Penalty method is not effective for nonnegative least square

Iterative Rewieighted Least Square

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft : Feburary 19, 2020
Last update : February 19, 2020

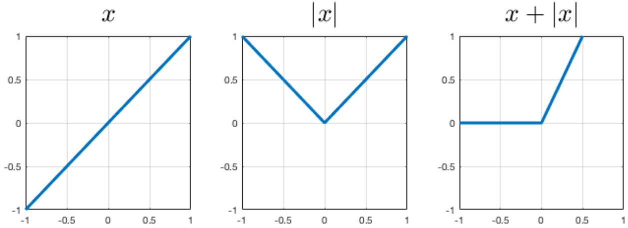
The max function and absolute value

The max function can be seen as the sum of identity and absolute function

$$\max(x, y) = \frac{x + y + |x - y|}{2}.$$

Hence, we have

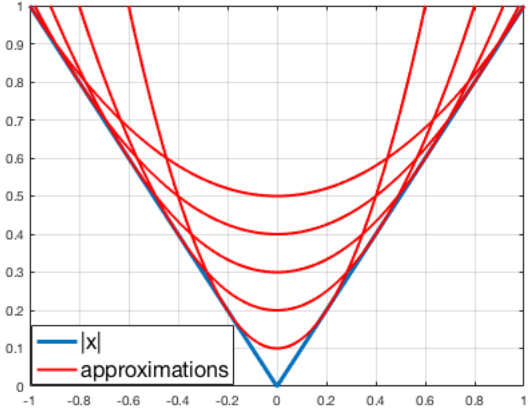
$$\max(x, 0) = \frac{x + |x|}{2}.$$



Quadratic characterization of absolute value

$$|x| = \min_{\alpha} \frac{1}{2}\frac{x^2}{\alpha} + \frac{1}{2}\alpha$$

where the optimal $\alpha = |x|$.



Penalty method

The problem

$$\underset{x \geq 0}{\operatorname{argmin}}\ f(x),$$

has the following penalized form

$$\underset{x}{\operatorname{argmin}}\ f_\lambda(x) = f(x) + \lambda g(x), \quad g(x) = \max\{-x, 0\}.$$

See here for smoothing the max operator using softmax.

Penalty method

By using the quadratic characterization of absolute value, we have

$$
\begin{aligned}
f_\lambda(x) &= f(x) + \lambda \max\{-x, 0\} \\
&= f(x) + \lambda \frac{-x + |-x|}{2} \\
&= f(x) + \frac{\lambda}{2}\big(|x| - x\big) \\
&= f(x) + \frac{\lambda}{2}\left(\min_\alpha \Big(\frac{1}{2}\frac{x^2}{\alpha} + \frac{1}{2}\alpha\Big) - x\right) \\
&= \min_\alpha f(x) + \frac{\lambda}{2}\left(\frac{1}{2}\frac{x^2}{\alpha} + \frac{1}{2}\alpha - x\right) \\
&= \min_\alpha f(x) + \frac{\lambda}{4}\frac{x^2}{\alpha} + \frac{\lambda}{4}\alpha - \frac{\lambda}{2}x
\end{aligned}
$$

Hence, we have

$$
\operatorname*{argmin}_{x,\alpha} \; f(x) + \frac{\lambda}{4}\frac{x^2}{\alpha} + \frac{\lambda}{4}\alpha - \frac{\lambda}{2}x.
$$

Application to nonnegative least squares

Non**n**egative **L**east **S**quares (NNLS) : given $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, find $\mathbf{x} \in \mathbb{R}^n_+$ by solving

$$(\mathcal{P}) \ : \ \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} \ f(\mathbf{x}) := \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2.$$

Using quadratic characterization, we have

$$(\mathcal{P}') \ : \ \underset{\mathbf{x}, \alpha_i}{\operatorname{argmin}} \ f_{\lambda_i}(\mathbf{x}) := \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \sum_{i=1}^{n} \left(\frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i \right).$$

