THE GENERALIZED SOLUTION TO A LINEAR OPERATOR EQUATION. THE CASE OF FINITE DIMENSION

HORSKÝ Richard (CZ)

Abstract. It is well-known that a system of linear algebraic equations need not have a solution. There is a way allowing to define a solution to such a system which is not dependent on the properties of the linear operator (injective, surjective) represented by the matrix of the system. This leads in special case to the known Least Square Method, however here in rather unusual point of view.

Keywords: system of linear algebraic equations, normal solution, Moore Penrose pseudoinverse

Mathematics Subject Classification: Primary 15A09, 47A05; Secondary 15A18

1 Basic Notions and Facts

In this section we recall some basic notions and facts and accept suitable notation. Suppose we have a linear mapping (operator)

$$A:\Re^n \to \Re^m, \tag{1}$$

where $\Re^n$ is the n-th dimensional Euclidean space. This mapping is represented by a matrix $\mathbf{A}=(a_{ij})$ of the type (m,n). For our purposes it is not necessary to distinguish strictly between the linear operator and its matrix representation. So we will use the letter *A* instead of **A**. and the word operator may be exchanged by the word matrix whenever it occurs in the sequel. We will consider the operator equation

$$Ax=b, \tag{2}$$

which is in fact the system of *m* linear algebraic equations for *n* unknowns ξ_i, $x=(\xi_1,\dots,\xi_n)$. Two important sets are connected with the operator (1). They are the *kernel* and the *range of the operator* (1):

$$\ker A = \{x \in \Re^n : Ax = o_m\}, \quad \operatorname{ran} A = \{y \in \Re^m : \exists (x \in \Re^n) y = Ax\}.$$

The symbol o_m is used for the zero vector of the space $\Re^m$. The Euclidean space is equipped by the inner product which is here taken in general sense, see [5] (page 110). It is not necessary to consider the standard Euclidean inner product. An inner product of the vectors x and y is usually denoted (x, y) and to distinguish among the inner products in different Euclidean spaces we write $(x, y)_k$ for the inner product in $\Re^k$. The vector norm is derived from the inner product in usual way: $\|x\| = \sqrt{(x, x)}$. We distinguish again among the norms in the different Euclidean spaces by the index. The inner product allows to define the so called *adjoint operator (conjugate transpose) of A*

$$A^* : \Re^m \to \Re^n \tag{3}$$

by the equality

$$(Ax, y)_m = (x, A^* y)_n$$

for any $x \in \Re^n, y \in \Re^m$. Its existence and uniqueness follow from Riesz representation theorem, see [5](page 237).

In the end of this section let us recall the relations among kernels and ranges of the operator (1) and its adjoint (3):

$$\Re^n = \ker A \oplus \operatorname{ran} A^*, \quad \Re^m = \operatorname{ran} A \oplus \ker A^*. \tag{4}$$

The direct decompositions (4) are even orthogonal. It means for example that any vector $x \in \Re^n$ can be expressed in the form $x = v + w$, where $Av = o_m$, $w = A^* u$ for some $u \in \Re^m$ and $(v, w)_n = 0$. The vectors v and w are determined by x uniquely.

2 Generalized Solution to the System of Linear Algebraic Equations

We define

$$F(x) = (Ax - b, Ax - b)_m = \|Ax - b\|_m^2. \tag{5}$$

The function (5) is called the *price functional* (terminology used in [1]). It is a nonnegative function $F : \Re^n \to \Re$. We will consider the problem of its minimum. The *generalized solution* to the system of linear algebraic equations (2) is any point at which the price functional (5) achieves its minimum.

Proposition 2.1
There is at least one point at which the price functional achieves its minimum.

