Three methods for solving systems of linear equations:
Comparing the advantages and disadvantages

Haoyu Luo1,a,*, †, Simeng Wu2, b, *,†, Ningyun Xie3, c, *,†
1Cheshire Academy, Cheshire, Connecticut, 06410, China;
2School of Dulwich International High School Suzhou, Suzhou, 215000, China;
3School of Guangdong Country Garden School, Foshan, 528500, China;
*Corresponding author Email:
Haoyu.Luo@cheshireacademy.org, Grace.Wu22@dchigh-suzhou.cn,
guanghua.ren@gecacademy.cn
†These authors contributed equally.

Abstract. This paper aims to introduce some prevalent techniques which have been used to solve linear systems. Firstly, we introduce the simplest method: how to eliminate variables in which the original systems would not be changed. We then introduce Gaussian eliminations, which work on the augmented matrix derived from a linear system. To the end, we present Cramer’s rule, which computes the solution to a linear system based on matrix determinants. These methods have their application scenarios. For instance, eliminating variables has no limited condition for its operations. However, Cramer's rule needs to be under the condition of square matrices. At the same time, Gaussian elimination requires three elementary row operations, and Gaussian elimination paves the way for computing the rank of matrices. The choice between these methods depends on coefficient matrices in terms of dimensions and matrix properties such as singularity.

1. Introduction
How to find the solutions for simultaneous linear equations? Linear algebra is the branch of mathematics concerned with linear equations, which have been studied for a very long time. The whole map for linear algebra did not progress until the 17th century: matrices, an array of numbers, appeared [1]. It is surprising that as linear algebra rapidly developed, the study of coefficients in linear equations inspired mathematicians to discover determinants. In the late 17th century, Leibnitz contributed to linear algebra by relating his work to square matrices with a unique solution.

Moreover, in 1750 AD, Lagrange adopted matrices to his work, called "characterize the maxima and minima multivariate functions." However, mathematicians had not stopped their research. Cramer's rule is one of the essential parts of utilizing the determinants to find the solution for simultaneous linear equations after the results of Leibnitz. Whereas, this rule concluded a potential problem which is that n×n linear systems were not able to be proved. This drew to Euler's idea that the systems may not have a solution, so he found out the conditions required: square matrices and unique solution.

In this context, in the 19th century, Gauss made significant progress: he solved the systems without matrices. Moreover, Gaussian Elimination is a famous and ingenious method for solving linear systems since Gauss used adding, multiplying, and swapping and put much effort into unknown variables [2].
As there are many ways to solve linear equations, mathematicians may wonder about efficiency, the advantages, and the disadvantages of each method. Reviewing the research done by other experts, we choose to focus on three main techniques: Eliminating variables, Cramer's rule, and Gaussian Elimination.

2. Elimination of variables

The first introduced method in this paper aims to indicate one important approach to solve linear problems: eliminating variables. The most basic of this part is the mutual conversion of matrix and linear systems. A table of $m$ rows and $n$ columns arranged by a series of $m \times n$ numbers $a_{ij}$ ($i = 1, 2, 3, \ldots, m; ~ j = 1, 2, 3, \ldots, n$) is called an $m$ rows and $n$ columns matrix, an $m \times n$ matrix for short. For a matrix denoted as $(a_{ij})_{m \times n}$ or $(a_{ij})$, a capital letter is used to represent it. And for a linear system, it has something in common with matrices. This is a linear system with $m$ equations and $n$ variables, both $m$ and $n$ unknown:

\[
\begin{align*}
\begin{cases}
    a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= k_1 \\
    a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= k_2 \\
    a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= k_3 \\
    \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= k_m
\end{cases}
\]

(1)

So, using a matrix to represent this linear system, which can be written as $A \cdot x = k$ and $A$ is the coefficient matrix. The matrix composed of the coefficients and constant terms of the system is called an augmented matrix. [3]

\[
A = \begin{bmatrix}
    a_{11} & \cdots & a_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{m1} & \cdots & a_{mn}
\end{bmatrix}, ~ x = \begin{bmatrix}
    x_1 \\
    \vdots \\
    x_n
\end{bmatrix}, ~ k = \begin{bmatrix}
    k_1 \\
    \vdots \\
    k_m
\end{bmatrix}
\]  

