n!} \left[f^n(t)(x-t)^n \Big|_{t=\alpha}^{t=x} + n \int_{\alpha}^x f^{(n)}(t)(x-t)^{n-1} dt \right] \\ &= -\frac{1}{n!} f^{(n)}(\alpha)(x-\alpha)^n + E_{n-1}(x). \end{aligned}$$

با ادامه این روند خواهیم داشت

$$\begin{aligned} E_n(x) &= -\sum_{k=1}^n \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + E_0(x) \\ &= -\sum_{k=1}^n \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + \int_{\alpha}^x f'(t)dt \\ &= -\sum_{k=1}^n \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + f(x) - f(\alpha). \end{aligned}$$

ازاین‌رو،

$$f(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + E_n(x).$$

تعریف ۷.۱ توابع $f, g : \mathbb{R} \rightarrow \mathbb{R}$ را در نظر بگیرید. برای هر $L \in \mathbb{R} \cup \{\pm\infty\}$ ، می‌نویسیم $f(x) = \mathcal{O}_{x \rightarrow L}(g(x))$ و می‌خوانیم «وقتی که x به L میل می‌کند، تابع f از مرتبه g است»، اگر یک ثابت نامنفی مثل C وجود داشته باشد به طوری که به ازای هر x در یک همسایگی از L داشته باشیم $|f(x)| \leq C|g(x)|$. اگر بیم خطا نرود، گاهی برای سادگی از نماد $\mathcal{O}(\cdot)$ برای $\mathcal{O}_{x \rightarrow L}(\cdot)$ استفاده می‌شود.

مثال ۱۰.۱ تابع $f(t) = \cos t$ را در نظر بگیرید. داریم

$$f'(t) = -\sin t, \quad f''(t) = -\cos t, \quad f'''(t) = \sin t, \quad f^{(4)}(t) = \cos t.$$

در رابطه (۵.۱)، فرض کنید $x = 0$. در این صورت، خواهیم داشت

$$\begin{aligned} f(0+h) &= f(0) + f'(0)h + \frac{f''(0)}{2!}h^2 + \frac{f'''(0)}{3!}h^3 + \frac{f^{(4)}(\xi)}{4!}h^4 \\ &= 1 - \frac{h^2}{2} + \frac{h^4}{24} \cos \xi, \end{aligned}$$

که در آن $|x| < \xi$. بنابراین

$$\cos h - 1 + \frac{h^2}{2} = \left| \frac{h^4}{24} \right| |\cos \xi| \leq \frac{1}{24} |h^4|.$$

در نتیجه، داریم

$$\cos h - 1 + \frac{h^2}{2} = \mathcal{O}_{h \rightarrow 0}(h^4).$$

مثال ۱۱.۱ فرض کنید $h > 0$. اگر $x, x \pm h \in [a, b]$ ، آنگاه با استفاده از قضیه ۵.۱ به‌ازای

$n = 3$ داریم

$$f(x-h) = f(x_0) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{6} f^{(3)}(x) + \frac{h^4}{24} f^{(4)}(\xi_2),$$

$$f(x+h) = f(x_0) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f^{(3)}(x) + \frac{h^4}{24} f^{(4)}(\xi_1),$$

که در آن، ξ_1 بین x و $x+h$ و ξ_2 بین $x-h$ و x است. با جمع کردن طرفین رابطهٔ اخیر، خواهیم داشت

$$f(x-h) + f(x+h) = 2f(x) + h^2 f''(x) + \frac{h^4}{24} (f^{(4)}(\xi_1) + f^{(4)}(\xi_2)).$$

در نتیجه

$$f''(x) = \frac{f(x-h) - 2f(x) + f(x+h)}{h^2} + E,$$

که در آن

$$E = -\frac{h^4}{24} (f^{(4)}(\xi_1) + f^{(4)}(\xi_2)).$$

از طرفی، داریم

$$\min_{\xi_1 \leq x \leq \xi_2} f^{(4)}(x) \leq \frac{1}{2} (f^{(4)}(\xi_1) + f^{(4)}(\xi_2)) \leq \max_{\xi_1 \leq x \leq \xi_2} f^{(4)}(x).$$

در نتیجه، با استفاده از قضیهٔ مقدار میانی، یک ξ در بازهٔ $[\xi_1, \xi_2]$ وجود دارد به‌طوری‌که

$$f^{(4)}(\xi) = \frac{1}{2} (f^{(4)}(\xi_1) + f^{(4)}(\xi_2)).$$

از این‌رو،

$$E = -\frac{h^4}{24} f^{(4)}(\xi) = \mathcal{O}(h^4).$$

فرمول

$$f''(x) \approx \frac{f(x-h) - 2f(x) + f(x+h)}{h^2}, \quad (6.1)$$

به فرمول تفاضل مرکزی برای مشتق دوم f معروف است.

قضیه ۷.۱ (بسط تیلر برای توابع برداری) فرض کنید $f = (f_1, f_2, \dots, f_m)^T$ که در آن f_i ، $i = 1, 2, \dots, m$ توابعی از \mathbb{R}^n به \mathbb{R} هستند. اگر مشتقات مرتبه اول و دوم تابع f پیوسته باشند، آنگاه به ازای هر $x, p \in \mathbb{R}^n$ داریم

$$f(x+p) = f(x) + J(x)p + \mathcal{O}(\|p\|^2),$$

که در آن $J(x)$ ماتریس ژاکوبی است که به صورت

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

تعریف می‌شود که $x = (x_1, x_2, \dots, x_n)^T$. بنابراین، اگر $p = (p_1, p_2, \dots, p_n)^T$ آنگاه

$$f_i(x+p) = f_i(x) + \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} p_j + \mathcal{O}(\|p\|^2), \quad i = 1, 2, \dots, m.$$

قضیه ۸.۱ (بسط تیلر توابع چندمتغیره) ([۶]) فرض کنید $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع چندمتغیره است به طوری که مشتقات آن تا مرتبه سوم پیوسته هستند. اگر $x = (x_1, x_2, \dots, x_n)^T$ و آنگاه $p = (p_1, p_2, \dots, p_n)^T$

$$\varphi(x+p) = \varphi(x) + \nabla\varphi(x)^T p + \frac{1}{2} p^T \nabla^2\varphi(x) p + \mathcal{O}(\|p\|^3), \quad (7.1)$$

که در آن $\nabla\varphi(x)$ بردار گرادیان و $\nabla^2\varphi(x)$ ماتریس هسیان هستند که به صورت

$$\nabla\varphi(x) = \begin{pmatrix} \frac{\partial\varphi}{\partial x_1} \\ \frac{\partial\varphi}{\partial x_2} \\ \vdots \\ \frac{\partial\varphi}{\partial x_n} \end{pmatrix}, \quad \nabla^2\varphi(x) = \begin{pmatrix} \frac{\partial^2\varphi}{\partial x_1^2} & \frac{\partial^2\varphi}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2\varphi}{\partial x_1\partial x_n} \\ \frac{\partial^2\varphi}{\partial x_2\partial x_1} & \frac{\partial^2\varphi}{\partial x_2^2} & \cdots & \frac{\partial^2\varphi}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\varphi}{\partial x_n\partial x_1} & \frac{\partial^2\varphi}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2\varphi}{\partial x_n^2} \end{pmatrix},$$

تعریف می‌شوند.

مثال ۱۲.۱ با کمک بسط تیلر، یک تقریب برای $\sqrt{1.02^2 + 1.97^2}$ محاسبه می‌کنیم. فرض کنید $\varphi(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$ و

$$\tilde{x} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \tilde{p} = \begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \end{pmatrix} = \begin{pmatrix} 0.02 \\ -0.03 \end{pmatrix}.$$

اکنون، با کمک قضیه ۸.۱، مقدار $\varphi(\tilde{x} + \tilde{p})$ را با استفاده از سه جمله اول بسط داده‌شده در رابطه (۷.۱) محاسبه می‌کنیم. داریم

$$\begin{aligned} \varphi(\tilde{x}) &= 3, \\ \nabla\varphi(\tilde{x}) &= \begin{pmatrix} \frac{\partial\varphi}{\partial x_1}(\tilde{x}) \\ \frac{\partial\varphi}{\partial x_2}(\tilde{x}) \end{pmatrix} = \begin{pmatrix} \frac{\tilde{x}_1}{\sqrt{\tilde{x}_1^2 + \tilde{x}_2^2}} \\ \frac{2\tilde{x}_2}{2\sqrt{\tilde{x}_1^2 + \tilde{x}_2^2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 2 \end{pmatrix}, \\ \nabla^2\varphi(\tilde{x}) &= \begin{pmatrix} \frac{\partial^2\varphi}{\partial x_1^2}(\tilde{x}) & \frac{\partial^2\varphi}{\partial x_1\partial x_2}(\tilde{x}) \\ \frac{\partial^2\varphi}{\partial x_2\partial x_1}(\tilde{x}) & \frac{\partial^2\varphi}{\partial x_2^2}(\tilde{x}) \end{pmatrix} = \begin{pmatrix} \frac{\tilde{x}_2^2}{(\tilde{x}_1^2 + \tilde{x}_2^2)^{\frac{3}{2}}} & \frac{-2\tilde{x}_1\tilde{x}_2}{2(\tilde{x}_1^2 + \tilde{x}_2^2)^{\frac{3}{2}}} \\ \frac{-2\tilde{x}_1\tilde{x}_2}{2(\tilde{x}_1^2 + \tilde{x}_2^2)^{\frac{3}{2}}} & \frac{12\tilde{x}_1^2\tilde{x}_2 + 2\tilde{x}_2^3}{4(\tilde{x}_1^2 + \tilde{x}_2^2)^{\frac{3}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{pmatrix}. \end{aligned}$$

بنابراین

$$\begin{aligned} \varphi(\tilde{x} + \tilde{p}) &\approx \varphi(\tilde{x}) + \nabla\varphi(\tilde{x})^T \tilde{p} + \frac{1}{2} \tilde{p}^T \nabla^2\varphi(\tilde{x}) \tilde{p} \\ &= 3 + \begin{pmatrix} \frac{1}{\sqrt{3}} & 2 \end{pmatrix} \begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \tilde{p}_1 & \tilde{p}_2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \end{pmatrix} \\ &= 3 + \frac{1}{\sqrt{3}}\tilde{p}_1 + 2\tilde{p}_2 + \frac{1}{2} \left(\frac{1}{\sqrt{3}}\tilde{p}_1^2 - \frac{4}{3}\tilde{p}_1\tilde{p}_2 + \frac{2}{3}\tilde{p}_2^2 \right). \end{aligned} \quad (8.1)$$

از این رو

$$\begin{aligned} \sqrt{1.02^2 + 1.97^2} &= \varphi(\tilde{x}_1 + \tilde{p}_1, \tilde{x}_2 + \tilde{p}_2) \\ &\approx 3 + \frac{1}{\sqrt{3}}\tilde{p}_1 + 2\tilde{p}_2 + \frac{1}{2} \left(\frac{1}{\sqrt{3}}\tilde{p}_1^2 - \frac{4}{3}\tilde{p}_1\tilde{p}_2 + \frac{2}{3}\tilde{p}_2^2 \right) \\ &= 3 + \frac{1}{\sqrt{3}} \times 0.02 + 2 \times (-0.03) + \frac{1}{2} \left(\frac{1}{\sqrt{3}} \times (0.02)^2 - \frac{4}{3} \times (0.02 \times (-0.03)) + \frac{2}{3} \times (-0.03)^2 \right) \\ &= 2.9471593. \end{aligned}$$

لازم به ذکر است که مقدار دقیق تابع برابر با $2.9471636\dots$ است. ملاحظه می‌کنیم که تقریب به دست آمده مناسب است.

قضیه ۹.۱ ([۶]) فرض کنید $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع چندمتغیره باشد به طوری که مشتقات آن تا مرتبه دوم پیوسته هستند و $x^* \in \mathbb{R}^n$.
الف) اگر $\nabla f(x^*) = 0$ و $\nabla^2\varphi(x^*)$ معین مثبت متقارن باشد، آنگاه x^* یک نقطه مینیمم موضعی

تابع φ است.

(ب) اگر x^* یک نقطهٔ مینیمم موضعی تابع φ باشد، آنگاه $\nabla f(x^*) = 0$ و $\nabla^2 \varphi(x^*)$ نیمه‌معین مثبت متقارن است.

مثال ۱۳.۱ فرض کنید

$$\varphi(x_1, x_2) = 2x_1^2 - x_1x_2 + x_2^2 - 11x_1 + x_2 + 5.$$

داریم

$$\nabla \varphi = \begin{pmatrix} \frac{\partial \varphi}{\partial x_1} \\ \frac{\partial \varphi}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 4x_1 - x_2 - 11 \\ -x_1 + 2x_2 + 1 \end{pmatrix}.$$

با حل معادلهٔ $\nabla f(x_1, x_2) = 0$ به جواب یکتای $(x_1^*, x_2^*) = (3, 1)$ می‌رسیم. از طرفی، داریم

$$\nabla^2 \varphi = \begin{pmatrix} \frac{\partial^2 \varphi}{\partial x_1^2} & \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \varphi}{\partial x_2 \partial x_1} & \frac{\partial^2 \varphi}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix}.$$

ماتریس به‌دست آمده، معین مثبت متقارن است. بنابراین، نقطهٔ $(x_1, x_2) = (3, 1)$ یک نقطهٔ مینیمم تابع φ است.

تمرین

(۱.۱) نشان دهید که به‌ازای هر $\alpha \in \mathbb{R}$ ، ماتریس

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 5 & 2\alpha - 1 \\ 1 & 2\alpha - 1 & \alpha^2 + 2 \end{pmatrix},$$

معین مثبت متقارن است.

(۲.۱) فرض کنید $n > 1$ و $a, b \in \mathbb{R}$ ، به‌طوری که $a > b \geq 0$. نشان دهید که ماتریس

$$A = \begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{pmatrix} \in \mathbb{R}^{n \times n},$$

معین مثبت متقارن است. راهنمایی: ماتریس A را به‌صورت $A = (b - a)I + bee^T$ بنویسید که در آن I ماتریس همانی است و $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$.

۳.۱ فرض کنید تابع f به اندازه کافی هموار باشد. عبارت خطا برای تقریب

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h},$$

را بیابید.

۴.۱ فرض کنید $A \in \mathbb{R}^{n \times n}$ و $b \in \mathbb{R}^n$. تابع $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت

$$\varphi(x) = \frac{1}{2}x^T A x - x^T b,$$

تعریف می‌کنیم. نشان دهید بردار گرادینان و هسیان تابع φ به ترتیب برابر با $\nabla \varphi(x) = Ax - b$ و $\nabla^2 \varphi(x) = A$ هستند.

۵.۱ فرض کنید $A \in \mathbb{R}^{m \times n}$ و $b \in \mathbb{R}^m$ ، که در آن $m \geq n$ و $\text{rank}(A) = n$. تابع $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ را به صورت $\varphi(x) = \|b - Ax\|_2^2$ تعریف می‌کنیم. نشان دهید x^* یک مینیمم سراسری برای φ است اگر و تنها اگر $A^T A x^* = A^T b$.

۶.۱ فرض کنید $f(x_1, x_2) = e^{x_1 x_2} (x_1 - x_2)$. سه جمله اول بسط تیلر برای تابع f را در نقطه $(1, 1)$ بنویسید.

فصل ۲

خطاها

برای بررسی یک پدیده، معمولاً یک مدل برای آن نوشته می‌شود. سپس، این مدل را با استفاده از روش‌های مختلف ریاضی حل می‌کنند. با توجه به اینکه بسیاری از این مدل‌ها را نمی‌توان به‌طور دقیق حل کرد، یک روش عددی برای حل آن انتخاب می‌شود. به‌کارگیری یک روش عددی برای حل یک مدل واقعی و با دقت کافی بدون استفاده از یک رایانه تقریباً غیرممکن است. با توجه به اینکه نمایش رقم‌های قسمت کسری یک عدد حقیقی در رایانه با محدودیت همراه است، تقریباً تمام اعداد با خطا در رایانه ذخیره می‌شوند. بنابراین، این خطاها ممکن است بر نتایج به‌دست آمده تأثیر قابل توجهی بگذارد. در این بخش، به بررسی این موضوع می‌پردازیم.

۱.۲ اعداد ممیز شناور در دستگاه IEEE

هر عدد حقیقی ناصفر را می‌توان به‌صورت

$$x = \pm q \times 2^m,$$

نمایش داد که در آن $1 \leq q < 2$ و m یک عدد صحیح است. در اینجا مقدار q را ماتیس عدد m را نمای عدد گویند. توجه کنید که در دستگاه دودویی، نمایش q به‌صورت

$$q = (1.b_1b_2b_3\dots)_2 = 1 + b_12^{-1} + b_22^{-2} + b_32^{-3} + \dots,$$

است، که در آن b_i ها صفر یا یک هستند. برای مثال

$$۱۳.۱۲۵ = (۱۱۰۱.۰۰۰۱)_۲ = (۱.۱۰۱۰۰۰۱)_۲ \times ۲^۳,$$

و

$$-\frac{۵۷}{۳۰} = -۲.۸۵ = -(۱۰.۱۱۰\overline{۱۱۰۰})_۲ = -(۱.۰۱۱۰\overline{۱۱۰۰})_۲ \times ۲^۱.$$

فرض کنید عدد حقیقی x به صورت

$$x = \pm q \times ۲^m = \pm (۱.b_۱b_۲ \dots b_{۲۳})_۲ \times ۲^m, \quad (۱.۲)$$

باشد (در ادامه، در صورتی که رقم‌های بیست و چهارم به بعد مانتیس صفر نباشند، بحث خواهد شد). در این بخش، شیوه ذخیره‌سازی یک عدد حقیقی مثل (۱.۲) را با دقت ساده در دستگاه IEEE^۱ بررسی می‌کنیم. هر عدد حقیقی مثل (۱.۲) با دقت ساده در ۳۲ بیت ذخیره می‌شود و برای ذخیره‌سازی آن در رایانه نیاز به ذخیره‌سازی سه بخش مختلف آن است: علامت، مانتیس و نمای عدد. اختصاص بیت‌ها برای قسمت‌های مختلف یک عدد حقیقی به صورت زیر است

۱	۸	۲۳
s	e	f
علامت	نما به صورت افزودنی	قسمت کسری مانتیس

این جدول نشان می‌دهد که یک بیت برای ذخیره‌سازی علامت استفاده می‌شود (اگر $s = ۰$ ، یعنی، عدد مثبت و در غیر این صورت عدد منفی است). برای ذخیره‌سازی قسمت مانتیس عدد از ۲۳ بیت استفاده می‌شود و با توجه به اینکه بیت قبل از اعشار همیشه یک است، این ۲۳ بیت برای رقم ابتدایی بعد از اعشار اختصاص داده می‌شود، یعنی، داریم $f = b_۱b_۲ \dots b_{۲۳}$. همان‌گونه که جدول بالا نشان می‌دهد، هشت بیت برای نما در نظر گرفته شده است. در این بیت‌ها، $۲^۸ = ۲۵۶$ تا ۲۵۵ را می‌توان ذخیره کرد (در مبنای ۲). اما با توجه به اینکه نمای اعداد می‌تواند منفی باشد، معقول به نظر می‌رسد که محدوده نما به صورت $۱۲۸ \leq m \leq -۱۲۷$ در نظر گرفته شود. برای ذخیره‌سازی این اعداد از شیوه افزودنی استفاده می‌شود؛ به این ترتیب که یک عدد ثابت به نام بایاس به مقدار واقعی نما اضافه می‌شود و سپس این عدد در مبنای دو نوشته شده و در هشت بیت مربوطه ذخیره می‌شود. در دقت ساده، این مقدار ثابت برابر با ۱۲۷ است. در این صورت، با توجه به اینکه داریم $۲۵۵ \leq m + ۱۲۷ \leq ۰$ ، عدد $e = m + ۱۲۷$ در هشت بیت

مورد نظر ذخیره می‌شود. همان‌گونه که خواهیم دید، شیوهٔ ذخیره‌سازی اعداد، زمانی که $m = 128$ و $m = -127$ با سایرین متفاوت است. در این صورت، در دقت ساده اعداد نرمالیزه به صورت

$$x = (-1)^s (1.f)_2 \times 2^m, \quad -126 \leq m \leq 127,$$

نمایش داده می‌شوند. تعداد اعداد مثبتی که به این صورت نمایش داده می‌شوند برابر با

$$254 \times 2^{23} = 2130706432 \approx \text{دو میلیارد},$$

است. تعداد این اعداد را با تعداد اعداد مجموعهٔ اعداد حقیقی مقایسه کنید!

مثال ۱.۲ عدد 11.703125 را در نظر بگیرید. داریم

$$11.703125 = (1011.101101)_2 = (1.011101101)_2 \times 2^3.$$

در این صورت، $f = 01110110100000000000000000$ ، $m = 3$ و $s = 0$. از طرفی، داریم

$$e = m + 127 = 130 = (10000010)_2.$$

بنابراین، نمایش (sef) به صورت زیر است

o	۱۰۰۰۰۰۱۰	۰۱۱۱۰۱۱۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰
s	e	f

تعریف ۱.۲ یک عدد حقیقی را عدد ممیز شناور می‌گویند، هرگاه بتوان آن را به صورت دقیق در (sef) ذخیره کرد.

می‌دانیم که مجموعهٔ اعداد حقیقی، چگال است، ولی در ادامه می‌بینیم که مجموعه اعداد ممیز شناور چگال نیست. فرض کنید $x = (1.f)_2 \times 2^m$ یک عدد ممیز شناور است. در این صورت، عدد ممیز شناور بعدی برابر است با

$$x' = \left((1.f)_2 + 2^{-23} \right) \times 2^m = x + 2^{m-23},$$

بنابراین، فاصلهٔ بین x و x' برابر با 2^{m-23} است و بین این دو عدد هیچ عدد ممیز شناوری وجود ندارد.

روشن است که کوچک‌ترین عدد ممیز شناور نرمالیزهٔ مثبت برابر با

$$x = (1.000\dots 0)_2 \times 2^{-126} = 2^{-126},$$

و بزرگ‌ترین عدد ممیز شناور نرمالیزه (مثبت) برابر با

$$x = (1.11\dots 1)_2 \times 2^{-127} = 2^{128} - 2^{104} = 3.4028 \times 10^{38},$$

است.

در ادامه، خواهیم دید که در دستگاه نمایش ارائه‌شده عدد صفر و $\pm\infty$ چگونه نمایش داده می‌شوند.

• اگر $m = 128$ (یعنی، $(e = (11111111)_2)$ و $f = 0$)، آنگاه عدد

$$x = (-1)^s (1.f)_2 \times 2^{128},$$

نمایشی برای $\pm\infty$ خواهد بود. اگر $s = 0$ ، آنگاه نمایشی برای $+\infty$ و در غیر این صورت، نمایشی برای $-\infty$ خواهد بود. در صورتی که $m = 128$ و $f \neq 0$ ، آنگاه

$$x = (-1)^s (1.f)_2 \times 2^{128},$$

هیچ عددی را مشخص نمی‌کند و با «NaN»^۱ نمایش داده می‌شود.

• اگر $m = -127$ ، آنگاه عدد به صورت

$$x = (-1)^s (0.f)_2 \times 2^{-126},$$

ذخیره می‌شود. در صورتی که $f = 0$ ، آنگاه $x = \pm 0$ و در این صورت، اگر $s = 0$ ، آنگاه داریم $x = +0$ و در غیر این صورت، $x = -0$. بنابراین، برای عدد صفر دو نمایش مختلف وجود دارد. اگر $f \neq 0$ ، آنگاه این اعداد را اعداد زیرنرمال گویند.

یادداشت ۱.۲ با توجه به تعریف ۱.۲، اعداد ممیز شناور شامل دو نمایش برای عدد صفر، اعداد زیرنرمال (منفی و مثبت)، اعداد نرمالیزه‌شده (منفی و مثبت) و نمایش‌های $+\infty$ و $-\infty$ هستند.

در دقت مضاعف IEEE، ۶۴ بیت برای ذخیره‌سازی یک عدد در رایانه در نظر گرفته می‌شود. در این صورت یک بیت برای علامت عدد، ۱۱ بیت برای نما به شیوهٔ افزودنی و ۵۲ بیت برای ذخیره‌سازی قسمت کسری مانتیس اختصاص داده می‌شود. همانند دقت ساده، تمام بحث‌های انجام‌شده را می‌توان برای دقت مضاعف هم انجام داد.

^۱ Not a Number

۲.۲ حساب IEEE

می‌خواهیم ببینیم که اگر یک عمل حسابی بین دو عدد ممیز شناور انجام شود، آنگاه چه چیزی در (sef) ذخیره می‌شود. در اینجا دقت ساده را مطالعه می‌کنیم و به‌طور مشابه می‌توان دقت مضاعف را بررسی کرد. نخست، تابع fl را از مجموعه \mathbb{R} به FP به صورت

$$fl : \mathbb{R} \rightarrow FP$$

$$x \rightarrow fl(x)$$

تعریف می‌کنیم که در آن $fl(x)$ نمایش ممیز شناور عدد x و FP بیانگر مجموعه اعداد ممیز شناور است. اگر اعداد ممیز شناور را به‌عنوان نقاطی روی محور اعداد حقیقی در نظر بگیریم، آنگاه ضابطه تابع fl به صورت زیر تعریف می‌شود

(الف) اگر $|x|$ از بزرگ‌ترین عدد ممیز شناور نرمالیزه شده بزرگ‌تر باشد، آنگاه در صورتی که $x > 0$ داریم $fl(x) = +\infty$ ، و در غیر این صورت، $fl(x) = -\infty$.
 (ب) اگر x بین دو عدد ممیز شناور (نرمالیزه و زیرنرمال) قرار گیرد، آنگاه $fl(x)$ برابر با عدد ممیز شناوری خواهد بود که به x نزدیک‌تر است.
 دقت ساده IEEE را در نظر بگیرید. فرض کنید عدد مثبت x بین دو عدد ممیز شناور نرمالیزه شده باشد، یعنی،

$$q \times 2^m < x < (q + 2^{-23}) \times 2^m,$$

که در آن $1 \leq q < 2$. در این صورت، داریم $x = q \times 2^m + \delta$ ، که در آن $0 < \delta < 2^{m-23}$. اکنون، اگر $\delta < \frac{1}{4} \times 2^{m-23}$ ، آنگاه

$$fl(x) = q \times 2^m,$$

و اگر $\delta > \frac{1}{4} \times 2^{m-23}$ ، آنگاه

$$fl(x) = (q + 2^{-23}) \times 2^m.$$

همچنین، حالتی را در نظر بگیرید که در آن $\delta = \frac{1}{4} \times 2^{m-23}$. در این صورت

$$x = (1.b_1 b_2 \dots b_{23} 1)_2 \times 2^m.$$

اینک، اگر $b_{23} = 0$ ، آنگاه

$$fl(x) = q \times 2^m,$$

و در صورتی که $b_{23} = 1$ ، خواهیم داشت

$$fl(x) = (q + 2^{-23}) \times 2^m.$$

کوچک‌ترین عدد زیرنرمال مثبت برابر با

$$x = (0.000001)_2 \times 2^{-126} = 2^{-23} \times 2^{-126} = 2^{-149},$$

است. اکنون، داریم

$$- \text{اگر } x > 0 \text{ و } 0 < |x| < 2^{-149}, \text{ آنگاه } fl(x) = +0.$$

$$- \text{اگر } x < 0 \text{ و } 0 < |x| < 2^{-149}, \text{ آنگاه } fl(x) = -0.$$

$$\text{به علاوه، اگر } x \in FP, \text{ آنگاه } fl(x) = x.$$

در ادامه، با توجه به مطالب ارائه شده، به بررسی اعمال حسابی در مجموعه اعداد ممیز شناور می‌پردازیم. فرض کنید “ \odot ” یک عمل در مجموعه اعداد حقیقی باشد. در این صورت، عمل متناظر در مجموعه اعداد ممیز شناور را با “ \odot ” نمایش می‌دهیم. برای دو عدد ممیز شناور x و y قرار می‌دهیم

$$x \odot y = fl(x \cdot y).$$

ولی این تعریف نیاز به تعمیم دارد، زیرا ممکن است x یا y بینهایت یا NaN باشد. این تعمیم به صورت زیر انجام می‌شود.

- اگر x یا y NaN باشد، آنگاه $x \odot y$ نیز NaN خواهد بود. توجه کنید که NaN به این معنا است که در بخش نمای (sef)، همه رقم‌ها برابر یا یک هستند و $f \neq 0$.

- داریم $\lim_{x,y \rightarrow +\infty} (x+y) = +\infty$ ، بنابراین، $(+\infty) \oplus (+\infty) = +\infty$. همچنین، $\lim_{x,y \rightarrow +\infty} (x-y)$ موجود نیست، لذا $(+\infty) \ominus (+\infty) = \text{NaN}$.

- تفاضل دو عدد بسیار کوچک منفی، عدد بسیار کوچکی است، اما ممکن است منفی یا مثبت باشد. بنابراین، $(-0) \ominus (-0)$ برابر با -0 یا $+0$ خواهد بود. به همین ترتیب، $(+0) \oplus (-0)$ نیز برابر با $+0$ یا -0 خواهد بود.

در دقت ساده IEEE، نخستین عدد ممیز شناور بزرگ‌تر از یک برابر با

$$(1.000001)_2 \times 2^0 = 1 + 2^{-23},$$

است. نقطه وسط ۱ و $1 + 2^{-23}$ ، عدد

$$1 + \frac{1}{4} \times 2^{-23} = 1 + 2^{-24},$$

است. اینک، اگر $0 \leq x \leq 2^{-24}$ ، آنگاه

$$1 \oplus x = 1,$$

و اگر $2^{-24} < x \leq 2^{-23}$ ، آنگاه

$$1 \oplus x = 1 + 2^{-23}.$$

تعریف ۲.۲ اپسیلون ماشین را با ϵ_{mach} نشان می‌دهیم و بنا به تعریف کوچک‌ترین عدد ممیز شناوری است که $1 \oplus \epsilon_{mach} > 1$. همچنین، خطای روند عدد یک را با u نشان می‌دهیم و بنا به تعریف، بزرگ‌ترین عدد ممیز شناوری است که $1 \oplus u = 1$.

با توجه به مطالبی که قبل از این تعریف آمده است، به روشنی، داریم $u = 2^{-24}$ و همچنین ϵ_{mach} نخستین عدد ممیز شناوری است که پس از 2^{-24} قرار دارد. به این ترتیب، داریم

$$1.0 \dots 01 \times 2^{-24} = (1 + 2^{-23})2^{-24} \approx 5.9605 \times 10^{-8}.$$

بعضی از خواص مجموعه اعداد حقیقی ممکن است در مجموعه اعداد ممیز شناور برقرار نباشد. برای مثال، داریم

$$(1 \oplus 2^{-24}) \oplus 2^{-24} = 1 \oplus 2^{-24} = 1,$$

$$1 \oplus (2^{-24} \oplus 2^{-24}) = 1 \oplus 2^{-23} = 1 + 2^{-23}.$$

این نتیجه نشان می‌دهد که مجموعه اعداد ممیز شناور، خاصیت شرکت‌پذیری ندارد.

۳.۲ آنالیز خطا

فرض کنید عدد حقیقی x با عدد حقیقی دیگری مثل x^* تقریب زده شود. در این صورت

(۱) مقدار $x - x^*$ را خطا می‌گوییم؛

(۲) مقدار $|x^* - x|$ را خطای مطلق می‌گوییم؛

(۳) مقدار $\frac{|x^* - x|}{|x|}$ را خطای نسبی می‌گوییم (با فرض اینکه $x \neq 0$).

توجه کنید که در مسائل عملی، خطای نسبی نسبت به دو خطای دیگر از اهمیت بیشتری برخوردار است؛ زیرا بزرگی خطا را بهتر توصیف می‌کند. اگر قرار دهیم

$$\delta = \frac{x^* - x}{x},$$

آنگاه داریم $x^* = x(1 + \delta)$ ، که در آن $|\delta|$ خطای نسبی است. نخست، خطای نسبی تولیدشده در یک گردکردن ساده را بررسی می‌کنیم. فرض می‌کنیم $x > 0$ (در حالت $x < 0$ به‌طور مشابه عمل می‌شود) و دقت ساده را در نظر می‌گیریم. به‌علاوه، فرض کنید که x بین دو عدد ممیز شناور متوالی $q \times 2^m$ و $(q + 2^{-23}) \times 2^m$ باشد، به‌طوری که $1 \leq q < 2$ ، یعنی،

$$q \times 2^m \leq x < (q + 2^{-23}) \times 2^m.$$

در این صورت، خواهیم داشت

$$|x - fl(x)| \leq \frac{1}{4} \left((q + 2^{-23}) \times 2^m - q \times 2^m \right) = \frac{1}{4} 2^{-23} 2^m = 2^{m-24}.$$

بنابراین، خطای نسبی در رابطه

$$\left| \frac{x - fl(x)}{x} \right| \leq \frac{2^{m-24}}{q \times 2^m} \leq \frac{2^{m-24}}{2^m} = 2^{-24} = u,$$

صدق می‌کند. از این رو، داریم

$$fl(x) = x(1 + \delta), \quad |\delta| \leq u = 2^{-24}.$$

این رابطه نشان می‌دهد که حداکثر خطای نسبی تولیدشده در یک گردکردن در دقت ساده برابر با خطای روند عدد یک (u) است.

یادداشت ۲.۲ اگر $fl(x)$ یک عدد زیرنرمال باشد، آنگاه خطای مطلق حداکثر برابر با 2^{-150} خواهد بود؛ زیرا اگر $r = (0.f)_2 \times 2^{-126}$ یک عدد زیرنرمال باشد، آنگاه عدد زیرنرمال بعدی برابر با

$$s = ((0.f)_2 + 2^{-23}) \times 2^{-126} = (0.f)_2 \times 2^{-126} + 2^{-149} = r + 2^{-149},$$

خواهد بود. از این رو، اگر $r \leq x \leq s$ ، آنگاه

$$|x - fl(x)| \leq \frac{1}{4}(s - r) = 2^{-150}.$$

همان‌گونه که مشاهده می‌شود، حداکثر خطای مطلق عدد کوچکی است، ولی خطای نسبی می‌تواند بسیار بزرگ باشد. برای مثال، فرض کنید

$$x = 2^{-149} + 2^{-150} = (0.0 \dots 011)_2 \times 2^{-126}.$$

در این صورت، داریم $fl(x) = (0.0 \dots 01)_2 \times 2^{-126} = 2^{-148}$ و بنابراین،

$$\frac{|x - fl(x)|}{|x|} = \frac{2^{-148} - 2^{-149} - 2^{-150}}{2^{-149} + 2^{-150}} = \frac{1}{3},$$

و این یعنی، خطای نسبی تقریباً برابر با ۳۳٪ است.

تعریف ۳.۲ فرض کنید x و y دو عدد ممیزشناور باشند. می‌گوییم نتیجه عمل $x.y$ زیرریز خواهد بود، هرگاه مقدار دقیق آن ناصفر و قدرمطلق آن از کوچک‌ترین عدد نرمالیزه مثبت، کوچک‌تر باشد. به همین ترتیب، نتیجه عمل $x.y$ سرریز خواهد بود، هرگاه مقدار دقیق آن ناصفر و قدرمطلق آن از بزرگ‌ترین عدد نرمالیزه، بزرگ‌تر باشد.

اگر نتیجه یک ضرب زیرریز شود، آنگاه خطای نسبی می‌تواند بسیار بزرگ شود. برای مثال، داریم

$$\begin{aligned} x &= (1.1)_2 \times 2^{-74}, \\ y &= (1.0)_2 \times 2^{-75}, \\ x \times y &= (1.1)_2 \times 2^{-149} = 2^{-149} + 2^{-150}, \\ x \otimes y &= (0.0 \dots 010)_2 \times 2^{-126} = 2^{-148}. \end{aligned}$$

بنابراین، همان‌گونه که پیش از این دیدیم، داریم

$$\frac{|x \times y - x \otimes y|}{|x \times y|} = \frac{1}{3}.$$

در نتیجه، خطای نسبی تقریباً ۳۳٪ است.

اکنون فرض کنید x و y دو عدد حقیقی و $x.y$ و $fl(x).fl(y)$ در محدوده ماشین باشند (قدرمطلق آن بین کوچک‌ترین و بزرگ‌ترین عدد ممیزشناور نرمالیزه قرار داشته باشد). در این صورت، نتیجه‌ای که در رایانه ذخیره می‌شود را با $(x.y)^*$ نشان می‌دهیم و داریم

$$(x.y)^* = fl(fl(x).fl(y)) = fl(x) \odot fl(y).$$

ابتدا عمل ضرب را در نظر می‌گیریم. در این حالت، داریم

$$\begin{aligned}(x \times y)^* &= fl(x) \otimes fl(y) \\ &= x(1 + \delta_1) \otimes y(1 + \delta_2) \\ &= x(1 + \delta_1)y(1 + \delta_2)(1 + \delta_3) \\ &= x \times y(1 + \delta),\end{aligned}$$

که در آن $|\delta_i| \leq u$ و

$$\delta = \delta_1 + \delta_2 + \delta_3 + \delta_1\delta_2 + \delta_1\delta_3 + \delta_2\delta_3 + \delta_1\delta_2\delta_3.$$

با توجه به اینکه $|\delta_i\delta_j| \leq u^2$ و $|\delta_1\delta_2\delta_3| \leq u^3$ و u^3 در مقایسه با u بسیار کوچک هستند، $|\delta| \leq 3u$ تقریب خوبی است. بنابراین، خطای نسبی در ضرب خیلی کوچک است، مگر اینکه یک زیرریز یا سرریز رخ دهد.

اینک، عمل جمع را در نظر می‌گیریم. در این حالت، داریم

$$\begin{aligned}(x + y)^* &= fl(x) \oplus fl(y) \\ &= x(1 + \delta_1) \oplus y(1 + \delta_2) \\ &= (x(1 + \delta_1) + y(1 + \delta_2))(1 + \delta_3) \\ &= (x + y)(1 + \delta_3) + (x\delta_1 + y\delta_2)(1 + \delta_3).\end{aligned}$$

بنابراین، خواهیم داشت

$$\left| \frac{(x + y) - (x + y)^*}{x + y} \right| \leq |\delta_3| + \frac{|x\delta_1 + y\delta_2|}{|x + y|} |1 + \delta_3|.$$

از طرفی، داریم

$$\frac{|x\delta_1 + y\delta_2|}{|x + y|} \leq \frac{|x||\delta_1| + |y||\delta_2|}{|x + y|} \leq \frac{|x| + |y|}{|x + y|} u.$$

اکنون، دو حالت مختلف را در نظر می‌گیریم. ابتدا فرض کنید x و y هم علامت هستند. در این صورت، داریم

$$|\delta_3| + \frac{|x\delta_1 + y\delta_2|}{|x + y|} |1 + \delta_3| \leq u + \frac{|x| + |y|}{|x + y|} u(1 + |\delta_3|) \leq u(2 + u) = 2u + u^2.$$

این رابطه نشان می‌دهد که $2u$ تقریب خوبی برای کران بالای خطای نسبی در جمع است. اینک، حالتی را در نظر بگیرید که در آن x و y مختلف‌العلامه باشند. در این صورت، خطای نسبی

می‌تواند بسیار بزرگ باشد. برای مثال، فرض کنید $y \in FP$ و $\alpha \neq 0$ یک عدد حقیقی است به طوری که

$$x = -y + \alpha, \quad fl(-y + \alpha) = -y.$$

در این صورت، خواهیم داشت

$$x + y = \alpha, \quad (x + y)^* = fl(x) \oplus fl(y) = (-y) \oplus y = 0.$$

در نتیجه،

$$\left| \frac{(x + y)^* - (x + y)}{x + y} \right| = 1,$$

که نشان می‌دهد خطای نسبی برابر با ۱۰۰٪ است.

به طور کلی، در تفریق دو عدد نزدیک به هم، دقت زیادی (رقم‌های با معنای زیادی) از دست می‌رود که به آن حذف فاجعه‌انگیز می‌گویند. برای درک بهترین موضوع، یک مثال ارائه می‌کنیم. فرض کنید $x = 3.14160$ و $y = 3.14158$ دو تقریب برای عدد $\pi = 3.14159265358979323846264338327950288419716939937510582097494459230781640628620899862803482534211706798214808651328230664709384460955058223172535940812813622691447110623$ هستند. اگر با استفاده از نرم‌افزار متلب و با دقت مضاعف مقدار $z = x - y$ را محاسبه کنیم، خواهیم دید که این مقدار برابر با $10^{-5} \times 13102$ می‌شود. مقدار به دست آمده نامعقول است؛ زیرا انتظار داشتیم که مقدار z برابر با $10^{-5} \times 2.0000000000000000$ باشد. ولی با توجه به اینکه هیچ کدام از دو عدد x و y را نمی‌توان به طور دقیق در `sef` ذخیره کرد، چنین اتفاقی رخ داده است. توجه کنید که رقم‌های عدد 13102 که در انتهای مانتیس عدد ظاهر شده‌اند، در تمام محاسبات بعدی به کار گرفته می‌شوند که باعث ایجاد خطای قابل توجهی می‌شوند.

مثال ۲.۲ توابع

$$f(x) = \frac{1 - \cos x}{\sin^2 x}, \quad g(x) = \frac{1}{1 + \cos x},$$

را در نظر بگیرید. در جدول زیر، مقادیر این دو تابع به ازای $x_i = 10^{-i}$ ، $i = 0, 1, \dots, 10$ داده شده است. انتظار می‌رود که مقادیر f و g به ازای این مقادیر برابر باشند، ولی می‌بینیم که به ازای 10^{-8} ، 10^{-9} و 10^{-10} مقادیر f و g بسیار متفاوت هستند. دلیل این اتفاق، این است که در صورت کسر f عبارت $1 - \cos x$ وجود دارد و به ازای مقادیر کوچک x ، مقادیر $\cos x$ به ۱ نزدیک هستند و در تفریق دو عدد نزدیک به هم، دقت زیادی از دست می‌رود.

x_i	$f(x_i)$	$g(x_i)$
۱.۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۶۴۹۲۲۳۲۰۵۲۰۴۷۶۲	۰.۶۴۹۲۲۳۲۰۵۲۰۴۷۶۲
۰.۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۱۲۵۲۰۸۶۲۸۸۵۷۷	۰.۵۰۱۲۵۲۰۸۶۲۸۸۵۷۱
۰.۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۱۲۵۰۰۲۰۸۴۸۱	۰.۵۰۰۰۱۲۵۰۰۲۰۸۳۳۶
۰.۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۱۲۴۹۹۲۱۸۹	۰.۵۰۰۰۰۰۱۲۵۰۰۰۰۲۱
۰.۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۴۹۹۹۹۹۹۹۸۶۲۷۹۳۱	۰.۵۰۰۰۰۰۰۱۲۵۰۰۰۰۰
۰.۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۰۰۴۱۳۸۶۸۵۲	۰.۵۰۰۰۰۰۰۰۰۱۲۵۰۰۰
۰.۰۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۴۴۴۵۰۲۹۱۳۳۷	۰.۵۰۰۰۰۰۰۰۰۰۱۲۵
۰.۰۰۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۴۹۹۶۰۰۳۶۱۰۸۱۳۲۲	۰.۵۰۰۰۰۰۰۰۰۰۰۰۰۰۱
۰.۰۰۰۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰
۰.۰۰۰۰۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰۰	۰.۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰
۰.۰۰۰۰۰۰۰۰۰۱۰۰۰۰۰۰۰۰۰۰۰	۰.۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	۰.۵۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰

مثال ۳.۲ فرض کنید مشتق دوم تابع f در یک همسایگی از نقطه x_0 پیوسته است. در این صورت، با استفاده از بسط تیلر تابع f در نقطه x_0 و به ازای هر $h > 0$ ، داریم

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{h^2}{2}f''(\xi), \quad |\xi - x_0| \leq h.$$

در نتیجه، خواهیم داشت

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{h}{2}f''(\xi), \quad |\xi - x_0| \leq h.$$

بنابراین، داریم

$$f'(x_0) \approx \frac{f(x_0 + h) - f(x_0)}{h}, \tag{۲.۲}$$

و خطای حاصل از این تقریب به صورت

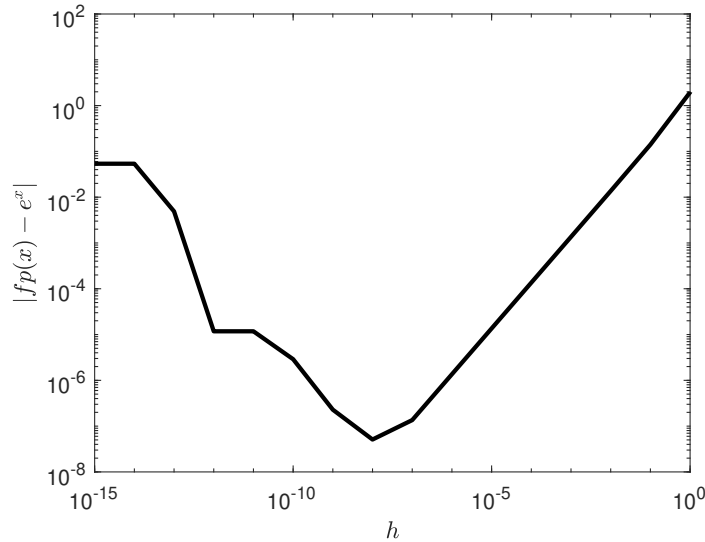
$$E(h) = -\frac{h}{2}f''(\xi),$$

است. اینک، داریم

$$|E(h)| = \frac{h}{2}|f''(\xi)| \leq Ch,$$

که در آن

$$C = \max_{\{x: |x-x_0| \leq h\}} \frac{1}{2}|f''(x)|.$$



شکل ۱.۲: نمایش خطا برای مثال ۳.۲.

از این رو، $E(h) = \mathcal{O}_{h \rightarrow 0}(h)$. فرض کنید $f(x) = e^x$ و $x_0 = 1$. می‌خواهیم با استفاده از رابطه (۲.۲) مقدار $f'(x_0)$ را تقریب زده و به ازای مقادیر کوچک h ، خطای حاصل را مورد مطالعه قرار دهیم. برای این کار، برنامهٔ متلب زیر را تهیه می‌کنیم

```
h=10.^(-15:0);
f=@(x) exp(x);
x0=1;
fp=(f(x0+h)-f(x0))./h;
loglog(h,abs(fp-exp(x0)));
```

این برنامه، تابع

$$e(x) = \frac{f(x+h) - f(x)}{h} - f'(x) = \frac{f(x+h) - f(x)}{h} - e^x,$$

را در بازهٔ $[10^{-15}, 10^0]$ رسم می‌کند. با اجرای این برنامه، شکل ۱.۲ به دست می‌آید. این شکل نشان می‌دهد که به ازای h های کوچک، دقت زیادی در محاسبهٔ

$$fp(x) = \frac{f(x+h) - f(x)}{h},$$

از دست می‌رود.

مثال ۴.۲ فرض کنید $x_0, x_1, \dots, x_n, n+1$ عدد حقیقی مثبت و در محدوده ماشین هستند. اگر این اعداد به ماشین داده شوند و سپس جمع آنها به طور دقیق انجام شود، حداکثر خطای نسبی چه خواهد بود؟

حل: می‌دانیم که برای هر i ، مقدار $fl(x_i)$ مقداری است که در ماشین ذخیره می‌شود. بنابراین، خواهیم داشت $fl(x_i) = x_i(1 + \delta_i)$ به طوری که $|\delta_i| \leq u$. از این رو، داریم

$$\begin{aligned} S^* &= \sum_{i=0}^n fl(x_i) = \sum_{i=0}^n x_i(1 + \delta_i) \\ &= \sum_{i=0}^n x_i + \sum_{i=0}^n x_i \delta_i = S + \sum_{i=0}^n x_i \delta_i, \end{aligned}$$

که در آن $S = \sum_{i=0}^n x_i$. در نتیجه، خواهیم داشت

$$\left| \frac{S^* - S}{S} \right| = \frac{1}{S} \left| \sum_{i=0}^n x_i \delta_i \right| \leq \frac{1}{S} \sum_{i=0}^n |x_i| |\delta_i| \leq \frac{1}{S} \sum_{i=0}^n x_i u = u.$$

بنابراین، حداکثر خطای نسبی برابر با u خواهد بود.

مثال ۵.۲ فرض کنید $x_0, x_1, \dots, x_n, n+1$ عدد ممیز شناور مثبت هستند. در این صورت، خطای نسبی در محاسبه $S = \sum_{i=0}^n x_i$ را محاسبه کنید.

حل: قرار می‌دهیم $S_k = x_0 + x_1 + \dots + x_k$ و فرض می‌کنیم S_k^* مقدار متناظر با آن باشد که در رایانه ذخیره می‌شود. داریم

$$\begin{cases} S_0 = x_0, \\ S_{k+1} = S_k + x_{k+1}, \end{cases} \quad \begin{cases} S_0^* = x_0, \\ S_{k+1}^* = fl(S_k^* + x_{k+1}), \end{cases}$$

که در آن $k = 0, 1, \dots, n-1$. قرار می‌دهیم

$$\rho_k = \frac{S_k^* - S_k}{S_k}, \quad \delta_k = \frac{S_{k+1}^* - (S_k^* + x_{k+1})}{S_k^* + x_{k+1}}.$$

در این صورت، $S_{k+1}^* = (S_k^* + x_{k+1})(1 + \delta_k)$ و $S_k^* = S_k(1 + \rho_k)$ که در آن $|\delta_k| \leq u$. داریم

$$\rho_{k+1} = \frac{S_{k+1}^* - S_{k+1}}{S_{k+1}}$$

$$\begin{aligned}
&= \frac{(S_k^* + x_{k+1})(1 + \delta_k) - (S_k + x_{k+1})}{S_{k+1}} \\
&= \frac{(S_k(1 + \rho_k) + x_{k+1})(1 + \delta_k) - (S_k + x_{k+1})}{S_{k+1}} \\
&= \frac{(S_k + x_{k+1} + S_k \rho_k)(1 + \delta_k) - (S_k + x_{k+1})}{S_{k+1}} \\
&= \rho_k \frac{S_k}{S_{k+1}} (1 + \delta_k) + \frac{S_k + x_{k+1}}{S_{k+1}} \delta_k.
\end{aligned}$$

در نتیجه، خواهیم داشت

$$\rho_{k+1} = \rho_k \frac{S_k}{S_{k+1}} (1 + \delta_k) + \delta_k.$$

اکنون، با توجه به اینکه $S_k < S_{k+1}$ داریم

$$|\rho_{k+1}| \leq |\rho_k| \left| \frac{S_k}{S_{k+1}} \right| (1 + |\delta_k|) + |\delta_k| \leq |\rho_k| (1 + u) + u.$$

اگر قرار دهیم $\theta = 1 + u$ ، آنگاه داریم $|\rho_{k+1}| \leq u + \theta |\rho_k|$ و در نتیجه، خواهیم داشت

$$|\rho_0| = \frac{|S_0^* - S_0|}{|S_0|} = \frac{|x_0 - x_0|}{|x_0|} = 0,$$

$$|\rho_1| \leq u,$$

$$|\rho_2| \leq u + \theta |\rho_1| = u(1 + \theta),$$

$$|\rho_3| \leq u + \theta(u + \theta u) = u(1 + \theta + \theta^2),$$

⋮

$$|\rho_n| \leq u(1 + \theta + \theta^2 + \dots + \theta^{n-1}) = u \frac{\theta^n - 1}{\theta - 1} = (1 + u)^n - 1.$$

اکنون، با کمک بسط دوجمله‌ای نیوتن-خیام، داریم

$$\begin{aligned}
(1 + u)^n - 1 &= \sum_{k=0}^n \binom{n}{k} u^k - 1 \\
&= (1 + nu + \frac{n(n-1)}{2} u^2 + \dots) - 1 \\
&= nu + \frac{n(n-1)}{2} u^2 + \dots.
\end{aligned}$$

در این صورت، اگر n در مقایسه با $\frac{1}{u}$ کوچک باشد، آنگاه خواهیم داشت

$$(1 + u)^n - 1 \approx nu,$$

که نشان می‌دهد nu تقریب مناسبی برای خطای نسبی است.

یکی دیگر از منابع خطا، جمع یک عدد بسیار کوچک با یک عدد بسیار بزرگ است. به مثال زیر توجه کنید.

مثال ۶.۲ فرض کنید $x = 2^{18} = 262144$ و $y = 2^{-6} = 0.015625$. در این صورت، داریم

$$\begin{aligned} x \oplus y &= fl(x + y) \\ &= fl(2^{18} + 2^{-6}) \\ &= fl\left(\left(1.000000\right)_2 \times 2^{18} + \left(0.000001\right)_2 \times 2^{18}\right) \\ &= fl\left(\left(1.000001\right)_2 \times 2^{18}\right) \\ &= \left(1.000000\right)_2 \times 2^{18} = 2^{18}. \end{aligned}$$

این نتیجه نشان می‌دهد که عدد 0.015625 در برابر 262144 در دقت ساده IEEE نادیده گرفته شده است.

گاهی می‌توان با ارائه یک الگوریتم مناسب از جمع یک عدد بسیار کوچک با یک عدد بسیار بزرگ جلوگیری کرد و نتیجه عددی را بهبود داد. به مثال بعدی توجه کنید.

مثال ۷.۲ فرض کنید $x = y = 2^{-6}$ و $z = 2^{18}$. می‌خواهیم در دقت ساده IEEE، مقدار $x + y + z$ را محاسبه کنیم. اگر از $x \oplus (y \oplus z)$ استفاده کنیم، به روشنی با توجه به تمرین ۶.۲ خواهیم داشت $x \oplus (y \oplus z) = z$. اکنون، مقدار $x + y + z$ را به این ترتیب محاسبه می‌کنیم که نخست مجموع دو عدد کوچک‌تر را محاسبه کرده و سپس حاصل را با عدد بزرگ‌تر جمع می‌کنیم. در این صورت، خواهیم داشت

$$\begin{aligned} (x \oplus y) \oplus z &= fl(2^{-6} + 2^{-6}) \oplus z \\ &= 2^{-5} \oplus 2^{18} \\ &= fl\left(\left(0.000001\right)_2 \times 2^{18} + \left(1.000000\right)_2 \times 2^{18}\right) \\ &= fl\left(\left(1.000001\right)_2 \times 2^{18}\right) \\ &= \left(1.000001\right)_2 \times 2^{18}. \end{aligned}$$

پیدااست که این مقدار، تقریب بهتری نسبت به $x \oplus (y \oplus z)$ برای $x + (y + z)$ است.

۴.۲ جلوگیری از حذف فاجعه‌انگیز

گاهی می‌توان با ارائه یک الگوریتم مناسب، از تفریق دو عدد نزدیک به هم و یا به عبارت دیگر، از حذف فاجعه‌انگیز جلوگیری کرد. برای روشن شدن موضوع، چند مثال ارائه می‌کنیم.

مثال ۸.۲ می‌خواهیم مقدار $y = \sqrt{1+x^2} - 1$ را به‌ازای x های بسیار کوچک محاسبه کنیم. اگر x بسیار کوچک باشد، آنگاه داریم $\sqrt{1+x^2} \approx 1$ و مقدار y حاصل تفریق دو عدد نزدیک به هم خواهد بود. برای این رفع این مشکل، داریم

$$\sqrt{1+x^2} - 1 = \frac{x^2}{\sqrt{1+x^2} + 1} \approx \frac{x^2}{2}.$$

برای مثال، اگر $x = 10^{-6}$ ، آنگاه 0.5×10^{-12} تقریب خوبی برای y خواهد بود. برای محاسبه مقدار y با دقت بیشتر می‌توان از

$$(1+y)^n = \sum_{k=0}^n \binom{n}{k} u^k = 1 + ny + \frac{n(n-1)}{2}y^2 + \frac{n(n-1)(n-2)}{6}y^3 + \dots,$$

نیز استفاده کرد. در واقع، داریم

$$(1+x^2)^{\frac{1}{2}} = 1 + \frac{1}{2}x^2 + \frac{\frac{1}{2}(-\frac{1}{2})}{2}x^4 + \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})}{6}x^6 + \dots.$$

برای مثال، اگر $x = 10^{-6}$ ، آنگاه

$$y \approx \frac{1}{2} \times 10^{-12} - \frac{1}{8} \times 10^{-24} + \frac{1}{16} \times 10^{-36}.$$

مثال ۹.۲ فرض کنید x عددی کوچک باشد و می‌خواهیم $y = x - \sin x$ را محاسبه کنیم. به‌ازای x های کوچک داریم $x \approx \sin x$ و بنابراین، y از تفریق دو عدد نزدیک به هم به‌دست می‌آید. ولی، داریم

$$y = x - \sin x = \frac{1}{3!}x^3 - \frac{1}{5!}x^5 + \frac{1}{7!}x^7 - \dots \approx \frac{1}{3!}x^3.$$

بنابراین، $\frac{1}{3!}x^3$ تقریب خوبی برای y خواهد بود. برای مثال، فرض کنید $x = 10^{-8}$. در این صورت، مقداری که نرم‌افزار متلب برای $y = x - \sin x$ به‌دست می‌دهد برابر با صفر است، درحالی‌که مقدار $\frac{1}{3!}x^3$ برابر با $1.6666666666666667 \times 10^{-25}$ است.

مثال ۱۰.۲ می‌خواهیم تابع

$$f(x) = \frac{1}{1+2x} - \frac{1-x}{1+x},$$

را به ازای $|x| \ll 1$ محاسبه کنیم. به ازای مقادیر بسیار نزدیک به صفر، هر دو کسر موجود در f به مقدار ۱ بسیار نزدیک خواهند بود و تفاضل دو عدد نزدیک به هم باعث می‌شود که تعداد ارقام بامعناى زیادى (و در نتیجه دقت) از بین برود. برای مثال، به ازای $x = 10^{-10}$ نرم افزار متلب (با دقت مضاعف) مقدار صفر را برمی‌گرداند که منطقی نیست. توجه کنید که تنها ریشه f عدد صفر است. برای رفع این مشکل، تابع را به صورت

$$f(x) = \frac{2x^2}{(1+2x)(1+x)},$$

می‌نویسیم. اگر از تابع بالا برای محاسبه مقدار $f(10^{-10})$ استفاده کنیم، آنگاه نرم افزار متلب مقدار $10^{-20} \times 1.999999999940000000$ را به دست می‌دهد. همچنین، توجه کنید که به ازای مقادیر بسیار نزدیک به صفر، تابع f را می‌توان با عبارت $2x^2$ تقریب زد. این عبارت به ازای 10^{-10} برابر با $10^{-20} \times 2.000000000000000000$ خواهد بود که مقدار مناسبی است.

۵.۲ خوش وضعی، پایداری و حساسیت

ابتدا تعریف زیر را ارائه می‌کنیم. این تعریف نخستین بار توسط هادامارد در سال ۱۹۲۳ ارائه شده است [۸].

تعریف ۴.۲ یک مسئله خطی را خوش وضع می‌گویند، هرگاه شرایط زیر برقرار باشد

(۱) مسئله حداقل یک جواب داشته باشد؛

(۲) جواب مسئله، یکتا باشد؛

(۳) جواب مسئله به طور پیوسته به داده‌ها وابسته باشد.

اگر مسئله‌ای حداقل یکی از این شرایط را نداشته باشد، آن مسئله را بدوضع می‌گویند.

همه مسائل که در این درس مورد بررسی قرار می‌گیرند، مسائل خوش وضع هستند.

تعریف ۵.۲ یک الگوریتم برای حل یک مسئله پایدار گفته می‌شود، اگر خطای کوچک یک مرحله به طور اساسی رشد نکرده باشد و دقت نتیجه نهایی را به طور جدی تحت تأثیر قرار ندهد. الگوریتمی را که پایدار نباشد، ناپایدار می‌گویند.

با توجه به این تعریف، می‌بینیم که پایداری، به الگوریتم حل مسئله برمی‌گردد و ممکن است الگوریتمی برای حل یک مسئله، پایدار و الگوریتمی دیگر برای حل همان مسئله، ناپایدار باشد. مثال زیر این موضوع را روشن می‌کند.

مثال ۱۱.۲ دنباله

$$x_0 = 1, \quad x_1 = \frac{1}{3}, \quad (3.2)$$

$$x_n = \frac{13}{3}x_{n-1} - \frac{4}{3}x_{n-2}, \quad n \geq 2, \quad (4.2)$$

را در نظر بگیرید. می‌دانیم که جواب عمومی رابطه بازگشتی (۴.۲) به صورت $x_n = qr^n$ است [۷]. با توجه به مقادیر اولیه (۳.۲)، مقادیر q و r ناصفر هستند. با جای‌گذاری $x_n = qr^n$ در (۴.۲) خواهیم داشت

$$3r^2 - 13r + 4 = 0.$$

ریشه‌های این معادله عبارتند از $r_1 = \frac{1}{3}$ و $r_2 = 4$. بنابراین، جواب عمومی مسئله به صورت

$$x_n = \alpha r_1^n + \beta r_2^n = \alpha \left(\frac{1}{3}\right)^n + \beta \times 4^n,$$

خواهد بود. به‌ازای $x_0 = 1$ و $x_1 = \frac{1}{3}$ داریم

$$\begin{cases} \alpha + \beta = 1, \\ \alpha \left(\frac{1}{3}\right) + 4\beta = \frac{1}{3}. \end{cases}$$

با حل این دستگاه، خواهیم داشت $\alpha = 1$ و $\beta = 0$ که دنباله

$$x_n = \left(\frac{1}{3}\right)^n, \quad (5.2)$$

را نتیجه می‌دهد.

با استفاده از یک برنامه به زبان متلب، جملات دنباله را محاسبه کرده و در جدول بعد نشان داده‌ایم. در این جدول، x_n با استفاده از رابطه (۵.۲) و مقدار x_n^* ، با استفاده از روند بازگشتی (۴.۲) محاسبه شده است. ملاحظه می‌شود که x_{16}^* منفی است. دلیل آن را می‌توان به صورت زیر توجیه کرد. در گام دوم، با توجه به اینکه عدد $\frac{1}{3}$ را نمی‌توان به‌طور دقیق در رایانه ذخیره کرد، خطای کوچکی در آن وارد شده است که این خطا رشد کرده و در x_{16}^* جمع شده است (انباشتگی خطا) و در نتیجه، باعث شده است که تفاوت زیادی با مقدار واقعی داشته باشد (البته توجه کنید که این اتفاق در گام چهاردهم نیز رخ داده است). در ادامه، دلیل رشد خطا را مطالعه می‌کنیم.

