① $\lambda_{min}\left(\frac{1}{n}X^TX\right) = \lambda_{min} > 0$

② $Y = X\beta^* + \varepsilon$, support$(\beta^*) = S$, $|S| = s$

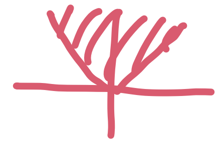
③ $\left\|\frac{1}{n}X^T\varepsilon\right\|_{max} \le \sqrt{\frac{\log p}{n}}$, where $\|u\|_{max} = \max_j \|V_j\|$

$\hat\beta \Rightarrow \frac{1}{2n}\|Y - X\beta\|^2 + \lambda\|\beta\|_1$

Q1. $\frac{1}{2n}\|X(\hat\beta - \beta^*)\|^2 \le \left|\frac{1}{n}\varepsilon^T X(\hat\beta - \beta^*)\right| - \lambda\|\hat\beta\|_1 + \lambda\|\beta^*\|_1$

Q2. if $\lambda = 2\sqrt{\frac{\log p}{n}}$, $\frac{1}{2}\lambda_{min}\|\hat\beta - \beta^*\|^2 \le \frac{1}{2}\lambda\|\hat\beta - \beta^*\|_1 - \lambda\|\hat\beta\|_1 + \lambda\|\beta^*\|_1$

Q3. $\|(\hat\beta - \beta^*)_{S^c}\|_1 \le 3\|(\hat\beta - \beta^*)_S\|_1$: cone property



if $x \in$ Cone, $ax \in$ Cone, $\forall a > 0$

Q4. $\|\hat\beta - \beta^*\| \le \frac{16}{\lambda_{min}}\sqrt{\frac{s\log p}{n}}$

① $\frac{1}{2n}\|Y - X\hat\beta\|^2 + \lambda\|\hat\beta\|_1 \le \frac{1}{2n}\|Y - X\beta^*\|^2 + \lambda\|\beta^*\|_1$ by property of LASSO

$$\frac{1}{2n}\|Y - X\beta^* + X\beta^* - X\hat\beta\|^2 + \lambda\|\hat\beta\|_1$$

$$= \frac{1}{2n}\|Y - X\beta^*\|^2 + \frac{1}{2n}\|X(\hat\beta - \beta^*)\|^2 - \frac{1}{n}\langle Y - X\beta^*, X(\hat\beta - \beta^*)\rangle + \lambda\|\hat\beta\|_1$$

$$= \frac{1}{2n}\|Y - X\beta^*\|^2 + \frac{1}{2n}\|X(\hat\beta - \beta^*)\|^2 - \left|\frac{1}{n}\varepsilon^T X(\hat\beta - \beta^*)\right| + \lambda\|\hat\beta\|_1$$

$$\le \frac{1}{2n}\|Y - X\beta^*\|^2 + \lambda\|\beta^*\|_1$$

$$\therefore \frac{1}{2n}\|X(\hat\beta - \beta^*)\|^2 \le \left|\frac{1}{n}\varepsilon^T X(\hat\beta - \beta^*)\right| - \lambda\|\hat\beta\|_1 + \lambda\|\beta^*\|_1$$

② $\frac{1}{2n}\|X(\hat\beta - \beta^*)\|^2 = \frac{1}{2}\left|(\hat\beta - \beta^*)^T \underline{\frac{X^TX}{n}}(\hat\beta - \beta^*)\right| \ge \frac{1}{2}\lambda_{min}\|X(\hat\beta - \beta^*)\|^2$

③ $0 \le 0.5\|\hat\beta - \beta^*\|_1 + \|\beta^*\|_1 - \|\hat\beta\|_1$

$$= 0.5\|(\hat\beta - \beta^*)_S\|_1 + 0.5\|(\hat\beta - \beta^*)_{S^c}\|_1 + \|\beta^*_S\|_1 - \|\hat\beta_S\|_1 - \|(\hat\beta - \beta^*)_{S^c}\|_1$$

$$\le 0.5\|(\hat\beta - \beta^*)_S\|_1 - 0.5\|(\hat\beta - \beta^*)_{S^c}\|_1 + \|(\hat\beta - \beta^*)_S\|_1$$

$$0.5\|(\hat\beta - \beta^*)_{S^c}\|_1 \le 1.5\|(\hat\beta - \beta^*)_S\|_1$$

$$\|(\hat\beta - \beta^*)_{S^c}\|_1 \le 3\|(\hat\beta - \beta^*)_S\|_1$$

Q3 $\Rightarrow$ $\|\hat{\beta}-\beta^*\|_1 \le 4\|(\hat{\beta}-\beta^*)_S\|_1$

Q4. $\le 4\sqrt{s}\,\|(\hat{\beta}-\beta^*)_S\|_2$

$\le 4\sqrt{s}\,\|\hat{\beta}-\beta^*\|_2$

Q2 $\frac{1}{2}\lambda_{min}\|\hat{\beta}-\beta^*\|^2 \le 0.5\lambda\|\hat{\beta}-\beta^*\|_1 + \lambda\|\beta^*\|_1 - \lambda\|\hat{\beta}\|_1$

$\le 1.5\lambda\|\hat{\beta}-\beta^*\|_1$

$\le 1.5\lambda 4\sqrt{s}\,\|\hat{\beta}-\beta^*\|_2$

$\|\hat{\beta}-\beta^*\|^2 \le \frac{12\lambda}{\lambda_{min}}\sqrt{s} = \frac{24}{\lambda_{min}}\sqrt{\frac{s\log p}{n}}$

Q3 $\|\hat{\beta}-\beta^*\|_1 \le 4\|(\hat{\beta}-\beta^*)_S\|_1$

$\le 4\sqrt{s}\,\|(\hat{\beta}-\beta^*)_S\|_1$

$\le 4\sqrt{s}\,\|\hat{\beta}-\beta^*\|_1$

$\frac{1}{2}\lambda_{min}\|\hat{\beta}-\beta^*\|^2 \le \frac{1}{2}\lambda\|\hat{\beta}-\beta^*\|_1 - \lambda\|\hat{\beta}\| + \lambda\|\beta^*\|$

$\le 2\sqrt{s}\,\lambda\|\hat{\beta}-\beta^*\|_1 - \lambda\|\hat{\beta}\| + \lambda\|\beta^*\|$

$\le (2\sqrt{s}+1)\lambda\|\hat{\beta}-\beta^*\|_1$

$\|\hat{\beta}-\beta^*\|^2 \le \frac{2(2\sqrt{s}+1)}{\lambda_{min}}\lambda\|\hat{\beta}-\beta^*\|_1$