Matrix Analysis

Part. 1

Chang Jae Ho
jaehochang@konkuk.ac.kr

February 28, 2020

Contents

Contents

Rank I

Theorem

If A is an $m \times n$ matrix of rank $r > 0$, then there **exists** *nonsingular $m \times m$ and $n \times n$ matrices B and C, such that $H = BAC$ and $A = B^{-1}HC^{-1}$ where H is given by*

a. I_r if $r = m = n$
b. $\begin{bmatrix} I_r & (0) \end{bmatrix}$ if $r = m < n$
c. $\begin{bmatrix} I_r \\ (0) \end{bmatrix}$ if $r = n < m$
d. $\begin{bmatrix} I_r & (0) \\ (0) & (0) \end{bmatrix}$ if $r < m, r < n$.

Rank II

Following is straightforward from the theorem.

Corollary

Let A be an $m \times n$ matrix with rank $r > 0$. Then, there **exist** *an $m \times r$ matrix F and an $r \times n$ matrix G such that $rank(F) = rank(G) = r$ and $A = FG$.*

Contents

Complex Matrix I

Any complex number c can be written in the form $c = a + bi$ where a and b are real numbers. Put $a = r\cos\partial$ and $b = r\sin\partial$, or, simply $c = re^{i\partial}$. Then,

$$|c| = \sqrt{a^2 + b^2} = r.$$

Likewise, $|c_1 c_2| = |c_1| \cdot |c_2|$ and $|c\bar{c}| = |c|^2$. Hence,

$$|c_1 + c_2| \leq |c_1| + |c_2|.$$

Proof.

(Exercise) □

Complex Matrix II

Complex matrix is defined as

$$C = A + iB$$
$$\bar{C} = A - iB$$

where A and B are real-valued matrices. If C is square and $\bar{C}' = C$, so that $\bar{c}_{ij} = c_{ji}$, then C is said to be **Hermitian**.

Remark If C is Hermitian and real-valued, then C is symmetric.

The square matrix C is said to be **unitary** if $\bar{C}'C = I$.

Contents

Random Vectors I

EXPECTATIONS

(Exercises)

1. $cov(\alpha' x, \beta' x) = \alpha' \Sigma \beta$ where $\Sigma = cov(x)$
2. $cor(x) = D\Sigma D$ where $D = diag(\sigma_1^{-1/2}, \cdots, \sigma_p^{-1/2})$
3. If $x \sim \mathcal{N}(0, I)$, then $Tx + \mu \sim \mathcal{N}(\mu, \Sigma)$ where $\Sigma = TT'$.

Contents

Definitions I

Definition

A collection of m-dimensional vectors that is closed under addition and scalar multiplication is called a **vector space** in m-dimensional space.

Theorem

Let $\{x_1, \cdots, x_p\}$ be a set of $n \times 1$ vectors in the vector space S, and let $\mathcal{Y}$ be the set of all possible linear combinations of these vectors. That is,

$$\mathcal{Y} = \left\{ y; y = \sum_{i=1}^{p} x_i \beta_i, |\beta_i| < \infty \text{ for all } i \right\}.$$

Then, $\mathcal{Y}$ is a vector subspace of S.

Definitions II

Definition

Let S be a vector space. A function, $< x, y >$, defined for all $x \in S$ and $y \in S$, is an **inner product** if for any x, y and z in S, and any scalar c,

a. $< x, x > \geq 0$ and $< x, x >= 0$ iff $x = 0$
b. $< x, y >=< y, x >$
c. $< x + y, z >=< x, z > + < y, z >$
d. $< cx, y >= c < x, y >$.

