L03: Eigenvalue decomposition

1. Eigenvalue decomposition

   (1) Eigenvalues and eigenvectors

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

   $$\begin{aligned} \lambda \text{ is an eigenvalue for } A &\Longleftrightarrow Ax = \lambda x,\, x \neq 0 \Longleftrightarrow (A-\lambda I)x = 0,\, x \neq 0 \\ &\Longleftrightarrow \text{The columns of } A - \lambda I \in R^{n\times n} \text{ are LD} \\ &\overset{*}{\Longleftrightarrow} |A - \lambda I| = 0 \\ &\Longleftrightarrow \lambda \text{ is a solution to the characteristic equation.} \end{aligned}$$

   With eigenvalue λ, $S_\lambda(A) \overset{def}{=\!=} \{x \in R^n :\, Ax = \lambda x\}$ is a subspace of R^n since

   $$\begin{aligned} x,\, y \in S_\lambda(A) &\Longrightarrow Ax = \lambda x \text{ and } Ay = \lambda y \Longrightarrow A(\alpha x + \beta y) = \lambda(\alpha x + \beta y) \\ &\Longrightarrow \alpha x + \beta y \in S_\lambda(A). \end{aligned}$$

   All vectors in $S_\lambda(A)$ but 0 are the eigenvectors of A wrt to eigenvalue λ. This space is called the eigen-space of A wrt to λ.

   (2) Eigenvalue decomposition

   For $A' = A \in R^{n\times n}$, there are P and Λ such that $A = P\Lambda P'$ called the EVD of A. Here $P \in R^{n\times n}$ with $P' = P^{-1}$ so that $P'P = I = PP'$, i.e., the columns and the rows of P are orthonormal. Such P is called an orthogonal matrix. Λ is a diagonal matrix. From

   $$A = P\Lambda P' \Longleftrightarrow AP = P\Lambda \Longleftrightarrow A(P_1, .., P_n) = (P_1, ..., P_n)\Lambda \Longleftrightarrow AP_i = \lambda_i P_i\, \forall\, i$$

   the diagonal elements of Λ are the eigenvalues of A, and columns of P are n orthonormal eigenvectors.

   (3) A is a LC of orthonormal matrices

   $A = P\Lambda P' = \sum_{i=1}^n \lambda_i P_i P_i'$ is a LC of matrices in $D = [P_1P_1', ..., P_nP_n'] \subset R^{n\times n}$. From $\langle P_iP_i',\, P_jP_j'\rangle = \text{tr}(P_jP_j'P_iP_i') = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, D is an orthonormal basis for Span(D). Clearly $A \in \text{Span}(D)$.

   (4) Compact form

   $\text{rank}(A) = \text{rank}(P\Lambda P') = \text{rank}(\Lambda) = \#$ of non-zero eigenvalues.

   With $\text{rank}(A) = r$, $\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix}$ where $\Lambda_r = \text{diag}(\lambda_1, .., \lambda_r)$, $\lambda_i \neq 0$, $i = 1, ..., r$.

   $A = (P_I,\, P_{II}) \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} (P_I,\, P_{II})' = P_I \Lambda_r P_I' = \sum_{i=1}^r \lambda_i P_i P_i'$ is the compact form of the EVD where $P_I \in R^{n\times r}$ has orthonormal columns.

   Ex1: Real symmetric matrix $A \in R^{n\times n}$ has real eigenvalues and n orthgonal eigenvectors.

   Ex2: A and B are similar if $A = XBX^{-1}$. Similar matrices share characteristic polynomial and hence share all eigenvalues.

   $|A - \lambda I| = |X(B - \lambda I)X^{-1}| = |X|\,|B - \lambda I|\,|X^{-1}| = |B - \lambda I|$. By EVD $A = P\Lambda P'$.

   In EVD $A = P\Lambda P'$, A and Λ are similar.

