



Soviet-era science, translated into English

MATHEMATICS

Corresponding Member of the Academy of Sciences of the USSR A.
N. TIKHONOV

1965

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196501.45331>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

MATHEMATICS

Corresponding Member of the Academy of Sciences of the USSR A. N. TIKHONOV

ON ILL-POSED PROBLEMS OF OPTIMAL PLANNING AND STABLE METHODS FOR THEIR SOLUTION

1. Consider the problem of linear programming (see, for example, (1)). Find an element $z = \{z_j\}$ ($j = 1, \dots, n$) of the n -dimensional space R_n satisfying m conditions

$$Az = \bar{u}; \quad A = \{a_{ij}\}, \quad \bar{u} = \{\bar{u}_i\} \quad (i = 1, \dots, m; j = 1, \dots, n), \quad (1')$$

the additional constraints

$$z_j \geq 0 \quad (1'')$$

and minimizing the linear form

$$C[z] = \sum_j c_j z_j, \quad c_j \geq 0. \quad (1''')$$

We shall consider this problem without the assumption that the rows of the matrix A are linearly independent, as is often done. Verification of this assumption is practically impossible, since the elements of the matrix A are usually given with errors.

In Sec. 2 an example is constructed of a problem (1) which is ill-posed, i.e. a problem in which arbitrarily small changes in the input data correspond to arbitrarily large changes in the minimum C_0 of the functional C . The problem under consideration may have a nonunique solution. In Sec. 3 an additional condition is given (the condition of minimum cost of organizational restructuring), which makes it possible to single out a unique solution. In Sec. 4 the existence of a solution of the problem is proved. Sec. 5 is devoted to the construction and study of an algorithm for obtaining the normal solution of problem (1), stable with respect to small perturbations of the input data A , \bar{u} , and C . This algorithm is a development of the method set forth in (2, 3).

The entire study is carried out not only for linear functions $C[z]$, but also for nonnegative functions $C(z)$, continuous in the domain $z_j \geq 0$ (cf. (4)).

2. Let us give an example of a linear programming problem which is an ill-posed problem in the sense of Hadamard. Suppose that in the space $R_4 = (z_1, z_2, z_3, z_4)$ the conditions $Az = \bar{u}$ and the function C are given by

$$z_1 - z_2 = \bar{u}_1, \quad z_3 - z_2 = \bar{u}_2, \quad \xi z_1 + \eta z_2 + \zeta z_3 = \bar{u}_3,$$

$$C = z_1 + z_4, \quad \bar{u}_1 \geq 0, \quad \bar{u}_2 \geq 0.$$

Let the determinant of this system be $\Delta = -(\xi + \eta + \zeta) = 0$, and let the consistency condition $u_3 = \zeta u_2 + \xi u_1$ be satisfied. In this case the admissible set A is a one-dimensional parametric family

$$z_1 = \bar{u}_1 + s, \quad z_2 = s, \quad z_3 = \bar{u}_2 + s, \quad z_4 = 0, \quad s = z_2 \geq 0,$$

and the minimum value is $C_0 = \bar{u}_1$.

Suppose, however, that instead of $\xi, \eta, \zeta, \bar{u}_1, \bar{u}_2, \bar{u}_3$, their approximate values $\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}, \tilde{u}_1, \tilde{u}_2, \tilde{u}_3$ are given with some error not exceeding the accuracy δ . This is inevitable, for example, if ξ, η, ζ are irrational and the computations are carried out—

on a finite-precision machine. In this case, in general, $\tilde{\Delta} = -(\tilde{\xi} + \tilde{\eta} + \tilde{\zeta}) \neq 0$ and

$$\tilde{z}_1 = (\tilde{u}_3 - \tilde{\zeta}\tilde{u}_2 - \tilde{\xi}\tilde{u}_1) / (\tilde{\xi} + \tilde{\eta} + \tilde{\zeta}).$$

The value \tilde{z}_1 , as the ratio of two small numbers, may turn out to be equal to any prescribed positive quantity for any δ -accuracy of specifying the input data, so that $\tilde{C}_0 = \tilde{z}_1$. This shows that the problem under consideration is ill-posed in Hadamard's sense.

In practice, in linear programming problems with a large number of constraints, one can easily encounter problems in which, when the coefficients of the matrix A are varied within the accuracy of their specification, the system of constraints becomes degenerate. Such problems are close to the one considered, and no increase in the accuracy of computations will help in solving them.

