بسم الله الرحمن الرحیم

دانشگاه اصفهان

دانشکده ریاضی و آمار

گروه ریاضی

پروژه ی کارشناسی رشته ی ریاضیات و کاربردها

جبر خطی برای علوم داده

استاد راهنما :

دکتر ملیحه یوسف زاده

پژوهشگر :

پرهام پیشرو

شهریور ماه ۱۴۰۱

تشکر و قدردانی

با سپاس و تقدیر ویژه از استاد فرهیخته سرکار خانم دکتر ملیحه یوسف‌زاده که در مراحل مختلف این پژوهش، راهنمایی‌های ارزنده و سازنده خود را بر من ارزانی داشتند و همچنین با انتخاب موضوعی فوق‌العاده، انگیزه بنده را در تهیه و تدوین این پروژه، دو چندان کردند.

چکیده

در این پروژه کارشناسی، قسمت‌هایی از جبر خطی که در حوزه علوم داده کاربرد دارند بررسی و توضیح داده شده‌اند. هدف تهیه منبعی مناسب در سطح دانشجویان کارشناسی است که به دنبال کاربردهای جبر خطی در علوم داده هستند.

در ابتدا، مباحث اولیه و مورد نیاز که پیش‌نیاز کاربردهای جبر خطی در علوم داده هستند، ذکر شده است. پس از آن روش‌های مختلف به دست آوردن ماتریس وارون و همچنین دستگاه معادلات خطی که کاربرد نسبتاً کمتری در علوم داده دارند، آورده شده است. پس از آن با بررسی مسئله کمترین مربعات و روش تجزیه مقدار تکین، کاربرد بسیار این روش در فشرده سازی تصاویر دیجیتالی و کاهش ابعاد آورده شده است.

در پایان، پس از بررسی مقادیر ویژه و بردارهای ویژه و همچنین روش پیدا کردن آن‌ها (با کمک روش‌های تکراری و روش‌های تبدیلی)، روش تحلیل مولفه اساسی ذکر شده و کاربرد آن در علوم داده که همان کاهش ابعاد داده می‌باشد، بررسی شده است.

کلمات کلیدی: مسئله کمترین مربعات، تجزیه مقدار تکین، مقدار ویژه و بردار ویژه، تحلیل مولفه اساسی، علوم داده

فهرست مطالب

عنوان **صفحه**

عنوان **صفحه**

عنوان **صفحه**

عنوان **صفحه**

فهرست شکل‌ها

عنوان **صفحه**

عنوان **صفحه**

فهرست جدول‌ها

مقدمه

در گذشته، داده‌ها اغلب ساختار یافته بودند و در حجم بسیار کم وجود داشتند. این ویژگی‌ها امکان تحلیل آن‌ها را با استفاده از ابزارهای ساده هوش تجاری[1] فراهم می‌کرد؛ اما طی سال‌های اخیر با گسترش روز افزون داده‌ها، رشد تکنولوژی های دیجیتال، همه‌گیر شدن اینترنت و توسعه دسترسی به اینترنت اشیا[2] داده‌های فراوانی ذخیره شده‌اند. داده‌هایی که در صورت مدیریت صحیح کاربردهای زیادی خواهند داشت. حوزه های مختلف علمی، تجاری، کسب و کار و هر اقدامی که امروزه در جهان در حال اجراست، می‌تواند داده محور[3] انجام شود. این امر ضرورت استفاده از ابزارهای علمی برای تحلیل و استخراج دانش از مجموعه داده‌ها را بسیار بیشتر می‌کند. البته شایان ذکر است که این وضعیت حال حاضر، چالش‌هایی را بر سر راه تحلیلگران و دانشمندان داده قرار داده است که به طور مثال می‌توان به دو مورد ابعاد و حجم بالای داده ها اشاره کرد:

- ابعاد بالای داده: هنگامی که ابعاد داده از ۳ بعد بیشتر می‌شود، نمایش آن بر روی صفحه امکان پذیر نیست. پس باید از تکنیک‌های کاهش بعد در جبر خطی و هندسه استفاده کرد.
- حجم بالای داده: هنگامی که حجم داده ها به گونه ای زیاد شود که کامپیوتر نتواند اعمال لازم را بر روی داده‌ها انجام دهد، از تکنیک‌های نمونه برداری[4]، انتقال جریانی از داده‌ها[5] و باز طراحی داده‌ها[6] در احتمال و فرآیندهای تصادفی (شامل توزیع داده، زنجیر مارکف و قانون اعداد بزرگ)، ترکیبیات (شامل مدل های گراف های تصادفی) و علوم کامپیوتر (شامل یادگیری ماشین و خوشه بندی) استفاده خواهد شد.

در واقع به علم داده باید به چشم یادگیری دامنه جدیدی از ریاضیات نگریست که با فراگیری چشم انداز نویی از علوم کامپیوتر همراه است. بر اساس تحقیقات دانشگاه هاروارد جذاب‌ترین شغل قرن ۲۱، دانشمند داده[7] معرفی شد.

همان‌گونه که در بالا ذکر شد، جبر خطی از اهمیت بالایی در مباحث علوم داده برخوردار است. برخی از متخصصان از جبر خطی به عنوان زبان ریاضیات در تحلیل داده یاد می‌کنند؛ چرا که پایه هر مجموعه داده[8] یک ماتریس است و هر نوع محاسباتی بر روی این مجموعه داده صورت خواهد گرفت. در مجموع باید گفت یادگیری جبر خطی برای درک مبانی این حوزه الزامی است.

---

[1] Business Intelligence (BI)
[2] Internet of Things (IoT)
[3] Data-Driven
[4] Sampling
[5] Streaming
[6] Sketching
[7] Data Scientist
[8] Data Set

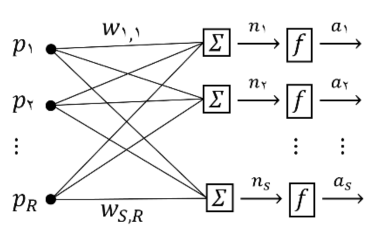
قطعا با یادگیری این مباحث، دانشجو چیزی پیرامون چارچوب‌های مدرن[۱]، جریان کاری یک پروژه علم داده و یا دیگر مواردی که برای موفقیت در یک پروژه علم داده نیاز است نمی‌آموزد؛ اما یک نقطه شروع فوق‌العاده جهت طی کردن مسیر تبدیل شدن به تحلیلگر داده[۲] خواهد داشت.

در ادامه چند مثال برای پی بردن به اهمیت یادگیری جبر خطی و کاربرد آن در علوم داده اشاره شده است:

۱. تصویر یک شی را مجسم کنید. انسان در صورت مشاهده تصویر یک شی به راحتی می‌تواند تشخیص دهد که تصویر موجود نشان‌دهنده کدام شی می‌باشد؛ اما آیا نوشتن کد و ساختن برنامه‌ای که بتواند شی درون تصویر را تشخیص دهد، کار آسانی است؟ جواب آن قطعاً منفی است، چرا که حتی بیان این مسئله به صورت کد نیز کار بسیار مشکلی برای کامپیوتر است. علت سادگی این امر برای انسان این است که مغز انسان در طی میلیون‌ها سال تکامل یافته و اکنون قادر به تشخیص چنین چیزی است. آموزش پردازش تصویر[۳] برای کامپیوتر کار دشواری است و این مسئله یکی از حوزه های تحقیقاتی فعال در علوم داده و یادگیری ماشین[۴] است. کامپیوترهای امروزی برای پردازش دو عدد۰ و ۱ طراحی شده‌اند. پس برای ذخیره‌سازی تصویر با ویژگی های مختلف چه باید کرد؟ باید شدت پیکسل‌ها را در ساختاری به نام ماتریس ذخیره نمود. به این ترتیب، انجام هر گونه عملی بر روی تصویر، به قواعدی از جبر خطی و ماتریس‌ها نیازمند است.

۲. الگوریتم XGBOOST یکی از الگوریتم‌هایی است که توسط بسیاری از افراد موفق در علوم داده، به کار گرفته می‌شود. این الگوریتم، برای ارائه پیش‌بینی‌ها، داده‌های عددی را در فرم ماتریسی ذخیره می‌کند. این کار، باعث تقویت سرعت پردازش داده و نتایج دقیق‌تر آن می‌گردد. به علاوه، بسیاری از الگوریتم‌های مختلف دیگر نیز از ماتریس‌ها برای ذخیره‌سازی و پردازش داده‌ها استفاده می‌کنند.

۳. یادگیری عمیق[۵]، یک عبارت فراگیر جدید در حوزه به‌کارگیری ماتریس‌ها به منظور ذخیره ورودی‌هایی مانند تصویر، گفتار، متن و ارائه یک راه‌حل برای مسائل این‌چنینی است. حتی وزن‌های به‌دست‌آمده توسط یک شبکه عصبی مصنوعی[۶] نیز در ماتریس‌ها ذخیره می‌شوند. نمونه زیر، مثالی از نمایش گرافیکی وزن‌های ذخیره شده در یک ماتریس است.



$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix}$$

---

[۱] Modern Frameworks
[۲] Data Analyst
[۳] Image Processing
[۴] Machine Learning (ML)
[۵] Deep Learning (DL)
[۶] Artificial Neural Networks (ANN)

۴. یکی دیگر از زمینه‌های تحقیقاتی فعال در حوزه یادگیری ماشین، نحوه کار با متن است. رایج‌ترین تکنیک-های مورد استفاده در این زمینه، کیسه‌ی کلمات[1] و ماتریس لغت-سند[2] هستند. تمام این تکنیک‌ها، به صورت بسیار مشابهی عمل می‌کنند. به این صورت که شمارش کلمات در اسناد را انجام داده و به منظور اجرای عملیاتی مانند تحلیل معنایی، ترجمه زبان، خلق زبان و... تعداد تکرار کلمات را در یک فرم ماتریسی ذخیره می‌کنند.

نمونه‌های بالا تنها بخشی از موارد جبر خطی در علوم داده هستند. پس به درستی می‌توان گفت که جبر خطی اهمیت و کاربرد بسیاری در مباحث مختلف علوم داده دارد.

برای آموزش مباحث جبر خطی مورد نیاز در علم داده، دسته‌بندی‌های مختلفی از سوی دانشگاه‌ها، موسسات آموزش عالی و سایت‌های آموزشی ارائه شده است که پس از بررسی و مقایسه بین دسته‌بندی‌های موجود، به یک فهرست جامع و کامل از مباحث رسیده و سپس هر کدام از مباحث به تفکیک توضیح داده خواهد شد.

[1] Bag of Words

[2] Term Document Matrix

فصل اول

مباحث اولیه جبر خطی

فصل اول: مباحث اولیه جبر خطی

قبل از شروع تذکر داده می‌شود که مطالب این فصل را می‌توان با جزئیات بیشتر در [۱] و [۴] پیدا نمود. در این فصل مقدمات و پیش‌نیازهای لازم برای تهیه و تنظیم پروژه قرار داده شده است. به طور مثال می‌توان به ماتریس‌ها اشاره کرد. ماتریس‌ها یکی از پرکاربردترین ابزارهای ریاضی در شاخه‌های مختلف علوم می‌باشد. به طور خاص می‌توان حوزه علوم داده را نام برد که در آن هر مجموعه داده به‌صورت یک ماتریس نمایش داده می‌شود. بدین جهت، ابتدای پروژه با مفهوم ماتریس شروع خواهد شد.

۱-۱ ماتریس

تعریف ۱-۱) فرض کنید $\mathbb{F}$ یک میدان باشد، آرایه مستطیل شکل زیر که عناصر آن متعلق به $\mathbb{F}$ هستند را یک ماتریس می‌نامند.

$$A = \begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m۱} & a_{m۲} & \cdots & a_{mn} \end{bmatrix} = \left[a_{ij}\right]_{m\times n}$$

به ازای هر $i = ۱,۲,\cdots,m$ و $j = ۱,۲,\cdots,n$، عنصر a_{ij} را درایه ij ام ماتریس A می‌نامند. همچنین اگر m تعداد سطرهای ماتریس و n تعداد ستون‌های ماتریس A باشند، ماتریس را از اندازه $m \times n$ می‌خوانند. هر ماتریس $m \times ۱$ را یک ستون m تایی و هر ماتریس $۱ \times n$ را یک سطر n تایی می‌نامند. همچنین اگر $m = n$ باشد، ماتریس A ، یک ماتریس مربعی نامیده می‌شود.

قضیه ۱-۲) اگر A و B دو ماتریس مربعی دلخواه باشند و به گونه‌ای باشد که $AB = BA$ ، آنگاه به ازای هر مقدار $n \in \mathbb{N}$، $(AB)^n = A^nB^n$.

برهان) برای اثبات باید از استقرا روی n استفاده کرد. پس برای $n = ۱$ می‌توان نوشت:

$$(AB)^۱ = A^۱B^۱$$

حال با فرض $n = k$ (یعنی $(AB)^k = A^kB^k$) باید حکم را برای $n = k + ۱$ اثبات کرد:

$$(AB)^{k+۱} = (AB)^k(AB) = A^kB^kAB = A^k \underbrace{BB\cdots B}_{k\text{ بار}} AB = A^k \underbrace{BB\cdots B}_{k-۱\text{ بار}} ABB \xRightarrow{\text{ادامه همین روند}}$$

$$A^kA \underbrace{B\cdots BB}_{k\text{ بار}} B = A^{k+۱}B^{k+۱}$$

بدین ترتیب اثبات پایان می‌یابد. ■

تعریف ۱-۳) در ماتریس مربعی $A = \left[a_{ij}\right]_{n\times n}$ ، درایه‌های $a_{۱۱}, a_{۲۲}, \cdots, a_{nn}$ درایه‌های قطری یا قطر اصلی ماتریس A نامیده می‌شود. اگر در یک ماتریس تمام درایه‌ها به غیر از درایه‌های قطری آن صفر باشند، آن را ماتریس قطری نامیده و به صورت $A = diag(a_{۱۱}, a_{۲۲}, \cdots, a_{nn})$ نمایش داده می‌شود. فرم کلی یک ماتریس قطری به صورت زیر می‌باشد:

$$\begin{bmatrix} a_{۱۱} & ۰ & \cdots & ۰ \\ ۰ & a_{۲۲} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & a_{nn} \end{bmatrix}$$

قضیه ۱-۴) اگر ماتریس A یک ماتریس قطری $n \times n$ و $A = diag(a_{۱}, a_{۲}, \cdots, a_{n})$ باشد، آنگاه به ازای هر $k \in \mathbb{N}$ می‌توان نوشت $A^k = diag(a_{۱}^k, a_{۲}^k, \cdots, a_{n}^k)$.

برهان) به سادگی و با استقرا روی k به دست می‌آید. حکم برای $k = ۱$ ، به وضوح برقرار است؛ یعنی می‌توان نوشت $A = diag(a_{۱}^{۱}, a_{۲}^{۱}, \cdots, a_{n}^{۱})$. حال فرض کنید حکم برای k برقرار باشد. پس باید حکم را برای $k + ۱$ اثبات کرد.

$$A^{k+۱} = AA^k = diag(a_{۱}, \cdots, a_n) diag(a_{۱}^k, \cdots, a_n^k) \Rightarrow$$

$$A^{k+۱} = \begin{bmatrix} a_{۱} & ۰ & \cdots & ۰ \\ ۰ & a_{۲} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & a_n \end{bmatrix} \begin{bmatrix} a_{۱}^k & ۰ & \cdots & ۰ \\ ۰ & a_{۲}^k & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & a_n^k \end{bmatrix} = \begin{bmatrix} a_{۱}^{k+۱} & ۰ & \cdots & ۰ \\ ۰ & a_{۲}^{k+۱} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & a_n^{k+۱} \end{bmatrix} \Rightarrow$$

$$A^{k+۱} = diag(a_{۱}^{k+۱}, \cdots, a_n^{k+۱})$$

بدین ترتیب، با توجه به استقرا، اثبات تمام است. ■

تعریف ۱-۵) ترانهاده ماتریس $A_{m\times n}$، یک ماتریس $n \times m$ است که هر ستون A یک سطر آن است و آن را با A^T نمایش می‌دهند ($A = \left[a_{ij}\right]_{m\times n} \to A^T = \left[a_{ji}\right]_{n\times m}$).

تعریف ۱-۶) ماتریس مربعی A را متقارن گویند، هرگاه $A^T = A$ و پاد متقارن گویند هرگاه $A^T = -A$ باشد.

اعمال جبری روی ماتریس‌ها و بردارها:

جمع و تفریق دو ماتریس فقط در صورتی امکان‌پذیر است که هر دو ماتریس هم اندازه باشند؛ به صورت زیر:

$$A = \begin{bmatrix} a_{۱۱} & \cdots & a_{۱n} \\ \vdots & \ddots & \vdots \\ a_{m۱} & \cdots & a_{mn} \end{bmatrix}, B = \begin{bmatrix} b_{۱۱} & \cdots & b_{۱n} \\ \vdots & \ddots & \vdots \\ b_{m۱} & \cdots & b_{mn} \end{bmatrix}$$

$$\Rightarrow A \pm B = \begin{bmatrix} a_{۱۱} \pm b_{۱۱} & \cdots & a_{۱n} \pm b_{۱n} \\ \vdots & \ddots & \vdots \\ a_{m۱} \pm b_{m۱} & \cdots & a_{mn} \pm b_{mn} \end{bmatrix}$$

ضرب یک عدد یا همان اسکالر $c \in \mathbb{F}$ در ماتریس A روی میدان $\mathbb{F}$ به صورت زیر می‌باشد:

$$cA = \begin{bmatrix} ca_{۱۱} & ca_{۱۲} & \cdots & ca_{۱n} \\ ca_{۲۱} & ca_{۲۲} & \cdots & ca_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m۱} & ca_{m۲} & \cdots & ca_{mn} \end{bmatrix}$$

ضرب دو ماتریس $A_{m\times n}$ و $B_{p\times q}$ فقط در صورتی امکان‌پذیر است که $n = p$ باشد. یعنی تعداد ستون‌های ماتریس اول با تعداد سطرهای ماتریس دوم برابر باشد. در این صورت حاصل‌ضرب AB به صورت زیر خواهد بود:

$$AB = \left[c_{ij}\right]_{m\times q}, c_{ij} = a_{i۱}b_{۱j} + a_{i۲}b_{۲j} + \cdots + a_{in}b_{nj}, i = ۱, \cdots, m, j = ۱, \cdots, p$$

تعریف ۱-۷) تابع $(.,.): \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ را یک ضرب داخلی[۱] روی $\mathbb{C}^n$ گویند، هرگاه:

۱) $\forall x \in \mathbb{C}^n: (x,x) \geq ۰$؛

۲) $(x,x) = ۰ \Leftrightarrow x = \vec{۰}$؛

۳) $\forall x,y \in \mathbb{C}^n: (x,y) = \overline{(y,x)}$؛

۴) $\forall \alpha,\beta \in \mathbb{C}, \forall x,y,z \in \mathbb{C}^n : (\alpha x + \beta y, z) = \bar{\alpha}(x,z) + \bar{\beta}(y,z)$

تذکر ۱-۸) اگر در تعریف (۱-۷) به جای $\mathbb{C}^n$ از $\mathbb{R}^n$ استفاده شود، آنگاه ویژگی‌های ۲، ۳ و ۴ به صورت زیر خواهند بود:

۲) $(x,x) = ۰ \Leftrightarrow x = ۰$؛

۳) $\forall x,y \in \mathbb{R}^n: (x,y) = (y,x)$؛

۴) $\forall \alpha,\beta \in \mathbb{R}, \forall x,y,z \in \mathbb{R}^n : (\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z)$

ضرب داخلی دو بردار a و b از فرمول‌های زیر محاسبه می‌شود:

$$a \cdot b = a_۱b_۱ + a_۲b_۲ + \cdots + a_nb_n \tag{۱-۱}$$

$$a \cdot b = |a| \times |b| \times \cos(\theta) \tag{۲-۱}$$

که در آن $|a|$ و $|b|$ به ترتیب اندازه بردارهای a و b می‌باشد.

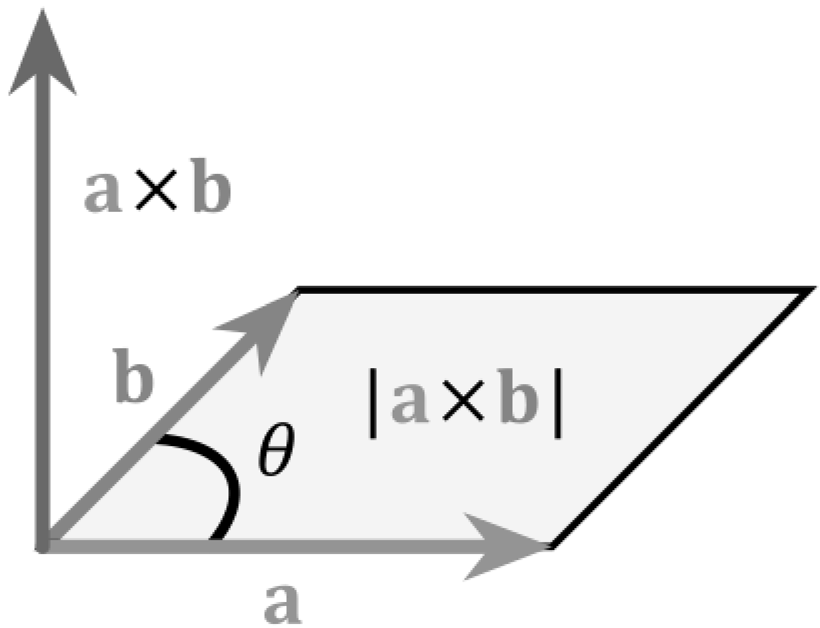
تعریف ۱-۹) ضرب خارجی[۲] یک عمل دوتایی (با نماد $\times$) بین دو بردار اقلیدسی در فضای سه‌بعدی است که نتیجه آن برداری است که بر هر دو بردار اولیه عمود است.

ضرب خارجی دو بردار نیز به صورت زیر محاسبه خواهد شد و تعبیر هندسی آن، شکل (۱-۱) خواهد بود:

$$\vec{a} \times \vec{b} = \begin{vmatrix} i & j & k \\ a_۱ & a_۲ & a_۳ \\ b_۱ & b_۲ & b_۳ \end{vmatrix}$$

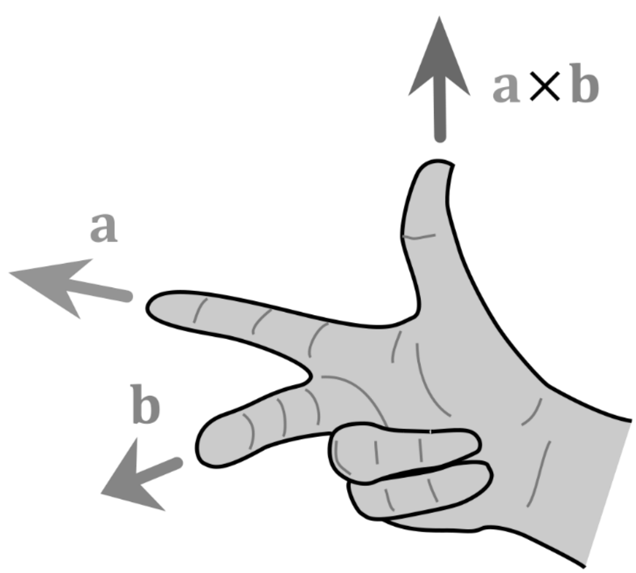
[۱] Inner Product

[۲] Cross Product



شکل (۱-۱) تعبیر هندسی ضرب خارجی

تذکر۱-۱۰) جهت بردار حاصل از ضرب خارجی، عمود بر هر دو بردار $\vec{a}$ و $\vec{b}$ می‌باشد و به کمک قانون دست راست همانند شکل (۱-۲) قابل تشخیص است:



شکل (۱-۲) قانون دست راست

تذکر ۱-۱۱) طول بردار $\vec{a} \times \vec{b}$ برابر مساحت متوازی‌الاضلاعی با اضلاع بردارهای اولیه همانند شکل (۱-۱) می‌باشد؛ پس اگر θ زاویه بین دو بردار باشد، طول بردار $\vec{a} \times \vec{b}$ از رابطه (۱-۳) به دست خواهد آمد:

$$|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}| \sin\theta \qquad (۱-۳)$$

قضیه ۱-۱۲) فرض کنید A ، B و C ماتریس‌هایی $m \times n$ بوده و O یک ماتریس $m \times n$ که تمام درایه‌های آن صفر است و c و d اسکالر باشند؛ در این صورت:

الف) $A + B = B + A$ ب) $A + (B + C) = (A + B) + C$ ج) $O + A = A$

د) $c(A + B) = cA + cB$ ه) $(c + d)A = cA + dA$ و) $(cd)A = c(dA)$

برهان) رجوع شود به [۱].

تعریف ۱-۱۳) دو ماتریس A و B را متشابه گویند، اگر ماتریس وارون‌پذیر P موجود باشد که $B = P^{-۱}AP$.

تعریف ۱-۱۴) ماتریس مربعی I که درایه‌های قطر اصلی آن ۱ و بقیه عناصر صفر می‌باشد را ماتریس همانی می‌نامند. یعنی $I_n = diag(\underbrace{۱,۱,\cdots,۱}_{n \text{ بار}})$ و به صورت زیر نمایش می‌دهند:

$$I = \begin{bmatrix} ۱ & ۰ & \cdots & ۰ \\ ۰ & ۱ & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & ۱ \end{bmatrix}$$

قضیه ۱-۱۵) فرض کنید A ، B و C ماتریس‌هایی باشند که اعمال زیر روی آن‌ها قابل تعریف باشد، در این صورت:

الف) $A(BC) = (AB)C$ ب) $(A + B)C = AC + BC$ ج) $IA = AI = A$

د) $(AB)^T = B^T A^T$ ه) $(A + B)^T = A^T + B^T$

برهان) رجوع شود به [۱].

تعریف ۱-۱۶) اگر تمام درایه‌های زیر قطر اصلی یک ماتریس صفر باشند، یعنی ماتریسی که در آن به ازای هر $i > j$ عنصر $a_{ij} = ۰$ باشد، به آن ماتریس بالا مثلثی گویند. به طور مشابه اگر تمام درایه‌های بالای قطر اصلی یک ماتریس برابر با صفر باشند، یعنی ماتریسی که در آن به ازای هر $i < j$ ، عنصر $a_{ij} = ۰$ باشد، به آن ماتریس پایین مثلثی می‌گویند. به طور مثال فرم کلی ماتریس بالا مثلثی ۳ × ۳ در سمت راست و فرم کلی ماتریس پایین مثلثی ۳ × ۳ در سمت چپ نمایش داده شده است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & a_{۱۳} \\ ۰ & a_{۲۲} & a_{۲۳} \\ ۰ & ۰ & a_{۳۳} \end{bmatrix} \qquad \begin{bmatrix} a_{۱۱} & ۰ & ۰ \\ a_{۲۱} & a_{۲۲} & ۰ \\ a_{۳۱} & a_{۳۲} & a_{۳۳} \end{bmatrix}$$

تذکر ۱-۱۷) حاصل‌ضرب هر تعداد متناهی از ماتریس‌های بالا مثلثی یک ماتریس بالا مثلثی و حاصل‌ضرب هر تعداد متناهی از ماتریس‌های پایین مثلثی یک ماتریس پایین مثلثی است.

تعریف ۱-۱۸) اگر A یک ماتریس $n \times n$ باشد، آنگاه اثر A که با $tr(A)$ یا $trace(A)$ نمایش داده می‌شود؛ برابر است با مجموع تمام درایه‌های قطر اصلی آن، یعنی:

$$tr(A) = a_{۱۱} + a_{۲۲} + \cdots + a_{nn} = \sum_{i=۱}^{n} a_{ii} \qquad (۱-۴)$$

قضیه ۱-۱۹) اگر A و B دو ماتریس مربعی $n \times n$ دلخواه باشند، آنگاه $tr(AB) = tr(BA)$.

برهان) رجوع شود به [۱].

قضیه ۱-۲۰) اگر A و B دو ماتریس مربعی و $a, b \in \mathbb{R}$ دو اسکالر دلخواه باشند، آنگاه

$$tr(aA + bB) = a \times tr(A) + b \times tr(B)$$

برهان) فرض کنید $A = [a_{ij}]_{n\times n}$ و $B = [b_{ij}]_{n\times n}$ باشند. در این صورت:

$$aA + bB = a[a_{ij}]_{n\times n} + b[b_{ij}]_{n\times n} = [aa_{ij} + bb_{ij}]_{n\times n} \xRightarrow{\text{طبق تعریف}}$$

$$tr(aA + bB) = \sum_{i=۱}^{n} (aa_{ii} + bb_{ii}) = a\sum_{i=۱}^{n} a_{ii} + b\sum_{i=۱}^{n} b_{ii} = a \times tr(A) + b \times tr(B) \;\blacksquare$$

تعریف ۱-۲۱) ماتریس مربعی A را پوچ توان گویند، هرگاه عدد صحیح k وجود داشته باشد؛ به طوری که $A^k = ۰$ و کوچکترین عدد صحیح و مثبت k که به ازای آن $A^k = ۰$ شود را مرتبه‌ی پوچی ماتریس A گویند.

تعریف ۱-۲۲) مجموعه همه ماتریس‌های $m \times n$ با درایه‌هایی در میدان $\mathbb{F}$ را با $M_{m\times n}(\mathbb{F})$ نمایش می‌دهند.

۱-۱-۱ اعمال سطری مقدماتی

تعریف ۱-۲۳) اعمال زیر را اعمال سطری مقدماتی گویند:

۱) تعویض دو سطر با یک‌دیگر ($R_i \leftrightarrow R_j$)
۲) ضرب اسکالر ناصفر در یک سطر ($cR_i \to R_i$)
۳) اضافه کردن مضربی از یک سطر به سطر دیگر ($cR_i + R_j \to R_j$)

تعریف ۱-۲۴) به ماتریسی که تنها با انجام یک عمل سطری مقدماتی به I_n تبدیل شود، ماتریس مقدماتی گویند.

تعریف ۱-۲۵) دو ماتریس A و B را که با انجام اعمال سطری مقدماتی به یک‌دیگر تبدیل شوند را هم‌ارز سطری گویند و با $A \sim B$ نمایش می‌دهند.

تعریف ۱-۲۶) ماتریس A را سطری-پلکانی گویند، هرگاه:

۱) همه سطرهای صفر (در صورت وجود) در انتها (پایین) ماتریس باشند.
۲) اگر دو سطر متوالی ناصفر باشند، سطر دوم با تعداد صفرهای بیشتری شروع شود.
۳) در هر سطر اولین درایه ناصفر برابر ۱ باشد که به آن درایه پیشرو گویند.

تعریف ۱-۲۷) به ماتریس سطری-پلکانی با شرایط فوق، ماتریس سطری-پلکانی تحویل یافته گویند هرگاه علاوه بر شرایط موجود در تعریف (۱-۲۱)، شرط زیر را داشته باشند:

۴) همه درایه‌های واقع در ستون درایه پیشرو به جز خود درایه پیشرو صفر باشند.

تذکر ۱-۲۸) در تعاریف (۱-۲۳) تا (۱-۲۷) می‌توان به‌جای سطر از ستون استفاده کرد و تعاریف برقرار خواهد بود.

۱-۱-۲ دترمینان[۱]

دترمینان عدد مخصوصی است که به ماتریس‌های مربعی اختصاص داده می‌شود و در یافتن معکوس ماتریس و همچنین حل دستگاه معادلات خطی و در شرایط بسیار دیگری کاربرد دارد. اگر دترمینان ماتریسی مخالف صفر باشد، آنگاه می‌توان متوجه شد که آن ماتریس معکوس‌پذیر است. از این رو از طریق دترمینان می‌توان مقادیر ویژه یک ماتریس را تعیین کرد. برای محاسبه دترمینان یک ماتریس باید از قضیه زیر که به بسط لاپلاس معروف است، استفاده کرد. همچنین نماد دترمینان ماتریس A به صورت $\det(A)$ یا $|A|$ می‌باشد.

قضیه ۱-۲۹) اگر $A = [a_{ij}]_{n\times n}$ یک ماتریس $n \times n$ روی میدان $\mathbb{F}$ باشد، آنگاه:

$$\begin{cases} \det(A) = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} & \text{بسط لاپلاس روی سطر } i \text{ ام} \quad , \quad i = 1,2,\cdots,n \\ \det(A) = a_{1j}c_{1j} + a_{2j}c_{2j} + \cdots + a_{nj}c_{nj} & \text{بسط لاپلاس روی ستون } j \text{ ام} \quad , \quad j = 1,2,\cdots,n \end{cases}$$

به طور مثال دترمینان ماتریس $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ برابر خواهد شد با $\det(A) = ad - bc$. یا دترمینان ماتریس 3×3 که به صورت $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ می‌باشد برابر خواهد شد با:

$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg) \tag{۱-۵}$$

برهان) رجوع شود به [۱].

[۱] Determinant

انجام اعمال سطری (و یا ستونی) مقدماتی بر دترمینان ماتریس‌ها تاثیر خواهد گذاشت. اگر ماتریس A ، ماتریس اولیه باشد و ماتریس B ماتریس حاصل از انجام یک عمل سطری (و یا ستونی) مقدماتی بر ماتریس A باشد، آنگاه:

۱) $R_i \leftrightarrow R_j \Rightarrow \det(B) = -\det(A)$

۲) $cR_i \rightarrow R_i \Rightarrow \det(B) = c\det(A)$

۳) $cR_i + R_j \rightarrow R_j \Rightarrow \det(B) = \det(A)$

تذکر ۱-۳۰) اگر ماتریسی مثلثی یا قطری باشد، آنگاه دترمینان آن برابر با حاصل‌ضرب درایه‌های روی قطر اصلی آن است.

تذکر ۱-۳۱) اگر ماتریسی دارای یک سطر (و یا ستون) صفر باشد و همچنین اگر دو سطر (و یا ستون) یکسان داشته باشد، آنگاه دترمینان آن صفر خواهد بود.

قضیه ۱-۳۲) اگر A و B دو ماتریس دلخواه مربعی باشند، احکام زیر برقرارند (برهان: رجوع شود به [۱]).

الف) $|A^{-۱}| = \frac{۱}{|A|}$ ب) $|A^T| = |A|$ ج) $|AB| = |A| \times |B|$

۱-۱-۳ ماتریس وارون[۱]

تعریف ۱-۳۳) ماتریس $A^{-۱}$ را وارون ماتریس A گویند، هرگاه $AA^{-۱} = A^{-۱}A = I$ باشد.

قضیه ۱-۳۴) ماتریس A وارون‌پذیر (نامنفرد) است اگر و تنها اگر دترمینان آن مساوی صفر باشد.

برهان) رجوع شود به [۱].

تذکر ۱-۳۵) اگر دو ماتریس دلخواه A و B موجود باشند، آنگاه $(AB)^{-۱} = B^{-۱}A^{-۱}$.

چندین روش برای یافتن ماتریس وارون وجود دارد که در فصل دوم به‌طور مفصل راجع به آن‌ها بحث خواهد شد.

۱-۱-۴ ماتریس قطر غالب[۲]

تعریف ۱-۳۶) ماتریس مربعی $A_{n\times n}$ را قطر غالب گویند هرگاه:

$$|a_{ii}| \geq \sum_{\substack{j=۱ \\ j\neq i}}^{n} |a_{ij}| \quad , \quad i = ۱,۲,\cdots,n \tag{۱-۶}$$

چنانچه در نامساوی فوق علامت بزرگتر مساوی ($\geq$) به علامت بزرگتر ($>$) تبدیل شود، ماتریس $A_{n\times n}$ را اکیداً قطر غالب گویند.

---

[۱] Inverse Matrix

[۲] Diagonally Dominant Matrix

مثال ۱-۳۷) ماتریس‌های زیر نمونه‌هایی از ماتریس‌های اکیداً قطر غالب هستند.

$$A=\begin{bmatrix} ۴ & ۱ & ۱ & -۱ \\ -۱ & ۳ & ۰ & -۱ \\ -۱ & ۰ & ۴ & -۱ \\ ۰ & -۱ & ۲ & ۴ \end{bmatrix}\ ,\ B=\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۳ & ۱ \\ ۰ & ۲ & ۴ \end{bmatrix}\ ,\ C=\begin{bmatrix} ۸ & ۲ & -۳ & ۱ & ۰ \\ -۱ & ۴ & -۱ & ۰ & ۱ \\ ۱ & ۰ & ۹ & -۱ & -۵ \\ ۰ & -۱ & ۲ & ۴ & ۰ \\ ۵ & ۰ & -۱ & ۱ & ۱۰ \end{bmatrix}$$

همان‌طور که مشاهده می‌شود درایه‌های قطر اصلی، از مجموع قدرمطلق درایه‌های هر سطر اکیداً بزرگتر است. پس تمامی ماتریس‌های مثال (۱-۳۷) اکیداً قطر غالب هستند.

قضیه ۱-۳۸) هر ماتریس اکیداً قطر غالب وارون‌پذیر است.

برهان) رجوع شود به [۱].

۱-۱-۵ ماتریس معین مثبت[۱]

تعریف ۱-۳۹) ماتریس $Q_{n\times n}$ را معین مثبت گویند هرگاه به ازای هر بردار ناصفر $X_{n\times ۱}$ بتوان گفت:

$$X^TQX \geq ۰ \qquad (۱-۷)$$

همچنین اگر به ازای هر بردار ناصفر $X_{n\times ۱}$، $X^TQX > ۰$ باشد، به ماتریس Q معین مثبت اکید گویند.

تذکر ۱-۴۰) به ازای هر ماتریس دلخواه Q، همواره QQ^T و Q^TQ معین مثبت هستند.

مثال ۱-۴۱) برای بررسی معین مثبت بودن ماتریس $Q=\begin{bmatrix} ۲ & ۲ & ۱ \\ ۲ & ۵ & ۱ \\ ۱ & ۱ & ۲ \end{bmatrix}$ به صورت زیر باید عمل کرد.

بردار $X^T=[x_۱\ x_۲\ x_۳]^T$ مفروض است. حال رابطه (۱-۷) را باید به صورت زیر تشکیل داد:

$$X^TAX=[x_۱\ x_۲\ x_۳]\begin{bmatrix} ۲ & ۲ & ۱ \\ ۲ & ۵ & ۱ \\ ۱ & ۱ & ۲ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix}=۲x_۱^۲+۵x_۲^۲+۲x_۳^۲+۴x_۱x_۲+۲x_۱x_۳+۲x_۲x_۳$$

$$\Rightarrow X^TAX=(x_۱+۲x_۲)^۲+(x_۱+x_۳)^۲+(x_۲+x_۳)^۲>۰$$

ماتریس فوق نه تنها معین مثبت، بلکه معین مثبت اکید می‌باشد.

۱-۱-۶ زیرماتریس‌های اصلی پیشرو[۲]

تعریف ۱-۴۲) فرض کنید $A_{n\times n}$ داده شده باشد، زیرماتریس‌های اصلی پیشروی مرتبه k ($k=۱,۲,\cdots,n$) از ماتریس A به صورت زیر تعریف می‌شوند:

---
[۱] Positive Definite Matrix
[۲] Leading Principal Submatrix

$$A_k = \begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱k} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k۱} & a_{k۲} & \cdots & a_{kk} \end{bmatrix}$$

در واقع زیرماتریس‌های اصلی پیشرو عبارت است از ماتریس‌هایی که از یک سطر و یک ستون اول، دو سطر و دو ستون اول، $\cdots$ و n سطر و n ستون اول ماتریس A تولید می‌شوند.

مثال ۱-۴۳) زیرماتریس‌های اصلی پیشروی ماتریس $A = \begin{bmatrix} ۱ & ۲ & ۵ \\ ۳ & ۲ & ۱ \\ ۱ & ۱ & ۷ \end{bmatrix}$ به صورت زیر هستند.

$$A_۱ = [۱]\,, A_۲ = \begin{bmatrix} ۱ & ۲ \\ ۳ & ۲ \end{bmatrix}\,, A_۳ = \begin{bmatrix} ۱ & ۲ & ۵ \\ ۳ & ۲ & ۱ \\ ۱ & ۱ & ۷ \end{bmatrix} = A$$

در قضیه بعد، ارتباط بین ماتریس معین مثبت اکید و زیرماتریس‌های اصلی پیشرو مطرح شده است.

قضیه ۱-۴۴) فرض کنید $A_{n\times n}$ ماتریسی متقارن و معین مثبت اکید باشد؛ در این صورت احکام زیر واقع هستند:

۱) ماتریس A وارون‌پذیر است و $A^{-۱}$ متقارن و معین مثبت اکید است.
۲) همه زیرماتریس‌های اصلی پیشروی A نیز معین مثبت اکید هستند.
۳) دترمینان همه زیرماتریس‌های اصلی پیشروی A نیز مثبت هستند.
۴) مولفه‌های روی قطر اصلی A مثبت می‌باشند.
۵) مقادیر ویژه ماتریس A حقیقی و مثبت هستند.

برهان) رجوع شود به [۱]

۱-۱-۷ ماتریس متعامد[۱]

تعریف ۱-۴۵) دو بردار بر هم عمودند (یا متعامدند) اگر حاصل‌ضرب داخل آن‌ها صفر شود.

تعریف ۱-۴۶) مجموعه بردارهای ناصفر $\{v_۱, v_۲, \cdots, v_n\}$ متعامد نامیده می‌شوند اگر این بردارها دو به دو بر هم عمود باشند؛ یعنی برای $i \neq j$ بتوان نوشت $(v_i, v_j) = ۰$. همچنین اگر برای $i = ۱, \cdots, n$ بتوان نوشت که $(v_i, v_j) = ۱$ ، در این صورت مجموعه بردارهای ناصفر $\{v_۱, v_۲, \cdots, v_n\}$ متعامد یکه می‌نامند.

قضیه ۱-۴۷) بردارهای متعامد ناصفر، مستقل خطی هستند.

برهان) رجوع شود به [۴].

[۱] Orthogonal Matrix

مثال ۱–۴۸) بردارهای $v_۱, v_۲, v_۳$ متعامد و بردارهای $u_۱, u_۲, u_۳$ متعامد یکه هستند.

$$v_۱ = \begin{bmatrix} ۰ \\ ۴ \\ ۲ \end{bmatrix}, v_۲ = \begin{bmatrix} -۱ \\ -\frac{۱}{۵} \\ \frac{۲}{۵} \end{bmatrix}, v_۳ = \begin{bmatrix} \frac{۱}{۶} \\ -\frac{۱}{۶} \\ \frac{۱}{۳} \end{bmatrix}$$

$$u_۱ = \frac{v_۱}{\|v_۱\|_۲} = \begin{bmatrix} ۰ \\ \frac{۲\sqrt{۵}}{۵} \\ \frac{\sqrt{۵}}{۵} \end{bmatrix}, v_۲ = \frac{v_۲}{\|v_۲\|_۲} = \begin{bmatrix} -\frac{\sqrt{۳۰}}{۶} \\ -\frac{\sqrt{۳۰}}{۳۰} \\ \frac{\sqrt{۳۰}}{۱۵} \end{bmatrix}, v_۳ = \frac{v_۳}{\|v_۳\|_۲} = \begin{bmatrix} \frac{\sqrt{۶}}{۶} \\ -\frac{\sqrt{۶}}{۶} \\ \frac{\sqrt{۶}}{۳} \end{bmatrix}$$

پس به راحتی می‌توان یک مجموعه متعامد از بردارهای ناصفر را به یک مجموعه متعامد یکه تبدیل کرد.

تعریف ۱–۴۹) یک ماتریس مربعی و حقیقی را متعامد گویند هرگاه ستون‌های آن متعامد یکه باشند. به طور مشابه یک ماتریس مربعی و مختلط را ماتریس یکانی[1]گویند هرگاه ستون‌های آن متعامد یکه باشند.

قضیه ۱–۵۰) برای ماتریس $A_{n\times n}$ احکام زیر معادلند:

۱) A ماتریس متعامد است.
۲) $A^T = A^{-۱}$ یا $AA^T = ۱$.
۳) برای هر بردار $x \in \mathbb{R}^n$، می‌توان نوشت $\|Ax\|_۲ = \|x\|_۲$.

برهان) رجوع شود به [۴].

۱–۱–۸ نرم برداری[2]

تعریف ۱–۵۱) نگاشت $\|\cdot\|: \mathbb{R}^n \to \mathbb{R}$ با ویژگی‌های

۱) $\forall x \in \mathbb{R}^n, \|x\| \geq ۰$؛
۲) $\|x\| = ۰ \leftrightarrow x = \vec{۰}$؛
۳) $\forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n, \|\lambda x\| = |\lambda|\|x\|$؛
۴) $\forall \lambda \in \mathbb{R}, \forall y \in \mathbb{R}^n, \|x + y\| = \|x\| + \|y\|$

را یک نرم برداری روی $\mathbb{R}^n$ می‌نامند.

[1] Unitary Matrix
[2] Vector Norm

فرض کنید $x = [x_۱, x_۲, \dots, x_n]^T$ باشد، در این صورت نرم‌های زیر تعریف خواهند شد:

$$\|x\|_۱ = \sum_{i=۱}^{n} |x_i|$$

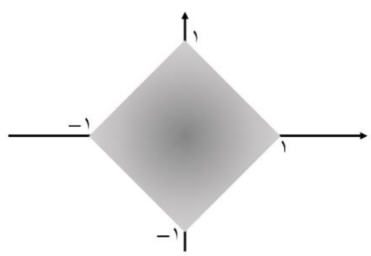
$$\|x\|_p = \left(\sum_{i=۱}^{n} |x_i|^p\right)^{\frac{۱}{p}}, p \geq ۱$$

$$\|x\|_\infty = \max_{۱\leq i\leq n} |x_i|$$

در حالی که $p = ۲$ باشد ، نرم برداری حاصل را نرم اقلیدسی x می‌نامند زیرا که این نرم معرف فاصله x از مبدا مختصات می باشد. همچنین اگر $p = \infty$ باشد، نرم حاصل را نرم ماکزیمم یا نرم چبیشف می‌نامند.

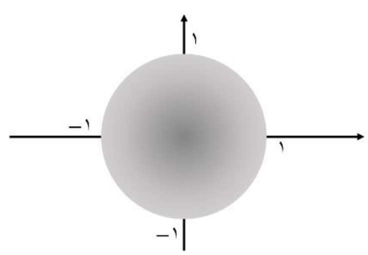
مثال ۱-۵۲) مثال: اگر $x \in \mathbb{R}^۲$ آنگاه $\|x\|_p \leq ۱$ گوی یکه (گوی واحد) نامیده می‌شود. برای نرم‌های گفته‌شده یعنی $p = ۱, p = ۲, p = \infty$ شکل گوی یکه به صورت زیر رسم خواهد شد.

فرض کنید $x = [x_۱, x_۲]^T \in \mathbb{R}^۲$ باشد؛ در این صورت $|x_۱| + |x_۲| \leq ۱$. در واقع تصویر گوی یکه در نرم $p = ۱$ به صورت محیط و داخل چهارضلعی همانند شکل (۱-۳) خواهد بود:



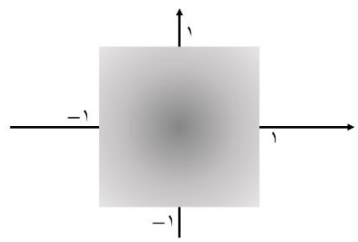
شکل (۱-۳) تصویر گوی واحد در نرم ۱

حال برای $p = ۲$ می‌توان نوشت که $x_۱^۲ + x_۲^۲ \leq ۱$. پس در واقع تصویر گوی واحد در نرم ۲ محیط و داخل دایره به مرکز مبدا و شعاع ۱ به صورت شکل (۱-۴) خواهد بود:



شکل (۱-۴) تصویر گوی واحد در نرم ۲

در انتها برای $p=\infty$ می‌توان نوشت $\|x\|_\infty = max\{|x_1| و |x_2|\} \le 1$. پس در واقع گوی واحد در نرم∞ محیط و داخل مستطیلی به رئوس(۱,۱) و (۱,۱−) و (۱−,۱) و (۱−,۱−) یعنی شکل (۱-۵) خواهد بود:



شکل (۱-۵) تصویر گوی واحد در نرم ∞

لم زیر به لم کوشی-شوارتز معروف است.

لم ۱-۵۳) اگر x و $y \in \mathbb{R}^n$ آنگاه $|x^T y| \le \|x\|_2 \|y\|_2$.

برهان) همان‌طور که در قبل اشاره شده است $\|x\|_2^2 = x^T x = x_1^2 + x_2^2 + \cdots + x_n^2$ که $x = [x_1 \cdots x_n]^T$

حال فرض کنید $0 \neq \lambda \in \mathbb{R}$ باشد، در این صورت طبق خاصیت اول:

$$0 \le \|x+\lambda y\|_2^2 = (x+\lambda y)^T (x+\lambda y) = x^T x + 2\lambda x^T y + \lambda^2 y^T y \Rightarrow$$

$$\|x+\lambda y\|_2^2 = \|x\|_2^2 + 2\lambda x^T y + \lambda^2 \|y\|_2^2$$

حال اگر قرار داده شود $\|y\|_2^2 = \mathrm{a} \ge 0$ ، $\|x\|_2^2 = \mathrm{c} \ge 0$ و $2x^T y = b$ که اعداد حقیقی اند، می‌توان نوشت:

$$a\lambda^2 + b\lambda + c \ge 0$$

یک معادله درجه دو بر حسب λ به وجود می‌آید، چون $a \ge 0$ است، زمانی معادله بالا همواره برقرار است که مقدار $\Delta = b^2 - 4ac$ کوچکتر یا مساوی صفر شود؛ بنابراین می‌توان نوشت:

$$(2x^T y)^2 - 4\|x\|_2^2 \|y\|_2^2 \le 0 \Rightarrow |x^T y| \le \|x\|_2 \|y\|_2 \quad \blacksquare$$

قضیه ۱-۵۴) اگر x و $y \in \mathbb{R}^n$ آنگاه: $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$.

برهان) می‌توان نوشت:

$$\begin{aligned}\|x+y\|_2^2 = (x+y)^T (x+y) &= \|x\|_2^2 + 2x^T y + \|y\|_2^2 \\ &\le \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2 \\ &= (\|x\|_2 + \|y\|_2)^2\end{aligned}$$

بنابراین با جذرگیری می‌توان نوشت $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$. $\blacksquare$

قضیه ۱-۵۵) چنانچه $p, q \in \mathbb{N}$ و $\frac{۱}{p} + \frac{۱}{q} = ۱$ و $x, y \in \mathbb{R}^n$. آنگاه می‌توان نوشت $x^T y = \|x\|_p \|y\|_q$ و سپس با استفاده از رابطه فوق می‌توان نشان داد که $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

برهان) رجوع شود به [۴].

تعریف ۱-۵۶) اگر $x, y \in \mathbb{R}^n$ آنگاه فاصله این دو بردار به صورت $\|x - y\|$ تعریف می شود.

قضیه ۱-۵۷) به ازای هر نرم برداری دلخواه $\|\cdot\|$ ، می‌توان نوشت $\|x - y\| \leq \big|\|x\| - \|y\|\big|$.

برهان) رجوع شود به [۴].

تعریف ۱-۵۸) دو نرم $\|\cdot\|_a$, $\|\cdot\|_b$ را معادل گویند، هرگاه همگرایی به یک نرم، همگرایی به نرم دیگر را نتیجه دهد. به عبارت دیگر هرگاه اعداد ثابت و مثبت M , m موجود باشند به طوری که:

$$\forall x :\ m\|x\|_b \leq \|x\|_a \leq M\|x\|_b$$

گویند دو نرم $\|\cdot\|_a$, $\|\cdot\|_b$ معادلند.

قضیه ۱-۵۹) به ازای هر $x \in \mathbb{R}^n$ می‌توان نوشت:

$$\|x\|_\infty \leq \|x\|_۲ \leq \|x\|_۱ \leq \sqrt{n}\|x\|_۲ \leq n\|x\|_\infty$$

برهان) فرض کنید x_k درایه‌ای از x باشند به طوری که:

$$\|x\|_\infty = \max_{۱\leq \cdot \leq n} |x_i| = |x_k|$$

بنابراین:

$$\begin{aligned}
|x_k|^۲ &\leq \sum_{i=۱}^{n} |x_i|^۲ \leq \left(\sum_{i=۱}^{n} |x_i|\right)^۲ = \left(\sum_{i=۱}^{n} ۱|x_i|\right)^۲ \\
&\leq \left(\sum_{i=۱}^{n} ۱\right) \cdot \left(\sum_{i=۱}^{n} |x_i|^۲\right) = n\sum_{i=۱}^{n} |x_i|^۲ \\
&\leq n\sum_{i=۱}^{n} |x_k|^۲ = n^۲|x_k|^۲
\end{aligned}$$

پس می‌توان نوشت:

۱) $|x_k| \leq (\sum_{i=۱}^{n} |x_i|^۲)^{\frac{۱}{۲}} = \|x\|_۲$

۲) $|x_k| \leq \sum_{i=۱}^{n} |x_i| = \|x\|_۱$

۳) $\|x\|_\infty \leq \|x\|_۲ \leq \|x\|_۱ \leq \sqrt{n}\|x\|_۲ \leq n\|x\|_\infty$ ■

قضیه ۱-۶۰) تمامی نرم‌ها در $\mathbb{R}^n$ معادل هستند. به بیان دیگر تمام نرم‌ها روی فضای متناهی‌البعد معادل هستند.

برهان) رجوع شود به [۴].

در ادامه مفهوم نرم‌های برداری به ماتریس‌های مربعی تعمیم یافته است.

۱-۱-۹ نرم ماتریسی[۱]

تعریف ۱-۶۱) فرض کنید $M_n(\mathbb{R})$ معرف فضای برداری همه ماتریس های $n \times n$ روی $\mathbb{R}$ باشند. در این صورت نگاشت $\|\cdot\|: M_n(\mathbb{R}) \to \mathbb{R}$ را نرم ماتریسی می نامند ، هرگاه در شرایط زیر صدق کند:

۱) $\forall A \in M_n(\mathbb{R}) : \ \|A\| \geq 0$

۲) $\forall A \in M_n(\mathbb{R}), \|A\| = 0 \Leftrightarrow A = \bar{0}$

۳) $\forall \lambda \in \mathbb{R}. \forall A \in M_n(\mathbb{R}), \|\lambda A\| = |\lambda| \|A\|$

۴) $\forall A. B \in M_n(\mathbb{R}), \|A + B\| \leq \|A\| + \|B\|$

۵) $\forall A. B \in M_n(\mathbb{R}), \|AB\| \leq \|A\| \cdot \|B\|$

تعریف ۱-۶۲) یک نرم ماتریسی $\|A\|$ با یک نرم برداری $\|x\|$ "سازگار" نامیده می‌شود هرگاه:

$$\forall x \in \mathbb{R}^n: \|Ax\| \leq \|A\| \cdot \|x\|$$

با توجه به تعریف فوق متناظر با هر نرم برداری مفروض، یک نرم ماتریسی سازگار با آن به صورت:

$$\|A\| = sup \frac{\|Ax\|}{\|x\|} \ , x \neq 0 \qquad (۱-۸)$$

تعریف می‌شود. این نرم را نرم طبیعی، نرم ماتریسی وابسته به نرم برداری و یا نرم القایی می نامند. واضح است که نرم طبیعی، یک نرم ماتریسی سازگار است.

اگر قرار داده شود $z = \frac{x}{\|x\|}$ آنگاه برای $z \neq \bar{0}$ می‌توان نوشت:

$$\begin{cases} \|z\| = 1 \\ \dfrac{\|Ax\|}{\|x\|} = \|AZ\| \end{cases}$$

بنابراین رابطه (۱-۸) معادل است با $\|A\| = sup\|AZ\|$. حال چون مجموعه $\{Z: \|Z\| = 1\}$ مجموعه‌ای بسته و کراندار است و $\|\cdot\|$ تابعی پیوسته است بنابراین رابطه (۱-۸) معادل است با:

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

بنابراین رابطه فوق به عنوان تعریف نرم طبیعی در نظر گرفته می‌شود.

---

[۱] Matrix Norm

قضیه ۱-۶۳) اگر$\|\cdot\|$ نرمی در$\mathbb{R}^n$ باشد ، آنگاه رابطه $\|A\| = \max_{\|x\|=1} \|Ax\|$ یک نرم (طبیعی) بر روی فضای خطی همه ماتریس های $n \times n$ تعریف می کند.

برهان) باید نشان دهیم ویژگی های ۱ تا ۵ را داراست:

۱) از آنجا که$\|Ax\|$ یک نرم برداری است، بنابراین $\|Ax\| \geq 0$ در نتیجه $\|A\| = \max_{\|x\|=1} \|Ax\| \geq 0$.

۲) اگر $A = \bar{0}$ بنابرایم برای هر بردار x ، $Ax = \bar{0}$ و بنابراین $\|Ax\| = 0$ و در نتیجه $\|A\| = 0$. حال فرض کنید $\|A\| = 0$ بنابراین برای هر x که $\|x\| = 1$ ، $\max_{\|x\|=1} \|Ax\| = 0$ یعنی $\|Ax\| = 0$.
حال فرض کنید $x = e_i$ بنابراین $\|x\| = 1$ و Ax برابر ستون i ام ماتریس A خواهد بود. بنابراین چون $\|Ax\| = 0$ در نتیجه نرم بردار ستون i ام برابر صفر است و طبق خاصیت نرم برداری، ستون i ام ماتریسA صفر خواهد بود. حال چون i می‌تواند از ۱ تا n تغییر کند، بنابراین $A = \bar{0}$.
حال هدف این است که نشان داده شود برای $A \neq 0$ ، $\|A\| \geq 0$. چون $A = \bar{0}$ بنابراین حداقل یک ستون مخالف دارد. فرض کنید این ستون مخالف صفر، ستون j ام ماتریس A باشد که با $A^{(j)}$ نمایش داده می‌شود. حال چنانچه قرار داده شود $x = e_j$ ، آنگاه $\|x\| = 1$ و می‌توان نوشت:

$$\|A\| = \max_{\|x\|=1} \|Ax\| \geq \|Ax\| > 0$$

۳) فرض کنید $\lambda \in \mathbb{R}$ بنابراین:

$$\|\lambda A\| = \max_{\|x\|=1} \|\lambda Ax\| = \max_{\|x\|=1} |\lambda| \|Ax\| = |\lambda| \max_{\|x\|=1} \|Ax\| = |\lambda| \|Ax\|$$

۴)
$$\|A + B\| = \max_{\|x\|=1} \|(A + B)x\| = \max_{\|x\|=1} \|(Ax + Bx)\| \leq \max_{\|x\|=1} (\|Ax\| + \|Bx\|)$$
$$= \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\| + \|B\|$$

۵)
$$\|AB\| = \max_{\|x\|=1} \|ABx\| = \max_{\|x\|=1} \|A(Bx)\| \leq \max_{\|x\|=1} (\|A\| \|Bx\|)$$
$$\leq \max_{\|x\|=1} (\|A\| \|B\| \|X\|) = \|A\| \|B\| \quad \blacksquare$$

قضیه ۱-۶۴) نرم طبیعی متناظر با نرم برداری $\|\cdot\|_\infty$ به صورت زیر

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| \qquad (۹-۱)$$

مشخص می‌شود.

برهان) طبق تعریف نرم ماتریسی متناسب با نرم برداری ، $\|\cdot\|_\infty$داریم:

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{\|x\|_\infty=1} \left(\max_{1 \leq i \leq n} |(Ax)_i| \right) = \max_{1 \leq i \leq n} \left(\max_{\|x\|_\infty=1} |(Ax)_i| \right) \Rightarrow$$

$$\|A\|_\infty = \max_{1\le i\le n}\left(\max_{\|x\|_\infty=1}\left|\sum_{j=1}^{n} a_{ij}x_j\right|\right) \le \max_{1\le i\le n}\left(\max_{\|x\|_\infty=1}\sum_{j=1}^{n}|a_{ij}x_j|\right) \le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$$

حال فرض کنید p عدد صحیحی باشد به طوری که $1 \le p \le n$ و

$$\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| = \sum_{j=1}^{n}|a_{pj}| \quad \text{(I)}$$

حال بردار $x = [x_1, x_2, \dots, x_n]^T$ را به صورت زیر در نظر بگیرید:

$$x_j = \begin{cases} 1 \ , & a_{pj} \ge 0 \\ -1 \ , & a_{pj} < 0 \end{cases}$$

به وضوح $\|x\|_\infty = 1$ و همچنین برای $j = 1,2,\cdots,n$ می‌توان نوشت:

$$a_{pj}x_j = |a_{pj}|$$

از طرفی می‌توان نوشت:

$$\|Ax\|_\infty = \max_{\|x\|_\infty=1}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \ge \left|\sum_{j=1}^{n} a_{pj}x_j\right| = \left|\sum_{j=1}^{n}|a_{pj}|\right| = \sum_{j=1}^{n}|a_{pj}| = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$$

بنابراین:

$$\|A\|_\infty = max\|Ax\|_\infty \ge \|Ax\|_\infty \ge max\sum_{j=1}^{n}|a_{ij}| \quad \text{(II)}$$

از (I) و (II) حکم به دست می‌آید. ■

قضیه ۱-۶۵) نرم طبیعی متناظر با نرم برداری $\|\cdot\|_1$ به صورت زیر

$$\|A\|_1 = \max_{1\le j\le n}\sum_{j=1}^{n}|a_{ij}| \qquad (1\text{-}10)$$

مشخص می‌شود.