2. Definite and semi-definite matrices

   (1) Definitions and notations

   With $A' = A \in R^{n\times n}$, $x'Ax$ is a quadratic form of $x \in R^n$.

$$A \text{ is } \begin{cases} \text{positive definite denoted as } A > 0 \\ \text{non-negative definite denoted as } A \geq 0 \\ \text{negative definite denoted as } A < 0 \\ \text{non-positive definite denoted as } A \leq 0 \end{cases} \text{ if } x'Ax \begin{cases} > 0 \\ \leq 0 \\ < 0 \\ \leq 0 \end{cases} \text{ for all } 0 \neq x \in R^n$$

(2) A sufficient and necessary condition
Let $A = P\Lambda P'$ be the EVD for A where $\Lambda = \text{diag}(\lambda_1, ..., \lambda_n)$. Then
$A > 0 \Longleftrightarrow \lambda_i > 0$ for all i; $A \geq 0 \Longleftrightarrow \lambda_i \geq 0$ for all i;
$A < 0 \Longleftrightarrow \lambda_i < 0$ for all i; $A \leq 0 \Longleftrightarrow \lambda_i \leq 0$ for all i.

Proof. Only show the first one.
$\Rightarrow$: $A > 0 \Longrightarrow 0 < P_i'AP_i = P_i'P\Lambda P'P_i = e_i'\Lambda e_i = \lambda_i$ for al i.
$\Leftarrow$: $0 \neq x \in R^n \Longrightarrow y = P'x \neq 0 \Longrightarrow x'Ax = y'\Lambda y = \sum_i \lambda_i y_i^2 > 0$. So $A > 0$.

(3) $A^{1/2}$ and $A^{-1/2}$
For $A = P\Lambda P' \geq 0$, define $\Lambda^{1/2} = \text{diag}(\lambda_1^{1/2}, ..., \lambda_n^{1/2})$ and $A^{1/2} = P\Lambda^{1/2}P'$. Then $A^{1/2} \geq 0$ and $(A^{1/2})^2 = A$.
For $A = P\Lambda P' > 0$, define $\Lambda^{-1/2} = \text{diag}(\lambda_1^{-1/2}, ..., \lambda_n^{-1/2})$ and $A^{-1/2} = P\Lambda^{-1/2}P'$. Then $A^{-1/2} > 0$ and $(A^{-1/2})^2 = A^{-1}$.

(4) Simple relations
$A > 0 \Longleftrightarrow -A < 0$; $A \geq 0 \Longleftrightarrow -A \leq 0$; $A < 0 \Longleftrightarrow -A > 0$; $A \leq 0 \Longleftrightarrow -A \geq 0$.
Proof. $A > 0 \Longleftrightarrow x'Ax > 0$ for all $x \neq 0 \Longleftrightarrow x'(-A)x < 0$ for all $x \neq 0 \Longleftrightarrow -A < 0$.

3. Some properties

(1) Extended notations

$$\begin{aligned} A > B \Longleftrightarrow A - B > 0 \Longleftrightarrow B - A < 0 \Longleftrightarrow B < A \\ A \geq B \Longleftrightarrow A - B \geq 0 \Longleftrightarrow B - A \leq 0 \Longleftrightarrow B \leq A \end{aligned}$$

(2) $A \leq A$; $A \leq B$ and $B \leq C \Longrightarrow A \leq C$.
Proof. Show the second one. For $0 \neq x \in R^n$,
$x'(A - C)x = x'[(A - B) + (B - C)]x = x'(A - B)x + x'(B - C)x \leq 0$. So $A \leq C$.
Comments: $\leq$ is reflexive and transitive, and hence is a pre-order.

(3) $A_1 \leq B_1$ and $A_1 \leq B_2 \Longrightarrow \alpha A_1 + \beta B_1 \leq \alpha A_2 + \beta B_2$ for all α, $\beta \geq 0$.
Proof. Skipped. **Comment:** $\leq$ is preservable under LC with non-negative coefficients.

(4) $A \leq B \Longrightarrow CAC' \leq CBC'$.
Proof. For $0 \neq x \in R^n$, let $y = C'x$. Then $x'(CAC' - CBC')x = y'(A - B)y \leq 0$.
So $CAC' - CBC' \leq 0$, i.e., $CAC' \leq CBC'$.
Comment: In (2), (3) and (4) $\leq$ can be replaced by $\geq$.

