L06 Unitary, Hermitian and Idempotent matrices

1 Unitary matrices

(1) Frobenius inner product
$\langle x, y\rangle$ is an inner product if (i) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ (ii) $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ (iii) $\langle x, x\rangle \geq 0$ for all x and $\langle x, x\rangle = 0 \Longleftrightarrow x = 0$.
For $x, y \in C^n$, $\langle x, y\rangle = y^*x$, for $x, y \in R^n$, $\langle x, y\rangle = y'x$ are Frobenius inner product.

(2) Norm and angle
$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x^*x}$ is the norm of x; So $x \in C^n$ is a unit vector $\Longleftrightarrow x^*x = 1$.
θ is the angle formed by $x \neq 0$ and $y \neq 0 \overset{def}{\Longleftrightarrow} \cos(\theta) = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}$.
x and y are orthogonal $\Longleftrightarrow x \perp y \Longleftrightarrow \langle x, y\rangle = 0 \Longrightarrow \|x \pm y\|^2 = \|x\|^2 + \|y\|^2$.

(3) Interpretation of A^*B
For $A = (A_1, ..., A_p) \in C^{m\times p}$ and $B = (B_1, ..., B_q) \in C^{m\times q}$ $A^*B = (A_i^*B_j)_{p\times q}$ is the matrix of inner products of the columns of A and that of B.
$A^*B = 0 \Longleftrightarrow B^*A = 0 \Longleftrightarrow$ The columns of A are perpendicular to the columns of B

(4) Matrix with orthonormal columns
For $A = (A_1, ..., A_p) \in C^{m\times p}$, $A^*A = (A_i^*A_j)_{p\times p} = (\langle A_j\, A_i\rangle)_{p\times p}$. So
$A^*A = \text{diag}(d_1, ..., d_p) \Longleftrightarrow A_i \perp A_j$ for all $i \neq j$.

$$\begin{aligned} A^*A = \text{diag}(d_1, ..., d_p),\ d_i > 0\,\forall\, i \quad &\Longleftrightarrow \quad A_i \neq 0 \text{ for all } i,\ A_i \perp A_j \text{ for all } i \neq j \\ &\overset{def}{\Longleftrightarrow} \quad A \text{ has orthogonal columns} \\ &\Longrightarrow \quad A \text{ has full column rank.} \end{aligned}$$

$A^*A = I_p \Longleftrightarrow \|A_i\| = 1$ for all i, $A_i \perp A_j$ for all $i \neq j \Longleftrightarrow A$ has orthonormal columns

(5) Unitary matrices
$A \in C^{n\times n}$.

$$\begin{aligned} A \text{ is unitary} \quad &\overset{def}{\Longleftrightarrow} \quad A^{-1} = A^* \Longleftrightarrow A^*A = I_n \Longleftrightarrow A \text{ has orthonormal columns} \\ &\Longleftrightarrow \quad AA^* = I \Longleftrightarrow (A')^*A' = I_n \Longleftrightarrow A \text{ has orthonormal rows} \end{aligned}$$

$A \in R^{n\times n}$.

$$\begin{aligned} A \text{ is unitary} \quad &\Longleftrightarrow \quad A^{-1} = A' \Longleftrightarrow A'A = I_n \Longleftrightarrow A \text{ has orthonormal columns} \\ &\Longleftrightarrow \quad AA' = I \Longleftrightarrow A \text{ has orthonormal rows} \overset{def}{\Longleftrightarrow} A \text{ is orthogonal} \end{aligned}$$

Ex1: Show that if A has orthogonal columns, then A has full column rank.
$Ax = 0 \Longrightarrow A^*Ax = 0 \Longrightarrow \text{diag}(d_1, .., d_p)x = 0 \Longrightarrow x = [\text{diag}(d_1, .., d_p)]^{-1}0 = 0$.

2. Hermitian matrices

(1) Hermitian matrix and quadratic form
$A \in C^{n\times n}$
A is Hermitian $\overset{def}{\Longleftrightarrow} A^* = A \overset{def}{\Longleftrightarrow} z^*Az$ is a quadratic form of $z \in C^n \Longrightarrow z^*Az$ is real.
$A \in R^{n\times n}$

$$\begin{aligned} A \text{ is Hermitian} \quad &\Longleftrightarrow \quad A^* = A \Longleftrightarrow A' = A \Longleftrightarrow A \text{ is real symmetric} \\ &\Longleftrightarrow \quad x'Ax \text{ is a quadratic form of } x \in R^n. \end{aligned}$$

Ex2: Show that if $A^* = A$, then z^*Az is real.
$\overline{z^*Az} = (z^*Az)^* = z^*A^*z = z^*Az \Longrightarrow z^*Az$ is real.

(2) Definite and semi-definite matrices

$A > 0$ A is positive definite $\overset{def}{\Longleftrightarrow}$ $A = A^*$ and $z^*Az > 0 \,\forall\, 0 \neq z \in C^n$

$A \geq 0$ A is semi p.d. $\overset{def}{\Longleftrightarrow}$ $A = A^*$ and $z^*Az \geq 0 \,\forall\, 0 \neq z \in C^n$

$A < 0$ A is negative definite $\overset{def}{\Longleftrightarrow}$ $A = A^*$ and $z^*Az < 0 \,\forall\, 0 \neq z \in C^n$

$A \leq 0$ A is semi n. d. $\overset{def}{\Longleftrightarrow}$ $A = A^*$ and $z^*Az \leq 0 \,\forall\, 0 \neq z \in C^n$.

