



The application of positive definite matrix in finding function extremum

Liang Fang^a, Rui Chen^b

College of Mathematics and Statistics, Taishan University, 271000, Tai'an, Shandong, China

^aEmail: fangliang3@163.com

^bEmail: chenruimengting@163.com

Abstract—Function extremum is an important problem in mathematical analysis. In production and daily life, we all hope to reduce the consumption rate of production process and increase the utilization rate of products. We can attribute these practical problems to the extremum of function. It is easy to solve the extremum of simple function, but it is difficult to solve the extremum of multivariate function. Positive definite matrix plays a very important role in solving function extremum. It has a very broad prospect to use positive definite matrix to solve function extremum. This paper discusses the application of positive definite matrix in solving the extremum problem of multivariate function.

Index Terms—Function extremum, gradient, quadratic form, positive definite matrix, Hesse matrix.

I. INTRODUCTION

Function plays a very important role in mathematics. At the same time, function is also a problem often encountered in solving mathematical problems [1]. The extreme value of function is a very important tool to solve mathematical problems.

Positive definite matrix plays a very important role in solving function extremum. It has a very broad prospect to use positive definite matrix to solve function extremum. This paper mainly discusses the application of positive definite matrix in solving the extremum problem of multivariate function, and obtains the method of using the gradient of function and Hesse matrix to find the extremum of multivariate function [2-8].

II. THE RELATIONSHIP BETWEEN POSITIVE DEFINITE MATRIX AND FUNCTION EXTREMUM

In the process of solving some mathematical problems, we can realize that matrix is a very useful tool. Matrix plays a very important role in many fields. For example, modern mathematics, physics, engineering technology, computing technology and national economy. In fact, the matrix, especially the positive definite matrix, plays a more important role in finding the function extremum. Next, we discuss the relationship between positive definite matrix and function extremum.

In our learning process, we know that positive definite matrix plays a very important role in finding the function

extreme value. Through the positive definiteness of Hesse matrix, we can get the extremum of function and the existence point of extremum.

A. Positive definite matrix and function extremum

Lemma 1 [6]. Assume $f(x)$ is a function with n variables x_1, x_2, \dots, x_n , where $X = (x_1, x_2, \dots, x_n)^T$. If $f(x)$ has a first-order continuous partial derivative for each variable and $p_0 = (x_1^0, x_2^0, \dots, x_n^0) \in R^n$ is a stationary point of $f(x)$, then the necessary condition for $f(x)$ to obtain the extreme value at point p_0 is $\text{grad}f(p_0) = 0$.

Lemma 2 [5]. Assume $Z = f(x, y)$ has a continuous second-order partial derivative in a neighborhood of point (x_0, y_0) , and $f'_x(x_0, y_0) = 0$, $f'_y(x_0, y_0) = 0$. Let $f''_{xx}(x_0, y_0) = A$, $f''_{xy}(x_0, y_0) = B$, $f''_{yy}(x_0, y_0) = C$. Then we have the following results.

(1) When $AC - B^2 > 0$, the extreme value is taken at point (x_0, y_0) . Furthermore, the minimum value is obtained when $A > 0$, while the maximum value is obtained when $A < 0$.

(2) When $AC - B^2 = 0$, it is impossible to determine whether (x_0, y_0) is the extreme point of $f(x, y)$ or not.

We can extend the above to multivariate functions as follows:

Lemma 3 [4] (Sufficient condition of extremum). Suppose that function $f(x)$ satisfies

$$\nabla f(x) = \left(\frac{\partial f(x_0)}{\partial x_1}, \frac{\partial f(x_0)}{\partial x_2}, \dots, \frac{\partial f(x_0)}{\partial x_n} \right) = 0$$

in a neighborhood of point $x_0 \in R^n$, then we have the following results.

(1) When $H(x_0)$ is a positive definite matrix, $f(x_0)$ is the minimum of $f(x)$.

(2) When $H(x_0)$ is a negative definite matrix, $f(x_0)$ is



the maximum of $f(x)$.

