תרגיל 1

שם: תמר שבי
ת.ז: 315315523

2.1.1 (1) תהי העתקה $T: V \to W$, נתון כי A מטריצה אורתוגונלית.
צ"ל $\forall x \in V \quad \|Ax\| = \|x\|$

יהי $x \in V$
לפי ש A מטריצה אורתוגונלית מתקיים כי לכל $u, v \in V$

$$\langle Av, Au \rangle = \langle v, u \rangle$$

כלומר מתקיים

$$\|Ax\| = \sqrt{\langle Ax, Ax \rangle} = \sqrt{\langle x, x \rangle} = \|x\|$$

2.1.1 (2) נבנה פירוק U, Σ, V^t

נמצא ע"ע של A^TA

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

$$|xI - A^TA| = \begin{vmatrix} x-2 & 0 & -2 \\ 0 & x-2 & 2 \\ -2 & 2 & x-4 \end{vmatrix} = (x-2)\begin{vmatrix} x-2 & 2 \\ 2 & x-4 \end{vmatrix} - 2\begin{vmatrix} 0 & -2 \\ x-2 & 2 \end{vmatrix}$$

$$= (x-2)\Big((x-2)(x-4) - 4\Big) - 2\Big((0\cdot 2) - (x-2)\cdot(-2)\Big)$$

$$= (x-2)^2(x-4) - 4(x-2) - 2\big(2(x-2)\big)$$

$$= (x^2 - 4x + 4)(x-4) - 4x + 8 - 4(x-2)$$

$$= x^3 - 4x^2 - 4x^2 + 16x + \cancel{4x} - \cancel{16} - 4x + \cancel{8} - \cancel{4x} + \cancel{8}$$

$$= x^3 - 8x^2 + 12x$$

$$\Rightarrow x(x^2 - 8x + 12) = 0 \Rightarrow x(x-2)(x-6) = 0$$

קיבלנו ש
$\lambda_1 = 0$
$\lambda_2 = 2$
$\lambda_3 = 6$

נמצא ו"ע. עבור $\lambda_1 = 0$

$$\begin{bmatrix} -2 & 0 & -2 \\ 0 & -2 & 2 \\ -2 & 2 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{matrix} x_1 = -1 \\ x_2 = 1 \\ x_3 = 1 \end{matrix} \Rightarrow v_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

עבור $\lambda_2 = 2$

$$\begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 2 \\ -2 & 2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{matrix} x_1 - x_2 + x_3 = 0 \\ x_2 = t \\ x_3 = 0 \end{matrix} = \begin{matrix} x_1 = 1 \\ x_2 = 1 \\ x_3 = 0 \end{matrix} \Rightarrow V_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

עבור $\lambda_3 = 6$

$$\begin{bmatrix} 4 & 0 & -2 \\ 0 & 4 & 2 \\ -2 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & 2 & 0 \\ 0 & 4 & 2 \\ -2 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & 2 & 0 \\ 0 & 4 & 2 \\ 0 & 4 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \to \begin{matrix} x_1 + x_2 = 0 \\ x_2 + \frac{x_3}{2} = 0 \\ x_3 = 2 \end{matrix} \Rightarrow \begin{matrix} x_1 = 1 \\ x_2 = -1 \\ x_3 = 2 \end{matrix} \Rightarrow V_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

המשך:

כעת נבצע גרם שמידט על מנת לנרמל הוקטורים

$$V_3' = \frac{V_3}{\|V_3\|} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \cdot \frac{1}{\sqrt{6}} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} \qquad V_2' = \frac{V_2}{\|V_2\|} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \qquad V_1' = \frac{V_1}{\|V_1\|} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{3}} = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$\sqrt{1+1+4} = \sqrt{6} \qquad \sqrt{2}$$

קיבלנו כי:

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

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

נמצא את $U \in M_{2\times2}$

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

$$u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{1}{2} \\ \frac{1}{2} - \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

קיבלנו כי:

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

2.1.1) 3)

יהי $v\in\mathbb{R}^n$, $u\in\mathbb{R}^m$ $v=\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}$, $u=\begin{pmatrix}u_1\\ \vdots\\ u_m\end{pmatrix}$

