תרגיל בלומדות – ת ג'

3 נ'

1. תהי $A_{d\times d}$ מט' אורתוגונלית. אזי $\forall x$

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

A אורתוגונלית לכן הע"ע שלה ב-$\{1,-1\}$.

נסמן ב $\bar{V}_1,\dots,\bar{V}_d$ הו"ע של A.

אלו מהווים בסיס אורתונורמלי פורש למרחב הקואורדינטות R^d

(הרצ מאוחר ש A אורתוגונלית נימצע ש $d = \dim V = \dim W$)

נפרוש את x על V_i:

$$\boxed{x = \sum_{i=1}^{d} a_i \bar{V}_i}$$

ולכן

$$Ax = \sum_{i=1}^{d} a_i A\bar{V}_i =$$

$$= \boxed{\sum_{i=1}^{d} a_i \lambda_i \bar{V}_i}$$

V_i בסיס אורתונורמלי פורש

ולכן

$$\|Ax\| = \sqrt{\sum_{i=1}^{d} a_i^2 \lambda_i^2} = \sqrt{\sum_{i=1}^{d} a_i^2} = \|x\|$$

מ.ש.ל.

$A=\begin{bmatrix}1&1&0\\1&-1&2\end{bmatrix}$ $\quad A=U\Sigma V^T$ .2

$$AA^T=U\Sigma V^TV\Sigma^TU^T=U\Sigma\Sigma^TU^T$$

מאחר ש- U אורתוגונלית אזי הע"ע של AA^T

ושל $\Sigma\Sigma^T$ הם שווים:

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

ולכן הע"ע של $\Sigma\Sigma^T$ הם 2 ו-6:

ולכן מאחר ש-Σ אלכסונית: $\Sigma=\begin{bmatrix}\sqrt2&0&0\\0&\sqrt6&0\end{bmatrix}$

נמצא את U הן הו"ע של AA^T:

$$AA^T-2I=\begin{bmatrix}0&0\\0&4\end{bmatrix}\xRightarrow{\text{דירוג}}\bar u_1=\begin{pmatrix}1\\0\end{pmatrix}$$

$$AA^T-6I=\begin{bmatrix}-4&0\\0&0\end{bmatrix}\xRightarrow{\text{דירוג}}\bar u_2=\begin{pmatrix}0\\1\end{pmatrix}$$

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

נמצא: $A^TA=V\Sigma^T\Sigma V^T$

ולכן משיקולים דומים נמצא את הו"ע של A^TA.

המשך בעמוד הבא

2. המשך.

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

$$|A^TA-\lambda I| = \begin{vmatrix}2-\lambda&0&2\\0&2-\lambda&-2\\2&-2&4-\lambda\end{vmatrix} = \begin{vmatrix}2-\lambda&2-\lambda&2\\0&2-\lambda&-2\\2&0&4-\lambda\end{vmatrix} =$$

$$= \begin{vmatrix}2-\lambda&0&4\\0&2-\lambda&-2\\2&0&4-\lambda\end{vmatrix} = (2-\lambda)^2(4-\lambda)-8(2-\lambda) =$$

$$= (2-\lambda)\left[(2-\lambda)(4-\lambda)-8\right] = (2-\lambda)\left[8-6\lambda+\lambda^2-8\right] =$$

$$= (2-\lambda)(\lambda-6)\lambda \Longrightarrow \boxed{\lambda_1=2 \quad \lambda_2=6 \quad \lambda_3=0}$$

נמצא הו"ע של A^TA הם העמודות של V:

$$A^TA-2I = \begin{bmatrix}0&0&2\\0&0&-2\\2&-2&2\end{bmatrix} \longrightarrow \begin{bmatrix}2&-2&0\\0&0&1\\0&0&0\end{bmatrix} \Longrightarrow V_1 = \begin{pmatrix}\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}\\0\end{pmatrix}$$

$$A^TA-6I = \begin{bmatrix}-4&0&2\\0&-4&-2\\2&-2&-2\end{bmatrix} \longrightarrow \begin{bmatrix}2&-2&-2\\0&-4&-2\\0&-4&-2\end{bmatrix} \longrightarrow \begin{bmatrix}2&-2&-2\\0&2&1\\0&0&0\end{bmatrix} \rightarrow \begin{bmatrix}2&0&-1\\0&2&1\\0&0&0\end{bmatrix}$$

$$V_2 = \begin{pmatrix}\frac12\alpha\\-\frac12\alpha\\\alpha\end{pmatrix} \underset{(\text{נורמה } 1)}{=} \sqrt{\frac23}\begin{pmatrix}\frac12\\-\frac12\\1\end{pmatrix}$$ לכן:

