SIAM Review
Vol. 3, No. 1, January, 1961

A NOTE ON SYSTEMS OF LINEAR EQUATIONS[1]

David L. Elliott[2]

This note is a comment on reference [1] and a generalization of the method there presented. We consider a system of m linear equations in n unknowns $x_1, x_2, \cdots, x_n$,

$$\sum_{j=1}^{n} a_{ij}x_j = c_i \qquad i = 1, 2, \cdots, m, a_{ij}, c_i \text{ real} \tag{1}$$

or $A \cdot x = c$ in matrix notation. We distinguish three cases:

(I) There is *no* finite vector x satisfying (1) (inconsistent case);

(II) There is a *unique* vector x satisfying (1);

(III) There are an *infinity* of vectors x satisfying (1), such that their end-points lie on some line, plane, or higher-dimensional linear manifold.

Semarne considers Case II for $m = n$, but the presentation in reference [1] requires modification, as will be shown below.

Semarne's method of solving (1) combines the two classical ideas of the augmented-matrix equation and Gram-Schmidt orthogonalization. (The method does *not* involve computing a matrix inverse to A.) As modified in this report, the method will handle any system (1) whether or not A has an inverse, furnishing a solution if one exists and informing us if one does not exist.

We shall use vectors (or points)

$$x = (x_1, x_2, \cdots, x_n)$$

in E^n (euclidean n-space),

$$y = (y_0, y_1, \cdots, y_n)$$

in E^{n+1}. We extend E^n by introducing, in addition to the x_i, an artificial variable t. Multiplying (1) by t and transposing, we obtain:

$$-c_it + \sum_{j=1}^{n} a_{ij}tx_j = 0 \qquad i = 1, 2, \cdots, m, \tag{1'}$$

or

$$\sum_{j=0}^{n} b_{ij}y_j = 0 \qquad i = 1, 2, \cdots, m, \tag{2}$$

where

$$b_{i0} = -c_i, \quad b_{ij} = a_{ij} \qquad j = 1, 2, \cdots, n,$$

and

$$y_j = tx_j, 1 \leqq j \leqq n, \text{ and } y_0 = t. \tag{3}$$

[1] Received by the editors August 10, 1959.

[2] U. S. Naval Ordnance Test Station, Pasadena Annex.

(This notation differs from that of reference [1].) If x is a solution of (1), then for arbitrary t (3) gives a solution of (2) lying on a line through the origin of E^{n+1}; $(1, x_1, \cdots, x_n)$ is the intersection of this line with the hyperplane $y_0 = 1$. Conversely, if we find any point y^* satisfying (2), then every point on the line $y = sy^*$, $-\infty < s < \infty$, also satisfies (2). If $y_0^* \neq 0$, let $s^* = 1/y_0^*$; for the point $y = s^*y^*$, $y_0 = 1$. Then in (3) we have $t = y_0 = 1$,

$$s^*y_j^* = y_j = x_j, \qquad 1 \leqq j \leqq n, \tag{4}$$

and x satisfies (1).

If $y_0^* = 0$, the line $y = sy^*$ lies in the hyperplane $y_0 = 0$ and can have no finite intersection with $y_0 = 1$, so no corresponding solution of (1) can be found. If $y_0 = 0$ for every y satisfying (2), we classify (1) as Case I, inconsistent. We shall give explicit forms for solutions of (2) and (1) below.

Equation (2),

$$B \cdot y = 0,$$

where B is the m by $n + 1$ matrix (b_{ij}), can be interpreted geometrically in two ways. The first is that the vector y lies in the intersection of m hyperplanes (each passing through the origin). That is, the "solution" of (2) is a vector subspace, U, of E^{n+1}: the origin itself, or some line, plane, etc., through the origin. Defining the vector $(b_{i1}, b_{i2}, \cdots, b_{in}) = b^i$ for any $i = 1, 2, \cdots, m$, each hyperplane has equation

$$b^i \cdot y = 0 \qquad i = 1, 2, \cdots, m. \tag{2'}$$

Then the vector b^i is perpendicular to the ith hyperplane. Our second interpretation of (2) is "any solution-vector y is perpendicular to the m vectors $b^1, \cdots, b^m$," and a "solution" is the subspace U which can be described as the orthogonal complement in E^{n+1} of the subspace V spanned by $b^1, \cdots, b^m$. (That is, any point in V can be written as a linear combination of the b^i.) Then we need to find a set of vectors which span U, in order to define the solution of (2) completely.

From the viewpoint of the second interpretation, above, a good solution method is evident, namely, the Gram-Schmidt orthogonalization procedure. We note these facts:

a. One complete basis of orthogonal vectors for E^{n+1} is the set

$$\{u\} = \begin{cases} u^0 = (1, 0, \cdots, 0, 0), \\ u^1 = (0, 1, \cdots, 0, 0), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ u^n = (0, 0, \cdots, 0, 1). \end{cases}$$

b. We can form an orthogonal basis for V, consisting of M mutually orthogonal vectors $(M \leqq m)\, f^1, f^2, \cdots f,^M$, from the (not necessarily independent) vectors $b^1, \cdots, b^m$ by using the Gram-Schmidt process.

c. If we continue the Gram-Schmidt process by adjoining vectors—say, from $\{u\}$—which do not belong to V, we eventually will obtain a new complete orthogonal basis for E^{n+1},

$$f^1, f^2, \cdots, f^M, f^{M+1}, \cdots, f^{n+1},$$

and the vectors (if any) $f^{M+1}, \cdots, f^{n+1}$ will be a complete orthogonal basis for U, hence a "solution". If $M = n + 1$, U contains only the origin.

