Lec 19 11/16/17 Math 621

1

Vector r.v.'s $\in \mathbb{R}^{n\times 1}$

$$\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad E[\vec{X}] = \vec{\mu} \in \mathbb{R}^{n\times 1}, \quad E[\vec{X}^T] = \mu^T$$

Matrix of r.v.'s $\in \mathbb{R}^{n\times m}$

$$X = \begin{bmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & & \\ X_{n1} & \cdots & X_{nm} \end{bmatrix}, \quad E(X) = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{n1} & & \mu_{nm} \end{bmatrix} = \mu \in \mathbb{R}^{n\times m}$$

We will define covariance more formally now

$$\Sigma := \mathrm{Cov}[\vec{X}] := E\Big[\underbrace{(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})^T}_{\text{"outer product"}}\Big] = E\begin{bmatrix} X_1 - \mu_1 \\ \vdots \\ X_n - \mu_n \end{bmatrix}\begin{bmatrix} X_1 - \mu_1 & \cdots & X_n - \mu_n \end{bmatrix} = \begin{bmatrix} E[(X_1-\mu_1)^2] & E[(X_1-\mu_1)(X_2-\mu_2)] & \cdots \\ E[(X_2-\mu_2)(X_1-\mu_1)] & E[(X_2-\mu_2)^2] & \\ \vdots & & \ddots \end{bmatrix}$$

$aa^T \to$ matrix, inner product $a^Ta \to$ scalar

$$= \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \\ \vdots & & \ddots \end{bmatrix}$$

Rule $(a+b)^T = a^T + b^T$

$$\Sigma := \mathrm{Cov}[\vec{X}] = E[(X-\mu)(X^T-\mu^T)] = E[XX^T - \mu X^T - X\mu^T + \mu\mu^T]$$

$(n\times 1)\times(1\times n)$

... we need more tools...

let $X \in \mathbb{R}^{n\times m}$, $A \in \mathbb{R}^{p\times n}$ $\quad AX \in \mathbb{R}^{p\times m}$

$$E(AX) = E\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{p1} & \cdots & a_{pn} \end{bmatrix}\begin{bmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & & \vdots \\ X_{n1} & \cdots & X_{nm} \end{bmatrix} = \begin{bmatrix} E\begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix}\begin{bmatrix} X_{11} \\ \vdots \\ X_{n1} \end{bmatrix} & \\ & \end{bmatrix}$$

$$a_{1\cdot}\,\mu_{\cdot 1} = a_{11}E[X_{11}] + \dots + a_{1n}E[X_{n1}]$$

$$= \begin{bmatrix} a_{1\cdot}\mu_{\cdot 1} & a_{1\cdot}\mu_{\cdot 2} & \cdots & a_{1\cdot}\mu_{\cdot m} \\ a_{2\cdot}\mu_{\cdot 1} & a_{2\cdot}\mu_{\cdot 2} & \cdots & a_{2\cdot}\mu_{\cdot m} \\ \vdots & & & \\ a_{p\cdot}\mu_{\cdot 1} & & \cdots & a_{p\cdot}\mu_{\cdot m} \end{bmatrix} \underbrace{\begin{bmatrix} a_{1\cdot} \\ a_{2\cdot} \\ \vdots \\ a_{p\cdot} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mu_{\cdot 1} & \mu_{\cdot 2} & \cdots & \mu_{\cdot m} \end{bmatrix}}_{E[X]} = A\,E[X]$$

let $X \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{n \times m}$

$$E[X+B] = E\begin{bmatrix} X_{11}+b_{11} & \cdots & X_{1m}+b_{1m} \\ \vdots & & \\ X_{n1}+B_{n1} & \cdots & X_{nm}+b_{nm} \end{bmatrix} = \begin{bmatrix} \mu_{11}+b_{11} & \cdots & \mu_{1m}+b_{1m} \\ \vdots & & \\ \mu_{n1}+b_{n1} & \cdots & \mu_{nm}+b_{nm} \end{bmatrix} = \mu + B = E(X) + B$$

$\Rightarrow E[AX+B] = A\,E(X) + B$ if dim's conform (o/w not yet defined)

Similarly $E[B+XA] = B + E(X)A$ if dim's conform

Back to the story...

$$\Sigma := Cov[\vec{X}] = E[XX^T] + E[-\mu X^T] + E(-X\mu^T) + E(\mu\mu^T)$$

$$= E[XX^T] - \mu \underbrace{E(X^T)}_{\mu^T} - \underbrace{E(X)}_{\mu}\mu^T + \mu\mu^T = E(XX^T) - \overbrace{\mu\mu^T}^{E(X)E(X^T)}$$

(Similar to $E(X^2) - \mu^2$ in the univariate setting)

