L11 Singular values

1. Eigenvalue decomposition of Hermitian matrices

   (1) Orthonormal matrices
   For $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ $\langle A,\, B\rangle = \sum_{ij} \bar{b}_{ij} a_{ij} = \mathrm{tr}(B^*A)$ is an inner product.
   With this inner product

   A is a unit matrix $\overset{def}{\Longleftrightarrow} \|A\| = 1 \Longleftrightarrow \|A\|^2 = 1 \Longleftrightarrow \langle A,\, A\rangle = 1$.

   A and B are orthogonal $\overset{def}{\Longleftrightarrow} \langle A,\, B\rangle = 0$ but $A \neq 0$ and $B \neq 0$.

   Matrices $M_1, ..., M_n$ are orthonormal$\overset{def}{\Longleftrightarrow} \langle M_i,\, M_j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

   Ex1: If $U = (u_1, ..., u_n) \in C^{n\times n}$ is unitary, let $M_i = u_i u_i^* \in C^{n\times n}$, $i = 1, ..., n$.
   Then (i) $M_i^* = M_i = M_i^2$ (ii) $\mathrm{rank}(M_i) = 1$ (iii) $M_1, ..., M_n$ are orthonormal.
   (i) $M_i^* = (u_i u_i^*)^* = u_i u_i^* = M_i$; $M_i M_i = (u_i u_i^*)(u_i u_i^*) = u_i u_i^* = M_i$;
   (ii) $\mathrm{rank}(M_i) = \mathrm{tr}(M_i) = \mathrm{tr}(u_i u_i^*) = \mathrm{tr}(u_i^* u_i) = \mathrm{tr}(1) = 1$
   (iii) $\langle M_i,\, M_i\rangle = \mathrm{tr}(M_i^* M_i) = \mathrm{tr}(M_i) = 1$; $\langle M_i,\, M_j\rangle = \mathrm{tr}(u_j u_j^* u_i u_i^*) = \mathrm{tr}(0) = 0$.

   (2) Spectral decomposition
   EVD of $A = A^* \in C^{n\times n}$ expresses A as a LC of n orthonormal matrices
   $A = U\Lambda U^* = (u_1, .., u_n)\mathrm{diag}(\lambda_1, ..., \lambda_n)(u_1, ..., u_n)^* = \lambda_1(u_1u_1^*) + \cdots + \lambda_n(u_nu_n^*)$.
   This expression is also called a spectral decomposition of A.

   (3) Compact form of EVD
   If $A^* = A \in C^{n\times n}$ has rank r, then $A = U_I \Lambda_r U_I^*$ where $U_I \in C^{n\times r}$ with $U_I^* U_I = I_r$ and $\Lambda_r = \mathrm{diag}(\lambda_1, .., \lambda_r)$ with $\lambda_i \neq 0$ for $i = 1, ..., r$.
   If $A = U\Lambda U^*$ has rank r, then Λ has rank r. So $\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix}$ where $\Lambda_r = \mathrm{diag}(\lambda_1, .., \lambda_r)$ with $\lambda_i \neq 0$ for $i = 1, .., r$. Thus $A = (U_I,\, U_{II}) \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_I^* \\ U_{II}^* \end{pmatrix} = U_I \Lambda_r U_I^*$.

   Ex2: With $A = U_I \Lambda_r U_I^*$, by expanding U_I to unitary $U = (U_I,\, U_{II})$ and Λ_r to $\begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix}$

   $A = U_I \Lambda_r U_I^* = U\Lambda U^*$. This is a full form of EVD of A.

2. Definite and semi-definite matrices

   (1) Recall definitions
   $A^* = A \in C^{n\times n}$.
   $A > 0 \overset{def}{\Longleftrightarrow} x^*Ax > 0$ for all $0 \neq x \in C^n$; $A < 0 \overset{def}{\Longleftrightarrow} x^*Ax < 0$ for all $0 \neq x \in C^n$
   $A \geq 0 \overset{def}{\Longleftrightarrow} x^*Ax \geq 0$ for all $0 \neq x \in C^n$; $A \leq 0 \overset{def}{\Longleftrightarrow} x^*Ax \leq 0$ for all $0 \neq x \in C^n$.

   (2) Iff-conditions
   $A^* = A = X\Lambda X^*$ where $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ and X is unitary by EVD.
   $A > 0 \Longleftrightarrow \lambda_i > 0$ for all i $A < 0 \Longleftrightarrow \lambda_i < 0$ for all i
   $A \geq 0 \Longleftrightarrow \lambda_i \geq 0$ for all i $A \leq 0 \Longleftrightarrow \lambda_i \leq 0$ for all i.

