CAMBRIDGE

Cambridge University Press
978-1-108-47909-7 — Matrix, Numerical, and Optimization Methods in Science and Engineering
Kevin W. Cassel
Excerpt
More Information

1 Vector and Matrix Algebra

Mathematics is the music of reason.

The world of ideas which [mathematics] discloses or illuminates, the contemplation of divine beauty and order which it induces, the harmonious connexion of its parts, the infinite hierarchy and absolute evidence of the truths with which it is concerned, these, and such like, are the surest grounds of the title of mathematics to human regard, and would remain unimpeached and unimpaired were the plan of the universe unrolled like a map at our feet, and the mind of man qualified to take in the whole scheme of creation at a glance. (James Joseph Sylvester)

Approximately 50 years before publication of this book, Neil Armstrong was the first human to step foot on the Moon and – thankfully – return safely to Earth. While the Apollo 11 mission that facilitated this amazing accomplishment in July of 1969 was largely symbolic from a scientific perspective, it represented the culmination of a decade of focused effort to create, develop, test, and implement a whole host of new technologies that were necessary to make such a feat possible. Many of these technologies are now fully integrated into modern life.

The march to the Moon during the 1960s was unprecedented in modern technological history for its boldness, aggressive time frame, scale, scientific and technological developments, and ultimate success. The state of the art in space travel in the early 1960s was to put a human in orbit 280 kilometers (km) (175 miles) above the Earth for several hours. It was in this context on September 12, 1962, that President Kennedy boldly – and some argued foolishly – set the United States on a course "to go to the Moon in this decade and do the other things, not because they are easy, but because they are hard." The Moon is 385,000 km (240,000 miles) from Earth – orbital flights were a mere baby step toward this aggressive goal. It was not that the technology was available and we simply needed to marshal the financial resources and political will to accomplish this bold task, it was not clear *if* the technology could be developed at all, let alone in the aggressive time frame proposed.

This problem appeared straightforward on paper. Newton's laws of motion had been articulated nearly 300 years before, and they had served remarkably well as the basis for an uncountable array of terrestrial applications and cosmological predictions. But to successfully land an object the size of a small truck on a moving body 385,000 km from its origin presented a whole host of issues. The mathematical model in the form of Newton's laws could not be fully trusted to accurately navigate such extreme distances. Nor could noisy and inaccurate measurements made on board the spacecraft as well as back on Earth be relied upon either. How could a variety of measurements, each with various levels of reliability and accuracy, be combined with mathematical predictions of the spacecraft's trajectory to provide the best estimate of its location and determine the necessary course corrections to keep it on its precise trajectory toward its target – and back?

The obstacle was not that the equations were not correct, or that they could not be accurately calculated; the problem was that there were numerous opportunities for the spacecraft to deviate slightly from its intended path. Each of these could have led to large deviations from the target trajectory over such large distances. A mechanism was required to correct the trajectory of the spacecraft along its route, but that was the easy part. How does one determine precisely where you are when so far from, well, anything? How does one adjust for the inevitable errors in making such a *state estimate*?

While so-called *filtering* techniques were available at the time, they were either not accurate enough for such a mission or too computationally intensive for on-board computers. It turned out that the methodology to address just such a predicament had been published by Kalman (1960) only 18 months before Kennedy's famous speech. In fact, a small group of researchers from NASA Ames Research Center had been in contact with Kalman in the late 1950s to discuss his new method and how it could be used for midcourse navigation correction of the Apollo spacecraft.[1] Kalman's original method applied to linear estimation problems, whereas NASA scientists were faced with a nonlinear problem. The modifications made to the original approach came to be known as the "extended Kalman filter," which remains one of the dominant approaches for treating nonlinear estimation problems today.

