Recitation Week 12

Ashwin Bhola

CDS, NYU

Nov 20th , 2019

Linear Regression

1. When solving the least squares problem, the optimization problem is: $\min_x ||Ax - b||^2$
    1. Geometrically, what does this mean?
    2. How can we obtain the equation $A^T Ax = A^T b$ from this geometric intuition?
    3. If $\mathrm{x} = P_{Im(A)} b$, what is $P_{Im(A)}$?
    4. Under what conditions is $A^T A$ invertible? If $A^T A$ is not invertible, must the normal equations still have a solution?
    5. The objective of linear regression is to find a hyperplane of the form $f(x) = w_0 + w^T x = w_0 + w_1 x_1 + \cdots w_d x_d$. Solving least squares gives us $w \in \mathbb{R}^d$. How do we find $w_0 \in \mathbb{R}$?
    6. One of the things we expect from machine learning models is regularity in prediction i.e. neighboring points in feature space should have similar prediction. How does regularization induce this regularity?
    7. In ridge regression, we add a penalty $\lambda||x||_2^2$. Is it better than adding $\lambda||x||_2$?
    8. What are the disadvantages of using lasso regularization?

Linear Regression

1. Let $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times 2}$ have the form

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}$$

   Let $x \in \mathbb{R}^2$ be the minimizer of the least squares objective function $\min_x ||Ax - b||^2$. Prove that $\sum_{i=1}^{n} (Ax - b)_i = 0$

2. Suppose $M \in \mathbb{R}^{2 \times 2}$ has positive eigenvalues. Describe geometrically the contour lines of $f: \mathbb{R}^2 \to \mathbb{R}$ given by $f(x) = x^T M x$. Recall that the contour line for value γ is given by $\{x \in \mathbb{R}^2 \mid f(x) = \gamma\}$

Linear Regression

Suppose $A \in \mathbb{R}^{n\times 2}$ is your data matrix. The empirical risk or the least squares loss is given by $\mathrm{R(w)} = ||Aw - y||^2$. Suppose $\widehat{w} = argmin_w\, R(w)$. Let $\mathrm{M} = A^T A \in \mathbb{R}^{2\times 2}$ be invertible. It can be easily verified that

$$x^T Mx - 2b^T x = (x - M^{-1}b)M(x - M^{-1}b) - b^T M^{-1} b$$

This leads to $R(w) - R(\widehat{w}) = (w - \widehat{w})^T A^T A(w - \widehat{w})$

Describe geometrically the contour lines of $f: \mathbb{R}^2 \to \mathbb{R}$ given by $f(w) = R(w) - R(\widehat{w})$