2×2 MATRIX ALGEBRA

James T. Smith
San Francisco State University

These notes introduce some fundamental notions of linear algebra: definitions and properties of algebraic operations on vectors and matrices. They're confined to two dimensions for two reasons. First, this eases your transition from elementary to higher mathematics. Two dimensional matrix algebra is mainly a notational device. There's little content here beyond elementary algebra. Second, two dimensional matrix algebra is sufficient for plane Cartesian coordinate geometry.

Linear algebra applies to both real and complex numbers. When you read these notes, you may interpret the word *scalar* always to mean *real number*, or always to mean *complex number*. Although there are some differences between real and complex linear algebra, they don't appear at the level covered here. Scalars will be denoted by small Latin letters $a, b, c, \ldots$.

You can write a pair of scalars as a row or column. A column is called a *vector*, and denoted by a small Greek letter $\alpha, \beta, \gamma, \ldots$. Its entries are identified by the corresponding Latin letter, with subscripts. The corresponding row is indicated by a prime. For example, consider

$$\xi = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \xi' = [x_1, x_2].$$

The prime is also used to convert a row back to the corresponding column, so that $\xi'' = \xi$ for any vector ξ.

You can *add* two vectors:

$$\xi + \eta = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}.$$

Vectors satisfy *commutative* and *associative* laws for addition:

$$\xi + \eta = \eta + \xi$$
$$\xi + (\eta + \zeta) = (\xi + \eta) + \zeta.$$

Therefore, as in scalar algebra, you can rearrange repeated sums at will and omit parentheses.

The *zero vector* is

$$o = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Every vector has a *negative*:

$$-\xi = -\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ -x_2 \end{bmatrix}.$$

Clearly,

$$-o = o \qquad \xi + o = \xi \qquad \xi + (-\xi) = o \qquad -(-\xi) = \xi.$$

You can *multiply a vector by a scalar*:

$$a\xi = a\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \end{bmatrix}.$$

Verify the following manipulative rules:

$$1\xi = \xi \qquad 0\xi = o \qquad (-a)\xi = -(a\xi) = a(-\xi)$$
$$(-1)\xi = -\xi \qquad ao = o$$

$$a(b\xi) = (ab)\xi \qquad \text{—\textit{associative law}}$$

$$(a + b)\xi = a\xi + b\xi \qquad \text{—\textit{distributive laws}}$$
$$a(\xi + \eta) = a\xi + a\eta.$$

Similarly, you can add two rows, and define the zero row, the negative of a row, and the product of a row by a scalar. Moreover,

$$\xi' + \eta' = (\xi + \eta)' \qquad -(\xi') = (-\xi)' \qquad a(\xi') = (a\xi)'.$$

You can *multiply a row by a vector*:

$$\xi'\eta = [x_1, x_2]\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 + x_2y_2.$$

This is often called a *dot* or *scalar* product. With a little algebra you can verify the following manipulation rules:

$$o'\xi = 0 = \xi'o \qquad (a\xi')\eta = a(\xi'\eta) = \xi'(a\eta)$$
$$\xi'\eta = \eta'\xi \qquad (-\xi')\eta = -(\xi'\eta) = \xi'(-\eta)$$

$$(\xi' + \eta')\zeta = \xi'\zeta + \eta'\zeta \qquad \text{—\textit{distributive} laws}$$
$$\xi(\eta + \zeta) = \xi'\eta + \xi'\zeta.$$

A *matrix* is a 2×2 array of scalars. Matrices are denoted by large Latin letters $A, B, C, \ldots,$ and their entries, by corresponding small letters with subscripts:

$$A = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{2 \text{ columns.}} \left.\vphantom{\begin{bmatrix} a \\ a \end{bmatrix}}\right\} \quad 2 \text{ rows}$$

You can *add* matrices:

$$A + B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}.$$

The *zero matrix* is

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

You can also define the *negative* of a matrix, and the *product of a matrix by a scalar*:

$$-A = \begin{bmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{bmatrix} \qquad cA = \begin{bmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{bmatrix}.$$

Manipulation rules analogous to those derived earlier for vectors and rows hold for matrices as well. Check them yourself.

You can regard *subtraction* of two vectors, rows, or matrices as composition of negation and addition. For example, $\xi - \eta = \xi + (-\eta)$. You should state and verify appropriate manipulation rules.

You can *multiply a matrix A by a vector ξ*. The product $A\xi$ is the vector whose entries are the products of the rows of A by ξ:

$$A\xi = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}.$$

This equation defines a mapping $\xi \to A\xi$ from $\mathbb{R}^2$ to $\mathbb{R}^2$. You can verify the following manipulation rules:

$$O\xi = o = Ao \qquad\qquad (cA)\xi = c(A\xi) = A(c\xi)$$
$$(-A)\xi = -(A\xi) = A(-\xi)$$

$(A + B)\xi = A\xi + B\xi$ —*distributive laws*
$A(\xi + \eta) = A\xi + A\eta.$

The definition of the product of a matrix by a column vector was motivated by the appearance of a system of two linear equations in two unknowns x_1 and x_2. You can rewrite the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad \text{as} \quad A\xi = \beta.$$

Similarly, you can *multiply a row* ξ' *by a matrix* A. The product $\xi'A$ is the row whose entries are the products of ξ' by the columns of A:

$$\xi'A = [x_1, x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = [x_1a_{11} + x_2a_{21},\ x_1a_{12} + x_2a_{22}].$$

