Chapter ①

- $AA^T = A^TA = I \rightarrow$ orthogonal matrix
- Eigen value $Ax = \lambda x$

$(A-\lambda I)x = 0$

ch. eqn $|A-\lambda I| = 0$

- Spectral decomposition

$$A = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \dots + \lambda_k e_k e_k^T$$

- $\Sigma = E[(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})^T]$

- $Cov(X^{(1)}, X^{(2)}) = \Sigma_{12} \rightarrow$ not necessarily symmetric or squre.

- Sum and difference of two random variable with same variance are uncorrelated.

- $Var(X_i) = E[(X_i - \mu_i)^2] = \sigma_{ii}$

- $Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = \sigma_{ij}$

- $\rho = I$, when all r.v are independent

- $\rho = V^{-1/2} \Sigma V^{-1/2}$ & $\Sigma = V^{1/2} \rho V^{1/2}$

- $\sigma_{11} = E[(X_1 - \mu_1)^2] = \sum_{\text{all } x_1} (x_1 - \mu_1)^2 P_1(x_1)$

- For variable

$E(cX_1) = c\mu_1$, $var(cX_1) = c^2 var(X_1)$

$Cov(aX_1, bX_2) = ab\, cov(X_1, X_2)$

$Var(aX_1 + bX_2) = a^2 var(X_1) + b^2 Var(X_2) + 2ab\, Cov(X_1, X_2)$

- For matrix & vector

$E(\underset{\sim}{c}^T \underset{\sim}{X}) = \underset{\sim}{c}^T \underset{\sim}{X}$, $Var(\underset{\sim}{c}^T \underset{\sim}{X}) = \underset{\sim}{c}^T \Sigma \underset{\sim}{c}$

$E(\underline{C}\underset{\sim}{X}) = \underline{C}\underset{\sim}{\mu}$, $Cov(\underline{C}\underset{\sim}{X}) = \underline{C}\Sigma\underline{C}^T$

- $Cov(AX^{(1)}, BX^{(2)}) = A\, Cov(X^{(1)}, X^{(2)}) B^T$

Chapter ②

$\underline{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$

- $E(S_n) = \frac{n-1}{n}\Sigma$

$= \Sigma - \underbrace{\frac{1}{n}\Sigma}_{\text{bias}}$

- $S = \frac{n}{n-1} S_n$ is unbiased estimator of Σ

- Sample mean matrix $\underset{p\times 1}{\underset{\sim}{\bar{X}}} = \frac{1}{n}\underline{X}^T \mathbb{1}$

- Sample variance matrix $S = \frac{1}{n-1}\underline{X}'(I - \frac{1}{n}\mathbb{1}\mathbb{1}^T)\underline{X}$

$S_n = \frac{1}{n}\underline{X}'(I - \frac{1}{n}\mathbb{1}\mathbb{1}^T)\underline{X}$

- sample standard deviation matrix $D^{1/2} = \begin{bmatrix} \sqrt{s_{11}} & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & & \sqrt{s_{pp}} \end{bmatrix}$

- sample correlation matrix

$R = D^{-1/2} S D^{-1/2}$ | $S = D^{1/2} R D^{1/2}$

- $A^TA = A \Rightarrow A$ is idempotent matrix

- $E(\underset{\sim}{c}^T \underset{\sim}{X}) = \underset{\sim}{c}^T \underset{\sim}{\mu}$ & $Var(\underset{\sim}{c}^T \underset{\sim}{X}) = \underset{\sim}{c}^T \Sigma c$

- For sample
  - An estimate of mean of $\underset{\sim}{c}^T \underset{\sim}{\mu}$ is $\underset{\sim}{c}^T \underset{\sim}{\bar{X}}_{p\times 1}$
  - An estimate of variance $\underset{\sim}{c}^T \Sigma \underset{\sim}{c}$ is $\underset{\sim}{c}^T S_n \underset{\sim}{c}$
  - An estimate of the population covariance of $\underset{\sim}{b}'\underset{\sim}{X}$, $\underset{\sim}{c}'\underset{\sim}{X}$ is $\underset{\sim}{b}^T S_n \underset{\sim}{c}$.
- The observation on linear combination are obtained by replacing x_1, x_2 & x_3 by...