ולבסוף

$$A^TA = \begin{bmatrix}2&0&2\\0&2&-2\\2&-2&4\end{bmatrix} \rightarrow \begin{bmatrix}2&0&2\\0&2&-2\\0&-2&2\end{bmatrix} \rightarrow \begin{bmatrix}1&0&1\\0&1&-1\\0&0&0\end{bmatrix} \Rightarrow \frac{1}{\sqrt3}\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

ולכן

$$V = \begin{pmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt6}&-\frac{1}{\sqrt3}\\\frac{1}{\sqrt2}&\frac{1}{\sqrt6}&\frac{1}{\sqrt3}\\0&\sqrt{\frac23}&\frac{1}{\sqrt3}\end{pmatrix}$$

וסה"כ

$$A = \begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}\sqrt2&0&0\\0&\sqrt6&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}&-\frac{1}{\sqrt3}\\\frac{1}{\sqrt2}&\frac{1}{\sqrt6}&\frac{1}{\sqrt3}\\0&\sqrt{\frac23}&\frac{1}{\sqrt3}\end{bmatrix}$$

3

$A\in\mathbb{R}^{m\times n}$ $\quad C_0=A^TA \quad \lambda_1\geq\lambda_2\geq\dots\geq\lambda_n$

C_0 סימטרית ולכן ניתנת לפירוק EVD.

$$C_0=UDU^T$$

כאשר D אלכסונית ו- U אורתוגונלית

$$C_0 b_k \underset{\text{ראינו בתרגול}}{\overset{(*)}{=}} \sum_{i=1}^{d}\lambda_i\langle b_k,v_i\rangle v_i$$

נניח $\lambda_1>\lambda_2$. נשים לב:

$$b_k=\frac{C_0 b_{k-1}}{\|C_0 b_{k-1}\|} \overset{\dots}{=} \frac{C_0^k b_0}{\prod_{i=0}^{k-1}\|C_0 b_i\|} \underset{\substack{\text{תכונות}\\ \text{EVD}\\ \text{ומט' אורתוגונליות}}}{=} \frac{UD^kU^Tb_0}{\prod_{i=0}^{k-1}\|C_0 b_i\|} =$$

$$\underset{\text{לפי }(*)}{=} \frac{\sum_{i=1}^{d}\lambda_i^k\langle b_k,v_i\rangle v_i}{\prod_{i=0}^{k-1}\|C_0 b_i\|} = \frac{\sum_{i=1}^{d}\lambda_i^k\langle b_k,v_i\rangle v_i}{\prod_{i=0}^{k-1}\sqrt{\sum_{j=1}^{d}\lambda_j^2\langle b_i,v_j\rangle^2}} =$$

$$= \frac{\lambda_1^k\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_1}\right)^k\langle b_k,v_i\rangle v_i}{\prod_{i=0}^{k-1}\sqrt{\sum_{j=1}^{d}\lambda_j^2\langle b_i,v_j\rangle^2}} = \frac{\lambda_1^k\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_1}\right)^k\langle b_k,v_i\rangle v_i}{\lambda_1^k\prod_{i=0}^{k-1}\sqrt{\sum_{j=1}^{d}\left(\frac{\lambda_j}{\lambda_1}\right)^2\langle b_i,v_j\rangle^2}} =$$

$$= \frac{\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_1}\right)^k\langle b_k,v_i\rangle v_i}{\prod_{i=0}^{k-1}\sqrt{\sum_{j=1}^{d}\left(\frac{\lambda_j}{\lambda_1}\right)^2\langle b_i,v_j\rangle}}$$

