חלק תיאורטי - אלגברה לינארית

1. יהי $A\in M_{d\times n}$ מטריצה אורתוגונלית, $x\in\mathbb{R}^d$

$$\|Ax\| = \sqrt{(Ax)^TAx} = \sqrt{x^TA^TAx} = \sqrt{x^TI_nx}$$

$$\sqrt{x^Tx} = \|x\|$$

2.

$$A^TA = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

$$A^TA = UDU^T$$

$$\begin{vmatrix} 2-\lambda & 0 & 2 \\ 0 & 2-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{vmatrix} =$$

$$(2-\lambda)\begin{vmatrix} 2-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} + 2\begin{vmatrix} 0 & 2-\lambda \\ 2 & -2 \end{vmatrix} =$$

$$(2-\lambda)\big((2-\lambda)(4-\lambda)-(-2\cdot -2)\big)+2\big(-2(2-\lambda)\big)$$

$$=(2-\lambda)(8-6\lambda+\lambda^2-4)-8+4\lambda=$$

$$(2-\lambda)(4-6\lambda+\lambda^2)-4(2-\lambda)=$$

$$(2-\lambda)(4-6\lambda+\lambda^2-4)=(2-\lambda)(\lambda-6)\lambda$$

ולכן הערכים העצמיים של המטריצה הם 2,6,0

$$\Sigma=\begin{bmatrix}6&0&0\\0&2&0\end{bmatrix}$$

$\Leftarrow$

אלכסונית

$$A=\underbrace{U}_{\text{אורתוגונליות}}\Sigma \underbrace{V^T}\Rightarrow AV=U\Sigma$$

V - מטריצה של וקטורים עמודה עצמיים מנורמלים

Σ היא המטריצה האלכסונית של A^TA.

AA^T – וקטורים עצמיים נכם המנורמלות של U.

$$\begin{bmatrix} 2-\lambda & 0 & 2 \\ 0 & 2-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{bmatrix}$$

נציב $\lambda=2$

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 2 & -2 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\begin{matrix} x-y=0 \\ z=0 \end{matrix} \Rightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

ו"ע המתאים לע"ע 2.

נציב $\lambda=0$

$$\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \to \quad x=-z=-y \Rightarrow \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

ו"ע המתאים ל-0

נציב $\lambda=6$

$$\begin{pmatrix} -4 & 0 & 2 \\ 0 & -4 & -2 \\ 2 & -2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 1 & -1 & -1 \end{pmatrix} \to \left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \to$$

$$\left.\begin{array}{l} x=\frac{z}{2} \\ y=-\frac{z}{2} \end{array}\right\} \Rightarrow \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

ו"ע מתאים לע"ע 6,

נרמל את הוקטורים העצמ״ם:

$$\hat{v}_0 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix} \quad v_1 = \hat{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \quad v_2 = \hat{v}_6 = \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

$$V^T = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{bmatrix}$$

נמצא את u

$$u_i = \frac{1}{\sigma_i} A \cdot v_i$$

$$\sigma_1 = \sqrt{2} \qquad \sigma_2 = \sqrt{6}$$

$$u_1 = \frac{1}{\sqrt{2}} A \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$u_2 = \frac{1}{6}\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow$$

$$\boxed{A = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, & \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, & \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \end{bmatrix}^T}_{V^T}}$$

3.

$$A = V \otimes U = \begin{bmatrix} v_1 u_1 & \cdots & v_n u_1 \\ \vdots & & \vdots \\ v_1 u_m & \cdots & v_n u_m \end{bmatrix} = \begin{bmatrix} | & | & & | \\ v_1 U, & v_2 U & \cdots & v_n U \\ | & | & & | \end{bmatrix}$$

קל לראות שמרחב העמודות של המטריצה

נפרש ע״י U $\quad \operatorname{Im} A = (R(A)) = \operatorname{span}(U)$

$$X = a_1 u_1 + a_2 u_2 + \dots + a_n u_n = \qquad . U$$

$$x = a_1 u$$

$$\sum_{i=1}^{n} \langle x, u_i \rangle u_i =$$

$$\langle x, u_1 \rangle u_1 + \langle x, u_2 \rangle u_2 + \dots + \langle x, u_n \rangle u_n =$$

$$\sum_{i=1}^{n} a_i u_i = X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \quad x_j = \sum_{i=1}^{n} a_i u_{i_j}$$

הקורדינטה ה-j

בוקטור ה-i.

