چاپ دوم

جهاد دانشگاهی

جبر خطی عددی

تألیف:

دکتر داود خجسته سالکویه

استاد دانشکده علوم ریاضی دانشگاه گیلان

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

جبر خطی عددی

مؤلف :

دکتر داود خجسته‌سالکویه

استاد دانشکده علوم ریاضی دانشگاه گیلان

سرشناسنامه	: خجسته‌سالکویه، داود، ۱۳۴۹
عنوان و نام پدیدآور	: جبر خطی عددی/ مولف داود خجسته سالکویه.
مشخصات نشر	: رشت: جهاد دانشگاهی، واحد گیلان، ۱۳۹۵.
مشخصات ظاهری	: ۲۵۲ ص.
شابک	: ۲–۰۳–۶۱۸۳–۹۶۴–۹۷۸
وضعیت فهرست نویسی	: فیپا
یادداشت	: کتابنامه: ص. ۲۲۳–۲۲۴.
موضوع	: جبر خطی – راهنمای آموزشی (عالی)
موضوع	: Linear algebra - Study and teaching (Higher)
موضوع	: جبر خطی – مسائل، تمرین‌ها و غیره (عالی)
موضوع	: Linera algebra - Problems, exercises, etc. (Higher)
موضوع	: آنالیز عددی – راهنمای آموزشی (عالی)
موضوع	: Numerical analysis - Study and teaching
موضوع	: آنالیز عددی – مسائل، تمرین‌ها و غیره (عالی)
موضوع	: Numerical analysis - Problems, exercises, etc. (Higher)
شناسه افزوده	: جهاد دانشگاهی. واحد گیلان
رده‌بندی کنگره	: ۱۳۹۵ ۲ج۳خ/۲/QA۱۸۴
رده‌بندی دیویی	: ۵۱۲/۵
شماره کتابشناسی ملی	: ۴۳۴۶۲۵۹

شناسنامه کتاب:

نام کتاب : جبر خطی عددی

مولف : دکتر داود خجسته سالکویه

ناشر: انتشارات جهاد دانشگاهی استان گیلان

شابک : ۲–۰۳–۶۱۸۳–۹۶۴–۹۷۸

نوبت چاپ : دوم– ۱۳۹۷

قیمت : ۱۷۰۰۰۰ ریال

تقدیم به همسرم

و

به دخترانم فاطمه و فائزه

پیش‌گفتار

بررسی بسیاری از مباحث در شاخه‌های مختلف علوم و مهندسی منجر به حل مسائل مختلف در جبرخطی با ابعاد بزرگ می‌شود. از مهمترین مسائل حل دستگاه معادلات خطی، مسالهٔ مقدار ویژه و مسالهٔ کمترین توان‌های دوم می‌باشند. برای حل این گونه مسائل در ابعاد بزرگ استفاده از رایانه ضروری به نظر می‌رسد. در حل یک مسأله توسط رایانه دو موضوع اساسی می‌بایستی مورد توجه قرارگیرد: استفاده بهینه از حافظه کامپیوتر و کنترل خطاهای ناشی از گردکردن اعداد. برای دست‌یابی به این دو هدف لازم است که الگوریتم‌های مناسب برای حل مسائل جبرخطی ارائه نمود. در این کتاب گزیده‌ای از روش‌ها و الگوریتم‌های عددی برای سه مسألهٔ اشاره شده در جبرخطی ارائه می‌گردد.

این کتاب مطابق با سرفصل درس "جبرخطی عددی" دوره کارشناسی تهیه و تنظیم شده و مشتمل بر شش فصل است. در فصل اول این کتاب سعی شده است که مروری بر پیش‌نیازها ارائه گردد. فصل دوم به مباحث مربوط به ضرب داخلی و نرم اختصاص داده شده است. در فصل سوم چند الگوریتم مربوط به حل دستگاه معادلات خطی معرفی شده است. در فصل چهارم ابتدا مسألهٔ مقدار ویژه معرفی، مبانی نظری آن مرور و سپس چند روش عددی برای محاسبهٔ تقریبی مقدار ویژه ارائه می‌شود. فصل پنجم به معرفی چند روش تکراری ایستا برای حل دستگاه معادلات خطی اختصاص داده شده است و سرانجام در فصل ششم به مسألهٔ کمترین توان‌های دوم پرداخته شده است.

در این کتاب سعی شده است که مطالب به صورت ساده و روان ارائه گردد و برای هر مورد مثال‌هایی آورده شود. همچنین برای درک بهتر بعضی از مطالب از ارائهٔ شکل‌های مرتبط با آن استفاده شده است. در پایان هر فصل تعدادی تمرین مطابق با مباحث همان فصل گنجانده شده است که حل آنها می‌تواند به دانشجو در درک مفاهیم

همان فصل کمک شایانی نماید. برای درک بهتر مطالب هر فصل، به دانشجویان گرامی پیشنهاد می‌شود که برای الگوریتم‌های ارائه شده در هر فصل برنامه‌ای با زبان متلب (MATLAB) یا هر زبان دیگر تهیه نموده و با مثال‌های مختلف مورد آزمایش قرار دهند.

بدیهی است تذکرات خوانندگان گرامی در مورد لغزش‌های احتمالی موجب تعالی خواهد بود و لذا پیشاپیش از دریافت نظرات و پیشنهادات این عزیزان سپاسگزاری می‌کنم.

لازم می‌دانم از تمامی افرادی که بنده را در تهیه این کتاب یاری نموده‌اند، قدردانی کنم. از جناب آقای دکتر حسین امینی خواه به خاطر نظرات اصلاحی و پیشنهادات‌شان سپاسگزاری می‌نمایم. از جناب آقای محمد رزاقی سیاهرودی که طرح روی جلد کتاب را طراحی نموده‌اند، تشکر می‌نمایم. از انتشارات جهاد دانشگاهی گیلان به دلیل پذیرش چاپ این کتاب سپاسگزاری می‌نمایم. در پایان از خانواده‌ام که در طول تهیهٔ این کتاب همواره از حمایت‌های آنها برخوردار بوده‌ام، تشکر و قدردانی می‌نمایم.

داود خجسته سالکویه

دانشکده علوم ریاضی دانشگاه گیلان

مرداد ۱۳۹۵

فهرست مندرجات

فصل ۱

مروری بر پیش‌نیازها

در این فصل، ابتدا بعضی از مفاهیم، تعاریف و قضایای مورد نیاز در فصل‌های بعدی را مرور می‌کنیم. سپس دو مدل ریاضی، که حل آنها منجر به دو مسألهٔ اصلی در جبرخطی می‌شود را بررسی می‌کنیم.

۱.۱ مفاهیم اولیه

در سراسر این کتاب از نمادهای $\mathbb{R}$ و $\mathbb{C}$ به ترتیب برای نمایش مجموعهٔ اعداد حقیقی و مجموعهٔ اعداد مختلط استفاده می‌شوند. برای عدد مختلط z، قسمت حقیقی و مختلط آن را به ترتیب با $\Re(z)$ و $\Im(z)$ نمایش می‌دهیم. مجموعهٔ $\mathbb{C}^n$ را به صورت

$$\mathbb{C}^n = \left\{ \begin{pmatrix} x_۱ \\ x_۲ \\ \vdots \\ x_n \end{pmatrix} : \ x_i \in \mathbb{C}, \ i = ۱, ۲, \ldots, n \right\},$$

نشان می‌دهیم. به همین ترتیب مجموعهٔ $\mathbb{R}^n$ تعریف می‌شود. مجموعهٔ ماتریس‌های $m \times n$ با درایه‌های مختلط (حقیقی) را با $\mathbb{C}^{m\times n}$ ($\mathbb{R}^{m\times n}$) نشان می‌دهیم. از اینجا به

بعد، اگر درایه‌هایی از یک ماتریس نمایش داده نشوند، به این معنی است که این درایه‌ها برابر با صفر هستند. فرض کنید $A = (a_{ij})$ یک ماتریس $n \times n$ باشد. در این صورت، ماتریس A را

- قطری می‌گویند، هرگاه به ازای $i \neq j$، $a_{ij} = ۰$. یک ماتریس قطری معمولاً با $\mathrm{diag}_n(a_{۱۱}, a_{۲۲}, \ldots, a_{nn})$ نشان داده می‌شود؛
- همانی می‌گویند هرگاه $A = \mathrm{diag}_n(۱, ۱, \ldots, ۱)$. یک ماتریس همانی معمولاً با I یا I_n نشان داده می‌شود؛
- بالامثلثی می‌گویند، هرگاه به ازای هر $i > j$، $a_{ij} = ۰$؛
- پایین‌مثلثی می‌گویند، هرگاه به ازای هر $i < j$، $a_{ij} = ۰$؛
- سه‌قطری می‌گویند، هرگاه به ازای هر $|i - j| > ۱$، $a_{ij} = ۰$. ماتریس سه‌قطری

$$T = \begin{pmatrix} b & c & & & \\ a & b & c & & \\ & \ddots & \ddots & \ddots & \\ & & a & b & c \\ & & & a & b \end{pmatrix} \in \mathbb{C}^{n \times n},$$

را با $T = \mathrm{tridiag}_n(a, b, c)$ نمایش می‌دهیم؛

- بالاهسنبرگی می‌گویند، هرگاه به ازای هر $i - j > ۱$، $a_{ij} = ۰$. در واقع یک ماتریس بالاهسنبرگی به شکل

$$A = \begin{pmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱,n-۱} & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲,n-۱} & a_{۲n} \\ & \ddots & \ddots & \vdots & \vdots \\ & & a_{n-۱,n-۲} & a_{n-۱,n-۱} & a_{n-۱,n} \\ & & & a_{n,n-۱} & a_{nn} \end{pmatrix},$$

است؛

- پایین‌هسنبرگی می‌گویند، هرگاه به ازای هر $j - i > ۱$، $a_{ij} = ۰$.

• ماتریس $A \in \mathbb{C}^{n\times n}$ را نیمه‌بالامثلثی می‌گوییم هرگاه به صورت

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ & A_{22} & \cdots & A_{2p} \\ & & \ddots & \vdots \\ & & & A_{pp} \end{pmatrix},$$

باشد که در آن A_{ii}، $i = 1,2,\dots,p$ ماتریس‌های ۱ × ۱ یا ۲ × ۲ هستند.

برای مثال فرض کنید

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

در این صورت، ماتریس A یک ماتریس بالامثلثی و ماتریس B یک ماتریس بالاهسنبرگی است. همچنین ماتریس

$$C = \begin{pmatrix} -1 & 3 & 5 & 1 & 3 \\ 0 & 3 & 2 & 2 & 4 \\ 0 & -2 & 3 & 3 & 1 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix},$$

یک ماتریس نیمه‌بالامثلثی است.

ترانهادهٔ ماتریس $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ را با A^T نشان می‌دهیم و درایهٔ (i,j) ام آن برابر با درایهٔ (j,i) ام ماتریس A است، یعنی $(A^T)_{ij} = a_{ji}$. ماتریس A را متقارن می‌گوییم هرگاه $A^T = A$ و متقارن‌کج می‌گوییم هرگاه $A^T = -A$. برای مثال، ماتریس‌های A و B که به صورت

$$A = \begin{pmatrix} 5 & 2 & -1 \\ 2 & 6 & 1 \\ -1 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix},$$

تعریف می‌شوند، به ترتیب متقارن و متقارن‌کج هستند.

فرض کنید $A = (a_{ij}) \in \mathbb{C}^{n\times n}$. در این صورت، ترانهادهٔ هرمیتی آن را با A^H نشان داده و درایهٔ (i,j) ام آن برابر با مزدوج درایهٔ (j,i) ام ماتریس A است، یعنی $(A^H)_{ij} = \overline{a_{ji}}$. یادآوری می‌کنیم که اگر $z = \alpha+i\beta$ یک عدد مختلط باشد، آنگاه مزدوج آن به صورت $\overline{z} = \alpha - i\beta$ تعریف می‌شود که در آن $i = \sqrt{-۱}$ واحد موهومی است. ماتریس A را هرمیتی می‌گوییم هرگاه $A^H = A$ و هرمیتی‌کج گوییم هرگاه $A^H = -A$. برای مثال، ماتریس‌های A و B که به صورت

$$A = \begin{pmatrix} ۵ & ۲-۳i & ۳+۲i \\ ۲+۳i & ۶ & ۱-i \\ ۳-۲i & ۱+i & -۲ \end{pmatrix},$$
$$B = \begin{pmatrix} ۲i & ۵+۲i & ۲-۳i \\ -۵+۲i & -۳i & ۱+۵i \\ -۲-۳i & -۱+۵i & ۵i \end{pmatrix},$$

تعریف می‌شوند، به ترتیب هرمیتی و هرمیتی‌کج هستند. همچنین ماتریس

$$A = \begin{pmatrix} -۳ & ۲-i \\ ۲-i & ۲ \end{pmatrix},$$

متقارن است، اما هرمیتی نیست.

انتظار می‌رود که خواننده با اعمال ماتریسی از قبیل ضرب یک اسکالر در یک ماتریس، جمع دو یا چند ماتریس و ضرب یک ماتریس در ماتریس دیگر آشنایی کامل دارد. با این وجود ذکر یک نکته در خصوص ضرب یک ماتریس در بردار لازم است. فرض کنید

$$A = (a_۱, \ldots, a_n) \in \mathbb{C}^{m\times n},$$

که در آن $a_i \in \mathbb{C}^m$، $i = ۱,۲,\ldots,n$ و $x = (x_۱, x_۲, \ldots, x_n)^T \in \mathbb{C}^n$. در این صورت،

$$Ax = x_۱ a_۱ + \cdots + x_n a_n \in \mathbb{C}^m.$$

برای مثال اگر

$$A = \begin{pmatrix} ۱ & ۲ & ۳ \\ ۲ & ۱ & ۴ \end{pmatrix}, \quad x = \begin{pmatrix} ۵ \\ -۱ \\ -۲ \end{pmatrix},$$

آنگاه

$$Ax = ۵\begin{pmatrix} ۱ \\ ۲ \end{pmatrix} + (-۱)\begin{pmatrix} ۲ \\ ۱ \end{pmatrix} + (-۲)\begin{pmatrix} ۳ \\ ۴ \end{pmatrix} = \begin{pmatrix} -۳ \\ ۱ \end{pmatrix}.$$

برای ضرب دو ماتریس، اگر $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ و $B = (b_{ij}) \in \mathbb{C}^{n\times p}$، آنگاه خواهیم داشت $C = AB = (c_{ij}) \in \mathbb{C}^{m\times p}$ که در آن

$$c_{ij} = \sum_{k=۱}^{n} a_{ik}b_{kj} = a_{i*}b_{*j},$$

که در آن a_{i*} و b_{*j} به ترتیب سطر i ام ماتریس A و ستون j ام ماتریس B می‌باشند. همچنین داریم:

$$AB = A(b_{*۱}, \ldots, b_{*p}) = (Ab_{*۱}, \ldots, Ab_{*p}) \in \mathbb{C}^{m\times p}.$$

فرض کنید:

$$A = \begin{pmatrix} ۱ & ۲ & -۱ \\ -۲ & ۳ & -۱ \end{pmatrix}, \quad B = \begin{pmatrix} ۴ & -۳ \\ ۱ & ۵ \\ ۱ & ۱ \end{pmatrix} = (b_۱, b_۲).$$

داریم:

$$Ab_۱ = \begin{pmatrix} ۵ \\ -۶ \end{pmatrix}, \quad Ab_۲ = \begin{pmatrix} ۶ \\ ۲۰ \end{pmatrix}.$$

از این‌رو

$$AB = (Ab_۱, Ab_۲) = \begin{pmatrix} ۵ & ۶ \\ -۶ & ۲۰ \end{pmatrix}.$$

نکتهٔ قابل ذکر دیگر، ضرب ماتریسی بلوکی است. فرض کنید ماتریس‌های A و B به صورت

$$A = \begin{pmatrix} A_{۱۱} & A_{۱۲} & \cdots & A_{۱r} \\ A_{۲۱} & A_{۲۲} & \cdots & A_{۲r} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s۱} & A_{s۲} & \cdots & A_{sr} \end{pmatrix}, \quad B = \begin{pmatrix} B_{۱۱} & B_{۱۲} & \cdots & B_{۱t} \\ B_{۲۱} & B_{۲۲} & \cdots & B_{۲t} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r۱} & B_{r۲} & \cdots & B_{rt} \end{pmatrix},$$

بلوک‌بندی شده باشند به طوری که تعداد ستون‌های A_{ik} با تعداد سطرهای B_{kj} برابر است. در این صورت، همانند ضرب معمولی ماتریس‌ها، ماتریس $C = AB$ یک ماتریس بلوک‌بندی شده است به طوری‌که بلوک (i,j) ام آن به صورت

$$C_{ij} = \sum_{k=1}^{r} A_{ik}B_{kj}, \quad i = 1,\dots,s,\ j = 1,\dots,t,$$

می‌باشد.

مثال ۱.۱ فرض کنید:

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} C & I \\ I & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{pmatrix} = \begin{pmatrix} I & 0 \\ C & C \end{pmatrix},$$

که در آن

$$C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

در این صورت،

$$AB = \begin{pmatrix} C & I \\ I & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ C & C \end{pmatrix} = \begin{pmatrix} 2C & C \\ I & 0 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 1 & 2 \\ 6 & 8 & 3 & 4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

ماتریس $A \in \mathbb{R}^{n\times n}$ را متعامد می‌گوییم، هرگاه $A^TA = AA^T = I$ که در آن I ماتریس همانی است. به همین ترتیب ماتریس $A \in \mathbb{C}^{n\times n}$ را یک ماتریس یکانی می‌گوییم، هرگاه $A^HA = AA^H = I$. برای مثال ماتریس‌های A و B که به صورت

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad B = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix},$$

تعریف می‌شوند به ترتیب متعامد و یکانی هستند.

فرض کنید ماتریس $A = (a_۱, a_۲, \ldots, a_n) \in \mathbb{C}^{n\times n}$ یکانی باشد. در این صورت،

$$(A^H A)_{ij} = (A^H)_{i*} A_{*j} = a_i^H a_j = \begin{cases} ۱, & i = j, \\ ۰, & i \neq j. \end{cases} \qquad (۱.۱)$$

ماتریس $A \in \mathbb{C}^{n\times n}$ را نرمال می‌گوییم، هرگاه $A^H A = A A^H$. در صورتی که ماتریس A حقیقی باشد، شرط نرمال بودن به $A^T A = A A^T$ تبدیل می‌شود. به روشنی ماتریس‌های قطری، هرمیتی (متقارن)، هرمیتی‌کج (متقارن‌کج) و یکانی (متعامد) نرمال هستند. برای مثال ماتریس‌های

$$A = \begin{pmatrix} ۱ & ۷ & ٦ \\ ۷ & ۴ & ۷ \\ ۳ & ۷ & ۱ \end{pmatrix}, \quad B = \begin{pmatrix} ۳ + ۴i & ۱ + ۲i & i \\ ۲ + ۳i & ۱ + ۲i & ۲ + i \\ ۲ + ۳i & ۳ + ۲i & ۴ + ۳i \end{pmatrix},$$

نرمال هستند.

اگر $A = (a_{ij}) \in \mathbb{C}^{n\times n}$، آنگاه اثر ماتریس A به صورت

$$\mathrm{trace}(A) = \sum_{i=۱}^{n} a_{ii},$$

تعریف می‌شود. به سادگی می‌توان دید، اگر $A, B \in \mathbb{C}^{n\times n}$ و $\alpha, \beta \in \mathbb{C}$، آنگاه

- $\mathrm{trace}(\alpha A + \beta B) = \alpha\ \mathrm{trace}(A) + \beta\ \mathrm{trace}(B)$؛
- $\mathrm{trace}(AB) = \mathrm{trace}(BA)$؛
- $\mathrm{trace}(A^H) = \overline{\mathrm{trace}(A)}$.

ماتریس $A \in \mathbb{R}^{n\times n}$ را یک ماتریس جایگشت گویند هرگاه در هر سطر و ستون آن تنها یک درایه برابر با ۱ و سایر درایه‌ها ۰ باشند. برای مثال،

$$P = \begin{pmatrix} ۰ & ۱ & ۰ \\ ۱ & ۰ & ۰ \\ ۰ & ۰ & ۱ \end{pmatrix}, \qquad (۲.۱)$$

یک ماتریس جایگشت است. به عبارت دیگر، فرض کنید $(i_۱, i_۲, \ldots, i_n)$ یک جایگشت از $(۱, ۲, \ldots, n)$ باشد. در این صورت، ماتریس

$$P = \begin{pmatrix} e_{i_۱}^T \\ e_{i_۲}^T \\ \vdots \\ e_{i_n}^T \end{pmatrix},$$

یک ماتریس جایگشت است که در آن e_k ستون k ام ماتریس همانی $n \times n$ است. اگر $A = (a_{ij}) \in \mathbb{C}^{n\times n}$، آنگاه

$$PA = \begin{pmatrix} A_{i_۱*} \\ A_{i_۲*} \\ \vdots \\ A_{i_n*} \end{pmatrix},$$

که در آن A_{i_k*} معرف سطر i_k ام ماتریس A است. به سادگی می‌توان دید که هر ماتریس جایگشت یک ماتریس متعامد است.

مثال ۱.۲ ماتریس جایگشت P تعریف شده در (۱.۲) را در نظر بگیرید. در این صورت، اگر

$$A = \begin{pmatrix} -۲ & ۲ & ۷ \\ ۳ & ۵ & ۸ \\ ۴ & ٦ & -۳ \end{pmatrix}, \qquad (۱.۳)$$

آنگاه

$$PA = \begin{pmatrix} ۳ & ۵ & ۸ \\ ۴ & ٦ & -۳ \\ -۲ & ۲ & ۷ \end{pmatrix}.$$

۱.۲ دترمینان و معکوس ماتریس

فرض می‌کنیم خواننده با مفهوم دترمینان و معکوس یک ماتریس آشنایی دارد. با این وجود، چند نکته مهم در این خصوص ذکر می‌کنیم که در این کتاب به وفور از آنها استفاده

می‌شوند. فرض کنید $A = (a_{ij}) \in \mathbb{C}^{n\times n}$، $A(i,j)$ برابر با دترمینان ماتریسی باشد که از حذف سطر i ام و j ام ماتریس A به دست آمده باشد و

$$\Delta_{ij} = (-1)^{i+j} A(i,j).$$

در این صورت، دترمینان ماتریس A به صورت زیر تعریف می‌شود

$$\det(A) = \begin{cases} a_{11}, & n = 1, \\ \sum_{j=1}^{n} a_{ij}\Delta_{ij}, & n > 1, \end{cases} \tag{4.1}$$

که در آن $1 \leqslant i \leqslant n$ یک مقدار دلخواه است. این رابطهٔ بازگشتی به قاعدهٔ لاپلاس معروف است (برای جزئیات بیشتر به [۹] مراجعه کنید). ماتریس A را نامنفرد (معکوس‌پذیر) می‌گویند، هرگاه یک ماتریس $n \times n$، مانند B وجود داشته باشد به طوری که $AB = BA = I$، که در آن I ماتریس همانی است و در غیر این صورت، ماتریس A را منفرد گویند. B را معکوس ماتریس A می‌گویند و با A^{-1} نشان می‌دهند. معکوس یک ماتریس، یکتا است. می‌توان دید که ماتریس A نامنفرد است اگر و تنها اگر $\det(A) \neq 0$ ([۹] را ببینید). اگر $\det(A) \neq 0$، آنگاه

$$A^{-1} = \frac{1}{\det(A)}\, \Delta^T,$$

که در آن $\Delta = (\Delta_{ij})$.

مثال ۳.۱ فرض کنید:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

داریم:

$$\Delta_{11} = (-1)^{1+1} d = d, \quad \Delta_{12} = (-1)^{1+2} c = -c,$$
$$\Delta_{21} = (-1)^{2+1} b = -b, \quad \Delta_{22} = (-1)^{2+2} a = a,$$

و بنا به قاعدهٔ لاپلاس

$$\det(A) = a\Delta_{11} + b\Delta_{12} = a(-1)^{1+1} d + b(-1)^{1+2} c = ad - bc.$$

بنابراین اگر $ad \neq bc$، آنگاه

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

مثال ۴.۱ فرض کنید:

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

برای محاسبهٔ دترمینان ماتریس A از قاعدهٔ لاپلاس استفاده می‌کنیم. اگر دترمینان را بر حسب سطر اول بسط دهیم، خواهیم داشت:

$$\begin{aligned}\det(A) &= (-1)^{1+1} \times 1 \times \begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} + (-1)^{1+2} \times 1 \times \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} \\ &\quad + (-1)^{1+3} \times 1 \times \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 1.\end{aligned}$$

از طرفی

$$\begin{aligned}
\Delta_{11} &= (-1)^{1+1}\begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} = 2, \quad \Delta_{12} = (-1)^{1+2}\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = -1, \\
\Delta_{13} &= (-1)^{1+1}\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0, \quad \Delta_{21} = (-1)^{2+1}\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = -1, \\
\Delta_{22} &= (-1)^{2+2}\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} = 2, \quad \Delta_{23} = (-1)^{2+3}\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = -1, \\
\Delta_{31} &= (-1)^{3+1}\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0, \quad \Delta_{32} = (-1)^{3+2}\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = -1, \\
\Delta_{33} &= (-1)^{3+3}\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1.
\end{aligned}$$

بنابراین

$$A^{-1} = \frac{1}{1}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}^{T} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

حال نکات زیر در خصوص خواص دترمینان و معکوس ماتریس را یادآوری می‌کنیم ([۳] را ببینید).

• اگر یک سطر یا یک ستون A در $\alpha \in \mathbb{C}$ ضرب شود، آنگاه دترمینان ماتریس حاصل برابر با $\alpha \det(A)$ خواهد بود. لذا داریم $\det(\alpha A) = \alpha^n \det(A)$.

• داریم $\det(A^H) = \overline{\det(A)}$. به عنوان یک نتیجه اگر A یک ماتریس حقیقی باشد، آنگاه $\det(A^T) = \det(A)$.

• اگر A یک ماتریس قطری، بالامثلثی یا پایین‌مثلثی باشد، آنگاه

$$\det(A) = a_{11}a_{22}\dots a_{nn}.$$

• اگر $A, B \in \mathbb{C}^{n\times n}$، آنگاه $\det(AB) = \det(A)\det(B)$. همچنین اگر A نامنفرد باشد، آنگاه

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

• اگر A یک ماتریس بالامثلثی (پایین‌مثلثی) نامنفرد باشد، آنگاه معکوس آن بالامثلثی (پایین مثلثی) است. به‌علاوه

$$(A^{-1})_{ii} = \frac{1}{a_{ii}}, \quad i = 1, 2, \dots, n.$$

• اگر $A, B \in \mathbb{C}^{n\times n}$ نامنفرد باشند، آنگاه $(AB)^{-1} = B^{-1}A^{-1}$.

۳.۱ فضای برداری و برخی از خواص آن

این بخش را با تعریف فضای برداری آغاز می‌کنیم.

تعریف ۱.۱ یک فضای برداری شامل دو مجموعهٔ V و $\mathbb{F}$ و دو عمل جمع برداری و ضرب اسکالر است. اعضای V را بردار و اعضای $\mathbb{F}$ را اسکالر می‌گویند. در این کتاب $\mathbb{F}$، میدان اعداد حقیقی $\mathbb{R}$ یا میدان اعداد مختلط $\mathbb{C}$ است. اگر عمل جمع برداری را با "+" و ضرب اسکالر را با "." نشان دهیم، آنگاه می‌گوییم V یک فضای برداری روی میدان $\mathbb{F}$ است، هرگاه

(۱) به ازای هر $u, v \in V$، $u + v \in V$.

(۲) به ازای هر $\alpha \in \mathbb{F}$ و $v \in V$، $\alpha.v \in V$.

(۳) به ازای هر $u, v \in V$، $u + v = v + u$.

(۴) به ازای هر $u, v, w \in V$، $(u + v) + w = u + (v + w)$.

(۵) عضوی مثل $۰ \in V$ وجود دارد به طوری که به ازای هر $v \in V$، $v + ۰ = v$.

(٦) برای هر $v \in V$ عضوی مثل $-v \in V$ وجود دارد به طوری که $v + (-v) = ۰$.

(۷) برای هر $\alpha, \beta \in \mathbb{F}$ و $v \in V$، $(\alpha\beta).v = \alpha(\beta.v)$.

(۸) برای هر $\alpha, \beta \in \mathbb{F}$ و $v \in V$، $(\alpha + \beta).v = \alpha.v + \beta.v$.

(۹) برای هر $\alpha \in \mathbb{F}$ و $v, w \in V$، $\alpha(v + w) = \alpha.v + \alpha.w$.

(۱۰) برای هر $v \in V$، $۱.v = v$ (دراینجا $۱ \in \mathbb{F}$).

یک فضای برداری معمولاً به صورت $(V, \mathbb{F})$ نشان داده می‌شود و در صورتی جای اشتباه وجود نداشته باشد گاهی اوقات با V نیز نشان داده می‌شود. همچنین می‌توان از نوشتن "." در تعریف ۱.۱ صرفنظر کرد، زیرا به راحتی می‌توان بین ضرب‌های اسکالر در اسکالر و ضرب اسکالر در بردار تمایز قائل شد.

مثال ۵.۱ قرار می‌دهیم $\mathbb{F} = \mathbb{R}$ و $V = \mathbb{R}^n$. در این صورت، $(V, \mathbb{F})$ با جمع برداری و ضرب اسکالر در بردار

$$v + w = \begin{pmatrix} v_۱ + w_۱ \\ v_۲ + w_۲ \\ \vdots \\ v_n + w_n \end{pmatrix}, \quad \alpha.v = \begin{pmatrix} \alpha v_۱ \\ \alpha v_۲ \\ \vdots \\ \alpha v_n \end{pmatrix},$$

یک فضای برداری است. به طور مشابه فضای برداری $(\mathbb{C}^n, \mathbb{C})$ نیز تعریف می‌شود.

مثال ٦.۱ قرار می‌دهیم $\mathbb{F} = \mathbb{R}$ و $V = \mathbb{R}^{m\times n}$. در این صورت $(V, \mathbb{F})$ با جمع

$$A + B = \begin{pmatrix} a_{۱۱} + b_{۱۱} & a_{۱۲} + b_{۱۲} & \cdots & a_{۱n} + b_{۱n} \\ a_{۲۱} + b_{۲۱} & a_{۲۲} + b_{۲۲} & \cdots & a_{۲n} + b_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m۱} + b_{m۱} & a_{m۲} + b_{m۲} & \cdots & a_{mn} + b_{mn} \end{pmatrix},$$

و ضرب اسکالری

$$\alpha.A = \begin{pmatrix} \alpha a_{۱۱} & \alpha a_{۱۲} & \cdots & \alpha a_{۱n} \\ \alpha a_{۲۱} & \alpha a_{۲۲} & \cdots & \alpha a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m۱} & \alpha a_{m۲} & \cdots & \alpha a_{mn} \end{pmatrix},$$

یک فضای برداری است. به طور مشابه فضای برداری $(\mathbb{C}^{m\times n}, \mathbb{C})$ نیز تعریف می‌شود.

تعریف ۲.۱ فرض کنید $(V, \mathbb{F})$ یک فضای برداری و W یک زیرمجموعهٔ ناتهی از V باشد. اگر برای هر $\alpha, \beta \in \mathbb{F}$ و $w_۱, w_۲ \in W$، داشته باشیم $\alpha w_۱ + \beta w_۲ \in W$، آنگاه $(W, \mathbb{F})$ را یک زیر فضای $(V, \mathbb{F})$ می‌گوییم.

مثال ۷.۱ فرض کنید $(V, \mathbb{F}) = (\mathbb{R}^{n\times n}, \mathbb{R})$ و $W = \{A \in \mathbb{R}^{n\times n} : A^T = A\}$. در این صورت، $(W, \mathbb{F})$ یک زیر فضای $(V, \mathbb{R})$ است. زیرا اگر $\alpha, \beta \in \mathbb{R}$ و $A, B \in W$، آنگاه $\alpha A + \beta B$ یک ماتریس متقارن است.

تعریف ۳.۱ مجموعهٔ $S = \{v_۱, v_۲, \ldots, v_m\}$ در فضای برداری $(V, \mathbb{F})$ یک مجموعهٔ وابستهٔ خطی نامیده می‌شود، هرگاه m اسکالر $\alpha_i \in \mathbb{F}$، $i = ۱, ۲, \ldots, m$ که همگی صفر نیستند، وجود داشته باشند به طوری که $\alpha_۱ v_۱ + \alpha_۲ v_۲ + \cdots + \alpha_m v_m = ۰$. همچنین، S یک مجموعهٔ مستقل خطی از بردارها نامیده می‌شود، اگر به ازای اسکالرهای $\alpha_i \in \mathbb{F}$، $i = ۱, ۲, \ldots, m$ داشته باشیم $\alpha_۱ v_۱ + \alpha_۲ v_۲ + \cdots + \alpha_m v_m = ۰$، آنگاه نتیجه شود که $\alpha_۱ = \alpha_۲ = \cdots = \alpha_m = ۰$.

مثال ۸.۱ مجموعهٔ بردارهای

$$S = \left\{ u_۱ = \begin{pmatrix} ۱ \\ -۲ \\ ۱ \end{pmatrix}, u_۲ = \begin{pmatrix} -۱ \\ ۲ \\ ۲ \end{pmatrix}, u_۳ = \begin{pmatrix} ۱ \\ -۲ \\ ۴ \end{pmatrix} \right\},$$

در $\mathbb{R}^۳$ وابستهٔ خطی می‌باشند، زیرا داریم:

$$۲u_۱ + ۱u_۲ + (-۱)u_۳ = ۰.$$

از طرفی مجموعهٔ بردارهای

$$S = \left\{ v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\},$$

در $\mathbb{R}^3$ مستقل خطی هستند، زیرا فرض کنید:

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0.$$

این رابطه معادل با دستگاه معادلات خطی

$$\begin{pmatrix} 1 & -1 & 1 \\ -2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

است. دترمینان ماتریس ضرایب این دستگاه برابر با ۵− است. لذا ماتریس ضرایب این دستگاه نامنفرد است و دستگاه فقط جواب بدیهی $\alpha_1 = \alpha_2 = \alpha_3 = 0$ دارد.

تعریف ۴.۱ فرض کنید $(V, \mathbb{F})$ یک فضای برداری باشد و $S = \{v_1, v_2, \ldots, v_k\}$ که در آن $v_i \in V$، $i = 1, \ldots, k$. در این صورت، فضای تولید شده توسط S با $\mathrm{span}(S)$ نشان داده می‌شود و به صورت

$$\mathrm{span}(S) = \{v : v = \alpha_1 v_1 + \cdots + \alpha_k v_k;\ \alpha_i \in \mathbb{F},\ v_i \in S\},$$

تعریف می‌شود.

تعریف ۵.۱ مجموعه بردارهای B یک پایه برای فضای برداری V نامیده می‌شود، اگر
(الف) B یک مجموعه از بردارهای مستقل خطی (بردارهای پایه) باشد؛
(ب) $\mathrm{span}(B) = V$.

تعریف ۶.۱ فرض کنید $\{b_1, b_2, \ldots, b_n\}$ یک پایه برای فضای برداری V باشد. در این صورت، برای هر $v \in V$ یک n- تایی مرتب یکتا مثل $(\xi_1, \xi_2, \ldots, \xi_n)$ وجود دارد به طوری که

$$v = \xi_1 b_1 + \xi_2 b_2 + \cdots + \xi_n b_n = Bx,$$

که در آن $x = (\xi_۱, \xi_۲, \dots, \xi_n)^T$ و $B = (b_۱, b_۲, \dots, b_n)$. در اینجا ξ_i ها را مختصات بردار v نسبت به پایهٔ $\{b_۱, b_۲, \dots, b_n\}$ می‌گویند.

تعریف ۷.۱ اگر پایهٔ B برای زیرفضای $V\ (\neq \emptyset)$ دارای n عضو باشد، آنگاه گفته می‌شود که این فضا n- بُعدی است (یا با بُعد n است) و می‌نویسیم $\dim(V) = n$. اگر $\dim(V) < \infty$، آنگاه فضا را با بُعد متناهی و در غیر این صورت با بُعد نامتناهی گویند.

مثال ۹.۱ سه زیرفضای برداری همراه با با بُعد هر کدام از آنها به صورت زیر معرفی می‌شوند.

(الف) مجموعهٔ $\{e_۱, e_۲, \dots, e_n\}$ یک پایه برای $\mathbb{R}^n$ است که در آن e_i یک بردار n- تایی است که تمام درایه‌های آن برابر با صفرند، بجز درایهٔ i ام که برابر با ۱ است. بنابراین $\dim(\mathbb{R}^n) = n$.

(ب) فرض کنید E_{ij} یک ماتریس $m \times n$ باشد که تمام درایه‌های آن صفرند، بجز درایهٔ (i, j) ام که برابر با ۱ است. در این صورت، مجموعهٔ $\{E_{ij} : ۱ \leqslant i \leqslant m, ۱ \leqslant j \leqslant n\}$ یک پایه برای $\mathbb{R}^{m\times n}$ است. بنابراین $\dim(\mathbb{R}^{m\times n}) = mn$.

(ج) می‌دانیم $W = \{A \in \mathbb{R}^{۲\times۲} : A^T = A\}$ یک زیرفضای $V = \mathbb{R}^{۲\times۲}$ است. مجموعهٔ $B = \{B_۱, B_۲, B_۳\}$ یک پایه برای W است که در آن

$$B_۱ = \begin{pmatrix} ۱ & ۰ \\ ۰ & ۰ \end{pmatrix}, \quad B_۲ = \begin{pmatrix} ۰ & ۰ \\ ۰ & ۱ \end{pmatrix}, \quad B_۳ = \begin{pmatrix} ۰ & ۱ \\ ۱ & ۰ \end{pmatrix}.$$

زیرا هر ماتریس متقارنِ ۲ × ۲ را می‌توان به صورت

$$\begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix} = \alpha B_۱ + \beta B_۲ + \gamma B_۳,$$

نمایش داد. بنابراین $W = \operatorname{span}(B)$. در اینجا α، β و γ را مختصات ماتریس نسبت به پایهٔ B نامیده می‌شوند. از طرفی اگر

$$\alpha B_۱ + \beta B_۲ + \gamma B_۳ = ۰,$$

آنگاه خواهیم داشت $\alpha = \beta = \gamma = ۰$. درنتیجه $\dim(W) = ۳$.

تعریف ۸.۱ برای هر ماتریس $A \in \mathbb{R}^{m\times n}$ دو زیرفضای بُرد A و فضای پوچ A به ترتیب به صورت

$$\mathcal{R}(A) = \{Ax : x \in \mathbb{R}^n\},$$
$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\},$$

تعریف می‌شوند.

به وضوح اگر $A = (a_1, a_2, \ldots, a_n)$، آنگاه $\mathcal{R}(A) = \operatorname{span}\{a_1, a_2, \ldots, a_n\}$. رتبهٔ ماتریس A را با $\operatorname{rank}(A)$ نشان داده و به صورت $\operatorname{rank}(A) = \dim(\mathcal{R}(A))$ تعریف می‌شود. می‌توان دید که ([۸] را ببینید)

$$\operatorname{rank}(A) = \operatorname{rank}(A^T),$$
$$\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = n. \tag{۵.۱}$$

مثال ۱۰.۱ فرض کنید:

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \end{pmatrix} = (a_1, a_2, a_3),$$

که در آن a_i ها ستون‌های ماتریس A هستند. داریم $\mathcal{R}(A) = \operatorname{span}\{a_1, a_2, a_3\}$ از طرفی بردارهای a_1 و a_2 مستقل خطی هستند، امّا $a_3 = 4a_1 - a_2$. از این‌رو، خواهیم داشت $\mathcal{R}(A) = \operatorname{span}\{a_1, a_2\}$ و در نتیجه $\dim(\mathcal{R}(A)) = 2$. از طرفی به ازای هر بردار $x = (x_1, x_2, x_3)^T$ در $\mathcal{N}(A)$ خواهیم داشت $Ax = 0$، که معادل با

$$\begin{cases} x_1 + x_2 + 3x_3 = 0, \\ x_1 + 2x_2 + 2x_3 = 0, \end{cases}$$

است. با فرض $x_3 = \alpha \in \mathbb{R}$، خواهیم داشت $x_1 = -4\alpha$ و $x_2 = \alpha$. بنابراین به ازای هر $\alpha \in \mathbb{R}$، بردار

$$x = \begin{pmatrix} -4\alpha \\ \alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix} = \alpha w,$$

عضوی از $\mathcal{N}(A)$ است و در نتیجه $\mathcal{N}(A) = \operatorname{span}\{w\}$ و $\dim(\mathcal{N}(A)) = 1$. با توجه به اینکه $\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = 3$ می‌بینیم که رابطهٔ (۵.۱) در اینجا برقرار است.

مثال ۱۱.۱ ماتریس

$$A = \begin{pmatrix} ۰ & ۱ & ۰ & \cdots & ۰ \\ ۰ & ۰ & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & ۰ & ۱ \\ ۰ & ۰ & ۰ & ۰ & ۰ \end{pmatrix} \in \mathbb{R}^{n\times n},$$

را در نظر بگیرید. به ازای هر $x \in \mathbb{R}^n$، داریم:

$$Ax = \begin{pmatrix} x_۲ \\ \vdots \\ x_n \\ ۰ \end{pmatrix}.$$

بنابراین $x \in \mathcal{N}(A)$ اگر و تنها اگر $x_۲ = \cdots = x_n = ۰$ و $x_۱ = \alpha$ که در آن $\alpha \in \mathbb{R}$ دلخواه است. از این‌رو،

$$\mathcal{N}(A) = \left\{ \begin{pmatrix} \alpha \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} : \alpha \in \mathbb{R} \right\} = \left\{ \alpha \begin{pmatrix} ۱ \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} : \alpha \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} ۱ \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} \right\}.$$

لذا $B = \{(۱, ۰, \ldots, ۰)^T\}$ یک پایه برای $\mathcal{N}(A)$ است و $\dim(\mathcal{N}(A)) = ۱$. از طرفی داریم $y = (y_۱, \ldots, y_n)^T \in \mathcal{R}(A)$ اگر و تنها اگر

$$\begin{pmatrix} y_۱ \\ \vdots \\ y_n \end{pmatrix} = Ax = \begin{pmatrix} x_۲ \\ \vdots \\ x_n \\ ۰ \end{pmatrix}.$$

بنابراین

$$\begin{aligned} \mathcal{R}(A) &= \left\{ \begin{pmatrix} x_۲ \\ \vdots \\ x_n \\ ۰ \end{pmatrix} : x_i \in \mathbb{R}, i = ۲, \ldots, n \right\} \\ &= \{x_۲ e_۱ + \cdots + x_n e_{n-۱} : x_i \in \mathbb{R}, i = ۲, \ldots, n\} \\ &= \operatorname{span}\{e_i : i = ۱, \ldots, n-۱\}, \end{aligned}$$

که درآن e_i ستون i ام ماتریس همانی $n \times n$ است. در نتیجه

$$B' = \{e_i : i = ۱, \ldots, n - ۱\},$$

یک پایه برای $\mathcal{R}(A)$ است و داریم $\dim(\mathcal{R}(A)) = n - ۱$. سرانجام توجه کنید که داریم:

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = ۱ + (n - ۱) = n.$$

۴.۱ منشأ ظهور دو مسألهٔ اساسی جبرخطی

حل دستگاه معادلات خطی و مسألهٔ مقدار ویژه، دو موضوع مهم در شاخه‌های مختلف علوم و مهندسی می‌باشند. در این بخش دو مثال از مدل‌هایی ارائه می‌کنیم که حل تقریبی آنها منجر به حل یکی از این دو مسأله می‌شود. روش‌های حل این دو مسأله با جزئیات بیشتر در فصل‌های بعدی بررسی می‌شوند. ابتدا نماد $\mathcal{O}$ را معرفی می‌کنیم.

تعریف ۹.۱ **(نماد $\mathcal{O}$)** فرض کنید توابع f و g در یک همسایگی از $x_۰$ تعریف شده باشند. می‌گوییم تابع f وقتی که x به $x_۰$ میل می‌کند از مرتبهٔ تابع g است، هرگاه یک ثابت نامنفی مثل C و یک همسایگی از $x_۰$ وجود داشته باشد به طوری که به ازای هر x در این همسایگی داشته باشیم $|f(x)| \leqslant C|g(x)|$. در این صورت، می‌نویسیم $f(x) = \mathcal{O}(g(x))$.

مثال ۱۲.۱ با استفاده از بسط تیلور تابع $f(x) = \sin x$، به ازای یک $\xi \in (۰, x)$، داریم:

$$\sin x = x - \frac{x^۳}{۳!} + \frac{x^۵}{۵!}\cos \xi.$$

در نتیجه

$$|\sin x - (x - \frac{x^۳}{۳!})| = |\frac{\cos \xi}{۵!}||x^۵| \leqslant \frac{۱}{۱۲۰}|x^۵|.$$

بنابراین وقتی که $x \to ۰$، خواهیم داشت:

$$\sin x - (x - \frac{x^۳}{۳!}) = \mathcal{O}(x^۵).$$

حال مدلی از معادلات دیفرانسیل معمولی را بررسی می‌کنیم که حل عددی آن به حل دستگاه معادلات خطی منجر می‌شود. معادلهٔ دیفرانسیل زیر با شرایط مرزی داده شدهٔ

$$\begin{cases} -y'' + xy = x^2, & x \in (0, 1), \\ y(0) = 0, \quad y(1) = 1, \end{cases}$$

را در نظر بگیرید. برای حل این مدل با استفاده از روش تفاضلات متناهی، بازهٔ $[0, 1]$ را به $n + 1$ بازهٔ مساوی تقسیم کرده و جواب معادله دیفرانسیل را در نقاط گرهی درونی به صورت تقریبی محاسبه می‌کنیم. برای این کار، قرار می‌دهیم $h = 1/(n + 1)$ و $x_i = ih$، $i = 0, 1, \ldots, n + 1$. معادلهٔ دیفرانسیل را در نقاط گرهی درونی x_i، $i = 1, 2, \ldots, n$ می‌نویسیم:

$$-y''(x_i) + x_i y(x_i) = x_i^2, \quad i = 1, 2, \ldots, n.$$

حال بجای $y''(x_i)$ مقدار

$$y''(x_i) = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + \mathcal{O}(h^2),$$

را جایگذاری می‌کنیم که در آن $y_i = y(x_i)$. با صرفنظر کردن از $\mathcal{O}(h^2)$، خواهیم داشت:

$$-\frac{y_{i-1} - 2y_i + y_{i+1}}{h^2} + x_i y_i = x_i^2, \quad i = 1, 2, \ldots, n.$$

این رابطه را می‌توان به صورت

$$-y_{i-1} + (2 + h^2 x_i) y_i - y_{i+1} = h^2 x_i^2, \quad i = 1, 2, \ldots, n,$$

نوشت. با توجه به شرایط مرزی داده شده، داریم $y_0 = 0$ و $y_{n+1} = 1$. بنابراین محاسبهٔ مقادیر تقریبی y_i، $i = 1, 2, \ldots, n$، توسط دستگاه معادلات خطی

$$Ay = b, \tag{6.1}$$

انجام می‌پذیرد که در آن

$$A = \begin{pmatrix} 2 + h^2 x_1 & -1 & & & \\ -1 & 2 + h^2 x_2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 + h^2 x_{n-1} & -1 \\ & & & -1 & 2 + h^2 x_n \end{pmatrix},$$

و

$$y=\begin{pmatrix} y_۱ \\ y_۲ \\ \vdots \\ y_{n-۱} \\ y_n \end{pmatrix}, \quad b=\begin{pmatrix} h^۲x_۱^۲ \\ h^۲x_۲^۲ \\ \vdots \\ h^۲x_{n-۱}^۲ \\ h^۲x_n^۲+۱ \end{pmatrix}.$$

ملاحظه می‌شود که حل عددی مدل در نظر گرفته شده به حل یک دستگاه معادلات خطی منجر شده است.

حال مدلی از معادلات دیفرانسیل معمولی را بررسی می‌کنیم که حل عددی آن منجر به حل یک مسألهٔ مقدار ویژه می‌شود. دستگاه معادلات دیفرانسیل معمولی همگن

$$\begin{cases} x'_۱(t)=a_{۱۱}x_۱(t)+a_{۱۲}x_۲(t)+\cdots+a_{۱n}x_n(t), \\ x'_۲(t)=a_{۲۱}x_۱(t)+a_{۲۲}x_۲(t)+\cdots+a_{۲n}x_n(t), \\ \qquad\vdots \\ x'_n(t)=a_{n۱}x_۱(t)+a_{n۲}x_۲(t)+\cdots+a_{nn}x_n(t), \end{cases} \tag{۷.۱}$$

با شرط اولیهٔ $x_i(a)=\alpha_i$، $i=۱,۲,\ldots,n$، را در نظر بگیرید که در آن مقادیر a_{ij} و α_i حقیقی و معلوم هستند. قرار می‌دهیم:

$$X(t)=\begin{pmatrix} x_۱(t) \\ x_۲(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad A=\begin{pmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} \end{pmatrix}.$$

در این صورت، معادلهٔ (۷.۱) را می‌توان به صورت

$$X'(t)=AX(t), \tag{۸.۱}$$

نوشت که در آن $X(a)=\alpha$ و $\alpha=(\alpha_۱,\ldots,\alpha_n)^T$. فرض می‌کنیم $\alpha\neq ۰$. برای حل دستگاه (۸.۱) فرض می‌کنیم که این دستگاه جوابی به صورت $X(t)=e^{\lambda t}u$ داشته باشد، که در آن u یک بردار ناصفر است. با جایگذاری این جواب در (۸.۱)، خواهیم داشت:

$$\frac{d}{dt}(e^{\lambda t}u)=Ae^{\lambda t}u.$$

در نتیجه $Ae^{\lambda t}u = \lambda e^{\lambda t}u$، یا

$$Au = \lambda u. \tag{۹.۱}$$

این مسأله به مسألهٔ مقدار ویژه معروف است. در واقع بایستی یک اسکالر مثل λ و یک بردار ناصفر مثل u بیابیم که در رابطهٔ (۹.۱) صدق کند (در این خصوص، در فصل ۴ به طور مفصل بحث خواهد شد). می‌بینیم که مسألهٔ اصلی منجر به حل یک مسألهٔ مقدار ویژه شده است. فرض کنید مقادیر ویژهٔ A متمایز و u_i، $i = ۱, ۲ \ldots, n$، بردارهای ویژهٔ A متناظر با مقادیر ویژهٔ λ_i، $i = ۱, ۲ \ldots, n$ باشند. در این صورت، $X_i(t) = e^{\lambda_i t}u_i$، $i = ۱, ۲ \ldots, n$، جواب‌های مسألهٔ (۷.۱) می‌باشند. از این‌رو جواب عمومی مسألهٔ (۷.۱) به صورت

$$X(t) = c_۱ e^{\lambda_۱ t}u_۱ + c_۲ e^{\lambda_۲ t}u_۲ + \cdots + c_n e^{\lambda_n t}u_n,$$

خواهد بود که در آن c_i ها ضرایبی هستند که بایستی محاسبه گردند. این ضرایب را نیز می‌توان از شرایط اولیهٔ داده شده به دست آورد.

تمرینات

۱.۱) نشان دهید اگر $A \in \mathbb{C}^{n\times n}$ بالامثلثی اکید یا پایین‌مثلثی اکید باشد، آنگاه داریم $A^n = ۰$.

۲.۱) فرض کنید $A = (a_{ij}) \in \mathbb{C}^{n\times n}$. نشان دهید اگر e_i و e_j به ترتیب ستون‌های i ام و j ام ماتریس همانی I_n باشند، آنگاه $a_{ij} = e_i^T A e_j$.

۳.۱) نشان دهید اگر A یک ماتریس بالاهسنبرگی و T یک ماتریس بالامثلثی باشد، آنگاه AT یک ماتریس بالاهسنبرگی است.

۴.۱) فرض کنید $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{n\times n}$ دو ماتریس بالامثلثی باشند. نشان دهید $(AB)_{ii} = a_{ii}b_{ii}$، $i = ۱, ۲, \ldots, n$.

۵.۱) فرض کنید $x \in \mathbb{R}^n$، $x^T x = ۱$ و $H = I - ۲xx^T$. نشان دهید H متقارن است، $H^۲ = I$ و $H^T H = I$.

۶.۱) فرض کنید $A, B \in \mathbb{C}^{m\times n}$. نشان دهید $A = B$ اگر و تنها اگر به ازای هر $x \in \mathbb{C}^n$، $Ax = Bx$.

۷.۱) نشان دهید
(الف) درایه‌های روی قطر هر ماتریس هرمیتی، حقیقی هستند.
(ب) درایه‌های روی قطر هر ماتریس متقارن‌کج برابر با صفر هستند.
(ج) درایه‌های روی قطر هر ماتریس هرمیتی‌کج، صفر یا موهومی محض هستند (قسمت حقیقی درایه‌های قطری صفرند).

۸.۱) فرض کنید $A \in \mathbb{C}^{n\times n}$ هرمیتی باشد و $x \in \mathbb{C}^n$. نشان دهید $x^H Ax$ حقیقی است.

۹.۱) فرض $A \in \mathbb{R}^{n\times n}$ یک ماتریس متقارن‌کج باشد. نشان دهید به ازای هر $x \in \mathbb{R}^n$، $x^T Ax = ۰$.

۱۰.۱) فرض کنید $A, B \in \mathbb{C}^{n\times n}$ دو ماتریس یکانی باشند. نشان دهید AB نیز یکانی است.

۱۱.۱) معکوس ماتریس‌های

$$A = \begin{pmatrix} ۳ & ۵ \\ ۴ & ۷ \end{pmatrix}, \qquad B = \begin{pmatrix} -۲ & ۱ & ۰ \\ -۱ & ۲ & -۱ \\ ۰ & -۱ & ۲ \end{pmatrix},$$

را بیابید.

۱۲.۱) نشان دهید $\mathcal{B} = \{v_۱, v_۲, v_۳\}$ و $\mathcal{B}' = \{w_۱, w_۲, w_۳\}$ دو پایه برای $\mathbb{R}^۳$ می‌باشند که در آن

$$v_۱ = \begin{pmatrix} ۱ \\ ۱ \\ ۱ \end{pmatrix}, \quad v_۲ = \begin{pmatrix} ۱ \\ ۰ \\ -۱ \end{pmatrix}, \quad v_۳ = \begin{pmatrix} ۱ \\ ۰ \\ ۱ \end{pmatrix},$$
$$w_۱ = \begin{pmatrix} ۱ \\ ۲ \\ ۱ \end{pmatrix}, \quad w_۲ = \begin{pmatrix} ۲ \\ ۳ \\ ۴ \end{pmatrix}, \quad w_۳ = \begin{pmatrix} ۳ \\ ۴ \\ ۳ \end{pmatrix}.$$

۱۳.۱) کدام یک از مجموعه‌های

$$V_1 = \{x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 0\},$$
$$V_2 = \{x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n : \sum_{i=1}^{n} i x_i = 0\},$$
$$V_3 = \{x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1\},$$

یک زیرفضا از فضای برداری $\mathbb{R}^n$ می‌باشند.

۱۴.۱) نشان دهید $\{(x, y-x, y)^T : x, y, \in \mathbb{R}\}$ یک زیرفضای برداری از $\mathbb{R}^3$ است. بعد و یک پایه برای این زیرفضا بیابید.

۱۵.۱) کدام یک از گزاره‌های زیر درست و کدام یک نادرست هستند.

(الف) مجموعهٔ $\{A \in \mathbb{C}^{n\times n} : A^H = A\}$ یک زیرفضای برداری $\mathbb{C}^{n\times n}$ است.

(ب) مجموعهٔ $\{A \in \mathbb{R}^{n\times n} : A^T = -A\}$ یک زیرفضای برداری $\mathbb{R}^{n\times n}$ است.

(ج) مجموعهٔ $\{A \in \mathbb{C}^{n\times n} : \mathrm{trace}(A) = 0\}$ یک زیرفضای برداری $\mathbb{C}^{n\times n}$ است.

۱۶.۱) ماتریس‌های P و A تعریف شده در روابط (۲.۱) و (۳.۱) را در نظر بگیرید. ماتریس‌های AP و PAP^T را محاسبه کرده و با ماتریس A مقایسه کنید. چه نتیجه‌ای می‌گیرید؟ این نتیجه را در حالت کلی بیان کنید.

۱۷.۱) فرض کنید

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

یک پایه برای فضاهای $\mathcal{R}(A)$ و $\mathcal{N}(A)$ بیابید.

۱۸.۱) فرض کنید $u, v \in \mathbb{R}^n$ دو بردار ناصفر باشند و $A = uv^T$. یک پایه برای فضاهای $\mathcal{R}(A)$ و $\mathcal{N}(A)$ بیابید.

۱۹.۱) فرض کنید $A \in \mathbb{R}^{n\times n}$، $x \in \mathcal{N}(A^T)$ و $y \in \mathcal{R}(A)$. نشان دهید $x^T y = 0$.

۲۰.۱) فرض کنید I ماتریس همانی $m \times m$ و

$$J = \begin{pmatrix} ۰ & ۱ & & \\ & ۰ & \ddots & \\ & & \ddots & ۱ \\ & & & ۰ \end{pmatrix} \in \mathbb{R}^{m\times m}.$$

یک پایه برای فضاهای $\mathcal{R}(A)$ و $\mathcal{N}(A)$ بیابید که در آن

$$A = \begin{pmatrix} I & ۰ \\ ۰ & J \end{pmatrix}.$$

۲۱.۱) فرض کنید $S \in \mathbb{C}^{n\times n}$ ماتریس همانی پسرو باشد، یعنی

$$S = \begin{pmatrix} ۰ & ۰ & \cdots & ۰ & ۱ \\ ۰ & ۰ & \cdots & ۱ & ۰ \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ ۰ & ۱ & \cdots & ۰ & ۰ \\ ۱ & ۰ & \cdots & ۰ & ۰ \end{pmatrix}.$$

نشان دهید $S^{-۱} = S^T = S$ یا به طور معادل $S^T = S$، $S^۲ = I$ و $S^T S = I$. دترمینان S را محاسبه کنید. به ازای $n = ۳$ و برای هر ماتریس $A \in \mathbb{C}^{۳\times ۳}$، ماتریس SAS را محاسبه کنید.

۲۲.۱) فرض کنید $A, B \in \mathbb{R}^{n\times n}$ دو ماتریس متقارن باشند. نشان دهید AB متقارن است اگر و تنها اگر $AB = BA$.

۲۳.۱) فرض کنید $A \in \mathbb{C}^{n\times n}$. نشان دهید که A را می‌توان به صورت $A = H + iK$ نمایش داد که در آن H و K ماتریس‌های هرمیتی هستند.

۲۴.۱) فرض کنید $A \in \mathbb{R}^{n\times n}$. نشان دهید اگر $A^T A = ۰$، آنگاه $A = ۰$.

۲۵.۱) نشان دهید اگر ماتریس $A \in \mathbb{R}^{n\times n}$ منفرد باشد، آنگاه بردار ناصفری مانند $x \in \mathbb{R}^n$ وجود دارد به طوری $Ax = ۰$.

۲۶.۱) فرض کنید

$$P_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

که در آن α یک عدد حقیقی است. نشان دهید

(الف) $P_\alpha P_\beta = P_{\alpha+\beta}$.

(ب) $P_\alpha^{-۱} = P_{-\alpha}$.

(ج) به ازای هر عدد صحیح k، $P_\alpha^k = P_{k\alpha}$.

۲۷.۱) فرض کنید $A, B \in \mathbb{R}^{n\times n}$. نشان دهید:

(الف) برای هر $\alpha, \beta \in \mathbb{R}$، $\mathrm{trace}(\alpha A + \beta B) = \alpha\ \mathrm{trace}(A) + \beta\ \mathrm{trace}(B)$.

(ب) نشان دهید $\mathrm{trace}(AB) = \mathrm{trace}(BA)$، حتی اگر $AB \neq BA$.

(ج) نشان دهید اگر ماتریس S متقارن‌کج باشد، آنگاه $\mathrm{trace}(S) = ۰$.

(د) $\mathrm{trace}(A^H) = \overline{\mathrm{trace}(A)}$.

۲۸.۱) ماتریس $A \in \mathbb{R}^{n\times n}$ را خودتوان گویند هرگاه $A^۲ = A$.

(الف) نشان دهید ماتریس

$$A = \frac{۱}{۲}\begin{pmatrix} ۲\cos^۲\theta & \sin ۲\theta \\ \sin ۲\theta & ۲\sin^۲\theta \end{pmatrix},$$

به ازای هر $\theta \in \mathbb{R}$ خود توان است.

(ب) اگر ماتریس $A \in \mathbb{R}^{n\times n}$ خودتوان باشد آنگاه یکی از دو ماتریس A یا $I - A$ منفرد می‌باشد.

۲۹.۱) برای هر دو ماتریس $A \in \mathbb{R}^{n\times k}$ و $B \in \mathbb{R}^{k\times n}$، نشان دهید برای

$$L = \begin{pmatrix} I - BA & B \\ ۲A - ABA & AB - I \end{pmatrix},$$

داریم $L^۲ = I$.

۳۰.۱) فرض کنید A یک ماتریس $n \times n$ باشد. نشان دهید هیچ ماتریسی مانند $X \in \mathbb{R}^{n\times n}$ وجود ندارد که $AX - XA = I$ که در آن I ماتریس همانی $n \times n$ است.
(راهنمایی: کافی است اثر طرفین معادله را محاسبه کنید)

۳۱.۱) برای چندجمله‌ای $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ و ماتریس $A \in \mathbb{R}^{n\times n}$، فرض کنید $p(A) = a_0 I + a_1 A + \cdots + a_n A^n$. نشان دهید $p(A^T) = p(A)^T$. به علاوه نشان دهید اگر q چندجمله‌ای دیگری از درجه‌ی m باشد، آنگاه $p(A)q(A) = q(A)p(A)$.

۳۲.۱) به ازای هر $a, b \in \mathbb{C}$ معکوس ماتریس

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

را به دست آورید.

۳۳.۱) فرض کنید ماتریس‌های بلوکی

$$\begin{pmatrix} I & 0 \\ C & I \end{pmatrix}, \quad \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}, \quad \begin{pmatrix} 0 & I \\ I & D \end{pmatrix},$$

نامنفرد باشند. معکوس آنها را محاسبه کنید. اگر ماتریس A نامنفرد باشد، نشان دهید:

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$

بعلاوه اگر ماتریس $S = D - CA^{-1}B$ نامنفرد باشد، معکوس ماتریس M را به دست آورید.

۳۴.۱) (الف) فرض کنید:

$$J = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{R}^{m\times m}.$$

به ازای $m = 4$، J^2، J^3، J^4 را محاسبه کنید. به ازای مقادیر دلخواه m و k، J^k را محاسبه کنید.

(ب) قرار می‌دهیم $e^J = I + \frac{1}{1!}J + \frac{1}{2!}J^2 + + \frac{1}{3!}J^3 + \cdots$ که در آن I ماتریس همانی $m \times m$ است. ماتریس e^J را محاسبه کنید.

فصل ۲

ضرب داخلی و نرم

برای تخمین خطا در حل دستگاه معادلات خطی به روش‌های مختلف نیاز به محاسبهٔ اندازهٔ یک بردار و یا یک ماتریس می‌باشد و محاسبهٔ این اندازه توسط نرم‌های مختلف انجام می‌گیرد. یکی از عملگرهای مهم در جبرخطی ضرب داخلی می‌باشد که در تعریف نرم، پایهٔ متعامد و ماتریس متعامد کاربرد دارند. این فصل به تعریف نرم، ضرب داخلی و مباحث مربوط به آنها اختصاص یافته است.

۱.۲ نرم

تعریف ۱.۲ یک نرم روی فضای برداری V تعریف شده روی میدان $\mathbb{F}$، تابعی است مثل $\|.\|$ از V به $\mathbb{R}$، به طوری که در شرایط زیر صدق کند:

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

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

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

در سراسر این کتاب میدان $\mathbb{F}$ برابر با $\mathbb{R}$ یا $\mathbb{C}$ می‌باشد. از اینجا به بعد نرم‌های تعریف شده روی $\mathbb{C}^n$ و $\mathbb{C}^{m\times n}$ را به ترتیب نرم برداری و نرم ماتریسی می‌نامیم.

مثال ۱.۲ فرض کنید $V = \mathbb{C}^n$. نرم-۱ بردار $x = (x_1, \ldots, x_n)^T \in V$ به صورت

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|,$$

تعریف می‌شود. نشان می‌دهیم که $\|.\|_1$ در شرایط تعریف ۱.۲ صدق می‌کند.

(الف) به روشنی داریم $\|x\| \geqslant 0$. از طرفی

$$\|x\|_1 = 0 \iff \sum_{i=1}^{n} |x_i| = 0 \iff |x_i| = 0,\ i = 1, \ldots, n \iff x = 0.$$

(ب) به ازای هر $\alpha \in \mathbb{C}$ و $x \in V$، داریم:

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

(ج) به ازای هر $x, y \in V$، داریم:

$$\|x+y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leqslant \sum_{i=1}^{n} (|x_i| + |y_i|) \leqslant \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1.$$

مثال ۲.۲ فرض کنید $V = \mathbb{C}^n$ و $x = (x_1, \ldots, x_n)^T \in V$. در این صورت،

(الف) نرم-۲ یا نرم اقلیدسی بردار x به صورت $\|x\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$, تعریف می‌شود.

(ب) نرم بینهایت بردار x به صورت $\|x\|_\infty = \max_i |x_i|$ تعریف می‌شود.

به عنوان یک تمرین نشان دهید که $\|.\|_2$ و $\|.\|_\infty$ در شرایط نرم صدق می‌کنند.