برهان) به طور مشابه با قضیه (۱-۶۴) اثبات می‌شود.

تعریف ۱-۶۶) برای ماتریس $A \in \mathbb{R}^{n\times n}$ ، نرم فروبنیوس[1]به صورت زیر تعریف می‌شود:

$$\|A\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{2}\right)^{\frac{1}{2}} \quad , \quad A = (a_{ij}) \tag{11-1}$$

قضیه ۱-۶۷) اثبات کنید که رابطه (۱-۱۱) یک نرم ماتریسی است.

برهان) خواص ۱ تا ۳ به سادگی قابل اثبات هستند. اثبات خواص ۴ و ۵ به صورت زیر می‌باشد:

$$\begin{aligned}\|A+B\|_F^{2} &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}+b_{ij}\right)^{2}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}^{2}+2a_{ij}b_{ij}+b_{ij}^{2}\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^{2}+\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}^{2}+2\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_{ij} \quad \text{(I)}\end{aligned}$$

از طرفی طبق نامساوی کوشی-شوارتز می‌توان نوشت:

$$\sum_{j=1}^{n}a_{ij}b_{ij} \leq \left(\sum_{j=1}^{n}a_{ij}^{2}\right)^{\frac{1}{2}}\left(\sum_{j=1}^{n}b_{ij}^{2}\right)^{\frac{1}{2}}$$

با استفاده مجدد از نامساوی کوشی-شوارتز می‌توان نوشت:

$$2\sum_{j=1}^{n}x_i y_i \leq \left(\sum_{j=1}^{n}x_i^{2}\right)^{\frac{1}{2}}\left(\sum_{j=1}^{n}y_i^{2}\right)^{\frac{1}{2}}$$

بنابراین:

$$2\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_{ij} \leq 2\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^{2}\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}^{2}\right)^{\frac{1}{2}} = 2\|A\|_F\|B\|_F \quad \text{(II)}$$

[1] Frobenius Norm

حال از رابطه (I) و (II) می‌توان نوشت:

$$\|A+B\|_F^۲ \leq \|A\|_F^۲ + \|B\|_F^۲ + ۲\|A\|_F\|B\|_F = (\|A\|_F + \|B\|_F)^۲$$
$$\Rightarrow \|A+B\|_F \leq \|A\|_F + \|B\|_F$$

حال برای اثبات خاصیت ۵ نرم ماتریسی، ابتدا باید یادآوری شود که درایه سطر i ام و ستون j ام ماتریس AB را با $(AB)_{ij}$ نمایش می‌دهند و می‌توان نوشت:

$$(AB)_{ij} = \sum_{k=۱}^{n} a_{ik}b_{kj}$$

از طرفی با استفاده از نامساوی کوشی-شوارتز می‌توان نوشت:

$$(AB)_{ij}^۲ = \left(\sum_{k=۱}^{n} a_{ik}b_{kj}\right)^۲ \leq \left(\sum_{k=۱}^{n} a_{ik}^۲\right)\left(\sum_{k=۱}^{n} b_{kj}^۲\right)$$

حال می‌توان نوشت:

$$\|AB\|_F^۲ = \sum_{i=۱}^{n}\sum_{j=۱}^{n}\left(\sum_{k=۱}^{n} a_{ik}b_{kj}\right)^۲ \leq \sum_{i=۱}^{n}\sum_{j=۱}^{n}\left(\sum_{k=۱}^{n} a_{ik}^۲ \sum_{k=۱}^{n} b_{kj}^۲\right)$$
$$= \left(\sum_{i=۱}^{n}\sum_{k=۱}^{n} a_{ik}^۲\right)\left(\sum_{k=۱}^{n}\sum_{j=۱}^{n} b_{kj}^۲\right) = \|A\|_F^۲\|B\|_F^۲$$

و در نهایت حکم ثابت خواهد شد. ■

تعریف ۱-۶۸) فرض کنید A یک ماتریس $n \times n$ باشد. شعاع طیفی ماتریس A را با $\rho(A)$ نمایش می‌دهند و به صورت $\rho(A) = \max_{۱\leq i \leq n} |\lambda_i|$ تعریف می‌شود که در آن λ_i ها، مقادیر ویژه ماتریس A می‌باشند.

قضیه ۱-۶۹) اگر A یک ماتریس $n \times n$ باشد. آنگاه $\|A\|_۲ = \sqrt{\rho(A^TA)}$ و به ازای هر نرم طبیعی $\|\cdot\|$ ، می‌توان گفت $\rho(A) \leq \|A\|$.

برهان) فرض کنید λ یک مقدار ویژه دلخواه ماتریس A باشد و x بردار ویژه نظیر λ باشد. بنابراین می‌توان نوشت $Ax = \lambda x$

پس:

$$\|\lambda x\| = \|Ax\| \Rightarrow |\lambda|\|x\| = \|Ax\| \leq \|A\|\|x\| \Rightarrow |\lambda| \leq \|A\|$$

چون λ دلخواه بوده است، بنابراین $\rho(A) = max\{|\lambda|\} \leq \|A\|$. ■

مثال ۱-۷۰) برای ماتریس $A=\begin{bmatrix} ۳ & -۱ \\ -۲ & ۲ \end{bmatrix}$ مقدار $\|A\|_۲$ به صورت زیر محاسبه می‌شود.

طبق قضیه (۱-۶۹) $\|A\|_۲^۲=\rho(A^TA)$. بنابراین:

$$A^TA=\begin{bmatrix} ۱۳ & -۷ \\ -۷ & ۵ \end{bmatrix}$$

حال باید مقدار ویژه هایA^TA را محاسبه کرده و سپس ماکزیمم این مقادیر را استخراج نمود. پس:

$$det(\lambda I-A^TA)=\cdot \Rightarrow \lambda^۲-۱۸\,\lambda+۱۶=\cdot \Rightarrow \lambda=۹\pm\sqrt{۶۵}$$

بنابراین:

$$\|A\|_۲=\sqrt{\rho(A^TA)}=\sqrt{۹+\sqrt{۶۵}}\approx ۴/۱۳$$

همچنین می توان دریافت که $\|A\|_\infty=۴ \quad , \quad \|A\|_۱=۵$.

۱-۲ دستگاه معادلات خطی[۱]

تعریف ۱-۷۱) گردایه‌ای متناهی از معادلات خطی بر حسب متغیرهای $x_۱,x_۲,\cdots,x_n$ را یک دستگاه معادلات خطی گویند.

فرم کلی یک دستگاه معادلات خطی با n معادله و n مجهول به صورت زیر می‌باشد.

$$\begin{cases} a_{۱۱}x_۱+a_{۱۲}x_۲+\cdots+a_{۱n}x_n=b_۱ \\ a_{۲۱}x_۱+a_{۲۲}x_۲+\cdots+a_{۲n}x_n=b_۲ \\ \vdots \\ a_{n۱}x_۱+a_{n۲}x_۲+\cdots+a_{nn}x_n=b_n \end{cases} \qquad (۱۲\text{-}۱)$$

که در آن x_i ها مجهول و b_i ها و a_{ij} ها معلوم هستند.

در واقع هدف از حل دستگاه (۱-۱۲)، محاسبه x_i ها (در صورت وجود) می‌باشد. فرم ساده تری نیز برای نمایش دستگاه (۱-۱۲) وجود دارد که همان فرم ماتریسی است. دستگاه فوق را می‌توان به فرم مناسب تر $AX=b$ بازنویسی کرد. فرم $AX=b$ در زیر نمایش داده شده است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & a_{۱۳} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & a_{۲۳} & \cdots & a_{۲n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & a_{n۳} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} b_۱ \\ b_۲ \\ \vdots \\ b_n \end{bmatrix}$$

[۱] System of Linear Equations

تذکر ۱-۷۲) یک دستگاه معادلات خطی یا اصلا جواب ندارد که در این حالت به آن دستگاه، ناسازگار گویند و یا دقیقاً یک جواب دارد یا بی‌شمار جواب دارد که در این دو حالت اخیر به این دستگاه، سازگار گویند.

تذکر ۱-۷۳) دستگاه $AX = O$ همگن نام دارد و $X = O$ همواره جواب بدیهی این دستگاه است.

تعریف ۱-۷۴) ماتریس افزوده دستگاه معادلات خطی (۱-۱۲)، که با $[A|b]$ نمایش داده می‌شود؛ به صورت زیر نمایش داده می‌شود:

$$[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

تعریف ۱-۷۵) به دو دستگاه $AX = b$ و $CY = d$ ، هم‌ارز گویند هرگاه جواب‌های دو دستگاه یکسان باشند؛ یا به عبارت دیگر $X = Y$.

قضیه ۱-۷۶) دستگاه $AX = b$ دارای جواب یکتاست اگر و تنها اگر $\det(A) \neq 0$.

برهان) رجوع شود به [۱].

قضیه ۱-۷۷) اگر A یک ماتریس $n \times n$ باشد، گزاره‌های زیر معادلند:

۱) A وارون‌پذیر است.
۲) A هم‌ارز سطری I_n است.
۳) دستگاه همگن $AX = o$ فقط دارای جواب بدیهی است.
۴) دستگاه $AX = b$ به ازای هر برداری ستونی b دارای جواب منحصر به فرد $X = A^{-1}b$ است.

برهان) رجوع شود به [۱].

فصل دوم

ماتریس وارون

فصل دوم: ماتریس وارون

مطالب این فصل از [۴] آورده شده است. در فصل قبل، ماتریس وارون به‌طور مختصر توضیح داده شد و گفته شد که $AA^{-1} = A^{-1}A = I_n$؛ همچنین ذکر شد که ماتریس A وارون‌پذیر است اگر و تنها اگر دترمینان آن مخالف صفر باشد. ولی روشی برای یافتن وارون یک ماتریس و همچنین کاربرد آن در علوم داده ذکر نشد. هدف این فصل اشاره به روش‌های پیدا کردن وارون ماتریس و کاربرد آن در علوم داده است. می‌توان روش‌های تعیین وارون ماتریس را به‌صورت زیر دسته‌بندی کرد:

روش تعیین وارون ماتریس:
- به کمک دترمینان
- به کمک اعمال سطری مقدماتی
- به کمک تجزیه مثلثی $A = LU$

۲-۱ دترمینان

ماتریس مربعی $A_{n\times n}$ با درایه‌هایی در میدان F موجود است.

تعریف ۲-۱) به ماتریس حاصل از حذف سطر i ام و ستون j ام ماتریس مربعی A، کهاد ij ام ماتریس A[۱] گویند و آن را با A_{ij} نشان می‌دهند.

تعریف ۲-۲) به مقدار حاصل از رابطه زیر، همسازه ij ام ماتریس A[۲] گویند و آن را با C_{ij} نمایش می‌دهند.

$$C_{ij} = (-1)^{i+j}\det(A_{ij}) \tag{۱-۲}$$

تعریف ۲-۳) به ماتریس زیر که درایه‌های آن c_{ij} هستند، ماتریس همسازه‌ها[۳] گویند.

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}$$

تعریف ۲-۴) به ترانهاده‌ی ماتریس همسازه‌ها، ماتریس الحاقی ماتریس A[۴] گویند و آن را با $adj(A)$ نمایش می‌دهند.

$$adj(A) = C^T \tag{۲-۲}$$

حال با توجه رابطه زیر، می‌توان وارون ماتریس A را پیدا کرد.

$$A^{-1} = \frac{adj(A)}{\det(A)} \tag{۳-۲}$$

---

[۱] Minor

[۲] Cofactor

[۳] Matrix of Cofactors

[۴] Adjugate Matrix

مثال ۲-۵) ماتریس وارون $A = \begin{bmatrix} ۱ & ۰ & ۲ \\ ۱ & -۱ & ۱ \\ ۰ & ۱ & ۲ \end{bmatrix}$ به صورت زیر به دست می‌آید.

ابتدا باید همسازه‌های ماتریس را طبق فرمول (۲-۱) به دست آورد:

$$C_{۱۱} = (-۱)^{۱+۱} \begin{vmatrix} -۱ & ۱ \\ ۱ & ۲ \end{vmatrix} = -۳ \qquad C_{۱۲} = (-۱)^{۱+۲} \begin{vmatrix} ۱ & ۱ \\ ۰ & ۲ \end{vmatrix} = -۲$$

$$C_{۱۳} = (-۱)^{۱+۳} \begin{vmatrix} ۱ & -۱ \\ ۰ & ۱ \end{vmatrix} = ۱ \qquad C_{۲۱} = (-۱)^{۲+۱} \begin{vmatrix} ۰ & ۲ \\ ۱ & ۲ \end{vmatrix} = ۲$$

$$C_{۲۲} = (-۱)^{۲+۲} \begin{vmatrix} ۱ & ۲ \\ ۰ & ۲ \end{vmatrix} = ۲ \qquad C_{۲۳} = (-۱)^{۲+۳} \begin{vmatrix} ۱ & ۰ \\ ۰ & ۱ \end{vmatrix} = -۱$$

$$C_{۳۱} = (-۱)^{۳+۱} \begin{vmatrix} ۰ & ۲ \\ -۱ & ۱ \end{vmatrix} = ۲ \qquad C_{۳۲} = (-۱)^{۳+۲} \begin{vmatrix} ۱ & ۲ \\ ۱ & ۱ \end{vmatrix} = ۱$$

$$C_{۳۳} = (-۱)^{۳+۳} \begin{vmatrix} ۱ & ۰ \\ ۱ & -۱ \end{vmatrix} = -۱$$

سپس باید ماتریس همسازه‌ها را تشکیل داد:

$$C(A) = \begin{bmatrix} -۳ & -۲ & ۱ \\ ۲ & ۲ & -۱ \\ ۲ & ۱ & -۱ \end{bmatrix}$$

حال ترانهاده ماتریس همسازه‌ها یعنی ماتریس الحاقی را باید پیدا کرد:

$$adj(A) = C^T(A) = \begin{bmatrix} -۳ & ۲ & ۲ \\ -۲ & ۲ & ۱ \\ ۱ & -۱ & -۱ \end{bmatrix}$$

برای به دست آوردن ماتریس وارون، ابتدا باید دترمینان ماتریس محاسبه شود:

$$\det(A) = |A| = ۱ \times \begin{vmatrix} -۱ & ۱ \\ ۱ & ۲ \end{vmatrix} - ۰ \times \begin{vmatrix} ۱ & ۱ \\ ۰ & ۲ \end{vmatrix} + ۲ \times \begin{vmatrix} ۱ & -۱ \\ ۰ & ۱ \end{vmatrix} \Rightarrow$$

$$|A| = \Big(۱ \times \Big(\big((-۱) \times ۲\big) - (۱ \times ۱)\Big)\Big) - ۰ + \Big(۲ \times \Big((۱ \times ۱) - \big((-۱) \times ۰\big)\Big)\Big) \Rightarrow$$

$$|A| = -۳ + ۰ + ۲ = -۱$$

در نهایت با توجه به فرمول (۲-۳) به مقدار ماتریس وارون می‌توان رسید:

$$A^{-۱} = \frac{adj(A)}{|A|} = \frac{\begin{bmatrix} -۳ & ۲ & ۲ \\ -۲ & ۲ & ۱ \\ ۱ & -۱ & -۱ \end{bmatrix}}{-۱} = \begin{bmatrix} ۳ & -۲ & -۲ \\ ۲ & -۲ & -۱ \\ -۱ & ۱ & ۱ \end{bmatrix}$$

۲-۲ اعمال سطری مقدماتی

برای پیدا کردن ماتریس وارون با استفاده از این روش، ابتدا باید ماتریس افزوده $[A|I_n]$ را تشکیل داد و سپس با انجام اعمال سطری مقدماتی روی ماتریس افزوده، ماتریس A را به ماتریس همانی I_n تبدیل کرد. با انجام این کار ماتریس I_n در ماتریس افزوده $[A|I_n]$ به $A^{-۱}$ تبدیل خواهد شد.

به طور کامل‌تر می‌توان به صورت زیر نوشت:

$$\left[\begin{array}{cccc|cccc} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} & ۱ & ۰ & \cdots & ۰ \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} & ۰ & ۱ & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} & ۰ & ۰ & \cdots & ۱ \end{array}\right] \xrightarrow{\text{اعمال سطری مقدماتی}}$$

$$\left[\begin{array}{cccc|cccc} ۱ & ۰ & \cdots & ۰ & a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & \cdots & a_{۱n}^{(-۱)} \\ ۰ & ۱ & \cdots & ۰ & a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & \cdots & a_{۲n}^{(-۱)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & ۱ & a_{n۱}^{(-۱)} & a_{n۲}^{(-۱)} & \cdots & a_{nn}^{(-۱)} \end{array}\right]$$

مثال ۲-۶) برای پیدا کردن وارون ماتریس $A = \begin{bmatrix} ۱ & -۱ & ۲ \\ ۳ & ۲ & ۴ \\ ۰ & ۱ & -۲ \end{bmatrix}$ به کمک اعمال سطری مقدماتی باید به صورت زیر عمل کرد.

ابتدا ماتریس افزوده $[A|I_n]$ را باید تشکیل داد:

$$\left[\begin{array}{ccc|ccc} ۱ & -۱ & ۲ & ۱ & ۰ & ۰ \\ ۳ & ۲ & ۴ & ۰ & ۱ & ۰ \\ ۰ & ۱ & -۲ & ۰ & ۰ & ۱ \end{array}\right]$$

سپس با کمک اعمال سطری مقدماتی باید ماتریس افزوده فوق را به صورت $[I_n|A^{-۱}]$ تبدیل کرد:

$$\left[\begin{array}{ccc|ccc} ۱ & -۱ & ۲ & ۱ & ۰ & ۰ \\ ۳ & ۲ & ۴ & ۰ & ۱ & ۰ \\ ۰ & ۱ & -۲ & ۰ & ۰ & ۱ \end{array}\right] \xrightarrow{-۳R_۱+R_۲\to R_۲} \left[\begin{array}{ccc|ccc} ۱ & -۱ & ۲ & ۱ & ۰ & ۰ \\ ۰ & ۵ & -۲ & -۳ & ۱ & ۰ \\ ۰ & ۱ & -۲ & ۰ & ۰ & ۱ \end{array}\right] \xrightarrow{{}^{۱}\!/_{۵}\, R_۲\to R_۲}$$

$$\left[\begin{array}{ccc|ccc} ۱ & -۱ & ۲ & ۱ & ۰ & ۰ \\ ۰ & ۱ & -\frac{۲}{۵} & -\frac{۳}{۵} & -\frac{۱}{۵} & ۰ \\ ۰ & ۱ & -۲ & ۰ & ۰ & ۱ \end{array}\right] \xrightarrow{R_۲+R_۱\to R_۱} \left[\begin{array}{ccc|ccc} ۱ & ۰ & \frac{۸}{۵} & \frac{۲}{۵} & \frac{۱}{۵} & ۰ \\ ۰ & ۱ & -\frac{۲}{۵} & -\frac{۳}{۵} & -\frac{۱}{۵} & ۰ \\ ۰ & ۱ & -۲ & ۰ & ۰ & ۱ \end{array}\right] \xrightarrow{-R_۲+R_۳\to R_۳}$$

$$\left[\begin{array}{ccc|ccc} ۱ & ۰ & \frac{۸}{۵} & \frac{۲}{۵} & \frac{۱}{۵} & ۰ \\ ۰ & ۱ & -\frac{۲}{۵} & -\frac{۳}{۵} & -\frac{۱}{۵} & ۰ \\ ۰ & ۰ & -\frac{۸}{۵} & \frac{۳}{۵} & -\frac{۱}{۵} & ۱ \end{array}\right] \xrightarrow{R_۳+R_۱\to R_۱} \left[\begin{array}{ccc|ccc} ۱ & ۰ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۱ & -\frac{۲}{۵} & -\frac{۳}{۵} & -\frac{۱}{۵} & ۰ \\ ۰ & ۰ & -\frac{۸}{۵} & \frac{۳}{۵} & -\frac{۱}{۵} & ۱ \end{array}\right] \xrightarrow{-{}^{۵}/_{۸}R_۳\to R_۳}$$

$$\left[\begin{array}{ccc|ccc} ۱ & ۰ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۱ & -\frac{۲}{۵} & -\frac{۳}{۵} & -\frac{۱}{۵} & ۰ \\ ۰ & ۰ & ۱ & -\frac{۳}{۸} & \frac{۱}{۸} & -\frac{۵}{۸} \end{array}\right] \xrightarrow{{}^{۲}/_{۵}R_۳+R_۲\to R_۲} \left[\begin{array}{ccc|ccc} ۱ & ۰ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۱ & ۰ & -\frac{۳}{۴} & \frac{۱}{۴} & -\frac{۱}{۴} \\ ۰ & ۰ & ۱ & -\frac{۳}{۸} & \frac{۱}{۸} & -\frac{۵}{۸} \end{array}\right]$$
$$= [I_n|A^{-۱}]$$

در نهایت وارون ماتریس A به دست آمد که برابر است با:

$$A^{-۱} = \begin{bmatrix} ۱ & ۰ & ۱ \\ -\frac{۳}{۴} & \frac{۱}{۴} & -\frac{۱}{۴} \\ -\frac{۳}{۸} & \frac{۱}{۸} & -\frac{۵}{۸} \end{bmatrix}$$

قضیه ۲-۷) اگر ماتریس A دارای یک سطر صفر باشد، آنگاه A وارون‌پذیر نیست.

برهان) فرض خلف بگیرید که A وارون‌پذیر است. پس ماتریس B موجود است به طوری که $BA = AB = I$. فرض کنید سطر i ام ماتریس A صفر باشد؛ لذا طبق تعریف ضرب ماتریس‌ها سطر i ام ماتریس AB برابر صفر است. لذا سطر i ام ماتریس همانی I_n صفر است که تناقض است. پس فرض خلف باطل و حکم برقرار است. ■

تذکر ۲-۸) اگر یکی از سطرها یا ستون‌های ماتریس A، با انجام اعمال سطری مقدماتی صفر شود، آنگاه ماتریس A منفرد است و وارون ندارد.

۲-۳ تجزیه مثلثی LU

تجزیه مثلثی یک ماتریس بدین صورت است که ماتریس A را به حاصل‌ضرب ماتریس‌های پایین مثلثی L و بالا مثلثی U (یا به عبارت دیگر $A = LU$) می‌توان تجزیه کرد. چنانچه این امکان وجود داشته باشد که ماتریس A را به صورت $A = LU$ بازنویسی کرد، در این صورت می‌توان به سادگی وارون ماتریس را از طریق حل $AA^{-۱} = I$ پیدا کرد. روش این کار در ادامه توضیح داده شده است.

اگر ماتریس A به حاصل‌ضرب LU تجزیه شده باشد در این صورت دستگاه $AA^{-۱} = I$ به صورت $LUA^{-۱} = I$ بازنویسی می‌شود. چنانچه قرار داده شود $UA^{-۱} = B$، آنگاه دستگاه $LUA^{-۱} = I$ به صورت $LB = I$ خواهد

شد. چون L ماتریسی مثلثی است بنابراین با استفاده از جایگذاری می‌توان B را به دست آورد. سپس با حل دستگاه $UA^{-1} = B$ و پس از جایگذاری‌های لازم، ماتریس وارون یعنی A^{-1} به دست خواهد آمد.

صورت کلی ماتریس $A_{n\times n}$ و ماتریس‌های پایین مثلثی L و بالا مثلثی U به صورت زیر است:

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

در اینجا هدف محاسبه u_{ij} و l_{ij} است. تعداد مجهول ها $n^2 + n$ و تعداد معلوم‌ها (یعنی همان a_{ij}) n^2 می‌باشد. بنابراین برای یافتن مجهول‌های u_{ij} و l_{ij}، n درجه آزادی وجود دارد. این بدین معنی است که می‌توان n مجهول را به دلخواه خود انتخاب کرد. سه راه متفاوت برای این انتخاب پیشنهاد شده است:

- تجزیه دولیتل[1]: اعضای قطری L، یک فرض می‌شوند (یعنی $l_{ii} = 1 , i = 1, \cdots , n$).
- تجزیه کروت[2]: اعضای قطری U، یک فرض می‌شوند (یعنی $u_{ii} = 1 , i = 1, \cdots , n$).
- تجزیه چولسکی[3]: اعضای قطری L و U برابر فرض می‌شوند (یعنی $l_{ii} = u_{ii} , i = 1, \cdots , n$).

با این انتخاب‌ها، در واقع تعداد معلومات و مجهولات برابر شده و تجزیه‌ای یکتا به دست می‌آید.

تعریف۲-۹) به تجزیه $A = LU$ ، تجزیه دولیتل گویند؛ هرگاه اعضای قطری ماتریس L برابر ۱ باشند.

تعریف۲-۱۰) به تجزیه $A = LU$ ، تجزیه کروت گویند؛ هرگاه اعضای قطری ماتریس U برابر ۱ باشند.

تعریف۲-۱۱) به تجزیه $A = LU$ ، تجزیه چولسکی گویند؛ هرگاه اعضای قطری ماتریس L با اعضای قطری ماتریس U برابر باشند.

قضیه ۲-۱۲) اگر ماتریس $A_{n\times n}$ ماتریسی اکیداً قطر غالب باشد، آنگاه A دارای تجزیه LU می‌باشد.

برهان) رجوع شود به [۴].

قضیه ۲-۱۳) اگر ماتریس $A_{n\times n}$ ماتریسی معین مثبت اکید باشد، آنگاه A دارای تجزیه LU می‌باشد.

برهان) رجوع شود به [۴].

[1] Doolittle
[2] Crout
[3] Cholesky

قضیه ۲-۱۴) اگر تمام زیرماتریس‌های اصلی پیشروی $A_{n\times n}$، وارون‌پذیر باشند آنگاه A دارای تجزیه LU می‌باشد.

برهان) اگر $A_{n\times n} = (a_{ij})$ که در آن $i,j = ۱,۲,\cdots,n$ باشند و $A_۱,\cdots,A_n$ نیز معرف زیرماتریس‌های اصلی پیشروی ماتریس A باشند، باید ثابت کرد که $A_n = A$ دارای تجزیه LU است. مسئله به روش استقراء روی n اثبات خواهد شد. برای $n = ۱$ حکم برقرار است. حال فرض کنید که برای $n = k - ۱$ نیز حکم برقرار باشد؛ یعنی $A_{k-۱}$ دارای تجزیه LU به فرم $A_{k-۱} = L_{k-۱}U_{k-۱}$ باشد که در آن $\det(L_{k-۱}) = ۱$ است. از آنجا که $A_{k-۱}$ وارون‌پذیر است، بنابراین $\det(A_{k-۱}) = \det(L_{k-۱})\det(U_{k-۱}) = \det(U_{k-۱}) \neq ۰$. یعنی $U_{k-۱}$ وارون‌پذیر است.

حال باید ثابت کرد که A_k نیز دارای تجزیه LU به فرم $A_k = L_kU_k$ می‌باشد. قرار دهید:

$$\sum_{s=۱}^{k-۱} l_{is}u_{sk} = a_{ik} \quad ۱ \leq i \leq k-۱ \tag{۲-۴}$$

که در این صورت $k-۱$ معادله ساخته می‌شود که $u_{k-۱,k},\cdots,u_{۲,k},u_{۱,k}$ در آن مجهول هستند. ماتریس ضرایب این دستگاه $L_{k-۱}$ است که وارون‌پذیر است. همچنین a_{ik} عناصر ستون k ام ماتریس A_k می‌باشد. جواب به دست آمده را در ستون k ام ماتریس $U_{k-۱}$ قرار دهید. حال قرار دهید:

$$\sum_{s=۱}^{k-۱} l_{ks}u_{sj} = a_{kj} \quad ۱ \leq j \leq k-۱ \tag{۲-۵}$$

به طریق مشابه $k-۱$ معادله ساخته می‌شود که $l_{k,۱},l_{k,۲},\cdots,l_{k,k-۱}$ مجهول می‌باشند. ماتریس ضرایب این بار $U_{k-۱}$ می‌باشد که وارون‌پذیر است. پس دستگاه بالا جواب یکتا دارد. جواب به دست آمده را در سطر k ام ماتریس $L_{k-۱}$ قرار دهید. حال مجدداً قرار دهید:

$$\sum_{s=۱}^{k} l_{ks}u_{sk} = a_{kk} \Rightarrow \sum_{s=۱}^{k-۱} l_{ks}u_{sk} + l_{kk}u_{kk} = a_{kk} \tag{۲-۶}$$

که در این معادله u_{kk} و l_{kk} مجهول می‌باشند. با فرض $l_{kk} = ۱$، u_{kk} از معادله بالا محاسبه شده و به این ترتیب سطر k ام در ماتریس $L_{k-۱}$ و ستون k ام در ماتریس $U_{k-۱}$ نیز تکمیل می‌شود. این دو ماتریس جدید را L_k و U_k نامگذاری کنید. با توجه به معادلات تولید این سطر و ستون جدید، به $A_k = L_kU_k$ خواهید رسید. پس حکم ثابت شد. ■

یکی از روش های تجزیه LU استفاده از روش حذفی گاوس می‌باشد؛ ولی این روش به طور کامل در فصل بعد یعنی فصل دستگاه معادلات خطی توضیح داده خواهد شد. در ادامه توضیح دقیق و حل مثال‌هایی از روش های مستقیم تجزیه LU (یعنی تجزیه‌های دولیتل، کروت و چولسکی) آورده شده است.

۲-۳-۱ تجزیه دولیتل

همانطور که قبلاً ذکر شد، در تجزیه دولیتل، قطر اصلی ماتریس L برابر با ۱ در نظر خواهد گرفته شد. پس ماتریس $A_{n\times n}$ و ماتریس‌های پایین مثلثی L و بالا مثلثی U به صورت زیر مفروض است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & ۱ & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & ۱ \end{bmatrix} \begin{bmatrix} u_{۱۱} & u_{۱۲} & u_{۱۳} & \cdots & u_{۱n} \\ ۰ & u_{۲۲} & u_{۲۳} & \cdots & u_{۲n} \\ ۰ & ۰ & u_{۳۳} & \cdots & u_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & u_{nn} \end{bmatrix}$$

از ضرب دو ماتریس L و U، $n^۲$ معادله به دست می‌آید.

$$u_{۱۱} = a_{۱۱}\,, \qquad u_{۱۲} = a_{۱۲}\,, \cdots, \qquad u_{۱n} = a_{۱n} \tag{۲-۷}$$
$$l_{۲۱}u_{۱۱} = a_{۲۱}\,, \quad l_{۲۱}u_{۱۲} + u_{۲۲} = a_{۲۲}\,, \cdots, \quad l_{۲۱}u_{۱n} + u_{۲n} = a_{۲n} \tag{۲-۸}$$
$$\vdots \qquad\qquad \vdots \qquad \vdots \qquad\qquad \vdots$$
$$l_{n۱}u_{۱۱} = a_{n۱}\,, l_{n۱}u_{۱۲} + l_{n۲}u_{۲۲} = a_{n۲}\,, \cdots, l_{n۱}u_{۱n} + \cdots + u_{nn} = a_{nn} \tag{۲-۹}$$

با توجه به معادلات (۲-۷) به سادگی می‌توان مقادیر $u_{۱۱}, u_{۱۲}, u_{۱۳}, \cdots, u_{۱n}$ را به دست آورد. پس از آن با جایگذاری مقادیر مذکور در معادلات (۲-۸) می‌توان به ترتیب مقادیر $l_{۲۱}, u_{۲۲}, u_{۲۳}, \cdots, u_{۲n}$ را به دست آورد. به همین ترتیب، پس از انجام همه معادلات تا معادلات (۲-۹)، همه مقادیر مجهول به دست می‌آیند و ماتریس A به ماتریس‌های L و U تجزیه خواهد شد.

سپس با جایگذاری فرم LU به جای ماتریس A در رابطه $AA^{-۱} = I$، به صورت $LUA^{-۱} = I$ بازنویسی می‌شود. حال باید $UA^{-۱}$ را مساوی ماتریس مفروض B قرار داد تا رابطه $LB = I$ تشکیل شود. یعنی:

$$\begin{bmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & ۱ & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & ۱ \end{bmatrix} \begin{bmatrix} b_{۱۱} & b_{۱۲} & b_{۱۳} & \cdots & b_{۱n} \\ b_{۲۱} & b_{۲۲} & b_{۲۳} & \cdots & b_{۲n} \\ b_{۳۱} & b_{۳۲} & b_{۳۳} & \cdots & b_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n۱} & b_{n۲} & b_{n۳} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ ۰ & ۱ & ۰ & \cdots & ۰ \\ ۰ & ۰ & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & ۱ \end{bmatrix}$$

به طور مشابه از ضرب دو ماتریس L و B، تعداد $n^۲$ معادله به دست می‌آید که به ترتیب می‌توان مقادیر b_{ij} ها را به دست آورد. سپس از رابطه $UA^{-۱} = B$، مقادیر $a_{ij}^{(-۱)}$ به دست خواهد آمد.

مثال ۲-۱۵) روش پیدا کردن وارون ماتریس $A = \begin{bmatrix} ۵ & ۴ & ۳ \\ ۱۰ & ۱۲ & ۹ \\ ۱۵ & ۲۰ & ۱۸ \end{bmatrix}$ با تجزیه دولیتل به صورت زیر است.

ابتدا ماتریس فوق باید به صورت روبرو تجزیه شود: $\begin{bmatrix} ۵ & ۴ & ۳ \\ ۱۰ & ۱۲ & ۹ \\ ۱۵ & ۲۰ & ۱۸ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ l_{۲۱} & ۱ & ۰ \\ l_{۳۱} & l_{۳۲} & ۱ \end{bmatrix} \begin{bmatrix} u_{۱۱} & u_{۱۲} & u_{۱۳} \\ ۰ & u_{۲۲} & u_{۲۳} \\ ۰ & ۰ & u_{۳۳} \end{bmatrix}$

با توجه به تجزیه مذکور، تعداد ۹ معادله به وجود می‌آید (R_iها معرف سطرهای L و C_iها معرف ستون‌های U هستند):

$$R_۱ \times C_۱ \to u_{۱۱} = ۵ \quad , \quad R_۱ \times C_۲ \to u_{۱۲} = ۴ \quad , \quad R_۱ \times C_۳ \to u_{۱۳} = ۳$$

$$R_۲ \times C_۱ \to l_{۲۱}u_{۱۱} = a_{۲۱} \Rightarrow l_{۲۱} = \frac{a_{۲۱}}{u_{۱۱}} = \frac{۱۰}{۵} = ۲$$

$$R_۲ \times C_۲ \to l_{۲۱}u_{۱۲} + u_{۲۲} = a_{۲۲} \Rightarrow u_{۲۲} = a_{۲۲} - l_{۲۱}u_{۱۲} = ۱۲ - ۸ = ۴$$

$$R_۲ \times C_۳ \to l_{۲۱}u_{۱۳} + u_{۲۳} = a_{۲۳} \Rightarrow u_{۲۳} = a_{۲۳} - l_{۲۱}u_{۱۳} = ۹ - ۶ = ۳$$

$$R_۳ \times C_۱ \to l_{۳۱}u_{۱۱} = a_{۳۱} \Rightarrow l_{۳۱} = \frac{a_{۳۱}}{u_{۱۱}} = \frac{۱۵}{۵} = ۳$$

$$R_۳ \times C_۲ \to l_{۳۱}u_{۱۲} + l_{۳۲}u_{۲۲} = a_{۳۲} \Rightarrow l_{۳۲} = \frac{a_{۳۲} - l_{۳۱}u_{۱۲}}{u_{۲۲}} = \frac{۲۰ - ۱۲}{۴} = ۲$$

$$R_۳ \times C_۳ \to l_{۳۱}u_{۱۳} + l_{۳۲}u_{۲۳} + u_{۳۳} = a_{۳۳} \Rightarrow u_{۳۳} = a_{۳۳} - l_{۳۱}u_{۱۳} - l_{۳۲}u_{۲۳} = ۱۸ - ۱۵ = ۳$$

پس ماتریس به فرم زیر تجزیه شد:

$$\begin{bmatrix} ۵ & ۴ & ۳ \\ ۱۰ & ۱۲ & ۹ \\ ۱۵ & ۲۰ & ۱۸ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۱ & ۰ \\ ۳ & ۲ & ۱ \end{bmatrix} \begin{bmatrix} ۵ & ۴ & ۳ \\ ۰ & ۴ & ۳ \\ ۰ & ۰ & ۳ \end{bmatrix}$$

حال با جایگذاری $A = LU$ در رابطه $AA^{-۱} = I$، رابطه $LUA^{-۱} = I$ به دست می‌آید:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۱ & ۰ \\ ۳ & ۲ & ۱ \end{bmatrix} \begin{bmatrix} ۵ & ۴ & ۳ \\ ۰ & ۴ & ۳ \\ ۰ & ۰ & ۳ \end{bmatrix} \begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

حال باید $UA^{-۱} = B$ قرار داد تا به رابطه $LB = I$ رسید:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۱ & ۰ \\ ۳ & ۲ & ۱ \end{bmatrix} \begin{bmatrix} b_{۱۱} & b_{۱۲} & b_{۱۳} \\ b_{۲۱} & b_{۲۲} & b_{۲۳} \\ b_{۳۱} & b_{۳۲} & b_{۳۳} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

مجدداً ۹ معادله به دست می‌آید (R_iها معرف سطرهای L و C_iها معرف ستون‌های B هستند):

$$R_۱ \times C_۱ \to b_{۱۱} = ۱ \quad , \quad R_۱ \times C_۲ \to b_{۱۲} = ۰ \quad , \quad R_۱ \times C_۳ \to b_{۱۳} = ۰$$

$$R_۲ \times C_۱ \to ۲b_{۱۱} + b_{۲۱} = ۰ \Rightarrow b_{۲۱} = -۲b_{۱۱} = -۲$$

$$R_۲ \times C_۲ \to ۲b_{۱۲} + b_{۲۲} = ۱ \Rightarrow b_{۲۲} = ۱ - ۲b_{۱۲} = ۱ - ۰ = ۱$$

$$R_۲ \times C_۳ \to ۲b_{۱۳} + b_{۲۳} = ۰ \Rightarrow b_{۲۳} = -۲b_{۱۳} = ۰$$

$$R_{۳} \times C_{۱} \to ۳b_{۱۱} + ۲b_{۲۱} + b_{۳۱} = ۰ \Rightarrow b_{۳۱} = -۲b_{۲۱} - ۳b_{۱۱} = ۴ - ۳ = ۱$$

$$R_{۳} \times C_{۲} \to ۳b_{۱۲} + ۲b_{۲۲} + b_{۳۲} = ۰ \Rightarrow b_{۳۲} = -۲b_{۲۲} - ۳b_{۱۲} = -۲ - ۰ = -۲$$

$$R_{۳} \times C_{۳} \to ۳b_{۱۳} + ۲b_{۲۳} + b_{۳۳} = ۱ \Rightarrow b_{۳۳} = ۱ - ۲b_{۲۳} - ۳b_{۱۳} = ۱ - ۰ - ۰ = ۱$$

حال مقادیر b_i باید در ماتریس B جایگذاری شوند:

$$\Rightarrow B = \begin{bmatrix} ۱ & ۰ & ۰ \\ -۲ & ۱ & ۰ \\ ۱ & -۲ & ۱ \end{bmatrix}$$

پس از به دست آوردن B طبق رابطه $UA^{-۱} = B$ می‌توان ماتریس وارون را به دست آورد:

$$\begin{bmatrix} ۵ & ۴ & ۳ \\ ۰ & ۴ & ۳ \\ ۰ & ۰ & ۳ \end{bmatrix} \begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ -۲ & ۱ & ۰ \\ ۱ & -۲ & ۱ \end{bmatrix}$$

به طور مشابه پس از حل ۹ معادله، به مقادیر $a_{ij}^{(-۱)}$ رسیده و ماتریس وارون به دست می‌آید:

$$R_{۳} \times C_{۳} \to a_{۳۳}^{(-۱)} = \frac{۱}{۳} \quad , \quad R_{۳} \times C_{۲} \to a_{۳۲}^{(-۱)} = -\frac{۲}{۳} \quad , \quad R_{۳} \times C_{۱} \to a_{۳۱}^{(-۱)} = \frac{۱}{۳}$$

$$R_{۲} \times C_{۳} \to ۴a_{۲۳}^{(-۱)} + ۳a_{۳۳}^{(-۱)} = ۰ \Rightarrow a_{۲۳}^{(-۱)} = \frac{-۳a_{۳۳}^{(-۱)}}{۴} = -\frac{۱}{۴}$$

$$R_{۲} \times C_{۲} \to ۴a_{۲۲}^{(-۱)} + ۳a_{۳۲}^{(-۱)} = ۱ \Rightarrow a_{۲۲}^{(-۱)} = \frac{۱ - ۳a_{۳۲}^{(-۱)}}{۴} = \frac{۱ + ۲}{۴} = \frac{۳}{۴}$$

$$R_{۲} \times C_{۱} \to ۴a_{۲۱}^{(-۱)} + ۳a_{۳۱}^{(-۱)} = -۲ \Rightarrow a_{۲۱}^{(-۱)} = \frac{-۲ - ۳a_{۳۱}^{(-۱)}}{۴} = -\frac{۳}{۴}$$

$$R_{۱} \times C_{۳} \to ۵a_{۱۳}^{(-۱)} + ۴a_{۲۳}^{(-۱)} + ۳a_{۳۳}^{(-۱)} = ۰ \Rightarrow a_{۱۳}^{(-۱)} = \frac{-۳a_{۳۳}^{(-۱)} - ۴a_{۲۳}^{(-۱)}}{۵} = ۰$$

$$R_{۱} \times C_{۲} \to a_{۱۲}^{(-۱)} = -\frac{۱}{۵} \quad , \quad R_{۱} \times C_{۱} \to a_{۱۱}^{(-۱)} = \frac{۳}{۵}$$

پس از جایگذاری مقادیر $a_{ij}^{(-۱)}$ در ماتریس $A^{-۱}$ مسئله تمام شده است:

$$\Rightarrow A^{-۱} = \begin{bmatrix} ۳/۵ & -۱/۵ & ۰ \\ -۳/۴ & ۳/۴ & -۱/۴ \\ ۱/۳ & -۲/۳ & ۱/۳ \end{bmatrix}$$

۲-۳-۲ تجزیه کروت

این تجزیه نیز همانند تجزیه دولیتل می‌باشد، فقط با این تفاوت که در تجزیه دولیتل، قطر ماتریس L برابر با ۱ در نظر گرفته می‌شود. ولی در تجزیه کروت، قطر ماتریس U برابر ۱ در نظر گرفته خواهد شد. پس ماتریس $A_{n\times n}$ و ماتریس‌های پایین مثلثی L و بالا مثلثی U به صورت زیر مفروض است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} ۱ & u_{۱۲} & u_{۱۳} & \cdots & u_{۱n} \\ ۰ & ۱ & u_{۲۳} & \cdots & u_{۲n} \\ ۰ & ۰ & ۱ & \cdots & u_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & ۱ \end{bmatrix}$$

از ضرب دو ماتریس L و U، $n^{۲}$ معادله به دست می‌آید.

$$l_{۱۱} = a_{۱۱}\,, \qquad l_{۲۱} = a_{۲۱}\,, \cdots, \qquad l_{n۱} = a_{n۱} \tag{۲-۱۰}$$

$$l_{۱۱}u_{۱۲} = a_{۱۲}\,, \qquad l_{۲۱}u_{۱۲} + l_{۲۲} = a_{۲۲}\,, \cdots, \qquad l_{n۱}u_{۱۲} + l_{n۲} = a_{n۲} \tag{۲-۱۱}$$

$$\vdots \qquad\qquad \vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$l_{۱۱}u_{۱n} = a_{۱n}\,, l_{۲۱}u_{۱n} + l_{۲۲}u_{۲n} = a_{۲n}\,, \cdots, l_{n۱}u_{۱n} + \cdots + l_{nn} = a_{nn} \tag{۲-۱۲}$$

همانند دفعه قبل، با توجه به معادلات (۲-۱۰) به سادگی می‌توان مقادیر $l_{۱۱}, l_{۲۱}, l_{۳۱}, \cdots, l_{n۱}$ را به دست آورد. پس از آن با جایگذاری مقادیر مذکور در معادلات (۲-۱۱) می‌توان به ترتیب مقادیر $u_{۱۲}, l_{۲۲}, l_{۳۲}, \cdots, l_{n۲}$ را به دست آورد. به همین ترتیب، پس از انجام همه معادلات تا معادلات (۲-۱۲)، همه مقادیر مجهول به دست می‌آیند و ماتریس A به ماتریس‌های L و U تجزیه خواهد شد.

در ادامه با جایگذاری فرم LU به جای ماتریس A در رابطه $AA^{-۱} = I$، به صورت $LUA^{-۱} = I$ بازنویسی می‌شود. حال باید $UA^{-۱}$ را مساوی ماتریس مفروض B قرار داد تا رابطه $LB = I$ تشکیل شود. یعنی:

$$\begin{bmatrix} l_{۱۱} & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} b_{۱۱} & b_{۱۲} & b_{۱۳} & \cdots & b_{۱n} \\ b_{۲۱} & b_{۲۲} & b_{۲۳} & \cdots & b_{۲n} \\ b_{۳۱} & b_{۳۲} & b_{۳۳} & \cdots & b_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n۱} & b_{n۲} & b_{n۳} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ ۰ & ۱ & ۰ & \cdots & ۰ \\ ۰ & ۰ & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & ۱ \end{bmatrix}$$

با توجه به اینکه درایه‌های روی قطر اصلی ماتریس L در این روش می‌توانند مقادیر مختلفی کسب کنند و صرفاً برابر با ۱ نیستند، به نظر می‌رسد که برای محاسبه کار دشوارتری به نسبت تجزیه دولیتل در پیش روی فرد است؛ اما در مرحله بعد قطر اصلی ماتریس U برابر ۱ خواهد بود و عملاً دو روش تجزیه دولیتل و کروت تفاوتی با هم ندارند. در مرحله بعد به طور مشابه از ضرب دو ماتریس L و B، تعداد $n^{۲}$ معادله به دست می‌آید که به ترتیب می‌توان مقادیر b_{ij} ها را به دست آورد. سپس از رابطه $UA^{-۱} = B$، مقادیر $a_{ij}^{(-۱)}$ به دست خواهد آمد.

مثال ۲-۱۶) روش پیدا کردن وارون ماتریس $\begin{bmatrix} ۲ & ۶ & ۴ \\ ۳ & ۱۴ & ۲۱ \\ ۴ & ۱۸ & ۲۷ \end{bmatrix}$ با کمک تجزیه کروت به صورت زیر است.

ابتدا ماتریس فوق باید تجزیه شود:

$$\begin{bmatrix} ۲ & ۶ & ۴ \\ ۳ & ۱۴ & ۲۱ \\ ۴ & ۱۸ & ۲۷ \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} \end{bmatrix} \begin{bmatrix} ۱ & u_{۱۲} & u_{۱۳} \\ ۰ & ۱ & u_{۲۳} \\ ۰ & ۰ & ۱ \end{bmatrix}$$

با توجه به تجزیه بالا، تعداد ۹ معادله به وجود می‌آید (R_iها معرف سطرهای L و C_iها معرف ستون‌های U هستند):

$$R_۱ \times C_۱ \to l_{۱۱} = ۲ \ \ , \ \ R_۲ \times C_۱ \to l_{۲۱} = ۳ \ \ , \ \ R_۳ \times C_۱ \to l_{۳۱} = ۴$$

$$R_۱ \times C_۲ \to l_{۱۱}u_{۱۲} = a_{۱۲} \Rightarrow u_{۱۲} = \frac{a_{۱۲}}{l_{۱۱}} = \frac{۶}{۲} = ۳$$

$$R_۲ \times C_۲ \to l_{۲۱}u_{۱۲} + l_{۲۲} = a_{۲۲} \Rightarrow l_{۲۲} = a_{۲۲} - l_{۲۱}u_{۱۲} = ۱۴ - ۹ = ۵$$

$$R_۳ \times C_۲ \to l_{۳۱}u_{۱۲} + l_{۳۲} = a_{۳۲} \Rightarrow l_{۳۲} = a_{۳۲} - l_{۳۱}u_{۱۲} = ۱۸ - ۱۲ = ۶$$

$$R_۱ \times C_۳ \to l_{۱۱}u_{۱۳} = a_{۱۳} \Rightarrow u_{۱۳} = \frac{a_{۱۳}}{l_{۱۱}} = \frac{۴}{۲} = ۲$$

$$R_۲ \times C_۳ \to l_{۲۱}u_{۱۳} + l_{۲۲}u_{۲۳} = a_{۲۳} \Rightarrow u_{۲۳} = \frac{a_{۲۳} - l_{۲۱}u_{۱۳}}{l_{۲۲}} = \frac{۲۱ - ۶}{۵} = ۳$$

$$R_۳ \times C_۳ \to l_{۳۱}u_{۱۳} + l_{۳۲}u_{۲۳} + l_{۳۳} = a_{۳۳} \Rightarrow l_{۳۳} = a_{۳۳} - l_{۳۱}u_{۱۳} - l_{۳۲}u_{۲۳} = ۲۷ - ۲۶ = ۱$$

پس ماتریس به فرم زیر تجزیه شد:

$$\begin{bmatrix} ۲ & ۶ & ۴ \\ ۳ & ۱۴ & ۲۱ \\ ۴ & ۱۸ & ۲۷ \end{bmatrix} = \begin{bmatrix} ۲ & ۰ & ۰ \\ ۳ & ۵ & ۰ \\ ۴ & ۶ & ۱ \end{bmatrix} \begin{bmatrix} ۱ & ۳ & ۲ \\ ۰ & ۱ & ۳ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

حال با جایگذاری $A = LU$ در رابطه $AA^{-۱} = I$، رابطه $LUA^{-۱} = I$ به دست می‌آید:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۳ & ۵ & ۰ \\ ۴ & ۶ & ۱ \end{bmatrix} \begin{bmatrix} ۱ & ۳ & ۲ \\ ۰ & ۱ & ۳ \\ ۰ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

حال باید $UA^{-۱} = B$ قرار داد تا به رابطه $LB = I$ رسید:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۳ & ۵ & ۰ \\ ۴ & ۶ & ۱ \end{bmatrix} \begin{bmatrix} b_{۱۱} & b_{۱۲} & b_{۱۳} \\ b_{۲۱} & b_{۲۲} & b_{۲۳} \\ b_{۳۱} & b_{۳۲} & b_{۳۳} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

مجدداً ۹ معادله به دست می‌آید (R_iها معرف سطرهای L و C_iها معرف ستون‌های B هستند):

$$R_۱ \times C_۱ \to b_{۱۱} = \frac{۱}{۲} \quad , \quad R_۱ \times C_۲ \to b_{۱۲} = ۰ \quad , \quad R_۱ \times C_۳ \to b_{۱۳} = ۰$$

$$R_۲ \times C_۱ \to ۳b_{۱۱} + ۵b_{۲۱} = ۰ \Rightarrow b_{۲۱} = \frac{-۳b_{۱۱}}{۵} = -\frac{۳}{۱۰}$$

$$R_۲ \times C_۲ \to ۳b_{۱۲} + ۵b_{۲۲} = ۱ \Rightarrow b_{۲۲} = \frac{۱ - ۲b_{۱۲}}{۵} = \frac{۱ - ۰}{۵} = \frac{۱}{۵}$$

$$R_۲ \times C_۳ \to ۳b_{۱۳} + ۵b_{۲۳} = ۰ \Rightarrow b_{۲۳} = \frac{-۳b_{۱۳}}{۵} = ۰$$

$$R_۳ \times C_۱ \to ۴b_{۱۱} + ۶b_{۲۱} + b_{۳۱} = ۰ \Rightarrow b_{۳۱} = -۶b_{۲۱} - ۴b_{۱۱} = \frac{۹}{۵} - ۲ = -\frac{۱}{۵}$$

$$R_۳ \times C_۲ \to ۴b_{۱۲} + ۶b_{۲۲} + b_{۳۲} = ۰ \Rightarrow b_{۳۲} = -۶b_{۲۲} - ۴b_{۱۲} = -\frac{۶}{۵} - ۰ = -\frac{۶}{۵}$$

$$R_۳ \times C_۳ \to ۴b_{۱۳} + ۶b_{۲۳} + b_{۳۳} = ۱ \Rightarrow b_{۳۳} = ۱ - ۶b_{۲۳} - ۴b_{۱۳} = ۱ - ۰ - ۰ = ۱$$

حال مقادیر b_i باید در ماتریس B جایگذاری شوند:

$$\Rightarrow B = \begin{bmatrix} \frac{۱}{۲} & ۰ & ۰ \\ -\frac{۳}{۱۰} & \frac{۱}{۵} & ۰ \\ -\frac{۱}{۵} & -\frac{۶}{۵} & ۱ \end{bmatrix}$$

پس از به دست آوردن B طبق رابطه $UA^{-۱} = B$ می‌توان ماتریس وارون را به دست آورد:

$$\begin{bmatrix} ۱ & ۳ & ۲ \\ ۰ & ۱ & ۳ \\ ۰ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix} = \begin{bmatrix} \frac{۱}{۲} & ۰ & ۰ \\ -\frac{۳}{۱۰} & \frac{۱}{۵} & ۰ \\ -\frac{۱}{۵} & -\frac{۶}{۵} & ۱ \end{bmatrix}$$

به طور مشابه پس از حل ۹ معادله، به مقادیر $a_{ij}^{(-۱)}$ رسیده و ماتریس وارون به دست می‌آید (R_iها معرف سطرهای U و C_iها معرف ستون‌های $A^{-۱}$ هستند:

$$R_۳ \times C_۳ \to a_{۳۳}^{(-۱)} = ۱ \quad , \quad R_۳ \times C_۲ \to a_{۳۲}^{(-۱)} = -\frac{۶}{۵} \quad , \quad R_۳ \times C_۱ \to a_{۳۱}^{(-۱)} = -\frac{۱}{۵}$$

$$R_۲ \times C_۳ \to a_{۲۳}^{(-۱)} + ۳a_{۳۳}^{(-۱)} = ۰ \Rightarrow a_{۲۳}^{(-۱)} = -۳a_{۳۳}^{(-۱)} = -۳$$

$$R_۲ \times C_۲ \to a_{۲۲}^{(-۱)} = \frac{۱۹}{۵} \quad , \quad R_۲ \times C_۱ \to a_{۲۱}^{(-۱)} = \frac{۳}{۱۰}$$

$$R_۱ \times C_۳ \to a_{۱۳}^{(-۱)} + ۳a_{۲۳}^{(-۱)} + ۲a_{۳۳}^{(-۱)} = ۰ \Rightarrow a_{۱۳}^{(-۱)} = -۲a_{۳۳}^{(-۱)} - ۳a_{۲۳}^{(-۱)} = -۲ + ۹ = ۷$$

$$R_۱ \times C_۲ \to a_{۱۲}^{(-۱)} = -۹ \quad , \quad R_۱ \times C_۱ \to a_{۱۱}^{(-۱)} = ۰$$

سپس مقادیر $a_{ij}^{(-۱)}$ باید در ماتریس $A^{-۱}$ جایگذاری شوند:

$$\Rightarrow A^{-۱} = \begin{bmatrix} ۰ & -۹ & ۷ \\ \frac{۳}{۱۰} & \frac{۱۹}{۵} & -۳ \\ -\frac{۱}{۵} & -\frac{۶}{۵} & ۱ \end{bmatrix}$$

۲-۳-۳ تجزیه چولسکی

در تجزیه چولسکی، قطر ماتریس‌های L و U با هم برابر گرفته می‌شوند. می‌توان درایه‌های قطر اصلی را به صورت u_{ii} و یا l_{ii} در نظر گرفت که هر دو حالت صحیح می‌باشد. در ادامه درایه‌های قطر اصلی به صورت l_{ii} در نظر گرفته شدند. یعنی ماتریس‌های A، L و U به صورت زیر خواهد بود:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{۱۱} & u_{۱۲} & u_{۱۳} & \cdots & u_{۱n} \\ ۰ & l_{۲۲} & u_{۲۳} & \cdots & u_{۲n} \\ ۰ & ۰ & l_{۳۳} & \cdots & u_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & l_{nn} \end{bmatrix}$$

با ضرب دو ماتریس پایین مثلثی L و بالا مثلثی U، $n^۲$ معادله به دست می‌آید که همگی آن‌ها را می‌توان به صورت‌های زیر نوشت:

$$\begin{array}{llll} l_{۱۱}^{۲} = a_{۱۱}\,, & l_{۱۱}u_{۱۲} = a_{۱۲}\,, \cdots, & l_{۱۱}u_{۱n} = a_{۱n} & (۱۳\text{-}۲) \\ l_{۱۱}l_{۲۱} = a_{۲۱}\,, & l_{۲۱}u_{۱۲} + l_{۲۲}^{۲} = a_{۲۲}\,, \cdots, & l_{۲۱}u_{۱n} + l_{۲۲}u_{۲n} = a_{۲n} & (۱۴\text{-}۲) \\ \vdots & \vdots \qquad \vdots & \vdots & \vdots \\ \multicolumn{3}{l}{l_{۱۱}l_{n۱} = a_{n۱}\,, l_{n۱}u_{۱۲} + l_{۲۲}l_{n۲} = a_{n۲}\,, \cdots, l_{n۱}u_{۱n} + \cdots + l_{nn}^{۲} = a_{nn}} & (۱۵\text{-}۲) \end{array}$$

همانند دفعات قبل، با توجه به معادلات (۲-۱۳) به ترتیب می‌توان مقادیر $l_{۱۱}, u_{۱۲}, u_{۱۳}, \cdots, u_{۱n}$ را به دست آورد. پس از آن با جایگذاری مقادیر مذکور در معادلات (۲-۱۴) می‌توان به ترتیب مقادیر $l_{۲۱}, l_{۲۲}, u_{۲۳}, \cdots, u_{۲n}$ را به دست آورد. به همین ترتیب، پس از انجام همه معادلات تا معادلات (۲-۱۵)، همه مقادیر مجهول به دست می‌آیند و ماتریس A به ماتریس‌های L و U تجزیه خواهد شد.

در ادامه همانند تجزیه‌های گذشته فرم LU باید به جای ماتریس A در رابطه $AA^{-1} = I$ جایگذاری شود. درنتیجه رابطه ذکر شده به صورت $LUA^{-1} = I$ بازنویسی می‌شود. حال باید UA^{-1} را مساوی ماتریس مفروض B قرار داد تا رابطه $LB = I$ تشکیل شود. یعنی:

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

به طور مشابه از ضرب دو ماتریس L و B، تعداد n^2 معادله به دست می‌آید که به ترتیب می‌توان مقادیر b_{ij} ها را به دست آورد. سپس از رابطه $UA^{-1} = B$ به صورت زیر، مقادیر $a_{ij}^{(-1)}$ به دست خواهد آمد:

$$\begin{bmatrix} l_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & l_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & l_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} a_{11}^{(-1)} & a_{12}^{(-1)} & \cdots & a_{1n}^{(-1)} \\ a_{21}^{(-1)} & a_{22}^{(-1)} & \cdots & a_{2n}^{(-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(-1)} & a_{n2}^{(-1)} & \cdots & a_{nn}^{(-1)} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1n} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2n} \\ b_{31} & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix}$$

مثال ۲-۱۷) روش یافتن وارون ماتریس $\begin{bmatrix} 4 & 4 & 4 \\ 2 & 3 & 3 \\ 6 & 9 & 18 \end{bmatrix}$ با روش تجزیه چولسکی به صورت زیر است.

