Problems of Eigenvalues/Eigenvectors

- ♣ Reveiw of Eigenvalues and Eigenvectors
- ♣ Gerschgorin's Disk Theorem
- ♣ Power and Inverse Power Methods
- ♣ Jacobi Transform for Symmetric Matrices
- ♣ Singular Value Decomposition with Applications
- ♣ QR Iterations for Computing Eigenvalues
- ♣ Other Topics with Applications

Definition and Examples

Let $A \in R^{n\times n}$. If $\exists$ $\mathbf{v} \neq \mathbf{0}$ such that $A\mathbf{v} = \lambda\mathbf{v}$, λ is called an eigenvalue of matrix A, and $\mathbf{v}$ is called an eigenvector corresponding to (or belonging to) the eigenvalue λ. Note that $\mathbf{v}$ is an eigenvector implies that $\alpha\mathbf{v}$ is also an eigenvector for all $\alpha \neq 0$. We define the Eigenspace(λ) as the vector space spanned by all of the eigenvectors corresponding to the eigenvalue λ.

Examples:

1. $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $\lambda_1 = 2$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda_2 = 1$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
2. $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, $\lambda_1 = 2$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda_2 = 1$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
3. $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, $\lambda_1 = 4$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda_2 = 2$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
4. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\lambda_1 = j$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ j \end{bmatrix}$, $\lambda_2 = -j$, $\mathbf{u}_2 = \begin{bmatrix} j \\ 1 \end{bmatrix}$, $j = \sqrt{-1}$.
5. $B = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$, then $\lambda_1 = 3$, $\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$; $\lambda_2 = -1$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
6. $C = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$, then $\tau_1 = 4$, $\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$; $\tau_2 = 2$, $\mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.

$$A\mathbf{x} = \lambda\mathbf{x} \;\Rightarrow\; (\lambda I - A)\mathbf{x} = \mathbf{0},\ \mathbf{x} \neq \mathbf{0} \;\Rightarrow\; det(\lambda I - A) = P(\lambda) = 0.$$

Gershgorin's Disk Theorem

Note that $\|\mathbf{u}_i\|_2 = 1$ and $\|\mathbf{v}_i\|_2 = 1$ for $i = 1, 2$. Denote $U = [\mathbf{u}_1, \mathbf{u}_2]$ and $V = [\mathbf{v}_1, \mathbf{v}_2]$, then

$$U^{-1}BU = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}, \quad V^{-1}CV = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

Note that $V^t = V^{-1}$ but $U^t \neq U^{-1}$.

Let $A \in R^{n\times n}$, then $det(\lambda I - A)$ is called the *characteristic polynomial* of matrix A.

♣ Fundamental Theorem of Algebra

A real polynomial $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$ of degree n has n roots $\{\lambda_i\}$ such that

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) = \lambda^n - \left(\sum_{i=1}^{n} \lambda_i\right)\lambda^{n-1} + \cdots + (-1)^n \left(\prod_{i=1}^{n} \lambda_i\right)$$

- $\sum_{i=1}^{n} \lambda_i \;=\; \sum_{i=1}^{n} a_{ii} \;=\; tr(A)$
- $\prod_{i=1}^{n} \lambda_i \;=\; det(A)$

♣ Gershgorin's Disk Theorem

Every eigenvalue of matrix $A \in R^{n\times n}$ lies in at least one of the disks

$$D_i = \{x \mid |x - a_{ii}| \; \leq \sum_{j\neq i} |a_{ij}|\}, \quad 1 \leq i \leq n$$

Example: $B = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 4 & 1 \\ 2 & 2 & 5 \end{bmatrix}$, $\lambda_1, \lambda_2, \lambda_3 \in D_1 \cup D_2 \cup D_3$, *where* $D_1 = \{z \mid \; |z-3| \leq 2\}$, $D_2 = \{z \mid \; |z-4| \leq 1\}$, $D_3 = \{z \mid \; |z-5| \leq 4\}$. Note that $\lambda_1 = 6.5616$, $\lambda_2 = 3.0000$, $\lambda_3 = 2.4383$.

