# Lecture #3

$X \sim N_m(\mu, \Sigma)$

$$\boxed{\begin{array}{l} \text{If} \quad Z \sim N_m(0, I_m) \quad \text{and} \quad X = \mu + \Sigma^{1/2} Z \\ \text{then} \quad X \sim N_m(\mu, \Sigma) \end{array}}$$

$X \sim N_m(\mu, \Sigma)$

Fact $\mathbb{E}X = \mu \qquad Var(X) = \Sigma$

$$\boxed{Z_i \sim N(0,1) \qquad \mathbb{E}Z_i = 0, \quad Var Z_i = 1}$$

$\underline{Z} \sim N_m(\underline{0}, I_m)$ then $\mathbb{E}\underline{Z} = \underline{0}$, $Var\underline{Z} = I_m$

$\underline{Z} = (Z_1, \dots, Z_m)$ each $Z_i$ indep $N(0,1)$

$$\mathbb{E}\underline{Z} = \begin{pmatrix} \mathbb{E}Z_1 \\ \vdots \\ \mathbb{E}Z_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$Cov(Z_i, Z_j) = \begin{cases} 0 & i \neq j \quad \text{by indep.} \\ Var(Z_i) = 1 & i = j \end{cases}$$

so $Var(Z) = I_m$

$\to X \sim N_m(\mu, \Sigma) \qquad X = \mu + \Sigma^{1/2} \cdot Z$

$\underline{Z} \sim N(0, I)$

$$\mathbb{E}X = \mathbb{E}(\mu + \Sigma^{1/2} Z) = \mu + \mathbb{E}(\Sigma^{1/2} Z)$$

$$= \mu + \Sigma^{1/2} \underbrace{EZ}_{0} = \mu$$

$$\mathrm{Var}(X) = \mathrm{Cov}(X, X)$$

$$= \mathrm{Cov}(\mu + \Sigma^{1/2} Z, \mu + \Sigma^{1/2} Z)$$

(invariance under shifts)

$$= \mathrm{Cov}(\Sigma^{1/2} Z, \Sigma^{1/2} Z)$$

$$= \Sigma^{1/2} \cdot \mathrm{Cov}(Z, Z) \cdot \Sigma^{1/2}$$

$$= \Sigma^{1/2} \mathrm{Var}(Z) \Sigma^{1/2}$$

$$= \Sigma^{1/2} \cdot I_m \cdot \Sigma^{1/2} = \Sigma$$

---

Shape of MVN

$$f_X(x) = \frac{1}{(2\pi)^{m/2}} \cdot \sqrt{\det \Sigma^{-1}} \, e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

$$E = \{ x : (x-\mu)^T \Sigma^{-1} (x-\mu) = c^2 \}$$

if $x \in E$

$$f_X(x) = \frac{1}{(2\pi)^{m/2}} \sqrt{\det \Sigma^{-1}} \, e^{-\frac{1}{2} c^2}$$

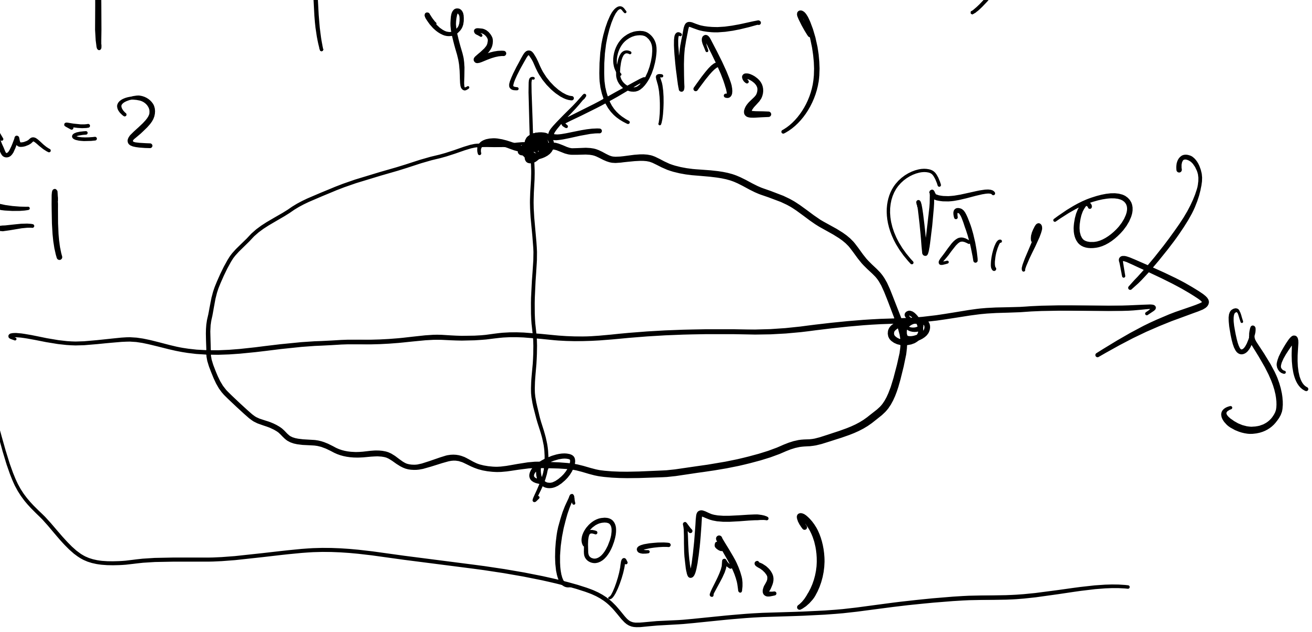
<u>Recall</u> Ellipsoid in $\mathbb{R}^m$

