11.2 Power Method

We now describe the power method for computing the dominant eigenpair. Its extension to the inverse power method is practical for finding any eigenvalue provided that a good initial approximation is known. Some schemes for finding eigenvalues use other methods that converge fast, but have limited precision. The inverse power method is then invoked to refine the numerical values and gain full precision. To discuss the situation, we will need the following definitions.

Definition 11.10. If λ_1 is an eigenvalue of $\boldsymbol{A}$ that is larger in absolute value than any other eigenvalue, it is called the ***dominant eigenvalue***. An eigenvector $\boldsymbol{V}_1$ corresponding to λ_1 is called a ***dominant eigenvector***. ▲

Definition 11.11. An eigenvector $\boldsymbol{V}$ is said to be ***normalized*** if the coordinate of largest magnitude is equal to unity (i.e., the largest coordinate in the vector $\boldsymbol{V}$ is the number 1). ▲

It is easy to normalize an eigenvector $\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}'$ by forming a new vector $\boldsymbol{V} = (1/c)\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}'$, where $c = v_j$ and $|v_j| = \max_{1 \le i \le n}\{|v_i|\}$.

Suppose that the matrix $\boldsymbol{A}$ has a dominant eigenvalue λ and that there is a unique normalized eigenvector $\boldsymbol{V}$ that corresponds to λ. This eigenpair λ, $\boldsymbol{V}$ can be found by the following iterative procedure called the ***power method***. Start with the vector

(1) $$\boldsymbol{X}_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}'.$$

Generate the sequence $\{\boldsymbol{X}_k\}$ recursively, using

(2) $$\begin{aligned} \boldsymbol{Y}_k &= \boldsymbol{A}\boldsymbol{X}_k, \\ \boldsymbol{X}_{k+1} &= \frac{1}{c_{k+1}}\boldsymbol{Y}_k, \end{aligned}$$

where c_{k+1} is the coordinate of $\boldsymbol{Y}_k$ of largest magnitude (in the case of a tie, choose the coordinate that comes first). The sequences $\{\boldsymbol{X}_k\}$ and $\{c_k\}$ will converge to $\boldsymbol{V}$ and λ, respectively:

(3) $$\lim_{k\to\infty} \boldsymbol{X}_k = \boldsymbol{V} \qquad \text{and} \qquad \lim_{k\to\infty} c_k = \lambda.$$

Remark. If $\boldsymbol{X}_0$ is an eigenvector and $\boldsymbol{X}_0 \neq \boldsymbol{V}$, then some other starting vector must be chosen.

Example 11.5. Use the power method to find the dominant eigenvalue and eigenvector for the matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 11 & -5 \\ -2 & 17 & -7 \\ -4 & 26 & -10 \end{bmatrix}.$$

Start with $\boldsymbol{X}_0 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}'$ and use the formulas in (2) to generate the sequence of vectors $\{\boldsymbol{X}_k\}$ and constants $\{c_k\}$. The first iteration produces

$$\begin{bmatrix} 0 & 11 & -5 \\ -2 & 17 & -7 \\ -4 & 26 & -10 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 12 \end{bmatrix} = 12\begin{bmatrix} \frac{1}{2} \\ \frac{2}{3} \\ 1 \end{bmatrix} = c_1\boldsymbol{X}_1.$$

The second iteration produces

$$\begin{bmatrix} 0 & 11 & -5 \\ -2 & 17 & -7 \\ -4 & 26 & -10 \end{bmatrix}\begin{bmatrix} \frac{1}{2} \\ \frac{2}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} \\ \frac{10}{3} \\ \frac{16}{3} \end{bmatrix} = \frac{16}{3}\begin{bmatrix} \frac{7}{16} \\ \frac{5}{8} \\ 1 \end{bmatrix} = c_2\boldsymbol{X}_2.$$

Table 11.1 Power Method Used in Example 11.5 to Find the Normalized Dominant Eigenvector $\boldsymbol{V} = \left[\frac{2}{5} \quad \frac{3}{5} \quad 1\right]'$ and Corresponding Eigenvalue $\lambda = 4$

