**دانشگاه صنعتی امیرکبیر**
**( پلی تکنیک تهران )**

# جبرخطی عددی

(کارشناسی)

## فصل پنجم:روش های مستقیم و تکراری برای محاسبه مقادیر ویژه و بردار ویژه یک ماتریس

**مدرس: مهدی دهقان**

دانشکده
ریاضی و علوم کامپیوتر

ترم دوم ۱۴۰۳-۱۴۰۲

# فهرست مطالب

# ۱ روش های محاسبه مقادیر ویژه و بردار ویژه

در فصل صفر دیدیم که $(\lambda, X)$ را یک جفت ویژه ماتریس مربعی $A$ گوییم هر گاه رابطه‌ی

$$AX = \lambda X, \qquad X \neq ۰$$

برقرار باشد. بعلاوه ثابت کردیم که هر ماتریس $n \times n$، $A$ دارای $n$ مقدار ویژه می‌باشد که درواقع ریشه‌های چندجمله‌ای

$$P_A(\lambda) = \det(\lambda I - A) \qquad or \qquad P_A(\lambda) = \det(A - \lambda I)$$

می‌باشند. بنابراین روش سرراست محاسبه‌ی مقادیر ویژه‌ی یک ماتریس شامل دو مرحله خواهد بود:

۱) محاسبه‌ی چندجمله‌ای ویژه $P_A(\lambda)$
۲) محاسبه‌ی ریشه‌های $P_A(\lambda)$

با توجه به اینکه مسأله‌ی ریشه‌یابی توابع چندجمله‌ای از جمله مسائل بدوضع است بنابراین محاسبه‌ی مقادیر ویژه‌ی یک ماتریس با استفاده از مراحل فوق توصیه نمی‌شود. برای دیدن این موضوع به مثال زیر توجه کنید.

**مثال ۵.۱**

چندجمله‌ای زیر را در نظر بگیرید.

$$P(\lambda) = (\lambda - ۱)(\lambda - ۲)(\lambda - ۳)(\lambda - ۴)(\lambda - ۵)$$

واضح است که ریشه‌های این چندجمله‌ای به صورت

$$\lambda_۱ = ۱, \quad \lambda_۲ = ۲, \quad \lambda_۳ = ۳, \quad \lambda_۴ = ۴, \quad \lambda_۵ = ۵$$

می‌باشند. با بسط دادن جملات این چندجمله‌ای خواهیم داشت:

$$P(\lambda) = \lambda^۵ - ۱۵\lambda^۴ + ۸۵\lambda^۳ - ۲۲۵\lambda^۲ + ۲۷۴\lambda - ۱۲۰$$

حال فرض کنید یک اختلال کوچک به اندازه‌ی $۰/۰۲$ به ضریب $\lambda^۴$ وارد گردد و ضریب $-۱۵$ به $-۱۵/۰۲$ تغییر یابد یعنی $P$ به $\tilde{P}$ تبدیل گردد:

$$\tilde{P}(\lambda) = \lambda^۵ - ۱۵/۰۲\lambda^۴ + ۸۵\lambda^۳ - ۲۲۵\lambda^۲ + ۲۷۴\lambda - ۱۲۰$$

حال با محاسبه‌ی ریشه‌های $\tilde{P}$ داریم $(i = \sqrt{-۱})$

$$\tilde{\lambda}_۱ = ۱/۰۰۰۸, \qquad \tilde{\lambda}_۲ = ۱/۹۵۳۲, \qquad \tilde{\lambda}_۳ = ۳/۳۵۶۳ + ۰/۴۴۹۴i$$
$$\tilde{\lambda}_۴ = ۳/۳۵۶۳ - ۰/۴۴۹۴i, \qquad \tilde{\lambda}_۵ = ۵/۳۵۳۳$$

مشاهده می‌شود که مقادیر بعضی از ریشه‌ها تغییرات بسیاری داشته است. حتی دو ریشه‌ی این چندجمله‌ای مختلط گردیده‌اند!

مثال قبل نشان می‌دهد که برای محاسبه‌ی مقادیر ویژه، اینکه ابتدا چندجمله‌ای ویژه را محاسبه کنیم و سپس ریشه‌های آن را به دست آوریم، نمی‌تواند راه مطمئنی باشد، به دلیل اینکه یک چندجمله‌ای ممکن است نسبت به تغییرات ضرایب خودش بسیار حساس می‌باشد.

به طور کلی روش‌هایی که مبنای آن‌ها ابتدا محاسبه‌ی چندجمله‌ای ویژه ماتریس $A$ و سپس ریشه‌یابی آن می‌باشد را می‌توانیم جز روش‌های مستقیم برای محاسبه‌ی مقادیر ویژه در نظر بگیریم.

قبل از ادامه‌ی بحث به مسائلی اشاره می‌کنیم که نیازمند حل یک مسأله مقدار ویژه

$$AX = \lambda X, \qquad X \neq 0$$

یعنی محاسبه‌ی $\lambda$ و $X$ می‌باشد.

**مثال ۵.۲**

(حل یک دستگاه معادلات دیفرانسیل) دستگاه معادلات دیفرانسیل زیر را در نظر بگیرید

$$\begin{cases} x'(t) = a_{11}x(t) + a_{12}y(t) \\ y'(t) = a_{21}x(t) + a_{22}y(t) \end{cases} \qquad (1)$$

اگر تعریف کنیم

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \qquad X'(t) = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

آنگاه دستگاه معادلات دیفرانسیل (۱) به فرم ماتریسی–برداری

$$X'(t) = AX(t) \qquad (2)$$

نوشته می‌شود. حال با فرض اینکه (۲) جوابی به صورت

$$X(t) = Ve^{\lambda t} \qquad (\text{یک بردار مخالف صفر است } V)$$

داشته باشد داریم

$$X'(t) = \lambda Ve^{\lambda t}$$

پس با قرار دادن $X(t)$ و $X'(t)$ در (۲) داریم

$$\lambda Ve^{\lambda t} = AVe^{\lambda t}, \qquad V \neq 0 \qquad (3)$$

چون برای هر $\lambda$ و $t$ داریم $e^{\lambda t} \neq 0$. پس با تقسیم طرفین (۳) بر $e^{\lambda t}$ داریم

$$\lambda V = AV$$

یا

$$AV = \lambda V, \qquad V \neq 0 \qquad (4)$$

که یک مسأله مقدار ویژه است. از حل (۴) و به دست آوردن جفت ویژه $(\lambda, V)$ می‌توان یک جواب دستگاه معادلات (۱) را به صورت $X(t) = Ve^{\lambda t}$ محاسبه نمود.

**مثال ۵.۳**

دستگاه معادلات دیفرانسیل داده شده را حل کنید.

$$\begin{cases} x'(t) = x(t) - ۳y(t) \\ y'(t) = ۴x(t) - ۶y(t) \end{cases}$$

**حل:** واضح است که

$$A = \begin{bmatrix} ۱ & -۳ \\ ۴ & -۶ \end{bmatrix}$$

باید مسأله (۴) را حل کنیم

$$AV = \lambda V, \qquad V \neq ۰$$

یعنی

$$\begin{bmatrix} ۱ & -۳ \\ ۴ & -۶ \end{bmatrix} V = \lambda V, \qquad V \neq ۰$$

برای حل مسأله فوق کافی است مقادیر ویژه $\lambda$ را از حل چندجمله‌ای ویژه‌ی زیر به دست آوریم:

$$\begin{aligned} P_A(\lambda) &= \det(\lambda I - A) = \det\left(\begin{bmatrix} \lambda & ۰ \\ ۰ & \lambda \end{bmatrix} - \begin{bmatrix} ۱ & -۳ \\ ۴ & -۶ \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} \lambda - ۱ & ۳ \\ -۴ & \lambda + ۶ \end{bmatrix}\right) = (\lambda - ۱)(\lambda + ۶) - (۳)(-۴) \\ &= \lambda^۲ + ۶\lambda - \lambda - ۶ + ۱۲ = \lambda^۲ + ۵\lambda + ۶ = (\lambda + ۲)(\lambda + ۳) \end{aligned}$$

بنابراین دو مقدار ویژه $A$ به صورت زیر می‌باشند

$$\lambda_۱ = -۲, \qquad \lambda_۲ = -۳$$

برای محاسبه‌ی بردار ویژه متناظر با $\lambda_۱ = -۲$ داریم:

$$\begin{aligned} &AV_۱ = \lambda_۱ V_۱ \Rightarrow \begin{bmatrix} ۱ & -۳ \\ ۴ & -۶ \end{bmatrix} \begin{bmatrix} v_۱ \\ v_۲ \end{bmatrix} = -۲ \begin{bmatrix} v_۱ \\ v_۲ \end{bmatrix} \\ &\Rightarrow \begin{cases} v_۱ - ۳v_۲ = -۲v_۱ \\ ۴v_۱ - ۶v_۲ = -۲v_۲ \end{cases} \Rightarrow \begin{cases} v_۱ - ۳v_۲ = -۲v_۱ \\ ۲v_۱ - ۳v_۲ = -v_۲ \end{cases} \Rightarrow v_۱ = v_۲ \end{aligned}$$

با انتخاب $v_۱ = v_۲ = ۱$ داریم

$$V_۱ = \begin{bmatrix} v_۱ \\ v_۲ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۱ \end{bmatrix}$$

بنابراین یک جفت ویژه $A$ به صورت زیر مشخص می‌شود

$$(\lambda_1, V_1) = \left(-2, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

با داشتن این جفت ویژه یک جواب دستگاه معادلات دیفرانسیل داده شده به صورت زیر محاسبه می‌گردد

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = V_1 e^{\lambda_1 t} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} = \begin{bmatrix} e^{-2t} \\ e^{-2t} \end{bmatrix}$$

لذا $x(t) = e^{-2t}, y(t) = e^{-2t}$ یک جواب مسأله داده شده است.

**تمرین ۵.۱**

با محاسبه‌ی جفت ویژه‌ی دیگر ماتریس $A$ جوابی دیگر برای دستگاه معادلات داده شده بیابید.

همانطور که در مثال قبل دیدیم برای حل مسأله داده شده نیاز به مقادیر دقیق مقادیر ویژه بود. در بعضی از مسائل تنها داشتن کران بالا و یا پایینی از مقادیر ویژه کفایت می‌کند. به مثال بعد توجه کنید:

**مثال ۵.۴**

روش تکراری زیر را

$$X^{(k+1)} = \underbrace{\begin{bmatrix} 0 & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{3} & -\frac{1}{3} & 0 \end{bmatrix}}_{M} X^{(k)} + \begin{bmatrix} \frac{2}{3} \\ \frac{1}{2} \\ \frac{2}{3} \end{bmatrix}$$

برای حل دستگاه

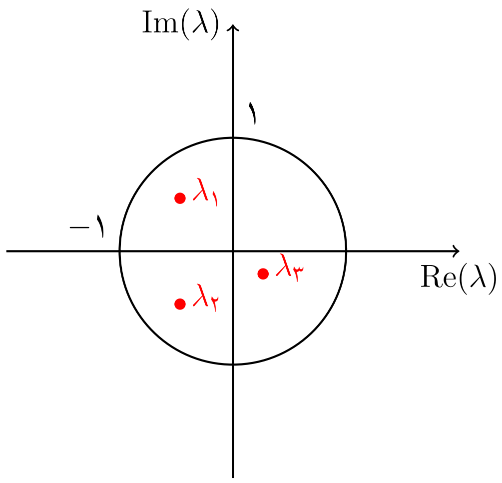
$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 4 & 1 \\ -1 & 1 & 3 \end{bmatrix} X = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

در نظر بگیرید. همگرایی روش فوق را بررسی نمایید.

حل: برای همگرایی روش فوق لازم است که شعاع طیفی ماتریس تکرار کمتر از یک باشد، به‌عبارتی اگر $\lambda_1, \lambda_2, \lambda_3$ مقادیر ویژه ماتریس $M$ باشند باید

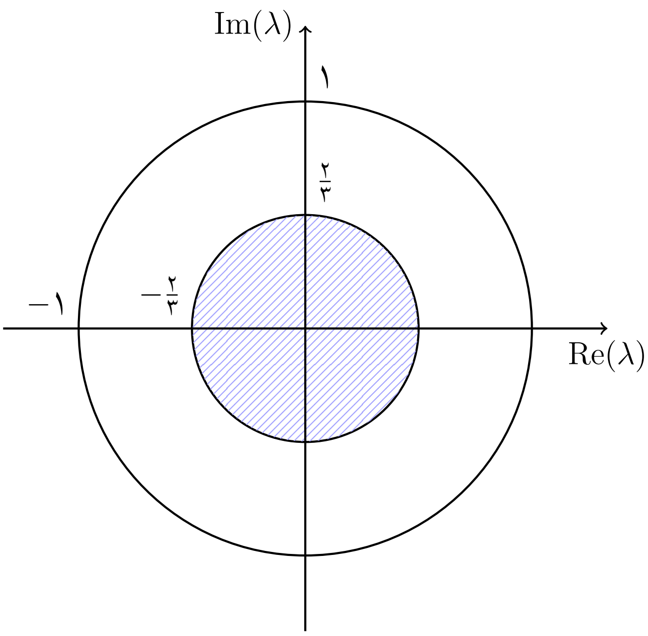
$$|\lambda_1| < 1, \qquad |\lambda_2| < 1, \qquad |\lambda_3| < 1$$

واضح است که شرایط فوق معادل با این است که مقادیر ویژه $M$ در داخل دایره‌ی یکه قرار بگیرند، مثلاً به شکل زیر باشند




$$\lambda = \mathrm{Re}(\lambda) + i\,\mathrm{Im}(\lambda)$$

(توجه کنید که منظور از $\mathrm{Re}(\lambda)$ محور حقیقی و $\mathrm{Im}(\lambda)$ محور موهومی است. حال با استفاده از قرص‌های (دیسک‌ها یا دوایر) گرشگورین می‌توان نشان داد که مقادیر ویژه ماتریس $M$ در دایره‌ای به مرکز مبدأ و شعاع حداکثر $\frac{۲}{۳}$ قرار می‌گیرند. (در مطلب بعدی درمورد قرص‌های گرشگورین توضیح می‌دهیم)




و لذا همواره در داخل دایره‌ی یکّه خواهند بود. این نشان می‌دهد که روش تکراری داده شده به جواب دستگاه داده شده همگرا خواهد بود.
همانطور که در مثال قبل دیدیم برای دانستن اینکه روش تکراری داده شده همگراست یا خیر نیاز به دانستن مقادیر دقیق مقادیر ویژه نیستیم و تنها دانستن این‌که آن‌ها در داخل دایره یکّه قرار می‌گیرند برای ما کافی است. از این‌رو گاهی اوقات فقط داشتن حدودی از مقادیر ویژه در بعضی مسائل کفایت می‌کند.
دانستن حدود مقادیر ویژه به طور زیبایی توسط دیسک‌های گرشگورین (Gershgorin circle) انجام می‌شود.

**قضیه ۵.۱**

فرض کنید $A = (a_{ij})$ ماتریس $n \times n$ (حقیقی یا مختلط) باشد و

$$r_i=\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|,\qquad i=1,2,\dots,n$$

آنگاه هر مقدار ویژه از $A$ حداقل در یکی از نامساویهای زیر صدق می‌کند:

$$|\lambda-a_{ii}|\leqslant r_i,\qquad i=1,2,\dots,n$$

بنابراین همه مقادیر ویژه $A$ در اجتماع دیسک‌ها قرار دارند

$$\cup_{i=1}^{n}\{z;\ |z-a_{ii}|\leqslant r_i\}$$

توجه کنید به هر مجموعه به صورت

$$D_i=\{z;\ |z-a_{ii}|\leqslant r_i\}$$

یک دیسک یا قرص گرشگورین گفته می‌شود.

**اثبات:** فرض کنید $(\lambda,X)$ یک جفت ویژه $A$ باشد پس $AX=\lambda X$ یا

$$\lambda X-AX=0$$

یا

$$(\lambda I-A)X=0$$

یا

$$\left(\begin{bmatrix}\lambda & & & 0\\ & \lambda & & \\ & & \ddots & \\ 0 & & & \lambda\end{bmatrix}-\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{bmatrix}\right)\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

پس

$$\begin{bmatrix}\lambda-a_{11} & -a_{12} & \dots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

در نتیجه

$$\begin{bmatrix}(\lambda-a_{11})x_1-a_{12}x_2-\cdots-a_{1n}x_n\\ -a_{21}x_1+(\lambda-a_{22})x_2\cdots-a_{2n}x_n\\ \vdots\\ -a_{n1}x_1-a_{n2}x_2-\cdots+(\lambda-a_{nn})x_n\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

پس باید داشته باشیم:

$$\begin{cases}(\lambda-a_{11})\,x_1-a_{12}x_2-\cdots-a_{1n}x_n=0\\ \quad -a_{21}x_1+(\lambda-a_{22})\,x_2-\cdots-a_{2n}x_n=0\\ \qquad\vdots\qquad\qquad\vdots\\ \quad -a_{n_1}x_1-a_{n2}x_2-\cdots-a_{n,n-1}x_{n-1}+(\lambda-a_{nn})\,x_n=0\end{cases}$$

یا

$$\begin{cases} (\lambda - a_{11})\, x_1 = a_{12}x_2 + \cdots + a_{1n}x_n = \sum\limits_{\substack{j=1\\ j\neq 1}}^{n} a_{1j}x_j \\ (\lambda - a_{22})\, x_2 = a_{21}x_1 + \cdots + a_{2n}x_n = \sum\limits_{\substack{j=1\\ j\neq 2}}^{n} a_{2j}x_j \\ \qquad\vdots \\ (\lambda - a_{nn})\, x_n = a_{n1}x_1 + \cdots + a_{n,n-1}x_{n-1} = \sum\limits_{\substack{j=1\\ j\neq n}}^{n} a_{nj}x_j \end{cases}$$

پس برای هر i می‌توان نوشت:

$$(\lambda - a_{ii})\, x_i = \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j, \qquad i = 1, 2, \ldots, n \tag{5}$$

با گرفتن قدر مطلق از طرفین رابطه (۵) خواهیم داشت:

$$|\lambda - a_{ii}|\, |x_i| = |\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j| \leqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\, |x_j|$$

با تقسیم طرفین بر $|x_i| \neq 0$ داریم

$$|\lambda - a_{ii}| \leqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\, \frac{|x_j|}{|x_i|}, \qquad i = 1, 2, \ldots, n \tag{6}$$

از طرفی $X = [x_1, x_2, \ldots, x_n]$ ، با فرض اینکه درایه $k$ام از لحاظ قدر مطلق ماکزیمم باشد یعنی

$$|x_k| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

خواهیم داشت $|x_k| \geq |x_j|$ برای هر $j$. پس

$$\frac{|x_j|}{|x_k|} \leqslant 1, \qquad j = 1, 2, \ldots, n$$

حال با قرار دادن $i = k$ در (۶) خواهیم داشت:

$$|\lambda - a_{kk}| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|\, \frac{|x_j|}{|x_k|} \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}| = r_k$$

یعنی $\lambda \in \cup_{i=1}^{n} D_i$ و این اثبات را کامل می‌کند.
توجه کنید که

$$|\lambda - a_{kk}| \leqslant r_k, \qquad k = ۱, ۲, \dots, n$$

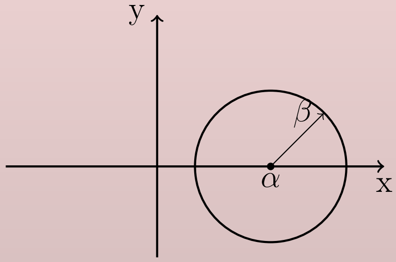
نشان می‌دهد که مقدار ویژه ی $\lambda$ حداقل در یک دیسک گرشگورین $D_i,\ i = ۱, ۲, ..., n$ قرار می‌گیرد.

**تذکر ۵.۱**

در ریاضیات عمومی دیده‌ایم که اگر $z \in \mathbb{C}$ آنگاه معادله $|z - \alpha| = \beta$ نشان دهنده‌ی یک دایره به مرکز $(\alpha, ۰)$ و شعاع $\beta$ است؛ زیرا با فرض اینکه $z = x + iy$ آنگاه

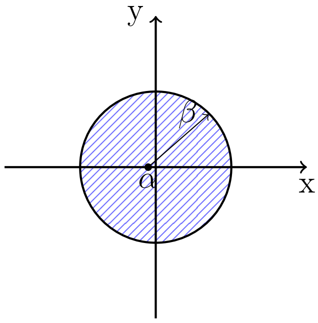
$$\beta^۲ = |z - \alpha|^۲ = |x + iy - \alpha|^۲ = |(x - \alpha) + iy|^۲ = (x - \alpha)^۲ + y^۲$$

یعنی $|z - \alpha| = \beta$ معادل است با $(x - \alpha)^۲ + y^۲ = \beta^۲$ که این همان دایره‌ای با شعاع $\beta$ و مرکز $(\alpha, ۰)$ است.

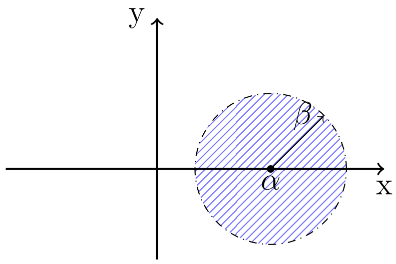



$\alpha > ۰$

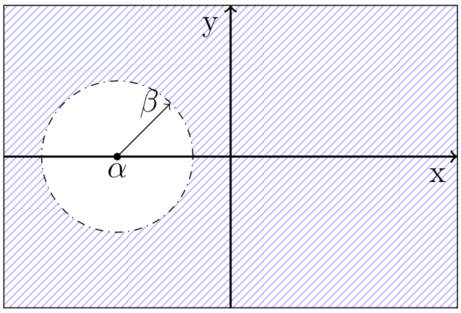
به طور مشابه می‌توان دید که عبارت $|z - \alpha| \leq \beta$ نشان دهنده‌ی درون یک دایره با شعاع $\beta$ و مرکز $(\alpha, ۰)$ به همراه محیطش است.



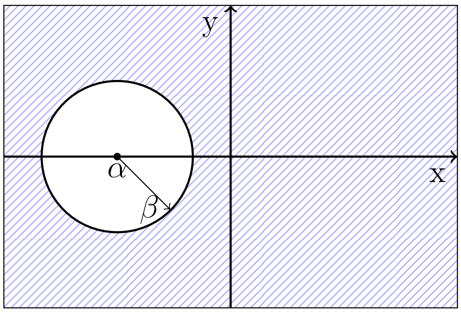

عبارت $|z - \alpha| < \beta$ نشان دهنده‌ی درون یک دایره به مرکز $(\alpha, ۰)$ و شعاع $\beta$ و بدون محیطش است. ( نقطه چین در محیط دایره نشان دهنده آن است که محیط دایره جزو ناحیه نمی‌باشد)




عبارت $|z - \alpha| > \beta$ نشان دهنده‌ی خارج یک دایره بدون احتساب محیطش است(و در حالت $|z - \alpha| \geq \beta$ نیز مرز ناحیه‌ی یعنی محیط دایره جزو مرز است).




$|z-\alpha| > \beta$




$|z-\alpha| \geq \beta$

**مثال ۵.۵**

برای ماتریس داده شده قرص‌های گرشگورین را به دست آورید.

$$A = \begin{bmatrix} ۱ & ۳ \\ -۴ & -۶ \end{bmatrix}$$

**حل:** داریم:

$$|z-a_{ii}| \leqslant r_i, \quad i = ۱, ۲$$

از طرفی $r_i = \sum_{\substack{j=۱ \\ j\neq i}}^{n} |a_{ij}|$ پس $r_۱$ و $r_۲$ به صورت زیر محاسبه می‌شوند:

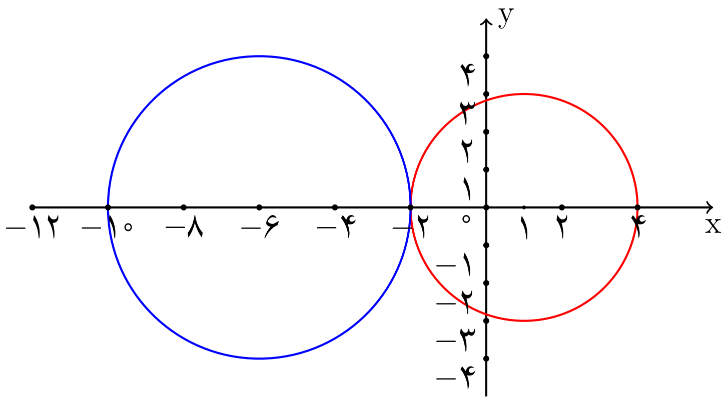
$$r_۱ = \sum_{\substack{j=۱ \\ j\neq ۱}}^{۲} |a_{۱j}| = |a_{۱۲}| = |۳| = ۳$$

$$r_۲ = \sum_{\substack{j=۱ \\ j\neq ۲}}^{۲} |a_{۲j}| = |a_{۲۱}| = |-۴| = ۴$$

پس

$$|z-a_{ii}| \leqslant r_i \xrightarrow{i=۱} |z-a_{۱۱}| \leqslant r_۱ \rightarrow |z-۱| \leqslant ۳$$
$$|z-a_{ii}| \leqslant r_i \xrightarrow{i=۲} |z-a_{۲۲}| \leqslant r_۲ \rightarrow |z+۶| \leqslant ۴$$

یعنی قرص‌های گرشگورین به صورت $|z-۱| \leqslant ۳$ و $|z+۶| \leqslant ۴$ می‌باشند که اولی دایره‌ای به مرکز (۰و۱) و شعاع ۳ است در حالی که دومی دایره‌ای به مرکز (۰و۶-) و شعاع ۴ است.




همان‌طور که در شکل‌های فوق می‌بینید دو دیسک بر هم مماس‌اند. از طرفی مقادیر ویژه در اجتماع این دو دیسک قرار خواهد گرفت.
البته حالتی می‌تواند رخ دهد و آن این است که در آن مرزی که دیسک‌ها مماس‌اند یک مقدار ویژه وجود داشته باشد که در این مثال نیز رخ داده است. درواقع مقادیر ویژه $A$ برابر

$$\lambda_1 = -۲, \qquad \lambda_2 = -۳$$

می‌باشند و همان‌طور که می‌بینید دو دیسک در $\lambda_1 = -۲$ بر هم مماس‌اند. همچنین دیسک دوم شامل مقدار ویژه‌ی $\lambda_2 = -۳$ می‌باشد.

**نکته ۵.۱**

مثال قبل نشان می‌دهد که ممکن است مقادیر ویژه‌ی یک ماتریس در نقطه‌ای که دیسک‌ها بر هم مماس هستند نیز قرار بگیرند زیرا مرزی که دیسک ها بر هم مماس اند نیز جزء اجتماع دیسک ها می باشد.

**مثال ۵.۶**

برای ماتریس داده شده دیسک‌های گرشگورین را به دست آورید.

$$A = \begin{bmatrix} ۱ & -۲ & ۳ \\ ۰ & ۲ & ۱ \\ ۲ & ۱ & -۳ \end{bmatrix}$$

**حل:** ابتدا $r_i$ها را محاسبه می‌کنیم.

$$r_1 = \sum_{\substack{j=1 \\ j\neq 1}}^{3} |a_{1j}| = |a_{12}| + |a_{13}| = |-۲| + |۳| = ۲ + ۳ = ۵$$

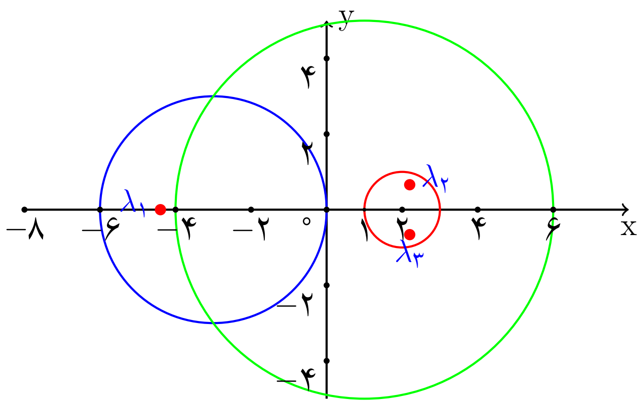
$$r_2 = \sum_{\substack{j=1 \\ j\neq 2}}^{3} |a_{2j}| = |a_{21}| + |a_{23}| = |۰| + |۱| = ۱$$

$$r_3 = \sum_{\substack{j=1 \\ j\neq 3}}^{3} |a_{3j}| = |a_{31}| + |a_{32}| = |۲| + |۱| = ۳$$

بنابراین

$$|z-a_{۱۱}| \leqslant r_۱ \quad \text{یا} \quad |z-۱| \leqslant ۵ \qquad : \text{دیسک اول}$$
$$|z-a_{۲۲}| \leqslant r_۲ \quad \text{یا} \quad |z-۲| \leqslant ۱ \qquad : \text{دیسک دوم}$$
$$|z-a_{۳۳}| \leqslant r_۳ \quad \text{یا} \quad |z+۳| \leqslant ۳ \qquad : \text{دیسک سوم}$$

اولی دایره‌ای است با شعاع ۵ و مرکز (۰و۱)
دومی دایره‌ای است با شعاع ۱ و مرکز (۰و۲)
سومی دایره‌ای است با شعاع ۳ و مرکز (۰و۳-)




بنابراین تمامی مقادیر ویژه $A$ دراجتماع ۳ دیسک فوق قرار دارند. توجه کنید که مقادیر ویژه $A$ تقریباً به صورت زیرند (با استفاده از نرم افزار متلب)

$$\lambda_۱ = -۴/۳۸۶۷, \qquad \lambda_۲ = ۲/۱۹۳۳ + ۰/۶۵۷۵i, \qquad \lambda_۳ = ۲/۱۹۳۳ - ۰/۶۵۷۵i$$

مشاهده می‌شود که $\lambda_۱$ در دیسک سوم قرار دارد و $\lambda_۲$ و $\lambda_۳$ در دیسک دوم (که در داخل دیسک اول قرار دارد) قرار دارند.