$$\left\{ y \in \mathbb{R}^m : \frac{y_1^2}{\lambda_1} + \dots + \frac{y_m^2}{\lambda_m} = c^2 \right\}$$

for fixed $\lambda_i > 0$, $c \in \mathbb{R}$

$m = 2$

$c = 1$




$$\Sigma = U \Lambda U^T$$

$$\Sigma^{-1} = U \Lambda^{-1} U^T$$

change coordinates:

$$y = U^T (x - \mu)$$

$$(x-\mu)^T \Sigma^{-1} (x-\mu) =$$

$$= (x-\mu)^T U \Lambda^{-1} \underbrace{U^T (x-\mu)}_{y}$$

$$= y^T \Lambda^{-1} y =$$

$$= \frac{y_1^2}{\lambda_1} + \dots + \frac{y_m^2}{\lambda_m}$$


$\Sigma \in S_+^m$

Mahalanobis Distance, $\mathbb{R}^m$

$$\boxed{\|x-y\|_\Sigma = \sqrt{(x-y)^T \Sigma^{-1} (x-y)}}$$

**Prop**: $X \sim N_m(\mu, \Sigma)$ then

$$\|X-\mu\|_\Sigma^2 \sim \chi_m^2$$

**Recall**: If $Z_1, \dots, Z_m \overset{iid}{\sim} N(0,1)$

$$Z_1^2 + \dots + Z_m^2 \sim \chi_m^2$$

**Proof** $\underline{Z} = \Sigma^{-1/2}(X-\mu)$

$$Z \sim N_m(0, I)$$

$$\|X-\mu\|_\Sigma^2 = (X-\mu)^T \Sigma^{-1} (X-\mu)$$

$$= (X-\mu)^T \Sigma^{-1/2} \underbrace{\Sigma^{-1/2}(X-\mu)}_{Z}$$

$$= Z^T Z = Z_1^2 + \dots + Z_m^2$$

$$\sim \chi_m^2$$

---

characteristic functions

$X$ random variable

$$\varphi_X(s) := \mathbb{E}\, e^{i \cdot s \cdot X}$$

$X \in \mathbb{R}^m$ random vector

$$\varphi_X(\underline{t}) := \mathbb{E}\, e^{i\, \underline{t}^T \cdot X}$$

$\underline{t} \in \mathbb{R}^m$

⚠ the characteristic function defines the distribution uniquely.

