Recap

linear model:

$$Y = X_{n \times p} \underbrace{\beta^0}_{\text{unknown par.}} + \varepsilon$$

"0" means "true" (parameter)

high-dimensional: $p \gg n$

$\rightsquigarrow$ need to make assumptions such that β^0 can be well estimated
we make a sparsity assumption on β^0, see also later

Lasso:

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

- sparse estimator
  (different from sparsity assumption in the model)