Estimation using the L_1 norm

December 10, 2024

1 Introduction

In estimation theory, the optimized cost functions are usually given in least squares term. In that case, the square of the so-called L_2-norm is applied.

If there is a given error vector $\epsilon = [\epsilon_1 \quad \epsilon_2 \dots \epsilon_N]^T$,the L_2 norm is defined as

$$\|\epsilon\|_2 = \sqrt{\sum_{i=1}^{N} \epsilon_i^2}.$$

However, there are very significant problems in the application of this norm. For example, if there is only one outlier in the dataset, it can totally distort the estimation. Therefore, L_2 is not the best choose for robust estimations.

Other norms can also be defined. In the literatire, the application of p-normas is very frequent. Its definition is as follows:

$$\|\epsilon\|_p = \sqrt[p]{\sum_{i=1}^{N} |\epsilon_i|^p}$$

For example, the L_1 norm is

$$\|\epsilon\|_1 = \sum_{i=1}^{N} |\epsilon_i|$$

Remark that it is also called the Taxicab norm or Manhattan norm.

Another interesting norm is the infinite norm. It is trivial that

$$\|\epsilon\|_\infty = max\,|\epsilon_i|$$

2 Weighted least squares fitting.

Another problem is the least squares estimation for inhomogenous linear systems:

$$\arg_x \min \|Ax - b\|_2^2$$

where

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_N^T \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_1 \\ \vdots \\ b_N \end{bmatrix}.$$

Thus, the problem can be written

$$\arg_x \min \sum \left(a_i^T x - b_i\right)^2$$

the optimal solution is well-known:

$$x = \left(A^T A\right)^{-1} A^T b$$

The weighted problem is similar to the original one:

$$\arg_x \min \sum \left(w_i a_i^T x - w_i b_i\right)^2$$

where $w_i$is the i-th weight, that is a positive real number.
In matrix-vector form, the error is

$$\|WAx - Wb\|_2^2$$

can be written, where

$$W = diag(w_1, w_2, \ldots, w_N) = \begin{bmatrix} w_1 & 0 & \ldots & 0 \\ 0 & w_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & w_N \end{bmatrix}$$

The problem remains linear, there the solution is as follows:

$$x = \left(A^T W^T W A\right)^{-1} A^T W^T W b$$

3 Iteratively Rewieighted Least Squares (IRLS) method

Now, we connect the p-th norm and the weighted least squares solution. Let us define the error

$$\epsilon_i = a_i^T x - b_i$$

and use the p-norm to define e_i':

$$\epsilon_i' \doteq |\epsilon_i|_p = \left(|\epsilon_i|^p\right)^{1/p}$$

The weighted error for the weighted least squares method:

$$\epsilon"_i = w_i \epsilon_i$$

Let make the square of the weighted problem and the p-th power of ϵ_i' be equal to each other.

$$|\epsilon_i'|^p = w_i^2 \epsilon_i^2$$

Then

$$w_i^2 = \frac{|\epsilon_i|^p}{\epsilon_i^2} = |\epsilon_i|^{p-2}$$

Therefore, the final value for the weight is given as

$$w_i = |\epsilon_i|^{\frac{p-2}{2}}$$

Special case. If the one-norm is used, then

$$w_i = \sqrt{\frac{1}{|\epsilon_i|}}$$

One norm is frequently used as it is a quite robust norm.