Closed-form Solutions for Penalty Methods

We want to penalize a single-variable x (using LASSO, SCAD, MCP) or a vector $\boldsymbol{x}$ (using gLASSO, gSCAD, gMCP). Let a be a fixed value and $\boldsymbol{a}$ be a fixed vector, we want to minimize the following function

$$\mathcal{L}(x;\lambda) = \frac{\theta}{2}(x-a)^2 + P(x;\lambda)$$

or

$$\mathcal{L}(\boldsymbol{x};\lambda) = \frac{\theta}{2}(\boldsymbol{x}-\boldsymbol{a})^2 + P(\boldsymbol{x};\lambda)$$

where λ and θ are positive fixed values and $P(\cdot)$ is a penalty term.

1 LASSO

Consider $L1$-norm penalty with a single variable x, that is

$$P(x;\lambda) = \lambda|x|.$$

We want to compute

$$x^* = \arg\min_x \mathcal{L}(x;\lambda) = \arg\min_\theta \left\{\frac{\theta}{2}(x-a)^2 + \lambda|x|\right\}.$$

We consider the following two cases.

- **CASE** I. If $a=0$, it is obvious that $x^*=0$.
- **CASE** II. If $a \neq 0$, we observe that
  - When $x=0$, we have $\mathcal{L}(0;\lambda) = \frac{1}{2}\theta a^2$.
  - When $x \neq 0$. WLOG, assume that $\operatorname{sign}(x) = \operatorname{sign}(a)$. (If $\operatorname{sign}(x) = -\operatorname{sign}(a)$, we can replace x by $-x$ which can make the objective function decrease.) Note that $x = |x|\operatorname{sign}(x) = |x|\operatorname{sign}(a)$, we obtain

$$\begin{aligned}\mathcal{L}(x;\lambda) &= \frac{\theta}{2}\left(|x|\operatorname{sign}(a) - |a|\operatorname{sign}(a)\right)^2 + \lambda|x| \\ &= \frac{\theta}{2}(|x|-|a|)^2 + \lambda|x| \\ &= \frac{\theta}{2}x^2 + (\lambda - \theta|a|)|x| + \frac{1}{2}\theta a^2.\end{aligned}$$

    * If $\lambda - \theta|a| \geq 0$, it obvious that $\mathcal{L}(x;\lambda) > \mathcal{L}(0;\lambda) = \frac{1}{2}\theta a^2$.
    * If $\lambda - \theta|a| < 0$, it obvious that $|x^*| = \frac{\theta|a|-\lambda}{\theta} = |a| - \frac{\lambda}{\theta}$ can minimize the objective function and the minimum is smaller than $\frac{1}{2}\theta a^2$.

Summarize the discussion above, we have

$$x^* = \operatorname{sign}(a)\left(|a| - \frac{\lambda}{\theta}\right)_+.$$

Let $S(x,t) = \operatorname{sign}(x)\left(|x| - t\right)_+ = x\left(1 - \frac{t}{|x|}\right)_+$ be the soft thresholding operator. The closed-form solution is

$$x^* = S\left(a, \frac{\lambda}{\theta}\right). \tag{1}$$

2 SCAD

Consider Smoothly Clipped Absolute Deviations (SCAD) penalty with a single variable x and parameter $\gamma > 1 + \theta^{-1}$, that is

$$P_\gamma(x;\lambda) = \begin{cases} \lambda|x| & \text{if } |x| \le \lambda \\ \frac{2\gamma\lambda|x| - \left(x^2+\lambda^2\right)}{2(\gamma-1)} & \text{if } \lambda < |x| \le \gamma\lambda \\ \frac{1}{2}(\gamma+1)\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases}$$

We want to compute

$$x^* = \arg\min_x \mathcal{L}_\gamma(x;\lambda),$$

where

$$\mathcal{L}_\gamma(x;\lambda) = \begin{cases} \frac{\theta}{2}(x-a)^2 + \lambda|x| & \text{if } |x| \le \lambda \\ \frac{\theta}{2}(x-a)^2 + \frac{2\gamma\lambda|x| - \left(x^2+\lambda^2\right)}{2(\gamma-1)} & \text{if } \lambda < |x| \le \gamma\lambda \\ \frac{\theta}{2}(x-a)^2 + \frac{1}{2}(\gamma+1)\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases}$$

We consider the following two cases.