נגדיר $U=\{u_1,\dots u_n\}$, .

$$X=\sum_{i=1}^{n} a_i u_i$$

$$\Rightarrow \langle X,u_i\rangle = \left\langle \sum_{i=1}^{n} a_i u_i, u_i \right\rangle =$$

$$= a_1\langle u_1,u_i\rangle+\dots+a_i\langle u_i,u_i\rangle+\dots+a_n\langle u_n,u_i\rangle$$

$$= a_i \Rightarrow X=\sum_{i=1}^{n}\langle x,u_i\rangle u_i$$

$\Uparrow$

U is orthogonal

$\blacksquare$

חלק אינורכי – וכינפ'

גם נגדיר

$$f(\sigma)=U\cdot \operatorname{diag}(\sigma)U^T x=\begin{pmatrix} f_1(\sigma)\\ f_2(\sigma)\\ \vdots\\ f_n(\sigma)\end{pmatrix}$$

$$U=\begin{bmatrix} | & & | \\ u_1, & \dots & u_n \\ | & & | \end{bmatrix}$$

$$U \cdot \mathrm{diag}(\sigma) U^T x = \begin{bmatrix} | & & | \\ \sigma_1 u_1 & \cdots & \sigma_n u_n \\ | & & | \end{bmatrix} U^T x =$$

$$\sum_{i=1}^{n} (\sigma_i u_i u_i^T) X \Rightarrow$$

$1 \leq j \leq n$

$$\frac{\partial f_j(\sigma_0)}{\partial \sigma_i} = u_i u_i^T x_j$$

$$\Rightarrow J(f(\partial)) = \begin{bmatrix} - \nabla f_1(\partial) - \\ \vdots \\ - \nabla f_n(\partial) - \end{bmatrix} =$$

$$\begin{bmatrix} x_1 u_1 u_1^T & \cdots & x_1 u_n u_n^T \\ \vdots & & \vdots \\ x_n u_1 u_1^T & \cdots & x_n u_n u_n^T \end{bmatrix}$$

$$= \begin{pmatrix} | & & | \\ \langle x, u_1 \rangle u_1 & \cdots & \langle x, u_n \rangle u_n \\ | & & | \end{pmatrix}$$

.6

$$h(\sigma) = \frac{1}{2}\|f(\sigma) - y\|^2 =$$

$$\frac{1}{2}\left(f^2(\sigma) - 2f(\sigma)y^T + \|y\|^2\right)$$

$$\frac{\partial h(\sigma)}{\partial \sigma_i} = \left(f(\sigma) - y\right)^T \cdot \frac{df_i(\sigma)}{d\sigma_i}$$

מסעיף הקודם

$$\Rightarrow \quad \nabla h = \left(f(\sigma) - y\right)^T J\left(f(\sigma)\right)$$

ד.

$$S(x)_j = \frac{e^{x_j}}{\sum_{l=1}^{K} e^{x_l}} \,.$$

נחשב את הגרדיאנט ה-i:

$$\nabla S_i(x) = \begin{pmatrix} \dfrac{\partial \frac{e^{x_i}}{\sum_{l=1}^{K} e^{x_l}}}{\partial x_1} \\ \vdots \\ \dfrac{\partial \frac{e^{x_i}}{\sum_{l=1}^{K} e^{x_l}}}{\partial x_d} \end{pmatrix} = \begin{bmatrix} \dfrac{e^{x_i}\left(\sum_{l=1}^{K} e^{x_l} - e^{x_1}\right)}{\left(\sum_{l=1}^{K} e^{x_l}\right)^2} \\ \vdots \\ \dfrac{e^{x_i}\left(\sum_{l=1}^{K} e^{x_l} - e^{x_d}\right)}{\left(\sum_{l=1}^{K} e^{x_l}\right)^2} \end{bmatrix}$$

