# Recap

$$\text{Lasso: } \hat{\beta}(\lambda) = \text{argmin}_{\beta}(\|Y - X\beta\|_2^2/n + \lambda\|\beta\|_1)$$

want to understand its asymptotic properties for high-dimensional linear model

$$Y = X\beta^0 + \varepsilon, \; p = p_n \gg n \text{ and } n \to \infty$$

Assumptions on the model:

- condition on "nice" errors:
  $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. $\mathcal{N}(0, \sigma^2)$
- scaled covariates:
  $n^{-1} \sum_{i=1}^{n} (X_i^{(j)})^2 \equiv 1$
- sparsity of regression coefficients w.r.t. $\ell_1$-norm:
  $\|\beta^0\|_1 = o(\sqrt{n/\log(p_n)})$ $(n \to \infty)$
  (implicit: dimensionality $p_n$: $\log(p_n)/n \to 0$ $(n \to \infty)$)

*Theorem*
Assume the model assumptions (above).
Assumption on the estimator: choose
$\lambda = \lambda_n = 4\sigma\sqrt{\frac{t_n^2 + 2\log(p_n)}{n}}$ with
$t_n^2 \to \infty,\ t_n^2 = O(\log(p_n))$, e.g $t_n^2 = \log(p_n)$

in short: $\lambda_n = C\sigma\sqrt{\log(p_n)/n}$ with $C > 0$ sufficiently large
(e.g. $C > 4\sqrt{3}$)

if $\sigma$ unknown: $\hat{\sigma}$ with $\mathbb{P}[C > \hat{\sigma} \geq \sigma] \to 1$ $(n \to \infty)$
Then:

$$\|X(\hat{\beta}(\lambda_n) - \beta^0)\|_2^2/n \to 0 \text{ in probability } (n \to \infty)$$