L01: Real matrices

1. Matrices and their operations
   In this course we only consider real matrices.

   (1) $R^{m\times n}$ is a linear space
   Matrix addition, scalar multiplication, a set of rules;
   Linear combination $\alpha_1 A + \alpha_2 B + \cdots + \alpha_k D$
   Zero matrix can always be written as a LC of other matrices

   (2) Matrix multiplication
   Condition for AB
   Condition for AB where A and B are matrices with blocks
   Interpretation of Ax
   Interpretation of AB
   Identity matrices. Left-inverse of A; right-inverse of A; inverse of A.
   Non-singular matrix. Show $B = A^{-1}$.

   (3) Trace of a square matrix
   Show $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
   Symmetric matrices

   (4) Frobenius inner product
   Inner product, induced norm, distance, angle, Pythagorean Theorem
   Frobenius inner product
   Matrix with orthogonal column, matrix with othonormal columns, orthogonal matrices

   Ex1: For A, B, C in $R^{n\times n}$, show $(I + ABC)^{-1} = I - A(B^{-1} + CA)^{-1}C$.
   Proof

$$
\begin{aligned}
& (I + ABC)[I - A(B^{-1} + CA)^{-1}C] \\
= \; & I - A(B^{-1} + CA)^{-1}C + ABC - ABCA(B^{-1} + CA)^{-1}C \\
= \; & I - AB[B^{-1}(B^{-1} + CA)^{-1}C - C + CA(B^{-1} + CA)^{-1}C] \\
= \; & I - AB[B^{-1}(B^{-1} + CA)^{-1} - I + CA(B^{-1} + CA)^{-1}]C \\
= \; & I - AB[(B^{-1} + CA)(B^{-1} + CA)^{-1} - I]C = I - 0 = I.
\end{aligned}
$$

2. Terminology and notation

   (1) For statements A and B the followings are the same
   (a) If A is true, then B is true (b) $A \Longrightarrow B$ (c) $B \Longleftarrow A$
   (d) A is a sufficient condition for B
   (e) B is a necessary condition for A

   (2) For statements A and B the followings are the same
   (a) A is defined by B (b) B is defined by A
   (c) A is true if and only if B is true (d) B is true if and only if A is true
   (e) A is false if and only if B is false (f) B is false if and only if A is false
   (g) $A \Longleftrightarrow B$ (h) A and B are equivalent
   (i) A is a sufficient and necessary condition for B
   (j) B is a sufficient and necessary condition for A

   (3) For sets A and B the followings are equivalent
   (a) $A \subset B$ (b) $B \supset A$ (c) $x \in A \Rightarrow x \in B$

   (4) For sets A and B the followings are equivalent
   (a) $A = B$ (b) $A \subset B$ and $A \supset B$

Ex2: (i) $\mathrm{tr}(XA) = 0$ for all X has a sufficient condition $A = 0$

Proof $A = 0 \Longrightarrow \mathrm{tr}(XA) = \mathrm{tr}(X0) = \mathrm{tr}(0) = 0$ for all X

(ii) $\mathrm{tr}(XA) = 0$ for all X has necessary condition $A = 0$

Proof $0 = \mathrm{tr}(XA)$ for all $X \Longrightarrow 0 = \mathrm{tr}(A'A) = \|A\|^2 \Longrightarrow A = 0$

(iii) So $\mathrm{tr}(AX) = 0$ for all $X \Longleftrightarrow A = 0$

Ex3 For $A = \begin{pmatrix} A_1 \\ 0 \end{pmatrix} \in R^{m\times r}$ where $A_1 \in R^{r\times r}$ is non-singular, denote the left-inverse of A by A^L. Show $A^L = (A_1^{-1}, H)$ for all $H \in R^{r\times m-r}$.

$\subset$**:** Suppose $B = (B_1, B_2) \in A^L$. Then $I_r = BA = B_1 A_1 \Rightarrow B_1 = A_1^{-1}$.
So $B = (A_1^{-1}, H)$ with $H = B_2$

$\supset$**:** $(A_1^{-1}, H)A = (A_1^{-1}, H)\begin{pmatrix} A_1 \\ 0 \end{pmatrix} = I_r$. So $(A_1^{-1}, H) \subset A^L$.

