Ex 1

$\psi : \mathbb{R} \to \mathbb{R}$, $k$-Lipschitz

- $F_0 = \{ x \mapsto w^T x \mid \|w\|_1 \leq B \}$, $0 \in F_0$
- $F_1 = \{ x \mapsto \psi ( v + \sum_{j=1}^{m} w_j f_j(x) ) \mid |v| \leq V, \|w\|_1 \leq B, f_j \in F_0 \}$
- $F_p = \{ x \mapsto \psi ( v + \sum_{j=1}^{m} w_j f_j(x) ) \mid |v| \leq V, \|w\|_1 \leq B, f_j \in F_{p-1} \}$

1) $$\hat{R}_n(F_1) \leq \frac{k}{n} \mathbb{E}\left[ \sup_{\substack{|v| \leq V \\ \|w\|_1 \leq B \\ f_j \in F_0}} \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \left( v + \sum_{j=1}^{m} w_j f_j(x) \right) \right]$$

$$= k \left( B \hat{R}_n(\text{absconv}(F_0)) + \mathbb{E}\left[ \sup_{|v| \leq V} \frac{1}{n} \sum \varepsilon_i v \right] \right.$$

$$\leq k \left( B \hat{R}_n[\text{conv}(F_0 \cup (-F_0))] \right) + \frac{V}{n} \mathbb{E}\left[ |\sum \varepsilon_i| \right]$$

$$= k \left( B \hat{R}_n ( F_0 \cup (-F_0) ) \right) + \frac{V}{n} \mathbb{E}\left[ \left( (\sum \varepsilon_i)^2 \right)^{\frac{1}{2}} \right]$$

$$\leq k \left( B \hat{R}_n ( F_0 + (-F_0) ) \right) + \frac{V}{n} \mathbb{E}\left[ (\sum \varepsilon_i)^2 \right]^{1/2}$$

$$= k \left( 2B \hat{R}_n(F_0) + \frac{V}{n} \sqrt{n} \right)$$

$$= k \left( 2 B \hat{R}(F_0) + \frac{V}{\sqrt{n}} \right)$$

(si $F_0 = -F_0 \Rightarrow k \hat{R}_n(F_0)$)

$(F_0 \cup (-F_0) \subset F_0 + (-F_0))$

$\mathbb{E}\left[ (\sum \varepsilon_i)^2 \right]$

$= \mathbb{V}(\sum \varepsilon_i) + \mathbb{E}\left[ \sum \varepsilon_i \right]^2$

$= \sum_{i=1}^{n} 1 + \left( \sum \mathbb{E}[\varepsilon_i] \right)$

$= n$

2) Supposons $\mathcal{X} = \{ x \in \mathbb{R}^d \mid \|x\|_\infty \leq 1 \}$

$$n \hat{\mathcal{R}}_n(F_0) = \mathbb{E}\left[\sup_w \sum_{i=1}^n \varepsilon_i w^T x_i\right]$$

$$= \mathbb{E}\left[\sup_w w^T \sum_{i=1}^n \varepsilon_i x_i\right]$$

$$\leq \mathbb{E}\left[\sup_w \|w\|_1 \left\|\sum_{i=1}^n \varepsilon_i x_i\right\|_\infty\right] \quad \text{(Hölder)}$$

$$\leq B\, \mathbb{E}\left[\max_{1 \leq j \leq d} \left|\sum_{i=1}^n \varepsilon_i x_i^{(j)}\right|\right]$$

$$= B\, \mathbb{E}\left[\max_{a \in A} \left|\sum_{i=1}^n \varepsilon_i a_i\right|\right]$$

où $A = \left\{ \{x_1^{(1)}, \dots, x_n^{(1)}\}, \dots, \{x_1^{(d)}, \dots, x_n^{(d)}\} \right\}$

$$= B\, \mathbb{E}\left[\max_{a \in A \cup (-A)} \sum_{i=1}^n \varepsilon_i a_i\right]$$

$$\leq B\, r \sqrt{2 \ln 2d}$$

avec $r = \max_{a \in A \cup (-A)} \|a\|_2 = \max_{a \in A} \left(\sum_{i=1}^n a_i^2\right)^{1/2} \leq \left(\sum_{i=1}^n 1^2\right)^{1/2} = \sqrt{n}$

$$\Rightarrow \quad \hat{\mathcal{R}}_n(F_0) \leq \frac{B\sqrt{2 \ln 2d}}{\sqrt{n}}$$