Definitions III

Theorem

If x and y are in the vector space S and $< x, y >$ is an inner product defined on S, then

$$< x, y >^2 \leq < x, x >< y, y >$$

Definitions IV

Definition

A function $||x||$ is a **vector norm** on the vector space S if, for any vectors x and y in S, we have

a. $||x|| \geq 0$
b. $||x|| = 0$ iff $x = 0$
c. $||cx|| = |c|||x||$ for any scalar c
d. $||x + y|| \leq ||x|| + ||y||$

Definitions V

Definition

A function $d(x, y)$ is a **distance function** defined on the vector space S if for any vectors x, y, and z in S, we have

a. $d(x, y) \geq 0$
b. $d(x, y) = 0$ iff $x = y$
c. $d(x, y) = d(y, x)$
d. $d(x, z) \leq d(x, y) + d(y, z)$

Remark $d_\Sigma(x, \mu) = \sqrt{(x-\mu)'\Sigma^{-1}(x-\mu)}$ is called the Mahalanobis distance between x and μ.

Contents

Basis and Dimension I

Definition

Let $\{x_1, \cdots, x_p\}$ be a set if $n \times 1$ vectors in a vector space S. This set is called a **basis** of S if it spans S and its vectors are linearly independent.

Definition

$dim(S)$ is the number of vectors in any basis for S.

Contents

Matrix Rank I

Definition

Let X be an $n \times p$ matrix. The subspace of $\mathbb{R}^n$ spanned by the n row vectors of X is called the **row space** of X. Likewise, the **column space(range)** of X can also be defined;

$$
\begin{aligned}
R(X) &= \left\{ y; y = \sum_{j=1}^{p} x_j \beta_j, |\beta_j| < \infty \text{ for all } j \right\} \\
&= \{ y; y = X\beta, \beta \in \mathbb{R}^p \}
\end{aligned}
$$

Remark $R(X')$ is a row space of X.

Matrix Rank II

Theorem

Let X be an $n \times p$ matrix. If r is the number of linearly independent rows of X and c is the number of linearly independent columns of X, then $rank(X) = r = c$.

Theorem

Let A, B be $m \times n$ matrices and C be $n \times p$ matrix. Then,

a. $rank(AC) \leq \min\{rank(A), rank(C)\}$
b. $rank(A + B) \leq rank(A) + rank(B)$
c. $rank(A) = rank(A') = rank(AA') = rank(A'A)$

Matrix Rank III

Theorem

Let A, B and C be any matrices for which the partitioned matrices below are defined. Then,

1. $rank\left(\begin{bmatrix} A & B \end{bmatrix}\right) \geq \max\{rank(A), rank(B)\}$
2. $rank\left(\begin{bmatrix} A & (0) \\ (0) & B \end{bmatrix}\right) = rank\left(\begin{bmatrix} (0) & B \\ A & (0) \end{bmatrix}\right) = rank(A) + rank(B)$
3. $rank\left(\begin{bmatrix} A & (0) \\ C & B \end{bmatrix}\right) = rank\left(\begin{bmatrix} C & B \\ A & (0) \end{bmatrix}\right) = rank\left(\begin{bmatrix} B & C \\ (0) & A \end{bmatrix}\right) = rank\left(\begin{bmatrix} (0) & A \\ B & C \end{bmatrix}\right) \geq rank(A) + rank(B)$

Matrix Rank IV

Theorem

Let A, B and C be any matrices for which ABC can be defined. Then,

$$rank(ABC) \geq rank(AB) + rank(BC) - rank(B)$$

Remark If $C = I_n$, for example, then the lower bound becomes $rank(A) + rank(B) - n$.

