SYSTEM OF LINEAR EQUATIONS

J. L. ARORA

(Received November 8, 1974)

1. INTRODUCTION

The system

(1.1) $$\sum_{j=1}^{n} \alpha_{ij} x_j = \beta_i \,, \quad i = 1, 2, \ldots, m$$

(α_{ij}'s and β_i's are real) of m equations in n unknowns can be written in the matrix form

(1.2) $$Ax^t = b^t \,,$$

where $A = (\alpha_{ij})$ is the coefficient matrix of the system (1.1), x is the unknown vector $(x_1, x_2, \ldots, x_n)$, b is the known vector $(\beta_1, \beta_2, \ldots, \beta_m)$ of scalars and t denotes the transpose.

In this paper we shall develop a method of solving the system (1.1) which is based upon the Gram-Schmidt orthogonalization process.

2. NOTATION

R: The field of real numbers.

R^n: The n-dimensional inner product space of n-tuples of real numbers with the inner product

(2.1) $$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \,,$$

where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$.

$\|x\|$: The norm of x.

r_i: The i-th row of the matrix A, i.e., the n-tuple of coefficients of the i-th equation of the system (1.1).

e_i: The vector $(0, 0, \ldots, 1, 0, \ldots, 0)$ of R^n (1 is in the i-th place).
$[r_1, r_2, \ldots, r_m]$: The span of the set of vectors $\{r_1, r_2, \ldots, r_m\}$.
(r_i, β_i): The vector $(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}, \beta_i)$.
(A, b): The augmented matrix

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} & \beta_2 \\ . & . & \ldots & . & . \\ . & . & \ldots & . & . \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} & \beta_m \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} (r_1, \beta_1) \\ (r_2, \beta_2) \\ . \\ . \\ (r_m, \beta_m) \end{pmatrix}.$$

3. AN INTERPRETATION OF THE SYSTEM

In this section an interpretation of the system (1.1) is given which is being used in developing the method. The system (1.1) can also be written as

(3.1) $$\langle r_1, x \rangle = \beta_i\,, \quad i = 1, 2, \ldots, m\,.$$

From the properties of the inner product, it follows that if we replace the k-th equation

(3.2) $$\langle r_k, x \rangle = \beta_k$$

by the equation

(3.3) $$\langle c_1 r_1 + c_2 r_2 + \ldots + c_k r_k, x \rangle = c_1 \beta_1 + c_2 \beta_2 + \ldots + c_k \beta_k\,,$$

c_i's not all zero, the solution of the system does not change. The solution set B of the system (3.1) is a linear variety $X_p + K$. A leader X_p is a particular solution of the system (3.1) and the base space K is the solution space of the associated homogeneous system

(3.4) $$\langle r_i, x \rangle = 0\,, \quad i = 1, 2, \ldots, m\,,$$

which is precisely the kernel of the matrix A.

We shall first develop a method of finding X_p, a particular solution of (3.1).

4. A PARTICULAR SOLUTION X_p

Consider the set

(4.1) $$\{r_1, r_2, \ldots, r_m\}$$

of row vectors of the matrix A. Applying the Gram-Schmidt orthogonalization process, [1], to it we get the set

(4.2) $$\{s_1, s_2, \ldots, s_m\}\,,$$

where s_i's are given by

(4.3)
$$s_1 = r_1 ,$$
$$s_i = r_i - \sum_{j=1}^{i-1} \frac{\langle r_i, s_j \rangle}{\langle s_j, s_j \rangle} s_j , \quad i = 2, 3, \ldots, m .$$

If a vector r_i, say r_{i_0}, is dependent on $r_1, r_2, \ldots, r_{i_0 - 1}$, then the vector s_{i_0} is the zero vector of R^n. In the process of finding s_{i_0+1} and onwards we shall ignore all such zero vectors.