While on-board Kalman filtering ended up only being used as a backup to ground-based measurements for midcourse navigation between the Earth and Moon, it was found to be essential for the rendezvous problem. The three Apollo astronauts traveled together in the Command Module until the spacecraft entered lunar orbit. The Lunar Module then detached from the Command Module with Neil Armstrong and Buzz Aldrin aboard, while Michael Collins stayed in the Command Module as it remained in lunar orbit. After their historic landing on the Moon's surface, the Lunar Module then had to take off from the lunar surface and rendezvous with the Command Module in lunar orbit.[2] This was the most difficult navigational challenge of the entire mission. Because both spacecraft were moving, it presented a unique navigational challenge to facilitate their rendezvous with the necessary precision to dock the two spacecraft and reunite the three history-making astronauts. Kalman filtering was used to ingest the model predictions and on-board measurements to obtain relative state estimates

[1] NASA Technical Memorandum 86847, "Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry," by L. A. McGee and S. F. Schmidt (1985).

[2] The picture on the cover of the book was taken by Michael Collins in the Command Module just before the rendezvous was executed.

between the two orbiting spacecraft. The calculations were performed redundantly on both the Lunar and Command Modules' guidance computers for comparison.[3]

Before we can introduce state estimation in Section 11.9, there is a great deal of mathematical machinery that we need to explore. This includes material from this book on matrix methods as well as variational methods (see, for example, Cassel 2013). There are numerous additional examples of new technologies and techniques developed for the specific purpose of landing a man on the Moon – many of them depending upon the matrix, numerical, and optimization methods that are the subjects of this text.

1.1 Introduction

For most scientists and engineers, our first exposure to vectors and matrices is in the context of mechanics. Vectors are used to represent quantities, such as velocity or force, that have both a magnitude and direction in contrast to scalar quantities, such as pressure and temperature, that only have a magnitude. Likewise, matrices generally first appear when stress and strain tensors are introduced, once again in a mechanics setting. At first, such threedimensional vectors and matrices appear to be simply a convenient way to tabulate such quantities in an orderly fashion. In large part, this is true. In contrast to many areas of mathematics, for which certain operations would not be possible without it, vectors and matrices are not mathematical elements of necessity; rather they are constructs of convenience. It could be argued that there is not a single application in this book that *requires* matrix methods. However, this ubiquitous framework not only supplies a convenient way of representing large and complex data sets, it also provides a common formalism for their analysis; the topics contained in this book would be far more complex and confusing without the machinery of matrix methods.

The benefit of first being exposed to vectors and matrices through mechanics is that we naturally develop a strong geometric interpretation of them from the start. Because it appeals to our visual sensibilities, therefore, we hardly realize that we are learning the basics of *linear algebra*. In such a mechanics context, however, there is no reason to consider vectors larger than three dimensions corresponding to the three directions in our various coordinate systems that represent physical space. Unfortunately, our visual interpretation of vectors and matrices does not carry over to higher dimensions; we cannot even sketch a vector larger than three dimensions let alone impose a meaningful geometric interpretation. This is when linear algebra seems to lose its moorings in physically understandable reality and is simply a fun playground for mathematicians.

[3] MIT Report E-2411, "Apollo Navigation, Guidance, and Control: A Progress Report," by D. G. Hoag (1969). MIT Report R-649, "The Apollo Rendezvous Navigation Filter Theory, Description and Performance" (Volume 1), by E. S. Muller Jr. and P. M. Kachmar (1970).

Given its roots in algebra, it is not surprising that mathematicians have long viewed linear algebra as an essential weapon in their arsenal. For those of us who have benefited from a course in linear algebra taught by a mathematician, complementing the geometrically rich and physically practical exposure in mechanics with the formal framework of operations and methods provides a strong, and frankly necessary, foundation for research and practice in almost all areas of science and engineering. However, the mathematician's discussion of "singular matrices," "null spaces," and "vector bases" often leaves us with the notion that linear algebra, beyond our initial mechanics-driven exposure, is of very little relevance to the scientist or engineer. On the contrary, the formalism of linear algebra provides the mathematical foundation for three of the most far-reaching and widely applicable "applications" of matrix methods in science and engineering, namely dynamical systems theory, numerical methods, and optimization. Together they provide the tools necessary to solve, analyze, and optimize large-scale, complex systems of practical interest in both research and industrial practice. As we will see, even the analysis and prediction of continuous systems governed by differential equations ultimately reduces to solution of a matrix problem. This is because more often than not, numerical methods must be used, which convert the continuous governing equations into a discrete system of algebraic equations.

