PAUL KOSTRO
105-20 66th Road
Forest Hills, N. Y. 11375

Paul G. Kostro

3538 June, 1963 EM 3

Journal of the

ENGINEERING MECHANICS DIVISION

Proceedings of the American Society of Civil Engineers

MATRIX DIVISION

By George Kostro,[1] M. ASCE

SYNOPSIS

By adapting to classical matrixes one of the division methods of krakovian calculus, a scheme is presented for matrix division. It is shown that any matrix can be divided by a triangular one when divisor and dividend have the same number of rows. A rule is formulated for the division of a matrix by a square matrix, both matrixes having the same number of rows. The method yields a quotient matrix without previous inversion of any factor of the computation. One numerical example illustrates the procedure. This division method represents a general form of which Cholesky's scheme is a particular shortcut variant.

INTRODUCTION

Matrix algebra can be subdivided into three branches, each characterized by its own rules of matrix multiplication. Firstly, classical matrix calculus whose canons are built on multiplication of matrixes by rows into columns. Second, krakovian calculus, introduced in 1938 by T. Banachiewicz, of Krakow

Note.—Discussion open until November 1, 1963. To extend the closing date one month, a written request must be filed with the Executive Secretary, ASCE. This paper is part of the copyrighted Journal of the Engineering Mechanics Division, Proceedings of the American Society of Civil Engineers, Vol. 89, No. EM3, June, 1963.

1 Senior Structural Engr., Strobel & Rongved, New York, N. Y.

University, Poland, where matrixes are multiplied by columns into columns.[2] Third, determinant calculus which allows the multiplication of matrixes in any desired form—row into column, column into row, row into row, and column into row.[3]

The simplicity of division schemes in each of these branches of matrix algebra depends directly on the basic multiplication laws. Thus, in determinant or krakovian calculus the matrix division is a simple operation. However, in classical matrix calculus, it is assumed that matrixes cannot be divided.[4]

Rather, if in a given matrix equation,

$$[A]\,[X] = [W] \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

in which [ ] denotes a rectangular matrix, [A] is the square matrix, and [X] must be computed, one takes for the division-like operation, the pre-multiplication of both sides of the equation by the inverse of [A].

Thus

$$[X] = [A]^{-1}\,[W] \quad \ldots\ldots\ldots\ldots\ldots \quad (2)$$

in which $[\ ]^{-1}$ denotes the inverse of a matrix.

Some authors call this multiplication by an inverse a pre-division or post-division, depending on the operation order.[4] However, this is not division in an algebraic sense, which means a rule for finding how many times one quantity is contained in another. In the cited case, the division could be represented by the equation

$$[X] = [W] \div [A] \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

Eq. 3 is a form perfectly possible in determinant and krakovian calculus. The writer, spurred on by the mentioned divisibility limitation of classical matrixes, tried to adapt one of the division schemes of krakovian calculus to them. The present paper gives the results of this effort.

DIVISION BY A TRIANGULAR MATRIX

In order to define the division rules of any matrix by a triangular matrix (canonical or reduced), the following reasoning is applied. First consider a system of linear equations.

---

2 "Krakowiany," by Z. Dowgird, Panstwowe Wydawnictwo Naukowe, Warsaw, 1959.

3 "Determinants and Matrices," by A. C. Aitken, Oliver & Boyd, Edinburgh, 1959.

4 "Elementary Matrices," by R. A. Frazer, W. T. Duncan, and A. R. Collar, Cambridge Univ. Press, New York, N. Y., 1960.

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= w_1 \\ a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n &= w_2 \\ a_{33}x_3 + \dots + a_{3n}x_n &= w_3 \\ \dots\dots\dots\dots\dots & \\ a_{nn}x_n &= w_n \end{aligned}\right\} \quad \dots\dots \quad (4)$$

that is written in matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \dots \\ w_n \end{bmatrix} \quad \dots\dots \quad (5)$$

or abbreviating as

$$[A]\ \{X\} = \{W\} \quad \dots\dots\dots\dots\dots \quad (6)$$

with [A] being noted as an upper triangular matrix. The { } denotes a column matrix. If, instead of a single system of Eqs. 4, there is a set of systems corresponding to various columns of w_{ij}, Eq. 5 is represented

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{2n} & x_{3n} & \dots & x_{nm} \end{bmatrix}$$

$$= \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1m} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{nm} \end{bmatrix} \quad \dots\dots\dots\dots \quad (7)$$

or simply as

$$[A]\ [X] = [W] \quad \dots\dots\dots\dots\dots\dots \quad (8)$$

always stating that [A] is an upper triangular matrix.