(3) When $H(X_0)$ is an indefinite matrix, $f(x_0)$ is not the extreme value of $f(x)$.

From the lemma above, the sufficient condition for a binary function $z = f(x, y)$ to be minimal at point (x_0, y_0) is

$$\frac{\partial f(x_0, y_0)}{\partial x} = \frac{\partial f(x_0, y_0)}{\partial y} = 0.$$

Therefore, we have

$$D_1 = \frac{\partial^2 f(x_0, y_0)}{\partial x^2} > 0,$$

$$D_2 = \begin{vmatrix} \frac{\partial^2 f(x_0, y_0)}{\partial x^2} & \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} \\ \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} & \frac{\partial^2 f(x_0, y_0)}{\partial y^2} \end{vmatrix} \\ = \frac{\partial^2 f(x_0, y_0)}{\partial x^2} \cdot \frac{\partial^2 f(x_0, y_0)}{\partial y^2} - \left[\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} \right]^2 > 0.$$

B. Some examples

These are the important definitions and lemmas listed above. Here are some examples.

Example 1. Find the extreme value of function $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3 + 12x_1x_2 + 2x_3$.

Solution. From the condition, we have

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 12x_2, \frac{\partial f}{\partial x_2} = 2x_2 + 12x_1, \frac{\partial f}{\partial x_3} = 2x_3 + 2.$$

Let $\frac{\partial f}{\partial x_1} = 0, \frac{\partial f}{\partial x_2} = 0, \frac{\partial f}{\partial x_3} = 0$, then we get the standing

point $X_0 = (0, 0, 1)^T$, $X_1 = (24, -144, -1)^T$.

And because the second partial derivatives of $f(x)$ are

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 12, \frac{\partial^2 f}{\partial x_1 \partial x_3} = 2,$$

$$\frac{\partial^2 f}{\partial x_2^2} = 2, \frac{\partial^2 f}{\partial x_2 \partial x_3} = 0, \frac{\partial^2 f}{\partial x_3^2} = 2,$$

we get that Hesse matrix is

$$H(x) = \begin{pmatrix} 6x_1 & 12 & 2 \\ 12 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

At point X_0 , we can get

$$H(X_0) = \begin{pmatrix} 0 & 12 & 2 \\ 12 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

The order principal subexpression of $H(X_0)$ is

$$\det H_1 = 0, \det H_2 = \begin{vmatrix} 0 & 12 \\ 12 & 2 \end{vmatrix} = -144 < 0$$

,

$$\det H_3 = \det H(X_0) = -152 < 0.$$

So $H(X_0)$ indefinite, and then we get that X_0 is not the extreme point of the function.

At point X_1 , we can get

$$H(X_1) = \begin{pmatrix} 144 & 12 & 2 \\ 12 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

But the order principal formula of $H(X_1)$ is

$$\det H_1 = 144 > 0,$$

$$\det H_2 = \begin{vmatrix} 144 & 12 \\ 12 & 2 \end{vmatrix} = 144 > 0,$$

$$\det H_3 = \begin{vmatrix} 144 & 12 & 2 \\ 12 & 2 & 0 \\ 2 & 0 & 2 \end{vmatrix} = 280 > 0.$$

So $H(X_1)$ is a positive definite matrix.

$X_1 = (24, -144, -1)^T$ is a minimum, and the minimum is

$$f(X_1) = f(24, -144, -1) = -6913.$$

Example 2. Find the extreme value of

$$u = \sin x + \sin y + \sin z - \sin(x + y + z),$$

where $0 < x < \pi, 0 < y < \pi, 0 < z < \pi$.

