$$Y = \underset{\uparrow}{X} \beta^0 + \varepsilon$$

random

condition on X:

$Y|X$, X fixed

$\varepsilon|X$: assume that $\varepsilon \perp\!\!\!\perp X$

holds in causal models
if X are ancestors of Y

so far:

$$\|X(\hat\beta - \beta^0)\|_2^2 / n \le \dots \xrightarrow{P} 0$$

$\boxed{\text{interested now in estimation of } \beta^0:}$

$$\|\hat\beta - \beta^0\|_1 , \ \|\dots\|_2 , \ \|\dots\|_\infty$$

if $\lambda_{min}(\hat\Sigma) > 0 \Rightarrow \|\theta - \beta^0\|_2^2 = 0 \Rightarrow \theta = \beta^0$

$X\theta = X\beta^0$

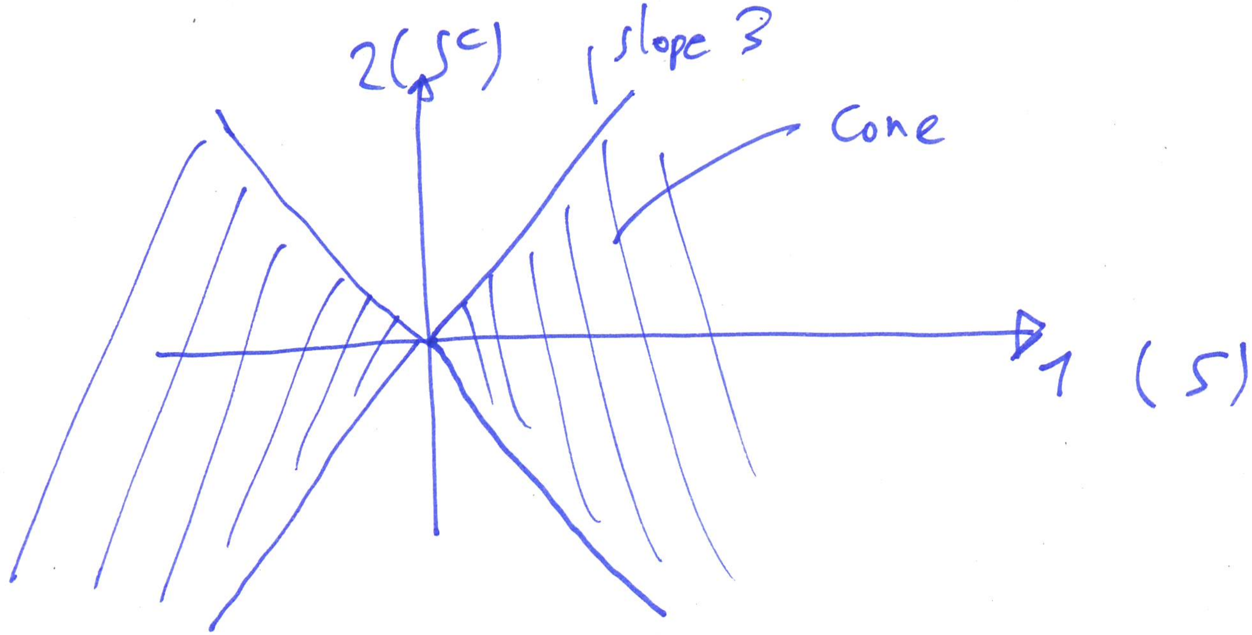
$$0 = \|X(\theta - \beta^0)\|_2^2 / n \geq \overbrace{\lambda_{min}^2(\hat{\Sigma})}^{\text{min. eigenvalue of } \hat{\Sigma} \text{ (notation with } \lambda_{min}^{2}\text{)}} \|\theta - \beta^0\|_2^2$$

if $\lambda_{min}^2(\hat{\Sigma}) > 0$: $\|\theta - \beta^0\|_2^2 = 0 \Longrightarrow \theta = \beta^0$