   Proof Only show $A > 0 \Longleftrightarrow \lambda_i > 0$ for all i

   $\Rightarrow$: $A = X\Lambda X^* > 0$ where $X = (x_1, .., x_n)$ is unitary and $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$.
   Then $0 < x_i^* A x_i = x_i^* X\Lambda X^* x_i = e_i'\Lambda e_i = \lambda_i$ for all i.

   $\Leftarrow$: $z \neq 0$, $z^*Az = (z^*X)\Lambda(X^*z) \overset{y=X^*z}{=\!=\!=} y^*\Lambda y = \sum_i \lambda_i |y_i|^2$.
   $z \neq 0 \Longrightarrow y = X^*z \neq 0 \Longrightarrow \sum_i \lambda_i |y_i|^2 > 0$. So $A > 0$.

(3) Application: Definition of $A^{1/2}$ and $A^{-1/2}$.
For $\Lambda = \text{diag}(\lambda_1, ..., \lambda_n)$ with $\lambda_i \ge 0$ for all i, define $\Lambda^{1/2} = \text{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})$;
For $\Lambda = \text{diag}(\lambda_1, ..., \lambda_n)$ with $\lambda_i > 0$ for all i, define $\Lambda^{-1/2} = \text{diag}(1/\sqrt{\lambda_i}, ..., 1/\sqrt{\lambda_n})$.
For $0 \le A = X\Lambda X^*$ define $A^{1/2} = X\Lambda^{1/2}X^*$. Then $A^{1/2} \ge 0$ and $A^{1/2}A^{1/2} = A$.
For $0 < A = X\Lambda X^*$ define $A^{-1/2} = X\Lambda^{-1/2}X^*$. Then $A^{-1/2} > 0$, $A^{-1/2}A^{-1/2} = A^{-1}$.

3. Singular values of $A \in C^{m\times n}$

(1) Eigenvalues of AA^* and A^*A
For $A \in C^{m\times n}$, all eigenvalues of $AA^* \in C^{m\times m}$ and $A^*A \in C^{n\times n}$ are ≥ 0.

Proof By EVD for Hermitian AA^* and A^*A, $AA^* = U\Lambda U^*$ and $A^*A = V\Gamma V^*$.
But $AA^* \ge 0$ and $A^*A \ge 0$ since $x^*AA^*x = \|A^*x\|^2 \ge 0$ for all $x \in C^m$ and $y^*A^*Ay = \|Ay\|^2 \ge 0$ for all $y \in C^n$. So all diagonal elements of Λ and Γ are ≥ 0.

(2) Number of positive eigenvalues of AA^* and A^*A
For $A \in C^{m\times n}$ with rank r, AA^* and A^*A have r positive eigenvalues.

Proof (3) of 3 in L07 shows that $\text{rank}(AA^*) = \text{rank}(A^*A) = \text{rank}(A) = r$.
Thus in EVDs $AA^* = U\Lambda U^*$ and $A^*A = V\Gamma V^*$, $\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} \in R^{m\times m}$ and $\Gamma = \begin{pmatrix} \Gamma_r & 0 \\ 0 & 0 \end{pmatrix} \in R^{n\times n}$, and all diagonal elements of $\Lambda_r \in R^{r\times r}$ and $\Gamma_r \in R^{r\times r}$ are positives.

(3) Distinct positive eigenvalues of AA^* and A^*A
AA^* and A^*A share the same set of distinct eigenvalues

Proof For "$\delta^2 > 0$ is an eigenvalue for $AA^* \Longleftrightarrow \delta^2 > 0$ is an eigenvalue for A^*A"
only show $\Rightarrow$: If $\delta^2 > 0$ is an eigenvalue for AA^*, then $AA^*x = \delta^2 x$, $x \ne 0$.
So $(A^*A)(A^*x) = \delta^2(A^*x)$. Here $A^*x \ne 0$ since otherwise

$$A^*x = 0 \Longrightarrow 0 = AA^*x = \delta^2 x \Longrightarrow x = 0 \quad \otimes .$$

Thus $\delta^2 > 0$ is an eigenvalue for A^*A.

(4) $\delta^2 > 0$ is an eigenvalue for AA^* and A^*A with multiplicity r and s. Then $r = s$.

Pf: $\delta^2 > 0$ is an eigenvalue of AA^* with multiplicity r. There exists $B \in C^{m\times r}$ with $B^*B = I_r$ such that $(AA^*)B = \delta^2 B$. So $(A^*A)(A^*B) = \delta^2(A^*B)$ where $A^*B \in C^{n\times r}$ with $(A^*B)^*(A^*B) = B^*AA^*B = B^*\delta^2 B = \delta^2 I_r$. Thus $r \le s$.
$\delta^2 > 0$ is an eigenvalue of A^*A with multiplicity s. There exists $D \in C^{n\times s}$ with $D^*D = I_s$ such that $(A^*A)D = \delta^2 D$. So $(AA^*)(AD) = \delta^2(AD)$ where $(AD)^*(AD) = D^*A^*AD = D^*\delta^2 D = \delta^2 I_s$. Thus $s \le r$. It follows that $r = s$.

