$\mu = E[X]$
$\sigma^2 = Var[X]$

Let $\vec{X}$ be a vector r.v. of dimension n and define $\vec{\mu} := E[\vec{X}]$

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

vector of constants

notice:

$$E[\vec{X} + c] = \vec{\mu} + \vec{c}$$

$$E[\vec{c}^T \vec{X}] = \vec{c}^T \vec{\mu}$$

instead of $\vec{\sigma}^2$ use variance-covariance matrix

$$\Sigma := Var[\vec{X}]$$

$$\Sigma = E[\vec{X}\vec{X}^T] - E[\vec{X}]E[\vec{X}]^T$$

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

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

← variance covariance matrix

notice

$$Var[\vec{X} + \vec{c}] = Var[\vec{X}] = \Sigma$$

$$Var[\vec{c}^T \vec{X}] = \vec{c}^T \Sigma \vec{c}$$

variance covariance matrix

Let $A \in \mathbb{R}^{m \times n}$ (all constants)

A is $m \times n$ matrix of constants

$$E[A\vec{X}] = E\begin{bmatrix} \vec{a}_1 . X \\ \vec{a}_2 . \vec{X} \\ \\ \vec{a}_n \vec{X} \end{bmatrix} = \begin{bmatrix} E[\vec{a}_1 . \vec{X}] \\ E[\vec{a}_2 \cdot \vec{X}] \\ \\ E[\vec{a}_n \vec{X}] \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \vec{\mu} \\ \vec{a}_2 \vec{\mu} \\ \\ \vec{a}_n \vec{\mu} \end{bmatrix}$$

$$\begin{bmatrix} A \end{bmatrix} \begin{bmatrix} \vec{x} \end{bmatrix} = \begin{bmatrix} \end{bmatrix}$$

$$E[\vec{X} B] = \vec{\mu} B$$

$$= \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_n \end{bmatrix} \vec{\mu} = A\vec{\mu}$$

$\vec{X}\vec{X}^T$ is outer product of $\vec{X}$ vector

$$Var[A\vec{X}] = E[(A\vec{X})(AX)^T] - E[A\vec{X}]\,E[A\vec{X}]^T$$

$$= E[A\vec{X}\vec{x}^TA^T] - A\vec{\mu}\vec{\mu}^TA^T$$

$$= AE[\vec{X}\vec{X}^T]A^T - A\vec{\mu}\vec{\mu}^TA^T$$

$$= A\left(E[\vec{X}\vec{X}^T]A^T - \vec{\mu}\vec{\mu}A^T\right)$$

$$= A\left(E[\vec{X}\vec{X}^T] - \vec{\mu}\vec{\mu}\right)A^T$$

$$Var[A\vec{X}] = A\Sigma A^T$$

$$\vec{Z} \sim \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}\vec{Z}^T\vec{Z}}$$ standard mutinomial normal

$$\vec{X} = A\vec{Z} + \vec{\mu} \sim f_{\vec{X}} = ?$$ want PDF of $\vec{X}$

$$g(\vec{Z}) = A\vec{Z} + \vec{\mu}$$

$$E[\vec{X}] = E[A\vec{Z} + \vec{\mu}]$$

$$= AE[\vec{Z}] + \vec{\mu}$$

$$= A\cdot\vec{O} + \vec{\mu}$$

$$E[\vec{X}] = \vec{\mu}$$

$$Var[\vec{X}] = Var[A\vec{Z} + \vec{\mu}]$$

$$= Var[A\vec{Z}]$$

$$= A\,Var[\vec{Z}]A^T$$

$$= AIA^T$$

$$= AA^T = \Sigma$$

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

Fact #4

$$\Sigma = AA^T$$

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

use rule $(AB)^{-1} = B^{-1}A^{-1}$

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

$$\Sigma^{-1} = (A^{-1})^TA^{-1}$$

$$\vec{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} \sim N_n(\vec{O}, I_n)$$

