**History of Matrices and Determinants**
Based upon an article by J J O'Connor and E F Robertson[[1]](#footnote-1)
<http://mathshistory.st-andrews.ac.uk/HistTopics/Matrices_and_determinants.html>

**Guiding Questions**

* When and where did linear algebra begin? By way of comparison, when/where was the calculus discovered/invented?
* Which section titles in the textbook do (and do not) appear in this article?
* What mathematicians play a prominent role (their name is mentioned at least three times)?
* What percent of the article is about determinants? What percent of our course is about determinants? Why do you think there is such a difference?
* What is the difference between proofs by Cayley, Hamilton, Kronecker, and Weierstrass and what we are learning in as exemplified by Frobenius?
* What is one insight you have after reading this history?

**Solving systems**

The beginnings of matrices and determinants goes back to the second century BC although traces can be seen back to the fourth century BC. However it was not until near the end of the 17th Century that the ideas reappeared and development really got underway.

It is not surprising that the beginnings of matrices and determinants should arise through the study of systems of linear equations. The Babylonians studied problems that led to simultaneous linear equations and some of these are preserved in clay tablets that survive. For example, a tablet dating from around 300 BC contains the following problem:

*There are two fields whose total area is*1800*square yards. One produces grain at the rate of*2*/*3*of a bushel per square yard while the other produces grain at the rate of*1*/*2*a bushel per square yard. If the total yield is*1100*bushels, what is the size of each field.*

The Chinese, between 200 BC and 100 BC, came much closer to matrices than the Babylonians. Indeed it is fair to say that the text Nine Chapters on the Mathematical Art written during the Han Dynasty gives the first known example of matrix methods. First a problem is set up which is similar to the Babylonian example given above:

*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?*

Now the author does something quite remarkable. He sets up the coefficients of the system of three linear equations in three unknowns as a table on a 'counting board'.

 1 2 3

 2 3 2

 3 1 1

 26 34 39

Our late 20th Century methods would have us write the linear equations as the rows of the matrix rather than the columns but the method is identical. Most remarkably the author, writing in 200 BC, instructs the reader to multiply the middle column by 3 and subtract the right column *as many times as possible*, the same is then done subtracting the right column *as many times as possible* from 3 times the first column. This gives

 0 0 3

 4 5 2

 8 1 1

 39 24 39

Next the left most column is multiplied by 5 and then the middle column is subtracted *as many times as possible*. This gives the column echelon form:

 0 0 3

 0 5 2

 36 1 1

 99 24 39

from which the solution can be found for the third type of corn, then for the second, then the first by back substitution. This method, now known as Gauss-Jordan Elimination, would not become well known until the early 19th Century.

Gaussian elimination, which first appeared in the text [Nine Chapters on the Mathematical Art](http://mathshistory.st-andrews.ac.uk/HistTopics/Nine_chapters.html) written in 200 BC, was used by [Gauss](http://mathshistory.st-andrews.ac.uk/Mathematicians/Gauss.html) in his work which studied the orbit of the asteroid Pallas. Using observations of Pallas taken between 1803 and 1809, [Gauss](http://mathshistory.st-andrews.ac.uk/Mathematicians/Gauss.html) obtained a system of six linear equations in six unknowns. [Gauss](http://mathshistory.st-andrews.ac.uk/Mathematicians/Gauss.html) gave a systematic method for solving such equations which is precisely Gauss-Jordan elimination on the coefficient matrix.

**“Modern” matrix notation and vocabulary**

[Eisenstein](http://mathshistory.st-andrews.ac.uk/Mathematicians/Eisenstein.html) in 1844 denoted linear substitutions by a single letter and showed how to add and multiply them like ordinary numbers except for the lack of commutativity. It is fair to say that [Eisenstein](http://mathshistory.st-andrews.ac.uk/Mathematicians/Eisenstein.html) was the first to think of linear substitutions as forming an algebra as can be seen in this quote from his 1844 paper:

*An algorithm for calculation can be based on this, it consists of applying the usual rules for the operations of multiplication, division, and exponentiation to symbolic equations between linear systems, correct symbolic equations are always obtained, the sole consideration being that the order of the factors may not be altered.*

The first to use the term 'matrix' was [Sylvester](http://mathshistory.st-andrews.ac.uk/Mathematicians/Sylvester.html) in 1850. [Sylvester](http://mathshistory.st-andrews.ac.uk/Mathematicians/Sylvester.html) defined a matrix to be *an oblong arrangement of terms* and saw it as something which led to various determinants from square arrays contained within it. After leaving America and returning to England in 1851, [Sylvester](http://mathshistory.st-andrews.ac.uk/Mathematicians/Sylvester.html) became a lawyer and met [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html), a fellow lawyer who shared his interest in mathematics. [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) quickly saw the significance of the matrix concept and by 1853 [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) had published a note giving, for the first time, the inverse of a matrix.

[Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) in 1858 published *Memoir on the theory of matrices* which is remarkable for containing the first abstract definition of a matrix. He shows that the coefficient arrays studied earlier for quadratic forms and for linear transformations are special cases of his general concept. [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) gave a matrix algebra defining addition, multiplication, scalar multiplication and inverses. He gave an explicit construction of the inverse of a matrix in terms of the determinant of the matrix.

An axiomatic definition of a determinant was used by [Weierstrass](http://mathshistory.st-andrews.ac.uk/Mathematicians/Weierstrass.html) in his lectures and, after his death, it was published in 1903 in the note *On determinant theory*. In the same year [Kronecker](http://mathshistory.st-andrews.ac.uk/Mathematicians/Kronecker.html)'s lectures on determinants were also published, again after his death. With these two publications the modern theory of determinants was in place but matrix theory took slightly longer to become a fully accepted theory. An important early text which brought matrices into their proper place within mathematics was *Introduction to higher algebra* by [Bôcher](http://mathshistory.st-andrews.ac.uk/Mathematicians/Bocher.html) in 1907. [Turnbull](http://mathshistory.st-andrews.ac.uk/Mathematicians/Turnbull.html) and [Aitken](http://mathshistory.st-andrews.ac.uk/Mathematicians/Aitken.html) wrote influential texts in the 1930's and [Mirsky](http://mathshistory.st-andrews.ac.uk/Mathematicians/Mirsky.html)'s *An introduction to linear algebra* in 1955 saw matrix theory reach its present major role in as one of the most important undergraduate mathematics topic.

**Eigenvalues and diagonalization**:

Many standard results of elementary matrix theory first appeared long before matrices were the object of mathematical investigation. For example [de Witt](http://mathshistory.st-andrews.ac.uk/Mathematicians/De_Witt.html) in *Elements of curves*, published as a part of the commentaries on the 1660 Latin version of [Descartes](http://mathshistory.st-andrews.ac.uk/Mathematicians/Descartes.html)' *Géométrie ,* showed how a transformation of the axes reduces a given equation for a conic to canonical form. This amounts to diagonalizing a symmetric matrix but [de Witt](http://mathshistory.st-andrews.ac.uk/Mathematicians/De_Witt.html) never thought in these terms. [This is covered in Chapter 5].

In 1826 [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html), in the context of quadratic forms in *n* variables, used the term 'tableau' for the matrix of coefficients. He found the eigenvalues and gave results on diagonalization of a matrix in the context of converting a form to the sum of squares. [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html) also introduced the idea of similar matrices (but not the term) and showed that if two matrices are similar they have the same characteristic equation. He also, again in the context of quadratic forms, proved that every real symmetric matrix is diagonalisable.

Jacques [Sturm](http://mathshistory.st-andrews.ac.uk/Mathematicians/Sturm.html) gave a generalisation of the eigenvalue problem in the context of solving systems of ordinary differential equations [Math 230]. In fact the concept of an eigenvalue appeared 80 years earlier, again in work on systems of linear differential equations, by [D'Alembert](http://mathshistory.st-andrews.ac.uk/Mathematicians/DAlembert.html) studying the motion of a string with masses attached to it at various points.

**Linear Transformations**

It should be stressed that neither [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html) nor Jacques [Sturm](http://mathshistory.st-andrews.ac.uk/Mathematicians/Sturm.html) realized the generality of the ideas they were introducing and saw them only in the specific contexts in which they were working. [Jacobi](http://mathshistory.st-andrews.ac.uk/Mathematicians/Jacobi.html) from around 1830 and then [Kronecker](http://mathshistory.st-andrews.ac.uk/Mathematicians/Kronecker.html) and [Weierstrass](http://mathshistory.st-andrews.ac.uk/Mathematicians/Weierstrass.html) in the 1850's and 1860's also looked at matrix results but again in a special context, this time the notion of a linear transformation. [Jacobi](http://mathshistory.st-andrews.ac.uk/Mathematicians/Jacobi.html) published three treatises on determinants in 1841. These were important in that for the first time the definition of the determinant was made in an algorithmic way and the entries in the determinant were not specified so his results applied equally well to cases were the entries were numbers or to where they were functions. These three papers by [Jacobi](http://mathshistory.st-andrews.ac.uk/Mathematicians/Jacobi.html) made the idea of a determinant widely known.

