Borna Nikolić
0036534977

DZ 4

V14 - Procjena parametra II

1.4.

- multivarijantna normalna razdioba

$\vec{x}^1 = (9.5, -0.7, -2.8)$
$\vec{x}^2 = (8.8, -0.8, -3.2)$
$\vec{x}^3 = (6.5, -0.2, -0.8)$
$\vec{x}^4 = (2.3, 0.3, 1.2)$
$\vec{x}^5 = (2.2, 0, 0)$
$\vec{x}^6 = (3.6, 0.3, 1.2)$

a)

$\hat{\vec{\mu}}_{MLE} = ?$

$\hat{\Sigma}_{MLE} = ?$

$$\hat{\vec{\mu}}_{MLE} = \frac{1}{N}\sum_{i=1}^{N}\vec{x}^i = \frac{1}{6}\sum_{i=1}^{6}\vec{x}^i = \left(\frac{329}{60}, -\frac{11}{60}, -\frac{11}{15}\right)$$
$$\approx (5.48, -0.18, -0.73)$$

$$\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N}(\vec{x}^i - \hat{\mu})(\vec{x}^i - \hat{\mu})^T$$
$$= \frac{1}{6}\left(\begin{bmatrix} 241/60 \\ -31/60 \\ -31/15 \end{bmatrix}\begin{bmatrix} \frac{241}{60} & \frac{-31}{60} & \frac{-31}{15} \end{bmatrix} + \dots\right)$$
$$= \begin{bmatrix} 8.771 & -1.198 & -4.792 \\ -1.198 & 0.191 & 0.766 \\ -4.792 & 0.766 & 3.06 \end{bmatrix}$$

b)

$\vec{x} = (-2, 1, 0)$

Je li ta gustoća dobro definirana?

$$p(\vec{x} \mid \hat{\mu}_{MLE}, \hat{\Sigma}_{MLE}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1}(\vec{x}-\vec{\mu})\right\}$$

$|\hat{\Sigma}| = 0$

Gustoća nije dobro definirana jer je determinanta MLE kovarijacijske matrice 0, tj. matrica je singularna
$\Rightarrow$ ne možemo izračunati gustoću vjerojatnosti (dijelimo s 0)

c)

$$\rho_{A,B} = \frac{Cov(A,B)}{\sigma_A \cdot \sigma_B}$$

$$\hat{\Sigma}_{MLE} = \begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline x_1 & 8.771 & -1.198 & -4.792 \\ x_2 & -1.198 & 0.191 & 0.766 \\ x_3 & -4.792 & 0.766 & 3.06 \end{array}$$

$$\rho_{x_1,x_2} = \frac{Cov(X_1,X_2)}{\sigma_{x_1} \cdot \sigma_{x_2}} = \frac{-1.198}{\sqrt{8.771 \cdot 0.191}} \approx -0.9256$$

$$\rho_{x_1,x_3} = \frac{Cov(X_1,X_3)}{\sigma_{x_1} \cdot \sigma_{x_3}} = \frac{-4.792}{\sqrt{8.771 \cdot 3.06}} \approx -0.92498$$

$$\rho_{x_2,x_3} = \frac{Cov(X_2,X_3)}{\sigma_{x_2} \cdot \sigma_{x_3}} = \frac{0.766}{\sqrt{0.191 \cdot 3.06}} \approx 1.002$$