مثال ۳.۲ (الف) اگر $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$، آنگاه

$$\|e\|_\infty = 1, \quad \|e\|_1 = n, \quad \|e\|_2 = \sqrt{n}.$$

(ب) اگر $e = (1, 2, \ldots, n)^T \in \mathbb{R}^n$، آنگاه

$$\|e\|_\infty = n, \quad \|e\|_1 = \frac{1}{2} n(n+1), \quad \|e\|_2 = \sqrt{\frac{1}{6} n(n+1)(2n+1)}.$$

به ازای هر نرم $\|.\|$ روی فضای برداری V و $x, y \in V$، داریم:

$$||\|x\| - \|y\|| \leqslant \|x - y\|. \tag{۱.۲}$$

زیرا با استفاده از نامساوی مثلث، داریم:

$$\|x\| = \|(x-y) + y\| \leqslant \|x - y\| + \|y\|.$$

بنابراین $\|x - y\| \geqslant \|x\| - \|y\|$. از طرفی با تعویض نقش x و y، داریم:

$$\|x - y\| = \|y - x\| \geqslant \|y\| - \|x\|,$$

که نتیجهٔ مطلوب به دست می‌آید.

برای نرم‌های اقلیدسی، نرم بینهایت و نرم-۱ روی $\mathbb{C}^n$، داریم

$$\|x\|_{۲} \leqslant \|x\|_{۱} \leqslant \sqrt{n}\|x\|_{۲}, \tag{۲.۲}$$

$$\|x\|_{\infty} \leqslant \|x\|_{۲} \leqslant \sqrt{n}\|x\|_{\infty}, \tag{۳.۲}$$

$$\|x\|_{\infty} \leqslant \|x\|_{۱} \leqslant n\|x\|_{\infty}. \tag{۴.۲}$$

در اینجا رابطهٔ اول را ثابت کرده و دو رابطه دیگر را به عنوان تمرین به خواننده واگذار می‌کنیم. ابتدا می‌بینیم که به ازای هر i، $|x_i| \leqslant \|x\|_{۱}$. بنابراین

$$\|x\|_{۲} = (\sum_{i=۱}^{n} |x_i||x_i|)^{\frac{۱}{۲}} \leqslant (\sum_{i=۱}^{n} |x_i|\|x\|_{۱})^{\frac{۱}{۲}} = \|x\|_{۱}.$$

از طرفی با استفاده از نامساوی کوشی− شوارتز، برای هر $a_i, b_i \in \mathbb{C}$، $i = ۱, \ldots, n$، داریم (تمرین ۱.۲ را ببینید):

$$(\sum_{i=۱}^{n} \overline{a_i} b_i)^{۲} \leqslant \sum_{i=۱}^{n} |a_i|^{۲} \sum_{i=۱}^{n} |b_i|^{۲},$$

با استفاده از نامساوی کوشی− شوارتز داریم:

$$\|x\|_{۱} = \sum_{i=۱}^{n} ۱.|x_i| \leqslant (\sum_{i=۱}^{n} ۱^{۲})^{\frac{۱}{۲}} (\sum_{i=۱}^{n} |x_i|^{۲})^{\frac{۱}{۲}} = \sqrt{n}\|x\|_{۲}.$$

یکی از نرم‌های ماتریسی متداول نرم فروبنیوس است. اگر $A = (a_{ij}) \in \mathbb{C}^{m\times n}$، آنگاه نرم فروبنیوس آن به صورت زیر تعریف می‌شود:

$$\|A\|_F = (\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2)^{\frac{1}{2}} = \mathrm{trace}(A^H A)^{\frac{1}{2}}.$$

اگر سطر و ستون k ام ماتریس A را به ترتیب با A_{k*} و A_{*k} نشان دهیم، آنگاه خواهیم داشت:

$$\|A\|_F^2 = \sum_{i=1}^{m} \|A_{i*}\|_2^2 = \sum_{j=1}^{n} \|A_{*j}\|_2^2.$$

مثال ۴.۲ (الف) فرض کنید

$$A = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix},$$

که در آن $i = \sqrt{-1}$. در این صورت،

$$\|A\|_F = \sqrt{\left|\frac{1+i}{2}\right|^2 + \left|\frac{1-i}{2}\right|^2 + \left|\frac{1-i}{2}\right|^2 + \left|\frac{1+i}{2}\right|^2} = \sqrt{2}.$$

(ب) اگر I یک ماتریس همانی $n \times n$ باشد، آنگاه $\|I\|_F = \sqrt{n}$.

نشان می‌دهیم که $\|.\|_F$ در شرایط نرم ماتریسی صدق می‌کند. برای این کار، ابتدا عملگر vec را تعریف می‌کنیم. به ازای هر $A = (a_1, a_2, \ldots, a_n) \in \mathbb{C}^{m\times n}$، $\mathrm{vec}(A)$ که برداری در $\mathbb{R}^{mn}$ است به صورت

$$\mathrm{vec}(A) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

تعریف می‌شود. چند خاصیت مهم vec در لم بعدی معرفی می‌شود.

لم ۱.۲ اگر $A, B \in \mathbb{C}^{m\times n}$، آنگاه

(الف) به ازای هر $\alpha \in \mathbb{C}$، $\mathrm{vec}(\alpha A) = \alpha\ \mathrm{vec}(A)$.

(ب) $\mathrm{vec}(A + B) = \mathrm{vec}(A) + \mathrm{vec}(B)$.

(ج) $\|A\|_F = \|\mathrm{vec}(A)\|_2$.

برهان: اثبات این لم ساده است و به خواننده واگذار می‌شود. □

قضیه ۱.۲ نرم فروبنیوس در شرایط نرم ماتریسی صدق می‌کند.

برهان: (الف) برای هر $A \in \mathbb{C}^{m\times n}$، به روشنی داریم $\|A\| \geqslant 0$. به‌علاوه

$$\begin{aligned}\|A\|_F = 0 &\iff \sum_{j=1}^{n} \|A_{*j}\|_2^2 = 0\\ &\iff A_{*j} = 0,\quad j = 1,2,\dots,n\\ &\iff A = 0.\end{aligned}$$

(ب) به ازای هر $A \in \mathbb{C}^{m\times n}$ و $\alpha \in \mathbb{C}$، داریم:

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

(ج) با توجه به لم ۱.۲ به ازای هر $A, B \in \mathbb{C}^{m\times n}$، داریم:

$$\|A+B\|_F = \|\mathrm{vec}(A) + \mathrm{vec}(B)\|_2 \leqslant \|\mathrm{vec}(A)\|_2 + \|\mathrm{vec}(B)\|_2 = \|A\|_F + \|B\|_F,$$

که بدین ترتیب اثبات قضیه کامل می‌شود. □

تعریف ۲.۲ نرم ماتریسی $\|.\|$ را نسبت به نرم‌های برداری $\|.\|_v$ روی $\mathbb{C}^n$ و $\|.\|_w$ روی $\mathbb{C}^m$ سازگار گویند هرگاه به ازای هر $x \in \mathbb{C}^n$ و $A \in \mathbb{C}^{m\times n}$، $\|Ax\|_w \leqslant \|A\|\|x\|_v$.

تعریف ۳.۲ گوییم نرم ماتریسی $\|.\|$ روی $\mathbb{C}^{n\times n}$ خاصیت ضربی دارد هرگاه به ازای هر $A, B \in \mathbb{C}^{n\times n}$، $\|AB\| \leqslant \|A\|\|B\|$.

با استفاده از نامساوی کوشی– شوارتز، برای هر بردار $x \in \mathbb{C}^n$، داریم

$$|A_{i*}x|^2 \leqslant \|A_{i*}\|_2^2\|x\|_2^2.$$

در نتیجه

$$\|Ax\|_2^2 = \sum_{i=1}^{m} |A_{i*}x|^2 \leqslant \sum_{i=1}^{m} \|A_{i*}\|_2^2\|x\|_2^2 = \|A\|_F^2\|x\|_2^2.$$

یعنی

$$\|Ax\|_۲ \leqslant \|A\|_F \|x\|_۲. \tag{۵.۲}$$

رابطهٔ (۵.۲) نشان می‌دهد که نرم ماتریسی فروبنیوس نسبت به نرم برداری اقلیدسی سازگار است. همچنین با استفاده از (۵.۲) برای هر دو ماتریس A و B، داریم:

$$\|AB\|_F^۲ = \sum_{j=۱}^{n} \|(AB)_{*j}\|_۲^۲ = \sum_{j=۱}^{n} \|AB_{*j}\|_۲^۲ \leqslant \sum_{j=۱}^{n} \|A\|_F^۲ \|B_{*j}\|_۲^۲ = \|A\|_F^۲ \|B\|_F^۲.$$

یعنی $\|AB\|_F \leqslant \|A\|_F \|B\|_F$. در نتیجه نرم فروبنیوس دارای خاصیت ضربی است.

تعریف ۴.۲ **(نرم طبیعی)** فرض کنید $\|.\|$ یک نرم برداری روی $\mathbb{C}^n$ باشد و $A \in \mathbb{C}^{m\times n}$. در این صورت، نرم طبیعی متناظر به این نرم برداری به صورت

$$\|A\| = \max_{\|x\|\neq ۰} \frac{\|Ax\|}{\|x\|}, \tag{۶.۲}$$

تعریف می‌شود.

با توجه به اینکه $\left\|\frac{x}{\|x\|}\right\| = ۱$، داریم:

$$\|A\| = \max_{x\neq ۰} \frac{\|Ax\|}{\|x\|} = \max_{x\neq ۰} \left\|A\frac{x}{\|x\|}\right\| = \max_{\|x\|=۱} \|Ax\|. \tag{۷.۲}$$

قضیه ۲.۲ نرم تعریف شده در (۶.۲) در شرایط نرم ماتریسی صدق می‌کند، خاصیت ضربی دارد و با نرم برداری تعریف کننده آن سازگار است.

برهان: (الف) به وضوح $\|A\| \geqslant ۰$. اگر $A = ۰$، آنگاه به ازای هر بردار x، $Ax = ۰$ و در نتیجه $\|A\| = ۰$. حال فرض کنید که $\|A\| = ۰$. با توجه به رابطهٔ (۶.۲)، به ازای هر $x \neq ۰$، $Ax = ۰$. با انتخاب $x = e_j$، $j = ۱,\dots,n$، داریم $A_{*j} = ۰$. یعنی $A = ۰$.

(ب) فرض کنید $\alpha \in \mathbb{C}$. در این صورت،

$$\|\alpha A\| = \max_{\|x\|=۱} \|\alpha Ax\| = \max_{\|x\|=۱} |\alpha|\|Ax\| = |\alpha|\|A\|.$$

(ج) فرض کنید $A, B \in \mathbb{C}^{m\times n}$. در این صورت،

$$\begin{aligned} \|A+B\| &= \max_{\|x\|=1} \|(A+B)x\| \leqslant \max_{\|x\|=1} (\|Ax\|+\|Bx\|) \\ &\leqslant \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\|+\|B\|. \end{aligned}$$

از طرفی به ازای هر بردار $0 \neq y \in \mathbb{C}^n$، داریم:

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \geqslant \frac{\|Ay\|}{\|y\|}.$$

در نتیجه $\|Ay\| \leqslant \|A\|\|y\|$. به وضوح این رابطه به ازای $y = 0$ نیز درست است. بنابراین سازگاری نرم طبیعی با نرم تولید کننده آن ثابت می‌شود. با استفاده از سازگاری، داریم:

$$\begin{aligned} \|AB\| &= \max_{\|x\|=1} \|A(Bx)\| \leqslant \max_{\|x\|=1} \|A\|\|Bx\| \\ &\leqslant \max_{\|x\|=1} \|A\|\|B\|\|x\| = \|A\|\|B\|. \end{aligned}$$

این رابطه خاصیت ضربی نرم طبیعی را نشان می‌دهد. □

نتیجه ۱.۲ اگر I ماتریس همانی باشد، آنگاه برای هر نرم طبیعی $\|.\|$، داریم $\|I\| = 1$.

برهان: داریم

$$\|I\| = \max_{x\neq 0} \frac{\|Ix\|}{\|x\|} = \max_{x\neq 0} 1 = 1,$$

که اثبات نتیجه را کامل می‌کند. □

مسأله ۱.۲ نشان دهید $\|.\|_F$ در حالت کلی نمی‌تواند یک نرم طبیعی باشد.

حل: با توجه به نتیجهٔ ۱.۲ نرم طبیعی ماتریس همانی برابر با ۱ است، اما نرم فروبنیوس ماتریس همانی $I \in \mathbb{C}^{n\times n}$ برابر با $\sqrt{n}$ است. بنابراین به ازای $n \geqslant 2$ نرم فروبنیوس نمی‌تواند یک نرم طبیعی باشد.

تعریف ۵.۲ فرض کنید $A \in \mathbb{C}^{n\times n}$. گوییم $\lambda \in \mathbb{C}$ یک مقدار ویژهٔ متناظر با بردار ویژهٔ x برای ماتریس A است، هرگاه $Ax = \lambda x$ و $x \neq 0$. در این صورت، (λ, x) را یک زوج ویژهٔ A می‌گویند.

تعریف ۶.۲ مجموعهٔ مقادیر ویژهٔ ماتریس A را طیف A نامیده و با $\sigma(A)$ نشان می‌دهند. به‌علاوه شعاع طیفی A را با $\rho(A)$ نشان می‌دهند و بنا به تعریف $\rho(A) = \max\limits_{\lambda\in\sigma(A)} |\lambda|$.

در حالت کلی محاسبهٔ نرم طبیعی یک ماتریس با استفاده از رابطهٔ (۶.۲) یا (۷.۲) به کمک روش‌های متداول بهینه‌سازی بسیار دشوار است. امّا با استفاده از قضیهٔ بعدی نرم‌های طبیعی تولید شده توسط نرم-۱، نرم−۲ و نرم بینهایت را می‌توان محاسبه کرد.

قضیه ۳.۲ برای هر ماتریس $A \in \mathbb{C}^{m\times n}$، داریم

$$\begin{aligned}
\|A\|_1 &= \max_{\|x\|_1=1} \|Ax\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|,\\
\|A\|_2 &= \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\rho(A^H A)},\\
\|A\|_\infty &= \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|.
\end{aligned}$$

برهان: اثبات رابطهٔ دوم در فصل ۴ (بعد از نتیجهٔ ۳.۴) ارائه خواهد شد و رابطهٔ سوم، به عنوان تمرین به خواننده واگذار می‌شود. برای اثبات رابطهٔ اول، فرض می‌کنیم $\|x\|_1 = 1$. در این صورت،

$$\begin{aligned}
\|Ax\|_1 &= \|\sum_{j=1}^{n} A_{*j}x_j\|_1 \leqslant \sum_{j=1}^{n} \|A_{*j}x_j\|_1\\
&\leqslant \sum_{j=1}^{n} \|A_{*j}\|_1 |x_j| \leqslant \left(\max_j \|A_{*j}\|_1\right)\left(\sum_{j=1}^{n} |x_j|\right)\\
&= \max_j \sum_{i=1}^{m} |a_{ij}|. \qquad (۸.۲)
\end{aligned}$$

از طرفی فرض کنید $\max\limits_j \sum\limits_{i=1}^{m} |a_{ij}| = \sum\limits_{i=1}^{m} |a_{ik}|$. قرار می دهیم $x = e_k$ که در آن e_k ستون k ام ماتریس همانی $n \times n$ است. در این صورت، خواهیم داشت:

$$\|Ax\|_1 = \|Ae_k\|_1 = \|A_{*k}\|_1 = \sum_{i=1}^{m} |a_{ik}| = \max_j \sum_{i=1}^{m} |a_{ij}|. \qquad (۹.۲)$$

حال با استفاده از رابطه‌های (۸.۲) و (۹.۲) نتیجهٔ لازم به دست می آید. □

مثال ۵.۲ فرض کنید:

$$A = \begin{pmatrix} -۱ & -۲ & ۰ & -۲ \\ ۲ & -۱ & ۰ & ۴ \\ ۰ & ۰ & ۱ & ۰ \\ -۲ & ۰ & ۰ & ۵ \end{pmatrix}.$$

داریم:

$$A^T A = \begin{pmatrix} ۹ & ۰ & ۰ & ۰ \\ ۰ & ۵ & ۰ & ۰ \\ ۰ & ۰ & ۱ & ۰ \\ ۰ & ۰ & ۰ & ۴۵ \end{pmatrix}.$$

بنابراین مقادیر ویژهٔ ماتریس $A^T A$ به صورت $\{۹, ۵, ۱, ۴۵\}$ است و در نتیجه $\|A\|_۲ = \sqrt{۴۵}$. همچنین داریم:

$$\begin{aligned} \|A\|_۱ &= \max\{|-۱| + |۲| + |۰| + |-۲|, |-۲| + |-۱| + |۰| + |۰|, \\ &\quad |۰| + |۰| + |۱| + |۰|, |-۲| + |۴| + |۰| + |۵|\} \\ &= \max\{۵, ۳, ۱, ۱۱\} = ۱۱, \\ \|A\|_\infty &= \max\{|-۱| + |-۲| + |۰| + |-۲|, |۲| + |-۱| + |۰| + |۴|, \\ &\quad |۰| + |۰| + |۱| + |۰|, |-۲| + |۰| + |۰| + |۵|\} \\ &= \max\{۵, ۷, ۱, ۷\} = ۷. \end{aligned}$$

نتیجه ۲.۲ اگر $A \in \mathbb{C}^{n\times n}$ هرمیتی باشد، آنگاه $\|A\|_۲ = \rho(A)$.

برهان: فرض کنید $A^H = A$. لذا

$$\|A\|_۲ = \sqrt{\rho(A^H A)} = \sqrt{\rho(A^۲)} = \sqrt{\rho(A)^۲} = \rho(A),$$

و بدین ترتیب اثبات نتیجه، کامل می‌شود. □

قضیه ۴.۲ اگر λ یک مقدار ویژهٔ ماتریس A باشد، آنگاه برای هر نرم طبیعی $\|.\|$، داریم $|\lambda| \leqslant \|A\|$.

برهان: فرض کنید (λ, x) یک زوج ویژهٔ ماتریس A باشد. در این صورت

$$\lambda x = Ax \Longrightarrow |\lambda| \|x\| = \|Ax\| \leqslant \|A\| \|x\| \Longrightarrow |\lambda| \leqslant \|A\|,$$

که اثبات قضیه را کامل می‌کند. $\square$

نتیجه ۳.۲ برای هر ماتریس $A \in \mathbb{C}^{n\times n}$ و هر نرم طبیعی $\|.\|$، $\rho(A) \leqslant \|A\|$.

قضیه ۵.۲ اگر $A \in \mathbb{C}^{m\times n}$، آنگاه

(الف) $\frac{1}{\sqrt{n}}\|A\|_\infty \leqslant \|A\|_2 \leqslant \sqrt{m}\|A\|_\infty$.

(ب) $\|A\|_2 \leqslant \|A\|_F \leqslant \sqrt{n}\|A\|_2$.

(ج) $\frac{1}{\sqrt{m}}\|A\|_1 \leqslant \|A\|_2 \leqslant \sqrt{n}\|A\|_1$.

برهان: در اینجا قسمت (الف) را اثبات کرده و دو قسمت دیگر را به عنوان تمرین به خواننده واگذار می‌کنیم.

(الف) با استفاده از رابطهٔ (۳.۲)، داریم:

$$\|Ax\|_\infty \leqslant \|Ax\|_2, \qquad \|x\|_2 \leqslant \sqrt{n}\|x\|_\infty.$$

پس از این دو رابطه نتیجه می‌شود که

$$\frac{\|Ax\|_\infty}{\sqrt{n}\|x\|_\infty} \leqslant \frac{\|Ax\|_2}{\|x\|_2}.$$

در نتیجه

$$\frac{1}{\sqrt{n}} \max_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \leqslant \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}.$$

لذا،

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leqslant \|A\|_2.$$

از طرفی با استفاده از معادل بودن نرم‌های برداری نتیجه می‌شود که

$$\|Ax\|_2 \leqslant \sqrt{m}\|Ax\|_\infty, \qquad \|x\|_\infty \leqslant \|x\|_2.$$

با استفاده از این دو رابطه می‌توان نوشت:

$$\frac{\|Ax\|_2}{\|x\|_2} \leqslant \sqrt{m}\frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

در نتیجه

$$\max_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} \leqslant \sqrt{m}\max_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

از این‌رو،

$$\|A\|_2 \leqslant \sqrt{m}\|A\|_\infty.$$

به این ترتیب اثبات قسمت (الف) کامل می‌شود. □

مسأله ۲.۲ به ازای هر $A \in \mathbb{C}^{n\times n}$ تعریف کنید $\|A\|_{\max} = \max_{i,j}|a_{ij}|$. نشان دهید $\|.\|_{\max}$ در شرایط نرم صدق می‌کند اما خاصیت ضربی ندارد.

حل: (الف) روشن است که به ازای هر ماتریس $A = (a_{ij})$، $\|A\|_{\max} \geqslant 0$. همچنین داریم:

$$\|A\|_{\max} = 0 \iff |a_{ij}| = 0 \quad \forall i,j \iff A = 0.$$

(ب) به ازای هر $\alpha \in \mathbb{C}$ و هر ماتریس $A = (a_{ij})$، داریم:

$$\|\alpha A\|_{\max} = \max_{ij}|\alpha a_{ij}| = |\alpha|\max_{ij}|a_{ij}| = |\alpha|\|A\|_{\max}.$$

(ج) به ازای هر دو ماتریس A و B، داریم:

$$\|A+B\| = \max_{ij}(|a_{ij}+b_{ij}|) \leqslant \max_{ij}|a_{ij}| + \max_{ij}|b_{ij}| = \|A\|_{\max} + \|B\|_{\max}.$$

حال فرض کنید:

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

داریم $\|A\|_{\max} = \|B\|_{\max} = 2$. از طرفی داریم:

$$AB = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \qquad \|AB\|_{\max} = 5.$$

بنابراین

$$\|AB\|_{\max} > \|A\|_{\max}\|B\|_{\max} = 4,$$

که نشان می‌دهد نرم تعریف شده خاصیت ضربی ندارد.

۲.۲ ضرب داخلی

این بخش را با تعریف ضرب داخلی شروع می‌کنیم.

تعریف ۷.۲ یک ضرب داخلی روی فضای برداری V تعریف شده روی میدان $\mathbb{F}$، تابعی است که به هر زوج از بردارهای x و y در V یک اسکالر مثل $\langle x,y\rangle$ در $\mathbb{F}$ را نسبت می‌دهد به طوری که در شرایط زیر صدق کند:

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

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

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

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

هر فضای برداری که یک ضرب داخلی در آن تعریف شده باشد را یک فضای ضرب داخلی می‌نامند. به ازای هر $\alpha \in \mathbb{F}$ و $x,y \in V$ داریم:

$$\langle x,\alpha y\rangle = \overline{\langle \alpha y,x\rangle} = \overline{\alpha\langle y,x\rangle} = \overline{\alpha}\,\overline{\langle y,x\rangle} = \overline{\alpha}\langle x,y\rangle,$$

و به ازای هر $x,y,z \in V$، داریم:

$$\langle x+y,z\rangle = \overline{\langle z,x+y\rangle} = \overline{\langle z,x\rangle + \langle z,y\rangle} = \overline{\langle z,x\rangle} + \overline{\langle z,y\rangle} = \langle x,z\rangle + \langle y,z\rangle.$$

مثال ۶.۲ ضرب داخلی استاندارد روی $\mathbb{R}^n$ و $\mathbb{C}^n$ به ترتیب به صورت $\langle x,y\rangle = x^T y$ و $\langle x,y\rangle = x^H y$ تعریف می‌شوند. نشان می‌دهیم که $\langle x,y\rangle = x^H y$ در شرایط ضرب داخلی صدق می‌کند.

(الف) برای هر $x \in \mathbb{C}^n$، داریم:

$$\langle x,x\rangle = x^H x = \sum_{i=1}^{n} \bar{x}_i x_i = \sum_{i=1}^{n} |x_i|^2 \geqslant 0.$$

همچنین

$$\langle x,x\rangle = 0 \iff \sum_{i=1}^{n} |x_i|^2 = 0 \iff x_i = 0,\ i = 1,\dots,n \iff x = 0.$$

(ب) برای هر $x, y \in \mathbb{C}^n$ و $\alpha \in \mathbb{C}$، داریم:

$$\langle x, \alpha y\rangle = x^H(\alpha y) = \alpha x^H y = \alpha\langle x, y\rangle.$$

(ج) برای هر $x, y, z \in \mathbb{C}^n$،

$$\langle x, y+z\rangle = x^H(y+z) = x^H y + x^H z = \langle x, y\rangle + \langle x, z\rangle.$$

(د) برای هر $x, y \in \mathbb{C}^n$، داریم:

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

مثال ۷.۲ اگر $A \in \mathbb{C}^{n\times n}$ یک ماتریس نامنفرد باشد، آنگاه $\langle x, y\rangle = x^H A^H A y$ یک ضرب داخلی روی $\mathbb{C}^n$ تعریف می‌کند که آن را $A^H A$- ضرب داخلی نامند.

حل: (الف) برای هر $x \in \mathbb{C}^n$، داریم:

$$\langle x, x\rangle = x^H A^H A x = (Ax)^H A x = \|Ax\|_2^2 \geqslant 0.$$

همچنین

$$\langle x, x\rangle = 0 \iff \|Ax\|_2^2 = 0 \iff Ax = 0 \iff x = 0.$$

(ب) برای هر $x, y \in \mathbb{C}^n$ و $\alpha \in \mathbb{C}$، داریم:

$$\langle x, \alpha y\rangle = x^H A^H A(\alpha y) = \alpha x^H A^H A y = \alpha\langle x, y\rangle.$$

(ج) برای هر $x, y, z \in \mathbb{C}^n$،

$$\langle x, y+z\rangle = x^H A^H A(y+z) = x^H A^H A y + x^H A^H A z = \langle x, y\rangle + \langle x, z\rangle.$$

(د) برای هر $x, y \in \mathbb{C}^n$،

$$\overline{\langle y, x\rangle} = \overline{y^H A^H A x} = (y^H A^H A x)^H = x^H A^H A y = \langle x, y\rangle.$$

مثال ۸.۲ توابع $\langle A, B\rangle = \mathrm{trace}(A^T B)$ و $\langle A, B\rangle = \mathrm{trace}(A^H B)$ به ترتیب ضرب داخلی (استاندارد) روی $\mathbb{R}^{m\times n}$ و $\mathbb{C}^{m\times n}$ تعریف می‌کنند.

حل: در اینجا نشان می‌دهیم که $\langle A, B\rangle = \mathrm{trace}(A^H B)$ در شرایط ضرب داخلی صدق می‌کند. روشن است که در این صورت، نتیجه روی $\mathbb{R}^{m\times n}$ نیز برقرار است.
(الف) برای هر $A \in \mathbb{C}^{m\times n}$، داریم:

$$\langle A, A\rangle = \mathrm{trace}(A^H A) = \|A\|_F^2 \geqslant 0.$$

همچنین

$$\langle A, A\rangle = 0 \Longleftrightarrow \|A\|_F^2 = 0 \Longleftrightarrow A = 0.$$

(ب) برای هر $A, B \in \mathbb{C}^{m\times n}$ و $\alpha \in \mathbb{C}$، داریم:

$$\langle A, \alpha B\rangle = \mathrm{trace}(A^H(\alpha B)) = \alpha\ \mathrm{trace}(A^H B) = \alpha\langle A, B\rangle.$$

(ج) برای هر $A, B, C \in \mathbb{C}^{m\times n}$، داریم

$$\begin{aligned}\langle A, B + C\rangle &= \mathrm{trace}(A^H(B + C)) = \mathrm{trace}(A^H B + A^H C)) \\ &= \mathrm{trace}(A^H B) + \mathrm{trace}(A^H C)) = \langle A, B\rangle + \langle A, C\rangle.\end{aligned}$$

(د) برای هر $A, B \in \mathbb{C}^{m\times n}$، و با توجه تعریف اثر و خواص آن که در فصل ۱ به آنها اشاره شد، داریم:

$$\begin{aligned}\overline{\langle B, A\rangle} &= \overline{\mathrm{trace}(B^H A)} = \overline{\mathrm{trace}(A^H B)^H} = \overline{\overline{\mathrm{trace}(A^H B)}} \\ &= \mathrm{trace}(A^H B) = \langle A, B\rangle.\end{aligned}$$

هر ضرب داخلی روی فضای برداری V یک نرم تعریف می‌کند. در واقع اگر $\langle .,.\rangle$ یک ضرب داخلی روی V باشد، آنگاه $\|.\| = \sqrt{\langle .,.\rangle}$ نرم متناظر روی V است. برای اثبات این ادعا ابتدا قضیهٔ زیر را ثابت می‌کنیم.

قضیه ۶.۲ (نامساوی کوشی- شوارتز). اگر V یک فضای ضرب داخلی باشد و $\|.\| = \sqrt{\langle .,.\rangle}$، آنگاه برای هر $x, y \in V$، $|\langle x, y\rangle| \leqslant \|x\|\|y\|$.

برهان: روشن است که اگر $x=0$ یا $y=0$، آنگاه نامساوی برقرار است. بنابراین فرض می‌کنیم هیچ‌کدام از این دو بردار صفر نباشند. در این صورت، داریم:

$$\begin{aligned}
0 \leqslant \|x-\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\,y\|^2 &= (x-\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\,y, x-\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\,y)\\
&= \langle x,x\rangle-\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\langle x,y\rangle-\frac{\langle x,y\rangle}{\|y\|^2}\langle y,x\rangle+\frac{\langle x,y\rangle}{\|y\|^2}\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\langle y,y\rangle\\
&= \|x\|^2-\frac{|\langle x,y\rangle|^2}{\|y\|^2}-\frac{|\langle x,y\rangle|^2}{\|y\|^2}+\frac{|\langle x,y\rangle|^2}{\|y\|^2}\\
&= \|x\|^2-\frac{|\langle x,y\rangle|^2}{\|y\|^2}.
\end{aligned}$$

این رابطه اثبات قضیه را کامل می‌کند. □

قضیه ۷.۲ اگر V یک فضای ضرب داخلی با ضرب داخلی $\langle x,y\rangle$ باشد، آنگاه $\|.\|=\sqrt{\langle .,.\rangle}$ یک نرم روی V تعریف می‌کند.

برهان: (الف) با توجه به $\langle x,x\rangle \geqslant 0$ نتیجه می‌شود $\|x\| \geqslant 0$. داریم $\langle x,x\rangle=0$ اگر و تنها اگر $x=0$، که نتیجه می‌شود $\|x\|=0$ اگر و تنها اگر $x=0$.
(ب) فرض کنید α یک اسکالر باشد. در این صورت،

$$\|\alpha x\|=\sqrt{\langle \alpha x,\alpha x\rangle}=\sqrt{\alpha\overline{\alpha}\langle x,x\rangle}=\sqrt{|\alpha|^2\langle x,x\rangle}=|\alpha|\sqrt{\langle x,x\rangle}=|\alpha|\|x\|.$$

(ج) داریم $|\langle x,y\rangle| \leqslant \|x\|\|y\|$ و $\Re\langle x,y\rangle \leqslant |\langle x,y\rangle|$. در نتیجه

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y,x+y\rangle=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle\\
&= \langle x,x\rangle+2\Re\langle x,y\rangle+\langle y,y\rangle \leqslant \|x\|^2+2|\langle x,y\rangle|+\|y\|^2\\
&\leqslant \|x\|^2+2\|x\|\|y\|+\|y\|^2=(\|x\|+\|y\|)^2,
\end{aligned}$$

که اثبات قضیه را کامل می‌کند. □

مثال ۹.۲ در این مثال نرم‌های تولید شده توسط ضرب‌های داخلی معرفی شده در مثال‌های ٦.۲، ۷.۲ و ۸.۲ را مورد مطالعه قرار می‌دهیم.

• نرم تولید شده توسط ضرب داخلی $\langle x,y\rangle = x^T y$ روی $\mathbb{R}^n$ و نرم تولید شده توسط $\langle x,y\rangle = x^H y$ روی $\mathbb{C}^n$ نرم‌های اقلیدسی $\|.\|_۲$ هستند.

• فرض کنید $A \in \mathbb{C}^{n\times n}$ یک ماتریس نامنفرد باشد. در این صورت، ضرب داخلی $\langle x,y\rangle = x^H A^H A y$ روی $\mathbb{C}^n$ نرم

$$\|x\|_{A^H A} = \sqrt{\langle x,x\rangle} = \sqrt{x^H A^H A x} = \|Ax\|_۲,$$

را تولید می‌کند.

• ضرب داخلی $(A,B) = \mathrm{trace}(A^H B)$ روی $\mathbb{C}^{m\times n}$ ، نرم

$$\|A\| = \sqrt{\langle A,A\rangle} = \sqrt{\mathrm{trace}(A^H A)} = \|A\|_F,$$

را تعریف می‌کند.

قضیه ۸.۲ (اتحاد متوازی الاضلاع). اگر $\|.\|$ نرم تولید شده توسط ضرب داخلی $\langle .,.\rangle$ روی V باشد، آنگاه

$$\|x+y\|^۲ + \|x-y\|^۲ = ۲(\|x\|^۲ + \|y\|^۲).$$

برهان: داریم:

$$\begin{aligned}\|x+y\|^۲ + \|x-y\|^۲ &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= ۲\langle x,x\rangle + ۲\langle y,y\rangle = ۲(\|x\|^۲ + \|y\|^۲),\end{aligned}$$

که اثبات قضیه را کامل می‌کند. □

تعریف ۸.۲ دو بردار x و y را در فضای ضرب داخلی V با ضرب داخلی $\langle .,.\rangle$ متعامد می‌گویند و با $x \perp y$ نشان می‌دهند هرگاه $\langle x,y\rangle = ۰$.

در فضای برداری $\mathbb{R}^n$ با ضرب داخلی استاندارد

$$x \perp y \Longleftrightarrow x^T y = ۰,$$

و در فضای برداری $\mathbb{C}^n$ با ضرب داخلی استاندارد

$$x \perp y \Longleftrightarrow x^H y = ۰.$$

مثال ۱۰.۲ دو بردار $x = \begin{pmatrix} ۱ \\ ۱ \end{pmatrix}$ و $y = \begin{pmatrix} ۱ \\ -۱ \end{pmatrix}$ در $\mathbb{R}^n$ و $\mathbb{C}^n$ با ضرب داخلی استاندارد بر هم عمودند، زیرا $x^H y = x^T y = ۰$. حال فرض کنید که $x = \begin{pmatrix} i \\ ۱ \end{pmatrix}$ و $y = \begin{pmatrix} i \\ ۱ \end{pmatrix}$ (در اینجا $i = \sqrt{-۱}$). با وجود اینکه $x^T y = ۰$ این دو بردار در $\mathbb{C}^n$ با ضرب داخلی استاندارد بر هم عمود نیستند، چون $x^H y = ۲ \neq ۰$.

تعریف ۹.۲ مجموعه بردارهای $B = \{u_۱, u_۲, \ldots, u_n\}$ را یک مجموعهٔ یکامتعامد نامند هرگاه به ازای هر i، $\|u_i\| = ۱$ و به ازای هر i و j با شرط $i \neq j$، داشته باشیم $u_i \perp u_j$. به عبارت دیگر

$$\langle u_i, u_j \rangle = \begin{cases} ۱, & i = j, \\ ۰, & i \neq j. \end{cases}$$

قضیه ۹.۲ هر مجموعه از بردارهای یکامتعامد، مستقل خطی‌اند. به‌علاوه هر مجموعه از بردارهای یکامتعامد شامل n بردار از فضای برداری V با بعد n، یک پایه یکامتعامد برای V تشکیل می‌دهند.

برهان: برای اثبات قسمت اول فرض کنید که $B = \{u_۱, u_۲, \ldots, u_n\}$ یک مجموعه از بردارهای یکامتعامد باشند و $\alpha_۱ u_۱ + \alpha_۲ u_۲ + \cdots + \alpha_n u_n = ۰$. در این صورت، با استفاده از خواص ضرب داخلی، داریم:

$$\begin{aligned} ۰ &= \langle u_i, ۰ \rangle = \langle u_i, \alpha_۱ u_۱ + \alpha_۲ u_۲ + \cdots + \alpha_n u_n \rangle \\ &= \alpha_۱ \langle u_i, u_۱ \rangle + \cdots + \alpha_i \langle u_i, u_i \rangle + \cdots + \alpha_n \langle u_i, u_n \rangle = \alpha_i \|u_i\|^۲ = \alpha_i. \end{aligned}$$

قسمت دوم قضیه، یک نتیجه از قسمت اول است. $\square$

فرض کنید که $B = \{u_۱, u_۲, \ldots, u_n\}$ یک پایهٔ یکامتعامد برای فضای برداری V باشد و $x \in V$. در این صورت اسکالرهای α_i، $i = ۱, ۲, \ldots, n$ وجود دارند به طوری که

$$x = \sum_{j=۱}^{n} \alpha_j u_j.$$

پس به ازای هر i، داریم:

$$\langle u_i, x\rangle = \langle u_i, \sum_{j=1}^{n} \alpha_j u_j\rangle = \sum_{j=1}^{n} \alpha_j \langle u_i, u_j\rangle = \alpha_i \|u_i\|^2 = \alpha_i.$$

بنابراین

$$x = \langle u_1, x\rangle u_1 + \langle u_2, x\rangle u_2 + \cdots + \langle u_n, x\rangle u_n.$$

این بسط را بسط فوریهٔ x نامند.

مثال ۱۱.۲ مجموعه بردارهای $\{u_1, u_2, u_3\}$ را در نظر بگیرید که

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

این بردارها متعامدند، اما یکه نیستند. با توجه به اینکه

$$\|u_1\|_2 = \sqrt{2}, \quad \|u_2\|_2 = \sqrt{3}, \quad \|u_3\|_2 = \sqrt{6}.$$

قرار می‌دهیم:

$$\tilde{u}_1 = \frac{1}{\sqrt{2}} u_1, \quad \tilde{u}_2 = \frac{1}{\sqrt{3}} u_2, \quad \tilde{u}_3 = \frac{1}{\sqrt{6}} u_3.$$

در این صورت، مجموعه بردارهای $\{\tilde{u}_1, \tilde{u}_2, \tilde{u}_3\}$ یکامتعامد هستند. برای مثال اگر فرض کنید $x = (-1, 2, 1)^T$، آنگاه

$$\alpha_1 = \langle \tilde{u}_1, x\rangle = \frac{-3}{\sqrt{2}}, \quad \alpha_2 = \langle \tilde{u}_2, x\rangle = \frac{2}{\sqrt{3}}, \quad \alpha_3 = \langle \tilde{u}_3, x\rangle = \frac{1}{\sqrt{6}}.$$

بنابراین $x = \alpha_1 \tilde{u}_1 + \alpha_2 \tilde{u}_2 + \alpha_3 \tilde{u}_3$.

در ادامه می‌خواهیم که چگونه می‌توان با داشتن یک پایهٔ دلخواه برای یک فضای برداری یک پایهٔ یکامتعامد برای آن ساخت.

قضیه ۱۰.۲ فرض کنید V یک فضای ضرب داخلی با ضرب داخلی $\langle .,. \rangle$ روی میدان $\mathbb{F}$ باشد. اگر مجموعهٔ $\mathcal{B} = \{x_1, x_2, \ldots, x_n\}$ یک پایه برای V باشد، آنگاه مجموعهٔ

$\mathcal{B}' = \{u_1, u_2, \dots, u_n\}$ یک پایهٔ یکامتعامد برای V تشکیل می‌دهد که v_i توسط رابطهٔ بازگشتی

$$u_1 = \frac{x_1}{\|x_1\|}, \quad u_k = \frac{x_k - \sum_{i=1}^{k-1} \langle x_i, u_i \rangle u_i}{\|x_k - \sum_{i=1}^{k-1} \langle x_i, u_i \rangle u_i\|}, \quad k = 2, \dots, n,$$

تولید می‌شوند.

برهان: با استقراء این قضیه را ثابت می‌کنیم. در واقع با استقراء نشان می‌دهیم که به ازای هر k، مجموعهٔ $\mathcal{B}'_k = \{u_1, u_2, \dots, u_k\}$ یک پایهٔ یکامتعامد برای مجموعهٔ $\mathrm{span}(\mathcal{B}_k)$ تشکیل می‌دهد که در آن $\mathcal{B}_k = \{x_1, x_2, \dots, x_k\}$. واضح است که به ازای $k = 1$، داریم $\mathrm{span}\{x_1\} = \mathrm{span}\{u_1\}$. حال فرض کنید که حکم برای $k = m$ برقرار باشد. نشان می‌دهیم که مجموعهٔ $\mathcal{B}'_{m+1}$ یک پایهٔ یکامتعامد برای $\mathrm{span}(\mathcal{B}_{m+1})$ تشکیل می‌دهد. ابتدا نشان می‌دهیم $\mathcal{B}'_{m+1}$ یک مجموعهٔ یکامتعامد است. برای این کار، کافی است نشان دهیم که بردار u_{m+1} یکه است و بر u_j، $j = 1, \dots, m$ عمود است. بنا به شیوهٔ ساخت u_kها، بردار u_{m+1} یکه است و برای هر $j = 1, \dots, m$، داریم:

$$\begin{aligned}
\langle u_j, u_{m+1} \rangle &= \langle u_j, x_{m+1} - \sum_{i=1}^{m} \langle u_i, x_{m+1} \rangle u_i \rangle \\
&= \langle u_j, x_{m+1} \rangle - \sum_{i=1}^{m} \langle u_i, x_{m+1} \rangle \langle u_j, u_i \rangle \\
&= \langle u_j, x_{m+1} \rangle - \langle u_j, x_{m+1} \rangle \langle u_j, u_j \rangle = 0.
\end{aligned}$$

برای تکمیل اثبات نشان می‌دهیم که $\mathrm{span}(\mathcal{B}_{m+1}) = \mathrm{span}(\mathcal{B}'_{m+1})$. برای این کار، کافی است نشان دهیم که $x_{m+1} \in \mathrm{span}(\mathcal{B}'_{k+1})$ و $u_{m+1} \in \mathrm{span}(\mathcal{B}_{k+1})$. با توجه به نحوهٔ ساخت u_k ها داریم

$$w_{m+1} u_{m+1} = x_{m+1} - \sum_{i=1}^{m} \langle u_i, x_{m+1} \rangle u_i,$$

یا

$$x_{m+1} = w_{m+1}u_{m+1} + \sum_{i=1}^{m}\langle u_i, x_{m+1}\rangle u_i \in \mathrm{span}(\mathcal{B}'_{m+1}),$$

که در آن $w_{m+1} = \|x_{m+1} - \sum_{i=1}^{m}\langle u_i, x_{m+1}\rangle u_i\|$. از طرفی بنا به فرض استقراء، داریم:

$$\sum_{i=1}^{m}\langle u_i, x_{m+1}\rangle u_i \in \mathrm{span}(\mathcal{B}'_m) = \mathrm{span}(\mathcal{B}_m) \subset \mathcal{B}_{m+1}.$$

این رابطه و اینکه $x_{m+1} \in \mathrm{span}(\mathcal{B}_{m+1})$، نشان می‌دهد:

$$u_{m+1} = \frac{1}{w_{m+1}}\langle x_{m+1} - \sum_{i=1}^{m}\langle u_i, x_{m+1}\rangle u_i\rangle \in \mathrm{span}(\mathcal{B}_{m+1}),$$

بدین ترتیب اثبات قضیه کامل می‌گردد. □

با استفاده از قضیهٔ ۱۰.۲ الگوریتم زیر را ارائه می‌نمایم که به الگوریتم متعامد سازی گرام- اشمیت معروف است.

الگوریتم ۱.۲: متعامدسازی گرام- اشمیت

1. $u_1 := \dfrac{x_1}{\|x_1\|}$
2. For $k = 2, \ldots, n$, Do
3. $\quad w_k := x_k - \sum_{i=1}^{k-1}\langle u_i, x_k\rangle u_i$
4. $\quad u_k := \dfrac{w_k}{\|w_k\|}$
5. EndDo

مثال ۱۲.۲ با استفاده از الگوریتم گرام- اشمیت یک پایهٔ یکامتعامد برای مجموعهٔ $V = \mathrm{span}\{x_1, x_2, x_3\}$ به دست آورید که در آن

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

ضرب داخلی را ضرب داخلی استاندارد بگیرید.

حل: داریم $\|x_1\|_2 = \sqrt{2}$. بنابراین

$$u_1 = \frac{x_1}{\|x_1\|_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

از طرفی داریم $u_1^T x_2 = \sqrt{2}$. لذا

$$w_2 = x_2 - (u_1^T x_2) u_1 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}.$$

در نتیجه

$$u_2 = \frac{w_2}{\|w_2\|_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

حال برای محاسبهٔ u_3 داریم $u_1^T x_3 = 2\sqrt{2}$ و $u_2^T x_3 = 1$. از این‌رو

$$w_3 = x_3 - (u_1^T x_3) u_1 - (u_2^T x_3) u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

با استفاده از این بردار، داریم:

$$u_3 = \frac{w_3}{\|w_3\|_2} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

مثال ۱۳.۲ یک پایهٔ یکامتعامد برای $V = \mathrm{span}\{x_1, x_2, x_3\}$ بیابید که در آن

$$x_1 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}.$$

ضرب داخلی را ضرب داخلی استاندارد بگیرید.

حل: داریم $\|x_1\|_2 = 3$. بنابراین

$$u_1 = \frac{x_1}{\|x_1\|_2} = \frac{1}{3}\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}.$$

از طرفی داریم $u_1^T x_2 = 2$. در نتیجه

$$w_2 = x_2 - (u_1^T x_2)u_1 = \frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

داریم $\|w_2\|_2 = 1$. لذا

$$u_2 = \frac{w_2}{\|w_2\|_2} = \frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

امّا برای محاسبهٔ u_3، داریم $u_1^T x_3 = 5$ و $u_2^T x_3 = 1$. از این‌رو

$$w_3 = x_3 - (u_1^T x_3)u_1 - (u_2^T x_3)u_2 = x_3 - 5u_1 - u_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

این رابطه نشان می‌دهد که $x_3 = 5u_1 + u_2$. بنابراین x_3 وابسته به دو بردار u_1 و u_2 و در نتیجه وابسته به x_1 و x_2 است. بنابراین $\{u_1, u_2\}$ یک پایهٔ یکامتعامد برای $\operatorname{span}\{x_1, x_2, x_3\}$ است.

مثال ۱۴.۲ فرض کنید

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} -1 & 2 & 1 \\ 3 & -6 & 1 \end{pmatrix}.$$

یک پایهٔ یکامتعامد برای $\mathcal{R}(S)$ بیابید. ضرب داخلی را $\langle x, y\rangle = x^H A^H A y$ بگیرید.

حل: می‌دانیم $\mathcal{R}(S)$ فضای تولید شده توسط ستون‌های ماتریس S است. بنابراین کافی است یک پایهٔ یکامتعامد برای $\operatorname{span}\{A_{*1}, A_{*2}, A_{*3}\}$ بیابیم. برای سادگی فرض کنید

$B = A^H A$. قرار می‌دهیم $x_1 = A_{*1}$. با توجه به قسمت دوم از مثال ۹.۲ داریم:

$$\|x_1\|_B = \sqrt{x^H B x} = 5, \quad u_1 = \frac{x_1}{\|x_1\|_B} = \frac{1}{5}\begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

با فرض $x_2 = A_{*2}$ داریم:

$$\langle u_1, x_2 \rangle = u_1^T B x_2 = -10, \quad w_2 = x_2 - \langle u_1, x_2 \rangle u_1 = x + 10 u_1 = 0.$$

این رابطه نشان می‌دهد که بردار A_{*2} و u_1 وابسته خطی هستند. از این رو بردار A_{*2} نمی‌تواند در پایه قرار گیرد. بنابراین قرار می‌دهیم $x_2 = A_{*3}$. لذا خواهیم داشت:

$$\begin{aligned} \langle u_1, x_2 \rangle &= u_1^T B x_2 = \frac{3}{5}, \\ w_2 &= x_2 - \langle u_1, x_2 \rangle u_1 = x_2 - \frac{3}{5} v_1 = \frac{1}{25}\begin{pmatrix} 28 \\ 16 \end{pmatrix}. \end{aligned}$$

در نتیجه

$$\|w_2\|_B = \sqrt{w_2^T B w_2} = \frac{4}{5}, \quad u_2 = \frac{w_2}{\|w_2\|_B} = \frac{1}{5}\begin{pmatrix} 7 \\ 4 \end{pmatrix}.$$

بنابراین $\{u_1, u_2\}$ یک پایهٔ یکامتعامد نسبت به ضرب داخلی $\langle x, y \rangle = x^T B y$ می‌باشد. توجه کنید که اگر قرار دهید $U = (u_1, u_2)$، آنگاه $U^T B U = I$ (چرا؟).

تمرینات

۱.۲) فرض کنید $a_i, b_i \in \mathbb{R}$، $i = 1, 2, \ldots, n$. با استفاده از اینکه به ازای هر $t \in \mathbb{R}$، داریم $\sum_{i=1}^n (a_i t - b_i)^2 \geqslant 0$ نتیجه بگیرید:

$$\left| \sum_{i=1}^n a_i b_i \right| \leqslant \left(\sum_{i=1}^n |a_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |b_i|^2 \right)^{\frac{1}{2}}.$$

با استفاده از این رابطه نشان دهید که به ازای هر $a_i, b_i \in \mathbb{C}$، $i = 1, 2, \ldots, n$، داریم:

$$\left| \sum_{i=1}^n \overline{a_i} b_i \right| \leqslant \left(\sum_{i=1}^n |a_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |b_i|^2 \right)^{\frac{1}{2}}.$$

۲.۲) نرم-۱ ، نرم- ۲ و نرم-∞ بردارهای

$$x = \begin{pmatrix} ۲ \\ ۱ \\ -۴ \\ -۲ \end{pmatrix}, \quad y = \begin{pmatrix} ۱+i \\ ۱-i \\ ۱ \\ ۴i \end{pmatrix},$$

را محاسبه کنید (i واحد موهومی است).

۳.۲) فرض کنید:

$$A = \begin{pmatrix} ۱ & ۱ & -۱ \\ ۰ & ۱ & ۲ \\ -۱ & ۱ & -۱ \end{pmatrix}.$$

مقادیر $\|A\|_F$، $\|A\|_\infty$، $\|A\|_۱$ و $\|A\|_۲$ را محاسبه کنید.

۴.۲) درایه‌های ماتریس $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ به صورت

$$a_{ij} = \begin{cases} -۱, & i > j, \\ ۱, & i = j, \\ ۰, & i < j, \end{cases}$$

تعریف شده‌اند. مقادیر $\|A\|_F$، $\|A\|_\infty$ و $\|A\|_۱$ را محاسبه کنید.

۵.۲) توضیح دهید که چرا برای هر نرم $\|.\|$، داریم $\|x-y\| = \|y-x\|$.

۶.۲) فرض کنید $A, E \in \mathbb{C}^{m\times n}$. نشان دهید برای هر نرم ماتریسی $\|.\|$، داریم:

$$|\|A+E\| - \|A\|| \leqslant \|E\|.$$

۷.۲) فرض کنید برای دو بردار $x, y \in \mathbb{R}^n$، داریم $\|x+y\|_۲ = \|x-y\|_۲$. نشان دهید $x^Ty = ۰$.

۸.۲) نشان دهید $\|.\|_۲$ و $\|.\|_F$ به ترتیب در شرایط نرم برداری و نرم ماتریسی صدق می‌کنند.

۹.۲) روابط (۳.۲) و (۴.۲) را ثابت کنید.

۱۰.۲) فرض کنید $A \in \mathbb{C}^{m\times n}$. نشان دهید:

$$\|A\|_\infty = \max_{1\leqslant i\leqslant m}\sum_{j=1}^{n}|a_{ij}|, \qquad \|A\|_1 = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{m}|a_{ij}|,$$

در شرایط نرم ماتریسی صدق می‌کند.

۱۱.۲) قسمت‌های (ب) و (ج) از قضیهٔ ۵.۲ را ثابت کنید.

۱۲.۲) نشان دهید ضرب‌های داخلی تعریف شده در مثال ۷.۲ در شرایط ضرب داخلی صدق می‌کنند.

۱۳.۲) نشان دهید $\rho(A)$ نمی‌تواند یک نرم ماتریسی باشد.

۱۴.۲) فرض کنید θ_i ها و λ_i ها اعداد مثبتی باشند به طوری که $\sum_{i=1}^{n}\theta_i = 1$. با کمک نامساوی کوشی– شوارتز، نشان دهید:

$$\left(\sum_{i=1}^{n}\theta_i\lambda_i\right)\left(\sum_{i=1}^{n}\frac{\theta_i}{\lambda_i}\right) \geqslant 1.$$

۱۵.۲) هر یک از مجموعه‌های زیر را در $\mathbb{R}^2$ نمایش دهید

$$\left\{x\in\mathbb{R}^2 : \|x\|_2 \leqslant 1\right\}, \quad \left\{x\in\mathbb{R}^2 : \|x\|_\infty \leqslant 1\right\}, \quad \left\{x\in\mathbb{R}^2 : \|x\|_1 \leqslant 1\right\}.$$

۱۶.۲) فرض کنید $A \in \mathbb{R}^{n\times n}$ یک ماتریس نامنفرد و $\|.\|$ یک نرم روی $\mathbb{R}^n$ باشد. نشان دهید $\|.\|_A$ که به صورت $\|x\|_A = \|Ax\|$ تعریف می‌شود، یک نرم روی $\mathbb{R}^n$ است.

۱۷.۲) نشان دهید برای هر دو بردار $x, y \in \mathbb{R}^n$، داریم:

$$\|xy^H\|_F = \|xy^H\|_2 = \|x\|_2\|y\|_2.$$

۱۸.۲) فرض کنید $x \in \mathbb{C}^n$ یک بردار دلخواه و ماتریس $P \in \mathbb{C}^{n\times n}$ یکانی باشد. نشان دهید $\|P\|_2 = 1$ و $\|Px\|_2 = \|x\|_2$.

۱۹.۲) فرض کنید $A \in \mathbb{R}^{n\times n}$. نشان دهید $\|A\|_2 \leqslant n\|A\|_{\max}$ و $\|A^TA\|_2 = \|A\|_2^2$.

۲۰.۲) فرض کنید $A \in \mathbb{C}^{n\times n}$ و به ازای هر $x, y \in \mathbb{C}^n$ داریم $\langle Ax, Ay\rangle = \langle x, y\rangle$، که در آن $\langle .,.\rangle$ ضرب داخلی استاندارد روی $\mathbb{C}^n$ است. نشان دهید ماتریس A یکانی است.

۲۱.۲) فرض کنید $A \in \mathbb{R}^{n\times n}$ و به ازای هر $x \in \mathbb{R}^n$ داریم $\|Ax\|_2 = \|x\|_2$. نشان دهید ماتریس A متعامد است. (راهنمایی: نشان دهید:

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|_2^2 - \|x-y\|_2^2),$$

و سپس از تمرین ۲۰.۲ استفاده کنید.)

۲۲.۲) فرض کنید $\{u_1, \dots, u_n\}$ یک مجموعهٔ یکامتعامد در $\mathbb{C}^n$ نسبت به ضرب داخلی استاندارد $\langle .,.\rangle$ باشد. نشان دهید به ازای هر $x \in \mathbb{C}^n$،

$$\left\|\sum_{i=1}^{n}\langle u_i, x\rangle u_i\right\|_2^2 = \sum_{i=1}^{n}|\langle u_i, x\rangle|^2$$

۲۳.۲) با استفاده از فرایند گرام- اشمیت یک پایهٔ یکامتعامد برای

$$S = \operatorname{span}\left\{\begin{pmatrix}0\\2\\0\end{pmatrix}, \begin{pmatrix}3\\0\\4\end{pmatrix}, \begin{pmatrix}3\\1\\4\end{pmatrix}\right\},$$

بیابید.

۲۴.۲) با استفاده از فرایند گرام- اشمیت یک پایهٔ یکامتعامد برای فضای تولید شده توسط ستون‌های هر یک از ماتریس‌های زیر را به دست آورید:

$$A = \begin{pmatrix}1 & 2 & 3\\2 & 1 & 1\\2 & 1 & 2\end{pmatrix}, \quad B = \begin{pmatrix}1 & 2 & 3\\2 & 4 & 1\end{pmatrix}.$$

۲۵.۲) فرض کنید $n \geqslant 1$ و $0 \neq s \in \mathbb{R}^n$. نشان دهید:

$$\left\|I - \frac{ss^T}{s^Ts}\right\|_F = \sqrt{n-1},$$

که در آن I ماتریس همانی $n \times n$ است.

فصل ۳

حل دستگاه معادلات خطی

دستگاه معادلات خطی

$$Ax = b, \tag{۱.۳}$$

را در نظر بگیرید که در آن

$$A = \begin{pmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{C}^{n\times n}, \quad b = \begin{pmatrix} b_۱ \\ b_۲ \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{C}^n,$$

معلوم هستند و بردار $x \in \mathbb{C}^n$ یک بردار مجهول می‌باشد. هدف، یافتن برداری مثل x است به طوری که در (۱.۳) صدق کند. به طور کلی روش‌های حل این دستگاه به دو دسته تقسیم می‌شوند: روش‌های مستقیم و روش‌های تکراری. در روش‌های مستقیم بعد از تعداد متناهی اعمال حسابی (در حساب دقیق) به جواب دستگاه می‌رسیم. در روش‌های تکراری با استفاده از یک حدس اولیه برای بردار جواب دنباله‌ای از بردارها تولید می‌شود که به جواب دستگاه همگرا است. در این فصل بعضی از روش‌های مستقیم و در فصل ۵ بعضی از روش‌های تکراری را معرفی می‌کنیم.

۱.۳ روش‌های حذفی گاوس و گاوس- جُردن

۱.۱.۳ روش حذفی گاوس

در روش حذفی گاوس، دستگاه معادلات خطی $Ax = b$ به دستگاه معادل $Ux = c$ تبدیل می‌شود که در آن U یک ماتریس بالامثلثی است. این دستگاه به سادگی قابل حل است زیرا اگر دستگاه بالامثلثی را به صورت

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix},$$

بنویسیم آنگاه با شرط اینکه برای هر i، $u_{ii} \neq 0$، داریم:

$$\begin{aligned} x_n &= \frac{c_n}{u_{nn}}, \\ x_k &= \frac{1}{u_{kk}}(c_k - \sum_{j=k+1}^{n} u_{kj}x_j), \quad k = n-1, n-2, \ldots, 1. \end{aligned}$$

این روند را جایگذاری پسرو می‌گویند. به همین ترتیب اگر ماتریس ضرایب دستگاه یک ماتریس پایین‌مثلثی باشد، یعنی اگر دستگاه به صورت

$$\begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix},$$

باشد، آنگاه با شرط اینکه برای هر i، $l_{ii} \neq 0$، داریم:

$$x_1 = \frac{c_1}{l_{11}}, \quad x_k = \frac{1}{l_{kk}}(c_k - \sum_{j=1}^{k-1} l_{kj}x_j), \quad k = 2, 3, \ldots, n.$$

این روند را جایگذاری پیشرو می‌گویند. دستگاه (۱.۳) را به صورت

$$E_1 : a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1,$$

$$E_2 : a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2,$$
$$\vdots$$
$$E_n : a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n,$$

می‌نویسیم، که درآن E_i نشان دهندهٔ معادلهٔ i ام است. فرایند روش حذفی گاوس در $n-1$ مرحله انجام می‌شود و ایدهٔ اصلی روش این است که با سه عمل

(۱) تعویض دو معادله؛

(۲) افزودن مضربی از طرفین یک معادله به طرفین معادلهٔ دیگر؛

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

که اعمال سطری مقدماتی نامیده می‌شوند، دستگاه را به دستگاه معادل دیگری تبدیل می‌کنیم (دستگاهی که جواب آن با جواب دستگاه اولیه یکسان است). برای این کار، ماتریس افزودهٔ

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

را تشکیل می‌دهیم. انجام اعمال سطری مقدماتی روی دستگاه $Ax = b$، معادل با انجام این اعمال روی سطرهای ماتریس افزوده است. $n - 1$ مرحلهٔ روش حذفی گاوس به صورت زیر انجام می‌گیرد.

گام ۱: متغیر x_1 در تمام معادلات بجز معادلهٔ اول حذف می‌شود. برای این کار، فرض کنید $a_{11} \neq 0$. قرار می‌دهیم:

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

سپس به ازای $i = 2, 3, \ldots, n$، $-l_{i1}$ برابر سطر اول ماتریس افزودهٔ (A, b) را به سطر i ام این ماتریس اضافه می‌کنیم. در این صورت، ماتریس افزودهٔ به ماتریس $(A^{(1)}, b^{(1)})$ تبدیل می‌شود که در آن

$$a_{ij}^{(1)} = a_{ij} - l_{i1}a_{1j}, \quad i = 2, 3, \ldots, n, \quad j = 2, 3, \ldots, n,$$
$$b_i^{(1)} = b_i - l_{i1}b_1, \quad i = 2, 3, \ldots, n.$$

در اینجا توجه کنید که $a_{i1}^{(1)} = 0$، $i = 2, 3, \dots, n$ و سایر درایه‌های ماتریس افزودهٔ $(A^{(1)}, b^{(1)})$ با درایه‌های متناظر در ماتریس (A, b) برابرند.

گام ۲: متغیر x_2 را از معادلات سوم تا n ام حذف می‌کنیم. فرض می‌کنیم $a_{22}^{(1)} \neq 0$ و قرار می‌دهیم:

$$l_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}, \quad i = 3, 4, \dots, n.$$

سپس به ازای $i = 3, 4, \dots, n$، $-l_{i2}$ برابر سطر دوم ماتریس افزودهٔ $(A^{(1)}, b^{(1)})$ را به سطر i ام این ماتریس اضافه می‌کنیم. در این صورت، ماتریس $(A^{(2)}, b^{(2)})$ به دست می‌آید که در آن

$$\begin{aligned}
a_{ij}^{(2)} &= a_{ij}^{(1)} - l_{i2} a_{2j}^{(1)}, \quad i = 3, 4, \dots, n, \quad j = 3, 4, \dots, n, \\
b_i^{(2)} &= b_i^{(1)} - l_{i2} b_2^{(1)}, \quad i = 3, 4, \dots, n, \\
a_{i2}^{(2)} &= 0, \quad i = 3, 4, \dots, n.
\end{aligned}$$

توجه کنید که سایر درایه‌های ماتریس $(A^{(2)}, b^{(2)})$ با درایه‌های متناظر در ماتریس $(A^{(1)}, b^{(1)})$ برابرند.

گام k: متغیر x_k را از معادلات $(k+1)$ ام تا n ام حذف می‌کنیم. برای این کار، فرض می‌کنیم $a_{kk}^{(k-1)} \neq 0$ و قرار می‌دهیم:

$$l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad i = k+1, k+2, \dots, n.$$

سپس به ازای $i = k+1, k+2, \dots, n$، $-l_{ik}$ برابر سطر k ام ماتریس افزودهٔ $(A^{(k-1)}, b^{(k-1)})$ را به سطر i ام این ماتریس اضافه می‌کنیم. در این صورت، ماتریس $(A^{(k)}, b^{(k)})$ به دست می‌آید که در آن

$$\begin{aligned}
a_{ij}^{(k)} &= a_{ij}^{(k-1)} - l_{ik} a_{kj}^{(k-1)}, \quad i = k+1, \dots, n, \quad j = k+1, \dots, n, \\
b_i^{(k)} &= b_i^{(k-1)} - l_{ik} b_k^{(k-1)}, \quad i = k+1, \dots, n, \\
a_{ik}^{(k)} &= 0, \quad i = k+1, \dots, n.
\end{aligned}$$

درایه‌های دیگر ماتریس $(A^{(k)}, b^{(k)})$ با درایه‌های متناظر در ماتریس $(A^{(k-1)}, b^{(k-1)})$ برابر می‌باشند.

به وضوح در پایان مرحلهٔ $(n-1)$ام ماتریس $A^{(n-1)}$ یک ماتریس بالامثلثی خواهد بود و دستگاه متناظر به سادگی حل می‌شود. روند فوق به روش حذفی گاوس بدون محورگیری (یا به طور خلاصه، روش حذفی گاوس) معروف است.

به درایه‌های a_{11}، $a_{22}^{(1)}$، …، $a_{nn}^{(n-1)}$ که در $n-1$ مرحله از روش حذفی گاوس ظاهر می‌شوند، درایه‌های محوری می‌گویند. در صورتی که یکی از درایه‌های محوری برابر با صفر باشد، آنگاه روش حذفی گاوس شکست خورده و الگوریتم متوقف می‌شود. همان‌طور که ملاحظه شد، اگر تمام درایه‌های محوری ناصفر باشند، آنگاه الگوریتم با موفقیت به پایان می‌رسد. الگوریتم روش حذفی گاوس را می‌توان به صورت الگوریتم ۱.۳ جمع‌بندی کرد. در اینجا لازم به ذکر است که تمام $a_{ij}^{(k)}$ ها و $b_i^{(k)}$ ها به ترتیب در A و b ذخیره می‌شوند. در واقع در پایان اجرای الگوریتم، قسمت بالامثلثی $A^{(n-1)}$ در قسمت بالامثلثی A و بردار $b^{(n-1)}$ در بردار b ذخیره شده‌اند.

الگوریتم ۱.۳: الگوریتم حذفی گاوس

1. Input $A = (a_{ij})$ and $b = (b_1, \ldots, b_n)^T$
2. For $k = 1, \ldots, n-1$, Do
3.   For $i = k+1, \ldots, n$ Do
4.   If $a_{kk} \neq 0$ then $l_{ik} := \dfrac{a_{ik}}{a_{kk}}$, else stop
5. $b_i := b_i - l_{ik} b_k$
6.   For $j = k+1, \ldots, n$ Do
7. $a_{ij} := a_{ij} - l_{ik} a_{kj}$
8.   EndDo
9.   EndDo
10. EndDo

مثال ۱.۳ دستگاه

$$\begin{pmatrix} 3 & 4 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 6 \end{pmatrix},$$

را به روش حذفی گاوس حل کنید.

حل: ماتریس افزودهٔ متناظر برای این دستگاه به صورت

$$(A,b) = \left(\begin{array}{ccc|c} ۲ & ۴ & ۳ & ۱ \\ ۳ & ۲ & ۱ & ٦ \\ ۲ & ۱ & ۲ & ٦ \end{array}\right),$$

است.

