עמרי תמאס 317704112 חלק מעשי'

1.

A אורתוגונלית $\iff$ $A^TA=I$

$$\|Av\| = \sqrt{\langle Av, Av\rangle} = \sqrt{[Av]^T \cdot Av} = \sqrt{v^T A^T A v} = \sqrt{v^T I v} = \sqrt{v^T v} = \sqrt{\langle v,v\rangle} = \|v\|$$

הגדרה נורמה משרית מ״פ

הגדרה מ״פ

$A^TA=I$ נתון

2.

$$A=\begin{bmatrix}1&1&0\\1&-1&2\end{bmatrix}, \quad A^T=\begin{bmatrix}1&1\\1&-1\\0&2\end{bmatrix}$$

$$AA^T=\begin{bmatrix}1&1&0\\1&-1&2\end{bmatrix}\begin{bmatrix}1&1\\1&-1\\0&2\end{bmatrix}=\begin{bmatrix}2&0\\0&6\end{bmatrix}$$

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

נעזר ב np.linalg.eig ונקבל ש $e_1 e_2$ ו 2, 6 הם ע״ע ו״ע של AA^T.

ו 6, 2, 0 הם ע״ע של A^TA.

נשים לב ש Σ היא

$$\begin{bmatrix}\sqrt{6}&0&0\\0&\sqrt{2}&0\end{bmatrix}$$

נמצא את הו״ע

$$\ker\left(\begin{bmatrix}2&0&2\\0&2&-2\\2&-2&4\end{bmatrix}-\begin{bmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{bmatrix}\right)$$

$\lambda=0$

$$\begin{bmatrix}2&0&2\\0&2&-2\\2&-2&4\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=0$$

$$\begin{cases}a+c=0\\b-c=0\\2a-2b+4c=0\end{cases}$$

$\lambda=2$

$$\begin{bmatrix}0&0&2\\0&0&-2\\2&-2&2\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=0$$

$$\begin{cases}2c=0\\-2c=0\\2a-2b+2c=0\end{cases}$$

$\lambda=6$

$$\begin{bmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=0$$

$$\begin{cases}-4a+2c=0\\-4b-2c=0\\2a-2b-2c=0\end{cases}$$

$c=2a$

$c = 2b$

$$\begin{cases} a = -c \\ b = c \end{cases}$$

$$v_1 = [-1, 1, 1]$$

$$\|v_1\| = 3$$

$$v_2 = [1, 1, 0] \qquad \|v_2\| = 2$$

$$v_3 = [1, -1, 2] \qquad \|v_3\| = 6$$

ונרמל את הו״ע ונקבל מאלכסנת את V

$$\begin{bmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \end{bmatrix}$$

נחשב את וקטורי העמודה של U

$$u_1 = \frac{1}{\sigma_1} \cdot A \cdot v_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{-\sqrt{6}}{6} \\ \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

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

[illegible]

3.

ונשאר אם z_0 כצ"ל של וקטורי הבסיס העצמיים!

(8) $z_0 = \sum_{i=1}^{n} c_i e_i$

<u>נוכיח באינדוקציה</u>

$$b_k = \frac{C_0^k \cdot b_0}{\|C_0^k b_0\|}$$

<u>בסיס:</u> עבור $k=0$ הטענה נכונה מאופן הגדרה b_0,

עבור k הטענה נכונה מהגדרת האיטרציה.

<u>הנחה:</u> הטענה

<u>צעד:</u>

$$b_{k+1} = \frac{C_0 b_k}{\|C_0 b_k\|} = \frac{C_0 \cdot \frac{C_0^k b_0}{\|C_0^k b_0\|}}{\left\| C_0 \cdot \frac{C_0^k b_0}{\|C_0^k b_0\|} \right\|} = \frac{C_0^{k+1} \cdot b_0 \cdot \cancel{\frac{1}{\|C_0^k b_0\|}}}{\|C_0^{k+1} b_0\| \cancel{\frac{1}{\|C_0^k b_0\|}}} = \frac{C_0^{k+1} b_0}{\|C_0^{k+1} b_0\|}$$

כנדרש.

הוצאת סקלר מנורמה לפי תכונות נורמה.

