Projection

(1) <u>Def:</u>

For $M = N_1 \oplus N_2$. $\forall z = x + \eta$. $x \in N_1, \eta \in N_2$. The linear transformation $P_{N_1|N_2} z = x$. is projection onto N_1 along subspace N_2.

<u>Rmk:</u> i) $P_{N_2|N_1} = I - P_{N_1|N_2}$. It's easy to check.

ii) If proj. is at the right angle. Then it's unique (Orthogonal). Otherwise it's not unique.

<u>Def:</u> $A \in M^{n \times n}$. A is proj. onto $C(A)$ along $N(A)$ if $\forall v \in C(A)$. $Av = v$.

① <u>Idempotent Matrix:</u>

<u>Def:</u> $A^2 = A$. A is called idempotent.

<u>Thm.</u> $A^2 = A \Leftrightarrow$ A is proj. matrix.

<u>Pf:</u> ($\Leftarrow$) $A(Ax) = Ax$. $\forall x \in \mathbb{R}^n$. $\therefore A^2 = A$.

($\Rightarrow$) $A^2 x = Ax$. $\forall x \in \mathbb{R}^n$. So for $v \in C(A)$.

$Av = v$.

<u>Thm.</u> $M = C(A) \oplus C(I-A) = N(A) \oplus N(I-A)$

$C(A) = N(I-A)$. $C(I-A) = N(A)$. if $A^2 = A$.

Pf: $C(A) \subseteq N(I-A)$. $C(I-A) \subseteq N(A)$.

$\forall \alpha \in M$. $\alpha = \alpha - A\alpha + A\alpha$.

$\in C(I-A) + C(A)$.

$\therefore M = C(I-A) + C(A)$.

If $\beta \in C(I-A) \cap C(A)$. check $\beta = 0$.

$\therefore C(I-A) = N(A)$. $C(A) = N(I-A)$.

Since similarly. $M = N(I-A) \oplus N(A)$

Thm. If A is proj on $C(A)$. Then it proj $N(A)^c$ as well.

Pf: $\forall v \in N(A)^c$. Since $v = u + w \in N(A) \oplus N(I-A)$.

$v - w = u \in N(A)^c \cap N(A)$ $\therefore u = 0$.

$Av = Aw = w = v$.

② Orthogonal Proj:

Def: A is orthogonal proj onto $C(X)$. if

$v \in C(X) \Rightarrow Av = v$. $v \in C(X)^\perp \Rightarrow Av = 0$.

Rmk: It's at right angle.

Thm: A is orthogonal proj on $C(A) \Leftrightarrow A^2 = A$

and $A^T = A$

Pf: Recall a fact: $C(A)^\perp = N(A^T)$

$C(A^T)^\perp = N(A)$.

Thm. (Uniqueness)

i) If M is orthogonal proj. on $C(X)$. Then $C(M) = C(X)$.

ii) If M, P are two orthorgonal proj. on M. Then $M = P$.

Pf: i) $C(X) \subseteq C(M)$ is trival.

Conversely. $\forall v \in C(M)$. $\exists u$. $v = Mu$.

$u = t_1 + t_2 \in C(X) + C(X)^\perp$ $\therefore Mu = t_1 \in C(X)$.

$\therefore v \in C(X)$.

ii) Check $Mv = Pv$. $\forall v \in \mathbb{R}^n$.

Thm. $M \in M^{n\times n}$. $r(M) = r$. Othomogonal projection. Then:

i) $\sigma_M \subseteq \{0, 1\}$. ii) $r(M) = tr(M) = r$.

iii) M is positive semidefinite matrix.

Thm. $X \in M^{n\times p}$ with rank $r \leq \min\{n, p\}$. $\{a_k\}_1^r$ is orthonormal basis of $C(X)$. $A = (a_1 \cdots a_r)$. Then $AA^T = \sum a_k a_k^T$ is the orthogonal projection on $C(X)$.

Pf: Check $(AA^T)^2 = AA^T$. $C(X) = C(A) = C(AA^T)$.

Rmk: $X = QR$. Then $A = Q$. QQ^T is the orthogonal proj on $C(X)$.