Although "linear" algebra is strictly speaking a special case of algebra,[4] it is in many ways a dramatic extension of the algebra that we learn in our formative years. This is certainly the case with regard to the applications that matrix methods address for scientists and engineers. Because linear methods are so well developed mathematically, with their wide-ranging set of tools, it is often tempting to reframe nonlinear problems in such a way as to allow for the utilization of linear methods. In many cases, this can be formally justified; however, one needs to be careful to do so in a manner that is faithful to the true nature of the underlying system and the information being sought. This theme will be revisited throughout the text as many of our applications exhibit nonlinear behavior.

Whereas vectors and matrices arise in a wide variety of applications and settings, the mathematics of these constructs is the same regardless of where the vectors or matrices have their origin. We will focus in Part I on the mathematics, but with little emphasis on formalism and proofs, and mention or illustrate many of the applications in science and engineering. First, however, let us motivate the need for such mathematical machinery using two simple geometric scenarios. After introducing some basic definitions and algebraic operations, we will then return in Section 1.4.1 to introduce some additional applications of matrix problems common in science and engineering.

1.1.1 Equation of a Line

We know intuitively that the shortest distance between two points is a straight line. Mathematically, this is reflected in the fact that there is a single unique straight line

[4] Although algebra can be traced back to the ancient Babylonians and Egyptians, matrix algebra was not formalized until the middle of the nineteenth century by Arthur Cayley.

that connects any two points. Consider the two points (x_1, y_1) and (x_2, y_2) in two dimensions. In order to determine the line,

$$a_0 + a_1 x = y,$$

passing through these two points, we could substitute the two points into this equation for the line as follows:

$$\begin{aligned} a_0 + a_1 x_1 &= y_1, \\ a_0 + a_1 x_2 &= y_2. \end{aligned} \tag{1.1}$$

Because the values of x and y are known for the two points, this is two equations for the two unknown constants a_0 and a_1, and we expect a unique solution. As will be shown in the next section, these coupled algebraic equations can conveniently be written in vector-matrix form as

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

In order to see how this corresponds to the system of equations (1.1), matrix multiplication will need to be defined. Working from left to right, we have a matrix that includes the x values multiplied by a vector containing the coefficients in the equation of the line set equal to a vector of the y values. In the present case, the matrix and right-hand-side vector are known, and a solution is sought for the coefficients a_0 and a_1 in the equation for the line. Performing the same exercise in three dimensions would result in the need to solve three equations for three unknowns, which again could be expressed in matrix form.

What if we have more than two points? Say that instead we have $N > 2$ such points. Certainly, it would not be possible to determine a single straight line that connects all N of the points (unless they so happen to all be collinear). Expressed as before, we would have N equations for the two unknown coefficients; this is called an *overdetermined system* as there are more equations than unknowns and no unique solution exists. While there is not a single line that connects all of the points in this general case, we could imagine that there is a single line that best represents the points as illustrated in Figure 1.1. This is known as *linear least-squares regression* and illustrates that there will be times when a "solution" of the system of equations is sought even when a unique solution does not exist. Least-squares methods will be taken up in Section 10.2 after we have covered the necessary background material.

