Løsningsforslag eksamen 2013H

Oppgave 1

Se læarebok og notater.

Oppgave 2

a) Se læarebok og notater for beskrivelse av metoden. Kortfattet utledning av likningssystemet:

$$p(\mathcal{X}|\bar{\theta}) = \prod_{k=1}^{n} p(\bar{x}_k|\bar{\theta}) \quad \text{der } \mathcal{X} = \{\bar{x}_1, \cdots, \bar{x}_n\}$$

er treningssettet.

$$\Downarrow$$

$$\mathcal{L}(\bar{\theta}) = \ln p(\mathcal{X}|\bar{\theta}) = \sum_{k=1}^{n} \ln p(\bar{x}_k|\bar{\theta})$$

$$\Downarrow$$

$$\nabla_{\bar{\theta}} \mathcal{L}(\bar{\theta}) = \sum_{k=1}^{n} \nabla_{\bar{\theta}} \ln p(\bar{x}_k|\bar{\theta})$$

som skal settes lik null. Likningssystemet blir da:

$$\boxed{\sum_{k=1}^{n} \nabla_{\bar{\theta}} \ln p(\bar{x}_k|\bar{\theta}) = 0}$$

b) Her er:

$$p(x|\bar{\theta}) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{der } \bar{\theta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix} \text{ er ukjent}$$

slik at

$$\ln p(x|\bar{\theta}) = -\frac{(x-\mu)^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2)$$

De partialderiverte mhp. komponentene i $\bar{\theta}$ blir da:

$$\left.\begin{aligned}\frac{\partial \ln p(x|\mu,\sigma^2)}{\partial\mu} &= \frac{x-\mu}{\sigma^2}\\ \frac{\partial \ln p(x|\mu,\sigma^2)}{\partial\sigma^2} &= \frac{(x-\mu)^2}{2\sigma^4} - \frac{1}{2\sigma^2}\end{aligned}\right\} \text{gradienten til logaritmen til } p(x|\bar{\theta})$$

Estimatet av μ blir da:

$$\sum_{k=1}^{n} \frac{x_k-\mu}{\sigma^2} = 0 \Rightarrow \hat{\mu} = \underline{\underline{\frac{1}{n}\sum_{k=1}^{n} x_k}}$$ (forutsetter at $\sigma^2 \neq 0$)

mens variansen blir:

$$\sum_{k=1}^{n} \frac{(x_k-\mu)^2}{2\sigma^4} - \frac{n}{2\sigma^2} = 0 \Rightarrow \hat{\sigma}^2 = \underline{\underline{\frac{1}{n}\sum_{k=1}^{n}(x_k-\hat{\mu})^2}}$$

Oppgave 3

a) Se lærebok og notater for beskrivelse.

Eksempel:

$$g_i(\bar{x}) = \ln p(\bar{x}|\omega_i) + \ln P(\omega_i), \quad i=1,\dots,c$$

b)

$$\mu_1 = \frac{1}{n_1}\sum_{x\in\mathcal{X}_1} x = \frac{1}{10}(1.3+0.4+\dots+0.6) = \underline{\underline{1.15}}$$

$$\mu_2 = \frac{1}{n_2}\sum_{x\in\mathcal{X}_2} x = \frac{1}{10}(3.4+2.4+\dots+2.1) = \underline{\underline{2.78}}$$

$$\sigma_1^2 = \frac{1}{n_1}\sum_{x\in\mathcal{X}_1}(x-\mu_1)^2 = \frac{1}{10}\sum_{x\in\mathcal{X}_1}(x-1.15)^2 = 0.1685 \Rightarrow \underline{\underline{\sigma_1 = 0.410}}$$

$$\sigma_2^2 = \frac{1}{n_2}\sum_{x\in\mathcal{X}_2}(x-\mu_2)^2 = \frac{1}{10}\sum_{x\in\mathcal{X}_2}(x-2.78)^2 = 0.4396 \Rightarrow \underline{\underline{\sigma_2 = 0.663}}$$

c) En mulig toklasse diskriminantfunksjon er:

$$g(x) = g_1(x) - g_2(x) = \ln p(x|\omega_1) + \ln P(\omega_1) - (\ln p(x|\omega_2) + \ln P(\omega_2))$$