מהגדרה של $\otimes$ מתקיים כי:

$$A=v\otimes u=\begin{bmatrix}v_1u_1 & \cdots & v_1u_m\\ \vdots & & \vdots\\ v_nu_1 & \cdots & v_nu_m\end{bmatrix}$$

נסתכל על עמודות המטריצה A.

$$\begin{bmatrix}v_1u_1\\ \vdots\\ v_nu_1\end{bmatrix},\begin{bmatrix}v_1u_2\\ \vdots\\ v_nu_2\end{bmatrix},\dots,\begin{bmatrix}v_1u_m\\ \vdots\\ v_nu_m\end{bmatrix}$$

נשים לב כי עבור כל וקטור עמודה i ב A ניתן לרשום אותו: u_i כמו סקלר כפול וקטור v
העמודות של A הן $u_1v, u_2v, \dots, u_mv$ כלומר לכל $i\in[m]$

$$colA_i\in span(v)\Rightarrow dim(span(v))=1$$

משום שהממד של span(v) הוא 1, כל קבוצה של וקטורים שגדולה מ 1 (גודל הבסיס) תהיה תלויה לינארית
כלומר מהגדרה של דרגת מטריצה, הדרגה של A היא 1.

2.1.1) 4) יהי בסיס א"נ $(u_1\dots u_n)$ יהי $x\in\mathbb{R}^n$ וידוע $\alpha_1\dots\alpha_n\in\mathbb{R}$ כך ש $x=\sum\alpha_iu_i$

כעת לכל $j\in[n]$

$$\langle x,u_j\rangle=\langle\sum_{i=1}^{n}\alpha_iu_i,u_j\rangle\underset{\text{תכונות מ"פ}}{=}\sum_{i=1}^{n}\alpha_i\langle u_i,u_j\rangle=\alpha_j$$

בסיס א"נ לכל $i\neq j$ $\langle u_i,u_j\rangle=0$ ו- $\langle u_j,u_j\rangle=1$

כנדרש $\boxed{N}$

(2.1.2

5) יהי $X\in\mathbb{R}^n$ ו-$U\in M_{n\times n}$ מטריצה אורתוגונלית

צריך לחשב את היעקוביאן של $f(\sigma)=U\,diag(\sigma)\,U^T x$

$$f(\sigma)=\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}\begin{bmatrix} — & u_1 & — \\ & \vdots & \\ — & u_n & — \end{bmatrix}\cdot x$$

$$=\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}\begin{bmatrix} — & u_1\sigma_1 & — \\ & \vdots & \\ — & u_n\sigma_n & — \end{bmatrix}\cdot x$$

$$=\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}\begin{bmatrix} \sigma_1 u_1^t x \\ \\ \sigma_n u_n^t x \end{bmatrix}=\sum_{i=1}^{n}\sigma_i(u_i^t x)u_i$$

כעת ישירות מהגדרת היעקוביאן

$$[J_\sigma(f)]_{ij}=\frac{\partial f_i(\sigma)}{\partial\sigma_j}=[(u_j^t x)u_j]_i$$

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

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

נגדיר הפרדה לצורך $g:\mathbb{R}^n\to\mathbb{R}^n$ $\quad g(\sigma)=f(\sigma)-y$

$p:\mathbb{R}^n\to\mathbb{R}$ $\quad p(x)=\frac{1}{2}\|x\|^2$

כעת מתקיים כי $h(\sigma)=p\circ g(\sigma)$ ולכן הנגזרת

$$h'(\sigma)=p'(g(\sigma))\cdot g'(\sigma)$$

נשים לב קודם לכל ש y קבוע לכן $g'(\sigma)=f'(\sigma)=J(f)$ כלומר $[g'(\sigma)]_{ij}=[J_\sigma(f)]_{ij}=[u_j^t x\,u_j]_i$