identity

Linear Algebra Facts

Fact #1

identity

$$I = AA^{-1}$$

$$\det[I] = \det[AA^{-1}]$$

$$1 = \det[A]\det[A^{-1}]$$

$$\Rightarrow \det[A^{-1}] = \frac{1}{\det[A]}$$

Fact #2

$$\Sigma = AA^T$$

$$\det[\Sigma] = \det[AA^T]$$

$$\det[\Sigma] = \det[A]\det[A^T]$$

$\det[A^T] = \det[A]$

$$\det[\Sigma] = \det[A]\det[A]$$

$$\det[\Sigma] = (\det[A])^2$$

$$\Rightarrow \det[A] = \pm\sqrt{\det[\Sigma]}$$

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

Fact #3

$$I = AA^{-1}$$

$$I^T = (AA^{-1})^T$$

$I^T = I$

$$I = (A^{-1})^T A^T \quad \text{and} \quad I = (A^T)^{-1} A^T$$

$$\Rightarrow (A^T)^{-1} = (A^{-1})^T$$

Fact #4

$$\Sigma = AA^T$$

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

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

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

$$f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}\left(A^{-1}(\vec{x}-\vec{\mu})\right) \left|\det[A^{-1}]\right|$$

$$= \ldots =$$

$$= \frac{\left|\det[A^{-1}]\right|}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T (A^{-1})^T A^{-1} (\vec{x}-\vec{\mu})}$$

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

this is pdf of

$$\sim N_n(\vec{\mu}, \Sigma)$$ General Multinomial Normal

using change of variables

$\vec{X} = A\vec{Z} + \vec{\mu}$ Want $f_{\vec{X}}(\vec{x})$ Let $A \in \mathbb{R}^{n \times n}$ matrix of constants
Let $\vec{\mu} \in \mathbb{R}^n$ vector of constants

Use $f_{\vec{Z}}(h(\vec{x}))|J_h|$ where $g(\vec{z}) = a\vec{z} + \vec{\mu}$ are using change of variables
Need $h(\vec{x}) = g^{-1}(\vec{x})$

$$\vec{X} = A\vec{Z} + \vec{\mu}$$
$\Rightarrow \vec{X} - \vec{\mu} = A\vec{Z}$ have $E[\vec{X}] = \mu$
$\Rightarrow A^{-1}(\vec{X} - \vec{\mu}) = A^{-1}(A\vec{Z})$ $Var[\vec{X}] = \Sigma$
$\Rightarrow A^{-1}(\vec{X} - \vec{\mu}) = \vec{Z}$

Let $B = A^{-1}$

$\Rightarrow \vec{Z} = A^{-1}(\vec{X} - \vec{\mu}) = B(\vec{X} - \vec{\mu}) = B\vec{X} - B\vec{\mu} = h(\vec{X})$

so $h(\vec{X}) = B\vec{X} - B\vec{\mu}$ where $B = A^{-1}$

need J_h

$$J_h = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \\ \vdots \\ \frac{\partial h_n}{\partial x_1} & \frac{\partial h_n}{\partial x_2} & \cdots & \frac{\partial h_n}{\partial x_n} \end{bmatrix} = \det\left(\begin{bmatrix} b_{11} & b_{12} & & b_{1n} \\ b_{21} & b_{22} & & b_{2n} \\ \\ b_{n1} & b_{n2} & & b_{nn} \end{bmatrix}\right)$$

$$J_h = \det[B] = \det[A^{-1}] = \frac{1}{\det[A]} = \frac{1}{\pm\sqrt{\det[\Sigma]}}$$

so $|J_h| = |\det[A^{-1}]| = \frac{1}{\det[A]} = \frac{1}{\sqrt{\det[\Sigma]}}$

this happens because $\Sigma = AA^T$

$\Sigma = AA^T$ (so $\det[\Sigma] = \det[AA^T]$)

