One-Variable Lasso

Let $\mathbf{r}_{n\times 1} = (r_1, \ldots, r_n)^t$ and $\mathbf{z}_{n\times 1} = (z_1, \ldots, z_n)^t$ denote the response and feature vectors, respectively. Consider the following one-variable Lasso problem

$$\min_b \frac{1}{2n} \sum_{i=1}^{n} (r_i - bz_i)^2 + \lambda|b| = \min_b \frac{1}{2n} \|\mathbf{r} - b \cdot \mathbf{z}\|^2 + \lambda|b| \tag{1}$$

where $\lambda > 0$ and b is the coefficient associated with the feature vector $\mathbf{z}$.

How to solve (1)? In class we have discussed how to find the minimizer of

$$f(x) = (x - a)^2 + \lambda|x|,$$

which is given by

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

We can write (1) in the form of $f(x)$ and then use the solution above.

First note that the term $\|\mathbf{r} - b \cdot \mathbf{z}\|^2$ in (1) looks like the RSS from a regression model with only one predictor (whose coefficient is b) without the intercept. The corresponding LS estimate is given by

$$\hat{b} = \mathbf{r}^t \mathbf{z} / \|\mathbf{z}\|^2.$$

Then we have

$$\begin{aligned} \|\mathbf{r} - b \cdot \mathbf{z}\|^2 &= \|\mathbf{r} - \hat{b} \cdot \mathbf{z} + \hat{b} \cdot \mathbf{z} - b \cdot \mathbf{z}\|^2 \\ &= \|\mathbf{r} - \hat{b} \cdot \mathbf{z} + (\hat{b} - b) \cdot \mathbf{z}\|^2 \\ &= \|\mathbf{r} - \hat{b} \cdot \mathbf{z}\|^2 + \|(\hat{b} - b) \cdot \mathbf{z}\|^2 + \\ &\qquad 2 \times \text{inner-product-between } (\mathbf{r} - \hat{b} \cdot \mathbf{z}) \text{ and } (\hat{b} - b) \cdot \mathbf{z} \\ &= \|\mathbf{r} - \hat{b} \cdot \mathbf{z}\|^2 + \|(\hat{b} - b) \cdot \mathbf{z}\|^2, \end{aligned} \tag{3}$$

where the last equality is due to the fact that the inner product term is zero since the two vectors are orthogonal[1].

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

Note that the first term of (3) has nothing do with b. So to minimize (1) with respect to b, we can ignore the first term and instead minimize

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

Now we can use (2), the solution we have derived for $f(x)$, with

$$a \leftarrow \mathbf{r}^t\mathbf{z}/\|\mathbf{z}\|^2, \quad \lambda \leftarrow 2n\lambda/\|\mathbf{z}\|^2.$$