(5) Singular values of A
For $A \in C^{m\times n}$ with rank r, $AA^* \in C^{m\times m}$ and $A^*A \in C^{n\times n}$ share positive-eigenvalue matrix $\Delta^2 = \text{diag}(\delta_1^2, ..., \delta_r^2)$ with $\delta_1 \ge \cdots \ge \delta_r > 0$. Here $\delta_1 \ge \cdots \ge \delta_r > 0$ are positive eigenvalues for A and Δ is the positive-singular value matrix.

Ex3: Find positive-singular matrix for $A = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Two methods.

$AA' = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \Longrightarrow |AA' - \lambda I| = \lambda(\lambda - 5) \Longrightarrow \delta^2 = 5 \Longrightarrow \Delta^2 = 5,\ \Delta = \sqrt{5}.$
$A'A = 5 \Longrightarrow \delta^2 = 5 \Longrightarrow \Delta^2 = 5,\ \Delta = \sqrt{5}.$

L12 Singular value decomposition and generalized inverse

1. Singular value decomposition

   (1) Singular value decomposition (SVD)
   For $A \in C^{m\times n}$ with $\text{rank}(A) = r$

   (i) Full form: $A = U\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}V^*$ where $U \in C^{m\times m}$ and $V \in C^{n\times n}$ are unitary and in $\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix} \in R^{m\times n}$ $\Delta = \begin{pmatrix}\delta_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \delta_r\end{pmatrix}$ with $\delta_1 \geq \cdots \geq \delta_r > 0$ is positive-singular value matrix for A.

   (ii) Compact form: $A = U_I\Delta V_I^*$ where $U_I \in C^{m\times r}$, $V_I \in C^{n\times r}$, $U_I^*U_I = I_r = V_I^*V_I$, and $\Delta \in R^{r\times r}$ is positive-singular value matrix for A.

   (2) Constructive proof approach I
   By EVD $AA^* = U\begin{pmatrix}\Delta^2 & 0\\ 0 & 0\end{pmatrix}U^* = (U_I,\, U_{II})\begin{pmatrix}\Delta^2 & 0\\ 0 & 0\end{pmatrix}(U_I,\, U_{II})^* = U_I\Delta^2U_I^*$.
   Let $V_I = A^*U_I\Delta^{-1}$. Then $V_I^*V_I = \Delta^{-1}U_I^*AA^*U_I\Delta^{-1} = \Delta^{-1}U_I^*U_I\Delta^2U_I^*U_I\Delta^{-1} = I_r$.

$$\begin{aligned} A &= UU^*A = U(A^*U)^* = U(A^*U_I,\, A^*U_{II})^* = U(V_I\Delta,\, A^*U_{II})^* \\ &= (U_I,\, U_{II})\begin{pmatrix}\Delta V_I^*\\ U_{II}^*A\end{pmatrix} = U_I\Delta V_I^* + U_{II}U_{II}^*A. \end{aligned}$$

   But $U_{II}^*A = 0$ since $(U_{II}^*A)(U_{II}^*A)^* = U_{II}^*AA^*U_{II} = U_{II}^*U_I\Delta^2U_I^*U_{II} = 0$.
   Hence $A = U_I\Delta V_I^*$. The compact form of SVD holds.
   Expand V_I to unitary $V = (V_I,\, V_{II})$. Then $A = U_I\Delta V_I^* = (U_I,\, U_{II})\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}(V_I,\, V_{II})^*$.
   The full form of SVD holds.

   (3) Constructive proof approach II
   By EVD, $A^*A = V\begin{pmatrix}\Delta^2 & 0\\ 0 & 0\end{pmatrix}V^* = (V_I\, V_{II})\begin{pmatrix}\Delta^2 & 0\\ 0 & 0\end{pmatrix}(V_I,\, V_{II})^* = V_I\Delta^2V_I^*$.
   Let $U_I = AV_I\Delta^{-1}$. Then $U_I^*U_I = \Delta^{-1}V_I^*A^*AV_I\Delta^{-1} = \Delta^{-1}V_I^*V_I\Delta^2V_I^*VI\Delta^{-1} = I_r$.

$$\begin{aligned} A &= AVV^* = (AV_I,\, AV_{II})V^* = (U_I\Delta,\, AV_{II})V^* = (U_I\Delta,\, AV_{II})\begin{pmatrix}V_I^*\\ V_{II}^*\end{pmatrix} \\ &= U_I\Delta V_I^* + AV_{II}V_{II}^*. \end{aligned}$$

   But $AV_{II} = 0$ since $(AV_{II})^*(AV_{II}) = V_{II}^*A^*AV_{II} = V_{II}^*V_I\Delta^2V_I^*V_{II} = 0$.
   Hence $A = U_I\Delta V_I^*$. The compact form of SVD holds.
   Expand U_I to unitary $(U_I,\, U_{II})$. Then $A = U_I\Delta V_I^* = (U_I,\, U_{II})\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}(V_I,\, V_{II})^*$.
   The full form of SVD holds.