Contents

Orthonormal Basis I

REGRESSION EXAMPLE

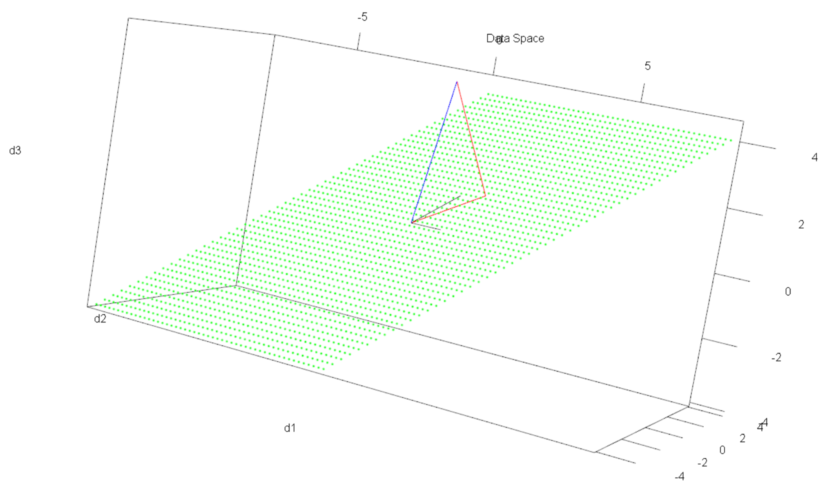
In simple linear regression, we find $\hat{\beta}$ such that minimizes $(y - \hat{y})'(y - \hat{y}) = (y - X\hat{\beta})'(y - X\hat{\beta})$. For a choice of $\hat{\beta}$, $\hat{y} = X\hat{\beta}$ gives a point in the subspaces of $\mathbb{R}^n$ spanned by 1_n and x. Thus, the point $\hat{y}$ that minimizes the distance from y will be given by the orthogonal projection of y onto this plane. That is,

$$(y - \hat{y})' 1_n = 0, \ (y - \hat{y})' x = 0$$

through which we obtain an LS estimator.

Orthonormal Basis II

REGRESSION EXAMPLE



Orthonormal Basis III

REGRESSION EXAMPLE

Definition

Let $S = X\beta$ be a (data) vector subspace of $\mathbb{R}^n$. The **orthogonal complement** of S, denoted by $S^\perp$, is the collection of all n-dimensional vectors that are orthogonal to every vector in S; that is,

$$S^\perp = \{x; x \in \mathbb{R}^n \text{ and } x'a = 0 \ \forall a \in S\}.$$

Theorem

If S is a vector subspace of $\mathbb{R}^n$ then $S^\perp$ is also a vector subspace of $\mathbb{R}^n$.

Orthonormal Basis IV

REGRESSION EXAMPLE

Theorem

Suppose that $\{x_1, \cdots, x_n\}$ is an orthonormal basis for $\mathbb{R}^n$ and $\{x_1, \cdots, x_s\}$ is an orthonormal basis for the vector subspace S. Then $\{x_{s+1}, \cdots, x_p\}$ is an orthonormal basis for $S^\perp$.

Contents

Orthogonal Projection I

Theorem

Suppose the columns of the $n \times p$ matrix X_1 forms an orthonormal basis for the vector subspace S of $\mathbb{R}^n$. if $x \in \mathbb{R}^n$, then the orthogonal projection of x onto S is given by $X_1 X_1' x$.

Example

By this theorem, the projection of y onto $R(XP\Lambda^{-1/2})$ is $X(X'X)^{-1}X'y$ where $X'X = P\Lambda P'$(s-decomposition). Here,

$$\Lambda^{-1/2}P'X'XP\Lambda^{-1/2} = \Lambda^{-1/2}P'P\Lambda P'P\Lambda^{-1/2} = \Lambda^{-1/2}\Lambda\Lambda^{-1/2} = I.$$

Contents

Null Space I

Theorem

Let $u = X\beta$ where $X \in \mathbb{R}^{n \times p}$ and $\beta \in \mathbb{R}^p$. The null space of X, given by the set

$$N(X) = \{\beta \in \mathbb{R}^p; X\beta = 0\}$$

is a vector space.

Theorem

Let X be an $n \times p$ matrix. Then,

$$p = rank(X) + dim\{R(X')^{\perp}\} = rank(X) + dim\{N(X)\}$$

Null Space II

Example

For an $n \times p$ matrix X, $rank(X) = rank(X'X)$.