נסמן $Z = \sum_{l=1}^{K} e^{x_l}$

$$= \begin{pmatrix} \dfrac{e^{x_i}(Z - e^{x_1})}{Z} \\ \vdots \\ \dfrac{e^{x_i}(Z - e^{x_d})}{Z} \end{pmatrix} = \begin{bmatrix} \dfrac{e^{x_i}}{Z}\left(1 - \dfrac{e^{x_1}}{Z}\right) \\ \vdots \\ \dfrac{e^{x_i}}{Z}\left(1 - \dfrac{e^{x_d}}{Z}\right) \end{bmatrix}$$

$$J(S(x)) = \begin{bmatrix} - \nabla S_1(x) - \\ \vdots \\ - \nabla S_K(x) - \end{bmatrix}$$

$$f(x,y) = x^3 - 5xy - y^5$$ 8

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 5y \\ -5x - 5y^4 \end{pmatrix}$$

$$\nabla(\nabla f(x,y)) = \begin{bmatrix} 6x\ , & -5 \\ -5\ , & -20y^3 \end{bmatrix}$$

$\Downarrow$

Hessian

<u>קמירות</u>

יהיו $x,y\in C$ $\Leftarrow$ לכל $i\in I$

$x,y\in C_i$. C_i קמור ולכן

לכל $0\le\alpha\le 1$ $\alpha x+(1-\alpha)y\in C_i$.

כיון שמתקיים לכל i:

$$\alpha x+(1-\alpha)y\in \bigcap_{i\in I} C_i = C$$

ולכן C קבוצה קמורה.

סוף.

$b\in C_1+C_2 \Rightarrow$

עבור $0\le b,\alpha\le 1$

$$b=c_1+c_2=(\alpha+1-\alpha)c_1+(\alpha+1-\alpha)c_2=$$

$$\alpha(c_1+c_2)+(1-\alpha)(c_1+c_2)$$

$$b = \alpha c_{1a} + (1-\alpha)c_{1b} + \alpha c_{2a} + (1-\alpha)c_{2b} =$$

$$\alpha(c_{1a}+c_{2a}) + (1-\alpha)(c_{1b}+c_{2b})$$

סו. יהיו $x,y \in C_1 + C_2$

ויהי $0 \le \alpha \le 1$

נסמן $x_1, y_1 \in C_1$

$x_2, y_2 \in C_2$

s.t $x = x_1 + x_2$, $y = y_1 + y_2$

$$\alpha x + (1-\alpha)y = \alpha(x_1 + x_2) + (1-\alpha)(y_1 + y_2)$$

$$= \underbrace{\alpha x_1 + (1-\alpha)y_1}_{\in C_1} + \underbrace{\alpha x_2 + (1-\alpha)y_2}_{\in C_2}$$

מהקמירות

$$\Rightarrow \alpha x + (1-\alpha)y \in C_1 + C_2$$

ס11

י ה'

$$x,y \in \lambda C$$

$$\alpha x + (1-\alpha) y = \alpha \lambda c_x + (1-\alpha)\lambda c_y =$$

$$\lambda(\underset{\underset{C}{\cap}}{\alpha c_x + (1-\alpha) c_y}) \in \lambda C$$

