

Least-Squares Method

From Wikiversity

Learning project summary

- **Project code:** Least-Squares Method
- **School:** Computer Science
- **Department:** Scientific_Computing

Content summary

A brief introduction to Least-Squares method, and its statistic meaning.

Goals

This learning project offers learning activities and some application for Least-Squares Method. With this project, one should understand the intention of Least-Squares Method, and what it means. Moreover, one should be able to apply some simple Least-Squares methods to find a good approximation for any functions. For more mathematical explanation, one should visit the following page: "Least squares" to obtain more information.

Learning materials

Texts

[1] [Numerical Mathematics and Computing](#) Chapter 12.1

[2] [Numerical Method for Engineers: With Software and Programming Applications](#) Chapter 17.3

[3] [Statistics for Management and Economics](#) Chapter 17.1

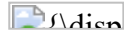
Contents

- 1 Learning project summary
- 2 Content summary
- 3 Goals
- 4 Learning materials
 - 4.1 Texts
 - 4.2 Lessons
 - 4.2.1 Lesson 1: *Introduction to Least-Squares Method*
 - 4.2.2 Lesson 2: *Linear Least-Squares Method*
 - 4.2.3 Lesson 3: *Linear Least-Squares Method in matrix form*
 - 4.2.4 Lesson 4: *Least-Squares Method in statistical view*
 - 4.3 Assignments
 - 4.3.1 Question1: Linear Least-Square Example
 - 4.3.1.1 Calculation
 - 4.3.2 Question2: Nonpolynomial example
- 5 References
- 6 External Links
- 7 Active participants

[4] T.Strutz: Data Fitting and Uncertainty. A practical introduction to weighted least squares and beyond. 2nd edition, Springer Vieweg, 2016, ISBN 978-3-658-11455-8.

Lessons

Lesson 1: *Introduction to Least-Squares Method*

The goal of Least-Squares Method is to find a good estimation of parameters that fit a function, $f(x)$, of a set of data, . The Least-Squares Method requires that the estimated function has to deviate as little as possible from $f(x)$ in the sense of a 2-norm. Generally speaking, Least-Squares Method has two categories, linear and non-linear. We can also classify these methods further: ordinary least squares (OLS), weighted least squares (WLS), and alternating least squares (ALS) and partial least squares (PLS).

To fit a set of data best, the least-squares method minimizes the sum of squared residuals (it is also called the Sum of Squared Errors, SSE.)

$$S = \sum_{i=1}^{i=n} r_i^2,$$

with, r_i , the residual, which is the difference between the actual points and the regression line, and is defined as

$$r_i = y_i - f(x_i)$$

where the m data pairs are (x_i, y_i) and the model function is $f(x_i)$.

At here, we can choose n different parameters for $f(x)$, so that the approximated function can best fit the data set.

For example, in the right graph, $R_3 = Y_3 - f(X_3)$, and $(R_1^2 + R_2^2 + R_3^2 + R_4^2 + R_5^2)$, the sum of the square of each red line's length, is what we want to minimize.



Lesson 2: *Linear Least-Squares Method*

Linear Least-Squares (LLS) Method assumes that the data set falls on a straight line. Therefore, $f(x) = ax + b$, where a and b are constants. However, due to experimental error, some data might not be on the line exactly. There must be error (residual) between the estimated function and real data. Linear Least-Squares

Method (or l_2 approximation) defined the best-fit function as the function that minimizes $S = \sum_{i=1}^{i=n} (y_i - (ax_i + b))^2$

The advantages of LLS:

1. If we assume that the errors have a normal probability distribution, then minimizing S gives us the best approximation of a and b.
2. We can easily use calculus to determine the approximated value of a and b.

To minimize S, the following conditions must be satisfied $\frac{\partial S}{\partial a} = 0$, and $\frac{\partial S}{\partial b} = 0$

Taking the partial derivatives, we obtain $\sum_{i=0}^{i=n} (2 * ((ax_i + b) - y_i)) * x_i = 0$, and $\sum_{i=0}^{i=n} (2 * ((ax_i + b) - y_i)) = 0$.

This system actually consists of two simultaneous linear equations with two unknowns a and b. (These two equations are so-called normal equations.)