گام ۱: داریم:

$$\begin{aligned}
l_{۲۱} &= \frac{۳}{۲}, \quad l_{۳۱} = \frac{۲}{۲} = ۱, \\
a_{۲۲}^{(۱)} &= a_{۲۲} - l_{۲۱}a_{۱۲} = ۲ - \frac{۳}{۲} \times ۴ = -۴, \\
a_{۲۳}^{(۱)} &= a_{۲۳} - l_{۲۱}a_{۱۳} = ۱ - \frac{۳}{۲} \times ۳ = -\frac{۷}{۲}, \\
a_{۳۲}^{(۱)} &= a_{۳۲} - l_{۳۱}a_{۱۲} = ۱ - ۱ \times ۴ = -۳, \\
a_{۳۳}^{(۱)} &= a_{۳۳} - l_{۳۱}a_{۱۳} = ۲ - ۱ \times ۳ = -۱, \\
b_{۲}^{(۱)} &= b_{۲} - l_{۲۱}b_{۱} = ٦ - \frac{۳}{۲} \times ۱ = \frac{۹}{۲}, \\
b_{۳}^{(۱)} &= b_{۳} - l_{۳۱}b_{۱} = ٦ - ۱ \times ۱ = ۵.
\end{aligned}$$

در نتیجه

$$(A^{(۱)}, b^{(۱)}) = \left(\begin{array}{ccc|c} ۲ & ۴ & ۳ & ۱ \\ ۰ & -۴ & -\frac{۷}{۲} & \frac{۹}{۲} \\ ۰ & -۳ & -۱ & ۵ \end{array}\right).$$

گام ۲: داریم:

$$\begin{aligned}
l_{۳۲} &= \frac{-۳}{-۴} = \frac{۳}{۴}, \\
a_{۳۳}^{(۲)} &= a_{۳۳}^{(۱)} - l_{۳۲}a_{۲۳}^{(۱)} = -۱ - \frac{۳}{۴} \times (-\frac{۷}{۲}) = \frac{۱۳}{۸}, \\
b_{۳}^{(۲)} &= b_{۳}^{(۱)} - l_{۳۲}b_{۲}^{(۱)} = ۵ - \frac{۳}{۴} \times \frac{۹}{۲} = \frac{۱۳}{۸}.
\end{aligned}$$

در نتیجه

$$(A^{(۲)}, b^{(۲)}) = \left(\begin{array}{ccc|c} ۲ & ۴ & ۳ & ۱ \\ ۰ & -۴ & -\frac{۷}{۲} & \frac{۹}{۲} \\ ۰ & ۰ & \frac{۱۳}{۸} & \frac{۱۳}{۸} \end{array}\right).$$

حال با استفاده از جایگذاری پسرو داریم:

$$\begin{aligned} x_۳ &= \frac{\frac{۱۳}{۸}}{\frac{۱۳}{۸}} = ۱, \\ x_۲ &= \frac{۱}{-۴}(\frac{۹}{۲} + \frac{۷}{۲}x_۳) = \frac{۱}{-۴}(\frac{۹}{۲} + \frac{۷}{۲}) = -۲, \\ x_۱ &= \frac{۱}{۲}(۱ - ۴x_۲ - ۳x_۳) = \frac{۱}{۲}(۱ - ۴ \times (-۲) - ۳ \times ۱) = ۳. \end{aligned}$$

۳.۱.۲ تعداد اعمال حسابی در روش حذفی گاوس

گام k ام از روش حذفی گاوس را در نظر بگیرید. برای محاسبهٔ l_{ik}، $i = k + ۱, \ldots, n$، نیاز به $n - k$ عمل تقسیم می‌باشد که در این صورت، در $n - ۱$ مرحله از روش حذفی گاوس تعداد کل اعمال حسابی برای محاسبهٔ l_{ik} ها عبارت است از

$$\sum_{k=۱}^{n-۱}(n-k) = \sum_{k=۱}^{n-۱} k = \frac{n(n-۱)}{۲}.$$

همچنین در گام k ام، برای $i = k + ۱, \ldots, n$، ضرب l_{ik} در سطر k ام نیاز به $(n-k)(n-k+۱)$ عمل ضرب می‌باشد که تعداد کل ضرب‌ها در $n - ۱$ مرحله عبارت است از

$$\sum_{k=۱}^{n-۱}(n-k)(n-k+۱) = \sum_{k=۱}^{n-۱} k(k+۱) = \frac{n^۳}{۳} - \frac{n}{۳}.$$

از طرفی در گام k ام، برای $i = k + ۱, \ldots, n$، l_{ik} برابر سطر k ام به سطرهای دیگر افزوده می‌شود که تعداد اعمال جمع برابر است با $(n-k)(n-k+۱)$ و بنابراین تعداد کل اعمال جمع در $n - ۱$ مرحله عبارت است از

$$\sum_{k=۱}^{n-۱}(n-k)(n-k+۱) = \sum_{k=۱}^{n-۱} k(k+۱) = \frac{n^۳}{۳} - \frac{n}{۳}.$$

در حل دستگاه بالامثلثی با استفاده از جایگذاری پسرو برای محاسبهٔ مجهول x_k به یک عمل تقسیم، $n-k$ عمل ضرب و $n-k$ عمل جمع نیاز است. بنابراین تعداد کل تقسیم‌ها برای محاسبهٔ تمامی مجهولات برابر با n، تعداد ضرب‌ها $n(n-۱)/۲$ و تعداد جمع‌ها $n(n-۱)/۲$ می‌باشند.

بنابراین تعداد کل اعمال حسابی در روش حذفی گاوس را می‌توان به صورت زیر جمع‌بندی کرد:

تقسیم :

$$\frac{n(n-۱)}{۲}+n=\frac{n(n+۱)}{۲},$$

ضرب:

$$\frac{n^۳}{۳}-\frac{n}{۳}+\frac{n(n-۱)}{۲}=\frac{n^۳}{۳}+\frac{n^۲}{۲}-\frac{۵n}{۶},$$

جمع:

$$\frac{n^۳}{۳}-\frac{n}{۳}+\frac{n(n-۱)}{۲}=\frac{n^۳}{۳}+\frac{n^۲}{۲}-\frac{۵n}{۶}.$$

به این ترتیب تعداد کل اعمال حسابی در روش حذفی گاوس برای حل دستگاه معادلات خطی برابر است با

$$\frac{n(n+۱)}{۲}+۲\times(\frac{n^۳}{۳}+\frac{n^۲}{۲}-\frac{۵n}{۶})=\frac{۲}{۳}n^۳+\frac{۳}{۲}n^۲-\frac{۷}{۶}n.$$

در عمل، در شمارش تعداد اعمال حسابی از جملات با درجات پایین چشم‌پوشی می‌شود. این کار، معقول به نظر می‌رسد، زیرا برای n های بزرگ جملهٔ با بزرگترین درجه، غالب بر سایر جملات است و به خوبی تعداد اعمال حسابی را تخمین می‌زند. بنابراین تعداد تقریبی اعمال حسابی در روش حذفی گاوس برابر با $۲n^۳/۳$ است.

به دلیل انباشتگی خطاهای حاصل از گرد کردن اعداد، ممکن است که روش حذفی گاوس برای حل دستگاه معادلات خطی جواب های غیر قابل قبولی را تولید کند (حتی برای دستگاه‌های معادلات خطی با ابعاد کوچک). در اینجا یک نمونه از این دستگاه‌ها را بررسی می‌کنیم.

مثال ۲.۳ دستگاه معادلات خطی

$$\begin{cases} -10^{-4}x_1 + x_2 = 1, \\ x_1 + x_2 = 2, \end{cases}$$

را در نظر بگیرید. جواب دقیق این دستگاه

$$x_1 = \frac{1}{1.0001} \approx 1, \quad x_2 = \frac{1.0002}{1.0001} \approx 1,$$

می‌باشد. حال فرض کنید که این دستگاه به روش حذفی گاوس با کامپیوتری حل شود که محاسبات را با سه رقم بامعنی در مبنای ۱۰ انجام می‌دهد. در این صورت، ماتریس افزودهٔ زیر را تشکیل می‌دهیم:

$$(A, b) = \left(\begin{array}{cc|c} -10^{-4} & 1 & 1 \\ 1 & 1 & 2 \end{array} \right).$$

اگر x یک عدد حقیقی باشد، آنگاه نمایش آن را در این کامپیوتر با $\mathrm{fl}(x)$ نشان می‌دهیم. قرار می‌دهیم:

$$l_{21} = \mathrm{fl}(\frac{1}{-10^{-4}}) = \mathrm{fl}(-10^{4}) = -0.1 \times 10^{5},$$

و

$$a_{22}^{(1)} = \mathrm{fl}(1 + 0.1 \times 10^{5}) = \mathrm{fl}(0.10001 \times 10^{5}) = 0.100 \times 10^{5} = 10^{4},$$

$$b_{2}^{(1)} = \mathrm{fl}(2 + 0.1 \times 10^{5}) = \mathrm{fl}(0.10002 \times 10^{5}) = 0.100 \times 10^{5} = 10^{4}.$$

بنابراین

$$(A^{(1)}, b^{(1)}) = \left(\begin{array}{cc|c} -10^{-4} & 1 & 1 \\ 0 & 10^{4} & 10^{4} \end{array} \right).$$

با استفاده از جایگذاری پسرو داریم $x_2 = 1$ و $x_1 = 0$. می‌بینیم که جواب به دست آمده با جواب واقعی دستگاه اولیه تفاوت فاحش دارد. به وضوح این تفاوت مربوط به این است که درایهٔ محوری در اینجا بسیار کوچک و در نتیجه $|l_{21}|$ بسیار بزرگ شده است. این عدد باعث شده است که در محاسبهٔ هریک از مقادیر $b_2^{(1)}$ و $a_{22}^{(1)}$، عددی بسیار بزرگ با

یک عدد کوچک جمع گردد که باعث خطا گردیده است. در واقع عدد کوچک در حاصل جمع آن با عدد بسیار بزرگ نادیده گرفته شده است. این مشکل را می‌توان با استفاده از محورگیری جزئی که در ادامه می‌آید، بر طرف کرد.

۳.۱.۳ محورگیری جزئی

در گام اول از روش حذفی گاوس با محورگیری جزئی، در ستون اول ماتریس افزوده، بزرگترین درایه از لحاظ قدرمطلق را یافته و سطر متناظر را با سطر اول (سطر درایهٔ محوری) تعویض می‌کنیم. با این عمل داریم:

$$|l_{i1}| \leqslant 1, \quad i = 2, 3, \ldots, n.$$

در گام k ام، از بین درایهٔ $a_{kk}^{(k-1)}$ و درایه‌های زیرآن، بزرگترین درایه از لحاظ قدرمطلق را یافته و سطر متناظر به این درایه را با سطر درایهٔ محوری جابجا می‌کنیم. با این عمل داریم:

$$|l_{ik}| \leqslant 1, \quad i = k+1, k+2, \ldots, n.$$

بنابراین محورگیری جزئی باعث می‌شود که از بزرگ شدن l_{ik} ها جلوگیری شود. ذکر این نکته ضروری است که اگر در مرحلهٔ k ام، درایهٔ محوری و تمام درایه‌هایی که در ستون k ام و در زیر درایهٔ محوری قرار دارند برابر با صفر باشند، آنگاه ماتریس ضرایب دستگاه منفرد است و الگوریتم متوقف می‌شود. حال مثال ۲.۳ را با استفاده از محورگیری جزئی حل می‌کنیم.

مثال ۳.۳ دستگاه

$$\begin{cases} -10^{-4}x_1 + x_2 = 1, \\ \qquad\quad x_1 + x_2 = 2, \end{cases}$$

را به روش حذفی گاوس با محورگیری جزئی با کامپیوتری که محاسبات را با سه رقم با معنی در مبنای ۱۰ انجام می‌دهد، حل کنید.

حل: ماتریس افزودهٔ

$$(A,b) = \left(\begin{array}{cc|c} -10^{-4} & 1 & 1 \\ 1 & 1 & 2 \end{array}\right).$$

را تشکیل می‌دهیم. در ستون اول درایهٔ a_{21} از لحاظ قدرمطلق از a_{11} بزرگتر است. بنابراین سطر دوم را با سطر اول جابجا می‌کنیم:

$$(A,b) = \left(\begin{array}{cc|c} 1 & 1 & 2 \\ -10^{-4} & 1 & 1 \end{array}\right).$$

قرار می‌دهیم:

$$l_{21} = \text{fl}(\frac{-10^{-4}}{1}) = \text{fl}(-10^{-4}) = -0.1 \times 10^{-3},$$

و

$$a_{22}^{(1)} = \text{fl}(1 + 0.1 \times 10^{-3}) = \text{fl}(0.10001 \times 10^{1}) = 0.100 \times 10^{1} = 1,$$
$$b_{2}^{(1)} = \text{fl}(1 + 2 \times 0.1 \times 10^{-3}) = \text{fl}(0.10002 \times 10^{1}) = 0.100 \times 10^{1} = 1.$$

بنابراین

$$(A^{(1)}, b^{(1)}) = \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right).$$

با استفاده از جایگذاری پسرو داریم $x_1 = x_2 = 1$. می‌بینیم که این جواب با جواب واقعی تفاوت چندانی ندارد.

مثال ۴.۳ دستگاه معادلات خطی

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

را به روش حذفی گاوس با محورگیری جزئی حل کنید.

حل: ماتریس افزودهٔ

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

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

$$\left(\begin{array}{ccc|c} ۳ & ۲ & ۲ & ۹ \\ ۲ & ۰ & ۱ & ۳ \\ ۱ & ۱ & ۵ & ۸ \end{array}\right).$$

با حذف درایه‌های زیر درایهٔ $(۱,۱)$، داریم:

$$(A^{(۱)}, b^{(۱)}) = \left(\begin{array}{ccc|c} ۳ & ۲ & ۲ & ۹ \\ ۰ & -\frac{۴}{۳} & -\frac{۱}{۳} & -۳ \\ ۰ & \frac{۱}{۳} & \frac{۱۳}{۳} & ۵ \end{array}\right).$$

حال از میان درایه‌های $(۲,۲)$ و $(۳,۲)$ درایهٔ $(۲,۲)$ از لحاظ قدرمطلق بزرگتر است و بنابراین نیازی به تعویض سطر نیست. با صفرکردن درایهٔ $(۳,۲)$ خواهیم داشت:

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

حال از جایگذاری پسرو استفاده می‌کنیم. معادلهٔ سوم به صورت

$$\frac{۱۷}{۴}x_۳ = \frac{۱۷}{۴},$$

است که با حل آن داریم $x_۳ = ۱$. با جایگذاری $x_۳ = ۱$ در معادلهٔ دوم، یعنی

$$-\frac{۴}{۳}x_۲ - \frac{۱}{۳}x_۳ = -۳,$$

مقدار $x_۲ = ۲$ به دست می‌آید. سرانجام با جایگذاری $x_۲$ و $x_۳$ در معادلهٔ اول، یعنی

$$۳x_۱ + ۲x_۲ + ۲x_۳ = ۹,$$

خواهیم داشت $x_۱ = ۱$.

۴.۱.۳ محورگیری کلی

در روش حذفی گاوس همراه با محورگیری کلی و در گام k ام، فرض کنید که

$$|a_{rs}^{(k-1)}| = \max_{k\leqslant i,j\leqslant n} |a_{ij}^{(k-1)}|.$$

در این صورت ابتدا سطر r ام را با سطر k ام و سپس ستون s ام را با ستون k ام جابجا می‌کنیم. با این عمل درایهٔ $a_{rs}^{(k-1)}$ به مکان (k,k) انتقال داده می‌شود. توجه کنید که با این عمل درایهٔ محوری تا حد امکان بزرگ خواهد شد. در اینجا باید دقت کرد که جابجایی ستون‌های ماتریس ضرایب باعث جابجایی متغیرهای متناظر نیز می‌شود که بعد از اتمام کار بایستی منظور گردد. برای این کار، بعد از ستون $(n+1)$ ام در ماتریس افزوده، بردار $J=(1,2,\ldots,n)^T$ که مؤلفه‌های آن نشانگر ستون‌های ماتریس A می‌باشند، اضافه می‌شود. در صورتی که ستون‌های i ام و j ام ماتریس A با یکدیگر جابجا شوند، درایه‌های i ام و j ام ستون آخر ماتریس افزوده با یکدیگر جابجا می‌شوند. لازم به ذکر است که تعویض ستون‌ها تاثیری روی بردار سمت راست ندارد.

مثال ۵.۳ دستگاه مثال ۴.۳ را به روش حذفی گاوس با محورگیری کلی حل کنید.

حل: ماتریس افزودهٔ

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

را تشکیل می‌دهیم. بزرگترین درایهٔ ماتریس از لحاظ قدرمطلق در مکان $(3,3)$ قرار دارد. بنابراین مطابق آنچه که گفته شد ابتدا سطر اول و سوم و سپس ستون اول و سوم ماتریس را با یکدیگر جابجا می‌کنیم. لذا ماتریس افزوده به صورت زیر تبدیل می‌شود:

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

با حذف درایه‌های زیر درایهٔ $(۱,۱)$، داریم:

$$(A^{(۱)}, b^{(۱)}, J^{(۱)}) = \left(\begin{array}{ccc|c|c} ۵ & ۱ & ۱ & ۸ & ۳ \\ ۰ & \frac{۸}{۵} & \frac{۱۳}{۵} & \frac{۲۹}{۵} & ۲ \\ ۰ & -\frac{۱}{۵} & \frac{۹}{۵} & \frac{۷}{۵} & ۱ \end{array}\right).$$

حال درایهٔ مکان $(۲,۳)$ به مکان $(۲,۲)$ انتقال داده می‌شود. برای این منظور کافی است ستون‌های دوم و سوم ماتریس را با هم جابجا کنیم. از این‌رو ماتریس افزوده به صورت

$$\left(\begin{array}{ccc|c|c} ۵ & ۱ & ۱ & ۸ & ۳ \\ ۰ & \frac{۱۳}{۵} & \frac{۸}{۵} & \frac{۲۹}{۵} & ۱ \\ ۰ & \frac{۹}{۵} & -\frac{۱}{۵} & \frac{۷}{۵} & ۲ \end{array}\right).$$

خواهد بود. با صفر کردن درایهٔ مکان $(۳,۲)$، خواهیم داشت:

$$(A^{(۲)}, b^{(۲)}, J^{(۲)}) = \left(\begin{array}{ccc|c|c} ۵ & ۱ & ۱ & ۸ & ۳ \\ ۰ & \frac{۱۳}{۵} & \frac{۸}{۵} & \frac{۲۹}{۵} & ۱ \\ ۰ & ۰ & -\frac{۱۷}{۱۳} & -\frac{۳۴}{۱۳} & ۲ \end{array}\right).$$

حال از جایگذاری پسرو استفاده می‌کنیم. معادلهٔ آخر به صورت

$$-\frac{۱۷}{۱۳}x_۲ = -\frac{۳۴}{۱۳},$$

است که با حل آن داریم $x_۲ = ۲$. با جایگذاری $x_۲ = ۲$ در معادلهٔ دوم، یعنی

$$\frac{۱۳}{۵}x_۱ + \frac{۸}{۵}x_۲ = \frac{۲۹}{۵},$$

مقدار $x_۱ = ۱$ به دست می‌آید. سرانجام با جایگذاری $x_۱$ و $x_۲$ در معادلهٔ اول، یعنی

$$۵x_۳ + x_۱ + x_۲ = ۸,$$

خواهیم داشت $x_۳ = ۱$.

۵.۱.۳ روش گاوس- جُردن

در اینجا با استفاده از اعمال سطری مقدماتی ماتریس افزودهٔ (A,b) به ماتریس (I,c) تبدیل می‌شود که در آن I ماتریس همانی است. در این صورت بردار c، جواب دستگاه $Ax=b$ است.

مثال ۶.۳ دستگاه مثال ۱.۳ را به روش گاوس- جُردن حل کنید.

حل: ماتریس افزودهٔ (A,b) برای این دستگاه به صورت

$$(A,b)=\left(\begin{array}{ccc|c} 2 & 4 & 3 & 1 \\ 3 & 2 & 1 & 6 \\ 2 & 1 & 2 & 6 \end{array}\right),$$

است. سطر اول این ماتریس را در $\frac{1}{2}$ ضرب کرده و سپس -3 و -2 برابر سطر اول را به ترتیب به سطرهای دوم و سوم اضافه می‌کنیم. در نتیجه

$$(A^{(1)},b^{(1)})=\left(\begin{array}{ccc|c} 1 & 2 & \frac{3}{2} & \frac{1}{2} \\ 0 & -4 & -\frac{7}{2} & \frac{9}{2} \\ 0 & -3 & -1 & 5 \end{array}\right).$$

حال سطر دوم را در $-\frac{1}{4}$ ضرب کرده و سپس -2 و 3 برابر سطر دوم را به ترتیب به سطرهای اول و سوم اضافه می‌کنیم. بنابراین

$$(A^{(2)},b^{(2)})=\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{4} & \frac{11}{4} \\ 0 & 1 & \frac{7}{8} & -\frac{9}{8} \\ 0 & 0 & \frac{13}{8} & \frac{13}{8} \end{array}\right).$$

سطر سوم را در $\frac{8}{13}$ ضرب کرده و سپس $-\frac{7}{8}$ و $\frac{1}{4}$ سطر سوم را به ترتیب به سطرهای دوم و اول اضافه می‌کنیم. در این صورت،

$$(A^{(3)}, b^{(3)}) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

در نتیجه $x_1 = 3$، $x_2 = -2$ و $x_3 = 1$.

در اینجا لازم به ذکر است که همانند روش حذفی گاوس می‌توان عمل محورگیری (جزئی یا کلی) را بکار گرفت. انتظار می‌رود که خواننده جزئیات روش را بررسی نماید. در بخش بعد تجزیهٔ LU یک ماتریس را معرفی می‌کنیم که کاربردهای متعددی در جبر خطی دارد.

۲.۳ تجزیهٔ LU یک ماتریس

فرض کنید $A \in \mathbb{C}^{n\times n}$. تجزیهٔ LU ماتریس A یک تجزیه به صورت $A = LU$ است که در آن L یک ماتریس پایین‌مثلثی با درایه‌های قطری یک و U یک ماتریس بالامثلثی با درایه‌های قطری ناصفر می‌باشد. چنین تجزیه‌ای کاربردهای متعددی در علوم و مهندسی دارد که بعضی از آنها در این کتاب مورد مطالعه قرار می‌گیرند. در ادامه ابتدا روشی برای محاسبهٔ تجزیهٔ LU (در صورت وجود) و سپس در خصوص شرایط وجود و یکتایی آن بحث می‌کنیم.

در روش حذفی گاوس دیدیم که اگر هیچکدام از درایه‌های محوری صفر نشوند، آنگاه روش حذفی گاوس بدون عمل محورگیری با موفقیت انجام می‌شود. در این بخش نشان می‌دهیم که اگر روش حذفی گاوس با موفقیت انجام شود، آنگاه بدون هیچ محاسبهٔ اضافی می‌توان تجزیهٔ LU ماتریس A را به دست آورد. همانند روش حذفی گاوس این عمل در $n-1$ مرحله انجام می‌گیرد.

فرض کنید $a_{11} \neq 0$. تعریف می‌کنیم:

$$\ell_{i1} = \frac{a_{i1}}{a_{11}}, \quad i = 2, \ldots, n,$$

و

$$T_1 = \begin{pmatrix} 1 & & & & \\ -\ell_{21} & 1 & & & \\ -\ell_{31} & & 1 & & \\ \vdots & & & \ddots & \\ -\ell_{n1} & & & & 1 \end{pmatrix}.$$

در اینجا سایر درایه‌های ماتریس که نوشته نشده‌اند، برابر با صفرند. با ضرب ماتریس T_1 از چپ در ماتریس A، خواهیم داشت:

$$A^{(1)} = T_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix},$$

که در آن $a_{ij}^{(1)}$ ها درایه‌هایی هستند که بعد از مرحلهٔ اول روش حذفی گاوس به دست می‌آیند. فرض کنید ماتریس $A^{(k-1)}$ به صورت

$$A^{(k-1)} = \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & a_{kk}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ 0 & \cdots & a_{k+1,k}^{(k-1)} & \cdots & a_{k+1,n}^{(k-1)} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{pmatrix},$$

باشد که در آن $a_{kk}^{(k-1)} \neq 0$. قرار می‌دهیم:

$$\ell_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad i = k+1, \ldots, n,$$

و

$$T_k = \begin{pmatrix} ۱ & \cdots & ۰ & ۰ & \cdots & ۰ \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ ۰ & \cdots & ۱ & ۰ & \cdots & ۰ \\ ۰ & \cdots & -\ell_{k+۱,k} & ۱ & \cdots & ۰ \\ \vdots & & \vdots & \vdots & & \vdots \\ ۰ & \cdots & -\ell_{n,k} & ۰ & \cdots & ۱ \end{pmatrix}.$$

با ضرب T_k در $A^{(k-۱)}$، خواهیم داشت:

$$A^{(k)} = T_k A^{(k-۱)} = \begin{pmatrix} a_{۱۱} & \cdots & a_{۱k} & a_{۱,k+۱} & \cdots & a_{۱n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ ۰ & \cdots & a_{kk}^{(k-۱)} & a_{k,k+۱}^{(k-۱)} & \cdots & a_{kn}^{(k-۱)} \\ ۰ & \cdots & ۰ & a_{k+۱,k+۱}^{(k)} & \cdots & a_{k+۱,n}^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ ۰ & \cdots & ۰ & a_{n,k+۱}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}. \quad (۲.۳)$$

با انجام این عمل، بعد از $n-۱$ مرحله، یعنی به ازای $k = ۱, ۲, \ldots, n-۱$، ماتریس $A^{(n-۱)}$ یک ماتریس بالامثلثی خواهد بود. این ماتریس را U می‌نامیم. در نتیجه، داریم:

$$\begin{aligned} U &= A^{(n-۱)} = T_{n-۱}A^{(n-۲)} = T_{n-۱}T_{n-۲}A^{(n-۳)} \\ &= \cdots = T_{n-۱}T_{n-۲}\ldots T_۱ A. \end{aligned} \quad (۳.۳)$$

بنابراین $A = T_۱^{-۱}T_۲^{-۱}\cdots T_{n-۱}^{-۱}U$. قرار می‌دهیم $L = T_۱^{-۱}T_۲^{-۱}\cdots T_{n-۱}^{-۱}$ و نشان می‌دهیم که L یک ماتریس پایین‌مثلثی با درایه‌های قطری یک است. به سادگی می‌توان دید که

$$T_k^{-۱} = \begin{pmatrix} ۱ & \cdots & ۰ & ۰ & \cdots & ۰ \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ ۰ & \cdots & ۱ & ۰ & \cdots & ۰ \\ ۰ & \cdots & \ell_{k+۱,k} & ۱ & \cdots & ۰ \\ \vdots & & \vdots & \vdots & & \vdots \\ ۰ & \cdots & \ell_{n,k} & ۰ & \cdots & ۱ \end{pmatrix}, \quad k = ۱, ۲, \ldots, n-۱,$$

و بنابراین

$$L = \begin{pmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ \ell_{۲۱} & ۱ & ۰ & \cdots & ۰ \\ \ell_{۳۱} & \ell_{۳۲} & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n۱} & \ell_{n۲} & \ell_{n۳} & \cdots & ۱ \end{pmatrix}.$$

این نتیجه نشان می‌دهد که برای هر $i > j$، درایهٔ (i, j) ام ماتریس L برابر با ℓ_{ij} است که در روش حذفی گاوس ظاهر می‌شوند. بدیهی است اگر درایه‌های محوری در طی مراحلِ روش ناصفر باشند، آنگاه تجزیهٔ LU ماتریس A موجود و با موفقیت محاسبه می‌شود و در غیر این صورت الگوریتم شکست می‌خورد.

مثال ۷.۳ تجزیهٔ LU ماتریس

$$A = \begin{pmatrix} ۲ & ۲ & ۲ \\ ۴ & ۷ & ۷ \\ ۶ & ۱۸ & ۲۲ \end{pmatrix},$$

را به دست آورید.

حل: داریم:

$$\ell_{۲۱} = \frac{۴}{۲} = ۲, \qquad \ell_{۳۱} = \frac{۶}{۲} = ۳,$$

و

$$T_۱ = \begin{pmatrix} ۱ & ۰ & ۰ \\ -۲ & ۱ & ۰ \\ -۳ & ۰ & ۱ \end{pmatrix}.$$

در نتیجه

$$A^{(۱)} = T_۱ A = \begin{pmatrix} ۲ & ۲ & ۲ \\ ۰ & ۳ & ۳ \\ ۰ & ۱۲ & ۱۶ \end{pmatrix}.$$

قرارمی‌دهیم:

$$\ell_{۳۲} = \frac{۱۲}{۳} = ۴,$$

و

$$T_{۲} = \begin{pmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & -۴ & ۱ \end{pmatrix}.$$

بنابراین

$$T_{۲} A^{(۱)} = \begin{pmatrix} ۲ & ۲ & ۲ \\ ۰ & ۳ & ۳ \\ ۰ & ۰ & ۴ \end{pmatrix} = U.$$

از طرفی

$$L = \begin{pmatrix} ۱ & ۰ & ۰ \\ \ell_{۲۱} & ۱ & ۰ \\ \ell_{۳۱} & \ell_{۳۲} & ۱ \end{pmatrix} = \begin{pmatrix} ۱ & ۰ & ۰ \\ ۲ & ۱ & ۰ \\ ۳ & ۴ & ۱ \end{pmatrix}.$$

قضیهٔ بعد نشان می‌دهد که تجزیهٔ LU در صورت وجود یکتا خواهد بود.

قضیه ۱.۳ تجزیهٔ LU در صورت وجود یکتاست.

برهان: فرض کنید A دارای دو تجزیهٔ LU به صورت

$$A = L_{۱}U_{۱} = L_{۲}U_{۲},$$

باشد. از این رابطه نتیجه می‌شود که

$$L_{۲}^{-۱}L_{۱} = U_{۲}U_{۱}^{-۱}. \qquad (۴.۳)$$

با توجه به اینکه $L_{۲}$ پایین‌مثلثی با درایه‌های قطری یک است، نتیجه می‌شود که $L_{۲}^{-۱}$ نیز چنین است. بنابراین $L_{۲}^{-۱}L_{۱}$ یک ماتریس پایین‌مثلثی با درایه‌های قطری یک است. از طرفی چون $U_{۲}U_{۱}^{-۱}$ بالامثلثی است، با توجه به رابطهٔ (۴.۳) نتیجه می‌شود:

$$L_{۲}^{-۱}L_{۱} = U_{۲}U_{۱}^{-۱} = I.$$

بنابراین $L_{۱} = L_{۲}$ و $U_{۱} = U_{۲}$. □

به روشنی نامنفرد بودن ماتریس A برای وجود تجزیهٔ LU لازم می‌باشد، امّا این شرط، کافی نیست. برای مثال ماتریس

$$A = \begin{pmatrix} ۰ & ۱ \\ ۱ & ۱ \end{pmatrix},$$

را در نظر بگیرید. این ماتریس نامنفرد است اما تجزیهٔ LU آن موجود نیست. زیرا اگر

$$\begin{pmatrix} ۰ & ۱ \\ ۱ & ۱ \end{pmatrix} = \begin{pmatrix} ۱ & ۰ \\ \ell_{۲۱} & ۱ \end{pmatrix} \begin{pmatrix} u_{۱۱} & u_{۱۲} \\ ۰ & u_{۲۲} \end{pmatrix},$$

آنگاه درایهٔ محوری $u_{۱۱}$ برابر با صفر می‌شود و ادامهٔ محاسبهٔ تجزیه امکان‌پذیر نخواهد بود. اما تحت شرایطی وجود تجزیهٔ LU تضمین می‌شود. در ادامه موضوع وجود تجزیهٔ LU را بررسی می‌کنیم.

تعریف ۱.۳ فرض کنید $A \in \mathbb{C}^{n\times n}$، $N = \{۱, ۲, \ldots, n\}$ و $N_k = \{i_۱, i_۲, \ldots, i_k\}$ یک زیرمجموعه از N باشد به طوری که $i_۱ < i_۲ < \cdots < i_k$. در این صورت، یک زیرماتریس اصلی ماتریس A به صورت زیر تعریف می‌شود:

$$A_k = \begin{pmatrix} a_{i_۱,i_۱} & a_{i_۱,i_۲} & \cdots & a_{i_۱,i_k} \\ a_{i_۲,i_۱} & a_{i_۲,i_۲} & \cdots & a_{i_۲,i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k,i_۱} & a_{i_k,i_۲} & \cdots & a_{i_k,i_k} \end{pmatrix}.$$

در صورتی که $N_k = \{۱, ۲, \ldots, k\}$، آنگاه زیرماتریس اصلی متناظر با N_k را یک زیرماتریس اصلی پیشرو ماتریس A می‌نامند. برای مثال، اگر $A = (a_{ij})$ یک ماتریس ۳×۳ باشد، آنگاه زیرماتریس‌های اصلی پیشرو A عبارتند از

$$A_۱ = (a_{۱۱}), \quad A_۲ = \begin{pmatrix} a_{۱۱} & a_{۱۲} \\ a_{۲۱} & a_{۲۲} \end{pmatrix}, \quad A_۳ = A.$$

قضیه ۲.۳ (وجود تجزیهٔ LU). ماتریس A دارای تجزیهٔ LU است اگر و تنها اگر تمام زیرماتریس‌های اصلی پیشرو A نامنفرد باشند.

برهان: فرض کنید ماتریس A دارای تجزیهٔ LU باشد. در این صورت، داریم:

$$A = \begin{pmatrix} A_k & * \\ * & * \end{pmatrix} = \begin{pmatrix} L_k & 0 \\ * & * \end{pmatrix} \begin{pmatrix} U_k & * \\ 0 & * \end{pmatrix} = \begin{pmatrix} L_k U_k & * \\ * & * \end{pmatrix},$$

که در آن علامت $*$ معرف درایه‌های ناصفر ماتریس می‌باشند. در نتیجه $A_k = L_k U_k$، که در آن درایه‌های قطری L_k برابر با یک و درایه‌های قطری U_k غیر صفرند. بنابراین A_k نامنفرد است.

برعکس فرض کنید که تمام زیرماتریس‌های اصلی پیشرو A نامنفرد باشند. به استقراء روی بعد ماتریس حکم را ثابت می‌کنیم. به ازای $n = 1$، $A_1 = (a_{11})$ نامنفرد است و بنابراین $a_{11} \neq 0$. در نتیجه تجزیهٔ LU ماتریس A_1 به صورت

$$A_1 = (1)(a_{11}) = L_1 U_1,$$

است. حال فرض کنید حکم برای $n = k$ برقرار باشد و $n = k + 1$. با توجه به رابطهٔ (۲.۳)، داریم:

$$\begin{aligned}
A &= T_1^{-1} T_2^{-1} \cdots T_k^{-1} \begin{pmatrix} U_{k+1} & S_{k+1} \\ P_{k+1} & Q_{k+1} \end{pmatrix} \\
&= \begin{pmatrix}
1 & & & & & & \\
\ell_{21} & 1 & & & & & \\
\vdots & \ddots & \ddots & & & & \\
\ell_{k+1,1} & \cdots & \ell_{k+1,k} & 1 & & & \\
\ell_{k+2,1} & \cdots & \ell_{k+2,k} & & 1 & & \\
\vdots & \ddots & \vdots & & & \ddots & \\
\ell_{n,1} & \cdots & \ell_{n,k} & & & & 1
\end{pmatrix} \begin{pmatrix} U_{k+1} & S_{k+1} \\ P_{k+1} & Q_{k+1} \end{pmatrix} \\
&= \begin{pmatrix} L_{k+1} & O \\ R_{k+1} & I \end{pmatrix} \begin{pmatrix} U_{k+1} & S_{k+1} \\ P_{k+1} & Q_{k+1} \end{pmatrix} \\
&= \begin{pmatrix} L_{k+1} U_{k+1} & L_{k+1} P_{k+1} \\ R_{k+1} U_{k+1} + P_{k+1} & R_{k+1} S_{k+1} + Q_{k+1} \end{pmatrix}.
\end{aligned}$$

که درآن U_{k+1} یک ماتریس بالامثلثی $(k+1) \times (k+1)$ است. لذا خواهیم داشت

$$A_{k+1} = L_{k+1} U_{k+1},$$

از اینکه تمام زیرماتریس‌های اصلی پیشروی ماتریس A نامنفرد هستند، نتیجه می‌شود $\det(A_{k+1}) \neq 0$ و بنابراین $\det(U_{k+1}) \neq 0$. لذا، $a^{(k)}_{k+1,k+1} \neq 0$ که نشان می‌دهد درایهٔ محوری گام بعدی ناصفر است. این نتیجه نشان می‌دهد که فرایند حذفی گاوس برای محاسبهٔ تجزیهٔ LU بدون شکست انجام می‌گیرد. □

نتیجه ۱.۳ اگر ماتریس A دارای تجزیهٔ LU باشد، آنگاه

$$u_{11} = a_{11}, \quad u_{kk} = \frac{\det(A_k)}{\det(A_{k-1})}, \quad k = 2, \ldots, n,$$

که در آن A_i، i امین زیرماتریس اصلی پیشرو ماتریس A است و u_{ii} ها درایه‌های قطری U می‌باشند.

برهان: روشن است که $u_{11} = a_{11}$. برای نشان دادن رابطهٔ دوم، با استفاده از قضیهٔ ۲.۳ داریم

$$\det(A_k) = \det(L_k)\det(U_k) = 1 \times u_{11} \cdots u_{k-1,k-1} u_{kk} = \det(A_{k-1}) u_{k,k},$$

که نتیجه لازم بدست می‌دهد. □

۱.۲.۳ الگوریتم دولیتل برای محاسبهٔ تجزیهٔ LU

فرض کنید تمام زیرماتریس‌های اصلی پیشرو ماتریس A نامنفرد باشند. در این صورت، ماتریس A دارای یک تجزیهٔ LU مثل

$$A = LU = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \ell_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix},$$

است. سطر اول ماتریس L کاملاً معلوم است. با ضرب این سطر در ستون j ام از ماتریس U، خواهیم داشت:

$$u_{1j} = a_{1j}, \quad j = 1, \ldots, n.$$

بنابراین سطر و ستون اول ماتریس U کاملاً معلوم می‌شوند. حال سطرهای دوم تا n ام ماتریس L را در ستون اول U ضرب می‌کنیم. در این صورت داریم:

$$a_{i۱} = \ell_{i۱} u_{۱۱}, \quad i = ۲, \ldots, n,$$

و بنابراین

$$\ell_{i۱} = \frac{a_{i۱}}{u_{۱۱}}, \quad i = ۲, \ldots, n.$$

در نتیجه به طور همزمان ستون اول و سطر دوم ماتریس L کاملاً معلوم می‌شوند. حال فرض کنید که $k-۱$ ستون اول از L و $k-۱$ سطر اول از U معلوم باشند. از اینکه L یک ماتریس پایین‌مثلثی است، سطر k ام ماتریس L به صورت

$$(\ell_{k۱}, \ell_{k۲}, \ldots, \ell_{k,k-۱}, ۱, ۰, \ldots, ۰),$$

است. درایه‌های $\ell_{k۱}$، $\ell_{k۲}$، ...، $\ell_{k,k-۱}$ در $k-۱$ ستون اول L قرار دارند و معلوم هستند. بنابراین سطر k ام L کاملاً معلوم است. این سطر را در ستون‌های k ام تا n ام U ضرب می‌کنیم. در این صورت داریم:

$$a_{kj} = \sum_{m=۱}^{k-۱} \ell_{km} u_{mj} + u_{kj}, \quad j = k, \ldots, n.$$

از این‌رو

$$u_{kj} = a_{kj} - \sum_{m=۱}^{k-۱} \ell_{km} u_{mj}, \quad j = k, \ldots, n.$$

در نتیجه سطر k ام از ماتریس U به طور کامل به دست می‌آید. با توجه به اینکه u_{kk} معلوم شده است، ستون k ام از U نیز کاملاً معلوم می‌شود. حال سطرهای $(k+۱)$ ام تا n ام ماتریس L را در ستون k ام ماتریس U ضرب می‌کنیم. در این صورت،

$$a_{ik} = \sum_{m=۱}^{k-۱} \ell_{im} u_{mk} + \ell_{ik} u_{kk}, \quad i = k+۱, \ldots, n.$$

لذا

$$\ell_{ik} = \frac{۱}{u_{kk}} (a_{ik} - \sum_{m=۱}^{k-۱} \ell_{im} u_{mk}), \quad i = k+۱, \ldots, n.$$

از این رابطه نتیجه می‌شود که ستون k ام از ماتریس L نیز معلوم می‌شود. با ادامهٔ این روند ماتریس‌های L و U به دست می‌آیند. با توجه به مطالب فوق الگوریتم دولیتل برای محاسبهٔ تجزیهٔ LU یک ماتریس را می‌توان به صورت زیر جمع‌بندی کرد.

الگوریتم ۲.۳: الگوریتم دولیتل

1. $u_{1j} := a_{1j}, \quad j = 1, \ldots, n$
2. $\ell_{i1} := a_{i1}/u_{11}, \quad i = 2, \ldots, n$
3. For $k = 2, \ldots, n$, Do
4. For $j = k, \ldots, n$ Do
5. $u_{kj} := a_{kj} - \sum_{m=1}^{k-1} \ell_{km} u_{mj}$
6. EndDo
7. For $i = k+1, \ldots, n$ Do
8. $\ell_{ik} := \frac{1}{u_{kk}}(a_{ik} - \sum_{m=1}^{k-1} \ell_{im} u_{mk})$
9. EndDo
10. EndDo

مثال ۸.۳ تجزیهٔ LU ماتریس

$$A = \begin{pmatrix} 2 & 4 & 2 \\ -2 & -5 & -3 \\ 4 & 7 & 6 \end{pmatrix},$$

را با استفاده از الگوریتم دولیتل به دست آورید.

حل: فرض کنید:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

با استفاده از الگوریتم دولیتل داریم:

$$\begin{aligned}
u_{۱۱} &= a_{۱۱} = ۲, \quad u_{۱۲} = a_{۱۲} = ۴, \quad u_{۱۳} = a_{۱۳} = ۲,\\
\ell_{۲۱} &= \frac{a_{۲۱}}{u_{۱۱}} = -۱, \quad \ell_{۳۱} = \frac{a_{۳۱}}{u_{۱۱}} = ۲,\\
u_{۲۲} &= a_{۲۲} - \ell_{۲۱}u_{۱۲} = -۵ - (-۱)\times ۴ = -۱,\\
u_{۲۳} &= a_{۲۳} - \ell_{۲۱}u_{۱۳} = -۳ - (-۱)\times ۲ = -۱,\\
\ell_{۳۲} &= \frac{۱}{u_{۲۲}}(a_{۳۲} - \ell_{۳۱}u_{۱۲}) = \frac{۱}{-۱}(۷ - ۲\times ۴) = ۱,\\
u_{۳۳} &= a_{۳۳} - \ell_{۳۱}u_{۱۳} - \ell_{۳۲}u_{۲۳} = ۶ - ۲\times ۲ - ۱(-۱) = ۳.
\end{aligned}$$

بنابراین

$$L = \begin{pmatrix} ۱ & ۰ & ۰\\ -۱ & ۱ & ۰\\ ۲ & ۱ & ۱ \end{pmatrix}, \qquad U = \begin{pmatrix} ۲ & ۴ & ۲\\ ۰ & -۱ & -۱\\ ۰ & ۰ & ۳ \end{pmatrix}.$$

۲.۲.۳ کاربردهای تجزیهٔ LU

(الف) **حل دستگاه معادلات خطی.** فرض کنید تجزیهٔ LU ماتریس A در دست باشد، یعنی $A = LU$. در این صورت، برای حل دستگاه معادلات خطی $Ax = b$ داریم $LUx = b$. قرار می‌دهیم $y = Ux$. در این صورت $Ly = b$ و این دستگاه با استفاده از الگوریتم جایگذاری پیشرو قابل حل است. بعد از محاسبهٔ بردار y دستگاه $Ux = y$ با استفاده از الگوریتم جایگذاری پسرو حل می‌شود.

مثال ۹.۳ دستگاه

$$\begin{pmatrix} ۲ & ۴ & ۲\\ -۲ & -۵ & -۳\\ ۴ & ۷ & ۶ \end{pmatrix}\begin{pmatrix} x_{۱}\\ x_{۲}\\ x_{۳} \end{pmatrix} = \begin{pmatrix} ۶\\ -۸\\ ۷ \end{pmatrix},$$

را با استفاده از تجزیهٔ LU ماتریس A به دست آورید.

حل: با استفاده از تجزیهٔ LU به دست آمده در مثال ۸.۳ ابتدا دستگاه

$$\begin{pmatrix} ۱ & ۰ & ۰ \\ -۱ & ۱ & ۰ \\ ۲ & ۱ & ۱ \end{pmatrix} \begin{pmatrix} y_۱ \\ y_۲ \\ y_۳ \end{pmatrix} = \begin{pmatrix} ٦ \\ -۸ \\ ۷ \end{pmatrix},$$

را حل می‌کنیم. با حل این دستگاه با استفاده از جایگذاری پیشرو خواهیم داشت $y = (٦, -۲, -۳)^T$. حال دستگاه

$$\begin{pmatrix} ۲ & ۴ & ۲ \\ ۰ & -۱ & -۱ \\ ۰ & ۰ & ۳ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = \begin{pmatrix} ٦ \\ -۲ \\ -۳ \end{pmatrix},$$

را حل می‌کنیم. با حل این دستگاه با استفاده از جایگذاری پسرو خواهیم داشت $x = (-۲, \ \ ۳, -۱)^T$.

(ب) **محاسبهٔ معکوس ماتریس.** فرض کنید B معکوس ماتریس A باشد و $A = LU$. در این صورت از $AB = I$، داریم:

$$A(b_۱, b_۲, \dots, b_n) = (e_۱, e_۲, \dots, e_n),$$

یا

$$Ab_i = e_i, \quad i = ۱, ۲, \dots, n. \tag{۵.۳}$$

بنابراین برای محاسبهٔ معکوس ماتریس A کافی است n دستگاه (۵.۳) حل شوند. با توجه به اینکه ماتریس ضرایب همه این دستگاه‌ها ماتریس A است، هر کدام را می‌توان با استفاده از روند معرفی شده در قسمت (۱) حل کرد. لازم به ذکر است که این دستگاه‌ها را می‌توان مستقل از هم و به طور هم‌زمان حل کرد.

مثال ۱۰.۳ همانند مثال ۹.۳ دستگاه‌های

$$Ab_۱ = \begin{pmatrix} ۱ \\ ۰ \\ ۰ \end{pmatrix}, \quad Ab_۲ = \begin{pmatrix} ۰ \\ ۱ \\ ۰ \end{pmatrix}, \quad Ab_۳ = \begin{pmatrix} ۰ \\ ۰ \\ ۱ \end{pmatrix},$$

را حل می‌کنیم. با حل این دستگاه‌ها خواهیم داشت:

$$A^{-۱} = (b_۱, b_۲, b_۳) = \begin{pmatrix} \frac{۳}{۲} & \frac{۵}{۳} & \frac{۱}{۳} \\ ۰ & -\frac{۲}{۳} & -\frac{۱}{۳} \\ -۱ & -\frac{۱}{۳} & \frac{۱}{۳} \end{pmatrix}.$$

(ج) **محاسبهٔ مقادیر ویژهٔ یک ماتریس.** یکی دیگر از کاربردهای تجزیهٔ LU محاسبهٔ مقادیر ویژهٔ یک ماتریس است که در فصل ۴ به آن اشاره خواهد شد (برای بحث اجمالی به [۱۱] مراجعه نمایید).

۳.۳ تجزیهٔ $\mathrm{LDL^H}$ ماتریس‌های هرمیتی

فرض کنید $A \in \mathbb{C}^{n\times n}$ یک ماتریس هرمیتی و تجزیهٔ LU آن $A = LU$ باشد. در این صورت،

$$\begin{aligned} U &= \begin{pmatrix} u_{۱۱} & u_{۱۲} & \cdots & u_{۱n} \\ ۰ & u_{۲۲} & \cdots & u_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & u_{nn} \end{pmatrix} \\ &= \begin{pmatrix} u_{۱۱} & ۰ & \cdots & ۰ \\ ۰ & u_{۲۲} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} ۱ & u_{۱۲}/u_{۱۱} & \cdots & u_{۱n}/u_{۱۱} \\ ۰ & ۱ & \cdots & u_{۲n}/u_{۲۲} \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & ۱ \end{pmatrix} \\ &=: D\tilde{U}. \end{aligned}$$

بنابراین $A = LD\tilde{U}$. در نتیجه از اینکه $A = A^H$، داریم $LD\tilde{U} = \tilde{U}^H D^H L^H$، یا

$$D^{-H}\tilde{U}^{-H}LD = L^H\tilde{U}^{-۱}.$$

به سادگی می‌توان دید که سمت چپ رابطهٔ اخیر یک ماتریس پایین‌مثلثی با درایه‌های قطری یک و سمت راست یک ماتریس بالامثلثی با درایه‌های قطری یک است. از این‌رو

داریم $L^H \tilde{U}^{-1} = I$. از اینجا نتیجه می‌شود که $\tilde{U} = L^H$. بنابراین اگر ماتریس A هرمیتی باشد و تجزیهٔ LU آن موجود باشد، آنگاه این تجزیه را می‌توان به صورت $A = LDL^H$ نوشت که در آن L یک ماتریس پایین‌مثلثی با درایه‌های قطری یک و D یک ماتریس قطری است. به عنوان یک نتیجه، اگر ماتریس $A \in \mathbb{R}^{n\times n}$ متقارن و دارای تجزیهٔ LU باشد، آنگاه این تجزیه را می‌توان به صورت $A = LDL^T$ نوشت.

مثال ۱۱.۳ تجزیهٔ $\mathrm{LDL^T}$ ماتریس

$$A = \begin{pmatrix} 4 & 8 & 12 \\ 8 & 19 & 27 \\ 12 & 27 & 41 \end{pmatrix},$$

را به دست آورید.

حل: داریم $A = LU$، که در آن

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & 8 & 12 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

در نتیجه داریم:

$$A = LU = L\begin{pmatrix} 4 & 8 & 12 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix} = L\begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} L^T = LDL^T.$$

۴.۳ ماتریس‌های معین مثبت

تعریف ۲.۳ ماتریس $A \in \mathbb{R}^{n\times n}$ را معین مثبت متقارن می‌گوییم، هرگاه

(الف) $A^T = A$ (A متقارن باشد)؛

(ب) برای هر $0 \neq x \in \mathbb{R}^n$، $x^T A x > 0$.

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

تعریف ۳.۳ ماتریس $A \in \mathbb{C}^{n\times n}$ را معین مثبت هرمیتی می‌گوییم، هرگاه
(الف) $A^H = A$ (A هرمیتی باشد)؛
(ب) برای هر $0 \neq x \in \mathbb{C}^n$، $x^H Ax > 0$.
ماتریس A را نیمه معین مثبت هرمیتی می‌گویند هرگاه (الف) برقرار باشد و برای هر $x \in \mathbb{C}^n$، $x^H Ax \geqslant 0$.

ذکر دو نکته در اینجا ضروری به نظر می‌رسد. ابتدا در تعریف ۳.۳ توجه کنید که اگر A هرمیتی باشد آنگاه به ازای هر $x \in \mathbb{C}^n$، اسکالر $z = x^H Ax$ حقیقی است، زیرا داریم:

$$\overline{z} = z^H = (x^H Ax)^H = x^H A^H x = x^H Ax = z.$$

همچنین اگر ماتریس $A \in \mathbb{R}^n$ معین مثبت متقارن باشد، آنگاه معین مثبت هرمیتی نیز می‌باشد. زیرا به ازای هر $0 \neq x = r + is \in \mathbb{C}^n$، داریم:

$$x^H Ax = r^T Ar + s^T As > 0.$$

مثال ۱۲.۳ نشان دهید که ماتریس

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

معین مثبت متقارن است.

حل: ماتریس A متقارن است. فرض کنید $0 \neq x = (x_1, x_2, x_3)^T \in \mathbb{R}^3$. در این صورت، با کمی محاسبه می‌بینیم که

$$x^T Ax = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 \geqslant 0.$$

از طرفی $x^T Ax \neq 0$، زیرا در غیر این صورت، داریم $x_1 = x_2 = x_3 = 0$ که با فرض $x \neq 0$ در تناقض است. □

مثال ۱۳.۳ نشان دهید که ماتریس

$$A = \begin{pmatrix} 4 & 3 + 4i \\ 3 - 4i & 25 \end{pmatrix},$$

معین مثبت هرمیتی است.

حل: ابتدا توجه کنید A هرمیتی است. حال فرض کنید $\mathbb{C}^{۲} \ni x = (x_{۱}, x_{۲})^T \neq ۰$. داریم

$$\begin{aligned} x^H A x &= (\overline{x_{۱}} \quad \overline{x_{۲}}) \begin{pmatrix} ۴ & ۳ + ۴i \\ ۳ - ۴i & ۲۵ \end{pmatrix} \begin{pmatrix} x_{۱} \\ x_{۲} \end{pmatrix} \\ &= ۴|x_{۱}|^{۲} + ۲\Re((۳ + ۴i)\overline{x_{۱}} x_{۲}) + ۲۵|x_{۲}|^{۲} \\ &\geqslant ۴|x_{۱}|^{۲} - ۲|۳ + ۴i||\overline{x_{۱}}||x_{۲}| + ۲۵|x_{۲}|^{۲} \\ &= ۴|x_{۱}|^{۲} - ۱۰|x_{۱}||x_{۲}| + ۲۵|x_{۲}|^{۲} \\ &= ۳|x_{۱}|^{۲} + (|x_{۱}| - ۵|x_{۲}|)^{۲}. \end{aligned}$$

لذا $x^H A x \geqslant ۰$. به روشنی اگر $x^H A x = ۰$، آنگاه $x = ۰$. بنابراین به ازای هر $x \neq ۰$، $x^H A x > ۰$. لازم به ذکر است که در نوشتن نامساوی در روابط فوق از $\Re(z) \geqslant -|z|$ استفاده نموده‌ایم.

تعریف ۴.۳ دترمینان هر زیرماتریس اصلی یک ماتریس را یک کهاد (مینور) می‌گویند.

قضیه ۳.۳ فرض کنید A یک ماتریس معین مثبت هرمیتی باشد. در این صورت
(الف) $A^{-۱}$ موجود و معین مثبت هرمیتی است.
(ب) تمام زیرماتریس‌های اصلی A معین مثبت هرمیتی هستند.
(ج) تمام کهادهای ماتریس A مثبت هستند.

برهان: (الف) به برهان خُلف، فرض کنید که ماتریس A معکوس‌پذیر نباشد. در این صورت، بردار ناصفری مثل $x \in \mathbb{C}^n$ وجود دارد به طوری که $Ax = ۰$. در نتیجه $x^H A x = ۰$، که یک تناقض است. برای نشان دادن اینکه $A^{-۱}$ معین مثبت هرمیتی است، ابتدا توجه کنید که

$$(A^{-۱})^H = (A^H)^{-۱} = A^{-۱}.$$

بنابراین $A^{-۱}$ هرمیتی است. از طرفی فرض کنید که $x \neq ۰$، در این صورت،

$$x^H A^{-۱} x = x^H A^{-۱} A A^{-۱} x = x^H A^{-H} A A^{-۱} x = (A^{-۱} x)^H A (A^{-۱} x).$$

اگر $y = A^{-۱}x$، آنگاه $y \neq ۰$ (چرا؟). در نتیجه

$$x^H A^{-۱} x = y^H A y > ۰.$$

بنابراین $A^{-۱}$ نیز معین مثبت هرمیتی است.

(ب) هر زیرماتریس اصلی ماتریس A به صورت

$$\tilde{A} = \begin{pmatrix} a_{i_۱ i_۱} & a_{i_۱ i_۲} & \cdots & a_{i_۱ i_k} \\ a_{i_۲ i_۱} & a_{i_۲ i_۲} & \cdots & a_{i_۲ i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k i_۱} & a_{i_k i_۲} & \cdots & a_{i_k i_k} \end{pmatrix},$$

است. واضح است که $\tilde{A}$ هرمیتی است. فرض کنید $۰ \neq \tilde{x} = (\tilde{x}_۱, \tilde{x}_۲, \ldots, \tilde{x}_k)^T \in \mathbb{C}^k$.
بردار $x \in \mathbb{C}^n$ را با مؤلفه‌های

$$\begin{cases} x_{i_j} = \tilde{x}_j, & j = ۱, ۲, \ldots, k, \\ x_{i_j} = ۰, & \text{در غیر این صورت}, \end{cases}$$

تعریف می‌کنیم. در این صورت، داریم:

$$\tilde{x}^H \tilde{A} \tilde{x} = x^H A x > ۰.$$

این رابطه نشان می‌دهد که ماتریس $\tilde{A}$ معین مثبت هرمیتی است.

(ج) کافی است نشان دهیم که اگر A معین مثبت هرمیتی باشد، آنگاه $\det(A) > ۰$. این مطلب را به استقراء ثابت می‌کنیم. فرض کنید $n = ۱$. در این صورت $A = (a_{۱۱})$ و

$$\det(A) = a_{۱۱} = (۱)^H A\ (۱) > ۰.$$

فرض کنید دترمینان هر ماتریس معین مثبت هرمیتی از مرتبهٔ $n - ۱$ مثبت باشد و $A = (a_{ij})$ یک ماتریس از مرتبهٔ n و معین مثبت هرمیتی باشد. به‌علاوه فرض کنید $A^{-۱} = (\alpha_{ij})$. داریم:

$$\alpha_{۱۱} = e_۱^H A^{-۱} e_۱ > ۰.$$

از طرفی

$$\alpha_{۱۱} = \frac{\Delta_{۱۱}}{\det(A)} = \frac{(-۱)^{۱+۱}}{\det(A)} \det \begin{pmatrix} a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \ddots & \vdots \\ a_{n۲} & \cdots & a_{nn} \end{pmatrix}.$$

بنابراین با توجه به فرض استقراء و اینکه $\alpha_{۱۱} > ۰$، نتیجه می‌گیریم $\det(A) > ۰$. □

قضیه ۴.۳ فرض کنید ماتریس $A \in \mathbb{C}^{n\times n}$ معین مثبت هرمیتی باشد. در این صورت،
(الف) درایه‌های قطری A مثبت هستند.
(ب) به ازای هر $i \neq j$، $|a_{ij}| < \sqrt{a_{ii}a_{jj}}$.
(ج) به ازای هر $i \neq j$، $\Re(a_{ij}) < (a_{ii} + a_{jj})/۲$.
(د) بزرگترین درایهٔ ماتریس از حیث قدرمطلق روی قطر ماتریس قرار دارد.

برهان: (الف) فرض کنید e_i ستون i ام ماتریس همانی $n \times n$ باشد. در این صورت،

$$a_{ii} = e_i^H A e_i > ۰, \quad i = ۱, \ldots, n.$$

(ب) زیرماتریس اصلی

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

را در نظر بگیرید. داریم $a_{ji} = \overline{a_{ij}}$. با توجه به قسمت (ج) از قضیهٔ ۳.۳ دترمینان ماتریس B مثبت است. بنابراین

$$a_{ii}a_{jj} - a_{ij}a_{ji} > ۰ \Longrightarrow a_{ii}a_{jj} - a_{ij}\overline{a_{ij}} > ۰ \Longrightarrow |a_{ij}| < \sqrt{a_{ii}a_{jj}}.$$

(ج) ماتریس B تعریف شده در قسمت (ب) را در نظر بگیرید. با توجه به قسمت (ب) از قضیهٔ ۳.۳ ماتریس B معین مثبت هرمیتی است. بنابراین

$$۰ < \begin{pmatrix} ۱ & -۱ \end{pmatrix} \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} \begin{pmatrix} ۱ \\ -۱ \end{pmatrix} = a_{ii} + a_{jj} - ۲\Re(a_{ij}),$$

که نتیجهٔ لازم را به دست می‌دهد.
(د) به برهان خلف، فرض کنید درایهٔ غیرقطری a_{ij} بزرگترین درایهٔ ماتریس از حیث

قدرمطلق باشد. در این صورت، داریم $a_{ii} \geqslant |a_{ij}|$ و $a_{jj} \geqslant |a_{ij}|$. در نتیجه

$$|a_{ij}| \geqslant \sqrt{a_{ii}a_{jj}},$$

که با قسمت (ب) در تناقض است. □

لازم به ذکر است که اگر ماتریس A معین مثبت متقارن باشد، آنگاه با توجه به قسمت (ج) از قضیهٔ ۴.۳ داریم $a_{ij} < (a_{ii} + a_{jj})/۲$.

مثال ۱۴.۳ ماتریس‌های

$$A = \begin{pmatrix} ۹ & -۱ & ۵ \\ -۱ & ۰ & ۱ \\ ۵ & ۱ & ۱۰ \end{pmatrix}, \quad B = \begin{pmatrix} ۸ & -۱ & ۹ \\ -۱ & ۸ & ۱ \\ ۹ & ۱ & ۷ \end{pmatrix}$$

معین مثبت نیستند، زیرا یکی از درایه‌های قطری A صفر و بزرگترین درایه از حیث قدرمطلق در ماتریس B روی قطر قرار ندارد.

قضیه ۵.۳ برای هر ماتریس معین مثبت هرمیتی $A \in \mathbb{C}^{n\times n}$ یک ماتریس یکتای پایین‌مثلثی L با درایه‌های قطری مثبت وجود دارد به طوری که $A = LL^H$.

برهان: با توجه به اینکه تمام زیرماتریس‌های اصلی پیشرو A نامنفرد هستند، این ماتریس دارای تجزیهٔ یکتای LU است. با توجه به اینکه A هرمیتی است، این تجزیه را می‌توان به صورت $A = \tilde{L}\tilde{D}\tilde{L}^H$ نوشت که در آن $\tilde{L}$ یک ماتریس پایین‌مثلثی با درایه‌های قطری ۱ است. حال داریم:

$$\tilde{D}_{ii} = e_i^H \tilde{D} e_i = e_i^H \tilde{L}^{-۱} A \tilde{L}^{-H} e_i = (\tilde{L}^{-H}e_i)^H A (\tilde{L}^{-H}e_i) > ۰, \quad i = ۱, \ldots, n.$$

این رابطه نشان می‌دهد که درایه‌های قطری $\tilde{D}$ مثبت هستند. قرار می‌دهیم:

$$\tilde{D}^{\frac{۱}{۲}} = \mathrm{diag}(\sqrt{\tilde{D}_{۱۱}}, \ldots, \sqrt{\tilde{D}_{nn}}),$$

و $L = \tilde{L}\tilde{D}^{\frac{۱}{۲}}$. در این صورت،

$$A = \tilde{L}\tilde{D}\tilde{L}^H = \tilde{L}\tilde{D}^{\frac{۱}{۲}}\tilde{D}^{\frac{۱}{۲}}\tilde{L}^H = LL^H,$$

که اثبات قضیه را کامل می‌کند. □

لازم به ذکر است که دو قضیهٔ قبل برای حالتی که A یک ماتریس حقیقی می‌باشد نیز برقرار است و کافیست بجای کلمه "هرمیتی" از کلمه "متقارن" و بجای اندیس "H" از "T" استفاده شود.

اگرچه قضیهٔ ۵.۳ یک روش برای محاسبهٔ تجزیهٔ $\mathrm{LL^H}$ یک ماتریس را می‌دهد، اما این روش در عمل مناسب نیست زیرا برای محاسبهٔ تجزیهٔ $\mathrm{LL^H}$ لازم است ابتدا تجزیهٔ LU ماتریس محاسبه گردد. در اینجا یک الگوریتم کارا، که به الگوریتم چولسکی معروف است، ارائه می‌شود که تجزیهٔ $\mathrm{LL^H}$ را به طور مستقیم محاسبه می‌کند. الگوریتم چولسکی را برای یک ماتریس حقیقی بیان می‌کنیم، اما به سادگی می‌توان این الگوریتم را برای حالت مختلط نیز تعمیم داد. با توجه به قضیهٔ ۵.۳ و اینکه ماتریس $A = (a_{ij})$ معین مثبت متقارن است، تجزیهٔ $\mathrm{LL^T}$ آن به صورت زیر است:

$$A = LL^T = \begin{pmatrix} \ell_{11} & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \cdots & \ell_{nn} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{21} & \cdots & \ell_{n1} \\ 0 & \ell_{22} & \cdots & \ell_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \ell_{nn} \end{pmatrix}.$$

با برابر قراردادن ستون اول طرفین این رابطه، داریم:

$$\ell_{11}^2 = a_{11} \Longrightarrow \ell_{11} = \sqrt{a_{11}},$$
$$a_{i1} = \ell_{11}\ell_{i1}, \;\; i = 2, \ldots, n \Longrightarrow \ell_{i1} = \frac{a_{i1}}{\ell_{11}}, \;\; i = 2, \ldots, n.$$

بنابراین ستون اول ماتریس L محاسبه شده است. فرض کنید $k-1$ ستون اول L ($k-1$ سطر اول L^T) محاسبه شده باشند. در این صورت،

$$a_{kk} = \ell_{k1}^2 + \ell_{k2}^2 + \cdots + \ell_{kk}^2, \quad \ell_{kk} > 0, \tag{۶.۳}$$

$$a_{ik} = \ell_{i1}\ell_{k1} + \ell_{i2}\ell_{k2} + \cdots + \ell_{ik}\ell_{kk}, \quad i = k+1, \ldots, n. \tag{۷.۳}$$

با استفاده از این دو رابطه ستون k ام ماتریس L محاسبه می‌گردد. با اجرای روند فوق به ازای $k = 2, \ldots, n$، ماتریس L محاسبه می‌گردد. با توجه به روابط (۶.۳) و (۷.۳)، داریم:

$$\ell_{kk} = \sqrt{a_{kk} - (\ell_{k1}^2 + \ell_{k2}^2 + \cdots + \ell_{k,k-1}^2)},$$

$$\ell_{ik} = \frac{1}{\ell_{kk}}(a_{ik} - (\ell_{i1}\ell_{k1} + \ell_{i2}\ell_{k2} + \cdots + \ell_{i,k-1}\ell_{k,k-1})), \quad i = k+1, \ldots, n.$$

بنابراین الگوریتم چولسکی را می‌توان به صورت زیر جمع‌بندی نمود.

