THE TEACHING OF MATHEMATICS
2005, Vol. VIII, 2, pp. 53–60

QUADRATIC FUNCTIONS IN SEVERAL VARIABLES

Vladimir Janković

Abstract. In secondary schools students learn to investigate the behavior of the quadratic function in one variable, and to find the point where the function reaches its extremal value. The purpose of this article is to demonstrate how the idea which is applied to functions in one variable can be extended and applied to functions in several variables. We present the procedure to determine whether a quadratic function in several variables has a minimum or a maximum, and if it has, to find points in which the extremal value is reached. This procedure leads to several theoretical results.

ZDM Subject Classification: I65; *AMS Subject Classification*: 00 A 35

Key words and phrases: Quadratic functions in several variables; quadratic forms; extremal values; definitness, Sylvester Criterion.

0. Introduction

In Section 1 we present an elementary procedure to determine extremal values of quadratic functions in two variables. Based on this procedure, we prove Theorem 2, which can be used to establish immediately whether a quadratic function in two variables has an extremal value, and if it has one, to find the points where the extremal value is reached.

In Section 2 we study the quadratic form in two variables. Based on the results of Section 1 we obtain necessary and sufficient conditions for positive semi-definiteness and for positive definiteness. The second condition is known in the mathematical literature as Silvester Criterion.

In Section 3 we show that the procedure applied in the second section to quadratic functions in two variables, can be applied to quadratic functions in several variables, too. Using this procedure, we prove a theoretical result: if a quadratic function in n variables is bounded from bellow, then it reaches its minimum.

The last section is written for advanced students. Here we prove a generalization of Theorem 2 from Section 1. This proof is not an elementary one, contrary to the previous ones in this article. It turns out that, for quadratic functions, the stationarity and the the nonnegativity of the Hessian is not only a necessary, but also a sufficient condition for a function to reach its global minimum at some point.

1. Quadratic function in two variables

The quadratic function in two variables is a polynomial of the second degree in two variables. Accordingly, the quadratic function is a function that can be represented as

$$K(x,y) = ax^2 + by^2 + 2cxy + 2dx + 2ey + f,$$

where a, b, c, d, e and f are coefficients such that at least one of the coefficients a, b and c is different from zero. Here we consider quadratic functions over the field of real numbers. So, coefficients a, b, c, d, e i f are real numbers and variables x and y take values in the set of reals. We shall show how to determine the extremes of such a function.

To every function K we correspond the following system of linear equations

$$(S_K) \qquad \begin{cases} ax + cy + d = 0, \\ cx + by + e = 0. \end{cases}$$

The reader, familiar with differential calculus, will recognize that the expressions on the left-hand sides of the system (S_K) are one halves of partial derivatives of K in x and in y. Therefore, the solutions of the system (S_K) are stationary points of K, and hence one finds amongst them all of the points in which K reaches extremal values. We shall prove this fact by elementary means.

If $a = b = 0$ and $c \neq 0$, then K is not bounded, either from bellow, or from above, and therefore it has neither maximal nor minimal value. If $d \neq 0$, this follows from

$$K(x, 0) = 2dx + f,$$

and if $d = 0$, this follows from

$$K(x, 1) = 2cx + 2e + f.$$

If $a > 0$, the function K is not bounded from above. This follows from the fact that the function

$$K(x, 0) = ax^2 + 2dx + f$$

is not bounded from above. The same holds in the case when $b > 0$. In the similar way it can be proved that K is not bounded from bellow if $a < 0$ or if $b < 0$.

By the way, notice that the following theorem is valid.

THEOREM 1. *If a quadratic function K is bounded from bellow, then the following inequalities hold: $a \geq 0$ and $b \geq 0$, where at least one of them is strict.*

Further, we investigate the case when $a > 0$. In this case, K is not bounded from above. It remains to find out whether K is bounded from bellow and, if it is, to determine whether it has the minimum, and if it has, to find points in which the minimum is reached. The function aK can be represented in the following way:

$$aK(x, y) = (ax + cy + d)^2 + [(ab - c^2)y^2 + 2(ae - cd)y + (af - d^2)].$$

We consider the following cases.