3. Rank of matrices

(1) Independence

$U_1, .., U_r$ are vectors in a linear space

$U_1, .., U_r$ are linearly independent (LI) $\stackrel{def}{\Longleftrightarrow}$ "$x_1U_1 + \cdots + x_rU_r = 0 \Longrightarrow x_i = 0$ for all i"

$U_1, .., U_r$ are LD $\stackrel{def}{\Longleftrightarrow}$ $x_1U_1 + \cdots + x_rU_r = 0$ for some $x = (x_1, ..., x_r)' \neq 0$
$\Longleftrightarrow$ $\exists U_i$ that is a LC of others

(2) Rank and dimension

Suppose D is a set in linear space V,

$\mathrm{rank}(D) = r \stackrel{def}{\Longleftrightarrow} \exists[x_1, .., x_r] \subset D;\ x_1, ..., x_r$ are LI; x is a LC of $x_1, ..., x_r$ for all $x \in D$.

$\dim(V) = r \stackrel{def}{\Longleftrightarrow} \exists[x_1, .., x_r] \subset V;\ x_1, ..., x_r$ are LI; x is a LC of $x_1, ..., x_r$ for all $x \in V$.

$[x_1, ..., x_r]$ is a basis of V.

(3) Matrix rank

The rank of n columns of $A \in R^{m\times n}$ is called the column rank of A

The rank of m rows of $A \in R^{m\times n}$ is called the row rank of A

The column rank and row rank of A are always equal and is called the rank of A denoted as $\mathrm{rank}(A)$.

(4) For $A \in R^{m\times n}$, $\mathrm{rank}(A) = \mathrm{rank}(A') \le \min(m, n)$.

Ex5: $A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix}$

$\mathrm{rank}(A) = 2$ since $A_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ are LI; $A_3 = A_1 + A_2$ and $A_4 = A_1 - A_2$.

Note that the first and the third rows are LI, and the second row is a LC of the first and third row.

Ex6: $\mathrm{rank}(A) = \mathrm{rank}\left[\begin{pmatrix} A \\ 0 \end{pmatrix}\right] = \mathrm{rank}\left[\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}\right]$

L02 dimension of space and QR-decomposition

1. Dimension of space

   (1) Subspace and span
   V is a LS. $S \subset V$.

   $$\begin{aligned} & S \text{ is a subspace of } V \\ \iff & S \text{ is closed under addition and is closed under scalar multiplication} \\ \iff & S \text{ is closed under LC} \end{aligned}$$

   For $D \subset V$, the collection of all LCs of vectors in D is closed under LCs and hence is a subspace of V called the span of D denoted as Span(D). So

   $$D \subset V \Longrightarrow D \subset \text{Span}(D) \subset V.$$

   (2) Dimension and rank
   Suppose $[x_1, ..., x_r] \subset D$, $x_1, ..., x_r$ are LI, and x is a LC of $x_1, ..., x_r$ for all $x \in D$.
   Then rank$(D) = r$. But one can see that $[x_1, ..., x_r]$ is a basis of Span(D).
   So dim[Span(D)] $= r$. Thus dim[Span(D)] = rank(D).

   (3) Column space of matrix
   For $A \in R^{m\times n}$, $\{Ax : x \in R^n\}$ contains all LCs of the columns of A and hence is the span of the columns of A, a subspace of R^m called the column space of A with notations

   $$C(A) = \text{Span}(A) = L(A) = \{Ax \in R^m : x \in R^n\}.$$

   Clearly rank$[L(A)]$ = rank(A).

   Ex1: For two sets D_1 and D_2 in LS V, $D_1 \subset D_2 \Longrightarrow \text{rank}(D_1) \le \text{rank}(D_2)$
   For two spaces S_1 and S_2 in LS V, $S_1 \subset S_2 \Longrightarrow \dim(S_1) \le \dim(S_2)$.

   Ex2: For matrices A, B and AB,
   (i) $y \in L(AB) \Longrightarrow y = ABx = A(Bx) \in L(A)$. So $L(AB) \subset L(A)$
   (ii) $$\begin{aligned} L(AB) \subset L(A) &\Longrightarrow \dim[L(AB)] \subset L(A) \Longrightarrow \dim[L(AB)] \le \dim[L(A)] \\ &\Longrightarrow \text{rank}(AB) \le \text{rank}(A). \end{aligned}$$
   (iii) $\text{rank}(AB) = \text{rank}[(AB)'] = \text{rank}(B'A') \le \text{rank}(B') = \text{rank}(B)$
   So we conclude that the rank product is $\le$ the rank of a factor.