3. Consider problem (1), without assuming that the rows of the matrix A are linearly independent. Suppose that the vector \bar{u} satisfies the solvability conditions of the system (1'). Denote $N_A = \{z : Az = 0\}$; it is a linear subspace of R_n . Let \bar{C} be the plane defined by the equation

$$\sum c_j z_j = 0,$$

and let Q be a linear subspace of R_n : $Q = N_A \cap \overline{C}$.

It is obvious that if $Q \neq 0$ and $\bar{z}^{(0)}$ is a solution of problem (1), then every element of the set

$$Q_0 = z^0 + Q = \{\bar{z} : \bar{z} = \bar{z}^{(0)} + q, q \in Q\},$$

satisfying the conditions (1''), is, along with $\bar{z}^{(0)}$, also its solution. Thus, problems of type (1) may have nonunique solutions. (For a nonlinear problem, \overline{C} is defined by means of $C = \{z : C(z) = C_0\}$.)

We shall call a **normal solution of problem (1)** a solution having minimal norm. In other words, if $\bar{z}^{(0)}$ is the normal solution of the problem, then

$$\|\bar{z}^{(0)}\| \leq \|\bar{z}\|,$$

where \bar{z} is any other solution of the problem. If the solution of the problem is unique, then the normal solution coincides with it.

The definition of the normal solution depends essentially on the chosen norm and origin of coordinates.

Put

$$\|\bar{z}\| = \Omega[\bar{z} - z_{(0)}]^{1/2},$$

where $z_{(0)}$ is some fixed element of R_n , and

$$\Omega[z] = \sum p_{ij} z_i z_j$$

is a positive definite form.

By a generalized normal solution (with respect to the given element $z_{(0)}$ and the quadratic form $\Omega[z]$) we shall mean such a solution \bar{z}_0 of problem (1) for which

$$\|\bar{z}_0\| \leq \|\bar{z}\|,$$

where \bar{z} is any other solution of this problem.

Since the set Q_0 of possible solutions of the problem is a linear manifold, it is obvious that the normal (generalized normal) solution is uniquely determined.

The economic meaning of the (generalized) normal solution is obvious. Let $z_{(0)}$ be a plan realizable in planning or production, and let problem (1) be posed in connection with a change in the assignment for planning (a change in A, \bar{u}, C). Let the cost of organizational restructuring under the trans-

transition from the plan $z_{(0)}$ to the plan \bar{z} is determined by a positive definite functional $\Omega[\bar{z} - z_{(0)}]$. If problem (1) has a set of solutions $Q_0 = \bar{z}^{(0)} + Q = \{\bar{z} = \bar{z}^{(0)} + q, q \in Q\}$, then it is natural to prefer that solution of the problem which, while optimizing C , will be associated with a minimum of organizational readjustments. The study of functionals Ω is in general necessary for the economic analysis of the problem.

Thus the question naturally arises of methods for determining the normal solution of the problem. It is this problem that we shall call problem (1).

4. There is (cf. (5))

Theorem 1. *If the vector \bar{u} satisfies the solvability conditions of the system $Az = \bar{u}$ and conditions (1') and (1'') are compatible, then problem (1) has at least one solution.*

Choose the numbering of the coordinates so that for $0 \leq i \leq n_1$ the values $c_i > 0$, and for $n_1 + 1 \leq i \leq n$ the values $c_i = 0$. Denote $R_{n_1} = \langle z_i, i \leq n_1 \rangle$.

Let $\{z^{(n)}\}$ be a minimizing sequence of admissible elements such that $C(z^{(n)}) = C_n \rightarrow C_0$, where C_0 is the minimal value of C under conditions (1') and (1''). Obviously, $0 \leq c_i z_i^{(n)} \leq C_n \leq C_1$, or $0 \leq z_i^{(n)} \leq \frac{1}{c_i} C$ for $i \leq n_1$. Hence it follows that $\hat{z}^{(n)}$ —the projections of $z^{(n)}$ onto R_{n_1} —form a compact set. We shall assume that the sequence $\hat{z}^{(n)}$ is such that $\hat{z}^{(n)} \rightarrow \hat{z}$ (as $n \rightarrow \infty$), where \hat{z} is some element of R_{n_1} .

