

---

**TRAVAUX DIRIGÉS N° 4 : Convexification du risque**


---

Stéphan CLÉMENÇON <stephan.clemencon@telecom-paristech.fr>

Ekhine IRUROZKI <irurozki@telecom-paris.fr>

**EXERCICE 1.** On se place dans le cadre de la classification binaire : on considère un descripteur aléatoire  $X$  de loi  $P_X$  à valeurs dans  $\mathbb{R}^d$  ( $d \in \mathbb{N}^*$ ) muni de sa tribu des Boréliens, et un label aléatoire  $Y$  valant  $-1$  ou  $1$ .

La fonction de régression (probabilité a posteriori) est notée  $\eta : x \in \mathbb{R}^d \mapsto \mathbb{P}(Y = 1 \mid X = x) \in [0, 1]$ . On suppose que  $0 < \eta(X) < 1$  presque-sûrement.

Soit  $\mathcal{G} := \{g : \mathbb{R}^d \rightarrow \{-1, 1\}\}$  l'ensemble des classifieurs adaptés à ce contexte. L'erreur de classification est définie comme l'application  $L : g \in \mathcal{G} \mapsto \mathbb{P}(Y \neq g(X)) \in [0, 1]$  et on note  $L^* := \inf_{g \in \mathcal{G}} L(g)$ .

On pose  $\mathcal{F} := \{f : \mathbb{R}^d \rightarrow \mathbb{R}\}$  puis  $\text{sgn} : a \in \mathbb{R} \mapsto 2\mathbb{1}_{\{a \geq 0\}} - 1 \in \{-1, 1\}$ . Dans cet exercice on s'intéresse spécifiquement à l'ensemble de classifieurs  $\mathcal{G}_{\mathcal{F}} := \{\text{sgn} \circ f : f \in \mathcal{F}\}$ .

Soit enfin  $\phi : \mathbb{R} \rightarrow \mathbb{R}_+$  une fonction dérivable, strictement convexe, croissante, satisfaisant  $\lim_{x \rightarrow -\infty} \phi(x) = 0$  et  $\phi(0) = 1$ . Pour tout  $f \in \mathcal{F}$  (et donc tout classifieur de  $\mathcal{G}_{\mathcal{F}}$ ) on considère la version convexifiée du risque de classification :

$$A(f) := \mathbb{E} \left[ \phi(-Yf(X)) \right].$$

- 1) Pour tout  $u \in [0, 1]$  on définit  $h_u : a \in \mathbb{R} \mapsto u\phi(-a) + (1-u)\phi(a)$ . Montrer que  $\min_{f \in \mathcal{F}} A(f)$  est atteint en

$$f^* : x \in \mathbb{R}^d \mapsto \arg \min_{a \in \mathbb{R}} h_{\eta(x)}(a). \quad (1)$$

Cette fonction est-elle bien définie pour tout  $x \in \mathbb{R}^d$  ?

- 2) Dériver la fonction  $h_{\eta(x)}$  pour  $x \in \mathbb{R}^d$  quelconque, puis vérifier que  $\text{sgn} \circ f^*$  coïncide avec classifieur de Bayes.

- 3) (Lemme de Zhang) On pose  $H : u \in ]0, 1[ \mapsto \min_{a \in \mathbb{R}} h_u(a)$  et on suppose qu'il existe des constantes  $s > 1$  et  $c > 0$  telles que

$$\forall u \in ]0, 1[, \quad \left| \frac{1}{2} - u \right|^s \leq c^s (1 - H(u)). \quad (2)$$

- a) Soit  $f \in \mathcal{F}$ . Montrer que  $L(\text{sgn} \circ f) - L^* \leq 2 \left( \mathbb{E} \left[ |\eta(X) - 1/2|^s \mathbb{1}_{\{(2\eta(X)-1)f(X) < 0\}} \right] \right)^{1/s}$ .
- b) On pose  $A^* := \inf_{f \in \mathcal{F}} A(f)$ . D  duire de la question pr  c  dente que pour toute fonction  $f \in \mathcal{F}$ ,

$$L(\text{sgn} \circ f) - L^* \leq 2c (A(f) - A^*)^{1/s}. \quad (3)$$

- c) Que vaut  $H$  dans le cas o    $\phi = \exp$ ? Pour quelles constantes  $s$  et  $c$  la condition (2) est-elle v  rifi  e?