3) Récurrence. $\left(\Psi(-u) = -\Psi(u)\right)$

Test $F_0$ :

$$\hat{\mathcal{R}}_n(F_0) \leq \frac{1}{\sqrt{n}} \; \mathcal{B}\sqrt{2\ln(2d)} = \frac{\mathcal{B}\sqrt{2\ln 2d}}{\sqrt{n}} \quad (\text{q}^{\circ}2)$$

Supposons $p$. Montrons $p+1$:

$$\hat{\mathcal{R}}_n(F_{p+1}) = \frac{1}{n}\mathbb{E}\Bigg[\sup_{\substack{|v|\leq V \\ \|w\|_1 \leq \mathcal{B} \\ f_j \in F_p}} \sum_{i=1}^{n} \varepsilon_i \Psi(v + w^T f(x_i))\Bigg]$$

$$= \frac{1}{n}\mathbb{E}\Bigg[\sup_{\substack{|v|\leq V \\ \|w\|_1 \leq \mathcal{B} \\ f_j \in F_p}} \sum_{i=1}^{n} \varepsilon_i \left(v + w^T f(x_i)\right)\Bigg] \quad (1\text{-Lipschitz})$$

$$\leq \frac{1}{n}\mathbb{E}\Bigg[\sup_{\substack{\|w\|_1 \leq \mathcal{B} \\ f_j \in F_p}} \sum_{i=1}^{n} \varepsilon_i w^T f(x_i)\Bigg] + \frac{V}{\sqrt{n}}$$

$$= \mathcal{B}\,\hat{\mathcal{R}}_n(\mathrm{absconv}(F_p)) + \frac{V}{\sqrt{n}}$$

$$= \mathcal{B}\,\hat{\mathcal{R}}_n(F_p \cup (-F_p)) + \frac{V}{\sqrt{n}}$$

Or, comme $F_0 = -F_0$ et que $\Psi$ est impaire, on peut facilement montrer par récurrence que, $\forall p \geq 0$, $F_p = -F_p$. On obtient donc :

$$\hat{\mathfrak{R}}_n(F_{p+1}) \leq B\,\mathfrak{R}_n(F_p) + \frac{V}{\sqrt{n}}$$

$$\leq \frac{B}{\sqrt{n}}\left(B^{p+1}\sqrt{2\ln 2d} + V\sum_{\ell=0}^{p-1} B^{\ell}\right) + \frac{V}{\sqrt{n}}$$

$$= \frac{1}{\sqrt{n}}\left(B^{p+2}\sqrt{2\ln 2d} + BV\sum_{\ell=0}^{p-1} B^{\ell} + V\right)$$

$$= \frac{1}{\sqrt{n}}\left(B^{p+2}\sqrt{2\ln 2d} + V\left(\sum_{\ell=1}^{p} B^{\ell} + 1\right)\right)$$

$$= \frac{1}{\sqrt{n}}\left(B^{p+2}\sqrt{2\ln 2d} + V\sum_{\ell=0}^{p} B^{\ell}\right)$$

$\square$

Ex 2)

$$\tilde{F}_\lambda = \left\{ f = \sum_{j=1}^{N} w_j g_j \ : N \in \mathbb{N}, g_j \in \mathcal{G}, w_j \in \mathbb{R}, \sum_{j=1}^{N} |w_j| \leq \lambda \right\}$$

