

Noiseless recovery condition for the supremum norm

Proposition 1. Let $\beta \in \mathbb{R}^p$ where $\beta \neq 0$ and $I := \{i \in [p] : |\beta_i| < \|\beta\|_\infty\}$. Let $\tilde{X} = (\tilde{X}_1 | X_I)$ where

$$\tilde{X}_1 = \sum_{i \notin I} \text{sign}(\beta_i) X_i.$$

The following equivalence occurs

$$\exists \lambda > 0 \exists \hat{\beta} \in S_{X, \lambda \|\cdot\|_\infty}(X\beta) \text{ such that } \hat{\beta} \stackrel{\|\cdot\|_\infty}{\sim} \beta \Leftrightarrow \tilde{X}'(\tilde{X}')^+ e_1 = e_1,$$

where e_1 is a vector whose first component is 1 and remaining components are 0.

Before giving the proof, we recall that the subdifferential of the ℓ_∞ norm at 0 is the unit ball of the ℓ_1 norm and for $x \neq 0$ this subdifferential is equal to

$$\begin{aligned} \partial_{\|\cdot\|_\infty}(x) &= \{s \in \mathbb{R}^p : \|s\|_1 \leq 1 \text{ and } s'x = \|x\|_\infty\} \\ &= \left\{ s \in \mathbb{R}^p : \|s\|_1 = 1 \text{ and } \begin{cases} s_i x_i \geq 0 & \text{if } |x_i| = \|x\|_\infty \\ s_i = 0 & \text{otherwise} \end{cases} \right\}. \end{aligned} \quad (1)$$

Proof. (\Rightarrow) Let us assume that there exists $\lambda > 0$ and $\hat{\beta} \in S_{X, \lambda \|\cdot\|_\infty}(X\beta)$ such that $\hat{\beta} \stackrel{\|\cdot\|_\infty}{\sim} \beta$. Then the following property holds

$$\frac{1}{\lambda} X'(X\beta - X\hat{\beta}) \in \partial_{\|\cdot\|_\infty}(\hat{\beta}) = \partial_{\|\cdot\|_\infty}(\beta). \quad (2)$$

Let us set $c = (\|\beta\|_\infty, \beta_I)'$ and $\hat{c} = (\|\hat{\beta}\|_\infty, \hat{\beta}_I)'$. By construction $X\beta = \tilde{X}c$. Moreover, since $\partial_{\|\cdot\|_\infty}(\beta) = \partial_{\|\cdot\|_\infty}(\hat{\beta})$, we also have $X\hat{\beta} = \tilde{X}\hat{c}$. Consequently, according to (2), we have

$$\frac{1}{\lambda} X' \tilde{X}(c - \hat{c}) \in \partial_{\|\cdot\|_\infty}(\beta).$$

Given a subgradient $s \in \partial_{\|\cdot\|_\infty}(\beta)$, according to (2), the following implications hold

$$\begin{cases} \forall i \in I & X'_i \tilde{X}(c - \hat{c}) = 0 \\ \frac{1}{\lambda} \sum_{i \notin I} \text{sign}(\beta_i) X'_i \tilde{X}(c - \hat{c}) = 1 \end{cases} \Rightarrow \frac{1}{\lambda} \tilde{X}' \tilde{X}(c - \hat{c}) = e_1 \Rightarrow \tilde{X}(c - \hat{c}) = \lambda(\tilde{X}')^+ e_1.$$

Using the last implication, one may deduce the following fact

$$\frac{1}{\lambda} X'(X\beta - X\hat{\beta}) = \frac{1}{\lambda} X' \tilde{X}(c - \hat{c}) = X'(\tilde{X}')^+ e_1.$$

Since $X'(\tilde{X}')^+ e_1 \in \partial_{\|\cdot\|_\infty}(\beta)$, let us prove that $\tilde{X}'(\tilde{X}')^+ e_1 = e_1$.

$$X'(\tilde{X}')^+ e_1 \in \partial_{\|\cdot\|_\infty}(\beta) \Rightarrow \begin{cases} \forall i \in I & X'_i(\tilde{X}')^+ e_1 = 0 \\ \sum_{i \notin I} \text{sign}(\beta_i) X'_i(\tilde{X}')^+ e_1 = X'_1(\tilde{X}')^+ e_1 = 1 \end{cases} \Rightarrow \tilde{X}'(\tilde{X}')^+ e_1 = e_1.$$

(\Leftarrow) Let us define $\hat{c} = c - \lambda \tilde{X}'(\tilde{X}')^+ e_1 = (\hat{c}_1, \hat{c}_{-1})'$ (the first component is \hat{c}_1 and remaining components are \hat{c}_{-1}) and let us set $\hat{\beta}$ as follows

$$\forall i \notin I \quad \hat{\beta}_i = \text{sign}(\beta_i) \hat{c}_1 \text{ and } \hat{\beta}_I = \hat{c}_{-1}.$$