Now consider the set of scalars $\{\delta_1, \delta_2, \ldots, \delta_m\}$ defined by

(4.4)
$$\delta_1 = \beta_1 ,$$
$$\delta_i = \beta_i - \sum_{j=1}^{i-1} \frac{\langle r_i, s_j \rangle}{\langle s_j, s_j \rangle} \delta_j , \quad i = 2, 3, \ldots, m .$$

It is clear from the process of getting the vectors s_i and the scalars δ_i, $i = 1, 2, \ldots$ $\ldots, m$, that the augmented matrix (A, b) is row equivalent to the matrix

(4.5)
$$\begin{pmatrix} (s_1, \delta_1) \\ (s_2, \delta_2) \\ \vdots \\ (s_m, \delta_m) \end{pmatrix} .$$

Hence the systems

(4.6)
$$\langle r_i, x \rangle = \beta_i , \quad i = 1, 2, \ldots, m$$

and

(4.7)
$$\langle s_i, x \rangle = \delta_i , \quad i = 1, 2, \ldots, m$$

have the same solutions. Since some of the s_i's are zero in (4.2) and (4.5) is row equivalent to the augmented matrix (A, b), it follows that the system (4.7) or equivalently (1.1) is consistent if and only if $\delta_i = 0$ whenever s_i is the zero vector of R^n. Throughout our discussions, we shall now assume that the system is consistent, i.e., $\delta_i = 0$ whenever s_i is the zero vector.

Since (4.6) and (4.7) have same solutions, we now consider the system given by (4.7) after ignoring those values of i for which we have the trivial identities $\langle 0, x \rangle = 0$.

(4.8) We rename the nonzero vectors s_i, $i = 1, 2, \ldots, m$ as v_i, $i = 1, 2, \ldots, m_0$ and the corresponding δ_i as μ_i.

Dividing each vector (v_i, μ_i), $i = 1, 2, \ldots, m_0$ by the norm of v_i, we get

(4.9)
$$(S_i, \lambda_i) , \quad i = 1, 2, \ldots, m_0 ,$$

where $S_i = v_i / \|v_i\|$ and $\lambda_i = \mu_i / \|v_i\|$.

Hence we get an equivalent system

$$(4.10) \qquad \langle S_i, x \rangle = \lambda_i \,, \quad i = 1, 2, \ldots, m_0 \,.$$

Using the properties of the inner product, we get from (4.10)

$$(4.11) \quad \langle c_1 S_1 + c_2 S_2 + \ldots + c_{m_0} S_{m_0}, x \rangle = c_1 \lambda_1 + c_2 \lambda_2 + \ldots + c_{m_0} \lambda_{m_0}$$

for any choice of scalars $c_1, c_2, \ldots, c_{m_0}$. (4.11) gives the value of x_i, if we can find a set of scalars $\{c_1, c_2, \ldots, c_{m_0}\}$ such that

$$(4.12) \qquad c_1 S_1 + c_2 S_2 + \ldots + c_{m_0} S_{m_0} = e_i \,.$$

Since the set $\{S_1, S_2, \ldots, S_{m_0}\}$ is orthonormal, it follows that

$$c_j = \langle e_i, S_j \rangle = v_{ji} \,,$$

where $S_j = (v_{j1}, v_{j2}, \ldots, v_{jn})$. Thus

$$x_i = \lambda_1 v_{1i} + \lambda_2 v_{2i} + \ldots + \lambda_{m_0} v_{m_0 i} = \sum_{j=1}^{m_0} \lambda_j v_{ji} \,.$$

Hence

$$(4.13) \qquad x = (\lambda_1, \lambda_2, \ldots, \lambda_{m_0}) (S_1, S_2, \ldots, S_{m_0})^t \,,$$

which is a particular solution X_p of the system (1.1). This particular solution is the unique solution if $m_0 = n$.

