If $\frac{1}{n}X^TX$ has min e'val $c_{min} > 0$ (so $p \le n$), then $\phi^2 \geq c_{min}$.

Indeed, $\|\beta_S\|_1 = \text{sgn}(\beta_S)^T \beta_S \overset{C\text{-}S}{\le} \sqrt{s}\,\|\beta_S\|_2 \le \sqrt{s}\,\|\beta\|_2$

So $\phi^2 \geq \inf_{\beta \neq 0} \frac{\frac{1}{n}\|X\beta\|_2^2}{\|\beta\|_2^2} = c_{min}$

**Theorem 23** Assume $\phi^2 > 0$ and let $\hat{\beta}$ be the Lasso soln with $\lambda = A\sigma\sqrt{\frac{\log p}{n}}$, $A > 0$.

Then with probability at least $1 - 2p^{-(A^2/8-1)}$ have

$$\frac{1}{n}\|X(\beta^0 - \hat{\beta})\|_2^2 + \lambda\|\beta^0 - \hat{\beta}\|_1 \le \frac{16\lambda^2 s}{\phi^2} = \frac{16 A^2 \log p}{\phi^2} \underbrace{\frac{\sigma^2 s}{n}}_{\text{MSPE of OLS applied to } X_S}$$

**Proof**: Start with the basic inequality

$$\frac{1}{2n}\|X(\beta^0 - \hat{\beta})\|_2^2 + \lambda\|\hat{\beta}\|_1 \le \frac{1}{n}\varepsilon^T X(\hat{\beta} - \beta^0) + \lambda\|\beta^0\|_1$$

(see proof of theorem 9). Work on $\Omega = \{2\|X^T\varepsilon\|_\infty / n \le \lambda\}$ where after applying Hölder, we get

$$\frac{1}{n}\|X(\beta^0 - \hat{\beta})\|_2^2 + 2\lambda\|\hat{\beta}\|_1 \le \lambda\|\beta^0 - \hat{\beta}\|_1 + 2\lambda\|\beta^0\|_1$$

Lemma 13 shows that $P(\Omega) \geq 1 - p^{-(A^2/8-1)}$.

[Idea: Have $\frac{1}{n}\|X(\beta^0 - \hat{\beta})\|_2^2 \le 3\lambda\|\beta^0 - \hat{\beta}\|_1$.

If we could show that $3\lambda\|\beta^0 - \hat{\beta}\|_1 \le c\lambda\frac{1}{\sqrt{n}}\|X(\beta^0 - \hat{\beta})\|_2$

then $\frac{1}{n}\|X(\beta^0 - \hat{\beta})\|_2^2 \le c^2\lambda^2$ and $3\lambda\|\beta^0 - \hat{\beta}\|_1 \le c^2\lambda^2$]

$$a = \frac{1}{n\lambda}\|X(\beta^0 - \hat{\beta})\|_2^2$$

$$a + 2(\|\hat{\beta}_N\|_1 + \|\hat{\beta}_S\|_1) \le \|\hat{\beta}_S - \beta^0_S\|_1 + \|\hat{\beta}_N\|_1 + 2\|\beta^0_S\|_1$$

$$a + \|\hat{\beta}_N\|_1 \le \|\beta^0_S - \hat{\beta}_S\|_1 + 2\|\beta^0_S\|_1 - 2\|\hat{\beta}_S\|_1$$

$$a + \|\beta^0_N - \hat{\beta}_N\|_1 \le 3\|\beta^0_S - \hat{\beta}_S\|_1$$

$$a + \|\beta^0 - \hat{\beta}\|_1 \le 4\|\beta^0_S - \hat{\beta}_S\|_1 \quad (1)$$

Using the compatibility condition on $\beta = \beta^0 - \hat{\beta}$, have

$$\phi^2 \le \frac{\frac{1}{n}\|X(\hat{\beta} - \beta^0)\|_2^2}{\frac{1}{s}\|\hat{\beta}_S - \beta^0_S\|_1^2}$$

Substitute into ①

$$\frac{1}{n}\|X(\beta^0-\hat\beta)\|_2^2 + \lambda\|\beta^0-\hat\beta\|_1 \le 4\lambda \frac{\frac{1}{\sqrt{n}}\|X(\beta^0-\hat\beta)\|_2}{\frac{1}{\sqrt{s}}\phi}$$

Thus $\frac{1}{\sqrt{n}}\|X(\beta^0-\hat\beta)\|_2 \le \frac{4\lambda\sqrt{s}}{\phi}$