1. $ab - c^2 < 0$. The function

$$K(-(cy + d)/a, y) = [(ab - c^2)y^2 + 2(ae - cd)y + (af - d^2)]/a$$

is not bounded from bellow, hence K does not have the minimum.

2. $ab - c^2 = 0$, $ae - cd \neq 0$. The function

$$K(-(cy + d)/a, y) = [2(ae - cd)y + (af - d^2)]/a$$

is not bounded from bellow, hence in this case K does not have the minimum, too.

3. $ab - c^2 = 0$, $ae - cd = 0$. In this case the equations of the system (S_K) are proportional to each other, and hence the system reduces to one equation. The set of solutions of (S_K) is infinite and, geometrically it is a line in the plane. Since

$$K(x, y) = [(ax + cy + d)^2 + (af - d^2)]/a,$$

it follows that K has the minimum $\hat{K} = (af - d^2)/a$, which is attained at points satisfying the equation $ax + cy + d = 0$, i.e. satisfying the system (S_K).

4. $ab - c^2 > 0$. We have that

$$(ax + cy + d)^2 \geq 0,$$
$$(ab - c^2)y^2 + 2(ae - cd)y + (af - d^2) \geq [(ab - c^2)(af - d^2) - (ae - cd)^2]/(ab - c)^2.$$

It follows that $K(x, y) \geq \hat{K}$, where

$$\hat{K} = [(ab - c^2)(af - d^2) - (ae - cd)^2]/a(ab - c)^2$$

The equality $K(x, y) = \hat{K}$ is satisfied if $y = (cd - ae)/(ab - c^2)$ and $ax + cy + d = 0$. It is not hard to see that these conditions are satisfied by the unique point $(\hat{x}, \hat{y})$, which is the solution of the system (S_K).

We can perform similar analysis of K under the assumption that $a < 0$, $b > 0$ or $b < 0$. Thus we obtain a procedure for finding the extremal values of the quadratic function K, and also a procedure for proving general statements concerning the question whether K has extremal values, to determine them and also the points in which they are reached.

THEOREM 2. *If $a > 0$ or $b > 0$, the quadratic function K has the minimum if and only if one of the following two conditions are satisfied:*

1. $ab - c^2 = 0$, $a : c : d = c : b : e$,
2. $ab - c^2 > 0$.

The minimum is reached in points satisfying the system (S_K). In the first case, the equations of the system are proportional, and hence the system (S_K) has infinitely many solutions. In the second case the system (S_K) has the unique solution.

It is possible to formulate theorems related to the maximum of quadratic function K, similar to Theorems 1 and 2. We leave this to the reader.

2. Quadratic form in two variables

A quadratic form in two variables is a homogeneous quadratic function in two variables, i.e. the function H which can be represented as

$$H(x, y) = ax^2 + by^2 + 2cxy,$$

where coefficients a, b and c are real numbers, not all equal to 0. The system of equations corresponding to the quadratic form H is the following

$$(S_H) \qquad \begin{cases} ax + cy = 0, \\ cx + by = 0. \end{cases}$$

There is one trivial solution of this system, the point $(0,0)$. Since $H(0,0) = 0$, from the consideration of the previous section, we conclude that if the quadratic form H is bounded from bellow, then its minimal value is $\hat{H} = 0$, and hence $H(x,y) \geq 0$ for every $x, y \in R$. In this case we say that H is positive semi-definite. If H is positive semi-definite, and if $H(x,y) = 0$ only when $(x,y) = (0,0)$, we say that H is positive definite. It is not difficult to obtain from Theorem 2 criteria for H to be positive definite and positive semi-definite. This is how they look like.