- **CASE** I. If $a = 0$, it is obvious that $x^* = 0$.
- **CASE** II. If $a \ne 0$, we observe that
  - When $x = 0$, we have $\mathcal{L}_\gamma(0;\lambda) = \frac{1}{2}\theta a^2$.
  - When $x \ne 0$. WLOG, assume that $\operatorname{sign}(x) = \operatorname{sign}(a)$. (If $\operatorname{sign}(x) = -\operatorname{sign}(a)$, we can replace x by $-x$ which can make the objective function decrease.) Note that $x = |x|\operatorname{sign}(x) = |x|\operatorname{sign}(a)$, we obtain

$$\mathcal{L}_\gamma(x;\lambda) = \begin{cases} \frac{\theta}{2}(|x|-|a|)^2 + \lambda|x| & \text{if } |x| \le \lambda \\ \frac{\theta}{2}(|x|-|a|)^2 + \frac{2\gamma\lambda|x| - \left(x^2+\lambda^2\right)}{2(\gamma-1)} & \text{if } \lambda < |x| \le \gamma\lambda \\ \frac{\theta}{2}(|x|-|a|)^2 + \frac{1}{2}(\gamma+1)\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases}$$

$$\overset{z=|x|}{=} \begin{cases} \frac{\theta}{2}z^2 + (\lambda - \theta|a|)\,z + \frac{1}{2}\theta a^2 \triangleq f_1(z) & \text{if } 0 < z \le \lambda \\ \frac{1}{2}\left(\theta - \frac{1}{\gamma-1}\right)z^2 + \left(\frac{\gamma\lambda}{\gamma-1} - \theta|a|\right)z + \frac{1}{2}\theta a^2 - \frac{1}{2}\frac{\lambda^2}{\gamma-1} \triangleq f_2(z) & \text{if } \lambda < z \le \gamma\lambda \\ \frac{\theta}{2}(z-|a|)^2 + \frac{1}{2}(\gamma+1)\lambda^2 \triangleq f_3(z) & \text{if } z > \gamma\lambda \end{cases}$$

$*$ If $|a| \le \lambda + \frac{\lambda}{\theta}$, we observe that $|a| - \frac{\lambda}{\theta} \le \lambda$ and $\frac{\theta|a| - \frac{\gamma\lambda}{\gamma-1}}{\theta - \frac{1}{\gamma-1}} \le \lambda$, hence

$$\begin{aligned} &\min_{z>\gamma\lambda} f_3(z) > f_3(\gamma\lambda) = f_2(\gamma\lambda) \ge \min_{\lambda<z\le\gamma\lambda} f_2(z) > f_2(\lambda) = f_1(\lambda) \\ \ge &\min_{0<z\le\lambda} f_1(z) : \begin{cases} > f_1(0) = \mathcal{L}_\gamma(0;\lambda) & \text{if } |a| - \frac{\lambda}{\theta} \le 0 \\ = f_1\left(|a| - \frac{\lambda}{\theta}\right) < \mathcal{L}_\gamma(0;\lambda) & \text{if } |a| - \frac{\lambda}{\theta} > 0 \end{cases} \end{aligned}$$

That's to say, $z^* = \left(|a| - \frac{\lambda}{\theta}\right)_+$. Hence $x^* = S\left(a, \frac{\lambda}{\theta}\right)$.

$*$ If $\lambda + \frac{\lambda}{\theta} < |a| \le \gamma\lambda$, we observe that $|a| - \frac{\lambda}{\theta} > \lambda$ and $\lambda < \frac{\theta|a| - \frac{\gamma\lambda}{\gamma-1}}{\theta - \frac{1}{\gamma-1}} \le \gamma\lambda$, hence

$$\min_{z>\gamma\lambda} f_3(z) > f_3(\gamma\lambda) = f_2(\gamma\lambda) \ge \min_{\lambda<z\le\gamma\lambda} f_2(z) = f_2\left(\frac{\theta|a| - \frac{\gamma\lambda}{\gamma-1}}{\theta - \frac{1}{\gamma-1}}\right)$$

$$\mathcal{L}_\gamma(0;\lambda) > \min_{0<z\le\lambda} f_1(z) = f_1(\lambda) = f_2(\lambda) > \min_{\lambda<z\le\gamma\lambda} f_2(z) = f_2\left(\frac{\theta|a| - \frac{\gamma\lambda}{\gamma-1}}{\theta - \frac{1}{\gamma-1}}\right)$$