$\boldsymbol{AX}_k =$		$\boldsymbol{Y}_k$		=	$c_{k+1}\boldsymbol{X}_{k+1}$			
$\boldsymbol{AX}_0 =$	[6.000000	8.000000	12.00000]′	=	12.00000[0.500000	0.666667	1]′	$= c_1\boldsymbol{X}_1$
$\boldsymbol{AX}_1 =$	[2.333333	3.333333	5.333333]′	=	5.333333[0.437500	0.625000	1]′	$= c_2\boldsymbol{X}_2$
$\boldsymbol{AX}_2 =$	[1.875000	2.750000	4.500000]′	=	4.500000[0.416667	0.611111	1]′	$= c_3\boldsymbol{X}_3$
$\boldsymbol{AX}_3 =$	[1.722222	2.555556	4.222222]′	=	4.222222[0.407895	0.605263	1]′	$= c_4\boldsymbol{X}_4$
$\boldsymbol{AX}_4 =$	[1.657895	2.473684	4.105263]′	=	4.105263[0.403846	0.602564	1]′	$= c_5\boldsymbol{X}_5$
$\boldsymbol{AX}_5 =$	[1.628205	2.435897	4.051282]′	=	4.051282[0.401899	0.601266	1]′	$= c_6\boldsymbol{X}_6$
$\boldsymbol{AX}_6 =$	[1.613924	2.417722	4.025316]′	=	4.025316[0.400943	0.600629	1]′	$= c_7\boldsymbol{X}_7$
$\boldsymbol{AX}_7 =$	[1.606918	2.408805	4.012579]′	=	4.012579[0.400470	0.600313	1]′	$= c_8\boldsymbol{X}_8$
$\boldsymbol{AX}_8 =$	[1.603448	2.404389	4.006270]′	=	4.006270[0.400235	0.600156	1]′	$= c_9\boldsymbol{X}_9$
$\boldsymbol{AX}_9 =$	[1.601721	2.402191	4.003130]′	=	4.003130[0.400117	0.600078	1]′	$= c_{10}\boldsymbol{X}_{10}$
$\boldsymbol{AX}_{10} =$	[1.600860	2.401095	4.001564]′	=	4.001564[0.400059	0.600039	1]′	$= c_{11}\boldsymbol{X}_{11}$

Iteration generates the sequence $\{\boldsymbol{X}_k\}$ (where $\boldsymbol{X}_k$ is a normalized vector):

$$12\begin{bmatrix}\frac{1}{2}\\ \frac{2}{3}\\ 1\end{bmatrix},\ \frac{16}{3}\begin{bmatrix}\frac{7}{16}\\ \frac{5}{8}\\ 1\end{bmatrix},\ \frac{9}{2}\begin{bmatrix}\frac{5}{12}\\ \frac{11}{18}\\ 1\end{bmatrix},\ \frac{38}{9}\begin{bmatrix}\frac{31}{76}\\ \frac{23}{38}\\ 1\end{bmatrix},\ \frac{78}{19}\begin{bmatrix}\frac{21}{52}\\ \frac{47}{78}\\ 1\end{bmatrix},\ \frac{158}{39}\begin{bmatrix}\frac{127}{316}\\ \frac{95}{158}\\ 1\end{bmatrix},\ \cdots$$

The sequence of vectors converges to $\boldsymbol{V} = \left[\frac{2}{3} \quad \frac{3}{5} \quad 1\right]'$, and the sequence of constants converges to $\lambda = 4$ (see Table 11.1). It can be proved that the rate of convergence is linear. ■

Theorem 11.18 (Power Method). Assume that the $n \times n$ matrix $\boldsymbol{A}$ has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and that they are ordered in decreasing magnitude; that is,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|. \tag{4}$$