الگوریتم ۳.۳: الگوریتم چولسکی

1. $\ell_{11} := \sqrt{a_{11}}$ and $\ell_{i1} := \dfrac{a_{i1}}{\ell_{11}}, \quad i = 2, \ldots, n$
2. For $k = 2, \ldots, n$, Do
3. $\qquad \ell_{kk} := \sqrt{a_{kk} - (\ell_{k1}^2 + \ell_{k2}^2 + \ldots + \ell_{k,k-1}^2)}$
4. $\qquad$ For $i = k+1, \ldots, n$, Do
5. $\qquad\qquad \ell_{ik} = \dfrac{1}{\ell_{kk}}\left(a_{ik} - (\ell_{i1}\ell_{k1} + \ell_{i2}\ell_{k2} + \ldots + \ell_{i,k-1}\ell_{k,k-1})\right)$
6. $\qquad$ EndDo
7. EndDo

نتیجه ۲.۳ از رابطهٔ (۶.۳)، داریم:

$$|\ell_{kj}| \leqslant \sqrt{a_{kk}}, \quad j = 1, \ldots, k, \quad k = 1, \ldots, n.$$

این رابطه نشان می‌دهد که درایه‌های ماتریس L نمی‌توانند خیلی بزرگ شوند.

مثال ۱۵.۳ تجزیهٔ چولسکی ماتریس

$$A = \begin{pmatrix} 4 & -2 & 4 & 2 \\ -2 & 10 & -2 & -7 \\ 4 & -2 & 8 & 4 \\ 2 & -7 & 4 & 7 \end{pmatrix},$$

را به دست آورید.

حل: داریم

$$\begin{aligned}
\ell_{11} &= \sqrt{4} = 2 \\
\ell_{21} &= \frac{a_{21}}{\ell_{11}} = \frac{-2}{2} = -1, \\
\ell_{31} &= \frac{a_{31}}{\ell_{11}} = 2, \\
\ell_{41} &= \frac{a_{41}}{\ell_{11}} = 1, \\
\ell_{22} &= \sqrt{a_{22} - \ell_{21}^2} = \sqrt{10 - (-1)^2} = 3, \\
\ell_{32} &= \frac{1}{\ell_{22}}(a_{32} - \ell_{31}\ell_{21}) = \frac{1}{3}(-2 - 2(-1)) = 0, \\
\ell_{42} &= -2, \\
\ell_{33} &= \sqrt{a_{33} - \ell_{31}^2 - \ell_{32}^2} = \sqrt{8 - 2^2 - 0^2} = 2, \\
\ell_{43} &= \frac{1}{\ell_{33}}(a_{43} - \ell_{41}\ell_{31} - \ell_{42}\ell_{32}) = \frac{1}{2}(4 - 1 \times 2 - (-2) \times 0) = 1, \\
\ell_{44} &= \sqrt{a_{44} - \ell_{41}^2 - \ell_{42}^2 - \ell_{43}^2} = \sqrt{7 - 1^2 - (-2)^2 - 1^2} = 1.
\end{aligned}$$

بنابراین

$$L = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{pmatrix}.$$

۵.۳ ماتریس‌های غالب قطری اکید

در این بخش ردهٔ دیگری از ماتریس‌ها به نام ماتریس‌های غالب قطری را بررسی می‌کنیم. این گونه ماتریس‌ها کاربردهای زیادی در زمینه‌های مختلف علوم و مهندسی دارند.

تعریف ۵.۳ ماتریس $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ را غالب قطری اکید می‌نامند، هرگاه

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|, \quad i = 1, \dots, n.$$

مثال ۱۶.۳ ماتریس

$$A = \begin{pmatrix} -۳ & ۱ & -۱ \\ -۲ & ۴ & ۱ \\ ۲ & ۲ & -۵ \end{pmatrix},$$

غالب قطری اکید است، زیرا

$$\begin{aligned} |-۳| &> |۱| + |-۳|, \\ |۴| &> |-۲| + |۱|, \\ |-۵| &> |۲| + |۲|. \end{aligned}$$

قضیه ۶.۳ اگر $A \in \mathbb{C}^{n\times n}$ یک ماتریس غالب قطری اکید باشد، آنگاه نامنفرد است.

برهان: به برهان خُلف فرض کنید که ماتریس $A = (a_{ij})$ غالب قطری اکید و منفرد باشد. در این صورت، یک بردار ناصفر مثل $\circ \neq x \in \mathbb{C}^n$ وجود دارد به طوری که $Ax = \circ$. بدون از دست دادن کلیت، می‌توان فرض کرد که $\|x\|_\infty = |x_i| = ۱$. با استفاده از سطر i ام تساوی $Ax = \circ$، داریم:

$$\begin{aligned} \circ = |\sum_{j=۱}^{n} a_{ij}x_j| &\geqslant |a_{ii}x_i + \sum_{j=۱, j\neq i}^{n} a_{ij}x_j| \\ &\geqslant |a_{ii}| - |\sum_{j=۱, j\neq i}^{n} a_{ij}x_j| \\ &\geqslant |a_{ii}| - \sum_{j=۱, j\neq i}^{n} |a_{ij}||x_j| \\ &\geqslant |a_{ii}| - \sum_{j=۱, j\neq i}^{n} |a_{ij}|, \end{aligned}$$

که این یک تناقض است. بنابراین ماتریس A نامنفرد است. $\square$

روشن است اگر $A \in \mathbb{C}^{n\times n}$ غالب قطری اکید باشد، آنگاه تمام زیرماتریس‌های اصلی پیشرو آن نیز غالب قطری اکید و بنابراین نامنفرد می‌باشند. این نتیجه نشان می‌دهد که تجزیهٔ LU آن موجود است.

۶.۳ خطای ناشی از اختلال در حل دستگاه معادلات خطی

دستگاه $Ax = b$ را در نظر بگیرید. می‌خواهیم ببینیم که اختلال در داده‌های این دستگاه، یعنی ماتریس A و بردار b چه تاثیری در جواب دستگاه دارد. بررسی این مسأله لازم است چون به طور طبیعی وقتی که داده‌ها به کامپیوتر داده می‌شوند با توجه به اینکه کامپیوتر برای ذخیره‌سازی قسمت کسری عدد محدودیت دارد، اعداد با خطا (هر چند ناچیز) ذخیره می‌شوند. دستگاه معادلات خطی

$$\begin{pmatrix} ۱ & ۲ & ۳ \\ ۱ & ۴ & ۲ \\ ۲ & ٦ & ۴.۹۹۹ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = \begin{pmatrix} ٦ \\ ۷ \\ ۱۲.۹۹۹ \end{pmatrix},$$

در نظر بگیرید. جواب واقعی این دستگاه $x_۱ = x_۲ = x_۳ = ۱$ می‌باشد . حال درایهٔ سوم از بردار سمت راست را به ۱۳ تبدیل می‌کنیم. در نتیجه، جواب دستگاه $x_۱ = ۵$، $x_۲ = ۰.۵$ و $x_۳ = ۰$ خواهد بود که با جواب قبلی تفاوت فاحش دارد. ملاحظه می‌شود که این تغییر ناشی از وارد کردن خطا به اندازه ۰.۰۰۱ در یک مؤلفه از بردار سمت راست می‌باشد. دستگاه فوق را یک دستگاه بدحالت نامند. می‌خواهیم ببینیم که بدحالتی یک دستگاه را چگونه می‌توان تشخیص داد. در این بخش فرض می‌کنیم که نرم‌های ماتریسی مورد استفاده، طبیعی هستند.

لم ۱.۳ فرض کنید $F \in \mathbb{R}^{n\times n}$ و $\|F\| < ۱$. آنگاه ماتریس $I + F$ (I ماتریس همانی) نامنفرد است و

$$\|(I+F)^{-۱}\| \leqslant \frac{۱}{۱ - \|F\|}.$$

برهان: به برهان خُلف، فرض کنید که $I+F$ منفرد باشد. در این صورت، برداری ناصفری مثل x موجود است به طوری که $(I+F)x = ۰$. در نتیجه $x = -Fx$ و بنابراین

$$\|x\| = \|-Fx\| \leqslant \|F\|\|x\|.$$

این رابطه معادل است با $\|F\| \geqslant ۱$، که یک تناقض است. برای اثبات قسمت دوم، قرار می‌دهیم $C = (I+F)^{-۱}$. در نتیجه

$$۱ = \|I\| = \|(I+F)C\| = \|C + FC\|$$

$$
\begin{aligned}
&\geqslant \|C\| - \|FC\| \\
&\geqslant \|C\| - \|F\|\|C\| = \|C\|(1 - \|F\|).
\end{aligned}
$$

بنابراین

$$\|C\| \leqslant \frac{1}{1 - \|F\|},$$

که اثبات لم را کامل می‌کند. □

قضیه ۷.۳ فرض کنید ماتریس A نامنفرد و x و $x + \Delta x$ به ترتیب جواب دستگاه‌های $Ax = b$ و $(A + \Delta A)(x + \Delta x) = b + \Delta b$ باشند. اگر $\|A^{-1}\|\|\Delta A\| < 1$، آنگاه

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \frac{\mathrm{cond}(A)}{1 - \mathrm{cond}(A)\frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right), \qquad (۸.۳)$$

که در آن $\mathrm{cond}(A) = \|A\|\|A^{-1}\|$.

حل: ابتدا نشان می‌دهیم که $A + \Delta A$ نامنفرد است. داریم $A + \Delta A = A(I + A^{-1}\Delta A)$ و

$$\|A^{-1}\Delta A\| \leqslant \|A^{-1}\|\|\Delta A\| < 1.$$

بنابراین با استفاده از لم ۱.۳ ماتریس $I + A^{-1}\Delta A$ نامنفرد و در نتیجه ماتریس $A + \Delta A$ نیز نامنفرد خواهد بود. برای اثبات رابطهٔ (۸.۳)، با تفریق طرفین دستگاه‌های $Ax = b$ و $(A + \Delta A)(x + \Delta x) = b + \Delta b$، داریم

$$(A + \Delta A)\Delta x = \Delta b - \Delta A x,$$

که با رابطهٔ

$$\Delta x = (A + \Delta A)^{-1}(\Delta b - \Delta A\, x),$$

معادل است. در نتیجه با توجه به لم ۱.۳، داریم:

$$
\begin{aligned}
\|\Delta x\| &\leqslant \|(A + \Delta A)^{-1}\| \left(\|\Delta b\| + \|\Delta A\|\|x\|\right) \\
&\leqslant \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\Delta A\|}\left(\|\Delta b\| + \|\Delta A\|\|x\|\right).
\end{aligned}
$$

در نتیجه

$$\begin{aligned}\frac{\|\Delta x\|}{\|x\|} &\leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|\Delta A\|}\left(\frac{\|\Delta b\|}{\|x\|}+\|\Delta A\|\right)\\ &\leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|\Delta A\|}\left(\frac{\|\Delta b\|\|A\|}{\|b\|}+\|\Delta A\|\right).\end{aligned}$$

توجه کنید که برای به دست آوردن آخرین نامساوی از

$$\frac{1}{\|x\|} \leqslant \frac{\|A\|}{\|b\|},$$

استفاده کرده‌ایم. در نتیجه

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \frac{\|A^{-1}\|\|A\|}{1-\|A^{-1}\|\|\Delta A\|}\left(\frac{\|\Delta b\|}{\|b\|}+\frac{\|\Delta A\|}{\|A\|}\right).$$

این رابطه اثبات قضیه را کامل می‌کند. □

در قضیهٔ قبل مقدار $\mathrm{cond}(A)$ را عدد حالت ماتریس A می‌گویند. اگر $\mathrm{cond}(A)$ کوچک باشد، آنگاه دستگاه $Ax = b$ را یک دستگاه خوش‌حالت و در صورتی که بزرگ باشد، دستگاه را بدحالت می‌نامند. چند خاصیت مهم $\mathrm{cond}(A)$ به صورت زیر هستند:

(۱) $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$.

(۲) برای هر اسکالر ناصفر c، داریم $\mathrm{cond}(cA) = \mathrm{cond}(A)$.

(۳) اگر $\|.\|$ یک نرم طبیعی باشد، آنگاه برای ماتریس همانی I، داریم $\mathrm{cond}(I) = 1$ و $\mathrm{cond}(A) \geqslant 1$.

اثبات (۱) و (۲) ساده است و برای اثبات (۳) داریم

$$1 = \|I\| = \|AA^{-1}\| \leqslant \|A\|\|A^{-1}\| = \mathrm{cond}(A).$$

بنابراین بهترین مقدار ممکن برای عدد حالت یک ماتریس برابر با ۱ و مقدار عدد حالت برای نرم‌های مختلف، معمولاً متفاوت است.

دوباره به رابطهٔ (۸.۳) برمی‌گردیم. روشن است که اگر عدد حالت ماتریس A، خطای نسبی در ماتریس A و خطای نسبی در بردار b اعداد کوچکی باشند، آنگاه خطای نسبی در بردار جواب نیز عدد کوچکی خواهد. اما اگر عدد حالت ماتریس A بزرگ باشد،

آنگاه ممکن است خطای نسبی نیز در بردار جواب، بزرگ باشد. با استفاده از قضیهٔ قبل دو نتیجهٔ بعدی را می‌توان بیان کرد.

نتیجه ۳.۳ فرض کنید ماتریس A نامنفرد باشد و $b \neq ۰$. اگر x جواب دستگاه معادلات خطی $Ax = b$ و $x + \Delta x$ جواب دستگاه $A(x + \Delta x) = b + \Delta b$ باشد، آنگاه

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \operatorname{cond}(A)\frac{\|\Delta b\|}{\|b\|}. \tag{۹.۳}$$

نتیجه ۴.۳ فرض کنید A یک ماتریس نامنفرد و x و $x+\Delta x$ به ترتیب جواب دستگاه‌های $Ax = b$ و $(A + \Delta A)(x + \Delta x) = b$ باشند. اگر $\|A^{-۱}\|\|\Delta A\| < ۱$، آنگاه

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \frac{\operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}}{۱ - \operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}}. \tag{۱۰.۳}$$

مثال ۱۷.۳ دستگاهی که در ابتدای این بخش آمده است را در نظر بگیرید. داریم:

$$A = \begin{pmatrix} ۱ & ۲ & ۳ \\ ۱ & ۴ & ۲ \\ ۲ & ۶ & ۴.۹۹۹ \end{pmatrix}, \quad b = \begin{pmatrix} ۶ \\ ۷ \\ ۱۲.۹۹۹ \end{pmatrix}.$$

همان‌طور که دیدیم جواب واقعی دستگاه $Ax = b$ برابر $x = (۱, ۱, ۱)^T$ است. بردار b را به بردار

$$b' = b + \Delta b = \begin{pmatrix} ۶ \\ ۷ \\ ۱۲.۹۹۹ \end{pmatrix} + \begin{pmatrix} ۰ \\ ۰ \\ ۰.۰۰۱ \end{pmatrix} = \begin{pmatrix} ۶ \\ ۷ \\ ۱۳ \end{pmatrix},$$

تبدیل می‌کنیم. داریم:

$$\frac{\|\Delta b\|_۲}{\|b\|_۲} = ۶.۲۷۵ \times ۱۰^{-۵},$$

جواب دستگاه $A(x + \Delta x) = b + \Delta b$ برابر است با

$$x' = x + \Delta x = \begin{pmatrix} ۵ \\ ۰.۵ \\ ۰ \end{pmatrix},$$

که تفاوت زیادی با جواب دستگاه اولیه دارد. خطای نسبی در جواب به صورت

$$\frac{\|\Delta x\|_۲}{\|x\|_۲} = ۲.۳۹۸,$$

است. ملاحظه می‌شود که با توجه به اینکه تغییر نسبی در بردار b ناچیز بوده، تغییر نسبی در جواب بسیار بزرگ است. دلیل به دست آمدن چنین جوابی این است که

$$\mathrm{cond}_۲(A) = \|A\|_۲\|A^{-۱}\|_۲ = ۷.۱۱۶ \times ۱۰^۴,$$

عدد بسیار بزرگی است. توجه کنید که داریم:

$$\frac{\|\Delta x\|_۲}{\|x\|_۲} = ۲.۳۹۸ \leqslant \mathrm{cond}_۲(A)\frac{\|\Delta b\|_۲}{\|b\|_۲} = ۴.۴۶۵,$$

و این رابطه نشان می‌دهد که نتیجهٔ ۳.۳ در اینجا صادق است.

مثال ۱۸.۳ ماتریس و بردار مثال قبل را به صورت

$$A = \begin{pmatrix} ۱ & ۲ & ۳ \\ ۱ & ۴ & ۲ \\ ۲ & ۶ & ۴ \end{pmatrix}, \quad b = \begin{pmatrix} ۶ \\ ۷ \\ ۱۲ \end{pmatrix},$$

تبدیل می‌کنیم. جواب واقعی دستگاه $Ax = b$ بردار $x = (۱, ۱, ۱)^T$ است. بردار b را به بردار

$$b' = b + \Delta b = \begin{pmatrix} ۶ \\ ۷ \\ ۱۲ \end{pmatrix} + \begin{pmatrix} ۰.۰۱ \\ -۰.۰۱ \\ -۰.۰۱ \end{pmatrix} = \begin{pmatrix} ۶.۰۱ \\ ۶.۹۹ \\ ۱۱.۹۹ \end{pmatrix},$$

تبدیل می‌کنیم. داریم:

$$\frac{\|\Delta b\|_۲}{\|b\|_۲} = ۰.۰۰۱.$$

جواب دستگاه $A(x + \Delta x) = b + \Delta b$ برابر است با

$$x' = x + \Delta x = \begin{pmatrix} ۰.۹۹ \\ ۰.۹۹۵ \\ ۱.۰۱ \end{pmatrix},$$

که تفاوت چندانی با جواب دستگاه اولیه ندارد. داریم:

$$\frac{\|\Delta b\|_۲}{\|b\|_۲} = ۰.۰۰۱, \quad \frac{\|\Delta x\|_۲}{\|x\|_۲} = ۰.۰۰۹.$$

می‌بینیم که تغییر نسبی در بردار b کوچک و در نتیجه تغییر نسبی در جواب نیز کوچک است. دلیل به دست آمدن چنین جوابی این است که

$$\operatorname{cond}_۲(A) = \|A\|_۲\|A^{-۱}\|_۲ = ۶۵.۸۱۱,$$

عدد کوچکی است و دستگاه خوش‌حالت است. توجه کنید که داریم:

$$\frac{\|\Delta x\|_۲}{\|x\|_۲} = ۰.۰۰۹ \leqslant \operatorname{cond}_۲(A)\frac{\|\Delta b\|_۲}{\|b\|_۲} = ۰.۰۷۵.$$

این رابطه نشان می‌دهد که نتیجهٔ ۳.۳ در اینجا نیز صادق است.

مثال ۱۹.۳ فرض کنید:

$$A = \begin{pmatrix} ۱ & ۱.۹۹۹۳ & ۳ \\ ۲ & ۴ & ۶ \\ ۳ & ۶ & ۸.۹۹۹۲ \end{pmatrix}, \quad b = \begin{pmatrix} ۵.۹۹۹۳ \\ ۱۲.۰۰۰۰ \\ ۱۷.۹۹۹۲ \end{pmatrix}.$$

جواب واقعی دستگاه $Ax = b$ بردار $x = (۱, ۱, ۱)^T$ است. درایهٔ $(۲, ۳)$ از ماتریس A را به ۵.۹۹۹۹ تبدیل می‌کنیم. یعنی ΔA را به صورت

$$\Delta A = \begin{pmatrix} ۰ & ۰ & ۰ \\ ۰ & ۰ & -۰.۰۰۰۱ \\ ۰ & ۰ & ۰ \end{pmatrix},$$

می‌گیریم. جواب دستگاه $(A + \Delta A)x' = b$ عبارت است از

$$x' = x + \Delta x = \begin{pmatrix} ۰.۱۳۱۹ \\ ۱.۰۸۷۹ \\ ۱.۲۳۰۸ \end{pmatrix}.$$

داریم:

$$\frac{\|\Delta A\|_۲}{\|A\|_۲} = ۷.۱۴۳۲ \times ۱۰^{-۶}.$$

با وجود اینکه تغییر نسبی در ماتریس A ناچیز است، داریم:

$$\frac{\|\Delta x\|_۲}{\|x\|_۲} = ۰.۹۵۱۴,$$

که مقداری بزرگ است. دلیل این امر این است که عدد حالت ماتریس A بزرگ است. توجه کنید که $\mathrm{cond}_۲(A) = ۱.۲۳۳۱ \times ۱۰^۵$.

مثال ۲۰.۳ دستگاه $Ax = b$ را در نظر بگیرید که در آن

$$A = \begin{pmatrix} ۱.۲۱ & ۰.۵۱ & ۱.۷۲ \\ ۰.۵۲ & ۱.۱۰ & ۱.۶۳ \\ -۱.۲۱ & ۱.۹۱ & ۰.۷۳ \end{pmatrix}, \quad b = \begin{pmatrix} ۳.۴۴ \\ ۳.۲۵ \\ ۱.۴۳ \end{pmatrix}.$$

جواب واقعی این دستگاه بردار $x = (۱, ۱, ۱)^T$ است. مولفهٔ $(۳, ۱)$ از ماتریس A را از -۱.۲۱ به $-۱.۲۱ - ۱۰^{-۶}$ و درایهٔ اول از بردار b را از ۳.۴۴ را به $۳.۴۴ + ۱۰^{-۳}$ تبدیل می‌کنیم. در این صورت جواب واقعی دستگاه به دست آمده به صورت

$$x + \Delta x = \begin{pmatrix} ۰.۱۴۱۷ \\ ۰.۱۲۶۳ \\ ۱.۸۶۳۴ \end{pmatrix},$$

خواهد بود. می‌بینیم که تغییرات جزئی در ماتریس ضرایب و بردار سمت راست دستگاه باعث تغییرات فاحش درجواب شده است. دلیل به دست آمدن چنین جوابی بدحالت بودن ماتریس ضرایب دستگاه است. در واقع داریم:

$$\mathrm{cond}(A) = \|A\|_۲ \|A^{-۱}\|_۲ = ۷.۸۶ \times ۱۰^۳ \gg ۱.$$

برای بررسی قضیهٔ ۷.۳ قرار می‌دهیم:

$$\Delta A = \begin{pmatrix} ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ \\ -۱۰^{-۶} & ۰ & ۰ \end{pmatrix}, \quad \Delta b = \begin{pmatrix} ۱۰^{-۳} \\ ۰ \\ ۰ \end{pmatrix}.$$

ابتدا توجه کنید که داریم

$$\|A^{-۱}\|_۲ \|\Delta A\| = ۰.۰۰۲۵ < ۱,$$

و شرط اصلی قضیه ۷.۳ برقرار است. از طرفی داریم:

$$\frac{\|\Delta A\|_2}{\|A\|_2} = 3.19 \times 10^{-7}, \qquad \frac{\|\Delta b\|_2}{\|b\|_2} = 2.02 \times 10^{-4}.$$

با توجه به قضیهٔ ۷.۳ باید داشته باشیم:

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leqslant u = \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)\dfrac{\|\Delta A\|_2}{\|A\|_2}} \left(\frac{\|\Delta A\|_2}{\|A\|_2} + \frac{\|\Delta b\|_2}{\|b\|_2}\right) = 1.60.$$

همچنین داریم:

$$\frac{\|\Delta x\|_2}{\|x\|_2} = 0.87 < u = 1.60.$$

تمرینات

۱.۳) دستگاه

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ \beta \end{pmatrix},$$

را در نظر بگیرید. به ازای چه مقادیری از α و β

(الف) دستگاه جواب ندارد.

(ب) دستگاه جواب یکتا دارد.

(ج) دستگاه بی‌شمار جواب دارد.

۲.۳) نشان دهید ماتریس

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -3 & 3 & -2 \\ 9 & 11 & 3 \end{pmatrix},$$

دارای تجزیهٔ LU است. سپس تجزیهٔ LU آن را با استفاده از روش حذفی گاوس و دولیتل به دست آورید.

۳.۳) مستقیماً با استفاده از تعریف نشان دهید که ماتریس

$$A = \begin{pmatrix} ۶ & ۲ & ۱ \\ ۲ & ۴ & ۱ \\ ۱ & ۱ & ۴ \end{pmatrix},$$

معین مثبت متقارن است.

۴.۳) دستگاه

$$\begin{pmatrix} ۲ & ۴ & -۶ \\ ۱ & ۵ & ۳ \\ ۱ & ۳ & ۲ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = \begin{pmatrix} -۴ \\ ۱۰ \\ ۵ \end{pmatrix},$$

را با استفاده از روش حذفی گاوس حل کنید.

۵.۳) دستگاه معادلات خطی

$$\begin{cases} ۲x_۱ + x_۲ + x_۳ + x_۴ = ۵, \\ ۴x_۱ + ۴x_۲ + ۳x_۳ + x_۴ = ۱۲, \\ -۲x_۱ + x_۲ + x_۳ - x_۴ = -۱, \\ ۲x_۲ + ۲x_۳ + ۲x_۴ = ۶, \end{cases}$$

را با استفاده از روش حذفی گاوس با محورگیری جزئی حل کنید.

۶.۳) بدون استفاده از تجزیهٔ چولسکی ماتریس ضرایب دستگاه

$$\begin{pmatrix} ۴ & ۲ & ۶ \\ ۲ & ۱۰ & ۶ \\ ۶ & ۶ & ۱۴ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = \begin{pmatrix} ۲۴ \\ ۱۲ \\ ۴۸ \end{pmatrix},$$

نشان دهید این ماتریس معین مثبت متقارن است. با استفاده از تجزیهٔ چولسکی ماتریس ضرایب دستگاه، این دستگاه را حل کنید. با استفاده از تجزیهٔ به دست آمده، دترمینان ماتریس را به دست آورید.

۷.۳) تجزیه‌ای به صورت $A = LDL^T$ برای ماتریس

$$A = \begin{pmatrix} ۴ & ۸ & ۸ \\ ۸ & ۱۱ & ۱ \\ ۸ & ۱ & -۲۲ \end{pmatrix},$$

بنویسید که در آن L یک پایین‌مثلثی با درایه‌های قطری یک و D یک ماتریس قطری است.

۸.۳) فرض کنید:

$$A = \begin{pmatrix} ۴ & ۲ & ۰ \\ ۰ & ۴ & ٦ \\ ۲ & ۵ & ۲ \end{pmatrix}, \quad b = \begin{pmatrix} ۲ \\ ۸ \\ ۱ \end{pmatrix}.$$

بدون محاسبهٔ تجزیهٔ LU ماتریس A نشان دهید که تجزیهٔ LU آن موجود است. به روش دولیتل تجزیهٔ LU ماتریس A را به دست آورید و با استفاده از آن دستگاه $Ax = b$ را حل کنید.

۹.۳) نشان دهید ماتریس

$$A = \begin{pmatrix} ۴ & -۲ & ۴ & ٦ \\ -۲ & ۲ & -۱ & -۳ \\ ۴ & -۱ & ۱۴ & ۳ \\ ٦ & -۳ & ۳ & ۱۴ \end{pmatrix},$$

معین مثبت متقارن است. سپس تجزیهٔ چولسکی آن را محاسبه کنید.

۱۰.۳) مقادیر a و b را طوری تعیین کنید که ماتریس

$$A = \begin{pmatrix} a & ۱ & b \\ ۱ & a & ۱ \\ ۱ & ۱ & a \end{pmatrix},$$

معین مثبت متقارن باشد.

۱۱.۳) حدودی برای α بیابید به طوری که ماتریس

$$A = \begin{pmatrix} \alpha & ۲ & ۱ \\ ۲ & \alpha + ۴ & ۴ \\ ۱ & ۴ & ۹ \end{pmatrix},$$

معین مثبت متقارن باشد. به ازای $\alpha = ۱$ تجزیهٔ چولسکی ماتریس A را به دست آورید.

۱۲.۳) نشان دهید اگر $\alpha > \sqrt{n - ۱}$، آنگاه ماتریس

$$A_\alpha = \begin{pmatrix} \alpha & -۱ & \cdots & -۱ \\ -۱ & \alpha & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ -۱ & ۰ & \cdots & \alpha \end{pmatrix} \in \mathbb{R}^{n\times n},$$

معین مثبت متقارن است.

۱۳.۳) به ازای چه مقادیری از α تجزیهٔ LU ماتریس زیر موجود است:

$$A = \begin{pmatrix} ۱ & \alpha & \alpha \\ \alpha & ۱ & \alpha \\ \alpha & \alpha & ۱ \end{pmatrix}.$$

۱۴.۳) فرض کنید A یک ماتریس پایین‌مثلثی $n \times n$ با درایه‌های قطری ۱ و درایه‌های زیر قطر ۱− باشد. مقدار $\mathrm{cond}_\infty(A) = \|A\|_\infty \|A^{-۱}\|_\infty$ را محاسبه کنید.

۱۵.۳) تجزیهٔ LU ماتریس

$$A = \begin{pmatrix} ۲ & -۱ & & \\ -۱ & ۲ & \ddots & \\ & \ddots & \ddots & -۱ \\ & & -۱ & ۲ \end{pmatrix} \in \mathbb{R}^{n\times n},$$

را به دست آورید.

۱۶.۳) فرض کنید

$$A = \begin{pmatrix} -۳ & ٦ & ۲ \\ ۱ & ۱ & ۴ \\ ۵ & ۲ & -۱ \end{pmatrix}.$$

یک ماتریس جایگشت مثل P بنویسید به طوری که ماتریس PA یک ماتریس غالب قطری اکید باشد.

۱۷.۳) فرض کنید $A, X \in \mathbb{R}^{n\times n}$ و $x, b \in \mathbb{R}^n$. کدامیک از گزاره‌های زیر درست و کدامیک نادرست هستند. دلیل درستی یا نادرستی را بیان کنید.

(الف) $x^T A x = x^T (A + A^T) x / ۲$.

(ب) اگر به ازای یک $k \in \mathbb{N}$، داشته باشیم $A^k = ۰$، آنگاه $A = ۰$.

(ج) اگر A معین مثبت متقارن و X نامنفرد باشد، آنگاه ماتریس $X^T A X$ معین مثبت متقارن است.

(د) داریم $\|I\|_۲ = \|I\|_F = ۱$.

۱۸.۳) برای هر دو بردار $u, v \in \mathbb{R}^n$،
(الف) مقدار $\det(uv^T)$ چیست؟
(ب) مقدار $\det(I + uv^T)$ چیست؟
(ج) اگر A یک ماتریس نامنفرد باشد، مقدار $\det(A + uv^T)$ چیست؟

۱۹.۳) نشان دهید اگر دترمینان تمام زیرماتریس‌های پیشرو اصلی یک ماتریس متقارن مثبت باشند، آنگاه ماتریس معین مثبت متقارن است.

۲۰.۳) فرض کنید ماتریس $A \in \mathbb{C}^{n\times n}$ معین مثبت هرمیتی باشد. برای هر $x, y \in \mathbb{R}^n$، تعریف می‌کنیم $\langle x, y\rangle_A = \langle Ax, y\rangle$. نشان دهید $\langle .,.\rangle_A$ یک ضرب داخلی روی $\mathbb{R}^n$ است. این ضرب داخلی روی $\mathbb{R}^n$ نرم $\|x\|_A = \sqrt{\langle x, x\rangle_A}$ را تولید می‌کند که به نام نرم$-A$ معروف است.

۲۱.۳) فرض کنید درایه‌های ماتریس $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ به صورت زیر تعریف شده باشند

$$a_{ij} = \begin{cases} n - ۱, & i = j, \\ ۱, & i \neq j, \end{cases}$$

که در آن $n \geqslant ۳$. نشان دهید ماتریس A معین مثبت متقارن است.

۲۲.۳) فرض کنید ماتریس A نامنفرد و $A = LU$ تجزیهٔ LU آن باشد. روشی برای حل دستگاه $A^k x = b$ ارائه نمایید، که در آن k یک عدد طبیعی است.

۲۳.۳) فرض کنید $A \in \mathbb{R}^{n\times n}$. نشان دهید به ازای هر $۰ \neq x \in \mathbb{R}^n$، $x^TAx > ۰$ اگر و تنها اگر $A + A^T$ معین مثبت متقارن باشد.

۲۴.۳) فرض کنید $A, B \in \mathbb{R}^{n\times n}$ دو ماتریس غالب قطری اکید باشند. با یک مثال نشان دهید که $A + B$ لزوماً غالب قطری اکید نیست. تحت چه شرایطی $A + B$ غالب قطری اکید است.

۲۵.۳) ماتریسی مثل $X \in \mathbb{R}^{۲\times ۲}$ بیابید به طوری که $X^۲ = A$، که در آن

$$A = \begin{pmatrix} ۱۷ & ۹ \\ ۹ & ۲۶ \end{pmatrix}.$$

۲۶.۳) فرض کنید

$$A = \begin{pmatrix} ۱ & -۱ & & & \\ -۱ & ۲ & -۱ & & \\ & \ddots & \ddots & \ddots & \\ & & -۱ & ۲ & -۱ \\ & & & -۱ & ۲ \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

نشان دهید که $A = LL^T$ که در آن

$$L = \begin{pmatrix} ۱ & & & & \\ -۱ & ۱ & & & \\ & \ddots & \ddots & & \\ & & -۱ & ۱ & \\ & & & -۱ & ۱ \end{pmatrix}.$$

ماتریس $L^{-۱}$ را محاسبه کرده و با استفاده از آن نشان دهید

$$A^{-۱} = (L^T)^{-۱}L^{-۱} = \begin{pmatrix} n & n-۱ & n-۲ & \cdots & ۲ & ۱ \\ n-۱ & n-۱ & n-۲ & \cdots & ۲ & ۱ \\ n-۲ & n-۲ & n-۲ & \cdots & ۲ & ۱ \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ ۲ & ۲ & ۲ & \cdots & ۲ & ۱ \\ ۱ & ۱ & ۱ & \cdots & ۱ & ۱ \end{pmatrix}.$$

۲۷.۳) فرض کنید

$$A = \begin{pmatrix} ۳۷۵ & ۳۷۴ \\ ۷۵۲ & ۷۵۰ \end{pmatrix}.$$

(الف) معکوس A و $\mathrm{cond}_\infty(A) = \|A\|_\infty \|A^{-۱}\|_\infty$ را محاسبه کنید.

(ب) بردارهای b، Δb، x و Δx را طوری بیابید که $Ax = b$ و $A(x+\Delta x) = b + \Delta b$ و مقدار $\|\Delta b\|_\infty / \|b\|_\infty$ کوچک، اما $\|\Delta x\|_\infty / \|x\|_\infty$ بزرگ باشد.

۲۸.۳) فرض کنید

$$A = \begin{pmatrix} ۱ & -۱ & -۱ & \cdots & -۱ & -۱ \\ ۰ & ۱ & -۱ & \cdots & -۱ & -۱ \\ ۰ & ۰ & ۱ & \cdots & -۱ & -۱ \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ ۰ & ۰ & ۰ & \cdots & ۱ & -۱ \\ ۰ & ۰ & ۰ & \cdots & ۰ & ۱ \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

(الف) نشان دهید

$$A^{-1} = \begin{pmatrix} 1 & 1 & 2 & \cdots & 2^{n-3} & 2^{n-2} \\ 0 & 1 & 1 & \cdots & 2^{n-4} & 2^{n-3} \\ 0 & 0 & 1 & \cdots & 2^{n-5} & 2^{n-4} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

(ب) نشان دهید $\mathrm{cond}_\infty(A) = n2^{n-1}$. روشن است که داریم $\det(A) = 1$. حال فرض کنید $D = \mathrm{diag}(10^{-1}, 10^{-1}, \ldots, 10^{-1}) \in \mathbb{R}^{n\times n}$. مقدار $\mathrm{cond}_\infty(D)$ را محاسبه کنید. روشن است که $\det(D) = 10^{-n}$. با توجه به دو حالت اشاره شده، در حالت کلی در خصوص رابطهٔ بین det و cond چه چیزی می توان گفت؟

فصل ۴

مسائل مقدار ویژه

در این فصل مباحثی نظری و پیش‌زمینه‌هایی از مسائل مقدار ویژه را مطالعه کرده و در ادامه، چند روش عددی برای محاسبهٔ یک یا چند مقدار ویژهٔ یک ماتریس و بردارهای ویژهٔ متناظر را بررسی می‌کنیم.

۱.۴ مباحث نظری

همان‌طور که در فصل ۱ تعریف کردیم بردار ناصفر x را یک بردار ویژهٔ ماتریس $A \in \mathbb{C}^{n\times n}$ می‌نامیم هرگاه اسکالری مثل λ وجود داشته باشد به طوری که $Ax = \lambda x$. در اینجا x را بردار ویژهٔ A، λ را یک مقدار ویژهٔ A متناظر با x و زوج مرتب (λ, x) را یک زوج ویژهٔ A می‌گوییم. مجموعهٔ مقادیر ویژهٔ یک ماتریس را با $\sigma(A)$ نشان می‌دهند و طیف A می‌نامند. واضح است که اگر x یک بردار ویژهٔ A متناظر با مقدار ویژهٔ λ باشد، آنگاه به ازای هر $c \in \mathbb{C} \neq ۰$، بردار cx یک بردار ویژهٔ A متناظر با λ است. ماتریس‌های A و B را متشابه گویند، هرگاه ماتریس نامنفردی مثل S وجود داشته باشد به طوری که $A = S^{-۱}BS$، و در این صورت می‌نویسیم $A \sim B$. به سادگی می‌توان دید که اگر (λ, x) یک زوج ویژهٔ B باشد و $A = S^{-۱}BS$ (یعنی $A \sim B$)، آنگاه $(\lambda, S^{-۱}x)$ یک زوج ویژهٔ A است.

اگر (λ, x) یک زوج ویژهٔ ماتریس A باشد، آنگاه از رابطهٔ $Ax = \lambda x$، خواهیم داشت $(\lambda I - A)x = ۰$. از اینکه بردار ویژهٔ x ناصفر است، باید داشته باشیم $\det(\lambda I - A) = ۰$. تعریف می‌کنیم:

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

می‌توان دید که $p(\lambda)$ یک چند جمله‌ای درجهٔ n است که ریشه‌های آن مقادیر ویژهٔ A هستند. $p(\lambda)$ را چندجمله‌ای مشخصهٔ A می‌گویند.

مثال ۱.۴ زوج‌های ویژهٔ ماتریس

$$A = \begin{pmatrix} ۳ & ۰ & ۱ \\ ۰ & -۳ & ۰ \\ ۱ & ۰ & ۳ \end{pmatrix},$$

را به دست آورید.

حل: با صفر قرار دادن چندجمله‌ای مشخصهٔ ماتریس A داریم:

$$p(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - ۳ & ۰ & -۱ \\ ۰ & \lambda + ۳ & ۰ \\ -۱ & ۰ & \lambda - ۳ \end{vmatrix} = (\lambda + ۳)(\lambda - ۲)(\lambda - ۴) = ۰.$$

بنابراین مقادیر ویژهٔ A عبارتند از $\lambda_۱ = ۴$، $\lambda_۲ = -۳$ و $\lambda_۳ = ۲$. برای محاسبهٔ بردار ویژهٔ متناظر با $\lambda_۱$، یعنی $x_۱ = (\alpha_۱, \alpha_۲, \alpha_۳)^T$، داریم:

$$\begin{pmatrix} ۱ & ۰ & -۱ \\ ۰ & ۷ & ۰ \\ -۱ & ۰ & ۱ \end{pmatrix} \begin{pmatrix} \alpha_۱ \\ \alpha_۲ \\ \alpha_۳ \end{pmatrix} = \begin{pmatrix} ۰ \\ ۰ \\ ۰ \end{pmatrix}.$$

با حل این دستگاه خواهیم داشت $\alpha_۲ = ۰$ و $\alpha_۳ = \alpha_۱$. بنابراین جواب این دستگاه به صورت

$$x_۱ = \begin{pmatrix} \alpha_۱ \\ ۰ \\ \alpha_۱ \end{pmatrix} = \alpha_۱ \begin{pmatrix} ۱ \\ ۰ \\ ۱ \end{pmatrix},$$

است و در نتیجه $v_1 = (1, 0, 1)^T$ یک بردار ویژهٔ ماتریس A می‌باشد. به همین ترتیب می‌توان دید که بردارهای ویژهٔ متناظر با مقادیر ویژهٔ λ_2 و λ_3 به ترتیب عبارتند از

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

مثال ۲.۴ زوج‌های ویژهٔ ماتریس

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

را به دست آورید.

حل: داریم:

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

ریشه‌های p (مقادیر ویژهٔ A) عبارتند از $\lambda_1 = \lambda_2 = 0$ و $\lambda_3 = 3$. برای محاسبهٔ بردار ویژهٔ متناظر با λ_1، یعنی $x_1 = (\alpha_1, \alpha_2, \alpha_3)^T$، داریم:

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

این دستگاه معادل با $\alpha_1 + \alpha_2 + \alpha_3 = 0$ است. در نتیجه داریم $\alpha_1 = -\alpha_2 - \alpha_3$. لذا

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} -\alpha_2 - \alpha_3 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} -\alpha_2 \\ \alpha_2 \\ 0 \end{pmatrix} + \begin{pmatrix} -\alpha_3 \\ 0 \\ \alpha_3 \end{pmatrix}$$
$$= \alpha_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

از این‌رو هر بردار در زیرفضای $\mathrm{span}\{v_۱, v_۲\}$ یک بردار ویژهٔ متناظر با $\lambda_۱ = \lambda_۲ = ۰$ است، که در آن $v_۱ = (-۱, ۰, ۱)^T$ و $v_۲ = (-۱, ۱, ۰)^T$. بردارهای $v_۱$ و $v_۲$ یک پایه برای این زیر فضا تشکیل می‌دهند. به سادگی می‌توان دید که $v_۳ = (۱, ۱, ۱)^T$ یک بردار ویژهٔ متناظر با مقدار ویژهٔ $\lambda_۳ = ۳$ است.

قضیه ۱.۴ فرض کنید که A یک ماتریس $n \times n$ باشد.
(الف) اگر λ_i، $i = ۱, \ldots, n$ مقادیر ویژهٔ A باشند، آنگاه

$$\mathrm{trace}(A) = \sum_{i=۱}^{n} \lambda_i, \qquad \det(A) = \prod_{i=۱}^{n} \lambda_i,$$

(ب) اگر $p(x) = a_۰ + a_۱ x + \cdots + a_m x^m$ و (λ_i, x_i) یک زوج ویژهٔ A باشد، آنگاه $(p(\lambda_i), x_i)$ یک زوج ویژهٔ $p(A)$ است که $p(A) = a_۰ I + a_۱ A + \cdots + a_m A^m$.
(ج) اگر A یک ماتریس بالامثلثی یا پایین‌مثلثی باشد، آنگاه مقادیر ویژهٔ آن درایه‌های قطری آن می‌باشند.
(د) اگر A یک ماتریس بالامثلثی بلوکی به صورت

$$M = \begin{pmatrix} A_{۱۱} & A_{۱۲} & \cdots & A_{۱m} \\ & A_{۲۲} & \cdots & A_{۲m} \\ & & \ddots & \vdots \\ & & & A_{mm} \end{pmatrix},$$

باشد، آنگاه مجموعهٔ مقادیر ویژهٔ M برابر با مجموعه مقادیر ویژهٔ $A_{۱۱}$، $\cdots$، A_{mm} است.

برهان: (الف) با بسط $p(\lambda) = \det(\lambda I - A)$ بر حسب سطر اول می‌توان دید:

$$\begin{aligned} p(\lambda) &= (\lambda - \lambda_۱)(\lambda - \lambda_۲) \cdots (\lambda - \lambda_n) \\ &= (\lambda - a_{۱۱})(\lambda - a_{۲۲}) \cdots (\lambda - a_{nn}) + q(\lambda), \end{aligned}$$

که در آن q یک چندجمله‌ای برحسب λ و حداکثر از درجهٔ $n - ۲$ است. با برابر قرار دادن ضریب $\lambda^{n-۱}$ در دو رابطهٔ اخیر، رابطهٔ اول اثبات می‌شود. برای اثبات رابطهٔ دوم، داریم:

$$p(\lambda) = \det(\lambda I - A) = (\lambda - \lambda_۱)(\lambda - \lambda_۲) \cdots (\lambda - \lambda_n).$$

با قرار دادن $\lambda = ۰$ در این رابطه، خواهیم داشت:

$$\det(-A) = \prod_{i=۱}^{n}(-\lambda_i),$$

یا

$$(-۱)^n \det(A) = (-۱)^n \prod_{i=۱}^{n} \lambda_i,$$

که رابطهٔ دوم را نتیجه می‌دهد.

(ب) به ازای هر عدد طبیعی k، داریم:

$$A^k x_i = A^{k-۱}(Ax_i) = A^{k-۱}(\lambda_i x_i) = \lambda_i A^{k-۱} x_i = \cdots = \lambda_i^k x_i.$$

این رابطه نشان می‌دهد که (λ_i^k, x_i) یک زوج ویژهٔ A^k است. از این‌رو

$$\begin{aligned} p(A)x_i &= (a_۰ I + a_۱ A + \cdots + a_m A^m)x_i = a_۰ x_i + a_۱ Ax_i + \cdots + a_m A^m x_i \\ &= a_۰ x_i + a_۱ \lambda_i x_i + \cdots + a_m \lambda_i^m x_i = (a_۰ + a_۱ \lambda_i + \cdots + a_m \lambda_i^m)x_i \\ &= p(\lambda_i)x_i, \end{aligned}$$

که نشان می‌دهد $(p(\lambda_i), x_i)$ یک زوج ویژهٔ $p(A)$ است.

(ج) فرض کنید ماتریس A یک ماتریس بالامثلثی باشد. در این صورت، داریم:

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{۱۱} & -a_{۱۲} & \cdots & -a_{۱n} \\ & \lambda - a_{۲۲} & \cdots & -a_{۲n} \\ & & \ddots & \vdots \\ & & & \lambda - a_{nn} \end{vmatrix} \\ &= (\lambda - a_{۱۱})(\lambda - a_{۲۲}) \cdots (\lambda - a_{nn}). \end{aligned}$$

بنابراین ریشه‌های p عبارتند از a_{ii}، $i = ۱, \ldots, n$. اثبات برای حالت پایین‌مثلثی به طور مشابه انجام می‌پذیرد.

(د) داریم:

$$p(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda I_۱ - A_{۱۱} & -A_{۱۲} & \cdots & -A_{۱m} \\ & \lambda I_۲ - A_{۲۲} & \cdots & -A_{۲m} \\ & & \ddots & \vdots \\ & & & \lambda I_m - A_{mm} \end{vmatrix}$$

$$= \det(\lambda I_۱ - A_{۱۱}) \det(\lambda I_۲ - A_{۲۲}) \cdots \det(\lambda I_m - A_{mm}),$$

که درآن به ازای $i = ۱, \ldots, m$، ماتریس I_i یک ماتریس همانی و هم‌بعد با A_{ii} است. از این‌رو $p(\lambda) = ۰$ اگر و تنها اگر حداقل به ازای یک i، $\det(\lambda I_i - A_{ii}) = ۰$. بدین ترتیب اثبات کامل می‌شود. □

مسأله ۱.۴ فرض کنید $\alpha = \lambda_۱$، $\lambda_۲ = ۲i$ و $\lambda_۳ = -۲i$ مقادیر ویژهٔ یک ماتریس متقارن‌کج $A = (a_{ij}) \in \mathbb{R}^{۳\times۳}$ باشند.

(الف) مقادیر $\mathrm{trace}(A)$، α و $\det(A)$ را محاسبه کنید.

(ب) چرا $I + A$ نامنفرد است؟ مقدار $\det(I + A)^{-۱}$ را محاسبه کنید.

حل: (الف) از اینکه $A^T = -A$، داریم $a_{ii} = -a_{ii}$. درنتیجه $a_{ii} = ۰$، $i = ۱, ۲, ۳$. بنابراین $\mathrm{trace}(A) = a_{۱۱} + a_{۲۲} + a_{۳۳} = ۰$. از طرفی داریم:

$$۰ = \mathrm{trace}(A) = \lambda_۱ + \lambda_۲ + \lambda_۳ = \alpha + (-۲i) + ۲i = \alpha.$$

همچنین داریم:

$$\det(A) = \lambda_۱\lambda_۲\lambda_۳ = (-۲i) \times (۲i) \times ۰ = ۰.$$

(ب) بنا به قسمت دوم قضیهٔ ۱.۴، اگر λ یک مقدار ویژهٔ A باشد، آنگاه $\mu = ۱ + \lambda$ یک مقدار ویژهٔ $I + A$ است. لذا مقادیر ویژهٔ ماتریس $I + A$ عبارتند از

$$\mu_۲ = ۱ - ۲i, \quad \mu_۱ = ۱ + ۰ = ۱, \quad \mu_۳ = ۱ + ۲i.$$

در نتیجه

$$\det(I + A) = \mu_۱\mu_۲\mu_۳ = ۱ \times (۱ - ۲i) \times (۱ + ۲i) = ۵ \neq ۰,$$

که نشان می‌دهد ماتریس $I + A$ نامنفرد است. سرانجام فرض کنید (λ, x) یک زوج ویژهٔ A باشد. در این صورت

$$(I + A)x = (۱ + \lambda)x \Longrightarrow (I + A)^{-۱}x = \frac{۱}{۱ + \lambda}x.$$

این رابطه نشان می‌دهد که $(x, 1/(1+\lambda))$ یک زوج ویژهٔ ماتریس $(I+A)^{-1}$ است. لذا

$$\det(I+A)^{-1} = \frac{1}{1+\lambda_1}\frac{1}{1+\lambda_2}\frac{1}{1+\lambda_3} = \frac{1}{1+0}\frac{1}{1-2i}\frac{1}{1+2i} = \frac{1}{5},$$

که حل مسأله را کامل می‌کند.

قضیه ۲.۴ فرض کنید $p(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n$ و A یک ماتریس $n \times n$ به صورت

$$A = \begin{pmatrix} -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_0 \\ 1 & 0 & 0 & \cdots & 0 \\ & 1 & 0 & \cdots & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 1 & 0 \end{pmatrix},$$

باشد. در این صورت، p چندجمله‌ای مشخصهٔ A است.

برهان: داریم:

$$p(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda + a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \\ -1 & \lambda & 0 & \cdots & 0 \\ & -1 & \lambda & \cdots & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & -1 & \lambda \end{vmatrix}.$$

با بسط دترمینان برحسب سطر اول خواهیم داشت (رابطه (۴.۱) را به خاطر بیاورید):

$$\begin{aligned} p(\lambda) &= (\lambda + a_{n-1})\lambda^{n-1} - a_{n-2}(-\lambda^{n-2}) + a_{n-3}(-1)^2\lambda^{n-3} + \\ &\quad \cdots + (-1)^{(1+n)}a_0(-1)^{n-1} \\ &= \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + a_{n-3}\lambda^{n-3} + \cdots + a_0, \end{aligned}$$

که اثبات قضیه را کامل می‌کند. □

قضیهٔ ۲.۴ نشان می‌دهد که مسألهٔ محاسبهٔ مقادیر ویژهٔ یک ماتریس معادل با یافتن ریشه‌های یک چندجمله‌ای است.

قضیه ۳.۴ گر $A \in \mathbb{C}^{n\times n}$، $u, v \in \mathbb{C}^n$ و $\langle .,. \rangle$ ضرب داخلی استاندارد در $\mathbb{C}^n$ باشد، آنگاه

$$\langle u, Av\rangle = \langle A^H u, v\rangle.$$

برهان: با توجه به تعریف ضرب داخلی استاندارد می‌توان نوشت:

$$\langle u, Av\rangle = u^H Av = (A^H u)^H v = \langle A^H u, v\rangle,$$

که بدین ترتیب اثبات قضیه کامل می‌گردد. □

قضیه ۴.۴ مقادیر ویژهٔ ماتریس‌های هرمیتی، حقیقی‌اند.

برهان: فرض کنید که ماتریس A هرمیتی و $\langle \lambda, x\rangle$ یک زوج ویژهٔ آن باشد. دراین صورت،

$$\lambda\langle x, x\rangle = \langle x, \lambda x\rangle = \langle x, Ax\rangle = \langle A^H x, x\rangle = \langle Ax, x\rangle = \langle \lambda x, x\rangle = \bar{\lambda}\langle x, x\rangle.$$

بنابراین $\bar{\lambda} = \lambda$. این رابطه نشان می‌دهد که λ حقیقی است. □

نتیجه ۱.۴ مقادیر ویژهٔ یک ماتریس متقارن حقیقی‌اند.

مثال ۳.۴ ماتریس مختلط

$$A = \begin{pmatrix} ۲ & ۷ - ۷i & -۷ + ۷i \\ ۷ + ۷i & ۵ & ۳ \\ -۷ - ۷i & ۳ & ۵ \end{pmatrix}$$

هرمیتی است و مجموعهٔ مقادیر ویژهٔ این ماتریس برابر با $\{-۱۲, ۸, ۱۶\}$ می‌باشد.

قضیه ۵.۴ بردارهای ویژهٔ یک ماتریس هرمیتی متناظر به مقادیر ویژهٔ متمایز، متعامدند.

برهان: فرض کنید که (λ_i, x_i) و (λ_j, x_j) دو زوج ویژهٔ ماتریس هرمیتی A باشند به طوری که $\lambda_i \neq \lambda_j$. داریم:

$$\begin{aligned} Ax_i = \lambda_i x_i \;\Rightarrow\; \langle Ax_i, x_j\rangle = \langle \lambda_i x_i, x_j\rangle \Rightarrow \langle x_i, A^H x_j\rangle = \overline{\lambda_i}\langle x_i, x_j\rangle \\ \Rightarrow\; \langle x_i, Ax_j\rangle = \overline{\lambda_i}\langle x_i, x_j\rangle \Rightarrow \langle x_i, \lambda_j x_j\rangle = \overline{\lambda_i}\langle x_i, x_j\rangle \\ \Rightarrow\; (\lambda_j - \overline{\lambda_i})\langle x_i, x_j\rangle = \circ. \end{aligned}$$

از اینکه A هرمیتی است، داریم $\overline{\lambda_i} = \lambda_i$ و با توجه به اینکه $\lambda_i \neq \lambda_j$ نتیجه می‌شود که $\langle x_i, x_j\rangle = \circ$. □

قضیه ۶.۴ بردارهای ویژهٔ متناظر به مقادیر ویژهٔ متمایز یک ماتریس مستقل خطی‌اند.

برهان: (اثبات به استقراء) فرض کنید $m = ۱$. در این صورت، مجموعهٔ $\{x_۱\}$ متناظر به مقدار ویژهٔ $\lambda_۱$ از ماتریس A یک مجموعهٔ مستقل خطی است. فرض کنید که حکم برای $m - ۱$ بردار ویژهٔ $x_۱$، $x_۲$، ...، $x_{m-۱}$ متناظر به مقادیر ویژهٔ دو به دو متمایز $\lambda_۱$، $\lambda_۲$، ...، $\lambda_{m-۱}$ از ماتریس A برقرار باشد. فرض کنید که (λ_m, x_m) یک زوج ویژهٔ ماتریس A باشد به طوری که $\lambda_m \neq \lambda_j$، $j = ۱, ۲, \ldots, m-۱$. به برهان خلف، اگر بردارهای x_i، $i = ۱, ۲, \ldots, m$ مستقل خطی نباشند آنگاه α_i، $i = ۱, ۲, \ldots, m-۱$ وجود دارند به طوری که

$$x_m = \sum_{i=۱}^{m-۱} \alpha_i x_i. \tag{۱.۴}$$

با ضرب طرفین این رابطه در A خواهیم داشت:

$$\lambda_m x_m = Ax_m = \sum_{i=۱}^{m-۱} \alpha_i A x_i = \sum_{i=۱}^{m-۱} \alpha_i \lambda_i x_i. \tag{۲.۴}$$

با ضرب طرفین رابطهٔ (۱.۴) در λ_m و کم کردن از طرفین رابطهٔ (۲.۴)، نتیجه می‌گیریم:

$$\circ = \sum_{i=۱}^{m-۱} \alpha_i(\lambda_i - \lambda_m)x_i.$$

حال با توجه به اینکه بردارهای x_i، $i = ۱, ۲, \ldots, m-۱$ مستقل خطی می‌باشند نتیجه می‌شود که $\alpha_i = \circ$، $i = ۱, ۲, \ldots, m-۱$. بنابراین با توجه به معادلهٔ (۱.۴) نتیجه می‌شود که $x_m = \circ$ و این یک تناقض است. □

قضیه ۷.۴ فرض کنید که λ_j، $j = ۱,۲,\ldots,n$، n مقدار ویژهٔ متناظر با n بردار ویژهٔ مستقل خطی x_j، $j = ۱,۲,\ldots,n$ برای ماتریس $A \in \mathbb{C}^{n\times n}$ باشند. اگر ماتریس C را به صورت $C = (x_۱, x_۲, \ldots, x_n)$ تعریف کنیم، آنگاه $C^{-۱}AC = D$، که در آن $D = \mathrm{diag}(\lambda_۱, \lambda_۲, \ldots, \lambda_n)$. برعکس اگر ماتریس نامنفردی مثل C وجود داشته باشد به طوری $C^{-۱}AC = \mathrm{diag}(\lambda_۱, \lambda_۲, \ldots, \lambda_n)$. آنگاه ستون‌های C بردارهای ویژهٔ متناظر با مقادیر ویژهٔ $\lambda_۱$، $\lambda_۲$، ...، λ_n از ماتریس A هستند.

برهان: برای اثبات قسمت اول، داریم:

$$\begin{aligned} AC &= A(x_۱, x_۲, \ldots, x_n) = (Ax_۱, Ax_۲, \ldots, Ax_n) = (\lambda_۱ x_۱, \lambda_۲ x_۲, \ldots, \lambda_n x_n) \\ &= (x_۱, x_۲, \ldots, x_n) \begin{pmatrix} \lambda_۱ & & & \\ & \lambda_۲ & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = CD. \end{aligned}$$

از اینکه ستون‌های C مستقل خطی هستند، نتیجه می‌شود که نامنفرد است، لذا نتیجهٔ لازم به دست می‌آید. قسمت دوم قضیه نیز به طور مشابه ثابت می‌شود. □

تعریف ۱.۴ ماتریس A را قطری‌شدنی می‌گویند هرگاه یک ماتریس نامنفرد مثل C وجود داشته باشد به طوری که $C^{-۱}AC = D$ که درآن D یک ماتریس قطری است، به عبارت دیگر با یک ماتریس قطری متشابه باشد. ماتریس A را به طوری یکانی قطری‌شدنی می‌گویند هرگاه ماتریس C یکانی باشد.

بنا به قضیهٔ ۷.۴، ماتریس $A \in \mathbb{C}^{n\times n}$ قطری‌شدنی است اگر و تنها اگر n بردار ویژهٔ مستقل خطی داشته باشد. قضیهٔ بعد می‌تواند در تشخیص اینکه آیا یک ماتریس قطری‌شدنی است یا نه، مفید باشد.

قضیه ۸.۴ فرض کنید $A \in \mathbb{C}^{n\times n}$.

(الف) اگر مقادیر ویژهٔ A متمایز باشند، آنگاه ماتریس A قطری‌شدنی است.

(ب) ماتریس A به طور یکانی قطری‌شدنی است اگر و تنها اگر نرمال باشد.

برهان: (الف) با توجه به قضیهٔ ۶.۴، اگر مقادیر ویژهٔ یک ماتریس متمایز باشند، آنگاه بردارهای ویژهٔ آن مستقل خطی و در نتیجه با توجه به قضیهٔ ۷.۴ ماتریس قطری‌شدنی است.
برای اثبات (ب) به [۱۱] مراجعه کنید. □

نتیجه ۲.۴ هر ماتریس هرمیتی، نرمال است، لذا با توجه به قضیهٔ ۸.۴ به طور یکانی قطری‌شدنی است.

به عنوان یک نتیجه از قضیهٔ ۸.۴ اگر ماتریس $A \in \mathbb{R}^{n\times n}$ متقارن، یا مقادیر ویژهٔ آن متمایز باشند، آنگاه قطری‌شدنی است. به‌علاوه ماتریس حقیقی A نرمال است، یعنی $A^TA = AA^T$، اگر و تنها اگر به طور متعامد قطری‌شدنی باشد.

مثال ۴.۴ نشان دهید که ماتریس‌های

$$A = \begin{pmatrix} ۲ & ۴ \\ -۴ & ۲ \end{pmatrix}, \quad B = \begin{pmatrix} ۴ & ۱ \\ ۶ & ۳ \end{pmatrix},$$

قطری‌شدنی هستند.

حل: داریم:

$$A^TA = AA^T = \begin{pmatrix} ۲۰ & ۰ \\ ۰ & ۲۰ \end{pmatrix}.$$

لذا ماتریس A یک ماتریس نرمال است و بنا به قسمت (ب) قضیهٔ ۸.۴، قطری‌شدنی است. توجه کنید که مقادیر ویژهٔ ماتریس A عبارتند از $\{۲ - ۴i, ۲ + ۴i\}$ و بنا به قسمت (الف) قضیهٔ ۸.۴ قطری‌شدنی است. در واقع داریم $A = CDC^{-۱}$، که در آن

$$C = \begin{pmatrix} \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \\ \frac{\sqrt{۲}}{۲}i & -\frac{\sqrt{۲}}{۲}i \end{pmatrix}, \quad D = \begin{pmatrix} ۲ + ۴i & ۰ \\ ۰ & ۲ - ۴i \end{pmatrix}.$$

ماتریس B نرمال نیست، زیرا داریم:

$$B^TB = \begin{pmatrix} ۵۲ & ۲۲ \\ ۲۲ & ۱۰ \end{pmatrix} \neq BB^T = \begin{pmatrix} ۱۷ & ۲۷ \\ ۲۷ & ۴۵ \end{pmatrix}.$$

اما مجموعهٔ مقادیر ویژهٔ ماتریس B به صورت $\{۱, ٦\}$ است و در نتیجه بنا به قسمت (الف) قضیهٔ ۸.۴ قطری‌شدنی است و داریم $B = CDC^{-۱}$ که در آن

$$C = \begin{pmatrix} \frac{۱}{\sqrt{۵}} & \frac{-۱}{\sqrt{۱۰}} \\ \frac{۲}{\sqrt{۵}} & \frac{۳}{\sqrt{۱۰}} \end{pmatrix}, \quad D = \begin{pmatrix} ٦ & ۰ \\ ۰ & ۱ \end{pmatrix}.$$

در اینجا توجه کنید اگرچه ماتریس A قطری‌شدنی است اما ماتریس C ای که A را قطری می‌کند، متعامد نیست و بنابراین نمی‌توان با استناد به قسمت (ب) قضیهٔ ۸.۴ نتیجه گرفت که ماتریس B نرمال است.

مثال ۵.۴ دستگاه معادلات دیفرانسیل

$$\begin{cases} \begin{pmatrix} x'_۱(t) \\ x'_۲(t) \end{pmatrix} = \begin{pmatrix} ۴ & ۱ \\ ۳ & ۲ \end{pmatrix} \begin{pmatrix} x_۱(t) \\ x_۲(t) \end{pmatrix}, \\ \begin{pmatrix} x_۱(۰) \\ x_۲(۰) \end{pmatrix} = \begin{pmatrix} ٦ \\ -۲ \end{pmatrix}, \end{cases} \tag{۳.۴}$$

را حل کنید.

حل: با فرض

$$X(t) = \begin{pmatrix} x_۱(t) \\ x_۲(t) \end{pmatrix}, \quad A = \begin{pmatrix} ۴ & ۱ \\ ۳ & ۲ \end{pmatrix}, \quad X_۰ = \begin{pmatrix} ٦ \\ -۲ \end{pmatrix},$$

دستگاه (۳.۴) را می‌توان به صورت

$$\begin{cases} X'(t) = AX(t), \\ X(۰) = X_۰, \end{cases} \tag{۴.۴}$$

نوشت. مقادیر ویژهٔ ماتریس A عبارتند از $\lambda_۱ = ۱$ و $\lambda_۲ = ۵$ و بردارهای ویژهٔ متناظر با آنها به صورت

$$x_۱ = \begin{pmatrix} ۱ \\ -۳ \end{pmatrix}, \quad x_۲ = \begin{pmatrix} ۱ \\ ۱ \end{pmatrix},$$

می‌باشند. بردارهای ویژهٔ A مستقل خطی هستند، زیرا مقادیر ویژهٔ A متمایز هستند. بنابراین اگر قرار دهیم $C = (x_۱, x_۲)$، آنگاه بنا به قضیهٔ ۷.۴، داریم:

$$C^{-۱}AC = D = \begin{pmatrix} ۱ & ۰ \\ ۰ & ۵ \end{pmatrix}.$$

حال معادلهٔ اول در (۴.۴) را به صورت

$$C^{-1}X'(t) = C^{-1}ACC^{-1}X(t), \qquad (۵.۴)$$

می‌نویسیم. با قرار دادن $Z(t) = C^{-1}X(t)$، معادلهٔ (۵.۴) به صورت $Z'(t) = DZ(t)$ نوشته می‌شود که معادل با

$$\begin{cases} z_1'(t) = z_1(t), \\ z_2'(t) = 5z_2(t), \end{cases} \qquad (۶.۴)$$

است، که در آن $Z(t) = (z_1(t), z_2(t))^T$. با حل معادلهٔ (۶.۴)، داریم $z_1(t) = c_1e^t$ و $z_2(t) = c_2e^{5t}$ که در آن c_1 و c_2 ثابت‌های دلخواه هستند. در نتیجه

$$X(t) = CZ(t) = \begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} c_1e^t \\ c_2e^{5t} \end{pmatrix} = \begin{pmatrix} c_1e^t + c_2e^{5t} \\ -3c_1e^t + c_2e^{5t} \end{pmatrix}. \qquad (۷.۴)$$

حال با استفاده از شرط اولیهٔ داده شده، داریم:

$$X(0) = \begin{pmatrix} c_1 + c_2 \\ -3c_1 + c_2 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \end{pmatrix}.$$

با حل این دستگاه داریم $c_1 = 2$ و $c_2 = 4$ و با جایگذاری در رابطهٔ (۷.۴) خواهیم داشت:

$$\begin{cases} x_1(t) = \ \ 2e^t + 4e^{5t}, \\ x_2(t) = -6e^t + 4e^{5t}. \end{cases}$$

قضیه ۹.۴ اگر مقادیر ویژهٔ ماتریس هرمیتی A به صورت $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ مرتب شده باشند، آنگاه

$$\lambda_1 = \max_{0 \neq x \in \mathbb{C}^n} \frac{x^HAx}{x^Hx}, \qquad \lambda_n = \min_{0 \neq x \in \mathbb{C}^n} \frac{x^HAx}{x^Hx}.$$

برهان: با استفاده از نتیجهٔ ۲.۴، یک ماتریس یکانی مثل U وجود دارد به طوری که $U^HAU = D$ که در آن $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. در نتیجه $A = UDU^H$. از این‌رو اگر $x \neq 0$، آنگاه

$$\frac{x^HAx}{x^Hx} = \frac{x^HUTU^Hx}{x^HUU^Hx} = \frac{(U^Hx)^HT(U^Hx)}{(U^Hx)^H(U^Hx)} = \frac{y^HTy}{y^Hy},$$

که در آن $y = U^H x = [\eta_1, \dots, \eta_n]^T \neq 0$. لذا

$$\frac{x^H A x}{x^H x} = \frac{\sum_{i=1}^{n} \lambda_i |\eta_i|^2}{\sum_{i=1}^{n} |\eta_i|^2} \leqslant \frac{\sum_{i=1}^{n} \lambda_1 |\eta_i|^2}{\sum_{i=1}^{n} |\eta_i|^2} = \lambda_1. \qquad (۸.۴)$$

از طرفی اگر x_1 یک بردار ویژهٔ ماتریس A متناظر با مقدار ویژهٔ λ_1 باشد، آنگاه

$$\frac{x_1^H A x_1}{x_1^H x_1} = \frac{x_1^H \lambda_1 x_1}{x_1^H x_1} = \lambda_1. \qquad (۹.۴)$$

بنابراین از روابط (۸.۴) و (۹.۴) رابطهٔ اول در قضیه به دست می‌آید. اثبات رابطهٔ دیگر به طور مشابه انجام می‌گیرد. □

قضیه ۱۰.۴ فرض کنید که A یک ماتریس هرمیتی باشد. A معین مثبت است اگر و تنها اگر مقادیر ویژهٔ آن مثبت باشند.

