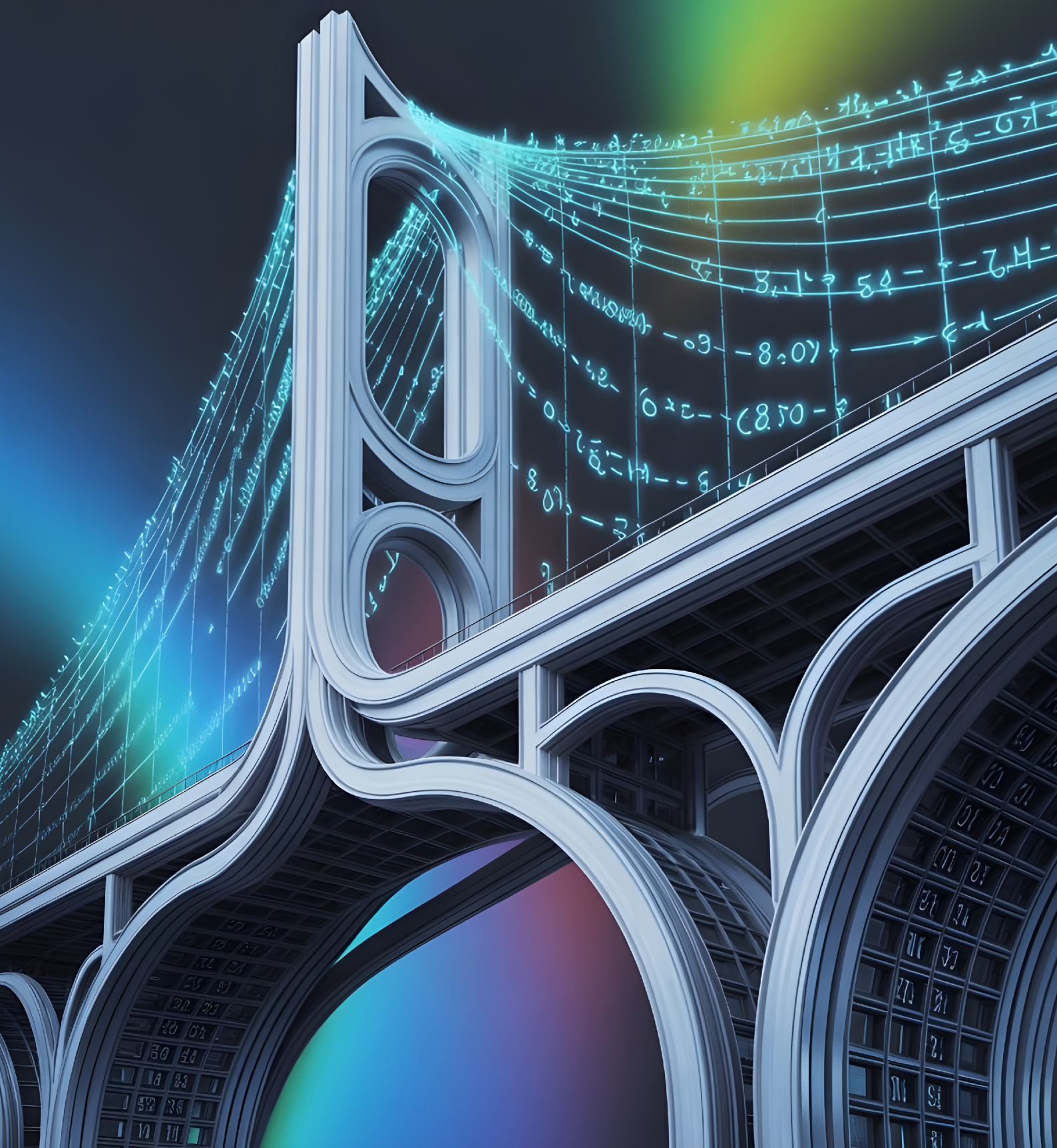


محاسبات علمی ۱

استاد درس: دکتر نجمه حسینی منجری
نگارش: یحیی ایزدی





دانشکده ریاضی و آمار
دانشگاه اصفهان

محاسبات علمی ۱

استاد درس: دکتر نجمه حسینی منجری

نگارش: یحیی ایزدی

نیمسال دوم ۱۴۰۳-۰۴

بہ نام خدا

فهرست مطالب

۱	خطا و پایداری - نمایش ماشینی اعداد	1
۲	1.1 خطا	1.1
۲	1.1.1 منابع خطا	1.1.1
۵	1.1.2 خطای مطلق و نسبی	1.1.2
۶	1.2 نُرم	1.2
۸	1.2.1 نُرم های ماتریسی	1.2.1
۱۰	1.3 ارقام با معنای یک عدد	1.3
۱۰	1.3.1 ارقام با معنای درست یک عدد	1.3.1
۱۲	1.4 نمایش ماشینی اعداد	1.4
۱۴	1.4.1 تغییر مبنا	1.4.1
۱۶	1.4.2 گرد کردن و بریدن	1.4.2
۱۷	1.4.3 نمایش ممیز ثابت و ممیز شناور	1.4.3
۲۱	1.4.4 اپسیلون ماشین	1.4.4
۲۱	1.4.5 نگاشت گرد کردن	1.4.5
۲۳	1.4.6 اعداد زیر نرمال	1.4.6
۲۴	1.4.7 استاندارد IEEE	1.4.7
۲۷	1.5 وضعیت مسئله و پایداری روش	1.5
۲۹	1.5.1 وضعیت یک مسئله ریاضی	1.5.1
۳۰	1.5.2 ضریب وضعیت	1.5.2
۳۳	1.5.3 سازگاری، ناپایداری و همگرایی	1.5.3
۳۵	1.5.4 هزینه محاسباتی	1.5.4
۳۹	2 روش های حل دستگاه های خطی	2
۴۰	2.1 روش های مستقیم	2.1
۴۰	2.1.1 روش کرامر	2.1.1
۴۰	2.1.2 روش حذفی گاوس	2.1.2
۴۰	2.1.3 روش گاوس-ژوردان (جردن)	2.1.3
۴۱	2.1.4 روش های تجزیه ماتریسی	2.1.4
۴۱	2.2 روش های تکراری	2.2

2.3	روش حذفی گاوس	۴۱
2.3.1	محاسبه معکوس یک ماتریس با استفاده از روش حذفی گاوس	۴۵
2.4	محاسبه تجزیه LU ماتریس A	۴۷
2.5	محورگیری جزئی و کلی	۴۹
2.5.1	محورگیری جزئی	۴۹
2.5.2	محورگیری کلی	۴۹
2.6	ماتریس جایگشت	۴۹
2.7	مقیاس‌دهی	۵۰
2.8	روش گاوس-ژوردان (چردن)	۵۰
2.8.1	صورت‌های دیگر تجزیه LU	۵۴
2.8.2	تجزیه LU ماتریس سه‌قطری	۵۸
2.9	تجزیه چولسکی	۵۹
2.9.1	ماتریس معین مثبت	۵۹
2.9.2	تجزیه چولسکی	۶۳
2.10	کمترین مربعات	۶۶
2.10.1	جواب‌های کمترین مربعات توان‌های دوم برای دستگاه‌های خطی	۶۷
2.10.2	تعبیر هندسی مسئله کمترین مربعات توان‌های دوم	۶۸
2.10.3	استفاده از تجزیه QR برای حل مسئله کمترین مربعات	۷۳
2.11	تجزیه QR	۷۴
2.12	تجزیه مقادیر تکین (SVD)	۸۵
2.12.1	کاربرد تجزیه SVD در حل مسئله کمترین مربعات	۸۸
2.13	فرآیند متعامدسازی گرام اشمیت	۹۴
2.13.1	الگوریتم فرآیند گرام اشمیت	۹۴
2.14	تجزیه شور (Schur)	۹۷
2.14.1	روش تکرار QR	۱۰۰
3	روش‌های تکراری حل دستگاه	۱۰۱
3.1	روش ژاکوبی	۱۰۱
3.1.1	فرم فشرده روش ژاکوبی	۱۰۴
3.2	روش گاوس-سایدل	۱۰۵
3.3	بررسی همگرایی روش‌های تکراری	۱۰۷
3.4	تسریع همگرایی روش گاوس-سایدل (روش SOR)	۱۱۵
4	حل معادلات غیرخطی	۱۱۹
4.0.1	رسم منحنی	۱۱۹
4.0.2	جدول مقادیر تابع	۱۲۱
4.1	سرعت همگرایی الگوریتم‌ها	۱۲۲
4.2	روش دوبخشی یا نصف کردن (تتصیف)	۱۲۳
4.2.1	همگرایی روش دوبخشی	۱۲۵
4.3	روش نابجایی	۱۲۶
4.4	روش تکرار ساده (نقطه ثابت)	۱۲۷
4.4.1	همگرایی روش تکرار ساده	۱۲۸

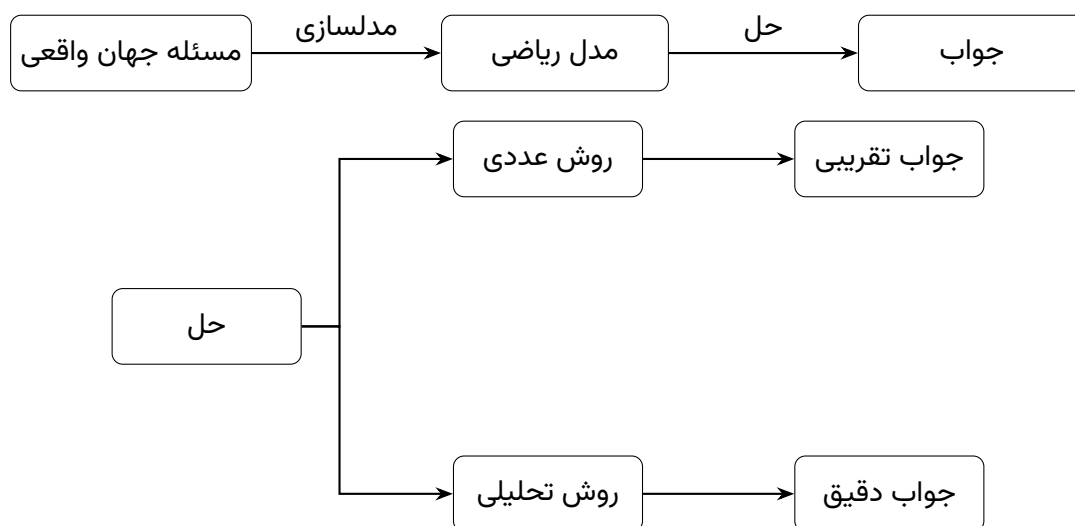
۱۳۰	مرتبۀ همگرایی روش تکرار ساده	4.4.2
۱۳۲	روش نیوتن	4.5
۱۳۲	فرمول تکراری روش نیوتن	4.5.1
۱۳۵	همگرایی روش نیوتن	4.5.2
۱۴۱	روش وتری (خط قاطع)	4.6
۱۴۲	تعبیر هندسی	4.6.1
۱۴۳	حل دستگاه‌های غیرخطی	4.7
۱۴۳	تعمیم روش نیوتن برای حل دستگاه غیرخطی دو معادله دومجهول	4.7.1

۱۴۷	روش‌های عددی مقادیر ویژه	5
۱۴۸	ویژگی‌های مقادیر ویژه و بردارهای ویژه	5.1
۱۵۵	تعیین چندجمله‌ای سرشت نما	5.2
۱۵۵	روش کریلف	5.2.1
۱۵۷	روش لوبریه	5.2.2
۱۵۹	محاسبه بزرگترین مقدار ویژه از لحاظ قدرمطلق	5.3
۱۵۹	روش توانی	5.3.1

خطا و پایداری - نمایش ماشینی اعداد

1

آنالیز عددی، محاسبات علمی و محاسبات عددی در واقع علم تقریب می‌باشند. در اکثر مسائلی که در جهان واقعی با آن‌ها مواجه می‌شویم، معمولاً یک مسئله ریاضی ظاهر می‌شود. گام ابتدایی برای حل هر مسئله این است که یک فرمول‌بندی ریاضی برای مسئله به دست آوریم. این مبحث را **مدلسازی** می‌گوییم. در این درس به مدل‌سازی نمی‌پردازیم.



بعد از اینکه مدل ریاضی مسئله به دست آمد، باید مسئله را حل کنیم و جواب مسئله را به دست آوریم. بعد از محاسبه جواب، آن را تحلیل می‌کنیم و جواب مسئله واقعی را به دست می‌آوریم. گاهی اوقات حل مسئله به سادگی امکان‌پذیر است، پس مسئله را به صورت تحلیلی حل می‌کنیم و جواب دقیق را به دست می‌آوریم. از طرف دیگر در موارد زیادی با مسائلی مواجه می‌شویم که حل آن‌ها به سادگی امکان‌پذیر نیست، ممکن است بسیار پیچیده و پرهزینه باشد و یا ممکن است حل تحلیلی آن امکان‌پذیر نباشد. در این‌جا با روش‌های عددی و تکراری، یک جواب تقریبی برای مسئله به دست می‌آوریم که موضوع اصلی بحث این درس نیز می‌باشد.

معادله $f(x) = 0$:

* مثال 1.1 | مثالی از تکرار

مقدار تقریبی تابع	مقدار
$f(x_0) \approx 10^{-1}$	x_0
$f(x_1) \approx 10^{-2}$	x_1
$f(x_2) \approx 10^{-3}$	x_2
$\vdots$	$\vdots$
$f(x_5) \approx 10^{-6}$	x_5
$\vdots$	$\vdots$
$f(\alpha) = 0$	$x_n \rightarrow \alpha$

نکته 1.1

از نقطه اولیه دلخواه x_0 شروع می‌کنیم و با هر تکرار، خطای تقریب کمتر می‌شود تا به جواب دقیق α همگرا شویم.

1.1 خطا

1.1.1 منابع خطا

خطاها ممکن است منشاهای متفاوتی داشته باشند که شناخت آن‌ها برای کنترل و کاهش‌شان ضروری است. در ادامه به مهم‌ترین این موارد اشاره می‌کنیم:

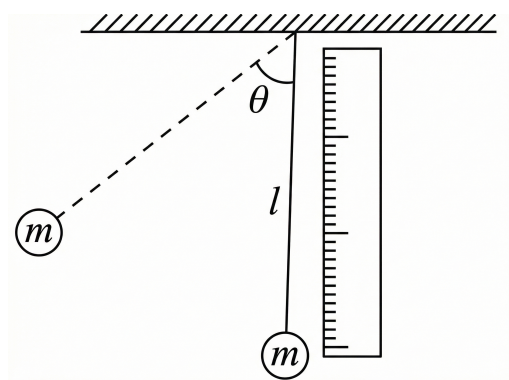
خطای مدل‌سازی

اگرچه در ریاضیات علاقه‌مندیم همه‌چیز را ایده‌آل ببینیم، در فیزیک و علوم کاربردی ناچاریم ساده‌سازی‌هایی را در نظر بگیریم. این ساده‌سازی‌ها باعث ایجاد خطا در جواب نهایی می‌شوند. خطای حاصل از این موارد را **خطای مدل‌سازی** می‌گوییم.

برای اینکه فرمول‌بندی مسئله امکان پذیر باشد لازم است فرضیاتی

* مثال 1.2 | مطالعه حرکت یک آونگ

برای مسئله در نظر بگیریم. به عنوان مثال:



- از مقاومت هوا چشم‌پوشی می‌کنیم.
 - نخ را بدون جرم در نظر می‌گیریم.
 - از اصطکاک محل اتصال چشم‌پوشی می‌کنیم.
- در انتها مسئله به صورت یک معادله دیفرانسیل مرتبه دوم فرمول‌بندی می‌شود.

* مثال 1.3

- در معادله گرما و میله، میله را بسیار نازک در نظر می‌گیریم، در صورتی که در واقعیت این‌چنین نیست و میله دارای سطح مقطع می‌باشد.
- در مطالعه مباحث اقتصادی، نرخ علاقه‌مندی مردم به یک کالای خاص در یک بازه زمانی ثابت در نظر گرفته می‌شود در صورتی که این شرط همواره برقرار نیست.

خطای داده ورودی

هر مسئله دارای تعدادی ورودی و خروجی می‌باشد. ورودی‌ها معمولاً مقادیری هستند که ناشی از اندازه‌گیری می‌باشند. اندازه‌گیری‌ها توسط وسایل اندازه‌گیری انجام می‌شود که ممکن است چندان دقیق نباشد و همراه با خطا باشد. خطاهای این‌چنینی را **خطای داده ورودی** می‌گوییم. مانند اندازه‌گیری θ و l و m در مثال آونگ.

خطای نمایش اعداد

گاهی اوقات نمی‌توانیم عدد را به صورت دقیق نمایش دهیم. به عنوان مثال، عدد π دارای بینهایت رقم است. در محاسبات نمی‌توانیم بینهایت رقم را به کار بگیریم و مجبور هستیم تنها تعداد متناهی از ارقام را به کار بگیریم. خطای حاصل از این‌گونه موارد را **خطای نمایش اعداد** می‌نامیم. گاهی اوقات اعداد ورودی دقیق می‌باشند اما انجام یکی از چهار عمل اصلی ریاضی ($+$, $-$, $\times$, $\div$) روی آن‌ها منجر به عددی با بینهایت رقم (یا تعداد زیادی رقم) می‌شود که قابل نمایش نمی‌باشد و ما مجبور هستیم تنها تعداد متناهی از ارقام را نگه داریم و بقیه ارقام را دور بریزیم. خطای حاصل از آن را نیز خطای نمایش اعداد می‌گوییم.

* مثال 1.4

۱. مثالی از بینهایت رقم اعشار:

$$22 \div 7 = 3.1428571428\dots$$

۲. فرض کنیم محدودیت دو رقم اعشار وجود دارد:

$$0.91 \div 2.08 = 0.4375 \rightarrow 0.44$$

خطای برشی یا خطای گسسته سازی

خطای برشی

برای محاسبه مقادیر مختلف، ممکن است از فرمول‌های تقریبی استفاده کنیم. به عنوان مثال، زمانی که می‌خواهیم مقدار تابع $f(x)$ را در نقطه β با استفاده از سری تیلور تابع محاسبه کنیم، مجبور هستیم تعداد متناهی جمله از سری تیلور را در نظر بگیریم. در اینجا خطای حاصل در جواب را **خطای برشی** می‌نامیم.

یادآوری 1.1 | سری تیلور

سری تیلور یک تابع را به صورت سری بی‌نهایت نمایش می‌دهد که بر اساس مشتقات تابع در یک نقطه خاص (مثلاً a) محاسبه می‌شود که در آن $f^{(i)}(a)$ مشتق i ام تابع در نقطه a است.

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!} (x-a)^i$$

* مثال 1.5 | خطای برشی در سری تیلور

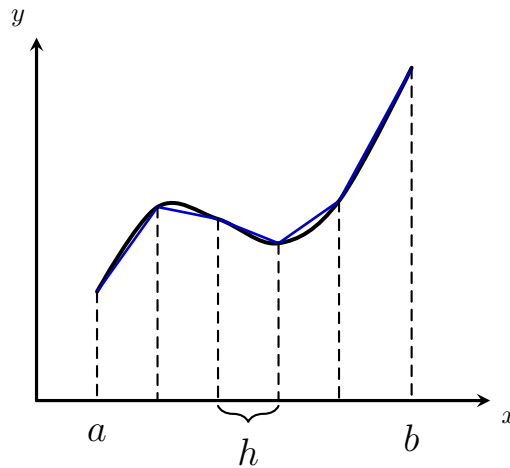
$$f(\beta) \simeq \sum_{i=0}^m \frac{f^{(i)}(a)}{i!} (\beta-a)^i$$

خطای گسسته‌سازی

* مثال 1.6 | روش ذوزنقه‌ای انتگرال گیری

به عنوان مثالی دیگر، زمانی که بخواهیم مقدار انتگرال معین f را در بازه $[a, b]$ با استفاده از مساحت زیر نمودار تابع تقریب بزنیم، دچار خطای گسسته‌سازی هستیم.

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a)$$



$$\begin{aligned}\int_a^b f(x)dx &\simeq \frac{h}{2}(f(x_0) + f(x_1)) + \frac{h}{2}(f(x_1) + f(x_2)) + \dots + \frac{h}{2}(f(x_{n-1}) + f(x_n)) \\ &= \frac{h}{2}(f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + f(x_n))\end{aligned}$$

نکته 1.2 | اشتباه

در همه کارهای عددی، احتمال اشتباه وجود دارد که ممکن است در بخش‌های مختلف انجام کار، مرتکب آن شویم. اشتباه‌ها را جزو منابع خطا در نظر نمی‌گیریم.

نکته 1.3

خطای برشی و گسسته‌سازی، کنترل‌شدنی است. با توجه به اهمیت مسئله، می‌توانیم میزان خطا را کنترل کنیم.

1.1.2 < خطای مطلق و نسبی

خطای مطلق

فرض کنیم $\hat{x}$ تقریبی از عدد x باشد. می‌خواهیم دو متر برای تعیین میزان خطای $\hat{x}$ معرفی کنیم. خطای مطلق $\hat{x}$ به صورت زیر تعریف می‌شود:

$$E_{abs}(\hat{x}) = |\hat{x} - x|$$

* مثال 1.7

$$x_1 = 1, \quad \hat{x}_1 = 3 \quad \rightarrow \quad E_{abs}(\hat{x}_1) = 2$$

$$x_2 = 10000, \hat{x}_2 = 10002 \rightarrow E_{abs}(\hat{x}_2) = 2$$

خطای نسبی

در مثال بالا خطای مطلق هر دو گروه یکسان است، در صورتی که می‌دانیم عدد $\hat{x}_2$ تقریب بهتری از x_2 نسبت به $\hat{x}_1$ برای x_1 است. در واقع، خطای مطلق بزرگی و کوچکی اعداد را در نظر نمی‌گیرد؛ پس در اینجا خطای نسبی را معرفی می‌کنیم:

$$E_{rel}(\hat{x}) = \frac{|\hat{x} - x|}{|x|}, x \neq 0$$

* مثال 1.8

$$x_1 = 1, \hat{x}_1 = 3 \rightarrow E_{rel}(\hat{x}_1) = 2$$

$$x_2 = 10000, \hat{x}_2 = 10002 \rightarrow E_{rel}(\hat{x}_2) = 2 \times 10^{-4}$$

نکته 1.4

زمانی که $\hat{x}, x \in \mathbb{R}^n$ و $\hat{x}$ تقریبی از x باشد، خطای مطلق و نسبی با استفاده از نُرم تعریف می‌شود.

$$E_{abs}(\hat{x}) = \|\hat{x} - x\|$$

$$E_{rel}(\hat{x}) = \frac{\|\hat{x} - x\|}{\|x\|}, \|x\| \neq 0$$

1.2 نُرم

تعریف 1.1

تابع $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ که دارای سه ویژگی زیر باشد را یک نرم می‌گوییم.

$$1. \quad \|x\| \geq 0, \forall x \in \mathbb{R}^n, \quad \|x\| = 0 \Leftrightarrow x = 0$$

$$2. \quad \|\alpha x\| = |\alpha| \|x\|, \forall \alpha \in \mathbb{R}, x \in \mathbb{R}^n$$

$$3. \quad \|x + y\| \leq \|x\| + \|y\|, \forall x, y \in \mathbb{R}^n$$

چهار نرم معروف روی $\mathbb{R}^n$ به صورت زیر تعریف می‌شوند:

$$1. \quad \|x\|_2 = \|x\| = (\sum_{i=1}^n x_i^2)^{1/2}$$

$$2. \quad \|x\|_1 = \sum_{i=1}^n |x_i|$$

$$3. \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

$$\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}, p \geq 1, p \in \mathbb{N} \quad ۴.$$

تمرین 1.1

نشان دهید چهار مورد معرفی شده در فوق، نُرم هستند.

یک خاصیت مهم، نرم‌ها p -نامساوی هولدر است:

نکته 1.5 | نامساوی هولدر

$$|x^T y| \leq \|x\|_p \|y\|_q \quad \frac{1}{p} + \frac{1}{q} = 1$$

در حالتی که $p = q = 2$ باشد، آنگاه نامساوی کوشی-شوارتز است:

نکته 1.6 | نامساوی کوشی-شوارتز

$$|x^T y| \leq \|x\| \|y\|$$

تعریف 1.2 | دو نرم معادل

گوییم دو نرم α, β روی $\mathbb{R}^n$ معادل هستند هرگاه مقادیر ثابت A, B وجود داشته باشد به طوری که:

$$A\|x\|_\alpha \leq \|x\|_\beta \leq B\|x\|_\alpha$$

• تمام نرم‌های برداری در فضای $\mathbb{R}^n$ معادل هستند.

۳ نرم معادل:

$$1. \quad \|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

$$2. \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

$$3. \quad \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

اثبات ۱

$$\|x\|_\infty = \max |x_i| = |x_k|, \quad 1 \leq k \leq n$$

$$|x_k| \leq |x_1| + |x_2| + \cdots + |x_n| = \|x\|_1$$

$$\leq |x_k| + |x_k| + \cdots + |x_k| = n|x_k| = n\|x_k\| \quad \blacksquare$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty \quad \checkmark \text{ اثبات } | ۲$$

$$\begin{aligned} \|x\|_\infty &= \max |x_i| = |x_k|, \quad 1 \leq k \leq n \\ |x_k| &\leq \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} = \|x\|_2 \\ &\leq \sqrt{|x_k|^2 + |x_k|^2 + \dots + |x_k|^2} \\ &= \sqrt{n \cdot |x_k|^2} \\ &= \sqrt{n} \cdot |x_k| = \sqrt{n} \cdot \|x\|_\infty \quad \blacksquare \end{aligned}$$

1.2.1 < نُرم های ماتریسی

فرض کنیم A یک ماتریس $m \times n$ باشد، هر تابع $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ که در ۳ خاصیت صدق کند، نرم ماتریسی است.

1.1 ویژگی‌ها

$$\begin{aligned} ۱. \quad & \|A\| \geq 0, \quad \|A\| = 0 \Leftrightarrow A = O \\ ۲. \quad & \|\alpha A\| = |\alpha| \|A\|, \quad \forall \alpha \in \mathbb{R}, \forall A \in \mathbb{R}^{m \times n} \\ ۳. \quad & \|A + B\| \leq \|A\| + \|B\| \end{aligned}$$

نُرم های ماتریسی وابسته

1.3 تعریف

فرض کنیم A یک ماتریس $m \times n$ و $\|\cdot\|_v$ یک نرم برداری باشد. متناظر با این نرم برداری، یک نرم ماتریسی به صورت زیر تعریف می‌کنیم.

$$\|A\|_M = \max_{x \in \{\mathbb{R}^n \setminus \{0\}\}} \frac{\|Ax\|_v}{\|x\|_v}$$

1.7 نکته

هر نرم ماتریسی وابسته به یک نرم برداری در خاصیت زیر صدق می‌کند:

$$\|Ax\|_v \leq \|A\|_M \cdot \|x\|_v$$

اثبات ✓

$$\|A\|_M = \max \frac{\|Ax\|_v}{\|x\|_v} \geq \frac{\|Ax\|_v}{\|x\|_v}$$

$$\rightarrow \|A\|_M \cdot \|x\|_v \geq \|Ax\|_v \quad \blacksquare$$

نکته 1.8 💡

تعریف های معادل زیر را برای نرم های ماتریسی داریم

$$\|A\|_1 = \max \sum_{i=1}^m |a_{ij}|$$

حداکثر جمع ستون ها

$$\|A\|_\infty = \max \sum_{j=1}^n |a_{ij}|$$

حداکثر جمع سطرها

$$\|A\|_2 = \sqrt{p(A^T A)}$$

 p بیشترین مقدار ویژه

$$(\det(A - \lambda I)) = 0, \quad \max \lambda = p$$

نرم فروبنیوس (مستقل)

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

نکته 1.9

فرض کنیم A یک ماتریس $m \times n$ باشد

$$\frac{1}{\sqrt{n}} \|A\|_{\infty} \leq \|A\|_2 \leq \sqrt{m} \|A\|_m$$

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$$

$$\frac{1}{\sqrt{m}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1$$

$$\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_{\infty}}$$

تعریف 1.4

ماتریس متعامد ماتریس $Q \in \mathbb{R}^{n \times n}$ را یک ماتریس متعامد گوییم هرگاه

$$Q^T Q = I$$

1.3 ارقام با معنای یک عدد

ارقام بامعنای یک عدد شامل اولین رقم غیرصفر از (سمت چپ) و ارقام بعد از آن (هم صفر و هم غیرصفر) است. به عنوان مثال، عدد 0.00123 دارای ۳ رقم بامعنا و 1.0541 دارای ۵ رقم بامعنا است.

* مثال 1.9

۱.

$$x_1 = 1.0000, \hat{x}_1 = 1.00499, E_{rel}(\hat{x}_1) = 4.99 \times 10^{-3}$$

$\hat{x}_1$ تا ۳ رقم با معنا با x_1 مساوی است.

۲.

$$x_2 = 9.0000, \hat{x}_2 = 8.99899, E_{rel}(\hat{x}_2) = 1.11 \times 10^{-4}$$

$\hat{x}_2$ تا ۵ رقم با معنا با x_2 مساوی است.

1.3.1 ارقام با معنای درست یک عدد

همانگونه که در مثال فوق مشاهده کردیم، ارقام با معنای یک عدد، معیار مناسبی برای تعیین تقریب عدد نمی‌باشد و در ادامه ارقام با معنای درست را معرفی می‌کنیم. گوییم $\hat{x}$ تقریبی از x با t رقم با معنای درست می‌باشد، هرگاه $\hat{x}$ و

x به یک عدد سوم یکسان با t رقم با معنا گرد شوند.

* مثال 1.10

ارقام با معنا	یک رقم	دو رقم	سه رقم	چهار رقم
$x_1 = 1.0000$	$\bar{x}_1 = 1$	$\bar{x}_1 = 1.0$	$\bar{x}_1 = 1.00$	$\bar{x}_1 = 1.000$
$\hat{x}_1 = 1.00499$	$\tilde{x}_1 = 1$	$\tilde{x}_1 = 1.0$	$\tilde{x}_1 = 1.00$	$\tilde{x}_1 = 1.005$
برابری	✓	✓	✓	×

* مثال 1.11

ارقام با معنا	یک رقم	دو رقم	سه رقم	چهار رقم
$x_2 = 9.00000$	$\bar{x}_2 = 9$	$\bar{x}_2 = 9.0$	$\bar{x}_2 = 9.00$	$\bar{x}_2 = 9.000$
$\hat{x}_2 = 8.99899$	$\tilde{x}_2 = 9$	$\tilde{x}_2 = 9.0$	$\tilde{x}_2 = 9.00$	$\tilde{x}_2 = 8.999$
برابری	✓	✓	✓	×

* مثال 1.12

ارقام با معنا	یک رقم	دو رقم	سه رقم
$x_3 = 0.9951$	$\bar{x}_3 = 1.0$	$\bar{x}_3 = 1.00$	$\bar{x}_3 = 0.995$
$\hat{x}_3 = 0.9949$	$\tilde{x}_3 = 1.0$	$\tilde{x}_3 = 0.99$	$\tilde{x}_3 = 0.995$
برابری	✓	×	✓

در مثال قبل مشاهده کردیم $\hat{x}$ و $\bar{x}$ دارای یک رقم و سه رقم با معنای درست هستند. حال می‌خواهیم تغییری در تعریف ایجاد کنیم که تعداد ارقام با معنای درست یک عدد به صورت یکتا تعیین شود.

تعریف 1.5

می‌گوییم $\hat{x}$ تقریبی از x با t رقم با معنای درست است، هرگاه t بزرگ‌ترین عدد طبیعی باشد که در رابطه زیر صدق کند:

$$\frac{|\hat{x} - x|}{|x|} \leq \frac{1}{2} \times 10^{-t}$$

برای مثال قبل داریم:

* مثال 1.13

$$\begin{aligned}\frac{|\bar{x} - \hat{x}|}{|\bar{x}|} &= \frac{|0.9951 - 0.9949|}{|0.9949|} \\ &\approx 2 \times 10^{-4} \\ &\leq \frac{1}{2} \times 10^{-3} \\ &\rightarrow t = 3\end{aligned}$$

1.4 نمایش ماشینی اعداد

در زندگی روزمره برای نمایش اعداد و محاسبات از مبنای ۱۰ استفاده می‌کنیم. بنابراین برای نمایش از ارقام $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ بهره می‌بریم. ارزش هر عدد به میزان کاراکترهایش و محل قرارگیری آن‌ها بستگی دارد. به عنوان مثال:

* مثال 1.14

$$1024.576 = 1 \times 10^3 + 0 \times 10^2 + 2 \times 10^1 + 4 \times 10^0 + 5 \times 10^{-1} + 7 \times 10^{-2} + 6 \times 10^{-3}$$

💡 نکته 1.10

هر عدد حقیقی دارای نمایش منحصر به فرد به صورت اعشاری است، مگر اینکه بی‌نهایت بار رقم 9 بعد از ممیز تکرار شود. یعنی مثلاً:

$$3.15, 3.149999\dots$$

نمایش عدد یکسانی می‌باشند.

نکته 1.11

هر عدد صحیح $\beta \geq 2$ ($\beta \leq -2$) می‌تواند به عنوان مبنای محاسبات و نمایش اعداد در نظر گرفته شود. اگر مبنا β باشد، ارقام مورد نیاز جهت نمایش اعداد به صورت:

$$\{0, 1, 2, \dots, \beta - 1\}$$

می‌باشند. هرچه مبنا کوچک‌تر باشد، محاسبات ساده‌تر است. بنابراین مبنای 2 ساده‌ترین حالت است که با استفاده از ارقام $\{0, 1\}$ قابل نمایش و محاسبه است. علاوه بر این، کامپیوتر و ماشین‌های محاسباتی همگی از مبنای 2 استفاده می‌کنند، هم به دلیل سادگی و هم اینکه نمایش دو حالت به سادگی در یک ماشین قابل نمایش است (لامپ روشن و خاموش می‌تواند دو حالت را نمایش دهد و...).

نکته 1.12

هر عدد حقیقی $a \in \mathbb{R}$ را می‌توان در مبنای β نمایش داد. یعنی:

$$\begin{aligned} a &= (d_{n-1}d_{n-2}\dots d_1d_0|d_{-1}d_{-2}\dots)_\beta \\ &= d_{n-1} \times \beta^{n-1} + d_{n-2} \times \beta^{n-2} + \dots + d_1 \times \beta^1 + d_0 \times \beta^0 + d_{-1} \times \beta^{-1} + d_{-2} \times \beta^{-2} + \dots \\ d_i &\in \{0, 1, \dots, \beta - 1\} \end{aligned}$$

* مثال 1.15 | تبدیل ۵۲۴.۱۷۶ به مبنای ۳

تبدیل قسمت صحیح (176) ابتدا عدد 176 را بر ۳ تقسیم می‌کنیم:

باقی مانده	خارج قسمت
2	58
1	19
1	6
0	2
2	0

پس از خواندن باقیمانده‌ها از پایین به بالا، نتیجه:

$$176_{10} = (20112)_3$$

محاسبه مقدار 20112_3 بر مبنای دسیمال (جهت اطمینان):

$$\begin{aligned}
 20112_3 &= 2 \times 3^4 + 0 \times 3^3 + 1 \times 3^2 + 1 \times 3^1 + 2 \times 3^0 \\
 &= 2 \times 81 + 0 \times 27 + 1 \times 9 + 1 \times 3 + 2 \times 1 \\
 &= 162 + 0 + 9 + 3 + 2 \\
 &= 176
 \end{aligned}$$

تبدیل قسمت اعشاری (0.524) برای قسمت اعشاری 0.524، هر بار در ۳ ضرب می‌کنیم و بخش صحیح را به عنوان رقم بعدی در مبنای ۳ برمی‌داریم:

قسمت اعشاری	قسمت صحیح	ضرب در ۳	مقدار اولیه
0.524	1	1.572	0.524
0.572	1	1.716	0.572
0.148	2	2.148	0.716
0.444	0	0.444	0.148
0.332	1	1.332	0.444
0.996	0	0.996	0.332
...	...	...	...

بنابراین قسمت اعشاری در مبنای ۳ به صورت زیر خواهد بود:

$$0.524_{10} = (0.11201\dots)_3$$

در نتیجه، عدد کامل در مبنای ۳:

$$176.524_{10} = (20112.11201\dots)_3$$

نکته 1.13

مثال فوق نشان می‌دهد یک عدد با تعداد ارقام متناهی در یک مبنای دیگر ممکن است بی‌نهایت رقم داشته باشد.

1.4.1 تغییر مبنا

هر عدد حقیقی $a \in \mathbb{R}$ را می‌توانیم به صورت زیر بر مبنای دلخواه β که $\beta \geq 2$ یا $\beta \leq -2$ و $\beta \in \mathbb{Z}$ است بنویسیم.

$$\begin{aligned}
 a &= \pm(d_{n-1}d_{n-2}\dots d_1d_0/d_{-1}d_{-2}\dots)_\beta \\
 &= d_{n-1}\beta^{n-1} + d_{n-2}\beta^{n-2} + \dots + d_1\beta^1 + d_0\beta^0 \\
 &\quad + d_{-1}\beta^{-1} + d_{-2}\beta^{-2} + \dots
 \end{aligned}$$

$$\text{where } d_i \in \{0, 1, \dots, \beta - 1\}$$

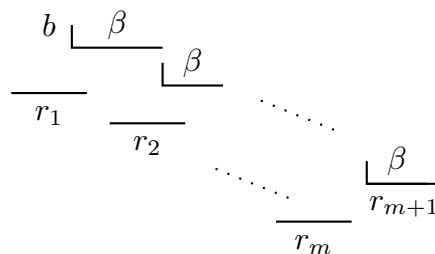
هر عدد حقیقی دارای یک نمایش منحصر به فرد به صورت بالا می‌باشد، مگر اینکه بعد از ممیز، بی‌نهایت رقم $\beta - 1$ تکرار شود.

هر عدد $a \in \mathbb{R}$ را از یک مبنا می‌توانیم به مبنای دلخواه β تبدیل کنیم. برای این منظور، قسمت اعشاری و قسمت صحیح عدد را به صورت جداگانه می‌نویسیم. هر کدام از آن‌ها را به صورت جداگانه بر مبنای β می‌نویسیم.

$$a = b + c, \quad b \in \mathbb{Z}, \quad 0 < c < 1$$

تبدیل قسمت صحیح (b)

برای تبدیل قسمت صحیح، تقسیم‌های متوالی بر β را انجام می‌دهیم تا جایی که خارج قسمت کوچکتر یا مساوی $\beta - 1$ باشد.



آنگاه رقم موجود در خارج قسمت را می‌نویسیم و از سمت راست به چپ رقم‌های باقی‌مانده را در کنار آن می‌نویسیم. عدد حاصل نمایش b بر مبنای β می‌باشد.

$$b = (r_{m+1}r_m \dots r_2r_1)_\beta$$

$$\text{where } r_i \in \{0, 1, \dots, \beta - 1\}$$

تبدیل قسمت اعشاری (c)

برای قسمت اعشاری، در واقع ضرب‌های متوالی در β را انجام می‌دهیم و هر بار قسمت صحیح حاصل از حاصل ضرب را جدا کرده و به عنوان رقم موجود در نمایش بر مبنای β می‌نویسیم و باقی‌مانده حاصل ضرب را مجدداً در β ضرب می‌کنیم و این فرایند را چندین مرحله انجام می‌دهیم.

• ممکن است بعد از چند تکرار متناهی، باقی‌مانده صفر شود. در این صورت نمایش قسمت اعشاری در مبنای β

مختوم می‌باشد و عملیات متوقف می‌شود.