□ A matrix is said to be *diagonally dominant* if $\sum_{j\neq i} |a_{ij}| < |a_{ii}|$, $\forall \; 1 \leq i \leq n$.

◇ A diagonally dominant matrix is invertible.

Theorem: Let $A, P \in R^{n\times n}$, with P nonsingular, then λ is an eigenvalue of A with eigenvector $\mathbf{x}$ iff λ is an eigenvalue of $P^{-1}AP$ with eigenvector $P^{-1}\mathbf{x}$.

Theorem: Let $A \in R^{n\times n}$ and let λ be an eigenvalue of A with eigenvector $\mathbf{x}$. Then

(a) $\alpha\lambda$ is an eigenvalue of matrix αA with eigenvector $\mathbf{x}$

(b) $\lambda - \mu$ is an eigenvalue of matrix $A - \mu I$ with eigenvector $\mathbf{x}$

(c) If A is nonsingular, then $\lambda \neq 0$ and λ^{-1} is an eigenvalue of A^{-1} with eigenvector $\mathbf{x}$

Definition: A matrix A is similar to B, denote by $A \sim B$, iff there exists an invertible matrix U such that $U^{-1}AU = B$. Furthermore, a matrix A is *orthogonally similar* to B, iff there exists an orthogonal matrix Q such that $Q^tAQ = B$.

Theorem: Two similar matrices have the same eigenvalues, i.e., $A \sim B \Rightarrow \lambda(A) = \lambda(B)$.

Diagonalization of Matrices

Theorem: Suppose $A \in R^{n\times n}$ has n linearly independent eigenvectors $\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_n$ corresponding to eigenvalues $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$. Let $V = [\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_n]$, then $V^{-1}AV = diag[\lambda_1, \lambda_2, \ldots, \lambda_n]$.

◇ If $A \in R^{n\times n}$ has n distinct eigenvalues, then their corresponding eigenvectors are linearly independent. Thus, any matrix with distinct eigenvalues can be diagonalized.

◇ Not all matrices have distinct eigenvalues, therefore not all matrices are diagonalizable.

Spectrum Decomposition Theorem*

Every real symmetric matrix can be diagonalized.

Nondiagonalizable Matrices

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$$

Diagonalizable Matrices

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Theorem: Let $\{(\lambda_i, \mathbf{v}_i),\ 1 \le i \le n\}$ be eigenvalues/eigenvectors of matrix $A \in R^{n\times n}$, then $A^k\mathbf{v}_j = \lambda_j^k\mathbf{v}_j$, $\forall\ k \ge 1$. Moreover, if $\{\mathbf{v}_i\}$ are linearly independent, then $\forall \mathbf{y} \in R^n$ can be written in the form

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Then

$$A^k\mathbf{y} = \lambda_1^k c_1\mathbf{v}_1 + \lambda_2^k c_2\mathbf{v}_2 + \cdots + \lambda_n^k c_n\mathbf{v}_n.$$

If $|\lambda_1| > |\lambda_i|,\ \forall 2 \le i \le n$, and $c_1 \ne 0$, then $A^k\mathbf{y} \to \alpha\mathbf{v}_1$ as $k \to \infty$.

A Markov Process

Suppose that 10% of the people outside Taiwan move in, and 20% of the people indside Taiwan move out in each year. Let y_k and z_k be the population at the end of the $k-th$ year, outside Taiwan and inside Taiwan, respectively. Then we have

$$\begin{bmatrix} y_k \\ \\ z_k \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_{k-1} \\ \\ z_{k-1} \end{bmatrix} \quad \Rightarrow \quad \lambda_1 = 1.0, \ \lambda_2 = 0.7$$

$$\begin{bmatrix} y_k \\ \\ z_k \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ \\ 0.1 & 0.8 \end{bmatrix}^k \begin{bmatrix} y_0 \\ \\ z_0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ \\ 0 & (0.7)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ \\ z_0 \end{bmatrix}$$

□ A *Markov* matrix A is nonnegative with each colume adding to 1.