נשים לב ש $C_0 v_i = \lambda_i v_i$ מהגדרת וקטורים עצמיים

$$\Downarrow$$

$$C_0^k \cdot v_i = \lambda$$

<u>כעת נמשיך הלאה:</u>

$$b_k = \frac{C_0^k \cdot b_0}{\|C_0^k \cdot b_0\|} = \frac{(UD^kU^{-1})^k \cdot b_0}{\|(UD^kU^{-1})^k \cdot b_0\|} = \frac{UD^k \sum c_i e_i}{\|UD^k \sum c_i e_i\|} = \frac{U \sum D^k c_i e_i}{\|UD^k \sum c_i e_i\|} =$$

פירוק EVD

לפי (4)

D היא מטריצה לכסינה אז ע״כ לכל $D_{i,j} = 0 \Rightarrow D^k_{i,j} = 0$, $D^k_{i,i} = (D_{i,i})^k = \lambda_i^k e_i$.
$i \neq j$

$$\frac{U \sum \lambda_i^k c_i e_i}{\|U \sum \lambda_i^k c_i e_i\|} = \frac{c_1 \lambda_1^k}{\|c_1 \lambda_1^k\|} \cdot \frac{U \cdot \sum \left(\frac{\lambda_i}{\lambda_1}\right)^k \frac{c_i}{c_1} \cdot e_i}{\|U \sum \left(\frac{\lambda_i}{\lambda_1}\right)^k \frac{c_i}{c_1} \cdot e_i\|}$$

נשים לב ש מקסימלי אז $\frac{\lambda_i}{\lambda} \to 0$ לכל $i \in N$ שונה מ1, כי $\lambda_1 > \lambda_2 \ldots$.

נוכל להשלים אם כך האיברים המתאפסים לאפס

ואז הביטוי שרצינו!

$$\frac{c_1 \lambda_1^k}{\|c_1 \lambda_1\|^k} \frac{v \cdot e_1}{\|v\, v e_1\|} = \frac{c_1^k}{c_1^k} \cdot \frac{\lambda_1^k}{\|\lambda_1\|^k} \cdot \frac{v_1}{\|v_1\|} \begin{cases} \boxed{v_1} & \lambda_1 > 0,\ c_1 > 0 \ \text{or}\ \lambda_1 < 0,\ c_1 < 0 \\ \boxed{-v_1} & \text{else} \end{cases}$$

4.

$f(\sigma)=U\cdot diag(\sigma)\cdot U^T\cdot x$

מטריצה אורת' nxn ~ U

$x\in\mathbb{R}^n$

דפוק

נתון

נשים לב ש $U\cdot diag(\sigma)\cdot U^T$ היא מטריצה לכסינה, U ו-U^T הן מטריצות מלכסנות
כי הן מורכבות מבסיס $(u_1,\dots,u_n)$ של המרחב.

נוכל לסמן כל וקטור ע"י צ"ל של וקטורי הבסיס $(u_1,\dots,u_n)$
בפרט את הוקטור x.

נסמן וקטור הייצוג של x בבסיס U:
$$[x]_U=\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}$$

לכן
$$x=\sum_{i=1}^n \underbrace{u_i}_{\in\mathbb{R}^n}\cdot\underbrace{\alpha_i}_{\in\mathbb{R}}$$

נשים לב שמכיוון ש U היא מטריצה אורת' ו σ_i הם ערכים עצמיים, ניתן לבחור U
כך ש σ_i הוא ע"ע של u_i במטריצה U.

ננסה לפרק את $f(\sigma)$ לפעולות מטריציוניות

$$f(\sigma)=U\cdot diag(\sigma)\cdot U^T\cdot x=U\cdot diag(\sigma)\cdot U^T\cdot\sum_{i=1}^n u_i\cdot\alpha_i=$$

$$\sum_{i=1}^n U\cdot diag(\sigma)\cdot U^T\cdot u_i\cdot\alpha_i$$

U מטריצה אורת' לכן לכל u_i: $u_i\cdot u_i=1$

$$\sum_{i=1}^n U\, diag(\sigma)\, e_i\cdot\alpha_i=$$

$$\sum_{i=1}^n\sum_{j=1}^n \alpha_i\sigma_i u_i^j=$$

$$\begin{bmatrix}\sum_{i=1}^n(\alpha_i\cdot u_i^1\cdot\sigma_i)\\ \vdots\\ \sum_{i=1}^n \alpha_i u_i^n\cdot\sigma_i\end{bmatrix}$$

נחשב את היעקוביאן ע"י גזירה של כל קורדינטה בנפרד?

