IML ex1:

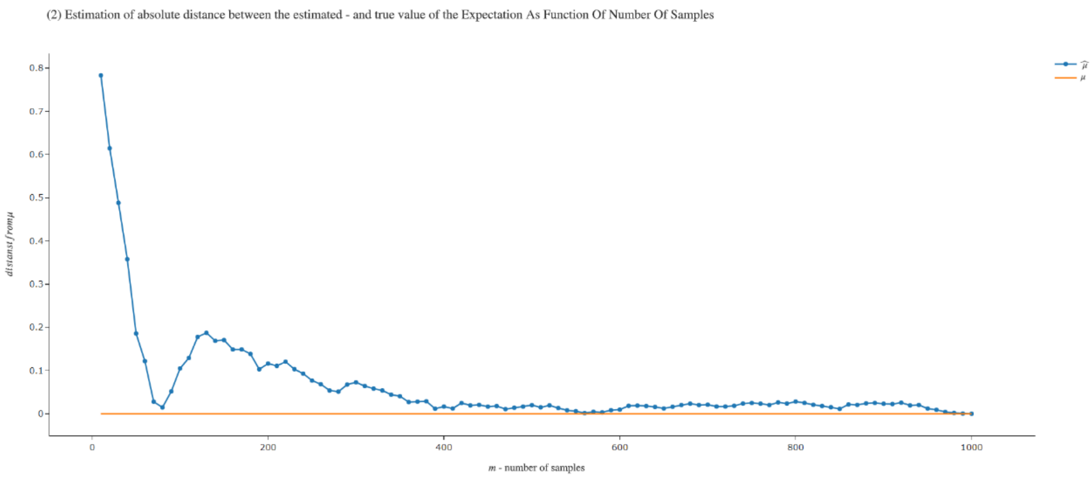
Yair Shtern

ID: 318442241

prints for practical part:

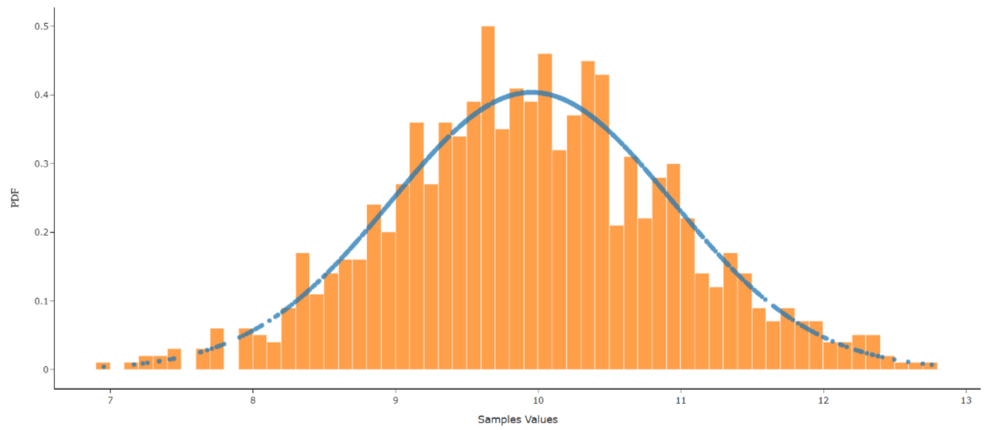
```
Question (1)
(9.954743292509804, 0.9752096659781323)

Question (4)
Expectation:
 [-0.02282878 -0.04313959  3.9932571  -0.02038981]
Covariance:
[[ 0.91667608  0.16634444 -0.03027563  0.46288271]
 [ 0.16634444  1.9741828  -0.00587789  0.04557631]
 [-0.03027563 -0.00587789  0.97960271 -0.02036686]
 [ 0.46288271  0.04557631 -0.02036686  0.9725373 ]]

Question (6)
The values of (f1, f3) that gets the maximum log-likelihood is:
(-0.05, 3.97)
```

