NONNEGATIVE SOLUTIONS OF LINEAR ALGEBRAIC SYSTEMS WITH RATIO SCALE COEFFICIENTS

Thomas L. Saaty
Luis G. Vargas

University of Pittsburgh

Abstract: Ratio scales are the life blood of the AHP. By definition ratio scales are nonnegative. Sometimes one is faced with solving a system of linear equations involving ratio scales and solving equations can give rise to negative numbers. This note provides a partial discussion of the relation between ratio scales, the solution of equations and the meaningfulness of negative numbers.

Keywords: Linear, equations, ratio scale

1. INTRODUCTION

There are areas of mathematics such as probability theory where negative numbers are not allowed, which make appropriate use of subtraction to always produce nonnegative numbers. Similarly, when all the variables of a system of equations belong to a ratio scale as do yards and kilograms, invariant under multiplication by a positive constant, a negative solution is meaningless. Negative numbers cannot belong to a ratio scale. What is needed is a way of generating positive numbers that can be associated with the solutions, and a way to interpret the meaning of negative numbers obtained from them. In his *Complete Introduction to Algebra (1770)*, Euler justified the operation of subtracting *-b* as equivalent to adding *b* because "to cancel a debt signifies the same as giving a gift." Apparently, more than a thousand years before Euler, the Indians used negative numbers to represent debts, leading to the acceptance of negative coefficients and negative solutions of equations [1]. Descartes called negative roots of equations false on the ground that they claim to represent numbers less than nothing. This sort of puzzling over the distinction between subtraction and negative numbers haunted mathematicians in the late 18th and early 19th century [3]. Resolution is still needed, for example, for those who deal with systems of equations involving ratio scales and must explain the result.

If we use mathematics to model a real life problem, and if inherent in that problem it is not possible to get a negative solution, then a negative solution from a model means that the problem is modelled incorrectly. One way to validate the model is to obtain a solution. If it is negative, we know that there is something wrong with the model, or the problem itself is infeasible.

Suppose a tailor makes two standard products, pants and shirts, using the same material. Suppose that in a certain week he makes five pants and two shirts using 20 yards of material. The week before, he made four pants and seven shirts using 25 yards of material. Now he has an order for pants and shirts and wants to buy material to make them, how much should he buy? The answer to this question requires knowledge of how much material the tailor uses to make a unit of each product, if each product requires the same amount each time. We have:

$$5x + 2y = 20$$
$$4x + 7y = 25$$

and x = 10/3 and y = 5/3. If the solution value of one of the variables were negative, we would conclude that something must be wrong in the statement of the problem. But in some problems, as mentioned by Euler, it can happen that a variable which represents giving rather that taking may be negative. How do we give meaning to it?

Our goal in this note is to provide a representation of the solution using nonnegative vectors. However, we may still use negative numbers as an auxiliary to derive positive solutions, just as complex numbers are used in Electrical Engineering to represent the magnitude and phase of signals, and Fourier series are used to represent the frequency spectrum. Our assumption is that we cannot start with negative numbers in our system of equations because they are meaningless on ratio scales or as probabilities. Negative numbers are an elegant way for representing solutions when we use our traditional one-sided approach to solve a system of linear equations by putting all the variables on one side and inverting the matrix of coefficients. An alternative way is to use a two-sided or balancing approach, thus placing only positive numbers on both sides of the system, and then imposing conditions to obtain nonnegative solutions for each side, whose combination yields the kind of solution with negative numbers derived by the one-sided approach. One may argue that a negative number can be written in an infinite number of ways as the difference of two positive numbers placed on two sides of an equation. It is this kind of observation that we want to address by directly deriving unique positive solutions.

2. NONNEGATIVE SOLUTIONS - THE MAIN THEOREM

A linear algebraic system

$$Ax=b \tag{1}$$

may or may not have a nonnegative solution $x = (x_1,\dots,x_m)^T$ even when $A = (a_{ij}) \geq 0$, $i=1,\dots,n$, $j=1,\dots,m$ and $b = (b_1,\dots,b_n)^T \geq 0$. The Minkowski-Farkas theorem [2] is concerned with the existence of a nonnegative solution:

"*Given an equation $Ax=b$, where b is an element of R^n, a necessary and sufficient condition for a solution $x \geq 0$ to exist is that $u^Tb \geq 0$ holds for any vector u such that $u^TA \geq 0$.*"

It is well known that a positive system of equations does not always have a positive solution. The system of 2 equations in 2 unknowns:

$$\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

has the solution

$$x = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

A well known characterization of positive solutions is:

Theorem 1. *A necessary and sufficient condition for (1) to have a positive solution $x>0$ is that b can be represented as a positive linear combination of the columns of A.*

<u>Proof</u>: (Necessity) Let $b = \sum_{j=1}^{m} \alpha_j A_j,\ A_j = Ae_j,\ e_j = (\delta_{ij}),\ \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases};\ \alpha_j > 0$, for all j. If we

write $x_j = \alpha_j$, for all j, we have $x > 0$ and the result follows.

