Numerical Analysis Lecture Notes

Peter J. Olver

3. Review of Matrix Algebra

Vectors and matrices are essential for modern analysis of systems of equations — algebrai, differential, functional, etc. In this part, we will review the most basic facts of matrix arithmetic. See [**42**] for full details.

3.1. Matrices and Vectors.

A *matrix* is a rectangular array of numbers. Thus,

$$\begin{pmatrix} 1 & 0 & 3 \\ -2 & 4 & 1 \end{pmatrix}, \qquad \begin{pmatrix} \pi & 0 \\ e & \frac{1}{2} \\ -1 & .83 \\ \sqrt{5} & -\frac{4}{7} \end{pmatrix}, \qquad (\,.2 \quad -1.6 \quad .32\,), \qquad \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 3 \\ -2 & 5 \end{pmatrix},$$

are all examples of matrices. We use the notation

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{3.1}$$

for a general matrix of size $m \times n$ (read "m by n"), where m denotes the number of *rows* in A and n denotes the number of *columns*. Thus, the preceding examples of matrices have respective sizes 2×3, 4×2, 1×3, 2×1, and 2×2. A matrix is *square* if $m = n$, i.e., it has the same number of rows as columns. A *column vector* is a $m \times 1$ matrix, while a *row vector* is a $1 \times n$ matrix. As we shall see, column vectors are by far the more important of the two, and the term "vector" without qualification will always mean "column vector". A 1×1 matrix, which has but a single entry, is both a row and a column vector.

The number that lies in the ith row and the jth column of A is called the (i, j) *entry* of A, and is denoted by a_{ij}. The row index always appears first and the column index second. Two matrices are equal, $A = B$, if and only if they have the same size, and *all* their entries are the same: $a_{ij} = b_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

A general linear system of m equations in n unknowns will take the form

$$
\begin{aligned}
a_{11} x_1 + a_{12} x_2 + \;\cdots\; + a_{1n} x_n &= b_1, \\
a_{21} x_1 + a_{22} x_2 + \;\cdots\; + a_{2n} x_n &= b_2, \\
\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \qquad\quad& \\
a_{m1} x_1 + a_{m2} x_2 + \;\cdots\; + a_{mn} x_n &= b_m.
\end{aligned}
\tag{3.2}
$$

As such, it is composed of three basic ingredients: the $m \times n$ *coefficient matrix* A, with entries a_{ij} as in (3.1), the column vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ containing the *unknowns*, and the column vector $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ containing *right hand sides*. As an example, consider the linear system

$$
\begin{aligned}
x + 2y + z &= 2, \\
2y + z &= 7, \\
x + y + 4z &= 3,
\end{aligned}
$$

The coefficient matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 4 \end{pmatrix}$ can be filled in, entry by entry, from the coefficients of the variables appearing in the equations. (Don't forget to put a zero when a avariable doesn't appear in an equation!) The vector $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ lists the variables, while the entries of $\mathbf{b} = \begin{pmatrix} 2 \\ 7 \\ 3 \end{pmatrix}$ are the right hand sides of the equations.

Remark: We will consistently use bold face lower case letters to denote vectors, and ordinary capital letters to denote general matrices.

Matrix Arithmetic

Matrix arithmetic involves three basic operations: *matrix addition*, *scalar multiplication*, and *matrix multiplication*. First we define *addition* of matrices. You are only allowed to add two matrices of the *same size*, and matrix addition is performed entry by entry. For example,

$$
\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & -5 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 1 \end{pmatrix}.
$$

Therefore, if A and B are $m \times n$ matrices, their sum $C = A + B$ is the $m \times n$ matrix whose entries are given by $c_{ij} = a_{ij} + b_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. When defined, matrix

addition is commutative, $A + B = B + A$, and associative, $A + (B + C) = (A + B) + C$, just like ordinary addition.