for $p > n$: $\lambda_{min}^2(\hat{\Sigma}) = 0$

Cone condition:

e.g. $S = \{1\}$, $S^c = \{2\}$, $p = 2$



in cone ///: $\|\beta_{S^c}\|_1 \leq 3\|\beta_S\|_1$

"Proof" for identifiability with Lasso assuming restricted eigenvalues

(i) Lasso fulfills cone condition on $\mathcal{T}$:

$$\|(\hat{\beta}-\beta^0)_{S_0^c}\|_1 \leq 3\|(\hat{\beta}-\beta^0)_{S_0}\|_1$$

$$u_S : \quad (u_S)_j = \begin{cases} u_j & j \in S \\ 0 & j \notin S \end{cases}$$

not easy to see $\rightsquigarrow$ p. 106 in book

(ii) assuming sparsity of β^0, we know that on $\mathcal{T}$:

$$a_n := \|X(\hat\beta-\beta^0)\|_2^2/n \quad \text{with} \quad a_n \longrightarrow 0$$

$$a_n = \|X(\hat\beta-\beta^0)\|_2^2/n = (\hat\beta-\beta^0)^T \hat\Sigma (\hat\beta-\beta^0)$$

$$\underset{(i)}{\geq} \kappa^2(s_0,3)\,\|(\hat\beta-\beta^0)_{S_0}\|_2^2$$

(iii) assuming $\kappa^2(s_0,3) \geq L > 0$

$$\Rightarrow \|(\hat\beta-\beta^0)_{S_0}\|_2^2 \leq \frac{a_n}{\kappa^2(s_0,3)} \leq \frac{a_n}{L} \longrightarrow 0$$

$$\|(\hat\beta-\beta^0)_{S_0^c}\|_1 \underset{(i)}{\leq} 3\,\|(\hat\beta-\beta^0)_{S_0}\|_1$$

$$= 3 \sum_{j\in S_0} |(\hat\beta-\beta^0)_j|$$

Cauchy-Schwarz

$$\leq 3 \sqrt{\sum_{j\in S_0} 1^2}\,\sqrt{\sum_{j\in S_0} (\hat\beta-\beta^0)_j^2}$$

$$= 3\sqrt{s_0}\cdot \cancel{\|(\hat\beta-\beta^0)\|_2}\ \|(\hat\beta-\beta^0)_{S_0}\|_2$$

not so easy : it will hold that $a_n \asymp \frac{s_0 \log(p)}{n}$ (see Th. 6.1)

$$\Rightarrow \|(\hat\beta-\beta^0)_{S_0^c}\|_2 \leq 3\sqrt{s_0}\,\mathcal{O}(\sqrt{a_n}) = \mathcal{O}\left(s_0\cdot\sqrt{\frac{\log(p)}{n}}\right)$$

$$\|(\hat{\beta}-\beta^0)_{S_0}\|_1 = O\left(s_0 \cdot \sqrt{\frac{\log(p)}{n}}\right)$$

$$\not\Longrightarrow \|\hat{\beta}-\beta^0\|_1 = O\left(s_0 \cdot \sqrt{\frac{\log(p)}{n}}\right)$$

if restricted eigenvalue condition $\kappa^2(s_0, 3) \geq L > 0$

restricted eigenvalues:

$$\|\hat{\beta}-\beta^0\|_2^2 = \mathcal{O}\left(\frac{s_0 \log(p)}{n}\right) \longrightarrow 0 \quad \text{if } s_0 = o\left(\frac{n}{\log(p)}\right)$$

$$\|\hat{\beta}-\beta^0\|_1 = \mathcal{O}\left(s_0 \cdot \sqrt{\frac{\log(p)}{n}}\right) \longrightarrow 0 \quad \text{if } s_0 = o\left(\sqrt{\frac{n}{\log(p)}}\right)$$

only needs: compatibility constant