ELE520 Exercise 2 Daniel Fylling

P1.

$$\hat{\mu} = \frac{1}{N} \sum_{k=1}^{N} \underline{x}_k$$

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^{N} (\underline{x}_k - \hat{\mu})(\underline{x}_k - \hat{\mu})^T$$

$$\hat{\mu}_1 = \left[\begin{pmatrix}2\\6\end{pmatrix} + \begin{pmatrix}3\\4\end{pmatrix} + \begin{pmatrix}3\\8\end{pmatrix} + \begin{pmatrix}4\\6\end{pmatrix}\right]\frac{1}{4} = \begin{pmatrix}3\\6\end{pmatrix}$$

$$\hat{\mu}_2 = \left[\begin{pmatrix}1\\-2\end{pmatrix} + \begin{pmatrix}2,7\\-4\end{pmatrix} + \begin{pmatrix}3,3\\0\end{pmatrix} + \begin{pmatrix}5\\-2\end{pmatrix}\right]\frac{1}{4} = \begin{pmatrix}3\\-2\end{pmatrix}$$

$$\hat{\Sigma}_1 = \frac{1}{4}\left[\begin{pmatrix}2\\6\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix}\right)^2 + \left(\begin{pmatrix}3\\4\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix}\right)^2 + \left(\begin{pmatrix}3\\8\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix}\right)^2 + \left(\begin{pmatrix}4\\6\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix}\right)^2\right]$$

$$= \frac{1}{4}\left(\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\4\end{pmatrix} + \begin{pmatrix}0\\4\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}0,5 & 0\\0 & 2\end{pmatrix}$$

$$\hat{\Sigma}_2 = \frac{1}{4}\left[\left(\begin{pmatrix}1\\-2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix}\right)^2 + \left(\begin{pmatrix}2,7\\-4\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix}\right)^2 + \left(\begin{pmatrix}3,3\\0\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix}\right)^2 + \left(\begin{pmatrix}5\\-2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix}\right)^2\right]$$

$$= \frac{1}{4}\left[\begin{pmatrix}4\\0\end{pmatrix} + \begin{pmatrix}0,09\\4\end{pmatrix} + \begin{pmatrix}0,09\\4\end{pmatrix} + \begin{pmatrix}4\\0\end{pmatrix}\right] = \begin{pmatrix}2,045 & 0,3\\0,3 & 2\end{pmatrix}$$

$$g_i(x) = \underline{x}^T \Theta_i \underline{x} + \theta_i^T \underline{x} + \theta_{i0}$$

$$\Theta_i = -\frac{1}{2}\Sigma_i^{-1} \quad i=1 \Rightarrow \begin{bmatrix}-1 & 0\\0 & -\frac{1}{4}\end{bmatrix} \quad i=2 \quad \begin{bmatrix}-0,25 & 0,0375\\0,0375 & -0,256\end{bmatrix}$$

$$\theta_i = \Sigma_i^{-1}\mu_i \quad \begin{bmatrix}6\\3\end{bmatrix} \quad \begin{bmatrix}1,65\\-1,25\end{bmatrix}$$

$$\theta_{i0} = -\frac{1}{2}\mu_i^T \Sigma_i^{-1}\mu_i - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i)$$