Proof.

Let $\beta \in N(X)$ so $X\beta = 0$. Therefore, $X'X\beta = 0$ and $\beta \in N(X'X)$. So, $dim\{N(X)\} \leq dim\{N(X'X)\}$, or

$$rank(X) \geq rank(X'X).$$

On the other hand, if $\beta \in N(X'X)$ then $\beta' X' X\beta = 0$ which is only satisfied only if $X\beta = 0$. Therefore, $\beta \in N(X)$ so $rank(X) \leq rank(X'X)$. $\square$

Linear Transformation I

Rotation of Axes

Example

Let

$$A := \begin{bmatrix} \cos\partial & \sin\partial & 0 \\ -\sin\partial & \cos\partial & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}.$$

The transformation given by $u = Ax$ rotates axes e_1, e_2, e_3 to the new axes x_1, x_2, x_3. This represents a roation of e_1 and e_2 through an angle of ∂.

Contents

Definitions I

Definition

Let S_1, S_2 be vector subspaces of $\mathbb{R}^m$. Define $\cup$ and $\cap$ of vector spaces by

$$S_1 \cap S_2 := \{x \in \mathbb{R}^m; x \in S_1 \text{ and } x \in S_2\}$$
$$S_1 \cup S_2 := \{x \in \mathbb{R}^m; x \in S_1 \text{ or } x \in S_2\}$$

and define the sum of vector spaces by

$$S_1 + S_2 := \{x_1 + x_2; x_1 \in, x_2 \in S_2\}.$$

Definitions II

Remark If $S_1 \cap S_2 = \{0_m\}$, then $S_1 \oplus S_2 := S_1 + S_2$ is called the direct sum of S_1, S_2.

Each $x \in S_1 \oplus S_2$ has a unique representation as $x = s_1 + s_2$ where $s_1 \in S_1$ and $s_2 \in S_2$.

Remark If S_1 and S_2 are orthogonal, then $x \in S_1 \oplus S_2$ has $x = P_{S_1}x + P_{S_2}x$.

For instance, for any vector subspace S of $\mathbb{R}^m$, $\mathbb{R}^m = S \oplus S^\perp$, and $\forall x \in \mathbb{R}^m$,

$$x = P_S x + P_{S^\perp} x.$$

Orthogonal Decomposition of $R(X)$ I

Theorem

Let X be $n \times p$ with $\mathrm{rank}(X) = p$ and

$$X = \begin{bmatrix} X_0 & X_1 \end{bmatrix}, P_X = X(X'X)^{-1}X', P_i = X_i(X_i'X_i)^{-1}X_i' \text{ for } i = 0, 1$$
$$\text{and } X_{1|0} = (I - P_0)X_1, P_{1|0} = X_{1|0}(X_{1|0}'X_{1|0})^{-1}X_{1|0}'.$$

Then,

1. $R(X) = R\left(\begin{bmatrix} X_0 & X_{1|0} \end{bmatrix}\right)$
2. $\{X_{1|0}\beta; \beta \in \mathbb{R}^{p_1}\} = \{X_0\beta; \beta \in \mathbb{R}^{p_0}\}^{\perp}$
3. $P_X = P_0 + P_{1|0}$ *and* $P_0P_{1|0} = 0$

Contents

Definition I

Convex Set

Definition

A set $S \subseteq \mathbb{R}^m$ is said to be convex if for any $x_1 \in S$ and $x_2 \in S$,

$$cx_1 + (1-c)x_2 \in S$$

for a constant $c \in (0,1)$.

Theorem

If $S_1, S_2 \subseteq \mathbb{R}^m$ are convex, then $S_1 \cap S_2$ and $S_1 + S_2$ are convex.

Theorem

If $S \subseteq \mathbb{R}^m$ is convex, then $\bar{S}$ is also convex.