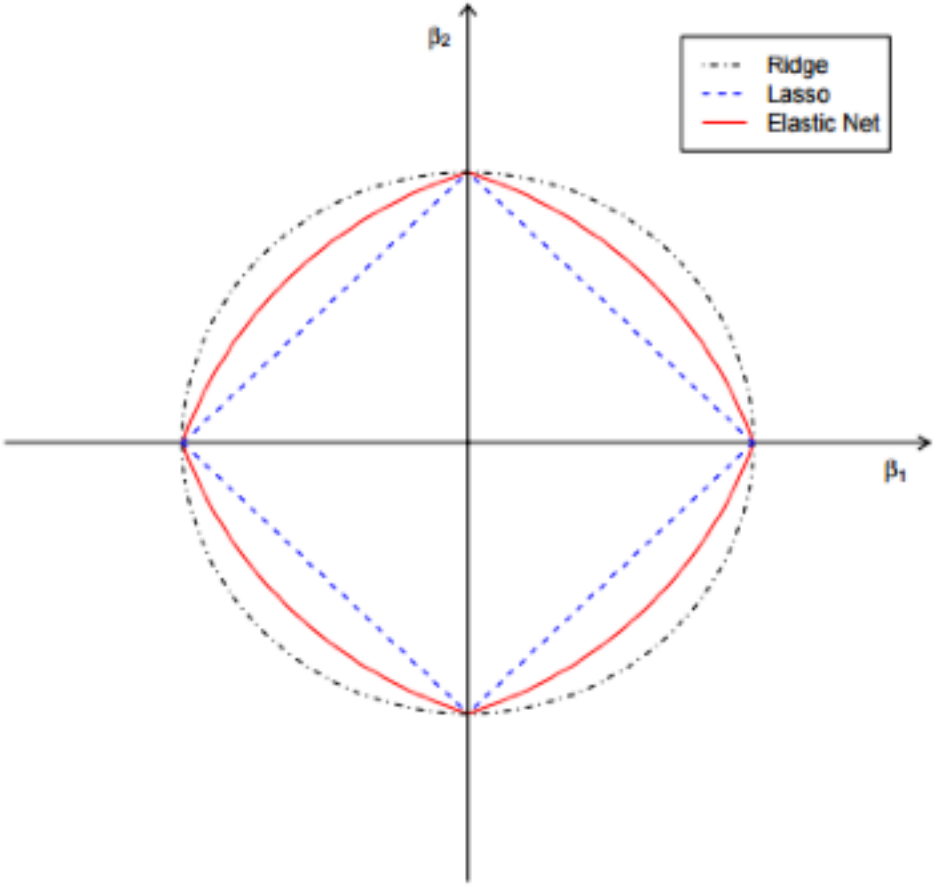
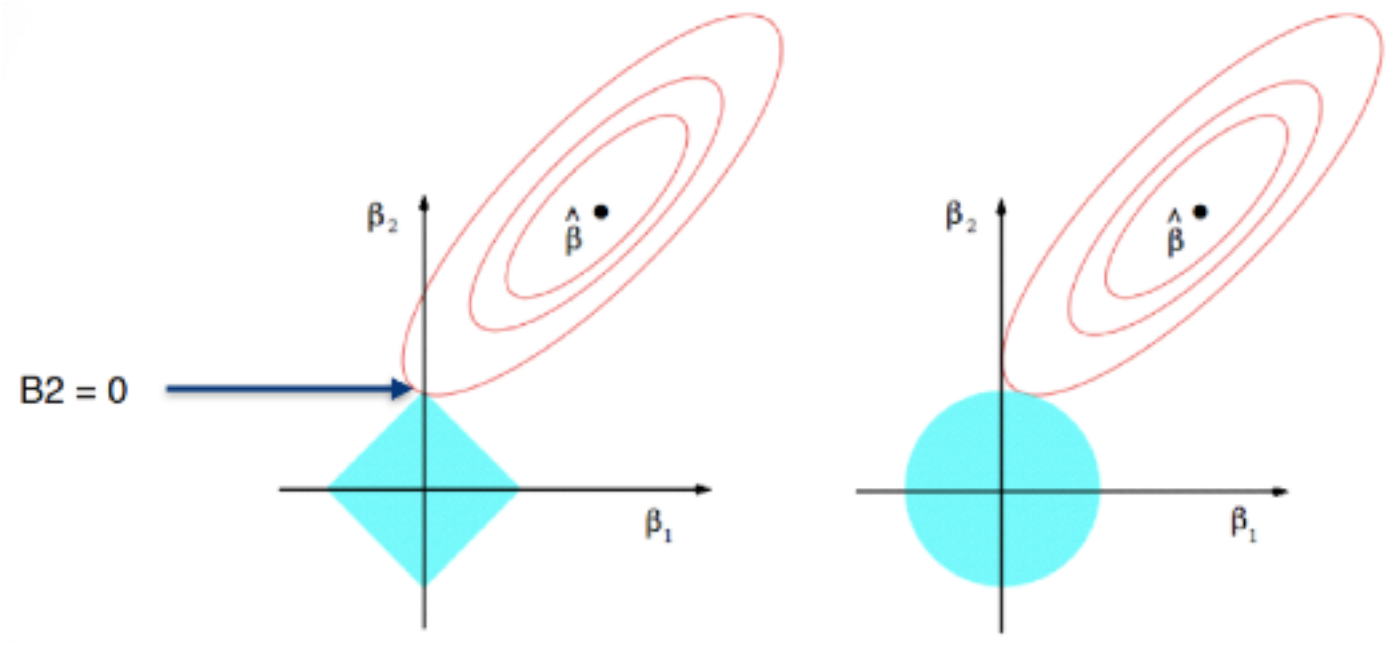