If $\boldsymbol{X}_0$ is chosen appropriately, then the sequences $\left\{\boldsymbol{X}_k = \left[x_1^{(k)} \quad x_2^{(k)} \quad \ldots \quad x_n^{(k)}\right]'\right\}$ and $\{c_k\}$ generated recursively by

$$\boldsymbol{Y}_k = \boldsymbol{AX}_k \tag{5}$$

and

$$\boldsymbol{X}_{k+1} = \frac{1}{c_{k+1}}\boldsymbol{Y}_k, \tag{6}$$

where

$$c_{k+1} = x_j^{(k)} \quad \text{and} \quad x_j^{(k)} = \max_{1\leq i\leq n}\{|x_i^{(k)}|\}, \tag{7}$$

will converge to the dominant eigenvector $\boldsymbol{V}_1$ and eigenvalue λ_1, respectively. That is,

$$\lim_{k\to\infty} \boldsymbol{X}_k = \boldsymbol{V}_1 \quad \text{and} \quad \lim_{k\to\infty} c_k = \lambda_1. \tag{8}$$

Proof. Since $\boldsymbol{A}$ has n eigenvalues, there are n corresponding eigenvectors $\boldsymbol{V}_j$, for $j = 1, 2, \ldots, n$, that are linearly independent, normalized, and form a basis for n-dimensional space. Hence the starting vector $\boldsymbol{X}_0$ can be expressed as the linear combination

$$\boldsymbol{X}_0 = b_1\boldsymbol{V}_1 + b_2\boldsymbol{V}_2 + \cdots + b_n\boldsymbol{V}_n. \tag{9}$$

Assume that $\boldsymbol{X}_0 = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}'$ was chosen in such a manner that $b_1 \neq 0$. Also, assume that the coordinates of $\boldsymbol{X}_0$ are scaled so that $\max_{1\le j\le n}\{|x_j|\} = 1$. Because $\{\boldsymbol{V}_j\}_{j=1}^n$ are eigenvectors of $\boldsymbol{A}$, the multiplication $\boldsymbol{A}\boldsymbol{X}_0$, followed by normalization, produces

$$\begin{aligned}
\boldsymbol{Y}_0 = \boldsymbol{A}\boldsymbol{X}_0 &= \boldsymbol{A}(b_1\boldsymbol{V}_1 + b_2\boldsymbol{V}_2 + \cdots + b_n\boldsymbol{V}_n) \\
&= b_1\boldsymbol{A}\boldsymbol{V}_1 + b_2\boldsymbol{A}\boldsymbol{V}_2 + \cdots + b_n\boldsymbol{A}\boldsymbol{V}_n \\
&= b_1\lambda_1\boldsymbol{V}_1 + b_2\lambda_2\boldsymbol{V}_2 + \cdots + b_n\lambda_n\boldsymbol{V}_n \\
&= \lambda_1\left(b_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)\boldsymbol{V}_n\right)
\end{aligned} \tag{10}$$

and

$$\boldsymbol{X}_1 = \frac{\lambda_1}{c_1}\left(b_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)\boldsymbol{V}_n\right).$$

After k iterations we arrive at

(11)

$$\begin{aligned}
\boldsymbol{Y}_{k-1} &= \boldsymbol{A}\boldsymbol{X}_{k-1} \\
&= \boldsymbol{A}\frac{\lambda_1^{k-1}}{c_1c_2\cdots c_{k-1}}\left(b_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\boldsymbol{V}_n\right) \\
&= \frac{\lambda_1^{k-1}}{c_1c_2\cdots c_{k-1}}\left(b_1\boldsymbol{A}\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}\boldsymbol{A}\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\boldsymbol{A}\boldsymbol{V}_n\right) \\
&= \frac{\lambda_1^{k-1}}{c_1c_2\cdots c_{k-1}}\left(b_1\lambda_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}\lambda_2\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\lambda_n\boldsymbol{V}_n\right) \\
&= \frac{\lambda_1^{k}}{c_1c_2\cdots c_{k-1}}\left(b_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k}\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k}\boldsymbol{V}_n\right)
\end{aligned}$$

and

$$\boldsymbol{X}_k = \frac{\lambda_1^{k}}{c_1c_2\cdots c_k}\left(b_1\boldsymbol{V}_1 + b_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}\boldsymbol{V}_2 + \cdots + b_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\boldsymbol{V}_n\right).$$

