

5.3. Bounds for eigenvalues. For some iterative numerical methods for finding eigenvalues, one needs to have an initial guess for the eigenvalue. In that case, the following result can sometimes be useful.

Theorem 18. (*Gershgorin*) Let $A = (a_{ij})$ and let C_i denote circles with centers a_{ii} and radii $r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$. Let $D = \cup_{i=1}^n C_i$. Then all the eigenvalues of A lie in the set D .

Proof. If λ is an eigenvalue of A , then there exists an eigenvector $x \neq 0$ such that $Ax = \lambda x$, i.e., $\sum_{j=1}^n a_{ij}x_j = \lambda x_i$, $i = 1, \dots, n$, which we rewrite as

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j, \quad i = 1, \dots, n.$$

Let x_k be the largest component of x in absolute value (so $x_k \neq 0$). Then applying the above with $i = k$, we have

$$|\lambda - a_{kk}||x_k| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}||x_j|.$$

Since $|x_j| \leq |x_k|$ for $j \neq k$, we get

$$|\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}| = r_k.$$

Hence, $\lambda \in C_k \subset D$. Since λ is an arbitrary eigenvalue, all the eigenvalues of A lie in D . \square

Corollary 4. The spectral radius of A , $\rho(A) \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

Proof. For any eigenvalue λ , there is some k such that

$$|\lambda| - |a_{kk}| \leq |\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}|,$$

and so

$$|\lambda| \leq \sum_{j=1}^n |a_{kj}| \leq \max_i \sum_{j=1}^n |a_{ij}|.$$

\square

One can also establish the following more precise version of Gershgorin's theorem.

Theorem 19. If k Gershgorin circles of the matrix A are disjoint from the others, then exactly k eigenvalues of A lie in the union of the k circles.

Example: Let

$$A = \begin{pmatrix} 1 & 10^{-4} & 10^{-4} \\ 10^{-4} & 1 & 10^{-4} \\ 10^{-4} & 10^{-4} & 2 \end{pmatrix}.$$

Then

$$C_1 = C_2 = \{\lambda : |\lambda - 1| \leq 2 \times 10^{-4}\}, \quad C_2 = \{\lambda : |\lambda - 2| \leq 2 \times 10^{-4}\}.$$

Hence two eigenvalues lie in C_1 and one eigenvalue lies in C_2 .

5.4. Power method. We begin the study of numerical methods for finding eigenvalues and eigenvectors of a matrix by considering the *power method*, which can be used to find approximations to the dominant eigenvalue and corresponding eigenvector of a matrix A . We will assume that A is real, has a complete set of eigenvectors, and that the eigenvalues of A satisfy:

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

In particular, λ_1 is real, since if it was complex, $\bar{\lambda}_1$ would also be an eigenvalue and since $|\lambda_1| = |\bar{\lambda}_1|$, our hypothesis would be violated.

Let x^1, \dots, x^n denote a set of linearly independent eigenvectors of A , with x^1 the eigenvector corresponding to λ_1 . Let q_0 denote an initial approximation to x^1 , where we assume that $\|x^1\| = \|q_0\| = 1$. We may choose any norm, but for simplicity, we take $\|x^1\|_\infty = \|q_0\|_\infty = 1$. The power method then defines a sequence $\{q_k\}$ by

$$q_{k+1} = Aq_k / \sigma_k, \quad \text{where } \sigma_k = \|Aq_k\|_\infty.$$

Hence, $\|q_{k+1}\|_\infty = 1$. We then have the following result.

Theorem 20. *Suppose that $q_0 = \gamma_1 x^1 + \gamma_2 x^2 + \cdots + \gamma_n x^n$, where $\gamma_1 \neq 0$. Then*

$$\lim_{k \rightarrow \infty} q_k \operatorname{sgn}(\lambda_1^k) = (\operatorname{sgn} \gamma_1) x_1,$$

i.e., $\{q_k\}$ converges to a unit eigenvector associated to λ_1 .

Although x^1 is not known, a random choice of q_0 will usually have the desired property.

