STA 437, Lecture 2

$X = (X_1, \dots, X_m) \in \mathbb{R}^m$

mean vector $\mu = \mathbb{E}X$

covariance matrix $\Sigma = \mathrm{Var}(X)$

$$\mathrm{Var}(X) = \mathbb{E}(X-\mu)(X-\mu)^T$$

$\Sigma_{ij} = \mathrm{Cov}(X_i, X_j)$

Σ is PD

$\odot$ $X \in \mathbb{R}^p$, $Y \in \mathbb{R}^q$

$$\Sigma_{XY} = \mathrm{Cov}(X, Y) = \mathbb{E}\underbrace{\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)^T\right]}_{p \times q}$$

$$(\Sigma_{XY})_{ij} = \mathrm{Cov}(X_i, Y_j)$$

note:

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

Lemma $X, X' \in \mathbb{R}^m$ r. vectors

$\rightarrow$ $\mathbb{E}(X+X') = \mathbb{E}(X) + \mathbb{E}X'$

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

$\mathbb{E}(AX+b) = A \cdot \mathbb{E}X + b$

Proof: fix i^{th} coord.

$$L_i = \mathbb{E}\left[(AX+b)_i\right]$$
$$= \mathbb{E}\left[\sum_{j=1}^{m} A_{ij}X_j + b_i\right]$$
$$\overset{1D}{=} \sum_{j=1}^{m} A_{ij}\,\mathbb{E}X_j + b_i$$
$$= (A \cdot \mathbb{E}X + b)_i = R_i$$

Lemma $S \in \mathbb{R}^{m \times n}$ r. matrix

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

$$\mathbb{E}[A \cdot S \cdot B] = A \cdot \mathbb{E}[S] \cdot B$$

Lemma $X, X' \in \mathbb{R}^p \quad Y \in \mathbb{R}^q$

1) $\underset{p \times q}{\mathrm{Cov}(X, Y)} = \mathrm{Cov}(Y, X)^T$

(symmetry)

2) $\mathrm{Cov}(X + X', Y) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X', Y)$

3) $A \in \mathbb{R}^{m \times p} \qquad B \in \mathbb{R}^{q \times n}$

$$\mathrm{Cov}(AX, BY) = A \cdot \mathrm{Cov}(X, Y) \cdot B^T$$

Data $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^m$

from distr. with

mean is μ and

covariance Σ,

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

sample mean

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i =$$

sample covariance matrix

$$\boxed{S_n = \frac{1}{n}\sum_{i=1}^{n} (\underline{x}_i - \bar{x}_n)(\underline{x}_i - \bar{x})^T}$$

$$\underline{X} = \begin{pmatrix} — \underline{x}_1 — \\ — \underline{x}_2 — \\ \vdots \\ — \underline{x}_n — \end{pmatrix} \in \mathbb{R}^{n \times m}$$

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i = \frac{1}{n} \underline{X}^T \cdot \mathbb{1}_n$$

$$\mathbb{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

the centering matrix

$$H = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^T$$

$\rightarrow$ $n \times n$, symmetric

$$H = H^T$$

$$\rightarrow H \cdot \mathbb{1}_n = \mathbb{1}_n - \frac{1}{n} \cdot \mathbb{1}_n \underbrace{(\mathbb{1}_n^T \cdot \mathbb{1}_n)}_{n}$$

$$= \mathbb{1}_n - \mathbb{1}_n = 0_n$$

$$\rightarrow \quad H^2 =$$

$$\left(I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^T\right)\left(I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^T\right)$$

$$= I_n - \frac{2}{n} \mathbb{1}_n \mathbb{1}_n^T + \frac{1}{n^2} \mathbb{1}_n \underbrace{(\mathbb{1}_n^T \cdot \mathbb{1}_n)}_{n} \cdot \mathbb{1}^T$$

$$= I_n - \frac{2}{n} \mathbb{1}_n \mathbb{1}_n^T + \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^T$$

$$= H \qquad (H^2 = H)$$

Why "centering"?