Based on the simple calculation on summation, we can easily find out that

$$a = \frac{1}{c} * [(n+1) * \sum_{i=0}^{i=n} (x_i * y_i) - (\sum_{i=0}^{i=n} (x_i))(\sum_{i=0}^{i=n} (y_i))] \text{ and } b = \frac{1}{c} * [(\sum_{i=0}^{i=n} ((x_i)^2)) * (\sum_{i=0}^{i=n} (y_i)) - (\sum_{i=0}^{i=n} (x_i))(\sum_{i=0}^{i=n} (x_i * y_i))]$$

where

$$c = (n+1) * (\sum_{i=0}^{i=n} ((x_i)^2)) - (\sum_{i=0}^{i=n} (x_i))^2$$

Thus, the best estimated function for data set (i, y_i) , for i is an integer between [1, n], is

$$y = ax + b, \text{ where } a = \frac{1}{n^3 - n} * [12 * \sum_{i=1}^{i=n} (i * y_i) - 6 * (n+1)(\sum_{i=1}^{i=n} (y_i))] \text{ and } b = \frac{1}{n^2 - n} * [(4 * n + 2) * \sum_{i=1}^{i=n} (y_i) - 6 * (\sum_{i=1}^{i=n} (i * y_i))].$$

Lesson 3: Linear Least-Squares Method in matrix form

We can also represent estimated linear function in the following model: $y_i = a_1 * x_1 + a_2 * x_2 + \dots + a_m * x_m + r$.

It can be also represented in the matrix form: $\mathbf{Y} = [\mathbf{X}]\mathbf{A} + \mathbf{R}$, where $[\mathbf{X}]$ is a matrix containing coefficients that are derived from the data set (It might not be a square matrix based on the number of variables (m), and data point (n).); Vector \mathbf{Y} contains the value of dependent variable, which is $\mathbf{Y}^T = [y_1 \ y_2 \ \dots \ y_n]$; Vector \mathbf{A} contains the unknown coefficients that we'd like to solve, which is $\mathbf{A}^T = [a_1 \ a_2 \ \dots \ a_m]$; Vector $\{\mathbf{R}\}$ contains the residuals, which is $\mathbf{R}^T = [r_1 \ r_2 \ \dots \ r_n]$.

To minimize \mathbf{R}^T , we follow the same method in lesson 2, obtaining partial derivative for each coefficient, and setting it equal zero. As a result, we have a system of normalized equations, and they can be represented in the following matrix form: $[[\mathbf{X}]^T[\mathbf{X}]]\mathbf{A} = [\mathbf{X}]^T[\mathbf{Y}]$.

To solve the system, we have many options, such as LU method, Cholesky method, inverse matrix, and Gauss-Seidel. (Generally, the equations might not result in diagonal dominated matrices, so Gauss-Seidel method is not recommended.)

182.56.115.73 (discuss) 11:40, 23 August 2014 (UTC)KB

Lesson 4: *Least-Squares Method in statistical view*

From equation $[[\mathbf{X}]^T[\mathbf{X}]]\mathbf{A} = [\mathbf{X}]^T[\mathbf{Y}]$, we can derive the following equation: $\mathbf{A} = [[\mathbf{X}]^T[\mathbf{X}]]^{-1}[\mathbf{X}]^T\mathbf{Y}$.

From this equation, we can determine not only the coefficients, but also the approximated values in statistic.

Using calculus, the following formulas for coefficients can be obtained:

$$a = \frac{S_{xy}}{S_x^2} \text{ and } b = \bar{y} - a * \bar{x}$$

where

$$S_{xy} = \frac{\sum_{i=1}^{i=n} ((x_i - \bar{x}) * (y_i - \bar{y}))}{n - 1}$$

$$S_x^2 = \frac{\sum_{i=1}^{i=n} ((x_i - \bar{x})^2)}{n - 1}$$

$$\bar{x} = \frac{\sum_{i=1}^{i=n} (x_i)}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^{i=n} (y_i)}{n}.$$

Moreover, the diagonal values and non-diagonal values matrix $[[X]^T [X]]^{-1}$ represents variances and covariances of coefficient a_i , respectively.

Assume the diagonal values of $[[X]^T [X]]^{-1}$ is $x_{i,i}$ and the corresponding coefficient is a_i , then

$$var(a_{i-1}) = x_{i,i} * s_{y/x}^2 \text{ and } cov(a_{i-1}, a_{j-1}) = x_{i,j} * s_{y/x}^2$$

where $s_{y/x}$ is called stand error of the estimate, and $s_{y/x} = \sqrt{\frac{S}{n-2}}$.

(Here, lower index, y/x, means that the error of certain x is caused by the inaccurate approximation of corresponding y.)

We have many application on these two information. For example, we can derive the upper and lower bound of intercept and slope.

Assignments

To better understand the application of Least-Squares application, the first question will be solved by applying the LLS equations, and the second one will be solved by Matlab program.

Question1: Linear Least-Square Example

The following are 8 data points that shows the relationship between the number of fishermen and the amount of fish (in thousand pounds) they can catch a day.