- ممکن است بعد از چند مرحله یک یا چند رقم به صورت نامتناهی تکرار شوند. در این صورت نمایش عدد نامختوم می‌شود و یک یا چند رقم تکراری داریم. در اینجا بعد از چند مرحله فرایند را متوقف می‌کنیم.
- در حالت آخر ممکن است تعداد ارقام متناهی نباشد و نظم خاصی در تکرار ارقام نیز نداشته باشیم. در این حالت بعد از چند مرحله الگوریتم را متوقف می‌کنیم.

◀ 1.4.2 گرد کردن و بریدن

اگر در محاسبات با عددی برخورد کردیم که تعداد ارقام بعد از ممیز آن از تعداد مورد نظر ما بیشتر است، مجبوریم به تعداد مورد نظر از ارقام عدد را نگه داریم و بقیه را دور بریزیم. برای این منظور از دو روش استفاده می‌کنیم:

📖 تعریف 1.6 | روش قطع کردن یا بریدن

در روش قطع کردن تا رقمی که می‌خواهیم نگه داریم جدا می‌کنیم و بقیه ارقام را بدون توجه به کوچکی و بزرگی آن‌ها دور می‌ریزیم.

* مثال 1.16

$$\pi = 3.141592\dots$$

$$\pi \simeq 3.14$$

$$\pi \simeq 3.141$$

$$\pi \simeq 3.1415$$

📖 تعریف 1.7 | روش گرد کردن

در روش گرد کردن تا رقمی که می‌خواهیم نگه داریم را جدا می‌کنیم و اولین رقم ناخواسته را مشخص می‌کنیم. اگر این رقم به صورت اکید از ۵ کوچکتر باشد، ارقام ناخواسته را دور می‌ریزیم و ارقام عدد را بدون تغییر نگه می‌داریم. اگر این رقم بزرگتر یا مساوی ۵ باشد، یک واحد به آخرین رقم نگه داشته شده اضافه می‌کنیم.

* مثال 1.17

$$\pi = 3.141592\dots$$

$$\pi \simeq 3.14$$

$$\pi \simeq 3.142$$

$$\pi \simeq 3.1416$$

◀ 1.4.3 نمایش ممیز ثابت و ممیز شناور

نمایش ممیز ثابت

در دستگاه نمایش ممیز ثابت به خاطر حافظه محدود کامپیوتر، مجبور هستیم تعداد t رقم برای ارقام بعد از ممیز و تعداد n رقم برای ارقام قبل از ممیز در نظر بگیریم. پس فرم کلی اعداد این دستگاه به صورت زیر است:

$$a = \pm (d_{n-1}d_{n-2}\dots d_1d_0/d_{-1}d_{-2}\dots d_{-t})_\beta$$

$$d_i \in \{0, 1, \dots, \beta - 1\}, \quad -t \leq i \leq n-1, \quad i \in \mathbb{Z}$$

ویژگی‌های دستگاه ممیز ثابت

این دستگاه را با $\mathbb{F}_0(\beta, n, t)$ نمایش می‌دهیم. ویژگی‌های این دستگاه به صورت زیر است:

- بزرگترین عدد در این دستگاه دارای نمایشی به صورت زیر است:

$$\begin{aligned} x_{max} &= \pm \underbrace{(\beta-1)(\beta-1)\dots(\beta-1)}_{n \text{ digits}} / \underbrace{(\beta-1)(\beta-1)\dots(\beta-1)}_{t \text{ digits}}_\beta \\ &= (\beta-1) \times (\beta^{n-1} + \beta^{n-2} + \dots + \beta^{-t}) \\ &= (\beta-1) \times \sum_{i=-t}^{n-1} \beta^i \\ &= (\beta-1) \times \frac{\beta^n - \beta^{-t}}{\beta - 1} = \beta^n - \beta^{-t} \simeq \beta^n \end{aligned}$$

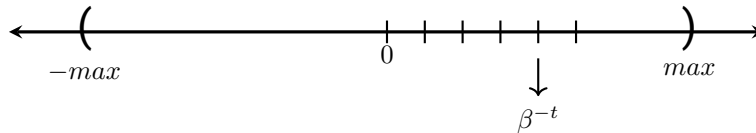
- کوچکترین عدد مثبت غیرصفر در این دستگاه دارای نمایشی به صورت زیر است:

$$x_{min} = + \underbrace{(00\dots 0)}_{n \text{ digits}} / \underbrace{(00\dots 0 1)}_{t-1 \text{ digits}}_\beta = \beta^{-t}$$

- تعداد اعداد قابل نمایش در این دستگاه:

$$2\beta^{n+t} - 1$$

• فاصله اعداد در این دستگاه مقدار ثابت β^{-t} می‌باشد و این امر باعث می‌شود حساسیت روی اعداد کوچک و بزرگ یکسان باشد. این امر باعث می‌شود تعداد اعداد قابل نمایش چندان بزرگ نباشد. فرض کنیم x و x_+ دو عدد متوالی در این دستگاه باشند می‌خواهیم فاصله این دو عدد را محاسبه کنیم.



$$\begin{aligned} x &= (d_{n-1}d_{n-2}\dots d_1d_0/d_{-1}d_{-2}\dots d_{-t})_\beta \\ x_+ &= (d_{n-1}d_{n-2}\dots d_1d_0/d_{-1}d_{-2}\dots (d_{-t} + 1))_\beta \\ x_+ - x &= \underbrace{(0\dots 0)}_n / \underbrace{(0\dots 0 1)}_{t-1}_\beta = \beta^{-t} \end{aligned}$$

همان‌گونه که که ملاحظه می‌کنیم فاصله دو عدد متوالی به عدد در نظر گرفته شده بستگی ندارد و تنها به تعداد ارقام بعد از ممیز بستگی دارد.

دستگاه ممیز ثابت $\mathbb{F}_0(2, 4, 2)$ را در نظر بگیرید. اعداد نیمه مثبت این دستگاه را بنویسید.

* مثال 1.18

$$\beta = 2, n = 4, t = 2$$

$$a \in \mathbb{F}_0(2, 4, 2), a = (d_3 d_2 d_1 d_0 \mid d_{-1} d_{-2})_2$$

$$d_i \in \{0, 1\}$$

$$x_{min} = x_1 = (0000/01)_2 = 1 \times 2^{-2} = \frac{1}{4}$$

$$x_2 = (0000/10)_2 = 1 \times 2^{-1} = \frac{1}{2}$$

$$x_3 = (0000/11)_2 = (1 \times 2^{-1} + 1 \times 2^{-2}) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$x_4 = (0001/00)_2 = 1 \times 2^0 = 1$$

$$x_5 = (0001/01)_2 = 1 \times 2^0 + 1 \times 2^{-2} = 1 + \frac{1}{4} = \frac{5}{4}$$

$$x_6 = (0001/10)_2 = 1 \times 2^0 + 1 \times 2^{-1} = 1 + \frac{1}{2} = \frac{3}{2}$$

$$x_7 = (0001/11)_2 = 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2} = 1 + \frac{1}{2} + \frac{1}{4} = \frac{7}{4}$$

$$x_8 = (0010/00)_2 = 1 \times 2^1 = 2$$

$$x_9 = (0010/01)_2 = 1 \times 2^1 + 1 \times 2^{-2} = 2 + \frac{1}{4} = \frac{9}{4}$$

$$x_{10} = (0010/10)_2 = 1 \times 2^1 + 1 \times 2^{-1} = 2 + \frac{1}{2} = \frac{5}{2}$$

⋮

$$x_{max} = x_{63} = (1111/11)_2 = 8 + 4 + 2 + 1 + \frac{1}{2} + \frac{1}{4} = \frac{63}{4}$$

نمایش ممیز شناور

در اولین کامپیوترهای طراحی شده از حساب ممیز ثابت استفاده می‌شد که در آن برای ذخیره‌سازی هر عدد به $n + t + 1$ بیت نیاز داریم. محدوده‌ای که توسط دستگاه $\mathbb{F}_0(\beta, n, t)$ قابل نمایش می‌باشد به صورت زیر است:

$$\mathbb{F}_0(\beta, n, t) \simeq [-\beta^n, \beta^n]$$

این محدوده چندان بزرگ نیست مگر اینکه n را خیلی بزرگ در نظر بگیریم که حافظه محدود کامپیوتر به ما اجازه انتخاب n خیلی بزرگ را نمی‌دهد. این امر به این علت اتفاق می‌افتد که توزیع اعداد در این دستگاه یکنواخت می‌باشد و حساسیت روی اعداد کوچک و بزرگ یکسان است. حال دستگاه ممیز شناور را معرفی می‌کنیم که در این دستگاه توزیع اعداد یکنواخت نیست و هرچه اعداد بزرگتر باشند فاصله بیشتری از همدیگر دارند. در دستگاه ممیز شناور، هر عدد به صورت زیر نمایش داده می‌شود:

تعریف 1.8

$$a = \pm M \times \beta^e$$

که در آن:

• M : مانیتیس، β : مبنا، e : نما

با شرایط:

$$\beta^{-1} \leq M < 1, \quad e \in \mathbb{Z}, L \leq e \leq U$$

دستگاه ممیز شناور در حالت کلی به صورت $\mathbb{F}(\beta, t, L, U)$ تعریف می‌شود که در آن:

• β : مبنا، نمایش اعداد

• t : تعداد ارقام مانیتیس

• L : کران پایین برای نما

• U : کران بالا برای نما

فرم کلی اعداد قابل نمایش

نکته 1.14 | فرم کلی

برای $a \in \mathbb{F}(\beta, t, L, U)$ داریم:

$$a = \pm (0/d_1 d_2 d_3 \dots d_t)_\beta \times \beta^e$$

با شرایط:

$$e \in \mathbb{Z} \cdot$$

$$d_i \in \{0, 1, \dots, \beta - 1\} \cdot$$

$$d_1 \neq 0 \cdot$$

$$L \leq e \leq U \cdot$$

ویژگی‌های دستگاه ممیز شناور

• تعداد اعضا:

$$2(\beta - 1)\beta^{t-1} \times (U - L + 1)$$

• نمایش عدد ۱:

اگر $1 \in [L, U]$ آنگاه:

$$1 = (0/100\dots 0)_\beta \times \beta^1 = \beta^{-1} \times \beta^1 = \beta^0$$

• کوچکترین عدد مثبت قابل نمایش:

$$x_{min} = (0/100\dots 0)_\beta \times \beta^L = \beta^{L-1}$$

• بزرگترین عدد قابل نمایش:

$$x_{max} = (0/(\beta - 1)(\beta - 1) \dots (\beta - 1))_{\beta} \times \beta^U = (1 - \beta^{-t})\beta^U = \beta^U - \beta^{U-t}$$

در این دستگاه با فرض $d_1 \neq 0$ ، عدد صفر قابل نمایش نیست و معمولاً به صورت قراردادی و جداگانه در سیستم تعریف می‌شود.

توزیع اعداد در دستگاه ممیز شناور

دو عدد متوالی $x, x_+ \in \mathbb{F}(\beta, t, L, U)$ را در نظر می‌گیریم، داریم:

$$x = (0/d_1 d_2 \dots d_t)_{\beta} \times \beta^e, \quad L \leq e \leq U$$

$$x_+ = (0/d_1 d_2 \dots (d_t + 1))_{\beta} \times \beta^e$$

$$x_+ - x = (0/00 \dots 01)_{\beta} \times \beta^e = \beta^{e-t}$$

فاصله بین اعداد به وضوح به e بستگی دارد پس فاصله بین اعداد یکنواخت نیست. برای اعداد کوچک فاصله کمتر است و برای اعداد بزرگ فاصله بیشتر است.

دستگاه ممیز شناور $\mathbb{F}(3, 3, -1, 2)$ را در نظر بگیرید.

* مثال 1.19

$$\beta = 3, \quad t = 3, \quad L = -1, \quad U = 2$$

تعداد اعضا:

$$2(2)3^2(2 + 1 + 1) = 144$$

فرم کلی اعداد قابل نمایش در این دستگاه به صورت زیر است:

$$a = \pm(0.d_1 d_2 d_3)_3 \times 3^e$$

$$d_i \in \{0, 1, 2\}, \quad d_1 \neq 0, \quad -1 \leq e \leq 2$$

بخش مثبت اعداد عبارتند از:

$$\begin{aligned} x_{min} = x_1 &= (0.100)_3 \times 3^{-1} = (3^{-1}) \times 3^{-1} = \frac{1}{9} \\ x_2 &= (0.101)_3 \times 3^{-1} = (3^{-1} + 3^{-3}) \times 3^{-1} = \frac{10}{81} = x_1 + \frac{1}{81} \\ &\vdots \\ x_{36} &= (0.222)_3 \times 3^0 \\ x_{37} &= (0.100)_3 \times 3^1 \\ &\vdots \\ x_{max} &= (0.222)_3 \times 3^2 = 8\frac{2}{3} \end{aligned}$$

1.4.4 < اپسیلون ماشین

اپسیلون ماشین در واقع فاصله بین عدد ماشینی ۱ و اولین عدد ماشینی بعد از آن می‌باشد:

$$\epsilon_M = 1_+ - 1$$

که برای دستگاه $\mathbb{F}(\beta, t, L, U)$ به صورت زیر است:

$$1 = (0.100\dots 00)_\beta \times \beta^1$$

$$1_+ = (0.100\dots 01)_\beta \times \beta^1 = \beta^1 \times (\beta^{-1} + \beta^{-t}) = 1 + \beta^{1-t}$$

$$\epsilon_M = \beta^{1-t}$$

اپسیلون دستگاه در واقع میزان دقت ماشین را نشان می‌دهد و در ادامه خواهیم دید در کران خطای گرد کردن و قطع کردن ظاهر می‌شود.

1.4.5 < نگاشت گرد کردن

واضح است که $\mathbb{F}(\beta, t, L, U) \subseteq \Omega$ به طوری که $\Omega = [-x_{max}, x_{max}]$ بنابراین ممکن است در طی محاسباتمان با عددی مثل a روبه‌رو شویم به طوری که:

$$a \in \Omega, a \notin \mathbb{F}(\beta, t, L, U)$$

اما ما نیاز داریم به جای عدد a (که قابل نمایش در دستگاه نمایش اعداد نمی‌باشد) عددی مثل $\bar{a}$ را جایگزین کنیم که $\bar{a} \in \mathbb{F}(\beta, t, L, U)$ این عدد را با $\bar{a} = fl(a)$ نمایش می‌دهیم و به آن گرد شده عدد a می‌گوییم. پس fl یک نگاشت به صورت زیر است:

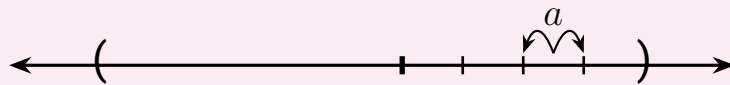
$$fl : \Omega \rightarrow \mathbb{F}(\beta, t, L, U)$$

و می‌توانیم به روش‌های مختلف آن را تعریف کنیم، به عنوان مثال:

- گرد کردن به صفر
- گرد کردن به نزدیک‌ترین عدد
- گرد کردن به $+\infty$
- گرد کردن به $-\infty$

تعریف 1.9 | روش های گرد کردن

- گرد کردن به صفر (قطع کردن): از نقطه a به سمت عدد صفر شروع به حرکت می‌کنیم. اولین عضوی از دستگاه که با آن مواجه شویم، همان گرد شده a به صفر یا قطع شده a می‌باشد.
- گرد کردن به $\pm\infty$: از نقطه a به سمت $\pm\infty$ حرکت می‌کنیم و اولین عدد ماشین که با آن برخورد می‌کنیم همان گرد شده a به $\pm\infty$ است.
- گرد کردن به نزدیک‌ترین: در این حالت فاصله عدد a و عدد ماشینی قبل از آن و بعد از آن را محاسبه می‌کنیم. هر کدام از فاصله‌ها که مقدارش کمتر بود، عدد ماشینی متناظر با آن به عنوان گرد شده عدد a انتخاب می‌شود.



گرد کردن به صفر

فرض کنیم $a \in \Omega$ و دارای نمایشی به صورت زیر است:

$$a = (0.d_1d_2\dots d_t d_{t+1}d_{t+2}\dots)_\beta \times \beta^e, \quad d_i \in 1, \dots, \beta-1, \quad L \leq e \leq u$$

و واضح است که $a \notin \mathbb{F}(\beta, t, L, U)$

گرد شده a به صفر با نماد $\bar{a} = fl_c(a)$ نمایش داده می‌شود و به صورت زیر است:

$$\bar{a} = fl_c(a) = (0.d_1\dots d_t)_\beta \times \beta^e$$

به راحتی می‌توان مشاهده کرد خطای مطلق قطع کردن به صورت زیر است:

$$E_{abs}(\bar{a}) = |a - \bar{a}| = |a - fl_c(a)| = |(0.\underbrace{0\dots 0}_t d_{t+1}d_{t+2}\dots)_\beta \times \beta^e| \leq (0.\underbrace{0\dots 0}_{t-1} 1)_\beta \times \beta^e = \beta^{e-t}$$

و خطای نسبی آن به صورت زیر است:

$$\begin{aligned} E_{rel}(\bar{a}) &= \frac{|a - \bar{a}|}{|a|} \leq \frac{\beta^{e-t}}{|a|} \\ &= \frac{\beta^{e-t}}{|(0/d_1, \dots, d_t, d_{t+1}, \dots)_\beta \times \beta^e|} \\ &\leq \frac{\beta^{e-t}}{\beta^{e-1*}} = \beta^{1-t} = \epsilon_M \end{aligned}$$

$$|(0/d_1, d_2, \dots, d_t, d_{t+1}, \dots)_\beta \times \beta^e| \geq (0/1 \underbrace{0 \dots 0}_{t-1})_\beta \times \beta^e = \beta^{e-1} \quad *$$

گرد کردن به نزدیکترین عدد

فرض کنیم a عضو Ω به صورت زیر تعریف شده باشد:

$$a = (0/d_1, d_2, \dots, d_t, d_{t+1}, \dots)_\beta \times \beta^e$$

اعداد ماشینی قبل و بعد از a به صورت زیر می‌باشد:

$$a_- = (0/d_1, d_2, \dots, d_t)_\beta \times \beta^e$$

$$a_+ = (0/d_1, d_2, \dots, (d_t + 1))_\beta \times \beta^e$$

حال $|a_- - a|$ و $|a_+ - a|$ را محاسبه می‌کنیم و هر کدام مقدار کمتری داشت، عدد متناظر با آن گرد شده a به نزدیکترین عدد می‌باشد. و واضح است که کران خطاهای مطلق و نسبی آن به صورت زیر می‌باشد:

$$\bar{a} = fl_{rn}(a)$$

$$E_{abs}(\bar{a}) = |a - fl_{rn}(a)| \leq \frac{1}{2}\beta^{e-t}$$

$$E_{rel}(\bar{a}) = \frac{|a - fl_{rn}(a)|}{|a|} \leq \frac{1}{2}\beta^{1-t} = \epsilon_M$$

تمرین 1.2

روابط بالا را اثبات کنید.

پس خطای مطلق و نسبی گرد کردن نصف خطای مطلق و نسبی قطع کردن است. کران خطاهای نسبی و مطلق مربوط به گرد کردن به $\pm\infty$ کاملاً مشابه با کران‌های مربوط به قطع کردن می‌باشد.

◀ 1.4.6 اعداد زیر نرمال

همانگونه که مشاهده کردیم که عدد صفر توسط دستگاه $\mathbb{F}(\beta, t, L, U)$ قابل نمایش نمی‌باشد. علاوه بر این یک گپ خالی اطراف صفر داریم. حال می‌خواهیم اعداد زیر نرمال را معرفی کنیم که توسط آن‌ها می‌توانیم عدد صفر را نمایش دهیم و فضای خالی اطراف صفر در بازه $[-x_{min}, x_{min}]$ را نیز پر کنیم. این اعداد را اعداد زیر نرمال می‌نامیم و با اضافه کردن آن به دستگاه $\mathbb{F}(\beta, t, L, U)$ یک دستگاه کامل‌تر شامل اعداد نرمال و زیر نرمال برای ماشین خواهیم داشت.

فرم کلی اعداد زیر نرمال به صورت زیر است:

$$a = \pm(0/0d_2d_3\dots d_t)_\beta \times \beta^L, \quad d_i \in \{0, 1, \dots, \beta - 1\}$$

بنابراین عدد صفر به صورت زیر قابل نمایش است:

$$0 = (0/\underbrace{00\dots 0}_t)_\beta \times \beta^L$$

تعداد اعداد زیر نرمال:

$$2\beta^{t-1} - 1$$

اعداد نرمال و زیر نرمال دستگاه زیر را مشخص کنید.

* مثال 1.20

$$\mathbb{F}(2, 3, -1, 2) \longrightarrow \beta = 2, t = 3, L = -1, U = 2$$

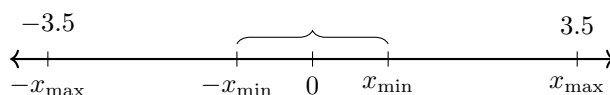
$$a = \pm(0.d_1d_2d_3)_2 \times 2^e, \quad -1 \leq e \leq 2$$

$$x_{\min} = x_1 = +(0.100)_2 \times 2^{-1} = 2^{-2} = \frac{1}{4}$$

$$x_2 = +(0.101)_2 \times 2^{-1} = (2^{-1} + 2^{-3}) \times 2^{-1} = \frac{1}{4} + \frac{1}{16}$$

$\vdots$

$$x_{\max} = +(0.111)_2 \times 2^2 = 2 + 1 + \frac{1}{2}$$



$$0 = (0.000) \times 2^{-1} = 0$$

$$\bar{x}_1 = (0.001) \times 2^{-1} = \frac{1}{16}$$

$$\bar{x}_2 = (0.010) \times 2^{-1} = \frac{1}{8}$$

$$\bar{x}_3 = (0.011) \times 2^{-1} = \frac{1}{8} + \frac{1}{16} = \frac{3}{16}$$

$$\text{در کل: } 2(2^2) - 1 = 7$$

1.4.7 < استاندارد IEEE

اگرچه در برخی از ماشین حساب‌های قدیمی از دستگاه‌های ده‌دهی استفاده می‌شد، امروزه برای تمام کامپیوترها و ماشین‌های محاسباتی از دستگاه دودویی استفاده می‌شود. در اوایل عصر توسعه کامپیوترها از دستگاه‌هایی با ویژگی‌های متفاوت (u, L, t) استفاده می‌شد. این امر منجر به انتقال ناپذیری نرم‌افزارهای مختلف روی ماشین‌های

متفاوت می‌شد. بنابراین وجود یک دستگاه ماشینی استاندارد کاملاً احساس می‌شد. در نتیجه در سال ۱۹۸۵ دانشمندان علوم کامپیوتر به سرپرستی ویلیام کامان از دانشگاه کالیفرنیا یک استاندارد برای نمایش اعداد در انجمن مهندسان برق و کامپیوتر مشخص کردند که به استاندارد IEEE754 مشهور می‌باشد و از آن پس تمام کامپیوترها بر اساس این استاندارد طراحی و ساخته شدند.

استاندارد IEEE754 دارای دو دقت معمولی و دابل می‌باشد که منجر به کامپیوترهای ۳۲ و ۶۴ بیتی می‌شود و دستگاه‌های متناظر به صورت زیر می‌باشند:

$$\mathbb{F}(2, 24, -126, 127) \rightarrow 32\text{-bit}$$

$$\underbrace{1}_{+or-} + \underbrace{23}_M + \underbrace{8}_e = 32$$

$$\mathbb{F}(2, 53, -1022, 1023) \rightarrow 64\text{-bit}$$

$$\underbrace{1}_{+or-} + \underbrace{52}_M + \underbrace{11}_e = 64$$

توجه داشته باشیم برای ذخیره نما از ذخیره‌سازی اعداد مثبت و منفی استفاده نمی‌کنیم بلکه از روش مکمل دو بهره می‌بریم. به این معنا که یک توان اریبی محاسبه می‌کنیم و آن را با نمای عدد جمع می‌کنیم و مجموع حاصل را به عنوان نما ذخیره می‌کنیم.

$$\text{computer } 32\text{-bit} \rightarrow = 2^{8-1} - 1 = 127$$

$$\text{as stored is (e) Exponent } e + 127$$

$$\text{computer } 64\text{-bit} \rightarrow = 2^{11-1} - 1 = 1023$$

$$\text{as stored is (e) Exponent } e + 1023$$

عدد 89.75 در استاندارد IEEE با دقت معمولی دارای چه نمایشی می‌باشد؟

*** مثال 1.21**

$$89.75 = (?)_2$$

بخش صحیح

$$89 \div 2 = 44 \quad R = 1$$

$$44 \div 2 = 22 \quad R = 0$$

$$22 \div 2 = 11 \quad R = 0$$

$$11 \div 2 = 5 \quad R = 1$$

$$5 \div 2 = 2 \quad R = 1$$

$$2 \div 2 = 1 \quad R = 0$$

$$1 \div 2 = 0 \quad R = 1$$

$$\Rightarrow 89 = (1011001)_2$$

بخش اعشاری

$$0.75 \times 2 = 1.5$$

$$0.5 \times 2 = 1.0$$

$$0.75 = (0.11)_2$$

$$89.75 = (1011001.11)_2$$

$$= (\underbrace{1}_{\text{ذخیره نمی شود}}.01100111)_2 \times 2^6$$

$$\rightarrow M = (0.01100111)_2, \quad e = 6$$

$$e + (2^{8-1} - 1) = 6 + 127 = 133$$

$$133 \div 2 = 66 \quad R = 1$$

$$66 \div 2 = 33 \quad R = 0$$

$$33 \div 2 = 16 \quad R = 1$$

$$16 \div 2 = 8 \quad R = 0$$

$$8 \div 2 = 4 \quad R = 0$$

$$4 \div 2 = 2 \quad R = 0$$

$$2 \div 2 = 1 \quad R = 0$$

$$1 \div 2 = 0 \quad R = 1$$

$$133 = (10000101)_2$$

نمایش a در IEEE با دقت معمولی (32-bit)

$$\underbrace{0}_{\text{علامت}} \quad \underbrace{10000101}_{\text{نما}} \quad \underbrace{011001110 \dots 0}_{\text{مانتیس تا ۱۵}}$$

۳۲ بیت زیر نمایش چه عددی در مبنای ۱۰ می باشد؟

* مثال 1.22

$$a = \underbrace{1}_{\text{علامت منفی}} \quad \underbrace{01011001}_{\text{نما}} \quad \underbrace{0111010 \dots 0}_{\text{مانتیس}}$$

نما:

$$e + 127 = (01011001)_2 = 2^6 + 2^4 + 2^3 + 2^0 = 89$$

$$e = -38$$

مانتیس:

$$M = \left(\underset{\substack{\uparrow \\ \text{باید ۱ اضافه بشه}}}{1}, 0111010 \dots 0 \right)_2 = 2^0 + 2^{-2} + 2^{-3} + 2^{-4} + 2^{-6}$$

$$= \frac{93}{64} = 1.4531 \dots$$

$$a = -1.4531 \dots \times 2^{-38} \approx -0.5286 \dots \times 10^{-11}$$

1.5 وضعیت مسئله و پایداری روش

در حل یک مسئله ریاضی توسط یک روش عددی لازم است به دو موضوع توجه کنیم:

۱. خوش وضعی مسئله

۲. پایداری روش عددی

در ادامه با این دو موضوع آشنا می‌شویم.

در اکثر مسائلی که به صورت عددی حل می‌کنیم با یک دستگاه خطی مواجه هستیم. یعنی لازم است یک دستگاه $Ax = b$ را حل کنیم که A یک ماتریس $m \times n$ و b یک بردار m تایی می‌باشد. می‌خواهیم $x \in \mathbb{R}^n$ را چنان محاسبه کنیم که در تساوی $Ax = b$ صدق کند.

حال فرض می‌کنیم ماتریس A که در دستگاه $Ax = b$ ظاهر می‌شود یک ماتریس هیلبرت باشد.

تعریف 1.10 | ماتریس هیلبرت

$$H_n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \dots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \dots & \frac{1}{2n-1} \end{pmatrix}_{n \times n}$$

این ماتریس در حل مسائل مختلف ظاهر می‌شود، از جمله در یکی از روش‌های حل مسئله کمترین مربعات با این ماتریس برخورد می‌کنیم.

حال دستگاه $Ax = b$ را چنان تعریف می‌کنیم که:

$$A = H_{11}, \quad \bar{x} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$b = A\bar{x} = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{11} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{12} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{11} & \frac{1}{12} & \cdots & \frac{1}{21} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 + \frac{1}{2} + \cdots + \frac{1}{11} \\ \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{12} \\ \vdots \\ \frac{1}{11} + \frac{1}{12} + \cdots + \frac{1}{21} \end{pmatrix}$$

ماتریس A و بردار b را به کامپیوتر می‌دهیم و به کمک یک نرم‌افزار (متلب با دستور $\text{linsolve}(A, b)$) جواب دستگاه معادلات را به دست می‌آوریم. جواب دستگاه (با دستور $\text{linsolve}(A, b)$ در متلب) به صورت زیر به دست می‌آید:

The screenshot shows the MATLAB environment. The 'Files' pane on the left displays 'Hilbert.m'. The 'Workspace' pane shows variables: A (11x11 double), X (11x1 double), X_bar (11x1 double), b (11x1 double), and n (11, 1x1 double). The 'Command Window' shows the execution of the script 'Hilbert', displaying the solution X as a column vector of 11 values.

```

1 format long;
2
3 n = 11;
4 A = hilb(n);
5
6 X_bar = ones(n, 1);
7
8 b = A * X_bar;
9
10 X = linsolve(A, b);
11
12 disp('Solution X:');
13 disp(X);

```

Command Window

```

>> Hilbert
Solution X:
0.999999994843199
1.000000538857933
0.999986041359532
1.000155875384751
0.999072324686194
1.003258717986100
0.992910160557468
1.009658807364932
0.991981907673492
1.003707653813361
0.999267974328987

```

برای حل این مسئله از یک کامپیوتر معمولی مثلاً ۶۴ بیتی استفاده کرده‌ایم و طبق اپسیلون ماشین انتظار داریم خطا به اندازه 10^{-16} باشد ولی در اینجا در بعضی درایه‌ها در رقم سوم جواب خطا داریم. این اتفاق به دلیل روش محاسباتی نمی‌باشد بلکه مشکل از خود مسئله می‌باشد و در واقع مسئله یک مسئله بدوضع است. (در ادامه به

صورت دقیق خوش‌وضع و بدوضع مسئله تعریف می‌شود).
اگر خطای نسبی جواب را محاسبه کنیم، داریم:

$$\frac{\|x - \bar{x}\|_{\infty}}{\|\bar{x}\|_{\infty}} = 0.00130933 \dots$$

پس می‌توانیم نتیجه بگیریم حل دستگاه $Ax = b$ زمانی که A یک ماتریس هیلبرت باشد یک مسئله بد وضع می‌باشد پس باید از ایجاد دستگاهی خطی با ماتریس هیلبرت تا جای ممکن جلوگیری کنیم.

*** مثال 1.23** می‌خواهیم $I_n = \int_0^1 \frac{t^n}{t+10} dt$ را برای $n \geq 1$ محاسبه کنیم.

$$I_0 = \int_0^1 \frac{dt}{t+10} = \ln 1.1$$

$$I_1 = \int_0^1 \frac{t}{t+10} dt = \int_0^1 1 - \frac{10}{t+10} dt = 1 - 10I_0$$

و رابطه بازگشتی زیر را داریم:

$$\begin{cases} I_k = -10I_{k-1} + \frac{1}{k} & k = 1, 2, \dots, n \\ I_0 = \ln 1.1 \end{cases}$$

اگر قرار دهیم $n = 16$ و مقدار I_{16} را به روش بالا حساب کنیم داریم:

$$I_{16} = -0.069122123 \dots$$

که قطعاً مقداری نادرست است. در اینجا خود مسئله مشکل ندارد اما روشی که برای حل آن در نظر گرفته‌ایم مناسب نمی‌باشد و نیاز به تغییر دارد.

در دو مثال فوق، دو مفهوم، یکی بدوضع یک مسئله تحت مطالعه و دیگری ناپایداری الگوریتم عددی را مشاهده کردیم. دقت کنیم قبل از سپردن یک مسئله به یک روش عددی، ابتدا باید از خوش‌وضع مسئله مطمئن باشیم. پس از آن لازم است روشی را برای حل آن به کار گیریم که مطمئن باشیم به جواب‌های معتبر منجر شود.

1.5.1 < وضعیت یک مسئله ریاضی

یک مسئله ریاضی معمولاً دارای ورودی و خروجی می‌باشد. ورودی را با x و خروجی را با y نمایش می‌دهیم.

$$x \longrightarrow \boxed{f} \longrightarrow y$$

و توسط یک تابع، ورودی و خروجی با هم در ارتباط هستند. اگر رابطه ورودی و خروجی به صورت ضمنی باشد، آنگاه تابع F وجود دارد به طوری که

$$F(x, y) = 0$$

و ممکن است y به صورت صریح از x بدست آید. در این صورت تابع f وجود دارد به طوری که

$$y = f(x)$$

می‌خواهیم حساسیت مسئله ریاضی $y = f(x)$, $F(x, y) = 0$ را نسبت به ایجاد اختلال در ورودی x مشخص کنیم. اگر اختلال کوچک در x منجر به ایجاد اختلالی از همان مرتبه در y شود، گوییم مسئله خوش‌وضع است و اگر اختلال کوچک در x منجر به اختلال بزرگی در y شود، گوییم مسئله بد وضع است.

تعریف 1.11 | خوش وضعی

فرض کنیم:

$$f : D \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$$

$$x \mapsto y = f(x)$$

تعریف: فرض کنیم $x \in D$ باشد. به x اختلالی به اندازه δx وارد می‌کنیم به طوری که $x + \delta x \in D$ و جواب مسئله یعنی $y + \delta y$ را بدست می‌آوریم، یعنی:

$$F(x + \delta x, y + \delta y) = 0$$

or

$$y + \delta y = f(x + \delta x)$$

گوییم مسئله خوش‌وضع است هرگاه ثابت کوچک c وجود داشته باشد به طوری که:

$$\|\delta y\| \leq c \|\delta x\| \quad \text{or} \quad \frac{\|\delta y\|}{\|y\|} \leq c \frac{\|\delta x\|}{\|x\|}$$

خطاهای بالا یکی به صورت مطلق و یکی به صورت نسبی بیان شده است. در تعاریف بالا می‌توانیم نرم‌های مختلفی استفاده کنیم و چون x و y ممکن است متعلق به فضاهای متفاوتی باشند، ممکن است لازم باشد با نرم‌های متفاوتی کار کنیم.

نکته 1.15

هر مسئله‌ای که خوش‌وضع نباشد، یک مسئله بدوضع است. مسئله بدوضع را با هر روش عددی حل کنیم، نمی‌توان به جواب حاصل اعتماد کرد.

1.5.2 ضریب وضعیت

توسط ضریب وضعیت می‌توانیم خوش‌وضع یا بدوضع یک مسئله را مشخص کنیم. فرض کنیم $x \neq 0$, $y \neq 0$, $m = n = 1$ توسط تابع f به هم مرتبط باشند، یعنی:

$$f: \mathbb{R} \mapsto \mathbb{R}$$

$$x \mapsto y = f(x), \quad x \neq 0, y \neq 0$$

به ورودی x اختلالی به اندازه δx وارد می‌کنیم و خروجی $y + \delta y$ حاصل می‌شود و داریم:

$$y + \delta y = f(x + \delta x)$$

نکته 1.16

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!} (x-a)^i$$

$$\begin{aligned} y + \delta y = f(x + \delta x) &= \sum_{i=0}^{\infty} \frac{f^{(i)}(x)}{i!} (x + \delta x - x)^i \\ &= \sum_{i=0}^{\infty} \frac{f^{(i)}(x)}{i!} (\delta x)^i \\ &= f(x) + f'(x)\delta x + \frac{f''(x)}{2!}(\delta x)^2 + \frac{f'''(x)}{3!}(\delta x)^3 + \dots \\ &= f(x) + f'(x)\delta x + O(|\delta x|^2) \end{aligned}$$

$$y + \delta y = f(x) + f'(x)\delta x + O(|\delta x|^2)$$

$$\Rightarrow y + \delta y \simeq \underbrace{f(x)}_y + f'(x)\delta x$$

$$\delta y \simeq f'(x)\delta x$$

اگر از طرفین رابطه قدر مطلق بگیریم و به ازای $|f'(x)|$ کران بالای آن را قرار دهیم، داریم:

$$|\delta y| \leq c_1 |\delta x|, \quad |f'(x)| \leq c_1$$

بنابراین اگر کران بالای مشتق اول تابع f عدد کوچکی باشد، مسئله $y = f(x)$ یک مسئله خوش‌وضعی (در حالت خطای مطلق) می‌باشد.

و در حالت خطای نسبی داریم:

$$\begin{aligned} \frac{\delta y}{y} &\simeq \frac{f'(x)}{f(x)} \frac{x}{x} \frac{\delta x}{x} \\ &= \frac{f'(x)}{f(x)} \frac{x}{x} \frac{\delta x}{x} \end{aligned}$$

حال از طرفین رابطه قدر مطلق می‌گیریم و به ازای:

$$\frac{|f'(x)||x|}{|f(x)|} \leq c_2$$

کران بالای آن را قرار می‌دهیم. یعنی:

$$\frac{|\delta y|}{|y|} \leq c_2 \frac{|\delta x|}{|x|}$$

اگر c_2 کوچک باشد، مسئله خوش‌وضع است و اگر عدد بزرگی باشد، مسئله بدوضع است. بنابراین متناظر با مسئله $y = f(x)$ ضریب وضعیت مسئله در حالت نسبی به صورت زیر تعریف می‌شود:

تعریف 1.12 | ضریب وضعیت در حالت نسبی

$$(cond f)(x) = \left| \frac{f'(x) x}{f(x)} \right|$$

در حالتی که $x \in \mathbb{R}^m$ و $y \in \mathbb{R}^n$ باشد و داشته باشیم:

$$f : D \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$$

$$x \mapsto y = f(x)$$

آنگاه ضریب وضعیت مسئله به صورت زیر تعریف می‌شود:

$$(cond f)(x) = \frac{\|f'(x)\| \|x\|}{\|f(x)\|}$$

ضریب وضعیت دستگاه معادلات خطی $Ax = b$ را مشخص کنید زمانی که A یک ماتریس

* مثال 1.24

$n \times n$ معکوس پذیر و $x, b \in \mathbb{R}^n$ است. فرض کنیم اختلالی به اندازه δb به b وارد می‌شود و ورودی A ثابت و بدون اختلال می‌باشد.

❓ حل:

$$A, b \mapsto x$$

$$Ax = b \implies x = A^{-1}b = f(b)$$

$$x = f(b), \quad f'(b) = A^{-1}$$

$$* \quad \|Ax\| \leq \|A\| \|x\|$$

$$\begin{aligned} (cond f)(b) &= \frac{\|f'(b)\| \|b\|}{\|f(b)\|} = \frac{\|A^{-1}\| \|Ax\|}{\|x\|} \\ &\leq \frac{\|A^{-1}\| \|A\| \|x\|}{\|x\|} = \|A^{-1}\| \|A\| = cond A \end{aligned}$$

مشاهده می‌کنیم ضریب وضعیت مسئله $Ax = b$ کاملاً به ضریب وضعیت ماتریس A بستگی دارد. به عنوان مثال برای ماتریس هیلبرت داریم:

n	$\text{cond}_2 H_n$
10	1.6×10^3
20	2.45×10^8
30	7.65×10^{58}

اکنون دلیل خطاهای ایجاد شده در مسئله $H_{11}x = b$ را متوجه می‌شویم.