3. המשך. כעת:

$$\lim_{k\to\infty} b_k = \lim_{k\to\infty} \frac{\sum_{i=1}^{d} \left(\frac{\lambda_i}{\lambda_1}\right)^k \langle b_k, v_i\rangle v_i}{\prod_{i=0}^{k} \sqrt{\sum_{j=1}^{d} \left(\frac{\lambda_j}{\lambda_1}\right)^2 \langle b_i, v_j\rangle^2}} = \lim_{k\to\infty} \frac{\langle b_k, v_1\rangle v_1}{\prod_{i=0}^{k-1} \sqrt{\sum_{j=1}^{d} \left(\frac{\lambda_j}{\lambda_1}\right)^2 \langle b_i, v_j\rangle^2}} =$$

$$\boxed{\text{לכל } i\neq 1 \quad \left(\frac{\lambda_i}{\lambda_1}\right)^k \xrightarrow{k\to\infty} 0}$$

נשים לב ש $\|b_k\|=1$ לכל $k>0$ (כי הוא תמיד מנורמל בחישוב), ומכך ש הכיוון של b_k הוא v_1 כאשר ייתכן בסימן מינוס וייתכן בסימן פלוס, ובנוסף $\|v_1\|=1$

ולכן אם ל b_k יש גבול אז הוא או $+v_1$ או $-v_1$.

בנוסף נשים לב שלכל סביבה של $\bar{v}_1$ ניתן לקחת k גדול מספיק כך ששאר רכיבי הביטוי $\sum_{i=1}^{d} \left(\frac{\lambda_i}{\lambda_1}\right)^k \langle b_k, v_i\rangle v_i$ (שאינם בכיוון $\bar{v}_1$) יהיו קטנים כ"כ כך שנוותר בתוך הסביבה (כי $|b_k|=|v_1|=1$ ממילא הכן ניתן רק לסטות בכיוון). לכן הגבול קיים!

$$\boxed{\lim_{k\to\infty} b_k = \pm v_1}$$

$$f: \mathbb{R}^n \to \mathbb{R}^n, \quad f(\bar{\sigma}) = U \cdot diag(\bar{\sigma}) \cdot U^T X \qquad .4$$

$$U^T = \begin{pmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{pmatrix}$$

$$f(\bar{\sigma}) = U \cdot diag(\bar{\sigma}) \cdot \begin{bmatrix} \langle \bar{u}_1, \bar{x} \rangle \\ \vdots \\ \langle \bar{u}_n, \bar{x} \rangle \end{bmatrix} =$$

$$= U \cdot \begin{bmatrix} \sigma_1 \langle u_1, x \rangle \\ \vdots \\ \sigma_n \langle u_n, x \rangle \end{bmatrix} = \sum_{i=1}^{n} \sigma_i \langle u_{i}, x \rangle \bar{u}_i$$

הגדרה

$$\equiv T(x)$$

$$J_x(f) = \begin{pmatrix} \frac{\partial f_1(\bar{u})}{\partial \sigma_1} & \cdots & \frac{\partial f_1(\bar{u})}{\partial \sigma_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\bar{u})}{\partial \sigma_1} & \cdots & \frac{\partial f_n(\bar{u})}{\partial \sigma_n} \end{pmatrix} = \begin{pmatrix} \langle u_1, x \rangle u_{11} & \cdots & \langle u_n, x \rangle u_{1n} \\ \langle u_2, x \rangle u_{21} & & \vdots \\ \vdots & & \vdots \\ \langle u_n, x \rangle u_{n1} & \cdots & \langle u_n, x \rangle u_{nn} \end{pmatrix}$$

$$h(\bar{\sigma}) = \frac{1}{2} \| f(\sigma) - y \|^2 \qquad .5$$

$$J_\sigma(h) = J_f\left(\frac{1}{2} \| f(\sigma) - y \|^2\right) \cdot J_\sigma(f) =$$

$$= J_x\left(\frac{1}{2}(x_1^2 + \dots + x_n^2)\right)\Big|_{x=f(\sigma)} \cdot J_\sigma(f) =$$