$\Rightarrow$ iz ulaznog prostora izbacujemo značajku X_2 jer ima najveću korelaciju sa ostalim varijablama

nakon izbaciv. značajke

$$\hat{\Sigma}_{MLE} = \begin{bmatrix} 8.771 & -4.792 \\ -4.792 & 3.06 \end{bmatrix}$$

$$|\hat{\Sigma}_{MLE}| = 3.8945$$

$$\hat{\Sigma}_{MLE}^{-1} = \begin{bmatrix} 0.7863 & 1.2304 \\ 1.2304 & 2.2522 \end{bmatrix}$$

ne gledamo u računu

$$\vec{x} = (-2, 1, 0) \qquad n = 2 \qquad \hat{\mu}_{MLE} = (5.48, -0.73)$$

$$p(\vec{x} \mid \hat{\mu}_{MLE}, \hat{\Sigma}_{MLE}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right\}$$

$$= \ldots$$

$$= 1.0346 \cdot 10^{-8}$$

1.5

a) $\hat{\theta}_{MAP} = \underset{\theta}{\mathrm{argmax}}\, p(\vec{\theta}|D) = \underset{\theta}{\mathrm{argmax}}\, p(D|\vec{\theta})p(\vec{\theta})$

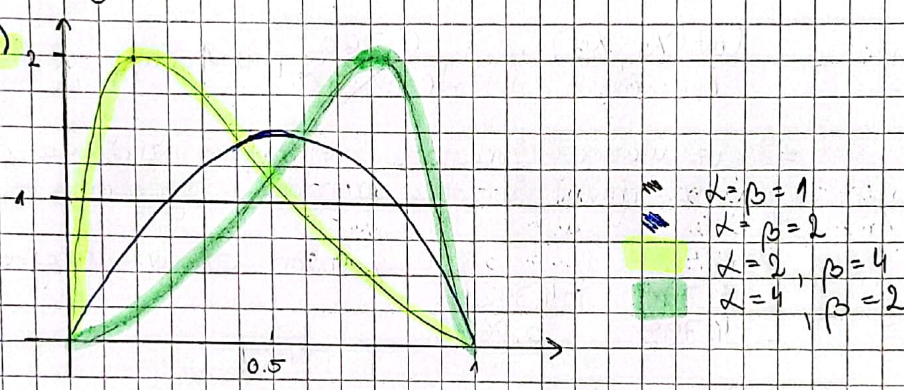
Procjenitelj MAP bolji je od procjenitelja MLE jer kombinira apriorno znanje i informacije dobivene iz podataka.

b) Za distribucije $p(\vec{\theta}|D)$ i $p(\vec{\theta})$ kažemo da su konjugatne ako su to distribucije iste vrste (npr. obje Gauss.)

Konjugatna apriorna distribucija je $p(\vec{\theta})$ (ona koja kada pomnožena izglednošću daje distribuciju iste vrste kao aposteriorna.

Svojstvo konjugatnosti nam je bitno jer nam omogućava „online" učenje.

c)



d) $p(D|\mu) = \mu^m (1-\mu)^{N-m}$

– izglednost

$p(\mu|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} \mu^{\alpha-1}(1-\mu)^{\beta-1}$

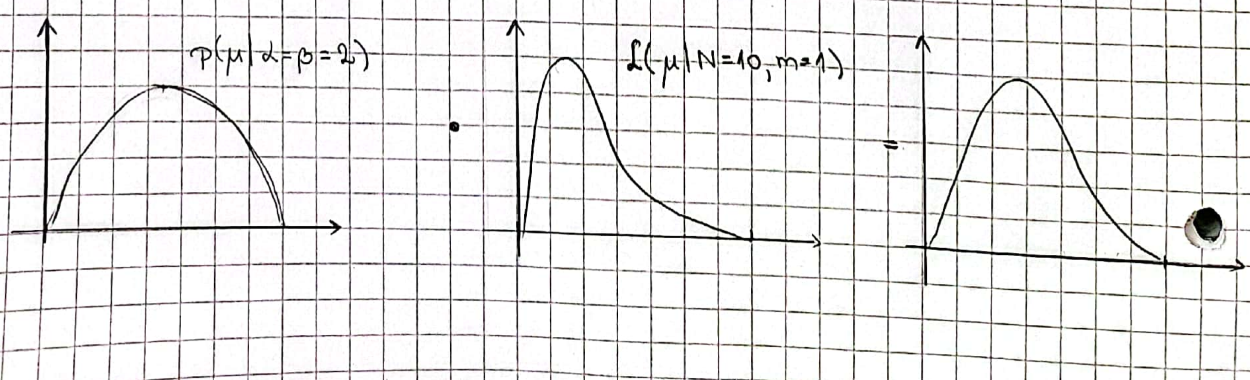
– apriorna dist.