Comparing Eq. 5 with Eq. 7 and Eq. 6 with Eq. 8, it is concluded that Eq. 8 represents a shorthand form of various independent systems of equations of the type of Eq. 6, with a factor [A] common to all of them.

$$\left.\begin{array}{l} [A]\ \{X_1\} = \{W_1\} \\ [A]\ \{X_2\} = \{W_2\} \\ [A]\ \{X_m\} = \{W_m\} \end{array}\right\} [A]\ [X] = [W] \quad \dots\dots\dots \quad (9)$$

The system of Eq. 4 is recognized as a simple, triangular, or reduced form. Its solution involves the determination of x_1 to x_w by the successive application of the equation

$$x_i = \frac{w_i - \left(a_{in}\ x_n + a_{i,n-1}\ x_{n-1} + \dots\ a_{i,i+1}\ x_{i+1}\right)}{a_{ii}} \quad \dots \quad (10a)$$

beginning with the x_n. In view of Eq. 9, Eq. 10a, rewritten as

$$x_{ij} = \frac{w_{ij} - \left(a_{in}\ x_{n,} + a_{i,n-1}\ x_{n-1,i} + \dots + a_{ij+1}\ x_{j+1,i}\right)}{a_{ij}} \quad \dots\dots \quad (10b)$$

can be applied for the solution of the system of Eq. 7 as well.

At this point the method for the matrical division of the system of Eq. 5 abbreviated to the Eq. 6 may be established. Writing

$$\{X\} = \frac{\{W\}}{[A]} \quad \dots\dots\dots\dots\dots\dots \quad (11)$$

because

$$\{X\} \equiv [X]^T \quad \dots\dots\dots\dots\dots\dots \quad (12)$$

Eq. 12 is

$$\{W\} \div [A] = [X]^T \quad \dots\dots\dots\dots\dots \quad (13)$$

in which $[\ ]^T$ denotes the transpose of a matrix.

After expanding and regrouping Eq. 11 becomes

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \cdot\cdot \\ w_n \end{bmatrix} \div \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}^T \quad \dots\dots\dots \quad (14)$$

From the system of Eq. 14, it is seen that by multiplying the row of $[X]^T$ into rows of $[A]$ the values of w_i are generated in order, one by one, starting with the w_n, and

$$w_i = a_{in}\, x_n + a_{i,n-1}\, x_{n-1} + \dots a_{ii}\, x_i \quad \dots\dots\dots \quad (15)$$

A transformation of Eq. 15 to extract x_i permits the definition of all x in order, beginning with x_n, or

$$x_i = \frac{w_i - \left(a_{in}\, x_n + a_{i,n-1}\, x_{n-1} + \dots + a_{i,i+1}\, x_{i+1}\right)}{a_{ii}} \quad \dots \quad (16)$$

Realizing that Eq. 16 is identical with Eq. 10a or with Eq. 10b the following is formulated:

Any nonsingular matrix $[W]$ can bedivided by a nonsingular and triangular matrix $[A]$, if both matrixes have the same number of rows. The elements of the quotient matrix $[X]^T$ are determined one by one from the multiplication of rows of $[X]^T$ into rows of $[A]$ for an orderly generation of the terms of $[W]$.

$$\frac{[W]}{[A]} = [X]^T \quad \dots\dots\dots\dots\dots\dots \quad (17)$$

Shown in the expanded form, the corresponding equations are as follows:

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & w_{1m} \\ w_{21} & w_{22} & w_{23} & \cdots & w_{2m} \\ w_{31} & w_{32} & w_{33} & \cdots & w_{3m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & w_{n3} & \cdots & w_{nm} \end{bmatrix} \div \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & x_{31} & \cdots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2} \\ x_{13} & x_{23} & x_{33} & \cdots & x_{n3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{1m} & x_{2m} & x_{3m} & \cdots & x_{nm} \end{bmatrix}^T \qquad \text{(14a)}$$