Recall: $Z \sim N(0,1)$

$$\psi_Z(s) = e^{-\frac{1}{2}s^2}$$

$$\underline{Z} \sim N_m(\underline{0}_m, I_m)$$

$$\psi_{\underline{Z}}(\underline{t}) = \mathbb{E}\, e^{i\, t^T \cdot Z}$$

$$= \mathbb{E}\, e^{i(t_1 Z_1 + \dots + t_m Z_m)}$$

$$= \mathbb{E} \prod_{j=1}^{m} e^{i t_j Z_j}$$

$$\overset{\text{indep}}{=} \prod_{j=1}^{m} \underbrace{\mathbb{E}\, e^{i t_j Z_j}}_{e^{-\frac{1}{2}t_j^2}} = e^{-\frac{1}{2}\sum_{j=1}^{m} t_j^2}$$

$$\mathbb{E}\, e^{i t_j Z_j} = \psi_{Z_j}(t_j)$$

$$= e^{-\frac{1}{2}t^T t} = e^{-\frac{1}{2}\|t\|^2}$$

$\underline{10}$ $X \sim N(\mu, \sigma^2)$

$$X = \mu + \sigma \cdot Z \qquad Z \sim N(0,1)$$

$$\psi_X(s) = \mathbb{E} e^{isX} = \mathbb{E} e^{is(\mu + \sigma Z)}$$

$$= e^{is\mu} \cdot \underbrace{\mathbb{E}\, e^{is\sigma Z}}_{\psi_Z(s \cdot \sigma) = e^{-\frac{1}{2}s^2\sigma^2}}$$

$$= e^{is\mu - \frac{1}{2}s^2\sigma^2}$$

---

$$\boxed{\begin{aligned} &X \sim N_m(\mu, \Sigma) \\ &\phi_X(\underline{t}) = e^{i\mu^T t - \frac{1}{2}t^T \Sigma t} \end{aligned}}$$

---

$$X \sim N_m(\mu, \Sigma)$$

$A \in \mathbb{R}^{p \times d} \qquad b \in \mathbb{R}^p$

$$AX + b \sim N_p( \quad , \quad )$$

mean $= A\mu + b$

covariance $= A \Sigma A^T$

check the characteristic function

$$\Psi_{AX+b}(\underset{\substack{\cup \\ \mathbb{R}^p}}{\underline{t}}) = \mathbb{E} e^{i \underline{t}^T \cdot (AX+b)}$$

$$= e^{i t^T \cdot b} \cdot \mathbb{E} e^{i t^T A X}$$

$$= e^{i t^T b} \underbrace{\mathbb{E} e^{i (A^T t)^T X}}_{\Psi_X(A^T t)}$$

$$= e^{it^T b} e^{i\mu^T (A^T t) - \frac{1}{2}(A^T t)^T \Sigma (A^T t)}$$

$$= e^{i \underbrace{(b + A\mu)^T}_{} t - \frac{1}{2} t^T \underbrace{A \Sigma A^T}_{} t}$$

characteristic functio of

$$N_p\left(b + A\mu, A\Sigma A^T\right)$$

---

MARGINAL AND CONDITIONAL DISTR.

$$X = (X_1, \dots, X_m)$$

split $X$ into $X_A, X_B$

eg $X = (X_1, X_2, X_3, X_4)$

$X_A = (X_1, X_3)$

$X_B = (X_2, X_4)$

the corresponding split of $\mu$

$\mu_A, \mu_B$

decomposition of $\Sigma$

$\Sigma_{AA}, \Sigma_{AB}, \Sigma_{BA} = \Sigma_{AB}^T$

$\Sigma_{BB}$

e.g. $\Sigma_{AB} = \mathrm{Cov}(X_A, X_B)$

$= \begin{pmatrix} \Sigma_{12} & \Sigma_{14} \\ \Sigma_{32} & \Sigma_{34} \end{pmatrix}$

e.g. $X_1, X_2, X_3, X_4$

$|A| + |B| = m$

$k \qquad m-k$

$\mu_A \in \mathbb{R}^k \qquad \mu_B \in \mathbb{R}^{m-k}$

$\Sigma_{AA} \in \mathbb{R}^{k \times k}$

$\Sigma_{AB} \in \mathbb{R}^{k \times (m-k)}$

$\Sigma_{BB} \in \mathbb{R}^{(m-k) \times (m-k)}$

<u>Prop</u> $X \sim N_m(\mu, \Sigma)$

split $X_A^{k}$, $X_B^{m-k}$

- the marginal distr. of $X_A$ is $k$-dim Gaussian with mean

$\mathbb{E}(X_A) = \mu_A$

$\mathrm{Var}(X_A) = \Sigma_{AA}$ $k \times k$

- the conditional distribution of $X_B \mid X_A = x_A$ is $(m-k)$-variate Gaussian with

$$\mathbb{E}(X_B \mid X_A = x_A) = \mu_B + \boxed{\Sigma_{BA}\Sigma_{AA}^{-1}}(x_A - \mu_A)$$

$$\boxed{\operatorname{Var}(X_B \mid X_A = x_A) = \Sigma_{BB} - \Sigma_{BA}\Sigma_{AA}^{-1}\Sigma_{AB}}$$