A *scalar* is a fancy name for an ordinary number — the term merely distinguishes it from a vector or a matrix. For the time being, we will restrict our attention to real scalars and matrices with real entries, but eventually complex scalars and complex matrices must be dealt with. We will consistently identify a scalar $c \in \mathbb{R}$ with the 1×1 matrix (c) in which it is the sole entry, and so will omit the redundant parentheses in the latter case. *Scalar multiplication* takes a scalar c and an $m \times n$ matrix A and computes the $m \times n$ matrix $B = c\,A$ by multiplying each entry of A by c. For example,

$$3 \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ -3 & 0 \end{pmatrix}.$$

In general, $b_{ij} = c\,a_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Basic properties of scalar multiplication are summarized at the end of this section.

Finally, we define *matrix multiplication*. First, the product between a row vector $\mathbf{a}$ and a column vector $\mathbf{x}$ having the *same* number of entries is the *scalar* or 1×1 matrix defined by the following rule:

$$\mathbf{a}\,\mathbf{x} = (\, a_1 \; a_2 \; \ldots \; a_n \,) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = \sum_{k=1}^{n} a_k x_k. \qquad (3.3)$$

More generally, if A is an $m \times n$ matrix and B is an $n \times p$ matrix, so that the number of *columns* in A equals the number of *rows* in B, then the matrix product $C = A\,B$ is defined as the $m \times p$ matrix whose (i, j) entry equals the vector product of the ith row of A and the jth column of B. Therefore,

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}. \qquad (3.4)$$

Note that our restriction on the sizes of A and B guarantees that the relevant row and column vectors will have the same number of entries, and so their product is defined.

For example, the product of the coefficient matrix A and vector of unknowns $\mathbf{x}$ for our original system (4.1) is given by

$$A\,\mathbf{x} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ 2x + 6y + z \\ x + y + 4z \end{pmatrix}.$$

The result is a column vector whose entries reproduce the left hand sides of the original linear system! As a result, we can rewrite the system

$$A\,\mathbf{x} = \mathbf{b} \qquad (3.5)$$

as an equality between two column vectors. This result is general; a linear system (3.2) consisting of m equations in n unknowns can be written in the matrix form (3.5) where A

is the $m \times n$ coefficient matrix (3.1), $\mathbf{x}$ is the $n \times 1$ column vector of unknowns, and $\mathbf{b}$ is the $m \times 1$ column vector containing the right hand sides. This is one of the principal reasons for the non-evident definition of matrix multiplication. Component-wise multiplication of matrix entries turns out to be almost completely useless in applications.

Now, the bad news. Matrix multiplication is *not* commutative — that is, BA is not necessarily equal to AB. For example, BA may not be defined even when AB is. Even if both are defined, they may be different sized matrices. For example the product $s = \mathbf{r}\,\mathbf{c}$ of a row vector $\mathbf{r}$, a $1 \times n$ matrix, and a column vector $\mathbf{c}$, an $n \times 1$ matrix with the same number of entries, is a 1×1 matrix or scalar, whereas the reversed product $C = \mathbf{c}\,\mathbf{r}$ is an $n \times n$ matrix. For instance,

$$(1 \quad 2)\begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3, \qquad \text{whereas} \qquad \begin{pmatrix} 3 \\ 0 \end{pmatrix}(1 \quad 2) = \begin{pmatrix} 3 & 6 \\ 0 & 0 \end{pmatrix}.$$

In computing the latter product, don't forget that we multiply the *rows* of the first matrix by the *columns* of the second. Moreover, even if the matrix products AB and BA have the same size, which requires both A and B to be square matrices, we may still have $AB \neq BA$. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 5 \\ -4 & 11 \end{pmatrix} \neq \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

On the other hand, matrix multiplication is associative, so $A(BC) = (AB)C$ whenever A has size $m \times n$, B has size $n \times p$, and C has size $p \times q$; the result is a matrix of size $m \times q$. The proof of associativity is a tedious computation based on the definition of matrix multiplication that, for brevity, we omit. Consequently, the one difference between matrix algebra and ordinary algebra is that you need to be careful not to change the order of multiplicative factors without proper justification.