(Sufficiency) $b = \sum_{j=1}^{m} x_j A_j \equiv \sum_{j=1}^{m} \alpha_j A_j$ and the proof is complete.

If we normalize the columns of A and b to unity we obtain what is called a stochastic system:

$$A'x' = b'$$

where

$$A' = A \begin{bmatrix} (\sum_{i=1}^{n} a_{i1})^{-1} & 0 & \cdots & 0 \\ 0 & (\sum_{i=1}^{n} a_{i2})^{-1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & (\sum_{i=1}^{n} a_{im})^{-1} \end{bmatrix}$$

and

$$b_j' = b_j (\sum_{i=1}^{n} b_i)^{-1}.$$

with solution:

$$x_j' = x_j (\sum_{i=1}^{n} a_{ij}) (\sum_{i=1}^{n} b_i)^{-1},\ j=1, 2, \ldots, m$$

and we now have:

Corollary 1. *A necessary and sufficient condition for (1) to have a positive solution $x>0$ is that b belongs to the convex hull of the columns of A'.*

Corollary 2. *A necessary condition for the vector b' to belong to the convex hull defined by the columns of the matrix A' is that:*

$$\min_j \{a_{ij}'\} \le b_i' \le \max_j \{a_{ij}'\},\ \textit{for all } i.$$

Proof: Assume that there exists $\alpha_j > 0$, for all j, $\sum_{j=1}^{m} \alpha_j = 1$, such that $b'_i = \sum_{j=1}^{m} a_{ij}' \alpha_j$.

We have, for all i, $\min_j\{a_{ij}'\} = \sum_{j=1}^{m} \min_j\{a_{ij}'\}\alpha_j \le b_i' \le \sum_{j=1}^{m} \max_j\{a_{ij}'\}\alpha_j = \max_j\{a_{ij}'\}$.

The foregoing theory is descriptive rather than constructive. As for the latter we note that while matrix inversion is the standard procedure for deriving solutions to a linear system of equations, it relies on the use of negative numbers and does not ensure nonnegative solutions. A constructive procedure for obtaining a nonnegative solution relies on the Perron-Frobenius theorem [4]. It identifies the principal eigenvector of a nonnegative matrix of coefficients of a homogeneous system as a nonnegative solution. That vector can be estimated from the limiting powers of the matrix, and it is given by:

$$\lim_{k \to \infty} \frac{A^k e}{e^T A^k e}.$$

where $e = (1, \ldots, 1)^T$.

If a positive solution exists, it can be easily shown that the solution of $Ax = b$ is also a solution of a related eigenvalue problem.

Define $Diag(x) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & x_m \end{bmatrix}$ and $Diag(b) = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_n \end{bmatrix}$

Theorem 2. *If (1) has a nonnegative solution, there exists a matrix*

$$B = Diag(x)A^T\, Diag(b)^{-1} \tag{2}$$

such that x is the principal right eigenvector of the eigenvalue problem:

$$BAx = x. \tag{3}$$

Proof: If we write (1) as

$$ADiag(x)e = Diag(b)e \tag{4}$$

and multiplying by both sides on the left by $Diag(b)^{-1}$ we obtain:

$$Diag(b)^{-1}ADiag(x)e = e. \tag{5}$$

Existence of a solution of (1) implies that there is a matrix B such that $x = Bb$ which in turn can be written as:

$$Diag(x)e = BDiag(b)e. \tag{6}$$

Substituting for (4) in (6) yields:

$$Diag(x)e = BADiag(x)e \text{ or } BAx = x.$$

Similarly, it can be shown that $ABb = b$ holds. Thus, we have:

$$\begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \begin{bmatrix} b \\ x \end{bmatrix} = \begin{bmatrix} b \\ x \end{bmatrix}.$$

This system has a solution if and only if [5]:

$$A = Diag(b)B^TDiag(x)^{-1}$$

from which (2) follows.

Thus, if we know that the system has a positive solution we can find it by solving an eigenvalue problem using Perron-Frobenius theory thus avoiding the occurrence of both subtraction and negative numbers which appear in matrix inversion. We still have for a general system of linear equations that:

Theorem 3. *The solution of $Ax = b$ is the principal right eigenvector of*

$$\left[A^TDiag(b)^{-1}ADiag(x)\right]^T.$$

Proof: From Theorem 2 and $BAx = x$ we have

$$Diag(x)A^TDiag(b)^{-1}Ax = x \tag{7}$$

or

$$x^T\left[A^TDiag(b)^{-1}ADiag(x)\right] = x^T$$

and the result follows.