*** مثال 1.25** یکی دیگر از ماتریس‌های بدوضع، ماتریس واندروند می‌باشد که برای مقادیر متمایز $t_1, \dots, t_n$ به صورت زیر تعریف می‌شود:

$$V_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \\ \vdots & \vdots & \ddots & \vdots \\ t_1^{n-1} & t_2^{n-1} & \cdots & t_n^{n-1} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

این ماتریس در بحث درونیابی به کمک چندجمله‌ای ظاهر می‌شود.

$\text{cond}_\infty V_n$	n
1.36×10^4	10
1.05×10^9	20
6.93×10^{18}	40

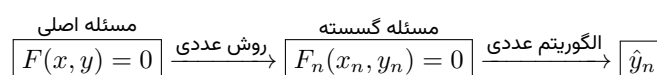
آنچه گفتیم، در ارتباط با وضعیت یک مسئله ریاضی بود. حال به مفهوم پایداری الگوریتم عددی می‌پردازیم.

1.5.3 سازگاری، ناپایداری و همگرایی

حل عددی یک مسئله ریاضی شامل دو فاز است.

۱. در فاز اول، مسئله اصلی $F(x, y) = 0$ توسط یک روش عددی به یک مسئله (عمدتاً) گسسته تبدیل می‌شود که ما به آن **مسئله گسسته** یا **میانی** می‌گوییم و آن را با $0 = F_n(x_n, y_n)$ نشان می‌دهیم که در آن x_n, y_n به ترتیب ورودی و جواب مسئله گسسته هستند. در بسیاری از مواقع $x_n = x$ است. یعنی ورودی‌های مسئله گسسته با ورودی‌های مسئله اصلی یکسان هستند.

۲. در فاز دوم، الگوریتم در کامپیوتر اجرا می‌شود و جواب هر مسئله گسسته به دست می‌آید. دقت کنیم ممکن است جواب محاسبه شده نهایی به خاطر خطاهای گرد کردن و خطای الگوریتم با y_n متفاوت باشد که در اینجا، آن را با $\hat{y}_n$ نشان می‌دهیم.



دقت کنیم مسئله گسسته به صورت $F_n(x_n, y_n) = 0, \quad n \geq 0$ می‌باشد و انتظار داریم $y_n \rightarrow y$ باشد. در این صورت لازم است:

$$x_n \rightarrow x, \quad F_n \rightarrow F$$

و اگر مسئله به صورت صریح مشخص شده باشد یعنی داریم $y_n = f(x_n)$ گاهی برای گسسته‌سازی ما به یک پارامتر $h > 0$ وابسته هستیم یعنی داریم $F_h(x_h, y_h) = 0, \quad h > 0$ در این صورت انتظار داریم وقتی $h \rightarrow 0$ داشته باشیم $y_h \rightarrow y$

سازگاری

اگر یک داده x از مسئله $F(x, y) = 0$ برای F_n نیز به عنوان ورودی مجاز باشد (در دامنه آن باشد) گوییم مسئله $F_n(x_n, y_n) = 0$ سازگار است اگر

$$F_n(x, y) \rightarrow 0 \text{ when } n \rightarrow \infty$$

و گوییم روش قویاً سازگار است اگر نه فقط وقتی $n \rightarrow \infty$ بلکه برای هر n داشته باشیم

$$F_n(x, y) = 0$$

روش تکراری نیوتن برای تقریب ریشه غیر تکراری α از تابع $g: \mathbb{R} \mapsto \mathbb{R}$ به صورت زیر می‌باشد

* مثال 1.26

$$y_{n+1} = y_n - \frac{g(y_n)}{g'(y_n)} \quad n \geq 0, \text{ with given } y_0$$

نشان دهید این روش قویاً سازگار است.

$$F_n(g, y_{n+1}, y_n) = y_{n+1} - y_n - \frac{g(y_n)}{g'(y_n)}$$

با توجه به اینکه برای هر n داریم $F_n(g, \alpha, \alpha) = 0$ ، روش قویاً سازگار می‌باشد.

پایداری روش عددی

برای اینکه یک روش پایدار باشد لازم است برای هر n ثابت یک جواب یکتای y_n متناظر با داده x_n وجود داشته باشد و به علاوه اختلال جزئی در ورودی x_n منجر به اختلال جزئی از همان مرتبه در جواب y_n باشد یعنی وابستگی پیوسته به داده x_n داشته باشد.

تعریف 1.13

گیریم D_n مجموعه داده‌های مجاز برای رابطه $F_n(x_n, y_n) = 0$ باشد و فرض کنیم x_n عضوی دلخواه از D_n باشد. گیریم δx_n یک اختلال ورودی برای مسئله باشد به گونه‌ای که $x_n + \delta x_n \in D_n$ و δy_n اختلال متناظر آن در جواب باشد به گونه‌ای که:

$$F_n(x_n + \delta x_n, y_n + \delta y_n) = 0 \quad \text{or} \quad y_n + \delta y_n = f(x_n + \delta x_n)$$

روش عددی پایدار است اگر این مسئله به ازای هر n دارای جواب یکتا باشد، ثابت نه چندان بزرگ c_n موجود باشد به طوری که

$$|\delta y_n| \leq c_n |\delta x_n| \quad \text{or} \quad \frac{|\delta y_n|}{|y_n|} \leq c_n \frac{|\delta x_n|}{|x_n|}$$

همگرایی

تعریف 1.14

گوییم جواب y_n از مسئله گسسته $F_n(x_n, y_n) = 0$ در $\lim ||.||$ همگرا به جواب مسئله خوش وضع $0 = F(x, y)$ است اگر و فقط اگر

$$\frac{|y - y_n|}{|y|} \leq g(n) \quad n \rightarrow \infty$$

که در آن $g(n)$ دنباله‌ای مثبت و همگرا به صفر است.

1.5.4 هزینه محاسباتی

هزینه محاسباتی یا پیچیدگی یک روش، تعداد کل اعمال ریاضی (جمع، ضرب، تقسیم و تفریق) و همچنین ارزیابی‌های توابع در سرتاسر الگوریتم است. اگر برای حل یک مسئله دو روش مناسب (سازگار، پایدار و همگرا) موجود باشد، روشی که هزینه محاسباتی کمتری داشته باشد ارجحیت خواهد داشت.

شاید تصور شود با وجود کامپیوترهای سریع که اعمال ریاضی را در کسر ناچیزی از ثانیه انجام می‌دهند، پرداختن به این موضوع بی‌اهمیت باشد. برای نشان دادن اهمیت این موضوع، حل یک دستگاه خطی از بعد 11×11 را در نظر بگیرید. اگرچه روش کرامر برای حل چنین دستگاهی پایدار نیست، با این حال اگر از این روش استفاده شود، نیاز به محاسبه 12 دترمینان است. اگر دترمینان‌ها را با روش بسط حول یک سطر یا ستون به صورت بازگشتی محاسبه کرد، حل این دستگاه به بیش از 13 میلیارد عمل محاسباتی (جمع، تفریق، ضرب و تقسیم) نیاز دارد که مدت زمان قابل توجهی برای پردازش این تعداد محاسبات مورد نیاز است.

اما اگر همین دستگاه را با روش حذفی گاوس حل کنیم، به 900 عمل محاسباتی احتیاج داریم که کامپیوترهای امروزی در کسری از ثانیه آن را انجام می‌دهند. اگر ابعاد مسئله بزرگتر باشد، اهمیت استفاده از روش‌هایی با پیچیدگی محاسباتی کم آشکارتر می‌شود.

* مثال 1.27

هزینه محاسباتی $p(x_0)$ را بدست آورید.

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

$$x_0 \in \mathbb{R}$$

$$P(x_0) = ?$$

$$A = [a_1, a_2, \dots, a_n]_{1 \times n}$$

$$S = a_0$$

for $i = 1$ to n

$$S = S + A(1, i) * x_0^i$$

for end

حل:💡

ضرب	جمع	i
1	1	i=1
2	1	i=2
⋮	⋮	⋮
n	1	i=n

$$\sum_{i=1}^n 1 + \sum_{i=1}^n i$$

$$= n + \frac{n(n+1)}{2} = \frac{n^2}{2} + \frac{3}{2}n = O(n^2)$$

* مثال 1.28

هزینه محاسباتی $p(x_0)$ را بدست آورید.

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

$$x_0 \in \mathbb{R}$$

$$P(x_0) = ?$$

$$A = [a_1, a_2, \dots, a_n]_{1 \times n}$$

$$S = a_0$$

$$P = 1$$

for $i = 1$ to n

$$S = S + A(1, i) * P * x_0$$

$$P = P * x_0$$

for end

💡 حل:

ضرب	جمع	i
3	1	$\forall i$

$$\begin{aligned} \sum_{i=1}^n 1 + \sum_{i=1}^n 3 \\ = n + 3n = 4n = O(n) \end{aligned}$$

هزینه محاسباتی $p(x_0)$ را بدست آورید.

* مثال 1.29

$$p(x) = (\dots (a_n x_0 + a_{n-1})x_0 + a_{n-2})x_0 + a_{n-3} + \dots + a_0$$

$$p_1(x) = a_n x_0 + a_{n-1}$$

$$p_2(x) = a_n x_0^2 + a_{n-1} x_0 + a_{n-2}$$

$$p_3(x) = a_n x_0^3 + a_{n-1} x_0^2 + a_{n-2} x_0 + a_{n-3}$$

⋮

$$p(x) = p_n(x) = a_n x_0^n + a_{n-1} x_0^{n-1} + \dots + a_1 x_0 + a_0$$

$$A = [a_1, a_2, \dots, a_n]$$

$$c = a_0$$

$$p = A(1, n)$$

for $i = n - 1 : -1 : 1$

$$p = p * x_0 + A(1, i)$$

end

$$p = p + c$$

💡 حل:

ضرب	جمع	i
1	1	$\forall i$

و یک جمع دیگر در آخر

$$\sum_{i=1}^n 1 + \sum_{i=1}^n 1 + 1$$

$$= n + n + 1 = 2n + 1 = O(n)$$

روش‌های حل دستگاه‌های خطی

2

یک دستگاه خطی در حالت کلی به فرم زیر می‌باشد:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

$$a_{ij} \in \mathbb{R}, b_j \in \mathbb{R}, x_1, x_2, \dots, x_n = ?$$

$$1 \leq i \leq m, \quad 1 \leq j \leq n$$

این دستگاه را می‌توانیم به فرم ماتریس به صورت زیر نیز نمایش دهیم:

$$Ax = b$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m \times n}$$

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m \times 1}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1}$$

نکته 2.1

حالت‌های مختلف برای m, n در نظر می‌گیریم، داریم:

حالت ۱: $m < n$ در حالت کلی ممکن است بی‌نهایت جواب داشته باشیم.

حالت ۲: $m = n$ معمولاً یک جواب داریم و جواب یکتاست.

حالت ۳: $m > n$ در حالت کلی ممکن است جواب نداشته باشیم.

برای حل دستگاه مربعی فوق می‌توانیم از روش‌های مستقیم و یا روش‌های تکراری استفاده کنیم.

2.1 روش‌های مستقیم

در این روش‌ها جواب مسئله به صورت دقیق محاسبه می‌شود که در ادامه به برخی از این روش‌ها اشاره می‌کنیم:

2.1.1 روش کرامر

روشی ناپایدار است و هزینه محاسباتی زیادی دارد. در این درس به این روش نمی‌پردازیم و تنها به آن اشاره می‌کنیم. در این روش ماتریس‌های A_i را به صورت زیر تعریف می‌کنیم. ماتریس A_i همان ماتریس A است که ستون i ام آن با بردار b جایگزین شده است برای $i = 1, 2, \dots, n$ و جواب دستگاه به صورت زیر محاسبه می‌شود:

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad i = 1, \dots, n, \quad \det(A) \neq 0$$

2.1.2 روش حذفی گاوس

جواب $\xrightarrow[\text{پسرو}]{\text{جای‌گذاری}} \left[\begin{array}{c|c} U & \bar{b} \end{array} \right] \xrightarrow[\text{مقدماتی}]{\text{اعمال سطری}} [A, b] \xrightarrow[\text{ماتریس افزوده}]{} \text{ماتریس افزوده}$

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_n \end{pmatrix}$$

2.1.3 روش گاوس-ژوردان (جردن)

$$[A, b] \xrightarrow[\text{ماتریس افزوده}]{\text{مقدماتی}} [I, \hat{b}] \longrightarrow x = \hat{b}$$

2.1.4 < روش‌های تجزیه ماتریسی

تجزیه LU

$$A = \underbrace{L}_{\text{بالامثلثی پایین‌مثلثی}} \underbrace{U}_{\text{بالامثلثی}}$$

$$Ax = b \longrightarrow L(\underbrace{Ux}_y) = b \begin{cases} Ux = y & \text{②} \longrightarrow \text{جای‌گذاری پس‌رو} \\ Ly = b & \text{①} \longrightarrow \text{جای‌گذاری پیش‌رو} \end{cases}$$

تجزیه QR

$$A = \underbrace{Q}_{\text{بالامثلثی متعامد}} \underbrace{R}_{\text{بالامثلثی}}$$

$$Ax = b \longrightarrow QRx = b \longrightarrow Rx = Q^T b \longrightarrow \text{جای‌گذاری پس‌رو}$$

2.2 ≡ روش‌های تکراری

در این روش‌ها معمولاً جواب به صورت تقریبی بدست می‌آید. با یک جواب اولیه x^0 شروع می‌کنیم و توسط یک الگوریتم تکراری دنباله $\{x^k\}$ را بدست می‌آوریم که این دنباله تحت شرایطی به جواب مسئله همگراست (در فصل بعد این روش‌ها را مطالعه خواهیم کرد).

$$x^0 \Rightarrow x^1 \Rightarrow x^2 \Rightarrow \dots$$

$$\{x^k\}, x^k \rightarrow x^* \text{ (جواب)}$$

2.3 ≡ روش حذفی گاوس

همانگونه که در بخش قبل اشاره کردیم در مرحله آخر روش حذفی گاوس یک جای‌گذاری پس‌رو داریم. ابتدا می‌خواهیم هزینه محاسباتی جای‌گذاری پس‌رو را مشخص کنیم. در جای‌گذاری پس‌رو ما یک دستگاه مثلثی به فرم زیر را داریم:

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & \ddots & & \ddots & & \vdots & & \vdots \\ & & & & a_{n-1\ n-1}x_{n-1} & + & a_{n-1\ n}x_n & = & b_{n-1} \\ & & & & & & a_{nn}x_n & = & b_n \end{array}$$

$$x_n = \frac{b_n}{a_{nn}}, \quad a_{nn} \neq 0$$

$$x_{n-1} = \frac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,n}x_n)$$

و به همین ترتیب پیش می‌رویم و x_1 را در انتها حساب می‌کنیم. بنابراین در حالت کلی فرمول زیر را داریم:

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^n a_{ij}x_j \right), \quad i = n, n-1, \dots, 1$$

● هزینه محاسباتی

i	ضرب	جمع
$i = n$	1	0
$i = n-1$	2	1
$\vdots$	$\vdots$	$\vdots$
$i = 1$	n	$n-1$

• جمع:

$$\sum_{i=1}^n (n-i) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$$

• ضرب:

$$\sum_{i=1}^n (n-i+1) = n^2 - \frac{n(n+1)}{2} + n = \frac{n(n+1)}{2}$$

روش جای‌گذاری پیشرو نیز به همین ترتیب می‌باشد.

مشاهده کردیم در روش حذفی گاوس ابتدا لازم است ماتریس افزوده دستگاه را مشخص کنیم و بعد با اعمال سطری مقدماتی، ماتریس سمت چپ در ماتریس افزوده را به یک ماتریس بالامثلثی تبدیل کنیم.

≡ ویژگی‌ها 2.1 | اعمال سطری مقدماتی

۱. جابه‌جایی دو سطر i و j

$$E_{i,j}$$

۲. ضرب $\alpha \neq 0$ در سطر i ام

$$E_i(\alpha)$$

۳. افزودن α برابر سطر i ام به سطر j ام

$$E_{i,j}(\alpha)$$

تعریف 2.1 | دو دستگاه معادل

گوییم دو دستگاه $Ax = b$ و $\bar{A}x = \bar{b}$ معادل هستند هرگاه جواب یکسان داشته باشند.

قضیه 2.1

اگر دستگاه $\bar{A}x = \bar{b}$ از روی دستگاه $Ax = b$ با انجام اعمال سطری مقدماتی بدست آمده باشد، دو دستگاه معادل می‌باشند.

فرض کنیم دستگاه $Ax = b$ را داریم

$$[A, b] = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array} \right] \xrightarrow[\text{مرحله اول}]{\text{ستون اول}} \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right]$$

$$\begin{cases} m_{i1}^{(1)} = +\frac{a_{i1}^{(1)}}{a_{11}^{(1)}} & i = 2, 3, \dots, n \\ a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1}^{(1)} a_{1j}^{(1)} \\ b_i^{(2)} = b_i^{(1)} - m_{i1}^{(1)} b_1^{(1)} \end{cases}$$

$$[A, b] = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right] \xrightarrow[\text{مرحله دوم}]{\text{ستون دوم}} \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array} \right]$$

$$\begin{cases} m_{i2}^{(2)} = +\frac{a_{i2}^{(2)}}{a_{22}^{(2)}} & i = 3, 4, \dots, n \\ a_{ij}^{(3)} = a_{ij}^{(2)} - m_{i2}^{(2)} a_{2j}^{(2)} \\ b_i^{(3)} = b_i^{(2)} - m_{i2}^{(2)} b_2^{(2)} \end{cases}$$

به همین ترتیب تا مرحله $n - 1$ پیش می‌رویم و در مرحله آخر، ماتریس زیر را خواهیم داشت:

$$\left[\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & b_{n-1}^{(n-1)} \\ 0 & 0 & \dots & 0 & a_{nn}^{(n)} & b_n^{(n)} \end{array} \right]$$

در واقع خلاصه الگوریتم فوق را به صورت زیر می‌توان نوشت:
در مرحله k داریم:

$$m_{ik}^{(k)} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k+1, \dots, n, \quad a_{kk}^{(k)} \neq 0$$

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}^{(k)} \times a_{kj}^{(k)}, \quad i, j = k+1, \dots, n$$

$$b_i^{(k+1)} = b_i^{(k)} - m_{ik}^{(k)} \times b_k^{(k)}, \quad i = k+1, \dots, n$$

جواب دستگاه زیر را به روش حذفی گaus بدست آورید.

* مثال 2.1

$$x_1 + 3x_2 + 2x_3 = 7$$

$$5x_1 + x_2 + 3x_3 = 10$$

$$x_1 - x_2 + 2x_3 = 5$$

💡 حل:

$$\left[\begin{array}{ccc|c} 1 & 3 & 2 & 7 \\ 5 & 1 & 3 & 10 \\ 1 & -1 & 2 & 5 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 3 & 2 & 7 \\ 0 & -14 & -7 & -25 \\ 0 & -4 & 0 & -2 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 3 & 2 & 7 \\ 0 & -14 & -7 & -25 \\ 0 & 0 & 2 & \frac{36}{7} \end{array} \right]$$

جای‌گذاری پسرو:

$$2x_3 = \frac{36}{7} \rightarrow x_3 = \frac{18}{7}, \quad -14x_2 - 7 \times \frac{18}{7} = -25 \rightarrow x_2 = \frac{1}{2}$$

$$x_1 + 3 \times \frac{1}{2} + 2 \times \frac{18}{7} = 7 \rightarrow x_1 = \frac{5}{14} \quad x = \begin{bmatrix} \frac{5}{14} \\ \frac{1}{2} \\ \frac{18}{7} \end{bmatrix}$$

● هزینه محاسباتی

هزینه محاسباتی روش حذفی گاوس - محاسبه $m_{i,k}$

مرحله	ضرب
1	$n - 1$
2	$n - 2$
$\vdots$	$\vdots$
$n - 1$	1

• محاسبه ماتریس U

مرحله	ضرب	جمع
1	$(n - 1)^2$	$(n - 1)^2$
2	$(n - 2)^2$	$(n - 2)^2$
$\vdots$	$\vdots$	$\vdots$
$n - 1$	1	1

• محاسبه b

مرحله	ضرب	جمع
1	$(n - 1)$	$(n - 1)$
2	$(n - 2)$	$(n - 2)$
$\vdots$	$\vdots$	$\vdots$
$n - 1$	1	1

• تعداد کل ضربها:

$$2 \sum_{k=1}^{n-1} (n - k) + \sum_{k=1}^{n-1} (n - k)^2 = \frac{n(n - 1)(2n + 5)}{6}$$

• تعداد کل جمعها:

$$\sum_{k=1}^{n-1} (n - k) + \sum_{k=1}^{n-1} (n - k)^2 = \frac{n(n - 1)(n + 1)}{3}$$

◀ 2.3.1 محاسبه معکوس یک ماتریس با استفاده از روش حذفی گاوس

روش حذفی گاوس را برای حل دستگاه معرفی کردیم اما یکی دیگر از کاربردهای این روش محاسبه معکوس یک ماتریس می‌باشد. برای محاسبه معکوس ماتریس A با استفاده از روش حذفی گاوس به ترتیب زیر عمل می‌کنیم. دستگاه‌های زیر را تشکیل می‌دهیم:

$$Ax_1 = e_1 \quad [A, e_1] \rightarrow [U, b_1]$$

$$Ax_2 = e_2 \quad [A, e_2] \rightarrow [U, b_2]$$

$$\vdots$$

$$Ax_n = e_n \quad [A, e_n] \rightarrow [U, b_n]$$

با حل این n دستگاه، x_i ها را محاسبه می‌کنیم. آنگاه:

$$A^{-1} = [x_1, x_2, \dots, x_n]_{n \times n}$$

درستی رابطه:

$$AA^{-1} = [Ax_1, Ax_2, \dots, Ax_n]$$

$$= [e_1, e_2, \dots, e_n]$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I$$

توجه کنید برای محاسبه هر کدام از x_i ها لازم است روش حذفی گاوس را روی ماتریس افزوده متناظر اعمال کنیم، یعنی:

$$[A, e_i] \rightarrow [U, b_i] \rightarrow x_i$$

اما دقت کنیم در اینجا ماتریس ضرایب تمام دستگاه‌های فوق یکسان می‌باشد. در اینجا به جای پیاده‌سازی n بار روش حذفی گاوس برای محاسبه x_i ها، بهتر است ابتدا تجزیه LU را برای ماتریس A به صورت جداگانه محاسبه کنیم و پس از آن با جای‌گذاری پسرو و پیشرو x_i ها را به روش زیر به دست آوریم:

$$A = LU$$

$$\underbrace{L U x_i}_{y_i} = e_i \implies \begin{cases} Ly_i = e_i \\ Ux_i = y_i \end{cases}$$

و با این روش هزینه محاسباتی برای محاسبه معکوس ماتریس A از مرتبه n^3 می‌باشد. اما اگر روش حذفی گاوس را n مرتبه تکرار اجرا کنیم هزینه محاسباتی از مرتبه n^4 می‌شود. در ادامه نحوه محاسبه تجزیه LU را بیان می‌کنیم.

نکته 2.2

نکته از روش ارائه شده در بالا می‌توان هر زمانی که لازم باشد چند دستگاه با ماتریس ضرایب یکسان حل شوند بهره برد.

2.4 محاسبه تجزیه LU ماتریس A

ماتریس A را در نظر می‌گیریم. با استفاده از الگوریتم گفته شده برای روش حذفی گاوس پیش می‌رویم و ماتریس U را به دست می‌آوریم.

$$A \mapsto U$$

و ماتریس L را به صورت زیر تعریف می‌کنیم:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ m_{21} & 1 & \cdots & 0 & 0 \\ m_{31} & m_{32} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & 1 \end{pmatrix}$$

قضیه 2.2

قضیه با تعاریف L و U ارائه شده در بالا داریم:

$$A = LU$$

عنصر سطر i -ام L و ستون j -ام U را محاسبه می‌کنیم. با توجه به تعریف ضرب ماتریسی داریم:

اثبات ✓

$$(LU)_{ij} = \sum_{k=1}^n L_{ik} U_{kj}$$

از آنجا که L پایین‌مثلثی ($L_{ik} = 0$ for $k > i$) و U بالامثلثی ($U_{kj} = 0$ for $k > j$) است، کران بالای سیگما برابر با $\min(i, j)$ خواهد بود.

رابطه بنیادی حذف گاوسی برای به‌روزرسانی درایه‌ها عبارت است از:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)} \implies m_{ik} a_{kj}^{(k)} = a_{ij}^{(k)} - a_{ij}^{(k+1)} \quad (*)$$

که در آن $U_{kj} = a_{kj}^{(k)}$ برای $k \leq j$. دو حالت را بررسی می‌کنیم:

حالت اول: $i \leq j$

در این حالت حاصل ضرب سطر i -ام L در ستون j -ام U تا $k = i$ ادامه دارد (چون $L_{ii} = 1$ و برای $k > i$ ، $L_{ik} = 0$):

$$\begin{aligned}(LU)_{ij} &= \sum_{k=1}^{i-1} m_{ik} u_{kj} + \underbrace{L_{ii}}_1 u_{ij} \\&= \sum_{k=1}^{i-1} m_{ik} a_{kj}^{(k)} + a_{ij}^{(i)} \\&\xrightarrow{(*)} \sum_{k=1}^{i-1} (a_{ij}^{(k)} - a_{ij}^{(k+1)}) + a_{ij}^{(i)} \\&= (a_{ij}^{(1)} - \cancel{a_{ij}^{(2)}}) + (\cancel{a_{ij}^{(2)}} - \cancel{a_{ij}^{(3)}}) + \cdots + (\cancel{a_{ij}^{(i-1)}} - \cancel{a_{ij}^{(i)}}) + \cancel{a_{ij}^{(i)}} \\&= a_{ij}^{(1)} = a_{ij}\end{aligned}$$

حالت دوم: $i > j$

در این حالت مجموع تا $k = j$ ادامه می‌یابد (چون برای $k > j$ ، $U_{kj} = 0$):

$$\begin{aligned}(LU)_{ij} &= \sum_{k=1}^j m_{ik} u_{kj} = \sum_{k=1}^{j-1} m_{ik} a_{kj}^{(k)} + m_{ij} a_{jj}^{(j)} \\&\xrightarrow{m_{ij} = a_{ij}^{(j)} / a_{jj}^{(j)}} \sum_{k=1}^{j-1} (a_{ij}^{(k)} - a_{ij}^{(k+1)}) + \underbrace{\frac{a_{ij}^{(j)}}{a_{jj}^{(j)}} a_{jj}^{(j)}}_{a_{ij}^{(j)}} \\&= (a_{ij}^{(1)} - \cancel{a_{ij}^{(2)}}) + \cdots + (\cancel{a_{ij}^{(j-1)}} - \cancel{a_{ij}^{(j)}}) + \cancel{a_{ij}^{(j)}} \\&= a_{ij}^{(1)} = a_{ij}\end{aligned}$$

در هر دو حالت، حاصل ضرب برابر با درایه اصلی ماتریس A شد.

تجزیه LU ماتریس A را بدست آورید.

*** مثال 2.2**

💡 حل:

$$\begin{aligned}A &= \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ -1 & -3 & 4 \end{pmatrix} \quad m_{21} = \frac{2}{1} \quad m_{31} = \frac{-1}{1} \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \frac{1}{3} & 1 \end{pmatrix} \\&\rightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & 1 \\ 0 & -1 & 5 \end{pmatrix} \xrightarrow{m_{32} = \frac{-1}{-3}} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & 1 \\ 0 & 0 & \frac{14}{3} \end{pmatrix} = U\end{aligned}$$

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & 1 \\ 0 & 0 & \frac{14}{3} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ -1 & -3 & 4 \end{pmatrix} \checkmark$$

2.5 محوری جزئی و کلی

2.5.1 محوری جزئی

در روش حذفی گاوس برای حل دستگاه در هر مرحله k ضرایب m_{ik} ($i = k + 1, \dots, n$) را محاسبه می‌کنیم و m_{ik} برابر سطر k ام را به سطر i ام اضافه می‌کنیم. اگر m_{ik} ها مقدار بزرگی داشته باشند باعث ایجاد خطا در جواب نهایی می‌شوند. ما می‌خواهیم از این امر جلوگیری کنیم و در هر مرحله $|m_{ik}| \leq 1$ باشد. در محورگیری جزئی در ابتدای مرحله k ام ستون k ام و درایه‌های زیر عنصر قطری را بررسی می‌کنیم و درایه‌ای که بزرگترین قدر مطلق را دارد را مشخص می‌کنیم. فرض کنیم در سطر r ام اتفاق می‌افتد. آنگاه سطر r ام را با سطر k ام جایگزین می‌کنیم. در ابتدای هر مرحله از روش حذفی گاوس محورگیری جزئی را انجام می‌دهیم.

2.5.2 محوری کلی

در این روش در هر مرحله قبل از شروع روش حذفی گاوس در بین درایه‌هایی که قرار است در این مرحله به‌روز شوند جستجو می‌کنیم و بزرگترین درایه از لحاظ قدر مطلق را مشخص می‌کنیم. فرض کنیم در مرحله k ام هستیم و این درایه در سطر p ام و ستون q ام قرار دارد. حال سطر p ام را با سطر k ام و ستون q ام را با ستون k ام جابجا می‌کنیم. بنابراین بزرگترین درایه از لحاظ قدر مطلق اکنون در جایگاه عضو قطری مورد نظر قرار دارد. حال عملیات حذفی گاوس را ادامه می‌دهیم. عملیات محورگیری کلی در حالت کلی هزینه محاسباتی زیادی دارد (برای مشخص کردن بزرگترین درایه) و معمولاً در محاسبات دستی یا با استفاده از کامپیوتر از این روش استفاده نمی‌شود.

نکته 2.3

نکته در الگوریتم روش حذفی گاوس در هر مرحله k ، احتیاج داریم شرط زیر برقرار باشد:

$$a_{kk}^{(k)} \neq 0, \quad k = 1, 2, \dots, n-1$$

اگر در مرحله‌ای این شرط برقرار نباشد، درایه‌های زیر عنصر $a_{kk}^{(k)}$ در ستون k ام را بررسی می‌کنیم و عضو غیر صفری در بین این درایه‌ها تعیین می‌کنیم. فرض کنیم در سطر r ام عضو غیر صفری داریم، در این صورت سطر r ام را با سطر k ام جابجا می‌کنیم (مجاز به انجام این کار هستیم چون جزء اعمال سطری مقدماتی است) و عملیات را ادامه می‌دهیم.

2.6 ماتریس جایگشت

- ماتریسی است که از جابجایی سطرها یا ستون‌های ماتریس همانی حاصل می‌شود.
- به عبارت دیگر، P ماتریسی جایگشت است هرگاه در هر سطر و ستون آن تنها یک درایه ۱ داشته باشیم و بقیه درایه‌ها صفر باشند.

≡ ویژگی‌ها 2.2 | ماتریس جایگشت

۱. اگر P, Q ماتریس‌های جایگشت باشند، آنگاه PQ, QP نیز جایگشت می‌باشند.
۲. دترمینان هر ماتریس جایگشت ± 1 می‌باشد.
۳. فرض کنیم ماتریس جایگشت P از جابجایی سطر i ام و سطر j ام ماتریس همانی حاصل شده باشد. آنگاه PA همان ماتریس A می‌باشد که سطر i ام و سطر j ام آن جابجا شده‌اند.
۴. اگر P یک ماتریس جایگشت باشد، آنگاه $P^2 = I$.

* مثال 2.3

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$PA = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

≡ 2.7 مقیاس‌دهی

عملیات مقیاس‌دهی یک‌بار و قبل از شروع روش حذفی گاوس انجام می‌شود. به این ترتیب که برای تمام سطرها عملیات زیر را انجام می‌دهیم: در سطر i ام جستجو می‌کنیم و بزرگترین درایه از لحاظ قدر مطلق را مشخص می‌کنیم و بعد همه درایه‌های سطر i ام را به این درایه تقسیم می‌کنیم (این عملیات را در ماتریس افزوده انجام می‌دهیم). با این عملیات تمام درایه‌ها در بازه $[-1, 1]$ قرار می‌گیرند و بعد روش حذفی گاوس را آغاز می‌کنیم.

≡ 2.8 روش گاوس-ژوردان (جردن)

در این روش مشابه با روش حذفی گاوس ماتریس افزوده دستگاه را تشکیل می‌دهیم و بعد با انجام عملیات سطری مقدماتی آن را به فرم ماتریسی که سمت چپ آن ماتریس همانی است تبدیل می‌کنیم.

$$[A, b] \mapsto [I, \bar{b}] \Rightarrow X = b$$

و همانند روش حذفی گاوس، روش گاوس-ژوردان می‌تواند با محورگیری جزئی، محورگیری کلی یا عملیات مقیاس‌دهی انجام می‌شود.

مثال: جواب دستگاه زیر را با روش گاوس-ژوردان و محورگیری جزئی بدست آورید.

* مثال 2.4

$$x_1 + x_2 + x_3 + x_4 = -2$$

$$x_1 + 2x_2 + 3x_3 + 4x_4 = -10$$

$$2x_1 + 3x_2 + 4x_3 + x_4 = 4$$

$$3x_1 + 4x_2 + x_3 + 2x_4 = -10$$

💡 حل:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ -10 \\ 4 \\ -10 \end{bmatrix}$$

$$\rightsquigarrow \left[\begin{array}{cccc|c} 3 & 4 & 1 & 2 & -10 \\ 1 & 2 & 3 & 4 & -10 \\ 2 & 3 & 4 & 1 & 4 \\ 1 & 1 & 1 & 1 & -2 \end{array} \right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & \frac{-10}{3} \\ 0 & \frac{2}{3} & \frac{8}{3} & \frac{10}{3} & \frac{-20}{3} \\ 0 & \frac{1}{3} & \frac{10}{3} & -\frac{1}{3} & \frac{32}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} & \frac{4}{3} \end{array} \right]$$

$$\rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -5 & -6 & 10 \\ 0 & 1 & 4 & 5 & -10 \\ 0 & 0 & 2 & -2 & 14 \\ 0 & 0 & 2 & 2 & -2 \end{array} \right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & -11 & 45 \\ 0 & 1 & 0 & 9 & -38 \\ 0 & 0 & 1 & -1 & 7 \\ 0 & 0 & 0 & 4 & -16 \end{array} \right]$$

$$\rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -4 \end{array} \right] \rightsquigarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \\ -4 \end{bmatrix}$$

نکته 2.4

یکی از کاربردهای روش گاوس-ژوردان محاسبه معکوس ماتریس می‌باشد. برای این منظور ماتریس افزوده زیر را تشکیل می‌دهیم. و با انجام اعمال سطری مقدماتی و عملیات مربوط به گاوس-ژوردان، ماتریس سمت چپ را به همانی تبدیل می‌کنیم.

$$[A, I] \mapsto [I, B] \Rightarrow B = A^{-1}$$

معکوس ماتریس زیر را به روش گاوس-ژوردان به دست آورید.

* مثال 2.5

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

حل:

$$\begin{aligned} & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ -10 \\ 4 \\ -10 \end{bmatrix} \\ & \rightsquigarrow \left[\begin{array}{cccc|c} 3 & 4 & 1 & 2 & -10 \\ 1 & 2 & 3 & 4 & -10 \\ 2 & 3 & 4 & 1 & 4 \\ 1 & 1 & 1 & 1 & -2 \end{array} \right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & \frac{-10}{3} \\ 0 & \frac{2}{3} & \frac{8}{3} & \frac{10}{3} & \frac{-20}{3} \\ 0 & \frac{1}{3} & \frac{10}{3} & -\frac{1}{3} & \frac{32}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} & \frac{4}{3} \end{array} \right] \\ & \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & -5 & -6 & 10 \\ 0 & 1 & 4 & 5 & -10 \\ 0 & 0 & 2 & -2 & 14 \\ 0 & 0 & 2 & 2 & -2 \end{array} \right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & -11 & 45 \\ 0 & 1 & 0 & 9 & -38 \\ 0 & 0 & 1 & -1 & 7 \\ 0 & 0 & 0 & 4 & -16 \end{array} \right] \\ & \rightsquigarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -4 \end{array} \right] \rightsquigarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \\ -4 \end{bmatrix} \end{aligned}$$

نکته 2.5

هزینه محاسباتی گاوس-ژوردان نسبت به روش حذفی گاوس برای حل دستگاه $Ax = b$ بیشتر است.

تعریف 2.2 | ماتریس مثلثی یکه

ماتریس مثلثی (بالا یا پایین) که روی قطر اصلی آن همه درایه‌ها برابر با ۱ باشند را ماتریس مثلثی یکه می‌نامیم.

نکته 2.6 | نکته

• حاصل ضرب دو ماتریس پایین (بالا) مثلثی یک ماتریس پایین (بالا) مثلثی است. • حاصل ضرب دو ماتریس پایین (بالا) مثلثی یکه، یک ماتریس پایین (بالا) مثلثی یکه است.

قضیه 2.3

اگر ماتریس‌های پیشروی اصلی A برای $1 \leq k \leq n-1$ معکوس‌پذیر باشند، تجزیه LU وجود دارد و اگر خود ماتریس A معکوس‌پذیر باشد، تجزیه LU یکتاست. (یکتایی قضیه را با این شرط اثبات می‌کنیم که روی قطر اصلی L همه درایه‌ها ۱ هستند)

اثبات

ماتریس ضرایب در هر مرحله از تجزیه LU را با $A^{(k)}$ نشان می‌دهیم برای $k = 1, 2, \dots, n$ می‌دانیم:

$$\det(A) = \det(A^{(k)}), \quad k = 1, 2, \dots, n$$

بنابراین هر کدام از زیر ماتریس‌های پیشروی اصلی A را در نظر بگیریم در شرط زیر صدق می‌کند

$$\det(A_k) = \det(A_k^{(k)}), \quad k = 1, 2, \dots, n$$

• $k = 1$

$$0 \neq \det(A_1) = a_{11} = \det(A_1^{(1)}) = a_{11}^{(1)} \Rightarrow a_{11}^{(1)} \neq 0$$

• $k = 2$

$$0 \neq \det(A_2) = \det(A_2^{(2)}) = a_{11}^{(1)} \cdot a_{22}^{(2)} \Rightarrow a_{22}^{(2)} \neq 0$$

• اگر به همین ترتیب ادامه دهیم، در مرحله k ام داریم:

$$\det(A_k) = \det(A_k^{(k)}) = a_{11}^{(1)} \cdot a_{22}^{(2)} \cdot \dots \cdot a_{kk}^{(k)} \Rightarrow a_{kk}^{(k)} \neq 0$$

• و در مرحله $(n-1)$ ام داریم:

$$\det(A_{n-1}) = \det(A_{n-1}^{(n-1)}) = a_{11}^{(1)} \cdot a_{22}^{(2)} \cdot \dots \cdot a_{(n-1)(n-1)}^{(n-1)} \Rightarrow a_{(n-1)(n-1)}^{(n-1)} \neq 0$$

نشان دادیم که:

$$a_{11}^{(1)}, a_{22}^{(2)}, \dots, a_{(n-1)(n-1)}^{(n-1)} \neq 0$$

پس شرط لازم برای انجام تجزیه LU برقرار است پس تجزیه LU وجود دارد. حال فرض کنیم ماتریس A معکوس‌پذیر است، می‌خواهیم یکتایی تجزیه LU را اثبات کنیم. به خلاف فرض می‌کنیم تجزیه یکتا نباشد پس:

$$\exists L_1, L_2, U_1, U_2 \quad s.t. \quad A = L_1 U_1 = L_2 U_2 \quad and \quad L_1 \neq L_2, U_1 \neq U_2$$

ماتریس L_1, L_2 پایین‌مثلثی یک‌ه هستند یعنی:

$$\det(L_1) = \det(L_2) = 1$$

$$A = L_1 U_1 \Rightarrow 0 \neq \det(A) = \det(L_1) \det(U_1) = \det(U_1) \neq 0$$

به همین ترتیب:

$$\det(U_2) \neq 0$$

$$L_1 U_1 = L_2 U_2$$

$$L_1^{-1} L_1 U_1 U_2^{-1} = L_1^{-1} L_2 U_2 U_2^{-1}$$

$$U_1 U_2^{-1} = L_1^{-1} L_2$$

ماتریس‌های $U_1 U_2^{-1}$ بالامثلثی هستند، پس ضرب آنها نیز بالامثلثی است. و ماتریس‌های $L_1^{-1} L_2$ پایین‌مثلثی یک‌ه هستند، پس ضرب آنها نیز پایین‌مثلثی یک‌ه است. پایین‌مثلثی یک‌ه = بالامثلثی $\Leftarrow$ ماتریس یک‌ه I

$$U_1 U_2^{-1} = L_1^{-1} L_2 = I$$

$$\Rightarrow U_1 = U_2, L_1 = L_2 \quad \blacksquare$$

2.8.1 < صورت‌های دیگر تجزیه LU

• تجزیه‌ای که تاکنون برای LU بیان کردیم به این ترتیب بود که درایه‌های روی قطر اصلی L همگی برابر 1 بودند. این تجزیه به تجزیه دولیتل معروف می‌باشد.
• تجزیه LU را می‌توان به ترتیبی دیگر نیز نوشت که روی قطر اصلی U همه درایه‌ها یک باشند. این تجزیه را تجزیه کروت می‌گوییم.