**مثال ۵.۷**

قرص‌های گرشگورین ماتریس زیر را به دست آورید.

$$A = \begin{bmatrix} ۱ & -\frac{۱}{۴} & ۰ \\ ۱ & ۴ & -۱ \\ \frac{۱}{۴} & ۰ & ۷ \end{bmatrix}$$

**حل:**

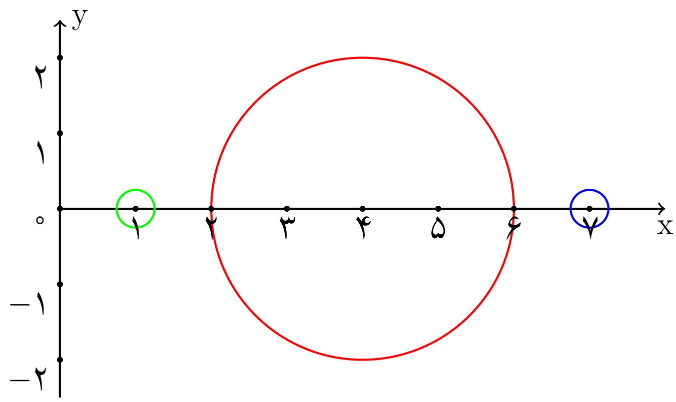
$$r_۱ = \left|-\frac{۱}{۴}\right| + |\,۰\,| = \frac{۱}{۴}$$
$$r_۲ = |۱| + |-۱| = ۲$$
$$r_۳ = \left|\frac{۱}{۴}\right| + |\,۰\,| = \frac{۱}{۴}$$

بنابراین قرص‌های گرشگورین به صورت زیر به دست می‌آیند:

$$|z-1| \leqslant \frac{1}{4} \longrightarrow (1,0)$$ دیسک اول : دایره‌ای به شعاع $\frac{1}{4}$ و مرکز $(1,0)$

$$|z-4| \leqslant 2 \longrightarrow (4,0)$$ دیسک دوم : دایره‌ای به شعاع ۲ و مرکز $(4,0)$

$$|z-7| \leqslant \frac{1}{4} \longrightarrow (7,0)$$ دیسک سوم : دایره‌ای به شعاع $\frac{1}{4}$ و مرکز $(7,0)$




از طرفی مقادیر ویژه $A$ به صورت زیرند(با استفاده از نرم افزار متلب)

$$\lambda_1 \approx 1/09, \quad \lambda_2 \approx 3/91, \quad \lambda_3 \approx 7$$

همانطور که می‌بینید هر دیسک شامل یک مقدار ویژه است.

**مثال ۵.۸**

ماتریس مختلط مقدار زیر را در نظر بگیرید.

$$A = \begin{bmatrix} 0 & i & -1 \\ 1 & 0 & -1 \\ 1-2i & 0 & 2 \end{bmatrix}$$

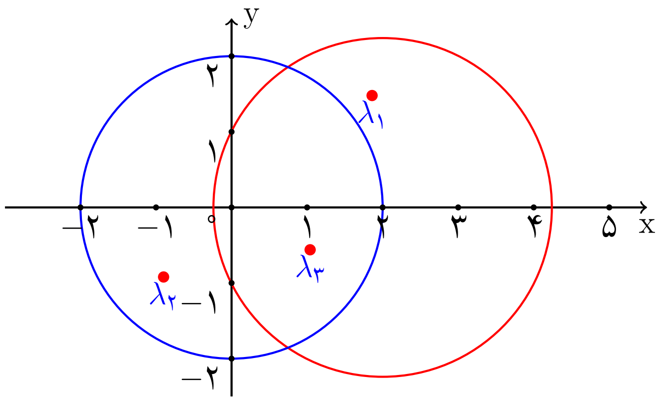
دیسک‌های گرشگورین به صورت زیر به دست می‌آیند.

$$r_1 = |i| + |-1| = 1+1 = 2$$
$$r_2 = |1| + |-1| = 1+1 = 2$$
$$r_3 = |1-2i| + 0 = \sqrt{1+4} = \sqrt{5}$$

دیسک اول : $|z-0| \leqslant 2$

دیسک دوم : $|z-0| \leqslant 2$

دیسک سوم : $|z-2| \leqslant \sqrt{5}$




دو دیسک اول و دوم برای این مثال یکی هستند( همان‌طور که در شکل زیر نیز مشخص است).
حال با محاسبه‌ی مقادیر ویژه $A$ داریم (با استفاده از نرم افزار متلب)

$$\lambda_۱ \approx ۱/۸۶ + ۱/۴۸i, \qquad \lambda_۲ \approx -۰/۹۰ - ۰/۹۲i, \qquad \lambda_۳ \approx ۱/۰۴ - ۰/۵۶i$$

**مثال ۵.۹**

ماتریس زیر را در نظر بگیرید.

$$A = \begin{bmatrix} ۹ & ۸ & ۱۰ \\ ۴ & ۱ & ۷ \\ ۴ & ۱ & ۷ \end{bmatrix}$$

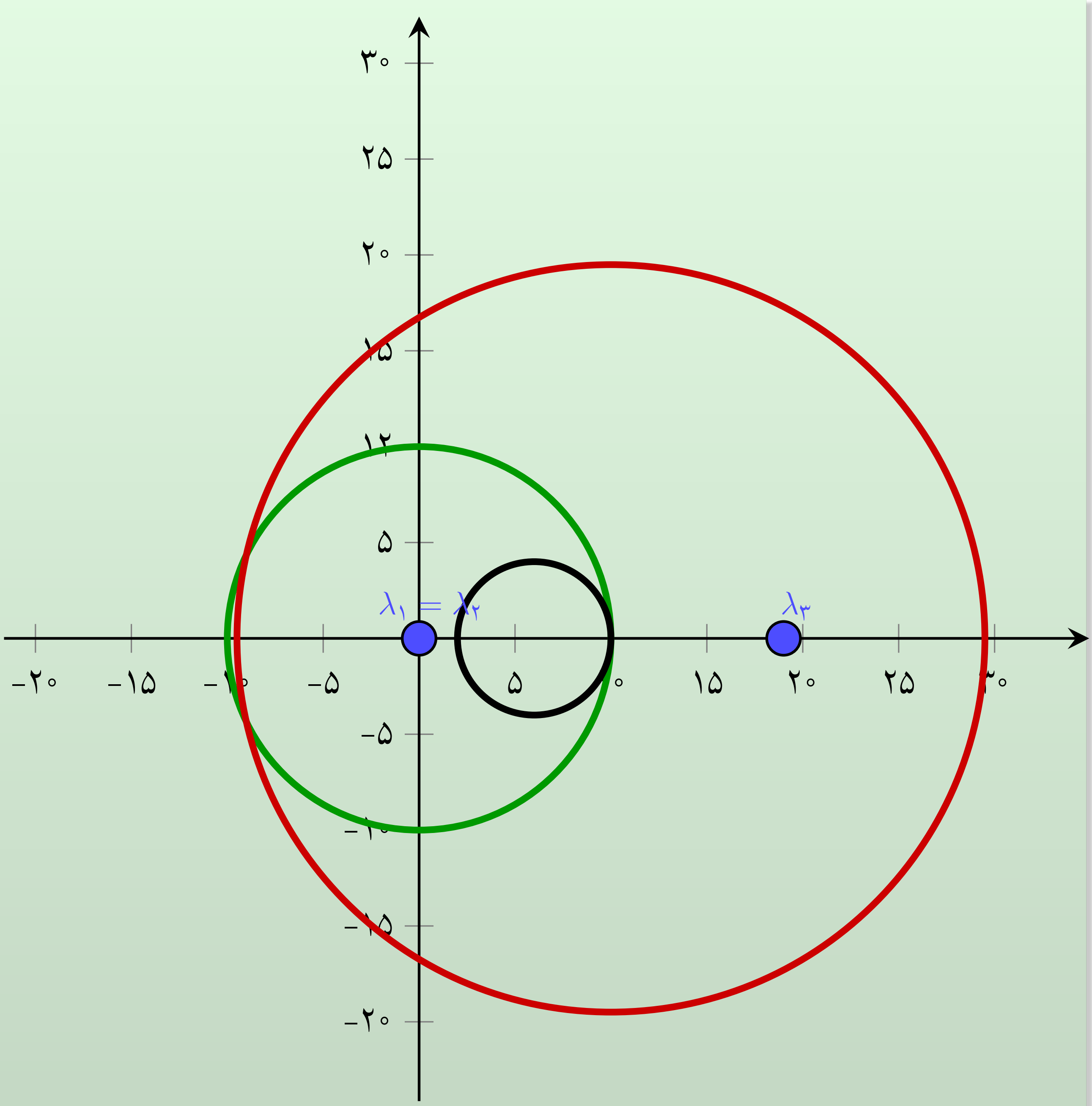
مقادیر ویژه ی این ماتریس برابرند با

$$\lambda_۱ = \lambda_۲ = ۰, \qquad \lambda_۳ = ۱۷$$

از طرفی دیسک‌های گرشگورین به صورت زیر به دست می‌آیند

دیسک اول : $|z - ۹| \leqslant ۱۸$
دیسک دوم : $|z - ۱| \leqslant ۱۱$
دیسک سوم : $|z - ۷| \leqslant ۵$

با رسم آنها داریم




شکل فوق نشان می دهد که ممکن است یک دیسک وجود داشته باشد که شامل هیچ مقدار ویژه ای نباشد. اگرچه مقادیر ویژه ی $A$ در اجتماع این دیسک ها قرار دارند.

## ۲ خواص اساسی مقادیر ویژه و بردار ویژه

(۱) فرض کنید $A$ ماتریسی $n \times n$ با مقادیر ویژه‌ی $\lambda_1, \cdots, \lambda_n$ باشد، آنگاه

$$\det(A) = \lambda_1 \times \lambda_2 \times \cdots \times \lambda_n$$
$$\operatorname{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

**توجه ۵.۱**

علاقمندان برای دیدن اثبات به به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**مثال ۵.۱۰**

برای ماتریس

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

حاصل‌ضرب و مجموع مقادیر ویژه را (بدون محاسبه‌ی مقادیر ویژه ) به دست‌آورید.

**حل:** بنا به روابط گفته شده اگر $\lambda_1, \lambda_2, \lambda_3$ مقادیر ویژه $A$ باشند آنگاه

$$\lambda_1 + \lambda_2 + \lambda_3 = \operatorname{trace}(A) = 1 + 2 + 1 = 4$$

و

$$\lambda_1\lambda_2\lambda_3 = \det(A) = 1 \times \det\begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix} - 2\det\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$$
$$+ 3\det\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = 2 - 3 - 2(3-2) + 3(9-4) = 12$$

از طرفی مقادیر ویژه $A$ به صورت

$$\lambda_1 = 6, \quad \lambda_2 = -1 + i, \quad \lambda_3 = -1 - i$$

می‌باشند. واضح است که

$$\lambda_1 + \lambda_2 + \lambda_3 = 4, \quad \lambda_1\lambda_2\lambda_3 = 6(-1+i)(-1-i) = 6 \times 2 = 12$$

(۲) مقادیر ویژهِ $A$ و $A^{-1}$ عکس یکدیگرند.

**اثبات:** فرض کنید $(\lambda, X)$ جفت ویژه‌ی ماتریس نامنفرد $A$ باشد پس $\lambda \neq 0$ همچنین $X \neq 0$. اکنون داریم

$$AX = \lambda X \Rightarrow A^{-1}AX = \lambda A^{-1}X \Rightarrow X = \lambda A^{-1}X \Rightarrow A^{-1}X = \frac{1}{\lambda}X$$

این نشان می‌دهد که $(\frac{1}{\lambda}, X)$ جفت ویژه‌ی $A^{-1}$ خواهند بود. توجه کنید که بردار ویژه‌ی $X$ هیچ تغییری نخواهد کرد.

(۳) مقادیر ویژه‌ی $A$ و $A^T$ یکی هستند.

**اثبات:** چون $\det(B) = \det\left(B^T\right)$ برای هر ماتریس $B$ برقرار است پس می‌توان نوشت

$$P_{A^T}(\lambda) = \det\left(A^T - \lambda I\right) = \det\left((A^T - \lambda I)^T\right) = \det(A - \lambda I) = P_A(\lambda)$$

این نشان می‌دهد که چند جمله ای ویژه ماتریس های $A$ و $A^T$ یکی هستند لذا اگر $\lambda$ مقدار ویژه‌ی $A$ باشد آنگاه مقدار ویژه‌ی $A^T$ نیز خواهد بود.

(۴) فرض کنید $\lambda(A)$ نشان دهنده‌ی مقدار ویژه‌ی $A$ باشد آنگاه $\lambda(A^۲) = (\lambda(A))^۲$. درواقع برای محاسبهِ مقادیر ویژه‌ی ماتریس $A^۲$ کافی است مقادیر ویژهِ $A$ را به توان ۲ برسانیم.

**اثبات:** از $AX = \lambda X$ و ضرب طرفین از چپ در $A$ داریم:

$$A^۲X = \lambda AX \overset{AX=\lambda X}{\longrightarrow} A^۲X = \lambda(\lambda X) = \lambda^۲X$$

رابطه بالا نشان می‌دهد که $\lambda^۲$ مقدار ویژه‌ی $A^۲$ است. توجه کنید که بردار ویژه‌ها تغییری نخواهند کرد.

**مثال ۵.۱۱**

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۱ & -۳ \\ ۲ & ۶ \end{bmatrix}$$

که دارای مقادیر ویژه‌ی $\lambda_۱ = ۳, \lambda_۲ = ۴$ و بردار ویژه‌های

$$X_۱ = \begin{bmatrix} -\frac{۳}{۲} \\ ۱ \end{bmatrix}, \quad X_۲ = \begin{bmatrix} -۱ \\ ۱ \end{bmatrix}$$

می‌باشد.
حال با محاسبه‌ی $A^۲$ داریم

$$A^۲ = \begin{bmatrix} -۵ & -۲۱ \\ ۱۴ & ۳۰ \end{bmatrix}$$

که دارای چندجمله‌ای ویژه‌ی $\lambda^۲ - ۲۵\lambda + ۱۴۴ = (۹ - \lambda)(۱۶ - \lambda)$ می‌باشد. بنابراین مقادیر ویژه‌ی $A^۲$ همان $\lambda = \lambda_۱^۲ = ۳^۲ = ۹, \lambda = \lambda_۲^۲ = ۴^۲ = ۱۶$ می‌باشند. از طرفی داریم

$$A^۲X_۱ = \begin{bmatrix} -۵ & -۲۱ \\ ۱۴ & ۳۰ \end{bmatrix}\begin{bmatrix} -\frac{۳}{۲} \\ ۱ \end{bmatrix} = \begin{bmatrix} -\frac{۲۷}{۲} \\ ۹ \end{bmatrix} = ۹\begin{bmatrix} -\frac{۳}{۲} \\ ۱ \end{bmatrix} = \lambda_۱^۲X_۱$$
$$A^۲X_۲ = \begin{bmatrix} -۵ & -۲۱ \\ ۱۴ & ۳۰ \end{bmatrix}\begin{bmatrix} -۱ \\ ۱ \end{bmatrix} = \begin{bmatrix} -۱۶ \\ ۱۶ \end{bmatrix} = ۱۶\begin{bmatrix} -۱ \\ ۱ \end{bmatrix} = \lambda_۲^۲X_۲$$

که نشان می‌دهد $X_۱, X_۲$ همچنان بردار ویژه‌ی $A^۲$ می‌باشند.

(۵) برای هر عدد صحیح $k$ داریم
$$\lambda\left(A^k\right) = (\lambda(A))^k$$
و بردار ویژه‌های $A^k$ همان بردار ویژه‌های $A$ هستند.
**اثبات: تمرین**

(۶) مقادیر ویژه‌ی یک ماتریس بالامثلثی (یا پایین مثلثی ) اعضای روی قطر آن هستند.

**اثبات:** فرض کنید $A$ ماتریسی بالامثلثی باشد

$$A = \begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ & a_{۲۲} & \cdots & a_{۲n} \\ & & \ddots & \vdots \\ \mathbb{O} & & & a_{nn} \end{bmatrix}$$

آنگاه ماتریس $A - \lambda I$ به صورت زیر خواهد بود

$$A - \lambda I = \begin{bmatrix} a_{۱۱} - \lambda & a_{۱۲} & \cdots & a_{۱n} \\ & a_{۲۲} - \lambda & \cdots & a_{۲n} \\ & & \ddots & \vdots \\ \mathbb{O} & & & a_{nn} - \lambda \end{bmatrix}$$

بنابراین از اینکه دترمینان یک ماتریس بالامثلثی حاصل‌ضرب عناصر روی قطر اصلی است خواهیم داشت:

$$\det(A - \lambda I) = (a_{۱۱} - \lambda)\,(a_{۲۲} - \lambda)\cdots(a_{nn} - \lambda)$$

که نشان می‌دهد مقادیر ویژه‌ی $A$ به صورت $\lambda_۱ = a_{۱۱}, \lambda_۲ = a_{۲۲}, \dots, \lambda_n = a_{nn}$ می باشند. اثبات برای یک ماتریس پایین مثلثی به طور مشابه انجام می‌گردد.

**مثال ۵.۱۲**

مقادیر ویژه ماتریس‌های

$$A = \begin{bmatrix} ۲ & ۰ & -۳ \\ ۰ & ۰ & ۴ \\ ۰ & ۰ & ۷ \end{bmatrix}, \quad B = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۲ & -۱ & ۰ \\ ۳ & ۸ & ۱۱ \end{bmatrix}, \quad C = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۶ & ۰ \\ ۰ & ۰ & ۲ \end{bmatrix}$$

را بیابید.

**حل:** بنا به نکته‌ی قبل عناصر روی قطر اصلی مقادیر ویژه خواهند بود.

$$\lambda(A) = \{۰, ۲, ۷\}, \quad \lambda(B) = \{-۱, ۱, ۱۱\}, \quad \lambda(C) = \{۱, ۲, ۶\}$$

(۷) اگر $\lambda$ یک مقدار ویژه‌ی ماتریس $A$ باشد آنگاه مقادیر ویژه‌ی ماتریس $A - \alpha I$ برابر $\lambda - \alpha$ خواهند بود. و بردار ویژه‌ی $A - \alpha I$ همان بردار ویژه‌ی $A$ است.

**اثبات:** فرض کنید $(\lambda, X)$ یک جفت ویژه‌ی $A$ باشد پس $AX = \lambda X$. از طرفی با در نظر گرفتن $(A-\alpha I)X$ داریم:

$$(A-\alpha I)X = AX - \alpha X = \lambda X - \alpha X = (\lambda - \alpha)X$$

که نشان می‌دهد $(\lambda - \alpha, X)$ یک جفت ویژه ماتریس $A - \alpha I$ است.
توجه کنید که به خاصیت $\lambda(A-\alpha I) = \lambda(A) - \alpha$، خاصیت انتقال (shift property) گفته می‌شود.

**قضیه ۵.۲**

دو ماتریس متشابه دارای مقادیر ویژه‌ی یکسان هستند. یعنی مقادیر ویژه $A$ و $TAT^{-1}$ برای همه ماتریس‌های نامنفرد $T$ یکسان هستند.

**اثبات:** نشان می‌دهیم که $A$ و $TAT^{-1}$ دارای چندجمله‌ای مشخصه‌های یکسان هستند و در نتیجه دارای مقادیر ویژه یکسان هستند.

$$\begin{aligned}\det\left(TAT^{-1} - \lambda I\right) &= \det\left(T(A-\lambda I)T^{-1}\right) = \det(T)\cdot\det(A-\lambda I)\cdot\det\left(T^{-1}\right)\\ &= \det(T)\cdot\det\left(T^{-1}\right)\cdot\det(A-\lambda I) = \det\left(T\cdot T^{-1}\right)\cdot\det(A-\lambda I)\\ &= \det(A-\lambda I)\end{aligned}$$

بنابراین $TAT^{-1}$ و $A$ دارای چندجمله‌ای مشخصه‌های یکسان هستند و درنتیجه دارای مقادیر ویژه یکسان می‌باشند.

**توجه ۵.۲**

از نماد $B \sim A$ برای وقتی که $A$ و $B$ متشابه اند استفاده می شود.

**تذکر ۵.۲**

اگر دو ماتریس دارای مقادیر ویژه‌ی یکسان باشند لزوماً متشابه نمی‌باشند.

برای مثال دو ماتریس

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

دارای مقادیر ویژه‌ی ۱ می‌باشند. اما $A$ و $B$ متشابه نمی‌باشند زیرا اگر $T$ یک ماتریس وارون‌پذیر باشد به قسمی که $TBT^{-1} = A$ آنگاه از $B = I_2$ می توان نوشت

$$A = TBT^{-1} = TI_2T^{-1} = I_2$$

اما $A \neq I_2$. پس هیچ ماتریس $T$ای وجود ندارد که $A$ و $B$ با هم متشابه گردند.

**تذکر ۵.۳**

نکتهِ قبل یک راهکار برای محاسبه‌ی مقادیر ویژه ارائه می‌دهد. بدین صورت که اگر بتوانیم ماتریسی مانند $B$ به دست آوریم که با $A$ متشابه باشد با این تفاوت که محاسبه‌ی مقادیر ویژهِ $B$ نسبت به $A$ راحت‌تر باشد.
مثلا $B$ یک ماتریس قطری، مثلثی و یا سه قطری باشد آنگاه تبدیل $TAT^{-1} = B$ می تواند مفید باشد. اینکه به چه صورت $A$ را با یک ماتریس مناسب مثل $B$ متشابه نماییم در مطالب بعدی آورده می‌شود.

**نکته ۵.۲**

همان‌طور که دیدیم $\lambda(A) = \lambda(B)$ وقتی $A$ و $B$ متشابه‌اند.

حال سوالی که مطرح می‌گردد این است که بردارهای ویژه‌ی دو ماتریس متشابه به چه صورت با هم مرتبطاند؟!

(۸) اگر $A$ و $B$ متشابه باشند یعنی ماتریس وارون‌پذیر $T$ وجود داشته باشد به قسمی که $TAT^{-1} = B$ و اگر $X$ بردار ویژه‌ی $A$ باشد آنگاه $TX$ بردار ویژه‌ی $B$ است.

**اثبات:** اگر $(\lambda, X)$ جفت ویژه‌ی $A$ باشد آنگاه $AX = \lambda X$ یا

$$TAX = \lambda TX \Rightarrow \left(TAT^{-1}\right)(TX) = \lambda(TX) \Rightarrow B(TX) = \lambda(TX)$$

که نشان می‌دهد $TX$ بردار ویژه‌ی $B$ است.

(۹) اگر $A$ ماتریسی متقارن باشد آنگاه مقادیر ویژه‌ی $A$ حقیقی‌اند.

اثبات: داریم

$$AX = \lambda X \Rightarrow (AX)^H = (\lambda X)^H \Rightarrow X^H A^H = \lambda^H X^H$$

چون $A$ متقارن است پس $A^H = A^T = A$ لذا

$$X^H A = \lambda^H X^H \xrightarrow{\text{ضرب از راست در } X} X^H AX = \lambda^H X^H X \qquad (۷)$$

حال تساوی $AX = \lambda X$ را از چپ در $X^H$ ضرب می‌کنیم

$$X^H AX = \lambda X^H X \qquad (۸)$$

اکنون از مقایسه‌ی (۷) و (۸) درمی‌یابیم که می‌بایست

$$\lambda X^H X = \lambda^H X^H X$$

چون $X \neq ۰$ پس $X^H X > ۰$ و لذا با تقسیم طرفین تساوی فوق بر $X^H X$ داریم $\lambda = \lambda^H$. اگر $\lambda = \alpha + i\beta$ آنگاه $\lambda^H = \alpha - i\beta$ پس باید $\alpha + i\beta = \alpha - i\beta$ یا $۲i\beta = ۰$ یا $\beta = ۰$. پس $\lambda = \alpha + i\beta = \alpha + i \times ۰ = \alpha$ یعنی $\lambda$ حقیقی مقدار است و اثبات تمام می‌شود.

(۱۰) اگر $A$ ماتریسی پادمتقارن باشد $(A^T = -A)$ آنگاه مقادیر ویژه‌ی $A$ به صورت $\lambda = i\alpha$ که $\alpha \in \mathbb{R}$ می‌باشد.
بنابراین مقادیر ویژه‌ی $A$ یا صفرند و یا موهومی محض.
**اثبات: تمرین**

**در ادامه به معرفی یک روش برای محاسبه‌ی چندجمله‌ای ویژه‌ی ماتریس $A$ می پردازیم که‌آن را جزو دسته روش های مستقیم برای حل مسأله مقدار ویژه در نظر می‌گیریم.**

همان‌طور که از تعریف چندجمله‌ای ویژه برمی‌آید

$$P_A(\lambda) = \det(\lambda I - A)$$

محاسبه‌ی $P_A(\lambda)$ از تعریف (چون وابسته به محاسبه‌ی دترمینان است) کاری دشوار و هزینه‌بر خواهد بود. بنابراین می‌بایست راهکار دیگری را برای محاسبه‌ی آن پی گرفت.

ابتدا به یک قضیه‌ی مهم و اساسی در ارتباط با چندجمله‌ای ویژه یک ماتریس توجه کنید.

**قضیه ۵.۳**

**(کیلی - همیلتون) (Cayley-Hamilton)**:
هر ماتریس مربعی در معادله‌ی مشخصه‌اش صدق می‌کند، یعنی اگر $A$ ماتریسی $n \times n$ و دارای معادله مشخصه

$$P_A(\lambda) = \lambda^n + c_{n-۱}\lambda^{n-۱} + c_{n-۲}\lambda^{n-۲} + \cdots + c_۱\lambda + c_۰$$

باشد آنگاه

$$P_A(A) = A^n + c_{n-۱}A^{n-۱} + c_{n-۲}A^{n-۲} + \cdots + c_۱A + c_۰I = \mathbb{O}$$

جایی که $\mathbb{O}$ نشان دهنده‌ی ماتریس $n \times n$ صفر است و $I$ ماتریسی همانی $n \times n$ است.

**توجه ۵.۳**

علاقمندان برای دیدن اثبات به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**مثال ۵.۱۳**

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۲ & ۵ \\ -۱ & ۶ \end{bmatrix}$$

چند جمله ای مشخصه $A$ چنین است

$$P_A(\lambda) = \lambda^۲ - ۸\lambda + ۱۷$$

واضح است که

$$P_A(A) = A^۲ - ۸A + ۱۷I = \begin{bmatrix} -۱ & ۴۰ \\ -۸ & ۳۱ \end{bmatrix} - ۸\begin{bmatrix} ۲ & ۵ \\ -۱ & ۶ \end{bmatrix} + \begin{bmatrix} ۱۷ & ۰ \\ ۰ & ۱۷ \end{bmatrix} = \begin{bmatrix} ۰ & ۰ \\ ۰ & ۰ \end{bmatrix} = \mathbb{O}$$

یعنی$A$ در معادله مشخصه‌اش صدق می‌کند.

در ادامه خواهیم دید که به کمک این قضیه روش جالب و زیبایی برای استخراج ضرایب چندجمله ای ویژه یک ماتریس دلخواه می توان ساخت.

## ۳ روش کرایلف

حال آماده‌ایم روش کرایلف (Krylov) را برای محاسبه‌ی چندجمله‌ای ویژه یک ماتریس بیان نماییم.
فرض کنید

$$P_A(\lambda) = \det(\lambda I - A) = \lambda^n + c_{n-۱}\lambda^{n-۱} + \cdots + c_۱\lambda + c_۰ = ۰$$

بنا به قضیه‌ی کیلی_ همیلتون $A$ در معادله مشخصه‌اش صدق می‌کند پس

$$A^n + c_{n-۱}A^{n-۱} + c_{n-۲}A^{n-۲} + \cdots + c_۱A + c_۰I = \mathbb{O} \tag{۹}$$

فرض کنید $V$ یک بردار دلخواه ناصفر $n \times 1$ باشد. با ضرب $V$ از راست در (۹) داریم:

$$A^nV + c_{n-1}A^{n-1}V + c_{n-2}A^{n-2}V + \cdots + c_1AV + c_0V = 0$$

یا

$$c_{n-1}A^{n-1}V + c_{n-2}A^{n-2}V + \cdots + c_1AV + c_0V = -A^nV$$

یا

$$\begin{bmatrix} A^{n-1}V & A^{n-2}V & \cdots & AV & V \end{bmatrix} \begin{bmatrix} c_{n-1} \\ c_{n-2} \\ \vdots \\ c_1 \\ c_0 \end{bmatrix} = -A^nV \tag{10}$$

اگر تعریف کنیم

$$\begin{aligned} M &= \left[A^{n-1}V \quad A^{n-2}V \cdots AV \quad V\right] \\ C &= \begin{bmatrix} c_{n-1}, & c_{n-2}, & \cdots & c_1, & c_0 \end{bmatrix}^T \\ b &= -A^nV \end{aligned}$$

آنگاه (۱۰) معادل با دستگاه خطی $MC = b$ است که از حل آن بردار مجهولات $C$ حاصل می‌شود. واضح است که با یافتن $C = \begin{bmatrix} c_{n-1}, & c_{n-2}, & \cdots & c_1, & c_0 \end{bmatrix}^T$ چندجمله‌ای مشخصه $A$ به شکل زیر در دسترس خواهد بود:

$$P_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_1\lambda + c_0$$

**مثال ۵.۱۴**

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

با استفاده از روش کرایلف و با $V = [0, \ 1]^T$، چندجمله‌ای مشخصه $A$ را تعیین نمایید.

**حل:** ابتدا ماتریس $M$ و بردار $b$ را تشکیل می‌دهیم:

$$M = \begin{bmatrix} AV & V \end{bmatrix}, \quad b = -A^2V$$

$$AV = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$A^2V = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 22 \end{bmatrix}$$

پس

$$M = \begin{bmatrix} AV & V \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$$

$$b = -A^۲V = \begin{bmatrix} -۱۰ \\ -۲۲ \end{bmatrix}$$

با حل دقیق دستگاه $MC = b$ داریم $C = [-۵,\ -۲]^T$
پس چندجمله‌ای ویژه $A$ به صورت زیر حاصل می‌شود:

$$P_A(\lambda) = \lambda^۲ - ۵\lambda - ۲$$

حال با حل معادله‌ی $P_A(\lambda) = ۰$ می‌توان مقادیر ویژه $A$ را یافت. بعد از محاسبه ی مقادیر ویژه می توان از رابطه $AX = \lambda X$ به محاسبه ی مقادیر ویژه پرداخت.

**مثال ۵.۱۵**

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۵ & -۳ \\ ۲ & ۷ \end{bmatrix}$$

با استفاده از روش کرایلف و با $V = [۲, ۳]^T$، چندجمله‌ای مشخصه $A$ را تعیین نمایید.

**حل:** ابتدا ماتریس $M$ و بردار $b$ را تشکیل می‌دهیم:

$$M = [\ AV \quad V\ ], \quad b = -A^۲V$$

$$AV = \begin{bmatrix} ۱ \\ ۲۵ \end{bmatrix}, \qquad A^۲V = \begin{bmatrix} -۷۰ \\ ۱۷۷ \end{bmatrix}$$

پس

$$M = [\ AV \quad V\ ] = \begin{bmatrix} ۱ & ۲ \\ ۲۵ & ۳ \end{bmatrix}$$

$$b = -A^۲V = \begin{bmatrix} ۷۰ \\ -۱۷۷ \end{bmatrix}$$

با حل دقیق دستگاه $MC = b$ داریم $C = [-۱۲, ۴۱]^T$. پس چندجمله‌ای ویژه $A$ به صورت زیر حاصل می‌شود:

$$P_A(\lambda) = \lambda^۲ - ۱۲\lambda + ۴۱$$

حال با حل معادله‌ی $P_A(\lambda) = ۰$ می‌توان مقادیر ویژه $A$ را یافت. بعد از محاسبه ی مقادیر ویژه می توان از رابطه $AX = \lambda X$ به محاسبه ی مقادیر ویژه پرداخت.

**نکته ۵.۳**

باید بردار دلخواه $V$ طوری باشد که ماتریس $M = [A^{n-۱}V \quad A^{n-۲}V \cdots AV \quad V]$ منفرد نگردد تا بتوان دستگاه $MC = b$ را حل کرد. در این حالت باید بردار دلخواه $V$ را تعویض نمود.