עבור המקדם הכללי $[J_\sigma f]^i_j$

$$\frac{\partial f_i(\sigma)}{\partial \sigma_j} = \left(\sum_{i=1}^{n} (\sigma_i \cdot u_i^i \cdot \sigma_i)\right)' = \frac{\cancel{\sum_{\substack{i=1 \\ i\neq j}}^{n} \sigma_i \cdot u_i^i \cdot \sigma_i} + \sigma_j \cdot u_j^i \cdot (\sigma_j + h) - \cancel{\sum_{i=1}^{n} \sigma_i \cdot u_j^i \cdot \sigma_i}}{h} =$$

$$\frac{h\, \sigma_j \cdot u_j^i}{h} = \underline{\sigma_j \cdot u_j^i}$$

$$[J_\sigma f]^i_j = \sigma_i \cdot u_j^i$$

ננסה למצוא דרך להציג את σ_i ע"י נתוני השאלה.

נשים לב שמתקיים $U^T x = \sigma$

ומכלל ש U א"נ $U^{-1} = U^T$

לכן $\sigma_j = \sum_{k=1}^{n} U_j^{ik} \cdot x_k$

$$[J_\sigma f]^i_j = u_j^i \cdot \sum_{k=1}^{n} u_j^k \cdot x_k =$$

$$\sum_{k=1}^{n} u_j^k \cdot x_k \cdot u_j^i =$$

$$\underline{\langle u_j, x\rangle \cdot u_j^i}$$

$\mathcal{J}_\circ$

סימון

$\mathbb{R}^d \to \mathbb{R}^d \quad g(\sigma) = f(\sigma) - y$

$\mathbb{R}^d \to \mathbb{R} \quad l(\sigma) = \frac{1}{2}\|\sigma\|^2$

$J_\sigma(g) = J_\sigma(f)$ מכיון ש g לינארית ב f -

$$J_a(l): \quad (J_a(l))_i = \frac{\partial a}{\partial a_i} = \frac{\frac{1}{2}\left(\sum_{k \neq i} a^2 + (a_i + h)^2 - \sum a^2\right)}{\partial a_i} =$$

$$\frac{\frac{1}{2}\left(\cancel{a_i^2} + 2a_i h + h^2 - \cancel{a_i^2}\right)}{\partial a_i} = \frac{1}{2} \cdot 2a_i = \underline{\underline{a_i}}$$

נשים לב:

$$h(\sigma) = l \circ g(\sigma)$$

ומכלל השרשרת -

$$\nabla h(\sigma) = J_\sigma(h) = J_\sigma(l \circ g) =$$

$$J_{g(\sigma)}(l) \cdot J_\sigma(g) =$$

$$(f(\sigma) - y)^T \cdot J_\sigma(f)$$

$\mathcal{G}.$

[illegible] סקלרית

הנגזרת המכנה

$$\frac{\partial f(x_i)}{\partial x_j} = \begin{cases} i \neq j: & \dfrac{0 \cdot \sum_j e^{x_j} - e^{x_i} \cdot e^{x_j}}{(\sum e^{x_i})^2} = -\dfrac{e^{x_i}}{\sum e^{x_i}} \cdot \dfrac{e^{x_j}}{\sum e^{x_i}} = -S(x)_i \cdot S(x)_j \\[2ex] i = j: & \dfrac{e^{x_i} \sum e^{x_j} - e^{x_i} \cdot e^{x_j}}{(\sum e^{x_j})^2} = \dfrac{e^{x_j} \cdot \cancel{\sum e^{x_j}}}{(\sum e^{x_j})^2} - \dfrac{(e^{x_j})^2}{(\sum e^{x_j})^2} = S(x)_i - (S(x)_i)^2 \end{cases}$$

7.

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

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

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

8.

