Dantzig Selector : étude des performances (Tao-Candès) 4

⇨ preuve du résultat principal

Rappel du modèle

$$\boxed{\underset{(m\times 1)}{y} = \underset{(m\times p)}{X}\ \underset{(p\times 1)}{\beta} + \underset{(m,1)}{z}} \qquad \beta \text{ à estimer}$$

* β est S _ sparse $\quad \left(S = \text{card}\left\{ i ; \beta_i \neq 0 \right\}\right)$
* X vérifie U.U.P

Dantzig Selector : $\hat{\beta}$ solution de $\left[\min_{\tilde{\beta}\in \mathbb{R}^p} \|\tilde{\beta}\|_{\ell_1} \text{ tel que } \|{}^tX(y - X\tilde{\beta})\|_{\ell_\infty} \leq \dots \right.$

Théorème Soit $K = \sigma(1 + t^{-1})\sqrt{2 \operatorname{Log} p}$

Si β est S - sparse tel que $\delta_{2S} + \theta_{2S} < 1 - t$ alors

$$\|\hat{\beta} - \beta\|_{\ell_2}^2 \leq C \cdot (\operatorname{Log} p)\left[\sigma^2 + \sum_{i=1}^{p} \min(\beta_i^2, \sigma^2)\right]$$

avec probabilité $\geqslant 1 - \left(\sqrt{\pi \operatorname{Log} p}\right)^{-1}$

Notations * Sans perte de généralités : $\delta_{2S} = \delta \quad \theta_{2S} = \theta$

$\lambda = \sqrt{2 \operatorname{Log} p}$; $\sigma = 1$ et $|\beta_1| \geqslant |\beta_2| \geqslant \dots \geqslant |\beta_p|$

* S_0 plus petit entier vérifiant $\sum_{i=1}^{p} \min(\beta_i^2, \lambda^2) \leq S_0 \lambda^2$
* $T_0 = \{1, 2, \dots, S_0\}$
* $T_1 = \left\{ \text{position des } S_0 \text{ plus grandes valeurs } |\beta_j| \in T_0^c \right\}$
* $T_{01} = T_0 \cup T_1 \quad (\rightarrow |T_{01}| = 2S_0)$
* $\beta = \beta^{(1)} + \beta^{(2)} \in \mathbb{R}^p$

$$\left(\beta^{(1)}\right)_j = \beta_j \cdot \mathbb{1}_{\{1 \leqslant j \leqslant S_0\}}$$

$$\left(\beta^{(2)}\right)_j = \beta_j \cdot \mathbb{1}_{\{j > S_0\}}$$

$\boxed{\text{idée preuve}}$ $\|\hat\beta - \beta\|_2^2$ borne supérieure ?

(avec grande proba)

Préliminaires $S_0 \lambda^2 \leq \lambda^2 + \sum_i \min(\beta_i^2, \lambda^2)$

- étape 1 $\|\beta^{(2)}\|_{\ell_2}^2 \leq S_0 \lambda^2$
- étape 2 $\|\hat\beta - \beta^{(1)}\|_{\ell_2}^2 \leq C^2 . S_0 . \lambda^2$

$\rightarrow$ $\|\hat\beta - \beta\|_2^2 \leq 2\|\hat\beta - \beta^{(1)}\|_{\ell_2}^2 + 2\|\beta^{(2)}\|_{\ell_2}^2$

$\leq 2(C^2+1) S_0 \lambda^2 \leq K(\log p)\left[1 + \sum_i \min(\beta_i^2, \lambda^2)\right]$

Préliminaires : So vérifie

$* \; 0 \leq S_0 \leq S$ car β est S sparse

$* \; S_0 \lambda^2 \leq \lambda^2 + \sum_{i=1}^p \min(\beta_i^2, \lambda^2)$

Etape 1

$* \; \forall j > S_0 \; |\beta_j| < \lambda$

$$\|\beta^{(2)}\|_{\ell_2}^2 = \sum_{j > S_0} \beta_j^2 \leq \sum_{i=1}^p \min(\beta_i^2, \lambda^2) \leq S_0 . \lambda^2$$

Etape 2 nécessite l'utilisation de lemmes techniques

Lemme $\boxed{1}$ Pour tout $h \in \mathbb{R}^p$

a) $\|h\|_{\ell_2}^2 \leq 2\left(\|h\|_{\ell_2(T_{01})}^2 + S_0^{-1}\|h\|_{\ell_1(T_0^c)}^2\right)$

b) $\|h\|_{\ell_2(T_{01})} \leqslant (1-\delta)^{-1}\|{}^tX_{T_{01}} X h\|_{\ell_2} + \Theta(1-\delta)^{-1} S_0^{-\frac{1}{2}}\|h\|_{\ell_1(T_0^c)}$ $\boxtimes$

Lemme $\boxed{2}$ Pour tout $u \in \mathbb{R}^p$

$$\|Xu\|_{\ell_2} \leq \sqrt{1+\delta}\left[\|u\|_{\ell_2} + (2S)^{-\frac{1}{2}}\|u\|_{\ell_1}\right]$$