Question (2):



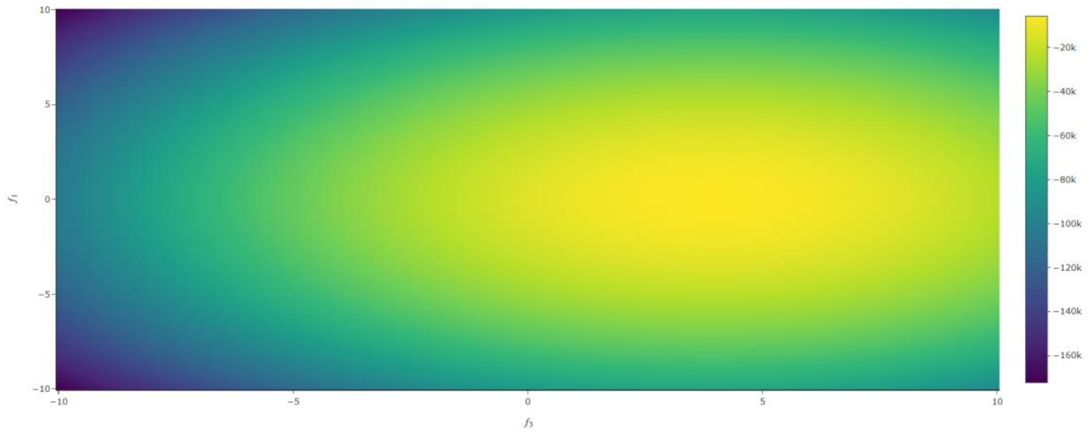
Question (3):

(3) Histograms of PDF samples - It can be seen from the graph that the PDF ~ $\mathcal{N}(10, 1)$ as expected



Question (5):

(5) Heatmap of the log_likelihood - It can be seen from the most likelihood is around f1 = 0 and f3 = 4 which make sense because the expectation is [0,0,4,0]



IML תרגיל 1 - חלק תיאורטי. יאיר שטאן 318442241

א) נסמן $x = a_1 v_1^T + \dots + a_n v_n^T$

נסתכל על השאלה דריבוע:

$$\|x\|^2 = \langle x, x\rangle = \left\langle \sum_{i=1}^{n} a_i v_i^T , \sum_{i=1}^{n} a_i v_i^T \right\rangle =$$

$$= \sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j \langle v_i^T, v_j^T\rangle = \sum_{i=1}^{n} a_i^2 \langle v_i^T, v_i^T\rangle =$$

[עבור $(i \neq j)$ מהאורתוגונליות זה מתאפס]

$$= \sum_{i=1}^{n} a_i^2 \|v_i^T\|^2 = \sum_{i=1}^{n} a_i^2$$

נורמה 1 אורתונורמליות.

$$(Ax)_i = A_i x = v_i x = v_i (a_1 v_1^T \dots a_n v_n^T) =$$

$$= a_i v_i v_i^T = a_i \|v_i\|^2 = a_i$$

$$\|Ax\|^2 = (Ax)^T (Ax) = (a_1, \dots a_n)(a_1 \dots a_n)^T = \sum_{i=1}^{n} a_i^2 \Leftarrow$$

אז קיבלנו:

$$\|Ax\|^2 = \sum_{i=1}^{n} a_i^2 = \|x\|^2$$

בגלל ש מדובר במספרים חיוביים נפ״אור לכם ש

$$\|Ax\| = \|x\|$$

ב) לאחר SVD נקבל $A = U \Sigma V^T$

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T \underbrace{U^T U}_{I} \Sigma V^T = V\Sigma^T\Sigma V^T$$

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

נחשב את הע"ע:

$$\det(A^TA - \lambda I) = \det\begin{bmatrix} 2-\lambda & 0 & 2 \\ 0 & 2-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{bmatrix} =$$