Because $c_1 = \|\beta\|_\infty$, then for $\lambda > 0$ small enough, we have $\{i \in [p] : |\hat{\beta}_i| < \|\hat{\beta}\|_\infty\} = \{i \in [p] : |\beta_i| < \|\beta\|_\infty\} = I$ and for $i \notin I$ we have $\beta_i \hat{\beta}_i > 0$. Consequently, for λ small enough we have

$\partial_{\|\cdot\|_\infty}(\hat{\beta}) = \partial_{\|\cdot\|_\infty}(\beta)$. Therefore, to conclude the proof, it is enough to show that $\hat{\beta} \in S_{X,\lambda\|\cdot\|_\infty}(X\beta)$, i.e. $\frac{1}{\lambda}X'(X\beta - X\hat{\beta}) \in \partial_{\|\cdot\|_\infty}(\hat{\beta})$. Because $\text{col}((\tilde{X}')^+) = \text{col}(\tilde{X})$ and since $\tilde{X}\tilde{X}^+$ is the projection onto $\text{col}(\tilde{X})$ we have the following equalities

$$\frac{1}{\lambda}X'(X\beta - X\hat{\beta}) = \frac{1}{\lambda}X'(\tilde{X}c - \tilde{X}\hat{c}) = X'\tilde{X}\tilde{X}^+(\tilde{X}')^+e_1 = X'(\tilde{X}')^+e_1.$$

Finally, let us prove that $X'(\tilde{X}')^+e_1 \in \partial_{\|\cdot\|_\infty}(\beta)$.

$$\begin{aligned} \tilde{X}'(\tilde{X}')^+e_1 = e_1 &\Rightarrow \begin{cases} X'_1(\tilde{X}')^+e_1 = 1 \\ \forall i \in I X'_i(\tilde{X}')^+e_1 = 0 \end{cases} \\ &\Rightarrow \begin{cases} \sum_{i \notin I} \text{sign}(\beta_i) X'_i(\tilde{X}')^+e_1 = 1 \\ \forall i \in I X'_i(\tilde{X}')^+e_1 = 0 \end{cases} \Rightarrow X'(\tilde{X}')^+e_1 \in \partial_{\|\cdot\|_\infty}(\beta). \end{aligned}$$

Consequently, for $\lambda > 0$ small enough, $\hat{\beta} \in S_{X,\lambda\|\cdot\|_\infty}(X\beta)$. □

Linear span of a pattern equivalent class

In Proposition 2 we prove that the linear span of a pattern equivalent class is equal to the model subspace as defined in Vaiter.

Proposition 2. *Let $\phi(x) = \max\{u'_1x \dots, u'_lx\}$ for some $u_1, \dots, u_l \in \mathbb{R}^p$ be a convex function then*

$$\text{lin}(C_x) = \overrightarrow{\partial\phi(x)}^\perp$$

The proof of Proposition 2 is mainly based on the following lemmas

Lemma 1. *Let P be the polyhedron $\{x \in \mathbb{R}^p : s'_1x \leq r_1, \dots, s'_nx \leq r_n\}$ let $\bar{x} \in P$ and $\bar{I} := \{i \in [n] : s'_i\bar{x} = r_i\}$ then*

$$N_P(\bar{x}) = \text{cone}\{s_i\}_{i \in \bar{I}}$$

The proof of Lemma 1 is given in Gruber (Proposition 14.1 page 250).

Lemma 2. *Let P be the polyhedron $\{x \in \mathbb{R}^p : s'_1x \leq r_1, \dots, s'_nx \leq r_n\}$ let $\bar{x} \in P$ and $\bar{I} := \{i \in [n] : s'_i\bar{x} = r_i\}$ then*

$$\text{lin}(N_P(\bar{x})) = \overrightarrow{\text{aff}(F)}^\perp$$

where F is the smallest face of P containing \bar{x} :

$$F = \{x \in P : s'_ix = r_i \quad \forall i \in \bar{I}\}$$

Proof. The face F can be rewritten as the following polyhedron:

$$F = \left\{ x \in \mathbb{R}^p : \begin{cases} s'_ix \leq r_i \text{ and } -s'_ix \leq -r_i & \forall i \in \bar{I} \\ s'_ix \leq r_i & \forall i \notin \bar{I} \end{cases} \right\}.$$

Consequently, according to Lemma 1, the following identity occurs.

$$N_F(\bar{x}) = \text{cone}\{s_i, -s_i : i \in \bar{I}\} = \text{lin}\{s_i : i \in \bar{I}\} = \text{lin}(\text{cone}\{s_i : i \in \bar{I}\}) = \text{lin}(N_P(\bar{x})).$$