1) $$\mathcal{R}_n(F_\lambda) = \mathbb{E}\left[ \sup_{\substack{N \in \mathbb{N} \\ g_j \in \mathcal{G} \\ \|w_j\|_1 \leq \lambda}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \sum_{j=1}^{N} w_j g_j(X) \right\} \middle| D_n \right]$$

$$= \lambda \mathbb{E}\left[ \sup_{\substack{N \in \mathbb{N} \\ g_j \in \mathcal{G} \\ \|w_j\| \leq 1}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \sum_{j=1}^{N} w_j g_j(X) \right\} \middle| D_n \right]$$

$$= \lambda \mathcal{R}_n(\mathrm{abs\,conv}(\mathcal{G}))$$

$$= \lambda \mathcal{R}_n(\mathrm{conv}(\mathcal{G} \cup (-\mathcal{G})))$$

$$= \lambda \mathcal{R}_n(\mathcal{G} \cup (-\mathcal{G}))$$

On suppose que $\exists g \in \mathcal{G}, \forall x \; g(x) = 0$.

Ainsi, $\mathcal{G} \cup (-\mathcal{G}) \subseteq \mathcal{G} + (-\mathcal{G})$.

On en déduit $\lambda \mathcal{R}_n(\mathcal{G} \cup (-\mathcal{G})) \leq \lambda (\mathcal{R}_n(\mathcal{G}) + \mathcal{R}_n(-\mathcal{G}))$

$$= \lambda 2 \mathcal{R}_n(\mathcal{G})$$

$$\leq 2\lambda \left( \frac{2 \log \gamma(\mathcal{G}, n)}{n} \right)^{1/2} \quad \text{(Corollary 2.4)}$$

$$\leq 2\lambda \left( \frac{2 V \log\left(\frac{en}{V}\right)}{n} \right)^{1/2} \quad \text{(Sauer's lemma)}$$

2) $\left.\begin{array}{l} \varphi(x) = \log_2 (1 + e^x) \\ A(f) = \mathbb{E}[\varphi(-Y f(X))] \end{array}\right\} \quad f^*(X) = \ln\left(\frac{\eta(X)}{1-\eta(X)}\right)$ (ex 7.4 TD1) p. 11

$$A(f^*) = \mathbb{E}\left[ \eta(X) \underbrace{\log_2\left(1 + \frac{1-\eta(X)}{\eta(X)}\right)}_{y=1} + (1-\eta(X)) \underbrace{\log_2\left(1 + \frac{\eta(X)}{1-\eta(X)}\right)}_{y=-1} \right]$$

$$= \mathbb{E}\left[ \eta(X) \log_2\left(\frac{1}{\eta(X)}\right) + (1-\eta(X)) \log_2\left(\frac{1}{1-\eta(X)}\right) \right]$$

$$= \mathbb{E}\left[ -\eta(X) \log_2(\eta(X)) - (1-\eta(X)) \log_2(1-\eta(X)) \right]$$

$$= \mathbb{E}\left[ -\eta(X) \log_2(\eta(X)) - \log(1-\eta(X)) + \eta(X) \log(1-\eta(X)) \right]$$

$$= \mathbb{E}\left[ \eta(X) \log_2\left(\frac{1-\eta(X)}{\eta(X)}\right) - \log(1-\eta(X)) \right]$$

$$\Rightarrow H(t) = t \log_2\left(\frac{1-t}{t}\right) - \log_2(1-t)$$

3) $H(t) = H\left(\frac{1}{2}\right) + H'\left(\frac{1}{2}\right)\left(t - \frac{1}{2}\right) + \frac{H''(v)}{2}\left(t - \frac{1}{2}\right)^2$

On a :

$$H\left(\frac{1}{2}\right) = 1 \quad , \quad H'(t) = \log_2\left(\frac{1-t}{t}\right) - \frac{1}{1-t} + \frac{1}{1-t}$$

