Eksamen i TTK4205 - 2011H

Oppgave 1

Se lærebok.

Oppgave 2

Se lærebok.

Oppgave 3 - Diskriminantfunksjoner

a) Se læreboka

b) Diskriminantfunksjoner for c klasser:

$$g_i(\bar{x}) = \ln P(\omega_i|\bar{x}) = \ln p(\bar{x}|\omega_i) + \ln P(\omega_i) - \ln p(\bar{x})$$

For to klasser:

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

Her velges ω_1 dersom $g(\bar{x}) > 0$ og ω_2 ellers.

c) Her er:

$$p(\bar{x}|\omega_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_i)^t \Sigma_i^{-1}(\bar{x}-\bar{\mu}_i)}, \quad i=1,2.$$

med:

$$\Sigma_1 = \Sigma_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \bar{\mu}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \bar{\mu}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

og $P(\omega_1) = P(\omega_2)$.

Dette gir:

$$p(\bar{x}|\omega_i) = \frac{1}{(2\pi)^{d/2} \cdot 2} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_i)^t \Sigma^{-1}(\bar{x}-\bar{\mu}_i)}$$

der

$$\Sigma^{-1} = \Sigma_1^{-1} = \Sigma_2^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{2} I$$

Diskriminantfunksjonen blir da:

$$g(\bar{x}) = \ln p(\bar{x}|w_1) - \ln p(\bar{x}|w_2) \qquad \left(\text{siden } \ln\frac{P(w_1)}{P(w_2)} = 0\right)$$

$$= -\frac{1}{4}(\bar{x}-\bar{\mu}_1)^t I(\bar{x}-\bar{\mu}_1) + \frac{1}{4}(\bar{x}-\bar{\mu}_2)^t I(\bar{x}-\bar{\mu}_2)$$

$$= \frac{1}{4}\left[(\bar{x}-\bar{\mu}_2)^t(\bar{x}-\bar{\mu}_2) - (\bar{x}-\bar{\mu}_1)^t(\bar{x}-\bar{\mu}_1)\right]$$

$$= \frac{1}{4}\left[\bar{x}^t\bar{x} - \bar{\mu}_2^t\bar{x} - \bar{x}^t\bar{\mu}_2 + \bar{\mu}_2^t\bar{\mu}_2 - (\bar{x}^t\bar{x} - \bar{\mu}_1^t\bar{x} - \bar{x}^t\bar{\mu}_1 + \bar{\mu}_1^t\bar{\mu}_1)\right]$$

$$= \frac{1}{4}\left[-2\bar{\mu}_2^t\bar{x} + 2\bar{\mu}_1^t\bar{x} + \bar{\mu}_2^t\bar{\mu}_2 - \bar{\mu}_1^t\bar{\mu}_1\right]$$

$$= \frac{1}{4}\left[2(\bar{\mu}_1-\bar{\mu}_2)^t\bar{x} + \bar{\mu}_2^t\bar{\mu}_2 - \bar{\mu}_1^t\bar{\mu}_1\right]$$

Dropper faktoren $\frac{1}{4}$ og setter inn for $\bar{\mu}_1$ og $\bar{\mu}_2$:

$$g(\bar{x}) = 2[-2,1]\bar{x} + 13 - 10$$

$$= \underline{\underline{2[-2,1]\bar{x} + 3}}$$

3d) Ukjent objekt $\bar{x}_0 = \begin{bmatrix} 2,0 \\ 1,0 \end{bmatrix}$

Her blir:

$$g(\bar{x}_0) = 2[-2,1]\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3 = -6 + 3 = -3$$

Siden $g(\bar{x}_0) < 0$ blir $\bar{x}_0$ klassifisert til $\underline{\underline{w_2}}$.