Ex3: p64 1. $A' = A \in R^{n\times n}$, $B \in R^{n\times m}$ with $\text{rank}(B) = n$ such that $(BB')^{-1}$ exists.
Show that $I_m \geq B'AB \Longleftrightarrow (BB')^{-1} \geq A$.
⇒: $I_m \geq B'AB \Longrightarrow [(BB')^{-1}B]I_m[(BB')^{-1}B]' \leq [(BB')^{-1}B]B'AB[(BB')^{-1}B]'$
$\Longrightarrow (BB')^{-1} \geq A$.
⇐: $(BB')^{-1} \geq A \Longrightarrow D = B'(BB')^{-1}B \geq B'AB$. $D' = D = D^2 = P\Lambda P'$ by EVD.
$D^2 = D \Longrightarrow \Lambda^2 = \Lambda \Longrightarrow \lambda_i = 0, 1 \Longrightarrow I \geq \Lambda \Longrightarrow I = PP' \geq P\Lambda P' = D$.
So $I \geq D \geq B'AB$, i.e., $I \geq B'AB$ by transitivity.

Notes: In p64 1, $A > 0$ is not needed. In p65 5, $\lambda_i > 0$ should be $\lambda_i \neq 0$.

L04 Singular value decomposition and generalized inverse

1. Singular value decomposition

   (1) Singular value decomposition

   For $A \in R^{m\times n}$ with $\text{rank}(A) = r$, $A = (U_I, U_{II}) \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} (V_I, V_{II})' = U_I \Delta V_{II}'$ is the SVD where $U = (U_I, U_{II}) \in R^{m\times m}$ and $V = (V_I, V_{II}) \in R^{n\times n}$ are orthogonal ($U' = U^{-1}$ and $V' = V^{-1}$), $U_I \in R^{m\times r}$ and $V_I \in R^{n\times r}$ ($U_I'U_I = I_r = V_I'V_{II}$), $\begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} \in R^{m\times n}$ with $\Delta = \text{diag}(\delta_1, ..., \delta_r)$, $\delta_1 \geq \cdots \geq \delta_r > 0$.

   (2) Singular values

   $0 \leq AA' \in R^{m\times m}$, $0 \leq A'A \in R^{n\times n}$ and A share ranks r.

   The EVD $AA' = (U_I, U_{II}) \begin{pmatrix} \Delta^2 & 0 \\ 0 & 0 \end{pmatrix} (U_I, U_{II})' = U_I \Delta^2 U_I'$ and

   the EVD $A'A = (V_I, V_{II}) \begin{pmatrix} \Delta^2 & 0 \\ 0 & 0 \end{pmatrix} (V_I, V_{II})' = V_I \Delta^2 V_I'$ share positive eigenvalues $\delta_1^2 \geq \cdots \geq \delta_r^2 > 0$ in $\Delta^2 = \text{diag}(\delta_1^2, ..., \delta_r^2)$.

   $\delta_1 \geq \cdots \geq \delta_r > 0$ are called the positive singular values of A and $\Delta = \text{diag}(\delta_1, ..., \delta_r)$ can be obtained by either EVD of AA' or EVD of $A'A$.

   (3) Construct SVD for A

   (i) Method I

   By EVD for AA', $AA' = (U_I\, U_{II}) \begin{pmatrix} \Delta^2 & 0 \\ 0 & 0 \end{pmatrix} (U_I, U_{II})'$. Let $V_I = A'U_I\Delta^{-1}$. Then $V_I'V_I = I_r$ and $A = U_I \Delta V_I'$.

   (ii) Method II

   By EVD for $A'A$, $A'A = (V_I, V_{II}) \begin{pmatrix} \Delta^2 & 0 \\ 0 & 0 \end{pmatrix} (V_I\, V_{II})'$. Let $U_I = AV_I\Delta^{-1}$. Then $U_I'U_I = I_r$ and $A = U_I \Delta V_I'$.