That's to say, $z^* = \frac{1}{1-\frac{1}{(\gamma-1)\theta}}\left(a - \frac{\gamma\lambda}{(\gamma-1)\theta}\right)$. Hence $x^* = \frac{S\left(a, \frac{\gamma\lambda}{(\gamma-1)\theta}\right)}{1-\frac{1}{(\gamma-1)\theta}}$.

$*$ If $|a| > \gamma\lambda$, we observe that $|a| - \frac{\lambda}{\theta} > \lambda$ and $\frac{\theta|a| - \frac{\gamma\lambda}{\gamma-1}}{\theta - \frac{1}{\gamma-1}} > \gamma\lambda$, hence

$$\begin{aligned} f_3(|a|) = \min_{z>\gamma\lambda} f_3(z) < f_3(\gamma\lambda) = f_2(\gamma\lambda) = \min_{\lambda<z\le\gamma\lambda} f_2(z) \\ < f_2(\lambda) = f_1(\lambda) = \min_{0<z\le\lambda} f_1(z) < \mathcal{L}_\gamma(0;\lambda). \end{aligned}$$

That's to say, $z^* = |a|$. Hence $x^* = a$.

Summarize the discussion above, the closed-form solution is

$$x^* = \begin{cases} S\left(a, \frac{\lambda}{\theta}\right) & \text{if } |a| \le \lambda + \frac{\lambda}{\theta} \\ \frac{S\left(a, \frac{\gamma\lambda}{(\gamma-1)\theta}\right)}{1-\frac{1}{(\gamma-1)\theta}} & \text{if } \lambda + \frac{\lambda}{\theta} < |a| \le \gamma\lambda \\ a & \text{if } |a| > \gamma\lambda \end{cases} \tag{2}$$

3 MCP

Consider Minimax Concave Penalty (MCP) with a single variable x and parameter $\gamma > \frac{1}{\theta}$, that is

$$P_\gamma(x;\lambda) = \begin{cases} \lambda|x| - \frac{x^2}{2\gamma} & \text{if } |x| \le \gamma\lambda \\ \frac{1}{2}\gamma\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases}$$

We want to compute

$$x^* = \arg\min_x \mathcal{L}_\gamma(x;\lambda),$$

where

$$\mathcal{L}_\gamma(x;\lambda) = \begin{cases} \frac{\theta}{2}(x-a)^2 + \lambda|x| - \frac{x^2}{2\gamma} & \text{if } |x| \le \gamma\lambda \\ \frac{\theta}{2}(x-a)^2 + \frac{1}{2}\gamma\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases}$$

We consider the following two cases.

- **CASE** I. If $a = 0$, it is obvious that $x^* = 0$.
- **CASE** II. If $a \neq 0$, we observe that
  - When $x = 0$, we have $\mathcal{L}_\gamma(0;\lambda) = \frac{1}{2}\theta a^2$.
  - When $x \neq 0$. WLOG, assume that $\operatorname{sign}(x) = \operatorname{sign}(a)$. (If $\operatorname{sign}(x) = -\operatorname{sign}(a)$, we can replace x by $-x$ which can make the objective function decrease.) Note that $x = |x|\operatorname{sign}(x) = |x|\operatorname{sign}(a)$, we obtain

$$\begin{aligned} \mathcal{L}_\gamma(x;\lambda) &= \begin{cases} \frac{\theta}{2}(|x|-|a|)^2 + \lambda|x| - \frac{x^2}{2\gamma} & \text{if } |x| \le \gamma\lambda \\ \frac{\theta}{2}(|x|-|a|)^2 + \frac{1}{2}\gamma\lambda^2 & \text{if } |x| > \gamma\lambda \end{cases} \\ &\overset{z=|x|}{=} \begin{cases} \frac{1}{2}\left(\theta - \frac{1}{\gamma}\right)z^2 + (\lambda - \theta|a|)\,z + \frac{1}{2}\theta a^2 \triangleq g_1(z) & \text{if } 0 < z \le \gamma\lambda \\ \frac{\theta}{2}(z-|a|)^2 + \frac{1}{2}\gamma\lambda^2 \triangleq g_2(z) & \text{if } z > \gamma\lambda \end{cases} \end{aligned}$$