**تمرین ۵.۲**

نشان دهید اگر بردار دلخواه $V$ به اتفاق یک بردار ویژه ی $A$ باشد آنگاه ماتریس

$$M = \left[A^{n-1}V \quad A^{n-2}V \cdots AV \quad V\right]$$

منفرد خواهد بود و در این حالت باید بردار دلخواه $V$ را تعویض نمود.

**توجه ۵.۴**

علاقمندان برای دیدن روش های مستقیم دیگر از جمله روش لوربیرـ فادیو و ضرایب نامعین به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

# ۴ روش های تکراری برای محاسبه ی جفت ویژه ها

در بخش های قبلی توانستیم بدون استفاده از دترمینان چندجمله ای ویژه یک ماتریس را به دست آوریم. سپس با حل چند جمله ای ویژه با روش هایی چون نیوتن‌ـ رافسون به محاسبه مقادیر ویژه پرداختیم. در این بخش می خواهیم بدون محاسبه چند جمله ایی ویژه به محاسبه مقدار ویژه بپردازیم. با توجه به اینکه مساله یافتن مقادیر ویژه یک ماتریس معادل با حل چند جمله ایی ویژه است و اینکه هر چند جمله ایی از درجه ی $n \geq 5$ در حالت کلی به طور دقیق قابل حل نیست پس ماهیت روش های یافتن مقادیر ویژه می بایست تکراری باشد.

## ۱.۴ روش توانی برای محاسبه بزرگترین ریشه از لحاظ اندازه

(Power Method)

در برخی از کاربردهای مربوط به مقادیر ویژه لازم است تنها بزرگترین مقدار ویژه از لحاظ اندازه در دسترس باشد. به عنوان یادآوری یک موضوع خاص مساله حل $AX = b$ با روش تکراری

$$X^{(k+1)} = MX^{(k)} + c, \quad k = 0, 1, 2, \ldots$$

تنها لازم است بدانیم که

$$\rho(M) = \max\{|\lambda|; \text{است}M\text{یک مقدار ویژه ی ماتریس}\lambda\} < 1$$

در نتیجه دانستن شعاع طیفی ماتریس $M$ از لحاظ اندازه بوجود خواهد آمد. روش توانی تحت شرایطی می تواند این مقدار ویژه را به طور تکراری تقریب نماید.

قبل از بیان روش توانی یادآوری می کنیم که اگر $(\lambda, X)$ جفت ویژه ی $A$ باشد آنگاه $(\lambda^k, X)$ جفت ویژه ی $A^k$ خواهد بود. به عبارتی

$$AX = \lambda X \Rightarrow A^k X = \lambda^k X$$

حال فرض کنید $\lambda_1, \lambda_2, \ldots, \lambda_n$ مقادیر ویژه ماتریس $A$ و $X_1, X_2, \ldots, X_n$ بردار ویژه های متناظر با آنها باشند. بعلاوه فرض کنید دو شرط زیر برقرار باشند

بردار ویژه های $A$ مستقل خطی باشند

و

مقدار ویژه های $A$ در رابطه زیر صدق کنند:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \tag{11}$$

توجه کنید در این حالت گفته می شود $\lambda_1$ یک مقدار ویژه غالب $A$ است.

بعلاوه به قضایای مهم زیر توجه کنید(یادآوری)

**قضیه ۵.۴**

هرگاه $\{V_1, \dots, V_n\}$ یک مجموعه مستقل خطی در $\mathbb{R}^n$ باشد آنگاه یک پایه برای $\mathbb{R}^n$ است. یعنی هر عضو از $\mathbb{R}^n$ مثل $X$ برحسب ترکیب خطی $V_i$ ها نوشته می شود

$$X = \alpha_1 V_1 + \cdots + \alpha_n V_n \quad , \quad \alpha_n \in \mathbb{R}$$

**قضیه ۵.۵**

ماتریس $A$ قطری شدنی است $\Leftrightarrow$ بردارهای ویژه $A$ مستقل خطی باشد.

اکنون آماده ایم روش توانی را بیان کنیم، فرض کنید $V^{(0)} \in \mathbb{R}^n \neq 0$ برداری دلخواه باشد و $A$ ماتریس قطری شدنی باشد. بنابراین بردار ویژه های $A$ مجموعه ی $\mathbb{R}^n$ را تولید می کند. پس هر عضو دلخواه از $\mathbb{R}^n$ مثل $V^{(0)}$ برحسب ترکیب خطی بردار ویژه های $A$ قابل نمایش است:

$$V^{(0)} = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n$$

تعریف می کنیم:

$$V^{(k)} = A^k V^{(0)}, \quad k = 1, 2, \dots$$

بنابراین

$$\begin{aligned} V^{(k)} &= A^k(\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n) \\ &= \alpha_1 A^k X_1 + \alpha_2 A^k X_2 + \cdots + \alpha_n A^k X_n \end{aligned}$$

از طرفی $A^k X_i = \lambda_i^k X_i$. پس می توان نوشت:

$$\begin{aligned} V^{(k)} &= \alpha_1 \lambda_1^k X_1 + \alpha_2 \lambda_2^k X_2 + \cdots + \alpha_n \lambda_n^k X_n \\ &= \lambda_1^k(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k X_n) \end{aligned} \tag{12}$$

با تبدیل $k$ به $k+1$ خواهیم داشت

$$V^{(k+1)} = \lambda_1^{k+1}(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k+1} X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k+1} X_n) \tag{13}$$

با تقسیم مولفه $i$ ام (۱۳) بر مولفه $i$ ام (۱۲) داریم:

$$\begin{aligned} \frac{(V^{(k+1)})_i}{(V^{(k)})_i} &= \frac{\lambda_1^{k+1}}{\lambda_1^k} . \frac{(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k+1} X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k+1} X_n)_i}{(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k} X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k} X_n)_i} \\ &= \lambda_1 \times \frac{(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k+1} X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k+1} X_n)_i}{(\alpha_1 X_1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k} X_2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k} X_n)_i} \end{aligned} \tag{14}$$

بنابه شرط (۱۱) خواهیم داشت:

$$\left|\frac{\lambda_2}{\lambda_1}\right| < 1, \quad \left|\frac{\lambda_3}{\lambda_1}\right| < 1, \quad \dots, \quad \left|\frac{\lambda_n}{\lambda_1}\right| < 1 \tag{15}$$

پس

$$\begin{cases} \lim_{k\to\infty}\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1} = \lim_{k\to\infty}\left(\frac{\lambda_2}{\lambda_1}\right)^{k} = 0 \\ \lim_{k\to\infty}\left(\frac{\lambda_3}{\lambda_1}\right)^{k+1} = \lim_{k\to\infty}\left(\frac{\lambda_3}{\lambda_1}\right)^{k} = 0 \\ \vdots \\ \lim_{k\to\infty}\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} = \lim_{k\to\infty}\left(\frac{\lambda_n}{\lambda_1}\right)^{k} = 0 \end{cases} \tag{16}$$

حال با حد بی نهایت گرفتن از (۱۴) و لحاظ کردن (۱۶) داریم:

$$\begin{aligned} \lim_{k\to\infty}\frac{(V^{(k+1)})_i}{(V^{(k)})_i} &= \lambda_1 \lim_{k\to\infty}\frac{(\alpha_1 X_1 + \alpha_2(\frac{\lambda_2}{\lambda_1})^{k+1}X_2 + \dots + \alpha_n(\frac{\lambda_n}{\lambda_1})^{k+1}X_n)_i}{(\alpha_1 X_1 + \alpha_2(\frac{\lambda_2}{\lambda_1})^{k}X_2 + \dots + \alpha_n(\frac{\lambda_n}{\lambda_1})^{k}X_n)_i} \\ &= \lambda_1 \frac{(\alpha_1 X_1 + \alpha_2 \lim_{k\to\infty}(\frac{\lambda_2}{\lambda_1})^{k+1}X_2 + \dots + \alpha_n \lim_{k\to\infty}(\frac{\lambda_n}{\lambda_1})^{k+1}X_n)_i}{(\alpha_1 X_1 + \alpha_2 \lim_{k\to\infty}(\frac{\lambda_2}{\lambda_1})^{k}X_2 + \dots + \alpha_n \lim_{k\to\infty}(\frac{\lambda_n}{\lambda_1})^{k}X_n)_i} \\ &= \lambda_1 \frac{(\alpha_1 X_1)_i}{(\alpha_1 X_1)_i} = \lambda_1 \end{aligned} \tag{17}$$

به شرطی که $\alpha_1 \neq 0$. این نشان می دهد که برای $k$ به اندازه کافی بزرگ:

$$\frac{(V^{(k+1)})_i}{(V^{(k)})_i} \approx \lambda_1$$

و لذا بدین طریق می توان مقدار ویژه ی غالب $\lambda_1$ را تقریب زد. برای محاسبه ی بردار ویژه ی متناظر با $\lambda_1$ یعنی $X_1$ با توجه به (۱۲) یعنی

$$V^{(k)} = \lambda_1^k(\alpha_1 X_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k X_2 + \dots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k X_n)$$

داریم:

$$\lim_{k\to\infty}\frac{V^{(k)}}{\lambda_1^k} = \alpha_1 X_1 \quad , \quad \alpha_1 \neq 0$$

پس بردار $\frac{V^{(k)}}{\lambda_1^k}$ به مضربی از بردار $X_1$ میل می کند در نتیجه $V^{(k)}$ به مضربی از $X_1$ میل خواهد کرد.

**نکته ۵.۴**

برای محاسبه ی $V^{(k)}$ از $V^{(k)} = A^k V^{(0)}$ لازم نیست که ماتریس $A$ را به توان $k$ برسانیم. برای اینکار می توانیم بنویسیم:

$$V^{(k-1)} = A^{k-1}V^{(0)} \Rightarrow AV^{(k-1)} = AA^{k-1}V^{(0)} = A^k V^{(0)}$$

از طرفی داریم $A^k V^{(0)} = V^{(k)}$ پس به دست می آوریم:

$$AV^{(k-1)} = V^{(k)}$$

یا

$$\boxed{V^{(k)} = AV^{(k-1)} \quad , \quad k = 1, 2, \dots}$$

بنابراین در عمل برای محاسبه ی $V^{(k)}$ ها از رابطه ی فوق استفاده می کنیم نه رابطه ی $V^{(k)} = A^k V^{(0)}$.

**توجه ۵.۵**

همانطور که می بینید در روش توانی به طور ضمنی نیاز به محاسبه ی توان های ماتریس $A$ است از اینرو آن را روش توانی می نامند.

**مثال ۵.۱۶**

مقدار ویژه و بردار ویژه غالب ماتریس زیر را بیابید.

$$A = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix}$$

**حل:** فرض کنید $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = V^{(0)}$ آنگاه داریم(محاسبات در نرم افزار متلب انجام شده است و اعداد ۴ رقم بعد از اعشار گرد شده اند)

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \to \frac{(V^{(1)})_1}{(V^{(0)})_1} = 4/0000$$

$$V^{(2)} = AV^{(1)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 19 \\ 10 \end{bmatrix} \to \frac{(V^{(2)})_1}{(V^{(1)})_1} = 4/7500$$

$$V^{(3)} = AV^{(2)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 19 \\ 10 \end{bmatrix} = \begin{bmatrix} 106 \\ 79 \end{bmatrix} \to \frac{(V^{(3)})_1}{(V^{(2)})_1} = 5/5789$$

$$V^{(4)} = AV^{(3)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 106 \\ 79 \end{bmatrix} = \begin{bmatrix} 661 \\ 580 \end{bmatrix} \to \frac{(V^{(4)})_1}{(V^{(3)})_1} = 6/2358$$

$$V^{(5)} = AV^{(4)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 661 \\ 580 \end{bmatrix} = \begin{bmatrix} 4384 \\ 4141 \end{bmatrix} \rightarrow \frac{(V^{(5)})_1}{(V^{(4)})_1} = 6/6324$$
$$V^{(6)} = AV^{(5)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 4384 \\ 4141 \end{bmatrix} = \begin{bmatrix} 29959 \\ 29230 \end{bmatrix} \rightarrow \frac{(V^{(6)})_1}{(V^{(5)})_1} = 6/8332$$
$$V^{(7)} = AV^{(6)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 29959 \\ 29230 \end{bmatrix} = \begin{bmatrix} 207526 \\ 205339 \end{bmatrix} \rightarrow \frac{(V^{(7)})_1}{(V^{(6)})_1} = 6/9270$$
$$V^{(8)} = AV^{(7)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 207526 \\ 205339 \end{bmatrix} = \begin{bmatrix} 1446121 \\ 1439560 \end{bmatrix} \rightarrow \frac{(V^{(8)})_1}{(V^{(7)})_1} = 6/9684$$
$$V^{(9)} = AV^{(8)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 1446121 \\ 1439560 \end{bmatrix} = \begin{bmatrix} 10103164 \\ 10083481 \end{bmatrix} \rightarrow \frac{(V^{(9)})_1}{(V^{(8)})_1} = 6/9864$$
$$V^{(10)} = AV^{(9)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 10103164 \\ 10083481 \end{bmatrix} = \begin{bmatrix} 70663099 \\ 70604050 \end{bmatrix} \rightarrow \frac{(V^{(10)})_1}{(V^{(9)})_1} = 6/9942$$

همانطور که مشاهده می گردد نسبت $\frac{(V_1^{(k+1)})_1}{(V_1^{(k)})_1}$ به ۷ همگراست.

این قابل پیش بینی بود زیرا مقادیر ویژه $A$ به صورت $\lambda_2 = 3$ , $\lambda_1 = 7$ می باشند. بنابراین روش توانی به مقدار ویژه غالب $A$ همگرا می باشد. توجه کنید که $V^{(10)}$ را به عنوان تقریبی از بردار ویژه متناظر با $\lambda_1 = 7$ در نظر می گیریم.

**توجه ۵.۶**

همانطور که مشاهده کردید درایه های بردارهای $V^{(k)}$ با افزایش تکرارها بزرگ خواهند شد یا ممکن است خیلی کوچک شوند. برای جلوگیری از این کار می توان در هر گام روش توانی بردار $V^{(k)}$ را به صورت

$$\widetilde{V}^{(k)} = \frac{V^{(k)}}{\|V^{(k)}\|_\infty}$$

نرمال (یا مقیاس کردن) نمود. دراین صورت بزرگترین درایه $\widetilde{V}^{(k)}$ از لحاظ اندازه برابر یک خواهد بود. به این نسخه از روش توانی، **روش توانی مقیاس شده** (scaled power method) گوییم.

**تمرین ۵.۳**

نشان دهید در روش توانی مقیاس شده $\|V^{(k)}\|_\infty \rightarrow \lambda_1$ و $\widetilde{V}^{(k)}$ یک بردار ویژه متناظر با $\lambda_1$ خواهد بود.

**مثال ۵.۱۷**

مثال قبل را با روش توانی مقیاس شده حل کنید.

**حل:** داریم $V^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. در تکرار اول داریم(محاسبات در نرم افزار متلب انجام شده است و اعداد ۴ رقم بعد از اعشار گرد شده اند)

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} 4 & 3 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

حال با نرمال کردن $V^{(1)}$ به دست می آوریم:

$$\tilde{V}^{(1)} = \frac{V^{(1)}}{\|V^{(1)}\|_\infty} = \frac{\begin{bmatrix} 4 \\ 1 \end{bmatrix}}{4} = \begin{bmatrix} 1 \\ \frac{1}{4} \end{bmatrix}$$

در تکرار دوم داریم

$$V^{(2)} = A\tilde{V}^{(1)} = \begin{bmatrix} 3 & 4 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 4/7500 \\ 2/5000 \end{bmatrix}$$

و

$$\tilde{V}^{(2)} = \frac{V^{(2)}}{\|V^{(2)}\|_\infty} = \frac{\begin{bmatrix} 4/7500 \\ 2/5000 \end{bmatrix}}{4/7500} = \begin{bmatrix} 1/0000 \\ 0/5263 \end{bmatrix}$$

با ادامه ی این روند داریم

$$V^{(3)} = A\tilde{V}^{(2)} = \begin{bmatrix} 5/5789 \\ 4/1579 \end{bmatrix}$$

$$\tilde{V}^{(3)} = \frac{V^{(3)}}{\|V^{(3)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/7453 \end{bmatrix}$$

$$V^{(4)} = A\tilde{V}^{(3)} = \begin{bmatrix} 6/2358 \\ 5/4717 \end{bmatrix}$$

$$\tilde{V}^{(4)} = \frac{V^{(4)}}{\|V^{(4)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/8775 \end{bmatrix}$$

$$V^{(5)} = A\tilde{V}^{(4)} = \begin{bmatrix} 6/6324 \\ 6/2648 \end{bmatrix} \quad , \quad \tilde{V}^{(5)} = \frac{V^{(5)}}{\|V^{(5)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9446 \end{bmatrix}$$

$$V^{(6)} = A\tilde{V}^{(5)} = \begin{bmatrix} 6/8337 \\ 6/6674 \end{bmatrix} \quad , \quad \tilde{V}^{(6)} = \frac{V^{(6)}}{\|V^{(6)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9757 \end{bmatrix}$$

$$V^{(7)} = A\tilde{V}^{(6)} = \begin{bmatrix} 6/9270 \\ 6/8540 \end{bmatrix} \quad , \quad \tilde{V}^{(7)} = \frac{V^{(7)}}{\|V^{(7)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9895 \end{bmatrix}$$

$$V^{(8)} = A\tilde{V}^{(7)} = \begin{bmatrix} 6/9684 \\ 6/9368 \end{bmatrix} \quad , \quad \tilde{V}^{(8)} = \frac{V^{(8)}}{\|V^{(8)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9955 \end{bmatrix}$$

$$V^{(9)} = A\tilde{V}^{(8)} = \begin{bmatrix} 6/9864 \\ 6/9728 \end{bmatrix} \quad , \quad \tilde{V}^{(9)} = \frac{V^{(9)}}{\|V^{(9)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9981 \end{bmatrix}$$

$$V^{(10)} = A\tilde{V}^{(9)} = \begin{bmatrix} 6/9942 \\ 6/9883 \end{bmatrix} \quad , \quad \tilde{V}^{(10)} = \frac{V^{(10)}}{\|V^{(10)}\|_\infty} = \begin{bmatrix} 1/0000 \\ 0/9992 \end{bmatrix}$$

$$\vdots$$

$$V^{(17)} = \begin{bmatrix} 7/0000 \\ 7/000 \end{bmatrix} \quad , \quad \tilde{V}^{(17)} = \begin{bmatrix} 1/0000 \\ 1/0000 \end{bmatrix}$$

توجه کنید که در تکرار ۱۷ ام داریم $\|V^{(17)}\|_\infty = 7/0000$ که همان مقدار ویژه غالب $A$ خواهد بود و $\tilde{V}^{(17)} = \begin{bmatrix} 1/0000 \\ 1/0000 \end{bmatrix}$ بردار ویژه متناظر با این مقدار ویژه غالب خواهد بود.

## ۲.۴ تشخیص علامت مقدار ویژه غالب

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} -۱ & ۱ \\ ۴ & -۱ \end{bmatrix}$$

مقادیر ویژه $A$ به صورت $\lambda_۱ = -۳,\ \lambda_۲ = ۱$ می باشند. همانطور که می بینید مقدار ویژه غالب $A$ برابر $\lambda_۱ = -۳$ که از لحاظ اندازه از دیگری بزرگتر است. حال روش توانی مقیاس شده را با $V^{(۰)} = [۰,\ ۱]^T$ اعمال می کنیم.

$$\begin{aligned}
V^{(۱)} &= AV^{(۰)} = \begin{bmatrix} -۱ & ۱ \\ ۴ & -۱ \end{bmatrix}\begin{bmatrix} ۰ \\ ۱ \end{bmatrix} = \begin{bmatrix} ۱ \\ -۱ \end{bmatrix}\\
\tilde{V}^{(۱)} &= \frac{V^{(۱)}}{\|V^{(۱)}\|_\infty} = \begin{bmatrix} ۱ \\ -۱ \end{bmatrix}\\
V^{(۲)} &= A\tilde{V}^{(۱)} = \begin{bmatrix} -۱ & ۱ \\ ۴ & -۱ \end{bmatrix}\begin{bmatrix} ۱ \\ -۱ \end{bmatrix} = \begin{bmatrix} -۲ \\ ۵ \end{bmatrix}\\
\tilde{V}^{(۲)} &= \frac{V^{(۲)}}{\|V^{(۲)}\|_\infty} = \begin{bmatrix} -۰/۴۰۰۰ \\ ۱/۰۰۰۰ \end{bmatrix}\\
V^{(۳)} &= A\tilde{V}^{(۲)} = \begin{bmatrix} ۱/۴۰۰۰ \\ -۲/۶۰۰۰ \end{bmatrix}\\
\tilde{V}^{(۳)} &= \frac{V^{(۳)}}{\|V^{(۳)}\|_\infty} = \begin{bmatrix} ۰/۵۳۸۵ \\ -۱/۰۰۰۰ \end{bmatrix}\\
V^{(۴)} &= A\tilde{V}^{(۳)} = \begin{bmatrix} -۱/۵۳۸۵ \\ ۳/۱۵۳۸ \end{bmatrix}\\
\tilde{V}^{(۴)} &= \frac{V^{(۴)}}{\|V^{(۴)}\|_\infty} = \begin{bmatrix} -۰/۴۸۷۵ \\ ۱/۰۰۰۰ \end{bmatrix}\\
&\vdots\\
V^{(۱۱)} &= \begin{bmatrix} ۱/۵۰۰۰۰ \\ -۲/۹۹۹۹ \end{bmatrix}, \quad \tilde{V}^{(۱۱)} = \begin{bmatrix} ۰/۵۰۰۰ \\ -۱/۰۰۰۰ \end{bmatrix}\\
V^{(۱۲)} &= \begin{bmatrix} -۱/۵۰۰۰ \\ ۳/۰۰۰۰ \end{bmatrix}, \quad \tilde{V}^{(۱۲)} = \begin{bmatrix} -۰/۵۰۰۰ \\ ۱/۰۰۰۰ \end{bmatrix}\\
V^{(۱۳)} &= \begin{bmatrix} ۱/۵۰۰۰ \\ -۳/۰۰۰۰ \end{bmatrix}, \quad \tilde{V}^{(۱۳)} = \begin{bmatrix} ۰/۵۰۰۰ \\ -۱/۰۰۰۰ \end{bmatrix}
\end{aligned}$$

توجه دارید که $\|V^{(۱۳)}\|_\infty = ۳/۰۰۰۰$ بنابراین مقدار ویژه غالب یا ۳/۰۰۰۰ و یا ۳/۰۰۰۰− خواهد بود.

از طرفی طبق رابطه

$$\frac{(V^{(k+۱)})_i}{(V^{(k)})_i} \approx \lambda_۱$$

اگر علامت درایه های دو بردار متوالی $V^{(k+۱)}$ و $V^{(k)}$ یکسان باشد آنگاه علامت $\lambda_۱$ مثبت خواهد بود زیرا در تقسیم مولفه های $i$ ام دو مولفه هم علامت اند و در صورتی که علامت درایه های دو بردار متوالی $V^{(k+۱)}$ و $V^{(k)}$ متفاوت باشند، مقدار ویژه $\lambda_۱$ باید منفی باشد زیرا در تقسیم مولفه های $i$ ام دو مولفه مختلف علامت اند.

از طرفی در این مثال چون علامت درایه های دو بردار متوالی $V^{(12)}$ و $V^{(13)}$ مخالف هم می باشد پس مقدار ویژه غالب $\lambda_1 = -3/0000$ خواهد بود. بعلاوه بردار $\tilde{V}^{(13)}$ به دست آمده بردار ویژه متناظر با $\lambda_1$ می باشد.

**مثال ۵.۱۸**

مقدار ویژه غالب ماتریس داده شده را با روش توانی مقیاس شده تقریب کنید.

$$A = \begin{bmatrix} 4 & 1 & 2 \\ -3 & 0 & 5 \\ 1 & 3 & 8 \end{bmatrix} \quad , \quad V^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

**حل:** داریم(محاسبات در نرم افزار متلب انجام شده است و اعداد ۴ رقم بعد از اعشار گرد شده اند)

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \quad \tilde{V}^{(1)} = \begin{bmatrix} 0/3333 \\ 0 \\ 1/0000 \end{bmatrix}$$

$$V^{(2)} = \begin{bmatrix} 3/3333 \\ 4/0000 \\ 8/3333 \end{bmatrix}, \quad \tilde{V}^{(2)} = \begin{bmatrix} 0/4000 \\ 0/4800 \\ 1/0000 \end{bmatrix}$$

$$V^{(3)} = \begin{bmatrix} 4/0800 \\ 3/8000 \\ 9/8400 \end{bmatrix}, \quad \tilde{V}^{(3)} = \begin{bmatrix} 0/4146 \\ 0/3862 \\ 1/0000 \end{bmatrix}$$

$$V^{(4)} = \begin{bmatrix} 4/0447 \\ 3/7561 \\ 9/5732 \end{bmatrix}, \quad \tilde{V}^{(4)} = \begin{bmatrix} 0/4225 \\ 1/3924 \\ 1/0009 \end{bmatrix}$$

با ادامه این روند داریم:

$$V^{(10)} = \begin{bmatrix} 4/0961 \\ 3/7188 \\ 0/5904 \end{bmatrix}, \quad \tilde{V}^{(10)} = \begin{bmatrix} 0/4271 \\ 0/3878 \\ 1/0000 \end{bmatrix}$$

$$V^{(11)} = \begin{bmatrix} 4/0962 \\ 3/7187 \\ 9/5904 \end{bmatrix}, \quad \tilde{V}^{(11)} = \begin{bmatrix} 0/4271 \\ 0/3878 \\ 1/0000 \end{bmatrix}$$

$$V^{(12)} = \begin{bmatrix} 4/0962 \\ 3/7187 \\ 9/5904 \end{bmatrix}, \quad \tilde{V}^{(12)} = \begin{bmatrix} 0/4271 \\ 0/3877 \\ 1/0000 \end{bmatrix}$$

$$V^{(13)} = \begin{bmatrix} 4/0962 \\ 3/7187 \\ 9/5904 \end{bmatrix}, \quad \tilde{V}^{(13)} = \begin{bmatrix} 0/4271 \\ 0/3877 \\ 1/0000 \end{bmatrix}$$

مشاهده می شود که درایه های دو بردار متوالی $V^{(12)}$ , $V^{(13)}$ تا ۴ رقم بعد از اعشار با هم برابرند و بعلاوه درایه های این دو بردار متوالی هم علامت اند.

بنابراین درمی یابیم که مقدار ویژه غالب $A$ برابر $\|V^{(13)}\|_\infty = 9/5904$ و با علامت مثبت است. همچنین $\tilde{V}^{(13)}$ را به عنوان تقریبی از بردار ویژه متناظر با این مقدار ویژه ی غالب در نظر می گیریم:

$$\lambda_1 \approx 9/5904 \quad , \quad X_1 \approx \begin{bmatrix} 0/4271 \\ 0/3877 \\ 1/0000 \end{bmatrix}$$

**قضیه ۵.۶**

فرض کنید $A$ ماتریس مربعی $n \times n$ با مقادیر ویژه‌ی $\lambda_1, \lambda_2, \dots, \lambda_n$ که در شرط

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

نیز صدق می‌کند باشد و $X_1, X_2, \dots, X_n$ بردارهای ویژه آن باشد. فرض کنید $\{V^{(k)}\}_{k=0}^{\infty}$ دنباله‌ی تولید شده توسط روش توانی باشد آنگاه

$$\left\| \frac{V^{(k)}}{\lambda_1^{(k)}} - \alpha_1 X_1 \right\| \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^k, \quad k = 1, 2, \dots$$

جایی که

$$C = \sum_{i=2}^{n} |\alpha_i| \|X_i\|$$

و $\|.\|$ یک نرم برداری دلخواه است. در واقع هر چه نسبت $|\frac{\lambda_2}{\lambda_1}|$ کوچکتر باشد دنباله‌ی حاصل از روش توانی سریعتر همگرا می‌گرد.

**توجه ۵.۷**

علاقمندان برای دیدن اثبات قضیه همگرایی روش توانی به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**نکته ۵.۵**

مشاهده می‌شود $\frac{V^{(k)}}{\lambda_1^k}$ به مضربی از بردار ویژه‌ی $X_1$ (یعنی $\alpha_1 X_1$ ) میل می‌کند بنابراین سرعت همگرایی روش توانی به میزان کوچکی نسبت $|\frac{\lambda_2}{\lambda_1}|$ بستگی دارد.

برای مثال دو ماتریس $A$ و $B$ با مقادیر ویژه را در نظر بگیرید.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}, \quad \lambda_1^A = 5, \quad \lambda_2^A = 1$$

$$B = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix}, \quad \lambda_1^B = 3, \quad \lambda_2^B = 2$$

نسبت $|\frac{\lambda_2}{\lambda_1}|$ برای ماتریس $A$ برابر $\frac{1}{5} = 0/2$ و برای $B$ برابر $\frac{2}{3} = 0/6667$ می‌باشد. قضیه‌ی فوق پیش‌بینی می‌کند که روش توانی برای ماتریس $A$ سریعتر از ماتریس $B$ همگرا خواهد بود.

این را می‌توان به طور مستقیم نیز بررسی کرد. با فرض اینکه $V^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ بردار شروع روش توانی باشد آنگاه بعد از ۵ تکرار برای ماتریس‌های $A$ و $B$ داریم

$$\lambda_1^A \approx 4/9979 \qquad \text{برای ماتریس } A$$

$$\lambda_1^B \approx 3/1096 \qquad \text{برای ماتریس } B$$

مشاهده می‌شود که $\lambda_1^A$ دارای دقت بیشتری نسبت به $\lambda_1^B$ است.