As Semarne points out, the beauty of this approach is that the Gram-Schmidt computation is quite simple; the basis vectors are:

$$\begin{aligned} f^1 &= b^1, \\ f^2 &= b^2 - \left(\frac{f^1 \cdot b^2}{f^1 \cdot f^1}\right) f^1, \\ &\cdots\cdots\cdots \\ f^m &= b^m - \left(\frac{f^1 \cdot b^m}{f^1 \cdot f^1}\right) f^1 - \left(\frac{f^2 \cdot b^m}{f^2 \cdot f^2}\right) f^2 - \cdots - \left(\frac{f^{m-1} \cdot b^m}{f^{m-1} \cdot f^{m-1}}\right) f^{m-1}. \end{aligned} \tag{5}$$

Now M ($\leqq m$) of these vectors are non-zero; we may omit the others and renumber the non-zero ones as $f^1, \cdots, f^M$; they span V. Now we continue computing until we obtain a total of $n + 1$ non-zero vectors:

$$\begin{aligned} f^{M+1} &= u^0 - \left(\frac{f^1 \cdot u^0}{f^1 \cdot f^1}\right) f^1 - \cdots - \left(\frac{f^M \cdot u^0}{f^M \cdot f^M}\right) f^M \\ &= u^0 - \left(\frac{f_0^{\ 1}}{f^1 \cdot f^1}\right) f^1 - \cdots - \left(\frac{f_0^{\ M}}{f^M \cdot f^M}\right) f^M, \\ f^{M+2} &= u^1 - \left(\frac{f_1^{\ 1}}{f^1 \cdot f^1}\right) f^1 - \cdots - \left(\frac{f_1^{\ M+1}}{f^{M+1} \cdot f^{M+1}}\right) f^{M+1} \\ &\cdots\cdots\cdots \\ f^{n+1} &= u^n - \left(\frac{f_n^{\ 1}}{f^1 \cdot f^1}\right) f^1 - \cdots - \left(\frac{f_n^{\ n}}{f^n \cdot f^n}\right) f^n, \end{aligned} \tag{5'}$$

where again we have dropped all f's which are zero. The "solution" of (2) is then the space U consisting of all y such that:

$$y = \alpha_{M+1} f^{M+1} + \cdots + \alpha_{n+1} f^{n+1}, \text{ for any real } \alpha_i. \tag{6}$$

We can obtain the solution of (1) as the intersection (if any) of U with the hyperplane $y_0 = 1$; that is, points of U having $y_0 = 1$ satisfy (1). We can now, from the vectors of (5′), decide which case (of the first paragraph) we have:

(I) If $f_0^{\ K} = 0$ for all the f^K in (5′), U lies in the hyperplane $y_0 = 0$, and no finite solution of (1) exists.

(II) If it happens that $M = n$ and that $f_0^{\ n+1} \neq 0$, (6) becomes $y = \alpha_{n+1} f^{n+1}$; we take $s_0 = 1/f_0^{\ n+1}$ and write:

$$s_0 y = (1, x_1, \cdots, x_n);$$

x solves (1) uniquely. (See note below.)

(III) If $M < n$ and some $f_0^K \neq 0$, $K > M$, (2) is satisfied, from (6), by

$$\bar{y} = \frac{\sum_{K=M+1}^{n+1} \alpha_K f^K}{\sum_{K=M+1}^{n+1} \alpha_K f_0^K} \quad \text{whenever} \quad \sum_{K=M+1}^{n+1} \alpha_K f_0^K \neq 0. \tag{7}$$

It is evident that $\bar{y}_0 = 1$, so $\bar{y} = (1, x_1, \cdots, x_n)$ and x solves (1) if the conditions of (7) are satisfied. Thus, (1) is completely solved.

NOTE. Reference [1] considers only Case II, for $m = n$, but the author forms f^{n+1} *not* from u^0, but from his vector:

$$e^{n+1} = (1, 1, \cdots, 1, 1).$$

This vector is unsatisfactory for the following reason. In reference [1], the matrix A of (1) is assumed n-by-n non-singular; then no combination of its row vectors a^i can satisfy:

$$\sum_{i=1}^{n} \beta_i a^i = 0, \qquad \sum_{i=1}^{n} \beta_i \neq 0.$$

Then no combination of the b^i of the matrix B can equal u^0, so $f^{n+1} \neq 0$ and we obtain our solution. But if we replace u^0 with e^{n+1} in the orthogonalization it may occur that, for a non-singular A, some combination of the rows of B will equal e^{n+1}; then we get a zero solution of (2), and the solution of (1) is indeterminate, as in the example:

$$\begin{aligned} 2x_1 + x_2 &= -2, \\ x_1 + 0 &= -1. \end{aligned} \tag{8}$$

In such a case, reference [1] gives $f^3 = (0, 0, 0)$, whereas our $f^3 = (\frac{1}{2}, -\frac{1}{2}, 0)$, giving the solution $x_1 = -1$, $x_2 = 0$.

REFERENCE

1. H. MANVEL SEMARNE, *New direct method of solution of a system of simultaneous linear equations*, SIAM Review, 1, 1(1959), pp. 53–54.