برهان: فرض کنید که A معین مثبت هرمیتی و $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ مقادیر ویژهٔ آن باشند. همچنین فرض کنید (λ_n, x_n) یک زوج ویژهٔ A باشد. از این‌رو با توجه به قضیهٔ قبل داریم:

$$\lambda_n = \min_{0 \neq x \in \mathbb{C}^n} \frac{x^H A x}{x^H x} = \frac{x_n^H A x_n}{x_n^H x_n} > 0.$$

بنابراین $\lambda_i > 0$، $i = 1, 2, \dots, n$. برعکس فرض کنید که $y \neq 0$. در این صورت،

$$0 < \lambda_n = \min_{0 \neq x \in \mathbb{C}^n} \frac{x^H A x}{x^H x} \leqslant \frac{y^H A y}{y^H y},$$

و لذا $y^H A y > 0$، که اثبات قضیه را کامل می‌کند. □

نتیجه ۳.۴ فرض کنید که A یک ماتریس هرمیتی باشد. A نیمه معین مثبت هرمیتی است اگر و تنها اگر مقادیر ویژهٔ آن نامنفی باشند.

حال آماده‌ایم که قسمت دوم از قضیهٔ ۳.۲ را ثابت کنیم. ابتدا توجه کنید که ماتریس $A^H A$ نیمه معین مثبت متقارن و با توجه به قضیهٔ ۱۰.۴ مقادیر ویژهٔ آن نامنفی‌اند. از

این‌رو با استفاده از قضیهٔ ۹.۴ داریم:

$$\rho(A^HA) = \max_{\lambda_i\in\sigma(A^HA)}\lambda_i = \max_{x\neq 0}\frac{x^HA^HAx}{x^Hx} = \max_{x\neq 0}\frac{(Ax)^H(Ax)}{x^Hx}$$
$$= \max_{x\neq 0}\frac{\|Ax\|_2^2}{\|x\|_2^2} = \left(\max_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2}\right)^2 = \|A\|_2^2.$$

بنابراین $\|A\|_2 = \sqrt{\rho(A^HA)}$.

قضیه ۱۱.۴ (قضیهٔ دیسک‌های گرشگورین). فرض کنید $A\in\mathbb{C}^{n\times n}$.

(الف) مقادیر ویژهٔ ماتریس A در اجتماع n دیسک گرشگورین زیر قرار دارند:

$$C_r : |z-a_{ii}| \leqslant r_i, \qquad r_i = \sum_{j=1, j\neq i}^{n}|a_{ij}|, \quad i = 1,2,\dots,n.$$

(ب) مقادیر ویژهٔ ماتریس A در مجموعه دیسک‌های زیر قرار دارند:

$$C_c : |z-a_{jj}| \leqslant c_j, \qquad c_i = \sum_{i=1, i\neq j}^{n}|a_{ij}|, \quad j = 1,2,\dots,n.$$

(ج) مقادیر ویژهٔ A در $C_r\bigcap C_c$ قرار دارند.

برهان: (الف) فرض کنید که (λ, x) یک زوج ویژهٔ ماتریس A باشد. بدون از دست دادن کلیت فرض می‌کنیم که $\|x\|_\infty = |x_i| = 1$ که درآن x_i مؤلفهٔ i ام x است. در این صورت، با در نظر گرفتن سطر i ام از $Ax=\lambda x$، داریم:

$$\lambda x_i = \sum_{j=1}^{n}a_{ij}x_j \Longrightarrow (\lambda - a_{ii})x_i = \sum_{j=1, j\neq i}^{n}a_{ij}x_j.$$

بنابراین

$$|\lambda - a_{ii}| = |\lambda - a_{ii}||x_i| = |\sum_{j=1, j\neq i}^{n}a_{ij}x_j| \leqslant \sum_{j=1, j\neq i}^{n}|a_{ij}||x_j|$$
$$\leqslant \sum_{j=1, j\neq i}^{n}|a_{ij}| = r_i.$$

در نتیجه λ در یکی از دیسک‌های گرشگورین قرار دارد. بنابراین اجتماع این دیسک‌ها شامل مقادیر ویژهٔ A هستند.

(ب) از اینکه مجموعهٔ مقادیر ویژهٔ A و A^T یکسانند، نتیجه لازم به دست می‌آید.

(ج) این یک نتیجه از (الف) و (ب) است. □

مثال ۶.۴ حدودی برای مقادیر ویژهٔ ماتریس

$$A = \begin{pmatrix} ۵ & ۱ & ۱ \\ ۰ & ۶ & ۱ \\ ۱ & ۰ & -۵ \end{pmatrix},$$

به دست آورید.

حل: با استفاده از قضیهٔ ۴.۲، داریم:

$$|\lambda| \leqslant \|A\|_\infty = \max\{۷, ۷, ۶\} = ۷.$$

برای به دست آوردن یک تقریب بهتر، از قضیهٔ دیسک‌های گرشگورین استفاده می‌کنیم. مقادیر ویژهٔ ماتریس A در دیسک‌هایی که از سطرهای ماتریس به دست می‌آیند، قرار دارند:

$$|z - ۵\,| \leqslant ۲, \quad |z - ۶\,| \leqslant ۱, \quad |z + ۵\,| \leqslant ۱.$$

همچنین مقادیر ویژه در اشتراک دیسک‌های مربوط به سطر و ستون‌های ماتریس A قرار دارند. بنابراین یک تقریب بهتر به صورت زیر به دست می‌آید:

$$|z - ۵\,| \leqslant ۱, \quad |z - ۶\,| \leqslant ۱, \quad |z + ۵\,| \leqslant ۱.$$

مثال ۷.۴ ماتریس

$$A = \begin{pmatrix} -۶ & ۲ & -۰.۳ & ۱+i \\ ۱ & ۱-i & ۲ & ۱ \\ -i & ۱ & -۱+۳i & ۳ \\ ۱ & ۰ & ۳ & ۱۰ \end{pmatrix},$$

را در نظر بگیرید. مقادیر ویژهٔ این ماتریس عبارتند از

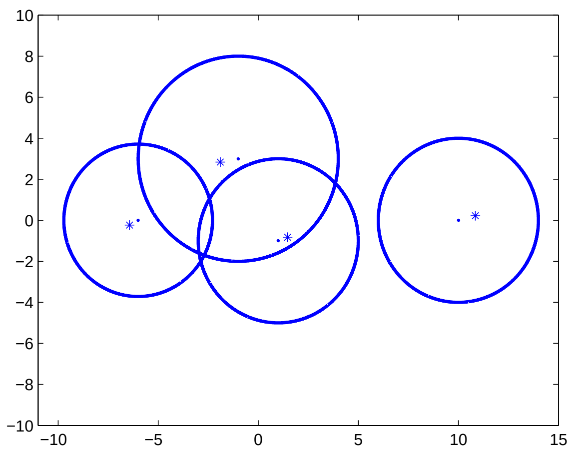
$$\lambda_۱ = ۱۰.۸۵۱۲ + ۰.۲۱۹۵i, \qquad \lambda_۲ = -۶.۴۲۷۴ - ۰.۲۳۲۵i,$$

$$\lambda_۳ = ۱.۴۷۰۵ - ۰.۸۲۱۳i, \qquad \lambda_۴ = -۱.۸۹۴۲ + ۲.۸۳۴۳i.$$

دیسک‌های گرشگورین ماتریس A به صورت

$$\begin{aligned} |z+6| &\leqslant 2.3+\sqrt{2}, & |z-(1-i)| &\leqslant 4,\\ |z-(-1+3i)| &\leqslant 5, & |z-10| &\leqslant 4. \end{aligned}$$

می‌باشند. این دیسک‌ها به همراه مقادیر ویژهٔ A در شکل ۱.۴ نمایش داده شده‌اند. نقاطی که با "*" مشخص شده‌اند، مقادیر ویژهٔ A هستند. همان‌طور که ملاحظه می‌شود تمام مقادیر ویژهٔ A در دیسک‌های گرشگورین ماتریس قرار دارند.



شکل ۱.۴: نمایش دیسک‌های گرشگورین و مقادیر ویژه برای مثال ۷.۴.

قضیه ۱۲.۴ نشان دهید اگر ماتریس $A=(a_{ij})\in\mathbb{C}^{n\times n}$ هرمیتی و غالب قطری اکید با درایه‌های قطری مثبت باشد، آنگاه معین مثبت هرمیتی است.

برهان: از اینکه ماتریس A هرمیتی است، بنا به قضیهٔ ۴.۴ مقادیر ویژهٔ A حقیقی‌اند. از اینکه درایه‌های قطر اصلی A مثبت می‌باشند، مرکز دایره‌های گرشگورین ماتریس A، روی قسمت مثبت محور x ها قرار دارد. از اینکه ماتریس A غالب قطری اکید است، داریم:

$$|a_{ii}|<\sum_{j=1,j\neq i}^{n}|a_{ij}|,\qquad i=1,2,\ldots,n.$$

این رابطه نشان می‌دهد که شعاع دایره‌های گرشگورین ماتریس A از فاصلهٔ بین مبدأ مختصات تا مرکز دایره کمتر است. بنابراین دایره‌های گرشگورین ماتریس کاملاً در ربع

اول و چهارم محورهای مختصات قرار دارند.لذا تمام مقادیر ویژهٔ آن مثبت می‌باشند و با توجه به قضیهٔ ۱۰.۴، ماتریس A معین مثبت هرمیتی است.

مثال ۸.۴ ماتریس

$$A = \begin{pmatrix} ۶ & ۱+۲i & ۱-i \\ ۱-۲i & ۴ & i \\ ۱+i & -i & ۸ \end{pmatrix},$$

غالب قطری اکید است زیرا

$$\begin{aligned} ۶ &> |۱+۲i| + |۱-i| = \sqrt{۵} + \sqrt{۲}, \\ ۴ &> |۱-۲i| + |i| = \sqrt{۵} + ۱, \\ ۸ &> |۱+i| + |-i| = \sqrt{۲} + ۱. \end{aligned}$$

از طرفی، A یک ماتریس هرمیتی با درایه‌های قطری مثبت است. لذا بنا به قضیهٔ ۱۲.۴ ماتریس A معین مثبت هرمیتی است.

۲.۴ روش توانی و توانی‌معکوس

۱.۲.۴ روش توانی

روش توانی برای محاسبهٔ بزرگترین مقدار ویژه از حیث قدر مطلقِ ماتریس $A \in \mathbb{R}^{n\times n}$ و بردار ویژهٔ متناظر به کار می‌رود. از اینجا به بعد این مقدار ویژه را مقدار ویژهٔ غالب می‌نامیم. فرض کنید (λ_i, x_i)، زوج‌های ویژهٔ A باشند به طوری که x_i ها مستقل خطی هستند و

$$|\lambda_۱| > |\lambda_۲| \geqslant |\lambda_۳| \geqslant \cdots \geqslant |\lambda_n|.$$

روشن است که $\lambda_۱$ حقیقی است (چرا؟). بردار دلخواه $v_۰ \in \mathbb{R}^n$ را انتخاب می‌کنیم. با توجه به اینکه x_i ها مستقل خطی هستند، یک پایه برای $\mathbb{R}^n$ تشکیل می‌دهند. بنابراین اعداد حقیقی α_i، $i = ۱, ۲, \ldots, n$ وجود دارند به طوری که

$$v_۰ = \sum_{i=۱}^{n} \alpha_i x_i.$$

به‌علاوه فرض کنید $\alpha_1 \neq 0$. به ازای هر $m \geqslant 1$، داریم:

$$\begin{aligned} A^m v_0 &= \sum_{i=1}^{n} \alpha_i A^m x_i = \sum_{i=1}^{n} \alpha_i \lambda_i^m x_i \\ &= \lambda_1^m \left(\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^m x_i \right). \end{aligned}$$

لذا از اینکه $|\lambda_i|/|\lambda_1| < 1$، $i = 2, \ldots, n$ خواهیم داشت:

$$\lim_{m\to\infty} \frac{A^m v_0}{\lambda_1^m} = \lim_{m\to\infty} \left(\alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^m x_i \right) = \alpha_1 x_1 =: x. \quad (10.4)$$

از اینکه $\alpha_1 \neq 0$، نتیجه می‌گیریم که $x \neq 0$. روشن است که (λ_1, x) یک زوج ویژهٔ A است. رابطهٔ (۱۰.۴) نشان می‌دهد که بردار $A^m v_0$ وقتی که $m \to \infty$، به سمت یک راستا از بردار x (در نتیجه به سمت یک بردار ویژهٔ متناظر با λ_1) میل می‌کند.

با توجه به رابطهٔ (۱۰.۴)، به ازای m های بزرگ داریم $A^m v_0 \approx \alpha_1 \lambda_1^m x_1$. این رابطه نشان می‌دهد که اگر $|\lambda_1| > 1$، آنگاه مؤلفه‌های $A^m v_0$ به سمت ∞ یا $-\infty$ میل خواهند کرد و در صورتی که $|\lambda_1| < 1$، آنگاه $A^m v_0 \to 0$. در هر یک از این حالت‌ها خطاهای محاسباتی غیر قابل کنترل خواهند بود. برای رفع این مشکل ابتدا توجه کنید که اگر

$$v_i = A^i v_0, \quad i = 1, 2, \ldots,$$

آنگاه خواهیم داشت:

$$\begin{aligned} \lim_{m\to\infty} \frac{\|v_m\|_\infty}{\|v_{m-1}\|_\infty} &= \lim_{m\to\infty} \frac{\|A v_{m-1}\|_\infty}{\|v_{m-1}\|_\infty} \\ &= |\lambda_1| \lim_{m\to\infty} \frac{\left\| \alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^m x_i \right\|_\infty}{\left\| \alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^{m-1} x_i \right\|_\infty} \\ &= |\lambda_1|. \end{aligned} \quad (11.4)$$

به روشنی یکه بودن یا نبودن بردار v_{m-1} تاثیری در رابطهٔ اخیر ندارد. بنابراین در هر مرحله، قبل از محاسبهٔ بردار v_m بردار v_{m-1} را یکه می‌کنیم. در این صورت، از بزرگ

شدن درایه‌های v_m یا صفر شدن این بردار جلوگیری می‌شود. با توجه به مطالب ارائه شده الگوریتم روش توانی را می‌توان به صورت زیر جمع‌بندی کرد.

الگوریتم ۱.۴: روش توانی

1. Choose a vector $v_0 \in \mathbb{R}^n$ such that $\|v_0\|_\infty = 1$
2. For $m = 0, 1, \ldots,$ until convergence, Do
3. $\quad w_m := Av_m$
4. $\quad \mu_m := \|w_m\|_\infty$
5. $\quad v_{m+1} := w_m/\mu_m$
6. EndDo

با توجه به رابطهٔ (۱۱.۴)، می‌بینیم که $\lim_{m\to\infty} \mu_m = |\lambda_1|$. بنابراین لازم است علامت λ_1 را تشخیص دهیم. برای این کار، رابطهٔ $A^m v_0 \approx \lambda_1^m x$ نشان می‌دهد که اگر $\lambda_1 < 0$، آنگاه به ازای m های بزرگ مؤلفه‌های متناظر دو بردار v_m و v_{m+1} قرینهٔ یکدیگر و اگر داشته باشیم $\lambda_1 > 0$، آنگاه هم‌علامت خواهند بود. نکتهٔ قابل ذکر دیگر این است که با توجه به رابطهٔ (۱۰.۴) سرعت همگرایی روش توانی به $|\lambda_2|/|\lambda_1|$ بستگی دارد. در واقع هرچه این مقدار به صفر نزدیکتر باشد سرعت همگرایی بیشتر و هرچه به یک نزدیک‌تر باشد، کُندتر خواهد بود. آخرین نکتهٔ قابل ذکر در اینجا شرط توقف در الگوریتم ۱.۴ است که برای ϵ داده شده، می‌توان از $|\mu_m - \mu_{m-1}| < \epsilon$ استفاده کرد.

مثال ۹.۴ با استفاده از روش توانی مقدار ویژهٔ غالب و بردار ویژهٔ متناظر با ماتریس

$$A = \begin{pmatrix} -7 & 13 & -16 \\ 13 & -10 & 13 \\ -16 & 13 & -7 \end{pmatrix},$$

را با حداکثر خطای $\epsilon = 0.05$ به دست آورید.

حل: برنامه‌ای به زبان متلب با دقت ساده برای روش توانی نوشته شده و با استفاده از آن و با بردار آغازین $v_0 = (1, 0, 0)^T$ نتایج عددی زیر به دست آمده است:

$$m = 0, \; w_0 = \begin{pmatrix} -7 \\ 13 \\ -16 \end{pmatrix}, \qquad \mu_0 = 16, \qquad v_1 = \begin{pmatrix} -0.4375 \\ 0.8125 \\ -1.0000 \end{pmatrix},$$

$$m = ۱,\ w_۱ = \begin{pmatrix} ۲۹.۶۲۵۰ \\ -۲۶.۸۱۲۵ \\ ۲۴.۵۶۲۵ \end{pmatrix},\ \mu_۱ = ۲۹.۶۲۵۰,\ v_۲ = \begin{pmatrix} ۱.۰۰۰۰ \\ -۰.۹۰۵۱ \\ ۰.۸۲۹۱ \end{pmatrix},$$

$$m = ۲,\ w_۲ = \begin{pmatrix} -۳۲.۰۳۱۶ \\ ۳۲.۸۲۹۱ \\ -۳۳.۵۶۹۶ \end{pmatrix},\ \mu_۲ = ۳۳.۵۶۹۶,\ v_۳ = \begin{pmatrix} -۰.۹۵۴۲ \\ ۰.۹۷۷۹ \\ -۱.۰۰۰۰ \end{pmatrix},$$

$$m = ۳,\ w_۳ = \begin{pmatrix} ۳۵.۳۹۲۵ \\ -۳۵.۱۸۳۸ \\ ۳۴.۹۸۰۲ \end{pmatrix},\ \mu_۳ = ۳۵.۳۹۲۵,\ v_۴ = \begin{pmatrix} ۱.۰۰۰۰ \\ -۰.۹۹۴۱ \\ ۰.۹۸۸۳ \end{pmatrix},$$

$$m = ۴,\ w_۴ = \begin{pmatrix} -۳۵.۷۳۶۹ \\ ۳۵.۷۸۹۶ \\ -۳۵.۸۴۱۸ \end{pmatrix},\ \mu_۴ = ۳۵.۸۴۱۸,\ v_۵ = \begin{pmatrix} -۰.۹۹۷۱ \\ ۰.۹۹۸۵ \\ -۱.۰۰۰۰ \end{pmatrix},$$

$$m = ۵,\ w_۵ = \begin{pmatrix} ۳۵.۹۶۰۶ \\ -۳۵.۹۴۷۴ \\ ۳۵.۹۳۴۳ \end{pmatrix},\ \mu_۵ = ۳۵.۹۶۰۶,\ v_۶ = \begin{pmatrix} ۱.۰۰۰۰ \\ -۰.۹۹۹۶ \\ ۰.۹۹۹۳ \end{pmatrix},$$

$$m = ۶,\ w_۶ = \begin{pmatrix} -۳۵.۹۸۳۵ \\ ۳۵.۹۸۶۸ \\ -۳۵.۹۹۰۱ \end{pmatrix},\ \mu_۶ = ۳۵.۹۹۰۱,\ v_۷ = \begin{pmatrix} -۰.۹۹۹۸ \\ ۰.۹۹۹۹ \\ -۱.۰۰۰۰ \end{pmatrix}.$$

توجه کنید که داریم:

$$|\mu_۶ - \mu_۵| = ۰.۰۲۹۵ < \epsilon.$$

با توجه به اینکه مؤلفه‌های متناظر دو بردار $v_۵$ و $v_۶$ قرینهٔ یکدیگر هستند، داریم:

$$\lambda_۱ \approx -\mu_۶ = -۳۵.۹۹۰۱, \quad x_۱ \approx v_۶ = \begin{pmatrix} -۰.۹۹۹۸ \\ ۰.۹۹۹۹ \\ -۱.۰۰۰۰ \end{pmatrix}.$$

لازم به ذکر است که زوج ویژهٔ غالب ماتریس A به صورت زیر است:

$$\lambda_۱ = -۳۶, \quad x_۱ = \begin{pmatrix} -۱ \\ ۱ \\ -۱ \end{pmatrix}.$$

ملاحظه می‌شود که روش توانی به خوبی تقریب مناسبی از مقدار ویژهٔ غالب A و بردار ویژهٔ متناظر فراهم نموده است.

۲.۲.۴ روش توانی معکوس

همان‌طور که دیدیم در روش توانی یک حدس اولیه مثل v_0 برای بردار ویژهٔ متناظر به مقدار ویژهٔ غالب انتخاب می‌گردد و در هر تکرار این بردار بهبود می‌یابد و مقدار ویژهٔ تقریبی متناظر با آن به دست می‌آید. اما در روش توانی معکوس علاوه بر بردار اولیهٔ v_0، تقریبی از یک مقدار ویژهٔ دلخواه داده می‌شود و دنباله‌ای از اعداد ساخته می‌شود به طوری که به این مقدار ویژه همگرا شود. به‌علاوه می‌توان سرعت همگرایی را بهبود داد.

همانند روش توانی فرض کنید (λ_i, x_i)، $i = 1, 2, \ldots, n$ زوج‌های ویژهٔ A باشند به طوری که x_i ها مستقل خطی هستند. به‌علاوه فرض کنید σ به اندازه کافی به مقدار ویژهٔ λ_i نزدیک باشد، به عبارت دقیق‌تر فرض کنید:

$$0 \neq |\lambda_i - \sigma| < |\lambda_j - \sigma|, \quad \forall j \neq i. \tag{۱۲.۴}$$

در اینجا σ پارامتر انتقال نامیده می‌شود. به ازای $k = 1, 2, \ldots, n$، داریم:

$$A_\sigma x_k = Ax_k - \sigma x_k = \lambda_k x_k - \sigma x_k = (\lambda_k - \sigma)x_k. \tag{۱۳.۴}$$

این رابطه نشان می‌دهد که

$$(\lambda_k - \sigma, x_k), \quad k = 1, 2, \ldots, n,$$

زوج‌های ویژهٔ A_σ هستند. با توجه به رابطهٔ (۱۲.۴)، داریم $\lambda_k - \sigma \neq 0$، $k = 1, 2, \ldots, n$ و در نتیجه ماتریس A_σ نامنفرد است. از این‌رو از رابطهٔ (۱۳.۴) نتیجه می‌گیریم که

$$A_\sigma^{-1} x_k = \gamma_k x_k, \quad k = 1, 2, \ldots, n,$$

که در آن

$$\gamma_k = \frac{1}{\lambda_k - \sigma}.$$

این رابطه نشان می‌دهد که (γ_k, x_k)، $k = 1, 2, \ldots, n$ زوج‌های ویژهٔ A_σ^{-1} می‌باشند و به‌علاوه با توجه به رابطهٔ (۱۲.۴) داریم:

$$|\gamma_i| \gg |\gamma_j|, \quad \forall j \neq i.$$

لازم به ذکر است که نماد $\gg$ برای بسیار بزرگتر بودن استفاده می‌شود. بنابراین (γ_i, x_i) زوج ویژهٔ غالب ماتریس A_σ^{-1} است. لذا برای محاسبهٔ زوج ویژهٔ غالب ماتریس A_σ^{-1} می‌توان از روش توانی استفاده کرد. الگوریتمی که بدین ترتیب حاصل می‌شود، به الگوریتم روش توانی معکوس معروف است. با توجه به اینکه از حیث قدرمطلق، γ_i از سایر مقادیر ویژهٔ A_σ^{-1} بسیار بزرگتر است، روش توانی برای محاسبهٔ مقدار ویژهٔ غالب A_σ^{-1} از سرعت همگرایی بالایی برخوردار خواهد بود. الگوریتم روش توانی را می توانیم به صورت الگوریتم ۲.۴ جمع‌بندی کنیم.

الگوریتم ۲.۴: روش توانی معکوس

1. Choose σ and a vector $v_\circ \in \mathbb{R}^n$ such that $\|v_\circ\|_\infty = ۱$
2. For $m = \circ, ۱, \ldots,$ until convergence, Do
3. $\quad w_m := A_\sigma^{-۱} v_m$
4. $\quad \mu_m := \|w_m\|_\infty$
5. $\quad v_{m+۱} := w_m/\mu_m$
6. EndDo

در گام سومِ الگوریتم برای محاسبهٔ w_m کافی است دستگاه $A_\sigma w_m = v_m$ را حل کنیم. برای این کار، می‌توان از تجزیهٔ LU ماتریس A_σ استفاده کرد (در صورت وجود). توجه کنید که این تجزیه فقط یک بار قبل از اجرای الگوریتم محاسبه می‌گردد. با اجرای الگوریتم یک مقدار تقریبی برای x_i و $\gamma_i = ۱/(\lambda_i - \sigma)$ به دست می‌آید که با استفاده از آن می‌توان یک تقریب برای λ_i به دست آورد.

مثال ۱۰.۴ ماتریس مثال ۹.۴ را در نظر می‌گیریم:

$$A = \begin{pmatrix} -۷ & ۱۳ & -۱۶ \\ ۱۳ & -۱۰ & ۱۳ \\ -۱۶ & ۱۳ & -۷ \end{pmatrix}.$$

نزدیک ترین مقدار ویژهٔ ماتریس A به عدد ۱۰ را با استفاده از روش توانی معکوس با حداکثر خطای ۰.۰۵ به دست آورید.

حل: برنامه‌ای به زبان متلب با دقت ساده برای روش توانی معکوس نوشته شده و با استفاده از آن و با بردار آغازین $v_۰ = (۱, ۰, ۰)^T$ و پارامتر انتقال $\sigma = ۱۰$ نتایج عددی زیر به دست آمده است:

$$m = ۰, \quad w_۰ = \begin{pmatrix} -۰.۵۳۱۱ \\ -۰.۰۴۰۴ \\ ۰.۴۶۸۹ \end{pmatrix}, \quad \mu_۰ = ۰.۵۳۱۱, \quad v_۱ = \begin{pmatrix} -۱.۰۰۰۰ \\ -۰.۰۷۶۰ \\ ۰.۸۸۳۰ \end{pmatrix},$$

$$m = ۱, \quad w_۱ = \begin{pmatrix} ۰.۹۴۸۲ \\ ۰.۰۱۲۵ \\ -۰.۹۳۴۸ \end{pmatrix}, \quad \mu_۱ = ۰.۹۴۸۲, \quad v_۲ = \begin{pmatrix} ۱.۰۰۰۰ \\ ۰.۰۱۳۲ \\ -۰.۹۸۵۹ \end{pmatrix},$$

$$m = ۲, \quad w_۲ = \begin{pmatrix} -۰.۹۹۳۹ \\ -۰.۰۰۱۹ \\ ۰.۹۹۲۰ \end{pmatrix}, \quad \mu_۲ = ۰.۹۹۳۹, \quad v_۳ = \begin{pmatrix} -۱.۰۰۰۰ \\ -۰.۰۰۱۹ \\ ۰.۹۹۸۰ \end{pmatrix}.$$

توجه کنید که در اینجا داریم

$$|\mu_۲ - \mu_۱| = ۰.۰۴۵۷ < ۰.۰۵.$$

با توجه به اینکه مولفه‌های $v_۳$ در مقایسه با مولفه‌های $v_۲$ تغییر علامت می‌دهند، مقدار $-\mu_۲$ یک تقریب از بزرگترین مقدار ویژهٔ (از حیث قدرمطلق) ماتریس $A_\sigma^{-۱}$ است. برای محاسبهٔ نزدیکترین مقدار ویژهٔ A به σ از رابطهٔ

$$\frac{۱}{\lambda_i - \sigma} \approx -\mu_۲$$

استفاده می‌کنیم که در این صورت خواهیم داشت $\lambda_i \approx ۸.۹۹۳۹$. بنابراین زوج ویژهٔ محاسبه شده به صورت زیر خواهد بود:

$$\lambda_i \approx ۸.۹۹۳۹, \quad x_i \approx v_۳ = \begin{pmatrix} -۱.۰۰۰۰ \\ -۰.۰۰۱۹ \\ ۰.۹۹۸۰ \end{pmatrix}.$$

لازم به ذکر است که یک زوج ویژهٔ ماتریس A (مقدار ویژهٔ مربوطه نزدیکترین مقدار ویژه به $\sigma = ۱۰$) به صورت زیر است:

$$\lambda_۱ = ۹, \quad x_۱ = \begin{pmatrix} -۱ \\ ۰ \\ ۱ \end{pmatrix}.$$

ملاحظه می‌شود که روش توانی به خوبی یک تقریب مناسب از یک مقدار ویژهٔ A و بردار ویژهٔ متناظر فراهم نموده است.

۳.۴ محاسبهٔ مقادیر ویژه با استفاده از روش تکراری LU

ماتریس $A \in \mathbb{R}^{n\times n}$ را در نظر بگیرید. در این روش ابتدا قرار می‌دهیم $A_\circ = A$. در تکرار k ام ($k \geqslant ۱$)، تجزیهٔ LU ماتریس $A_{k-۱}$ را محاسبه می‌کنیم (در صورت وجود)، یعنی

$$A_{k-۱} = L_{k-۱}U_{k-۱}. \tag{۱۴.۴}$$

سپس قرار می‌دهیم:

$$A_k = U_{k-۱}L_{k-۱}. \tag{۱۵.۴}$$

به این ترتیب دنباله‌ای مثل $\{A_k\}$ ساخته می‌شود. یکی از خواص جالب این دنباله این است که هر جمله از این دنباله با جمله‌های دیگرِ دنباله متشابه است. زیرا از (۱۴.۴)، داریم:

$$U_{k-۱} = L_{k-۱}^{-۱}A_{k-۱}.$$

با جایگذاری $R_{k-۱}$ در رابطهٔ (۱۵.۴) خواهیم داشت:

$$A_k = L_{k-۱}^{-۱}A_{k-۱}L_{k-۱}.$$

این رابطه نشان می‌دهد که $A_k \sim A_{k-۱}$. به همین ترتیب می‌توان دید که

$$A_k \sim A_{k-۱} \sim \cdots \sim A_\circ = A.$$

بنابراین مقادیر ویژهٔ ماتریس A با مقادیر ویژهٔ هریک از جملات دنباله برابر است.

قضیه ۱۳.۴ اگر مقادیر ویژهٔ A در رابطهٔ

$$|\lambda_۱| > |\lambda_۲| > \cdots > |\lambda_n|, \tag{۱۶.۴}$$

صدق کرده و تجزیهٔ LU تمامی A_k ها موجود باشند، آنگاه دنبالهٔ $\{A_k\}$ که در روش تکراری LU تولید می‌شود به یک ماتریس بالامثلثی همگرا است.

برهان: به [۱۲] مراجعه کنید. □

این قضیه نشان می‌دهد اگر شرط (۱۶.۴) برقرار باشد، آنگاه $\lim_{k\to\infty} A_k = U$ که در آن U یک ماتریس بالامثلثی است. بنابراین درایه‌های روی قطر اصلی U مقادیر ویژهٔ A خواهند بود. در عمل به ازای یک k به اندازه کافی بزرگ، مقادیر روی قطر اصلی A_k را مقادیر ویژهٔ تقریبی A می‌گیریم. در واقع به ازای یک $\epsilon > ۰$ داده شده، مقدار k را طوری انتخاب می‌کنیم که

$$\max_{i>j} |a_{ij}^{(k)}| < \epsilon,$$

که در آن $A_k = (a_{ij}^{(k)})$.

مثال ۱۱.۴ با استفاده از روش تکراری LU، تقریبی از مقادیر ویژهٔ ماتریس

$$A = \begin{pmatrix} ۵ & ۱ & ۲ \\ -۶ & ۶ & ۲ \\ ۶ & -۳ & ۱ \end{pmatrix},$$

را با دقت $\epsilon = ۱۰^{-۳}$ به دست آورید.

حل: مقادیر ویژهٔ A عبارتند از $\lambda_۱ = ۷$، $\lambda_۲ = ۳$ و $\lambda_۳ = ۲$. قرار می‌دهیم $A_۰ = A$. در تکرار اول داریم $A_۰ = L_۰U_۰$، که در آن

$$L_۰ = \begin{pmatrix} ۱.۰۰۰۰ & ۰ & ۰ \\ -۱.۲۰۰۰ & ۱.۰۰۰۰ & ۰ \\ ۱.۲۰۰۰ & -۰.۵۸۳۳ & ۱.۰۰۰۰ \end{pmatrix},$$

$$U_۰ = \begin{pmatrix} ۵.۰۰۰۰ & ۱.۰۰۰۰ & ۲.۰۰۰۰ \\ ۰ & ۷.۲۰۰۰ & ۴.۴۰۰۰ \\ ۰ & ۰ & ۱.۱۶۶۷ \end{pmatrix},$$

و

$$A_۱ = U_۰L_۰ = \begin{pmatrix} ۶.۲۰۰۰ & -۰.۱۶۶۷ & ۲.۰۰۰۰ \\ -۳.۳۶۰۰ & ۴.۶۳۳۳ & ۴.۴۰۰۰ \\ ۱.۴۰۰۰ & -۰.۶۸۰۶ & ۱.۱۶۶۷ \end{pmatrix}.$$

در تکرار دوم داریم $A_1 = L_1U_1$، که در آن

$$L_1 = \begin{pmatrix} 1.0000 & 0 & 0 \\ -0.5419 & 1.0000 & 0 \\ 0.2258 & -0.1415 & 1.0000 \end{pmatrix},$$
$$U_1 = \begin{pmatrix} 6.2000 & -0.1667 & 2.0000 \\ 0.0000 & 4.5430 & 5.4839 \\ 0.0000 & 0 & 1.4911 \end{pmatrix},$$

و

$$A_2 = U_1L_1 = \begin{pmatrix} 6.7419 & -0.4497 & 2.0000 \\ -1.2237 & 3.7669 & 5.4839 \\ 0.3367 & -0.2110 & 1.4911 \end{pmatrix}.$$

در تکرار سوم خواهیم داشت $A_2 = L_2U_2$، که در آن

$$L_2 = \begin{pmatrix} 1.0000 & 0 & 0 \\ -0.1815 & 1.0000 & 0 \\ 0.0499 & -0.0512 & 1.0000 \end{pmatrix},$$
$$U_2 = \begin{pmatrix} 6.7419 & -0.4497 & 2.0000 \\ 0 & 3.6853 & 5.8469 \\ 0 & 0 & 1.6904 \end{pmatrix},$$

و

$$A_3 = U_2L_2 = \begin{pmatrix} 6.9234 & -0.5520 & 2.0000 \\ -0.3769 & 3.3862 & 5.8469 \\ 0.0844 & -0.0865 & 1.6904 \end{pmatrix}.$$

با ادامهٔ این روند در تکرار دوازدهم خواهیم داشت:

$$A_{12} = \begin{pmatrix} 7.0000 & -0.6652 & 2.0000 \\ -0.0000 & 3.0043 & 6.0000 \\ 0.0000 & -0.0007 & 1.9957 \end{pmatrix}.$$

همان‌طور که ملاحظه می‌شود درایه‌های قطری این ماتریس تقریب‌های خوبی از مقادیر ویژهٔ A و درایه‌های زیر قطر اصلی از حیث قدر مطلق کوچکتر ϵ هستند.

۴.۴ تجزیهٔ QR و کاربردهای آن

در این بخش فرض می‌کنیم $A \in \mathbb{R}^{m\times n}$ به طوری که $m \geqslant n$ و $\operatorname{rank}(A) = n$. به سادگی نتایج به دست آمده را می‌توان برای ماتریس‌های دلخواه نیز تعمیم داد. تجزیهٔ QR ماتریس A به صورت $A = QR$ است که در آن Q یک ماتریس متعامد می‌باشد ($Q^TQ = I$) و ماتریس R یک ماتریس بالامثلثی یا به صورت

$$R = \begin{pmatrix} \tilde{R} \\ \circ \end{pmatrix}, \qquad (۱۷.۴)$$

است که در آن $\tilde{R}$ یک ماتریس بالامثلثی است. در حالت اول Q یک ماتریس $m\times n$ و R یک ماتریس $n\times n$ است و در حالت دوم Q یک ماتریس $m\times m$ و R یک ماتریس $m\times n$ است ($\tilde{R}$ یک ماتریس $n\times n$ و ۰ ماتریس صفر با بعد $(m-n)\times n$ است). ماتریس R در رابطهٔ (۱۷.۴) را یک ماتریس بالامثلثی قائم می‌نامیم. در ادامه خواهیم دید که از هر کدام از این تجزیه‌ها، دیگری را می‌توان نتیجه گرفت.

تجزیهٔ QR یک ماتریس کاربردهای مختلفی دارد، از جمله حل دستگاه معادلات خطی، حل مسألهٔ کمترین توان‌های دوم و محاسبهٔ زوج‌های ویژهٔ یک ماتریس. روش‌های مختلفی برای محاسبهٔ تجزیهٔ QR یک ماتریس وجود دارد. در این بخش سه روش مهم زیر را مطالعه می‌کنیم:

(۱) تجزیهٔ QR با استفاده از الگوریتم گرام-اشمیت؛
(۲) تجزیهٔ QR با استفاده از ماتریس‌های هوسهولدر؛
(۳) تجزیهٔ QR با استفاده از ماتریس‌های گیونز.

۱.۴.۴ تجزیهٔ QR با استفاده از الگوریتم گرام-اشمیت

فرض کنید $A = (a_1, a_2, \ldots, a_n)$ یک ماتریس $m\times n$ باشد به طوری که ستون‌های آن، یعنی a_i ها، مستقل خطی هستند. با توجه به الگوریتم گرام-اشمیت که شرح آن در فصل ۲ آمده است، قرار می‌دهیم:

$$q_1 = \frac{a_1}{v_1},$$

$$q_k = \frac{a_k - \sum_{i=۱}^{k-۱} (q_i^T a_k) q_i}{v_k}, \quad k = ۲, ۳, \dots, n,$$

که در آن

$$v_۱ = \|a_۱\|_۲, \quad v_k = \left\| a_k - \sum_{i=۱}^{k-۱} (q_i^T a_k) q_i \right\|_۲.$$

لذا $\{q_۱, q_۲, \dots, q_n\}$ یک مجموعهٔ یکامتعامد است و داریم:

$$\begin{aligned} a_۱ &= v_۱ q_۱, \\ a_k &= (q_۱^T a_k) q_۱ + \dots + (q_{k-۱}^T a_k) q_{k-۱} + v_k q_k, \quad k = ۲, \dots, n. \end{aligned}$$

با نوشتن این روابط به صورت ماتریسی، خواهیم داشت:

$$A = (q_۱, q_۲, \dots, q_n) \begin{pmatrix} v_۱ & q_۱^T a_۲ & q_۱^T a_۳ & \cdots & q_۱^T a_n \\ ۰ & v_۲ & q_۲^T a_۳ & \cdots & q_۲^T a_n \\ ۰ & ۰ & v_۳ & \cdots & q_۳^T a_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ۰ & ۰ & ۰ & \cdots & v_n \end{pmatrix} = QR. \qquad (۱۸.۴)$$

بنابراین ماتریس A به صورت QR تجزیه شده است. توجه کنید $Q = (q_۱, q_۲, \dots, q_n)$ یک ماتریس متعامد است و درایه‌های روی قطر ماتریس R، یعنی v_i ها مثبت می‌باشند. حال می‌توان $n - m$ بردار $q_{n+۱}$، ...، q_m را طوری ساخت که مجموعهٔ بردارهای $\{q_۱, \dots, q_n, q_{n+۱}, \dots, q_m\}$ یک مجموعهٔ یکامتعامد باشند. در این صورت، ماتریس $\tilde{Q} = (q_۱, \dots, q_m)$ متعامد است و با افزودن $(m - n)$ سطر صفر به ماتریس R که در رابطهٔ (۱۸.۴) به دست آمده است، داریم:

$$A = \tilde{Q} \begin{pmatrix} R \\ ۰ \end{pmatrix}.$$

به سادگی می‌توان دید که عکس این مطلب نیز به همین ترتیب امکان پذیر است. قضیهٔ زیر یکتایی تجزیهٔ QR را نشان می‌دهد.

قضیه ۱۴.۴ تجزیهٔ QR ماتریس A، یعنی $A = QR$ که در آن درایه‌های روی قطر R مثبت‌اند، یکتاست.

برهان: بدون از دست دادن کلیت فرض کنید ماتریس R در تجزیهٔ QR بالامثلثی است. همچنین فرض کنید ماتریس A دو تجزیهٔ QR به صورت $A = Q_۱R_۱ = Q_۲R_۲$ داشته باشد به طوری که عناصر روی قطر $R_۱$ و $R_۲$ مثبت باشند. در این صورت، داریم $Q_۲^TQ_۱ = R_۲R_۱^{-۱}$. با توجه به اینکه ماتریس $Q_۲^TQ_۱$ متعامد است (چرا؟)، نتیجه می‌شود که ماتریس $S = R_۲R_۱^{-۱}$ نیز متعامد است. از طرفی چون $R_۱^{-۱}$ و $R_۲$ بالامثلثی هستند نتیجه می‌شود که S بالامثلثی است. بنابراین از اینکه S بالامثلثی و متعامد است نتیجه می‌شود که $S = I$ (چرا؟). بنابراین $R_۱ = R_۲$ و $Q_۱ = Q_۲$. □

مثال ۱۲.۴ تجزیهٔ QR ماتریس

$$A = \begin{pmatrix} ۱ & ۲ & -۱ \\ ۱ & -۱ & ۲ \\ ۱ & -۱ & ۲ \\ -۱ & ۱ & ۱ \end{pmatrix},$$

را به دست آورید.

حل: فرض کنید $A = (a_۱, a_۲, a_۳)$. در این صورت $v_۱ = \|a_۱\|_۲ = ۲$ و

$$q_۱ = \frac{a_۱}{v_۱} = \frac{۱}{۲}\begin{pmatrix} ۱ \\ ۱ \\ ۱ \\ -۱ \end{pmatrix},$$

$$q_۱^Ta_۲ = -\frac{۱}{۲},\ w_۲ = a_۲ - (q_۱^Ta_۲)q_۱ = \frac{۳}{۴}\begin{pmatrix} ۳ \\ -۱ \\ -۱ \\ ۱ \end{pmatrix},\ v_۲ = \|w_۲\|_۲ = \frac{۳\sqrt{۳}}{۲},$$

$$q_۲ = \frac{w_۲}{v_۲} = \frac{۱}{۲\sqrt{۳}}\begin{pmatrix} ۳ \\ -۱ \\ -۱ \\ ۱ \end{pmatrix},$$

$$q_1^T a_3 = 1, \quad q_2^T a_3 = -\sqrt{3}, \quad w_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix},$$

$$v_3 = \|w_3\|_2 = \sqrt{6}, \quad q_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

بنابراین

$$Q = \frac{1}{6} \begin{pmatrix} 3 & 3\sqrt{3} & 0 \\ 3 & -\sqrt{3} & \sqrt{6} \\ 3 & -\sqrt{3} & \sqrt{6} \\ -3 & \sqrt{3} & 2\sqrt{6} \end{pmatrix}, \quad R = \begin{pmatrix} 2 & -\frac{1}{2} & 1 \\ 0 & \frac{3\sqrt{3}}{2} & -\sqrt{3} \\ 0 & 0 & \sqrt{6} \end{pmatrix}.$$

یکی از کاربردهای تجزیهٔ QR حل دستگاه معادلات خطی است. فرض کنید که ماتریس A نامنفرد و $A = QR$ تجزیهٔ QR آن باشد. در این صورت، با جایگذاری تجزیهٔ QR ماتریس A در دستگاه $Ax = b$ داریم $QRx = b$. با ضرب طرفین این رابطه در Q^T داریم

$$Q^T QRx = Q^T b \Rightarrow Rx = Q^T b.$$

حال با توجه به اینکه که ماتریس R یک ماتریس بالامثلثی است، کافی این دستگاه با استفاده از جایگذاری پسرو حل شود.

مثال ۱۳.۴ دستگاه معادلات خطی

$$\begin{pmatrix} 2 & 4 & 3 \\ 2 & 1 & 3 \\ 1 & -1 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \\ -7 \end{pmatrix},$$

را با استفاده از تجزیهٔ QR ماتریس ضرایب حل کنید.

حل: همانند مثال قبل می‌توان دید که $A = QR$، که در آن

$$Q = \frac{1}{3} \begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}, \qquad R = \begin{pmatrix} 3 & 3 & 6 \\ 0 & 3 & -3 \\ 0 & 0 & 3 \end{pmatrix}.$$

حال کافی است دستگاه

$$Rx = Q^T b = \begin{pmatrix} ۳ \\ ۹ \\ -۳ \end{pmatrix},$$

را با استفاده الگوریتم جایگذاری پسرو حل کنیم. با حل این دستگاه خواهیم داشت:

$$x_۱ = ۱, \quad x_۲ = ۲, \quad x_۳ = -۱.$$

۲.۴.۴ تجزیهٔ QR با استفاده از ماتریس‌های هوسهولدر

تعریف ۲.۴ فرض کنید که $v \in \mathbb{R}^n$ یک بردار ناصفر باشد. در این صورت، ماتریس

$$H = I - \frac{۲}{v^T v} v v^T,$$

را یک ماتریس هوسهولدر و v را بردار هوسهولدر گویند.

قضیه ۱۵.۴ ماتریس هوسهولدر H متقارن و متعامد است. به‌علاوه اگر $\circ \neq \alpha \in \mathbb{R}$، آنگاه بردارهای v و αv ماتریس‌های هوسهولدر یکسانی را تولید می‌کنند.

برهان: به سادگی می‌توان دید که $H^T = H$. از طرفی داریم:

$$H^T H = H^۲ = (I - \frac{۲}{v^T v} v v^T)^۲ = I - \frac{۴}{v^T v} v v^T + \frac{۴}{(v^T v)^۲} v (v^T v) v^T = I.$$

بنابراین ماتریس H متعامد است. قسمت دوم قضیه واضح است. □

قضیه ۱٦.۴ اگر $\circ \neq x \in \mathbb{R}^n$ و $v = x \pm \|x\|_۲ e_۱$ و H ماتریس هوسهولدر متناظر با بردار v باشد، آنگاه $Hx = \mp \|x\|_۲ e_۱$ که در آن $e_۱$ ستون اول ماتریس همانی $n \times n$ است.

برهان: هدف یافتن برداری مثل $v \in \mathbb{R}^n$ است به طوری که $Hx = \alpha e_۱$. از اینکه $\|Hx\|_۲ = |\alpha|$ و با توجه به تمرین ۱۸.۲ و اینکه ماتریس H متعامد است خواهیم داشت $\alpha = \mp \|x\|_۲$. از طرفی می‌خواهیم بردار v را طوری بیابیم که $Hx = \alpha e_۱$، یا

$$Hx = (I - \frac{۲}{v^T v} v v^T) x = x - ۲ \frac{v^T x}{v^T v} v = \alpha e_۱,$$

یا

$$x - \alpha e_۱ = ۲\frac{v^T x}{v^T v}v.$$

این رابطه نشان می‌دهد که بردار v با بردار $x - \alpha e_۱$ هم‌راستا است. از این‌رو با توجه به قسمت دوم قضیهٔ ۱۵.۴ می‌توانیم قرار دهیم:

$$v = x - \alpha e_۱ = x \pm \|x\|_۲ e_۱,$$

که اثبات قضیه را کامل می‌کند. $\square$

در عمل برای جلوگیری از خطای ناشی از تفریق دو عدد نزدیک به هم اگر $\xi_۱ = ۰$، آنگاه قرار می‌دهیم $v = x + \|x\|_۲ e_۱$، که در این صورت خواهیم داشت $Hx = -\|x\|_۲ e_۱$. در صورتی که $\xi_۱ \neq ۰$، قرار می‌دهیم $v = x + \mathrm{sign}(\xi_۱)\|x\|_۲ e_۱$ که در این صورت، نتیجه می‌شود $Hx = -\mathrm{sign}(\xi_۱)\|x\|_۲ e_۱$.

مثال ۱۴.۴ فرض کنید $x = (۲, ٦, -۳)^T$. داریم $\xi_۱ = ۲$ و $\|x\|_۲ = ۷$. در نتیجه

$$v = x + \mathrm{sign}(\xi_۱)\|x\|_۲ e_۱ = \begin{pmatrix} ۲ \\ ٦ \\ -۳ \end{pmatrix} + ۷\begin{pmatrix} ۱ \\ ۰ \\ ۰ \end{pmatrix} = \begin{pmatrix} ۹ \\ ٦ \\ -۳ \end{pmatrix}.$$

از این‌رو

$$H = I - \frac{۲}{v^T v}vv^T = \frac{۱}{۷}\begin{pmatrix} -۲ & -٦ & ۳ \\ -٦ & ۳ & ۲ \\ ۳ & ۲ & ٦ \end{pmatrix}.$$

لذا $Hx = -\mathrm{sign}(\xi_۱)\|x\|_۲ e_۱ = (-۷, ۰, ۰)^T$. توجه کنید که $v = ۳(۳, ۲, -۱)^T$، بنابراین با استناد به قضیهٔ ۱۵.۴، اگر بردار v را $(۳, ۲, -۱)^T$ اختیار کنیم، نتیجهٔ حاصل یکسان خواهد بود.

حال چگونگی به دست آوردن تجزیهٔ QR یک ماتریس با استفاده از ماتریس‌های هوسهولدر را شرح می‌دهیم. ماتریس $A = (a_۱, a_۲, \ldots, a_n) \in \mathbb{R}^{m\times n}$ را در نظر بگیرید که در آن a_i ستون i ام ماتریس A است. یک ماتریس هوسهولدر مثل $H_۱$ می‌سازیم به طوری که

$Ha_1 = (r_{11}, 0, \dots, 0)^T$. در این صورت، خواهیم داشت:

$$H_1 A = (H_1 a_1, H_1 a_2, \dots, H_1 a_n) = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} = \begin{pmatrix} r_{11} & r_1^T \\ 0 & A_2 \end{pmatrix}.$$

حال به طور مشابه یک ماتریس هوسهولدر مثل $\tilde{H}_2$ می‌سازیم به طوری که درایه‌های زیر مکان $(1, 1)$ از ماتریس A_2 را صفر کند. سپس ماتریس

$$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{H}_2 \end{pmatrix},$$

را تعریف می‌کنیم. روشن است که H_2 یک ماتریس متقارن و متعامد است. لذا

$$H_2 H_1 A = \begin{pmatrix} r_{11} & r_1^T \\ 0 & \tilde{H}_2 A_2 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & * & & * \\ 0 & 0 & * & \cdots & * \end{pmatrix}.$$

با ادامهٔ این روند بعد از $k - 1$ مرحله خواهیم داشت:

$$H_{k-1} \cdots H_2 H_1 A = \begin{pmatrix} R_{k-1} & \tilde{R}_{k-1} \\ 0 & A_k \end{pmatrix},$$

که در آن ماتریس R_{k-1} بالامثلثی است. در مرحلهٔ k ام یک ماتریس هوسهولدر مثل $\tilde{H}_k$ ساخته می‌شود به طوری که درایه‌های زیر مکان $(1, 1)$ از ماتریس A_k را صفر کند. حال ماتریس H_k را به صورت

$$H_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & \tilde{H}_k \end{pmatrix},$$

تعریف می‌کنیم. به سادگی می‌توان دید که ماتریس H_k متقارن و متعامد است. از این‌رو، داریم:

$$H_k \cdots H_2 H_1 A = \begin{pmatrix} R_k & \tilde{R}_k \\ 0 & A_{k+1} \end{pmatrix},$$

که R_k بالامثلثی است. با ادامهٔ این روند نتیجهٔ مطلوب به دست می‌آید. در واقع دو حالت زیر را در نظر می‌گیریم.

- اگر $m > n$ باشد، آنگاه بعد از n مرحله، خواهیم داشت:

$$H_n \cdots H_۲ H_۱ A = \begin{pmatrix} R_n \\ ۰ \end{pmatrix}.$$

در این صورت، قرار می‌دهیم $Q = (H_n \cdots H_۲ H_۱)^T = H_۱ H_۲ \cdots H_n$ و در نتیجه

$$A = Q \begin{pmatrix} R_n \\ ۰ \end{pmatrix},$$

که در آن R_n یک ماتریس $n \times n$ بالامثلثی و ۰ ماتریس صفر است.

- در صورتی که $m = n$ باشد، آنگاه بعد از $n - ۱$ مرحله، داریم:

$$H_{n-۱} \cdots H_۲ H_۱ A = R_n,$$

که در آن R_n یک ماتریس بالامثلثی است. با قرار دادن

$$Q = (H_{n-۱} \cdots H_۲ H_۱)^T = H_۱ H_۲ \cdots H_{n-۱},$$

خواهیم داشت $A = QR_n$.

انتظار می‌رود که خواننده حالت $m < n$ را بررسی نماید.

مثال ۱۵.۴ به کمک ماتریس‌های هوسهولدر تجزیهٔ QR ماتریس

$$A = \begin{pmatrix} ۱ & ۵ & ۱۰ \\ -۲ & ۰ & ۵ \\ ۲ & -۱ & ۶ \end{pmatrix},$$

را به دست آورید.

حل: فرض کنید $x = (۱, -۲, ۲)^T$. لذا $v = x + \|x\|_۲ e_۱ = (۴, -۲, ۲)^T$. در این صورت، ماتریس هوسهولدر متناظر با این بردار برابر با

$$H_۱ = \frac{۱}{۳} \begin{pmatrix} -۱ & ۲ & -۲ \\ ۲ & ۲ & ۱ \\ -۲ & ۱ & ۲ \end{pmatrix},$$

است و

$$A_۱ = H_۱ A = \begin{pmatrix} -۳ & -۱ & -۴ \\ ۰ & ۳ & ۱۲ \\ ۰ & -۴ & -۱ \end{pmatrix}.$$

حال قرار می‌دهیم $x = (۳, -۴)^T$ و $v = x - \|x\|_۲ e_۱ = (۸, -۴)^T$. ماتریس هوسهولدر متناظر به v را $\tilde{H}_۲$ می‌نامیم و داریم:

$$\tilde{H}_۲ = \frac{۱}{۵} \begin{pmatrix} -۳ & ۴ \\ ۴ & ۳ \end{pmatrix}.$$

ماتریس $H_۲$ را به صورت

$$H_۲ = \begin{pmatrix} ۱ & ۰ \\ ۰ & \tilde{H}_۲ \end{pmatrix},$$

تعریف می‌کنیم. در نتیجه

$$A_۲ = H_۲ A_۱ = H_۲ H_۱ A = \begin{pmatrix} -۳ & -۱ & -۴ \\ ۰ & -۵ & -۸ \\ ۰ & ۰ & ۹ \end{pmatrix} =: R.$$

از طرفی داریم

$$Q = H_۱ H_۲ = \frac{۱}{۱۵} \begin{pmatrix} -۵ & -۱۴ & ۲ \\ ۱۰ & -۲ & ۱۱ \\ -۱۰ & ۵ & ۱۰ \end{pmatrix},$$

و $A = QR$.

مثال ۱۶.۴ با استفاده از ماتریس‌های هوسهولدر تجزیهٔ QR ماتریس

$$A = \begin{pmatrix} ۱ & ۲ & -۱ \\ ۱ & -۱ & ۱ \\ ۱ & -۱ & ۰ \\ -۱ & ۲ & ۰ \end{pmatrix},$$

را به دست آورید.

حل: فرض کنید $x = (1, 1, 1, -1)^T$. در نتیجه $v = x + \|x\|_2 e_1 = (3, 1, 1, -1)^T$.
در این صورت ماتریس هوسهولدر متناظر با این بردار عبارت است از

$$H_1 = \frac{1}{6}\begin{pmatrix} -3 & -3 & -3 & 3 \\ -3 & 5 & -1 & 1 \\ -3 & -1 & 5 & 1 \\ 3 & 1 & 1 & 5 \end{pmatrix},$$

و

$$A_1 = H_1 A = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -\frac{4}{3} & \frac{4}{3} \\ 0 & -\frac{4}{3} & \frac{1}{3} \\ 0 & \frac{7}{3} & -\frac{1}{3} \end{pmatrix}.$$

حال قرار می‌دهیم $x = (-\frac{4}{3}, -\frac{4}{3}, \frac{7}{3})^T$ و $v = x - \|x\|_2 e_1 = (-\frac{13}{3}, -\frac{4}{3}, \frac{7}{3})^T$. ماتریس هوسهولدر متناظر به v را $\tilde{H}_2$ می‌نامیم و داریم:

$$\tilde{H}_2 = \frac{1}{117}\begin{pmatrix} -52 & -52 & 91 \\ -52 & 101 & 28 \\ 91 & 28 & 68 \end{pmatrix}.$$

ماتریس H_2 را به صورت زیر تعریف می‌کنیم:

$$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{H}_2 \end{pmatrix}.$$

با استفاده از این ماتریس خواهیم داشت:

$$A_2 = H_2 H_1 A = \begin{pmatrix} -2 & 1 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & -\frac{5}{13} \\ 0 & 0 & \frac{12}{13} \end{pmatrix}.$$

در نهایت قرار می‌دهیم $x = (-\frac{5}{13}, \frac{12}{13})^T$ و در نتیجه $v = x - \|x\|_2 e_1 = (-\frac{18}{13}, \frac{12}{13})^T$.
از این‌رو، ماتریس هوسهولدر متناظر به v به صورت

$$\tilde{H}_3 = \frac{1}{13}\begin{pmatrix} -5 & 12 \\ 12 & 5 \end{pmatrix},$$

است و قرار می‌دهیم:

$$H_۳ = \begin{pmatrix} I_۲ & ۰ \\ ۰ & \tilde{H}_۳ \end{pmatrix}.$$

لذا

$$H_۳ H_۲ H_۱ A = \begin{pmatrix} -۲ & ۱ & ۰ \\ ۰ & ۳ & -۱ \\ ۰ & ۰ & ۱ \\ ۰ & ۰ & ۰ \end{pmatrix} = R,$$

و

$$Q = H_۱ H_۲ H_۳ = \frac{۱}{۶} \begin{pmatrix} -۳ & ۵ & -۱ & -۱ \\ -۳ & -۱ & ۵ & -۱ \\ -۳ & -۱ & -۱ & ۵ \\ ۳ & ۳ & ۳ & ۳ \end{pmatrix},$$

به طوری که $A = QR$.

قضیه ۱۷.۴ هر ماتریس $n \times n$، A با یک ماتریس بالاهسنبرگی به طور متعامد متشابه است، به عبارت دیگر یک ماتریس متعامد مثل P وجود دارد به طوری که $PAP^T = H$ که در آن H یک ماتریس بالاهسنبرگی است.

برهان: فرض کنید $A = (a_{ij})$. یک ماتریس هوسهولدر مثل $\tilde{H}_۱$ می‌سازیم به طوری که

$$\tilde{H}_۱ \begin{pmatrix} a_{۲۱} \\ a_{۳۱} \\ \vdots \\ a_{n۱} \end{pmatrix} = \begin{pmatrix} * \\ ۰ \\ \vdots \\ ۰ \end{pmatrix}.$$

حال تعریف می‌کنیم:

$$H_۱ = \begin{pmatrix} ۱ & ۰ \\ ۰ & \tilde{H}_۱ \end{pmatrix}.$$

در این صورت،

$$H_۱ A H_۱^T = A^{(۱)},$$

که درایه‌های زیر مکان $(۲,۱)$ از ماتریس $A^{(۱)}$ صفرند (چرا؟). فرض کنید $A^{(۱)} = (a_{ij}^{(۱)})$. یک ماتریس هوسهولدر مثل $\tilde{H}_۲$ می‌سازیم به طوری که

$$\tilde{H}_۲ \begin{pmatrix} a_{۳۲}^{(۱)} \\ a_{۴۲}^{(۱)} \\ \vdots \\ a_{n۲}^{(۱)} \end{pmatrix} = \begin{pmatrix} * \\ ۰ \\ \vdots \\ ۰ \end{pmatrix}.$$

تعریف می‌کنیم:

$$H_۲ = \begin{pmatrix} I_۲ & ۰ \\ ۰ & \tilde{H}_۲ \end{pmatrix}.$$

قرار می‌دهیم:

$$H_۲ H_۱ A H_۱^T H_۲^T = A^{(۲)}.$$

در این صورت، می‌توان دید که درایه‌های زیر مکان‌های $(۲,۱)$ و $(۳,۲)$ از ماتریس $A^{(۲)}$ صفرند. با ادامهٔ این روند بعد از $n-۲$ مرحله ماتریس‌های $H_۳$، ...، $H_{n-۲}$ ساخته می‌شوند به طوری که

$$H_{n-۲} \dots H_۳ H_۲ H_۱ A H_۱^T H_۲^T H_۳^T \dots H_{n-۲}^T = A^{(n-۲)} = H,$$

که H یک ماتریس بالاهسنبرگی است. بنابراین با قرار دادن $P = H_{n-۲} \dots H_۳ H_۲ H_۱$، نتیجهٔ مطلوب به دست می‌آید. □

نتیجه ۴.۴ هر ماتریس متقارن $A \in \mathbb{R}^{n\times n}$ با یک ماتریس سه‌قطری به طور متعامد متشابه است.

برهان: فرض کنید A متقارن باشد. بنا بر قضیهٔ ۱۷.۴، یک ماتریس متعامد مثل P وجود دارد به طوری که $PAP^T = H$ و H بالاهسنبرگی است. اما

$$H^T = (PAP^T)^T = PA^TP^T = PAP^T = H.$$

بنابراین H متقارن است و از اینکه بالاهسنبرگی است نتیجه می‌شود که سه‌قطری است. □

مثال ۱۷.۴ فرض کنید

$$A=\begin{pmatrix} ۱ & ۱ & -۲ & ۲ \\ ۱ & -۱ & -۱ & ۲ \\ -۲ & -۱ & ۱ & ۲ \\ ۲ & ۲ & ۲ & ۱ \end{pmatrix}.$$

یک ماتریس متعامد مثل P بیابید به طوری که $PAP^T=H$ که در آن H یک ماتریس سه‌قطری باشد.

حل: با استفاده از قضیهٔ ۱۷.۴ و نتیجهٔ ۴.۴ ماتریس P را می‌سازیم. برای این کار، قرار می‌دهیم $x=(۱,-۲,۲)^T$. در نتیجه

$$v=x+\|x\|_۲ e_۱=\begin{pmatrix} ۴ \\ -۲ \\ ۲ \end{pmatrix}.$$

در این صورت، ماتریس هوسهولدر متناظر با این بردار عبارت است از

$$\tilde{H}_۱=\frac{۱}{۳}\begin{pmatrix} -۱ & ۲ & -۲ \\ ۲ & ۲ & ۱ \\ -۲ & ۱ & ۲ \end{pmatrix}.$$

تعریف می‌کنیم:

$$H_۱=\begin{pmatrix} ۱ & ۰ \\ ۰ & \tilde{H}_۱ \end{pmatrix}.$$

در این صورت،

$$A^{(۱)}=H_۱AH_۱^T=\begin{pmatrix} ۱ & -۳ & ۰ & ۰ \\ -۳ & \frac{۱}{۳} & -\frac{۴}{۳} & ۱ \\ ۰ & -\frac{۴}{۳} & ۱ & \frac{۸}{۳} \\ ۰ & ۱ & \frac{۸}{۳} & -\frac{۱}{۳} \end{pmatrix}.$$

حال قرار می‌دهیم $x = (-\frac{۴}{۳}, ۱)^T$ و

$$v = x - \|x\|_۲ e_۱ = \begin{pmatrix} -۳ \\ ۱ \end{pmatrix}.$$

ماتریس هوسهولدر متناظر با v را $\tilde{H}_۲$ می‌نامیم. لذا

$$\tilde{H}_۲ = \frac{۱}{۵} \begin{pmatrix} -۴ & ۳ \\ ۳ & ۴ \end{pmatrix}.$$

ماتریس $H_۲$ را به صورت

$$H_۲ = \begin{pmatrix} I_۲ & ۰ \\ ۰ & \tilde{H}_۲ \end{pmatrix},$$

می‌سازیم. با استفاده از این ماتریس خواهیم داشت:

$$A^{(۲)} = H_۲ A^{(۱)} H_۲^T = \begin{pmatrix} ۱ & -۳ & ۰ & ۰ \\ -۳ & \frac{۱}{۳} & -\frac{۵}{۳} & ۰ \\ ۰ & \frac{۵}{۳} & -\frac{۵۱}{۲۵} & -\frac{۱۰۴}{۷۵} \\ ۰ & ۰ & -\frac{۱۰۴}{۷۵} & \frac{۲۰۳}{۷۵} \end{pmatrix} = H.$$

در نتیجه

$$H \;=\; A^{(۲)} = H_۲ A^{(۱)} H_۲^T = H_۲ H_۱ A H_۲^T H_۱^T.$$

لذا با فرض $P = H_۲ H_۱$، خواهیم داشت:

$$P = \frac{۱}{۱۵} \begin{pmatrix} ۱۵ & ۰ & ۰ & ۰ \\ ۰ & -۵ & ۱۰ & -۱۰ \\ ۰ & -۱۴ & -۵ & ۲ \\ ۰ & -۲ & ۱۰ & ۱۱ \end{pmatrix},$$

و $PAP^T = H$.