Thm (Construction)

$X \in M^{n \times p}$ with rank p. $M = X(X^TX)^{-1}X^T$ is orthogonal projection on $C(X)$.

Pf: 1°) Check $M^T = M$. $M^2 = M$.

2°) $C(M) \subset C(X)$. Conversely. Note $MX = X$.

$\therefore C(X) = C(MX) \subset C(M)$.

Cor. $M = VV^T$ is unique orthogonal projection on $C(X)$. $X = U \Sigma V^T$.

Thm (Othogonal Space)

$I - M$ is unique orthogonal proj on $C(X)^\perp$.

$M = X(X^TX)^{-1}X^T$.

Pf: Check $(I-M)^T = I-M$. $(I-M)^2 = I-M$.

Show: $C(I-M) = C(X)^\perp$. $\Longleftrightarrow$

$(I-M)X = 0$. $\therefore \eta^T(I-M)X = 0$.

$\therefore (I-M)\eta \in C(X)^\perp$. $\forall \eta \in \mathbb{R}^n$.

Thm. $M = (m_{ij})_{n\times n}$ orthogonal proj. Then $m_{ii} \in [0,1]$.

Lemma. If $A = A^T$. Then $\lambda_1 \le a_{ii} \le \lambda_n$. where $\lambda_1 \le \lambda_2 \cdots \lambda_n$. $\sigma_A = \{\lambda_k\}_1^n$

Pf: $A = Q\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}Q^T$. $a_{ii} = \sum q_{ij}^2 \lambda_j$

Thm. (Decomposition)

$X \in M^{n\times p}$ with rank p. $X = (X_1\ X_2)$. $X_1 \in M^{n\times k}$, $X_2 \in M^{n\times(p-k)}$. $r(X_1) = k$. $r(X_2) = p-k$. Let $M = X(X^TX)^{-1}X^T$.

$M_i = X_i(X_i^TX_i)^{-1}X_i^T$. $X_i^* = (I - M_{3-i})X_i$. $i = 1.2$.

$M_i^* = X_i^*(X_i^{*T}X_i^*)^{-1}X_i^{*T}$. Then $M = M_1 + M_2^* = M_1^* + M_2$.

Pf: 1°) $X_1^{*T}X_2 = X_1 X_2^{*T} = 0$. since $M_2X_2 = X_2$. $M_1X_1 = X_1$.

2°) $(M_1 + M_2^*)^T = M_1 + M_2^*$

$(M_1 + M_2^*)^2 = M_1^2 + M_2^{*2} = M_1 + M_2^*$

since: $M_1 M_2^* = \square\, X_1^T X_2^*\, \square = 0$.

3°) M is orthogonal proj on $C(X_1\ X_2)$.

M_1 is orthogonal proj on $C(X_1)$.

M_2^* is on $C((I - M_1)X_2)$

$\because C(X) = C(X_1) \oplus C(X_2) = C(X_1) \oplus C(C(X_2) \cap C^{\perp}(X_1))$

$M = C(X) \oplus C(X)^{\perp}$ $v = u_1 + u_2 + u_3 \in M$

$\therefore Mv = u_1 + u_2 = M_1u_1 + M_2^*u_2 = M_1v + M_2^*v$.

since $u_2 + u_3 \in C(X_1)^{\perp}$. $u_1 + u_3 \in C(X_2)$.

Thm. M_0, M are orthogonal proj. $C(M_0) \subset C(M)$. Then

i) $M_0 M = M M_0 = M_0$

ii) $M - M_0$ is orthogonal proj on $C(M) \cap C(M_0)^\perp$

iii) $C(M - M_0) \perp C(M_0)$.

Pf: i) $M_0 M v = M_0 v \quad \forall v \in \mathbb{R}^n$.

Besides, $(M_0 M)^T = M M_0 = M_0$

ii) Check $(M - M_0)^2 = (M - M_0) = (M - M_0)^T$.

$\because (M - M_0) M_0 = 0$. $\therefore C(M - M_0) \perp C(M_0)$.

$\forall v \in C(M - M_0) \Rightarrow M_0 v = 0$. $(M - M_0) v = 0$.

$\therefore v \in C(M_0)^\perp \cap C(M)$. Converse is trivial.