ابتدا ماتریس فوق باید تجزیه شود:

$$\begin{bmatrix} 4 & 4 & 4 \\ 2 & 3 & 3 \\ 6 & 9 & 18 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & u_{12} & u_{13} \\ 0 & l_{22} & u_{23} \\ 0 & 0 & l_{33} \end{bmatrix}$$

با توجه به تجزیه بالا، تعداد ۹ معادله به وجود می‌آید (R_iها معرف سطرهای ماتریس L و C_iها معرف ستون‌های U هستند):

$$R_1 \times C_1 \rightarrow l_{11}^2 = a_{11} \Rightarrow l_{11} = \sqrt{a_{11}} = \sqrt{4} = \pm 2$$

برای درایه l_{11}، دو مقدار ۲+ و ۲- وجود دارد. در صورت انتخاب هر یک از مقادیر ماتریس‌های متفاوتی برای تجزیه ایجاد می‌شوند. پس اگر یکی از مقادیر هم انتخاب شود، کافی می‌باشد. در اینجا مقدار $l_{11} = 2$ انتخاب شده است.

$$R_1 \times C_2 \rightarrow l_{11}u_{12} = a_{12} \Rightarrow u_{12} = \frac{a_{12}}{l_{11}} = \frac{4}{2} = 2$$

$$R_1 \times C_3 \rightarrow l_{11}u_{13} = a_{13} \Rightarrow u_{13} = \frac{a_{13}}{l_{11}} = \frac{4}{2} = 2$$

$$R_۲ \times C_۱ \rightarrow l_{۱۱}l_{۲۱} = a_{۲۱} \Rightarrow l_{۲۱} = \frac{a_{۲۱}}{l_{۱۱}} = \frac{۲}{۲} = ۱$$

$$R_۲ \times C_۲ \rightarrow l_{۲۱}u_{۱۲} + l_{۲۲}^۲ = a_{۲۲} \Rightarrow l_{۲۲} = \sqrt{a_{۲۲} - l_{۲۱}u_{۱۲}} = \sqrt{۳ - ۲} = \pm ۱$$

همانند $l_{۱۱}$ دو مقدار برای $l_{۲۲}$ به دست می‌آید که در اینجا مقدار مثبت آن یعنی $l_{۲۲} = ۱$ در نظر گرفته شده است.

$$R_۲ \times C_۳ \rightarrow l_{۲۱}u_{۱۳} + l_{۲۲}u_{۲۳} = a_{۲۳} \Rightarrow u_{۲۳} = \frac{a_{۲۳} - l_{۲۱}u_{۱۳}}{l_{۲۲}} = \frac{۳ - ۲}{۱} = ۱$$

$$R_۳ \times C_۱ \rightarrow l_{۱۱}l_{۳۱} = a_{۳۱} \Rightarrow l_{۳۱} = \frac{a_{۳۱}}{l_{۱۱}} = \frac{۶}{۲} = ۳$$

$$R_۳ \times C_۲ \rightarrow l_{۳۱}u_{۱۲} + l_{۲۲}l_{۳۲} = a_{۳۲} \Rightarrow l_{۳۲} = \frac{a_{۳۲} - l_{۳۱}u_{۱۲}}{l_{۲۲}} = \frac{۹ - ۶}{۱} = ۳$$

$$R_۳ \times C_۳ \rightarrow l_{۳۱}u_{۱۳} + l_{۳۲}u_{۲۳} + l_{۳۳}^۲ = a_{۳۳} \Rightarrow l_{۳۳} = \sqrt{a_{۳۳} - l_{۳۱}u_{۱۳} - l_{۳۲}u_{۲۳}} = \sqrt{۹} = \pm ۳$$

همانند $l_{۱۱}$ و $l_{۲۲}$ دو مقدار برای $l_{۳۳}$ به دست می‌آید که در اینجا مقدار مثبت آن یعنی $l_{۳۳} = ۱$ در نظر گرفته شده است. پس ماتریس به فرم زیر تجزیه شد:

$$\begin{bmatrix} ۴ & ۴ & ۴ \\ ۲ & ۳ & ۳ \\ ۶ & ۹ & ۱۸ \end{bmatrix} = \begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۱ & ۰ \\ ۳ & ۳ & ۳ \end{bmatrix} \begin{bmatrix} ۲ & ۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۳ \end{bmatrix}$$

برای درایه‌های $l_{۱۱}$، $l_{۲۲}$ و $l_{۳۳}$ دو مقدار ذکر شد. به ازای دیگر مقادیر، ماتریس‌های جدیدی به دست می‌آیند:

$$(l_{۱۱} = +۲, l_{۲۲} = +۱, l_{۳۳} = +۳) \quad , \quad (l_{۱۱} = +۲, l_{۲۲} = +۱, l_{۳۳} = -۳)$$

$$\begin{bmatrix} +۲ & ۰ & ۰ \\ ۱ & +۱ & ۰ \\ ۳ & ۳ & +۳ \end{bmatrix} \begin{bmatrix} +۲ & ۲ & ۲ \\ ۰ & +۱ & ۱ \\ ۰ & ۰ & +۳ \end{bmatrix} \; , \; \begin{bmatrix} +۲ & ۰ & ۰ \\ ۱ & +۱ & ۰ \\ ۳ & ۳ & -۳ \end{bmatrix} \begin{bmatrix} +۲ & ۲ & ۲ \\ ۰ & +۱ & ۱ \\ ۰ & ۰ & -۳ \end{bmatrix}$$

$$(l_{۱۱} = +۲, l_{۲۲} = -۱, l_{۳۳} = +۳) \quad , \quad (l_{۱۱} = +۲, l_{۲۲} = -۱, l_{۳۳} = -۳)$$

$$\begin{bmatrix} +۲ & ۰ & ۰ \\ ۱ & -۱ & ۰ \\ ۳ & -۳ & +۳ \end{bmatrix} \begin{bmatrix} +۲ & ۲ & ۲ \\ ۰ & -۱ & -۱ \\ ۰ & ۰ & +۳ \end{bmatrix} \; , \; \begin{bmatrix} +۲ & ۰ & ۰ \\ ۱ & -۱ & ۰ \\ ۳ & -۳ & -۳ \end{bmatrix} \begin{bmatrix} +۲ & ۲ & ۲ \\ ۰ & -۱ & -۱ \\ ۰ & ۰ & -۳ \end{bmatrix}$$

$$(l_{۱۱} = -۲, l_{۲۲} = +۱, l_{۳۳} = +۳) \quad , \quad (l_{۱۱} = -۲, l_{۲۲} = +۱, l_{۳۳} = -۳)$$

$$\begin{bmatrix} -۲ & ۰ & ۰ \\ -۱ & +۱ & ۰ \\ -۳ & ۳ & +۳ \end{bmatrix} \begin{bmatrix} -۲ & -۲ & -۲ \\ ۰ & +۱ & ۱ \\ ۰ & ۰ & +۳ \end{bmatrix} \; , \; \begin{bmatrix} -۲ & ۰ & ۰ \\ -۱ & +۱ & ۰ \\ -۳ & ۳ & -۳ \end{bmatrix} \begin{bmatrix} -۲ & -۲ & -۲ \\ ۰ & +۱ & ۱ \\ ۰ & ۰ & -۳ \end{bmatrix}$$

$$(l_{۱۱} = -۲, l_{۲۲} = -۱, l_{۳۳} = +۳) \quad , \quad (l_{۱۱} = -۲, l_{۲۲} = -۱, l_{۳۳} = -۳)$$

$$\begin{bmatrix} -۲ & ۰ & ۰ \\ -۱ & -۱ & ۰ \\ -۳ & -۳ & +۳ \end{bmatrix}\begin{bmatrix} -۲ & -۲ & -۲ \\ ۰ & -۱ & -۱ \\ ۰ & ۰ & +۳ \end{bmatrix} , \begin{bmatrix} -۲ & ۰ & ۰ \\ -۱ & -۱ & ۰ \\ -۳ & -۳ & -۳ \end{bmatrix}\begin{bmatrix} -۲ & -۲ & -۲ \\ ۰ & -۱ & -۱ \\ ۰ & ۰ & -۳ \end{bmatrix}$$

۸ فرم مختلف با کمک تجزیه چولسکی به دست می‌آید که در ادامه همان فرم اول یعنی ماتریس‌های L و U با درایه‌های مثبت روی قطر اصلی در نظر گرفته می‌شوند. حال با جایگذاری $A = LU$ در رابطه $AA^{-۱} = I$، رابطه $LUA^{-۱} = I$ به دست می‌آید:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۱ & ۰ \\ ۳ & ۳ & ۳ \end{bmatrix}\begin{bmatrix} ۲ & ۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۳ \end{bmatrix}\begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

حال باید $UA^{-۱} = B$ قرار داد تا به رابطه $LB = I$ رسید:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۱ & ۰ \\ ۳ & ۳ & ۳ \end{bmatrix}\begin{bmatrix} b_{۱۱} & b_{۱۲} & b_{۱۳} \\ b_{۲۱} & b_{۲۲} & b_{۲۳} \\ b_{۳۱} & b_{۳۲} & b_{۳۳} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

مجدداً ۹ معادله به دست می‌آید (R_iها معرف سطرهای L و C_iها معرف ستون‌های B هستند):

$$R_۱ \times C_۱ \to ۲b_{۱۱} = ۱ \Rightarrow b_{۱۱} = \frac{۱}{۲}$$

$$R_۱ \times C_۲ \to ۲b_{۱۲} = ۰ \Rightarrow b_{۱۲} = ۰$$

$$R_۱ \times C_۳ \to ۲b_{۱۳} = ۰ \Rightarrow b_{۱۳} = ۰$$

$$R_۲ \times C_۱ \to b_{۱۱} + b_{۲۱} = ۰ \Rightarrow b_{۲۱} = -b_{۱۱} = -\frac{۱}{۲}$$

$$R_۲ \times C_۲ \to b_{۱۲} + b_{۲۲} = ۱ \Rightarrow b_{۲۲} = ۱ - b_{۱۲} = ۱ - ۰ = ۱$$

$$R_۲ \times C_۳ \to b_{۱۳} + b_{۲۳} = ۰ \Rightarrow b_{۲۳} = -b_{۱۳} = ۰$$

$$R_۳ \times C_۱ \to ۳b_{۱۱} + ۳b_{۲۱} + ۳b_{۳۱} = ۰ \Rightarrow b_{۳۱} = \frac{-۳b_{۲۱} - ۳b_{۱۱}}{۳} = \frac{\frac{۳}{۲} - \frac{۳}{۲}}{۳} = ۰$$

$$R_۳ \times C_۲ \to ۳b_{۱۲} + ۳b_{۲۲} + ۳b_{۳۲} = ۰ \Rightarrow b_{۳۲} = \frac{-۳b_{۲۲} - ۳b_{۱۲}}{۳} = \frac{-۳ - ۰}{۳} = -۱$$

$$R_۳ \times C_۳ \to ۳b_{۱۳} + ۳b_{۲۳} + ۳b_{۳۳} = ۱ \Rightarrow b_{۳۳} = \frac{۱ - ۳b_{۲۳} - ۳b_{۱۳}}{۳} = \frac{۱ - ۰ - ۰}{۳} = \frac{۱}{۳}$$

پس از جایگذاری مقادیر b_{ij} ماتریس B به صورت زیر خواهد بود:

$$\Rightarrow B=\begin{bmatrix} \frac{۱}{۲} & ۰ & ۰ \\ -\frac{۱}{۲} & ۱ & ۰ \\ ۰ & -۱ & \frac{۱}{۳} \end{bmatrix}$$

پس از به دست آوردن B طبق رابطه $UA^{-۱}=B$ می‌توان ماتریس وارون را به دست آورد:

$$\begin{bmatrix} ۲ & ۲ & ۲ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۳ \end{bmatrix}\begin{bmatrix} a_{۱۱}^{(-۱)} & a_{۱۲}^{(-۱)} & a_{۱۳}^{(-۱)} \\ a_{۲۱}^{(-۱)} & a_{۲۲}^{(-۱)} & a_{۲۳}^{(-۱)} \\ a_{۳۱}^{(-۱)} & a_{۳۲}^{(-۱)} & a_{۳۳}^{(-۱)} \end{bmatrix}=\begin{bmatrix} \frac{۱}{۲} & ۰ & ۰ \\ -\frac{۱}{۲} & ۱ & ۰ \\ ۰ & -۱ & \frac{۱}{۳} \end{bmatrix}$$

به طور مشابه پس از حل ۹ معادله، به مقادیر $a_{ij}^{(-۱)}$ رسیده و ماتریس وارون به دست می‌آید:

$$R_۳\times C_۳\rightarrow a_{۳۳}^{(-۱)}=\frac{۱}{۹} \quad , \quad R_۳\times C_۲\rightarrow a_{۳۲}^{(-۱)}=-\frac{۱}{۳} \quad , \quad R_۳\times C_۱\rightarrow a_{۳۱}^{(-۱)}=۰$$

$$R_۲\times C_۳\rightarrow a_{۲۳}^{(-۱)}+a_{۳۳}^{(-۱)}=۰\Rightarrow a_{۲۳}^{(-۱)}=-a_{۳۳}^{(-۱)}=-\frac{۱}{۹}$$

$$R_۲\times C_۲\rightarrow a_{۲۲}^{(-۱)}=\frac{۴}{۳} \quad , \quad R_۲\times C_۱\rightarrow a_{۲۱}^{(-۱)}=-\frac{۱}{۲}$$

$$R_۱\times C_۳\rightarrow ۲a_{۱۳}^{(-۱)}+۲a_{۲۳}^{(-۱)}+۲a_{۳۳}^{(-۱)}=۰\Rightarrow a_{۱۳}^{(-۱)}=\frac{-۲a_{۳۳}^{(-۱)}-۲a_{۲۳}^{(-۱)}}{۲}=۰$$

$$R_۱\times C_۲\rightarrow a_{۱۲}^{(-۱)}=-۱ \quad , \quad R_۱\times C_۱\rightarrow a_{۱۱}^{(-۱)}=\frac{۳}{۴}$$

و در نهایت ماتریس وارون با جایگذاری مقادیر فوق به دست خواهد آمد:

$$\Rightarrow A^{-۱}=\begin{bmatrix} ۳/۴ & -۱ & ۰ \\ -۱/۲ & ۴/۳ & -۱/۹ \\ ۰ & -۱/۳ & ۱/۹ \end{bmatrix}$$

۲-۴ کاربرد ماتریس وارون در علوم داده

از معکوس ماتریس، برای محاسبه بردار پارامتر توسط معادله نرمال[1]در معادله‌ای خطی استفاده می‌شود. به طور مثال فرض کنید که مجموعه داده زیر موجود باشد:

جدول (۲-۱) پیش‌بینی پیروزی تیم‌ها در بازی بیسبال

Team	League	Year	RS	RA	W	OBP	SLG	BA	G	OOBP	OSLG
ARI	NL	۲۰۱۲	۷۳٤	٦۸۸	۸۱	۰/۳۲۸	۰/٤۱۸	۰/۲٥۹	۱	۰/۳۱۷	۰/٤۱٥
ATL	NL	۲۰۱۲	۷۰۰	٦۰۰	۹٤	۰/۳۲	۰/۳۸۹	۰/۲٤۷	۱	۰/۳۰٦	۰/۳۷۸
BAL	AL	۲۰۱۲	۷۱۲	۷۰٥	۹۳	۰/۳۱۱	۰/٤۱۷	۰/۲٤۷	۱	۰/۳۱٥	۰/٤۰۳
BOS	AL	۲۰۱۲	۷۳٤	۸۰٦	٦۹	۰/۳۱٥	۰/٤۱٥	۰/۲٦	۱	۰/۳۳۱	۰/٤۲۸
CHC	NL	۲۰۱۲	٦۱۳	۷٥۹	٦۱	۰/۳۰۲	۰/۳۷۸	۰/۲٤	۱	۰/۳۳٥	۰/٤۲٤
CHW	AL	۲۰۱۲	۷٤۸	٦۷٦	۸٥	۰/۳۱۸	۰/٤۲۲	۰/۲٥٥	۱	۰/۳۱۹	۰/٤۰٥
CIN	NL	۲۰۱۲	٦٦۹	٥۸۸	۹۷	۰/۳۱٥	۰/٤۱۱	۰/۲٥۱	۱	۰/۳۰٥	۰/۳۹
CLE	AL	۲۰۱۲	٦٦۷	۸٤٥	٦۸	۰/۳۲٤	۰/۳۸۱	۰/۲٥۱	۱	۰/۳۳٦	۰/٤۳
COL	NL	۲۰۱۲	۷٥۸	۸۹۰	٦٤	۰/۳۳	۰/٤۳٦	۰/۲۷٤	۱	۰/۳٥۷	۰/٤۷
DET	AL	۲۰۱۲	۷۲٦	٦۷۰	۸۸	۰/۳۳٥	۰/٤۲۲	۰/۲٦۸	۱	۰/۳۱٤	۰/٤۰۲
HOU	NL	۲۰۱۲	٥۸۳	۷۹٤	٥٥	۰/۳۰۲	۰/۳۷۱	۰/۲۳٦	۱	۰/۳۳۷	۰/٤۲۷
KCR	AL	۲۰۱۲	٦۷٦	۷٤٦	۷۲	۰/۳۱۷	۰/٤	۰/۲٦٥	۱	۰/۳۳۹	۰/٤۲۳
LAA	AL	۲۰۱۲	۷٦۷	٦۹۹	۸۹	۰/۳۳۲	۰/٤۳۳	۰/۲۷٤	۱	۰/۳۱	۰/٤۰۳
LAD	NL	۲۰۱۲	٦۳۷	٥۹۷	۸٦	۰/۳۱۷	۰/۳۷٤	۰/۲٥۲	۱	۰/۳۱	۰/۳٦٤

جدول (۲-۱) متغیرهای تیم‌های مختلف بیسبال برای پیش‌بینی پیروزی یا شکست آن‌ها را نشان می‌دهد. برای تبدیل مسئله به یک مسئله‌ی رگرسیونی[2]، فعلاً فرض کنید که متغیر OOBP، متغیر هدف می‌باشد. برای حل این مسئله با رگرسیون خطی، باید بردار پارامتر را پیدا کرده و از روش معادله نرمال استفاده کرد. روش معادله نرمال نیز از ماتریس‌ها استفاده می‌کند.

اگر متغیرهای مستقل لازم برای حل مسئله را به صورت ماتریس $X_{m\times n}$، بردار هدف را به صورت بردار $Y_{m\times 1}$ و همچنین بردار پارامتر را به صورت $\theta_{n\times 1}$ در نظر بگیرید، ماتریس X، بردار Y و بردار θ به صورت زیر نمایش داده می‌شوند:

---

[1] روشی شبیه به الگوریتم گرادیان کاهشی در یادگیری ماشین که برای مینیمم کردن تابع از آن استفاده می‌شود. تفاوت معادله نرمال با گرادیان کاهشی در بدون حلقه تکرار بودن آن است.

[2] Regression

$$X=\begin{bmatrix} ۷۳۴ & ۶۸۸ & ۸۱ & ۰/۳۲۸ & ۰/۴۱۸ & ۰/۲۵۹ \\ ۷۰۰ & ۶۰۰ & ۹۴ & ۰/۳۲ & ۰/۳۸۹ & ۰/۲۴۷ \\ ۷۱۲ & ۷۰۵ & ۹۳ & ۰/۳۱۱ & ۰/۴۱۷ & ۰/۲۴۷ \\ ۷۳۴ & ۸۰۶ & ۶۹ & ۰/۳۱۵ & ۰/۴۱۵ & ۰/۲۶ \\ ۶۱۳ & ۷۵۹ & ۶۱ & ۰/۳۰۲ & ۰/۳۷۸ & ۰/۲۴ \\ ۷۴۸ & ۶۷۶ & ۸۵ & ۰/۳۱۸ & ۰/۴۲۲ & ۰/۲۵۵ \\ ۶۶۹ & ۵۸۸ & ۹۷ & ۰/۳۱۵ & ۰/۴۱۱ & ۰/۲۵۱ \\ ۶۶۷ & ۸۴۵ & ۶۸ & ۰/۳۲۴ & ۰/۳۸۱ & ۰/۲۵۱ \\ ۷۵۸ & ۸۹۰ & ۶۴ & ۰/۳۳ & ۰/۴۳۶ & ۰/۲۷۴ \\ ۷۲۶ & ۶۷۰ & ۸۸ & ۰/۳۳۵ & ۰/۴۲۲ & ۰/۲۶۸ \\ ۵۸۳ & ۷۹۴ & ۵۵ & ۰/۳۰۲ & ۰/۳۷۱ & ۰/۲۳۶ \\ ۶۷۶ & ۷۴۶ & ۷۲ & ۰/۳۱۷ & ۰/۴ & ۰/۲۶۵ \\ ۷۶۷ & ۶۹۹ & ۸۹ & ۰/۳۳۲ & ۰/۴۳۳ & ۰/۲۷۴ \\ ۶۳۷ & ۵۹۷ & ۸۶ & ۰/۳۱۷ & ۰/۳۷۴ & ۰/۲۵۲ \end{bmatrix} \;,\; Y=\begin{bmatrix} ۰/۳۱۷ \\ ۰/۳۰۶ \\ ۰/۳۱۵ \\ ۰/۳۳۱ \\ ۰/۳۳۵ \\ ۰/۳۱۹ \\ ۰/۳۰۵ \\ ۰/۳۳۶ \\ ۰/۳۵۷ \\ ۰/۳۱۴ \\ ۰/۳۳۷ \\ ۰/۳۳۹ \\ ۰/۳۱ \\ ۰/۳۱ \end{bmatrix} \;,\; \theta=\begin{bmatrix} \theta_۱ \\ \theta_۲ \\ \theta_۳ \\ \theta_۴ \\ \theta_۵ \\ \theta_۶ \end{bmatrix}$$

در نهایت برای پیدا کردن بردار پارامتر نهایی یعنی $\theta_{n\times ۱}$، با فرض اینکه تابع اولیه توسط θ و X به صورت پارامتری باشد، باید معکوس X^TX را به دست آورد. زیرا بردار پارامتر از رابطه زیر به دست می‌آید:

$$\theta=[(X^TX)^{-۱}X^T]Y \qquad (۱۶\text{-}۲)$$

تصور کنید که مجبور بودید این دستگاه معادلات را بدون استفاده از جبر خطی حل کنید. باید توجه داشت که مجموعه داده بررسی‌شده در این مثال، تنها یک درصد از مجموعه داده اصلی را شامل می‌شد. حال تصور کنید که مجبور به پیدا کردن بردار پارامتر، بدون استفاده از جبر خطی بودید. حل مسئله بدون استفاده از جبر خطی، نیاز به زمان و توان بسیار زیادی دارد و گاهی اوقات این کار غیرممکن می‌شود.

فصل سوم

دستگاه معادلات خطی

فصل سوم : دستگاه معادلات خطی

مطالب این فصل از [۵] و [۶] آورده شده است. حل بسیاری از مسائل در مهندسی و علوم، اکثراً منجر به حل یک دستگاه معادلات خطی می‌شود. به عنوان مثال در تحلیل‌های آماری ممکن است که به یک دستگاه معادلات خطی رسید که با حل آن دستگاه، در واقع جواب مسائل استخراج می‌شود. حوزه علوم داده نیز از این قاعده مستثنی نیست. در فصل اول به صورت مختصر دستگاه معادلات خطی معرفی شد؛ در این فصل به روش‌های حل یک دستگاه معادلات خطی و کاربردهای آن در علوم داده پرداخته خواهد شد.

به طور کلی می‌توان روش‌های حل دستگاه معادلات خطی $AX = b$ را به صورت زیر دسته‌بندی کرد:

حل دستگاه $AX = b$:
- روش مستقیم کرامر
- روش‌های تولید دستگاه مثلثی معادل
  - روش حذفی گاوس
  - روش حذفی گاوس − جردن
- روش‌های تجزیه مثلثی LU
  - تجزیه دولیتل
  - تجزیه کروت
  - تجزیه چولسکی
- روش‌های تکراری
  - روش تصفیه تکراری
  - روش ژاکوبی
  - روش گاوس − سایدل
  - روش فوق (تحت) تخفیف

۳-۱ روش کرامر[۱]

فرض کنید دستگاه $AX = b$ دستگاهی وارون‌پذیر باشد، در این‌صورت می‌توان جواب این دستگاه را به صورت زیر محاسبه کرد:

$$x_i = \frac{\det(A_i)}{\det(A)} \; , i = ۱,۲,\cdots,n \qquad (۱\text{-}۳)$$

که در آن A_i همان ماتریس A می‌باشد که به جای ستون i ام آن، بردار b قرار گرفته است. پس برای به دست آوردن جواب دستگاه معادلات خطی به روش کرامر باید تعداد $(n+۱)$ دترمینان به دست آورد.

---
۱ Cramer

مثال ۳-۱) جواب دستگاه $\begin{cases} x_۱ + ۲x_۲ + ۳x_۳ = ۱ \\ ۲x_۱ - x_۲ + ۳x_۳ = -۱ \\ x_۱ - ۳x_۲ + ۲x_۳ = ۲ \end{cases}$ با روش کرامر به صورت زیر به دست خواهد آمد.

ابتدا باید ماتریس‌های $A_{۳\times۳}$، $X_{۳\times۱}$ و $b_{۳\times۱}$ را تشکیل داد:

$$A = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۲ & -۱ & ۳ \\ ۲ & -۳ & ۲ \end{bmatrix} \ , \ X = \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} \ , \ b = \begin{bmatrix} ۱ \\ -۱ \\ ۲ \end{bmatrix}$$

برای محاسبه x_i ها باید دترمینان ماتریس‌های $A_۱$، $A_۲$، $A_۳$ و A را محاسبه کرد:

$$\det(A) = -۱۰$$

$$A_۱ = \begin{bmatrix} ۱ & ۲ & ۳ \\ -۱ & -۱ & ۳ \\ ۲ & -۳ & ۲ \end{bmatrix} \Rightarrow \det(A_۱) = ۳۸ \Rightarrow x_۱ = \frac{\det(A_۱)}{\det(A)} = \frac{۳۸}{-۱۰} = -\frac{۱۹}{۵}$$

$$A_۲ = \begin{bmatrix} ۱ & ۱ & ۳ \\ ۲ & -۱ & ۳ \\ ۲ & ۲ & ۲ \end{bmatrix} \Rightarrow \det(A_۲) = ۶ \Rightarrow x_۲ = \frac{\det(A_۲)}{\det(A)} = \frac{۶}{-۱۰} = -\frac{۳}{۵}$$

$$A_۳ = \begin{bmatrix} ۱ & ۲ & ۱ \\ ۲ & -۱ & -۱ \\ ۲ & -۳ & ۲ \end{bmatrix} \Rightarrow \det(A_۳) = -۲۰ \Rightarrow x_۳ = \frac{\det(A_۳)}{\det(A)} = \frac{-۲۰}{-۱۰} = ۲$$

در نهایت جواب دستگاه معادلات خطی با جایگذاری مقادیر x_i در بردار X دستگاه معادلات خطی حل خواهد شد:

$$X = \begin{bmatrix} -\frac{۱۹}{۵} \\ -\frac{۳}{۵} \\ ۲ \end{bmatrix}$$

۳-۲ تولید دستگاه مثلثی معادل

تعریف ۳-۲) اگر A و B ماتریس‌هایی مربعی باشند و همچنین $AX = b$ و $BX = d$، دو دستگاه مذکور را معادل گویند هرگاه مجموعه جواب هر دو دستگاه یکسان باشند.

یعنی برای حل یک دستگاه می‌توان دستگاه معادل با آن را حل کرد، در حالی که جوابی از آن کم یا به آن اضافه نشده باشد. این خاصیت دستگاه‌های معادل باعث می‌شود که بتوان یک دستگاه با ماتریس ضرایب پُر (یعنی ماتریسی که تعداد عناصر صفر کمی دارد) را به ماتریسی مثلثی تبدیل کرد.

در روش‌های تولید دستگاه مثلثی معادل، ابتدا به کمک روش حذفی گاوس[1]، ماتریس ضرایب به ماتریسی مثلثی تبدیل خواهد شد و پس از آن به کمک جایگذاری پیشرو و یا پسرو معادلات خطی حل می‌شوند. همچنین به کمک روش حذفی گاوس-جردن[2]، ماتریس ضرایب به ماتریس سطری -پلکانی تحویل یافته تبدیل خواهد شد.

۳-۲-۱ روش حذفی گاوس

برای حل دستگاه $AX = b$، ابتدا ماتریس افزوده $[A|b]$ را باید تشکیل داد، سپس باید با استفاده از عملیات سطری مقدماتی بر روی ماتریس افزوده، ماتریس ضرایب دستگاه (یعنی A) را بالا مثلثی یا پایین مثلثی کرد و پس از آن با جایگذاری پسرو یا پیشرو، جواب‌های دستگاه را به دست آورد.

گام اول (صفر کردن عناصر زیر $a_{۱۱}$): فرض کنید که $a_{۱۱} \neq ۰$، در این صورت این درایه درایه محوری و سطر شامل $a_{۱۱}$ سطر محوری خوانده می‌شود (اگر $a_{۱۱} = ۰$ آنگاه سطری از ماتریس افزوده که اولین درایه آن صفر نیست باید با این سطر جابه‌جا شود). حال ضربگرهای $m_{i۱}$ به صورت زیر تعریف می‌شوند:

$$m_{i۱} = \frac{a_{i۱}}{a_{۱۱}}\ , i = ۲,۳,\cdots,n \qquad (۲\text{-}۳)$$

حال $-m_{i۱}$ برابر سطر اول به سطر i ام که $i = ۲,۳,\cdots,n$ اضافه می‌شود. یا به عبارت دیگر:

$$-m_{i۱}R_۱ + R_i \to R_i$$

در این صورت ماتریس افزوده دستگاه مذکور به صورت زیر خواهد بود:

$$[A|b]^{(۱)} = \left[\begin{array}{cccc|c} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} & b_۱ \\ ۰ & a_{۲۲}^{(۱)} & \cdots & a_{۲n}^{(۱)} & b_۲^{(۱)} \\ ۰ & a_{۳۲}^{(۱)} & \cdots & a_{۳n}^{(۱)} & b_۳^{(۱)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ ۰ & a_{n۲}^{(۱)} & \cdots & a_{nn}^{(۱)} & b_n^{(۱)} \end{array}\right]$$

که در آن:

$$\begin{cases} a_{ij}^{(۱)} = a_{ij} - m_{i۱}a_{۱j}\ , i,j = ۲,\cdots,n \\ b_i^{(۱)} = b_i - m_{i۱}b_۱\ , i = ۲,\cdots,n \end{cases}$$

گام دوم (صفر کردن عناصر زیر $a_{۲۲}^{(۱)}$): فرض کنید $a_{۲۲}^{(۱)} \neq ۰$، این عنصر دومین درایه محوری خوانده می‌شود (مجدداً اگر $a_{۲۲}^{(۱)} = ۰$، آنگاه سطر دوم را با یکی از سطرهای بعدی که در آن دومین درایه ناصفر است، جابه‌جا می‌کنیم). حال ضربگرهای $m_{i۲}$ به صورت زیر تعریف می‌شوند:

[1] Gaussian Elimination Method
[2] Gauss-Jordan Elimination Method

$$m_{i۲} = \frac{a_{i۲}^{(۱)}}{a_{۲۲}^{(۱)}} \ , i = ۳,۴,\cdots,n \qquad (۳\text{-}۳)$$

حال $-m_{i۲}$ برابر سطر دوم به سطر i ام که $i = ۳,۴,\cdots,n$ اضافه می‌شود. یا به عبارت دیگر:

$$-m_{i۲}R_۲ + R_i \to R_i$$

در این صورت ماتریس افزوده دستگاه مذکور به صورت زیر خواهد بود:

$$[A|b]^{(۲)} = \left[\begin{array}{ccccc|c} a_{۱۱} & a_{۱۲} & a_{۱۳} & \cdots & a_{۱n} & b_۱ \\ ۰ & a_{۲۲}^{(۱)} & a_{۲۳}^{(۱)} & \cdots & a_{۲n}^{(۱)} & b_۲^{(۱)} \\ ۰ & ۰ & a_{۳۳}^{(۲)} & \cdots & a_{۳n}^{(۲)} & b_۳^{(۲)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ ۰ & ۰ & a_{n۳}^{(۲)} & \cdots & a_{nn}^{(۲)} & b_n^{(۲)} \end{array}\right]$$

که در آن:

$$\begin{cases} a_{ij}^{(۲)} = a_{ij}^{(۱)} - m_{i۲}a_{۲j}^{(۱)} \ , i,j = ۳,\cdots,n \\ b_i^{(۲)} = b_i^{(۱)} - m_{i۲}b_۲^{(۱)} \ , i = ۳,\cdots,n \end{cases}$$

گام $n - ۱$ ام (گام نهایی): با ادامه روند فوق، در گام $(n - ۱)$ ام ماتریس افزوده به صورت زیر خواهد بود:

$$[A|b]^{(n-۱)} = \left[\begin{array}{ccccc|c} a_{۱۱} & a_{۱۲} & a_{۱۳} & \cdots & a_{۱n} & b_۱ \\ ۰ & a_{۲۲}^{(۱)} & a_{۲۳}^{(۱)} & \cdots & a_{۲n}^{(۱)} & b_۲^{(۱)} \\ ۰ & ۰ & a_{۳۳}^{(۲)} & \cdots & a_{۳n}^{(۲)} & b_۳^{(۲)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ ۰ & ۰ & ۰ & \cdots & a_{nn}^{(n-۱)} & b_n^{(n-۱)} \end{array}\right]$$

که در آن:

$$\begin{cases} a_{nn}^{(n-۱)} = a_{nn}^{(n-۲)} - m_{n,n-۱}a_{n-۱,n}^{(n-۲)} \\ b_n^{(n-۱)} = b_n^{(n-۲)} - m_{n,n-۱}b_{n-۱}^{(n-۲)} \end{cases}$$

همان‌طور که مشاهده می‌شود، ماتریس حاصل یک ماتریس بالا مثلثی است و دستگاه متناظر با آن نیز یک دستگاه بالا مثلثی است که می‌توان آن را با روش جایگذاری پسرو (از آخر به اول) به صورت زیر حل کرد:

$$x_n = \frac{b_n}{a_{nn}} \qquad (۴\_۳)$$

$$x_i = \frac{۱}{a_{ii}}\left(b_i - \sum_{j=i+۱}^{n} a_{ij}x_j\right) \quad , \quad i = n-۱, n-۲, \cdots, ۱ \qquad (۵\_۳)$$

تذکر ۳-۳) روش حذفی گاوس را می‌توان به طریقی مشابه، جهت تبدیل دستگاه مذکور به یک دستگاه پایین مثلثی نیز به کار برد و در نهایت دستگاه حاصی را با روش جایگذاری پیشرو (از اول به آخر) به صورت زیر حل کرد:

$$x_۱ = \frac{b_۱}{a_{۱۱}} \qquad (۶\_۳)$$

$$x_i = \frac{۱}{a_{ii}}\left(b_i - \sum_{j=۱}^{i} a_{ij}x_j\right) \quad , \quad i = ۲,۳, \cdots, n \qquad (۷\_۳)$$

تذکر ۳-۴) اگر در یکی از گام‌های روش حذفی گاوس، نتوان درایه محوری ناصفر پیدا کرد، آنگاه دترمینان ماتریس ضرایب دستگاه مذکور صفر بوده و دستگاه یا جواب ندارید یا بی‌شمار جواب دارد.

مثال ۳-۵) دستگاه معادلات خطی $\begin{cases} x_۱ + ۲x_۲ + x_۳ + ۴x_۴ = ۱۳ \\ ۲x_۱ + ۴x_۳ + ۳x_۴ = ۲۸ \\ ۴x_۱ + ۲x_۲ + ۲x_۳ + x_۴ = ۲۰ \\ -۳x_۱ + x_۲ + ۳x_۳ + ۲x_۴ = ۶ \end{cases}$ با روش حذفی گاوس به صورت زیر حل خواهد شد.

گام اول) ابتدا ماتریس افزوده متناظر با این دستگاه باید تشکیل شود و سپس با کمک اعمال سطری مقدماتی باید ماتریس را بالا مثلثی کرد (درایه‌ای که زیر آن خط کشیده شده، عنصر محوری است):

$$[A|b] = \left[\begin{array}{cccc|c} \underline{۱} & ۲ & ۱ & ۴ & ۱۳ \\ ۲ & ۰ & ۴ & ۳ & ۲۸ \\ ۴ & ۲ & ۲ & ۱ & ۲۰ \\ -۳ & ۱ & ۳ & ۲ & ۶ \end{array}\right] \xrightarrow{-۲R_۱+R_۲\to R_۲} \left[\begin{array}{cccc|c} \underline{۱} & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & -۴ & ۲ & -۵ & ۲ \\ ۴ & ۲ & ۲ & ۱ & ۲۰ \\ -۳ & ۱ & ۳ & ۲ & ۶ \end{array}\right] \xrightarrow{-۴R_۱+R_۳\to R_۳}$$

$$\left[\begin{array}{cccc|c} \underline{۱} & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & -۴ & ۲ & -۵ & ۲ \\ ۰ & -۶ & -۲ & -۱۵ & -۳۲ \\ -۳ & ۱ & ۳ & ۲ & ۶ \end{array}\right] \xrightarrow{۳R_۱+R_۴\to R_۴} \left[\begin{array}{cccc|c} \underline{۱} & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & -۴ & ۲ & -۵ & ۲ \\ ۰ & -۶ & -۲ & -۱۵ & -۳۲ \\ ۰ & ۷ & ۶ & ۱۴ & ۴۵ \end{array}\right]$$

گام دوم) حال باید با در نظر گرفتن عنصر محوری جدید، به ادامه محاسبات پرداخت:

$$\left[\begin{array}{cccc|c} ۱ & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & \underline{-۴} & ۲ & -۵ & ۲ \\ ۰ & -۶ & -۲ & -۱۵ & -۳۲ \\ ۰ & ۷ & ۶ & ۱۴ & ۴۵ \end{array}\right] \xrightarrow{-\frac{۳}{۲}R_۲+R_۳\to R_۳} \left[\begin{array}{cccc|c} ۱ & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & \underline{-۴} & ۲ & -۵ & ۲ \\ ۰ & ۰ & -۵ & -\frac{۱۵}{۲} & -۳۵ \\ ۰ & ۷ & ۶ & ۱۴ & ۴۵ \end{array}\right] \xrightarrow{\frac{۷}{۴}R_۲+R_۴\to R_۴}$$

$$\left[\begin{array}{cccc|c} ۱ & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & \underline{-۴} & ۲ & -۵ & ۲ \\ ۰ & ۰ & -۵ & -\frac{۱۵}{۲} & -۳۵ \\ ۰ & ۰ & \frac{۱۹}{۲} & \frac{۲۱}{۴} & \frac{۹۷}{۲} \end{array}\right]$$

گام سوم) این بار عنصر محوری ۵- است و ضربگر نظیر آن نیز $m_{۴۳} = -{}^{۱۹}\!/_{۱۰}$ می‌باشد:

$$\left[\begin{array}{cccc|c} ۱ & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & -۴ & ۲ & -۵ & ۲ \\ ۰ & ۰ & \underline{-۵} & -\frac{۱۵}{۲} & -۳۵ \\ ۰ & ۰ & \frac{۱۹}{۲} & \frac{۲۱}{۴} & \frac{۹۷}{۲} \end{array}\right] \xrightarrow{\frac{۱۹}{۱۰}R_۳+R_۴\to R_۴} \left[\begin{array}{cccc|c} ۱ & ۲ & ۱ & ۴ & ۱۳ \\ ۰ & -۴ & ۲ & -۵ & ۲ \\ ۰ & ۰ & -۵ & -\frac{۱۵}{۲} & -۳۵ \\ ۰ & ۰ & ۰ & -۹ & -۱۸ \end{array}\right] \Rightarrow$$

$$x_۴ = ۲, x_۳ = ۴, x_۲ = -۱, x_۱ = ۳$$

مثال ۳-۶) جواب دستگاه $\begin{cases} x_۱ + x_۲ + x_۳ + x_۴ = ۷ \\ x_۱ + x_۲ + ۲x_۴ = ۸ \\ ۲x_۱ + ۲x_۲ + ۳x_۳ = ۱۰ \\ -x_۱ - x_۲ - ۲x_۳ + ۲x_۴ = ۰ \end{cases}$ با روش حذفی گاوس به صورت زیر استخراج خواهد شد.

گام اول) ابتدا ماتریس افزوده متناظر با دستگاه تشکیل خواهد شد و سپس با اعمال سطری مقدماتی ماتریس باید بالا مثلثی شود (درایه‌ای که زیر آن خط کشیده شده، عنصر محوری است):

$$[A|b] = \left[\begin{array}{cccc|c} \underline{۱} & ۱ & ۱ & ۱ & ۷ \\ ۱ & ۱ & ۰ & ۲ & ۸ \\ ۲ & ۲ & ۳ & ۰ & ۱۰ \\ -۱ & -۱ & -۲ & ۲ & ۰ \end{array}\right] \xrightarrow{-R_۱+R_۲\to R_۲} \left[\begin{array}{cccc|c} \underline{۱} & ۱ & ۱ & ۱ & ۷ \\ ۰ & ۰ & -۱ & ۱ & ۱ \\ ۲ & ۲ & ۳ & ۰ & ۱۰ \\ -۱ & -۱ & -۲ & ۲ & ۰ \end{array}\right] \xrightarrow{-۲R_۱+R_۳\to R_۳}$$

$$\left[\begin{array}{cccc|c} \underline{1} & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -2 & -4 \\ -1 & -1 & -2 & 2 & 0 \end{array}\right] \xrightarrow{R_1+R_4\to R_4} \left[\begin{array}{cccc|c} \underline{1} & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & -1 & 3 & 7 \end{array}\right]$$

گام دوم) همان‌طور که مشاهده می‌شود $a_{22}^{(1)} = 0$ و در سطرهای زیرین آن هیچ عنصر ناصفری جهت جابه‌جایی با عنصر محوری وجود ندارد. بنابراین گام دوم تغییری در ماتریس ایجاد نمی‌کند. یعنی:

$$[A|b]^{(2)} = [A|b]^{(1)}$$

پس بدون تغییر به سراغ گام سوم رفته و از سطر سوم باید ادامه محاسبات را انجام داد.

گام سوم) اولین درایه سطر سوم ناصفر بوده، پس به عنوان عنصر محوری انتخاب خواهد شد و ضربگر نظیر آن نیز می‌شود $m_{43} = -1$. در ادامه با کمک اعمال سطری مقدماتی ماتریس مذکور به ماتریس مثلثی تبدیل خواهد شد:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & \underline{1} & -2 & -4 \\ 0 & 0 & -1 & 3 & 7 \end{array}\right] \xrightarrow{R_3+R_4\to R_4}$$

$$[A|b]^{(3)} = \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]$$

حال از طریق جایگذاری پسرو مقادیر x_i ها را باید محاسبه کرد:

$$x_4 = 3 \quad , \quad x_3 - 2x_4 = -4 \xRightarrow{x_4=3} x_3 = 2$$

$$R_2:\ -x_3 + x_4 = 1 \quad (\text{برقرار است})$$

$$R_1:\ x_1 + x_2 + x_3 + x_4 = 7 \xRightarrow{x_4=3, x_3=2} x_1 + x_2 = 2 \Rightarrow x_2 = 2 - x_1$$

پس در نهایت جواب این دستگاه به صورت زیر است:

$$X = \begin{bmatrix} t \\ 2-t \\ 2 \\ 3 \end{bmatrix}, \quad t \in \mathbb{R}$$

به ازای مقادیر مختلف t، جواب‌های مختلفی استخراج می‌شود. بنابراین دستگاه مذکور بی‌شمار جواب دارد.

مثال ۳-۷) دستگاه معادلات خطی
$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 7 \\ x_1 + x_2 + 2x_4 = 5 \\ 2x_1 + 2x_2 + 3x_3 = 10 \\ -x_1 - x_2 - 2x_3 + 2x_4 = 0 \end{cases}$$
با روش حذفی گاوس به صورت زیر حل خواهد شد.

گام اول) ماتریس افزوده تشکیل شده و a_{11} به عنوان عنصر محوری انتخاب شده است:

$$[A|b] = \left[\begin{array}{cccc|c} \underline{1} & 1 & 1 & 1 & 7 \\ 1 & 1 & 0 & 2 & 5 \\ 2 & 2 & 3 & 0 & 10 \\ -1 & -1 & -2 & 2 & 0 \end{array}\right]$$

حال درایه‌های زیر عنصر محوری با کمک اعمال سطری مقدماتی باید صفر شوند:

$$\left[\begin{array}{cccc|c} \underline{1} & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & -1 & 3 & 7 \end{array}\right]$$

گام دوم) چون که عنصر محوری جدید و درایه‌های زیر آن همگی صفر هستند، پس مرحله دوم با مرحله اول ماتریسی یکسان دارند.

گام سوم) عنصر محوری جدید a_{33} می‌باشد؛ پس:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & \underline{1} & -2 & -4 \\ 0 & 0 & -1 & 3 & 7 \end{array}\right] \xrightarrow{R_3+R_4\rightarrow R_4} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] = [A|b]^{(3)}$$

حال از طریق جایگذاری پسرو مقادیر x_i ها را باید محاسبه کرد:

$$x_4 = 3 \quad , \quad x_3 - 2x_4 = -4 \xRightarrow{x_4=3} x_3 = 2$$

$$R_2: \; -x_3 + x_4 = -2 \xRightarrow{x_4=3, x_3=2} -2 + 3 = -2$$

که به وضوح تناقض دارد. یعنی دستگاه مذکور فاقد جواب است.

۳-۲-۲ روش حذفی گاوس جردن

این روش همانند روش حذفی گاوس است؛ با این تفاوت که ماتریس ضرایب دستگاه $AX = b$ به یک ماتریس قطری تبدیل خواهد شد. به عبارت دیگر، در این روش در هر ستون از ماتریس ضرایب A، غیر از عنصر محوری، تمامی درایه‌های بالا و پایین هر ستون صفر می‌شود که در این حالت جواب دستگاه به صورت جایگذاری مستقیم حاصل می‌شود. برای تبدیل به ماتریس قطری، باید از ستون اول شروع کرده، ابتدا درایه‌های زیر عنصر محوری و سپس درایه‌های بالایی عنصر محوری صفر شوند. در واقع باید ستون به ستون پیش رفت.

مثال ۳-۸) جواب دستگاه معادلات خطی $\begin{cases} ۲x_۱ + ۳x_۲ + ۴x_۳ = ۷ \\ ۶x_۱ + ۵x_۲ + ۷x_۳ = ۱۴ \\ -۸x_۱ + ۱۲x_۲ - x_۳ = ۲۹ \end{cases}$ با روش گاوس-جردن به صورت زیر به دست خواهد آمد.

ابتدا ماتریس افزوده متناظر با دستگاه فوق باید نوشته شود:

$$[A|b] = \left[\begin{array}{ccc|c} \underline{۲} & ۳ & ۴ & ۷ \\ ۶ & ۵ & ۷ & ۱۴ \\ -۸ & ۱۲ & -۱ & ۲۹ \end{array}\right] \xrightarrow{-۳R_۱+R_۲\to R_۲} \left[\begin{array}{ccc|c} \underline{۲} & ۳ & ۴ & ۷ \\ ۰ & -۴ & -۵ & -۷ \\ -۸ & ۱۲ & -۱ & ۲۹ \end{array}\right] \xrightarrow{۴R_۱+R_۳\to R_۳}$$

$$[A|b]^{(۱)} = \left[\begin{array}{ccc|c} ۲ & ۳ & ۴ & ۷ \\ ۰ & \underline{-۴} & -۵ & -۷ \\ ۰ & ۲۴ & ۱۵ & ۵۷ \end{array}\right] \xrightarrow{۶R_۲+R_۳\to R_۳} \left[\begin{array}{ccc|c} ۲ & ۳ & ۴ & ۷ \\ ۰ & \underline{-۴} & -۵ & -۷ \\ ۰ & ۰ & -۱۵ & ۱۵ \end{array}\right] \xrightarrow{\frac{۳}{۴}R_۲+R_۱\to R_۱}$$

$$[A|b]^{(۲)} = \left[\begin{array}{ccc|c} ۲ & ۰ & \frac{۱}{۴} & \frac{۷}{۴} \\ ۰ & -۴ & -۵ & -۷ \\ ۰ & ۰ & \underline{-۱۵} & ۱۵ \end{array}\right] \xrightarrow{\frac{-۱}{۳}R_۳+R_۲\to R_۲} \left[\begin{array}{ccc|c} ۲ & ۰ & \frac{۱}{۴} & \frac{۷}{۴} \\ ۰ & -۴ & ۰ & ۱۲ \\ ۰ & ۰ & \underline{-۱۵} & ۱۵ \end{array}\right] \xrightarrow{\frac{۱}{۶۰}R_۳+R_۱\to R_۱}$$

$$[A|b]^{(۳)} = \left[\begin{array}{ccc|c} ۲ & ۰ & ۰ & ۲ \\ ۰ & -۴ & ۰ & ۱۲ \\ ۰ & ۰ & -۱۵ & ۱۵ \end{array}\right]$$

حال با جایگذاری مستقیم، مقادیر x_i به دست خواهد آمد:

$$x_۱ = \frac{۲}{۲} = ۱ \;\; , \;\; x_۲ = \frac{۱۲}{-۴} = -۳ \;\; , \;\; x_۳ = \frac{۱۵}{-۱۵} = -۱$$

تذکر ۳-۹) اگر در روش حذفی گاوس-جردن، ماتریس ضرایب به همانی تبدیل شود، سمت راست ماتریس افزوده در واقع جواب دستگاه است. در برخی مراجع روش حذفی گاوس-جردن برای دستگاه $AX = b$ به صورت زیر بیان می‌شود: $[A|b] \xrightarrow{\text{دنباله‌ای از اعمال سطری مقدماتی}} [I|A^{-۱}b]$

۳-۳ روش‌های تجزیه مثلثی

هدف این بخش بازنویسی (تجزیه) ماتریس ضرایب دستگاه $AX = b$ به حاصل‌ضرب ماتریس‌های پایین مثلثی L و بالا مثلثی U می‌باشد به طوری که $A = LU$. اگر این امکان وجود داشته باشد که ماتریس A به صورت گفته شده تجزیه شود، به آسانی می‌توان دستگاه $AX = b$ را حل کرد.

فرض کنید A به حاصل‌ضرب LU تجزیه شده باشد، در این صورت دستگاه $AX = b$ به صورت $LUX = b$ بازنویسی می‌شود. اگر UX مساوی با Y قرار داده شود، رابطه $LUX = b$ به صورت $LY = b$ تبدیل خواهد شد. چون L ماتریسی پایین مثلثی است، پس با استفاده از روش جایگذاری پیشرو می‌توان دستگاه $LY = b$ را حل کرد و Y را به دست آورد. سپس دستگاه $UX = Y$ را با توجه به بالامثلثی بودن U می‌توان به روش جایگذاری پسرو حل کرد و درنهایت X را به دست آورد.

تذکر ۳-۱۰) یکی از روش‌های تجزیه ماتریس A به حاصل ضرب LU، با استفاده از روش حذفی گاوس می‌باشد. اگر $U = \left(u_{ij}\right)$ و $L = \left(l_{ij}\right)$ آنگاه ماتریس‌ها به صورت زیر تعریف خواهند شد:

$$u_{ij} = \begin{cases} a_{ij}^{(i-۱)}, & i = ۱,۲,\cdots,j \\ ۰, & i = j+۱,\cdots,n \end{cases} \quad , \quad l_{ij} = \begin{cases} ۰, & i = ۱,۲,\cdots,j-۱ \\ ۱, & i = j \\ m_{ij}, & i = j+۱,\cdots,n \end{cases}$$

که در آن m_{ij} ها ضریبگرهای روش حذفی گاوس می‌باشند. همچنین ماتریس U در واقع همان ماتریس $A^{(n-۱)}$ می‌باشد (ماتریس آخرین مرحله روش حذفی گاوس).

مثال ۳-۱۱) دستگاه $\begin{cases} x_۱ + x_۲ + x_۳ = ۶ \\ ۹x_۱ + ۲x_۲ + x_۳ = ۱۶ \\ x_۱ + ۲x_۲ - ۵x_۳ = -۱۰ \end{cases}$ به روش تجزیه LU و با کمک روش حذفی گاوس به صورت زیر حل خواهد شد.

ابتدا ماتریس ضرایب دستگاه فوق با روش حذفی گاوس به ماتریس بالا مثلثی تبدیل خواهد شد:

$$\begin{bmatrix} ۱ & ۱ & ۱ \\ ۹ & ۲ & ۱ \\ ۱ & ۲ & -۵ \end{bmatrix} \xrightarrow{-۹R_۱+R_۲\to R_۲, m_{۲۱}=۹} \begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۱ & ۲ & -۵ \end{bmatrix} \xrightarrow{-R_۱+R_۳\to R_۳, m_{۳۱}=۱}$$

$$\begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۰ & ۱ & -۶ \end{bmatrix} \xrightarrow{\frac{۱}{۷}R_۲+R_۳\to R_۳, m_{۳۲}=\frac{-۱}{۷}} \begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۰ & ۰ & -\frac{۵۰}{۷} \end{bmatrix}$$

بنابراین $U = \begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۰ & ۰ & -\frac{۵۰}{۷} \end{bmatrix}$ و $L = \begin{bmatrix} ۱ & ۰ & ۰ \\ m_{۲۱} & ۱ & ۰ \\ m_{۳۱} & m_{۳۲} & ۱ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۹ & ۱ & ۰ \\ ۱ & -\frac{۱}{۷} & ۱ \end{bmatrix}$.

به وضوح رابطه $A = LU$ برقرار است. حال LU را جایگزین ماتریس A در رابطه $AX = b$ باید کرد:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۹ & ۱ & ۰ \\ ۱ & -\frac{۱}{۷} & ۱ \end{bmatrix} \begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۰ & ۰ & -\frac{۵۰}{۷} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۶ \\ ۱۶ \\ -۱۰ \end{bmatrix}$$

با در نظر گرفتن UX برابر با Y، رابطه $LY = b$ ایجاد خواهد شد:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۹ & ۱ & ۰ \\ ۱ & -\frac{۱}{۷} & ۱ \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \end{bmatrix} = \begin{bmatrix} ۶ \\ ۱۶ \\ -۱۰ \end{bmatrix}$$

با روش جایگذاری پیشرو، مقادیر y_i به دست خواهد آمد:

$$y_۱ = ۶$$

$$(۹ \times ۶) + y_۲ = ۱۶ \Rightarrow y_۲ = ۱۶ - ۵۴ = -۳۸$$

$$۶ - \frac{۱}{۷}(-۳۸) + y_۳ = -۱۰ \Rightarrow y_۳ = -۱۰ - \frac{۳۸}{۷} - ۶ = -\frac{۱۵۰}{۷}$$

پس از جایگذاری مقادیر فوق در بردار Y، باید دستگاه $UX = Y$ را حل کرد:

$$\begin{bmatrix} ۱ & ۱ & ۱ \\ ۰ & -۷ & -۸ \\ ۰ & ۰ & -\frac{۵۰}{۷} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۶ \\ -۳۸ \\ -\frac{۱۵۰}{۷} \end{bmatrix}$$

با جایگذاری پسرو می‌توان جواب نهایی دستگاه معادلات خطی را پیدا کرد:

$$-\frac{۵۰}{۷} x_۳ = -\frac{۱۵۰}{۷} \Rightarrow x_۳ = ۳$$

$$-۷x_۲ - (۸ \times ۳) = -۳۸ \Rightarrow x_۲ = \frac{-۳۸ + ۲۴}{-۷} = ۲$$

$$x_۱ + ۲ + ۳ = ۶ \Rightarrow x_۱ = ۶ - ۵ = ۱$$

در نهایت ماتریس جواب به صورت زیر خواهد بود:

$$X = \begin{bmatrix} ۳ \\ ۲ \\ ۱ \end{bmatrix}$$

یادآوری ۳-۱۲) اگر ماتریس A اکیداً قطر غالب یا معین مثبت اکید باشد و یا تمام زیرماتریس‌های اصلی و پیشروی ماتریس A معکوس‌پذیر باشند، آنگاه A دارای تجزیه LU است.

۳-۳-۱ تجزیه دولیتل

روش تجزیه ماتریس A به ماتریس‌های پایین مثلثی L و بالا مثلثی U که در آن درایه‌های قطر L برابر با ۱ باشند، در بخش (۳-۲-۱) گفته شد. در این بخش توضیحات مربوط به استفاده از این تجزیه و حل مثالی از این روش گردآوری شده است. پس از به دست آوردن تجزیه LU، با جایگذاری فرم LU به جای ماتریس A در $AX = b$، به صورت $LUX = b$ بازنویسی می‌شود. حال باید UX را مساوی بردار مفروض Y قرار داد تا رابطه $LY = b$ تشکیل شود. یعنی:

$$\begin{bmatrix} ۱ & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & ۱ & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & ۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & ۱ \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_۱ \\ b_۲ \\ b_۳ \\ \vdots \\ b_n \end{bmatrix}$$

حال از ضرب ماتریس L و بردار Y، تعداد $n^۲$ معادله به دست می‌آید که با روش جایگذاری پیشرو به ترتیب می‌توان مقادیر y_i ها را به دست آورد. سپس از رابطه $UX = Y$، مقادیر x_i به دست خواهد آمد.

مثال ۳-۱۳) با استفاده از تجزیه دولیتل دستگاه معادلات خطی $\begin{cases} ۲x_۱ + ۴x_۲ - ۶x_۳ = -۴ \\ x_۱ + ۵x_۲ + ۳x_۳ = ۱۰ \\ x_۱ + ۳x_۲ + ۲x_۳ = ۵ \end{cases}$ به صورت زیر حل خواهد شد.

ابتدا با کمک تجزیه دولیتل، ماتریس ضرایب به LU تجزیه می‌شود:

$$A = \begin{bmatrix} ۲ & ۴ & -۶ \\ ۱ & ۵ & ۳ \\ ۱ & ۳ & ۲ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ l_{۲۱} & ۱ & ۰ \\ l_{۳۱} & l_{۳۲} & ۱ \end{bmatrix} \begin{bmatrix} u_{۱۱} & u_{۱۲} & u_{۱۳} \\ ۰ & u_{۲۲} & u_{۲۳} \\ ۰ & ۰ & u_{۳۳} \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۱/۲ & ۱ & ۰ \\ ۱/۲ & ۱/۳ & ۱ \end{bmatrix} \begin{bmatrix} ۲ & ۴ & -۶ \\ ۰ & ۳ & ۶ \\ ۰ & ۰ & ۳ \end{bmatrix}$$

پس رابطه $AX = b$ به فرم $LUX = b$ تبدیل خواهد شد:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ \frac{۱}{۲} & ۱ & ۰ \\ \frac{۱}{۲} & \frac{۱}{۳} & ۱ \end{bmatrix} \begin{bmatrix} ۲ & ۴ & -۶ \\ ۰ & ۳ & ۶ \\ ۰ & ۰ & ۳ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۱۰ \\ ۵ \end{bmatrix}$$

با قرار دادن $UX = Y$، باید ابتدا دستگاه $LY = b$ را حل کرده و مقادیر y_i را به دست آورد:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ \frac{۱}{۲} & ۱ & ۰ \\ \frac{۱}{۲} & \frac{۱}{۳} & ۱ \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۱۰ \\ ۵ \end{bmatrix} \Rightarrow y_۱ = -۴ \,, y_۲ = ۱۲ \,, y_۳ = ۳$$

سپس باید رابطه $UX = Y$ را با جایگذاری پسرو حل کرد:

$$\begin{bmatrix} ۲ & ۴ & -۶ \\ ۰ & ۳ & ۶ \\ ۰ & ۰ & ۳ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۱۲ \\ ۳ \end{bmatrix}$$

$$۳x_۳ = ۳ \Rightarrow x_۳ = \frac{۳}{۳} = ۱$$

$$۳x_۲ + ۶ = ۱۲ \Rightarrow x_۲ = \frac{۱۲ - ۶}{۳} = \frac{۶}{۳} = ۲$$

$$۲x_۱ + (۴ \times ۲) - ۶ = -۴ \Rightarrow x_۱ = \frac{-۴ - ۲}{۲} = \frac{-۶}{۲} = -۳$$

در نهایت جواب دستگاه معادلات خطی مذکور به صورت زیر خواهد بود:

$$\begin{cases} x_۱ = ۱ \\ x_۲ = ۲ \\ x_۳ = -۳ \end{cases} \Rightarrow X = \begin{bmatrix} ۱ \\ ۲ \\ -۳ \end{bmatrix}$$

۳-۳-۲ تجزیه کروت

روش تجزیه ماتریس A به ماتریس‌های پایین مثلثی L و بالا مثلثی U که در آن درایه‌های قطر U برابر با ۱ باشند، در بخش (۲-۳-۲) گفته شد. در این بخش توضیحات مربوط به استفاده از این تجزیه و حل مثالی از این روش گردآوری شده است. به طور مشابه پس از به دست آوردن تجزیه LU، با جایگذاری فرم LU به جای ماتریس A در $AX = b$، به صورت $LUX = b$ بازنویسی می‌شود. حال باید UX را مساوی بردار مفروض Y قرار داد تا رابطه $LY = b$ تشکیل شود. یعنی:

$$\begin{bmatrix} l_{۱۱} & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_۱ \\ b_۲ \\ b_۳ \\ \vdots \\ b_n \end{bmatrix}$$

حال از ضرب ماتریس L و بردار Y، تعداد $n^۲$ معادله به دست می‌آید که با روش جایگذاری پیشرو به ترتیب می‌توان مقادیر y_i ها را به دست آورد. سپس از رابطه $UX = Y$ یعنی رابطه زیر، مقادیر x_i به دست خواهد آمد.

$$\begin{bmatrix} ۱ & u_{۱۲} & u_{۱۳} & \cdots & u_{۱n} \\ ۰ & ۱ & u_{۲۳} & \cdots & u_{۲n} \\ ۰ & ۰ & ۱ & \cdots & u_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & ۱ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \\ \vdots \\ y_n \end{bmatrix}$$

مثال ۳-۱۴) دستگاه معادلات خطی مثال (۳-۱۳) با تجزیه کروت به صورت زیر حل خواهد شد.

