2021 Spring MAS 365
Chapter 9: Approximating Eigenvalues

Donghwan Kim

KAIST

June/8,10, 2021

Eigenvalues

- The eigenvalues of an $n \times n$ matrix A, corresponds to the zeros of the characteristic polynomial

$$p(\lambda) = \det\{A - \lambda I\}\,.$$

- Finding the determinant and the roots of $p(\lambda)$ is computationally expensive and difficult.

- Sometimes knowing only the region of the complex plane in which the eigenvalues lie is good enough. (When?)

Geršgorin Circle

Theorem 1 (Geršgorin Circle)

Let A be an $n \times n$ matrix and R_i denote the circle in the complex plane with center a_{ii} and radius $\sum_{j=1, j\neq i}^{n} |a_{ij}|$; that is,

$$R_i = \left\{ z \in \mathcal{C} \;\middle|\; |z - a_{ii}| \leq \sum_{\substack{j=1,\\ j\neq i}}^{n} |a_{ij}| \right\}$$

where $\mathcal{C}$ denotes the complex plane. The eigenvalues of A are contained within the union of these circles, $R = \cup_{i=1}^{n} R_i$. Moreover, the union of any k of the circles that do not intersect the remaining $(n-k)$ contains precisely k (counting multiplicities) of the eigenvalues.

Geršgorin Circle (cont'd)

Proof Suppose that λ is an eigenvalue of A with associated eigenvector $\boldsymbol{x}$, where $||\boldsymbol{x}||_\infty = 1$. Since $A\boldsymbol{x} = \lambda\boldsymbol{x}$, the equivalent component representation is

$$\sum_{j=1}^{n} a_{ij} x_j = \lambda x_i, \quad \text{for each } i = 1, 2, \ldots, n.$$

Let k be an integer with $|x_k| = ||\boldsymbol{x}||_\infty = 1$. When $i = k$, we have

$$\sum_{j=1}^{n} a_{kj} x_j = \lambda x_k.$$

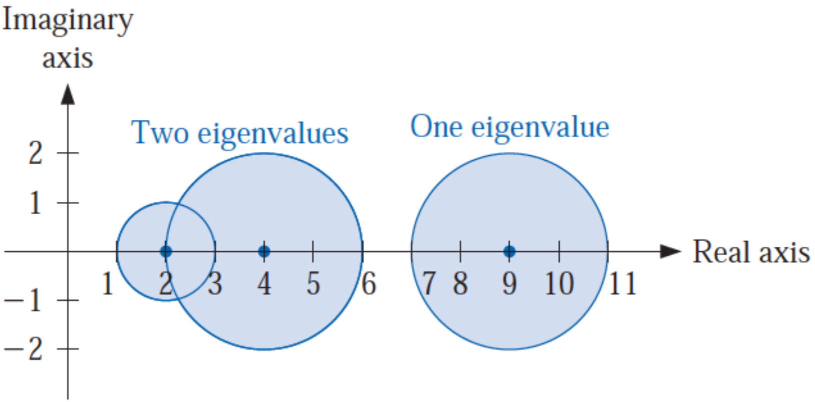
Geršgorin Circle (cont'd)

Ex. Determine the Geršgorin circles for the matrix

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$$

and use these to find bounds for the spectral radius of A.

Geršgorin Circle (cont'd)



- This technique is useful even when we need to find the eigenvalues. When?

The Power Method

- The **power method** is an iterative technique that determines the dominant eigenvalue of a matrix. A modified version (such as inverse power method and deflation methods) can also find other eigenvalues.

- We assume that the $n \times n$ matrix A has n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with an associated collection of linearly independent eigenvectors $\{\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \ldots, \boldsymbol{v}^{(n)}\}$. When they are linearly dependent, the power method is not guaranteed to work well.