$$= \log_2\left(\frac{1-t}{t}\right)$$

$$H'\left(\frac{1}{2}\right) = 0$$

$$H''(t) = \frac{1}{\ln 2}\left(\ln\left(\frac{1-t}{t}\right)\right)'$$

$$= \frac{1}{\ln 2}\left(\frac{t}{1-t} \times -\frac{1}{t^2}\right)$$

$$= \frac{-1}{\ln 2\, t(1-t)} .$$

Or, $\forall t \in [0,1]$, $t(1-t) \leq 1/4$. Ainsi,

$$H''(t) \leq -\frac{4}{\ln 2}$$

On en déduit

$$H(t) \leq 1 - \frac{2}{\ln 2}\left(t - \frac{1}{2}\right)^2$$

$$= 1 - \frac{2}{\ln 2}\left(\frac{1-t}{2}\right)^2 = 1 - \left(\sqrt{\frac{2}{\ln 2}}\ \frac{1-t}{2}\right)^2$$

$$\Rightarrow H(t) \leq 1 - \left(\frac{1-2t}{2c}\right)^2, \quad c \triangleq \sqrt{\frac{\ln 2}{2}}.$$

h) $\varphi$ est une fonction de coût et :

$$\inf_{\alpha \in \mathbb{R}} \left( \eta \varphi(-\alpha) + (1-\eta) \varphi(\alpha) \right)$$

$$= \eta \varphi(-f^*) + (1-\eta) \varphi(f^*)$$

On a donc

$$H(t) = \inf_{\alpha \in \mathbb{R}} \eta \varphi(-\alpha) + (1-\eta) \varphi(\alpha)$$

De plus, $\quad H(t) \leq 1 - \left(\frac{1-2t}{2c}\right)^2$

$$\Rightarrow H(t) \leq 1 - c^{-2} \left(\frac{1}{2} - t\right)^2$$

$$\Rightarrow \left(\frac{1}{2} - t\right)^2 \leq c^2 \left(1 - H(t)\right)$$

$$\Rightarrow \left(\frac{1}{2} - t^2\right) \leq \frac{\ln 2}{2} \left(1 - H(t)\right)$$

L'inégalité de Zhang permet de conclure :

$\forall f,\quad L(f) - L^* \leq 2c\,(A(f) - A^*)^{1/2}$

$$\Rightarrow \frac{L(f) - L^*}{(A(f) - A^*)^{1/2}} \leq \ln 2 .$$

5) On pose $F = \varphi \circ \tilde{F}$, $\tilde{F} = \{(x,y) \mapsto -y f(x) \mid f \in F_\lambda\}$

On peut appliquer le théorème 2.10 : avec probabilité $1-\delta$ :

$$\sup_{h \in F} \left( \mathbb{E}[h(z)] - \frac{1}{n}\sum h(z_i) \right) \leq 2 R_n(F) + c\sqrt{\frac{\log 1/\delta}{2n}}$$

On cherche c. Il suffit de borner toute fonction de F :

$$\forall h \in F,\ |h(z)| = |\varphi(-y f(x))| = \log_2\left(1 + e^{-y \sum w_i g_i(x)}\right)$$

$$\leq \log_2(1 + e^{\lambda}) \triangleq c$$

$$(g_i \in \{-1, 1\})$$

De plus, $\varphi'(x) = \underbrace{\frac{e^x}{1+e^x}}_{<1} \frac{1}{\ln 2} \leq \frac{1}{\ln 2} \Rightarrow \varphi(x)$ est $\frac{1}{\ln 2}$-Lipschitz.

On a :

$$\sup_{f \in F_\lambda} |\hat{A}_n(f) - A(f)| \leq \frac{2}{\ln 2} R_n(F_\lambda) + \log_2(1+e^{\lambda})\sqrt{\frac{\log 1/\delta}{2n}}$$

$$\Rightarrow \sup_{f\in F_\lambda} \left| \hat{A}_n(f) - A(f) \right| \leq \underbrace{\frac{4\lambda\sqrt{2}}{\ln 2}}_{c_1(\lambda)} \sqrt{\frac{V\ln\left(\frac{en}{V}\right)}{n}} + \underbrace{\log_2(1+e^\lambda)}_{c_2(\lambda)} \sqrt{\frac{\ln 1/\delta}{2n}} \quad (1)$$

q° 1

6) $L(\hat{f}_{n,\lambda}) - L^* \leq \ln 2 \left( A(\hat{f}_{n,\lambda}) - A^* \right)^{1/2}$ q° 4

$$\leq \ln 2 \left[ \left( A(\hat{f}_{n,\lambda}) - \inf_{f\in F_\lambda} A(f) \right)^{1/2} + \left( \inf_{f\in F_\lambda} A(f) - A^* \right)^{1/2} \right]$$

