Regularization

- AdaBoost minimizes

$$L(\boldsymbol{\lambda}) = \sum_i \exp\left(-y_i \sum_j \lambda_j g_j(x_i)\right)$$

- to avoid overfitting, want to constrain $\boldsymbol{\lambda}$ to make solution "smoother"
- (ℓ_1) regularization:

$$\begin{aligned} &\text{minimize: } && L(\boldsymbol{\lambda}) \\ &\text{subject to: } && \|\boldsymbol{\lambda}\|_1 \leq B \end{aligned}$$

- or:

$$\text{minimize: } \quad L(\boldsymbol{\lambda}) + \beta \|\boldsymbol{\lambda}\|_1$$

- other norms possible
  - ℓ_1 ("lasso") currently popular since encourages sparsity [Tibshirani]