כעת עבור $p'(g(\sigma))$ לפי הגדרה מתקיים כי $p'(g(\sigma))=(2\,g(\sigma))^t$

קיבלנו כי

$$\nabla h(\sigma)=((2\,g(\sigma))^t\cdot J_\sigma(f))^t$$

$$=2\,J_\sigma(f)^t\cdot(f(\sigma)-y)$$

$$=2\cdot(U\cdot diag(U^t x))^t\cdot(U\,diag(\sigma)\,U^T x-y)$$

$$=2\,diag(U^t x)\cdot\underbrace{U^t\cdot U}_{I_n}diag(\sigma)\,U^t x-2\,diag(U^t x)\,U^t y$$

$$=2\,diag(U^t x\,\sigma)\,U^t x-2\,diag(U^t x)\,U^t y$$

2.1.2)

7)

$$S(x)_j = e^{x_j} \cdot \frac{1}{\sum_{l=1}^{k} e^{x_l}}$$

הראינו ש $J_x(S)$ הוא $k \times d$

$i \neq j$
$1 \leq i \leq k$
$1 \leq j \leq d$

$$\left[J_x(S)\right]_{ij} = \frac{\partial S_i(x)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{e^{x_i}}{\sum_{l=1}^{k} e^{x_l}}\right) =$$

$$= \frac{\partial S_i(x)}{\partial x_j} = \frac{f'(x_j)\,g(x_j) - g'(x_j)\,f(x_j)}{g(x_j)^2} = \frac{0 \cdot g(x_j) - e^{x_j} e^{x_i}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{-e^{x_j} e^{x_i}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

נגדיר $f(x_j) = e^{x_i}$, $g(x_j) = \sum_{l=1}^{k} e^{l}$

כעת נסתכל $i = j$

$$\left[J_x(S)\right]_{ii} = \frac{f'(x_i)\,g(x_i) - g'(x_i)f(x_i)}{g(x_i)^2} = \frac{e^{x_i}\sum_{l=1}^{k} e^{x_l} - e^{x_i} \cdot e^{x_i}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} = \frac{\sum_{\substack{l=1 \\ l \neq i}}^{k} e^{x_i + x_l}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

8)

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

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

$$\frac{\partial^2 f(x)}{\partial^2 x} = (3x^2 - 5y)' = 6x$$

$$\frac{\partial^2 f(x)}{\partial x \partial y} = \frac{\partial}{\partial y}\left(3x^2 - 5y\right) = -5$$

$$\frac{\partial^2 f(x)}{\partial y \partial x} = \frac{\partial}{\partial x}\left(-5x - 5y^4\right) = -5$$

$$\frac{\partial^2 f(x)}{\partial^2 y} = \frac{\partial}{\partial y}\left(-5x - 5y^4\right) = -20y^3$$

2.1.3)

9) יהיו $u,v\in C$ אז צ"ל שלכל $\alpha\in[0,1]$ מתקיים כי $\alpha v+(1-\alpha)u\in C$

תהי $\alpha\in[0,1]$, ונראה שלכל $i\in[l]$ אם $u,v\in C_i$ אז $u,v\in\bigcap_{i\in l}C_i$ ומאחר ש-C_i קמורה

$$\alpha v+(\alpha-1)u\in C_i$$

לכל C_i אז

$$\alpha v+(\alpha-1)u\in\bigcap_{i=1}^{l}C_i=C$$

כנדרש ☒

10) יהיו $x,y\in C_1+C_2$ אזי $\alpha\in[0,1]$ צ"ל $\alpha x+(1-\alpha)y\in C_1+C_2$

מאחר ש $x\in C_1+C_2$ קיימים $u_1\in C_1,\ u_2\in C_2$ כך ש $x=u_1+u_2$

" " $y\in C_1+C_2$ " $v_1\in C_1,\ v_2\in C_2$ " " $y=v_1+v_2$

כעת

$$\alpha x+(1-\alpha)y=\alpha(u_1+u_2)+(1-\alpha)(v_1+v_2)$$

$$=\underbrace{\alpha u_1+(1-\alpha)v_1}_{\in C_1\text{ כי היא קמורה}}+\underbrace{\alpha u_2+(1-\alpha)v_2}_{\in C_2\text{ כי היא קמורה}}\in C_1+C_2$$