Since matrix multiplication acts by multiplying rows by columns, one can compute the columns in a matrix product AB by multiplying the matrix A and the individual columns of B. For example, the two columns of the matrix product

$$\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 8 & 6 \end{pmatrix}$$

are obtained by multiplying the first matrix with the individual columns of the second:

$$\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

In general, if we use $\mathbf{b}_k$ to denote the kth column of B, then

$$AB = A\begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_p \end{pmatrix} = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \dots & A\mathbf{b}_p \end{pmatrix}, \tag{3.6}$$

indicating that the kth column of their matrix product is $A\mathbf{b}_k$.

There are two special matrices. The first is the *zero matrix*, all of whose entries are 0. We use $\mathrm{O}_{m \times n}$ to denote the $m \times n$ zero matrix, often written as just O if the size is clear

from the context. The zero matrix is the additive unit, so $A + O = A = O + A$ when O has the same size as A. In particular, we will use a bold face $\mathbf{0}$ to denote a column vector with all zero entries.

The role of the multiplicative unit is played by the square *identity matrix*

$$\mathrm{I} = \mathrm{I}_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

of size $n \times n$. The entries along the *main diagonal* (which runs from top left to bottom right) are equal to 1, while the *off-diagonal* entries are all 0. As you can check, if A is any $m \times n$ matrix, then $\mathrm{I}_m A = A = A \, \mathrm{I}_n$. We will sometimes write the preceding equation as just $\mathrm{I} A = A = A \, \mathrm{I}$, since each matrix product is well-defined for exactly one size of identity matrix.

The identity matrix is a particular example of a *diagonal matrix*. In general, a square matrix A is diagonal if all its off-diagonal entries are zero: $a_{ij} = 0$ for all $i \neq j$. We will sometimes write $D = \mathrm{diag}\,(c_1, \ldots, c_n)$ for the $n \times n$ diagonal matrix with diagonal entries $d_{ii} = c_i$. Thus, $\mathrm{diag}\,(1, 3, 0)$ refers to the diagonal matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, while the $n \times n$ identity matrix can be written as $\mathrm{I}_n = \mathrm{diag}\,(1, 1, \ldots, 1)$.

Let us conclude this section by summarizing the basic properties of matrix arithmetic. In the accompanying table, A, B, C are matrices; c, d are scalars; O is a zero matrix; and I is an identity matrix. All matrices are assumed to have the correct sizes so that the indicated operations are defined.

Matrix Inverses

The inverse of a matrix is analogous to the reciprocal $a^{-1} = 1/a$ of a scalar. We already encountered the inverses of matrices corresponding to elementary row operations. In this section, we will study inverses of general square matrices. We begin with the formal definition.

Definition 3.1. Let A be a square matrix of size $n \times n$. An $n \times n$ matrix X is called the *inverse* of A if it satisfies

$$X A = \mathrm{I} = A X, \tag{3.7}$$

where $\mathrm{I} = \mathrm{I}_n$ is the $n \times n$ identity matrix. The inverse is commonly denoted by $X = A^{-1}$.

Remark: Noncommutativity of matrix multiplication requires that we impose both conditions in (3.7) in order to properly define an inverse to the matrix A. The first condition, $X A = \mathrm{I}$, says that X is a *left inverse*, while the second, $A X = \mathrm{I}$, requires that X also be a *right inverse*. Rectangular matrices might have either a left inverse or a right inverse, but, as we shall see, *only* square matrices have both, and so only square matrices

Matrix Addition:	Commutativity	$A + B = B + A$
	Associativity	$(A + B) + C = A + (B + C)$
	Zero Matrix	$A + O = A = O + A$
	Inverse	$A + (-A) = O, \quad -A = (-1)A$
Scalar Multiplication:	Associativity	$c(dA) = (cd)A$
	Distributivity	$c(A + B) = (cA) + (cB)$ $(c + d)A = (cA) + (dA)$
	Unit	$1A = A$
	Zero	$0A = O$
Matrix Multiplication:	Associativity	$(AB)C = A(BC)$
	Distributivity	$A(B + C) = AB + AC,$ $(A + B)C = AC + BC,$
	Identity Matrix	$AI = A = IA$
	Zero Matrix	$AO = O, \quad OA = O$

can have full-fledged inverses. However, not every square matrix has an inverse. Indeed, not every scalar has an inverse: $0^{-1} = 1/0$ is not defined since the equation $0x = 1$ has no solution.