Number of Fishermen	Fish Caught
18	39
14	9
9	9

10	7
5	8
22	35
14	36
12	22

According to this data set, what is the function between the number of fishermen and the amount of fish caught? *hint: let the number of fisherman be x , and the amount of fish caught be y , and use LLS to find the coefficients.*

Calculation

By the simple calculation and statistic knowledge, we can easily find out:

1. $\bar{X} = 13$
2. $\bar{Y} = 20.625$, and
3. the following chart

X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X}) * (Y - \bar{Y})$	$(X - \bar{X})^2$
18	39	5	18.375	91.875	25
14	9	1	-11.625	-11.625	1
9	9	-4	-11.625	46.5	16
10	7	-3	-13.625	40.875	9
5	8	-8.0	-12.625	101	64
22	35	9	14.375	129.375	81
14	36	1	15.375	15.375	1
12	22	-1	1.375	-1.375	1

Thus, we have $\Sigma(X - \bar{X}) * (Y - \bar{Y}) = 412$, and $\Sigma(X - \bar{X})^2 = 198$, so the slope, a , = $\frac{412}{198} = 2.08$.

And last the intercept, b , = $20.625 - 2.08 * 13 = -6.4255$.

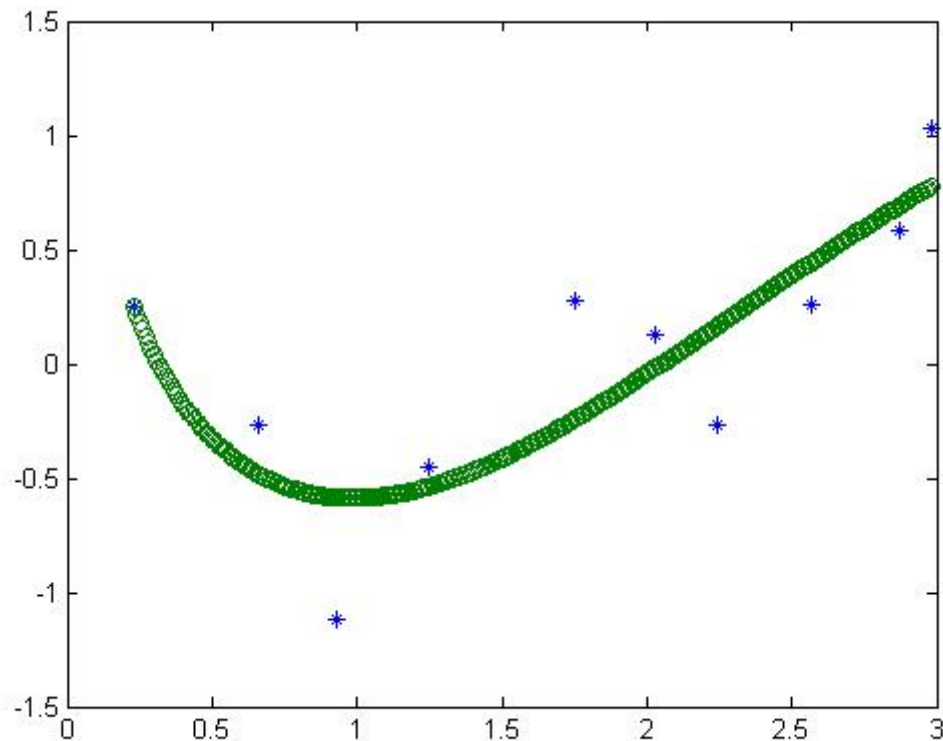
Therefore, the linear least-squares line is $Y = f(X) = -6.4255 + 2.08 * X$.

Question2: Nonpolynomial example

We have the following data (x_i, y_i) , where $1 \leq i \leq n$, by a function of the form $a * e^x + b * \ln x + c * \sin x + d * \cos x$.

x	0.23	0.66	0.93	1.25	1.75	2.03	2.24	2.57	2.87	2.98
y	0.25	-0.27	-1.12	-0.45	0.28	0.13	-0.27	0.26	0.58	1.03

Write a Matlab program that uses Least-Squares method to obtain the estimated function. *hint: input the data in the matrix form, and solve the system to obtain the coefficients.*



The blue spots are the data, the green spots are the estimated nonpolynomial function.

References

- [1] Cheney, Ward and Kincaid, David. Numerical Mathematics and Computing Fifth Edition. Belmont: Thomson Learning, 2004
- [2] Chapra, Steven and Canale, Raymond. Numerical Method for Engineers: With Software and Programming Applications Fourth Edition. McGraw-Hill, 2005
- [3] Keller, Gerald. Statistics for Management and Economics Seventh Edition. Thomson Higher Education, 2005

External Links

- [1] Least Squares Fitting (<http://mathworld.wolfram.com/LeastSquaresFitting.html>) at MathWorld
- [2] Least Squares at wikipedia

Active participants

Active participants in this Learning Group

- Itseng3

Retrieved from "https://en.wikiversity.org/w/index.php?title=Least-Squares_Method&oldid=1561500"

-
- This page was last modified on 25 April 2016, at 11:22.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy.