Multivariate Random Vector : Some Basic concepts

Multivariate random vector : $\underset{\sim}{X} = (X_1, \ldots, X_K)'$

X_is are r.v. on $(\Omega, \mathcal{F}, P)$

joint distribution f^n

$$F_{X_1,\ldots,X_K}(x_1,\ldots,x_K) = P(X_1 \le x_1, \ldots\ldots, X_K \le x_K)$$

$$\forall (x_1,\ldots,x_K) \in \mathbb{R}^K$$

For discrete multivariate distributions

$$F_{X_1,\ldots,X_K}(x_1,\ldots,x_K) = \sum_{i_1 \le x_1} \cdots \sum_{i_K \le x_K} P(X_1 = i_1, \ldots, X_K = i_K)$$

$P(X_1 = x_1, \ldots, X_K = x_K)$ – jt p.m.f.

$\forall (x_1, \ldots, x_K) \in$ Support of mult dist^n.

Marginal p.m.f. of X_1

$$P(X_1 = x_1) = \sum_{x_2} \cdots \sum_{x_K} P(X_1 = x_1, X_2 = x_2, \ldots X_K = x_K)$$

sly marginal for X_i $(i = 2(1)K)$

joint marginal of X_i, X_j (or any subset)

$$\underset{\text{except } x_i, x_j}{\sum \cdots \sum} P(X_1 = x_1, \ldots, X_K = x_K)$$

For continuous multivariate dist^{n}s

jt. d.f.

$$F_{X_1,\ldots X_K}(x_1,\ldots x_K) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_K} f_{X_1,\ldots X_K}(u_1,\ldots u_K)\prod_1^K du_i$$

$f_{X_1,\ldots X_K}(x_1,\ldots,x_K)$: jt p.d.f.

Marginal p.d.f. of X_i

$$f_{X_i}(x_i) = \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{K-1 \text{ fold, except over } x_i} f(u_1,\ldots u_K)\prod_{\substack{j=1\\ j\neq i}}^{K} du_j$$

jt marginal p.d.f. of (X_i, X_j) (or any subset)

$$f_{X_i,X_j}(x_i,x_j) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(u_1,\ldots u_K)\prod_{\substack{l=1\\ l\neq i;\, l\neq j}}^{K} du_l$$

K-2 fold, except over x_i & x_j

Expectation vector of $\underset{\sim}{X}$

$$\underset{K\times 1}{E(\underset{\sim}{X})} = \begin{pmatrix} EX_1 \\ \vdots \\ EX_K \end{pmatrix} = \underset{\sim}{\mu}$$

Let $\underset{q\times 1}{\underset{\sim}{Y}} = \underset{q\times K}{A}\,\underset{K\times 1}{\underset{\sim}{X}} + \underset{q\times 1}{\underset{\sim}{b}}$; A : matrix of constants

$\underset{\sim}{b}$: vector of constants

$$E(\underset{\sim}{Y}) = A\,E(\underset{\sim}{X}) + \underset{\sim}{b}$$

$$E(\underset{\sim}{\alpha}'\underset{\sim}{X}) = \underset{\sim}{\alpha}'E(\underset{\sim}{X})$$

$\underset{\sim}{\alpha} \in \mathbb{R}^K$

Note:

More generally, for a random matrix

$\underset{p\times q}{Z} = ((Z_{ij}))$ $\quad Z_{ij}$'s are r.v.s. on $(\Omega, \mathcal{F}, P)$

p.d.f. of Z is jt p.d.f of Z_{ij}'s $\left(\begin{array}{l} i = 1(1)p \\ j = 1(1)q \end{array} \right)$

$$E(Z) = ((E Z_{ij}))$$

If $Z \to \underset{m\times p}{B} Z C^{q\times n} + D^{m\times n}$

then $E Z = B E(Z) C + D$

Covariance matrix of $\underset{\sim}{X}$

$$Cov(\underset{\sim}{X}) = E(\underset{\sim}{X} - E(\underset{\sim}{X}))(\underset{\sim}{X} - E(\underset{\sim}{X}))'$$

$$\Sigma = E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})'$$

$(i, j)^{th}$ element of Σ is

$$\sigma_{ij} = E(X_i - \mu_i)(X_j - \mu_j)$$

i.e $\sigma_{ij} = Cov(X_i, X_j)$

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

Note: (i) Total variation in $\underset{\sim}{X} = tr\, \Sigma = \sum_{i=1}^{K} \sigma_{ii}$

