1

Historical Elements of Matrices and Tensors

COPYRIGHTED MATERIAL

The objective of this introduction is by no means to outline a rigorous and comprehensive historical background of the theory of matrices and tensors. Such a historical record should be the work of a historian of mathematics and would require thorough bibliographical research, including reading the original publications of authors cited to analyze and reconstruct the progress of mathematical thinking throughout years and collaborations. A very interesting illustration of this type of analysis is provided, for example, in the form of a "representation of research networks"[1], over the period 1880–1907, in which are identified the interactions and influences of some mathematicians, such as James Joseph Sylvester (1814–1897), Karl Theodor Weierstrass (1815–1897), Arthur Cayley (1821–1895), Leopold Kronecker (1823–1891), Ferdinand Georg Frobenius (1849–1917), or Eduard Weyr (1852–1903), with respect to the theory of matrices, the theory of numbers (quaternions, hypercomplex numbers), bilinear forms, and algebraic structures.

Our modest goal here is to locate in time the contributions of a few mathematicians and physicists[2] who have laid the foundations for the theory of matrices and tensors, and to whom we will refer later in our presentation. This choice is necessarily very incomplete.

1 F. Brechenmacher, "Les matrices : formes de représentation et pratiques opératoires (1850–1930)", Culture MATH - Expert site ENS Ulm / DESCO, December 20, 2006.

2 For more information on the mathematicians cited in this introduction, refer to the document "Biographies de mathématiciens célèbres", by Johan Mathieu, 2008, and the remarkable site Mac Tutor History of Mathematics Archive (http://www-history.mcs.st-andrews.ac.uk) of the University of St. Andrews, in Scotland, which contains a very large number of biographies of mathematicians.

The first studies of determinants that preceded those of matrices were conducted independently by the Japanese mathematician Seki Kowa (1642–1708) and the German mathematician Gottfried Leibniz (1646–1716), and then by the Scottish mathematician Colin Maclaurin (1698–1746) for solving 2×2 and 3×3 systems of linear equations. These works were then generalized by the Swiss mathematician Gabriel Cramer (1704–1752) for the resolution of $n \times n$ systems, leading, in 1750, to the famous formulae that bear his name, whose demonstration is due to Augustin-Louis Cauchy (1789–1857).

In 1772, Théophile Alexandre Vandermonde (1735–1796) defined the notion of determinant, and Pierre-Simon Laplace (1749–1827) formulated the computation of determinants by means of an expansion according to a row or a column, an expansion which will be presented in section 4.11.1. In 1773, Joseph-Louis Lagrange (1736–1813) discovered the link between the calculation of determinants and that of volumes. In 1812, Cauchy used, for the first time, the determinant in the sense that it has today, and he established the formula for the determinant of the product of two rectangular matrices, a formula which was found independently by Jacques Binet (1786–1856), and which is called nowadays the Binet–Cauchy formula.

In 1810, Johann Carl Friedrich Gauss (1777–1855) introduced a notation using a table, similar to matrix notation, to write a 3×3 system of linear equations, and he proposed the elimination method, known as Gauss elimination through pivoting, to solve it. This method, also known as Gauss–Jordan elimination method, was in fact known to Chinese mathematicians (first century). It was presented in a modern form, by Gauss, when he developed the least squares method, first published by Adrien-Marie Legendre (1752–1833), in 1805.

Several determinants of special matrices are designated by the names of their authors, such as Vandermonde's, Cauchy's, Hilbert's, and Sylvester's determinants. The latter of whom used the word "matrix" for the first time in 1850, to designate a rectangular table of numbers. The presentation of the determinant of an nth-order square matrix as an alternating n-linear function of its n column vectors is due to Weierstrass and Kronecker, at the end of the 19th century.

The foundations of the theory of matrices were laid in the 19th century around the following topics: determinants for solving systems of linear equations, representation of linear transformations and quadratic forms (a topic which will be addressed in detail in Chapter 4), matrix decompositions and reductions to canonical forms, that is to say, diagonal or triangular forms such as the Jordan (1838–1922) normal form with Jordan blocks on the diagonal, introduced by Weierstrass, the block-triangular form of Schur (1875–1941), or the Frobenius normal form that is a block-diagonal matrix, whose blocks are companion matrices.