$$\underset{n\times n}{H}\;\underset{n\times m}{\underline{X}} = \left(I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^T\right)\underline{X}$$

$$= \underline{X} - \frac{1}{n}\mathbb{1}\mathbb{1}^T\underline{X}$$

$$= \underline{X} - \mathbb{1}\cdot\bar{x}_n^T$$

$$\text{centered data} = \begin{pmatrix} \cdots & \underline{x}_1 - \bar{x}_n & \cdots \\ \cdots & \underline{x}_2 - \bar{x}_n & \cdots \\ & \vdots & \\ \cdots & x_n - \bar{x}_n & \cdots \end{pmatrix}$$

<u>Note:</u> $y_i := \underline{x}_i - \bar{x}_n$

$$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}_n\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n}\bar{x}_n$$

$$= \bar{x}_n - \bar{x}_n = \underline{0}$$

$$\frac{1}{n}(HX)^T(HX) =$$

earlier:

$$\frac{1}{n}X^TX = \frac{1}{n}\sum x_i x_i^T$$

n

$$= \frac{1}{n}\sum_{i=1}^{n}(\underset{\sim}{x}_i - \bar{x})(x_i - \bar{x})^T$$

$$= S_n$$

$$= \frac{1}{n} X^T H^T H X$$

$$= \frac{1}{n} X^T H \cdot H X$$

$$= \boxed{\frac{1}{n} X^T H X}$$

Properties of $\bar{X}_n, S_n$

$$\mathbb{E}\underset{\sim}{X}_i = \mu$$

$$\mathbb{E}\bar{X}_n = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \mu$$

$$\mathrm{Var}(\bar{X}_n) =$$

$$= \mathrm{Cov}(\bar{X}_n, \bar{X}_n)$$

$$= \mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} \underline{X}_i, \frac{1}{n}\sum_{j=1}^{n} X_j\right)$$

(bilinearity)

$$= \frac{1}{n}\sum_{i=1}^{n} \mathrm{Cov}\left(\underline{X}_i, \frac{1}{n}\sum_{j=1}^{n} X_j\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(\underline{X}_i, \underline{X}_j),$$

equal to $0_{m\times m}$ unless $i=j$ Σ

$$= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Cov}(\underline{X}_i, \underline{X}_i)$$

$$= \frac{1}{n^2}\, n\cdot\Sigma \quad = \frac{1}{n}\Sigma$$

<u>Lemma</u> $\mathbb{E}S_n = \frac{n-1}{n}\cdot\Sigma$

<u>Proof</u>,

$$\mathbb{E}S_n = \mathbb{E}\frac{1}{n}\sum_{i=1}^{n}(\underline{X}_i-\bar{X}_n)(\underline{X}_i-\bar{X}_n)^T$$

$$= \mathbb{E}\frac{1}{n}\sum_{i=1}^{n}(\underbrace{\underline{X}_i-\mu}_{①}+\underbrace{\mu-\bar{X}_n}_{②})(\underbrace{\underline{X}_i-\mu}_{③}+\underbrace{\mu-\bar{X}_n}_{④})^T$$

$$= \mathbb{E}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)(\underline{X}_i-\mu)^T \quad Ⓐ$$

n

$$- \mathbb{E} \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(\bar{x}_n - \mu)^T \quad \text{B}$$

$$- \mathbb{E} \frac{1}{n} \sum_{i=1}^{n} (\bar{x}_n - \mu)(x_i - \mu)^T \quad \text{C}$$

$$+ \mathbb{E} \frac{1}{n} \sum_{i=1}^{n} (\bar{x}_n - \mu)(\bar{x}_n - \mu)^T \quad \text{D}$$

$$\text{A} = \frac{1}{n} \sum_{i=1}^{n} \underbrace{\boxed{\mathbb{E}(\underline{x}_i - \mu)(\underline{x}_i - \mu)^T}}_{} \ \Sigma \ = \Sigma$$