## ۵ افزایش سرعت همگرایی روش توانی

قضیه‌ی قبل هشدار می دهد که چنان چه نسبت $|\frac{\lambda_2}{\lambda_1}|$ به قدر کافی کوچک نباشد آنگاه روش توانی به طور ناامیدکننده‌ای می‌تواند کند باشد. بعلاوه کند بودن روش توانی را وقتی می توان متوجه شد که تعدادی از تکرارهای آن را محاسبه کرده باشیم و ببینیم که تکرارها به کندی همگرایند. سوالی که به ذهن خطور می‌کند این است که آیا می‌توان با یک سری اعمال ماتریسی بر روی ماتریس $A$ باعث افزایش سرعت روش توانی شد؟ پاسخ سوال اخیر مثبت است و می‌توان از دو ایده‌ی جالب برای افزایش سرعت روش توانی استفاده نمود:

۱- **روش توان ماتریسی**

۲- **روش توانی انتقال یافته**

### ۱.۵ روش توان ماتریسی

ماتریس $B$ را در نظر بگیرید

$$B=\begin{bmatrix}1 & -2\\ 1 & 4\end{bmatrix},\quad |\frac{\lambda_2}{\lambda_1}|=\frac{2}{3}=0/6667$$

فرض کنید ماتریس $B$ را به توان پنج برسانیم آنگاه

$$B^5=\begin{bmatrix}-179 & -422\\ 211 & 454\end{bmatrix}$$

حال با توجه به این حقیقت که $\lambda(B^k)=(\lambda(B))^k$ برای هر $k$، می‌توان دید که نسبت به مقادیر ویژه $\frac{\lambda_2^{B^5}}{\lambda_1^{B^5}}$ برای ماتریس $B^5$ برابر است با

$$|\frac{\lambda_2^{B^5}}{\lambda_1^{B^5}}|=(\frac{2}{3})^5=0/1317 \qquad \text{برای}\ B^5$$

که مقدار کوچکی (نسبت به $|\frac{\lambda_2^B}{\lambda_1^B}|=\frac{2}{3}=0/6667$) می‌باشد. از اینرو اگر روش توانی را برای $B^5$ بکار بگیریم آنگاه سرعت همگرایی این روش به $0/1317$ وابسته خواهد بود و لذا خیلی سریعتر همگرا خواهد شد. اکنون اگر $\lambda_1^{B^5}$ تقریبی از مقدار ویژه‌ی غالب $B^5$ باشد آنگاه با گرفتن ریشه‌ی پنجم از $\lambda^{B^5}$ تقریبی از مقدار ویژه‌ی $B$ حاصل می‌شود.
با اعمال روش توانی با $V^{(0)}=\begin{bmatrix}0\\ 1\end{bmatrix}$ بر روی $B^5$ بعد از سه تکرار داریم

$$\lambda^{B^5}\approx 244/8455$$

از اینرو با گرفتن ریشه‌ی پنجم داریم

$$\lambda^B\approx(244/8455)^{\frac{1}{5}}\approx 3/0045$$

که به مقدار قابل قبولی به ۳ نزدیک است.
به ایده‌ی گفته شده، روش توان ماتریسی می‌گوییم. بنابراین می‌توان گفت که:

**توجه ۵.۸**

در روش توان ماتریسی ابتدا ماتریس $A$ را به توان عددی صحیح مثل $k$ می‌رسانیم و سپس روش توانی (یا توانی مقیاس شده) را برای $A^k$ اعمال کرده و در نهایت از مقدار ویژه‌ی به دست آمده برای $A^k$ ریشه‌ی $k$ ام می‌گیریم.

**توجه ۵.۹**

با توجه به اینکه برای $k$ های زوج $\sqrt[k]{\lambda}$ و $-\sqrt[k]{\lambda}$ ریشه ی $k$ ام $\lambda$ هستند پس برای اینکه بتوانیم علامت مقدار ویژه را تشخیص دهیم می‌بایست $k$ را عددی فرد در نظر بگیریم تا مقدار ویژه ی غالب $A$ تعیین گردد. البته باید تاکید کنیم هیچ الگوریتم یا روشی برای تعیین $k$ مناسب وجود ندارد. هرچه $k$ بزرگتر باشد الگوریتم توانی برای $A^k$ زودتر همگرا می گردد.

**توجه ۵.۱۰**

در عمل تشخیص کند بودن روش توانی ممکن نیست مگر اینکه تعدادی از تکرارها را به دست آوریم و ببینیم که دنباله ی حاصل همگراست ولی کند است تا بخواهیم از روش توان ماتریسی استفاده کنیم.

**مثال ۵.۱۹**

ماتریس زیر را در نظر بگیرید

$$A=\begin{bmatrix} ۵ & -۱ & -۳ \\ ۰ & ۵ & ۳ \\ ۰ & -۱ & -۱ \end{bmatrix}$$

**حل-** این ماتریس دارای مقادیر ویژه‌ی ( $\lambda_۱$ مقدار ویژه‌ی غالب است )

$$\lambda_۱=۵, \quad \lambda_۲=۴/۴۴۹۵, \quad \lambda_۳=-۰/۴۴۹۵$$

می‌باشد. با توجه به اینکه $|\frac{\lambda_۲}{\lambda_۱}|=|\frac{۴/۴۴۹۵}{۵}|=۰/۸۸۹۹$ مقداری نزدیک به ۱ است پس روش توانی (یا مقیاس شده) برای ماتریس $A$ به کندی همگرا می‌شود. حال اگر ماتریس $A^{۱۱}$ را در نظر بگیریم این نسبت مقادیر ویژه برای ماتریس $A^{۱۱}$ برابر است با

$$|\frac{\lambda_۲}{\lambda_۱}|^{۱۱}=(۰/۸۸۹۹)^{۱۱}=۰/۲۷۷۲,$$

که مقداری نسبتا کوچک است. از اینرو انتظار داریم روش توانی (یا مقیاس شده) برای $A^{۱۱}$ با سرعت بیشتری همگرا گردد. داریم

$$\hat{A}=A^{۱۱}=\begin{bmatrix} ۴۸۸۲۸۱۲۵ & -۳۶۵۳۶۲۸۵ & -۴۲۰۶۱۳۴۱ \\ ۰ & ۱۵۰۵۴۳۶۸ & ۸۲۸۷۵۸۴ \\ ۰ & -۲۷۶۲۵۲۸ & -۱۵۲۰۸۰۰ \end{bmatrix}$$

حال با اعمال روش توانی مقیاس شده و با $V^{(۰)}=[۰,\ ۰,\ ۱]^T$ داریم( تمام محاسبات را تا ۶ رقم بعد از اعشار گرد می‌کنیم) $k=۱$ :

$$V^{(۱)}=\hat{A}V^{(۰)}, \quad \|V^{(۱)}\|_\infty=۴۲۰۶۱۳۴۱$$

$$\tilde{V}^{(1)} = \frac{V^{(1)}}{\|V^{(1)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/197035 \\ -0/036156 \end{bmatrix}$$

توجه کنید به علت بزرگی درایه‌های $V^{(1)}$، این بردار نمایش داده نشده و نرمال شده‌ی آن یعنی $\tilde{V}^{(1)}$ را نشان داده‌ایم. در تکرار بعدی داریم

$k = 2$ :

$$V^{(2)} = \hat{A}V^{(1)}, \quad \|V^{(2)}\|_\infty = 5/450627 \times 10^7$$

$$\tilde{V}^{(2)} = \frac{V^{(2)}}{\|V^{(2)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/048922 \\ -0/008977 \end{bmatrix}$$

$k = 3$ :

$$V^{(3)} = \hat{A}V^{(2)}, \quad \|V^{(3)}\|_\infty = 5/023797 \times 10^7$$

$$\tilde{V}^{(3)} = \frac{V^{(3)}}{\|V^{(3)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/013179 \\ -0/002418 \end{bmatrix}$$

$$\vdots$$

$k = 15$ :

$$V^{(15)} = \hat{A}V^{(14)}, \quad \|V^{(15)}\|_\infty = 4/882812 \times 10^7$$

$$\tilde{V}^{(15)} = \frac{V^{(15)}}{\|V^{(15)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/000000 \\ -0/000000 \end{bmatrix}$$

$k = 16$ :

$$V^{(16)} = \hat{A}V^{(15)}, \quad \|V^{(16)}\|_\infty = 4/882812 \times 10^7$$

$$\tilde{V}^{(16)} = \frac{V^{(16)}}{\|V^{(16)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/000000 \\ 0/000000 \end{bmatrix}$$

بنابراین می توان $4/882812 \times 10^7$ را به عنوان تقریبی از مقدار ویژه‌ی غالب ماتریس $A^{11}$ در نظر گرفت. اکنون با محاسبه‌ی ریشه‌ی یازدهم این مقدار داریم

$$\lambda_1 \approx (4/882812 \times 10^7)^{\frac{1}{11}} \approx 5/000000$$

که مقدار مناسبی از مقدار ویژه‌ی غالب $A$ را نشان می‌دهد.
توجه کنید روش توانی مقیاس شده و با اعمال روی ماتریس $A$ بعد از ۱۰۰ تکرار نتایجی به صورت زیر خواهد داشت:

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix}$$

$$\|V^{(1)}\|_\infty = 3, \quad \tilde{V}^{(1)} = \frac{V^{(1)}}{\|V^{(1)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 1/000000 \\ -0/333333 \end{bmatrix}$$

$$V^{(2)} = AV^{(1)} = \begin{bmatrix} -5/000000 \\ 4/00000 \\ -0/666667 \end{bmatrix}$$

$$\|V^{(2)}\|_\infty = 5, \quad \tilde{V}^{(2)} = \frac{V^{(2)}}{\|V^{(2)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/800000 \\ -0/133333 \end{bmatrix}$$

$$\vdots$$

$$V^{(100)} = AV^{(99)} = \begin{bmatrix} -5/000003 \\ 0/000026 \\ -0/000005 \end{bmatrix}$$

$$\|V^{(100)}\|_\infty = 5/000003, \quad \tilde{V}^{(100)} = \frac{V^{(100)}}{\|V^{(100)}\|_\infty} = \begin{bmatrix} -1/000000 \\ 0/000005 \\ -0/000001 \end{bmatrix}$$

بنابراین بعد از ۱۰۰ تکرار روش توانی مقیاس شده تقریب ۵/۰۰۰۰۰۳ را برای مقدار ویژه‌ی غالب $A$ به دست می‌دهد. این بدین معنی است که توانی مقیاس شده تنها با ۱۶ تکرار به جوابی بهتر در مقایسه با توانی با صد تکرار منجر می شود.

**نکته ۵.۶**

توجه کنید که بردار ویژه‌ی به دست آمده برای هر دو ماتریس $A^{11}$ و $A$ یکسان بوده بنابراین می‌توان دید که بردار ویژه‌ی غالب $A$ به صورت زیر است.

$$\begin{bmatrix} -1/000000 \\ 0/000000 \\ 0/000000 \end{bmatrix}$$

که در واقع تقریب دقیقی از بردار ویژه‌ی غالب $X_1 = [-1, \ 0, \ 0]^T$ است.

**تمرین ۵.۴**

با استفاده از روش توان ماتریسی به ازای $k = 5$ مقادیر ویژه غالب ماتریس زیر را تقریب کنید.

$$A = \begin{bmatrix} -2 & -1 & -4 \\ -3 & 0 & -2 \\ -2 & 1 & 4 \end{bmatrix}, \qquad \lambda_1 \approx 5/0517$$

- با توجه به اینکه $\lambda_1 \approx 5/0517$ تعداد تکرارها را تا جایی ادامه دهید که جواب تقریبی $\tilde{\lambda}$ در $|\tilde{\lambda} - \lambda_1| \leq 5 \times 10^{-4}$ صدق کند.
- با استفاده از روش توانی برای $A$ برای برقراری شرط توقف بیان شده به چند تکرار نیاز داریم؟

## ۲.۵ روش توانی با انتقال

(shifted power method)

قبلا جهت افزایش سرعت همگرایی روش توانی، روش توان ماتریسی را معرفی کردیم. در اینجا از یک ایده دیگر برای افزایش سرعت همگرایی روش توانی کمک میگیریم که بر پایه قضیه زیر است.

**قضیه ۵.۷**

فرض کنید $\lambda$ یک مقدار ویژه ماتریس مربعی $A$ باشد و $\alpha$ یک اسکالر دلخواه باشد آنگاه مقادیر ویژه ماتریس جدید $B = A - \alpha I$ به صورت $\lambda - \alpha$ خواهند بود.

**اثبات:** فرض کنید $(\lambda, X)$ یک جفت ویژه $A$ باشد آنگاه

$$AX = \lambda X \tag{۱۸}$$

با کم کردن $\alpha X$ از طرفین (۱۸) داریم

$$AX - \alpha X = \lambda X - \alpha X$$

یا

$$(A - \alpha I)X = (\lambda - \alpha)X \Rightarrow BX = (\lambda - \alpha)X \tag{۱۹}$$

اکنون رابطه (۱۹) نشان می دهد که $(\lambda - \alpha, X)$ یک جفت ویژه $B$ خواهد بود.

**توجه ۵.۱۱**

به خاصیت فوق برای مقادیر ویژه، خاصیت انتقال (shift property) گفته می شود و به طور خلاصه به صورت زیر نمایش داده می شود

$$\lambda(A - \alpha I) = \lambda(A) - \alpha$$

حال فرض کنید روش توانی برای ماتریس $A$ کند باشد یعنی اگر $\lambda_۱, \lambda_۲$ دو مقدار ویژه بزرگ $A$ از لحاظ قدرمطلق باشند، نسبت $|\frac{\lambda_۲}{\lambda_۱}|$ به یک نزدیک باشد. با توجه به اینکه مقادیر ویژه $B = A - \alpha I$ به صورت $\lambda_۱ - \alpha, \lambda_۲ - \alpha, \cdots$ می باشند اگر $|\lambda_۱ - \alpha| > |\lambda_۲ - \alpha| \geq \cdots \geq |\lambda_n - \alpha|$ برقرار باشد آنگاه نرخ همگرایی روش توانی برای ماتریس $B$ به نسبت $|\frac{\lambda_۲ - \alpha}{\lambda_۱ - \alpha}|$ وابسته خواهد بود.

هدف از روش توانی با انتقال، یافتن $\alpha$ ای است که

$$|\frac{\lambda_۲ - \alpha}{\lambda_۱ - \alpha}| < |\frac{\lambda_۲}{\lambda_۱}| \tag{۲۰}$$

یعنی $\alpha$ ای بیابیم که روش توانی برای ماتریس $B = A - \alpha I$ سریعتر همگرا گردد. واضح است که اگر چنین $\alpha$ یافت شود آنگاه کافی است روش توانی برای ماتریس $B$ اعمال شود و مقدار ویژه غالب $B$ را یافت و از آنجایی که

$$\lambda(B) = \lambda(A) - \alpha \Rightarrow \lambda(A) = \lambda(B) + \alpha$$

کافی است مقدار $\alpha$ به مقدار ویژه غالب $B$ اضافه گردد تا مقدار ویژه غالب $A$ محاسبه گردد.

در ادامه نشان می دهیم که چنین $\alpha$ ای می تواند پیدا شود. برای مثال فرض کنید $A$ ماتریس $۵ \times ۵$ با مقادیر ویژه

$$\lambda_۱^A = ۵, \quad \lambda_۲^A = ۴, \quad \lambda_۳^A = ۳, \quad \lambda_۴^A = ۲, \quad \lambda_۵^A = ۱$$

باشد اگر $\alpha = ۱$ فرض شود آنگاه مقادیر ویژه $B = A - \alpha I = A - I$ به صورت زیر خواهند بود

$$\lambda_۱^B = ۴, \quad \lambda_۲^B = ۳, \quad \lambda_۳^B = ۲, \quad \lambda_۴^B = ۱, \quad \lambda_۵^B = ۰$$

حال نسبت $|\frac{\lambda_2^A}{\lambda_1^A}|$ برای $A$ برابر $0/8 = \frac{4}{5}$ است در حالی که برای $B$ برابر است با $0/75 = \frac{3}{4} = |\frac{\lambda_2^B}{\lambda_1^B}|$ و

$$0/75 = |\frac{\lambda_2^B}{\lambda_1^B}| < |\frac{\lambda_2^A}{\lambda_1^A}| = 0/8$$

حال اگر $\alpha = 2/5$ فرض شود آنگاه مقادیر ویژه ماتریس $B = A - 2/5I$ برابرند با

$$\lambda_1^B = 2/5, \quad \lambda_2^B = 1/5, \quad \lambda_3^B = 0/5, \quad \lambda_4^B = -0/5, \quad \lambda_5^B = -1/5$$

آنگاه $0/6 = \frac{1/5}{2/5} = |\frac{\lambda_2^B}{\lambda_1^B}|$ و

$$0/6 = |\frac{\lambda_2^B}{\lambda_1^B}| < |\frac{\lambda_2^A}{\lambda_1^A}| = 0/8.$$

بنابراین اگر روش توانی روی ماتریس $B = A - 2/5I$ اعمال شود دارای سرعت همگرایی بیشتری خواهد بود. با توجه به مثال فوق ۲ سوال اساسی به ذهن خطور خواهد کرد:

- نحوه انتخاب $\alpha$ به چه صورت است؟
- بهترین انتخاب $\alpha$ چگونه خواهد بود؟

باید گفت که در حالت کلی نمی توان پارامتر $\alpha$ را محاسبه نمود و تنها در حالت های خاصی مقادیری برای $\alpha$ مشخص شده است. مثلا فرض کنید مقادیر ویژه $A$ همگی مثبت باشند و $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n > 0$ آنگاه $|\frac{\lambda_2}{\lambda_1}| = \frac{\lambda_2}{\lambda_1}$. حال اگر مقادیر ویژه $B = A - \alpha I$ در شرط $|\lambda_1 - \alpha| > |\lambda_2 - \alpha| \geq \cdots \geq |\lambda_n - \alpha| > 0$ صدق کنند آنگاه $\alpha$ مناسب باید در شرط $|\frac{\lambda_2 - \alpha}{\lambda_1 - \alpha}| < \frac{\lambda_2}{\lambda_1}$ صدق کند. یعنی $-\frac{\lambda_2}{\lambda_1} < \frac{\lambda_2 - \alpha}{\lambda_1 - \alpha} < \frac{\lambda_2}{\lambda_1}$ . بنابراین باید نامعادلات زیر حل گردند

$$\begin{cases} \frac{\lambda_2 - \alpha}{\lambda_1 - \alpha} < \frac{\lambda_2}{\lambda_1} \qquad (1) \\ \frac{\lambda_2 - \alpha}{\lambda_1 - \alpha} > -\frac{\lambda_2}{\lambda_1} \qquad (2) \end{cases}$$

اکنون می توان دید که (۱) فقط برای $0 < \alpha < \lambda_1$ برقرار است و (۲) نیز وقتی برقرار است که $\alpha > \lambda_1$ یا $\alpha < \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2}$. بنابراین باید $\alpha$ در شرایط زیر صدق کند

$$\begin{cases} 0 < \alpha < \lambda_1 \\ \alpha < \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2} \end{cases}$$

که معادل با این است که $\alpha$ در شرط $0 < \alpha < \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2}$ صادق باشد.
بنابراین قضیه زیر را ثابت کردیم:

**قضیه ۵.۸**

فرض کنید $A$ ماتریسی با مقادیر ویژه مثبت که در شرط $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ صدق می کنند باشد و $\alpha \in (0, \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2})$. اگر مقادیر ویژه $B = A - \alpha I$ در شرط $|\lambda_1 - \alpha| > |\lambda_2 - \alpha| \geq \cdots \geq |\lambda_n - \alpha| > 0$ صدق کنند آنگاه نرخ همگرایی روش توانی برای ماتریس $B = A - \alpha I$ سریعتر از نرخ همگرایی روش توانی برای $A$ خواهد بود.

**توجه ۵.۱۲**

با توجه به اینکه وقتی $A$ متقارن معین مثبت باشد مقادیر ویژه اش مثبت اند، پس اگر $A$ متقارن معین مثبت و با مقادیر ویژه غالب $\lambda_1$ که در شرط

$$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$$

نیز صدق می کند باشد آنگاه می توان قضیه فوق را برای $A$ بکار گرفت. یعنی برای ماتریس های متقارن مثبت معین می توان مقدار پارامتر $\alpha$ را پیدا کرد

**تمرین ۵.۵**

نشان دهید به ازای $\alpha = \lambda_n$ و با شرط اینکه $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n > 0$ نیز روش توانی سریعتر همگرا می گردد.

## ۳.۵ انتخاب بهینه پارامتر $\alpha$

در ادامه قصد داریم بهترین پارامتر $\alpha$ را از لحاظ تئوری تعیین نماییم.

فرض کنید $A$ ماتریسی با مقادیر ویژه $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n > 0$ باشد. اگر مقادیر ویژه $B = A - \alpha I$ در شرط

$$|\lambda_1 - \alpha| > |\lambda_2 - \alpha| \geq |\lambda_3 - \alpha| \geq \cdots \geq |\lambda_n - \alpha| \qquad (21)$$

صدق کنند می خواهیم مقدار بهینه $\alpha$ را بیابیم. با تقسیم طرفین (۲۱) بر $|\lambda_1 - \alpha|$ (فرض می کنیم $\alpha \neq \lambda_1$ ) داریم

$$|\frac{\lambda_n - \alpha}{\lambda_1 - \alpha}| \leq \cdots \leq |\frac{\lambda_3 - \alpha}{\lambda_1 - \alpha}| \leq |\frac{\lambda_2 - \alpha}{\lambda_1 - \alpha}| < 1 \qquad (22)$$

همانطور که میبینید به ازای $\alpha$ های مختلف مقادیر مختلفی برای $|\frac{\lambda_2-\alpha}{\lambda_1-\alpha}|$ داریم. اما می خواهیم $\alpha$ را طوری بیابیم که مقدار $|\frac{\lambda_2-\alpha}{\lambda_1-\alpha}|$ کمترین مقدار خود را داشته باشد.

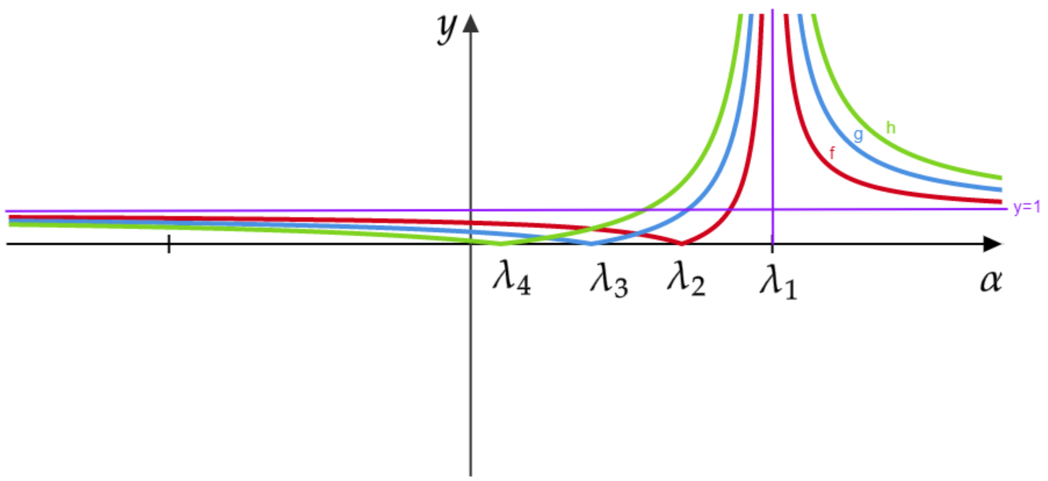
اکنون برای حل این مسئله برای راحتی $n = 4$ بگیرید پس

$$|\frac{\lambda_4 - \alpha}{\lambda_1 - \alpha}| \leq |\frac{\lambda_3 - \alpha}{\lambda_1 - \alpha}| \leq |\frac{\lambda_2 - \alpha}{\lambda_1 - \alpha}| < 1 \qquad (23)$$

تعریف می کنیم

$$f(\alpha) = |\frac{\lambda_2 - \alpha}{\lambda_1 - \alpha}|, \quad g(\alpha) = |\frac{\lambda_3 - \alpha}{\lambda_1 - \alpha}|, \quad h(\alpha) = |\frac{\lambda_4 - \alpha}{\lambda_1 - \alpha}|$$

پس طبق (۲۳) داریم $h(\alpha) \leq g(\alpha) \leq f(\alpha)$ واضح است که $h(\lambda_4) = 0, g(\lambda_3) = 0, f(\lambda_2) = 0$ اکنون نمودار توابع $f, g, h$ را برای $\lambda_2 \geq \lambda_3 \geq \lambda_4$ رسم می کنیم. توجه کنید که خط $y = 1$ مجانب افقی و خط $\alpha = \lambda_1$ مجانب عمودی برای هر سه تابع است. همانطور که در شکل نیز مشخص شده است، تابع $f$ رنگ قرمز، $g$ با رنگ آبی و $h$ با رنگ سبز مشخص شده اند. حال باید از توابع در شکل ، ماکزیمم بگیریم. مطابق این شکل، ماکزیمم نمودار ها برای $\alpha > \lambda_1$ همان تابع سبز رنگ یعنی $h$ است در حالی که به ازای $\alpha < \lambda_1$ از دو تابع سبز و قرمز رنگ یعنی $h$ و $f$ تشکیل شده است. بعلاوه مشاهده می شود که در این شکل ماکزیمم زمانی کمینه است که دو تابع $f$ و $h$ همدیگر را قطع کنند که نقطه تقاطع همان پارامتر بهینه $\alpha$ است. دقت کنید که به ازای $\alpha > \lambda_1$ هیچ یک از توابع یکدیگر را قطع نمی کند بنابراین پارامتر بهینه $\alpha$ در بازه $(0, \lambda_1)$ صدق می کند. با توجه به بحث های انجام شده کافی است نقطه تقاطع توابع $f$ و $h$ را بیابیم.




یعنی باید مسئله زیر حل شود:

$$f(\alpha) = g(\alpha) \Rightarrow |\frac{\lambda_۲ - \alpha}{\lambda_۱ - \alpha}| = |\frac{\lambda_۴ - \alpha}{\lambda_۱ - \alpha}|$$

یا $|\lambda_۲ - \alpha| = |\lambda_۴ - \alpha|$. بنابراین $\lambda_۲ - \alpha = \lambda_۴ - \alpha$ یا $\lambda_۲ - \alpha = -(\lambda_۴ - \alpha)$ این نتیجه می دهد که $\alpha = \frac{۱}{۲}(\lambda_۲ + \lambda_۴)$ قابل قبول است. بنابراین پارامتر بهینه به صورت زیر حاصل می شود:

$$\alpha_{opt} = \frac{۱}{۲}(\lambda_۲ + \lambda_۴)$$

اکنون اگر استدلال فوق را برای هر $n \geq ۲$ بکار ببریم می توانیم دریابیم که پارامتر بهینه $\alpha$ به صورت

$$\alpha_{opt} = \frac{۱}{۲}(\lambda_۲ + \lambda_n)$$

محاسبه خواهد شد.

البته در عمل فرمول بالا کاربردی نخواهد داشت زیرا در حالت کلی اطلاعاتی از مقادیر ویژه $\lambda_۲$ و $\lambda_n$ در دسترس نیست!

**مثال ۵.۲۰**

ماتریس سه قطری زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۵ & -۱ & ۰ \\ -۱ & ۵ & -۱ \\ ۰ & -۱ & ۵ \end{bmatrix}$$

که دارای مقادیر ویژه زیر است:

$$\lambda_۱ = ۶/۴۱۴۲, \qquad \lambda_۲ = ۵/۰۰۰۰, \qquad \lambda_۳ = ۳/۵۸۵۸$$

با توجه به اینکه $|\frac{\lambda_۲}{\lambda_۱}| = \frac{۵/۰۰۰}{۶/۴۱۴۲} = ۰/۷۷۹۵$ پس روش توانی برای این ماتریس می تواند به کندی همگرا شود. جدول زیر نشان دهنده روش توانی (مقیاس شده) برای ماتریس $A$ است.

| $k$ | $V^{(k)}$ | $\tilde{V}^{(k)}$ | $\lambda^{(k)}$ |
|---|---|---|---|
| ۰ | $[۰/۰۰۰۰, ۰/۰۰۰۰, ۱/۰۰۰۰]^T$ | - | - |
| ۱ | $[۰/۰۰۰۰, -۱/۰۰۰۰, ۵/۰۰۰۰]^T$ | $[۰/۰۰۰۰, -۰/۲۰۰۰, ۱/۰۰۰۰]^T$ | ۵/۰۰۰۰ |
| ۲ | $[۰/۲۰۰۰, -۲/۰۰۰۰, ۵/۲۰۰۰]^T$ | $[۰/۰۳۸۵, -۰/۳۸۴۶, ۱/۰۰۰۰]^T$ | ۵/۲۰۰۰ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| ۲۱ | $[۴/۴۸۷۱, -۶/۴۱۴۲, ۴/۵۸۴۱]^T$ | $[۰/۶۹۹۵, -۱/۰۰۰۰, ۰/۷۱۴۷]^T$ | ۶/۴۱۴۲ |
| ۲۲ | $[۴/۴۹۷۷, -۶/۴۱۴۲, ۴/۵۷۳۴]^T$ | $[۰/۷۰۱۲, -۱/۰۰۰۰, ۰/۷۱۳۰]^T$ | ۶/۴۱۴۲ |

لذا بعد از ۲۲ تکرار تقریبی از مقدار ویژه غالب $A$ به صورت ۶/۴۱۴۲ به دست می آید. حال با توجه به اینکه

$$\alpha = \frac{۱}{۲}(\lambda_۲ + \lambda_۳) = \frac{۱}{۲}(۵/۰۰۰۰ + ۳/۵۸۵۸) = ۴/۲۹۲۹$$

داریم

$$B = A - \alpha I = A - ۴/۲۹۲۹ I = \begin{bmatrix} ۰/۷۰۷۱ & -۱/۰۰۰۰ & ۰/۰۰۰۰ \\ -۱/۰۰۰۰ & ۰/۷۰۷۱ & -۱/۰۰۰۰ \\ ۰/۰۰۰۰ & -۱/۰۰۰ & ۰/۷۰۷۱ \end{bmatrix}$$

از طرفی مقادیر ویژه ماتریس $B$ به صورت زیر به دست می آیند:

$$\lambda_۱^B \approx ۲/۱۲۱۳, \qquad \lambda_۲^B \approx ۰/۷۰۷۱, \qquad \lambda_۳^B \approx -۰/۷۰۷۱$$

بنابراین روش توانی با نرخ $|\frac{\lambda_۲^B}{\lambda_۱^B}| = \frac{۰/۷۰۷۱}{۲/۱۲۱۳} = ۰/۳۳۳۳$ برای ماتریس $B$ همگرا خواهد بود و این یعنی این روش برای ماتریس $B$ سریعتر از ماتریس $A$ همگرا می شود. جدول زیر نتایج روش توانی مقیاس شده را برای ماتریس $B$ نشان می دهد.

| $k$ | $V^{(k)}$ | $\tilde{V}^{(k)}$ | $\lambda^{(k)}$ |
|---|---|---|---|
| ۰ | $[۰/۰۰۰۰, ۰/۰۰۰۰, ۱/۰۰۰۰]^T$ | - | - |
| ۱ | $[۰/۰۰۰۰, -۱/۰۰۰۰, ۰/۷۰۷۱]^T$ | $[۰/۰۰۰۰, -۱/۰۰۰۰, ۰/۷۰۷۱]^T$ | ۱/۰۰۰۰ |
| ۲ | $[۱/۰۰۰۰, -۱/۴۱۴۲, ۱/۵۰۰۰]^T$ | $[۰/۶۶۶۷, -۰/۹۴۲۸, ۱/۰۰۰۰]^T$ | ۱/۵۰۰۰ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| ۱۲ | $[۱/۵۰۰۰, -۲/۱۲۱۳, ۱/۵۰۰۰]^T$ | $[۰/۷۰۷۱, -۱/۰۰۰۰, ۰/۷۰۷۱]^T$ | ۲/۱۲۱۳ |
| ۱۳ | $[۱/۵۰۰۰, -۲/۱۲۱۳, ۱/۵۰۰۰]^T$ | $[۰/۷۰۷۱, -۱/۰۰۰۰, ۰/۷۰۷۱]^T$ | ۲/۱۲۱۳ |