(ii) Generalized variance of $\underset{\sim}{X} = |\Sigma|$

Let $\underset{\sim}{X}_{k\times 1}$ & $\underset{\sim}{Y}_{p\times 1}$ be two random vectors $\ni$

$E(\underset{\sim}{X}) = \underset{\sim}{\mu}_X$ & $E(\underset{\sim}{Y}) = \underset{\sim}{\mu}_Y$

$$\mathrm{Cov}(\underset{\sim}{X}, \underset{\sim}{Y}) = E(\underset{\sim}{X} - \underset{\sim}{\mu}_X)(\underset{\sim}{Y} - \underset{\sim}{\mu}_Y)'$$

<u>Note</u>: $\mathrm{Cov}(\underset{\sim}{X}) = \Sigma$ is always a sym matrix

<u>Result</u>: Characterization of a Covariance matrix

Any $p\times p$ real sym matrix Σ is a Covariance matrix iff Σ is positive semi definite (i.e. $\underline{\alpha}'\Sigma\underline{\alpha} \geq 0$ $\forall\ \underline{\alpha} \in \mathbb{R}^p$).

<u>Pf</u>: Suppose Σ is Cov matrix of $\underset{\sim}{X}$, then

$\forall\ \underline{\alpha} \in \mathbb{R}^p$

$$0 \leq V(\underline{\alpha}'\underset{\sim}{X}) = E(\underline{\alpha}'\underline{X} - \underline{\alpha}'\underline{\mu})^2 \quad ; E(\underline{X}) = \underline{\mu}$$

$$= E(\underline{\alpha}'(\underline{X} - \underline{\mu}))^2$$

$$= E(\underline{\alpha}'(\underline{X} - \underline{\mu}))(\underline{\alpha}'(\underline{X} - \underline{\mu}))'$$

$$= E(\underline{\alpha}'(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'\underline{\alpha})$$

$$= \underline{\alpha}'\Sigma\underline{\alpha}$$

$\Rightarrow$ $\underline{\alpha}'\Sigma\underline{\alpha} \geq 0$ $\forall\ \underline{\alpha} \in \mathbb{R}^p$

$\Rightarrow$ Σ is p.s.d.

Alternatively, suppose Σ is p.s.d. with rank r $(\leq p)$.

Then, $\Sigma = CC'$; C is $p \times r$ matrix of rank r

Let $\underset{\sim}{Y}$ be $r \times 1$ vector of indep r.v, s $\ni$

$E(\underset{\sim}{Y}) = \underset{\sim}{0}$ and $Cov(\underset{\sim}{Y}) = I_r$

Transform $\underset{\sim}{Y} \to \underset{\sim}{X} = C\underset{\sim}{Y}$

then $E(\underset{\sim}{X}) = C\,E(\underset{\sim}{Y}) = \underset{\sim}{0}$

$Cov(\underset{\sim}{X}) = E\,\underset{\sim}{X}\,\underset{\sim}{X}' = E\,(C\underset{\sim}{Y})(C\underset{\sim}{Y})'$

$= C\,E(\underset{\sim}{Y}\,\underset{\sim}{Y}')\,C' = CC' = \Sigma$

$= \cancel{C\Sigma C'}$

$\Rightarrow$ Σ is a covariance matrix

Note: $\underset{\sim}{X} \ni E(\underset{\sim}{X}) = \underset{\sim}{\mu}$; $Cov(\underset{\sim}{X}) = \Sigma$

Transform $\underset{\sim}{X} \to \underset{\sim}{Y} = A\underset{\sim}{X} + \underline{b}$

$Cov(\underset{\sim}{Y}) = E\,(\underset{\sim}{Y} - E(\underline{Y}))(\underline{Y} - E(\underline{Y}))'$

$= E\left(A\underline{X} + \underset{\sim}{b} - (A\underset{\sim}{\mu} + \underline{b})\right)\left(A\underline{X} + \underline{b} - (A\underset{\sim}{X} + \underline{b})\right)'$

$= E\,(A\underline{X} - A\underline{\mu})(A\underline{X} - A\underline{\mu})'$

$= A\,E\,(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'A' = A\Sigma A'$

Sp. case : Suppose $V = \text{diag}(\sigma_{11}, \ldots, \sigma_{kk})$

with $\sigma_{ii} > 0$

Take $A = (V^{1/2})^{-1}$

$$\underset{\sim}{X} \to \underset{\sim}{y} = A \underset{\sim}{X}$$

$$\text{Cov}(\underset{\sim}{y}) = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1} = \Sigma_y$$

$(i,j)^{th}$ entry of $\Sigma_y = \dfrac{\sigma_{ij}}{(\sigma_{ii}\sigma_{jj})^{1/2}} = \rho_{ij}$

ρ_{ij} is corrl^n (X_i, X_j)

i.e. Σ_y = corrl^n matrix of $\underset{\sim}{X}$

i.e. $\text{Corrl}^n(\underset{\sim}{X}) = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1} = \rho$.

