$$\Rightarrow \boxed{(n-1)S = X^T H X}$$

computational cost $\downarrow$

H: symmetric idempotent matrix.

$\rightarrow$ centering matrix.

$\boxed{14-08-25}$

$$S = X^T H X \quad , \quad H = I - \frac{1}{n} \vec{1} \cdot \vec{1}^T \quad , \quad \vec{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n \times 1}$$

centering matrix

$H^T = H$ & $H^2 = H$.

if $n \leq p$, then $|S| = 0 \Rightarrow$ The sample indicates that the features have degeneracy or dependency.

proof :- $\Delta = [d_1 \; \cdots \; d_p]_{n \times p}$, $n \leq p$.

$\Rightarrow \text{rank}(\Delta) \leq n.$

$$\sum_{i=1}^{n} d_{ij} = 0 \quad \forall \, 1 \leq j \leq p.$$

$\Rightarrow$ by row transformation, the last row of Δ can be reduced to zero row.

$\Rightarrow \text{rank}(\Delta) \leq n-1 < p.$

Now,

$$S = \frac{1}{n-1} \Delta^T \Delta$$

$\text{rank}(S) = \text{rank}(\Delta^T \Delta) = \text{rank}(\Delta)$ [$\text{rank}(M) = \text{rank}(M^T M)$ can be derived from rank-nullity thm]

$\Rightarrow \text{rank}(S) < p$

Prop.

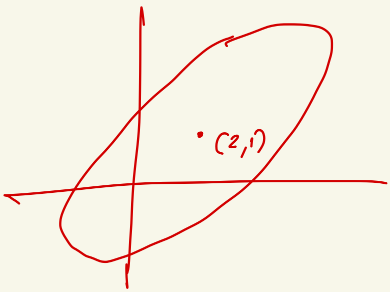
$\Rightarrow |S| = 0.$

A limitation of $|S|$ for variability in data is that it does not capture data orientation.

Let $p=2$ & we have n sample data pts.

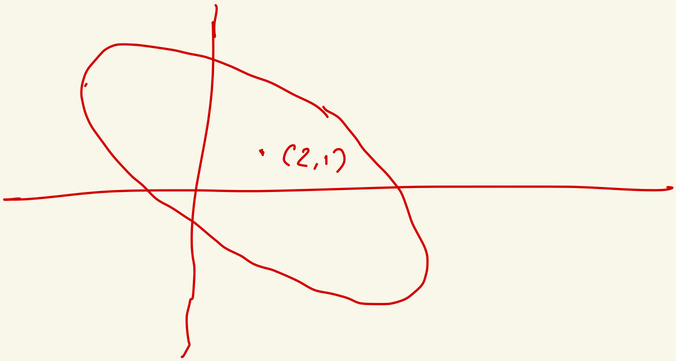
Suppose sample mean, $\bar{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ & let $S = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

$R = \begin{bmatrix} 1 & 4/5 = 0.8 \\ 0.8 & 1 \end{bmatrix}$, $R = [r_{ij}]$, $r_{ij} = \dfrac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}$



I take same features but second sample & calculate sample mean, $\bar{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ & variance-covariance matrix is suppose $\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$

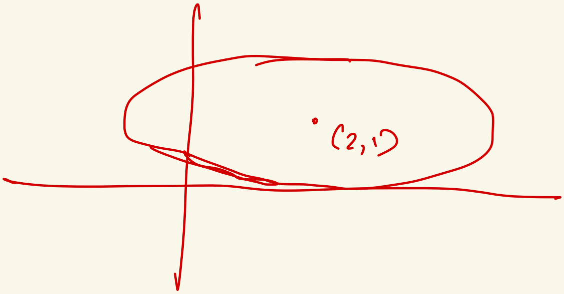
$\therefore R = \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}$



Note: $|S| = 9$.

Take a third sample with say some $\bar{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ &

$S = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, $|S| = 9$.