(The arrows point to the systems of the type of Eq. 14 of which the system of Eq. 14a is a part.) And

$$\left.\begin{aligned} x_{nm} &= \frac{w_{nm}}{a_{nn}} \\ x_{n-1,m} &= \frac{w_{n-1,m} - a_{n-1,n}\, x_{nm}}{a_{n-1,n-1}} \\ x_{ij} &= \frac{w_{ij} - \left(a_{in}\, x_{ni} + a_{i,n-1}\, x_{n-1,i} + \cdots + a_{i,j+1}\, x_{j+1,i}\right)}{a_{ij}} \end{aligned}\right\} \qquad \text{(16a)}$$

If the matrix [A] is a lower triangular one, instead of upper triangular, as has been assumed, nothing will change in the reasoning but the order of computing x terms. The x_{11} will be taken first.

Eqs. 14a and 16a shall transform to 14b and 16b, as follows:

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & w_{1m} \\ w_{21} & w_{22} & w_{23} & \cdots & w_{2m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & w_{n3} & \cdots & w_{nm} \end{bmatrix} \div \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix}^T \quad \ldots (14b)$$

$$\left.\begin{aligned} x_{11} &= \frac{w_{11}}{a_{11}} \\ x_{21} &= \frac{w_{21} - a_{21}\, x_{11}}{a_{22}} \\ x_{ij} &= \frac{w_{ij} - \left(a_{i1}\, x_{1i} + a_{i2}\, x_{2i} + \cdots + a_{i,j-1}\, x_{j-1,i}\right)}{a_{ij}} \end{aligned}\right\} \quad \ldots\ldots (16b)$$

The provision concerning the nonsingularity of the dividend and the divisor matrixes comes from the rules of matrix multiplication. The same explanation applies to the requirement that the number of rows of the matrixes must be equal.

DIVISION BY A SQUARE MATRIX

Many authors have already established that any square matrix [A] can be considered as the product of two other matrixes [B] and [C], of which one must be a lower triangular matrix and the other an upper triangular matrix. Moreover one of the factor-matrixes should have the unit terms on the principal diagonal.[5] The unknown elements of the matrixes [B] and [C] are determined one by one by generating the elements of [A] in order.

---

5 "Computational Methods of Linear Algebra," by V. N. Faddeeva, Dover Publications, Inc., New York, N. Y., 1959.

In view of the preceding statement, a division by a square matrix may be substituted by two consecutive divisions by triangular matrixes, which are factors of the considered square matrix.

If

$$[A]\ [X] = [W] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

and

$$[A] = [B]\ [C] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (18)$$

then,

$$[X]^T = \frac{[W]}{[A]} = \frac{[W]}{[B]\,[C]} = \left(\frac{[W]}{[B]}\right) \div [C] \quad \ldots\ldots\ldots \quad (19)$$

Because matrix multiplication is not commutative, it is understood that matrix division is not either.

$$\left(\frac{[W]}{[B]}\right) \div [C] \neq \left(\frac{[W]}{[C]}\right) \div [B] \quad \ldots\ldots\ldots\ldots \quad (20)$$

Furthermore, because of the law for the product of three or more matrixes, "that the order of multiplication cannot be altered"[5], the order of division by the factors cannot be altered either.

$$\left.\begin{array}{l}[A]\ [X] = [W] \\ [A] = [B]\ [C]\end{array}\right\} \quad [B]\ [C]\ [X] = [W]$$

$$[C]\ [X] = \frac{[W]}{[B]}$$

$$[X] = \left(\frac{[W]}{[B]}\right) \div [C] \quad \ldots\ldots \quad (21)$$

Therefore, the following rule can be formulated:

Any nonsingular matrix [W] can be divided by a nonsingular and square matrix [A], if both matrixes have the same number of rows.
Matrix [A] is first factored into a product of two, one lower and one upper, triangular matrixes [B] and [C] (one of them with the unit terms on the principal diagonal). The division of matrix [W] by the factors [B] and [C] proceeds in order in conformity with the rule established previously for the division by triangular matrixes.