$$\theta_{10} = -18 \quad 0 + \ln\frac{1}{2} = -18,69$$

$$\theta_{20} = -3,72 - \frac{1}{2}\ln 4 + \ln\frac{1}{2} = -5,10$$

$$g_1(x) = -x_1^2 - \frac{1}{4}x_2^2 + 6x_1 + 3x_2 - 18,69$$

$$g_2(x) = -0,25x_1^2 + 0,075x_1x_2 - 0,256x_2^2 + 1,65x_1 - 1,25x_2 - 5,1$$

ELE 520 Exercise 2 Daniel Fylling

P7

a, cont

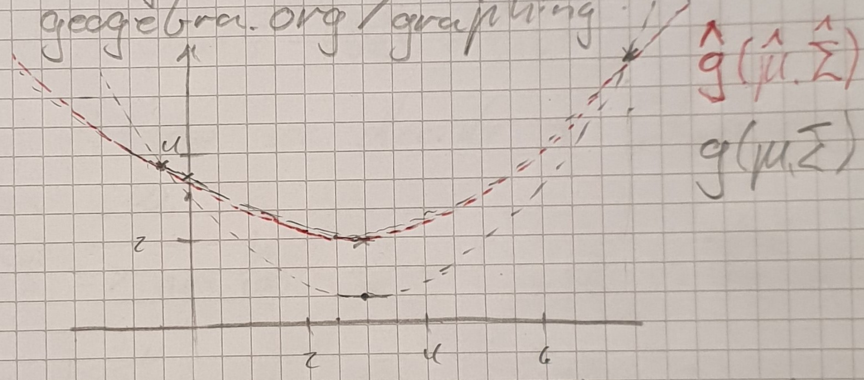
$$g_1(x) - g_2(x) = 0$$

$$-0.75x_1^2 + 0.006x_2^2 - 0.075x_1x_2 + 4.35x_1 + 4.25x_2 - 13.59 = 0$$

solved by symbolab.com/solver:

$$x_2 = \frac{0.075x_1 - 4.25 \pm \sqrt{0.0296x_1^2 - 0.75x_1 + 18.39}}{0.012}$$

b) Plotted against solution from E1 at geogebra.org/graphing



$\hat{\mu}_1$ has a much higher x_2 than μ_1 - 6 vs 3

$\hat{\Sigma}_2$ contains covariance terms not found in Σ_2

- it is expected that an estimate from such a low sampling would deviate significantly from the "true" distribution

c) Increase number of samples.

ELE 520 Exercise 2 Daniel Fylling

P2

a) for $g_1 > g_2$ decide ω_1
otherwise decide ω_2

$X = (2.5, 2)$

$g_1 = -x_1^2 - \frac{1}{4}x_2^2 + 6x_1 + 3x_2 - 18.69$
$= -6.25 - 1 + 15 + 6 - 18.69 = -4.94$

$g_2 = -0.25x_1^2 + 0.075x_1x_2 - 0.256x_2^2 + 1.65x_1 - 1.25x_2 - 5.1$
$= -1.56 + 0.375 - 1.024 + 4.13 - 2.5 - 5.1 = -5.68$

$g_1 > g_2$ so we decide X belongs in class ω_1

b) $h_1 = 0.5 \Rightarrow h_n = \frac{h_1}{\sqrt{n}} = \frac{0.5}{2} = \frac{1}{4}$ $d = 2$ $V_n = h_n^d$

$$p_n(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{V_n (2\pi)^{\frac{d}{2}} |I|^{\frac{1}{2}}} e^{-\frac{1}{2}\frac{(x-x_i)^t}{h_n} I^{-1} \frac{(x-x_i)}{h_n}}$$

$$p_n(x) = \frac{1}{n}\sum_{1}^{n} \frac{1}{2\pi\, h_n^2} e^{-\frac{1}{2h_n^2}(x-x_i)^t(x-x_i)}$$

1: $g_1(x) = \frac{16}{4 \cdot 2 \cdot \pi}\left[e^{-8(0.5^2+4^2)} + e^{-8(0.5^2+2^2)} + e^{-8(0.5^2+6^2)} + e^{-8(1.5^2+4^2)}\right]$

$= \frac{2}{\pi}\left[3.5 \cdot 10^{-57} + 1.7 \cdot 10^{-15} + 1.3 \cdot 10^{-126} + 3.9 \cdot 10^{-64}\right]$

$= \frac{2}{\pi} 1.7 \cdot 10^{-15}$

2: $g_2(x) = \frac{2}{\pi}\left[e^{-8(1.5^2+4^2)} + e^{-8(0.2^2+6^2)} + e^{-8(0.8^2+2^2)} + e^{-8(2.5^2+4^2)}\right]$

$= \frac{2}{\pi}\left[3.9 \cdot 10^{-64} + 6 \cdot 10^{-126} + 7.6 \cdot 10^{-17} + 5 \cdot 10^{-78}\right]$

$= \frac{2}{\pi} \cdot 7.6 \cdot 10^{-17}$

$g_1 > g_2 \rightarrow$ same conclusion

ELE520 Exercise 2

P2 c)