Solution. Let

$$u_x = \cos x - \cos(x + y + z) = 0,$$

$$u_y = \cos y - \cos(x + y + z) = 0,$$

$$u_z = \cos z - \cos(x + y + z) = 0,$$

then we can easily get the only standing point

$$P\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right).$$

From



$$u_{xx} = -\sin x + \sin(x+y+z),$$

$$u_{xy} = \sin(x+y+z),$$

$$u_{xz} = \sin(x+y+z),$$

$$u_{yy} = -\sin y + \sin(x+y+z),$$

$$u_{yz} = \sin(x+y+z),$$

$$u_{zz} = -\sin z + \sin(x+y+z),$$

we get $F = \begin{pmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix}$ at P , so P is the maximum

point, and the maximum value is $u(P) = 4$.

The extremum of multivariate function can be determined by the positive definite form of Hesse matrix. This method can be used for ternary and above functions. However, by learning the sufficient conditions for the determination of extreme values of unary and bivariate functions, we can know that this method can also be used to determine the extreme values of unary and bivariate functions. Hesse matrix of one variable function and two variable function are first-order matrix and second-order matrix respectively. By using the positive definiteness of Hesse matrix of one variable function and two variable function, we can get the same conclusion as above.

Example 3. Discusses the extreme value of function

$$Z = \frac{x^2}{2p} - \frac{y^2}{2q} \quad (p > 0, q > 0).$$

Solution. It is easy to know that

$$Z_x = \frac{x}{p}, \quad Z_y = -\frac{y}{q}.$$

Since the first partial derivative of function Z at point $(0,0)$ is zero. Therefore matrix

$$A = \begin{pmatrix} Z_{xx}(0,0) & Z_{xy}(0,0) \\ Z_{yx}(0,0) & Z_{yy}(0,0) \end{pmatrix} = \begin{pmatrix} \frac{1}{p} & 0 \\ 0 & -\frac{1}{q} \end{pmatrix}.$$

Therefore, from the discriminant law, function Z has no extremum.

III. USING POSITIVE DEFINITE MATRIX TO SOLVE CONDITIONAL EXTREMUM PROBLEM

In the following discussion, we will list the methods of using positive definite matrix to solve the problem of conditional extremum according to Hesse matrix of function.

The following is the solution process of this kind of problem:

If we want to find the extremum of function $f(x_1, x_2, \Lambda, x_n)$ with n arguments under condition $g_k(x_1, x_2, \Lambda, x_n) = 0$, $(k=1, 2, \Lambda, m)$, we can follow the following steps:

(1) Constructing Lagrange auxiliary function

$$L(x_1, x_2, \Lambda, x_n, \lambda_1, \lambda_2, \Lambda, \lambda_m) = f(x_1, x_2, \Lambda, x_n) + \sum_{k=1}^m \lambda_k g_k(x_1, x_2, \Lambda, x_n).$$

(2) Suppose $\frac{\partial L}{\partial x_i} = 0$ ($i=1, 2, \Lambda, n$), and

$\frac{\partial L}{\partial \lambda_k} = 0$ ($k=1, 2, \Lambda, m$). These equations are combined,

and the standing point $(x_1^0, x_2^0, \Lambda, x_n^0, \lambda_1^0, \lambda_2^0, \Lambda, \lambda_m^0)$ is obtained by solving the equations, then $P_0 = (x_1^0, x_2^0, \Lambda, x_n^0)$.

(3) Assume $F(x_1, x_2, \Lambda, x_n)$, the Hesse matrix of the constructor, be recorded as H , that is,

$$H = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{pmatrix}.$$

(4) The coordinates of $P_0 = (x_1^0, x_2^0, \Lambda, x_n^0)$ are substituted into H , and the corresponding numerical matrix is recorded as $H(P_0)$.

So we can get the following results.

① If $H(P_0)$ is a positive definite matrix, then under condition $g_k(x_1, x_2, \Lambda, x_n) = 0$, $f(x_1, x_2, \Lambda, x_n)$ gets a minimum at P_0 .

② If $H(P_0)$ is a negative definite matrix, then under condition $g_k(x_1, x_2, \Lambda, x_n) = 0$, $f(x_1, x_2, \Lambda, x_n)$ gets the maximum at P_0 .