(2)

When the transformation of matrices and linear systems are clear, their operational correspondence is also the same. The elementary operation is the heart of the elimination of variables.[4]

In this case, the following three operations are considered:

i. Interchange the order of the equations, denoted as $E_i \leftrightarrow E_j$

ii. Replace an equation by multiplying it by a nonzero number, denoted as $(cE_i) \rightarrow E_i$

iii. Replace an equation by adding $c$ times another equation, denoted as $(cE_i + E_j) \rightarrow E_j$

All three of these transformations are reversible, so the system of equations before and after the transformation are of the same solution. Moreover, in the above transformations, only the coefficients and constants of the equations are calculated, so the operations of a linear system can be completely converted into a row transformation of an augmented matrix, also known as the row elementary operations:[5]

1. Interchanging the order of two rows: $R_i \leftrightarrow R_j$
2. Multiplying any row of the matrix by a nonzero number: $R_i \rightarrow kR_i$
3. Adding a multiple of a row to another row: $R_i \rightarrow R_i + cR_j$

Similarly, as long as replacing row with the column, operations on matrix columns can correspond one-to-one with row elementary operations and are called elementary column operations.

a) Examples

\[
\begin{align*}
2x_1 - x_2 + 2x_3 &= 4 \quad (1) \\
x_1 + x_2 + 2x_3 &= 1 \quad (2) \\
4x_3 + x_2 + 4x_3 &= 2 \quad (3)
\end{align*}
\]

First, transferring the simultaneous linear equations into augmented matrix form, denoted as the capital letter B.

\[
B = \begin{bmatrix}
    2 & -1 & 2 & 1 \\
    1 & 1 & 2 & 4 \\
    4 & 1 & 4 & 2
\end{bmatrix}
\]
Then, interchanging the order of (1) and (2):

\[
\begin{bmatrix}
1 & 1 & 2 & 1 \\
2 & -1 & 2 & 4 \\
4 & 1 & 4 & 2
\end{bmatrix}
\]

Next, in order to ease the calculations, let (2)-2(1), (3)-4(1):

\[
\begin{bmatrix}
1 & 1 & 2 & 1 \\
0 & -3 & -2 & 2 \\
0 & -3 & 4 & -2
\end{bmatrix}
\]

And then, calculating (3)-(2):

\[
\begin{bmatrix}
1 & 1 & 2 & 1 \\
0 & -3 & -2 & 2 \\
0 & 0 & -2 & 4
\end{bmatrix}
\]

In the end, using back-substitution to get the final solutions:

1) \((3) \times \frac{1}{2}:
\]

\[
\begin{bmatrix}
1 & 1 & 2 & 1 \\
0 & -3 & -2 & 2 \\
0 & 0 & 1 & 2
\end{bmatrix}
\]

2) \((1-2) \times -\frac{1}{3} + 2(2):
\]

\[
\begin{bmatrix}
1 & 0 & 0 & -1 \\
0 & 1 & 0 & -2 \\
0 & 0 & 1 & 2
\end{bmatrix}
\]

So, now is the simplest row echelon form, and the solutions of this system are:

\[
\begin{align*}
x_1 &= -1, \\
x_2 &= -2, \\
x_3 &= 2
\end{align*}
\]

3. Gaussian Elimination

Gaussian elimination or we can call it the row reduction, is an algorithm for solving systems of linear equations. It is made of a sequence of operations performed on the corresponding matrix of coefficients. To perform row reduction on a matrix, one uses a sequence of elementary row operations to modify the matrix until the lower left-hand corner of the matrix is filled with zeros, as much as possible:

According to repeat those three elementary row operations that won’t affect the solution but able to simplify the augmented matrix, we can finally get the matrix like:

\[
\begin{bmatrix}
a_{11}x_1 & a_{12}x_2 & \ldots & a_{1r}x_r & a_{1r+1}x_{r+1} & \ldots & a_{1m}x_m & k_1 \\
0 & a_{22}x_2 & \ldots & a_{2r}x_r & a_{2r+1}x_{r+1} & \ldots & a_{2m}x_m & k_2 \\
0 & 0 & \ldots & a_{rr}x_r & a_{rr+1}x_{r+1} & \ldots & a_{rm}x_m & k_3 \\
0 & 0 & 0 & \ldots & a_{1m}x_m & \ldots & k_m
\end{bmatrix}
\]

We call the left part of the separate line “Row echelon matrix” [6], which has the character: the column mark of the first element in the non-zero row \(A_j\) must be smaller than or equal to the column mark of the first element in the non-zero row \(A_{j+1}\).