Since we assumed that $|\lambda_j|/|\lambda_1| < 1$ for each $j = 2, 3, \ldots, n$, we have

$$\lim_{k\to\infty} b_j \left(\frac{\lambda_j}{\lambda_1}\right)^k \mathbf{V}_j = \mathbf{0} \quad \text{for each} \quad j = 2,\ 3,\ \ldots,\ n. \tag{12}$$

Hence it follows that

$$\lim_{k\to\infty} \mathbf{X}_k = \lim_{k\to\infty} \frac{b_1\lambda_1^k}{c_1c_2\cdots c_k}\mathbf{V}_1. \tag{13}$$

We have required that both $\mathbf{X}_k$ and $\mathbf{V}_1$ be normalized and their largest component be 1. Hence the limiting vector on the left side of (13) will be normalized, with its largest component being 1. Consequently, the limit of the scalar multiple of $\mathbf{V}_1$ on the right side of (13) exists and its value must be 1; that is,

$$\lim_{k\to\infty} \frac{b_1\lambda_1^k}{c_1c_2\cdots c_k} = 1. \tag{14}$$

Therefore, the sequence of vectors $\{\mathbf{X}_k\}$ converges to the dominant eigenvector:

$$\lim_{k\to\infty} \mathbf{X}_k = \mathbf{V}_1. \tag{15}$$

Replacing k with $k - 1$ in the terms of the sequence in (14) yields

$$\lim_{k\to\infty} \frac{b_1\lambda_1^{k-1}}{c_1c_2\cdots c_{k-1}} = 1,$$

and dividing both sides of this result into (14) yields

$$\lim_{k\to\infty} \frac{\lambda_1}{c_k} = \lim_{k\to\infty} \frac{b_1\lambda_1^k/(c_1c_2\cdots c_k)}{b_1\lambda_1^{k-1}/(c_1c_2\cdots c_{k-1})} = \frac{1}{1} = 1.$$

Therefore, the sequence of constants $\{c_k\}$ converges to the dominant eigenvalue:

$$\lim_{k\to\infty} c_k = \lambda_1, \tag{16}$$

and the proof of the theorem is complete. •

Speed of Convergence

In the light of equation (12) we see that the coefficient of $\mathbf{V}_j$ in $\mathbf{X}_k$ goes to zero in proportion to $(\lambda_j/\lambda_1)^k$ and that the speed of convergence of $\{\mathbf{X}_k\}$ to $\mathbf{V}_1$ is governed by the terms $(\lambda_2/\lambda_1)^k$. Consequently, the rate of convergence is linear. Similarly, the

Table 11.2 Comparison of the Rate of Convergence of the Power Method and Acceleration of the Power Method Using Aitken's $\mathbf{\Delta}^2$ Technique