فرض کنید مقادیر اولیه به صورت $x_0 = 1$ و $x_1 = \frac{1}{3} + \epsilon$ باشد (خطای کوچکی به اندازه ϵ در x_1 وارد شده است). در این صورت، جمله عمومی دنباله به صورت

$$x_n = \left(1 - \frac{3}{11}\epsilon\right)\left(\frac{1}{3}\right)^n + \frac{3}{11}\epsilon \times 4^n,$$

خواهد بود. این رابطه نشان می‌دهد که خطای ϵ در محاسبه x_n ، 4^n برابر می‌شود و به سرعت رشد می‌کند. نتایج نشان می‌دهد که روند تکراری مورد استفاده برای محاسبه x_n ناپایدار است.

n	x_n	x_n^*
۲	۰.۱۱۱۱۱۱۱۱۱۱۱۱۱۱	۰.۱۱۱۱۱۱۱۱۱۱۱۱۱۱
۳	۰.۰۳۷۰۳۷۰۳۷۰۳۷۰۴	۰.۰۳۷۰۳۷۰۳۷۰۳۷۰۴
۴	۰.۰۱۲۳۴۵۶۷۹۰۱۲۳۵	۰.۰۱۲۳۴۵۶۷۹۰۱۲۳۴
۵	۰.۰۰۴۱۱۵۲۲۶۳۳۷۴۵	۰.۰۰۴۱۱۵۲۲۶۳۳۷۴۴
⋮	⋮	⋮
۱۴	۰.۰۰۰۰۰۰۰۲۰۹۰۷۵۱۶	۰.۰۰۰۰۰۰۰۲۰۵۷۵۱۹۵
۱۵	۰.۰۰۰۰۰۰۰۶۹۶۹۱۷۲	۰.۰۰۰۰۰۰۰۵۶۳۹۸۸۸
۱۶	۰.۰۰۰۰۰۰۰۲۳۲۳۰۵۷	-۰.۰۰۰۰۰۰۰۲۹۹۴۰۸۰

مثال ۱۲.۲ محاسبه انتگرال‌های

$$I_n = \int_0^1 \frac{x^n}{x+5} dx, \quad n = 0, 1, 2, \dots, 23,$$

را در نظر بگیرید. به روشنی، دنباله I_n نزولی و مثبت است. داریم

$$I_n + 5I_{n-1} = \int_0^1 \frac{x^n + 5x^{n-1}}{x+5} dx = \int_0^1 x^{n-1} dx = \frac{1}{n}.$$

بنابراین

$$I_n = \frac{1}{n} - 5I_{n-1}, \quad n = 1, 2, \dots, 23. \quad (6.2)$$

از طرفی، داریم

$$I_0 = \int_0^1 \frac{1}{x+5} dx = \ln \frac{6}{5}.$$

یک برنامه رایانه‌ای به زبان متلب نوشته‌ایم و با استفاده از رابطه بازگشتی (۶.۲)، مقادیر I_n را محاسبه کرده‌ایم. این مقادیر در جدول زیر آورده شده است. همان‌گونه که می‌بینیم، ناپایداری رخ داده است؛ زیرا در تکرار ۱۹، نزولی بودن دنباله به هم خورده است و به علاوه، $I_{22} < 0$.

n	I_n
۰	۰.۱۸۲۳۲۱۵۵۶۷۹۳۹۵
۱	۰.۰۸۸۳۹۲۲۱۶۰۳۰۲۳
۲	۰.۰۵۸۰۳۸۹۱۹۸۴۸۸۶
⋮	⋮
۱۷	۰.۰۰۹۳۷۱۹۰۶۸۵۶۸۶
۱۸	۰.۰۰۸۶۹۶۰۲۱۲۷۱۲۷
۱۹	۰.۰۰۹۱۵۱۴۷۲۵۹۱۰۲
۲۰	۰.۰۰۴۲۴۲۶۳۷۰۴۴۹۲
۲۱	۰.۰۲۶۴۰۵۸۶۲۳۹۴۴۴
۲۲	-۰.۰۸۶۵۷۴۷۶۷۶۵۱۷۶۵

برای تحلیل ناپایداری، فرض کنید I_n^* ها مقادیری هستند که برنامه رایانه‌ای محاسبه می‌کند. فرض کنید ϵ خطای حاصل از گرد کردن عدد در I_0^* باشد، یا به عبارت دیگر $I_0 = I_0^* + \epsilon$. در این صورت، با استفاده از رابطه بازگشتی داریم

$$\begin{aligned}
 I_1 &= 1 - 5I_0 = 1 - 5(I_0^* + \epsilon) = I_1^* - 5\epsilon, \\
 I_2 &= \frac{1}{4} - 5I_1 = \frac{1}{4} - 5(I_1^* - 5\epsilon) = \frac{1}{4} - 5I_1^* + 25\epsilon = I_2^* + 5^2\epsilon, \\
 &\vdots \\
 I_n &= \frac{1}{n} - 5I_{n-1} = \frac{1}{n} - 5(I_{n-1}^* + (-1)^{n-1}5^{n-1}\epsilon) = I_n^* + (-1)^n 5^n \epsilon.
 \end{aligned}$$

تساوی آخر نشان می‌دهد که خطای کوچک ϵ در I_0^* وقتی به I_n^* می‌رسد، بزرگ شده و به میزان $(-1)^n 5^n \epsilon$ می‌رسد، که به‌ازای n های بزرگ، بسیار بزرگ خواهد بود.

اکنون، برای این مثال یک الگوریتم پایدار ارائه می‌کنیم. با توجه به اینکه $I_n \rightarrow 0$ داریم $I_{22} \approx I_{21}$. بنابراین، با استفاده از رابطه بازگشتی (۶.۲)، خواهیم داشت

$$I_{22} = \frac{1}{44} - 5I_{21} \approx \frac{1}{44} - 5I_{22}.$$

در نتیجه، داریم

$$I_{22} \approx \frac{1}{6 \times 22} = 0.00757575757576.$$

شایان ذکر است که با استفاده از دستور quad در نرم‌افزار متلب مقدار I_{22} را محاسبه کرده‌ایم که برابر است با $I_{22} \approx 0.00729738381631$. با مشخص شدن مقدار تقریبی I_{22} ، رابطه بازگشتی را

به صورت

$$I_{n-1} = \frac{1}{\delta} \left(\frac{1}{n} - I_n \right),$$

می نویسیم. اکنون، می توانیم مقادیر I_n را به صورت بازگشتی پسرو محاسبه کنیم. با استفاده از یک برنامه در نرم افزار متلب نتایج جدول زیر به دست آمده است. این جدول نشان می دهد که با وجود اینکه مقدار I_n از دقت زیادی برخوردار نیست، همه رقم های اعشاری I_0 با مقدار واقعی آن، یعنی، $\ln \frac{7}{5} = 0.18232155679395$ یکسان است. این نتیجه نشان می دهد که الگوریتم مورد استفاده برای محاسبه I_n ها پایدار است. به عنوان یک تمرین، پایداری این الگوریتم را تحلیل نمایید.

n	I_n
۰	۰.۱۸۲۳۲۱۵۵۶۷۹۳۹۵
۱	۰.۰۸۸۳۹۲۲۱۶۰۳۰۲۳
۲	۰.۰۵۸۰۳۸۹۱۹۸۴۸۸۷
⋮	⋮
۱۶	۰.۰۰۹۸۹۶۳۵۰۱۴۴۵۲
۱۷	۰.۰۰۹۳۴۱۷۷۸۶۸۹۱۵
۱۸	۰.۰۰۸۸۴۶۶۶۲۱۰۹۸۲
۱۹	۰.۰۰۸۳۹۸۲۶۸۳۹۸۲۷
۲۰	۰.۰۰۸۰۰۸۶۵۸۰۰۸۶۶
۲۱	۰.۰۰۷۵۷۵۷۵۷۵۷۵۷۶
۲۲	۰.۰۰۷۵۷۵۷۵۷۵۷۵۷۶

مثال ۱۳.۲ این مثال از مرجع [۶] استخراج شده است. فرض کنید n عدد حقیقی x_1, x_2, \dots, x_n داده شده باشند. میانگین این اعداد برابر با

$$\mu_n = \frac{1}{n} \sum_{k=1}^n x_k, \quad (7.2)$$

و واریانس آنها برابر با

$$var(x) = \frac{1}{n} \sum_{k=1}^n (x_k - \mu_n)^2, \quad (8.2)$$

است. همچنین جذر واریانس، انحراف معیار نامیده می شود که برابر با

$$\sigma_n = \sqrt{var(x)}, \quad (9.2)$$

است. اگر داده‌ها را در بردار $x = [x_1, x_2, \dots, x_n]$ ذخیره کنیم، آنگاه با کمک دستورات زیر در متلب می‌توانیم مقادیر μ_n ، $var(x)$ و σ_n را محاسبه کنیم

```
n=length(x);
mux=sum(x)/n;
varx=sum((x-mux).^2)/n;
sx=sqrt(varx);
```

برای محاسبه واریانس به کمک رابطه (۸.۲)، لازم است کل داده‌ها دو بار پیمایش شوند، یک بار برای محاسبه μ_n و بار دیگر برای محاسبه $var(x)$. ولی می‌توانیم فرمولی دیگر برای محاسبه واریانس به دست آوریم که نیاز به دو بار پیمایش داده‌ها نباشد. داریم

$$\begin{aligned} var(x) &= \frac{1}{n} \sum_{k=1}^n (x_k^2 - 2\mu_n x_k + \mu_n^2) \\ &= \frac{1}{n} \sum_{k=1}^n x_k^2 - 2\mu_n \frac{1}{n} \sum_{k=1}^n x_k + \mu_n^2 \frac{1}{n} \sum_{k=1}^n 1 \\ &= \frac{1}{n} \sum_{k=1}^n x_k^2 - \mu_n^2. \end{aligned} \quad (10.2)$$

با توجه به رابطه به دست آمده، می‌توان از دستورات زیر و تنها با یک بار پیمایش داده مقدار واریانس را محاسبه کرد

```
s=0; sq=0; n=0;
while n<length(x)
    n=n+1;
    s=s+x(n);
    sq=sq+x(n)^2;
end
mux=s/n;
varx=sq/n-mux^2;
sx=sqrt(varx);
```

بردار x را به صورت

$$x = 100 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + 10^{-5} \left(v - 0.5 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right),$$

تعریف می‌کنیم که در آن $v \in \mathbb{R}^n$ یک بردار تصادفی با توزیع یکنواخت است. می‌توان این بردار را در متلب توسط دستور زیر تولید کرد

جدول ۱.۲: محاسبه میانگین، واریانس و انحراف معیار به کمک رابطه‌های (۵.۲) و (۷.۲).

	رابطه (۸.۲)	رابطه (۱۰.۲)
μ_n	۹۹.۹۹۹۹۹۹۵۶۴۵۳۹۰۱	۹۹.۹۹۹۹۹۹۵۶۴۵۳۹۰۱
$var(x)$	$۸.۳۱۵۵۵۵۴۰۴۰۷۱۵۷۴e - ۱۲$	$۵.۴۵۶۹۶۸۲۱۰۶۳۷۵۶۹e - ۱۲$
σ_n	$۲.۸۸۳۶۷۰۴۷۴۲۵۱۷۹۵e - ۰۶$	$۲.۳۳۶۰۱۵۴۵۵۹۹۲۸۶۸e - ۰۶$

$$x=100*\text{ones}(100,1)+1e-5*(\text{rand}(100,1)-0.5);$$

در جدول ۱.۲، مقادیری از μ_n ، $var(x)$ و σ_n که توسط رابطه‌های (۸.۲) و (۱۰.۲) محاسبه شده‌اند، ارائه کرده‌ایم. این جدول نشان می‌دهد که مقدار $var(x)$ محاسبه‌شده توسط رابطه (۱۰.۲) از دقت مناسبی برخوردار نیست و دلیل آن، ازدست‌دادن دقت به دلیل تفریق دو عدد نزدیک به هم است. اگرچه محاسبه واریانس توسط رابطه (۱۰.۲) تنها به یک بار پیمایش در داده‌ها نیاز دارد، ولی از لحاظ عددی ناپایدار است.

در ادامه، یک روش پایدار ارائه می‌کنیم. داریم

$$\begin{aligned} n\sigma_n^2 &= \sum_{k=1}^n (x_k - \mu_n)^2 \\ &= \sum_{k=1}^{n-1} (x_k - \mu_n)^2 + (x_n - \mu_n)^2 \\ &= \sum_{k=1}^{n-1} ((x_k - \mu_{n-1}) - (\mu_n - \mu_{n-1}))^2 + (x_n - \mu_n)^2 \\ &= \sum_{k=1}^{n-1} (x_k - \mu_{n-1})^2 - 2(\mu_n - \mu_{n-1}) \underbrace{\sum_{k=1}^{n-1} (x_k - \mu_{n-1})}_{=0} \\ &\quad + (n-1)(\mu_n - \mu_{n-1})^2 + (x_n - \mu_n)^2 \\ &= (n-1)\sigma_{n-1}^2 + (n-1)(\mu_n - \mu_{n-1})^2 + (x_n - \mu_n)^2. \end{aligned} \quad (11.2)$$

از طرفی، داریم

$$n\mu_n = (n-1)\mu_{n-1} + x_n,$$

که رابطه

$$\mu_{n-1} = \frac{n}{n-1}\mu_n - \frac{1}{n-1}x_n,$$

را نتیجه می دهد. بنابراین

$$(n-1)(\mu_n - \mu_{n-1})^2 = (n-1)\left(\mu_n - \frac{n}{n-1}\mu_n + \frac{1}{n-1}x_n\right)^2 = \frac{1}{n-1}(x_n - \mu_n)^2.$$

با جای گذاری این رابطه در رابطه (۱۱.۲)، خواهیم داشت

$$n\sigma_n^2 = (n-1)\sigma_{n-1}^2 + \frac{n}{n-1}(x_n - \mu_n)^2,$$

یا

$$\sigma_n^2 = \frac{n-1}{n}\sigma_{n-1}^2 + \frac{1}{n-1}(x_n - \mu_n)^2.$$

اکنون، با کمک این رابطه بازگشتی، برنامه متلب زیر را می نویسیم

```
s=x(1); mux=s; varx=0; n=1;
while n<length(x)
    n=n+1;
    s=s+x(n);
    mux=s/n;
    varx=(n-1)*varx/n+(x(n)-mux)^2/(n-1);
end
sx=sqrt(varx);
```

با اجرای این برنامه، مقادیر زیر را به دست می آوریم

$$\begin{aligned} \mu_n &= 99.9999999564453901, \\ var(x) &= 8.315555416278121e-12, \\ \sigma_n &= 2.883670476368290e-06. \end{aligned}$$

همان گونه که ملاحظه می شود، مقدار $var(x)$ و σ_n از دقت کافی برخوردار هستند.

تعریف ۶.۲ یک مسئله را حساس یا بدحالت می گویند، هرگاه یک تغییر کوچک در داده های مسئله باعث تغییر اساسی در جواب آن گردد؛ در غیر این صورت، مسئله را غیرحساس یا خوش حالت می گویند. حساسیت مربوط به ماهیت مسئله و مستقل از الگوریتمی است که برای حل مسئله به کار می رود.

یادداشت ۳.۲ باید توجه کرد که بین بدحالتی و بدوضعی، تفاوت زیادی وجود دارد. چنان که خواهیم دید، یک مسئله ممکن است خوش وضع، ولی بدحالت باشد.

مثال ۱۴.۲ دستگاه معادلات خطی

$$\begin{pmatrix} 1 & 1 \\ 1.1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad (12.2)$$

را در نظر بگیرید. جواب واقعی این دستگاه $x_1 = 10$ و $x_2 = -9$ است. اکنون، اگر ضریب $(2, 1)$ را از 1.1 به 1.01 تغییر دهیم، جواب واقعی دستگاه به صورت $x_1 = 100$ و $x_2 = -99$ خواهد شد. می‌بینیم که یک تغییر کوچک در یک داده مسئله باعث تغییر اساسی در جواب واقعی مسئله شده است. این نتیجه نشان می‌دهد که مسئله نسبت به ضریب $(2, 1)$ حساس است.

در ادامه، خواهیم دید که چگونه می‌توان حساسیت (بدحالتی) یک مسئله را برآورد کرد.

تعریف ۷.۲ فرض کنید $E_{rel,in}$ و $E_{rel,out}$ به ترتیب خطای نسبی در داده‌ها و خروجی یک مسئله هستند. در این صورت، عدد حالت مسئله به صورت

$$\kappa = \frac{E_{rel,out}}{E_{rel,in}},$$

تعریف می‌شود.

مسئله محاسبه مقدار تابع f را در نظر بگیرید. در اینجا x داده و $f(x)$ خروجی مسئله است. اگر بخواهیم f را در $x+h$ محاسبه کنیم (h عددی کوچک است)، خواهیم داشت

$$E_{rel,in} = \left| \frac{(x+h) - x}{x} \right| = \left| \frac{h}{x} \right|,$$

$$E_{rel,out} = \left| \frac{f(x+h) - f(x)}{f(x)} \right| \approx \left| \frac{hf'(x)}{f(x)} \right|.$$

بنابراین

$$\kappa = \frac{E_{rel,out}}{E_{rel,in}} \approx \left| \frac{xf'(x)}{f(x)} \right|.$$

مثال ۱۵.۲ می‌خواهیم مقدار تابع $f(x) = x - \sin x$ را به‌ازای یک مقدار کوچک و ناصفر x محاسبه کنیم. داریم

$$\frac{xf'(x)}{f(x)} = \frac{x(1 - \cos x)}{x - \sin x}$$

$$= \frac{x \left(1 - \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots \right) \right)}{x - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots \right)}$$

$$= \frac{1 - \frac{x^2}{4!} + \dots}{1 - \frac{x^2}{5!} + \dots} \approx \frac{\frac{1}{3}}{\frac{1}{3}} = 3.$$

بنابراین، $\kappa \approx 3$ و در نتیجه، مسئله محاسبه f به ازای مقادیر کوچک ناصفر خوش حالت است.

مثال ۱۶.۲ می خواهیم تابع

$$f(x) = (x - 1)^2 + 2,$$

را در $x = 1$ محاسبه کنیم. داریم $f(1) = 2$ و $f'(1) = 0$ و بنابراین $\kappa = \infty$. در نتیجه، مسئله محاسبه تابع f در $x = 1$ خوش حالت است. توجه کنید که اگر $x = 1 + h$ ، آنگاه $f(x) = 2 + h^2$. بنابراین، دقت $f(x)$ خیلی بیشتر از دقت x است. اکنون، اگر فرض کنیم

$$f(x) = 2 + ((x - 1)^2)^{0.51},$$

آنگاه مثل تابع قبل، محاسبه تابع f در $x = 1$ خوش حالت است. در اینجا توجه کنید که به ازای $x = 1 + h$ ، داریم $f(1 + h) \approx 2 + |h|$. یعنی، خطای ایجاد شده در $f(1 + h)$ تقریباً برابر با خطای وارد شده در x است.

مثال ۱۷.۲ تابع $f(x) = \sqrt{x+4} - 2$ را در نظر بگیرید. در یک نقطه دلخواه x در بازه $(-4, \infty)$ داریم

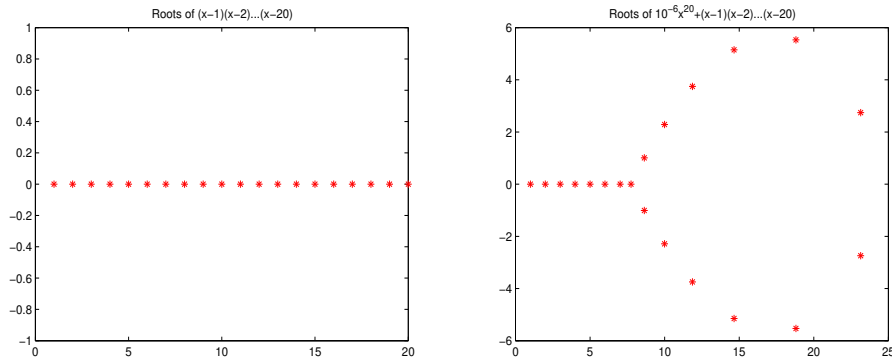
$$\frac{xf'(x)}{f(x)} = \frac{2 + \sqrt{4+x}}{2\sqrt{4+x}}.$$

به ازای x های نزدیک به -4 ، تابع f بد حالت و در سایر نقاط خوش حالت است. برای مثال، به کمک نرم افزار متلب، داریم

$$f(-4 + 10^{-14}) = -1.9999999898935408,$$

چنان که مشاهده می شود، خطایی به اندازه 10^{-14} در داده (متغیر x)، تغییر بسیار زیادی در مقدار تابع ایجاد کرده است. توجه کنید که مقدار تابع در نقطه -4 برابر با -2 است.

این بخش را با مثال دیگری به پایان می بریم.



شکل ۲.۲: ریشه‌های $p(x)$ و $q(x)$ در مثال ۱۸.۲.

مثال ۱۸.۲ چندجمله‌ای

$$p(x) = \prod_{i=1}^{20} (x - i) = x^{20} + \dots,$$

را در نظر بگیرید. ریشه‌های این چندجمله‌ای برابر با $x_1 = 1, \dots, x_{20} = 20$ هستند. ضریب x^{20} را از ۱ به $1 + 10^{-6}$ تبدیل می‌کنیم و چندجمله‌ای حاصل را $q(x)$ می‌نامیم. در واقع، داریم

$$q(x) = p(x) + 10^{-6}x^{20} = (1 + 10^{-6})x^{20} + \dots.$$

با محاسبه ریشه‌های q خواهیم دید که ۱۲ ریشه آن مختلط هستند و با ریشه‌های p تفاوت زیادی دارند. به عبارت دیگر، ریشه‌های p نسبت به ضریب x^{20} بسیار حساس هستند (بدحالت است). در شکل ۲.۲ توزیع ریشه‌های p و q نمایش داده شده است.

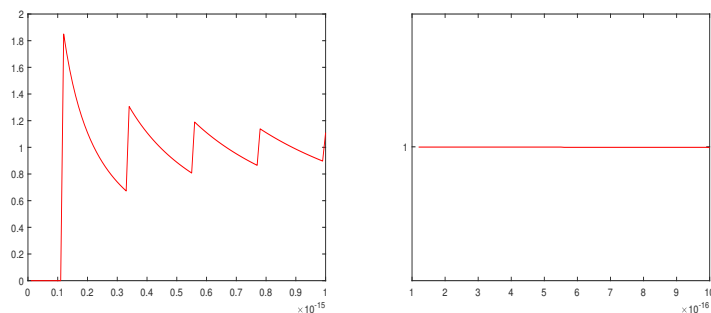
تمرین

(۱.۲) اعداد $\frac{5}{3}$ و 13.43 را در مبنای ۲ نمایش دهید.

(۲.۲) مباحث اشاره‌شده در بخش ۱.۲ درباره دقت ساده را برای دقت مضاعف نیز بنویسید.

(۳.۲) نشان دهید که فاصله بین هر دو عدد زیرنرمال در دقت ساده برابر با 2^{-149} است.

(۴.۲) در دقت مضاعف IEEE، مقادیر اپسیلون ماشین و خطای روند عدد یک را بیابید.



شکل ۲.۲: شکل تمرین ۷.۲.

(۵.۲) فرض کنید $i = 1, 2, 3, 4, x_i > 0$. یک کران بالا برای خطای مطلق در محاسبه مقدار $z = x_1 x_2 + x_3 x_4$ به دست آورید.

(۶.۲) نشان دهید که روند تکراری

$$x_0 = 1, \quad x_n = \frac{1}{3} x_{n-1} \quad n \geq 1,$$

برای محاسبه $x_n = \left(\frac{1}{3}\right)^n$ پایدار است.

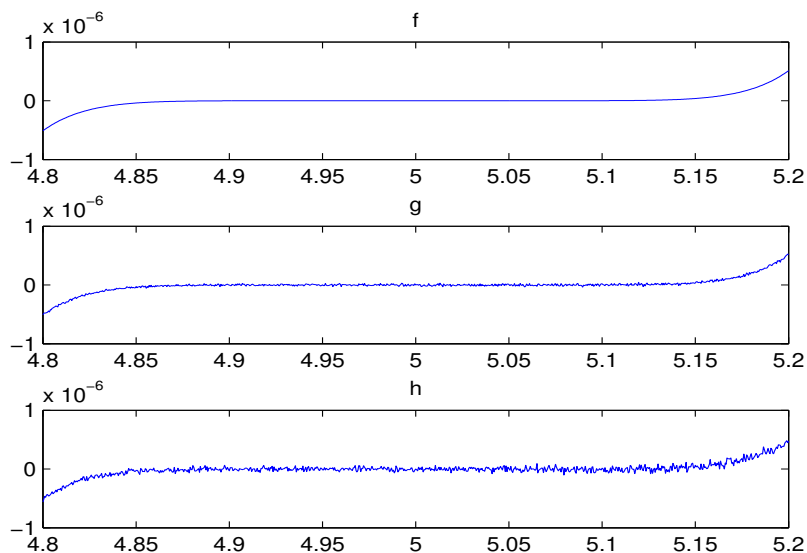
(۷.۲) تابع

$$f(x) = \frac{\ln(1+x)}{x},$$

را در نظر بگیرید. این تابع را به دو روش مختلف در بازه $(10^{-15}, 10^{-17})$ با طول گام 10^{-17} و با استفاده از نرم افزار متلب رسم کرده ایم. در روش نخست، با استفاده از $y = \ln(1+x)/x$ تابع f را رسم کرده ایم (شکل ۳.۲، چپ) و در روش دوم، قرار می دهیم $d = 1+x$ و سپس نمودار $y = \ln(d)/(d-1)$ را نمایش داده ایم (شکل ۳.۲، راست). با اینکه هر دو نمودار برای نمایش یک تابع رسم شده اند، چرا با یکدیگر متفاوتند؟

(۸.۲) چند جمله ای های

$$p(x) = (x-5)^9,$$



شکل ۴.۲: چندجمله‌ای‌های p ، q و r در تمرین ۸.۲.

$$\begin{aligned}
 q(x) &= x^9 - 45x^8 + 900x^7 - 10500x^6 + 78750x^5 - 393750x^4 \\
 &\quad + 1312500x^3 - 2812500x^2 + 3515625x - 1953125, \\
 r(x) &= x(x(x(x(x(x(x(x - 45) + 900) - 10500) + 78750) \\
 &\quad - 393750) + 1312500) - 2812500) + 3515625) - 1953125,
 \end{aligned}$$

را در نظر بگیرید. می‌توان دید که این سه چندجمله‌ای با هم برابرند. تعداد محاسبات هریک از این چندجمله‌ای‌ها را محاسبه کرده و با هم مقایسه کنید. با استفاده از دستور plot در متلب این سه چندجمله‌ای را در بازه $[4.8, 5.2]$ با استفاده از طول گام 0.0005 رسم کرده‌ایم. نمودار این چندجمله‌ای‌ها در شکل ۴.۲ نشان داده شده است. دلیل تفاوت بین شکل‌ها را بیان کنید.

۹.۲ تابع $f_1(x, \delta) = \cos(x + \delta) - \cos x$ را در نظر بگیرید. با استفاده از فرمول

$$\cos \phi - \cos \psi = -2 \sin \frac{\phi + \psi}{2} \sin \frac{\phi - \psi}{2},$$

عبارت دیگری برای f_1 یافته و $f_2(x, \delta)$ بنامید.

(الف) نشان دهید، اگر δ به اندازه کافی کوچک باشد، دو عبارت $f_1(x, \delta)/\delta$ و $f_2(x, \delta)/\delta$ تقریب

مناسبی برای $\sin x - \sin x$ به دست می دهد.

(ب) به ازای $x = 3$ و $\delta = 1e - 11$ و با کمک نرم افزار متلب، مقادیر $g_1(x, \delta) = f_1(x, \delta) + \sin x$ و $g_2(x, \delta) = f_2(x, \delta) + \sin x$ را محاسبه کنید. دلیل تفاوت این دو مقدار را بیان کنید.

۱۰.۲ نشان دهید

$$\ln(x - \sqrt{x^2 - 1}) = -\ln(x + \sqrt{x^2 - 1}).$$

کدام یک از این دو عبارت برای محاسبات عددی مناسب تر است؟ چرا؟ با یک مثال عددی ادعای خود را نشان دهید.

۱۱.۲ برای هر یک از عبارتهای زیر، نشان دهید که ممکن است مشکلات عددی به وجود آید. سپس عبارتهای را به صورتی بنویسید که مشکلات نوشته شده رفع شود.

(الف) عبارت $\sqrt{x + \frac{1}{x}} - \sqrt{x - \frac{1}{x}}$ به ازای $x \gg 1$.

(ب) عبارت $\sqrt{\frac{1}{a^2} + \frac{1}{b^2}}$ به ازای $a \approx 0$ و $b \approx 1$.

۱۲.۲ دستگاه معادلات خطی

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

را در نظر بگیرید که در آن $a, b > 0$.

(الف) اگر $a \approx b$ ، چه مشکل عددی در حل این دستگاه به وجود می آید؟

(ب) یک روش عددی پایدار برای محاسبه $z = x + y$ ارائه کنید.

۱۳.۲ در محاسبه تابع

$$f(x) = \frac{e^x - 1 - x}{x^2},$$

به وسیله رایانه، به ازای مقادیر $x \approx 0$ خطای نسبی زیادی ایجاد می شود؛ دلیل آن چیست؟ عبارت دیگری برای f بنویسید که محاسبه f را برای مقادیر $x \approx 0$ بهبود دهد. با یک برنامه متلب درستی موضوع را بررسی کنید.

۱۴.۲ به ازای هر $x_0 > -1$ دنباله ای که به کمک رابطه بازگشتی

$$x_{n+1} = 2^{n+1}(\sqrt{1 + 2^{-n}x_n} - 1),$$

تعریف می‌شود، به $\ln(1 + x + \frac{1}{4}x^2)$ همگرا می‌شود. ولی برنامه متلب

```
x=4;
for n=1:100
    x=2^(n+1)*(sqrt(1+x/2^n)-1);
end
x
```

جوابی کاملاً نادرست به دست می‌دهد. دلیل آن چیست؟ چگونه می‌توان این مشکل را حل کرد؟

۱۵.۲) فرض کنید b و c دو ضلع یک مثلث قائم‌الزاویه باشند، به طوری که $c > 0$ و $b \gg c$. در محاسبه وتر مثلث به وسیله رایانه خطای نسبی زیادی ایجاد می‌شود. دلیل آن چیست؟ برای رفع این مشکل، چه راهی پیشنهاد می‌کنید؟

۱۶.۲) برنامه متلب

```
d=sqrt(b^2-4*a*c);
r1=-(b-d)/(2*a);
r2=-(b+d)/(2*a);
```

ریشه‌های معادله درجه دوم $ax^2 + bx + c = 0$ را محاسبه می‌کند. از این برنامه برای حل معادله درجه دوم $x^2 - 10^9x + 1 = 0$ ، که دو ریشه مثبت دارد، استفاده می‌کنیم. کدامیک از ریشه‌ها از دقت کافی برخوردار نیست و چرا؟ اگر از برنامه

```
d=sqrt(b^2-4*a*c);
r1=-(b+sign(b)*d)/(2*a);
r2=c/(a*r1);
```

استفاده کنیم، چه نتیجه‌ای به دست می‌آید؟

فصل ۳

درونیابی

فرض کنید مقادیری از یک تابع (معلوم یا نامعلوم) در تعدادی از نقاط با استفاده از آزمایش یا نمونه‌گیری به دست آمده است. می‌خواهیم مقدار این تابع را در نقاط دیگری بدون مراجعه به خود تابع تقریب بزنیم. این کار با استفاده از درونیابی امکان‌پذیر است.

۱.۳ درونیابی با استفاده از چندجمله‌ای‌ها

فرض کنید Π_n مجموعه تمام چندجمله‌ای‌ها با ضرایب حقیقی یا مختلط از درجه حداکثر n است، یعنی

$$\Pi_n = \left\{ P(x) : P(x) = \sum_{i=0}^n a_i x^i, a_i \in \mathbb{C} \right\}.$$

تعریف ۱.۳ می‌گوییم چندجمله‌ای $P \in \Pi_n$ تابع f را در نقاط متمایز $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ ، درونیابی می‌کند، هرگاه

$$P(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

را چندجمله‌ای درونیاب می‌گویند.

قضیه ۱.۳ چندجمله‌ای درونیاب یکتا است.

برهان: به برهان خُلف، فرض کنید P_1 و P_2 دو چندجمله‌ای درونیاب برای تابع f در نقاط متمایز $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ هستند، یعنی

$$P_1(x_i) = P_2(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

قرار می‌دهیم $P \equiv P_1 - P_2$. در این صورت، P یک چندجمله‌ای از درجه حداکثر n است و

$$P(x_i) = P_1(x_i) - P_2(x_i) = f(x_i) - f(x_i) = 0, \quad i = 0, 1, \dots, n.$$

این رابطه نشان می‌دهد که چندجمله‌ای P ، $n + 1$ ریشه دارد که یک تناقض است. □

در ادامه، وجود چندجمله‌ای درونیاب را مطالعه می‌کنیم. قرار می‌دهیم

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, 1, \dots, n. \quad (1.3)$$

روشن است که داریم

$$L_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

اکنون، قرار می‌دهیم

$$P(x) = \sum_{i=0}^n f(x_i) L_i(x).$$

در این صورت، روشن است که $P \in \Pi_n$ و

$$P(x_k) = \sum_{i=0}^n f(x_i) L_i(x_k) = \sum_{i=0}^n f(x_i) \delta_{ik} = f(x_k), \quad k = 0, 1, \dots, n.$$

یعنی P ، چندجمله‌ای درونیاب تابع f در نقاط داده است. چندجمله‌ای P را چندجمله‌ای درونیاب لاگرانژ می‌گویند.

مثال ۱.۳ سه مقدار از تابع f به صورت $\{(0, 1), (1, 2), (2, 4)\}$ داده شده است. چندجمله‌ای درونیاب تابع f را در این نقاط بنویسید و با کمک آن، مقدار تقریبی $f(1.5)$ را محاسبه کنید.

حل: داریم

$$\begin{aligned} L_0(x) &= \frac{(x-1)(x-2)}{(0-1)(0-2)}, \\ L_1(x) &= \frac{(x-0)(x-2)}{(1-0)(1-2)}, \\ L_2(x) &= \frac{(x-0)(x-1)}{(2-0)(2-1)}. \end{aligned}$$

در این صورت

$$P(x) = 1 \times L_0(x) + 2 \times L_1(x) + 4 \times L_2(x) = \frac{1}{4}(x^2 + x + 2),$$

چند جمله‌ای درونیاب مورد نظر است. بنابراین، داریم

$$f(1.5) \approx P(1.5) = \frac{1}{4}(1.5^2 + 1.5 + 2) = 2.875.$$

فرض کنید $w(x) = \prod_{j=0}^n (x - x_j)$ به سادگی می‌توان دید که

$$w'(x_i) = \prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j),$$

و بنابراین، داریم

$$L_i(x) = \frac{w(x)}{(x - x_i)w'(x_i)}.$$

با کمک این رابطه، چند جمله‌ای درونیاب لاگرانژ را می‌توان به صورت

$$P(x) = w(x) \sum_{i=0}^n \frac{f(x_i)}{(x - x_i)w'(x_i)}, \quad (2.3)$$

نوشت.

برای محاسبه هر کدام از L_i ها در نقطه داده شده x ، به $(n-1)$ عمل حسابی نیاز داریم. در نتیجه، برای تمام L_i ها به $(n-1)(n+1)$ عمل حسابی، و برای محاسبه چند جمله‌ای درونیاب لاگرانژ به $2n - 2 + 2n + 1 + n = 3n^2 + 2n - 2$ عمل حسابی نیاز داریم. نسخه‌ای از درونیابی لاگرانژ به نام فرمول گرانیگاهی [۶، ۱] وجود دارد که به طور چشمگیری حجم محاسبات را کاهش می‌دهد. در اینجا، این فرمول را ارائه می‌کنیم. فرض کنید

$$\lambda_i = \frac{1}{w'(x_i)}, \quad i = 0, 1, \dots, n.$$

در این صورت، به کمک رابطه (۲.۳) خواهیم داشت

$$P(x) = w(x) \sum_{i=0}^n \frac{\lambda_i}{x - x_i} f(x_i). \quad (3.3)$$

فرض کنید $f(x) = 1$. در این صورت، چند جمله‌ای درونیاب تابع f با تابع f برابر خواهد بود. بنابراین، داریم

$$1 = f(x) = P(x) = w(x) \sum_{i=0}^n \frac{\lambda_i}{x - x_i},$$

که نتیجه می‌دهد

$$w(x) = \frac{1}{\sum_{i=0}^n \frac{\lambda_i}{x - x_i}}.$$

با جای‌گذاری $w(x)$ در رابطه (۳.۳)، چندجمله‌ای P به صورت

$$P(x) = \frac{\sum_{i=0}^n \frac{\lambda_i}{x - x_i} f(x_i)}{\sum_{i=0}^n \frac{\lambda_i}{x - x_i}},$$

نوشته می‌شود. به این ترتیب، می‌توان دید که تعداد اعمال حسابی برای محاسبه این چندجمله‌ای در یک نقطه داده‌شده برابر با $4 + 7n + 2n^2$ است. با مقایسه این مقدار و تعداد اعمال حسابی در روش لاگرانژ می‌بینیم که استفاده از شکل گرانیگاهی کم‌هزینه‌تر است.

در ادامه، الگوریتم دیگری با نام الگوریتم نویل را معرفی می‌کنیم. فرض کنید $n + 1$ نقطه متمایز $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ ، از تابع f داده شده باشد. همچنین، فرض کنید $P_{i_0 i_1 \dots i_{k-1}}$ چندجمله‌ای درونیاب تابع f در نقاط $(x_{i_j}, f(x_{i_j}))$ ، $j = 0, 1, \dots, k-1$ ، و $P_{i_1 i_2 \dots i_k}$ چندجمله‌ای درونیاب تابع f در نقاط $(x_{i_j}, f(x_{i_j}))$ ، $j = 1, \dots, k$ ، قرار می‌دهیم

$$P_{i_0 i_1 \dots i_k}(x) = \frac{(x - x_{i_0})P_{i_1 i_2 \dots i_k}(x) - (x - x_{i_k})P_{i_0 i_1 \dots i_{k-1}}(x)}{x_{i_k} - x_{i_0}}. \quad (4.3)$$

روشن است که $P_{i_0 i_1 \dots i_k}(x)$ یک چندجمله‌ای از درجه حداکثر k است و داریم

$$P_{i_0 i_1 \dots i_k}(x_{i_0}) = P_{i_0 i_1 \dots i_{k-1}}(x_{i_0}) = f(x_{i_0}),$$

$$P_{i_0 i_1 \dots i_k}(x_{i_k}) = P_{i_1 i_2 \dots i_k}(x_{i_k}) = f(x_{i_k}),$$

و برای $j = 1, 2, \dots, k-1$ ،

$$\begin{aligned} P_{i_0 i_1 \dots i_k}(x_{i_j}) &= \frac{(x_{i_j} - x_{i_0})P_{i_1 i_2 \dots i_k}(x_{i_j}) - (x_{i_j} - x_{i_k})P_{i_0 i_1 \dots i_{k-1}}(x_{i_j})}{x_{i_k} - x_{i_0}} \\ &= \frac{(x_{i_j} - x_{i_0})f(x_{i_j}) - (x_{i_j} - x_{i_k})f(x_{i_j})}{x_{i_k} - x_{i_0}} = f(x_{i_j}). \end{aligned}$$

بنابراین، $P_{i_0 i_1 \dots i_k}(x)$ چندجمله‌ای درونیاب تابع f در نقاط $(x_{i_j}, f(x_{i_j}))$ ، $j = 0, 1, \dots, n$ ، است. در الگوریتم نویل، نخست چندجمله‌ای درونیاب تابع (درجه صفر) در هر یک از نقاط نوشته می‌شود، سپس با استفاده از رابطه بازگشتی (۴.۳) چندجمله‌ای درونیاب در تمام نقاط به دست می‌آید. جدول زیر (به‌ازای $n = 3$) شیوه به‌کارگیری الگوریتم نویل را نشان می‌دهد.

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
x_0	$f_0 = P_0(x) \searrow$			
x_1	$f_1 = P_1(x) \swarrow \searrow$	$P_{0,1}(x) \searrow$		
x_2	$f_2 = P_2(x) \swarrow \searrow$	$P_{1,2}(x) \swarrow \searrow$	$P_{0,1,2}(x) \searrow$	
x_3	$f_3 = P_3(x) \rightarrow$	$P_{2,3}(x) \rightarrow$	$P_{1,2,3}(x) \rightarrow$	$P_{0,1,2,3}(x)$

مثال ۲.۳ مقدار تقریبی $f(2)$ را با استفاده از سه مقدار تابع f که در جدول زیر آمده است، محاسبه کنید.

i	0	1	2
x_i	0	1	3
$f(x_i)$	1	3	2

حل: با استفاده از الگوریتم نویل، جدول زیر به دست می آید.

	$k = 0$	$k = 1$	$k = 2$
$x_0 = 0$	$f_0 = P_0(2) = 1 \searrow$		
$x_1 = 1$	$f_1 = P_1(2) = 3 \swarrow \searrow$	$P_{0,1}(2) = 5 \searrow$	
$x_2 = 3$	$f_2 = P_2(2) = 2 \rightarrow$	$P_{1,2}(2) = \frac{5}{3} \rightarrow$	$P_{0,1,2}(2) = \frac{10}{3}$

توجه کنید که مقادیر جدول به صورت زیر به دست آمده است

$$P_{0,1}(2) = \frac{(2 - x_0)P_1(2) - (2 - x_1)P_0(2)}{x_1 - x_0} = \frac{(2 - 0) \times 3 - (2 - 1) \times 1}{1 - 0} = 5,$$

$$P_{1,2}(2) = \frac{(2 - x_1)P_2(2) - (2 - x_2)P_1(2)}{x_2 - x_1} = \frac{(2 - 1) \times 2 - (2 - 3) \times 3}{3 - 1} = \frac{5}{2},$$

$$P_{0,1,2}(2) = \frac{(2 - x_0)P_{1,2}(2) - (2 - x_2)P_{0,1}(2)}{x_2 - x_0} = \frac{(2 - 0) \times \frac{5}{2} - (2 - 3) \times 5}{3 - 0} = \frac{10}{3}.$$

یک صورت دیگر الگوریتم نویل را که برای برنامه نویسی مناسبتر است، می توان به صورت زیر جمع بندی کرد. از رابطه (۴.۳) داریم

$$P_{j,j+1,\dots,j+k-1,j+k}(x) = \frac{(x - x_j)P_{j+1,\dots,j+k} - (x - x_{j+k})P_{j,\dots,j+k-1}}{x_{j+k} - x_j}. \quad (5.3)$$

قرار می دهیم

$$T_{j+k,k} := P_{j,j+1,\dots,j+k}. \quad (6.3)$$

در این صورت، رابطه (۵.۳) به صورت زیر نوشته می‌شود

$$T_{j+k,k}(x) = \frac{(x-x_j)T_{j+k,k-1} - (x-x_{j+k})T_{j+k-1,k-1}}{x_{j+k} - x_j}$$

اینک قرار می‌دهیم $j = i - k$. بنابراین، خواهیم داشت

$$T_{i,k}(x) = \frac{(x-x_{i-k})T_{i,k-1} - (x-x_i)T_{i-1,k-1}}{x_i - x_{i-k}} \quad (۷.۳)$$

در این صورت، جدول الگوریتم نویل به صورت زیر خواهد بود (به‌ازای $n = ۳$).

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
x_0	$f_0 = T_{00} \searrow$			
x_1	$f_1 = T_{10} \searrow$	$T_{11} \searrow$		
x_2	$f_2 = T_{20} \searrow$	$T_{21} \searrow$	$T_{22} \searrow$	
x_3	$f_3 = T_{30} \rightarrow$	$T_{31} \rightarrow$	$T_{32} \rightarrow$	T_{33}

با توجه به رابطه (۷.۳) و الگوریتم نویل، داریم

$$\begin{aligned} T_{i0} &= f(x_i), \\ T_{ik} &= \frac{(x-x_{i-k})T_{i,k-1} - (x-x_i)T_{i-1,k-1}}{x_i - x_{i-k}} \\ &= T_{i,k-1} + \frac{T_{i,k-1} - T_{i-1,k-1}}{\frac{x-x_{i-k}}{x-x_i} - 1}. \end{aligned} \quad (۸.۳)$$

در محاسبه T_{ik} با استفاده از رابطه اخیر، تنها دو عمل تقسیم نیاز است، در حالی که در رابطه (۷.۳) به دو عمل ضرب و یک عمل تقسیم نیاز است. بنابراین، توصیه می‌شود که از رابطه (۸.۳) برای محاسبه T_{ik} ها استفاده شود.

شکل دیگر الگوریتم نویل را می‌توان به صورت زیر طرح‌ریزی کرد، که دقت مقدار چندجمله‌ای درونیاب را کمی بهبود می‌بخشد. رابطه (۷.۳) را در نظر بگیرید. برای هر $i = 0, 1, \dots, n$ ، قرار می‌دهیم

$$\begin{aligned} Q_{i0} &= D_{i0} = f(x_i), \\ Q_{ik} &= T_{ik} - T_{i,k-1}, \quad k = 1, 2, \dots, i, \\ D_{ik} &= T_{ik} - T_{i-1,k-1} \quad k = 1, 2, \dots, i. \end{aligned} \quad (۹.۳)$$

شکل زیر، ارتباط بین T_{ij} و Q_{ij} ، D_{ij} را نشان می‌دهد.

$$\begin{array}{ccc}
 T_{i-1,k-1} & \searrow^{D_{ik}} & \\
 T_{i,k-1} & \xrightarrow{Q_{ik}} & T_{ik}
 \end{array}$$

در این صورت، داریم

$$\begin{aligned}
 D_{i,k-1} - Q_{i-1,k-1} &= T_{i,k-1} - T_{i-1,k-2} - T_{i-1,k-1} + T_{i-1,k-2} \\
 &= T_{i,k-1} - T_{i-1,k-1} \\
 &= (T_{ik} - T_{i,k-1}) \left(\frac{x - x_{i-k}}{x - x_i} - 1 \right) \\
 &= Q_{ik} \frac{x_i - x_{i-k}}{x - x_i}.
 \end{aligned}$$

بنابراین، خواهیم داشت

$$Q_{ik} = (D_{i,k-1} - Q_{i-1,k-1}) \frac{x_i - x}{x_{i-k} - x_i}. \quad (10.3)$$

به همین ترتیب، می‌توان نشان داد

$$D_{ik} = (D_{i,k-1} - Q_{i-1,k-1}) \frac{x_{i-k} - x}{x_{i-k} - x_i}. \quad (11.3)$$

در نتیجه، با استفاده از رابطه‌های (۱۰.۳) و (۱۱.۳) به صورت بازگشتی می‌توان Q_{ik} ها و D_{ik} ها را محاسبه کرد. با استفاده از رابطه (۹.۳)، داریم

$$\begin{aligned}
 T_{nn} &= Q_{nn} + T_{n,n-1} \\
 &= Q_{nn} + Q_{n,n-1} + T_{n,n-2} \\
 &\vdots \\
 &= \sum_{k=1}^n Q_{nk} + T_{n0} \\
 &= f(x_n) + \sum_{k=1}^n Q_{nk}.
 \end{aligned} \quad (12.3)$$

اگر $f(x_i)$ ها نزدیک به هم باشند، آنگاه مقادیر Q_{ik} ها و D_{ik} ها کوچک خواهد بود. بنابراین، برخلاف الگوریتم نویل، در محاسبه T_{nn} ، نخست Q_{nk} ها با یکدیگر جمع شده و نتیجه نهایی به $f(x_n)$ اضافه می‌شود.

به‌طور مشابه، داریم

$$\begin{aligned}
 T_{nn} &= D_{nn} + T_{n-1, n-1} \\
 &= D_{nn} + D_{n-1, n-1} + T_{n-2, n-2} \\
 &\vdots \\
 &= f(x_0) + \sum_{k=1}^n D_{kk}.
 \end{aligned} \tag{۱۳.۳}$$

مثال ۳.۳ مثال ۲.۳ را با کمک الگوریتم ارائه‌شدهٔ اخیر حل می‌کنیم. جدول‌های مربوط به مقادیر D_{ik} و Q_{ik} به صورت زیر هستند.

	Q_{i0}	Q_{i1}	Q_{i2}		D_{i0}	D_{i1}	D_{i2}
x_0	$Q_{00} \searrow$			x_0	$D_{00} \searrow$		
x_1	$Q_{10} \swarrow \searrow$	$Q_{11} \searrow$		x_1	$D_{10} \swarrow \searrow$	$D_{11} \searrow$	
x_2	$Q_{20} \rightarrow$	$Q_{21} \rightarrow$	Q_{22}	x_2	$D_{20} \rightarrow$	$D_{21} \rightarrow$	D_{22}

قرار می‌دهیم $Q_{0,0} = D_{0,0} = 1$ ، $Q_{1,0} = D_{1,0} = 3$ و $Q_{2,0} = D_{2,0} = 2$. سایر D_{ik} و Q_{ik} ‌ها را به صورت زیر محاسبه می‌کنیم

$$\begin{aligned}
 Q_{1,1} &= (D_{1,0} - Q_{0,0}) \frac{x_1 - x_0}{x_0 - x_1} = (3 - 1) \times \frac{1 - 0}{0 - 1} = 2, \\
 D_{1,1} &= (D_{1,0} - Q_{0,0}) \frac{x_0 - x_1}{x_0 - x_1} = (3 - 1) \times \frac{0 - 1}{0 - 1} = 4, \\
 Q_{2,1} &= (D_{2,0} - Q_{1,0}) \frac{x_2 - x_1}{x_1 - x_2} = (2 - 3) \times \frac{3 - 1}{1 - 3} = \frac{1}{2}, \\
 D_{2,1} &= (D_{2,0} - Q_{1,0}) \frac{x_1 - x_2}{x_1 - x_2} = (2 - 3) \times \frac{1 - 3}{1 - 3} = -\frac{1}{2}, \\
 Q_{2,2} &= (D_{2,1} - Q_{1,1}) \frac{x_2 - x_1}{x_0 - x_2} = (-\frac{1}{2} - 2) \times \frac{3 - 1}{0 - 3} = \frac{5}{6}, \\
 D_{2,2} &= (D_{2,1} - Q_{1,1}) \frac{x_0 - x_2}{x_0 - x_2} = (-\frac{1}{2} - 2) \times \frac{0 - 3}{0 - 3} = -\frac{5}{3}.
 \end{aligned}$$

بنابراین، جدول‌های متناظر با D_{ik} و Q_{ik} ‌ها به صورت زیر هستند.

	Q_{i0}	Q_{i1}	Q_{i2}		D_{i0}	D_{i1}	D_{i2}
$x_0 = 0$	$1 \searrow$			$x_0 = 0$	$1 \searrow$		
$x_1 = 1$	$3 \swarrow \searrow$	$2 \searrow$		$x_1 = 1$	$3 \swarrow \searrow$	$4 \searrow$	
$x_2 = 3$	$2 \rightarrow$	$\frac{1}{2} \rightarrow$	$\frac{5}{6}$	$x_2 = 3$	$2 \rightarrow$	$-\frac{1}{2} \rightarrow$	$-\frac{5}{3}$

در نتیجه، با استفاده از رابطه (۱۲.۳)، داریم

$$f(2) \approx f(x_2) + Q_{2,1} + Q_{2,2} = 2 + \frac{1}{4} + \frac{5}{6} = \frac{10}{3}.$$

به طور مشابه، با کمک رابطه (۱۳.۳)، خواهیم داشت

$$f(2) \approx f(x_0) + D_{1,1} + D_{2,2} = 1 + 4 - \frac{5}{3} = \frac{10}{3}.$$

صورت دیگر الگوریتم نویل، الگوریتم ایتنکن است که به شیوه دیگری از چندجمله‌ای‌های درونیاب میانی استفاده می‌کند. جدول زیر گویای همه چیز است (به ازای $n = 4$).

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
x_0	$f_0 = P_0(x)$				
x_1	$f_1 = P_1(x)$	$P_{0,1}(x)$			
x_2	$f_2 = P_2(x)$	$P_{0,2}(x)$	$P_{0,1,2}(x)$		
x_3	$f_3 = P_3(x)$	$P_{0,3}(x)$	$P_{0,1,3}(x)$	$P_{0,1,2,3}(x)$	
x_4	$f_4 = P_4(x)$	$P_{0,4}(x)$	$P_{0,1,4}(x)$	$P_{0,1,2,4}(x)$	$P_{0,1,2,3,4}(x)$

در این جدول، مقدار $P_{0,1,2,\dots,nk}(x)$ از رابطه زیر محاسبه می‌شود

$$P_{0,1,2,\dots,nk}(x) = \frac{(x - x_n)P_{0,1,\dots,(n-1)k}(x) - (x - x_k)P_{0,1,\dots,n}(x)}{x_k - x_n}.$$

۲.۳ تفاضلات تقسیم‌شده و چندجمله‌ای درونیاب نیوتن

در اینجا چندجمله‌ای درونیاب را به صورت

$$P(x) = P_{0,1,2,\dots,n}(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0) \cdots (x - x_{n-1}),$$

در نظر می‌گیریم. کافی است a_i ها را به گونه‌ای محاسبه کنیم که P در شرایط درونیابی

$$P(x_i) = f(x_i), \quad i = 0, 1, \dots, n,$$

صدق کند. پیش از هر چیز، این نکته را یادآوری می‌کنیم که برای محاسبه P در نقطه‌ای مثل $x = \xi$ ، این چندجمله‌ای را به صورت

$$P(\xi) = (\dots (a_n(\xi - x_{n-1}) + a_{n-1})(\xi - x_{n-2}) + \dots + a_1)(\xi - x_0) + a_0,$$

می‌نویسیم. سپس برای محاسبه $P(\xi)$ می‌توان از الگوریتم زیر استفاده کرد که به الگوریتم هورنر معروف است

$$p := a_n(\xi - x_{n-1}) + a_{n-1}$$

For $k := n - 2, n - 3, \dots, 0$, Do

$$p := p * (\xi - x_k) + a_k$$

EndDo

اکنون، ببینیم که چگونه می‌توان a_i ها را محاسبه کرد. هریک از چندجمله‌ای‌های $P_{i_0 i_1 \dots i_k}(x)$ و $P_{i_0 i_1 \dots i_{k-1}}(x)$ درونیاب تابع f در نقاط $x_{i_0}, \dots, x_{i_{k-1}}$ هستند. در نتیجه، داریم

$$P_{i_0 i_1 \dots i_k}(x) = P_{i_0 i_1 \dots i_{k-1}}(x) + \alpha(x - x_{i_0}) \dots (x - x_{i_{k-1}}).$$

فرض کنید $\alpha = f[x_{i_0}, x_{i_1}, \dots, x_{i_k}]$ در این صورت

$$P_{i_0 i_1 \dots i_k}(x) = f[x_{i_0}] + f[x_{i_0}, x_{i_1}](x - x_{i_0}) + f[x_{i_0}, x_{i_1}, x_{i_2}](x - x_{i_0})(x - x_{i_1}) + \dots + f[x_{i_0}, x_{i_1}, \dots, x_{i_k}](x - x_{i_0})(x - x_{i_1}) \dots (x - x_{i_{k-1}}). \quad (14.3)$$

عدد $f[x_{i_0}, x_{i_1}, \dots, x_{i_k}]$ را k امین تفاضل تقسیم‌شده می‌گویند. در ادامه، یک رابطه بازگشتی برای محاسبه k امین تفاضل تقسیم‌شده به دست می‌آوریم. در الگوریتم نوئل، دیدیم که

$$P_{i_0 i_1 \dots i_k}(x) = \frac{(x - x_{i_0})P_{i_1 i_2 \dots i_k}(x) - (x - x_{i_k})P_{i_0 i_1 \dots i_{k-1}}(x)}{x_{i_k} - x_{i_0}}. \quad (15.3)$$

اکنون، با توجه به رابطه (۱۴.۳)، ضرب x^k برابر با $f[x_{i_0}, x_{i_1}, \dots, x_{i_k}]$ ، و در سمت راست رابطه (۱۵.۳) برابر با

$$\frac{f[x_{i_1}, x_{i_2}, \dots, x_{i_k}] - f[x_{i_0}, x_{i_1}, \dots, x_{i_{k-1}}]}{x_{i_k} - x_{i_0}},$$

است. بنابراین، داریم

$$f[x_{i_0}, x_{i_1}, \dots, x_{i_k}] = \frac{f[x_{i_1}, x_{i_2}, \dots, x_{i_k}] - f[x_{i_0}, x_{i_1}, \dots, x_{i_{k-1}}]}{x_{i_k} - x_{i_0}}.$$

با استفاده از جدول زیر، همه مقادیر تفاضلات تقسیم‌شده را می‌توان محاسبه کرد ($n = 3$).

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
x_0	$f[x_0]$			
		$f[x_0, x_1]$		
x_1	$f[x_1]$		$f[x_0, x_1, x_2]$	
		$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$
x_2	$f[x_2]$		$f[x_1, x_2, x_3]$	
		$f[x_2, x_3]$		
x_3	$f[x_3]$			

توجه کنید که در این جدول، داریم $f_i = f[x_i]$ ، $i = 0, 1, \dots, n$. اکنون، با استفاده از این جدول، چندجمله‌ای $P_n(x)$ را می‌توان به صورت

$$P(x) = P_{0,1,\dots,n}(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots \\ + f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}),$$

به دست آورد. این چندجمله‌ای به چندجمله‌ای درونیاب نیوتن معروف است.

یادداشت ۱.۳ تفاضلات تقسیم‌شده f_{i_0, i_1, \dots, i_k} تحت هر جایگشت از i_0, i_1, \dots, i_k بدون تغییر باقی می‌ماند. یعنی، اگر (j_0, j_1, \dots, j_k) یک جایگشت از (i_0, i_1, \dots, i_k) باشد، آنگاه

$$f[x_{j_0}, x_{j_1}, \dots, x_{j_k}] = f[x_{i_0}, x_{i_1}, \dots, x_{i_k}].$$

این واقعیت از آنجایی که چندجمله‌ای درونیاب یکتا است، نتیجه می‌شود.

قضیه ۲.۳ اگر f یک چندجمله‌ای از درجه n باشد، آنگاه

$$f[x_0, \dots, x_k] = 0, \quad k > n.$$

برهان: چندجمله‌ای $P_{0,1,\dots,k}(x)$ را در نظر بگیرید. این چندجمله‌ای از درجه حداکثر k است و با توجه به یکتایی چندجمله‌ای درونیاب، داریم $P_{0,1,\dots,k}(x) = f(x)$. از طرفی، ضریب x^k در $P_{0,1,\dots,k}(x)$ برابر با $f[x_0, x_1, \dots, x_k]$ و این ضریب در f ، برابر با صفر است. در نتیجه، $f[x_0, x_1, \dots, x_k] = 0$.

مثال ۴.۳ برای روشن شدن قضیه ۲.۳، جدول تفاضلات تقسیم‌شده تابع $f(x) = x^2$ را در پنج نقطه محاسبه کرده‌ایم که در جدول زیر نمایش داده شده است. همان‌گونه که ملاحظه می‌شود، به ازای $k > 2$ ، داریم $f[x_0, \dots, x_k] = 0$.

x_i	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
0	0				
1	1	1			
2	4	3	1	0	0
3	9	5	1	0	0
4	16	7	1	0	0

مثال ۵.۳ را با استفاده از روش نیوتن حل می‌کنیم. جدول تفاضلات تقسیم‌شده برای داده‌ها به صورت زیر است.

	$k = 0$	$k = 1$	$k = 2$
$x_0 = 0$	$f[x_0] = 1$	$f[x_0, x_1] = 2$	$f[x_0, x_1, x_2] = -\frac{5}{6}$
$x_1 = 1$	$f[x_1] = 3$	$f[x_1, x_2] = -\frac{1}{6}$	
$x_2 = 3$	$f[x_2] = 2$		

توجه کنید که در این جدول، تفاضلات تقسیم‌شده به صورت زیر محاسبه شده‌اند

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{3 - 1}{1 - 0} = 2,$$

$$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{2 - 3}{3 - 1} = -\frac{1}{2},$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{-\frac{1}{2} - 2}{3} = -\frac{5}{6}.$$

در نتیجه، چندجمله‌ای درونیاب به صورت

$$P(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)$$

$$= 1 + 2x - \frac{5}{6}x(x - 1).$$

خواهد بود و بنابراین، داریم $f(2) \approx P(2) = \frac{10}{3}$.

۳.۳ خطای درونیابی

قضیه ۳.۳ فرض کنید $P \in \Pi_n$ چندجمله‌ای درونیاب تابع f در نقاط متمایز $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ باشد و تابع f مشتق مرتبه $(n+1)$ ام دارد. همچنین، فرض کنید $\mathcal{I} = I[x_0, x_1, \dots, x_n]$ کوچک‌ترین بازه‌ای باشد که شامل تمام نقاط گرهی است. در این صورت، به ازای هر $\bar{x} \in \mathcal{I}$ عددی مثل $\xi_{\bar{x}} \in I[x_0, x_1, \dots, x_n, \bar{x}]$ وجود دارد به طوری که

$$f(\bar{x}) - P(\bar{x}) = \frac{f^{(n+1)}(\xi_{\bar{x}})}{(n+1)!} w(\bar{x}), \quad (16.3)$$

$$w(x) = \prod_{i=0}^n (x - x_i)$$

که در آن

برهان: برای محاسبه خطا، از تابع $E(x) = f(x) - P(x)$ استفاده می‌کنیم. هدف، محاسبه $E(\bar{x})$ است. قرار می‌دهیم

$$g(x) = f(x) - P(x) - E(\bar{x}) \frac{(x - x_0) \cdots (x - x_n)}{(\bar{x} - x_0) \cdots (\bar{x} - x_n)}.$$

به سادگی می‌توان دید که

$$\begin{cases} g(x_i) = 0, & i = 0, 1, \dots, n, \\ g(\bar{x}) = f(\bar{x}) - P(\bar{x}) - E(\bar{x}) = 0. \end{cases}$$

بنابراین، تابع g دارای $n+2$ ریشه است. در نتیجه، با استفاده از قضیه رُل داریم

$$\exists \xi_{\bar{x}} \in I[x_0, x_1, \dots, x_n, \bar{x}] : g^{(n+1)}(\xi_{\bar{x}}) = 0,$$

که در آن اندیس \bar{x} برای ξ ، وابستگی ξ به \bar{x} را نشان می‌دهد. ولی، داریم

$$0 = g^{(n+1)}(\xi_{\bar{x}}) = f^{(n+1)}(\xi_{\bar{x}}) - 0 - E(\bar{x}) \frac{(n+1)!}{w(\bar{x})},$$

□ که نتیجه لازم را به دست می‌دهد.

شکل دیگر نمایش خطا را می‌توان به صورت زیر به دست آورد. فرض کنید $P(x)$ چندجمله‌ای درونیاب تابع f در نقاط x_0, \dots, x_n و $\bar{P}(x)$ چندجمله‌ای درونیاب تابع f در نقاط x_0, \dots, x_n, \bar{x} باشد. داریم

$$\bar{P}(x) = P(x) + f[x_0, \dots, x_n, \bar{x}] (x - x_0) \cdots (x - x_n).$$

در این صورت

$$\bar{P}(\bar{x}) = P(\bar{x}) + f[x_0, \dots, x_n, \bar{x}](\bar{x} - x_0) \cdots (\bar{x} - x_n).$$

با استفاده از این رابطه و اینکه $\bar{P}(\bar{x}) = f(\bar{x})$ نتیجه می‌گیریم

$$E(\bar{x}) = f(\bar{x}) - P(\bar{x}) = f[x_0, \dots, x_n, \bar{x}](\bar{x} - x_0) \cdots (\bar{x} - x_n).$$

با مقایسه این رابطه و رابطه (۱۶.۳)، داریم

$$f[x_0, \dots, x_n, \bar{x}] = \frac{f^{(n+1)}(\xi_{\bar{x}})}{(n+1)!}, \quad (17.3)$$

و

$$E(\bar{x}) = f[x_0, \dots, x_n, \bar{x}]w(\bar{x}).$$

مثال ۶.۳ تابع $f(x) = \sin \frac{\pi}{4}x$ را در بازه $[-1, 1]$ و در نقاط گرهی

$$x_i = -1 + \frac{2i}{n}, \quad i = 0, 1, \dots, n,$$

درونیابی می‌کنیم. نشان دهید که اگر $P_n(x)$ چندجمله‌ای درونیاب تابع باشد، آنگاه خواهیم داشت

$$\lim_{n \rightarrow \infty} P_n(x) = f(x), \quad x \in [-1, 1].$$

حل: داریم

$$f^{(n+1)}(x) = \left(\frac{\pi}{4}\right)^{n+1} \sin\left(\frac{(n+1)\pi}{4} + \frac{\pi}{4}x\right),$$

و بنابراین

$$|f^{(n+1)}(x)| \leq \left(\frac{\pi}{4}\right)^{n+1}, \quad x \in [-1, 1].$$

از طرفی، با توجه به اینکه $|x - x_i| \leq 2$ ، نتیجه می‌گیریم که $|w(x)| \leq 2^{n+1}$. بنابراین، با توجه به قضیه ۳.۳ خواهیم داشت

$$|E_n(x)| = |P_n(x) - f(x)| \leq \frac{\left(\frac{\pi}{4}\right)^{n+1}}{(n+1)!} \times 2^{n+1} = \frac{\pi^{n+1}}{(n+1)!} \rightarrow 0.$$

در اینجا دو نکته را یادآور می‌شویم. نخست اینکه تا حد امکان باید از محاسبه چندجمله‌ای درونیاب به ازای یک مقدار در خارج از بازه $I[x_0, x_1, \dots, x_n]$ پرهیز کرد و دوم اینکه با توجه به

قضیه ۳.۳، به طور کلی خطای درونیابی به $w(x)$ و مشتق $(n+1)$ ام تابع f بستگی دارد. بنابراین، در حالت کلی نمی توان انتظار داشت که با افزایش تعداد نقاط درونیابی، خطای درونیابی نیز کاهش یابد. یک نمونه از این گونه توابع، تابع رونگه است که به صورت

$$f(x) = \frac{1}{1+x^2}, \quad x \in [-5, 5],$$

تعریف می شود. در شکل ۱.۳، تابع f به همراه چند جمله ای درونیاب متناظر به ازای n های مختلف (زیربازه های مساوی) نمایش داده شده است. چنان که ملاحظه می شود، به ازای n های بزرگ تر، اختلاف تابع f و چند جمله ای درونیاب متناظر زیاد می شود.