Proof. Due to the orthogonal decompositions of Euclidean spaces (4) we may write

$$x = v + A^* u\,,\ Av = o_m\,,$$
$$b = Ac + d\,,\ A^* d = o_n\,.$$

Hence we obtain

$$\begin{aligned} F(x) = \left(A\left(v + A^* u\right) - Ac - d, A\left(v + A^* u\right) - Ac - d\right)_m = \\ = \left(A\left(A^* u - c\right), A\left(A^* u - c\right)\right)_m + \left(d, d\right)_m\,. \end{aligned} \tag{6}$$

Since the vector $d \in \Re^m$ is determined by the fixed vector $b \in \Re^m$ we may affect only the first term in (6). Thus we set $x = c$. Due to $Ac = AA^* u$ we obtain $F(x) = (d, d)_m$. □

The minimal value $F(x) = (d, d)_m$ is so called the *price of the generalized solution.* It is obvious that in the case there exists the classical solution to (2) the price is zero.

There may exist more than one generalized solution for the given operator *A* and the vector *b* in (2). Let us denote

$$S = \left\{x \in \Re^n : F(x) = (d, d)_m\right\}.$$

The set *S* is not empty as follows from the Proposition 2.1. Further *S* is a closed set because the function (5) is continuous. Finally *S* is convex. Suppose $x \in S,\ y \in S,\ t \in [0,1]$. If we set $z(t) = tx + (1 - t)y$ then we obtain $F(z(t)) = (d, d)_m$ and hence $z(t) \in S$ for any $t \in [0,1]$. In the case that (2) has a classical solution the set *S* is a linear set of the form $c + \ker A,\ Ac = b$. However, if the price of the generalized solution is positive, the structure of *S* is the same as in the case of the classical solution.

Proposition 2.2
There is a unique solution among the generalized solutions which has the least norm. This solution is called normal solution.

Proof. The claim of this Proposition says that there is a unique vector $x^* \in S$ such that

$$\left\|x^*\right\|_n = \min\left\{\|x\|_n : x \in S\right\}.$$

However it follows from the fact that *S* is nonempty, closed and convex set, see [5](page 232). □

Remark 2.3
Let us notice that we have no requirements as to the inner products in Euclidean spaces. If we consider the standard inner product (Euclidean norm is derived from this product) we come to the well-known Least Square Method. Particularly, set $x = (\xi_1, \ldots, \xi_k),\ y = (\eta_1, \ldots, \eta_k)$. The

standard inner product in $\Re^k$ is $(x,y)_k = \sum_{i=1}^{k} \xi_i \eta_i$ and hence the price functional (5) is of the form

$$F(x) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \xi_j - b_i \right)^2 .$$

3 Moore-Penrose Pseudoinverse Operator

Let us recall one concept of a generalization of the notion of the inverse operator (matrix). We consider the operator (1) and the following system of equations

(i) $AXA = A$,
(ii) $XAX = X$,
(iii) $(AX)^* = AX$,
(iv) $(XA)^* = XA$.

The operator $X : \Re^m \to \Re^n$ (the matrix is of the type (*n,m*)). However, it is not clear whether such an operator exists and whether it is unique.

Remark 3.1
If there is $m = n$ in (1) and the operator (1) is injective then $X = A^{-1}$. It is obvious and we can verify this by direct substitution.

Proposition 3.2
For any operator (1) there is unique operator X satisfying (i)-(iv).

Proof of uniqueness. Suppose *X* and *Y* are matrices satisfying (i)-(iv). It holds

$$AX = (AX)^* = X^* A^* = X^* (AYA)^* = X^* A^* (AY)^* = (AX)^* (AY)^* = AXAY = AY .$$

It also holds

$$XA = (XA)^* = A^* X^* = (AYA)^* X^* = (YA)^* A^* X^* = YA(XA)^* = YAXA = YA .$$

We can conclude

$$X = X(AX) = X(AY) = (XA)Y = (YA)Y = Y .$$

Proof of existence. Consider the operator

$$T = A^* A : \Re^n \to \Re^n . \tag{7}$$

The operator (7) is self-adjoint (its matrix is symmetric) and nonnegative in the sense that $(Tx, x)_n = \|Ax\|_n^2 \geq 0$ for any $x \in \Re^n$. It has all eigenvalues nonnegative. We denote each of the different positive eigenvalues of (7) by the symbol λ_j, $j = 1,\ldots,s$. The operator (7) may have a non-trivial kernel (its matrix is singular) and then $\lambda_0 = 0$ is also its eigenvalue.
There are the spectral projections corresponding to positive eigenvalues λ_j, $j = 1,\ldots,s$