ابتدا با کمک تجزیه کروت، ماتریس ضرایب به LU تجزیه می‌شود:

$$A = \begin{bmatrix} ۲ & ۴ & -۶ \\ ۱ & ۵ & ۳ \\ ۱ & ۳ & ۲ \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} \end{bmatrix} \begin{bmatrix} ۱ & u_{۱۲} & u_{۱۳} \\ ۰ & ۱ & u_{۲۳} \\ ۰ & ۰ & ۱ \end{bmatrix} = \begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۳ & ۰ \\ ۱ & ۱ & ۳ \end{bmatrix} \begin{bmatrix} ۱ & ۲ & -۳ \\ ۰ & ۱ & ۲ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

پس رابطه $AX = b$ به فرم $LUX = b$ تبدیل خواهد شد:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۳ & ۰ \\ ۱ & ۱ & ۳ \end{bmatrix} \begin{bmatrix} ۱ & ۲ & -۳ \\ ۰ & ۱ & ۲ \\ ۰ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۱۰ \\ ۵ \end{bmatrix}$$

حال با تبدیل UX به بردار Y، رابطه زیر به دست می‌آید:

$$\begin{bmatrix} ۲ & ۰ & ۰ \\ ۱ & ۳ & ۰ \\ ۱ & ۱ & ۳ \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۱۰ \\ ۵ \end{bmatrix}$$

با کمک جایگذاری پیشرو می‌توان درایه‌های بردار Y را پیدا کرد:

$$۲y_۱ = -۴ \Rightarrow y_۱ = -\frac{۴}{۲} = -۲$$

$$-۲ + ۳y_۲ = ۱۰ \Rightarrow y_۲ = \frac{۱۰ + ۲}{۳} = ۴$$

$$-۲ + ۴ + ۳y_۳ = ۵ \Rightarrow y_۳ = \frac{۵ - ۴ + ۲}{۳} = ۱$$

سپس باید رابطه $UX = Y$ را با جایگذاری پسرو حل کرد:

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

$$x_۳ = ۱ \quad , \quad x_۲ + ۲ = ۴ \Rightarrow x_۲ = ۴ - ۲ = ۲$$

$$x_۱ + (۲ \times ۲) - ۳ = -۲ \Rightarrow x_۱ = -۲ - ۱ = -۳$$

در نهایت جواب دستگاه معادلات خطی مذکور به صورت زیر خواهد بود:

$$\begin{cases} x_۱ = ۱ \\ x_۲ = ۲ \\ x_۳ = -۳ \end{cases} \Rightarrow X = \begin{bmatrix} ۱ \\ ۲ \\ -۳ \end{bmatrix}$$

همان‌طور که مشاهده می‌شود، جواب یک دستگاه با دو روش تجزیه متفاوت، یکسان خواهد بود.

۳-۳-۳ تجزیه چولسکی

روش تجزیه ماتریس A به ماتریس‌های پایین مثلثی L و بالا مثلثی U که در آن درایه‌های قطر U برابر با درایه‌های قطر L باشند، در بخش (۳-۳-۲) گفته شد. در این بخش توضیحات مربوط به استفاده از این تجزیه و حل مثالی از این روش گردآوری شده است. به طور مشابه پس از به دست آوردن تجزیه LU، با جایگذاری فرم LU به جای ماتریس A در $AX = b$، به صورت $LUX = b$ بازنویسی می‌شود. حال باید UX را مساوی بردار مفروض Y قرار داد تا رابطه $LY = b$ تشکیل شود. یعنی:

$$\begin{bmatrix} l_{۱۱} & ۰ & ۰ & \cdots & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ & \cdots & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} & \cdots & ۰ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n۱} & l_{n۲} & l_{n۳} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_۱ \\ b_۲ \\ b_۳ \\ \vdots \\ b_n \end{bmatrix}$$

حال از ضرب ماتریس L و بردار Y، تعداد $n^۲$ معادله به دست می‌آید که با روش جایگذاری پیشرو به ترتیب می‌توان مقادیر y_i ها را به دست آورد. سپس از رابطه $UX = Y$ یعنی رابطه زیر، مقادیر x_i به دست خواهد آمد.

$$\begin{bmatrix} l_{۱۱} & u_{۱۲} & u_{۱۳} & \cdots & u_{۱n} \\ ۰ & l_{۲۲} & u_{۲۳} & \cdots & u_{۲n} \\ ۰ & ۰ & l_{۳۳} & \cdots & u_{۳n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & ۰ & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \\ \vdots \\ y_n \end{bmatrix}$$

مثال ۳-۱۵) دستگاه معادلات خطی مثال (۳-۱۳) با تجزیه چولسکی به صورت زیر حل خواهد شد.

ابتدا با کمک تجزیه چولسکی، ماتریس ضرایب به LU تجزیه می‌شود:

$$A = \begin{bmatrix} ۲ & ۴ & -۶ \\ ۱ & ۵ & ۳ \\ ۱ & ۳ & ۲ \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} \end{bmatrix} \begin{bmatrix} l_{۱۱} & u_{۱۲} & u_{۱۳} \\ ۰ & l_{۲۲} & u_{۲۳} \\ ۰ & ۰ & l_{۳۳} \end{bmatrix} \Rightarrow$$

$$A = \begin{bmatrix} \sqrt{۲} & ۰ & ۰ \\ \frac{\sqrt{۲}}{۲} & \sqrt{۳} & ۰ \\ \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۳}}{۳} & \sqrt{۳} \end{bmatrix} \begin{bmatrix} \sqrt{۲} & ۲\sqrt{۲} & -۳\sqrt{۲} \\ ۰ & \sqrt{۳} & ۲\sqrt{۳} \\ ۰ & ۰ & \sqrt{۳} \end{bmatrix}$$

پس رابطه $AX = b$ به فرم $LUX = b$ تبدیل خواهد شد. حال با تبدیل UX به بردار Y، رابطه $LY = b$ به دست می‌آید که با کمک جایگذاری پیشرو می‌توان مقادیر y_i را پیدا کرد:

$$y_۱ = -\frac{۴}{\sqrt{۲}} = -۲\sqrt{۲} \;\; , \;\; y_۲ = \frac{۱۰ + ۲}{\sqrt{۳}} = ۴\sqrt{۳} \;\; , \;\; y_۳ = \frac{۵ - ۴ + ۲}{\sqrt{۳}} = \sqrt{۳}$$

سپس باید رابطه $UX = Y$ را با جایگذاری پسرو حل کرد:

$$\begin{bmatrix} \sqrt{۲} & ۲\sqrt{۲} & -۳\sqrt{۲} \\ ۰ & \sqrt{۳} & ۲\sqrt{۳} \\ ۰ & ۰ & \sqrt{۳} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} -۲\sqrt{۲} \\ ۴\sqrt{۳} \\ \sqrt{۳} \end{bmatrix}$$

$$\sqrt{۳}x_۳ = \sqrt{۳} \Rightarrow x_۳ = \frac{\sqrt{۳}}{\sqrt{۳}} = ۱$$

$$\sqrt{۳}x_۲ + ۲\sqrt{۳} = ۴\sqrt{۳} \Rightarrow x_۲ = \frac{۴\sqrt{۳} - ۲\sqrt{۳}}{\sqrt{۳}} = ۲$$

$$\sqrt{۲}x_۱ + \left(۲ \times ۲\sqrt{۲}\right) - ۳\sqrt{۲} = -۲\sqrt{۲} \Rightarrow x_۱ = \frac{-۲\sqrt{۲} - \sqrt{۲}}{\sqrt{۲}} = -۳$$

در نهایت جواب دستگاه معادلات خطی مذکور به صورت زیر خواهد بود:

$$\begin{cases} x_۱ = ۱ \\ x_۲ = ۲ \\ x_۳ = -۳ \end{cases} \Rightarrow X = \begin{bmatrix} ۱ \\ ۲ \\ -۳ \end{bmatrix}$$

همان‌طور که مشاهده می‌شود، جواب تجزیه چولسکی نیز با دو روش تجزیه قبل، یکسان بوده است.

قضیه ۳-۱۶) اگر ماتریس A یک ماتریس حقیقی، متقارن و معین مثبت اکید باشد، در این‌صورت A تجزیه‌ای یکتا به صورت $A = LL^T$ دارد که در آن L یک ماتریس پایین مثلثی با عناصر قطری مثبت می‌باشد.

برهان) ماتریس A متقارن است، یعنی $A^T = A$ و همچنین این ماتریس نیز یک ماتریس معین مثبت اکید است، یعنی $X \in \mathbb{R}^{n\times ۱}, X^T AX > ۰$، طبق یادآوری (۳-۱۲) چون A معین مثبت اکید است بنابراین تجزیه LU دارد و چون ماتریس متقارن است، رابطه زیر برقرار است:

$$A = LU = A^T = U^T L^T \Rightarrow LU = U^T L^T \xRightarrow{\text{معین مثبت اکید}} U(L^T)^{-۱} = L^{-۱}U^T$$

چون A معین مثبت اکید است، بنابراین $\det(A) \neq ۰$ یعنی $\det(U) \neq ۰, \det(L) \neq ۰$. همچنین از طرفی $U(L^T)^{-۱}$ یک ماتریس بالا مثلثی است (زیرا حاصل‌ضرب دو ماتریس بالا مثلثی، بالا مثلثی خواهد بود) و همچنین $L^{-۱}U^T$ یک ماتریس پایین مثلثی است (زیرا حاصل ضرب دو ماتریس پایین مثلثی، پایین مثلثی خواهد بود). بنابراین در واقع این دو ماتریس باید به ناچار قطری باشند، یعنی ماتریس D با درایه‌های قطری (d_{ii}) هست که:

$$U(L^T)^{-۱} = L^{-۱}U^T = D \Rightarrow U = DL^T \xRightarrow{A=LU} A = LDL^T$$

از طرفی A معین مثبت اکید است، بنابراین $۰ \neq X \in \mathbb{R}^{n\times ۱}$ موجود است که:

$$۰ < X^T AX = X^T(LDL^T)X = (L^T X)^T D(L^T X)$$

پس خود ماتریس D نیز معین مثبت اکید است. بنابراین درایه‌های روی قطر D همگی مثبت می‌باشند. بنابراین می‌توان نوشت $D = D^{۱/۲}.D^{۱/۲}$ که درایه‌های روی قطر اصلی $D^{۱/۲}$ همان $\sqrt{d_{ii}}$ می‌باشند. پس:

$$A = LDL^T \Rightarrow A = LD^{۱/۲}D^{۱/۲}L^T = \left(LD^{۱/۲}\right)\left(LD^{۱/۲}\right)^T$$

پس اگر $\tilde{L} = LD^{۱/۲}$ قرار داده شود، بنابراین A به صورت $A = \tilde{L}\tilde{L}^T$ تجزیه می‌شود. ■

قضیه ۳-۱۷) اگر ماتریس A یک ماتریس حقیقی، متقارن و معین مثبت اکید باشد، در این‌صورت A تجزیه‌ای یکتا به صورت $A = U^TU$ دارد که در آن U یک ماتریس بالا مثلثی با عناصر قطری مثبت می‌باشد.

برهان) مشابه با قضیه (۳-۱۶) می‌توان این قضیه را نیز اثبات کرد. ■

مثال ۳-۱۸) دستگاه $\begin{cases} x_۱ + ۲x_۲ + ۴x_۳ = ۰ \\ ۲x_۱ + ۱۳x_۲ + ۲۳x_۳ = ۶ \\ ۴x_۱ + ۲۳x_۲ + ۲۷x_۳ = ۴۶ \end{cases}$ با تجزیه متقارن چولسکی به صورت زیر خواهد بود.

ماتریس ضرایب دستگاه، یک ماتریس متقارن و معین مثبت اکید می‌باشد. حال باید $A = LL^T$ را باز نویسی کرد:

$$A = \begin{bmatrix} ۱ & ۲ & ۴ \\ ۲ & ۱۳ & ۲۳ \\ ۴ & ۲۳ & ۲۷ \end{bmatrix} = \begin{bmatrix} l_{۱۱} & ۰ & ۰ \\ l_{۲۱} & l_{۲۲} & ۰ \\ l_{۳۱} & l_{۳۲} & l_{۳۳} \end{bmatrix} \begin{bmatrix} l_{۱۱} & l_{۲۱} & l_{۳۱} \\ ۰ & l_{۲۲} & l_{۳۲} \\ ۰ & ۰ & l_{۳۳} \end{bmatrix}$$

با ضرب سطرهای ماتریس L در ستون‌های ماتریس L^T به معادلات زیر خواهید رسید:

$$R_۱ \times C_۱ \rightarrow l_{۱۱}^۲ = ۱ \Rightarrow l_{۱۱} = \pm ۱$$

همانند مثال (۲-۱۴) دو مقدار برای $l_{۱۱}$ پیدا شده که مقدار مثبت آن در نظر گرفته خواهد شد.

$$R_۱ \times C_۲ \rightarrow l_{۱۱}l_{۲۱} = ۲ \Rightarrow l_{۲۱} = ۲ \quad , \quad R_۱ \times C_۳ \rightarrow l_{۱۱}l_{۳۱} = ۴ \Rightarrow l_{۳۱} = ۴$$

$$R_۲ \times C_۲ \rightarrow l_{۲۱}^۲ l_{۲۲}^۲ = ۱۳ \Rightarrow l_{۲۲}^۲ = ۱۳ - ۴ = ۹ \Rightarrow l_{۲۲} = \sqrt{۹} = \pm ۳$$

به طور مشابه مقدار مثبت در نظر گرفته می‌شود.

$$R_۲ \times C_۳ \rightarrow l_{۲۱}l_{۳۱} + l_{۲۲}l_{۳۲} = ۲۳ \Rightarrow l_{۳۲} = \frac{۲۳ - ۸}{۳} = \frac{۱۵}{۳} = ۵$$

$$R_۳ \times C_۳ \rightarrow l_{۳۱}^۲ + l_{۳۲}^۲ + l_{۳۳}^۲ = ۷۷ \Rightarrow l_{۳۳} = \sqrt{۷۷ - ۱۶ - ۲۵} = \sqrt{۳۶} = \pm ۶$$

به طور مشابه مقدار مثبت در نظر گرفته می‌شود. به علت تقارن معادلات، ضرب‌های $R_i \times C_j$ با $R_j \times C_i$ یکسان هستند؛ پس می‌توان آن‌ها را از محاسبات حذف کرد. حال می‌توان فرم $A = LL^T$ را به صورت دقیق نوشت:

$$A = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۳ & ۰ \\ ۴ & ۵ & ۶ \end{bmatrix} \begin{bmatrix} ۱ & ۲ & ۴ \\ ۰ & ۳ & ۵ \\ ۰ & ۰ & ۶ \end{bmatrix}$$

پس رابطه $AX = b$ به فرم $LL^T X = b$ تبدیل خواهد شد:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۳ & ۰ \\ ۴ & ۵ & ۶ \end{bmatrix} \begin{bmatrix} ۱ & ۲ & ۴ \\ ۰ & ۳ & ۵ \\ ۰ & ۰ & ۶ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۶ \\ ۴۶ \end{bmatrix}$$

حال با تبدیل UX به بردار Y، رابطه زیر به دست می‌آید:

$$\begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & ۳ & ۰ \\ ۴ & ۵ & ۶ \end{bmatrix} \begin{bmatrix} y_۱ \\ y_۲ \\ y_۳ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۶ \\ ۴۶ \end{bmatrix}$$

با کمک جایگذاری پیشرو می‌توان درایه‌های بردار Y را پیدا کرد:

$$y_۱ = ۰$$

$$۳y_۲ = ۶ \Rightarrow y_۲ = \frac{۶}{۳} = ۲$$

$$۱۰ + ۶y_۳ = ۴۶ \Rightarrow y_۳ = \frac{۴۶ - ۱۰}{۶} = ۶$$

سپس باید رابطه $UX = Y$ را با جایگذاری پسرو حل کرد:

$$\begin{bmatrix} ۱ & ۲ & ۴ \\ ۰ & ۳ & ۵ \\ ۰ & ۰ & ۶ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۲ \\ ۶ \end{bmatrix}$$

$$۶x_۳ = ۶ \Rightarrow x_۳ = \frac{۶}{۶} = ۱$$

$$۳x_۲ + ۵ = ۲ \Rightarrow x_۲ = \frac{۲ - ۵}{۳} = \frac{-۳}{۳} = -۱$$

$$x_۱ - ۲ + ۴ = ۰ \Rightarrow x_۱ = ۲ - ۴ = -۲$$

در نهایت جواب دستگاه معادلات خطی مذکور به صورت زیر خواهد بود:

$$\begin{cases} x_۱ = -۲ \\ x_۲ = -۱ \\ x_۳ = ۱ \end{cases} \Rightarrow X = \begin{bmatrix} -۲ \\ -۱ \\ ۱ \end{bmatrix}$$

نکته مثبت روش تجزیه متقارن چولسکی $A = LL^T$، در کاهش محاسبات آن است. این امر باعث می‌شود که از انجام محاسبات تکراری پرهیز کرده و بسیار سریع‌تر به جواب رسید. همچنین با توجه به مقدارهای متفاوت $l_{۱۱}$، $l_{۲۲}$ و $l_{۳۳}$ می‌توان تجزیه‌های متفاوتی را برای ماتریس A در نظر گرفت.

۳-۴ روش‌های تکراری

در دنیای واقعی، همه‌ی داده‌ها لزوماً به صورت یک عدد صحیح یا گویا نیستند؛ بلکه ممکن است داده‌ها به صورت اعدادی گنگ باشند و یا حتی اعداد گویایی باشند که تعداد ارقام اعشارشان بیشتر از حد مجاز کامپیوتر است. در این حالت داده‌ها به طور دقیق در کامپیوتر ذخیره نمی‌شوند، بلکه مقدار تقریبی و بسیار نزدیک به آن‌ها ذخیره خواهد شد. حال یک دستگاه معادلات خطی را در نظر بگیرید که ضرایب آن به همین صورت باشند. در این حالت نمی‌توان جواب دقیقی برای دستگاه مذکور پیدا کرد، بلکه جوابی تقریبی با استفاده از روش‌های تکراری به دست می‌آید. البته لازم به ذکر است که از روش‌های تکراری برای محاسبه جواب دستگاه با ضرایب گویا نیز می‌توان استفاده کرد.

روش‌های تکراری در واقع برای تعیین ریشه $f(x) = ۰$ استفاده می‌شوند. به طوری که معادله $f(x) = ۰$ به صورت معادله $x = g(x)$ بازنویسی می‌شود. به تابع $g(x)$ نیز تابع تکرار می‌گویند. سپس با در نظر گرفتن یک حدس اولیه $x.$ و با انجام مکرر روش، به ریشه $f(x) = ۰$ می‌توان رسید. پس انتخاب حدس اولیه $x.$ و تابع تکرار $g(x)$ در روش‌های تکراری بسیار حائز اهمیت است.

ایده اصلی کلیه روش‌های تکراری برای حل دستگاه $AX = b$ به این صورت است که جواب دستگاه مذکور را باید به صورت جواب معادله $f(X) = AX - b = ۰$ در نظر گرفت. سپس برای حل این معادله $f(X) = ۰$ ، باید یک معادله $X = g(X)$ تولید کرد. پس از آن یک روش تکراری به صورت $X^{(K+۱)} = g\left(X^{(K)}\right)$ تعریف کرده و حدس اولیه‌ای نظیر $X^{(۰)}$ نیز باید تعیین کرد. حال برای $K = ۰,۱,۲,\cdots$، بردارهای تقریبی $X^{(۱)}$، $X^{(۲)}$، $X^{(۳)}$، ...، $X^{(K+۱)}$ و... را باید تولید کرد.

یک فرم ساده برای تابع $g(X)$ عبارت است از:

$$g(X) = MX + d \qquad (۸-۳)$$

که در آن M ماتریسی $n \times n$ و d برداری $n \times ۱$ و هر دو معلوم هستند. بنابراین با این انتخاب برای تابع $g(X)$ و انتخاب یک حدس اولیه نظیر $X^{(۰)}$ دنباله روش‌های تکراری برای حل دستگاه $AX = b$ به صورت زیر با شرط اولیه $X^{(۰)}$ تولید می‌شود:

$$X^{(K+۱)} = MX^{(K)} + d \quad , \quad K = ۰,۱,۲,\cdots \qquad (۹\_۳)$$

تذکر ۳-۱۹) در روش تکراری $X^{(K+۱)} = MX^{(K)} + d$ انتخاب ماتریس M و بردار d، نوع روش تکراری برای حل دستگاه $AX = b$ را تعیین می‌کند.

چنانچه ماتریس A به صورت $A = B + C$ تفکیک شده باشد که در آن ماتریس B معکوس‌پذیر باشد، می‌توان رابطه $AX = b$ را به صورت زیر بازنویسی کرد:

$$AX = b \Rightarrow (B + C)X = b \Rightarrow BX = -CX + b \xRightarrow{\times B^{-۱}} X = -B^{-۱}CX + B^{-۱}b$$

با انتخاب $d = B^{-۱}b$ و $M = -B^{-۱}C$ در واقع معادله $AX = b$ به صورت $X = MX + d$ بازنویسی شد.

بنابراین برای حل دستگاه $AX = b$ بی‌نهایت روش تکراری به صورت $X^{(K+۱)} = MX^{(K)} + d$ تولید خواهد شد. ولی باید توجه داشت که کدام یک از روش‌های تکراری و تحت چه شرایطی به جواب دستگاه $AX = b$ همگرا خواهد بود. با استفاده از قضیه زیر که به قضیه بنیادی همگرایی روش‌های تکراری خطی معروف است، یک شرط لازم و کافی برای همگرایی روش تکراری $X^{(K+۱)} = MX^{(K)} + d$ به ازای هر انتخاب اولیه $X^{(۰)}$ به دست آورد.

لم ۳-۲۰) احکام زیر معادل هستند:

۱- A یک ماتریس همگراست (یعنی $\lim\limits_{K\to\infty} A^K = ۰$).

۲- به ازای هر نرم $\|.\|$ داریم: $\lim\limits_{K\to\infty} \|A^K\| = ۰$.

۳- $\rho(A) < ۱$.

قضیه ۳-۲۱) فرض کنید M ماتریس حقیقی $n \times n$ باشد و معادله $X = MX + d$ دارای جواب منحصر به فرد X^* باشد. معادله بازگشتی (تکراری) $X^{(K+۱)} = MX^{(K)} + d$ به ازای هر حدس اولیه $X^{(۰)}$ به X^* همگراست اگر و تنها اگر $\rho(M) < ۱$.

برهان) فرض کنیم دنباله $\{X^{(K)}\}_{K\in\mathbb{N}}$ به ازای هر حدس اولیه $X^{(۰)}$ به X^* همگرا باشد. از طرفی X^* جواب دقیق و یکتای معادله $X = MX + d$ است. یعنی:

$$\left.\begin{array}{c} X^* = MX^* + d \\ X^{(K+۱)} = MX^{(K+۱)} + d \end{array}\right\} \Longrightarrow X^{(K+۱)} - X^* = M(X^{(K)} - X^*) \Rightarrow$$

$$X^{(K+۱)} - X^* = M\left(M\left(X^{(K-۱)} - X^*\right)\right) = M^۲\left(X^{(K-۱)} - X^*\right) = \cdots = M^{K+۱}\left(X^{(۰)} - X^*\right)$$

پس $X^{(K+۱)} - X^* = M^{K+۱}\left(X^{(۰)} - X^*\right)$. حال با حدگیری از طرفین رابطه اخیر می‌توان نوشت که:

$$\lim_{K\to\infty} M^{K+۱}\left(X^{(K+۱)} - X^*\right) = ۰ \xRightarrow{\lim\limits_{K\to\infty} X^{(K+۱)} - X^* = ۰} \lim_{K\to\infty} M^{K+۱} = ۰$$

یعنی ماتریس M همگراست و بنابر لم (۳-۲۰)، $\rho(M) < ۱$.

حال عکس برهان، یعنی فرض کنید $\rho(M) < ۱$ و باید ثابت کرد که معادله بازگشتی $X^{(K+۱)} = MX^{(K)} + d$ به ازای هر جواب آغازین $X^{(۰)}$ به جواب منحصر به فرد X^* همگراست. با توجه به فرض، چون $\rho(M) < ۱$ پس $\lim\limits_{K\to\infty} M^{K+۱} = ۰$. بنابراین $\lim\limits_{K\to\infty} M^{K+۱}(X^{(۰)} - X^*) = ۰$، در نتیجه $\lim\limits_{K\to\infty} M^K(X^{(۱)} - X^*) = ۰$ و در نهایت با ادامه همین روند به $\lim\limits_{K\to\infty} (X^{(K+۱)} - X^*) = ۰$ خواهید رسید. زیرا:

$$\lim_{K\to\infty}\left(X^{(K+۱)} - X^*\right) = \lim_{K\to\infty} M^{K+۱}\left(X^{(۰)} - X^*\right) \xrightarrow{\rho(M)<۱} ۰ \ \blacksquare$$

برای توقف تکرارها می‌توان شرایط مختلفی را در نظر گرفت:

۱) می‌توان تعداد تکرارها را از قبل مشخص کرد؛ به طور مثال روش تکراری مذکور تا n تکرار ادامه پیدا کند که در این حالت با محاسبه $X^{(n)}$، محاسبات پایان می‌یابد.

۲) می‌توان محاسبات را تا مرحله‌ای انجام داد که به ازای یک نرم دلخواه خطای مطلق دو تقریب متوالی از عدد داده شده‌ی ε کمتر باشد. یعنی $\left\|X^{(K+۱)} - X^{(K)}\right\| < \varepsilon$ که در آن $\varepsilon > ۰$. به بیان دیگر می‌توان شرط توقف تکرارها را کوچک بودن اختلاف تقریب‌ها در دو مرحله (تکرار) متوالی در نظر گرفت. دقت کنید که هر چه قدر ε کوچک‌تر باشد، تعداد تکرارها افزایش خواهد یافت و جواب در صورت همگرایی، دقت بیشتری خواهد داشت.

۳) می‌توان تکرارها را تا جایی ادامه داد که شرط $\frac{\|X^{(K+۱)} - X^{(K)}\|}{\|X^{(K+۱)}\|} < \varepsilon$ با $\varepsilon > ۰$ دلخواه برقرار باشد.

۳-۴-۱ روش تصفیه تکراری[۱]

قبل از ارائه روش تصفیه تکراری، باید تعریف بردار مانده نظیر دستگاه $AX = b$ را دانست.

تعریف ۳-۲۲) فرض کنید $\tilde{X}$ یک جواب تقریبی برای دستگاه $AX = b$ باشد؛ در این‌صورت بردار مانده نظیر این جواب تقریبی به صورت $r = b - A\tilde{X}$ تعریف خواهد شد (واضح است اگر $\tilde{X}$ جواب دقیق دستگاه $AX = b$ باشد، آنگاه $r = ۰$).

حال اگر $r = b - A\tilde{X}$ بردار مانده جواب تقریبی $\tilde{X}$ باشد، آنگاه به وضوح:

$$r = b - A\tilde{X} = AX - A\tilde{X} = A\left(X - \tilde{X}\right)$$

اگر $X - \tilde{X}$ را مساوی با بردار دلخواه Y گرفت، دستگاه جدید $AY = r$ تولید خواهد شد. حال با حل این دستگاه جواب تقریبی $\tilde{Y}$ به دست خواهد آمد، بنابراین $\tilde{Y} \approx X - \tilde{X}$ پس می‌توان نتیجه گرفت $X = \tilde{X} + \tilde{Y}$ که تقریبی بهتر از $\tilde{X}$ برای حل دستگاه $AX = b$ خواهد بود.

مثال ۳-۲۳) فرض کنید دستگاه $\begin{cases} x_۱ - ۰/۳۳۳۳x_۲ + ۳/۳۳۳۳x_۳ = ۴ \\ -x_۱ - ۰/۳۳۳۳x_۲ + ۶/۶۶۶۷x_۳ = ۵/۳۳۳۳ \\ ۳/۳۳۳۳x_۲ + ۱۳/۳۳۳۳x_۳ = ۱۶/۶۶۶۷ \end{cases}$ با دقت ۵ رقم اعشار توسط روش حذفی گاوس حل شده. حال برای بهبود بخشیدن جواب (بهتر کردن جواب تقریبی) باید از روش تصفیه تکراری استفاده کرد. در مرحله آخر روش حذفی گاوس ماتریس افزوده زیر به دست می‌آید:

$$\left[\begin{array}{ccc|c} ۱ & -۰/۳۳۳۳ & ۳/۳۳۳۳ & ۴ \\ ۰ & -۰/۶۶۶۶ & ۱۰ & ۹/۳۳۳۳ \\ ۰ & ۰ & ۶۳/۳۳۸۳ & ۶۳/۳۳۷۹ \end{array}\right] \Rightarrow \tilde{X} = \begin{bmatrix} ۱ \\ ۱/۰۰۰۲ \\ ۱/۰۰۰۱ \end{bmatrix} \Rightarrow r = b - A\tilde{X} = \begin{bmatrix} ۰ \\ ۰/۰۰۰۱ \\ -۰/۰۰۰۶ \end{bmatrix}$$

[۱] Iterative Refinement Method

سپس باید دستگاه $AY = r$ را حل کرد تا جواب تقریبی $\tilde{Y}$ پیدا شود:

$$\tilde{Y} = \begin{bmatrix} -۰/۰۰۰۱ \\ -۰/۰۰۰۲ \\ ۰ \end{bmatrix}$$

پس یک تقریب بهتر برای جواب دستگاه $AX = b$ عبارت است از:

$$X = \tilde{X} + \tilde{Y} = \begin{bmatrix} ۱/۰۰۰۱ \\ ۱/۰۰۰۲ \\ ۱ \end{bmatrix} + \begin{bmatrix} -۰/۰۰۰۱ \\ -۰/۰۰۰۲ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \end{bmatrix}$$

دقت کنید که جواب به دست آمده همان جواب دقیق دستگاه $AX = b$ شده است.

۳-۴-۲ روش تکراری ژاکوبی[۱]

فرض کنید دستگاه زیر داده شده باشد و $a_{ii} \neq ۰$ (چنانچه این شرط برقرار نباشد، می‌توان با جابه‌جایی معادلات آن را به دست آورد.):

$$\begin{cases} a_{۱۱}x_۱ + a_{۱۲}x_۲ + \cdots + a_{۱n}x_n = b_۱ \\ a_{۲۱}x_۱ + a_{۲۲}x_۲ + \cdots + a_{۲n}x_n = b_۲ \\ \vdots \\ a_{n۱}x_۱ + a_{n۲}x_۲ + \cdots + a_{nn}x_n = b_n \end{cases}$$

حال از معادله i ام، مقدار x_i به صورت زیر به دست می‌آید:

$$x_i = \frac{۱}{a_{ii}}\left(b_i - \sum_{\substack{j=۱ \\ j\neq i}}^{n} a_{ij}x_j \right), \quad i = ۱,۲,۳,\cdots,n \qquad (۱۰\text{-}۳)$$

در واقع روش تکراری ژاکوبی در مرحله K-ام به صورت زیر خواهد بود:

$$x_i^{(K)} = \frac{۱}{a_{ii}}\left(b_i - \sum_{\substack{j=۱ \\ j\neq i}}^{n} a_{ij}x_j^{(K-۱)} \right), \quad i = ۱,۲,۳,\cdots,n\,, K = ۱,۲,\cdots \qquad (۱۱\text{-}۳)$$

که در آن $x_i^{(۰)}$ (حدس اولیه) نیز معلوم است.

---
[۱] Jacobi Iterative Method

محاسبات انجام شده به فرم ماتریسی به صورت زیر خواهد بود:

چنانچه ماتریس ضرایب دستگاه $AX = b$، به صورت $A = B + D$ تفکیک شود که در آن ماتریس قطری D برابر باشد با $diag(a_{۱۱}, a_{۲۲}, \cdots, a_{nn})$، B همان ماتریس A باشد که فقط درایه‌های روی قطر اصلی آن صفر شده‌اند و $a_{ii} \neq ۰$؛ بنابراین:

$$A = B + D \Rightarrow AX = b \Rightarrow (B + D)X = b \Rightarrow DX = -BX + b$$

چون $a_{ii} \neq ۰$ پس D معکوس‌پذیر است و رابطه زیر برقرار است:

$$X = -D^{-۱}BX + D^{-۱}b \qquad (۱۲\text{-}۳)$$

حال روش تکراری ژاکوبی را می‌توان به صورت ماتریسی زیر بازنویسی کرد:

$$X^{(K)} = -D^{-۱}BX^{(K-۱)} + D^{-۱}b\,, \quad K = ۱,۲,\cdots \qquad (۱۳\text{-}۳)$$

مثال ۳-۲۴) دستگاه معادلات $\begin{cases} ۸x_۱ + ۲x_۲ + x_۳ = ۷ \\ ۳x_۱ + ۹x_۲ + ۲x_۳ = ۱ \\ ۲x_۱ + x_۲ + ۷x_۳ = -۵ \end{cases}$ با روش ژاکوبی به صورت زیر حل خواهد شد.

ابتدا باید از معادله اول $x_۱$، از معادله دوم $x_۲$ و از معادله سوم $x_۳$ را به دست آورد. پس:

$$x_۱ = \frac{۱}{۸}(۷ - ۲x_۲ - x_۳)\,, x_۲ = \frac{۱}{۹}(۱ - ۳x_۱ - ۲x_۳)\,, x_۳ = \frac{۱}{۷}(-۵ - ۲x_۱ - x_۲)$$

حال باید فرم تکراری ژاکوبی را تشکیل داد. یعنی:

$$x_۱^{(K)} = \frac{۱}{۸}\left(۷ - ۲x_۲^{(K-۱)} - x_۳^{(K-۱)}\right)$$

$$x_۲^{(K)} = \frac{۱}{۹}\left(۱ - ۳x_۱^{(K-۱)} - ۲x_۳^{(K-۱)}\right)$$

$$x_۳^{(K)} = \frac{۱}{۷}\left(-۵ - ۲x_۱^{(K-۱)} - x_۲^{(K-۱)}\right)$$

با انتخاب حدس اولیه $X^{(۰)} = \left(x_۱^{(۰)}, x_۲^{(۰)}, x_۳^{(۰)}\right)^T = (۰,۰,۰)^T$ ادامه محاسبات به صورت زیر خواهد بود:

$$\begin{cases} x_۱^{(۱)} = \frac{۱}{۸}(۷ - ۲ \times ۰ - ۰) = \frac{۷}{۸} \\ x_۲^{(۱)} = \frac{۱}{۹}(۱ - ۳ \times ۰ - ۰) = \frac{۱}{۹} \\ x_۱^{(۱)} = \frac{۱}{۷}(-۵ - ۲ \times ۰ - ۰) = \frac{-۵}{۷} \end{cases} \Rightarrow X^{(۱)} = \begin{bmatrix} \frac{۷}{۸} \\ \frac{۱}{۹} \\ -\frac{۵}{۷} \end{bmatrix} = \begin{bmatrix} x_۱^{(۱)} \\ x_۲^{(۱)} \\ x_۳^{(۱)} \end{bmatrix}$$

حال با قرار دادن $X^{(۱)}$ در معادله‌های $x_i^{(K)}$ به روابط زیر می‌توان رسید:

$$\begin{cases} x_۱^{(۲)} = \frac{۱}{۸}\left(۷ - ۲x_۲^{(۱)} - x_۳^{(۱)}\right) \\ x_۲^{(۲)} = \frac{۱}{۹}\left(۱ - ۳x_۱^{(۱)} - ۲x_۳^{(۱)}\right) \\ x_۳^{(۲)} = \frac{۱}{۷}\left(-۵ - ۲x_۱^{(۱)} - x_۲^{(۱)}\right) \end{cases} \Rightarrow \begin{cases} x_۱^{(۲)} = \frac{۱}{۸}\left(۷ - ۲ \times \frac{۱}{۹} - \left(\frac{-۵}{۷}\right)\right) = \frac{۵۹}{۶۳} \\ x_۲^{(۲)} = \frac{۱}{۹}\left(۱ - ۳ \times \frac{۷}{۸} - ۲\left(\frac{-۵}{۷}\right)\right) = \frac{-۱۱}{۵۰۴} \\ x_۳^{(۲)} = \frac{۱}{۷}\left(-۵ - ۲ \times \frac{۷}{۸} - \frac{۱}{۹}\right) = -\frac{۲۴۷}{۲۵۲} \end{cases}$$

$$\Rightarrow X^{(۲)} = (۰/۹۳۶۵, -۰/۰۲۱۸, -۰/۹۸۰۲)^T$$

با تکرار همین روش، در ادامه محاسبات به صورت زیر خواهد بود:

$$X^{(۳)} = (۱/۰۰۳۰, ۰/۰۱۶۸, -۰/۹۸۸۷)^T$$

$$X^{(۴)} = (۰/۹۹۳۲, -۰/۰۰۵۷, -۱/۰۰۳۲)^T$$

$$X^{(۵)} = (۱/۰۰۱۸, ۰/۰۰۳۰, -۰/۹۹۷۲)^T$$

$$X^{(۶)} = (۰/۹۹۸۹, -۰/۰۰۱۲, -۱/۰۰۱۰)^T$$

$$\vdots$$

$$X^{(۱۱)} = (۱/۰۰۰۰, ۰/۰۰۰۰, -۱/۰۰۰۰)^T$$

قضیه ۳-۲۵) فرض کنید A ماتریسی اکیداً قطر غالب (سطری و یا ستونی) باشد، در این صورت روش تکراری ژاکوبی برای حل دستگاه $AX = b$ به ازای هر انتخاب اولیه $X^{(۰)}$ به جواب این دستگاه همگراست.

برهان) ماتریس $A = (a_{ij})$ اکیداً قطر غالب سطری است، هرگاه:

$$\sum_{\substack{j=۱ \\ j \neq i}}^{n} |a_{ij}| < |a_{ii}|, \quad i = ۱,۲,۳,\cdots,n$$

به همین ترتیب این ماتریس اکیداً قطر غالب ستونی است، هرگاه:

$$\sum_{\substack{i=۱ \\ i \neq j}}^{n} |a_{ij}| < |a_{jj}|, \quad j = ۱,۲,۳,\cdots,n$$

حال چون ماتریس A اکیداً قطر غالب سطری است، بنابراین $\sum \left|\frac{a_{ij}}{a_{ii}}\right| < ۱$. حال ماتریس تکراری ژاکوبی یعنی M_J را تشکیل داده و باید ثابت کرد که $\rho(M_J) < ۱$.

با توجه به فرم ماتریسی روش ژاکوبی می‌توان گفت:

$$M_J = -D^{-۱}B \Rightarrow M_J = -\begin{bmatrix} \frac{۱}{a_{۱۱}} & ۰ & \cdots & ۰ \\ ۰ & \frac{۱}{a_{۲۲}} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & \frac{۱}{a_{nn}} \end{bmatrix} \begin{bmatrix} ۰ & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & ۰ & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n۱} & a_{n۲} & \cdots & ۰ \end{bmatrix}$$

$$\Rightarrow M_J = -\begin{bmatrix} ۰ & \frac{a_{۱۲}}{a_{۱۱}} & \cdots & \frac{a_{۱n}}{a_{۱۱}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n۱}}{a_{nn}} & \frac{a_{n۲}}{a_{nn}} & \cdots & ۰ \end{bmatrix}$$

بنابراین:

$$\|M_J\|_\infty = \max_i \sum_{\substack{j=۱ \\ j\neq i}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| \xrightarrow{\text{قطر غالب سطری اکید}} \|M_J\|_\infty < ۱$$

در نتیجه $\rho(M_J) < ۱$. یعنی روش ژاکوبی برای هر تکرار اولیه $X^{(۰)}$ همگرا خواهد بود. روند اثبات برای اکیداً قطر غالب ستونی به صورت مشابه است. ■

۳-۴-۳ روش تکراری گاوس-سایدل[۱]

این روش مشابه روش ژاکوبی است با این تفاوت که به محض محاسبه $x_۱^{(۱)}$ از معادله اول، برای محاسبه $x_۲^{(۱)}$ در معادله دوم، دیگر از $x_۱^{(۰)}$ استفاده نمی‌شود و به جای آن از مقدار $x_۱^{(۱)}$ که از معادله قبل به دست آمده استفاده می‌کنیم. به همین ترتیب برای محاسبه $x_۳^{(۱)}$، به جای استفاده از $x_۱^{(۰)}$ و $x_۲^{(۰)}$، باید از مقادیر به دست آمده از معادله اول و دوم یعنی $x_۱^{(۱)}$ و $x_۲^{(۱)}$ استفاده کرد. به همین ترتیب با محاسبه هر $x_i^{(K)}$، در محاسبات مربوط به $x_{i+۱}^{(K)}$ دیگر، باید از مقادیر $x_i^{(K)}$ جدید به دست آمده از معادلات قبلی استفاده کرد. این روش می‌تواند همگرایی را سرعت ببخشد.

در حالت کلی روش گاوس-سایدل در مرحله K-ام به صورت زیر می‌باشد:

$$x_i^{(K)} = \frac{۱}{a_{ii}}\left(b - \sum_{j=۱}^{i-۱} a_{ij}x_j^{(K)} - \sum_{j=i+۱}^{n} a_{ij}x_j^{(K-۱)}\right), \quad i = ۱,۲,\cdots,n \qquad (۱۴\text{-}۳)$$

---
[۱] Gauss-Seidel Iterative Method

محاسبات انجام شده به فرم ماتریسی به صورت زیر خواهد بود:

چنانچه ماتریس ضرایب دستگاه $AX = b$، به صورت $A = L + D + U$ تفکیک شود که در آن ماتریس قطری D برابر باشد با $diag(a_{۱۱}, a_{۲۲}, \cdots, a_{nn})$، L ماتریسی شامل درایه‌های زیر قطر اصلی A و U ماتریسی شامل درایه‌های بالای قطر اصلی A باشد؛ بنابراین:

$$AX = b \Rightarrow (L + D + U)X = b \Rightarrow (L + D)X = -UX + b$$

رابطه فوق در مرحله K-ام به صورت زیر تبدیل خواهد شد:

$$(L + D)X^{(K)} = -UX^{(K-۱)} + b \Rightarrow DX^{(K)} = -LX^{(K)} - UX^{(K-۱)} + b$$

حال روش تکراری ژاکوبی را می‌توان به صورت ماتریسی زیر بازنویسی کرد:

$$X^{(K)} = -D^{-۱}\left[b - LX^{(K)} - UX^{(K-۱)}\right], \quad K = ۱,۲,\cdots \qquad (۱۵\text{-}۳)$$

مثال ۳-۲۶) دستگاه مثال (۳-۲۴) با روش گاوس سایدل به صورت زیر حل خواهد شد.

ابتدا $x_۱$، $x_۲$ و $x_۳$ را به ترتیب از معادلات اول، دوم و سوم باید به صورت زیر به دست آورد:

$$\begin{cases} x_۱ = \frac{۱}{۸}(۷ - ۲x_۲ - x_۳) \\ x_۲ = \frac{۱}{۹}(۱ - ۳x_۱ - ۲x_۳) \\ x_۳ = \frac{۱}{۷}(-۵ - ۲x_۱ - x_۲) \end{cases} \Rightarrow \begin{cases} x_۱^{(K)} = \frac{۱}{۸}\left(۷ - ۲x_۲^{(K-۱)} - x_۳^{(K-۱)}\right) \\ x_۲^{(K)} = \frac{۱}{۹}\left(۱ - ۳x_۱^{(K)} - ۲x_۳^{(K-۱)}\right) \\ x_۳^{(K)} = \frac{۱}{۷}\left(-۵ - ۲x_۱^{(K)} - x_۲^{(K)}\right) \end{cases}$$

با انتخاب حدس اولیه $X^{(۰)} = \begin{bmatrix} ۰ \\ ۰ \\ ۰ \end{bmatrix}$ ادامه محاسبات به صورت زیر خواهد بود:

$$x_۱^{(۱)} = \frac{۱}{۸}(۷ - ۲ \times ۰ - ۰) = \frac{۷}{۸} = ۰/۸۷۵۰$$

حال با قرار دادن این مقدار $x_۱^{(۱)}$ در معادله دوم، معادله زیر به دست می‌آید:

$$x_۲^{(۱)} = \frac{۱}{۹}(۱ - ۳ \times ۰/۸۷۵۰ - ۰) = -۰/۱۸۰۶$$

حال با قرار دادن این مقادیر $x_۱^{(۱)}$ و $x_۲^{(۱)}$ در معادله سوم، معادله زیر به دست می‌آید:

$$x_۳^{(۱)} = \frac{۱}{۷}(-۵ - ۲ \times ۰/۸۷۵۰ + ۰/۱۸۰۶) = -۰/۹۳۸۵$$

لذا $X^{(1)} = \begin{bmatrix} ۰/۸۷۵۰ \\ -۰/۱۸۰۶ \\ -۰/۹۳۸۵ \end{bmatrix}$ و ادامه محاسبات به صورت زیر خواهد بود:

$$X^{(۲)} = \begin{bmatrix} ۱/۰۳۷۵ \\ -۰/۰۲۶۲ \\ -۱/۰۰۶۰ \end{bmatrix}, X^{(۳)} = \begin{bmatrix} ۱/۰۰۷۴ \\ -۰/۰۰۰۹ \\ -۱/۰۰۲۰ \end{bmatrix}, \cdots, X^{(۶)} = \begin{bmatrix} ۱/۰۰۰۰ \\ ۰/۰۰۰۰ \\ -۱/۰۰۰۰ \end{bmatrix}$$

تذکر ۳-۲۷) با مقایسه مثال‌های (۳-۲۵) و (۳-۲۴) متوجه خواهید شد که روش گاوس-سایدل در تکرارهای کمتری به جواب دقیق خواهد رسید.

مثال ۳-۲۸) جواب تقریبی دستگاه $\begin{cases} ۱۰x_۱ + x_۲ + x_۳ = ۱۲ \\ x_۱ + ۱۰x_۲ + x_۳ = ۱۲ \\ x_۱ + x_۲ + ۱۰x_۳ = ۱۲ \end{cases}$ با روش‌های ژاکوبی و گاوس-سایدل با حدس اولیه $X^{(۰)} = (۰,۰,۰)^T$ تا ۳ تکرار به صورت زیر خواهد بود.

ماتریس ضرایب $A = \begin{bmatrix} ۱۰ & ۱ & ۱ \\ ۱ & ۱۰ & ۱ \\ ۱ & ۱ & ۱۰ \end{bmatrix}$ می‌باشد.

همچنین ماتریس تکراری ژاکوبی به صورت $A = B + D$ می‌باشد، یعنی:

$$B = \begin{bmatrix} ۰ & ۱ & ۱ \\ ۱ & ۰ & ۱ \\ ۱ & ۱ & ۰ \end{bmatrix}, D = \begin{bmatrix} ۱۰ & ۰ & ۰ \\ ۰ & ۱۰ & ۰ \\ ۰ & ۰ & ۱۰ \end{bmatrix}$$

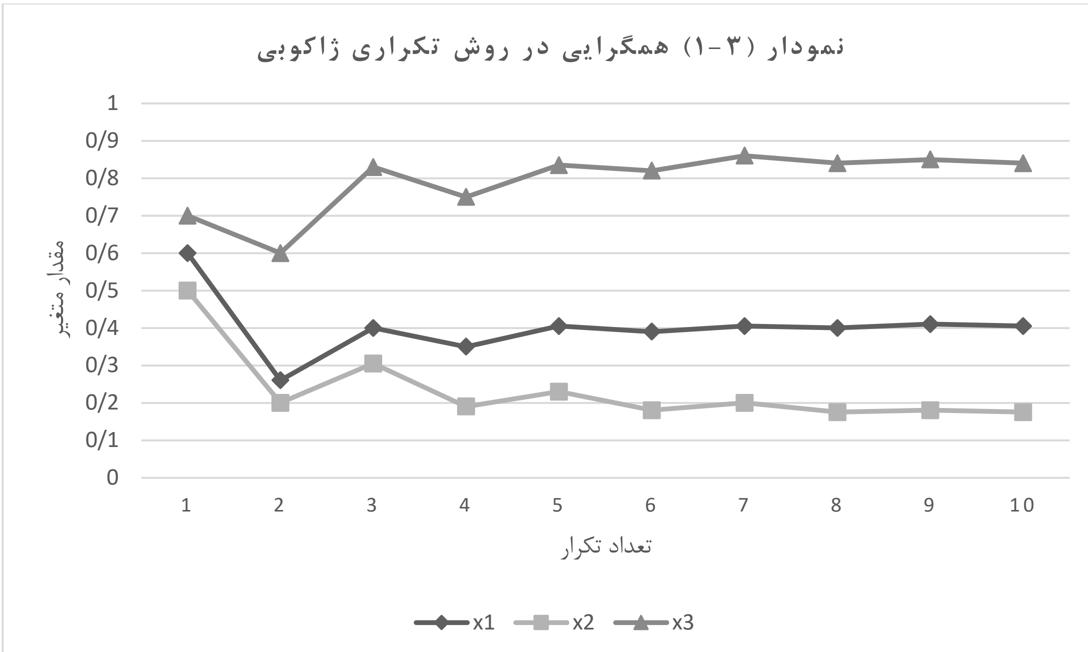
و ماتریس تکراری گاوس-سایدل نیز به صورت $A = L + D + U$ می‌باشد. یعنی:

$$L = \begin{bmatrix} ۰ & ۰ & ۰ \\ ۱ & ۰ & ۰ \\ ۱ & ۱ & ۰ \end{bmatrix}, D = \begin{bmatrix} ۱۰ & ۰ & ۰ \\ ۰ & ۱۰ & ۰ \\ ۰ & ۰ & ۱۰ \end{bmatrix}, U = \begin{bmatrix} ۰ & ۱ & ۱ \\ ۰ & ۰ & ۱ \\ ۰ & ۰ & ۰ \end{bmatrix}$$

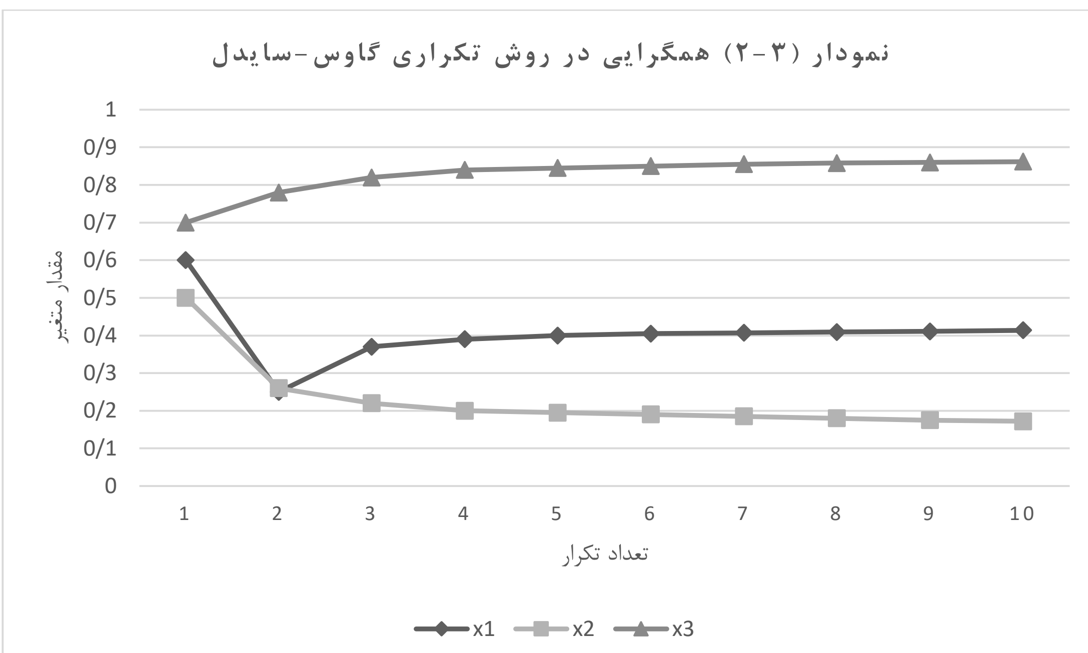
با روش ژاکوبی جواب تقریبی به دست آمده پس از ۳ تکرار برابر خواهد بود با $X_J^{(۳)} = \begin{bmatrix} ۱/۰۰۸ \\ ۱/۰۰۸ \\ ۱/۰۰۸ \end{bmatrix}$ و با روش گاوس-سایدل جواب تقریبی به دست آمده بعد از ۳ تکرار برابر خواهد بود با $X_G^{(۳)} = \begin{bmatrix} ۰/۹۹۹۶ \\ ۱/۰۰۰۰ \\ ۱/۰۰۰۰ \end{bmatrix}$. پس می‌توان نشان داد که:

$$\left\| X - X_J^{(۳)} \right\|_\infty = ۰/۰۰۸ \quad > \quad \left\| X - X_G^{(۳)} \right\|_\infty = ۰/۰۰۴$$

مثال ۳-۲۹) نمودار همگرایی دستگاه معادلات خطی $\begin{cases} ۴x_۱ + ۲x_۲ = ۲ \\ ۲x_۱ + ۱۰x_۲ + ۴x_۳ = ۶ \\ ۴x_۲ + ۵x_۳ = ۵ \end{cases}$ در روش‌های ژاکوبی و گاوس-سایدل پس از ۱۰ تکرار با حدس اولیه $X^{(۰)} = (۰/۶,۰/۵,۰/۷)$ در ادامه آورده شده است.



نمودار (۳-۱) همگرایی در روش تکراری ژاکوبی



نمودار (۳-۲) همگرایی در روش تکراری گاوس-سایدل

تذکر ۳-۳۰) در صورت همگرایی هر دو روش، روش گاوس-سایدل سرعت همگرایی بیشتری به نسبت روش ژاکوبی خواهد داشت.

قضیه ۳-۳۱) فرض کنید A ماتریسی اکیداً قطر غالب باشد. در این صورت روش تکراری گاوس-سایدل برای حل دستگاه $AX = b$ به ازای هر انتخاب اولیه $X^{(\circ)}$ به جواب دقیق دستگاه همگراست.

برهان) باید اثبات کرد که $\rho(M_G) < ۱$. ابتدا اثبات برای ماتریس اکیداً قطر غالب سطری اثبات خواهد شد و سپس با روشی مشابه می‌توان حکم را برای ماتریس اکیداً قطر غالب ستونی، ثابت کرد. فرض کنید λ یک مقدار ویژه‌ی دلخواه ماتریس تکراری گاوس-سایدل یعنی M_G باشد و $V = (v_۱, v_۲, \cdots, v_n)$ بردار ویژه ناصفر نظیر λ باشد، به طوری که $\|V\|_\infty = ۱$. بنابراین می‌توان نوشت که:

$$M_G V = \lambda V \Rightarrow -(D+L)^{-۱} UV = \lambda V \Rightarrow -UV = (D+L)\lambda V \Rightarrow$$

$$-\sum_{j=i+۱}^{n} a_{ij} v_j = \lambda \sum_{j=۱}^{i} a_{ij} v_j \ , \quad \forall i = ۱,۲,\cdots,n$$

حال فرض کنید که $\|V\|_\infty = |V_k| = ۱$ به ازای $۱ \le k \le n$. حال چنانچه رابطه قبل برای $i = k$ نوشته شد، در ادامه:

$$\lambda a_{kk} v_k = -\lambda \sum_{j=۱}^{k-۱} a_{kj} v_j - \sum_{j=k+۱}^{n} a_{kj} v_j \Rightarrow |\lambda||a_{kk} v_k| \le |\lambda| \sum_{j=۱}^{k-۱} |a_{kj}| + \sum_{j=k+۱}^{n} |a_{kj}|$$

$$\xRightarrow{|a_{kk} v_k| = |a_{kk}|\overbrace{|v_k|}^{=۱} = |a_{kk}|} |\lambda| \le \frac{\sum_{j=k+۱}^{n} |a_{kj}|}{|a_{kk}| - \sum_{j=۱}^{k-۱} |a_{kj}|} < ۱$$

زیرا بنابر فرض ماتریس A اکیداً قطر غالب سطری می‌باشد و این بدین معنی است که:

$$|a_{kk}| > \sum_{j=۱}^{k-۱} |a_{kj}| + \sum_{j=k+۱}^{n} |a_{kj}|$$

چون λ دلخواه بود، بنابراین $\rho(M_G) < ۱$. این بدین معنی است که روش تکراری گاوس-سایدل به ازای هر تکرار اولیه $X^{(\circ)}$ به جواب دستگاه $AX = b$ همگراست.

به طور مشابه می‌توان قضیه را برای ماتریس اکیداً قطر غالب ستونی اثبات کرد. ■

لم ۳-۳۲) فرض کنید A متقارن و معین مثبت اکید باشد و P ماتریسی نامنفرد که $B = P + P^T - A$ معین مثبت اکید باشد؛ در این صورت ماتریس $M = I - P^{-۱} A$ همگراست.

قضیه ۳-۳۳) فرض کنید A ماتریسی متقارن و معین مثبت اکید باشد، در این صورت روش تکراری گاوس-سایدل برای حل دستگاه $AX = b$، به ازای هر انتخاب اولیه $X^{(۰)}$ به جواب این دستگاه همگراست.