نکته 2.7

شرایط بیان شده در قضیه برای وجود تجزیه LU شرایط کافی می‌باشند اما لازم نیستند. به عنوان مثال:

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

شرایط قضیه فوق را ندارد اما با جابه‌جایی سطرها می‌توان آن را به فرمی نوشت که شرایط قضیه را داشته باشد.

$$A = \begin{pmatrix} 3 & 4 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix}$$

قضیه 2.4

فرض کنیم $n \geq 2$, $A \in \mathbb{R}^{n \times n}$ باشد آنگاه همواره ماتریس جایگشت P وجود دارد به طوری که

$$PA = LU$$

و L یک ماتریس پایین‌مثلثی یک‌ه و U یک ماتریس بالامثلثی می‌باشد.

اثبات

با استقرا اثبات می‌کنیم. ابتدا حالت $n = 2$ را در نظر می‌گیریم.

$$A_{2 \times 2} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$L_{2 \times 2} = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$$

$$U_{2 \times 2} = \begin{pmatrix} \beta & \gamma \\ 0 & \lambda \end{pmatrix}$$

$$1. \ a \neq 0, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$PA = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} \beta & \gamma \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \beta & \gamma \\ \alpha\beta & \alpha\gamma + \lambda \end{pmatrix}$$

$$\Rightarrow \beta = a, \quad \gamma = b, \quad \alpha = \frac{c}{a}, \quad \lambda = d - \frac{c}{a}b$$

$$\Rightarrow L = \begin{pmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} a & b \\ 0 & d - \frac{c}{a}b \end{pmatrix} \quad \square$$

$$2. \ a = 0, \quad c \neq 0, \quad P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & b \end{pmatrix}$$

$$\Rightarrow L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} c & d \\ 0 & b \end{pmatrix} \quad \square$$

$$3. \ a = c = 0, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$PA = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix}$$

$$\Rightarrow L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \quad \square$$

فرض می‌کنیم قضیه به‌ازای $1 \leq k \leq n-1$ برقرار باشد. می‌خواهیم نشان دهیم که برای ماتریس A با ابعاد $(k+1) \times (k+1)$ نیز برقرار است.

فرض کنیم A یک ماتریس $(k+1) \times (k+1)$ باشد و بزرگ‌ترین درایه از لحاظ قدر مطلق در ستون اول، در سطر r ام قرار داشته باشد. سطر r ام را با سطر اول ماتریس جابه‌جا می‌کنیم، یعنی ماتریس A را در یک ماتریس جایگشت ضرب می‌کنیم. این ماتریس جایگشت را با P^r نشان می‌دهیم.

$$P^r A = \begin{pmatrix} \alpha_{1 \times 1} & w_{1 \times k}^T \\ q_{k \times 1} & B_{k \times k} \end{pmatrix}_{(k+1) \times (k+1)}$$

$$= \begin{pmatrix} 1 & 0 \\ m & I \end{pmatrix} \begin{pmatrix} \alpha & v^T \\ 0 & C \end{pmatrix}$$

$$\Rightarrow \alpha = \alpha, \quad v^T = w^T$$

$$\alpha m = q, \quad mv^T + C = B$$

$$\text{if } \alpha = 0 \text{ then } q = 0 \Rightarrow m = 0, B = C$$

$$\text{if } \alpha \neq 0 \Rightarrow m = \frac{1}{\alpha}q, \quad C = B - mv^T$$

ماتریس C یک ماتریس $k \times k$ است؛ پس طبق فرضِ استقرا، ماتریس‌های L^*, P^* و U^* که همگی از بعد $k \times k$ هستند، وجود دارند به‌طوری‌که:

$$P^*C = L^*U^*$$

حال داریم:

$$\begin{aligned} P^r A &= \begin{pmatrix} 1 & 0 \\ 0 & P^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ P^*m & L^* \end{pmatrix} \begin{pmatrix} \alpha & v^T \\ 0 & U^* \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & P^* \end{pmatrix} \begin{pmatrix} \alpha & v^T \\ \alpha P^*m & P^*mv^T + L^*U^* \end{pmatrix} \\ &= \begin{pmatrix} \alpha & v^T \\ \alpha P^*P^*m & P^*P^*mv^T + P^*P^*C \end{pmatrix} \\ &= \begin{pmatrix} \alpha & v^T \\ \alpha m & mv^T + C \end{pmatrix} \\ &= \begin{pmatrix} \alpha & v^T \\ q & B \end{pmatrix} \end{aligned}$$

بنابراین:

$$P^r A = \begin{pmatrix} 1 & 0 \\ 0 & P^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ P^*m & L^* \end{pmatrix} \begin{pmatrix} \alpha & v^T \\ 0 & U^* \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & P^* \end{pmatrix}}_P P^r A = \begin{pmatrix} 1 & 0 \\ P^* m & L^* \end{pmatrix} \begin{pmatrix} \alpha & v^T \\ 0 & U^* \end{pmatrix}$$

$$PA = LU \quad \blacksquare$$

2.8.2 < تجزیه LU ماتریس سه قطری

یک ماتریس سه قطری در حالت کلی به صورت زیر تعریف می‌شود:

$$T = \begin{pmatrix} b_1 & c_1 & 0 & \cdots & 0 & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 & 0 \\ 0 & a_3 & b_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \cdots & a_n & b_n \end{pmatrix}_{n \times n}$$

$$a_1 = c_n = 0$$

می‌خواهیم برای این ماتریس تجزیه LU معرفی کنیم که L یک ماتریس پایین‌مثلثی یک‌ه و U یک ماتریس بالامثلثی می‌باشد. یعنی داریم:

$$\begin{pmatrix} b_1 & c_1 & 0 & \cdots & 0 & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 & 0 \\ 0 & a_3 & b_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \cdots & a_n & b_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \ell_2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \ell_3 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & \ell_n & 1 \end{pmatrix} \begin{pmatrix} u_1 & v_1 & 0 & \cdots & 0 & 0 \\ 0 & u_2 & v_2 & \cdots & 0 & 0 \\ 0 & 0 & u_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{n-1} & v_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & u_n \end{pmatrix}$$

کافی است مقادیر l_i, u_i, v_i را مشخص کنیم. ضرب سطر اول L در ستون اول U :

$$b_1 = u_1$$

ضرب سطر اول L در ستون دوم U :

$$c_1 = v_1$$

ضرب سطر دوم L در ستون اول U :

$$l_2 u_1 = a_2 \Rightarrow l_2 = \frac{a_2}{u_1}, b \neq 0$$

ضرب سطر دوم L در ستون دوم U :

$$b_2 = l_2 v_1 + u_2 \Rightarrow u_2 = b_2 - l_2 v_1$$

تعریف 2.3 | ماتریس قطر غالب

فرض می‌کنیم A یک ماتریس $n \times n$ باشد.

$$A = (a_{ij}), \quad 1 \leq i, j \leq n$$

گوییم A یک ماتریس قطر غالب است هرگاه:

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}|, \quad \forall i = 1, 2, \dots, n$$

و اگر رابطه فوق به صورت اکید برقرار باشد، گوییم A غالب قطر اکید است.

قضیه 2.5

فرض کنیم T یک ماتریس $n \times n$ سه‌قطری باشد. اگر T غالب قطر اکید باشد آنگاه تجزیه LU برای T وجود دارد و این تجزیه بدون جابه‌جایی سطر قابل انجام است.

$$|b_i| > |a_i| + |c_i|, \quad i = 1, \dots, n$$

2.9 تجزیه چولسکی

2.9.1 < ماتریس معین مثبت

تعریف 2.4

فرض کنیم A یک ماتریس متقارن $n \times n$ باشد. گوییم ماتریس A معین مثبت است هرگاه:

$$\forall x \in \mathbb{R}^n, x \neq 0: \quad x^T A x > 0$$

و A را ماتریس نیمه معین مثبت می‌گوییم هرگاه:

$$\forall x \in \mathbb{R}^n: \quad x^T A x \geq 0$$

نشان دهید ماتریس زیر معین مثبت است

* مثال 2.6

$$\begin{pmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

💡 حل:

$$\begin{aligned} x^T A x &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^T \begin{pmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 + 2x_2 + x_3 & 2x_1 + 5x_2 + x_3 & x_1 + x_2 + 2x_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= 2x_1^2 + 2x_1x_2 + x_1x_3 + 2x_1x_2 + 5x_2^2 + x_2x_3 + x_1x_3 + x_2x_3 + 2x_3^2 \\ &= 2x_1^2 + 5x_2^2 + 2x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3 \\ &= (x_1^2 + 4x_2^2 + 4x_1x_2) + (x_1^2 + x_3^2 + 2x_1x_3) + (x_2^2 + x_3^2 + 2x_2x_3) \\ &= (x_1 + 2x_2)^2 + (x_1 + x_3)^2 + (x_2 + x_3)^2 > 0 \end{aligned}$$

ویژگی‌های ماتریس معین مثبت ($A_{n \times n}$)

۱. تمام درایه‌های قطری ماتریس‌های معین مثبت، مثبت می‌باشد.
۲. تمام مقادیر ویژه ماتریس‌های معین مثبت، مثبت می‌باشد.
۳. دترمینان هر ماتریس معین مثبت، مثبت است.
۴. هر زیر ماتریس از ماتریس معین مثبت، خود معین مثبت است.
۵. به ازای هر i, j داریم

$$a_{ij}^2 < a_{ii}a_{jj}$$

۶. بزرگترین درایه از لحاظ قدرمطلق در ماتریس معین مثبت روی قطر قرار دارد.

۱ | اثبات ✓

$$\forall x \in \mathbb{R}^n, x \neq 0 \quad x^T A x > 0$$

$$x = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_1^T A e_1 = a_{11} > 0$$

$$x = e_i \quad e_i^T A e_i = a_{ii} > 0$$

۲ | اثبات ✓

$$\exists x \in \mathbb{R}^n, x \neq 0 \text{ s.t. } Ax = \lambda x$$

$$x^T A x = \lambda x^T x \Rightarrow \lambda = \frac{x^T A x}{\|x\|^2} > 0$$

می‌دانیم دترمینان ماتریس، حاصل ضرب مقادیر ویژه آن ماتریس است. در قسمت قبل نشان

۳ | اثبات ✓

دادیم که تمام مقادیر ویژه، مثبت هستند؛ پس حاصل ضرب آن‌ها، یعنی دترمینان، نیز مثبت است.

۴ | اثبات ✓

فرض کنیم $B_{m \times m}$ یک زیرماتریس از $A_{n \times n}$ باشد، به طوری که $m < n$. برای مثال:

$$A_{5 \times 5} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

$$B = \begin{pmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{pmatrix}$$

- بردار y را به دلخواه در $\mathbb{R}^m$ انتخاب می‌کنیم. بردار $x \in \mathbb{R}^n$ را به صورت زیر تعریف می‌کنیم:
- درایه‌هایی از x که سطر و ستون متناظر با آن‌ها در زیرماتریس ظاهر نشده‌اند، برابر با صفر قرار می‌دهیم.
 - درایه‌هایی از x که سطر و ستون متناظر با آن‌ها در زیرماتریس ظاهر شده‌اند، با درایه‌های y پر می‌کنیم.

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Rightarrow x = \begin{pmatrix} y_1 \\ 0 \\ y_2 \\ y_3 \\ 0 \end{pmatrix}$$

در این صورت:

$$y^T B y = x^T A x > 0$$

$$\Rightarrow \forall y \in \mathbb{R}^m \quad y^T B y > 0$$

در نتیجه، B نیز معین مثبت است.

تمام سطرها و ستون‌های ماتریس معین مثبت A را حذف می‌کنیم به جز سطرهای i ام و ستون‌های j ام. پس ماتریس B حاصل به صورت زیر است (با فرض $i < j$):

$$B = \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$$

طبق بخش ۴، B معین مثبت است. طبق بخش ۳:

$$\det(B) > 0$$

از رابطه دترمینان ماتریس 2×2 داریم:

$$\det(B) = a_{ii}a_{jj} - a_{ij}a_{ji} = a_{ii}a_{jj} - a_{ij}^2 > 0$$

اثبات ۶

$$M = \max_{1 \leq i \leq n} |a_{ii}|$$

طبق بخش ۵ داریم:

$$\forall i, j = 1, \dots, n \quad a_{ij}^2 \leq a_{ii}a_{jj} \leq M^2$$

$$\Rightarrow a_{ij} \leq |a_{ij}| \leq M$$

$$\Rightarrow |a_{ij}| \leq M \quad \forall i, j = 1, \dots, n$$

2.9.2 < تجزیه چولسکی

قضیه 2.6 | تجزیه چولسکی

فرض کنیم A یک ماتریس معین مثبت است، آنگاه یک ماتریس بالامثلثی با درایه‌های قطری مثبت (R) وجود دارد به‌طوری‌که:

$$A = R^T R$$

تجزیه چولسکی را به‌صورت مستقیم تعریف می‌کنیم. فرض کنیم A یک ماتریس $n \times n$ باشد، ماتریس R را در فرم کلی می‌نویسیم و درایه‌های آن را تعیین می‌کنیم.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} r_{11} & 0 & 0 & \cdots & 0 \\ r_{12} & r_{22} & 0 & \cdots & 0 \\ r_{13} & r_{23} & r_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{1n} & r_{2n} & r_{3n} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

$$a_{11} = r_{11}^2 \Rightarrow r_{11} = \sqrt{a_{11}}$$

$$a_{12} = r_{11}r_{12} \Rightarrow r_{12} = \frac{a_{12}}{r_{11}}$$

$$a_{1i} = r_{11}r_{1i} \Rightarrow r_{1i} = \frac{a_{1i}}{r_{11}} \quad i = 2, \dots, n$$

$$r_{11}r_{12} = a_{21} = a_{12}$$

$$r_{12}^2 + r_{22}^2 = a_{22} \Rightarrow r_{22} = \sqrt{a_{22} - r_{12}^2}$$

$$r_{12}r_{13} + r_{22}r_{23} = a_{23} \Rightarrow r_{23} = \frac{a_{23} - r_{12}r_{13}}{r_{22}}$$

$$r_{2i} = \frac{a_{2i} - r_{12}r_{1i}}{r_{22}} \quad i = 3, \dots, n$$

در حالت کلی الگوریتم تجزیه چولسکی به صورت زیر می‌باشد:

for $k = 1, 2, \dots, n$

$$r_{kk} = \sqrt{a_{kk} - \sum_{j=1}^{k-1} r_{jk}^2}$$

for $i = k+1, \dots, n$

$$r_{ki} = \frac{1}{r_{kk}} \left(a_{ki} - \sum_{j=1}^{k-1} r_{ij}r_{kj} \right)$$

end for

end for

تجزیه چولسکی ماتریس زیر را بدست آورید.

* مثال 2.7

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} r_{11} & 0 & 0 \\ r_{12} & r_{22} & 0 \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}$$

$$r_{11}^2 = 2 \Rightarrow r_{11} = \sqrt{2}$$

$$r_{11}r_{12} = 2 \Rightarrow r_{12} = \sqrt{2}$$

$$r_{11}r_{13} = 1 \Rightarrow r_{13} = \frac{\sqrt{2}}{2}$$

$$r_{12}^2 + r_{22}^2 = 5 \Rightarrow r_{22} = \sqrt{3}$$

$$r_{12}r_{13} + r_{22}r_{23} = 1 \Rightarrow r_{23} = 0$$

$$r_{13}^2 + r_{23}^2 + r_{33}^2 = 2 \Rightarrow r_{33} = \sqrt{\frac{3}{2}}$$

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{\frac{3}{2}} \end{pmatrix}$$

قضیه 2.7 | قضیه چولسکی

اگر $A \in \mathbb{R}^{n \times n}$ یک ماتریس معین مثبت باشد، ماتریس یکتای بالامتثلی با درایه‌های قطری مثبت R وجود دارد، به طوری که:

$$A = R^T R$$

اثبات ✓

با استقرا روی n (بعد ماتریس)

برای $n = 1$ واضح است. در این حالت، $A = (a_{11})$. چون در فرض بیان شده که A معین مثبت است، پس داریم:

$$a_{11} > 0$$

$$\Rightarrow n = 1 \quad A = (a_{11}) = (\sqrt{a_{11}})(\sqrt{a_{11}})$$

فرض می‌کنیم حکم برای $n - 1$ برقرار است، و می‌خواهیم آن را برای ماتریسی A با ابعاد $n \times n$ اثبات کنیم.

$$A_{n-1} = A(1 : n-1, 1 : n-1) \quad A_{n-1} \in \mathbb{R}^{(n-1)(n-1)}$$

$$A = \begin{pmatrix} A_{n-1} & c \\ c^T & \alpha \end{pmatrix}_{n \times n}$$

می‌دانیم A_{n-1} معین مثبت است؛ پس طبق فرض، تجزیه فوق را داریم:

$$A_{n-1} = R_{n-1}^T R_{n-1}$$

$$A = \begin{pmatrix} A_{n-1} & c \\ c^T & \alpha \end{pmatrix} = \begin{pmatrix} R_{n-1}^T & 0 \\ r^T & \beta \end{pmatrix} \begin{pmatrix} R_{n-1} & r \\ 0 & \beta \end{pmatrix} := R^T R$$

کافی است بردار r و عدد β را تعیین کنیم.

$$R_{n-1}^T r = c$$

از آن جا که ماتریس R_{n-1}^T معکوس‌پذیر است، بردار r به صورت یکتا به دست می‌آید.

$$r^T r + \beta^2 = \alpha \Rightarrow \beta^2 = \alpha - r^T r$$

در این جا β را می‌توان هم مثبت و هم منفی در نظر گرفت، اما برای تضمین یکتایی، همواره آن را مثبت در نظر می‌گیریم. حال اثبات می‌کنیم که $\beta = 0$ اتفاق نمی‌افتد،

$$0 < \det(A) = \det(R_{n-1})^2 \cdot \beta^2 \Rightarrow \beta^2 > 0 \Rightarrow \beta > 0$$

و در نتیجه، خوش تعریفی حاصل می‌شود.

2.10 $\equiv$ کمترین مربعات

دو مثال برای درک بهتر روش کمترین مربعات

معادله خطی را بیابید که از چهار نقطه زیر عبور کند.

*** مثال 2.8**

x_i	1	4	-1	-2
y_i	-1	11	-9	-13

💡 حل:

فرم کلی معادله خط یعنی $y(x) = mx + n$ را در نظر می‌گیریم. برای اینکه خط مورد نظر از این نقاط عبور کند باید مختصات نقاط در آن صدق کند. با قراردادن هر یک از نقاط بالا در معادله خط داریم:

$$\begin{aligned} m + n &= -1 \\ 4m + n &= 11 \\ -m + n &= -9 \\ -2m + n &= -13 \end{aligned} \Rightarrow \begin{pmatrix} 1 & 1 \\ 4 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} -1 \\ 11 \\ -9 \\ -13 \end{pmatrix}$$

با حل معادلات فوق جواب $m = 4$ و $n = -5$ را بدست می‌آوریم بنابراین معادله خط به صورت $y(x) = 4x - 5$ می‌شود. این نمونه از دستگاه‌ها را سازگار می‌نامیم. سازگار بودن یک دستگاه به صورت زیر بررسی می‌شود.

$$Ax = b \Rightarrow \text{rank}(A) = \text{rank}(A, b) = 2$$

معادله خطی را بیابید که از چهار نقطه زیر عبور کند.

*** مثال 2.9**

x_i	-3	1	-7	5
y_j	70	21	110	-35

💡 حل:

همانند آنچه در مثال قبل انجام دادیم در اینجا نیز انجام می‌دهیم. داریم:

$$\begin{aligned} -3m + n &= 70 \\ m + n &= 21 \\ -7m + n &= 110 \\ 5m + n &= -35 \end{aligned} \Rightarrow \begin{pmatrix} -3 & 1 \\ 1 & 1 \\ -7 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} 70 \\ 21 \\ 110 \\ -35 \end{pmatrix}$$

این دستگاه ناسازگار است و پاسخی برای آن وجود ندارد. پس بر خلاف مثال قبل نمی‌توان یک خط بدست آورد که از این چهار نقطه عبور کند. دقت کنیم:

$$\text{rank}(A) = 2 \quad \& \quad \text{rank}(A, b) = 3$$

در اینجا خطی را بدست می‌آوریم که از این چهار نقطه کمترین فاصله را داشته باشد.

2.10.1 < جواب‌های کمترین مربعات توان‌های دوم برای دستگاه‌های خطی

در بخش قبلی چندین روش برای حل دستگاه خطی $Ax = b$ مورد بحث قرار دادیم و فرض کردیم A مربعی و نامنفرد باشد. ولی در برخی مسائل علمی نظیر کاربرد های آماری، مدلسازی هندسی و پردازش سیگنال نیاز به حل یک دستگاه داریم که در آن ماتریس A نامربعی و یا منفرد است. در چنین حالت‌هایی ممکن است به هیچ وجه جواب وجود نداشته باشد و در حالت‌هایی که جواب داریم ممکن است بی‌نهایت جواب وجود داشته باشد. برای مثال هنگامی که A یک ماتریس $m \times n$ با $m > n$ است یک دستگاه فرامعین داریم (یعنی تعداد معادلات بزرگتر از تعداد مجهولات است) و یک دستگاه فرامعین معمولاً جواب ندارد بر عکس یک دستگاه فرومعین $m < n$ معمولاً یک یا تعداد نامتناهی جواب دارد.

در این حالت‌ها بهترین وضعیتی که انتظارش را میتوان داشت این است که یک بردار x پیدا کنیم که Ax را تا حد ممکن به بردار b نزدیک سازد. به بیان دیگر، یک بردار x را جستجو می‌کنیم به قسمی که $\|r(x)\| = \|Ax - b\|$ را کمینه سازد. هنگامی که نرم اقلیدسی $\|\cdot\|$ استفاده می‌شود به این جواب به عنوان یک جواب کمترین توان‌های دوم برای دستگاه $Ax = b$ اطلاق می‌شود.

وجود و یکتایی جواب

نظیر حالت حل دستگاه خطی، طبیعتاً سوال های زیر مطرح می‌شود:

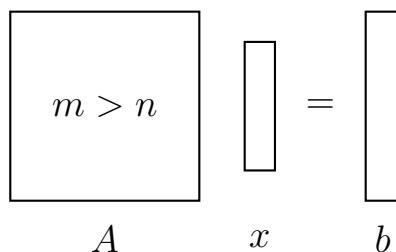
۱. آیا همیشه یک جواب کمترین توان‌های دوم برای $Ax = b$ وجود دارد؟

۲. اگر وجود دارد، آیا یکتاست؟

۳. چگونه چنین جوابی را بدست آوریم؟

فرض کنید که دستگاهی $Ax = b$ فرامعین باشد یعنی A از مرتبه $m \times n$ باشد که در آن $m \geq n$. یک دستگاه

فرامعین $Ax = b$ را میتوان به صورت نموداری به صورت زیر نمایش دهیم:



قضیه 2.8 | وجود و یکتایی جواب

یک جواب برای مسئله کمترین توان های دوم وجود دارد. این جواب یکتاست اگر و فقط اگر A دارای رتبه کامل باشد. یعنی

$$\text{rank}(A) = n$$

اگر A دارای رتبه ناقص باشد، آنگاه مسئله کمترین توان های دوم دارای جواب‌های بی شماری خواهد بود.

لم 2.1

x یک جواب کمترین توان های دوم $Ax = b$ است اگر و فقط اگر در رابطه زیر صدق می‌کند. فرض کنیم A از رتبه کامل n باشد. ($A^T A$ معین مثبت است و تجزیه چولسکی را داریم)

$$A^T A x = A^T b$$

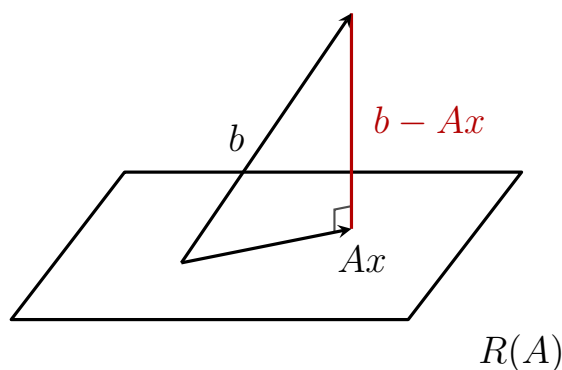
2.10.2 < تعبیر هندسی مسئله کمترین توان های دوم

فرض کنید A یک ماتریس $m \times n$ با $m > n$ باشد. آن گاه A یک نگاشت خطی $R^n \mapsto R^m$ است. $R(A)$ یک زیر فضای R^m است.

$$R(A) = \{Ax \mid x \in R^n\}$$

$$= \{\alpha_1 x_1 + \dots + \alpha_n x_n, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n, A = (\alpha_1, \dots, \alpha_n)\}$$

هر بردار $u \in R(A)$ می‌تواند به صورت $u = Ax$ به ازای برداری مانند $x \in R^n$ نوشت. فرض کنید $b \in R^m$. چون $\|\cdot\|_2$ یک نرم اقلیدسی است، $\|b - Ax\|_2$ فاصله بین نقاط انتهایی b و Ax باشد. واضح است که این فاصله مینیمم است اگر و فقط اگر $b - Ax$ بر $R(A)$ عمود باشد. در این حالت $\|b - Ax\|_2$ فاصله از نقطه انتهایی b تا صفحه $R(A)$ است.



از این تعبیر به آسانی می‌فهمیم که یک جواب مسئله کمترین توان های دوم برای دستگاه خطی $Ax = b$ همواره وجود دارد. این جواب وجود دارد زیرا می‌توانیم b را برای صفحه $R(A)$ تصویر کنیم و یک بردار u در $R(A)$ بدست آوریم و یک بردار x در R^n وجود دارد به‌طوری که $u = Ax$ این x یک جواب است. چون $b - Ax$ بر $R(A)$ عمود است و هر

بردار بر $R(A)$ یک ترکیب خطی از بردارهای ستونی A است. آنگاه $b - Ax$ بر هر بردار ستونی A عمود است یعنی

$$A^T(b - Ax) = 0 \Rightarrow A^T b = A^T Ax$$

دستگاه معادلات $A^T Ax = A^T b$ را معادلات نرمال می‌نامیم.

یک مثال آشنا از چگونه مطرح شدن معادلات نرمال برازش یک چندجمله‌ای به یک مجموعه از داده‌های تجربی است. مهندسان و دانشمندان از آزمایش‌ها، داده‌ها را جمع‌آوری می‌کنند و علاقمند هستند نمایش با معنایی از داده‌های جمع‌آوری شده بدست آورند.

فرض کنید $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ یک مجموعه از جفت مشاهدات باشند. فرض کنید توسط یک چندجمله‌ای از درجه m که $m \leq n$ باشد، می‌خواهیم بهترین برازش برای این مجموعه داده را بدست آوریم:

$$y(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$$

یک استراتژی برای بهترین برازش مینیم کردن مجموع توان‌های دوم مانده‌هاست یعنی

$$E = \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_mx_i^m)^2$$

می‌خواهیم $a_0, a_1, \dots, a_m$ را چنان بدست آوریم که مجموع بالا کمترین مقدار ممکن باشد پس باید داشته باشیم:

$$\frac{\partial E}{\partial a_i} = 0, \quad i = 0, \dots, m$$

$$\frac{\partial E}{\partial a_0} = 0 \Rightarrow -2 \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_mx_i^m) = 0$$

$$\frac{\partial E}{\partial a_1} = 0 \Rightarrow -2 \sum_{i=1}^n x_i (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_mx_i^m) = 0$$

⋮

$$\frac{\partial E}{\partial a_m} = 0 \Rightarrow -2 \sum_{i=1}^n x_i^m (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_mx_i^m) = 0$$

بنابراین:

$$na_0 + (\sum x_i)a_1 + (\sum x_i^2)a_2 + \dots + a_m(\sum x_i^m) = \sum y_i$$

$$(\sum x_i)a_0 + (\sum x_i^2)a_1 + (\sum x_i^3)a_2 + \dots + a_m(\sum x_i^{m+1}) = \sum x_i y_i$$

⋮

$$(\sum x_i^m)a_0 + (\sum x_i^{m+1})a_1 + (\sum x_i^{m+2})a_2 + \dots + a_m(\sum x_i^{2m}) = \sum x_i^m y_i$$

با قرار دادن $S_k = \sum x_i^k$, $k = 0, 1, \dots, 2m$ و نمایش عناصر سمت راست با $b_0, b_1, \dots, b_m$ دستگاه زیر نتیجه می‌شود:

$$\begin{pmatrix} S_0 & S_1 & \cdots & S_m \\ S_1 & S_2 & \cdots & S_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_m & S_{m+1} & \cdots & S_{2m} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{pmatrix}$$

این دستگاه یک دستگاه $(m+1) \times (m+1)$ می‌باشد که $m+1$ مجهول داریم. این دستگاه در واقع یک دستگاه نرمال است. در واقع داریم:

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{pmatrix}_{n \times (m+1)}$$

آنگاه دستگاه فوق همان $V^T V a = b = V^T y$ می‌شود که

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

که یک دستگاه معادلات نرمال است. بعلاوه اگر x_i ها متمایز باشند آنگاه ماتریس دارای رتبه کامل است. ماتریس $A^+ = (A^T A)^{-1} A^T$ هنگامی که A یک ماتریس $m \times n$ ($m \geq n$) بوده و دارای رتبه n می‌باشد یا ماتریس $A^+ = A^T (A A^T)^{-1}$ هنگامی که A یک ماتریس $m \times n$ ($m < n$) بوده و دارای رتبه m می‌باشد، شبه‌معکوس A نامیده می‌شود

بنابراین $x = A^+ b$ جواب مسئله کمترین توان های دوم با استفاده از شبه‌معکوس می‌باشد:

$$A^T A x = A^T b \Rightarrow x = (A^T A)^{-1} A^T b = A^+ b$$

روش معادلات نرمال در حل مسئله کمترین مربعات

یکی از متداول ترین روش‌ها برای محاسبه جواب کمترین توان‌های دوم روش معادلات نرمال است. این روش بر جواب دستگاه معادلات نرمال $A^T A x = A^T b$ مبتنی است. فرض می‌کنیم که A یک ماتریس $m \times n$ ($m < n$) با رتبه کامل باشد. چون در این حالت $A^T A$ متقارن و معین مثبت است تجزیه چولسکی را می‌پذیرد و داریم

$$A^T A = R^T R$$

بنابراین الگوریتم زیر را برای حل آن داریم:

❖ الگوریتم | جواب کمترین توان‌های دوم با استفاده از معادلات نرمال

فرض می‌کنیم A یک ماتریس $m \times n$ ($m > n$) و $\text{rank}(A) = n$ و $c = A^T b$ را تشکیل دهید.

۲. با استفاده از متقارن بودن $A^T A$ ، ماتریس R را محاسبه کنید.

۳. دستگاه‌های مثلثی زیر را به ترتیب حل کنید.

$$R^T y = c$$

$$Rx = y$$

$$A^T A x = A^T b$$

$$A^T A x = \underbrace{A^T b}_c \Rightarrow \underbrace{A^T A}_{R^T R} x = c \Rightarrow \begin{cases} R^T R x = c \\ R^T y = c \Rightarrow y \\ R x = y \Rightarrow x \end{cases}$$

💡 نکته 2.8

دقت کنیم همواره برای حل هر مسئله کمترین مربعات می‌توانی از معادلات نرمال استفاده کنیم و دستگاه $A^T A x = A^T b$ را تشکیل دهیم و سپس برای حل این دستگاه از هر روشی می‌توان استفاده کرد (دقت کنیم در بالا به تجزیه چولسکی اشاره کردیم ولی روش‌های دیگر را نیز می‌توان به کار گرفت).

دستگاه معادلات مثال 2.9 را به روش معادلات نرمال حل کنید.

* مثال 2.10

$$A_n = b \Rightarrow \begin{pmatrix} -3 & 1 \\ 1 & 1 \\ -7 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} 70 \\ 21 \\ 110 \\ -35 \end{pmatrix}$$

$$\Rightarrow A^T A x = A^T b \Rightarrow \begin{pmatrix} -3 & 1 & -7 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 1 & 1 \\ 7 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} -3 & 1 & -7 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 70 \\ 21 \\ 110 \\ -35 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 84 & -4 \\ -4 & 4 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} -1134 \\ 166 \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix} = \begin{pmatrix} -12.1 \\ 29.4 \end{pmatrix} \Rightarrow y(x) = -12.1x + 29.4$$

حال می‌توان در صورت نیاز خطای تقریبی را نیز محاسبه کرد:

$$\epsilon = b - Ax \Rightarrow \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix} = \begin{pmatrix} 70 \\ 21 \\ 110 \\ -35 \end{pmatrix} - \begin{pmatrix} -3 & 1 \\ 1 & 1 \\ -7 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix}$$

با در نظر گرفتن $\hat{x} = \begin{pmatrix} \hat{m} \\ \hat{n} \end{pmatrix} = \begin{pmatrix} -12.1 \\ 29.4 \end{pmatrix}$ داریم

$$\epsilon = \begin{pmatrix} 4.3 \\ 3.7 \\ -4.1 \\ -3.9 \end{pmatrix}$$

و مقدار نرم را می‌توان با نرم ۲ محاسبه کرد:

$$\|\epsilon\|^2 = (4.3)^2 + (3.7)^2 + (4.1)^2 + (3.9)^2 = 64.2 \Rightarrow \|\epsilon\| = \sqrt{64.2} = 8.0125$$

در واقع هر مقدار دیگری برای m و n انتخاب می‌کردیم خطا بزرگتر از ۱۲۵.۸ می‌شد.

برای دستگاه زیر مسئله کمترین مربعات را با استفاده از تجزیه چولسکی حل کنید.

* مثال 2.11

$$Ax = b \Rightarrow \frac{1}{5} \begin{pmatrix} 3 & -6 & 26 \\ 4 & -8 & -7 \\ 0 & 4 & 4 \\ 0 & -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

سیستم ناسازگار است $\Rightarrow rank(A) = 3 \text{ \& } rank(A, b) = 4$

$$\underbrace{A^T A}_C = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 5 & -3 \\ 2 & -3 & 30 \end{pmatrix}, \underbrace{A^T b}_d = \begin{pmatrix} \frac{7}{5} \\ \frac{-13}{5} \\ 4 \end{pmatrix}$$

سپس معادلات نرمال را حل می‌کنیم

$$A^T A x = A^T b \Rightarrow C x = d \Rightarrow \begin{pmatrix} 1 & -2 & 2 \\ -2 & 5 & -3 \\ 2 & -3 & 30 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{5} \\ \frac{-13}{5} \\ 4 \end{pmatrix}$$

سپس تجزیه چولسکی برای C را بدست می‌آوریم:

$$C = R^T R = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

حال باید دو دستگاه معادله را حل کنیم

$$R^T \underbrace{R x}_y = d \Rightarrow R^T y = d, \quad R x = y$$

$$R^T y = d \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & 1 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{5} \\ \frac{-13}{5} \\ 4 \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{pmatrix}$$

$$R x = y \Rightarrow \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{4}{25} \\ \frac{4}{25} \\ \frac{1}{25} \end{pmatrix}$$

2.10.3 استفاده از تجزیه QR برای حل مسئله کمترین مربعات

همانگونه که مشاهده کردیم برای حل مسئله کمترین مربعات $\min_{x \in \mathbb{R}^n} \|Ax - b\|$ لازم است جواب دستگاه حاصل از معادلات نرمال را بدست آوریم یعنی باید $A^T A x = A^T b$ را حل کنیم. یک روش برای حل این دستگاه استفاده از تجزیه QR می‌باشد که در آن Q یک ماتریس متعامد و R یک ماتریس بالامثلثی می‌باشد یعنی داریم:

$$A^T A = Q R$$

$$Q R x = A^T b = c$$

$$Q R x = c$$

$$\underbrace{Q^T Q}_I R x = Q^T c$$

$$R x = Q^T c \Rightarrow x = ?$$

حال سوال این است که تجزیه QR یک ماتریس را چگونه بدست آوریم که در ادامه به این امر می‌پردازیم.

$$A_{m \times n} x = b, A = Q_{m \times m} R_{m \times n}$$

2.11 تجزیه QR

تعریف 2.5 | ماتریس‌های هوسهلدر

فرض کنیم $u \in \mathbb{R}^n$ برداری غیر صفر باشد، ماتریس تعریف شده به صورت

$$H = I_{n \times n} - \frac{2u_{n \times 1} u_{1 \times n}^T}{u_{1 \times n}^T u_{n \times 1}}$$

را یک ماتریس هوسهلدر متناظر با u گوئیم.

$$H = I - \left(\frac{2}{u^T u} \right) u u^T = I - \beta u u^T, \quad \beta \in \mathbb{R}$$

بدون کاستن از کلیت می‌توان $u^T u = 1$ در نظر گرفت.

$$u \in \mathbb{R}^n$$

$$\frac{u}{\|u\|} \in \mathbb{R}^n \rightarrow H$$

قضیه 2.9

فرض کنیم $H = I - \frac{2uu^T}{u^T u}$ که $u \in \mathbb{R}^n$ و $u \neq 0$

۱- ماتریس H متقارن است.

۲- ماتریس H متعامد است.

۳-

$$H^2 = I$$

۴-

$$Hu = -u$$

۵- اگر v برداری باشد که $v^T u = 0$ آنگاه $Hv = v$

۶-

$$\det(H) = -1$$

۷- اگر u موازی $x - y$ باشد و $\|x\|_2 = \|y\|_2$ و $x \neq y$ آنگاه $Hx = y$

۱ | اثبات ✓

$$\begin{aligned}
 H^T &= \left(I - \frac{2}{u^T u} u u^T \right)^T \\
 &= I - \frac{2}{u^T u} (u^T)^T u^T \\
 &= I - \frac{2}{u^T u} u u^T = H \quad \blacksquare
 \end{aligned}$$

۲ | اثبات ✓

$$\begin{aligned}
 H^T H &= \left(I - \frac{2}{u^T u} u u^T \right) \left(I - \frac{2}{u^T u} u u^T \right) \\
 &= I - \frac{4}{u^T u} u u^T + \left(\frac{2}{u^T u} \right)^2 u u^T u u^T \\
 &= I - \frac{4}{u^T u} u u^T + \frac{4}{u^T u} u u^T = I \quad \blacksquare
 \end{aligned}$$

۳ | اثبات ✓

$$\begin{aligned}
 H^2 &= H H \stackrel{(1)}{=} H^T H \stackrel{(2)}{=} I \quad \blacksquare \\
 H^{-1} &= H = H^T
 \end{aligned}$$

۴ | اثبات ✓

$$\begin{aligned}
 H u &= \left(I - \frac{2}{u^T u} u u^T \right) u \\
 &= u - \frac{2}{u^T u} u (u^T u) = -u \\
 H u &= -u \quad \blacksquare
 \end{aligned}$$

از این ویژگی نتیجه می‌گیریم $\lambda = -1$ یک مقدار ویژه برای H است و بردار ویژه متناظر با آن $u \neq 0$ است.