$$P_j : \Re^n \to V_j,\ j = 1,\ldots,s, \tag{8.1}$$

where V_j are the eigenspaces corresponding to λ_j. We define also the mapping

$$P_0 : \Re^n \to V_0, \tag{8.2}$$

which is the spectral projection for the eigenvalue $\lambda_0 = 0$ with eigenspace V_0 in the case (7) is singular and $P_0 = 0$ in the other case. We may write

$$T = \sum_{j=0}^{s} \lambda_j P_j. \tag{9}$$

The mappings (8.1) and (8.2) are orthogonal (symmetric) projections:

$$P_j P_k = 0,\ j \neq k,\ P_j^2 = P_j,\ P_j^* = P_j,\ I - P_0 = \sum_{j=1}^{s} P_j,\ j,k = 1,\ldots,s. \tag{10}$$

The expansion (9) is the *spectral decomposition of the operator* (7) and (10) is the well-known *spectral decomposition of unit.* Now we set

$$X = \sum_{j=1}^{s} \frac{1}{\lambda_j} P_j A^*. \tag{11}$$

We prove that (11) is desired pseudoinverse operator, i.e. it satisfies the properties (i)-(iv) above. Prove only (i). The remaining properties may be proved similarly.

$$\begin{aligned} AXA &= A\sum_{j=1}^{s} \frac{1}{\lambda_j} P_j A^* A = A\sum_{j=1}^{s}\sum_{k=1}^{s} \frac{\lambda_k}{\lambda_j} P_j P_k = \\ &= A\sum_{j=1}^{s} P_j = A(I - P_0) = A - AP_0. \end{aligned}$$

It suffices to show that $AP_0 = 0$. However it is obvious since the range of P_0 is the kernel of the operator T and $\ker A \subseteq \ker T$. □

We have showed that there is a unique operator satisfying the properties (i)-(iv). This is operator is called *Moore-Penrose pseudoinverse operator (matrix)* and it is usually denoted by the symbol A^+. As we noted in Remark 3.1 if A is injective operator then $A^+ = A^{-1}$. In this case $A^+ A = I$, however in general case it holds $A^+ A = I - P_0$.

4 The normal solution and the Moore-Penrose Pseudoinverse Operator

The pseudoinverse operator (11) is in fact $A^+ : \Re^m \to \Re^n$. Let us denote

$$x^+ = A^+ b. \tag{12}$$

We explain the relation between the vector (12) and the normal solution defined in Proposition 2.2.

Proposition 4.1
The normal solution of the system (2) is the vector (12).

Proof. Firstly we show that (12) is a generalized solution to (2). It means that the price functional (5) achieves its minimum at the point (12). Let us compute

$$F(x^+) = (Ax^+ - b, Ax^+ - b)_m = (AA^+b - b, AA^+b - b)_m = (AA^+Ac - Ac - d, AA^+Ac - Ac - d)_m =$$
$$= (A(I - P_0)c - Ac - d, A(I - P_0)c - Ac - d)_m = (d, d)_m.$$

The point of the minimum of (5) with the least norm has to be orthogonal to the kernel of the operator (1) because we observed that the set S of the minima of (5) is of the form $x^* + \ker A$. Using (11) we obtain

$$A^+ b = \sum_{j=1}^{s} \frac{1}{\lambda_j} P_j A^* b \in \operatorname{ran}(T).$$

Using the relations (4) applied on the self-adjoint operator T we have that $A^+ b$ is orthogonal to any vector lying in $\ker T$. Since $\ker A \subseteq \ker T$ (in fact even equality holds) we have proved the claim of the proposition.