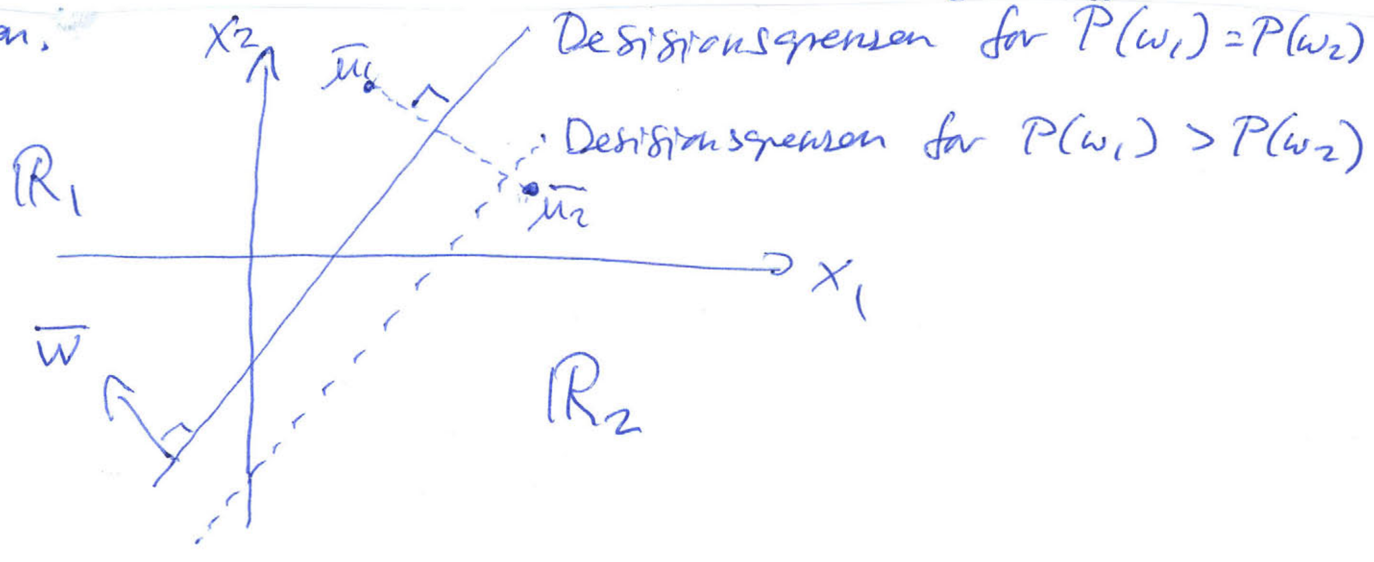
e) Siden $g(\bar{x})$ er lineær (på formen $g(\bar{x}) = \bar{w}^t\bar{x} + w_0$) er desisjonsgrensen et hyperplan generelt, og i dette tilfellet en rett linje.

3f)

Dersom $P(\omega_1) > P(\omega_2)$ får diskriminantfunksjonen et tillegg $\sim \ln \frac{P(\omega_1)}{P(\omega_2)}$ (positivt), se pkt. 3G.

Dette fører til at desisjonsgrensen flyttes „vekk" fra ω_1 „mot" ω_2, dvs. at desisjonsregionen $\mathbb{R}_1$ vokser på bekostning av $\mathbb{R}_2$. Årsaken er at klassifikatoren i dette tilfellet oftere vil velge ω_1, siden denne klassen i utgangspunktet er mer sannsynlig.

Se figuren.



3g)

Her er $\bar{w} = 2\begin{bmatrix}-2\\1\end{bmatrix}$ og $\bar{\mu}_1 - \bar{\mu}_2 = \begin{bmatrix}1\\3\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}-2\\1\end{bmatrix}$

Derav ser man at $\bar{w}$ er parallell med $\bar{\mu}_1 - \bar{\mu}_2$. Siden $\bar{w}$ er normalvektor til desisjonsgrensen, vil differansevektoren stå normalt på hyperplanet.

Oppgave 4

a) Likelihoodfunksjonen : $p(\mathcal{X}|\bar{\theta}) = \prod_{k=1}^{n} p(\bar{x}_k|\bar{\theta})$

Log-likelihood funksjonen : $\mathcal{L}(\bar{\theta}) = \sum_{k=1}^{n} \ln p(\bar{x}_k|\bar{\theta})$

$\mathcal{X} = \{\bar{x}_1, \dots, \bar{x}_n\}$ er treningssettet.