	$c_k \boldsymbol{Y}_k$			$\widehat{c}_k \widehat{\boldsymbol{X}}_k$			
$c_1\boldsymbol{X}_1$	$= 12.000000[0.5000000$	0.6666667	$1]'$;	$4.3809524[0.4062500$	0.6041667	$1]' =$	$\widehat{c}_1\widehat{\boldsymbol{X}}_1$
$c_2\boldsymbol{X}_2$	$= 5.3333333[0.4375000$	0.6250000	$1]'$;	$4.0833333[0.4015152$	0.6010101	$1]' =$	$\widehat{c}_2\widehat{\boldsymbol{X}}_2$
$c_3\boldsymbol{X}_3$	$= 4.5000000[0.4166667$	0.6111111	$1]'$;	$4.0202020[0.4003759$	0.6002506	$1]' =$	$\widehat{c}_3\widehat{\boldsymbol{X}}_3$
$c_4\boldsymbol{X}_4$	$= 4.2222222[0.4078947$	0.6052632	$1]'$;	$4.0050125[0.4000938$	0.6000625	$1]' =$	$\widehat{c}_4\widehat{\boldsymbol{X}}_4$
$c_5\boldsymbol{X}_5$	$= 4.1052632[0.4038462$	0.6025641	$1]'$;	$4.0012508[0.4000234$	0.6000156	$1]' =$	$\widehat{c}_5\widehat{\boldsymbol{X}}_5$
$c_6\boldsymbol{X}_6$	$= 4.0512821[0.4018987$	0.6012658	$1]'$;	$4.0003125[0.4000059$	0.6000039	$1]' =$	$\widehat{c}_6\widehat{\boldsymbol{X}}_6$
$c_7\boldsymbol{X}_7$	$= 4.0253165[0.4009434$	0.6006289	$1]'$;	$4.0000781[0.4000015$	0.6000010	$1]' =$	$\widehat{c}_7\widehat{\boldsymbol{X}}_7$
$c_8\boldsymbol{X}_8$	$= 4.0125786[0.4004702$	0.6003135	$1]'$;	$4.0000195[0.4000004$	0.6000002	$1]' =$	$\widehat{c}_8\widehat{\boldsymbol{X}}_8$
$c_9\boldsymbol{X}_9$	$= 4.0062696[0.4002347$	0.6001565	$1]'$;	$4.0000049[0.4000001$	0.6000001	$1]' =$	$\widehat{c}_9\widehat{\boldsymbol{X}}_9$
$c_{10}\boldsymbol{X}_{10}$	$= 4.0031299[0.4001173$	0.6000782	$1]'$;	$4.0000012[0.4000000$	0.6000000	$1]' =$	$\widehat{c}_{10}\widehat{\boldsymbol{X}}_{10}$

convergence of the sequence of constants $\{c_k\}$ to λ_1 is linear. The Aitken $\mathbf{\Delta}^2$ method can be used for any linearly convergent sequence $\{p_k\}$ to form a new sequence,

$$\left\{ \widehat{p}_k = \frac{(p_{k+1} - p_k)^2}{p_{k+2} - 2p_{k+1} + p_k} \right\},$$

that converges faster. In Example 11.4 this Aitken $\mathbf{\Delta}^2$ method can be applied to speed up convergence of the sequence of constants $\{c_k\}$, as well as the first two components of the sequence of vectors $\{\boldsymbol{X}_k\}$. A comparison of the results obtained with this technique and the original sequences is shown in Table 11.2.

Shifted-Inverse Power Method

We will now discuss the shifted inverse power method. It requires a good starting approximation for an eigenvalue, and then iteration is used to obtain a precise solution. Other procedures such as the QM and Givens' method are used first to obtain the starting approximations. Cases involving complex eigenvalues, multiple eigenvalues, or the presence of two eigenvalues with the same magnitude or approximately the same magnitude will cause computational difficulties and require more advanced methods. Our illustrations will focus on the case where the eigenvalues are distinct. The shifted inverse power method is based on the following three results (the proofs are left as exercises).

Theorem 11.19 (Shifting Eigenvalues). Suppose that λ, $\boldsymbol{V}$ is an eigenpair of $\boldsymbol{A}$. If α is any constant, then $\lambda - \alpha$, $\boldsymbol{V}$ is an eigenpair of the matrix $\boldsymbol{A} - \alpha\boldsymbol{I}$.

Theorem 11.20 (Inverse Eigenvalues). Suppose that λ, $\boldsymbol{V}$ is an eigenpair of $\boldsymbol{A}$. If $\lambda \neq 0$, then $1/\lambda$, $\boldsymbol{V}$ is an eigenpair of the matrix $\boldsymbol{A}^{-1}$.

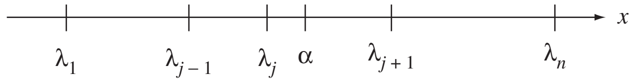

Figure 11.2 The location of α for the shifted-inverse power method.

