Which penalty functions can we use?

- ***L*$_2$ penalty:** $$\text{penalty}(\boldsymbol{\theta}) = ||\boldsymbol{\theta}||_2 = \sum_d \theta_d^2$$
  - Easy to optimize (strongly convex)
  - Closed form solutions exists
  - Redundant parameters will be close to 0, but never 0
- ***L*$_1$ penalty:** $$\text{penalty}(\boldsymbol{\theta}) = ||\boldsymbol{\theta}||_1 = \sum_d |\theta_d|$$
  - Induces **sparse** solutions
  - Called “Lasso” regularization
  - Much harder to optimize (not in this lecture)