Proceedings of the International Conference IMCS-60, MSU
October 10–13, 2024, Chisinau, Republic of Moldova

Using the Ellipsoid Method to Find Parameters of Lasso and Ridge Regressions

Petro Stetsyuk, Olha Khomiak

Abstract

We consider the optimization problem for finding the parameters of a linear regression according to the criterion of the least moduli powered to p $(1 \leq p \leq 2)$ with the regularization of parameters according to the criterion of the least moduli powered to q $(1 \leq q \leq 2)$. Its partial cases are lasso regression and ridge regression, as well as least squares method and the least moduli method. An algorithm for solving the problem is developed based on the well-known ellipsoid method.

Keywords: Lasso regression, ridge regression, linear regression, least moduli criterion, convex function, ellipsoid method.

1 Introduction

Lasso, ridge, and other linear regression models [1] belong to machine learning area and used in statistics, system identification, artificial intelligence problems, medicine, finance, economics, etc. These models are formulated as convex programming problems to which methods of minimizing non-smooth convex functions can be applied.

In the paper [2], the problem of minimizing a non-smooth convex function for finding the parameters of a linear regression model with the criterion of the least moduli powered to $1 \leq p \leq 2$ with L_1-regularization was formulated. Here we present a more general formulation of this problem, where the criterion of the least moduli powered to q $(1 \leq q \leq 2)$ will be used to regularize the parameters. The algorithm for solving the problem will use the ellipsoid method [3].

©2024 by Petro Stetsyuk, Olha Khomiak

2 Formulation of the optimization problem

Let matrix $A = \{a_{ij}\}_{i=1,\dots,m}^{j=1,\dots,n}$, vector $y = \{y_i\}_{i=1,\dots,m}$, and scalars $\lambda \geq 0$, $1 \leq p, q \leq 2$ be given. Consider the convex optimization problem

$$f_{pq}^* = f(x_{pq}^*) = \min_{x \in \mathbb{R}^n} \left\{ f(x) = \sum_{i=1}^{m} \left| y_i - \sum_{j=1}^{n} a_{ij} x_j \right|^p + \lambda \sum_{i=1}^{n} |x_i|^q \right\}, \quad (1)$$

where $x = \{x_j\}_{j=1,\dots,n}$ is a vector of unknown parameters, and $|t|$ is the modulus (absolute value) of the number t. The function $f(x)$ is non-smooth if $p = 1$ or $q = 1$ and smooth if $p > 1$ and $q > 1$. Its subgradient at the point $\overline{x}$ is calculated by the formula

$$g_f(\overline{x}) = p \begin{pmatrix} \sum_{i=1}^{m} sign\left(\sum_{j=1}^{n} a_{ij}\overline{x}_j - y_i\right) \left|\sum_{j=1}^{n} a_{ij}\overline{x} - y_i\right|^{p-1} a_{i1} \\ \cdots \\ \sum_{i=1}^{m} sign\left(\sum_{j=1}^{n} a_{ij}\overline{x}_j - y_i\right) \left|\sum_{j=1}^{n} a_{ij}\overline{x} - y_i\right|^{p-1} a_{in} \end{pmatrix} + \\ + \lambda q \begin{pmatrix} \sum_{i=1}^{n} sign\left(\overline{x}_1\right) |\overline{x}_1|^{q-1} \\ \cdots \\ \sum_{i=1}^{n} sign\left(\overline{x}_n\right) |\overline{x}_n|^{q-1} \end{pmatrix}. \quad (2)$$

If $p = 2$, $q = 1$ or $q = 2$, then partial cases of the problem (1) is the well-known problem of minimizing a non-smooth function

$$f_{LASSO}^* = f_1(x_{2,1}^*) = \min_{x \in \mathbb{R}^n} \left\{ f_1(x) = \|y - Ax\|^2 + \lambda \sum_{i=1}^{n} |x_i| \right\}, \quad (3)$$

and the problem of minimizing a strictly convex smooth function

$$f_{RIDGE}^* = f_2(x_{2,2}^*) = \min_{x \in \mathbb{R}^n} \left\{ f_2(x) = \|y - Ax\|^2 + \lambda \sum_{i=1}^{n} x_i^2 \right\}. \quad (4)$$

Problem (3) uses L_1-regularization of unknown parameters and corresponds to lasso regression, and problem (4) uses L_2-regularization of unknown parameters and corresponds to ridge regression. If $\lambda = 0$, then the problem (1) corresponds to the least squares method ($p = 2$) and the least moduli method ($p = 1$).