Or, $A(\hat{f}_{n\lambda}) - A(\bar{f})$

$A(f) - \hat{A}_n(f)$

$\hat{A}_n(\hat{f}_\lambda) - \hat{A}_n(\bar{f}) \leq 0$

$$= A(\hat{f}_{n\lambda}) - A(\bar{f}) - \hat{A}_n(\hat{f}_{\lambda n}) + \hat{A}_n(\hat{f}_{\lambda n})$$

$$= A(\hat{f}_{n\lambda}) - \hat{A}_n(\hat{f}_{\lambda n}) + \hat{A}_n(\bar{f}) - A(\bar{f}) + \underbrace{\hat{A}_n(\hat{f}_{\lambda n}) - \hat{A}_n(\bar{f})}_{\leq 0}$$

$$\leq 2 \sup_{f\in F_\lambda} \left| A(f) - \hat{A}_n(f) \right|$$

$(\bar{f}, \hat{f}_{\lambda n} \in F_\lambda)$

Cela nous donne la borne suivante :

$$L(\hat{f}_{n,\lambda}) - L^* \leq \ln 2 \left[ \left( 2 \sup_{f\in F_\lambda} \left| A(f) - \hat{A}_n(f) \right| \right)^{1/2} + \left( A(\bar{f}) - A^* \right)^{1/2} \right]$$

Utiliser (1) permet de conclure.

$$f^* = \underset{f}{\operatorname{argmin}}\ A(f)$$

$$= \underset{f}{\operatorname{argmin}}\ \mathbb{E}\left[\varphi(-Y f(X)) \mid X = x\right]$$

$$= \underset{f}{\operatorname{argmin}}\ \underbrace{\eta(x)\, \varphi(-f(x)) + (1-\eta(x))\, \varphi(f(x))}_{h_{\eta(x)} f(x)}$$

Ainsi,

$$f^*(x) = \underset{\alpha}{\operatorname{argmin}}\ h_{\eta(x)}(\alpha).$$

Comme $\varphi$ est convexe, $h_{\eta(x)}$ l'est aussi. Si $\varphi$ est dérivable, on a :

$$\alpha^* \in \underset{\alpha}{\operatorname{argmin}}\ h_{\eta(x)}(\alpha) \iff h'_{\eta(x)}(\alpha^*) = 0.$$

$$\iff \frac{\eta(x)}{1-\eta(x)} = \frac{\varphi'(\alpha^*)}{\varphi'(-\alpha^*)}.$$

Posons $\varphi(x) = \log_2(1 + e^x)$.

On a $\varphi'(x) = \dfrac{e^x}{e^x \ln(2) + \ln(2)}$. On cherche $\alpha^*$ tq

$$\frac{\eta(x)}{1-\eta(x)} = \frac{e^{\alpha^*}}{e^{\alpha^*}\ln 2 + \ln 2} \quad \frac{e^{-\alpha^*}\ln 2 + \ln 2}{e^{-\alpha^*}}$$

$$= \frac{e^{\alpha^*} e^{-\alpha^*} \ln 2 + e^{\alpha^*} \ln 2}{e^{-\alpha^*} e^{\alpha^*} \ln 2 + \ln 2 \, e^{-\alpha^*}}$$

$$= \frac{\ln 2 + e^{\alpha^*} \ln 2}{\ln 2 + e^{-\alpha^*} \ln 2}$$

$$= \frac{1 + e^{\alpha^*}}{1 + e^{-\alpha^*}}$$

$$= e^{\alpha^*} \quad \frac{1 + e^{\alpha^*}}{e^{\alpha^*} + 1}$$

$$= e^{\alpha^*}$$

$$\Leftrightarrow \quad \alpha^* = \ln\left(\frac{\eta(x)}{1-\eta(x)}\right)$$