- Moreover, we assume that λ_1 is largest in magnitude, so that

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \geq 0.$$

The Power Method (cont'd)

- For any vector $\boldsymbol{x}$ in $\mathbb{R}^n$, constants $\beta_1, \beta_2, \ldots, \beta_n$ exist with

$$\boldsymbol{x} = \sum_{j=1}^{n} \beta_j \boldsymbol{v}^{(j)}.$$

- Multiplying both sides of this equation by A^k gives

$$A^k \boldsymbol{x} = \sum_{j=1}^{n} \beta_j \lambda_j^k \boldsymbol{v}^{(j)} = \lambda_1^k \sum_{j=1}^{n} \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^k \boldsymbol{v}^{(j)}$$

- Since $|\lambda_1| > |\lambda_j|$ for $j = 2, 3, \ldots, n$, we have $\lim_{k\to\infty} (\lambda_j/\lambda_1)^k = 0$, and

$$\lim_{k\to\infty} A^k \boldsymbol{x} = \lim_{k\to\infty} \lambda_1^k \beta_1 \boldsymbol{v}^{(1)},$$

  which converges to 0 if $|\lambda_1| < 1$ and diverges if $|\lambda_1| > 1$, provided that $\beta_1 \neq 0$.

The Power Method (cont'd)

- We scale the powers of $A^k\boldsymbol{x}$ in an appropriate way to ensure that the limit is finite and nonzero.
- This begins by choosing $\boldsymbol{x}$ to be a unit vector $\boldsymbol{x}^{(0)}$ relative to $||\cdot||_\infty$ and choosing a component $x_{p_0}^{(0)}$ of $\boldsymbol{x}^{(0)}$ with
$$x_{p_0}^{(0)} = 1 = ||\boldsymbol{x}^{(0)}||_\infty.$$
- Let $\boldsymbol{y}^{(1)} = A\boldsymbol{x}^{(0)}$ and define $\mu^{(1)} = y_{p_0}^{(1)}$. Then

$$\mu^{(1)} = y_{p_0}^{(1)} = \frac{y_{p_0}^{(1)}}{x_{p_0}^{(0)}} =$$

The Power Method (cont'd)

- Let p_1 be the smallest integer such that

$$|y_{p_1}^{(1)}| = ||\boldsymbol{y}^{(1)}||_\infty$$

and define $\boldsymbol{x}^{(1)}$ by

$$\boldsymbol{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}}\boldsymbol{y}^{(1)} = \frac{1}{y_{p_1}^{(1)}}A\boldsymbol{x}^{(0)}.$$

Then $x_{p_1}^{(1)} = 1 = ||\boldsymbol{x}^{(1)}||_\infty$, and define

$$\boldsymbol{y}^{(2)} = A\boldsymbol{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}}A^2\boldsymbol{x}^{(0)}$$

and

$$\mu^{(2)} = y_{p_1}^{(2)} =$$

The Power Method (cont'd)

- Let p_2 be the smallest integer such that

$$|y_{p_2}^{(2)}| = ||\boldsymbol{y}^{(2)}||_\infty$$

and define

$$\boldsymbol{x}^{(2)} = \frac{1}{y_{p_2}^{(2)}}\boldsymbol{y}^{(2)} = \frac{1}{y_{p_2}^{(2)}}A\boldsymbol{x}^{(1)} = \frac{1}{y_{p_2}^{(2)}y_{p_1}^{(1)}}A^2\boldsymbol{x}^{(0)}.$$

The Power Method (cont'd)

- Similarly, define sequences of vectors $\{\boldsymbol{x}^{(m)}\}_{m=0}^{\infty}$ and $\{\boldsymbol{y}^{(m)}\}_{m=0}^{\infty}$ and a sequence of scalars $\{\mu^{(m)}\}_{m=0}^{\infty}$ inductively by

$$\begin{aligned}
\boldsymbol{y}^{(m)} &= A\boldsymbol{x}^{(m-1)}, \\
\mu^{(m)} &= y_{p_{m-1}}^{(m)} = \lambda_1 \left[\frac{\beta_1 \boldsymbol{v}_{p_{m-1}}^{(1)} + \sum_{j=2}^{n} (\lambda_j/\lambda_1)^m \beta_j \boldsymbol{v}_{p_{m-1}}^{(j)}}{\beta_1 \boldsymbol{v}_{p_{m-1}}^{(1)} + \sum_{j=2}^{n} (\lambda_j/\lambda_1)^{m-1} \beta_j \boldsymbol{v}_{p_{m-1}}^{(j)}} \right] \\
\boldsymbol{x}^{(m)} &= \frac{\boldsymbol{y}^{(m)}}{y_{p_m}^{(m)}} = \frac{A^m \boldsymbol{x}^{(0)}}{\prod_{k=1}^{m} y_{p_k}^{(k)}},
\end{aligned}$$

where at each step, p_m is used to represent the smallest integer for which

$$|y_{p_m}^{(m)}| = ||\boldsymbol{y}^{(m)}||_\infty.$$

The Power Method (cont'd)

- Since $|\lambda_j/\lambda_1| < 1$ for each $j = 2, 3, \ldots, n$,

$$\lim_{m\to\infty} \mu^{(m)} = \lambda_1,$$

provided that $\boldsymbol{x}^{(0)}$ is chosen so that $\beta_1 \neq 0$. Moreover, the sequence of vectors $\{\boldsymbol{x}^{(m)}\}_{m=0}^{\infty}$ converges to an eigenvector associated with λ_1 that has l_∞ norm equal to one.