A history of the theory of matrices in the 19th century was published by Thomas Hawkins[3] in 1974, highlighting, in particular, the contributions of the British mathematician Arthur Cayley, seen by historians as one of the founders of the theory of matrices. Cayley laid the foundations of the classical theory of determinants[4] in 1843. He then developed matrix computation[5] by defining certain matrix operations as the product of two matrices, the transposition of the product of two matrices, and the inversion of a 3×3 matrix using cofactors, and by establishing different properties of matrices, including, namely, the famous Cayley–Hamilton theorem which states that every square matrix satisfies its characteristic equation. This result highlighted for the fourth order by William Rowan Hamilton (1805–1865), in 1853, for the calculation of the inverse of a quaternion, was stated in the general case by Cayley in 1857, but the demonstration for any arbitrary order is due to Frobenius, in 1878.

An important part of the theory of matrices concerns the spectral theory, namely, the notions of eigenvalue and characteristic polynomial. Directly related to the integration of systems of linear differential equations, this theory has its origins in physics, and more particularly in celestial mechanics for the study of the orbits of planets, conducted in the 18th century by mathematicians, physicists, and astronomers such as Lagrange and Laplace, then in the 19th century by Cauchy, Weierstrass, Kronecker, and Jordan.

The names of certain matrices and associated determinants are those of the mathematicians who have introduced them. This is the case, for example, for Alexandre Théophile Vandermonde (1735–1796) who gave his name to a matrix whose elements on each row (or each column) form a geometric progression and whose determinant is a polynomial. It is also the case for Carl Jacobi (1804–1851) and Ludwig Otto Hesse (1811–1874), for Jacobian and Hessian matrices, namely, the matrices of first- and second-order partial derivatives of a vector function, whose determinants are called Jacobian and Hessian, respectively. The same is true for the Laplacian matrix or Laplace matrix, which is used to represent a graph. We can also mention Charles Hermite (1822–1901) for Hermitian matrices, related to the so-called Hermitian forms (see section 4.15). Specific matrices such as Fourier (1768–1830) and Hadamard (1865–1963) matrices are directly related to the transforms of the same name. Similarly, Householder (1904–1993) and Givens (1910–1993) matrices are associated with transformations corresponding to reflections and rotations, respectively. The so-called structured matrices, such as Hankel (1839–1873) and Toeplitz (1881–1943) matrices, play a very important role in signal processing.

3 Thomas Hawkins, "The theory of matrices in the 19th century", *Proceedings of the International Congress of Mathematicians, Vancouver*, 1974.

4 Arthur Cayley, "On a theory of determinants", *Cambridge Philosophical Society* 8, l–16, 1843.

5 Arthur Cayley, "A memoir on the theory of matrices", *Philosophical Transactions of the Royal Society of London* 148, 17–37, 1858.

Matrix decompositions are widely used in numerical analysis, especially to solve systems of equations using the method of least squares. This is the case, for example, of EVD (eigenvalue decomposition), SVD (singular value decomposition), LU, QR, UD, Cholesky (1875–1918), and Schur (1875–1941) decompositions, which will be presented in Volume 2.

Just as matrices and matrix computation play a fundamental role in linear algebra, tensors and tensor computation are at the origin of multilinear algebra. It was in the 19th century that tensor analysis first appeared, along with the works of German mathematicians Georg Friedrich Bernhard Riemann[6] (1826–1866) and Elwin Bruno Christoffel (1829–1900) in geometry (non-Euclidean), introducing the index notation and notions of metric, manifold, geodesic, curved space, curvature tensor, which gave rise to what is today called Riemannian geometry and differential geometry.

It was the Italian mathematician Gregorio Ricci-Curbastro (1853–1925) with his student Tullio Levi-Civita (1873–1941) who were the founders of the tensor calculus, then called absolute differential calculus[7], with the introduction of the notion of covariant and contravariant components, which was used by Albert Einstein (1879–1955) in his theory of general relativity, in 1915.

Tensor calculus originates from the study of the invariance of quadratic forms under the effect of a change of coordinates and, more generally, from the theory of invariants initiated by Cayley[8], with the introduction of the notion of hyperdeterminant which generalizes matrix determinants to hypermatrices. Refer to the article by Crilly[9] for an overview of the contribution of Cayley on the invariant theory. This theory was developed by Jordan and Kronecker and involved controversy[10] between these two authors, then continued by David Hilbert (1862–1943), Elie Joseph Cartan (1869–1951), and Hermann Klaus Hugo Weyl (1885–1955), for algebraic forms

6 A detailed analysis of Riemann's contributions to tensor analysis has been made by Ruth Farwell and Christopher Knee, "The missing link: Riemann's Commentatio, differential geometry and tensor analysis", *Historia Mathematica* 17, 223–255, 1990.

7 G. Ricci and T. Levi-Civita, "Méthodes de calcul différentiel absolu et leurs applications", *Mathematische Annalen* 54, 125–201, 1900.