After we get the row echelon matrix, the amounts of variables are decreasing from top to bottom. Thus, using the back substitution method: getting the solution of the bottom first, then repeating plug in the known value of variables the above row until the top. Therefore, we will get the solution.
3.1 The basic usage (Introduced example) of Gaussian Elimination

Assume that we have three variables: x, y, z. And we have the equation set of them (the highest power here is one):

\[
\begin{align*}
1x + 4y + 7z &= 10 \quad (r1) \\
2x + 5y + 8z &= 11 \quad (r2), \\
3x + 6y + 9z &= 12 \quad (r3)
\end{align*}
\]

Traditionally, we can solve it by:

\[
\begin{align*}
\frac{1}{2} r_2 - r_1: \quad \left( \frac{2}{5} - 4 \right) y + (4 - 7) z &= \frac{11}{2} - 10, \text{ which } \frac{-18}{5} y - 3z = \frac{-9}{2} \quad (r4); \\
\frac{1}{3} r_3 - r_1: \quad (2 - 4) y + (3 - 7) z &= 4 - 10, \text{ which } -y - 2z = -3 \quad (r5); \\
\frac{18}{5} r_5 - r_4: \quad \left( -\frac{36}{5} - (-3) \right) z &= -\frac{18}{5} x_3 - \left( -\frac{9}{2} \right), \text{ which } \frac{-21}{5} z = -\frac{63}{10}, z = \frac{3}{2};
\end{align*}
\]

Thus, we plug the value of z into the r5 to get the value of y;

Plug the value of z and y into the r1/r2/r3 to get the value of x.

Therefore, we will have the unique solution

\[
\begin{align*}
x &= -\frac{1}{2} \\
y &= 0 \\
z &= \frac{3}{2}
\end{align*}
\]

Or we can use the Gaussian Elimination.

First, we can extract all the coefficients and got the “coefficient matrix”:

\[
\begin{bmatrix}
1 & 4 & 7 \\
2 & 5 & 8 \\
3 & 6 & 9
\end{bmatrix}
\]

then extract all the constants on the right side of the equation, and use the line to separate them. Thus, we got:

\[
\begin{bmatrix}
1 & 4 & 7 & |10 \\
2 & 5 & 8 & |11 \\
3 & 6 & 9 & |12
\end{bmatrix}
\]

In this form, we call it the “Augmented matrix”.

Then by repeating the three elementary row operations:

1. (1st rule) Row A2 replace by \( \frac{1}{2} A2 \),
2. (3rd rule) Row A2 replace by \(-1A1+A2\)
3. (2nd rule) Swap the position of Row A3 and A2

We will get:

\[
\begin{bmatrix}
1 & 4 & 7 & |10 \\
0 & -\frac{18}{5} & -3 & |\frac{-9}{2} \\
0 & 0 & -\frac{21}{5} & |\frac{63}{10}
\end{bmatrix}
\]

Which we can get \( z = \frac{3}{2} \) from the last row, and by making back substitution, we can get \( y = 0 \) and \( x = -\frac{1}{2} \).

Compared to the traditional way, obviously, the Gaussian Elimination shows the process of solving more clearly and directly.