۵ | اثبات ✓

فرض کنیم $v \in \mathbb{R}^n$ برداری باشد که $v^T u = u^T v = 0$. آنگاه داریم:

$$\begin{aligned}
 H v &= \left(I - \frac{2 u u^T}{u^T u} \right) v \\
 &= v - \frac{2}{u^T u} u \underbrace{(u^T v)}_0 \\
 &\Rightarrow H v = v \quad \blacksquare
 \end{aligned}$$

اثبات | ۶

از ویژگی ۵ نتیجه می‌گیریم $\lambda = 1$ یک مقدار ویژه برای H است و هر برداری که بر u عمود باشد (یعنی هر v که $v^T u = 0$ باشد) یک بردار ویژه متناظر با مقدار ویژه ۱ است. از طرفی می‌دانیم هر بردار در صفحه p بر u عمود است و صفحه p در فضای $\mathbb{R}^n$ یک زیرفضا از بعد $n - 1$ می‌باشد. پس یک پایه به فرم $\{v_1, v_2, \dots, v_{n-1}\}$ برای p داریم و v_i ها مستقل خطی هستند و همگی بردارهای ویژه برای $\lambda = 1$ هستند:

$$Hv_i = v_i, \quad i = 1, 2, \dots, n-1, \quad v_i \text{ مستقل خطی}$$

و از اینجا نتیجه می‌گیریم $\lambda = 1$ یک مقدار ویژه H با مرتبه تکرار $n - 1$ می‌باشد. از طرف دیگر در ویژگی ۴ نشان دادیم $\lambda = -1$ یک مقدار ویژه H است، پس مقادیر ویژه H برابر $\{1, -1\}$ هستند و نتیجه می‌گیریم $\det H = -1$. ■

اثبات | ۷

بردار x را می‌توان به صورت زیر نوشت:

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

و بردار u را موازی بردار $x - y$ انتخاب می‌کنیم، یعنی:

$$\exists \alpha > 0 : u = \alpha(x - y)$$

ابتدا نشان می‌دهیم $(x - y) \perp (x + y)$

$$\begin{aligned} (x+y)^T(x-y) &= (x^T + y^T)(x-y) \\ &= x^T x - x^T y + y^T x - y^T y \\ &= \|x\|^2 - \|y\|^2 \\ &= 0 \end{aligned}$$

حال Hx را بدست می‌آوریم:

$$\begin{aligned} Hx &= H \left(\frac{1}{2}(x+y) + \frac{1}{2}(x-y) \right) \\ &= \frac{1}{2} \underbrace{H(x+y)}_{\perp u} + \frac{1}{2} \underbrace{H(x-y)}_{\parallel u} \\ &\stackrel{(5)}{=} \frac{1}{2}(x+y) - \frac{1}{2}(x-y) = y \quad \blacksquare \end{aligned}$$

تجزیه QR ماتریس بر اساس ماتریس‌های هوسهلدر

فرض کنیم A یک ماتریس $n \times n$ باشد. می‌خواهیم ماتریس بالامثلثی R و ماتریس متعامد Q را چنان بدست آوریم که:

$$A = QR$$

قضیه 2.10

اگر بردار x چنان باشد که $x \neq e_1$ ، آنگاه همواره ماتریس هوسهولدر H وجود دارد به طوری که:

$$Hx = ke_1$$

که در آن k عددی ثابت است.

اثبات ✓

بردار u را به صورت زیر تعریف می‌کنیم:

$$u = x \pm \|x\|e_1$$

و ماتریس H را بر اساس این u تعریف می‌کنیم.

$$1. u = x - \underbrace{\|x\|e_1}_y = x - y$$

$$\Rightarrow Hx = y = \underbrace{\|x\|}_k e_1$$

$$2. u = x + \|x\|e_1 = x - (-\|x\|e_1) = x - \underbrace{(-\|x\|e_1)}_y$$

$$\Rightarrow Hx = y = \underbrace{-\|x\|}_k e_1$$

$$Hx = \mp \|x\|e_1, \quad k = \mp \|x\|$$

تعریف 2.6 | تابع $\text{sign}(z)$

$$\text{sign}(z) = \begin{cases} + & z \geq 0 \\ - & z < 0 \end{cases}$$

2.1 قرارداد

همانگونه که در اثبات فوق مشاهده کردیم، به ازای هر x از دو طریق می‌توان بردار u را تعریف کرد. برای اینکه در ادامه به صورت یکسان عمل کنیم، قرارداد می‌کنیم که u را به صورت زیر تعریف کنیم:

$$u = x + \text{sign}(x_1)\|x\|e_1$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow Hx = -\text{sign}(x_1)\|x\|e_1$$

* مثال 2.12 فرض کنید $x = \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$. ماتریس هوسهولدری بسازید که $Hx = ke_1$.

$$\|x\| = \sqrt{0^2 + 3^2 + 4^2} = 5$$

$$u = x + \text{sign}(x_1)\|x\|e_1$$

$$= \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} + 5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$$

$$H = I - \frac{2uu^T}{u^T u}$$

$$u^T u = \begin{pmatrix} 5 & 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} = 50$$

$$\begin{aligned}
 H &= I - \frac{2}{50} \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} 5 & 3 & 4 \end{pmatrix} \\
 H &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{25} \begin{pmatrix} 25 & 15 & 20 \\ 15 & 9 & 12 \\ 20 & 12 & 16 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & \frac{3}{5} & \frac{4}{5} \\ \frac{3}{5} & \frac{9}{25} & \frac{12}{25} \\ \frac{4}{5} & \frac{12}{25} & \frac{16}{25} \end{pmatrix} \\
 &= \begin{pmatrix} 0 & -\frac{3}{5} & -\frac{4}{5} \\ -\frac{3}{5} & \frac{16}{25} & -\frac{12}{25} \\ -\frac{4}{5} & -\frac{12}{25} & \frac{9}{25} \end{pmatrix} \\
 Hx &= \begin{pmatrix} 0 & -\frac{3}{5} & -\frac{4}{5} \\ -\frac{3}{5} & \frac{16}{25} & -\frac{12}{25} \\ -\frac{4}{5} & -\frac{12}{25} & \frac{9}{25} \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} \\
 &= \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}
 \end{aligned}$$

قضیه 2.11

فرض کنیم A یک ماتریس $n \times n$ باشد. همواره یک ماتریس بالامثلثی R و یک ماتریس متعامد Q وجود دارد به طوری که

$$A = QR$$

در طی اثبات نحوه ساختن ماتریس‌های Q و R را مشخص می‌کنیم.

فرض کنیم

اثبات ✓

$$A^{(0)} = A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

بردار ستون اول یعنی x_1 را در نظر می‌گیریم و تعریف می‌کنیم:

$$u_1 = x_1 + \text{sign}(a_{11}) \|x_1\| e_1$$

و بر اساس u_1 ماتریس هوسهولدر H_1 را می‌سازیم.

$$A^{(1)} = H_1 A^{(0)} = H_1 A = \begin{pmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{pmatrix}$$

درایه‌های ماتریس $A^{(1)}$ را به صورت زیر نامگذاری می‌کنیم

$$A^{(1)} = H_1 A^{(0)} = H_1 A$$

$$= \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

در مرحله دوم قرار می‌دهیم

$$x_2 = \begin{pmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{pmatrix}, \quad u_2 = x_2 + \text{sign}(a_{22}^{(1)}) \|x_2\| e_1$$

و ماتریس هوسهولدر $\hat{H}_2$ که یک ماتریس $(n-1) \times (n-1)$ می‌باشد را بر اساس u_2 می‌سازیم و تعریف می‌کنیم:

$$H_2 = \begin{pmatrix} 1 & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & \hat{H}_2 \end{pmatrix}_{n \times n}$$

و قرار می‌دهیم

$$A^{(2)} = H_2 A^{(1)} = H_2 H_1 A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(2)} \end{pmatrix}$$

به همین ترتیب ادامه می‌دهیم و در مرحله k داریم

$$x_k = \begin{pmatrix} a_{kk}^{(k-1)} \\ a_{(k+1)k}^{(k-1)} \\ \vdots \\ a_{nk}^{(k-1)} \end{pmatrix} \quad u_k = x_k + \text{Sign}(a_{kk}^{(k-1)}) \|x_k\| e_1$$

و ماتریس هوسهولدر $\hat{H}_k$ که یک ماتریس $(n-k+1) \times (n-k+1)$ بعدی هست را از روی u_k می‌سازیم و تعریف می‌کنیم

$$H_k = \begin{pmatrix} I_{(k-1)(k-1)} & 0 \\ 0 & \hat{H}_k \end{pmatrix}_{n \times n}$$

و قرار می‌دهیم

$$A^{(k)} = H_k H_{k-1} \cdots H_2 H_1 A$$

و این مراحل را تا مرحله $n-1$ ادامه می‌دهیم و داریم

$$A^{(n-1)} = H_{n-1} H_{n-2} \cdots H_2 H_1 A = H_{n-1} A^{(n-2)} = \begin{pmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & * & * \\ 0 & 0 & \cdots & 0 & * \end{pmatrix} = R$$

ماتریس‌های H_i همگی متعامد هستند و حاصلضرب ماتریس‌های متعامد، متعامد است (چرا؟) پس $(H_{n-1} H_{n-2} \cdots H_2 H_1)$ ماتریسی متعامد است.

$$Q^T = H_{n-1} H_{n-2} \cdots H_2 H_1$$

$$Q = H_1 H_2 \cdots H_{n-1}$$

و داریم:

$$Q^T A = R$$

$$QQ^T A = QR$$

$$\Rightarrow A = QR \quad \blacksquare$$

برای ماتریس زیر تجزیه QR را به روش ماتریس‌های هوسهولدر بدست آورید.

* مثال 2.13

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

💡 حل:

$$A = A^{(0)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \quad x_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$u_1 = x_1 + \text{sign}(a_{11}^{(0)}) \|x_1\| e_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix}$$

$$u_1^T u_1 = \begin{pmatrix} \sqrt{2} & 1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix} = 4$$

$$H_1 = I - \frac{2u_1 u_1^T}{4} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix} (\sqrt{2}, 1, 1)$$

$$\begin{aligned}
&= \begin{pmatrix} 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
A^{(1)} &= H_1 A = \begin{pmatrix} 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -\sqrt{2} & -\frac{3\sqrt{2}}{2} & -2\sqrt{2} \\ 0 & \frac{1-\sqrt{2}}{2} & \frac{2-\sqrt{2}}{2} \\ 0 & \frac{-1-\sqrt{2}}{2} & \frac{-2-\sqrt{2}}{2} \end{pmatrix} \approx \begin{pmatrix} -1.414 & -2.1213 & -2.8284 \\ 0 & -0.2071 & 0.2928 \\ 0 & -1.2071 & -1.7071 \end{pmatrix} \\
x_2 &= \begin{pmatrix} -0.2071 \\ -1.2071 \end{pmatrix} \\
u_2 = x_2 - \|x_2\|e_1 &= \begin{pmatrix} -0.2071 \\ -1.2071 \end{pmatrix} - \begin{pmatrix} 1.2247 \\ 0 \end{pmatrix} = \begin{pmatrix} -1.4318 \\ -1.2071 \end{pmatrix} \\
u_2^T u_2 &= 3.5071 \\
\hat{H}_2 &= I - \frac{2u_2 u_2^T}{3.5071} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - (0.5702) \begin{pmatrix} -1.4318 \\ -1.2071 \end{pmatrix} \begin{pmatrix} -1.4318 & -1.2071 \end{pmatrix} \\
&= \begin{pmatrix} -0.1689 & -0.9854 \\ -0.9854 & 0.1692 \end{pmatrix} \\
H_2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \hat{H}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.1689 & -0.9854 \\ 0 & -0.9854 & 0.1692 \end{pmatrix} \\
A^{(2)} &= H_2 A^{(1)} = \begin{pmatrix} -1.4142 & -2.1213 & -2.8284 \\ 0 & 1.2274 & 1.6330 \\ 0 & 0 & -0.5774 \end{pmatrix} = R
\end{aligned}$$

$$Q = H_1 H_2 = \begin{pmatrix} 0 & -0.7071 & -0.7071 \\ -0.7071 & 0.5 & -0.5 \\ -0.7071 & -0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.1689 & -0.9854 \\ 0 & -0.9854 & 0.1692 \end{pmatrix}$$

تجزیه QR برای ماتریس‌های غیر مربعی

فرض کنیم A یک ماتریس $m \times n$ باشد به طوری که $m \neq n$ آنگاه برای این ماتریس به صورت مشابه می‌توانیم تجزیه QR را با استفاده از ماتریس‌های هوسهولدر بدست آوریم. لازم است مشابه روشی که برای ماتریس‌های مربعی انجام دادیم در اینجا دنبال کنیم و s تکرار انجام دهیم که

$$s = \min\{n, m - 1\}$$

ماتریس‌های H_i همگی $m \times m$ و ماتریس R ، $m \times n$ و به شکل زیر است.

$$\begin{bmatrix} * & & & \\ & * & & \\ 0 & & & \\ & 0 & \cdots & 0 \\ & & \ddots & \\ & 0 & \cdots & 0 \end{bmatrix}$$

$$m > n$$

$$\begin{bmatrix} * & & \\ & * & \\ 0 & & \end{bmatrix}$$

$$m = n$$

$$\begin{bmatrix} * & & \\ & * & \\ 0 & & \end{bmatrix}$$

$$m < n$$

تجزیه QR ماتریس زیر را بدست آورید.

*** مثال 2.14**

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

2.12 تجزیه مقادیر تکین (SVD)

تعریف 2.7

فرض کنیم A یک ماتریس $m \times n$ باشد، ماتریس $A^T A$ یک ماتریس $n \times n$ می‌باشد که متقارن و نیمه معین مثبت می‌باشد. و فرض کنیم $\lambda_1, \lambda_2, \dots, \lambda_n$ مقادیر ویژه $A^T A$ باشند. آنگاه مقادیر تکین A عبارت‌اند از:

$$\sigma_i = \sqrt{\lambda_i} \quad \text{for } i = 1, \dots, n$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

مقادیر تکین ماتریس زیر را حساب کنید.

* مثال 2.15

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 1 \end{pmatrix}$$

حل: 

$$A^T A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\det(\lambda I - A) = \det \left(\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} - \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} \right)$$

$$= \det \begin{pmatrix} \lambda - 5 & -2 \\ -2 & \lambda - 2 \end{pmatrix}$$

$$= (\lambda - 5)(\lambda - 2) - 4 = 0$$

$$= \lambda^2 - 7\lambda + 6 = 0$$

$$\Rightarrow \lambda_1 = 6, \quad \lambda_2 = 1$$

$$\Rightarrow \sigma_1 = \sqrt{6}, \quad \sigma_2 = 1$$

❏ قضیه 2.12

فرض کنیم A یک ماتریس $m \times n$ باشد. آنگاه ماتریس‌های متعامد $U_{m \times m}$ و $V_{n \times n}$ و ماتریس $\Sigma_{m \times n}$ وجود دارند به طوری که

$$A = U \Sigma V^T$$

و ماتریس Σ به صورت زیر است:

$$D = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix}_{r \times r}, \quad \Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{m \times n}$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n$$

و σ_i ها مقادیر تکین ماتریس A هستند.

همچنین فرض می‌کنیم $\bar{v}_i$ ها بردارهای ویژه ماتریس $A^T A$ متناظر با مقادیر ویژه λ_i باشند. برای محاسبه V ستون‌های آن را (که همان بردارهای ویژه نرمال‌سازی شده هستند) به صورت زیر بدست می‌آوریم:

$$V = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}, \quad v_i = \frac{\bar{v}_i}{\|\bar{v}_i\|}$$

و ستون‌های U (متناظر با مقادیر تکین غیرصفر) به صورت زیر بدست می‌آیند:

$$u_i = \frac{1}{\sigma_i} A v_i, \quad i = 1, \dots, r$$

تجزیه مقادیر تکین (تجزیه SVD) ماتریس زیر را بدست آورید.

* مثال 2.16

$$A = \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix}$$

❏ حل:

$$A^T A = \begin{pmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix} = \begin{pmatrix} 40 & 34 & 14 \\ 34 & 37 & 20 \\ 14 & 20 & 13 \end{pmatrix}$$

$$\det(I - A) = \det \begin{pmatrix} \lambda - 40 & -34 & -14 \\ -34 & \lambda - 37 & -20 \\ -14 & -20 & \lambda - 13 \end{pmatrix}$$

$$= (\lambda - 40)((\lambda - 37)(\lambda - 13) - 400) + 34(-34(\lambda - 13) - 280) - 14(680 + 14(\lambda - 37))$$

$$= \lambda(\lambda - 81)(\lambda - 9)$$

$$\lambda_1 = 81 \Rightarrow \sigma_1 = 9$$

$$\lambda_2 = 9 \Rightarrow \sigma_2 = 3$$

$$\lambda_3 = 0 \Rightarrow \sigma_3 = 0$$

$$\Sigma = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

$$m = 2, n = 3, r = 2$$

• محاسبه V :

$$\bar{v}_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \bar{v}_2 = \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \quad \bar{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

$$\|\bar{v}_1\| = \|\bar{v}_2\| = \|\bar{v}_3\| = 3$$

$$V = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

• محاسبه U :

$$u_1 = \frac{1}{9} \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix} \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$u_2 = \frac{1}{3} \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix} \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$$U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

◀ 2.12.1 کاربرد تجزیه SVD در حل مسئله کمترین مربعات

فرض کنیم دستگاه معادلات خطی به فرم زیر را داریم:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

$$A_{m \times n} x_{n \times 1} = b_{m \times 1}$$

فرض می‌کنیم $m > n$ باشد. در این صورت معمولاً دستگاه فوق جواب ندارد و لازم است جواب کمترین مربعات آن را بدست آوریم، یعنی $x^* \in \mathbb{R}^n$ را محاسبه کنیم به طوری که

$$\|Ax^* - b\| = \min_{x \in \mathbb{R}^n} \|Ax - b\|$$

💡 نکته 2.9

ماتریس‌های متعامد طول را حفظ می‌کنند.

$$\|x\| = \|Ux\|$$

✅ اثبات

$$\|Ux\|_2^2 = (Ux)^T(Ux) = x^T \underbrace{U^T U}_I x = \|x\|_2^2$$

$$\Rightarrow \|x\| = \|Ux\|$$

تجزیه SVD ماتریس A را بدست می‌آوریم:

$$A = U\Sigma V^T$$

$$\begin{aligned} \|Ax - b\| &= \|U\Sigma V^T x - b\| \\ &= \|U^T(U\Sigma V^T x - b)\| \\ &= \|\underbrace{\Sigma V^T x}_z - \underbrace{U^T b}_d\| \\ &= \left\| \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} - \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix} \right\| \end{aligned}$$

دو حالت ممکن است:

حالت اول

فرض کنیم $r = n$ (یعنی تعداد مقادیر ویژه غیرصفر $A^T A$ ، n باشد).

$$\begin{aligned} \Sigma z - d &= \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} - \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \\ \vdots \\ d_m \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1 z_1 - d_1 \\ \sigma_2 z_2 - d_2 \\ \vdots \\ \sigma_n z_n - d_n \\ -d_{n+1} \\ \vdots \\ -d_m \end{pmatrix} \end{aligned}$$

$$\|Ax - b\|^2 = \|\Sigma z - d\|^2 = \sum_{i=1}^n (\sigma_i z_i - d_i)^2 + \sum_{i=n+1}^m d_i^2$$

برای مینیمم‌سازی $\|Ax - b\|^2$ ، با توجه به اینکه d_i ها مقادیر ثابت و معلومی هستند، کافی است مجموع $\sum_{i=1}^n (\sigma_i z_i - d_i)^2$ را مینیمم کنیم. از آنجا که این مجموع، جمع مربعات جملات مستقل است، برای مینیمم‌سازی آن، هر یک از جملات $(\sigma_i z_i - d_i)^2$ باید به طور جداگانه مینیمم شود. بنابراین:

$$\sigma_i z_i - d_i = 0 \quad \Rightarrow \quad z_i = \frac{d_i}{\sigma_i}, \quad \forall i = 1, \dots, n, \quad \sigma_i \neq 0$$

با تعیین تمام z_i ها، بردار z کاملاً مشخص می‌شود. در نتیجه، از رابطه $V^T x = z$ ، بردار x نیز به طور یکتا تعیین می‌گردد.

حالت دوم

فرض کنیم $r < n$ (یعنی تعداد مقادیر ویژه غیرصفر $A^T A$ ، کمتر از n باشد).

$$\Sigma z - d = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \\ 0 & \dots & 0 & \dots & 0 & \end{pmatrix}_{m \times n} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} - \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \\ \vdots \\ d_m \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 z_1 - d_1 \\ \sigma_2 z_2 - d_2 \\ \vdots \\ \sigma_r z_r - d_r \\ -d_{r+1} \\ \vdots \\ -d_n \\ \vdots \\ -d_m \end{pmatrix}$$

$$\|Ax - b\|^2 = \|\Sigma z - d\|^2 = \sum_{i=1}^r (\sigma_i z_i - d_i)^2 + \sum_{i=r+1}^m d_i^2$$

برای مینیمم سازی $\|Ax - b\|^2$ ، با توجه به اینکه d_i ها مقادیر ثابت و معلومی هستند، کافی است مجموع $\sum_{i=1}^r (\sigma_i z_i - d_i)^2$ را مینیمم کنیم. از آنجا که این مجموع، جمع مربعات جملات مستقل است، برای مینیمم سازی آن، هر یک از جملات $(\sigma_i z_i - d_i)^2$ باید به طور جداگانه مینیمم شود.
بنابراین:

$$\sigma_i z_i - d_i = 0 \Rightarrow z_i = \frac{d_i}{\sigma_i}, \quad \forall i = 1, \dots, r, \quad \sigma_i \neq 0$$

در اینجا کلیه اجزای بردار z بدست نمی آید، اما مقادیر تعیین نشده (z_n تا z_{r+1}) بر مقدار کمترین مربعات تأثیری ندارند (طبق فرمول بالاتر) و می توان مقادیر دلخواهی برای آنها در نظر گرفت. سپس از رابطه $V^T x = z$ بردار x را بدست آورد.

*** مثال 2.17** جواب مسئله کمترین مربعات زیر را با استفاده از تجزیه SVD بدست آورید و مقدار خطا (باقی مانده) را تعیین کنید.

$$-2x_1 + 6x_2 = 3$$

$$x_1 + 6x_2 = 0$$

$$2x_1 + 3x_2 = -3$$

💡 حل:

$$\begin{pmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix}, \quad m = 3, n = 2$$

$$A = U\Sigma V^T$$

$$A^T A = \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix} \begin{pmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 0 \\ 0 & 81 \end{pmatrix}$$

$$\lambda_1 = 81, \quad \sigma_1 = 9$$

$$\lambda_2 = 9, \quad \sigma_2 = 3$$

$$\bar{v}_1 = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

$$A^T A \bar{v}_1 = 81 \bar{v}_1 = \begin{pmatrix} 9 & 0 \\ 0 & 81 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = 81 \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

$$\omega_1 = 0 \quad , \quad \omega_2 = 1 \quad \Rightarrow \quad \bar{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = v_1$$

$$A^T A \bar{v}_2 = \lambda_2 \bar{v}_2 = 9 \bar{v}_2$$

$$\bar{v}_2 = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

$$\begin{pmatrix} 9 & 0 \\ 0 & 81 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = 9 \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

$$\Rightarrow \begin{cases} 9\omega_1 = 9\omega_1 \\ 81\omega_2 = 9\omega_2 \Rightarrow 72\omega_2 = 0 \Rightarrow \omega_2 = 0 \end{cases}$$

$$\omega_1 = 1 \Rightarrow \bar{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = v_2$$

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 9 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

برای محاسبه ماتریس U باید بردارهای ویژه AA^T را محاسبه کنیم.

$$AA^T = \begin{pmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{pmatrix} = \begin{pmatrix} 40 & 34 & 14 \\ 34 & 37 & 20 \\ 14 & 20 & 13 \end{pmatrix}$$

$$\lambda_1 = 81, \quad \lambda_2 = 9, \quad \lambda_3 = 0$$

$$\bar{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \bar{u}_2 = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \quad \bar{u}_3 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

$$U = \begin{pmatrix} \frac{\bar{u}_1}{\|\bar{u}_1\|} & \frac{\bar{u}_2}{\|\bar{u}_2\|} & \frac{\bar{u}_3}{\|\bar{u}_3\|} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

$$d = U^T b$$

$$z_i = \frac{d_i}{\sigma_i}, \quad i \in 1, 2$$

$$U^T x = z \Rightarrow x = ?$$

$$d = U^T b$$

$$= \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ -1 \end{pmatrix}$$

$$z_1 = \frac{d_1}{\sigma_1} = \frac{1}{9}$$

$$z_2 = \frac{d_2}{\sigma_2} = \frac{-4}{3}$$

$$V^T x = z \Rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{9} \\ -\frac{4}{3} \end{pmatrix}$$

$$x_2 = \frac{1}{9}$$

$$x_1 = -\frac{4}{3}$$

$$x^* = \begin{pmatrix} -\frac{4}{3} \\ \frac{1}{9} \end{pmatrix}$$

و مقدار خطا:

$$\|Ax^* - b\| = \left(\sum_{i=3}^3 d_i^2\right)^{0.5} = (-1^2)^{0.5} = 1$$

2.13 $\equiv$ فرآیند متعامدسازی گرام اشمیت

فرض کنیم یک مجموعه از بردارهای مستقل خطی به صورت زیر داریم

$$\{a_1, a_2, \dots, a_n\}, \quad a_i \in \mathbb{R}^n, \quad i = 1, 2, \dots, n$$

می‌خواهیم مجموعه فوق را به یک مجموعه با مشخصات زیر تبدیل کنیم

$$\{q_1, q_2, \dots, q_n\}, \quad q_i \in \mathbb{R}^n, \quad i = 1, 2, \dots, n$$

$$\begin{aligned} \|q_i\| &= q_i^T q_i = 1, \quad i = 1, 2, \dots, n \\ \langle q_i, q_j \rangle &= q_i^T q_j = 0, \quad i, j = 1, 2, \dots, n, \quad i \neq j \end{aligned}$$

از این فرآیند استفاده کنیم و برای ماتریس A تجزیه QR بدست آوریم. فرض کنیم A یک ماتریس $n \times n$ باشد و آن را به صورت زیر تعریف می‌کنیم

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$$

$\downarrow$

$$Q = \begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix}$$

$$Q^T Q = \begin{pmatrix} q_1^T \\ \vdots \\ q_n^T \end{pmatrix} \begin{pmatrix} q_1 & \dots & q_n \end{pmatrix}$$

$$= \begin{pmatrix} q_1^T q_1 & q_1^T q_2 & \dots & q_1^T q_n \\ q_2^T q_1 & q_2^T q_2 & \dots & q_2^T q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_n^T q_1 & q_n^T q_2 & \dots & q_n^T q_n \end{pmatrix} = I$$

2.13.1 $\leftarrow$ الگوریتم فرآیند گرام اشمیت

فرض کنیم $a_1, a_2, \dots, a_n$ بردارهای مستقل خطی باشند.

❖ الگوریتم

• مرحله اول

قرار می‌دهیم

$$q_1 = \frac{a_1}{\|a_1\|} = \frac{a_1}{r_{11}}$$

• مرحله دوم

قرار می‌دهیم

$$\hat{q}_2 = a_2 - r_{12}q_1$$

می‌خواهیم $\hat{q}_2$ بر q_1 عمود باشد پس

$$0 = q_1^T \hat{q}_2 = q_1^T (a_2 - r_{12}q_1) = q_1^T a_2 - r_{12} \underbrace{q_1^T q_1}_1$$

و از اینجا مقدار r_{12} را تعیین می‌کنیم، یعنی

$$r_{12} = q_1^T a_2 = \langle q_1, a_2 \rangle$$

به منظور اینکه q_2 یک‌ه باشد، قرار می‌دهیم

$$q_2 = \frac{\hat{q}_2}{\|\hat{q}_2\|} = \frac{\hat{q}_2}{r_{22}}$$

• مرحله سوم

قرار می‌دهیم

$$\hat{q}_3 = a_3 - r_{13}q_1 - r_{23}q_2$$

$$q_1^T \hat{q}_3 = 0 \Rightarrow q_1^T a_3 - r_{13}q_1^T q_1 - r_{23}q_1^T q_2 = 0$$

$$q_2^T \hat{q}_3 = 0 \Rightarrow q_2^T a_3 - r_{13}q_2^T q_1 - r_{23}q_2^T q_2 = 0$$

$$r_{13} = q_1^T a_3$$

$$r_{23} = q_2^T a_3$$

و قرار می‌دهیم

$$q_3 = \frac{\hat{q}_3}{\|\hat{q}_3\|}$$

• مرحله k ام

$$\hat{q}_k = a_k - \sum_{i=1}^{k-1} r_{ik}q_i$$

$$r_{ik} = q_i^T a_k, \quad i = 1, 2, \dots, k-1$$

$$q_k = \frac{\hat{q}_k}{\|\hat{q}_k\|}$$

و این فرآیند را تا مرحله n ام ادامه می‌دهیم. در واقع هر مرحله یک ستون از Q و یک ستون از R ساخته می‌شوند.

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots \\ 0 & r_{22} & r_{23} & \cdots \\ 0 & 0 & r_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & \cdots \end{pmatrix}$$

$$Q = \begin{pmatrix} q_1 & q_2 & q_3 & \cdots \end{pmatrix}$$

برای ماتریس زیر با روش گرام اشمیت تجزیه QR را بدست آورید.

*** مثال 2.18**

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

تعریف 2.8 | دو ماتریس متشابه

دو ماتریس $A_{n \times n}, B_{n \times n}$ را متشابه گوئیم هرگاه ماتریس نامفرد $S_{n \times n}$ وجود داشته باشد به طوری که

$$A = S^{-1}BS$$

$$SAS^{-1} = SS^{-1}BSS^{-1} = B$$

$$SAS^{-1} = B$$

*** مثال 2.19**

$$A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix}$$

دو ماتریس A, B متشابه هستند زیرا که S را به صورت زیر تعریف کنیم

$$S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

در شرط متشابه بودن صدق می‌کند:

$$S^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$S^{-1}BS = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} = A$$

نکته 2.10

اگر A, B دو ماتریس متشابه باشند و λ یک مقدار ویژه متناظر با بردار ویژه x برای A باشد آنگاه λ یک مقدار ویژه برای B می‌باشد و بردار ویژه نظیر آن $y = Sx$ است.

اثبات

$$Ax = \lambda x$$

$$\xrightarrow{S} S Ax = S A S^{-1}(Sx) = \lambda(Sx)$$

$$\underbrace{B(Sx)}_y = \lambda \underbrace{(Sx)}_y$$

2.14 تجزیه شور (Schur)

ابتدا چند تعریف:

تعریف 2.9 | ترانهاد مزدوج

به صورت زیر تعریف می‌شود:

$$U^* = (\bar{U})^T$$

مزدوج: اگر درایه‌های U به صورت $\alpha + \beta i$ باشند، درایه‌های $\bar{U}$ به صورت $\alpha - \beta i$ تعریف می‌شوند.

* مثال 2.20

$$U = \begin{pmatrix} 1+i & 2 \\ 3i & 4-i \end{pmatrix}$$

حل:💡

$$U^* = (\overline{U})^T = \begin{pmatrix} 1-i & -3i \\ 2 & 4+i \end{pmatrix}$$

💡 نکته 2.11

اگر درایه‌های ماتریس حقیقی باشند، ترانهاده مزدوج، همان ترانهاده است.

$$A \in \mathbb{R}^{n \times n} \Rightarrow A^* = A^T$$

📖 تعریف 2.10

• ماتریس یکانی

$$U^*U = I$$

• ماتریس متعامد

$$A^T A = I$$

• ماتریس هرمیتی

$$U^* = U$$

• ماتریس متقارن

$$A^T = A$$

قضیه 2.13 | قضیه مثلثی‌سازی شور

فرض کنیم A یک ماتریس $n \times n$ باشد، آنگاه ماتریس یکانی U و ماتریس مثلثی T وجود دارد به طوری که

$$U^*AU = T$$

که روی قطر ماتریس T همان مقادیر ویژه ماتریس A قرار دارند.

$$T = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

قضیه 2.14 | قضیه مثلثی‌سازی حقیقی شور

فرض کنیم A یک ماتریس $n \times n$ با درایه‌های حقیقی باشد، آنگاه ماتریس متعامد Q و ماتریس مثلثی R وجود دارد به طوری که

$$Q^T A Q = R$$

که در آن R_{ii} یا یک اسکالر یا یک ماتریس 2×2 می‌باشد.

$$R = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ 0 & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{nn} \end{pmatrix}$$

* مثال 2.21

$$R = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 4 & -1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

$$R = \begin{pmatrix} \boxed{1} & 2 & 3 & 4 \\ 0 & \boxed{3} & 4 & -1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 0 & \boxed{3} \end{pmatrix}$$

◀ 2.14.1 روش تکرار QR

در ادامه می‌خواهیم روشی را بیان کنیم که توسط آن می‌توانیم ماتریس‌های Q, R در تجزیه مثلثی‌سازی حقیقی شور را برای ماتریس A بدست آوریم.

$$\begin{aligned} A_0 &= A & A_0 &= Q_0 R_0 \Rightarrow R_0 = Q_0^T A_0 \\ \text{define } A_1 &= R_0 Q_0 & A_1 &= Q_1 R_1 \Rightarrow R_1 = Q_1^T A_1 \\ \text{define } A_2 &= R_1 Q_1 & A_2 &= Q_2 R_2 \Rightarrow R_2 = Q_2^T A_2 \\ &\vdots \\ \text{define } A_k &= R_{k-1} Q_{k-1} & A_k &= Q_k R_k \Rightarrow R_k = Q_k^T A_k \end{aligned}$$

$$\begin{aligned} A_k &= R_{k-1} Q_{k-1} \\ &= Q_{k-1}^T A_{k-1} Q_{k-1} \\ &= Q_{k-1}^T Q_{k-2}^T A_{k-2} Q_{k-2} Q_{k-1} \\ &\vdots \\ &= (Q_{k-1}^T \cdots Q_0^T) A_0 (Q_0 \cdots Q_{k-1}) \\ &= Q^T A_0 Q = Q^T A Q \end{aligned}$$

☰ قضیه 2.15

اگر مقادیر ویژه $\lambda_1, \dots, \lambda_n$ از A به‌گونه‌ای باشند که

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

و همچنین $(X^T)^{-1}$ که X متشکل از بردارهای ویژه A است، دارای زیر ماتریس‌های پیشرو اصلی نامنفرد باشد، آنگاه دنباله $\{A_k\}$ تولید شده با روش تکرار QR به یک ماتریس شبه مثلثی مانند R همگراست.

3

روش های تکراری حل دستگاه

فرض کنیم می‌خواهیم جواب دستگاه $Ax = b$ را به دست آوریم. اگر ابعاد ماتریس A بزرگ باشد و یک ماتریس خلوت (sparse) باشد، آنگاه استفاده از روش‌های مستقیم که در فصل قبل به آن‌ها اشاره شد، بسیار پرهزینه می‌باشد. در این‌جا از روش‌های تکراری برای حل دستگاه استفاده می‌کنیم.

توسط این روش‌ها و با شروع از یک بردار دلخواه اولیه، یک دنباله $\{x^{(k)}\}$ به دست می‌آوریم. علاوه بر این، به یک شرط توقف در الگوریتم نیاز داریم. این شرط توقف برای دقت مورد نظر $\varepsilon > 0$ به صورت زیر تعریف می‌شود:

شرط توقف: اگر $\|Ax^{(k)} - b\| \leq \varepsilon$ ، متوقف می‌شویم و $x^{(k)}$ را به عنوان جواب می‌پذیریم. اگر الگوریتم به صورت متناهی متوقف نشود، یک دنباله $\{x^{(k)}\}$ به دست می‌آوریم و در صورتی که شرایط لازم برقرار باشند، نتیجه می‌گیریم این دنباله به جواب همگرا می‌باشد.

3.1 روش ژاکوبی

فرض کنیم دستگاه $Ax = b$ به صورت زیر باشد، که در آن A یک ماتریس $n \times n$ و b یک بردار $n \times 1$ باشد و مجهول ما برداری $n \times 1$ می‌باشد. فرض کنیم درایه‌های قطری، یعنی a_{ii} ها، همگی غیرصفر باشند. اگر این شرط برقرار نباشد، سطرها را چنان جابه‌جا می‌کنیم که این شرط برقرار شود. از سطر اول درایه x_1 ، از سطر دوم درایه x_2 و ... از سطر n ام درایه x_n را به دست می‌آوریم، یعنی:

$$Ax = b$$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 &\Rightarrow x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - \cdots - a_{1n}x_n) \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 &\Rightarrow x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n) \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n &\Rightarrow x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - \cdots - a_{n,n-1}x_{n-1}) \end{aligned}$$

یعنی داریم:

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \right) \quad \text{for } i = 1, \dots, n$$

بردار (نقطه) اولیه دلخواه $x^{(0)}$ را داریم.

$$x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix} \Rightarrow x^{(1)} = \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}$$

در سمت راست معادلات فوق، به جای x_i ها، مقادیر $x_i^{(0)}$ را قرار می‌دهیم و در سمت چپ به جای x_i ها، مقادیر $x_i^{(1)}$ را به دست می‌آوریم، یعنی:

$$\begin{aligned} x_1^{(1)} &= \frac{1}{a_{11}} (b_1 - a_{12}x_2^{(0)} - \dots - a_{1n}x_n^{(0)}) \\ x_2^{(1)} &= \frac{1}{a_{22}} (b_2 - a_{21}x_1^{(0)} - \dots - a_{2n}x_n^{(0)}) \\ &\vdots \\ x_n^{(1)} &= \frac{1}{a_{nn}} (b_n - a_{n1}x_1^{(0)} - \dots - a_{n,n-1}x_{n-1}^{(0)}) \end{aligned}$$

با این روش $x^{(1)}$ حاصل می‌شود و با انجام فرایند مشابه، $x^{(2)}$ ، $x^{(3)}$ و ... را به دست می‌آوریم. در واقع، فرمول تکراری زیر را داریم:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^n a_{ij}x_j^{(k)} \right), \quad j \neq i, \quad i = 1, 2, \dots, n$$

$$x^{(k+1)} = Bx^{(k)} + c$$

این روش را روش ژاکوبی می‌نامیم.

دو تکرار از روش ژاکوبی را برای دستگاه زیر انجام دهید.