Estimation theory

ש12

$\star$ $E(x_i)=\theta$ הוכחנו עקביות שכוללת הממוצעים בינינו מוכח ונכון:

$$E\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} E(x_i) = n\theta$$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} x_i - \theta\right| > \varepsilon\right) = P\left(\left|\sum x_i - \theta n\right| > n\varepsilon\right) \le$$

$$P\left(\left|\sum_{i=1}^{n} x_i - E\left(\sum_{i=1}^{n} x_i\right)\right| \ge n\varepsilon\right) \underset{\star}{\le} \frac{Var(\sum x_i)}{(n\varepsilon)^2}$$

$\star$ אי שוויון צ'בישב.

$$\overset{\text{i.d}}{=} \frac{n \, Var(x_i)}{n^2 \varepsilon^2} = \frac{Var(x_i)}{n \varepsilon^2}$$

$$\Rightarrow$$

$$0 \leq \lim_{n \to \infty} \left(|\hat{\mu}_n - \mu| \right) \leq \lim_{n \to \infty} \frac{Var(x_i)}{n \varepsilon^2} = 0$$

ולכן האומד קונסיסטנטי

19, נסמן $X_i \in \mathbb{R}^d$

$$h((\mu, \Sigma) \mid x_1, \dots, x_m) = \prod_{i=1}^{m} \frac{\exp\left(-\frac{1}{2}[X_i - \mu]^T \Sigma^{-1} [X_i - \mu]\right)}{\sqrt{(2\pi)^d \cdot \det(\Sigma)}}$$

$$= \frac{\exp\left(-\frac{1}{2} \sum_{i=1}^{m} (X_i - \mu)^T \Sigma^{-1} (X_i - \mu)\right)}{\left(\sqrt{(2\pi)^d \det(\Sigma)}\right)^m} \Rightarrow$$

log של מכפלה

$$\log(h(\mu, \Sigma \mid X_1, \dots, X_m)) = \left(-\frac{1}{2} \sum_{i=1}^{m} [X_i - \mu]^T \Sigma^{-1} [X_i - \mu]\right) + \left(-\frac{m}{2} \log\left((2\pi)^d \det(\Sigma)\right)\right)$$

$$= -\frac{1}{2}\left(\sum_{i=1}^{m} [X_i - \mu]^T \Sigma^{-1} [X_i - \mu] + m \left(d \log(2\pi) + \log \det(\Sigma) \right) \right)$$

Parctical Part

1.