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

$$= (2-\lambda)\left[(2-\lambda)(4-\lambda) - 4\right] + 2\left[-2(2-\lambda)\right] =$$

$$= -\lambda^3 + 8\lambda^2 - 12\lambda$$

$$-\lambda^3 + 8\lambda^2 - 12\lambda = 0$$ נשווה ל- 0

$$\lambda(-\lambda^2 + 8\lambda - 12) = 0 \qquad \lambda_1 = 0$$

$$-\lambda^2 + 8\lambda - 12 = 0 \rightarrow \lambda_2 = 2, \quad \lambda_3 = 6$$

נחשב וקטורים עצמיים:

$$(2I - A^TA) = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 2 & -2 & 2 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \begin{matrix} x_1 - x_2 = 0 \\ x_3 = 0 \end{matrix}$$

לכן: $x_1 = x_2 = t \quad , \quad x_3 = 0$

לכן הוקטור העצמי עבור הע״ע 2 הוא $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

לאחר נרמול נקבל $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

עבור $\lambda = 6$:

$$(6I - A^TA) = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 4 & 2 \\ -2 & 2 & 2 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{matrix} x_1 = \frac{1}{2}x_3 \\ x_2 = -\frac{1}{2}x_3 \end{matrix}$$

סה״כ $x_3 = t \quad v_1 = \frac{1}{2}t \quad x_2 = -\frac{1}{2}t$

הוקטור העצמי שקבל עבור $\lambda = 6$ הוא: $\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

נרמול: $\|v\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1} = \frac{1}{\sqrt{1.5}}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

עבור $\lambda = 0$

$$(-A^TA) = \begin{pmatrix} -2 & 0 & -2 \\ 0 & -2 & 2 \\ -2 & 2 & -4 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\to \begin{matrix} x_1 = -x_3 \\ x_2 = x_3 \end{matrix}$$

נרמ׳: $\frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

נציר לפי הגדרת Σ

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

באלכסון הע"ע שגדולים מאפס (בסדר יורד).

V אורתוגונלית שעמודותיה הוקטורים העצמיים נציר:

$$V = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

כעת נמצא את U:

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma\Sigma^T U^T$$

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

$$\det(AA^T - \lambda I) = \det\begin{pmatrix} 2-\lambda & 0 \\ 0 & 6-\lambda \end{pmatrix} = (2-\lambda)(6-\lambda)$$

$\longleftarrow$ ע"ע $2,6$.

נמצא ו"ע: עבור $\lambda=2$

$$\begin{pmatrix} 0 & 0 \\ 0 & -4 \end{pmatrix} \to \begin{matrix} x_1 = 0 \\ x_2 = t \end{matrix}$$

הו"ע השייך ל- 2 הוא $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (מנורמל)

הו"ע השייך ל- $\lambda=6$: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (מנורמל)

$$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} \to \begin{matrix} x_1 = t \\ x_2 = 0 \end{matrix}$$

מהגדרת U: מורכבת מהוקטורים העצמיים

$$U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

ולכן:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

3) נוכיח באינדוקציה שלכל $k\in\mathbb{N}$ מתקיים
$$b_k=\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}$$

בסיס:

עבור $k=0$:

$$b_1=\frac{C_0b_0}{\|C_0b_0\|}=\frac{C_0\sum_{i=1}^{n}(a_iv_i)}{\|C_0\sum_{i=1}^{n}(a_iv_i)\|}=$$

$$=\frac{\sum_{i=1}^{n}a_i(C_0v_i)}{\|\sum_{i=1}^{n}a_i(C_0v_i)\|}=\frac{\sum_{i=1}^{n}(a_i\lambda_iv_i)}{\|\sum_{i=1}^{n}(a_i\lambda_iv_i)\|}=\frac{\sum_{i=1}^{n}a_i\lambda_iv_i}{\sqrt{\langle\sum_{i=1}^{n}a_i\lambda_iv_i,\sum_{i=1}^{n}a_i\lambda_iv_i\rangle}}$$

v וקטורים עצמיים
הם וקטורים עצמיים

$$=\frac{\sum_{i=1}^{n}a_i\lambda_iv_i}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_i\lambda_ia_j\lambda_j\langle v_i,v_j\rangle}}=\frac{\sum_{i=1}^{n}a_i\lambda_iv_i}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^2}}$$

