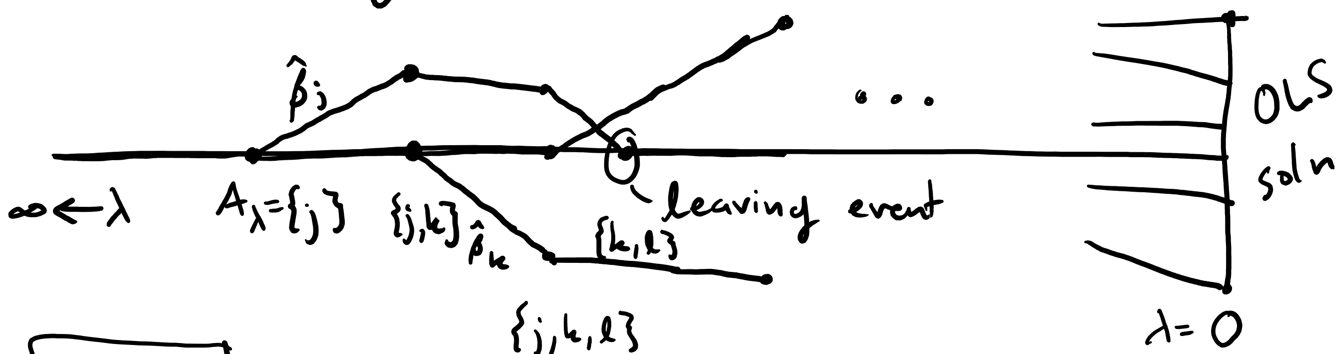
More about the Lasso

Wednesday, April 19, 2017 5:47 PM

Recall the lasso: $\min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1$

Define $A_\lambda = \operatorname{supp}(\hat{\beta}_\lambda) = \{ j : \hat{\beta}_{\lambda j} \neq 0 \}$.

As $\lambda \to \infty$, $\beta \to 0$ so start at $\lambda = \infty$ and eventually as λ decreases $A_\lambda \neq \{\}$.



$\boxed{\text{fact 1}}$ $\hat{\beta}_j$ is continuous, piecewise linear in λ

$\boxed{\text{fact 2}}$ Slope of $\hat{\beta}_j$ is regression coef on residual for OLS on A_λ.

$\boxed{\text{fact 3}}$ LAR with lasso modification solves the lasso.

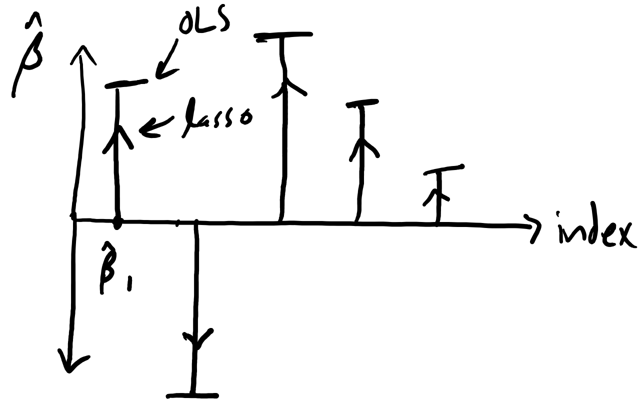
We can think of the lasso path as a sequence of models $A_{\lambda_1}, A_{\lambda_2}, \ldots$ for $\lambda_1 > \lambda_2 \ldots$ knots.

Lasso introduces a bias

So, for selected

$\hat{\beta} \neq \hat{\beta}^{OLS}$ T

model A ...



model A, solve restricted OLS,

$$\tilde{\beta} = (X_A^T X_A)^{-1} X_A^T y$$

<u>Clarification</u>: $\beta_j = \beta_{j+} - \beta_{j-}$ for $\beta_{j+}, \beta_{j-} \geq 0$

QP: $$\min_{\beta_+, \beta_-} \|y - X(\beta_+ - \beta_-)\|_2^2 \quad \text{s.t.} \quad \sum_j \beta_{j+} + \beta_{j-} \leq C$$

$$\beta_{j+}, \beta_{j-} \geq 0$$

if β_+, β_- feasible then $\sum_j |\beta_j| \leq \sum_j \beta_{j+} + \beta_{j-} \leq C$

for $\beta := \beta_+ - \beta_- \Rightarrow$ QP $\geq$ Lasso

if β is Lasso sol$^{\underline{n}}$ $\beta_{j+} = (\beta_j)_+$, $\beta_{j-} = (-\beta_j)_+$

$\sum_j \beta_{j+} + \beta_{j-} = \sum_j |\beta_j| \leq C \Rightarrow$ QP $\leq$ Lasso