(a) $\lambda_1 = 1$ is an eigenvalue with a nonnegative eigenvector $\mathbf{x}_1$.

(b) The other eigenvalues satisfy $|\lambda_i| \leq 1$.

(c) If any power of A has all positive entries, and the other $|\lambda_i| < 1$. Then $A^k\mathbf{u}_0$ approaches the steady state of $\mathbf{u}_\infty$ which is a multiple of $\mathbf{x}_1$ as long as the projection of $\mathbf{u}_0$ in $\mathbf{x}_1$ is not zero.

◇ Check Perron-Fröbenius theorem in Strang's book.

Differential Equations and e^A

♣ $e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!} + \cdots$

♣ $\frac{du}{dt} = -\lambda u \;\Rightarrow\; u(t) = e^{-\lambda t}u(0)$

♣ $\frac{d\mathbf{u}}{dt} = -A\mathbf{u} \;=\; \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}\mathbf{u} \;\Rightarrow\; \mathbf{u}(t) = e^{-tA}\mathbf{u}(0)$

♣ $A = U\Lambda U^t$ for an orthogonal matrix U, then

$$e^A = Ue^{\Lambda}U^{=}Udiag[e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}]U^t$$

♣ Solve $x''' - 3x'' + 2x' = 0$.

Let $y = x'$, $z = y' = x''$, and let $\mathbf{u} = [x, y, z]^t$. The problem is reduced to solving

$$\mathbf{u}' = A\mathbf{u} \;=\; \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{bmatrix}\mathbf{u}$$

Then

$$\mathbf{u}(t) = e^{tA}\mathbf{u}(0) \;=\; \begin{bmatrix} \frac{1}{\sqrt{21}} & \frac{1}{\sqrt{3}} & 1 \\ \frac{2}{\sqrt{21}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{4}{\sqrt{21}} & \frac{1}{\sqrt{3}} & 0 \end{bmatrix} \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^{t} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -2.2193 & 2.2193 \\ 0 & 3.4641 & -1.7321 \\ 1 & 1.5000 & 0.5000 \end{bmatrix}\mathbf{u}(0)$$

Similarity transformation and triangularization

Schur's Theorem: $\forall\, A \in R^{n\times n}$, $\exists$ an orthogonal matrix U such that $U^tAU = T$ is upper-Δ. The eigenvlues must be shared by the similarity matrix T and appear along its main diagonal.

Hint: By induction, suppose that the theorem has been proved for all matrices of order $n-1$, and consider $A \in R^{n\times n}$ with $A\mathbf{x} = \lambda\mathbf{x}$ and $\|\mathbf{x}\|_2 = 1$, then $\exists$ a Householder matrix H_1 such that $H_1\mathbf{x} = \beta\mathbf{e}_1$, *e.g.*, $\beta = -\|\mathbf{x}\|_2$, hence

$$H_1AH_1^t\mathbf{e}_1 = H_1A(H_1^{-1}\mathbf{e}_1) = H_1A(\beta^{-1}\mathbf{x}) = H_1\beta^{-1}A\mathbf{x} = \beta^{-1}\lambda(H_1\mathbf{x}) = \beta^{-1}\lambda(\beta\mathbf{e}_1) = \lambda\mathbf{e}_1$$

Thus,

$$H_1AH_1^t = \begin{bmatrix} \lambda & | & * \\ --- & | & --- \\ O & | & A^{(1)} \end{bmatrix}$$

Spectrum Decomposition Theorem: Every real symmetric matrix can be diagonalized by an orthogonal matrix.

$\diamond$ $Q^tAQ = \Lambda$ or $A = Q\Lambda Q^t = \sum_{i=1}^{n} \lambda_i \mathbf{q}_i\mathbf{q}_i^t$

Definition: A symmetric matrix $A \in R^{n\times n}$ is nonnegative definite if $\mathbf{x}^tA\mathbf{x} \geq 0\ \forall\ \mathbf{x} \in R^n$, $\mathbf{x} \neq \mathbf{0}$.