NUMERICAL EXAMPLE

Given

$$[A]\ \{X\} = \{W\}; \quad \begin{bmatrix} 4 & 4 & 8 \\ 1 & 2 & 0 \\ 2 & 6 & 16 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -5 \\ 16 \end{bmatrix}$$

find $\{x\}$ by the method of matrix division.

I. $[A] = [B]\ [C]$:

$$\begin{bmatrix} 4 & 4 & 8 \\ 1 & 2 & 0 \\ 2 & 6 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ b_{21} & 1 & 0 \\ b_{31} & b_{32} & 1 \end{bmatrix} \times \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{bmatrix}$$

$C_{11} = 4;\ C_{12} = 4;\ C_{13} = 8;$ $\quad b_{31} \times 4 + 0 + 0 = 2 \rightarrow b_{31} = 0.5$

$b_{21} \times 4 + 0 + 0 = 1 \rightarrow b_{21} = 0.25;$ $\quad 0.5 \times 4 + b_{32} \times 1 + 0 = 6 \rightarrow b_{32} = 4$

$0.25 \times 4 + C_{22} + 0 = 2 \rightarrow C_{22} = 1;$ $\quad 0.5 \times 8 - 4 \times 2 + C_{33} = 6 \rightarrow C_{33} = 20$

$0.25 \times 8 + C_{23} + 0 = 0 \rightarrow C_{23} = -2;$

and

$$[B] = \begin{bmatrix} 1 & 0 & 0 \\ 0.25 & 1 & 0 \\ 0.5 & 4 & 1 \end{bmatrix} \quad [C] = \begin{bmatrix} 4 & 4 & 8 \\ 0 & 1 & -2 \\ 0 & 0 & 20 \end{bmatrix}$$

II. $[C]\ \{X\} = \{X^1\} = \dfrac{[W]}{[B]}$

$$\begin{bmatrix} 8 \\ -5 \\ 16 \end{bmatrix} \div \begin{bmatrix} 1 & 0 & 0 \\ 0.25 & 1 & 0 \\ 0.5 & 4 & 1 \end{bmatrix}$$

$x_1^1 = \frac{8}{1} = 8$

$8 \times 0.25 + x_2^1 \times 1 + 0 = -5 \rightarrow x_2^1 = -7$

$8 \times 0.5 - 7 \times 4 + x_3^1 \times 1 = 16 \rightarrow x_3^1 = 40$

$$= \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 \end{bmatrix}^T$$

and

$$\{x^1\} = \begin{bmatrix} 8 \\ -7 \\ 40 \end{bmatrix}$$

III. $[x]^T = \left(\frac{[W]}{[B]}\right) \div [C] = \frac{\{x^1\}}{[C]}$;

$$\begin{bmatrix} 8 \\ -7 \\ 40 \end{bmatrix} \div \begin{bmatrix} 4 & 4 & 8 \\ 0 & 1 & -2 \\ 0 & 0 & 20 \end{bmatrix} \qquad \begin{array}{l} x_3 = \frac{40}{20} = 2; \\ -2 \times 2 + x_2 \times 1 + 0 = -7 \rightarrow x_2 = -3; \\ 2 \times 8 - 3 \times 4 + 4x_1 = 8 \rightarrow x_1 = 1; \end{array}$$

$$= [x_1 \quad x_2 \quad x_3]^T$$

and

$$\{x\} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix};$$

IV. Check: $[A] \{x\} = \{W\}$

$$\begin{bmatrix} 4 & 4 & 8 \\ 1 & 2 & 0 \\ 2 & 6 & 16 \end{bmatrix} \times \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4x1-4x3+8x2 = 8 \\ 1x1-3x2+0x2 = -5 \\ 1x2-3x6+2x16 = 16 \end{bmatrix} \equiv \begin{bmatrix} 8 \\ -5 \\ 16 \end{bmatrix}$$

CONCLUSIONS

The rules of matrix division formulated herein will be compared with the existing schemes, invented for the algebraical solution of Eq. 8. First consider the method of pre-multiplying both sides of the equation by the inverse of [A]. Thus,

$$[X] = [A]^{-1} [W] \quad \ldots\ldots\ldots\ldots\ldots \quad (22)$$

But the operation of inverting matrix [A] is in itself equal to the solution of a system of equations of an order described by the number of columns of [A].