Note: If Cov matrix of $\underset{\sim}{X}$ is not p.d., then w.p. 1, components of $\underset{\sim}{X}$ are linearly related.

Pf. If $\Sigma \not> 0$, then $\exists$ an $\underline{\alpha} \in \mathbb{R}^p$ $(\underline{\alpha} \neq \underline{0})$ $\ni$

$$0 = \underline{\alpha}' \Sigma \underline{\alpha} = V(\underline{\alpha}' \underline{X})$$

i.e. $P(\underline{\alpha}' \underset{\sim}{X} = \underline{\alpha}' \underline{\mu}) = 1$

i.e $P(\underline{\alpha}'(\underset{\sim}{X} - \underline{\mu}) = 0) = 1$

i.e. $\sum_i \alpha_i (X_i - \mu_i) = 0$ w.p. 1 for not all $\alpha_i = 0$

i.e. w.p 1 X_i s are linearly related.

Note: Partitions of Covariance matrix

$\underset{k\times 1}{\underset{\sim}{X}}$ is $\ni$ $E(\underline{X}) = \underline{\mu}$ & $Cov(\underline{X}) = \Sigma$

Consider the partition

$$\underset{\sim}{X} = \left(\begin{array}{c} \underset{\sim}{X}^{(1)}_{p\times 1} \\ \hline \underset{\sim}{X}^{(2)}_{k-p\times 1} \end{array} \right); \quad \underset{\sim}{\mu} = \left(\begin{array}{c} \underset{\sim}{\mu}^{(1)} \\ \hline \underset{\sim}{\mu}^{(2)} \end{array} \right).$$

$$Cov(\underline{X}) = E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'$$

$$= E \left(\begin{array}{c|c} (\underline{X}^{(1)} - \underline{\mu}^{(1)})(\underline{X}^{(1)} - \underline{\mu}^{(1)})' & (\underline{X}^{(1)} - \underline{\mu}^{(1)})(\underline{X}^{(2)} - \underline{\mu}^{(2)})' \\ \hline (\underline{X}^{(2)} - \underline{\mu}^{(2)})(\underline{X}^{(1)} - \underline{\mu}^{(1)})' & (\underline{X}^{(2)} - \underline{\mu}^{(2)})(\underline{X}^{(2)} - \underline{\mu}^{(2)})' \end{array} \right)$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

$$\Sigma_{11} = Cov(\underline{X}^{(1)}); \quad \Sigma_{22} = Cov(\underline{X}^{(2)})$$

$$\Sigma_{12} = Cov(\underline{X}^{(1)}, \underset{\sim}{X}^{(2)}) = \Sigma_{21}'$$

$$= \left(Cov(\underline{X}^{(2)}, X^{(1)})\right)'$$

If elements of $\underset{\sim}{X}^{(1)}$ are indep of the elements of $\underline{X}^{(2)}$ then $\Sigma_{12} = 0$, i.e $\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$

Converse is NOT true in general.

Random sampling from multivariate pop^{n}s

Multivariate pop^n with mean vector $\underset{\sim}{\mu}$ and covariance Σ

$\underset{\sim}{\mu}$ & Σ usually unknown

\# of unknowns $p + \frac{p(p+1)}{2}$ (for p dimensional pop^n)

Let $\underset{\sim}{X}_1 = \begin{pmatrix} X_{11} \\ X_{21} \\ \vdots \\ X_{p1} \end{pmatrix}, \ldots\ldots \underset{\sim}{X}_n = \begin{pmatrix} X_{1n} \\ X_{2n} \\ \vdots \\ X_{pn} \end{pmatrix}$ be n random samples from the pop^n (i.i.d)