مشاهده می شود که روش توانی مقیاس شده بعد از ۱۳ تکرار تقریبی از مقدار ویژه غالب $B$ به صورت ۲/۱۲۱۳ به دست می دهد. در اینصورت مقدار ویژه غالب ماتریس $A$ را می توان به صورت زیر بدست آورد:

$$\lambda_۱^A \approx ۲/۱۲۱۳ + \alpha = ۲/۱۲۱۳ + ۴/۲۹۲۹ = ۶/۴۱۴۲$$

**مثال ۵.۲۱**

ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۷ & ۲ & ۱ & ۸ \\ ۱۲ & ۲۳ & ۱۳ & ۳ \\ ۴ & ۱۰ & ۲۴ & -۶ \\ ۱۶ & -۸ & ۱۰ & ۲۷ \end{bmatrix}$$

مقادیر ویژه $A$ به صورت زیر هستند:

$$\lambda_۱ \approx ۳۵/۸۲۵۷, \qquad \lambda_۲ \approx ۲۵/۶۴۵۵, \qquad \lambda_۳ \approx ۱۸/۵۷۱۷, \qquad \lambda_۴ \approx ۰/۹۵۷۰$$

نرخ همگرایی روش توانی برای $A$ متناسب با $\frac{\lambda_۲}{\lambda_۱} = ۰/۷۱۵۸$ می باشد که نشان می دهد این روش می تواند کند باشد. در واقع اگر تکرار های روش توانی (مقایس شده) را با $V^{(۰)} = [۰,\ ۰,\ ۰,\ ۱]^T$ شروع کنیم داریم

| $k$ | $V^{(k)}$ | $\tilde{V}^{(k)}$ | $\lambda^{(k)}$ |
|---|---|---|---|
| ۰ | $[۰/۰۰۰۰, ۰/۰۰۰۰, ۰/۰۰۰۰, ۱/۰۰۰۰]^T$ | - | - |
| ۱ | $[۸/۰۰۰۰, ۳/۰۰۰۰, -۶/۰۰۰۰, ۲۷/۰۰۰۰]^T$ | $[۰/۲۹۶۳, ۰/۱۱۱۱, -۰/۲۲۲۲, ۱/۰۰۰۰]^T$ | ۲۷/۰۰۰۰ |
| ۲ | $[۱۰/۰۷۴۱, ۶/۲۲۲۲, -۹/۰۳۷۰, ۲۸/۶۲۹۶]^T$ | $[۰/۳۵۱۹, ۰/۲۱۷۳, -۰/۳۱۵۷, ۱/۰۰۰۰]^T$ | ۲۸/۶۲۹۶ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| ۴۴ | $[۶/۱۳۹۹, ۳۵/۸۲۵۷, ۲۷/۴۳۰۵, ۹/۷۳۷۸]^T$ | $[۰/۱۷۱۴, ۱/۰۰۰۰, ۰/۷۶۵۷, ۰/۲۷۱۸]^T$ | ۳۵/۸۲۵۷ |
| ۴۵ | $[۶/۱۳۹۸, ۳۵/۸۲۵۷, ۲۷/۴۳۰۷, ۹/۷۳۷۷]^T$ | $[۰/۱۷۱۴, ۱/۰۰۰۰, ۰/۷۶۵۷, ۰/۲۷۱۸]^T$ | ۳۵/۸۲۵۷ |

بنابراین روش توانی مقیاس شده برای $A$ بعد از ۴۵ تکرار به جواب تقریبی $\lambda \approx ۳۵/۸۲۵۷$ به عنوان مقدار ویژه غالب $A$ خواهد رسید. حال فرض کنید هدف بکار گیری روش توانی انتقال یافته با پارامتر انتقال

$$\alpha = \frac{۱}{۲}(\lambda_۲ + \lambda_۴) = \frac{۱}{۲}(۲۵/۶۴۵۵ + ۰/۹۵۷۰) \approx ۱۳/۳۰۱۳$$

می باشد. در اینصورت داریم

$$B = A - \alpha I = A - ۱۳/۳۰۱۳I = \begin{bmatrix} -۶/۳۰۱۳ & ۲/۰۰۰ & ۱/۰۰۰ & ۸/۰۰۰ \\ ۱۲/۰۰۰ & ۹/۶۹۸۷ & ۱۳/۰۰۰ & ۳/۰۰۰ \\ ۴/۰۰۰ & ۱۰/۰۰۰ & ۱۰/۶۹۸۷ & -۶/۰۰۰ \\ ۱۶/۰۰۰ & -۸/۰۰۰ & ۱۰/۰۰۰ & ۱۳/۶۹۸۷ \end{bmatrix}$$

در این صورت مقادیر ویژه ماتریس $B$ به صورت زیرند:

$$\lambda_۱^B \approx ۲۲/۵۲۴۴, \quad \lambda_۲^B \approx ۱۲/۳۴۴۲, \quad \lambda_۳^B \approx -۱۲/۳۴۴۲, \quad \lambda_۴^B \approx ۵/۲۷۰۵$$

لذا روش توانی برای $B$ با نرخ همگرایی $|\frac{\lambda_۲^B}{\lambda_۱^B}| \approx \frac{۱۲/۳۴۴۲}{۲۲/۵۲۴۴} \approx ۰/۵۴۸۰$ همگرا می گردد. همانطور که مشاهده می شود نرخ همگرایی از ۰/۷۱۵۸ به ۰/۵۴۸۰ تبدیل شده است پس انتظار داریم روش توانی برای ماتریس $B$ با سرعت بیشتری همگرا گردد. جدول زیر نتایج روش توانی مقیاس شده را برای ماتریس $B$ نشان می دهد:

| $k$ | $V^{(k)}$ | $\tilde{V}^{(k)}$ | $\lambda^{(k)}$ |
|---|---|---|---|
| ۰ | $[۰/۰۰۰۰, ۰/۰۰۰۰, ۰/۰۰۰۰, ۱/۰۰۰۰]^T$ | - | - |
| ۱ | $[۸/۰۰۰۰, ۳/۰۰۰۰, -۶/۰۰۰۰, ۱۳/۶۹۸۷]^T$ | $[۰/۵۸۴۰, ۰/۲۱۹۰, -۰/۴۳۸۰, ۱/۰۰۰۰]^T$ | ۱۳/۶۹۸۷ |
| ۲ | $[۴/۳۲۰۱, ۶/۴۳۸۰, -۶/۱۶۰۰, ۱۶/۹۱۰۷]^T$ | $[۰/۲۵۵۵, ۰/۳۸۰۷, -۰/۳۶۴۳, ۱/۰۰۰۰]^T$ | ۱۶/۹۱۰۷ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| ۲۶ | $[۳/۸۶۰۲, ۲۲/۵۲۴۴, ۱۷/۲۴۶۴, ۶/۱۲۲۲]^T$ | $[۰/۱۷۱۴, ۱/۰۰۰۰, ۰/۷۶۵۷, ۰/۲۷۱۸]^T$ | ۲۲/۵۲۴۴ |
| ۲۷ | $[۳/۸۶۰۲, ۲۲/۵۲۴۴, ۱۷/۲۴۶۴, ۶/۱۲۲۲]^T$ | $[۰/۱۷۱۴, ۱/۰۰۰۰, ۰/۷۶۵۷, ۰/۲۷۱۸]^T$ | ۲۲/۵۲۴۴ |

مشاهده می شود که روش توانی مقیاس شده بعد از ۲۷ تکرار تقریبی از مقدار ویژه غالب به صورت ۲۲/۵۲۴۴ به دست آورده است. حال با افزودن مقدار $\alpha = ۱۳/۳۰۱۳$ به این مقدار می توان تقریبی از مقدار ویژه غالب ماتریس $A$ به صورت $\lambda_۱^A \approx ۲۲/۵۲۴۴ + ۱۳/۱۳۰۱۳ = ۳۵/۸۲۵۷$ که تا ۴ رقم بعد از اعشار با مقدار دقیق برابری می کند را بدست آورد.
توجه کنید که در اینجا تعداد تکرارهای لازم کمتر از ۴۵ می باشد.

## ۶ مقادیر ویژه مختلط

چون مقادیر ویژه یک ماتریس ریشه های چند جمله ای مشخصه هستند، پس اگر $\lambda_۱$ ریشه مختلط چند جمله ای ویژه ماتریس حقیقی $A$ باشد آنگاه $\bar{\lambda_۱}$ نیز ریشه آن می باشد. بنابراین اگر مقدار ویژه غالب یک ماتریس حقیقی مختلط باشد آنگاه این مقدار ویژه یکتا نخواهد بود زیرا از آنجایی که $\bar{\lambda_۱}$ نیز مقدار ویژه است پس $|\lambda_۱| = |\bar{\lambda_۱}|$ بنابراین می توان گفت که $\lambda_۲ = \bar{\lambda_۱}$ نیز مقدار ویژه غالب $A$ خواهد بود. در این حالت می توان نشان داد که روش توانی همگرا نمی گردد.
برای مثال ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} ۵ & ۵ & -۳ \\ -۳ & ۰ & ۰ \\ ۵ & ۴ & ۵ \end{bmatrix}$$

که دارای مقادیر ویژه زیر است

$$\lambda_۱ \approx ۳/۲۹۳۲ + ۴/۶۵۵۳i, \qquad \lambda_۲ \approx ۳/۲۹۳۲ - ۴/۶۵۵۳i, \qquad \lambda_۳ \approx ۳/۴۱۳۷$$

واضح است که

$$|\lambda_۱| = |\lambda_۲| > |\lambda_۳|$$

اکنون اگر روش توانی مقیاس شده را بر $A$ با $V^{(۰)} = [۰,\ ۰,\ ۱]^T$ نتایج زیر را به دست می آوریم

| $k$ | $\lambda^{(k)}$ | $k$ | $\lambda^{(k)}$ |
|---|---|---|---|
| ۱ | ۵/۰۰۰۰ | ۵ | ۵/۹۲۹۱ |
| ۲ | ۶/۰۰۰۰ | ۶ | ۶/۵۰۲۰ |
| ۳ | ۴/۵۰۰۰ | ۷ | ۵/۳۲۳۷ |
| ۴ | ۴/۷۰۳۷ | ۸ | ۴/۵۱۰۸ |

نتایج جدول حاکی از این است که دنباله $\lambda^{(k)}$ رفتار همگرایی از خود نشان نمی دهد حتی اگر تعداد تکرار ها را افزایش دهیم نیز نتیجه تغییری نخواهد کرد. بنابراین روش توانی به این شکلی که بیان کردیم قادر نیست مقادیر ویژه مختلط $A$ را بیابد. می توان به شکلی ساختار روش توانی را تغییر داد که بتواند مقادیر ویژه مختلط $A$ (که مقادیر ویژه غالب نیز هستند) را پیدا کند.

**توجه ۵.۱۳**

علاقمندان برای دیدن محاسبه مقادیر ویژه مختلط با روش توانی به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

## ۷ محک توقف در محاسبه‌ی مقادیر ویژه

سوالی که در محاسبه‌ی مقادیر ویژه به وسیله‌ی یک روش تکراری مطرح می‌شود این است که تکرارهای روش می‌بایست تا کجا ادامه یابند؟!

فرض کنید $\lambda^{(k)}$ تقریب به دست آمده برای مقادیر ویژه‌ی غالب ماتریس $A$ توسط روش توانی باشد آنگاه می‌توان از هر یک از محک‌های توقف زیر استفاده کرد:

$$\left|\lambda^{(k+1)}-\lambda^{(k)}\right| \leqslant \epsilon \qquad (۲۴)$$

یا

$$\frac{\left|\lambda^{(k+1)}-\lambda^{(k)}\right|}{\left|\lambda^{(k+1)}\right|} \leqslant \epsilon \qquad (۲۵)$$

که $\epsilon > 0$ از قبل داده شده است. البته توجه کنید که چون در روش توانی بردار ویژه غالب نیز محاسبه می‌شود می‌توان از تقریب $V^{(k)}$ به صورت زیر استفاده کرد:

$$\left\|V^{(k+1)}-V^{(k)}\right\| \leqslant \epsilon \qquad (۲۶)$$

یا

$$\frac{\left\|V^{(k+1)} - V^{(k)}\right\|}{\left\|V^{(k+1)}\right\|} \leqslant \epsilon \qquad (۲۷)$$

که نرم در بالا یک نرم برداری است.

## ۱.۷ محک توقف باقی مانده

بنا به تعریف مقادیر ویژه و بردار ویژه می‌توان از محک توقف باقی مانده زیر نیز استفاده نمود

$$\left\|AV^{(k)} - \lambda^{(k)}V^{(k)}\right\| \leqslant \epsilon \qquad (۲۸)$$

توجه کنید که برای روش توانی مقیاس شده از بردارهای نرمال شده $\widetilde{V}^{(k)}$ استفاده می‌شود.

**مثال ۵.۲۲**

مقدار ویژه‌ی غالب ماتریس

$$A = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۳ & ۱ & ۲ \\ ۲ & ۳ & ۱ \end{bmatrix}, \qquad \lambda_{\max} = ۶$$

را با روش توانی (مقیاس شده) به قسمی محاسبه کنید که دو تقریب متوالی $\lambda^{(k)}, \lambda^{(k+1)}$ در شرط توقف

$$|\lambda^{(k+1)} - \lambda^{(k)}| \leqslant ۵ \times ۱۰^{-۴}$$

صدق نمایند.

**حل:** مقدار ویژه غالب $A$ برابر ۶ است. حال با $V^{(۰)} = [۱, \ ۰, \ ۰]^T$ شروع می‌کنیم (توجه کنید که محاسبات در نرم افزار متلب انجام شده است اما تا ۴ رقم بعد از اعشار نمایش داده شده است):

$$V^{(۱)} = AV^{(۰)} = [۱, \ ۳, \ ۲]^T, \ \lambda^{(۱)} = ۳$$
$$\widetilde{V}^{(۱)} = \frac{V^{(۱)}}{\|V^{(۱)}\|_\infty} = [۰/۳۳۳۳, \ ۱, \ ۰/۶۶۶۷]^T$$
$$V^{(۲)} = A\widetilde{V}^{(۱)} = [۴/۳۳۳۳, \ ۳/۳۳۳۳, \ ۴/۳۳۳۳]^T, \ \lambda^{(۲)} = ۴/۳۳۳۳$$
$$\widetilde{V}^{(۲)} = \frac{V^{(۲)}}{\|V^{(۲)}\|_\infty} = [۱/۰۰۰۰, \ ۰/۷۶۹۲, \ ۱/۰۰۰۰]^T$$

در دو تکرار اول داریم $\left|\lambda^{(۲)} - \lambda^{(۱)}\right| = |۴/۳۳۳۳ - ۳| = ۱/۳۳۳۳$ که هنوز شرط توقف برقرار نشده است. در تکرار سوم داریم:

$$V^{(۳)} = A\widetilde{V}^{(۲)} = [۵/۵۳۸۵, \ ۵/۷۶۹۲, \ ۵/۳۰۷۷]^T, \ \lambda^{(۳)} = ۵/۷۶۹۲$$
$$\widetilde{V}^{(۳)} = \frac{V^{(۳)}}{\|V^{(۳)}\|_\infty} = [۰/۹۶۰۰, \ ۱/۰۰۰۰, \ ۰/۹۲۰۰]^T$$

بنابراین

$$\left|\lambda^{(3)} - \lambda^{(2)}\right| = |5/7692 - 4/3333| = 1/4359$$

لذا می‌بایست تکرارها را ادامه دهیم. اگر محاسبات فوق را ادامه دهیم در تکرار نهم داریم:

$$V^{(9)} = [5/9994,\ \ 5/9993,\ \ 5/9996]^T, \qquad \lambda^{(9)} = 5/9996$$
$$\widetilde{V}^{(9)} = [1/0000,\ \ 1/0000,\ \ 1/0000]^T$$

از این‌رو

$$\left|\lambda^{(9)} - \lambda^{(8)}\right| = |5/9996 - 5/9988| = 7/2309 \times 10^{-4}$$

مشاهده می‌شود که هنوز شرط $\left|\lambda^{(k+1)} - \lambda^{(k)}\right| \leq 5 \times 10^{-4}$ حاصل نشده است.
در تکرار دهم به دست می‌اوریم:

$$V^{(10)} = [5/9999,\ \ 5/9999,\ \ 5/9998]^T, \qquad \lambda^{(10)} = 5/9999$$
$$\widetilde{V}^{(10)} = [1/0000,\ \ 1/0000,\ \ 1/0000]^T$$

با توجه به اینکه

$$\left|\lambda^{(10)} - \lambda^{(9)}\right| = |5/9999 - 5/9996| = 3/1343 \times 10^{-4} \leq 5 \times 10^{-4}$$

که نشان می‌دهد محک توقف داده شده برقرار است. بنابراین ۱۰ تکرار روش توانی مقیاس شده نیاز است و به ازای این ۱۰ تکرار مقدار ویژه‌ی غالب به صورت $\lambda^{(10)} = 5/9999$ حاصل می‌شود.

## ۸ روش معکوس توانی

دیدیم که از روش توانی برای بزرگترین مقدار ویژه ی یک ماتریس بهره جستیم حال فرض کنید هدف محاسبه‌ی کوچک‌ترین مقدار ویژه‌ی ماتریس $A$ است. آیا می‌توان روش توانی را به شکلی تغییر داد که بتواند کوچک‌ترین مقدار ویژه‌ی $A$ را تقریب نماید؟
درواقع با تغییری کوچک در روش توانی می‌توان کوچک‌ترین مقدار ویژه‌ی $A$ را یافت و اینکار بر این حقیقت که مقادیر ویژه ماتریس های $A$ و $A^{-1}$ معکوس یکدیگرند استوار است.
اگر مقادیر ویژه‌ی $A$ در شرط

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_{n-1}| > |\lambda_n| > 0$$

صدق نمایند (یعنی کوچک‌ترین مقدار ویژه یکتا باشد) آنگاه مقادیر ویژه‌ی $A^{-1}$ به صورت

$$\frac{1}{|\lambda_n|} > \frac{1}{|\lambda_{n-1}|} \geqslant \frac{1}{|\lambda_{n-2}|} \geqslant \cdots \geqslant \frac{1}{|\lambda_2|} \geqslant \frac{1}{|\lambda_1|}$$

خواهند بود. بنابراین $\frac{1}{\lambda_n}$ مقدار ویژه‌ی غالب $A^{-1}$ خواهد بود. لذا اگر روش توانی بر $A^{-1}$ اعمال گردد می‌توانیم $\frac{1}{\lambda_n}$ و از آن $\lambda_n$ را به دست آوریم. برای اینکه روش توانی بر $A^{-1}$ اعمال گردد به طور مشابه باید تعریف کنیم

$$V^{(۱)} = A^{-۱}V^{(۰)}$$
$$V^{(۲)} = A^{-۱}V^{(۱)}$$
$$\vdots$$

که $V^{(۰)}$ برداری دلخواه است. دقت کنید که برای محاسبه‌ی $V^{(۱)},\ \ V^{(۲)},\ \ \cdots$ در عمل لازم نیست $A^{-۱}$ به طور صریح محاسبه شود بلکه می‌توان نوشت

$$AV^{(۱)} = V^{(۰)}$$
$$AV^{(۲)} = V^{(۱)}$$
$$\vdots$$

این نشان می‌دهد که در روش توانی معکوس باید مکرراً دستگاه‌هایی خطی با ماتریس ضرایب $A$ و بردارهای سمت راست متغیر حل گردند تا بردارهای $V^{(۱)},\ \ V^{(۲)},\ \ \cdots$ به دست آیند. بنابراین با توجه به ثابت بودن ماتریس ضرایب دستگاه‌های فوق توصیه می‌گردد از روش‌های تجزیه مثل تجزیه $LU$ برای حل دستگاه‌ها استفاده شود.

باتوجه به اینکه در روش توانی معکوس نیز ممکن است درایه‌های بردارهای $V^{(۱)},\ \ V^{(۲)},\ \ \cdots$ از اعداد بزرگ یا کوچکی تشکیل شوند می توان از نسخه‌ی مقیاس شده‌ی روش توانی استفاده کرد. بنابراین با داشتن بردار $V^{(۰)}$ دلخواه می‌توان روش معکوس توانی مقیاس شده را به شکل زیر بیان نمود:

$$\begin{aligned}
k = ۱ \qquad & AV^{(۱)} = V^{(۰)} \\
& \widetilde{V}^{(۱)} = \frac{V^{(۱)}}{\|V^{(۱)}\|_\infty} \\
k = ۲ \qquad & AV^{(۲)} = \widetilde{V}^{(۱)} \\
& \widetilde{V}^{(۲)} = \frac{V^{(۲)}}{\|V^{(۲)}\|_\infty} \\
& \quad \vdots
\end{aligned}$$

**مثال ۵.۲۳**

با استفاده از روش معکوس توانی مقیاس شده، کوچک‌ترین مقدار ویژه‌ی ماتریس داده شده را تقریب نمایید.

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

**حل:** مقدار دقیق مقادیر ویژه $A$ تا ۴ رقم بعد از اعشار برابرند با:

$$\lambda_۱ \approx ۶/۷۱۲۴, \qquad \lambda_۲ \approx -۵/۶۱۷۲, \qquad \lambda_۳ \approx -۲/۰۹۵۲$$

لذا کوچک‌ترین مقدار ویژهِ $A$ برابر $\lambda_۳ \approx -۲/۰۹۵۲$ می‌باشد. بعلاوه می توان دید که شرط مقدار ویژه

$$|\lambda_۱| \geq |\lambda_۲| > |\lambda_۳|$$

نیز برای $A$ برقرار است. با انتخاب $V^{(0)} = [1, \ 0, \ 0]^T$ داریم:

$$AV^{(1)} = V^{(0)} \rightarrow \begin{bmatrix} -4 & -3 & -4 \\ -2 & -2 & 5 \\ -3 & 0 & 5 \end{bmatrix} V^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \rightarrow V^{(1)} = \begin{bmatrix} -0/1266 \\ -0/0633 \\ -0/0759 \end{bmatrix}$$

حال با نرمال کردن $V^{(1)}$ داریم:

$$\widetilde{V}^{(1)} = \frac{V^{(1)}}{\|V^{(1)}\|_\infty} = [-1/0000, \ \ -0/5000, \ \ -0/6000]^T$$

اکنون با حل دستگاه $AV^{(2)} = \widetilde{V}^{(1)}$ می‌توان $V^{(2)}$ را یافت:

$$\begin{bmatrix} -4 & -3 & -4 \\ -2 & -2 & 5 \\ -3 & 0 & 5 \end{bmatrix} V^{(2)} = \begin{bmatrix} -1/0000 \\ -0/5000 \\ -0/6000 \end{bmatrix} \rightarrow V^{(2)} = \begin{bmatrix} 0/2063 \\ 0/0532 \\ 0/0038 \end{bmatrix}$$

و با نرمال کردن $V^{(2)}$ داریم:

$$\widetilde{V}^{(2)} = \frac{V^{(2)}}{\|V^{(2)}\|_\infty} = [1/0000, \ \ 0/2577, \ \ 0/0184]^T$$

با ادامهِ این روند داریم :

$$V^{(10)} = \begin{bmatrix} -0/3969 \\ 0/4769 \\ -0/1680 \end{bmatrix}, \quad \widetilde{V}^{(10)} = \begin{bmatrix} -0/8322 \\ 1/0000 \\ -0/3522 \end{bmatrix}$$
$$V^{(11)} = \begin{bmatrix} 0/3978 \\ -0/4772 \\ 0/1682 \end{bmatrix}, \quad \widetilde{V}^{(11)} = \begin{bmatrix} 0/8335 \\ -1/0000 \\ 0/3525 \end{bmatrix}$$

با توجه به اینکه علامت درایه‌های دو بردار متوالی $V^{(10)}, V^{(11)}$ مختلف است پس باید مقدار ویژه‌ی غالب $A^{-1}$ منفی باشد لذا می‌توان نوشت:

$$\lambda^{(10)} = -\left\|V^{(10)}\right\|_\infty = -0/4769$$
$$\lambda^{(11)} = -\left\|V^{(11)}\right\|_\infty = -0/4772$$

بعلاوه اگر تکرار بعدی را نیز محاسبه کنیم داریم:

$$V^{(12)} = [-0/3980, \ \ 0/4772, \ \ -0/1683]^T$$
$$\widetilde{V}^{(12)} = [-0/8340, \ \ 1/0000, \ \ -0/3527]^T$$

و از این‌جا $\lambda^{(12)} = -\|V^{(12)}\|_\infty = -0/4772$ پس $-0/4772$ می‌بایست مقدار ویژه‌ی $A^{-1}$ باشد. بنابراین کوچک‌ترین مقدار ویژه‌ی $A$ به صورت

$$\frac{1}{-0/4772} \approx -2/0954$$

به دست می‌آید که تقریب قابل قبولی از مقدار دقیق $\lambda_3 \approx -2/0952$ می‌باشد.

**توجه ۵.۱۴**

همان‌طور که دیدید روش توانی معکوس چیزی نیست جز اعمال روش توانی بر $A^{-1}$. از این‌رو اگر برای یک مسأله روش توانی معکوس کند باشد همچنان می‌توان از ایده‌های

۱. الف. روش توان ماتریسی

۲. ب. روش توانی با انتقال

برای افزایش سرعت روش توانی معکوس به طور مشابه استفاده کرد.

## ۱.۸ یافتن نزدیک‌ترین مقدار ویژه به مقدار داده شده مشخص

فرض کنید اسکالر $\beta$ داده شده است. می‌خواهیم از بین $n$ مقدار ویژه‌ی $A$، آن مقدار ویژه‌ای که به $\beta$ نزدیک است را بیابیم. برای مثال فرض کنید $A$ ماتریسی $۴ \times ۴$ با مقادیر ویژه‌ی

$$\lambda_۱ = -۳, \qquad \lambda_۲ = ۲, \qquad \lambda_۳ = ۱, \qquad \lambda_۴ = ۰/۵$$

است و $\beta = ۴$ داده شده است. با توجه به اینکه

$$d_۱ = |\lambda_۱ - \beta| = |-۳ - ۴| = ۷$$
$$d_۲ = |\lambda_۲ - \beta| = |۲ - ۴| = ۲$$
$$d_۳ = |\lambda_۳ - \beta| = |۱ - ۴| = ۳$$
$$d_۴ = |\lambda_۴ - \beta| = |۰/۵ - ۴| = ۳/۵$$

بنابراین از بین این مقادیر ویژه، $\lambda_۲$ نزدیک‌ترین به $\beta = ۴$ می‌باشد. بنابراین جواب مسأله گفته شده $\lambda_۲ = ۲$ خواهد بود. دقت کنید در صورتی که مقادیر ویژه‌ی $A$ مختلطی باشند منظور از فاصله در این‌جا اندازه(مدول) خواهد بود(توجه کنید که برای عدد مختلط $z = x + iy$، اندازه‌ی $z$ به صورت $|z| = \sqrt{x^۲ + y^۲}$ می‌باشد.)

در ادامه نشان می دهیم که روش معکوس توانی قادر به حل مسأله یاد شده می‌باشد. فرض کنید $\lambda$ مقدار ویژه‌ای از $A$ باشد آنگاه مقادیر ویژه‌ی $B = A - \beta I$ به صورت $\lambda - \beta$ خواهند بود. اکنون کوچکترین مقدار ویژه ماتریس $B$ به رسیدن به این پاسخ کمک می کند. اگر روش معکوس توانی بر $B$ اعمال شود آنگاه کوچک‌ترین مقدار ویژه‌ی $B$ از لحاظ اندازه پیدا می‌شود. این یعنی $\lambda$ای به دست می‌آید که $\lambda - \beta$ کمینه باشد و این همان چیزی است که به دنبال آن هستیم. توجه کنید که اگر $\mu$ کوچک‌ترین مقدار ویژه‌ی $B$ باشد یعنی $\mu = \lambda - \beta$ آنگاه $\lambda = \mu + \beta$ جواب مسأله‌ی فوق خواهد بود.

**مثال ۵.۲۴**

مقدار ویژه‌ای از $A$ بیابید که با $\beta = ۱$ کمترین فاصله را دارد.

$$A = \begin{bmatrix} ۱ & ۲ & ۱ \\ ۰ & -۴ & -۴ \\ -۴ & -۴ & ۴ \end{bmatrix}$$

**حل:** ابتدا ماتریس $B = A - \beta I = A - I$ را تشکیل می‌دهیم

$$B = \begin{bmatrix} ۰ & ۲ & ۱ \\ ۰ & -۵ & -۴ \\ -۴ & -۴ & ۳ \end{bmatrix}$$

با اعمال روش معکوس توانی بر $B$ با بردار شروع اولیه $V^{(۰)} = [۰, \quad ۱, \quad ۱]^T$ داریم:

| $k$ | $\lambda^{(k)}$ | $\mu = \frac{۱}{\lambda^{(k)}}$ | علامت اولین درایه‌ی $V^{(k)}$ |
|---|---|---|---|
| ۱ | ۱/۰۸۳۳ | ۰/۹۲۳۱ | $-$ |
| ۲ | ۲/۴۸۰۸ | ۰/۴۰۳۱ | $+$ |
| ۳ | ۲/۳۱۷۲ | ۰/۴۳۱۶ | $-$ |
| ۴ | ۲/۳۰۷۴ | ۰/۴۳۳۴ | $+$ |
| ۵ | ۲/۳۰۶۶ | ۰/۴۳۳۵ | $-$ |
| ۶ | ۲/۳۰۶۶ | ۰/۴۳۳۵ | $+$ |
| ۷ | ۲/۳۰۶۶ | ۰/۴۳۳۵ | $-$ |

نتایج جدول فوق نشان می‌دهد که $|\mu| = |۰/۴۳۳۵|$ کوچک‌ترین مقدار ویژه‌ی $B$ از لحاظ اندازه است. بعلاوه می‌توان دید که علامت این مقدار ویژه نیز منفی است زیرا علامت درایه‌های دو بردار متوالی $V^{(k)}, V^{(k+۱)}$ مخالف هم می‌باشند(توجه کنید در اینجا از علامت درایه ی اول بردار $V^{(k)}$ استفاده شده است)به ستون آخر جدول فوق دقت کنید. بنابراین $\mu$ منفی است لذا:

$\mu = -۰/۴۳۳۵$ و از آنجا نزدیک‌ترین مقدار ویژه‌ی $A$ به ازای $\beta = ۱$ به صورت زیر به دست می‌آید:

$$\lambda = \mu + \beta = ۱ - ۰/۴۳۳۵ = ۰/۵۶۶۵$$

نتیجه‌ی به دست آمده صحیح است زیرا مقادیر ویژه‌ی $A$ در حقیقت چنین‌اند:

$$\lambda_۱ = ۵/۵۳۵۸, \quad \lambda_۲ = -۵/۱۰۲۳, \quad \lambda_۳ = ۰/۵۶۶۵$$

پس

$$d_۱ = |\lambda_۱ - \beta| = |۵/۵۳۵۸ - ۱| = ۴/۵۳۵۸$$
$$d_۲ = |\lambda_۲ - \beta| = |-۵/۱۰۲۳ - ۱| = ۶/۱۰۲۳$$
$$d_۳ = |\lambda_۳ - \beta| = |۰/۵۶۶۵ - ۱| = ۰/۴۳۳۵$$

بنابراین $\lambda_۳ = ۰/۵۶۶۵$ با $\beta = ۱$ کمترین فاصله را دارد. این همان نتیجه‌ای است که با روش معکوس توانی به دست آوردیم.