**Solution.**

- 1) Soit  $f \in \mathcal{F}$ . On a

$$\begin{aligned} A(f) &:= \mathbb{E}[\phi(-Yf(X))] \\ &= \mathbb{E}[\phi(-f(X))\mathbb{1}_{\{Y=1\}} + \phi(f(X))\mathbb{1}_{\{Y=-1\}}] \\ &= \mathbb{E}[\phi(-f(X))\eta(X) + \phi(f(X))(1 - \eta(X))] \quad (\text{esp  rance totale}) \\ &= \mathbb{E}[h_{\eta(X)}(f(X))] = \int_{\mathbb{R}^d} h_{\eta(x)}(f(x)) P_X(dx) \end{aligned}$$

Ainsi, par construction, si  $f^*$  est bien d  finie pour  $P_X$ -presque tout  $x$ , elle minimise effectivement le risque convexifi  .

V  rifions maintenant cette condition. Pour cela, commen  ons par rappeler les hypoth  ses faites sur la fonction  $\phi : \mathbb{R} \rightarrow \mathbb{R}^+$  :

- (i)  $\phi$  est d  rivable,
- (ii)  $\phi$  est strictement convexe,
- (iii)  $\phi$  est croissante,
- (iv)  $\lim_{x \rightarrow -\infty} \phi(x) = 0$ ,
- (v)  $\phi(0) = 1$ .

On pourra remarquer au passage que  $x \in \mathbb{R} \mapsto e^x$  est un exemple de fonction qui satisfait toutes ces hypoth  ses. On en d  duit un certain nombre de propri  t  s suppl  mentaires.

(vi)  $\phi$  est strictement croissante. **Preuve** :  $\phi$  étant strictement convexe sur  $\mathbb{R}$ , elle y est continue. Elle y est en outre croissante. Par conséquent, si elle n'était pas strictement croissante, elle admettrait au moins un palier, ce qui contredirait la stricte convexité.

(vii)  $\phi > 0$ . **Preuve** : Conséquence immédiate des hypothèses (iv) et (vi).

(viii)  $\phi'(0) > 0$ . **Preuve** :  $\phi$  étant (strictement) convexe et dérivable, elle vérifie pour tout  $x \in \mathbb{R}^*$  l'inégalité  $\phi(x) > \phi(0) + \phi'(0)x$  (i.e., la courbe représentative de  $\phi$  est au-dessus de toutes ses tangentes, en particulier de celle en 0). Comme  $\phi$  est strictement croissante, sa dérivée est positive ou nulle. Par l'absurde, si on avait  $\phi'(0) = 0$ , alors pour tout  $x < 0$ , on obtiendrait  $\phi(x) > \phi(0)$ , ce qui contredirait la croissance de  $\phi$ . On en conclut que  $\phi'(0) > 0$ .

(ix)  $\lim_{x \rightarrow +\infty} \phi(x) = +\infty$ . **Preuve** : Comme précédemment, par (i) et (ii), on a pour tout  $x \in \mathbb{R}^*$  l'inégalité  $\phi(x) > \phi(0) + \phi'(0)x$ . En utilisant (viii), on obtient alors directement que  $\lim_{x \rightarrow +\infty} \phi(x) = +\infty$ .

Prenons maintenant  $u \in [0, 1]$  et étudions la fonction  $h_u$ .

– Cas où  $u \in \{0, 1\}$ . Les propriétés (vi) et (vii) impliquent que  $\phi$  n'admet pas de minimum, mais seulement un infimum en  $-\infty$ . Il vient que lorsque  $u \in \{0, 1\}$ , les fonctions  $h_0 : a \in \mathbb{R} \mapsto \phi(a)$  et  $h_1 : a \in \mathbb{R} \mapsto \phi(-a)$  n'admettent pas de minimum. Ainsi, la fonction  $f^*$  n'est pas définie sur l'ensemble  $\mathcal{X}_0 := \{x \in \mathbb{R}^d : \eta(x) \in \{0, 1\}\}$ .