It remains to prove that $N_F(\bar{x}) = \overrightarrow{\text{aff}(F)}^\perp$. Let $v \in \overrightarrow{\text{aff}(F)}^\perp$. Because $\overrightarrow{\text{aff}(F)} = \text{lin}\{w - \bar{x} : w \in F\}$ then for all $w \in F$ we have $v'(w - \bar{x}) = 0$. Therefore, $v \in N_F(\bar{x})$. Conversely, let $v \in N_F(\bar{x})$. Because $\bar{x} \in \text{ri}(F)$ then for $x \in F$ and $\delta \neq 0$ small enough $\bar{x} + \delta(x - \bar{x}) \in F$. Consequently,

$$v'(\bar{x} + \delta(x - \bar{x}) - \bar{x}) = \delta v'(x - \bar{x}) \leq 0. \quad (3)$$

Taking δ arbitrarily small and positive (resp. negative) in (3) leads to $v'(x - \bar{x}) \leq 0$ (resp. $v'(x - \bar{x}) \geq 0$). Therefore, $v \in \overrightarrow{\text{aff}(F)}^\perp$ □

Proof of Proposition 2. According to Theorem 4 (in main) we have $\text{lin}(C_x) = \text{lin}(\text{ri}(N_{B^*}(\partial\phi(x))))$, where B^* is the polytope $\text{conv}\{u_1, \dots, u_l\}$. Since a linear space is closed one may deduce that $N_{B^*}(\partial\phi(x)) \subset \text{lin}(\text{ri}(N_{B^*}(\partial\phi(x))))$. Consequently, $\text{lin}(\text{ri}(N_{B^*}(\partial\phi(x)))) = \text{lin}(N_{B^*}(\partial\phi(x)))$. Let $u \in \text{ri}(\partial\phi(x))$. Because $\text{lin}(N_{B^*}(\partial\phi(x))) = \text{lin}(N_{B^*}(u))$ (see *e.g.* Ewald page 16) and since $\partial\phi(x)$ is the smallest face of B^* containing u one may deduce, according to Lemma 1, that $\text{lin}(N_{B^*}(u)) = \overrightarrow{\partial\phi(x)}^\perp$. Therefore $\text{lin}(C_x) = \overrightarrow{\partial\phi(x)}^\perp$. \square

Geometrical characterisation of the noiseless recovery condition

Theorem 1 we provides a geometrical characterisation of the noiseless recovery condition.

Theorem 1. *Let pen be a polyhedral norm.*

$$\exists \lambda > 0 \exists \hat{\beta} \in S_{X, \lambda \text{pen}}(X\beta), \hat{\beta} \stackrel{\text{pen}}{\approx} \beta \Leftrightarrow X'X \text{lin}(C_\beta) \cap \partial_{\text{pen}}(\beta) \neq \emptyset \Leftrightarrow X'X \overrightarrow{\text{aff}(\partial_{\text{pen}}(\beta))}^\perp \cap \partial_{\text{pen}}(\beta) \neq \emptyset.$$

Proof. According to Proposition 2, $\text{lin}(C_\beta) = \overrightarrow{\text{aff}(\partial_{\text{pen}}(\beta))}^\perp$. Therefore $X'X \text{lin}(C_\beta) \cap \partial_{\text{pen}}(\beta) \neq \emptyset \Leftrightarrow X'X \overrightarrow{\text{aff}(\partial_{\text{pen}}(\beta))}^\perp \cap \partial_{\text{pen}}(\beta) \neq \emptyset$. Let us prove the first equivalence

(\Rightarrow): Let $\hat{\beta} \in S_{X, \lambda \text{pen}}(X\beta)$ such that $\hat{\beta} \stackrel{\text{pen}}{\approx} \beta$ then

$$\frac{1}{\lambda} X'X(\beta - \hat{\beta}) \in \partial_{\text{pen}}(\beta).$$

Since $\hat{\beta} \in C_\beta$ then $\frac{\beta - \hat{\beta}}{\lambda} \in \text{lin}(C_\beta)$. Consequently,

$$X'X \text{lin}(C_\beta) \cap \partial_{\text{pen}}(\beta) \neq \emptyset.$$

(\Leftarrow): Let us assume that $X'X \text{lin}(C_\beta) \cap \partial_{\text{pen}}(\beta) \neq \emptyset$, i.e. there exists such $z \in \text{lin}(C_\beta)$ such that $X'Xz \in \partial_{\text{pen}}(\beta)$. Let us set $\hat{\beta} = \beta - \lambda z$. Because, according to Proposition 3 (in main), C_β is relatively open and because $z \in \text{lin}(C_\beta) = \text{aff}(C_\beta)$ (since 0 lies in the relative boundary of C_β), then for λ small enough we have $\hat{\beta} \stackrel{\text{pen}}{\approx} \beta$. Consequently

$$\frac{1}{\lambda} X'(X\beta - X\hat{\beta}) = X'Xz \in \partial_{\text{pen}}(\hat{\beta}).$$

Therefore $\hat{\beta} \in S_{X, \lambda \text{pen}}(X\beta)$, which achieves the proof. \square