For the marginal distribution

$$(\underbrace{X_1, \dots, X_k}_{X_A}, \dots, X_m)$$

CLAIM

$$X_A \sim N_k\left(\begin{pmatrix}\mu_1\\ \vdots \\ \mu_k\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \cdots & \Sigma_{1k}\\ \\ \Sigma_{k1} & \cdots & \Sigma_{kk}\end{pmatrix}\right)$$

$$\begin{pmatrix}X_1\\ \vdots \\ X_k\end{pmatrix} = \left(I_k \;\vdots\; 0_{k\times(m-k)}\right) \cdot \begin{pmatrix}X_1\\ \vdots \\ X_k \\ \vdots \\ X_m\end{pmatrix}$$

$$\underbrace{\phantom{M}}_{k \times (m-k)}$$

$M$

$$X_A = M \cdot X$$

$$X_A \sim N_k\left(M \cdot \mu,\ M \Sigma M^T\right)$$

$$M \cdot \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}, \qquad M \Sigma M^T = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1k} \\ \vdots & & \\ \Sigma_{k1} & \cdots & \Sigma_{kk} \end{pmatrix}$$

---

$$X = (X_1, \ldots, X_m)$$

$$X_i \perp\!\!\!\perp X_j \implies \mathrm{Cov}(X_i, X_j) = 0$$

$$\cancel{\Longleftarrow} \qquad Z \sim N(0,1)$$

$\mathrm{cov}(\varepsilon, \varepsilon^T) \neq$

⊙ If $X \sim N_m(\mu, \Sigma)$ and $\Sigma_{ij} = 0$ then $X_i \perp\!\!\!\perp X_j$.

<u>Proof</u>:

$$(X_i, X_j) \sim N_m\left( \begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix}, \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} \right)$$

$$\begin{pmatrix} \Sigma_{ii} & 0 \\ 0 & \Sigma_{jj} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\Sigma_{ii}} & 0 \\ 0 & \frac{1}{\Sigma_{jj}} \end{pmatrix}$$

$$f(x_i, x_j) = \frac{1}{2\pi} \frac{1}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \cdot$$

$$\cdot\, e^{\underbrace{-\frac{1}{2\Sigma_{ii}}(x_i - \mu_i)^2 - \frac{1}{2\Sigma_{jj}}(x_j - \mu_j)^2}}$$

$$\left(\begin{pmatrix} x_i \\ x_j \end{pmatrix} - \begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix}\right)^T \begin{pmatrix} \frac{1}{\Sigma_{ii}} & 0 \\ 0 & \frac{1}{\Sigma_{jj}} \end{pmatrix} \left(\begin{pmatrix} x_i \\ x_j \end{pmatrix} - \begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix}\right)$$

$$= \left( \frac{1}{\sqrt{2\pi \Sigma_{ii}}} e^{-\frac{1}{2\Sigma_{ii}}(x_i - \mu_i)^2} \right) \cdot$$

$$\left( \frac{1}{\sqrt{2\pi \Sigma_{jj}}} e^{-\frac{1}{2\Sigma_{jj}}(x_j - \mu_j)^2} \right)$$

$$\Rightarrow X_i \perp\!\!\!\perp X_j$$

---

$$X_i \perp\!\!\!\perp X_j \mid \text{rest}$$

$X_A = (X_i, X_j)$ $X_B =$ all the other $X_k$'s

$$\mathrm{Var}(X_A | X_B) = \underbrace{\Sigma_{AA} - \Sigma_{AB} \Sigma_{BB}^{-1} \Sigma_{BA}}_{2\times 2}$$

off-diagonal entry

$$\Sigma_{ij} - \Sigma_{iB} \Sigma_{BB}^{-1} \Sigma_{Bj}$$

so $X_i \perp\!\!\!\perp X_j \mid X_B$

$$\Sigma_{ij} = \Sigma_{iB} \Sigma_{BB}^{-1} \Sigma_{Bj}$$