

Method of finding initial solution for the geometric programming problem

Abstract: The paper proposes a new method for finding an initial approximation (initial solution) in the problem of geometric programming. The paper describes the conditions, under which the geometric programming problem of obtaining a positive solution of the matrix equation is solved. This equation describes the orthogonality and normalization conditions.

Key words: geometric programming, initial solution

A number of practical problems are well approximated to geometric programming [1]. We can solve them by quantifying into a mathematical optimization model. For example, geometric programming technique is extremely useful in engineering analysis and design of electrical circuits (e.g., VLSI circuit component sizing) [2] and in solving water resources optimization problems [3]. It has been applied to a wide variety of problems in economics as well [4].

Let us consider the problem of geometric programming [5] for finding the minimum of the objective function (criterion)

$$y = f(\mathbf{x}) = u_1(\mathbf{x}) + u_2(\mathbf{x}) + \dots + u_n(\mathbf{x}). \quad (1)$$

Let $u_i(\mathbf{x})$ be a *posynomial*

$$u_i(\mathbf{x}) = C_i \cdot \prod_{j=1}^m x_j^{a_{ij}}, \quad C_i > 0, \quad i = 1, 2, \dots, n, \quad a_{ij} \in \mathbb{R},$$

and the vector $\mathbf{x} = (x_1, x_2, \dots, x_m)$ of some components (parameters) x_j is positive. An example of this would be a common risk of the technical object safety violation; the criterion would be expressed as a sum of particular risks $u_i(\mathbf{x})$ on a certain threats of safety violation. In this case, $\mathbf{x}$ would be the vector of parameters of a business object protection system. In the theory of reliability, the function (1) is met as a common rate of failure, which expressed as a sum of separate rates of failure $u_i(\mathbf{x})$. The coefficients a_{ij} and C_i are got by methods of linear regression analysis (see [6]).

The matrix $A = (a_{ij})$ is called an exponent matrix. Suppose that the matrix A has the rank $r(A) = m$. This can be written as a block matrix

$$A = \begin{pmatrix} B \\ H \end{pmatrix}. \quad (2)$$

Here, basis B is a $m \times m$ matrix whose determinant $\det B \neq 0$, and a submatrix H contains d rows of the matrix A , which doesn't belong to the basis B . The difficulty level is characterized by this number $d = n - m$. In the paper, we consider the case $d > 1$. The case $d = 1$ was considered in [7].

Let us consider a $d \times m$ matrix $Q = -H \cdot B^{-1}$ and a $d \times n$ matrix $S = (Q, I_d)$, where I_d is the identity $d \times d$ matrix.

According to [8], the minimum $y_* = f(x_*)$ of the objective function (1) and components x_{*j} of optimal vector $x_* = (x_{*1}, x_{*2}, \dots, x_{*m})$ are given in analytical form:

$$y_* = \prod_{i=1}^n \left(\frac{C_i}{\delta_i^*} \right)^{\delta_i^*}, x_{*j} = \prod_{v=1}^m \left(\frac{\delta_v^* y_*}{C_v} \right)^{k_{jv}}, j = 1, 2, \dots, m. \quad (3)$$

Here, numbers $\delta_i = \delta_i^*$ are elements of vector $\delta = (\delta_1, \delta_2, \dots, \delta_m, \delta_{m+1}, \dots, \delta_n) = (\delta_{(m)}, \delta_{(d)})$, where $\delta_{(m)} = (\delta_1, \delta_2, \dots, \delta_m)$ and $\delta_{(d)} = (\delta_{m+1}, \delta_{m+2}, \dots, \delta_{m+d})$, which are found from the equation system:

$$\delta_{(m)} = (\delta_1, \delta_2, \dots, \delta_m) = \delta_{(d)} \cdot Q; \delta_{(d)} \cdot \mu = 1; \delta_{m+i} = \frac{C_{m+i}}{\lambda} \cdot \prod_{j=1}^m \left(\frac{C_j}{\lambda \cdot \delta_j} \right)^{S_{ij}}; \quad (4)$$

$$i = 1, 2, \dots, d.$$

The column vector $\mu = \begin{pmatrix} \mu_1 \\ \dots \\ \mu_d \end{pmatrix}$ consists of components μ_i , each of which is the sum of elements of row with number i of the matrix $S = (s_{ij})$.

In expression (4), the first and the second ratios are called the orthogonality and normalization conditions respectively, and the last one – third – expresses the condition of optimality.