Cor. $C(M) = C(M_0) \perp C(M - M_0)$

Thm. M_1, M_2 are two orthogonal proj. Then $M_1 + M_2$ is orthogonal proj. on $C(M_1, M_2) \iff C(M_1) \perp C(M_2)$

Rmk: $C(M_1) \perp C(M_2) \iff M_1 M_2 = M_2 M_1 = 0$.

Thm. (Converse)

If M_1, M_2 are symmetric, $C(M_1) \perp C(M_2)$, and $M_1 + M_2$ is orthogonal proj. Then M_1, M_2 are orthogonal proj's.

Pf: Directly decompose the space.

Generalized Inverse

· Note that if $X \in M^{n\times p}$. $r(X) < \min\{n, p\}$. Then $(X^TX)^{-1}$ doesn't exist. We need G. I. of X^TX. so that the estimate can be computed.

(1) <u>Definition:</u>

<u>Def.</u> $A \in M^{n\times p}$. $A^- \in M^{p\times n}$ is generalized inverse of A if $AA^-A = A$.

<u>Rmk:</u> It's equi with: $AA^-\eta = \eta$. $\forall \eta \in C(A)$.

① <u>Prop.</u>

i) A^-A is idempotent (so a proj).

ii) For G_1, G_2 are generalized inverse of A. Then so is G_1AG_2.

iii) A is symmetric $\Rightarrow \exists A^-$ is symmetric.

② <u>Existence:</u>

$\forall A \in M^{n\times p}$. A^- exists. but not nessecary to be unique.

<u>Pf:</u> $A = P\begin{pmatrix} I_r & \\ & \end{pmatrix}Q$. Solve $AXA = A$.

$$\Rightarrow X = Q^{-1}\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix}P^{-1}. \quad Q\times P \triangleq T.$$

Rmk: If $|A| \neq 0$. Then $A^- = A^{-1}$. unique.

③ Properties:

i) $r(A^-) \geq r(A) = r(AA^-) = r(A^-A) = tr(AA^-)$.

Pf: $AA^- = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q\, Q^{-1}\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} P^{-1}$

$= P\begin{pmatrix} I_r & T_{12} \\ 0 & 0 \end{pmatrix} P^{-1}$

ii) $A^TA(A^TA)^-A^T = A^T$. $A(A^TA)^-A^TA = A$

Pf: By $AA^Tx = 0 \iff Ax = 0 \iff A^Tx = 0$

iii) $A(A^TA)^-A^T$ is orthonormal proj. indept with the choice of $(A^TA)^-$.

Pf: By ii) $A^TA(A^TA)^-_1 A^T = A^T = A^TA(A^TA)^-_2 A^T$

$\Rightarrow A(A^TA)^-_1A^T = A(A^TA)^-_2 A^T$

1') Check $A(A^TA)^-A^T$ is symmetric:

Choose P: $A^TA = P^T\begin{pmatrix} \lambda_1^2 & & & \\ & \ddots & & \\ & & \lambda_r^2 & \\ & & & 0 \ddots 0 \end{pmatrix} P$

$\begin{pmatrix} \lambda_1^2 & & & \\ & \ddots & & \\ & & \lambda_r^2 & \\ & & & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 1 \ddots 1 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 1 \ddots 1 \end{pmatrix}$

Set $Q = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 1 \ddots 1 \end{pmatrix} P$ $\therefore A^TA = Q^T\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$.

Choose: $(A^TA)^- = Q^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}(Q^{-1})^T$. sym.

2') $(A(A^TA)^-A^T)^2 = A(A^TA)^-A^T$.

<u>Cor.</u> $A(A^TA)^-A^T$ projects on $C(A)$.

(2) <u>Moore-Penrose G.I.</u>:

<u>Def</u>: $A \in M^{n\times p}$. M-P g-inverse of A is $A^+ \in M^{p\times n}$ satisfies: i) $(A^+A)^T = A^+A$. $(AA^+)^T = AA^+$

ii) $AA^+A = A$. $A^+AA^+ = A^+$.