$h_1 = 5 \qquad h_N = 2{,}5 \qquad h_N^2 = 6{,}25$

$$g_1(x) = \frac{1}{4 \cdot 2 \cdot \pi \cdot 6{,}25}\left[e^{-\frac{2}{25}(0{,}5^2+4^2)} + e^{-\frac{2}{25}(0{,}5+2)} + e^{-\frac{2}{25}(0{,}8^2+6^2)} + e^{-\frac{2}{25}(1{,}5^2+4^2)}\right]$$

$$= \frac{1}{50\pi}\left[0{,}27 + 0{,}71 + 0{,}05 + 0{,}23\right]$$

$$= \frac{1}{50\pi}\left[1{,}26\right]$$

$$g_2(x) = \frac{1}{50\pi}\left[e^{-\frac{2}{25}(1{,}5^2+4^2)} + e^{-\frac{2}{25}(0{,}2^2+6^2)} + e^{-\frac{2}{25}(0{,}8^2+2^2)} + e^{-\frac{2}{25}(2{,}5^2+4^2)}\right]$$

$$= \frac{1}{50\pi}\left[0{,}23 + 0{,}06 + 0{,}69 + 0{,}17\right]$$

$$= \frac{1}{50\pi}\left[1{,}15\right]$$

$g_1 > g_2$ still same conclusion

When $h_1 = 0{,}5$ it was clear that the class that contained the closest point to x would "win" as all other points had contributions at many orders of magnitude less. For $h_1 = 5$ we see that all sample points have relevant contributions toward g-value. $h_1 = 5$ seems more appropriate for this scale of problem.

ELE520 Exercise 2

P2

d) Distances between x and sample points of each class:

$$R_1 = [(0.5^2+4^2)^{\frac{1}{2}}, (0.5^2+2^2)^{\frac{1}{2}}, (0.5^2+6^2)^{\frac{1}{2}}, (1.5^2+4^2)^{\frac{1}{2}}]$$
$$= [4.03, 2.06, 6.02, 4.27]$$

$$R_2 = [(1.5^2+4^2)^{\frac{1}{2}}, (0.2^2+6^2)^{\frac{1}{2}}, (0.8^2+2^2)^{\frac{1}{2}}, (2.5^2+4^2)^{\frac{1}{2}}]$$
$$= [4.27, 6.0, 2.15, 4.72]$$

$k_N = 1 \qquad p_N(x) = \frac{k_N/N}{V_N} = \frac{k_N}{N\pi R_N^2}$

$g_1 = p_1 = \frac{1}{4\pi\, 2.06^2}$

$g_2 = \frac{1}{4\pi\, 2.15^2}$

$g_1 > g_2 \rightarrow$ x still identifies as ω_1.

*similar to $h_1 = 0.5$, $k=1$ is only sensitive to the closest sample to x.

e) $k_N = 3$

$g_1 = \frac{3}{4\pi} \cdot \frac{1}{4.27^2}$

$g_2 = \frac{3}{4\pi} \cdot \frac{1}{4.72^2}$

$g_1 > g_2 \rightarrow$ same conclusion

... if x is truly ω_1 we will never know... Even though all methods we tried gave the same result

ELE520 Exercise 2

3 $p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)} = V \cdot e^{u}$

$p(x_1)\cdot p(x_2)\cdot \dots \cdot p(x_n) = Ve^{u_1} \cdot Ve^{u_2} \cdot \dots \cdot Ve^{u_n}$

$= V^n e^{\sum_{i=1}^{n} u_i}$

$\sum_{i=1}^{n} \ln(p(x_i)) = n \ln V + \sum_{i=1}^{n} u_i$

$= n \ln V - \frac{1}{2} \sum (x_i - \mu)^t \Sigma^{-1} (x_i - \mu)$

$= n \ln V - \frac{1}{2}\left[\sum x_i^t \Sigma^{-1} x_i - \sum x_i^t \Sigma^{-1} \mu - \sum \mu \Sigma^{-1} x_i + \sum \mu^t \Sigma^{-1} \mu\right]$

Differentiating with respect to μ (V contains no μ) and setting to zero:

$-\frac{1}{2}\left[-\sum x_i \Sigma^{-1} - \sum x_i \Sigma^{-1} + n \cdot 2 \mu \Sigma^{-1}\right] = 0 \quad | \cdot \Sigma$

$-\sum x_i + n\mu = 0. \Rightarrow \mu = \frac{1}{n} \sum_{i=1}^{n} x_i$