Example 3.2. Since

$$
\begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix},
$$

we conclude that when $A = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$, then $A^{-1} = \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix}$. Observe that there is no obvious way to anticipate the entries of A^{-1} from the entries of A.

Example 3.3. Let us compute the inverse $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$, when it exists, of a general 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The right inverse condition

$$
AX = \begin{pmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I
$$

holds if and only if x, y, z, w satisfy the linear system

$$a\,x + b\,z = 1, \qquad a\,y + b\,w = 0,$$
$$c\,x + d\,z = 0, \qquad c\,y + d\,w = 1.$$

Solving by Gaussian Elimination (or directly), we find

$$x = \frac{d}{a\,d - b\,c}, \qquad y = -\frac{b}{a\,d - b\,c}, \qquad z = -\frac{c}{a\,d - b\,c}, \qquad w = \frac{a}{a\,d - b\,c},$$

provided the common denominator $a\,d - b\,c \neq 0$ does not vanish. Therefore, the matrix

$$X = \frac{1}{a\,d - b\,c} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

forms a right inverse to A. However, a short computation shows that it also defines a left inverse:

$$X\,A = \begin{pmatrix} x\,a + y\,c & x\,b + y\,d \\ z\,a + w\,c & z\,b + w\,d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I},$$

and hence $X = A^{-1}$ is the inverse to A.

The denominator appearing in the preceding formulae has a special name; it is called the *determinant* of the 2×2 matrix A, and denoted

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,d - b\,c. \tag{3.8}$$

Thus, the determinant of a 2×2 matrix is the product of the diagonal entries minus the product of the off-diagonal entries. Thus, the 2×2 matrix A is invertible, with

$$A^{-1} = \frac{1}{a\,d - b\,c} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{3.9}$$

if and only if $\det A \neq 0$. For example, if $A = \begin{pmatrix} 1 & 3 \\ -2 & -4 \end{pmatrix}$, then $\det A = 2 \neq 0$. We conclude that A has an inverse, which, by (3.9), is $A^{-1} = \dfrac{1}{2} \begin{pmatrix} -4 & -3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -\frac{3}{2} \\ 1 & \frac{1}{2} \end{pmatrix}$.

Lemma 3.4. *The inverse of a square matrix, if it exists, is unique.*

Proof: Suppose both X and Y satisfy (3.7), so $X\,A = \mathrm{I} = A\,X$ and $Y\,A = \mathrm{I} = A\,Y$. Then, by associativity, $X = X\,\mathrm{I} = X(A\,Y) = (X\,A)Y = \mathrm{I}\,Y = Y$, and hence $X = Y$. *Q.E.D.*

Inverting a matrix twice brings us back to where we started.

Lemma 3.5. *If A is invertible, then A^{-1} is also invertible and $(A^{-1})^{-1} = A$.*

Proof: The matrix inverse equations $A^{-1}\,A = \mathrm{I} = A\,A^{-1}$ are sufficient to prove that A is the inverse of A^{-1}. *Q.E.D.*

Lemma 3.6. *If A and B are invertible matrices of the same size, then their product, AB, is invertible, and*

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{3.10}$$

Note that the order of the factors is reversed under inversion.

Proof: Let $X = B^{-1}A^{-1}$. Then, by associativity,

$$X(AB) = B^{-1}A^{-1}AB = B^{-1}B = \mathrm{I}, \qquad (AB)X = ABB^{-1}A^{-1} = AA^{-1} = \mathrm{I}.$$

Thus X is both a left and a right inverse for the product matrix AB and the result follows. $\hspace{10cm}$ *Q.E.D.*

Example 3.7. One verifies, directly, that the inverse of $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ is $A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$, while the inverse of $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is $B^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Therefore, the inverse of their product $C = AB = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix}$ is given by $C^{-1} = B^{-1}A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}$.