برهان) چون A متقارن است پس $L^T = U$. از طرفی چون A معین مثبت اکید است بنابراین درایه‌های روی قطر اصلی آن مثبت می‌باشد و در نتیجه $L + D$ معکوس پذیر است. حال چنانچه در لم (۳-۳۲) ماتریس P را برابر با $L + D$ قرار داد، $D = P + P^T - A$. زیرا:

$$P + P^T - A = (L + D) + (L + D)^T - A = L + D + L^T + D^T - A = D$$

چون درایه‌های روی قطر اصلی D مثبت بوده‌اند پس D معین مثبت اکید است. بنابر لم (۳-۳۲) $M = I - P^{-۱}A$ همگراست. سپس باید نشان داد که M همان ماتریس تکراری گاوس-سایدل است. یعنی:

$$M = I - P^{-۱}A = I - (L + D)^{-۱}A = I - (L + D)^{-۱}[L + D + U] \Rightarrow$$
$$M = I - [I + (L + D)^{-۱}U] = -(L + D)^{-۱}U = M_G \quad \blacksquare$$

۳-۴-۴ روش فوق (تحت) تخفیف متوالی[۱]

در روش گاوس-سایدل، روش تکراری بر مبنای تفکیک ماتریس A به صورت $A = L + D + U$ بود. یعنی به صورت $(L + D)X^{(K+۱)} = -UX^{(K)} + b$ بازنویسی می‌شود. حال چنانچه روش گاوس-سایدل همگرا باشد، برای سرعت بخشی به این روش، ماتریس A به صورت زیر تفکیک خواهد شد:

$$A = L + \lambda D + (۱ - \lambda)D + U\,, \quad \lambda \in \mathbb{R}, \lambda \neq ۰ \qquad (۳\text{-}۱۶)$$

حال بر این اساس، روش تکراری جدیدی که شامل پارامتر λ می‌باشد به شرح زیر تولید می‌شود:

$$(L + \lambda D)X^{(K+۱)} = [(\lambda - ۱)D - U]X^{(K)} + b \qquad (۳\text{-}۱۷)$$

با فرض $\omega = \frac{۱}{\lambda}$ رابطه فوق به رابطه زیر تبدیل خواهد شد:

$$(D + \omega L)X^{(K+۱)} = [(۱ - \omega)D - \omega U]X^{(K)} + \omega b \qquad (۳\text{-}۱۸)$$

دقت کنید که در رابطه (۳-۱۸)، L ماتریسی شامل درایه‌های زیر قطر اصلی، U ماتریس شامل درایه‌های بالای قطر اصلی و D شامل قطر اصلی ماتریس A است. شایان ذکر است که رابطه (۳-۱۸) برای $\omega = ۱$ به روش تکراری گاوس-سایدل تبدیل خواهد شد. درایه‌های بردار جواب به روش فوق تخفیف متوالی نیز به شرح زیر خواهد بود:

$$x_i^{(K+۱)} = (۱ - \omega)x_i^{(K)} + \frac{\omega}{a_{ii}}[b_i - \sum_{j=۱}^{i-۱} a_{ij}x_j^{(K+۱)} - \sum_{j=i+۱}^{n} a_{ij}x_j^{(K)} \qquad (۳\text{-}۱۹)$$

---

[۱] Successive Over (Under) Relaxation

در واقع رابطه اخیر از بازنویسی فرمول (۳-۱۸) به صورت زیر استخراج می‌شود:

$$DX^{(K+1)} = (1-\omega)DX^{(K)} + \omega\left(b - LX^{(K+1)} - UX^{(K)}\right) \qquad (۳-۲۰)$$

با ضرب طرفین رابطه (۳-۲۰) در D^{-1} به رابطه زیر خواهید رسید:

$$X^{(K+1)} = (1-\omega)X^{(K)} + \omega D^{-1}\left(b - LX^{(K+1)} - UX^{(K)}\right) \qquad (۳-۲۱)$$

ماتریس تکرار روش مذکور نیز با استفاده از رابطه (۳-۱۸) به دست خواهد آمد. بدین صورت که از همان رابطه ذکر شده به رابطه زیر می‌توان رسید:

$$X^{(K+1)} = (D+\omega L)^{-1}[(1-\omega)D - \omega U]X^{(K)} + \omega(D+\omega L)^{-1}b \qquad (۳-۲۲)$$

یعنی ماتریس تکرار روش فوق تخفیف و تحت تخفیف متوالی عبارتست از:

$$\underset{(SUR)}{M_{SOR}} = (D+\omega L)^{-1}[(1-\omega)D - \omega U] \qquad (۳-۲۳)$$

تذکر ۳-۳۴) اگر $0 < \omega < 1$ روش (۳-۱۸) را روش تحت تخفیف متوالی (SUR) و اگر $1 < \omega < 2$ به آن روش فوق تخفیف متوالی (SOR) گویند.

قضیه ۳-۳۵) فرض کنید $M = (D+\omega L)^{-1}[(1-\omega)D - \omega U]$ معرف ماتریس تکرار SOR و SUR باشد. چنانچه ماتریس ضرایب A و همچنین $A - M^TAM$ ماتریسی معین مثبت اکید باشند، آنگاه $\rho(M) < 1$. یعنی روش‌های فوق تخفیف و تحت تخفیف متوالی به ازای هر انتخاب اولیه $X^{(0)}$ به جواب دستگاه $AX = b$ همگرا خواهند بود.

برهان) بنابر فرض A و $A - M^TAM$ معین مثبت اکید می‌باشند. فرض کنید λ مقدار ویژه دلخواه ماتریس M باشد و بردار ناصفر X بردار ویژه‌ی نظیر مقدار ویژه λ باشد؛ بنابراین:

$$MX = \lambda X \Rightarrow (MX)^T = \lambda X^T \Rightarrow (MX)^TA(MX) = (\lambda X^T)A(\lambda X) = \lambda^2 X^TAX$$
$$\Rightarrow X^T(M^TAM)X = \lambda^2 X^TAX \Rightarrow -\lambda^2 X^TAX = X^T(A - M^TAM)X - X^TAX$$

چون $X^TAX > 0$ و $X^T(A - M^TAM)X > 0$ بنابراین می‌توان نوشت:

$$(1-\lambda^2)\underbrace{X^TAX}_{>0} = \underbrace{X^T(A - M^TAM)X}_{>0} \Rightarrow 1 - \lambda^2 > 0$$
$$\Rightarrow \lambda^2 < 1 \Rightarrow |\lambda| < 1 \Rightarrow \rho(M) < 1 \quad \blacksquare$$

قضیه ۳-۳۶) در روش‌های تکراری فوق تخفیف و تحت تخفیف متوالی، چنانچه $|\omega - 1| \geq 1$ یا به عبارت دیگر مقدار ω در خارج از بازه $(0,2)$ باشد، آنگاه این روش‌ها به ازای هر شروع اولیه $X^{(0)}$ همگرا نخواهد بود.

برهان) رجوع شود به [۵].

قضیه ۳-۳۷) اگر A ماتریسی متقارن و معین مثبت اکید باشد. چنانچه $\omega \in (۰,۲)$، آنگاه روش‌های فوق تخفیف و تحت تخفیف متوالی برای هر انتخاب اولیه $X^{(۰)}$ همگرا خواهد بود.

برهان) رجوع شود به [۵].

مثال ۳-۳۸) دستگاه معادلات خطی $\begin{cases} x_۱ + ۲x_۲ - ۲x_۳ = ۱ \\ x_۱ + x_۲ + x_۳ = ۳ \\ ۲x_۱ + ۲x_۲ + x_۳ = ۵ \end{cases}$ با روش تحت تخفیف متوالی و $\omega = \frac{۱}{۴}$ تا ۳ تکرار تقریبی به صورت زیر خواهد داد.

حدس اولیه را به صورت $X^{(۰)} = \begin{bmatrix} ۰ \\ ۰ \\ ۰ \end{bmatrix}$ در نظر بگیرید. حال باید فرم کلی معادلات را تشکیل داد:

$$\begin{cases} x_۱^{(K)} = \frac{۳}{۴}x_۱^{(K)} + \frac{۱}{۴}\left(۱ - ۲x_۲^{(K-۱)} + ۲x_۳^{(K-۱)}\right) \\ x_۲^{(K)} = \frac{۳}{۴}x_۲^{(K)} + \frac{۱}{۴}\left(۳ - x_۱^{(K)} - x_۳^{(K-۱)}\right) \\ x_۳^{(K)} = \frac{۳}{۴}x_۳^{(K)} + \frac{۱}{۴}\left(۵ - ۲x_۱^{(K)} - ۲x_۲^{(K)}\right) \end{cases} \Rightarrow X^{(۱)} = \begin{bmatrix} \frac{۱}{۴} \\ \frac{۱۱}{۱۶} \\ \frac{۲۵}{۳۲} \end{bmatrix} = \begin{bmatrix} ۰/۲۵ \\ ۰/۶۸۷۵ \\ ۰/۷۸۱۲۵ \end{bmatrix}$$

$$\Rightarrow X^{(۲)} = \begin{bmatrix} \frac{۳۷}{۱۲۸} \\ \frac{۵۱۱}{۵۱۲} \\ \frac{۱۲۲۱}{۱۰۲۴} \end{bmatrix} = \begin{bmatrix} ۰/۲۸۹۰۶ \\ ۰/۹۹۸۰۴ \\ ۱/۱۹۲۳۸ \end{bmatrix} \Rightarrow X^{(۳)} = \begin{bmatrix} \frac{۴۰۳۳}{۴۰۹۶} \\ \frac{۱۶۳۷۳}{۱۶۳۸۴} \\ \frac{۳۲۷۱۱}{۳۲۷۶۸} \end{bmatrix} = \begin{bmatrix} ۰/۹۸۴۶۱ \\ ۰/۹۹۹۳۲ \\ ۰/۹۹۸۲۶ \end{bmatrix}$$

مثال ۳-۳۹) دستگاه معادلات خطی $\begin{cases} ۴x_۱ + ۳x_۲ = ۲۴ \\ ۳x_۱ + ۴x_۲ - x_۳ = ۳۰ \\ -x_۲ + ۴x_۳ = -۲۴ \end{cases}$ با روش فوق تخفیف متوالی و $\omega = \frac{۵}{۴}$ تا ۷ تکرار تقریبی به صورت زیر خواهد داد (جواب دقیق دستگاه $X = [۳,۴,-۵]^T$ می‌باشد).

حدس اولیه را به صورت $X^{(۰)} = \begin{bmatrix} ۰ \\ ۰ \\ ۰ \end{bmatrix}$ در نظر بگیرید. حال باید فرم کلی معادلات را تشکیل داد:

$$\begin{cases} x_۱^{(K)} = -۰/۲۵x_۱^{(K-۱)} - ۰/۹۳۷۵x_۲^{(K-۱)} + ۷/۵ \\ x_۲^{(K)} = -۰/۹۳۷۵x_۱^{(K)} - ۰/۲۵x_۲^{(K-۱)} + ۰/۳۱۲۵x_۳^{(K-۱)} + ۹/۳۷۵ \\ x_۳^{(K)} = ۰/۳۱۲۵x_۲^{(K)} - ۰/۲۵x_۳^{(K-۱)} - ۷/۵ \end{cases}$$

$$X^{(۱)} = \begin{bmatrix} ۰/۱۳۱۲۵ \\ ۳/۵۱۹۵۳۱۳ \\ -۶/۶۵۰۱۴۶۵ \end{bmatrix} \Rightarrow X^{(۲)} = \begin{bmatrix} ۲/۶۲۳۳۱۴۵ \\ ۳/۹۵۸۵۲۶۶ \\ -۴/۶۰۰۴۲۳۸ \end{bmatrix} \Rightarrow \cdots \Rightarrow X^{(۷)} = \begin{bmatrix} ۳/۰۰۰۰۴۹۸ \\ ۴/۰۰۰۲۵۸۶ \\ -۵/۰۰۰۳۴۸۶ \end{bmatrix}$$

میزان خطای این روش نیز برابر است با:

$$\left\| X - X^{(۷)} \right\|_{\infty} = \left\| \begin{bmatrix} -۰/۰۰۰۰۴۹۸ \\ -۰/۰۰۰۲۵۸۶ \\ -۰/۰۰۰۳۴۸۶ \end{bmatrix} \right\|_{\infty} = ۰/۰۰۰۳۴۸۶$$

۳-۵ کاربرد دستگاه معادلات خطی در علوم داده

همان‌طور که قبلاً نیز ذکر شد، دستگاه معادلات خطی در اکثر زمینه‌های علمی و تجاری حضور دارند. در علوم داده نیز برای حل مسئله کمترین مربعات[۱]و تجزیه مقدار تکین[۲]که در فصل ۴ به طور کامل بررسی شده و همچنین تعیین چندجمله‌ای مشخصه ماتریس، تحلیل مولفه‌های اساسی[۳]و... که به طور کامل در فصل ۵ بررسی شده‌اند، باید از تکنیک‌های حل دستگاه معادلات خطی بهره برد. پس قطعاً می‌توان گفت که یکی از بخش‌های مهم در علوم داده، دستگاه معادلات خطی و روش‌های حل اینگونه دستگاه‌ها می‌باشد. جواب نهایی دستگاه‌ها می‌تواند دقیق یا به طور تقریبی تعیین شود.

رگرسیون خطی نمونه‌ای خاص از دستگاه معادلات خطی است. در دستگاه معادلات خطی به جای کار با اعداد، باید با ماتریس‌ها و بردارها کار کرد. جبر خطی و مخصوصاً مبحث حل دستگاه معادلات خطی، کلید درک محاسبات و مباحث آماری است که در یادگیری ماشین به آن نیاز دارید. درک تکنیک‌های یادگیری ماشین در سطح بردارها و ماتریس‌ها، بسیار به شهود فرد کمک خواهد کرد.

[۱] Least Square

[۲] Singular Value Decomposition (SVD)

[۳] Principal Component Analysis (PCA)

فصل چهارم

مسئله کمترین مربعات

فصل چهارم : مسئله کمترین مربعات

مطالب این فصل از [۲] و [۳] آورده شده است. در فصل گذشته همواره ماتریس ضرایب در دستگاه $AX = b$ به صورت مربعی و وارون‌پذیر در نظر گرفته می‌شد. حال اگر ماتریس A، ماتریسی منفرد و یا غیر مربعی باشد، منظور از پیدا کردن جواب دستگاه $AX = b$ در واقع یافتن مانده دستگاه $r = b - AX$ است که تا حد امکان کوچک باشد.

در ماتریس ضرایب غیر مربعی جدید یعنی $A_{m\times n}$، اگر $m > n$ باشد آنگاه تعداد معادلات از تعداد مجهولات بیشتر خواهد بود و دستگاه را فرا معین[۱] می‌نامند. همچنین اگر $m < n$ باشد، آنگاه تعداد مجهولات از تعداد معادلات بیشتر خواهد بود و دستگاه را زیر معین[۲] می‌نامند.

تذکر ۴-۱) چنانچه A ماتریسی غیر مربعی باشد، آنگاه منظور از یافتن جواب دستگاه $AX = b$، یافتن برداری است مانند X که نرم $b - AX$ حداقل شود. به عبارت دیگر مسئله یافتن جواب دستگاه $AX = b$ معادل با مسئله مینیمم‌سازی مقابل خواهد بود: $\min\limits_X f(X),\ \ f(X) = \|b - AX\|^{۲}$.

اگر در مسئله مینیمم‌سازی از نرم ۲ (نرم اقلیدسی) استفاده شود، جواب مسئله $\min\limits_X \|b - AX\|_{۲}^{۲}$، به جواب کمترین مربعات دستگاه $AX = b$ موسوم است. چنانچه از نرم ماتریس قطری D با درایه‌های مثبت که به صورت $\|Y\|_D^{۲} = Y^TDY$ تعریف می‌شود، استفاده گردد، جواب مسئله $\min\limits_X \|b - AX\|_D^{۲}$ را جواب کمترین مربعات وزن‌دار دستگاه $AX = b$ می‌خوانند. تنها زمانی از این نرم استفاده خواهد شد که اهمیت یک یا چند معادله از سایر معادلات دستگاه بیشتر باشد. به عبارت دیگر چنانچه در حل دستگاه $AX = b$ معادلات از اهمیت غیر یکسانی برخوردار باشند، آنگاه از این نرم استفاده خواهد شد. چنانچه در مسئله مینیمم‌سازی مذکور از نرم ماتریس معین مثبت اکید S که به صورت $\|Y\|_S^{۲} = Y^TSY$ تعریف می‌شود، استفاده گردد، جواب مسئله مینیمم‌سازی را جواب کمترین مربعات وزن دار تعمیم یافته می‌خوانند.

تعریف ۴-۲) فرض کنید ماتریس A یک ماتریس $m \times n$ باشد. اگر $m < n$ باشد، ماتریس A را رتبه کامل گویند هرگاه m سطر ماتریس A مستقل خطی باشند ($rank(A) = m$). اگر $m > n$ باشد، ماتریس A را رتبه کامل گویند هرگاه n ستون ماتریس A مستقل خطی باشند ($rank(A) = n$). اگر $m = n$ باشد، ماتریس A را رتبه کامل گویند هرگاه یا سطرهای آن و یا ستون‌های آن مستقل خطی باشند (اگر سطرها مستقل خطی باشند، ستون‌ها نیز مستقل خطی هستند).

تذکر ۴-۳) چنانچه $A \in \mathbb{R}^{m\times n}$ و $m > n$ و A از رتبه کامل باشد، آنگاه دستگاه $AX = b$ و ماتریس افزوده مربوط به آن را به دستگاه زیر تحویل کرد:

$$[A|b] \to \left[\begin{matrix} R_{n\times n} \\ O_{(m-n)\times n}\end{matrix}\middle|\begin{matrix} C \\ d\end{matrix}\right],\ \ C \in \mathbb{R}^{n\times ۱}, d \in \mathbb{R}^{(m-n)\times ۱}$$

---

۱ Over Determined

۲ Under Determined

حال اگر $d \neq \bar{O}_{(m-n)\times 1}$ آنگاه دستگاه $AX = b$ فاقد جواب است. اگر $d = \bar{O}_{(m-n)\times 1}$ آنگاه دستگاه مذکور دارای جواب یکتای $X = R^{-1}C$ خواهد بود.

در ادامه به ارائه روش حل مسئله مینیمم‌سازی $\min_X \|AX - b\|_\eta^2$ که در آن η معرف نرم ۲، D و یا S می‌باشد. برای حل مسئله مینیمم‌سازی $\min_X f(X), \ \ f(X) = \|b - AX\|_{2,D,S}^2$ به قضیه آزمون مشتق دوم برای تابع درجه دوم[1] نیاز است.

قضیه ۴-۴) تابع $f: \mathbb{R}^n \to \mathbb{R}$ با ضابطه $f(X) = X^TMX + b^TX + c$ که در آن $X, b \in \mathbb{R}^{n\times 1}$ می‌باشد و همچنین $M \in \mathbb{R}^{n\times n}$ و $c \in \mathbb{R}$ است؛ دارای مینیمم است هرگاه M معین مثبت اکید باشد و یا دارای ماکزیمم می‌باشد هرگاه M معین منفی اکید باشد (برهان: رجوع شود به [۳])

قضیه۴-۵) اگر $A \in \mathbb{R}^{m\times n}$، $m > n$ و $rank(A) = n$ و ماتریس S یک ماتریس معین مثبت اکید باشد، آنگاه ماتریس‌های A^TA و A^TSA معین مثبت اکید هستند و ماتریس AA^T نیمه معین مثبت خواهد بود.

برهان) چون $A \in \mathbb{R}^{m\times n}$، $m > n$ و A رتبه کامل است؛ در نتیجه اگر $0 \neq Y \in \mathbb{R}^{n\times 1}$ بنابراین $AY \neq 0$ و در نتیجه $\|AY\|_2^2 > 0$. این بدین معنی است که:

$$0 < \|AY\|_2^2 = Y^T(A^TA)Y \Rightarrow \text{معین مثبت اکید است } A^TA$$

به همین ترتیب $\|AY\|_S^2$ و این بدین معنی است که:

$$0 < \|AY\|_S^2 = (AY)^TS(AY) = Y^T(A^TSA)Y \Rightarrow \text{معین مثبت اکید است } A^TSA \ \blacksquare$$

تذکر ۴-۶) فرض کنید $g: \mathbb{R}^n \to \mathbb{R}$ با ضابطه $g(X) = V^TX$ باشد که در آن $X, V \in \mathbb{R}^{n\times 1}$ در این صورت می‌توان نوشت $\nabla_X g(X) = V$ که در آن:

$$\nabla_X g(X) = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \\ \vdots \\ \frac{\partial g}{\partial x_n} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

تذکر ۴-۷) فرض کنید $g: \mathbb{R}^n \to \mathbb{R}$ با ضابطه $g(X) = X^TAX$ باشد که در آن $X \in \mathbb{R}^{n\times 1}$ و $A \in \mathbb{R}^{n\times n}$. در این صورت $\nabla_X g(X) = (A + A^T)X$.

[1] Quadratic Form

با ارائه این مقدمات می‌توان جواب‌های مسئله کمترین مربعات، کمترین مربعات وزن‌دار و کمترین مربعات تعمیم‌یافته را به دست آورد.

قضیه ۴-۸) فرض کنید A یک ماتریس $m \times n$ با $m > n$ باشد. چنانچه $rank(A) = n$ در این صورت جواب مسئله مینیمم‌سازی $\min\limits_X f(X)$ و $f(X) = \|b - AX\|_۲^۲$ معادل است با جواب دستگاه $A^T AX = A^T b$.

برهان) ابتدا باید ضابطه تابع $f(X)$ را به دست آورد. یعنی:

$$f(X) = \|b - AX\|_۲^۲ = (b - AX)^T (b - AX) \Rightarrow$$

$$f(X) = (b^T - X^T A^T)(b - AX) = X^T A^T AX - b^T AX - X^T A^T b + b^T b$$

$$f(X) = (b^T - X^T A^T)(b - AX) = X^T A^T AX - b^T AX - X^T A^T b + b^T b \qquad (۱\text{-}۴)$$

برای محاسبه مینیمم تابع n متغیره $f(X)$ کافیست از این تابع گرادیان گرفته و به صورت زیر مساوی صفر قرار داد:

$$\nabla_X f(X) = (A^T A + A^T A)X - (b^T A)^T - A^T b = ۰ \Rightarrow A^T AX = A^T b$$

از طرفی چون A یک ماتریس $m \times n$ بوده که در آن $m > n$ و $rank(A) = n$، بنابراین $A^T A$ معین مثبت اکید می‌باشد. در نتیجه فرم درجه دوم (۴-۱) در جواب دستگاه $A^T AX = A^T b$ مینیمم دارد. ضمناً چون دستگاه مذکور یکتاست، این مینیمم (جواب مسئله مینیمم‌سازی کمترین مربعات) نیز یکتا است. ■

قضیه ۴-۹) فرض کنید A ماتریسی $m \times n$ با $m > n$ باشد. چنانچه $rank(A) = n$ و S ماتریسی معین مثب اکید و متقارن باشد، در این صورت جواب مسئله مینیمم‌سازی کمترین مربعات وزن‌دار $\min\limits_X f(X)$ که در آن $f(X) = \|b - AX\|_۲^۲$ معادل است با جواب دستگاه $A^T SAX = A^T Sb$.

برهان) ابتدا ضابطه تابع $f(X)$ را باید به دست آورد. یعنی:

$$f(X) = X^T A^T SAX - b^T SAX - X^T A^T Sb + b^T Sb \Rightarrow$$

$$f(X) = (b - AX)^T S(b - AX) \qquad (۲\text{-}۴)$$

برای یافتن مینیمم تابع $f(X)$، باید از آن گرادیان گرفت. یعنی:

$$\nabla_X f(X) = ۲A^T SAX - A^T S^T b - A^T Sb = ۲A^T SAX - ۲A^T Sb = ۰ \Rightarrow$$

$$\nabla_X f(X) = ۰ \Rightarrow A^T SAX = A^T Sb$$

حال چون A ماتریسی $m \times n$ می‌باشد که در آن $m > n$ و $rank(A) = n$. در نتیجه $A^T SA$ معین مثبت اکید است و فرم درجه دوم (۴-۲) در جواب دستگاه $A^T SAX = A^T Sb$ مینیمم خواهد بود. از طرفی چون دستگاه مذکور جواب یکتا دارد، پس مینیمم به دست آمده نیز یکتاست. این بدین معنی است که جواب مسئله مینیمم‌سازی مذکور یکتا است. ■

مثال ۴–۱۰) تقریب کمترین مربعات دستگاه $AX = b$ که در آن $A = \begin{bmatrix} ۰ & ۱ \\ ۲ & -۱ \\ ۲ & ۱ \end{bmatrix}$ و $b = \begin{bmatrix} ۰ \\ ۰ \\ ۴ \end{bmatrix}$ می‌باشد، به صورت زیر به دست خواهد آمد.

طبق مطالب یاد شده جواب کمترین مربعات دستگاه مذکور معادل است با یافتن جواب دستگاه $A^TAX = A^Tb$. بنابراین:

$$A^TA = \begin{bmatrix} ۰ & ۲ & ۲ \\ ۱ & -۱ & ۱ \end{bmatrix}\begin{bmatrix} ۰ & ۱ \\ ۲ & -۱ \\ ۲ & ۱ \end{bmatrix} = \begin{bmatrix} ۸ & ۰ \\ ۰ & ۳ \end{bmatrix}, A^TA = \begin{bmatrix} ۰ & ۲ & ۲ \\ ۱ & -۱ & ۱ \end{bmatrix}\begin{bmatrix} ۰ \\ ۰ \\ ۴ \end{bmatrix} = \begin{bmatrix} ۸ \\ ۴ \end{bmatrix}$$

در نتیجه جواب کمترین مربعات دستگاه اخیر معادل جواب دستگاه $\begin{bmatrix} ۸ & ۰ \\ ۰ & ۳ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۸ \\ ۴ \end{bmatrix}$ خواهد بود. پس حال می‌توان مقادیر $x_۱ = ۱$ و $x_۲ = \frac{۴}{۳}$ را به دست آورد. به ازای این مقادیر $\|r\|_۲ = \|b - AX\|_۲$ به صورت زیر محاسبه خواهد شد:

$$\|r\|_۲ = \|b - AX\|_۲ = \left\| \begin{bmatrix} ۰ \\ ۰ \\ ۴ \end{bmatrix} - \begin{bmatrix} ۰ & ۱ \\ ۲ & -۱ \\ ۲ & ۱ \end{bmatrix}\begin{bmatrix} ۱ \\ \frac{۴}{۳} \end{bmatrix} \right\|_۲ = \left\| \begin{bmatrix} -\frac{۴}{۳} \\ -\frac{۲}{۳} \\ \frac{۲}{۳} \end{bmatrix} \right\|_۲ = \frac{\sqrt{۲۴}}{۳} = \frac{۲\sqrt{۶}}{۳}$$

مثال ۴–۱۱) با مفروضات مثال قبل، چنانچه ارزش (اهمیت) معادله اول دستگاه ۳ برابر معادله سوم و ارزش معادله دوم دستگاه، ۲ برابر معادله سوم باشد؛ در این صورت جواب کمترین مربعات وزن دار متناظر با دستگاه مثال قبل به صورت زیر به دست خواهد آمد.

طبق مطالب گفته شده $D = diag(۳,۲,۱)$ و در نتیجه، جواب کمترین مربعات وزن‌دار دستگاه مذکور معادل با یافتن جواب دستگاه $A^TDAX = A^TDB$ خواهد بود. یعنی:

$$\begin{bmatrix} ۰ & ۲ & ۲ \\ ۱ & -۱ & ۱ \end{bmatrix}\begin{bmatrix} ۳ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}\begin{bmatrix} ۰ & ۱ \\ ۲ & -۱ \\ ۲ & ۱ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۰ & ۲ & ۲ \\ ۱ & -۱ & ۱ \end{bmatrix}\begin{bmatrix} ۳ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}\begin{bmatrix} ۰ \\ ۰ \\ ۴ \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} ۰ & ۴ & ۲ \\ ۳ & -۲ & ۱ \end{bmatrix}\begin{bmatrix} ۰ & ۱ \\ ۲ & -۱ \\ ۲ & ۱ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۰ & ۴ & ۲ \\ ۳ & -۲ & ۱ \end{bmatrix}\begin{bmatrix} ۰ \\ ۰ \\ ۴ \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} ۱۲ & -۲ \\ -۲ & ۶ \end{bmatrix}\begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۸ \\ ۴ \end{bmatrix} \Rightarrow \begin{cases} x_۱ = \frac{۱۴}{۱۷} \\ x_۲ = \frac{۱۶}{۱۷} \end{cases}$$

به ازای این مقادیر $\|r\|_۲ = \|b - AX\|_۲$ برابر خواهد شد با $\frac{۴\sqrt{۶۱}}{۱۷}$.

قضیه ۴-۱۲) فرض کنید $A \in \mathbb{R}^{m\times n}$ و $b \in \mathbb{R}^{m\times 1}$. در این صورت احکام زیر معادل هستند:

۱) دستگاه $AX = b$ جواب کمترین مربعات یکتا دارد.

۲) ستون‌های ماتریس A مستقل خطی هستند. ۳) ماتریس A^TA وارون‌پذیر است.

برهان) رجوع شود به [۳].

تعریف ۴-۱۳) فرض کنید $A \in \mathbb{R}^{m\times n}$. در این صورت ماتریس یکتای $A^+ \in \mathbb{R}^{n\times m}$ که در روابط چهارگانه زیر صدق می‌کند را شبه‌وارون[1]ماتریس A می‌خوانند:

الف) $A^+AA^+ = A^+$ ب) $AA^+A = A$

ج) $(AA^+)^+ = AA^+$ د) $(A^+A)^+ = A^+A$

قضیه ۴-۱۴) با استفاده از تعریف شبه‌وارون یک ماتریس، احکام زیر برقرارند:

۱) چنانچه A ماتریسی مربعی و معکوس‌پذیر باشد؛ در این صورت $A^+ = A^{-1}$.

۲) اگر $A \in \mathbb{R}^{m\times n}$، $m > n$ و $rank(A) = n$. در این صورت روابط موجود در تعریف اخیر برای $(A^TA)^{-1}A^T$ برقرار است. بنابریان به عنوان نمایشی صریح برای شبه وارون ماتریس غیر مربعی A می‌توان نوشت $A^+ = (A^TA)^{-1}A^T$.

۳) اگر $m < n$ و $rank(A) = m$، آنگاه شبه‌وارون A عبارت است از $A^+ = A^T(AA^T)^{-1}$.

۴) از یکتایی موجود در تعریف شبه‌وارون می‌توان فهمید که $(A^+)^+ = A \,, (A^+)^T = (A^T)^+$.

برهان) رجوع شود به [۳].

تذکر ۴-۱۵) با استفاده از تعریف شبه‌وارون می‌توان نوشت که جواب مسئله کمترین مربعات $\min\limits_X \|b - AX\|_2^2$ عبارت است از:

$$\begin{cases} X = A^+b \\ A^+ = (A^TA)^{-1}A^T \end{cases}$$

مثال ۴-۱۶) اگر دستگاه $AX = b$ که در آن $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{bmatrix}$ و $B = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ می‌باشد، حل شود (یعنی جواب کمترین مربعات دستگاه محاسبه شود) متوجه خواهید شد که این دستگاه بی‌شمار جواب دارد. علت این امر، منفرد بودن ماتریس A^TA است. این اتفاق می‌تواند منجر به این مطلب شود که جواب مسئله کمترین مربعات یکتا نباشد.

تذکر ۴-۱۷) در صورت وارون‌پذیر نبودن ماتریس A^TA در دستگاه $AX = b$، آنگاه مسئله جواب کمترین مربعات یکتا نخواهد داشت.

---
[1] Pseudoinverse

حال برای نشان دادن تعبیر هندسی جواب کمترین مربعات دستگاه $AX = b$ می‌توان به صورت زیر عمل کرد: فرض کنید A ماتریسی $m \times n$ باشد که در آن $m > n$ و $rank(A) = n$ ؛ در این صورت برای $m = ۳$ و $n = ۲$ به صورت زیر بازنویسی می‌شوند:

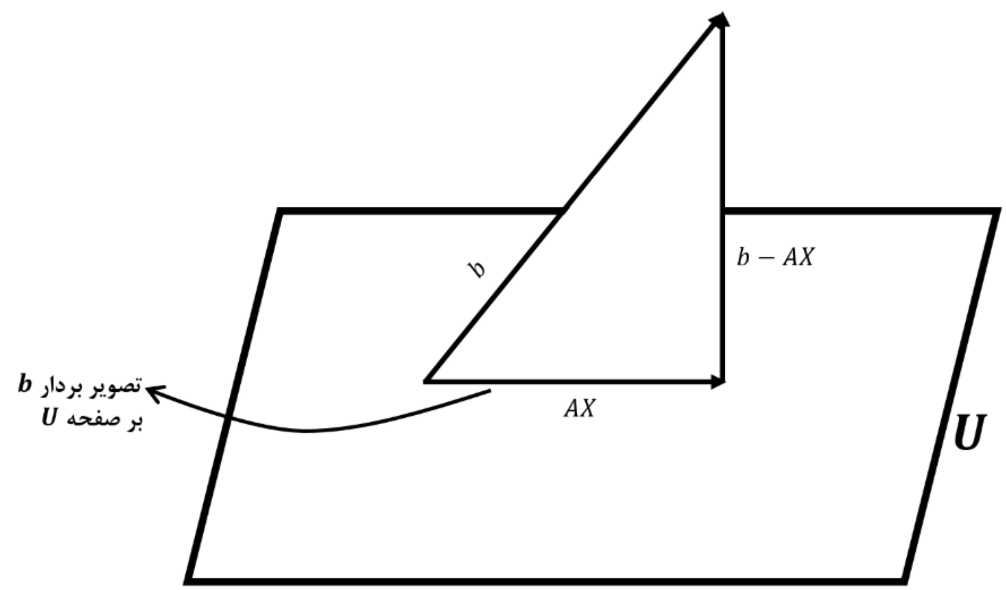
$$b \in \mathbb{R}^{۳\times۱}\ , X \in \mathbb{R}^{۲\times۱}\ , A \in \mathbb{R}^{۳\times۲}\ , \begin{cases} A = [a_۱\ a_۲], a_i \in \mathbb{R}^{۳\times۱} \\ X = \begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} \end{cases}$$

از طرفی چون ستون‌های ماتریس A مستقل خطی هستند، پس اگر فضای تولید شده توسط ستون‌های ماتریس A با U نمایش داده شود، در این صورت:

$$U = span\{a_۱, a_۲\} = c_۱a_۱ + c_۲a_۲$$

از طرفی $AX \in U$ ؛ زیرا:

$$AX = [a_۱\ a_۲]\begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = a_۱x_۱ + a_۲x_۲$$



شکل (۴-۱) تعبیر هندسی $b - AX$

طبق شکل (۴-۱)، $b - AX$ بر فضای U عمود است، بنابراین بر تک تک مولفه‌های سازنده فضا نیز عمود است. یعنی $b - AX$ بر $a_۱$ و $a_۲$ عمود می‌باشد. یعنی:

$$\begin{cases} a_۱ \perp (b - AX) \\ a_۲ \perp (b - AX) \end{cases}$$

طبق تعریف ضرب داخلی، ضرب داخلی آن‌ها صفر خواهد بود. این بدین معنی است که:

$$\left.\begin{array}{l}(a_۱, b - AX) = ۰ \xRightarrow{\text{تعریف ضرب داخلی}} a_۱^T(b - AX) = ۰ \\ (a_۲, b - AX) = ۰ \xRightarrow{\text{تعریف ضرب داخلی}} a_۲^T(b - AX) = ۰\end{array}\right\} A^T(b - AX) = ۰ \Rightarrow A^T AX = A^T b$$

این بدین معنی است که جواب مسئله کمترین مربعات حل دستگاه $AX = b$، معادل حل دستگاه گفته شده خواهد بود. جهت حل دستگاه $A^T AX = A^T b$ می‌توان از دو روش حذفی گاوس و تجزیه چولسکی LL^T استفاده کرد. برای روش حذفی گاوس، باید ابتدا $A^T A$ و $A^T b$ را به دست آورد و سپس دستگاه جدید را با کمک روش حذفی گاوس حل نمود. برای روش تجزیه چولسکی LL^T نیز باید به صورت زیر جواب را استخراج کرد:

$$A^T A = LL^T \Rightarrow LL^T X = A^T b \Rightarrow \begin{cases} LY = A^T b \rightarrow \text{جایگذاری پیشرو} \\ \text{جایگذاری پسرو} \leftarrow L^T X = Y \\ X \text{ به دست می‌آید} \end{cases}$$

تذکر ۴–۱۸) معمولاً از روش تجزیه چولسکی LL^T برای حل دستگاه $A^T AX = A^T b$ استفاده نمی‌شود.

در ادامه روش‌های حل دستگاه $A^T AX = A^T b$ معرفی شده است:

$$\text{حل دستگاه } A^T AX = A^T b \left\{\begin{array}{l} \text{تجزیه متعامد } QR \left\{\begin{array}{l}\text{روش گرام – اشمیت} \\ \text{روش دوران گیونز} \\ \text{روش بازتاب هاوس هُلدر}\end{array}\right. \\ \text{تجزیه مقدار تکین } (SVD)\end{array}\right.$$

۴–۱ تجزیه متعامد[1]

تعریف ۴–۱۹) به تجزیه یک ماتریس نامنفرد و حقیقی A به صورت حاصل‌ضرب یک ماتریس متعامد در یک ماتریس مثلثی که درایه‌های روی قطر اصلی آن مثبت است، تجزیه متعامد گویند.

تذکر ۴–۲۰) چنانچه ماتریس مثلثی از نوع بالامثلثی باشد، به آن تجزیه QR و اگر ماتریس مثلثی از نوع پایین‌مثلثی باشد، به آن تجزیه LQ می‌گویند.

[1] Orthogonal Factorisation

قضیه ۴-۲۱) فرض کنید $A \in \mathbb{R}^{m\times n}$ و $m > n$ و $rank(A) = n$، در این صورت A دارای تجزیه متعامد $A = QR$ که در آن $Q \in \mathbb{R}^{m\times n}$ و Q یک ماتریس متعامد و همچنین $R \in \mathbb{R}^{n\times n}$ که در آن R یک ماتریس بالامثلثی با درایه‌های روی قطر اصلی مثبت می‌باشند (برهان: رجوع شود به [۳]).

حال اگر فرض کنید که $A \in \mathbb{R}^{m\times n}$ و $m > n$ و $rank(A) = n$، در این صورت چنانچه $A = QR$ که در آن $Q_{m\times n}$ یک ماتریس متعامد و همچنین $R_{n\times n}$ یک ماتریس بالامثلثی با درایه‌های روی قطر اصلی مثبت باشند، آنگاه نحوه حل دستگاه $A^TAX = A^Tb$ به روش تجزیه QR به صورت زیر خواهد بود:

$$A^TAX = A^Tb \xRightarrow{A=QR} (QR)^T(QR)X = (QR)^Tb \Rightarrow R^T\underbrace{Q^TQ}_{=I}RX = R^TQ^Tb \Rightarrow$$

$$R^TRX = R^TQ^Tb \xRightarrow{\text{وارون‌پذیر است } R} RX = Q^Tb \tag{۴-۳}$$

پس جواب دستگاه $A^TAX = A^Tb$ به روش تجزیه QR معادل با حل دستگاه $RX = Q^Tb$ می‌باشد.

۴-۱-۱ روش گرام-اشمیت[۱]

الگوریتم متعامد سازی روش گرام-اشمیت بدین صورت است که فرض کنید $A \in \mathbb{R}^{m\times n}$ و $m > n$ و $rank(A) = n$. چنانچه $A = [a_1 \cdots a_n]$ که a_i ها معرف ستون‌های ماتریس A هستند و $Q = [q_1 \cdots q_n]$ که q_i ها نیز معرف ستون‌های ماتریس Q هستند و همچنین $R = [r_{ij}]$. در این صورت:

گام اول) باید v_j را به صورت زیر قرار داد:

$$v_j = a_j - \sum_{k=1}^{j-1}(a_j, q_k)q_k \tag{۴-۴}$$

$$q_j = \frac{v_j}{\|v_j\|_2} \tag{۴-۵}$$

به ازای $j = 1$ از رابطه (۴-۳)، v_1 و از رابطه (۴-۴)، q_1 را باید به دست آورد. سپس روابط (۴-۳) و (۴-۴) به ازای $j = 2$ تا $j = n$ باید تکرار شوند. بدین ترتیب ستون‌های ماتریس Q به دست می‌آید.

گام دوم) حال باید $R = [r_{ij}]$ را قرار داد که درایه‌های ماتریس R به صورت زیر تولید می‌شود:

$$r_{ij} = (a_j, q_i) = a_j^T . q_i \quad , \quad i = 1,2,\cdots,j \ , j = 1,2,\cdots,n \tag{۴-۶}$$

شایان ذکر است که R یک ماتریس بالامثلثی با درایه‌های مثبت روی قطر اصلی می‌باشد. یعنی $r_{ii} > 0$.

---

[۱] Gram-Schmidt Method

مثال ۴-۲۲) تجزیه QR ماتریس $A = \begin{bmatrix} ۰ & ۱ & ۰ \\ ۱ & ۰ & ۱ \\ ۱ & ۱ & ۰ \end{bmatrix}$ با روش گرام-اشمیت به صورت زیر خواهد بود.

گام اول) ابتدا ستون‌های ماتریس Q از روابط (۴-۳) و (۴-۴) را باید محاسبه کرد. اگر $j = ۱$، پس:

$$v_۱ = a_۱ = \begin{bmatrix} ۰ \\ ۱ \\ ۱ \end{bmatrix} \Rightarrow q_۱ = \frac{v_۱}{\|v_۱\|_۲} = \frac{۱}{\sqrt{۲}}\begin{bmatrix} ۰ \\ ۱ \\ ۱ \end{bmatrix} = \begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix}$$

برای $j = ۲$ می‌توان نوشت:

$$v_۲ = a_۲ - (a_۲, q_۱).\, q_۱ = \begin{bmatrix} ۱ \\ ۰ \\ ۱ \end{bmatrix} - \left(\begin{bmatrix} ۱ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} \right) \begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix}$$

$$\Rightarrow v_۲ = \begin{bmatrix} ۱ \\ ۰ \\ ۱ \end{bmatrix} - \frac{۱}{\sqrt{۲}} \begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} = \begin{bmatrix} ۱ \\ -\frac{۱}{۲} \\ \frac{۱}{۲} \end{bmatrix} \Rightarrow q_۲ = \frac{v_۲}{\|v_۲\|_۲} = \sqrt{\frac{۲}{۳}} \begin{bmatrix} ۱ \\ -\frac{۱}{۲} \\ \frac{۱}{۲} \end{bmatrix} = \begin{bmatrix} \frac{۲}{\sqrt{۶}} \\ -\frac{۱}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۶}} \end{bmatrix}$$

برای $j = ۳$ نیز می‌توان نوشت:

$$v_۳ = a_۳ - (a_۳, q_۱)q_۱ - (a_۳, q_۲)q_۲ = \begin{bmatrix} ۰ \\ ۱ \\ ۰ \end{bmatrix} - \frac{\sqrt{۲}}{۲} \begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} + \frac{۱}{\sqrt{۶}} \begin{bmatrix} \frac{۲}{\sqrt{۶}} \\ -\frac{۱}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۶}} \end{bmatrix} = \begin{bmatrix} \frac{۱}{۳} \\ \frac{۱}{۳} \\ -\frac{۱}{۳} \end{bmatrix}$$

$$\Rightarrow q_۳ = \frac{v_۳}{\|v_۳\|_۲} = \sqrt{۳} \begin{bmatrix} \frac{۱}{۳} \\ \frac{۱}{۳} \\ -\frac{۱}{۳} \end{bmatrix} = \begin{bmatrix} \frac{۱}{\sqrt{۳}} \\ \frac{۱}{\sqrt{۳}} \\ -\frac{۱}{\sqrt{۳}} \end{bmatrix}$$

گام دوم) حال باید ماتریس R را محاسبه کرد:

$$j = ۱ \Rightarrow r_{۱۱} = (a_۱, q_۱) = a_۱^T q_۱ = [۰ \quad ۱ \quad ۱]\begin{bmatrix} ۰ \\ \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} = \sqrt{۲}$$

$$j = ۲ \Rightarrow r_{۱۲} = (a_۲, q_۱) = a_۲^T q_۱ = -\frac{۱}{\sqrt{۶}}\ , r_{۲۲} = (a_۲, q_۲) = a_۲^T q_۲ = \sqrt{\frac{۳}{۲}}$$

$$j = ۳ \Rightarrow r_{۱۳} = (a_۳, q_۱) = \frac{۱}{\sqrt{۲}}\ , r_{۲۳} = (a_۳, q_۲) = -\frac{۱}{\sqrt{۶}}\ , r_{۳۳} = (a_۳, q_۳) = \frac{۱}{\sqrt{۳}}$$

بنابراین ماتریس A به صورت زیر تجزیه خواهد شد:

$$A = \begin{bmatrix} ۰ & ۱ & ۰ \\ ۱ & ۰ & ۱ \\ ۱ & ۱ & ۰ \end{bmatrix} = \begin{bmatrix} ۰ & \frac{۲}{\sqrt{۶}} & \frac{۱}{\sqrt{۳}} \\ \frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۶}} & \frac{۱}{\sqrt{۳}} \\ \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۶}} & -\frac{۱}{\sqrt{۳}} \end{bmatrix} \begin{bmatrix} \sqrt{۲} & \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \\ ۰ & \sqrt{\frac{۳}{۲}} & -\frac{۱}{\sqrt{۶}} \\ ۰ & ۰ & \frac{۱}{\sqrt{۳}} \end{bmatrix}$$

مثال ۴-۲۳) جواب کمترین مربعات دستگاه $AX = b$ که در آن $A = \begin{bmatrix} ۰ & ۱ \\ ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}$ و $b = \begin{bmatrix} \frac{۳}{۲} \\ ۱ \\ ۱ \end{bmatrix}$ می‌باشد، با روش گرام-اشمیت به صورت زیر به دست می‌آید.

ابتدا همانند مثال قبل باید تجزیه متعامد ماتریس A به صورت $A = QR$ را پیدا کرد. این تجزیه به صورت زیر خواهد بود:

$$A = \begin{bmatrix} ۰ & ۱ \\ ۱ & ۰ \\ ۱ & ۱ \end{bmatrix} = \begin{bmatrix} ۰ & \frac{۲}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۶}} \end{bmatrix} \begin{bmatrix} \sqrt{۲} & \frac{\sqrt{۲}}{۲} \\ ۰ & \sqrt{\frac{۳}{۲}} \end{bmatrix}$$

پس از تجزیه ماتریس A به دو ماتریس متعامد Q و ماتریس بالامثلثی با درایه‌های روی قطر اصلی مثبت R، حال باید طبق فرمول $RX = Q^T b$، X را محاسبه کرد. یعنی:

$$\begin{bmatrix} \sqrt{۲} & \frac{\sqrt{۲}}{۲} \\ ۰ & \sqrt{\frac{۳}{۲}} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۰ & \frac{۲}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۶}} \end{bmatrix}^T \begin{bmatrix} \frac{۳}{۲} \\ ۱ \\ ۱ \end{bmatrix} \Rightarrow \begin{bmatrix} \sqrt{۲} & \frac{\sqrt{۲}}{۲} \\ ۰ & \sqrt{\frac{۳}{۲}} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} ۰ & \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۲}} \\ \frac{۲}{\sqrt{۶}} & -\frac{۱}{\sqrt{۶}} & \frac{۱}{\sqrt{۶}} \end{bmatrix} \begin{bmatrix} \frac{۳}{۲} \\ ۱ \\ ۱ \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} \sqrt{۲} & \frac{\sqrt{۲}}{۲} \\ ۰ & \sqrt{\frac{۳}{۲}} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} \sqrt{۲} \\ \sqrt{\frac{۳}{۲}} \end{bmatrix} \Rightarrow \begin{cases} x_۱ = \frac{۱}{۲} \\ x_۲ = ۱ \end{cases}$$

به ازای مقادیر $x_۱$ و $x_۲$، مقدار $\|r\|_۲ = \|b - AX\|_۲$ را نیز می‌توان به صورت زیر محاسبه نمود:

$$\|r\|_۲ = \|b - AX\|_۲ = \left\| \begin{bmatrix} \frac{۳}{۲} \\ ۱ \\ ۱ \end{bmatrix} - \begin{bmatrix} ۰ & ۱ \\ ۱ & ۰ \\ ۱ & ۱ \end{bmatrix} \begin{bmatrix} \frac{۱}{۲} \\ ۱ \end{bmatrix} \right\|_۲ = \left\| \begin{bmatrix} \frac{۳}{۲} \\ ۱ \\ ۱ \end{bmatrix} - \begin{bmatrix} \frac{۱}{۲} \\ ۱ \\ \frac{۳}{۲} \end{bmatrix} \right\|_۲ = \left\| \begin{bmatrix} ۱ \\ ۰ \\ -\frac{۱}{۲} \end{bmatrix} \right\|_۲ = \sqrt{\frac{۵}{۴}}$$

۴-۱-۲ روش دوران گیونز[۱]

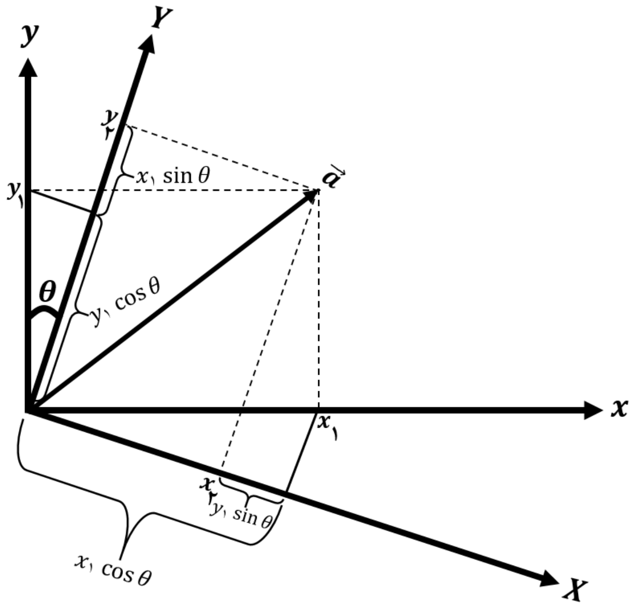
قبل از ارائه روش گیونز نیاز به ارائه یک سری مقدمات می‌باشد.

تعریف ۴-۲۴) فرض کنید بردار $\overrightarrow{a}$ به مختصات $\begin{bmatrix} x_۱ \\ y_۱ \end{bmatrix}$ در صفحه xoy داده شده باشد. چنانچه دستگاه xoy به دستگاه XOY تبدیل شود، مختصات بردار $\overrightarrow{a}$ در دستگاه جدید با $\begin{bmatrix} x_۲ \\ y_۲ \end{bmatrix}$ نمایش داده می‌شود. حال با فرض این که دستگاه جدید از دوران دستگاه قدیم به اندازه θ درجه (در جهت ساعت‌گرد) به دست آمده باشد، ماتریس R_θ را ماتریس دوران می‌خوانند، هرگاه $\begin{bmatrix} x_۲ \\ y_۲ \end{bmatrix} = R_\theta \begin{bmatrix} x_۱ \\ y_۱ \end{bmatrix}$.

طبق شکل (۴-۲) می‌توان گفت:

$$\begin{cases} x_۲ = x_۱ \cos\theta - y_۱ \sin\theta \\ y_۲ = x_۱ \sin\theta - y_۱ \cos\theta \end{cases}$$

[۱] Givens Rotation Method



شکل (۴-۲) ماتریس دوران

بنابراین:

$$\begin{bmatrix} x_۲ \\ y_۲ \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_۱ \\ y_۱ \end{bmatrix}$$

این بدین معنی است که ماتریس دوران R_θ عبارت است از:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \qquad (۷\text{-}۴)$$

که در آن $C = \cos\theta$ و $S = \sin\theta$ می‌باشد. این نحوه نمایش صرفاً جهت سهولت در نوشتار می‌باشد.

تذکر ۴-۲۵) ماتریس دوران R_θ یک ماتریس متعامد است. یعنی $R_\theta . R_\theta^T = I$.

به علت خاصیت تعامد ماتریس دوران R_θ، می‌توان در تجزیه متعامد QR، از این ماتریس‌ها استفاده کرد. فرم کلی ماتریس دوران گیونز به صورت زیر است:

$$G(i,j,\theta) = \begin{array}{c} \begin{bmatrix} ۱ & \cdots & ۰ & \cdots & ۰ & \cdots & ۰ \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ ۰ & \cdots & C & \cdots & -S & \cdots & ۰ \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ ۰ & \cdots & S & \cdots & C & \cdots & ۰ \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ ۰ & \cdots & ۰ & \cdots & ۰ & \cdots & ۱ \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \mapsto i\text{ام} \\ \\ \mapsto j\text{ام} \\ \\ \\ \end{array}$$

به اختصار می‌توان درایه‌های ناصفر $G(i,j,\theta)$ را به صورت زیر نمایش داد:

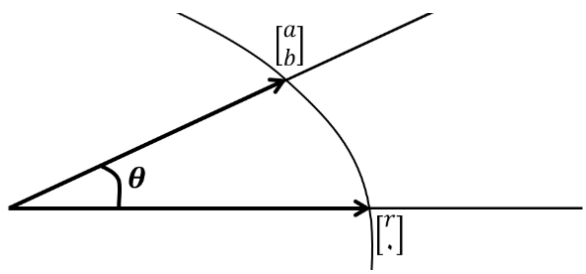
$$G(i,j,\theta) = (g_{ij}) \Rightarrow \begin{cases} g_{kk} = ۱\ , & k \neq i,j \\ g_{kk} = C\ , & k = i,j \\ g_{ij} = -g_{ji} = -S \end{cases}$$

با ضرب ماتریس $G(i,j,\theta)$ از سمت چپ در ماتریس دلخواه A، فقط سطرهای i ام و j ام ماتریس A تغییر خواهد کرد. روش کار ماتریس دوران گیونز $G(i,j,\theta)$ به این صورت است که در هر بار استفاده از این ماتریس، تنها یک درایه از ماتریس A صفر می‌شود. حال اگر هدف تبدیل $\begin{bmatrix} a \\ b \end{bmatrix}$ به $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ باشد، با کمک ماتریس R_θ می‌توان اعمال زیر را انجام داد:

$$R_\theta \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ ۰ \end{bmatrix} \Rightarrow \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ ۰ \end{bmatrix}$$

با حل دستگاه فوق می‌توان به آسانی فهمید که:

$$\begin{cases} C = \dfrac{a}{r} \\ S = -\dfrac{b}{r} \end{cases}, r = \sqrt{a^۲ + b^۲}$$



شکل (۴-۳) تبدیل $\begin{bmatrix} a \\ b \end{bmatrix}$ **به** $\begin{bmatrix} r \\ ۰ \end{bmatrix}$

مثال ۴-۲۶) بردار $\begin{bmatrix} ۳ \\ ۲ \end{bmatrix}$ با استفاده از ماتریس دوران به صورت زیر به $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ تبدیل خواهد شد.

$$r = \sqrt{۳^۲ + ۲^۲} = \sqrt{۹ + ۴} = \sqrt{۱۳}$$

$$\begin{cases} C = \dfrac{۳}{\sqrt{۱۳}} \\ S = -\dfrac{۲}{\sqrt{۱۳}} \end{cases} \Rightarrow R_\theta = \begin{bmatrix} \dfrac{۳}{\sqrt{۱۳}} & \dfrac{۲}{\sqrt{۱۳}} \\ -\dfrac{۲}{\sqrt{۱۳}} & \dfrac{۳}{\sqrt{۱۳}} \end{bmatrix} \begin{bmatrix} ۳ \\ ۲ \end{bmatrix} = \begin{bmatrix} \sqrt{۱۳} \\ ۰ \end{bmatrix}$$

این تبدیل گفته شده برای تجزیه ماتریس A به صورت $A = QR$ کاربرد دارد. فرض کنید با ضرب ماتریس‌های گیونز $G_۱$، $G_۲$، $G_۳$ و... تا G_m در ماتریس A (از سمت چپ)، ماتریس بالامثلثی R ایجاد شود. در این صورت می‌توان نوشت:

$$R = G_m G_{m-۱} \cdots G_۲ G_۱ A \Rightarrow A = (G_m \cdots G_۲ G_۱)^{-۱} R \Rightarrow$$

$$Q = (G_m \cdots G_۲ G_۱)^{-۱} = G_۱^{-۱} G_۲^{-۱} \cdots G_m^{-۱} \qquad (۸\text{-}۴)$$

پس برای صفر کردن عنصر a_{ij} در ماتریس A، بردار $\begin{bmatrix} a_{i-۱\, j} \\ a_{ij} \end{bmatrix}$ در نظر گرفته شده و سپس طبق توضیحات، به بردار $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ تبدیل می‌شود. حال ماتریس دوران گیونز نظیر این تغییر که به صورت $G(i-۱, i, \theta)$ می‌باشد، تولید شده و از سمت چپ در ماتریس A ضرب خواهد شد تا درایه a_{ij} به صفر تبدیل شود. با ادامه همین روند یک ماتریس بالامثلثی ایجاد خواهد شد.

مثال ۴-۲۷) تجزیه QR ماتریس $\begin{bmatrix} ۴ & ۵ & ۰ \\ ۳ & ۱ & ۴ \\ ۰ & ۴ & ۲ \end{bmatrix}$ به کمک روش گیونز به صورت زیر محاسبه می‌شود.

ابتدا باید عنصر $a_{۲۱} = ۳$ صفر شود. بنابراین بردار $\begin{bmatrix} ۴ \\ ۳ \end{bmatrix}$ در نظر گرفته شده و به $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ تبدیل می‌شود. یعنی:

$$\begin{cases} C = \dfrac{۴}{۵} \\ S = -\dfrac{۳}{۵} \end{cases}, r = \sqrt{۳^۲ + ۴^۲} = ۵$$

بنابراین ماتریس گیونز نظیر این مرحله عبارت است از:

$$G(۱,۲,\theta_۱) = \begin{bmatrix} C & -S & ۰ \\ S & C & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} = \begin{bmatrix} \dfrac{۴}{۵} & \dfrac{۳}{۵} & ۰ \\ -\dfrac{۳}{۵} & \dfrac{۴}{۵} & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} := G_۱$$

بنابراین:

$$G_۱ A = \begin{bmatrix} \dfrac{۴}{۵} & \dfrac{۳}{۵} & ۰ \\ -\dfrac{۳}{۵} & \dfrac{۴}{۵} & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} ۴ & ۵ & ۰ \\ ۳ & ۱ & ۴ \\ ۰ & ۴ & ۲ \end{bmatrix} = \begin{bmatrix} ۵ & \dfrac{۲۳}{۵} & \dfrac{۱۲}{۵} \\ ۰ & -\dfrac{۱۱}{۵} & \dfrac{۱۶}{۵} \\ ۰ & ۴ & ۲ \end{bmatrix} := A_۱$$

هدف صفر کردن عنصر $a_{۳۲}^{(۱)} = ۴$ می‌باشد. بنابراین درمورد بردار $\begin{bmatrix} -\frac{۱۱}{۵} \\ ۴ \end{bmatrix}$ روند بالا باید تکرار شود و به بردار $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ تبدیل شود. یعنی:

$$\begin{cases} C = -۰/۴۸۱۹۲ \\ S = -۰/۸۷۶۲۱ \end{cases}, r = \sqrt{(-\frac{۱۱}{۵})^۲ + ۴^۲} = ۴/۵۶۵۱$$

پس ماتریس دوران گیونز نظیر این تغییر عبارت است از:

$$G(۲,۳,\theta_۲) = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & -۰/۴۸۱۹۲ & ۰/۸۷۶۲۱ \\ ۰ & -۰/۸۷۶۲۱ & -۰/۴۸۱۹۲ \end{bmatrix} := G_۲$$

در نتیجه می‌توان نوشت:

$$G_۲ A_۱ = G_۲ G_۱ A = \begin{bmatrix} ۵ & ۴/۶ & ۲/۴ \\ ۰ & ۴/۵۶۵۰۶ & ۰/۲۱۰۲۸ \\ ۰ & ۰ & -۳/۷۶۷۷۱ \end{bmatrix} := A_۲ = R$$

یعنی:

$$R = \overbrace{G_۲ G_۱}^{\text{معکوس‌پذیر}} A \Rightarrow A = G_۱^T G_۲^T R = QR \Rightarrow Q = G_۱^T G_۲^T$$

و ماتریس Q نیز برابر است با:

$$Q = G_۱^T G_۲^T = \begin{bmatrix} ۰/۸ & ۰/۲۸۹۱۵ & ۰/۵۲۵۷۳ \\ ۰/۶ & -۰/۳۸۵۵۴ & -۰/۷۰۰۹۷ \\ ۰ & ۰/۸۷۶۲۱ & -۰/۴۸۱۹۲ \end{bmatrix}$$

مثال ۴-۲۸) جواب کمترین مربعات دستگاه $AX = b$ که در آن $A = \begin{bmatrix} ۱ & ۱ \\ ۰ & ۱ \\ ۱ & ۰ \end{bmatrix}$ و $b = \begin{bmatrix} ۱ \\ \frac{۳}{۲} \\ ۱ \end{bmatrix}$ می‌باشد، با روش گیونز به صورت زیر به دست می‌آید.