*** مثال 3.1**

نقطه شروع اولیه را $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ در نظر بگیرید.

$$10x_1 - 2x_2 - x_3 = 3$$

$$-x_1 + 5x_2 - x_3 = 6$$

$$-x_1 - x_2 + 2x_3 = 3$$

💡 حل:

$$x_1^{(k+1)} = \frac{1}{10} (3 + 2x_2^{(k)} + x_3^{(k)})$$

$$x_2^{(k+1)} = \frac{1}{5} (6 + x_1^{(k)} + x_3^{(k)})$$

$$x_3^{(k+1)} = \frac{1}{2} (3 + x_1^{(k)} + x_2^{(k)})$$

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{pmatrix}$$

برای $k = 0$:

$$x_1^{(1)} = \frac{1}{10} (3 + 2x_2^{(0)} + x_3^{(0)}) = \frac{3}{10}$$

$$x_2^{(1)} = \frac{1}{5} (6 + x_1^{(0)} + x_3^{(0)}) = \frac{6}{5}$$

$$x_3^{(1)} = \frac{1}{2} (3 + x_1^{(0)} + x_2^{(0)}) = \frac{3}{2}$$

$$x^{(1)} = \begin{pmatrix} \frac{3}{10} \\ \frac{6}{5} \\ \frac{3}{2} \end{pmatrix}$$

برای $k = 1$:

$$\begin{aligned} x_1^{(2)} &= \frac{1}{10} (3 + 2x_2^{(1)} + x_3^{(1)}) \\ &= \frac{1}{10} \left(3 + 2 \left(\frac{6}{5} \right) + \frac{3}{2} \right) \\ &= \frac{1}{10} \left(3 + \frac{12}{5} + \frac{3}{2} \right) \\ &= \frac{1}{100} (30 + 24 + 15) \\ &= \frac{69}{100} \end{aligned}$$

$$\begin{aligned} x_2^{(2)} &= \frac{1}{5} (6 + x_1^{(1)} + x_3^{(1)}) \\ &= \frac{1}{5} \left(6 + \frac{3}{10} + \frac{3}{2} \right) \\ &= \frac{1}{5} \left(6 + \frac{3}{10} + \frac{15}{10} \right) \\ &= \frac{1}{50} (60 + 3 + 15) \\ &= \frac{78}{50} = \frac{39}{25} \end{aligned}$$

$$\begin{aligned}
 x_3^{(2)} &= \frac{1}{2} \left(3 + x_1^{(1)} + x_2^{(1)} \right) \\
 &= \frac{1}{2} \left(3 + \frac{3}{10} + \frac{6}{5} \right) \\
 &= \frac{1}{20} (30 + 3 + 12) \\
 &= \frac{45}{20} = \frac{9}{4}
 \end{aligned}$$

$$x^{(2)} = \begin{pmatrix} \frac{69}{100} \\ \frac{39}{25} \\ \frac{9}{4} \end{pmatrix}$$

$$\vdots$$

$$x^{(10)} = (0.99836, 1.99757, 2.99647)$$

◀ 3.1.1 فرم فشرده روش ژاکوبی

در ابتدا بیان کردیم روش های تکراری به صورت زیر می باشند:

$$x^{(k+1)} = Bx^{(k)} + c$$

که $x^{(0)}$ اولیه است.

می خواهیم ماتریس B و بردار c را برای روش ژاکوبی تعیین کنیم.

$$Ax = b \quad \text{where} \quad A = L + D + U$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$\begin{aligned}
 Ax &= b \\
 (L + D + U)x &= b \\
 Dx &= b - (L + U)x \\
 x &= D^{-1}b - D^{-1}(L + U)x \\
 x^{(k+1)} &= \underbrace{(-D^{-1}(L + U))}_{B_J} x^{(k)} + \underbrace{D^{-1}b}_{c_J} \\
 x^{(k+1)} &= B_J x^{(k)} + c_J
 \end{aligned}$$

3.2 روش گاوس-سایدل

این روش کاملاً مشابه روش ژاکوبی است با این تفاوت که در هر مرحله، از درایه‌های جدید محاسبه‌شده در همان مرحله استفاده می‌کنیم. یعنی داریم:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$

$$i = 1, 2, \dots, n$$

فرم فشرده روش گاوس-سایدل به صورت زیر می‌باشد:

$$\begin{aligned}
 x^{(k)} &= \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} \\
 x_1^{(k+1)} &= \frac{1}{a_{11}} \left(b_1 - a_{12} x_2^{(k)} - a_{13} x_3^{(k)} \right) \\
 x_2^{(k+1)} &= \frac{1}{a_{22}} \left(b_2 - a_{21} x_1^{(k+1)} - a_{23} x_3^{(k)} \right) \\
 x_3^{(k+1)} &= \frac{1}{a_{33}} \left(b_3 - a_{31} x_1^{(k+1)} - a_{32} x_2^{(k+1)} \right)
 \end{aligned}$$

$$\begin{aligned}
 Ax &= b \\
 (L + D + U)x &= b \\
 Lx + Dx + Ux &= b \\
 Lx^{(k+1)} + Dx^{(k+1)} + Ux^{(k)} &= b \\
 (L + D)x^{(k+1)} &= -Ux^{(k)} + b \\
 x^{(k+1)} &= -(L + D)^{-1}Ux^{(k)} + (L + D)^{-1}b \\
 x^{(k+1)} &= B_{GS}x^{(k)} + c_{GS}
 \end{aligned}$$

دو تکرار به روش گاوس-سایدل برای دستگاه زیر بدست آورید. نقطه شروع اولیه را $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

* مثال 3.2

در نظر بگیرید.

$$10x_1 - 2x_2 - x_3 = 3$$

$$-x_1 + 5x_2 - x_3 = 6$$

$$-x_1 - x_2 + 2x_3 = 3$$

❗ حل:

$$x_1^{(k+1)} = \frac{1}{10} (3 + 2x_2^{(k)} + x_3^{(k)})$$

$$x_2^{(k+1)} = \frac{1}{5} (6 + x_1^{(k+1)} + x_3^{(k)})$$

$$x_3^{(k+1)} = \frac{1}{2} (3 + x_1^{(k+1)} + x_2^{(k+1)})$$

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{pmatrix}$$

برای $k = 0$:

$$x_1^{(1)} = \frac{1}{10} (3 + 2x_2^{(0)} + x_3^{(0)}) = \frac{3}{10}$$

$$x_2^{(1)} = \frac{1}{5} (6 + x_1^{(1)} + x_3^{(0)}) = \frac{63}{50}$$

$$x_3^{(1)} = \frac{1}{2} (3 + x_1^{(1)} + x_2^{(1)}) = \frac{228}{100}$$

$$x^{(1)} = \begin{pmatrix} \frac{3}{10} \\ \frac{63}{50} \\ \frac{228}{100} \end{pmatrix}$$

برای $k = 1$:

$$x_1^{(2)} = \frac{1}{10} (3 + 2x_2^{(1)} + x_3^{(1)}) = \frac{780}{1000}$$

$$x_2^{(2)} = \frac{1}{5} (6 + x_1^{(2)} + x_3^{(1)}) = \frac{1812}{1000}$$

$$x_3^{(2)} = \frac{1}{2} (3 + x_1^{(2)} + x_2^{(2)}) = \frac{2796}{1000}$$

$$x^{(2)} = \begin{pmatrix} 0.78 \\ 1.812 \\ 2.796 \end{pmatrix}$$

$$\vdots$$

$$x^{(10)} = (0.99999336, 1.99999407, 2.99999372)$$

3.3 بررسی همگرایی روش های تکراری

روش های تکراری:

$$x^{(k+1)} = Bx^{(k)} + c, \quad x^{(0)} \text{ اولیه}$$

قضیه 3.1

فرض کنیم $B \in \mathbb{R}^{n \times n}$ و $\rho(B) < 1$ (بزرگترین مقدار ویژه از لحاظ قدرمطلق) آنگاه $(I - B)$ وارون پذیر است.

فرض کنیم λ یک مقدار ویژه B باشد،

اثبات ✓

$$|\lambda| \leq \rho(B) < 1$$

چون λ یک مقدار ویژه B است، یک بردار ویژه $x \neq 0$ وجود دارد به طوری که

$$Bx = \lambda x$$

حال نشان می‌دهیم $1 - \lambda$ یک مقدار ویژه برای $I - B$ می‌باشد

$$(I - B)x = x - Bx$$

$$= x - \lambda x$$

$$= (1 - \lambda)x$$

$$\Rightarrow (I - B)x = (1 - \lambda)x$$

بنابراین تمام مقادیر ویژه $I - B$ از روی مقادیر ویژه B بدست می‌آیند و داریم

$$|1 - \lambda| > 0$$

پس تمام مقادیر ویژه $I - B$ غیر صفر هستند. پس دترمینان $I - B$ ناصفر است یعنی $I - B$ وارون پذیر است. □

قضیه 3.2

فرض کنیم $B \in \mathbb{R}^{n \times n}$ آنگاه $\rho(B) < 1$ اگر و تنها اگر

$$B^n \rightarrow 0, \quad n \rightarrow \infty$$

اثبات

($\Leftarrow$) ابتدا فرض کنیم $\rho(B) < 1$ اگر λ یک مقدار ویژه B باشد آنگاه

$$\exists x \in \mathbb{R}, x \neq 0 : Bx = \lambda x$$

$$B(Bx) = B(\lambda x) = \lambda Bx = \lambda^2 x$$

$\vdots$

$$B^n x = \lambda^n x \quad (*)$$

با توجه به این شرط که $\rho(B) < 1$ برای λ داریم

$$|\lambda| \leq \rho(B) < 1 \Rightarrow |\lambda| < 1$$

حال از طرفین رابطه (*) حد می‌گیریم، داریم

$$\lim_{n \rightarrow \infty} B^n x = \lim_{n \rightarrow \infty} \lambda^n x = 0$$

$$\lim_{n \rightarrow \infty} B^n x = \left(\lim_{n \rightarrow \infty} B^n \right) x \stackrel{x \neq 0}{\Rightarrow} \lim_{n \rightarrow \infty} B^n = 0 \quad \square$$

($\Rightarrow$) فرض کنیم $\lim_{n \rightarrow \infty} B^n = 0$ می‌خواهیم نشان دهیم $\rho(B) < 1$.

فرض کنیم $\rho(B) \not< 1$ (برهان خلف) پس

$$\rho(B) \geq 1$$

$$\exists \bar{\lambda} \text{ eigenvalue, } \rho(B) = |\bar{\lambda}| \geq 1$$

$$\exists x \in \mathbb{R}^n, x \neq 0, Bx = \bar{\lambda} x$$

$$\Rightarrow B^n x = \bar{\lambda}^n x$$

$$0 = \lim_{n \rightarrow \infty} B^n x = \lim_{n \rightarrow \infty} \bar{\lambda}^n x \neq 0$$

و به تناقض می‌رسیم پس فرض خلف باطل است یعنی

$$\rho(B) < 1 \quad \square$$

قضیه 3.3

فرض کنیم $B \in \mathbb{R}^{n \times n}$ و $\rho(B) < 1$ آنگاه

$$(I - B)^{-1} = I + B + B^2 + \dots = \sum_{i=0}^{\infty} B^i$$

قرار می‌دهیم

اثبات ☒

$$C_n = \sum_{i=0}^n B^i$$

$$\begin{aligned} (I - B)C_n &= (I - B) \left(\sum_{i=0}^n B^i \right) \\ &= (I - B)(I + B + B^2 + \dots + B^n) \\ &= I - B + B - B^2 + B^2 - \dots - B^{n+1} \\ &= I - B^{n+1} \end{aligned}$$

$$\lim_{n \rightarrow \infty} (I - B)C_n = \lim_{n \rightarrow \infty} I - B^{n+1} = I$$

$$(I - B) \left(\sum_{i=0}^{\infty} B^i \right) = I$$

$$(I - B)^{-1} = \sum_{i=0}^{\infty} B^i \quad \square$$

نکته 3.1

در روش های تکراری داریم

$$x^{(k+1)} = Bx^{(k)} + c$$

اگر x جواب باشد

$$Ax = b \rightarrow x = Bx + c$$

$$\Rightarrow x - Bx = (I - B)x = c$$

$$\text{if } \rho(B) < 1, c \neq 0 \Rightarrow x = (I - B)^{-1}c$$

قضیه 3.4

اگر در روش تکراری $x^{(k+1)} = Bx^{(k)} + c$ داشته باشیم $c \neq 0$ آنگاه به ازای هر نقطه شروع دلخواه $x^{(0)}$ دنباله $\{x^{(k)}\}$ تولید شده توسط روش بالا به جواب $\bar{x} = (I - B)^{-1}c$ از دستگاه همگراست اگر و تنها اگر $\rho(B) < 1$.

اثبات

($\Rightarrow$) ابتدا فرض می کنیم $\rho(B) < 1$ می خواهیم نشان دهیم دنباله $\{x^{(k)}\}$ تولید شده توسط روش تکراری به جواب همگراست.

$$\begin{aligned} x^{(k+1)} &= Bx^{(k)} + C \\ &= B(Bx^{(k-1)} + C) + C = B^2x^{(k-1)} + BC + C \\ &= B^2(Bx^{(k-2)} + C) + BC + C = B^3x^{(k-2)} + B^2C + BC + C \\ &\vdots \\ &= B^{k+1}x^{(0)} + B^kC + B^{k-1}C + \dots + BC + C \\ &= B^{k+1}x^{(0)} + (B^k + B^{k-1} + \dots + B + I)C \end{aligned}$$

از طرفین رابطه بالا حد می گیریم، داریم:

$$\begin{aligned} \lim_{k \rightarrow \infty} x^{(k+1)} &= 0 + \lim_{k \rightarrow \infty} \sum_{i=0}^k B^i C \\ &= \sum_{i=0}^{\infty} B^i C \\ &= (I - B)^{-1}C = \bar{x} \quad \square \end{aligned}$$

($\Leftarrow$) فرض کنیم دنباله $\{x^{(k)}\}$ به $\bar{x}$ همگراست. می‌خواهیم نشان دهیم $\rho(B) < 1$

$$\begin{aligned} x^{(k+1)} - \bar{x} &= Bx^{(k)} + C - \bar{x} \\ &= B(Bx^{(k-1)} + C) + C - \bar{x} = B^2x^{(k-1)} + BC + C - \bar{x} \\ &= B^2(Bx^{(k-2)} + C) + BC + C - \bar{x} = B^3x^{(k-2)} + B^2C + BC + C - \bar{x} \\ &\vdots \\ &= B^{k+1}x^{(0)} + B^kC + B^{k-1}C + \cdots + BC + C - \bar{x} \\ &= B^{k+1}x^{(0)} + (B^k + B^{k-1} + \cdots + B + I)C - \bar{x} \end{aligned}$$

از طرفین رابطه بالا حد می‌گیریم، داریم:

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} x^{(k+1)} - \bar{x} = \lim_{k \rightarrow \infty} B^{k+1}x^{(0)} + \underbrace{\lim_{k \rightarrow \infty} \left(\sum_{i=0}^k B^i \right) C - \bar{x}}_0 \\ &\Rightarrow \lim_{k \rightarrow \infty} B^{k+1} \underbrace{x^{(0)}}_{\neq 0} = 0 \quad \Rightarrow \lim_{k \rightarrow \infty} B^{k+1} = 0 \quad \Leftrightarrow \quad \rho(B) < 1 \quad \square \end{aligned}$$

نکته 3.2

همانگونه که مشاهده کردیم همگرایی روش های تکراری کاملاً به ماتریس B وابسته می‌باشد و برای همگرایی باید

$$\rho(B) < 1$$

هرچه $\rho(B)$ کوچک تر باشد (به صفر نزدیک تر باشد) سرعت همگرایی بیشتر می‌باشد.

نکته 3.3

اگر برای یک نرم القایی از نرم برداری داشته باشیم $\|B\| < 1$ آنگاه میتوانیم نتیجه بگیریم دنباله $\{x^{(k)}\}$ تولید شده توسط روش تکراری $x^{(k+1)} = Bx^{(k)} + c$ به ازای هر نقطه $x^{(0)}$ دلخواه اولیه همگراست. چون

$$\rho(B) \leq \|B\| < 1$$

فرض کنیم دستگاه $Ax = b$ را داریم به طوری که

*** مثال 3.3**

$$A = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$$

درباره همگرایی یا واگرایی روش های تکراری ژاکوبی و گاوس-سایدل بحث کنید.

💡 حل:

$$L = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$B_J = -D^{-1}(L + U) = \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{pmatrix}$$

$$\det(\lambda I - B_J) = \det \begin{pmatrix} \lambda & 2 & -2 \\ 1 & \lambda & 1 \\ 2 & 2 & \lambda \end{pmatrix} = 0$$

$$\lambda(\lambda^2 - 2) - 2(\lambda - 2) - 2(2 - 2\lambda) = 0$$

$$\lambda^3 - 2\lambda - 2\lambda + 4 - 4 + 4\lambda = 0$$

$$\lambda^3 = 0 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0 \Rightarrow \rho(B_J) = 0 < 1$$

روش ژاکوبی همگراست.

$$B_{GS} = -(D + L)^{-1}U$$

$$(D + L)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

$$B_{GS} = - \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\Rightarrow \lambda_1 = \lambda_2 = 2, \lambda_3 = 0 \Rightarrow \rho(B_{GS}) = 2 \not< 1$$

روش گاوس-سایدل همگرا نیست.

قضیه 3.5

فرض کنیم $A \in \mathbb{R}^{n \times n}$ و غالب قطری اکید باشد دنباله حاصل از روش ژاکوبی یا گاوس-سایدل به جواب دستگاه $Ax = b$ همگرا می‌باشد.

قضیه 3.6

فرض کنیم ماتریس A غالب قطری اکید سطری باشد، دنباله $\{x^{(k)}\}$ تولید شده توسط روش ژاکوبی یا گاوس-سایدل به ازای هر نقطه شروع دلخواه همگرا به جواب دستگاه

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n$$

می‌باشد.

دستگاه $Ax = b$ را در نظر می‌گیریم جواب x از این دستگاه به صورت زیر می‌باشد و آنگاه طبق

اثبات ✓

$Ax = b$ داریم

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i, \quad i = 1, \dots, n$$

و می‌توانیم درایه‌های x_i را به صورت زیر توصیف کنیم

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij}x_j \right) (*), \quad i = 1, 2, \dots, n$$

حال از روش ژاکوبی دنباله $\{x^{(k)}\}$ به صورت زیر تعریف می‌شود

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right) (**), \quad i = 1, 2, \dots, n$$

تعریف می‌کنیم

$$e^{(k)} = x - x^{(k)}$$

و بر اساس درایه‌ها داریم

$$e_i^{(k)} = x_i - x_i^{(k)}, \quad i = 1, 2, \dots, n$$

رابطه $(**)$ را از $(*)$ کم می‌کنیم، داریم

$$\underbrace{x_i - x_i^{(k+1)}}_{e_i^{(k+1)}} = - \left(\sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} \underbrace{(x_j - x_j^{(k)})}_{e_j^{(k)}} \right)$$

$$e_i^{(k+1)} = - \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} e_j^{(k)}, \quad i = 1, \dots, n$$

$$\Rightarrow |e_i^{(k+1)}| = \left| \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} e_j^{(k)} \right|$$

$$\leq \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} |e_j^{(k)}|$$

$$\leq \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} \|e^{(k)}\|_{\infty}$$

$$= \|e^{(k)}\|_{\infty} \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|}$$

حال به توجه به اینکه ماتریس A غالب قطری اکید می‌باشد داریم

$$\sum_{j=1, j \neq i}^n |a_{ij}| < |a_{ii}|, \quad i = 1, \dots, n$$

$$\Rightarrow \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} < 1, \quad i = 1, \dots, n$$

$$\mu = \max_{1 \leq i \leq n} \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} < 1$$

بنابراین داریم

$$|e_i^{(k+1)}| \leq \|e^{(k)}\|_{\infty} \mu, \quad i = 1, \dots, n$$

$$\|e^{(k+1)}\|_{\infty} = \max |e_i^{(k+1)}| \leq \mu \|e^{(k)}\|_{\infty}$$

و داریم

$$\begin{aligned}\|e^{(k+1)}\|_{\infty} &\leq \mu^2 \|e^{(k-1)}\|_{\infty} \\ &\leq \mu^3 \|e^{(k-2)}\|_{\infty} \\ &\vdots \\ &\leq \mu^{k+1} \|e^{(0)}\|_{\infty}\end{aligned}$$

از طرفین رابطه فوق حد می‌گیریم و با توجه به اینکه $\mu < 1$ داریم

$$\begin{aligned}\lim_{k \rightarrow \infty} \|e^{(k+1)}\|_{\infty} &= 0 \\ \Rightarrow x^{(k)} &\rightarrow x\end{aligned}$$

و از ابتدا x را جواب دستگاه $Ax = b$ در نظر گرفته بودیم. بنابراین دنباله حاصل از روش ژاکوبی به جواب دستگاه $Ax = b$ همگرا می‌باشد. به صورت مشابه برای معادله حاصل از روش گاوس-سایدل نیز می‌توان اثبات کرد $\square$

نکته 3.4

همانگونه که در قضیه فوق مشاهده کردیم اگر ماتریس ضرایب دستگاه $Ax = b$ غالب قطری اکید باشد دنباله حاصل از روش ژاکوبی و گاوس-سایدل همگرا می‌باشد. اگر این شرط برقرار نبود با جابه‌جایی سطرها گاهی اوقات می‌توان این شرط را برقرار کرد.

3.4 تسريع همگرایی روش گاوس-سایدل (روش SOR)

می‌دانیم دنباله حاصل از روش گاوس-سایدل به صورت زیر می‌باشد:

$$\begin{aligned}x_i^{(k+1)} &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \quad i = 1, 2, \dots, n \\ x_i^{(k+1)} &= x_i^{(k)} + \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \quad i = 1, 2, \dots, n \\ x_i^{(k+1)} - x_i^{(k)} &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \quad i = 1, 2, \dots, n\end{aligned}$$

رابطه بالا تفاضل دو جمله $x_i^{(k+1)}$ ، $x_i^{(k)}$ می‌باشد و به ازای هر $i = 1, 2, \dots, n$ برقرار است. در صورتی که گاوس-سایدل همگرا باشد با ضرب رابطه بالا در یک ضریب مانند $\omega > 0$ می‌توانیم سرعت همگرایی روش را افزایش دهیم. بنابراین داریم.

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^n a_{ij} x_j^{(k)} \right) \quad i = 1, 2, \dots, n$$

فرمول فوق روش SOR برای حل تکراری دستگاه $Ax = b$ را ارائه می‌دهد و به ازای $\omega = 1$ همان روش گاوس-سایدل است. حال فرمول را به صورت ساده‌تری می‌نویسیم.

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right) \quad i = 1, 2, \dots, n \quad (***)$$

حال می‌خواهیم B_{SOR}, c_{SOR} را تعیین کنیم، به طوری که

$$x^{(k+1)} = B_{SOR}x^{(k)} + c_{SOR}$$

ابتدا رابطه $(***)$ را در a_{ii} ضرب می‌کنیم

$$a_{ii}x_i^{(k+1)} = (1 - \omega)a_{ii}x_i^{(k)} + \omega \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right) \quad i = 1, 2, \dots, n$$

$$Dx^{(k+1)} = (1 - \omega)Dx^{(k)} + \omega b - \omega Lx^{(k+1)} - \omega Ux^{(k)}$$

$$(D + \omega L)x^{(k+1)} = ((1 - \omega)D - \omega U)x^{(k)} + \omega b$$

$$x^{(k+1)} = \underbrace{(D + \omega L)^{-1}((1 - \omega)D - \omega U)}_{B_{SOR}} x^{(k)} + \underbrace{\omega(D + \omega L)^{-1}b}_{c_{SOR}}$$

دو تکرار از روش SOR را برای حل دستگاه زیر بدست آورید.

*** مثال 3.4**

$$A = \begin{pmatrix} 5 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 5 \end{pmatrix}, \quad b = \begin{pmatrix} 7 \\ 7 \\ 7 \end{pmatrix}, \quad \omega = 1.2, \quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

💡 حل:

$$\begin{aligned} x_1^{(1)} &= (1 - \omega)x_1^{(0)} + \frac{\omega}{a_{11}}(b_1 - a_{12}x_2^{(0)} - a_{13}x_3^{(0)}) \\ &= (1 - 1.2) \cdot 0 + \frac{1.2}{5}(7 - 0 - 0) = 1.68 \end{aligned}$$

$$\begin{aligned} x_2^{(1)} &= (1 - \omega)x_2^{(0)} + \frac{\omega}{a_{22}}(b_2 - a_{21}x_1^{(1)} - a_{23}x_3^{(0)}) \\ &= (1 - 1.2) \cdot 0 + \frac{1.2}{5}(7 - 1.68 - 0) = 1.2768 \end{aligned}$$

$$\begin{aligned} x_3^{(1)} &= (1 - \omega)x_3^{(0)} + \frac{\omega}{a_{33}}(b_3 - a_{31}x_1^{(1)} - a_{32}x_2^{(1)}) \\ &= (1 - 1.2) \cdot 0 + \frac{1.2}{5}(7 - 1.68 - 1.2768) = 0.9704 \end{aligned}$$

$$x^{(1)} = \begin{pmatrix} 1.68 \\ 1.2768 \\ 0.9704 \end{pmatrix}$$

$$\begin{aligned} x_1^{(2)} &= (1 - \omega)x_1^{(1)} + \frac{\omega}{a_{11}}(b_1 - a_{12}x_2^{(1)} - a_{13}x_3^{(1)}) \\ &= (1 - 1.2) \cdot 1.68 + \frac{1.2}{5}(7 - 1.2768 - 0.9704) = 0.8047 \end{aligned}$$

$$\begin{aligned} x_2^{(2)} &= (1 - \omega)x_2^{(1)} + \frac{\omega}{a_{22}}(b_2 - a_{21}x_1^{(2)} - a_{23}x_3^{(1)}) \\ &= (1 - 1.2) \cdot 1.2768 + \frac{1.2}{5}(7 - 0.8047 - 0.9704) = 0.9986 \end{aligned}$$

$$\begin{aligned} x_3^{(2)} &= (1 - \omega)x_3^{(1)} + \frac{\omega}{a_{33}}(b_3 - a_{31}x_1^{(2)} - a_{32}x_2^{(2)}) \\ &= (1 - 1.2) \cdot 0.9704 + \frac{1.2}{5}(7 - 0.8047 - 0.9986) = 1.0531 \end{aligned}$$

$$x^{(2)} = \begin{pmatrix} 0.8047 \\ 0.9986 \\ 1.0531 \end{pmatrix}$$

قضیه 3.7

در روش SOR اگر ω خارج از بازه $(0, 2)$ باشد ممکن است به ازای یک نقطه شروع دلخواه، روش همگرا نباشد.

ماتریس B_{SOR} را در نظر بگیریم داریم

اثبات ✓

$$B_{SOR} = (D + \omega L)^{-1}((1 - \omega)D - \omega U)$$

- ماتریس D قطری با قطر a_{ii} - ماتریس L پایین مثلثی با قطر صفر - ماتریس U بالا مثلثی با قطر صفر
 $((1 - \omega)D - \omega U)$ - ماتریس $(D + \omega L)^{-1}$ پایین مثلثی با درایه‌های قطری $\frac{1}{a_{ii}}$ - ماتریس $(1 - \omega)a_{ii}$ یک ماتریس بالا مثلثی با قطر $(1 - \omega)$
 یک ماتریس با قطر $(1 - \omega)$

$$\det B_{SOR} = \det((D + \omega L)^{-1}) \det((1 - \omega)D - \omega U)$$

$$= \frac{1}{a_{11}a_{22}\dots a_{nn}} (1 - \omega)^n (a_{11}a_{22}\dots a_{nn})$$

$$\det B_{SOR} = (1 - \omega)^n$$

اگر $\lambda_1, \lambda_2, \dots, \lambda_n$ مقادیر ویژه B_{SOR} باشند، داریم:

$$(p(B_{SOR}))^n \geq |\lambda_1 \lambda_2 \cdots \lambda_n| = |\det B_{SOR}| = |(1 - \omega)|^n$$

$$1 > p(B_{SOR}) \geq |1 - \omega|$$

$$|1 - \omega| < 1$$

$$-1 < 1 - \omega < 1$$

$$0 < \omega < 2 \quad \square$$

نکته 3.5

اگر روش گاوس-سایدل همگرا نباشد با انتخاب $0 < \omega < 1$ میتوان روش SOR همگرا را بدست آورد.

نکته 3.6

اگر روش گاوس-سایدل همگرا باشد با انتخاب $1 < \omega < 2$ میتوان روش گاوس-سایدل را تسریع کرد.



حل معادلات غیرخطی

در بسیاری از مسائل اجتماعی، اقتصادی، مهندسی و ... با حل یک مسئله به صورت $f(x) = 0$ روبه‌رو می‌شویم و برای محاسبه جواب مسئله لازم است ریشه تابع $f(x)$ را محاسبه کنیم. اگر $f(x)$ یک چند جمله‌ای از درجه ۱ باشد یعنی

$$f(x) = ax + b$$

آنگاه حل معادله $f(x) = 0$ به سادگی امکان‌پذیر می‌باشد و همه می‌دانیم که ریشه این تابع به صورت $-\frac{b}{a}$ می‌باشد پس در این حالت مسئله حل شده است. علاوه بر این اگر $f(x)$ یک چند جمله‌ای از درجه ۲ باشد یعنی

$$f(x) = ax^2 + bx + c$$

آنگاه برای حل مسئله $f(x) = 0$ نیز یک راه حل کلی داریم و تنها لازم است مقدار $\Delta = b^2 - 4ac$ را حساب کنیم و با توجه به اینکه Δ ، مثبت، منفی یا صفر باشد می‌توانیم به صورت دقیق تعداد ریشه‌ها و خود ریشه‌ها را محاسبه کنیم. به غیر از این دو حالت بالا حل مسئله $f(x) = 0$ زمانی که f تابعی دلخواه باشد به سادگی امکان‌پذیر نیست. به همین دلیل در این فصل می‌خواهیم روش‌های عددی برای تعیین ریشه تابع (به صورت تقریبی) را معرفی کنیم. روش‌هایی که در اینجا معرفی می‌کنیم به دو دسته تقسیم می‌شوند. برخی از این روش‌ها به یک بازه شامل ریشه تابع احتیاج دارند و دیگر روش‌ها به تقریبی (نادقیق) از ریشه تابع نیاز دارند. پس ابتدا باید به این سوال پاسخ دهیم که چگونه تقریبی از ریشه یا یک بازه شامل ریشه به دست آوریم؟ برای این منظور دو روش داریم:

۱. رسم منحنی
۲. جدول مقادیر تابع

4.0.1 رسم منحنی

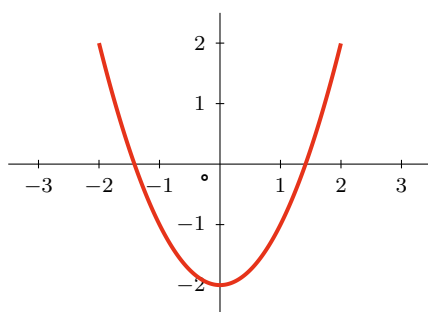
$$f : \mathbb{R} \mapsto \mathbb{R}$$

در این روش، منحنی تابع را رسم می‌کنیم. لازم نیست منحنی به صورت دقیق ترسیم شود. با رسم منحنی می‌توانیم تقریبی از ریشه یا یک بازه شامل ریشه را تعیین کنیم.

* مثال 4.1

$$f(x) = x^2 - 2$$

x	0	1	-1	2	-2
y	-2	-1	-1	2	2

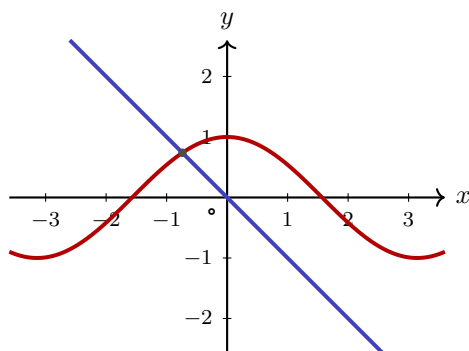


یک ریشه در بازه $[-2, -1]$ و ریشه دیگر در $[1, 2]$ قرار دارد.

* مثال 4.2

$$f(x) = \cos(x) + x$$

$$\cos(x) + x = 0 \Rightarrow \cos(x) = -x$$

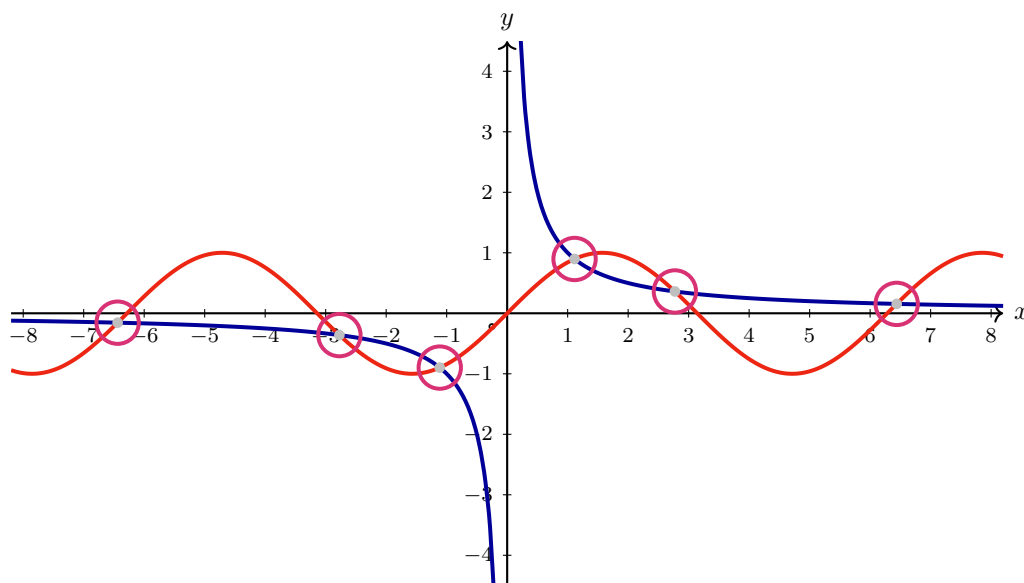


ریشه در بازه $[-\frac{\pi}{2}, 0]$ قرار دارد.

* مثال 4.3

$$f(x) = x \sin(x) - 1$$

$$x \sin(x) - 1 = 0 \xRightarrow{x=0 \text{ یک ریشه نیست}} \sin(x) = \frac{1}{x}$$



◀ 4.0.2 جدول مقادیر تابع

در این روش به ازای مقادیر مختلف مقدار تابع را محاسبه می‌کنیم و مقادیر محاسبه شده را در جدولی قرار می‌دهیم. در هر جایی از جدول که تغییر علامت دو مقدار تابع داشته باشیم با توجه به پیوستگی تابع نتیجه می‌گیریم که در آن بازه حتماً یک ریشه داریم و بازه شامل ریشه تعیین می‌شود.

☰ قضیه 4.1 | مقدار میانی

اگر f در بازه $[a, b]$ پیوسته باشد و

$$f(a)f(b) < 0$$

آنگاه (حداقل یک) نقطه $c \in (a, b)$ وجود دارد به طوری که $f(c) = 0$. در صورتی که f در این بازه اکیداً صعودی یا اکیداً نزولی باشد آنگاه می‌توان گفت f در این بازه ریشه یکتا دارد.

$$\forall x \in (a, b) : f'(x) > 0 \text{ or } f'(x) < 0$$

تعریف 4.1 | شرایط (*)

شرایط (*) را به صورت زیر تعریف می‌کنیم. فرض کنیم تابع f و بازه $[a, b]$ داده شده باشند. اگر شرایط زیر برقرار باشند گوییم شرایط (*) برای تابع f در بازه $[a, b]$ برقرار است:

۱. تابع f در بازه $[a, b]$ پیوسته باشد.
- ۲.

$$f(a)f(b) < 0$$

۳.

$$\forall x \in (a, b) : f'(x) > 0 \text{ or } f'(x) < 0$$

اگر شرایط (*) برقرار باشد می‌دانیم f در بازه $[a, b]$ تنها یک ریشه دارد. اگر فقط دو شرط ابتدایی شرایط (*) برقرار باشد نتیجه می‌گیریم f در بازه $[a, b]$ حداقل یک ریشه دارد.

در اینجا چند روش عددی را برای محاسبه ریشه ارائه می‌دهیم. ساختار کلی تمام این روش‌ها به صورت زیر است ابتدا باید تابع f و بازه $[a, b]$ را داشته باشیم که در شرایط (*) صدق کند. از یک نقطه دلخواه $x^{(0)} \in [a, b]$ الگوریتم را آغاز می‌کنیم. با یک روش تکراری از روی $x^{(0)}$ نقطه جدید $x^{(1)}$ را به دست می‌آوریم. مجدد با روش تکراری از $x^{(1)}$ نقطه جدید $x^{(2)}$ را محاسبه می‌کنیم و این فرایند را ادامه می‌دهیم و دنباله $x^{(k)}$ را به دست می‌آوریم که

$$\lim_{k \rightarrow \infty} x^{(k)} = \alpha, \quad f(\alpha) = 0$$

4.1 سرعت همگرایی الگوریتم‌ها

فرض کنیم برای ریشه‌یابی چندین الگوریتم همگرا معرفی کرده‌ایم. حال سوالی که در اینجا مطرح می‌شود این است که از بین این الگوریتم‌ها کدام الگوریتم برای حل مسئله مناسب‌تر می‌باشد؟ مشخصاً الگوریتمی که سرعت همگرایی آن بیشتر باشد و سریع‌تر به ریشه برسد مناسب‌تر است. برای مطالعه سرعت همگرایی تعریف‌های زیر را داریم.

تعریف 4.2

گوییم دنباله $\{x_n\}$ همگرا به α با مرتبه حداقل خطی (مرتبه یک) است هرگاه $c \in [0, 1)$ وجود داشته باشد به طوری که

$$|x_{n+1} - \alpha| \leq c|x_n - \alpha| \quad \forall n \geq \bar{n}, \bar{n} \in \mathbb{N}$$

تعریف 4.3

می‌گوییم دنباله x_n همگرا به α با مرتبه حداقل ابرخطی است هرگاه دنباله

$$c_n \subseteq [0, 1), \quad c_n \rightarrow 0$$

وجود داشته باشد به طوری که

$$|x_{n+1} - \alpha| \leq c_n |x_n - \alpha| \quad \forall n \geq \bar{n}, \bar{n} \in \mathbb{N}$$

تعریف 4.4

تعریف گوییم دنباله $\{x_n\}$ همگرا به α با مرتبه p است هرگاه ثابت‌های $\bar{n} \in \mathbb{N}$, $c \in [0, 1)$, $p \geq 1$ وجود داشته باشند به طوری که

$$|x_{n+1} - \alpha| \leq c |x_n - \alpha|^p, \quad \forall n \geq \bar{n}$$

در عمل برای اینکه الگوریتم‌ها قابل کاربرد باشند لازم است یک یا چند شرط توقف برای آنها در نظر بگیریم. برای مسئله مورد مطالعه در اینجا می‌توانیم یک یا چند شرط از شرایط زیر را به صورت همزمان در نظر بگیریم: اگر در مرحله n ام داشته باشیم

$$|f(x_n)| < \epsilon, \epsilon > 0$$

۱.

$$|x_{n+1} - x_n| < \epsilon$$

۲.

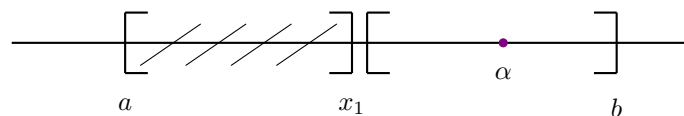
$$|f(x_{n+1}) - f(x_n)| < \epsilon$$

۳.

۴. تعداد تکرارها داده شده باشد.

4.2 روش دوبخشی یا نصف کردن (تنصیف)

در این روش با تابع $f(x)$ شروع می‌کنیم که روی بازه $[a, b]$ در شرایط (*) صدق می‌کند. نقطه وسط بازه را مشخص می‌کنیم در واقع بازه را به دو قسمت تقسیم می‌کنیم و بعد زیر بازه‌ای که ریشه در آن قرار دارد را مشخص می‌کنیم و زیربازه دیگر را حذف می‌کنیم.