1.1.2 Linear Transformation

As a second example, consider geometric transformations. When scientists and engineers communicate their ideas, theories, and results, they must always draw attention to their *reference frame*. For example, when describing the motion of a passenger walking down the aisle of an airplane in flight, is the description from the point of view of the passenger, the airplane, a fixed point on the Earth's surface, the center of the Earth, the center of the Sun, or some other point in the universe? Obviously, the

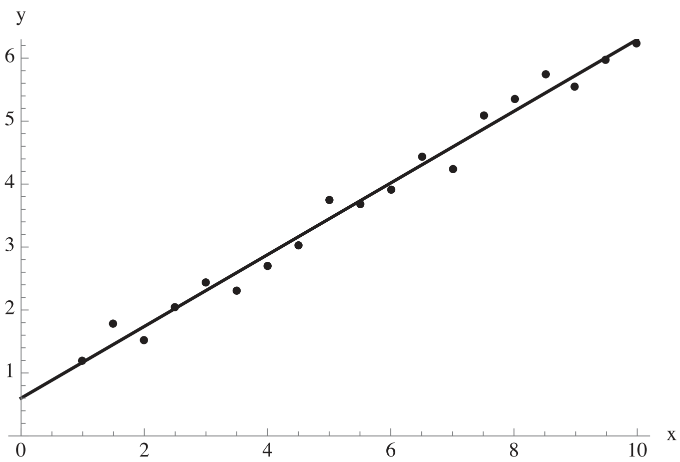

Figure 1.1 Best-fit line for $N = 19$ data points.

description will be quite different depending on which of these reference frames is used. Mathematically, we formally define the reference frame and how we will locate positions relative to it using a *coordinate system* or *vector basis*.[5] This is comprised of specifying the location of the coordinate system's origin – the location from which each coordinate and vector is measured – and each of the coordinate directions emanating from this origin.

Of course, one may specify a problem with respect to a different coordinate system than someone else, or different aspects of the problem may be best analyzed using different coordinate systems. In order to communicate information between the two coordinate systems, a *transformation* is necessary. For example, let us say that a point in the two-dimensional coordinate system X is given by $\mathbf{x}_1 = (2, 5)$, and the same point with respect to another coordinate system Y is defined by $\mathbf{y}_1 = (-3, 4)$. If this is a *linear* transformation, then we must be able to multiply $\mathbf{x}_1$ times something, say $\mathbf{A}$, to produce $\mathbf{y}_1$ in the form

$$\mathbf{A}\mathbf{x}_1 = \mathbf{y}_1. \tag{1.2}$$

Because both $\mathbf{x}_1$ and $\mathbf{y}_1$ involve two values, or coordinates, $\mathbf{A}$ is clearly not simply a single scalar value under general circumstances. Instead, we will need to include some combination of both coordinates to get from one coordinate system to the other. This could involve, for example, each coordinate of $\mathbf{y}_1$ being some linear combination of the two coordinates of $\mathbf{x}_1$, such that

[5] We generally call it a *coordinate system* when referring to a system with one, two, or three spatial dimensions, for which we can illustrate the coordinate system geometrically. For systems with more than three dimensions, that is, where the dimensions do not correspond to spatial coordinates, we use the more general *vector basis*, or simply *basis*, terminology. Consequently, a coordinate system is simply a vector basis for a two- or three-dimensional vector space.

$$\begin{aligned} 2a + 5b &= -3, \\ 2c + 5d &= 4. \end{aligned} \tag{1.3}$$

If we suitably define multiplication, (1.3) can be written in the compact form of (1.2) if **A** is defined as

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

which is a *transformation matrix*.[6] It just remains to determine the constants a, b, c, and d that comprise the transformation matrix. What we have in (1.3) is a system of two coupled algebraic equations for these four unknown constants. Clearly, there is no unique solution as there are many, in fact infinite, ways to transform the single point $\mathbf{x}_1$ into the point $\mathbf{y}_1$. For example, we could simply translate the coordinate system or rotate it clockwise or counterclockwise to transform one point to the other.

