L08: Eigenvalues and eigenvectors

1. Eigenvalues

   (1) Eigenvalue and eigenvector

   $x \in C^n$ is an eigenvector wrt the eigenvalue $\lambda \in C$ for $A \in C^{n\times n} \overset{def}{\Longleftrightarrow} Ax = \lambda x,\ x \neq 0$.

   Comments: If $x = 0$ is allowed to be an eigenvector, all λ would have been eigenvalues.
   When $\lambda = 0$, $Ax = 0,\ x \neq 0 \Longleftrightarrow 0 \neq x \in \mathcal{N}(A)$ gives all eigenvectors for $\lambda = 0$.

   (2) Distinct eigenvalues and multiplicity

   $$\begin{aligned} \lambda \text{ is an eigenvalue for } A &\Longleftrightarrow Ax = \lambda x,\ x \neq 0 \Longleftrightarrow (A - \lambda I)x = 0,\ x \neq 0 \\ &\Longleftrightarrow \text{rank}(A - \lambda I) < n \Longleftrightarrow |A - \lambda I| = 0 \\ &\Longleftrightarrow \lambda \text{ is a solution to the equation } |A - \lambda I| = 0. \end{aligned}$$

   $$\begin{aligned} p(\lambda) &= |A - \lambda I| = (-\lambda)^n + c_{n-1}(-\lambda)^{n-1} + \cdots + c_1(-\lambda) + c_0 = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda) \\ &= (\lambda_1 - \lambda)^{r_1}\cdots(\lambda_k - \lambda)^{r_k} \end{aligned}$$

   is the characteristic polynomial, and $P(\lambda) = 0$ is characteristic equation.
   $\lambda_1, ..., \lambda_k$ are distinct eigenvalues with multiplicity $r_1, ..., r_k$. $r_1 + \cdots + r_k = n$.
   An eigenvalue is a simple one if its multiplicity is 1.

   (3) Coefficients of characteristic polynomial

   For A with n eigenvalues $\lambda_1, ..., \lambda_n$, (may not be distinct), in

   $$p(\lambda) = |A - \lambda I| = (-\lambda)^n + c_{n-1}(-\lambda)^{n-1} + \cdots + c_1(-\lambda) + c_0 = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda),$$

   $c_0 = |A| = \lambda_1 \cdots \lambda_n$ and $c_{n-1} = \text{tr}(A) = \lambda_1 + \cdots + \lambda_n$.

   Pf: $P(0) = |A| = c_0 = \lambda_1 \cdots \lambda_n$.
   The term $(-\lambda)^{n-1}$ in $|A - \lambda I|$ is from $(a_{11} - \lambda)\cdots(a_{nn} - \lambda)$ with coefficient $\text{tr}(A)$.
   The term $(-\lambda)^{n-1}$ in $p(\lambda)$ has coefficient c_{n-1}.
   The term $(-\lambda)^{n-1}$ in $(\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$ has coefficient $\lambda_1 + \cdots + \lambda_n$.
   So $\text{tr}(A) = c_{n-1} = \sum_i \lambda_i$.

Ex1: $A = \begin{pmatrix} 0 & -1 \\ 4 & 0 \end{pmatrix}$, $|A - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 4 & -\lambda \end{vmatrix} = \lambda^2 + 4 = (\lambda - 2i)(\lambda + 2i) \overset{set}{=\!=} 0 \Rightarrow \lambda_{1,2} = \pm 2i$
are two simple eigenvalues. $\lambda_1 + \lambda_2 = 0 = \text{tr}(A)$ and $\lambda_1\lambda_2 = 4 = |A|$.

Comment: The eigenvalues of a real matrix might be complex numbers.

2. Eigenvectors

   (1) Eigenspace

   $$\begin{aligned} x \text{ is an eigenvector of } A \text{ wrt } \lambda_i &\Longleftrightarrow Ax = \lambda_i x,\ x \neq 0 \Longleftrightarrow (A - \lambda_i I)x = 0,\ x \neq 0 \\ &\Longleftrightarrow 0 \neq x \in \mathcal{N}(A - \lambda_i I) \overset{def}{=\!=} S_A(\lambda_i). \end{aligned}$$

   Every vector but 0 in the eigenspace $S_A(\lambda_i) = \mathcal{N}(A - \lambda_i I)$ is an eigenvector of A wrt λ_i.

   (2) Dimension of eigenspace

   Let $d_i = \dim[S_A(\lambda_i)] = \dim[\mathcal{N}(A - \lambda_i I)] = n - \text{rank}(A - \lambda_i I) \geq 1$.
   A has d_i LI eigenvectors wrt λ_i. It has d_i orthogonal eigenvectors wrt λ_i.
   It has d_i orthonormal eigenvectors wrt λ_i.
   So there exists $P_i \in C^{n\times d_i}$ with $P_i^* P_i = I_{d_i}$ and $AP_i = \lambda_i P_i$.

(3) $1 \le d_i \le r_i$.

Proof $A \in C^{n\times n}$ has eigenvalue λ_i with multiplicity r_i and $d_i = \dim[S_A(\lambda_i)] \ge 1$.