Definition: A symmetric matrix $A \in R^{n\times n}$ is positive definite if $\mathbf{x}^tA\mathbf{x} > 0\ \forall\ \mathbf{x} \in R^n$, $\mathbf{x} \neq \mathbf{0}$.

Singular Value Decomposition Theorem: Each matrix $A \in R^{m\times n}$ can be decomposed as $A = U\Sigma V^t$, where both $U \in R^{m\times m}$ and $V \in R^{n\times n}$ are orthogonal. Moreover, $\Sigma \in R^{m\times n} = diag[\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0]$ is essentially diagonal with the singular values satisfying $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$.

$\diamond$ $A = U\Sigma V^t = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i\mathbf{v}_i^t$

Example:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix},\quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

A Jacobi Transform (Givens Rotation)

$$J(i,k;\theta) = \begin{bmatrix} 1 & \cdot & \cdot & \cdots & \cdot & \cdot & 0 \\ 0 & \ddots & \cdot & \cdots & \cdot & \vdots & 0 \\ 0 & \cdot & c & \cdots & s & \cdot & 0 \\ \cdot & \vdots & \cdot & \ddots & \cdot & \vdots & \cdot \\ 0 & \cdot & -s & \cdots & c & \cdot & 0 \\ 0 & \vdots & \cdot & \cdots & \cdot & \ddots & 0 \\ \cdot & \cdot & 0 & \cdots & 0 & \cdot & 1 \end{bmatrix}$$

$J_{hh} = 1$ if $h \neq i$ or $h \neq k$, where $i < k$

$J_{ii} = J_{kk} = c = \cos\theta$

$J_{ki} = -s = -\sin\theta,\ J_{ik} = s = \sin\theta$

Let $\mathbf{x}, \mathbf{y} \in R^n$, then $\mathbf{y} = J(i,k;\theta)\mathbf{x}$ implies that

$y_i = cx_i + sx_k$

$y_k = -sx_i + cx_k$

$c = \frac{x_i}{\sqrt{x_i^2 + x_k^2}},\ s = \frac{x_k}{\sqrt{x_i^2 + x_k^2}},$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, \quad then \quad J(2,4;\theta)\mathbf{x} = \begin{bmatrix} 1 \\ \sqrt{20} \\ 3 \\ 0 \end{bmatrix}$$

Jacobi Transforms (Givens Rotations)

The Jacobi method consists of a sequence of orthogonal similarity transformations such that

$$J_K^t J_{K-1}^t \cdots J_2^t J_1^t A J_1 J_2 \cdots J_{K-1} J_K = \Lambda$$

where each J_i is orthogonal, so is $Q = J_1 J_2 \cdots J_{K-1} J_K$.

Each Jacobi transform (Given rotation) is just a plane rotation designed to annihilate one of the off-diagonal matrix elements. Let $A = (a_{ij})$ be symmetric, then

$B = J^t(p, q, \theta) A J(p, q, \theta)$, where

$b_{rp} = c a_{rp} - s a_{rq} \quad for \quad r \neq p,\ r \neq q$

$b_{rq} = s a_{rp} + c a_{rq} \quad for \quad r \neq p,\ r \neq q$

$b_{pp} = c^2 a_{pp} + s^2 a_{qq} - 2sc a_{pq}$

$b_{qq} = s^2 a_{pp} + c^2 a_{qq} + 2sc a_{pq}$

$b_{pq} = (c^2 - s^2) a_{pq} + sc(a_{pp} - a_{qq})$

To set $b_{pq} = 0$, we choose c, s such that

$$\alpha = \cot(2\theta) = \frac{c^2 - s^2}{2sc} = \frac{a_{qq} - a_{pp}}{2a_{pq}} \tag{1}$$

For computational convenience, let $t = \frac{s}{c}$, then $t^2 + 2\alpha t - 1 = 0$ whose smaller root (in absolute sense) can be computed by

$$t = \frac{sgn(\alpha)}{\sqrt{\alpha^2 + 1} + |\alpha|}, \quad and \quad c = \frac{1}{\sqrt{1 + t^2}}, \quad s = ct, \quad \tau = \frac{s}{1 + c} \tag{2}$$