ניעזר בא"ש צ'בישב

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

יהי $\varepsilon>0$, ונחשב את ההסתברות לכך שההסתברות המשוערת לפי הדגימות הנתונות
תהיה רחוקה יותר מ ε מהתוחלת האמיתית

$$P(|\hat{\mu}_n - \theta| > \varepsilon) \leq \frac{Var(\hat{\mu}_n)}{\varepsilon^2} = \frac{\frac{\sigma^2}{n}}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

א"ש צ'בישב

הדגימות שונות מקריות

$$\lim_{n \to \infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

ההסתברות לכך שהתוחלת המשוערת תהיה רחוקה מ ε שואפת
לאפס כאשר גודל המדגם גדל.

a.

ניעזר בהגדרה לנראות למ"מ נורמלי מימד אחד

$$L(\mu, var, x_1 ... x_n) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{\frac{-1}{2\sigma^2}\cdot\sum_{i=1}^{n}(x_i-\mu)^2}$$

מכיוון שהמ"מ ב"ת אז ההסתברות למאורע מינהם תהיה הסתברות מכפלה.

$$L(\mu, \Sigma, x_1 ... x_n) = \prod_{i=1}^{n} f_X(x_i;\mu,V) = \prod (2\pi\sigma^2)^{-\frac{n}{2}} e^{\frac{-1}{2\sigma^2}\cdot\sum_{i=1}^{n}(x_i-\mu)^2} =$$

$$(2\pi\sigma^2)^{-\frac{n^2}{2}}$$

פונקציית הצפיפות למ"מ רב נורמלי היא

$$f_X(\bar{x}) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \cdot e^{-0.5(x-\mu)^T\cdot\Sigma^{-1}(x-\mu)}$$

פונקציית נראות היא ההסתברות המשותפת לתצפיות $x_1 ... x_n$ כאשר נתון Σ, μ.

$$L(\mu, \Sigma, x_1 ... x_n) = P(x_1, x_2 ... x_n \mid \mu, \Sigma) =$$

$$\prod^{n} f(x_i;\mu,\Sigma) =$$

$$\prod^{n} \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \cdot e^{-0.5(x-\mu)^T\cdot\Sigma^{-1}(x-\mu)} =$$

מכיוון שהמ"מ הם מקובצים ממאורעים ב"ת

$$\frac{1}{\sqrt{2\pi^{n^2}|\Sigma|^n}} \cdot e^{-0.5\sum_{i=1}^{n}(x_i-\mu)^T(\Sigma^{-1})(x_i-\mu)}$$

[illegible]

