Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

TDA231 - Homework 1

Theoretical Problems

Problem 1.1

TDA231 - Homework 1 - Elias Hult Pappas & Heiko Joshua Jungen

Problem 1.1

$N(\mu, \sigma^2 I)$

$$L = p(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \, |\sigma^2 I|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x-\mu)^T (\sigma^2 I)^{-1} (x-\mu)\right)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \cdot \sigma^n} \cdot \exp\left(-\frac{1}{2}(x-\mu)^T \left(\frac{1}{\sigma^2} I\right)(x-\mu)\right)$$

$$\ln L = \ln(1) - \left(\ln(2\pi)^{\frac{n}{2}} + \ln(\sigma^n)\right) + \left(-\frac{1}{2}(x-\mu)^T\left(\frac{1}{\sigma^2} I\right)(x-\mu)\right)$$

$$= -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \frac{1}{2\sigma^2}(x-\mu)^T I (x-\mu)$$

$$\frac{d \ln L}{d\sigma} = 0 - \frac{n}{\sigma} + \frac{1}{\sigma^3}(x-\mu)^T I (x-\mu) = 0$$

$$\Leftrightarrow \quad \frac{1}{\sigma^2}(x-\mu)^T I (x-\mu) = n$$

$$\Leftrightarrow \quad \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_i)^2 = \sigma^2$$

$$\Rightarrow \quad \hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_i)^2}$$

Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

Problem 1.2

Problem 1.2 TDA 231 - Homework 1 Elias Hult Pappas & Heiko Joshua Jungen

a) $\mathcal{D} = x_1, \dots, x_n$

$$p(\sigma^2 = s \mid \mathcal{D}; \alpha, \beta) \propto \prod_{i=1}^{n} p(X = x_i \mid \sigma^2) \cdot p(\sigma^2 = s \mid \alpha, \beta)$$

$$= \frac{1}{(2\pi s)^n} \exp\left(\frac{\sum_{i=1}^{n} (x_i - \mu)^T (x_i - \mu)}{-2s}\right) \cdot \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \cdot s^{-\alpha_0 - 1} \exp\left(-\frac{\beta_0}{s}\right)$$

$$= \frac{\beta_0^{\alpha_0}}{(2\pi)^n \Gamma(\alpha_0)} \cdot s^{-\alpha_0 - 1 - n} \exp\left(\frac{\sum_{i=1}^{n} (x_i - \mu)^T (x_i - \mu) + 2\beta_0}{-2s}\right)$$

$$p(\sigma^2 = s \mid \mathcal{D}; \alpha_n, \beta_n) = \frac{\beta_n^{\alpha_n}}{\Gamma(\alpha_n)} \cdot s^{-\alpha_n - 1} \exp\left(-\frac{\beta_n}{s}\right)$$

$$\Rightarrow \quad -\alpha_n - 1 = -\alpha_0 - 1 - n$$

$$\Leftrightarrow \quad \underline{\underline{\alpha_n}} = \alpha_0 + n$$

$$-\frac{1}{2}\left(\sum_{i=1}^{n} (x_i - \mu)^T (x_i - \mu) + 2\beta_0\right) = -\beta_n$$

$$\Leftrightarrow \quad \underline{\underline{\beta_n}} = \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^T (x_i - \mu) + \beta_0$$

b)

$$BF = \frac{p(D \mid M_A)}{p(D \mid M_B)} = \frac{\int p(\sigma^2 = s \mid \mathcal{D}; \alpha_{n,A}, \beta_{n,A}) \, ds}{\int p(\sigma^2 = s \mid \mathcal{D}; \alpha_{n,B}, \beta_{n,B}) \, ds}$$

Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

Problem 1.2 TDA231 - Homework 2 Elias Hult Pappas & Heiko Joshua Jungen

c)

$$\hat{s}_{MAP} := \frac{d}{ds} \operatorname{argmax} \log p(\sigma^2 = s \mid D; \alpha_n, \beta_n) = 0$$

$$\log p(\sigma^2 = s \mid D; \alpha_n, \beta_n) = \log\left(\frac{\beta_n^{\alpha_n}}{\Gamma(\alpha_n)}\right) + \log\left(s^{-\alpha_n - 1}\right) + \left(-\frac{\beta_n}{s}\right)$$

$$\frac{d}{ds} \operatorname{argmax} \log p(\sigma^2 = s \mid D; \alpha_n, \beta_n) = 0$$

$$\Leftrightarrow \frac{-\alpha_n - 1}{s} + \frac{\beta_n}{s^2} = 0$$

$$\Leftrightarrow s = \frac{\beta_n}{\alpha_n + 1}$$

$$\hat{s}_{MAP} = \frac{\beta_n}{\alpha_n + 1}$$

$$P(M \mid D) = \frac{P(D \mid M)\, P(M)}{P(D)}$$

$$\Leftrightarrow P(D \mid M) = \frac{P(M \mid D)\, P(D)}{P(M)}$$

$$\Rightarrow BF = \frac{P(D \mid M_A)}{P(D \mid M_B)} = \frac{P(M_A \mid D)\, P(M_B)}{P(M_B \mid D)\, P(M_A)}$$

$$P(M_A) = P(M_B) = \frac{1}{2}$$

$$\Rightarrow BF = \frac{P(M_A \mid D)}{P(M_B \mid D)}$$

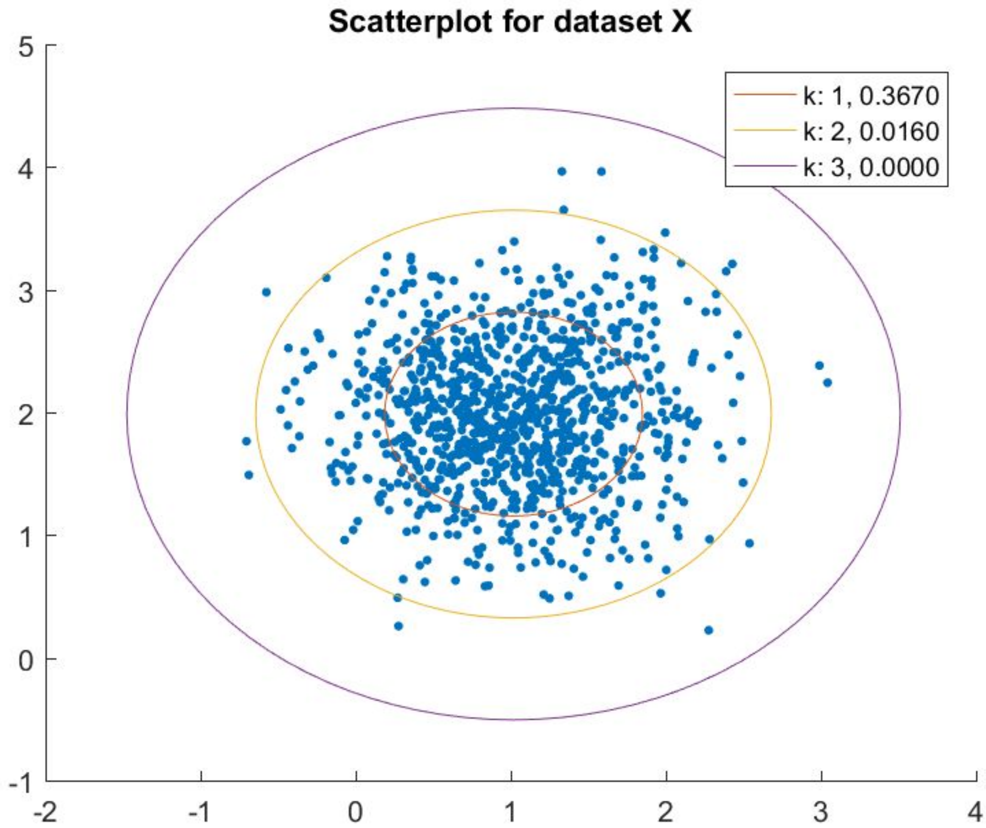
Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