$\det[\Sigma] = \det[A]\det[A^T]$, $\det[A^T] = \det[A]$

$\det[\Sigma] = \det[A]\det[A] \Rightarrow \det[\Sigma] = (\det[A])^2$

$\det[A] = \pm\sqrt{\det[\Sigma]}$

$$\vec{Z} = A^{-1}(\vec{X}-\vec{\mu}) \overset{\text{Let } B=A^{-1}}{=} B(\vec{X}-\vec{\mu}) = B(\vec{X}-\vec{\mu}) = B\vec{X} - B\vec{x} = h(\vec{x})$$

$$h(\vec{x}) = B\vec{X} - B\vec{\mu} \quad \text{where } B = A^{-1}$$

$$h(\vec{X}) = \begin{bmatrix} \vec{b}_{1\cdot}\vec{X} - \vec{b}_{1\cdot}\vec{\mu} \\ \vec{b}_{2\cdot}\vec{X} - \vec{b}_{2\cdot}\vec{\mu} \\ \\ \vec{b}_{n\cdot}\vec{X} - \vec{b}_{n\cdot}\vec{\mu} \end{bmatrix} = \underbrace{\begin{bmatrix} h_1(\vec{X}) \\ h_2(\vec{X}) \\ \vdots \\ h_n(\vec{X}) \end{bmatrix}}_{\text{this is } h(\vec{X})}$$

Let
$A \in \mathbb{R}^{n\times m}$ constants
$\vec{\mu} \in \mathbb{R}$ constants

notice $h_1(\vec{x}) = \vec{b}_{1\cdot}\vec{X} - \vec{b}_{1\cdot}\vec{\mu}$

so $h_1(\vec{X}) = b_{11}X_1 + b_{12}X_2 + \dots + b_{1n}X_n - b_{11}\mu_1 - b_{12}\mu_2 - \dots b_{1n}\mu_n$

and so on for h_2, h_3, etc.

$h(\vec{X}) = (h_1(\vec{x}), h_2(\vec{X}), \dots, h_n(\vec{X}))$

$$h(\vec{X}) = A^{-1}(\vec{X}-\vec{\mu})$$

$$f_{\vec{Z}}(z) = \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}Z^TZ}$$

know: $(AB)^T = B^TA^T$

$$\vec{X} = A\vec{Z} + \mu = f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}(A^{-1}(\vec{X}-\vec{\mu}))\left|\det[A^{-1}]\right|$$

$$f_{\vec{X}}(\vec{X}) = \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}(A^{-1}(\vec{X}-\vec{\mu}))^T(A^{-1}(\vec{X}-\vec{\mu}))}$$

know: $(AB)^{-1} = B^{-1}A^{-1}$

Used
$\Sigma = AA^T$
$\Sigma^{-1} = (A^{-1})^TA$

$$f_{\vec{x}}(\vec{X}) = \frac{\left|\det[A^{-1}]\right|}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}(\vec{X}-\vec{\mu})^T(A^{-1})^TA^{-1}(\vec{X}-\vec{\mu})}$$

$$f_{\vec{x}}(\vec{X}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} e^{-\frac{1}{2}(\vec{X}-\vec{\mu})^T\Sigma^{-1}(\vec{X}-\vec{\mu})}$$

This defines $N_n(\vec{\mu}, \Sigma)$ general multivariate normal

(General) Multivariate Normal

$X \sim N_n(\vec{\mu}, \Sigma)$ has pdf (JDF) $f_{\vec{x}}(x) = e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$

Sanity Check

① let $\vec{\mu} = \vec{0}$, $\Sigma = I_n$

$$f_{\vec{x}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}\vec{x}^T\vec{x}}$$

this confirms standard multivariate normal

$$Z \sim f_z(\vec{z}) = \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}\vec{z}^T\vec{z}} \sim N_n(\vec{0}, I_n)$$

✓

② try $n=1$ (to get one variable normal)

