$Y = X\beta^0 + \varepsilon$ Rows of X iid $N_p(0, \Sigma)$, $\Omega = \Sigma^{-1}$

β^0: sparsity s

$X_j = X_{-j}\gamma^{(j)} + \varepsilon^{(j)}$

$\gamma^{(j)}$: sparsity s_j

$s_{max} = \max(s, \max_j s_j)$

We consider an asymptotic regime where X, s_{max} etc. are allowed to change as $n \to \infty$. We will treat σ as constant.

<u>Theorem 40</u> Suppose the min e'val of Σ is always at least some fixed const $c_{min} > 0$ and $\max_j \Sigma_{jj} < A$, const A. Suppose $s_{max}\sqrt{\log(p)/n} \to 0$. Then $\exists$ constants A_1, A_2 s.t. if we take $\lambda = \lambda_j = A_1\sqrt{\log(p)/n}$, we have

$$\sqrt{n}(\hat{b} - \beta^0) = W + \Delta$$

$$W \mid X \sim N_p(0, \sigma^2 \hat{\Theta}\hat{\Sigma}\hat{\Theta}^T)$$

and as $n, p \to \infty$, $\mathbb{P}(\|\Delta\|_\infty > A_2 s \log(p)/\sqrt{n}) \to 0$.

<u>Remark</u> We see in particular that

$$\sqrt{n}(\hat{b}_j - \beta^0_j) \simeq W_j \quad \text{where } W_j \sim N(0, \sigma^2 \underbrace{(\hat{\Theta}\hat{\Sigma}\hat{\Theta}^T)_{jj}}_{= \frac{\frac{1}{n}\|X_j - X_{-j}\hat{\gamma}^{(j)}\|_2^2}{\hat{\tau}_j^4}})$$

This suggests the following approximate $(1-\alpha)$-level confidence interval for β^0_j

$$\left[\hat{b}_j - z_{\alpha/2}\,\sigma\frac{\sqrt{v_j}}{\sqrt{n}},\ \hat{b}_j + z_{\alpha/2}\,\sigma\frac{\sqrt{v_j}}{\sqrt{n}}\right]$$

where z_α is the upper α pt of a standard normal. The only unknown quantity is σ, for which $\exists$ estimation techniques e.g. the scaled Lasso of Sun & Zhang 2012.

<u>Proof</u>

Consider the sequence of events Λ_n described by the following properties:

⌈ Notation: For a matrix $M \in \mathbb{R}^{p\times p}$ and $m < p$ define

$$\phi^2_{M,m} = \min_{I \subseteq \{1,\dots,p\},\ |I| \le m} \phi^2_M(I)$$

where recall

$$\phi^2_M(I) = \inf_{\substack{\beta : \|\beta_I\|_1 \neq 0 \\ \|\beta_{I^c}\|_1 \le 3\|\beta_I\|_1}} \frac{\beta^T M \beta}{\|\beta_I\|_1^2/|I|}$$ ⌋

- $\phi^2_{\hat{\Sigma},s} \ge c_{min}/2$ and $\phi^2_{\hat{\Sigma}_{-j,-j},s_j} \ge c_{min}/2 \quad \forall j$
- $2\|X^T\varepsilon\|_\infty/n \le \lambda$ and $2\|X_{-j}^T\varepsilon^{(j)}\|_\infty/n \le \lambda \quad \forall j$
- $\|\varepsilon^{(j)}\|_2^2/n \ge \Omega_{jj}^{-1}(1 - 4\sqrt{\log(p)/n}) \quad \forall j$

$\mathbb{P}(\Lambda_n) \to 1$. We work on Λ_n.

By thm 23,

$$\|\beta^0 - \hat\beta\|_1 \le c_1 s \sqrt{\log(p)/n} \qquad \text{const } c_1$$

Consider the linear models

$$X_j = X_{-j}\gamma^{(j)} + \varepsilon^{(j)} \qquad (\varepsilon_i^{(j)} \overset{iid}{\sim} N(0, \Omega_{jj}^{-1})) \qquad -(\ast)$$

Note $\Omega_{jj}^{-1} = \mathrm{Var}(\cancel{\ldots} X_{ij} \mid X_{i,-j}) \le \mathrm{Var}(X_{ij}) = \Sigma_{jj} \le A$

Also max e'val of Ω is at most c_{min}^{-1}, so

$$\Omega_{jj} \le c_{min}^{-1} \qquad (\Omega_{jj} = e_j^T \Omega e_j)$$

jth unit vector

Then

$$\hat\varsigma_j^2 = \frac{1}{n}\|X_j - X_{-j}\hat\gamma^{(j)}\|_2^2 + \lambda\|\hat\gamma^{(j)}\|_1$$

$$\ge \frac{1}{n}\|\underbrace{X_j - X_{-j}\gamma^{(j)}}_{\varepsilon^{(j)}} + X_{-j}\gamma^{(j)} - X_{-j}\hat\gamma^{(j)}\|_2^2$$

$$\ge \frac{1}{n}\|\varepsilon^{(j)}\|_2^2 + \frac{2}{n}\varepsilon^{(j)T}X_{-j}(\gamma^{(j)} - \hat\gamma^{(j)})$$

$$\ge \Omega_{jj}^{-1}\Big(1 - \underbrace{4\sqrt{\log(p)/n}}_{\to 0}\Big) - \underbrace{\frac{2}{n}\|X_{-j}^T\varepsilon^{(j)}\|_\infty \|\hat\gamma^{(j)} - \gamma^{(j)}\|_1}_{\le \lambda c_2 s_j \sqrt{\log(p)/n} \;=\; c\frac{s_j \log p}{n} \to 0}$$

(thm 23 applied to $(\ast)$)

$$\ge c_{min}/2$$

forall j, n suff. large.

Thus,

$$\|\Delta\|_\infty \le \lambda\sqrt{n}\,\|\beta^0 - \hat\beta\|_1 \max_j \hat\varsigma_j^{-2}$$

$$\le c_1 s\sqrt{\log p}\; 2c_{min}^{-1}\; A_1\sqrt{\frac{\log p}{n}}$$

$$\le A_2 \frac{s\log p}{n}$$

$\square$