3.2 General and special

3.2.1 The general process:

According to the shown example above, the solving mode can be concluded as below.
By extracting the related condition from the question, we can have an expression:

\[
\begin{align*}
\sum_{i=1}^{m} a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \cdots + a_{im}x_m &= k_i \\
\sum_{i=1}^{m} a_{i2}x_2 + a_{i3}x_3 + \cdots + a_{im}x_m &= k_i \\
\sum_{i=1}^{m} a_{i3}x_3 + \cdots + a_{im}x_m &= k_i \\
\vdots \\
\sum_{i=1}^{m} a_{im}x_1 + a_{2m}x_2 + a_{3m}x_3 + \cdots + a_{nm}x_n &= k_m
\end{align*}
\]

Thus, transfer to the augmented matrix form:

\[
\begin{bmatrix}
a_{11}x_1 & a_{12}x_2 & a_{13}x_3 & \cdots & a_{1m}x_m & k_1 \\
a_{21}x_1 & a_{22}x_2 & a_{23}x_3 & \cdots & a_{2m}x_m & k_2 \\
a_{31}x_1 & a_{32}x_2 & a_{33}x_3 & \cdots & a_{3m}x_m & k_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1}x_1 & a_{m2}x_2 & a_{m3}x_3 & \cdots & a_{nm}x_m & k_m
\end{bmatrix}
\]

Therefore, we can analogy the steps in the example, to get the row echelon matrix:

\[
\begin{bmatrix}
b_{11}x_1 & b_{12}x_2 \cdots b_{1r}x_r & b_{1r+1}x_{r+1} \cdots b_{1m}x_m & t_1 \\
0 & b_{22}x_2 \cdots b_{2r}x_r & b_{2r+1}x_{r+1} \cdots b_{2m}x_m & t_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 \cdots b_{rr}x_r & b_{r+1}x_{r+1} \cdots b_{mm}x_m & t_m
\end{bmatrix}
\]

Then use the \( x_n = \frac{t_m}{b_{1m}} \) to make the back substitution for all the row above and get the solution.

3.2.2 Rank:

“In linear algebra, the rank of a matrix is the highest order of its non-zero subform.” [6] Claim that we use \( R(A) \) to represent the rank of the matrix \( A \). During the process of Gaussian Elimination in our Research Question, the rank of the coefficient matrix \( u \) equal to the rank that in the form of row echelon matrix. Credit to the character of row echelon, \( R(u) = m \); if there is another matrix \( a \) with \( n \) variables, getting a row echelon form after the process of Gaussian Elimination, \( R(a) \) will equal to the row mark of the highest \( x_n \) with the coefficient 0.

Rank can help us to define the range/feature of the solution of the linear equations: let the coefficient matrix of the linear equation system be \( A \), and the augmented matrix is \( B = (A, b) \), then we will have:

i. \( R(A) = R(B) = n, (n \in \mathbb{N}^+) \), the system of equations has a unique solution;  
ii. \( R(A) = R(B) < n, (n \in \mathbb{N}^+) \), the equations have infinite solutions;  
iii. \( R(A) = R(B) > n, (n \in \mathbb{N}^+) \), the equation has no solution.

4. Cramer’s rule

Cramer’s rule provides another method other than using augmented matrices.

4.1 The basic knowledge of Cramer’s rule(conditions)

Cramer’s rule is another method of solving linear system problems, which are transformed into matrix form. Determinants play an important role in dealing with such simultaneous linear equations. Usually, the determinant of matrix \( A \) is denoted \( \text{det}(A) \) or \( |\text{matrix } A| \), which only calculated of square matrices. When computing a determinant, the matrix always has a unique inverse. This can be proved by assuming the matrix \( A \) has two different inverse-matrix \( B \) and \( C \). \( AB = AC = I(\text{identity matrix}) \). However, after multiplying \( B \) and using the associative law of matrix multiplication, \( (BA)B = (BA)C \). This results in \( B = C \), which contradicts the precondition. The determinant of this matrix could be any real values. Because of this, more and more mathematicians like Kronecker remark the properties of determinants:[7]  
(1) If one row or column in the square matrix, \( \text{det}(A) \) will be zero.
(2) If one row or column in a square matrix is proportional to another row or column, \( \det(A) \) will always be zero.

In geometric ways, the essence of determinants is that they give the factor by which area changes. To simplify, using an example of a transforming triangle \(
\begin{pmatrix}
1 & 5 \\
1 & 4
\end{pmatrix}
\) by matrix \(
\begin{pmatrix}
3 & 0 \\
0 & 1
\end{pmatrix}
\) helps understand the concept:

\[
\begin{pmatrix}
3 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 5 \\
1 & 4
\end{pmatrix} = \begin{pmatrix}
3 & 15 \\
1 & 15
\end{pmatrix}
\]

where the original area is 6 unit\(^2\) and the transformed area is 18 unit\(^2\).