   (4) Comment:
   For SVD Δ and $\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}$ are unique. But U, U_I, V and V_I are not unique.

Ex1: If A has full row rank, then by SVD $A = U(\Delta,\, 0)V^* = U\Delta V_I^*$.
If A has full column rank, then by SVD $A = U\begin{pmatrix}\Delta\\ 0\end{pmatrix}V^* = U_I\Delta V^*$.

2. Generalized inverse

   (1) Definition
   For $A \in C^{m\times n}$ $G \in C^{n\times m}$ is a generalized inverse of A denoted as A^- if $AGA = A$.

   (2) Existence, non-uniqueness and expressions
   If $A = U\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}V^*$, then $A^- = \left\{V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U^* : H_{12},\, H_{21} \text{ and } H_{22}\right\}$.

   Proof $G \in A^-$ $\iff$ $AGA = A \iff \begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}V^*GU\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}$

   $$\overset{V^*GU=H}{\iff} \begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}$$

   $$\iff \Delta H_{11}\Delta = \Delta \iff H_{11} = \Delta^{-1} \iff V^*GU = \begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}$$

   $$\iff G = V\begin{pmatrix}\Delta^{-1} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}U^*$$

   (3) AA^- and A^-A are idempotent.
   $(AA^-)^2 = (AA^-A)A^- = AA^-$ and $(A^-A)^2 = A^-(AA^-A) = A^-A$.
   A^- represents the set of all generalized inverses of A, or any individual $G \in A^-$.

Ex2: A has full row rank$\Longrightarrow A = U(\Delta,\, 0)V^* \Longrightarrow A^- = V\begin{pmatrix}\Delta^{-1}\\ H_{21}\end{pmatrix}U^*$.

   A has full column rank $\Longrightarrow A = U\begin{pmatrix}\Delta\\ 0\end{pmatrix}V^* \Longrightarrow A^- = V(\Delta^{-1},\, H_{12})U^*$.

   $A \in C^{n\times n}$ is nonand $\text{rank}(A) = n \Longrightarrow A = U\Delta V^* \Longrightarrow A^- = V\Delta^{-1}U^*$.

3. Generalized inverse and ordinary inverse

   (1) Left-inverse and A^-
   A^L is the collection of all left-inverses of A. If $A^L \neq \emptyset$, then $A^L = A^-$.

   Proof $\subset$**:** $G \in A^L \neq \emptyset \Longrightarrow GA = I \Longrightarrow AGA = A \Longrightarrow G \in A^-$.

   $\supset$**:** $A^L \neq \emptyset \Longrightarrow A$ has full column rank $\Longrightarrow$ By SVD $A = U\begin{pmatrix}\Delta\\ 0\end{pmatrix}V^*$.

   $G \in A^- \Longrightarrow G = V(\Delta^{-1},\, H_{12})U^* \Longrightarrow GA = I \Longrightarrow G \in A^L$.

   (2) Right-inverse and A^-
   A^R is the collection of all right-inverses of A. If $A^R \neq \emptyset$, then $A^R = A^-$.

   Proof $\subset$**:** $G \in A^R \neq \emptyset \Longrightarrow AG = I \Longrightarrow AGA = A \Longrightarrow G \in A^-$.

   $\supset$**:** $A^R \neq \emptyset \Longrightarrow A$ full row rank $\Longrightarrow$ By SVD $A = U(\Delta,\, 0)V^*$.

   $G \in A^- \Longrightarrow G = V\begin{pmatrix}\Delta^{-1}\\ H_{21}\end{pmatrix}U^* \Longrightarrow AG = I \Longrightarrow G \in A^R$.

   (3) Inverse and A^-
   Suppose A has inverse A^{-1}. Then $A^{-1} = A^-$.

   Proof $AA^{-1}A = A$. So $A^{-1} \in A^-$.
   In Ex2 we see $A = U\Delta V^*$ and $A^- = V\Delta^{-1}U^*$. So $AA^- = A^-A = I$.
   Hence A^- contains only one matrix that is A^{-1}.