Let $P_I \in C^{n\times d_i}$ with $P_I^* P_I = I_{d_i}$ and $AP_I = \lambda_i P_I$. Non-singular $P = (P_I,\, P_{II})$ has $P^{-1} = (Q_I,\, Q_{II})'$. So $\begin{pmatrix} I_{d_i} & 0 \\ 0 & I \end{pmatrix} = Q'P = \begin{pmatrix} Q_I'P_I & Q_I'P_{II} \\ Q_{II}'P_I & Q_{II}'P_{II} \end{pmatrix}$.

$$\begin{aligned} |A-\lambda I| &= \left|P^{-1}\right| |A-\lambda I| \, |P| = \left|P^{-1}(AP_I,\, AP_{II}) - \lambda I_n\right| \\ &= \left| \begin{pmatrix} Q_I' \\ Q_{II}' \end{pmatrix} (\lambda_i P_I, AP_{II}) - \lambda I_n \right| = \begin{vmatrix} (\lambda_i - \lambda) I_{d_i} & Q_I' A P_{II} \\ 0 & Q_{II}' A P_{II} - \lambda I \end{vmatrix} \\ &= (\lambda_i - \lambda)^{d_i} \, |Q_{II}' A P_{II} - \lambda I_{n-d_i}| . \quad \text{So, } d_i \le r_i. \end{aligned}$$

Comment: $k \le \sum_{i=1}^k d_i \le n$.

Ex2: For $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $|A - \lambda I| = \lambda^2$ So $\lambda = 0$ with $r = 2$.

$d = \dim[\mathcal{N}(A - \lambda I)] = \dim[\mathcal{N}(A)] = 2 - \text{rank}(A) = 2 - 1 = 1 < 2 = r$.

3. Total number of LI eigenvectors

(1) Vectors from different eigenspaces

$\lambda_1, .., \lambda_k$ are distinct eigenvalues for $A \in C^{n\times n}$. Let $x_i \in S_A(\lambda_i) = \mathcal{N}(A - \lambda_i I)$.

Then $x_1 + \cdots + x_k = 0 \Longrightarrow x_1 = \cdots = x_k = 0$.

Proof We show by induction. $x_1 = 0 \Longrightarrow x_1 = 0$. So the statement holds for $k = 1$.

Assuming that the statement holds for $k = t - 1$, now consider $k = t$.

$$x_1 + \cdots + x_{t-1} + x_t = 0 \Longrightarrow \begin{cases} \lambda_t x_1 + \cdots + \lambda_t x_{t-1} + \lambda_t x_t &= 0 \\ \lambda_1 x_1 + \cdots + \lambda_{t-1} x_{t-1} + \lambda_t x_t &= 0 \end{cases}$$

$$\begin{aligned} &\Longrightarrow (\lambda_t - \lambda_1)x_1 + \cdots + (\lambda_t - \lambda_{t-1})x_{t-1} = 0 \\ &\Longrightarrow (\lambda_t - \lambda_1)x_1 = \cdots = (\lambda_t - \lambda_{t-1})x_{t-1} = 0 \\ &\Longrightarrow x_1 = \cdots = x_{t-1} = 0 \Longrightarrow x_1 = \cdots = x_{t-1} = x_t = 0 \end{aligned}$$

(2) If A has distinct eigenvalues $\lambda_1, ..., \lambda_k$, then A has $d_1 + \cdots + d_k$ LI eigenvectors.

Proof A has d_i LI eigenvectors $x_{i1}, .., x_{id_i}$ in $S_A(\lambda_i)$ wrt to λ_i.

We show $d_1 + \cdots + d_k$ eigenvectors $x_{11}, ..., d_{kd_k}$ are LI.

$$\begin{aligned} \textstyle\sum_{ij} \alpha_{ij} x_{ij} = 0 &\Longrightarrow y_1 + \cdots + y_k = 0 \text{ where } y_i = \textstyle\sum_j \alpha_{ij} x_{ij} \in S_A(\lambda_i) \\ &\Longrightarrow y_i = 0 \text{ for all } i \text{ by (1)} \\ &\Longrightarrow \alpha_{ij} = 0 \text{ for all } i,\, j \text{ since in } y_i \text{ is a LC of LI vectors.} \end{aligned}$$

(3) Conditions for $A \in C^{n\times n}$ to have n LI eigenvectors.

$$\begin{aligned} A \in C^{n\times n} \text{ has } n \text{ LI eigenvectors } \Longleftrightarrow d_1 + \cdots + d_k = n &\Longleftrightarrow d_i = r_i \text{ for all } i \\ &\Longleftrightarrow d_i = r_i \text{ for all } r_i > 1. \end{aligned}$$

If all eigenvalues for A are simple ones, then A has n LI eigenvectors.

Ex3: $A \in C^{2\times 2}$ in Ex1 has 2 LI eigenvectors since it has two simple eigenvalues.

$A \in C^{2\times 2}$ in Ex2 does not have 2 LI eigenvectors since $d = 1 < 2 = r$.

Comment: A may not have $d_1 + \cdots + d_k$ orthogonal eigenvectors since even if $x_{11}, ..., x_{1d_1}$ are orthogonal eigenvector wrt to λ_1, and $x_{21}, ..., x_{2d_2}$ are orthogonal eigenvectors wrt to λ_2, but $x_{11}, ..., x_{2d_2}$ may not be orthogonal.