**توجه ۵.۱۵**

در صورتی که $\beta$ به اندازه ی کافی به یک مقدار ویژه $A$ چون $\lambda$ نزدیک باشد آنگاه نتیجه ی روش معکوس توانی به صورت گفته شده فوق جفت ویژه ای چون $(\lambda, V)$ خواهد بود. یعنی در این حالت می توان بردار ویژه ی متناظر با $\beta$ (که تقریبی خوبی از یک مقدار ویژه $\lambda$ است) را نیز به دست آورد.

**توجه ۵.۱۶**

تمامی جفت ویژه های یک ماتریس را می توان به کمک الگوریتمی که تقلیل توانی نام دارد محاسبه کرد. علاقمندان به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**توجه ۵.۱۷**

یک روش سریع برای محاسبه جفت ویژه روش خارج قسمت ریلی است. علاقمندان برای دیدن این روش به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

# ۹ روش‌های تجزیه جهت محاسبه جفت ویژه‌های یک ماتریس

## ۱.۹ روش تجزیه LU

فرض کنید هدف محاسبه‌ی همه‌ی مقادیر ویژه $A$ است. اگر بتوان $A$ را به صورت $A = LU$ تجزیه کرد که $L$ ماتریسی پایین مثلثی واحد و $U$ ماتریسی بالا مثلثی (در واقع تجزیه دولیتل) باشد آنگاه می‌توان طبق یک روند تکراری به محاسبه مقادیر ویژه $A$ پرداخت. ابتدا فرض کنید $A_1 = A$ و $A_1$ را به صورت $A_1 = L_1U_1$ تجزیه می کنیم. سپس می‌توان نوشت :

$$L_1^{-1}A_1L_1 = L_1^{-1}(L_1U_1)L_1 = (L_1^{-1}L_1)U_1L_1 = IU_1L_1 = U_1L_1$$

پس نتیجه می‌شود که $A_1$ با $U_1L_1$ متشابه است

$$A_1 \sim U_1L_1 \tag{۲۹}$$

حال تعریف می کنیم $A_2 = U_1L_1$. پس داریم :

$$A_1 \sim U_1L_1 = A_2 \Longrightarrow A_1 \sim A_2 \tag{۳۰}$$

اکنون اگر بتوان $A_2$ را به صورت $A_2 = L_2U_2$ تجزیه کرد مطابق (۲۹) خواهیم داشت

$$A_2 \sim U_2L_2 \tag{۳۱}$$

و با تعریف $A_3 = U_2L_2$ از (۳۰) و (۳۱) در می‌یابیم که

$$A_1 \sim A_2 \sim A_3 = U_2L_2$$

با ادامه این روند دنباله‌ی ماتریسی $\{A_k\}_{k=1}^{\infty}$ را می‌توان ساخت که

$$A = A_1 \sim A_2 \sim A_3 \sim \cdots \sim A_k \sim A_{k+1} \sim \dots \tag{۳۲}$$

از طرفی می‌دانیم که ماتریس‌های متشابه دارای مقادیر ویژه یکسانند، لذا از (۳۲) می‌توان نوشت :

$$\lambda(A) = \lambda(A_1) = \lambda(A_2) = \lambda(A_3) = \cdots = \lambda(A_k) = \lambda(A_{k+1}) = \dots$$

در ادامه نشان خواهیم داد که تحت شرایط خاصی برای مقادیر ویژه $A$ دنباله‌ی تولید شده توسط روش تجزیه LU به یک ماتریس بالا مثلثی $R$ همگراست :

$$\lim_{k\to\infty} A_k = R$$

پس می‌توان نتیجه گرفته که $\lambda(A) = \lambda(R)$ و ازآنجاییکه مقادیر ویژه ماتریس بالا مثلثی، عناصر روی قطر اصلی آن هستند بنابراین به راحتی می‌توان مقادیر ویژه $A$ را یافت.
قبل از ادامه‌ی بحث ابتدا با یک مثال از روش تجزیه توجه کنید

**مثال ۵.۲۵**

با استفاده از روش تجزیه LU ، مقادیر ویژه‌ی ماتریس $A = \begin{bmatrix} ۲ & -۳ \\ -۵ & ۱ \end{bmatrix}$ را بیابید.

**حل:** ابتدا تجزیه LU را برای $A_۱ = A$ محاسبه می‌کنیم :

$$A_۱ = L_۱U_۱ = \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ -۲/۵۰۰۰ & ۱/۰۰۰۰ \end{bmatrix} \begin{bmatrix} ۲/۰۰۰۰ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۶/۵۰۰۰ \end{bmatrix}$$

حال قرار می‌دهیم $A_۲ = U_۱L_۱$ داریم :

$$A_۲ = \begin{bmatrix} ۲/۰۰۰۰ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۶/۵۰۰۰ \end{bmatrix} \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ -۲/۵۰۰۰ & ۱/۰۰۰۰ \end{bmatrix} = \begin{bmatrix} ۹/۵۰۰۰ & -۳/۰۰۰۰ \\ ۱۶/۲۵۰۰ & -۶/۵۰۰۰ \end{bmatrix}$$

اکنون می‌بایست تجزیه LU را برای $A_۲$ محاسبه کنیم :

$$A_۲ = L_۲U_۲ = \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ ۱/۷۱۰۵ & ۱/۰۰۰۰ \end{bmatrix} \begin{bmatrix} ۹/۵۰۰۰ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۱/۳۶۸۴ \end{bmatrix}$$

اکنون قرار می‌دهیم $A_۳ = U_۲L_۲$ داریم :

$$A_۳ = \begin{bmatrix} ۹/۵۰۰۰ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۱/۳۶۸۴ \end{bmatrix} \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ ۱/۷۱۰۵ & ۱/۰۰۰۰ \end{bmatrix} = \begin{bmatrix} ۴/۳۶۸۴ & -۳/۰۰۰۰ \\ -۲/۳۴۰۷ & -۱/۳۶۸۴ \end{bmatrix}$$

با محاسبه تجزیه LU ماتریس $A_۳$ داریم :

$$A_۳ = L_۳U_۳ = \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ -۰/۵۳۵۸ & ۱/۰۰۰۰ \end{bmatrix} \begin{bmatrix} ۴/۳۶۸۴ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۹۷۵۹ \end{bmatrix}$$

حال قرار می‌دهیم $A_۴ = U_۳L_۳$ داریم :

$$A_۴ = \begin{bmatrix} ۴/۳۶۸۴ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۹۷۵۹ \end{bmatrix} \begin{bmatrix} ۱/۰۰۰۰ & ۰/۰۰۰۰ \\ -۰/۵۳۵۸ & ۱/۰۰۰۰ \end{bmatrix} = \begin{bmatrix} ۵/۹۷۵۹ & -۳/۰۰۰۰ \\ ۱/۵۹۴۶ & -۲/۹۷۵۹ \end{bmatrix}$$

با ادامه این روند داریم :

$$A_۵ = \begin{bmatrix} ۵/۱۷۵۴ & -۳/۰۰۰۰ \\ -۰/۵۸۰۵ & -۲/۱۷۵۴ \end{bmatrix}$$

$$A_۶ = \begin{bmatrix} ۵/۵۱۱۹ & -۳/۰۰۰۰ \\ ۰/۲۸۱۷ & -۲/۵۱۱۹ \end{bmatrix}$$

$$A_۷ = \begin{bmatrix} ۵/۳۵۸۵ & -۳/۰۰۰۰ \\ -۰/۱۲۰۶ & -۲/۳۵۸۵ \end{bmatrix}$$

$$A_۸ = \begin{bmatrix} ۵/۴۲۶۰ & -۳/۰۰۰۰ \\ ۰/۰۵۴۶ & -۲/۴۲۶۰ \end{bmatrix}$$

$$A_{۹} = \begin{bmatrix} ۵/۳۹۵۹ & -۳/۰۰۰۰ \\ -۰/۰۲۴۱ & -۲/۳۹۵۹ \end{bmatrix}$$

$$A_{۱۰} = \begin{bmatrix} ۵/۴۰۹۳ & -۳/۰۰۰۰ \\ ۰/۰۱۰۸ & -۲/۴۰۹۳ \end{bmatrix}$$

$$A_{۱۱} = \begin{bmatrix} ۵/۴۰۳۳ & -۳/۰۰۰۰ \\ -۰/۰۰۴۸ & -۲/۴۰۳۳ \end{bmatrix}$$

$$A_{۱۲} = \begin{bmatrix} ۵/۴۰۵۹ & -۳/۰۰۰۰ \\ ۰/۰۰۲۱ & -۲/۴۰۵۹ \end{bmatrix}$$

$$A_{۱۳} = \begin{bmatrix} ۵/۴۰۴۸ & -۳/۰۰۰۰ \\ -۰/۰۰۰۹ & -۲/۴۰۴۸ \end{bmatrix}$$

$$A_{۱۴} = \begin{bmatrix} ۵/۴۰۵۳ & -۳/۰۰۰۰ \\ ۰/۰۰۰۴ & -۲/۴۰۵۳ \end{bmatrix}$$

$$A_{۱۵} = \begin{bmatrix} ۵/۴۰۵۲ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۴۰۵۲ \end{bmatrix}$$

$$A_{۱۶} = \begin{bmatrix} ۵/۴۰۵۱ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۴۰۵۱ \end{bmatrix}$$

$$A_{۱۷} = \begin{bmatrix} ۵/۴۰۵۱ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۴۰۵۱ \end{bmatrix}$$

از آنجاییکه $A_{۱۷}$ تقریباً بالا مثلثی شده است پس تکرارها را در همین جا متوقف کرده و چون $A \sim A_{۱۷}$ بنابراین :

$$\lambda(A) = \lambda(A_{۱۷}) = \lambda\left(\begin{bmatrix} ۵/۴۰۵۱ & -۳/۰۰۰۰ \\ ۰/۰۰۰۰ & -۲/۴۰۵۱ \end{bmatrix}\right) = \{۵/۴۰۵۱,\ -۲/۴۰۵۱\}$$

**تمرین ۵.۶**

با استفاده از روش تجزیه LU مقادیر ویژه ماتریس $A = \begin{bmatrix} -۱ & ۵ \\ ۶ & ۷ \end{bmatrix}$ را بیابید.
توجه کنید که مقادیر ویژه دقیق ماتریس عبارتند از :

$$\lambda_{۱} = ۹/۷۸۲۳, \quad \lambda_{۲} = -۳/۷۸۲۳$$

## ۲.۹ نحوه محاسبه‌ی بردار ویژه در روش تجزیه $LU$

همانطور که دیدیم در روش تجزیه $LU$ با استفاده از دنباله‌ی ماتریسی

$$A = A_{۱},\ A_{۲},\ A_{۳}, \ldots, A_k, \ldots$$

توانستیم یک ماتریس بالامثلثی $R$ (البته در حد) بدست آوریم که با $A$ متشابه باشد یعنی $R$ به صورت

$$R = \lim_{k\to\infty} A_k$$

وجود دارد که $\lambda(A) = \lambda(R)$. البته تحت شرایطی این دنباله ماتریسی همگرا می گردد. اکنون سوالی که مطرح می شود این است که بردارهای ویژه $A$ چگونه قابل محاسبه هستند؟

**یادآوری ۵.۱**

فرض کنید $A$ و $B$ دو ماتریس متشابه باشند یعنی ماتریس معکوس پذیری چون $P$ باشد که

$$A = PBP^{-1} \tag{۳۳}$$

می‌خواهیم ارتباط بردارهای ویژه $A$ و $B$ را بیابیم. البته می‌دانیم که مقادیر ویژه‌ی $A$ و $B$ یکسان اند. فرض کنیم $(\lambda, X)$ یک جفت ویژه‌ی $A$ باشد پس

$$AX = \lambda X \tag{۳۴}$$

اگر عبارت معادل $A$ را از (۳۳) در (۳۴) قرار دهیم به دست می‌آوریم

$$PBP^{-1}X = \lambda X$$

با ضرب طرفین رابطه فوق در $P^{-1}$ به دست می آویم

$$B(P^{-1}X) = \lambda(P^{-1}X)$$

این نشان می‌دهد که $(\lambda, P^{-1}X)$ جفت ویژه‌ی $B$ است. بنابراین:
اگر $A$ و $B$ به صورت $A = PBP^{-1}$ متشابه باشند و $X$ یک بردار ویژه‌ی $A$ باشد آنگاه $P^{-1}X$ بردار ویژه‌ی $B$ خواهد بود.

در روش تجزیه $LU$ دیدیم که $A = A_1$ با ماتریس $A_2 = U_1L_1$ متشابه است، از طرفی $A_2$ با $A_3 = U_2L_2$ متشابه است و این روند ادامه دارد. بنابر این می‌توان نوشت

$$\begin{aligned}
A_1 &= L_1U_1 \Rightarrow L_1^{-1}A_1L_1 = U_1L_1 = A_2 \Rightarrow \qquad A_1 \sim A_2\\
A_2 &= L_2U_2 \Rightarrow L_2^{-1}A_2L_2 = U_2L_2 = A_3 \Rightarrow \qquad A_2 \sim A_3\\
A_3 &= L_3U_3 \Rightarrow L_3^{-1}A_3L_3 = U_3L_3 = A_4 \Rightarrow \qquad A_3 \sim A_4\\
&\vdots\\
A_{k-1} &= L_{k-1}U_{k-1} \Rightarrow L_{k-1}^{-1}A_{k-1}L_{k-1} = U_{k-1}L_{k-1} = A_k \Rightarrow \qquad A_{k-1} \sim A_k
\end{aligned}$$

از روابط فوق به دست می‌آوریم

$$\begin{aligned}
A = A_1 &= L_1A_2L_1^{-1}\\
A_2 &= L_2A_3L_2^{-1}\\
&\vdots\\
A_{k-1} &= L_{k-1}A_kL_{k-1}^{-1}
\end{aligned}$$

بنابراین

$$\begin{aligned}
A = A_1 &= L_1A_2L_1^{-1} = L_1(L_2A_3L_2^{-1})L_1^{-1}\\
&= (L_1L_2)A_3(L_2^{-1}L_1^{-1})
\end{aligned}$$

$$\begin{aligned}
&= (L_1L_2)(L_3A_4L_3^{-1})(L_2^{-1}L_1^{-1}) \\
&= (L_1L_2L_3)A_4(L_3^{-1}L_2^{-1}L_1^{-1}) \\
&\vdots \\
&= (L_1L_2\dots L_{k-1})A_k(L_{k-1}^{-1}\dots L_2^{-1}L_1^{-1}) \\
&= (L_1L_2\dots L_{k-1})A_k(L_1L_2\dots L_{k-1})^{-1}
\end{aligned}$$

اگر انتخاب کنیم $P = L_1L_2\dots L_{k-1}$ آنگاه به دست می‌آوریم:

$$A = PA_kP^{-1} \tag{۳۵}$$

که نشان می‌دهد $A$ با $A_k$ متشابه است. طبق نکته‌ی قبل اگر $X$ بردار ویژه‌ی $A$ باشد آنگاه $P^{-1}X$ بردار ویژه‌ی $A_k$ است. اگر تعریف کنیم $Y = P^{-1}X$ آنگاه $X = PY$. لذا اگر $Y$ یک بردار ویژه‌ی $A_k$ باشد آنگاه $X = PY$ بردار ویژه‌ی $A$ خواهد بود یعنی بردار زیر

$$X = PY = L_1L_2L_3\dots L_{k-1}Y$$

به طور خلاصه داریم

فرض کنید $A_k$ ماتریسی تقریبا بالامثلثی باشد که از روش تجزیه LU به دست آمده است و $Y$ بردار ویژه ای از آن است. آنگاه یک بردار ویژه ی $A$ از $X = L_1L_2L_3\dots L_{k-1}Y$ به دست می آید که در آن $L_1, L_2, L_3, \dots, L_{k-1}$ ماتریس های پایین مثلثی حاصل در تکرارهای روش تجزیه LU می باشند.

**توجه ۵.۱۸**

با توجه به اینکه مقادیر ویژه ماتریس بالا مثلثی $A_k$ در دسترس هستند. بردار ویژه های این ماتریس را با کمک روش معکوس توانی نیز می توان محاسبه نمود.

**توجه ۵.۱۹**

چون $A_k$ تقریبا بالامثلثی است محاسبه‌ی جفت ویژه‌هایش براحتی انجام می‌پذیرد. در ادامه بحثی مختصر در ارتباط با محاسبه ی بردار ویژه های ماتریس های بالامثلثی ارایه می دهیم.

## ۳.۹ محاسبه‌ی بردارهای ویژه‌ی ماتریس‌های مثلثی

ماتریس بالامثلثی زیر را در نظر بگیرید

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \ddots & \vdots \\ & & \ddots & u_{n-1,n} \\ 0 & \dots & & u_{nn} \end{bmatrix}$$

همانطور که می‌دانیم مقادیرویژه‌ی $U$ به صورت زیر هسستند

$$\lambda_1 = u_{11}, \quad \lambda_2 = u_{22}, \quad \cdots, \quad \lambda_n = u_{nn}$$

فرض کنید که $\lambda_i$ ها متمایز باشند.
می‌خواهیم روشی برای محاسبه‌ی بردارهای ویژه‌ی ماتریس $U$ بسازیم که به راحتی بتوان آن را بکار گرفت. توجه کنید که در حالت کلی اگر مقدار ویژه‌ی $\lambda$ از ماتریسی چون $A$ را داشته باشیم برای محاسبه‌ی بردارویژه‌ی متناظر باید دستگاه همگن $(A-\lambda I)X=0$ را حل کرد که دارای مشکلات خاص خودش است.
فرض کنید $e_j$ نشان دهنده‌ی ستون j ام ماتریس همانی باشد. همچنین فرض کنید $U_j$ نشان دهنده‌ی ستون $j$ ماتریس $U$ باشد. واضح است که

$$U_j = Ue_j, \quad j=1,2,...,n$$

از طرفی تساوی

$$Ue_1 = U_1 = \begin{bmatrix} u_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = u_{11}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = u_{11}e_1$$

نشان می‌دهد که $(u_{11},e_1)$ یک جفت ویژه‌ی $U$ است. فرض کنید $V_j$ نشان دهنده‌ی بردار ویژه‌ی $U$ باشد. بنابراین یک جفت ویژه‌ی $U$ برابر است با

$$(\lambda_1, V_1) = (u_{11}, e_1)$$

می‌خواهیم دیگر بردارهای ویژه‌ی $U$ یعنی $V_2, V_3, \ldots, V_n$ را محاسبه کنیم. ابتدا فرض کنیم که $\alpha_1$ای وجود دارد به طوری که

$$V_2 = e_2 + \alpha_1 V_1 \tag{36}$$

پس باید $UV_2 = \lambda_2 V_2$ برقرار باشد:

$$UV_2 = \lambda_2 V_2 \Rightarrow UV_2 = u_{22}V_2$$
$$\Rightarrow U(e_2+\alpha_1 V_1) = u_{22}(e_2+\alpha_1 V_1)$$
$$\Rightarrow Ue_2 + \alpha_1 UV_1 = u_{22}e_2 + \alpha_1 u_{22}V_1 \tag{37}$$

از طرفی داریم

$$Ue_2 = U_2 = \begin{bmatrix} u_{12} \\ u_{22} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} u_{12} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ u_{22} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = u_{12}e_1 + u_{22}e_2 \tag{38}$$

با قرار دادن (۳۸) در (۳۷) و اینکه $UV_1 = u_{11}e_1$ و $V_1 = e_1$ داریم

$$u_{12}e_1 + u_{22}e_2 + \alpha_1 u_{11}e_1 = u_{22}e_2 + \alpha_1 u_{22}e_1$$

یا

$$(u_{12} + \alpha_1 u_{11})e_1 = \alpha_1 u_{22}e_1 \tag{39}$$

چون $e_1 = [1,0,\ldots,0]^T$ پس (۳۹) نتیجه می‌دهد.

$$u_{12} + \alpha_1 u_{11} = \alpha_1 u_{22} \Rightarrow \alpha_1 = \frac{u_{12}}{u_{22}-u_{11}}$$

بنابراین دومین بردار ویژه‌ی $U$ به صورت $V_2 = e_2 + \alpha_1 e_1$ خواهد بود که $\alpha_1$ در بالا مشخص شده است. حال به محاسبه‌ی $V_3$ می‌پردازیم. فرض کنید $\beta_1$ و $\beta_2$ ای وجود دارند که:

$$V_3 = e_3 + \beta_1 V_1 + \beta_2 V_2$$

پس باید $UV_3 = \lambda_3 V_3$ یا $UV_3 = u_{33} V_3$. پس داریم

$$U(e_3 + \beta_1 V_1 + \beta_2 V_2) = u_{33}(e_3 + \beta_1 V_1 + \beta_2 V_2)$$

یا

$$Ue_3 + \beta_1 UV_1 + \beta_2 UV_2 = u_{33} e_3 + \beta_1 u_{33} V_1 + \beta_2 u_{33} V_2$$

یا

$$Ue_3 + \beta_1 \lambda_1 V_1 + \beta_2 \lambda_2 V_2 = u_{33} e_3 + \beta_1 u_{33} V_1 + \beta_2 u_{33} V_2 \tag{40}$$

حال توجه کنید که

$$Ue_3 = U_3 = \begin{bmatrix} u_{13} \\ u_{23} \\ u_{33} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = u_{13} e_1 + u_{23} e_2 + u_{33} e_3 \tag{41}$$

از طرفی از $V_1 = e_1, \lambda_1 = u_{11}, V_2 = e_2 + \alpha_1 e_1, \lambda_2 = u_{22}$ و معادلات (۴۰) و (۴۱) خواهیم داشت.

$$\begin{aligned} &u_{13} e_1 + u_{23} e_2 + u_{33} e_3 + \beta_1 u_{11} e_1 + \beta_2 u_{22} \left(e_2 + \alpha_1 e_1\right) \\ &= u_{33} e_3 + \beta_1 u_{33} e_1 + \beta_2 u_{33} \left(e_2 + \alpha_1 e_1\right) \end{aligned} \tag{42}$$

با ساده کردن (۴۲) داریم

$$u_{13} e_1 + u_{23} e_2 + \beta_1 u_{11} e_1 + \beta_2 u_{22} e_2 + \alpha_1 \beta_2 u_{22} e_1 = \beta_1 u_{33} e_1 + \beta_2 u_{33} e_2 + \alpha_1 \beta_2 u_{33} e_1$$

یا

$$(u_{13} + \beta_1 u_{11} + \alpha_1 \beta_2 u_{22}) e_1 + (u_{23} + \beta_2 u_{22}) e_2 = (\beta_1 u_{33} + \alpha_1 \beta_2 u_{33}) e_1 + \beta_2 u_{33} e_2 \tag{43}$$

واضح است که (۴۳) وقتی برقرار است که داشته باشیم

$$u_{23} + \beta_2 u_{22} = \beta_2 u_{33} \tag{44}$$

$$u_{13} + \beta_1 u_{11} + \alpha_1 \beta_2 u_{22} = \beta_1 u_{33} + \alpha_1 \beta_2 u_{33} \tag{45}$$

حال از (۴۴) داریم

$$\beta_2 = \frac{u_{23}}{u_{33} - u_{22}}$$

همچنین از (۴۵) داریم

$$\beta_1 (u_{33} - u_{11}) = u_{13} + \alpha_1 \beta_2 u_{22} - \alpha_1 \beta_2 u_{33}$$

یا

$$\beta_۱ = \frac{۱}{u_{۳۳} - u_{۱۱}}(u_{۱۳} + \alpha_۱ \beta_۲ (u_{۲۲} - u_{۳۳}))$$
$$= \frac{۱}{u_{۳۳} - u_{۱۱}}(u_{۱۳} + \alpha_۱(-u_{۲۳})) = \frac{u_{۱۳} - \alpha_۱ u_{۲۳}}{u_{۳۳} - u_{۱۱}}$$

بنابراین

$$\beta_۱ = \frac{u_{۱۳} - \alpha_۱ u_{۲۳}}{u_{۳۳} - u_{۱۱}}$$

اکنون با ادامه ی روند فوق می توان تمامی بردارهای ویژه این ماتریس را محاسبه کنید. اینکار از شما در تمرین زیر خواسته شده است.

**تمرین ۵.۷**

به طور متشابه نشان دهید که هر بردار ویژه‌ی ماتریس $U$ را می‌توان برحسب بردار ویژه‌های قبلی به صورت زیر محاسبه نمود

$$V_k = e_k + \mu_۱ V_۱ + \mu_۲ V_۲ + \ldots + \mu_{k-۱} V_{k-۱}, \quad k = ۲, ۳, \ldots, n$$

جایی که $\mu_k$ ها مقادیری ثابت هستند که برحسب درایه‌های $U$ مشخص می‌شوند.

**تمرین ۵.۸**

همانطور که مشاهده گردید برای محاسبه‌ی بردارهای ویژه ماتریس $U$ با روش گفته شده شرط متمایز بودن عناصر روی قطر $U$ لازم است. آیا در حالتی که این شرط برقرار نباشد می‌توان روشی برای محاسبه‌ی بردارهای ویژه ارائه داد؟

**تمرین ۵.۹**

تمامی بحث‌های گفته شده را برای وقتی که هدف محاسبه‌ی بردارهای ویژه ماتریس پایین مثلثی زیر است انجام دهید.

$$L = \begin{bmatrix} l_{۱۱} & & & ۰ \\ l_{۲۱} & l_{۲۲} & & \\ \vdots & & \ddots & \\ l_{n۱} & \ldots & l_{n,n-۱} & l_{nn} \end{bmatrix}$$

**مثال ۵.۲۶**

بردارهای ویژه‌ی ماتریس بالامثلثی زیر را محاسبه کنید.

$$U = \begin{bmatrix} ۲ & -۳ & ۴ \\ ۰ & ۱ & ۵ \\ ۰ & ۰ & ۳ \end{bmatrix}$$

حل: مطابق مطالب گفته شده $V_۱ = e_۱ = [۱, ۰, ۰]^T$ یک بردار ویژه است. حال داریم

$$\alpha_۱ = \frac{u_{۱۲}}{u_{۲۲} - u_{۱۱}} = \frac{-۳}{۱ - ۲} = ۳$$

پس بردارویژه‌ی $V_2$ متناظر با $\lambda_2 = 1$ چنین است

$$V_2 = e_2 + \alpha_1 e_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

برای محاسبه‌ی $V_3$ ابتدا $\beta_1$ و $\beta_2$ را محاسبه می‌کنیم

$$\beta_1 = \frac{u_{13} - \alpha_1 u_{23}}{u_{33} - u_{11}} = \frac{4 - 3 \times 5}{3 - 2} = -11$$
$$\beta_2 = \frac{u_{23}}{u_{33} - u_{22}} = \frac{5}{3 - 1} = \frac{5}{2}$$

پس داریم

$$V_3 = e_3 + \beta_1 V_1 + \beta_2 V_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 11 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{5}{2} \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{7}{2} \\ \frac{5}{2} \\ 1 \end{bmatrix}$$

بنابراین بردارهای ویژه‌ی $U$ چنین خواهند بود:

$$V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \quad V_3 = \begin{bmatrix} -\frac{7}{2} \\ \frac{5}{2} \\ 1 \end{bmatrix}$$

**مثال ۵.۲۷**

بردارهای ویژه‌ی ماتریس داده شده را به دست آورید.

$$U = \begin{bmatrix} 4 & -1 & 0 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

حل: داریم

$$V_1 = e_1 = [1, 0, 0, 0]^T$$

برای محاسبه $V_2$ ابتدا $\alpha_1$ را محاسبه می‌کنیم

$$\alpha_1 = \frac{u_{12}}{u_{22} - u_{11}} = \frac{1}{3}$$

$$V_2 = e_2 + \alpha_1 e_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

همچنین $\beta_1$ و $\beta_2$ را به صورت زیر محاسبه می‌کنیم

$$\beta_1 = \frac{u_{13} - \alpha_1 u_{23}}{u_{33} - u_{11}} = \frac{0 - \frac{1}{3} \times 3}{2 - 4} = \frac{1}{2}$$

$$\beta_2 = \frac{u_{23}}{u_{33} - u_{22}} = \frac{3}{2 - 1} = 3$$

پس $V_3$ چنین است

$$V_3 = e_3 + \beta_1 V_1 + \beta_2 V_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} \frac{1}{3} \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ 3 \\ 1 \\ 0 \end{bmatrix}$$

برای محاسبه‌ی $V_4$ می‌توان فرض کرد که

$$V_4 = e_4 + \gamma_1 V_1 + \gamma_2 V_2 + \gamma_3 V_3$$

اکنون به کمک تمرین قبل نشان دهید که

$$\gamma_1 = -\frac{7}{3}, \ \gamma_2 = -6, \ \gamma_3 = \frac{10}{3}$$

پس

$$V_4 = e_4 - \frac{7}{3} V_1 - 6 V_2 + \frac{10}{3} V_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 1 \end{bmatrix}$$

لذا بردار ویژه های ماتریس داده شده چنین اند:

$$V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad V_2 = \begin{bmatrix} \frac{1}{3} \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad V_3 = \begin{bmatrix} \frac{3}{2} \\ 3 \\ 1 \\ 0 \end{bmatrix}, \quad V_4 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 1 \end{bmatrix}$$

اکنون آماده ایم تا با استفاده از روش تجزیه LU ، جفت ویژه های یک ماتریس را محاسبه کنیم.

**مثال ۵.۲۸**

با استفاده از روش تجزیه LU ، جفت ویژه‌ی ماتریس $A = \begin{bmatrix} -3 & 2 \\ 1 & -5 \end{bmatrix}$ را بیابید.