$$= \ln p(x|\omega_1) - \ln p(x|\omega_2) + \ln P(\omega_1) - \ln P(\omega_2)$$

Siden det er like mange sampler i $\mathcal{X}_1$ som i $\mathcal{X}_2$ er det naturlig å anta $P(\omega_1) = P(\omega_2) = \frac{1}{2}$. De to siste leddene faller da bort, og diskriminant-funksjonen blir (med innsetting av normalfordelingene):

$$g(x) = -\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{1}{2}\ln(2\pi) - \ln\sigma_1 + \frac{(x-\mu_2)^2}{2\sigma_2^2} + \frac{1}{2}\ln(2\pi) + \ln\sigma_2$$

$$= \frac{x^2 - 2\mu_2 x + \mu_2^2}{2\sigma_2^2} - \frac{x^2 - 2\mu_1 x + \mu_1^2}{2\sigma_1^2} + \ln\frac{\sigma_2}{\sigma_1}$$

$$= \frac{x^2}{2\sigma_2^2} - \frac{\mu_2}{\sigma_2^2}x + \frac{\mu_2^2}{2\sigma_2^2} - \frac{x^2}{2\sigma_1^2} + \frac{\mu_1}{\sigma_1^2}x - \frac{\mu_1^2}{2\sigma_1^2} + \ln\frac{\sigma_2}{\sigma_1}$$

$$= \underbrace{\left(\frac{1}{2\sigma_2^2} - \frac{1}{2\sigma_1^2}\right)}_{a = -1{,}83}x^2 + \underbrace{\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)}_{b = 0{,}501}x + \underbrace{\frac{\mu_2^2}{2\sigma_2^2} - \frac{\mu_1^2}{2\sigma_1^2} + \ln\frac{\sigma_2}{\sigma_1}}_{c = 5{,}35}$$

$$= \underline{\underline{ax^2 + bx + c}}$$ kvadratisk diskr. funk.

Desisjonsgrensene er da gitt ved:

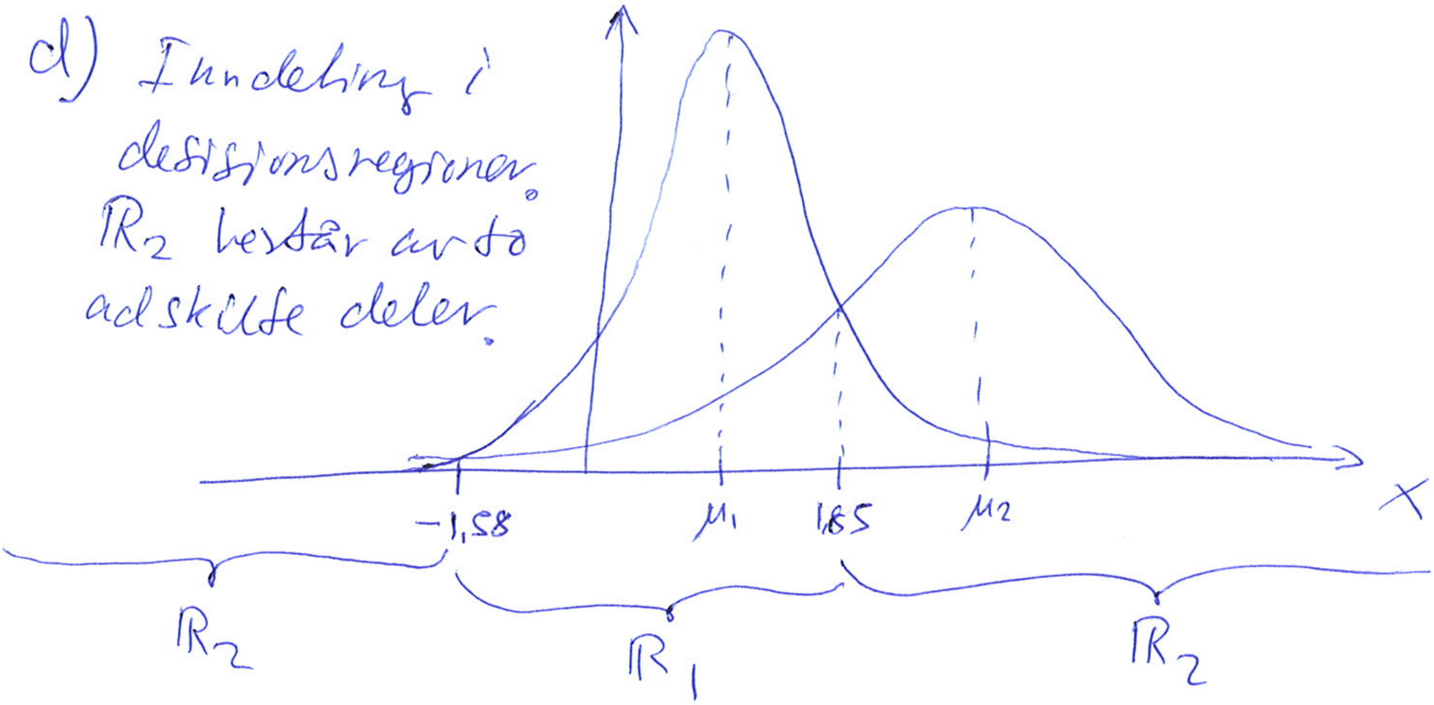
$$ax^2 + bx + c = 0$$

$$\Downarrow$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \begin{cases} -1{,}58 \\ 1{,}85 \end{cases}$$

Desisjonsregionen for ω_1 vil ligge mellom disse to terskelverdiene (i origo er $g(0) = c > 0$, dvs. R_1).

d) Inndeling i desisjonsregioner. $\mathbb{R}_2$ består av to adskilte deler.



e) Ukjent sample gitt ved $x_0 = 2.0$.

Her blir:

$$g(x_0) = g(2,0) = -0,968 < 0$$

Desisjonsregelen velger da klasse $\underline{\underline{\omega_2}}$.

Oppgave 4

a) Se lærebok og notater.

b) —— " ——

c) Estimatet er her:

$$p_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{V_n} \varphi\left(\frac{x - x_i}{h}\right) \quad \text{der } V_n = h = 1.$$

De estimerte tetthetene i punktet x_0 blir da for hver av klassene:

$$p_n(x_0 | \omega_1) = \frac{1}{10} \sum_{i=1}^{n} \varphi(2 - x_i) = \frac{2}{10} = \underline{\underline{0,2}} \quad (x_i \in \mathcal{X}_1)$$

$$p_n(x_0 | \omega_2) = \text{—— " ——} = \frac{4}{10} = \underline{\underline{0,4}} \quad (x_i \in \mathcal{X}_2)$$

d) Á posteriori sannsynlighetene blir da i h.h.t. Bayes regel:

$$\hat{P}(\omega_1|x_0) = \frac{p_n(x_0|\omega_1)P(\omega_1)}{p(x)} = \frac{0{,}2 \cdot \frac{1}{2}}{0{,}2 \cdot \frac{1}{2} + 0{,}4 \cdot \frac{1}{2}} = \underline{\underline{\frac{1}{3}}}$$

$$\hat{P}(\omega_2|x_0) = \frac{p_n(x_0|\omega_2)P(\omega_2)}{p(x)} = \frac{0{,}4 \cdot \frac{1}{2}}{0{,}2 \cdot \frac{1}{2} + 0{,}4 \cdot \frac{1}{2}} = \underline{\underline{\frac{2}{3}}}$$

Minimum feilrateprinsippet sier at man skal velge klassen med størst á posteriori sannsynlighet. Her velges derfor klasse $\underline{\underline{\omega_2}}$.

e) Se lærebok og notater.

f) Den nærmeste naboen til $x_0 = 2.0$ er samplet „2.1" i $\mathcal{X}_2$, NNR vil da velge $\underline{\underline{\omega_2}}$.

De tre nærmeste naboene til $x_0 = 2.0$ er samplene 2.1 og 2.4 (eller 1.6) fra ω_2 og 1.8 fra ω_1. k-NNR med $k=3$ velger da også $\underline{\underline{\omega_2}}$, siden denne klassen har to av de tre nærmeste samplene.

Oppgave 5

Se lærebok og forelesningsnotater.