In this case, the new area is 3 times the original, where 3 is the determinant of the transformation matrix.

By adopting the essence, the Leibniz formula computes the standard forms of measuring determinants.

\[
\begin{vmatrix}
a & b \\
c & d
\end{vmatrix} = ad - bc \quad (7)
\]

\[
\begin{vmatrix}
a & b & c \\
d & e & f \\
g & h & i
\end{vmatrix} = cde - caf - bdi + bfg + iae - fbd \quad (8)
\]

With the given formulae, we present the whole process step by step. In setting the examples, we just mention the nonzero cases.

Use 2×2 matrix as an example:

\[
\begin{vmatrix}
3 & 4 \\
6 & 7
\end{vmatrix} = 3 \times 7 - 4 \times 6 = -3
\]

Use 3×3 matrix as an example:

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{vmatrix} = 1 \times 5 \times 9 - 3 \times 4 \times 7 - 2 \times 5 \times 6 = -72
\]

What is known from these examples is that determinants can be negative. Indeed, Cramer’s rule utilizes the determinants.

4.2 The general process of Cramer’s rule

Cramer’s rule includes a few steps. First, transforming the linear equations into argument matrix form is an excellent way to understand the concept of this rule since this transformation eliminates some recurring unknowns. The argument matrix, which is a better choice to illustrate the relationship between different variables, is a better choice to further improve.

Let \( T \) be an \( n \times n \) matrix, defined as matrix \( T \). \( \det(T) \) is the determinant of this original matrix \( T \). In this case, \( X \) is assumed as the unique solution of linear system \( TX = N \). \( N \) is the constant vector (\( N \in \mathbb{R} \)). Then, replacing one \( T \) column with \( N \), \( \det(T_1) \) is the determinant of this new matrix. Repeating this replacement with all the columns, then, we apply this formula:

\[
X_n = \frac{\det(T_n)}{\det(T)} \quad \text{for} \ n = 1, 2, 3, \ldots, n \quad (9)
\]

The denominator is the determinant of the original matrix. And, the numerator is the determinant of the matrix whose one column has been replaced by the constant vector(\( X \)).

4.3 Examples of applying Cramer’s rule

\[
\begin{align*}
x + y + 2z &= 25 \\
x + 4y - 3z &= 8 \\
2x + y - z &= 9
\end{align*}
\]

Matrix \( T = \begin{pmatrix}
1 & 1 & 2 & 25 \\
1 & 4 & -3 & 8 \\
2 & 1 & -1 & 9
\end{pmatrix}
\]
Check the results by inputting these to the original linear equations:

(1) \(5+6+14=25\)

(2) \(5+24-21=8\)

(3) \(10+6-7=9\)

5. Comparison between methods

When the elimination method is used to solve linear systems, it can be found that there are no applied conditions for its operation. So, which means that the elimination method can crack any linear system as long as the transformation of an augmented matrix to a linear system is correctly applied. So that is why this is one fundamental way for us to get started with linear algebra.

However, since elementary operations are based on row to row’s, column to column’s, or equation to equation’s transformations, this method will require us to carry out many operations when we encounter more significant equations or matrices. Eventually, it will become a very cumbersome and complex process. Therefore, when solving more complex linear systems, there are other more straightforward methods than using the basic elimination rules, which is Gaussian Elimination.

Compare to the basic eliminations. Gaussian Elimination provides a solid mode to solve the problem. Without considering all seven fundamental matrix transferring methods, we can use Gaussian Elimination to only use three of them, including scalar multiplication of a row, row swap, and scalar multiplication and addition. Therefore, Gaussian Elimination is an effective method with low complexity. Much less computation is required for more significant problems, conveniently get the rank of the given matrix. While the drawback for it is not relatively easy to remember the procedure for handwork, it takes a long time without a computer.

