A Note on the Uniqueness of the Lasso Solution

Alessandro Rinaldo*
Department of Statistics
Carnegie Mellon University

Abstract

In this note we show that, if β_1 and β_2 are two distinct solutions to the lasso problem $\min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 + \lambda\|\beta\|_1$ for some $n \times p$ matrix X with $p > n$, then $\mathrm{X}\beta_1 = \mathrm{X}\beta_2$.

Consider the lasso problem,

$$\min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 + \lambda\|\beta\|_1, \tag{1}$$

where X is a $n \times p$ design matrix. For each solution $\widehat{\beta}$ to the unconstrained problem (1), there exists a value $t > 0$ such that $\widehat{\beta}$ solves the constrained program

$$\begin{array}{c} \min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 \\ \text{s.t } \ \|\beta\|_1 \leq t, \end{array}$$

with $\|\widehat{\beta}\|_1 = t$ (see Osborne et al., 2000). If $p > n$, then there may not be a unique solution.

Proposition *Let β_1, β_2 be two distinct solutions to (1). Then,* $\mathrm{X}\beta_1 = \mathrm{X}\beta_2$.

Proof. Using results from Osborne et al. (2000), one can verify that

$$\|r_1\|_2^2 = \|r_2\|_2^2 \tag{2}$$

and

$$r_i^\top \mathrm{X}\beta_i = \lambda\|\beta_i\|_1 = \lambda t, \tag{3}$$

where $r_i = y - \mathrm{X}\beta_i$, for $i = 1, 2$. Because of (2), we have

$$0 = \langle y - \mathrm{X}\beta_1, y - \mathrm{X}\beta_1\rangle - \langle y - \mathrm{X}\beta_2, y - \mathrm{X}\beta_2\rangle,$$

which implies

$$\beta_1^\top \mathrm{X}^\top \mathrm{X}\beta_1 - \beta_2^\top \mathrm{X}^\top \mathrm{X}\beta_2 = 2y^\top \mathrm{X}(\beta_1 - \beta_2). \tag{4}$$

*Email: arinaldo@stat.cmu.edu

Rewrite (2) in a different form as

$$\begin{aligned} 0 &= \langle y - \mathrm{X}\beta_1 - y + \mathrm{X}\beta_2, y - \mathrm{X}\beta_1 + y - \mathrm{X}\beta_2 \rangle \\ &= \langle \mathrm{X}(\beta_2 - \beta_1), r_1 + r_2 \rangle \\ &= r_1^\top \mathrm{X}\beta_2 - r_1^\top \mathrm{X}\beta_1 + r_2^\top \mathrm{X}\beta_2 - r_2^\top \mathrm{X}\beta_1 \\ &= r_1^\top \mathrm{X}\beta_2 - r_2^\top \mathrm{X}\beta_1 \\ &= y^\top \mathrm{X}(\beta_1 - \beta_2), \end{aligned} \tag{5}$$

where, in the forth equality, we use $r_1^\top \mathrm{X}\beta_1 = r_2^\top \mathrm{X}\beta_2$, stemming from (3). Combining (4) and (5), we get

$$\beta_1^\top \mathrm{X}^\top \mathrm{X}\beta_1 = \beta_2^\top \mathrm{X}^\top \mathrm{X}\beta_2, \tag{6}$$

Because of the convexity of the set of all lasso solutions (see, again, Osborne et al., 2000), for any $0 < \lambda < 1$, the vector $\beta_3 = \lambda\beta_1 + (1-\lambda)\beta_2$ is also a solution to (1). By the same arguments leading to (6), we obtain

$$\beta_1^\top \mathrm{X}^\top \mathrm{X}\beta_1 = \beta_2^\top \mathrm{X}^\top \mathrm{X}\beta_2 = \beta_3^\top \mathrm{X}^\top \mathrm{X}\beta_3.$$

Then, letting $y_i = \mathrm{X}\beta_i$ for $i = 1, 2, 3$ and $K = \|y_1\|_2^2$, we get $y_3 = \lambda y_1 + (1-\lambda)y_2$ and

$$\|y_1\|_2^2 = \|y_2\|_2^2 = \|y_3\|_2^2 = K,$$

which gives a contradiction unless $y_1 = y_2$, because the $n-1$ sphere $\{x \in \mathbb{R}^n \colon \|x\|_2^2 = K\}$ is not a convex set. The claim is proved.

■

References

Osbirne, M., Presnell, B. and Turlach, B. A. (2000). On the LASSO and Its Dual. Journal of Computational and Graphical Statistics, 9 (2), 319–337.