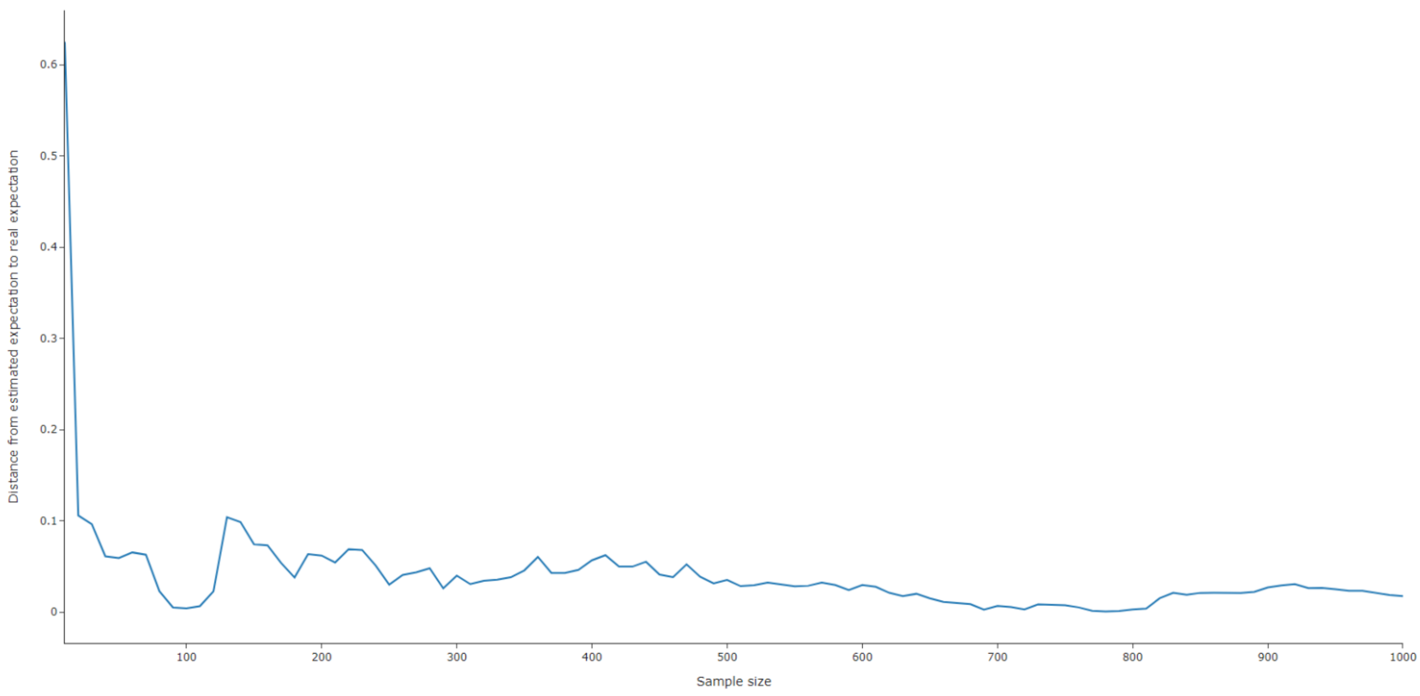
Univariate Gaussian

1.

```
(9.97639160436368,1.0070307407539743)
```

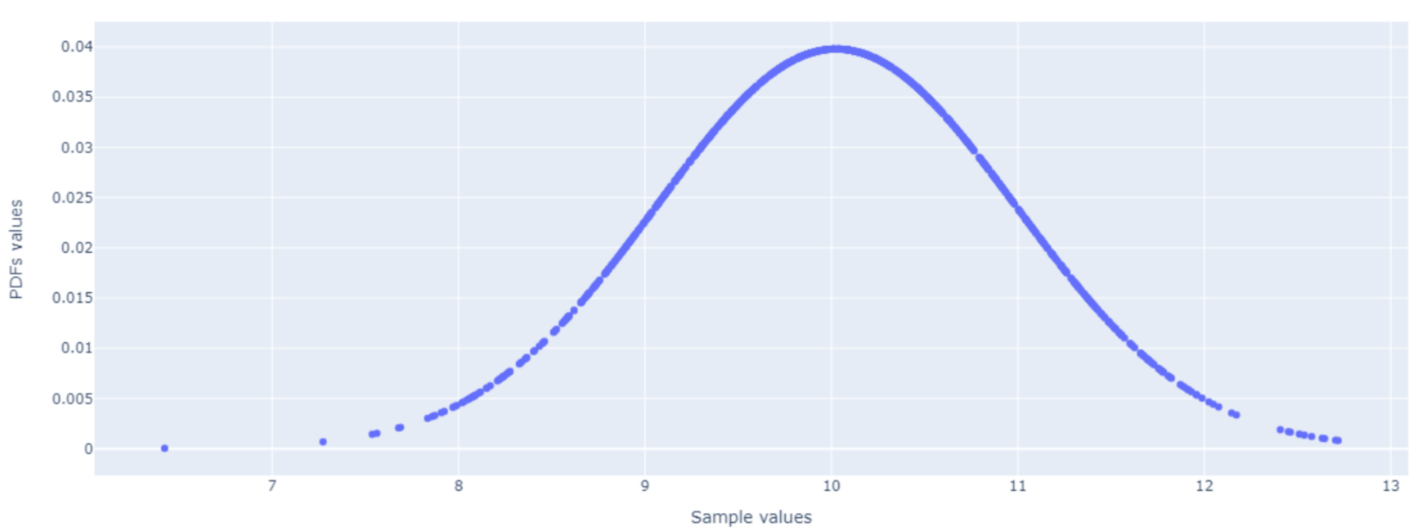
2.

Distance from estimated expectation to real expectation as an output of sample sizefor Univariate Gaussian with expectation 10, variance 1



3.