۴.۳ مسئله اقل اکثر

همان گونه که در بخش قبل دیدیم، خطای درونیابی به صورت

$$E_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} w(x), \quad (18.3)$$

است. بنابراین، داریم

$$|E_n(x)| \leq \frac{1}{(n+1)!} \max_{x \in \mathcal{I}} |f^{(n+1)}(x)| \max_{x \in \mathcal{I}} |w(x)| \leq \frac{M}{(n+1)!} \max_{x \in \mathcal{I}} |w(x)|,$$

که در آن M یک کران بالا برای $\max_{x \in \mathcal{I}} |f^{(n+1)}(x)|$ است. اکنون می خواهیم x_0, \dots, x_n را طوری بیابیم که $\max_{x \in \mathcal{I}} |w(x)|$ کمترین مقدار ممکن را داشته باشد. این مسئله، به مسئله اقل اکثر معروف است. پاسخ به این پرسش با استفاده از چند جمله ای های چبیشف امکان پذیر است. نخست به معرفی این چند جمله ای ها می پردازیم و سپس خواص آن را بررسی می کنیم.

ابتدا چند جمله ای های چبیشف را برای بازه $[-1, 1]$ در نظر می گیریم که به صورت

$$T_n(x) = \cos(n \cos^{-1}(x)), \quad n \geq 0,$$

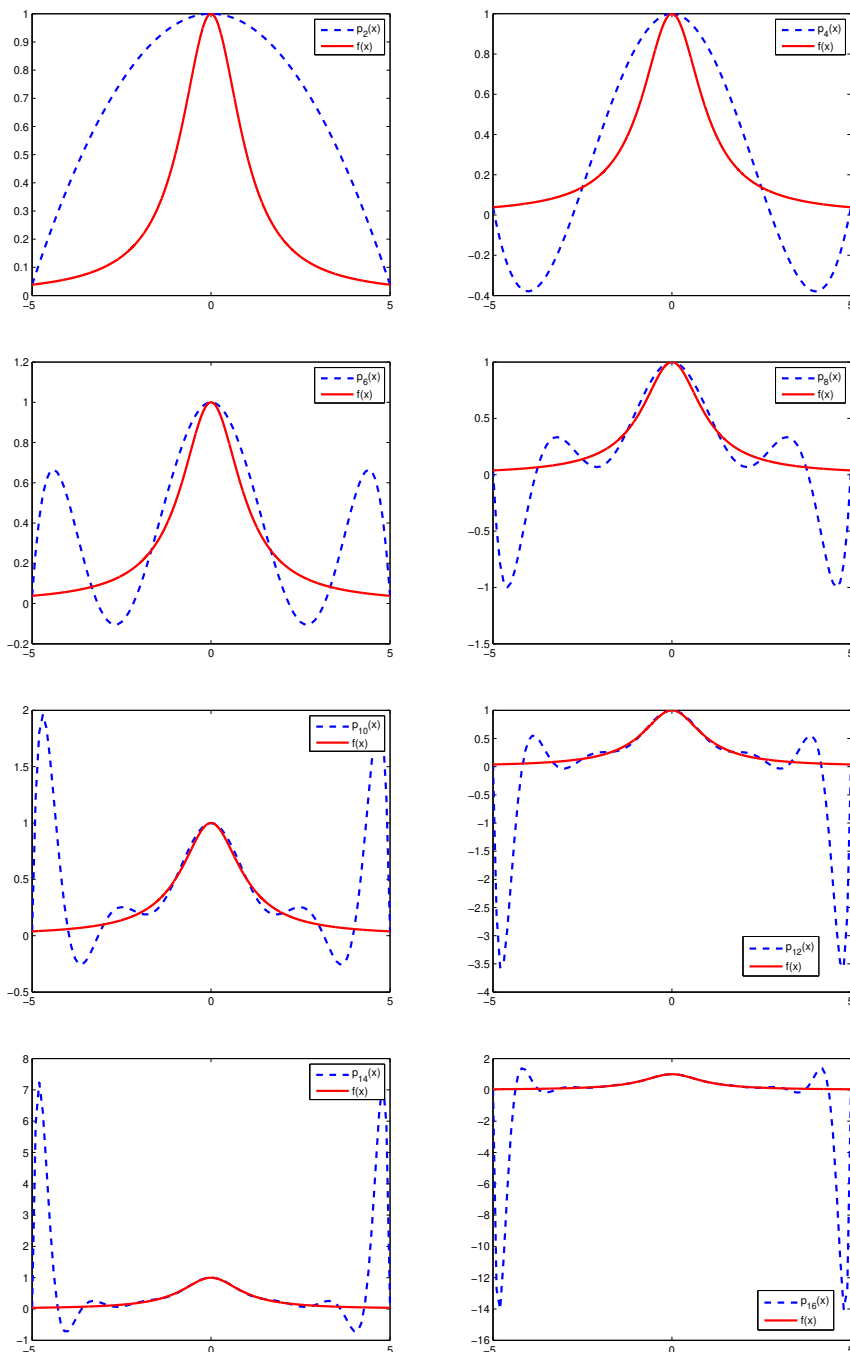
تعریف می شوند. فرض کنید $x = \cos^{-1} \theta$. در این صورت، با جمع طرفین رابطه های مثلثاتی

$$\cos(n+1)\theta = \cos \theta \cos n\theta - \sin \theta \sin n\theta,$$

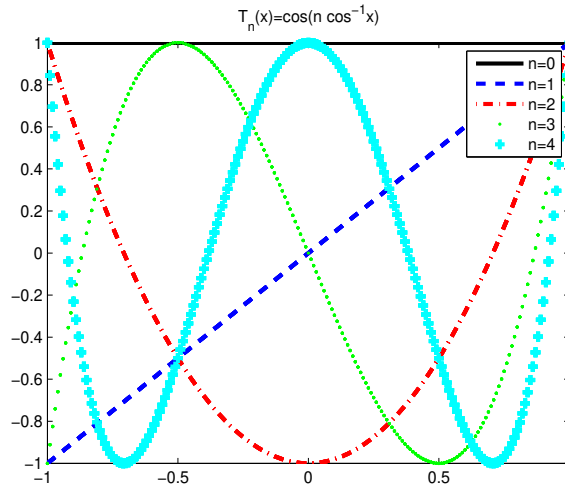
$$\cos(n-1)\theta = \cos \theta \cos n\theta + \sin \theta \sin n\theta,$$

داریم

$$\cos(n+1)\theta = 2 \cos \theta \cos n\theta - \cos(n-1)\theta,$$



شکل ۱.۳: درونیایی تابع رونگه.



شکل ۲.۳: نمایش چند جمله از چندجمله‌ای‌های چبیشف.

که با استفاده از $\cos \theta = x$ خواهیم داشت

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

این رابطه، یک رابطه بازگشتی برای محاسبه چندجمله‌ای‌های چبیشف است. توجه کنید که دو جمله نخست به صورت $T_0(x) = 1$ و $T_1(x) = x$ هستند. شکل ۲.۳، تعدادی از T_i ها را نمایش می‌دهد.

چند ویژگی مهم چندجمله‌ای‌های چبیشف به صورت زیر هستند.

(۱) T_n یک چندجمله‌ای از درجه n و ضریب x^n در T_n برابر با 2^{n-1} است و بنابراین، چندجمله‌ای $2^{1-n}T_n$ تکین است.

(۲) روشن است که $-1 \leq x \leq 1$ ، $|T_n(x)| \leq 1$.

(۳) ریشه‌های T_n : از $T_n(x) = 0$ نتیجه می‌شود که $\cos n\theta = 0$ و در نتیجه

$$\theta_k = \frac{(2k+1)\pi}{2n}, \quad k = 0, 1, \dots, n-1,$$

که با توجه به $x_k = \cos \theta_k$ ، ریشه‌های T_n به صورت

$$x_k = \cos \frac{(2k+1)\pi}{2n}, \quad k = 0, 1, \dots, n-1.$$

خواهد بود.

(۴) نقاط ماکسیمم و مینیمم T_n : با فرض $\theta = \cos^{-1} x$ داریم $T_n(x) = \cos(n\theta)$. با محاسبهٔ مشتق T_n نسبت به x ، داریم

$$\frac{dT_n}{dx} = \frac{dT_n}{d\theta} \cdot \frac{d\theta}{dx} = \frac{-n \sin n\theta}{-\sin \theta} = \frac{n \sin n\theta}{\sin \theta}.$$

اگر $\frac{dT_n}{dx} = 0$ ، آنگاه $\sin n\theta = 0$ و $\theta \neq 0$. در نتیجه

$$\theta_k = \frac{k\pi}{n}, \quad k = 1, 2, \dots, n-1,$$

و بنابراین، خواهیم داشت

$$x_k = \cos \frac{k\pi}{n}, \quad k = 1, 2, \dots, n-1.$$

به ازای x_k ، $k = 1, 2, \dots, n-1$ ، داریم

$$T(x_k) = T_n(\cos \frac{k\pi}{n}) = \cos n(\cos^{-1}(\cos \frac{k\pi}{n})) = \cos k\pi = (-1)^k.$$

از طرفی، داریم

$$T_n(\cos \frac{0 \times \pi}{n}) = T_n(1) = \cos(n \cos^{-1} 1) = \cos(n \times 0) = 1,$$

و

$$T_n(\cos \frac{n\pi}{n}) = T_n(\cos \pi) = T_n(-1) = \cos(n \cos^{-1}(-1)) = \cos(n\pi) = (-1)^n.$$

با جمع بندی مطالب خواهیم داشت

$$T_n(\cos \frac{k\pi}{n}) = (-1)^k, \quad k = 0, 1, \dots, n.$$

این رابطه نشان می دهد که $k = 0, 1, \dots, n$ ، $x_k = \cos \frac{k\pi}{n}$ نقاط ماکسیمم و مینیمم T_n را به دست می دهند.

قضیهٔ ۴.۳ اگر p یک چندجمله ای تکین از درجهٔ n باشد، آنگاه

$$\max_{-1 \leq x \leq 1} |p(x)| \geq 2^{1-n}.$$

برهان: به برهان خلف، فرض کنید

$$|p(x)| < 2^{1-n}, \quad -1 \leq x \leq 1.$$

قرار می‌دهیم $q \equiv 2^{1-n}T_n$ و $x_i = \cos \frac{i\pi}{n}$ ، $i = 0, 1, \dots, n$. در این صورت، q یک چندجمله‌ای تکین است و

$$q(x_i) = 2^{1-n}T_n(x_i) = 2^{1-n}(-1)^i,$$

یا $(-1)^i q(x_i) = 2^{1-n}$ داریم، از طرفی،

$$(-1)^i p(x_i) \leq |p(x_i)| < 2^{1-n} = (-1)^i q(x_i),$$

و در نتیجه

$$(-1)^i (q(x_i) - p(x_i)) > 0.$$

قرار می‌دهیم $r \equiv q - p$. در نتیجه، روشن است که درجه r حداکثر $n - 1$ است و به علاوه،

$$(-1)^i r(x_i) > 0, \quad i = 0, 1, \dots, n.$$

این رابطه نشان می‌دهد که r دارای n ریشه متمایز است. این نتیجه با اینکه درجه چندجمله‌ای r حداکثر $n - 1$ است، تناقض دارد. \square

اکنون به مسئله انتخاب بهترین نقاط گرهی پاسخ می‌دهیم. ابتدا توجه کنید که $w(x)$ یک چندجمله‌ای تکین از درجه $n + 1$ است و بنابراین، با توجه به قضیه ۴.۳، داریم

$$\max_{-1 \leq x \leq 1} |w(x)| \geq 2^{-n}.$$

از طرفی، $2^{-n}T_{n+1}$ یک چندجمله‌ای تکین از درجه $n + 1$ است و

$$\max_{-1 \leq x \leq 1} |2^{-n}T_{n+1}| = 2^{-n}.$$

از این رو، کمترین مقدار $|w(x)|$ به ازای ریشه‌های $2^{-n}T_{n+1}$ ، یا به عبارت دیگر، ریشه‌های T_{n+1} به دست می‌آید که این ریشه‌ها به صورت زیر هستند

$$x_k = \cos \frac{(2k+1)\pi}{2n+2}, \quad k = 0, 1, \dots, n,$$

اکنون، با استفاده از مطالب ذکر شده، قضیه زیر را بیان می‌کنیم.

قضیه ۵.۳. اگر x_j ، $j = 0, 1, \dots, n$ ، ریشه‌های چندجمله‌ای چبیشف T_{n+1} باشند، آنگاه

$$|E_n(x)| \leq \frac{M}{(n+1)!} \max_{-1 \leq x \leq 1} \left| \prod_{j=0}^n (x - x_j) \right| = \frac{M}{(n+1)!} 2^{-n} = \frac{M}{2^n(n+1)!}.$$

در ادامه، فرض می‌کنیم که بازه درونیابی به صورت $[a, b]$ باشد، که در آن a و b مقادیر دلخواه هستند. از تغییر متغیر

$$t = \frac{2x - (b + a)}{b - a}, \quad (19.3)$$

استفاده می‌کنیم. داریم

$$-1 \leq t \leq 1 \iff a \leq x \leq b.$$

اگر فرض کنیم t_j ، $j = 0, 1, \dots, n$ ، ریشه‌های چندجمله‌ای T_{n+1} در بازه $[-1, 1]$ باشند، آنگاه با تغییر متغیر بالا، این مقادیر در بازه $[a, b]$ به صورت

$$t_j = \frac{2x_j - (b + a)}{b - a}, \quad j = 0, 1, \dots, n. \quad (20.3)$$

خواهند بود. با استفاده از رابطه‌های (۱۹.۳) و (۲۰.۳)، داریم

$$x - x_j = \frac{b - a}{2}(t - t_j), \quad j = 0, 1, \dots, n.$$

در نتیجه، برای بازه $[a, b]$ ، داریم

$$\begin{aligned} |E_n(x)| &\leq \frac{M}{(n+1)!} \max_{a \leq x \leq b} \left| \prod_{j=0}^n (x - x_j) \right| \\ &= \frac{M}{(n+1)!} \max_{-1 \leq t \leq 1} \left| \prod_{j=0}^n \frac{b-a}{2}(t - t_j) \right| \\ &= \frac{M}{2^{n+1}(n+1)!} (b-a)^{n+1} \max_{-1 \leq t \leq 1} \left| \prod_{j=0}^n (t - t_j) \right| \\ &\leq \frac{M}{2^{n+1}(n+1)!} (b-a)^{n+1} 2^{-n} \\ &= \frac{M}{2^{2n+1}(n+1)!} (b-a)^{n+1}. \end{aligned}$$

با جمع‌بندی مطالب ذکرشده برای بازه $[a, b]$ ، بهترین نقاط برای درونیابی (با توجه به مسئله اقل اکثر) به صورت

$$x_j = \frac{b-a}{2}t_j + \frac{b+a}{2} = \frac{b-a}{2} \cos \frac{(2j+1)\pi}{2n+2} + \frac{b+a}{2}, \quad j = 0, 1, \dots, n, \quad (21.3)$$

است و حداکثر خطا در رابطه

$$|E_n(x)| \leq \frac{M}{2^{2n+1}(n+1)!} (b-a)^{n+1}, \quad (22.3)$$

صدق می‌کند که در آن M یک کران بالا برای مشتق $(n+1)$ ام تابع f در بازه $[a, b]$ است.

مثال ۷.۳ تابع

$$f(x) = \frac{1}{1 + (x - 1.5)^3}, \quad 1 \leq x \leq 2,$$

را در نظر بگیرید. می‌خواهیم این تابع را در سه نقطه در بازه $[1, 2]$ درونیابی کنیم. اگر نقاط را متساوی‌فاصله بگیریم، خواهیم داشت $x_0 = 1$ ، $x_1 = 1.5$ و $x_2 = 2$. در این صورت، چندجمله‌ای درونیاب به صورت

$$P_2(x) = 1.5238 - 0.4444x + 0.0635x^2,$$

خواهد بود و

$$\max_{1 \leq x \leq 2} |w(x)| = \max_{1 \leq x \leq 2} |(x - 1)(x - 1.5)(x - 2)| = 0.0481.$$

در نتیجه

$$|E_2(x)| \leq \frac{M}{3!} \max_{1 \leq x \leq 2} |w(x)| = 0.0080M,$$

که در آن M ماکسیمم تابع f''' در بازه $[1, 2]$ است.

اکنون، با کمک رابطه (۲۱.۳) نقاط درونیابی را به صورت

$$x_j = \frac{1}{3} \cos \frac{(2j+1)\pi}{2n+2} + \frac{2}{3}, \quad j = 0, 1, 2,$$

در نظر می‌گیریم. در این صورت، مجموعه نقاط درونیابی به صورت

$$x_0 = \frac{2}{3} - \frac{\sqrt{3}}{4} = 1.0670, \quad x_1 = 1.5, \quad x_2 = \frac{2}{3} + \frac{\sqrt{3}}{4} = 1.9330,$$

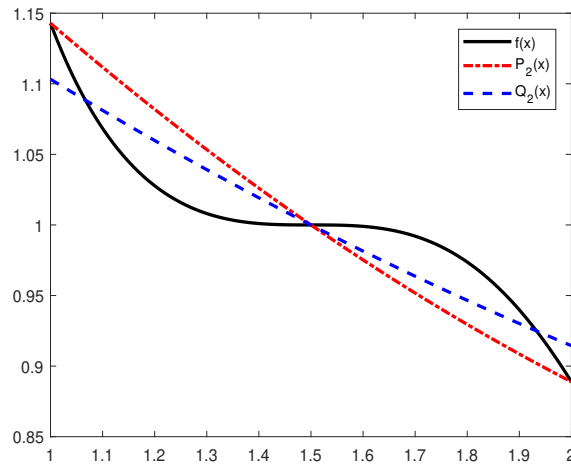
و چندجمله‌ای درونیاب به صورت

$$Q_2(x) = 1.3627 - 0.2949x + 0.0354x^2,$$

خواهد بود. با کمک رابطه (۲۲.۳)، کران بالای خطای درونیابی در رابطه

$$|E_2(x)| \leq \frac{M}{2^{2 \times 2 + 1} (2 + 1)!} (2 - 1)^{2+1} = 0.0052M,$$

صدق می‌کند. همان‌گونه که می‌بینیم، کران بالای خطای درونیابی با کمک ریشه‌های چندجمله‌ای چیبیشف، کوچک‌تر از کران بالای خطای درونیابی با استفاده از نقاط متساوی‌فاصله است. در شکل ۳.۳ نمودار تابع f به همراه چندجمله‌ای‌های P_2 و Q_2 در بازه $[1, 2]$ نمایش داده شده است. این شکل، به خوبی برتری دقت چندجمله‌ای Q_2 بر P_2 را نمایش می‌دهد.



شکل ۳.۳: تابع f و چندجمله‌ای‌های P_2 و Q_2 برای مثال ۷.۳.

۵.۳ درونیابی هرمیت

پیش از تعریف شکل کلی چندجمله‌ای درونیاب هرمیت، نخست حالتی خاص از آن را که کاربردهای متعددی دارد، معرفی می‌کنیم.

قضیه ۶.۳ فرض کنید $n + 1$ مقدار متمایز x_i , $i = 0, 1, \dots, n$ ، در بازه $[a, b]$ داده شده باشند. همچنین، فرض کنید $2n + 2$ مقدار y_i و y'_i در \mathbb{R} داده شده باشند. در این صورت، چندجمله‌ای یکتایی مثل P ، حداکثر از درجه $2n + 1$ وجود دارد چنان‌که

$$P(x_i) = y_i, \quad P'(x_i) = y'_i, \quad i = 0, 1, \dots, n.$$

برهان: قرار می‌دهیم

$$H_k^0(x) = [1 - 2L'_k(x_k)(x - x_k)][L_k(x)]^2,$$

$$H_k^1(x) = (x - x_k)[L_k(x)]^2,$$

که در آن L_k ها چندجمله‌ای‌هایی هستند که توسط رابطه (۱.۳) داده شده‌اند. روشن است که

$$H_k^0(x_j) = \delta_{jk}, \quad H_k^1(x_j) = 0.$$

همچنین

$$\frac{d}{dx} H_k^0(x) = -2L'_k(x_k)[L_k(x)]^2 + [1 - 2L'_k(x_k)(x - x_k)] \times 2L'_k(x)L_k(x),$$

و بنابراین، داریم

$$\frac{d}{dx} H_k^\circ(x_j) = 0, \quad \forall k, j.$$

از طرفی

$$\frac{d}{dx} H_k^\circ(x) = [L_k(x)]^2 + (x - x_k) \times 2L_k'(x)L_k(x),$$

و در نتیجه

$$\frac{d}{dx} H_k^\circ(x_j) = \delta_{jk}, \quad \forall k, j.$$

اکنون، قرار می‌دهیم

$$P(x) = \sum_{k=0}^n \left(y_k H_k^\circ(x) + y_k' H_k^\circ(x) \right).$$

به سادگی می‌توان دید که P در شرایط قضیه صدق می‌کند. برای نشان دادن یکتایی، فرض کنید P و Q دو چندجمله‌ای درونیاب باشند. قرار می‌دهیم $R \equiv P - Q$. در این صورت، داریم

$$\begin{cases} R(x_i) = P(x_i) - Q(x_i) = y_i - y_i = 0, \\ R'(x_i) = P'(x_i) - Q'(x_i) = y_i' - y_i' = 0, \end{cases} \quad i = 0, 1, \dots, n,$$

بنابراین، چندجمله‌ای R دارای $2n + 2$ ریشه با احتساب تکرار است. با توجه به اینکه R یک چندجمله‌ای حداکثر از درجه $2n + 1$ است، نتیجه می‌گیریم که $R \equiv 0$ و بنابراین، یکتایی نتیجه می‌شود. \square

قضیه ۷.۳ فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ تابعی باشد که $2n + 2$ بار در بازه $[a, b]$ مشتق پذیر است. در این صورت، اگر P چندجمله‌ای درونیاب هرमित تابع f در $n + 1$ نقطه متمایز $x_i \in [a, b]$ ، $i = 0, 1, \dots, n$ باشد، آنگاه یک $\xi_x \in I[x_0, \dots, x_n]$ موجود است چنان که

$$f(x) - P(x) = \frac{f^{(2n+2)}(\xi_x)}{(2n+2)!} \prod_{j=0}^n (x - x_j)^2, \quad \forall x \in [a, b].$$

برهان: فرض کنید $E(t) = f(t) - P(t)$ و x یک مقدار ثابت در بازه $[a, b]$ باشد. قرار می‌دهیم

$$g(t) = f(t) - P(t) - E(x) \frac{(t - x_0)^2 \dots (t - x_n)^2}{(x - x_0)^2 \dots (x - x_n)^2}.$$

روشن است که

$$\begin{cases} g(x_i) = g'(x_i) = 0, & i = 0, 1, \dots, n, \\ g(x) = f(x) - P(x) - E(x) = 0. \end{cases}$$

بنابراین، تابع g دارای $2n + 3$ ریشه است (با احتساب تکرار). در نتیجه، با استفاده از قضیهٔ رُل، داریم

$$\exists \xi_x \in I[x_0, x_1, \dots, x_n, x] : g^{(2n+2)}(\xi_x) = 0.$$

از این رو

$$0 = g^{(2n+2)}(\xi_x) = f^{(2n+2)}(\xi_x) - 0 - E(x) \frac{(2n+2)!}{w(x)^2},$$

□ که اثبات قضیه را کامل می‌کند.

مثال ۸.۳. مقادیری از توابع f و f' در جدول زیر داده شده‌اند. با استفاده از چندجمله‌ای درونیاب هرमित تابع f ، مقدار تقریبی $f(\frac{1}{2})$ را محاسبه کنید.

k	۰	۱	۲
x_k	-۱	۰	۱
y_k	۱	۱	۳
y'_k	-۵	۱	۷

سپس، یک کران بالا برای خطا به دست آورید.

حل: با محاسبهٔ چندجمله‌ای‌های لاگرانژ، خواهیم داشت

$$L_0(x) = \frac{(x-0)(x-1)}{(-1-0)(-1-1)} = \frac{1}{2}(x^2-1),$$

$$L_1(x) = \frac{(x+1)(x-1)}{(0+1)(0-1)} = 1-x^2,$$

$$L_2(x) = \frac{(x+1)(x-0)}{(1+1)(1-0)} = \frac{1}{2}(x^2+x).$$

همچنین، داریم

$$L'_0(x) = x, \quad L'_1(x) = -2x, \quad L'_2(x) = \frac{1}{2}(2x+1).$$

در نتیجه

$$H_0^\circ(x) = [1 - 2L'_0(x_0)(x-x_0)](L_0(x))^2 = \frac{1}{4}(2x+3)(x^2-1)^2,$$

$$H_1^\circ(x) = [1 - 2L'_1(x_1)(x-x_1)](L_1(x))^2 = (1-x^2)^2,$$

$$H_2^\circ(x) = [1 - 2L'_2(x_2)(x-x_2)](L_2(x))^2 = \frac{1}{4}(4-3x)(x^2+x)^2,$$

$$H_0^{\setminus}(x) = (x-x_0)(L_0(x))^2 = \frac{1}{4}(x+1)(x^2-1)^2,$$

$$H_1^{\setminus}(x) = (x - x_1)(L_1(x))^2 = x(1 - x^2)^2,$$

$$H_2^{\setminus}(x) = (x - x_2)(L_2(x))^2 = \frac{1}{4}(x - 1)(x^2 + x)^2.$$

بنابراین، خواهیم داشت

$$\begin{aligned} f\left(\frac{1}{4}\right) &\approx P\left(\frac{1}{4}\right) = H_0^{\circ}\left(\frac{1}{4}\right) + H_1^{\circ}\left(\frac{1}{4}\right) + 3H_2^{\circ}\left(\frac{1}{4}\right) - 5H_3^{\circ}\left(\frac{1}{4}\right) + H_4^{\setminus}\left(\frac{1}{4}\right) + 7H_5^{\setminus}\left(\frac{1}{4}\right) \\ &= \frac{9}{16} + \frac{9}{16} + 3 \times \frac{45}{128} - 5 \times \frac{27}{128} + \frac{9}{32} + 7 \times \left(-\frac{9}{128}\right) \\ &= \frac{323}{128} = 2.6015625. \end{aligned}$$

با توجه به قضیه ۷.۳، خطای درونیابی برابر است با

$$E(x) = \frac{f^{(6)}(\xi)}{6!} x^2(x-1)^2(x-2)^2, \quad \xi \in (0, 2).$$

اکنون، اگر $M = \max_{x \in (0, 2)} |f^{(6)}(x)|$ ، آنگاه

$$|E\left(\frac{1}{4}\right)| \leq \frac{M}{6!} \left(\frac{1}{4}\right)^2 \left(\frac{1}{4} - 1\right)^2 \left(\frac{1}{4} - 2\right)^2 = 0.0002M.$$

اینک به معرفی چندجمله‌ای درونیاب هرمیت در حالت کلی می‌پردازیم. اعداد حقیقی

$$x_i, f_i^{(k)}, \quad i = 0, 1, \dots, m, \quad k = 0, 1, \dots, n_i - 1,$$

را در نظر بگیرید، به طوری که $x_0 < x_1 < \dots < x_m$. قرار می‌دهیم

$$n = \sum_{i=0}^m n_i - 1.$$

در حالت کلی، چندجمله‌ای درونیاب هرمیت P ، یک چندجمله‌ای از درجه حداکثر n است به طوری که

$$P^{(k)}(x_i) = f_i^{(k)}, \quad i = 0, 1, \dots, m, \quad k = 0, 1, \dots, n_i - 1. \quad (23.3)$$

قضیه ۸.۳ یک چندجمله‌ای یکتا از درجه حداکثر n وجود دارد که در شرایط (۲۳.۳) صدق می‌کند.

برهان: کافی است نشان دهیم که یک و تنها یک چندجمله‌ای از درجه حداکثر n مثل

$$P(x) = c_0 + c_1 x + \dots + c_n x^n,$$

وجود دارد به گونه‌ای که در شرایط (۲۳.۳) صدق می‌کند. شرایط

$$P^{(k)}(x_i) = 0, \quad i = 0, 1, \dots, m, \quad k = 0, 1, \dots, n_i - 1, \quad (24.3)$$

را در نظر بگیرید. روشن است که ماتریس ضرایب دستگاه حاصل از این شرایط و شرایط (۲۳.۳) یکسان است. اگر P در شرایط (۲۴.۳) صدق کند، آنگاه P دارای $n + 1$ ریشه با احتساب تکرار است و در نتیجه، $P \equiv 0$. به عبارت دیگر، این دستگاه جواب یکتای صفر دارد. بنابراین، ماتریس ضرایب دستگاه مربوطه نامنفرد است و این اثبات قضیه را کامل می‌کند. \square

اگرچه با حل دستگاه حاصل از شرایط (۲۳.۳) می‌توان چند جمله‌ای درونیاب را به دست آورد، این روش کارا نیست. از این رو، همانند چند جمله‌ای‌های لاگرانژ و نیوتن روش‌هایی برای محاسبه چند جمله‌ای درونیاب هرمیت ابداع شده است که در ادامه این روش‌ها را معرفی می‌کنیم (مرجع [۱۵] را ببینید).

چند جمله‌ای‌های $L_{ik} \in \Pi_n$ را که به چند جمله‌ای‌های لاگرانژ تعمیم یافته معروف هستند، به شرح زیر تعریف می‌کنیم. قرار می‌دهیم

$$l_{ik}(x) = \frac{(x - x_i)^k}{k!} \prod_{\substack{j=0 \\ j \neq i}}^n \left(\frac{x - x_j}{x_i - x_j} \right)^{n_j}, \quad 0 \leq i \leq m, \quad 0 \leq k \leq n_i.$$

سپس، تعریف می‌کنیم

$$L_{i, n_i - 1}(x) = l_{i, n_i - 1}, \quad i = 0, 1, \dots, m,$$

و به طور بازگشتی برای $k = n_i - 2, n_i - 3, \dots, 0$ قرار می‌دهیم

$$L_{ik} = l_{ik} - \sum_{\nu=k+1}^{n_i-1} l_{ik}^{(\nu)}(x_i) L_{i\nu}(x).$$

با استقرا می‌توان نشان داد که

$$L_{ik}^{(\sigma)}(x_j) = \begin{cases} 1, & i = j, \quad k = \sigma, \\ 0, & \text{otherwise.} \end{cases}$$

اکنون، چند جمله‌ای P را به صورت

$$P(x) = \sum_{i=0}^m \sum_{k=0}^{n_i-1} f_i^{(k)} L_{ik}(x),$$

تعریف می‌کنیم. در این صورت، $P(x)$ چندجمله‌ای درونیاب هرमित تابع f خواهد بود. روند ذکرشده برای محاسبه چندجمله‌ای $P(x)$ ، به الگوریتم هرमित-بیرکف معروف است. در ادامه، روش نیوتن را برای محاسبه چندجمله‌ای درونیاب هرमित تعمیم می‌دهیم. نخست، موضوع را در یک حالت خاص مطالعه می‌کنیم. می‌خواهیم یک چندجمله‌ای از درجه حداکثر ۲ مثل P بیابیم، به طوری که

$$P(x_0) = f(x_0), \quad P'(x_0) = f'(x_0), \quad P''(x_0) = f''(x_0).$$

فرض کنید α_i ، $i = 0, 1, 2$ اعداد متمایزی باشند، به طوری که $\alpha_i \neq x_0$. جدول درونیاب نیوتن برای نقاط $(\alpha_i, f(\alpha_i))$ ، $i = 0, 1, 2$ به صورت زیر است

	$k = 0$	$k = 1$	$k = 2$
α_0	$f[\alpha_0]$	$f[\alpha_0, \alpha_1]$	$f[\alpha_0, \alpha_1, \alpha_2]$
α_1	$f[\alpha_1]$	$f[\alpha_1, \alpha_2]$	
α_2	$f[\alpha_2]$		

بنابراین، چندجمله‌ای درونیاب متناظر با این جدول به صورت زیر است

$$P(x) = f[\alpha_0] + f[\alpha_0, \alpha_1](x - \alpha_0) + f[\alpha_0, \alpha_1, \alpha_2](x - \alpha_0)(x - \alpha_1).$$

در این صورت، با توجه به رابطه (۱۷.۳)، داریم

$$\begin{aligned} \lim_{\alpha_0, \alpha_1 \rightarrow x_0} f[\alpha_0, \alpha_1] &= \lim_{\alpha_0, \alpha_1 \rightarrow x_0} f'(\xi) \quad \xi \in I[\alpha_0, \alpha_1] \\ &= f'(x_0). \end{aligned}$$

به همین ترتیب

$$\lim_{\alpha_1, \alpha_2 \rightarrow x_0} f[\alpha_1, \alpha_2] = f'(x_0).$$

همچنین

$$\begin{aligned} \lim_{\alpha_0, \alpha_1, \alpha_2 \rightarrow x_0} f[\alpha_0, \alpha_1, \alpha_2] &= \lim_{\alpha_0, \alpha_1, \alpha_2 \rightarrow x_0} \frac{1}{2!} f''(\xi) \quad \xi \in I[\alpha_0, \alpha_1, \alpha_2] \\ &= \frac{1}{2!} f''(x_0). \end{aligned}$$

در این صورت، جدول درونیاب نیوتن به صورت زیر خواهد بود.

	$k = 0$	$k = 1$	$k = 2$
x_0	$f[x_0]$	$f'(x_0)$	
x_0	$f[x_0]$	$f'(x_0)$	$\frac{1}{2!}f''(x_0)$
x_0	$f[x_0]$	$f'(x_0)$	

بنابراین، چندجمله‌ای درونیاب مورد نظر به صورت

$$P(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2,$$

به دست می‌آید. شایان ذکر است که تقریب به دست آمده، سه جمله اول بسط تیلر تابع f در نقطه x_0 است.

اکنون، موضوع را در حالت کلی بررسی می‌کنیم. نخست، دنباله $x_0 < x_1 < \dots < x_m$ را به صورت

$$\underbrace{x_0 = x_0 = \dots = x_0}_{\text{تا } n_0} < \underbrace{x_1 = x_1 = \dots = x_1}_{\text{تا } n_1} < \dots < \underbrace{x_m = x_m = \dots = x_m}_{\text{تا } n_m},$$

توسعه می‌دهیم. j امین جمله این دنباله را با t_j نشان می‌دهیم. در این صورت، دنباله

$$t_0 = x_0 \leq t_1 \leq t_2 \leq \dots \leq t_n = x_m,$$

را خواهیم داشت، که در آن $n = \sum_{i=0}^m n_i - 1$. در اینجا، t_j ها را نقاط مجازی می‌گویند. متناظر با t_j ها مجموعه

$$\mathcal{F} = \{(t_j, f(t_j)) : j = 0, 1, \dots, n\},$$

را تعریف می‌کنیم. کافی است چندجمله‌ای درونیاب نیوتن را برای این نقاط بنویسیم (چرا؟). ولی، با توجه به اینکه بعضی از t_j ها برابر هستند، نیاز به تعمیم تفاضلات تقسیم شده برای این نقاط داریم. مجموعه $\tilde{\mathcal{F}}$ را به صورت

$$\tilde{\mathcal{F}} = \{(\alpha_j, f(\alpha_j)) : j = 0, 1, \dots, n\},$$

تعریف می‌کنیم، که در آن

$$\alpha_0 < \alpha_1 < \dots < \alpha_n.$$

نخست، چندجمله‌ای درونیاب نیوتن تابع f را در مجموعه نقاط \tilde{F} می‌نویسیم. این چندجمله‌ای به صورت

$$P(x) = \sum_{j=0}^n f[\alpha_0, \alpha_1, \dots, \alpha_j](x - \alpha_0) \cdots (x - \alpha_{j-1}),$$

خواهد بود. با توجه به جدول تفاضلات تقسیم‌شده، برای محاسبه $f[\alpha_0, \alpha_1, \dots, \alpha_j]$ ، به ازای i و k های مختلف، نیاز به محاسبه

$$f[\alpha_i, \alpha_{i+1}, \dots, \alpha_{i+k}],$$

داریم. با استفاده از رابطه (۱۷.۳)، می‌بینیم که

$$f[\alpha_i, \alpha_{i+1}, \dots, \alpha_{i+k}] = \frac{f^{(k)}(\xi)}{k!}, \quad \xi \in (\alpha_i, \alpha_{i+k}).$$

اکنون آماده‌ایم که چندجمله‌ای درونیاب هرمیت $P(x)$ را برای مجموعه نقاط \mathcal{F} تعریف کنیم. طبیعی است که تعریف کنیم

$$\begin{cases} f[x_i, x_{i+1}, \dots, x_{i+k}] = \lim_{\substack{\alpha_j \rightarrow x_i \\ \forall j}} f[\alpha_i, \alpha_{i+1}, \dots, \alpha_{i+k}] = \frac{f^{(k)}(x_i)}{k!}, & x_i = \cdots = x_{i+k}, \\ f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}, & x_i \neq x_{i+k}, \end{cases} \quad (25.3)$$

و

$$P(x) = \sum_{j=0}^n f[t_0, t_1, \dots, t_j](x - t_0) \cdots (x - t_{j-1}).$$

مثال ۹.۳ فرض کنید $n_0 = 2$ ، $n_1 = 3$ و

$$\begin{aligned} x_0 &= 0, & f_0^{(0)} &= -1, & f_0^{(1)} &= -2, \\ x_1 &= 1, & f_1^{(0)} &= 0, & f_1^{(1)} &= 10, & f_1^{(2)} &= 40. \end{aligned}$$

چندجمله‌ای درونیاب هرمیت متناظر با این نقاط را بنویسید.

حل: نخست با استفاده از رابطه (۲۵.۳) جدول تفاضلات تقسیم‌شده تعمیم‌یافته را می‌نویسیم.

t_i	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$t_0 = 0$	$f[t_0] = -1$				
		$f[t_0, t_1] = -2$			
$t_1 = 0$	$f[t_1] = -1$		$f[t_0, t_1, t_2] = 3$		
		$f[t_1, t_2] = 1$		$f[t_0, \dots, t_3] = 6$	
$t_2 = 1$	$f[t_2] = 0$		$f[t_1, t_2, t_3] = 9$		$f[t_0, \dots, t_4] = 5$
		$f[t_2, t_3] = 10$		$f[t_1, \dots, t_4] = 11$	
$t_3 = 1$	$f[t_3] = 0$		$f[t_2, t_3, t_4] = 20$		
		$f[t_3, t_4] = 10$			
$t_4 = 1$	$f[t_4] = 0$				

برای روشن شدن شیوه محاسبه اعداد ارائه شده در این جدول، سه مقدار از تفاضلات تقسیم شده را محاسبه می‌کنیم. داریم

$$f[t_0, t_1] = \frac{f^{(1)}(t_0)}{1!} = f'(0) = -2,$$

$$f[t_2, t_3, t_4] = \frac{f''(t_2)}{2!} = \frac{f''(1)}{2} = \frac{40}{2} = 20,$$

$$f[t_0, \dots, t_4] = \frac{f[t_1, \dots, t_4] - f[t_0, \dots, t_3]}{t_4 - t_0} = \frac{11 - 6}{1 - 0} = 5.$$

اینک، با استفاده از جدول بالا، چندجمله‌ای درونیاب هرمیت متناظر، به صورت

$$\begin{aligned} P(x) &= f[t_0] + f[t_0, t_1](x - t_0) + f[t_0, t_1, t_2](x - t_0)(x - t_1) + \dots \\ &\quad + f[t_0, t_1, t_2, t_3, t_4](x - t_0)(x - t_1)(x - t_2)(x - t_3) \\ &= -1 - 2(x - 0) + 3(x - 0)(x - 0) + 6(x - 0)(x - 0)(x - 1) \\ &\quad + 5(x - 0)(x - 0)(x - 1)(x - 1) \\ &= -1 - 2x + 3x^2 + 6x^2(x - 1) + 5x^2(x - 1)^2, \end{aligned}$$

خواهد بود.

همانند قضیه ۷.۳، می‌توان دید که خطای چندجمله‌ای درونیاب هرمیت در حالت کلی به صورت

$$f(x) - P(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} w(x), \quad (26.3)$$

است، که در آن $n = \sum_{i=0}^m n_i - 1$ و $w(x) = (x - x_0)^{n_0} (x - x_1)^{n_1} \dots (x - x_m)^{n_m}$ و $n = \sum_{i=0}^m n_i - 1$ با توجه به رابطه (۲۶.۳)، خطای درونیابی در مثال ۹.۳ به صورت زیر است

$$E(x) = \frac{f^{(\delta)}(\xi)}{\delta!} x^\nu (x-1)^\nu, \quad \xi \in (0, 1).$$

به راحتی می توان دید که

$$\max_{x \in [0, 1]} |x^\nu (x-1)^\nu| = \frac{108}{3125}.$$

بنابراین، اگر M یک کران بالا برای تابع $f^{(\delta)}$ در بازه $[0, 1]$ باشد، آنگاه

$$|E(x)| \leq \frac{M}{\delta!} \max_{x \in [0, 1]} |x^\nu (x-1)^\nu| = \frac{M}{\delta!} \times \frac{108}{3125} = 2.88 \times 10^{-4}.$$

۶.۳ درونیابی توسط توابع گویا

فرض کنید $n + 1$ نقطه $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ ، از تابع f داده شده باشد. می خواهیم یک عبارت گویا به شکل

$$\Phi^{\mu, \nu}(x) = \frac{P^{\mu, \nu}(x)}{Q^{\mu, \nu}(x)} = \frac{a_0 + a_1 x + \dots + a_\mu x^\mu}{b_0 + b_1 x + \dots + b_\nu x^\nu},$$

بیابیم، به طوری که

$$\Phi^{\mu, \nu}(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

برای تعیین $\Phi^{\mu, \nu}$ باید $\mu + \nu + 2$ ضریب $a_0, a_1, \dots, a_\mu, b_0, b_1, \dots, b_\nu$ مشخص شوند. روشن است که یکی از ضرایب را می توان برابر با یک فرض کرد و بنابراین، $\mu + \nu + 1$ ضریب نامعلوم داریم. در نتیجه، $n = \mu + \nu$. اکنون، مسئله

$$\Phi^{\mu, \nu}(x_i) = f(x_i), \quad i = 0, 1, \dots, \mu + \nu, \quad (A^{\mu, \nu} \text{ مسئله})$$

را در نظر بگیرید. به روشنی، ضرایب در $\Phi^{\mu, \nu}$ ، جواب معادله همگن

$$P^{\mu, \nu}(x_i) - f(x_i)Q^{\mu, \nu}(x_i) = 0, \quad i = 0, 1, \dots, \mu + \nu,$$

یا به عبارت دیگر، مسئله

$$a_0 + a_1 x_i + \dots + a_\mu x_i^\mu - f(x_i)(b_0 + b_1 x_i + \dots + b_\nu x_i^\nu) = 0, \quad i = 0, 1, \dots, \mu + \nu,$$

است. این مسئله را مسئله $S^{\mu, \nu}$ می‌نامیم. این دستگاه را می‌توان به شکل ماتریسی به صورت

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^m & -f(x_0) & -f(x_0)x_0 & \cdots & -f(x_0)x_0^n \\ 1 & x_1 & \cdots & x_1^m & -f(x_1) & -f(x_1)x_1 & \cdots & -f(x_1)x_1^n \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^m & -f(x_n) & -f(x_n)x_n & \cdots & -f(x_n)x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_\mu \\ b_0 \\ \vdots \\ b_\mu \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (27.3)$$

نمایش داد. همان‌گونه که اشاره شد، هر جواب $A^{\mu, \nu}$ یک جواب $S^{\mu, \nu}$ نیز هست، ولی عکس این مطلب در حالت کلی درست نیست. به مثال زیر توجه کنید.

مثال ۱۰.۳ نقاط داده‌شده در جدول زیر را در نظر بگیرید.

i	۰	۱	۲
x_i	۰	۱	۲
$f(x_i)$	۱	۲	۲

فرض کنید $\mu = \nu = 1$. در این صورت، مسئله $S^{\mu, \nu}$ به صورت

$$\begin{cases} a_0 - b_0 = 0, \\ a_0 + a_1 - 2(b_0 + b_1) = 0, \\ a_0 + 2a_1 - 2(b_0 + 2b_1) = 0, \end{cases}$$

است. می‌توان این دستگاه را به شکل ماتریسی به صورت

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & -2 & -2 \\ 1 & 2 & -2 & -4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

نمایش داد. با استفاده از عملیات سطری مقدماتی می‌بینیم که این دستگاه معادل با دستگاه

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

است. بنابراین، با فرض $b_1 = \alpha \in \mathbb{R}$ ، جواب عمومی دستگاه به صورت

$$(a_0, a_1, b_0, b_1) = (0, 2\alpha, 0, \alpha),$$

خواهد بود. به ازای $\alpha = 1$ ، جواب $a_0 = 0$ ، $b_0 = 0$ ، $a_1 = 2$ و $b_1 = 1$ به دست می آید و در این صورت، تابع گویای $\Phi^{1,1}$ به صورت

$$\Phi^{1,1}(x) = \frac{2x}{x},$$

خواهد بود. می بینیم که $\Phi^{1,1}(0) \neq 1$.

با همان داده های جدول، فرض کنید $\mu = 0$ و $\nu = 2$. در این صورت

$$\Phi^{0,2}(x) = \frac{a_0}{b_0 + b_1 x + b_2 x^2},$$

و مسئله $S^{\mu,\nu}$ مربوط به صورت

$$\begin{cases} a_0 - 1b_0 & = 0, \\ a_0 - 2(b_0 + b_1 + b_2) & = 0, \\ a_0 - 2(b_0 + 2b_1 + 4b_2) & = 0, \end{cases}$$

است. این دستگاه به شکل ماتریسی به صورت

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -2 & -2 & -2 \\ 1 & -2 & -4 & -8 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

نمایش داده می شود. با استفاده از عملیات سطری مقدماتی، این دستگاه را می توان به صورت

$$\begin{pmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

نوشت. با فرض $b_2 = \alpha \in \mathbb{R}$ ، جواب عمومی این دستگاه به صورت

$$(a_0, b_0, b_1, b_2) = (4\alpha, 4\alpha, -3\alpha, \alpha),$$

خواهد بود. همچنین، با فرض $\alpha = 1$ ، یک جواب نابدیهی برای این دستگاه به صورت

$$a_0 = 4, b_0 = 4, b_1 = -3, b_2 = 1,$$

است. به ازای این مقادیر، داریم

$$\Phi^{0,2}(x) = \frac{4}{4 - 3x + x^2}.$$

اکنون می خواهیم ببینیم که چه زمانی یک جواب از مسئله $S^{\mu,\nu}$ ، جوابی برای $A^{\mu,\nu}$ است.

تعریف ۲.۳ دو عبارت گویای

$$\Phi_1(x) \equiv \frac{P_1(x)}{Q_1(x)}, \quad \Phi_2(x) \equiv \frac{P_2(x)}{Q_2(x)}, \quad Q_1(x) \not\equiv 0, \quad Q_2(x) \not\equiv 0,$$

معادل گفته می‌شود و با $\Phi_1 \sim \Phi_2$ نمایش می‌دهیم، هرگاه

$$P_1(x)Q_2(x) \equiv P_2(x)Q_1(x).$$

تعریف ۳.۳ یک عبارت گویا به طور نسبی اول گفته می‌شود، هرگاه صورت و مخرج این عبارت، نسبت به هم اول باشند.

مثال ۱۱.۳ مثال ۱۰.۳ را در نظر بگیرید. دیدیم که

$$\Phi_1(x) = \frac{2x}{x},$$

یک جواب برای دستگاه $S^{\mu,\nu}$ است. قرار می‌دهیم

$$\Phi_2(x) = 2.$$

داریم $\Phi_1 \sim \Phi_2$. به علاوه، Φ_2 به طور نسبی اول است، ولی Φ_1 این گونه نیست.

قضیه ۹.۳ دستگاه معادلات خطی همگن $S^{\mu,\nu}$ همیشه جواب‌های نابدیهی دارد و برای هر جواب نابدیهی

$$\Phi^{\mu,\nu}(x) = \frac{P^{\mu,\nu}(x)}{Q^{\mu,\nu}(x)},$$

داریم $Q^{\mu,\nu}(x) \not\equiv 0$. به عبارت دیگر، هر جواب نابدیهی یک عبارت گویا تعریف می‌کند.

برهان: دستگاه معادلات خطی همگن $S^{\mu,\nu}$ دارای $\mu + \nu + 1$ معادله و $\mu + \nu + 2$ مجهول است و بنابراین، بی‌شمار جواب دارد. فرض کنید یک جواب نابدیهی برای $S^{\mu,\nu}$ ، به صورت

$$(a_0, a_1, \dots, a_\mu, b_0, b_1, \dots, b_\nu) \neq (0, 0, \dots, 0),$$

باشد. به برهان خلف، فرض کنید $Q^{\mu,\nu} \equiv 0$. در این صورت، از

$$P^{\mu,\nu}(x_i) - f_i Q^{\mu,\nu}(x_i) = 0, \quad i = 0, 1, \dots, \mu + \nu,$$

خواهیم داشت

$$P^{\mu,\nu}(x_i) = 0, \quad i = 0, 1, \dots, \mu + \nu,$$

و این یعنی، $P^{\mu,\nu}$ دارای $\mu + \nu + 1$ ریشه است که یک تناقض است؛ زیرا $P^{\mu,\nu}$ از درجه حداکثر μ است. بنابراین، $Q^{\mu,\nu} \not\equiv 0$. \square

قضیه ۱۰.۳ اگر Φ_1 و Φ_2 دو جواب نابدیهی برای دستگاه $S^{\mu,\nu}$ باشند، آنگاه $\Phi_1 \sim \Phi_2$ ، یعنی، Φ_1 و Φ_2 معادل هم هستند.

برهان: فرض کنید

$$\Phi_1(x) = \frac{P_1(x)}{Q_1(x)}, \quad \Phi_2(x) = \frac{P_2(x)}{Q_2(x)},$$

دو جواب برای دستگاه $S^{\mu,\nu}$ باشند. در این صورت، درجه چندجمله‌ای

$$P(x) \equiv P_1(x)Q_2(x) - P_2(x)Q_1(x),$$

حداکثر $\mu + \nu$ است و داریم

$$\begin{aligned} P(x_i) &= P_1(x_i)Q_2(x_i) - P_2(x_i)Q_1(x_i) \\ &= f(x_i)Q_1(x_i)Q_2(x_i) - f(x_i)Q_2(x_i)Q_1(x_i) \\ &= 0, \quad i = 0, 1, \dots, \mu + \nu. \end{aligned}$$

بنابراین، P دارای $\mu + \nu + 1$ ریشه است که نشان می‌دهد $P(x) \equiv 0$ و این یعنی، $\Phi_1 \sim \Phi_2$.
شایان ذکر است که عکس این قضیه درست نیست، به این معنی که اگر Φ_1 جوابی برای $S^{\mu,\nu}$ و Φ_2 عبارت گویای دیگری باشد که $\Phi_1 \sim \Phi_2$ ، آنگاه دلیلی ندارد که Φ_2 نیز جوابی برای $S^{\mu,\nu}$ باشد. در مثال ۱۱.۳ دو تابع Φ_1 و Φ_2 چنین هستند.

با استفاده از دو قضیه پیشین، برای هر مسئله درونیابی $A^{\mu,\nu}$ یک تابع گویای یکتای $\Phi^{\mu,\nu}$ متناظر با دستگاه $S^{\mu,\nu}$ موجود است. این تابع گویا به ازای هر i در شرط $\Phi^{\mu,\nu}(x_i) = f(x_i)$ صدق می‌کند یا نقاطی مثل $(x_i, f(x_i))$ موجود است که $\Phi^{\mu,\nu}(x_i) \neq f(x_i)$. در صورتی که $\Phi^{\mu,\nu}(x_i) \neq f(x_i)$ ، آنگاه نقطه $(x_i, f(x_i))$ را دست‌نیافتنی می‌گوییم. اکنون، فرض کنید

$$\Phi^{\mu,\nu}(x) = \frac{P^{\mu,\nu}(x)}{Q^{\mu,\nu}(x)},$$

یک جواب برای $S^{\mu,\nu}$ باشد. به ازای هر $i \in \{0, 1, \dots, \mu + \nu\}$ دو حالت

(الف) $Q^{\mu,\nu}(x_i) \neq 0$ ،

(ب) $Q^{\mu,\nu}(x_i) = 0$ ،

را در نظر می‌گیریم. در حالت نخست، به وضوح داریم $\Phi^{\mu,\nu}(x_i) = f(x_i)$. در حالت دوم ممکن است $(x_i, f(x_i))$ دست‌نیافتنی باشد. از اینکه $Q^{\mu,\nu}(x_i) = 0$ نتیجه می‌شود $P^{\mu,\nu}(x_i) = 0$. یعنی، $P^{\mu,\nu}$ و $Q^{\mu,\nu}$ دارای عامل مشترک $x - x_i$ هستند و نسبت به هم اول نیستند. با جمع‌بندی این مطالب، نتیجه بعد را بیان می‌کنیم.

نتیجه ۱.۳ اگر $S^{\mu,\nu}$ جوابی مثل $\Phi^{\mu,\nu}$ داشته باشد که به طور نسبی اول باشد، آنگاه هیچ نقطه دست نیافتنی وجود ندارد و مسئله $A^{\mu,\nu}$ حل پذیر است.

همچنین، می توانیم قضیه زیر را بیان کنیم.

قضیه ۱۱.۳ فرض کنید $\Phi^{\mu,\nu}$ یک عبارت گویا و $\tilde{\Phi}^{\mu,\nu}$ یک عبارت گویای به طور نسبی اول و معادل با آن باشد به طوری که $\Phi^{\mu,\nu}$ جوابی برای $S^{\mu,\nu}$ است. در این صورت، $A^{\mu,\nu}$ حل پذیر است و $\Phi^{\mu,\nu}$ یک جواب برای آن است اگر و تنها اگر $\tilde{\Phi}^{\mu,\nu}$ جوابی برای $S^{\mu,\nu}$ باشد.

مثال ۱۲.۳ برای نقاط داده شده

i	۰	۱	۲	۳	۴
x_i	-۱	۰	۱	۲	۳
$f(x_i)$	۲	۱	۲	۲	-۱

فرض کنید $\mu = \nu = ۲$. در این صورت، تابع $\Phi^{۲,۲}$ به صورت

$$\Phi^{۲,۲}(x) = \frac{a_0 + a_1x + a_2x^2}{b_0 + b_1x + b_2x^2},$$

دستگاه (۲۷.۳) برای محاسبه a_i ها و b_i ها به صورت

$$\begin{pmatrix} ۱ & -۱ & ۱ & -۲ & ۲ & -۲ \\ ۱ & ۰ & ۰ & -۱ & ۰ & ۰ \\ ۱ & ۱ & ۱ & -۲ & -۲ & -۲ \\ ۱ & ۲ & ۴ & -۲ & -۴ & -۸ \\ ۱ & ۳ & ۹ & ۱ & ۳ & ۹ \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} ۰ \\ ۰ \\ ۰ \\ ۰ \\ ۰ \\ ۰ \end{pmatrix},$$

خواهد بود. با استفاده از عملیات سطری مقدماتی، این دستگاه معادل با دستگاه

$$\begin{pmatrix} ۱ & ۰ & ۰ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ & ۰ & ۶ \\ ۰ & ۰ & ۱ & ۰ & ۰ & -۲ \\ ۰ & ۰ & ۰ & ۱ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۰ & ۱ & ۳ \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} ۰ \\ ۰ \\ ۰ \\ ۰ \\ ۰ \\ ۰ \end{pmatrix},$$

است. با فرض $b_2 = \alpha \in \mathbb{R}$ ، جواب عمومی دستگاه به صورت

$$(a_0, a_1, a_2, b_0, b_1, b_2) = (0, -6\alpha, 2\alpha, 0, -3\alpha, \alpha),$$

خواهد بود و به ازای $\alpha = 1$ ، داریم

$$a_0 = 0, a_1 = -6, a_2 = 2, b_0 = 0, b_1 = -3, b_2 = 1.$$

در نتیجه

$$\Phi^{2,2}(x) = \frac{-6x + 2x^2}{-3x + x^2} = \frac{2x(x-3)}{x(x-3)}.$$

با توجه به اینکه صورت و مخرج این کسر، هر دو عامل $x - x_0$ و $x - x_1$ را دارند، نتیجه می‌گیریم که مسئله $A^{\mu,\nu}$ حل پذیر نیست.

اینک، فرض می‌کنیم $\mu = 3$ و $\nu = 1$. دستگاه (۲۷.۳) به صورت

$$\begin{pmatrix} 1 & -1 & 1 & -1 & -2 & 2 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 & -2 & -2 \\ 1 & 2 & 4 & 8 & -2 & -4 \\ 1 & 3 & 9 & 27 & 1 & 3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

خواهد بود. با استفاده از عملیات سطری مقدماتی، این دستگاه معادل با دستگاه

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & 0 & -\frac{13}{3} \\ 0 & 0 & 1 & 0 & 0 & -9 \\ 0 & 0 & 0 & 1 & 0 & \frac{9}{3} \\ 0 & 0 & 0 & 0 & 1 & -9 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

است. با فرض $b_1 = \alpha \in \mathbb{R}$ ، جواب عمومی دستگاه به صورت

$$(a_0, a_1, a_2, b_0, b_1, b_2) = (9\alpha, \frac{13}{3}\alpha, 9\alpha, -\frac{9}{3}\alpha, 9\alpha, \alpha),$$

خواهد بود و به ازای $\alpha = 2$ ، خواهیم داشت

$$a_0 = 18, a_1 = 13, a_2 = 18, a_3 = -9, b_0 = 18, b_2 = 2.$$

در نتیجه، داریم

$$\Phi^{3,1}(x) = \frac{18 + 13x + 18x^2 - 9x^3}{18 + 2x}.$$

با توجه به اینکه صورت و مخرج کسر این نسبت به هم اول هستند، نتیجه می‌گیریم که $\Phi^{3,1}$ جواب مسئله $A^{\mu,\nu}$ است.

۷.۳ تفاضلات معکوس و کسرهای مسلسل تیل

برای مجموعه نقاط $(x_i, f(x_i))$ ، $i = 0, 1, \dots$ ، جدول تفاضلات معکوس به صورت زیر تعریف می‌شود.

i	x_i	$\varphi(x_i)$	$\varphi(x_0, x_i)$	$\varphi(x_0, x_1, x_i)$	$\varphi(x_0, x_1, x_2, x_i)$
۰	x_0	$f(x_0)$			
۱	x_1	$f(x_1)$	$\varphi(x_0, x_1)$		
۲	x_2	$f(x_2)$	$\varphi(x_0, x_2)$	$\varphi(x_0, x_1, x_2)$	
۳	x_3	$f(x_3)$	$\varphi(x_0, x_3)$	$\varphi(x_0, x_1, x_3)$	$\varphi(x_0, x_1, x_2, x_3)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

در این جدول، داریم

$$\begin{aligned} \varphi(x_i) &= f(x_i), \\ \varphi(x_i, x_j) &= \frac{x_i - x_j}{f(x_i) - f(x_j)}, \\ \varphi(x_i, x_j, x_k) &= \frac{x_j - x_k}{\varphi(x_i, x_j) - \varphi(x_i, x_k)}, \\ \varphi(x_i, \dots, x_\ell, x_m, x_n) &= \frac{x_m - x_n}{\varphi(x_i, \dots, x_\ell, x_m) - \varphi(x_i, \dots, x_\ell, x_n)}. \end{aligned}$$

در اینجا توجه کنید که مخرج کسرهای بالا ممکن است صفر شود که در این صورت، مقدار کسر را برابر با ∞ می‌گیریم. همچنین، توجه کنید که تفاضلات معکوس نسبت به آرگومان‌های خود متقارن نیستند. اکنون، فرض کنید $P^\mu \in \Pi_\mu$ و $Q^\nu \in \Pi_\nu$. می‌خواهیم یک عبارت گویا مثل

$$\Phi^{n,n}(x) = \frac{P^n(x)}{Q^n(x)},$$

بیابیم، به طوری که

$$\Phi^{n,n}(x_i) = f(x_i), \quad i = 0, 1, \dots, 2n.$$

برای محاسبه $\Phi^{n,n}$ به شرح زیر عمل می‌کنیم. داریم

$$\frac{P^n(x)}{Q^n(x)} = f_0 + \frac{P^n(x)}{Q^n(x)} - f_0 = f_0 + \frac{P^n(x) - f_0 Q^n(x)}{Q^n(x)}.$$

ولی، $P^n(x_0) - f_0 Q^n(x_0) = 0$ ، و بنابراین، یک چندجمله‌ای $P^{n-1} \in \Pi_{n-1}$ وجود دارد به طوری که

$$P^n(x) - f_0 Q^n(x) = (x - x_0) P^{n-1}(x).$$

از این رو

$$\frac{P^n(x)}{Q^n(x)} = f(x_0) + (x - x_0) \frac{P^{n-1}(x)}{Q^n(x)} = f_0 + \frac{x - x_0}{\frac{Q^n(x)}{P^{n-1}(x)}}. \quad (28.3)$$

ولی، داریم

$$f(x_i) = \frac{P^n(x_i)}{Q^n(x_i)} = f(x_0) + \frac{x_i - x_0}{\frac{Q^n(x_i)}{P^{n-1}(x_i)}}, \quad i = 1, 2, \dots, 2n,$$

و در نتیجه

$$\frac{Q^n(x_i)}{P^{n-1}(x_i)} = \frac{x_i - x_0}{f(x_i) - f(x_0)} = \varphi(x_0, x_i), \quad i = 1, 2, \dots, 2n.$$

همچنین، داریم

$$\begin{aligned} \frac{Q^n(x)}{P^{n-1}(x)} &= \varphi(x_0, x_1) + \frac{Q^n(x)}{P^{n-1}(x)} - \varphi(x_0, x_1) \\ &= \varphi(x_0, x_1) + \frac{Q^n(x) - \varphi(x_0, x_1)P^{n-1}(x)}{P^{n-1}(x)}. \end{aligned}$$

از طرفی

$$Q^n(x_1) - \varphi(x_0, x_1)P^{n-1}(x_1) = 0.$$

در نتیجه، به ازای یک $Q^{n-1} \in \Pi_{n-1}$ ، داریم

$$\frac{Q^n(x)}{P^{n-1}(x)} = \varphi(x_0, x_1) + (x - x_1) \frac{Q^{n-1}(x)}{P^{n-1}(x)}. \quad (29.3)$$

ولی، داریم

$$\varphi(x_0, x_i) = \frac{Q^n(x_i)}{P^{n-1}(x_i)} = \varphi(x_0, x_1) + (x_i - x_1) \frac{Q^{n-1}(x_i)}{P^{n-1}(x_i)}, \quad i = 2, 3, \dots, 2n,$$

و بنابراین

$$\frac{P^{n-1}(x_i)}{Q^{n-1}(x_i)} = \frac{x_i - x_1}{\varphi(x_0, x_i) - \varphi(x_0, x_1)} = \varphi(x_0, x_1, x_i), \quad i = 2, 3, \dots, 2n.$$

به این ترتیب، با استفاده از رابطه‌های (۲۸.۳) و (۲۹.۳)، داریم

$$\Phi^{n,n}(x) = \frac{P^n(x)}{Q^n(x)} = f_0 + \frac{x - x_0}{\varphi(x_0, x_1) + \frac{x - x_1}{\frac{P^{n-1}(x)}{Q^{n-1}(x)}}}.$$

با ادامه روند بالا، خواهیم داشت

$$\Phi^{n,n}(x) = \frac{P^n(x)}{Q^n(x)} = f_0 + \frac{x - x_0}{\varphi(x_0, x_1) + \frac{x - x_1}{\varphi(x_0, x_1, x_2) + \dots + \frac{x - x_{2n-1}}{\varphi(x_0, \dots, x_{2n})}}$$

$\Phi^{n,n}$ را با استفاده از کسرهای مسلسل به صورت زیر نمایش می دهند

$$\Phi^{n,n}(x) = f_0 + \frac{x - x_0}{\varphi(x_0, x_1)} + \frac{x - x_1}{\varphi(x_0, x_1, x_2)} + \frac{x - x_2}{\varphi(x_0, x_1, x_2, x_3)} + \dots + \frac{x - x_{2n-1}}{\varphi(x_0, \dots, x_{2n})}$$

مثال ۱۳.۳ با استفاده از جدول تفاضلات معکوس، تابع f را در نقاط جدول

i	۰	۱	۲	۳
x_i	۰	۱	۲	۳
$f(x_i)$	۰	-۱	$-\frac{2}{3}$	۹

درنیابی کنید.