**Proofs**

[Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) also proved that, in the case of 2 × 2 matrices, that a matrix satisfies its own characteristic equation. He stated that he had checked the result for 3 × 3 matrices, indicating its proof, but says:-

*I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree.*

That a matrix satisfies its own characteristic equation is called the [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html)-[Hamilton](http://mathshistory.st-andrews.ac.uk/Mathematicians/Hamilton.html) theorem so it’s reasonable to ask what it has to do with [Hamilton](http://mathshistory.st-andrews.ac.uk/Mathematicians/Hamilton.html). In fact he also proved a special case of the theorem, the 4 × 4 case, in the course of his investigations into quaternions.

[Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html), in 1878, wrote an important work on matrices *On linear substitutions and bilinear forms* although he seemed unaware of [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html)'s work. [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) in this paper deals with coefficients of forms and does not use the term matrix. However he proved important results on canonical matrices as representatives of equivalence classes of matrices. He cites [Kronecker](http://mathshistory.st-andrews.ac.uk/Mathematicians/Kronecker.html) and [Weierstrass](http://mathshistory.st-andrews.ac.uk/Mathematicians/Weierstrass.html) as having considered special cases of his results in 1874 and 1868 respectively. [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) also proved the general result that a matrix satisfies its characteristic equation. This 1878 paper by [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) also contains the definition of the rank of a matrix which he used in his work on canonical forms and the definition of orthogonal matrices.

In 1896 [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) became aware of [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html)'s 1858 *Memoir on the theory of matrices* and after this started to use the term matrix. Despite the fact that [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) only proved the [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html)-[Hamilton](http://mathshistory.st-andrews.ac.uk/Mathematicians/Hamilton.html) theorem for 2 × 2 and 3 × 3 matrices, [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) generously attributed the result to [Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html) despite the fact that [Frobenius](http://mathshistory.st-andrews.ac.uk/Mathematicians/Frobenius.html) had been the first to prove the general theorem.

**Determinants**

[Cardan](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cardan.html), in *Ars Magna* (1545), gives a rule for solving a system of two linear equations which he calls *regula de modo* and which one author called the *mother of rules*! This rule gives what essentially is [Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html)'s rule for solving a 2 × 2 system although [Cardan](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cardan.html) does not make the final step. [Cardan](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cardan.html) therefore does not reach the definition of a determinant but, with the advantage of hindsight, we can see that his method does lead to the definition. [We will not cover Cramer’s rule in this course, but do discuss Determinants in Chapter 3].

The idea of a determinant appeared in Japan before it appeared in Europe. In 1683 [Seki](http://mathshistory.st-andrews.ac.uk/Mathematicians/Seki.html) wrote *Method of solving the dissimulated problems* which contains matrix methods written as tables in exactly the way the Chinese methods described above were constructed. Without having any word which corresponds to 'determinant' [Seki](http://mathshistory.st-andrews.ac.uk/Mathematicians/Seki.html) still introduced determinants and gave general methods for calculating them based on examples. Using his 'determinants' [Seki](http://mathshistory.st-andrews.ac.uk/Mathematicians/Seki.html) was able to find determinants of 2 × 2, 3 × 3, 4 × 4 and 5 × 5 matrices and applied them to solving equations but not systems of linear equations. [Determinants are in Chapter 3].

The first appearance of a determinant in Europe was ten years later. In 1693 [Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) wrote to [de l'Hôpital](http://mathshistory.st-andrews.ac.uk/Mathematicians/De_LHopital.html). He explained that the system of equations

10 + 11*x* + 12*y* = 0
20 + 21*x* + 22*y* = 0
30 + 31*x* + 32*y* = 0

had a solution because

10.21.32 + 11.22.30 + 12.20.31 = 10.22.31 + 11.20.32 + 12.21.30

which is exactly the condition that the coefficient matrix has determinant 0. Notice that here [Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) is not using numerical coefficients but

*two characters, the first marking in which equation it occurs, the second marking which letter it belongs to.*

Hence 21 denotes what we might write as *a*21.

[Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) was convinced that good mathematical notation was the key to progress so he experimented with different notation for coefficient systems. His unpublished manuscripts contain more than 50 different ways of writing coefficient systems which he worked on during a period of 50 years beginning in 1678. Only two publications (1700 and 1710) contain results on coefficient systems and these use the same notation as in his letter to [de l'Hôpital](http://mathshistory.st-andrews.ac.uk/Mathematicians/De_LHopital.html) mentioned above.

[Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) used the word 'resultant' for certain combinatorial sums of terms of a determinant. He proved various results on resultants including what is essentially [Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html)'s rule. He also knew that a determinant could be expanded using any column - what is now called the [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) expansion. As well as studying coefficient systems of equations which led him to determinants, [Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) also studied coefficient systems of quadratic forms which led naturally towards matrix theory.

