The History of Matrices and Modern Applications

Nova G. Casillo

SUNY Oswego, Class of 2018

Abstract

This paper discusses early knowledge and development of linear systems of equations, matrices, and their determinants, as well as a few modern applications including the Second Partials Test and Markov Chains.

Keywords: linear systems of equations, matrices, determinant, application

The History of Matrices and Modern Applications

Systems of linear equations are pivotal to society, for ancient civilizations and modern civilizations alike. Matrices that can be used to solve such systems are of incalculable wealth to society since they allow fast solving of incredibly large systems.

Ancient Linear Systems

The Babylonian civilization shows the first evidence of solving linear systems, specifically linear systems of two equations with two unknowns. The Babylonians solved such systems by systematic elimination. The first step was to replace the variables, say x and y, in the first equation, with say x_1 and y_1. Then one would let $x_1 = y_1$ to solve the first equation. This answer would then be substituted into the second equation with an added error, that is let $x = x_1 + d$ and $y = y_1 - d$, since it was known that x_1 and y_1 were not the exact answers. This allows x to be a little over the average and y a little below the average. The second equation could then be solved for d and the values for the unknowns found by substituting d into the equations for x and y. A tablet from approximately the third century BCE contained the problem:

> *There are two fields whose total area is* 1800 *square yards. One produces grain at the rate of* $\frac{2}{3}$ *of a bushel per square yard while the other produces grain at the rate of* $\frac{1}{2}$ *a bushel per square yard. If the total yield is* 1100 *bushels, what is the size of each field.*
>
> (O'Connor, J., & Robertson, E., 1996).

This can be translated into modern notation, where x represents the first field and y represents the second, as:

$$\text{Eq. 1: } x + y = 1800 \text{ and Eq. 2: } \frac{2}{3}x + \frac{1}{2}y = 1100.$$

Suppose $x = x_1, y = y_1,$ $\quad x_1 + y_1 = 1800$ and let $x_1 = y_1$

$$\Rightarrow 2x_1 = 1800.$$

Now we can solve the first equation by dividing by 2:

$$\frac{2x_1}{2} = \frac{1800}{2} \quad \Rightarrow \quad x_1 = 900.$$

Since here $x_1 = y_1$, $y_1 = 900$. If we let $x = x_1 + d$ and $y = y_1 - d$, then $x = 900 + d$ and $y = 900 - d$. Substituting into the second equation gives:

$$\frac{2}{3}(900 + d\,) + \frac{1}{2}(900 - d) = 1100$$

So, we find that d =300. Which implies:

$$x = 900 + 300 = 1200 \quad \text{and} \quad y = 900 - 300 = 600$$

So the area of the field that produces grain at a rate of $^2/_3$ of a bushel per square yard is 1200 square yards and the field that produces grain at a rate of $^1/_2$ of a bushel per square yard is 600 square yards.

Modern techniques to solve systems of two linear equations with two unknowns are similar to the Babylonian method. The modern technique is referred to as ‘elimination’ or more formally, Gaussian Elimination, named after German mathematician, Carl Friedrich Gauss. The method involves solving the first equation for one unknown and substituting into the second equation. This leaves only one equation with one unknown to be solved. Solving the above problem by modern elimination:

$$x + y = 1800 \text{ and } \tfrac{2}{3}x + \tfrac{1}{2}y = 1100.$$

Then, $x = 1800 - y.$

By substitution we see, $\frac{2}{3}(1800 - y) + \frac{1}{2}y = 1100.$

After simplifying we are left with, $y = 600.$

Thus, plugging back into $x = 1800 - y$ we get: $x = 1800 - 600 = 1200.$

Therefore, the area of the first field is 1200 square yards and the area of the second is 600 square yards.

It is clear that modern elimination method is faster than the Babylonian method, but it is still evident that the Babylonians were very clever in discovering a method to solve systems at all.

The ancient Chinese discovered a method that could be used on any system, unware that it could be used on any system, they only used it to solve systems of up to six equations and six unknowns. In *Nine Chapters on the Mathematical Art,* a method to solve such systems is described using a matrix, although the term matrix was not coined until much later (O'Connor, J., & Robertson, E., 1996). The matrices and the method used by the ancient Chinese were identical to modern practices aside from the orientation; coefficients were placed in columns rather than rows as in modern notation. One such problem, quite similar to the Babylonian example is:

> *There are three types of corn, of which three bundles of the first, two of the second, and one of the third make* 39 *measures. Two of the first, three of the second and one of the third make* 34*measures. And one of the first, two of the second and three of the third make* 26 *measures. How many measures of corn are contained of one bundle of each type?* (O'Connor, J., & Robertson, E.)

Let x be the number of bundles of the first type of corn, y be the second, and z be the third. Then we can symbolically express this system as:

$$3x + 2y + z = 39, \quad 2x + 3y + z = 34, \quad x + 2y + 3z = 26$$

The astonishing thing is that the text, *Nine Chapters on the Mathematical Art,* does not work with these equations in this form. Rather the coefficients are placed on a “counting board” with an augmented row that contains the coefficients on the right hand side of the equations:

$$\begin{matrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{matrix}$$

From here the author instructs the reader to multiply the middle column by 3 and then subtract the right column from the middle column as many times as possible. Since the middle column when multiplied by 3 is $\begin{matrix} 6 \\ 9 \\ 3 \\ 102 \end{matrix}$ and when $\begin{matrix} 3 \\ 2 \\ 1 \\ 39 \end{matrix}$ is subtracted from this twice the result is $\begin{matrix} 0 \\ 5 \\ 1 \\ 24 \end{matrix}$. So, this gives:

$$\begin{matrix} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{matrix}$$

The reader is then instructed to multiply the first column by 3 and subtract the right column from it as many times as possible, resulting in:

$$\begin{matrix} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{matrix}$$

Finally, the reader is told to multiply the left column by 5 and subtract the middle column from it as many times as possible, giving:

$$\begin{matrix} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{matrix}$$

From here one can find the third type of corn, z, since the left column gives $0x + 0y + 36z = 99$ which implies $z = 2.75$. Then, when substituted back into the middle column equation the second type of corn, y, is found, as follows:

$0x + 5y + 1(2.75) = 24$ which implies $y = 4.25$. To find the remaining (first) type of corn, substitute into the right column equation:

$3x + 2(4.25) + 2.75 = 39$ which implies $x = 9.25$.