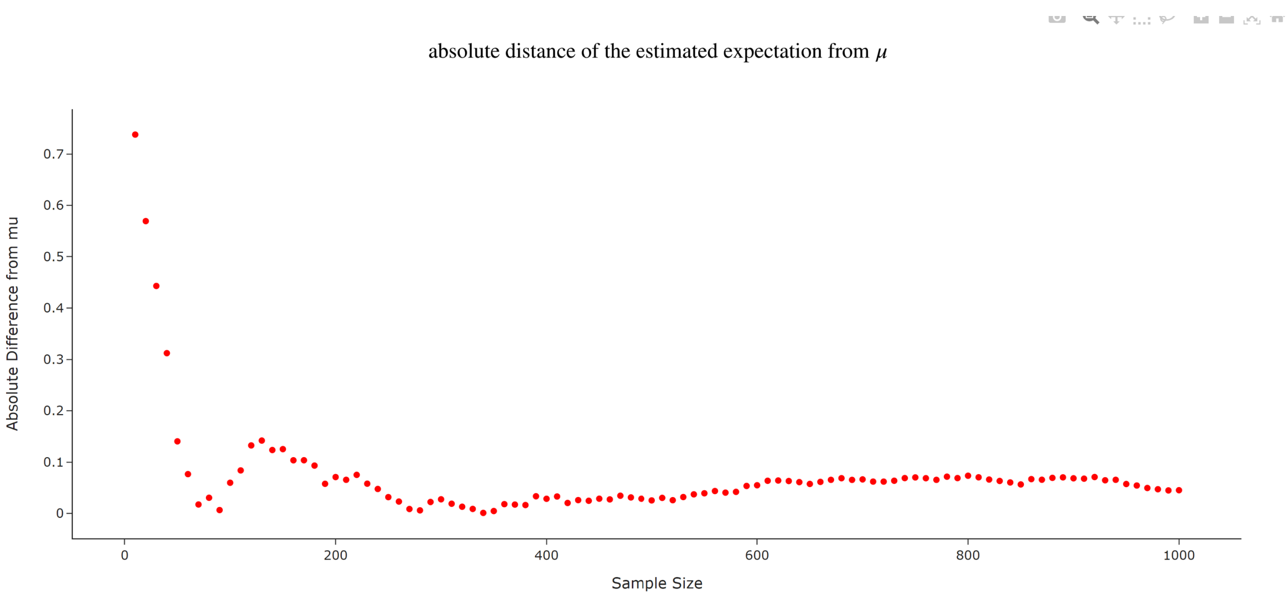
משתנה univariant עבור $X \sim N(10,1)$

$|X| = 1000$

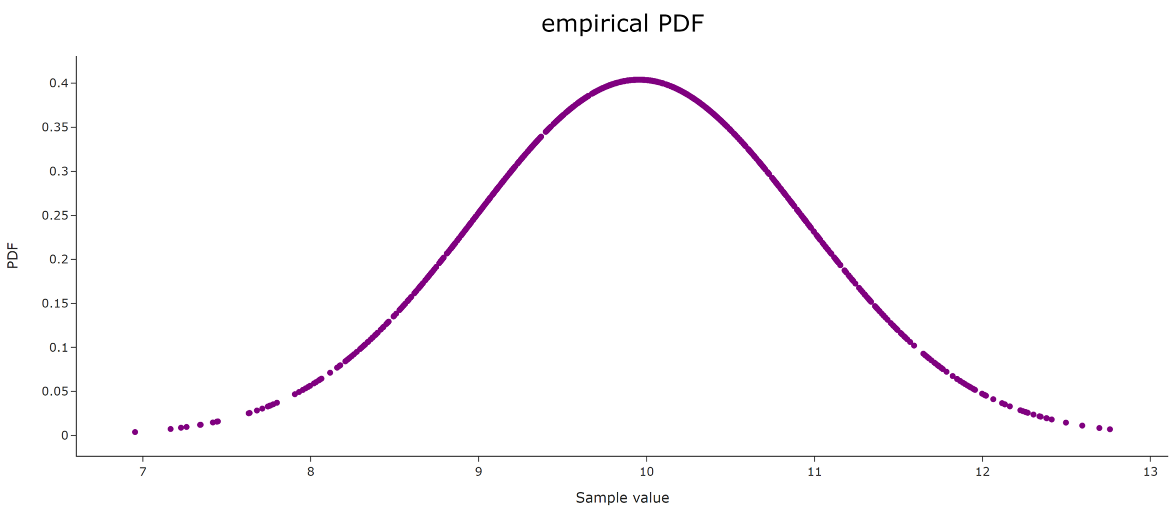
מאומד - E var

```
(9.955, 0.975)
```

2.



3.



עמוד משתנה multivariant

4. עמוד

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0.2 & 0 & 0.5 \\ 0.2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.5 & 0 & 0 & 1 \end{bmatrix}$$

$$X \sim N(\mu, \Sigma), \quad |X| = 1000$$

$$\hat{\mu} =$$

```
[-2.000e-03  5.200e-02  3.993e+00 -6.000e-03]
```

$\hat{\Sigma} =$

```
[[ 0.95   0.168 -0.013  0.468]
 [ 0.168  1.968  0.022  0.059]
 [-0.013  0.022  0.98  -0.026]
 [ 0.468  0.059 -0.026  0.945]]
```

ל כ

ניתן ללמוד שהסבירות המרבית היא:

ט כ

f_1 $\quad$ f_3

```
[-0.05  3.97]
```

וכן שאכן שההתפלגויות נורמליות כמו כן השונות של f_3 גדולה משל f_1.