حل: ابتدا تجزیه LU را برای $A_1 = A$ محاسبه می‌کنیم :

$$A_1 = L_1 U_1 = \begin{bmatrix} 1/0000 & 0/0000 \\ -2/5000 & 1/0000 \end{bmatrix} \begin{bmatrix} 2/0000 & -3/0000 \\ 0/0000 & -6/5000 \end{bmatrix}$$

حال قرار می‌دهیم $A_2 = U_1L_1$ داریم :

$$A_2 = \begin{bmatrix} 2.0000 & -3.0000 \\ 0.0000 & -6.5000 \end{bmatrix} \begin{bmatrix} 1.0000 & 0.0000 \\ -2.5000 & 1.0000 \end{bmatrix} = \begin{bmatrix} 9.5000 & -3.0000 \\ 16.2500 & -6.5000 \end{bmatrix}$$

اکنون می‌بایست تجزیه LU را برای $A_2$ محاسبه کنیم :

$$A_2 = L_2U_2 = \begin{bmatrix} 1.0000 & 0.0000 \\ 1.7105 & 1.0000 \end{bmatrix} \begin{bmatrix} 9.5000 & -3.0000 \\ 0.0000 & -1.3684 \end{bmatrix}$$

اکنون قرار می‌دهیم $A_3 = U_2L_2$ داریم :

$$A_3 = \begin{bmatrix} 9.5000 & -3.0000 \\ 0.0000 & -1.3684 \end{bmatrix} \begin{bmatrix} 1.0000 & 0.0000 \\ 1.7105 & 1.0000 \end{bmatrix} = \begin{bmatrix} 4.3684 & -3.0000 \\ -2.3407 & -1.3684 \end{bmatrix}$$

با محاسبه تجزیه LU ماتریس $A_3$ داریم :

$$A_3 = L_3U_3 = \begin{bmatrix} 1.0000 & 0.0000 \\ -0.5358 & 1.0000 \end{bmatrix} \begin{bmatrix} 4.3684 & -3.0000 \\ 0.0000 & -2.9759 \end{bmatrix}$$

حال قرار می‌دهیم $A_4 = U_3L_3$ داریم :

$$A_4 = \begin{bmatrix} 4.3684 & -3.0000 \\ 0.0000 & -2.9759 \end{bmatrix} \begin{bmatrix} 1.0000 & 0.0000 \\ -0.5358 & 1.0000 \end{bmatrix} = \begin{bmatrix} 5.9759 & -3.0000 \\ 1.5946 & -2.9759 \end{bmatrix}$$

با ادامه این روند داریم :

$$A_5 = \begin{bmatrix} 5.1754 & -3.0000 \\ -0.5805 & -2.1754 \end{bmatrix}, \quad A_6 = \begin{bmatrix} 5.5119 & -3.0000 \\ 0.2817 & -2.5119 \end{bmatrix}$$

$$A_7 = \begin{bmatrix} 5.3585 & -3.0000 \\ -0.1206 & -2.3585 \end{bmatrix}, \quad A_8 = \begin{bmatrix} 5.4260 & -3.0000 \\ 0.0546 & -2.4260 \end{bmatrix}$$

$$A_9 = \begin{bmatrix} 5.3959 & -3.0000 \\ -0.0241 & -2.3959 \end{bmatrix}, \quad A_{10} = \begin{bmatrix} 5.4093 & -3.0000 \\ 0.0108 & -2.4093 \end{bmatrix}$$

$$A_{11} = \begin{bmatrix} 5.4033 & -3.0000 \\ -0.0048 & -2.4033 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5.4059 & -3.0000 \\ 0.0021 & -2.4059 \end{bmatrix}$$

$$A_{13} = \begin{bmatrix} 5.4048 & -3.0000 \\ -0.0009 & -2.4048 \end{bmatrix}, \quad A_{14} = \begin{bmatrix} 5.4053 & -3.0000 \\ 0.0004 & -2.4053 \end{bmatrix}$$

$$A_{15} = \begin{bmatrix} 5.4052 & -3.0000 \\ 0.0001 & -2.4052 \end{bmatrix}, \quad A_{16} = \begin{bmatrix} 5.4051 & -3.0000 \\ 0.0000 & -2.4051 \end{bmatrix}$$

$$A_{17} = \begin{bmatrix} 5.4051 & -3.0000 \\ 0.0000 & -2.4051 \end{bmatrix}$$

از آنجاییکه $A_{17}$ تقریباً بالا مثلثی شده است پس تکرارها را در همین جا متوقف کرده و چون $A \sim A_{17}$ بنابراین :

$$\lambda(A) = \lambda(A_{17}) = \lambda\left(\begin{bmatrix} 5.4051 & -3.0000 \\ 0.0000 & -2.4051 \end{bmatrix}\right) = \{5.4051,\ -2.4051\}$$

مطابق مطالب بیان شده برای ماتریس های بالا مثلثی بردار ویژه‌ی نظیر $\lambda_1$ ماتریس $A_{17}$ برابر است با

$$Y_1 = e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

و بردار ویژه نظیر $\lambda_2$ به صورت زیر قابل محاسبه است:

$$\alpha_1 = \frac{u_{12}}{u_{22} - u_{11}} = 0/3841$$

$$Y_2 = e_2 + \alpha_1 Y_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0/3841 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0/3841 \\ 1/0000 \end{bmatrix}$$

با محاسبه‌ی ماتریس $P$ داریم

$$P = L_1 L_2 L_3 \dots L_{k-1} = L_1 L_2 L_3 \dots L_{16}$$

$$= \begin{bmatrix} 1 & 0 \\ -2/5000 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1/7105 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -0/5358 & 1 \end{bmatrix} \times \dots \times \begin{bmatrix} 1 & 0 \\ 0/0000 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/1350 & 1 \end{bmatrix}$$

بنابراین بردار ویژه‌ی متناظر با $\lambda_1$ ماتریس $A$ برابر است با

$$X_1 = PY_1 = \begin{bmatrix} 1 & 0 \\ -1/1350 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1/1350 \end{bmatrix}$$

و بردار ویژه متناظر با $\lambda_2$ چنین است:

$$X_2 = PY_2 = \begin{bmatrix} 1 & 0 \\ -1/1350 & 1 \end{bmatrix} \begin{bmatrix} 0/3841 \\ 1/0000 \end{bmatrix} = \begin{bmatrix} 0/3841 \\ 0/5640 \end{bmatrix}$$

توجه کنید که دقت جفت ویژه‌های $(\lambda_1, X_1)$ و $(\lambda_2, X_2)$ را می‌توان به صورت زیر محاسبه کرد:

$$||AX_1 - \lambda_1 X_1||_\infty = 3/0258 \times 10^{-5}$$

$$||AX_2 - \lambda_2 X_2||_\infty = 8/3867 \times 10^{-6}$$

توجه شود که از محک توقف باقی مانده برای کنترل دقت جواب استفاده کرده ایم. بنظر شما برای بررسی دقت جفت ویژه ها از چه محک توقف دیگری می توان استفاده نمود؟

## ۴.۹ همگرایی روش تجزیه $LU$

**قضیه ۵.۹**

فرض کنید $(\lambda_i, X_i)$ ، که $i = 1, 2, \dots, n$ جفت ویژه‌های ماتریس $A$ باشند به قسمی که

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0 \tag{46}$$

بعلاوه فرض کنید $X = [X_1, X_2, \dots, X_n]$ ماتریس حاصل از بردارهای ویژه باشد. اگر تجزیه $LU$ ماتریس‌های $X, X^{-1}$ موجود باشد آنگاه دنباله $\{A_k\}_{k=1}^{\infty}$ به یک ماتریس بالامثلثی همگراست:

$$R = \lim_{k\to\infty} A_k$$

( $R$ یک ماتریس بالامثلثی است )

**توجه ۵.۲۰**

علاقمندان برای دیدن اثبات به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**توجه ۵.۲۱**

در قضیه ی فوق شرط وجود تجزیه LU برای هر دو ماتریس $X$ و $X^{-1}$ لازم است. برای مثال ماتریس $X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ تجزیه LU دارد:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

اما $X^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$ دارای تجزیه LU نیست. زیرا در صورت وجود چنین تجزیه ای می بایست داشته باشیم

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & u_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & l_{22} + u_{22} \end{bmatrix}$$

که تناقض است.

**توجه ۵.۲۲**

شرایط قضیه فوق شرایط لازم همگرایی نیستند. برای مثال ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}, \quad \lambda_1 = \lambda_2 = 1$$

واضح است که شرط $|\lambda_1| > |\lambda_2| >$ برقرار نیست، اما با محاسبه ی دنباله ی $\{A_k\}$ حاصل از روش تجزیه LU می توان دید که

$$\lim_{k \longrightarrow \infty} A_k = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

که نشان از همگرایی دنباله $\{A_k\}$ دارد.

**توجه ۵.۲۳**

علاقمندان برای دیدن روش تجزیه $LU$ با محورگیری جزئی به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**نکته ۵.۷**

در صورتی که بخواهیم تمامی جفت ویژه ها را با روش توانی (و با کمک تقلیل توانی) به دست آوریم شرط (۴۶) کافی است. بنابراین شرایط روش تجزیه LU به مراتب قوی ترند.

## ۵.۹ روش تجزیه QR

فرض کنید در الگوریتم تجزیه LU از تجزیه QR استفاده شود، آنگاه الگوریتم حاصل را الگوریتم تجزیه QR می‌نامیم. با توجه به اینکه قبلاً بحث کردیم که تجزیه QR نسبت به LU هم پایدارتر بوده و هم همواره برای هر ماتریس دلخواهی وجود دارد پس در عمل استفاده از الگوریتم تجزیه QR به الگوریتم LU ارجعیت دارد.

الگوریتم تجزیه QR به صورت خلاصه‌ی زیر است :

$$k = ۱, ۲, ۳, \dots$$

تجزیه $\qquad A_k = Q_k R_k$

تعریف $\qquad A_{k+۱} = R_k Q_k$

در استفاده از الگوریتم تجزیه QR باید به دو نکته‌ی مهم زیر توجه نمود :

۱ - آیا ماتریس $A$ با هر ماتریس $A_k$ متشابه است؟

۲ - آیا ماتریس های $A_k$ به یک ماتریس که محاسبه ی مقادیر ویژه‌اش راحت‌تر باشد همگراند ؟

در پاسخ به سوال ۱ می‌بایست گفت چون $A = A_۱ = Q_۱ R_۱$ پس $Q_۱^{-۱} A = R_۱$ لذا

$$A_۲ = R_۱ Q_۱ = (Q_۱^{-۱} A) Q_۱ = Q_۱^{-۱} A Q_۱$$

این نشان می‌دهد که $A$ با $A_۲$ متشابه است. با همین استدلال می‌توان نشان داد که $A$ با هر $A_k$ متشابه می‌باشد.

حال در پاسخ به سوال دوم قضیه زیر مطرح می‌گردد:

**قضیه ۵.۱۰**

فرض کنید $i = ۱, ۲, \dots n$ ; $(\lambda_i, X_i)$ جفت ویژه‌های ماتریس $A$ باشند و $X = [X_۱, X_۲, \dots X_n]$ اگر شرط

$$|\lambda_۱| > |\lambda_۲| > \cdots > |\lambda_n|$$

برقرار باشد و تجزیه $LU$ ماتریس $X^{-۱}$ وجود داشته باشد، آنگاه دنباله‌ی $\{A_k\}$ به یک ماتریس بالا مثلثی همگراست.

**اثبات :** تمرین

**نکته ۵.۸**

فرض کنید شرایط قضیه ی قبل برقرار باشند. آنگاه هرچه نسبت های

$$\frac{|\lambda_i|}{|\lambda_{i-۱}|}, \quad i = ۲, ۳, \dots, n,$$

کوچکتر باشند روش QR سریع تر همگرا خواهد بود. برای اثبات کتاب زیر را ببینید

J. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford. (1965) .

همان‌طور که از الگوریتم QR برمی‌آید در هر تکرار این الگوریتم نیاز است که تجزیه QR ماتریس $A$ را محاسبه کنیم. محاسبه تجزیه QR یک ماتریس مربعی می‌تواند به یکی از روشهای زیر انجام پذیرد :

۱- الگوریتم گرام_اشمیت

۲- ماتریس‌های دوران گیونز

۳- ماتریس‌های انعکاسی هاوس هولدر

در فصل اول با الگوریتم گرام_اشمیت و نحوه‌ی محاسبه تجزیه QR ماتریس $A$ با آن آشنا شدید. در این بخش از روش‌های دوم و سوم برای محاسبه‌ی تجزیه QR استفاده نمی‌شود و این تجزیه را با همان الگوریتم گرام اشمیت محاسبه می‌کنیم، هر چند ماتریس‌های گیونز و هاوس هولدر ابزارهایی هستند که کاربردهای دیگری نیز دارند، در مطالب بعدی و در صورت نیاز به تشریح آنها می‌پردازیم.

**مثال ۵.۲۹**

با استفاده از الگوریتم $QR$ مقادیر ویژه‌ی ماتریس زیر را به دست آورید

$$A = \begin{bmatrix} -۶ & ۲ \\ ۵ & ۸ \end{bmatrix}$$

**حل :** ابتدا تجزیه $QR$ ماتریس $A_۱ = A$ را به دست می‌آوریم

$$\text{تجزیه} \qquad A_۱ = Q_۱R_۱$$

$$Q_۱ = \begin{bmatrix} -۰/۷۶۸۲ & ۰/۶۴۰۲ \\ ۰/۶۴۰۲ & ۰/۷۶۸۲ \end{bmatrix}, \quad R_۱ = \begin{bmatrix} ۷/۸۱۰۲ & ۳/۵۸۵۰ \\ ۰ & ۷/۴۲۶۱ \end{bmatrix}$$

$$\text{تعریف} \qquad A_۲ = R_۱Q_۱ = \begin{bmatrix} -۳/۷۰۴۹ & ۷/۷۵۴۱ \\ ۴/۷۵۴۱ & ۵/۷۰۴۹ \end{bmatrix}$$

$$\text{تجزیه} \qquad A_۲ = Q_۲R_۲$$

$$Q_۲ = \begin{bmatrix} -۰/۶۱۴۷ & ۰/۷۸۸۸ \\ ۰/۷۸۸۸ & ۰/۶۱۴۷ \end{bmatrix}, \quad R_۲ = \begin{bmatrix} ۶/۰۲۷۳ & -۰/۲۶۶۵ \\ ۰ & ۹/۶۲۲۹ \end{bmatrix}$$

$$\text{تعریف} \qquad A_۳ = R_۲Q_۲ = \begin{bmatrix} -۳/۹۱۵۲ & ۴/۵۹۰۳ \\ ۷/۵۹۰۳ & ۵/۹۱۵۲ \end{bmatrix}$$

$$\text{تجزیه} \qquad A_۳ = Q_۳R_۳$$

$$Q_۳ = \begin{bmatrix} -۰/۴۵۸۴ & ۰/۸۸۸۷ \\ ۰/۸۸۸۷ & ۰/۴۵۸۴ \end{bmatrix}, \quad R_۳ = \begin{bmatrix} ۸/۵۴۰۵ & ۳/۱۵۲۷ \\ ۰ & ۶/۷۹۱۲ \end{bmatrix}$$

$$\text{تعریف} \qquad A_۴ = R_۳Q_۳ = \begin{bmatrix} -۱/۱۱۳۲ & ۹/۰۳۵۵ \\ ۶/۰۳۵۵ & ۳/۱۱۳۲ \end{bmatrix}$$

$$\vdots$$

$$\text{تجزیه} \qquad A_{۵۱} = Q_{۵۱}R_{۵۱}$$

$$Q_{۵۱} = \begin{bmatrix} -۱/۰۰۰۰ & -۰/۰۰۰۰ \\ -۰/۰۰۰۰ & ۱/۰۰۰۰ \end{bmatrix}, \quad R_{۵۱} = \begin{bmatrix} -۸/۶۸۱۲ & ۳/۰۰۰۰ \\ ۰ & -۶/۶۸۱۱ \end{bmatrix}$$

$$\text{تعریف} \qquad A_{۵۲} = R_{۵۱}Q_{۵۱} = \begin{bmatrix} ۸/۶۸۱۱ & ۳/۰۰۰۱ \\ ۰/۰۰۰۱ & -۶/۶۸۱۱ \end{bmatrix}$$

$$A_{52} = Q_{52}R_{52} \qquad \text{تجزیه}$$

$$Q_{52} = \begin{bmatrix} -1/0000 & -0/0000 \\ -0/0000 & 1/0000 \end{bmatrix}, \quad R_{52} = \begin{bmatrix} -8/6811 & -3/0000 \\ 0 & -6/6812 \end{bmatrix}$$

$$A_{53} = R_{52}Q_{52} = \begin{bmatrix} 8/6812 & -2/9999 \\ 0/0001 & -6/6812 \end{bmatrix} \qquad \text{تعریف}$$

$$A_{53} = Q_{53}R_{53} \qquad \text{تجزیه}$$

$$Q_{53} = \begin{bmatrix} -1/0000 & -0/0000 \\ -0/0000 & 1/0000 \end{bmatrix}, \quad R_{53} = \begin{bmatrix} -8/6812 & 3/0000 \\ 0 & -6/6811 \end{bmatrix}$$

$$A_{54} = R_{53}Q_{53} = \begin{bmatrix} 8/6811 & 3/0000 \\ 0/0000 & -6/6811 \end{bmatrix} \qquad \text{تعریف}$$

چون ماتریس $A_{54}$ به یک ماتریس بالا مثلثی نزدیک شده است پس اعضای روی قطر این ماتریس را می‌توان تقریبی از مقادیر ویژه‌ی $A$ پذیرفت

$$\lambda_1 \approx 8/6811, \qquad \lambda_2 \approx -6/6811$$

البته اینها تا ۴ رقم بعد از اعشار دقیق هستند زیرا مقادیر ویژه‌ی $A$ تا ۶ رقم بعد از اعشار به صورت زیر می‌باشد

$$\lambda_1 = 8/681146, \qquad \lambda_2 = -6/681146$$

به طور مشابه با روش LU اگر $Y$ یک بردار ویژه‌ی $A_k$ باشد آنگاه $X = QY = Q_1Q_2 \dots Q_{k-1}Y$ بردار ویژه‌ی $A$ خواهد بود. حال می توان دید که بردارهای ویژه ماتریس بالا مثلثی $A_{54}$ چنین اند:

$$Y_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad Y_2 = \begin{bmatrix} -0/1917 \\ 0/9815 \end{bmatrix}$$

و چون

$$Q = Q_1Q_2 \dots Q_{53} = \begin{bmatrix} -0/1350 & -0/9908 \\ -0/9908 & 0/1350 \end{bmatrix}$$

پس بردارهای ویژه ی $A$ به صورت زیر به دست می آیند

$$X_1 = QY_1 = \begin{bmatrix} -0/1350 \\ -0/9908 \end{bmatrix}, \qquad X_2 = QY_2 = \begin{bmatrix} -0/9466 \\ 0/3224 \end{bmatrix}$$

**توجه ۵.۲۴**

شرایط قضیه روش QR شرایط لازم همگرایی نیستند. برای مثال ماتریس زیر را در نظر بگیرید

$$A = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}, \quad \lambda_1 = \lambda_2 = 1$$

واضح است که شرط $|\lambda_1| > |\lambda_2| >$ برقرار نیست، اما با محاسبه ی دنباله ی $\{A_k\}$ حاصل از روش تجزیه QR می توان دید که

$$\lim_{k\longrightarrow\infty} A_k = \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}$$

که نشان از همگرایی دنباله $\{A_k\}$ دارد.

# ۱۰ الگوریتم هسنبرگ - QR

الگوریتم QR برای محاسبه ی مقادیر ویژه از دو عیب زیر برخوردار است:
۱. با توجه به اینکه هزینه ی محاسباتی الگوریتم گرام- اشمیت از $O(n^3)$ است(تمرینات فصل اول را ببینید) پس هر تکرار الگوریتم QR هزینه محاسباتی حدود $O(n^3)$ خواهد داشت و چنانچه تعداد تکرارهای لازم برای الگوریتم QR در حدود $n$ باشد آنگاه کل هزینه محاسباتی الگوریتم QR از $O(n^4)$ خواهد بود که مقرون به صرفه نیست.

۲. دیدیم که اگر دو مقدار ویژه ی غالب $A$ به هم نزدیک باشند آنگاه الگوریتم QR کند خواهد بود.

۱- برطرف کردن مشکل اول : ابتدا ماتریس دلخواه $A$ را به یک ماتریس بالاهسنبرگی (Hessenberg matrix) مثل $\widehat{A}$ تبدیل می کنیم و سپس الگوریتم QR را برای ماتریس $\widehat{A}$ اعمال می کنیم(می توان ثابت کرد که در دنباله $\{\widehat{A}\}_{k=1}^{\infty}$ حاصل از الگوریتم QR بر ماتریس $\widehat{A}$ تمامی ماتریس ها بالا هسنبرگی هستند ) با اینکار کل هزینه ی محاسباتی از $O(n^3)$ خواهد بود. زیرا اولا تبدیل کردن ماتریس $A$ به یک ماتریس بالا هسنبرگی هزینه محاسباتی از $O(n^3)$ دارد بعلاوه اعمال الگوریتم QR بر دنباله ی ماتریسی $\{\widehat{A}\}_{k=1}^{\infty}$ در هر تکرار هزینه ی محاسباتی از $O(n^2)$ خواهد رسید و در صورتی که حداقل به $n$ تکرار نیاز باشد آنگاه هزینه محاسباتی QR بر دنباله ی ماتریسی $\{\widehat{A}\}_{k=1}^{\infty}$ از $O(n^3)$ خواهد بود. بنابراین کل هزینه ی محاسباتی از $O(n^3)+O(n^3)=O(n^3)$ خواهد ماند و این خیلی مقرون به صرفه تر از $O(n^4)$ است.
**بنابراین با این تکنیک حجم عملیات از $O(n^4)$ به $O(n^3)$ کاهش می یابد.**

۲- برطرف کردن مشکل دوم: معمولا سعی می شود با یک سری تبدیلات ماتریسی بجای محاسبه ی مقادیر ویژه ی A به محاسبه ی مقادیر ویژه ی ماتریس انتقال یافته ی $B=A-\alpha I$ بپردازند که $\alpha$ طوری تعیین می گردد که مقادیر ویژه ی غالب B از هم فاصله داشته باشند در واقع سعی می شود تا $\alpha$ به طوری تعیین گردد که:

$$\left|\frac{\lambda_2^B}{\lambda_1^B}\right| < \left|\frac{\lambda_2^A}{\lambda_1^A}\right|$$

و سپس با این انتظار که الگوریتم QR برای B سریع همگرا گردد به محاسبه ی مقادیر ویژه ی $B$ و پس از رابطه ی $\lambda(A)=\lambda(B)+\alpha$ به محاسبه ی مقادیر ویژه ی ماتریس A می پردازند. در این درس مبحث اصلی بر روی برطرف کردن مشکل اول خواهد بود زیرا همان طور که ذکر کردیم با اینکار می توان هزینه ی محاسباتی الگوریتم QR را از $O(n^4)$ به $O(n^3)$ کاهش داد درحالی که با برطرف کردن مشکل دوم همچنان انتظار داریم الگوریتم دارای هزینه ی محاسباتی بالایی باشد. بویژه اینکه در مبحث روش توانی با ایده ی انتقال کم و بیش آشنا شدید. بنابراین برای بحث در مورد برطرف کردن مشکل دوم خواننده را به کتاب زیر ارجاع می دهیم.
J. H. Wilkinson, The algebraic eigenvalue problem, Clarendon Press Oxford University Press (1988).
برای برطرف کردن مشکل اول لازم است که تبدیل یک ماتریس دلخواه به یک ماتریس بالاهسنبرگی را بدانیم. برای اینکار از ۲ روش زیر می توان استفاده کرد.
۱. استفاده از ماتریس های دوران گیونز (Givens rotation)
۲. استفاده از ماتریس های تبدیل هاوس هولدر (Householder transformation)

**توجه ۵.۲۵**

در این درس تنها با ماتریس های هاوس هولدر آشنا می شوید. علاقمندان جهت آشنایی با ماتریس های گیونز به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

# ۱۱ ماتریس های هاوس هولدر

ماتریس های هاوس هولدر نیز همانند ماتریس های دوران گیونز از کاربرد های مختلفی برخوردارند برای مثال محاسبه تجزیه $QR$ ماتریس دلخواه $A$را می توان به کمک این ماتریس ها به دست آورد.

**تعریف ۵.۱**

یک ماتریس هاوس هولدر به صورت زیر تعریف می شود

$$H_{n\times n} = I - \frac{2}{u^T u} u u^T \ , \ I \in \mathbb{R}^{n\times n}, u \in \mathbb{R}^n$$

که در آن $u$ یک بردار غیر صفر $1 \times n$ است.

**مثال ۵.۳۰**

ماتریس هاوس هولدر (Householder) متناظر با بردار $u = [1, 2]^T$ را بیابید.

**حل :**

$$u u^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [1, 2] = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$u^T u = [1, 2] \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (1)(1) + (2)(2) = 5$$

$$H = I - \frac{2}{u^T u} u u^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{2}{5} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{-4}{5} & \frac{-3}{5} \end{bmatrix}$$

با توجه به اینکه $H = H^T$ داریم

$$H H^T = H^2 = \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{-4}{5} & \frac{-3}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{-4}{5} & \frac{-3}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

می توان دریافت که $H$ ماتریسی متعامد است. این را می توان در حالت کلی نیز بررسی نمود.

پس برای هر بردار $u$ داریم

$$\boxed{H H^T = H^T H = H^2 = I} \qquad (47)$$

از طرفی با محاسبه $Hu$ داریم

$$Hu = \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{-4}{5} & \frac{-3}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = -u$$

که این موضوع نیز در حالت کلی برقرار است لذا :

برای هر بردار ناصفر $u$ داریم

$$\boxed{Hu=-u} \tag{۴۸}$$

تساوی بالا نشان می دهد که ماتریس $H$ یک مقدار ویژه $\lambda=-۱$ دارد و در واقع $(-۱,u)$ یک جفت ویژه آن است.

حال فرض کنید $X\in\begin{bmatrix}x_۱\\x_۲\end{bmatrix}\in\mathbb{R}^۲$ برداری دلخواه باشد. با استفاده از ماتریس هاوس هولدر مثال قبل داریم

$$u^THX=[۱,\ ۲]\begin{bmatrix}\frac{۳}{۵} & \frac{-۴}{۵}\\ \frac{-۴}{۵} & \frac{-۳}{۵}\end{bmatrix}\begin{bmatrix}x_۱\\x_۲\end{bmatrix}=[-۱,\ -۲]\begin{bmatrix}x_۱\\x_۲\end{bmatrix}=-x_۱-۲x_۲$$

و

$$-u^TX=-[۱,\ ۲]\begin{bmatrix}x_۱\\x_۲\end{bmatrix}=-x_۱-۲x_۲$$

در واقع برای هر بردار دلخواه $X$ داریم

$$\boxed{u^THX=-u^TX} \tag{۴۹}$$

این موضوع نیز در حالت کلی برقرار است در واقع برای هر $X,u\in\mathbb{R}^n$ ابتدا از (۴۸) داریم

$$Hu=-u\rightarrow u^TH^T=-u^T\rightarrow u^TH=-u^T$$

حال با محاسبه $u^THX$ داریم

$$u^THX=(u^TH)X=(-u^T)X=-u^TX$$

که نشان می دهد (۴۹) برقرار است. به کمک رابطه (۴۹) یک ویژگی جالب ماتریس های هاوس هولدر را می توان نشان داد. فرض کنید $\theta_۱$ زاویه بین دو بردار ناصفر $X$ و $u$ و $\theta_۲$ زاویه بین دو بردار $u$ و $HX$ باشد آنگاه باید

$$\cos\theta_۱=\frac{<u,X>}{||u||_۲\ ||X||_۲}=\frac{u^TX}{||u||_۲\ ||X||_۲} \tag{۵۰}$$

$$\cos\theta_۲=\frac{<u,HX>}{||u||_۲\ ||HX||_۲}=\frac{u^THX}{||u||_۲\ ||X||_۲}\overset{(۴۹)}{=}\frac{-u^TX}{||u||_۲\ ||X||_۲} \tag{۵۱}$$

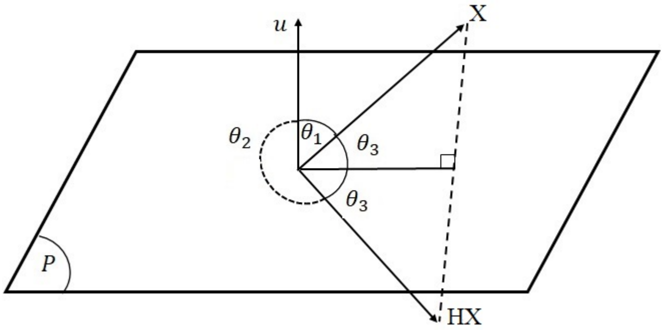
توجه کنید چون $H$ متعامد است در (۵۱) از $||HX||_۲=||X||_۲$ استفاده شده است. حال با تقسیم طرفین (۵۰) و (۵۱) بر هم به دست می آوریم

$$\frac{\cos\theta_۱}{\cos\theta_۲}=-۱\Longrightarrow\cos\theta_۲=-\cos\theta_۱$$

پس $\theta_۲=\pi+\theta_۱$ خواهد بود. این نشان می دهد که اگر $\theta_۱$ زاویه بین دو بردار $X$ و $u$ باشد آنگاه زاویه بین دو بردار $u$ و $HX$ برابر $\theta_۲=\pi+\theta_۱$ خواهد بود. این موضوع در شکل ۱ دیده می شود. توجه کنید که در این شکل داریم

$$\theta_۱+\theta_۲+۲\theta_۳=۲\pi$$

و چون $\theta_۱+\theta_۳=\frac{\pi}{۲}$ پس $\theta_۲=\theta_۱+\pi$ که همان نتیجه ی از قبل به دست آمده است.




شکل ۱: زاویه بین دو بردار $u$ و $HX$

بنابراین $HX$ در واقع بازتاب(انعکاس) بردار $X$ بر صفحه $P$ است.

از اینرو گاهی اوقات به ماتریس های هاوس هولدر بازتاب های هاوس هولدر نیز گفته می شود. توجه کنید که طبق شکل متوجه می شویم که سه بردار $X, u, HX$ در یک صفحه قرار دارند. این را به طور جبری نیز می توان نشان داد زیرا داریم

$$HX = (I - \frac{۲}{u^T u} uu^T)X = X - \frac{۲}{u^T u} uu^T X$$

اما از اینجایی که $u^T X$ یک اسکالر است پس $u(u^T X) = (u^T X)u$ یعنی $uu^T X = u^T X u$ پس

$$HX = X - \frac{۲}{u^T u} u^T X u = X + \alpha u \tag{۵۲}$$

که در آن $\alpha = -\frac{۲u^T X}{u^T u}$. واضح است که (۵۲) نشان می دهد سه بردار $X, u, HX$ بر یک صفحه اند.
اما خاصیت مهمی که ماتریس های هاوس هولدر دارند این است که می توانند برای صفر کردن درایه های یک بردار مفروض بکار روند.

## ۱.۱۱ صفر کردن درایه های دوم تا آخر یک بردار مفروض به کمک ماتریس های هاوس هولدر

فرض کنید بردار $X = [x_۱,\ x_۲,\ x_۳, \cdots, x_n]^T$ داده شده است و می خواهیم ماتریس هاوس هولدر H را به قسمی بیابیم که

$$HX = \begin{bmatrix} \beta \\ ۰ \\ ۰ \\ \vdots \\ ۰ \end{bmatrix}$$

که $\beta$ یک اسکالر باشد. یعنی می خواهیم $H$ را به قسمی محاسبه کنیم که با ضرب آن در $X$ درایه های $x_۲, x_۳, \cdots, x_n$ به صفر تبدیل شوند. برای اینکار ابتدا بردار $u$ را به صورت

$$u = X + \text{sign}(x_۱)||X||_۲ e_۱$$

تعریف می کنیم که در آن $e_۱$ ستون اول ماتریس همانی $n \times n$ و sign به صورت زیر تعریف می شود:

$$\text{sign}(x_۱) = \begin{cases} ۱, & if\ x_۱ \geq ۰ \\ -۱, & if\ x_۱ < ۰ \end{cases}$$

با استفاده از بردار $u$ فوق است ماتریس $H$ را به صورت

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

بسازیم آنگاه خواهیم دید که

$$HX = [\beta,\ 0,\ 0,\ \cdots,\ 0]^T$$

بعلاوه $\beta$ به صورت زیر خواهد بود

$$\beta = -\operatorname{sign}(x_1)||X||_2$$

از اثبات این موضوع در حالت کلی اجتناب می ورزیم و آن را در حالت ساده تر $n = 2$ بررسی می کنیم

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

**حالت اول:** $x_1 \geq 0$ آنگاه $\operatorname{sign}(x_1) = 1$ پس

$$u = X + \operatorname{sign}(x_1)||X||_2 e_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} ||X||_2 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + ||X||_2 \\ x_2 \end{bmatrix}$$

پس داریم

$$u^T u = [x_1 + ||X||_2,\ x_2] \begin{bmatrix} x_1 + ||X||_2 \\ x_2 \end{bmatrix} = (x_1 + ||X||_2)^2 + x_2^2 = 2x_1^2 + 2x_1||X||_2 + 2x_2^2$$

$$u^T X = [x_1 + ||X||_2,\ x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1 + ||X||_2)x_1 + x_2^2 = x_1^2 + x_1||X||_2 + x_2^2$$

از طرفی دیدیم که $HX = X + \alpha u$ که در آن

$$\alpha = \frac{-2u^T X}{u^T u}$$

در اینجا

$$\alpha = -\frac{2x_1^2 + 2x_1||X||_2 + 2x_2^2}{2x_1^2 + 2x_1||X||_2 + 2x_2^2} = -1$$

پس

$$HX = X - u = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} x_1 + ||X||_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -||X||_2 \\ 0 \end{bmatrix} = \begin{bmatrix} -\operatorname{sign}(x_1)||X||_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix}$$

که درستی حکم را نشان می دهد. اثبات در حالت $x_1 < 0$ به طور مشابه انجام می شود.