Proof. Using the expansion for q_0 , we have

$$\begin{aligned} q_1 &= [\gamma_1 \lambda_1 x^1 + \gamma_2 \lambda_2 x^2 + \cdots + \gamma_n \lambda_n x^n] / \sigma_0, \\ q_2 &= [\gamma_1 \lambda_1^2 x^1 + \gamma_2 \lambda_2^2 x^2 + \cdots + \gamma_n \lambda_n^2 x^n] / [\sigma_0 \sigma_1], \\ &\dots = \dots \\ q_k &= [\gamma_1 \lambda_1^k x^1 + \gamma_2 \lambda_2^k x^2 + \cdots + \gamma_n \lambda_n^k x^n] / [\sigma_0 \sigma_1 \cdots \sigma_{k-1}], \\ &= \lambda_1^k [\gamma_1 x^1 + \gamma_2 (\lambda_2 / \lambda_1)^k x^2 + \cdots + \gamma_n (\lambda_n / \lambda_1)^k x^n] / [\sigma_0 \sigma_1 \cdots \sigma_{k-1}]. \end{aligned}$$

Since $\|q_k\|_\infty = 1$ for all k , $|\lambda_j / \lambda_1| < 1$ for $j = 2, \dots, n$, and $\sigma_i > 0$ for all i ,

$$1 = \lim_{k \rightarrow \infty} \frac{|\lambda_1^k|}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} \|\gamma_1 x^1 + \gamma_2 (\frac{\lambda_2}{\lambda_1})^k x^2 + \cdots + \gamma_n (\frac{\lambda_n}{\lambda_1})^k x^n\|_\infty = |\gamma_1| \lim_{k \rightarrow \infty} \frac{|\lambda_1^k|}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}}.$$

Hence,

$$\lim_{k \rightarrow \infty} q_k \operatorname{sgn}(\lambda_1^k) = \lim_{k \rightarrow \infty} \frac{|\lambda_1^k|}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} [\gamma_1 x^1 + \gamma_2 (\frac{\lambda_2}{\lambda_1})^k x^2 + \cdots + \gamma_n (\frac{\lambda_n}{\lambda_1})^k x^n] = \frac{\gamma_1}{|\gamma_1|} x^1.$$

Thus, except for possible sign changes, $\{q_k\}$ converges to $\pm x^1$. □

The rate of convergence depends on the ratio $|\lambda_2/\lambda_1|$. To see this more precisely, we observe that

$$\begin{aligned} \left\| \frac{\sigma_0 \sigma_1 \cdots \sigma_{k-1}}{\gamma_1 \lambda_1^k} q_k - x^1 \right\|_\infty &= \left\| \sum_{j=2}^n (\gamma_j / \gamma_1) (\lambda_j / \lambda_1)^k x^j \right\|_\infty \\ &\leq \sum_{j=2}^n |\gamma_j / \gamma_1| |\lambda_j / \lambda_1|^k \leq |\lambda_2 / \lambda_1|^k \sum_{j=2}^n |\gamma_j / \gamma_1| \leq C |\lambda_2 / \lambda_1|^k, \end{aligned}$$

where we have assumed that the eigenvectors x^j are normalized so that $\|x^j\|_\infty = 1$.

There are a number of ways to obtain a sequence of approximate eigenvalues. One method is to suppose that u is any vector such that $u^T x^1 \neq 0$. Although x^1 is not known, a random choice of u will usually have this property. Now consider the sequence of eigenvalue approximations given by $\mu_k = u^T A q_k / (u^T q_k)$. Then

$$u^T A q_k = \sigma_k u^T q_{k+1} = \sigma_k \frac{\lambda_1^{k+1}}{\sigma_0 \sigma_1 \cdots \sigma_k} u^T \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^{k+1} x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^{k+1} x^n \right],$$

while

$$u^T q_k = \frac{\lambda_1^k}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} u^T \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x^n \right].$$

Hence,

$$\mu_k = \frac{u^T A q_k}{u^T q_k} = \lambda_1 \frac{u^T \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^{k+1} x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^{k+1} x^n \right]}{u^T \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x^n \right]},$$

and it easily follows that $\lim_{k \rightarrow \infty} \mu_k = \lambda_1$.