Lemme $\boxed{3}$ $\beta^{(2)}$ peut s'écrire $\beta^{(2)} = \beta' + \beta''$ avec

a) $\|\beta'\|_{\ell_2} \leq \dfrac{1+\delta}{1-\delta-\theta}\, \lambda\, S_0^{\frac{1}{2}}$

b) $\|\beta'\|_{\ell_1} \leq \dfrac{1+\delta}{1-\delta-\theta} . \lambda . S_0$

c) $\|{}^tXX\beta''\|_{\ell^\infty} \leq \dfrac{1-\delta^2}{1-\delta-\theta} . \lambda$

Notons $h = \hat{\beta} - \beta^{(1)}$. On va montrer que

(i) * $\|h\|_{\ell_1(T_0^c)} \leq \|h\|_{\ell_1(T_0)} + \dfrac{1+\delta}{1-\delta-\theta} . S_0 . \lambda$

(ii) * $\|h\|_{\ell_2(T_{01})} \leq C_1(\delta,\theta) . S_0^{\frac{1}{2}} . \lambda$ et $\|h\|_{\ell_1(T_0)} \leq C_2(\delta,\theta) S_0 . \lambda$

➡ par le lemme $\boxed{1}$ a), on pourra conclure $\|h\|_{\ell_2}^2 \leq K(\delta,\theta) . S_0 . \lambda^2$

(i) Considérons $\beta^{(2)} = \beta' + \beta''$ avec $\|\beta^{(2)}\|_{\ell_2}^2 \leq S_0 . \lambda^2$

$$ {}^tX\left(X(\beta^{(1)}+\beta') - y\right) = -{}^tXX\beta'' - {}^tXz$$

inégalité triangulaire

$\boxed{3}$ c) ➡ $\left[\|{}^tX(X(\beta^{(1)}+\beta')-y)\|_{\ell_\infty} \leq \left(1 + \dfrac{1-\delta^2}{1-\delta-\theta}\right).\lambda\right]$ I0

Comme $(1-\delta-\theta)^{-1} \leq t^{-1}$, $\beta^{(1)} + \beta'$ vérifie contrainte [C], donc

$$\|\hat{\beta}\|_{\ell_1} \leq \|\beta^{(1)} + \beta'\|_{\ell_1} \leq \|\beta^{(1)}\|_{\ell_1} + \frac{1+\delta}{1-\delta-\theta} . S_0 . \lambda$$

(Inégalité triangulaire) $\|\hat{\beta}\|_{\ell_1} \geq \|\beta^{(1)}\|_{\ell_1} - \|h\|_{\ell_1(T_0)} + \|h\|_{\ell_1(T_0^c)}$

➡ $\left[\|h\|_{\ell_1(T_0^c)} \leq \|h\|_{\ell_1(T_0)} + \dfrac{1+\delta}{1-\delta-\theta} . S_0 . \lambda\right]$ I1

(ii) [D'après I0 $\| {}^t X X(\beta' - h) \|_{\ell_\infty} \leq 2 \left(1 + \frac{1-\delta^2}{1-\delta-\theta}\right) \lambda$] I2

Par le lemme $\boxed{1}$ b), d'une part :

I3 [$\| h \|_{\ell_2(T_{01})} \leq (1-\delta)^{-1} \| {}^t X_{T_{01}} X h \|_{\ell_2} + \theta (1-\delta)^{-1} S_0^{-\frac{1}{2}} \| h \|_{\ell_1(T_0^c)}$]

D'autre part, par I2

$$\| {}^t X_{T_{01}} X(\beta' - h) \|_{\ell_\infty} \leq 2\sqrt{2} \left(1 + \frac{1-\delta^2}{1-\delta-\theta}\right) . S_0^{\frac{1}{2}} . \lambda$$

De plus, via lemme $\boxed{2}$ avec $u = \beta'$

$$\| X \beta' \|_{\ell_2} \leq \left(1 + \frac{1}{\sqrt{2}}\right) \frac{(1+\delta)^{3/2}}{1-\delta-\theta} . S_0^{\frac{1}{2}} . \lambda$$

+ hypothèse [PI] $\Rightarrow$ $\| {}^t X_{T_{01}} X \beta' \|_{\ell_2} \leq \left(1 + \frac{1}{\sqrt{2}}\right) \frac{(1+\delta)^2}{1-\delta-\theta} . S_0^{\frac{1}{2}} . \lambda$

Bref $\| {}^t X_{T_{01}} X h \|_{\ell_2} \leq C_0(\delta, \theta) . S_0^{\frac{1}{2}} . \lambda$

I3 $\Rightarrow$ $\| h \|_{\ell_2(T_{01})} \leq \frac{C_0(\delta,\theta)}{1-\delta} S_0^{\frac{1}{2}} . \lambda + \frac{\theta}{(1-\delta) S_0^{1/2}} \| h \|_{\ell_1(T_0^c)}$

$\Rightarrow$ reste à montrer que $\| h \|_{\ell_1(T_0^c)} \leq S_0^{\frac{1}{2}} \| h \|_{\ell_2(T_{01})} + \frac{1+\delta}{1-\delta-\theta} S_0 . \lambda$

vrai par Cauchy-Schwartz + I1

Finalement, on a bien $\| h \|_{\ell_2(T_{01})} \leq C_1(\delta, \theta) . S_0^{\frac{1}{2}} . \lambda$

et $\| h \|_{\ell_1(T_0)} \leq C_2(\delta, \theta) . S_0 . \lambda$