STAT 542/CS 598

Solution for One-Variable Lasso Problem

Consider two vectors:

$$\mathbf{v}_{n\times 1} = \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix}, \quad \mathbf{z}_{n\times 1} = \begin{pmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{pmatrix},$$

where $\mathbf{v}_{n\times 1}$ is the response vector and $\mathbf{z}_{n\times 1}$ is the feature vector. We are interested in solving the following one-variable Lasso optimization problem:

$$\boxed{\min_b \frac{1}{2n} \|\mathbf{v} - b \cdot \mathbf{z}\|^2 + \lambda |b|} \tag{1}$$

where $\lambda > 0$ and b is the coefficient we seek to determine.

1. Relationship with Earlier Optimization Problem

In a previous class, we discussed how to minimize the function

$$\boxed{f(x) = (x - a)^2 + \eta |x|.}$$

The minimizer, denoted by x^*, is:

$$x^* = \arg\min_x f(x) = \mathrm{sign}(a)(|a| - \eta/2)_+ = \begin{cases} a - \eta/2, & \text{if } a > \eta/2 \\ 0, & \text{if } |a| \leq \eta/2 \\ a + \eta/2, & \text{if } a < -\eta/2 \end{cases} \tag{2}$$

Next we will write (1) in the form of $f(x)$ and then use the solution above.

2. Connection to Least Squares Estimate

Let's first acknowledge that the term $\|\mathbf{v} - b \cdot \mathbf{z}\|^2$ in (1) is the Residual Sum of Squares (RSS) for a simple regression model without intercept, where the response $\mathbf{v}$ is regressed against the feature $\mathbf{z}$, and b is the coefficient for this feature.

The Least Squares (LS) estimate for b is then given by

$$\hat{b} = \frac{\mathbf{v}^t \mathbf{z}}{\|\mathbf{z}\|^2}.$$

3. Expanding and Simplifying the Objective

We can rewrite and simplify the RSS term as follows:

$$\|\mathbf{v} - b \cdot \mathbf{z}\|^2 = \|\mathbf{v} - \hat{b} \cdot \mathbf{z}\|^2 + \|(\hat{b} - b) \cdot \mathbf{z}\|^2. \tag{3}$$

The detailed derivation for this simplification is provided below.

$$\begin{aligned}
\|\mathbf{v} - b \cdot \mathbf{z}\|^2 &= \|\mathbf{v} - \hat{b} \cdot \mathbf{z} + \hat{b} \cdot \mathbf{z} - b \cdot \mathbf{z}\|^2 \\
&= \|\mathbf{v} - \hat{b} \cdot \mathbf{z} + (\hat{b} - b) \cdot \mathbf{z}\|^2 \\
&= \|\mathbf{v} - \hat{b} \cdot \mathbf{z}\|^2 + \|(\hat{b} - b) \cdot \mathbf{z}\|^2 + \\
&\qquad 2 \times \text{inner-product-between } (\mathbf{v} - \hat{b} \cdot \mathbf{z}) \text{ and } (\hat{b} - b) \cdot \mathbf{z}
\end{aligned}$$

where the inner-product term disappears due to the orthogonality between the residual and the feature vector $\mathbf{z}$[1].

4. Reduced Optimization Problem

Since the first term in (3) does not depend on b, it doesn't affect the minimization. Our problem is then equivalent to minimizing:

$$\boxed{(b - \hat{b})^2 + \frac{2n\lambda}{\|\mathbf{z}\|^2}|b|.}$$

This is because

$$\begin{aligned}
\frac{1}{2n}\|(\hat{b} - b) \cdot \mathbf{z}\|^2 + \lambda|b| &= \frac{\|\mathbf{z}\|^2}{2n}(\hat{b} - b)^2 + \lambda|b| \\
&= \frac{\|\mathbf{z}\|^2}{2n}\left((b - \hat{b})^2 + \frac{2n\lambda}{\|\mathbf{z}\|^2}|b|\right) \\
&\propto (b - \hat{b})^2 + \frac{2n\lambda}{\|\mathbf{z}\|^2}|b|.
\end{aligned}$$

Now, we can utilize (2), our earlier solution for $f(x)$, by setting:

$$a = \frac{\mathbf{v}^t\mathbf{z}}{\|\mathbf{z}\|^2}, \quad \eta = \frac{2n\lambda}{\|\mathbf{z}\|^2}.$$

By doing so, we can immediately find the minimizer b for our original one-variable Lasso problem.

[1] Recall that the residual vector is orthogonal to each column of the design matrix. Here, $(\mathbf{v} - \hat{b} \cdot \mathbf{z})$ represents the residual vector from a regression model with $\mathbf{z}$ being a column (actually the only column) of the design matrix, and therefore $(\mathbf{v} - \hat{b} \cdot \mathbf{z})$ is orthogonal to $\mathbf{z}$.