In [8], it was proven that the condition (4) expresses the uniqueness of the solution $\delta > 0$ if it exists. To get δ in [8], the iteration method of solving was applied, which uses simple matrix operations at each step. The convergence of this method has been proven.

In the formulas (4), λ denotes the maximum value

$$V^* = V(\delta^*) = \prod_{i=1}^n \left(\frac{C_i}{\delta_i^*} \right)^{\delta_i^*} \quad (5)$$

of the dual function $V(\delta) = \prod_{i=1}^n \left(\frac{C_i}{\delta_i} \right)^{\delta_i}$ over the area $\delta_+ = \{\delta_{(d)} > 0: \delta_{(m)} = \delta_{(d)} \cdot Q, \delta_{(d)} \cdot \mu = 1\}$.

The following two systems of equations are used for finding the initial solution:

$$\delta_{(m)} = \delta_{(d)} \cdot Q \text{ and } \delta_{(d)} \cdot \mu = 1. \quad (6)$$

That is, we want to find a positive solution to the system of linear algebraic equations. It is known that this problem still has no general solution, which would also be acceptable computationally. In this paper, we solve the problem under the assumptions that reflect its specific features.

Consider the equation

$$\delta \cdot A = 0 \Rightarrow \delta \cdot \begin{pmatrix} B \\ H \end{pmatrix} = 0. \quad (7)$$

According to the theory of generalized inverse matrices [9], the general solution of equation (7) is given by the formula

$$\delta = \eta' \cdot P(A), \text{ where } P(A) = (I_n - A \cdot A^C) \quad (8)$$

by selecting all possible values of an arbitrary vector η' of dimension n . Here, A^C represents the so-called S - inverse matrix for A (see [9]). And in the our case (see (2)), we use the relation $A^C =$

(B^{-1}, O) to determine A^C . Because $A^C \cdot A = (B^{-1}, O) \cdot \begin{pmatrix} B \\ H \end{pmatrix} = I_m$, A^C is the left inverse of A . From this and (8), it follows that

$$\begin{aligned} \delta &= \eta' \cdot \left(I_n - \begin{pmatrix} B \\ H \end{pmatrix} (B^{-1}, O_{m \times d}) \right) = \eta' \cdot \left(I_n - \begin{pmatrix} I_m & O_{m \times d} \\ H \cdot B^{-1} & O_d \end{pmatrix} \right) = \\ &= \eta' \cdot \begin{pmatrix} O_m & O_{m \times d} \\ -H \cdot B^{-1} & I_d \end{pmatrix} \text{ or } \delta = \eta \cdot (Q \quad I_d)_{d \times n}, \end{aligned} \quad (9)$$

where $\eta = (\eta_1, \eta_2, \dots, \eta_d)$ is an arbitrary vector.

From (9), it follows that $\delta = (\eta \cdot Q, \eta)$ and, therefore,

$$\eta = (\delta_{m+1}, \delta_{m+2}, \dots, \delta_{m+d}) \equiv (\delta_{(d)}). \quad (10)$$

This is an arbitrary vector η with a set of components $\delta_{m+1}, \delta_{m+2}, \dots, \delta_{m+d}$ with numbers $m+1, m+2, \dots, m+d = n$ of the vector δ .

Let us say that numbers $\delta_1, \delta_2, \dots, \delta_m$ and $\delta_{m+1}, \delta_{m+2}, \dots, \delta_{m+d}$ are the basic and free elements of the vector δ respectively.

Theorem 1. Let the vector δ be selected from the orthogonality condition (7). Then the following statements will be true:

$$\text{a) Vector } \delta \text{ is written as } \delta = \delta_{(d)} \cdot S \quad (11)$$

$$\text{or } \delta_{(m)} = \delta_{(d)} \cdot Q, \quad (12)$$

where $\delta_{(m)}$ and $\delta_{(d)}$ are composed of basic and free (arbitrary) elements of the vector δ . In the expression (11) and (12), an arbitrary row $\delta_{(d)}$ can be selected only as a positive one: $(\delta = \delta_{(d)} S > 0) \Rightarrow \delta_{(d)} > 0$.

b) Normalization conditions type of $\delta_1 + \delta_2 + \dots + \delta_n = 1$ is written by free elements of the vector δ as

$$\mu_1 \cdot \delta_{m+1} + \mu_2 \cdot \delta_{m+2} + \dots + \mu_d \cdot \delta_{m+d} = \delta_{(d)} \cdot \mu = 1. \quad (13)$$

Moreover, the number $\delta_{m+i} > 0$ and the coefficient μ_i is equals to the sum of row elements s_i of the matrix S , i.e., $\mu_i = s_{i1} + s_{i2} + \dots + s_{im} + 1$. (14)

c) The set δ_+ of all admissible vectors δ is given by (11) by specifying all possible values of its free elements $\delta_{m+i} > 0$ satisfying the condition (13).