- For $\boldsymbol{x}^{(0)} = \sum_{j=1}^{n} \beta_j \boldsymbol{v}^{(j)}$, we know that

$$\lim_{k\to\infty} A^k \boldsymbol{x}^{(0)} = \lim_{k\to\infty} \lambda_1^k \beta_1 \boldsymbol{v}^{(1)},$$

so

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \lim_{k\to\infty} \frac{A^k \boldsymbol{x}^{(0)}}{||A^k \boldsymbol{x}^{(0)}||_\infty} = \frac{\boldsymbol{v}^{(1)}}{||\boldsymbol{v}^{(1)}||_\infty}.$$

The Power Method (cont'd)

Ex. Approximate the dominant eigenvalue and its associated eigenvector of

$$A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix},$$

using two iterations of the power method with $\boldsymbol{x}_0 = (1, 1)^t$.

The Power Method (cont'd)

- Choosing the smallest integer p_m for which $|y_{p_m}^{(m)}| = ||\boldsymbol{y}^{(m)}||_\infty$ will generally ensure that this index eventually becomes invariant.
- The rate at which $\{\mu^{(m)}\}_{m=1}^{\infty}$ converges to λ_1 is determined by the ratios $\left|\frac{\lambda_j}{\lambda_1}\right|^m$, for $j = 2, 3, \ldots, n$, and in particular by $\left|\frac{\lambda_2}{\lambda_1}\right|^m$.

Inverse Power Method

- **Inverse Power method** is a modification of the Power method that determines the eigenvalue of A that is closest to a specified number q.

- Suppose that A has eigenvalues $\lambda_1, \ldots, \lambda_n$ with linearly independent eigenvectors $\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(n)}$.

Inverse Power Method (cont'd)

- The eigenvalues of $(A - qI)^{-1}$, where $q \neq \lambda_i$, for $i = 1, 2, \ldots, n$ are

$$\frac{1}{\lambda_1 - q}, \quad \frac{1}{\lambda_2 - q}, \quad \ldots, \quad \frac{1}{\lambda_n - q},$$

with the same eigenvectors $\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \ldots, \boldsymbol{v}^{(n)}$.

Inverse Power Method (cont'd)

- Applying the Power method to $(A - qI)^{-1}$ gives

$$\begin{aligned}
\boldsymbol{y}^{(m)} &= (A - qI)^{-1}\boldsymbol{x}^{(m-1)}, \\
\mu^{(m)} &= y^{(m)}_{p_{m-1}} = \frac{y^{(m)}_{p_{m-1}}}{x^{(m-1)}_{p_{m-1}}} = \\
\boldsymbol{x}^{(m)} &= \frac{\boldsymbol{y}^{(m)}}{y^{(m)}_{p_m}},
\end{aligned}$$

  where p_m represents the smallest integer for which $|y^{(m)}_{p_m}| = ||\boldsymbol{y}^{(m)}||_\infty$.

- The sequence $\mu^{(m)}$ converges to $\frac{1}{\lambda_k - q}$, where

$$\frac{1}{|\lambda_k - q|} = \max_{1 \le i \le n} \frac{1}{|\lambda_i - q|},$$

  where $\lambda_k \approx q + \frac{1}{\mu^{(m)}}$ is the eigenvalue of A closest to q.

Inverse Power Method (cont'd)

- With known k, we have

$$\mu^{(m)} = \frac{1}{\lambda_k - q}\left[\frac{\beta_k v_{p_{m-1}}^{(k)} + \sum_{\substack{j=1 \\ j\neq k}}^{n} \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q}\right]^m v_{p_{m-1}}^{(j)}}{\beta_k v_{p_{m-1}}^{(k)} + \sum_{\substack{j=1 \\ j\neq k}}^{n} \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q}\right]^{m-1} v_{p_{m-1}}^{(j)}}\right]$$

- The choice of q determines the convergence, provided that $1/(\lambda_k - q)$ is a unique dominant eigenvalue of $(A - qI)^{-1}$; the convergence is determined by the ratio

$$\left|\frac{(\lambda - q)^{-1}}{(\lambda_k - q)^{-1}}\right|^m = \left|\frac{\lambda_k - q}{\lambda - q}\right|^m$$

  where λ represents the eigenvalue of A that is second closest to q.