An algorithm for deriving the solution is based on the following iterations:

Let $x^{(k)}$ be the kth estimate of the solution and let

$$x^{(k+1)} = x^{(k)} + [I - A^TDiag(b)^{-1}ADiag(x^{(k)})]e \tag{8}$$

It can be shown that $x^{(k+1)}$ converges to $[A^TDiag(b)^{-1}A]^{-1}e$, which is the desired positive solution.

When positive solutions do not exist, one can decompose the system and the solution into two corresponding nonnegative components. When the solution components are added, they yield the solution of the entire system. This stronger approach using subtraction but not negative numbers which bedevil one in the context of ratio scales, is effected by transforming the original system as we now show.

3. General Systems of Linear Equations

Consider a system of equations $Ax = b$ where A is an arbitrary real matrix. We write this system as:

$$A_1x = b + A_2x \tag{9}$$

where $A = A_1 - A_2$, $A_1 \geq 0$ and $A_2 \geq 0$. It is convenient to take A_1 as the positive part of A and A_2 as its negative part. If this system has a solution, then there exist non-negative vectors x_1, x_2, b_1 and b_2 such that $Ax_i = b_i$, $i = 1, 2$, $A(x_1 - x_2) = b$, and $b = b_1 - b_2$.

A negative solution (a solution with some negative components) can be interpreted as a pair of nonnegative vectors x_1 and x_2 that satisfy:

$$\begin{aligned} A_1x_1 &= A_2x_1 + b_1 \\ A_1x_2 &= A_2x_2 + b_2 \\ b_1 &= b + b_2 \end{aligned} \tag{10}$$

The solution of this system can be obtained by solving the linear program:

$$\text{Minimize} \sum_{j=1}^{m} (x_{1j}+x_{2j})$$

subject to:

$$\begin{bmatrix} A_1 & -A_2 & 0 & 0 & -I & 0 \\ 0 & 0 & A_1 & -A_2 & 0 & -I \\ 0 & 0 & 0 & 0 & I & -I \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \\ x_2 \\ x_2 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix} \tag{11}$$

$$x_1,\ x_2,\ b_1,\ b_2 > 0$$

where $x_k = (x_{k1},\ldots,x_{km})$, k=1,2, and $b_k = (b_{k1},\ldots,b_{kn})$, k=1,2.

The solution of the system can now be written as a 4-tuple of positive vectors:

$$x = (x_1,\ b_1; x_2,\ b_2).$$

As an illustration, the system:

$$\underset{\mathbf{A}}{\begin{bmatrix} 3 & 3 & 2 \\ 2 & -3 & -2 \\ 4 & -1 & -2 \end{bmatrix}} \underset{\mathbf{x}}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} = \underset{\mathbf{b}}{\begin{bmatrix} 3 \\ 7 \\ 5 \end{bmatrix}}$$

can be solved by decomposing the matrix **A** into

$$A_1 = \begin{pmatrix} 3 & 3 & 2 \\ 2 & 0 & 0 \\ 4 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

and solving the corresponding linear programming problem given by (11). We have:

$$x_1 = \begin{pmatrix} 6.5 \\ 0 \\ 3 \end{pmatrix},\ x_2 = \begin{pmatrix} 4.5 \\ 3 \\ 0 \end{pmatrix},\ b_1 = \begin{pmatrix} 25.5 \\ 7 \\ 20 \end{pmatrix},\ \text{and}\ b_2 = \begin{pmatrix} 22.5 \\ 0 \\ 15 \end{pmatrix}.$$

from which, if we allow negative (nonratio scale) numbers, the solution of the original system is obtained by writing $x = x_1 - x_2 = (2, -3, 3)^T$.

This formulation is equivalent to the well known linear programming problem in which we ensure a nonnegative solution by replacing constraints by:

$$A(x^1 - x^2) = b$$

where x^1 *and* x^2 are nonnegative, which can be written as:

$$A_1x^1 + A_2x^2 = b + A_1x^2 + A_2x^1$$

where as before $A = A_1 - A_2$.

References

[1] S. Hollingsdale, *Makers of Mathematics*, Penguin Books, 1989.

[2] J. Farkas, "Uber die Theorie der Einfachen Ungleichungen," J. Reine Agnew. Math. 124 (1902) 1-27.

[3] M. Kline, *Mathematics: The Loss of Certainty*, Oxford University Press, 1980.

[4] T.L. Saaty, *The Analytic Hierarchy Process*, RWS Publications, 4922 Ellsworth Ave., Pittsburgh, PA 15213, 1992.

[5] T.L. Saaty and L.G. Vargas, "Diagnosis with Dependent Symptoms: Bayes Theory Derived from the Analytic Hierarchy Process," in *Fundamentals of Decision Making* by Thomas L. Saaty, RWS Publications, 1994.