Ex2: By the definition one can show that $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} > 0$; $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \geq 0$. But But $C = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$ is not a definite nor semi-definite matrix.

(3) Notation extension

$A > 0 \Longleftrightarrow 0 < A$, $A \geq 0 \Longleftrightarrow 0 \leq A$, $A < 0 \Longleftrightarrow 0 > A$, $A \leq 0 \Longleftrightarrow 0 \geq A$.

$A > B \Longleftrightarrow A - B > 0$, $A \geq B \Longleftrightarrow A - B \geq 0$, $A < B \Longleftrightarrow A - B < 0$ and $A \leq 0 \Longleftrightarrow A - B \leq 0$.

(4) Properties

$>$, $\geq$, $<$ and $\leq$ are all transitive. For example, $A > B$ and $B > C \Longrightarrow A > C$.

$>$, $\geq$, $<$ and $\leq q$ are all preserved under scalar multiplication with positive scalar. For example $A < 0 \Longrightarrow aA < 0$ for all $a > 0$

But with $a < 0$, $A > 0 \Longrightarrow aA < 0$, $A \geq 0 \Longrightarrow aA \leq 0$, $A < 0 \Longrightarrow aA > 0$ and $A \leq 0 \Longrightarrow aA \geq 0$.

(5) Properties for $\geq$ and $\leq$

$\geq$ and $\leq$ are reflexive, i.e., $A \geq A$ and $A \leq A$ for all $A = A^*$

$\geq$ and $\leq$ are preserved under scalar multiplication with non-negative scalars. For example $A \leq 0 \Longrightarrow aA \leq 0$ for all $a \geq 0$.

Ex3: Show $A \geq 0 \Longrightarrow BAB^* \geq 0$.

For vector z, let $y = B^*z$. Then $z^*BAB^*z = y^*Ay \geq 0$. So $BAB^* \geq 0$.

3. Idempotent matrices

(1) A decomposition based on rank

For $A \in C^{m\times n}$ with $\text{rank}(A) = r$ there exist non-singular $P \in C^{m\times m}$ and $Q \in C^{n\times n}$ such that $A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q'$.

Proof: $A \overset{e.r.}{\longrightarrow} \begin{pmatrix} H \\ 0 \end{pmatrix}$, $(H', 0) \overset{e.r.}{\Longrightarrow} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \Longrightarrow P^{-1}A = \begin{pmatrix} H \\ 0 \end{pmatrix}$, $Q^{-1}(H', 0) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

$\Longrightarrow A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q'$.

Compact form: $A = P_I Q_I'$ where $P_I \in C^{m\times r}$, $Q_I \in C^{n\times r}$, $\text{rank}(P_I) = \text{rank}(Q_I) = r$.

(2) Idempotent matrices: $A \in C^{n\times n}$ is idempotent $\overset{def}{\Longleftrightarrow}$ $A^2 = A$.

(3) If A is idempotent, then $\text{rank}(A) = \text{tr}(A)$.

Pf: A with $\text{rank}(A) = r$ has decomposition $A = P_I Q_I'$ based on rank r.

$A^2 = A \Longrightarrow P_I Q_I' P_I Q_I' = P_I Q_I' \Longrightarrow Q_I' P_I = I_r$.

So $\text{tr}(A) = \text{tr}(P_I Q_I') = \text{tr}(Q_I' P_I) = \text{tr}(I_r) = r = \text{rank}(A)$.

L07 QR-decomposition and decompositions based on ranks

1. Gram-Schmidt process

   (1) From LI vectors to orthogonal vectors
   For LI $x_i \in C^n$, $i = 1, ..., r$, there are orthogonal $y_i \in C^n$, $i = 1, ..., r$ such that x_i is a LC of $y_1, .., y_i$, $i = 1, ..., r$.
   Process: Let $y_i \propto x_i - \frac{\langle x_i, y_1 \rangle}{\|y_1\|^2} y_1 - \cdots - \frac{\langle x_i, y_{i-1} \rangle}{\|y_{i-1}\|^2} y_{i-1}$, $i = 1, ..., r$.
   Proof. Skipped
   Ex1: $x_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ and $x_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ are LI.

   $y_1 \propto x_1$. Let $y_1 = x_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $x_1 = y_1$;

   $y_2 \propto x_2 - \frac{\langle x_2, y_1 \rangle}{\|y_1\|^2} y_1 = x_2$. Let $y_2 = x_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $x_2 = y_2$;

   $y_3 \propto x_3 - \frac{\langle x_3, y_1 \rangle}{\|y_1\|^2} y_1 - \frac{\langle x_3, y_2 \rangle}{\|y_2\|^2} y_2 = x_3 - \frac{5}{6} y_1 + \frac{1}{2} y_2 = \frac{1}{6}(6x_3 - 5y_1 + 3y_2) = \frac{1}{6}\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix}$.