حل: برای داده‌های بالا، جدول تفاضلات معکوس به صورت زیر است.

i	x_i	$f(x_i)$	$\varphi(x_0, x_i)$	$\varphi(x_0, x_1, x_i)$	$\varphi(x_0, x_1, x_2, x_i)$
۰	۰	۰			
۱	۱	-۱	-۱		
۲	۲	$-\frac{2}{3}$	-۳	$-\frac{1}{3}$	
۳	۳	۹	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$

بعضی از مقادیر این جدول به صورت زیر محاسبه می شوند

$$\varphi(x_0, x_1, x_2) = \frac{x_1 - x_2}{\varphi(x_0, x_1) - \varphi(x_0, x_2)} = \frac{1 - 2}{-1 + 3} = -\frac{1}{2}$$

$$\varphi(x_0, x_1, x_3) = \frac{x_1 - x_3}{\varphi(x_0, x_1) - \varphi(x_0, x_3)} = \frac{1 - 3}{-1 - \frac{1}{3}} = \frac{3}{2}$$

در این صورت، داریم

$$\Phi^{2,1}(x) = 0 + \frac{x - 0}{-1} + \frac{x - 1}{-1/2} + \frac{x - 2}{1/2} = \frac{4x^2 - 9x}{-2x + 7}$$

اگر قرار دهیم

$$\begin{aligned}\Phi^{0,0}(x) &= f_0, \\ \Phi^{1,0}(x) &= f_0 + \frac{x - x_0}{x_1 - x_0} \varphi(x_0, x_1), \\ \Phi^{2,0}(x) &= f_0 + \frac{x - x_0}{x_1 - x_0} \varphi(x_0, x_1) + \frac{x - x_1}{x_2 - x_1} \varphi(x_1, x_2), \\ &\vdots\end{aligned}$$

آنگاه ملاحظه می‌شود که کسرهای جزئی بالا را می‌توان به صورت $\Phi^{\mu,\mu}$ یا $\Phi^{\mu+1,\mu}$ نشان داد، که در آن $\mu = 0, 1, \dots, n-1$. توجه کنید که اگر تعداد نقاط فرد باشد، آنگاه درجه صورت و مخرج عبارت گویا برابر خواهد بود، یعنی به صورت $\Phi^{\mu,\mu}$ و اگر تعداد نقاط زوج باشد، آنگاه درجه صورت یک واحد از درجه مخرج بیشتر خواهد بود، یعنی به صورت $\Phi^{\mu+1,\mu}$.

تفاضلات معکوس نسبت به آرگومان‌های خود متقارن نیستند، ولی در ادامه تفاضلات متقابل را معرفی می‌کنیم که نسبت به آرگومان‌های خود متقارن است. یادآوری می‌کنیم که تابع f را متقارن می‌گویند، هرگاه هر مقدار آن به ازای هر جایگشت از آرگومان‌های آن بدون تغییر باقی بماند [۱۱]. تفاضلات متقابل را به صورت $\rho(x_i, x_{i+1}, \dots, x_{i+k})$ نشان می‌دهند و رابطه بازگشتی برای محاسبه آن به صورت زیر تعریف می‌شود

$$\begin{aligned}\rho(x_i) &= f(x_i), \\ \rho(x_i, x_k) &= \frac{x_i - x_k}{f(x_i) - f(x_k)}, \\ &\vdots \\ \rho(x_i, x_{i+1}, \dots, x_{i+k}) &= \frac{x_i - x_{i+k}}{\rho(x_i, \dots, x_{i+k-1}) - \rho(x_{i+1}, \dots, x_{i+k})} \\ &\quad + \rho(x_{i+1}, \dots, x_{i+k-1}).\end{aligned}\tag{30.3}$$

قضیه ۱۲.۳ اگر فرض کنیم $\rho(x_0, x_{-1}) = 0$ ، آنگاه برای $p = 1, 2, \dots$ داریم

$$\varphi(x_0, x_1, \dots, x_p) = \rho(x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-2}).$$

برهان: اثبات به استقرا روی p انجام می‌شود. اگر $p = 1$ ، آنگاه

$$\varphi(x_0, x_1) = \rho(x_0, x_1) - 0 = \rho(x_0, x_1),$$

که همیشه برقرار است. فرض کنید حکم برای p برقرار باشد. طبق تعریف تابع φ ، داریم

$$\varphi(x_0, x_1, \dots, x_{p+1}) = \frac{x_p - x_{p+1}}{\varphi(x_0, \dots, x_p) - \varphi(x_0, \dots, x_{p-1}, x_{p+1})}.\tag{31.3}$$

اکنون، می‌توان مخرج این کسر را به صورت زیر نوشت

$$\begin{aligned} \varphi(x_0, \dots, x_p) - \varphi(x_0, \dots, x_{p-1}, x_{p+1}) &= (\rho(x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-2})) \\ &\quad - (\rho(x_0, \dots, x_{p-1}, x_{p+1}) - \rho(x_0, \dots, x_{p-2})) \\ &= \rho(x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-1}, x_{p+1}). \end{aligned}$$

شایان ذکر است که آخرین تساوی رابطه بالا با استفاده از فرض استقرا نوشته شده است. اینک، با جای‌گذاری رابطه اخیر در رابطه (۳۱.۳)، خواهیم داشت

$$\varphi(x_0, x_1, \dots, x_{p+1}) = \frac{x_p - x_{p+1}}{\rho(x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-1}, x_{p+1})}.$$

ولی با توجه به این رابطه و رابطه (۳۰.۳)، داریم

$$\begin{aligned} \rho(x_{p+1}, x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-1}) &= \frac{x_{p+1} - x_p}{\rho(x_{p+1}, x_0, \dots, x_{p-1}) - \rho(x_0, \dots, x_p)} \\ &= \frac{x_p - x_{p+1}}{\rho(x_0, \dots, x_p) - \rho(x_0, \dots, x_{p-1}, x_{p+1})} \\ &= \varphi(x_0, x_1, \dots, x_{p+1}). \end{aligned}$$

اکنون، از اینکه $\rho(\cdot)$ متقارن است، داریم

$$\varphi(x_0, x_1, \dots, x_{p+1}) = \rho(x_0, \dots, x_p, x_{p+1}) - \rho(x_0, \dots, x_{p-1}),$$

□ که اثبات قضیه را کامل می‌کند.

جدول تفاضلات متقابل را می‌توان به صورت زیر نمایش داد.

x_i	$f(x_i)$	$\rho(\cdot, \cdot)$	$\rho(\cdot, \cdot, \cdot)$	$\rho(\cdot, \cdot, \cdot, \cdot)$
x_0	$f(x_0)$			
		$\rho(x_0, x_1)$		
x_1	$f(x_1)$		$\rho(x_0, x_1, x_2)$	
		$\rho(x_1, x_2)$		$\rho(x_0, x_1, x_2, x_3)$
x_2	$f(x_2)$		$\rho(x_1, x_2, x_3)$	
		$\rho(x_2, x_3)$		
x_3	$f(x_3)$			
\vdots	\vdots	\vdots	\vdots	\vdots

با استفاده از مقادیر این جدول، داریم

$$\Phi^{n,n}(x) = f_0 + \frac{x - x_0}{\rho(x_0, x_1)} + \frac{x - x_1}{\rho(x_0, x_1, x_2) - \rho(x_0)} + \dots + \frac{x - x_{n-1}}{\rho(x_0, \dots, x_n) - \rho(x_0, \dots, x_{n-2})}.$$

این کسر را کسر مسلسل تیل می‌گویند.

مثال ۱۴.۳ با استفاده از جدول تفاضلات متقابل تابع f را در نقاط جدول

i	۰	۱	۲	۳
x_i	۰	۱	۲	۳
$f(x_i)$	۰	-۱	$-\frac{2}{3}$	۹

درونیابی کنید.

حل: برای داده‌های بالا، جدول تفاضلات متقابل به صورت زیر است.

i	x_i	$f(x_i)$	$\rho(.,.)$	$\rho(.,.,.)$	$\rho(.,.,.,.)$
۰	۰	۰			
۱	۱	-۱	-۱		
۲	۲	$-\frac{2}{3}$	۳	$-\frac{1}{3}$	
۳	۳	۹	$\frac{2}{29}$	$-\frac{19}{14}$	$-\frac{1}{3}$

برخی از مقادیر این جدول به صورت زیر محاسبه شده‌اند

$$\begin{aligned} \rho(x_2) &= -\frac{2}{3}, & \rho(x_2, x_3) &= \frac{x_2 - x_3}{f(x_2) - f(x_3)} = \frac{2 - 3}{-\frac{2}{3} - 9} = \frac{3}{29}, \\ \rho(x_1, x_2, x_3) &= \frac{x_1 - x_3}{\rho(x_1, x_2) - \rho(x_2, x_3)} + \rho(x_2) = \frac{1 - 3}{3 - \frac{2}{29}} + \left(-\frac{2}{3}\right) = -\frac{19}{14}, \\ \rho(x_0, x_1, x_2, x_3) &= \frac{x_0 - x_3}{\rho(x_0, x_1, x_2) - \rho(x_1, x_2, x_3)} + \rho(x_1, x_2, x_3) = \frac{0 - 3}{-\frac{1}{3} + \frac{19}{14}} + 3 = -\frac{1}{3}. \end{aligned}$$

در این صورت، خواهیم داشت

$$\Phi^{3,1}(x) = 0 + \frac{x - 0}{-1} + \frac{x - 1}{(-1/2) - (-1)} + \frac{x - 2}{(-1/2) - (-1)} = \frac{4x^2 - 9x}{-2x + 7}.$$

انتظار می‌رود که خواننده بتواند مزیت مقارن بودن $\rho(\cdot)$ نسبت به $\varphi(\cdot)$ را تشخیص دهد.

۸.۳ درونیابی مثلثاتی

درونیابی توابع متناوب با استفاده از چندجمله‌ای‌ها مناسب نیست. برای این توابع از درونیابی مثلثاتی استفاده می‌شود. فرض کنید تابع f متناوب و با دوره تناوب $T > 0$ باشد، یعنی

$$f(t+T) = f(t), \quad t \in \mathbb{R}.$$

از اینجا به بعد فرض می‌کنیم $T = 2\pi$. این فرض از کلیت مسئله کم نمی‌کند.

تعریف ۴.۳ برای هر $n \in \mathbb{N}$ فضای برداری T_n مجموعه همه چندجمله‌ای‌های مثلثاتی به شکل

$$q(t) = \sum_{k=0}^n a_k \cos kt + \sum_{k=1}^n b_k \sin kt,$$

است، که در آن a_0, \dots, a_n و b_1, \dots, b_n اعداد حقیقی یا مختلط هستند. چندجمله‌ای $q \in T_n$ را از درجه n می‌گویند، هرگاه $|a_n| + |b_n| > 0$.

به سادگی می‌توان دید که اگر $q_1 \in T_{n_1}$ و $q_2 \in T_{n_2}$ ، آنگاه $q_1 q_2 \in T_{n_1+n_2}$.

قضیه ۱۳.۳ اگر یک چندجمله‌ای در T_n بیش از $2n$ ریشه متمایز در فاصله تناوبی $[0, 2\pi]$ داشته باشد، آنگاه معادل با صفر خواهد بود.

برهان: فرض کنید $q \in T_n$ و

$$q(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt). \quad (32.2)$$

با جای گذاری رابطه‌های مثلثاتی

$$\cos kt = \frac{e^{ikt} + e^{-ikt}}{2}, \quad \sin kt = \frac{e^{ikt} - e^{-ikt}}{2i},$$

در رابطه (۳۲.۳)، خواهیم داشت

$$q(t) = \frac{a_0}{2} + \sum_{k=1}^n \left(\frac{1}{2} (a_k - ib_k) e^{ikt} + \frac{1}{2} (a_k + ib_k) e^{-ikt} \right).$$

اکنون، با فرض $b_0 = 0$ و

$$\gamma_k = \frac{1}{2} (a_k - ib_k), \quad \gamma_{-k} = \frac{1}{2} (a_k + ib_k), \quad k = 0, 1, \dots, n,$$

می‌توان $q(t)$ را به صورت

$$q(t) = \sum_{k=-n}^n \gamma_k e^{ikt},$$

نوشت. اگر قرار دهیم $z = e^{it}$ و

$$p(z) = \sum_{k=-n}^n \gamma_k z^{n+k},$$

آنگاه

$$q(t) = z^{-n} p(z).$$

اینک، فرض کنید که چندجمله‌ای مثلثاتی $q \in T_n$ بیش از $2n$ ریشه متمایز در فاصله $[0, 2\pi]$ داشته باشد. در این صورت، p یک چندجمله‌ای از درجه $2n$ است و بیش از $2n$ ریشه متمایز روی دایره واحد دارد. در نتیجه، داریم

$$p(z) = 0,$$

و این یعنی به ازای هر k ، $\gamma_k = 0$ و از اینجا نتیجه می‌شود که به ازای هر k ، $a_k = b_k = 0$. □

قضیه ۱۴.۳ توابع $c_k(t) = \cos kt$ ، $k = 0, 1, \dots, n$ و $s_k(x) = \sin kt$ ، $k = 1, 2, \dots, n$ در $C[0, 2\pi]$ مستقل خطی هستند.

برهان: فرض کنید

$$\sum_{k=0}^n a_k \cos kt + \sum_{k=1}^n b_k \sin kt = 0, \quad \forall t \in [0, 2\pi].$$

این رابطه نشان می‌دهد که یک چندجمله‌ای مثلثاتی با ضرایب $a_0, \dots, a_n, b_1, \dots, b_n$ بیش از $2n$ ریشه در بازه $[0, 2\pi]$ دارد. در نتیجه، بنابر قضیه ۱۳.۳، همه این ضرایب برابر با صفر هستند. □

این قضیه نشان می‌دهد که c_k ها و s_k ها یک پایه برای T_n تشکیل می‌دهند.

قضیه ۱۵.۳ برای $2n + 1$ نقطه متمایز داده شده $t_0, t_1, \dots, t_{2n} \in [0, 2\pi]$ و $2n + 1$ مقدار $y_0, y_1, \dots, y_{2n} \in \mathbb{R}$ یک و تنها یک چندجمله‌ای مثلثاتی $q_n \in T_n$ موجود است به گونه‌ای که

$$q_n(t_j) = y_j, \quad j = 0, 1, \dots, 2n.$$

برهان: قرار می‌دهیم

$$\ell_k(t) = \prod_{\substack{i=0 \\ i \neq k}}^{2n} \frac{\sin \frac{t-t_i}{2}}{\sin \frac{t_k-t_i}{2}}, \quad k = 0, 1, \dots, 2n,$$

و

$$q_n(t) = \sum_{k=0}^{2n} y_k \ell_k(t).$$

روشن است که داریم $\ell_k(t_j) = \delta_{kj}$ و بنابراین، $q_n(t_j) = y_j$ ، $j = 0, 1, \dots, 2n$. در اینجا، کافی است نشان دهیم که $q_n \in T_n$ داریم

$$\sin \frac{t-t_0}{2} \sin \frac{t-t_1}{2} = \frac{1}{2} \cos \frac{t_1-t_0}{2} - \frac{1}{2} \cos(t - \frac{t_1+t_0}{2}),$$

و با قراردادن

$$a_0 = \frac{1}{2} \cos \frac{t_1-t_0}{2}, \quad a_1 = -\frac{1}{2} \cos \frac{t_1+t_0}{2}, \quad b_1 = -\frac{1}{2} \sin \frac{t_1+t_0}{2},$$

می‌بینیم که

$$s(t) := \sin \frac{t-t_0}{2} \sin \frac{t-t_1}{2} = a_0 + a_1 \cos t + b_1 \sin t.$$

این رابطه نشان می‌دهد که عبارت $s(t)$ به صورت یک چندجمله‌ای مثلثاتی از درجه یک است. از طرفی، ℓ_k به صورت حاصل ضرب n چندجمله‌ای مثلثاتی از درجه یک است و بنابراین، درجه ℓ_k برابر با n است. این نتیجه نشان می‌دهد که درجه q_n حداکثر n است.

برای نشان دادن یکتایی جواب، فرض کنید p_n و q_n دو چندجمله‌ای مثلثاتی هستند که در شرایط داده شده صدق می‌کنند. قرار می‌دهیم $r \equiv p_n - q_n$. در این صورت

$$r(t_j) = p_n(t_j) - q_n(t_j) = y_j - y_j = 0, \quad j = 0, 1, \dots, 2n.$$

یعنی r دارای $2n+1$ ریشه متمایز در بازه $[a, b]$ است و بنابر قضیه ۱۳.۳، داریم $r \equiv 0$. به این ترتیب، اثبات قضیه کامل می‌شود. \square

نقاط متمایز

$$t_j = \frac{2\pi j}{2n+1}, \quad j = 0, 1, \dots, 2n,$$

را در نظر بگیرید. ابتدا توجه کنید که اگر $k \neq 0$ ، آنگاه (در اینجا $\sqrt{-1}$)

$$\sum_{j=0}^{2n} e^{ikt_j} = \sum_{j=0}^{2n} e^{ijtk} = \sum_{j=0}^{2n} (e^{itk})^j = \frac{1 - e^{i(2n+1)tk}}{1 - e^{itk}} = \frac{1 - e^{2i\pi k}}{1 - e^{itk}} = 0,$$

و اگر $k = 0$ ، آنگاه

$$\sum_{j=0}^{2n} e^{ikt_j} = 2n + 1.$$

بنابراین، داریم

$$\sum_{j=0}^{2n} e^{ikt_j} = \begin{cases} 2n + 1, & k = 0, \\ 0, & k = \pm 1, \pm 2, \dots, \pm 2n. \end{cases} \quad (33.3)$$

اکنون، به دنبال یک چندجمله‌ای یکتا به صورت

$$q_n(t) = \sum_{k=-n}^n \gamma_k e^{ikt},$$

هستیم که در شرایط

$$q_n(t_j) = y_j, \quad j = 0, 1, \dots, 2n,$$

صدق کند. حل این مسئله معادل با حل مسئله

$$\sum_{k=-n}^n \gamma_k e^{ikt_j} = y_j, \quad j = 0, 1, \dots, 2n,$$

است. فرض کنید ضرایب γ_k جواب این دستگاه باشد. در این صورت، با استفاده از (۳۳.۳)، داریم

$$\sum_{j=0}^{2n} y_j e^{-imt_j} = \sum_{k=-n}^n \left(\gamma_k \sum_{j=0}^{2n} e^{i(k-m)t_j} \right) = (2n + 1)\gamma_m,$$

و در نتیجه

$$\gamma_k = \frac{1}{2n + 1} \sum_{j=0}^{2n} y_j e^{-ikt_j}, \quad k = -n, \dots, n. \quad (34.3)$$

از طرفی

$$\begin{aligned} \sum_{k=-n}^n \gamma_k e^{ikt_j} &= \sum_{k=-n}^n \left(\frac{1}{2n + 1} \sum_{p=0}^{2n} y_p e^{-ikt_p} \right) e^{ikt_j} \\ &= \frac{1}{2n + 1} \sum_{p=0}^{2n} y_p \left(\sum_{k=-n}^n e^{ik(t_j - t_p)} \right) \\ &= \frac{1}{2n + 1} y_j (2n + 1) = y_j. \end{aligned}$$

بنابراین، γ_k هایی که در (۳۴.۳) ارائه شده‌اند، جواب یکتای مسئله هستند.

قضیه ۱۶.۳ یک چندجمله‌ای یکتا مثل

$$q_n(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt), \quad (35.3)$$

وجود دارد چنان که

$$q_n\left(\frac{2\pi j}{2n+1}\right) = y_j, \quad j = 0, \dots, 2n.$$

به علاوه، ضرایب این چندجمله‌ای به صورت زیر محاسبه می‌شود

$$a_k = \frac{2}{2n+1} \sum_{j=0}^{2n} y_j \cos \frac{2\pi jk}{2n+1}, \quad k = 0, \dots, n,$$

$$b_k = \frac{2}{2n+1} \sum_{j=0}^{2n} y_j \sin \frac{2\pi jk}{2n+1}, \quad k = 1, \dots, n.$$

برهان: با توجه به قضیه ۱۳.۳، چندجمله‌ای مثلثاتی $q_n(t)$ را می‌توان به صورت

$$q_n(t) = \sum_{k=-n}^n \gamma_k e^{ikt},$$

نمایش داد که در آن

$$\gamma_k = \frac{1}{2} (a_k - ib_k), \quad \gamma_{-k} = \frac{1}{2} (a_k + ib_k).$$

از اینجا این نتیجه به دست می‌آید که

$$a_k = \gamma_k + \gamma_{-k}, \quad b_k = \frac{1}{i} (\gamma_{-k} - \gamma_k).$$

بنابراین، با استفاده از رابطه (۳۴.۳) خواهیم داشت

$$\begin{aligned} a_k &= \frac{1}{2n+1} \sum_{j=0}^{2n} y_j (e^{-ikt_j} + e^{ikt_j}) = \frac{1}{2n+1} \sum_{j=0}^{2n} (y_j \times 2 \cos kt_j) \\ &= \frac{2}{2n+1} \sum_{j=0}^{2n} y_j \cos \frac{2\pi jk}{2n+1}, \end{aligned}$$

$$b_k = \frac{1}{i} \times \frac{1}{2n+1} \sum_{j=0}^{2n} y_j (e^{ikt_j} - e^{-ikt_j}) = \frac{2}{2n+1} \sum_{j=0}^{2n} y_j \sin \frac{2\pi jk}{2n+1},$$

□

که اثبات قضیه را کامل می‌کند.

مثال ۱۵.۳ پنج مقدار از تابع $f(t) = \sin t + \cos^2 t$ در جدول زیر آورده شده است. یک چندجمله‌ای مثلثاتی از درجه دو مثل q_2 بیابید به طوری که $q_2(t_j) = f(t_j) =: y_j$ که $j = 0, 1, \dots, 4$.

j	۰	۱	۲	۳	۴
t_j	۰	$\frac{2\pi}{5}$	$\frac{4\pi}{5}$	$\frac{6\pi}{5}$	$\frac{8\pi}{5}$
$y_j = f(t_j)$	۱	۰.۹۸۱	۰.۰۵۸	-۱.۱۱۷	-۰.۹۲۲

حل: با استفاده از قضیه ۱۶.۳، داریم

$$a_0 = \frac{2}{5} \sum_{j=0}^4 y_j = 0,$$

$$a_1 = \frac{2}{5} \sum_{j=0}^4 y_j \cos \frac{2\pi j}{5} = 0.75,$$

$$a_2 = \frac{2}{5} \sum_{j=0}^4 y_j \cos \frac{4\pi j}{5} = 0.25,$$

$$b_1 = \frac{2}{5} \sum_{j=0}^4 y_j \sin \frac{2\pi j}{5} = 1,$$

$$b_2 = \frac{2}{5} \sum_{j=0}^4 y_j \sin \frac{4\pi j}{5} = 0.$$

بنابراین

$$q_2(t) = 0.75 \cos t + 0.25 \cos 2t + \sin t.$$

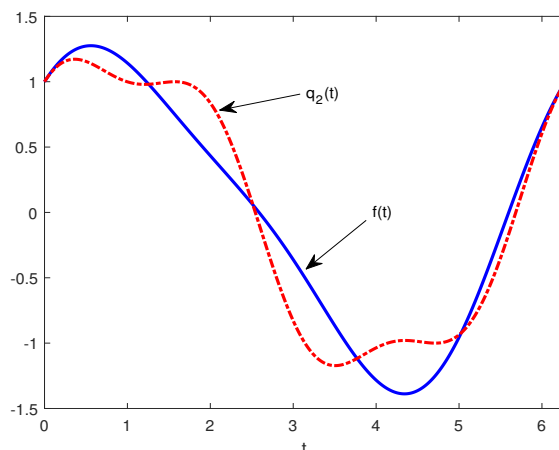
در شکل ۴.۳، نمودار چندجمله‌ای مثلثاتی q_2 به همراه تابع f در بازه $[0, 2\pi]$ رسم شده است. همان‌گونه که می‌بینیم، چندجمله‌ای q_2 یک تقریب مناسب برای تابع f است.

قضیه ۱۷.۳ یک چندجمله‌ای یکتا مثل

$$q_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n-1} (a_k \cos kt + b_k \sin kt) + \frac{a_n}{2} \cos nt,$$

وجود دارد به طوری که

$$q_n\left(\frac{\pi j}{n}\right) = y_j, \quad j = 0, \dots, 2n-1.$$

شکل ۴.۳: نمودار q_2 و f در مثال ۱۵.۳.

به علاوه، ضرایب این چندجمله‌ای از رابطه‌های زیر به دست می‌آید

$$a_k = \frac{1}{n} \sum_{j=0}^{n-1} y_j \cos \frac{\pi j k}{n}, \quad k = 0, \dots, n,$$

$$b_k = \frac{1}{n} \sum_{j=0}^{n-1} y_j \sin \frac{\pi j k}{n}, \quad k = 1, \dots, n-1.$$

برهان: اثبات این قضیه همانند اثبات قضیه ۱۶.۳ است و به خواننده واگذار می‌شود. □

۹.۳ درونیابی توابع چندمتغیره

فرض کنید $(n+1) \times (m+1)$ نقطه متمایز (x_i, y_i) ، $i = 0, 1, \dots, n$ ، $j = 0, 1, \dots, m$ در $\mathbb{R} \times \mathbb{R}$ و مقادیر تابع دومتغیره f در این نقاط داده شده باشند. می‌خواهیم یک چندجمله‌ای مثل $P(x, y)$ که درجه آن نسبت به x و y به ترتیب n و m است، بیابیم به طوری که

$$P(x_i, y_j) = f(x_i, y_j), \quad i = 0, 1, \dots, n, \quad j = 0, 1, \dots, m.$$

برای این کار، قرار می‌دهیم

$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{x - x_k}{x_i - x_k}, \quad i = 0, 1, \dots, n,$$

$$L_j(y) = \prod_{\substack{k=0 \\ k \neq j}}^m \frac{y - y_k}{y_j - y_k}, \quad j = 0, 1, \dots, m.$$

روشن است که

$$L_i(x_k)L_j(y_l) = \delta_{ik}\delta_{jl} = \begin{cases} 1, & i = k, j = l, \\ 0, & \text{otherwise.} \end{cases}$$

اکنون، تعریف می‌کنیم

$$P(x, y) = \sum_{i=0}^n \sum_{j=0}^m f(x_i, y_j) L_i(x) L_j(y).$$

واضح است که درجه چندجمله‌ای P نسبت به x و y به ترتیب برابر با n و m است و به علاوه، در شرط درونیابی نیز صدق می‌کند. این چندجمله‌ای یکتا است؛ زیرا فرض کنید Q چندجمله‌ای دیگری باشد که تابع f را در نقاط داده شده درونیابی کند. تابع $R = P - Q$ تعریف می‌کنیم. داریم

$$\begin{aligned} R(x_r, y_s) &= P(x_r, y_s) - Q(x_r, y_s) \\ &= f(x_r, y_s) - f(x_r, y_s) = 0, \quad r = 0, 1, \dots, m, \quad s = 0, 1, \dots, n. \end{aligned}$$

اکنون، فرض کنید

$$R(x, y) = \sum_{i=0}^n \sum_{j=0}^m \alpha_{ij} x^i y^j.$$

به‌ازای هر $0 \leq r \leq n$ و $0 \leq s \leq m$ ، داریم

$$0 = R(x_r, y_s) = \sum_{i=0}^n \sum_{j=0}^m \alpha_{ij} x_r^i y_s^j = \sum_{i=0}^n x_r^i \underbrace{\left(\sum_{j=0}^m \alpha_{ij} y_s^j \right)}_{:=\beta_i} = \sum_{i=0}^n \beta_i x_r^i.$$

این رابطه نشان می‌دهد که چندجمله‌ای $h(t) = \sum_{i=0}^n \beta_i t^i$ دارای $(n+1)$ ریشه متمایز x_r ، مگر اینکه $h \equiv 0$ است که ناممکن است (توجه کنید که درجه چندجمله‌ای حداکثر n است)، مگر این‌که $h \equiv 0$. بنابراین

$$0 = \beta_i = \sum_{j=0}^m \alpha_{ij} y_s^j, \quad i = 0, 1, \dots, n.$$

این رابطه نشان می‌دهد که به ازای هر i ، چندجمله‌ای $\sum_{j=0}^m \alpha_{ij} y^j$ دارای $(m+1)$ ریشه متمایز $y_s, s = 0, 1, \dots, m$ است که ناممکن است، مگر اینکه $\alpha_{ij} = 0, j = 0, 1, \dots, m$. بنابراین

$$\alpha_{ij} = 0, \quad i = 0, 1, \dots, n, \quad j = 0, 1, \dots, m.$$

از این رو، داریم $R \equiv 0$ ، که یکتایی را نتیجه می‌دهد. □
روند بالا را می‌توان برای توابع چندمتغیره نیز تعمیم داد.

مثال ۱۶.۳ مقادیر تابع دومتغیره f در شش نقطه در جدول زیر داده شده است. چندجمله‌ای درونیاب تابع را در این نقاط بنویسید.

	$y_0 = 0$	$y_1 = 2$
$x_0 = 0$	$f(x_0, y_0) = 2$	$f(x_0, y_1) = -1$
$x_1 = 1$	$f(x_1, y_0) = 1$	$f(x_1, y_1) = 3$
$x_2 = 2$	$f(x_2, y_0) = 2$	$f(x_2, y_1) = 1$

حل: داریم

$$L_0(x) = \frac{(x-1)(x-2)}{(0-1)(0-2)} = \frac{1}{2}(x-1)(x-2),$$

$$L_1(x) = \frac{(x-0)(x-2)}{(1-0)(1-2)} = -x(x-2),$$

$$L_2(x) = \frac{(x-0)(x-1)}{(2-0)(2-1)} = \frac{1}{2}x(x-1),$$

$$L_0(y) = \frac{y-2}{0-2} = -\frac{1}{2}(y-2),$$

$$L_1(y) = \frac{y-0}{2-0} = \frac{1}{2}y.$$

بنابراین، چندجمله‌ای درونیاب تابع در نقاط داده شده به صورت

$$\begin{aligned} p(x, y) &= 2L_0(x)L_0(y) - L_0(x)L_1(y) + L_1(x)L_0(y) + 3L_1(x)L_1(y) \\ &\quad + 2L_2(x)L_0(y) + L_2(x)L_1(y) \\ &= \frac{9}{2}xy - \frac{3}{2}y - 2x - 2x^2y + x^2 + 2, \end{aligned}$$

است.

۱۰.۳ درونیابی با کمک توابع اسپلاین

فرض کنید $a = x_0 < x_1 < \dots < x_n = b$ یک افراز از $[a, b]$ است.

تعریف ۵.۳ یک اسپلاین S_Δ از درجه k روی بازه $[a, b]$ ، یک تابع حقیقی با خواص زیر است:
 (الف) S_Δ در هر زیربازه $[x_i, x_{i+1}]$ ، $i = 0, \dots, n-1$ ، یک چندجمله‌ای حداکثر از درجه k است؛

(ب) $S_\Delta \in C^{k-1}[a, b]$. به عبارت دیگر، تابع S_Δ و مشتقات آن تا مرتبه $k-1$ در بازه $[a, b]$ پیوسته هستند.

اینک، مجموعه نقاط $Y = \{y_0, y_1, \dots, y_n\}$ را در نظر بگیرید. در این صورت، $S_\Delta(Y; \cdot)$ تابع اسپلاین درونیاب می‌گوییم، هرگاه

$$S_\Delta(Y; x_i) = y_i, \quad i = 0, 1, \dots, n.$$

مثال ۱۷.۳ تابع

$$S_\Delta(x) = \begin{cases} 2x^3 - x^2, & x \in [0, 1), \\ 6x^3 - 13x^2 + 12x - 4, & x \in [1, 2], \end{cases}$$

یک تابع اسپلاین درجه ۳ روی $[0, 2]$ است؛ زیرا این تابع در بازه $[0, 2]$ پیوسته است. از طرفی، داریم

$$S'_\Delta(x) = \begin{cases} 6x^2 - 2x, & x \in [0, 1), \\ 18x^2 - 26x + 12, & x \in [1, 2]. \end{cases}$$

اکنون، با توجه به اینکه $S'_\Delta(1) = 4$ ، $\lim_{x \rightarrow 1^-} S'_\Delta(x) = \lim_{x \rightarrow 1^+} S'_\Delta(x) = 4$ ، نتیجه می‌گیریم که S'_Δ در بازه $[0, 2]$ پیوسته است. همچنین، داریم

$$S''_\Delta(x) = \begin{cases} 12x - 2, & x \in [0, 1), \\ 36x - 26, & x \in [1, 2], \end{cases}$$

و $S''_\Delta(1) = 10$ ، $\lim_{x \rightarrow 1^-} S''_\Delta(x) = \lim_{x \rightarrow 1^+} S''_\Delta(x) = 10$ ، که پیوستگی تابع S''_Δ در بازه $[0, 2]$ را نتیجه می‌دهد. بنابراین، S_Δ یک تابع اسپلاین درجه ۳ است.

متداول‌ترین تابع اسپلاین، تابع اسپلاین مکعبی است. بنابراین، در ادامه روی این نوع از اسپلاین متمرکز می‌شویم. با توجه به تعریف ۵.۳، یک تابع اسپلاین مکعبی مثل S_Δ روی بازه $[a, b]$ ، تابعی است که روی هر زیربازه یک چندجمله‌ای از درجه حداکثر ۳ است به طوری که S_Δ ، S'_Δ و S''_Δ در نقاط گرهی پیوسته هستند.

توجه کنید که در حالت کلی، یک تابع اسپلاین مکعبی درونیاب $S_\Delta(Y; \cdot)$ که در شرایط

$$S_\Delta(Y; x_i) = y_i, \quad i = 0, 1, \dots, n,$$

صدق کند، یکتا نیست. هر چند جمله‌ای درجه سه دارای چهار ضریب نامعلوم است و بنابراین، برای تعیین یک تابع اسپلاین مکعبی درونیاب، $4n$ ضریب نامعلوم باید تعیین شود. به دلیل اینکه در بازه $[x_i, x_{i+1}]$ داریم

$$S_{\Delta}(Y; x_i) = y_i, \quad S_{\Delta}(Y; x_{i+1}) = y_{i+1}, \quad i = 0, 1, \dots, n-1,$$

به تعداد $2n$ معادله به دست خواهیم آورد. با توجه به اینکه مشتقات مرتبه اول و دوم در نقاط گرهی پیوسته هستند، به تعداد $2(n-1)$ معادله دیگر خواهیم داشت. بنابراین، تعداد کل معادلات برابر با $4n-2$ خواهد بود. از این رو، برای تعیین $4n$ ضریب نامعلوم، به دو معادله دیگر نیاز خواهیم داشت. ذکر این نکته ضروری است که پیوستگی اسپلاین محدودیت اضافی تولید نمی‌کند. برای دو معادله دیگر معمولاً یکی از شرایط زیر اعمال می‌شود

(الف) $S''_{\Delta}(Y, a) = S''_{\Delta}(Y, b) = 0$ (اسپلاین طبیعی)؛

(ب) $S^{(k)}_{\Delta}(Y, a) = S^{(k)}_{\Delta}(Y, b)$ ، $k = 0, 1, 2$ (اسپلاین متناوب)؛

(ج) $S'_{\Delta}(Y, b) = y'_n$ و $S'_{\Delta}(Y, a) = y'_0$ که در آن y'_n و y'_0 مقادیر داده شده هستند (اسپلاین مقید).

توجه کنید که شرط $S_{\Delta}(Y, a) = S_{\Delta}(Y, b)$ در (ب) تنها زمانی برقرار می‌شود که $y_0 = y_n$ ؛ زیرا در غیر این صورت، S_{Δ} موجود نخواهد بود. خواهیم دید که تحت هر کدام از شرایط بالا، تابع اسپلاین درونیاب مکعبی یکتا خواهد بود.

تعریف ۶.۳ تابع حقیقی $f : [a, b] \rightarrow \mathbb{R}$ را بر بازه $[a, b]$ مطلقاً پیوسته می‌گوییم، هرگاه برای هر $\epsilon > 0$ ، یک $\delta > 0$ موجود باشد، به طوری که برای هر مجموعه متناهی از بازه‌های $[a_i, b_i]$ ، $i = 1, 2, \dots, n$ که در شرط

$$a \leq a_1 < b_1 < a_2 < b_2 < \dots < a_n < b_n = b,$$

$$\sum_{i=1}^n |f(b_i) - f(a_i)| < \epsilon \quad \text{صدق می‌کند، داشته باشیم} \quad \sum_{i=1}^n |b_i - a_i| < \delta$$

اگر تابع در بازه بسته $[a, b]$ مشتق پذیر باشد یا در حالت کلی در شرط لپشیتز

$$|f(x_1) - f(x_2)| \leq \theta |x_1 - x_2|, \quad \forall x_1, x_2 \in [a, b]$$

که در آن، θ یک مقدار ثابت است، صدق کند، آنگاه مطلقاً پیوسته خواهد بود. اگر تابع f مطلقاً پیوسته باشد، آنگاه پیوسته است و مشتق آن، یعنی f' ، تقریباً همه جا موجود است و به علاوه، خواهیم داشت

$$f(x) = f(a) + \int_a^x f'(t) dt, \quad \forall x \in [a, b].$$

همچنین، اگر f و g دو تابع مطلقاً پیوسته باشند، آنگاه از قاعده انتگرال گیری جزء به جزء برای محاسبه

$$\int_a^b f(x)g'(x)dx$$

می توان استفاده کرد، یعنی

$$\int_a^b f(x)g'(x)dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx.$$

تعریف ۷.۳ مجموعه توابع $\mathbb{R} : [a, b] \rightarrow \mathbb{R}$ که مربع آنها بر $[a, b]$ انتگرال پذیر هستند، یعنی

$$\int_a^b |f(x)|^2 dx$$

موجود و متناهی هستند، را با $L^2[a, b]$ نشان می دهیم.

تعریف ۸.۳ فرض کنید m یک عدد طبیعی است. مجموعه توابع $\mathbb{R} : [a, b] \rightarrow \mathbb{R}$ به طوری که $f^{(m-1)}$ روی $[a, b]$ مطلقاً پیوسته و $f^{(m)} \in L^2[a, b]$ را با $\mathcal{K}^m[a, b]$ نشان می دهیم. همچنین، مجموعه همه توابع $\mathbb{R} : [a, b] \rightarrow \mathbb{R}$ به طوری که $f \in \mathcal{K}^m[a, b]$ و

$$f^{(k)}(a) = f^{(k)}(b), \quad k = 1, \dots, m-1,$$

را با $\mathcal{K}_p^m[a, b]$ نشان می دهیم.

یادداشت ۲.۳ توجه کنید که اگر $S_\Delta(Y; \cdot)$ یک تابع اسپلاین مکعبی درونیاب باشد، آنگاه $S_\Delta(Y; \cdot) \in \mathcal{K}^3[a, b]$ و در صورتی که شرط (ب) برقرار باشد، داریم $S_\Delta(Y; \cdot) \in \mathcal{K}_p^3[a, b]$.

برای هر $f \in \mathcal{K}^2[a, b]$ ، تعریف می کنیم

$$\|f\|^2 = \int_a^b |f''(x)|^2 dx.$$

داریم $\|f\| \geq 0$. شایان ذکر است که $\|\cdot\|$ یک نرم نیست؛ زیرا برای هر تابع خطی $f(x) = \alpha x + \beta$ ، داریم $\|f\| = 0$. اینک، با توجه به مطالب ارائه شده، می توانیم اتحاد هولادی را ارائه کنیم.

قضیه ۱۸.۳ (اتحاد هولادی) فرض کنید $f \in \mathcal{K}^2[a, b]$ و

$$\Delta = \{a = x_0 < x_1 < \dots < x_n = b\},$$

یک افراز از $[a, b]$ است. اگر S_Δ یک تابع اسپلاین مکعبی با گره های $x_i \in \Delta$ باشد، آنگاه

$$\|f - S_\Delta\|^2 = \|f\|^2 - \|S_\Delta\|^2 - 2 \left[(f'(x) - S'_\Delta(x)) S''_\Delta(x) \Big|_a^b - \sum_{i=1}^n (f(x) - S_\Delta(x)) S'''_\Delta(x) \Big|_{x_{i-1}^+}^{x_i^-} \right].$$

توجه کنید که S'''_Δ لزوماً پیوسته نیست و مقادیر آن در انتهای هر بازه و ابتدای بازه بعدی یکسان نیست. بنابراین، از نماد $S_\Delta|_{x_{i-1}^+}^{x_i^-}$ استفاده شده است.

برهان: داریم

$$\begin{aligned} \|f - S_\Delta\|^2 &= \int_a^b |f''(x) - S''_\Delta(x)|^2 dx \\ &= \|f\|^2 - 2 \int_a^b f''(x) S''_\Delta(x) dx + \|S_\Delta\|^2 \\ &= \|f\|^2 - \|S_\Delta\|^2 - 2 \int_a^b (f''(x) - S''_\Delta(x)) S''_\Delta(x) dx. \end{aligned}$$

از طرفی، داریم

$$\begin{aligned} \int_{x_{i-1}}^{x_i} (f''(x) - S''_\Delta(x)) S''_\Delta(x) dx &= (f'(x) - S'_\Delta(x)) S''_\Delta(x) \Big|_{x_{i-1}}^{x_i} \\ &\quad - \int_{x_{i-1}}^{x_i} (f'(x) - S'_\Delta(x)) S'''_\Delta(x) dx \\ &= (f'(x) - S'_\Delta(x)) S''_\Delta(x) \Big|_{x_{i-1}}^{x_i} \\ &\quad - (f(x) - S_\Delta(x)) S'''_\Delta(x) \Big|_{x_{i-1}^+}^{x_i^-} \\ &\quad + \int_{x_{i-1}}^{x_i} (f(x) - S_\Delta(x)) S^{(4)}_\Delta(x) dx. \end{aligned}$$

در هر زیربازه $[x_{i-1}, x_i]$ ، داریم $S_\Delta^{(4)} \equiv 0$. از طرفی، توابع f' ، S'_Δ و S''_Δ در بازه $[a, b]$ پیوسته هستند. بنابراین، خواهیم داشت

$$\sum_{i=1}^n (f'(x) - S'_\Delta(x)) S''_\Delta(x) \Big|_{x_{i-1}}^{x_i} = (f'(x) - S'_\Delta(x)) S''_\Delta(x) \Big|_a^b.$$

□ به این ترتیب، اثبات قضیه کامل می‌شود.

قضیه ۱۹.۳ (خاصیت مینیمم نرم) فرض کنید $\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}$ یک افراز از $[a, b]$ باشد و مقادیر $Y = \{y_0, y_1, \dots, y_n\}$ و تابع $f \in \mathcal{K}^2[a, b]$ داده شده باشند، به طوری

که $f(x_i) = y_i$ ، $i = 0, 1, \dots, n$. در این صورت، اگر یکی از شرایط

(الف) $S''_\Delta(Y, a) = S''_\Delta(Y, b) = 0$ ؛

(ب) $f \in \mathcal{K}_p^2[a, b]$ و $S_\Delta(Y; \cdot)$ متناوب باشد؛

(ج) $S'_\Delta(Y, a) = f'(a)$ و $S'_\Delta(Y, b) = f'(b)$ ؛

برقرار باشد، آنگاه $\|f\|^2 \geq \|S_\Delta(Y; \cdot)\|^2$ ، یا به طور دقیق‌تر

$$\|f - S_\Delta(Y; \cdot)\|^2 = \|f\|^2 - \|S_\Delta(Y; \cdot)\|^2 \geq 0. \quad (۳۶.۳)$$

به علاوه، در هر یک از حالت‌های (الف)، (ب) و (ج) تابع $S_\Delta(Y; \cdot)$ به طور یکتا تعیین می‌شود.

برهان: در هریک از سه حالت (الف)، (ب) و (ج)، داریم

$$\begin{aligned} f(x_{i-1}^+) - S_{\Delta}(Y; x_{i-1}^+) &= 0, \\ f(x_i^-) - S_{\Delta}(Y; x_i^-) &= 0. \end{aligned}$$

بنابراین، در هریک از این سه حالت، خواهیم داشت

$$\begin{aligned} g &:= (f'(x) - S'_{\Delta}(Y; x)) S''_{\Delta}(Y; x) \Big|_a^b - \sum_{i=1}^n (f(x) - S_{\Delta}(Y; x)) S'''_{\Delta}(Y; x) \Big|_{x_{i-1}^+}^{x_i^-} \\ &= (f'(x) - S'_{\Delta}(Y; x)) S''_{\Delta}(Y; x) \Big|_a^b \\ &= (f'(b) - S'_{\Delta}(Y; b)) S''_{\Delta}(Y; b) - (f'(a) - S'_{\Delta}(Y; a)) S''_{\Delta}(Y; a). \end{aligned}$$

در حالت (الف)، داریم $S''_{\Delta}(Y; a) = S''_{\Delta}(Y; b) = 0$. بنابراین، در این حالت، داریم $g = 0$. اکنون، حالت (ب) را در نظر می‌گیریم. از اینکه $f \in \mathcal{K}_p^{\chi}[a, b]$ ، خواهیم داشت $f(a) = f(b)$ و $f'(a) = f'(b)$. با توجه به اینکه S_{Δ} متناوب است، نتیجه می‌گیریم که $S'_{\Delta}(Y; a) = S'_{\Delta}(Y; b)$ و $S''_{\Delta}(Y; a) = S''_{\Delta}(Y; b)$. با استفاده از این رابطه‌ها، روشن است که $g = 0$. در حالت (ج)، پیداست که $g = 0$. اکنون، با استفاده از قضیه ۱۸.۳ در هریک از سه حالت (الف)، (ب) و (ج)، رابطه (۳۶.۳) نتیجه می‌شود.

برای اثبات یکتایی، فرض کنید $\bar{S}_{\Delta}(Y; \cdot)$ تابع اسپلاین درونیاب مکعبی دیگری باشد. در رابطه (۳۶.۳) به جای f ، تابع $\bar{S}_{\Delta}(Y; \cdot)$ را قرار می‌دهیم

$$\|\bar{S}_{\Delta}(Y; \cdot) - S_{\Delta}(Y; \cdot)\|^2 = \|\bar{S}_{\Delta}(Y; \cdot)\|^2 - \|S_{\Delta}(Y; \cdot)\|^2 \geq 0. \quad (37.3)$$

با تعویض نقش S_{Δ} و \bar{S}_{Δ} ، خواهیم داشت

$$\|S_{\Delta}(Y; \cdot) - \bar{S}_{\Delta}(Y; \cdot)\|^2 = \|S_{\Delta}(Y; \cdot)\|^2 - \|\bar{S}_{\Delta}(Y; \cdot)\|^2 \geq 0.$$

در نتیجه، با استفاده از این رابطه و رابطه (۳۷.۳)، خواهیم داشت

$$\|S_{\Delta}(Y; \cdot) - \bar{S}_{\Delta}(Y; \cdot)\|^2 = \int_a^b |S''_{\Delta}(Y; x) - \bar{S}''_{\Delta}(Y; x)|^2 dx = 0.$$

این رابطه نشان می‌دهد که به ازای هر $x \in [a, b]$

$$\bar{S}''_{\Delta}(Y; x) - S''_{\Delta}(Y; x) = 0 \implies \bar{S}_{\Delta}(Y; x) = S_{\Delta}(Y; x) + cx + d.$$

اکنون، از اینکه $\bar{S}_\Delta(Y; a) = S_\Delta(Y; a)$ و $\bar{S}_\Delta(Y; b) = S_\Delta(Y; b)$ داریم

$$\begin{cases} ca + d = 0, \\ cb + d = 0. \end{cases}$$

با توجه به اینکه $a \neq b$ ، از حل این دستگاه خواهیم داشت $c = d = 0$ ، که یکتایی اسپلاین را نتیجه می‌دهد. \square

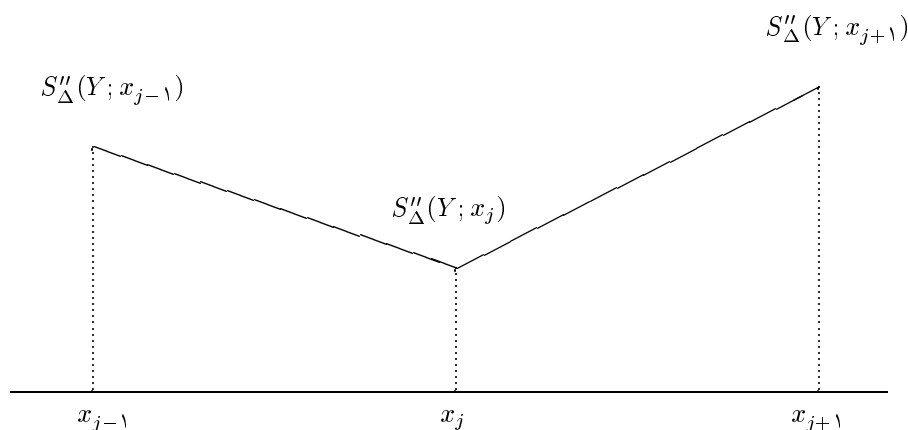
۱۱.۳ محاسبه اسپلاین مکعبی درونیاب

فرض کنید $\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}$ یک افراز از $[a, b]$ و مقادیر $Y = \{y_0, y_1, \dots, y_n\}$ داده شده باشند. قرار می‌دهیم $h_{j+1} = x_{j+1} - x_j$ ، $j = 0, 1, \dots, n-1$ و

$$M_j := S''_\Delta(Y; x_j), \quad j = 0, 1, \dots, n.$$

M_j را ممان j ام می‌گویند. با توجه به اینکه $S''_\Delta(Y; x)$ در هر زیربازه $[x_j, x_{j+1}]$ خطی است، داریم (به شکل زیر توجه کنید)

$$S''_\Delta(Y; x) = M_j \frac{x_{j+1} - x}{h_{j+1}} + M_{j+1} \frac{x - x_j}{h_{j+1}}, \quad x \in [x_j, x_{j+1}].$$



با انتگرال‌گیری از این تابع، داریم

$$S'_{\Delta}(Y; x) = -M_j \frac{(x_{j+1} - x)^2}{2h_{j+1}} + M_{j+1} \frac{(x - x_j)^2}{2h_{j+1}} + A_j,$$

که در آن A_j یک مقدار ثابت است. با انتگرال‌گیری از طرفین این رابطه، خواهیم داشت

$$S_{\Delta}(Y; x) = M_j \frac{(x_{j+1} - x)^3}{6h_{j+1}} + M_{j+1} \frac{(x - x_j)^3}{6h_{j+1}} + A_j(x - x_j) + B_j,$$

که در آن B_j یک مقدار ثابت است. با توجه به اینکه $S_{\Delta}(Y; x_j) = y_j$ و $S_{\Delta}(Y; x_{j+1}) = y_{j+1}$ نتیجه می‌گیریم

$$\begin{cases} M_j \frac{h_{j+1}^3}{6} + B_j = y_j, \\ M_{j+1} \frac{h_{j+1}^3}{6} + A_j h_{j+1} + B_j = y_{j+1}. \end{cases}$$

با حل این دستگاه، خواهیم داشت

$$\begin{cases} A_j = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{h_{j+1}}{6}(M_{j+1} - M_j), \\ B_j = y_j - M_j \frac{h_{j+1}^3}{6}. \end{cases}$$

بنابراین

$$S_{\Delta}(Y; x) = \alpha_j + \beta_j(x - x_j) + \gamma_j(x - x_j)^2 + \delta_j(x - x_j)^3, \quad x \in [x_j, x_{j+1}], \quad (38.3)$$

که در آن

$$\begin{aligned} \alpha_j &= y_j, \\ \beta_j &= -\frac{M_j h_{j+1}}{2} + A_j = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{2M_j + M_{j+1}}{6} h_{j+1}, \\ \gamma_j &= \frac{M_j}{2}, \\ \delta_j &= \frac{M_{j+1} - M_j}{6h_{j+1}}. \end{aligned}$$

با جای‌گذاری این مقادیر در (38.3) و سپس مشتق‌گیری از آن، خواهیم داشت

$$S'_{\Delta}(Y; x) = -M_j \frac{(x_{j+1} - x)^2}{2h_{j+1}} + M_{j+1} \frac{(x - x_j)^2}{2h_{j+1}} + \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{h_{j+1}}{6}(M_{j+1} - M_j).$$

اکنون، با توجه به پیوستگی $S'_{\Delta}(Y; x)$ در گره x_j داریم

$$S'_{\Delta}(Y; x_j^-) = S'_{\Delta}(Y; x_j^+), \quad j = 1, 2, \dots, n-1. \quad (39.3)$$

بنابراین، با توجه به

$$S'_{\Delta}(Y; x_j^-) = \frac{y_j - y_{j-1}}{h_j} + \frac{h_j}{3}M_j + \frac{h_j}{6}M_{j-1}, \quad (40.3)$$

$$S'_{\Delta}(Y; x_j^+) = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{h_{j+1}}{3}M_j - \frac{h_{j+1}}{6}M_{j+1}, \quad (41.3)$$

و رابطه (۳۹.۳)، خواهیم داشت

$$\frac{h_j}{6}M_{j-1} + \frac{h_j + h_{j+1}}{3}M_j + \frac{h_{j+1}}{6}M_{j+1} = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j},$$

یا

$$\frac{h_j}{h_j + h_{j+1}}M_{j-1} + 2M_j + \frac{h_{j+1}}{h_j + h_{j+1}}M_{j+1} = \frac{6}{h_j + h_{j+1}} \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j} \right), \quad (42.3)$$

که در آن $j = 1, 2, \dots, n-1$. بنابراین، با فرض

$$\begin{cases} \mu_j = \frac{h_j}{h_j + h_{j+1}}, & \lambda_j = \frac{h_{j+1}}{h_j + h_{j+1}}, \\ d_j = \frac{6}{h_j + h_{j+1}} \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j} \right), \end{cases} \quad j = 1, 2, \dots, n-1,$$

دستگاه (۴۲.۳) را می‌توان به صورت زیر نوشت

$$\mu_j M_{j-1} + 2M_j + \lambda_j M_{j+1} = d_j, \quad j = 1, 2, \dots, n-1, \quad (43.3)$$

که در آن $\mu_j + \lambda_j = 1$. این دستگاه شامل $n-1$ معادله و $n+1$ مجهول است. برای دو معادله دیگر، از شرایط اضافی (الف)، (ب) یا (ج) استفاده می‌کنیم.

حالت (الف): در این حالت، داریم $S''_{\Delta}(Y; b) = M_n = 0 = M_0 = S''_{\Delta}(Y; a)$. در نتیجه، معادله (۴۳.۳) به صورت

$$\begin{pmatrix} 2 & \lambda_1 & & & & \\ \mu_2 & 2 & \lambda_2 & & & \\ & \mu_3 & 2 & \ddots & & \\ & & \ddots & \ddots & \lambda_{n-2} & \\ & & & \mu_{n-1} & 2 & \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ \vdots \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{n-1} \end{pmatrix}, \quad (44.3)$$

نوشته می‌شود. با حل این دستگاه، M_j ها و سپس با استفاده از آنها، $S_{\Delta}(Y; \cdot)$ در هر زیربازه $[x_j, x_{j+1}]$ توسط (۳۸.۳) تعیین می‌شود.

حالت (ب): از شرط $S''_{\Delta}(Y; a) = S''_{\Delta}(Y; b)$ نتیجه می‌شود که $M_0 = M_n$ ، و با استفاده از آن، معادله اول از مجموعه معادلات (۴۳.۳) به صورت

$$2M_1 + \lambda_1 M_2 + \mu_1 M_n = d_1. \quad (45.3)$$

نوشته می‌شود. به همین ترتیب، از شرط $S'_{\Delta}(Y; a) = S'_{\Delta}(Y; b)$ داریم

$$\frac{h_n}{6} M_{n-1} + \frac{h_n + h_1}{3} M_n + \frac{h_1}{6} M_1 = \frac{y_1 - y_n}{h_1} - \frac{y_n - y_{n-1}}{h_n},$$

یا

$$\frac{h_n}{h_1 + h_n} M_{n-1} + 2M_n + \frac{h_1}{h_1 + h_n} M_1 = \frac{6}{h_1 + h_n} \left(\frac{y_1 - y_n}{h_1} - \frac{y_n - y_{n-1}}{h_n} \right).$$

این معادله با قراردادن

$$\begin{cases} \mu_n = \frac{h_n}{h_1 + h_n}, & \lambda_n = \frac{h_1}{h_1 + h_n}, \\ d_n = \frac{6}{h_1 + h_n} \left(\frac{y_1 - y_n}{h_1} - \frac{y_n - y_{n-1}}{h_n} \right), \end{cases}$$

به صورت

$$\mu_n M_{n-1} + 2M_n + \lambda_n M_1 = d_n, \quad (46.3)$$

خواهد بود که در آن $\mu_n + \lambda_n = 1$. بنابراین، با استفاده از رابطه‌های (۴۳.۳)، (۴۵.۳) و (۴۶.۳)، دستگاه حاصل برای محاسبه M_j ها به صورت زیر خواهد بود

$$\begin{pmatrix} 2 & \lambda_1 & & & \mu_1 \\ \mu_2 & 2 & \lambda_2 & & \\ & \mu_3 & 2 & \ddots & \\ & & \ddots & \ddots & \lambda_{n-1} \\ \lambda_n & & & \mu_n & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ \vdots \\ M_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_n \end{pmatrix}. \quad (47.3)$$

حالت (ج): با استفاده از شرط $S'_{\Delta}(Y; a) = y'_0$ ، خواهیم داشت

$$\frac{h_1}{3} M_0 + \frac{h_1}{6} M_1 = \frac{y_1 - y_0}{h_1} - y'_0,$$

یا

$$2M_0 + M_1 = \frac{6}{h_1} \left(\frac{y_1 - y_0}{h_1} - y'_0 \right).$$

این معادله با فرض

$$\lambda_0 = 1, \quad d_0 = \frac{1}{h_1} \left(\frac{y_1 - y_0}{h_1} - y'_0 \right),$$

به صورت

$$2M_0 + \lambda_0 M_1 = d_0, \quad (48.3)$$

خواهد بود. به همین ترتیب، با استفاده از $y'_n = S'_\Delta(Y; b)$ داریم

$$\frac{h_n}{1} M_{n-1} + \frac{h_n}{3} M_n = y'_n - \frac{y_n - y_{n-1}}{h_n},$$

یا

$$M_{n-1} + 2M_n = \frac{1}{h_n} \left(y'_n - \frac{y_n - y_{n-1}}{h_n} \right).$$

این معادله با فرض

$$\mu_n = 1, \quad d_n = \frac{1}{h_n} \left(y'_n - \frac{y_n - y_{n-1}}{h_n} \right),$$

به صورت

$$\mu_n M_{n-1} + 2M_n = d_n, \quad (49.3)$$

خواهد بود. اکنون، با استفاده از معادله‌های (۴۳.۳)، (۴۸.۳) و (۴۹.۳)، دستگاه زیر نتیجه می‌شود

$$\begin{pmatrix} 2 & \lambda_0 & & & & \\ \mu_1 & 2 & \lambda_1 & & & \\ & \mu_2 & 2 & \ddots & & \\ & & \ddots & \ddots & \lambda_{n-1} & \\ & & & \mu_n & 2 & \end{pmatrix} \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ \vdots \\ M_n \end{pmatrix} = \begin{pmatrix} d_0 \\ d_1 \\ \vdots \\ \vdots \\ d_n \end{pmatrix}. \quad (50.3)$$

همان‌گونه که ملاحظه شد، در هر سه حالت، داریم $\lambda_i + \mu_i = 1$ ، $\lambda_i \geq 0$ و $\mu_i \geq 0$. توجه کنید که در هر سه حالت، ماتریس ضرایب دستگاه تنها به x_i ها بستگی دارد و مستقل از y_i ها است.

تعریف ۹.۳ ماتریس $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ را غالب قطری اکید می‌گوییم، هرگاه

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, \quad i = 1, 2, \dots, n.$$

قضیه ۲۰.۳ اگر ماتریس A غالب قطری اکید باشد، آنگاه نامنفرد است.

برهان: به برهان خُلف فرض کنید که ماتریس $A = (a_{ij})$ غالب قطری اکید و منفرد است. در این صورت، یک بردار ناصفر مثل $x \in \mathbb{C}^n$ وجود دارد به طوری که $Ax = 0$. بدون از دست دادن کلیت، می‌توان فرض کرد که $\|x\|_\infty = |x_i| = 1$ داریم

$$\begin{aligned} 0 &= \left| \sum_{j=1}^n a_{ij}x_j \right| = |a_{ii}x_i + \sum_{j=1, j \neq i}^n a_{ij}x_j| \\ &\geq |a_{ii}| - \left| \sum_{j=1, j \neq i}^n a_{ij}x_j \right| \geq |a_{ii}| - \sum_{j=1, j \neq i}^n |a_{ij}||x_j| \\ &\geq |a_{ii}| - \sum_{j=1, j \neq i}^n |a_{ij}|. \end{aligned}$$

در نتیجه

$$|a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}|,$$

□ که با غالب قطری بودن ماتریس A در تناقض است. بنابراین، ماتریس A نامنفرد است.

در هر یک از حالت‌های (الف)، (ب) و (ج)، دستگاه‌های به دست آمده در درونیابی اسپلاین غالب قطری اکید هستند. در نتیجه، با توجه به قضیه ۲۰.۳ نامنفردند.

ماتریس ضرایب دستگاه‌های (۴۴.۳) و (۵۰.۳)، سه قطری هستند. در ادامه، الگوریتم توماس برای حل شکل کلی دستگاه معادلات خطی با ماتریس ضرایب سه قطری را ارائه می‌کنیم [۱۵].

دستگاه معادلات خطی زیر را در نظر بگیرید

$$\begin{pmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & \ddots & & \\ & & \ddots & \ddots & c_{n-1} & \\ & & & a_n & b_n & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_n \end{pmatrix}.$$

ماتریس افزوده زیر را تشکیل می‌دهیم

$$\left(\begin{array}{cccccc|c} b_1 & c_1 & & & & & d_1 \\ a_2 & b_2 & c_2 & & & & d_2 \\ & a_3 & b_3 & \ddots & & & d_3 \\ & & \ddots & \ddots & c_{n-1} & & \vdots \\ & & & a_n & b_n & & d_n \end{array} \right).$$

با کمک روش حذفی گاوس، ماتریس ضرایب این دستگاه را بالامثلثی می‌کنیم. با توجه به ساختار ماتریس، کافی است درایه‌های a_i ، $i = 2, 3, \dots, n$ را صفر کنیم. در اینجا، کمی متفاوت با الگوریتم حذفی گاوس عمل می‌کنیم. به این صورت که، پس از حذف هر درایه a_i ، ماتریس و بردار سمت راست به دست آمده را روی همان دستگاه قبلی کپی می‌کنیم. در مرحله نخست، قرار می‌دهیم

$$m = \frac{a_2}{b_1}.$$

سپس، m برابر سطر نخست را از سطر دوم کم می‌کنیم. در این صورت، برای به‌هنگام کردن ماتریس افزوده قرار می‌دهیم

$$b_2 \leftarrow b_2 - mc_1,$$

$$d_2 \leftarrow d_2 - md_1.$$

با انجام این کار، ماتریس افزوده به صورت زیر خواهد بود

$$\left(\begin{array}{cccc|c} b_1 & c_1 & & & d_1 \\ \circ & b_2 & c_2 & & d_2 \\ & a_2 & b_2 & \ddots & d_2 \\ & & \ddots & \ddots & \vdots \\ & & & a_n & b_n & d_n \end{array} \right).$$

در مرحله بعد، قرار می‌دهیم

$$m = \frac{a_3}{b_2}.$$

و سپس، m برابر سطر دوم را از سطر سوم کم می‌کنیم. در این صورت، برای به‌هنگام کردن ماتریس افزوده، قرار می‌دهیم

$$b_3 \leftarrow b_3 - mc_2,$$

$$d_3 \leftarrow d_3 - md_2.$$

با انجام این کار، ماتریس افزوده به صورت زیر خواهد بود

$$\left(\begin{array}{cccc|c} b_1 & c_1 & & & d_1 \\ \circ & b_2 & c_2 & & d_2 \\ & \circ & b_3 & \ddots & d_3 \\ & & \ddots & \ddots & \vdots \\ & & & a_n & b_n & c_{n-1} \\ & & & & & d_n \end{array} \right).$$

با ادامه این روند، در $n - 1$ مرحله، ماتریس افزوده به صورت زیر به دست خواهد آمد

$$\left(\begin{array}{cccc|c} b_1 & c_1 & & & d_1 \\ & b_2 & c_2 & & d_2 \\ & & b_3 & \ddots & d_3 \\ & & & \ddots & \vdots \\ & & & & c_{n-1} \\ & & & & b_n & d_n \end{array} \right).$$

در واقع، با کمک حلقه زیر می توانیم b_i ها و d_i ها را به هنگام کنیم

for $k = 2, 3, \dots, n$

$$m \leftarrow a_k / b_{k-1}$$

$$b_k \leftarrow b_k - mc_{k-1}$$

$$d_k \leftarrow d_k - md_{k-1}$$

end

سپس، دستگاه بالادوقطری به دست آمده بالا را به سادگی با جای گذاری پسرو می توان حل کرد. از معادله آخر، داریم

$$x_n = \frac{d_n}{b_n}.$$

معادله k ام دستگاه به دست آمده به صورت زیر است

$$b_k x_k + c_k x_{k+1} = d_k.$$

بنابراین، داریم

$$x_k = \frac{d_k - c_k x_{k+1}}{b_k}, \quad k = n-1, n-2, \dots, 1.$$

الگوریتم مشابه دیگری با نام الگوریتم رابنش برای حل دستگاه (۴۷.۳) وجود دارد که کمی پیچیده تر از الگوریتم توماس است و از ارائه آن صرف نظر می کنیم [۲].