3 Algorithm for solving the problem (1)

The problem (1) is a problem of unconditional minimization of a convex function $f(x)$. For its solving, Shor's ellipsoid method can be used, which is implemented as an **emshor** program [3]. We will show its application for the problem of the function $f(x)$ minimization, providing that its minimum point x^*_{pq} is localized in n-dimensional ball with radius r_0 that is centered at the point $x_0 \in \mathbb{R}^n$, i.e., $\|x_0 - x^*_{pq}\| \le r_0$.

Algorithm 1 – Algorithm **emshor** for solving the problem (1)

Step 0. Choose $x_0 \in \mathbb{R}^n$ and $r_0 > 0$ so that $\|x_0 - x^*_{pq}\| \le r_0$. Moreover, choose $\varepsilon > 0$, set $B_0 := I_n$ and $k := 0$.

Step 1. Calculate $g_f(x_k)$ at the point x_k using formula (2).

Step 2. If $\|B_k^\top g_f(x_k)\|\, r_k \le \varepsilon$, then set $k^* := k$, $x^*_\varepsilon := x_k$ and STOP.

Step 3. Calculate

$$x_{k+1} := x_k - h_k B_k \xi_k, \quad \text{where} \quad \xi_k := \frac{B_k^\top g_f(x_k)}{\|B_k^\top g_f(x_k)\|}, \quad h_k := \frac{1}{n+1} r_k.$$

Step 4. Update

$$B_{k+1} := B_k + \left(\sqrt{\frac{n-1}{n+1}} - 1\right)(B_k \xi_k)\, \xi_k^\top \quad \text{and} \quad r_{k+1} := \frac{n}{\sqrt{n^2-1}} r_k.$$

Step 5. Set $k := k + 1$ and go to Step 1.

Theorem 1. *Let x_k and x_{k+1} be generated by Algorithm 1. Then, the ratio of volumes of the ellipsoids $\mathcal{E}_k = \{x \mid \|B_k^{-1}(x_k - x)\| \le r_k\}$ and $\mathcal{E}_{k+1} = \{x \mid \|B_{k+1}^{-1}(x_{k+1} - x)\| \le r_{k+1}\}$ does not depend on k and is equal to*

$$q_n := \frac{vol(\mathcal{E}_{k+1})}{vol(\mathcal{E}_k)} = \frac{n}{n+1}\left(\frac{n}{\sqrt{n^2-1}}\right)^{n-1} < \exp\left\{-\frac{1}{2(n+1)}\right\} < 1.$$

Moreover, $x^*_{pq} \in \mathcal{E}_k$ holds for all $k = 0, 1, \ldots, k^*$, $x^*_\varepsilon \in \mathcal{E}_{k^*}$ and, if Algorithm 1 stops, $f(x^*_\varepsilon) \le f^*_{pq} + \varepsilon$ is valid.

4 Conclusion

In this paper, the convex optimization problem for finding the parameters of a linear regression according to the criterion of the least moduli powered to p $(1 \leq p \leq 2)$ with the regularization of parameters according to the criterion of the least moduli powered to q $(1 \leq q \leq 2)$ is formulated. An algorithm for solving the problem was developed based on the well-known ellipsoid method.

Acknowledgments. The paper is supported by Volkswagen Foundation grant № 97775, National Research Foundation of Ukraine grant № 2021.01/0136 and DTT TS KNU NASU grant № 2M-2024.

References

[1] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction.* Springer, New York, 2009.

[2] P. Stetsyuk, V. Stovba, and M. Korablov. *Using of Ellipsoid Method for Finding Linear Regression Parameters with L1-regularization* 2nd International Conference on Information Technologies and Their Applications (ITTA 2024), Baku, Azerbaijan, April 23-25, 2024, to appear.

[3] A. Fischer, O. Khomyak, and P. Stetsyuk. *The ellipsoid method and computational aspects.* Commun. Optim. Theory, no. 21, 2023, pp. 1–14.

Petro Stetsyuk[1], Olha Khomiak[2]

[1]V. M. Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine
E–mail: stetsyukp@gmail.com
ORCID: https://orcid.org/0000-0003-4036-2543

[2]V. M. Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine
E–mail: khomiak.olha@gmail.com
ORCID: https://orcid.org/0000-0002-5384-9070