حال فرایند فوق را روی زیر بازه جدید تکرار می‌کنیم و این عملیات را ادامه می‌دهیم تا جایی که شرط یا شرایط توقف برقرار باشد و الگوریتم را متوقف می‌کنیم.

❖ الگوریتم | روش دو بخشی

گام 0: تابع f و بازه $[a, b]$ که در شرایط (*) صدق می‌کند در دسترس باشند. $n = 1$
 گام 1: شرط توقف را بررسی می‌کنیم اگر شرط توقف برقرار باشد متوقف می‌شویم در غیر اینصورت قرار دهید

$$x_n = \frac{a+b}{2}$$

گام 2: مقدار $f(a) \cdot f(x_n)$ را محاسبه می‌کنیم. • اگر $\underbrace{f(a) \cdot f(x_n)}_{\neq 0} = 0$ آنگاه x_n همان ریشه مورد نظر است و الگوریتم متوقف می‌شود.

• اگر $f(a) \cdot f(x_n) > 0$ آنگاه قرار دهید $a = x_n$ و $n = n + 1$ و به گام 1 بروید.
 • اگر $f(a) \cdot f(x_n) < 0$ آنگاه قرار دهید $b = x_n$ و $n = n + 1$ و به گام 1 بروید.

* مثال 4.4

نشان دهید معادله $f(x) = x + \cos x = 0$ در بازه $[-1, 0]$ یک ریشه یکتا دارد. تقریبی از ریشه به روش دو بخشی را به دست آورید.

💡 حل:

بررسی شرایط (*)

۱. تابع $\cos x$ یک تابع مثلثاتی است و پیوسته است پس در بازه $[-1, 0]$ پیوسته هست. تابع x یک چند جمله‌ای از درجه ۱ هست و چون تمام چند جمله‌ای‌ها همه جا پیوسته هستند پس تابعی پیوسته هست. از طرفی مجموع دو تابع پیوسته، پیوسته است پس

$$f(x) = x + \cos x$$

در بازه $[-1, 0]$ تابعی پیوسته است.
 ۲.

$$f(a)f(b) = f(-1)f(0) \stackrel{?}{<} 0$$

$$f(0) = 0 + \cos(0) = 1 > 0$$

$$f(-1) = -1 + \cos(-1) < 0$$

$$f(a)f(b) = f(-1)f(0) < 0$$

۳.

$$f'(x) = 1 - \sin x$$

$$\Rightarrow -1 \leq \sin x \leq 1$$

$$\Rightarrow -1 \leq -\sin x \leq 1$$

$$\Rightarrow 0 \leq 1 - \sin x \leq 2$$

$$\Rightarrow f'(x) = 1 - \sin x \geq 0$$

n	a	b	$x_n = \frac{a+b}{2}$	$f(a)f(x_n)$
1	-1	0	-0.5	-
2	-1	-0.5	-0.75	+
3	-0.75	-0.5	-0.625	-
4	-0.75	-0.625	-0.6875	-
5	-0.6875	-0.625	-0.65625	+
6	-0.65625	-0.625	-0.640625	+

4.2.1 < همگرایی روش دوبخشی

روش دوبخشی همگرایی تضمین شده دارد. چون طبق فرآیند الگوریتم داریم:

$$\begin{aligned} |x_1 - \alpha| &\leq \frac{b-a}{2} \\ |x_2 - \alpha| &\leq \frac{b-a}{2^2} \\ &\vdots \\ 0 \leq |x_n - \alpha| &\leq \frac{b-a}{2^n} \end{aligned}$$

با گرفتن حد از دو طرف داریم:

$$\begin{aligned} \lim_{n \rightarrow \infty} |x_n - \alpha| &\leq \lim_{n \rightarrow \infty} \frac{b-a}{2^n} = 0 \\ \Rightarrow \lim_{n \rightarrow \infty} x_n &= \alpha \end{aligned}$$

چند تکرار از روش دوبخشی لازم است تا تقریبی از ریشه تابع $f(x) = x^3 + \cos x$ در بازه $[-1, 0]$

* مثال 4.5

به دست آوریم به طوری که

$$|x_n - \alpha| \leq 10^{-3}$$

❗ حل:

برای روش دو بخشی داریم:

$$\begin{aligned} |x_n - \alpha| &\leq \frac{b-a}{2^n} \leq 10^{-3} \Rightarrow n = ? \\ \frac{0 - (-1)}{2^n} &\leq 10^{-3} \Rightarrow 2^n \geq 1000 \\ \Rightarrow n &= 10 \end{aligned}$$

همانگونه که اشاره کردیم روش دوبخشی همگرایی تضمین شده دارد و دنباله $\{x_n\}$ تولید شده توسط این الگوریتم

حتما به ریشه α در $[a, b]$ همگراست و علاوه بر این یک کران خطا نیز برای $|x_n - \alpha|$ به دست آوریم. اما این روش، روشی کند است و برای رسیدن به دقت مطلوب معمولا تعداد تکرار زیادی احتیاج داریم. علاوه بر این تعداد تکرارها به تابع f وابسته نیست (یعنی برای مسائل ساده یا پیچیده این الگوریتم یکسان عمل می کند).

4.3 روش نابجایی

تا حدودی مشابه به روش دوبخشی عمل می کند فقط در این روش نحوه تعیین x_n متفاوت می باشد. مطابق قبل فرض می کنیم تابع f و بازه $[a, b]$ در دسترس می باشند. به طوری که شرایط (*) برقرار است. معادله خط گذرنده از نقاط $(a, f(a))$ و $(b, f(b))$ به صورت زیر می باشد:

$$y - f(b) = \frac{f(b) - f(a)}{b - a}(x - b)$$

حال برخورد این خط و محور x ها را به دست می آوریم:

$$0 - f(b) = \frac{f(b) - f(a)}{b - a}(x_n - b)$$

$$x_n = -f(b) \frac{b - a}{f(b) - f(a)} + b$$

$$\Rightarrow x_n = \frac{f(b)a - f(a)b}{f(b) - f(a)}$$

❖ الگوریتم | روش نابجایی

گام 0: تابع f و بازه $[a, b]$ که در شرایط (*) صدق می کند در دسترس باشند. $n = 1$
گام 1: شرط توقف را بررسی می کنیم اگر شرط توقف برقرار باشد متوقف می شویم در غیر این صورت قرار دهید

$$x_n = \frac{f(b)a - f(a)b}{f(b) - f(a)}$$

گام 2: مقدار $f(a) \cdot f(x_n)$ را محاسبه می کنیم.

• اگر $f(a) \cdot f(x_n) = 0$ آنگاه x_n همان ریشه مورد نظر است و الگوریتم متوقف می شود.

• اگر $f(a) \cdot f(x_n) > 0$ آنگاه قرار دهید $a = x_n$ و $n = n + 1$ و به گام 1 بروید.

• اگر $f(a) \cdot f(x_n) < 0$ آنگاه قرار دهید $b = x_n$ و $n = n + 1$ و به گام 1 بروید.

تقریبی از ریشه $f(x) = x + \cos(x)$ را در بازه $[-1, 0]$ به دست آورید.

* مثال 4.6

4.4 روش تکرار ساده (نقطه ثابت)

معادله $f(x) = 0$ را در نظر می‌گیریم و آن را به فرم $g(x) = x$ می‌نویسیم به طوری که ریشه تابع f همان نقطه ثابت تابع g باشد. با یک نقطه دلخواه x_0 شروع می‌کنیم و دنباله $\{x_n\}$ را به صورت زیر به دست می‌آوریم.

$$x_1 = g(x_0)$$

$$x_2 = g(x_1)$$

$$\vdots$$

$$x_{n+1} = g(x_n)$$

در اینجا سوالاتی مطرح می‌شود:

۱. آیا دنباله $\{x_n\}$ به ازای هر انتخابی از g و x_0 همگراست؟
۲. آیا تابع g در سرعت همگرایی دنباله تاثیر گذار است؟
۳. آیا می‌توان شرایط لازم برای همگرایی دنباله را مشخص کرد؟
۴. اگر g را ثابت در نظر بگیریم و نقاط شروع را متفاوت در نظر بگیریم در همگرایی یا واگرایی دنباله تاثیر گذار است؟

* مثال 4.7

معادله $x^2 + x - 1 = 0$ را در نظر بگیرید. می‌خواهیم با روش تکرار ساده تقریبی از ریشه مثبت

این تابع در بازه $[0, 1]$ به دست آوریم. $x_0 = 0.5$ در نظر بگیرید.

💡 حل:

پاسخ دقیق به روش دلتا:

$$\Delta = b^2 - 4ac = 1 - (4)(1)(-1) = 5$$

$$\alpha_1, \alpha_2 = \frac{-b \pm \sqrt{\Delta}}{2a} \xrightarrow{\text{positive root}} \alpha_1 = 0.6180$$

حال چند g پیدا می‌کنیم و همگرایی آنها را به $\alpha_1 = 0.6180$ بررسی می‌کنیم

$$x^2 + x - 1 = 0 \Rightarrow 1 - x^2 = x = g_1(x)$$

$$x^2 + x - 1 = 0 \Rightarrow x^2 = 1 - x \Rightarrow x = \pm\sqrt{1-x} \Rightarrow g_2(x) = \sqrt{1-x}$$

$$x^2 + x - 1 = 0 \Rightarrow x^2 + x = 1 \Rightarrow x(x+1) = 1 \Rightarrow x = \frac{1}{1+x} \Rightarrow g_3(x) = \frac{1}{1+x}$$

$$\xrightarrow{+x^2+1} 2x^2 + x = x^2 + 1 \Rightarrow x(2x+1) = x^2 + 1 \Rightarrow \frac{x^2+1}{2x+1} = x = g_4(x)$$

n	x_n	$x_{n+1} = g_1(x_n)$	$x_{n+1} = g_2(x_n)$	$x_{n+1} = g_3(x_n)$	$x_{n+1} = g_4(x_n)$
0	0.5	0.75	0.7071	0.6666	0.625
1	-	0.4375	0.5412	0.6000	0.6180
2	-	0.8085	0.6773	0.625	
3	-	0.3463	0.5686	0.615	
4	-		0.6580	0.6154	
⋮	⋮	⋮	⋮	⋮	
8	-			0.6180	
⋮	⋮	⋮	⋮		
34	-	واگرا	0.6180		

با توجه به نتایجی که در بالا به دست آوردیم می‌توانیم نتیجه بگیریم که همگرایی یا واگرایی دنباله حاصل از روش تکرار ساده کاملاً به انتخاب g بستگی دارد. علاوه بر آن با توجه به نتایج حاصل از g_4, g_3, g_2 نتیجه می‌گیریم انتخاب g در سرعت همگرایی دنباله تاثیر دارد.

اگر تابع g_2 را در نظر بگیریم و به جای نقطه شروع 0.5 قرار دهیم $x_0 = 0$ آنگاه دنباله حاصل از روش تکرار ساده واگرا خواهد بود. این مثال به وضوح نشان می‌دهد همگرایی یا واگرایی دنباله حاصل از تکرار ساده کاملاً به نقطه شروع وابسته است.

4.4.1 < همگرایی روش تکرار ساده

همانگونه که ملاحظه کردیم روش تکرار ساده کاملاً به تابع g وابسته است. حال می‌خواهیم شرایط لازم روی g را بیان کنیم به طوری که مطمئن باشیم حتماً دنباله حاصل همگرا می‌شود.

قضیه 4.2

فرض می‌کنیم تابع f روی بازه $[a, b]$ پیوسته است و روی بازه (a, b) مشتق‌پذیر است آنگاه

$$\exists \beta \in (a, b) \quad s.t. \quad f(b) - f(a) = f'(\beta)(b - a)$$

قضیه بالا را برای تابع g در نظر می‌گیریم و قرار می‌دهیم $a = \alpha$ و $b = x_n$ که همان

اثبات | همگرایی

نقطه ثابت g است یعنی $g(\alpha) = \alpha$ ($f(\alpha) = 0$).

$$\underbrace{g(x_n)}_{x_{n+1}} - \underbrace{g(\alpha)}_{\alpha} = g'(\beta_n)(x_n - \alpha)$$

$$x_{n+1} - \alpha = g'(\beta_n)(x_n - \alpha)$$

نقطه β بین α ، x_n

$$|x_{n+1} - \alpha| = |g'(\beta_n)| |x_n - \alpha| \quad (**)$$

$$|g'(x)| \leq L < 1 \quad \forall x \in [a, b] \quad (\#)$$

طبق رابطه $(**)$ برای اینکه دنباله $\{x_n\}$ همگرا به ریشه یعنی α شود مقادیر $|x_n - \alpha|$ باید در هر مرحله کوچک شوند. واضح است اگر شرایط $(\#)$ برقرار باشد این نتیجه حاصل می‌شود. حال یکتایی را اثبات می‌کنیم، به خلاف فرض کنیم دو ریشه α_1, α_2 در بازه $[a, b]$ برای h داریم.

$$h(\alpha_1) = 0 \Rightarrow g(\alpha_1) = \alpha_1$$

$$h(\alpha_2) = 0 \Rightarrow g(\alpha_2) = \alpha_2$$

$$\alpha_1 - \alpha_2 = g(\alpha_1) - g(\alpha_2) = g'(\beta)(\alpha_1 - \alpha_2)$$

$$\alpha_1, \alpha_2 \in [a, b], \quad \beta \text{ between } \alpha_1, \alpha_2 \Rightarrow \beta \in [a, b]$$

$$|\alpha_1 - \alpha_2| = |g'(\beta)| |\alpha_1 - \alpha_2| \leq L |\alpha_1 - \alpha_2| < |\alpha_1 - \alpha_2| \quad \square$$

قضیه 4.3

با توجه به شرایط قضیه قبل به ازای هر $x_0 \in [a, b]$ دنباله حاصل از $x_{n+1} = g(x_n)$ به جواب معادله $x = g(x)$ همگرا می‌باشد.

$$g : [a, b] \mapsto [a, b]$$

$$|g'(x)| \leq L < 1 \quad \forall x \in [a, b]$$

$$\Rightarrow x_0 \in [a, b], \quad x_{n+1} = g(x_n) \Rightarrow x_n \rightarrow \alpha$$

چون g تابعی از $[a, b]$ به $[a, b]$ می‌باشد به ازای هر $x_0 \in [a, b]$ داریم

اثبات 

$$\{x_n\} \subseteq [a, b]$$

$$\begin{aligned} x_0 \in [a, b] &\xrightarrow{g} x_1 \in [a, b] \\ &\xrightarrow{g} x_2 \in [a, b] \\ &\vdots \end{aligned}$$

و طبق روش نقطه ثابت داریم

$$x_{n+1} - \alpha = g(x_n) - g(\alpha) = g'(\beta_n)(x_n - \alpha)$$

نقطه β_n بین α ، x_n پس $\beta_n \in [a, b]$

$$|x_{n+1} - \alpha| = |g'(\beta_n)| |x_n - \alpha| \leq L |x_n - \alpha|$$

$$\begin{aligned} \Rightarrow |x_{n+1} - \alpha| &\leq L |x_n - \alpha| \\ &\leq L^2 |x_{n-1} - \alpha| \\ &\vdots \\ &\leq L^{n+1} |x_0 - \alpha| \end{aligned}$$

بنابراین داریم

$$0 \leq |x_{n+1} - \alpha| \leq L^{n+1} |x_0 - \alpha|$$

$$\Rightarrow \lim_{n \rightarrow \infty} |x_{n+1} - \alpha| = 0$$

$$\Rightarrow x_{n+1} \rightarrow \alpha \quad \square$$

نکته 4.1

در اثبات قضیه فوق دیدیم که همگرایی x_n به α به حد $\lim_{n \rightarrow \infty} L^{n+1} = 0$ وابسته می‌باشد و این حد به مقدار L وابسته است. هرچه L کوچکتر باشد (به صفر نزدیک‌تر باشد) سرعت همگرایی بیشتر است و هرچه به یک نزدیک‌تر باشد سرعت همگرایی کندتر است.

4.4.2 مرتبه همگرایی روش تکرار ساده

قبلاً مرتبه همگرایی یک دنباله را تعریف کردیم حال یک تعریف معادل دیگر در اینجا بیان می‌کنیم که برای تعیین همگرایی دنباله‌ها این تعریف مناسب‌تر می‌باشد.

تعریف 4.5

فرض کنیم $\{x_n\}$ یک دنباله همگرا به α باشد. گوئیم x_n همگرا به α با مرتبه همگرایی p است هرگاه اعداد ثابت، حقیقی و مثبت p, c وجود داشته باشد که

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = c$$

اگر $p = 1$, $c \neq 0$ آنگاه مرتبه همگرایی را خطی گوئیم و در صورتی که $p = 1$, $c = 0$ آنگاه مرتبه همگرایی را ابرخطی گوئیم.

قضیه 4.4

فرض کنیم $\{x_n\}$ دنباله حاصل از تکرار ساده باشد که به α همگراست و α نقطه ثابت $x = g(x)$ می‌باشد. اگر $g'(\alpha) \neq 0$ آنگاه مرتبه همگرایی دنباله 1 است.

اثبات

باید نشان دهیم رابطه زیر برقرار است

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = c \neq 0$$

سری تیلور تابع g را حول نقطه α می‌نویسیم، داریم

$$g(x) = g(\alpha) + \frac{g'(\alpha)}{1!}(x - \alpha) + \frac{g''(\eta_n)}{2!}(x - \alpha)^2$$

که η_n نقطه‌ای بین x و α می‌باشد. رابطه فوق را به ازای x_n می‌نویسیم، داریم

$$\underbrace{g(x_n)}_{x_{n+1}} = \underbrace{g(\alpha)}_{\alpha} + \frac{g'(\alpha)}{1!}(x_n - \alpha) + \frac{g''(\eta_n)}{2!}(x_n - \alpha)^2 \quad (*)$$

که η_n نقطه‌ای بین x_n و α می‌باشد.

$$x_{n+1} - \alpha = \frac{g'(\alpha)}{1!}(x_n - \alpha) + \frac{g''(\eta_n)}{2!}(x_n - \alpha)^2$$

$$\Rightarrow \frac{x_{n+1} - \alpha}{x_n - \alpha} = g'(\alpha) + \frac{1}{2}g''(\eta_n)(x_n - \alpha)$$

با گرفتن حد از طرفین رابطه بالا داریم

$$\lim_{n \rightarrow \infty} \left| \frac{x_{n+1} - \alpha}{x_n - \alpha} \right| = \lim_{n \rightarrow \infty} \left| g'(\alpha) + \frac{1}{2}g''(\eta_n)(x_n - \alpha) \right| = |g'(\alpha)| \neq 0$$

پس مرتبه همگرایی دنباله حاصل از روش تکرار ساده، یک (خطی) است. $\square$

4.5 قضیه

در صورتی که $g'(\alpha) = 0$ آنگاه مرتبه همگرایی حداقل 2 است.

اثبات 

طبق رابطه (*) از قضیه قبل داریم

$$x_{n+1} - \alpha = \underbrace{g'(\alpha)(x_n - \alpha)}_0 + \frac{g''(\eta_n)}{2!}(x_n - \alpha)^2$$

$$x_{n+1} - \alpha = \frac{g''(\eta_n)}{2}(x_n - \alpha)^2$$

حال طرفین رابطه را به $(x_n - \alpha)^2$ تقسیم می‌کنیم

$$\frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{g''(\eta_n)}{2}$$

$$\lim_{n \rightarrow \infty} \left| \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} \right| = \lim_{n \rightarrow \infty} \left| \frac{g''(\eta_n)}{2} \right|$$

چون η_n بین x_n و α است و $x_n \rightarrow \alpha$ پس

$$\eta_n \rightarrow \alpha$$

بنابراین

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{(x_n - \alpha)^2} = \frac{g''(\alpha)}{2}$$

اگر $g''(\alpha) \neq 0$ آنگاه مرتبه همگرایی 2 است در غیر این صورت مرتبه همگرایی بزرگتر از 2 است $\square$

4.5 روش نیوتن

روش نیوتن یکی از سریعترین روش‌هایی است که در این درس برای ریشه‌یابی با آنها آشنا شدیم. برای این روش لازم است تقریب نزدیکی از ریشه داشته باشیم. معمولاً از روش نیوتن برای یافتن تقریبی دقیق استفاده می‌شود و نقطه شروع اولیه آن استفاده از روش‌های دیگر از جمله روش دو بخشی یا روش وتری به دست می‌آید.

4.5.1 فرمول تکراری روش نیوتن

فرمول مربوط به روش نیوتن را به دو روش می‌توان به دست آورد.

روش اول

طبق فرض تابع f و بازه $[a, b]$ که شرایط (*) در آن برقرار است را داریم. نقطه دلخواه x_0 در این بازه را در نظر می‌گیریم و خط مماس بر تابع در این نقطه را رسم می‌کنیم. محل برخورد خط مماس با محور x ها را x_1 می‌نامیم. حال

خط مماس بر تابع در نقطه x_1 را رسم می‌کنیم و این فرایند را ادامه می‌دهیم. حال در ادامه می‌خواهیم فرمولی برای این دنباله به‌دست آوریم. در مرحله اول خط مماس از نقطه $(x_0, f(x_0))$ عبور می‌کند و شیب آن برابر با $m = f'(x_0)$ می‌باشد.

پس فرمول خط مماس به سادگی به‌صورت زیر است:

$$y - f(x_0) = f'(x_0)(x - x_0)$$

و با توجه به توصیفی که برای x_1 داریم:

$$0 - f(x_0) = f'(x_0)(x_1 - x_0) \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}, \quad f'(x_0) \neq 0$$

و در حالت کلی فرمول زیر را برای روش نیوتن داریم:

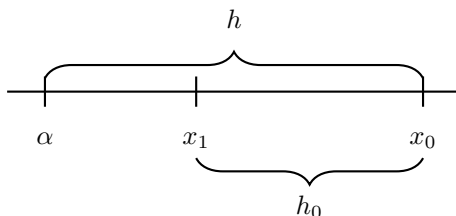
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0, \quad n = 0, 1, 2, \dots$$

روش دوم

فرض کنیم x_0 تقریبی از α باشد به طوری که اختلاف α و x_0 به اندازه h باشد.

$$\alpha = x_0 + h$$

اگر بتوانیم h را (به‌صورت تقریبی) تعیین کنیم و آن را به x_0 اضافه کنیم می‌توانیم (تقریبی از) α را به‌دست آوریم.



برای تعیین h از سری تیلور تابع استفاده می‌کنیم، داریم

$$0 = f(\alpha) = f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{f''(\eta)}{2!}h^2$$

از جمله آخر صرف نظر می‌کنیم و رابطه فوق را به‌صورت تقریبی می‌نویسیم

$$0 \simeq f(x_0) + f'(x_0)h$$

و h را به‌صورت تقریبی به‌دست می‌آوریم

$$h = -\frac{f(x_0)}{f'(x_0)}$$

حال h را به x_0 اضافه می‌کنیم و تقریب جدید یعنی x_1 را به‌دست می‌آوریم

$$x_1 = x_0 + h = x_0 - \frac{f(x_0)}{f'(x_0)}, \quad f'(x_0) \neq 0$$

و به صورت کلی فرمول زیر را داریم

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0, \quad n = 0, 1, 2, \dots$$

* مثال 4.8

الف) فرض کنیم a عددی مثبت است، فرمول حاصل از روش نیوتن برای تقریب $\sqrt{a}$ را تعیین کنید.
 ب) قرار دهید $a = 2$ و تقریبی از $\sqrt{2}$ را با فرمول بالا به دست آورید ($x_0 = 1$).

💡 حل:

الف)

$$x = \sqrt{a} \Rightarrow x^2 = a \Rightarrow x^2 - a = 0$$

$$f(x) = x^2 - a$$

$$f'(x) = 2x$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{x_n^2 + a}{2x_n} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$$

ب)

$$x_0 = 1$$

$$x_1 = \frac{1}{2} \left(1 + \frac{2}{1} \right) = 1.5$$

$$x_2 = \frac{1}{2} \left(1.5 + \frac{2}{1.5} \right) = 1.4166\dots$$

$$x_3 = \frac{1}{2} \left(1.4166 + \frac{2}{1.4166} \right) = 1.414215\dots$$

$$x_4 = \frac{1}{2} \left(1.4142 + \frac{2}{1.4142} \right) = 1.4142135624\dots$$

$$\sqrt{2} = 1.414213562\dots$$

4.5.2 < همگرایی روش نیوتن

4.1 یادآوری

تابع $h(x)$ پیوسته است.

$$h(\alpha) = 0$$

$$\lim_{x \rightarrow \alpha} h(x) = h(\alpha) = 0$$

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad |x - \alpha| < \delta \Rightarrow |h(x) - \underbrace{h(\alpha)}_{=0}| < \epsilon$$

$$x \in (\alpha - \delta, \alpha + \delta) \Rightarrow |h(x)| < \epsilon$$

در روش تکرار ساده داریم

$$x_{n+1} = g(x_n), n = 0, 1, 2, \dots$$

از طرف دیگر در روش نیوتن داریم

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0$$

بنابراین نتیجه می‌گیریم روش نیوتن حالت خاصی از روش تکرار ساده است و g متناظر آن به صورت زیر است

$$g(x) = x - \frac{f(x)}{f'(x)}$$

فرض کنیم α ریشه ساده f باشد (یعنی ریشه مرتبه اول باشد) (مشتق f در α ناصفر باشد) پس برای بحث درباره همگرایی روش نیوتن می‌توانیم از نتایج همگرایی مربوط به روش تکرار ساده استفاده کنیم.

$$g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$$

از طرفی α ریشه تابع f همان نقطه ثابت تابع g و مجهول ما می‌باشد

$$f(\alpha) = 0 \Rightarrow \alpha = ?$$

$$g(\alpha) = \alpha$$

پس داریم

$$g'(\alpha) = \frac{\overbrace{f(\alpha)f''(\alpha)}^{=0}}{\underbrace{(f'(\alpha))^2}_{\neq 0}}$$

چون α ریشه ساده f است پس $f'(\alpha) \neq 0$ بنابراین

$$g'(\alpha) = 0$$

حال چون g تابعی پیوسته است داریم

$$\forall 0 < \epsilon < 1, \quad \exists \delta > 0 \quad x \in (\alpha - \delta, \alpha + \delta) \Rightarrow |g'(x)| \leq \epsilon < 1$$

پس شرایط لازم برای همگرایی روش تکرار ساده برقرار است.
پس روش نیوتن به ازای هر نقطه شروعی که در این بازه باشد همگراست.
چون $g'(\alpha) = 0$ است مرتبه همگرایی حداقل 2 است.

$$g''(x) = \frac{(f'(x)f''(x) + f(x)f'''(x))(f'(x))^2 - (2f''(x)f'(x))(f(x)f''(x))}{(f'(x))^4}$$

$$g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$$

اگر $f''(\alpha) \neq 0$ آنگاه مرتبه همگرایی روش نیوتن دقیقاً 2 است در غیر این صورت بیشتر از 2 است.

نکته 4.2

تمام نتایج بالا را برای زمانی گفتیم که $f'(\alpha) \neq 0$ یعنی α ریشه ساده f باشد.

تعیین ریشه‌های تکراری

حال در ادامه می‌خواهیم حالتی را بررسی کنیم که α ریشه تکراری f باشد. فرض کنیم α ریشه تکراری مرتبه m تابع f باشد و $\{x_n\}$ دنباله حاصل از روش نیوتن باشد و این دنباله همگرا به ریشه باشد. در این حالت مرتبه همگرایی دنباله چقدر است؟

قضیه 4.6

اگر α ریشه مرتبه m تابع f باشد و دنباله حاصل از نیوتن همگرا باشد آنگاه

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = \frac{m-1}{m}$$

اثبات

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

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

α ریشه مرتبه m تابع f است بنابراین

$$f(x) = (x - \alpha)^m h(x), \quad h(\alpha) \neq 0$$

$$f'(x) = m(x - \alpha)^{m-1}h(x) + (x - \alpha)^m h'(x)$$

$$x_{n+1} - \alpha = (x_n - \alpha) - \frac{(x_n - \alpha)^m h(x_n)}{m(x_n - \alpha)^{m-1}h(x_n) + (x_n - \alpha)^m h'(x_n)}$$

$$x_{n+1} - \alpha = (x_n - \alpha) - \frac{(x_n - \alpha)h(x_n)}{mh(x_n) + (x_n - \alpha)h'(x_n)}$$

$$\frac{x_{n+1} - \alpha}{x_n - \alpha} = 1 - \frac{h(x_n)}{mh(x_n) + (x_n - \alpha)h'(x_n)}$$

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = 1 - \lim_{n \rightarrow \infty} \frac{h(x_n)}{mh(x_n) + (x_n - \alpha)h'(x_n)}$$

$$= 1 - \lim_{n \rightarrow \infty} \frac{h(x_n)}{mh(x_n)} = \frac{m-1}{m} \quad \square$$

نکته ۴.۳

با قضیه فوق، روش نیوتن برای محاسبه ریشه تکراری از مرتبه همگرایی ۱ است.

نکته ۴.۴

هر چه m بزرگتر باشد سرعت همگرایی نیوتن برای محاسبه ریشه تکراری کندتر است چون $\frac{m-1}{m}$ به ۱ نزدیکتر می‌شود.

نکته ۴.۵

در قضیه ثابت کردیم $\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = \frac{m-1}{m}$ اما از این حد برای تعیین m نمی‌توانیم استفاده کنیم چون α را نداریم (در واقع α مجهول مسئله است)

قضیه ۴.۷

با شرایط قضیه قبل داریم

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - x_n}{x_n - x_{n-1}} = \frac{m-1}{m}$$

اثبات ✓

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{x_{n+1} - x_n}{x_n - x_{n-1}} &= \lim_{n \rightarrow \infty} \frac{(x_{n+1} - a) - (x_n - a)}{(x_n - a) - (x_{n-1} - a)} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{x_{n+1}-a}{x_n-a} - 1}{1 - \frac{x_{n-1}-a}{x_n-a}} \\ &= \frac{\frac{m-1}{m} - 1}{1 - \frac{m}{m-1}} = \frac{m-1-m}{m} \cdot \frac{m-1}{m-1-m} = \frac{-1}{m} \cdot \frac{m-1}{-1} = \frac{m-1}{m} \quad \square\end{aligned}$$

با استفاده از جملات دنباله حاصل از نیوتن ریشه $x - \sin(x) = 0$ مرتبه همگرایی ریشه α را

* مثال 4.9

تعیین کنید.

💡 حل:

در اینجا درباره تکراری بودن α صحبتی نشده است. اگر α ریشه تکراری باشد، مرتبه همگرایی ۱ است و غیر این صورت حداقل ۲ است. ما با محاسبه چند جمله از دنباله $\{x_n\}$ مقدار $\frac{x_{n+1}-x_n}{x_n-x_{n-1}}$ را تقریب می‌زنیم و با توجه به اینکه این مقدار برابر با $\frac{m-1}{m}$ است و $m \in \mathbb{N}$ را تخمین می‌زنیم.

$$\begin{aligned}\frac{x_{n+1} - x_n}{x_n - x_{n-1}} &\simeq \frac{m-1}{m} \\ f(x) &= x - \sin(x) \\ f'(x) &= 1 - \cos(x) \\ x_{n+1} &= x_n - \frac{x_n - \sin(x_n)}{1 - \cos(x_n)} = \frac{\sin(x_n) - x_n \cos(x_n)}{1 - \cos(x_n)} \\ x_0 &= 0.5 \\ x_1 &= \frac{\sin(x_0) - x_0 \cos(x_0)}{1 - \cos(x_0)} \simeq 0.331932 \\ x_2 &= \frac{\sin(x_1) - x_1 \cos(x_1)}{1 - \cos(x_1)} \simeq 0.220880 \\ x_3 &= \frac{\sin(x_2) - x_2 \cos(x_2)}{1 - \cos(x_2)} \simeq 0.147133 \\ x_4 &= \frac{\sin(x_3) - x_3 \cos(x_3)}{1 - \cos(x_3)} \simeq 0.098054 \\ x_5 &= \frac{\sin(x_4) - x_4 \cos(x_4)}{1 - \cos(x_4)} \simeq 0.065359 \\ x_6 &= \frac{\sin(x_5) - x_5 \cos(x_5)}{1 - \cos(x_5)} \simeq 0.043569\end{aligned}$$

$$x_7 = \frac{\sin(x_6) - x_6 \cos(x_6)}{1 - \cos(x_6)} \simeq 0.029045$$

$$\frac{x_7 - x_6}{x_6 - x_5} = \frac{0.029045 - 0.043569}{0.043569 - 0.065359} \simeq \frac{-0.014524}{-0.021790} \simeq 0.666675 \simeq \frac{m-1}{m} \Rightarrow m = 3$$

در این بخش می‌خواهیم روش نیوتن اصلاح شده را معرفی کنیم. این روش برای یافتن ریشه‌های تکراری از مرتبه همگرایی حداقل 2 می‌باشد.

قضیه 4.8

اگر α ریشه تکراری از مرتبه m تابع f باشد و y_0 به اندازه کافی نزدیک ریشه انتخاب شود و $\{y_n\}$ با فرمول زیر ساخته شود

$$y_{n+1} = y_n - m \frac{f(y_n)}{f'(y_n)}$$

آنگاه $\{y_n\}$ از مرتبه حداقل 2 است.

چون α ریشه مرتبه m است داریم

اثبات ✓

$$f(x) = (x - \alpha)^m h(x), h(\alpha) \neq 0$$

$$f'(x) = m(x - \alpha)^{m-1} h(x) + (x - \alpha)^m h'(x)$$

$$\begin{aligned} y_{n+1} &= y_n - m \frac{f(y_n)}{f'(y_n)} \\ &= y_n - m \frac{(y_n - \alpha)^m h(y_n)}{m(y_n - \alpha)^{m-1} h(y_n) + (y_n - \alpha)^m h'(y_n)} \\ &= y_n - m \frac{(y_n - \alpha) h(y_n)}{m h(y_n) + (y_n - \alpha) h'(y_n)} \end{aligned}$$

$$y_{n+1} - \alpha = y_n - \alpha - m \frac{(y_n - \alpha) h(y_n)}{m h(y_n) + (y_n - \alpha) h'(y_n)}$$

$$\begin{aligned} \frac{y_{n+1} - \alpha}{y_n - \alpha} &= 1 - \frac{m h(y_n)}{m h(y_n) + (y_n - \alpha) h'(y_n)} \\ &= \frac{m h(y_n) + (y_n - \alpha) h'(y_n) - m h(y_n)}{m h(y_n) + (y_n - \alpha) h'(y_n)} \end{aligned}$$

$$\frac{y_{n+1} - \alpha}{(y_n - \alpha)^2} = \frac{h'(y_n)}{m h(y_n) + (y_n - \alpha) h'(y_n)}$$

از طرفین حد می‌گیریم

$$\lim_{n \rightarrow \infty} \frac{y_{n+1} - \alpha}{(y_n - \alpha)^2} = \frac{h'(\alpha)}{m h(\alpha)}$$

می‌دانیم $h(\alpha) \neq 0$ ، اگر $h'(\alpha) \neq 0$ آنگاه مرتبه همگرایی دقیقاً برابر 2 است در غیر این صورت بیشتر از 2 است □

* مثال 4.10

در مثال قبل مشاهده کردیم مرتبه تکرار ریشه تابع $f(x) = x - \sin x$ برابر 3 می‌باشد. حال با استفاده از فرمول بالا و با شروع از نقطه $x_0 = 0.5$ می‌خواهیم چند مرحله از روش نیوتن اصلاح شده را برای تقریب ریشه را انجام دهیم.

$$\alpha = 0, \quad m = 3, \quad x_0 = 0.5$$

$$x_{n+1} = x_n - m \frac{f(x_n)}{f'(x_n)} = x_n - 3 \frac{f(x_n)}{f'(x_n)}$$

$$f'(x) = 1 - \cos x$$

$$x_{n+1} = x_n - 3 \frac{x_n - \sin(x_n)}{1 - \cos(x_n)}$$

$$x_0 = 0.5$$

$$x_1 = 0.5 - 3 \frac{0.5 - \sin(0.5)}{1 - \cos(0.5)} = -0.042$$

$$x_2 = -0.042 - 3 \frac{-0.042 - \sin(-0.042)}{1 - \cos(-0.042)} = 5.38 \times 10^{-7}$$

فرض کنیم α ریشه مرتبه $m > 1$ تابع f می‌باشد و می‌دانیم روش نیوتن برای حل این مسئله از مرتبه 1 می‌باشد. برای اینکه روش نیوتن برای حل این مسئله از مرتبه حداقل 2 باشد، تابع $g(x)$ را معرفی می‌کنیم که α ریشه ساده تابع g باشد. آنگاه روش نیوتن معمولی را روی g پیاده‌سازی می‌کنیم. حال سوال این است g را چه تابعی در نظر بگیریم؟

$$f(x) = (x - \alpha)^m h(x), \quad h(x) \neq 0$$

α ریشه مرتبه m $f(x)$

α ریشه مرتبه $m-1$ $f'(x)$

α ریشه مرتبه $m-2$ $f''(x)$

⋮

α ریشه ساده $f^{(m-1)}(x)$

$$\Rightarrow g(x) := f^{(m-1)}(x)$$

بنابراین α ریشه ساده $f^{(m-1)}(x) = 0$ می‌باشد. حال روش نیوتن را برای این تابع پیاده‌سازی می‌کنیم:

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} = x_n - \frac{f^{(m-1)}(x_n)}{f^{(m)}(x_n)}$$

قضيه 4.9

اگر α ریشه مرتبه $m > 1$ از معادله $f(x) = 0$ باشد آنگاه α ریشه ساده $f^{(m-1)}(x) = 0$ است و بنابراین. دنباله $\{x_n\}$ حاصل از فرمول

$$x_{n+1} = x_n - \frac{f^{(m-1)}(x_n)}{f^{(m)}(x_n)}$$

دارای مرتبه همگرایی حداقل 2 است.