$$p(\mu|D,\alpha,\beta) = p(\mu|N,m,\alpha,\beta) = \mu^m(1-\mu)^{N-m} \frac{1}{B(\alpha,\beta)} \mu^{\alpha-1}(1-\mu)^{\beta-1} \cdot \frac{1}{p(D)}$$

$$= \mu^{\underbrace{m+\alpha}_{\alpha'}-1}(1-\mu)^{\underbrace{N-m+\beta}_{\beta'}-1} \cdot \frac{1}{B(\alpha,\beta)\cdot p(D)}$$

$$= \frac{1}{B(\alpha',\beta')} \mu^{\alpha'-1}(1-\mu)^{\beta'-1}$$

e)



f) $\hat{\mu}_{MAP} = \frac{m+\alpha-1}{\alpha+N+\beta-2} = \frac{2}{12} \approx 0.167$

$\hat{\mu}_{MLE} = \frac{1}{10} = 0.1$

Porastom broja primjera N, razlika između $\hat{\mu}_{MAP}$ i $\hat{\mu}_{MLE}$ se smanjuje jer sve više vjerujemo podacima

g) $\alpha = \beta = 2$

$\Rightarrow \hat{\mu}_{MAP} = \frac{m+\alpha-1}{N+\alpha+\beta-2} = \frac{m+1}{N+2}$ što je Laplaceov procjen.

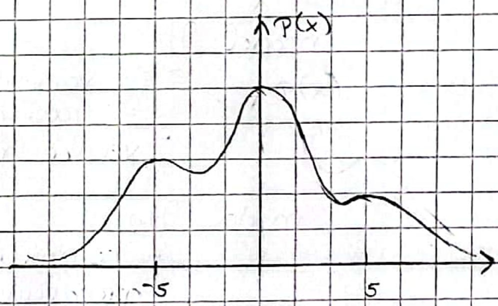
V15 - Bayesov klasifikator I

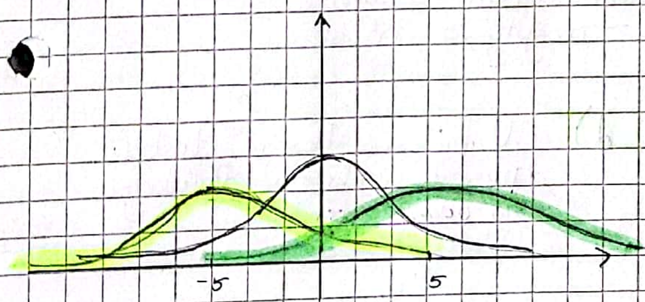
1.3.

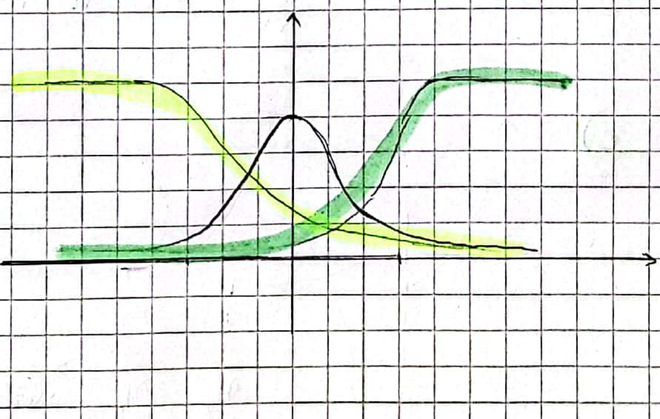
$P(y=1) = 0.3$ $\quad \mu_1 = -5$ $\quad \sigma_1^2 = 5$