Random Sample matrix

$$X_{p\times n} = \begin{pmatrix} X_{11} & \cdots & X_{1n} \\ X_{21} & \cdots & X_{2n} \\ - & - & - \\ X_{p1} & \cdots & X_{pn} \end{pmatrix}$$

Observation matrix / data matrix

$$\mathfrak{X}_{p\times n} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ x_{21} & \cdots & x_{2n} \\ - & - & - \\ x_{p1} & \cdots & x_{pn} \end{pmatrix}$$

$$= (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n) = \begin{pmatrix} \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_p' \end{pmatrix}$$

$\underset{\sim}{x}_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{pj} \end{pmatrix}$ j^{th} obsn vector
$j = 1(1)n$

$\underset{\sim}{y}_i' = (y_{i1}, \ldots, y_{in})$ - i^{th} variable obsns
$i = 1(1)p$

$\bar{\underset{\sim}{x}}$: observed sample mean vector

$$\bar{\underset{\sim}{x}} = \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \underline{y}_1' \underline{1}_n \\ \vdots \\ \underline{y}_p' \underline{1}_n \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \sum_{j=1}^{n} x_{1j} \\ \vdots \\ \sum_{j=1}^{n} x_{pj} \end{pmatrix} = \frac{1}{n}\mathcal{X}\,\underline{1}_n$$

Observed sample variance covariance matrix:

$$S_n = \frac{1}{n}\begin{pmatrix} \sum_{j=1}^{n}(x_{1j}-\bar{x}_1)^2 & \cdots & \cdots & \sum_1^n (x_{1j}-\bar{x}_1)(x_{pj}-\bar{x}_p) \\ & \sum_{j=1}^{n}(x_{2j}-\bar{x}_2)^2 & \cdots & \sum_1^n (x_{2j}-\bar{x}_2)(x_{pj}-\bar{x}_p) \\ - & - & - & - \\ & & & \sum_1^n (x_{pj}-\bar{x}_p)^2 \end{pmatrix}$$

or $S_{n-1} = \frac{n}{n-1} S_n$

Generalized sample variance : $|S_n|$ (or $|S_{n-1}|$)

Total sample variation : $\operatorname{tr} S_n$ (or $\operatorname{tr} S_{n-1}$)

Note that $(n-1)S_{n-1} = n S_n$

$$= \begin{pmatrix} \sum x_{1j}^2 & \sum x_{1j}x_{2j} & \cdots & \sum x_{1j}x_{pj} \\ & \sum x_{2j}^2 & \cdots & \sum x_{2j}x_{pj} \\ & & - - - & \\ & & & \sum x_{pj}^2 \end{pmatrix}$$

$$- n \begin{pmatrix} \bar{x}_1^2 & \bar{x}_1\bar{x}_2 & \cdots & \bar{x}_1\bar{x}_p \\ & \bar{x}_2^2 & \cdots & \bar{x}_2\bar{x}_p \\ - & - & \cdot & \cdot \\ & & & \bar{x}_p^2 \end{pmatrix}$$

i.e. $n\,\mathcal{S}_n = \mathcal{X}\mathcal{X}' - n\,\bar{\underset{\sim}{x}}\,\bar{\underset{\sim}{x}}'$

$$= \mathcal{X}\mathcal{X}' - n\left(\frac{1}{n}\mathcal{X}\underline{1}_n\right)\left(\frac{1}{n}\mathcal{X}\underline{1}_n\right)'$$

$$= \mathcal{X}\mathcal{X}' - \frac{1}{n}\mathcal{X}\underline{1}_n\underline{1}_n'\mathcal{X}'$$

$$= \mathcal{X}\left(I_n - \frac{1}{n}1_n1_n'\right)\mathcal{X}' = (n-1)\,\mathcal{S}_{n-1}$$

Alternately,

$$n\,\mathcal{S}_n = \mathcal{X}\mathcal{X}' - n\,\bar{\underline{x}}\,\bar{\underline{x}}'$$

$$= (\underset{\sim}{x}_1 : \cdots : \underset{\sim}{x}_n)\begin{pmatrix} \underset{\sim}{x}_1' \\ \vdots \\ \underline{x}_n' \end{pmatrix} - n\,\bar{\underline{x}}\,\bar{\underline{x}}'$$

$$= \sum \underline{x}_j\underline{x}_j' - n\,\bar{\underline{x}}\,\bar{\underline{x}}'$$

$$= \sum_{j=1}^{n}(\underline{x}_j - \bar{\underset{\sim}{x}})(\underline{x}_j - \bar{\underset{\sim}{x}})' = (n-1)\,\mathcal{S}_{n-1}$$