We can straightforwardly generalize the preceding result. The inverse of a k-fold product of invertible matrices is the product of their inverses, *in the reverse order*:

$$(A_1 A_2 \cdots A_{k-1} A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \cdots A_2^{-1} A_1^{-1}. \tag{3.11}$$

Warning: In general, $(A + B)^{-1} \neq A^{-1} + B^{-1}$. This equation is not even true for scalars (1×1 matrices)!

Transposes and Symmetric Matrices

Another basic operation on matrices is to interchange their rows and columns. If A is an $m \times n$ matrix, then its *transpose*, denoted A^T, is the $n \times m$ matrix whose (i, j) entry equals the (j, i) entry of A; thus

$$B = A^T \qquad \text{means that} \qquad b_{ij} = a_{ji}.$$

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad \text{then} \qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Observe that the rows of A become the columns of A^T and vice versa. In particular, the transpose of a row vector is a column vector, while the transpose of a column vector is a row vector; if $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, then $\mathbf{v}^T = (\,1 \;\; 2 \;\; 3\,)$. The transpose of a scalar, considered as a 1×1 matrix, is itself: $c^T = c$.

Remark: Most vectors appearing in applied mathematics are column vectors. To conserve vertical space in this text, we will often use the transpose notation, e.g., $\mathbf{v} = (v_1, v_2, v_3)^T$, as a compact way of writing column vectors.

In the square case, transposition can be viewed as "reflecting" the matrix entries across the main diagonal. For example,

$$\begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 5 \\ -2 & -4 & 8 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 0 & -4 \\ -1 & 5 & 8 \end{pmatrix}.$$

In particular, the transpose of a lower triangular matrix is upper triangular and vice-versa.

Transposing twice returns you to where you started:

$$(A^T)^T = A. \tag{3.12}$$

Unlike inversion, transposition *is* compatible with matrix addition and scalar multiplication:

$$(A + B)^T = A^T + B^T, \qquad (cA)^T = cA^T. \tag{3.13}$$

Transposition is also compatible with matrix multiplication, but with a twist. Like the inverse, the transpose *reverses* the order of multiplication:

$$(AB)^T = B^T A^T. \tag{3.14}$$

Indeed, if A has size $m \times n$ and B has size $n \times p$, so they can be multiplied, then A^T has size $n \times m$ and B^T has size $p \times n$, and so, in general, one has no choice but to multiply $B^T A^T$ in that order. Formula (3.14) is a straightforward consequence of the basic laws of matrix multiplication. An important special case is the product between a row vector $\mathbf{v}^T$ and a column vector $\mathbf{w}$ with the same number of entries. In this case,

$$\mathbf{v}^T \mathbf{w} = (\mathbf{v}^T \mathbf{w})^T = \mathbf{w}^T \mathbf{v}, \tag{3.15}$$

because their product is a scalar and so, as noted above, equals its own transpose.

Lemma 3.8. *If A is a nonsingular matrix, so is A^T, and its inverse is denoted*

$$A^{-T} = (A^T)^{-1} = (A^{-1})^T. \tag{3.16}$$

Thus, transposing a matrix and then inverting yields the same result as first inverting and then transposing.

Proof: Let $X = (A^{-1})^T$. Then, according to (3.14),

$$X A^T = (A^{-1})^T A^T = (A A^{-1})^T = \mathbf{I}^T = \mathbf{I}.$$

The proof that $A^T X = \mathbf{I}$ is similar, and so we conclude that $X = (A^T)^{-1}$.   *Q.E.D.*

A particularly important class of square matrices is those that are unchanged by the transpose operation.

Definition 3.9. A square matrix is called *symmetric* if it equals its own transpose: $A = A^T$.

Thus, A is symmetric if and only if its entries satisfy $a_{ji} = a_{ij}$ for all i, j. In other words, entries lying in "mirror image" positions relative to the main diagonal must be equal. For example, the most general symmetric 3×3 matrix has the form

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}.$$

Note that any diagonal matrix, including the identity, is symmetric. A lower or upper triangular matrix is symmetric if and only if it is, in fact, a diagonal matrix.