Lecture 05

* Suppose $X = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$,

Define: X is said to be normally distributed if every linear projection $a^T X$ is a 1-dim normal variable (i.e. for every $a \in \mathbb{R}^d$)

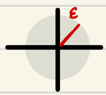
Result:

If X is mvN, then its distn is fully determined by $\mu = \mathbb{E}(X)$ and $\Sigma = Var(X)$.

* Multivariate MGF: $\phi_X(t) = \mathbb{E} e^{t^T X} = \mathbb{E} e^{t_1 X_1 + \cdots + t_n X_n}$

$$t = \begin{pmatrix} t_1 \\ \vdots \\ t_d \end{pmatrix} \in \mathbb{R}^d$$

$\underbrace{\phi_X(t) = \phi_Y(t)}_{\text{finite \& equal}}$, $\forall \|t\| < \varepsilon$



$\Rightarrow X \overset{d}{=} Y$.

* $\mathbb{E}(e^{t^T X})$, let $W = t^T X = t_1 X_1 + \cdots + t_d X_d$

by defn, W is 1-dim normal.

$$\varphi_X(t) = \mathbb{E}(e^W) = \varphi_W(1) = e^{\mathbb{E}(W) + \frac{1}{2} Var(W)}$$

Since:

$$\mathbb{E}(W) = \mathbb{E}(t^T X) = t^T \mu, \quad Var(W) = Var(t^T X) = t^T \Sigma t,$$

$$\varphi_X(t) = e^{t^T \mu + \frac{1}{2} t^T \Sigma t}.$$

* $\Sigma = Var(X)$ is PSD $[a^T \Sigma a \geq 0, \forall a \in \mathbb{R}^d]$

Further assumption: Σ is PD $\Rightarrow \Sigma = AA^T$ for some $\underset{d \times d}{A}$ with rank $= d$.

* Consider $Z_1, \dots, Z_d \overset{iid}{\sim} N_1(0,1)$, $Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_d \end{bmatrix}$

$$f_Z(z) = \prod_{j=1}^{d} f_{Z_j}(z_j)$$

$$= \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z_j^2}{2}\right\} = \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{\sum_{j=1}^{d} z_j^2}{2}\right\} = \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{z^T z}{2}\right\}$$

Now consider $Y = \mu + AZ$

$$\varphi_Y(t) = \mathbb{E}[e^{t^T Y}] = \mathbb{E}[e^{t^T\mu} \cdot e^{t^T A Z}] = e^{t^T\mu} * \mathbb{E}[e^{(t^T A)\cdot Z}]$$

Let $b = A^T t$, $b^T Z \sim N(0, b^T b)$,

$$\varphi_Y(t) = e^{t^T\mu} * \mathbb{E}(e^{b^T Z}) = e^{t^T\mu} \cdot e^{\frac{1}{2} b^T b} = e^{t^T\mu + \frac{1}{2} t^T A A^T t} = e^{t^T\mu + \frac{1}{2} t^T \Sigma t}$$

Conclusion : $X \overset{d}{=} Y$.

* Now we need the property A is full rank $\Rightarrow$ $|A| \neq 0$ and A^{-1} exists

Get the pdf of Y:

$Y = \mu + AZ$, $\quad f_Z(z) = \frac{1}{(2\pi)^{d/2}} \exp\{-\frac{z^T z}{2}\}$, $z \in \mathbb{R}^d$

Use change of variables :

$z = A^{-1}(Y - \mu)$, $\quad \frac{\partial z}{\partial Y} = A^{-1}$, $\quad \left|\frac{\partial z}{\partial Y}\right| = |A^{-1}| = |A|^{-1} = |\Sigma|^{-1/2}$

$\quad (|\Sigma| = |AA^T| = |A| \cdot |A^T| = |A|^2)$

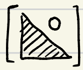
$$f_Y(y) = f_Z(A^{-1}(Y-\mu)) \cdot \left|\frac{\partial z}{\partial y}\right|$$

$$= \frac{1}{(2\pi)^{d/2}} \cdot \exp\{-\frac{1}{2}(y-\mu)^T (A^T)^{-1} A^{-1} (y-\mu)\} \cdot |\Sigma|^{-\frac{1}{2}}$$

$$= \frac{1}{(2\pi)^{d/2} \cdot |\Sigma|^{1/2}} \exp\{-\frac{1}{2}(y-\mu)^T \Sigma^{-1} (y-\mu)\}$$

* How to get A ? $\Sigma = AA^T$

- spectral decomposition : $\Sigma = P \Delta P^T$, then $A = P\Delta^{1/2}$
- Cholesky decomposition :

$A =$  lower trianglar matrix

* If $X_1 \sim N(0,1)$, $X_2 \sim N(0,1)$, is $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ (jointly) normal ?