כנדרש ☒

11) יהיו $u,v\in\lambda C$ ו-$\alpha\in[0,1]$ צ"ל $\alpha u+(\alpha-1)v\in\lambda C$

מאחר ש $u\in\lambda C$ קיים $u'\in C$ כך ש $\lambda u'=u$

" " $v\in\lambda C$ " $v'\in C$ " " $\lambda v'=v$

כעת

$$\alpha u+(1-\alpha)v=\lambda\alpha u'+(1-\alpha)\lambda v'=\lambda(\underbrace{\alpha u'+(1-\alpha)v'}_{\in C\text{ כי C קמורה}})\implies\in\lambda C$$

כנדרש ☒

2.2)

1.2) יהי $X_1,\dots \overset{\text{בת"ל}}{\sim} P$ יהי $\varepsilon>0$ ויהי $n\in\mathbb{N}$ נגדיר $X\sim P$ כך שתוחלת μ

כאשר התוחלת האמפירית, נראה כי

$$\mathbb{P}(|\mu-\hat{\mu}_n|\geq\varepsilon)\xrightarrow[n\to\infty]{}0$$

נשים ♡ ש sample mean הוא unbiased מתקיים כי $\mathbb{E}[\hat{\mu}_n]=\mu$ ולכן נוכל להשתמש בצ'בישב

$$Var(\hat{\mu}_n)=Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\overset{\text{בת"ל}}{=}\frac{1}{n^2}\sum_{i=1}^{n}Var(X_i)\overset{\text{שווי התפלגות}}{=}\frac{n\cdot\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

כעת נציב בצ'בישב ומתקיים

$$\mathbb{P}(|\mu-\hat{\mu}_n|\geq\varepsilon)\leq\frac{Var(\hat{\mu}_n)}{\varepsilon^2}=\frac{\sigma^2}{\varepsilon^2 n}\xrightarrow[n\to\infty]{}0$$

סיימנו מה נדרש קיבלנו ש האסטימטור קונסיסטנטי.

1.3) $X_1\dots X_m\overset{\text{בת"ל}}{\sim}N(\mu,\Sigma)$ נסמן $\theta=\{\mu,\Sigma\}$ כעת מהגדרה

$$\mathcal{L}(\theta|X_1\dots X_m)=f_\theta(X_1\dots X_m)$$

$$\overset{\text{בת"ל}}{=}\prod_{i=1}^{m}f_\theta(X_i)$$

$$=\prod_{i=1}^{m}\frac{1}{\sqrt{(2\pi)^d|\Sigma|}}\cdot\exp\left(-\frac{1}{2}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

$$=\prod_{i=1}^{m}\frac{1}{(2\pi)^{\frac{d}{2}}\cdot|\Sigma|^{\frac{1}{2}}}\cdot\exp\left(-\frac{1}{2}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

$$=\frac{1}{(2\pi)^{\frac{dm}{2}}\cdot|\Sigma|^{\frac{m}{2}}}\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

$$=\frac{1}{((2\pi)^d\cdot|\Sigma|)^{\frac{m}{2}}}\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

כעת ה log likelihood יהיה:

$$\ln\left(\frac{1}{((2\pi)^d\cdot|\Sigma|)^{\frac{m}{2}}}\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)\right)=-\frac{m}{2}\ln\left((2\pi)^d\cdot|\Sigma|\right)-\frac{1}{2}\left(\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

$$=-\frac{md}{2}\ln(2\pi)-\frac{m}{2}\left(\ln|\Sigma|\right)-\frac{1}{2}\left(\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\right)$$

