Lecture 20 Math 621 11/28/17

$\vec{Z} \sim N_n(\vec{0}_n, I_n)$

Is Σ symmetric?

$\Sigma \stackrel{?}{=} \Sigma^T = (AA^T)^T = A^{T^T}A^T = AA^T$ ✓

$E[\vec{Z}] = \vec{0}_n$, $Var[\vec{Z}] = I_n$

$\vec{\mu} \in \mathbb{R}^n$

if $\vec{X} = \vec{Z} + \vec{\mu} \sim N(\vec{\mu}, I_n)$

if $\vec{X} = A\vec{Z}$, $\mu = E[\vec{X}] = A\,E[\vec{Z}] = A\,\vec{0}_n = \vec{0}_m$, $\Sigma = Var[\vec{X}] = AA^T \in \mathbb{R}^{m \times m}$

$A \in \mathbb{R}^{m \times n}$

$f_{\vec{X}}(\vec{x})$? Note $\vec{X} = g(\vec{Z}) = A\vec{Z}$. Assume $A \in \mathbb{R}^{n \times n}$ for now.

$\vec{Z} = h(\vec{X}) = A^{-1}\vec{X}$, g is only 1:1 if A is full rank.

Let's do multivariable change of variables like before...

$$f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}\big(\underbrace{h(\vec{x})}_{A^{-1}x}\big)\,|J_h(\vec{x})| \qquad h(\vec{x}) = h_1(\vec{x}), \dots h_n(\vec{x})$$

$$J_h = \begin{bmatrix} \frac{\partial}{\partial x_1} h_1(\vec{x}) & \cdots & \frac{\partial}{\partial x_n} h_1(\vec{x}) \\ \vdots & & \\ \frac{\partial}{\partial x_1} h_n(\vec{x}) & \cdots & \frac{\partial}{\partial x_n} h_n(\vec{x}) \end{bmatrix}$$

$$h(\vec{x}) = \begin{bmatrix} h_1(\vec{x}) \\ h_2(\vec{x}) \\ \vdots \\ h_n(\vec{x}) \end{bmatrix} = B\vec{x}$$

let $B = A^{-1}$ just for notational simplicity

$$B = \begin{bmatrix} \vec{b}_{1\cdot} \\ \vdots \\ \vec{b}_{n\cdot} \end{bmatrix} \text{ or } \begin{bmatrix} \vec{b}_{\cdot 1} & \cdots & \vec{b}_{\cdot n} \end{bmatrix} \text{ or } \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}$$

$$h_1(\vec{x}) = \vec{b}_{1\cdot}\vec{x} = b_{11}x_1 + \dots + b_{1n}x_n$$

$$\left.\begin{aligned} &\frac{\partial}{\partial x_1}[h_1(\vec{x})] = b_{11} \\ &\frac{\partial}{\partial x_2}[h_1(\vec{x})] = b_{12} \\ &\quad\vdots \\ &\frac{\partial}{\partial x_n}[h_1(\vec{x})] = b_{1n} \end{aligned}\right\} \text{top row of } J_h$$

$$h_2(\vec{x}) = \vec{b}_{2\cdot}\vec{x} = b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n$$

$$\frac{\partial}{\partial x_1}[h_2(\vec{x})] = b_{21}$$

$$J_h = \det\left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & \cdots & & b_{nn} \end{bmatrix}\right) = \det(B) = \det(A^{-1})$$

Sidebar... vector derivatives...

$$\frac{\partial}{\partial \vec{x}}[C\vec{x}] = C$$

Many different definitions... no time for this...

$$\Rightarrow f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}(A^{-1}x)\,|\det(A^{-1})|$$

Recall $\det(A^{-1}) = \frac{1}{\det(A)}$ since $AA^{-1} = I$, $\det(AA^{-1}) = \det(I)$, $\det(A)\det(A^{-1}) = 1$

$$= \frac{1}{(2\pi)^{n/2}|\det(A)|} e^{-\frac{1}{2}\underbrace{(A^{-1}\vec{x})^T(A^{-1}\vec{x})}_{\vec{x}^T(A^{-1})^T A^{-1}\vec{x}}}$$

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

since $(AB)(AB)^{-1} = I$

$A\underbrace{BB^{-1}}_{I}A^{-1} = AA^{-1} = I$ ✓

Recall: $\Sigma = AA^T \Rightarrow \Sigma^{-1} = (AA^T)^{-1} = (A^T)^{-1}A^{-1}$

Does $(A^\top)^{-1} = (A^{-1})^\top$?