– Cas où  $u \in ]0, 1[$ . Commençons par remarquer que

$$(x) \lim_{a \rightarrow +\infty} h_u(a) = \lim_{a \rightarrow -\infty} h_u(a) = +\infty \quad (\text{d'après (iv) et (ix)}),$$

$$(xi) h_u(0) = 1, \quad (\text{d'après (v)})$$

(xii)  $h_u$  est continue car  $\phi$ , donc  $h_u$  est dérivable. Les points précédents impliquent qu'il existe  $\alpha > 0$  tel que pour tout  $a \in ]-\infty, -\alpha[ \cup ]\alpha, +\infty[$ , on a  $h_u(a) > 1 = h_u(0)$ . Le point (xii) dit alors que  $h_u$  est continue sur  $[-\alpha, \alpha]$ . Elle y est donc bornée et atteint ses bornes (théorème des valeurs extrêmes), i.e. elle y admet au moins un minimum  $m_u$ . Or,  $\phi$  étant strictement convexe,  $h_u$  l'est aussi, et  $m_u$  est unique sur cet intervalle. Par construction,  $m_u \leq 1 = h_u(0)$ , donc  $m_u$  est même un minimum global, sur tout  $\mathbb{R}$ , de  $h_u$ .

Ainsi,  $f^*$  est bien définie sur l'ensemble  $\mathcal{X}_1 := \{x \in \mathbb{R}^d : \eta(x) \in ]0, 1[ \}$ .

Rappelons-nous finalement que par hypothèse,  $\eta(X) \in ]0, 1[$   $P_X$ -p.s. En d'autres termes,  $P_X(\mathcal{X}_0) = 0$  et  $P_X(\mathcal{X}_1) = 1$ . On en conclut que  $f^*$  est bien définie pour  $P_X$ -presque tout  $x \in \mathbb{R}^d$ .

2) Soit  $x \in \mathbb{R}^d$ . Comme  $\phi$  est dérivable (i), la fonction  $h_{\eta(x)} : \mathbb{R} \rightarrow \mathbb{R}^+$  l'est aussi, et pour tout  $a \in \mathbb{R}$ , on a

$$h'_{\eta(x)}(a) = -\eta(x)\phi'(-a) + (1 - \eta(x))\phi'(a).$$

Vérifions maintenant que la fonction  $\text{sgn} \circ f^*$  coïncide avec le classifieur de Bayes. Prenons pour cela  $x \in \mathcal{X}_1$ . Nous avons vu qu'alors  $h_{\eta(x)}$  est dérivable, strictement convexe, admettant un minimum  $m$  en  $f^*(x)$ . Par conséquent,  $h'_{\eta(x)}(a) = 0$  ssi  $a = f^*(x)$ . Or  $h'_{\eta(x)}(a) = 0$  signifie

$$\frac{\eta(x)}{1 - \eta(x)} = \frac{\phi'(a)}{\phi'(-a)}.$$

(comme on a pris  $x \in \mathcal{X}_1$ , on a bien  $1 - \eta(x) \neq 0$ , puisque  $\phi$  est strictement convexe et strictement croissante, sa dérivée est positive ou nulle et strictement croissante sur tout  $\mathbb{R}$ , donc strictement positive). On a donc l'égalité des rapports

$$\frac{\eta(x)}{1 - \eta(x)} = \frac{\phi'(f^*(x))}{\phi'(-f^*(x))}.$$

d'où

$$\eta(x) > \frac{1}{2} \Leftrightarrow \frac{\eta(x)}{1 - \eta(x)} > 1 \Leftrightarrow \frac{\phi'(f^*(x))}{\phi'(-f^*(x))} > 1 \Leftrightarrow \phi(f^*(x)) > \phi(-f^*(x)) \Leftrightarrow f^*(x) > 0.$$

( $\phi'$  strictement croissante car  $\phi$  strictement convexe)

Finalement, on a bien

$$\text{sgn} \circ f^*(x) := 2\mathbb{1}_{\{f^*(x) > 0\}} - 1 = 2\mathbb{1}_{\{\eta(x) > \frac{1}{2}\}} - 1,$$

pour  $P_X$ -presque tout  $x \in \mathbb{R}^d$ . La fonction  $\text{sgn} \circ f^*$  coïncide effectivement avec le classifieur de Bayes.