Example. The lag operator B is of the basic meaning in the time series analysis, e.g. see [2]. In the finite dimensional case this operator is represented by the matrix

$$\mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \dots & \ddots & 1 \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix} \in M(n),$$

where $M(n)$ is the space of all matrix of the order *n*. It is clear that 0 is the unique eigenvalue of *B*. Its adjoint operator B^* is obviously represented by the matrix

$$\mathbf{B}^* = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \in M(n).$$

Finally the operator *T* defined by (7) has the matrix representation

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \dots & \ddots & 0 \\ 0 & 0 & \dots & \dots & 1 \end{bmatrix} \in M(n).$$

Its eigenvalues are $\lambda_0 = 0$ with eigenspace $V_0 = \mathrm{span}(e_1)$ and $\lambda_1 = 1$ with eigenspace $V_1 = \mathrm{span}(e_2, \dots, e_n)$ respectively. Here $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 on the *k*-th coordinate. In order to express the pseudoinverse operator B^+ in the form (11) it is necessary to find the projection P_1. However it is seen that $P_1 = T$. Hence we easily compute that $B^+ = B^*$.

5 Some notes to generalizations

The introduced approach to the generalized solution to the linear operator equation (2) is heavily dependent on the fact we consider the Euclidean space, i.e. the finite dimensional space equipped with an inner product. We may generalized in the direction of more generalized structure in a finite dimensional space: to take a normed linear space of the type $\ell_p(n)$, $p \in [1, +\infty]$, $p \neq 2$, see [4] or [5]. Another generalization is to relax the requirement of the finite dimension. The infinite dimensional analogue to the Euclidean space is the Hilbert space, i.e. the space of infinite dimension with an inner product which is complete with respect to the norm derived from this product. The typical example is the space of all square summable scalar sequences ℓ_2. This is the model of any separable Hilbert space. The price functional (5) corresponding to the linear operator $A : U \to H$, where *U* is a linear space and *H* is a Hilbert space achieves its minimum if *A* has the closed range. The proof is given in [4] (page 141). This proof is based again on Riesz projection theorem which is referred in the proof of Proposition 2.2 above. Unfortunately the range of *A* is not closed in many important situations.. It turns out that the sufficient and necessary condition for existence of the generalized solution to (2) is that $b \in \mathrm{ran}\, A \oplus (\mathrm{ran}\, A)^{\perp}$, where $(\mathrm{ran}\, A)^{\perp}$ is the orthogonal complement of $\mathrm{ran}\, A$.

The natural generalization of the operator (1) to infinite dimension is the compact operator, see [3](page 29). It is a consequence of the open mapping theorem that any compact operator of infinite rank between Hilbert spaces has non-closed range. Another example of an operator with non-closed range is given in [2]. It is the difference operator. In these cases there always exists vectors for which the price functional has no minimum and thus there is no generalized solution to the operator equation (2). Other generalizations and description of pseudoinverse operator can be found in [4].

References

[1] URL, http://mat.fsv.cvut.cz/marekivo/

[2] HORSKÝ, R., The lag operator and its spectral properties. *Mathematical Methods in Economics 2013.* Proceedings Part I (2013), 285-290

[3] LUKEŠ, J., Zápisky z funkcionální analýzy. Univerzita Karlova v Praze-Nakladatelství Karolinum, 2012, ISBN 978-80-246-2069-5

[4] NAIR, M. T., Linear operator equations; approximation and regularization. World Scientific Publishing Company, 2009

[5] TAYLOR, A. E., Úvod do funkcionální analýzy. Academia, Praha, 1973, ISBN 04-018-80, 216 s.

Current address

Horský Richard, RNDr., Ph.D.
Department of Mathematics , Faculty of Informatics and Statistics
University of Economics
Ekonomická 957, 148 00 Praha 4, Czech Republic
E-mail: rhorsky@vse.cz