تقریبی از ریشه $\alpha = 0$ از تابع

* مثال 4.11

$$f(x) = x - \sin(x)$$

را با شروع از نقطه $x_0 = 0.5$ و با استفاده از فرمول بالا به دست آورید

💡 حل:

از قبل می دانیم $m = 3$ پس

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

$$f(x) = x - \sin(x)$$

$$f'(x) = 1 - \cos(x)$$

$$f''(x) = \sin(x)$$

$$f'''(x) = \cos(x)$$

$$x_{n+1} = x_n - \frac{f''(x_n)}{f'''(x_n)} = x_n - \frac{\sin(x_n)}{\cos(x_n)} = x_n - \tan(x_n)$$

$$x_0 = 0.5$$

$$x_1 = 0.004630$$

$$x_2 = 0.00003312$$

$$x_3 = -1.2 \times 10^{-14}$$

4.6 روش وترى (خط قاطع)

در روش نیوتن در هر مرحله به $f'(x_n)$ احتیاج داریم. پس باید تابع $f'(x_n)$ غیرصفر باشد. گاهی اوقات دسترسی به f' غیرممکن یا پرهزینه می باشد و یا کلاً می خواهیم از اطلاعات مشتق استفاده نکنیم. در اینجا همان فرمول روش نیوتن را در نظر می گیریم و به جای $f'(x_n)$ مقدار تقریبی زیر را قرار می دهیم.

$$f'(x_n) \simeq \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}$$

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

$$= x_n - \frac{f(x_n)(x_{n-1} - x_n)}{f(x_{n-1}) - f(x_n)}$$

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

بنابراین فرمول روش وترى به صورت زیر به دست می‌آید:

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

پس برای شروع روش وترى به دو نقطه x_0, x_1 نیاز داریم.

$$x_0, x_1 \in \underbrace{[a, b]}_{*, f}$$

$$x_0, x_1 \rightarrow x_2$$

$$x_1, x_2 \rightarrow x_3$$

$$x_2, x_3 \rightarrow x_4$$

$$\vdots$$

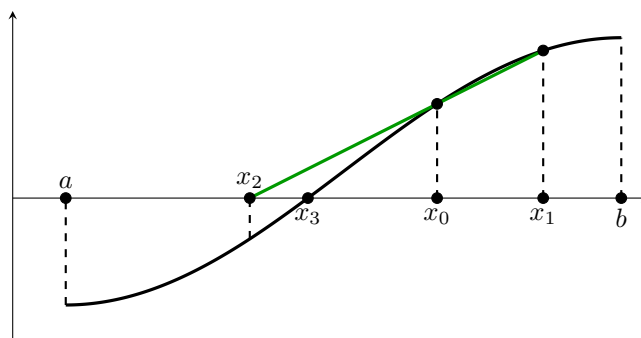
برای بازه $[a, b]$

$$x_n = \frac{bf(a) - af(b)}{f(a) - f(b)}$$

$$f(a)f(x_n) > 0 \Rightarrow [x_n, b] = [a, b]$$

$$f(a)f(x_n) < 0 \Rightarrow [a, x_n] = [a, b]$$

4.6.1 < تعبیر هندسی



فرمول روش وتری را با استفاده از تعبیر هندسی بالا نیز می‌توانیم به‌دست آوریم. تابع f و بازه $[a, b]$ که شرایط (*) در آن برقرار است را در نظر می‌گیریم دو نقطه دلخواه $x_0, x_1 \in [a, b]$ را انتخاب می‌کنیم. خط واصل $(x_0, f(x_0))$ و $(x_1, f(x_1))$ را رسم می‌کنیم. محل برخورد این خط با محور x ها را x_2 می‌نامیم. حال پاره خط $(x_1, f(x_1))$ و $(x_2, f(x_2))$ را رسم کرده و فرایند را ادامه می‌دهیم.

نکته 4.6

روش وتری کندتر از روش نیوتن است و معمولاً از روش دوبخشی و نابجایی سریعتر است.

4.7 حل دستگاه‌های غیرخطی

در این بخش می‌خواهیم درباره حل دستگاه‌های غیرخطی صحبت کنیم یعنی حل دستگاه‌هایی به فرم

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, \dots, x_n) = 0 \end{cases}$$

یعنی یک دستگاه n معادله و n مجهول را در نظر می‌گیریم که توابع درگیر در مسئله یعنی $f_1, f_2, \dots, f_n$ نسبت به متغیرهای $x_1, x_2, \dots, x_n$ غیرخطی باشند. در حالت کلی حل این مسئله به سادگی امکان‌پذیر نیست. در اینجا برای حالت دو معادله دو مجهول یک روش عددی معرفی خواهیم کرد. بنابراین دستگاه زیر را در نظر می‌گیریم

$$\begin{cases} f(x, y) = 0 \\ g(x, y) = 0 \end{cases}$$

که x, y متغیرهای مسئله هستند و f, g توابع غیرخطی نسبت به x, y هستند.

4.7.1 تعمیم روش نیوتن برای حل دستگاه غیرخطی دو معادله دو مجهول

فرض کنیم هدف تعیین یک جواب است که در دستگاه زیر صدق کند

$$\begin{cases} f(x, y) = 0 \\ g(x, y) = 0 \end{cases}$$

فرض کنیم (α, β) جواب دستگاه فوق باشد که مجهول مسئله است. فرض کنیم (x_0, y_0) تقریبی از جواب است و

$$\alpha = x_0 + h_0$$

$$\beta = y_0 + k_0$$

اگر بتوانیم h_0, k_0 را تقریب بزنیم و بعد آن‌ها را به ترتیب به x_0, y_0 اضافه کنیم، می‌توانیم تقریب‌های بهتر x_1, y_1 را

به صورت زیر به دست آوریم

$$x_1 = x_0 + h_0$$

$$y_1 = y_0 + k_0$$

یعنی در حالت کلی داریم

$$x_{n+1} = x_n + h_n$$

$$y_{n+1} = y_n + k_n$$

بسط تیلور $f(x), a$:

$$Taylor f(x), a \Rightarrow f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots$$

بسط تیلور $f(x, y), (a, b)$:

$$\nabla f(a, b) = \begin{pmatrix} \frac{\partial f}{\partial x}(a, b) \\ \frac{\partial f}{\partial y}(a, b) \end{pmatrix}$$

$$H(a, b) = \begin{pmatrix} \frac{\partial^2 f}{(\partial x)^2}(a, b) & \frac{\partial^2 f}{\partial y \partial x}(a, b) \\ \frac{\partial^2 f}{\partial x \partial y}(a, b) & \frac{\partial^2 f}{(\partial y)^2}(a, b) \end{pmatrix}$$

$$f(x, y) = f(a, b) + \nabla f(a, b)^T \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x-a \\ y-b \end{pmatrix}^T H(a, b) \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \dots$$

$$= f(a, b) + \frac{\partial f}{\partial x}(a, b)(x-a) + \frac{\partial f}{\partial y}(a, b)(y-b) + \frac{1}{2} \frac{\partial^2 f}{(\partial x)^2}(a, b)(x-a)^2 + \frac{1}{2} \frac{\partial^2 f}{(\partial y)^2}(a, b)(y-b)^2 + \frac{\partial^2 f}{\partial x \partial y}(a, b)(x-a)(y-b) + \dots$$

$$\alpha = x_0 + h_0$$

$$\beta = y_0 + k_0$$

فرض کنیم در مرحله ابتدایی هستیم و (x_0, y_0) را داریم. می‌خواهیم تقریبی برای (h_0, k_0) به دست آوریم. برای این منظور سری‌های تیلور f, g را حول نقطه (x_0, y_0) می‌نویسیم.

$$(\alpha, \beta) \Rightarrow \begin{matrix} \alpha = x_0 + h_0 & h_0 \simeq? \\ \beta = y_0 + k_0 & k_0 \simeq? \end{matrix}$$

$$\begin{aligned}
 0 &= f(\alpha, \beta) = f(x_0 + h_0, y_0 + k_0) \\
 &\simeq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x_0 + h_0 - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y_0 + k_0 - y_0)
 \end{aligned}$$

$$\begin{aligned}
 0 &= g(\alpha, \beta) = g(x_0 + h_0, y_0 + k_0) \\
 &\simeq g(x_0, y_0) + \frac{\partial g}{\partial x}(x_0, y_0)(x_0 + h_0 - x_0) + \frac{\partial g}{\partial y}(x_0, y_0)(y_0 + k_0 - y_0)
 \end{aligned}$$

بنابراین داریم

$$(x_0, y_0) \Rightarrow (h_0, k_0) = ?$$

$$\begin{cases} \frac{\partial f}{\partial x}(x_0, y_0)h_0 + \frac{\partial f}{\partial y}(x_0, y_0)k_0 = -f(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0)h_0 + \frac{\partial g}{\partial y}(x_0, y_0)k_0 = -g(x_0, y_0) \end{cases} \Rightarrow \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} h_0 \\ k_0 \end{pmatrix} = \begin{pmatrix} -f(x_0, y_0) \\ -g(x_0, y_0) \end{pmatrix}$$

به شرطی که:

$$\left(\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \right) \Big|_{(x_0, y_0)} \neq 0$$

دستگاه بالا قابل حل است و با حل آن مقدارهای تقریبی برای h_0, k_0 به دست می‌آوریم و قرار می‌دهیم.

$$\begin{cases} x_1 = x_0 + h_0 \\ y_1 = y_0 + k_0 \end{cases}$$

شرط توقف را بررسی می‌کنیم اگر برقرار بود متوقف می‌شویم در غیر اینصورت روند فوق را با (x_1, y_1) تکرار می‌کنیم و این فرآیند را ادامه می‌دهیم.

نکته 4.7

برای شرط توقف می‌توان ترکیبی از شرط‌های توقف گفته شده از قبل را در نظر گرفت به عنوان مثال:

۱. اگر $|f(x_n, y_n)| \leq \epsilon$ و $|g(x_n, y_n)| \leq \epsilon$ توقف کنید (برای $\epsilon > 0$ داده شده)
۲. اگر $|x_{n+1} - x_n| \leq \epsilon$ و $|y_{n+1} - y_n| \leq \epsilon$ توقف کنید (برای $\epsilon > 0$ داده شده)

⋮

با شروع از نقطه $(۴, ۳)$ دو تکرار از روش نیوتن تعمیم یافته برای حل دستگاه زیر را انجام

* مثال 4.12

دهید.

$$\begin{cases} x^2 - y^2 = 5 \\ x^2 + y^2 = 25 \end{cases}$$

حل:💡

$$\begin{cases} f(x, y) = x^2 - y^2 - 5 = 0 \\ g(x, y) = x^2 + y^2 - 25 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{\partial f}{\partial x}(x_n, y_n)h_n + \frac{\partial f}{\partial y}(x_n, y_n)k_n = -f(x_n, y_n) \\ \frac{\partial g}{\partial x}(x_n, y_n)h_n + \frac{\partial g}{\partial y}(x_n, y_n)k_n = -g(x_n, y_n) \end{cases}$$

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = -2y, \quad \frac{\partial g}{\partial x} = 2x, \quad \frac{\partial g}{\partial y} = 2y$$

$$\begin{cases} 2x_n h_n - 2y_n k_n = -x_n^2 + y_n^2 + 5 \\ 2x_n h_n + 2y_n k_n = -x_n^2 - y_n^2 + 25 \end{cases}$$

به شرطی که

$$4x_n y_n + 4x_n y_n = 8x_n y_n \neq 0$$

$$n = 0, (x_0, y_0) = (4, 3) \bullet$$

$$8 \times 4 \times 3 \neq 0$$

$$\begin{cases} 8h_0 - 6k_0 = -2 \\ 8h_0 + 6k_0 = 0 \end{cases}$$

$$\Rightarrow h_0 = -\frac{1}{8}, \quad k_0 = \frac{1}{6}$$

$$x_1 = x_0 + h_0 = 4 - \frac{1}{8} = 3.875$$

$$y_1 = y_0 + k_0 = 3 + \frac{1}{6} = 3.167$$

$$n = 1, (x_1, y_1) = (3.875, 3.167) \bullet$$

$$8 \times 3.875 \times 3.167 \neq 0$$

$$\begin{cases} 7.75h_1 - 6.334k_1 = -(3.875)^2 + (3.167)^2 + 5 = 0.0142 \\ 7.75h_1 + 6.334k_1 = -(3.875)^2 - (3.167)^2 + 25 = -0.045 \end{cases}$$

$$\Rightarrow h_1 = -0.0019, \quad k_1 = -0.0047$$

$$x_2 = x_1 + h_1 = 3.875 - 0.0019 = 3.8731$$

$$y_2 = y_1 + k_1 = 3.167 - 0.0047 = 3.1623$$

5

روش‌های عددی مقادیر ویژه

در این فصل می‌خواهیم روش‌های عددی برای تعیین چندجمله‌ای مشخصه یک ماتریس و روش توانی برای یافتن بزرگترین مقدار ویژه ماتریس را معرفی کنیم.

یادآوری 5.1 | مقدار ویژه، بردار ویژه

فرض کنیم A یک ماتریس $n \times n$ باشد، λ را یک مقدار ویژه برای ماتریس A گوییم هرگاه بردار غیرصفر n تایی x وجود داشته باشد بطوری که

$$Ax = \lambda x$$

در این صورت x را بردار ویژه نظیر λ گوییم.

یادآوری 5.2 | چندجمله‌ای مشخصه

$$Ax = \lambda x, \quad x \neq 0$$

$$\lambda x - Ax = 0$$

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

$$\det(\lambda I - A) = 0$$

اگر λ را مجهول فرض کنیم آنگاه $\det(\lambda I - A)$ یک چندجمله‌ای از درجه n بر اساس λ می‌باشد یعنی

$$P_n(\lambda) = \det(\lambda I - A) = \lambda^n + p_1 \lambda^{n-1} + \dots + p_{n-1} \lambda + p_n$$

این چندجمله‌ای را چندجمله‌ای مشخصه ماتریس A می‌نامیم که مقادیر ویژه ریشه‌های آن‌ها هستند.

۵.۱ ویژگی‌های مقادیر ویژه و بردارهای ویژه

۵.۱ ویژگی‌ها

۱. اگر λ یک مقدار ویژه ماتریس A باشد و x بردار ویژه نظیر آن باشد آنگاه هر ضریب غیر صفر از x نیز بردار ویژه است.
۲. اگر λ یک مقدار ویژه مختلط باشد آنگاه بردار ویژه نظیر آن نیز مختلط می‌باشد.
۳. اگر $\lambda_1, \dots, \lambda_n$ مقادیر ویژه ماتریس A باشند و $x_1, \dots, x_n$ بردارهای ویژه نظیر آن‌ها باشند آنگاه $\lambda_1^k, \dots, \lambda_n^k$ مقادیر ویژه A^k هستند با همان بردارهای ویژه.
۴. اگر λ مقدار ویژه A باشد و x بردار ویژه نظیر آن باشد و A^{-1} موجود باشد آنگاه $\frac{1}{\lambda}$ مقدار ویژه A^{-1} و x بردار ویژه آن است.
۵. اگر λ مقدار ویژه A باشد و p یک عدد حقیقی باشد آنگاه $\lambda - p$ یک مقدار ویژه $A - pI$ می‌باشد.
۶. فرض کنیم A, B معکوس پذیرند. آنگاه مقادیر ویژه AB, BA مساوی هستند.
۷. فرض کنیم A, B ماتریس‌های متشابه باشند آنگاه مقادیر ویژه آن‌ها یکسان است.

اثبات ۱

$$x \neq 0, \quad Ax = \lambda x$$

$$\alpha \in \mathbb{R}, \quad \alpha \neq 0$$

$$\xrightarrow{\times \alpha} \alpha Ax = \alpha \lambda x$$

$$A \underbrace{\alpha x}_y = \lambda \underbrace{\alpha x}_y \Rightarrow Ay = \lambda y$$

پس y یک بردار ویژه نظیر λ می‌باشد. چون α دلخواه بود پس نتیجه اثبات شد. از این ویژگی نتیجه می‌گیریم که بردار ویژه نظیر یک مقدار ویژه مشخص، یکتا نیست.

فرض کنیم λ یک مقدار ویژه مختلط باشد و x بردار ویژه نظیر آن حقیقی باشد (یعنی تمام درایه‌های آن حقیقی هستند)

اثبات ۲

$$Ax = \lambda x$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

عبارت فوق به صورت زیر قابل نوشتن است

$$\sum_{j=1}^n a_{ij}x_j = \lambda x_i, \quad i = 1, 2, \dots, n$$

بردار ویژه x یک بردار غیرصفر است پس حداقل یک اندیس $\bar{i} \in \{1, 2, \dots, n\}$ وجود دارد به طوری که

$$x_{\bar{i}} \neq 0$$

سطر مربوط به این اندیس را در $Ax = \lambda x$ در نظر می‌گیریم، داریم

$$\sum_{j=1}^n a_{\bar{i}j}x_j = \lambda x_{\bar{i}}$$

$$\Rightarrow \lambda = \frac{\sum_{j=1}^n a_{\bar{i}j}x_j}{x_{\bar{i}}} \begin{cases} a_{\bar{i}j}x_j \in \mathbb{R} \\ x_j \in \mathbb{R} \\ x_{\bar{i}} \in \mathbb{R} \end{cases} \Rightarrow \lambda \in \mathbb{R}$$

پس به تناقض رسیدیم یعنی x نمی‌تواند حقیقی باشد.

اثبات ✓ | ۳

$$\forall i = 1, 2, \dots, n$$

$$\lambda_i, x_i \Rightarrow Ax_i = \lambda_i x_i$$

$$\Rightarrow A^2 x_i = A \lambda_i x_i = \lambda_i^2 x_i$$

$$\vdots$$

$$\Rightarrow A^k x_i = \lambda_i^k x_i$$

اثبات ✓ | ۴

$$Ax = \lambda x \Rightarrow \underbrace{A^{-1}A}_I x = A^{-1}\lambda x = \lambda A^{-1}x$$

$$\stackrel{\lambda \neq 0}{\Rightarrow} \frac{1}{\lambda} x = A^{-1}x$$

اثبات | ۵

$$\begin{aligned}
 Ax &= \lambda x \\
 Ax - px &= \lambda x - px \\
 (A - pI)x &= (\lambda - p)x
 \end{aligned}$$

پس $\lambda - p$ مقدار ویژه $A - pI$ است

اثبات | ۶

فرض کنیم λ یک مقدار ویژه AB باشد، داریم

$$\begin{aligned}
 ABx &= \lambda x \\
 \xrightarrow{B \times} BABx &= \lambda Bx
 \end{aligned}$$

چون B وارون‌پذیر و $x \neq 0$:

$$\xrightarrow{Bx=y} BAy = \lambda y$$

مقدار λ مقدار ویژه نظیر BA می‌باشد.

اثبات | ۷

چون A, B متشابه هستند یک ماتریس معکوس پذیر P وجود دارد به طوری که

$$P^{-1}AP = B \Rightarrow PBP^{-1} = A$$

و فرض کنیم λ یک مقدار ویژه A است بنابر این

$$\begin{aligned}
 \exists x \neq 0 \Rightarrow Ax &= \lambda x \\
 PBP^{-1}x &= \lambda x \\
 \xrightarrow{P^{-1} \times} BP^{-1}x &= \lambda P^{-1}x \\
 B \underbrace{(P^{-1}x)}_y &= \lambda \underbrace{(P^{-1}x)}_y \\
 By &= \lambda y, \quad y \neq 0
 \end{aligned}$$

قضیه 5.1

مقادیر ویژه ماتریس هرمیتی، حقیقی می‌باشند.

اثبات ✓

فرض کنیم λ یک مقدار ویژه ماتریس A باشد و A یک ماتریس هرمیتی باشد. طبق تعریف داریم:

$$\exists x \neq 0 \quad Ax = \lambda x$$

$$(Ax)^* = (\lambda x)^*$$

$$x^* A^* = x^* \bar{\lambda} = \bar{\lambda} x^*$$

$$x^* A = \bar{\lambda} x^*$$

$$x^* Ax = \bar{\lambda} x^* x \Rightarrow x^* \lambda x = \bar{\lambda} x^* x$$

$$\lambda x^* x = \bar{\lambda} x^* x \Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}$$

قضیه 5.2

فرض کنیم A یک ماتریس هرمیتی باشد، نشان دهید بردارهای ویژه نظیر مقادیر ویژه متمایز، متعامدند.

اثبات ✓

فرض کنیم λ_1, λ_2 دو مقدار ویژه متمایز A باشند یعنی $\lambda_1 \neq \lambda_2$ و $x \neq 0$ بردار ویژه نظیر λ_1 و $y \neq 0$

بردار ویژه نظیر λ_2 باشد

$$Ax = \lambda_1 x \Rightarrow x^* A^* = x^* A = \lambda_1 x^*$$

$$x^* \underbrace{Ay}_{\lambda_2 y} = \lambda_1 x^* y$$

$$\Rightarrow \lambda_2 x^* y = \lambda_1 x^* y$$

$$\Rightarrow \underbrace{(\lambda_2 - \lambda_1)}_{\neq 0} x^* y = 0$$

$$\Rightarrow x^* y = 0 \Rightarrow x \perp y \quad \square$$

قضیه 5.3

بردارهای ویژه مقادیر ویژه متمایز، مستقل خطی‌اند.

✓ اثبات

فرض کنیم بردارهای ویژه نظیر مقادیر ویژه متمایز A مستقل خطی نباشند. فرض می‌کنیم r بزرگترین عددی باشد که بردارهای ویژه A با این تعداد مستقل خطی هستند یعنی مجموعه

$$\{x_1, x_2, \dots, x_r\}$$

مستقل خطی است و x_i ها بردار ویژه نظیر $\lambda_1, \dots, \lambda_r$ از A هستند و λ_i ها متمایزند حال مقدار ویژه λ_{r+1} را در نظر می‌گیریم و طبق تعریف r می‌دانیم مجموعه

$$\{x_1, x_2, \dots, x_r, x_{r+1}\}$$

وابسته خطی است پس ضرایب $c_1, c_2, \dots, c_{r+1}$ که همگی همزمان صفر نیستند وجود دارند به طوری که

$$c_1 x_1 + c_2 x_2 + \dots + c_r x_r + c_{r+1} x_{r+1} = 0 \quad (I)$$

طرفین رابطه (I) را از سمت چپ در A ضرب می‌کنیم، داریم

$$c_1 A x_1 + c_2 A x_2 + \dots + c_r A x_r + c_{r+1} A x_{r+1} = 0$$

$$c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \dots + c_r \lambda_r x_r + c_{r+1} \lambda_{r+1} x_{r+1} = 0 \quad (II)$$

طرفین رابطه (I) را در λ_{r+1} ضرب می‌کنیم، داریم

$$c_1 \lambda_{r+1} x_1 + c_2 \lambda_{r+1} x_2 + \dots + c_r \lambda_{r+1} x_r + c_{r+1} \lambda_{r+1} x_{r+1} = 0 \quad (III)$$

رابطه (II) و (III) را از هم کم می‌کنیم

$$c_1 (\lambda_1 - \lambda_{r+1}) x_1 + c_2 (\lambda_2 - \lambda_{r+1}) x_2 + \dots + c_r (\lambda_r - \lambda_{r+1}) x_r = 0$$

چون $x_1, x_2, \dots, x_r$ مستقل خطی اند و $\lambda_1, \lambda_2, \dots, \lambda_r, \lambda_{r+1}$ متمایز هستند:

$$c_1 = c_2 = \dots = c_r = 0$$

از رابطه (I) داریم

$$c_{r+1} \underbrace{x_{r+1}}_{\neq 0} = 0 \Rightarrow c_{r+1} = 0$$

پس

$$c_1 = c_2 = \dots = c_r = c_{r+1} = 0$$

که با فرض ما تناقض دارد $\square$

قضیه 5.4 | کیلی-همیلتون

فرض کنیم A یک ماتریس n بار n باشد و $P_n(\lambda)$ چندجمله‌ای مشخصه (چندجمله‌ای سرشت نما) A باشد آنگاه $P_n(A) = 0$.

فرض کنیم $p_n \neq 0$ آنگاه

$$P_n(A) = A^n + p_1 A^{n-1} + \dots + p_{n-1} A + p_n I = 0$$

$$\begin{aligned} A(A^{n-1} + p_1 A^{n-2} + \dots + p_{n-1} A) &= -p_n I \\ \Rightarrow -\frac{1}{p_n}(A^{n-1} + p_1 A^{n-2} + \dots + p_{n-1} A) &= A^{-1} \end{aligned}$$

$$\lambda = 0 \Rightarrow P_n(0) = p_n = \det(0 - A) = \det(-A) = (-1)^n \det(A)$$

قضیه 5.5 | گرشگورین

فرض کنیم A یک ماتریس $n \times n$ باشد و c_i ها دایره‌هایی به مرکز a_{ii} و به شعاع r_i باشند. آنگاه مقادیر ویژه A متعلق به C هستند.

$$\begin{aligned} r_i &= \sum_{j=1, j \neq i}^n |a_{ij}| \\ C &= \bigcup_{i=1}^n c_i \end{aligned}$$

* مثال 5.1

$$A = \begin{pmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{pmatrix}$$

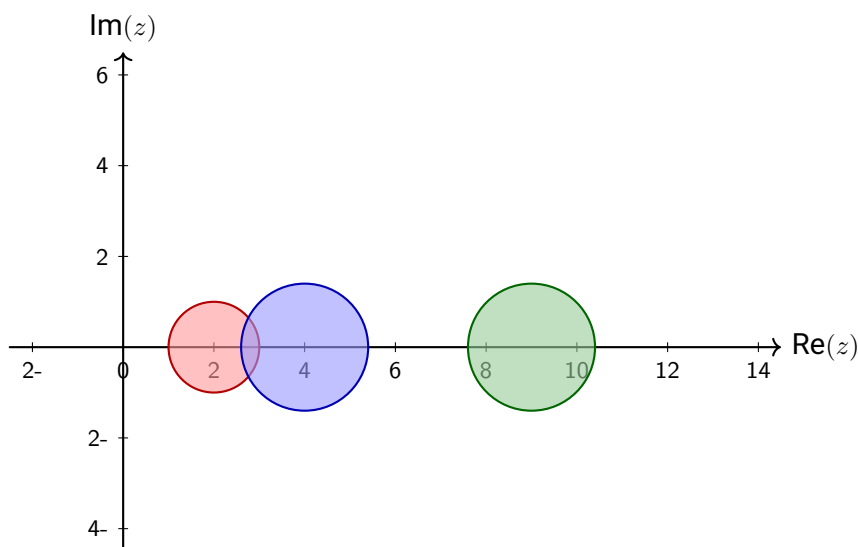
حل: 

$$c_1 = \{z \in \mathbb{C} : |z - 4| \leq 2\}$$

$$c_2 = \{z \in \mathbb{C} : |z - 2| \leq 1\}$$

$$c_3 = \{z \in \mathbb{C} : |z - 9| \leq 2\}$$

محور افق بخش حقیقی، و محور عمود بخش موهومی



مقادیر ویژه در اجتماع این دایره‌ها قرار دارند.
(ولی ممکن هست که در یک c_i مقدار ویژه‌ای موجود نباشد).

فرض کنیم λ یک مقدار ویژه ماتریس A باشد بنابراین بردار ویژه غیرصفر x وجود دارد

✓ اثبات | گرشگورین

به طوری که

$$Ax = \lambda x$$

بدون کم شدن از کلیت مسئله فرض می‌کنیم $\|x\|_\infty = 1$ طبق تعریف نرم بینهایت داریم

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

$$\exists \hat{i} \in \{1, 2, \dots, n\} \quad \text{s.t.} \quad \|x\|_\infty = |x_{\hat{i}}|$$

حال سطر $\hat{i}$ در رابطه

$$Ax = \lambda x$$

در نظر می‌گیریم

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\hat{i}1} & a_{\hat{i}2} & \dots & a_{\hat{i}n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\hat{i}} \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\hat{i}} \\ \vdots \\ x_n \end{pmatrix}$$

$$\begin{aligned}
\sum_{j=1}^n a_{ij} x_j &= \lambda x_i \\
\sum_{j=1, j \neq i}^n a_{ij} x_j + a_{ii} x_i &= \lambda x_i \\
\Rightarrow (\lambda - a_{ii}) x_i &= \sum_{j=1, j \neq i}^n a_{ij} x_j \\
|(\lambda - a_{ii}) x_i| &= \left| \sum_{j=1, j \neq i}^n a_{ij} x_j \right| \\
|(\lambda - a_{ii})| \underbrace{|x_i|}_{= \|x\|_\infty = 1} &= \left| \sum_{j=1, j \neq i}^n a_{ij} x_j \right| \\
|(\lambda - a_{ii})| &= \left| \sum_{j=1, j \neq i}^n a_{ij} x_j \right| \\
&\leq \sum_{j=1, j \neq i}^n |a_{ij} x_j| \\
&= \sum_{j=1, j \neq i}^n |a_{ij}| |x_j| \\
&\leq \sum_{j=1, j \neq i}^n |a_{ij}| \|x\|_\infty = \sum_{j=1, j \neq i}^n |a_{ij}|
\end{aligned}$$

بنابراین داریم

$$|(\lambda - a_{ii})| \leq \sum_{j=1, j \neq i}^n |a_{ij}| = r_i \quad \square$$

5.2 تعیین چندجمله‌ای سرشت نما

5.2.1 روش کریلف

فرض کنیم A یک ماتریس $n \times n$ باشد. در حالت کلی برای محاسبه چندجمله‌ای مشخصه باید دترمینان زیر را محاسبه کنیم:

$$P_n(\lambda) = \det(\lambda I - A)$$

و با توجه به تعریف این چندجمله‌ای می‌توانیم فرم زیر را برای آن در نظر بگیریم:

$$P_n(\lambda) = \lambda^n + p_1 \lambda^{n-1} + \dots + p_{n-1} \lambda + p_n$$

اگر بتوانیم ضرایب $p_1, p_2, \dots, p_n$ را مشخص کنیم، چندجمله‌ای موردنظر به دست می‌آید. فرض کنیم $y^{(0)} \in \mathbb{R}^n$ یک بردار دلخواه غیرصفر باشد. از طرفی طبق قضیه کیلی-همیلتون داریم:

$$P_n(A) = 0$$

$$A^n + p_1 A^{n-1} + \dots + p_{n-1} A + p_n I = 0$$

عبارت فوق را از سمت راست دو بردار $y^{(0)}$ ضرب می‌کنیم. داریم:

$$A^n y^{(0)} + p_1 A^{n-1} y^{(0)} + \dots + p_{n-1} A y^{(0)} + p_n y^{(0)} = 0$$

$$p_1 A^{n-1} y^{(0)} + \dots + p_{n-1} A y^{(0)} + p_n y^{(0)} = -A^n y^{(0)}$$

$$\begin{pmatrix} A^{n-1} y^{(0)} & A^{n-2} y^{(0)} & \dots & A y^{(0)} & y^{(0)} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = -A^n y^{(0)}$$

$$\begin{pmatrix} y^{(n-1)} & y^{(n-2)} & \dots & y^{(1)} & y^{(0)} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = -y^{(n)}$$

با حل دستگاه خطی $n \times n$ فوق مقادیر $p_1, p_2, \dots, p_n$ به دست می‌آیند و چندجمله‌ای مشخصه حاصل می‌شود. این روش برای تعیین چندجمله‌ای مشخصه را روش **کریلف** گوئیم. اگر $y^{(0)}$ بردار ویژه باشد آنگاه

$$A y^{(0)} = \lambda y^{(0)}$$

و بنابراین تمام ستون‌ها ضربی از ستون آخر می‌باشند. و این به این معناست که تمام ستون‌ها وابسته خطی می‌شوند و ماتریس ضرایب فوق معکوس‌پذیر نیست و نمی‌توان جواب دستگاه را به دست آورد. پس باید بردار اولیه را تغییر دهیم.

با روش کریلف چندجمله‌ای مشخصه ماتریس زیر را تعیین کنید.

* مثال 5.2

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix} \quad y^{(0)} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

💡 حل:

$$y^{(1)} = Ay^{(0)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \\ 3 \end{pmatrix}$$

$$y^{(2)} = A^2y^{(0)} = Ay^{(1)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 16 \\ 9 \\ 2 \\ -3 \end{pmatrix}$$

$$y^{(3)} = A^3y^{(0)} = Ay^{(2)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 16 \\ 9 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 28 \\ 36 \\ 62 \\ 90 \end{pmatrix}$$

$$y^{(4)} = A^4y^{(0)} = Ay^{(3)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 28 \\ 36 \\ 62 \\ 90 \end{pmatrix} = \begin{pmatrix} 646 \\ 486 \\ 398 \\ 372 \end{pmatrix}$$

$$\begin{pmatrix} 28 & 16 & -2 & 1 \\ 36 & 9 & 0 & 0 \\ 62 & 2 & 2 & -1 \\ 90 & -3 & 3 & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \begin{pmatrix} -646 \\ -486 \\ -398 \\ -372 \end{pmatrix}$$

با حل دستگاه فوق $p_1, \dots, p_4$ را محاسبه می‌کنیم

$$p_1 = -4, \quad p_2 = -40, \quad p_3 = -56, \quad p_4 = -20$$

$$P_4(\lambda) = \lambda^4 - 4\lambda^3 - 40\lambda^2 - 56\lambda - 20$$

◀ 5.2.2 روش لوبریه

این روش بر پایه قضیه نیوتن زیر است.

قضیه 5.6

اگر $\alpha_1, \alpha_2, \dots, \alpha_n$ ریشه‌های

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$$

باشند و تعریف کنیم

$$s_k = \alpha_1^k + \alpha_2^k + \dots + \alpha_n^k$$

آنگاه

$$s_k + a_1 s_{k-1} + \dots + a_{k-1} s_1 = -k a_k, \quad k = 1, 2, \dots, n \quad (*)$$

$$k = 1 \Rightarrow s_1 = -a_1 \Rightarrow a_1 = -s_1$$

$$k = 2 \Rightarrow s_2 + a_1 s_1 = -2a_2 \Rightarrow a_2 = -\frac{1}{2}(s_2 + a_1 s_1)$$

$\vdots$

حال می‌دانیم $\lambda_1, \dots, \lambda_n$ ریشه‌های چندجمله‌ای مشخصه

$$P_n(\lambda) = \lambda^n + p_1 \lambda^{n-1} + \dots + p_{n-1} \lambda + p_n$$

می‌باشد.

با توجه به قضیه فوق تعریف می‌کنیم

$$s_k = \lambda_1^k + \lambda_2^k + \dots + \lambda_n^k$$

و بعد از روی معادلات $(*)$ می‌توانیم بنویسیم

$$s_k + p_1 s_{k-1} + \dots + p_{k-1} s_1 = -k p_k, \quad k = 1, 2, \dots, n$$

و به این ترتیب $p_1, p_2, \dots, p_n$ را مشخص کنیم.

حال سوال این است بدون داشتن مقادیر ویژه چگونه مجموع آن‌ها را حساب کنیم؟ که پاسخ سوال ساده است

چون می‌دانیم

$$s_1 = \text{trace}(A) = \sum_{i=1}^n \lambda_i$$

$$s_2 = \text{trace}(A^2) = \sum_{i=1}^n \lambda_i^2$$

$\vdots$

$$s_k = \text{trace}(A^k) = \sum_{i=1}^n \lambda_i^k$$

* مثال 5.3

چند جمله ای مشخصه ماتریس زیر را به دست آورید.

$$A = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ -7 & 2 & 3 \end{pmatrix}$$

💡 حل:

$$s_k + p_1 s_{k-1} + \dots + p_{k-1} s_1 = -k p_k, \quad k = 1, 2, \dots, n$$

$$s_1 = \text{trace}(A) = 2 + 1 + 3 = 6$$

$$s_1 = -p_1 \Rightarrow p_1 = -6$$

$$A^2 = \begin{pmatrix} 8 & 6 & 2 \\ -1 & 7 & 4 \\ -31 & -6 & 11 \end{pmatrix}$$

$$s_2 = \text{trace}(A^2) = 8 + 7 + 11 = 26$$

$$s_2 + p_1 s_1 = -2p_2 \Rightarrow p_2 = 5$$

$$A^3 = \begin{pmatrix} 14 & 26 & 12 \\ -16 & 13 & 19 \\ -151 & -46 & 27 \end{pmatrix}$$

$$s_3 = \text{trace}(A^3) = 14 + 13 + 27 = 54$$

$$s_3 + p_1 s_2 + p_2 s_1 = -3p_3 \Rightarrow p_3 = 24$$

$$\begin{aligned} P_3(\lambda) &= \lambda^3 + p_1 \lambda^2 + p_2 \lambda + p_3 \\ &= \lambda^3 - 6\lambda^2 + 5\lambda + 24 \end{aligned}$$

5.3 محاسبه بزرگترین مقدار ویژه از لحاظ قدرمطلق

5.3.1 روش توانی

روش توانی برای محاسبه بزرگترین مقدار ویژه یک ماتریس از لحاظ قدرمطلق با این روش می‌توانیم بزرگترین مقدار ویژه یک ماتریس از لحاظ قدرمطلق را به صورت عددی به دست آوریم.

قضیه 5.7

فرض کنیم A یک ماتریس $n \times n$ حقیقی باشد. مقدارهای ویژه آن $\lambda_1, \lambda_2, \dots, \lambda_n$ باشند به طوری که

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

و بردارهای ویژه نظیر آن‌ها یعنی $x_1, x_2, \dots, x_n$ مستقل خطی باشند. اگر $u_0 \in \mathbb{R}^n$ بردار دلخواه غیرصفر باشد و دنباله $\{u_r\}$ را به صورت زیر بسازیم

$$u_{r+1} = Au_r$$

آنگاه

$$|\lambda_1| = \lim_{r \rightarrow \infty} \left| \frac{(u_{r+1})_j}{(u_r)_j} \right|$$

که $(u_r)_j$ مولفه j ام بردار u_r می‌باشد. همچنین داریم $\lim_{r \rightarrow \infty} u_r$ به ضریبی از x_1 همگرا شود.

اثبات

چون $x_1, x_2, \dots, x_n$ مستقل خطی هستند، ضرایب $\alpha_1, \alpha_2, \dots, \alpha_n$ وجود دارند به طوری که

$$u_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

و فرض می‌کنیم $\alpha_1 \neq 0$ (اگر $\alpha_1 = 0$ باشد با توجه به اینکه u_0 دلخواه است u_0 دیگری انتخاب می‌کنیم که این شرط برقرار باشد).

$$u_1 = Au_0 = A \left(\sum_{i=1}^n \alpha_i x_i \right) = \sum_{i=1}^n \alpha_i (Ax_i) = \sum_{i=1}^n \alpha_i \lambda_i x_i$$

$$u_2 = Au_1 = A^2 u_0 = \sum_{i=1}^n \alpha_i \lambda_i^2 x_i$$

و در حالت کلی داریم

$$u_r = Au_{r-1} = \sum_{i=1}^n \alpha_i \lambda_i^r x_i$$

$$\frac{(u_{r+1})_j}{(u_r)_j} = \frac{(\sum_{i=1}^n \alpha_i \lambda_i^{r+1} x_i)_j}{(\sum_{i=1}^n \alpha_i \lambda_i^r x_i)_j}$$

$$\left| \frac{(u_{r+1})_j}{(u_r)_j} \right| = \left| \frac{(\alpha_1 \lambda_1^{r+1} x_1 + \alpha_2 \lambda_2^{r+1} + \dots + \alpha_n \lambda_n^{r+1} x_n)_j}{(\alpha_1 \lambda_1^r x_1 + \alpha_2 \lambda_2^r + \dots + \alpha_n \lambda_n^r x_n)_j} \right| = \left| \frac{\lambda_1 (\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^{r+1} x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^{r+1} x_n)_j}{(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^r x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^r x_n)_j} \right|$$

$$\lim_{r \rightarrow \infty} \frac{(u_{r+1})_j}{(u_r)_j} = \left| \lambda_1 \frac{\alpha_1 x_1}{\alpha_1 x_1} \right| = |\lambda_1|$$

$$u_r \Rightarrow \bar{u}_r = \frac{u_r}{\lambda_1^r}$$

$$\lim_{r \rightarrow \infty} \bar{u}_r = \lim_{r \rightarrow \infty} \frac{u_r}{\lambda_1^r} = \lim_{r \rightarrow \infty} \frac{\lambda_1^r}{\lambda_1^r} (\alpha_1 x_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^r x_i) = \alpha_1 x_1 \quad \square$$

نکته 5.1

چون در عمل ممکن است مولفه های u_r خیلی ضعیف رشد کنند، در هر مرحله این بردار را به نرمش تقسیم می کنیم.

فرض کنیم $u_0 \in \mathbb{R}^n$ به دلخواه انتخاب شده باشد و $u_1 = Au_0$.

❖ الگوریتم | روش توانی

For $r = 1, 2, \dots$ do:

$$\mu_r = \|u_r\|_{\infty}$$

$$w_r = \frac{u_r}{\mu_r}$$

$$u_{r+1} = Aw_r$$

End

در این صورت $w_r \rightarrow x_1$ و $\mu_r \rightarrow |\lambda_1|$.

در الگوریتم فوق به شرط توقف نیز احتیاج داریم که می توان شرط های توقف مختلفی در نظر گرفت. به عنوان مثال:

$$|\mu_{r+1} - \mu_r| < \epsilon \cdot$$

$$\|u_{r+1} - u_r\|_{\infty} < \epsilon \cdot$$