Printable Topics in Mathematics

Benjamin Grimmer
Johns Hopkins, AMS

A Collection of SCAD Level Sets

An Example of Nonconvex Regularization for Inducing Sparsity in Solutions

Printable Topics in Mathematics Project #9 - First Posted 1/3/2023

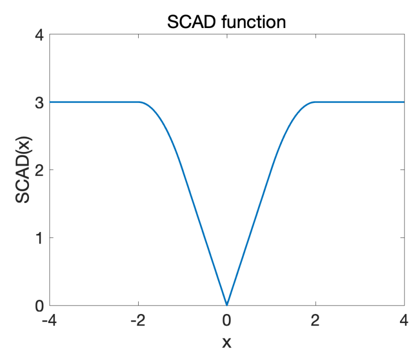
REGULARIZATION is typically incorporated into optimization problems either by adding a simple function $\lambda r(x)$ to the objective or equivalently by constraining $r(x) \leq \eta$. Common convex regularizers include 2-norm regularization to improve problem conditioning, 1-norm regularization to induce sparsity, and Schatten matrix 1-norm (i.e., the nuclear norm) regularization to induce low-rank solutions. Nonconvex regularizers have been employed to great success in statistics and machine learning and are almost always star-convex.

SCAD Regularization

One of the simplest nonconvex regularizers is the Smoothly Clipped Absolute Deviation (SCAD) function which sums up piecewise quadratic clipped absolute values in each coordinate

$$SCAD(x_i) = \begin{cases} 2|x_i| & 0 \leq |x_i| \leq 1, \\ -x_i^2 + 4|x_i| - 1 & 1 < |x_i| \leq 2, \\ 3 & |x_i| > 2. \end{cases}$$

Adding this to an objective being minimized in an optimization problem encourages coordinates to be set to zero while not excessively penalizing large nonzero coordinates. Then for any vector x, we say $SCAD(x) = \sum_i SCAD(x_i)$.



Other Nonconvex Regularizers

As further examples, q-norm and Schatten matrix q-norm regularization for any $0 < q < 1$ tend to induce sparse or low-rank solutions respectively. Figure 1 shows examples of the star-convex constraint sets $\{(x, y, z) \mid \|(x, y, z)\|_q \leq 1\}$ and $\left\{(x, y, z) \mid \left\| \begin{bmatrix} x & y \\ y & z \end{bmatrix} \right\|_q \leq 1 \right\}$. The extreme points of these printed model sets verify their sparsity/low-rank inducing behaviors as the vector norm ball has extreme points given by all 1-sparse vectors and the matrix norm ball has extreme points given by all rank one matrices (here two rings, one composed of all positive semidefinite rank one matrices and one composed of all negative semidefinite rank one matrices).

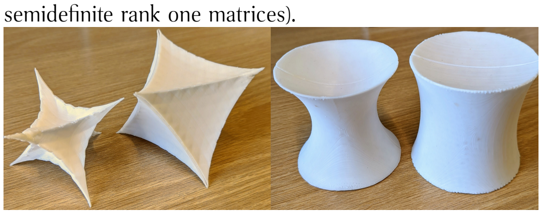


Figure 1. $q = 0.5,\ 0.75$–Vector "Norm" Unit Balls (Left) and Schatten Matrix "Norm" Unit Balls (Right).

3D Model Viewer and Links

- ❑ `.stl` files and 3D model viewer at `printables.com/model/358708`
- ❑ This `.pdf` is available at `ams.jhu.edu/~grimmer/SCAD.pdf`

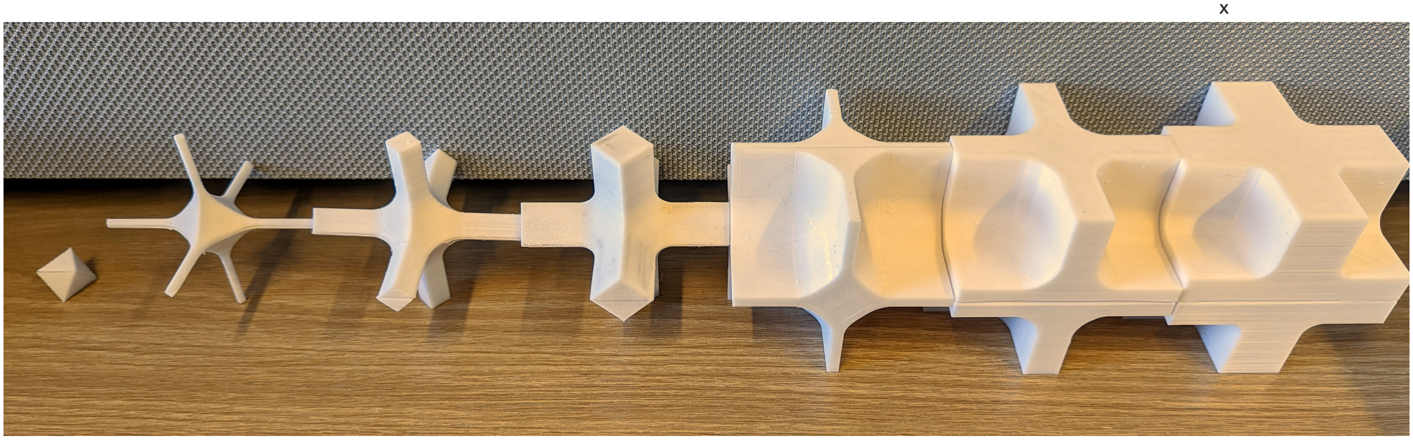


Figure 2. Constraint sets $\{(x, y, z) \mid SCAD(x, y, z) \leq \eta\}$ for increasing values of $\eta = 2.5,\ 3.5,\ \ldots,\ 7.5,\ 8.5$. This demonstration shows the set as η grows: the set is bounded for $\eta \in [0, 3)$, then contains all coordinate axis allowing one large component for $\eta \in [3, 6)$, then contains all coordinate planes allowing two large components for $\eta \in [6, 9)$, and finally contains all of $\mathbb{R}^3$ for $\eta \geq 9$.