

Exercices 3

TABLE DES MATIÈRES

1	Bayes estimator and Bayes risk	1
2	Logistic regression	1
3	OLS	2

1 BAYES ESTIMATOR AND BAYES RISK

Consider the following joint random variable (X, Y) .

- $\mathcal{X} = \{0, 1, 2\}$
- $\mathcal{Y} = \{0, 1\}$.
- X follows a uniform law on \mathcal{X} .
-

$$Y = \begin{cases} B(1/5) & \text{if } X = 0 \\ B(3/4) & \text{if } X = 1 \\ B(2/3) & \text{if } X = 2 \end{cases}$$

With $B(p)$ a Bernoulli law with parameter p .

Compute the Bayes estimator and the bayes risk.

2 LOGISTIC REGRESSION

Summary of the setting : in the context of binary classification, we consider the following setting.

- $\mathcal{X} = \mathbb{R}^d$
- $\mathcal{Y} = \{0, 1\}$ (sometimes $\mathcal{Y} = \{-1, 1\}$)
- $\ell_{0-1}(y, z) = 1_{y \neq z}$ ("0-1" loss)

Note that we can extend these definitions to non-binary classification. We would like a predictor that minimizes the binary loss.

Definition 1. Binary loss function

$$\hat{R}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{y_i \neq \hat{f}(x_i)}$$

However, as we have seen in the class, it is hard to minimize the binary loss as it is neither differentiable nor convex in θ . We can replace it by a **convex, differentiable surrogate loss (substitut convexe)**. Several possibilities exist instead of using $\mathbb{1}_{y_i \neq \hat{f}(x_i)}$ as ℓ (binary loss). The **logistic loss** is one of them

Definition 2. Logistic loss

$$l(\hat{y}, y) = y \log(1 + e^{-\hat{y}}) + (1 - y) \log(1 + e^{\hat{y}}) \quad (1)$$

We can define the corresponding empirical risk.

$$R_n(\theta) = \frac{1}{n} \sum_{i=1}^n l(X_i^T \theta, Y_i) \quad (2)$$

Definition 3. Logistic regression estimator

If l is the logistic loss, it is defined as

$$\hat{\theta}_{\text{logit}} = \arg \min_{\theta \in \mathbb{R}^d} R_n(\theta)$$

Compute the gradient of $R_n(\theta)$, $\nabla_{\theta} R_n$.

3 OLS

Show the first part of proposition 15 in FTML.pdf (Risk decomposition for OLS).

$$R_X(\theta) - R_X(\theta^*) = \|\theta - \theta^*\|_{\Sigma}^2$$