8 A. Cayley, "On the theory of linear transformations", *Cambridge Journal of Mathematics* 4, 193–209, 1845. A. Cayley, "On linear transformations", *Cambridge and Dublin Mathematical Journal* 1, 104–122, 1846.

9 T. Crilly, "The rise of Cayley's invariant theory (1841–1862)", *Historica Mathematica* 13, 241–254, 1986.

10 F. Brechenmacher, "La controverse de 1874 entre Camille Jordan et Leopold Kronecker: Histoire du théorème de Jordan de la décomposition matricielle (1870–1930)", *Revue d'histoire des Mathématiques, Society Math De France* 2, no. 13, 187–257, 2008 (hal-00142790v2).

(or homogeneous polynomials), or for symmetric tensors[11]. A historical review of the theory of invariants was made by Dieudonné and Carrell[12].

This property of invariance vis-à-vis the coordinate system characterizes the laws of physics and, thus, mathematical models of physics. This explains that tensor calculus is one of the fundamental mathematical tools for writing and studying equations that govern physical phenomena. This is the case, for example, in general relativity, in continuum mechanics, for the theory of elastic deformations, in electromagnetism, thermodynamics, and so on.

The word tensor was introduced by the German physicist Woldemar Voigt (1850–1919), in 1899, for the geometric representation of tensions (or pressures) and deformations in a body, in the areas of elasticity and crystallography. Note that the word tensor was introduced independently by the Irish mathematician, physicist and astronomer William Rowan Hamilton (1805–1865), in 1846, to designate the modulus of a quaternion[13].

As we have just seen in this brief historical overview, tensor calculus was used initially in geometry and to describe physical phenomena using tensor fields, facilitating the application of differential operators (gradient, divergence, rotational, and Laplacian) to tensor fields[14].

Thus, we define the electromagnetic tensor (or Maxwell's (1831–1879) tensor) describing the structure of the electromagnetic field, the Cauchy stress tensor, and the deformation tensor (or Green–Lagrange deformation tensor), in continuum mechanics, and the fourth-order curvature tensor (or Riemann–Christoffel tensor) and the third-order torsion tensor (or Cartan tensor[15]) in differential geometry.

---

11 M. Olive, B. Kolev, and N. Auffray, "Espace de tenseurs et théorie classique des invariants", *21ème Congrès Français de Mécanique*, Bordeaux, France, 2013 (hal-00827406).

12 J. A. Dieudonné and J. B. Carrell, *Invariant Theory, Old and New*, Academic Press, 1971.

13 See page 9 in E. Sarrau, *Notions sur la théorie des quaternions*, Gauthiers-Villars, Paris, 1889, http://rcin.org.pl/Content/13490.

14 The notion of tensor field is associated with physical quantities that may depend on both spatial coordinates and time. These variable geometric quantities define differentiable functions on a domain of the physical space. Tensor fields are used in differential geometry, in algebraic geometry, general relativity, and in many other areas of mathematics and physics. The concept of tensor field generalizes that of vector field.

15 E. Cartan, "Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion", *Comptes rendus de l'Académie des Sciences* 174, 593–595, 1922. Elie Joseph Cartan (1869–1951), French mathematician and student of Jules Henri Poincaré (1854–1912) and Charles Hermite (1822–1901) at the Ecole Normale Supérieure. He brought major contributions concerning the theory of Lie groups, differential geometry, Riemannian geometry, orthogonal polynomials, and elliptic functions. He discovered spinors, in 1913, as part of his work on the

After their introduction as computational and representation tools in physics and geometry, tensors have been the subject of mathematical developments related to polyadic decomposition (Hitchcock 1927) aiming to generalize dyadic decompositions, that is to say, matrix decompositions such as SVD.

Then emerged their applications as tools for the analysis of three-dimensional data generalizing matrix analysis to sets of matrices, viewed as arrays of data characterized by three indices. We can mention here the works of pioneers in factor analysis by Cattell[16] and Tucker[17] in psychometrics (Cattell 1944; Tucker 1966), and Harshman[18] in phonetics (Harshman 1970) who have introduced Tucker's and PARAFAC (parallel factors) decompositions. This last one was proposed independently by Carroll and Chang (1970), under the name of canonical decomposition (CANDECOMP), following the publication of an article by Wold (1966), with the objective to generalize the (Eckart and Young 1936) decomposition, that is, SVD, to arrays of order higher than two. This decomposition was then called CP (for CANDECOMP/PARAFAC) by Kiers (2000). For an overview of tensor methods applied to data analysis, the reader should consult the books by Coppi and Bolasco (1989) and Kroonenberg (2008).