Theorem 11.21. Suppose that λ, $\boldsymbol{V}$ is an eigenpair of $\boldsymbol{A}$. If $\alpha \neq \lambda$, then $1/(\lambda - \alpha)$, $\boldsymbol{V}$ is an eigenpair of the matrix $(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}$.

Theorem 11.22 (Shifted-Inverse Power Method). Assume that the $n \times n$ matrix $\boldsymbol{A}$ has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and consider the eigenvalue λ_j. Then a constant α can be chosen so that $\mu_1 = 1/(\lambda_j - \alpha)$ is the dominant eigenvalue of $(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}$. Furthermore, if $\boldsymbol{X}_0$ is chosen appropriately, then the sequences $\{\boldsymbol{X}_k = [x_1^{(k)} \; x_2^{(k)} \; \ldots \; x_n^{(k)}]'\}$ and $\{c_k\}$ are generated recursively by

$$\boldsymbol{Y}_k = (\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_k \tag{17}$$

and

$$\boldsymbol{X}_{k+1} = \frac{1}{c_{k+1}}\boldsymbol{Y}_k, \tag{18}$$

where

$$c_{k+1} = x_j^{(k)} \quad \text{and} \quad x_j^{(k)} = \max_{1 \le j \le n}\{|x_i^{(k)}|\} \tag{19}$$

will converge to the dominant eigenpair μ_1, $\boldsymbol{V}_j$ of the matrix $(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}$. Finally, the corresponding eigenvalue for the matrix $\boldsymbol{A}$ is given by the calculation

$$\lambda_j = \frac{1}{\mu_1} + \alpha. \tag{20}$$

Remark. For practical implementations of Theorem 11.22, a linear system solver is used to compute $\boldsymbol{Y}_k$ in each step by solving the linear system $(\boldsymbol{A} - \alpha\boldsymbol{I})\boldsymbol{Y}_k = \boldsymbol{X}_k$.

Proof. Without loss of generality, we may assume that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. Select a number α ($\alpha \neq \lambda_j$) that is closer to λ_j than any of the other eigenvalues (see Figure 11.2), that is,

$$|\lambda_j - \alpha| < |\lambda_i - \alpha| \quad \text{for each } i = 1, 2, \ldots, j-1, j+1, \ldots, n. \tag{21}$$

According to Theorem 11.21, $1/(\lambda_j - \alpha)$, $\boldsymbol{V}$ is an eigenpair of the matrix $(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}$. Relation (21) implies that $1/|\lambda_i - \alpha| < 1/|\lambda_j - \alpha|$ for each $i \neq j$ so that $\mu_1 = 1/(\lambda_j - \alpha)$ is the dominant eigenvalue of the matrix $(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}$. The shifted-inverse power method uses a modification of the power method to determine the eigenpair μ_1, $\boldsymbol{V}_j$. Then the calculation $\lambda_j = 1/\mu_1 + \alpha$ produces the desired eigenvalue of the matrix $\boldsymbol{A}$. •

Table 11.3 Shifted-Inverse Power Method for the Matrix $(\boldsymbol{A} - 4.2\boldsymbol{I})^{-1}$ in Example 11.6: Convergence to the Eigenvector $\boldsymbol{V} = \left[\frac{2}{5} \;\; \frac{3}{5} \;\; 1\right]'$ and $\mu_1 = -5$

$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_k$ =	$c_{k+1}\boldsymbol{X}_{k+1}$				
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_0$	$= -23.18181818$	[0.4117647059	0.6078431373	1]′	$= c_1\boldsymbol{X}_1$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_1$	$= -5.356506239$	[0.4009983361	0.6006655574	1]′	$= c_2\boldsymbol{X}_2$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_2$	$= -5.030252609$	[0.4000902120	0.6000601413	1]′	$= c_3\boldsymbol{X}_3$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_3$	$= -5.002733697$	[0.4000081966	0.6000054644	1]′	$= c_4\boldsymbol{X}_4$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_4$	$= -5.000248382$	[0.4000007451	0.6000004967	1]′	$= c_5\boldsymbol{X}_5$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_5$	$= -5.000022579$	[0.4000000677	0.6000000452	1]′	$= c_6\boldsymbol{X}_6$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_6$	$= -5.000002053$	[0.4000000062	0.6000000041	1]′	$= c_7\boldsymbol{X}_7$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_7$	$= -5.000000187$	[0.4000000006	0.6000000004	1]′	$= c_8\boldsymbol{X}_8$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_8$	$= -5.000000017$	[0.4000000001	0.6000000000	1]′	$= c_9\boldsymbol{X}_9$

