STAT 8004 Lecture 1

Thursday, January 15, 2015 4:34 PM

- Y $F(y)$ ← response, central interest

- $X_1, X_2 \cdots X_p$ } predictors / covariates

  $(Y, X_1, X_2, \cdots X_p)$ ← joint distribution

- $$\boxed{\begin{array}{c} Y \mid X_1, X_2, \ldots, X_p \\ E(Y \mid X_1, \ldots X_p) \\ Var(Y \mid X_1, \ldots X_p) \end{array}}$$

- linear model

  $$Y = \beta_0 + \beta_1 \underline{\underline{X_1}} + \cdots + \beta_p X_p + \underline{\underline{\epsilon}}$$

  - estimation
  - hypothesis testing
  - new obs. X_{new}, ?

  ...

- prediction

- more general distribution for Y

conditional mean of Y given $\underline{X_1 \dots X_p}$

is linear in

up to some "known" transformation

- "GLM"

- $\underline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix}$ $\qquad cov(Y \mid X_1 \dots X_p)$

$cov(Y_1, Y_2 \mid X_1 \dots X_p)$

- variance - covariance structure

mixed model.

- $Y = \underline{g}(X_1, X_2, \dots X_p, \epsilon)$

- nonparametric $\hat{g}(\qquad)$

$$\left\{ \begin{array}{l} \text{- estimation (point)} \\ \text{- hypothesis testing} \\ \text{- confidence set estimation} \end{array} \right.$$

- Assessment of level of uncertainties

$$\hat{X} \pm \boxed{z_\alpha \frac{\hat{\sigma}}{\sqrt{n}}}$$

bootstrap.

- working with vectors & matrices. (R&S, Ch 2 - Ch 3)

Definitions:

vector : $a = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

- inner product: $a, b \in \mathbb{R}^n$

$$a^T b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

- length (Euclidian distance)

$$\|a\| = \sqrt{a^T a} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

$$\begin{cases} l_2\text{-norm} \\ l_1\text{-norm} \\ l_p\text{-norm} \\ l_\infty\text{-norm} \end{cases}$$

- matrix $A_{m\times k} = \begin{pmatrix} a_{11}, & \cdots & a_{1k} \\ \vdots & & \\ a_{m1}, & \cdots & a_{mk} \end{pmatrix}$

$\in \mathbb{R}^{m\times k}$

- matrix addition $A + B = C$

$c_{ij} = a_{ij} + b_{ij}$

matrix substraction $A - B = C$

- scalar product, a, B

$aB = (a\, b_{ij})$

- transpose A^T, or A'

$$A = (a_{ij}) \qquad A^T = (a_{ji})$$

- square matrix $\quad A_{m \times m}$

- symmetric matrix $\quad A = A^T$

- matrix multiplication

$$A_{m \times k} B_{k \times n} = C_{m \times n}$$

$$c_{ij} = \sum_{l=1}^{k} a_{ik} b_{kj}$$

$$= a_{i \cdot}^T b_{\cdot j}$$

- Hadamard product (element product)

$$A_{m \times k} \quad B_{m \times k} = C_{m \times k}$$

$$c_{ij} = a_{ij} b_{ij}$$

- Kronecker product

$$A_{m \times k} \otimes B_{n \times l} = C_{mn \times kl}$$

$$\begin{pmatrix} a_{11}B, & a_{12}B, & \cdots & a_{1k}B \\ \vdots & & & \vdots \\ a_{m1}B & \cdots & \cdots & a_{mk}B \end{pmatrix}$$

- R

- $A_{n \times n}$

determinant $|A| = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} |M_{ij}|$

M_{ij} "minor" is A deleting the ith row & jth column

$= \sum_{j=1}^{n} a_{ij} (-1)^{i+j} |M_{ij}|$

rank$(A) = 2$ $\boxed{A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}}$ $|A| = 0$

properties :

$|A| = |A^T|$

$|AB| = |A||B| = |B||A|$
$= |BA|$

$|cA| = c^n |A|$

- linear independence

$Y_1, Y_2, \cdots Y_k$. k vectors

$a_1, a_2, \cdots a_k$

if $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 = \sum_{i=1}^{k} a_i Y_i \Rightarrow a_1, a_2 \ldots a_k$ are all zero

then, $Y_1, Y_2 \cdots Y_k$ are linearly independent

- rank: A

rank(A)