From the early 1990s, tensor analysis, also called multi-way analysis, has also been widely used in chemistry, and more specifically in chemometrics (Bro 1997). Refer to, for example, the book by Smilde *et al.* (2004) for a description of various applications in chemistry.

In parallel, at the end of the 1980s, statistic "objects," such as moments and cumulants of order higher than two, have naturally emerged as tensors (McCullagh 1987). Tensor-based applications were then developed in signal processing for solving the problem of blind source separation using cumulants (Cardoso 1990, 1991; Cardoso and Comon 1990). The book by Comon and Jutten (2010) outlines an overview of methods for blind source separation.

In the early 2000s, tensors were used for modeling digital communication systems (Sidiropoulos *et al.* 2000a), for array processing (Sidiropoulos *et al.* 2000b), for multi-dimensional harmonics recovery (Haardt *et al.* 2008; Jiang *et al.* 2001; Sidiropoulos 2001), and for image processing, more specifically for face recognition

representations of groups. Like tensor calculus, spinor calculus plays a major role in quantum physics. His name is associated with Albert Einstein (1879–1955) for the classical theory of gravitation that relies on the model of general relativity.

16 Raymond Cattell (1905–1998), Anglo-American psychologist who used factorial analysis for the study of personality with applications to psychotherapy.

17 Ledyard Tucker (1910–2004), American mathematician, expert in statistics and psychology, and more particularly known for tensor decomposition which bears his name.

18 Richard Harshman (1943–2008), an expert in psychometrics and father of three-dimensional PARAFAC analysis which is the most widely used tensor decomposition in applications.

(Vasilescu and Terzopoulos 2002). The field of wireless communication systems has then given rise to a large number of tensor models (da Costa *et al.* 2018; de Almeida and Favier 2013; de Almeida *et al.* 2008; Favier *et al.* 2012a; Favier and de Almeida 2014b; Favier *et al.* 2016). These models will be covered in a chapter of Volume 3. Tensors have also been used for modeling and parameter estimation of dynamic systems both linear (Fernandes *et al.* 2008, 2009a) and nonlinear, such as Volterra systems (Favier and Bouilloc 2009a, 2009b, 2010) or Wiener-Hammerstein systems (Favier and Kibangou 2009a, 2009b; Favier *et al.* 2012b; Kibangou and Favier 2008, 2009, 2010), and for modeling and estimating nonlinear communication channels (Bouilloc and Favier 2012; Fernandes *et al.* 2009b, 2011; Kibangou and Favier 2007). These different tensor-based applications in signal processing will be addressed in Volume 4.

Many applications of tensors also concern speech processing (Nion *et al.* 2010), MIMO radar (Nion and Sidiropoulos 2010), and biomedical signal processing, particularly for electroencephalography (EEG) (Cong *et al.* 2015; de Vos *et al.* 2007; Hunyadi *et al.* 2016), and electrocardiography (ECG) signals (Padhy *et al.* 2018), magnetic resonance imaging (MRI) (Schultz *et al.* 2014), or hyperspectral imaging (Bourennane *et al.* 2010; Velasco-Forero and Angulo 2013), among many others. Today, tensors viewed as multi-index tables are used in many areas of application for the representation, mining, analysis, and fusion of multi-dimensional and multi-modal data (Acar and Yener 2009; Cichocki 2013; Lahat *et al.* 2015; Morup 2011).

A very large number of books address linear algebra and matrix calculus, for example: Gantmacher (1959), Greub (1967), Bellman (1970), Strang (1980), Horn and Johnson (1985, 1991), Lancaster and Tismenetsky (1985), Noble and Daniel (1988), Barnett (1990), Rotella and Borne (1995), Golub and Van Loan (1996), Lütkepohl (1996), Cullen (1997), Zhang (1999), Meyer (2000), Lascaux and Théodor (2000), Serre (2002), Abadir and Magnus (2005), Bernstein (2005), Gourdon (2009), Grifone (2011), and Aubry (2012).

For multilinear algebra and tensor calculus, there are much less reference books, for example: Greub (1978), McCullagh (1987), Coppi and Bolasco (1989), Smilde *et al.* (2004), Kroonenberg (2008), Cichocki *et al.* (2009), and Hackbusch (2012). For an introduction to multilinear algebra and tensors, see Ph.D. theses by de Lathauwer (1997) and Bro (1998). The following synthesis articles can also be consulted: (Bro 1997; Cichocki *et al.* 2015; Comon 2014; Favier and de Almeida 2014a; Kolda and Bader 2009; Lu *et al.* 2011; Papalexakis *et al.* 2016; Sidiropoulos *et al.* 2017).