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 \( \beta_1 \) and \( \beta_2 \) 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 \( X\beta_1 = X\beta_2 \).

Consider the lasso problem,
\[
\min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1
\] (1)
where \( X \) is a \( n \times p \) design matrix. For each solution \( \hat{\beta} \) to the unconstrained problem (1), there exists a value \( t > 0 \) such that \( \hat{\beta} \) solves the constrained program
\[
\min_{\beta \in \mathbb{R}^p} \|y - X\beta\|_2^2 \\
\text{s.t. } \|\beta\|_1 \leq t,
\]
with \( \|\hat{\beta}\|_1 = t \) (see Osborne et al., 2000). If \( p > n \), then there may not be a unique solution.

**Proposition** Let \( \beta_1, \beta_2 \) be two distinct solutions to (1). Then, \( X\beta_1 = X\beta_2 \).

**Proof.** Using results from Osborne et al. (2000), one can verify that
\[
\|r_1\|_2^2 = \|r_2\|_2^2
\] (2)
and
\[
r_i^\top X\beta_i = \lambda \|\beta_i\|_1 = \lambda t,
\] (3)
where \( r_i = y - X\beta_i \), for \( i = 1, 2 \). Because of (2), we have
\[
0 = \langle y - X\beta_1, y - X\beta_1 \rangle - \langle y - X\beta_2, y - X\beta_2 \rangle,
\]
which implies
\[
\beta_1^\top X^\top X\beta_1 - \beta_2^\top X^\top X\beta_2 = 2y^\top X(\beta_1 - \beta_2).
\] (4)

∗Email: arinaldo@stat.cmu.edu
Rewrite (2) in a different form as

\[
\begin{align*}
0 &= \langle y - X\beta_1 - y + X\beta_2, y - X\beta_1 + y - X\beta_2 \rangle \\
&= \langle X(\beta_2 - \beta_1), r_1 + r_2 \rangle \\
&= r_1^T X\beta_2 - r_1^T X\beta_1 + r_2^T X\beta_2 - r_2^T X\beta_1 \\
&= r_1^T X\beta_2 - r_2^T X\beta_1 \\
&= y^T X(\beta_1 - \beta_2),
\end{align*}
\]

(5)

where, in the fourth equality, we use \( r_1^T X\beta_1 = r_2^T X\beta_2 \), stemming from (3). Combining (4) and (5), we get

\[
\beta_1^T X^T X\beta_1 = \beta_2^T X^T X\beta_2,
\]

(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

\[
\beta_1^T X^T X\beta_1 = \beta_2^T X^T X\beta_2 = \beta_3^T X^T X\beta_3.
\]

Then, letting \( y_i = X\beta_i \) 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 \( \{ x \in \mathbb{R}^n : \|x\|_2^2 = K \} \) is not a convex set. The claim is proved.

\[\blacksquare\]

References