ابتدا باید تجزیه QR ماتریس A را پیدا کرد. پس درایه $a_{۳۱}$ در ماتریس $A = \begin{matrix} \\ ۲ \\ ۳ \end{matrix}\begin{bmatrix} ۱ & ۱ \\ \begin{bmatrix} ۰ \\ ۱ \end{bmatrix} & \begin{matrix} ۱ \\ ۰ \end{matrix} \end{bmatrix}$ صفر شود. یعنی بردار $\begin{bmatrix} ۰ \\ ۱ \end{bmatrix}$ باید به بردار $\begin{bmatrix} ۱ \\ ۰ \end{bmatrix}$ تبدیل شود. پس:

$$\begin{bmatrix} ۰ \\ ۱ \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \Rightarrow \begin{cases} C = ۰ \\ S = -۱ \end{cases}, r = ۱ \Rightarrow G(۲,۳,\theta_۱) = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۰ & ۱ \\ ۰ & -۱ & ۰ \end{bmatrix} := G_۱ \Rightarrow$$

$$G_۱A=\begin{bmatrix}۱&۰&۰\\۰&۰&۱\\۰&-۱&۰\end{bmatrix}\begin{bmatrix}۱&۱\\۰&۱\\۱&۰\end{bmatrix}=\begin{array}{c}۱\\۲\\ \\ \end{array}\begin{bmatrix}\begin{bmatrix}۱\\۱\end{bmatrix}&\begin{matrix}۱\\۰\end{matrix}\\ ۰&-۱\end{bmatrix}\coloneqq A_۱$$

حال بردار $\begin{bmatrix}۱\\۱\end{bmatrix}$ باید به $\begin{bmatrix}\sqrt{۲}\\۰\end{bmatrix}$ تبدیل شود. پس:

$$\begin{bmatrix}۱\\۱\end{bmatrix}=\begin{bmatrix}a\\b\end{bmatrix}\Rightarrow\begin{cases}C=\dfrac{\sqrt{۲}}{۲}\\[2ex] S=-\dfrac{\sqrt{۲}}{۲}\end{cases},r=\sqrt{۲}\Rightarrow G(۱,۲,\theta_۲)=\begin{bmatrix}\dfrac{\sqrt{۲}}{۲}&\dfrac{\sqrt{۲}}{۲}&۰\\-\dfrac{\sqrt{۲}}{۲}&\dfrac{\sqrt{۲}}{۲}&۰\\۰&۰&۱\end{bmatrix}\coloneqq G_۲\Rightarrow$$

$$G_۲A_۱=\begin{bmatrix}\dfrac{\sqrt{۲}}{۲}&\dfrac{\sqrt{۲}}{۲}&۰\\-\dfrac{\sqrt{۲}}{۲}&\dfrac{\sqrt{۲}}{۲}&۰\\۰&۰&۱\end{bmatrix}\begin{bmatrix}۱&۱\\۱&۰\\۰&-۱\end{bmatrix}=\begin{array}{c} \\ ۲\\ ۳\end{array}\begin{bmatrix}\sqrt{۲}&\dfrac{\sqrt{۲}}{۲}\\ \begin{matrix}۰\\۰\end{matrix}&\begin{bmatrix}-\dfrac{\sqrt{۲}}{۲}\\-۱\end{bmatrix}\end{bmatrix}\coloneqq A_۲$$

حال باید بردار $\begin{bmatrix}-\frac{\sqrt{۲}}{۲}\\-۱\end{bmatrix}$ به $\begin{bmatrix}\sqrt{\frac{۳}{۲}}\\۰\end{bmatrix}$ تبدیل شود. پس:

$$\begin{cases}C=-\dfrac{\sqrt{۳}}{۳}\\[2ex] S=\sqrt{\dfrac{۲}{۳}}\end{cases},r=\sqrt{\dfrac{۳}{۲}}\Rightarrow G(۲,۳,\theta_۳)=\begin{bmatrix}۱&۰&۰\\۰&-\dfrac{\sqrt{۳}}{۳}&-\sqrt{\dfrac{۳}{۲}}\\۰&\sqrt{\dfrac{۳}{۲}}&-\dfrac{\sqrt{۳}}{۳}\end{bmatrix}\coloneqq G_۲\Rightarrow$$

$$G_۳A_۲=\begin{bmatrix}۱&۰&۰\\۰&-\dfrac{\sqrt{۳}}{۳}&-\sqrt{\dfrac{۳}{۲}}\\۰&\sqrt{\dfrac{۳}{۲}}&-\dfrac{\sqrt{۳}}{۳}\end{bmatrix}\begin{bmatrix}\sqrt{۲}&\dfrac{\sqrt{۲}}{۲}\\۰&-\dfrac{\sqrt{۲}}{۲}\\۰&-۱\end{bmatrix}=\begin{bmatrix}\sqrt{۲}&\dfrac{\sqrt{۲}}{۲}\\۰&\sqrt{\dfrac{۳}{۲}}\\۰&۰\end{bmatrix}\coloneqq R$$

بنابراین:

$$R=G_۳G_۲G_۱A\Rightarrow A=\underbrace{G_۱^TG_۲^TG_۳^T}_{=Q}R\Rightarrow Q=G_۱^TG_۲^TG_۳^T$$

حال می‌توان ماتریس A را به صورت تجزیه QR نوشت. یعنی:

$$A=\begin{bmatrix}\frac{\sqrt{۲}}{۲} & \frac{\sqrt{۶}}{۶} & \frac{\sqrt{۳}}{۳}\\ ۰ & \frac{\sqrt{۶}}{۳} & -\frac{\sqrt{۳}}{۳}\\ \frac{\sqrt{۲}}{۲} & -\frac{\sqrt{۶}}{۶} & -\frac{\sqrt{۳}}{۳}\end{bmatrix}\begin{bmatrix}\sqrt{۲} & \frac{\sqrt{۲}}{۲}\\ ۰ & \sqrt{\frac{۳}{۲}}\\ ۰ & ۰\end{bmatrix}$$

پس از به دست آوردن تجزیه، باید طبق فرمول (۴-۳) یعنی $RX=Q^Tb$ ، ماتریس X را به صورت زیر به دست آورد:

$$\begin{bmatrix}\sqrt{۲} & \frac{\sqrt{۲}}{۲}\\ ۰ & \sqrt{\frac{۳}{۲}}\\ ۰ & ۰\end{bmatrix}\begin{bmatrix}x_۱\\ x_۲\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{۲}}{۲} & ۰ & \frac{\sqrt{۲}}{۲}\\ \frac{\sqrt{۶}}{۶} & \frac{\sqrt{۶}}{۳} & -\frac{\sqrt{۶}}{۶}\\ \frac{\sqrt{۳}}{۳} & -\frac{\sqrt{۳}}{۳} & -\frac{\sqrt{۳}}{۳}\end{bmatrix}\begin{bmatrix}۱\\ \frac{۳}{۲}\\ ۱\end{bmatrix}\Rightarrow$$

$$\begin{bmatrix}\sqrt{۲} & \frac{\sqrt{۲}}{۲}\\ ۰ & \sqrt{\frac{۳}{۲}}\\ ۰ & ۰\end{bmatrix}\begin{bmatrix}x_۱\\ x_۲\end{bmatrix}=\begin{bmatrix}\sqrt{۲}\\ \frac{\sqrt{۶}}{۲}\\ -\frac{\sqrt{۳}}{۲}\end{bmatrix}\Rightarrow\begin{cases}x_۱=\frac{۱}{۲}\\ x_۲=۱\end{cases}$$

به ازای مقادیر $x_۱$ و $x_۲$، مقدار $\|r\|_۲=\|b-AX\|_۲$ را نیز می‌توان به صورت زیر محاسبه نمود:

$$\|r\|_۲=\|b-AX\|_۲=\left\|\begin{bmatrix}\frac{۳}{۲}\\ ۱\\ ۱\end{bmatrix}-\begin{bmatrix}۰ & ۱\\ ۱ & ۰\\ ۱ & ۱\end{bmatrix}\begin{bmatrix}\frac{۱}{۲}\\ ۱\end{bmatrix}\right\|_۲\Rightarrow$$

$$\|r\|_۲=\left\|\begin{bmatrix}\frac{۳}{۲}\\ ۱\\ ۱\end{bmatrix}-\begin{bmatrix}\frac{۱}{۲}\\ ۱\\ \frac{۳}{۲}\end{bmatrix}\right\|_۲=\left\|\begin{bmatrix}۱\\ ۰\\ -\frac{۱}{۲}\end{bmatrix}\right\|_۲=\sqrt{\frac{۵}{۴}}$$

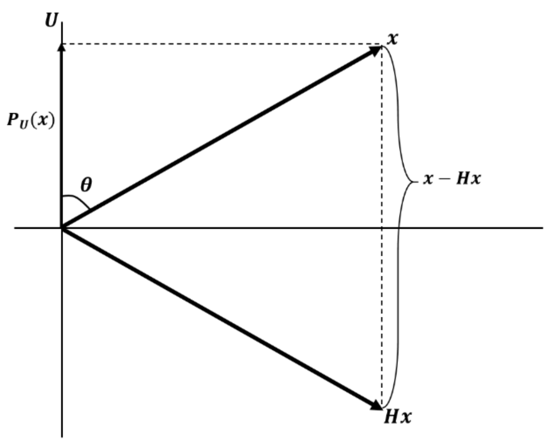
اگر دقت کنید، متوجه خواهید شد که مقادیر $x_۱$ ، $x_۲$ و $\|r\|_۲$ در مثال فوق با مثال (۴-۲۳) یکسان شده است. زیرا ماتریس‌های A و b در مثال فوق، همان ماتریس‌های موجود در مثال (۴-۲۳) می‌باشند، تنها با این تفاوت که سطرهای آن با یک‌دیگر جابه‌جا شده‌اند.

۴-۱-۳ روش بازتاب هاوس هلدر[۱]

قبل از ارائه روش هاوس هلدر نیاز به ارائه یک سری مقدمات می‌باشد.

تعریف ۴-۲۹) بردار $x \in \mathbb{R}^{n\times ۱}$ را فرض کنید. منظور از تبدیل هاوس هلدر یک بردار دلخواه مانند x عبارت است از قرینه این بردار نسبت به صفحه‌ای گذرا از مرکز (نقطه ابتدایی بردار x).

مطابق شکل (۴-۴) تبدیل هاوس هلدر با Hx نمایش داده می‌شود.



شکل (۴-۴) تبدیل هاوس هلدر

طبق شکل (۴-۴) می‌توان تبدیل هاوس هلدر Hx را مساوی با $x - ۲P_U(x)$ در نظر گرفت. از طرفی طبق تعریف تصویر یک بردار بر بردار دیگر می‌توان نوشت:

$$P_U(x) = \frac{UU^T}{U^TU}x \quad , \quad U \in \mathbb{R}^{n\times ۱} , x \in \mathbb{R}^{n\times ۱} , U \neq ۰$$

بنابراین تبدیل هاوس هلدر بردار x عبارت است از:

$$Hx = x - ۲\frac{UU^T}{U^TU}x \quad , \quad U \in \mathbb{R}^{n\times ۱} , x \in \mathbb{R}^{n\times ۱} , U \neq ۰ \qquad (۴-۹)$$

تعریف ۴-۳۰) ماتریسی که قرینه‌ی برداری دلخواه نسبت به صفحه‌ای گذرا از ابتدای این بردار را تولید می‌کند، ماتریس هاوس هلدر می‌نامند.

طبق تعریف (۴-۲۹) چنانچه ماتریس هاوس هلدر با $H \in \mathbb{R}^{n\times n}$ نمایش داده شود، آنگاه:

$$H = I_{n\times n} - ۲\frac{UU^T}{U^TU} \quad , \quad U \in \mathbb{R}^{n\times ۱} , U \neq ۰ \qquad (۴-۱۰)$$

[۱] Householder Reflection Method

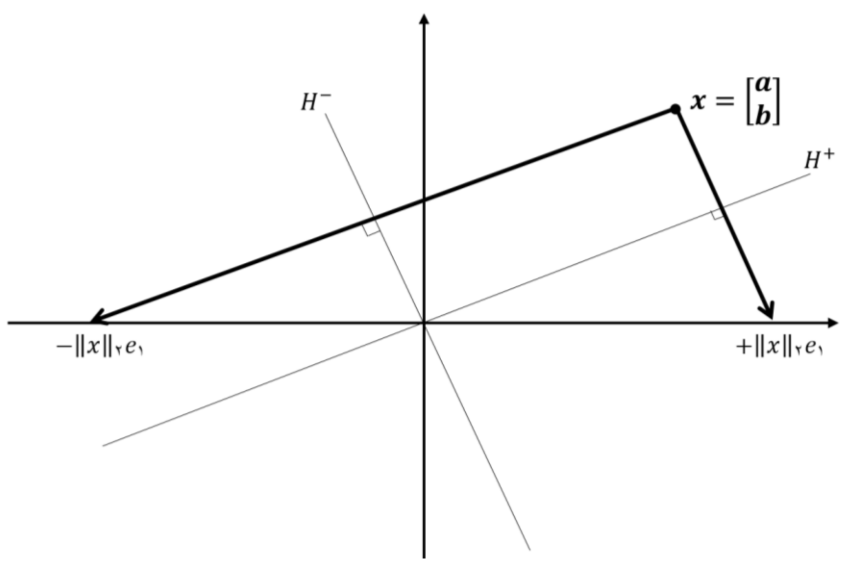
خواص ماتریس مربعی هاوس هلدر H:

۱) $H^T = H$ (ماتریس هاوس هلدر متقارن است).
۲) $H^۲ = I$ (ماتریس هاوس هلدر، متعامد است).
۳) $Hu = -u$.
۴) اگر $(u,v) = ۰$ آنگاه $Hv = v$.
۵) اگر u طوری انتخاب شود که موازی دو بردار $x - y$ و یا $x + y$ باشد به طوری که $\|x\|_۲ = \|y\|_۲$ آنگاه $Hx = y$.

قضیه ۴-۳۱) برای بردار ناصفر $x \neq e_۱$ که $e_۱ = [۱\ ۰\ ۰ \cdots ۰]^T$ چنانچه ماتریس هاوس هلدر با بردار U به صورت $U = x \pm \|x\|_۲ e_۱$ تعریف شود؛ در این صورت $Hx = \mp\|x\|_۲ e_۱ = [\mp\|x\|_۲ e_۱, ۰, \cdots, ۰]^T$.

برهان) طبق رابطه (۵) از خواص ماتریس مربعی هاوس هلدر H، فرض کنید که $y = \pm\|x\|_۲ e_۱$. اولاً چون $x \neq e_۱$ پس به وضوح $y \neq x$ و از طرفی $\|x\|_۲ = \|y\|_۲$. حال چنانچه $U = x - y$ تعریف شود، آنگاه حتماً $Hx = y$. از طرفی $U = x - y = x - (\pm\|x\|_۲ e_۱) = x \mp \|x\|_۲ e_۱$ و حکم ثابت می‌شود. ■

روش کار ماتریس هاوس هلدر برای تبدیل بردار $\begin{bmatrix} a \\ b \end{bmatrix}$ به $\begin{bmatrix} r \\ ۰ \end{bmatrix}$ بدین صورت است که ابتدا با توجه به شکل (۴-۵) باید بازتاب نقطه x نسبت به صفحات H^- و H^+ به دست آورده شود. سپس کافیست ماتریس هاوس هلدری تولید گردد که عملاً کار انعکاس بردار $x = \begin{bmatrix} a \\ b \end{bmatrix}$ به بردار $\begin{bmatrix} \pm\|x\|_۲ \\ ۰ \end{bmatrix}$ را با توجه به قضیه (۴-۳۱) انجام دهد.



شکل (۴-۵) بازتاب نقطه x نسبت به صفحات H^- و H^+

تذکر ۴-۳۲) به لحاظ عددی مطلوب آن است که U در قضیه (۴-۳۱) به صورت زیر انتخاب شود:

$$U = x + sign\{x_۱\}\|x\|_۲ e_۱ \qquad (۱۱\text{-}۴)$$

تذکر ۴-۳۳) اگر مولفه اول بردار x صفر بود، در فرمول (۴-۱۱)، $sign\ \{x_۱\} = +$ قرار می‌گیرد.

مثال ۴-۳۴) فرض کنید $x = \begin{bmatrix} ۰ \\ ۱ \\ ۱ \end{bmatrix}$ ، در این صورت ماتریس هاوس هلدری که بردار x را به برداری به صورت $\begin{bmatrix} r \\ ۰ \\ ۰ \end{bmatrix}$ تبدیل کند، به صورت زیر محاسبه خواهد شد (یعنی $Hx = \begin{bmatrix} r \\ ۰ \\ ۰ \end{bmatrix}$).

طبق رابطه (۴-۱۱) بردار U به صورت زیر خواهد بود:

$$U = \begin{bmatrix} ۰ \\ ۱ \\ ۱ \end{bmatrix} + \sqrt{۲}\begin{bmatrix} ۱ \\ ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} \sqrt{۲} \\ ۱ \\ ۱ \end{bmatrix}$$

حال با استفاده از این بردار U ، ماتریس هاوس هلدر را می‌توان به صورت زیر به دست آورد:

$$H = I - ۲\frac{UU^T}{U^TU} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} - \frac{۲}{۴}\begin{bmatrix} ۲ & \sqrt{۲} & \sqrt{۲} \\ \sqrt{۲} & ۱ & ۱ \\ \sqrt{۲} & ۱ & ۱ \end{bmatrix} = \begin{bmatrix} ۰ & -\frac{\sqrt{۲}}{۲} & -\frac{\sqrt{۲}}{۲} \\ -\frac{\sqrt{۲}}{۲} & \frac{۱}{۲} & -\frac{۱}{۲} \\ -\frac{\sqrt{۲}}{۲} & -\frac{۱}{۲} & \frac{۱}{۲} \end{bmatrix}$$

بنابراین $Hx = \begin{bmatrix} -\sqrt{۲} \\ ۰ \\ ۰ \end{bmatrix}$.

برای انجام تجزیه QR به کمک تبدیل هاوس هلدر باید به صورت زیر عمل کرد:

فرض کنید A ماتریسی $m \times n$ باشد.

گام اول) اگر $x = \begin{bmatrix} a_{۱۱} \\ a_{۲۱} \\ \vdots \\ a_{m۱} \end{bmatrix}$ و $U_۱ = x + sign\{a_{۱۱}\}\|x\|_۲ e_۱ \in \mathbb{R}^{m\times ۱}$ باشد؛ ماتریس هاوس هلدر $H_۱$ به صورت $\widetilde{H}_۱ = H_۱ = I_m - ۲\frac{U_۱U_۱^T}{U_۱^TU_۱}$ ساخته می‌شود. با ضرب $H_۱$ از سمت چپ در ماتریس A ، درایه‌های زیر عنصر محوری $a_{۱۱}$، همگی صفر می‌شوند. ماتریس جدید $A_۱$ نامگذاری می‌شود؛ یعنی $A_۱ = H_۱A$.

گام دوم) اگر $x=\begin{bmatrix}a_{۲۲}^{(۱)}\\a_{۳۲}^{(۱)}\\\vdots\\a_{m۲}^{(۱)}\end{bmatrix}$ و $U_۲=x+sign\left\{a_{۲۲}^{(۱)}\right\}\|x\|_۲e_۱\in\mathbb{R}^{(m-۱)\times ۱}$ باشد؛ ماتریس جدیدی به صورت $\widetilde{H}_۲=I_{m-۱}-۲\frac{U_۲U_۲^T}{U_۲^TU_۲}$ تعریف می‌شود. حال ماتریس هاوس هلدر $H_۱$ نظیر این گام عبارت است از:

$$H_۲=\begin{bmatrix}۱&O\\O&\widetilde{H}_۲\end{bmatrix}$$

حال با ضرب $H_۲$ از سمت چپ در ماتریس $A_۱$ ، درایه‌های زیر عنصر محوری $a_{۲۲}^{(۱)}$، همگی صفر می‌شوند. ماتریس جدید $A_۲$ نامگذاری می‌شود؛ یعنی $A_۲=H_۲A_۱$.

گام k ام) اگر $x=\begin{bmatrix}a_{kk}^{(k-۱)}\\\vdots\\a_{mk}^{(k-۱)}\end{bmatrix}$ و $U_k=x+sign\left\{a_{kk}^{(k-۱)}\right\}\|x\|_۲e_۱\in\mathbb{R}^{(m-k+۱)\times ۱}$ باشد؛ ماتریس جدیدی به صورت $\widetilde{H}_k=I_{m-k+۱}-۲\frac{U_kU_k^T}{U_k^TU_k}$ تعریف می‌شود. حال ماتریس هاوس هلدر H_k نظیر این گام عبارت است از:

$$H_k=\begin{bmatrix}I_{k-۱}&O\\O&\widetilde{H}_k\end{bmatrix}$$

حال با ضرب H_k از سمت چپ در ماتریس $A_{k-۱}$ ، درایه‌های زیر عنصر محوری $a_{kk}^{(k-۱)}$، همگی صفر می‌شوند. این گام‌ها تا جایی ادامه پیدا می‌کند که حاصل‌ضرب $H_kA_{k-۱}$ یک ماتریس بالا مثلثی شود.

مثال ۴-۳۵) تجزیه QR ماتریس $A=\begin{bmatrix}۱&-۴\\۲&۳\\۲&۲\end{bmatrix}$ با روش هاوس هلدر به صورت زیر خواهد بود.

گام اول) ابتدا $x=\begin{bmatrix}۱\\۲\\۲\end{bmatrix}$ و $U_۱=x+sign\{۱\}\|x\|_۲e_۱=\begin{bmatrix}۴\\۲\\۲\end{bmatrix}$ قرار داده می‌شود. حال با استفاده از $U_۱$ ماتریس $H_۱$ باید به صورت زیر تولید شود:

$$\left.\begin{array}{c}U_۱U_۱^T=\begin{bmatrix}۱۶&۸&۸\\۸&۴&۴\\۸&۴&۴\end{bmatrix}\\U_۱^TU_۱=۲۴\end{array}\right\}\Rightarrow H_۱=I-\frac{۱}{۱۲}U_۱U_۱^T=\begin{bmatrix}-\frac{۱}{۳}&-\frac{۲}{۳}&-\frac{۲}{۳}\\-\frac{۲}{۳}&\frac{۲}{۳}&-\frac{۱}{۳}\\-\frac{۲}{۳}&-\frac{۱}{۳}&\frac{۲}{۳}\end{bmatrix}$$

حال با ضرب $H_۱$ در A ماتریس جدید $A_۱$ به صورت $A_۱=\begin{bmatrix}-۳&-۲\\۰&\begin{bmatrix}۴\\۳\end{bmatrix}\\۰&\end{bmatrix}$ به دست می‌آید.

گام دوم) حال $x = \begin{bmatrix} ۴ \\ ۳ \end{bmatrix}$ و $U_۲ = x + sign\{۴\}\|x\|_۲ e_۱ = \begin{bmatrix} ۹ \\ ۳ \end{bmatrix}$ قرار داده می‌شود. حال با استفاده از $U_۲$ ماتریس $\widetilde{H}_۲$ باید به صورت زیر تولید شود:

$$\widetilde{H}_۲ = I - ۲\frac{U_۲U_۲^T}{U_۲^TU_۲} = \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \end{bmatrix} - \frac{۱}{۴۵}\begin{bmatrix} ۸۱ & ۲۷ \\ ۲۷ & ۹ \end{bmatrix} = \begin{bmatrix} -\frac{۴}{۵} & -\frac{۳}{۵} \\ -\frac{۳}{۵} & \frac{۴}{۵} \end{bmatrix}$$

سپس ماتریس $H_۲$ با کمک $\widetilde{H}_۲$ به صورت زیر تولید می‌شود:

$$H_۲ = \begin{bmatrix} ۱ & ۰ \\ ۰ & \widetilde{H}_۲ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & -\frac{۴}{۵} & -\frac{۳}{۵} \\ ۰ & -\frac{۳}{۵} & \frac{۴}{۵} \end{bmatrix}$$

حال با ضرب $H_۲$ در $A_۱$ ماتریس جدید $A_۲$ به صورت زیر به دست می‌آید:

$$H_۲A_۱ = H_۲H_۱A = \begin{bmatrix} -۳ & -۲ \\ ۰ & -۵ \\ ۰ & ۰ \end{bmatrix} = R$$

بنابراین:

$$Q = (H_۲H_۱)^{-۱} \xRightarrow{H_۱ \text{ و } H_۲ \text{ متعامد و متقارن}} Q = H_۱H_۲ = \begin{bmatrix} -\frac{۱}{۳} & \frac{۱۴}{۱۵} & -\frac{۲}{۱۵} \\ -\frac{۲}{۳} & -\frac{۱}{۳} & -\frac{۲}{۳} \\ -\frac{۲}{۳} & -\frac{۲}{۱۵} & \frac{۱۱}{۱۵} \end{bmatrix}$$

۴-۲ تجزیه مقدار تکین

اگر ماتریس A یک ماتریس $m \times n$ باشد، آنگاه ماتریس A^TA معین مثبت است؛ زیرا اگر $X \neq ۰$ و $X \in \mathbb{R}^{n\times ۱}$ آنگاه:

$$X^TA^TAX = (AX)^TAX = \|AX\|_۲^۲ \geq ۰$$

بنابراین کلیه مقادیر ویژه ماتریس A^TA نامنفی می‌باشند؛ زیرا چنانچه λ مقدار ویژه دلخواهی از A^TA باشد، در این صورت $X \neq ۰$ نظیر این مقدار ویژه موجود است که:

$$A^TAX = \lambda X \xRightarrow{\text{ضرب طرفین در } X^T} X^TA^TAX = \lambda X^TX \Rightarrow \lambda = \frac{\|AX\|_۲^۲}{\|X\|_۲^۲} \geq ۰$$

تعریف ۴-۳۶) فرض کنید $A \in \mathbb{R}^{m\times n}$. منظور از مقادیر تکین (منفرد) ماتریس A عبارت است از ریشه دوم مقادیر ویژه ماتریس A^TA . چنانچه مقادیر ویژه ماتریس A^TA با λ_i ($i = ۱,۲,\cdots,n$) نمایش داده شود به طوری که $\lambda_۱ \geq \lambda_۲ \geq \cdots \geq \lambda_n \geq ۰$ و همچنین $\sigma_i = \sqrt{\lambda_i}$ قرار داده شود، در این صورت به σ_i مقادیر تکین ماتریس A گفته می‌شود.

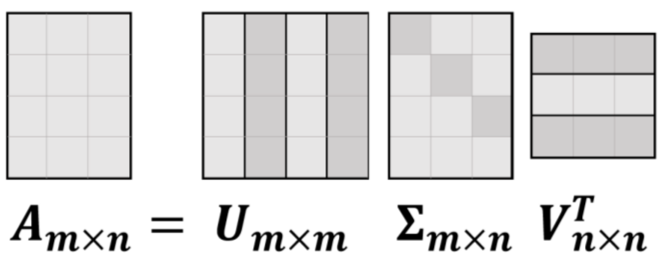
قضیه ۴-۳۷) فرض کنید $A \in \mathbb{R}^{m\times n}$ و $m \geq n$ باشد. در این صورت ماتریس‌های متعامد $U \in \mathbb{R}^{m\times m}$ و $V \in \mathbb{R}^{n\times n}$ و ماتریس قطری $\Sigma \in \mathbb{R}^{m\times n}$ موجودند که:

$$A = U\Sigma V^T \tag{۴-۱۲}$$

و در آن $\Sigma = \begin{bmatrix} D_{r\times r} & \cdot \\ \cdot & \cdot \end{bmatrix}$ و $D_{r\times r} = diag(\sigma_۱,\cdots,\sigma_r)$ که $\sigma_۱,\sigma_۲,\cdots,\sigma_r$ مقدارهای تکین ناصفر ماتریس A می‌باشند.

برهان) رجوع شود به [۳].

در واقع رابطه (۴-۱۲) در صورت نمایش تصویری به صورت شکل (۴-۶) خواهد بود.



شکل (۴-۶) فرم کلی رابطه تجزیه مقادیر تکین

قضیه ۴-۳۸) اگر $\sigma_۱ \geq \sigma_۲ \geq \cdots \geq \sigma_r$ مقادیر تکین (منفرد) ماتریس $A \in \mathbb{R}^{m\times n}$ ($m \geq n$) باشند، در این صورت احکام زیر برقرارند:

۱) $\|A\|_۲ = \sigma_۱ - \max\{\sigma_۱,\cdots,\sigma_r\}$.
۲) $\|A\|_F^۲ = \sigma_۱^۲ + \sigma_۲^۲ + \cdots + \sigma_r^۲$.
۳) اگر A ماتریسی مربعی $n \times n$ و نامنفرد باشد، در این صورت $\|A^{-۱}\|_۲ = \frac{۱}{\sigma_r} = \frac{۱}{\sigma_{min}}$.
۴) اگر A ماتریسی مربعی $n \times n$ و نامنفرد باشد، در این صورت $K_۲(A) = \frac{\sigma_۱}{\sigma_r} = \frac{\sigma_{max}}{\sigma_{min}}$.
۵) رتبه A برابر است با تعداد مقادیر تکین غیر صفر خود.

برهان) رجوع شود به [۳].

تذکر ۴-۳۹) تجزیه مقدار تکین ماتریس A به صورت $A = U\Sigma V^T$ می‌باشد. به بیان دیگر $U^T AV = \Sigma$ که از طرفی:

$$\Sigma^T\Sigma = (V^T A^T U)(U^T AV) = V^T A^T AV = diag(\sigma_1^2, \sigma_2^2, \cdots) \in \mathbb{R}^{n\times n}$$
$$\Sigma\Sigma^T = (U^T AV)(V^T A^T U) = U^T AA^T U = diag(\sigma_1^2, \sigma_2^2, \cdots) \in \mathbb{R}^{m\times m}$$

این بدین معنی است که:

$$\begin{cases} A^T AV = \Sigma^2 V \\ AA^T U = \Sigma^2 U \end{cases}$$

یعنی ستون‌های V، n بردار ویژه $A^T A$ می‌باشند و ستون‌های U، m بردار ویژه AA^T هستند. ستون‌های V را بردارهای تکین راست و ستون‌های U را بردارهای تکین چپ ماتریس A می‌نامند. این بردارها یکتا نیستند ولی مقادیر تکین یکتا هستند.

تذکر ۴-۴۰) فرض کنید A یک ماتریس دلخواه باشد، چون مقادیر ویژه ناصفر ماتریس $A^T A$ و AA^T برابر می‌باشند؛ بنابراین مقادیر تکین ناصفر ماتریس‌های A و A^T نیز برابر خواهند بود.

مثال ۴-۴۱) تجزیه مقدار تکین ماتریس $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ به صورت زیر به دست خواهد آمد.

ابتدا باید $A^T A$ را به دست آورد، سپس مقادیر ویژه آن تولید شود و براساس آن مقدار تکین A را محاسبه کرد:

$$A^T A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

حال می‌توان نوشت:

$$\det(\lambda I - A^T A) = 0 \Rightarrow$$
$$\begin{vmatrix} \lambda - 1 & -1 & 0 \\ -1 & \lambda - 2 & -1 \\ 0 & -1 & \lambda - 1 \end{vmatrix} = (\lambda - 1)\begin{vmatrix} \lambda - 2 & -1 \\ -1 & \lambda - 1 \end{vmatrix} + \begin{vmatrix} -2 & -1 \\ 0 & \lambda - 1 \end{vmatrix} \Rightarrow$$
$$(\lambda - 1)(\lambda^2 - 3\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = 1 \\ \lambda_3 = 0 \end{cases}$$

نحوه نامگذاری مقادیر ویژه باید صعودی باشد؛ یعنی بزرگترین مقدار ویژه λ_1 و کوچکترین مقدار ویژه λ_r نامیده می‌شود. در نتیجه مقادیر تکین ماتریس A عبارت است از:

$$\sigma_1 = \sqrt{3} \ , \ \sigma_2 = 1 \ , \ \sigma_3 = 0$$

حال باید بردار ویژه نظیر λ_1 ، λ_2 و λ_3 را به دست آورد.

$$A^T A V_۱ = \lambda_۱ V_۱ \Rightarrow \begin{bmatrix} ۱ & ۱ & ۰ \\ ۱ & ۲ & ۱ \\ ۰ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} v_۱ \\ v_۲ \\ v_۳ \end{bmatrix} = ۳ \begin{bmatrix} v_۱ \\ v_۲ \\ v_۳ \end{bmatrix} \Rightarrow \begin{cases} v_۱ + v_۲ = ۳v_۱ \\ v_۱ + ۲v_۲ + v_۳ = ۳v_۲ \\ v_۲ + v_۳ = ۳v_۳ \end{cases}$$

$$\Rightarrow \begin{cases} v_۲ = ۲v_۱ \\ v_۲ = ۲v_۳ \end{cases} \Rightarrow v_۱ = v_۳, v_۲ = ۲v_۱ \Rightarrow V_۱ = \begin{bmatrix} ۱ \\ ۲ \\ ۱ \end{bmatrix}$$

به همین ترتیب بردارهای ویژه نظیر $\lambda_۲$ و $\lambda_۳$ عبارت است از $V_۲ = \begin{bmatrix} -۱ \\ ۰ \\ ۱ \end{bmatrix}$ و $V_۳ \begin{bmatrix} ۱ \\ -۱ \\ ۱ \end{bmatrix}$. واضح است که این بردارها متعامدند ولی یکه نیستند. بنابراین پس از یکه‌سازی بردارهای فوق، ماتریس V به دست خواهد آمد:

$$V = [V_۱\ V_۲\ V_۳] = \begin{bmatrix} \frac{۱}{\sqrt{۶}} & -\frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۳}} \\ \frac{۲}{\sqrt{۶}} & ۰ & -\frac{۱}{\sqrt{۳}} \\ \frac{۱}{\sqrt{۶}} & \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۳}} \end{bmatrix}$$

حال باید ماتریس U را تولید کرد. طبق $U = [U_۱\ U_۲]$ و $AV = \Sigma U$ می‌توان برای $i = ۱,۲$ نوشت:

$$\sigma_i U_i = AV_i \Rightarrow U_I = \frac{۱}{\sigma_i} AV_i\ , i = ۱,۲ \Rightarrow$$

$$U_۱ = \frac{۱}{\sqrt{۳}} \begin{bmatrix} ۰ & ۱ & ۱ \\ ۱ & ۱ & ۰ \end{bmatrix} \begin{bmatrix} \frac{۱}{\sqrt{۶}} \\ \frac{۲}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۶}} \end{bmatrix} = \begin{bmatrix} \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} \quad , \quad U_۲ = \frac{۱}{۱} \begin{bmatrix} ۰ & ۱ & ۱ \\ ۱ & ۱ & ۰ \end{bmatrix} \begin{bmatrix} -\frac{۱}{\sqrt{۲}} \\ ۰ \\ \frac{۱}{\sqrt{۲}} \end{bmatrix} = \begin{bmatrix} \frac{۱}{\sqrt{۲}} \\ -\frac{۱}{\sqrt{۲}} \end{bmatrix}$$

$$\Rightarrow U = \begin{bmatrix} \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۲}} \end{bmatrix}$$

در نهایت رابطه $A = U\Sigma V^T$ به صورت زیر خواهد بود:

$$A = \begin{bmatrix} \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۲}} \end{bmatrix} \begin{bmatrix} \sqrt{۳} & ۰ & ۰ \\ ۰ & ۱ & ۰ \end{bmatrix} \begin{bmatrix} \frac{۱}{\sqrt{۶}} & \frac{۲}{\sqrt{۶}} & \frac{۱}{\sqrt{۶}} \\ -\frac{۱}{\sqrt{۲}} & ۰ & \frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۳}} & -\frac{۱}{\sqrt{۳}} & \frac{۱}{\sqrt{۳}} \end{bmatrix} = U\Sigma V^T$$

تذکر ۴-۴۲) در فرایند تجزیه مقادیر تکین، اگر $r = m \leq n$ باشد (همانند مثال قبل) در ساخت بردار U مشکلی ایجاد نخواهد شد. ولی اگر $r < n \leq m$ باشد، جهت ساخت بردار U به روش متعامد سازی گرام-اشمیت نیاز است. همچنین اگر $r = n \leq m$ باشد، با توجه به این‌که $A^T = U\Sigma V^T$ آنگاه $A = V\Sigma^T U^T$ و در نتیجه نیازی به روش متعامد سازی گرام-اشمیت نیست.

مثال ۴-۴۳) تجزیه مقدار تکین ماتریس $A = \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \\ ۱ & ۱ \end{bmatrix}$ به صورت زیر به دست می‌آید.

طبق تذکر (۴-۴۰) مقادیر ویژه A^TA و AA^T یکی هستند؛ بنابراین مقادیر تکین AA^T عبارت است از $\sigma_۱ = \sqrt{۳}$ و $\sigma_۲ = ۱$. در صورتی که تجزیه مقادیر تکین ماتریس A^T به صورت $A^T = U\Sigma V^T$ باشد، آنگاه تجزیه مقادیر تکین A به صورت زیر خواهد بود:

$$A = V\Sigma^T U^T = \begin{bmatrix} \frac{۱}{\sqrt{۶}} & \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۳}} \\ \frac{۱}{\sqrt{۶}} & -\frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۳}} \\ \frac{۲}{\sqrt{۶}} & ۰ & -\frac{۱}{\sqrt{۳}} \end{bmatrix} \begin{bmatrix} \sqrt{۳} & ۰ \\ ۰ & ۱ \\ ۰ & ۰ \end{bmatrix} \begin{bmatrix} \frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۲}} \\ -\frac{۱}{\sqrt{۲}} & \frac{۱}{\sqrt{۲}} \end{bmatrix}$$

در صورتی که این تجزیه به طور مستقیم و با استفاده از روش متعامد سازی گرام-اشمیت به دست آید، می‌توان نوشت:

$$A^TA = \begin{bmatrix} ۲ & ۱ \\ ۱ & ۲ \end{bmatrix} \Rightarrow \det(\lambda I - A^TA) = ۰ \Rightarrow \begin{cases} \lambda_۱ = ۳ \\ \lambda_۲ = ۱ \end{cases}$$

بردارهای ویژه نظیر $\lambda_۱$ و $\lambda_۲$ عبارتند از $V_۱ = \begin{bmatrix} ۱ \\ ۱ \end{bmatrix}$ و $V_۲ = \begin{bmatrix} ۱ \\ -۱ \end{bmatrix}$. از طرفی چون $A = U\Sigma V^T$ بنابراین:

$$U\Sigma = AV$$

حال چنانچه $U = [U_۱ \ U_۲ \ U_۳]$ در نظر گرفته شود، می‌توان نوشت:

$$U_۱ = \frac{۱}{\sigma_۱} AV_۱ = \begin{bmatrix} \frac{۱}{\sqrt{۶}} \\ \frac{۱}{\sqrt{۶}} \\ \frac{۲}{\sqrt{۶}} \end{bmatrix} \quad , \quad U_۲ = \frac{۱}{\sigma_۲} AV_۲ = \begin{bmatrix} -\frac{۱}{\sqrt{۲}} \\ \frac{۱}{\sqrt{۲}} \\ ۰ \end{bmatrix}$$

با انتخاب مجموعه مستقل خطی $\left\{ U_۱, U_۲, \begin{bmatrix} ۱ \\ ۰ \\ ۰ \end{bmatrix} \right\}$ و استفاده از روش گرام-اشمیت، می‌توان $U_۳$ را به دست آورد.

قضیه ۴-۴۴) فرض کنید $A = U\Sigma V^T$ معرف تجزیه مقادیر تکین ماتریس $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ باشد. اگر $V = [V_1\ V_2 \cdots V_n]$ و $U = [U_1\ U_2 \cdots U_n]$ در نظر گرفته شود؛ در این صورت می‌توان نوشت:

$$A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i U_i V_i^T$$

و همچنین $\|A - A_k\|_2 = \sigma_{k+1}$ (برای سهولت می‌توان A_k را به صورت $A_k = U\Sigma_k V^T$ نمایش داد که در آن $\Sigma_k = diag(\sigma_1, \sigma_2, \cdots, \sigma_k, 0, 0, \cdots, 0)$).

برهان) رجوع شود به [۳].

حل مسئله کمترین مربعات به کمک تجزیه مقادیر تکین:

فرض کنید $A \in \mathbb{R}^{m\times n}$ و $m \geq n$ و $rank(A) = n$. در این صورت هدف حل مسئله کمترین مربعات $\min\limits_X \|b - AX\|_2^2$ می‌باشد. در ادامه فرض کنید که تجزیه مقادیر تکین ماتریس A به صورت $A = U\Sigma V^T$ باشد. در این صورت می‌توان نوشت:

$$\|b - AX\|_2^2 = \|b - U\Sigma V^T X\|_2^2 = \|U(U^T b - \Sigma V^T X)\|_2^2$$

طبق پایا بودن ماتریس‌های متعامد نسبت به نرم ۲، می‌توان نوشت:

$$\|b - AX\|_2^2 = \|U^T b - \Sigma V^T X\|_2^2$$

حال چنانچه $V^T X = Y$ و $U^T b = c$ قرار داده شود، که در آن:

$$U^T b = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ c_{n+1} \\ \vdots \\ c_m \end{bmatrix} \quad , \quad U = [\underbrace{U_1}_{\text{ستون اول } n} \quad \overbrace{U_2}^{\text{ستون آخر } m-n}] \quad , \quad \begin{cases} U_1^T b = d \\ U_2^T b = e \end{cases}$$

بنابراین:

$$\|b - AX\|_2^2 = \|c - \Sigma Y\|_2^2 = \|d - \Sigma Y\|_2^2 + \|e\|_2^2$$

واضح است برای کمینه شدن عبارت فوق باید $\Sigma Y = d$ شود که در این حالت میزان خطا عبارت است از $\|e\|_2^2$. با محاسبه Y از دستگاه $\Sigma Y = d$ ، برای پیدا کردن جواب $AX = b$ به روش کمترین مربعات کافیست که بردار X را به صورت $X = VY$ قرار داد (زیرا $V^T X = Y$).

ارتباط تجزیه متعامد و تجزیه مقادیر تکین با شبه معکوس ماتریس A:

۱) اگر تجزیه متعامد ماتریس مستطیلی A به صورت $A = QR$ باشد، در این صورت:

$$A^+ = \Big(\underbrace{R^T Q^T Q R}_{A^T A}\Big)^{-۱} \underbrace{R^T Q^T}_{A^T} = R^{-۱} Q^T$$

۲) اگر $m < n$ و $rank(A) = m$ و تجزیه متعامد ماتریس A^T به صورت $A^T = QR$ باشد؛ در این صورت با توجه به این‌که $(A^+)^T = (A^T)^+$ می‌توان نوشت $A^+ = Q(R^T)^{-۱}$.

۳) برای ماتریس $A \in \mathbb{R}^{m\times n}$ با رتبه دلخواه، چنانچه تجزیه مقادیر تکین آن به صورت $A = U\Sigma V^T$ باشد. در این صورت چنانچه قرار داده شود $\Sigma^T = diag\left(\frac{۱}{\sigma_۱}, \frac{۱}{\sigma_۲}, \cdots, \frac{۱}{\sigma_r}, ۰, \cdots, ۰\right)_{n\times m}$ آنگاه می‌توان نوشت:

$$A^+ = V\Sigma^{-T} U^T$$

۴-۳ کاربرد مسئله کمترین مربعات در علوم داده

یکی از مهم‌ترین کاربردهای مسئله کمترین مربعات در حوزه آمار و تحلیل رگرسیون می‌باشد. به طور دقیق‌تر کاربرد این روش در برازش منحنی بر داده‌ها است. هدف رگرسیون خطی یافتن ارتباطی بین متغیرهای موجود می‌باشد؛ یا به عبارت دیگر هدف ترسیم یک خط مستقیم بر روی داده‌ها می‌باشد که بهترین تناسب را با داده‌ها و کمترین خطا را داشته باشد. البته متخصصین آمار با استفاده از برخی تکنیک‌های آماری، روش کمترین مربعات را بهبود داده و میزان خطا را در رگرسیون خطی کاهش داده اند. ولی نمی‌توان اهمیت به کار گیری جبر خطی در رگرسیون را منکر شد، چرا که ایده روش رگرسیون از جبر خطی یا به طور دقیق‌تر مسئله کمترین مربعات گرفته شده است.

جدول (۴-۱) اجزای مختلف و ارزش ماشین‌های کارکرده

St	Rf	Qs	Pk	
۲۸/۷	۲	۰	۲	۱
۲۴/۸	۳	۱	۲	۲
۲۶/۰	۳	۲	۶	۳
۳۰/۵	۳	۳	۸	۴
۲۳/۸	۴	۳	۱۰	۵
۲۴/۶	۴	۴	۱۲	۶
۲۳/۸	۵	۵	۱۶	۷
۲۰/۴	۵	۶	۱۸	۸
۲۱/۶	۵	۶	۱۸	۹
۲۲/۱	۶	۷	۱۸	۱۰

مثال ۴-۴۵) با توجه به جدول (۴-۱)، ارتباط بین دو متغیر Pk و Qs، با در نظر گرفتن صرفاً ۵ داده اول آن‌ها (سطر ۱ تا ۵) به کمک مسئله کمترین مربعات به صورت زیر به دست خواهد آمد.

ابتدا مسئله $AX = b$ تشکیل داده می‌شود. بدین صورت که A همان ۵ سطر اول متغیر Pk و b همان ۵ سطر اول متغیر Qs می‌باشد. هدف پیدا کردن بردار X می‌باشد که در این قبیل مسائل به صورت ۱×۱ خواهد بود.

$$A = \begin{bmatrix} ۲ \\ ۲ \\ ۶ \\ ۸ \\ ۱۰ \end{bmatrix} \quad , \quad X = [x] \quad , \quad b = \begin{bmatrix} ۰ \\ ۱ \\ ۲ \\ ۳ \\ ۳ \end{bmatrix}$$

حال طبق فرمول $A^TAX = A^Tb$، مقادیر A^TA و A^Tb باید محاسبه شوند.

$$A^TA = [۲ \quad ۲ \quad ۶ \quad ۸ \quad ۱۰]\begin{bmatrix} ۲ \\ ۲ \\ ۶ \\ ۸ \\ ۱۰ \end{bmatrix} = ۲۰۸ \quad , \quad A^Tb = [۲ \quad ۲ \quad ۶ \quad ۸ \quad ۱۰]\begin{bmatrix} ۰ \\ ۱ \\ ۲ \\ ۳ \\ ۳ \end{bmatrix} = ۶۸$$

پس از محاسبه مقادیر A^TA و A^Tb، می‌توان به سادگی مقدار x که همان ارتباط بین دو متغیر است را به صورت زیر محاسبه کرد و ارتباط بین آن‌ها را شرح داد:

$$x = \frac{۶۸}{۲۰۸} = ۰/۳۲۶۹ \Rightarrow ۰/۳۲۶۹A = b \Rightarrow ۰/۳۲۶۹Pk = Qs$$

میزان خطا در روش گفته شده برابر خواهد شد با:

$$\|r\|_۲ = \|b - AX\|_۲ = \left\| \begin{bmatrix} -۰/۶۵۳۸ \\ ۰/۳۴۶۲ \\ ۰/۰۳۸۶ \\ ۰/۳۸۴۸ \\ ۰/۲۶۹۰ \end{bmatrix} \right\|_۲ = ۰/۸۷۷۲$$

ولی در روش رگرسیون باید از فرمول ذکر شده در ادامه استفاده کرد. فرض کنید که متغیر Pk با x و متغیر Qs با y نمایش داده شود. هدف یافتن ارتباط بین x و y است. فرمول رگرسیون خطی به صورت زیر است:

$$\hat{y} = \hat{\beta}_۱ x + \hat{\beta}_۰$$

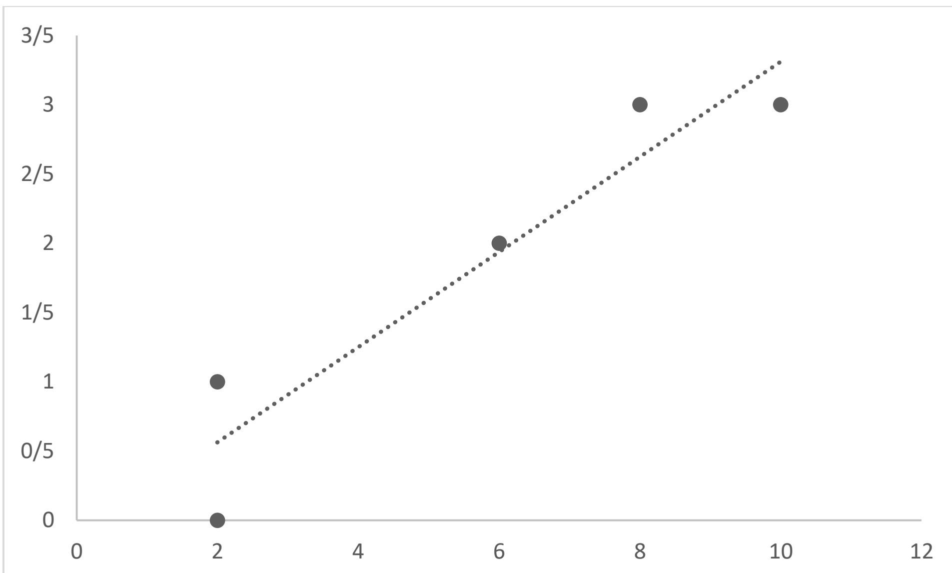
$$\hat{\beta}_۱ = \frac{SS_{xy}}{SS_{xx}} = \frac{\overbrace{\sum xy}^{A^Tb} - \frac{۱}{n}(\sum x)(\sum y)}{\underbrace{\sum x^۲}_{A^TA} - \frac{۱}{n}(\sum x)^۲} \quad , \quad \hat{\beta}_۰ = \bar{y} - \hat{\beta}_۱ \bar{x}$$

همان‌طور که مشاهده می‌شود قسمت‌هایی از فرمول مربوط به رگرسیون خطی دقیقاً از جبر خطی بهره گرفته است. در نهایت، پس از انجام محاسبات لازم رابطه خطی بین دو متغیر x و y به صورت زیر به دست می‌آید:

$$\hat{y} = ۰/۳۴۳۷۵x - ۰/۱۲۵$$

میزان خطا در روش رگرسیون برابر با ۰/۸۶۶۰ خواهد بود. در نمودار (۴-۲) خط برازش بین متغیرهای گفته شده نیز رسم شده است.

نمودار (۴-۲) ارتباط بین متغیرهای Pk و Qs



مثال ۴-۴۶) با توجه به جدول (۴-۱)، ارتباط بین دو متغیر Rf و St، به کمک مسئله کمترین مربعات به صورت زیر به دست خواهد آمد.

ابتدا مسئله $AX = b$ تشکیل داده می‌شود. بدین صورت که A همان متغیر Rf و b همان متغیر St می‌باشد. هدف پیدا کردن بردار X می‌باشد که در این قبیل مسائل به صورت ۱ × ۱ خواهد بود.

$$A = \begin{bmatrix} ۲ \\ ۳ \\ ۳ \\ ۳ \\ ۴ \\ ۴ \\ ۵ \\ ۵ \\ ۵ \\ ۶ \end{bmatrix} \quad , \quad X = [x] \quad , \quad b = \begin{bmatrix} ۲۸/۷ \\ ۲۴/۸ \\ ۲۶ \\ ۳۰/۵ \\ ۲۳/۸ \\ ۲۴/۶ \\ ۲۳/۸ \\ ۲۰/۴ \\ ۲۱/۶ \\ ۲۲/۱ \end{bmatrix}$$

حال طبق فرمول $A^TAX = A^Tb$ ، مقادیر A^TA و A^Tb باید محاسبه شوند.

$$A^TA = A^T\begin{bmatrix} ۲ \\ ۳ \\ ۳ \\ ۳ \\ ۴ \\ ۴ \\ ۵ \\ ۵ \\ ۵ \\ ۶ \end{bmatrix} = ۱۷۴ \quad , \quad A^Tb = A^T\begin{bmatrix} ۲۸/۷ \\ ۲۴/۸ \\ ۲۶ \\ ۳۰/۵ \\ ۲۳/۸ \\ ۲۴/۶ \\ ۲۳/۸ \\ ۲۰/۴ \\ ۲۱/۶ \\ ۲۲/۱ \end{bmatrix} = ۹۵۶/۵$$

پس از محاسبه مقادیر A^TA و A^Tb ، می‌توان به سادگی مقدار x که همان ارتباط بین دو متغیر است را به صورت زیر محاسبه کرد و ارتباط بین آن‌ها را شرح داد:

$$x = \frac{۱۷۴}{۹۵۶/۵} = ۵/۴۹۷۱ \Rightarrow ۵/۴۹۷۱A = b \Rightarrow ۵/۴۹۷۱Rf = St$$

میزان خطا در روش گفته شده برابر خواهد شد با:

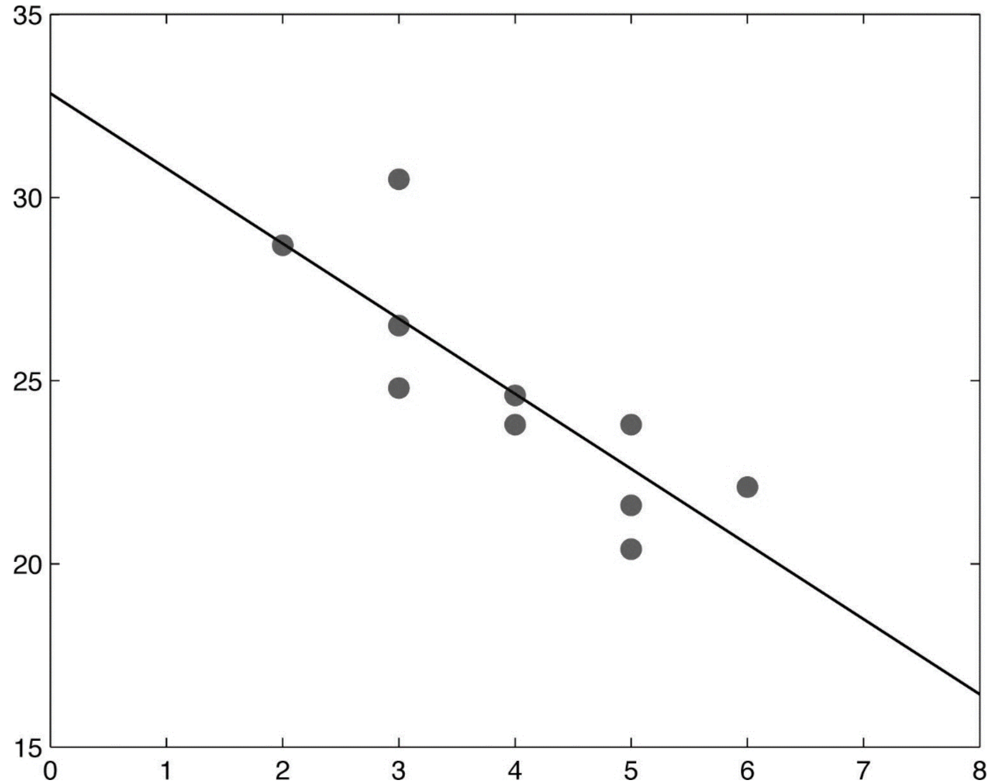
$$\|r\|_۲ = \|b - AX\|_۲ = \left\| \begin{bmatrix} ۱۷/۷۰۵۸ \\ ۸/۳۰۸۷ \\ ۹/۵۰۸۷ \\ ۱۴/۰۰۸۷ \\ ۱/۸۱۱۶ \\ ۲/۶۱۱۶ \\ -۳/۶۸۵۵ \\ -۷/۰۸۵۵ \\ -۵/۸۸۵۵ \\ -۱۰/۸۸۲۶ \end{bmatrix} \right\|_۲ = ۲۹/۹۳۵۷$$

همان‌طور که مشاهده می‌شود میزان خطا به شدت بالاست و عملاً روش کمترین مربعات در این مثال کارساز نیست. به علت این‌که ارتباط بین دو متغیر به صورت معکوس می‌باشد، نه مستقیم. یا به عبارت دیگر رگرسیون این دو متغیر منفی است. حال در روش رگرسیون خط رگرسیونی بین دو متغیر مذکور به صورت زیر خواهد بود:

$$\hat{y} = -۲/۰۵x - ۳۲/۸۳$$

میزان خطا در روش رگرسیون نیز کمتر از ۱ خواهد بود. پس برای رسیدن به جواب بین متغیرهایی که رابطه بین آن‌ها معکوس است، حتما باید از رگرسیون استفاده کرد. ولی برای کاهش محاسبات در متغیرهایی که ارتباط مستقیمی با یک‌دیگر دارند، می‌توان از روش کمترین مربعات استفاده نمود.

نمودار (۴-۳) ارتباط بین متغیرهای **Rf** و **St**



۴-۴ کاربرد تجزیه مقدار تکین در علوم داده

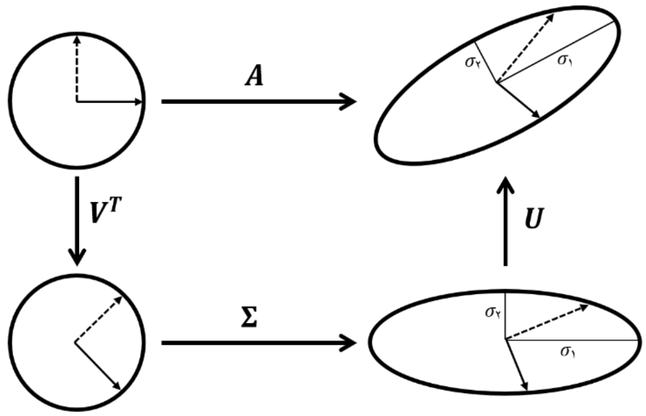
دو روش تجزیه مقادیر تکین و تحلیل مولفه‌های اساسی از مهم‌ترین و پرکاربردترین بخش‌های جبر خطی در علوم داده می‌باشند. با توجه به ارتباط ذکر شده بین روش تجزیه مقدار تکین و مسئله کمترین مربعات، کاربرد این روش در علوم داده در همین فصل توضیح داده خواهد شد و روش تحلیل مولفه‌های اساسی با توجه به ارتباطی که با مقادیر ویژه و بردارهای ویژه دارد، در فصل بعد توضیح داده خواهد شد.

تجزیه مقادیر تکین یا SVD با نام‌های مختلفی شناخته می‌شود. این روش، ابتدا با عنوان تجزیه عامل[1]شناخته می‌شد. همچنین تحلیل مولفه‌های اساسی و تجزیه توابع متعامد تجربی[2]نیز از دیگر اصطلاحات نزدیک به این روش است.

برای درک شهودی تجزیه مقادیر تکین، می‌توان ماتریس A را به عنوان یک تبدیل خطی در نظر گرفت. این تبدیل را نیز می‌توان به سه زیرتبدیل تجزیه کرد: ۱) دوران یا چرخش ۲) تغییر مقیاس ۳) دوران. این سه گام به ترتیب متناظر با سه ماتریس U ، Σ و V هستند. شکل (۴-۷) این موضوع را به خوبی نشان می‌دهد.

[1] Factor Analysis

[2] Empirical Orthogonal Function (EOF)



شکل (۴-۷) نمایش تجزیه مقادیر تکین

دو مورد از اصلی‌ترین کاربردهای تجزیه مقادیر تکین در علوم داده، کاربرد آن در کلان داده[1] و پردازش تصویر می‌باشد. بدین صورت که در حوزه کلان داده، می‌توان با استفاده از این تجزیه ابعاد داده‌ها را کاهش داد و در حوزه پردازش تصویر، می‌توان تصاویر دیجیتالی را فشرده سازی کرد.

۴-۴-۱ فشرده سازی تصاویر دیجیتالی

همان‌طور که در قضیه (۴-۴۴) مشاهده شد، A_k بهترین تقریب برای ماتریس $A \in \mathbb{R}^{m\times n}$ است که رتبه آن k است. حال فرض کنید یک تصویر دیجیتالی با $m \times n$ پیکسل موجود باشد. این تصویر در واقع حکم یک ماتریس $m \times n$ را داراست. هرچقدر کیفیت تصویر بالاتر باشد، m و n افزایش خواهد یافت و در نتیجه ذخیره کردن آن دشوار خواهد شد. در حالی که می‌توان از قضیه (۴-۴۴) بهره برد و یک تقریب بسیار نزدیک به این تصویر پیدا کرد.

طبق فرض گرفته شده، تصویر موجود با $m \times n$ پیکسل، یک ماتریس $m \times n$ مانند A خواهد بود. با توجه به تجزیه $A = U\Sigma V^T$ ، تصویر جدید را می‌توان به صورت ماتریسی به فرم زیر با رتبه k در نظر خواهد گرفت:

$$A_k = \sum_{i=1}^{k} \sigma_i U_i V_i^T$$

که در آن $U = [U_1\ U_2 \cdots U_n]$ و $V = [V_1\ V_2 \cdots V_3]$ می‌باشد. طبق قضیه یاد شده عکس جدید A_k نزدیک-ترین عکس به A در نرم ۲ (یا نرم فروبنیوسی) خواهد بود. حال تنها نکته‌ای که باقی می‌ماند این است که عکس جدید A_k باید بهینه‌ترین حافظه ممکن را در رایانه اشغال کند. یعنی عکس باید کمترین حجم با بهترین کیفیت را داشته باشد.