$\nabla_{\theta}\mathcal{L}(\bar{\theta}) = 0 \Rightarrow \boxed{\sum_{k=1}^{n} \nabla_{\theta} \ln p(\bar{x}_k|\bar{\theta}) = 0}$ dvs. likn.systemet som gir maksimum likelihood løsn.

b) Har er :

$$p(x|\theta) = \begin{cases} 2\theta x e^{-\theta x^2}, & x \geq 0 \\ 0 & \text{ellers} \end{cases} \quad \text{der } \theta > 0$$

Treningssett : $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$

Likelihoodfunksjonen blir da:

$$p(\mathcal{X}|\theta) = \prod_{k=1}^{n} p(x_k|\theta) = \prod_{k=1}^{n} 2\theta x_k e^{-\theta x_k^2}$$

og log-likelihood funksjonen :

$$\mathcal{L}(\theta) = \ln(2\theta)^n + \sum_{k=1}^{n} x_k - \theta \sum_{k=1}^{n} x_k^2$$

Gradienten blir da:

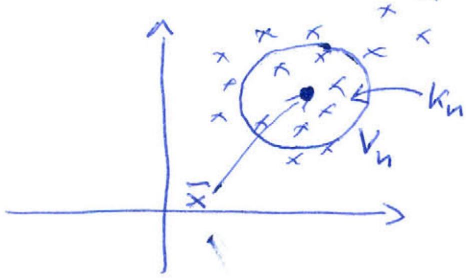
$$\nabla_{\theta}\mathcal{L}(\bar{\theta}) = \frac{n}{\theta} - \sum_{k=1}^{n} x_k^2$$

Maksimum likelihood løsningen finnes ved å sette $\nabla_{\theta}\mathcal{L}(\theta) = 0$, dvs.

$$\frac{n}{\theta} - \sum_{k=1}^{n} x_k^2 = 0 \Rightarrow \underline{\underline{\hat{\theta} = \frac{n}{\sum_{k=1}^{n} x_k^2}}}$$

Oppgave 5

a)

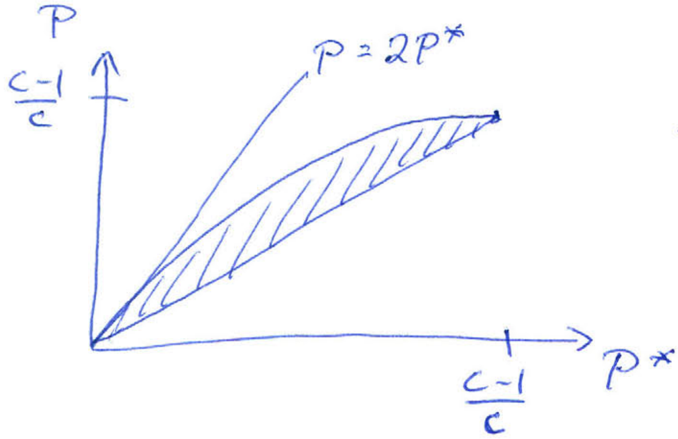


$$p_n(\bar{x}) = \frac{k_n/n}{V_n}$$

b) NNR: Velg samme klasse som den nærmeste naboen i treningssettet.

K-NNR: Velg den klassen som har flest representanter blant de k nærmeste naboene.

c)



NNR:

$$P^* \leq P \leq P^*\left(2 - \frac{c}{c-1}P^*\right)$$

K-NNR: Den øvre grensen $\longrightarrow P = P^*$ når $k \to \infty$

d) $\mathcal{X}_1 = \{0.1, 0.9, 1.3, 1.5, 1.8\}$ fra ω_1

$\mathcal{X}_2 = \{1.6, 1.9, 2.1, 2.3, 2.5\}$ fra ω_2

$x = 1.65$ (ukjent sample)

Klassifisering med NNR:

Nærmeste nabo er „1.6" fra ω_2

Nærmeste nabo-regelen velger da klasse $\underline{\underline{\omega_2}}$.

Klassifisering med k-NNR der $k=3$:

Nærmeste naboer er „1.5" og „1.8" fra ω_1 og „1.6" fra ω_2. k-NNR velger da $\underline{\omega_1}$.