A given case will be examined in which the number of columns of the matrix [A] is greater than the number of columns of the matrix [W]. In this circumstance, the auxiliary operation of the inversion is more involved than the primary problem of finding the elements of the matrix [X]. In such cases, instead of inverting first matrix [A], the methods of Doolittle,[6] Crout,[7] and so forth, or Cholesky[8] or the now formulated matrix division scheme will be used. The first two procedures (Doolittle and Crout) and other similar ones are all developed from Gauss' algorithmns.

Judging by the criterion of simplicity and automacity of operations, the writer considers these methods inferior to the rules of matrix division. The writer agrees that if the number of columns of matrix [A] is smaller than the number of columns of matrix [W], it is more convenient to invert [A] first than to find [x] by a direct matrix division. Cholesky's scheme, differs from the Gauss' elimination procedure in that it involves the determination of an auxiliary matrix [B], of the lower triangular type, that is capable of reducing the original Eq. 6 to

$$[C]\ \{X\} = \{V\} \qquad (23)$$

when [C] is a triangular matrix, by virtue of

$$[B]\ [C \;\vdots\; - V] = [A \;\vdots\; - W] \qquad (24)$$

in which [¦] denotes a partitioned matrix. The unknown terms of the matrixes [B] and [C ¦- V] are computed one by one, by generating the elements of the matrix [A ¦-W] in order. The values of the x_{ij} are obtained from the triangular system (Eq. 23) by "back substitution."

A similarity between operations of the Cholesky's scheme and the matrix division rule is immediately recognizable. Both methods have the steps of factorization and of the "back substitution." The total number of operations in both methods is approximately the same. But, when the matrix division rules are general ones, built on the basic canons of matrix algebra, the Cholesky's procedure is offered as a device for the solution of linear equations. Therefore, the writer classifies the Cholesky's scheme as a particular short-cut of the general matrix division method.

SUMMARY

A general method for matrix division has been developed and corresponding rules have been presented.

1. Any nonsingular matrix can be divided by a nonsingular and square matrix if both have the same number of rows.
2. The divisor matrix is first factorized in a product of two triangular matrixes.
3. The dividend matrix is divided consecutively, in order, by the factors of the divisor matrix.

---

6 Special Publication No. 28, U. S. Coast and Geodetic Survey, Washington, D. C., 1915.

7 "A Short Method for Evaluating Determinants," by P. D. Crout, Transactions, AIEE, Vol. 60, 1941.

8 "Numerical Methods in Engineering," by Mario G. Salvadori, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1956.

4. In dividing a matrix by a triangular matrix the elements of the quotient matrix are determined from the multiplication of the rows of the quotient matrix into the rows of the divisor matrix for an orderly generation of the terms of the dividend matrix.

ACKNOWLEDGMENTS

The writer wishes to express his indebtedness to E. Cohen, F. ASCE, and to F. R. Berman, for reading the paper and offering helpful suggestions.

KEY WORDS: engineering mechanics; matrixes; mathematics; simultaneous equations

ABSTRACT: Rules are developed for a direct division of matrixes without previous inversion of any factor of the computation. Any nonsingular matrix may be divided by a nonsingular and square matrix, if both matrixes have the same number of rows. The divisor matrix is first factored in a product of two triangular matrixes. The dividend matrix is divided consecutively by the factors of the divisor matrix. In dividing a matrix by a triangular matrix, the elements of the quotient matrix are determined from the multiplication of the rows of the quotient matrix into rows of the divisor matrix, for an orderly generation of the terms of the dividend matrix. One numerical example illustrates the procedure. This division method represents a general form of which Cholesky's scheme is a particular short-cut varient.

REFERENCE: "Matrix Division," by G. Kostro, Journal of the Engineering Mechanics Division, ASCE, Vol. 89, No. EM3, Proc. Paper 3538, June, 1963, pp. 9-20.