Thus, there are 9.25 bundles of the first type of corn, 4.25 of the second, and 2.75 of the third.

This revolutionary method developed by the ancient Chinese was not discovered elsewhere until early the 1800s, which was about 2000 years later.

Determinants

The matrices used to solve linear systems can give more information about the nature of systems of linear equations than the numerical values of the unknowns. The determinant of a matrix can be used to conclude whether or not there is a unique solution to a system, that is if there is only one possible set of values for a set of unknowns in a given linear system, if the system is linearly independent, and so on. Determinants also make the required calculations much faster.

The first major appearances of determinants occurred in Japan and Europe at about the same time. Seki Takakazu was a Japanese mathematician credited as being the first person to study determinants. In 1683 he wrote the *Method of Solving Dissimulated Problems*, which introduced a general method for finding the determinant of a matrix and used them to solve equations, but not entire systems (O'Connor, J., & Robertson, E.). The same year, 1683, Gottfried Leibniz, a German mathematician penned a letter to French mathematician Guillaume De L'Hôpital. Leibniz argued that the following system

$$10 + 11x + 12y = 0, \qquad 20 + 21x + 22y = 0, \qquad 30 + 31x + 32y = 0$$

has a solution since

$$10 \cdot 21 \cdot 32 + 11 \cdot 22 \cdot 30 + 12 \cdot 20 \cdot 31 = 10 \cdot 22 \cdot 31 + 11 \cdot 20 \cdot 32 + 12 \cdot 21 \cdot 30.$$

We would recognize this today as the condition that states that a homogenous system with determinant 0 will have infinitely many solutions.

This depicts the mere beginnings of determinants. Many other mathematicians have now studied and contributed to the knowledge base of determinants.

Modern Applications

In calculus a Hessian matrix, or more specifically, the determinant of it, can be used to find extrema of a function, by use of the Second Partials Test. A Hessian matrix is a matrix of the second partial derivatives of a function evaluated at a point, usually evaluated at a critical point. Suppose $f(x, y)$ is differentiable in the neighborhood of a critical point (x_0, y_0). Then the corresponding determinant of the Hessian matrix is:

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx} \cdot f_{yy} - (f_{xy})^2, \qquad \text{evaluated at } (x_0, y_0).$$

This value, represented by D (say), is used to determine if the point (x_0, y_0), already known to be a critical point, is a relative extrema or saddle point. If the value, D, is greater than 0, then the point is a point of relative extremum. Whether it is a point of maximum or minimum is determined by the sign of f_{xx} or f_{yy}. If $f_{xx}, f_{yy} > 0$ then the point is a relative minimum, and if $f_{xx}, f_{yy} < 0$ then the point is a relative maximum. If $D < 0$, then $f(x, y)$ has a saddle point at (x_0, y_0). If $D = 0$, then unfortunately the test is inconclusive. This has applications in many fields, including meteorology and physics.

In probability and various other fields, stochastic matrices are used to describe the progression of Markov Chains. Each entry in a stochastic matrix is a real number on the closed interval from 0 to 1 that represents the probability of an event. The entries must also sum to 1 in each column, each row, or both if the matrix is doubly stochastic. An example of a right stochastic matrix is seen in the matrix representation of the following scenario: a row of five boxes, with a cat in the first and a mouse in the fifth, in which the cat and mouse jump to an adjacent box for each state. This scenario ends when they are in the same box and the cat catches the mouse. The probability after the first advance that the mouse will be in the fifth box and the cat in the second is 1 since those are the only adjacent boxes to jump to. The pattern continues in this fashion considering which boxes it is possible for them to jump into, e.g. if the cat is in box two and the mouse in box four, they each have a ½ chance of jumping into box three, so the overall probability is ¼. This gives the stochastic matrix:

$$\begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

which can be used to visually show the transition probability of the cat catching the mouse. The cat will always catch the mouse, there is probability 1 that the cat will jump in the same box as the mouse.

The work of many mathematicians and the development of matrices have allowed modern society to solve and learn about large systems of equations, which appear in a multitude of fields. Indeed, matrices have become a critical tool in representing, manipulating and solving linear systems that occur in such fields as partial differential equations, graph theory, computer graphics and many more. They will continue to enrich our study of applied mathematics for a long time to come!

References

Debnath, L. (2014). A brief historical introduction to matrices and their applications. *International Journal of Mathematical Education in Science and Technology, 45*(3), 360-377. doi:10.1080/0020739X.2013.837521

O'Connor, J., & Robertson, E. (1996, February 1). Matrices and determinants. Retrieved November 23, 2015, from http://www-groups.dcs.st-andrews.ac.uk/~history/HistTopics/Matrices_and_determinants.html

Stochastic matrix. (2015, September 6). Retrieved November 23, 2015, from https://en.wikipedia.org/wiki/Stochastic_matrix