When the matrix A is symmetric, a better method is to consider the *Rayleigh quotient* sequence $\mu_k = q_k^T A q_k / q_k^T q_k$. Then, using the fact that the eigenvectors can be chosen to be orthogonal, we have

$$\begin{aligned} q_k^T A q_k &= \sigma_k q_k^T q_{k+1} \\ &= \sigma_k \frac{\lambda_1^k}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x^n \right]^T \\ &\quad \cdot \frac{\lambda_1^{k+1}}{\sigma_0 \sigma_1 \cdots \sigma_k} \left[\gamma_1 x^1 + \gamma_2 \left(\frac{\lambda_2}{\lambda_1} \right)^{k+1} x^2 + \cdots + \gamma_n \left(\frac{\lambda_n}{\lambda_1} \right)^{k+1} x^n \right] \\ &= \lambda_1 \left(\frac{\lambda_1^k}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} \right)^2 \left[\gamma_1^2 \|x^1\|_2^2 + \gamma_2^2 \left(\frac{\lambda_2}{\lambda_1} \right)^{2k+1} \|x^2\|_2^2 + \cdots + \gamma_n^2 \left(\frac{\lambda_n}{\lambda_1} \right)^{2k+1} \|x^n\|_2^2 \right], \end{aligned}$$

while

$$q_k^T q_k = \left(\frac{\lambda_1^k}{\sigma_0 \sigma_1 \cdots \sigma_{k-1}} \right)^2 \left[\gamma_1^2 \|x^1\|_2^2 + \gamma_2^2 \left(\frac{\lambda_2}{\lambda_1} \right)^{2k} \|x^2\|_2^2 + \cdots + \gamma_n^2 \left(\frac{\lambda_n}{\lambda_1} \right)^{2k} \|x^n\|_2^2 \right].$$

Hence, it easily follows that

$$\lim_{k \rightarrow \infty} \mu_k = \lambda_1$$

and the rate of convergence is proportional to $(\lambda_2/\lambda_1)^{2k}$, twice the rate of the eigenvector convergence, and an improvement over the nonsymmetric case.

5.5. Inverse power method. Using the ideas presented above, we next consider a method for finding an approximation to an eigenvector when a good approximation to an eigenvalue is known. The method uses the following result.

Lemma 3. *Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of the matrix A and assume that $\lambda \neq \lambda_i$ for $i = 1, \dots, n$. Then if x^i is an eigenvector of A corresponding to the eigenvalue λ_i , x^i is an eigenvector of $(\lambda I - A)^{-1}$ corresponding to the eigenvalue $(\lambda - \lambda_i)^{-1}$.*

Proof. Since $Ax^i = \lambda_i x^i$, $(\lambda I - A)x^i = (\lambda - \lambda_i)x^i$ and so $(\lambda I - A)^{-1}x^i = (\lambda - \lambda_i)^{-1}x^i$. \square

We now consider the power method applied to the matrix $B = (\lambda I - A)^{-1}$, i.e., $q_{k+1} = Bq_k/\sigma_k$, where $\sigma_k = \|Bq_k\|_\infty$. The resulting method is called the *inverse power method*. We know that this method converges to the eigenvector corresponding to the largest eigenvalue of B . But this is the eigenvalue of A that is closest to λ . Note that we may write this iteration in the form:

$$(\lambda I - A)\tilde{q}_{k+1} = q_k, \quad q_{k+1} = \tilde{q}_{k+1}/\|\tilde{q}_{k+1}\|_\infty.$$

Hence, at each step we need to solve a linear system of equations. However, since we always use the same matrix, only one LU decomposition is needed and then one backsolve per iteration. In fact, if λ is close to λ_i , the convergence is very rapid (a couple of iterations).

5.6. Rayleigh quotient iteration. For symmetric matrices, we can combine the use of the Rayleigh quotient and inverse power method to produce a fast converging method for the computation of an eigenvalue and the corresponding eigenvector. Given an approximate eigenvalue λ^0 and approximate eigenvector \mathbf{x}^0 , do the following steps for $n = 1, 2, \dots$ until convergence:

- (i) Compute a new approximate eigenvector \mathbf{x}^n by applying the inverse power method to the eigenvector \mathbf{x}^{n-1} using the approximate eigenvalue λ^{n-1} .
- (ii) Use the eigenvector \mathbf{x}^n to compute a new eigenvalue λ^n by the Rayleigh quotient.