2. Sum and intersection of subspaces

   (1) Sum, intersection and their dimensions
   S_1 and S_2 are two subspaces of V. Then $S_1 + S_2 = \{x + y : x \in S_1 \text{ and } y \in S_2\}$ and $S_1 \cap S_2$ are subspaces of V.

   $$S_1 \cap S_2 \subset \left\{ \begin{array}{c} S_1 \\ S_2 \end{array} \right. \subset S_1 \cup S_2 \subset S_1 + S_2 \subset S$$

   all but $S_1 \cup S_2$ are LSs. It can be shown that

   $$\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2).$$

   (2) Direct sum
   $S_1 + S_2$ is a direct sum denoted by $S_1 \oplus S_2$ if $S_1 \cap S_2 = \{0\}$.

   $$\dim(S_1 \oplus S_2) = \dim(S_1) + \dim(S_2)$$

(3) Orthogonal sum

$S_1 + S_2$ is an orthogonal sum denoted by $S_1 \dot{\oplus} S_2$ if $S_1 \perp S_2$.

But $S_1 \perp S_2 \Longrightarrow S_1 \cap S_2 = \{0\}$. So an orthogonal sum is a direct sum.

Ex3: $C[(A,\, B)] = C(A) + C(B)$.

$\subset$**:** $z \in C[(A,\, B)] \Rightarrow z = (A,\, B)\begin{pmatrix} x \\ y \end{pmatrix} = Ax + By \in C(A) + C(B)$

$\supset$**:** $z \in C(A) + C(B) \Rightarrow z = Ax + By = (A,\, B)\begin{pmatrix} x \\ y \end{pmatrix} \in C[(A,\, B)]$

$$\begin{aligned} \text{rank}[(A,\, B)] &= \dim[C(A,\, B)] = \dim[C(A) + C(B)] \\ &= \dim[C(A)] + \dim[C(B)] - \dim[C(A) \cap C(B)] \\ &= \text{rank}(A) + \text{rank}(B) - \dim[C(A) \cap C(B)] \leq \text{rank}(A) + \text{rank}(B) \end{aligned}$$

3. QR decomposition

(1) QR-decomposition

For $A \in R^{n \times t}$ with full column rank $\text{rank}(A) = t$ there exist $Q \in R^{n \times t}$ with orthonormal columns $Q'Q = I_t$ and non-singular upper diagonal $R \in R^{t \times t}$ such that $A = QR$.

(2) Gram-Schmidt process

The essence of the QR-decomposition is the Gram-Schmidt process.

For given LI columns vectors $A_1, ..., A_t$ in $A = (A_1, ..., A_t)$, the process produces orthonormal $Q_1, ..., Q_t$ such that $A_i = r_{1i}Q_1 + \cdots + r_{ii}Q_i$, $i = 1, ..., t$.

With $Q = (Q_1, ..., Q_t)$ and $R = \begin{pmatrix} r_{11} & \cdots & r_{1t} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_{tt} \end{pmatrix}$, $A = QR$.

(3) Additional requirement

The QR-decomposition is not unique. One can let the each of the diagonal elements of R have designated signs.

Ex4: (a), (b) and (c) below are equivalent.

(a) A has a L-inverse (b) A has full column rank (c) $\text{rank}(AB) = \text{rank}(B)\, \forall\, B$

Proof. Suppose $A \in R^{n \times r}$.

(a)$\Longrightarrow$(b): A has L-inverse A^L. Then

$r = \text{rank}(I_r) = \text{rank}(A^L A) \leq \text{rank}(A) \leq r$. So (b) $\text{rank}(A) = r$ holds.

(a)$\Longleftarrow$(b): If (b), then by QR-decomposition, $A = QR$. Let $B = R^{-1}Q'$.

Then $BA = R^{-1}Q'QR = I_r$, i.e., A has a L-inverse. So (a) holds.

(a)$\Longrightarrow$(c): A has L-inverse A^L. Then

$\text{rank}(AB) \leq \text{rank}(B) = \text{rank}(I_r B) = \text{rank}(A^L AB) \leq \text{rank}(AB)$. So (c) holds.

(b)$\Longleftarrow$(c: $\text{rank}(A) = \text{rank}(AI_r) = \text{rank}(I_r) = r$. So (b) holds. $\square$.