✡ what if $X_1 \perp\!\!\!\perp X_2$?

$a_1 X_1 + a_2 X_2$ is normal $\longrightarrow$ satisfies the defn

✡ what if $X_1 \not\perp\!\!\!\perp X_2$?

$P(X_1 < X_2) = P(\underbrace{X_1 - X_2}_{\text{linear proj of } \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}} < 0)$

Find the case : X_1 and X_2 are marginally normal, but $P(X_1 - X_2 < 0)$ is different from a normal would require to be.

A counterexample:

Let $X \sim N(0,1)$, $Y = S \cdot X$, where $S \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

Then $\varphi_Y(t) = \mathbb{E}\left[e^{tS\cdot X}\right] = \mathbb{E}\left[\mathbb{E}(e^{tS\cdot X} \mid S)\right]$

$$= \mathbb{E}\left[\varphi_X(tS)\right] = \mathbb{E}\left[e^{\frac{1}{2}t^2S^2}\right] = \frac{1}{2}\cdot e^{\frac{1}{2}t^2} + \frac{1}{2}e^{\frac{1}{2}t^2}$$

$$= e^{\frac{1}{2}t^2}.$$

So $Y \sim N(0,1)$.

However, $X+Y = (1+S)\cdot X$

$$\mathbb{P}(X+Y=0) = \mathbb{P}((1+S)\cdot X = 0) = \mathbb{P}(S=-1) = \frac{1}{2}.$$

$\hookrightarrow$ This is a property that cannot be achieved by Gaussian variable.

So $X+Y$ is not Gaussian $\Rightarrow$ Defn of MVG is violated.

* Suppose $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ random sample

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

sample mean   sample variance

① $\tilde{X} \sim N(\mu, \frac{\sigma^2}{n})$

② $(n-1)S^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2 \sim Ga\left(\frac{n-1}{2}, \frac{\sigma^2}{2}\right)$

$$\frac{(n-1)S^2}{\sigma^2} \sim Ga\left(\frac{n-1}{2}, \frac{1}{2}\right) = \chi^2_{n-1}$$

③ $\bar{X} \perp\!\!\!\perp S^2$.

Special case: $\mu = 0$, $\sigma = 1$ $\tilde{X}_i = \frac{X_i - \mu}{\sigma}$

$$\bar{X} = \frac{1}{n}\mathbb{1}^T X \sim N\left(\frac{1}{n}\mathbb{1}^T \cdot 0, \frac{1}{n}\mathbb{1}^T \cdot I \cdot \mathbb{1}\frac{1}{n}\right) = N\left(0, \frac{1}{n}\right)$$

Consider

$$P = \begin{bmatrix} \frac{1}{\sqrt{n}}, & \frac{1}{\sqrt{n}}, & \frac{1}{\sqrt{n}}, & \cdots, & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n(n-1)}}, & \frac{1}{\sqrt{n(n-1)}}, & \frac{1}{\sqrt{n(n-1)}}, & \cdots, & -\frac{n-1}{\sqrt{n(n-1)}} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{3\times4}}, & \frac{1}{\sqrt{3\times4}}, & \frac{1}{\sqrt{3\times4}}, & \cdots, & 0 \\ \frac{1}{\sqrt{2\times3}}, & \frac{1}{\sqrt{2\times3}}, & -\frac{1}{\sqrt{2\times3}}, & \cdots, & 0 \\ \frac{1}{\sqrt{1\times2}}, & -\frac{1}{\sqrt{1\times2}}, & 0, & \cdots, & 0 \end{bmatrix}$$