Remark

$b_{pp} = a_{pp} - t a_{pq}$

$b_{qq} = a_{qq} + t a_{pq}$

$b_{rp} = a_{rp} - s(a_{rq} + \tau a_{rp})$

$b_{rq} = a_{rq} + s(a_{rp} - \tau a_{rq})$

Algorithm of Jacobi Transforms to Diagonalize A

$A^{(0)} \leftarrow A$

for $k = 0, 1, \cdots,$ until convergence

Let $|a_{pq}^{(k)}| = Max_{i<j}\{|a_{ij}^{(k)}|\}$

Compute

$\alpha_k = \frac{a_{qq}^{(k)} - a_{pp}^{(k)}}{2a_{pq}^{(k)}}$, solve $\cot(2\theta_k) = \alpha_k$ for θ_k.

$t = \frac{sgn(\alpha)}{\sqrt{\alpha^2+1}+|\alpha|}$

$c = \frac{1}{\sqrt{1+t^2}}$, , $s = ct$

$\tau = \frac{s}{1+c}$

$A^{(k+1)} \leftarrow J_k^t A^{(k)} J_k$, where $J_k = J(p, q, \theta_k)$

endfor

Convergence of Jacobi Algorithm to Diagonalize A

Proof:

Since $|a_{pq}^{(k)}| \geq |a_{ij}^{(k)}|$ for $i \neq j,\ p \neq q$, then

$|a_{pq}^{(k)}|^2 \geq off(A^{(k)})/2N$, where $N = \frac{n(n-1)}{2}$, and

$off(A^{(k)}) = \sum_{i \neq j}^{n} \left(a_{ij}^{(k)}\right)^2$, the sum of square off-diagonal elements of $A^{(k)}$

Furthermore,

$$
\begin{aligned}
off(A^{(k+1)}) &= off(A^{(k)}) - 2\left(a_{pq}^{(k)}\right)^2 + 2\left(a_{pq}^{(k+1)}\right)^2 \\
&= off(A^{(k)}) - 2\left(a_{pq}^{(k)}\right)^2, \quad since \quad a_{pq}^{(k+1)} = 0 \\
&\leq off(A^{(k)})\left(1 - \tfrac{1}{N}\right), \quad since |a_{pq}^{(k)}|^2 \geq off(A^{(k)}/2N
\end{aligned}
$$

Thus

$$off(A^{(k+1)}) \leq \left(1 - \frac{1}{N}\right)^{k+1} off(A^{(0)}) \quad \rightarrow\ 0 \ \ as \ \ k\ \rightarrow\ \infty$$

Example:

$$A = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad J(1,2;\theta) = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$A^{(1)} = J^t(1,2;\theta) A J(1,2;\theta) = \begin{bmatrix} 4c^2 - 4cs + 3s^2 & 2c^2 + cs - 2s^2 & -s \\ 2c^2 + cs - 2s^2 & 3c^2 + 4cs + 4s^2 & c \\ -s & c & 1 \end{bmatrix}$$

Note that $off(A^{(1)}) = 2 < 10 = off(A^{(0)}) = off(A)$

Example for Convergence of Jacobi Algorithm

$$A^{(0)} = \begin{bmatrix} 1.0000 & 0.5000 & 0.2500 & 0.1250 \\ 0.5000 & 1.0000 & 0.5000 & 0.2500 \\ 0.2500 & 0.5000 & 1.0000 & 0.5000 \\ 0.1250 & 0.2500 & 0.5000 & 1.0000 \end{bmatrix}, \quad A^{(1)} = \begin{bmatrix} 1.5000 & 0.0000 & 0.5303 & 0.2652 \\ 0.0000 & 0.5000 & 0.1768 & 0.0884 \\ 0.5303 & 0.1768 & 1.0000 & 0.5000 \\ 0.2652 & 0.0884 & 0.5000 & 1.0000 \end{bmatrix}$$