Solving the penalized least squares

$$(\mathcal{P}') \;:\; \underset{\mathbf{x},\alpha_i}{\operatorname{argmin}}\; f_{\lambda_i}(\mathbf{x}) := \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2 + \sum_{i=1}^{n}\left(\frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i\right).$$

Denote the column of $\mathbf{A}$ as $\mathbf{a}_i$, we have

$$f_\lambda(\mathbf{x}) = \frac{1}{2}\|\sum_{i=1}^{n} x_i\mathbf{a}_i - \mathbf{b}\|_2^2 + \sum_{i=1}^{n}\left(\frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i\right).$$

Focusing on the i^{th} term, we have

$$f_\lambda(x_i) = \frac{1}{2}\left\|x_i\mathbf{a}_i - \mathbf{b}_{-i}\right\|_2^2 + \frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i + c.$$

where $\mathbf{b}_{-i} = \sum_{j\neq i} x_j\mathbf{a}_j - \mathbf{b}$. Expand it we get

$$f_\lambda(x_i) = \frac{1}{2}\|\mathbf{a}_i\|_2^2 x_i^2 - \mathbf{a}_i^\top\mathbf{b}_{-i}x_i + \frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i + c.$$

Solving the penalized least squares

We now arrive at a coordinate descent with componentwise subproblem

$$x_i = \operatorname*{argmin}_{x,\alpha} \, f_\lambda(x) = \frac{1}{2}\|\mathbf{a}_i\|_2^2 x^2 - \mathbf{a}_i^\top \mathbf{b}_{-i} x + \frac{\lambda_i}{4}\frac{x^2}{\alpha} + \frac{\lambda_i}{4}\alpha - \frac{\lambda_i}{2}x,$$

in which the objective function is easy to solve on x, it has close form solution. To solve it, we can solve $\frac{\partial}{\partial x} f_\lambda(x) = 0$. The derivative is

$$\frac{\partial}{\partial x} f_\lambda(x) = \|\mathbf{a}_i\|_2^2 x - \mathbf{a}_i^\top \mathbf{b}_{-i} + \frac{\lambda_i}{2}\frac{x}{\alpha} - \frac{\lambda_i}{2}.$$

One can see that solving $\frac{\partial}{\partial x} f_{\lambda,\mu}(x) = 0$ requires to find the root of a linear equation, which is very simple :

$$x = \frac{\mathbf{a}_i^\top \mathbf{b}_{-i} + \frac{\lambda_i}{2}}{\|\mathbf{a}_i\|_2^2 + \frac{\lambda_i}{2}}$$

Solving NNLS by iterative reweighted least squares

$$(\mathcal{P}') \ : \ \underset{\mathbf{x},\alpha_i}{\operatorname{argmin}} \ f_{\lambda_i}(\mathbf{x}) := \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \sum_{i=1}^{n} \left(\frac{\lambda_i}{4}\frac{x_i^2}{\alpha_i} + \frac{\lambda_i}{4}\alpha_i - \frac{\lambda_i}{2}x_i \right).$$

Algorithm 1: Iterative reweighted least square for NNLS

Result: A solution $\mathbf{x}$ that approximately solves $(\mathcal{P})$

Initialization Set $\mathbf{x}_0 \in \mathbb{R}^n_+$, $\lambda_i > 0$, $\alpha_i > 0$

while *stopping condition is not met* **do**

  for $i = 1 \ldots n$ **do**

    Compute $\mathbf{b}_{-i} = \sum_{j \neq i} \mathbf{x}_j \mathbf{a}_i - \mathbf{b}$

    $x_i = \dfrac{\mathbf{a}_i^\top \mathbf{b}_{-i} + \frac{\lambda_i}{2}}{\|\mathbf{a}_i\|_2^2 + \frac{\lambda_i}{2}}$

    Update $\alpha_i = |x_i|$

  end

end

As $\mathbf{b}_{-i}$ has to be recomputed n times per iteration k, thus the complexity of this algorithm is high. Once again, penalty method on NNLS may not be a good choice.

Last page - summary

- Max function and absolute value.
- Quadratic characterization of absolute value.
- Penalty term on nonnegative constrained problem.
- Solving NNLS using iteratively reweighted least squares.

End of document