$$\text{D} = \frac{1}{n} \sum_{i=1}^{n} \boxed{\mathbb{E}(\bar{x}_n - \mu)(\bar{x}_n - \mu)^T} \ \frac{1}{n}\Sigma$$

$$= \frac{1}{n} \Sigma$$

$$\text{B} = \mathbb{E} \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(\bar{x}_n - \mu)^T$$

$$\left(\sum_{i=1}^{n} a_i \cdot b = \left(\sum a_i\right) \cdot b \right)$$

$$= \mathbb{E}\left[\boxed{\left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu) \right)} \cdot (\bar{x}_n - \mu)^T \right]$$

$$\| \quad \frac{1}{n} \sum_{i=1}^{n} x_i - \frac{1}{n} \sum_{i=1}^{n} \mu = \bar{x}_n - \mu$$

$$= \mathbb{E}(\bar{X}_n - \mu)(\bar{X}_n - \mu)^T = \frac{1}{n}\Sigma$$

$$A - B - C + D =$$

$$\Sigma - \frac{1}{n}\Sigma - \frac{1}{n}\Sigma + \frac{1}{n}\Sigma$$

$$= \Sigma - \frac{1}{n}\Sigma = \frac{n-1}{n}\Sigma$$

$\square$

MULTIVARIATE NORMAL DIS.

Recall: Univariate normal

$$X \sim N(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad x \in \mathbb{R}$$

$$X_i \sim N(\mu_i, \sigma_i^2) \quad \text{independent}\ i = 1, \ldots, m$$

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

$$f_X(\underline{x}) = \prod_{i=1}^{m} f_i(x_i)$$

$\hookrightarrow$ by independence,

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_m^2 \end{pmatrix}$$

$$= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}$$

$$= \frac{1}{(2\pi)^{m/2}} \frac{1}{\sqrt{\sigma_1^2 \cdots \sigma_m^2}} e^{-\frac{1}{2}\sum_{i=1}^{m} \frac{1}{\sigma_i^2} \cdot (x_i-\mu_i)^2}$$

$$\frac{1}{\cdots} \frac{1}{\cdots} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

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

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

$$= \sum_i \sum_j (\Sigma^{-1})_{ij} (x_i - \mu_i)(x_j - \mu_j)$$

$$\Sigma^{-1} = \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_m^2 \end{pmatrix}$$

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

$\mu \in \mathbb{R}^m$, $\Sigma \in S_+^m$ (PD)

if X has density

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

Z is standard normal

if $Z \sim N_m(0_m, I_m)$

<u>Prop</u> : If $Z \sim N_m(0_m, I_m)$

$$X = \mu + \Sigma^{1/2} \cdot Z \qquad \Sigma \in \mathbb{S}^m_+$$

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

recall $\Sigma \in \mathbb{S}^m$

$$\Sigma = U \Lambda U^T \qquad \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

$$\Sigma^{1/2} = U \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_m} \end{pmatrix} U^T$$

$$\Sigma^{1/2} \cdot \Sigma^{1/2} = U \Lambda^{1/2} \underbrace{U^T \cdot U}_{I_m} \Lambda^{1/2} U^T$$

$$= U \cdot \Lambda^{1/2} \cdot \Lambda^{1/2} \cdot U^T$$

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

<u>Proof</u>; density of Z is

$$\boxed{f_Z(z) = \frac{1}{(2\pi)^{m/2}} e^{-\frac{1}{2} z^T z}}$$

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

$$\boxed{f_X(\underline{x}) = f_Z(G^{-1}(x)) \left|\det \nabla G^{-1}\right|}$$

$$G^{-1}(x) = \Sigma^{-1/2}(x-\mu)$$

Jacobian $\nabla G^{-1} = \Sigma^{-1/2}$

direct check that

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