**مثال ۵.۳۱**

ماتریس هاوس هولدر $H$ را به قسمی بیابید که درایه دوم $X = \begin{bmatrix} ۳ \\ ۴ \end{bmatrix}$ به صفر تبدیل شود.

**حل :** داریم $||X||_۲ = \sqrt{۱۶ + ۹} = ۵$. حال از اینکه $\operatorname{sign}(x_۱) = \operatorname{sign}(۳) = +۱$ داریم

$$u = X + \operatorname{sign}(x_۱)||X||_۲ e_۱ = X + ||X||_۲ e_۱ = \begin{bmatrix} ۳ \\ ۴ \end{bmatrix} + ۵ \begin{bmatrix} ۱ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۸ \\ ۴ \end{bmatrix}$$

اما $u^T u = ۸۰$ و

$$uu^T = \begin{bmatrix} ۶۴ & ۳۲ \\ ۳۲ & ۱۶ \end{bmatrix}$$

پس

$$H = I - \frac{۲}{u^T u} uu^T = \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \end{bmatrix} - \frac{۲}{۸۰} \begin{bmatrix} ۶۴ & ۳۲ \\ ۳۲ & ۱۶ \end{bmatrix} = \begin{bmatrix} -\frac{۳}{۵} & -\frac{۴}{۵} \\ -\frac{۴}{۵} & \frac{۳}{۵} \end{bmatrix}$$

بنابراین

$$HX = \begin{bmatrix} -\frac{۳}{۵} & -\frac{۴}{۵} \\ -\frac{۴}{۵} & \frac{۳}{۵} \end{bmatrix} \begin{bmatrix} ۳ \\ ۴ \end{bmatrix} = \begin{bmatrix} -۵ \\ ۰ \end{bmatrix}$$

توجه کنید که $\beta = -۵$ که با $\beta = -\operatorname{sign}(x_۱)||X||_۲ = -||X||_۲ = -۵$ مطابقت دارد.

**مثال ۵.۳۲**

درایه های دوم و سوم بردار $X = \begin{bmatrix} -۱ \\ ۲ \\ -۲ \end{bmatrix}$ را با کمک یک ماتریس هاوس هولدر به صفر تبدیل نمایید.

**حل :** ابتدا توجه کنید که $||X||_۲ = \sqrt{۱ + ۴ + ۴} = ۳$ و

$$\operatorname{sign}(x_۱) = \operatorname{sign}(-۱) = -۱$$

پس

$$u = X + \operatorname{sign}(x_۱)||X||_۲ e_۱ = X - ||X||_۲ e_۱ = \begin{bmatrix} -۱ \\ ۲ \\ -۲ \end{bmatrix} - ۳ \begin{bmatrix} ۱ \\ ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} -۴ \\ ۲ \\ -۲ \end{bmatrix}$$

بعلاوه

$$u^T u = [-۴,\ ۲,\ -۲] \begin{bmatrix} -۴ \\ ۲ \\ -۲ \end{bmatrix} = ۲۴$$

$$uu^T = \begin{bmatrix} -۴ \\ ۲ \\ -۲ \end{bmatrix} [-۴,\ ۲,\ -۲] = \begin{bmatrix} ۱۶ & -۸ & ۸ \\ -۸ & ۴ & -۴ \\ ۸ & -۴ & ۴ \end{bmatrix}$$

پس

$$H = I - \frac{2}{u^T u} u u^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{24} \begin{bmatrix} 16 & -8 & 8 \\ -8 & 4 & -4 \\ 8 & -4 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

بنابراین

$$HX = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

مشاهده می شود که با ماتریس هاوس هولدر توانستیم دو درایه $X$ را به صفر تبدیل کنیم.

**مثال ۵.۳۳**

اگر $X = [2, 1, -1, 3] \in \mathbb{R}^4$ آنگاه ماتریس $H$ را به قسمی بیابید که $HX = [\beta,\ 0,\ 0,\ 0]^T$. مقدار $\beta$ را تعیین کنید.

**حل :**

$$u = X + \operatorname{sign}(x_1)||X||_2 e_1 = \begin{bmatrix} 5/8730 \\ 1/0000 \\ -1/0000 \\ 3/0000 \end{bmatrix}$$

$$u^T u = 45/4919$$

$$uu^T = \begin{bmatrix} 34/4919 & 5/8730 & -5/8730 & 17/6190 \\ 5/8730 & 1/0000 & -1/0000 & 3/0000 \\ -5/8730 & -1/0000 & 1/0000 & -3/0000 \\ 17/6190 & 3/0000 & -3/0000 & 9/0000 \end{bmatrix}$$

$$H = \begin{bmatrix} -0/5164 & -0/2582 & 0/2582 & -0/7746 \\ -0/2582 & 0/9560 & 0/0440 & -0/1319 \\ 0/2582 & 0/0440 & 0/9560 & 0/1319 \\ -0/7746 & -0/1319 & 0/1319 & 0/6043 \end{bmatrix}$$

لذا داریم

$$HX = \begin{bmatrix} -0/5164 & -0/2582 & 0/2582 & -0/7746 \\ -0/2582 & 0/9560 & 0/0440 & -0/1319 \\ 0/2582 & 0/0440 & 0/9560 & 0/1319 \\ -0/7746 & -0/1319 & 0/1319 & 0/6043 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -3/8730 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

بنابراین $\beta = -3/8730$ می‌باشد.

# ۱۲ تبدیل یک ماتریس دلخواه به یک ماتریس به شکل هسنبرگی با استفاده از ماتریس‌های هاوس‌هولدر

به طور کلی تبدیل یک ماتریس به شکل هسنبرگی می تواند مزایای مفیدی به همراه داشته باشد. برای مثال در بهبود الگوریتم QR دیدیم که نیاز است چنین فرآیندی اجرا شود.

در پیوست با استفاده از ماتریس‌های گیونز نحوه‌ی به دست آمدن تجزیه هسنبرگی ماتریس دلخواه $A$ را به صورت

$$Q^T A Q = H$$

که $H$ ماتریس بالاهسنبرگی و $Q$ متعامد است ، را بیان نمودیم. تبدیل ماتریس $A$ به یک ماتریس بالاهسنبرگی $H$ با استفاده از ماتریس‌های هاوس‌هولدر سریع‌تر از ماتریس‌های گیونز خواهد بود. زیرا همان‌طور که با استفاده از ماتریس‌های گیونز دیدیم برای صفر نمودن هر درایه به یک ماتریس گیونز نیاز داریم ، در حالی‌که در استفاده از ماتریس‌های هاوس‌هولدر چنین محدودیتی نداریم و می‌توان به یکباره این کار را انجام داد. قبل از بیان بحث اصلی به برخی مطالب مقدماتی نیازمندیم.

## ۱.۱۲ صفر کردن درایه‌های سوم تا آخر یک بردار

قبل از وارد شدن به بحث اصلی ، ابتدا فرض کنید بردار $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ مفروض است و هدف صفر کردن درایه‌های سوم تا آخر $X$ است. یعنی می‌خواهیم $x_3, x_4, \ldots, x_n$ را به صفر تبدیل نماییم. دیدیم که ماتریس $H$ را می‌توان به قسمی ساخت که از درایه‌ی دوم تا آخر $X$ به صفر تبدیل شود. اما در اینجا می‌خواهیم $H$ را طوری بسازیم که

$$HX = \begin{bmatrix} * \\ * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

برای این کار ابتدا بردار $\widetilde{X}$ را به صورت زیر در نظر می‌گیریم :

$$\widetilde{X} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

که در واقع از درایه‌های دوم تا آخر بردار $X$ ساخته شده است. اکنون می‌توانیم ماتریس $\widetilde{H}$ را به قسمی به دست آوریم که

$$\widetilde{H}\widetilde{X} = \begin{bmatrix} * \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{۵۳}$$

که $\widetilde{H}$ یک ماتریس هاوس‌هولدر با اندازه ی $(n-1) \times (n-1)$ است. حال اگر تعریف کنیم

$$H = \begin{bmatrix} I_1 & 0 \\ 0 & \widetilde{H} \end{bmatrix}$$

که $I_۱$ ماتریس همانی ۱ × ۱ است. در اینصورت $H$ ماتریس $n \times n$ خواهد بود. آنگاه می‌توان دید که

$$HX = \begin{bmatrix} I_۱ & ۰ \\ ۰ & \widetilde{H} \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \\ x_۳ \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} * \\ * \\ ۰ \\ \vdots \\ ۰ \end{bmatrix} \tag{۵۴}$$

بنابراین ماتریس $H$ ای که به دنبالش بودیم را به دست آوردیم. اثبات رابطه (۵۴) سخت نیست زیرا

$$HX = \begin{bmatrix} I_۱ & ۰ \\ ۰ & \widetilde{H} \end{bmatrix} \begin{bmatrix} x_۱ \\ \widetilde{X} \end{bmatrix} = \begin{bmatrix} I_۱ \times x_۱ + ۰ \times \widetilde{X} \\ ۰ \times x_۱ + \widetilde{H}\widetilde{X} \end{bmatrix} = \begin{bmatrix} x_۱ \\ \widetilde{H}\widetilde{X} \end{bmatrix} \overset{(۵۳)}{=} \begin{bmatrix} x_۱ \\ * \\ ۰ \\ \vdots \\ ۰ \end{bmatrix}$$

بنابراین ماتریس $H$ به قسمی ساخته می‌شود که در بردار $HX$ درایه ی اول همان درایه‌ی اول بردار $X$ یعنی $x_۱$ است، درایه ی دوم تغییر می‌کند (که می‌تواند مقداری صفر باشد یا نباشد) و درایه سوم تا آخر به صفر تبدیل می شوند.

**مثال ۵.۳۴**

درایه‌ی سوم تا آخر بردار $X = [۵, ۱, -۲, ۲]^T$ را به صفر تبدیل نمایید.

**حل :** قرار می‌دهیم

$$\widetilde{X} = \begin{bmatrix} ۱ \\ -۲ \\ ۲ \end{bmatrix}$$

که از عناصر دوم تا آخر $X$ ساخته می شود. حال باید ماتریس $\widetilde{H}$ را به قسمی بسازیم که

$$\widetilde{H}\widetilde{X} = \begin{bmatrix} * \\ ۰ \\ ۰ \end{bmatrix}$$

اما قبلاً دیدیم که چنین ماتریسی می‌تواند به صورت زیر ساخته شود

$$\left\|\widetilde{X}\right\|_۲ = \sqrt{۱ + ۴ + ۴} = ۳, \quad \text{sign}\left(\widetilde{x}_۱\right) = \text{sign}(۱) = ۱$$

$$u = \widetilde{X} + \text{sign}\left(\widetilde{x}_۱\right)\left\|\widetilde{X}\right\|_۲ e_۱ = \widetilde{X} + \left\|\widetilde{X}\right\|_۲ e_۱ = \begin{bmatrix} ۱ \\ -۲ \\ ۲ \end{bmatrix} + ۳\begin{bmatrix} ۱ \\ ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۴ \\ -۲ \\ ۲ \end{bmatrix}$$

$$u^T u = [۴, -۲, ۲]\begin{bmatrix} ۴ \\ -۲ \\ ۲ \end{bmatrix} = ۴^۲ + ۲^۲ + ۲^۲ = ۲۴$$

$$uu^T = \begin{bmatrix} ۴ \\ -۲ \\ ۲ \end{bmatrix}[۴, -۲, ۲] = \begin{bmatrix} ۱۶ & -۸ & ۸ \\ -۸ & ۴ & -۴ \\ ۸ & -۴ & ۴ \end{bmatrix}$$

$$\widetilde{H} = I_3 - \frac{2}{u^T u} u u^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{24} \begin{bmatrix} 16 & -8 & 8 \\ -8 & 4 & -4 \\ 8 & -4 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

بنابراین داریم

$$\widetilde{H}\widetilde{X} = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix}$$

در نتیجه

$$H = \begin{bmatrix} I_1 & 0 \\ 0 & \widetilde{H} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ 0 & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

بنابراین

$$HX = \begin{bmatrix} x_1 \\ \widetilde{H}\,\widetilde{x} \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 0 \\ 0 \end{bmatrix}$$

البته این را می‌توان به صورت مستقیم با ضرب $H$ در $X$ نیز بررسی نمود :

$$HX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ 0 & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 0 \\ 0 \end{bmatrix}$$

## ۲.۱۲ صفر کردن درایه‌های $k$ ام تا آخر یک بردار

اما ممکن است که بخواهیم برای بردار مفروض

$$X = \begin{bmatrix} x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n \end{bmatrix}^T$$

درایه‌های $x_k, x_{k+1}, \ldots, x_n$ را به صفر تبدیل نماییم که $k$ می‌تواند مقادیر $k = 2, 3, \ldots, n$ را اختیار نماید. آیا برای این کار نیز می‌توان از یک ماتریس هاوس‌هولدر استفاده نمود؟

در واقع باید گفت همواره این کار امکان‌پذیر است ، ابتدا بردار $X$ را به دو بخش به صورت زیر تقسیم می‌کنیم :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-2} \\ \hline x_{k-1} \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} X_1 \\ \hline X_2 \end{bmatrix}$$

می‌خواهیم از درایه‌ی $k$ - ام تا آخر $X$ را به صفر تبدیل کنیم. ماتریس $\widetilde{H}$ از مرتبه $(n-k+2) \times (n-k+2)$ را به قسمی محاسبه می‌کنیم که $\widetilde{H} X_2 = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ آنگاه با تعریف ماتریس $H$ به صورت

$$H = \begin{bmatrix} I_{k-2} & 0 \\ 0 & \widetilde{H} \end{bmatrix}$$

داریم

$$HX = \begin{bmatrix} I_{k-2} & 0 \\ 0 & \widetilde{H} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ \widetilde{H} X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

رابطه ی فوق نشان‌دهنده این است که درایه های $x_k, x_{k+1}, \dots, x_n$ بردار $X$ به صفر تبدیل خواهند شد.

**مثال ۵.۳۵**

با استفاده از تبدیلات هاوس هولدر درایه های چهارم و به بعد بردار داده شده را به صفر تبدیل کنید.

$$X = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \quad k = 4$$

**حل :** داریم

$$X_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, \quad \|X_2\|_2 = 7/0711$$

$$u = X_۲ + \operatorname{sign}\left((X_۲)_۱\right) \|X_۲\|_۲ \, e_۱ = \begin{bmatrix} ۱۰/۰۷۱۱ \\ ۴ \\ ۵ \end{bmatrix}$$

$$\widetilde{H} = I_۳ - \frac{۲}{u^T u} u u^T = \begin{bmatrix} -۰/۴۲۴۳ & -۰/۵۶۵۷ & -۰/۷۰۷۱ \\ -۰/۵۶۵۷ & ۰/۷۷۵۳ & -۰/۲۸۰۸ \\ -۰/۷۰۷۱ & -۰/۲۸۰۸ & ۰/۶۴۸۹ \end{bmatrix}$$

$$\widetilde{H} X_۲ = \begin{bmatrix} -۷/۰۷۱۱ \\ ۰ \\ ۰ \end{bmatrix}$$

به‌علاوه طبق مطالب بیان شده ماتریس $H$ این چنین است

$$H = \begin{bmatrix} I_{k-۲} & ۰ \\ ۰ & \widetilde{H} \end{bmatrix} = \begin{bmatrix} I_۲ & ۰ \\ ۰ & \widetilde{H} \end{bmatrix} = \left[\begin{array}{cc|ccc} ۱ & ۰ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ & ۰ \\ \hline ۰ & ۰ & -۰/۴۲۴۳ & -۰/۵۶۵۷ & -۰/۷۰۷۱ \\ ۰ & ۰ & -۰/۵۶۵۷ & ۰/۷۷۵۳ & -۰/۲۸۰۸ \\ ۰ & ۰ & -۰/۷۰۷۱ & -۰/۲۸۰۸ & ۰/۶۴۸۹ \end{array}\right]$$

پس

$$HX = \begin{bmatrix} I_۲ & ۰ \\ ۰ & \widetilde{H} \end{bmatrix} \begin{bmatrix} X_۱ \\ X_۲ \end{bmatrix} = \begin{bmatrix} X_۱ \\ \widetilde{H} X_۲ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۲ \\ -۷/۰۷۱۱ \\ ۰ \\ ۰ \end{bmatrix}$$

توجه کنید که درایه‌های چهارم و پنجم $X$ به صفر تبدیل شده‌اند.

## ۳.۱۲ تبدیل یک ماتریس به یک ماتریس هسنبرگی

اکنون آماده‌ایم تبدیل ماتریس $A$ به یک ماتریس هسنبرگی را با استفاده از ماتریس های هاوس هولدر بیان کنیم. در واقع باید تجزیه‌ای به صورت

$$QAQ^T = H$$

را بیابیم که $Q$ ماتریسی متعامد و $H$ ماتریسی بالاهسنبرگی است.

$$A = \begin{bmatrix} a_{۱۱} & a_{۱۲} & a_{۱۳} & \cdots & a_{۱,n-۲} & a_{۱,n-۱} & a_{۱n} \\ a_{۲۱} & a_{۲۲} & a_{۲۳} & \cdots & a_{۲,n-۲} & a_{۲,n-۱} & a_{۲n} \\ a_{۳۱} & a_{۳۲} & a_{۳۳} & \cdots & a_{۳,n-۲} & a_{۳,n-۱} & a_{۳n} \\ a_{۴۱} & a_{۴۲} & a_{۴۳} & \cdots & a_{۴,n-۲} & a_{۴,n-۱} & a_{۴n} \\ a_{۵۱} & a_{۵۲} & a_{۵۳} & \cdots & a_{۵,n-۲} & a_{۵,n-۱} & a_{۵n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n-۱,۱} & a_{n-۱,۲} & a_{n-۱,۳} & \cdots & a_{n-۱,n-۲} & a_{n-۱,n-۱} & a_{n-۱,n} \\ a_{n,۱} & a_{n,۲} & a_{n,۳} & \cdots & a_{n,n-۲} & a_{n,n-۱} & a_{n,n} \end{bmatrix}$$

درایه‌هایی که باید صفر شوند را مشخص کرده‌ایم.

**گام ۱ :**

ابتدا یک ماتریس هاوس‌هولدر $\widetilde{H}_1$ از مرتبه‌ی $n-1$ به قسمی تولید شود که

$$\widetilde{H}_1 \begin{bmatrix} a_{21} \\ a_{31} \\ a_{41} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

سپس از $\widetilde{H}_1$ ماتریس $H_1$ را به شکل زیر می‌سازیم

$$H_1 = \begin{bmatrix} I_1 & 0 \\ 0 & \widetilde{H}_1 \end{bmatrix}$$

از اینجا ماتریس $A^{(1)}$ به صورت

$$A^{(1)} = H_1 A H_1^T = \begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

ساخته می‌شود. همان‌طور که می‌بینید از سطر سوم تا آخر ستون اول ماتریس جدید به صفر تبدیل شده اند و بعلاوه متعامد بودن $H_1$ تضمین می‌کند که $H_1^T = H_1^{-1}$ لذا $A \sim A^{(1)}$.

**گام ۲ :**

یک ماتریس هاوس‌هولدر $\widetilde{H}_2$ از مرتبه‌ی $n-2$ به قسمی تولید شود که

$$\widetilde{H}_2 \begin{bmatrix} a_{32}^{(1)} \\ a_{42}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

توجه کنید که علامت (۱) بالای درایه‌ها نشان‌دهنده این است که درایه‌ها مربوط به ماتریس $A^{(1)}$ است.
اینک ماتریس $H_2$ را به صورت زیر تعریف می‌کنیم

$$H_2 = \begin{bmatrix} I_2 & 0 \\ 0 & \widetilde{H}_2 \end{bmatrix}$$

از آن ماتریس $A^{(2)}$ به صورت زیر ساخته می‌شود

$$A^{(2)} = H_2 A^{(1)} H_2^T = \begin{bmatrix} * & * & * & \dots & * \\ * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix}$$

همان‌طور که می‌بینید از سطر چهارم تا آخر ستون دوم ماتریس جدید به صفر تبدیل شده‌اند. بعلاوه $A \sim A^{(1)} \sim A^{(2)}$.

با ادامه‌ی این روند در گام $(n-۲)$ - ام ماتریس $A^{(n-۲)}$ که به صورت

$$A^{(n-۲)} = H_{n-۲}AH_{n-۲}^T = \begin{bmatrix} * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ ۰ & * & * & \dots & * & * & * \\ ۰ & ۰ & * & \dots & * & * & * \\ ۰ & ۰ & ۰ & \dots & * & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & * & * & * \\ ۰ & ۰ & ۰ & \dots & ۰ & * & * \end{bmatrix}$$

به دست می‌آید که یک ماتریس بالاهسنبرگی است.
پس قرار می‌دهیم $H = A^{(n-۲)}$ از طرفی می‌توان نوشت

$$\begin{aligned} H &= A^{(n-۲)} = H_{n-۲}A^{(n-۳)}H_{n-۲}^T \\ &= H_{n-۲}\left(H_{n-۳}A^{(n-۴)}H_{n-۳}^T\right)H_{n-۲}^T \\ &= H_{n-۲}A_{n-۳}\left(H_{n-۴}A^{(n-۵)}H_{n-۴}^T\right)H_{n-۳}^TH_{n-۲}^T \\ &\vdots \\ &= \left(H_{n-۲}A_{n-۳}H_{n-۴}\cdots H_۱\right)A\left(H_۱^TH_۲^T\dots H_{n-۴}^TH_{n-۳}^TH_{n-۲}^T\right) \\ &= \left(\underline{H_{n-۲}H_{n-۳}H_{n-۴}\cdots H_۱}\right)A\left(\underline{H_{n-۲}H_{n-۳}H_{n-۴}\dots H_۱}\right)^T \end{aligned} \tag{۵۵}$$

اگر تعریف کنیم $Q = H_{n-۲}H_{n-۳}H_{n-۴}\dots H_۱$ آنگاه از (۵۵) داریم $H = QAQ^T$ که همان تجزیه هسنبرگی ماتریس $A$ می‌باشد. توجه کنید که متعامد بودن $Q$ نتیجه می‌دهد که $A \sim H$

**توجه ۵.۲۶**

ماتریس $A$ که مربعی و $n \times n$ است ، با استفاده از روش فوق بعد از حداکثر $n-۲$ گام به یک ماتریس بالا هسنبرگی تبدیل می‌شود.

**مثال ۵.۳۶**

با استفاده از ماتریس‌های هاوس‌هولدر ماتریس زیر را به یک ماتریس هسنبرگی تبدیل کنید.

$$A = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۳ & -۵ & ۱ \\ ۴ & ۷ & ۲ \end{bmatrix}, \qquad n = ۳$$

**حل :**
تنها به $۱ = ۲ - ۳ = ۲ - n$ گام نیاز است تا $A$ به یک ماتریس بالاهسنبرگی تبدیل شود. پس $\widetilde{H}_۱$ را تشکیل می‌دهیم به قسمی که

$$\widetilde{H}_۱\begin{bmatrix} a_{۲۱} \\ a_{۳۱} \end{bmatrix} = \begin{bmatrix} * \\ ۰ \end{bmatrix} \Longrightarrow \widetilde{H}_۱\begin{bmatrix} ۳ \\ ۴ \end{bmatrix} = \begin{bmatrix} * \\ ۰ \end{bmatrix}$$

از مطالب قبلی می‌دانیم که $\widetilde{H}_1$ باید به صورت زیر ساخته شود

$$\left\|\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right\|_2 = \sqrt{9+16} = 5$$

$$u = \begin{bmatrix} 3 \\ 4 \end{bmatrix} + 5e_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\widetilde{H}_1 = I_2 - \frac{2}{u^T u} uu^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{2}{80}\begin{bmatrix} 64 & 32 \\ 32 & 16 \end{bmatrix} = \begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

با توجه به $\widetilde{H}_1$ ماتریس $H_1$ را می‌سازیم

$$H_1 = \begin{bmatrix} I_1 & 0 \\ 0 & \widetilde{H}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

لذا

$$A^{(1)} = H_1 A H_1^T = \begin{bmatrix} 1 & -\frac{18}{5} & \frac{1}{5} \\ -5 & \frac{83}{25} & \frac{19}{25} \\ 0 & -\frac{131}{25} & -\frac{158}{25} \end{bmatrix} = H$$

که به یک ماتریس بالاهسنبرگی تبدیل شده است. توجه کنید که در اینجا $Q = H_1$ و همچنین $A$ با $H$ متشابه خواهد بود.

**توجه ۵.۲۷**

وقتی با ماتریس های هاوس هولدر یک ماتریس را به فرم هسنبرگی کاهش دهیم آنگاه می توان الگوریتم $QR$ را بر این ماتریس هسنبرگی پیاده کرده و به محاسبه ی مقادیر ویژه آن پرداخت. علاقمندان برای جزییات بیشتر به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان ارایه می شود) مراجعه نمایند.

**تمرین ۵.۱۰**

کد متلب کاهش هسنبرگی با ماتریس‌های هاوس هولدر را بنویسید.

**توجه ۵.۲۸**

در نرم‌افزار متلب با استفاده از دستور $[Q, H] = hess(A)$ می‌توان تجزیه‌ای به صورت $P^T AP = H$ به دست آورد. توجه کنید که در این دستور آماده‌ی متلب ماتریس $Q$ در تجزیهِ هسنبرگ همان $P = Q^T$ خواهد بود.

**توجه ۵.۲۹**

یک روش مناسب جهت یافتن تمامی جفت ویژه های یک ماتریس متقارن روش ژاکوبی است. علاقمندان برای دیدن این روش به درس تحت عنوان روش های عددی در جبرخطی(درسنامه مقطع کارشناسی ارشد که توسط مهدی دهقان

ارایه می شود) مراجعه نمایند. همچنین برای ماتریس های متقارن می توان از الگوریتم لانزوس استفاده نمود.

# واژه‌نامه انگلیسی به فارسی

## C

Cayley-Hamilton theorem ...... قضیه کیلی-همیلتون
characteristic equation ............. معادله مشخصه
complex ................................ مختلط

## D

deflation ................................ تقلیل
dominant ............................... غالب
dominant eigenvalue .............. مقدارویژه غالب

## E

eigenvalue ........................... مقدار ویژه
eigenvector ......................... بردارویژه

## G

Gersgorin disks ............. قرص های گرشگورین
Givens matrix ..................... ماتریس گیونز
Givens rotation .................... دوران گیونز

## H

Hessenberg matrix ............... ماتریس هسنبرگ
Hessenberg reduction ............. کاهش هسنبرگ
Householder matrix ........... ماتریس هاوس هولدر
Householder transformation ..... تبدیل هاوس هولدر

## I

inverse power method / inverse iteration ... روش توانی معکوس

## J

Jacobi method ...................... روش ژاکوبی

## L

Lanczos algorithm ................ الگوریتم لانزوس
Lower triangular ..................... پایین مثلثی
LR algorithm ........................ الگوریتم LR

## P

partial pivoting .................. محورگیری جزئی
power method / direct iteration ......... روش توانی
power method with deflation ...... روش تقلیل توانی

## Q

QR iteration ........................ تکرار QR

## R

Rayleigh quotient iteration method ... روش خارج قسمت ریلی
rotation matrix .................... ماتریس دوران

## S

shift property ..................... خاصیت انتقال
shifted inverse iteration ... روش توانی معکوس با انتقال
shifted power method ........... روش توانی با انتقال
similar matrices ................ ماتریس های متشابه
stopping criterion ................... محک توقف

## T

the rayleigh quotient ............. خارج قسمت ریلی
tridiagonal ............................ سه قطری

## U

Upper Hessenberg ................. بالا هسنبرگ

# واژه‌نامه فارسی به انگلیسی

## ا

الگوریتم لانزوس ........ Lanczos algorithm
الگوریتم LR ........ LR algorithm

## ب

بالا هسنبرگ ........ Upper Hessenberg
بردارویژه ........ eigenvector

## پ

پایین مثلثی ........ Lower triangular

## ت

تبدیل هاوس هولدر ........ Householder transformation
تقلیل ........ deflation
تکرار QR ........ QR iteration

## خ

خارج قسمت ریلی ........ the rayleigh quotient
خاصیت انتقال ........ shift property

## د

دوران گیونز ........ Givens rotation

## ر

روش تقلیل توانی ........ power method with deflation
روش توانی ........ power method / direct iteration
روش توانی با انتقال ........ shifted power method
روش توانی معکوس ........ inverse power method / inverse iteration
روش توانی معکوس با انتقال ........ shifted inverse iteration
روش خارج قسمت ریلی ........ Rayleigh quotient iteration method
روش ژاکوبی ........ Jacobi method

## س

سه قطری ........ tridiagonal

## غ

غالب ........ dominant

## ق

قرص های گرشگورین ........ Gersgorin disks
قضیه کیلی- همیلتون ........ Cayley-Hamilton theorem

## ک

کاهش هسنبرگ ........ Hessenberg reduction

## م

ماتریس دوران ........ rotation matrix
ماتریس گیونز ........ Givens matrix
ماتریس هاوس هولدر ........ Householder matrix
ماتریس های متشابه ........ similar matrices
ماتریس هسنبرگ ........ Hessenberg matrix
محک توقف ........ stopping criterion
محورگیری جزئی ........ partial pivoting
مختلط ........ complex
معادله مشخصه ........ characteristic equation
مقدار ویژه ........ eigenvalue
مقدارویژه غالب ........ dominant eigenvalue

**دانشگاه صنعتی امیرکبیر**
**( پلی تکنیک تهران )**

# جبرخطی عددی

(کارشناسی)

## فصل پنجم:روش های مستقیم و تکراری برای محاسبه مقادیر ویژه و بردار ویژه یک ماتریس

**مدرس: مهدی دهقان**

دانشکده
ریاضی و علوم کامپیوتر

ترم دوم ۱۴۰۳-۱۴۰۲