مثال ۱۸.۳ با استفاده از یک تابع اسپلاین مکعبی طبیعی، تابع $f(x) = |x|$ را در نقاط $\Delta = \{x_0 = -1 < x_1 = 0 < x_2 = 1 < x_3 = 3\}$ درونیابی کنید.

حل: داریم $(x_0, y_0) = (-1, 1)$ ، $(x_1, y_1) = (0, 0)$ ، $(x_2, y_2) = (1, 1)$ و $(x_3, y_3) = (3, 3)$. بنابراین

$$h_1 = x_1 - x_0 = 1, \quad h_2 = x_2 - x_1 = 1, \quad h_3 = x_3 - x_2 = 2.$$

با توجه به اینکه $M_0 = M_3 = 0$ ، دستگاه حاصل برای محاسبه M_1 و M_2 به صورت زیر خواهد بود

$$\begin{pmatrix} 2 & \lambda_1 \\ \mu_2 & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix},$$

که در آن

$$\begin{aligned} \lambda_1 &= \frac{h_2}{h_1 + h_2} = \frac{1}{2}, & \mu_2 &= \frac{h_2}{h_2 + h_3} = \frac{1}{3}, \\ d_1 &= \frac{6}{h_1 + h_2} \left(\frac{y_2 - y_1}{h_2} - \frac{y_1 - y_0}{h_1} \right) = 6, \\ d_2 &= \frac{6}{h_2 + h_3} \left(\frac{y_3 - y_2}{h_3} - \frac{y_2 - y_1}{h_2} \right) = 0. \end{aligned}$$

بنابراین، با حل دستگاه

$$\begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{3} & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix},$$

خواهیم داشت $M_1 = 3.1304$ و $M_2 = -0.5217$. در این صورت، برای محاسبه S_Δ در بازه $[-1, 0]$ داریم

$$\begin{aligned} A_0 &= \frac{y_1 - y_0}{h_1} - \frac{h_1}{6}(M_1 - M_0) = -1.5217, \\ B_0 &= y_0 - M_0 \frac{h_1^2}{6} = 1, \\ \alpha_0 &= y_0 = 1, \\ \beta_0 &= -\frac{M_0 h_1}{6} + A_0 = -1.5217, \\ \gamma_0 &= \frac{M_0}{6} = 0, \\ \delta_0 &= \frac{M_1 - M_0}{6h_1} = 0.5217. \end{aligned}$$

از این رو

$$\begin{aligned} S_{\Delta}|_{[-1,0]} &= \alpha_0 + \beta_0(x - x_0) + \gamma_0(x - x_0)^2 + \delta_0(x - x_0)^3 \\ &= 1 - 1.5217(x + 1) + 0 \times (x + 1)^2 + 0.5217(x + 1)^3. \end{aligned}$$

به همین ترتیب، در بازه $[0, 1]$ داریم

$$\begin{aligned} A_1 &= \frac{y_2 - y_1}{h_2} - \frac{h_2}{6}(M_2 - M_1) = 1.6087, \\ B_1 &= y_1 - M_1 \frac{h_2^2}{6} = -0.5217, \\ \alpha_1 &= y_1 = 0, \\ \beta_1 &= -\frac{M_1 h_2}{2} + A_1 = 0.0435, \\ \gamma_1 &= \frac{M_1}{2} = 1.5652, \\ \delta_1 &= \frac{M_2 - M_1}{6h_2} = -0.6087. \end{aligned}$$

در نتیجه

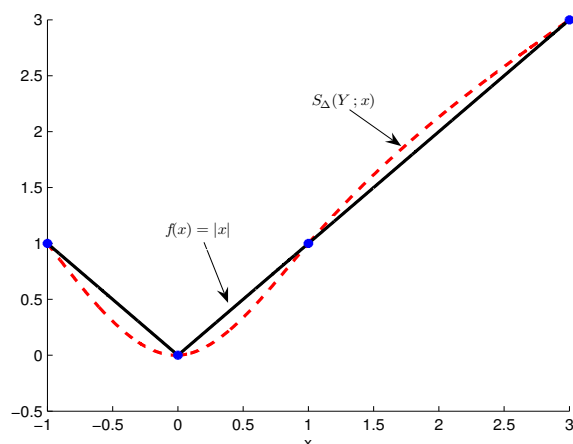
$$\begin{aligned} S_{\Delta}|_{[0,1]} &= \alpha_1 + \beta_1(x - x_1) + \gamma_1(x - x_1)^2 + \delta_1(x - x_1)^3 \\ &= 0 + 0.0435x + 1.5652x^2 - 0.6087x^3. \end{aligned}$$

به همین ترتیب، داریم

$$\begin{aligned} A_2 &= \frac{y_2 - y_2}{h_2} - \frac{h_2}{6}(M_2 - M_2) = 0.8261, \\ B_2 &= y_2 - M_2 \frac{h_2^2}{6} = 1.3478, \\ \alpha_2 &= y_2 = 1, \\ \beta_2 &= -\frac{M_2 h_2}{2} + A_2 = 1.3478, \\ \gamma_2 &= \frac{M_2}{2} = -0.2609, \\ \delta_2 &= \frac{M_2 - M_2}{6h_2} = 0.0435. \end{aligned}$$

بنابراین

$$\begin{aligned} S_{\Delta}|_{[1,2]} &= \alpha_2 + \beta_2(x - x_2) + \gamma_2(x - x_2)^2 + \delta_2(x - x_2)^3 \\ &= 1 + 1.3478(x - 1) - 0.2609(x - 1)^2 + 0.0435(x - 1)^3. \end{aligned}$$



شکل ۵.۳: نمودار تابع $f(x) = |x|$ و تابع اسپلاین مکعبی طبیعی درونیاب (مثال ۱۸.۳).

با جمع بندی نتایج به دست آمده، خواهیم داشت

$$S_{\Delta}(Y; x) = \begin{cases} 1 - 1.5217(x+1) + 0.5217(x+1)^3, & -1 \leq x \leq 0, \\ 0.0435x + 1.5652x^2 - 0.6087x^3, & 0 \leq x \leq 1, \\ 1 + 1.3478(x-1) - 0.2609(x-1)^2 + 0.0435(x-1)^3, & 1 \leq x \leq 3. \end{cases}$$

در شکل ۵.۳، نمودار S_{Δ} به همراه f نمایش داده شده است.

در صورتی که نقاط درونیابی متساوی الفاصله باشند، فرمول های درونیابی شکل ساده تری خواهند داشت. فرض کنید $h = x_{j+1} - x_j$, $j = 0, 1, \dots, n-1$. در این صورت، مقادیر μ_j و λ_j از رابطه های زیر به دست خواهند آمد

$$\begin{cases} \mu_j = \frac{1}{4}, & \lambda_j = \frac{1}{4}, \\ d_j = \frac{3}{h^3} (y_{j+1} - 2y_j + y_{j-1}). \end{cases} \quad j = 1, 2, \dots, n-1,$$

در نتیجه، برای محاسبه M_i ها در اسپلاین مکعبی طبیعی کافی است دستگاه معادلات خطی

$$\begin{pmatrix} 4 & 1 & & \\ 1 & 4 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 4 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} 2d_1 \\ 2d_2 \\ \vdots \\ 2d_{n-1} \end{pmatrix},$$

را حل کنیم. به همین ترتیب، می توان سایر رابطه ها را ساده کرد.

۱۲.۳ یکی از کاربردهای مهم توابع اسپلاین

یکی از کاربردهای مهم توابع اسپلاین، حل تقریبی معادلات دیفرانسیل می‌باشد. در اینجا، مثالی از این کاربرد را ارائه می‌کنیم. مسئله مقدار مرزی

$$\begin{cases} x^2 y'' + xy' - y = 0, \\ y(1) = 1, \quad y(2) = 0.5. \end{cases}$$

را در نظر بگیرید. بازه $[1, 2]$ را به n زیربازه مساوی با نقاط گرهی $x_j = 1 + jh$, $j = 0, 1, \dots, n$ تقسیم می‌کنیم که در آن $h = 1/n$. معادله دیفرانسیل را در هر کدام از نقاط گرهی x_j , $j = 0, 1, \dots, n$ می‌نویسیم

$$y_j'' = -\frac{1}{x_j} y_j' + \frac{1}{x_j^2} y_j. \quad (51.3)$$

با استفاده از رابطه‌های (۴۰.۳) و (۴۱.۳)، داریم

$$S_{\Delta}'(Y; x_j^-) = \frac{y_j - y_{j-1}}{h} + \frac{h}{3} M_j + \frac{h}{6} M_{j-1}, \quad (52.3)$$

$$S_{\Delta}'(Y; x_j^+) = \frac{y_{j+1} - y_j}{h} - \frac{h}{3} M_j - \frac{h}{6} M_{j+1}. \quad (53.3)$$

اکنون، با جای‌گذاری $y_j'' = M_j$ و رابطه‌های (۵۲.۳) و (۵۳.۳) در (۵۱.۳)، خواهیم داشت

$$\begin{cases} M_j = -\frac{1}{x_j} \left(\frac{y_{j+1} - y_j}{h} - \frac{h}{3} M_j - \frac{h}{6} M_{j+1} \right) + \frac{1}{x_j^2} y_j, & j = 0, 1, 2, \dots, n-1, \\ M_j = -\frac{1}{x_j} \left(\frac{y_j - y_{j-1}}{h} + \frac{h}{3} M_j + \frac{h}{6} M_{j-1} \right) + \frac{1}{x_j^2} y_j, & j = 1, 2, \dots, n, \end{cases}$$

با توجه به اینکه $y_0 = 1$ و $y_n = 0.5$ ، رابطه‌های اخیر یک دستگاه معادلات خطی را می‌سازد که شامل $2n$ معادله و $2n$ مجهول است. با حل این دستگاه، نه تنها مقادیر تقریبی y_j ، بلکه مقادیر M_j ها نیز به دست می‌آیند. برای مثال، بازه $[1, 2]$ را به دو زیربازه تقسیم می‌کنیم. در این صورت، دستگاه معادلات خطی حاصل به صورت

$$\begin{cases} 10M_0 - M_1 + 24y_1 = 36, \\ 16M_1 - M_2 - 32y_1 = -12, \\ M_0 + 20M_1 + 16y_1 = 24, \\ M_1 + 26M_2 - 24y_1 = -9, \end{cases}$$

خواهد بود. با حل این دستگاه، خواهیم داشت $y_1 = 0.6560$. مقدار دقیق y در 1.5 ، برابر است با

$$y(1.5) = \frac{2}{3} \approx 0.6667.$$

همان‌گونه که مشاهده می‌شود، تقریب به‌دست آمده مناسب است.

۱۳.۳ همگرایی تابع اسپلاین مکعبی درونیاب

به‌ازای هر $x \in [a, b]$ ، فرض کنید

$$E(x) = f(x) - S_{\Delta}(Y; x),$$

$$E'(x) = f'(x) - S'_{\Delta}(Y; x).$$

همچنین، فرض کنید $h = \max_j (x_{j+1} - x_j)$ و $\bar{x} \in [a, b]$ یک مقدار ثابت باشد. در این صورت، بازه‌ای مثل $[x_j, x_{j+1}]$ وجود دارد به‌طوری که $\bar{x} \in [x_j, x_{j+1}]$. با توجه به اینکه

$$E(x_j) = E(x_{j+1}) = 0,$$

طبق قضیه رُل به‌ازای یک $c \in (x_j, x_{j+1})$ داریم $E'(c) = 0$. در نتیجه

$$\int_c^{\bar{x}} E''(t) dt = E'(\bar{x}) - E'(c) = E'(\bar{x}).$$

بدون از دست دادن کلیت، فرض می‌کنیم $c \leq \bar{x}$. بنابراین

$$\begin{aligned} |E'(\bar{x})|^2 &= \left| \int_c^{\bar{x}} E''(t) dt \right|^2 \leq \int_c^{\bar{x}} |E''(t)|^2 dt \int_c^{\bar{x}} 1^2 dt \\ &= (\bar{x} - c) \int_c^{\bar{x}} |E''(t)|^2 dt \leq h \int_c^{\bar{x}} |E''(t)|^2 dt \\ &\leq h \int_a^b |E''(t)|^2 dt = h \|f - S_{\Delta}(Y; \cdot)\|^2 \\ &= h \left(\|f\|^2 - \|S_{\Delta}(Y; \cdot)\|^2 \right) \leq h \|f\|^2. \end{aligned}$$

از این رو، داریم

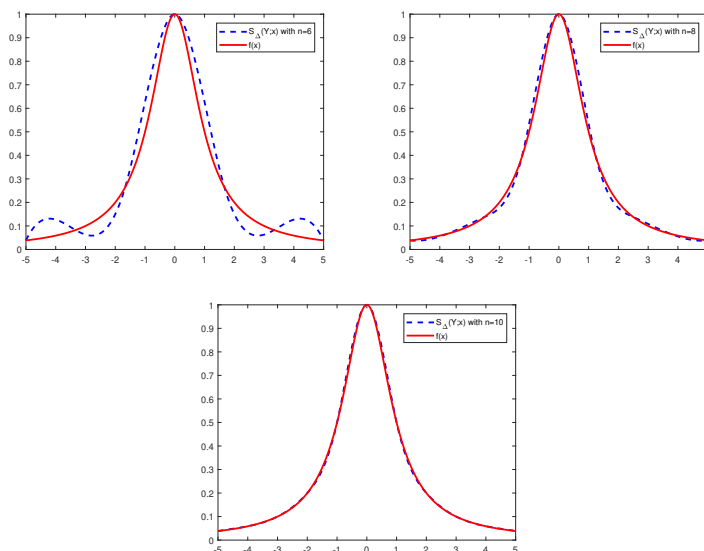
$$|E'(\bar{x})| \leq \sqrt{h} \|f\|. \quad (54.3)$$

از طرفی

$$\int_{x_j}^{\bar{x}} E'(t) dt = E(\bar{x}) - E(x_j) = E(\bar{x}).$$

در نتیجه، با استفاده از (54.3) خواهیم داشت

$$\begin{aligned} |E(\bar{x})| &= \left| \int_{x_j}^{\bar{x}} E'(t) dt \right| \leq \int_{x_j}^{\bar{x}} |E'(t)| dt \leq (\bar{x} - x_j) \max_{t \in [x_j, \bar{x}]} |E'(t)| \\ &\leq h \max_{t \in [a, b]} |E'(t)| \leq h \sqrt{h} \|f\|, \end{aligned}$$



شکل ۶.۳: درونیابی تابع رونگه با استفاده از توابع اسپلاین مکعبی.

یا

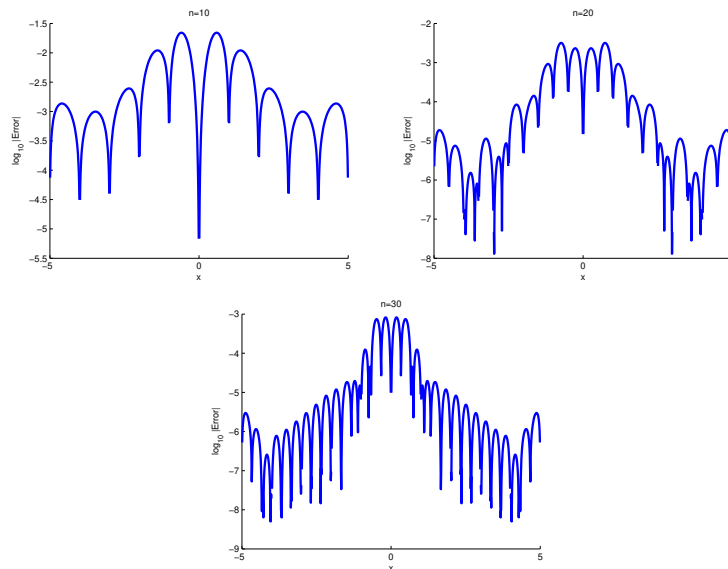
$$|E(\bar{x})| \leq h^{\frac{1}{3}} \|f\| = h^{\frac{1}{3}} \left(\int_a^b |f''(t)|^2 dt \right)^{\frac{1}{2}}.$$

این رابطه نشان می‌دهد که اگر $f'' \in L^2[a, b]$ و $h \rightarrow 0$ ، آنگاه $E(\bar{x}) \rightarrow 0$. این نتیجه، مهم‌ترین تفاوت درونیابی با استفاده از توابع اسپلاین و چندجمله‌ای‌ها است. برای مثال، تابع رونگه را به‌ازای n های مختلف با استفاده از تابع اسپلاین مکعبی درونیابی کرده‌ایم و در شکل ۶.۳ نمایش داده شده است. همچنین، در شکل ۷.۳، خطای حاصل از درونیابی تابع رونگه با استفاده از توابع اسپلاین مکعبی به‌ازای $n = 10, 20, 30$ نمایش داده شده است.

تمرین

(۱.۳) در رابطه (۸.۳)، چه تفاوتی بین دو عبارت مختلفی که برای T_{ik} ارائه شده است، وجود دارد؟

(۲.۳) مثال (۲.۳) را با استفاده از رابطه (۱۲.۳) حل کنید.



شکل ۷.۳: نمایش خطا برای درونیابی تابع رونگه با استفاده از توابع اسپلاین مکعبی.

(۳.۳) مثال (۲.۳) را با استفاده از الگوریتم ایتکن حل کنید.

(۴.۳) تابع $f(x) = x^{n+1}$ را در n نقطه متمایز $x_i, i = 0, 1, \dots, n$ درونیابی می‌کنیم. نشان

دهید درجه چندجمله‌ای درونیاب کمتر از n است اگر و تنها اگر $\sum_{i=0}^n x_i = 0$.

(۵.۳) فرض کنید $i = 0, 1, \dots, n, L_i$ چندجمله‌ای‌های لاگرانژ و نقاط گرهی $x_i, i = 0, 1, \dots, n$ متمایز باشند. نشان دهید اگر $c_i = L_i(0)$ ، آنگاه

$$\sum_{i=0}^n c_i x_i^j = \begin{cases} 1, & j = 0, \\ 0, & j = 1, 2, \dots, n, \\ (-1)^n x_0 x_1 \cdots x_n, & j = n+1. \end{cases}$$

(۶.۳) یک چندجمله‌ای مثل P از درجه حداکثر ۳ بیابید به طوری که

$$P(0) = 1, \quad P'(0) = 2, \quad P(1) = 3, \quad P'(1) = 0.$$

فرض کنید تابع f دارای مشتق چهارم پیوسته در بازه $[0, 1]$ است و $M = \max_{x \in [0, 1]} |f^{(4)}(x)|$ نشان

دهید

$$|f(x) - P(x)| \leq Mc,$$

که در آن c یک مقدار ثابت است. مقدار c را بیابید.

(۷.۳) یک چندجمله‌ای مثل P از درجه حداکثر ۳ بیابید به طوری که

$$P(0) = 0, \quad P'(0) = 5, \quad P''(0) = 0,$$

$$P(1) = 6.$$

اگر f یک تابع دلخواه باشد که در شرایط بالا صدق کند و $E(x) = f(x) - P(x)$ ، آنگاه یک کران بالا برای $|E(x)|$ در بازه $[0, 1]$ به دست آورید. در صورتی که تابع f یک چندجمله‌ای تکین از درجه ۴ باشد، مقدار دقیق خطا در $x = \frac{1}{4}$ را محاسبه کنید.

(۸.۳) سه نقطه x_0, x_1, x_2 و مقادیر y_0, y_1, y_2 در \mathbb{R} داده شده است. نشان دهید که یک و تنها یک چندجمله‌ای مثل P از درجه حداکثر ۳ وجود دارد به طوری که

$$P(x_0) = y_0,$$

$$P(x_1) = y_1, \quad P'(x_1) = y'_1,$$

$$P(x_2) = y_2.$$

اگر f تابعی باشد که

$$f(x_i) = y_i, \quad i = 1, 2, 3,$$

$$f'(x_1) = y'_1,$$

آنگاه عبارت خطای $E(x) = f(x) - P(x)$ را به دست آورید.

(۹.۳) با استفاده از جدول تفاضلات معکوس، یک چندجمله‌ای گویا بیابید که تابع f را در نقاط داده شده در جدول زیر درونیابی کند.

x_i	0	1	2	4
$f(x_i)$	1	$\frac{3}{4}$	-1	-27

(۱۰.۳) با استفاده از جدول تفاضلات متقابل (کسرهای مسلسل تیل)، یک چندجمله‌ای گویا بیابید که تابع f را در نقاط داده شده در جدول زیر درونیابی کند.

x_i	0	1	2	3
$f(x_i)$	0	1	5	21

(۱۱.۳) چهار مقدار از تابع $f(t) = \sin t + \cos^3 t$ در جدول زیر داده شده است. یک چندجمله‌ای مثلثاتی از درجه دو مثل q_2 بیابید به طوری که $q_2(t_j) = y_j$ که $j = 0, 1, \dots, 3$.

j	0	1	2	3
t_j	0	$\frac{\pi}{4}$	π	$\frac{3\pi}{4}$
$y_j = f(t_j)$	1	1	-1	-1

۱۲.۳) نشان دهید که تابع زیر یک اسپلاین مکعبی طبیعی است.

$$S(x) = \begin{cases} 2(x+1) + (x+1)^3, & x \in [-1, 0], \\ 3 + 5x + 3x^2, & x \in [0, 1], \\ 11 + 11(x-1) + 3(x-1)^2 - (x-1)^3, & x \in [1, 2]. \end{cases}$$

۱۳.۳) با استفاده از مقادیر تابع f که در جدول زیر داده شده است، معادلات اسپلاین مکعبی متناوب بر بازه $[-1, 1]$ و متناظر با گره‌های $-1, 0, 1$ را به دست آورید.

x_i	-1	0	1
$f(x_i)$	1	2	1

۱۴.۳) فرض کنید نقاط $x_0 < x_1 < \dots < x_n$ متساوی‌فاصله و به فاصله h از یکدیگر باشند و $Y = \{y_0, y_1, \dots, y_n\}$. همچنین، فرض کنید $S_{\Delta}(Y; x)$ یک اسپلاین درجه دوم با شرط

$$S_{\Delta}(Y; x_i) = y_i, \quad i = 0, 1, \dots, n,$$

باشد. نخست، نشان دهید که با توجه به تعریف یک اسپلاین درجه دوم، برای تعیین یک اسپلاین درجه دوم یکتا نیاز به یک شرط اضافی است. دوم، این شرط اضافی را به صورت $S'_{\Delta}(Y; x_0) = S'_{\Delta}(Y; x_1)$ بگیرید و اسپلاین درجه دوم متناظر را به دست آورید. آیا با شرط اضافی داده شده، این اسپلاین به طور یکتا تعیین می‌شود؟ چرا؟ راهنمایی: فرض کنید $m_j = S'_{\Delta}(Y; x_j)$ ، $j = 0, 1, \dots, n$ را در هر زیربازه برحسب m_j ها نوشته و سپس یک دستگاه معادلات خطی برای تعیین m_j ها بنویسید.

۱۵.۳) تابع $S(x)$ را به صورت

$$S(x) = \begin{cases} 3 + x - 9x^2, & x \in [0, 1], \\ a + b(x-1) + c(x-1)^2 + d(x-1)^3, & x \in [1, 2], \end{cases}$$

در نظر بگیرید. مقادیر a, b, c را طوری تعیین کنید که $S(x)$ یک اسپلاین مکعبی طبیعی با گره‌های $0, 1, 2$ باشد.

۱۶.۳) فرض کنید f یک تابع به اندازه کافی هموار روی بازه $[0, 2]$ باشد و

$$\Delta = \{x_0 = 0 < x_1 = 1 < x_2 = 2\}, \quad Y = \{y_0, y_1, y_2\},$$

که در آن $y_i = f(x_i)$ ، $i = 0, 1, 2$. فرض کنید چند جمله‌ای درونیاب اسپلاین مکعبی

$$S_{\Delta}(Y; x) = \begin{cases} 3 + x - 9x^3, & x \in [0, 1), \\ a + b(x-1) + c(x-1)^2 + d(x-1)^3, & x \in [1, 2], \end{cases}$$

به گونه‌ای باشد که $S_{\Delta}(Y; x_i) = f(x_i)$ ، $i = 0, 1, 2$. مقادیر a, b, c, d را طوری بیابید که

$$\int_0^2 |S_{\Delta}''(Y; x)|^2 dx \leq \int_0^2 |f''(x)|^2 dx.$$

۱۷.۳) با استفاده از قضیه ۱۹.۳ نشان دهید اگر $\Delta' \subset [a, b]$ شامل نقاط گرهی Δ باشد، یعنی $\Delta \subset \Delta'$ ، آنگاه برای هر یک از حالت‌های (الف)، (ب) و (ج)، داریم

$$\|f\| \geq \|S_{\Delta'(Y; \cdot)}\| \geq \|S_{\Delta(Y; \cdot)}\|.$$

۱۸.۳) مقادیر a, b, c, d را طوری تعیین کنید که تابع

$$S(x) = \begin{cases} 6 + 3x - 8x^3, & x \in [0, 1), \\ a + b(x-1) + c(x-1)^2 + d(x-1)^3, & x \in [1, 2], \end{cases}$$

یک اسپلاین مکعبی طبیعی باشد. اکنون، فرض کنید S یک اسپلاین درونیاب تابع f و همچنین، $T(x)$ یک اسپلاین مکعبی طبیعی و درونیاب تابع f روی $[0, 0.5]$ ، $[0.5, 1]$ و $[1, 2]$ است. نشان

دهید

$$\int_0^2 |S'''(x)|^2 dx \leq \int_0^2 |T'''(x)|^2 dx.$$

راهنمایی: از تمرین ۱۷.۳ استفاده کنید.

۱۹.۳) مجموعه نقاط گرهی $\Delta = \{x_0 < x_1 < \dots < x_n\}$ و مقادیر $Y := \{y_0, y_1, \dots, y_n\}$ داده شده است. بدون استفاده از قضیه ۱۹.۳، ثابت کنید که تابع اسپلاین $S_{\Delta}(Y; \cdot)$ با شرط $S_{\Delta}''(Y; x_0) = S_{\Delta}''(Y; x_n) = 0$ یکتاست.

۲۰.۳) فرض کنید $S_{\Delta}(x)$ یک تابع اسپلاین درونیاب تابع $f \in \mathcal{K}_p^4(a, b)$ در نقاط گرهی زیر است

$$\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}.$$

نشان دهید اگر یکی از شرایط

(الف) به‌ازای $x = a, b$ ، $f'(x) = S_{\Delta}'(x)$ ؛

(ب) به‌ازای $x = a, b$ ، $f''(x) = S_{\Delta}''(x)$ ؛

(ج) S_{Δ} متناوب باشد و $f \in \mathcal{K}_p^4(a, b)$ ؛

برقرار باشد، آنگاه خواهیم داشت

$$\|f - S_{\Delta}\|^2 = \int_a^b (f(x) - S_{\Delta}(x))^2 f^{(4)}(x) dx.$$

فصل ۴

انتگرال گیری عددی

۱.۴ مقدمه

اگر F تابع اولیه f باشد، یعنی $F' = f$ ، آنگاه داریم

$$\int_a^b f(x)dx = F(b) - F(a).$$

ولی، اغلب محاسبه تابع اولیه یک تابع دلخواه کاری دشوار یا حتی گاهی ناممکن است. از این رو، برای محاسبه مقدار این انتگرال از روش های عددی استفاده می شود. در این فصل، برخی از این روش ها را مطالعه می کنیم.

۲.۴ فرمول های نیوتن - کاتس

این فرمول ها با جای گذاری تابع f با یک چندجمله ای درونیاب آن به دست می آید. در واقع، اگر $P(x)$ یک چندجمله ای درونیاب برای تابع f در بازه $[a, b]$ باشد، آنگاه $\int_a^b P(x)dx$ به عنوان یک تقریب برای $\int_a^b f(x)dx$ در نظر گرفته می شود. فرض کنید $h = (b - a)/n$ و

$$x_i = a + ih, \quad i = 0, 1, \dots, n.$$

همچنین، فرض کنید P_n چندجمله‌ای درونیاب تابع f در نقاط $(x_i, f(x_i))$ ، $i = 0, 1, \dots, n$ است، یعنی

$$P_n(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

همان‌گونه که در فصل درونیابی دیدیم، چندجمله‌ای درونیاب لاگرانژ به صورت زیر است

$$P_n(x) = \sum_{i=0}^n f(x_i) L_i(x),$$

که در آن

$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{x - x_k}{x_i - x_k}.$$

برای هر $x \in [a, b]$ ، قرار می‌دهیم $t = (x - a)/h$. در این صورت، داریم $x = a + th$ و

$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{t - k}{i - k} =: \varphi_i(t).$$

اکنون، با توجه به اینکه $dx = h dt$ ، داریم

$$\begin{aligned} \int_a^b P_n(x) dx &= \sum_{i=0}^n f_i \int_a^b L_i(x) dx \\ &= h \sum_{i=0}^n f_i \int_0^n \varphi_i(t) dt \\ &= h \sum_{i=0}^n f_i \alpha_i. \end{aligned}$$

توجه کنید که در اینجا مقدار

$$\alpha_i = \int_0^n \varphi_i(t) dt,$$

تنها به n وابسته است و مستقل از f و کران‌های بازه، یعنی a و b است. بنابراین، داریم

$$\int_a^b f(x) dx \approx h \sum_{i=0}^n f_i \alpha_i.$$

این فرمول به فرمول نیوتن-کاتس معروف است. با فرض $w_i = h \alpha_i$ ، این رابطه به صورت

$$\int_a^b f(x) dx \approx \sum_{i=0}^n w_i f(x_i), \quad (۱.۴)$$

خواهد بود. در این فرمول، w_i ها وزن‌های فرمول نیوتن-کاتس نام دارند.

قضیه ۱.۴ به ازای $n + 1$ گره متمایز انتگرال گیری x_0, x_1, \dots, x_n ، وزن های w_i در فرمول انتگرال گیری (۱.۴) به طور یکتا تعیین می شوند اگر و تنها اگر این فرمول برای هر چند جمله ای از درجه حداکثر n دقیق باشد.

برهان: فرض کنید $f \in \Pi_n$. در این صورت، چند جمله ای درونیاب f ، یعنی P ، در نقاط x_0, x_1, \dots, x_n ، برابر با f خواهد بود. به عبارت دیگر، $f \equiv P$. در نتیجه

$$\int_a^b f(x) dx = \sum_{i=0}^n w_i f(x_i), \quad f \in \Pi_n.$$

برعکس، فرض کنید که فرمول انتگرال گیری (۱.۴) به ازای هر $P \in \Pi_n$ دقیق باشد. در نتیجه، این فرمول برای چند جمله ای های $1, x, \dots, x^n$ دقیق است که با جای گذاری آنها در رابطه (۱.۴)، دستگاه معادلات خطی

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^n & x_1^n & \dots & x_n^n \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \int_a^b 1 dx \\ \int_a^b x dx \\ \vdots \\ \int_a^b x^n dx \end{pmatrix},$$

به دست می آید. این دستگاه جواب یکتا دارد؛ زیرا x_i ها متمایز هستند و ماتریس ضرایب آن یک ماتریس واندرموند است (رابطه (۳.۱) را ببینید) و به این ترتیب اثبات قضیه کامل می شود. □

اکنون، به محاسبه چند فرمول انتگرال گیری در حالت خاص می پردازیم. اگر $n = 1$ ، آنگاه

$$\alpha_0 = \int_0^1 \frac{t-1}{0-1} dt = \frac{1}{2},$$

$$\alpha_1 = \int_0^1 \frac{t-0}{1-0} dt = \frac{1}{2},$$

و در نتیجه، خواهیم داشت

$$\int_a^b f(x) dx \approx \int_a^b P_1(x) dx = h \left(\frac{1}{2} f_0 + \frac{1}{2} f_1 \right) = \frac{h}{2} (f_0 + f_1),$$

که به قاعده ذوزنقه ای معروف است.

اگر $n = 2$ ، آنگاه

$$\alpha_0 = \int_0^2 \frac{t-1}{0-1} \frac{t-2}{0-2} dt = \frac{1}{6},$$

$$\alpha_1 = \int_0^2 \frac{t - 0}{1 - 0} \frac{t - 2}{1 - 2} dt = \frac{4}{3},$$

$$\alpha_2 = \int_0^2 \frac{t - 0}{2 - 0} \frac{t - 1}{2 - 1} dt = \frac{1}{3},$$

و در نتیجه، خواهیم داشت

$$\int_a^b f(x) dx \approx \int_a^b P_2(x) dx = h \left(\frac{1}{3} f_0 + \frac{4}{3} f_1 + \frac{1}{3} f_2 \right) = \frac{h}{3} (f_0 + 4f_1 + f_2),$$

که به قاعدهٔ سیمپسون معروف است.

یادداشت ۱.۴ حاصل جمع α_i ها برابر با تعداد گره‌ها در فرمول نیوتن-کاتس است؛ زیرا

$$\sum_{i=0}^n \alpha_i = \sum_{i=0}^n \int_0^n \varphi_i(t) dt = \int_0^n \left(\sum_{i=0}^n \varphi_i(t) \right) dt = \int_0^n 1 dt = n.$$

اکنون، فرض کنید $P(x)$ چند جمله‌ای درونیاب هرmit تابع f در نقاط $x_0 = a$ و $x_1 = b$ باشد، یعنی

$$P(a) = f(a), \quad P'(a) = f'(a),$$

$$P(b) = f(b), \quad P'(b) = f'(b).$$

برای سادگی، فرض می‌کنیم $a = 0$ و $b = 1$. با توجه به مطالب ارائه‌شده در فصل گذشته دربارهٔ درونیابی هرmit، داریم $L_0(x) = -x + 1$ و $L_1(x) = x$

$$H_0^{\circ}(x) = (1 - 2(x - x_0)L_0'(x_0))[L_0(x)]^2 = (1 + 2x)(x - 1)^2,$$

$$H_1^{\circ}(x) = (1 - 2(x - x_1)L_1'(x_1))[L_1(x)]^2 = (1 - 2(x - 1))x^2 = (3 - 2x)x^2,$$

$$H_0^{\backslash}(x) = (x - x_0)[L_0(x)]^2 = (x - 0)(-x + 1)^2 = x(x - 1)^2,$$

$$H_1^{\backslash}(x) = (x - x_1)[L_1(x)]^2 = (x - 1)x^2.$$

در نتیجه

$$P(x) = f(0)H_0^{\circ}(x) + f(1)H_1^{\circ}(x) + f'(0)H_0^{\backslash}(x) + f'(1)H_1^{\backslash}(x),$$

و بنابراین

$$\int_0^1 f(x) dx \approx \int_0^1 P(x) dx = f(0) \int_0^1 H_0^{\circ}(x) dx + f(1) \int_0^1 H_1^{\circ}(x) dx$$

$$\begin{aligned}
 & +f'(\circ) \int_{\circ}^1 H_{\circ}'(x)dx + f'(1) \int_{\circ}^1 H_1'(x)dx \\
 & = \frac{1}{4}(f(\circ) + f(1)) + \frac{1}{12}(f'(\circ) - f'(1)).
 \end{aligned}$$

اکنون، رابطه انتگرال گیری به دست آمده را برای بازه $[a, b]$ به کار می‌بریم. برای این کار، قرار می‌دهیم $t = (x - a)/(b - a)$. در این صورت، خواهیم داشت

$$dt = \frac{1}{b-a}dx, \quad x = (b-a)t + a.$$

در نتیجه

$$\begin{aligned}
 \int_a^b f(x)dx &= \int_{\circ}^1 f((b-a)t+a)(b-a)dt \\
 &= (b-a) \int_{\circ}^1 g(t)dt \quad (g(t) = f((b-a)t+a)) \\
 &\approx \frac{b-a}{4}(g(\circ) + g(1)) + \frac{b-a}{12}(g'(\circ) - g'(1)) \\
 &= \frac{b-a}{4}(f(a) + f(b)) + \frac{(b-a)^2}{12}(f'(a) - f'(b)).
 \end{aligned}$$

بنابراین، داریم

$$\int_a^b f(x)dx \approx \frac{b-a}{4}(f(a) + f(b)) + \frac{(b-a)^2}{12}(f'(a) - f'(b)),$$

که به قاعده دوزنقه‌ای اصلاح شده معروف است.

اینک، به بحث درباره خطای انتگرال گیری می‌پردازیم. در حالت کلی، اگر P چندجمله‌ای درونیاب تابع f و $E(x) = f(x) - P(x)$ خطای درونیابی باشد، آنگاه خطای انتگرال گیری به صورت

$$E_I = \int_a^b E(x)dx,$$

خواهد بود. با توجه به این نکته، عبارت خطای سه روش ارائه شده را محاسبه می‌کنیم.

خطا در قاعده دوزنقه‌ای: فرض کنید $f \in C^2[a, b]$. همان‌گونه که دیدیم، در این حالت تابع f را در نقاط $x_0 = a$ و $x_1 = b$ درونیابی کردیم. بنابراین، داریم

$$E(x) = \frac{f''(\xi_x)}{2!}(x-a)(x-b),$$

که در آن عددی در بازه $[a, b]$ است. در نتیجه، با توجه به اینکه در بازه $[a, b]$ داریم $(x-a)(x-b) \leq 0$ ، از قضیه ۳.۱ نتیجه می‌گیریم که

$$E_I = \int_a^b \frac{f''(\xi_x)}{2!}(x-a)(x-b)dx$$

$$\begin{aligned} &= \frac{f''(\alpha)}{2!} \int_a^b (x-a)(x-b)dx \\ &= -\frac{(b-a)^2}{12} f''(\alpha), \end{aligned}$$

که در آن $\alpha \in [a, b]$.

خطا در روش سیمپسون: فرض کنید $f \in C^4[a, b]$ همانند روش ذوزنقه‌ای، داریم

$$E_I = \int_a^b \frac{f^{(4)}(\xi_x)}{4!} (x-x_0)(x-x_1)(x-x_2)dx,$$

که در آن

$$x_0 = a, \quad x_1 = \frac{a+b}{2}, \quad x_2 = b.$$

با توجه به اینکه عبارت $(x-x_0)(x-x_1)(x-x_2)$ در بازه $[a, b]$ تغییر علامت می‌دهد، نمی‌توانیم از قضیه ۳.۱ برای ساده‌سازی آن استفاده کنیم. برای رفع این مشکل، به شرح زیر عمل می‌کنیم. فرض کنید Q چندجمله‌ای درونیاب تابع f در نقاط $a, x_1 = (a+b)/2, x_2 = b$ و $x_3 = (a+b)/2$ است. با توجه به شیوه ساخت چندجمله‌ای درونیاب نیوتن، داریم

$$Q(x) = P(x) + f[x_0, x_1, x_2, x_3](x-x_0)(x-x_1)(x-x_2),$$

به سادگی می‌توان دید که (چرا؟)

$$E_I = \int_a^b (f(x) - P(x))dx = \int_a^b (f(x) - Q(x))dx.$$

بنابراین

$$\begin{aligned} E_I &= \int_a^b (f(x) - Q(x))dx \\ &= \int_a^b f[x_0, x_1, x_2, x_3, x](x-x_0)(x-x_1)^2(x-x_2)dx \\ &= \int_a^b \frac{f^{(4)}(\xi_x)}{4!} (x-x_0)(x-x_1)^2(x-x_2)dx. \end{aligned}$$

اکنون، از اینکه در بازه $[a, b]$ داریم $(x-x_0)(x-x_1)^2(x-x_2) \leq 0$ نتیجه می‌گیریم

$$\begin{aligned} E_I &= \frac{f^{(4)}(\alpha)}{4!} \int_a^b (x-x_0)(x-x_1)^2(x-x_2)dx \\ &= -\frac{(b-a)^5}{2880} f^{(4)}(\alpha), \end{aligned}$$

که در آن $\alpha \in [a, b]$.

خطا در قاعدهٔ دوزنقه‌ای اصلاح‌شده: فرض کنید $f \in C^4[a, b]$. همان‌طور که در فصل درونیابی دیدیم، خطای درونیابی هرمیت به صورت

$$E(x) = \frac{f^{(4)}(\xi_x)}{4!} (x-a)^2 (x-b)^2,$$

است. بنابراین، با توجه به اینکه در بازهٔ $[a, b]$ داریم

$$(x-a)^2 (x-b)^2 \geq 0,$$

نتیجه می‌گیریم

$$\begin{aligned} E_I &= \int_a^b \frac{f^{(4)}(\xi_x)}{4!} (x-a)^2 (x-b)^2 dx \\ &= \frac{f^{(4)}(\alpha)}{4!} \int_a^b (x-a)^2 (x-b)^2 dx \\ &= \frac{(b-a)^5}{720} f^{(4)}(\alpha), \end{aligned}$$

که در آن $\alpha \in [a, b]$.

۳.۴ دستوره‌ای مرکب انتگرال‌گیری

در بخش قبل، ایدهٔ روش‌های انتگرال‌گیری عددی ارائه شد. دیدیم که فرمول‌های انتگرال‌گیری نیوتن-کاتس به وسیلهٔ انتگرال‌گیری از چندجمله‌ای درونیاب انتگرالده (تابع زیر انتگرال) در تعدادی از نقاط بازه به دست می‌آید. برای بالا بردن دقت انتگرال می‌توان تعداد نقاط انتگرال‌گیری را افزایش داد. به دو دلیل مهم این کار مناسب نیست. نخست، حجم محاسبات بسیار زیاد خواهد بود و دوم اینکه، با افزایش تعداد نقاط درونیابی لزوماً دقت چندجمله‌ای درونیاب بالا نخواهد رفت (مثال رونگه را به خاطر بیاورید). برای رفع این مشکل، از فرمول‌های مرکب انتگرال‌گیری استفاده می‌شود. ایدهٔ اصلی این است که بازهٔ انتگرال‌گیری را به چند زیربازه تقسیم کنیم و انتگرال تابع در کل بازه را به صورت مجموع انتگرال‌ها در زیربازه‌ها بنویسیم و برای هر زیربازه از فرمول انتگرال‌گیری استفاده کنیم. در ادامه، فرمول‌های انتگرال‌گیری مرکب برای سه روش انتگرال‌گیری معرفی شده در بخش قبل را به دست می‌آوریم.

فرمول مرکب ذوزنقه‌ای

بازه $[a, b]$ را به n قسمت مساوی با طول گام $h = (b - a)/n$ و نقاط گرهی $x_i = a + ih$ $i = 0, 1, \dots, n$ تقسیم می‌کنیم. داریم

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \\ &= \sum_{i=0}^{n-1} \left(\frac{x_{i+1} - x_i}{2} (f(x_i) + f(x_{i+1})) - \frac{(x_{i+1} - x_i)^3}{12} f''(\alpha_i) \right) \\ &= \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right) - \sum_{i=0}^{n-1} \frac{h^3}{12} f''(\alpha_i). \end{aligned}$$

بنابراین، فرمول ذوزنقه‌ای مرکب به صورت

$$\int_a^b f(x)dx \approx \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right),$$

است و خطای آن برابر با

$$E_{CI} = - \sum_{i=0}^{n-1} \frac{h^3}{12} f''(\alpha_i) = - \frac{h^3}{12} \sum_{i=0}^{n-1} f''(\alpha_i),$$

خواهد بود، که در آن $\alpha \in (a, b)$ از طرفی، داریم

$$\min_i f''(\alpha_i) \leq \frac{1}{n} \sum_{i=0}^{n-1} f''(\alpha_i) \leq \max_i f''(\alpha_i).$$

در نتیجه، با استفاده از قضیه مقدار میانی ۲.۱، یک α در $[a, b]$ وجود دارد به طوری که

$$\frac{1}{n} \sum_{i=0}^{n-1} f''(\alpha_i) = f''(\alpha).$$

بنابراین

$$E_{CI} = - \frac{h^3}{12} \sum_{i=0}^{n-1} f''(\alpha_i) = - \frac{h^3}{12} n f''(\alpha) = - \frac{b-a}{12} h^2 f''(\alpha).$$

فرمول مرکب ذوزنقه‌ای اصلاح شده

همانند قاعده ذوزنقه‌ای، می‌توان دید که فرمول مرکب ذوزنقه‌ای اصلاح شده به صورت

$$\int_a^b f(x)dx \approx \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right) + \frac{h^2}{12} (f'(a) - f'(b)),$$

و خطای آن به صورت زیر است

$$E_{CI} = \frac{b-a}{720} h^4 f^{(4)}(\alpha), \quad \alpha \in (a, b).$$

فرمول مرکب سیمپسون

برای به دست آوردن فرمول مرکب سیمپسون، کمی متفاوت با دو روش پیشین عمل می‌کنیم. بازه $[a, b]$ را به $2n$ زیربازه مساوی با طول گام $h = (b-a)/(2n)$ و نقاط گرهی x_i ، $i = 0, 1, \dots, 2n$ تقسیم می‌کنیم. اکنون، داریم

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{n-1} \int_{x_{2i}}^{x_{2i+2}} f(x) dx \\ &= \sum_{i=0}^{n-1} \left(\frac{x_{2i+2} - x_{2i}}{6} (f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})) \right. \\ &\quad \left. - \frac{(x_{2i+2} - x_{2i})^5}{720} f^{(4)}(\alpha_i) \right) \\ &= \frac{h}{3} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_{2i}) + 4 \sum_{i=0}^{n-1} f(x_{2i+1}) + f(b) \right) \\ &\quad - \frac{(2h)^5}{720} \sum_{i=0}^{n-1} f^{(4)}(\alpha_i). \end{aligned}$$

در این صورت، فرمول انتگرال گیری سیمپسون مرکب به صورت

$$\int_a^b f(x) dx \approx \frac{h}{3} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_{2i}) + 4 \sum_{i=0}^{n-1} f(x_{2i+1}) + f(b) \right),$$

و خطای مربوطه به صورت

$$E_{CI} = -\frac{(2h)^5}{720} \sum_{i=0}^{n-1} f^{(4)}(\alpha_i) = -\frac{b-a}{180} h^4 f^{(4)}(\alpha), \quad (2.4)$$

است، که در آن $\alpha \in (a, b)$.

مثال ۱.۴ برای محاسبه یک مقدار تقریبی از انتگرال $I = \int_0^1 e^{-x} \sin x dx$ ، از روش سیمپسون با $n = 4$ استفاده می‌کنیم. در این صورت، داریم $a = 0$ ، $b = 1$ و

$$h = \frac{b-a}{2n} = \frac{1-0}{8} = 0.125.$$

مقادیر تابع $f(x) = e^{-x} \sin x$ در نقاط گرهی $x_k = kh = 0.125k$ ، $k = 0, 1, \dots, 8$ به صورت زیر است

$$\begin{aligned} f(x_0) &= f(0) = 0, & f(x_1) &= f(0.125) = 0.110025066, \\ f(x_2) &= f(0.25) = 0.192678397, & f(x_3) &= f(0.375) = 0.251735182, \\ f(x_4) &= f(0.5) = 0.290786288, & f(x_5) &= f(0.625) = 0.313180002, \\ f(x_6) &= f(0.75) = 0.321983351, & f(x_7) &= f(0.875) = 0.319959735, \\ f(x_8) &= f(1) = 0.309559876. \end{aligned}$$

بنابراین، داریم

$$\begin{aligned} I &\approx \frac{h}{3}(f(x_0) + 4(f(x_1) + f(x_3) + f(x_5) + f(x_7)) \\ &\quad + 2(f(x_2) + f(x_4) + f(x_6)) + f(x_8)) \\ &= 0.24583566. \end{aligned}$$

مقدار دقیق I تا نه رقم بامعنا برابر با $I = 0.245837007$ است. چنان که مشاهده می شود با اینکه مقدار n عدد کوچکی است، مقدار تقریبی به دست آمده پنج رقم بامعنا درست دارد. اکنون، بینیم که اگر بخواهیم مقدار تقریبی I را با دقت 10^{-8} به دست آوریم، کوچک ترین مقدار n چه خواهد بود. با توجه به رابطه (۲.۴) می بینیم که

$$|E_{CI}| = \left| \frac{b-a}{180} h^4 f^{(4)}(\alpha) \right| \leq \frac{h^4}{180} \max_{0 \leq x \leq 1} |f^{(4)}(x)|. \quad (3.4)$$

داریم $g(x) := f^{(4)}(x) = -4e^{-x} \sin x$ با محاسبه مشتق تابع g ، داریم

$$g'(x) = 4e^{-x}(\sin x - \cos x).$$

تنها ریشه g' در بازه $[0, 1]$ ، عدد $x = \frac{\pi}{4}$ است و $g(\frac{\pi}{4}) = -1.29$ در نتیجه، داریم

$$M = \max_{0 \leq x \leq 1} |g(x)| = \max \left\{ |g(0)|, |g(\frac{\pi}{4})|, |g(1)| \right\} = \max \{ 0, 1.29, 1.24 \} = 1.29.$$

از این رو، با توجه به رابطه (۳.۴)، خواهیم داشت

$$|E_{CI}| \leq \frac{1.29}{180} h^4.$$

بنابراین، کافی است داشته باشیم

$$\frac{1.29}{180} h^4 \leq 10^{-8},$$

یا $h \leq 0.034$ اکنون، با توجه به

$$h = \frac{b-a}{2n} = \frac{1}{2n},$$

خواهیم داشت $n \geq 14.71$. در نتیجه، کافی است فرض کنیم $n = 15$.

۴.۴ نمایش خطای پئانو

قاعده انتگرال گیری

$$I(f) = \sum_{k=0}^{m_0} a_{k0} f(x_{k0}) + \sum_{k=0}^{m_1} a_{k1} f'(x_{k1}) + \cdots + \sum_{k=0}^{m_n} a_{kn} f^{(n)}(x_{kn}),$$

را در نظر بگیرید. قرار می دهیم

$$R(f) = I(f) - \int_a^b f(x) dx.$$

در حقیقت، خطای انتگرال گیری را با $R(f)$ نشان می دهیم. این تابع، خطی است، یعنی برای هر دو تابع f و g داریم

$$R(\alpha f + \beta g) = \alpha R(f) + \beta R(g), \quad \alpha, \beta \in \mathbb{R}.$$

اینک، قضیه بعدی را معرفی می کنیم که توسط پئانو ارائه شده است ([۹، ۱۲، ۱۳، ۱۵] را ببینید).

قضیه ۲.۴ فرض کنید برای تمام چند جمله ای های $P \in \Pi_n$ داشته باشیم $R(P) = 0$. در این صورت، برای هر $f \in C^{n+1}[a, b]$ داریم

$$R(f) = \int_a^b f^{(n+1)}(t) K(t) dt,$$

که در آن

$$K(t) = \frac{1}{n!} R_x[(x-t)_+^n], \quad (x-t)_+^n = \begin{cases} (x-t)^n, & x \geq t, \\ 0, & x < t, \end{cases}$$

و $R_x[(x-t)_+^n]$ به این معنا است که خطای انتگرال گیری تابع $(x-t)_+^n$ ، نسبت به متغیر x است. تابع $K(t)$ را هسته پئانو می گویند.

برهان: بسط تیلر تابع f را در نقطه $x = a$ در نظر بگیرید

$$f(x) = p_n(x) + r_n(x), \quad (4.4)$$

که در آن

$$p_n(x) = f(a) + f'(a)(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

و مانده تقریب به صورت

$$r_n(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x-t)^n dt = \frac{1}{n!} \int_a^b f^{(n+1)}(t)(x-t)_+^n dt,$$

است (قضیه ۶.۱ را ببینید). با توجه به اینکه p_n یک چندجمله‌ای از درجه حداکثر n است، داریم $R(p_n) = 0$. اکنون، اگر عملگر خطی R را روی رابطه (۴.۴) اثر دهیم، خواهیم داشت

$$R(f) = R(r_n) = \frac{1}{n!} R_x \left(\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right).$$

کافی است نشان دهیم که عملگر R_x و انتگرال را می‌توانیم با یکدیگر جابه‌جا کنیم. برای این کار، ابتدا نشان می‌دهیم که به‌ازای هر $1 \leq k \leq n$ ، داریم

$$\frac{d^k}{dx^k} \left[\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right] = \int_a^b f^{(n+1)}(t) \left[\frac{d^k}{dx^k} (x-t)_+^n \right] dt. \quad (5.4)$$

به‌ازای $k \leq n-1$ ، تابع $(x-t)_+^n$ ، $(n-1)$ بار پیوسته-مشتق‌پذیر است. از این رو، رابطه (۵.۴) به‌ازای هر $k \leq n-1$ برقرار است. بنابراین، کافی است رابطه را به‌ازای $k = n$ ثابت کنیم. به‌ازای $k = n-1$ داریم

$$\frac{d^{n-1}}{dx^{n-1}} \left[\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right] = \int_a^b f^{(n+1)}(t) \frac{d^{n-1}}{dx^{n-1}} [(x-t)_+^n] dt$$

بنابراین، خواهیم داشت

$$\begin{aligned} \frac{d^{n-1}}{dx^{n-1}} \left[\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right] &= n! \int_a^b f^{(n+1)}(t)(x-t)_+ dt \\ &= n! \int_a^x f^{(n+1)}(t)(x-t) dt. \end{aligned}$$

انتگرال آخر به‌عنوان یک تابع نسبت به متغیر x ، مشتق‌پذیر است؛ زیرا انتگرالده نسبت به هر دو متغیر x و t پیوسته است. در نتیجه

$$\frac{d^n}{dx^n} \left[\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right] = \frac{d}{dx} \left[\frac{d^{n-1}}{dx^{n-1}} \int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right]$$

$$\begin{aligned}
 &= \frac{d}{dx} \left[n! \int_a^x f^{(n+1)}(t)(x-t) dt \right] \\
 &= n! f^{(n+1)}(x)(x-x) + n! \int_a^x f^{(n+1)}(t) dt \\
 &= \int_a^b f^{(n+1)}(t) \left[\frac{d^n}{dx^n} (x-t)_+^n \right] dt.
 \end{aligned}$$

بنابراین، رابطه (۵.۴) به ازای $k = n$ برقرار است.
 رابطه (۵.۴) نشان می دهد که عملگرهای مشتق

$$\frac{d^k}{dx^k}, \quad k = 1, \dots, n,$$

با انتگرال جابه جا می شوند. بنابراین، با توجه به اینکه $I(f) = I_x(f)$ یک ترکیب خطی از عملگرهای مشتق است، با انتگرال جابه جا می شود.

سرانجام، پیوستگی انتگرالده $f^{(n+1)}(t)(x-t)_+^n$ نتیجه می دهد که انتگرال های زیر با هم جابه جا می شوند

$$\int_a^b \left[\int_a^b f^{(n+1)}(t)(x-t)_+^n dt \right] dx = \int_a^b f^{(n+1)}(t) \left[\int_a^b (x-t)_+^n dx \right] dt.$$

در نتیجه، عملگر R_x با انتگرال جابه جا می شود و نتیجه می گیریم که رابطه

$$R(f) = \frac{1}{n!} \int_a^b f^{(n+1)}(t) R_x((x-t)_+^n) dt$$

برقرار است. □

برای تعداد زیادی از فرمول های انتگرال گیری (از جمله قاعده سیمپسون) علامت هسته پتانو $K(t)$ در بازه $[a, b]$ ثابت است. در این حالت، با استفاده از قضیه مقدار میانگین برای انتگرال ۳.۱، به ازای یک $\xi \in (a, b)$ داریم

$$R(f) = f^{(n+1)}(\xi) \int_a^b K(t) dt.$$

این رابطه نشان می دهد که انتگرال $K(t)$ در بازه $[a, b]$ به تابع f وابسته نیست و می توان آن را با به کار بردن عملگر R روی یک تابع خاص مثل $f(x) = x^{n+1}$ به دست آورد. به عبارت دیگر، به ازای یک $\xi \in (a, b)$ داریم

$$R(f) = \frac{R(x^{n+1})}{(n+1)!} f^{(n+1)}(\xi).$$

به طور کلی، می‌توان دید زمانی که n فرد است، فرمول‌های نیوتن-کاتس برای هر چندجمله‌ای از درجه n و زمانی که n زوج است، برای هر چندجمله‌ای از درجه $n + 1$ دقیق هستند ([۱۵] را ببینید). بنابراین، به ازای یک $\xi \in (a, b)$ ، داریم

$$R_n(f) = \begin{cases} \frac{R(x^{n+1})}{(n+1)!} f^{(n+1)}(\xi), & \text{اگر } n \text{ فرد باشد,} \\ \frac{R(x^{n+2})}{(n+2)!} f^{(n+2)}(\xi), & \text{اگر } n \text{ زوج باشد.} \end{cases}$$

مثال ۲.۴ با استفاده از قضیه ۲.۴، عبارت خطا را برای روش سیمپسون به دست آورید.

حل: برای سادگی ابتدا فرض می‌کنیم $x_0 = -1$ ، $x_1 = 0$ و $x_2 = 1$. در این صورت، داریم

$$R(f) = \frac{1}{3}f(-1) + \frac{4}{3}f(0) + \frac{1}{3}f(1) - \int_{-1}^1 f(x)dx.$$

به سادگی می‌توان دید که $R(x^i) = 0$ ، $i = 0, 1, 2, 3$. همچنین، داریم

$$R(x^4) = \frac{1}{3} \left((-1)^4 + 4 \times (0)^4 + 1^4 \right) - \int_{-1}^1 x^4 dx = \frac{4}{15} \neq 0.$$

با توجه به اینکه R خطی است، نتیجه می‌گیریم که برای هر $P \in \Pi_3$ ، داریم $R(P) = 0$. پس، $n = 3$ و در نتیجه

$$\begin{aligned} K(t) &= \frac{1}{3!} R_x[(x-t)_+^3] \\ &= \frac{1}{6} \left(\frac{1}{3}(-1-t)_+^3 + \frac{4}{3}(0-t)_+^3 + \frac{1}{3}(1-t)_+^3 - \int_{-1}^1 (x-t)_+^3 dx \right). \end{aligned}$$

اکنون، برای هر $-1 \leq t \leq 1$ ، داریم

$$\begin{aligned} \int_{-1}^1 (x-t)_+^3 dx &= \int_t^1 (x-t)^3 dx = \frac{1}{4}(1-t)^4, \\ (-1-t)_+^3 &= 0, \\ (1-t)_+^3 &= (1-t)^3, \\ (0-t)_+^3 &= \begin{cases} 0, & t \geq 0, \\ -t^3, & t < 0. \end{cases} \end{aligned}$$

در نتیجه، اگر $-1 \leq t \leq 0$ ، آنگاه

$$K(t) = \frac{1}{6} \left(-\frac{4}{3}t^3 + \frac{1}{3}(1-t)^3 - \frac{1}{4}(1-t)^4 \right) = \frac{1}{\sqrt{24}}(1+t)^3(1-3t),$$

و اگر $۰ \leq t \leq ۱$ ، آنگاه

$$K(t) = \frac{1}{6} \left(\frac{1}{3}(1-t)^3 - \frac{(1-t)^4}{4} \right) = \frac{1}{72}(1-t)^3(1+3t).$$

در نتیجه، داریم

$$K(t) = \begin{cases} \frac{1}{72}(1+t)^3(1-3t), & -1 \leq t \leq 0, \\ \frac{1}{72}(1-t)^3(1+3t), & 0 \leq t \leq 1. \end{cases}$$

از اینکه $K(t)$ در بازه $[-1, 1]$ تغییر علامت نمی دهد، داریم

$$R(f) = \int_{-1}^1 f^{(4)}(t)K(t)dt = f^{(4)}(\alpha) \int_{-1}^1 K(t)dt,$$

که در آن $\alpha \in [-1, 1]$. با توجه به اینکه $\int_{-1}^1 K(t)dt$ مستقل از تابع f است، برای محاسبه آن می توان فرض کرد $f(x) = x^4$. در این صورت، خواهیم داشت

$$R(x^4) = 4! \int_{-1}^1 K(t)dt.$$

همان گونه که دیدیم، $R(x^4) = \frac{4}{15}$. بنابراین

$$\int_{-1}^1 K(t)dt = \frac{R(x^4)}{4!} = \frac{1}{90},$$

و از این رو، داریم

$$R(f) = \frac{1}{90} f^{(4)}(\alpha).$$

اکنون، با کمک تغییر متغیر

$$x = \frac{b+a}{2} + \frac{b-a}{2}t,$$

فرمول خطا را برای بازه $[a, b]$ محاسبه می کنیم. روشن است که $t \in [-1, 1]$ اگر و تنها اگر $x \in [a, b]$ داریم

$$dx = \frac{b-a}{2} dt.$$

بنابراین

$$\int_a^b f(x)dx = \frac{b-a}{2} \int_{-1}^1 f \left(\frac{b+a}{2} + \frac{b-a}{2}t \right) dt = \frac{b-a}{2} \int_{-1}^1 g(t)dt,$$

که در آن $g(t) = f\left(\frac{b+a}{2} + \frac{b-a}{2}t\right)$ از این رو، داریم

$$\begin{aligned} \int_a^b f(x)dx &= \frac{b-a}{2} \left(\frac{1}{3}g(-1) + \frac{4}{3}g(0) + \frac{1}{3}g(1) - \frac{1}{90} \frac{d^2}{dt^2}g(t) \Big|_{t=\alpha} \right) \\ &= \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) - \frac{1}{90} \frac{b-a}{2} \frac{d^2}{dt^2}g(t) \Big|_{t=\alpha}. \end{aligned}$$

در نتیجه، خطای فرمول مرکب سیمپسون به صورت زیر به دست می آید

$$\begin{aligned} E_{CI} &= -\frac{1}{90} \frac{b-a}{2} \frac{d^2}{dt^2}g(t) \Big|_{t=\alpha} \\ &= -\frac{1}{90} \frac{b-a}{2} \frac{d^2}{dt^2}f\left(\frac{b+a}{2} + \frac{b-a}{2}t\right) \Big|_{t=\alpha} \\ &= -\frac{b-a}{180} \left(\frac{b-a}{2}\right)^2 f^{(2)}\left(\frac{b+a}{2} + \frac{b-a}{2}\alpha\right) \\ &= -\frac{b-a}{180} h^2 f^{(2)}(\xi), \end{aligned}$$

که در آن $\xi = \frac{b+a}{2} + \frac{b-a}{2}\alpha$.

۵.۴ روش رامبرگ

فرض کنید تابع f ، تحلیلی (بینهایت بار مشتق پذیر) و α و β دو عدد صحیح نامنفی هستند. در این صورت، فرمول مجموع اولر-مکلوران [۵، ۶] برای محاسبه مجموع $\sum_{i=\alpha}^{\beta} f(i)$ به صورت

$$\sum_{i=\alpha}^{\beta} f(i) = \int_{\alpha}^{\beta} f(x)dx + \frac{f(\alpha) + f(\beta)}{2} + \sum_{j=1}^{\infty} \frac{B_{2j}}{(2j)!} \left(f^{(2j-1)}(\beta) - f^{(2j-1)}(\alpha) \right), \quad (6.4)$$

است که در آن، B_k ها اعداد برنولی هستند. این اعداد در بسط تابع $\frac{x}{e^x - 1}$ ، به صورت

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} x^k,$$

نمایان می شوند. چند مقدار ابتدایی B_k ها به صورت زیر هستند

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0.$$

می توان دید که به ازای هر $k \geq 1$ ، $B_{\nu_{k+1}} = 0$. اینک، برای محاسبه

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

بازه $[a, b]$ را به n قسمت مساوی تقسیم می کنیم و قرار می دهیم $h = (b-a)/n$. همچنین، تعریف می کنیم

$$x_i = a + ih, \quad i = 0, 1, 2, \dots, n.$$

تغییر متغیر

$$t = \frac{1}{h}(x - a),$$

را در نظر می گیریم. روشن است که $x \in [a, b]$ ، اگر و تنها اگر $t \in [0, n]$. اکنون، با توجه به اینکه $dx = h dt$ داریم

$$\int_a^b f(x) dx = h \int_0^n f(a + ht) dt = h \int_0^n g(t) dt, \quad (7.4)$$

که در آن $g(t) = f(a + ht)$ ، از طرفی، داریم

$$\frac{d^k}{dx^k} g(t) = h^k f^{(k)}(a + ht), \quad k = 0, 1, 2, \dots,$$

و در نتیجه

$$\left. \frac{d^k}{dx^k} g(t) \right|_{t=i} = h^k f^{(k)}(a + ih) = h^k f^{(k)}(x_i), \quad k = 0, 1, 2, \dots$$

بنابراین، از رابطه (7.4)، خواهیم داشت

$$\begin{aligned} I &= \int_a^b f(x) dx = h \int_0^n g(t) dt \\ &= h \left[\sum_{i=0}^n g(i) - \frac{g(0) + g(n)}{2} - \sum_{j=1}^{\infty} \frac{B_{\nu_j}}{(\nu_j)!} \left(g^{(\nu_{j-1})}(n) - g^{(\nu_{j-1})}(0) \right) \right] \\ &= h \left[\sum_{i=0}^n f(x_i) - \frac{f(a) + f(b)}{2} - \sum_{j=1}^{\infty} \frac{B_{\nu_j}}{(\nu_j)!} h^{\nu_{j-1}} \left(f^{(\nu_{j-1})}(b) - f^{(\nu_{j-1})}(a) \right) \right] \\ &= \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right) - \sum_{j=1}^{\infty} \frac{B_{\nu_j}}{(\nu_j)!} h^{\nu_j} \left(f^{(\nu_{j-1})}(b) - f^{(\nu_{j-1})}(a) \right) \\ &= T(h) - \sum_{j=1}^{\infty} \frac{B_{\nu_j}}{(\nu_j)!} h^{\nu_j} \left(f^{(\nu_{j-1})}(b) - f^{(\nu_{j-1})}(a) \right), \quad (8.4) \end{aligned}$$

که در آن

$$T(h) = \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right),$$

فرمول انتگرال ذوزنقه‌ای مرکب با طول گام h است. با فرض

$$c_{2j} = -\frac{B_{2j}}{(2j)!} \left(f^{(2j-1)}(b) - f^{(2j-1)}(a) \right), \quad j = 1, 2, \dots,$$

رابطه (۸.۴) را می‌توان به صورت

$$\int_a^b f(x) dx = T(h) + c_2 h^2 + c_4 h^4 + c_6 h^6 + \dots, \quad (9.4)$$

نوشت، که در آن c_j ها مستقل از h هستند. رابطه (۹.۴) نقش اساسی در روش رامبرگ بازی می‌کند.