THEOREM 3. *The quadratic form H is positive semi-definite if and only if $a \geq 0$, $b \geq 0$ and $ab - c^2 \geq 0$. The set of points in which it vanishes coincides with the set of solution of the system (S_H).*

THEOREM 4. *The quadratic form H is positive definite if and only if $a > 0$, and $ab - c^2 > 0$.*

The last theorem is a special case of Silvester Criterion for positive definiteness of a quadratic form on finite dimensional vector spaces.

3. Quadratic functions in n variables

The quadratic function K in n variables is a polynomial of the second degree in n variables. Hence, the quadratic function in n variables is a function that can be represented in the following way

$$K(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + 2 \sum_{k=1}^{n} b_k x_k + c,$$

where at least one of the coefficients a_{ij}, $i,j = 1,2,\ldots,n$, is different from zero. We suppose also that $a_{ij} = a_{ji}$, $i,j = 1,2,\ldots,n$, holds, which guarantees the uniqueness of the representation above. We shall show that, using a procedure similar to the one from the first section, it is possible to determine extremal values for quadratic function with real coefficients.

If $a_{kk} = 0$, $k = 1,2,\ldots,n$, then K is not bounded either from above, or from bellow. This fact is proved in the same way as in the case $n = 2$. There exist i and j, $1 \leq i < j \leq n$, for which $a_{ij} \neq 0$. We have

$$K(0,\ldots,0,x_i,0,\ldots,0,x_j,0,\ldots,0) = 2a_{ij}x_ix_j + 2b_ix_i + 2b_jx_j + c.$$

If $b_i \neq 0$, then the case $x_j = 0$ should be considered:

$$K(0,\ldots,0,x_i,0,\ldots,0,0,0,\ldots,0) = 2b_ix_i + c,$$

and if $b_i = 0$ then the case $x_j = 1$ is to be considered:

$$K(0,\dots,0,x_i,0,\dots,0,1,0,\dots,0) = 2a_{ij}x_i + 2b_j + c.$$

If $a_{kk} > 0$ for some k, $k = 1,2,\dots,n$, then K is not bounded from above. This follows from the fact that

$$K(0,\dots,0,x_k,0,\dots,0) = a_{kk}x_k^2 + 2b_kx_k + c.$$

Using the same argument we conclude that K is not bounded from bellow, if $a_{kk} < 0$ for some k, $k = 1,2,\dots,n$.

By the way, notice that the following theorem is valid.

THEOREM 5. *If the quadratic function K is bounded from bellow, then $a_{kk} \geq 0$, $k = 1,2,\dots,n$, is valid, moreover at least one of these inequalities is strict.*

Let us consider the case $a_{11} > 0$. Then K is not bounded from above. It remains to find out whether it is bounded from bellow and if the answer is positive, to see whether it has the minimum and at which points it is reached. The function $a_{11}K$ can be represented in the following way:

(R) $a_{11}K(x_1,x_2,\dots,x_n) = (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1)^2 + K_1(x_2,\dots,x_n)$,

where K_1 is a quadratic function, a linear function, or a constant. Given $x_2,\dots,x_n$, we can choose x_1 such that

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 = 0.$$

For such x_1 we have

$$a_{11}K(x_1,x_2,\dots,x_n) = K_1(x_2,\dots,x_n).$$

It follows that K is not bounded from bellow if K_1 is not bounded from bellow. If K_1 reaches its minimum at the point $(\hat{x}_2,\dots,\hat{x}_n)$, then K reaches its minimum at the point $(\hat{x}_1,\hat{x}_2,\dots,\hat{x}_n)$, where $\hat{x}_1$ is chosen in such a way that

$$a_{11}\hat{x}_1 + a_{12}\hat{x}_2 + \dots + a_{1n}\hat{x}_n + b_1 = 0.$$

Really,

$$\begin{aligned} a_{11}K(x_1,x_2,\dots,x_n) &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1)^2 + K_1(x_2,\dots,x_n) \\ &\geq K_1(x_2,\dots,x_n) \geq K_1(\hat{x}_2,\dots,\hat{x}_n) \\ &= (a_{11}\hat{x}_1 + a_{12}\hat{x}_2 + \dots + a_{1n}\hat{x}_n + b_1)^2 + K_1(\hat{x}_2,\dots,\hat{x}_n) \\ &= a_{11}K(\hat{x}_1,\hat{x}_2,\dots,\hat{x}_n). \end{aligned}$$

We can proceed in the similar way if $a_{kk} > 0$ for arbitrary k, $k = 2,\dots,n$, or if $a_{kk} < 0$, for arbitrary k, $k = 1,2,\dots,n$.