$$\begin{cases} AA^{-1} = I \Rightarrow (AA^{-1})^\top = I^\top = I \Rightarrow (A^{-1})^\top A^\top = I \\ (A^\top)^{-1} A^\top = I \end{cases}$$

by def of matrix inverse $\Rightarrow (A^\top)^{-1} = (A^{-1})^\top$

symmetric?

$$\Rightarrow \Sigma^{-1} = (A^{-1})^\top A^{-1} \qquad \Sigma^{-1\top} \overset{?}{=} \Sigma^{-1} = \left((A^{-1})^\top A^{-1}\right)^\top = (A^{-1})^\top (A^{-1})^{\top\top} = (A^{-1})^\top A^{-1} \checkmark$$

$$\Rightarrow f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}|\det(A)|} e^{-\frac{1}{2}\vec{x}^\top \Sigma^{-1}\vec{x}}$$

Just to get into canonical form...

Note $|\det(\Sigma)| = |\det(AA^\top)| = |\underbrace{\det(A)\det(A^\top)}_{=}| = |\det(A)^2|$

$$\Rightarrow \sqrt{|\det(\Sigma)|} = |\det(A)|$$

$$\Rightarrow f_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n |\det(\Sigma)|}} e^{-\frac{1}{2}\vec{x}^\top \Sigma^{-1}\vec{x}} = N_n(\vec{0}, \Sigma)$$

If $\vec{X} = A\vec{Z} + \vec{\mu} \sim N_n(\vec{\mu}, \Sigma) = \dfrac{1}{\sqrt{(2\pi)^n |\det(\Sigma)|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^\top \Sigma^{-1}(\vec{x}-\vec{\mu})}$

This is the generalization of $\left(\frac{x-\mu}{\sigma}\right)^2$ in the univariate normal i.e. "standardized distance" sqd

This "standardized distance" sqd is called "Mahalanobis distance" sqd (1936) by P.C. Mahalanobis (very famous Indian statistician). This is super useful!! I used this in my thesis to measure distance between two people for a clinical trial

If $\vec{X} \sim N_n(\vec{\mu}, \Sigma)$, $B \in \mathbb{R}^{m \times n}$, $B\vec{X} \sim ?$ We can't use multi. change of var's here easily...

How to prove?

Recall $\phi_X(t) = E[e^{itX}]$. What about for vector $\vec{X}$? $\phi_{\vec{X}}(\vec{t}) = E[e^{i\vec{t}^\top \vec{X}}]$

Rules... again... Same... for ch.f.'s

$$\phi_{\vec{X}_1+\vec{X}_2}(\vec{t}) = E\left[e^{i\vec{t}^T(\vec{X}_1+\vec{X}_2)}\right] = E\left[e^{i\vec{t}^T\vec{X}_1 + i\vec{t}^T\vec{X}_2}\right] = E\left[e^{i\vec{t}^T\vec{X}_1}e^{i\vec{t}^T\vec{X}_2}\right] \overset{\text{if indep.}}{=} \phi_{\vec{X}_1}(\vec{t})\,\phi_{\vec{X}_2}(\vec{t})$$

$$\phi_{A\vec{X}+\vec{c}}(\vec{t}) = E\left[e^{i\vec{t}^T(A\vec{X}+\vec{c})}\right] = E\left[e^{i\vec{t}^TA\vec{X} + i\vec{t}^T\vec{c}}\right] = e^{i\vec{t}^T\vec{c}}\,E\left[e^{i\vec{t}^TA\vec{X}}\right]$$

for appropriately sized $\vec{c}, A$

$$= e^{i\vec{t}^T\vec{c}}\,\phi_X(A^T\vec{t})$$

let $t'^T = t^TA$
$\Rightarrow t' = t'^{TT} = (t^TA)^T = A^Tt$

What is ch.f. for mvn $\vec{Z} \sim N_n(\vec{0}, I_n)$

$$E\left[e^{i\vec{t}\vec{Z}}\right] = \underbrace{\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}}_{\mathbb{R}^n} e^{i\vec{t}^T\vec{Z}} f_{\vec{Z}}(\vec{z})\,d\vec{z} = \int\cdots\int e^{i(t_1z_1+t_2z_2+\ldots+t_nz_n)}\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}(z_1^2+\ldots+z_n^2)}\,dz_1\ldots dz_n$$

$$= \prod_{i=1}^n \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{it_iz_i - \frac{1}{2}z_i^2}\,dz_i$$

we did a very similar exercise to those of the limiting ch.f. during the CLT proof

$$-\frac{1}{2}z^2 + itz = -\frac{1}{2}(z^2 - 2itz) = -\frac{1}{2}\left((z-it)^2 - i^2t^2\right) = -\frac{1}{2}\left((z-it)^2+t^2\right) = -\frac{1}{2}(z-it)^2 - \frac{t^2}{2}$$

$$\prod_{i=1}^n e^{-\frac{t_i^2}{2}}\underbrace{\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(z_i-it_i)^2}\,dz_i}_{\text{PDF for } N(it_i,1)\ =1} = e^{-\frac{1}{2}\sum t_i^2} = \boxed{e^{-\frac{1}{2}\vec{t}^T\vec{t}} = \phi_{\vec{Z}}(t)}$$