در رابطه (۹.۴) به جای h ، مقدار $h/2$ را قرار می‌دهیم

$$\int_a^b f(x) dx = T\left(\frac{h}{2}\right) + c_2 \left(\frac{h}{2}\right)^2 + c_4 \left(\frac{h}{2}\right)^4 + c_6 \left(\frac{h}{2}\right)^6 + \dots,$$

طرفین رابطه (۹.۴) را از ۴ برابر طرفین رابطه اخیر کم می‌کنیم. در این صورت، خواهیم داشت

$$\int_a^b f(x) dx = \frac{4T\left(\frac{h}{2}\right) - T(h)}{4 - 1} + c'_4 h^4 + c'_6 h^6 + \dots, \quad (10.4)$$

که در آن

$$c'_{2j} = -\frac{1 - 2^{2(j-1)}}{2^{2(j-1)}}, \quad j = 2, 3, \dots$$

رابطه (۱۰.۴) نشان می‌دهد

$$\int_a^b f(x) dx \approx \frac{4T\left(\frac{h}{2}\right) - T(h)}{4 - 1}, \quad (11.4)$$

و خطای تقریب از مرتبه $\mathcal{O}(h^4)$ است. با تکرار روند بالا روی رابطه (۱۰.۴)، می‌توان یک تقریب از مرتبه $\mathcal{O}(h^6)$ برای محاسبه مقدار انتگرال به دست آورد. با ادامه این روند، می‌توان تقریب‌هایی با دقت مراتب بالاتر نیز به دست آورد.

اکنون، با توجه به نتایج ارائه شده، چگونگی استفاده از روش رامبرگ را برای به دست آوردن یک تقریب مناسب برای مقدار انتگرال ارائه می‌کنیم. نخست مقدار انتگرال را به ازای

$$h_i = \frac{b-a}{2^i}, \quad i = 0, 1, 2, \dots,$$

محاسبه می کنیم و قرار می دهیم

$$T_{i,0} = T(h_i), \quad i = 0, 1, 2, \dots$$

با توجه به اینکه $T_{i,0}$ و $T_{i+1,0}$ ، به ترتیب به ازای h_i و $h_i/2$ به دست آمده اند، از رابطه (۱۱.۴) داریم

$$\int_a^b f(x) dx \approx \frac{4T_{i+1,0} - T_{i,0}}{4 - 1} =: T_{i,1}, \quad i = 0, 1, \dots,$$

و خطای تقریب از مرتبه $O(h_i^4)$ است. در حالت کلی، تعریف می کنیم

$$T_{i,p} = \frac{4^p T_{i+1,p-1} - T_{i,p-1}}{4^p - 1}.$$

در این صورت، می توان دید که $\int_a^b f(x) dx \approx T_{i,p}$ و خطای تقریب از مرتبه $O(h_i^{2p+2})$ است. اکنون، با توجه به مطالب بالا، جدول زیر (جدول رامبرگ) را تشکیل می دهیم.

$O(h_i^2)$	$O(h_i^4)$	$O(h_i^6)$	$O(h_i^8)$
$T_{0,0} \rightarrow$	$T_{0,1} \rightarrow$	$T_{0,2} \rightarrow$	$T_{0,3}$
$T_{1,0} \nearrow$	$T_{1,1} \nearrow$	$T_{1,2} \nearrow$	
$T_{2,0} \nearrow$	$T_{2,1} \nearrow$		
$T_{3,0} \nearrow$			

برای روشن شدن مطلب، مثال زیر را ارائه می کنیم.

مثال ۳.۴ برای محاسبه مقدار $\int_0^2 x^5 dx$ ، با استفاده از قاعده دوزنقه ای مرکب به ازای

$$h_i = \frac{b-a}{2^i} = \frac{2}{2^i}, \quad i = 0, 1, 2,$$

مقادیر تقریبی انتگرال را محاسبه می کنیم. به ازای مقادیر $i = 0, 1, 2$ داریم

$$i = 0, \quad h_0 = \frac{b-a}{2^0} = \frac{2-0}{2^0} = 2,$$

$$T_{0,0} = T(h_0) = \frac{h_0}{2} (f(a) + f(b)) = \frac{2}{2} (f(0) + f(2)) = 0^5 + 2^5 = 32,$$

$$i = 1, \quad h_1 = \frac{b-a}{2^1} = \frac{2-0}{2^1} = 1,$$

$$T_{1,0} = T(h_1) = \frac{h_1}{2} (f(a) + 2f(x_1) + f(b)) = \frac{1}{2} (f(0) + 2f(1) + f(2))$$

$$= \frac{1}{2} (0^5 + 2 \times 1^5 + 2^5) = 17,$$

$$\begin{aligned}
 i = 2, \quad h_2 &= \frac{b-a}{22} = \frac{2-0}{22} = \frac{1}{11}, \\
 T_{2,0} &= T(h_2) = \frac{h_2}{4} (f(a) + 2(f(x_1) + f(x_2) + f(x_3)) + f(b)) \\
 &= \frac{1}{4} \left(f(0) + 2 \left(f\left(\frac{1}{11}\right) + f(1) + f\left(\frac{3}{11}\right) \right) + f(2) \right) \\
 &= \frac{1}{4} \left(0.5 + 2 \times \left(\left(\frac{1}{11}\right)^5 + 1^5 + \left(\frac{3}{11}\right)^5 \right) + 2^5 \right) = \frac{197}{16}.
 \end{aligned}$$

اینک، جدول رامبرگ به صورت زیر خواهد بود.

$\mathcal{O}(h_i^2)$	$\mathcal{O}(h_i^4)$	$\mathcal{O}(h_i^6)$
$T_{0,0} = 32 \rightarrow$	$T_{0,1} = \frac{4 \times 17 - 32}{4-1} = 12 \rightarrow$	$T_{0,2} = \frac{4^2 \times \frac{197}{16} - 12}{4^2-1} = \frac{32}{3}$
$T_{1,0} = 17 \nearrow$	$T_{1,1} = \frac{4 \times \frac{197}{16} - 17}{4-1} = \frac{43}{4} \nearrow$	
$T_{2,0} = \frac{197}{16} \nearrow$		

هریک از اعداد این جدول، یک تقریب برای مقدار انتگرال به دست می‌دهد و بهترین تقریب برابر با $T_{0,2} = \frac{32}{3}$ (البته در اینجا مقدار دقیق انتگرال) است.

۶.۴ فرمول‌های انتگرال گیری گاوس

با توجه به قضیه ۱.۴، در فرمول‌های انتگرال گیری نیوتن-کاتس، $n+1$ گره x_i ، $i = 0, 1, \dots, n$ ، داده شده است و وزن‌های w_i ، $i = 0, 1, \dots, n$ ، به گونه‌ای محاسبه می‌شوند که فرمول انتگرال گیری

$$\int_a^b f(x) dx \approx \sum_{i=0}^n w_i f(x_i),$$

برای هر چند جمله‌ای از درجه n دقیق باشد. در این بخش، می‌خواهیم موضوع یافتن نقاط گرهی x_i و وزن‌های w_i ، $i = 0, 1, \dots, n$ ، را بررسی کنیم به گونه‌ای که این فرمول انتگرال گیری برای تمام چند جمله‌ای‌های حداکثر از درجه $2n+1$ دقیق باشد؛ به عبارت دیگر، داشته باشیم

$$\sum_{i=0}^n w_i f(x_i) = \int_a^b x^i dx, \quad i = 0, 1, \dots, 2n+1.$$

خواهیم دید که این دستگاه جواب یکتا دارد و نقاط x_i ، $i = 0, 1, \dots, n$ متمایز هستند. در ادامه، فرمول انتگرال گیری را برای شکل کلی‌تری از انتگرال‌ها مثل

$$I(f) = \int_a^b w(x) f(x) dx,$$

مورد بررسی قرار می‌دهیم، که در آن $w(x)$ تابع وزن نامیده می‌شود. فرض می‌کنیم تابع $w : (a, b) \rightarrow \mathbb{R}$ ، پیوسته و مثبت باشد، به گونه‌ای که $\int_a^b w(x) dx$ موجود است. یک تابع وزن، معمولاً به صورت یکی از توابع زیر در نظر گرفته می‌شود

$$w(x) = 1, \quad w(x) = \sqrt{1-x^2}, \quad w(x) = \frac{1}{\sqrt{1-x^2}}.$$

توجه کنید که برای تابع دوم، داریم $[a, b] = [-1, 1]$ و برای تابع سوم $(a, b) = (-1, 1)$.

تعریف ۱.۴ فرمول انتگرال گیری

$$\int_a^b w(x)f(x) dx \approx \sum_{k=0}^n w_k f(x_k),$$

که در آن x_k ها متمایز هستند، یک فرمول انتگرال گیری گاوس نامیده می‌شود، هرگاه برای هر $p \in \Pi_{2n+1}$ دقیق باشد، یعنی

$$\int_a^b w(x)p(x) dx = \sum_{k=0}^n w_k p(x_k).$$

x_k ها را نقاط کوادراتور گاوس و w_k ها را وزن‌های کوادراتور گاوس می‌نامند.

برای به دست آوردن فرمول‌های انتگرال گیری گاوس نخست برخی از مفاهیم را یادآوری می‌کنیم. فرض کنید

$$\mathcal{V} = L_w^2(a, b) = \left\{ f \in L^2(a, b) : \int_a^b w(x)f^2(x) dx < \infty \right\}.$$

اکنون، برای هر دو تابع $f, g \in L_w^2(a, b)$ ، ضرب داخلی این دو تابع را به صورت

$$\langle f, g \rangle = \int_a^b w(x)f(x)g(x) dx, \quad (12.4)$$

تعریف می‌کنیم. نرم تولیدشده توسط این ضرب داخلی روی \mathcal{V} ، به صورت

$$\|f\|_w = \left(\int_a^b w(x)f^2(x) dx \right)^{\frac{1}{2}},$$

خواهد بود. همچنین، دو تابع $f, g \in \mathcal{V}$ را متعامد می‌گوییم، هرگاه $\langle f, g \rangle = 0$.

تعریف ۲.۴ مجموعه توابع $\mathcal{V} \subset \{\varphi_1, \varphi_2, \dots, \varphi_n\}$ را نسبت به تابع وزن w متعامد می‌گوییم، هرگاه

$$\langle \varphi_i, \varphi_j \rangle = \int_a^b w(x)\varphi_i(x)\varphi_j(x) dx = 0, \quad i \neq j.$$

قضیه ۳.۴ هر مجموعه متعامد نسبت به ضرب داخلی (۱۲.۴) از توابع ناصفر مثل $\{\varphi_1, \varphi_2, \dots, \varphi_n\} \subset \mathcal{V}$ یک مجموعه مستقل خطی تشکیل می‌دهد.

برهان: فرض کنید

$$\sum_{i=1}^n \alpha_i \varphi_i = 0.$$

در این صورت، برای هر $1 \leq k \leq n$ داریم

$$0 = \langle \varphi_k, \sum_{i=1}^n \alpha_i \varphi_i \rangle = \sum_{i=1}^n \alpha_i \langle \varphi_k, \varphi_i \rangle = \alpha_k \langle \varphi_k, \varphi_k \rangle = \alpha_k \|\varphi_k\|_w^2.$$

در نتیجه، $\alpha_k = 0$ که اثبات قضیه را کامل می‌کند. \square

قضیه ۴.۴ فرض کنید $\mathcal{V} = \{p_j \in \Pi_n : j = 0, 1, \dots, n\}$ به طوری که

$$\langle p_i, p_j \rangle \begin{cases} = 0, & i \neq j, \\ \neq 0 & i = j. \end{cases}$$

در این صورت، \mathcal{V} یک پایه برای Π_n تشکیل می‌دهد.

برهان: با توجه به قضیه ۴.۴، مجموعه \mathcal{V} یک مجموعه مستقل خطی است. بنابراین، با توجه به

اینکه $\dim(\Pi_n) = n + 1$ ، مجموعه \mathcal{V} یک پایه برای Π_n تشکیل می‌دهد. \square

قضیه ۵.۴ فرض کنید $\bar{\Pi}_n$ مجموعه تمام چندجمله‌ای‌های تکین از درجه حداکثر n است. در این

صورت، چندجمله‌ای‌های $p_j \in \bar{\Pi}_j$ ، $j = 0, 1, 2, \dots$ موجودند به طوری که

$$\langle p_j, p_k \rangle = 0, \quad j \neq k.$$

این چندجمله‌ای‌ها به طور یکتا با کمک رابطه بازگشتی

$$p_0(x) = 1,$$

$$p_{i+1}(x) = (x - \delta_{i+1})p_i(x) - \gamma_{i+1}^2 p_{i-1}(x),$$

محاسبه می‌شوند، که در آن $p_{-1}(x) = 0$ و

$$\delta_{i+1} = \frac{\langle xp_i, p_i \rangle}{\langle p_i, p_i \rangle}, \quad i \geq 0$$

$$\gamma_{i+1}^2 = \begin{cases} 0, & i = 0, \\ \frac{\langle p_i, p_i \rangle}{\langle p_{i-1}, p_{i-1} \rangle} & i \geq 1. \end{cases}$$

برهان: اثبات به استقرا روی i انجام می شود. اگر $i = 0$ ، آنگاه

$$p_1(x) = (x - \delta_1)p_0(x),$$

واضح است که $p_1 \in \bar{\Pi}_1$ از طرفی، داریم

$$\begin{aligned} \langle p_1, p_0 \rangle &= \langle (x - \delta_1)p_0, p_0 \rangle \\ &= \langle xp_0, p_0 \rangle - \delta_1 \langle p_0, p_0 \rangle \\ &= \langle xp_0, p_0 \rangle - \frac{\langle xp_0, p_0 \rangle}{\langle p_0, p_0 \rangle} \langle p_0, p_0 \rangle = 0. \end{aligned}$$

اکنون، فرض کنید $k \in \bar{\Pi}_k$ ، $k = 0, 1, \dots, i$ ، دو به دو متعامد باشند. نشان می دهیم که p_{i+1} که توسط روند بازگشتی به دست می آید، بر $\bar{\Pi}_k$ ، $k = 0, 1, \dots, i$ عمود است. برای این کار، داریم

$$\begin{aligned} \langle p_{i+1}, p_i \rangle &= \langle (x - \delta_{i+1})p_i - \gamma_{i+1}^2 p_{i-1}, p_i \rangle \\ &= \langle (x - \delta_{i+1})p_i, p_i \rangle \\ &= \langle xp_i, p_i \rangle - \delta_{i+1} \langle p_i, p_i \rangle \\ &= \langle xp_i, p_i \rangle - \frac{\langle xp_i, p_i \rangle}{\langle p_i, p_i \rangle} \langle p_i, p_i \rangle = 0. \end{aligned}$$

همچنین، داریم

$$\begin{aligned} \langle p_{i+1}, p_{i-1} \rangle &= \langle (x - \delta_{i+1})p_i - \gamma_{i+1}^2 p_{i-1}, p_{i-1} \rangle \\ &= \langle xp_i, p_{i-1} \rangle - \gamma_{i+1}^2 \langle p_{i-1}, p_{i-1} \rangle \\ &= \langle p_i, xp_{i-1} \rangle - \gamma_{i+1}^2 \langle p_{i-1}, p_{i-1} \rangle. \end{aligned} \quad (13.4)$$

روشن است که اعدادی مثل $\alpha_j \in \mathbb{R}$ ، $j = 0, 1, \dots, i$ وجود دارند به گونه ای که $\alpha_i = 1$ و

$$xp_{i-1} = \sum_{j=0}^i \alpha_j p_j.$$

با جای گذاری این رابطه در (۱۳.۴)، خواهیم داشت

$$\begin{aligned} \langle p_{i+1}, p_{i-1} \rangle &= \langle p_i, \sum_{j=0}^i \alpha_j p_j \rangle - \gamma_{i+1}^2 \langle p_{i-1}, p_{i-1} \rangle \\ &= \langle p_i, p_i \rangle - \gamma_{i+1}^2 \langle p_{i-1}, p_{i-1} \rangle = 0. \end{aligned}$$

سرانجام، فرض کنید $2 - i \leq j$. در این صورت، داریم

$$\begin{aligned} \langle p_{i+1}, p_j \rangle &= \langle (x - \delta_{i+1})p_i - \gamma_{i+1}^\vee p_{i-1}, p_j \rangle \\ &= \langle xp_i, p_j \rangle - \delta_{i+1} \langle p_i, p_j \rangle - \gamma_{i+1}^\vee \langle p_{i-1}, p_j \rangle \\ &= \langle xp_i, p_j \rangle = \langle p_i, xp_j \rangle. \end{aligned} \quad (14.4)$$

کاملاً پیداست که $xp_j \in \bar{\Pi}_{i-1}$ و بنابراین، اعدادی مثل $\alpha_k \in \mathbb{R}$ ، $k = 0, 1, \dots, i-1$ وجود دارند به طوری که

$$xp_j = \sum_{k=0}^{i-1} \alpha_k p_k,$$

با جای‌گذاری این رابطه در (۱۴.۴)، داریم

$$\begin{aligned} \langle p_{i+1}, p_j \rangle &= \langle p_i, xp_j \rangle = \langle p_i, \sum_{k=0}^{i-1} \alpha_k p_k \rangle \\ &= \sum_{k=0}^{i-1} \alpha_k \langle p_i, p_k \rangle = 0. \end{aligned}$$

برای اثبات یکتایی، فرض کنید $q_j \in \bar{\Pi}_j$ ، $j = 0, 1, 2, \dots$ ، مجموعه دیگری از چند جمله‌ای‌های متعامد است. به روشنی، اعدادی مثل $\alpha_k \in \mathbb{R}$ ، $k = 0, 1, \dots, j$ وجود دارند به طوری که $\alpha_j = 1$

و

$$q_j = \sum_{k=0}^j \alpha_k p_k. \quad (15.4)$$

به همین ترتیب، اعدادی مثل $\beta_r \in \mathbb{R}$ ، $r = 0, 1, \dots, k$ وجود دارند به طوری که $\beta_{kk} = 1$ و

$$p_k = \sum_{r=0}^k \beta_{rk} q_r.$$

با جای‌گذاری در رابطه (۱۵.۴)، داریم

$$q_j = \sum_{k=0}^j \left(\sum_{r=0}^k \alpha_k \beta_{rk} q_r \right) = \sum_{r=0}^j \left(\sum_{k=r}^j \alpha_k \beta_{rk} \right) q_r.$$

با توجه به اینکه q_r ها مستقل خطی هستند، خواهیم داشت

$$\begin{cases} \sum_{k=r}^j \alpha_k \beta_{rk} = 0, & r = 0, 1, \dots, j-1, \\ \sum_{k=r}^j \alpha_k \beta_{rk} = 1, & r = j. \end{cases}$$

این رابطه‌ها به صورت ماتریسی به شکل

$$\begin{pmatrix} \beta_{00} & \beta_{01} & \cdots & \beta_{0j} \\ 0 & \beta_{11} & \cdots & \beta_{1j} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \beta_{jj} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_j \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

نوشته می‌شوند. با توجه به اینکه درایه‌های روی قطر ماتریس ضرایب برابر یک هستند، این دستگاه جواب یکتای $\alpha_j = 1$ و $\alpha_r = 0$ ، $r = 0, 1, \dots, j-1$ دارد. با جای‌گذاری این مقادیر در رابطه (۱۵.۴)، خواهیم داشت $p_j = q_j$ و به این ترتیب اثبات قضیه کامل می‌شود. \square

نتیجه ۱.۴ برای هر $p \in \Pi_{n-1}$ داریم $\langle p, p_n \rangle = 0$.

قضیه ۶.۴ ریشه‌های p_n ، یعنی x_i ، $i = 1, \dots, n$ حقیقی و ساده هستند و همگی در بازه (a, b) قرار دارند.

برهان: ریشه‌هایی از p_n را در نظر بگیرید که در بازه (a, b) قرار دارند و مرتبه تکرار آنها فرد است. فرض کنید این ریشه‌ها به صورت

$$a < x_1 < x_2 < \cdots < x_l < b,$$

هستند. قرار می‌دهیم

$$q(x) = \prod_{j=1}^l (x - x_j) \in \bar{\Pi}_l.$$

روشن است که چندجمله‌ای $p_n(x)q(x)$ در بازه (a, b) ، تغییر علامت نمی‌دهد. بنابراین، داریم

$$\langle q, p_n \rangle = \int_a^b w(x)p_n(x)q(x)dx \neq 0,$$

\square که با نتیجه ۱.۴ در تناقض است. در نتیجه، خواهیم داشت $\deg(q) = l = n$.

قضیه ۷.۴ فرض کنید x_i ، $i = 0, 1, \dots, n$ نقاط کوادراتوریک فرمول انتگرال گیری گاوس باشند،

یعنی

$$\int_a^b w(x)p(x)dx = \sum_{k=0}^n w_k p(x_k), \quad \forall p \in \Pi_{n+1}.$$

در این صورت، برای هر $q \in \Pi_n$ داریم

$$\int_a^b w(x)q_{n+1}(x)q(x)dx = 0,$$

$$q_{n+1}(x) = \prod_{j=0}^n (x - x_j)$$

برهان: با توجه به اینکه درجه چندجمله‌ای q_{n+1} حداکثر $n + 1$ است، داریم

$$\int_a^b w(x)q_{n+1}(x)q(x)dx = \sum_{k=0}^n w_k q_{n+1}(x_k)q(x_k) = 0.$$

□ به این ترتیب، اثبات قضیه کامل می‌شود.

قضیه ۸.۴ فرض کنید $n + 1$ نقطه متمایز باشند و

$$q_{n+1}(x) = \prod_{j=0}^n (x - x_j),$$

به طوری که

$$\int_a^b w(x)q_{n+1}(x)q(x)dx = 0, \quad \forall q \in \Pi_n.$$

همچنین، فرض کنید

$$w_k = \int_a^b w(x)L_k(x)dx, \quad k = 0, 1, \dots, n,$$

که در آن

$$L_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}.$$

در این صورت

$$\int_a^b w(x)f(x)dx \approx \sum_{k=0}^n w_k f(x_k),$$

یک فرمول کوادراتور گاوس است.

برهان: فرض کنید $p \in \Pi_{2n+1}$ و $p_n(x)$ چندجمله‌ای درونیاب تابع p در نقاط $x_i, i = 0, 1, \dots, n$ باشد. در نتیجه

$$p(x_i) - p_n(x_i) = 0, \quad i = 0, 1, \dots, n.$$

این رابطه نشان می‌دهد که $i = 0, 1, \dots, n, x_i$ ریشه‌های $p - p_n$ هستند. بنابراین، یک چندجمله‌ای مثل $q \in \Pi_n$ وجود دارد به طوری که $p - p_n = q_{n+1}q$. در نتیجه، داریم

$$\int_a^b w(x)p(x)dx = \int_a^b w(x)p_n(x)dx + \int_a^b w(x)q_{n+1}(x)q(x)dx$$

$$\begin{aligned} &= \int_a^b w(x)p_n(x)dx \\ &= \int_a^b w(x) \left(\sum_{k=0}^n p(x_k)L_k(x) \right) dx \\ &= \sum_{k=0}^n w_k p(x_k), \end{aligned}$$

□ که اثبات قضیه را کامل می کند.

قضیه ۹.۴ به ازای هر $n = 0, 1, 2, \dots$ یک فرمول کوادراتور گاوس یکتا وجود دارد.

برهان: چندجمله ای q_{n+1} از چندجمله ای های متعامد تکین نسبت به وزن w بر بازه $[a, b]$ را در نظر بگیرد. ریشه های q_{n+1} همگی ساده و در بازه $[a, b]$ قرار دارند. از این رو، داریم

$$q_{n+1}(x) = \prod_{j=0}^n (x - x_j),$$

که در آن $x_j, j = 0, 1, \dots, n$ ریشه های q_{n+1} هستند. با توجه به اینکه چندجمله ای های q_j ، $j = 0, 1, \dots, n, n+1$ متعامد هستند، داریم

$$\int_a^b w(x)q_{n+1}(x)q(x)dx = 0, \quad \forall q \in \Pi_n.$$

در نتیجه، با توجه به قضیه قبل، x_k ها نقاط کوادراتور گاوس یک فرمول انتگرال گیری گاوس هستند.

□

قضیه ۱۰.۴ وزن های کوادراتور گاوس همگی مثبت هستند.

برهان: به ازای هر $0 \leq k \leq n$ قرار می دهیم

$$f_k(x) = \left(\frac{q_{n+1}(x)}{x - x_k} \right)^2, \quad k = 0, 1, \dots, n,$$

که در آن $q_{n+1}(x) = \prod_{j=0}^n (x - x_j)$. در این صورت، با توجه به اینکه $f_k \in \Pi_{2n+1}$ ، داریم

$$\int_a^b w(x)f_k(x)dx = \sum_{j=0}^n w_j f_k(x_j) = w_k f_k(x_k). \quad (16.4)$$

از طرفی، داریم

$$\begin{aligned} f_k(x_k) &= \lim_{x \rightarrow x_k} f_k(x) = \lim_{x \rightarrow x_k} \left(\frac{q_{n+1}(x) - q_{n+1}(x_k)}{x - x_k} \right)^2 \\ &= \left(\lim_{x \rightarrow x_k} \frac{q_{n+1}(x) - q_{n+1}(x_k)}{x - x_k} \right)^2 = (q'_{n+1}(x_k))^2. \end{aligned}$$

با جای گذاری $f_k(x_k)$ در رابطه (۱۶.۴)، خواهیم داشت

$$\circ < \int_a^b w(x) f_k(x) dx = w_k (q'_{n+1}(x_k))^2.$$

این رابطه نشان می‌دهد که $w_k > \circ$. □

قضیه ۱۱.۴ فرض کنید $f \in C^{(2n+2)}[a, b]$. در این صورت، خطای فرمول کوادراتور گاوس از درجه n به صورت

$$\int_a^b w(x) f(x) dx - \sum_{k=\circ}^n w_k f(x_k) = \frac{f^{(2n+2)}(\alpha)}{(2n+2)!} \int_a^b w(x) (q_{n+1}(x))^2 dx,$$

است، که در آن $\alpha \in (a, b)$.

برهان: فرض کنید H_n چند جمله‌ای درونیاب هرمیت با شرایط

$$\begin{cases} H_n(x_k) = f(x_k), \\ H'_n(x_k) = f'(x_k), \end{cases} \quad k = \circ, 1, \dots, n,$$

است. در این صورت، $H_n \in \Pi_{2n+1}$ و بنابراین

$$\int_a^b w(x) H_n(x) dx = \sum_{k=\circ}^n w_k H_n(x_k) = \sum_{k=\circ}^n w_k f(x_k). \quad (17.4)$$

اکنون، فرض کنید $f \in C^{2n+2}[a, b]$. در این صورت، با استفاده از رابطه (۱۷.۴)، داریم

$$\begin{aligned} E_n(f) &= \int_a^b w(x) f(x) dx - \sum_{k=\circ}^n w_k f(x_k) \\ &= \int_a^b w(x) f(x) dx - \int_a^b w(x) H_n(x) dx \\ &= \int_a^b w(x) (f(x) - H_n(x)) dx. \end{aligned}$$

بنابراین، با استفاده از فرمول خطای درونیابی هر میت، خواهیم داشت

$$\begin{aligned} E_n(f) &= \int_a^b w(x)(f(x) - H_n(x)) dx \\ &= \int_a^b w(x) \frac{f^{(\nu_{n+2})}(\xi_x)}{(\nu_{n+2})!} (q_{n+1}(x))^{\nu} dx \\ &= \frac{f^{(\nu_{n+2})}(\alpha)}{(\nu_{n+2})!} \int_a^b w(x) (q_{n+1}(x))^{\nu} dx. \end{aligned}$$

توجه کنید که رابطه آخر با استفاده قضیه مقدار میانگین به دست آمده است. □

در ادامه، موضوع محاسبه وزن های کوادراتور گاوس را بررسی می کنیم.

قضیه ۱۲.۴ فرض کنید چند جمله ای های $p_j \in \Pi_j$ ، $j = 0, 1, \dots, n$ ، دویبه دو متعامد باشند و t_j ، $j = 0, 1, \dots, n$ ، عدد حقیقی متمایز هستند. در این صورت، ماتریس

$$A = \begin{pmatrix} p_0(t_0) & p_0(t_1) & \cdots & p_0(t_n) \\ p_1(t_0) & p_1(t_1) & \cdots & p_1(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_n(t_0) & p_n(t_1) & \cdots & p_n(t_n) \end{pmatrix},$$

نامنفرد است.

برهان: به برهان خلف، فرض کنید ماتریس A منفرد است. در این صورت، A^T نیز منفرد است و بنابراین، برداری مثل $v = (v_0, v_1, \dots, v_n)^T \in \mathbb{R}^{n+1}$ وجود دارد به طوری که $A^T v = 0$ ، که این خود با

$$\sum_{j=0}^n v_j p_j(t_i) = 0, \quad i = 0, 1, \dots, n,$$

معادل است. با قراردادن $q(x) = \sum_{j=0}^n v_j p_j(t)$ ، رابطه اخیر نشان می دهد که q دارای $n+1$ ریشه متمایز است. این نتیجه با اینکه $\deg(q) \leq n$ ، در تناقض است. □

قضیه ۱۳.۴ وزن های کوادراتور گاوس w_j ، $j = 0, 1, \dots, n$ ، جواب دستگاه معادلات خطی

$$\sum_{j=0}^n p_k(x_j) w_j = \begin{cases} \langle p_0, p_0 \rangle, & k = 0, \\ 0, & k = 1, 2, \dots, n, \end{cases}$$

هستند، که در آن x_j ، $j = 0, 1, \dots, n$ ، ریشه های چند جمله ای p_{n+1} هستند.

برهان: همان‌گونه که دیدیم، فرمول کوادراتور گاوس

$$\int_a^b w(x)f(x)dx \approx \sum_{j=0}^n w_j f(x_j),$$

برای تمام چندجمله‌ای‌های $p \in \Pi_{2n+1}$ دقیق است. در این صورت، با فرض $p = p_k$ داریم، $k = 0, 1, \dots, n$

$$\sum_{j=0}^n p_k(x_j)w_j = \int_a^b w(x)p_k(x)dx = \langle p_k, p_0 \rangle = \begin{cases} \langle p_0, p_0 \rangle, & k = 0, \\ 0, & k = 1, 2, \dots, n. \end{cases}$$

با توجه به قضیه قبل، ماتریس ضرایب این دستگاه نامنفرد است و بنابراین، دستگاه جواب یکتا دارد \square که همان وزن‌های کوادراتور گاوس هستند.

۷.۴ فرمول‌های انتگرال‌گیری گاوس-چبیشف و گاوس-لژاندر

با کمک قضیه ۵.۴ و انتخاب یک تابع وزن مناسب، می‌توان یک مجموعه از چندجمله‌ای‌های متعامد به دست آورد و سپس با کمک آنها و نتایج ارائه شده در بخش پیشین، فرمول‌های کوادراتور گاوس متناظر را تولید کرد. در این بخش، فرمول‌های انتگرال‌گیری گاوس-چبیشف و گاوس-لژاندر را معرفی می‌کنیم که به ترتیب، با کمک چندجمله‌ای‌های چبیشف و لژاندر تولید می‌شوند. نخست فرمول انتگرال‌گیری گاوس-چبیشف را معرفی می‌کنیم.

فرمول انتگرال‌گیری گاوس-چبیشف

فرض کنید $(a, b) = (-1, 1)$ و

$$w(x) = \frac{1}{\sqrt{1-x^2}}.$$

همان‌گونه که در بخش ۴.۳ دیدیم، چندجمله‌ای‌های چبیشف به صورت

$$T_n(x) = \cos(n \cos^{-1} x), \quad -1 \leq x \leq 1,$$

تعریف می‌شوند. با تغییر متغیر $x = \cos t$ داریم

$$\langle T_n, T_m \rangle = \int_{-1}^1 w(x)T_n(x)T_m(x)dx$$

$$\begin{aligned}
 &= \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx \\
 &= \int_0^\pi \cos nt \cos mtdt \\
 &= \frac{1}{2} \int_0^\pi (\cos(n-m)t + \cos(m+n)t) dt \\
 &= \begin{cases} \pi, & n = m = 0, \\ \frac{\pi}{2}, & n = m > 0, \\ 0, & n \neq m. \end{cases} \quad (18.4)
 \end{aligned}$$

این رابطه نشان می‌دهد که T_n ها دوجه دو متعامد هستند. با قراردادن $q_n = 2^{1-n} T_n$ و $q_0 = T_0$ ، $n = 1, 2, \dots$ می‌بینیم که q_n ها تکین هستند و در شرایط قضایای بخش قبل صدق می‌کنند. بنابراین، به‌ازای هر $n = 0, 1, 2, \dots$ ، چند جمله‌ای q_n یک فرمول کوادراتور گاوس تولید می‌کند. این فرمول‌ها، به فرمول‌های کوادراتور گاوس-چبیشف معروف هستند.

ریشه‌های q_{n+1} (ریشه‌های T_{n+1})، یعنی

$$x_k = \cos\left(\frac{(2k+1)\pi}{2(n+1)}\right), \quad k = 0, 1, \dots, n,$$

نقاط کوادراتور گاوس-چبیشف هستند و وزن‌های کوادراتور از حل دستگاه معادلات خطی

$$\sum_{k=0}^n q_m(x_k) w_k = \begin{cases} \langle q_0, q_0 \rangle, & m = 0, \\ 0, & m = 1, 2, \dots, n. \end{cases} \quad (19.4)$$

به‌دست می‌آیند. با استفاده از رابطه (۱۸.۴)، داریم

$$\langle q_0, q_0 \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} q_0(x)^2 dx = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_0(x)^2 dx = \pi.$$

در نتیجه، دستگاه (۱۹.۴) را می‌توان به‌صورت

$$\sum_{k=0}^n T_m(x_k) w_k = \begin{cases} \pi, & m = 0, \\ 0, & m = 1, 2, \dots, n, \end{cases}$$

نوشت. بنابراین، با توجه به اینکه

$$T_m(x_k) = \cos\left(\frac{(2k+1)m\pi}{2(n+1)}\right),$$

دستگاه اخیر به‌صورت

$$\sum_{k=0}^n w_k \cos\left(\frac{(2k+1)m\pi}{2(n+1)}\right) = \begin{cases} \pi, & m = 0, \\ 0, & m = 1, 2, \dots, n, \end{cases} \quad (20.4)$$

درمی‌آید. برای حل این دستگاه، نخست می‌بینیم که (تمرین ۸.۴ را ببینید)

$$\sum_{k=0}^n \cos\left(\frac{(2k+1)m\pi}{2(n+1)}\right) = \begin{cases} n+1, & m=0, \\ 0, & m=1, 2, \dots, n, \end{cases} \quad (21.4)$$

و بنابراین

$$\sum_{k=0}^n \frac{\pi}{n+1} \cos\left(\frac{(2k+1)m\pi}{2(n+1)}\right) = \begin{cases} \pi, & m=0, \\ 0, & m=1, 2, \dots, n. \end{cases}$$

با مقایسه این رابطه و رابطه (۲۰.۴)، نتیجه می‌گیریم که

$$w_k = \frac{\pi}{n+1}, \quad k=0, 1, \dots, n,$$

جواب دستگاه (۲۰.۴) است.

با جمع‌بندی مطالب ارائه‌شده، فرمول‌های کوادراتور گاوس-چبیشف به صورت زیر خواهند بود

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n+1} \sum_{k=0}^n f\left(\cos \frac{(2k+1)\pi}{2(n+1)}\right),$$

و خطای انتگرال‌گیری آن به صورت زیر است

$$\begin{aligned} E(f) &= \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx - \frac{\pi}{n+1} \sum_{k=0}^n f\left(\cos \frac{(2k+1)\pi}{2(n+1)}\right) \\ &= \frac{f^{(2n+2)}(\alpha)}{(2n+2)!} \int_{-1}^1 w(x)(q_{n+1}(x))^2 dx \\ &= \frac{f^{(2n+2)}(\alpha)}{(2n+2)!} (2^{-n})^2 \int_{-1}^1 w(x)(T_{n+1}(x))^2 dx \\ &= \frac{f^{(2n+2)}(\alpha)}{(2n+2)!} (2^{-n})^2 \frac{\pi}{2} \\ &= \frac{\pi}{2^{2n+1}(2n+2)!} f^{(2n+2)}(\alpha). \end{aligned}$$

به‌ازای $n=0$ ، فرمول انتگرال‌گیری یک‌نقطه‌ای گاوس-چبیشف به دست می‌آید. در این صورت، تنها ریشه چندجمله‌ای $x = T_1(x) = x$ برابر با $x_0 = 0$ است. همچنین، تنها وزن w_0 برابر با π است. در نتیجه، فرمول یک‌نقطه‌ای گاوس-چبیشف به صورت

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x) dx \approx \pi f(0),$$

و خطای روش برابر با

$$E(f) = \frac{\pi}{4} f''(\alpha),$$

خواهد بود. اکنون، فرض می‌کنیم $n = 1$. در این صورت، ریشه‌های T_2 را می‌یابیم. ریشه‌های T_2 عبارتند از

$$x_0 = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad x_1 = \cos \frac{3\pi}{4} = -\frac{\sqrt{2}}{2}.$$

همچنین، داریم

$$w_0 = w_1 = \frac{\pi}{4}.$$

بنابراین، فرمول دونقطه‌ای گاوس-چبیشف به صورت زیر خواهد بود

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x) dx \approx \frac{\pi}{4} \left(f\left(-\frac{\sqrt{2}}{2}\right) + f\left(\frac{\sqrt{2}}{2}\right) \right).$$

خطای این روش برابر است با

$$E(f) = \frac{\pi}{192} f^{(4)}(\alpha).$$

اینک، فرض کنید $n = 2$. در این صورت، ریشه‌های چندجمله‌ای T_3 عبارتند از

$$x_0 = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad x_1 = 0, \quad x_2 = \cos \frac{5\pi}{6} = -\frac{\sqrt{3}}{2}.$$

همچنین، داریم

$$w_0 = w_1 = w_2 = \frac{\pi}{3}.$$

از این رو، فرمول سه نقطه‌ای گاوس-چبیشف به صورت زیر خواهد بود

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x) dx \approx \frac{\pi}{3} \left(f\left(-\frac{\sqrt{3}}{2}\right) + f(0) + f\left(\frac{\sqrt{3}}{2}\right) \right).$$

خطای این روش برابر است با

$$E(f) = \frac{\pi}{23040} f^{(6)}(\alpha).$$

مثال ۴.۴ مقدار تقریبی

$$I = \int_{-1}^1 \frac{\cos x}{\sqrt{1-x^2}} dx,$$

را با استفاده از فرمول‌های انتگرال‌گیری دونقطه‌ای و سه‌نقطه‌ای گاوس-چبیشف به دست می‌آوریم. با فرض $f(x) = \cos x$ و فرمول دونقطه‌ای گاوس-چبیشف، داریم

$$\begin{aligned} \int_{-1}^1 f(x) dx &\approx \frac{\pi}{2} \left(f\left(-\frac{\sqrt{2}}{2}\right) + f\left(\frac{\sqrt{2}}{2}\right) \right) \\ &= \frac{\pi}{2} \left(\cos\left(-\frac{\sqrt{2}}{2}\right) + \cos\left(\frac{\sqrt{2}}{2}\right) \right) \\ &= 2.388379. \end{aligned}$$

اگر از فرمول سه‌نقطه‌ای گاوس-چبیشف استفاده کنیم، خواهیم داشت

$$\begin{aligned} \int_{-1}^1 f(x) dx &\approx \frac{\pi}{3} \left(f\left(-\frac{\sqrt{3}}{2}\right) + f(0) + f\left(\frac{\sqrt{3}}{2}\right) \right) \\ &= \frac{\pi}{3} \left(\cos\left(-\frac{\sqrt{3}}{2}\right) + \cos 0 + \cos\left(\frac{\sqrt{3}}{2}\right) \right) \\ &= 2.404071. \end{aligned}$$

شایان ذکر است که مقدار دقیق I با ۶ رقم اعشار درست، برابر با 2.403939 است.

فرمول انتگرال‌گیری گاوس-لژاندر

فرمول‌های انتگرال‌گیری گاوس-لژاندر، معروف‌ترین و مهم‌ترین فرمول‌های انتگرال‌گیری گاوس هستند؛ زیرا تابع وزن این فرمول‌ها برابر با $w(x) = 1$ است. فرض کنید بازه انتگرال‌گیری به صورت $[a, b] = [-1, 1]$ باشد. این فرمول‌ها به کمک چند جمله‌ای‌های متعامد لژاندر به دست می‌آیند که به صورت

$$P_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n} (x^2 - 1)^n,$$

تعریف می‌شوند. به سادگی می‌توان دید که P_n ها تکین هستند و $\deg(P_n) = n$. به ازای چند مقدار از n ، داریم

$$\begin{aligned} P_0(x) &= 1, & P_1(x) &= x, \\ P_2(x) &= x^2 - \frac{1}{3}, & P_3(x) &= x^3 - \frac{3}{5}x, \\ P_4(x) &= x^4 - \frac{6}{7}x^2 + \frac{3}{35}, & P_5(x) &= x^5 - \frac{10}{9}x^3 + \frac{5}{21}x. \end{aligned}$$

نشان می‌دهیم که P_n ها نسبت به وزن $w(x) = 1$ روی بازه $[-1, 1]$ متعامد هستند؛ به عبارت دیگر

$$\langle P_m, P_n \rangle = \int_{-1}^1 P_m(x)P_n(x)dx \begin{cases} = 0, & n \neq m, \\ \neq 0, & n = m. \end{cases}$$

برای این کار، کافی است نشان دهیم که به ازای $m < n$

$$I_{mn} := \int_{-1}^1 \frac{d^m}{dx^m}(x^2 - 1)^m \frac{d^n}{dx^n}(x^2 - 1)^n dx = 0.$$

با فرض $u = \frac{d^m}{dx^m}(x^2 - 1)^m$ و $dv = \frac{d^n}{dx^n}(x^2 - 1)^n dx$ و قاعده انتگرال گیری جزء به جزء، داریم

$$\begin{aligned} I_{mn} &= uv \Big|_{-1}^1 - \int_{-1}^1 v du \\ &= \left[\underbrace{\frac{d^m}{dx^m}(x^2 - 1)^m \frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n}_A \right] \Big|_{-1}^1 \\ &\quad - \int_{-1}^1 \frac{d^{m+1}}{dx^{m+1}}(x^2 - 1)^{m+1} \frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n dx. \end{aligned}$$

روشن است که $\frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n$ بر $x^2 - 1$ بخش پذیر است. بنابراین، مقدار A در عبارت اخیر برابر با صفر خواهد بود. از این رو

$$I_{mn} = - \int_{-1}^1 \frac{d^{m+1}}{dx^{m+1}}(x^2 - 1)^{m+1} \frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n dx.$$

با ادامه این روند به رابطه

$$I_{mn} = (-1)^{m+1} \int_{-1}^1 \frac{d^{m+1}}{dx^{m+1}}(x^2 - 1)^m \frac{d^{n-m-1}}{dx^{n-m-1}}(x^2 - 1)^n dx.$$

می‌رسیم. اکنون، با توجه به اینکه عبارت $(x^2 - 1)^m$ ، یک چندجمله‌ای از درجه $2m$ است، مشتق $(2m + 1)$ ام آن برابر با صفر خواهد بود. در نتیجه، $I_{mn} = 0$.

همان‌گونه که دیدیم، مجموعه چندجمله‌ای‌های لژاندر، یک مجموعه متعامد و تکین هستند. بنابراین، با کمک هریک از این چندجمله‌ای‌ها، می‌توان یک فرمول انتگرال گیری گاوس تولید کرد (فرمول انتگرال گیری گاوس-لژاندر). به ازای $n = 0, 1, 2$ ، فرمول‌های انتگرال گیری گاوس-لژاندر و عبارت خطای هریک را به دست می‌آوریم.

فرض کنید $n = 0$. در این صورت، ریشه $P_1(x) = x = 0$ (یعنی، $x_0 = 0$) تنها نقطه کوادراتور گاوس-لژاندر است. بنابراین، با کمک قضیه ۱۳.۴، تنها وزن کوادراتور گاوس در رابطه

$$P_0(x_0)w_0 = \langle P_0, P_0 \rangle = \int_{-1}^1 P_0^2(x) dx = \int_{-1}^1 dx = 2,$$

صدق می‌کند. اکنون، با توجه به اینکه $P_0(x_0) = 1$ ، نتیجه می‌گیریم که $w_0 = 2$. در این صورت، فرمول انتگرال‌گیری گاوس-لژاندر یک نقطه‌ای به صورت

$$\int_{-1}^1 f(x) dx \approx 2f(0),$$

خواهد بود. این روش به قاعده نقطه میانی معروف است. با استناد به قضیه ۱۱.۴، فرمول خطا برای قاعده نقطه میانی به صورت

$$\begin{aligned} E(f) &= \frac{f''(\alpha)}{2!} \int_{-1}^1 (P_1(x))^2 dx \\ &= \frac{f''(\alpha)}{2!} \int_{-1}^1 x^2 dx \\ &= \frac{1}{3} f''(\alpha), \end{aligned}$$

است.

فرض کنید $n = 1$. در این صورت، ریشه‌های $\frac{1}{\sqrt{3}} - x^2 = 0$ ، یا به عبارتی نقاط کوادراتور گاوس متناظر عبارتند از $x_0 = -\frac{\sqrt{3}}{2}$ و $x_1 = \frac{\sqrt{3}}{2}$. برای محاسبه وزن‌ها کافی است دستگاه

$$\begin{cases} P_0(x_0)w_0 + P_0(x_1)w_1 = \langle P_0, P_0 \rangle, \\ P_1(x_0)w_0 + P_1(x_1)w_1 = 0, \end{cases}$$

یا

$$\begin{cases} w_0 + w_1 = 2, \\ -\frac{\sqrt{3}}{2}w_0 + \frac{\sqrt{3}}{2}w_1 = 0, \end{cases}$$

را حل کنیم. با حل این دستگاه، خواهیم داشت $w_0 = w_1 = 1$. بنابراین، فرمول دونقطه‌ای انتگرال‌گیری گاوس-لژاندر به صورت

$$\int_{-1}^1 f(x) dx \approx f\left(-\frac{\sqrt{3}}{2}\right) + f\left(\frac{\sqrt{3}}{2}\right),$$

خواهد بود. به علاوه، خطای این فرمول به صورت

$$E(f) = \frac{f^{(4)}(\alpha)}{4!} \int_{-1}^1 (P_2(x))^2 dx$$

$$\begin{aligned} &= \frac{f^{(4)}(\alpha)}{4!} \int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx \\ &= \frac{1}{135} f^{(4)}(\alpha), \end{aligned}$$

است.

فرض کنید $n = 2$. در این صورت، ریشه‌های چندجمله‌ای $P_2(x) = x^2 - \frac{3}{5}x$ عبارتند از $x_2 = \sqrt{\frac{3}{5}}$ و $x_1 = 0$ ، $x_0 = -\sqrt{\frac{3}{5}}$ برای محاسبهٔ وزن‌ها، کافی است دستگاه

$$\begin{cases} P_0(x_0)w_0 + P_0(x_1)w_1 + P_0(x_2)w_2 = \langle P_0, P_0 \rangle, \\ P_1(x_0)w_0 + P_1(x_1)w_1 + P_1(x_2)w_2 = 0, \\ P_2(x_0)w_0 + P_2(x_1)w_1 + P_2(x_2)w_2 = 0, \end{cases}$$

یا

$$\begin{cases} w_0 + w_1 + w_2 = 2, \\ -\sqrt{\frac{3}{5}}w_0 + \sqrt{\frac{3}{5}}w_2 = 0, \\ \frac{4}{15}w_0 - \frac{1}{3}w_1 + \frac{4}{15}w_2 = 0, \end{cases}$$

را حل کنیم. با حل این دستگاه، خواهیم داشت $w_0 = w_2 = \frac{5}{9}$ و $w_1 = \frac{4}{9}$. بنابراین، فرمول سه نقطه‌ای انتگرال گیری گاوس-لژاندر به صورت

$$\int_{-1}^1 f(x) dx \approx \frac{1}{9} \left(5f\left(-\sqrt{\frac{3}{5}}\right) + 4f(0) + 5f\left(\sqrt{\frac{3}{5}}\right) \right),$$

خواهد بود. همچنین، خطای این فرمول به صورت

$$\begin{aligned} E(f) &= \frac{f^{(6)}(\alpha)}{6!} \int_{-1}^1 (P_2(x))^2 dx \\ &= \frac{f^{(6)}(\alpha)}{4!} \int_{-1}^1 \left(x^2 - \frac{3}{5}x\right)^2 dx \\ &= \frac{8}{175} f^{(6)}(\alpha), \end{aligned}$$

است.

مثال ۵.۴ می‌خواهیم به کمک فرمول‌های دونقطه‌ای و سه نقطه‌ای انتگرال گیری گاوس-لژاندر، مقدار تقریبی انتگرال

$$I = \int_{-1}^1 e^{-\frac{x^2}{4}} dx,$$

را محاسبه کنیم. مقدار دقیق این انتگرال تا ۶ رقم اعشار درست، برابر با ۱.۷۱۱۲۴۹ است. با فرض $f(x) = e^{-\frac{x^2}{3}}$ ، مقدار تقریبی این انتگرال با کمک فرمول انتگرال‌گیری گاوس-لژاندر دونقطه‌ای، برابر با

$$I \approx f\left(-\frac{\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right) = 1.692963,$$

و با کمک فرمول انتگرال‌گیری گاوس-لژاندر سه‌نقطه‌ای، برابر با

$$I \approx \frac{1}{9} \left(5f\left(-\sqrt{\frac{3}{5}}\right) + 8f(0) + 5f\left(\sqrt{\frac{3}{5}}\right) \right) = 1.712020,$$

است. اگر مقدار I را با استفاده از روش‌های ذوزنقه‌ای و سیمپسون ساده محاسبه کنیم، به ترتیب، مقادیر تقریبی ۱.۲۱۳۰۶۱ و ۱.۷۳۷۶۸۷ به دست می‌آیند. همان‌گونه که می‌بینیم، مقادیر تقریبی به دست آمده توسط فرمول‌های گاوس-لژاندر، به مراتب نسبت به قاعده‌های ذوزنقه‌ای و سیمپسون از دقت بیشتری برخوردار هستند.

اغلب، فرمول‌های کوادراتور گاوس برای بازه $[-1, 1]$ نوشته می‌شود، ولی به راحتی می‌توان این فرمول‌ها را برای هر بازه دلخواه تعمیم داد. در بخش ۴.۳، دیدیم که اگر $x \in [a, b]$ و

$$x = \frac{b+a}{2} + \frac{b-a}{2}t,$$

آنگاه $t \in [-1, 1]$ بنابراین

$$dx = \frac{b-a}{2}dt,$$

و

$$\int_a^b f(x)dx = \frac{b-a}{2} \int_{-1}^1 g(t)dt,$$

که در آن

$$g(t) = f\left(\frac{b+a}{2} + \frac{b-a}{2}t\right).$$

از این رو، خواهیم داشت

$$\int_a^b f(x)dx \approx \frac{b-a}{2} \sum_{k=0}^n w_k g(x_k) = \frac{b-a}{2} \sum_{k=0}^n w_k f\left(\frac{b+a}{2} + \frac{b-a}{2}x_k\right).$$

مثال ۶.۴ داریم

$$I = \int_0^2 x \sqrt{8-x^3} dx = \frac{16\pi}{9\sqrt{3}} = 3.224532\dots$$

این انتگرال را با استفاده از فرمول سه نقطه‌ای انتگرال گیری گاوس-لژاندر تقریب می‌زنیم. برای این کار، قرار می‌دهیم $a = 0$ و $b = 2$. اکنون، با استفاده از تغییر متغیر

$$x = \frac{b+a}{2} + \frac{b-a}{2}t = 1+t,$$

خواهیم داشت

$$I = \int_{-1}^1 f(1+t)dt,$$

که در آن $f(x) = x\sqrt{8-x^3}$. بنابراین، با کمک فرمول سه نقطه‌ای انتگرال گیری گاوس-لژاندر، داریم

$$\begin{aligned} I &= \int_{-1}^1 f(1+t)dt \approx \frac{1}{9} \left(5f\left(1 - \sqrt{\frac{3}{5}}\right) + 8f(1) + 5f\left(1 + \sqrt{\frac{3}{5}}\right) \right) \\ &= 3.272779. \end{aligned}$$

چنان‌که مشاهده می‌شود، جواب به دست آمده، از دقت زیادی برخوردار نیست. برای محاسبه تقریبی انتگرال با یک دقت مناسب، می‌توان از ایده فرمول‌های مرکب برای هر یک از روش‌های انتگرال گیری گاوس استفاده کرد.

تمرین

۱.۴) مقادیر α_i ، $i = 1, 2, 3$ را طوری محاسبه کنید که فرمول انتگرال گیری

$$\int_{-1}^1 w(x)f(x)dx \approx \alpha_1 f(-1) + \alpha_2 f(0) + \alpha_3 f(1),$$

برای هر چند جمله‌ای از درجه حداکثر ۳ دقیق باشد. مسئله را برای توابع وزن

$$w(x) = 1, \quad w(x) = \sqrt{1-x^2}, \quad w(x) = \frac{1}{1-x^2},$$

حل کنید.

(۲.۴) فرض کنید $a \leq x_0 < x_1 < x_2 < \dots < x_n \leq b$ یک افراز دلخواه و ثابت از بازه $[a, b]$ باشد. نشان دهید مقادیر یکتای $\gamma_0, \gamma_1, \dots, \gamma_n$ وجود دارند به طوری که به ازای هر چند جمله‌ای از درجه حداکثر n ، مثل p ، داریم

$$\int_a^b p(x) dx = \sum_{i=0}^n \gamma_i p(x_i).$$

(۳.۴) به کمک فرمول مجموع اولر-مکلوران (۶.۴) و انتخاب یک تابع مناسب، مقدار $\sum_{k=1}^n k^2$ را محاسبه کنید.

(۴.۴) با استفاده نمایش خطای پئانو، خطای دستور دوزنقه‌ای اصلاح شده را به دست آورید.

(۵.۴) فرض کنید $h = (b-a)/4$ و $x_i = a + ih$ ، $i = 0, 1, 2, 3$ ، چهار نقطه متساوی الفاصله در بازه $[a, b]$ باشند. در این صورت، فرمول انتگرال‌گیری زیر را ثابت کنید

$$\int_a^b f(x) dx = \frac{3}{8} h (f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)) - \frac{3h^5}{80} f^{(4)}(\alpha), \quad \alpha \in [a, b].$$

(۶.۴) فرمول انتگرال‌گیری زیر را در نظر بگیرید

$$\int_0^1 f(x) dx \approx \frac{1}{4} (f(0) + f(1)).$$

(الف) با استفاده از قضیه نمایش خطای پئانو، عبارت خطای این فرمول را به دست آورید.

(ب) با استفاده از این قاعده انتگرال‌گیری، فرمولی برای $\int_a^b f(x) dx$ به دست آورید و عبارت خطای آن را مشخص کنید.

(۷.۴) (الف) فرض کنید $h > 0$. نقاط $x_0, x_1 = x_0 + h$ و $x_2 = x_1 + 2h$ را در نظر بگیرید. چند جمله‌ای درونیاب تابع f را در نقاط x_1 و x_2 نوشته و p بنامید. نشان دهید

$$\int_{x_0}^{x_2} f(x) dx \approx \int_{x_0}^{x_2} p(x) dx = \frac{1}{4} (9f(x_1) + 3f(x_2)).$$

(ب) عبارت خطای این فرمول انتگرال‌گیری را با استفاده از قضیه نمایش خطای پئانو به دست آورید.

(ج) بدون استفاده از قضیه نمایش خطای پئانو، عبارت خطا را به دست آورید.

۸.۴ رابطه (۲۱.۴) را ثابت کنید. راهنمایی: با فرض $t = e^{i\frac{m\pi}{(n+1)}}$ مجموع خواسته شده را به صورت $\sum_{k=0}^n t^{k+1}$ بنویسید، که در آن i واحد موهومی است.

۹.۴ چندجمله‌ای‌های لاگر به صورت زیر تعریف می‌شوند

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x}), \quad n = 0, 1, 2, \dots$$

می‌توان دید که این چندجمله‌ای‌ها نسبت به تابع وزن $w(x) = e^{-x}$ روی $(0, \infty)$ متعامد هستند؛ به عبارت دیگر،

$$\int_0^{\infty} w(x) L_n(x) L_m(x) dx = \begin{cases} 0, & m \neq n, \\ (n!)^2, & m = n. \end{cases}$$

فرض کنید $t_0, \dots, t_n, n+1$ عدد حقیقی متمایز هستند. با ذکر دلیل بیان کنید که چرا ماتریس

$$A = \begin{pmatrix} L_0(t_0) & L_0(t_1) & \dots & L_0(t_n) \\ L_1(t_0) & L_1(t_1) & \dots & L_1(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ L_n(t_0) & L_n(t_1) & \dots & L_n(t_n) \end{pmatrix}$$

نامنفرد (معکوس پذیر) است. فرض کنید

$$\int_0^{\infty} w(x) f(x) dx \approx \sum_{k=0}^n w_k f(x_k)$$

یک فرمول کوادراتور گاوس باشد، که در آن x_k ها ریشه‌های چندجمله‌ای لاگر L_{n+1} هستند (فرمول کوادراتور گاوس-لاگر). نشان دهید که w_k ها جواب دستگاه معادلات خطی

$$\sum_{i=0}^n L_k(x_i) w_i = \begin{cases} 1, & k = 0, \\ 0, & k = 1, \dots, n, \end{cases}$$

هستند. با استفاده از مطالب بالا، فرمول دونقطه‌ای گاوس-لاگر را به دست آورید و عبارت خطای آن را مشخص کنید. با استفاده از فرمول به دست آمده، مقدار تقریبی انتگرال

$$\int_0^{\infty} \frac{e^{-2x}}{1+4x} dx$$

را به دست آورید.

(۱۰.۴) چندجمله‌ای‌های هرمیت به صورت

$$H_0(x) = 1, \quad H_1(x) = 2x,$$

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x), \quad n = 1, 2, \dots,$$

تعریف می‌شوند. می‌توان دید که این چندجمله‌ای‌ها نسبت به تابع وزن $w(x) = e^{-x^2}$ روی $(-\infty, \infty)$ متعامد هستند؛ به عبارت دیگر

$$\int_{-\infty}^{\infty} w(x) H_n(x) H_m(x) dx = \begin{cases} 0, & m \neq n, \\ 2^n n! \sqrt{\pi}, & m = n. \end{cases}$$

فرض کنید

$$\int_{-\infty}^{\infty} w(x) f(x) dx \approx \sum_{k=0}^n w_k f(x_k),$$

یک فرمول کوادراتور گاوس باشد، که در آن x_k ها ریشه‌های چندجمله‌ای هرمیت H_{n+1} هستند (فرمول کوادراتور گاوس-هرمیت). نشان دهید که w_k ها جواب دستگاه معادلات خطی

$$\sum_{i=0}^n H_k(x_i) w_i = \begin{cases} \sqrt{\pi}, & k = 0, \\ 0, & k = 1, \dots, n. \end{cases}$$

هستند. با استفاده از مطالب بالا، فرمول دو نقطه‌ای گاوس-هرمیت را به دست آورده و عبارت خطای آن را مشخص کنید.

(۱۱.۴) فرض کنید f و g ، دو چندجمله‌ای از درجه حداکثر n و x_k ، $k = 0, 1, 2, \dots, n$ ریشه‌های چندجمله‌ای لژاندر از درجه $n+1$ باشند. نشان دهید

$$\int_{-1}^1 f(x)g(x) dx = \sum_{k=0}^n \gamma_k f(x_k)g(x_k),$$

که در آن

$$\gamma_k = \int_{-1}^1 L_k(x) dx = \int_{-1}^1 \prod_{\substack{j=0 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j} dx, \quad k = 0, 1, \dots, n.$$

(۱۲.۴) فرض کنید $f \in C^2[0, 1]$. یک چندجمله‌ای (هرمیت) مثل P از درجه حداکثر دو بیابید به طوری که

$$\begin{cases} P(0) = f(0), & P'(0) = f'(0), \\ P(1) = f(1). \end{cases}$$

با انتگرال گیری از چند جمله‌ای P روی بازه $[0, 1]$ ، یک فرمول انتگرال گیری به صورت زیر بیابید

$$\int_0^1 f(x) dx \approx \frac{1}{3} (4f(0) + f'(0) + 2f(1)).$$

با استفاده از عبارت درونیابی هرمیت، نشان دهید خطای این فرمول به صورت زیر است

$$E(f) = -\frac{1}{72} f^{(3)}(\alpha) \quad \alpha \in (0, 1).$$

با استفاده از قضیه نمایش خطای پتانو، عبارت خطای فرمول انتگرال گیری بالا را به دست آورید. با یک تغییر متغیر مناسب، فرمول انتگرال گیری به دست آمده و خطای آن را برای $\int_a^b f(x) dx$ تعمیم دهید.

۱۳.۴ فرض کنید که w_k و x_k ها به ترتیب، وزن‌ها و گره‌ها در فرمول انتگرال گیری گاوس-لژاندر

$$\int_{-1}^1 f(x) dx \approx \sum_{k=0}^n w_k f(x_k),$$

هستند.

$$\sum_{k=0}^n w_k = 2 \quad \text{(الف) نشان دهید}$$

(ب) فرض کنید p یک چند جمله‌ای از درجه حداکثر n باشد. نشان دهید

$$\|p\|_2 \leq \sqrt{2} \max_{0 \leq k \leq n} |p(x_k)|,$$

که در آن $\|\cdot\|_2$ نرم L^2 است.

۱۴.۴ با کمک یکی از فرمول‌های انتگرال گیری، یک مقدار تقریبی برای انتگرال

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx,$$

به دست آورید. مقدار تقریبی به دست آمده را با مقدار واقعی انتگرال، یعنی π ، مقایسه کنید.

۱۵.۴ فرمول دونقطه‌ای گاوس-لژاندر را برای محاسبه تقریبی

$$\int_{-1}^1 \int_{-1}^1 f(x, y) dx dy,$$

تعمیم دهید.

۱۶.۴) با استفاده از قضیه ۵.۴، مجموعه چندجمله‌ای‌های تکین $\{P_0, P_1, P_2\}$ را بیابید به طوری که $i = 0, 1, 2, \deg(P_i) = i$ و

$$\langle P_i, P_j \rangle = \int_0^1 w(x)P_i(x)P_j(x)dx = 0, \quad i \neq j,$$

که در آن $w(x) = \sqrt{x}$ فرض کنید

$$I(f) = \int_0^1 \sqrt{x}f(x)dx,$$

که در آن f یک تابع هموار است. با کمک چندجمله‌ای P_2 ، یک فرمول انتگرال‌گیری گاوس دونقطه‌ای برای محاسبه $I(f)$ به دست آورید.

فصل ۵

حل عددی معادلات غیرخطی

حل بسیاری از مدل‌های ریاضی، به حل دستگاه معادلات غیرخطی منجر می‌شود که حل مستقیم و دقیق اغلب آنها ناممکن است. بنابراین، برای حل آنها از روش‌های تقریبی استفاده می‌شود. در این فصل، به معرفی برخی از این روش‌ها می‌پردازیم.