$P(y=2) = 0.2$ $\quad \mu_2 = 0$ $\quad \sigma_2^2 = 1$

$P(y=3) = 0.5$ $\quad \mu_3 = 5$ $\quad \sigma_3^2 = 10$







1.5.

a) $\mathcal{H}_1$ - dijeljena kovar. matrica

$$\#param = \frac{n}{2}(n+1) + n \cdot K + K + 1 = 34$$

$K = 2$, $n = 6$

$\mathcal{H}_2$ - dijag. kovar. matrica

$$\#param = n + n \cdot K + K + 1$$
$$= 19$$

$\mathcal{H}_3$ - izotropna Σ

$$\#param = K \cdot n + K =$$
$$= 13$$

b) Očekujemo da će najbolje generalizirati model $\mathcal{H}_2$

model $\mathcal{H}_1$

- modelirane korelacije svih značajki, međutim zbog velike korelacije značajki u primjeru može doći do loše kondicije matrice Σ^{-1}
- model sklon prenaučenosti

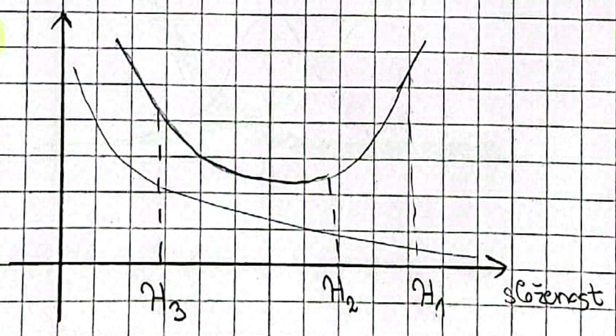
model $\mathcal{H}_3$

- sve varijance modela modelira kao jednake
  $\Rightarrow$ vodi na podnaučenost
- velike razlike u skalama

model $\mathcal{H}_2$

- ne modelira međusobne kolinearnosti $\rightarrow$ teže prenaučiti nego $\mathcal{H}_1$
- koristi normaliz. euklid. udaljenosti
  $\rightarrow$ neosjetljivost na skale

c)



d) U praksi bismo trebali provesti unakrsnu validaciju za odabir modela.

V16 - Bayesov klas. II

1.2.

a)

$P(x_1 = \text{Istra} \mid y=\text{ne}) = 0$
$P(x_1 = \text{Dalmacija} \mid y=\text{ne}) = 1/3$
$P(x_1 = \text{Kvarner} \mid y=\text{ne}) = 2/3$

$P(x_1 = \text{Istra} \mid y=\text{da}) = 1/2$
$P(x_1 = \text{Dalm.} \mid y=\text{da}) = 1/2$
$P(x_1 = \text{Kvarn.} \mid y=\text{da}) = 0$

$P(x_2 = \text{ne} \mid y=\text{ne}) = 1$
$P(x_2 = \text{da} \mid y=\text{ne}) = 0$

$P(x_2 = \text{ne} \mid y=\text{da}) = 1/4$
$P(x_2 = \text{da} \mid y=\text{da}) = 3/4$

$P(x_3 = \text{privatni} \mid y=\text{ne}) = 1/3$
$P(x_3 = \text{kamp} \mid y=\text{ne}) = 2/3$
$P(x_3 = \text{hotel} \mid y=\text{ne}) = 0$

$P(x_3 = \text{privatni} \mid y=\text{da}) = 1/2$
$P(x_3 = \text{kamp} \mid y=\text{da}) = 0$
$P(x_3 = \text{hotel} \mid y=\text{da}) = 1/2$

$P(x_4 = \text{auto} \mid y=\text{ne}) = 0$
$P(x_4 = \text{bus} \mid y=\text{ne}) = 2/3$
$P(x_4 = \text{avion} \mid y=\text{ne}) = 1/3$