- 3) (a) Soit  $f \in \mathcal{F}$ . Nous avons vu à l'exercice 2 du TD 2 que l'on pouvait écrire dans le cas présent

$$L(\text{sgn} \circ f) - L^* = 2\mathbb{E} \left[ \left| \eta(X) - \frac{1}{2} \right| \left| \mathbb{1}_{\{f(X) \leq 0\}} - \mathbb{1}_{\{\eta(X) \leq \frac{1}{2}\}} \right| \right].$$

Or on peut remarquer que pour tout  $x \in \mathbb{R}^d$ ,

$$\left| \mathbb{1}_{\{f(X) \leq 0\}} - \mathbb{1}_{\{\eta(X) \leq \frac{1}{2}\}} \right| = \begin{cases} 0 & \text{si } \text{sgn}(f(x)) = \text{sgn}(2\eta(x) - 1) \\ 1 & \text{sinon} \end{cases}$$

$$\leq \mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}}$$

d' ou

$$\begin{aligned}
L(\text{sgn} \circ f) - L^* &\leq 2\mathbb{E} \left[ \left| \eta(X) - \frac{1}{2} \right| \mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}} \right] \\
&= 2\mathbb{E} \left[ \left( \left| \eta(X) - \frac{1}{2} \right| \mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}} \right)^{\frac{s}{s}} \right] \\
&\leq \left( 2\mathbb{E} \left[ \left| \eta(X) - \frac{1}{2} \right|^s \mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}} \right] \right)^{\frac{1}{s}}
\end{aligned}$$

en utilisant l'inégalité de Jensen pour la fonction concave  $x \in \mathbb{R}_+^* \mapsto x^{\frac{1}{s}}$  qui, comme la fonction convexe intérieure  $x \in \mathbb{R}_+^* \mapsto x^s$ , peut être prolongée par continuité en 0 pour  $s > 1$  (ainsi toutes les écritures ci-dessus ont bien du sens).

(b) Par hypothèse, il existe des réels  $s > 1$  et  $c > 0$  tels que pour tout  $u \in ]0, 1[$  on a

$$\left| \frac{1}{2} - u \right|^s \leq c^s(1 - H(u)).$$

Par conséquent, pour tout  $x \in \mathcal{X}_1$  on a

$$\left| \eta(x) - \frac{1}{2} \right|^s \leq c^s(1 - H(\eta(x))).$$

avec  $H(\eta(x)) := \min h_{\eta(x)}(a) = h_{\eta(x)}(f^*(x))$  d'après la question 1.

Soit maintenant  $f \in \mathcal{F}$ . En repartant du résultat de la question 3.a, cette dernière inégalité donne

$$\begin{aligned}
L(\text{sgn} \circ f) - L^* &\leq 2\mathbb{E}[c^s(1 - h_{\eta(X)}(f^*(X)))\mathbb{1}_{\{(2\eta(X)-1)f(X) \leq 0\}}]^{\frac{1}{s}} \\
&= 2c\mathbb{E}[(1 - h_{\eta(X)}(f^*(X)))\mathbb{1}_{\{(2\eta(X)-1)f(X) \leq 0\}}]^{\frac{1}{s}}
\end{aligned}$$

Puisque la fonction  $x \in \mathbb{R}_+^* \mapsto x^{\frac{1}{s}}$  prolongée par continuité en 0 est croissante pour  $s > 1$  et puisque  $c > 0$ , il nous reste à vérifier que  $A(f) - A^*$  majore cette dernière espérance.

Nous avons vu en question 1 que  $A(f) - A^* = \mathbb{E}[h_{\eta(X)}(f(X)) - h_{\eta(X)}(f^*(X))]$ . Prenons maintenant  $x \in \mathcal{X}_1$ .