Matlab problems

Problem 2.1

$$\mu = \frac{\sum x}{n}$$

$$\sigma = \frac{1}{n}\sqrt{\sum_{i=1}^{n} (x_i - \mu_i) \cdot (x_i - \mu_i)^T}$$

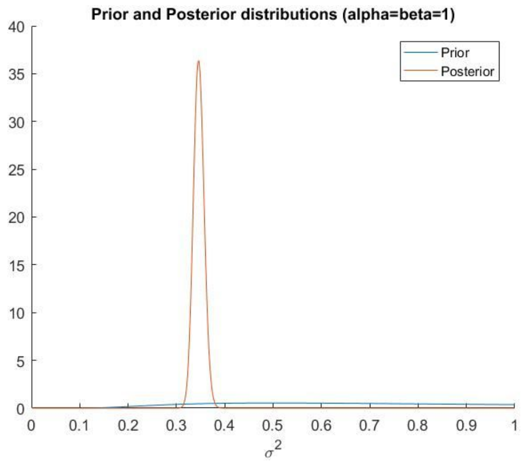


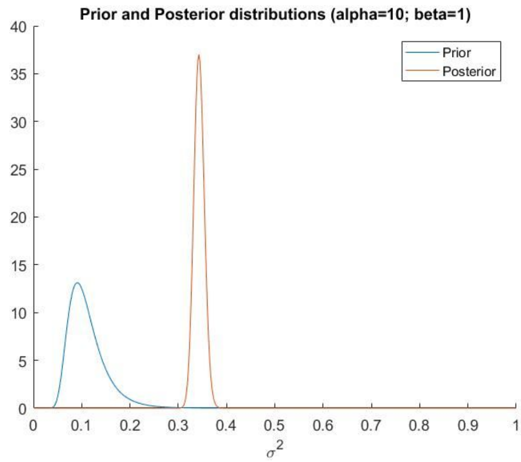
Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

Problem 2.2

a) Prior and posterior distributions

(Left plot) The blue line shows the prior distribution. It is very flat, which means that the variance is somewhere between 0.2 and 1. The posterior distribution, shown by the red line, has a peak between 0.3 and 0.4, where the variance is most likely to be.





(Right plot) Adapting the alpha and beta value results in a noticeably different prior distribution. Nevertheless, the posterior distribution changes only little. The prior suggests that the variance is roughly 0.05 and 0.25. That influences the posterior distribution by shifting its peak slightly to the left.

b) Map estimates for σ^2 = s

The analytical model for the map estimate has been derived in Task 1.2 c). The formulas for α_n and β_n have been derived in 1.2 a).

$$\alpha_n = \alpha_0 + n$$

$$\beta_n = 0.5 * \sum_{i=1}^{n} [(x_i - \mu)^T (x_i - \mu)] + \beta_0 \; ; \mu = empirical\ mean$$

$$s_{MAP} = \frac{\beta_n}{\alpha_n + 1}$$

Model	s_{MAP}
Model A (alpha=beta=1)	0.3457
Model B (alpha=10; beta=1)	0.3426

Heiko Joshua Jungen jungen@student.chalmers.se 950418-T075
Elias Hult Pappas pappas@student.chalmers.se 940711-2417

c) Model selection with bayes factor

Using $\mu = empirical\ mean$, $\sigma^2 = s_{MAP}$ as well as the respective α and β values the posterior distribution of the models can be calculated. With the assumption of $P(M_A) = P(M_B) = 0.5$, their ratio is equal to the bayes factor as shown in task 1.2 c).

$$BF = \frac{P(M_A|D)}{P(M_B|D)} = 0.9867 < 1$$

Since the bayes factor BF is less than 1, Model B is the better model.