5. THE BASE SPACE K

The base space K of the linear variety, which is the solution of the system (3.1), is the solution of the associated homogeneous system (3.4). The system

$$(5.1) \qquad \langle v_i, x \rangle = 0 \,, \quad i = 1, 2, \ldots, m_0 \,,$$

where v_i, $i = 1, 2, \ldots, m_0$ are defined in (4.8), is equivalent to the system (3.4). Hence the solution set of (5.1) is the required subspace K of R^n. If $m_0 = n$, then K is the zero subspace of R^n, if $m_0 \neq n$, then we proceed as follows:

The solution set K of (5.1) is the orthogonal complement of $[v_1, v_2, \ldots, v_{m_0}]$, because

$$(5.2) \qquad \langle c_1 v_1 + c_2 v_2 + \ldots + c_{m_0} v_{m_0}, x \rangle = 0$$

for any choice of scalars c_i's. Since R^n is finite dimensional, it follows that

$$(5.3) \qquad K \oplus [v_1, v_2, \ldots, v_{m_0}] = R^n \,.$$

In order to find this orthogonal complement K, it is sufficient to find an orthogonal set W of $(n - m_0)$ generators of K such that each member of W is orthogonal to each member of the set $\{v_1, v_2, \ldots, v_{m_0}\}$. To obtain this set W, we consider the set

$$J = \{v_1, v_2, \ldots, v_{m_0}, e_1, e_2, \ldots, e_n\} . \tag{5.4}$$

J contains $(n + m_0)$ vectors of R^n and spans R^n because $\{e_1, e_2, \ldots, e_n\}$ is a basis of R^n. We now apply the Gram-Schmidt orthogonalization process to the set J and obtain the set

$$\{v_1, v_2, \ldots, v_{m_0}, u_{m_0+1}, u_{m_0+2}, \ldots, u_{m_0+n}\} , \tag{5.5}$$

where

$$u_i = v_i , \quad i = 1, 2, \ldots, m_0 \tag{5.6}$$

$$u_{m_0+i} = e_i - \sum_{j=1}^{m_0+i-1} \frac{\langle e_i, u_j\rangle}{\langle u_j, u_j\rangle} u_j , \quad i = 1, 2, \ldots, n .$$

If at any stage of the process a vector, say $u_{m_0+i_0}$, comes out to be the zero vector, we shall ignore this in the further steps of the process. The set (5.5) obtained by this process contains exactly n nonzero vectors. By deleting the zero vectors and renaming the remaining nonzero vectors, we get the orthogonal set $\{v_1, v_2, \ldots, v_{m_0}, v_{m_0+1}, v_{m_0+2}, \ldots, v_n\}$. Hence the set $W = \{v_{m_0+1}, v_{m_0+2}, \ldots, v_n\}$. Therefore

$$K = [v_{m_0+1}, v_{m_0+2}, \ldots, v_n] . \tag{5.7}$$

Hence the solution set of the system (1.1), if consistent, is the linear variety

$$(\lambda_1, \lambda_2, \ldots, \lambda_{m_0})(S_1, S_2, \ldots, S_{m_0})^t + [v_{m_0+1}, v_{m_0+2}, \ldots, v_n] . \tag{5.8}$$

6. EXAMPLE

Consider the system

$$\begin{aligned} x_1 \qquad + x_3 - x_4 + x_5 &= 1 \\ 2x_1 \qquad + x_3 - x_4 + x_5 &= 2 \\ 6x_1 + x_2 + 4x_3 \qquad + x_5 &= 6 . \end{aligned} \tag{6.1}$$

The augmented matrix of the system (6.1) is

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & -1 & 1 & 1 \\ 2 & 0 & 1 & -1 & 1 & 2 \\ 6 & 1 & 4 & 0 & 1 & 6 \end{array}\right). \tag{6.2}$$

Calculating s_i's and δ_i's, we get

$$(6.3)\qquad \begin{aligned} s_1 &= (1, 0, 1, -1, 1), & \delta_1 &= 1, \\ s_2 &= (\tfrac{3}{4}, 0, -\tfrac{1}{4}, \tfrac{1}{4}, -\tfrac{1}{4}), & \delta_2 &= \tfrac{3}{4}, \\ s_3 &= (0, 1, \tfrac{7}{3}, \tfrac{5}{3}, -\tfrac{2}{3}), & \delta_3 &= 0. \end{aligned}$$

Since none of the s_i's is the zero vector, it follows that the system is consistent and $v_i = s_i$, $\mu_i = \delta_i$, $i = 1, 2, 3$; $m_0 = 3$.