Example 11.6. Employ the shifted-inverse power method to find the eigenpairs of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 11 & -5 \\ -2 & 17 & -7 \\ -4 & 26 & -10 \end{bmatrix}.$$

Use the fact that the eigenvalues of $\boldsymbol{A}$ are $\lambda_1 = 4$, $\lambda_2 = 2$, and $\lambda_3 = 1$, and select an appropriate α and starting vector for each case.

Case (i): For the eigenvalue $\lambda_1 = 4$, we select $\alpha = 4.2$ and the starting vector $\boldsymbol{X}_0 = \begin{bmatrix}1 & 1 & 1\end{bmatrix}'$. First, form the matrix $\boldsymbol{A} - 4.2\boldsymbol{I}$, compute the solution to

$$\begin{bmatrix} -4.2 & 11 & -5 \\ -2 & 12.8 & -7 \\ -4 & 26 & -14.2 \end{bmatrix} \boldsymbol{Y}_0 = \boldsymbol{X}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

and get the vector $\boldsymbol{Y}_0 = \begin{bmatrix}-9.545454545 & -14.09090909 & -23.18181818\end{bmatrix}'$. Then compute $c_1 = -23.18181818$ and $\boldsymbol{X}_1 = \begin{bmatrix}0.4117647059 & 0.6078431373 & 1\end{bmatrix}'$. Iteration generates the values given in Table 11.3. The sequence $\{c_k\}$ converges to $\mu_1 = -5$, which is the dominant eigenvalue of $(\boldsymbol{A} - 4.2\boldsymbol{I})^{-1}$, and $\{\boldsymbol{X}_k\}$ converges to $\boldsymbol{V}_1 = \left[\frac{2}{5} \;\; \frac{3}{5} \;\; 1\right]'$. The eigenvalue λ_1 of $\boldsymbol{A}$ is given by the computation $\lambda_1 = 1/\mu_1 + \alpha = 1/(-5) + 4.2 = -0.2 + 4.2 = 4$.

Case (ii): For the eigenvalue $\lambda_2 = 2$, we select $\alpha = 2.1$ and the starting vector $\boldsymbol{X}_0 = \begin{bmatrix}1 & 1 & 1\end{bmatrix}'$. Form the matrix $\boldsymbol{A} - 2.1\boldsymbol{I}$, compute the solution to

$$\begin{bmatrix} -2.1 & 11 & -5 \\ -2 & 14.9 & -7 \\ -4 & 26 & -12.1 \end{bmatrix} \boldsymbol{Y}_0 = \boldsymbol{X}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

and obtain the vector $\boldsymbol{Y}_0 = \begin{bmatrix}11.05263158 & 21.57894737 & 42.63157895\end{bmatrix}'$. Then $c_1 = 42.63157895$ and vector $\boldsymbol{X}_1 = \begin{bmatrix}0.2592592593 & 0.5061728395 & 1\end{bmatrix}'$. Iteration produces the

Table 11.4 Shifted-Inverse Power Method for the Matrix $(\boldsymbol{A} - 2.1\boldsymbol{I})^{-1}$ in Example 11.6: Convergence to the Dominant Eigenvector $\boldsymbol{V} = \left[\frac{1}{4} \;\; \frac{1}{2} \;\; 1\right]$ and $\mu_1 = -10$