The dual problem (5) has a feasible solution δ if and only if there exists a row $\delta_{(d)} > 0$, for which $\delta_{(m)} > 0$.

d) If matrix Q comprises of at least one column $q^i < 0, i = 1, 2, \dots, m$ then the dual problem (5) hasn't solution.

Proof. a) Relation (11) follows directly from (9) and formula (13) follows from the fact that $(\delta = \delta_{(d)} \cdot S) \Leftrightarrow (\delta = \delta_{(d)} \cdot (Q, I) = (\delta_{(d)} \cdot Q, \delta_{(d)})) \Rightarrow (\delta_{(m)} = \delta_{(d)} \cdot Q)$.

Here, an arbitrary vector $\delta_{(d)}$ can be selected only as a positive one, because the task is to find out the conditions of orthogonality of vectors $\delta = \delta_{(d)} \cdot S > 0$. Hence, we conclude that $\delta_{(d)} > 0$.

b) The normalization conditions $\sum_{i=1}^n \delta_i = 1$ can be written as

$\delta \cdot \mathbf{1} = 1 = \delta_{(d)} S \cdot \mathbf{1} = \delta_{(d)} \cdot \mu$, where $\mathbf{1}$ is a column of ones from which we obtain the relation (14).

c) The set of all vectors $\delta > 0$ is found by the formula (11) by giving all possible values of its free elements $\delta_{m+i} > 0$, satisfying the condition (13), because the expression (12) is the general solution of the homogeneous equation $\delta \cdot A = 0$.

From this it follows, that the task (7) has a positive solution if and only if there exists a row $\delta_{(d)} > 0$, for which $\delta_{(m)} = \delta_{(d)} \cdot Q > 0$.

d) Let the matrix Q contains a column $q^j < 0$. Then according to (11), we obtain the expression $\delta_j = \delta_{(d)} \cdot q^j$, in which, as shown above, $\delta_{(d)} > 0$, hence, in view of the $q^j < 0$ it follows, that $\delta_j < 0$. Hence, we conclude that if the matrix Q comprises at least one column $q^j < 0$, then the problem (7) has no solutions.

Theorem 2. Let in the equation $\delta_{(d)} \cdot \mu = 1$, vector $\mu > 0$ and each column q^j of matrix Q have at least one positive element. Then for the dual problem there exists a positive solution δ , given by the conditions:

1. We define the free variables row as

$$\delta_{(d)} = \frac{1}{\sigma} \cdot \mu^T, \quad (15)$$

where number $\sigma = \mu^T \cdot \mu = \mu_1^2 + \mu_2^2 + \dots + \mu_d^2 > 0$, (T denotes transposition). In this case, the following conditions are true: $\delta_{(d)} > 0$, $\delta_{(d)} \cdot \mu = 1$.

2. Let the row $\frac{1}{\sigma} \mu^T \cdot Q \equiv (\omega_1, \omega_2, \dots, \omega_m) > 0$, i.e., all $\omega_j > 0, j = 1, 2, \dots, m$. Then the following inequality is true:

$$\delta_{(m)} = \frac{1}{\sigma} \cdot \mu^T \cdot Q > 0, \quad (16)$$

which means that the theorem is carried out.

3. If in the row $(\omega_1, \omega_2, \dots, \omega_m)$ there is the element $\omega_j < 0$, and q^j is corresponding to its column of the matrix $Q \equiv (q^1 q^2 \dots q^m)$, then we use the expression for the element δ_j of the row $\delta_{(m)}$ in the form of

$$\delta_j = \omega_j + \eta \cdot q^j - \eta \cdot \mu \cdot \omega_j = -\omega_j \cdot (\eta \cdot \mu - 1) + \eta \cdot q^j, j = 1, 2, \dots, m, \quad (17)$$

where $\eta = (\eta_1, \eta_2, \dots, \eta_d)$ is an arbitrary vector.