<u>Rmk</u>: By def: A^+A. AA^+ are orthogonal proj

① <u>Existense</u>:

<u>Thm.</u> $\forall A \in M^{n\times p}$. It has a M-P g-inverse A^+

<u>Pf</u>: By SVD-decompose: $A = U\Sigma V^T$. $\Sigma \in M^{r\times r}$ where $r = r(A)$. Define $A^+ = V\Sigma^{-1}U^T$. Check

② <u>Uniqueness</u>:

<u>Thm.</u> A^+ is unique for A

<u>Pf</u>: $A^+ = A^+AA^+ = (A^+A)^TA^+ = A^T(A^+)^TA^+$

$= (AA^\#A)^T(A^+)^TA^+ = (AA^\#)^T(A^+A)^TA^+$

$= A^\#AA^+$. By symmetric. $A^T = A^\#$.

<u>Rmk</u>: i) $|A| \neq 0 \Rightarrow A^+ = A^{-1}$

ii) $A \in M^{n\times p}$. $r(A) = p \Rightarrow A^+ = (A^TA)^{-1}A^T$

iii) $A \in M^{n\times p}$. $r(A) = n \Rightarrow A^+ = A^T(AA^T)^{-1}$

③ property:

i) $(A^+)^+ = A$

Pf: By uniqueness. sym of A.

ii) $A^+ = A^T(AA^T)^+ = (A^TA)^+A^T$

Pf: By uniqueness

iii) $(A^TA)^+ = A^+(A^+)^T$

Pf: By uniqueness

iv) $(A^+)^T = (A^T)^+$. $\therefore$ A sym $\Rightarrow$ A^+ sym

v) $A = PQ^T$. $r(P) = r(Q) = r(A). = r$. $P \in M^{n \times r}$. $Q \in M^{r \times q}$

$\Rightarrow A^+ = (Q^+)^T P^+$.

Rmk: generally. $(AB)^+ \neq B^+A^+$.

vi) A is orthogonal proj $\Rightarrow$ $A^+ = A$.

Pf: $AAA = A$. with $A^T = A$.

vii) $A \in M^{n \times n}$. $A^T = A$. $A = P^T \operatorname{diag}\{\lambda_1 \cdots \lambda_n\} P$.

$\lambda_i^+ = \begin{cases} \lambda_i^{-1}, & \lambda_i \neq 0 \\ 0, & \lambda_i = 0 \end{cases}$. Then $A^+ = P^T \operatorname{diag}\{\lambda_i^+\} P$.

viii) $C(AA^+) = C(A)$

Pf: $r(AA^+) = r(A)$.

(3) Charaterization of Solutions:

- Consider $Y = X\beta$. $Y \in M^{n\times 1}$. $X \in M^{n\times p}$. Known. $\beta \in M^{p\times 1}$. unknown.

① $p = n$. X is nonsingular: $X^{-1}Y = \beta$.

② $p < n$. $Y \in C(X)$.

i) $r(X) = p$. Then $\beta = X^{-}Y$. $(XX^{-}Y = Y.\ \forall Y \in C(X))$ which is unique.

ii) $r(X) < p$. Then $\beta = X^{-}Y + (I - X^{T}(XX^{T})^{-}X)z$. $z \in \mathbb{R}^{p}$. since $C(X^{T}) \perp C(I - X^{T}(XX^{T})^{-}X)$

③ $p < n$. $Y \notin C(X)$.

Then β has no solution. We will look for the vector in $C(X)$ closest to Y. (i.e. $MY = X\beta$)

$\Rightarrow \beta = X^{-}MY + (I - X^{T}(XX^{T})^{-}X)z$. $z \in \mathbb{R}^{p}$.

If use MP-g.i. $X^{+}MY = X^{+}Y$.

④ $p = n$. $Y \notin C(X)$.

As above: $\beta = X^{-}MY$. choosedly. $\beta = X^{+}Y$.

Kronecker Product

(1) Definition:

Def: i) $A = (a_1 \cdots a_n) \in M^{p\times n}$. $Vec(A) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{np\times 1}$.

ii) $A \in M^{n\times m}$. $B \in M^{p\times q}$. $A\otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ & \vdots & \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix} \in \mathbb{R}^{np\times mq}$.