$\Sigma = [\sigma^2]$ $\quad \vec{\mu} = [\mu]$ $\quad \Rightarrow \Sigma^{-1} = \left[\frac{1}{\sigma^2}\right]$

$$f_{\vec{x}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^1}} e^{-\frac{1}{2}(x-\mu)\frac{1}{\sigma^2}(x-\mu)}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

confirms, at $n=1$, get Normal Dist. in 1 variable

✓

$$\text{Let } U \sim \chi_k^2 \quad \leftarrow \text{chi squared, k deg. of freedom}$$

$$Z_1, Z_2, \ldots, Z_n \overset{i.i.d.}{\sim} N(0,1)$$

$$Z_1, Z_2, \ldots, Z_n \overset{i.i.d}{\sim} N(0,1)$$

$$U = Z_1^2 + Z_2^2 + \ldots + Z_n^2$$

$$E[U] = E[Z_1^2] + E[Z_2]^2 + \ldots + E[Z_k^2]$$

$$\downarrow \; Z^2\text{'s are i.i.d.}$$

$$E[U] = kE[Z^2]$$

Notice $E[Z] = 0$, $Var[Z] = 1$

$$Var[Z] = E[Z^2] - (E[Z])^2$$

$$1 = E[Z^2] - 0^2$$

$$1 = E[Z^2] \quad \text{so } E[Z^2] = 1$$

so $E[U] = kE[Z^2] = K \cdot 1 = K$

$$f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}(h(\vec{x}))\,|J_h|$$

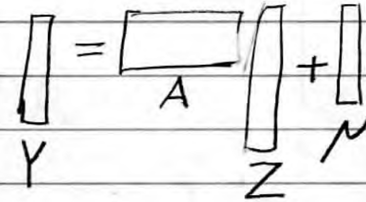
A is constant matrix im $n \times m$

Let $A \in \mathbb{R}^{n \times m}$, $\mu \in \mathbb{R}^m$, $m < n$

A has rank m (Full Rank)

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

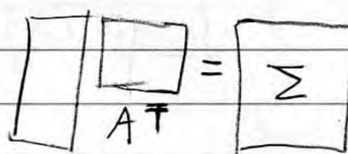
Can't invert A



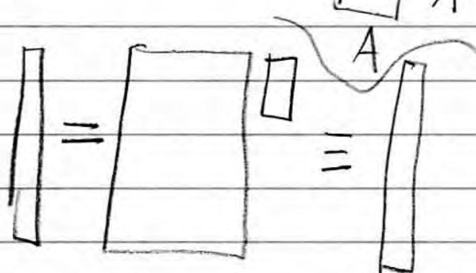
consider $m > n \Rightarrow A$ has at most rank n

$\Rightarrow \Sigma = \underset{m \times n}{A}\,\underset{n \times m}{A^T}$

get $m \times m$



Σ is at most rank $n < m$



$\Rightarrow \Sigma$ is not full rank (not all rows are lin. indep.)

$\Rightarrow \Sigma$ is non-invertible

$\Rightarrow \det[\Sigma] = 0$

$$\tilde{A} = \begin{bmatrix} A \\ \vec{v}_1 \\ \vdots \\ \vec{v}^{n+m} \end{bmatrix} \qquad \vec{\mu} = \begin{bmatrix} \vec{\mu} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

add extra rows ($\vec{v}$'s)

where $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-m}$ are chosen so $\tilde{A}$ is full rank

$\Rightarrow \tilde{A}^{-1}$ exists

$$\tilde{\vec{X}} = \tilde{A}\vec{Z} + \tilde{\vec{\mu}} \sim N_n(\tilde{\vec{\mu}}, \tilde{A}\tilde{A}^T)$$

would want

$$f_{\vec{X}}(\vec{x}) = \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}}_{n-m \text{ of these}} \frac{1}{\sqrt{(2\pi)^n \det[\tilde{A}\tilde{A}^T]}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T(\tilde{A}\tilde{A}^T)(\vec{x}-\vec{\mu})} dx_1 \ldots dx_{m-n}$$

margining out the variables we don't want

we won't do this

Use characteristic functions instead

Let $\phi_{\vec{X}}(\vec{t}) = E[e^{i\vec{t}^T\vec{X}}]$

$$= E\left[e^{i(t_1X_1 + t_2X_2 + \cdots + t_nX_n)}\right]$$

Multivariate Normal Characteristic Function

$X \sim N_n(\vec{\mu}, \Sigma)$ has PDF $f_{\vec{x}}(x) = e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$

acteristic
unction

$$\phi_{\vec{x}}(\vec{t}) = E\left[e^{i\vec{t}^T\vec{X}}\right]$$

$\vec{t}^T\vec{X}$ is dot product $\vec{t}\cdot\vec{X}$

$$= E\left[e^{i(t_1X_1 + t_2X_2 + \dots + t_nX_n)}\right] = E\left[e^{i\sum_{k=1}^{n} t_kX_k}\right]$$

if $X_1, X_2, \dots, X_n$ are indep.