Note: there is no normal centered at "it". Thus this integral needs to be proved to be 1 some other way. To do so requires complex analysis and Cauchy's theorem. See https://www.dsprelated.com/freebooks/sasp/Gaussian_Integral_Complex_Offset.html for a proof.

$A \in \mathbb{R}^{n\times n}$

if $X = A\vec{Z} + \vec{\mu}$

$$\phi_{\vec{X}}(t) = e^{i\vec{t}^T\vec{\mu}}\,\phi_{\vec{Z}}(A^Tt) = e^{i\vec{t}^T\vec{\mu}}e^{-\frac{1}{2}t^TAA^T\vec{t}}$$

$\Rightarrow X \sim N_n(\vec{\mu}, \Sigma)$

s.t. $\Sigma = AA^T$

$$= e^{i\vec{t}^T\vec{\mu} - \frac{1}{2}t^T\Sigma\vec{t}}$$

$\vec{Y} = B\vec{X}$ How is $Y \sim$? $\phi_{\vec{Y}}(t) = \phi_{\vec{X}}(B^Tt) = e^{it^TB\mu - \frac{1}{2}t^TB\Sigma B^Tt} \Rightarrow Y \sim N_m(B\mu, B\Sigma B^T)$

$B \in \mathbb{R}^{m\times n}$ $m \neq n$ necessarily

Let's do another one... $\vec{Z} \sim N_n(\vec{0}, I_n)$

Now $X = B\vec{Z} + \vec{c}$ $\quad \phi_X(\vec{t}) = e^{i\vec{t}^T\vec{c}} \phi_Z(B^T\vec{t}) = e^{i\vec{t}^T\vec{c} - \frac{1}{2}\vec{t}^T BB^T \vec{t}}$ $\Rightarrow X \sim N_m(\vec{c}, \Sigma)$

where $B \in \mathbb{R}^{m \times n}$, $\vec{c} \in \mathbb{R}^m$ $\quad$ $\Sigma = BB^T$

dimension $\neq n$ necessarily

But Σ must be full rank!

Given $\vec{X}$ how do we standardize back to $\vec{Z}$?

$X = AZ + \mu \Rightarrow X - \mu = AZ \Rightarrow A^{-1}(X - \mu) = Z$ $\quad$ You can only get back if A is invertible!

$Z = A^{-1}X - A^{-1}\mu$

Why? Imagine $A = \vec{1}$, $\vec{c} = \vec{0}$

$X = \sum Z_i \sim N(0, n\sigma^2)$

$X \in \mathbb{R}^1 \Rightarrow$ no way to know what all n Z_i's would be! Makes sense...

$$\vec{Z}^T\vec{Z} = \sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$$

$$\Rightarrow \left(A^{-1}(X - \vec{\mu})\right)^T \left(A^{-1}(X - \mu)\right) = \vec{Z}^T\vec{Z} \sim \chi^2_n$$

$$(\vec{X}^T - \vec{\mu})\underbrace{(A^{-1})^T A^{-1}}_{\underbrace{(A^T)^{-1}A^{-1}}_{\Sigma^{-1}}}(\vec{X} - \vec{\mu})$$

$$\Rightarrow (\vec{X} - \vec{\mu})^T \Sigma^{-1} (\vec{X} - \vec{\mu}) \sim \chi^2_n$$

Distr. of Mahalanobis distance squared

Standardizing and squaring! "Sphering"

$$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \vec{X} \sim N_n(\mu\vec{1}, \sigma^2 I_n)$$

$$\Sigma = \sigma^2 I = AA^T \Rightarrow A = \sigma I$$

$\vec{X} = \sigma I \vec{Z} + \vec{\mu} = \sigma\vec{Z} + \vec{\mu}$ $\quad$ Same as univariate case!

Same as before...

$$(\vec{X} - \vec{\mu})^T \frac{1}{\sigma^2} (\vec{X} - \mu) \sim \chi^2_n$$

$$= \frac{1}{\sigma^2} (\vec{X} - \vec{\mu})^T (\vec{X} - \vec{\mu})$$

$$\frac{1}{\sigma^2} \sum (X_i - \mu)^2 \sim \chi^2_n$$