$\underset{\sim}{b}'\underset{\sim}{x}_1 = 2x_{11} + 2x_{12} - x_{13}$

- $X \sim N(\mu, \sigma^2)$. Then $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

- $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$

1) If $z_1, \dots, z_m \overset{iid}{\sim} N_p(\underset{\sim}{0}, \Sigma)$. Then

$\sum_{j=1}^{m} z_j z_j^T \sim W_m(\Sigma)$

2) $A_1 \sim W_{m_1}(\Sigma)$ indep of $A_2 \sim W_{m_2}(\Sigma)$. then

$A_1 + A_2 \sim W_{m_1+m_2}(\Sigma)$

3) If $A \sim W_m(\Sigma)$, then $CAC^T \sim W_m(C\Sigma C^T)$

CamScanner

Chapter-4

- Univariate normal distribution pdf

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

- Test for normality: shapiro wilk, Kolmogorov smirnov
- pdf for a p-dimensional MVN r.v $\underset{\sim}{X} = (x_1, \dots, x_p)$

$$f(\underset{\sim}{X}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} (\underset{\sim}{X} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{X} - \underset{\sim}{\mu}) \right\}$$

We say $\boxed{\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)}$ - main

- pdf for bivariate normal distⁿ

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12})^2}} \exp\left\{ \frac{-1}{2(1-\rho_{12})^2} \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) \right] \right\}$$

- Properties of MVN distⁿ

$\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$. Then $\underset{\sim}{a}^T \underset{\sim}{X} = a_1 x_1 + \dots + a_p x_p \sim N(\underset{\sim}{a}^T \underset{\sim}{\mu}, \underset{\sim}{a}^T \Sigma \underset{\sim}{a})$, converse also

- For q linear combination $\bar{A}\underset{\sim}{X}$, $\bar{A}\underset{\sim}{X} \sim N_p(\bar{A}\underset{\sim}{\mu}, \bar{A}\Sigma\bar{A}^T)$
- $\underset{\sim}{X}_{p\times1} + \underset{\sim}{d}_{p\times1} \sim N(\underset{\sim}{\mu} + \underset{\sim}{d}, \Sigma)$
- All subset of $\underset{\sim}{X}$ are multivariate normally distributed

$\underset{\sim}{X} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}$, $\underset{\sim}{\mu} = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}$, $\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$, $X^{(1)} \sim N_q(\underset{\sim}{\mu}^{(1)}, \Sigma_{11})$

- If $\Sigma_{12} = \Sigma_{21}^T = 0 \Leftrightarrow X^{(1)}, X^{(2)}$ are independent.
- Univariate $Z^2 \Leftrightarrow Z^T Z$ multivariate
- If $Z_p \sim N_p(\underset{\sim}{0}, I)$. Then $ZZ^T \sim \chi^2_{(p)}$
- If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, then $\bar{\underset{\sim}{X}} \sim N_p(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$
- $(n-1)\underset{\sim}{S} \sim W_{(n-1)}$
- Law of Large number $Y_1, \dots, Y_n$ indepⁿ observation from population with $E(Y_i) = \mu$. Then $\bar{Y} \xrightarrow{P} \mu$ i.e $P(|\bar{Y} - \mu| < \varepsilon) = 1 \;\forall \varepsilon > 0$

So, $\bar{\underset{\sim}{X}}_i \xrightarrow{P} \underset{\sim}{\mu}$ and $\underset{\sim}{S} \xrightarrow{P} \Sigma$

- CLT: let $\underset{\sim}{X}_1, \dots, X_n$ be Independent observation from a popⁿ with mean $\underset{\sim}{\mu}$ and finite covariance Σ, then $\sqrt{n}(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}) \sim N_p(0, \Sigma)$ when n is large

also $n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}) \Sigma^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}) \sim \chi^2_{(p)}$

- χ^2-distⁿ in univariate & wishart distⁿ is multivariate

4 1 0
1 3 0
0 0 2

CS CamScanner