— Si  $(2\eta(x) - 1)f(x) > 0$ , alors par définition de  $f^*$  on a

$$(1 - h_{\eta(x)}(f^*(x)))\mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}} = 0 \leq h_{\eta(x)}(f(x)) - h_{\eta(x)}(f^*(x)).$$

— Si  $(2\eta(x) - 1)f(x) \leq 0$ , alors

$$\begin{aligned}
h_{\eta(x)}(f(x)) &= \eta(x)\phi(-f(x)) + (1 - \eta(x))\phi(f(x)) \text{ (par définition de } h_{\eta(x)}) \\
&> \phi(-\eta(x)f(x) + (1 - \eta(x))f(x)) \text{ (}\phi \text{ strictement convexe)} \\
&= \phi(-f(x)(2\eta(x) - 1)) \\
&\geq \phi(0) \text{ (}\phi \text{ croissante et } -f(x)(2\eta(x) - 1) \geq 0) \\
&= 1, \text{ par (v)}
\end{aligned}$$

puis

$$\begin{aligned}
h_{\eta(x)}(f(x)) - h_{\eta(x)}(f^*(x)) &\geq 1 - h_{\eta(x)}(f^*(x)) \\
&= (1 - h_{\eta(x)}(f^*(x)))\mathbb{1}_{\{(2\eta(x)-1)f(x) \leq 0\}}.
\end{aligned}$$

Nous avons donc démontré que

$$h_{\eta(X)}(f(X)) - h_{\eta(X)}(f^*(X)) \geq (1 - h_{\eta(X)}(f^*(X)))\mathbb{1}_{\{(2\eta(X)-1)f(X) \leq 0\}}$$

$P_X$ -p.s. et l'inégalité reste donc vraie en passant à l'espérance des deux côtés. In fine, nous avons bien montré que

$$L(\text{sgn} \circ f) - L^* \leq 2c(A(f) - A^*)^{\frac{1}{s}}$$

(c) Soient maintenant  $\phi : x \in \mathbb{R} \mapsto e^x$  et  $u \in ]0, 1[$ . Alors pour tout  $a \in \mathbb{R}$ , on a  $\phi(a) = ue^{-a} + (1 - u)e^a$ , qui atteint son minimum en  $a_u \in \mathbb{R}$  tel que

$$\frac{u}{1 - u} = \frac{e^{a_u}}{e^{-a_u}} = e^{2a_u} \quad \text{i.e., en} \quad a_u = \frac{1}{2} \ln \left( \frac{u}{1 - u} \right),$$

et

$$h_u(a_u) = u\sqrt{\frac{1 - u}{u}} + (1 - u)\sqrt{\frac{u}{1 - u}} = 2\sqrt{u}\sqrt{1 - u}$$

La fonction  $H$  s'écrit alors  $H : u \in ]0, 1[ \mapsto 2\sqrt{u}\sqrt{1 - u}$ . Elle est dérivable, de dérivée  $H'(u) :$

$$u \in ]0, 1[ \mapsto \frac{1-2u}{\sqrt{u(1-u)}} \text{ et } H'(u) = 0 \Leftrightarrow H(u) = 1 \Leftrightarrow u = \frac{1}{2} \text{ for } u \in ]0, 1[.$$

On remarque alors que pour tout  $u \in ]0, 1[$ , on a  $\sqrt{u}\sqrt{1 - u} \leq \frac{1}{2}$ , d'où

$$\begin{aligned}
1 - H(u) &= 1 - 2\sqrt{u}\sqrt{1-u} = 2\left(\frac{1}{2} - \sqrt{u}\sqrt{1-u}\right) \\
&\geq 2\left(\frac{1}{2} - \sqrt{u}\sqrt{1-u}\right)\left(\frac{1}{2} + \sqrt{u}\sqrt{1-u}\right) \quad (\text{la dernière parenthèse est } \leq 1) \\
&\geq 2\left(\frac{1}{4} - u(1-u)\right) = 2\left(\frac{1}{4} + u^2 - u\right) \\
&= 2\left(\frac{1}{2} - u\right)^2
\end{aligned}$$

qui donne finalement

$$\left|\frac{1}{2} - u\right|^2 \leq \frac{1}{2}(1 - H(u)).$$

Les constantes  $s = 2$  et  $c = \sqrt{\frac{1}{2}}$  permettent donc de vérifier la condition (2).