---

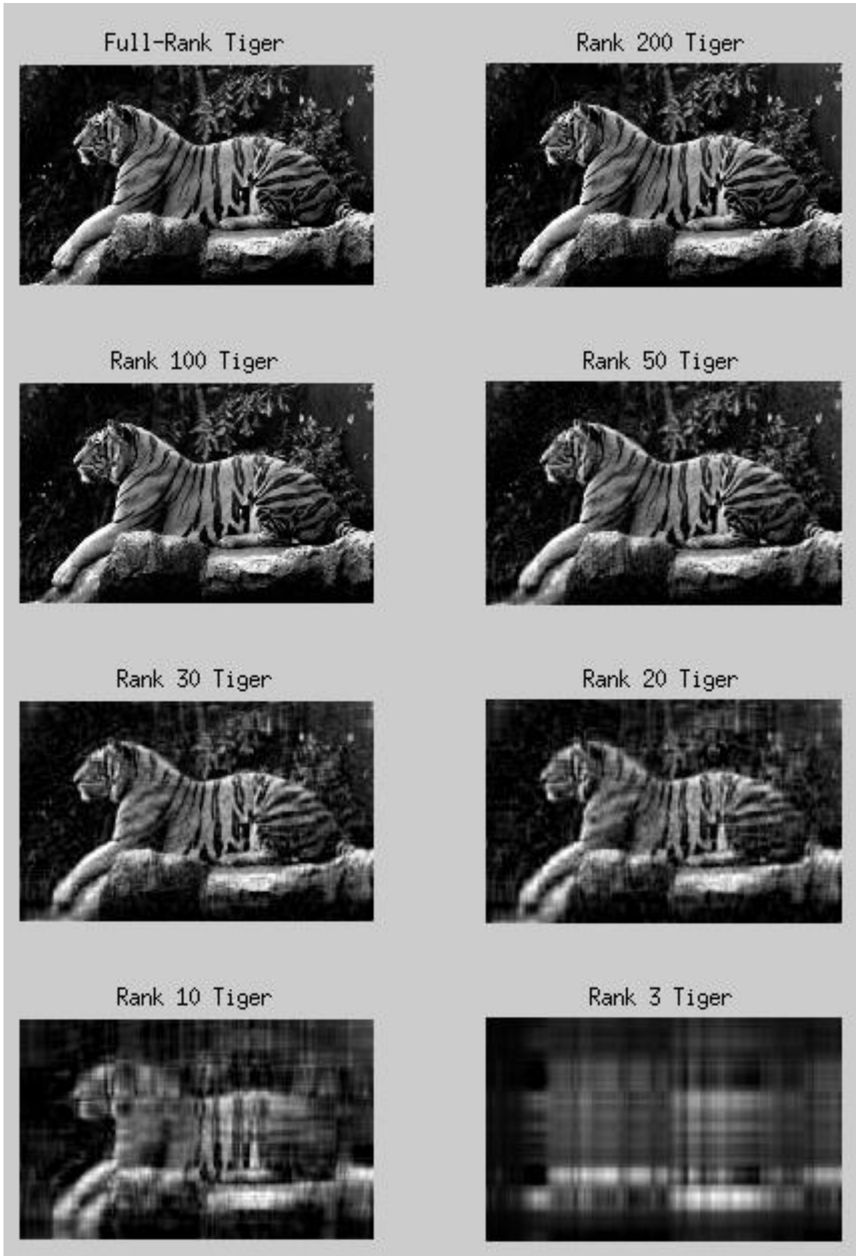
[1] Big Data

یادآوری ۴-۴۷) یک ماتریس $m \times n$ را می‌توان به صورت حاصل‌ضرب یک ستون $m \times ۱$ در یک سطر $۱ \times n$ بازنویسی کرد. با این کار یک ماتریس با mn درایه به صورت $m + n$ درایه ذخیره می‌شود. یعنی:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & \cdots & a_{۲n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m۱} & a_{m۲} & \cdots & a_{mn} \end{bmatrix}_{m\times n} = \begin{bmatrix} b_۱ \\ b_۲ \\ \vdots \\ b_m \end{bmatrix}_{m\times ۱} \begin{bmatrix} c_۱ & c_۲ & \cdots & c_n \end{bmatrix}_{۱\times n}$$

با این مقدمات، برای ذخیره A_k در کامپیوتر، k تا $m + n$ درایه باید ذخیره شود یعنی $k(m + n)$ خانه از حافظه توسط A_k اشغال می‌شود (زیرا $U_i \in \mathbb{R}^{m\times ۱}$ و $V_i \in \mathbb{R}^{n\times ۱}$ می‌باشد؛ پس ضرب U_iV_i ، $m + n$ درایه خواهد داشت). در ادامه با ارائه یک مثال چگونگی ذخیره یک تصویر با حجم کمتر توضیح داده خواهد شد.

شکل (۴-۸) تصویر یک ببر



شکل (۴-۹) تصویر یک ببر در رتبه‌های مختلف

مثال ۴-۴۸) عکس یک ببر یعنی شکل (۴-۸) متناظر با یک ماتریس ۵۰۰ × ۸۰۰ در نظر گرفته شده است. با استفاده از تجزیه مقدار تکین، عکس‌هایی با حجم کمتر و کیفیتی مشابه عکس اصلی همانند شکل (۴-۹) تولید می‌شود که حافظه کمتری از ماشین را درگیر می‌کند. فرایند تجزیه عکس مذکور به صورت زیر خواهد بود.

ابتدا ماتریس $A_{m\times n}$ به صورت $U\Sigma V^T$ تجزیه می‌شود. حال تقریب‌هایی از تصویر اصلی در نظر گرفته می‌شود که به صورت $A_k = \sum_{i=1}^{k} \sigma_i U_i V_i^T$ هستند و در آن σ_i معرف مقادیر تکین ماتریس A می‌باشند و U_i و V_i ستون‌های ماتریس U و V می‌باشند. این تقریب‌ها در واقع تنها $k(m+n)$ مکان از حافظه ماشین را درگیر خواهد کرد.

در این مثال $n = ۵۰۰$ و $m = ۸۰۰$ می‌باشد، پس $mn = ۴۰۰۰۰۰$ درایه وجود دارد. حال عکس‌هایی جدید با $k = ۲۰۰$ ، $k = ۱۰۰$ ، $k = ۵۰$ ، $k = ۳۰$ ، $k = ۲۰$ ، $k = ۱۰$ و $k = ۳$ تولید خواهد شد. در هر حالت می‌توان نوشت:

$k = ۲۰۰ \longrightarrow k(m+n) = ۲۶۰۰۰۰$ تقریباً ۳۵٪ از فضا صرفه‌جویی خواهد شد

$k = ۱۰۰ \longrightarrow k(m+n) = ۱۳۰۰۰۰$ تقریباً ۷۰٪ از فضا صرفه‌جویی خواهد شد

$k = ۵۰ \longrightarrow k(m+n) = ۶۵۰۰۰$ تقریباً ۸۵٪ از فضا صرفه‌جویی خواهد شد

$k = ۳۰ \longrightarrow k(m+n) = ۳۹۰۰۰$ تقریباً ۹۰٪ از فضا صرفه‌جویی خواهد شد

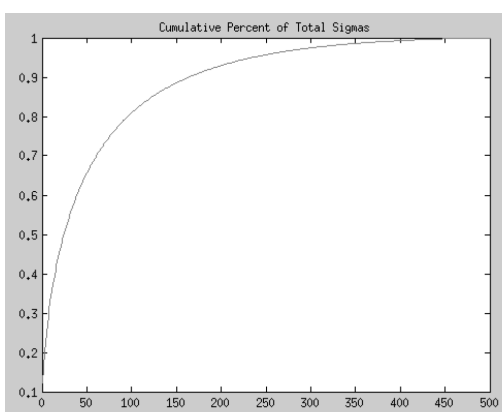
$k = ۲۰ \longrightarrow k(m+n) = ۲۶۰۰۰$ تقریباً ۹۳٪ از فضا صرفه‌جویی خواهد شد

$k = ۱۰ \longrightarrow k(m+n) = ۱۳۰۰۰$ تقریباً ۹۷٪ از فضا صرفه‌جویی خواهد شد

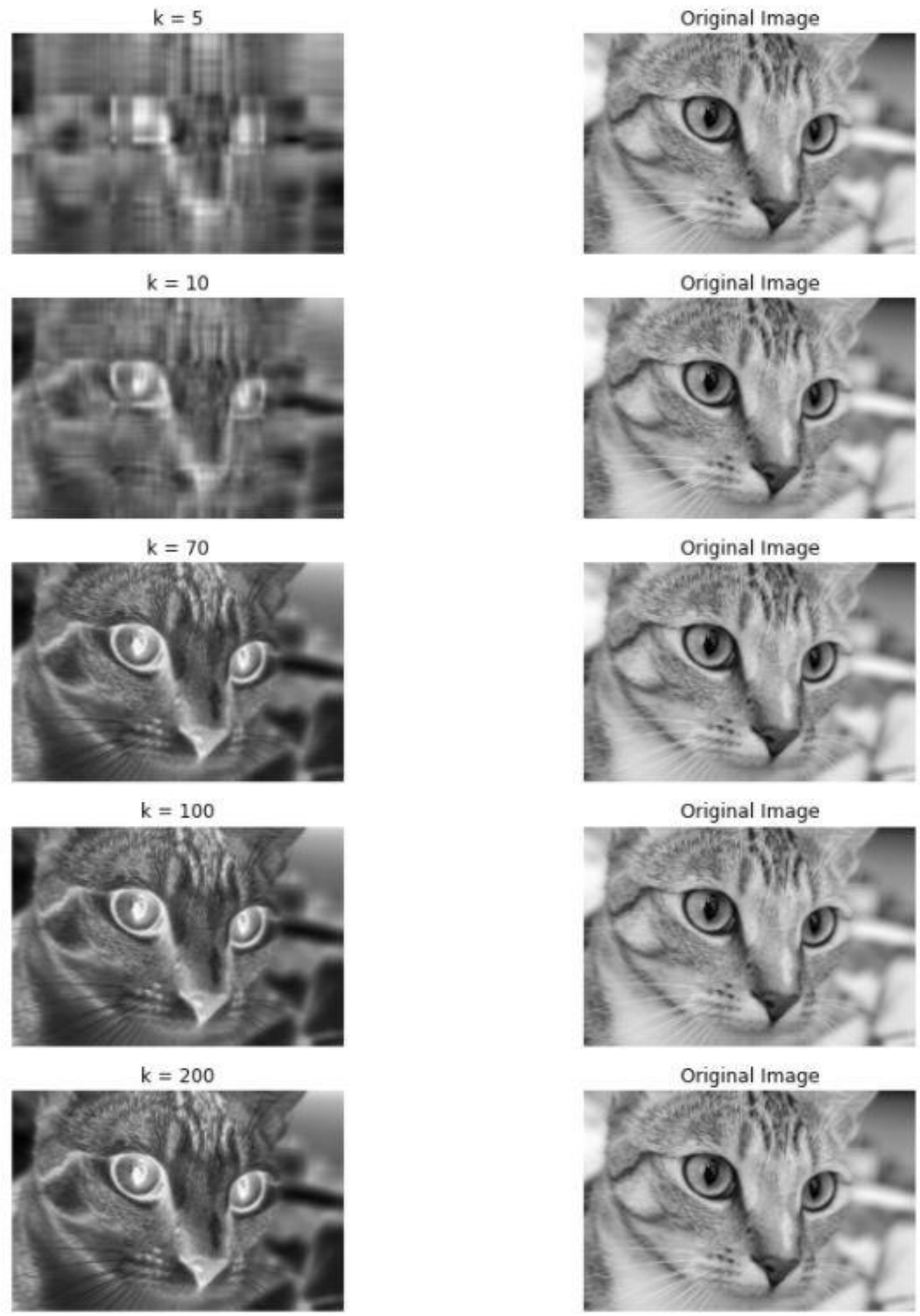
$k = ۳ \longrightarrow k(m+n) = ۳۹۰۰$ تقریباً ۹۹٪ از فضا صرفه‌جویی خواهد شد

همان‌طور که مشاهده می‌شود با هربار کاهش رتبه ماتریس، میزان قابل توجهی از فضای اشغال شده کم خواهد شد و البته از کیفیت عکس‌ها نیز کاسته می‌شود. بهترین تقریب برای عکس فشرده شده رتبه‌های ۳۰ و ۵۰ می‌باشند. نمودار (۴-۴) درصد تجمعی اطلاعات حفظ شده از تصویر ببر در رتبه‌های مختلف را نشان می‌دهد.

نمودار (۴-۴) درصد تجمعی اطلاعات حفظ شده عکس



مثال ۴-۴۹) تصویر (۴-۱۰) مثالی دیگر از کاربرد تجزیه مقادیر تکین در فشرده‌سازی تصویر یک گربه می‌باشد. می‌توان تقریب‌های مختلفی که از عکس گربه در نظر گرفته شده با رتبه‌های مختلف را با یک‌دیگر مقایسه و بررسی کرد.



شکل (۴-۱۰) عکس یک گربه در رتبه‌های مختلف

۴-۴-۲ کاهش ابعاد داده

یکی از چالش‌های متداول در یادگیری ماشین، وجود چند صد متغیر است، در حالی که بسیاری از الگوریتم‌های یادگیری ماشین، اگر با تعدادی بیش از یک مقدار مشخص متغیر کار کنند، با شکست مواجه می‌شوند. این موضوع، استفاده از تجزیه مقادیر تکین را برای کاهش متغیر در یادگیری ماشین ضروری می‌کند.

اگر تعداد و انواع مناسبی از ویژگی‌ها برای حل یک مسئله خاص به الگوریتم‌های یادگیری ماشین داده شود، این الگوریتم‌ها به خوبی کار می‌کنند. اما در صورتی که تعداد ویژگی‌ها (متغیرها) بسیار زیاد باشد، اغلب الگوریتم‌های یادگیری ماشین در حل مسئله دچار مشکل می‌شوند، زیرا با مسئله داده‌های ابعاد بالا[۱]مواجه خواهند بود. در اینجا بحث کاهش ابعاد[۲]مطرح می‌شود.

مثال ۴-۵۰) فرض کنید ماتریس A با ابعاد $m \times n$ ، مجموعه‌ای از داده‌های آموزش باشد که n یعنی طول هر بردار عددی بسیار بزرگ است. هدف وارد کردن ماتریس A به الگوریتم خوشه‌بندی[۳]است که خروجی آن، یک عدد ثابت از مراکز خوشه است. به علت بزرگ بودن n ، اجرای الگوریتم بسیار طول خواهد کشید و یا حتی ممکن است که ناپایدار باشد. بنابراین، باید تعداد متغیرها با استفاده از تجزیه مقادیر تکین کاهش پیدا کند. مراحل زیر، فرایند انجام این کار را نشان می‌دهد.

پس از فراخوانی داده‌ها، باید با کمک تجزیه مقادیر تکین، تقریب خوبی از ماتریس A پیدا کرد. انجام این کار با استفاده از فرمول $A = U\Sigma V^T$ امکان‌پذیر خواهد بود. بدین صورت که ماتریس A با کمک روش‌های گفته شده تجزیه خواهد شد. پس از کاهش ابعاد داده، با کمک الگوریتم خوشه‌بندی، داده‌ها در خوشه‌های مختلف دسته‌بندی شده و سپس مراکز هر خوشه محاسبه می‌شوند.

---
[۱] High Dimensional Data
[۲] Dimensionality Reduction
[۳] Clustering

فصل پنجم

مقدار ویژه و بردار ویژه

فصل پنجم : مقدار ویژه و بردار ویژه

مطالب این فصل از [۳] و [۶] آورده شده است. در فصل اول به‌طور مختصر در مورد مقدار ویژه و بردار ویژه توضیح داده شد. در این فصل ابتدا به معرفی چندجمله‌ای مشخصه و قضیه مهم کیلی-همیلتون پرداخته می‌شود و پس از آن روش‌های تعیین چندجمله‌ای مشخصه و همچنین روش‌های تعیین مقادیر و بردارهای ویژه شرح داده می‌شود.

تعریف ۵-۱) فرض کنید A یک ماتریس $n \times n$ باشد. منظور از چند جمله‌ای مشخصه ماتریس A ، چند جمله‌ای از درجه‌ی n برحسب λ به صورت زیر می‌باشد:

$$P(\lambda) = \det(\lambda I - A) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n \qquad (۵-۱)$$

مقادیر ویژه‌ی ماتریس A همان ریشه‌های $P(\lambda) = 0$ هستند.

قضیه زیر به قضیه کیلی-همیلتون معروف است.

قضیه ۵-۲) هر ماتریس در معادله مشخصه خود صدق می‌کند. یعنی اگر $P(\lambda)$ چند جمله‌ای مشخصه ماتریس A باشد، در این صورت $P(A) = 0$.

برهان) رجوع شود به [۶].

در ادامه روش‌های تعیین چند جمله‌ای مشخصه یک ماتریس و سپس روش‌های محاسبه مقادیر ویژه آورده شده است.

روش‌های تعیین چند جمله‌ای مشخصه یک ماتریس:
- روش بسط دترمینان
- روش کریلوف
- روش لوریير (روش اثر ماتریس)

روش بسط دترمینان بدین صورت است که فرض کنید $A \in \mathbb{R}^{n\times n}$ باشد؛ در این صورت می‌توان چند جمله‌ای مشخصه ماتریس A را از طریق محاسبه دترمینان زیر به دست آورد:

$$P(\lambda) = \det(\lambda I - A)$$

ایراد اساسی این روش آن است که وقتی ابعاد ماتریس A بالا رود، محاسبه دترمینان بسیار هزینه‌بر خواهد بود. بنابراین برای ابعاد بالای ماتریس A ، این روش مطرود است.

در روش کریلوف[۱]، همانند روش بسط دترمینان فرض کنید که A یک ماتریس $n \times n$ باشد و معادله مشخصه آن نیز به صورت رابطه (۵-۱) باشد. چون هر ماتریس در معادله مشخصه خود صدق می‌کند، بنابراین:

$$A^n + p_۱ A^{n-۱} + \cdots + p_{n-۱} A + p_n I_n = ۰ \qquad (۵\text{-}۲)$$

حال بردار دلخواه $Y^{(۰)}$ به صورت زیر تعریف می‌شود و از سمت چپ در رابطه (۵-۲) ضرب خواهد شد:

$$Y^{(۰)} = \begin{bmatrix} y_۱^{(۰)} \\ y_۲^{(۰)} \\ \vdots \\ y_n^{(۰)} \end{bmatrix} \ , \ \ A^n Y^{(۰)} + p_۱ A^{n-۱} Y^{(۰)} + \cdots + p_{n-۱} A Y^{(۰)} + p_n Y^{(۰)} = ۰$$

اگر قرار داده شود $Y^{(K)} = A^K Y^{(۰)}$ در این صورت رابطه بالا عبارت است از:

$$Y^{(n)} + p_۱ Y^{(n-۱)} + \cdots + p_{n-۱} Y^{(۱)} + p_n Y^{(۰)} = ۰ \Rightarrow$$
$$p_۱ Y^{(n-۱)} + \cdots + p_{n-۱} Y^{(۱)} + p_n Y^{(۰)} = -Y^{(n)} \qquad (۵\text{-}۳)$$

فرم ماتریسی رابطه (۵-۳) به صورت زیر خواهد بود:

$$\begin{bmatrix} y_۱^{(n-۱)} & y_۱^{(n-۲)} & \cdots & y_۱^{(۰)} \\ y_۲^{(n-۱)} & y_۲^{(n-۲)} & \cdots & y_۲^{(۰)} \\ \vdots & \vdots & \ddots & \vdots \\ y_n^{(n-۱)} & y_n^{(n-۲)} & \cdots & y_n^{(۰)} \end{bmatrix} \begin{bmatrix} p_۱ \\ p_۲ \\ \vdots \\ p_n \end{bmatrix} = - \begin{bmatrix} y_۱^{(n)} \\ y_۲^{(n)} \\ \vdots \\ y_n^{(n)} \end{bmatrix} \ , \ \ Y^{(K)} = \begin{bmatrix} y_۱^{(K)} \\ y_۲^{(K)} \\ \vdots \\ y_n^{(K)} \end{bmatrix}, K = ۰, ۱, \cdots, n$$

با حل دستگاه بالا (در صورت دارا بودن جواب یکتا) ضرایب معادله مشخصه به دست می‌آید. چنانچه $Y^{(۰)}$ (به طور اتفاقی) بردار ویژه A^K باشد، ماتریس ضرایب دستگاه اخیر منفرد خواهد شد و در نتیجه $Y^{(۰)}$ باید تغییر کند. فرم ساده‌تر دستگاه روش کریلوف عبارت است از:

$$\left[A^{n-۱} Y^{(۰)} \ A^{n-۲} Y^{(۰)} \cdots A Y^{(۰)} \ Y^{(۰)}\right] \begin{bmatrix} p_۱ \\ \vdots \\ p_۲ \end{bmatrix} = -A^n Y^{(۰)} \qquad (۵\text{-}۴)$$

همچنین توجه داشته باشید که برای محاسبه ستون‌های ماتریس ضرایب نیازی به محاسبه توان‌های A نیست؛ فقط کافیست از ضرب متوالی A در بردار $Y^{(۰)}$ استفاده شود. یعنی:

$$A Y^{(۰)} \xrightarrow{\times A} A^۲ Y^{(۰)} \xrightarrow{\times A} A^۳ Y^{(۰)} \xrightarrow{\times A} \cdots$$

---

[۱] Krylov

در روش لوریير[۱] یا همان روش اثر ماتریس برای محاسبه معادله مشخصه ماتریس A باید توجه داشت که اگر $\lambda_۱$ ، $\lambda_۲$ و... تا λ_n مقادیر ویژه ماتریس A باشند، آنگاه:

$$tr(A) = \sum_{i=۱}^{n} \lambda_i$$

حال اگر فرض کنید که معادله مشخصه ماتریس A به صورت زیر باشد:

$$P(\lambda) = \lambda^n - a_۱\lambda^{n-۱} - a_۲\lambda^{n-۲} - \cdots - a_{n-۱}\lambda - a_n$$

در این صورت ضرایب a_i را می‌توان به صورت بازگشتی از روابط زیر تولید کرد:

$$a_i = \frac{۱}{i} tr(A_i) \quad , \quad A_i = \begin{cases} A & i = ۱ \\ AB_{i-۱} & i \geq ۲ \end{cases} \quad , \quad B_i = A_i - a_i I \qquad (۵\text{-}۵)$$

مثال ۵-۳) چند جمله‌ای مشخصه ماتریس $A = \begin{bmatrix} ۳ & ۰ \\ ۸ & -۱ \end{bmatrix}$ به روش بسط دترمینان، روش کریلوف و روش لوریير به صورت زیر محاسبه خواهد شد.

با استفاده از بسط دترمینان می‌توان نوشت:

$$P(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - ۳ & ۰ \\ -۸ & \lambda + ۱ \end{vmatrix} \Rightarrow$$

$$p(\lambda) = \lambda^۲ - ۲\lambda - ۳ = (\lambda - ۳)(\lambda + ۱)$$

حال با کمک روش کریلوف، ابتدا بردار $Y^{(۰)} = \begin{bmatrix} ۱ \\ ۰ \end{bmatrix}$ به صورت دلخواه در نظر گرفته می‌شود؛ در این حالت:

$$AY^{(۰)} = \begin{bmatrix} ۳ & ۰ \\ ۸ & -۱ \end{bmatrix} \begin{bmatrix} ۱ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۳ \\ ۸ \end{bmatrix} \Rightarrow A^۲Y^{(۰)} = A \begin{bmatrix} ۳ \\ ۸ \end{bmatrix} = \begin{bmatrix} ۳ & ۰ \\ ۸ & -۱ \end{bmatrix} \begin{bmatrix} ۳ \\ ۸ \end{bmatrix} = \begin{bmatrix} ۹ \\ ۱۶ \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} AY^{(۰)} & Y^{(۰)} \end{bmatrix} \begin{bmatrix} p_۰ \\ p_۱ \end{bmatrix} = -A^۲Y^{(۰)} \Rightarrow \begin{bmatrix} ۳ & ۱ \\ ۸ & ۰ \end{bmatrix} \begin{bmatrix} p_۰ \\ p_۱ \end{bmatrix} = \begin{bmatrix} -۹ \\ -۱۶ \end{bmatrix} \Rightarrow$$

$$p_۰ = -۲, p_۱ = -۳$$

بنابراین:

$$P(\lambda) = \lambda^۲ + p_۰\lambda + p_۱ = \lambda^۲ - ۲\lambda - ۳$$

---

[۱] Laurier

در روش لوریر، ابتدا قرار داده می‌شود:

$$P(\lambda) = \lambda^{۲} - a_{۱}\lambda - a_{۲} \Rightarrow$$

$$a_{۱} = \frac{۱}{۱} tr(A_{۱}) = tr(A) = ۲ \quad ,$$

$$a_{۲} = \frac{۱}{۲} tr(A_{۲}) = \frac{۱}{۲} tr(AB_{۱}) = \frac{۱}{۲} tr\big(A(A_{۱} - a_{۱}I)\big) = \frac{۱}{۲} tr\big(A(A - ۲I)\big) \Rightarrow$$

$$a_{۲} = \frac{۱}{۲} tr\left(\begin{bmatrix} ۳ & ۰ \\ ۸ & -۱ \end{bmatrix}\begin{bmatrix} ۱ & ۰ \\ ۸ & -۳ \end{bmatrix}\right) = \frac{۱}{۲} tr\left(\begin{bmatrix} ۳ & ۰ \\ ۰ & ۳ \end{bmatrix}\right) = \frac{۶}{۲} = ۳$$

بنابراین می‌توان نوشت:

$$P(\lambda) = \lambda^{۲} - a_{۱}\lambda - a_{۲} = \lambda^{۲} - ۲\lambda - ۳$$

با مقدمات یاد شده، تا کنون می‌توان چند جمله‌ای مشخصه یک ماتریس را استخراج کرد. اما با توجه به این که چند جمله‌ای مشخصه ماتریس A یک چند جمله‌ای از درجه n می‌باشد؛ به ازای $n > ۳$ فرمول تحلیلی برای به دست آوردن ریشه‌های $P(\lambda)$ موجود نیست. بنابراین باید از روش‌های دیگری برای دستیابی به ریشه‌ها استفاده نمود. در ادامه روش‌های یافتن مقادیر ویژه ماتریس $A \in \mathbb{R}^{n\times n}$ ارائه شده است:

- روش‌های محاسبه مقادیر ویژه:
  - روش‌های تکراری
    - روش توانی $\xLeftarrow{\text{تسریع}}$
      - روش ویلکینسون
      - روش توان ماتریسی
    - روش معکوس توانی
    - روش تقلیل توانی
      - روش مستقیم
      - روش هاوس هلدر
  - روش‌های تبدیلی
    - روش ژاکوبی
    - روش گیونز
    - روش هاوس هلدر
    - تجزیه LU
    - تجزیه QR

۵-۱ روش‌های تکراری

در فصل‌های قبل، برخی از روش‌های تکراری آموزش داده شد. حال در این بخش، کاربرد روش‌های تکراری برای مسائلی که در آن نیاز به محاسبه مقدار ویژه می‌باشد، اشاره خواهد شد. در حالت کلی روش‌های تکراری برای یافتن یک یا دو مقدار ویژه مناسب هستند؛ البته روش‌هایی نیز برای تعمیم این روش‌ها جهت یافتن تمام مقادیر ویژه ماتریس وجود دارد. در این بخش فرض بر آن است که ماتریس A حقیقی و از مرتبه n بوده و دارای n بردار ویژه مستقل خطی است.

۵-۱-۱ روش توانی

در این روش، هدف یافتن بزرگترین مقدار ویژه A از نظر قدر مطلق می‌باشد. فرض کنید $\lambda_1, \lambda_2, \cdots, \lambda_n$ مقادیر ویژه‌ی ماتریس $A \in \mathbb{R}^{n\times n}$ باشد و $X^{(1)}, X^{(2)}, \cdots, X^{(n)}$ بردارهای ویژه‌ی نظیر این مقادیر ویژه باشند. دو حالت به وجود می‌آید یا بزرگترین مقدار ویژه A از لحاظ قدر مطلق، حقیقی بوده و تکراری نمی‌باشد و یا این‌که تکراری است. در هر دو حالت روش یافتن مقدار ویژه یکسان است.

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n| \left(\text{بدون تکرار}\right) , |\lambda_1| = \cdots = |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n| \left(\text{تکراری}\right)$$

فرض کنید $V^{(0)} \neq 0$ برداری دلخواه باشد؛ آنگاه طبق فرض مستقل خطی بودن بردارهای ویژه A، ضرایب α_1 ، α_2 و... تا α_n موجودند به طوری که:

$$V^{(0)} = \alpha_1 X^{(1)} + \alpha_2 X^{(2)} + \cdots + \alpha_n X^{(n)} = \sum_{i=1}^{n} \alpha_i X^{(i)} \qquad (۵-۶)$$

حال چنانچه $V^{(k)}$ به صورت زیر تعریف شود:

$$V^{(k)} = A^k V^{(0)}$$

با ضرب طرفین رابطه (۵-۶) در A^k از سمت چپ، می‌توان نوشت:

$$A^k V^{(0)} = \sum_{i=1}^{n} \alpha_i A^k X^{(i)} \ , A^k X^{(i)} = \lambda_i^k X^{(i)}$$

بنابراین:

$$V^{(k)} = \sum_{i=1}^{n} \alpha_i \lambda_i^k X^{(i)}$$

چون طبق فرض $\lambda_۱$ بزرگترین مقدار ویژه A است. پس:

$$V^{(k)} = \lambda_۱^k \left[\alpha_۱ X^{(۱)} + \alpha_۱ \left(\frac{\lambda_۲}{\lambda_۱}\right)^k X^{(۲)} + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_۱}\right)^k X^{(n)}\right] \Rightarrow$$

$$\frac{V^{(k)}}{\lambda_۱^k} = \alpha_۱ X^{(۱)} + \sum_{i=۲}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_۱}\right)^k X^{(i)} \tag{۷-۵}$$

چون طبق فرض $|\lambda_i| > |\lambda_۱|$ ($i = ۲,۳,\cdots,n$) بنابراین با حد گیری از طرفین رابطه (۵-۷) وقتی $k \to \infty$ می‌توان نوشت:

$$\lim_{k\to\infty} \frac{V^{(k)}}{\lambda_۱^k} = \alpha_۱ X^{(۱)}$$

بنابراین هنگامی که k به اندازه کافی بزرگ باشد، تمام جملات سری موجود در رابطه (۵-۷) قابل اغماض خواهند بود. اگر $\alpha_۱ \neq ۰$ آنگاه کسر $\frac{V^{(k)}}{\lambda_۱^k}$ به مضربی از بردار ویژه‌ی نظیر $\lambda_۱$ میل می‌کند و اگر $\alpha_۱ = ۰$ در این صورت $V^{(۰)}$ روی $X^{(۱)}$ مولفه‌ای ندارد، در این صورت دنباله $V^{(k)}$ به $X^{(۱)}$ همگرا نخواهد بود.

حال فرض کنید Y برداری ناصفر و غیر عمود بر بردار $V^{(k)}$ باشد. در این صورت:

$$\lim_{k\to\infty} \frac{Y^T V^{(k+۱)}}{Y^T V^{(k)}} = \lambda_۱$$

حال برای انتخاب بردار Y باید آن را به صورت $Y = e_i$ در نظر گرفت که i متناظر با مکان بزرگترین عضو $V^{(k)}$ از نظر قدر مطلق است و در این حالت Y بر $V^{(k)}$ عمود نخواهد بود.

$$Y = e_i = \begin{bmatrix} ۰ \\ \vdots \\ ۱ \\ \vdots \\ ۰ \end{bmatrix} \mapsto \text{سطر } i \text{ ام}$$

تذکر ۵-۴) در روند محاسبات، جهت جلوگیری از رشد خطا، در هر مرحله بردار $V^{(k)}$ را بر بزرگترین مولفه‌اش (از لحاظ قدر مطلق) باید تقسیم کرد تا آن درایه تبدیل به ۱ یا ۱- شود.

تذکر ۵-۵) برای تمایز در محاسبات، بردارهای نرمال شده (بردارهایی که بزرگترین عضو آن‌ها بر خود بردار تقسیم شده است) را با $U^{(k)}$ نمایش داده خواهد شد.

شرح الگوریتم روش توانی:

۱) بردار $V^{(۰)}$ را انتخاب کرده و نرمال شده $V^{(۰)}$ را باید معادل با $U^{(۰)}$ در نظر گرفت.
۲) حال باید نوشت $V^{(۱)} = AU^{(۰)}$. بزرگترین مولفه بردار $V^{(۱)}$، تقریبی از $\lambda_۱$ خواهد بود ($\lambda_۱ = \lambda_۱^{(۱)}$). در ادامه نرمال شده $V^{(۱)}$ را باید معادل با $U^{(۱)}$ در نظر گرفت.
۳) حال باید نوشت $V^{(۲)} = AU^{(۱)}$. بزرگترین مولفه بردار $V^{(۲)}$، تقریبی از $\lambda_۱$ خواهد بود ($\lambda_۱ = \lambda_۱^{(۲)}$). در ادامه نرمال شده $V^{(۲)}$ را باید معادل با $U^{(۲)}$ در نظر گرفت.
۴) با ادامه همین روند تقریبی از $\lambda_۱$ استخراج خواهد شد ($\lambda_۱$ به دست آمده در گام k ام با $\lambda_۱^{(k)}$ نمایش داده می‌شود).

مثال ۵-۶) بزرگترین مقدار ویژه ماتریس $A = \begin{bmatrix} ۰ & ۱۱ & -۵ \\ -۲ & ۱۷ & -۷ \\ -۴ & ۲۶ & -۱۰ \end{bmatrix}$ با روش توانی به صورت زیر به دست خواهد آمد.

ابتدا برداری دلخواه برای $V^{(۰)}$ در نظر گرفته می‌شود. فرضاً $V^{(۰)} = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \end{bmatrix}$. نرمال شده آن را $U^{(۰)}$ گویند، پس $U^{(۰)} = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \end{bmatrix}$. حال باید قرار داده شود $V^{(۱)} = AU^{(۰)}$. بنابراین $V^{(۱)} = \begin{bmatrix} ۶ \\ ۸ \\ ۱۲ \end{bmatrix}$ خواهد شد و بزرگترین مولفه آن تقریبی از $\lambda_۱$ می‌باشد؛ یعنی $\lambda_۱^{(۱)} = ۱۲$. نرمال شده $V^{(۱)}$ را مساوی با $U^{(۱)}$ باید در نظر گرفت. حال در ادامه $V^{(۲)} = AU^{(۱)}$ می‌باشد، پس:

$$U^{(۱)} = \begin{bmatrix} \frac{۱}{۲} \\ \frac{۱}{۳} \\ ۱ \end{bmatrix} \Rightarrow V^{(۲)} = \begin{bmatrix} \frac{۷}{۳} \\ \frac{۱۰}{۳} \\ \frac{۱۶}{۳} \end{bmatrix} \Rightarrow \lambda_۱^{(۲)} = \frac{۱۶}{۳}$$

$$U^{(۲)} = \begin{bmatrix} \frac{۷}{۱۶} \\ \frac{۵}{۸} \\ ۱ \end{bmatrix} \Rightarrow V^{(۳)} = \begin{bmatrix} \frac{۱۵}{۸} \\ \frac{۱۱}{۴} \\ \frac{۹}{۲} \end{bmatrix} \Rightarrow \lambda_۱^{(۳)} = \frac{۹}{۲}$$

با ادامه همین روند، بعد از ۸ تکرار، ماتریس $V^{(۸)}$ به صورت $\begin{bmatrix} ۱/۶۰۶۹ \\ ۲/۴۰۸۸ \\ ۴/۰۱۲۶ \end{bmatrix}$ خواهد شد و $\lambda_۱^{(۸)} = ۴/۰۱۲۶$.

سرعت همگرایی در روش توانی به $\left|\frac{\lambda_۲}{\lambda_۱}\right|$ وابسته است. هر چه این نسبت کوچکتر باشد، سرعت همگرایی روش توانی نیز بیشتر خواهد بود. در ادامه به دو روش جهت سرعت بخشیدن به روش توانی می‌توان اشاره کرد:

تسریع همگرایی روش توانی: $\begin{cases} \text{روش ویلکینسون} \\ \text{روش توان ماتریسی} \end{cases}$

۵-۱-۱-۱ روش ویلکینسون[۱]

فرض کنید $\lambda_۱$ و $\lambda_۲$ اولین و دومین مقدار ویژه‌ی A از نظر بزرگی باشند، به طوری که $|\lambda_۱| > |\lambda_۲|$ باشد؛ در این صورت $\lambda_۱ - p$ و $\lambda_۲ - p$ مقادیر ویژه ماتریس $A - pI$ می‌باشند. حال چنانچه p طوری انتخاب شود که:

$$|\lambda_۱ - p| > |\lambda_۲ - p| \geq |\lambda_i - p|\,, i = ۳,۴,\cdots,n \qquad (۸-۵)$$

در این حالت سرعت همگرایی روش توانی برای یافتن مقادیر ویژه $A - pI$ به $\frac{\lambda_۲ - p}{\lambda_۱ - p}$ وابسته است. برای حداکثر نمودن سرعت همگرایی روش توانی، باید p را طوری انتخاب کرد که اولاً رابطه (۵-۸) برقرار باشد و ثانیاً $\frac{\lambda_۲ - p}{\lambda_۱ - p}$ مینیمم گردد. بنابراین برای محاسبه بزرگترین مقدار ویژه A (از لحاظ قدر مطلق)، می‌توان روش توانی را برای ماتریس $A - pI$ انجام داد و بعد از تعیین بزرگترین مقدار ویژه $A - pI$ ، این مقدار را با p جمع کرد.

تذکر ۵-۷) از لحاظ تئوری حداکثر سرعت همگرایی روش توانی زمانی حاصل می‌شود که $p = \frac{\lambda_۲ + \lambda_n}{۲}$ باشد.

مثال ۵-۸) سرعت همگرایی روش توانی و روش ویلکینسون برای ماتریس $A = \begin{bmatrix} -۱۳ & ۲۳ & ۱۷ & ۱۲ \\ ۲۳ & -۱۳ & ۱۷ & ۱۳ \\ ۱۷ & ۱۷ & -۷ & ۱۳ \\ ۱۲ & ۱۳ & ۱۳ & ۱ \end{bmatrix}$ به صورت زیر خواهد بود.

با استفاده از روش توانی بر روی ماتریس A، مقدار ویژه $\lambda_۱^{(۵۹)} = ۳۹/۵۰۳۹$ به دست می‌آید؛ یعنی بعد از ۵۹ تکرار به مقدار گفته شده می‌توان رسید. حال چنانچه در روش ویلکینسون $p = -۲۴$ در نظر گرفته شود، باید به دنبال یافتن بزرگترین مقدار ویژه $A - pI = A + ۲۴I$ بود. حال با به کار بردن روش توانی برای $A + ۲۴I$ ، بعد از ۷ تکرار به مقدار $\lambda_۱^{(۷)} = ۶۳/۵۰۳۹$ می‌توان رسید. با کم کردن ۲۴ از این مقدار، بزرگترین مقدار ویژه A ، تولید خواهد شد؛ یعنی:

$$\lambda_۱^{(۷)} - ۲۴ = ۳۹/۵۰۳۹$$

[۱] Wilkinson

۵-۱-۱-۲ روش توان ماتریسی

فرض کنید $|\lambda_۱| > |\lambda_۲| \geq \cdots \geq |\lambda_n|$ مقادیر ویژه ماتریس A باشند، در این صورت می‌توان نوشت:

$$|\lambda_۱^m| > |\lambda_۲^m| \geq \cdots \geq |\lambda_n^m|$$

از طرفی اگر $\lambda_۱, \lambda_۲, \cdots, \lambda_n$ مقادیر ویژه A باشند، $\lambda_۱^m, \lambda_۲^m, \cdots, \lambda_n^m$ مقادیر ویژه A^m می‌باشند. از طرفی سرعت همگرایی روش توانی برای A و A^m به ترتیب به $\left|\frac{\lambda_۲}{\lambda_۱}\right|$ و $\left|\frac{\lambda_۲}{\lambda_۱}\right|^m$ وابسته خواهد بود. این بدین معنی است که چنانچه روش توانی به جای پیاده‌سازی برای ماتریس A، برای ماتریس A^m پیاده‌سازی شود، سرعت همگرایی روش توانی بسیار بیشتر خواهد بود. در نهایت ریشه m ام گرفتن از مقدار ویژه‌ی تولید شده‌ی از روش توانی برای A^m، بزرگترین مقدار ویژه A خواهد بود.

تذکر ۵-۹) لازم به ذکر است که توان m باید فرد در نظر گرفته شود؛ زیرا اگر m زوج باشد، در این صورت مثبت یا منفی بودن $\lambda_۱$ را نمی‌توان از روی $\lambda_۱^m$ تشخیص داد.

مثال ۵-۱۰) ماتریس $A = \begin{bmatrix} ۱ & ۲ \\ ۳ & ۲ \end{bmatrix}$ را در نظر بگیرید. نتایج به کار بردن روش توان ماتریسی برای ماتریس‌های $A^۲$ و $A^۳$ به فرم زیر خواهد بود.

جدول (۵-۱) نتایج استفاده از روش توان ماتریسی

تکرارها k	A	A^2	A^3
	مقادیر ویژه	$\sqrt{\text{مقادیر ویژه}}$	$\sqrt[۳]{\text{مقادیر ویژه}}$
۱	۵	۴/۳۵۸۹	۴/۲۵۲۳
۲	۳/۸	۴/۰۱۹۷	۳/۹۹۶۵
۳	۴/۰۵۲۶	۴/۰۰۱۲	۴/۰۰۰۰

همان‌طور که مشاهده می‌شود استفاده از $A^۲$ و $A^۳$ سرعت همگرایی را بیشتر کرده است. زیرا مقدار دقیق بزرگترین مقدار ویژه‌ی این ماتریس $\lambda_۱ = ۴$ است.

۵-۱-۲ روش معکوس توانی

روش معکوس توانی نیز همانند روش توانی می‌باشد، با این تفاوت که در این روش می‌توان کوچکترین مقدار ویژه به لحاظ قدر مطلق (نزدیکترین مقدار به عدد صفر) را پیدا کرد. همچنین پس از ذکر روش مذکور، روش دیگری برای پیدا کردن مقدار ویژه نزدیک به عددی دلخواه گفته می‌شود.

فرض کنید A ماتریسی معکوس‌پذیر باشد. چنانچه $\lambda_۱, \lambda_۲, \cdots, \lambda_n$ مقادیر ویژه A باشند، به طوری که:

$$|\lambda_۱| \geq |\lambda_۲| \geq \cdots \geq |\lambda_{n-۱}| > |\lambda_n|$$

در این صورت محاسبه کوچکترین مقدار ویژه ماتریس A ، معادل با محاسبه بزرگترین مقدار ویژه ماتریس $A^{-۱}$ و سپس معکوس کردن این مقدار خواهد بود. زیار مقادیر ویژه $A^{-۱}$ عبارتند از:

$$\frac{۱}{|\lambda_n|} > \frac{۱}{|\lambda_{n-۱}|} \geq \cdots \geq \frac{۱}{|\lambda_۱|}$$

برای این منظور بردار ناصفر $V_{n\times ۱}^{(۰)}$ انتخاب شده، نرمال شده‌ی آن را مساوی با $U^{(۰)}$ قرار داده و می‌توان نوشت:

$$V^{(۱)} = A^{-۱}U^{(۰)} \tag{۵-۹}$$

می‌توان به صورت مستقیم با محاسبه $A^{-۱}$ طبق روش‌های گفته شده در فصل دوم، معادله (۵-۹) را حل کرد؛ ولی چون محاسبه $A^{-۱}$ دشوار است، می‌توان به جای رابطه (۵-۹)، از رابطه معادل با آن یعنی (۵-۱۰) به صورت زیر استفاده کرد:

$$AV^{(۱)} = U^{(۰)} \tag{۵-۱۰}$$

با حل رابطه (۵-۱۰)، $V^{(۱)}$ استخراج خواهد شد. بزرگترین درایه $V^{(۱)}$ تقریبی از $\frac{۱}{\lambda_n}$ خواهد بود که با $\frac{۱}{\lambda_n^{(۱)}}$ نمایش داده می‌شود. دوباره نرمال شده $V^{(۱)}$ را مساوی با $U^{(۱)}$ قرار داده و همین روند ادامه پیدا خواهد کرد تا بعد از k مرحله، تقریبی برای $\frac{۱}{\lambda_n^{(k)}}$ به دست آید و بدین ترتیب تقریبی از خود $\lambda_n^{(k)}$ (کوچکترین مقدار ویژه A به لحاظ قدر مطلق) به دست می‌آید.

حال فرض کنید که هدف مسئله یافتن نزدیکترین مقدار ویژه به عدد مفروض p است (یعنی λ ای که کمترین فاصله را از عدد مفروض p دارد یا به عبارت دیگر هدف یافتن λ ای است که $|p-\lambda|$ کمترین مقدار خود را دارا باشد). از طرفی $\lambda - p$ مقدار ویژه‌ی $A - pI$ است؛ پس هدف پیدا کردن کوچکترین مقدار ویژه همین ماتریس می‌باشد. طبق روش معکوس توانی، ابتدا کوچکترین مقدار ویژه ماتریس $A - pI$ را باید به دست آورد، سپس آن را با p جمع نمود تا نزدیکترین مقدار ویژه A به عدد مفروض p محاسبه گردد.

مثال ۵-۱۱) نزدیکترین مقدار ویژه ماتریس $A = \begin{bmatrix} ۳ & ۶ & ۸ \\ ۴ & ۱۱ & ۱۶ \\ ۶ & ۱۵ & ۴۰ \end{bmatrix}$ را می‌توان به صورت زیر به دست آورد. همچنین بردارهای ویژه نرمال تولید شده در هر گام نیز در ادامه نمایش داده شده است.

ابتدا ماتریس $A - ۵I$ ساخته می‌شود که به صورت زیر خواهد بود:

$$A - ۵I = \begin{bmatrix} -۲ & ۶ & ۸ \\ ۴ & ۶ & ۱۶ \\ ۶ & ۱۵ & ۳۵ \end{bmatrix}$$

حال با استفاده از روش معکوس توانی، کوچکترین مقدار ویژه ماتریس فوق را باید پیدا کرد. یا به عبارت دیگر بزرگترین مقدار ویژه وارون ماتریس مذکور باید پیدا شود. این مقدار را با μ می‌توان به صورت زیر نمایش داد:

جدول (۵-۲) نتایج استفاده از روش معکوس توانی

تکرارها k	$\left(V^{(k)}\right)^T$	μ
۱	(۰/۵۵۱۰- و ۱ و ۰/۷۳۴۷)	۱۵/۴۶۶۰
۲	(۰/۵۵۷۵- و ۱ و ۰/۷۴۶۲)	۱۵/۵۴۲۸
۳	(۰/۵۵۷۵- و ۱ و ۰/۷۴۶۰)	۱۵/۵۴۲۱
۴	(۰/۵۵۷۵- و ۱ و ۰/۷۶۶۰)	۱۵/۵۴۲۱

یعنی بزرگترین مقدار ویژه‌ی $(A - ۵I)^{-۱}$ عبارت است از:

$$\mu = ۱۵/۵۴۲۱$$

لذا کوچکترین مقدار ویژه $A - ۵I$ عبارت است از:

$$\lambda = \frac{۱}{۱۵/۵۴۲۱} = ۰/۰۶۴۳$$

در نتبجه نزدیکترین مقدار ویژه ماتریس A به عدد ۵ عبارت است از:

$$\lambda = ۵ + ۰/۰۶۴۳ = ۵/۰۶۴۳$$

روش‌های توانی و معکوس توانی، برای پیدا کردن بزرگترین مقدار ویژه، کوچکترین مقدار ویژه و مقدار ویژه خاص مورد نظر، روش‌های بسیار مناسبی هستند. در ادامه روشی تکراری جهت پیدا کردن تمام مقادیر ویژه ماتریس A ذکر شده است.

۵-۱-۳ روش تقلیل توانی

اساس این روش بر این مبناست که ابتدا به روش توانی $\lambda_۱$ و $X_۱$ محاسبه می‌شود، سپس ماتریس $A_{n\times n}$ به یک ماتریس $(n-۱)\times(n-۱)$ تقلیل پیدا کرده که مقادیر ویژه این ماتریس همان $(n-۱)$ مقدار ویژه‌ی A (به جز $\lambda_۱$) خواهد بود. حال دوباره با روش توانی، بزرگترین مقدار ویژه‌ی این ماتریس تقلیل یافته، یعنی $\lambda_۲$ به دست می‌آید. سپس به طور مشابه، این ماتریس به یک ماتریس $(n-۲)\times(n-۲)$ تقلیل می‌یابد و به همین ترتیب، با ادامه این روند برای ماتریس‌های تقلیل یافته، کلیه مقادیر ویژه ماتریس A تولید خواهد شد. در ادامه دو روش جهت استخراج کلیه مقادیر ویژه‌ی A ارائه شده است.

۵-۱-۳-۱ روش مستقیم

روش مستقیم تقلیل توانی در قضیه زیر یعنی قضیه (۵-۱۲) بیان شده است.

قضیه ۵-۱۲) فرض کنید $A\in\mathbb{R}^{n\times n}$ و دارای مقادیر ویژه $\lambda_۱,\cdots\lambda_n$ باشد. چنانچه بزرگترین مقدار ویژه A به لحاظ قدر مطلق $\lambda_۱$ باشد و بردار ویژه‌ی نظیر آن $X_۱$ باشد و همچنین $Y_۱$ مساوی با بردار نرمال شده‌ی $X_۱$ باشد، آنگاه این بردار حداقل دارای یک عنصر (درایه) برابر ۱ است. اگر این درایه ۱ در مکان k ام بردار $Y_۱$ قررا داشته باشد و W^T نیز سطر k ام ماتریس A باشد؛ در این صورت ماتریس B به صورت:

$$B = A - Y_۱ W^T \qquad (۵\text{-}۱۱)$$

دارای مقادیر ویژه $۰,\lambda_۲,\lambda_۳,\cdots,\lambda_n$ می‌باشد.

برهان) رجوع شود به [۶].

مثال ۵-۱۳) مقادیر ویژه ماتریس $A=\begin{bmatrix}۲ & ۳ & ۲\\ ۱۰ & ۳ & ۴\\ ۳ & ۶ & ۱\end{bmatrix}$ با روش تقلیل توانی مستقیم به صورت زیر خواهد بود.

ابتدا با استفاده از روش توانی بزرگترین مقدار ویژه‌ی A را باید به دست آورد. این مقدار ویژه، $\lambda_۱ = ۱۱$ می‌باشد و بردار ویژه نظیر آن $X_۱ = \left[\frac{۱}{۲},۱,\frac{۳}{۴}\right]^T$ است. نرمال شده $X_۱$ مساوی با $Y_۱ = \left[\frac{۱}{۲},۱,\frac{۳}{۴}\right]^T$ قرار خواهد گرفت. بزرگترین درایه $Y_۱$ در مکان دوم قرار دارد، یعنی $k = ۲$. بنابراین سطر دوم ماتریس A باید در W^T قرار گیرد؛ یعنی $W^T = [۱۰\ ۳\ ۴]$.

حال ماتریس B به صورت زیر تولید خواهد شد:

$$B = A - Y_۱ W^T = \begin{bmatrix} -۳ & ۱/۵ & ۰ \\ ۰ & ۰ & ۰ \\ -۴/۵ & -۳/۷۵ & -۲ \end{bmatrix}$$

حال با حذف سطر دوم و ستون دوم ماتریس B، ماتریس جدید به صورت زیر خواهد بود:

$$C = \begin{bmatrix} -۳ & ۰ \\ -۴/۵ & -۲ \end{bmatrix}$$

مقادیر ویژه ماتریس C همان مقادیر ویژه A می‌باشد که $\lambda_۱$ از آن حذف شده است. چون C ماتریسی مثلثی است، بنابراین مقادیر ویژه‌ی آن، همان عناصر روی قطر اصلی ماتریس می‌باشند، یعنی $\lambda_۲ = -۳$ و $\lambda_۳ = -۲$.

۵-۱-۳-۲ روش هاوس هلدر

اساس کار این روش در قضیه زیر بیان شده است.

قضیه ۵-۱۴) فرض کنید $A \in \mathbb{R}^{n\times n}$ و دارای مقادیر ویژه $\lambda_۱, \cdots, \lambda_n$ باشد. چنانچه بزرگترین مقدار ویژه‌ی A به لحاظ قدر مطلق $\lambda_۱$ باشد و بردار ویژه نظیر آن $X_۱$ باشد و همچنین H معرف ماتریس هاوس هلدری باشد که از روی $X_۱$ ساخته شده است؛ یعنی:

$$U = X_۱ + sign\{x_۱\}\|x_۱\|_۲ e_۱ \in \mathbb{R}^{n\times ۱}$$

$$H = I_n - ۲\frac{UU^T}{U^T U}, H \in \mathbb{R}^{n\times n}$$

آنگاه:

$$A_۱ = HAH = \begin{bmatrix} \lambda_۱ & b^T \\ O & A_۲ \end{bmatrix}$$

که در آن $A_۲ \in \mathbb{R}^{(n-۱)\times(n-۱)}$ می‌باشد و مقادیر ویژه $A_۲$ همان مقادیر ویژه ماتریس A هستند به جز $\lambda_۱$ یعنی اگر $|\lambda_۱| > |\lambda_۲| > |\lambda_۳| \geq \cdots \geq |\lambda_n|$ آنگاه $\lambda_۲$ بزرگترین مقدار ویژه $A_۲$ است.

برهان) رجوع شود به [۶].

مثال ۵-۱۵) چنانچه بزرگترین مقدار ویژه ماتریس $A=\begin{bmatrix} ۲ & ۳ & ۲ \\ ۱۰ & ۳ & ۴ \\ ۳ & ۶ & ۱ \end{bmatrix}$ ، $\lambda_۱ = ۱۱$ باشد و بردار ویژه نظیر آن به صورت $X_۱ = \left[\frac{۱}{۲}, ۱, \frac{۳}{۴}\right]^T$ باشد؛ در این صورت سایر مقادیر ویژه A با کمک روش هاوس هلدر به صورت زیر به دست خواهد آمد.

ابتدا U را باید به صورت زیر تشکیل داد:

$$U = X_۱ + sign\left(\frac{۱}{۲}\right) \|X_۱\|_۲ e_۱ \Rightarrow$$

$$U = \begin{bmatrix} \frac{۱}{۲} + \frac{\sqrt{۲۹}}{۴} \\ ۱ \\ \frac{۳}{۴} \end{bmatrix}$$

حال باید H را به صورت زیر محاسبه کرد:

$$H = I_۳ - ۲\frac{UU^T}{U^TU} \Rightarrow$$

$$H = \begin{bmatrix} -۰/۳۷۱۴ & -۰/۷۴۲۸ & -۰/۵۵۷۱ \\ -۰/۷۴۲۸ & ۰/۵۹۷۸ & -۰/۳۰۱۷ \\ -۰/۵۵۷۱ & -۰/۳۰۱۷ & ۰/۷۷۳۷ \end{bmatrix}$$

از طرفی:

$$HAH = \begin{bmatrix} ۱۱ & ۴/۶۰۳۵ & ۴/۱۹۵۴ \\ ۰ & -۴/۱۸۸۰ & -۰/۸۵۳۳ \\ ۰ & ۳/۰۴۶۳ & -۰/۸۱۲۰ \end{bmatrix}$$

حال کافیست مقادیر ویژه $A_۲$ که به صورت زیر می‌باشد را به دست آورد:

$$A_۲ = \begin{bmatrix} -۴/۱۸۸۰ & -۰/۸۵۳۳ \\ ۳/۰۴۶۳ & -۰/۸۱۲۰ \end{bmatrix}$$

مقادیر ویژه $A_۲$ عبارتند از:

$$\lambda_۲ = -۴ \ \ , \ \ \lambda_۳ = -۲$$

۵-۲ روش‌های تبدیلی

در این روش با استفاده از تبدیلات مشابه، ماتریسی با مقادیر ویژه یکسان با ماتریس اولیه، اما با فرم ساده‌تر تولید می‌گردد و سپس مقادیر ویژه و بردارهای ویژه‌ی ماتریس جدید محاسبه می‌شوند. ماتریس‌های تبدیلی به کار رفته در این روش، ماتریس‌هایی متعامد می‌باشند؛ زیرا این دست از ماتریس‌ها، برای کاهش و یا کنترل خطا مناسب‌ترین هستند. در ادامه ۵ روش تبدیلی برای تولید مقادیر ویژه و بردارهای ویژه ارائه می‌گردد. قبل از ارائه روش‌ها، نیاز به ارائه تعاریف ماتریس ۳ قطری و ماتریس بالا (پایین) هسنبرگی و همچنین چگونگی تعیین مقادیر ویژه این دست از ماتریس‌ها می‌باشد.

تعریف ۵-۱۶) ماتریس مربعی A را ۳ قطری گویند، هرگاه $\forall |i-j| > ۱ : a_{ij} = ۰$ باشد. فرم کلی یک ماتریس ۳ قطری به صورت زیر است:

$$\begin{bmatrix} b_{۱} & c_{۱} & ۰ & \cdots & \cdots & ۰ \\ a_{۲} & b_{۲} & c_{۲} & \ddots & \ddots & \vdots \\ ۰ & a_{۳} & b_{۳} & c_{۳} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-۱} & b_{n-۱} & c_{n-۱} \\ ۰ & \cdots & \cdots & ۰ & a_n & b_n \end{bmatrix}_{n\times n}$$

روش تعیین مقادیر ویژه ماتریس ۳ قطری متقارن به صورت زیر می‌باشد:

ابتدا فرض کنید که A یک ماتریس ۳ قطری متقارن همانند زیر باشد:

$$A = \begin{bmatrix} b_{۱} & c_{۱} & ۰ & \cdots & \cdots & ۰ \\ c_{۱} & b_{۲} & c_{۲} & \ddots & \ddots & \vdots \\ ۰ & c_{۲} & b_{۳} & c_{۳} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{n-۲} & b_{n-۱} & c_{n-۱} \\ ۰ & \cdots & \cdots & ۰ & c_{n-۱} & b_n \end{bmatrix}_{n\times n}$$

هدف این است که یک فرمول بازگشتی برای تولید چند جمله‌ای مشخصه ماتریس A ارائه شود. در این صورت چنانچه چندجمله‌ای مشخصه ماتریس A با $f_n(\lambda)$ نمایش داده شود، آنگاه از رابطه بازگشتی زیر به دست خواهد آمد:

$$\begin{cases} f_r(\lambda) = (\lambda - b_r) f_{r-۱}(\lambda) - c_{r-۱}^{۲} f(\lambda)\ , \quad r = ۲,۳,\cdots,n \\ f_{۰}(\lambda) = ۱ \quad , \quad f_{۱}(\lambda) = \lambda - b_{۱} \end{cases}$$

حال می‌توان مقادیر ویژه ماتریس A را از این چند جمله‌ای مشخصه استخراج کرد.

مثال ۵-۱۷) معادله مشخصه ماتریس ۳ قطری و متقارن $A = \begin{bmatrix} ۱ & ۲ & ۰ \\ ۲ & ۳ & ۱ \\ ۰ & ۱ & ۴ \end{bmatrix}$ با استفاده از فرمول بازگشتی به صورت زیر به دست خواهد آمد.

$$f_{۰}(\lambda) = ۱ \quad , \quad f_{۱}(\lambda) = \lambda - ۱$$

$$r = ۲ \Rightarrow f_{۲}(\lambda) = (\lambda - b_{۲})f_{۱}(\lambda) - c_{۱}^{۲} f_{۰}(\lambda) = (\lambda - ۳)(\lambda - ۱) - ۲^{۲} = \lambda^{۲} - ۴\lambda - ۱$$

$$r = ۳ \Rightarrow f_{۳}(\lambda) = (\lambda - ۴)f_{۲}(\lambda) - c_{۲}^{۲} f_{۱}(\lambda) = (\lambda - ۴)(\lambda^{۲} - ۴\lambda - ۱) - ۱(\lambda - ۱) \Rightarrow$$

$$f_{۳}(\lambda) = \lambda^{۳} - ۸\lambda^{۲} + ۱۴\lambda + ۵$$

تعریف ۵-۱۸) ماتریس مربعی A را یک ماتریس پایین هسنبرگی گویند، هرگاه:

$$\forall j > i + ۱ : a_{ij} = ۰$$

به عبارت دیگر یک ماتریس پایین مثلثی که بالای قطر اصلی آن یک قطر دیگر قرار گرفته است.

فرم کلی یک ماتریس پایین هسنبرگی به صورت زیر است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & ۰ & \cdots & \cdots & ۰ \\ a_{۲۱} & a_{۲۲} & a_{۲۳} & \ddots & \ddots & \vdots \\ a_{۳۱} & a_{۳۲} & a_{۳۳} & a_{۳۴} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & ۰ \\ \vdots & \ddots & \ddots & \ddots & \ddots & a_{n-۱,n} \\ a_{n۱} & \cdots & \cdots & \cdots & \cdots & a_{nn} \end{bmatrix}$$

تعریف ۵-۱۹) ماتریس مربعی A را یک ماتریس بالا هسنبرگی گویند، هرگاه:

$$\forall i > j + ۱ : a_{ij} = ۰$$

به عبارت دیگر یک ماتریس بالا مثلثی که پایین قطر اصلی آن یک قطر دیگر قرار گرفته است.