Consider the system of equations

$$A_2 \hat{z} = \bar{u} - A_1 \hat{z},$$

where

$$\hat{z} = \{\hat{z}_j : j = n_1 + 1, \dots, n\}, \quad A_1 = \{a_{ij}\}, \quad i = 1, \dots, m; \quad j = 1, \dots, n_1,$$

$$A_2 = \{a_{ij}\}, \quad i = 1, \dots, m; \quad n_1 + 1 \leq j \leq n.$$

This system has solutions $\hat{z}^{(n)}$ for any $\hat{z}^{(n)}$; consequently, the right-hand sides of this system $\bar{u} - A_1 \hat{z}^{(n)}$ satisfy the solvability conditions for the system with matrix A_2 . Hence it follows that the right-hand sides $\bar{u} - A_1 \hat{z}$ also satisfy the solvability conditions and there exists at least one solution of this system

$\hat{z} = \{z_i, i = n_1 + 1, \dots, n\}$. It is not difficult to prove that there exists a solution of the system under consideration satisfying conditions (1''). If the solution of the system $A_2 \hat{z} = \bar{u} - A_1 \hat{z}$ had no solutions satisfying (1''), then the solution of the system $A_2 \hat{z} = \bar{u} - A_1 \hat{z}^{(n)}$ also could not have solutions possessing this property, and this contradicts the definition $z^{(n)} = (\hat{z}^{(n)}, \hat{z}^{(n)})$. The element $\bar{z}^{(0)} = (\hat{z}, \hat{z})$ is a solution of the problem under consideration.

However, it is inappropriate, when specifying approximate input data, to use the algorithm for constructing an exact solution, since incorrect problems may occur. The next section is devoted to the construction of a stable algorithm.

5. The main purpose of the present article is to construct a stable algorithm for solving problem (1). This algorithm is defined by means of the function

$$M_\lambda^\alpha[z, A, \bar{u}, C] = \|Az - \bar{u}\|^2 + \alpha(C^2[z] + \lambda\Omega[z]) \quad (\alpha, \lambda > 0),$$

where

$$\Omega[z] = \|z\|^2 \quad \left(\text{or } \Omega[z] = \sum_{k,j} p_{kj}(z_k - z_{k0})(z_j - z_{j0}) \right)$$

($z^{(0)}$ is the prescribed element of R_n and $\sum p_{kj}\xi_k\xi_j$ is a positive definite form).

Let us denote $R_1^{(n)} = \{z_i : z_i \geq 0\}$, and let z^α be the element realizing the minimum of the function M_λ^α in R_1^n .

We shall measure the deviations of the input data and of the solution by means of the norms

$$\|A\| = \left(\sum_{i,j} a_{ij}^2 \right)^{1/2}, \quad \|u\| = \left(\sum_i u_i^2 \right)^{1/2}, \quad \|c\| = \left(\sum_i c_i^2 \right)^{1/2}, \quad \|z\| = \left(\sum_j z_j^2 \right)^{1/2}.$$

The following stability theorem holds for the algorithm constructed.

Theorem 2. *Suppose that problem (1) with input data A, \bar{u}, C has a normal solution $\bar{z}^{(0)}$. Let $\tilde{A}, \tilde{u}, \tilde{C}$ be arbitrary δ -approximations to A, \bar{u}, C ; let $\varepsilon(\delta), \alpha_0(\delta)$ be arbitrary decreasing functions of δ , tending to zero as $\delta \rightarrow 0$, and such that $\delta^2 \leq \varepsilon(\delta)\alpha_0(\delta)$.*

For any $\varepsilon > 0$, there exist $\lambda_0(\varepsilon)$ and $\delta_0(\varepsilon, \lambda)$ such that the element \tilde{z}^α realizing the minimum of the functional

$$M_\lambda^\alpha[z, \tilde{A}, \tilde{u}, \tilde{C}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha[\tilde{C}^2(z) + \lambda\Omega(z)] \quad (\text{in } R_1^{(n)}),$$

where α is any number such that

$$\frac{1}{\varepsilon(\delta)} \delta^2 \leq \alpha \leq \alpha_0(\delta),$$

satisfies the inequality

$$\|\tilde{z}^\alpha - \bar{z}^{(0)}\| \leq \varepsilon,$$

provided only that $\delta \leq \delta_0(\varepsilon, \lambda)$.

Received
20 IV 1964

REFERENCES

1. D. B. Yudin, E. G. Gol' dshtein, *Linear Programming*, Moscow, 1963; S. Vajda, in the collection *Linear Inequalities and Related Systems*, IL, 1959; L. V. Kantorovich, *Mathematical Methods in the Organization and Planning of Production*, Leningrad, 1939.
2. A. N. Tikhonov, DAN, 151, No. 3 (1963).
3. A. N. Tikhonov, DAN, 153, No. 1 (1963).
4. A. N. Tikhonov, DAN, 161, No. 5 (1965).
5. A. J. Goldman, A. W. Tucker, in the collection *Linear Inequalities and Related Systems*, IL, 1959.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.