CS/ECE/ME532 Activity 9

Estimated Time: 25 minutes for P1, 30 minutes for P2

1. Consider the system of linear equations $\boldsymbol{X}\boldsymbol{w} = \boldsymbol{y}$ where $\boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$, $\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$.

a) Sketch the set of all $\boldsymbol{w}$ that satisfy $\boldsymbol{X}\boldsymbol{w} = \boldsymbol{y}$ in the w_1-w_2 plane. Is the solution unique? What is the value of the squared error $\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$?

b) Use your sketch to find the $\boldsymbol{w}$ of minimum norm that satisfies the system of equations: $\min_{\boldsymbol{w}} ||\boldsymbol{w}||_2^2$ subject to $\boldsymbol{X}\boldsymbol{w} = \boldsymbol{y}$. Is this solution unique? What makes it unique? What is the value of the squared error $||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$ at this solution? What is the value of $||\boldsymbol{w}||^2$? *Hint:* The equation $||\boldsymbol{w}||_2^2 = c$ describes a circle in $\mathbb{R}^2$ with radius $\sqrt{c}$.

c) Algebraically find the $\hat{\boldsymbol{w}}$ that solves the Tikhonov-regularized (or ridge regression) problem $\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \{||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2 + \lambda||\boldsymbol{w}||_2^2\}$ as a function of λ. *Hint:* Recall that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

d) Sketch the set solution to the Tikhonov-regularized problem in the w_1-w_2 plane as a function of λ for $0 < \lambda < \infty$. (Consider the solution for different values of λ in that range.) Find the squared error $||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$ and norm squared of the solution, $||\boldsymbol{w}||_2^2$ for $\lambda = 0$ and $\lambda = 5$. Compare the squared error and norm squared of the solution to those in part **b)**.

2. Let $\boldsymbol{X} = \begin{bmatrix} 1 & \gamma \\ 1 & -\gamma \\ 1 & -\gamma \\ 1 & \gamma \end{bmatrix}$.

a) Show that the columns of $\boldsymbol{X}$ are orthogonal to each other for any γ.

b) Express $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}$ where $\boldsymbol{U}$ is a 4-by-2 matrix with orthonormal columns and $\boldsymbol{\Sigma}$ is a 2-by-2 diagonal matrix (the non-diagonal entries are zero).

c) Express the solution to the least-squares problem $\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$ as a function of $\boldsymbol{U}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{y}$.

d) Let $\boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$. Find the weights $\boldsymbol{w}$ as a function of γ. What happens to $||\boldsymbol{w}||_2^2$ as $\gamma \to 0$?

e) The ratio of the largest to the smallest diagonal values in $\boldsymbol{\Sigma}$ is termed the condition number of $\boldsymbol{X}$. Find the condition number if $\gamma = 0.1$ and $\gamma = 10^{-8}$. Also find $||\boldsymbol{w}||_2^2$ for these two values of γ.

f) A system of linear equations with a large condition number is said to be "ill-conditioned". One consequence of an ill-conditioned system of equations is solutions with large norms as you found in the previous part of this problem. A second consequence is that the solution is very sensitive to small errors in $\boldsymbol{y}$ such as may result from measurement error or numerical error. Suppose $\boldsymbol{y} = \begin{bmatrix} 1+\epsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$. Write $\boldsymbol{w} = \boldsymbol{w}_o + \boldsymbol{w}_\epsilon$ where $\boldsymbol{w}_o$ is the solution for arbitrary γ when $\epsilon = 0$ and $\boldsymbol{w}_\epsilon$ is the perturbation in that solution due to some error $\epsilon \neq 0$. How does the norm of the perturbation due to $\epsilon \neq 0$, $||\boldsymbol{w}_\epsilon||_2^2$, depend on the condition number? Find $||\boldsymbol{w}_\epsilon||_2^2$ for $\epsilon = 0.01$ and $\gamma = 0.1$ and $\gamma = 10^{-8}$.

g) Now apply ridge regression, i.e., Tikhonov reqularization. Solve for $\boldsymbol{w}_o$ and $\boldsymbol{w}_\epsilon$ as a function of λ. Find $||\boldsymbol{w}_o||_2^2$ and $||\boldsymbol{w}_\epsilon||_2^2$ for $\lambda = 0.1$, $\epsilon = 0.01$ and $\gamma = 0.1$ and $\gamma = 10^{-8}$. Comment on the impact of regularization.