It seems that there remains only one case that we haven't yet discussed: the case when K_1 is bounded from bellow, but does not reach its minimum. But, this cannot happen. This is the content of the following theorem.

THEOREM 6. *If the quadratic function is bounded from bellow, then it reaches its minimum.*

Proof. We apply the mathematical induction in the number of variables n. That the statement is true for $n = 1$, follows from investigation about the quadratic trinomial, studied in the secondary school. Suppose that $n > 1$ and that the statement of the theorem is true for $n - 1$. Let the quadratic function K in n variables be bounded from bellow. According to Theorem 3, we conclude that $a_{kk} > 0$ for some k, $k = 1, 2, \dots, n$. Let, for example be $k = 1$. We represent the function K in the form (R). Then K_1 is also bounded from bellow. This is why it is either a constant or a quadratic function. In both cases it reaches its minimum. It follows that K reaches its minimum, and the proof is completed. ■

4. For advanced students

A careful reader has probably noticed that the generalization of Theorem 2, related to quadratic functions in two variables, was not proved in the previous section. In order to prove such a theorem, one needs some preliminary knowledge about quadratic forms. They are studied in the textbooks on Algebra (see, for example, [1]) or on Linear Algebra (see, for example, [2]). First, we present some facts concerning quadratic forms, which are necessary for further investigation of quadratic forms.

A quadratic form in n variables ia a homogeneous polynomial of the second degree in n variables. It follows that the quadratic form in n variables is a function that can be represented in the following way

$$H(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where at least one of the coefficients a_{ij}, $i, j = 1, 2, \dots, n$, is different from zero. Of course, here we assume also that $a_{ij} = a_{ji}$, $i, j = 1, 2, \dots, n$, which enables us to obtain the uniqueness of representation of the quadratic form.

The matrix of a quadratic form H is a symmetric quadratic matrix of order n whose entries are coefficients of the corresponding quadratic form: $A = \|a_{ij}\|$. We denote by A the matrix as well as the corresponding linear operator. The quadratic form can be represented using this operator in the following way: $H(x) = \langle Ax, x\rangle$.

The quadratic form H is positive definite if $H(x) > 0$ for every $x \in R$, $x \neq 0$, and it is positive semi-definite if $H(x) \geq 0$ for every $x \in R$. Note that $H(0) = 0$. In the similar way the negative definiteness and the negative semi-definiteness of quadratic forms are defined. A symmetric matrix A is positive definite, positive semi-definite, negative definite or negative semi-definite, and we write $A > 0$, $A \geq 0$, $A < 0$ or $A \leq 0$, if the corresponding quadratic form H is positive definite, positive semi-definite, negative definite or negative semi-definite.

Here we quote two classical theorems, giving necessary and sufficient conditions for positive definiteness and for positive semi-definiteness of quadratic forms.

Theorem A. *The quadratic form H is positive definite if and only if the principal minors of its matrix are positive.*

Remark. The principal minors of a quadratic matrix are determinants of sub-matrices located in the upper left corner of this matrix. The matrix of order n has n principal minors.

Theorem B. *The quadratic form H is positive semi-definite if and only if the diagonal minors of its matrix are nonnegative.*

Theorem A is known as Silvester Criterion, and it can be found in many books on Algebea and Linear Algebra (see [1], [2]). Theorem B is less present in the mathematical textbooks. Amongst the rare books where it can be found, is the book [3]. Theorems given in the second section are special cases of these theorems, when $n = 2$.

