The Georeg Regularizer

Guillermo Angeris Shane Barratt Jonathan Tuck*†

March 4, 2021

Abstract

In this paper, we introduce the *georeg* regularizer as a natural generalization of the elastic net regularizer by combining *all the norms*. We present a simple example showing some interesting properties of the regularizer and introduce a new `cvxpy` atom, `geo_reg`, which implements the georeg regularizer.

1 Introduction

We define the p-norm of a vector $x \in \mathbf{R}^n$ to be[1]

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

Note that any p-norm (and its pth power) is always convex [1]. We will define the mth georeg regularizer as the following convex combination of powers of norms:

$$g_m(x, \alpha) = \left(\sum_{k=1}^{m} \alpha^{-k} \right)^{-1} \sum_{k=1}^{m} \frac{\|x\|_k^k}{\alpha^k} = \left(\frac{1-\alpha^{m+1}}{1-\alpha^{-1}} - 1 \right)^{-1} \sum_{i=1}^{m} \left(\frac{1 - \left(\frac{|x_i|}{\alpha}\right)^{m+1}}{1 - \frac{|x_i|}{\alpha}} - 1 \right), \quad \alpha \neq 1.$$

We note (a) that $g_m(\cdot, \alpha)$ is convex in its domain for any $\alpha > 0$ as it is a sum of convex functions and (b) that the georeg regularizer reduces to the elastic net regularizer if we choose $m = 2$. Additionally, the series is absolutely convergent as $m \uparrow \infty$ whenever $|x_i| < \alpha$ and $\alpha > 1$ for all $i = 1, \ldots, m$.

In this paper, we focus on the $m = \infty$ case which reduces to

$$g_\infty(x, \alpha) = \sum_{i=1}^{m} (\alpha - 1) \left(\frac{1}{1 - \frac{|x_i|}{\alpha}} - 1 \right),$$

*Authors listed in alphabetical order.

†`{angeris, sbarratt, jonathantuck}@stanford.edu`

[1] This is well known, but we needed to make the paper longer.

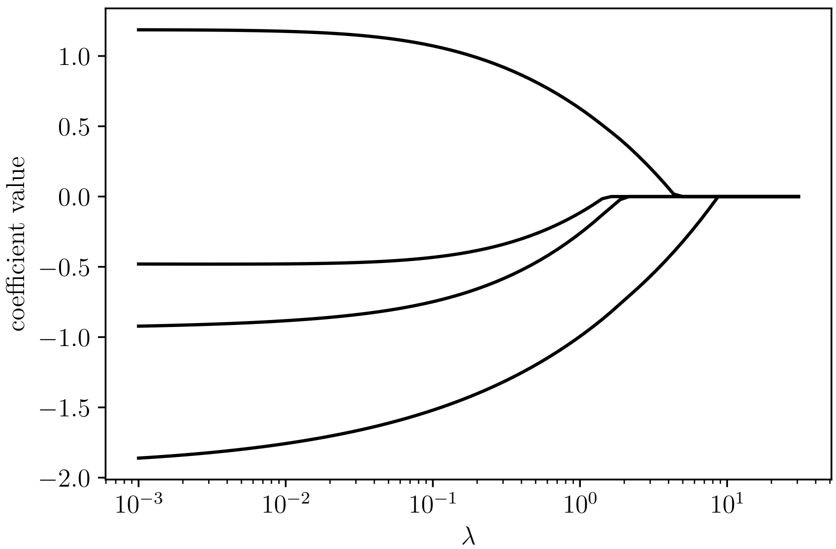

Figure 1: Regularization path.

and is easily representable by a set of second-order cone constraints (SOCs). Let $f : \mathbf{R}^n \to \mathbf{R}$ be a loss function, then the learning problem for x can be written as the following problem with SOCs and variables $x, r, t \in \mathbf{R}^n$:

$$\begin{array}{ll} \text{minimize} & f(x) + \mathbf{1}^T r - n, \\ \text{subject to} & \|(2, r_i - 1 + t_i/\alpha)\|_2 \le r_i + 1 - t_i/\alpha, \quad i = 1, \ldots, n \\ & -t \le x \le t. \end{array} \tag{1}$$

If f is convex and easily representable in conic form, the problem can be written as a conic program, which is often efficiently solvable [1].

2 Example

As an example, we solve the problem

$$\text{minimize} \quad \|Ax - b\|_2^2 + \lambda g_\infty(x, \alpha), \tag{2}$$

where $x \in \mathbf{R}^4$ is the optimization variable, and $A \in \mathbf{R}^{100\times 4}$ and $b \in \mathbf{R}^{100}$ are problem data. We randomly generate A and b and plot the values of the optimal x for varying values of $\lambda > 0$ in figure 1.

References

[1] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.