The method of Least Squares

March 14, 2013

In numerical analysis, the method of least squares is a classical method that used to obtain a approximation for some mathematical problems.

A classical problem is given A and b , solve x

$$Ax = b$$

But sometimes the true x cannot be found, since the system is inconsistent / contradictory, for example

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \end{cases}$$

Equation 1 and 2 contradict to each other. This kind of inconsistent system does not have exact true solution. Then it is ok to find a *numerical approximation*

In matrix form $Ax = b$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \qquad \Longleftrightarrow \qquad Ax = b$$

Perform the following

$$A^T A x = A^T b \qquad \Longrightarrow \qquad \bar{x} = \left(A^T A\right)^{-1} A^T b \qquad \Longleftrightarrow \bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.75 \\ 0.75 \end{bmatrix}$$

Define the residue $r = A\bar{x} - b$

$$r = \begin{bmatrix} -0.5 \\ 0.5 \\ 0 \end{bmatrix}$$

The L^2 norm of the r , which is also called the *Eulclidean Distance* is thus

$$||r||_2 = \sqrt{(-0.5)^2 + (0.5)^2 + 0^2} = 0.707$$

The L^2 norm indicate the "*size*" of the error, and in this case, the error is "samll"

Why use L^2 norm ? Actually, we can use L^1 , L^3, L^∞. But L^2 norm is most simple. Further more, L^2 is larger then L^∞ norm

Then what is the *best* approximation for the problem $Ax = b$? Obviously, it is the x_{best} that minimize $||r||_2$

Thus , the general problem for a inconsistent system is thus

For a given system A , and meaurement vecotr b , and the unknown vector x having the relation as

$$Ax = b$$

The best approximation of x can be found by

$$\bar{x}_{best} = \min ||r||_2 = \min ||A\bar{x}_{best} - b||_2 = \min ||A\left(\left(A^T A\right)^{-1} A^T b\right) - b||_2$$

$$\bar{x}_{best} = \min ||(AA^{\dagger} - I)b||_2$$

Where$A^{\dagger}$ **is the** $pseudo - inverse$**, or called** $generalized\ inverse$ **of** A **:** $A^{\dagger} = \left(A^T A\right)^{-1} A^T$

- One basic property of generalized inverse of matrix A is that such relation is unique (How to prove it ? By standard method , assume 2 matrix then prove them are equal)

$-END-$