فرم کلی یک ماتریس بالا هسنبرگی به صورت زیر است:

$$\begin{bmatrix} a_{۱۱} & a_{۱۲} & a_{۱۳} & \cdots & \cdots & a_{۱n} \\ a_{۲۱} & a_{۲۲} & a_{۲۳} & \ddots & \ddots & \vdots \\ ۰ & a_{۳۲} & a_{۳۳} & \ddots & \ddots & \vdots \\ \vdots & \ddots & a_{۴۳} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ ۰ & \cdots & \cdots & ۰ & a_{n,n-۱} & a_{nn} \end{bmatrix}$$

قضیه ۵-۲۰) فرض کنید ماتریس پایین هسنبرگی A داده شده باشد. چنانچه λ مقدار ویژه‌ی A و X بردار ویژه‌ی نظیر λ باشد؛ در این صورت می‌توان نوشت:

$$AX = \lambda X \,, X \neq ۰ \,, X = \begin{bmatrix} x_۱ \\ x_۲ \\ \vdots \\ x_n \end{bmatrix} \tag{۵-۱۲}$$

برهان) رجوع شود به [۶].

فرم باز شده دستگاه (۵-۱۲) به صورت زیر خواهد بود ($(A - \lambda I)X = ۰$):

$$\begin{cases} (a_{۱۱} - \lambda)x_۱ + a_{۱۲}x_۲ = ۰ \\ a_{۲۱}x_۱ + (a_{۲۲} - \lambda)x_۲ + a_{۲۳}x_۳ = ۰ \\ \vdots \\ a_{n۱}x_۱ + \cdots + (a_{nn} - \lambda)x_n = ۰ \end{cases}$$

که در آن $x_۱, \cdots, x_n$ و λ مجهول هستند. حال چنانچه قرار داده شود $x_۱ = ۱$ ، از معادل اول $x_۲$ ، از معادله دوم $x_۳$ و به همین ترتیب از معادله $(n - ۱)$ ام x_n به دست می‌آید. چنانچه قرار داده شود:

$$F(\lambda) = a_{n۱}x_۱ + \cdots + (a_{nn} - \lambda)x_n \tag{۵-۱۳}$$

آنگاه:

$$\det(A - \lambda I) = cF(\lambda) \;\; , \;\; c = (-۱)^{n+۱} \prod_{i=۱}^{n-۱} a_{i,i+۱} \tag{۵-۱۴}$$

تذکر ۵-۲۱) با روشی کاملاً مشابه با قضیه (۵-۲۰) می‌توان مقادیر ویژه یک ماتریس بالا هسنبرگی را محاسبه کرد.

مثال ۵-۲۲) معادله مشخصه ماتریس پایین هسنبرگی $A = \begin{bmatrix} ۲ & ۱ & ۰ \\ ۳ & ۴ & ۱ \\ ۶ & ۷ & ۸ \end{bmatrix}$ به صورت زیر به دست خواهد آمد.

ابتدا دستگاه متناظر با ماتریس A با شرط ناصفر بودن X و $(A - \lambda I)X = ۰$ باید تولید شود. پس:

$$\begin{cases} (۲ - \lambda)x_۱ + x_۲ = ۰ \\ ۳x_۱ + (۴ - \lambda)x_۲ + x_۳ = ۰ \\ ۶x_۱ + ۷x_۲ + (۸ - \lambda)x_۳ = ۰ \end{cases}$$

حال $x_۱$ مساوی با ۱ قرار داده می‌شود. بنابراین اگر روابط بالا به ترتیب (۱)، (۲) و (۳) نامگذاری شوند، می‌توان آن‌ها را با توجه به $x_۱ = ۱$ به صورت زیر بازنویسی کرد:

$$(۱) \xRightarrow{x_۱=۱} (۲-\lambda) + x_۲ = ۰ \Rightarrow x_۲ = \lambda - ۲$$

$$(۲) \xRightarrow{x_۲=\lambda-۲} ۳ + (۴-\lambda)(\lambda-۲) + x_۳ = ۰ \Rightarrow x_۳ = \lambda^۲ - ۶\lambda + ۵$$

$$(۳) \Longrightarrow F(\lambda) = ۶ + ۷(\lambda - ۲) + (۸-\lambda)(\lambda^۲ - ۶\lambda + ۵) = -\lambda^۳ + ۱۴\lambda^۲ - ۴۶\lambda - ۳۲$$

در ادامه روش‌های تبدیلی جهت استخراج مقادیر ویژه $A \in \mathbb{R}^{n\times n}$ آورده شده است.

روش‌های تبدیلی جهت استخراج مقادیر ویژه ماتریس $A_{n\times n}$:
- روش ژاکوبی
- روش گیونز
- روش هاوس هلدر
- تجزیه LU
- تجزیه QR

۵-۲-۱ روش ژاکوبی

فرض کنید $A \in \mathbb{R}^{n\times n}$ و متقارن باشد. بدون شک ساده‌ترین فرمی که می‌توان ماتریس A را به آن تبدیل کرد، فرم قطری است؛ زیرا در این صورت مقادیر ویژه‌ی ماتریس A ، همان اعضای روی قطر اصلی ماتریس قطری خواهد بود. این روش در واقع جهت تبدیل یک ماتریس متقارن $A \in \mathbb{R}^{n\times n}$ ، به یک ماتریس قطری به‌وسیله حذف عناصر غیر واقع بر قطر اصلی ماتریس A بنا شده است.این فرایند، یک فرایند تکراری است و در هر مرحله به‌وسیله یک ماتریس متعامد، یک صفر غیر قطری در A تولید می‌شود. البته متاسفانه در این روش با تولید هر صفر جدید، عنصر غیر صفری در محل صفرهای قبلی ایجاد خواهد شد.

روش ژاکوبی با صفر نمودن بزرگترین عنصر غیر قطری از نظر قدر مطلق در هر مرحله ادامه می‌یابد و هنگامی که تمام عناصر غیر قطری از نظر قدر مطلق از عدد کوچک مشخصی (همانند ε) کمتر باشند، عملیات خاتمه یافته و عناصر روی قطر اصلی به عنوان مقادیر ویژه ماتریس A انتخاب می‌شوند.

ایده کار روش ژاکوبی بدین صورت است که:

فرض کنید $A \in \mathbb{R}^{n\times n}$ و متقارن باشد. طبق قضایای مطرح شده، ماتریس متعامدی چون P موجود است که رابطه‌ی $P^TAP = D$ برقرار باشد که در آن D یک ماتریس قطری است و اعضای روی قطر اصلی آن معرف مقادیر ویژه A می‌باشند.

در واقع در این روش، هدف یافتن دنباله S_k از ماتریس‌های متعامد است به طوری که:

$$\lim_{k\to\infty} S_۱ S_۲ \cdots S_k = P$$

حال اگر قرار داده شود:

$$A_k = S_k^T \cdots S_۱^T A S_۱ \cdots S_k$$

آنگاه به صورت ساده‌تر می‌توان نوشت:

$$A_k = S_k^T A_{k-۱} S_k \quad , \quad A_. = A$$

واضح است که A_k با A متشابه بوده و در نتیجه دارای مقادیر ویژه یکسان می‌باشند. حال فرض کنید اعضای A_k با $\left(a_{ij}^{(k)}\right)$ و درایه‌های S_k با $\left(s_{ij}^{(k)}\right)$ نمایش داده شود. همچنین فرض کنید که بزرگترین عنصر غیر قطری در ماتریس $A_{k-۱}$، $\begin{cases} a_{pq}^{(k-۱)} \\ p < q \end{cases}$ باشد؛ در این صورت چنانچه ماتریس S_k به صورت زیر تعریف شود:

$$\begin{cases} s_{pp}^{(k)} = s_{qq}^{(k)} = \cos\theta_k \\ s_{pq}^{(k)} = -s_{qp}^{(k)} = \sin\theta_k \\ s_{ii}^{(k)} = ۰ \quad , \quad s_{ij}^{(k)} = ۰ \end{cases}$$

ماتریس S_k یک ماتریس متعامد است که ماتریس دوران نامیده می‌شود. در این ماتریس، θ_k طوری انتخاب می‌شود که در مرحله بعد $a_{pq}^{(k)} = ۰$ شود. در این صورت می‌توان نوشت:

$$\tan(۲\theta_k) = \frac{-۲a_{pq}^{(k-۱)}}{a_{pp}^{(k-۱)} - a_{qq}^{(k-۱)}}$$

و θ_k طوری انتخاب می‌شود که $|\theta_k| \le \frac{\pi}{۴}$. در واقع S_k همان ماتریس گیونز سطر p ام و ستون q ام می‌باشد. یعنی:

$$S_k = G(p, q, \theta_k)$$

حال با ضرب S_k از سمت چپ و راست در ماتریس $A_{k-۱}$ به ماتریسی باید رسید که:

$$A_k = S_k^T A_{k-۱} S_k$$

و درایه $a_{pq}^{(k)}$ و $a_{qp}^{(k)}$ صفر شده‌اند.

تذکر ۵-۲۳) p و q طوری انتخاب می‌شوند که $(p < q)$ $a_{pq}^{(k-۱)}$ بزرگترین عنصر به لحاظ قدر مطلق غیر قطری $A_{k-۱}$ باشد. با ادامه همین روند، می‌توان نشان داد ماتریس A_k به سمت یک ماتریس قطری میل خواهد کرد که اعضای قطری آن همان مقادیر ویژه‌ی ماتریس A می‌باشند.

مثال ۵-۲۴) مقادیر ویژه‌ی ماتریس $A = \begin{bmatrix} ۱۰ & ۷ & ۸ & ۷ \\ ۷ & ۵ & ۶ & ۵ \\ ۸ & ۶ & ۱۰ & ۹ \\ ۷ & ۵ & ۹ & ۱۰ \end{bmatrix}$ با روش ژاکوبی به صورت زیر به دست می‌آید.

ابتدا متقارن بودن A بررسی می‌شود که در این مثال قطری است. بزرگترین عنصر غیر قطری A عبارت است از $a_{۳۴} = ۹$. پس:

$$p = ۳, q = ۴ \Rightarrow \tan(۲\theta_۱) = \frac{-۲a_{۳۴}}{a_{۳۳} - aa_{۴۴}} = \infty \Rightarrow ۲\theta_۱ = \frac{\pi}{۲} \Rightarrow \theta_۱ = \frac{\pi}{۴}$$

بنابراین:

$$S_۱ = G\left(۳,۴,\frac{\pi}{۴}\right) = \begin{bmatrix} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ \\ ۰ & ۰ & \frac{\sqrt{۲}}{۲} & -\frac{\sqrt{۲}}{۲} \\ ۰ & ۰ & \frac{\sqrt{۲}}{۲} & \frac{\sqrt{۲}}{۲} \end{bmatrix} \Rightarrow$$

$$A_۱ = S_۱^T A S_۱ = \begin{bmatrix} ۱۰ & ۷ & \frac{۱۵}{\sqrt{۲}} & -\frac{۱}{\sqrt{۲}} \\ ۷ & ۵ & \frac{۱۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۲}} \\ \frac{۱۵}{\sqrt{۲}} & \frac{۱۱}{\sqrt{۲}} & ۱۹ & ۰ \\ -\frac{۱}{\sqrt{۲}} & -\frac{۱}{\sqrt{۲}} & ۰ & ۱ \end{bmatrix}$$

حال بزرگترین عنصر غیر قطری $A_۱$ باید پیدا شود، در این حالت $p = ۱, q = ۳, a_{pq} = \frac{۱۵}{\sqrt{۲}}$ می‌باشد. پس:

$$\tan(۲\theta_۲) = \frac{-۲a_{۱۳}}{a_{۱۱} - a_{۳۳}} = \frac{۳۰}{۹\sqrt{۲}} = \frac{۲}{۳۵۷۰} \Rightarrow \theta_۲ = ۳۳/۵۰۵°$$

بنابراین قرار داده می‌شود:

$$S_۲ = G(۱,۳,\theta_۲)$$

حال می‌توان نوشت:

$$A_۲ = S_۲^T A_۱ S_۲ = \begin{bmatrix} ۲/۹۷۸۳ & ۱/۵۴۳۲ & ۰ & -۰/۵۸۹۷ \\ ۱/۵۴۳۲ & ۵ & ۱۰/۳۴۹۸ & -۰/۷۰۷۱ \\ ۰ & ۱۰/۳۴۹۸ & ۲۶/۰۲۱۷ & -۰/۳۹۰۳ \\ -۰/۵۸۹۶ & -۰/۷۰۷۱ & -۰/۳۹۰۳ & ۱ \end{bmatrix}$$

با ادامه همین روند، بعد از ۱۰ تکرار به ماتریس زیر می‌توان رسید:

$$A_{۱۰} = \begin{bmatrix} ۳/۸۵۸۱ & ۰ & -۰/۰۰۰۷ & -۰/۰۰۱۷ \\ ۰ & ۰/۰۱۰۲ & ۰/۰۰۰۴ & ۰ \\ -۰/۰۰۰۷ & ۰/۰۰۰۴ & ۳۰/۲۸۸۷ & ۰/۰۰۱۶ \\ -۰/۰۰۱۷ & ۰ & ۰/۰۰۱۶ & ۰/۸۴۳۱ \end{bmatrix}$$

با اندکی دقت در ماتریس فوق، می‌توان فهمید که $i \neq j \quad \left|a_{ij}^{(۱۰)}\right| < ۰/۰۰۲$. لذا عناصر ر.ی قطر اصلی $A_{۱۰}$ در واقع تقریب‌هایی از مقادیر ویژه ماتریس A می‌باشند. یعنی:

$$\lambda_۱ = ۳۰/۲۸۸۷$$
$$\lambda_۲ = ۳/۸۵۸۱$$
$$\lambda_۳ = ۰/۸۴۳۱$$
$$\lambda_۴ = ۰/۰۱۰۲$$

تذکر ۵-۲۵) دقت کنید که در عمل، در به کارگیری روش ژاکوبی، برای تولید ماتریس‌های S_k نیازی به محاسبه θ_k نیست؛ زیرا با فرض $-\frac{\pi}{a} \leq \theta_k \leq \frac{\pi}{۴}$ و فرض‌های زیر:

$$\begin{cases} r = -a_{pq}^{(k-۱)} \\ s = \dfrac{a_{pp}^{(k-۱)} - a_{qq}^{(k-۱)}}{۲} \\ t = \sqrt{r^۲ + s^۲} \end{cases}$$

می‌توان نوشت:

$$\begin{cases} \cos\theta_k = \sqrt{\dfrac{t + |s|}{۲t}} \\ \sin\theta_k = \dfrac{|r|}{\sqrt{۲t(t + |s|)}} \end{cases}$$

۵-۲-۲ روش گیونز

روش گیونز بر تولید تبدیلات متعامد از نوع روش ژاکوبی استوار است. ولی این روش چنان طراحی شده است که بر خلاف روش ژاکوبی، صفرهای تولید شده در هر مرحله، در مراحل دیگر حفظ خواهد شد. فرض کنید $A \in \mathbb{R}^{n\times n}$ ماتریسی متقارن باشد. هدف این روش، تبدیل ماتریس A به یک ماتریس ۳ قطری است. در حالت کلی محاسبه مقادیر ویژه یک ماتریس ۳ قطری به آسانی صورت نمی‌گیرد. اما چون A متقارن است، ماتریس ۳ قطری نهایی، متقارن خواهد بود و محاسبه مقادیر ویژه‌ی این دست از ماتریس‌ها نسبتاً ساده است.

روش کار تبدیلات گیونز:

فرض کنید $A \in \mathbb{R}^{n\times n}$ متقارن باشد و $A = (a_{ij})$. هدف صفر کردن عناصر a_{ij} برای $|i-j| > ۱$ به کمک ماتریس‌های متعامد گیونز خواهد بود.

اولین گام جهت ۳ قطری کردن A، صفر کردن عناصر سطر اول به جز $a_{۱۱}$ و $a_{۱۲}$ می‌باشد. برای صفر کردن $a_{۱k}$ که $k = ۳,۴,\cdots,n$ از ماتریس دوران $G_{۲k}$ باید استفاده کرد. فرض کنید:

$$A_۱ = A \quad , \quad A_۲ = G_{۲k}^{-۱} A_۱ G_{۲k} = G_{۲k}^{T} A_۱ G_{۲k} \; , k = ۳,۴,\cdots,n$$

که در آن:

$$\begin{cases} G_{۲k} = G(۲,k,\theta_k) \; , k = ۳,۴,\cdots,n \\ \tan\theta_k = \dfrac{a_{۱k}^{(۱)}}{a_{۱۲}^{(۱)}} \; , k = ۳,۴,\cdots,n \; , a_{ij}^{(۱)} = a_{ij} \end{cases}$$

در گام دوم، باید برخی از عناصر سطر دوم ماتریس $A_۲$ صفر شود. به طور مشابه برای این منظور از ماتریس‌های $G_{۳k}$ استفاده می‌شود که در آن $k = ۴,۵,\cdots,n$. همچنین در این حالت $\tan\theta_k = \frac{a_{۲k}^{(۲)}}{a_{۲۳}^{(۲)}}$ می‌باشد. پس در نهایت برای صفر کردن برخی از عناصر سطر r ام ماتریس A_r که به صورت $a_{rk}^{(r)}$ می‌باشد و $k = r+۲,\cdots,n$، باید از ماتریس‌های دوران $G_{r+۱,k}$ که $G_{r+۱,k} = G(r+۱,k,\theta_k)$ می‌باشد و همچنین $\tan\theta_k = \frac{a_{rk}^{(r)}}{a_{r,r+۱}^{(r)}}$ استفاده کرد. در نهایت:

$$A_{r+۱} = G_{r+۱,k}^{T} A_r G_{r+۱,k}$$

تذکر ۵-۲۶) چنانچه ماتریس A متقارن نباشد، تبدیلات گیونز، ماتریس A را به یک ماتریس پایین هسنبرگی تبدیل خواهد کرد.

تذکر ۵-۲۷) دقت کنید که در عمل برای تولید ماتریس‌های گیونز نیازی به محاسبه θ_k نخواهد بود؛ بلکه مقادیر مثلثاتی ماتریس مذکور به صورت $\cos\theta_k = \frac{۱}{\sqrt{۱+\tan^۲\theta_k}}$ و $\sin\theta_k = \cos\theta_k \tan\theta_k$ به دست خواهد آمد.

مثال ۵-۲۸) با استفاده از تبدیلات گیونز ماتریس $A = \begin{bmatrix} ۱ & ۱ & ۰/۵ \\ ۱ & ۱ & ۰/۲۵ \\ ۰/۵ & ۰/۲۵ & ۲ \end{bmatrix}$ به صورت زیر به یک ماتریس ۳ قطری تبدیل خواهد شد.

برای این کار کافیست که $a_{۱۳}$ صفر شود (چون ماتریس قطری است، پس $a_{۳۱}$ نیز خودبه‌خود صفر خواهد شد). بنابراین $k = ۳$ و $r = ۱$ ، چون $a_{۱۲} \neq ۰$ و

$$\tan\theta_{۱} = \frac{a_{rk}}{a_{r,r+۱}} = \frac{a_{۱۳}}{a_{۱۲}} = \frac{۱}{۲}$$

طبق روابط یاد شده در تذکر (۵-۲۷) می‌توان نوشت:

$$\cos\theta_{۱} = \frac{۲}{\sqrt{۵}} \quad , \quad \sin\theta_{۱} = \frac{۱}{\sqrt{۵}}$$

بنابراین:

$$G_{۲۳} = G(۲,۳,\theta_{۱}) = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۰/۸۹۴۴ & -۰/۴۴۷۲ \\ ۰ & ۰/۴۴۷۲ & ۰/۸۹۴۴ \end{bmatrix}$$

در نتیجه:

$$A_{۲} = G_{۲۳}^{T} A_{۱} G_{۲۳} = G_{۲۳}^{T} A G_{۲۳} = \begin{bmatrix} ۱ & ۱/۱۱۸ & ۰ \\ ۱/۱۱۸ & ۱/۴۰ & ۰/۵۵ \\ ۰ & ۰/۵۵ & ۱/۶۰ \end{bmatrix}$$

که حاصل یک ماتریس ۳ قطری متقارن است.

۵-۲-۳ روش هاوس هلدر

این روش نیز همانند روش گیونز تبدیلات متعامد را برای تبدیل یک ماتریس متقارن به فرم ۳ قطری و یا تبدیل یک ماترسی نامتقارن به فرم ماتریس هسنبرگی تولید می‌کند.

مزیت این روش بر روش گیونز این است که تمام صفرهای ممکن در هر سطر، تنها و تنها به وسیله یک تبدیل تولید می‌شود. در واقع برای تبدیل یک ماتریس متقارن $n \times n$ به یک ماتریس ۳ قطری و یا حتی تبدیل یک ماتریس نامتقارن $n \times n$ به یک ماتریس هسنبرگی، به $n - ۲$ تبدیل هاوس هلدر نیاز است. با این‌که این روش به لحاظ محاسباتی پیچیده‌تر از روش گیونز می‌باشد؛ اما به علت استفاده از $n - ۲$ ماتریس تبدیل هاوس هلدر، در نهایت باعث صرفه‌جویی قابل ملاحظه‌ای در وقت رایانه خواهد شد.

روش کار تبدیلات هاوس هلدر:

فرض کنید $A \in \mathbb{R}^{n\times n}$ و $A = (a_{ij})$ باشد. حال می‌توان نوشت:

$$V_k = \begin{bmatrix} \cdot \\ \vdots \\ \cdot \\ v_k^{(k)} \\ \vdots \\ v_k^{(n)} \end{bmatrix}, k = ۲,۳,\cdots,n-۱$$

که در آن:

$$\begin{cases} v_k^{(k)} = \sqrt{\dfrac{۱}{۲}\left|۱ \pm \dfrac{a_{k-۱,k}}{\sqrt{s}}\right|} \quad , \quad v_k^{(j)} = \pm \dfrac{a_{k-۱,j}}{۲v_k^{(k)}\sqrt{s}} \quad , \quad k = ۲,۳,\cdots,n-۱ \\ s = \displaystyle\sum_{j=k}^{n} a_{k-۱,j}^{۲} \quad , \quad k = ۲,۳,\cdots,n-۱ \end{cases}$$

در روابط بالا، اگر $a_{k-۱,k}$ مثبت باشد، علامت (+) در نظر گرفته می‌شود و در غیر این‌صورت علامت (-) انتخاب می‌شود. با این انتخاب، می‌توان نوشت $V_k^T V_k = ۱$. همچنین می‌توان قرار داد:

$$H_k = I_n - ۲V_k V_k^T$$

و در نهایت:

$$\begin{cases} A_k = H_k^T A_{k-۱} H_k \quad , \quad k = ۲,۳,\cdots\backslash n-۱ \\ A_۱ = A \end{cases}$$

با این تعریف، عناصر مورد نظر در A_k صفر می‌شوند و در مرحله $(n-۱)$ ام، $A_{n-۱}$ یک ماتریس ۳ قطری یا هسنبرگی است که با A متشابه است و لذا مقادیر ویژه آن‌ها یکسان خواهند بود.

مثال ۵-۲۹) ماتریس $A = \begin{bmatrix} ۶ & -۲ & ۱ & ۲ \\ -۲ & -۱۱/۸ & -۵ & -۱۲۵ \\ ۱ & -۵ & ۱۳ & -۴ \\ ۲ & -۱۲ & -۴ & ۲ \end{bmatrix}$ با کمک روش هاوس هلدر به صورت زیر به یک ماتریس ۳ قطری تبدیل خواهد شد.

ابتدا $k = ۲$ قرار داده می‌شود. بنابراین $v_۲^{(۱)} = ۰$ خواهد شد.

در ادامه:

$$s = (-۲)^۲ + ۱^۲ + ۲^۲ = ۹ \quad , \quad v_۲^{(۲)} = \sqrt{\frac{۱}{۲}\left|۱ - \frac{-۲}{۳}\right|} = \sqrt{\frac{۵}{۶}}$$

چون $a_{۱۲}$ منفی است؛ بنابراین در محاسبه $v_۲^{(۲)}$ از علامت (-) استفاده شده است. سپس:

$$v_۲^{(۳)} = -\frac{a_{۱۳}}{۲v_۲^{(۲)}\sqrt{s}} = -\frac{۱}{۲(۳)\sqrt{\frac{۵}{۶}}} = -\frac{۱}{\sqrt{۳۰}} \, , v_۲^{(۴)} = -\frac{a_{۱۴}}{۲v_۲^{(۲)}\sqrt{s}} = -\frac{۲}{۲(۳)\sqrt{\frac{۵}{۶}}} = -\frac{۲}{\sqrt{۳۰}}$$

لذا بردار $V_۲$ به صورت زیر خواهد بود:

$$V_۲ = \begin{bmatrix} ۰ \\ \sqrt{\frac{۵}{۶}} \\ -\frac{۱}{\sqrt{۳۰}} \\ -\frac{۲}{\sqrt{۳۰}} \end{bmatrix}$$

حال می‌توان نوشت:

$$H_۲ = I - ۲V_۲V_۲^T \Rightarrow H_۲ = \begin{bmatrix} ۱ & ۰ & ۰ & ۰ \\ ۰ & -\frac{۲}{۳} & \frac{۱}{۳} & \frac{۲}{۳} \\ ۰ & \frac{۱}{۳} & \frac{۱۴}{۱۵} & -\frac{۲}{۱۵} \\ ۰ & \frac{۲}{۳} & -\frac{۲}{۱۵} & \frac{۱۱}{۱۵} \end{bmatrix}$$

سپس می‌توان نوشت:

$$A_۲ = H_۲^T A_۱ H_۲ = \begin{bmatrix} ۶ & ۳ & ۰ & ۰ \\ ۳ & ۸/۲۳ & ۳ & ۴ \\ ۰ & ۳ & ۹ & -۱۲ \\ ۰ & ۴ & -۱۲ & -۱۴ \end{bmatrix}$$

حال به ازای $k = ۳$ ، $v_۳^{(۱)} = v_۳^{(۲)} = ۰$ خواهد بود. پس در ادامه می‌توان نوشت:

$$s = ۳^۲ + ۴^۲ = ۲۵$$

$$v_۳^{(۳)} = \sqrt{\frac{۱}{۲}\left|۱ + \frac{۳}{۵}\right|} = \sqrt{\frac{۴}{۵}} \quad , \quad v_۳^{(۴)} = +\frac{۴}{۲(\sqrt{\frac{۴}{۵}})۵} = \frac{۱}{\sqrt{۵}}$$

بنابراین:

$$V_۳ = \frac{۱}{\sqrt{۵}}\begin{bmatrix} ۰ \\ ۰ \\ ۲ \\ ۱ \end{bmatrix}$$

در نتیجه $H_۳$ عبارت است از:

$$H_۳ = I - ۲V_۳V_۳^T = \begin{bmatrix} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ \\ ۰ & ۰ & -\frac{۳}{۵} & -\frac{۴}{۵} \\ ۰ & ۰ & -\frac{۴}{۵} & \frac{۳}{۵} \end{bmatrix}$$

و در نهایت $A_۳$ به صورت زیر خواهد بود:

$$A_۳ = H_۳^T A_۲ H_۳ = \begin{bmatrix} ۶ & ۳ & ۰ & ۰ \\ ۳ & ۸/۲ & -۵ & ۰ \\ ۰ & -۵ & -۱۷/۲۴ & ۷/۶۸ \\ ۰ & ۰ & ۷/۶۸ & ۱۲/۲۴ \end{bmatrix}$$

که حاصل ماتریسی ۳ قطری و متقارن خواهد بود.

۵-۲-۴ روش تجزیه LU

اساس این روش تجزیه ماتریس A ، به حاصل‌ضرب LU استوار است. هدف تبدیل A به ماتریسی متشابه و مثلثی است که مقادیر ویژه‌ی آن همان مقادیر ویژه‌ی A خواهد بود. فرض کنید عناصر روی قطر اصلی L ، ۱ باشند؛ در این صورت اگر $A_۱ = A$ آنگاه می‌توان نوشت که $A_۱ = L_۱U_۱$. اگر $A_۲ = U_۱L_۱$ تعریف شود، آنگاه ماتریس $A_۱$ و $A_۲$ متشابه‌اند و مقادیر ویژه یکسانی دارند، زیرا:

$$A_۲ = U_۱L_۱ = (L_۱^{-۱}L_۱)U_۱L_۱ = L_۱^{-۱}(L_۱U_۱)L_۱ = L_۱^{-۱}(A_۱)L_۱$$

دوباره $A_۲$ باید به حاصل‌ضرب LU تجزیه شود؛ یعنی $A_۲ = L_۲U_۲$. به همین ترتیب فرض کنید A_k به صورت L_kU_k تجزیه شود. باید قرار داد $A_{k+۱} = U_kL_k$ و ماتریس‌های A_k و $A_{k+۱}$ متشابه‌اند. در نهایت:

$$A_{k+۱} = L_k^{-۱}A_kL_k$$

مثال ۵-۳۰) با استفاده از تجزیه LU ، مقادیر ویژه ماتریس $A = \begin{bmatrix} ۳۰ & ۲ & ۴ \\ ۳ & ۷ & -۴ \\ ۶ & -۳ & ۸ \end{bmatrix}$ به صورت زیر به دست خواهد آمد.

گام اول) باید تجزیه LU ماتریس A را به دست آورد. پس:

$$\begin{cases} A = A_۱ \\ A_۱ = L_۱U_۱ \end{cases} \Rightarrow$$

$$L_۱ = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰/۱ & ۱ & ۰ \\ ۰/۲ & -۰/۵ & ۱ \end{bmatrix}, U_۱ = \begin{bmatrix} ۳۰ & ۲ & ۴ \\ ۰ & ۶/۸ & -۴/۴ \\ ۰ & ۰ & ۵ \end{bmatrix}$$

حال قرار داده می‌شود $A_۲ = U_۱L_۱$. پس:

$$A_۲ = U_۱L_۱ = \begin{bmatrix} ۳۱ & ۰ & ۴ \\ -۰/۲ & ۹ & -۴/۴ \\ ۱ & -۲/۵ & ۵ \end{bmatrix}$$

حال مجدداً $A_۲$ باید به صورت $L_۲U_۲$ تجزیه شود. یعنی $A_۲ = L_۲U_۲$ و به همین ترتیب باید ادامه داد. چنانچه محاسبات تا ۴ رقم بعد از اعشار صورت گیرد، آنگاه:

$$A_۹ = \begin{bmatrix} ۳۱/۱۵۹۴ & -۱/۶۷۴۲ & ۴/۰۰۰۰ \\ -۰/۰۰۰۲ & ۱۰/۸۱۳۱ & -۴/۳۳۲۳ \\ ۰/۰۰۰۰ & -۰/۰۰۰۲ & ۳/۰۲۷۴ \end{bmatrix} \Rightarrow$$

$$A_{۱۰} = \begin{bmatrix} ۳۱/۱۵۹۴ & -۱/۶۷۴۲ & ۴/۰۰۰۰ \\ -۰/۰۰۰۱ & ۱۰/۸۱۳۱ & -۴/۳۳۲۳ \\ ۰/۰۰۰۰ & ۰/۰۰۰۰ & ۳/۰۲۷۳ \end{bmatrix} \Rightarrow$$

$$A_{۱۱} = \begin{bmatrix} ۳۱/۱۵۹۴ & -۱/۶۷۴۳ & ۴/۰۰۰۰ \\ ۰/۰۰۰۰ & ۱۰/۸۱۳۳ & -۴/۳۳۲۳ \\ ۰/۰۰۰۰ & ۰/۰۰۰۰ & ۳/۰۲۷۳ \end{bmatrix}$$

ماتریس $A_{۱۱}$ یک ماتریس بالا مثلثی است. بنابراین عناصر روی قطر اصلی آن همان مقادیر ویژه ماتریس A می‌باشند. یعنی $\lambda_۳ = ۳/۰۲۷۳$ و $\lambda_۲ = ۱۰/۸۱۳۳$ و $\lambda_۱ = ۳۱/۱۵۹۴$.

۵-۲-۵ روش تجزیه QR

اساس این روش بر مبنای تجزیه QR ماتریس A بنا شده است. همانند بخش (۵-۲-۴)، باید $A_1 = A$ قرار داده شود و A_1 به صورت Q_1R_1 تجزیه شود. حال ماتریس جدید A_2 باید به صورت زیر تعریف شود:

$$A_2 = R_1Q_1$$

ماتریس‌های A_1 و A_2 متشابه هستند، پس مقادیر ویژه‌ی آن‌ها یکسان خواهد بود. زیرا:

$$\left.\begin{array}{r} A_2 = R_1Q_1 \\ Q_1^TQ_1 = I \end{array}\right\} \Rightarrow A_2 = IR_1Q_1 = Q_1^T \underbrace{Q_1R_1}_{A_1} Q_1 = Q_1^TA_1Q_1$$

با ادامه همین روند در مرحله k ام، A_k باید به صورت $A_k = Q_kR_k$ تجزیه شود. حال باید قرار داده شود:

$$A_{k+1} = R_kQ_k$$

در نتیجه:

$$A_{k+1} = Q_k^TA_kQ_k$$

می‌توان نشان داد که دنباله A_{k+1} ، وقتی $k \to \infty$ به یک ماتریس بالامثلثی میل می‌کند؛ لذا برای k های به قدر کافی بزرگ، مقادیر ویژه‌ی A همان عناصر روی قطر اصلی A_{k+1} خواهد بود.

۵-۳ کاربرد مقدار ویژه در علوم داده

بردارهای ویژه، کاربردهای بسیار زیادی در زمینه‌های مختلف مانند بینایی رایانه[1]و یادگیری ماشین دارند. مفهوم بردار ویژه در زمینه یادگیری ماشین، برای الگوریتم تحلیل مؤلفه‌های اساسی به کار برده می‌شود. اگر در مورد یادگیری ماشین مطالعه کرده و با الگوریتم تحلیل مؤلفه‌های اساسی آشنا باشید، حتماً از اهمیت این الگوریتم در هنگام کنترل یک مجموعه داده بزرگ باخبر هستید. فرض کنید داده‌ای با ویژگی‌های زیاد دارید (ابعاد بسیار زیاد در داده). امکان وجود ویژگی‌های اضافی در داده وجود دارد. به غیر از این، تعداد بسیار زیاد ویژگی‌ها، باعث کاهش کارآمدی خواهد شد و فضای اشغال شده از رایانه بیشتر می‌شود. کاری که تحلیل مؤلفه‌های اساسی انجام می‌دهد، کنار گذاشتن ویژگی‌های کم‌اهمیت و ناکارآمد است. برای مشخص کردن ویژگی‌های قابل حذف نیز می‌توان از بردارهای ویژه استفاده کرد. در بخش (۴-۴) به تحلیل مولفه‌های اساسی اشاره شد و در ادامه به‌طور کامل در مورد آن بحث خواهد شد.

[1] Computer Vision

۵-۳-۱ کاربرد تحلیل مولفه‌های اساسی در علوم داده

همان‌طور که گفته شد اندازه داده‌ها در عصر مدرن، نه تنها یک چالش برای سخت‌افزارهای کامپیوتر، بلکه تنگنایی برای کارای بسیاری از الگوریتم‌های یادگیری ماشین محسوب می‌شود. هدف اصلی تحلیل مولفه‌های اساسی، شناسایی الگوهای موجود در داده‌ها است؛ یا به عبارت دیگر تحلیل مولفه اساسی قصد دارد که همبستگی بین متغیرها را شناسایی کند. اگر یک همبستگی قوی بین متغیرها وجود داشت، تلاش‌ها برای کاهش ابعاد داده معنادار خواهد بود. به‌طور کلی، آنچه در تحلیل مولفه اساسی به وقوع می‌پیوندد پیدا کردن جهت واریانس بیشینه در داده‌های ابعاد بالا و طرح‌ریزی کردن آن در زیرفضایی با ابعاد کمتر است به طوری که بیشترین اطلاعات حفظ شود.

تحلیل مولفه‌های اساسی یک روش تبدیل خطی[۱]ساده و در عین حال محبوب و کارآمد محسوب می‌شود. از این روش در کاربردهایی مانند پیش‌بینی بازار بورس، تحلیل داده‌های بیان ژن و بسیاری از دیگر موارد استفاده می‌شود. از جمله روش‌های مشابه تحلیل مولفه اساسی، می‌توان به تحلیل تشخیصی خطی[۲]و تجزیه مقادیر تکین ذکر شده در فصل قبل اشاره کرد.

روش تحلیل مولفه‌های اساسی جهت‌هایی که واریانس[۳]داده‌ها را بیشینه می‌کنند (مولفه اساسی) می‌یابد. به بیان دیگر، تحلیل مولفه‌های اساسی کل مجموعه داده را در یک (زیر)فضای دیگر طرح‌ریزی می‌کند. اغلب، هدف مورد انتظار کاهش ابعاد یک مجموعه داده d بعدی با طرح‌ریزی آن در یک زیرفضای k بعدی (که در آن $k < d$ است)، به منظور افزایش بازدهی محاسباتی (با حفظ بخش مهم اطلاعات) می‌باشد.

بردارهای ویژه (مولفه‌های اساسی) برای مجموعه داده محاسبه شده و همه آن‌ها در یک ماتریس تصویر[۴]گردآوری می‌شوند. به هر یک از این بردارهای ویژه یک مقدار ویژه تخصیص داده می‌شود که می‌تواند به عنوان طول یا بزرگنمایی بردار ویژه متناظر در نظر گرفته شود. اگر برخی از مقدارهای ویژه دارای بزرگنمایی به طور قابل توجهی بزرگ‌تر از دیگر موارد باشند، کاهش مجموعه داده با تحلیل مولفه‌های اساسی به یک زیرفضای ابعاد کوچک‌تر با حذف جفت ویژه‌هایی با اطلاعات کمتر معقول است. تحلیل مولفه‌های اساسی برای داده‌های دارای سه یا تعداد بیشتری بُعد کاربردپذیرتر است. همچنین این روش همیشه روی ماتریس کوواریانس[۵]یا ماتریس همبستگی[۶]اعمال می‌شود؛ پس داده‌ها باید عددی و استاندارد شده باشند.

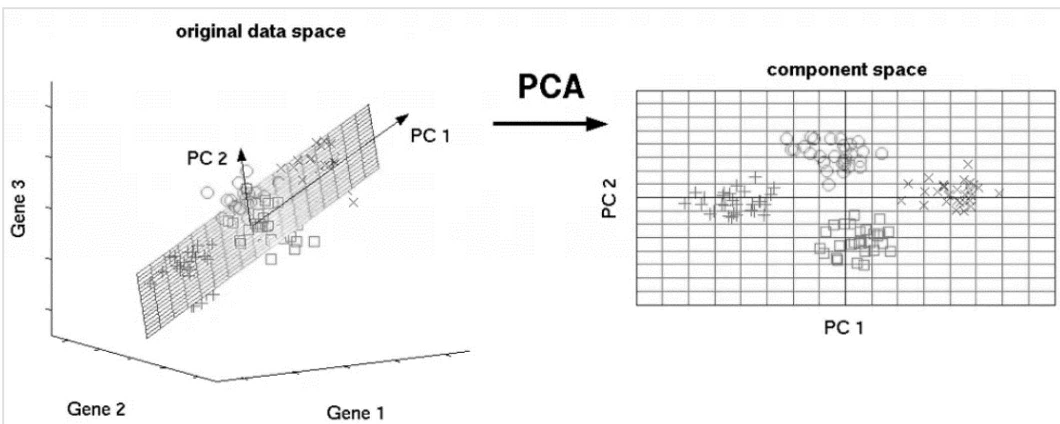
برای درک بهتر این روش، در ادامه یک مثال بیان شده است.

---

[۱] Linear Transformation Technique
[۲] Linear Discriminant Analysis (LDA)
[۳] Variance
[۴] Projection Matrix
[۵] Covariance Matrix
[۶] Correlation Matrix

مثال ۵-۳۱) فرض کنید که یک مجموعه داده با ابعاد $۳۰۰(n) \times ۵۰(p)$ موجود باشد. در این مجموعه داده n تعداد کل نمونه‌ها و p تعداد متغیرهای پیش‌بینی (یا همان متغیرهای هدف) است. به دلیل تعداد ابعاد زیاد ($p =$ ۵۰)، می‌توان $\frac{p(p-۱)}{۲}$ نمودار پراکندگی[1]برای آن رسم کرد، این یعنی بیش از ۱۰۰۰ نمودار برای انجام تحلیل روی روابط بین متغیرها وجود دارد و در نتیجه تحلیل آن‌ها کاری بسیار خسته‌کننده، دشوار و پیچیده خواهد بود.

در این شرایط یک رویکرد صحیح می‌تواند آن باشد که یک زیر مجموعه از پیش‌بین‌ها (متغیرهای پیش‌بینی) که حاوی بیشترین اطلاعات درباره داده‌ها هستند، انتخاب شود. این امر موجب می‌شود نمودار پراکندگی داده‌ها در ابعاد پایین‌تری قابل ترسیم باشد. تصویر زیر نگاشت داده‌های دارای ابعاد بالا (۳ بُعد) را به داده‌های با ابعاد پایین‌تر (۲ بُعد) با استفاده از روش تحلیل مولفه اساسی (اصلی) نشان می‌دهد. لازم به ذکر است هر بُعد حاصل شده در فضای جدید، یک ترکیب خطی از p ویژگی اصلی است.



شکل (۵-۱) کاهش ابعاد داده‌ها با استفاده از روش تحلیل مولفه اساسی

ایده اصلی این است که k مولفه از یک مجموعه داده متشکل از n نمونه و p متغیر هدف به طریقی انتخاب شوند که بیشتر اطلاعات موجود در p متغیر اصلی، در k مولفه انتخاب شده موجود باشند. در آمار این k مولفه، اولین k مولفه اصلی نامیده می‌شوند. شناخت مقادیر ویژه و بردارهای ویژه ماتریس کوواریانس برای پیدا کردن این مولفه‌های اصلی لازم می‌باشد. به ویژه اگر Σ ماتریس کوواریانس متناظر با بردارهای تصادفی زیر باشد:

$$X = (X_۱, X_۲, \cdots, X_p)$$

و همچنین $\lambda_۱ \geq \lambda_۲ \geq \cdots \geq \lambda_p$ مقادیر ویژه باشند و $x_۱$ تا x_p بردارهای ویژه ماتریس Σ باشند، آنگاه i امین مولفه اصلی به صورت $Y_i = x_i^T X$ که در آن $i = ۱,۲, \cdots, p$ می‌باشد، ارائه می‌شود.

[1] Scatter Plot

به‌علاوه قسمتی از کل واریانس جامعه که از i امین مولفه اصلی ناشی می‌شود توسط نسبت زیر بیان می‌گردد:

$$\frac{\lambda_i}{\lambda_۱ + \lambda_۲ + \cdots + \lambda_p} = \frac{\lambda_i}{trace(\Sigma)}\ ,\ \ i = ۱,۲,\cdots,p$$

تذکر ۵-۳۲) ماتریس کوواریانس ماتریسی معین مثبت متقارن است؛ بنابراین مقادیر ویژه آن همگی نامنفی هستند.

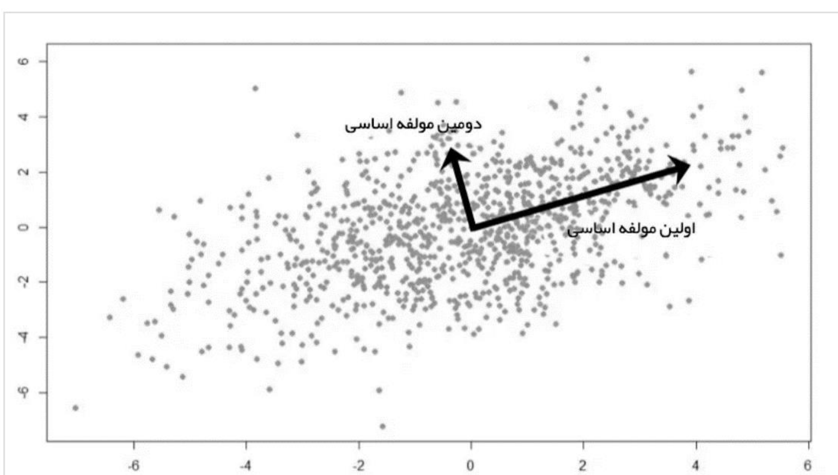
تذکر ۵-۳۳) اگر k نسبت اول بیشترین مقدار کل واریانس جامعه را تشکیل دهند، آنگاه k مولفه اساسی می‌توانند در تحلیل‌های آماری مورد استفاده قرار گیرند.

تعریف ۵-۳۴) اولین مولفه اساسی، یک ترکیب خطی از متغیرهای پیش‌بینی است که بیشترین واریانس موجود در مجموعه داده‌ها را در بر می‌گیرد.

اولین مولفه اساسی، جهت بیشترین تغییرات در داده‌ها را تعیین می‌کند. هرچه دامنه تغییرات موجود در اولین مولفه بالاتر باشد، اطلاعات موجود در این مولفه بیشتر است. هیچ مولفه دیگری نمی‌تواند بیش از مولفه اساسی اول دامنه تغییرات داشته باشد. نتیجه محاسبه اولین مولفه اساسی، خطی است که نزدیک‌ترین خط به داده‌ها محسوب می‌شود. در واقع این خط مجموع مربع فواصل را بین یک نقطه داده و خط، به کمینه مقدار می‌رساند.

تعریف ۵-۳۵) دومین مولفه اساسی، یک ترکیب خطی از متغیرهای پیش‌بینی است که واریانس باقی‌مانده در مجموعه داده را در خود حفظ می‌کند.

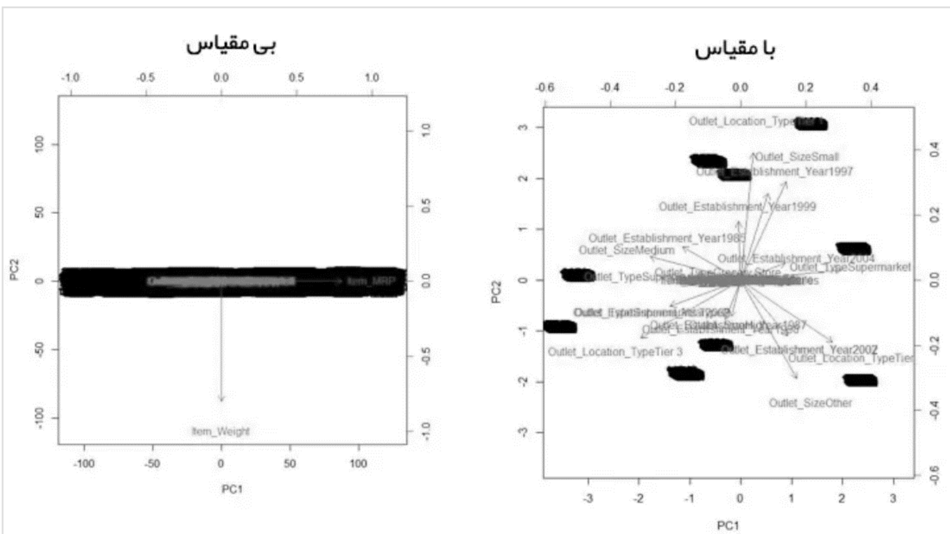
تذکر ۵-۳۶) همبستگی بین اولین و دومین مولفه اساسی صفر است. یا به عبارت دیگر، جهت‌های آن‌ها همانند شکل (۵-۲) باید متعامد باشند.



شکل (۵-۲) همبستگی مولفه اساسی اول و دوم

نکته دیگری که حائز اهمیت می‌باشد این است که تحلیل مولفه‌های اساسی روی نسخه نرمال شده‌ی متغیرها (نسخه‌ای که داده‌ها با حفظ نسبت بین صفر و ۱ قرار خواهند گرفت) قابل انجام است. این امر به جهت احتمال وجود مقیاس‌های گوناگون در میان داده‌ها باید انجام گیرد. به عنوان مثال می‌توان به یک مجموعه داده که شامل متغیرهایی با یکاهای میلی‌متر، کیلومتر، سال نوری و انواع دیگر واحدها است، اشاره کرد. واضح است که مقدار واریانس این متغیرها اعداد بزرگی خواهد بود. انجام روش تحلیل مولفه اساسی روی متغیرهای نرمال نشده منجر به ایجاد مقادیر بسیار بزرگی برای متغیرهای دارای واریانس بالا می‌شود و این امر می‌تواند منجر به وابستگی مولفه‌های اساسی شود که این امر بسیار نامطلوب است.

در شکل (۵-۳)، روش تحلیل مولفه‌های اساسی دو بار روی مجموعه داده‌ای مفروض اجرا شده است. در قسمت سمت راست، این روش روی داده‌های نرمال شده و در قسمت سمت چپ روی داده‌های نرمال نشده اجرا شده است. همان‌طور که مشاهده می‌شود، هنگامی‌که متغیرها نرمال شوند، بصری‌سازی آن‌ها در فضای دو بُعدی به شکل بهتری انجام‌پذیر است.



شکل (۵-۳) تحلیل مولفه اساسی بر روی داده‌های نرمال‌شده و غیرنرمال

تذکر ۵-۳۷) روش تحلیل مولفه‌های اساسی تنها روی مجموعه داده‌هایی با ویژگی‌های عددی قابل اعمال است. پس اگر متغیر موجود از نوع کیفی اسمی و یا کیفی ترتیبی بود، باید آن را تبدیل به داده‌های کمی کرد.

با توجه به مطالبی که تا این‌جا گفته شد، اول از همه باید داده‌ها را نرمال‌سازی کرد تا در صورت وجود مقیاس‌های گوناگون، همگی به یک مقیاس واحد برسند. پس باید از روش زیر پیروی کرد تا روش تحلیل مولفه‌های اساسی تکمیل گردد.

فرض کنید، یک مجموعه داده‌ی n بعدی موجود باشد. هدف کاهش ابعاد این مجوعه داده به k بعد می‌باشد. باید مراحل زیر را دنبال کرد:

۱) پیدا کردن ماتریس کوواریانس مجموعه داده‌ها (به حداقل رساندن از دست رفتن اطلاعات معادل است با حداکثر کردن واریانس)

۲) یافتن بردارهای ویژه ماتریس کوواریانس

۳) مرتب‌سازی مقادیر ویژه به‌ترتیب نزولی و انتخاب k بردار ویژه مربوط به k تا از بزرگ‌ترین مقادیر ویژه انتخاب‌شده

۴) ساخت ماتریس W از k بردار ویژه انتخاب‌شده

۵) تبدیل مجموعه داده اصلی X به‌وسیله W، به زیرفضای k بعدی Y.

مثال ۵-۳۸) فرض کنید که ماتریس کوواریانس برای نرخ‌های هفتگی برگشت برای اوراق بهادار پنج شرکت بزرگ به نام‌های Allied Chemical، Dupont، Union Carbide، Exxon و Texaco در یک دوره مفروض از زمان به صورت زیر باشد.

$$R=\begin{bmatrix} ۱/۰۰۰ & ۰/۵۷۷ & ۰/۵۰۹ & ۰/۳۸۷ & ۰/۴۶۲ \\ ۰/۵۷۷ & ۱/۰۰۰ & ۰/۵۹۹ & ۰/۳۸۹ & ۰/۳۲۲ \\ ۰/۵۰۹ & ۰/۵۹۹ & ۱/۰۰۰ & ۰/۴۳۶ & ۰/۴۲۶ \\ ۰/۳۸۷ & ۰/۳۸۹ & ۰/۴۳۶ & ۱/۰۰۰ & ۰/۵۲۳ \\ ۰/۴۶۲ & ۰/۳۲۲ & ۰/۴۲۶ & ۰/۵۲۳ & ۱/۰۰۰ \end{bmatrix}$$

دو مقدار ویژه اول R عبارتند از $\lambda_۱ = ۲/۸۵۷$ و $\lambda_۲ = ۰/۸۰۹$.

قسمتی از کل واریانس جامعه که از اولین مولفه نتیجه می‌گردد برابر است با

$$\frac{۲/۸۵۷}{۵} = ۵۷٪$$

قسمتی از کل واریانس جامعه که از دومین مولفه نتیجه می‌گردد برابر است با

$$\frac{۰/۸۰۹}{۵} = ۱۶٪$$

بنابراین دو مولفه اصلی اول ۷۳٪ از کل واریانس جامعه را بیان می‌کنند. بردارهای ویژه متناظر با این مولفه‌های اصلی عبارتند از:

$$x_۱^T = (۰/۴۶۴, ۰/۴۵۷, ۰/۴۷۰, ۰/۴۲۱, ۰/۴۲۱)$$
$$x_۲^T = (۰/۲۴۰, ۰/۵۰۹, ۰/۲۶۰, -۰/۵۲۶, -۰/۵۸۲)$$

این بردارهای ویژه تعبیرهای جالبی دارند. از عبارت x_1 ملاحظه می‌شود که اولین مولفه (به طور تقریبی) مجموع موزون به طور مساوی پنج اوراق بهادار می‌باشد. این مولفه معمولاً مولفه بازار نامیده می‌شود. ولیکن عبارت مربوط به x_2 نشان‌دهنده این است که مولفه دوم یک تضاد را بین سهام مواد شیمیایی و سهام صنایع نفتی ظاهر می‌سازد. این مولفه معمولاً مولفه صنعتی نامیده می‌شود. بنابراین نتیجه می‌توان گرفت که حدود ۵۷٪ از کل تغییرات در این برگشت‌های سهام از عملکرد بازار و ۱۶٪ از عملکرد صنعت ناشی می‌شود.

واژه‌نامه انگلیسی به فارسی

((A))

ماتریس الحاقی................................Adjugate Matrix
الگوریتم...Algorithm
ماتریس پادمتقارن................Antisymmetric Matrix
شبکه عصبی مصنوعی....Artifical Neural Network
ماتریس افزوده............................Augmented Matrix

((B))

جایگذاری پسرو.............................Back Substitution
کیسه‌ی کلمات....................................Bag of Words
کلان داده..Big Data
هوش تجاری............................Business Intelligence

((C))

چندجمله‌ای مشخصه....Characteristic Polynomial
خوشه‌بندی...Clustering
همسازه...Cofactor
ماتریس ضرایب..............................Cofficient Matrix
ستون..Column
سازگاری...Compatibility
رتبه کامل..Complete Rank
بینایی ماشین..................................Computer Vision
سازگاری...Consistency
همگرایی..Convergence
ضرب خارجی..Cross Product
همبستگی...Correlation
ماتریس کوواریانس......................Covariance Matrix

((D))

داده...Data
تحلیلگر داده..Data Analyst
داده محور..Data Driven
مجموعه داده..Data Set
دانشمند داده...Data Scientist
دترمینان..Determinant
ماتریس قطری................................Diagonal Matrix
قطری‌سازی.......................................Diagonalisation
ماتریس قطر غالب..Diagonally Dominant Matrix
بُعد..Dimension
کاهش ابعاد....................Dimensionality Reduction

((E))

مقدار ویژه...Eigenvalue
بردار ویژه..Eigenvector

Elementary Matrix..........................ماتریس مقدماتی

Elimination Method...............................روش حذفی

Entry..درایه (عنصر)

Euclidean Norm.....................................نرم اقلیدسی

Expansion...بسط

((F))

Factor Analysis..تجزیه عامل

Factorization...تجزیه

Forward Substitution......................جایگذاری پیشرو

((H))

High Dimensional Data............داده‌های با ابعاد بالا

((I))

Identity Matrix....................................ماتریس همانی

Image Processing...............................پردازش تصویر

Independence...استقلال

Inner Product...ضرب داخلی

Internet of Things..................................اینترنت اشیا

Iterative Method...................................روش تکراری

Inverse Matrix......................................ماتریس وارون

Invertible Matrix..........................ماتریس وارون‌پذیر

Iterative Refinement...........................تصفیه تکراری

((L))

Leading Principal Submatrixزیرماتریس‌اصلی‌پیشرو

Least square..کمترین مربعات

Linear Algebra..جبر خطی

Linear Discriminant Analysisتحلیل تشخیصی خطی

Linear Regression.............................رگرسیون خطی

Linear System......................................دستگاه خطی

Linear Transformation..........................تبدیل خطی

Lower Hesenberg Matrix.ماتریس پایین‌هسنبرگی

Lower Triangular Matrix.......ماتریس پایین مثلثی

((M))

Machine Learning............................یادگیری ماشین

Main Diagonal...قطر اصلی

Matrix..ماتریس

Matrix Norm...نرم ماتریسی

Matrix of Cofactors......................ماتریس همسازه‌ها

Matrix Power Method..............روش توان ماتریسی

Method...روش

Minor..کهاد

((N))

Nilpotent Matrix..............................ماتریس پوچ‌توان

Norm...نرم

((O))

Orthogonal..متعامد

Orthogonal Matrix.............................ماتریس متعامد

Orthogonality...تعامد

Orthogonalization..................................متعامدسازی

Over Determined...فرا معین

((P))

Positive Definite Matrix.........ماتریس معین مثبت

Power Method...روش توانی

Power Inverse Method.............روش معکوس توانی

Power Reduction Method............روش تقلیل توانی

Principal Component Analysis..تحلیل مولفه اساسی

Projection...طرح‌ریزی

Projection Matrix................................ماتریس تصویر

Pseudoinverse..شبه‌وارون

((Q))

Quadratic form......................................فرم درجه دوم

((R))

Rank..رتبه

Rate of Convergence.......................سرعت همگرایی

Reflection...بازتاب

Regression..رگرسیون

Rotation..دوران (چرخش)

Rotation Matrix...................................ماتریس دوران

Row...سطر

Row Equivalence..................................هم‌ارز سطری

Row Reduced..سطری-پلکانی

Row Reduced Echelon...سطری-پلکانی تحویل یافته

((S))

Sampling..نمونه‌برداری

Scatter Plot..نمودار پراکندگی

Similarity Matrix...............................ماتریس متشابه

Singular..منفرد (تکین)

Singular Value Decomposition..تجزیه مقدار تکین

Sketching..بازطراحی داده‌ها

Spectral Radius..شعاع طیفی

Square Matrix......................................ماتریس مربعی

Streaming................................انتقال جریانی از داده‌ها

Subspace..زیرفضا

Substitution..جایگذاری

Successive Over Relaxation...فوق تخفیف متوالی

Successive Under Relaxation..تحت تخفیف متوالی

Symmetry..تقارن

Symmetric Matrix.............................ماتریس متقارن

System of Linear Equations.دستگاه معادلات خطی

((T))

Term Document Matrix..............ماتریس لغت-سند

Trace...اثر

Transformation..تبدیل

Transformation Method.......................روش تبدیلی

Transpose...ترانهاده

Tridigonal Matrix...........................ماتریس ۳ قطری

Tringular Matrix.................................ماتریس مثلثی

((U))

Under Determined...................زیر معین (فرو معین)

Unitary Matrix.....................................ماتریس یکانی

Upper Hesenberg Matrix....ماتریس بالا هسنبرگی

Upper Tringular Matrix..............ماتریس بالا مثلثی

((V))

Variance...واریانس

Vector...بردار

Vector Norm..نرم برداری

Vector Space...فضای برداری

مراجع

[۱] فرقانی، مجید.، رئیسی دهکردی، هنگامه.، مبانی ماتریس‌ها و جبر خطی کارشناسی ارشد، تهران، انتشارات مدرسان شریف، ۱۳۹۳.

[۲] نعمت‌الهی، نادر.، روش‌های آماری، تهران، انتشارات دانشگاه علامه طباطبائی، ۱۳۹۷.

[۳] Datta, B. N., Numerical Linear Algebra and Applications, Philadelphia, Society for Industrial and Applied Mathematics, ۲۰۱۰.

[۴] Higham, N. J., Accuracy and Stability of Numerical Algorithms, Philadelphia, Society for Industrial and Applied Mathematics, ۱۹۶۱.

[۵] Saad, Y., Iterative Methods for Sparse Linear Systems, Philadelphia, Society for Industrial and Applied Mathematics, ۲۰۰۳.

[۶] Trefethen, L. N., Bau, D., Numerical Linear Algebra, Philadelphia, Society for Industrial and Applied Mathematics, ۱۹۹۷.

University of Isfahan

Faculty of Mathematics and Statistics

Department of Mathematics

B.Sc. Project in Mathematics and its Applications

Linear Algebra for Data Science

Supervisor:

Dr. Yousofzadeh

By:

Parham Pishro