$P(x_4 = \text{auto} \mid y=\text{da}) = 3/4$
$P(x_4 = \text{bus} \mid y=\text{da}) = 0$
$P(x_4 = \text{avion} \mid y=\text{da}) = 1/4$

$$h(\text{Istra}, \text{ne}, \text{kamp}, \text{bus}) = \underset{y}{\text{argmax}}\, P(y) \prod P(x_k \mid y)$$

$y=\text{ne}$: $h = \frac{3}{7} \cdot 0 \cdot 1 \cdot \frac{2}{3} \cdot \frac{2}{3} = 0$

$y=\text{da}$: $h = \frac{4}{7} \cdot \frac{1}{2} \cdot \frac{1}{4} \cdot 0 \cdot 0 = 0$

} ne možemo klasific.

$$h(\text{Dalm.}, \text{da}, \text{hotel}, \text{bus}) = \underset{y}{\text{argmax}}\, P(y) \prod P(x_k \mid y)$$

$y=\text{ne}$: $h = \frac{3}{7} \cdot \frac{1}{3} \cdot 0 \cdot 0 \cdot \frac{2}{3} = 0$

$y=\text{da}$: $h = \frac{4}{7} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot 0 = 0$

} ne možemo klasificirati

b)

$P(x_1 = \text{Istra} \mid y=\text{ne}) = 1/6$
$P(x_1 = \text{Dalmacija} \mid y=\text{ne}) = 2/6$
$P(x_1 = \text{Kvarner} \mid y=\text{ne}) = 3/6$

$P(x_1 = \text{Istra} \mid y=\text{da}) = 3/7$
$P(x_1 = \text{Dalm.} \mid y=\text{da}) = 3/7$
$P(x_1 = \text{Kvarner} \mid y=\text{da}) = 1/7$

$P(x_2 = \text{ne} \mid y=\text{ne}) = 4/5$
$P(x_2 = \text{da} \mid y=\text{ne}) = 1/5$

$P(x_2 = \text{ne} \mid y=\text{da}) = 2/6$
$P(x_2 = \text{da} \mid y=\text{da}) = 4/6$

$P(x_3 = \text{priv.} \mid y=\text{ne}) = 2/6$
$P(x_3 = \text{kamp} \mid y=\text{ne}) = 3/6$
$P(x_3 = \text{hotel} \mid y=\text{ne}) = 1/6$

$P(x_3 = \text{priv.} \mid y=\text{da}) = 3/7$
$P(x_3 = \text{kamp} \mid y=\text{da}) = 1/7$
$P(x_3 = \text{hotel} \mid y=\text{da}) = 3/7$

$P(x_4 = \text{auto} \mid y=\text{ne}) = 1/6$
$P(x_4 = \text{bus} \mid y=\text{ne}) = 3/6$
$P(x_4 = \text{avion} \mid y=\text{ne}) = 2/6$

$P(x_4 = \text{auto} \mid y=\text{da}) = 4/7$
$P(x_4 = \text{bus} \mid y=\text{da}) = 1/7$
$P(x_4 = \text{avion} \mid y=\text{da}) = 2/7$

$x = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$

$y=\text{ne}$: $h(x) = 3/7 \cdot 1/6 \cdot 4/5 \cdot 3/6 \cdot 3/6 = 1/70$

$y=\text{da}$: $h(x) = 4/7 \cdot 3/7 \cdot 2/6 \cdot 1/7 \cdot 1/7 = 4/2401$

} $h(x) = \text{ne}$

$x = (\text{Dalmacija}, \text{da}, \text{hotel}, \text{bus})$

$y=\text{ne}$: $h(x) = 3/7 \cdot 2/6 \cdot 1/5 \cdot 1/6 \cdot 3/6 = 1/420$

$y=\text{da}$: $h(x) = 4/7 \cdot 3/7 \cdot 4/6 \cdot 3/7 \cdot 1/7 = 96/2401$

} $h(x) = \text{da}$