For the last method we mentioned above, Cramer's rule has the advantages: Determinants can be quickly calculated using the formula of different types of matrices. Also, they have another property: if a column or row is multiplied by \(k\) to result in matrix \(T'\), then \(\text{det}(T')=k\text{det}(T)\), which is more efficient when figuring out the pattern of the matrices. Cramer's rule shows its benefits of solving a linear system, in which under the condition of inhomogeneous linear systems with \(n\) variables, \(n\) equations, and \(\text{rank}=n\). However, inhomogeneous linear systems whose numbers inside the constant vector are not all zeros, so the non-zero constant vector is used in previous examples. In addition, every unknown variable can be worked out independently. Unlike the Elimination of variables, Cramer's rule does not need to plug in one variable to find other variables.

Moreover, Cramer's rule is relatively convenient when dealing with small matrices, like \((2\times2, 3\times3, \text{maybe } 4\times4)\). However, contrary to the previous statement, as we compute the whole process ourselves, it is evident that when the matrix can be extensive, Cramer's rule works more slowly. Besides, the example given above is artificially created that determinant is non-zero. If the determinant is zero, Cramer's rule will not work since zero is on the denominator. This kind of calculation will lead to the uncertainty of whether it is infinite solutions or no solutions.

Thus, compared to two extensive ways, in addition to being faster for larger matrices, the Gaussian elimination method has the advantage of telling more information than Cramer's rule. For example, it
tells the rank of the matrix, while Cramer's rule only tells whether the matrix is invertible. The Gaussian elimination method is also applicable to non-square matrices (i.e. equations with unequal numbers of variables and equations), but Cramer's rule does not work. In addition, if there are multiple solutions (i.e. the number of solutions is unlimited), then the Gaussian elimination method will provide you with all the solutions. So instead, Cramer's rules cannot work well.

6. Conclusion
When faced with a math problem, finding the answer is always the goal, but solving the answer can be varied. The same is true of linear algebra. This paper identified three solutions for different linear systems: The elimination method, Gaussian Elimination, and Cramer's rule.

For each method, we provide detailed calculations, practical applications, and advantages and disadvantages. In summary, the elimination method has a wide range of applications and no use conditions. Still, its operation is cumbersome, and it is not suitable for complex and large linear systems. Other than eliminating variables, the Gaussian elimination method requires less computation for larger linear systems and is more convenient to obtain the rank of a given matrix. Nevertheless, without a computer, doing calculations by hand is time-consuming and hard to remember. In addition, Cramer's rule is a good choice for inhomogeneous linear systems because each unknown variable is computed independently. However, when it comes to large matrices, it gets complicated. Cramer's rule is based on the condition that the determinant is non-zero; otherwise, it leads to the uncertainty of whether there is an infinite solution or no solution.

Each method has its characteristics and shortcomings, so how to choose the best method to solve different linear algebra is what we want to appeal to. For different linear systems, it is necessary to apply various methods flexibly and finally successfully solve the problem. In terms of these problems, combining and unifying with other multi-disciplines and making these methods widely used in computer programming models is also a major development trend in the future. It's not just about solving problems manually, it's about computer science playing an increasingly important role in the future of algorithm optimization, and that's what we want to see.

References
[1] AYDIN S. (2017). On the history of some Linear Algebra Concepts: From Babylon to Pre-Technology. The Online Journal of Science and Technology-January, 7(1).
[2] Christensen, J., Gustafson, G. (2012). A Brief History of Linear Algebra. Grant Gustafson Univ. Utah.
[3] Tucker, A. (1993). The growing importance of linear algebra in undergraduate mathematics. The college mathematics journal, 24(1), 3-9.
[4] Leon, S. J., Bica, I., & Hohn, T. (1998). Linear algebra with applications (Vol. 6). Upper Saddle River, NJ: Prentice Hall.
[5] Olver, P. J., Shakiban, C., & Shakiban, C. (2006). Applied linear algebra (Vol. 1). Upper Saddle River, NJ: Prentice Hall.
[6] Szabo PhD, 2015. In The Linear Algebra Survival Guide.
[7] Lancaster, P., & Tismenetsky, M. (1985). The theory of matrices: with applications. Elsevier.