או"נ לכן $\langle v_i,v_j\rangle=\delta_{ij}$

$\blacksquare$

כעת נניח את נכונות הטענה עבור k ונוכיח עבור $k+1$.

$$b_{k+1}=\frac{C_0b_k}{\|C_0b_k\|}\underset{\text{מההנחה}}{=}\frac{C_0\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}}{\left\|C_0\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}\right\|}=\frac{\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}C_0v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}}{\left\|\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}C_0v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}\right\|}=$$

ו"ע

$$=\frac{\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}\lambda_iv_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}}{\left\|\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k}\lambda_iv_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}\right\|}=\frac{\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k+1}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}}{\left\|\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k+1}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2k}}}\right\|}\underset{\text{לינאריות הנורמה}}{=}\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k+1}v_i)}{\|\sum_{i=1}^{n}(a_i\lambda_i^{k+1}v_i)\|}$$

$$=\frac{\sum_{i=1}^{n}(a_i\lambda_i^{k+1}v_i)}{\sqrt{\sum_{i=1}^{n}a_i^2\lambda_i^{2(k+1)}}}$$ ומאותם טיעונים שטענו על הנחת האינדוקציה נקבל

כעת:

$$b_k = \frac{\sum_{i=1}^{n} a_i \lambda_i^k v_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} = \frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} V_1 + \dots + \frac{a_n \lambda_n^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} V_n$$

בהינתן: $a_1 \neq 0$ נקבל:

$$\lim_{k\to\infty} \frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} \leq \frac{a_1 \lambda_1^k}{\sqrt{a_1^2 \lambda_1^2}} \xrightarrow[k\to\infty]{} |1|$$

מצד שני:

$$\lim_{k\to\infty} \frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} \geq \lim_{k\to\infty} \frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_1^{2k}}} = \lim_{k\to\infty} \frac{a_1 \lambda_1^k}{\lambda_1^n \sqrt{\sum_{i=1}^{n} a_i^2}} \xrightarrow[k\to\infty]{} |1|$$

$$\lim_{k\to\infty} \frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} = 1$$ מסנדוויץ'

כעת לכל אינדקס בין 2 ל- n

$$\lim_{k\to\infty} \frac{a_i \lambda_i^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} \leq \lim_{k\to\infty} \frac{a_i \lambda_i^k}{\sqrt{a_1^2 \lambda_1^{2k}}} = \lim_{k\to\infty} \frac{a_i \lambda_i^k}{a_1 \lambda_1^k} =$$

$$= \frac{a_j}{a_1} \lim \left(\frac{\lambda_j}{\lambda_1}\right)^k = 0$$

$\lambda_1 > \lambda_i$ לכל $1 < i$

כמו כן:

$$\lim_{k\to\infty} \frac{a_i \lambda_0^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} \geq \lim_{k\to\infty} \frac{a_i \lambda_i^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_1^{2k}}}$$

$\lambda_i < \lambda_1$

$$= \lim_{k\to\infty} \frac{a_i \lambda_i^k}{\lambda_1^k \sqrt{\sum_{i=1}^{n} a_i^2}} = \frac{a_i}{\sqrt{\sum_{i=1}^{n} a_i^2}} \lim_{k\to\infty} \frac{\lambda_i^k}{\lambda_1^k} = 0$$

ומסנדוויץ' נקבל שואפים לאפס.