Example 4. Find the extreme value of function $f(x, y) = x^2 + y^2 - 3$ under condition $y = 1 + x$.

Solution. The first step is to construct the Lagrange function $L(x, y, \lambda) = x^2 + y^2 - 3 + \lambda(1 + x - y)$. Solve the following system of equations

$$\begin{cases} \frac{\partial L}{\partial x} = 2x + \lambda = 0 \\ \frac{\partial L}{\partial y} = 2y - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = 1 + x - y = 0 \end{cases}$$

to get $x = -\frac{1}{2}$, $y = \frac{1}{2}$, $\lambda = 1$.

Next, determine whether $p_0\left(-\frac{1}{2}, \frac{1}{2}\right)$ is the extreme point or not. According to

$$\begin{aligned} F(x, y) &= x^2 + y^2 - 3 + 1(1 + x - y) \\ &= x^2 + y^2 + x - y - 2, \end{aligned}$$

we can get

$$\begin{aligned} \frac{\partial F}{\partial x} &= 2x + 1, \quad \frac{\partial F}{\partial y} = 2y - 1, \quad \frac{\partial^2 F}{\partial x^2} = 2, \\ \frac{\partial^2 F}{\partial x \partial y} &= \frac{\partial^2 F}{\partial y \partial x} = 0, \quad \frac{\partial^2 F}{\partial y^2} = 2. \end{aligned}$$

Through these second-order partial derivatives, we can get

that Hesse matrix $H(p_0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ is a positive definite

matrix. Therefore, the function obtains the minimum value at point $P_0\left(-\frac{1}{2}, \frac{1}{2}\right)$, and can find that the minimum value is

$$f\left(-\frac{1}{2}, \frac{1}{2}\right) = -\frac{5}{2}.$$

IV. CONCLUSION

Using the positive definiteness of Hesse matrix to judge the extremum of function, this method can be applied to univariate function, bivariate function and multivariate function. In this paper, we discuss how to use positive definite matrix to solve the problem of conditional extremum, and the extension of positive definite matrix decision theorem of multivariate function extremum. This makes us realize that using positive definite matrix to solve function extremum can be extended to other aspects. These conclusions are of great benefit to the future study.

V. ACKNOWLEDGMENT

The work is supported by Tai'an Science and Technology Development Project: Research on case teaching of linear algebra combined with learners practical ability (2019ZC251) and Teaching Reform and Research Project of Taishan University in 2018: The Theory of Probability Combining Real Life Application and Computer Simulation and the Teaching Reform of Mathematical Statistics Course.

REFERENCES

- [1] Department of Beijing University of mathematics. *Higher Algebra* (3rd Edition). Higher education press, 2003.
- [2] Department of mathematics, East China Normal University. *Mathematical analysis* (4th Edition), higher education press, 2010.
- [3] G. Lu. *Linear algebra*. Higher education press, 2002.
- [4] S. Zhao. *Fundamentals of economic applied mathematics*. China Renmin University Press, 1988.
- [5] X.Y. Ma, L. Fang. *Selected lectures on higher algebra*. Tianjin Education Press, 2010.
- [6] W.J. Yang, J. Sun. *Extremum problems of multivariate functions*. Journal of Liaoning Institute of technology, 24 (1): 27-30, 2004.
- [7] Y.W. Zhu, et al. *Application of positive definite matrix in extreme value problem of function*. Practice and understanding of mathematics, 40 (21): 249-250, 2010.
- [8] W.L. Zhang. *Simple application of positive definite matrix*. Journal of Jindongnan normal college, 21 (2): 67-69, 2004.

Authors' biography with photos



Liang Fang is a professor at Taishan University. He obtained his PhD from Shanghai Jiaotong University in June, 2010. His research interests are in the areas of cone optimization, and complementarity problems.



Rui Chen is a lecturer at Taishan University. She obtained her master's degree from Shandong University in December, 2009. Her research interests are in the areas of application of probability theory, and applied statistics in recent years.