H2O.ai

Constraints for L1, L2 and L1&L2


β2
β1
Ridge
Lasso
Elastic Net

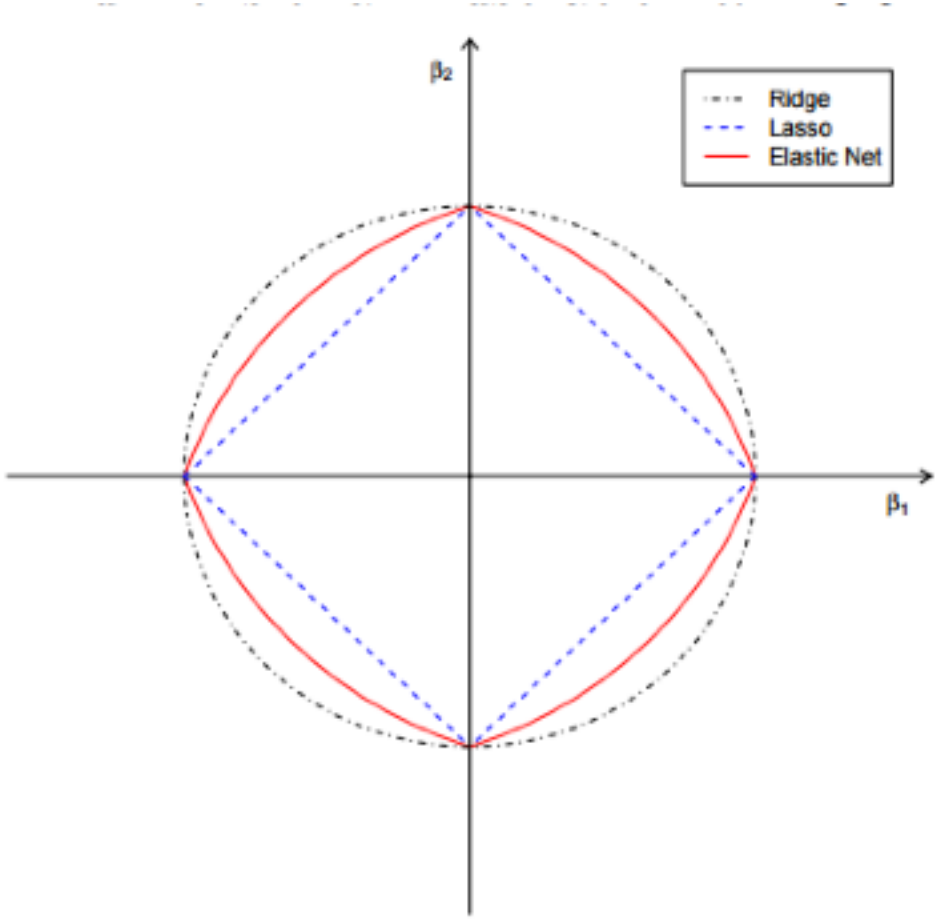

B2 = 0
β2
β̂
β1
β2
β̂
β1

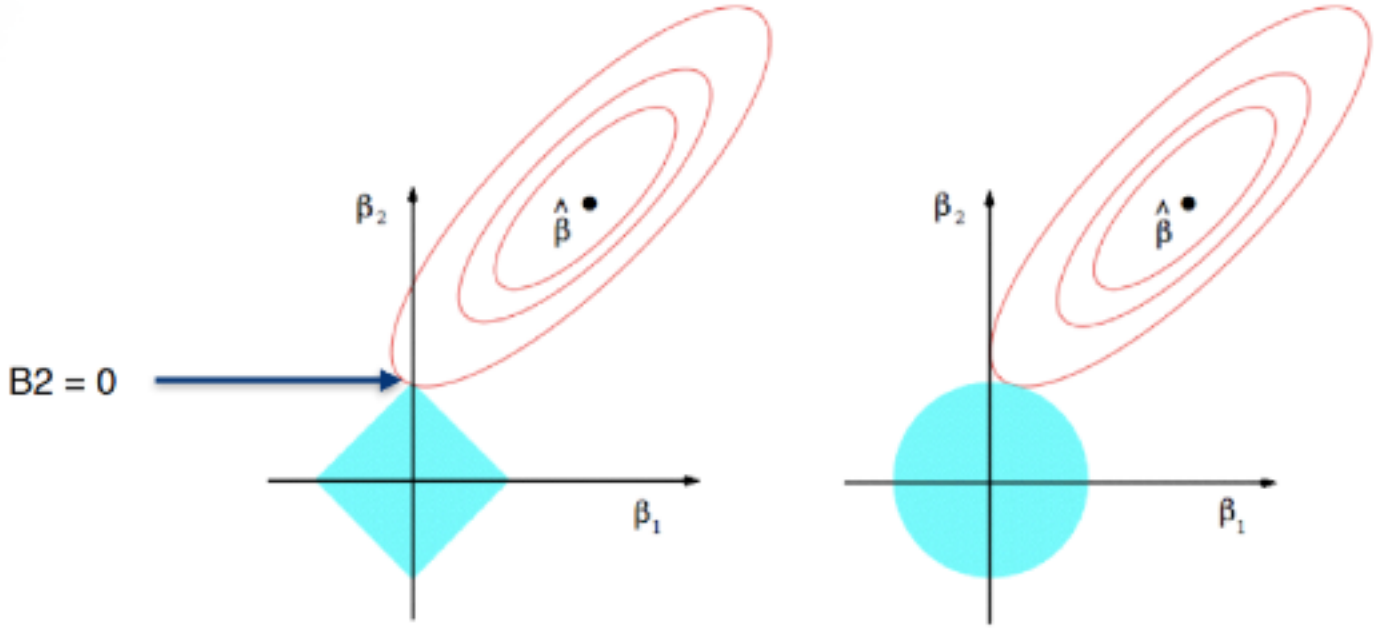
LASSO Regression vs. Ridge Regression

GLM Best Practices

- Regularization Selection
  - Explore a few values for alpha, e.g. 0.01, 0.25, 0.5, 0.75, 0.99
- Wide Data Sets (10K+ columns)
  - Iteratively Reweighted Least Squares (IRLS) fails with lambda = 0
    - IRLS requires p x p Hessian matrix, where p = # of coefficients
    - Could use Limited-memory BFGS (L-BFGS)
  - IRLS + lambda search works and is recommended
    - Use alpha >> 0
    - Can produce 1K+ non-zero coefficients
  - L-BFGS + L2 penalty works
  - L-BFGS + L1 penalty works, but may take a long time

Constraints for L1, L2 and L1&L2





LASSO Regression vs. Ridge Regression