ומאריתמטיקה נקבל:

$$\lim_{k\to\infty} b_k = \lim_{k\to\infty} \left(\frac{a_1 \lambda_1^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} v_1 + \dots + \frac{a_n \lambda_n^k}{\sqrt{\sum_{i=1}^{n} a_i^2 \lambda_i^{2k}}} v_n \right) =$$

$$= v_1 + 0 v_2 + \dots + 0 v_n = v_1$$

$\boxtimes$

4)

$$f\left(\begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{bmatrix}\right) = \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix} \cdot \begin{bmatrix} - & u_1^T & - \\ & \vdots & \\ - & u_n^T & - \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 u_1^T x \\ \vdots \\ \sigma_n u_n^T x \end{bmatrix} = \begin{bmatrix} \vdots & & \vdots \\ u_1 \sigma_1 u_1^T x & \dots & u_n \sigma_n u_n^T x \\ \vdots & & \vdots \end{bmatrix} =$$

$$J_\sigma(f) = \begin{bmatrix} | & & \vdots \\ u_1 u_1^T x & \dots & u_n u_n^T x \\ | & & \vdots \end{bmatrix}$$

בנוסף

$$\frac{d f(\sigma)}{d \sigma_j} = \frac{d}{dx}\left(\sum_{i=1}^{n} \sigma_i u_i u_i^T x \right) = u_j u_j^T x$$

(5

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

נסמן $g(\sigma) = f(\sigma) - y$, $f(\ell) = \frac{1}{2}\|\sigma\|$

מכלל השרשרת:

הגדרת J

$$(\nabla h(\sigma))^T = J_\sigma(h) = J_{g(\sigma)}(f) \cdot J_\sigma(g)$$

כלל השרשרת
וגזירות f,g

כי לכל $x \in \mathbb{R}^n$ ראינו - $\nabla\|x\|^2 = 2x$

ונקבל

$$J_{g(\sigma)}(f) = \frac{1}{2}(\Delta(\|g(\sigma)\|^2))^T = (g(\sigma))^T = (f(\sigma) - y)^T$$

o (וקטור קבוע)

$$J_\sigma(g) = J_\sigma(f(\sigma) - y) = J_\sigma(f\sigma) - J_\sigma(y) = A \quad (*)$$

((*) כאשר A היא המטריצה:

$$\begin{bmatrix} | & & | \\ u_1 u_1^T x & \cdots & u_n u_n^T x \\ | & & | \end{bmatrix}$$
)

מטבלה 4 $J_\sigma f(\sigma) = A$

וסה"כ נקבל

$$\nabla h(\sigma) = (J_{g(\sigma)}(f) J_\sigma(g))^T = ((f(\sigma) - y)^T A)^T =$$

$$= A^T(f(\sigma) - y) = J_\sigma(f)^T(f(\sigma) - y)$$

6) נחשב את M_{ij} לכל תא במטריצת היעקוביאן

עבור $i=j$:

(כלל נגזרת מנה, וגזירה לפי i: $\left\{\begin{matrix} (e^{x_i})^{(x_i)} = e^{x_i} \\ \left(\sum_{l=1}^{k} e^{x_l}\right)^{x_i} = e^{x_i} \\ \text{[השאר 0]} \end{matrix}\right\}$)

$$M_{ii} = \frac{d S(x)_i}{d x_i} = \frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} \quad \left(\left(\frac{f}{g}\right)' = \frac{f'g - g'f}{g^2}\right) = \frac{\sum_{l=1}^{k} e^{x_l} \cdot e^{x_i} - e^{x_i} \cdot e^{x_i}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

$$= \frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} - \left(\frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}}\right)^2 =$$