A unique transformation is obtained if we supply an additional pair of *image* points in the two coordinate systems, say $\mathbf{x}_2 = (-1, 6)$ and $\mathbf{y}_2 = (8, -7)$. This results in the two additional equations

$$\begin{aligned} -a + 6b &= 8, \\ -c + 6d &= -7, \end{aligned} \tag{1.4}$$

which in compact form is

$$\mathbf{A}\mathbf{x}_2 = \mathbf{y}_2. \tag{1.5}$$

Note that the transformation matrix is the same in (1.2) and (1.5). Equations (1.3) and (1.4) now provide four equations for the four unknowns, and the transformation is unique.

Is there a way to represent the *system of linear algebraic equations* given by (1.3) and (1.4) in a convenient mathematical form? How do we solve such a system? Can we be sure that this solution is unique? Can we represent *linear transformations* in a general fashion for systems having any number of coordinates? These questions are the subject of this chapter. As we will see, the points and the coordinate directions with respect to which they are defined will be conveniently represented as *vectors*. The transformation will be denoted by a *matrix* as will be the coefficients in the system of linear algebraic equations to be solved for the transformation matrix. The unknowns in the transformation matrix will be combined to form the *solution vector*.

Such changes of coordinate system (basis) and the transformation matrices that accomplish them will be a consistent theme throughout the text. We will often have occasion to transform a problem into a more desirable basis in order to facilitate interpretation, expose features, diagnose attributes, or ease solution. We will return

[6] Given their ubiquity across so many areas of mathematics, science, and engineering, one might be surprised to learn that the term "matrix" was not coined until 1850 by James Joseph Sylvester. This is more than a century after such classical fields as complex variables, differential calculus, and variational calculus had their genesis.

to this important topic in Section 1.8. Along the way, we will encounter numerous physical applications that lend themselves to similar mathematical representation as this geometric example despite their vastly different physical interpretations. As is so often the case in mathematics, this remarkable ability to unify many disparate applications within the same mathematical constructs and operations is what renders mathematics so essential to the scientist or engineer.

1.2 Definitions

Let us begin by defining vectors and matrices and several common types of matrices. One can think of a matrix as the mathematical analog of a table in a text or a spreadsheet on a computer.

Matrix: A *matrix* is an ordered arrangement of numbers, variables, or functions comprised of a rectangular grouping of elements arranged in rows and columns as follows:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix} = [A_{mn}] .$$

The size of the matrix is denoted by the number of rows, M, and the number of columns, N. We say that the matrix $\mathbf{A}$ is $M \times N$, which we read as "M by N." If $M = N$, then the matrix is said to be *square*. Each element A_{mn} in the matrix is uniquely identified by two subscripts, with the first, m, being its row and the second, n, being its column. Thus, $1 \leq m \leq M$ and $1 \leq n \leq N$. The elements A_{mn} may be real or complex numbers, variables, or functions.

The *main diagonal* of the matrix $\mathbf{A}$ is given by $A_{11}, A_{22}, \ldots, A_{MM}$ or A_{NN}; if the matrix is square, then $A_{MM} = A_{NN}$. Two matrices are said to be equal, that is $\mathbf{A} = \mathbf{B}$, if their sizes are the same and $A_{mn} = B_{mn}$ for all m and n.

Vector: A (column) *vector* is an $N \times 1$ matrix. For example,

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} .$$

The vector is said to be N-dimensional and can be considered a point in an N-dimensional coordinate system. By common convention, matrices are denoted by bold capital letters and vectors by bold lowercase letters.

Matrix Transpose (Adjoint): The *transpose* of matrix $\mathbf{A}$ is obtained by interchanging its rows and columns as follows:

$$\mathbf{A}^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{M1} \\ A_{12} & A_{22} & \cdots & A_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1N} & A_{2N} & \cdots & A_{MN} \end{bmatrix} = [A_{nm}],$$

which results in an $N \times M$ matrix. If $\mathbf{A}^T = \mathbf{A}$, then $\mathbf{A}$ is said to be *symmetric* ($A_{nm} = A_{mn}$). Note that a matrix must be square to be symmetric. If instead the matrix is such that $\mathbf{A} = -\mathbf{A}^T$, it is called *skew-symmetric*. Note that for this to be true, the elements along the main diagonal of $\mathbf{A}$ must all be zero.

