

# Statistical learning theory

Joris LIMONIER

October 11, 2022

## Contents

<b>1 Proposed exercise lecture 02</b>	<b>1</b>
1.1 Home Exercise . . . . .	1
<b>2 Proposed exercise lecture 05</b>	<b>2</b>
2.1 In-class Exercise . . . . .	2

## 1 Proposed exercise lecture 02

### 1.1 Home Exercise

**Demonstrate that the MAP (Maximum A Posteriori) classification rule is quadratic with respect to  $x^*$ .** We have the following:

$$\hat{y}^* = \arg \max_{k=1,\dots,k} \mathbb{P}(Y = k|X = x^*) \quad (1)$$

We compute the probability:

$$\begin{aligned} \arg \max_{k=1,\dots,k} \mathbb{P}(Y = k|X = x^*) &\stackrel{\text{Bayes}}{=} \frac{\mathbb{P}(Y = k)\mathbb{P}(X = x^*|Y = k)}{\mathbb{P}(X = x^*)} \\ &\propto \mathbb{P}(Y = k)\mathbb{P}(X = x^*|Y = k) \\ &\propto \hat{\Pi}_k \mathcal{N}(x^*, \hat{\mu}_k, \hat{\Sigma}_k) \end{aligned}$$

$$\begin{aligned}
\hat{y}^* &= \arg \max_{k=1, \dots, K} \hat{\Pi}_k \mathcal{N}(x^*, \hat{\mu}_k, \hat{\Sigma}_k) \\
&= \arg \max_{k=1, \dots, K} \log [\hat{\Pi}_k] + \log [\mathcal{N}(x^*, \hat{\mu}_k, \hat{\Sigma}_k)] \\
&= \arg \max_{k=1, \dots, K} \log [\hat{\Pi}_k] + \log \left[ \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (x^* - \mu)^\top \Sigma^{-1} (x^* - \mu)} \right] \\
&= \arg \max_{k=1, \dots, K} \underbrace{\log [\hat{\Pi}_k]}_{\text{Constant w.r.t. } x^*} + \underbrace{\left[ -\frac{1}{2} (x^* - \mu)^\top \Sigma^{-1} (x^* - \mu) \right]}_{\text{Quadratic in } x^*} \\
&\quad - \underbrace{\log [(2\pi)^{N/2} |\Sigma|^{1/2}]}_{\text{Constant w.r.t. } x^*}
\end{aligned}$$

The second term is Quadratic in  $x^*$ , the others are constant w.r.t.  $x^*$ , which completes the proof.

## 2 Proposed exercise lecture 05

### 2.1 In-class Exercise

Compute the derivative with respect to the means.

$$\begin{aligned}
\nabla_{\mu_k} \mathcal{L}(\theta) &= \nabla_{\mu_k} \sum_{n=1}^N \log \left[ \sum_{k=1}^K \pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k) \right] \\
&= \sum_{n=1}^N \frac{\pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k)}{\sum_{k=1}^K \pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k)} \\
&= \sum_{n=1}^N \frac{\pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k)}{\sum_{k=1}^K \pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k)} \Sigma^{-1} (x_n - \mu_k)
\end{aligned}$$