<u>Note</u>: Corresponding random matrices

$$n\,S_n = (n-1)\,S_{n-1}$$

$$= X\left(I_n - \frac{1}{n}1_n1_n'\right)X'$$

$$= \sum(\underline{X}_j - \bar{\underline{X}})(\underline{X}_j - \bar{\underline{X}})'$$

Multivariate Normal distⁿ

Defⁿ: Let $\underset{\sim}{X}_{p\times 1}$ be a random vector with $E(\underset{\sim}{X}) = \underline{\mu}$ and $Cov(\underset{\sim}{X}) = \Sigma$. We say that $\underset{\sim}{X}$ has a multivariate normal distⁿ ($\underset{\sim}{X} \sim N_p(\underline{\mu}, \Sigma)$) iff $\forall \underline{\alpha} \in \mathbb{R}^p$ ($\underline{\alpha} \neq \underline{0}$), $\underline{\alpha}'\underset{\sim}{X}$ has univariate normal distⁿ

Some results

If $\underset{\sim}{X} \sim N_p(\underline{\mu}, \Sigma)$, then

(i) $\underset{\sim}{X} - \underline{\mu} \sim N_p(\underline{0}, \Sigma)$

(ii) each component of $\underset{\sim}{X} \sim N_1$

(iii) Any subvector of $\underset{\sim}{X}$ has mult normal

(iv) $\underset{q\times p}{A}\underset{\sim}{X} + \underset{q\times 1}{\underset{\sim}{b}} \sim N_q(A\underset{\sim}{\mu} + \underset{\sim}{b}, A\Sigma A')$

(v) If $\Sigma > 0$, then p.d.f. of $\underset{\sim}{X}$ is

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

(vi) If $\Sigma \not> 0$, $\nexists$ p.d.f. of $\underline{X}$

(v) If $\underset{\sim}{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \underset{\sim}{X}^{(2)} \end{pmatrix}\begin{matrix} q\times 1 \\ p-q\times 1 \end{matrix}$, $\underset{\sim}{\mu} = \begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

$\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}$ indep iff $\Sigma_{12} = 0$

Random sampling from multivariate normal

$\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ r.s. from $N_p(\underline{\mu}, \Sigma)$, $\Sigma > 0$

Unbiased estimators:

$$\bar{\underset{\sim}{X}} = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{X}_i \quad ; \quad E\,\bar{\underset{\sim}{X}} = \underset{\sim}{\mu}$$

$$S_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (\underline{X}_i - \bar{\underline{X}})(\underline{X}_i - \bar{\underset{\sim}{X}})'$$

$$E\,S_{n-1} = \frac{1}{n-1} E\left(\sum \underline{X}_i \underline{X}_i' - n \bar{\underset{\sim}{X}} \bar{\underset{\sim}{X}}' \right)$$

$$= \frac{1}{n-1}\left(\sum_{1}^{n} E\,\underline{X}_i \underline{X}_i' - n\, E\,\bar{\underset{\sim}{X}} \bar{\underset{\sim}{X}}' \right)$$

$$= \frac{1}{n-1}\left(n(\Sigma + \underline{\mu}\,\underline{\mu}') - n\left(\Sigma/_n + \underline{\mu}\,\underline{\mu}'\right)\right)$$

$$= \Sigma$$

Thus $\underset{\sim}{\mu} \stackrel{\wedge}{=} \bar{\underset{\sim}{X}}$ & $\Sigma \stackrel{\wedge}{=} S_{n-1}$ (This is true for random sampling from any multivariate pop^n)

Maximum likelihood estimators

$\bar{\underset{\sim}{X}}$ is MLE of $\underset{\sim}{\mu}$

$S_n = \frac{1}{n} \sum (\underline{X}_i - \bar{\underset{\sim}{X}})(\underline{X}_i - \bar{\underline{X}})'$ is MLE of Σ

Dist^n:

$$\bar{\underset{\sim}{X}} \sim N_p(\underset{\sim}{\mu}, \Sigma/_n)$$

$$n S_n = (n-1) S_{n-1} \sim W_p(n-1, \Sigma)$$

a Wishart dist^n of order p with d.f. $n-1$ and associated covariance matrix Σ

Also $\bar{\underset{\sim}{X}}$ & S_{n-1} (or S_n) are independently distributed.