3.1

1) $(9.955, 0.975)$

2) בסוף ה pdf

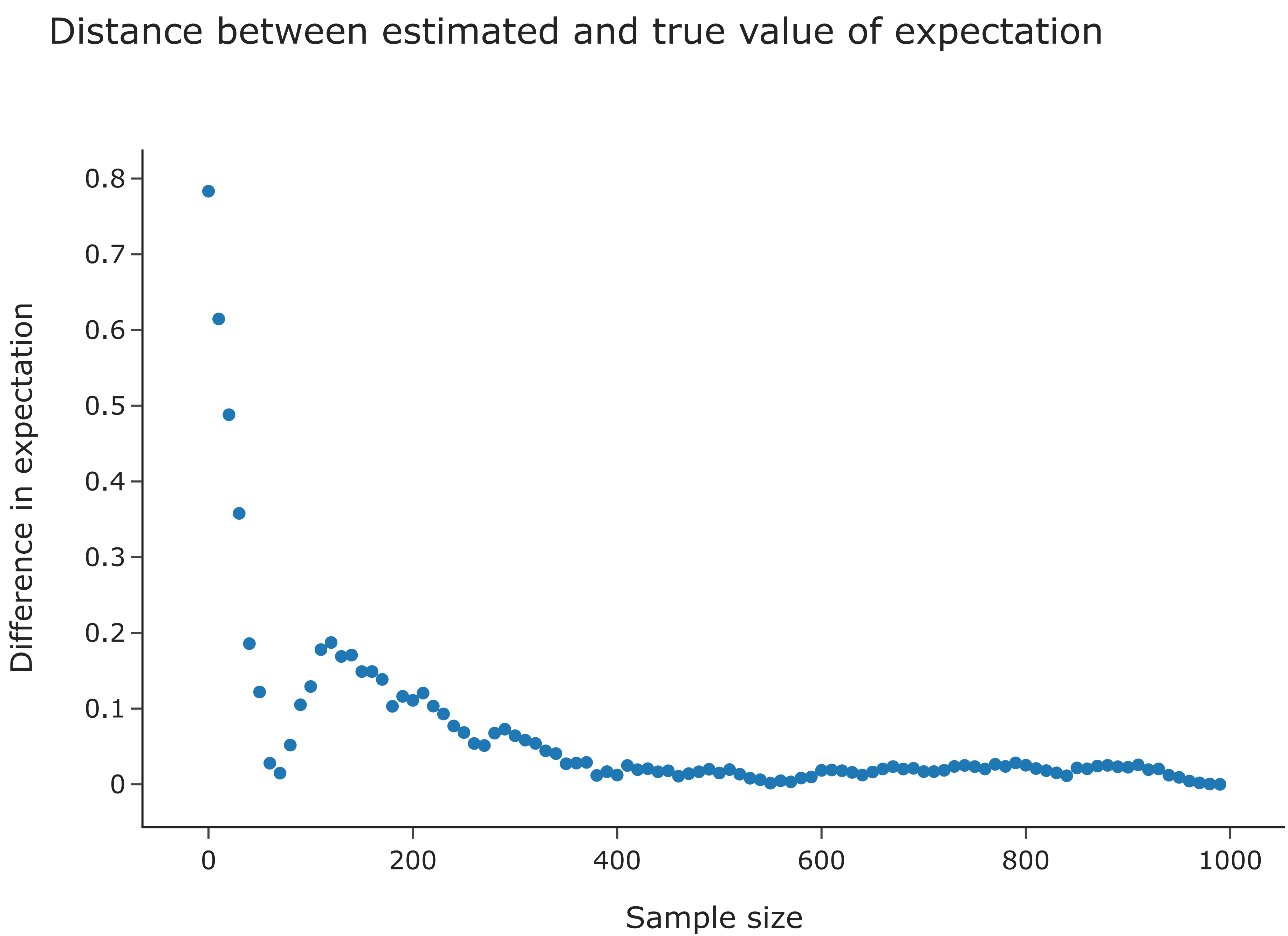
3) בסוף ה pdf

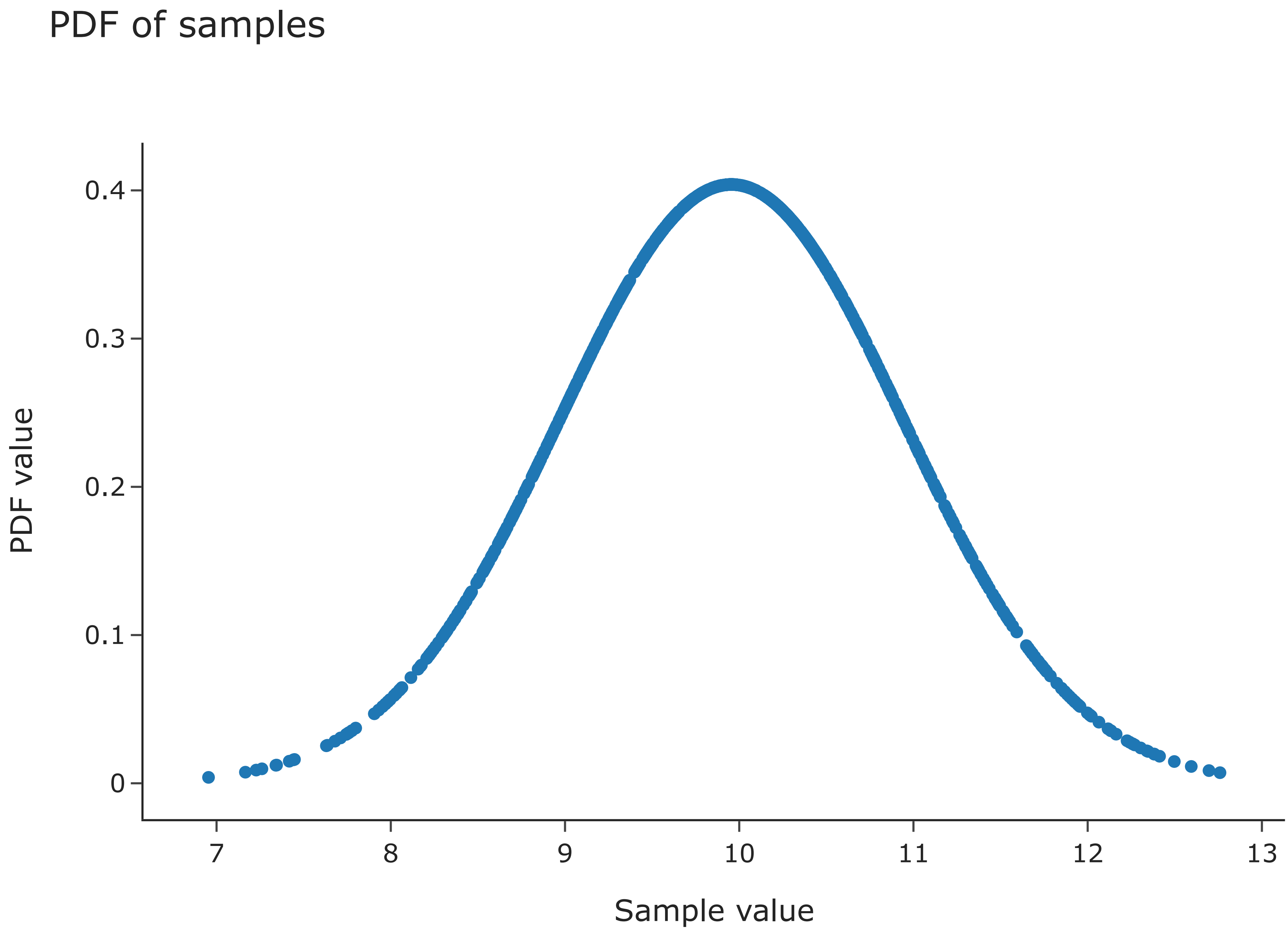
4)

$$\hat{\mu} = \begin{bmatrix} -0.022 \\ -0.043 \\ 3.993 \\ -0.020 \end{bmatrix} \qquad \hat{\Sigma} = \begin{bmatrix} 0.917 & 0.166 & -0.030 & 0.463 \\ 0.166 & 1.974 & -0.006 & 0.046 \\ -0.030 & -0.006 & 0.980 & -0.020 \\ 0.463 & 0.046 & -0.020 & 0.973 \end{bmatrix}$$