۱.۵ روش نیوتن - رافسون

دستگاه معادلات غیرخطی

$$f(x) = 0, \quad (1.5)$$

را در نظر بگیرید، که در آن $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ تابعی به صورت

$$f(x) = \begin{pmatrix} f_1(x^{(1)}, \dots, x^{(n)}) \\ f_2(x^{(1)}, \dots, x^{(n)}) \\ \vdots \\ f_n(x^{(1)}, \dots, x^{(n)}) \end{pmatrix},$$

است، به طوری که $x^T = (x^{(1)}, \dots, x^{(n)})^T$ در حالت کلی، نمی‌توان معادله (۱.۵) را در تعداد متناهی گام حل کرد، بنابراین از روش‌های تکراری برای حل آن استفاده می‌شود. یک روش تکراری برای حل (۱.۵)، به صورت

$$x_{i+1} = \Phi(x_i), \quad i = 0, 1, 2, \dots,$$

نوشته می‌شود، که در آن $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ تابع تکرار نامیده می‌شود و x_0 یک حدس اولیه برای ریشه تابع f است. بردار ξ را یک نقطه ثابت تابع Φ می‌گویند، هرگاه $\Phi(\xi) = \xi$. اگر هر نقطه ثابت Φ یک ریشه f و تابع Φ در یک همسایگی از هر نقطه ثابت Φ پیوسته باشد، آنگاه حد دنباله $\{x_i\}$ در صورت وجود یک نقطه ثابت Φ و در نتیجه، ریشه f خواهد بود. در این ارتباط، سه پرسش اساسی مطرح می‌شود:

(۱) چگونه می‌توان یک تابع تکرار مناسب Φ یافت؟

(۲) در چه شرایطی دنباله $\{x_i\}$ همگرا خواهد بود؟

(۳) با چه سرعتی دنباله $\{x_i\}$ همگراست؟

در آغاز، ببینیم چگونه می‌توان یک تابع تکرار مثل Φ ساخت. روش‌های مختلفی برای این کار وجود دارد. برای مثال، اگر بخواهیم معادله $x - \cos x = 0$ را حل کنیم، آنگاه می‌توان روش تکراری

$$x_{i+1} = \cos x_i, \quad i = 0, 1, 2, \dots,$$

را مورد آزمایش قرار داد. در اینجا $\Phi(x) = \cos x$ تابع تکرار است.

در ادامه، یک قاعده کلی برای یافتن تابع تکرار Φ معرفی می‌شود. فرض کنید ξ یک ریشه تابع f ، و تابع f در یک همسایگی از ξ مثل $\mathcal{N}_\epsilon(\xi) = \{x \in \mathbb{R} : |x - \xi| < \epsilon\}$ به اندازه کافی مشتق‌پذیر باشد. بسط تیلر تابع f را حول نقطه $x_0 \in \mathcal{N}_\epsilon(\xi)$ می‌نویسیم

$$\begin{aligned} 0 &= f(\xi) = f(x_0) + f'(x_0)(\xi - x_0) + f''(x_0) \frac{(\xi - x_0)^2}{2!} \\ &\quad + \dots + \frac{(\xi - x_0)^k}{k!} f^{(k)}(x_0 + \theta(\xi - x_0)), \end{aligned}$$

که در آن $0 < \theta < 1$. اگر از جمله‌ای که شامل $(\xi - x_0)^k$ است، چشم‌پوشی کنیم، یک معادله برحسب ξ از درجه $k-1$ خواهیم داشت که با حل آن مقدار تقریبی ξ به دست خواهد آمد. به ازای $k=2$ ، خواهیم داشت

$$0 \approx f(x_0) + (\xi - x_0)f'(x_0),$$

که با حل آن داریم (در صورتی که $f'(x_0) \neq 0$)

$$\xi \approx x_0 - \frac{f(x_0)}{f'(x_0)}.$$

با فرض

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

داریم $\xi \approx x_1$. ممکن است x_1 دقت کافی را نداشته باشد. بنابراین، می‌توان فرایند بالا را با استفاده از رابطه بازگشتی

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

تکرار کرد تا جوابی با دقت کافی به دست آید. این روند تکراری به روش نیوتن-رافسون معروف است. این رابطه را می‌توان به صورت

$$x_{i+1} = \Phi(x_i), \quad \Phi(x) = x - \frac{f(x)}{f'(x)},$$

نوشت. در واقع، در اینجا $\Phi(x)$ تابع تکرار روش نیوتن-رافسون است. می‌توان روش نیوتن-رافسون را برای حل دستگاه معادلات غیرخطی تعمیم داد. دستگاه معادلات خطی زیر را در نظر بگیرید

$$f(x) = \begin{pmatrix} f_1(x^{(1)}, \dots, x^{(n)}) \\ f_2(x^{(1)}, \dots, x^{(n)}) \\ \vdots \\ f_n(x^{(1)}, \dots, x^{(n)}) \end{pmatrix} = \circ.$$

فرض کنید ξ یک ریشه تابع f باشد. در این صورت، اگر $x_0 = (x_0^{(1)}, \dots, x_0^{(n)})^T$ یک تقریب برای $\xi = (\xi^{(1)}, \dots, \xi^{(n)})^T$ باشد، همانند قبل داریم

$$\circ = f(\xi) = f(x_0) + Df(x_0)(\xi - x_0) + \frac{1}{2}(\xi - x_0)^T D^2 f(x_0)(\xi - x_0) + \dots, \quad (2.5)$$

که در آن $Df(x_0) = (g_{ij}) \in \mathbb{R}^{n \times n}$ به طوری که

$$g_{ij} = \frac{\partial f_i}{\partial x^{(j)}}(x_0), \quad i, j = 1, 2, \dots, n.$$

و $D^2 f(x_0) = (h_{ij}) \in \mathbb{R}^{n \times n}$ به طوری که

$$h_{ij} = \frac{\partial^2 f_i}{\partial x^{(i)} \partial x^{(j)}}(x_0), \quad i, j = 1, 2, \dots, n.$$

با چشم‌پوشی از جمله سوم به بعد در رابطه (۲.۵)، داریم

$$\circ \approx f(x_0) + Df(x_0)(\xi - x_0),$$

و در صورتی که ماتریس $Df(x_0)$ نامنفرد باشد، خواهیم داشت

$$\xi \approx x_0 - (Df(x_0))^{-1} f(x_0).$$

بنابراین، اگر قرار دهیم

$$x_1 = x_0 - (Df(x_0))^{-1} f(x_0),$$

آنگاه $x_1 \approx \xi$. ولی ممکن است دقت x_1 به اندازه کافی نباشد. از این رو، می‌توانیم روند بالا را تکرار کنیم تا یک تقریب مناسب به دست آید. در این صورت، خواهیم داشت

$$x_{i+1} = x_i - (Df(x_i))^{-1} f(x_i), \quad i = 0, 1, 2, \dots \quad (3.5)$$

به ازای مقدار کوچک داده شده $\epsilon > 0$ ، می‌توان از هریک از شرایط

$$\|x_{m+1} - x_m\| < \epsilon, \quad (4.5)$$

$$\|f(x_{m+1})\| < \epsilon, \quad (5.5)$$

$$\|x_{m+1} - x_m\| + \|f(x_{m+1})\| < \epsilon, \quad (6.5)$$

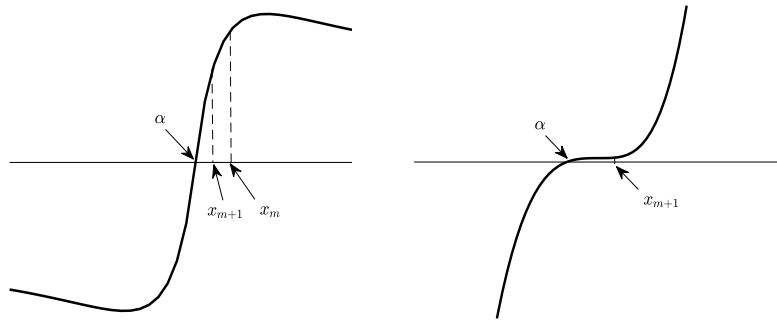
به عنوان یک معیار توقف استفاده کرد و x_{m+1} را یک جواب تقریبی در نظر گرفت. در اینجا، $\|\cdot\|$ یک نرم برداری دلخواه در \mathbb{R}^n است. ممکن است شرایط (۴.۵) و (۵.۵)، جوابی با دقت کافی به دست ندهند (البته این اتفاق به ندرت رخ می‌دهد)، ولی شرط (۶.۵) همیشه جواب مناسبی به دست خواهد داد. شکل ۱.۵ به خوبی دلیل اینکه ممکن است شرایط (۴.۵) و (۵.۵) جواب مناسبی به دست ندهند را نشان می‌دهد. نمودار سمت چپ نشان می‌دهد با اینکه فاصله بین x_m و x_{m+1} ناچیز است، مقدار x_{m+1} به اندازه کافی به α نزدیک نیست. همچنین تصویر سمت راست نشان می‌دهد با اینکه $|f(x_{m+1})|$ کوچک است، مقدار x_{m+1} به α نزدیک نیست. شایان ذکر است که برای به دست آوردن یک جواب تقریبی که در شرط سوم صدق کند، تعداد تکرارهای بیشتری نسبت به دو شرط دیگر نیاز است.

مثال ۱.۵ دستگاه معادلات غیرخطی

$$\begin{cases} f_1(x^{(1)}, x^{(2)}) = (x^{(1)})^2 - (x^{(2)})^2 + x^{(1)}x^{(2)} - 5 = 0, \\ f_2(x^{(1)}, x^{(2)}) = 2(x^{(1)})^2 + (x^{(2)})^2 - 9 = 0, \end{cases}$$

را در نظر بگیرید. می‌خواهیم یکی از جواب‌های این دستگاه را با استفاده از روش نیوتن-رافسون به صورت تقریبی محاسبه کنیم. برای سادگی، فرض کنید $r = x^{(1)}$ و $s = x^{(2)}$ ، $x = (r, s)^T$. در این صورت، معادله بالا را می‌توان به صورت

$$\begin{cases} f_1(r, s) = 4r^2 - s^2 + 2rs - 1 = 0, \\ f_2(r, s) = 4r^2 + s^2 + s - 3 = 0, \end{cases}$$



شکل ۱.۵: حالت‌هایی که ممکن است معیار توقف جواب مناسبی فراهم نکند.

نوشت. داریم

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial s} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial s} \end{pmatrix} = \begin{pmatrix} 18r + 2s & 2r - 2s \\ 18r & 2s + 1 \end{pmatrix}.$$

با حدس اولیه $x_0 = (r_0, s_0)^T = (1, 2)^T$ شروع می‌کنیم. در تکرار اول خواهیم داشت

$$\begin{aligned} x_1 &= x_0 - (Df(x_0))^{-1} f(x_0) \\ &= \begin{pmatrix} r_0 \\ s_0 \end{pmatrix} - \begin{pmatrix} 18r_0 + 2s_0 & 2r_0 - 2s_0 \\ 18r_0 & 2s_0 + 1 \end{pmatrix}^{-1} \begin{pmatrix} 4r_0^2 - s_0^2 + 2r_0s_0 - 1 \\ 4r_0^2 + s_0^2 + s_0 - 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 12 & -2 \\ 18 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 7 \end{pmatrix} \\ &= \begin{pmatrix} 0.6184 \\ 1.2105 \end{pmatrix}. \end{aligned}$$

بنابراین، $x_1 = (r_1, s_1)^T = (0.6184, 1.2105)^T$ سپس، در تکرار دوم خواهیم داشت

$$\begin{aligned} x_2 &= x_1 - (Df(x_1))^{-1} f(x_1) \\ &= \begin{pmatrix} r_1 \\ s_1 \end{pmatrix} - \begin{pmatrix} 18r_1 + 2s_1 & 2r_1 - 2s_1 \\ 18r_1 & 2s_1 + 1 \end{pmatrix}^{-1} \begin{pmatrix} 4r_1^2 - s_1^2 + 2r_1s_1 - 1 \\ 4r_1^2 + s_1^2 + s_1 - 3 \end{pmatrix} \\ &= \begin{pmatrix} 0.6184 \\ 1.2105 \end{pmatrix} - \begin{pmatrix} 7.3684 & -1.1842 \\ 4.9474 & 3.4211 \end{pmatrix}^{-1} \begin{pmatrix} 0.5616 \\ 1.2057 \end{pmatrix} \\ &= \begin{pmatrix} 0.5106 \\ 1.0140 \end{pmatrix}. \end{aligned}$$

به همین ترتیب، در تکرار سوم خواهیم داشت $x_3 = (r_3, s_3)^T = (0.5001, 1.0001)^T$ توجه کنید که

$$\left\| \begin{pmatrix} f_1(r_3, s_3) \\ f_2(r_3, s_3) \end{pmatrix} \right\|_2 = 8.3487 \times 10^{-4}.$$

چنان که می‌بینیم، روش تکراری نیوتن-رافسون در سه تکرار، یک تقریب مناسب فراهم کرده است. شایان ذکر است که $x^* = (0.5, 1)^T$ یکی از جواب‌های واقعی دستگاه است.

مثال ۲.۵ معادله دیفرانسیل غیرخطی

$$\begin{cases} u''(x) + e^{u(x)} = 0, & 0 < x < 1, \\ u(0) = u(1) = 0, \end{cases}$$

را در نظر بگیرید. بازه $[0, 1]$ را به $(n+1)$ قسمت مساوی تقسیم می‌کنیم و قرار می‌دهیم $h = 1/(n+1)$. با نوشتن معادله دیفرانسیل در نقاط گرهی $x^{(i)} = ih$ ، $i = 1, 2, \dots, n$ خواهیم داشت

$$u''(x^{(i)}) + e^{u(x^{(i)})} = 0, \quad i = 1, 2, \dots, n.$$

با استفاده از فرمول تفاضل مرکزی (۶.۱) برای مشتق دوم، داریم

$$\frac{u(x^{(i-1)}) - 2u(x^{(i)}) + u(x^{(i+1)})}{h^2} + \mathcal{O}(h^2) + e^{u(x^{(i)})} = 0, \quad i = 1, 2, \dots, n.$$

در رابطه اخیر از $\mathcal{O}(h^2)$ چشم‌پوشی می‌کنیم و $u^{(j)}$ را یک تقریب برای مقدار $u(x^{(j)})$ در نظر می‌گیریم ($u^{(j)} \approx u(x^{(j)})$ ، $j = 1, 2, \dots, n$). در این صورت، دستگاه معادلات غیرخطی

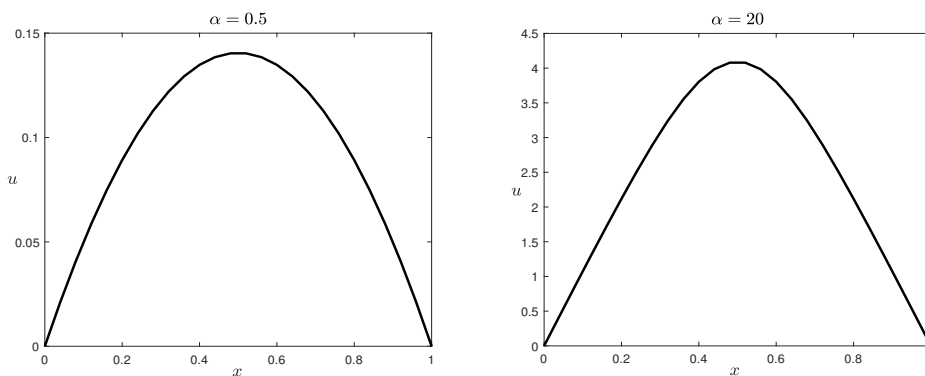
$$\frac{u^{(i-1)} - 2u^{(i)} + u^{(i+1)}}{h^2} + e^{u^{(i)}} = 0, \quad i = 1, 2, \dots, n.$$

به دست می‌آید. با توجه به اینکه

$$\begin{aligned} u^{(0)} &= u(x^{(0)}) = u(0) = 0, \\ u^{(n+1)} &= u(x^{(n+1)}) = u(1) = 0, \end{aligned}$$

و با فرض $u = (u^{(1)}, u^{(2)}, \dots, u^{(n)})^T$ ، این دستگاه معادل با دستگاه معادلات غیرخطی

$$f(u) = \begin{pmatrix} f_1(u) \\ f_2(u) \\ \vdots \\ f_n(u) \end{pmatrix} = 0,$$



شکل ۲.۵: نمودار مثال ۲.۵.

است، که در آن

$$f_i(u) = \frac{1}{h^2} \left(u^{(i-1)} - 2u^{(i)} + u^{(i+1)} + h^2 e^{u^{(i)}} \right).$$

در این صورت، خواهیم داشت

$$Df(u) = \frac{1}{h^2} \begin{pmatrix} -2 + h^2 e^{u^{(1)}} & 1 & & & & \\ 1 & -2 + h^2 e^{u^{(2)}} & 1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 + h^2 e^{u^{(n-1)}} & 1 \\ & & & & 1 & -2 + h^2 e^{u^{(n)}} \end{pmatrix}.$$

بنابراین، روش تکراری نیوتن - رافسون به صورت

$$u_{k+1} = u_k - (Df(u_k))^{-1} f(u_k), \quad k = 0, 1, \dots,$$

خواهد بود، که در آن $u_j = (u_j^{(1)}, u_j^{(2)}, \dots, u_j^{(n)})^T$ و u_0 یک حدس اولیه دلخواه است.

مؤلفه‌های حدس اولیه را به صورت

$$u_0^{(i)} = \alpha x^{(i)} (1 - x^{(i)}) = \alpha i h (1 - i h), \quad i = 1, 2, \dots, n,$$

در نظر می‌گیریم. به ازای $n = 24$ ($h = 0.04$) در شکل ۲.۵، سمت چپ، جواب معادله به ازای $\alpha = 0.5$ ، و سمت راست، جواب معادله به ازای $\alpha = 20$ نمایش داده شده است. با شرط توقف $\|u_{k+1} - u_k\|_\infty < 10^{-8}$ ، جواب‌های تقریبی به ترتیب در ۳ و ۶ تکرار به دست آمده‌اند.

مثال ۳.۵ ([۱۴]) قانون گاز ایده آل برای یک گاز در دمای پایین و فشار به صورت

$$PV = nRT, \quad (۷.۵)$$

است، که در آن P فشار گاز (برحسب atm)، V حجم گاز (برحسب L)، T دمای گاز (برحسب K)، n تعداد مول‌های گاز و $R = ۰.۰۸۲۰۵۷۸$ ثابت گاز مولار است. معادله واندروالس

$$\left(P + \frac{n^2 a}{V^2}\right)(V - nb) = nRT, \quad (۸.۵)$$

حالت ناپایده آل را پوشش می‌دهد که این مفروضات برقرار نیستند. می‌خواهیم با کمک معادله واندروالس، حجم یک مول از اکسیژن در دمای ۳۲°K و فشار ۱۵atm را به دست آوریم. برای گاز اکسیژن، مقادیر a و b به صورت $a = ۱.۳۶\text{L}^2 - \text{atm}/\text{mole}^2$ و $b = ۰.۰۰۳۱۸۳\text{L}/\text{mole}$ هستند.

تابع f را به صورت

$$\begin{aligned} f(V) &= \left(P + \frac{n^2 a}{V^2}\right)(V - nb) - nRT \\ &= \left(۱۵ + \frac{۱.۳۶}{V^2}\right)(V - ۰.۰۰۳۱۸۳) - ۲۶.۲۵۸۴۹۶, \end{aligned}$$

تعریف می‌کنیم. هدف، یافتن ریشه تابع f است. با توجه به اینکه f یک تابع غیرخطی نسبت به V است، از روش نیوتن-رافسون برای محاسبه ریشه تابع استفاده می‌کنیم. با محاسبه مشتق تابع f نسبت به V ، خواهیم داشت

$$f'(V) = ۱۵ - \frac{۱.۳۶}{V^2} + \frac{۰.۰۰۸۶۵۸}{V^3}.$$

بنابراین، تکرار روش نیوتن-رافسون به صورت زیر خواهد بود

$$V_{m+1} = V_m - \frac{f(V_m)}{f'(V_m)}, \quad m = ۰, ۱, \dots$$

با کمک معادله حاکم در حالت ایده آل، یعنی (۷.۵)، حدس اولیه V_0 را محاسبه می‌کنیم. در واقع، داریم

$$V_0 = \frac{nRT}{P} = \frac{۱ \times ۰.۰۸۲۰۵۸ \times ۳۲}{۱۵} = ۱.۷۵۰۵۶۶.$$

با شروع از V_0 و معیار توقف ۱۰^{-5} ، مقادیر زیر به دست می‌آید

$$V_1 = ۱.۷۰۰۵۷۷, \quad V_2 = ۱.۷۰۰۵۳۳, \quad V_3 = ۱.۷۰۰۵۳۳.$$

بنابراین، حجم یک مول اکسیژن در دمای ۳۲°K و فشار ۱۵atm ، تقریباً برابر با $V_3 = ۱.۷۰۰۵۳۳$ خواهد بود.

۲.۵ قضایای همگرایی در حالت کلی

در این بخش، همگرایی دنباله $\{x_i\}$ که توسط تابع تکرار Φ به صورت

$$x_{i+1} = \Phi(x_i), \quad i = 0, 1, 2, \dots,$$

تولید شده است را در یک همسایگی از نقطه ثابت ξ مطالعه می‌کنیم. ابتدا تعریف همگرایی یک دنباله در \mathbb{R}^n را یادآوری می‌کنیم.

تعریف ۱.۵ دنباله $\{x_i\}$ در \mathbb{R}^n را به x همگرا می‌گویند، هرگاه به ازای هر $\epsilon > 0$ ، یک عدد طبیعی مثل $N(\epsilon)$ وجود داشته باشد به طوری که

$$\forall l \geq N(\epsilon) : \|x_l - x\| < \epsilon,$$

که در آن $\|\cdot\|$ یک نرم در \mathbb{R}^n است.

یک نکته قابل توجه در اینجا این است که فضای \mathbb{R}^n کامل است، یعنی اینکه هر دنباله کوشی در \mathbb{R}^n همگرا است. این موضوع به طور دقیق‌تر توسط قضیه بعدی بیان می‌شود.

قضیه ۱.۵ دنباله $\{x_i\}$ در \mathbb{R}^n همگراست اگر و تنها اگر به ازای هر $\epsilon > 0$ ، یک عدد طبیعی مثل $N(\epsilon)$ وجود داشته باشد به طوری که برای هر $l, m \geq N(\epsilon)$ داریم $\|x_l - x_m\| < \epsilon$.

تعریف ۲.۵ فرض کنید $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع تکرار، ξ یک نقطه ثابت Φ و $\mathcal{N}(\xi)$ یک همسایگی ξ باشد. اگر به ازای هر $x_0 \in \mathcal{N}(\xi)$ ، دنباله تولیدشده $x_{i+1} = \Phi(x_i)$ ، $i = 0, 1, 2, \dots$ به ازای هر $i \geq 0$ در رابطه

$$\|x_{i+1} - \xi\| \leq C \|x_i - \xi\|^p,$$

صدق کند (به ازای $p = 1$ ، باید $C < 1$) آنگاه گفته می‌شود که روش تکراری تولیدشده توسط Φ ، روشی حداقل از مرتبه p برای تعیین ξ است.

قضیه ۲.۵ هر روش حداقل از مرتبه p برای محاسبه نقطه ثابت ξ ، به طور موضعی همگراست؛ به این معنا که یک همسایگی مثل $\mathcal{N}'(\xi)$ با این خاصیت وجود دارد به طوری که به ازای هر $x_0 \in \mathcal{N}'(\xi)$ ، دنباله تولیدشده توسط Φ به ξ همگرا است.

برهان: نخست، حالت $p \neq 1$ را در نظر می‌گیریم. فرض کنید $\mathcal{N}_r(\xi) = \{x : \|x - \xi\| < r\}$ یک همسایگی از ξ باشد که در تعریف (۲.۵) آمده است. قرار می‌دهیم

$$r' = \min\left\{r, \frac{1}{C^{\frac{1}{1-p}}}\right\}.$$

اکنون، فرض کنید $x_0 \in \mathcal{N}_{r'}(\xi) \subseteq \mathcal{N}_r(\xi)$. در این صورت، داریم

$$\begin{aligned} \|x_n - \xi\| &\leq C \|x_{n-1} - \xi\|^p \leq C^{1+p} \|x_{n-2} - \xi\|^{p^2} \\ &\leq \dots \leq C^{1+p+p^2+\dots+p^{n-1}} \|x_0 - \xi\|^{p^n} \\ &= C^{\frac{1-p^n}{1-p}} \|x_0 - \xi\|^{p^n} = C^{\frac{1}{1-p} - \frac{p^n}{1-p}} \|x_0 - \xi\|^{p^n} \\ &\leq C^{\frac{1}{1-p} - \frac{p^n}{1-p}} \left(\frac{1}{p} C^{\frac{1}{1-p}}\right)^{p^n} \\ &= C^{\frac{1}{1-p}} \frac{1}{p^{p^n}}. \end{aligned}$$

این رابطه نشان می‌دهد که $x_n \rightarrow \xi$.

اینک، حالت $p = 1$ را در نظر می‌گیریم. همانند حالت $p \neq 1$ ، برای هر $x_0 \in \mathcal{N}_{r'}(\xi)$ ، داریم

$$\|x_n - \xi\| \leq C^n \|x_0 - \xi\|.$$

بنابراین، با توجه به $C < 1$ ، همگرایی دنباله نتیجه می‌شود. \square

در قضیه ۲.۵، اگر بتوان $\mathcal{N}(\epsilon)$ را مجموعه \mathbb{R}^n در نظر گرفت، آنگاه همگرایی دنباله، سراسری نامیده می‌شود.

مثال ۴.۵ تابع تکرار

$$\Phi(x) = \frac{1}{2} \left(x + \frac{2}{x}\right),$$

را در نظر بگیرید. $\xi = \sqrt{2}$ یک نقطه ثابت Φ است. به سادگی می‌توان دید که اگر $x_{i+1} = \Phi(x_i)$ ، آنگاه

$$|x_{i+1} - \sqrt{2}| \leq \frac{1}{2x_i} |x_i - \sqrt{2}|^2. \quad (9.5)$$

به ازای هر $x > 0$ ، داریم

$$\Phi(x) = \frac{1}{2} \left(x + \frac{2}{x}\right) \geq \sqrt{2}.$$

بنابراین، اگر $x_0 > 0$ ، آنگاه به ازای هر $i \geq 1$ ، $x_i \geq 2$. در نتیجه

$$\frac{1}{2x_i} \leq \frac{1}{2\sqrt{2}}.$$

از این رو، با توجه به رابطه (۹.۵)، خواهیم داشت

$$|x_{i+1} - \sqrt{2}| \leq \frac{1}{2\sqrt{2}} |x_i - \sqrt{2}|^2,$$

و بنابراین، روش تکراری $x_{i+1} = \Psi(x_i)$ برای تعیین نقطه ثابت $\xi = \sqrt{2}$ ، از مرتبه ۲ است و بنا به قضیه ۲.۵ به طور موضعی همگراست.

در حالت یک بعدی، یعنی وقتی که $n = 1$ ، اگر تابع تکرار Φ به اندازه کافی هموار باشد، آنگاه به سادگی می توان مرتبه تکرار روش را در یک همسایگی $\mathcal{N}(\xi)$ که در قضیه ۲.۵ مشخص می شود، محاسبه کرد. فرض کنید $x_i \in \mathcal{N}(\xi)$

$$\Phi^{(k)}(\xi) = 0, \quad k = 0, 1, \dots, p-1,$$

و $\Phi^{(p)}(\xi) \neq 0$. در این صورت، با کمک بسط تیلر، داریم

$$x_{i+1} = \Phi(x_i) = \Phi(\xi) + \frac{(x_i - \xi)^p}{p!} \Phi^{(p)}(\xi + \theta(x_i - \xi)),$$

که در آن $0 < \theta < 1$. بنابراین

$$\lim_{i \rightarrow \infty} \frac{x_{i+1} - \alpha}{(x_i - \alpha)^p} = \frac{\Phi^{(p)}(\xi)}{p!}.$$

این رابطه نشان می دهد که برای هر $p \geq 2$ ، مرتبه روش دقیقاً p است (چرا؟) و برای $p = 1$ ، اگر $|\Phi'(\xi)| < 1$ ، آنگاه مرتبه همگرایی یک خواهد بود. شکل های ۳.۵ و ۴.۵ همگرایی دنباله را وقتی که $p = 1$ و $-1 < \Phi'(\xi) < 1$ ، نشان می دهد.

مثال ۵.۵ همان گونه که دیدیم، تابع تکرار روش نیوتن-رافسون برای حل $f(x) = 0$ به صورت

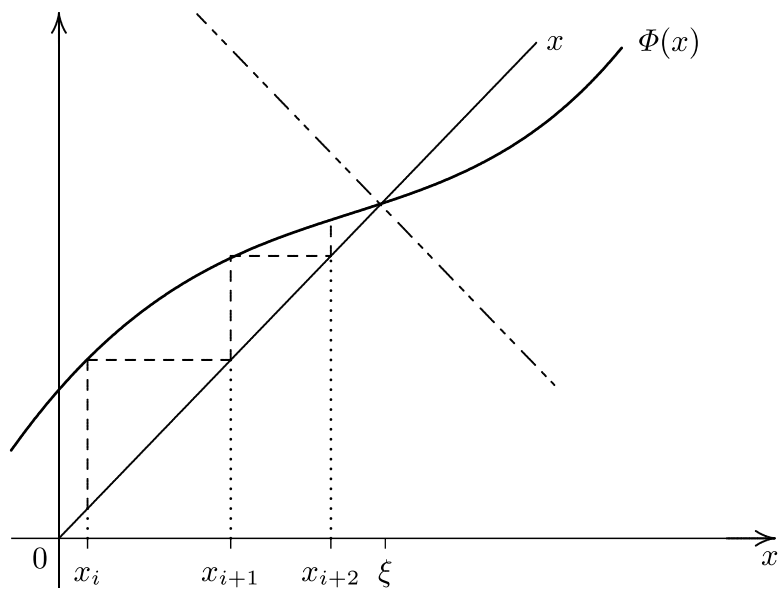
$$\Phi(x) = x - \frac{f(x)}{f'(x)},$$

است، که در آن $f: \mathbb{R} \rightarrow \mathbb{R}$. فرض کنید ξ یک ریشه ساده f است و توابع f ، f' و f'' در یک همسایگی از ξ پیوسته هستند. در این صورت، توجه کنید که $f'(\xi) \neq 0$ و داریم

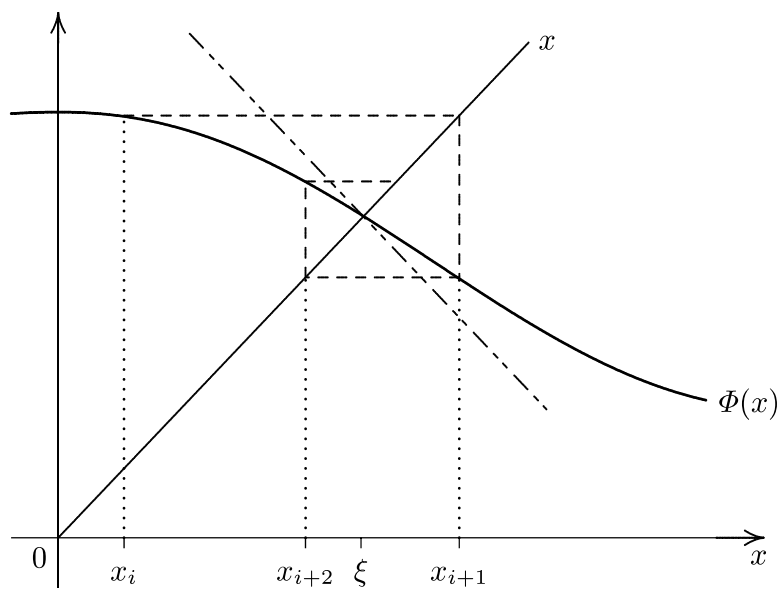
$$\Phi(\xi) = \xi, \quad \Phi'(\xi) = \frac{f(x)f''(x)}{(f'(x))^2} \Big|_{x=\xi} = 0, \quad \Phi''(\xi) = \frac{f''(\xi)}{f'(\xi)}.$$

اکنون، اگر $f''(\xi) \neq 0$ ، آنگاه $\Phi''(\xi) \neq 0$ و مرتبه همگرایی روش نیوتن-رافسون دقیقاً ۲ و در غیر این صورت، مرتبه همگرایی روش بزرگتر از ۲ خواهد بود. بنابراین، به طور کلی می توان گفت اگر ξ یک ریشه ساده f باشد، آنگاه مرتبه همگرایی روش نیوتن-رافسون حداقل برابر با ۲ خواهد بود.

تعریف ۳.۵ گفته می شود ξ یک ریشه تکراری با مرتبه تکرار m از تابع $f: \mathbb{R} \rightarrow \mathbb{R}$ است، هرگاه تابعی مثل $g: \mathbb{R} \rightarrow \mathbb{R}$ وجود داشته باشد به گونه ای که $f(x) = (x - \xi)^m g(x)$ و $g(\xi) \neq 0$.



شکل ۳.۵: نمایش همگرایی دنباله، به ازای $p = 1$ ، زمانی که $0 < \Phi'(\xi) < 1$.



شکل ۴.۵: نمایش همگرایی دنباله، به ازای $p = 1$ ، زمانی که $-1 < \Phi'(\xi) < 0$.

مثال ۶.۵ تابع $f(x) = x - \sin x$ را در نظر بگیرید. داریم

$$f(0) = f'(0) = f''(0) = 0, \quad f^{(3)}(0) \neq 0.$$

بنابراین، $\xi = 0$ یک ریشه تکراری f با مرتبه تکرار $m = 3$ است. توجه کنید که داریم

$$f(x) = x - \sin x = x - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots\right) = x^3 g(x),$$

که در آن

$$g(x) = \left(\frac{1}{3!} - \frac{x^2}{5!} + \frac{x^4}{7!} - \dots\right).$$

ملاحظه می‌شود که $g(0) = \frac{1}{3!} \neq 0$.

به سادگی می‌توان دید که اگر تابع f به اندازه کافی مشتق‌پذیر باشد، آنگاه ξ یک ریشه تابع f با مرتبه تکرار m است اگر و تنها اگر $f^{(k)}(\xi) = 0$ ، $k = 0, 1, \dots, m-1$ و $f^{(m)}(\xi) \neq 0$.

مثال ۷.۵ در مثال ۵.۵، فرض کنید ξ یک ریشه تکراری تابع f با مرتبه تکرار m است. در این صورت، داریم

$$f(x) = (x - \xi)^m g(x), \quad g(\xi) \neq 0,$$

$$f'(x) = m(x - \xi)^{m-1} g(x) + (x - \xi)^m g'(x),$$

که در آن g یک تابع مشتق‌پذیر است. از این رو، داریم

$$\Phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x - \xi)^m g(x)}{m(x - \xi)^{m-1} g(x) + (x - \xi)^m g'(x)},$$

با محاسبه مشتق تابع Φ در ξ ، خواهیم داشت

$$\Phi'(\xi) = 1 - \frac{1}{m}.$$

این رابطه نشان می‌دهد که اگر $m > 1$ ، آنگاه $\Phi'(\xi) \neq 0$ و بنابراین، مرتبه همگرایی روش نیوتن دقیقاً ۱ است.

قضیه ۳.۵ فرض کنید $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع تکرار و ξ یک نقطه ثابت آن است. همچنین، فرض کنید $S_r(\xi) = \{x: \|x - \xi\| < r\}$ یک همسایگی از ξ باشد. اگر Φ یک نگاشت انقباضی در $S_r(\xi)$ باشد، یعنی به ازای هر $x, y \in S_r(\xi)$ داشته باشیم

$$\|\Phi(x) - \Phi(y)\| \leq K \|x - y\|,$$

که در آن $0 < K < 1$ ، آنگاه به ازای هر $x_0 \in S_r(\xi)$ ، دنباله $\{x_i\}$ که توسط $x_{i+1} = \Phi(x_i)$ ، $i = 0, 1, \dots$ تولید می‌شود، دارای خواص زیر است

(الف) به ازای هر $i \geq 0$ ، داریم $x_i \in S_r(\xi)$ ؛

(ب) $\|x_{i+1} - \xi\| \leq K\|x_i - \xi\| \leq K^{i+1}\|x_0 - \xi\|$ ، یعنی دنباله $\{x_i\}$ به ξ همگراست.

برهان: به استقرا، قسمت (الف) را ثابت می‌کنیم. به ازای $i = 0$ ، طبق فرض، داریم $x_0 \in S_r(\xi)$. اگر $x_i \in S_r(\xi)$ ، آنگاه داریم

$$\|x_{i+1} - \xi\| = \|\Phi(x_i) - \Phi(\xi)\| \leq K\|x_i - \xi\| < \|x_i - \xi\| < r,$$

که نشان می‌دهد $x_{i+1} \in S_r(\xi)$.

برای اثبات (ب)، داریم

$$\|x_{i+1} - \xi\| = \|\Phi(x_i) - \Phi(\xi)\| \leq K\|x_i - \xi\|.$$

بنابراین

$$\|x_{i+1} - \xi\| \leq K\|x_i - \xi\| \leq K^2\|x_{i-1} - \xi\| \leq \dots \leq K^{i+1}\|x_0 - \xi\|.$$

□ از این رو، با توجه به اینکه $\lim_{i \rightarrow \infty} K^{i+1} = 0$ ، نتیجه می‌گیریم که $\{x_i\}$ به ξ همگراست.

قضیه ۴.۵ فرض کنید $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع تکرار، $x_0 \in \mathbb{R}^n$ یک نقطه شروع و

$$x_{i+1} = \Phi(x_i), \quad i = 0, 1, 2, \dots$$

همچنین، فرض کنید یک همسایگی از x_0 مثل $S_r(x_0) = \{x : \|x - x_0\| < r\}$ و یک ثابت K با شرط $0 < K < 1$ وجود داشته باشد به طوری که

$$(a) \|\Phi(x) - \Phi(y)\| \leq K\|x - y\|, \quad \forall x, y \in \overline{S_r(x_0)} := \{x : \|x - x_0\| \leq r\}$$

$$(b) \|x_1 - x_0\| = \|\Phi(x_0) - x_0\| \leq (1 - K)r < r.$$

در این صورت

(الف) به ازای هر $i = 0, 1, 2, \dots$ ، داریم $x_i \in S_r(x_0)$ ؛

(ب) Φ دقیقاً یک نقطه ثابت ξ در $\overline{S_r(x_0)}$ دارد و

$$\lim_{i \rightarrow \infty} x_i = \xi, \quad \|x_{i+1} - \xi\| \leq K\|x_i - \xi\|,$$

و علاوه بر آن

$$\|x_i - \xi\| \leq \frac{K^i}{1 - K}\|x_1 - x_0\|.$$

برهان: اثبات (الف) به استقرا روی i انجام می‌شود. به‌ازای $i = 1$ ، با توجه به (b)، می‌بینیم که $x_1 \in S_r(x_0)$. اکنون، فرض کنید $x_j \in S_r(x_0)$ ، $j = 0, 1, \dots, i$. در این صورت، با توجه به (a)، داریم

$$\|x_{i+1} - x_i\| = \|\Phi(x_i) - \Phi(x_{i-1})\| \leq K\|x_i - x_{i-1}\| \leq \dots \leq K^i\|x_1 - x_0\|.$$

از این رو، با استفاده از این رابطه و (b)، داریم

$$\begin{aligned} \|x_{i+1} - x_0\| &\leq \|x_{i+1} - x_i\| + \|x_i - x_{i-1}\| + \dots + \|x_1 - x_0\| \\ &\leq (K^i + K^{i-1} + \dots + 1)\|x_1 - x_0\| \\ &\leq (1 + K + \dots + K^i)(1 - K)r \\ &= (1 - K^{i+1})r < r, \end{aligned}$$

که نشان می‌دهد $x_{i+1} \in S_r(x_0)$.

برای اثبات قسمت (ب)، نخست نشان می‌دهیم $\{x_i\}$ یک دنباله کوشی است. فرض کنید $m > \ell$. در این صورت، داریم

$$\begin{aligned} \|x_m - x_\ell\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \dots + \|x_{\ell+1} - x_\ell\| \\ &\leq (K^{m-1} + K^{m-2} + \dots + K^\ell)\|x_1 - x_0\| \\ &\leq K^\ell(1 + K + \dots + K^{m-\ell-1})\|x_1 - x_0\| \\ &= \frac{K^\ell(1 - K^{m-\ell})}{1 - K}\|x_1 - x_0\| \\ &< K^\ell(1 - K^{m-\ell})r \leq K^\ell r. \end{aligned}$$

با توجه به اینکه $0 < K < 1$ ، به‌ازای هر $\epsilon > 0$ یک $N(\epsilon)$ (به اندازه کافی بزرگ) وجود دارد به‌طوری که برای هر $\ell > N(\epsilon)$ ، داریم $K^\ell r < \epsilon$. بنابراین، $\{x_i\}$ یک دنباله کوشی است و با توجه به کامل بودن \mathbb{R}^n ، نتیجه می‌گیریم که این دنباله همگرا است. فرض کنید

$$\lim_{i \rightarrow \infty} x_i = \xi.$$

با توجه به اینکه به‌ازای هر i ، $x_i \in S_r(x_0)$ ، نتیجه می‌شود که $\xi \in \overline{S_r(x_0)}$. اکنون، داریم

$$\begin{aligned} \|\Phi(\xi) - \xi\| &= \|\Phi(\xi) - \Phi(x_i) + \Phi(x_i) - \xi\| \\ &\leq \|\Phi(\xi) - \Phi(x_i)\| + \|\Phi(x_i) - \xi\| \\ &\leq K\|\xi - x_i\| + \|x_{i+1} - \xi\|. \end{aligned}$$

در نتیجه، با توجه به اینکه

$$\lim_{i \rightarrow \infty} \|\xi - x_i\| = \lim_{i \rightarrow \infty} \|x_{i+1} - \xi\| = 0,$$

خواهیم داشت $\|\Phi(\xi) - \xi\| = 0$ و بنابراین $\Phi(\xi) = \xi$.

برای نشان دادن یکتایی ξ ، به برهان خلف، فرض کنید $\hat{\xi}$ نقطه ثابت دیگری برای تابع Φ در $\overline{S_r(x_0)}$ باشد. در نتیجه، داریم

$$\|\xi - \hat{\xi}\| = \|\Phi(\xi) - \Phi(\hat{\xi})\| \leq K\|\xi - \hat{\xi}\| < \|\xi - \hat{\xi}\|,$$

که یک تناقض است و بنابراین، نقطه ثابت Φ در $\overline{S_r(x_0)}$ یکتا است. سرانجام، دیدیم

$$\|x_m - x_\ell\| \leq \frac{K^\ell}{1-K} \|x_1 - x_0\|,$$

و بنابراین

$$\lim_{m \rightarrow \infty} \|x_m - x_\ell\| \leq \frac{K^\ell}{1-K} \|x_1 - x_0\|,$$

یا

$$\|x_\ell - \xi\| \leq \frac{K^\ell}{1-K} \|x_1 - x_0\|.$$

همچنین، داریم

$$\|x_{i+1} - \xi\| = \|\Phi(x_i) - \Phi(\xi)\| \leq K\|x_i - \xi\|,$$

و به این ترتیب اثبات قضیه کامل می‌شود. \square

۳.۵ همگرایی روش نیوتن - رافسون

این بخش را با تعریف زیر شروع می‌کنیم.

تعریف ۴.۵ تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ در نقطه $x_0 \in \mathbb{R}^n$ مشتق پذیر گفته می‌شود، هرگاه یک ماتریس $n \times n$ مثل A موجود باشد به طوری که

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} = 0.$$

در این صورت، مشتق تابع برابر با A خواهد بود و با $Df(x_0)$ نشان داده می‌شود.

تعریف ۵.۵. ناحیه $C \subseteq \mathbb{R}^n$ را محدب می‌گویند، هرگاه به‌ازای هر $x, y \in C$ ، خط واصل بین x و y درون C قرار گیرد. به عبارت دیگر، مجموعه C شامل $\{z = \lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\}$ باشد.

لم ۱.۵. فرض کنید به‌ازای هر x در ناحیه محدب $C \subseteq \mathbb{R}^n$ ، $Df(x)$ موجود و یک ثابت مثل γ وجود داشته باشد به طوری که

$$\|Df(x) - Df(y)\| \leq \gamma \|x - y\|, \quad \forall x, y \in C.$$

در این صورت، به‌ازای هر $x, y \in C$

$$\|f(x) - f(y) - Df(y)(x - y)\| \leq \frac{\gamma}{2} \|x - y\|^2.$$

برهان: به‌ازای هر $x, y \in C$ ، تابع $\varphi: [0, 1] \rightarrow \mathbb{R}^n$ را به صورت $\varphi(t) = f(y + t(x - y))$ تعریف می‌کنیم. با استفاده از قاعده زنجیره‌ای مشتق، داریم

$$\varphi'(t) = Df(y + t(x - y))(x - y).$$

بنابراین

$$\begin{aligned} \|\varphi'(t) - \varphi'(0)\| &= \|(Df(y + t(x - y)) - Df(y))(x - y)\| \\ &\leq \|Df(y + t(x - y)) - Df(y)\| \|x - y\| \\ &\leq \gamma t \|x - y\|^2. \end{aligned}$$

از طرفی

$$\begin{aligned} \Delta &:= f(x) - f(y) - Df(y)(x - y) \\ &= \varphi(1) - \varphi(0) - \varphi'(0) \\ &= \int_0^1 (\varphi'(t) - \varphi'(0)) dt. \end{aligned}$$

در نتیجه

$$\|\Delta\| \leq \int_0^1 \|\varphi'(t) - \varphi'(0)\| dt \leq \gamma \|x - y\|^2 \int_0^1 t dt = \frac{\gamma}{2} \|x - y\|^2,$$

□ و به این ترتیب، اثبات لم کامل می‌شود.

قضیه ۵.۵ فرض کنید $C \subseteq \mathbb{R}^n$ یک مجموعه باز است. همچنین، فرض کنید C یک مجموعه محدب با خاصیت $\overline{C} \subseteq C$ و $f: C \rightarrow \mathbb{R}^m$ تابعی است که به ازای هر $x \in C$ مشتق پذیر و برای هر $x \in C$ پیوسته است. برای هر $x_0 \in C$ فرض کنید اعداد ثابت و مثبت r, α, β, γ و h وجود دارند به طوری که

$$\begin{aligned} S_r(x_0) &= \{x: \|x - x_0\| < r\} \subseteq C, \\ h &= \frac{\alpha\beta\gamma}{\gamma} < 1 \\ r &= \frac{\alpha}{1-h}, \end{aligned}$$

و تابع f دارای ویژگی های

$$\begin{aligned} (a) \text{ به ازای هر } x, y \in C, \|Df(x) - Df(y)\| &\leq \gamma\|x - y\| \\ (b) \text{ به ازای هر } x \in C, Df(x)^{-1} \text{ موجود باشد و } \|Df(x)^{-1}\| &\leq \beta \\ (c) \|Df(x_0)^{-1}f(x_0)\| &\leq \alpha \end{aligned}$$

است. در این صورت

(الف) با نقطه شروع x_0 ، هر نقطه

$$x_{k+1} = x_k - Df(x_k)^{-1}f(x_k), \quad k = 0, 1, \dots, \quad (10.5)$$

خوش تعریف است و به ازای هر $k \geq 0$ ، $x_k \in S_r(x_0)$.

(ب) حد دنباله $\{x_k\}$ وجود دارد و اگر $\lim_{k \rightarrow \infty} x_k = \xi$ ، آنگاه $\xi \in \overline{S_r(x_0)}$ و $f(\xi) = 0$.

(ج) به ازای هر $k \geq 0$ ، داریم

$$\|x_k - \xi\| \leq \alpha \frac{h^{\gamma k - 1}}{1 - h^{\gamma k - 1}}.$$

از اینکه $0 < h < 1$ ، مرتبه همگرایی روش نیوتن-رافسون حداقل ۲ است.

برهان: به استقرا (الف) را ثابت می کنیم. به ازای $k = 0$ ، داریم $x_0 \in S_r(x_0)$. به ازای $k = 1$ با استفاده از (c)، می بینیم که

$$\|x_1 - x_0\| = \| - Df(x_0)^{-1}f(x_0)\| \leq \alpha = (1-h)r < r,$$

که نشان می دهد $x_1 \in S_r(x_0)$. اکنون، فرض کنید $x_j \in S_r(x_0)$ ، $j = 0, 1, \dots, k$. نشان می دهیم $x_{k+1} \in S_r(x_0)$. برای این کار، ابتدا توجه کنید که با استفاده از رابطه (10.5)، داریم

$$-f(x_{k-1}) - Df(x_{k-1})(x_k - x_{k-1}) = 0.$$

در نتیجه، با استفاده از این رابطه، لم قبل و (b)، نتیجه می‌گیریم که

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|-Df(x_k)^{-1} f(x_k)\| \leq \beta \|f(x_k)\| \\ &= \beta \|f(x_k) - f(x_{k-1}) - Df(x_{k-1})(x_k - x_{k-1})\| \\ &\leq \frac{\beta\gamma}{\nu} \|x_k - x_{k-1}\|^\nu. \end{aligned}$$

اینک، داریم

$$\begin{aligned} \|x_1 - x_0\| &\leq \frac{\beta\gamma}{\nu} \|x_1 - x_0\|^\nu = \frac{\beta\gamma}{\nu} \|-Df(x_0)^{-1} f(x_0)\|^\nu \leq \frac{\beta\gamma}{\nu} \alpha^\nu = \alpha h, \\ \|x_2 - x_1\| &\leq \frac{\beta\gamma}{\nu} \|x_2 - x_1\|^\nu \leq \frac{\beta\gamma}{\nu} (\alpha h)^\nu = \alpha h^\nu, \\ \|x_3 - x_2\| &\leq \frac{\beta\gamma}{\nu} \|x_3 - x_2\|^\nu \leq \frac{\beta\gamma}{\nu} (\alpha h^\nu)^\nu = \alpha h^{\nu^2}, \\ &\vdots \\ \|x_{k+1} - x_k\| &\leq \alpha h^{\nu^k - 1}. \end{aligned} \tag{۱۱.۵}$$

سرانجام، داریم

$$\begin{aligned} \|x_{k+1} - x_0\| &= \|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \dots + \|x_1 - x_0\| \\ &= \alpha(1 + h + h^\nu + \dots + h^{\nu^k - 1}) < \frac{\alpha}{1 - h} = r, \end{aligned}$$

که نشان می‌دهد $x_{k+1} \in S_r(x_0)$

برای اثبات قسمت (ب)، نشان می‌دهیم که دنباله $\{x_k\}$ ، یک دنباله کوشی است. برای نشان دادن این مطلب، با استفاده از رابطه (۱۱.۵) برای هر $m \geq n$ ، داریم

$$\begin{aligned} \|x_{m+1} - x_n\| &\leq \|x_{m+1} - x_m\| + \|x_m - x_{m-1}\| + \dots + \|x_{n+1} - x_n\| \\ &\leq \alpha h^{\nu^n - 1} (1 + h^{\nu^n} + (h^{\nu^n})^2 + \dots) \\ &\leq \frac{\alpha h^{\nu^n - 1}}{1 - h^{\nu^n}}. \end{aligned}$$

اکنون، برای هر $\epsilon > 0$ داده‌شده، با توجه به اینکه $0 < h < 1$ ، یک $N(\epsilon)$ به اندازه کافی بزرگ وجود دارد به طوری که برای هر $n \geq N(\epsilon)$ ، داریم

$$\|x_{m+1} - x_n\| \leq \frac{\alpha h^{\nu^n - 1}}{1 - h^{\nu^n}} < \epsilon, \tag{۱۲.۵}$$

و این یعنی دنباله $\{x_k\}$ کوشی است و بنابراین، حد آن موجود است. اگر حد دنباله را با ξ نشان دهیم، آنگاه

$$\lim_{k \rightarrow \infty} x_k = \xi \in \overline{S_r(x_0)}.$$

در اینجا، توجه کنید که تمام جمله‌های دنباله در $S_r(x_0)$ قرار دارند و بنابراین، حد این دنباله در بستار این مجموعه قرار دارد.

اکنون، نشان می‌دهیم $f(\xi) = 0$. با توجه به اینکه به ازای هر $k \geq 0$ داریم $x_k \in S_r(x_0)$ ، با استفاده از (a) می‌بینیم که

$$\|Df(x_k) - Df(x_0)\| \leq \gamma \|x_k - x_0\| \leq \gamma r,$$

و بنابراین

$$\|Df(x_k)\| \leq \|Df(x_0)\| + \gamma r.$$

اینک، با فرض $K = \|Df(x_0)\| + \gamma r$ داریم

$$\|f(x_k)\| = \|-Df(x_k)(x_{k+1} - x_k)\| \leq K \|x_{k+1} - x_k\|.$$

در نتیجه، با محاسبه حد طرفین این رابطه، خواهیم داشت

$$0 = \lim_{k \rightarrow \infty} \|f(x_k)\| = \|f(\xi)\|.$$

این رابطه نشان می‌دهد که $f(\xi) = 0$.

برای اثبات (۳)، از رابطه (۱۲.۵)، داریم

$$\lim_{m \rightarrow \infty} \|x_{m+1} - x_m\| = \|\xi - x_m\| \leq \frac{\alpha h^{2^m - 1}}{1 - h^{2^m}},$$

و با توجه به اینکه $0 < h < 1$ ، نتیجه می‌گیریم که مرتبه همگرایی روش ۲ است. \square

۴.۵ روش بیرستو

فرض کنید $P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$ ، که در آن $a_i \in \mathbb{R}$ ، $i = 0, 1, \dots, n$ ، $a_0 \neq 0$ و $n \geq 3$. توجه کنید که به ازای $n = 1, 2$ ریشه‌های P به سادگی قابل محاسبه هستند. می‌دانیم که در حالت کلی، هیچ فرمول صریحی برای محاسبه ریشه‌های P وجود ندارد، ولی با استفاده

از روش بیرستو می‌توان ریشه‌های این چندجمله‌ای را به‌طور تقریبی محاسبه کرد. همان‌گونه که می‌دانیم، ریشه‌های مختلط P دوبه‌دو مزدوج هم هستند. بنابراین، P را می‌توان به‌صورت حاصل‌ضربی از عامل‌های درجه‌ اول و دوم تجزیه کرد. ایده اصلی روش بیرستو این است که در هر مرحله، یک عامل درجه دوم مثل $x^2 - rx - q$ از چندجمله‌ای P استخراج می‌شود. فرض کنید r و q دو عدد حقیقی هستند. در این صورت، برای هر چندجمله‌ای از درجه $n-2$ مثل Q ، داریم

$$P(x) = (x^2 - rx - q)Q(x) + Ax + B, \quad (13.5)$$

که در آن A و B اعدادی حقیقی هستند که به r و q وابسته‌اند، یعنی، داریم $A = A(r, q)$ و $B = B(r, q)$. اکنون، کافی است r و q ای بیابیم که

$$\begin{cases} A(r, q) = 0, \\ B(r, q) = 0. \end{cases} \quad (14.5)$$

در واقع، کافی است دستگاه معادلات غیرخطی (۱۴.۵) را حل کنیم. اگر $A(r, q)$ و $B(r, q)$ معلوم باشند، با انتخاب مقادیر اولیه r_0 و q_0 می‌توانیم از روش نیوتن-رافسون

$$\begin{pmatrix} r_{k+1} \\ q_{k+1} \end{pmatrix} = \begin{pmatrix} r_k \\ q_k \end{pmatrix} - \begin{pmatrix} \frac{\partial A}{\partial r}(r_k, q_k) & \frac{\partial A}{\partial q}(r_k, q_k) \\ \frac{\partial B}{\partial r}(r_k, q_k) & \frac{\partial B}{\partial q}(r_k, q_k) \end{pmatrix}^{-1} \begin{pmatrix} A(r_k, q_k) \\ B(r_k, q_k) \end{pmatrix}, \quad (15.5)$$

برای محاسبه مقادیر تقریبی r و q استفاده کنیم. بنابراین، برای محاسبه r و q با استفاده از رابطه بازگشتی (۱۵.۵)، کافی است روشی برای محاسبه $A(r_k, q_k)$ ، $B(r_k, q_k)$ ، و مشتقات A و B نسبت به r و q به‌ازای (r_k, q_k) داده‌شده، ارائه کنیم. برای این کار، فرض کنید

$$Q(x) = b_0 x^{n-2} + b_1 x^{n-3} + \dots + b_{n-2},$$

که b_0, b_1, \dots, b_{n-2} به r و q وابسته هستند. با جای‌گذاری Q در رابطه (۱۳.۵) و برابری‌دادن ضرایب توان‌های مختلف x در دو طرف رابطه، خواهیم داشت

$$\begin{cases} a_0 = b_0, \\ a_1 = b_1 - rb_0, \\ a_i = b_i - rb_{i-1} - qb_{i-2}, \quad i = 2, \dots, n-2, \\ a_{n-1} = A - rb_{n-2} - qb_{n-3}, \\ a_n = B - qb_{n-2}. \end{cases}$$

اکنون، با محاسبه b_i ها، A و B از رابطهٔ اخیر، داریم

$$\begin{cases} b_0 = a_0, \\ b_1 = a_1 + rb_0, \\ b_i = a_i + rb_{i-1} + qb_{i-2}, \quad i = 2, \dots, n-2, \\ A = a_{n-1} + rb_{n-2} + qb_{n-3}, \\ B = a_n + qb_{n-2}. \end{cases} \quad (16.5)$$

بنابراین، برای r و q داده شده، مقادیر $A(r, q)$ و $B(r, q)$ را می توان به صورت بازگشتی با استفاده از رابطهٔ (۱۶.۵) به دست آورد. برای محاسبهٔ مشتقات A و B نسبت به r و q ، کافی است از طرفین تساوی های رابطهٔ (۱۶.۵)، نسبت r و s مشتق بگیریم. اگر نسبت r مشتق بگیریم، خواهیم داشت (در اینجا اندیس r به معنای مشتق نسبت به r است)

$$\begin{cases} (b_0)_r = 0, \\ (b_1)_r = b_0, \\ (b_i)_r = b_{i-1} + r(b_{i-1})_r + q(b_{i-2})_r, \quad i = 2, \dots, n-2, \\ \frac{\partial A}{\partial r} = b_{n-2} + r(b_{n-2})_r + q(b_{n-3})_r, \\ \frac{\partial B}{\partial r} = q(b_{n-2})_r. \end{cases} \quad (17.5)$$

به همین ترتیب، داریم

$$\begin{cases} (b_0)_q = 0, \\ (b_1)_q = 0, \\ (b_i)_q = r(b_{i-1})_q + q(b_{i-2})_q + b_{i-2}, \quad i = 2, \dots, n-2, \\ \frac{\partial A}{\partial q} = r(b_{n-2})_q + q(b_{n-3})_q + b_{n-3}, \\ \frac{\partial B}{\partial q} = q(b_{n-2})_q + b_{n-2}. \end{cases} \quad (18.5)$$

بنابراین، با استفاده از رابطه های بازگشتی (۱۷.۵) و (۱۸.۵)، می توان مشتقات A و B را در یک نقطه دلخواه (r, q) محاسبه کرد. با جمع بندی مطالب ارائه شده، الگوریتم بیرستو را به صورت زیر جمع بندی می کنیم.

الگوریتم ۱.۵: روش بیرستو

(۱) حدس های اولیه r_0 و q_0 را انتخاب کنید.

(۲) A و B را از رابطهٔ (۱۶.۵) به ازای r_0 و q_0 محاسبه کنید.

(۳) مقادیر $\frac{\partial A}{\partial r}$ ، $\frac{\partial A}{\partial q}$ ، $\frac{\partial B}{\partial r}$ و $\frac{\partial B}{\partial q}$ را از رابطه های (۱۷.۵) و (۱۸.۵) به دست آورید.

(۴) با استفاده از رابطهٔ (۱۵.۵)، مقادیر r_1 و q_1 را محاسبه کنید.

(۵) اگر $|r_1 - r_0| < \epsilon$ و $|q_1 - q_0| < \epsilon$ ، آنگاه قرار دهید $r = r_0$ و $q = q_0$ و به گام ۶ بروید، در غیر این صورت، قرار دهید $r_0 := r_1$ و $q_0 := q_1$ و به گام ۲ بروید.

(۶) ریشه‌های $x^2 - rx - q$ را محاسبه کنید. اگر Q یک چندجمله‌ای درجه اول یا دوم باشد، ریشه‌های آن را محاسبه کنید و الگوریتم را خاتمه دهید، و در غیر این صورت، الگوریتم را برای چندجمله‌ای Q تکرار کنید.

مثال ۸.۵ ریشه‌های تقریبی چندجمله‌ای $P(x) = 8x^4 - 20x^3 + 29x^2 - 18x + 10$ را با کمک روش بیرستو محاسبه کنید. مقدار ϵ را 10^{-2} بگیرید.

حل: حدس اولیه برای r و q را به ترتیب $r_0 = 0.75$ و $q_0 = -0.5$ می‌گیریم. همه محاسبات با یک برنامه رایانه‌ای به زبان متلب و با دقت مضاعف انجام شده و نتایج با چهار رقم اعشار گرد شده است. در تکرار اول، داریم

$$a_0 = 8, \quad a_1 = -20, \quad a_2 = 29, \quad a_3 = -18, \quad a_4 = 10,$$

$$b_0 = a_0 = 8,$$

$$b_1 = a_1 + r_0 b_0 = -20 + 0.75 \times 8 = -14,$$

$$b_2 = a_2 + r_0 b_1 + q_0 b_0 = 29 + 0.75 \times (-14) + (-0.5) \times 8 = 14.5,$$

$$A = a_3 + r_0 b_2 + q_0 b_1 = -18 + 0.75 \times 14.5 + (-0.5) \times (-14) = -0.125,$$

$$B = a_4 + q_0 b_2 = 10 + (-0.5) \times 14.5 = 2.75,$$

$$(b_0)_r = 0,$$

$$(b_1)_r = b_0 = 8,$$

$$(b_2)_r = b_1 + r_0 (b_1)_r + q_0 (b_0)_r = -14 + 0.75 \times 8 + (-0.5) \times 0 = -8,$$

$$\frac{\partial A}{\partial r} = b_2 + r_0 (b_2)_r + q_0 (b_1)_r = 14.5 + 0.75 \times (-8) + (-0.5) \times 8 = 4.5,$$

$$\frac{\partial B}{\partial r} = q_0 (b_2)_r = -0.5 \times (-8) = 4,$$

$$(b_0)_q = 0$$

$$(b_1)_q = 0,$$

$$(b_2)_q = r_0 (b_1)_q + q_0 (b_0)_q + b_0 = 0.75 \times 0 + (-0.5) \times 0 + 8 = 8,$$

$$\frac{\partial A}{\partial q} = r_0 (b_2)_q + q_0 (b_1)_q + b_1 = 0.75 \times 8 + (-0.5) \times 0 + (-14) = -8,$$

$$\frac{\partial B}{\partial q} = q_0 (b_2)_q + b_2 = -0.5 \times 8 + 14.5 = 10.5.$$

بنابراین

$$\begin{aligned} d &= - \begin{pmatrix} \frac{\partial A}{\partial r}(r_0, q_0) & \frac{\partial A}{\partial q}(r_0, q_0) \\ \frac{\partial B}{\partial r}(r_0, q_0) & \frac{\partial B}{\partial q}(r_0, q_0) \end{pmatrix}^{-1} \begin{pmatrix} A(r_0, q_0) \\ B(r_0, q_0) \end{pmatrix} \\ &= - \begin{pmatrix} ۴.۵ & -۸ \\ ۴ & ۱۰.۵ \end{pmatrix}^{-1} \begin{pmatrix} -۰.۱۲۵ \\ ۲.۷۵ \end{pmatrix} \\ &= \begin{pmatrix} -۰.۲۶۱۱ \\ -۰.۱۶۲۵ \end{pmatrix}. \end{aligned}$$

در نتیجه

$$\begin{pmatrix} r_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} r_0 \\ q_0 \end{pmatrix} + d = \begin{pmatrix} ۰.۷۵ \\ -۰.۵ \end{pmatrix} + \begin{pmatrix} -۰.۲۶۱۱ \\ -۰.۱۶۲۵ \end{pmatrix} = \begin{pmatrix} ۰.۴۸۹ \\ -۰.۶۶۲۵ \end{pmatrix}.$$

با توجه به اینکه

$$\|d\|_{\infty} = \max\{|r_1 - r_0|, |q_1 - q_0|\} = ۰.۲۶۱۱ \not\leq ۱۰^{-۲},$$

تکرار را ادامه می‌دهیم. در تکرار دوم، قرار می‌دهیم $r_0 = ۰.۴۸۹$ و $q_0 = -۰.۶۶۲۵$. همانند تکرار اول، داریم

$$b_0 = ۸, \quad b_1 = -۱۶.۰۸۸۳, \quad b_2 = ۱۵.۸۳۳۸,$$

$$A = ۰.۴, \quad B = -۰.۴۸۹۳,$$

$$(b_0)_r = ۰, \quad (b_1)_r = ۸, \quad (b_2)_r = -۱۲.۱۷۶۷,$$

$$\frac{\partial A}{\partial r} = ۴.۵۸۰۲, \quad \frac{\partial B}{\partial r} = ۸.۰۶۶۶,$$

$$(b_0)_q = ۰, \quad (b_1)_q = ۰, \quad (b_2)_q = ۸,$$

$$\frac{\partial A}{\partial q} = -۱۲.۱۷۶۷, \quad \frac{\partial B}{\partial q} = ۱۰.۵۳۴۱,$$

$$d = \begin{pmatrix} ۰.۰۱۱۹ \\ ۰.۰۳۷۳ \end{pmatrix},$$

$$\begin{pmatrix} r_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} r_0 \\ q_0 \end{pmatrix} + d = \begin{pmatrix} ۰.۴۸۹۰ \\ -۰.۶۶۲۵ \end{pmatrix} + \begin{pmatrix} ۰.۰۱۱۹ \\ ۰.۰۳۷۳ \end{pmatrix} = \begin{pmatrix} ۰.۵۰۰۹ \\ -۰.۶۲۵۱ \end{pmatrix}.$$

اکنون، از اینکه

$$\|d\|_{\infty} = \max\{|r_1 - r_0|, |q_1 - q_0|\} = ۰.۰۳۷۳ \not\leq ۱۰^{-۲},$$

تکرار را ادامه می‌دهیم. در تکرار سوم، قرار می‌دهیم $r_0 = ۰.۵۰۰۹$ و $q_0 = -۰.۶۲۵۱$. همانند دو تکرار قبل، داریم

$$b_0 = ۸, \quad b_1 = -۱۵.۹۹۳۱, \quad b_2 = ۱۵.۹۸۸۵,$$

$$A = ۰.۰۰۶, \quad B = ۰.۰۰۵,$$

$$(b_0)_r = 0, \quad (b_1)_r = 8, \quad (b_2)_r = -11.9861,$$

$$\frac{\partial A}{\partial r} = 4.9840, \quad \frac{\partial B}{\partial r} = 7.4929,$$

$$(b_0)_q = 0, \quad (b_1)_q = 0, \quad (b_2)_q = 8,$$

$$\frac{\partial A}{\partial q} = -11.9861, \quad \frac{\partial B}{\partial q} = 10.9874,$$

$$d = 10^{-2} \begin{pmatrix} -0.869 \\ 0.1354 \end{pmatrix},$$

$$\begin{pmatrix} r_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} r_0 \\ q_0 \end{pmatrix} + d = \begin{pmatrix} 0.5009 \\ -0.6251 \end{pmatrix} + 10^{-2} \begin{pmatrix} -0.869 \\ 0.1354 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -0.625 \end{pmatrix}.$$

اینک، با توجه به

$$\|d\|_{\infty} = \max\{|r_1 - r_0|, |q_1 - q_0|\} = 0.8690 \times 10^{-2} < 10^{-2},$$

تکرار را پایان داده و قرار می‌دهیم $r = 0.5$ و $q = -0.625$. بنابراین، دو ریشه P از حل معادله درجه دوم

$$x^2 - rx - q = x^2 - 0.5x + 0.625 = 0,$$

به دست می‌آید. این دو ریشه به صورت تقریبی برابر با

$$x_1 = 0.25 + 0.75i,$$

$$x_2 = 0.25 - 0.75i,$$

هستند. دو ریشه دیگر P ، از حل معادله درجه دوم

$$b_0 x^2 + b_1 x + b_2 = 8x^2 - 15.9931x + 15.9885 = 0,$$

به دست می‌آید که برابر با

$$x_3 = 0.9996 + 0.9997i,$$

$$x_4 = 0.9996 - 0.9997i,$$

هستند. شایان ذکر است که ریشه‌های دقیق P عبارتند از

$$x_1 = 1 + i,$$

$$x_2 = 1 - i,$$

$$x_3 = 0.25 + 0.75i,$$

$$x_4 = 0.25 - 0.75i.$$

همان‌گونه که می‌بینیم، روش بیرستو تقریب‌های مناسبی از ریشه‌ها فراهم کرده است.