$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_k =$	$c_{k+1}\boldsymbol{X}_{k+1}$				
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_0 =$	42.63157895	[0.2592592593	0.5061728395	1]′	$= c_1\boldsymbol{X}_1$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_1 =$	−9.350227420	[0.2494788047	0.4996525365	1]′	$= c_2\boldsymbol{X}_2$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_2 =$	−10.03657511	[0.2500273314	0.5000182209	1]′	$= c_3\boldsymbol{X}_3$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_3 =$	−9.998082009	[0.2499985612	0.4999990408	1]′	$= c_4\boldsymbol{X}_4$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_4 =$	−10.00010097	[0.2500000757	0.5000000505	1]′	$= c_5\boldsymbol{X}_5$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_5 =$	−9.999994686	[0.2499999960	0.4999999973	1]′	$= c_6\boldsymbol{X}_6$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_6 =$	−10.00000028	[0.2500000002	0.5000000001	1]′	$= c_7\boldsymbol{X}_7$

Table 11.5 Shifted-Inverse Power Method for the Matrix $(\boldsymbol{A} - 0.875\boldsymbol{I})^{-1}$ in Example 11.6: Convergence to the Dominant Eigenvector $\boldsymbol{V} = \left[\frac{1}{2} \;\; \frac{1}{2} \;\; 1\right]'$ and $\mu_1 = 8$

$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_k =$	$c_{k+1}\boldsymbol{X}_{k+1}$				
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_0 =$	−30.40000000	[0.5052631579	0.4947368421	1]′	$= c_1\boldsymbol{X}_1$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_1 =$	8.404210526	[0.5002004008	0.4997995992	1]′	$= c_2\boldsymbol{X}_2$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_2 =$	8.015390782	[0.5000080006	0.4999919994	1]′	$= c_3\boldsymbol{X}_3$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_3 =$	8.000614449	[0.5000003200	0.4999996800	1]′	$= c_4\boldsymbol{X}_4$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_4 =$	8.000024576	[0.5000000128	0.4999999872	1]′	$= c_5\boldsymbol{X}_5$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_5 =$	8.000000983	[0.5000000005	0.4999999995	1]′	$= c_6\boldsymbol{X}_6$
$(\boldsymbol{A} - \alpha\boldsymbol{I})^{-1}\boldsymbol{X}_6 =$	8.000000039	[0.5000000000	0.5000000000	1]′	$= c_7\boldsymbol{X}_7$

values given in Table 11.4. The dominant eigenvalue of $(\boldsymbol{A} - 2.1\boldsymbol{I})^{-1}$ is $\mu_1 = -10$, and the eigenpair of the matrix $\boldsymbol{A}$ is $\lambda_2 = 1/(-10) + 2.1 = -0.1 + 2.1 = 2$ and $\boldsymbol{V}_2 = \left[\frac{1}{4} \;\; \frac{1}{2} \;\; 1\right]'$.

Case (iii): For the eigenvalue $\lambda_3 = 1$, we select $\alpha = 0.875$ and the starting vector $\boldsymbol{X}_0 = \left[0 \;\; 1 \;\; 1\right]'$. Iteration produces the values given in Table 11.5. The dominant eigenvalue of $(\boldsymbol{A} - 0.875\boldsymbol{I})^{-1}$ is $\mu_1 = 8$, and the eigenpair of matrix $\boldsymbol{A}$ is $\lambda_3 = 1/8 + 0.875 = 0.125 + 0.875 = 1$ and $\boldsymbol{V}_3 = \left[\frac{1}{2} \;\; \frac{1}{2} \;\; 1\right]'$. The sequence $\{\boldsymbol{X}_k\}$ of vectors with the starting vector $\left[0 \;\; 1 \;\; 1\right]'$ converged in seven iterations. (Computational difficulties were encountered when $\boldsymbol{X}_0 = \left[1 \;\; 1 \;\; 1\right]'$ was used, and convergence took significantly longer.) ■

Numerical Methods Using Matlab, 4th Edition, 2004
John H. Mathews and Kurtis K. Fink
ISBN: 0-13-065248-2

Prentice-Hall Inc.
Upper Saddle River, New Jersey, USA
http://vig.prenhall.com/