לכן

$$= S(x)_i - S(x)_i^2 = S(x)_i (1 - S(x)_i)$$

עבור $i \neq j$ נקבל

$\left\{\begin{matrix} (e^{x_i})^{(x_j)} = 0 \\ \left(\sum_{l=1}^{k} e^{x_l}\right)^{(x_j)} = e^{x_j} \\ \text{[השאר 0]} \end{matrix}\right\}_{i \neq j}$

$$M_{ij} = \frac{d S(x)_i}{d x_j} = \frac{e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} = \frac{\left(\sum_{k \neq j} e^{x_k} + e^{x_j}\right)^{x_j}}{\sum_{l=1}^{k} e^{x_l}} =$$

$$\frac{\sum_{l=1}^{k} e^{x_l} \cdot 0 - e^{x_i} \cdot e^{x_j}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{-e^{x_i}}{\sum_{l=1}^{k} e^{x_l}} \cdot \frac{e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} =$$

$$= -S(x)_i S(x)_j$$

סה"כ:

$$J_x(S) = \begin{pmatrix} S(x)_1(1 - S(x)_1) & \cdots & -S(x)_1 S(x)_k \\ \vdots & \ddots & \vdots \\ -S(x)_k S(x)_1 & \cdots & S(x)_k(1 - S(x)_k) \end{pmatrix}$$

(7

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

$$\frac{df(x,y)}{dx} = 3x^2 - 5y \qquad \frac{df(x,y)}{dy} = -5y^4 - 5x$$

$$\frac{df(x,y)}{d^2x} = 6x \qquad \frac{df(x,y)}{d^2y} = -20y^3$$

$$\frac{d^2 f(x,y)}{dy\,dx} = -5$$

סה"כ:

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

(8 זה רעיון החוק החלש של המספרים הגדולים, נוכיח אותו:

(רוצים) לכל $\varepsilon > 0$

$$\lim_{n\to\infty} \left(P\left(|\hat{\mu}_n - \mu| > \varepsilon\right)\right) = 0$$

יהי $\varepsilon > 0$, $X_1,\dots,X_n$ ב"ת לכן מתקיים לכל i: $E(X_i) = \mu$, $var(X_i) = \sigma^2$

(בת"ל ושווי התפלגות).

$$E(\hat{\mu}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{\text{לינאריות ושווי התפלגות}}{=} \frac{1}{n}\left(n \cdot E(X_i)\right) = \mu$$

כעת:

לפי הסבר [אי שוויון צ'בישב]

$$0 \le P\left(|\hat{\mu}_n - \mu| > \varepsilon\right) = P\left(|\hat{\mu}_n - E[\hat{\mu}_n]| > \varepsilon\right) \underset{\text{צ'בישב}}{\le} \frac{Var(\hat{\mu}_n)}{\varepsilon^2} =$$

$$= \frac{Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)}{\varepsilon^2} \underset{x_i \text{ בת"ל ולינאריות}}{=} \frac{\frac{1}{n^2}\sum_{i=1}^{n} var(x_i)}{\varepsilon^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \left(\frac{\sigma}{\varepsilon}\right)^2 \cdot \frac{1}{n} \xrightarrow{n\to\infty} 0$$

$\blacksquare$

9)

$$\log\text{-}L(\theta, x_1, \dots, x_m) = \log f_\theta(x_1, \dots, x_m)$$

$$= \log \prod_{i=1}^{m} f_\theta(x_i) = \sum_{i=1}^{m} \log(f(x_i))$$

וקטור שורה x_i

נקבל:

$$\sum_{i=1}^{m} \log\left(\frac{1}{\sqrt{2\pi^d\,|\Sigma|}} \cdot \exp\left(-\frac{1}{2}(x_i-\mu)^T \Sigma^{-1} (x_i-\mu)\right)\right) =$$

נשתמש עם חוקי לוגריתמים

כפל לוג ← חיבור

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

$$= -\frac{1}{2}\left(md\log(2\pi) + m\log(|\Sigma|) + \sum_{i=1}^{m}\left((x_i-\mu)^T \Sigma^{-1}(x_i-\mu)\right)\right)$$