Substitute into RHS of

$$\frac{1}{n}\|X(\hat\beta-\beta)\|_2^2 + \lambda\|\hat\beta-\beta^0\|_1 \le \frac{16\lambda^2 s}{\phi^2}. \quad \square$$

<u>The compatibility condition and random design</u>

Define $\phi_\Sigma^2(S) = \inf\limits_{\substack{\beta:\|\beta_N\|_1\le 3\|\beta_S\|_1 \\ \|\beta_S\|_1\neq 0}} \frac{\beta^T\Sigma\beta}{\frac{1}{|S|}\|\beta_S\|_1^2}$ for $\Sigma\in\mathbb{R}^{p\times p}$

Note $\phi^2 = \phi^2_{\hat\Sigma}(S)$ where $\hat\Sigma = \frac{1}{n}X^TX$ and $S$ is the support set of $\beta^0$.

<u>Lemma 24</u>. Let $\Theta, \Sigma \in \mathbb{R}^{p\times p}$. Suppose $\phi^2_\Theta(S) > 0$ and $\max\limits_{j,k}|\Theta_{jk} - \Sigma_{jk}| \le \frac{\phi^2_\Theta(S)}{32|S|}$. Then $\phi^2_\Sigma(S) \ge \frac{\phi^2_\Theta(S)}{2} > 0$.

<u>Proof</u> We'll supress dependence on $S$ and let $s=|S|$, $t = \frac{\phi^2_\Theta}{32s}$

We have $|\beta^T(\Sigma-\Theta)\beta| \overset{\text{Hölder}}{\le} \|\beta\|_1\|(\Sigma-\Theta)\beta\|_\infty$

$\le \|\beta\|_1^2 t$ (Hölder to inner product between rows of $\Sigma-\Theta$ and $\beta$)

If $\|\beta_N\|_1 \le 3\|\beta_S\|_1$ then

$$\|\beta\|_1 = \|\beta_N\|_1 + \|\beta_S\|_1 \le 4\|\beta_S\|_1 \le 4\frac{\sqrt{\beta^T\Theta\beta}}{\phi_\Theta/\sqrt{s}}$$

Thus $$\underbrace{\beta^T\Theta\beta - \underbrace{\frac{\phi^2_\Theta}{32s}}_{t}\frac{16\beta^T\Theta\beta}{\phi^2_\Theta/s}}_{=\frac{1}{2}\beta^T\Theta\beta} \le \frac{1}{2}\beta^T\Theta\beta \le \beta^T\Sigma\beta \quad \square$$

We now apply lemma 24 with $\Theta = \Sigma^0 = \mathbb{E}\frac{1}{n}X^TX$ and $\Sigma = \hat{\Sigma} = \frac{1}{n}X^TX$.

We will consider an asymptotic regime where the $X$ matrices can grow with $n$.

<u>Theorem 25</u> Suppose of $X$ are iid and each entry of $X$ is mean-zero sub-G with param $v$. Suppose $s\sqrt{\log(p)/n} \to 0$ as $n \to \infty$. Let

$$\phi^2_{\hat{\Sigma}} = \min_{S: |S| = s} \phi^2_{\hat{\Sigma}}(S)$$

$$\phi^2_{\Sigma^0} = \min_{S: |S| = s} \phi^2_{\Sigma^0}(S) > 0$$

Then $\mathbb{P}(\phi^2_{\hat{\Sigma}} \geq \phi^2_{\Sigma^0}/2) \to 1$ as $n \to \infty$.

<u>Proof</u> By lemma 24, ETS

$$\begin{aligned}
\mathbb{P}(\max_{jk} |\hat{\Sigma}_{jk} - \Sigma^0_{jk}| \geq t) &= \mathbb{P}\Big(\bigcup_{jk} \{|\hat{\Sigma}_{jk} - \Sigma^0_{jk}| \geq t\}\Big) \\
&\leq \sum_{jk} \mathbb{P}\big(|\underbrace{\hat{\Sigma}_{jk}}_{\frac{1}{n}\sum_{i=1}^n X_{ij}X_{ik}} - \Sigma^0_{jk}| \geq t\big) \\
&\overset{\text{lemma 15, prop 14}}{\leq} 2p^2 \exp\left(-\frac{nt^2}{2(64v^4 + 4v^2 t)}\right) \\
&\leq c_1 \exp\left(-c_2 \frac{n}{s^2} + c_3 \log p\right) \quad \text{for } c_1, c_2, c_3 \text{ const} \\
&\to 0
\end{aligned}$$