   Let $y_3 = \frac{6}{4} x_3 - \frac{5}{4} y_1 + \frac{3}{4} y_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $x_3 = \frac{5}{6} y_1 - \frac{3}{6} y_2 + \frac{4}{6} y_3$.

   Then y_1, y_2, y_3 are orthogonal and x_i is a LC of $y_1, .., y_i$.
   Comment: $y_1, ..., y_r$ are not unique.

   (2) A decomposition
   Writing (1) in matrix form: For full column rank $X \in C^{n\times r}$, $X = YT$ where $Y \in C^{n\times r}$ has orthogonal columns and $T \in C^{r\times r}$ is non-singular upper-triangular matrix.
   Ex2: For $X = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ $Y = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 0 & 5/6 \\ 0 & 1 & -3/6 \\ 0 & 0 & 4/6 \end{pmatrix}$ such that $Y'Y = \text{diag}(6,\ 2,\ 3)$ and $X = YT$.

2. QR-decomposition

   (1) QR-decomposition I
   For $X \in C^{n\times r}$ with rank r, there are $Q \in C^{n\times r}$ with orthonormal columns and non-singular upper-triangular $R \in R^{r\times r}$ such that $X = QR$.
   Proof. $X = YT = [Y(Y^*Y)^{-1/2}][(Y^*Y)^{1/2}T] = QR$ where $Y^*Y = \text{diag}(\|Y_1\|^2, ..., \|Y_r\|^2)$, $(Y^*Y)^{-1/2} = \text{diag}(\|Y_1\|^{-1}, ..., \|Y_r\|^{-1})$ and $(Y^*Y)^{1/2} = \text{diag}(\|Y_1\|, ..., \|Y_r\|)$. So $Q^*Q = I_r$ and R is non-singular upper-triangular.
   Ex3: For $X = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $Q = \begin{pmatrix} 1/\sqrt{5} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{5} & 0 & -1/\sqrt{3} \\ 1/\sqrt{5} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$, $R = \begin{pmatrix} \sqrt{5} & 0 & 5\sqrt{5}/6 \\ 0 & \sqrt{2} & -\sqrt{2}/2 \\ 0 & 0 & 2\sqrt{3}/3 \end{pmatrix}$ and $X = QR$.

(2) QR-decomposition II
In QR-decomposition $X = QR$ the diagonal elements of upper-triangular R can have designated signs.

Proof. Let $E \in R^{r\times r}$ be an elementary matrix obtained by multiplying the ith row of I_r by -1. Then ER changes the sign of the ith row of R. and QE changes the sign of ith column of Q. But $EE = I$. So $X = (QE)(ER)$ where the new Q, QE, still has orthonormal columns and new R, ER, has the signs of ith row changed.

3. Decompositions based on ranks

(1) Recall
For $A \in C^{m\times n}$ with $\mathrm{rank}(A) = r$, $A = P_I Q_I'$
where $P_I \in C^{m\times r}$, $Q_I \in C^{n\times r}$ and $\mathrm{rank}(P_I) = \mathrm{rank}(Q_I) = r$.
Also $\mathrm{rank}(A) = \mathrm{rank}(A') = \mathrm{rank}(\overline{A}) = \mathrm{rank}(A^*)$.
We claim that we can require one of P_I and Q_I to have orthonormal columns.

(2) Decomposition based on rank

For $A \in C^{m\times n}$ with $\mathrm{rank}(A) = r$, $A = UT' = HV'$
where U, $H \in C^{m\times r}$, T, $V \in C^{n\times r}$, $U^*U = V^*V = I_r$ and $\mathrm{rank}(H) = \mathrm{rank}(T) = r$.

Proof In $A = P_I Q_I'$, by QR-decomposition $P_I = UR_1$ and $Q_I = VR_2$.
So $A = (UR_1)Q_I' = U(Q_I R_1)' = UT'$ and $A = P_I(VR_2)' = (P_I R_2')V' = HV'$.

(3) Applications

$\mathrm{rank}(A^*A) = \mathrm{rank}(AA^*) = \mathrm{rank}(A)$.

Proof $A = UT'$ where $U^*U = I$ and T has full column rank $r = \mathrm{rank}(A)$. So,

$$\mathrm{rank}(A^*A) = \mathrm{rank}\left((UT')^*(UT')\right) = \mathrm{rank}(\overline{T}T') = \mathrm{rank}(T) = r = \mathrm{rank}(A).$$

$A = HV'$ where $V^*V = I$ and H has full column rank $r = \mathrm{rank}(A)$. So

$$\begin{aligned}\mathrm{rank}(AA^*) &= \mathrm{rank}\left((HV'\overline{V}H^*\right) = \mathrm{rank}\left(H\overline{(V^*V)}H^*\right) = \mathrm{rank}(HH^*)\\ &= \mathrm{rank}(H) = r = \mathrm{rank}(A).\end{aligned}$$

Comment: For $A \in R^{m\times n}$, $\mathrm{rank}(A'A) = \mathrm{rank}(AA') = \mathrm{rank}(A)$.