If the column q^j has its components $q_i^j < 0$ then in (17) we believe $\eta_i = 0$. The rest of the elements of the row $\eta > 0$ we believe are positive: $\eta_s > 0$. In this case, the number $\eta \cdot q^j > 0$ for any $\eta_s > 0$.

Numbers $\eta_s > 0$ are chosen by the condition $\eta \cdot \mu - 1 > 0$ and we arrive at the inequalities

$$\eta \cdot q^j > 0, -\omega_j \cdot (\eta \cdot \mu - 1) > 0,$$

whence follows the relation (16).

Proof. 1. According to the theory of generalized inverse matrices [9], the set of solutions of equation $\delta_{(d)} \cdot \mu = 1$ can be found by the formula

$$\delta_{(d)} = \frac{1}{\sigma} \cdot \mu^T + \eta \cdot (I_d - \frac{1}{\sigma} \cdot \mu \cdot \mu^T) \quad (18)$$

by specifying all possible values of the random vector η .

Provided that $\eta = 0$, we get one of the solutions $\delta_{(d)} = \frac{1}{\sigma} \cdot \mu^T$, satisfying the relations $\delta_{(d)} > 0$, $\delta_{(d)} \cdot \mu = 1$.

Further, according to (18), the arbitrary vector $\eta = (\eta_1, \eta_2, \dots, \eta_d) > 0$ is chosen.

2. For the row $\delta_{(m)}$ of basic variables δ_j , ($j = 1, 2, \dots, m$), we have

$$\delta_{(m)} = \delta_{(d)} \cdot Q = \mu^T \cdot Q \cdot \frac{1}{\sigma} + \eta \cdot Q - \eta \cdot \mu \cdot \mu^T \cdot Q \cdot \frac{1}{\sigma}, \quad (19)$$

$$\text{and } \delta_j = \omega_j + \eta \cdot q^j - \eta \cdot \mu \cdot \omega_j = -\omega_j \cdot (\eta \cdot \mu - 1) + \eta \cdot q^j.$$

Let the row $(\omega_1, \omega_2, \dots, \omega_m) > 0$. Then we have the inequality (16), and in this case the theorem is true.

3. If the row $(\omega_1, \omega_2, \dots, \omega_m)$ has elements $\omega_j < 0$ and q^j is corresponding to its column of the matrix Q , we obtain an expression for element δ_j of the row $\delta_{(m)}$ in the form of relation (17).

If the column q^j has component $q_i^j < 0$, then in (17) we believe $\eta_i = 0$. The remaining components of the row η remain positive: $\eta_s > 0$. Since according to the hypothesis, the column q^j contains at least one element $q_i^j > 0$, then in (17) the number $\eta \cdot q^j > 0$ for any $\eta_s > 0$. Choosing numbers η_s from the condition $\eta \cdot \mu - 1 > 0$ we arrive at the inequalities: $\eta \cdot q^j > 0$, $-\omega_j (\eta \cdot \mu - 1) > 0$, so, $\delta_j > 0$, whence follows the relation (16).

References

1. Boyd, S., Kim, S.-J., Vandenberghe, L., Hassibi, A. A Tutorial on Geometric Programming. Optimization and Engineering, 8(1): 67-127, 2007.
2. Chris Chu and D. F. Wong. VLSI Circuit Performance Optimization by Geometric Programming. Annals of Operations Research, vol. 105, pages 37-60, July 2001.
3. Meier, W. L.; Shih, C. S.; Wray, D. J. (1971). Water Resource System Optimization by Geometric Programming. Texas Water Resources Institute. Available electronically from <http://hdl.handle.net/1969.1/94877>.
4. J. K. Sengupta and J. H. Portillo-Campbell, The approach of geometric programming with economic applications. Z. Staatswirtschaft, 128 (1972), 437—455.
5. Daffin, R., Petersen, E., Zener, C.: Geometric programming, John Wiley, New York, 1969, 311 p.
6. Wilks S., *Mathematical Statistics*, John Wiley, New York, 1962, 402 p.
7. Yatsko, A. Akhlamov, A.: A new method for setting the acceptable restriction to technical object parameters, Scientific Issues, Mathematics, Częstochowa, 2008, p. 101-106.
8. Sudakov R., Yatsko, A.: Elements of applied theory of geometric programming, Znanie, Moscow, 2004, 220 p. ISBN 507-0029851 (in Russian).
9. Sudakov, R.: Tests of technical systems, Mashinostroenie, Moscow, 1988, 356 p. ISBN 5-217-00033-3 (in Russian).