Now considering the set $\{v_1, v_2, v_3, e_1, e_2, e_3, e_4, e_5\}$ and applying the Gram-Schmidt orthogonalization process to it, we get

$$\begin{aligned} v_1 &= s_1, \quad v_2 = s_2, \quad v_3 = s_3, \\ u_4 &= (0, 0, 0, 0, 0), \\ u_5 &= (0, \tfrac{26}{29}, -\tfrac{7}{29}, -\tfrac{5}{29}, \tfrac{2}{29}) = v_4, \\ u_6 &= (0, 0, \tfrac{1}{26}, -\tfrac{3}{26}, -\tfrac{4}{26}) = v_5, \\ u_7 &= (0, 0, 0, 0, 0), \\ u_8 &= (0, 0, 0, 0, 0). \end{aligned}$$

Dividing v_i and μ_i, $i = 1, 2, 3$ by the norm of v_i, we get

$$\begin{aligned} S_1 &= (\tfrac{1}{2}, 0, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}), & \lambda_1 &= \tfrac{1}{2}, \\ S_2 &= (\sqrt{(3)}/2, 0, -1/(2\sqrt{3}), 1/(2\sqrt{3}), -1/(2\sqrt{3})), & \lambda_2 &= \frac{\sqrt{3}}{2}, \\ S_3 &= (0, 3/\sqrt{87}, 7/\sqrt{87}, 5/\sqrt{87}, -2/\sqrt{87}), & \lambda_3 &= 0. \end{aligned}$$

Hence the required solutions is

$$(6.4)\qquad (\tfrac{1}{2}, \sqrt{(3)}/2, 0)\begin{pmatrix} \tfrac{1}{2} & 0 & \tfrac{1}{2} & -\tfrac{1}{2} & \tfrac{1}{2} \\ \sqrt{(3)}/2 & 0 & -1/(2\sqrt{3}) & 1/(2\sqrt{3}) & -1/(2\sqrt{3}) \\ 0 & 3/\sqrt{87} & 7/\sqrt{87} & 5/\sqrt{87} & -2/\sqrt{87} \end{pmatrix}$$

$+ [(0, \tfrac{26}{29}, -\tfrac{7}{29}, -\tfrac{5}{29}, \tfrac{2}{29}), (0, 0, \tfrac{1}{26}, -\tfrac{3}{26}, -\tfrac{4}{26})]$ or $(1, 0, 0, 0, 0) + [(0, 26, -7, -5, 2), (0, 0, 1, -3, -4)]$.

7. REMARK

If the scalars α_{ij}'s and β_i's in (1.1) are complex numbers, then the above method can be used with some modifications.

References

[1] *K. Hoffman, R. Kunze:* Linear Algebra (2nd ed.), Prentice Hall of India (1972).

Souhrn

SYSTÉM LINEÁRNÍCH ROVNIC

J. L. Arora

Článek popisuje metodu řešení soustavy lineárních algebraických rovnic s reálnou obdélníkovou maticí. Metoda je založena na dvojím užití Gramovy-Schmidtovy ortogonalizace. Řešení dané soustavy se hledá ve tvaru $x = x_p + y$, kde x_p je partikulární řešení soustavy a y je z prostoru řešení přidružené homogenní soustavy.

Author's address: Dr. *J. L. Arora*, Department of Mathematics, Birla Institute of Technology and Science, Pilani, Rajasthan, India.