$$\phi_{\vec{x}}(t) = E\left[e^{it_1X_1}\right]E\left[e^{it_2X_2}\right]\dots E\left[e^{it_nX_n}\right] = \prod_{k=1}^{n} \phi_{X_k}(t_k)$$

Verify (P0), (P1), (P2)

(P0) $\phi_{\vec{x}}(\vec{0}) = E\left[e^{i\vec{0}^T\vec{X}}\right] = E\left[e^{i\vec{0}}\right] = 1$ ✓

(P1) $\phi_{\vec{x}}(\vec{t}) = \phi_{\vec{Y}}(\vec{t}) \iff \vec{X} \stackrel{d}{=} \vec{Y}$

$\vec{X}, \vec{Y}$ have same dist.

won't prove this...

(P2) $\vec{Y} = A\vec{X} + \vec{b} \Rightarrow \phi_Y(\vec{t}) =$

$$\Rightarrow \phi_Y(\vec{t}) = E\left[e^{i\vec{t}^T(A\vec{X}+\vec{b})}\right]$$

$$= E\left[e^{i\vec{t}^TA\vec{X} + i\vec{t}^T\vec{b}}\right]$$

$$= E\left[e^{i\vec{t}^TA\vec{x}}\, e^{it\vec{b}}\right]$$

$$= E\left[e^{(A^T\vec{t})^T\vec{x}}\, e^{it\vec{b}}\right]$$

$$= E\left[e^{i(A^T\vec{t})^T\vec{X}}\right] = e^{it\vec{b}}\phi_{\vec{x}}(A^T\vec{t})$$

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

$$\Rightarrow \phi_{\vec{z}}(t) = \prod_{i=1}^{n} \phi_{z_i}(t_i) = \prod_{i=1}^{n} e^{-\frac{t_i^2}{2}}$$

$$= e^{-\frac{1}{2}\sum t_i^2} = e^{-\frac{1}{2}\vec{t}^T\vec{t}}$$

Let $A \in \mathbb{R}^{n\times n}$, $\vec{\mu} \in \mathbb{R}^n$

and invertible

A is $n\times n$ matrix of constants

$$\vec{X} = A\vec{Z} + \mu \sim N_n(\vec{\mu}, \Sigma)$$

$$\phi_{\vec{x}}(t) = e^{i\vec{t}^T\vec{\mu}}\, \phi_{\vec{z}}(A^T\vec{t})$$

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

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

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

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

$$\vec{Y} = B\vec{X} + \vec{c}$$

$$\phi_{\vec{Y}}(\vec{t}) = e^{i\vec{t}^T\vec{c}}\, \phi_{\vec{x}}(\vec{B}^T t)$$

$$= e^{i\vec{t}^T\vec{c}}\left(e^{i(B^T\vec{t})^T\vec{\mu} - \frac{1}{2}(B^T\vec{t})^T\Sigma(B^T\vec{t})}\right)$$

$$= e^{i\vec{t}^T\vec{c}}\left(e^{i\vec{t}^T B\vec{\mu} - \frac{1}{2}\vec{t}^T B\,\Sigma(B^T\vec{t})}\right)$$

$$= e^{i\vec{t}^T(\vec{c} + B\vec{\mu}) - \frac{1}{2}\vec{t}^T B\Sigma B^T\vec{t}}$$

$$\Rightarrow \vec{Y} \sim N_m(\vec{c} + B\vec{\mu}, B\Sigma B^T)$$