Consider

$$Cov[A^T X] = E[(A^TX)(A^TX)^T] - E(A^TX)\,E[(A^TX)^T] = E[A^T XX^T A] - E(A^TX)\,E[X^TA]$$

$$= A^T E[XX^T]A - A^T E(X)\,E(X^T)A$$

$$= A^T(E(XX^T) - \mu\mu^T)A$$

$$= A^T Cov(X) A = A^T \Sigma A$$

Note $E[A^TX] = A^T\mu$

$$E[(A^TX)^T] = \underbrace{(A^T\mu)^T = \mu^T A^{TT}}_{\text{why?}} = \mu^T A$$

$(AB)^T = B^TA^T$ proven above

Seen this before...

if $A^T = \vec{a}^T$

$Cov[\vec{a}^TX] = \vec{a}^T Cov(X)\vec{a} = \vec{a}^T\Sigma\vec{a}$

Above is the more general case...

$\Rightarrow Cov(AX) = A\Sigma A^T$ ← more canonical formula

just replace A^T w/ A...

(MVN) multivariate normal of dim n

Let $\vec{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$ where $Z_1, \ldots, Z_n \overset{iid}{\sim} N(0,1) \Rightarrow \vec{Z} \sim N_n(\vec{0}, I_n)$

$Supp[\vec{Z}] = \mathbb{R}^n$

$E[\vec{Z}] = \vec{0}$ $Cov[\vec{Z}]$? well all $Cov[Z_i, Z_j] = 0$ if $i \neq j$

$Var[Z_i] = 1 \ \forall i$

PDF? $\Rightarrow Cov[\vec{Z}] = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I_n \leftarrow$ identity matrix of size n

$$\Rightarrow f_{\vec{Z}}(\vec{z}) = f_{\vec{Z}}(z_1, \ldots z_n) = f_{Z_1}(z) \cdots f_{Z_n}(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n} z_i^2} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\vec{z}^T\vec{z}}$$

since indep — since identically distr. — $(2\pi)^{n/2} = (\sqrt{2\pi})^n$

let $\vec{X} = \vec{Z} + \vec{c}$ where $\vec{c} \in \mathbb{R}^n$, a constant

$E[\vec{X}] = E[\vec{Z}] + \vec{c} = \vec{0} + \vec{c} = \vec{c} \Rightarrow \vec{X} \sim N_n(\vec{c}, I_n)$

$Var[\vec{X}] = I_n$

$$= f_{X_1}(x) \cdots f_{X_n}(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - c_i)^2} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\sum(x_i - c_i)^2} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\vec{x}-\vec{c})^T(\vec{x}-\vec{c})}$$

let $\vec{X} = A\vec{Z}$, $A \in \mathbb{R}^{m \times n}$, $\vec{Z} \in \mathbb{R}^{n \times 1}$, $\vec{X} \in \mathbb{R}^{m \times 1}$

$$\vec{X} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix} = \begin{bmatrix} a_{11}Z_1 + a_{12}Z_2 + \ldots + a_{1n}Z_n \\ a_{21}Z_1 + a_{22}Z_2 + \ldots + a_{2n}Z_n \\ \vdots \\ a_{m1}Z_1 + a_{m2}Z_2 + \ldots + a_{mn}Z_n \end{bmatrix} \begin{matrix} \sim N(\overset{0}{\cancel{\sum a_{1i}}}, \sum a_{1i}^2) \\ \sim N(\overset{0}{\cancel{\sum a_{2i}}}, \sum a_{2i}^2) \\ \vdots \\ \sim N(\overset{0}{\cancel{\sum a_{mi}}}, \sum a_{mi}^2) \end{matrix}$$

proven via ch.f.'s

$E[\vec{X}] = AE[\vec{Z}] = A\vec{0}_n = \vec{0}_m \in \mathbb{R}^m$

$\Sigma = Cov[\vec{X}] = A\,Cov[\vec{Z}]\,A^T = A I_n A^T = AA^T \in \mathbb{R}^{m \times m}$ Is $Cov[X_1, X_2] = 0$? No... they are dependent!

since they contain the same Z_i's!

$f_{\vec{X}}(\vec{x})$? Note $\vec{X} = A\vec{Z} = g(\vec{Z})$ i.e. a multivariate change of variables!

Snag... $dim[\vec{X}] \overset{must}{=} dim[\vec{Z}]$ $m \neq n$

Assume $m = n$... we will prove general case later...

g^{-1} since 1:1 function it has to be!!

$\exists h$ s.t. $\vec{X} = h(\vec{Z})$ What is it? $\vec{X} = A\vec{Z} \Rightarrow \vec{Z} = A^{-1}\vec{X}$ Also A is invertible!!