۵.۵ روش نابجایی

فرض کنید تابع f روی بازه $[x_0, a_0]$ پیوسته است و $f(x_0)f(a_0) < 0$. در این صورت، تابع f در این بازه حداقل یک ریشه دارد. در تکرار نام روش نابجایی، دو مقدار x_i و a_i در بازه $[x_0, a_0]$ به گونه‌ای محاسبه می‌شوند که

$$f(x_i)f(a_i) < 0,$$

و بنابراین، بازه $[x_i, a_i]$ نیز شامل حداقل یک ریشه از f است. به علاوه، مقادیر x_i ها را طوری محاسبه می‌کنیم که دنباله $\{x_i\}$ همگرا باشد و اگر $\lim_{i \rightarrow \infty} x_i = \xi$ ، آنگاه $f(\xi) = 0$. برای تعریف مقادیر x_{i+1} و a_{i+1} ، چند جمله‌ای درونیاب خطی تابع f را در نقاط $(x_i, f(x_i))$ و $(a_i, f(a_i))$ می‌نویسیم. این چند جمله‌ای به صورت

$$P(x) = f(x_i) + (x - x_i) \frac{f(x_i) - f(a_i)}{x_i - a_i},$$

است. با توجه به اینکه $f(x_i)f(a_i) < 0$ ، چند جمله‌ای P در بازه $[x_i, a_i]$ حداقل یک ریشه خواهد داشت. اگر این ریشه را μ_i بنامیم، آنگاه خواهیم داشت

$$\mu_i = x_i - f(x_i) \frac{x_i - a_i}{f(x_i) - f(a_i)} = \frac{a_i f(x_i) - x_i f(a_i)}{f(x_i) - f(a_i)}. \quad (19.5)$$

در اینجا توجه کنید که از $f(x_i)f(a_i) < 0$ داریم، $f(x_i) - f(a_i) \neq 0$ و بنابراین، μ_i خوش تعریف است و به علاوه، $x_i < \mu_i < a_i$.

اکنون، اگر $f(\mu_i) = 0$ ، آنگاه μ_i یک ریشه f خواهد بود. بنابراین، فرض می‌کنیم $f(\mu_i) \neq 0$. در این صورت، x_{i+1} و a_{i+1} را به صورت زیر تعریف می‌کنیم

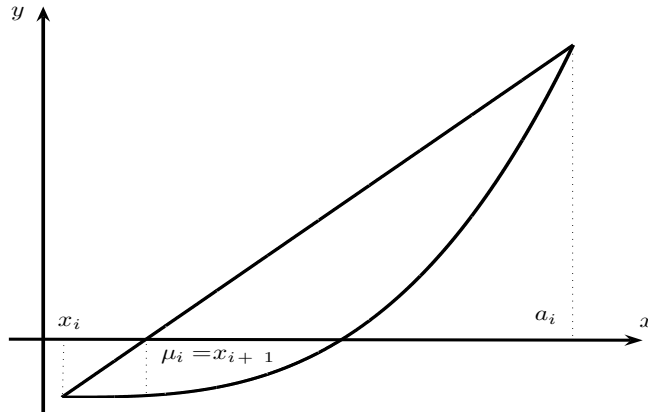
$$\left. \begin{aligned} x_{i+1} &= \mu_i \\ a_{i+1} &= a_i \end{aligned} \right\} \quad \text{if } f(\mu_i)f(x_i) > 0,$$

$$\left. \begin{aligned} x_{i+1} &= x_i \\ a_{i+1} &= \mu_i \end{aligned} \right\} \quad \text{if } f(\mu_i)f(x_i) < 0.$$

برای اثبات همگرایی روش نابجایی، فرض کنید f'' موجود باشد و به ازای یک i ،

$$\left\{ \begin{aligned} x_i &< a_i, \\ f(x_i) &< 0, \quad f(a_i) > 0, \\ f''(x) &\geq 0 \quad \forall x \in [x_i, a_i]. \end{aligned} \right. \quad (20.5)$$

شکل ۵.۵ به خوبی تابعی با خواص (۲۰.۵) را نمایش می‌دهد.



شکل ۵.۵: نمایش یک تابع با خواص (۲۰.۵).

قرار می‌دهیم

$$t = \frac{f(x_i)}{f(x_i) - f(a_i)}.$$

به سادگی می‌توان دید که $0 < t < 1$ و $\mu_i = tf(x_i) + (1-t)f(a_i)$. رابطهٔ اخیر نشان می‌دهد که نخست $x_i < \mu_i < a_i$ و دوم اینکه با توجه به شرط سوم در (۲۰.۵)، می‌بینیم که تابع f در بازهٔ $[x_i, a_i]$ محدب است و در نتیجه، داریم

$$f(\mu_i) \leq tf(x_i) + (1-t)f(a_i) = 0.$$

از این رو، $f(\mu_i) = 0$ یا $f(\mu_i)f(x_i) > 0$. در حالت اول، μ_i یک ریشهٔ تابع f خواهد بود و در حالت دوم، داریم

$$x_i < x_{i+1} = \mu_i < a_{i+1} = a_i.$$

با جمع‌بندی مطالب بالا می‌بینیم که اگر به ازای یک i ، شرایط (۲۰.۵) برقرار باشد، آنگاه به ازای هر $i \geq i_0$ خواهیم داشت $a_i = a$ ، که در آن a یک مقدار ثابت است. به علاوه، دنبالهٔ $\{x_i\}$ یک دنبالهٔ اکیداً صعودی و از بالا کراندار خواهد بود. بنابراین، این دنباله همگراست. فرض کنید $\lim_{i \rightarrow \infty} x_i = \xi$. نشان می‌دهیم ξ یک ریشهٔ f است.

با توجه به رابطهٔ (۱۹.۵) و مطالب ارائه‌شدهٔ اخیر، داریم

$$x_{i+1} = \frac{af(x_i) - x_i f(a)}{f(x_i) - f(a)}, \quad f(\xi) \leq 0, \quad f(a) > 0.$$

بنابراین

$$\xi = \frac{af(\xi) - \xi f(a)}{f(\xi) - f(a)},$$

و در نتیجه، $(\xi - a)f(\xi) = 0$ ولی، با توجه به اینکه $f(a) > 0$ و $f(\xi) \leq 0$ ، داریم $\xi \neq a$.
از این رو، $f(\xi) = 0$.

در ادامه، مرتبه همگرایی روش نابجایی را تعیین می‌کنیم. با توجه به مطالب ارائه‌شده، روش تکرار نابجایی را می‌توان به صورت

$$x_{i+1} = \Phi(x_i), \quad \Phi(x) = \frac{af(x) - xf(a)}{f(x) - f(a)},$$

نوشت. با محاسبه $\Phi'(\xi)$ و به کارگیری قضیه مقدار میانگین، داریم

$$\Phi'(\xi) = 1 - f'(\xi) \frac{\xi - a}{f(\xi) - f(a)} = 1 - \frac{f'(\xi)}{f'(\eta_1)}, \quad (21.5)$$

که در آن $a < \eta_1 < \xi$. از طرفی، داریم

$$\frac{f(x_i) - f(\xi)}{x_i - \xi} = f'(\eta_2),$$

که در آن $\xi < \eta_2 < x_i$. اکنون، با توجه به اینکه در بازه $[x_i, a]$ داریم $f''(x) \geq 0$ ، نتیجه می‌گیریم که f' در بازه $[x_i, a]$ صعودی است. از طرفی، داریم

$$f'(\eta_2) = \frac{f(x_i)}{x_i - \xi} > 0,$$

و بنابراین

$$0 < f'(\eta_2) \leq f'(\xi) \leq f'(\eta_1),$$

و

$$0 < \frac{f'(\eta_2)}{f'(\eta_1)} \leq \frac{f'(\xi)}{f'(\eta_1)} \leq 1.$$

از این رو

$$0 \leq 1 - \frac{f'(\xi)}{f'(\eta_1)} < 1,$$

که با توجه به رابطه (21.5)، خواهیم داشت $0 \leq \Phi'(\xi) < 1$ که نشان می‌دهد مرتبه همگرایی روش نابجایی حداقل ۱ است.

۶.۵ روش وتری

فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ پیوسته است و به ازای یک $\xi \in [a, b]$ داریم $f(\xi) = 0$. در روش وتری، تکرار با دو حدس اولیه x_0 و x_1 با شرط $f(x_0) \neq f(x_1)$ شروع می‌شود و در تکرار $(n+1)$ ام تقریب x_{n+1} با استفاده از دو تقریب x_n و x_{n-1} به صورت زیر ساخته می‌شود. چند جمله‌ای درونیاب خطی تابع f را در نقاط $(x_{n-1}, f(x_{n-1}))$ و $(x_n, f(x_n))$ در نظر بگیرید. اگر این چند جمله‌ای را $P(x)$ بنامیم، خواهیم داشت

$$P(x) = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x - x_n).$$

توجه کنید که در اینجا نیز شرط $f(x_{n-1}) \neq f(x_n)$ لازم است. محل برخورد این خط و محور x ها را $(x_{n+1}, 0)$ می‌گیریم. در این صورت، داریم

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n) = \frac{x_{n-1}f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})} \quad (22.5)$$

فرض کنید $\xi - x_\ell = e_\ell$ ، $\ell = 0, 1, 2, \dots$. در این صورت، با استفاده از رابطه (۲۲.۵)، داریم

$$\begin{aligned} e_{n+1} &= e_n - \frac{1}{f[x_{n-1}, x_n]} f(x_n) \\ &= e_n \left(1 - \frac{1}{f[x_{n-1}, x_n]} \frac{f(x_n) - f(\xi)}{x_n - \xi} \right) \\ &= e_n \left(1 - \frac{f[x_n, \xi]}{f[x_{n-1}, x_n]} \right) \\ &= e_n \frac{f[x_{n-1}, x_n] - f[x_n, \xi]}{f[x_{n-1}, x_n]} \\ &= e_n e_{n-1} \frac{f[x_{n-1}, x_n, \xi]}{f[x_{n-1}, x_n]}. \end{aligned} \quad (23.5)$$

قضیه بعدی، شرایطی کافی برای همگرایی روش وتری ارائه می‌کند.

قضیه ۶.۵ فرض کنید f, f', f'' در بازه‌ای شامل ξ پیوسته هستند و $f'(\xi) \neq 0$. اگر x_0 و x_1 به اندازه کافی نزدیک به ξ انتخاب شوند، آنگاه دنباله $\{x_n\}$ که توسط (۲۲.۵) تولید می‌شود، به ξ همگراست و مرتبه همگرایی دنباله برابر با $p = \frac{1+\sqrt{5}}{2}$ است.

برهان: با توجه به اینکه f' در یک همسایگی از ξ پیوسته است و $f'(\xi) \neq 0$ ، یک بازه مثل

$$J = [\xi - \epsilon, \xi + \epsilon],$$

وجود دارد به طوری که به ازای هر $x \in J$ ، داریم $f'(x) \neq 0$. قرار می‌دهیم

$$M = \frac{\max_{x \in J} |f''(x)|}{\forall \min_{x \in J} |f'(x)|}.$$

اکنون، فرض کنید $x_0, x_1 \in I$. در این صورت، با استفاده از رابطه (۲۳.۵)، داریم

$$e_2 = e_0 e_1 \frac{f''(\eta_2)}{2f'(\eta_1)},$$

که در آن، $\eta_1 \in I[x_0, x_1]$ و $\eta_2 \in I[x_0, x_1, \xi]$. در نتیجه

$$|e_2| = |e_0| |e_1| \left| \frac{f''(\eta_2)}{2f'(\eta_1)} \right| \leq M |e_0| |e_1|. \quad (24.5)$$

فرض کنید x_0 و x_1 به اندازه کافی به ξ نزدیک هستند به طوری که

$$\delta = \max\{M|e_0|, M|e_1|\} < 1.$$

بنابراین، با استفاده از رابطه (۲۳.۵)، خواهیم داشت

$$|e_2| \leq M|e_0||e_1| < |e_0|,$$

که نشان می‌دهد $x_2 \in I$ به همین ترتیب، به استقرا می‌توان دید که به ازای هر $i \geq 0$ ، داریم $x_i \in J$

برای اثبات همگرایی دنباله، نخست می‌بینیم که اگر $M = 0$ ، آنگاه با توجه به رابطه (۲۴.۵) داریم $e_2 = 0$ و این یعنی $x_2 = \xi$. اکنون، فرض کنید $M \neq 0$. در این صورت، با استفاده از رابطه (۲۳.۵)، داریم

$$e_{n+1} = e_{n-1} e_n \frac{f''(\eta_2)}{2f'(\eta_1)}, \quad (25.5)$$

که در آن، $\eta_1 \in I[x_{n-1}, x_n]$ و $\eta_2 \in I[x_{n-1}, x_n, \xi]$. در نتیجه

$$|e_{n+1}| \leq M|e_{n-1}||e_n|.$$

از این رو

$$M|e_2| = M|e_1|M|e_0| \leq \delta^2,$$

$$M|e_3| = M|e_2|M|e_1| \leq \delta^2 \delta = \delta^3,$$

$$M|e_4| = M|e_3|M|e_2| \leq \delta^3 \delta^2 = \delta^5,$$

⋮

$$M|e_n| \leq \delta^{q_n},$$

که در آن $q_0 = q_1 = 1$ و $q_n = q_{n-1} + q_{n-2}$. دنباله $\{q_n\}$ ، به دنباله فیوناچی معروف است و داریم $\lim_{n \rightarrow \infty} q_n = \infty$. بنابراین

$$\lim_{n \rightarrow \infty} M|e_n| \leq \lim_{n \rightarrow \infty} \delta^{q_n} = 0,$$

که همگرایی دنباله $\{x_n\}$ را نشان می‌دهد.

با توجه به رابطه (۲۵.۵)، روشن است که $\lim_{n \rightarrow \infty} f''(\eta_2) = f''(\xi)$ و $\lim_{n \rightarrow \infty} f'(\eta_1) = f'(\xi)$.

بنابراین

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n e_{n-1}} = \frac{f''(\xi)}{2f'(\xi)}.$$

این رابطه را می‌توان به صورت

$$|e_{n+1}| = c_n |e_n| |e_{n-1}|,$$

نوشت که در آن

$$\lim_{n \rightarrow \infty} c_n = \left| \frac{f''(\xi)}{2f'(\xi)} \right|.$$

مقدار $1 \geq p$ را طوری می‌یابیم که

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} = C.$$

برای این کار، داریم

$$\frac{|e_{n+1}|}{|e_n|^p} = c_n |e_n|^{1-p} |e_{n-1}| = c_n \left(\frac{|e_n|}{|e_{n-1}|^p} \right)^{1-p} |e_{n-1}|^{p(1-p)+1}$$

اکنون، مقدار p را به گونه‌ای تعیین می‌کنیم که $1 = 0 = p(1-p) + 1$. در این صورت، خواهیم داشت

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} = \lim_{n \rightarrow \infty} c_n \left(\frac{|e_n|}{|e_{n-1}|^p} \right)^{1-p},$$

یا

$$C = \left| \frac{f''(\xi)}{2f'(\xi)} \right| C^{1-p} \implies C = \left| \frac{f''(\xi)}{2f'(\xi)} \right|^{\frac{1}{p}}.$$

بنابراین

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} = C = \left| \frac{f''(\xi)}{2f'(\xi)} \right|^{\frac{1}{p}},$$

که در آن عدد p در معادله درجه دوم $1 = 0 = p(1-p) + 1$ صدق می‌کند. با محاسبه p ، خواهیم

داشت $p = \frac{1+\sqrt{5}}{2} \approx 1.618$. □

۷.۵ روش Δ^2 -ایتکن

روش Δ^2 -ایتکن، یکی از ساده‌ترین روش‌ها برای شتاب‌بخشیدن همگرایی یک دنباله است. فرض کنید دنباله $\{x_i\}$ به ξ همگرا است. ایده اصلی روش Δ^2 -ایتکن این است که دنباله $\{x_i\}$ را به دنباله دیگری مثل $\{x'_i\}$ تبدیل می‌کند، به گونه‌ای که سرعت همگرایی این دنباله به ξ بیشتر از دنباله اولیه باشد.

فرض کنید $\lim_{i \rightarrow \infty} x_i = \xi$ و مرتبه همگرایی دنباله برابر با یک باشد، یعنی

$$\lim_{i \rightarrow \infty} \frac{x_{i+1} - \xi}{x_i - \xi} = K,$$

به طوری که $|K| < 1$. به ازای مقادیر بزرگ i ، داریم

$$\frac{x_{i+1} - \xi}{x_i - \xi} \approx \frac{x_{i+2} - \xi}{x_{i+1} - \xi}.$$

با محاسبه ξ از این رابطه، داریم $x'_i \approx \xi$ ، که در آن

$$x'_i = x_i - \frac{(x_{i+1} - x_i)^2}{x_{i+2} - 2x_{i+1} + x_i}.$$

همان‌گونه که می‌دانیم $\Delta x_i = x_{i+1} - x_i$ و $\Delta^2 x_i = x_{i+2} - 2x_{i+1} + x_i$ که در آن Δ عملگر تفاضل پیشرو است و بنابراین

$$x'_i = x_i - \frac{(\Delta x_i)^2}{\Delta^2 x_i}. \quad (26.5)$$

قضیه ۷.۵ فرض کنید عددی مثل K با شرط $|K| < 1$ وجود داشته باشد به طوری که برای دنباله $\{x_i\}$ با شرط $x_i \neq \xi$ داشته باشیم

$$x_{i+1} - \xi = (K + \delta_i)(x_i - \xi), \quad \lim_{i \rightarrow \infty} \delta_i = 0.$$

در این صورت، دنباله $\{x'_i\}$ که در رابطه (۲۶.۵) تعریف شده است، موجود است و

$$\lim_{i \rightarrow \infty} \frac{x'_i - \xi}{x_i - \xi} = 0.$$

برهان: قرار می‌دهیم $e_i = x_i - \xi$ ، $i = 0, 1, \dots$. در اینجا، با توجه به فرض قضیه، داریم

$$e_{i+1} = (K + \delta_i)e_i.$$

در نتیجه

$$\begin{aligned} x_{i+2} - 2x_{i+1} + x_i &= e_{i+2} - 2e_{i+1} + e_i \\ &= e_i((K + \delta_{i+1})(K + \delta_i) - 2(K + \delta_i) + 1) \\ &= e_i((K - 1)^2 + \mu_i), \end{aligned} \quad (27.5)$$

که در آن $\mu_i = K(\delta_i + \delta_{i+1}) - 2\delta_i \rightarrow 0$ به علاوه، داریم

$$x_{i+1} - x_i = e_i((K - 1) + \delta_i), \quad \delta_i \rightarrow 0. \quad (28.5)$$

با توجه به اینکه به ازای مقادیر بزرگ i داریم $K \neq 1$ ، $\mu_i \rightarrow 0$ و $e_i \neq 0$ می بینیم که

$$x_{i+2} - 2x_{i+1} + x_i \neq 0,$$

و در نتیجه، دنباله $\{x'_i\}$ خوش تعریف است. با استفاده از رابطه های (27.5) و (28.5) به ازای مقادیر بزرگ i داریم

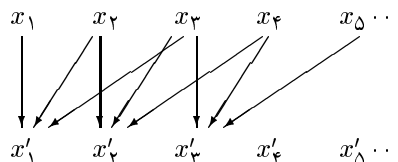
$$x'_i - \xi = e_i - e_i \frac{((K - 1) + \delta_i)^2}{(K - 1)^2 + \mu_i}.$$

بنابراین

$$\lim_{i \rightarrow \infty} \frac{x'_i - \xi}{x_i - \xi} = \lim_{i \rightarrow \infty} \left(1 - \frac{((K - 1) + \delta_i)^2}{(K - 1)^2 + \mu_i} \right) = 0,$$

□ که اثبات قضیه را کامل می کند.

این قضیه نشان می دهد که همگرایی دنباله $\{x'_i\}$ سریع تر از دنباله $\{x_i\}$ است. در عمل از رابطه (26.5) به گونه ای که شرح آن در ادامه می آید، استفاده می کنیم. نخست x_0 ، x_1 و x_2 را محاسبه می کنیم و سپس مقدار x'_0 را با استفاده از این سه مقدار به دست می آوریم. سپس x_3 را محاسبه کرده و با استفاده از x_1 ، x_2 و x_3 مقدار x'_3 را محاسبه می کنیم. به همین ترتیب، این روند را ادامه می دهیم تا شرط توقف برقرار شود. شایان ذکر است که شرط توقف روی دنباله $\{x'_i\}$ گذاشته می شود. برای مثال، شرطی به صورت $|f(x'_i)| < \epsilon$ ، $|x'_{i+1} - x'_i| < \epsilon$ یا به طور هم زمان هر دو شرط. شکل زیر، نمایی از شیوه به کارگیری روش Δ^2 -ایتکن را نشان می دهد.



روش دیگری برای به کارگیری رابطه (26.5) وجود دارد که به روش استیفنسن معروف است.

فرض کنید $i = 0, 1, \dots, x_{i+1} = \Phi(x_i)$ دنباله‌ای باشد که به ریشه f همگراست و مرتبه همگرایی آن ۱ است. در روش استیفنسن، نخست با استفاده از تابع تکرار Φ ، مقادیر x_0, x_1 و x_2 را محاسبه کرده و به کمک رابطه (۲۶.۵) مقدار x'_0 را محاسبه می‌کنیم. سپس، قرار می‌دهیم $x'_0 := x_1$ و با نقطه شروع x_1 و تابع تکرار Φ ، مقادیر x_2 و x_3 را محاسبه کرده و با استفاده از رابطه (۲۶.۵) مقدار x'_1 را محاسبه می‌کنیم. با ادامه این روند، دنباله $\{x'_i\}$ تولید می‌شود که ویژگی‌های آن در ادامه مورد بررسی قرار می‌گیرد. با توجه به مطالب بالا، روش استیفنسن را می‌توان به صورت زیر جمع‌بندی کرد

$$y_i = \Phi(x_i), \quad z_i = \Phi(y_i),$$

$$x_{i+1} := x_i - \frac{(y_i - x_i)^2}{z_i - 2y_i + x_i}, \quad i = 0, 1, 2, \dots,$$

که در اینجا، $\{x_i\}$ دنباله تولیدشده توسط روش استیفنسن است. این روش، دنباله زیر را تولید می‌کند

$$x_{i+1} = \Psi(x_i), \quad i = 0, 1, 2, \dots,$$

$$\Psi(x) = \frac{x\Phi(\Phi(x)) - \Phi(x)^2}{\Phi(\Phi(x)) - 2\Phi(x) + x}. \quad (29.5)$$

قضیه ۸.۵ اگر $\Psi(\xi) = \xi$ ، آنگاه $\xi = \Phi(\xi)$. برعکس، اگر $\xi = \Phi(\xi)$ و $\Phi'(\xi) \neq 1$ ، آنگاه $\Psi(\xi) = \xi$.

برهان: اگر $\xi = \Psi(\xi)$ ، آنگاه داریم

$$\frac{\xi\Phi(\Phi(\xi)) - \Phi(\xi)^2}{\Phi(\Phi(\xi)) - 2\Phi(\xi) + \xi} = \xi,$$

که رابطه

$$\Phi(\xi)^2 - 2\xi\Phi(\xi) + \xi^2 = 0,$$

را نتیجه می‌دهد و بنابراین، $\xi = \Phi(\xi)$. برعکس، فرض کنید $\xi = \Phi(\xi)$. در نتیجه، $\Psi(\xi) = \xi$. بنابراین، با استفاده از قاعده هوییتال، داریم

$$\Psi(\xi) = \frac{\Phi(\Phi(\xi)) + \xi\Phi'(\Phi(\xi))\Phi'(\xi) - 2\Phi(\xi)\Phi'(\xi)}{\Phi'(\Phi(\xi))\Phi'(\xi) - 2\Phi'(\xi) + 1} = \frac{\xi + \xi\Phi'(\xi)^2 - 2\xi\Phi'(\xi)}{1 + \Phi'(\xi)^2 - 2\Phi'(\xi)} = \xi,$$

□ که اثبات قضیه را کامل می‌کند.

برای اینکه رفتار همگرایی $x_{i+1} = \Psi(x_i)$ را در یک همسایگی از یک نقطه ثابت مثل ξ بررسی کنیم، فرض می‌کنیم تابع Φ در یک همسایگی از ξ ، $p+1$ بار مشتق‌پذیر باشد و یک روش از مرتبه p تعریف کند، یعنی

$$\Phi'(\xi) = \dots = \Phi^{(p-1)}(\xi) = 0, \quad \Phi^{(p)}(\xi) = p!A \neq 0,$$

و به علاوه، اگر $p = 1$ ، آنگاه $A = \Phi'(\xi) \neq 1$ بدون از دست دادن کلیت، می‌توان فرض کرد $\xi = 0$. بنابراین، برای مقادیر کوچک x داریم

$$\Phi(x) = Ax^p + \frac{x^{p+1}}{(p+1)!} \Phi^{(p+1)}(\theta x), \quad 0 < \theta < 1,$$

و بنابراین، داریم $\Phi(x) = Ax^p + \mathcal{O}(x^{p+1})$ از این رو،

$$\begin{aligned} \Phi(\Phi(x)) &= A \left(Ax^p + \mathcal{O}(x^{p+1}) \right)^p + \mathcal{O} \left(\left(Ax^p + \mathcal{O}(x^{p+1}) \right)^{p+1} \right) \\ &= \begin{cases} \mathcal{O}(x^{p^2}), & p > 1, \\ A^2 x^p + \mathcal{O}(x^2), & p = 1, \end{cases} \\ \Phi(x)^2 &= \left(Ax^p + \mathcal{O}(x^{p+1}) \right)^2 = A^2 x^{2p} + \mathcal{O}(x^{2p+1}). \end{aligned}$$

در نتیجه، اگر $p > 1$ ، آنگاه

$$\Psi(x) = \frac{\mathcal{O}(x^{p^2+1}) - A^2 x^{2p} + \mathcal{O}(x^{2p+1})}{\mathcal{O}(x^{p^2}) - 2Ax^p + \mathcal{O}(x^{p+1}) + x} = -A^2 x^{2p-1} + \mathcal{O}(x^{2p}),$$

و اگر $p = 1$ ، آنگاه

$$\Psi(x) = \frac{A^2 x^2 + \mathcal{O}(x^3) - A^2 x^2 + \mathcal{O}(x^2)}{A^2 x + \mathcal{O}(x^2) - 2Ax + \mathcal{O}(x^2) + x} = \mathcal{O}(x^2).$$

با جمع‌بندی مطالب بالا قضیه زیر را می‌توان بیان کرد.

قضیه ۹.۵ فرض کنید Φ تابع تکراریک روش برای تعیین نقطه ثابت ξ و از مرتبه p است. در این صورت، اگر $p > 1$ ، آنگاه روش تکراری تولیدشده توسط تابع تکرار Ψ برای تعیین ξ از مرتبه $2p-1$ و اگر $p = 1$ و شرط $\Phi(\xi) \neq 1$ برقرار باشد، آنگاه روش، حداقل از مرتبه ۲ خواهد بود.

در اینجا توجه کنید که روند تکراری تولیدشده توسط Ψ برای تعیین ξ ، حداقل از مرتبه ۲ و به‌طور موضعی همگرا با مرتبه همگرایی ۲ است، حتی اگر $|\Phi'(\xi)| > 1$ و دنباله تولیدشده توسط Φ واگرا باشد.

زمانی که $p = 1$ ، روش استیفنسن از اهمیت بسزایی برخوردار است؛ زیرا فرض کنید $p > 1$ و $\epsilon_i = x_i - \xi$ ، که در آن $|\epsilon_i|$ به اندازه کافی کوچک است. در این صورت، داریم

$$\begin{aligned}\Phi(x_i) &= \Phi(\xi + \epsilon_i) = \Phi(\xi) + \Phi'(\xi)\epsilon_i + \dots + \frac{\Phi^{(p-1)}(\xi)}{p!}\epsilon_i^p + \mathcal{O}(\epsilon^{p+1}) \\ &= \xi + A\epsilon_i^p + \mathcal{O}(\epsilon^{p+1})\end{aligned}$$

بنابراین

$$\Phi(x_i) - \xi = A\epsilon_i^p + \mathcal{O}(\epsilon^{p+1}), \quad (A = \frac{\Phi^{(p-1)}(\xi)}{p!}),$$

و در نتیجه، داریم

$$\Phi(\Phi(x_i)) - \xi = A^{p+1}\epsilon_i^{p^2} + \mathcal{O}(\epsilon^{p^2+1}). \quad (30.5)$$

در صورتی که از تابع تکرار Ψ استفاده کنیم و $x_{i+1} = \Psi(x_i)$ ، خواهیم داشت

$$\Psi(x_i) - \xi = x_{i+1} - \xi = -A^2\epsilon_i^{2p-1} + \mathcal{O}(\epsilon^{2p}). \quad (31.5)$$

ولی برای $p > 1$ ، داریم

$$\left| A^{p+1}\epsilon_i^{p^2} \right| \ll \left| A^2\epsilon_i^{2p-1} \right|.$$

با توجه به رابطه‌های (30.5) و (31.5)، رابطه اخیر نشان می‌دهد که $\Phi(\Phi(x_i))$ تقریب بهتری نسبت به $\Psi(x_i)$ برای ξ فراهم می‌کند. به همین دلیل، روش استیفنسن برای $p = 1$ توصیه می‌شود.

مثال 9.5 تابع تکرار $\Phi(x) = x^2$ دارای دو نقطه ثابت $\xi_1 = 0$ و $\xi_2 = 1$ است. به علاوه، داریم

$$\Phi'(\xi_1) = 0, \quad \Phi''(\xi_1) = 2,$$

$$\Phi'(\xi_2) = 2.$$

اگر $|x_0| < 1$ ، آنگاه روند تکراری $x_{i+1} = \Phi(x_i)$ به ξ_1 همگرا است و مرتبه همگرایی آن برابر با ۲ است. ولی اگر $|x_0| > 1$ ، آنگاه دنباله $\{x_i\}$ واگراست.

اکنون، با استفاده از (29.5)، داریم

$$\Psi(x) = \frac{x^3}{x^2 + x - 1} = \frac{x^3}{(x - r_1)(x - r_2)}, \quad (r_{1,2} = \frac{-1 \pm \sqrt{5}}{2})$$

نشان می‌دهیم که با انتخاب نقطه شروع مناسب x_0 ، دنباله $x_{i+1} = \Psi(x_i)$ به هر دو نقطه ثابت همگراست. به ازای $|x_0| \leq 0.5$ ، تابع Ψ یک نگاشت انقباضی است (چرا؟). اگر $|x_0| \leq 0.5$ ،

آنگاه $x_{i+1} = \Psi(x_i)$ به $\xi_1 = 0$ همگرا خواهد شد. در یک همسایگی به اندازه کافی کوچک از ξ_1 ، داریم

$$x_{i+1} = \Psi(x_i) \approx -x_i^3,$$

که نشان می‌دهد مرتبه همگرایی روش برابر با ۳ است. ولی به ازای هر $|x_0| < 1$ ، داریم

$$x_{i+1} = \Phi(\Phi(x_i)) = x_i^4,$$

که به $\xi_1 = 0$ همگرا است و مرتبه همگرایی دنباله برابر با ۴ است. به ازای هر $|x_0| > r_1$ ، دنباله $x_{i+1} = \Phi(x_i)$ به $\xi_2 = 1$ همگرا است. به سادگی می‌توان دید که

$$\Psi'(1) = 0, \quad \Psi''(1) \neq 0,$$

و بنابراین، مرتبه همگرایی دنباله برابر با ۲ است. در اینجا، توجه کنید که دنباله $x_{i+1} = \Phi(x_i)$ تحت هیچ شرایطی به $\xi_2 = 1$ همگرا نمی‌شود.

تمرین

۱.۵ فرض کنید

$$f(x) = f(r, s) = \begin{pmatrix} r + 2s + 3 \\ 4r + 5s + 6 \end{pmatrix}.$$

با استفاده از تعریف، $Df(0)$ را به دست آورید.

۲.۵ نشان دهید اگر ξ یک ریشه تکراری تابع f با مرتبه تکرار m است، آنگاه مرتبه همگرایی روش نیوتن-رافسون اصلاح شده

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

حداقل ۲ است.

۳.۵ فرض کنید تابع f در بازه $[a, b]$ به اندازه کافی هموار و

$$f(a) < 0, \quad f(b) > 0,$$

و به علاوه، به ازای هر $x \in [a, b]$ ، $f''(x) < 0$. نشان دهید که روش نابجایی به یک ریشه تابع f در بازه $[a, b]$ همگرا است.

(۴.۵) فرض کنید ξ یک ریشه ساده تابع $f: \mathbb{R} \rightarrow \mathbb{R}$ است. همچنین، فرض کنید f در گوی

$$S_r(\xi) = \{x : |x - \xi| < r\},$$

دو بار به طور پیوسته مشتق پذیر است. نشان دهید دنباله $\{x_n\}$ که به صورت

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)},$$

$$x_{n+1} = y_n - \frac{f(y_n)}{f'(y_n)},$$

تعریف می شود، به طور موضعی همگرا است. به علاوه، اگر $f(\alpha) = 0$ و $f'(\alpha) \neq 0$ ، آنگاه

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{1}{2} \left(\frac{f''(\alpha)}{f'(\alpha)} \right)^2.$$

(۵.۵) فرض کنید α یک نقطه ثابت تابع به طور پیوسته مشتق پذیر φ است، به گونه ای که $|\varphi'(\alpha)| < 1$. نشان دهید دنباله $x_{n+1} = \varphi(x_n)$ به طور موضعی همگراست.

(۶.۵) فرض کنید برای محاسبه $\sqrt[n]{\alpha}$ ، که در آن $\alpha \neq 0$ یک عدد مفروض است، از یک فرمول تکراری به صورت

$$x_{n+1} = Ax_n + \frac{B\alpha}{x_n^4}, \quad n = 0, 1, 2, \dots,$$

استفاده شود که $x_0 \neq 0$ یک تقریب اولیه برای $\sqrt[n]{\alpha}$ است. مقادیر A و B را طوری تعیین کنید که مرتبه همگرایی حداکثر باشد و اگر قرار دهیم $e_i = x_i - \sqrt[n]{\alpha}$ ، رابطه بین e_n و e_{n+1} را به دست آورید.

(۷.۵) فرض کنید برای محاسبه $\sqrt[n]{\alpha}$ ، که در آن $\alpha \neq 0$ یک عدد مفروض است، از یک فرمول تکراری به صورت

$$x_{n+1} = px_n + \frac{q\alpha}{x_n^2} + \frac{r\alpha^2}{x_n^5}, \quad n = 0, 1, 2, \dots,$$

استفاده شود که $x_0 \neq 0$ یک تقریب اولیه برای $\sqrt[n]{\alpha}$ است. مقادیر p ، q و r را طوری تعیین کنید که مرتبه همگرایی حداکثر باشد و اگر قرار دهیم $e_i = x_i - \sqrt[n]{\alpha}$ ، آنگاه رابطه بین e_n و e_{n+1} را به دست آورید.

۸.۵ نشان دهید روش تکراری

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)}{2f'(x_n)} \left(\frac{f(x_n)}{f'(x_n)} \right)^2,$$

برای حل معادله $f(x) = 0$ که f به اندازه کافی هموار، $f'(\alpha) \neq 0$ و $f(\alpha) = 0$ ، دارای مرتبه همگرایی حداقل ۳ است. نشان دهید اگر این روند تکراری برای حل معادله $f(x) = x^3 - 3 = 0$ به کار برده شود، آنگاه روند تکراری تمرین ۶.۵ به دست می آید. راهنمایی: فرض کنید

$$u(x) = \frac{f(x)}{f'(x)},$$

و با محاسبه مشتق u ، نشان دهید که می توان x_{n+1} را بر حسب u نوشت.

۹.۵ فرض کنید ماتریس $A \in \mathbb{R}^{n \times n}$ معین مثبت متقارن است و $x = (x_1, x_2, \dots, x_n)^T$ نشان دهید اگر روش نیوتن-رافسون را برای محاسبه مینیمم تابع

$$f(x) = x^T A x + b^T x + c,$$

به کار ببریم، در یک تکرار همگرا می شود، که در آن $b \in \mathbb{R}^n$ و $c \in \mathbb{R}$.

۱۰.۵ فرض کنید مشتقات تابع f تا مرتبه چهارم پیوسته و α یک ریشه ساده f است. روند تکراری

$$x_{n+1} = \frac{1}{4}(y_{n+1} + z_{n+1}), \quad n = 0, 1, 2, \dots,$$

را در نظر بگیرید، که در آن x_0 یک تقریب اولیه برای α است و

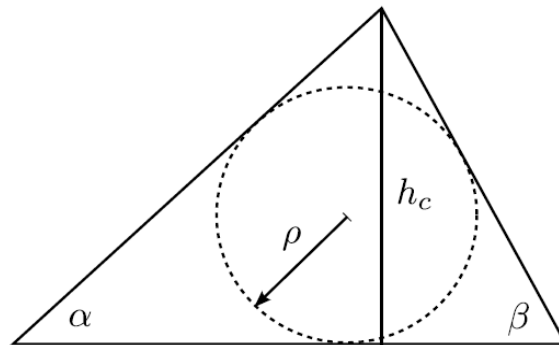
$$y_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad z_{n+1} = x_n - \frac{u(x_n)}{u'(x_n)},$$

به طوری که $u(x) = f(x)/f'(x)$.

۱۱.۵ فرض کنید $A \in \mathbb{R}^{n \times n}$. مسئله مقدار ویژه $Ax = \lambda x$ معادل با حل معادله $f(z) = 0$ است، به طوری که $f: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ و

$$f: \begin{pmatrix} x \\ \lambda \end{pmatrix} \rightarrow \begin{pmatrix} Ax - \lambda x \\ x^T x - 1 \end{pmatrix}.$$

روش نیوتن-رافسون را برای حل معادله $f(z) = 0$ بنویسید.



شکل ۶.۵: شکل تمرین ۱۲.۵.

۱۲.۵ شکل ۶.۵ را در نظر بگیرید. در مثلث داده شده، فرض کنید زاویه β دو برابر زاویه α است. به علاوه، فرض کنید ارتفاع مثلث h_c برابر با ۵ و شعاع دایره محاطی مثلث برابر با $\rho = ۲$ است. اندازه اضلاع مثلث را به دست آورید.

۱۳.۵ دنباله

$$x_{n+1} = \frac{1 + x_n}{1 + e^{x_n}}, \quad n = 0, 1, 2, \dots,$$

را در نظر بگیرید، که در آن x_0 یک حدس اولیه دلخواه است. یک تابع مثل f بیابید که این دنباله از روش نیوتن-رافسون برای حل $f(x) = 0$ به دست آمده باشد.

۱۴.۵ با استفاده از روش نیوتن-رافسون، یک مقدار تقریبی برای x بیابید که

$$\int_0^x e^{-t^2} dt = 0.5.$$

همگرایی روش تکراری

$$x_{n+1} = x_n + 0.5 - \int_0^{x_n} e^{-t^2} dt,$$

را برای حل معادله داده شده بررسی کنید.

۱۵.۵ با دو تکرار روش نیوتن-رافسون یک تقریب برای جواب دستگاه معادلات غیرخطی

$$\begin{cases} 4x^2 + y^2 = 2, \\ 2x + y + 2xy = 3, \end{cases}$$

به دست آورید. حدس اولیه برای جواب دستگاه را $(x^{(0)}, y^{(0)}) = (\frac{1}{4}, 1)$ بگیرید.

کتابنامه

- [1] U.M. Ascher, C. Greif, A First Course in Numerical Methods, SIAM, Philadelphia, 2011.
- [2] R. Bulirsch, H. Ratschauer, Interpolation and Genäherte Quadratur, In: Sauer, Szabó, 1968.
- [3] R.L. Burden, J.D. Faires, A.M. Burden, Numerical Analysis, Tenth Edition, Cengage Learning, 2016.
- [4] W. Cheney and D. Kincaid, Numerical Mathematics and Computing, Sixth Edition, Brooks/Cole Cengage Learning, Florence, Kentucky, 2008.
- [5] R.M. Corless, N. Fillion, A Graduate Introduction to Numerical Methods: From the Viewpoint of Backward Error Analysis, Springer, 2013.
- [6] W. Gander, M. Gander, F. Kwok, Scientific Computing, Springer-Verlag, New York, 2014.
- [7] R.P. Grimaldi, Discrete and Combinatorial Mathematics: An Applied Introduction, Fifth Edition 5th Edition, Pearson Education Limited, 2014.
- [8] J. Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Yale University Press, New Haven, 1923.

- [9] F.B. Hildebrand, Introduction to Numerical Analysis, Second Edition, New York: McGraw-Hill, 1974.
- [10] R. Kress, Numerical Analysis, Springer-Verlag, New York, 1998.
- [11] L.M. Milne-Thomson, The Calculus of Finite Differences, London: Macmillan, 1951.
- [12] G. Peano, *Resto nelle formula di quadratura espresso con un integrale definito*, Atti Della Reale Accademia Dei Lincei, Rendiconti (5), 22 (1913) 562-569.
- [13] A. Ralston, P. Rabinowitz, A First Course in Numerical Analysis, McGraw-Hill, 1978
- [14] T. Sauer, Numerical Analysis, Second Edition, Pearson Education, 2012.
- [15] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Second Edition, Springer-Verlag, New York, 1980.

واژه‌نامه فارسی به انگلیسی

	آ
spline	اسپلاین
Stefensen	استیفنسن
Bernoulli numbers	اعداد برنولی
perturbation	اختلال
min-max	اقل اکثر
error propagation	انباشتگی خطا
integral	انتگرال
integration	انتگرال‌گیری
integrand	انتگرالده
Aitken	ایتکن
algorithm	الگوریتم
	ب
bias	بایاس
ill-posed	بدوضع
ill-posedness	بدوضعی
ill-conditioned	بدحالت
gradient vector	بردار گرادیان
eigenvector	بردار ویژه
Bernoulli	برنولی

Taylor expansion	بسط تیلر
relatively	به‌طور نسبی
continuously	به‌طور پیوسته
absolutely continuous	به‌طور مطلق پیوسته
Bairstow	بیرستو
	پ
stable	پایدار
stability	پایداری
Peano	پئانو
continuous	پیوسته
	ت
primitive function	تابع اولیه
iteration function	تابع تکرار
weight function	تابع وزن
divided difference	تفاضل تقسیم‌شده
reciprocal differences	تفاضلات متقابل
inverse differences	تفاضلات معکوس
almost everywhere	تقریبا همه جا
monic	تکین
	چ
Chebyshev	چبیشف
dense	چگال
ploynomial	چند جمله‌ای
multiple variable	چند متغیره
	ح
initial guess	حدس اولیه
arithmetic	حساب

exact arithmetic	حساب دقیق
catastrophic cancelation	حذف فاجعه‌انگیز
sensitive	حساس
sensitivity	حساسیت
real	حقیقی
خ	
consistency property	خاصیت سازگاری
multiplicative property	خاصیت ضربی
error	خطا
Peano's error	خطای پیانو
absolute error	خطای مطلق
relative error	خطای نسبی
linear	خطی
well-defined	خوش‌تعریف
well-conditioned	خوش‌حالت
well-posed	خوش‌وضع
د	
inscribed circle, incircle	دایره محاطی
pivot element	درایه محوری
interpolation	درونیابی
Newton's interpolation	درونیابی نیوتن
trigonometric interpolation	درونیابی مثلثاتی
IEEE system	دستگاه IEEE
inaccessible	دست‌نیافتنی
single precision	دقت ساده
double precision	دقت مضاعف
Fibonacci sequence	دنباله فیبوناچی

trapezoid	ذوزنقه
modified trapezoidal	ذوزنقه اصلاح شده
	ر
Runge	رونگه
significant digit	رقم بامعنی
Rolle	رل
Bairstow method	روش بیرستو
iterative method	روش تکراری
root	ریشه
	ز
eigenpair	زوج ویژه
underflow	زیرریز
principle submatrix	زیرماتریس اصلی
leading principle submatrix	زیرماتریس اصلی پیشرو
subnormal	زیرنرمال
	ژ
Jacobi	ژاکوبی
	س
global	سراسری
overflow	سرریز
Taylor series	سری تیلر
tridiagonal	سه قطری
Simpson	سیمپسون
	ش
acceleration	شتاب
	ض

inner product	ضرب داخلی
multiplicative	ضربی
ح	
condition number	عدد حالت
science	علوم
floating point number	عدد ممیز شناور
خ	
strictly diagonally dominant	غالب قطری اکید
insensitive	غیر حساس
nonlinear	غیر خطی
ف	
composite formula	فرمول مرکب
vector space	فضای برداری
inner product space	فضای حاصل ضرب داخلی
Fibonacci	فیبوناچی
ق	
trapezoidal rule	قاعده ذوزنقه‌ای
Simpson rule	قاعده سیمپسون
fractional part	قسمت کسری
ک	
complete	کامل
Thiele's continued fraction	کسر مسلسل تیل
Cauchy-Schwarz	کوشی - شوارتز
گ	
Gauss	گوس
Gauss-Laguerre	گوس - لاگر

Gauss-Chebyshev	گائوس - چیشف
Gauss-Hermite	گائوس - هر میت
Gradient	گرادیان
barycentric	گرانیکاهی
rational	گویا
	ل
Laguerre	لاگر
Lagrange	لاگرانژ
generalized Laguerre	لاگرانژ تعمیم یافته
Lipschitz	لیپشیتز
	م
augmented matrix	ماتریس افزوده
Jacobian matrix	ماتریس ژاکوبی
Hessian matrix	ماتریس هسیان
mantissa	مانتیس
orthogonal	متعامد
trigonometric	مثلثاتی
convex	محدب
pivoting	محورگیری
order of convergence	مرتبه همگرایی
order of multiplicity	مرتبه تکرار
composite	مرکب
differential equations	معادلات دیفرانسیل
stopping criterion	معیار توقف
symmetric positive definite	معین مثبت متقارن
mean value	مقدار میانگین
intermediate	مقدار میانی
eigenvalue	مقدار ویژه

moment	ممان
floating point	ممیزشناور
local	موضعی
minimum norm	مینیمم نرم
field	میدان
	ن
regula falsi	نابجایی
unstable	ناپایدار
triangle inequality	نامساوی مثلث
nonsingular	نامنفرد
norm	نرم
vector norm	نرم برداری
infinity norm	نرم بینهایت
natural norm	نرم طبیعی
Frobenius norm	نرم فروبنیوس
matrix norm	نرم ماتریسی
norm-1	نرم-۱
norm-2	نرم-۲
induced norm	نرم القایی
Euclidean norm	نرم اقلیدسی
virtual abscissae	نقاط مجازی
fixed point	نقطه ثابت
exponent	نما
error representation	نمایش خطا
Neville	نویل
Newton	نیوتن
Newton-Raphson	نیوتن - رافسون
Newton-Cotes	نیوتن - کاتس

secant	وتری
	®
Hermite	هرمیت
Hermite-Birkhoff	هرمیت - بیرکف
kernel	هسته
Hessian	هسیان
neighborhood	همسایگی
Horner	هورنر
Holladay	هولادی

واژه‌نامه انگلیسی به فارسی

A

absolute error	خطای مطلق
absolutely continuous	به طور مطلق پیوسته
acceleration	شتاب
Aitken	ایتکن
algorithm	الگوریتم
almost everywhere	تقریباً همه جا
arithmetic	حساب
augmented matrix	ماتریس افزوده

B

Bairstow method	روش بیرستو
barycentric	گرانیکاهی
Bernoulli numbers	اعداد برنولی
bias	بایاس

C

catastrophic cancelation	حذف فاجعه‌انگیز
Cauchy-Schwarz	کوشی - شوارتز
Chebyshev	چبیشف
condition number	عدد حالت

consistency property	خاصیت سازگاری
continuous	پیوسته
convex	محدب
composite formula	فرمول مرکب
complete	کامل
D	
dense	چگال
differentiable	مشتق‌پذیر
divided difference	تفاضل تقسیم‌شده
double precision	دقت مضاعف
E	
eigenpair	زوج ویژه
eigenvalue	مقدار ویژه
eigenvector	بردار ویژه
error	خطا
error propagation	انباشتگی خطا
error representation	نمایش خطا
Euclidean norm	نرم اقلیدسی
exact arithmetic	حساب دقیق
exponent	نما
F	
Fibonacci sequence	دنباله فیبوناچی
field	میدان
fixed point	نقطه ثابت
floating point number	عدد ممیز شناور
fractional part	قسمت کسری

Frobenius norm	نرم فروبنیوس
G	
Gauss	گوس
Gauss-Chebyshev	گوس - چیشف
Gauss-Hermite	گوس - هرمیت
Gauss-Laguerre	گوس - لاگر
generalized Lagrange	لاگرانژ تعمیم یافته
global	سراسری
gradient vector	بردار گرادیان
H	
Hermite-Birkhoff	هرمیت - بیرکف
Hessian matrix	ماتریس هسیان
Holladay	هولادی
Horner	هورنر
I	
IEEE system	دستگاه IEEE
ill-posedness	بد وضعی
ill-conditioned	بد حالت
ill-posed	بد وضع
inaccessible	دست نیافتنی
induced norm	نرم القایی
infinity norm	نرم بینهایت
initial guess	حدس اولیه
inner product	ضرب داخلی
inner product space	فضای حاصل ضرب داخلی
insensitive	غیر حساس
integral	انتگرال
integration	انتگرال گیری

integrand	انتگرالده
intermediate	مقدار میانی
interpolation	درونیابی
inverse differences	تفاضلات معکوس
incircle, inscribed circle	محاطی دایره
iteration function	تابع تکرار
iterative method	روش تکراری

J

Jacobi	ژاکوبی
Jacobian matrix	ماتریس ژاکوبی

K

kernel	هسته
--------	------

L

Lagrange	لاگرانژ
Laguerre	لاگر
leading	پیشرو
linear	خطی
Lipschitz	لیپشیتز
local	موضعی

M

matrix norm	نرم ماتریسی
mantissa	مانتیس
mean value	مقدار میانگین
minimum norm	مینیمم نرم
min-max	اقل اکثر

modified trapezoidal	ذوزنقه اصلاح شده
multiple variable	چندمتغیره
multiplicative property	خاصیت ضربی
moment	ممان
monic	تکین
N	
natural norm	نرم طبیعی
neighborhood	همسایگی
Neville	نویل
Newton	نیوتن
Newton-Cotes	نیوتن - کاتس
Newton-Raphson	نیوتن - رافسون
Newton's interpolation	درون‌یای نیوتن
nonlinear	غیرخطی
nonsingular	نامنفرد
norm	نرم
norm-1	نرم-۱
norm-2	نرم-۲
O	
orthogonal	متعامد
order of multiplicity	مرتبه تکرار
order of convergence	مرتبه همگرایی
overflow	سرریز
P	
Peano's error	خطای پیانو
perturbation	اختلال

pivot element	درایه محوری
pivoting	محورگیری
polynomial	چندجمله‌ای
primitive function	تابع اولیه
principle submatrix	زیرماتریس اصلی

R

rational	گویا
real	حقیقی
reciprocal differences	تفاضلات متقابل
regula falsi	نابجایی
relative error	خطای نسبی
relatively	به‌طور نسبی
Rolle	رل
root	ریشه
Runge	رونگه

S

science	علوم
secant	وتری
sensitive	حساس
sensitivity	حساسیت
sequence	دنباله
significant digit	رقم بامعنی
Simpson rule	قاعده سیمپسون
single precision	دقت ساده
spline	اسپلاین
stable	پایدار
stability	پایداری

Stiefens	استیفنس
stopping criterion	معیار توقف
strictly diagonally dominant	غالب قطری اکید
subnormal	زیرنرمال
symmetric positive definite	معین مثبت متقارن
system	دستگاه

T

Taylor expansion	بسط تیلر
Taylor series	سری تیلر
Thiele's continued fraction	کسر مسلسل تیل
trapezoidal	ذوزنقه
trapezoidal rule	قاعده ذوزنقه‌ای
triangle inequality	نامساوی مثلث
tridiagonal	سه قطری
trigonometric interpolation	درونیابی مثلثاتی

U

underflow	زیرریز
unstable	ناپایدار

V

vector norm	نرم برداری
vector space	فضای برداری
virtual abscissae	نقاط مجازی

W

weight function	تابع وزن
well-conditioned	خوش حالت

well-defined

خوش تعریف

well-posed

خوش وضع

فهرست الفبایی

- بد
حالت، ۴۵
وضع، ۳۸
بدحالتی، ۴۵
بدوضعی، ۴۵
بردار گرادیان، ۱۷
بردار ویژه، ۱
بستار، ۱۹۲
بسط تیلر، ۱۷۴
به طور موضعی، ۱۸۱
به طور نسبی، ۸۶
بیرستو، ۱۹۳
- پایدار، ۳۸
پسرو، ۹
پیوسته، ۱۳
پیوسته - مشتق پذیر، ۱۴۰
پئانو، ۱۳۹
- تابع
اولیه، ۱۴، ۱۲۹
تکرار، ۱۷۴
وزن، ۱۴۹
تحلیلی، ۱۴۴
- اپسیلون ماشین، ۲۷
اسپلاین، ۱۰۵
درونیاب، ۱۰۵
طبیعی، ۱۰۶
متناوب، ۱۰۶
مقید، ۱۰۶
مکعبی، ۱۰۵
استیفنسن، ۲۰۵
اعداد برنولی، ۱۴۴
اعمال سطری مقدماتی، ۸
افزودنی، ۲۲
اقل اکثر، ۶۷
الگوریتم توماس، ۱۱۵
الگوریتم راینش، ۱۱۷
انباشتگی خطا، ۳۹
انتگرال، ۱۲۹
انتگرال گیری، ۱۲۹
انتگرالده، ۱۳۵، ۱۴۰
انحراف معیار، ۴۲
انقباضی، ۱۸۵، ۲۰۸
ایتکن، ۶۱، ۲۰۴
ایده آل، ۱۸۰
بایاس، ۲۲

مطلق، ۲۷	تفاضل تقسیم‌شده، ۶۲
نسبی، ۲۷	تفاضلات
خوش	متقابل، ۹۳
حالت، ۴۵	معکوس، ۹۰
وضع، ۳۸	تقریباً همه جا، ۱۰۶
خوش تعریف، ۱۹۰، ۲۰۵	تکین، ۶۹، ۱۵۰
دایره محاطی، ۲۱۲	تیل، ۹۵
دترمینان، ۱	تیلر، ۱۴
درایه محوری، ۹	جای‌گذاری پسرو، ۹
درونیاب، ۵۴، ۱۳۰	جزءه جزء، ۱۵
نیوتن، ۶۳	چبیشف، ۶۷، ۱۵۸
درونیابی، ۵۳	چگال، ۲۳
مثلثاتی، ۹۶	چند جمله‌ای، ۵۴
دستگاه IEEE، ۲۲	چند جمله‌ای درونیاب، ۵۳
دست‌نیافتنی، ۸۷	چند جمله‌ای مشخصه، ۲
دقت ساده، ۲۲	چندمتغیره، ۱۰۲
دقت مضاعف، ۲۴	حدس اولیه، ۱۷۴
دما، ۱۸۰	حذف فاجعه‌انگیز، ۳۱
دنباله، ۱۷۴	حذفی گاوس، ۸
دنباله فیبوناچی، ۲۰۳	حساب، ۱۳
دوبه‌دو متعامد، ۱۵۷	حساس، ۴۵
دیفرانسیل، ۱۳	حساسیت، ۴۵
رامبرگ، ۱۴۴، ۱۴۶	خاصیت
رتبه، ۱	سازگاری، ۴
رل، ۶۵	ضربی، ۴
روش	خاصیت مینیمم نرم، ۱۰۸
Δ^2 -ایتکن، ۲۰۴	خطا، ۲۷
استیفنسن، ۲۰۵	خطای
بیرستو، ۱۹۳	

- تکراری، ۱۷۳
 رامبرگ، ۱۴۴
 نابجایی، ۱۹۸
 نیوتن-رافسون، ۱۷۵
 وتری، ۲۰۱
 رونگه، ۶۷
 ریشه، ۱۷۴
 ریشه ساده، ۱۵۵
 زوج ویژه، ۱
 زیرریز، ۲۹
 زیرماتریس اصلی، ۶
 پیشرو، ۶
 زیرنرمال، ۲۴
 ژاکوبی، ۱۷
 سراسری، ۱۸۲
 سرریز، ۲۹
 سرعت همگرایی، ۱۷۴
 سری تیلر، ۱۴
 سطری مقدماتی، ۸
 سه قطری، ۱۱۵
 شتاب، ۲۰۴
 شعاع طیفی، ۱
 ضرب داخلی، ۴، ۱۴۹
 ضربی، ۴
 طیف، ۱
 عدد
- حالت، ۴۶
 ممیزشناور، ۲۳
 غالب قطری اکید، ۱۱۴
 غیرحساس، ۴۵
 غیرخطی، ۱۷۳
 فرمول مجموع اولر-مکلوران، ۱۴۴
 فرمول مرکب، ۱۳۵
 فروبنیوس، ۳
 فشار، ۱۸۰
 فضای
 برداری، ۱
 ضرب داخلی، ۴
 فیبوناچی، ۲۰۳
 قاعده
 دوزنقه‌ای، ۱۳۱، ۱۳۳
 دوزنقه‌ای اصلاح شده، ۱۳۳، ۱۳۶
 دوزنقه‌ای اصلاح شده، ۱۳۵
 سیمپسون، ۱۳۲
 نقطه میانی، ۱۶۴
 قاعده هوییتال، ۲۰۶
 قسمت کسری، ۲۱
 کامل، ۱۸۱
 کسر مسلسل تیلر، ۹۵
 کوادراتور گاوس، ۱۴۹
 کوشی-شوارتز، ۵
 گاز، ۱۸۰
 گاوس، ۸، ۱۴۹

- چیشف، ۱۵۹
 - لاگر، ۱۶۹
 - لژاندر، ۱۵۸، ۱۶۲
 - هرمیت، ۱۷۰
 گرادیان، ۱۷
 گرانیگاهی، ۵۵
 گویا، ۸۳
 لاگر، ۱۶۹
 لاگرانژ، ۵۴، ۱۳۰
 تعمیم یافته، ۷۸
 لژاندر، ۱۵۸، ۱۶۲
 لیپشیتز، ۱۰۶
 ماتریس
 ژاکوبی، ۱۷
 واندرموند، ۷، ۱۳۱
 هسیان، ۱۷
 ماتریس افزوده، ۸
 مانتیس، ۲۱
 متعامد، ۱۴۹
 متقارن، ۵
 متناوب، ۱۰۶
 مثلثاتی، ۹۶
 محاطی، ۲۱۲
 محدب، ۱۸۹
 محوری، ۹
 مدل، ۲۱
 مرتبه
 تکرار، ۱۵۳، ۱۸۳
 همگرایی، ۱۸۱
 مرکب، ۱۳۵، ۱۶۷
 مشتق پذیر، ۱۸۸
 مطلقا پیوسته، ۱۰۶
 معادل، ۸۶
 معادلات، ۱۷۳
 معیار توقف، ۱۷۶
 معین مثبت متقارن، ۵
 مقدار
 میانگین، ۱۳
 میانی، ۱۳، ۱۳۶
 ویژه، ۱
 مقدار میانگین، ۲۰۰
 مقید، ۱۰۶
 ممان، ۱۱۰
 ممیز شناور، ۲۳
 موضعی، ۱۸، ۱۸۱
 مولد، ۴
 میانگین، ۴۲
 میدان، ۱
 مینیمم، ۱۸
 موضعی، ۱۸
 نرم، ۱۰۸
 نابجایی، ۱۹۸
 ناپایدار، ۳۸
 نامساوی مثلث، ۲
 نامنفرد، ۱۷۵
 نرم، ۲
 -۱، ۳
 اقلیدسی، ۳

- برداری، ۳
 بینهایت، ۳
 طبیعی، ۴
 فروبنیوس، ۳
 ماتریسی، ۳
 نرم-۲، ۳
 نقاط مجازی، ۸۰
 نقطه ثابت، ۱۷۴
 نقطه میانی، ۱۶۴
 نگاشت انقباضی، ۱۸۵، ۲۰۸
 نما، ۲۱
 نمایش
 خطا، ۶۵
 خطای پئانو، ۱۳۹
 نویل، ۵۶
 نیمه معین مثبت متقارن، ۵
 نیوتن، ۶۳
 - رافسون، ۱۷۵
 - کاتس، ۱۲۹
 واریانس، ۴۲
 واندروالس، ۱۸۰
 وتری، ۲۰۱
 هرمیت، ۷۴، ۱۷۰
 - بیرکف، ۷۹
 هسته پئانو، ۱۳۹
 همسایگی، ۱۷۴، ۱۸۱
 همگرایی، ۱۷۴
 هوپیتال، ۲۰۶
 هورنر، ۶۲
 هولادی، ۱۰۷
 یکتا، ۳۸
 یکتایی، ۳۸

University of Guilan Press

Advanced Numerical Analysis

By:
Davod Khojasteh Salkuyeh, Ph.D