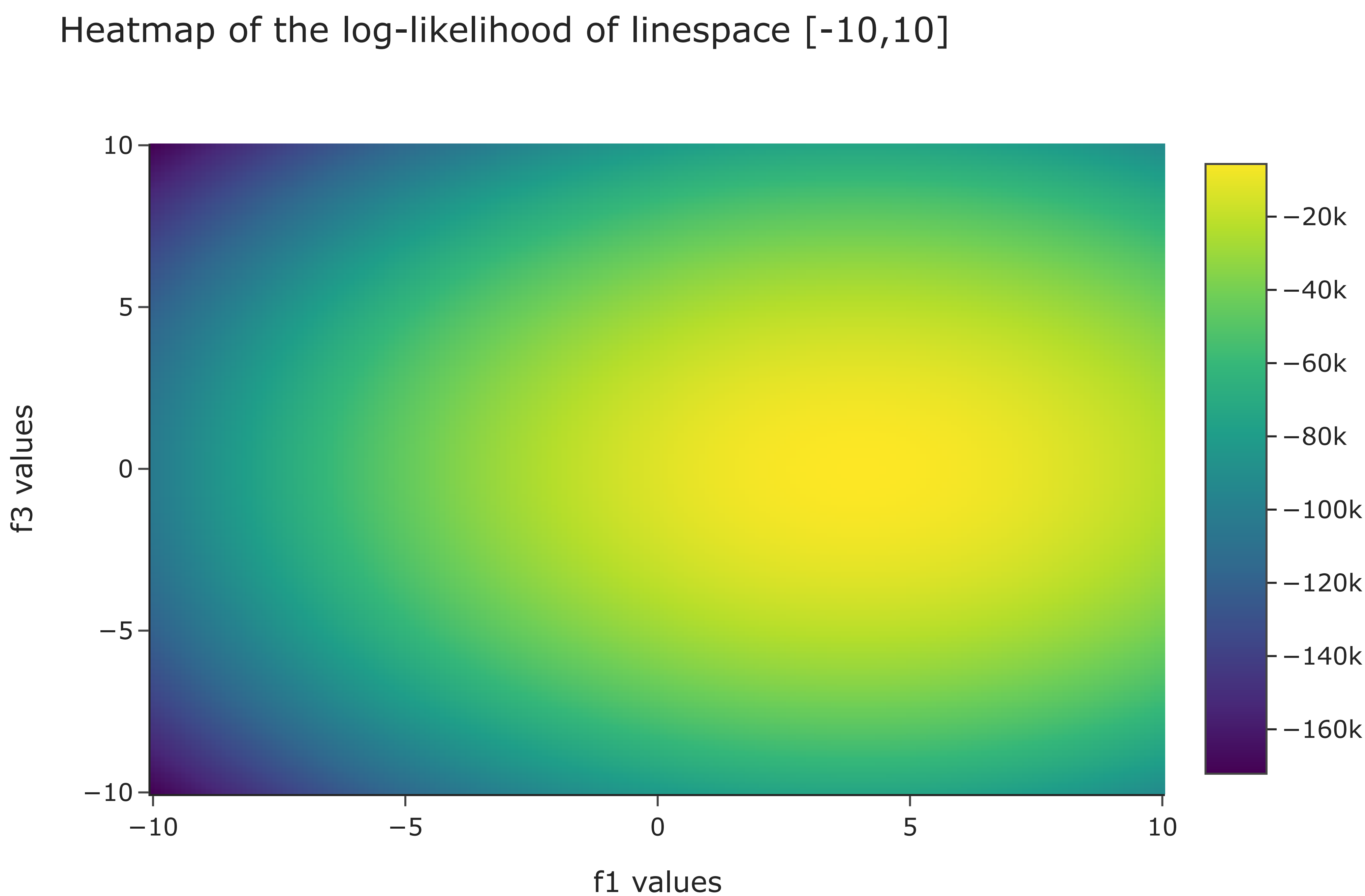
5) בסוף ה pdf

6) $(-0.050, 3.970)$


Distance between estimated and true value of expectation
Difference in expectation
Sample size
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
200
400
600
800
1000



PDF of samples
PDF value
0.4
0.3
0.2
0.1
0
7
8
9
10
11
12
13
Sample value


Heatmap of the log-likelihood of linespace [-10,10]
10
5
0
−5
−10
f3 values
−10
−5
0
5
10
f1 values
−20k
−40k
−60k
−80k
−100k
−120k
−140k
−160k