PDFs values as an output of sample values for Univariate Gaussian with expectation 10, variance 1

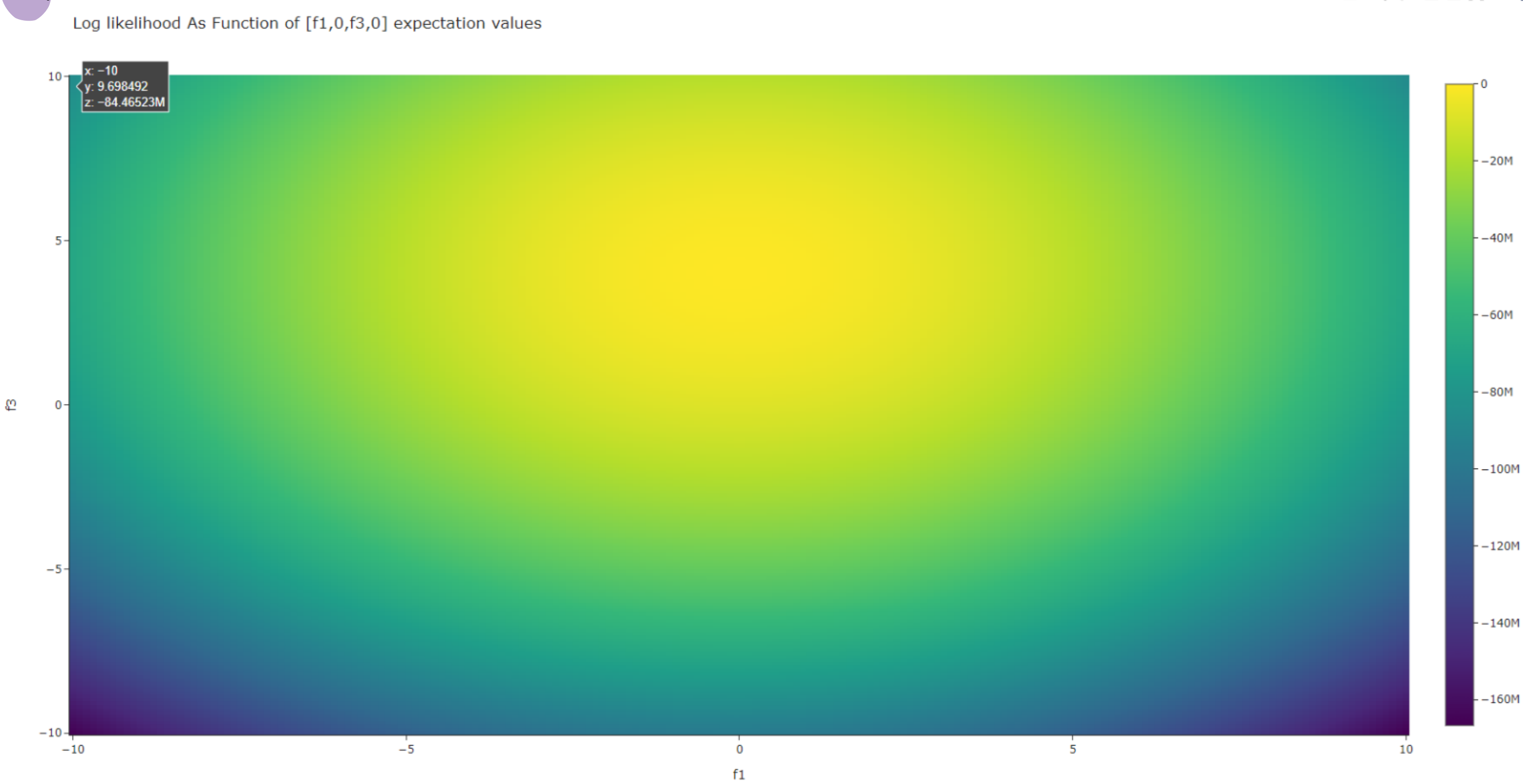


Multivariate Gaussian

4. μ: Σ:

```
[-0.03938902  0.11141852  3.98902775 -0.05205757]
[[ 0.97179968  0.22416488  0.01212463  0.51000642]
 [ 0.22416488  1.94581196 -0.12206255  0.019183  ]
 [ 0.01212463 -0.12206255  1.01683408 -0.0105179 ]
 [ 0.51000642  0.019183   -0.0105179   1.00318671]]
```

5.



6.

```
maximum log-likelihood achieved for f1 -0.050, f3 3.970
```