The Geometry of Uniqueness and Model Selection of Penalized Estimators including SLOPE, LASSO and Basis Pursuit

Patrick Tardivel, Institute of Mathematics, University of Wrocław, tardivel@math.uni.wroc.pl
Ulrike Schneider, TU Wien, ulrike.schneider@tuwien.ac.at

Uniqueness of the minimizer

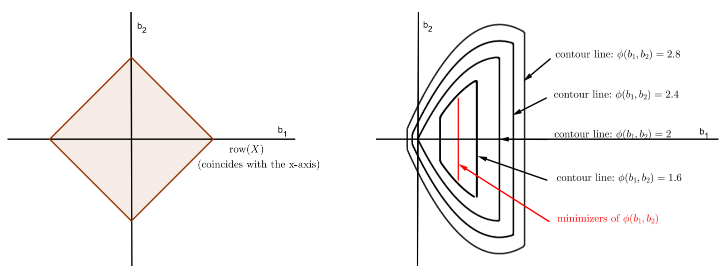
Let $X \in \mathbb{R}^{n\times p}$, $y \in \mathbb{R}^n$ and $\|.\|$ be a norm on $\mathbb{R}^p$ whose unit ball is a polytope.

$$S_{X,\|.\|}(y) := \underset{b\in\mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2}\|y - Xb\|_2^2 + \|b\|.$$

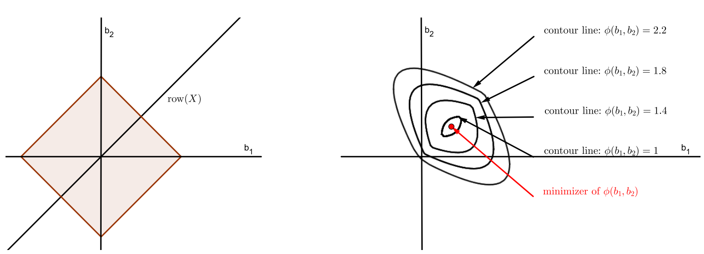
Note that $S_{X,\|.\|}(y) \neq \emptyset$ but $S_{X,\|.\|}(y)$ is not always a singleton. Theorem 1 provides a condition on X so that whatever y the set $S_{X,\|.\|}(y)$ is a singleton. We remind that for a norm $\|.\|$ on $\mathbb{R}^p$, the dual norm $\|.\|^*$ is defined by $\|x\|^* = \sup_{s\in\mathbb{R}^p:\|s\|\leq 1} s'x$.

Theorem 1 *Let $X \in \mathbb{R}^{n\times p}$ and $\|.\|$ be a norm for which the unit ball B is polytope. There exists $y \in \mathbb{R}^n$ for which $S_{X,\|.\|}(y)$ is not a singleton if and only if* $\text{row}(X)$ *intersects a face of the unit ball of the dual norm B^* whose dimension is $< dim(\ker(X))$.*

When $X = (1\ 0)$, the figure below illustrates that the set $S_{X,\|.\|_\infty}(2)$ is not a singleton by plotting the contour lines of the function $\phi(b_1, b_2) = 0.5(2 - b_1)^2 + \max\{|b_1|, |b_2|\}$.



When $X = (1\ 1)$, the figure below illustrates that $S_{X,\|.\|_\infty}(2)$ is a singleton through the contour lines of the function $\phi(b_1, b_2) = 0.5(2 - b_1 - b_2)^2 + \max\{|b_1|, |b_2|\}$.

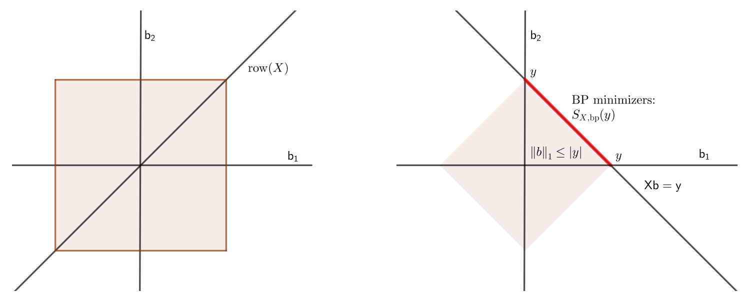


The set $S_{X,\text{bp}}(y)$ of BP minimizers is defined as

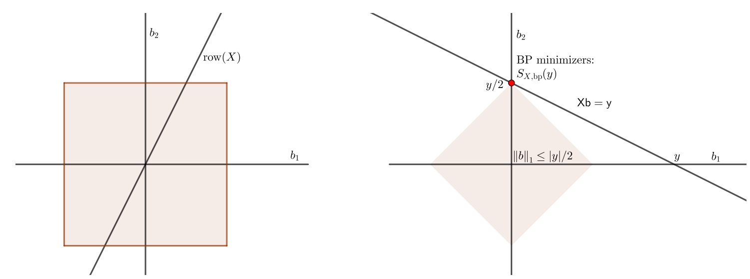
$$S_{X,\text{bp}}(y) = \arg\min \|b\|_1 \text{ subject to } Xb = y.$$

Theorem 2 *Let $X \in \mathbb{R}^{n\times p}$. There exists $y \in \text{col}(X)$ for which $S_{X,\text{bp}}(y)$ is not a singleton if and only if* $\text{row}(X)$ *intersects a face of the unit cube $[-1,1]^p$ whose dimension is $< dim(\ker(X))$.*

When $X = (1\ 1)$, the figure below illustrates that the set $S_{X,\text{bp}}(y)$ is not a singleton for some $y \in \mathbb{R}$.



When $X = (1\ 2)$, the figure below illustrates that the set $S_{X,\text{bp}}(y)$ is a singleton.



Accessible sign vectors for LASSO and BP

Definition 1 *Let $X \in \mathbb{R}^{n\times p}$, $\sigma \in \{-1, 0, 1\}^p$, and $\lambda > 0$. We say that σ is an accessible sign vector for LASSO (or BP), if there exists $y \in \mathbb{R}^n$ and $\hat\beta \in S_{X,\lambda\|.\|_1}(y)$ (or there exists $y \in \text{col}(X)$ and $\hat\beta \in S_{X,\text{bp}}(y)$), such that* $\text{sign}(\hat\beta) = \sigma$.

$$F(\sigma) = E_1 \times \cdots \times E_p \text{ with } E_j = \begin{cases} \{\sigma_j\} & |\sigma_j| = 1 \\ [-1, 1] & \sigma_j = 0 \end{cases}.$$

Theorem 3 *Let $X \in \mathbb{R}^{n\times p}$, $\lambda > 0$ and $\sigma \in \{-1, 0, 1\}^p$.*

Geometrical characterization: *The sign vector σ is accessible for LASSO (or BP) if and only if* $\text{row}(X)$ *intersects the face $F(\sigma)$.*

Analytical characterization: *The sign vector σ is accessible for LASSO (or BP) if and only if the following implication holds: $X b = X\sigma \Rightarrow \|b\|_1 \geq \|\sigma\|_1$.*

In the high-dimensional linear regression model, the accessibility condition is actually a necessary and sufficient for sign recovery by thresholded LASSO and by thresholded BP (Tardivel and Bogdan) and by thresholded justice pursuit (Descloux et al.).

Article and videos related to this poster: Schneider and Tardivel, part 1: uniqueness, part 2: model selection by SLOPE
Related articles: Tardivel and Bogdan, Ewald and Schneider