در ادامه خواهیم دید که چگونه قضیهٔ ۱۷.۴ و نتیجهٔ ۴.۴ در محاسبهٔ مقادیر ویژهٔ یک ماتریس کاربرد دارند.

۳.۴.۴ تجزیهٔ QR با استفاده از ماتریس‌های گیونز

تعریف ۳.۴ اعداد حقیقی c و s را در نظر بگیرید به طوری که $c^۲ + s^۲ = ۱$. همچنین فرض کنید $۱ \leqslant i < j \leqslant n$. درایه‌های مکان (i,i)، (i,j)، (j,i) و (j,j) از ماتریس همانی I_n را به ترتیب به c، s، $-s$ و c تبدیل می‌کنیم. ماتریس حاصل را با $J(i,j,c,s)$ نشان داده و یک ماتریس گیونز گویند. اینگونه ماتریس‌ها به صورت زیر هستند:

$$J(i,j,c,s) = \begin{array}{c} \begin{array}{cccccc} & & i\text{ ام} & & j\text{ ام} & \\ & & \downarrow & & \downarrow & \end{array} \\ \left(\begin{array}{cccccc} ۱ & & & & & \\ & \ddots & & & & \\ & & c & & s & \\ & & & \ddots & & \\ & & -s & & c & \\ & & & & \ddots & \\ & & & & & ۱ \end{array}\right) \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{ ام} \\ \\ \leftarrow j\text{ ام} \\ \\ \\ \end{array} . \qquad (۱۹.۴)$$

برای سادگی گاهی‌اوقات این ماتریس را با J_{ij} نشان می‌دهیم. ماتریس $J(i,j,c,s)$ متقارن نیست اما به سادگی می‌توان دید که متعامد است.

بردار $x = (x_۱, x_۲, \ldots, x_i, \ldots, x_j, \ldots, x_n)^T$ را در نظر بگیرید. مقادیر c و s را طوری می‌یابیم که

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} \alpha \\ ۰ \end{pmatrix}.$$

برای برقراری رابطهٔ اخیر کافی است داشته باشیم $-sx_i + cx_j = ۰$ یا $c = sx_i/x_j$. با جایگذاری c در $c^۲ + s^۲ = ۱$، خواهیم داشت:

$$c = \pm \frac{x_i}{\sqrt{x_i^۲ + x_j^۲}}.$$

با در نظر گرفتن علامت مثبت، داریم:

$$c = \frac{x_i}{\sqrt{x_i^۲ + x_j^۲}}, \qquad s = \frac{x_j}{\sqrt{x_i^۲ + x_j^۲}},$$

و

$$\alpha = cx_i + sx_j = \sqrt{x_i^{۲} + x_j^{۲}}.$$

حال ماتریس $J(i,j,c,s)$ را همانند رابطهٔ (۱۹.۴) تعریف می‌کنیم. با محاسبهٔ بردار $J(i,j,c,s)x$ ملاحظه می‌کنیم که درایهٔ j ام از این بردار برابر با صفر، درایهٔ i ام برابر است با $\sqrt{x_i^{۲} + x_j^{۲}}$ و سایر درایه‌های بردار x بدون تغییر باقی می‌مانند.

مثال ۱۸.۴ فرض کنید $x = (۳, ۱, ۴)^T$. می‌خواهیم درایهٔ سوم از این بردار را صفر کنیم. برای این کار، فرض می‌کنیم $i = ۱$ و $j = ۳$. در این صورت، داریم:

$$c = \frac{۳}{\sqrt{۳^{۲} + ۴^{۲}}} = \frac{۳}{۵}, \quad s = \frac{۴}{\sqrt{۳^{۲} + ۴^{۲}}} = \frac{۴}{۵}.$$

حال ماتریس $J(۱, ۳, \frac{۳}{۵}, \frac{۴}{۵})$ را به صورت زیر تعریف می‌کنیم:

$$J(۱, ۳, \frac{۳}{۵}, \frac{۴}{۵}) = \begin{pmatrix} \frac{۳}{۵} & ۰ & \frac{۴}{۵} \\ ۰ & ۱ & ۰ \\ -\frac{۴}{۵} & ۰ & \frac{۳}{۵} \end{pmatrix}.$$

لذا

$$J(۱, ۳, \frac{۳}{۵}, \frac{۴}{۵})\, x = \begin{pmatrix} ۵ \\ ۱ \\ ۰ \end{pmatrix}.$$

حال می‌خواهیم روندی ارائه کنیم که با استفاده از آن بتوان درایهٔ مشخصی در یک ماتریس را صفر کرد. ماتریس $A = (a_{ij})$ را در نظر بگیرید. می‌خواهیم درایهٔ مکان (j,i) ماتریس را صفر کنیم. برای این کار، مقادیر c و s را طوری می‌یابیم که

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{ii} \\ a_{ji} \end{pmatrix} = \begin{pmatrix} \alpha \\ ۰ \end{pmatrix}.$$

همان‌طور که قبلاً دیدیم کافی است قرار دهیم:

$$c = \frac{a_{ii}}{\sqrt{a_{ii}^{۲} + a_{ji}^{۲}}}, \quad s = \frac{a_{ji}}{\sqrt{a_{ii}^{۲} + a_{ji}^{۲}}}. \tag{۲۰.۴}$$

سپس با استفاده از (۱۹.۴) ماتریس $J(i,j,c,s)$ را می‌سازیم. با ضرب این ماتریس از سمت چپ در ماتریس A درایهٔ مکان (j,i) صفر می‌شود و به اصطلاح گوییم با استفاده از درایهٔ مکان (i,i) درایهٔ مکان (j,i) را صفر کرده‌ایم. با استفاده از این روند می‌توان تجزیهٔ QR یک ماتریس را به دست آورد.

ماتریس، $m \times n$، $A = (a_{ij})$ را در نظر بگیرید. در گام اول، ماتریس‌های گیونز $J_{۱۲}$، $J_{۱۳}$، $\cdots$، $J_{۱m}$ را می‌سازیم به طوری که درایه‌های زیر مکان $(۱,۱)$ از ماتریس

$$A^{(۱)} = J_{۱m} J_{۱,m-۱} \cdots J_{۱۲}\ A = Q_۱ A,$$

برابر با صفر باشند. این کار امکان پذیر است، زیرا با استفاده از ماتریس $J_{۱۲}$ درایهٔ مکان $(۲,۱)$ از ماتریس A صفر می‌شود. به عبارتی درایهٔ مکان $(۲,۱)$ از ماتریس $J_{۱۲}A$ صفر است. حال یک ماتریس گیونز مثل $J_{۱۳}$ می‌سازیم به طوری که درایهٔ مکان $(۳,۱)$ از ماتریس $J_{۱۲}A$ را صفر کند. توجه کنید که اگر $J_{۱۳}$ را از سمت چپ در $J_{۱۲}A$ ضرب کنیم، فقط دو سطر اول و سوم این ماتریس تغییر می‌کند. بنابراین درایهٔ $(۱,۲)$ ماتریس $J_{۱۳}J_{۱۲}A$ صفر خواهد ماند. با ادامهٔ این روند تمام درایه‌های زیر مکان $(۱,۱)$ از ماتریس A صفر می‌شوند. در گام k ام $(۲ \leqslant k \leqslant s = \min(n, m-۱))$ ماتریس‌های گیونز $J_{k,k+۱}$، $J_{k,k+۲}$، $\cdots$، J_{km} می‌سازیم به طوری که درایه‌های زیر مکان (k,k) از ماتریس

$$A^{(k)} = J_{km} J_{k,m-۱} \cdots J_{k,k+۱}\ A^{(k-۱)} = Q_k A^{(k-۱)},$$

برابر با صفر باشند. از این رو

$$R = A^{(s)} = Q_s A^{(s-۱)} = Q_s Q_{s-۱} A^{(s-۲)} = \cdots = Q_s Q_{s-۱} \ldots Q_۲ Q_۱ A.$$

قرار می‌دهیم $Q = (Q_s Q_{s-۱} \cdots Q_۲ Q_۱)^T$. دراین صورت خواهیم داشت $R = Q^T A$. لذا $A = QR$ و بنابراین نتیجهٔ مطلوب به دست می آید.

مثال ۱۹.۴ با استفاده از ماتریس‌های گیونز تجزیهٔ QR ماتریس

$$A = \begin{pmatrix} ۳ & ۲ \\ ۱ & ۳ \\ -۱ & ۱ \\ ۲ & ۲ \end{pmatrix},$$

را به دست آورید.

حل: در این مثال محاسبات را با چهار رقم اعشار انجام می‌دهیم. داریم $a_{۱۱} = ۳$ و $a_{۲۱} = ۱$. با استفاده از رابطهٔ (۴.۲۰) داریم $c = ۰.۹۴۸۷$ و $s = ۰.۳۱۶۲$. لذا

$$A := J_{۱۲}A = \begin{pmatrix} ۳.۱۶۲۳ & ۲.۸۴۶۰ \\ ۰.۰۰۰۰ & ۲.۲۱۳۶ \\ -۱.۰۰۰۰ & ۱.۰۰۰۰ \\ ۲.۰۰۰۰ & ۲.۰۰۰۰ \end{pmatrix}.$$

با توجه به اینکه $a_{۱۱} = ۳.۱۶۲۳$ و $a_{۳۱} = -۱.۰۰۰۰$ خواهیم داشت $c = ۰.۹۵۳۵$ و $s = -۰.۳۰۱۵$. در نتیجه

$$A := J_{۱۳}A = \begin{pmatrix} ۳.۳۱۶۶ & ۲.۴۱۲۱ \\ ۰.۰۰۰۰ & ۲.۲۱۳۶ \\ ۰.۰۰۰۰ & ۱.۸۱۱۶ \\ ۲.۰۰۰۰ & ۲.۰۰۰۰ \end{pmatrix}.$$

همچنین $a_{۱۱} = ۳.۳۱۶۶$ و $a_{۴۱} = ۲.۰۰۰۰$. در نتیجه $c = ۰.۸۵۶۳$ و $s = ۰.۵۱۶۴$. از این‌رو

$$A := J_{۱۴}A = \begin{pmatrix} ۳.۸۷۳۰ & ۳.۰۹۸۴ \\ ۰.۰۰۰۰ & ۲.۲۱۳۶ \\ ۰.۰۰۰۰ & ۱.۸۱۱۶ \\ ۰.۰۰۰۰ & ۰.۴۶۷۱ \end{pmatrix}.$$

تاکنون درایه‌های زیر مکان $(۱,۱)$ از ماتریس A صفر شده‌اند. حال درایه‌های زیر مکان $(۲,۲)$ را صفر می‌کنیم. داریم $a_{۲۲} = ۲.۲۱۳۶$ و $a_{۳۲} = ۱.۸۱۱۶$. در نتیجه $c = ۰.۷۷۳۹$ و $s = ۰.۶۳۳۳$. لذا

$$A := J_{۲۳}A = \begin{pmatrix} ۳.۸۷۳۰ & ۳.۰۹۸۴ \\ ۰.۰۰۰۰ & ۲.۸۶۰۴ \\ ۰.۰۰۰۰ & ۰.۰۰۰۰ \\ ۰.۰۰۰۰ & ۰.۴۶۷۱ \end{pmatrix}.$$

در نهایت داریم $a_{۲۲} = ۲.۸۶۰۴$ و $a_{۴۲} = ۰.۴۶۷۱$. در نتیجه $c = ۰.۹۸۶۹$ و $s = ۰.۱۶۱۲$. بنابراین

$$A := J_{۲۴}A = \begin{pmatrix} ۳.۸۷۳۰ & ۳.۰۹۸۴ \\ ۰.۰۰۰۰ & ۲.۸۹۸۳ \\ ۰.۰۰۰۰ & ۰.۰۰۰۰ \\ ۰.۰۰۰۰ & ۰.۰۰۰۰ \end{pmatrix} = R.$$

برای محاسبهٔ ماتریس Q در تجزیهٔ QR، داریم:

$$Q_۱ = J_{۱۴}J_{۱۳}J_{۱۲} = \begin{pmatrix} ۰.۷۷۴۶ & ۰.۲۵۸۲ & -۰.۲۵۸۲ & ۰.۵۱۶۴ \\ -۰.۳۱۶۲ & ۰.۹۴۸۷ & ۰.۰۰۰۰ & ۰.۰۰۰۰ \\ ۰.۲۸۶۰ & ۰.۰۹۵۳ & ۰.۹۵۳۵ & ۰.۰۰۰۰ \\ -۰.۴۶۷۱ & -۰.۱۵۵۷ & ۰.۱۵۵۷ & ۰.۸۵۶۳ \end{pmatrix},$$

$$Q_۲ = J_{۲۳}J_{۲۴} = \begin{pmatrix} ۱.۰۰۰۰ & ۰.۰۰۰۰ & ۰.۰۰۰۰ & ۰.۰۰۰۰ \\ ۰.۰۰۰۰ & ۰.۷۶۳۸ & ۰.۶۲۵۱ & ۰.۱۶۱۲ \\ ۰.۰۰۰۰ & -۰.۶۳۳۳ & ۰.۷۷۳۹ & ۰.۰۰۰۰ \\ ۰.۰۰۰۰ & -۰.۱۲۴۷ & -۰.۱۰۲۱ & ۰.۹۸۶۹ \end{pmatrix},$$

لذا

$$Q = (Q_۲Q_۱)^T = \begin{pmatrix} ۰.۷۷۴۶ & -۰.۱۳۸۰ & ۰.۴۲۱۶ & -۰.۴۵۰۷ \\ ۰.۲۵۸۲ & ۰.۷۵۹۱ & -۰.۵۲۷۰ & -۰.۲۸۱۷ \\ -۰.۲۵۸۲ & ۰.۶۲۱۱ & ۰.۷۳۷۹ & ۰.۰۵۶۳ \\ ۰.۵۱۶۴ & ۰.۱۳۸۰ & ۰.۰۰۰۰ & ۰.۸۴۵۲ \end{pmatrix}.$$

مثال ۲۰.۴ تجزیهٔ QR ماتریس بالاهسنبرگی زیر رابدست آورید:

$$A = \begin{pmatrix} ۳ & -۶ & -۵ & -۱ \\ ۴ & ۲ & -۱۰ & ۲ \\ ۰ & ۸ & ۴ & -۴ \\ ۰ & ۰ & ۳ & ۷ \end{pmatrix}.$$

حل: با توجه به اینکه ماتریس A بالاهسنبرگی است، کافی است درایه‌های $(۲,۱)$، $(۳,۲)$ و $(۴,۳)$ را صفر کنیم. ابتدا با استفاده از درایهٔ $a_{۱۱} = ۳$، درایهٔ $a_{۲۱} = ۴$ را صفر می‌کنیم.

به کمک رابطهٔ (۲۰.۴) داریم $c = 0.6$ و $s = 0.8$. لذا

$$A := J_{12}A = \begin{pmatrix} 5 & -2 & -11 & 1 \\ 0 & 6 & -2 & 2 \\ 0 & 8 & 4 & -4 \\ 0 & 0 & 3 & 7 \end{pmatrix}.$$

سپس با استفاده از درایهٔ $a_{22} = 6$، درایهٔ $a_{32} = 8$ را صفر می‌کنیم. برای این کار، داریم $c = 0.6$ و $s = 0.8$. از این‌رو داریم:

$$A := J_{23}A = \begin{pmatrix} 5 & -2 & -11 & 1 \\ 0 & 10 & 2 & -2 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & 3 & 7 \end{pmatrix}.$$

سرانجام با استفاده از درایهٔ $a_{33} = 4$، درایهٔ $a_{43} = 3$ را صفر می‌کنیم. در اینجا داریم $c = 0.8$ و $s = 0.6$. در نتیجه

$$A := J_{34}A = \begin{pmatrix} 5 & -2 & -11 & 1 \\ 0 & 10 & 2 & -2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

همچنین برای محاسبهٔ Q داریم:

$$Q = J_{12}^T J_{23}^T J_{34}^T = \frac{1}{125}\begin{pmatrix} 75 & -60 & 64 & -48 \\ 100 & 45 & -48 & 36 \\ 0 & 100 & 60 & -45 \\ 0 & 0 & 75 & 100 \end{pmatrix}.$$

با استفاده از ماتریس‌های گیونز نیز می‌توان قضیهٔ ۱۷.۴ را ثابت کرد. در حقیقت با کمک ماتریس‌های گیونز یک ماتریس متعامد مثل P می‌سازیم به طوری که $PAP^T = H$، که در آن ماتریس H بالاهسنبرگی است. در گام اول ماتریس‌های گیونز J_{23}، J_{24}، $\cdots$، J_{2n} را طوری می‌سازیم که درایه‌های زیر مکان $(2, 1)$ ماتریس

$$A^{(1)} = J_{2n} \cdots J_{24} J_{23}\ A\ J_{23}^T J_{24}^T \cdots J_{2n}^T = P_1 A P_1^T,$$

برابر با صفر باشند، که در آن ماتریس $P_۱$ به صورت $P_۱ = J_{۲n} \cdots J_{۲۴} J_{۲۳}$ تعریف می‌شود. در گام k ام $(۲ \leqslant k \leqslant n-۲)$ ماتریس‌های گیونز $J_{k+۱,k+۲}$، $J_{k+۱,k+۳}$، …، $J_{k+۱,n}$ را می‌سازیم به طوری که درایه‌های زیر مکان $(k+۱, k)$ از ماتریس

$$\begin{aligned} A^{(k)} &= J_{k+۱,n} \cdots J_{k+۱,k+۳} J_{k+۱,k+۲}\ A^{(k-۱)}\ J^T_{k+۱,k+۲} J^T_{k+۱,k+۳} \cdots J^T_{k+۱,n} \\ &= P_k A^{(k-۱)} P_k^T, \end{aligned}$$

برابر با صفر باشند، که در آن $P_k = J_{k+۱,n} \cdots J_{k+۱,k+۳} J_{k+۱,k+۲}$. بعد از $n-۲$، مرحله داریم:

$$\begin{aligned} H &= A^{(n-۲)} = P_{n-۲} A^{(n-۳)} P^T_{n-۲} \\ &= P_{n-۲} P_{n-۳} A^{(n-۴)} P^T_{n-۳} P^T_{n-۲} \\ &= P_{n-۲} P_{n-۳} \cdots P_۱ A P_۱^T \cdots P^T_{n-۳} P^T_{n-۲}. \end{aligned}$$

با قرار دادن $P = P_{n-۲} P_{n-۳} \cdots P_۱$ داریم $H = PAP^T$ که در آن ماتریس H یک ماتریس بالاهسنبرگی است. همان‌طور که قبلاً دیدیم اگر ماتریس A متقارن باشد، آنگاه H سه‌قطری است.

مثال ۲۱.۴ با استفاده از ماتریس‌های گیونز یک ماتریس متعامد مثل P بیابید به طوری که $T = PAP^T$ که در آن T یک ماتریس سه‌قطری است و

$$A = \begin{pmatrix} ۱ & ۲ & ۱ & -۱ \\ ۲ & ۱ & ۲ & ۱ \\ ۱ & ۲ & ۳ & ۲ \\ -۱ & ۱ & ۲ & ۱ \end{pmatrix}.$$

حل: داریم $a_{۲۱} = ۲$ و $a_{۳۱} = ۱$. لذا $c = ۰.۸۹۴۴$ و $s = ۰.۴۴۷۲$. با ساختن ماتریس $J_{۲۳}$، داریم:

$$A := J_{۲۳} A J^T_{۲۳} = \begin{pmatrix} ۱.۰۰۰۰ & ۲.۲۳۶۱ & ۰.۰۰۰۰ & -۱.۰۰۰۰ \\ ۲.۲۳۶۱ & ۳.۰۰۰۰ & ۲.۰۰۰۰ & ۱.۷۸۸۹ \\ ۰.۰۰۰۰ & ۲.۰۰۰۰ & ۱.۰۰۰۰ & ۱.۳۴۱۶ \\ -۱.۰۰۰۰ & ۱.۷۸۸۹ & ۱.۳۴۱۶ & ۱.۰۰۰۰ \end{pmatrix}.$$

به همین ترتیب با توجه اینکه $a_{21} = 2.2361$ و $a_{41} = -1.0000$ داریم $c = 0.9129$ و $s = -0.4082$. در نتیجه با استفاده از ماتریس J_{24}،

$$A := J_{24} A J_{24}^T = \begin{pmatrix} 1.0000 & 2.4495 & 0.0000 & 0.0000 \\ 2.4495 & 1.3333 & 1.2780 & 1.9379 \\ 0.0000 & 1.2780 & 1.0000 & 2.0412 \\ 0.0000 & 1.9379 & 2.0412 & 2.6667 \end{pmatrix}.$$

سرانجام داریم $a_{32} = 1.2780$ و $a_{42} = 1.9379$. لذا $c = 0.5505$ و $s = 0.8348$.
بنابراین با استفاده از ماتریس J_{34} می‌بینیم که

$$A := J_{34} A J_{34}^T = \begin{pmatrix} 1.0000 & 2.4495 & 0.0000 & 0.0000 \\ 2.4495 & 1.3333 & 2.3214 & 0.0000 \\ 0.0000 & 2.3214 & 4.0378 & -0.0379 \\ 0.0000 & 0.0000 & -0.0379 & -0.3711 \end{pmatrix} = T.$$

برای محاسبهٔ ماتریس P قرار می‌دهیم $P_1 = J_{24} J_{23}$ و $P_2 = J_{34}$. لذا

$$P \ := \ P_2 P_1 = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.8165 & 0.4082 & -0.4082 \\ 0.0000 & 0.0586 & 0.6448 & 0.7621 \\ 0.0000 & 0.5744 & -0.6462 & 0.5026 \end{pmatrix}.$$

در پایان این بخش، تجزیهٔ QR یک ماتریس $m \times n$ با استفاده از فرایند گرام- اشمیت، ماتریس‌های هوسهولدر و گیونز را از لحاظ تعداد اعمال حسابی و پایداری با یکدیگر مقایسه می‌کنیم. در جدول ۱.۴ مقدار تقریبی تعداد اعمال‌حسابی این روش‌ها و پایداری یا عدم پایداری آنها در حالت کلی ارائه شده است (برای جزئیات بیشتر به [۳] مراجعه کنید).

جدول ۱.۴

روش	تعداد تقریبی اعمال حسابی	پایداری
گرام — اشمیت	$2mn^2$	ناپایدار
هوسهولدر	$2n^2(m - \frac{n}{3})$	پایدار
گیونز	$4n^2(m - \frac{n}{3})$	پایدار

همان‌طور که جدول ۱.۴ نشان می‌دهد الگوریتم گرام- اشمیت در حالت کلی ناپایدار است. همچنین نشان می‌دهد که تجزیهٔ QR با استفاده از ماتریس‌های هوسهولدر و گیونز پایدارند. از لحاظ محاسباتی تعداد اعمال حسابی با استفاده از ماتریس‌های گیونز تقریباً دو برابر تعداد اعمال حسابی با استفاده از ماتریس‌های هوسهولدر است. لازم به ذکر است که برای بعضی از ماتریس‌هایی که ساختار خاص دارند ممکن است استفاده از ماتریس‌های گیونز مناسب‌تر باشد. به عنوان مثال، برای محاسبهٔ تجزیهٔ QR یک ماتریس $n \times n$ بالا هسنبرگی، کافی است $n - ۱$ درایهٔ زیر قطر اصلی را با استفاده از ماتریس‌های گیونز صفر کنیم (مثال ۲۰.۴ را به خاطر بیاورید).

در ادامه چند روش عددی برای محاسبهٔ زوج‌های ویژهٔ یک ماتریس را مطالعه می‌کنیم. اغلب روش‌های مورد بررسی روش‌های تکراری هستند. بعضی از این روش‌ها یک یا چند زوج ویژهٔ یک ماتریس و بعضی از آنها تمام زوج‌های ویژهٔ ماتریس را محاسبه می‌کنند.

۵.۴ روش تکراری QR

این روش برای محاسبهٔ تمام مقادیر ویژهٔ یک ماتریس به کار می‌رود و یک روش تکراری است. در این بخش این روش و یک نسخهٔ بهبود یافتهٔ آن را مطالعه می‌کنیم.

۱.۵.۴ روش تکراری QR پایه‌ای

فرض کنید $A \in \mathbb{R}^{n\times n}$. در این روش ابتدا قرار می‌دهیم $A_۰ = A$. در تکرار k ام ($k \geqslant ۱$)، تجزیهٔ QR ماتریس $A_{k-۱}$ را محاسبه می‌کنیم، یعنی

$$A_{k-۱} = Q_{k-۱}R_{k-۱}. \qquad (۲۱.۴)$$

سپس قرار می‌دهیم:

$$A_k = R_{k-۱}Q_{k-۱}. \qquad (۲۲.۴)$$

به این ترتیب دنباله‌ای مثل $\{A_k\}$ ساخته می‌شود. یکی از خواص جالب این دنباله این است که هر جمله از این دنباله به طور متعامد باجمله‌های دیگرِ دنباله متشابه است. زیرا

از (۲۱.۴)، داریم:

$$Q_{k-1}^T A_{k-1} = R_{k-1}.$$

با جایگذاری R_{k-1} در رابطهٔ (۲۲.۴) خواهیم داشت:

$$A_k = Q_{k-1}^T A_{k-1} Q_{k-1}.$$

این رابطه نشان می‌دهد که $A_k \sim A_{k-1}$. به همین ترتیب می‌توان دید که

$$A_k \sim A_{k-1} \sim \cdots \sim A_0 = A.$$

بنابراین مقادیر ویژهٔ ماتریس A با مقادیر ویژهٔ هریک از جملات دنباله برابر است. لذا اگر این دنباله به یک ماتریس بالامثلثی یا نیمه‌بالامثلثی همگرا شود، آنگاه مقادیر ویژهٔ ماتریس A با محاسبهٔ مقادیر ویژهٔ این ماتریس‌ها قابل محاسبه هستند. خوشبختانه همگرایی این دنباله تحت شرایطی تضمین می‌شود که در قضیهٔ بعد بیان می‌شوند.

قضیه ۱۸.۴ اگر مقادیر ویژهٔ A در رابطهٔ

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|, \qquad (۲۳.۴)$$

صدق کنند، آنگاه دنبالهٔ $\{A_k\}$ که در روش تکراری QR پایه‌ای تعریف می‌شود به یک ماتریس بالامثلثی همگرا می‌شود.

برهان: به [۱۱] مراجعه کنید. □

مثال ۲۲.۴ با استفاده از روش تکراری QR پایه‌ای، تقریبی از مقادیر ویژهٔ ماتریس

$$A = \begin{pmatrix} 5 & 1 & 2 \\ -6 & 6 & 2 \\ 6 & -3 & 1 \end{pmatrix},$$

را به دست آورید.

حل: مقادیر ویژهٔ A عبارتند از $\lambda_۱ = ۷$، $\lambda_۲ = ۳$ و $\lambda_۳ = ۲$. قرار می‌دهیم $A_۰ = A$. در تکرار اول، داریم $A_۰ = Q_۰R_۰$ به طوری که

$$Q_۰ = \begin{pmatrix} -۰.۵۰۷۷ & -۰.۷۶۴۹ & -۰.۳۹۶۵ \\ ۰.۶۰۹۲ & -۰.۶۴۴۱ & ۰.۴۶۲۶ \\ -۰.۶۰۹۲ & -۰.۰۰۶۷ & ۰.۷۹۳۰ \end{pmatrix},$$

$$R_۰ = \begin{pmatrix} -۹.۸۴۸۹ & ۴.۹۷۵۲ & -۰.۴۰۶۱ \\ ۰ & -۴.۶۰۹۵ & -۲.۸۲۴۷ \\ ۰ & ۰ & ۰.۹۲۵۱ \end{pmatrix},$$

و

$$A_۱ = R_۰Q_۰ = \begin{pmatrix} ۸.۲۷۸۴ & ۴.۳۳۱۴ & ۵.۸۸۴۳ \\ -۱.۰۸۷۳ & ۲.۹۸۸۰ & -۴.۳۷۲۲ \\ -۰.۵۶۳۶ & -۰.۰۰۶۲ & ۰.۷۳۳۶ \end{pmatrix}.$$

در تکرار دوم داریم $A_۱ = Q_۱R_۱$، که در آن

$$Q_۱ = \begin{pmatrix} -۰.۹۸۹۲ & -۰.۱۳۴۷ & ۰.۰۵۷۱ \\ ۰.۱۲۹۹ & -۰.۹۸۸۲ & -۰.۰۸۰۸ \\ ۰.۰۶۷۳ & -۰.۰۷۲۵ & ۰.۹۹۵۱ \end{pmatrix},$$

$$R_۱ = \begin{pmatrix} -۸.۳۶۸۴ & -۳.۸۹۷۰ & -۶.۳۳۹۶ \\ ۰ & -۳.۵۳۶۰ & ۳.۴۷۴۸ \\ ۰ & ۰ & ۱.۴۱۹۴ \end{pmatrix},$$

و

$$A_۲ = R_۱Q_۱ = \begin{pmatrix} ۷.۳۴۵۱ & ۵.۴۳۸۰ & -۶.۴۷۲۰ \\ -۰.۲۲۵۴ & ۳.۲۴۲۵ & ۳.۷۴۳۳ \\ ۰.۰۹۵۶ & -۰.۱۰۲۹ & ۱.۴۱۲۴ \end{pmatrix}.$$

در تکرار سوم داریم $A_۲ = Q_۲R_۲$، به طوری که

$$Q_۲ = \begin{pmatrix} -۰.۹۹۹۴ & -۰.۰۳۱۳ & -۰.۰۱۱۴ \\ ۰.۰۳۰۷ & -۰.۹۹۸۲ & ۰.۰۵۰۹ \\ -۰.۰۱۳۰ & ۰.۰۵۰۵ & ۰.۹۹۸۶ \end{pmatrix},$$

$$R_۲ = \begin{pmatrix} -۷.۳۴۹۱ & -۵.۳۳۴۲ & ۶.۵۶۴۸ \\ ۰ & -۳.۴۱۲۲ & -۳.۴۶۲۹ \\ ۰ & ۰ & ۱.۶۷۴۹ \end{pmatrix},$$

و

$$A_{۳} = R_{۲}Q_{۲} = \begin{pmatrix} ۷.۰۹۶۱ & ۵.۸۸۶۱ & ۶.۳۶۸۶ \\ -۰.۰۵۹۶ & ۳.۲۳۱۳ & -۳.۶۳۱۷ \\ -۰.۰۲۱۸ & ۰.۰۸۴۵ & ۱.۶۷۲۶ \end{pmatrix}.$$

با ادامهٔ این روند در تکرار دهم، خواهیم داشت:

$$A_{۱۰} = \begin{pmatrix} ۷.۰۰۰۰ & ۶.۴۲۵۷ & -۵.۹۳۸۵ \\ ۰.۰۰۰۰ & ۳.۰۱۴۹ & ۳.۶۶۲۶ \\ ۰.۰۰۰۰ & ۰.۰۰۴۱ & ۱.۹۸۵۱ \end{pmatrix}.$$

همان‌طور که ملاحظه می‌شود درایه‌های قطری این ماتریس تقریب‌های خوبی از مقادیر ویژهٔ A و درایه‌های زیر قطر اصلی از حیث قدر مطلق کوچک هستند.

۲.۵.۴ روش تکراری QR- هسنبرگ

یکی از مشکلات روش تکراری QR این است که در هر تکرار، تجزیهٔ QR یک ماتریس پُر محاسبه می‌گردد که بسیار پرهزینه است. خوشبختانه این مشکل را می‌توان تا حدودی حل نمود. برای این کار، ابتدا ماتریس A را با یک ماتریس بالاهسنبرگی متشابه می‌کنیم (قضیهٔ ۱۷.۴ را ببینید). سپس روش QR را برای ماتریس بالاهسنبرگی به دست آمده به کار می‌بریم. این روش به نام روش QR- هسنبرگ معروف است. به سادگی می‌توان دید که اگر ماتریس A بالاهسنبرگی باشد، آنگاه تمام A_kها بالاهسنبرگی خواهند بود. لازم به ذکر است تجزیهٔ QR یک ماتریس بالاهسنبرگی به سادگی قابل محاسبه است، زیرا کافی است $n-۱$ درایهٔ $(i, i-۱)$، $i = ۲, \dots, n$ را با استفاده از ماتریس‌های گیونز صفر کنیم. برای روشن شدن مطلب مثال بعدی را ارائه می‌کنیم.

مثال ۲۳.۴ مقادیر ویژهٔ ماتریس

$$A = \begin{pmatrix} ۱۱ & -۴ & ۴ & -۴ \\ -۲ & ۱ & -۲ & ۲ \\ -۱۲ & ۱۲ & ۳ & ۱۲ \\ -۲ & ۱۰ & ۶ & ۹ \end{pmatrix},$$

را با استفاده از روش QR- هسنبرگ با حداکثر خطای ۵×۱۰^{-۳} به دست آورید.

حل: همانند مثال ۲۱.۴ ابتدا با استفاده از ماتریس‌های گیونز این ماتریس را با ماتریس بالاهسنبرگی

$$H = \begin{pmatrix} 11.0000 & -2.5955 & 2.8110 & -5.7760 \\ 12.3288 & 7.8421 & 7.7519 & -15.9287 \\ 0.0000 & 8.1645 & 2.9239 & -8.0627 \\ 0.0000 & 0.0000 & -1.5739 & 2.2340 \end{pmatrix},$$

متشابه می‌کنیم. حال روش QR پایه‌ای را برای این ماتریس به کار می‌بریم. قرار می‌دهیم $H_0 = H$. در این صورت، در تکرار اول داریم $H_0 = Q_0 R_0$، که در آن

$$Q_0 = \begin{pmatrix} 0.6658 & -0.4919 & -0.1304 & 0.5457 \\ 0.7462 & 0.4389 & 0.1163 & -0.4869 \\ 0 & 0.7520 & -0.1532 & 0.6412 \\ 0 & 0 & -0.9726 & -0.2324 \end{pmatrix},$$

$$R_0 = \begin{pmatrix} 16.5227 & 4.1236 & 7.6557 & -15.7310 \\ 0.0000 & 10.8577 & 4.2180 & -10.2124 \\ 0.0000 & 0 & 1.6182 & -2.0377 \\ 0.0000 & 0 & 0 & -1.0851 \end{pmatrix},$$

و

$$H_1 = R_0 Q_0 = \begin{pmatrix} 14.0769 & -0.5609 & 12.4525 & 15.5736 \\ 8.1018 & 7.9370 & 10.5498 & -0.2088 \\ 0.0000 & 1.2168 & 1.7340 & 1.5111 \\ 0.0000 & -0.0000 & 1.0554 & 0.2522 \end{pmatrix}.$$

در تکرار دوم داریم $H_1 = Q_1 R_1$، به طوری که

$$Q_1 = \begin{pmatrix} 0.8667 & -0.4918 & 0.0632 & -0.0547 \\ 0.4988 & 0.8545 & -0.1098 & 0.0951 \\ 0 & 0.1676 & 0.7451 & -0.6455 \\ 0 & 0 & 0.6548 & 0.7558 \end{pmatrix},$$

$$R_1 = \begin{pmatrix} 16.2419 & 3.4730 & 16.0551 & 13.3935 \\ -0.0000 & 7.2615 & 3.1811 & -7.5838 \\ 0.0000 & -0.0000 & 1.6117 & 2.2978 \\ -0.0000 & -0.0000 & 0.0000 & -1.6572 \end{pmatrix},$$

و

$$H_{۲} = R_{۱}Q_{۱} = \begin{pmatrix} ۱۵.۸۰۹۳ & -۲.۳۲۹۴ & ۲۱.۳۷۷۹ & -۰.۸۰۰۳ \\ ۳.۶۲۲۲ & ۶.۷۳۷۶ & -۳.۳۹۲۸ & -۷.۰۹۴۸ \\ ۰.۰۰۰۰ & ۰.۲۷۰۱ & ۲.۷۰۵۵ & ۰.۶۹۶۳ \\ -۰.۰۰۰۰ & -۰.۰۰۰۰ & -۱.۰۸۵۱ & -۱.۲۵۲۵ \end{pmatrix}.$$

در تکرار سوم خواهیم داشت $H_{۲} = Q_{۲}R_{۲}$، که در آن

$$Q_{۲} = \begin{pmatrix} ۰.۹۷۴۷ & -۰.۲۲۳۲ & ۰.۰۰۸۰ & ۰.۰۰۲۹ \\ ۰.۲۲۳۳ & ۰.۹۷۴۰ & -۰.۰۳۴۹ & -۰.۰۱۲۶ \\ ۰ & ۰.۰۳۸۱ & ۰.۹۴۰۱ & ۰.۳۳۸۸ \\ ۰ & ۰ & -۰.۳۳۹۰ & ۰.۹۴۰۸ \end{pmatrix}$$

$$R_{۲} = \begin{pmatrix} ۱۶.۲۱۹۰ & -۰.۷۶۵۸ & ۲۰.۰۸۰۲ & -۲.۳۶۴۵ \\ ۰.۰۰۰۰ & ۷.۰۹۲۸ & -۷.۹۷۲۵ & -۶.۷۰۵۴ \\ ۰.۰۰۰۰ & ۰.۰۰۰۰ & ۳.۲۰۰۸ & ۱.۳۲۰۵ \\ -۰.۰۰۰۰ & -۰.۰۰۰۰ & ۰ & -۰.۸۵۵۵ \end{pmatrix},$$

و

$$H_{۳} = R_{۲}Q_{۲} = \begin{pmatrix} ۱۵.۶۳۸۳ & -۳.۶۰۰۹ & ۱۹.۸۳۵۵ & ۴.۶۳۴۳ \\ ۱.۵۸۴۰ & ۶.۶۰۵۱ & -۵.۴۶۹۴ & -۹.۰۹۸۴ \\ ۰.۰۰۰۰ & ۰.۱۲۱۹ & ۲.۵۶۱۵ & ۲.۳۲۶۶ \\ -۰.۰۰۰۰ & -۰.۰۰۰۰ & ۰.۲۹۰۰ & -۰.۸۰۴۸ \end{pmatrix}.$$

با ادامهٔ این روند در تکرار دهم، خواهیم داشت:

$$H_{۱۰} = \begin{pmatrix} ۱۵.۰۰۱۸ & -۴.۶۱۶۲ & ۱۸.۷۸۳۲ & ۱.۴۱۸۰ \\ ۰.۰۰۳۲ & ۶.۹۹۸۵ & -۹.۸۹۱۸ & -۸.۹۸۱۱ \\ ۰.۰۰۰۰ & ۰.۰۰۰۱ & ۲.۹۹۹۷ & ۲.۳۰۹۲ \\ -۰.۰۰۰۰ & -۰.۰۰۰۰ & ۰.۰۰۰۱ & -۱.۰۰۰۰ \end{pmatrix}.$$

بزرگترین درایه در قسمت پایین مثلثی ماتریس $H_{۱۰}$ برابر با ۰.۰۰۳۲ می‌باشد که از تولرانس داده شده کوچکتر است. بنابراین مقادیر ویژهٔ تقریبی ماتریس داده شده عبارتند از ۱۵.۰۰۱۸، ۶.۹۹۸۵، ۲.۹۹۹۷ و ۱.۰۰۰۰−، که به ترتیب تقریبی از مقادیر ویژهٔ $\lambda_{۱} = ۱۵$، $\lambda_{۲} = ۷$، $\lambda_{۳} = ۳$ و $\lambda_{۴} = -۱$ می‌باشند.

۶.۴ مقادیر ویژهٔ ماتریس‌های توپیلیتز سه‌قطری متقارن

یک ماتریس توپیلیتز ماتریسی است که روی هر قطر موازی با قطر اصلی ماتریس درایه‌های یکسانی قرار داشته باشد. برای مثال، ماتریس

$$T = \begin{pmatrix} a & b & c & d \\ e & a & b & c \\ f & e & a & b \\ g & f & e & a \end{pmatrix},$$

یک ماتریس توپیلیتز است. یکی از ماتریس‌هایی که به وفور در آنالیز پایداری روش تفاضلات متناهی برای حل معادلات دیفرانسیل ظاهر می‌شود، ماتریس توپیلیتز سه‌قطری متقارن، یعنی

$$T = \begin{pmatrix} b & a & & & \\ c & b & a & & \\ & \ddots & \ddots & \ddots & \\ & & c & b & a \\ & & & c & b \end{pmatrix} \in \mathbb{R}^{n\times n},$$

است که در آن a، b و c اعدادی حقیقی هستند و $ac > ۰$. در اینجا نشان می‌دهیم که زوج‌های ویژهٔ این ماتریس به سادگی محاسبه می‌شوند. فرض کنید (λ, x) یک زوج ویژهٔ T باشد و $x = (x_۱, x_۲, \ldots, x_n)^T$. در این صورت، $(T - \lambda I)x = ۰$ معادل با

$$cx_{k-۱} + (b-\lambda)x_k + ax_{k+۱} = ۰, \quad k = ۱, ۲, \ldots, n,$$

است که در آن $x_۰ = x_{n+۱} = ۰$. یا

$$\begin{cases} x_۰ = x_{n+۱} = ۰, \\ x_{k+۱} + (\dfrac{b-\lambda}{a})x_k + \dfrac{c}{a}x_{k-۱} = ۰, \quad k = ۱, ۲, \ldots, n. \end{cases} \qquad (۲۴.۴)$$

همان‌طور که می‌دانیم جواب‌هایی به صورت $x_k = \xi r^k$ برای این معادلات جستجو می‌شوند. در نتیجه با قراردادن x_k در رابطهٔ (۲۴.۴) معادلهٔ درجهٔ دوم زیر به دست می‌آید:

$$r^۲ + (\frac{b-\lambda}{a})r + \frac{c}{a} = ۰.$$

فرض کنید r_1 و r_2 ریشه‌های این معادله باشند. در این صورت، جواب عمومی رابطهٔ (۲۴.۴) به صورت زیر خواهد بود:

$$x_k = \begin{cases} \alpha r_1^k + \beta r_2^k, & r_1 \neq r_2, \\ \alpha \rho^k + \beta k \rho^k, & r_1 = r_2 = \rho. \end{cases}$$

ابتدا حالت $r_1 = r_2 = \rho$ را در نظر می‌گیریم. در این حالت، از اینکه $x_0 = x_{n+1} = 0$ داریم

$$\alpha = \alpha \rho^{n+1} + \beta(n+1)\rho^{n+1} = 0.$$

از این‌رو داریم $\alpha = \beta = 0$ و در نتیجه $x_k = 0$، $k = 1, 2, \ldots, n$. این نتیجه نمی‌تواند برقرار باشد، زیرا x یک بردار ویژه است. لذا $r_1 \neq r_2$. با توجه به اینکه $x_0 = x_{n+1} = 0$، داریم:

$$\begin{cases} \alpha + \beta = 0, \\ \alpha r_1^{n+1} + \beta r_2^{n+1} = 0. \end{cases}$$

از این دو رابطه داریم:

$$(\frac{r_1}{r_2})^{n+1} = -\frac{\beta}{\alpha} = 1.$$

در نتیجه

$$\frac{r_1}{r_2} = e^{i\frac{2\pi j}{n+1}}, \quad j = 1, 2, \ldots, n,$$

که در آن i واحد موهومی است. از طرفی داریم:

$$r^2 + (\frac{b-\lambda}{a})r + \frac{c}{a} = (r - r_1)(r - r_2) \Rightarrow \begin{cases} r_1 r_2 = \frac{c}{a}, \\ r_1 + r_2 = -\frac{b-\lambda}{a}. \end{cases}$$

بنابراین

$$r_1 = \sqrt{\frac{c}{a}} e^{i\frac{\pi j}{n+1}}, \quad r_2 = \sqrt{\frac{c}{a}} e^{-i\frac{\pi j}{n+1}}.$$

حال داریم:

$$\begin{aligned} \lambda &= b + a(r_1 + r_2) = b + a\sqrt{\frac{c}{a}}(e^{i\frac{\pi j}{n+1}} + e^{-i\frac{\pi j}{n+1}}) \\ &= b + 2a\sqrt{\frac{c}{a}} \cos(\frac{j\pi}{n+1}). \end{aligned}$$

بنابراین مقادیر ویژهٔ T به صورت زیر هستند:

$$\lambda_j = b + 2a\sqrt{\frac{c}{a}}\cos(\frac{j\pi}{n+1}), \quad j = 1, 2, \ldots, n. \qquad (25.4)$$

با توجه به اینکه $ac > 0$ و اعداد $\cos(\frac{j\pi}{n+1})$، $j = 1, 2, \ldots, n$ متمایزند نتیجه می‌گیریم که مقادیر ویژهٔ T نیز متمایزند و T قطری‌شدنی است. اما برای محاسبهٔ بردارهای ویژهٔ T، با توجه به اینکه $\alpha + \beta = 0$، داریم:

$$\begin{aligned} x_k &= \alpha r_1^k + \beta r_2^k = \alpha(r_1^k - r_2^k) \\ &= \alpha(\frac{c}{a})^{\frac{k}{2}}(e^{i\frac{\pi jk}{n+1}} - e^{-i\frac{\pi jk}{n+1}}) = 2i\alpha(\frac{c}{a})^{\frac{k}{2}}\sin(\frac{jk\pi}{n+1}). \end{aligned}$$

با توجه به اینکه α دلخواه است، قرار می‌دهیم $\alpha = 1/(2i)$. در این صورت

$$\begin{pmatrix} (\frac{c}{a})^{\frac{1}{2}}\sin(\frac{1j\pi}{n+1}) \\ (\frac{c}{a})^{\frac{2}{2}}\sin(\frac{2j\pi}{n+1}) \\ (\frac{c}{a})^{\frac{3}{2}}\sin(\frac{3j\pi}{n+1}) \\ \vdots \\ (\frac{c}{a})^{\frac{n}{2}}\sin(\frac{nj\pi}{n+1}) \end{pmatrix}, \quad j = 1, 2, \ldots, n,$$

بردارهای ویژهٔ T هستند.

مثال ۲۴.۴ اگر

$$T = \begin{pmatrix} 4 & -1 & & & \\ -4 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -4 & 4 & -1 \\ & & & -4 & 4 \end{pmatrix} \in \mathbb{R}^{10\times 10},$$

آنگاه زوج‌های ویژهٔ ماتریس $2I + 5T^2$ را محاسبه کنید.

حل: اگر (λ_j, x_j) یک زوج ویژهٔ ماتریس T باشد، آنگاه داریم:

$$\begin{aligned} (2I + 5T^2)x_j &= 2x_j + 5T^2x_j = 2x + 5\lambda_j^2 x_j \\ &= (2 + 5\lambda_j^2)x_j. \end{aligned}$$

این رابطه نشان می‌دهد که $(2+5\lambda_j^2, x_j)$ یک زوج ویژهٔ $2I+5T^2$ می باشد. از این‌رو با استفاده از رابطهٔ (۲۵.۴) مقادیر ویژهٔ $2I+5T^2$ عبارتند از

$$\begin{aligned}\mu_j &= 2+5\left(4+2\times(-1)\times\sqrt{\frac{-4}{-1}}\cos\frac{j\pi}{11}\right)^2\\ &= 4+90\sin^4\frac{j\pi}{22},\end{aligned}$$

به ازای $j=1,2,\dots,10$ و بردار ویژهٔ متناظر با μ_j به صورت زیر است:

$$x_j=\begin{pmatrix}4^{\frac{1}{2}}\sin(\frac{1j\pi}{11})\\ 4^{\frac{2}{2}}\sin(\frac{2j\pi}{11})\\ 4^{\frac{3}{2}}\sin(\frac{3j\pi}{11})\\ \vdots\\ 4^{\frac{11}{2}}\sin(\frac{10j\pi}{11})\end{pmatrix},\qquad j=1,2,\dots,10.$$

تمرینات

۱.۴) زوج‌های ویژهٔ ماتریس

$$A=\begin{pmatrix}-3&1&-3\\20&3&10\\2&-2&4\end{pmatrix},$$

را بیابید. آیا ماتریس A قطری‌شدنی است؟

۲.۴) چندجمله‌ای مشخصهٔ ماتریس زیر را نوشته و با استفاده از آن معکوس ماتریس A را به دست آورید.

$$A=\begin{pmatrix}4&5\\3&2\end{pmatrix}.$$

۳.۴) اگر

$$A=\begin{pmatrix}\frac{7}{5}&\frac{1}{5}\\-1&\frac{1}{2}\end{pmatrix},$$

آنگاه $\lim_{n\to\infty}A^n$ را محاسبه کنید.

۴.۴) نشان دهید که ماتریس $A=(a_{ij})\in\mathbb{R}^{n\times n}$ با درایه‌های
$$a_{ij}=\begin{cases} n, & i=j,\\ ۱, & i\neq j,\end{cases}$$
معین مثبت متقارن است.

۵.۴) زوج ویژهٔ غالب ماتریس
$$A=\begin{pmatrix} ۲ & ۱ & ۱ & ۰\\ ۱ & ۱ & ۰ & ۱\\ ۱ & ۰ & ۱ & ۱\\ ۰ & ۱ & ۱ & ۲\end{pmatrix},$$
را با دقت $۱۰^{-۳}$ با استفاده از روش توانی به دست آورید. به روش توانی معکوس، تقریبی از یک مقدار ویژهٔ ماتریس A با پارامتر انتقال $\sigma=۱.۷۵$ و با دقت $۱۰^{-۳}$ به دست آورید.

۶.۴) تجزیهٔ QR ماتریس زیر را بدست آورید:
$$A=\begin{pmatrix} -۱۲ & -۱۲ & ۰\\ ۴ & ۲۵ & ۳\\ -۳ & ۲۵ & ۴\end{pmatrix}.$$

۷.۴) تجزیهٔ QR ماتریس‌های زیر را بدست آورید:
$$A=\begin{pmatrix} ۴ & ۱ & ۳\\ ۳ & -۲ & ۱\end{pmatrix},\qquad B=\begin{pmatrix} ۳ & ۰\\ ۴ & -۵\\ ۰ & -۴\end{pmatrix}.$$

۸.۴) با استفاده از روش تکراری LU تقریبی از مقادیر ویژهٔ ماتریس
$$A=\begin{pmatrix} ۳ & ۱\\ ۱ & ۱\end{pmatrix},$$
را با دقت $۱۰^{-۳}$ به دست آورید.

۹.۴) به کمک روش QR- پایه‌ای مقادیر ویژهٔ ماتریس زیر را با دقت $۱۰^{-۳}$ به دست آورید.
$$A=\begin{pmatrix} ۱ & ۲ & ۳\\ ۲ & ۳ & ۴\\ ۳ & ۴ & ۵\end{pmatrix}.$$

۱۰.۴) ماتریس زیر را با استفاده از ماتریس‌های هوسهولدر با یک ماتریس سه‌قطری متشابه کنید.

$$A = \begin{pmatrix} ۴ & -۱ & -۲ & ۲ \\ -۱ & ۴ & -۱ & -۲ \\ -۲ & -۱ & ۴ & -۱ \\ ۲ & -۲ & -۱ & ۴ \end{pmatrix}.$$

۱۱.۴) با استفاده از قضیهٔ دیسک‌های گرشگورین حدودی برای مجموعه مقادیر ویژهٔ ماتریس $T = \text{tridiag}_n(۱, -۲, ۱)$ بیابید.

۱۲.۴) نشان دهید برای هر ماتریس نامنفرد $A \in \mathbb{R}^{n\times n}$ یک چندجمله‌ای حداکثر از درجهٔ $n-۱$ مثل P وجود دارد به طوری $A^{-۱} = P(A)$.

۱۳.۴) دستگاه معادلات دیفرانسیل

$$\begin{cases} x'_۱(t) = ۳x_۱(t) + x_۳(t), \\ x'_۲(t) = -۳x_۲(t), \\ x'_۳(t) = x_۱(t) + ۳x_۳(t), \end{cases}$$

با شرط اولیهٔ $(x_۱(۰), x_۲(۰), x_۳(۰)) = (۰, ۲, -۲)$ را حل کنبد.

۱۴.۴) فرض کنید $A, B \in \mathbb{R}^{m\times n}$. نشان دهید $\sigma(AB) = \sigma(BA)$ و از آنجا نتیجه بگیرید که $\rho(AB) = \rho(BA)$. فرض کنید $A \in \mathbb{C}^{m\times n}$. نشان دهید $\|A^T\|_۲ = \|A\|_۲$.

۱۵.۴) نشان دهید مقادیر ویژهٔ ماتریس‌های هرمیتی‌کج موهومی محض می‌باشند، یعنی قسمت حقیقی مقادیر ویژه برابر با صفرند.

۱۶.۴) فرض کنید $u, v \in \mathbb{R}^n$. مقادیر ویژه و بردارهای ویژهٔ ماتریس $A = I + uv^T$ را به دست آورید.

۱۷.۴) فرض کنید $A \in \mathbb{R}^{n\times n}$ و μ_i، $i = ۱, \ldots, n$ مقادیر ویژهٔ A^TA باشند. نشان دهید $\mu_i \geqslant ۰$، $i = ۱, \ldots, n$. سپس مقادیر ویژهٔ ماتریس

$$B = \begin{pmatrix} ۰ & -A^T \\ A & ۰ \end{pmatrix},$$

را بر حسب μ_i ها به دست آورید.

۱۸.۴) برای هر ماتریس $A \in \mathbb{R}^{n\times n}$، تعریف می‌کنیم:

$$\exp(A) = I + A + \frac{۱}{۲!}A^۲ + \cdots + \frac{۱}{m!}A^m + \cdots = \sum_{m=۰}^{\infty} \frac{۱}{m!}A^m.$$

فرض کنید:

$$A = \begin{pmatrix} ۲ & -۲ \\ -۱ & ۱ \end{pmatrix}.$$

یک ماتریس نامنفرد مثل P بیابید به طوری که $A = PDP^{-۱}$، که در آن D یک ماتریس قطری است. نشان دهید به ازای هر عدد صحیح نامنفی k، داریم $A^k = PD^kP^{-۱}$. سپس به کمک آن $\exp(A)$ را محاسبه کنید.

۱۹.۴) فرض کنید $T = \mathrm{tridiag}_n(-۱, ۲, -۱)$. نشان دهید به ازای هر $r > ۰$ ماتریس $۲I + rT$ نامنفرد است که در آن I ماتریس همانی می‌باشد. مقادیر ویژهٔ ماتریس

$$G = (۲I + rT)^{-۱}(۲I - rT),$$

را به دست آورید.

فصل ۵

روش‌های تکراری برای حل دستگاه معادلات خطی

دستگاه معادلات خطی

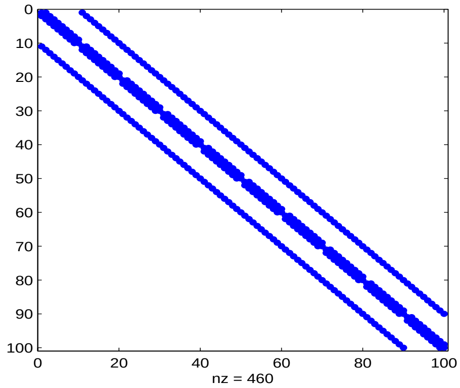
$$Ax = b, \tag{۱.۵}$$

را در نظر بگیرید که در آن $A \in \mathbb{R}^{n\times n}$ یک ماتریس نامنفرد، $b \in \mathbb{R}^n$ و $x \in \mathbb{R}^n$ یک بردار مجهول است. در فصل ۳ چند روش مستقیم برای حل دستگاه معادلات خطی (۱.۵) ارائه شد. از ایرادهای مهم روش‌های مستقیم این است که معمولاً حجم ذخیره‌سازی و انباشتگی خطا در به کارگیری آنها زیاد است. در روش‌هایی مثل حذفی گاوس یا تجزیهٔ LU ماتریس‌هایی که در طی فرایند روش تولید می‌شوند ممکن است ماتریس‌های پُری باشند، حتی اگر ماتریس ضرایب دستگاه تعداد درایه‌های ناصفر کمی داشته باشد. از طرفی با توجه به روند طولانی این گونه روش‌ها، خطاها روی هم انباشته شده و در نهایت ممکن است جوابی دور از انتظار به دست آید. برای رفع این ایرادات از روش‌های تکراری استفاده می‌شود. در این فصل برای درک بهتر این ایرادات مثال‌هایی را ارائه می‌کنیم، سپس چند روش تکراری برای حل دستگاه (۱.۵) را بررسی می‌کنیم. اغلب دستگاه‌هایی که در علوم و مهندسی ظاهر می‌شوند، بزرگ و تُنُک هستند. لذا جهت صرفه‌جویی در حافظه

کامپیوتری لازم است برای ذخیره‌سازی آنها از روش‌های مناسبی استفاده شود. در بخش بعدی مطالبی در این خصوص ذکر می‌شوند.

۱.۵ ماتریس‌های تُنُک و شیوه‌های ذخیره‌سازی آنها

ماتریس $A \in \mathbb{R}^{n\times n}$ را در نظر بگیرید. فرض می‌کنیم ماتریس A تُنُک باشد. لازم به ذکر است که یک ماتریس را تنک می‌گوییم هرگاه حداکثر پنج‌درصد از درایه‌های آن ناصفر باشند. در شکل ۱.۵ یک ماتریس تُنُکِ ۱۰۰ × ۱۰۰ نمایش داده شده است. این ماتریس تنها ۴۶۰ درایهٔ ناصفر دارد و بنابراین %۴.۶ از درایه‌های ماتریس ناصفرند.



شکل ۱.۵: نمایش یک ماتریس تُنُک.

در ادامه رایج‌ترین و ساده‌ترین شیوهٔ ذخیره‌سازی یک ماتریس در حافظهٔ کامپیوتر، یعنی قالب مختصاتی، را معرفی می‌کنیم. روش‌های دیگرِ ذخیره‌سازی را می‌توانید در [۱، ۱۰] بیابید. برای ذخیره‌سازی یک ماتریس تُنُک تنها درایه‌های ناصفر و اندیس‌های سطر و ستون آنها ذخیره می‌شوند. فرض کنید تعداد درایه‌های ناصفر ماتریس $A \in \mathbb{R}^{n\times n}$ برابر با nz باشد. در قالب مختصاتی، سه بردار ia، ja و a با طول nz برای ذخیره سازی A استفاده می‌شود. در بردار ia، اندیس سطری درایه‌ها، در بردار ja، اندیس ستونی درایه‌ها و

در بردار a، درایه‌های ماتریس ذخیره می‌شود. در واقع اگر $a(k)$ محل ذخیره‌سازی a_{ij} باشد، آنگاه $ia(k) = i$ و $ja(k) = j$. برای روشن شدن مطلب مثال بعدی را ارائه می‌کنیم.

مثال ۱.۵ ماتریس

$$A = \begin{pmatrix} ۱ & ۰ & ۰ & -۱ & ۰ \\ ۲ & ۰ & -۲ & ۰ & ۳ \\ ۰ & -۳ & ۰ & ۰ & ۰ \\ ۰ & ۴ & ۰ & -۴ & ۰ \\ ۵ & ۰ & -۵ & ۰ & ٦ \end{pmatrix},$$

را در نظر بگیرید. ذخیره‌سازی این ماتریس در قالب مختصاتی به صورت زیر است.

k	۱	۲	۳	۴	۵	٦	۷	۸	۹	۱۰	۱۱
ia	۱	۲	۲	۱	۵	۳	۴	۵	۲	۴	۵
ja	۴	۵	۱	۱	۵	۲	۴	۳	۳	۲	۱
a	−۱	۳	۲	۱	٦	−۳	−۴	−۵	−۲	۴	۵

در این مثال، درایه‌های ماتریس به صورت دلخواه و بدون هیچ ترتیبی چیده شده‌اند، ولی معمولاً به صورت سطر به سطر یا ستون به ستون ذخیره می‌شوند.

در قالب مختصاتی برای ذخیره‌سازی یک ماتریس $n \times n$، $۳nz$ عدد ذخیره می‌شود. توجه کنید که تعداد درایه‌های ماتریس $n^۲$ است که برای یک ماتریس تُنُک و در مقایسه با $۳nz$، عدد بسیار بزرگی است. برای مثال برای ذخیره سازی یک ماتریس سه‌قطری با بعد ۱۰۰۰ در قالب مختصاتی تنها ۸۹۹۴ عدد ذخیره می‌شوند، در حالی که تعداد درایه‌های ماتریس برابر با $۱۰^٦$ است.

در بسیاری از الگوریتم‌های تکراری برای حل دستگاه معادلات خطی نیاز به چندین ضرب ماتریس در بردار می‌باشد. در این صورت وقتی ماتریس ضرایب دستگاه در قالب مختصاتی ذخیره شده است، لازم است الگوریتمی برای ضرب ماتریس در یک بردار طراحی شود. الگوریتم ۱.۵ برای این منظور نوشته شده است.

الگوریتم ۱.۵: ضرب ماتریس در بردار در قالب مختصاتی

1. Input a, ia, ja, nnz and x
2. $y := ۰$
3. For $k = ۱, ۲, \ldots, nnz$, Do
4. $\quad y(ia(k)) := y(ia(k)) + a(k)\ x(ja(k))$
5. EndDo.

الگوریتم ۱.۵ به این صورت عمل می‌کند که ابتدا بردار y را برابر با صفر قرار داده، سپس یک پیمایش کامل در ia انجام می‌دهد، از $ia(۱)$ تا $ia(n)$. در مرحلهٔ k ام از پیمایش، درایهٔ $(ia(k), ja(k))$ از ماتریس A برابر با $a(k)$ خواهد بود. این درایه تنها روی $y(ia(k))$ اثر می‌گذارد و تاثیر آن به صورت

$$y(ia(k)) := y(ia(k)) + a(k)\ x(ja(k)),$$

خواهد بود.

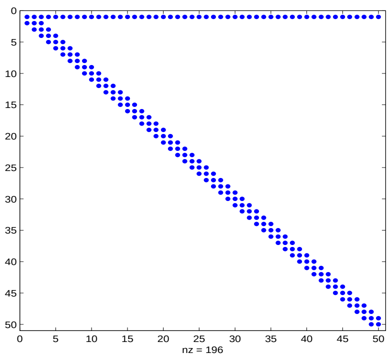
۲.۵ چرا روش تکراری؟

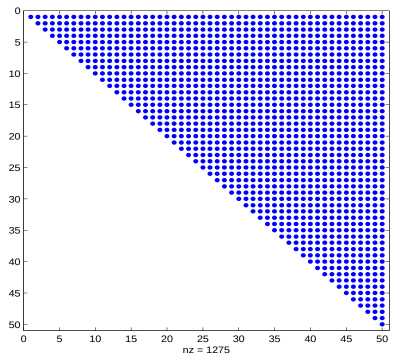
در این بخش با ارائه مثال‌هایی، دو مشکل اساسی روش‌های مستقیم و لزوم استفاده از روش‌های تکراری را مطالعه می‌نماییم. در فصل ۳ چند روش مستقیم برای حل دستگاه معادلات خطی، را معرفی نمودیم. از بین روش‌های مستقیم، مهمترین، پرکاربردترین و کم هزینه ترین روش مستقیم برای حل دستگاه معادلات خطی بزرگ و تنک، روش حذفی گاوس است. در ادامه با ارائه دو مثال، معایب این روش را نشان می‌دهیم.

مثال ۲.۵ دستگاه معادلات خطی $Ax = b$ را در نظر بگیرید که در آن ماتریس A به صورت

$$A = \begin{pmatrix} ۲ & -۱ & -۱ & -۱ & \cdots & -۱ \\ -۱ & ۲ & -۱ & & & \\ & -۱ & ۲ & -۱ & & \\ & & \ddots & \ddots & \ddots & \\ & & & -۱ & ۲ & -۱ \\ & & & & -۱ & ۲ \end{pmatrix} \in \mathbb{R}^{n\times n},$$

تعریف می‌شود. همچنین b یک بردار دلخواه انتخاب می‌شود. قرار می‌دهیم $n = ۵۰$. اگر روش حذفی گاوس را برای حل دستگاه به کار ببریم ماتریس ضرایب دستگاه نهایی یک ماتریس بالامثلثی U خواهد بود. در شکل ۲.۵ هر دو ماتریس A و U را نمایش داده‌ایم. همان‌طور که ملاحظه می‌شود، تعداد درایه‌های ناصفر ماتریس A برابر با $nz = ۱۹۶$ است اما تعداد درایه‌های U برابر با ۱۲۷۵ است. می‌بینیم که برای حل مسأله با استفاده از روش حذفی گاوس استفاده از حافظه کامپیوتر به میزان ۶.۵ برابر افزایش می‌یابد. این نسبت به ازای $n = ۸۰۰$ برابر با ۱۰۰.۲۵ است!





شکل ۲.۵: نمایش ماتریس A (سمت چپ) و U (سمت راست) برای مثال ۲.۵.

در بخش ۶.۳ اثر اختلال (از جمله اثر گرد کردن اعداد) روی جواب یک دستگاه را مطالعه نمودیم. دیدیم که عمل گردکردن اعداد چگونه ممکن است روی جواب یک دستگاه بدحالت اثر بگذارد. با این وجود مثال دیگری که به خوبی اثر گردکردن روی نتیجه حاصل از حل یک دستگاه را نشان می‌دهد، ارائه می‌کنیم.