$$= f(\bar{\sigma})^T \cdot J_\sigma(f) = U \cdot diag(\bar{\sigma}) \cdot U^T X \cdot T(x)$$

$$S(\bar{x})_j = \frac{e^{x_j}}{\sum_{l=1}^{k} e^{x_l}} \qquad .6$$

$$J_x(S) = \begin{pmatrix} \frac{\partial S_1}{\partial x_1} & \cdots & \frac{\partial S_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial S_n}{\partial x_1} & \cdots & \frac{\partial S_n}{\partial x_n} \end{pmatrix}$$

6. המשך.

$$\frac{\partial S_a}{\partial X_b} = \frac{\partial \frac{e^{x_a}}{\sum_{l=1}^{k} e^{x_l}}}{\partial X_b}$$

אם $a \neq b$:

$$\frac{\partial S_a}{\partial X_b} = \frac{-e^{x_a} e^{x_b}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

אם $a = b$:

$$\frac{\partial S_a}{X_a} = \frac{e^{x_a} \sum_{l=1}^{k} e^{x_l} - e^{2x_a}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2}$$

$$J_x(S) = \begin{pmatrix} \frac{e^{x_1} \sum_{l=1}^{k} e^{x_l} - e^{2x_1}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} & \cdots & \frac{-e^{x_1} e^{x_n}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} \\ \vdots & \ddots & \vdots \\ \frac{-e^{x_n} e^{x_1}}{\sum_{l=1}^{k} e^{x_l}} & \cdots & \frac{e^{x_n} \sum_{l=1}^{k} e^{x_l} - e^{2x_1}}{\left(\sum_{l=1}^{k} e^{x_l}\right)^2} \end{pmatrix}$$

7. $f(x,y) = x^3 - 5xy - y^5$ $\quad f: R^d \to R$

$$H_x(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial f}{\partial x \partial y} \\ \frac{\partial f}{\partial x \partial y} & \frac{\partial f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & -5 \\ -5 & -20y^3 \end{pmatrix}$$

8. צ"ל $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ הוא consistent.

ז"א צ"ל $\forall \varepsilon > 0 \quad \lim_{n\to\infty} P(|\hat{\mu}_n - \mu| > \varepsilon) = 0$

יהי $\varepsilon > 0$.

$$\cancel{P(|\hat{\mu}_n - \mu| > \varepsilon) \le \frac{E(|\mu - \hat{\mu}_n|)}{\varepsilon}}$$

$$E(\|\mu - \hat{\mu}_n\|)^2 \le E(\|\mu - \hat{\mu}_n\|^2) = E\left(\left[E(\hat{\mu}_n) - \hat{\mu}_n\right]^2\right) =$$

$$= V(\hat{\mu}_n)$$

$$P(\|\hat{\mu}_n - \mu\| > \varepsilon) \le \frac{E(\|\mu - \hat{\mu}_n\|)}{\varepsilon} \le \sqrt{V(\hat{\mu}_n)}$$

$$\cancel{\le E(\|\mu - \hat{\mu}_n\|) \le \sqrt{V(\hat{\mu}_n)}}$$

$$V(\hat{\mu}_n) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{i.i.d}{=} \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{\sigma^2}{n}$$

$$0 \xleftarrow{n\to\infty} 0 \le P(\|\hat{\mu}_n - \mu\| > \varepsilon) \le \sqrt{V(\hat{\mu}_n)} = \frac{\sigma}{\sqrt{n}} \xrightarrow{n\to\infty} 0$$

לכן ממשפט סנדוויץ' $\lim_{n\to\infty} P(\|\hat{\mu}_n - \mu\| > \varepsilon) = 0$

מ.ש.ל

$$\bar{X}_1,\dots,\bar{X}_m \overset{iid}{\sim} N(\bar{\mu},\Sigma) \quad , \mu\in R^d,\ \Sigma\in R^{d\times d}$$