   (iii) Comments

   In SVD Δ is unique. But U_I and V_I are not unique. Specifically V_I in (i) may not be V_I in (ii), and U_I in (ii) may not be U_I in (i)

   (4) A LC of orthonormal matrices

   By SVD $A = U_I \Delta V_I' = \sum_{i=1}^r \delta_i U_i V_i'$ is a LC of $U_1V_1', \cdots U_rV_r'$, a set of r orthonormal matrices since

$$\langle U_iV_i', U_jV_j' \rangle = \text{tr}(V_jU_j'U_iV_i) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

2. Generalized inverse

   (1) Generalized inverse A^- and Moore-Penrose inverse A^+

   For $A \in R^{m\times n}$ the conditions on $G \in R^{n\times m}$

$$\text{(i): } AGA = A, \text{ (ii): } GAG = G, \text{ (iii): } (AG)' = AG, \text{ (iv): } (GA)' = GA$$

   are Penrose conditions. Matrix G satisfying (i) is called a generalized inverse of A denoted by A^-. Matrix G satisfying all four conditions is Moore-Penrose inverse of A denoted by A^+.

(2) Generalized inverses A^-

Let $A = U\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}V'$ be the SVD for A. Then

$$A^- = \left\{V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U' : H_{12},\, H_{21},\, H_{22}\right\}.$$

Proof. $\subset$**:** $G \in A^- \Longrightarrow AGA = A$. Let $V'GU = \begin{pmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}$. Then $ABA = A$ implies that $\Delta H_{11}\Delta = \Delta$. Hence $H_{11} = \Delta^{-1}$. Thus $G = V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U'$.

$\supset$**:** With $G = V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U'$, direct computation shows $AGA = A$. So $G \in A^-$.

(3) Moore-Penrose inverse A^+

With the SVD form for A, $A^+ = V\begin{pmatrix}\Delta^{-1} & 0\\ 0 & 0\end{pmatrix}U'$.

Proof. $AGA = A \Longleftrightarrow G = V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U'$. Now

$$\begin{aligned} GAG = G,\ (AG)' = AG,\ (GA)' = GA \quad &\Longleftrightarrow \quad H_{12} = 0,\ H_{21} = 0,\ H_{22} = 0\\ &\Longleftrightarrow \quad G = V\begin{pmatrix}\Delta^{-1} & 0\\ 0 & 0\end{pmatrix}U. \end{aligned}$$

(3) Comment: $A^+ \in A^-$.

3. Some Moore-Penrose inverses

(1) $0^+_{m\times n} = 0_{n\times m}$, $(A')^+ = (A^+)'$ For $0 \neq x \in R^n$, $x^+ = \frac{x'}{x'x}$ and $(x')^+ = \frac{x}{x'x}$.

Ex1: $(AB)^+ \neq B^+A^+$ example

$$\left[(1,\,0)\begin{pmatrix}1\\2\end{pmatrix}\right]^+ = 1^+ = 1.\ \text{ But } \begin{pmatrix}1\\2\end{pmatrix}^+(1,\,0)^+ = \frac{(1,2)}{5}\begin{pmatrix}1\\0\end{pmatrix} = \tfrac{1}{5} \neq 1.$$

(2) $(A'A)^+ = A^+(A')^+$ and $(AA')^+ = (A')^+A^+$

Proof. For $(A'A)^+ = A^+(A')^+$ check (iv).

$A^+(A')^+A'A = A^+(A^+)'A'A = A^+(AA^+)'A = A^+AA^+A = A^+A$ is symmetric.

(3) Simplification: $(A'A)^+A' = A^+$ and $A'(AA')^+ = A^+$

Proof. Only show the first one.

$(A'A)^+A' = A^+(A')^+A' = A^+(A^+)'A' = A^+(AA^+)' = A^+AA^+ = A^+$.

Ex2: $A(A'A)^+A' = AA^+$ and $A'(AA')^+A = A^+A$.

(4) $(A'A)^-$, $(AA')^-$, $(A'A)^-A'$ and $A'(AA')^-$ exist but may not be unique.
But $A(A'A)^-A' = AA^+$ and $A'(AA')^-A = A^+A$.

Proof Show the first one only

$$\begin{aligned} A(A'A)^-A' &= AA^+A(A'A)^-(AA^+A)' = (AA^+)'A(A'A)^-A'(AA^+)'\\ &= (A')^+(A'A)(A'A)^-(A'A)A^+ = (A')^+(A'A)A^+ = (A^+)'A'AA^+\\ &= (AA^+)'AA^+ = AA^+AA^+ = AA^+ \end{aligned}$$