$$A^{(2)} = \begin{bmatrix} 1.8363 & 0.0947 & 0.0000 & 0.4917 \\ 0.0947 & 0.5000 & 0.1493 & 0.0884 \\ 0.0000 & 0.1493 & 0.6637 & 0.2803 \\ 0.4917 & 0.0884 & 0.2803 & 1.0000 \end{bmatrix}, \quad A^{(3)} = \begin{bmatrix} 2.0636 & 0.1230 & 0.1176 & 0.0000 \\ 0.1230 & 0.5000 & 0.1493 & 0.0405 \\ 0.1176 & 0.1493 & 0.6637 & 0.2544 \\ 0.0000 & 0.0405 & 0.2544 & 0.7727 \end{bmatrix}$$

$$A^{(4)} = \begin{bmatrix} 2.0636 & 0.1230 & 0.0915 & 0.0739 \\ 0.1230 & 0.5000 & 0.0906 & 0.1254 \\ 0.0915 & 0.0906 & 0.4580 & 0.0000 \\ 0.0739 & 0.1254 & 0.0000 & 0.9783 \end{bmatrix}, \quad A^{(5)} = \begin{bmatrix} 2.0636 & 0.1018 & 0.0915 & 0.1012 \\ 0.1018 & 0.4691 & 0.0880 & 0.0000 \\ 0.0915 & 0.0880 & 0.4580 & 0.0217 \\ 0.1012 & 0.0000 & 0.0217 & 1.0092 \end{bmatrix}$$

$$A^{(6)} = \begin{bmatrix} 2.0701 & 0.0000 & 0.0969 & 0.1010 \\ 0.0000 & 0.4627 & 0.0820 & -0.0064 \\ 0.0969 & 0.0820 & 0.4580 & 0.0217 \\ 0.1010 & -0.0064 & 0.0217 & 1.0092 \end{bmatrix}, \quad A^{(15)} = \begin{bmatrix} 2.0856 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.5394 & 0.0000 & -0.0000 \\ 0.0000 & 0.0000 & 0.3750 & 0.0000 \\ 0.0000 & -0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

Power of A Matrix and Its Eigenvalues

Theorem: Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be eigenvalues of $A \in R^{n\times n}$. Then $\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k$ are eigenvalues of $A^k \in R^{n\times n}$ with the same corresponding eigenvectors of A. That is,

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i \quad \rightarrow \quad A^k \mathbf{v}_i = \lambda_i^k \mathbf{v}_i \ \ \forall\ 1 \le i \le n$$

Suppose that the matrix $A \in R^{n\times n}$ has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_1, \cdots, \mathbf{v}_n$ corresponding to eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then any $\mathbf{x} \in R^n$ can be written as

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Then

$$A^k\mathbf{x} = \lambda_1^k c_1\mathbf{v}_1 + \lambda_2^k c_2\mathbf{v}_2 + \cdots + \lambda_n^k c_n\mathbf{v}_n$$

In particular, if $|\lambda_1| > |\lambda_j|$ for $2 \le j \le n$ and $c_1 \ne 0$, then $A^k\mathbf{x}$ will tend to lie in the direction $\mathbf{v}_1$ when k is *large enough.*

Power Method for Computing the Largest Eigenvalues

Suppose that the matrix $A \in R^{n\times n}$ is diagonalizable and that $U^{-1}AU = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ with $U = [\mathbf{v}_1,\ \mathbf{v}_2,\ \cdots,\ \mathbf{v}_n]$ and $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$. Given $\mathbf{u}^{(0)} \in R^n$, then power method produces a sequence of vectors $\mathbf{u}^{(k)}$ as follows.

for $k = 1, 2, \cdots$

$\mathbf{z}^{(k)} = A\mathbf{u}^{(k-1)}$

$r^{(k)} = z_m^{(k)} = \|\mathbf{z}^{(k)}\|_\infty$, for some $1 \leq m \leq n$.