$\bar{\theta} = (\bar{\mu},\Sigma)$

$$L(\bar{\theta}\,|\,\bar{x}_1,\dots,\bar{x}_m) = f_{\bar{\theta}}(\bar{x}_1,\dots,\bar{x}_m) \overset{iid}{=} \prod_{i=1}^{m} f_{\bar{\theta}}(\bar{x}_i) =$$

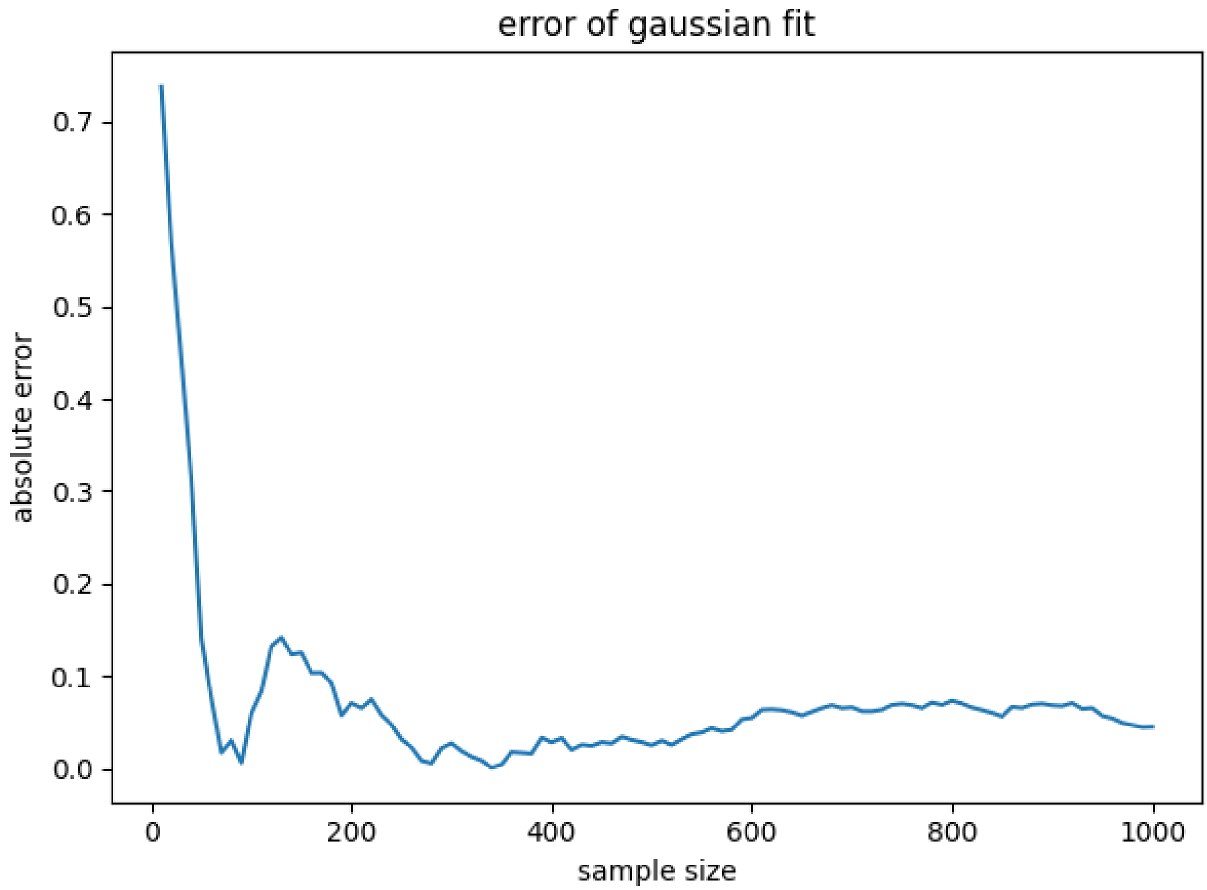
$$= \prod_{i=1}^{m} \frac{1}{\sqrt{(2\pi)^d|\Sigma|}}\, e^{\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\}}$$

$$\log L(\bar{\theta}\,|\,\bar{x}_1,\dots,\bar{x}_m) = \log\left(\frac{1}{((2\pi)^d|\Sigma|)^{\frac{m}{2}}}\right) - \frac{1}{2}\sum_{i=1}^{m}(\bar{x}_i-\bar{\mu})^T\Sigma^{-1}(\bar{x}_i-\bar{\mu}) =$$

