MOMENTS [H: 4.2]

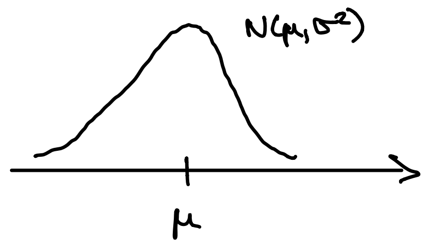
TMA 4267 [L2]
13.01.2017

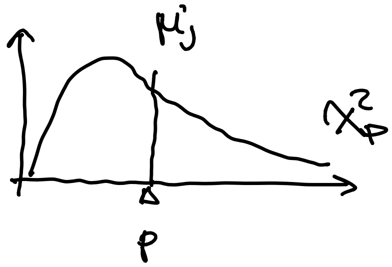
$\underset{p\times 1}{X}$ is a continuous random vector.

Mean (first order moment)

expected value
center of gravity
location of distribution

$$\underbrace{E(X_j)}_{\mu_j} = \int_{-\infty}^{\infty} x_j \, f_j(x_j) \, dx_j$$





DEF: Let X be a random vector

$$\underset{p\times 1}{\mu} = E(X) = \{E(X_j)\}$$

vector with this as the jth element

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

$\Rightarrow$ $E(X_j)$ is calculated from the marginal distribution of X_j and contains no information about dependencies between X_j and X_k, $j \neq k$

We also define a random matrix = matrix of RVs and $E(\underset{n\times p}{X}) = \{E(X_{ij})\}$ ← E of each element

Remember: $\underset{\text{scaler}}{X} \rightarrow E(aX+b) = \int_{-\infty}^{\infty} (ax+b)\, f(x)\, dx$

$$= a\int_{-\infty}^{\infty} x f(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx$$

$$= a\,E(X) + b$$

RULES FOR MEAN

$\underset{n\times p}{X}$ and $\underset{n\times p}{Y}$ are random matrices (n may be 1)

and $\underset{m\times n}{A}$ and $\underset{p\times q}{B}$ are constant matrices

1) $E(X+Y) = E(X) + E(Y)$

Proof: look at element $Z_{ij} = X_{ij} + Y_{ij}$ and see that $E(Z_{ij}) = E(X_{ij} + Y_{ij}) = E(X_{ij}) + E(Y_{ij})$

2) $E(AXB) = A\ E(X)\ B$

Proof: look at element (i,j) of $\overbrace{AXB}^{m\times q}$

$$e_{ij} = \sum_{k=1}^{n} a_{ik} \sum_{l=1}^{p} X_{kl}\, b_{lj}$$

and see that $E(e_{ij})$ is the element (ij) of $A\,E(X)\,B$.

Covariance matrix [H 4.2]

(1) From TMA4240/42 & ST1101: X and Y are scalar RVs with $E(X)=\mu_x$ and $E(Y)=\mu_Y$.

a) $\mathrm{Cov}(X,Y) = E\big((X-\mu_x)(Y-\mu_Y)\big)$

$$\overset{c)}{=} E(X \cdot Y) - \mu_x \mu_Y$$

Remember: $\mathrm{Cov}(X,Y)$ > 0 : high X and high Y
< 0 : high X and low Y
$= 0$: no linear pattern

b) $\mathrm{Var}(X) = \mathrm{Cov}(X,X) = E(X^2) - \mu_x^2$

d) $\mathrm{Cov}(X,Y) = 0$ if X and Y are independent, but not necessarily the opposite

(2) Now: random vectors

Ex
Chemical process
X = temp, pressure
Y = yield, quality

a) DEF [H 4.16]

$\underset{p\times 1}{X}$ and $\underset{q\times 1}{Y}$ are two random vectors.

$E(X) = \mu_x$ and $E(Y) = \mu_Y$

$$\underset{p\times q}{Cov(\mathbf{X}, \mathbf{Y})} = E\left[\underbrace{(\mathbf{X}-\mu_X)}_{p\times 1} \underbrace{(\mathbf{Y}-\mu_Y)^T}_{1\times q} \right]$$

$$= \Sigma_{\mathbf{X}\mathbf{Y}}$$

↑ sigma

and $Cov(\mathbf{X}, \mathbf{Y}) = 0$ if $\mathbf{X}$ and $\mathbf{Y}$ are independent.