$\mathbf{u}^{(k)} = \mathbf{z}^{(k)}/r^{(k)}$

endfor

λ_1 must be real since the complex eigenvalues must appear in a "relatively conjugate pair".

$$A = \begin{bmatrix} 2 & 1 \\ & \\ 1 & 2 \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} \lambda_1 = 3 \\ \\ \lambda_2 = 1 \end{matrix}, \quad \mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Let $\mathbf{u}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $\mathbf{u}^{(5)} = \begin{bmatrix} 1.0 \\ 0.9918 \end{bmatrix}$, and $r^{(5)} = 2.9756$.

QR Iterations for Computing Eigenvalues

```
%
% Script File: eigQR.m
% Solving Eigenvalues by QR factorization
%
n=4; Nrun=50;
fin=fopen('dataToeplitz.txt');
header1=fgetL(fin);
k=fscanf(fin,'%d');
A=fscanf(fin,'%f',[n n]);
A=A';
SaveA=A;
for k=1:Nrun,
  s=A(n,n);
  A=A-s*eye(n);
  [Q R]=qr(A);
  A=R*Q+s*eye(n);
end
for i=1:n,
  D(i)=A(i,i);
end
D=D';
E=sort(D,1);
E'
%
% Eigenvalues computed by Matlab Command
%
[U S]=eig(SaveA);
for i=1:n,
  D(i)=S(i,i);
end
D
```

Algebraic multiplicity and geometric multiplicity

♣ Albebraic Multiplicity

When the characteristic polynomial of $A \in R^{n \times n}$ is written as

$$det(\lambda I - A) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$$

with $\lambda_i \neq \lambda_j \ \forall \ i \neq j$ and $n_1 + n_2 + \ldots + n_k = n$. The positive integer n_i is called the *algebraic multiplicity* of the eigenvalue λ_i.

♣ Geometric Multiplicity

The geometric multiplicity m_i of the eigenvalue λ_i is defined as the maximum number of linearly independent eigenvectors associated with λ_i. That is, $m_i = \lambda_i(S)$, the dimension of the eigenspace. Note that $1 \leq m_i \leq n_i \ forall \ 1 \leq i \leq k$.

Example:

$$A = \begin{bmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

$$det(\lambda I - A) = (\lambda - 7)^2 (\lambda - 4)^3$$

$$n_1 = 2, \quad n_2 = 3$$

$$m_1 = 2, \quad m_2 = 2$$

$$\lambda_1 = 7, \quad \mathbf{v}_1 = a\mathbf{e}_1 + b\mathbf{e}_4$$

$$\lambda_2 = 4, \quad \mathbf{v}_2 = c\mathbf{e}_2 + d\mathbf{e}_5$$

Block Upper Triangular Matrices

Definition: The square matrix T is *block upper triangular* if it can be partitioned in the form

$$\begin{bmatrix} T_{11} & T_{12} & \cdot & \cdot & T_{1r} \\ O & T_{22} & \cdot & \cdot & T_{2r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ O & O & \cdot & \cdot & T_{rr} \end{bmatrix}$$

where each diagonal block T_{ii} is square. If each diagonal block is of order at most two, then T is said to be in a quasi-Δ form.

Theorem: $\lambda(T) \;=\; \cup_{i=1}^{r} \lambda(T_{ii})$

Theorem: Let $A \in R^{n\times n}$ have the characteristic polynomial

$$p(x) = (x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k}$$

where $\lambda_1,\ \lambda_2,\ \ldots,\lambda_k$ are distinct. Then A is similar to a matrix of the form

$$\begin{bmatrix} B_1 & O & \cdot & \cdot & O \\ O & B_2 & \cdot & \cdot & O \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ O & O & \cdot & \cdot & B_k \end{bmatrix}$$

where each B_i is an n_i by n_i upper-Δ matrix whose diagonal entries are λ_i.