Define $D^{-1} = \begin{bmatrix} \frac{1}{s_{11}} & & & \\ & \frac{1}{s_{22}} & & \\ & & \ddots & \\ & & & \frac{1}{s_{pp}} \end{bmatrix}_{p \times p}$

Check $\boxed{D^{1/2} \cancel{\times} R D^{1/2} = S}$

$R = [r_{ij}] = \left[\frac{s_{ji}}{\sqrt{s_{ii}} \sqrt{s_{jj}}} \right]$

$D^{1/2} = \text{diagonal} \left[\frac{1}{\sqrt{r_{ii}}} \right]_{p \times p}$

Multivariate Normal Distribution

(Gaussian)

Recall 1-dim normal pdf

$$X \sim N(\mu, \sigma) \quad , \quad p = 1$$

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

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.$$

Say $p = 2$, population

$$X = [X_1, X_2]$$

$$\downarrow$$

$$2\times 2 \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} = \sigma_{12} & \sigma_{22} \end{bmatrix}$$

$$\sigma_{11} = Var(X_1) = Cov(X_1, X_2) = \sigma_1^2$$

ally,
$$\sigma_{22} = \sigma_2^2$$

$$\sigma_{12} = Cov(X_1, X_2)$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}_{2\times 1}$$

Manhalobn's distance

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

- If $p=1$, $\Sigma : 1\times 1$, $\Sigma = [\sigma^2]$, $\sigma^2 = \mathrm{var}(x)$

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

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

$$\therefore (x-\mu)^T \Sigma^{-1} (x-\mu) = \left(\frac{x-\mu}{\sigma}\right)^2$$

- If $p=2$, $d(x,\mu) = d\big((x_1,x_2),(\mu_1,\mu_2)\big)$

$$= (x_1-\mu_1 \quad x_2-\mu_2) \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}$$

$$= [x_1-\mu_1 \quad x_2-\mu_2] \frac{1}{\sigma_1^2\sigma_2^2 - (\sigma_{12})^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}$$

$$= \frac{\sigma_2^2}{\sigma_1^2\sigma_2^2-(\sigma_{12})^2}(x_1-\mu_1)^2 + \frac{\sigma_1^2}{\sigma_1^2\sigma_2^2-(\sigma_{12})^2}(x_2-\mu_2)^2 - \frac{2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1^2\sigma_2^2-(\sigma_{12})^2}$$

$$\frac{\sigma_2^2}{\sigma_1^2\sigma_2^2-(\sigma_{12})^2} = \frac{1}{\sigma_1^2-\frac{(\sigma_{12})^2}{\sigma_2^2}} = \frac{1}{\sigma_1^2\left(1-\frac{(\sigma_{12})^2}{\sigma_1^2\sigma_2^2}\right)} = \frac{1}{\sigma_1^2(1-\rho^2)}.$$

$$\rho = \frac{\mathrm{Cov}(x_1, x_2)}{SD(x_1) * SD(x_2)}$$

$$\frac{2\sigma_{12}}{\sigma_1^2\sigma_2^2-(\sigma_{12})^2} = \frac{2\sigma_{12}}{\sigma_1^2\sigma_2^2\left(1-\frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}\right)} = \frac{2\left(\sigma_{12}/\sigma_1^2\sigma_2^2\right)}{1-\rho^2}$$

$$= \frac{2\rho}{\sigma_{12}(1-\rho^2)}.$$

$$\therefore\ d(\bar{x},\bar{\mu}) = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]$$

The pdf of $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_X(x_1, x_2)\, dx_1\, dx_2 = 1$$

In general p-dim,

$$f_X(\underbrace{x_1, \ldots, x_p}_{\in \mathbb{R}^p}) = \frac{1}{(2\pi)^{p/2}\,|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$P\begin{pmatrix} X_1 = x_1 \text{ and} \\ X_2 = x_2 \text{ and} \\ \vdots \\ X_p = x_p \end{pmatrix}$$

Result :- $\underset{p\times 1}{X} \sim N_p(\mu, \Sigma) \iff \forall \bar{a} \in \mathbb{R}^p,$

$$\underbrace{\vec{a}^T X}_{1\times 1} \sim N(a^T\mu, a^T\Sigma a)$$

proof later

① By taking $\vec{a}$ to be a standard basis vector of $\mathbb{R}^p$, we can see that $X \sim N_p(\mu, \Sigma) \Rightarrow$ each $X_i \sim N(\mu_i, \underset{\sigma_i^2.}{\underset{\|}{\sigma_{ii}}})$

②

Let $A : q\times p$, deterministic real matrix.

$b : q\times 1$ " " vector.

$$\underset{q\times 1}{Y} = \underset{q\times p}{A}\,\underset{p\times 1}{X} + \underset{q\times 1}{b} \Rightarrow Y \sim N_q(A\mu + b, A\Sigma A^T)$$

$$E(Y) = A\,E(X) + b\,E(1)$$

$$= A\mu + b$$

$$\mathrm{Var}(cX) = c^T\Sigma c$$