\# of linearly independent rows

\# of linearly independent columns

- non-singular $A_{n \times n}$

rank$(A) = n$

properties - rank$(AB) \le \min(\text{rank}(A), \text{rank}(B))$

. B, C nonsingular

rank$(BA) =$ rank$(CA) =$ rank(A)

-

$\cdot$ $\text{rank}(AA^T) = \text{rank}(A^TA)$

$= \text{rank}(A) = \text{rank}(A^T)$

- Identity matrix, $I = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}_{n\times n}$

- Inverse matix $A_{n\times n}$

$$A^{-1}A = AA^{-1} = I$$

- Equivaleme A is non-singular

$\Leftrightarrow |A| \neq 0$

$\Leftrightarrow A^{-1}$ exists

- properties $A_{n\times n}$

- $(A^T)^{-1} = (A^{-1})^T$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^{-1} = A$

- inverse of diagonal matrix

block diagonal matrix

- trace $A_{n\times n}$

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

- properties
  - $tr(cA) = c\,tr(A)$
  - $tr(AB) = tr(BA)$
  - $tr(B^{-1}AB) = tr(A)$

—x—

$$tr(BB^T) = \sum_i \sum_j b_{ij}^2$$

- matrix (Frobenius norm)

$$\|B\|_F = \sqrt{tr(BB^T)} = \sqrt{\sum_i \sum_j b_{ij}^2}$$

- R calculation

—x—

- Orthogonality - vector: a, b | $a \perp b$

$a^T b = 0$

$\Rightarrow$ a & b are othogonal

- matrix $A_{n\times n}$

$$AA^T = I \Rightarrow A \text{ is othogonal}$$

$$A^{-1} = A^T$$

$$A = (a_1, a_2, \cdots a_k)_{k\times k}$$

$a_i^T a_i = 1 \Leftrightarrow \|a\| = 1$

$a_i^T a_j = 0$

- Idempotent $A_{n\times n}$

$$AA = A$$

projection matrix

- positive definitness $B_{n\times n}$

$$a^T B a > 0$$

- nonnegative definitness $a^T B a \geq 0$

- Eigenvalues & eigenvectors $A_{n\times n}$

$$|A - \lambda I| = 0$$

$\lambda_1, \lambda_2, \ldots \lambda_n$ are roots

$\lambda_1, \lambda_2, \ldots \lambda_n$ are roots of $\}$

$\lambda_1, \ldots \lambda_n$ are called eigen values of A

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

$u_1, u_2, \ldots u_n \in R^n$

$$A u_j = \lambda u_j$$

$$\begin{cases} u_i^T u_i = 1 \\ u_i^T u_j = 0 \end{cases} \qquad U = (u_1\, u_2 \ldots u_n)$$

- $A_{n \times n}$ is symmetric,
  - eigenvalues of A are all real
  - rank(A) = # of nonzero eigen values
  - if A is non-negative definite then $\lambda_i \geq 0$

    if it p.d. the $\lambda_i > 0$

- $tr(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$

- $|A| = \prod_{i=1}^{n} \lambda_i$

- result: eigenvalues of idempotent matrix is either 1 or 0.

- Spectral decomposition $A_{n \times n}$ symmetric

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$U = (u_1, u_2 \cdots u_n)$$

$$A = \sum_{i=1}^{n} \underline{\underline{\lambda_i}}\, u_i u_i^T = UDU^T$$

- $A^{-1} = UD^{-1}U^T = \sum_{i=1}^{n} \lambda_i^{-1} u_i u_i^T$

- $A^{1/2} = \sum_{i=1}^{n} \lambda_i^{1/2} u_i u_i^T = UD^{1/2}U^T$

$$(A^{1/2} A^{1/2} = A)$$

$$A^{-1/2} = \sum_{i=1}^{n} \lambda_i^{-1/2} u_i u_i^T = UD^{-1/2}u^T$$

— x —

- random vector

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix}$$

$$\mu = E(Y) = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_p) \end{pmatrix}$$

$$\Sigma = Var(Y) = (\sigma_{ij})_{i,j=1}^{p}$$

$$\sigma_{ij} = E(Y_i - E(Y_i))(Y_j - E(Y_j))$$

$$a^TY = a_1Y_1 + \dots a_pY_p$$

$$E(a^TY) = a^T(E(Y))$$

$$\underline{0 \leq} Var(a^TY) = \underline{a^T\Sigma a}$$ nonnegative definite

$$Var(\underset{m\times p}{A}Y) = A\Sigma A^T_{m\times m}$$

— x —

- R calculation.