Cayley-Hamilton Theorem

Cayley-Hamilton Theorem: $p(A) = O$

Example:

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \quad \Rightarrow \quad p(x) = x^2 - 7x + 10 \quad \Rightarrow \quad A^2 - 7A + 10I = O$$

Example:

$$A = \begin{bmatrix} -3 & 2 & 1 & 1 \\ -6 & 3 & 3 & 1 \\ -3 & 2 & 0 & 2 \\ -2 & 2 & 1 & 0 \end{bmatrix} \quad \sim \quad T = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Block Diagonal Upper Triangular Form

Lemma: Suppose that the matrices T and B have the forms

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdot & \cdot & T_{1r} \\ O & T_{22} & \cdot & \cdot & T_{2r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ O & O & \cdot & \cdot & T_{rr} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & O & \cdot & \cdot & O \\ O & B_2 & \cdot & \cdot & O \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ O & O & \cdot & \cdot & B_r \end{bmatrix}$$

where $T_{ii} \in R^{n_i \times n_i}$ is upper-Δ, all of the main diagonal entries of T_{ii} equal λ_i, and $\lambda_i's$ are distinct for $1 \leq i \leq r$. Then T is similar to a block diagonal upper-Δ matrix B, where $B_i \in R^{n_i \times n_i}$ is upper-Δ whose main diagonal entries equal λ_i above.

Minimal Polynomial

Definition: The minimal polynomial of a matrix A over a field R is defined as the monic polynomial f with coefficients in R of least degree such that $f(A) = O$.

Example:

$$A = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix}, \quad \lambda_i \neq \lambda_j \; if \; i \neq j$$

Then

$$f_A(x) = (x-5)^2$$

$$f_B(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$$

Theorem: Similar matrices have the same minimal polynomial.

Jordan Canonical Form

A Jordan block having the eigenvalue λ of geometric multiplicity k has the form

$$J_\lambda^{(k)} = \begin{bmatrix} \lambda & 1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda & 1 & \cdot & \cdot & 0 \\ \cdot & 0 & \lambda & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \lambda & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & \lambda \end{bmatrix}$$

Theorem: Let $A \in R^{n\times n}$, then there are unique numbers $\lambda_1, \lambda_2, \ldots, \lambda_k \in \lambda(A)$ and $n_1, n_2, \ldots, n_k$ such that A is similar to the matrix

$$diag\left(J_{\lambda_1}^{n_1}, J_{\lambda_2}^{n_2}, \ldots, J_{\lambda_k}^{n_k}\right)$$

Example:

$$A = \begin{bmatrix} -2 & -1 & 2 \\ 2 & 2 & -1 \\ -3 & -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 1 & 3 & -1 \\ 3 & 0 & -2 & 2 \\ 1 & 1 & 2 & 1 \\ 1 & -1 & -3 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then A, B, C are silmilar to the following Jordan canonical forms.

$$J_A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad J_B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad J_C = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Computing A Jordan Canonical Form

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad J_A = U^{-1}AU = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$\lambda_1 = \lambda_2 = \lambda_3 = 0$. $dim(R(A)) = 2$, $Null(A) = \{[a, 0, 0]^t |\ a \in R\}$

(a) Find $\mathbf{w}_1, \mathbf{w}_2$ such that $R(A) = \{a\mathbf{w}_1 + b\mathbf{w}_2 |\ a, b \in R\}$ and $A\mathbf{w}_1 = \lambda_1\mathbf{w}_1,\ A\mathbf{w}_2 = \lambda_2\mathbf{w}_2 + \mathbf{w}_1$. Let

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} x \\ 1 \\ 0 \end{bmatrix}, \quad with \ \ x = 2$$

(b) Solve $A\mathbf{y} = \mathbf{w}_2$ to get

$$\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \quad with \ \ z = 1$$

(c) Since

$$\begin{aligned} A\mathbf{w}_1 &= \lambda_1\mathbf{w}_1 \\ A\mathbf{w}_2 &= \lambda_2\mathbf{w}_2 + \mathbf{w}_1 \\ A\mathbf{y} &= \mathbf{w}_2 \end{aligned}$$

(d) Let $U = [\mathbf{w}_1,\ \mathbf{w}_2,\ \mathbf{y}]$, then $U^{-1}AU = J_A$.