Assignment 9

Problem 1

Define

$$J_1(\boldsymbol{w}) = |\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}|^2 + \lambda_2|\boldsymbol{w}|^2 + \lambda_1|\boldsymbol{w}|_1$$

and

$$J_2(\boldsymbol{w}) = |\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{X}}\boldsymbol{w}|^2 + c\lambda_1|\boldsymbol{w}|_1$$

where $c = (1 + \lambda_2)^{-\frac{1}{2}}$ and

$$\tilde{\boldsymbol{X}} = c\begin{pmatrix}\boldsymbol{X} \\ \sqrt{\lambda_2}\boldsymbol{I}_d\end{pmatrix}, \tilde{\boldsymbol{y}} = \begin{pmatrix}\boldsymbol{y} \\ \boldsymbol{0}_{d\times 1}\end{pmatrix}$$

Show

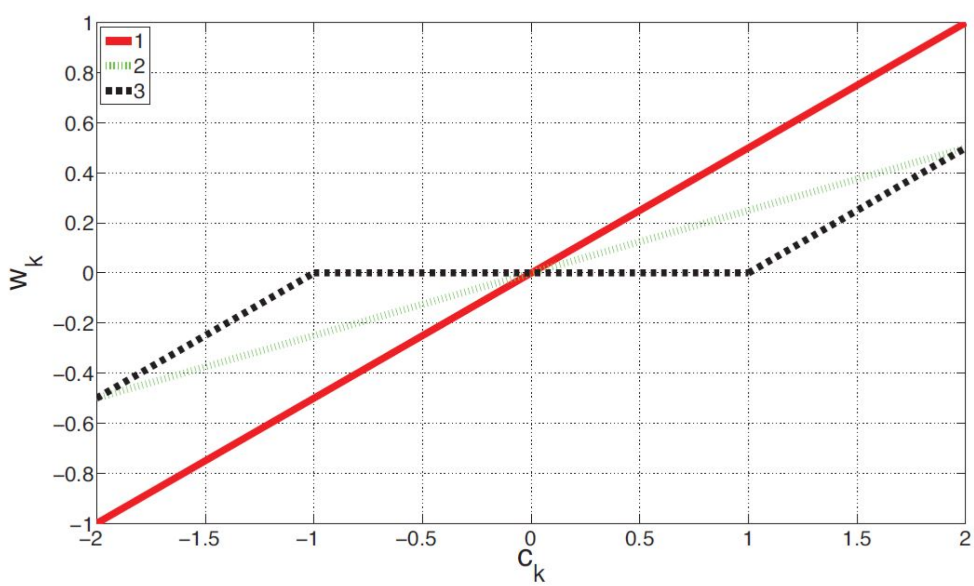
$$\arg\min J_1(\boldsymbol{w}) = c(\arg\min J_2(\boldsymbol{w}))$$

i.e. $J_1(c\boldsymbol{w}) = J_2(\boldsymbol{w})$ and hence that one can solve an elastic net problem using a lasso solver on modified data.

Problem 2

Consider performing linear regression with an orthonormal design matrix, so $||\boldsymbol{x}_{:,k}||_2^2 = 1$ for each column (feature) k, and $\boldsymbol{x}_{:,k}^T\boldsymbol{x}_{:,j} = 0$ so we can estimate each parameter w_k separately.

The following figure plots $\tilde{w_k}$ vs $c_k = 2\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{x}_{:,k}$, the correlation of feature k with the response, for 3 different esimation methods: ordinary least squares (OLS), ridge regression with parameter λ_2, and lasso with parameter λ_1.



Assignment 9

1. Unfortunately we forgot to label the plots. Which method does the solid (1), dotted (2) and dashed (3) line correspond to?

2. What is the value of λ_1?

3. What is the value of λ_2?

Problem 3

Please write one thing from this course so far that you found confusing, a topic you would like to hear more about, or something you found particularly interesting.