than one maximal component, let $m(\mathbf{v})$ be the *first* maximal component—e.g., $m(1,3,-2)=3$, and $m(-3,3,-2)=-3$. It's clear that $m(\alpha\mathbf{v})=\alpha m(\mathbf{v})$ for all scalars α. Suppose $m(\mathbf{A}^n\mathbf{x}_0/\lambda_1^n)\to\gamma$. Since $(\mathbf{A}^n/\lambda_1^n)\to\mathbf{G}_1$, we see that

$$\lim_{n\to\infty}\frac{\mathbf{A}^n\mathbf{x}_0}{m(\mathbf{A}^n\mathbf{x}_0)}=\lim_{n\to\infty}\frac{(\mathbf{A}^n/\lambda_1^n)\mathbf{x}_0}{m(\mathbf{A}^n\mathbf{x}_0/\lambda_1^n)}=\frac{\mathbf{G}_1\mathbf{x}_0}{\gamma}=\mathbf{x}$$

is an eigenvector associated with λ_1. But rather than successively powering $\mathbf{A}$, the sequence $\mathbf{A}^n\mathbf{x}_0/m(\mathbf{A}^n\mathbf{x}_0)$ is more efficiently generated by starting with $\mathbf{x}_0\notin R(\mathbf{A}-\lambda_1\mathbf{I})$ and setting

$$\mathbf{y}_n=\mathbf{A}\mathbf{x}_n,\quad \nu_n=m(\mathbf{y}_n),\quad \mathbf{x}_{n+1}=\frac{\mathbf{y}_n}{\nu_n},\quad \text{for } n=0,1,2,\ldots. \tag{7.3.17}$$

Not only does $\mathbf{x}_n\to\mathbf{x}$, but as a bonus we get $\nu_n\to\lambda_1$ because for all n, $\mathbf{A}\mathbf{x}_{n+1}=\mathbf{A}^2\mathbf{x}_n/\nu_n$, so if $\nu_n\to\nu$ as $n\to\infty$, the limit on the left-hand side is $\mathbf{A}\mathbf{x}=\lambda_1\mathbf{x}$, while the limit on the right-hand side is $\mathbf{A}^2\mathbf{x}/\nu=\lambda_1^2\mathbf{x}/\nu$. Since these two limits must agree, $\lambda_1\mathbf{x}=(\lambda_1^2/\nu)\mathbf{x}$, and this implies $\nu=\lambda_1$.

Summary. The sequence $(\nu_n,\mathbf{x}_n)$ defined by (7.3.17) converges to an eigenpair $(\lambda_1,\mathbf{x})$ for $\mathbf{A}$ provided that $\mathbf{G}_1\mathbf{x}_0\neq\mathbf{0}$ or, equivalently, $\mathbf{x}_0\notin R(\mathbf{A}-\lambda_1\mathbf{I})$.

- ▷ **Advantages.** Each iteration requires only one matrix–vector product, and this can be exploited to reduce the computational effort when $\mathbf{A}$ is large and sparse—assuming that a dominant eigenpair is the only one of interest.
- ▷ **Disadvantages.** Only a dominant eigenpair is determined—something else must be done if others are desired. Furthermore, it's clear from (7.3.16) that the rate at which (7.3.17) converges depends on how fast $(\lambda_2/\lambda_1)^n\to 0$, so convergence is slow when $|\lambda_1|$ is close to $|\lambda_2|$.

Example 7.3.8

Inverse Power Method. Given a real approximation $\alpha\notin\sigma(\mathbf{A})$ to any real $\lambda\in\sigma(\mathbf{A})$, this algorithm (also called the ***inverse iteration***) determines an eigenpair $(\lambda,\mathbf{x})$ for a diagonalizable matrix $\mathbf{A}\in\Re^{m\times m}$ by applying the power method[75] to $\mathbf{B}=(\mathbf{A}-\alpha\mathbf{I})^{-1}$. Recall from Exercise 7.1.9 that

$$\begin{gathered}\mathbf{x}\text{ is an eigenvector for }\mathbf{A}\Longleftrightarrow\mathbf{x}\text{ is an eigenvector for }\mathbf{B},\\ \lambda\in\sigma(\mathbf{A})\Longleftrightarrow(\lambda-\alpha)^{-1}\in\sigma(\mathbf{B}).\end{gathered} \tag{7.3.18}$$

If $|\lambda-\alpha|<|\lambda_i-\alpha|$ for all other $\lambda_i\in\sigma(\mathbf{A})$, then $(\lambda-\alpha)^{-1}$ is the dominant eigenvalue of $\mathbf{B}$ because $|\lambda-\alpha|^{-1}>|\lambda_i-\alpha|^{-1}$. Therefore, applying the power

[75] The relation between the power method and inverse iteration is clear to us now, but it originally took 15 years to make the connection. Inverse iteration was not introduced until 1944 by the German mathematician Helmut Wielandt (1910–2001).