The quadratic function K has the corresponding quadratic form H, obtained when monomials of the degree 1 and the constant from the expression by which the quadratic function is defined, are omitted. We denote by A the matrix of the quadratic form H, and by b we denote the vector with the coordinates b_j, $j = 1, 2, \dots, n$. The quadratic function can be represented in the following way:

$$K(x) = \langle Ax, x\rangle + 2\langle b, x\rangle + c.$$

Theorem 7. *The quadratic function K reaches its minimum at the point $\hat{x}$, if and only if $A \geq 0$ and $A\hat{x} + b = 0$.*

Proof. Suppose the quadratic function K reaches its minimum at the point $\hat{x}$. If $h \in R^n$, then we have

$$\begin{aligned} K(\hat{x} + th) &= \langle A\hat{x} + th, \hat{x} + th\rangle + 2\langle b, \hat{x} + th\rangle + c \\ &= \langle A\hat{x}, \hat{x}\rangle + 2\langle b, \hat{x}\rangle + c + \langle Ath, th\rangle + 2\langle A\hat{x}, th\rangle + 2\langle b, th\rangle \\ &= \langle Ah, h\rangle t^2 + 2\langle A\hat{x} + b, h\rangle t + K(\hat{x}). \end{aligned}$$

Since the function

$$\langle Ah, h\rangle t^2 + 2\langle A\hat{x} + b, h\rangle t + K(\hat{x})$$

reaches its minimum at the point $t = 0$ for every $h \in R^n$, then

$$\langle Ah, h\rangle \geq 0, \quad \langle A\hat{x} + b, h\rangle = 0,$$

for every $h \in R^n$. It follows that

$$A \geq 0, \quad A\hat{x} + b = 0.$$

Let $A\hat{x} + b = 0$ and $A \geq 0$. If $h \in R^n$, we have that

$$\begin{aligned} K(\hat{x} + h) &= \langle A\hat{x} + h, \hat{x} + h\rangle + 2\langle b, \hat{x} + h\rangle + c \\ &= \langle A\hat{x}, \hat{x}\rangle + 2\langle b, \hat{x}\rangle + c + \langle Ah, h\rangle + 2\langle A\hat{x}, h\rangle + 2\langle b, h\rangle \\ &= \langle Ah, h\rangle + 2\langle A\hat{x} + b, h\rangle + K(\hat{x}) \geq K(\hat{x}). \end{aligned}$$

It follows that the quadratic function K reaches its minimum at the point $\hat{x}$. ∎

Let D be an open set in R^n and let f be a two times differentiable function, which maps the set D to the real line. The following necessary and sufficient conditions for the function f to reach its local minimum at the point $\hat{x}$, can be found in the textbooks on Mathematical Analysis.

- If the function f reaches its local minimum at the point $\hat{x} \in D$, then

$$(N) \qquad f'(\hat{x}) = 0, \quad f''(\hat{x}) \geq 0.$$

- If the function f at the point $\hat{x} \in D$ satisfies

$$(S) \qquad f'(\hat{x}) = 0, \quad f''(\hat{x}) > 0,$$

  then it reaches its local minimum at the point $\hat{x}$.

For the quadratic function K, the first and the second derivatives are given by: $K'(\hat{x}) = A\hat{x} + b$ and $K''(\hat{x}) = 2A$. It follows that in the case of the quadratic function, the necessary conditions (N) are also sufficient in order to reach the global minimum at the point $\hat{x}$.

REFERENCES

1. А. Г. Курош, *Курс высшей алгебры*, Наука, Москва, 1971.
2. И. М. Гельфанд, *Лекции по линейой алгебре*, Наука, Москва, 1971.
3. И. В. Проскуряков, *Сборник задач по линейной алгебре*, Наука, Москва, 1970.

Faculty of Mathematics, University of Belgrade, Studentski trg 16/IV, 11000 Beograd, Serbia
E-mail: vjankovic@matf.bg.ac.yu