(2) Properties:

① $\forall A, B, C$.

i) $A\otimes B\otimes C = (A\otimes B)\otimes C = A\otimes(B\otimes C)$.

ii) $(A\otimes B)^T = A^T\otimes B^T$. iii) $r(A\otimes B) = r(A)r(B)$.

iv) $(A\otimes B)(C\otimes D) = AC\otimes BD$. if AC. BD. exists.

v) $(A+B)\otimes(C+D) = A\otimes C + A\otimes D + B\otimes C + B\otimes D$.

Rmk: $A\otimes B \neq B\otimes A$. $Vec(AXB) = (B^T\otimes A)Vec(X)$

② For $A \in M^{n\times n}$. $B \in M^{p\times p}$.

vi) $tr(A\otimes B) = tr(A)\,tr(B)$.

vii) $|A\otimes B| = |A|^p\,|B|^n$. ($A\otimes B = (A\otimes I_p)(B\otimes I_n)$)

viii) $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$. if A^{-1}. B^{-1}. exists.

ix). $\sigma_A = \{\lambda_i\}_1^n$. $\sigma_B = \{\mu_i\}_1^p$. Then: we have

$\sigma_{A\otimes B} = \{\lambda_i\mu_j\}_{i,j}$.

Pf: $(PAP^{-1})\otimes(QBQ^{-1}) = (P\otimes Q)(A\otimes B)(P\otimes Q)^{-1}$.

Differentiate on Matrix.

(1) For $Y = (y_{ij}(t)) \in M^{p\times q}$:

Def: $\frac{dY}{dt} = \left(\frac{dy_{ij}}{dt}\right)_{p\times q}$.

prop. i) $\frac{d(X+Y)}{dt} = \frac{dX}{dt} + \frac{dY}{dt}$. $\frac{d(XY)}{dt} = X\frac{dY}{dt} + \frac{dX}{dt}Y$.

ii) $\frac{\partial X}{\partial x_{ij}} = E_{ij}$

iii) $\frac{\partial AXB}{\partial x_{ij}} = AE_{ij}B$ (write $AXB = \sum_{i,j} x_{ij} AE_{ij}B$)

(2) For $f(X)$. $X \in M^{m\times n}$. $f \in \mathbb{R}^1$:

Def: $\frac{df(X)}{dX} = \left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}$.

prop. i) $\frac{\partial\, tr(AXB)}{\partial X} = A^T B^T$ (write in E_{ij})

ii) $\frac{\partial\, tr(AX)}{\partial X} = \begin{cases} A^T. & X \neq X^T \\ A + A^T - diag[a_{11} \cdots a_{nn}]. & X = X^T. \end{cases}$

iii) $\frac{\partial X^T A X}{\partial X} = (A + A^T)X$.

(3) For $\vec{f}(\vec{x})$. $X \in \mathbb{R}^n$. $\vec{f} \in \mathbb{R}^m$:

Def: $\frac{df}{dx} = \left(\frac{\partial f_i}{\partial x_j}\right)_{n\times m}$

Rmk: $\frac{\partial AX}{\partial X} = A$.

Multivariate Ch.f's

Properties:

i) $\varphi(t_1 \cdots t_n)$ uniformly conti. $|\varphi(\vec{t})| \le \varphi(\vec{0}) = 1$

ii) If $E(X_1^{k_1} \cdots X_n^{k_n})$ exists. Then:

$$E(X_1^{k_1} \cdots X_n^{k_n}) = i^{-\sum_1^n k_i} \frac{\partial^{\sum_1^n k_i}}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} \phi \Big|_{t_1 = t_2 \cdots = t_n = 0}$$

iii) $\phi(\vec{t}) = \phi(t_1 \cdots t_k, 0, 0 \cdots 0)$

iv) (Inversion Formula)

$$P(a_k < X_k \le b_k,\ 1 \le k \le n) = \lim_{T_k \to \infty, \forall k} \frac{1}{(2\pi)^n} \int_{-T_1}^{T_1} \int_{-T_2}^{T_2} \cdots \int_{-T_n}^{T_n} \prod_k^n \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} \phi(\vec{t})$$

Cor. d.f's $\Longleftrightarrow$ ch.f's one-to-one corresponds.

v) $\phi_K(\vec{t}) \longrightarrow \phi(\vec{t})$. which conti at $\vec{0}$.

$\Rightarrow \phi$ is ch.f of some random vector.