مثال ۳.۵ دستگاه معادلات خطی $H_n x = b$ رادر نظر بگیرید که در آن $H_n = (h_{ij})$ ماتریس هیلبرت $n \times n$ است. در اینجا لازم به ذکر است که درایه‌های ماتریس هیلبرت به صورت $h_{ij} = (۱/(i + j - ۱))$ تعریف می‌شوند. ماتریس هیلبرت معین مثبت متقارن است، اما یک ماتریس بدحالت می‌باشد[۴]. توجه کنید که به ازای $n = ۱۳$، داریم:

$$\mathrm{cond}_۲(A) = \|A\|_۲ \|A^{-۱}\|_۲ = ۹.۵۴ \times ۱۰^{۱۷} \gg ۱,$$

که نشان می‌دهد $H_n x = b$ یک دستگاه بدحالت است. قرار می‌دهیم $n = ۱۳$ و $b_i = \sum_{j=۱}^{۱۳} h_{ij}$، $i = ۱, \dots, ۱۳$. در این صورت، جواب واقعی دستگاه بردار $x^* = (۱, ۱, \dots, ۱)^T \in \mathbb{R}^{۱۳}$ خواهد بود. اگر این دستگاه را با کمک یک برنامه به زبان متلب و با دقت مضاعف با استفاده از روش حذفی گاوس حل کنیم جواب به دست آمده به صورت

$$\begin{aligned}\tilde{x} = (&۱.۰۰۰۰, ۱.۰۰۰۰, ۱.۰۰۰۶, ۰.۹۸۹۸, ۱.۰۹۲۰, ۰.۵۰۱۰, ۲.۷۴۰۹,\\ &-۳.۰۳۶۱, ۷.۲۸۳۹, -۵.۴۹۳۵, ۵.۲۷۰۹, -۰.۶۱۸۲, ۱.۲۶۸۸)^T,\end{aligned}$$

خواهد بود. خطای نسبی در جواب به دست آمده برابر با

$$\frac{\|x - \tilde{x}\|_۲}{\|x\|_۲} = ۳.۰۷,$$

است. بنابراین با استفاده از روش حذفی گاوس جوابی با ۳۰۷٪ خطا به دست آمده است. ملاحظه می‌شود با اینکه دستگاه بررسی‌شده بسیار کوچک است، جواب به دست آمده بسیار نامناسب است.

برای رفع مشکلات اشاره شده از روش‌های تکراری برای حل دستگاه معادلات خطی استفاده می‌شود. به طور کلی در یک روش تکراری یک حدس اولیه برای جواب انتخاب می‌شود و با تکرار یک فرایند این جواب بهبود می‌یابد. در بخش بعد دو روش تکراری شناخته شدهٔ ژاکوبی و گاوس- سایدل را معرفی می‌کنیم.

۳.۵ روش‌های تکراری ژاکوبی و گاوس- سایدل

۱.۳.۵ روش تکراری ژاکوبی

فرض می‌کنیم درایه‌های قطری ماتریس A ناصفر باشند. از معادلهٔ i ام دستگاه (۱.۵)، x_i را محاسبه می‌کنیم:

$$x_i = \frac{۱}{a_{ii}} \left(b_i - \sum_{\substack{j=۱\\ j\neq i}}^{n} a_{ij} x_j \right), \quad i = ۱, ۲, \dots, n.$$

حال با فرض $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \in \mathbb{R}^n$ به عنوان یک حدس اولیهٔ دلخواه برای جواب واقعی دستگاه $Ax = b$، روند تکراری ژاکوبی را به صورت

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(k)}\right), \quad i = 1, 2, \dots, n, \tag{۲.۵}$$

می‌سازیم، که در آن $k = 0, 1, 2, \dots$. با توجه به ساختار روند تکراری ژاکوبی، می‌بینیم که در هر تکرار مؤلفه‌های بردار $x^{(k+1)}$ توسط مؤلفه‌های بردار $x^{(k)}$ به روز می‌شوند. با توجه به اینکه روش ژاکوبی یک روش تکراری است، وقتی که $k \to \infty$، تحت شرایطی این روش به جواب واقعی دستگاه همگرا می‌شود. بنابراین یک تولرانس ϵ انتخاب می‌شود و به ازای یک k، اگر یکی از شرایط

$$e_k = \frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} < \epsilon,$$

یا

$$r_k = \frac{\|b - Ax^{(k)}\|}{\|b - Ax^{(0)}\|} < \epsilon$$

برقرار باشد توقف کرده و بردار $x^{(k)}$ را به عنوان جواب تقریبی دستگاه می‌پذیریم. در اینجا $\|.\|$ یک نرم برداری دلخواه است.

مثال ۴.۵ یک جواب تقریبی از دستگاه

$$\begin{pmatrix} 5 & 1 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ -1 \\ 9 \end{pmatrix},$$

را با استفاده از روش تکراری ژاکوبی که در شرط $e_k < 10^{-3}$ صدق کند، بیابید. از نرم بینهایت در شرط توقف استفاده کرده و حدس اولیه را بردار صفر بگیرید.

حل: روند تکراری ژاکوبی به صورت

$$\begin{cases} x_1^{(k+1)} = \frac{1}{5}(8 - x_2^{(k)} - 2x_3^{(k)}), \\ x_2^{(k+1)} = \frac{1}{4}(-1 - x_1^{(k)} - x_3^{(k)}), \\ x_3^{(k+1)} = \frac{1}{5}(9 - x_1^{(k)} - 2x_2^{(k)}), \end{cases}$$

است، که در آن $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T = (0, 0, 0)^T$. به ازای $k = 0$، داریم:

$$\begin{cases} x_1^{(1)} = \frac{1}{5}(8 - x_2^{(0)} - 2x_3^{(0)}) = \frac{1}{5}(8 - 0 - 2 \times 0) = 1.6000, \\ x_2^{(1)} = \frac{1}{4}(-1 - x_1^{(0)} - x_3^{(0)}) = \frac{1}{4}(-1 - 0 - 0) = -0.2500, \\ x_3^{(1)} = \frac{1}{5}(9 - x_1^{(0)} - 2x_2^{(0)}) = \frac{1}{5}(9 - 0 - 2 \times 0) = 1.800. \end{cases}$$

به همین ترتیب به ازای $k = 1$، داریم:

$$\begin{cases} x_1^{(2)} = \frac{1}{5}(8 - x_2^{(1)} - 2x_3^{(1)}) = \frac{1}{5}(8 + 0.25 - 2 \times 1.8) = 0.9300, \\ x_2^{(2)} = \frac{1}{4}(-1 - x_1^{(1)} - x_3^{(1)}) = \frac{1}{4}(-1 - 1.6 - 1.8) = -1.1000, \\ x_3^{(2)} = \frac{1}{5}(9 - x_1^{(1)} - 2x_2^{(1)}) = \frac{1}{5}(9 - 1.6 + 2 \times 0.25) = 1.5800. \end{cases}$$

با ادامهٔ این روند در تکرار دهم، خواهیم داشت:

$$x^{(10)} = \begin{pmatrix} 0.9978 \\ -1.0020 \\ 1.9978 \end{pmatrix},$$

و $e_{10} = 6.94 \times 10^{-4} < 10^{-3}$.

قضیهٔ بعدی یک شرط کافی برای همگرایی روش تکراری ژاکوبی را می‌دهد.

قضیه ۱.۵ اگر ماتریس A غالب قطری اکید باشد، آنگاه روش تکراری ژاکوبی با هر حدس اولیهٔ به جواب دستگاه همگرا می‌شود.

برهان: با فرض $x^* = A^{-1}b = (x_1^*, x_2^*, \ldots, x_n^*)^T$، داریم

$$x_i^* = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^* \right), \quad i = 1, 2, \ldots, n. \qquad (3.5)$$

با تفریق طرفین این رابطه و رابطهٔ (۲.۵)، داریم:

$$x_i^{(k+1)} - x_i^* = \frac{1}{a_{ii}} \left(-\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} (x_j^{(k)} - x_j^*) \right), \quad i = 1, 2, \ldots, n.$$

در نتیجه با فرض $e_i^{(l)} = x_i^{(l)} - x_i^*$، برای $i = ۱, ۲, \dots, n$ و $l = ۰, ۱, ۲, \dots$، داریم:

$$e_i^{(k+۱)} = \frac{۱}{a_{ii}} \left(- \sum_{\substack{j=۱ \\ j \neq i}}^{n} a_{ij} e_j^{(k)} \right), \quad i = ۱, ۲, \dots, n. \qquad (۴.۵)$$

از این‌رو، با توجه به رابطهٔ (۴.۵) می‌بینیم که

$$|e_i^{(k+۱)}| \leqslant \frac{۱}{|a_{ii}|} \left(\sum_{\substack{j=۱ \\ j \neq i}}^{n} |a_{ij}| |e_j^{(k)}| \right) \leqslant \left(\frac{۱}{|a_{ii}|} \sum_{\substack{j=۱ \\ j \neq i}}^{n} |a_{ij}| \right) \|e^{(k)}\|_\infty, \qquad (۵.۵)$$

که در آن $e^{(l)} = x^{(l)} - x^*$. حال قرار می‌دهیم:

$$\rho_i = \frac{۱}{|a_{ii}|} \sum_{\substack{j=۱ \\ j \neq i}}^{n} |a_{ij}|, \quad i = ۱, ۲, \dots, n,$$

و $\rho = \max_i \rho_i$. با توجه به غالب قطری بودن A داریم $\rho < ۱$. بنابراین از رابطهٔ (۵.۵) نتیجه می‌شود که

$$|e_i^{(k+۱)}| \leqslant \rho \|e^{(k)}\|_\infty, \quad i = ۱, ۲, \dots, n.$$

لذا

$$\|e^{(k+۱)}\|_\infty = \max_i |e_i^{(k+۱)}| \leqslant \rho \|e^{(k)}\|_\infty.$$

از این‌رو

$$\|e^{(k+۱)}\|_\infty \leqslant \rho \|e^{(k)}\|_\infty \leqslant \rho^۲ \|e^{(k-۱)}\|_\infty \leqslant \cdots \leqslant \rho^{k+۱} \|e^{(۰)}\|_\infty,$$

و بنابراین

$$\lim_{k \to \infty} \|e^{(k+۱)}\|_\infty \leqslant \lim_{k \to \infty} \rho^{k+۱} \|e^{(۰)}\|_\infty = ۰,$$

که همگرایی روش را نتیجه می‌دهد. □

در اینجا ذکر این نکته ضروری است که شرط غالب قطری بودن ماتریس ضرایب برای همگرایی یک شرط کافی است و ممکن است این شرط برقرار نباشد و همگرایی حاصل

گردد. با وجود اینکه ماتریس ضرایب دستگاه

$$\begin{pmatrix} ۴ & -۱ \\ -۲ & ۱ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix} = \begin{pmatrix} ۲ \\ ۰ \end{pmatrix}, \qquad (۶.۵)$$

غالب قطری اکید نیست، اما روش ژاکوبی با هر حدس اولیهٔ دلخواه همگرا است. برای اثبات این ادعا، ابتدا می بینیم که تکرار روش ژاکوبی برای حل این دستگاه به صورت

$$\begin{cases} x_۱^{(k+۱)} = \frac{۱}{۴}(x_۲^{(k)} + ۲), \\ x_۲^{(k+۱)} = ۲x_۱^{(k)}. \end{cases} \qquad (۷.۵)$$

است. از طرفی اگر $x^* = (x_۱^*, x_۲^*)^T$ جواب واقعی دستگاه باشد، آنگاه خواهیم داشت:

$$\begin{cases} x_۱^* = \frac{۱}{۴}(x_۲^* + ۲), \\ x_۲^* = ۲x_۱^*, \end{cases} \qquad (۸.۵)$$

حال با استفاده از تفریق طرفین تساوی‌های (۷.۵) و (۸.۵)، داریم:

$$\begin{cases} e_۱^{(k+۱)} = \frac{۱}{۴}e_۲^{(k)}, \\ e_۲^{(k+۱)} = ۲e_۱^{(k)}, \end{cases}$$

که در آن $e_۱^{(j)} = x_۱^{(j)} - x_۱^*$ و $e_۲^{(j)} = x_۲^{(j)} - x_۲^*$، $j = ۰, ۱, \ldots$. از این‌رو

$$e^{(k+۱)} = P^{k+۱}e^{(۰)},$$

که در آن

$$e^{(j)} = \begin{pmatrix} e_۱^{(j)} \\ e_۲^{(j)} \end{pmatrix}, \qquad P = \begin{pmatrix} ۰ & \frac{۱}{۴} \\ ۲ & ۰ \end{pmatrix}.$$

بنابراین برای همگرایی دنباله کافی است نشان دهیم $\lim_{k\to\infty} P^{k+۱} = ۰$. اما داریم:

$$P^m = \begin{cases} \frac{۱}{۲^r}I, & m = ۲r, \\ \frac{۱}{۲^r}P, & m = ۲r + ۱. \end{cases}$$

این رابطه نشان می‌دهد که $\lim_{k\to\infty} P^{k+۱} = ۰$ و بدین ترتیب همگرایی دنباله نتیجه می‌شود.

در ادامه به معرفی روش تکراری گاوس- سایدل می‌پردازیم.

۲.۳.۵ روش تکراری گاوس- سایدل

رابطهٔ (۲.۵) را به صورت

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right), \qquad (9.5)$$

به ازای $i = 1, 2, \ldots, n$، می‌نویسیم. فرض کنید بردار $x^{(k)}$ محاسبه شده و می‌خواهیم بردار $x^{(k+1)}$ را محاسبه کنیم. برای محاسبهٔ $x_i^{(k+1)}$ می‌بینیم که $i-1$ مولفهٔ اول $x^{(k+1)}$ محاسبه شده است. بنابراین از این مقادیر در محاسبهٔ $x_i^{(k+1)}$ استفاده کرده و رابطهٔ (۹.۵) را به صورت

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right), \qquad (10.5)$$

به ازای $i = 1, 2, \ldots, n$، می‌نویسیم. روند تکراری (۱۰.۵) معرف روش تکراری گاوس- سایدل است.

مثال ۵.۵ مثال ۴.۵ را با استفاده از روش گاوس- سایدل حل کنید.

حل: روند تکراری گاوس- سایدل برای حل دستگاه داده شده به صورت

$$\begin{cases} x_1^{(k+1)} = \frac{1}{5}(8 - x_2^{(k)} - 2x_3^{(k)}), \\ x_2^{(k+1)} = \frac{1}{4}(-1 - x_1^{(k+1)} - x_3^{(k)}), \\ x_3^{(k+1)} = \frac{1}{5}(9 - x_1^{(k+1)} - 2x_2^{(k+1)}). \end{cases}$$

است، که در آن $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T = (0, 0, 0)^T$. به ازای $k = 0$، داریم:

$$\begin{cases} x_1^{(1)} = \frac{1}{5}(8 - x_2^{(0)} - 2x_3^{(0)}) = \frac{1}{5}(8 - 0 - 2 \times 0) = 1.6000, \\ x_2^{(1)} = \frac{1}{4}(-1 - x_1^{(1)} - x_3^{(0)}) = \frac{1}{4}(-1 - 1.6 - 0) = -0.6500, \\ x_3^{(1)} = \frac{1}{5}(9 - x_1^{(1)} - 2x_2^{(1)}) = \frac{1}{5}(9 - 1.6 + 2 \times 0.6500) = 1.74. \end{cases}$$

به همین ترتیب به ازای $k = 1$، داریم:

$$\begin{cases} x_1^{(2)} = \frac{1}{5}(8 - x_2^{(1)} - 2x_3^{(1)}) = \frac{1}{5}(8 + 0.65 - 2 \times 1.74) = 1.0340, \\ x_2^{(2)} = \frac{1}{4}(-1 - x_1^{(2)} - x_3^{(1)}) = \frac{1}{4}(-1 - 1.034 - 1.74) = -0.9435, \\ x_3^{(2)} = \frac{1}{5}(9 - x_1^{(2)} - 2x_2^{(2)}) = \frac{1}{5}(9 - 1.034 + 2 \times 0.9435) = 1.9706. \end{cases}$$

با ادامهٔ این روند در تکرار چهارم، خواهیم داشت:

$$x^{(۴)} = \begin{pmatrix} ۰.۹۹۹۷ \\ -۰.۹۹۹۲ \\ ۱.۹۹۹۷ \end{pmatrix},$$

و $e_{۱۰} = ۷.۱۳ \times ۱۰^{-۴} < ۱۰^{-۳}$.

همان‌طور که ملاحظه می‌شود روش گاوس-سایدل در تعداد تکرارهای کمتری به یک جواب با دقت $e_k < ۱۰^{-۳}$ رسیده است. اگرچه این موضوع معمولاً درست است اما در حالت کلی صحیح نیست. برای مشاهده مثالی در این خصوص می‌توانید به [۱] مراجعه کنید. قضیهٔ بعدی یک شرط کافی برای همگرایی روش گاوس-سایدل ارائه می‌دهد.

قضیه ۲.۵ اگر ماتریس A غالب قطری اکید باشد، آنگاه روش تکراری گاوس-سایدل با هر حدس اولیهٔ دلخواه به جواب دستگاه همگرا می‌شود.

برهان: همانند اثبات قضیهٔ ۱.۵ و نمادهای استفاده شده، برای $i = ۱, ۲, \ldots, n$، داریم:

$$x_i^{(k+۱)} - x_i^* = \frac{۱}{a_{ii}} \left(-\sum_{j=۱}^{i-۱} a_{ij}(x_j^{(k+۱)} - x_j^*) - \sum_{j=i+۱}^{n} a_{ij}(x_j^{(k)} - x_j^*) \right),$$

یا

$$e_i^{(k+۱)} = \frac{۱}{a_{ii}} \left(-\sum_{j=۱}^{i-۱} a_{ij} e_j^{(k+۱)} - \sum_{j=i+۱}^{n} a_{ij} e_j^{(k)} \right), \quad i = ۱, ۲, \ldots, n. \quad (۱۱.۵)$$

فرض کنید $|e_r^{(k+۱)}| = \|e^{(k+۱)}\|_\infty$. بنابراین

$$|a_{rr}||e_r^{(k+۱)}| \leqslant \left(\sum_{j=۱}^{r-۱} |a_{rj}| \right) \|e^{(k+۱)}\|_\infty + \left(\sum_{j=r+۱}^{n} |a_{rj}| \right) \|e^{(k)}\|_\infty.$$

در نتیجه

$$\left(|a_{rr}| - \sum_{j=۱}^{r-۱} |a_{rj}| \right) \|e^{(k+۱)}\|_\infty \leqslant \left(\sum_{j=r+۱}^{n} |a_{rj}| \right) \|e^{(k)}\|_\infty,$$

که معادل با

$$\|e^{(k+۱)}\|_\infty \leqslant \rho_r \|e^{(k)}\|_\infty, \quad i = ۱, ۲, \ldots, n,$$

است، که در آن

$$\rho_r = \frac{\sum_{j=r+1}^{n} |a_{rj}|}{|a_{rr}| - \sum_{j=1}^{r-1} |a_{rj}|}.$$

به روشنی با توجه به اینکه ماتریس A غالب قطری اکید است داریم $\rho_r < 1$. بنابراین همانند اثبات قضیهٔ ۱.۵، نتیجهٔ لازم به دست می‌آید. □

مثال ۶.۵ دستگاه معادلات خطی $Ax = b$ را در نظر بگیرید که در آن

$$A = \begin{pmatrix} 4 & 1 & & & \\ -2 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -2 & 4 & 1 \\ & & & -2 & 4 \end{pmatrix} \in \mathbb{R}^{n\times n}, \quad b = \begin{pmatrix} 5 \\ 3 \\ \vdots \\ 3 \\ 2 \end{pmatrix} \in \mathbb{R}^n.$$

روشن است که ماتریس A غالب قطری اکید است. بنابراین هر دو روش ژاکوبی و گاوس-سایدل با هر حدس اولیهٔ دلخواه همگرا است. حدس اولیه را بردار صفر گرفته و این دو روش را برای حل دستگاه $Ax = b$ به کار می بریم. شرط توقف را

$$\|x^{(k+1)} - x^{(k)}\|_2 < 10^{-6}\|x^{(k+1)}\|_2,$$

می‌گیریم. به ازای مقادیر مختلف n، تعداد تکرارها و زمان‌های مربوطه برحسب ثانیه (داخل پرانتز) در جدول ۱.۵ ارائه شده‌اند. همان‌طور که می‌بینید هر دو روش در تکرارهای مناسبی به جواب تقریبی رسیده‌اند.

جدول ۱.۵: تعداد تکرارها و زمان لازم برای همگرایی برای مثال ۶.۵.

n	۲۰۰	۴۰۰	۶۰۰	۸۰۰	۱۰۰۰
ژاکوبی	۳۳(۰.۱۴)	۳۲(۰.۲۷)	۳۲(۰.۴۱)	۳۱(۰.۵۵)	۳۱(۰.۶۹)
گاوس - سایدل	۲۲(۰.۰۸)	۲۲(۰.۱۶)	۲۲(۰.۲۵)	۲۲(۰.۳۴)	۲۲(۰.۴۳)

روش‌های تکراری ژاکوبی و گاوس-سایدل در رده روش‌های تکراری ایستا هستند. در ادامه شکل کلی یک روش تکراری ایستا را معرفی خواهیم کرد.

۴.۵ روش‌های تکراری ایستا

این بخش را با تعریف شکافت شروع می‌کنیم.

تعریف ۱.۵ فرض کنید $A \in \mathbb{R}^{n\times n}$ و $A = M - N$ یک تجزیه از آن باشد. این تجزیه یک شکافت نامیده می‌شود هرگاه M یک ماتریس نامنفرد باشد.

مثال ۷.۵ فرض کنید:

$$A = \begin{pmatrix} ۲ & ۳ & -۱ \\ ۱ & ۲ & ۲ \\ ۱ & ۲ & ۴ \end{pmatrix},$$

و

$$M_۱ = \begin{pmatrix} ۲ & ۳ & ۰ \\ ۱ & ۲ & ۰ \\ ۰ & ۰ & ۴ \end{pmatrix}, \quad N_۱ = \begin{pmatrix} ۰ & ۰ & ۱ \\ ۰ & ۰ & -۲ \\ -۱ & -۲ & ۰ \end{pmatrix},$$
$$M_۲ = \begin{pmatrix} ۲ & ۳ & ۰ \\ ۱ & ۲ & ۰ \\ ۰ & ۰ & ۰ \end{pmatrix}, \quad N_۲ = \begin{pmatrix} ۰ & ۰ & ۱ \\ ۰ & ۰ & -۲ \\ -۱ & -۲ & -۴ \end{pmatrix}.$$

تجزیهٔ $A = M_۱ - N_۱$ یک شکافت است زیرا $M_۱$ نامنفرد است، اما تجزیهٔ ماتریس A به صورت $A = M_۲ - N_۲$ یک شکافت از A نیست چون $M_۲$ منفرد است.

حال به معرفی شکل کلی یک روش تکراری ایستا می‌پردازیم. فرض کنید $A = M - N$ یک شکافت از A باشد. دستگاه $Ax = b$ را به صورت $(M-N)x = b$ و بعد به صورت

$$x = M^{-۱}Nx + M^{-۱}b,$$

می‌نویسیم. سپس روند تکراری

$$x^{(k+۱)} = M^{-۱}Nx^{(k)} + M^{-۱}b, \quad k = ۰, ۱, \ldots, \qquad (۱۲.۵)$$

را تعریف می‌کنیم که در آن $x^{(۰)}$ یک حدس اولیهٔ دلخواه است. قرار می‌دهیم $G = M^{-۱}N$ و $f = M^{-۱}b$. ماتریس G را ماتریس تکرار روش می‌نامند. در این صورت

روند تکراری (۱۲.۵) به صورت

$$x^{(k+1)} = Gx^{(k)} + f, \qquad k = 0, 1, \ldots, \tag{۱۳.۵}$$

نوشته می‌شود. با توجه به اینکه ماتریس تکرار روش یک ماتریس ثابت است این گونه روش‌ها را روش‌های تکراری ایستا می‌نامند. قضیهٔ بعد شرط لازم و کافی برای همگرایی روش (۱۳.۵) به جواب دستگاه $Ax = b$ را بیان می‌کند.

قضیه ۳.۵ روند تکراری (۱۳.۵) به ازای هر حدس اولیهٔ $x^{(0)}$ به جواب دستگاه $Ax = b$ همگرا است اگر و تنها اگر $\rho(G) < 1$ که در آن $\rho(G) = \max\limits_{\lambda \in \sigma(G)} |\lambda|$.

برهان: در اینجا قضیه را برای حالتی که ماتریس G قطری‌شدنی باشد ثابت می‌کنیم. برای اثبات قضیه در حالت کلی به [۴، ۱] مراجعه کنید. فرض کنید P یک ماتریس نامنفرد به طوری که $P^{-1}GP = D = \mathrm{diag}_n(\lambda_1, \ldots, \lambda_n)$، که در آن $\lambda_i \in \sigma(G)$.

ابتدا فرض می‌کنیم $\rho(G) < 1$. در این صورت $|\lambda_i| < 1$، $i = 1, \ldots, n$. اگر $x^* = A^{-1}b$، آنگاه روشن است که

$$x^* = Gx^* + f. \tag{۱۴.۵}$$

از طرفی با توجه به رابطهٔ (۱۳.۵)، داریم:

$$x^{(k)} = Gx^{(k-1)} + f.$$

با تفریق طرفین این رابطه و رابطهٔ (۱۴.۵)، داریم:

$$x^{(k)} - x^* \;=\; G(x^{(k-1)} - x^*) = G^2(x^{(k-2)} - x^*) = \cdots = G^k(x^{(0)} - x^*).$$

حال با توجه به اینکه $|\lambda_i| < 1$، اگر $k \to \infty$، آنگاه $\lambda_i^k \to 0$. از این‌رو

$$\begin{aligned}\lim_{k\to\infty} G^k \;&=\; \lim_{k\to\infty} (P^{-1}DP)^k = \lim_{k\to\infty} P^{-1}D^kP = P^{-1}\lim_{k\to\infty} D^kP \\ &=\; P^{-1}\lim_{k\to\infty}\begin{pmatrix}\lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k\end{pmatrix}P = 0.\end{aligned}$$

بنابراین از رابطهٔ (۱۵.۵)، داریم:

$$\lim_{k\to\infty} x^{(k)} - x^* = \lim_{k\to\infty} G^k(x^{(\circ)} - x^*) = \circ,$$

که همگرایی را تضمین می‌کند. برعکس فرض می‌کنیم روند تکراری به ازای هر حدس اولیهٔ $x^{(\circ)}$ همگرا باشد. با توجه به همگرایی روش، از رابطهٔ (۱۵.۵) به ازای هر $x^{(\circ)}$، داریم:

$$\lim_{k\to\infty} G^k(x^{(\circ)} - x^*) = \circ.$$

حال فرض کنید (λ_r, x_r) یک زوج ویژه برای ماتریس G باشد به طوری که $\rho(G) = |\lambda_r|$. با فرض $x^{(\circ)} = x_r + x^*$ و قرار دادن آن در رابطهٔ اخیر داریم:

$$\begin{aligned} \circ &= \lim_{k\to\infty} G^k(x^{(\circ)} - x^*) = \lim_{k\to\infty} G^k(x_r + x^* - x^*) \\ &= \lim_{k\to\infty} G^k(x_r) = \lim_{k\to\infty} \lambda_r^k x_r = \left(\lim_{k\to\infty} \lambda_r^k\right) x_r. \end{aligned}$$

این رابطه نشان می‌دهد که $\rho(G) = |\lambda_r| < ۱$. □

یادداشت ۱.۵ با توجه به اثبات قضیهٔ ۳.۵ روشن است که هرچه شعاع طیفی ماتریس تکرار کوچکتر باشد انتظار می‌رود که سرعت همگرایی بیشتر باشد.

نتیجه ۱.۵ اگر به ازای نرم طبیعی $\|.\|$، $\|A\| < ۱$، آنگاه روند تکراری (۱۳.۵) به ازای هر حدس اولیهٔ $x^{(\circ)}$ به جواب دستگاه $Ax = b$ همگرا است.

برهان: فرض کنید λ یک مقدار ویژهٔ ماتریس G باشد. با استفاده از قضیهٔ ۴.۲، داریم $|\lambda| \leqslant \|A\|$. لذا $\rho(A) \leqslant \|A\| < ۱$ که همگرایی روش را نشان می‌دهد. □

مثال ۸.۵ دستگاه معادلات خطی $Ax = b$ را در نظر بگیرید که درآن A ماتریس مثال ۷.۵ می‌باشد و $b = (۱۲, ۱۰, ۱۲)^T$. همچنین شکافت $A = M_۱ - N_۱$ تعریف شده در مثال ۷.۵ را در نظر می‌گیریم. داریم:

$$G = M_۱^{-۱} N_۱ = \frac{۱}{۴}\begin{pmatrix} \circ & \circ & ۳۲ \\ \circ & \circ & -۲\circ \\ -۱ & -۲ & \circ \end{pmatrix}, \quad f = M_۱^{-۱} b = \begin{pmatrix} -٦ \\ ۸ \\ ۳ \end{pmatrix}.$$

با محاسبهٔ مقادیر ویژهٔ G، خواهیم داشت:

$$\sigma(G) = \left\{ ۰, -\frac{\sqrt{۲}}{۲}, \frac{\sqrt{۲}}{۲} \right\},$$

و از این‌رو $\rho(G) = \frac{\sqrt{۲}}{۲} \approx ۰.۷۰۷۱ < ۱$. این رابطه نشان می‌دهد که این روش همگرا است.

برای روشن شدن مطلب، صورت مؤلفه‌ای روند تکراری $x^{(k+۱)} = Gx^{(k)} + f$ را به صورت

$$\begin{pmatrix} x_۱^{(k+۱)} \\ x_۲^{(k+۱)} \\ x_۳^{(k+۱)} \end{pmatrix} = \frac{۱}{۴} \begin{pmatrix} ۰ & ۰ & ۳۲ \\ ۰ & ۰ & -۲۰ \\ -۱ & -۲ & ۰ \end{pmatrix} \begin{pmatrix} x_۱^{(k)} \\ x_۲^{(k)} \\ x_۳^{(k)} \end{pmatrix} + \begin{pmatrix} -۶ \\ ۸ \\ ۳ \end{pmatrix},$$

یا

$$\begin{cases} x_۱^{(k+۱)} = ۸x_۳^{(k)} - ۶, \\ x_۲^{(k+۱)} = -۵x_۳^{(k)} + ۸, \\ x_۳^{(k+۱)} = \frac{۱}{۴}(-x_۱^{(k)} - ۲x_۲^{(k)}) + ۳, \end{cases}$$

می‌نویسیم. حدس اولیه را $x^{(۰)} = (x_۱^{(۰)}, x_۲^{(۰)}, x_۳^{(۰)})^T = (۰, ۰, ۰)^T$ می‌گیریم. در تکرار سی‌ام خواهیم داشت:

$$x^{(۳۰)} = (۲.۰۰۱۰, ۲.۹۹۹۴, ۱.۰۰۰۰)^T.$$

لازم به ذکر است که جواب واقعی دستگاه برابر با $x^* = (۲, ۳, ۱)^T$ است. ملاحظه می‌شود که روش فوق جواب مناسبی را فراهم کرده است.

در مثال قبل ماتریس $M_۱$ یک ماتریس کوچک بود و معکوس آن به سادگی محاسبه گردید. اما در حالت کلی مسأله به سادگی این مثال نیست. در عمل ماتریس M باید طوری انتخاب گردد که دستگاه $Mz = y$ برای محاسبهٔ z به سادگی حل شود. دلیل این امر را بیان می‌کنیم. روند تکراری (۱۲.۵) را به صورت

$$Mx^{(k+۱)} = Nx^{(k)} + b,$$

می‌نویسیم. با فرض $f_k = Nx^{(k)} + b$، می‌بینیم که یک تکرار از روند تکراری (۱۳.۵) معادل با حل دستگاه معادلات خطی $Mx^{(k+1)} = f_k$، برای محاسبهٔ $x^{(k+1)}$ است. بنابراین ماتریس M بایستی طوری انتخاب شود که این دستگاه به سادگی حل شود. برای مثال اگر M یک ماتریس قطری، بالامثلثی یا پایین‌مثلثی باشد، آنگاه این دستگاه به سادگی حل می‌شود.

در بخش بعد نشان می‌دهیم که روش‌های ژاکوبی و گاوس- سایدل دو روش تکراری ایستا هستند.

۵.۵ شکل ماتریسی روش‌های ژاکوبی و گاوس- سایدل

فرض کنید که $A = D - E - F$ که در آن

$$
\begin{aligned}
D &= \operatorname{diag}_n(a_{11}, a_{22}, \ldots, a_{nn}),\\
E &= \begin{pmatrix} 0 & & & \\ -a_{21} & 0 & & \\ \vdots & \ddots & \ddots & \\ -a_{n,1} & \cdots & -a_{n,n-1} & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & -a_{12} & \cdots & -a_{1,n} \\ & 0 & \ddots & \vdots \\ & & \ddots & -a_{n-1,n} \\ & & & 0 \end{pmatrix}.
\end{aligned}
$$

توجه کنید که $-E$ $(-F)$ یک ماتریس پایین‌مثلثی‌اکید(بالامثلثی اکید) است به طوری که درایه‌های زیر قطر (بالای قطر) اصلی آن با قرینهٔ درایه‌های پایین قطر (بالای قطر) اصلی A یکی هستند. از اینجا به بعد فرض می‌کنیم $a_{ii} \neq 0$، $i = 1, \ldots, n$.

قرار می‌دهیم $M = D$ و $N = E + F$. توجه کنید که ماتریس D نامنفرد است. در این صورت، روند تکراری

$$x^{(k+1)} = D^{-1}(E + F)x^{(k)} + D^{-1}b, \quad k = 0, 1, 2, \ldots, \tag{۱۵.۵}$$

را به دست خواهیم آورد. شکل مؤلفه‌ای این روش را به دست می‌آوریم. رابطهٔ (۱۵.۵) را به صورت زیر می‌نویسیم:

$$Dx^{(k+1)} = (E + F)x^{(k)} + b.$$

معادلهٔ i ام رابطهٔ اخیر را می‌توان به صورت

$$a_{ii}x_i^{(k+1)} = b_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j^{(k)}, \quad i = 1, \ldots, n,$$

نوشت که معادل با

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

است. ملاحظه می‌شود که روند تکراری (۱۵.۵) چیزی نیست جز روش تکراری ژاکوبی.

قرار می‌دهیم $M = D - E$ و $N = F$. روشن است که درایه‌های قطری M با درایه‌های قطری A یکی هستند. با توجه به اینکه M یک ماتریس پایین مثلثی است، نامنفرد نیز خواهد بود. در این صورت روش تکراری

$$x^{(k+1)} = (D-E)^{-1}Fx^{(k)} + (D-E)^{-1}b, \quad k = 0, 1, 2, \ldots, \qquad (۱۶.۵)$$

را خواهیم داشت. برای به دست آوردن شکل مؤلفه‌ای این روش، رابطهٔ اخیر را به صورت

$$(D-E)x^{(k+1)} = Fx^{(k)} + b, \quad k = 0, 1, 2, \ldots,$$

می‌نویسیم. با در نظر گرفتن معادلهٔ i ام، داریم:

$$\sum_{j=1}^{i} a_{ij}x_j^{(k+1)} = -\sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i, \quad i = 1, \ldots, n.$$

یا

$$a_{ii}x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}, \quad i = 1, \ldots, n,$$

و در نتیجه

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

این رابطه نشان می‌دهد که رابطهٔ (۱٦.۵) همان روش تکراری گاوس- سایدل است.

در پایان همگرایی روش‌های تکراری ژاکوبی و گاوس- سایدل را برای ماتریس‌های معین مثبت متقارن بررسی می‌کنیم. از اینجا به بعد ماتریس تکرار روش‌های ژاکوبی و

گاوس- سایدل را به ترتیب با G_{J} و G_{GS} نشان می‌دهیم. ابتدا با ارائه یک مثال نشان می‌دهیم که روش تکراری ژاکوبی برای ماتریس‌های معین مثبت متقارن لزوماً همگرا نیست.

مثال ۹.۵ فرض کنید

$$A = \begin{pmatrix} ۱۵ & -۱۰ & -۴ \\ -۱۰ & ۱۴ & ٦ \\ -۴ & ٦ & ۷ \end{pmatrix}, \quad b = \begin{pmatrix} ۱ \\ ۱۰ \\ ۹ \end{pmatrix}.$$

مجموعهٔ مقادیر ویژهٔ ماتریس A برابر با $\sigma(A) = \{۳, ٦, ۲۷\}$ است. با توجه به قضیهٔ ۱۰.۴ از اینکه ماتریس A متقارن و مقادیر ویژهٔ آن مثبت هستند، معین مثبت است. ماتریس تکرار روش ژاکوبی برای حل دستگاه $Ax = b$ برابر با

$$G_{\mathrm{J}} = \begin{pmatrix} ۰ & \frac{۲}{۳} & \frac{۴}{۱۵} \\ \frac{۵}{۷} & ۰ & -\frac{۳}{۷} \\ \frac{۴}{۷} & -\frac{٦}{۷} & ۰ \end{pmatrix},$$

است و

$$\sigma(G_{\mathrm{J}}) = \{-۱.۱۳۳۲, ۰.۳۸۵۳, ۰.۷۴۷۹\},$$

بنابراین $\rho(G_{\mathrm{J}}) = ۱.۱۳۳۱ > ۱$. در نتیجه بنا به قضیهٔ ۳.۵، روش تکراری ژاکوبی برای حل دستگاه $Ax = b$ واگرا است.

قضیه ۴.۵ اگر ماتریس A معین مثبت متقارن باشد، آنگاه روش تکراری گاوس- سایدل با هر حدس اولیهٔ دلخواه به جواب دستگاه همگرا است.

برهان: فرض کنید (λ, x) یک زوج ویژهٔ ماتریس تکرار روش تکراری گاوس- سایدل، یعنی $G_{\mathrm{GS}} = (D-E)^{-۱}F$، باشد. در این صورت

$$\begin{aligned} (D-E)^{-۱}Fx = \lambda x &\iff Fx = \lambda(D-E)x \\ &\iff (D-E-A)x = \lambda(D-E)x \\ &\iff Ax = (\lambda - ۱)(D-E)x. \end{aligned}$$

با ضرب طرفین رابطهٔ اخیر در x^H خواهیم داشت:

$$x^H Ax = (۱ - \lambda)x^H(D - E)x. \qquad (۱۷.۵)$$

ترانهاده هرمیتی طرفین رابطهٔ (۱۷.۵) را محاسبه می‌کنیم:

$$x^H Ax = (۱ - \overline{\lambda})x^H(D - F)x. \qquad (۱۸.۵)$$

ابتدا توجه کنید که $\lambda \neq ۱$ (چرا؟). در نتیجه با استفاده از رابطه‌های (۱۷.۵) و (۱۸.۵)، داریم:

$$\left(\frac{۱}{۱ - \lambda} + \frac{۱}{۱ - \overline{\lambda}}\right) x^H Ax = x^H(D + A)x,$$

یا

$$\left(\frac{۱}{۱ - \lambda} + \frac{۱}{۱ - \overline{\lambda}} - ۱\right) x^H Ax = x^H Dx.$$

از اینکه $x^H Dx > ۰$ و $x^H Ax > ۰$، از رابطهٔ اخیر نتیجه می‌گیریم که

$$\frac{۱}{۱ - \lambda} + \frac{۱}{۱ - \overline{\lambda}} - ۱ > ۰,$$

که این خود معادل با

$$\frac{۱ - |\lambda|^۲}{۱ - ۲\Re(\lambda) + |\lambda|^۲} > ۰. \qquad (۱۹.۵)$$

است. اما داریم:

$$\begin{aligned} ۱ - ۲\Re(\lambda) + |\lambda|^۲ &= ۱ - ۲\Re(\lambda) + \Re(\lambda)^۲ + \Im(\lambda)^۲ \\ &= (۱ - \Re(\lambda))^۲ + \Im(\lambda)^۲ > ۰, \end{aligned}$$

و لذا از رابطهٔ (۱۹.۵) خواهیم داشت $۱ - |\lambda|^۲ > ۰$، یا $|\lambda| < ۱$. با توجه به اینکه λ یک مقدار ویژهٔ دلخواه ماتریس G_{GS} می‌باشد، خواهیم داشت $\rho(G_{\mathrm{GS}}) < ۱$ که همگرایی روش را نشان می‌دهد. □

مثال ۱۰.۵ دستگاه معادلات خطی مثال ۹.۵ را در نظر بگیرید. ماتریس تکرار روش گاوس- سایدل برای این دستگاه به صورت

$$G_{\mathrm{GS}} = \frac{۱}{۷۳۵}\begin{pmatrix} ۰ & ۴۹۰ & ۱۹۶ \\ ۰ & ۳۵۰ & -۱۷۵ \\ ۰ & -۲۰ & ۲۶۲ \end{pmatrix},$$

است. مجموعه مقادیر ویژهٔ ماتریس G_{GS} برابر با $\sigma(G_{\mathrm{GS}}) = \{۰, ۰.۵۱۶۶, ۰.۳۱۶\}$ است. بنابراین $\rho(G_{\mathrm{GS}}) = ۰.۵۱۶۶$ که همگرایی روش را نشان می‌دهد. اگر روش گاوس- سایدل را برای حل دستگاه با حدس اولیهٔ صفر به کار ببریم در تکرار دهم خواهیم داشت:

$$x^{(۱۰)} = \begin{pmatrix} ۱.۰۰۰۶ \\ ۱.۰۰۰۵ \\ ۰.۹۹۹۹ \end{pmatrix}.$$

لازم به ذکر است که جواب واقعی دستگاه $x^* = (۱, ۱, ۱)^T$ است.

مثال ۱۱.۵ دستگاه معادلات خطی $Ax = b$ را در نظر بگیرید که در آن

$$A = \begin{pmatrix} ۲ & ۱ & & & \\ ۱ & ۲ & ۱ & & \\ & \ddots & \ddots & \ddots & \\ & & ۱ & ۲ & ۱ \\ & & & ۱ & ۲ \end{pmatrix} \in \mathbb{R}^{n\times n}, \qquad b = \begin{pmatrix} ۳ \\ ۴ \\ \vdots \\ ۴ \\ ۳ \end{pmatrix} \in \mathbb{R}^n.$$

با استفاده از رابطهٔ (۲۵.۴) مقادیر ویژهٔ ماتریس A عبارتند از

$$\begin{aligned} \lambda_j &= ۲ + ۲ \times ۱ \times \sqrt{\frac{۱}{۱}} \cos \frac{j\pi}{n+۱} \\ &= ۲(۱ + \cos \frac{j\pi}{n+۱}) \\ &= ۴ \cos^۲ \frac{j\pi}{۲(n+۱)}, \qquad j = ۱, ۲, \ldots, n. \end{aligned}$$

این رابطه نشان می دهد که مقادیر ویژهٔ ماتریس A مثبت هستند. لذا با استفاده از قضیهٔ ۱۰.۴ و اینکه ماتریس A متقارن است نتیجه می‌گیریم که ماتریس A معین مثبت متقارن است. از این‌رو با توجه به قضیهٔ ۴.۵ روش تکراری گاوس- سایدل برای حل دستگاه همگرا است.

ماتریس تکرار روش ژاکوبی برابر است با

$$G_{\rm J} = D^{-1}(E+F) = -\frac{1}{2}\begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{pmatrix}.$$

دوباره با استفاده از رابطهٔ (۲۵.۴) مقادیر ویژهٔ ماتریس $G_{\rm J}$ به صورت

$$\lambda_j = -\cos\frac{j\pi}{n+1}, \quad j = 1, \ldots, n,$$

خواهند بود. بنابراین

$$\rho(G_{\rm J}) = \cos\frac{\pi}{n+1} < 1.$$

این رابطه نشان می‌دهد که روش تکراری ژاکوبی نیز همگرا است.

دستگاه $Ax = b$ را به ازای n های مختلف با استفاده از روش‌های تکراری ژاکوبی و گاوس-سایدل حل کرده و تعداد تکرارها و زمان‌های مربوطه را در جدول ۲.۵ ارائه کرده‌ایم. در این جدول تعداد تکرارها و زمان اجرای روش (به ثانیه و داخل پرانتز) داده شده‌اند. تمام مفروضات همانند مثال ۶.۵ می‌باشند. این جدول به خوبی نشان می‌دهد که هر دو روش همگرا هستند. با این وجود می‌بینیم که حل این دستگاه به سادگی دستگاه مثال ۶.۵ نیست.

جدول ۲.۵: نتایج عددی برای مثال ۱۱.۵.

n	۱۰	۲۰	۳۰	۴۰
ژاکوبی	۳۵۰(۰.۰۷)	۱۲۸۵(۰.۵۱)	۲۸۰۴(۱.۶۶)	۴۹۰۷(۳.۹۰)
گاوس - سایدل	۱۰۰(۰.۰۲)	۲۴۷(۰.۰۹)	۳۸۲(۰.۲۰)	۴۷۶(۰.۳۴)

ممکن است تصور شود که همگرایی روش گاوس-سایدل همیشه سریعتر از روش ژاکوبی است. مثال بعد نشان می‌دهد که ممکن است روش ژاکوبی برای حل یک دستگاه همگرا باشد، ولی روش گاوس-سایدل واگرا باشد.

مثال ۱۲.۵ فرض کنید:

$$A=\begin{pmatrix} -۱ & ۰ & -۱ \\ -۱ & ۱ & ۰ \\ ۱ & ۲ & -۳ \end{pmatrix}.$$

در این صورت ماتریس تکرار روش ژاکوبی به صورت

$$G_{\mathrm{J}}=\begin{pmatrix} ۰ & ۰ & -۱ \\ ۱ & ۰ & ۰ \\ \frac{۱}{۳} & \frac{۲}{۳} & ۰ \end{pmatrix},$$

و مجموعه مقادیر ویژهٔ آن به صورت $\sigma(G_{\mathrm{J}}) = \{۰.۳۷۳۷ \pm ۰.۸۶۷۴i, -۰.۷۴۷۴\}$ است. بنابراین

$$\rho(G_{\mathrm{J}}) = |۰.۳۷۳۷ \pm ۰.۸۶۷۴i| = ۰.۹۴۴۴,$$

که نشان می‌دهد که روش ژاکوبی با هر حدس اولیهٔ دلخواه همگرا است. از طرفی ماتریس تکرار روش گاوس- سایدل به صورت

$$G_{\mathrm{GS}}=\begin{pmatrix} ۰ & ۰ & -۱ \\ ۰ & ۰ & -۱ \\ ۰ & ۰ & -۱ \end{pmatrix},$$

و مقادیر ویژهٔ آن به صورت $\sigma(G_{\mathrm{GS}}) = \{۰, -۱\}$ می‌باشد. از این‌رو $\rho(G_{\mathrm{GS}}) = ۱$ که نشان می‌دهد روش گاوس- سایدل واگرا است.

تمرینات

۱.۵) نشان دهید که روش گاوس- سایدل برای حل دستگاه معادلات خطی (۶.۵) با هر حدس اولیهٔ دلخواه همگرا است.

۲.۵) دستگاه معادلات خطی

$$\begin{cases} ۴x_۱ - x_۲ + x_۳ = ۷, \\ ۴x_۱ - ۸x_۲ + x_۳ = -۲, \\ -۲x_۱ + x_۲ + ۵x_۳ = ۱۵, \end{cases}$$

را در نظر بگیرید. نشان دهید که روش‌های ژاکوبی و گاوس- سایدل برای حل این دستگاه با هر حدس اولیهٔ دلخواه همگرا است. سپس با چهار تکرار از هرکدام از این دو روش، یک تقریب برای جواب دستگاه بیابید.

۳.۵) (الف) نشان دهید روش تکراری ژاکوبی برای حل دستگاه معادلات خطی $Ax=b$ که در آن

$$A=\begin{pmatrix} ۲ & ۱ & -۲ \\ ۱ & ۱ & ۱ \\ ۳ & ۲ & ۱ \end{pmatrix},$$

همگرا و روش تکراری گاوس- سایدل واگرا است.

(ب) نشان دهید روش تکراری گاوس- سایدل برای حل دستگاه معادلات خطی $Ax=b$ که در آن

$$A=\begin{pmatrix} ۲ & ۱ & ۳ \\ ۱ & ۲ & ۱ \\ ۱ & ۱ & ۳ \end{pmatrix},$$

همگرا و روش تکراری ژاکوبی واگرا است.

۴.۵) نشان دهید که هر دو روش تکراری ژاکوبی و گاوس- سایدل برای حل دستگاه معادلات خطی (۶.۱) همگرا است.

۵.۵) نشان دهید که روش تکراری ژاکوبی برای حل دستگاهی که ماتریس ضرایب آن یک ماتریس معین مثبت متقارن ۲×۲ است، با هر حدس اولیهٔ دلخواه همگرا است.

۶.۵) حدود α را طوری بیابید که هریک از روش‌های تکراری ژاکوبی و گاوس- سایدل برای حل دستگاه معادلات خطی

$$\begin{pmatrix} ۵ & ۲ \\ ۱ & \alpha \end{pmatrix}\begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix}=\begin{pmatrix} ۵ \\ ۱۰ \end{pmatrix},$$

همگرا باشد.

۷.۵) حدود α را طوری بیابید که هریک از روش‌های تکراری ژاکوبی و گاوس- سایدل برای حل دستگاه معادلات خطی

$$\begin{pmatrix} ۱ & -\alpha \\ -\alpha & ۱ \end{pmatrix}\begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix}=\begin{pmatrix} ۱-\alpha \\ ۱-\alpha \end{pmatrix},$$

همگرا باشد.

۸.۵) فرض کنید

$$A = \begin{pmatrix} ۶\alpha & ۱ & ۱ \\ ۱ & -۳\alpha & ۴ \\ ۱ & ۱ & -۲\alpha \end{pmatrix}, \quad b = \begin{pmatrix} ۱ \\ ۲ \\ ۳ \end{pmatrix}.$$

که در آن $\alpha \in \mathbb{R}$. نشان دهید شرط لازم و کافی برای آن که روش تکراری ژاکوبی برای حل دستگاه $Ax = b$ همگرا باشد آن است که $\alpha > \frac{۵}{۶}$.

۹.۵) دستگاه $Ax = b$ را در نظر بگیرید که درآن

$$A = \text{tridiag}_{۷}(-۲, ۴, -۲), \quad b = (۲, ۰, \ldots, ۰, ۲)^T \in \mathbb{R}^{۷}.$$

نشان دهید ماتریس A معین مثبت متقارن است و نتیجه بگیرید که روش گاوس- سایدل برای حل این دستگاه همگرا است. ماتریس تکرار روش تکراری ژاکوبی را نوشته و مقادیر ویژهٔ آن را محاسبه کنید. سپس نشان دهید که روش ژاکوبی نیز برای حل این دستگاه همگرا است. با سه تکرار هرکدام از این روش‌ها یک جواب تقریبی برای دستگاه بیابید.

۱۰.۵) به ازای چه مقادیری از α روش‌های تکراری ژاکوبی و گاوس- سایدل برای حل دستگاه

$$\begin{pmatrix} ۱ & \alpha & \alpha \\ \alpha & ۱ & \alpha \\ \alpha & \alpha & ۱ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = \begin{pmatrix} ۲+\alpha \\ ۲+\alpha \\ ۲+\alpha \end{pmatrix},$$

همگرا است.

۱۱.۵) دستگاه $Ax = b$ را در نظر بگیرید که در آن

$$A = \begin{pmatrix} -۳ & ۶ & ۲ \\ ۱ & ۱ & ۴ \\ ۵ & ۲ & -۱ \end{pmatrix}, \quad b = \begin{pmatrix} ۵ \\ ۶ \\ ۶ \end{pmatrix}.$$

یک ماتریس جایگشت مثل P بیابید به طوری همگرایی روش‌های تکراری ژاکوبی و گاوس- سایدل برای حل دستگاه $PAx = Pb$ تضمین شود.

فصل ۶

مسألهٔ کمترین توان‌های دوم

مسائل مختلفی در علوم و مهندسی وجود دارند که برای مدل‌بندی آنها، تعداد مشاهدات بیشتر از تعداد پارامترهای مدل می‌باشند، اما مشاهدات همراه با خطا می‌باشند. بنابراین هدف این است که پارامترها با دقت بیشتری تقریب زده شوند. یکی از این مسائل، مسألهٔ کمترین توان‌های دوم است. در این فصل مسألهٔ کمترین توان‌های دوم که یکی از مسایل کاربردی است را مطالعه می‌نماییم. در بخش بعد دو مثال برای روشن شدن مسأله و در بخش‌های بعدی روش‌هایی برای حل مسأله ارائه می‌نماییم.

۱.۶ دو مثال مقدماتی

در این بخش دو مثال ارائه می‌نماییم که به حل مسألهٔ کمترین توان‌های دوم منجر می‌شود.

مثال ۱.۶ می‌خواهیم مقاومت یک قطعه سیم را اندازه‌گیری نماییم. برای این کار، جریانی را از سیم عبور داده و ولتاژ دو سر سیم را با استفاده از ولت‌سنج اندازه‌گیری می‌نماییم. همان‌طور که می‌دانیم بین ولتاژ دو سر سیم (V)، مقاومت سیم (R) و شدت جریان (I) عبور داده شده از سیم، رابطهٔ $V = RI$ برقرار است. این آزمایش را چهار بار تکرار می‌کنیم. نتایج به دست آمده در جدول ۱.۶ گزارش شده‌اند. در آزمایش اول مقدار مقاومت سیم ۱۰.۵۶ اهم به دست آمده است. اما در آزمایش‌های بعدی مقدار به دست آمده برای

مقاومت سیم کمی متفاوت است. این تفاوت مربوط به خطای اندازه‌گیری می‌باشد. حال می‌خواهیم ببینیم که کدام مقدار مورد قبول است.

جدول ۱.٦: داده‌های مثال ۱.٦.

i	I_i	V_i	$R_i = \frac{V_i}{I_i}$
۱	۰.۱۸	۱.۹	۱۰.۵٦
۲	۰.۳۱	۳.۰	۹.٦۸
۳	۰.٦۱	۵.۹	۹.٦۷
۴	۰.٦۹	۷.۰	۱۰.۱۴

با توجه به اینکه واقعاً نمی‌دانیم کدام آزمایش دقیق‌تر می‌باشد، بهتر است از تمام اطلاعات به دست آمده استفاده کنیم. برای این کار، قرار می‌دهیم:

$$A = \begin{pmatrix} ۰.۱۸ \\ ۰.۳۱ \\ ۰.٦۱ \\ ۰.٦۹ \end{pmatrix}, \qquad b = \begin{pmatrix} ۱.۹ \\ ۳.۰ \\ ۵.۹ \\ ۷.۰ \end{pmatrix}.$$

با توجه به رابطهٔ $V = RI$، عدد حقیقی x^* را طوری می‌یابیم که بردار Ax^* تا حد امکان به بردار b نزدیک باشد و سپس قرار می‌دهیم $R \approx x^*$. برای این منظور، x^* را طوری می‌یابیم که $\|Ax^* - b\|_۲$ کمترین مقدار ممکن در مجموعهٔ $\{\|Ax - b\|_۲ : x \in \mathbb{R}\}$ را داشته باشد. به عبارت دیگر، یک $x^* \in \mathbb{R}$ می‌یابیم به طوری که

$$\|Ax^* - b\|_۲ = \min_{x\in\mathbb{R}} \|Ax - b\|_۲. \qquad (۱.٦)$$

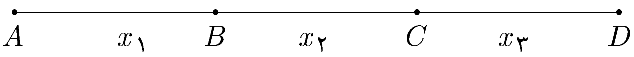
مسألهٔ (۱.٦) یک مسألهٔ کمترین توان‌های دوم است. قرار می‌دهیم $s(x) = \|Ax - b\|_۲^۲$. در نتیجه حل مسألهٔ (۱.٦)، معادل با مینیمم کردن تابع

$$s(x) = (۰.۱۸x - ۱.۹)^۲ + (۰.۳۱x - ۳)^۲ + (۰.٦۱x - ۵.۹)^۲ + (۰.٦۹x - ۷)^۲,$$

است. مشتق این تابع برابر با $s'(x) = ۱.۹۵۳۴x - ۱۹.۴۰۲$ و تنها ریشهٔ آن برابر با $x^* = ۹.۹۳۲۴$ است. از طرفی با توجه به اینکه $s''(x) = ۱.۹۵۳۴ > ۰$، نتیجه می‌گیریم

که مینیمم s به ازای x^* به دست می‌آید. از این‌رو می‌توان نتیجه گرفت که رابطهٔ بین شدت جریان و ولتاژ دو سر سیم تقریباً به صورت $V = ۹.۹۳I$ است. مثال بعدی نشان می‌دهد که حل مسألهٔ کمترین توان‌های دوم همیشه به سادگی این مثال نیست.

مثال ۲.۶ این مثال از [۴]، صفحهٔ ۲۶۲ استخراج شده است. شکل زیر جاده‌ای مستقیم از A به D را نشان می‌دهد.



فرض کنید پنج اندازه‌گیری (بر حسب متر) به صورت

$$AD = ۸۹, \quad AC = ۶۷, \quad BD = ۵۳, \quad AB = ۳۵, \quad CD = ۲۰,$$

انجام شده است. می‌خواهیم $x_۱ = AB$، $x_۲ = BC$ و $x_۳ = CD$ را محاسبه کنیم. با توجه به مقادیر به دست آمده داریم:

$$\begin{cases} x_۱ + x_۲ + x_۳ &= ۸۹, \\ x_۱ + x_۲ &= ۶۷, \\ x_۲ + x_۳ &= ۵۳, \\ x_۱ &= ۳۵, \\ x_۳ &= ۲۰, \end{cases} \tag{۲.۶}$$

که معادل با دستگاه $Ax = b$، که درآن

$$A = \begin{pmatrix} ۱ & ۱ & ۱ \\ ۱ & ۱ & ۰ \\ ۰ & ۱ & ۱ \\ ۱ & ۰ & ۰ \\ ۰ & ۰ & ۱ \end{pmatrix}, \qquad b = \begin{pmatrix} ۸۹ \\ ۶۷ \\ ۵۳ \\ ۳۵ \\ ۲۰ \end{pmatrix}, \qquad x = \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix}.$$

با حل سه معادلهٔ اول از دستگاه (۲.۶) خواهیم داشت $x_۱ = ۳۶$، $x_۲ = ۳۱$ و $x_۳ = ۲۲$. اما می‌بینیم که مقادیر به دست آمده برای $x_۱$ و $x_۳$ با مقادیری که از دو معادلهٔ آخر به دست می‌آیند در تناقض است. این نتیجه نشان می‌دهد که دستگاه (۲.۶) جواب ندارد. دلیل ظهور چنین اتفاقی به خطاهای اندازه‌گیری برمی‌گردد.

حال با توجه به اینکه امکان پیداکردن جواب برای دستگاه (۲.٦) وجود ندارد به دنبال یک جواب تقریبی هستیم به طوری که تا حد امکان طرف راست و چپ دستگاه $Ax = b$ نزدیک به هم باشند. یعنی برداری مثل x^* می‌یابیم به طوری که

$$\|Ax^* - b\|_2 = \min_{x\in\mathbb{R}^3} \|Ax - b\|_2.$$

همانند مثال قبل، قرار می‌دهیم $s(x) = \|Ax - b\|_2^2$. داریم:

$$\begin{aligned} s(x) &= (x_1 + x_2 + x_3 - 89)^2 + (x_1 + x_2 - 67)^2 + (x_2 + x_3 - 53)^2 \\ &\quad + (x_1 - 35)^2 + (x_3 - 20)^2, \end{aligned}$$

و در نتیجه

$$\nabla s = \begin{pmatrix} \dfrac{\partial s}{\partial x_1} \\ \dfrac{\partial s}{\partial x_2} \\ \dfrac{\partial s}{\partial x_3} \end{pmatrix} = 2\begin{pmatrix} 3x_1 + 2x_2 + x_3 - 191 \\ 2x_1 + 3x_2 + 2x_3 - 209 \\ x_1 + 2x_2 + 3x_3 - 162 \end{pmatrix}.$$

برای محاسبهٔ مینیمم s کافی است جواب‌های دستگاه معادلات خطی $\nabla s = (0,0,0)^T$ را به دست آوریم. با توجه به رابطهٔ اخیر محاسبهٔ جواب‌های $\nabla s = (0,0,0)^T$ معادل با حل دستگاه معادلات خطی

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 191, \\ 2x_1 + 3x_2 + 2x_3 = 209, \\ x_1 + 2x_2 + 3x_3 = 162, \end{cases} \tag{3.6}$$

است. تنها جواب این دستگاه

$$x^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 35.125 \\ 32.5 \\ 20.625 \end{pmatrix}.$$

از طرفی داریم:

$$\nabla^2 s = \begin{pmatrix} \frac{\partial^2 s}{\partial x_1^2} & \frac{\partial^2 s}{\partial x_1 \partial x_2} & \frac{\partial^2 s}{\partial x_1 \partial x_3} \\ \frac{\partial^2 s}{\partial x_2 \partial x_1} & \frac{\partial^2 s}{\partial x_2^2} & \frac{\partial^2 s}{\partial x_2 \partial x_3} \\ \frac{\partial^2 s}{\partial x_3 \partial x_1} & \frac{\partial^2 s}{\partial x_3 \partial x_2} & \frac{\partial^2 s}{\partial x_3^2} \end{pmatrix} = \begin{pmatrix} 6 & 4 & 2 \\ 4 & 6 & 4 \\ 2 & 4 & 6 \end{pmatrix}.$$

با توجه به اینکه $\nabla^2 s$ معین مثبت متقارن است (چرا؟)، نقطه x^* مینیمم s را می‌دهد.

در بخش بعد، ضمن معرفی شکل کلی مسألهٔ کمترین توان‌های دوم، روش‌های کاراتری برای حل آن ارائه می‌نماییم.

۲.۶ مسألهٔ کمترین توان‌های دوم و معادلات نرمال

با توجه به دو مثال ارائه شده در بخش قبل، فرض کنید $A \in \mathbb{R}^{m\times n}$ و $b \in \mathbb{R}^n$. در این صورت، مسألهٔ کمترین توان‌های دوم عبارت است از محاسبهٔ یک $x^* \in \mathbb{R}^n$ به طوری که

$$\|Ax^* - b\|_2 = \min_{x\in\mathbb{R}^n} \|Ax - b\|_2. \qquad (۴.۶)$$

قضیه ۱.۶ بردار x^* جواب مسألهٔ کمترین توانهای دوم (۴.۶) است اگر و تنها اگر x^* یک جواب دستگاه معادلات نرمال $A^TAx = A^Tb$ باشد.

برهان: فرض کنید x^* جواب دستگاه معادلات نرمال باشد و $f(x) = \|Ax - b\|_2^2$. در این صورت، به ازای هر $\delta \in \mathbb{R}^n$، داریم:

$$\begin{aligned} f(x^*+\delta) &= \|A(x^*+\delta) - b\|_2^2 \\ &= (Ax^* - b + A\delta)^T(Ax^* - b + A\delta) \\ &= \|Ax^* - b\|_2^2 + 2(A\delta)^T(Ax^* - b) + \|A\delta\|_2^2. \end{aligned}$$

در نتیجه از اینکه $A^TAx^* = A^Tb$، می‌بینیم که

$$f(x^*+\delta) - f(x^*) = 2\delta^T(A^TAx^* - A^Tb) + \|A\delta\|_2^2 = \|A\delta\|_2^2 \geqslant 0.$$

این رابطه نشان می‌دهد که x^* تابع f را مینیمم می‌کند.

حال فرض کنید x^* یک جواب مسألهٔ (۴.۶) باشد. در این صورت، داریم:

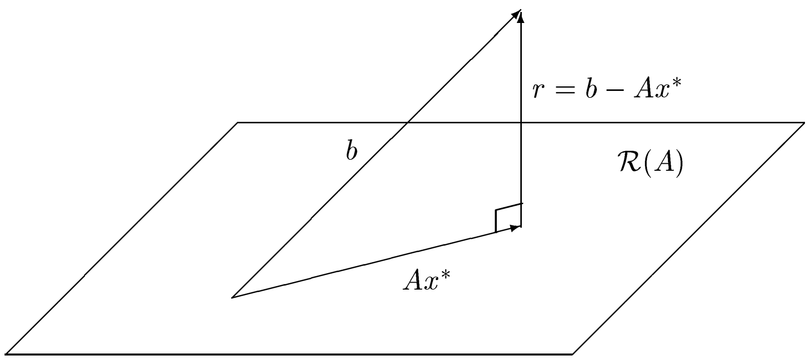
$$f(x^*+\delta) - f(x^*) = 2\delta^TA^Tr + \|A\delta\|_2^2 \geqslant 0 \qquad \forall \delta \in \mathbb{R}^n,$$

که در آن $r = Ax^* - b$. قرار می‌دهیم $\delta = -\alpha A^Tr$ که در آن $\alpha > 0$. در نتیجه

$$f(x^*+\delta) - f(x^*) = -2\alpha\|A^Tr\|_2^2 + \alpha^2\|AA^Tr\|_2^2 \geqslant 0.$$

وقتی که α عدد کوچکی است این رابطه نمی‌تواند درست باشد، مگر اینکه $A^Tr = ۰$، که اثبات قضیه را کامل می‌کند. $\square$

با توجه به قضیهٔ ۱.٦ می‌بینیم که x^* جواب مسألهٔ کمترین توان‌های دوم است اگر و تنها اگر $A^Tr = ۰$ به طوری که $r = b - Ax^*$. رابطهٔ $A^Tr = ۰$ نشان می‌دهد که بردار r بر $\mathcal{R}(A)$ عمود است. توجه کنید که بردار Ax به ازای هر x در $\mathcal{R}(A)$ قرار دارد و x^* برداری است که فاصله b تا Ax^* کمترین فاصلهٔ ممکن را دارد، زیرا r بر $\mathcal{R}(A)$ عمود است و فاصلهٔ عمود کوتاه ترین طول را دارد. شکل ۱.٦ این موضوع را به خوبی نشان می‌دهد.



شکل ۱.٦: تعبیر هندسی قضیهٔ ۱.٦.