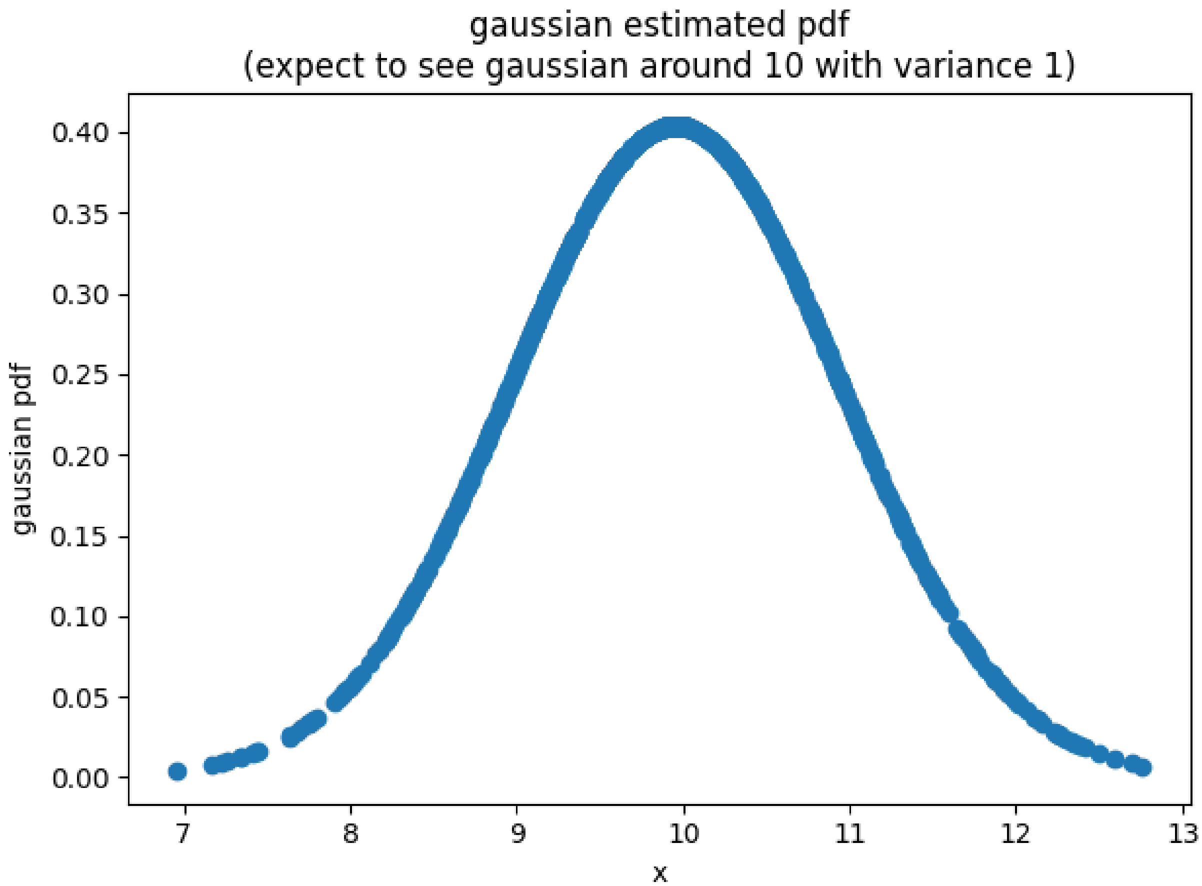
$$= -\log\left((2\pi)^{\frac{dm}{2}}\right) - \log\left(|\Sigma|^{\frac{m}{2}}\right) - \frac{1}{2}\sum_{i=1}^{m}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) =$$

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

תרגיל 1 שאלה 3.1.2



תרגיל 1 שאלה 3.1.3



תרגיל 1 שאלה 3.2.5