- Use Geršgorin Circle Theorem to initialize q, or choose q from $\boldsymbol{x}^{(0)}$ by

$$q = \frac{[\boldsymbol{x}^{(0)}]^t A \boldsymbol{x}^{(0)}}{[\boldsymbol{x}^{(0)}]^t \boldsymbol{x}^{(0)}}.$$

Deflation Methods

- How can we obtain other eigenvalues of matrix once an approximation to the dominant eigenvalue has been computed?
- **Deflation techniques** involve forming a new matrix B whose eigenvalues are the same as those of A, except that the dominant eigenvalue of A is replaced by the eigenvalue 0 in B.

Deflation Methods (cont'd)

Theorem 2

Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are eigenvalues of* A *with associated eigenvectors* $\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \ldots, \boldsymbol{v}^{(n)}$ *and that* λ_1 *has multiplicity* 1. *Let* $\boldsymbol{x}$ *be a vector with* $\boldsymbol{x}^t \boldsymbol{v}^{(1)} = 1$. *Then the matrix*

$$B = A - \lambda_1 \boldsymbol{v}^{(1)} \boldsymbol{x}^t$$

has eigenvalues $0, \lambda_2, \lambda_3, \ldots, \lambda_n$ *with associated eigenvectors* $\boldsymbol{v}^{(1)}, \boldsymbol{w}^{(2)}, \boldsymbol{w}^{(3)}, \ldots, \boldsymbol{w}^{(n)}$, *where* $\boldsymbol{v}^{(i)}$ *and* $\boldsymbol{w}^{(i)}$ *are related by the equation*

$$\boldsymbol{v}^{(i)} = (\lambda_i - \lambda_1) \boldsymbol{w}^{(i)} + \lambda_1 (\boldsymbol{x}^t \boldsymbol{w}^{(i)}) \boldsymbol{v}^{(1)},$$

for each $i = 2, 3, \ldots, n$.

Deflation Methods (cont'd)

- Wielandt deflation proceeds with

$$\boldsymbol{x} = \frac{1}{\lambda_1 v_i^{(1)}} (a_{i1}, a_{i2}, \ldots, a_{in})^t,$$

where $v_i^{(1)}$ is a nonzero coordinate of the eigenvector $\boldsymbol{v}^{(1)}$.

Deflation Method (cont'd)

- The ith row of B consists entirely of zero entries
- If $\lambda \neq 0$ is an eigenvalue with associated eigenvector $\boldsymbol{w}$, the relation $B\boldsymbol{w} = \lambda\boldsymbol{w}$ implies that the ith coordinate of $\boldsymbol{w}$ must also be zero.

Deflation Method (cont'd)

- The matrix B can be then replaced by $(n-1) \times (n-1)$ matrix B', by deleting the ith row and column from B. The matrix B' has eigenvalues $\lambda_2, \lambda_3, \ldots, \lambda_n$.
- If $|\lambda_2| > |\lambda_3|$, the Power method can be applied to B' to determine this new dominant eigenvalue λ_2 and an associated eigenvector $\boldsymbol{w}^{(2)'}$.
- To find $\boldsymbol{w}^{(2)}$ for B, insert a zero coordinate between the coordinates $w_{i-1}^{(2)'}$ and $w_i^{(2)'}$ of $\boldsymbol{w}^{(2)'}$ and then calculate

$$\boldsymbol{v}^{(2)} = (\lambda_2 - \lambda_1)\boldsymbol{w}^{(2)} + \lambda_1(\boldsymbol{x}^t \boldsymbol{w}^{(2)})\boldsymbol{v}^{(1)}.$$

Deflation Method (cont'd)

Ex The matrix

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

has the dominant eigenvalue $\lambda_1 = 6$ with associated unit eigenvector $\boldsymbol{v}^{(1)} = (1, -1, 1)^t$. Apply deflation to approximate the other eigenvalues and eigenvectors.