If $p=2$, $q=3$

$$Cov(\mathbf{X}, \mathbf{Y}) = \begin{bmatrix} Cov(X_1, Y_1) & Cov(X_1, Y_2) & Cov(X_1, Y_3) \\ Cov(X_2, Y_1) & Cov(X_2, Y_2) & Cov(X_2, Y_3) \end{bmatrix}$$

Not a square matrix, but $Cov(\mathbf{X}, \mathbf{Y}) = Cov(\mathbf{Y}, \mathbf{X})^T$

b) $$\underbrace{Cov(\mathbf{X}, \mathbf{X})}_{p\times p} = E\left[(\mathbf{X}-\mu_X)(\mathbf{X}-\mu_X)^T \right]$$

$$= \begin{bmatrix} \underbrace{E[(X_1-\mu_1)(X_1-\mu_1)]}_{Var(X_1)} & \cdots & \underbrace{E[(X_1-\mu_1)(X_p-\mu_p)]}_{Cov(X_1, X_p)} \\ \vdots & \ddots & \\ \underbrace{E[(X_p-\mu_p)(X_1-\mu_1)]}_{Cov(X_p, X_1)} & & \underbrace{E[(X_p-\mu_p)(X_p-\mu_p)]}_{Var(X_p)} \end{bmatrix}$$

$= \Sigma = \Sigma_{xx}$

$= \underline{Cov(\mathbb{X})}$ or $Var(\mathbb{X})$

= "variance - covariance matrix"

Example: 4x4 matrix

Observations: $Cov(\mathbb{X}_1, \mathbb{X}_2) = Cov(\mathbb{X}_2, \mathbb{X}_1)$

$$= E[(\mathbb{X}_1 - \mu_1)(\mathbb{X}_2 - \mu_2)]$$

$$= \int_{-\infty}^{\infty}\int (x_1 - \mu_1)(x_2 - \mu_2) f(x_1, x_2)\, dx_1 dx_2$$

Σ is a symmetric matrix.

Cont.

c) $Cov(\mathbb{X}, Y) = E(\mathbb{X} Y^T) - \mu_x \cdot \mu_Y^T$

$\uparrow$ Rec Ex 1, P3

$$\underset{p\times p}{Cov(\mathbb{X})} = E(\underset{p\times 1}{\mathbb{X}}\, \underset{1\times p}{\mathbb{X}^T}) - \underset{p\times 1}{\mu_x}\, \underset{1\times p}{\mu_x^T}$$

d) As before if $\mathbb{X}_i$ and Y_j are independent

$\Rightarrow Cov(\mathbb{X}_i, Y_j) = 0$.

And if all pairs of $\mathbb{X}$s and Ys are independent

$\Rightarrow Cov(\mathbb{X}, Y) = \underset{p\times q}{0}$

Correlation

X_1 and X_2: $\quad Corr(X_1, X_2) = \dfrac{Cov(X_1, X_2)}{\sqrt{Var(X_1) \cdot Var(X_2)}}$

Rec Ex1. P1i do this in R

See matrix formula on slides

Ex: Given Σ find ρ:

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

Q: can information i Σ be condenced?

$tr(\Sigma) = \sum_{i=1}^{p} Var(X_i)$ = "total varience"

trace = sum of diagonal elements

$\det(\Sigma)$ = "generalized varience"

E and Cov of linear combinations

$\underset{p\times 1}{X}$ RV $\qquad E(X) = \mu_X$

$\qquad\qquad Cov(X) = \Sigma_X$

$\underset{k\times p}{C}$ constants

Consider: $\underset{k\times 1}{Z} = CX = \begin{bmatrix} \sum_{j=1}^{p} c_{1j} X_j \\ \sum_{j=1}^{p} c_{2j} X_j \\ \vdots \\ \sum_{j=1}^{p} c_{kj} X_j \end{bmatrix}$

vector of k linear combinations

1) $\mu_Z = E(Z) = E(CX) = E(CXI)$

$E(AXB)$ where $B = I$

$= C\,E(X)\,I = C\,E(X) = \underline{\underline{C\mu_X}}$

2) $\Sigma_Z = Cov(Z) = Cov(CX) \overset{?}{=} C\,Cov(X)\,C^T$

$= C\,\Sigma_X\,C^T$

Proof: $Cov(Z) = E\left[(Z-\mu_Z)(Z-\mu_Z)^T\right]$

$$= E\left((C\mathbf{X} - C\mu_x)(C\mathbf{X} - C\mu_x)^T \right)$$

$$= E\Big(C \underbrace{(\mathbf{X}-\mu_x)(\mathbf{X}-\mu_x)^T}_{\substack{Y \\ p\times p}} C^T \Big)$$

use $E(\underbrace{A}_{C}\underbrace{\mathbf{X}}_{Y}\underbrace{B}_{C^T}) = A\,E(\mathbf{X})\,B$)

$$= C \underbrace{E\left((\mathbf{X}-\mu_x)(\mathbf{X}-\mu_x)^T \right)}_{Cov(\mathbf{X})} C^T = \underline{\underline{C\,Cov(\mathbf{X})\,C^T}}$$

Example: Cork data and C matrix

$$\mathbf{X} = \underbrace{\begin{bmatrix} \mathbf{X}_N \\ \mathbf{X}_E \\ \mathbf{X}_S \\ \mathbf{X}_W \end{bmatrix}}_{p=4}$$

, but want

$N-S$
$E-W$
$(E+W)-(N+S)$ $\Big|$ $k=3$

$Y = C\mathbf{X}$, where

$$\underset{3\times 4}{C} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

$E(Y_1) =$ 1th element of $C\mu$

$Cov(Y_1, Y_3) =$ element $(1,3)$ of $C\Sigma C^T$