    * If $|a| \le \gamma\lambda$, we observe that $\frac{\theta|a|-\lambda}{\theta-\frac{1}{\gamma}} \le \gamma\lambda$, hence

$$\begin{aligned} \min_{z>\gamma\lambda} g_2(z) &> g_2(\gamma\lambda) = g_1(\gamma\lambda) \\ &\ge \min_{0<z\le\gamma\lambda} g_1(z) : \begin{cases} > g_1(0) = \mathcal{L}_\gamma(0;\lambda) & \text{if } \frac{\theta|a|-\lambda}{\theta-\frac{1}{\gamma}} \le 0 \\ = g_1\left(\frac{\theta|a|-\lambda}{\theta-\frac{1}{\gamma}}\right) < \mathcal{L}_\gamma(0;\lambda) & \text{if } \frac{\theta|a|-\lambda}{\theta-\frac{1}{\gamma}} > 0 \end{cases} \end{aligned}$$

      That's to say, $z^* = \frac{(\theta|a|-\lambda)_+}{\theta-\frac{1}{\gamma}} = \frac{\left(|a|-\frac{\lambda}{\theta}\right)_+}{1-\frac{1}{\gamma\theta}}$. Hence $x^* = \frac{1}{1-\frac{1}{\gamma\theta}} S\left(a, \frac{\lambda}{\theta}\right)$.
    * If $|a| > \gamma\lambda$, we observe that $\frac{\theta|a|-\lambda}{\theta-\frac{1}{\gamma}} > \gamma\lambda$, hence

$$g_2(|a|) = \min_{z>\gamma\lambda} g_2(z) < g_2(\gamma\lambda) = g_1(\gamma\lambda) = \min_{0<z\le\gamma\lambda} g_1(z) < \mathcal{L}_\gamma(0;\lambda).$$

      That's to say, $z^* = |a|$. Hence $x^* = a$.

Summarize the discussion above, the closed-form solution is

$$x^* = \begin{cases} \frac{S\left(a,\frac{\lambda}{\theta}\right)}{1-\frac{1}{\gamma\theta}} & \text{if } |a| \le \gamma\lambda \\ a & \text{if } |a| > \gamma\lambda \end{cases} \tag{3}$$

4 Group Methods

Minimize

$$\mathcal{L}(\boldsymbol{x};\lambda) = \frac{\theta}{2}(\boldsymbol{x}-\boldsymbol{a})^2 + P(\boldsymbol{x};\lambda)$$

Let $S(\boldsymbol{x},t) = \text{sign}(\boldsymbol{x})\,(||x||-t)_+ = \boldsymbol{x}\left(1-\frac{t}{||x||}\right)_+$ be the soft thresholding operator. The closed-form solution for gLASSO is

$$\boldsymbol{x}^* = S\left(\boldsymbol{a},\frac{\lambda}{\theta}\right). \tag{4}$$

The closed-form solution for gSCAD penalty is

$$\boldsymbol{x}^* = \begin{cases} S\left(\boldsymbol{a},\frac{\lambda}{\theta}\right) & \text{if } ||\boldsymbol{a}|| \le \lambda + \frac{\lambda}{\theta} \\ \frac{S\left(\boldsymbol{a},\frac{\gamma\lambda}{(\gamma-1)\theta}\right)}{1-\frac{1}{(\gamma-1)\theta}} & \text{if } \lambda + \frac{\lambda}{\theta} < ||\boldsymbol{a}|| \le \gamma\lambda \\ \boldsymbol{a} & \text{if } ||\boldsymbol{a}|| > \gamma\lambda \end{cases} \tag{5}$$

The closed-form solution for gMCP is

$$\boldsymbol{x}^* = \begin{cases} \frac{S\left(\boldsymbol{a},\frac{\lambda}{\theta}\right)}{1-\frac{1}{\gamma\theta}} & \text{if } ||\boldsymbol{a}|| \le \gamma\lambda \\ \boldsymbol{a} & \text{if } ||\boldsymbol{a}|| > \gamma\lambda \end{cases} \tag{6}$$