قضیهٔ ۱.٦ بیان می‌کند که برای حل مسألهٔ کمترین توانهای دوم کافی است معادلات نرمال $A^TAx = A^Tb$ را حل کنیم. اما ممکن است ماتریس A^TA منفرد باشد. منفرد بودن A^TA وابسته به این است که ماتریس A با رتبهٔ کامل است یا نه. بنابراین این دو حالت را به طور جداگانه بررسی می‌نماییم.

۳.٦ مسألهٔ کمترین توان‌های دوم در حالت رتبهٔ کامل

فرض کنید $A \in \mathbb{R}^{m\times n}$ که در آن $m \geqslant n$. همچنین، فرض کنید رتبهٔ ماتریس A کامل باشد، یعنی $\text{rank}(A) = n$. در این صورت ماتریس A^TA معین مثبت متقارن خواهد بود،

زیرا اولاً متقارن است و ثانیاً به ازای هر $0 \neq x \in \mathbb{R}^n$، داریم:

$$x^T A^T A x = (Ax)^T (Ax) = \|Ax\|_2^2 > 0.$$

توجه کنید وقتی که رتبهٔ ماتریس A کامل است داریم $Ax \neq 0$. در این صورت، جواب دستگاه معادلات نرمال $A^T A x = A^T b$ به صورت

$$x^* = (A^T A)^{-1} A^T b, \tag{5.6}$$

خواهد بود.

با توجه به رابطهٔ (۵.۶) دو مثال ارائه شده در بخش ۱.۶ را حل می‌کنیم. ابتدا مثال ۱.۶ را در نظر می‌گیریم. با توجه به اینکه $\mathrm{rank}(A) = 1$ می‌توانیم از رابطهٔ (۵.۶) استفاده کنیم. داریم:

$$A^T A = 0.9767, \qquad A^T b = 9.701.$$

بنابراین

$$x^* = (A^T A)^{-1} A^T b = \frac{9.701}{0.9767} = 9.9324,$$

و ملاحظه می‌کنیم که این جواب با جواب به دست آمده در مثال ۱.۶ یکی است.

حال مثال ۲.۶ را در نظر می‌گیریم. داریم $\mathrm{rank}(A) = 3$ و

$$A^T A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}, \quad A^T b = \begin{pmatrix} 191 \\ 209 \\ 162 \end{pmatrix}.$$

بنابراین دستگاه معادلات نرمال متناظر، به صورت

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 191, \\ 2x_1 + 3x_2 + 2x_3 = 209, \\ x_1 + 2x_2 + 3x_3 = 162, \end{cases}$$

خواهد بود که با دستگاه (۳.۶) یکسان است و در نتیجه همان جواب به دست آمده در مثال ۲.۶ نتیجه می‌شود.

روش دیگر برای حل مسألهٔ در حالتی که ماتریس A با رتبهٔ کامل می‌باشد، استفاده از تجزیهٔ چولسکی ماتریس $A^T A$ است. فرض کنید $A = LL^T$ تجزیهٔ چولسکی ماتریس A

باشد. در این صورت ابتدا دستگاه $Ly = A^Tb$ با استفاده جایگذاری پیشرو و سپس دستگاه $L^Tx = y$ را با استفاده از جایگذاری پسرو حل می‌نماییم. مثال ۲.٦ را با استفاده از این روش حل می‌کنیم. با استفاده الگوریتم چولسکی (الگوریتم ۳.۳) ماتریس L برابر با

$$L = \begin{pmatrix} ۱.۷۳۲۱ & ۰ & ۰ \\ ۱.۱۵۴۷ & ۱.۲۹۱۰ & ۰ \\ ۰.۵۷۷۴ & ۱.۰۳۲۸ & ۱.۲٦۴۹ \end{pmatrix},$$

خواهد بود. با حل دستگاه

$$\begin{pmatrix} ۱.۷۳۲۱ & ۰ & ۰ \\ ۱.۱۵۴۷ & ۱.۲۹۱۰ & ۰ \\ ۰.۵۷۷۴ & ۱.۰۳۲۸ & ۱.۲٦۴۹ \end{pmatrix} \begin{pmatrix} y_۱ \\ y_۲ \\ y_۳ \end{pmatrix} = A^Tb = \begin{pmatrix} ۱۹۱ \\ ۲۰۹ \\ ۱٦۲ \end{pmatrix},$$

داریم $y = (۱۱۰.۲۷۳۹, ٦۳.۲۵۸۷, ۲٦.۰۸۸۸)^T$. حال با حل دستگاه $L^Tx = y$،

$$\begin{pmatrix} ۱.۷۳۲۱ & ۱.۱۵۴۷ & ۰.۵۷۷۴ \\ ۰ & ۱.۲۹۱۰ & ۱.۰۳۲۸ \\ ۰ & ۰ & ۱.۲٦۴۹ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = y = \begin{pmatrix} ۱۱۰.۲۷۳۹ \\ ٦۳.۲۵۸۷ \\ ۲٦.۰۸۸۸ \end{pmatrix},$$

خواهیم داشت $x = (۳۵.۱۲۵, ۳۲.۵, ۲۰.٦۲۵)^T$.

با توجه به اینکه محاسبهٔ A^TA و A^Tb پرهزینه است و اینکه دستگاه نرمال معمولاً بدحالت است ([۵، ۲] را ببینید)، در ادامه یک روش کاراتر برای حل دستگاه نرمال ارائه می‌کنیم. فرض کنید $A = QR$ تجزیهٔ QR ماتریس A باشد که در آن $Q \in \mathbb{R}^{m\times n}$ یک ماتریس متعامد و $R \in \mathbb{R}^{n\times n}$ یک ماتریس بالامثلثی است. توجه کنید که در اینجا R نامنفرد است. در نتیجه داریم:

$$A^TAx = A^Tb \Longleftrightarrow R^TQ^TQRx = R^TQ^Tb \Longleftrightarrow Rx = Q^Tb.$$

از این‌رو، کافی است دستگاه $Rx = Q^Tb$ را با استفاده از جایگذاری پسرو حل کنیم.

با توجه به نکته اشاره شده، مثال ۲.٦ را حل می‌کنیم. تجزیهٔ QR ماتریس A به صورت $A = QR$ است که در آن

$$Q = \begin{pmatrix} ۰.۵۷۷۴ & ۰.۲۵۸۲ & ۰.۳۱٦۲ \\ ۰.۵۷۷۴ & ۰.۲۵۸۲ & -۰.۴۷۴۳ \\ ۰ & ۰.۷۷۴٦ & ۰.۱۵۸۱ \\ ۰.۵۷۷۴ & -۰.۵۱٦۴ & ۰.۱۵۸۱ \\ ۰ & ۰ & ۰.۷۹۰٦ \end{pmatrix},$$

$$R = \begin{pmatrix} ۱.۷۳۲۱ & ۱.۱۵۴۷ & ۰.۵۷۷۴ \\ ۰ & ۱.۲۹۱۰ & ۱.۰۳۲۸ \\ ۰ & ۰ & ۱.۲۶۴۹ \end{pmatrix}.$$

بنابراین کافی است دستگاه

$$\begin{pmatrix} ۱.۷۳۲۱ & ۱.۱۵۴۷ & ۰.۵۷۷۴ \\ ۰ & ۱.۲۹۱۰ & ۱.۰۳۲۸ \\ ۰ & ۰ & ۱.۲۶۴۹ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \\ x_۳ \end{pmatrix} = Q^T b = \begin{pmatrix} ۱۱۰.۲۷۳۹ \\ ۶۳.۲۵۸۷ \\ ۲۶.۰۸۸۸ \end{pmatrix},$$

را حل کنیم. با حل این دستگاه داریم $x^* = (۳۵.۱۲۵, ۳۲.۵, ۲۰.۶۲۵)^T$.

حل مسألهٔ کمترین توان‌های دوم در حالت کلی نیاز به ابزار قوی‌تری دارد که مهمترین آنها تجزیهٔ مقدار تکین (SVD) یک ماتریس است که در بخش بعد به معرفی آن می‌پردازیم.

۴.۶ تجزیهٔ مقدار تکین

فرض کنید $A \in \mathbb{R}^{m\times n}$. قرار می‌دهیم $B = A^T A$. ماتریس B متقارن است و برای هر $x \in \mathbb{R}^n$، داریم:

$$x^T B x = x^T A^T A x = (Ax)^T Ax = \|Ax\|_۲^۲ \geqslant ۰.$$

بنابراین ماتریس B نیمه معین مثبت متقارن است. لذا با توجه به نتیجهٔ ۳.۴ مقادیر ویژهٔ ماتریس B نامنفی‌اند. اگر λ_i، $i = ۱, ۲, \ldots, n$ مقادیر ویژهٔ B باشند، قرار می‌دهیم:

$$\sigma_i = \sqrt{\lambda_i}, \quad i = ۱, ۲, \ldots, n.$$

σ_i ها را مقادیر تکین ماتریس A می‌گویند.

قضیه ۲.۶ (تجزیهٔ مقدار تکین، SVD) فرض کنید $A \in \mathbb{R}^{m\times n}$. در این صورت، ماتریس‌های متعامد $U \in \mathbb{R}^{m\times m}$ و $V \in \mathbb{R}^{n\times n}$ و ماتریس

$$\Sigma = \begin{pmatrix} D_r & ۰ \\ ۰ & ۰ \end{pmatrix} \in \mathbb{R}^{m\times n}, \quad D_r = \mathrm{diag}_r(\sigma_۱, \ldots, \sigma_r),$$

وجود دارند به طوری که $A = U\Sigma V^T$، که در آن $\sigma_۱ \geqslant \sigma_۲ \geqslant \cdots \geqslant \sigma_r \geqslant ۰$ مقادیر تکین A هستند. ستون‌های $U = (u_۱, \ldots, u_m)$ را بردارهای تکین چپ A و ستون‌های ماتریس

$V=(v_1,\ldots,v_n)$ را بردارهای تکین راست A می‌نامند. به‌علاوه اگر σ_r کوچکترین مقدار تکین ناصفر A باشد، آنگاه $\operatorname{rank}(A)=r$.

برهان: ماتریس A^TA یک ماتریس متقارن است و بنا به قضیهٔ ۲.۴ به طور متعامد قطری‌شدنی است و دارای n بردار ویژهٔ یکامتعامد می‌باشد. بنابراین یک ماتریس متعامد مثل V وجود دارد به طوری

$$V^TA^TAV=\operatorname{diag}_n(\lambda_1,\lambda_2,\ldots,\lambda_n). \qquad (۶.۶)$$

با توجه به تعریف مقادیر تکین ماتریس A، داریم $\sigma_i=\sqrt{\lambda_i}$، $i=1,2,\ldots,n$. فرض کنید:

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>0, \qquad \lambda_{r+1}=\cdots=\lambda_n=0.$$

حال از رابطهٔ (۶.۶)، داریم:

$$\begin{aligned} A^TAv_i &= \sigma_i^2v_i, \quad i=1,\ldots,r,\\ A^TAv_i &= 0, \quad i=r+1,\ldots,n. \end{aligned}$$

ابتدا توجه کنید که برای $i=r+1,\ldots,n$، داریم:

$$0=v^TA^TAv_i=\|Av_i\|_2^2.$$

بنابراین $Av_i=0$، $i=r+1,\ldots,n$. قرار می‌دهیم:

$$u_i=\frac{1}{\sigma_i}Av_i, \quad i=1,\ldots,r.$$

لذا

$$\|u_i\|_2^2=\frac{1}{\sigma_i^2}v_i^TA^TAv_i=\frac{1}{\sigma_i^2}v_i^T(\sigma_i^2v_i)=1, \quad i=1,2,\ldots,r.$$

حال بردارهای u_{r+1}، ...، u_m را طوری می‌سازیم که u_i، $i=1,\ldots,m$، یکامتعامد باشند. قرار می‌دهیم $U=(u_1,\ldots,u_m)$ و

$$\Sigma=\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}, \quad D=\operatorname{diag}_r(\sigma_1,\sigma_2,\ldots,\sigma_r),$$

که در آن Σ یک ماتریس $m \times n$ است. در این صورت ماتریس U متعامد است و

$$\begin{aligned} U\Sigma V^T &= (\sigma_۱ u_۱, \sigma_۲ u_۲, \ldots, \sigma_r u_r, ۰, \ldots, ۰)V^T \\ &= (Av_۱, Av_۲, \ldots, Av_r, ۰, \ldots, ۰)V^T \\ &= AVV^T = A. \end{aligned}$$

برای اثبات قسمت آخر قضیهٔ، می‌بینیم که

$$\operatorname{rank}(A) = \operatorname{rank}(U\Sigma V^T) = \operatorname{rank}(\Sigma) = r.$$

در اینجا توجه کنید که U و V ماتریس‌های متعامد (و بنابراین نامنفرد) هستند. $\square$

در ادامه با استفاده از قضیهٔ ۶.۲ روشی برای محاسبهٔ تجزیهٔ مقدار تکین یک ماتریس ارائه می‌دهیم. با توجه به قضیهٔ ۶.۲، داریم:

$$\begin{aligned} A^TA &= V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T \\ \Longrightarrow A^TAV &= \begin{pmatrix} \sigma_۱^۲ & & \\ & \ddots & \\ & & \sigma_n^۲ \end{pmatrix} V. \end{aligned}$$

از این‌رو،

$$A^TAv_i = \sigma_i^۲ v_i, \quad i = ۱, \ldots, n.$$

این رابطه نشان می‌دهد که $(\sigma_i^۲, v_i)$، $i = ۱, \ldots, n$ زوج‌های ویژهٔ A^TA هستند. یادآوری می‌کنیم که A^TA متقارن است و بنا به قضیهٔ ۴.۲ به طور یکانی قطری‌شدنی و دارای n بردار ویژهٔ یکامتعامد می‌باشد. همچنین اگر λ_i، $i = ۱, \ldots, n$ مقادیر ویژهٔ A^TA باشند، آنگاه

$$\sigma_i = \sqrt{\lambda_i}, \quad i = ۱, \ldots, n.$$

لازم به ذکر است که ماتریس A^TA نیمه معین مثبت متقارن است و بنابراین مقادیر ویژهٔ آن، یعنی λ_i ها، نامنفی هستند.

حال کافی است روشی برای محاسبهٔ ماتریس U بیابیم. برای این کار، با استفاده از تجزیهٔ مقدار تکین A، داریم:

$$AV = U\Sigma \Longleftrightarrow Av_i = \sigma_i u_i, \quad i = 1, \ldots, n.$$

در نتیجه

$$\begin{aligned} u_i &= \frac{1}{\sigma_i} A v_i, \quad i = 1, \ldots, r, \\ Av_i &= 0, \quad i = r+1, \quad , n. \end{aligned}$$

با استفاده از رابطهٔ اول می‌توان r ستون اول U را محاسبه کرد. توجه کنید که مجموعهٔ $\{u_1, \ldots, u_r\}$ یک مجموعهٔ یکامتعامد است، زیرا به ازای هر $1 \leqslant i, j \leqslant r$، داریم:

$$\begin{aligned} (u_i, u_j) &= \frac{1}{\sigma_i \sigma_j}(Av_i, Av_j) = \frac{1}{\sigma_i \sigma_j}(v_i, A^T A v_j) = \frac{1}{\sigma_i \sigma_j}(v_i, \sigma_i^2 v_j) \\ &= \frac{\sigma_i}{\sigma_j}(v_i, v_j) = \begin{cases} (v_i, v_i) = \|v_i\|_2^2 = 1, & i = j, \\ 0, & i \neq j. \end{cases} \end{aligned}$$

برای محاسبهٔ ستون‌های u_i، $i = r+1, \ldots, n$ از ماتریس U کافی است با استفاده از الگوریتم گرام- اشمیت این بردارها را طوری محاسبه کنیم که مجموعه بردارهای $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ یک مجموعه یکامتعامد باشد.

مثال ۳.٦ تجزیهٔ مقدار تکین ماتریس

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

را به دست آورید.

حل: داریم:

$$A^T A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

مقادیر ویژهٔ این ماتریس عبارتند از $\lambda_1 = 3$ و $\lambda_2 = 1$. بنابراین مقادیر تکین ماتریس A به صورت $\sigma_1 = \sqrt{3}$ و $\sigma_2 = 1$ خواهند بود. با توجه به اینکه هر دو مقدار تکین ماتریس A ناصفرند، داریم $\text{rank}(A) = 2$. به سادگی می‌توان دید که دو بردار

$$v_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \quad v_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix},$$

بردارهای ویژهٔ یکامتعامد برای ماتریس A^TA هستند. قرار می‌دهیم:

$$V = (v_1, v_2) = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

برای محاسبهٔ U، قرار می‌دهیم:

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{pmatrix},$$

و

$$u_2 = \frac{1}{\sigma_2} A v_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

بردار یکهٔ u_3 را طوری انتخاب می‌کنیم که بر u_1 و u_2 عمود باشد. برای این کار، از فرایند گرام- اشمیت استفاده می‌کنیم. فرض کنید $e_1 = (1, 0, 0)^T$. در این صورت،

$$w_3 = e_1 - (u_1, e_1)u_1 - (u_2, e_1)u_2 = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}.$$

با یکه کردن این بردار، خواهیم داشت:

$$u_3 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \end{pmatrix}.$$

بنابراین

$$U=(u_۱,u_۲,u_۳)=\begin{pmatrix} \frac{\sqrt{۶}}{۳} & ۰ & \frac{\sqrt{۳}}{۳} \\ \frac{\sqrt{۶}}{۶} & -\frac{\sqrt{۲}}{۲} & -\frac{\sqrt{۳}}{۳} \\ \frac{\sqrt{۶}}{۶} & \frac{\sqrt{۲}}{۲} & -\frac{\sqrt{۳}}{۳} \end{pmatrix}.$$

در نتیجه با قرار دادن

$$\Sigma=\begin{pmatrix} \sqrt{۳} & ۰ \\ ۰ & ۱ \\ ۰ & ۰ \end{pmatrix},$$

خواهیم داشت $A=U\Sigma V^T$.

نتیجه ۱.۶ فرض کنید $A = U\Sigma V^T$ تجزیهٔ مقدار تکین ماتریس A و U، V و Σ ماتریس‌هایی باشند که در قضیهٔ ۲.۶ تعریف شده اند. در این صورت

(الف) مجموعهٔ بردارهای $\{u_۱,u_۲,\dots,u_r\}$ یک پایهٔ یکامتعامد برای برد ماتریس A، یعنی $\mathcal{R}(A)=\{Ax : x\in\mathbb{R}^n\}$، است.

(ب) مجموعهٔ بردارهای $\{v_{r+۱},v_{r+۲},\dots,v_n\}$ یک پایهٔ یکامتعامد برای فضای پوچ A، یعنی $\mathcal{N}(A)=\{x\in\mathbb{R}^n : Ax=۰\}$، است.

برهان: با توجه به $AV=U\Sigma$، داریم:

$$Av_i = \sigma_i u_i, \quad i=۱,۲,\dots,r, \tag{۷.۶}$$

$$Av_i = ۰, \quad i=r+۱,\dots,n. \tag{۸.۶}$$

فرض کنید $x\in\mathbb{R}^n$ و $y=Ax$. با توجه به اینکه بردارهای v_i، $i=۱,۲,\dots,n$، یک پایه برای $\mathbb{R}^n$ تشکیل می‌دهند، اعدادی حقیقی مثل α_i، $i=۱,۲,\dots,n$، وجود دارند به طوری که $x=\sum_{i=۱}^{n}\alpha_i v_i$. در نتیجه با استفاده از (۷.۶)، خواهیم داشت:

$$y=Ax=\sum_{i=۱}^{n}\alpha_i A v_i=\sum_{i=۱}^{r}\alpha_i\sigma_i u_i.$$

بنابراین $y\in\operatorname{span}\{u_۱,u_۲,\dots,u_r\}$. از اینکه u_i ها مستقل خطی هستند نتیجه می‌شود که u_i، $i=۱,۲,\dots,r$، یک پایه برای برد A می‌باشند. از طرفی با استفاده از (۸.۶) داریم:

$$v_i \in \mathcal{N}(A), \quad i = r+1, \ldots, n.$$

حال داریم:

$$\dim(\mathcal{N}(A)) = n - \dim(\mathcal{R}(A)) = n - r.$$

با توجه به اینکه بردارهای v_i، $i = r+1, \ldots, n$ بردارهای مستقل خطی در $\mathcal{N}(A)$ و بعد فضای $\mathcal{N}(A)$ برابر با $n-r$ است، این بردارها یک پایه برای $\mathcal{N}(A)$ تشکیل می‌دهند. □

تعریف ۱.۶ فرض کنید $A = U\Sigma V^T$ تجزیهٔ مقدار تکین ماتریس A باشد و رتبهٔ ماتریس A برابر با r باشد. در این صورت، معکوس تعمیم‌یافته مور- پن‌رز ماتریس A را با A^+ نشان می‌دهند و به صورت

$$A^+ = V\Sigma^+ U^T, \tag{۹.۶}$$

تعریف می‌شود که در آن

$$\Sigma^+ = \begin{pmatrix} D_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times m}.$$

نتیجه ۲.۶ اگر ماتریس $A \in \mathbb{R}^{n\times n}$ نامنفرد باشد، آنگاه $A^+ = A^{-1}$.

برهان: فرض کنید $A = U\Sigma V^T$. از اینکه $\operatorname{rank}(A) = \operatorname{rank}(\Sigma) = n$، داریم $\Sigma^+ = \Sigma^{-1}$. بنابراین

$$A^+ = V\Sigma^+ U^T = V\Sigma^{-1}U^T = (U\Sigma V^T)^{-1} = A^{-1}.$$

توجه کنید که $V = (V^T)^{-1}$ و $U^{-1} = U^T$. □

یادداشت ۱.۶ فرض کنید $A \in \mathbb{R}^{m\times n}$ و $\operatorname{rank}(A) = r$. در این صورت، می‌توان تجزیهٔ مقدار تکین ماتریس A را به صورت

$$A = \begin{pmatrix} u_1 & u_2 & \cdots & u_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$
\begin{aligned}
&= \begin{pmatrix}\sigma_۱ u_۱ & \cdots & \sigma_r u_r & ۰ & \cdots & ۰\end{pmatrix}\begin{pmatrix} v_۱^T \\ v_۲^T \\ \vdots \\ v_n^T\end{pmatrix} \\
&= \sigma_۱ u_۱ v_۱^T + \cdots + \sigma_r u_r v_r^T \\
&= \sum_{k=۱}^{r} \sigma_k u_k v_k^T,
\end{aligned}
$$

نوشت. نمایش ماتریس A به صورت $A = \sum_{k=۱}^{r} \sigma_k u_k v_k^T$ را شکل کاهش یافتهٔ تجزیهٔ مقدار تکین A می‌گویند. به طور مشابه، داریم

$$A^+ = \sum_{k=۱}^{r} \frac{۱}{\sigma_k} v_k u_k^T. \tag{۱۰.٦}$$

این رابطه نشان می‌دهد که با استفاده از r مقدار تکین و r تا از بردارهای تکینِ ماتریس A می‌توان ماتریس A^+ را می‌توان تولید کرد. بنابراین اگر r کوچک باشد استفاده از فرمول (۱۰.٦) بسیار مقرون به صرفه خواهد بود.

مثال ۴.٦ معکوس تعمیم یافته مور- پن‌رز ماتریس

$$A = \begin{pmatrix} ۱ & -۱ \\ ۲ & -۲ \\ -۲ & ۲ \end{pmatrix},$$

را بیابید.

حل: داریم:

$$A^T A = \begin{pmatrix} ۹ & -۹ \\ -۹ & ۹ \end{pmatrix},$$

و مقادیر ویژهٔ آن عبارتند از $\lambda_۱ = ۱۸$ و $\lambda_۲ = ۰$. بنابراین $\sigma_۱ = ۳\sqrt{۲}$ و $\sigma_۲ = ۰$. از این‌رو

$$\Sigma = \begin{pmatrix} ۳\sqrt{۲} & ۰ \\ ۰ & ۰ \\ ۰ & ۰ \end{pmatrix}.$$

بردارهای ویژهٔ ماتریس A^TA به صورت

$$v_۱ = \begin{pmatrix} \frac{\sqrt{۲}}{۲} \\ -\frac{\sqrt{۲}}{۲} \end{pmatrix}, \quad v_۲ = \begin{pmatrix} \frac{\sqrt{۲}}{۲} \\ \frac{\sqrt{۲}}{۲} \end{pmatrix},$$

می‌باشند و در نتیجه

$$V = \begin{pmatrix} \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \\ -\frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \end{pmatrix}.$$

برای محاسبهٔ $U = (u_۱, u_۲, u_۳)$، داریم:

$$u_۱ = \frac{۱}{\sigma_۱} A v_۱ = \frac{۱}{۳} \begin{pmatrix} ۱ \\ ۲ \\ -۲ \end{pmatrix}.$$

با استفاده الگوریتم گرام- اشمیت فرض می‌کنیم $x_۲ = (۱, ۱, ۰)^T$ و قرار می‌دهیم:

$$w_۲ = x_۲ - (u_۱^T x_۲) u_۱ = \frac{۱}{۳} \begin{pmatrix} ۲ \\ ۱ \\ ۲ \end{pmatrix}.$$

با توجه به اینکه $\|w_۲\|_۲ = ۱$، خواهیم داشت $u_۲ = w_۲$. برای محاسبهٔ $u_۳$ قرار می‌دهیم $x_۳ = (۰, ۰, ۱)^T$ و

$$w_۳ = x_۳ - (u_۱^T x_۳) u_۱ - (u_۲^T x_۳) u_۲ = \frac{۱}{۹} \begin{pmatrix} -۲ \\ ۲ \\ ۱ \end{pmatrix}.$$

از اینکه $\|w_۳\|_۲ = \frac{۱}{۳}$، داریم:

$$u_۳ = \frac{w_۳}{\|w_۳\|_۲} = \frac{۱}{۳} \begin{pmatrix} -۲ \\ ۲ \\ ۱ \end{pmatrix}.$$

بنابراین

$$U = \frac{۱}{۳} \begin{pmatrix} ۱ & ۲ & -۲ \\ ۲ & ۱ & ۲ \\ -۲ & ۲ & ۱ \end{pmatrix}.$$

لذا

$$\begin{aligned} A^+ &= V\Sigma^+U^T \\ &= \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{3\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \\ &= \frac{1}{3\sqrt{2}} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 1 & 2 & -2 \\ -1 & -2 & 2 \end{pmatrix}. \end{aligned}$$

۵.۶ حل مسألهٔ کمترین توان‌های دوم در حالت کلی

در این بخش به حل مسألهٔ کمترین توان‌های دوم در حالت کلی، یعنی در حالتی که A یک ماتریس دلخواه است، می‌پردازیم. این کار با استفاده از تجزیهٔ مقدار تکین ماتریس A امکان‌پذیر است. فرض کنید $A = U\Sigma V^T$ تجزیهٔ مقدار تکین $A \in \mathbb{R}^{m\times n}$ باشد و $\mathrm{rank}(A) = r \leqslant \min\{m,n\}$. در این صورت، داریم:

$$\begin{aligned} \|Ax-b\|_2^2 &= |U\Sigma V^Tx-b\|_2^2 = \|U^T(U\Sigma V^Tx-b)\|_2^2 \\ &= \|\Sigma V^Tx - U^Tb\|_2^2 = \|\Sigma z-d\|_2^2, \end{aligned}$$

که در آن $z = V^Tx$ و $d = U^Tb$. با فرض $z = (z_1,\dots,z_n)^T$ و $d = (d_1,\dots,d_m)^T$، خواهیم داشت:

$$\|Ax-b\|_2^2 = \|\Sigma z-d\|_2^2 = \left\| \begin{pmatrix} \sigma_1z_1-d_1 \\ \vdots \\ \sigma_rz_r-d_r \\ -d_{r+1} \\ \vdots \\ -d_m \end{pmatrix} \right\|_2^2 = \sum_{i=1}^{r} |\sigma_iz_i-d_i|^2 + \sum_{i=r+1}^{n} |d_i|^2.$$

بنابراین بردار z عبارت اخیر را مینیمم می‌کند اگر و تنها اگر

$$\begin{cases} z_i = \dfrac{d_i}{\sigma_i}, & i = 1,\dots,r, \\ z_i \ \text{دلخواه}, & i = r+1,\dots,n, \end{cases}$$

و به ازای تمام جواب‌های مسألهٔ کمترین توان‌های دوم مقدار $\|Ax - b\|_۲$ برابر با $(\sum_{i=r+۱}^{n} d_i^۲)^{۱/۲}$ خواهد بود. در نتیجه جواب عمومی مسألهٔ کمترین توان‌های دوم به صورت

$$\begin{aligned} x &= Vz = V\begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ z_{r+۱} \\ \vdots \\ z_n \end{pmatrix} = V\begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} + V\begin{pmatrix} ۰ \\ \vdots \\ ۰ \\ z_{r+۱} \\ \vdots \\ z_n \end{pmatrix} \\ &= V\Sigma^+ d + V_۲ w = V\Sigma^+ U^T b + V_۲ w = A^+ b + V_۲ w, \end{aligned}$$

است که در آن $w = (z_{r+۱}, \ldots, z_n)^T$ و $V_۲ = (v_{r+۱}, \ldots, v_n)$. با توجه به قسمت (الف) از نتیجهٔ ۱.٦، داریم:

$$t \in \mathcal{N}(A) \Longleftrightarrow t = V_۲ w, \quad w \in \mathbb{R}^{n-r}.$$

از این‌رو مجموعهٔ جواب‌های مسألهٔ کمترین توان‌های دوم به صورت

$$\mathcal{S} = \{x \in \mathbb{R}^n : x = A^+ b + V_۲ w,\ w \in \mathbb{R}^{n-r}\} = A^+ b + \mathcal{N}(A). \qquad (۱۱.٦)$$

است. از بین مجموعهٔ جواب‌های مسألهٔ کمترین توان‌های دوم جواب $x^* = A^+ b$ که به ازای $w = ۰$ به دست می‌آید از اهمیت ویژه‌ای برخوردار است. برای این جواب خاص، داریم:

$$\begin{aligned} \|x^*\|_۲ &= \left\| V\begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} \right\|_۲ = \left\| \begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ ۰ \\ \vdots \\ ۰ \end{pmatrix} \right\|_۲ < \left\| \begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ z_{r+۱} \\ \vdots \\ z_n \end{pmatrix} \right\|_۲ = \left\| V\begin{pmatrix} \frac{d_۱}{\sigma_۱} \\ \vdots \\ \frac{d_r}{\sigma_r} \\ z_{r+۱} \\ \vdots \\ z_n \end{pmatrix} \right\|_۲ \\ &= \|x\|_۲. \end{aligned}$$

این رابطه نشان می‌دهد که

$$\|x^*\| = \|A^+b\|_۲ < \|x\|_۲ \quad \forall x \in \mathcal{S}.$$

و این یعنی از بین جواب‌های مسألهٔ کمترین توان‌های دوم، $x^* = A^+b$ کمترین نرم-۲ را دارد. این نتایج در قضیهٔ بعدی جمع‌بندی می‌شود.

قضیه ۳.٦ فرض کنید $A \in \mathbb{R}^{m\times n}$. در این‌صورت

$$\|A\bar{x} - b\|_۲ = \min_{x\in\mathbb{R}^n} \|Ax - b\|_۲ \Longleftrightarrow \bar{x} \in \mathcal{S} = A^+b + \mathcal{N}(A).$$

به‌علاوه به ازای هر $\bar{x} \in \mathcal{S}$، داریم $\|x^*\|_۲ < \|\bar{x}\|_۲$ که در آن $\bar{x} \neq x^* = A^+b$.

مثال ۵.٦ مسألهٔ کمترین توان‌های دوم

$$\begin{pmatrix} ۱ & -۱ \\ ۲ & -۲ \\ -۲ & ۲ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix} = \begin{pmatrix} ۱ \\ ۱ \\ ۱ \end{pmatrix},$$

را حل کنید.

حل: با استفاده از مثال ۴.٦ جواب مسأله کمترین توان‌های دوم با کمترین نرم-۲ برابر است با

$$x^* = A^+b = \frac{۱}{۱۸}\begin{pmatrix} ۱ & ۲ & -۲ \\ -۱ & -۲ & ۲ \end{pmatrix} \begin{pmatrix} ۱ \\ ۱ \\ ۱ \end{pmatrix} = \frac{۱}{۱۸}\begin{pmatrix} ۱ \\ -۱ \end{pmatrix}.$$

بنابراین

$$x_۱ = -x_۲ = \frac{۱}{۱۸}.$$

از طرفی به سادگی می‌توان دید که

$$\mathcal{N}(A) = \left\{\alpha \begin{pmatrix} ۱ \\ ۱ \end{pmatrix} : \alpha \in \mathbb{R}\right\}.$$

بنابراین جواب عمومی مسألهٔ کمترین توان‌های دوم به صورت

$$\begin{pmatrix} x_۱(\alpha) \\ x_۲(\alpha) \end{pmatrix} = \begin{pmatrix} \alpha + \frac{۱}{۱۸} \\ \alpha - \frac{۱}{۱۸} \end{pmatrix},$$

است که در آن $\alpha \in \mathbb{R}$.

مثال ٦.٦ مجموعهٔ نقاط

$$\{(۰.۳۲, ۰.۴۵), (۰.۴۵, ۰.۴۱), (۰.۵۳, ۰.٦۱), (۰.٦۲, ۰.۷۱), (۰.۷۲, ۰.۸۰), (۰.۸۵, ۰.۷۴)\},$$

را در نظر بگیرید. روشن است که نمی‌توان یک خط راست مثل $y = a_۰ + a_۱ x$ نوشت به طوری که از تمام این نقاط عبورکند، به عبارت دیگر، دستگاه

$$\begin{cases} a_۰ + ۰.۳۲ a_۱ = ۰.۴۵, \\ a_۰ + ۰.۴۵ a_۱ = ۰.۴۱, \\ a_۰ + ۰.۵۳ a_۱ = ۰.٦۱, \\ a_۰ + ۰.٦۲ a_۱ = ۰.۷۱, \\ a_۰ + ۰.۷۲ a_۱ = ۰.۸۰, \\ a_۰ + ۰.۸۵ a_۱ = ۰.۷۴, \end{cases}$$

یا معادل آن

$$Aa = \begin{pmatrix} ۱ & ۰.۳۲ \\ ۱ & ۰.۴۵ \\ ۱ & ۰.۵۳ \\ ۱ & ۰.٦۲ \\ ۱ & ۰.۷۲ \\ ۱ & ۰.۸۵ \end{pmatrix} \begin{pmatrix} a_۰ \\ a_۱ \end{pmatrix} = \begin{pmatrix} ۰.۴۵ \\ ۰.۴۱ \\ ۰.٦۱ \\ ۰.۷۱ \\ ۰.۸۰ \\ ۰.۷۴ \end{pmatrix} = b,$$

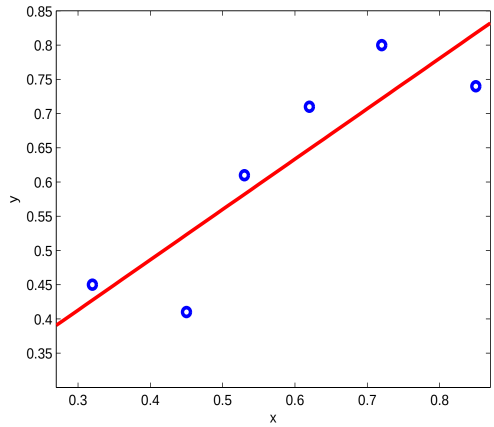
جواب ندارد. بنابراین بردار a را طوری می‌یابیم که Aa تا حد امکان به b نزدیک شود. برای این کار، کافی است $\|Ax - b\|_۲$ را مینیمم کنیم. با توجه به اینکه A یک ماتریس با رتبهٔ کامل است (چرا؟)، با استفاده از رابطهٔ (۵.٦) کافی است دستگاه $A^T A a = A^T b$ را حل کنیم. این دستگاه با

$$\begin{pmatrix} ٦.۰۰۰۰ & ۳.۴۹۰۰ \\ ۳.۴۹۰۰ & ۲.۲۱۱۱ \end{pmatrix} \begin{pmatrix} a_۰ \\ a_۱ \end{pmatrix} = \begin{pmatrix} ۳.۷۲۰۰ \\ ۳.۲۹۷۰ \end{pmatrix},$$

معادل است، که با حل آن داریم $a_۰ = ۰.۱۹۲۱$ و $a_۱ = ۰.۷۳۵٦$. بنابراین خط مورد نظر به صورت

$$y = ۰.۱۹۲۱ + ۰.۷۳۵٦ x,$$

است. در شکل ۲.٦ مجموعه نقاط داده شده و خطی که به داده‌ها برازش شده، نمایش داده شده است.



شکل ۲.٦: نمایش داده‌ها و خط برازش برای مثال ٦.٦.

در ادامه یک فرمول صریح برای معادلهٔ خط راستی که مجموعه‌ای از داده‌ها را برازش کند، می‌یابیم. فرض کنید مجموعه نقاط $\{(x_۱, y_۱), (x_۲, y_۲), \dots, (x_m, y_m)\}$ داده شده باشند. می‌خواهیم یک خط راست مثل $y = a_۰ + a_۱ x$ به این دادها برازش کنیم. در این صورت ماتریس‌های A و b در مسألهٔ کمترین توان‌های دوم (۴.٦) به صورت

$$A = \begin{pmatrix} ۱ & x_۱ \\ ۱ & x_۲ \\ \vdots & \vdots \\ ۱ & x_m \end{pmatrix}, \qquad b = \begin{pmatrix} y_۱ \\ y_۲ \\ \vdots \\ y_m \end{pmatrix},$$

می‌باشند. فرض می‌کنیم حداقل دو مقدار از x_i ها متمایز باشند. لذا $\operatorname{rank}(A) = ۲$ و رتبهٔ ماتریس کامل است. بنابراین با توجه به (۵.٦)، کافی است دستگاه معادلات خطی

$$A^T A \begin{pmatrix} a_۰ \\ a_۱ \end{pmatrix} = A^T b, \qquad (۱۲.٦)$$

را حل کنیم که در آن

$$A^T A = \begin{pmatrix} m & \sum_{i=۱}^{m} x_i \\ \sum_{i=۱}^{m} x_i & \sum_{i=۱}^{m} x_i^۲ \end{pmatrix}, \qquad A^T b = \begin{pmatrix} \sum_{i=۱}^{m} y_i \\ \sum_{i=۱}^{m} x_i y_i \end{pmatrix}.$$

با حل دستگاه معادلات خطی (۱۲.۶)، خواهیم داشت:

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2}, \tag{13.6}$$

$$a_1 = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2}. \tag{14.6}$$

مثال ۷.۶ خط راستی بنویسید که داده‌های جدول ۲.۶ را برازش کند.

جدول ۲.۶: داده‌های مثال ۷.۶.

i	۱	۲	۳
x_i	۰	۱	۲
y_i	۰.۱	۰.۹	۲

داریم $m = 3$ و

$$\begin{aligned}\sum_{i=1}^{3} x_i &= 0+1+2=3, \quad \sum_{i=1}^{3} y_i = 1+0.9+2=3,\\ \sum_{i}^{3} x_i^2 &= 0^2+1^2+2^2=5, \quad \sum_{i=1}^{3} x_i y_i = 0\times 1+1\times 0.9+2\times 2=4.9,\end{aligned}$$

و بنابراین با استفاده از روابط (۱۳.۶) و (۱۴.۶) خواهیم داشت:

$$\begin{aligned}a_0 &= \frac{5\times 3-4.9\times 3}{3\times 5-3^2}=0.05,\\ a_1 &= \frac{3\times 4.9-3\times 3}{3\times 5-3^2}=0.95.\end{aligned}$$

از این‌رو $y = 0.05 + 0.95x$ خط مورد نظر است.

مثال ۸.۶ مجموعهٔ نقاط

$$\{(0.3, 0.4), (0.5, 0.4), (0.7, 0.5), (0.8, 1.1), (0.9, 1.2)\},$$

را در نظر بگیرید. در این مثال می‌خواهیم یک چندجمله‌ای درجهٔ دوم به صورت $p(x) = a_0 + a_1x + a_2x^2$ به این داده‌ها برازش کنیم. همانند مثال قبل، کافی است مسألهٔ کمترین توان‌های دوم متناظر به مسألهٔ

$$\begin{cases} a_0 + 0.3a_1 + 0.3^2a_2 = 0.4, \\ a_0 + 0.5a_1 + 0.5^2a_2 = 0.4, \\ a_0 + 0.7a_1 + 0.7^2a_2 = 0.5, \\ a_0 + 0.8a_1 + 0.8^2a_2 = 1.1, \\ a_0 + 0.9a_1 + 0.9^2a_2 = 1.2, \end{cases}$$

یا معادل آن

$$Aa = \begin{pmatrix} 1 & 0.3 & 0.09 \\ 1 & 0.5 & 0.25 \\ 1 & 0.7 & 0.49 \\ 1 & 0.8 & 0.64 \\ 1 & 0.9 & 0.81 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.5 \\ 1.1 \\ 1.2 \end{pmatrix} = b,$$

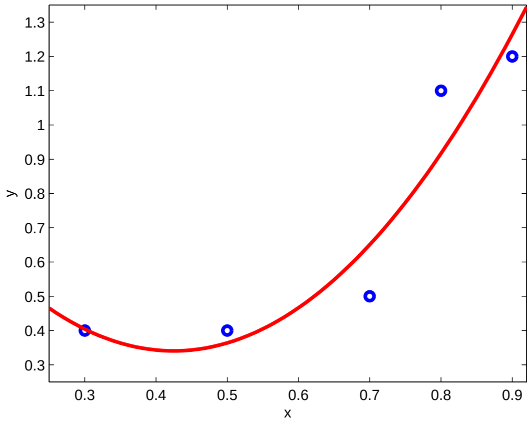
را حل کنیم. با توجه به اینکه رتبهٔ ماتریس ضرایب کامل است، کافی است دستگاه معادلات نرمال $A^TAa = A^Tb$ را حل کنیم. این دستگاه معادل با

$$\begin{pmatrix} 5.0000 & 3.2000 & 2.2800 \\ 3.2000 & 2.2800 & 1.7360 \\ 2.2800 & 1.7360 & 1.3764 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 3.6000 \\ 2.6300 \\ 2.0570 \end{pmatrix},$$

است، که با حل آن داریم $a_0 = 1.0783$، $a_1 = -3.4728$ و $a_2 = 4.0885$ و چندجمله‌ای مورد نظر به صورت

$$p(x) = 1.0783 - 3.4728x + 4.0885x^2,$$

است. در شکل ۳.۶ مجموعه نقاط داده شده و چندجمله‌ای درجهٔ دومی که به داده‌ها برازش شده، نمایش داده شده است.



شکل ۳.٦: نمایش‌داده‌ها و منحنی برازش برای مثال ۸.٦.

٦.٦ فشرده‌سازی تصاویر به کمک SVD

قبل از اینکه بحث اصلی در خصوص فشرده‌سازی تصاویر را شروع کنیم، یک نتیجهٔ مهم در مورد تجزیهٔ مقدار تکین یک ماتریس را بیان می‌کنیم. فرض کنید $A = U\Sigma V^T$ تجزیهٔ مقدار تکین ماتریس $A \in \mathbb{R}^{m\times n}$ باشد، که در آن

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \quad D = \mathrm{diag}_r(\sigma_1, \sigma_2, \ldots, \sigma_r), \quad \sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0.$$

همچنین برای $k \leqslant r$ تعریف می کنیم $A_k = U\Sigma_k V^T$، که در آن

$$\Sigma_k = \begin{pmatrix} D_k & 0 \\ 0 & 0 \end{pmatrix}, \quad D_k = \mathrm{diag}_k(\sigma_1, \sigma_2, \ldots, \sigma_k).$$

به روشنی داریم $\mathrm{rank}(A_k) = k$. ماتریس A_k را می‌توان به صورت یک تقریب از ماتریس A در نظر گرفت. قضیهٔ زیر کیفیت این تقریب را نشان می‌دهد.

قضیه ۴.٦ از بین تمام ماتریس‌های $m \times n$ ای که دارای رتبهٔ k هستند، ماتریس A_k بهترین تقریب برای ماتریس A است، به این مفهوم که به ازای هر $B \in \mathbb{C}^{m\times n}$ با $\mathrm{rank}(B) = k$، داریم $\|A - A_k\|_2 \leqslant \|A - B\|_2$. به‌علاوه $\|A - A_k\|_2 = \sigma_{k+1}$.

برهان: به [۲] رجوع کنید. □

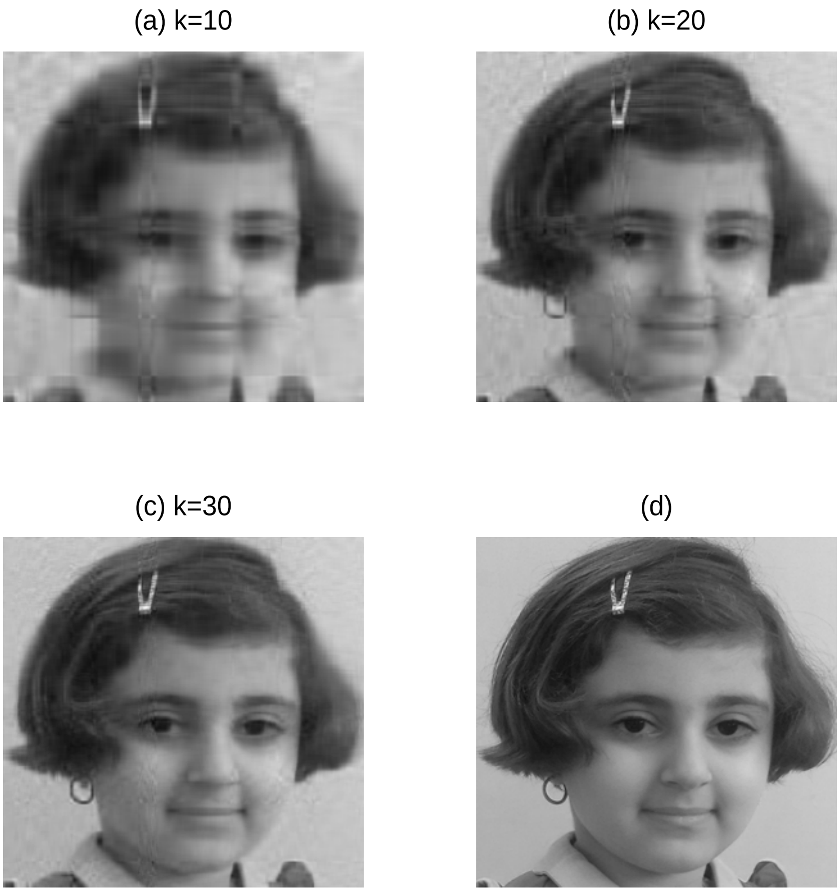
حال چگونگی استفاده از این قضیه برای فشرده‌سازی تصاویر را شرح می‌دهیم. هر تصویر سیاه و سفید را می‌توان با یک ماتریس مستطیلی مثل $A = (a_{ij})$ مشخص کرد که درآن ابعاد ماتریس برابر است با تعداد سلول‌های (پیکسل‌های) تصویر در طول و عرض تصویر، و a_{ij} ها در بازهٔ $[۰, ۱]$ تغییر می‌کنند، ۰ متناظر به سلول تصویری سفید و ۱ متناظر به سلول تصویری سیاه است. مقادیر $۰ < a_{ij} < ۱$، متناظر به سطوح مختلف خاکستری است. فرض کنید که ابعاد یک تصویر خیلی بزرگ باشند و امکان ذخیره‌سازی آن روی یک دیسک یا امکان ارسال آن توسط پُست الکترونیکی نباشد. در این صورت با استفاده از تجزیهٔ مقدار تکین A می‌توان تصویر را فشرده کرده و از حجم تصویر کاست بدون اینکه از کیفیت تصور به طور قابل توجه کاسته شود. برای این کار، تجزیهٔ مقدار تکین A را به دست آورده و به ازای یک مقدار کوچکِ k، ماتریس A_k را به دست آورده و تصویر متناظر را نمایش می‌دهیم. در واقع با این کار، یک تقریب از تصویر اصلی نمایش داده می‌شود. اگر $A \in \mathbb{R}^{m\times n}$، آنگاه برای ذخیره‌سازی تجزیهٔ مقدار تکین ماتریس A_k لازم است برای $i = ۱, ۲, \ldots, k$، بردارهای $\sigma_i u_i \in \mathbb{R}^m$ و $v_i \in \mathbb{R}^n$ ذخیره شوند. در این صورت بایستی $k(m+n)$ اسکالر ذخیره شوند.

در شکل ۴.۶، تصویر (d) مربوط به چهرهٔ یک دختر خانم است که شامل ۱۱۱۹×۱۲۰۶ سلول تصویری و در نتیجه متناظر به یک ماتریس ۱۱۱۹×۱۲۰۶ است. رتبهٔ این ماتریس ۱۱۱۹ است و بنابراین دارای ۱۱۱۹ مقدار تکین ناصفر است. در شکل ۴.۶، به ازای مقادیر مختلف k، یعنی $k = ۱۰, ۲۰, ۳۰$، تصاویر تقریبی مربوط به A_k نیز نمایش داده شده‌اند. به ازای k های مختلف نسبت اسکالرهای ذخیره شدهٔ تصاویر تقریبی به درایه‌های ماتریس A، یعنی $r_k = k(m+n)/(mn)$ به صورت زیرند:

$$r_{۱۰} = \frac{۱۰ \times (۱۲۰۶ + ۱۱۱۹)}{۱۲۰۶ \times ۱۱۱۹} = ۰.۰۲,$$
$$r_{۲۰} = \frac{۲۰ \times (۱۲۰۶ + ۱۱۱۹)}{۱۲۰۶ \times ۱۱۱۹} = ۰.۰۳,$$
$$r_{۴۰} = \frac{۳۰ \times (۱۲۰۶ + ۱۱۱۹)}{۱۲۰۶ \times ۱۱۱۹} = ۰.۰۵.$$

همان‌طور که ملاحظه می‌شود تصویر متناظر به $A_{۳۰}$ بسیار شبیه به تصویر اصلی است. این در حالی است که برای ذخیره‌سازی آن تنها به اندازهٔ ۵% حافظهٔ استفاده شده برای

ذخیره‌سازی تصویر اصلی، از حافظه استفاده شده است.



شکل ۲.٦: فشرده‌سازی تصویر به کمک SVD.

تمرینات

۱.٦) با استفاده از تجزیهٔ QR (به کمک الگوریتم گرام- اشمیت) مسألهٔ کمترین توان‌های دوم (۴.٦) را حل کنید که در آن

$$A = \begin{pmatrix} ۱ & -۳ \\ ۲ & ۴ \\ ۲ & ۵ \end{pmatrix}, \quad b = \begin{pmatrix} ۴ \\ ۳ \\ -۵ \end{pmatrix}.$$

۲.٦) تجزیهٔ مقدار تکین ماتریس‌های زیر را محاسبه کنید:

$$A = \begin{pmatrix} -۱ & ۱ \\ ۰ & -۱ \end{pmatrix}, \quad B = \begin{pmatrix} ۱ & ۰ \\ ۱ & ۰ \end{pmatrix}.$$

۳.٦) کمترین توان‌های دوم (۴.٦) را در نظر بگیرید که در آن

$$A = \begin{pmatrix} ۱ & ۱ & ۰ \\ ۱ & ۱ & ۰ \\ ۱ & ۰ & ۱ \end{pmatrix}, \quad b = \begin{pmatrix} ۴ \\ ۲ \\ ۱ \end{pmatrix}.$$

این مسأله را با استفاده از قضیهٔ ۱.٦ حل کنید. برای این کار، دستگاه معادلات نرمال متناظر، یعنی $A^TAx = A^Tb$، را تشکیل داده و با حل آن نشان دهید که جواب‌های مسألهٔ کمترین توان‌های دوم به صورت

$$x(t) = \begin{pmatrix} ۱ \\ ۲ \\ ۰ \end{pmatrix} + t \begin{pmatrix} -۱ \\ ۱ \\ ۱ \end{pmatrix}, \quad t \in \mathbb{R},$$

است. نشان دهید که بردار $(۱, ۲, ۰)^T$ یک جواب دستگاه $A^TAx = A^Tb$ است و داریم $(-۱, ۱, ۱)^T \in \mathcal{N}(A)$. با محاسبهٔ $x^* = A^+b$ نشان دهید

$$x(-\frac{۱}{۳}) = x^* = \frac{۱}{۳} \begin{pmatrix} ۴ \\ ۵ \\ -۱ \end{pmatrix}.$$

درستی قضیهٔ ۳.٦ را در اینجا تحقیق کنید.

۴.٦) تجزیهٔ مقدار تکین ماتریس

$$A = \begin{pmatrix} ۱ & ۱ \\ ۲ & ۲ \\ ۲ & ۲ \end{pmatrix},$$

را محاسبه کرده و با استفاده از آن نشان دهید:

$$A = ۳\sqrt{۲} \begin{pmatrix} \frac{۱}{۳} \\ \frac{۲}{۳} \\ \frac{۲}{۳} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \end{pmatrix}.$$

با استفاده از این رابطه A^+ را محاسبه کنید.

۵.۶) تجزیهٔ مقدار تکین ماتریس‌های زیر را به دست آورید:

$$A = \begin{pmatrix} ۱ \\ -۲ \\ ۲ \end{pmatrix}, \quad B = (\ ۱ \quad -۲ \quad ۲\ ).$$

۶.۶) فرض کنید $x, y \in \mathbb{R}^{n\times n}$ دو بردار ناصفر باشند. نشان دهید:

$$(xy^T)^+ = \frac{yx^T}{x^Tx\ y^Ty}.$$

با فرض

$$x = \begin{pmatrix} ۱ \\ ۳ \\ ۱ \end{pmatrix}, \quad y = \begin{pmatrix} ۲ \\ ۲ \\ ۲ \end{pmatrix},$$

و $A = xy^T$ ماتریس A^+ را محاسبه کنید.

۷.۶) فرض کنید $A \in \mathbb{R}^{n\times n}$ و $q \in \mathbb{R}^n \neq ۰$. نشان دهید عدد حقیقی یکتای

$$\rho = \frac{q^TAq}{q^Tq},$$

عبارت $\|Aq - \rho q\|_۲$ را مینیمم می‌کند.

۸.۶) فرض کنید $A \in \mathbb{R}^{m\times n}$. نشان دهید:

(الف) گر $m \geqslant n$ و $\mathrm{rank}(A) = n$، آنگاه $A^+ = (A^TA)^{-۱}A^T$.

(ب) اگر $m \leqslant n$ و $\mathrm{rank}(A) = m$، آنگاه $A^+ = A^T(AA^T)^{-۱}$.

۹.۶) فرض کنید σ_i، $i = ۱, \ldots, r$ مقادیر تکین ناصفر ماتریس A باشند. نشان دهید:

$$\|A\|_F = \sqrt{\sum_{i=۱}^{r} \sigma_i^۲}.$$

۱۰.۶) با فرض

$$A = \begin{pmatrix} ۳ & ۰ & ۳ \\ ۴ & -۱ & ۱ \end{pmatrix}, \quad b = \begin{pmatrix} ۶ \\ ۴ \end{pmatrix}.$$

مسألهٔ زیر را حل کنید:

$$\min \|x\|_۲$$
$$s.t. : Ax = b.$$

۱۱.۶) فرض کنید $A = PS$ که در آن

$$A = \begin{pmatrix} ۱ & -۱ \\ ۲ & ۱ \\ ۰ & ۱ \\ ۱ & -۱ \end{pmatrix}, \quad S = \begin{pmatrix} ۲ & -۱ \\ ۰ & ۱ \end{pmatrix}.$$

برداری مثل $x^* \in \mathbb{R}^۲$ بیابید به طوری که داشته باشیم $\|Ax^* - b\|_۲ = \min_{x\in\mathbb{R}^۲} \|Ax - b\|_۲$ که در آن $b = (۱, ۲, ۳, ۴)^T$. توجه کنید که ستون‌های ماتریس P برهم عمودند و نیازی به محاسبهٔ A نیست.

۱۲.۶) یک خط راست به داده‌های $\{(۲, ۷), (۳.۵, ۱۲), (۷, ۱۳)\}$ برازش کنید.

۱۳.۶) یک چندجمله‌ای درجه دوم به داده‌های $\{(۳, ۸), (۵, ۲۹), (۸, ۵۹), (۱۰, ۹۸)\}$ برازش کنید.

۱۴.۶) یک چندجمله‌ای درجهٔ دوم به صورت $p(x) = \alpha + \beta x^۲$ به مجموعهٔ داده‌های $\{(۲, ۰), (۱, ۱), (-۱, ۲)\}$ برازش کنید.

۱۵.۶) فرض کنید

$$A = \begin{pmatrix} ۱ & ۲ \\ ۳ & ۴ \\ ۵ & ۶ \end{pmatrix}.$$

یک ماتریس مثل B با رتبهٔ یک بیابید که تحت نرم-۲ نزدیکترین ماتریس به A باشد. مقدار $\|A - B\|_۲$ را محاسبه کنید.

۱۶.۶) نشان دهید اگر ستون‌های ماتریس A یک مجموعهٔ یکامتعامد تشکیل دهند، آنگاه $A^+ = A^T$.

۱۷.۶) فرض کنید:

$$A = \begin{pmatrix} ۱ & ۱ & ۱ \\ ۱ & ۱ & ۱ \\ ۱ & ۱ & ۱ \end{pmatrix}, \quad b = \begin{pmatrix} ۴ \\ ۲ \\ ۱ \end{pmatrix}.$$

فضای پوچ A و ماتریس A^+ را محاسبه کرده و به کمک قضیهٔ ۳.۶ جواب عمومی مسألهٔ کمترین توانهای دوم متناظر را بنویسید.

کتاب‌نامه

[۱] د. خجسته‌سالکویه، روش‌های عددی در جبرخطی، انتشارات آموزش‌های بنیادی، تهران، ۱۳۹۰.

[2] G. Allaire and S. M. Kaber, *Numerical linear algebra*, Springer, 2008.

[3] B. N. Datta, *Numerical linear algebra and applications*, Brooks/Cole Publishing Company, 1995.

[4] W. Gander, M. J. Gander and F. Kwok, *Scientific computing, an introduction using Maple and MATLAB*, Springer, 2015.

[5] G. H. Golub and C. F. Van Loan, *Matrix computation*, Third edition, John Hopkins Press, Baltimore, 1996.

[6] G. H. Golub and J. M. Ortega, *Scientific computing and differential equation*, Academic Press, New York, 1992.

[7] A. J. Laub, *Matrix analysis for scientists & engineers*, SIAM, Philadelphia, 2005.

[8] C. D. Meyer, *Matrix analysis and applied linear algebra*, SIAM, Philadelphia, 2000.

[9] A. Quarteroni, R. Sacco and F. Saleri, *Numerical mathematics*, Springer-Verlag, New York, 2000.

[10] Y. Saad, *Iterative methods for sparse linear systems*, Second Edition, SIAM, Philadelphia, 2003.

[11] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, Second Edition, Springer-Verlag, New York, 1980.

[12] J. H. Wilkinson, *The algebraic eigenvalue problem*, Oxford: Clarendon Press, London and New York, 1965.

واژه‌نامهٔ فارسی به انگلیسی

آ

آبلی	Abelian
اثر	trace
آرایه	array
اختلال	purturbation
استقراء	induction
اسکالر	scaler
اشتراک	intersection
اعمال حسابی	arithmetic operations
اعمال سطری مقدماتی	row elementary operations
الگوریتم	algorithm
انباشتگی خطا	error propagation
ایستا	stationary

ب

بالامثلثی	upper triangular
بالامثلثی بلوکی	block upper triangular
بالاهسنبرگی	upper Hessenberg
بدحالت	ill-conditioned

برازش داده‌ها	data fitting
برد	range
بردار	vector
بردار تکین چپ	left singular vector
بردار تکین راست	right singular vector
بردار ویژه	eigenvector
بسط فوریه	Fourier expansion
بعد	dimension
بعد فضای پوچ، پوچی	nullity

پ

پارامتر انتقال	shift parameter
پارامتر بهینه	optimal parameter
پایه	base
پایدار	stable
پایداری	stability
پایین‌مثلثی	lower triangular
پایین‌هسنبرگی	lower Hessenberg
پر	full
پوچ توان	nilpotent
پیکسل	pixel

ت

ترانهاده	transpose
تجزیه	decomposition
تجزیهٔ LU	LU factorization
تجزیهٔ مقدار تکین	singular value decomposition
تصویر	image

تعامد	orthogonality
تعمیم‌یافته	generalized
تفاضلات متناهی	finite differences
تکرار	iteration
تکراری	iterative
تکین	singular
تُنُک	sparse
توپلیتز	Toeplitz
تیلور	Taylor

ج

جایگشت	permutation
جایگذاری پسرو	backward substitution
جایگذاری پیشرو	forward substitution
جبر	algebra
جزئی	partial
جمع مستقیم	direct sum
جواب تقریبی	approximate solution
جواب عمومی	general solution
جواب واقعی	exact solution

چ

چگال	dense
چندجمله‌ای	polynomial
چولسکی	Cholesky

ح

حافظه	memory

حالت	condition
حساب دقیق	exact arithmetic
حدس اولیه	initial guess
حذفی گوس	Gaussian elimination
حقیقی	real

خ

خطا	error
خطی	linear
خودتوان	idempotent

د

دایره‌های گرشگورین	Gerschgorin circles
دترمینان	determinant
درایه	entry
درجه دوم	quadratic
دستگاه	system
دستگاه معادلات خطی	system of linear equations
دولیتل	Doolittle

ذ

ذخیره‌سازی	storage

ر

رتبه	rank
رتبهٔ کامل	full rank
رد	trace
رقم بامعنی	significant digit
روش توانی	power method

روش توانی معکوس	inverse power method
روش تکراری	iterative method
روش مستقیم	direct method

ز

زوج ویژه	eigenpair
زیرفضا	subspace
زیرماتریس اصلی	principle submatrix
زیرماتریس اصلی پیشرو	leading principle submatrix

ژ

ژاکوبی	Jacobi

س

سازگار	compatible
سری	series
سلول	pixel
سه‌قطری	tridiagonal

ش

شرایط مرزی	boundary conditions
شدت جریان	circuitry
شرمن	Sherman
شعاع طیفی	spectral radius
شکافت	splitting
شور	Schur

ض

ضرب داخلی	inner product
ضربی	multiplicative

ط

طیف	spectrum
طبیعی	natural

ع

عدد حالت	condition number
علوم	science
عمل حسابی	arithmetic operation
عنصر متناهی	finite element
عنصر محوری	pivot entry

غ

غالب قطری اکید	strictly diagonally dominant

ف

فروبنیوس	Frobenius
فشرده‌سازی	compression
فضای برداری	vector space
فضای پوچ	nullspace
فضای ضرب داخلی	inner product space
فوریه	Fourier

ق

قالب	format
قالب مختصاتی	coordinate format
قطری	diagonal
قطری شدنی	diagonalizable

ک

کراندار	bounded

کرونکر	Kronecker
کمترین توان‌های دوم	least squares
کمینه	minimum
کوشی- شوارتز	Cauchy-Schwarz
کهاد	minor

گ

گرادیان	gradient
گروه	group
گسسته	discrete
گرام- اشمیت	Gram-Schmidt
گرام- اشمیت اصلاح شده	modified Gram-Schmidt
گرشگورین	Gershgorin
گاوس- سایدل	Gauss-Seidel

م

ماتریس	matrix
ماتریس افزوده	augmented matrix
ماتریس بلوکی	block matrix
ماتریس تکرار	iteration matrix
ماتریس ضرایب	coefficient matrix
ماتریس گیونز	Givens matrix
ماتریس مقدماتی	elementary matrix
ماتریس هیلبرت	Hilbert matrix
ماکسیمم	maximum
ماینور	minor
متشابه	similar
متعامد	orthogonal

متقارن	symmetric
متقارن کج	skew symmetric
متناهی	finite
متوازی الاضلاع	parallelogram
مثبت	positive
مثلث	triangle
مجموعه گسترنده	spanning set
محورگیری	pivoting
محورگیری جزئی	partial pivoting
محورگیری کلی	complete pivoting
مختصات	coordinate
مختلط	complex
مدل	model
مزدوج	conjugate
مستقل خطی	linearly independent
معادلات	equations
معادلات دیفرانسیل جزئی	partial differential equations
معادلات دیفرانسیل معمولی	ordinary differential equations
معادلات نرمال	normal equations
معکوس	inverse
معکوس‌پذیر	invertible
معکوس تعمیم یافته	generalized inverse
معمولی	ordinary
معین	definite
معین مثبت متقارن	symmetric positive definite
معین مثبت هرمیتی	Hermitian positive definite
مقاومت	resistence

مقدار ویژه	eigenvalue
مقدار ویژهٔ غالب	dominant eigenvalue
مکمل	complement
مکمل شور	Schur complement
منفرد	singular
مور- پن‌رز	Moore-Penrose
موهومی	imaginary
مهندسی	engineering
میدان	field
مینیمم کننده	minimizer

ن

ناتهی	nonempty
ناصفر	nonzero
نامساوی	inequality
نامنفرد	nonsingular
نامنفی	nonnegative
نرم	norm
نرم اقلیدسی	Euclidean norm
نرمال	normal
نرم القایی	induced norm
نرم بینهایت	infinity norm
نرم طبیعی	natural norm
نرم ماتریسی	matrix norm
نرم ماکزیمم	maximum norm
نرم-۱	1-norm
نگاشت	map

نیمه معین مثبت متقارن	symmetric positive semidefinite

و

وابسته خطی	linearly dependent
واحد موهومی	imaginary unit
واگرا	divergent
وجود	existence
ولتاژ	voltage

ه

هرمیتی	Hermitian
همانی	identity
همگرا	convergent
همگرایی	convergence
همگن	homogeneous
هندسی	geometrical
هوسهولدر	Householder
هیلبرت	Hilbert

ی

یکانی	unitary
یکامتعامد	orthonormal
یک به یک	one-to-one
یکتا	unique
یکتایی	uniqueness
یکه	unit

فهرست الفبایی

Numerical Linear Algebra

Davod Khojasteh Salkuyeh (PhD)
University of Guilan

Numerical Linear Algebra

By:

Dr. Davod Khojasteh Salkuyeh

ISBN:978-964-6183-03-2