In the 1730's [Maclaurin](http://mathshistory.st-andrews.ac.uk/Mathematicians/Maclaurin.html) wrote *Treatise of algebra* although it was not published until 1748, two years after his death. It contains the first published results on determinants proving [Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html)'s rule for 2 × 2 and 3 × 3 systems and indicating how the 4 × 4 case would work. [Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html) gave the general rule for *n* × *n* systems in a paper *Introduction to the analysis of algebraic curves* (1750). It arose out of a desire to find the equation of a plane curve passing through a number of given points. The rule appears in an Appendix to the paper but no proof is given:-

*One finds the value of each unknown by forming n fractions of which the common denominator has as many terms as there are permutations of n things.*

[Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html) does go on to explain precisely how one calculates these terms as products of certain coefficients in the equations and how one determines the sign. He also says how the *n* numerators of the fractions can be found by replacing certain coefficients in this calculation by constant terms of the system.

Work on determinants now began to appear regularly. In 1764 [Bezout](http://mathshistory.st-andrews.ac.uk/Mathematicians/Bezout.html) gave methods of calculating determinants as did [Vandermonde](http://mathshistory.st-andrews.ac.uk/Mathematicians/Vandermonde.html) in 1771. In 1772 [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) claimed that the methods introduced by [Cramer](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cramer.html) and [Bezout](http://mathshistory.st-andrews.ac.uk/Mathematicians/Bezout.html) were impractical and, in a paper where he studied the orbits of the inner planets, he discussed the solution of systems of linear equations without actually calculating it, by using determinants. Rather surprisingly [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) used the word 'resultant' for what we now call the determinant: surprising since it is the same word as used by [Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html) yet [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) must have been unaware of [Leibniz](http://mathshistory.st-andrews.ac.uk/Mathematicians/Leibniz.html)'s work. [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) gave the expansion of a determinant which is now named after him.

[Lagrange](http://mathshistory.st-andrews.ac.uk/Mathematicians/Lagrange.html), in a paper of 1773, studied identities for 3 × 3 functional determinants. However this comment is made with hindsight since [Lagrange](http://mathshistory.st-andrews.ac.uk/Mathematicians/Lagrange.html) himself saw no connection between his work and that of [Laplace](http://mathshistory.st-andrews.ac.uk/Mathematicians/Laplace.html) and [Vandermonde](http://mathshistory.st-andrews.ac.uk/Mathematicians/Vandermonde.html). This 1773 paper on mechanics, however, contains what we now think of as the volume interpretation of a determinant for the first time. [Lagrange](http://mathshistory.st-andrews.ac.uk/Mathematicians/Lagrange.html) showed that the tetrahedron formed by O(0,0,0) and the three points *M*(*x*,*y*,*z*), *M*'(*x*',*y*',*z*'), *M*"(*x*",*y*",*z*") has volume

1/6 [*z*(*x*'*y*" - *y*'*x*") + *z*'(*yx*" - *xy*") + *z*"(*xy*' - *yx*')].

The term 'determinant' was first introduced by [Gauss](http://mathshistory.st-andrews.ac.uk/Mathematicians/Gauss.html) in *Disquisitiones arithmeticae* (1801) while discussing quadratic forms. He used the term because the determinant determines the properties of the quadratic form. However the concept is not the same as that of our determinant. In the same work [Gauss](http://mathshistory.st-andrews.ac.uk/Mathematicians/Gauss.html) lays out the coefficients of his quadratic forms in rectangular arrays. He describes matrix multiplication (which he thinks of as composition so he has not yet reached the concept of matrix algebra) and the inverse of a matrix in the particular context of the arrays of coefficients of quadratic forms.

It was [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html) in 1812 who used 'determinant' in its modern sense. [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html)'s work is the most complete of the early works on determinants. He reproved the earlier results and gave new results of his own on minors and adjoints. In the 1812 paper the multiplication theorem for determinants is proved for the first time although, at the same meeting of the Institut de France, [Binet](http://mathshistory.st-andrews.ac.uk/Mathematicians/Binet.html) also read a paper which contained a proof of the multiplication theorem but it was less satisfactory than that given by [Cauchy](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cauchy.html).

[Cayley](http://mathshistory.st-andrews.ac.uk/Mathematicians/Cayley.html), also writing in 1841, published the first English contribution to the theory of determinants. In this paper he used two vertical lines on either side of the array to denote the determinant, a notation which has now become standard.

1. I have taken the liberty of reordering the original article so that it is more thematic (rather than purely chronological). I have also cut out and reordered some sections. The purpose of all these edits was to make the article a bit more approachable for a reader new to linear algebra. [↑](#footnote-ref-1)