If the elements of $\mathbf{A}$ are complex and $\overline{\mathbf{A}}^T = \mathbf{A}$, then $\mathbf{A}$ is a *Hermitian matrix* ($\overline{A}_{nm} = A_{mn}$), where the overbar represents the *complex conjugate*, and $\overline{\mathbf{A}}^T$ is the *conjugate transpose* of $\mathbf{A}$. Note that a symmetric matrix is a special case of a Hermitian matrix.

Zero Matrix (**0**): Matrix of all zeros.

Identity Matrix (**I**): Square matrix with ones on the main diagonal and zeros everywhere else, for example,

$$\mathbf{I}_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = [\delta_{mn}],$$

where

$$\delta_{mn} = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}.$$

Triangular Matrix: All elements above (left triangular) or below (right triangular) the main diagonal are zero. For example,

$$\mathbf{L} = \begin{bmatrix} A_{11} & 0 & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 & 0 \\ A_{31} & A_{32} & A_{33} & 0 & 0 \\ A_{41} & A_{42} & A_{43} & A_{44} & 0 \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ 0 & A_{22} & A_{23} & A_{24} & A_{25} \\ 0 & 0 & A_{33} & A_{34} & A_{35} \\ 0 & 0 & 0 & A_{44} & A_{45} \\ 0 & 0 & 0 & 0 & A_{55} \end{bmatrix}.$$

Tridiagonal Matrix: All elements are zero except along the lower (first subdiagonal), main, and upper (first superdiagonal) diagonals as follows:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 & 0 \\ A_{21} & A_{22} & A_{23} & 0 & 0 \\ 0 & A_{32} & A_{33} & A_{34} & 0 \\ 0 & 0 & A_{43} & A_{44} & A_{45} \\ 0 & 0 & 0 & A_{54} & A_{55} \end{bmatrix}.$$

Hessenberg Matrix: All elements are zero below the lower diagonal, that is

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ 0 & A_{32} & A_{33} & A_{34} & A_{35} \\ 0 & 0 & A_{43} & A_{44} & A_{45} \\ 0 & 0 & 0 & A_{54} & A_{55} \end{bmatrix}.$$

Toeplitz Matrix: Each diagonal is a constant, such that

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{11} & A_{12} & A_{13} & A_{14} \\ A_{31} & A_{21} & A_{11} & A_{12} & A_{13} \\ A_{41} & A_{31} & A_{21} & A_{11} & A_{12} \\ A_{51} & A_{41} & A_{31} & A_{21} & A_{11} \end{bmatrix}.$$

Matrix Inverse: If a square matrix $\mathbf{A}$ is *invertible*, then its inverse $\mathbf{A}^{-1}$ is such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Note: If $\mathbf{A}$ is the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then its inverse is

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix},$$

where the diagonal elements A_{11} and A_{22} have been exchanged, the off-diagonal elements A_{12} and A_{21} have each switched signs, and $|\mathbf{A}| = A_{11}A_{22} - A_{12}A_{21}$ is the *determinant* of $\mathbf{A}$ (see Section 1.4.3).

Orthogonal Matrix: An $N \times N$ square matrix $\mathbf{A}$ is *orthogonal* if

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}.$$

It follows that

$$\mathbf{A}^T = \mathbf{A}^{-1}$$

for an orthogonal matrix. Such a matrix is called orthogonal because its column (and row) vectors are mutually orthogonal (see Sections 1.2 and 2.4).

Block Matrix: A block matrix is comprised of smaller submatrices, called *blocks*. For example, suppose that matrix $\mathbf{A}$ is a 2×2 block matrix comprised of four submatrices as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$