Similar manipulation rules hold. Moreover, you can check the *associative law*

$$\xi'(A\eta) = (\xi'A)\eta.$$

You can *multiply two matrices* A *and* B. The product AB is a matrix that can be described in two ways. Its columns are the products of A by the columns of B, and its rows are the products of the rows of A by B:

$$AB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

The ikth entry of AB is thus $a_{i1}b_{1k} + a_{i2}b_{2k}$. You can check the manipulation rules

$AO = O = OB$

$(aA)C = a(AC) = A(aC)$
$(-A)C = -(AC) = A(-C)$

$(A + B)C = AC + BC$ —*distributive laws*
$A(C + D) = AC + AD.$

The definition of the product of two matrices was motivated by the formulas for linear substitution. From

$$\begin{cases} z_1 = a_{11}y_1 + a_{12}y_2 \\ z_2 = a_{21}y_1 + a_{22}y_2 \end{cases} \qquad \begin{cases} y_1 = b_{11}x_1 + b_{12}x_2 \\ y_2 = b_{21}x_1 + b_{22}x_2 \end{cases}$$

you can derive

$$\begin{cases} z_1 = (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 \\ z_2 = (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2 \end{cases}.$$

That is, from $\zeta = A\eta$ and $\eta = B\xi$ you can derive $\zeta = (AB)\xi$. In short,

$$A(B\xi) = (AB)\xi \qquad \text{—}\textit{associative law}.$$

From this rule, you can deduce the *general* associative law:

$$A(BC) = (AB)C.$$

Proof: jth column of $A(BC)$ = $A(j$th column of $BC)$ = $A(B\cdot j$th column of $C)$
= $AB(j$th column of $C)$ = jth column of $(AB)C$.

The *commutative* law $AB = BA$ doesn't hold in general. For example,

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This also shows that the product of nonzero matrices can be O.

Every matrix A has a *transpose* A', the matrix whose jith entry is the ijth entry of A:

$$A' = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}.$$

The following manipulation rules hold:

$$A'' = A \qquad O' = O$$
$$(A + B)' = A' + B' \qquad (cA)' = c(A').$$

For any matrices A and B and any vector ξ,

$$(A\xi)' = \xi' A' \qquad (AB)' = B'A'.$$

Proof: ith entry of $A\xi$ = $(i$th row of $A)\xi$ = $\xi'(i$th row of $A)'$
= $\xi'(i$th column of $A')$ = ith entry of $\xi' A'$.

ith row of $(AB)'$ = ith column of AB = $A(i$th column of $B)$
= $(i$th column of $B)'A'$ = $(i$th row of $B')A'$
= ith row of $B'A'$.

The *unit* vectors are

$$u^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad u^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For any row ξ', $\xi' u^j$ is the jth entry of ξ'. For any matrix A, Au^j is the jth column of A. For example,

$$\xi' u^1 = [x_1, x_2] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = x_1$$

$$Au^1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}.$$

The *identity* matrix is constructed from the unit vectors:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Clearly, $I' = I$. Also, $\xi' I = \xi'$:

$$\xi' I = [x_1, x_2] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [x_1, x_2].$$

This implies that $AI = A$ for any matrix A.

Similarly, the unit rows $u^{i\prime}$ are the rows of I. You can verify that for any column ξ, $u^{i\prime}\xi$ is the ith entry of ξ and for any matrix A, $u^{i\prime}A$ is the ith row of A. This yields $I\xi = \xi$ for any column ξ and $IA = A$ for any matrix A.

A matrix A is called *invertible* if there's a matrix B such that $AB = I = BA$. Clearly, O isn't invertible because $OB = O \neq I$ for every B. Also, some nonzero matrices aren't invertible: for example, for every B,

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ b_{21} & b_{22} \end{bmatrix} \neq I.$$

When there is a B such that $AB = I = BA$, it's unique: if also $AC = I = CA$, then $B = BI = B(AC) = (BA)C = IC = C$. Thus an invertible matrix A has a unique *inverse* A^{-1} such that

$$AA^{-1} = I = A^{-1}A.$$

Clearly, I is invertible and $I^{-1} = I$.

The inverse and transpose of an invertible matrix are invertible and

$$(A^{-1})^{-1} = A \qquad (A')^{-1} = (A^{-1})'.$$

Proof: the first result follows from the equation $AA^{-1} = I = A^{-1}A$, and the second, from $A'(A^{-1})' = (A^{-1}A)' = I' = I$ and $(A^{-1})'A' = (AA^{-1})' = I' = I$.

Any product of invertible matrices is invertible:

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

Proof: $(AB)(B^{-1}A^{-1}) = ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = (AI)A^{-1} = AA^{-1} = I$. Similarly, $(B^{-1}A^{-1})(AB) = I$.

Matrix inversion is closely related to the solution of linear systems. If A is invertible, then a system $A\xi = \beta$ has exactly one solution $\xi = A^{-1}\beta$. This means that the mapping $\xi \to A\xi$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ is bijective. *Proof*: $A^{-1}\beta$ is a solution because $A(A^{-1}\beta) = (AA^{-1})\beta = I\beta = \beta$, and if ξ is any solution then $\xi = I\xi = (A^{-1}A)\xi = A^{-1}(A\xi) = A^{-1}\beta$.

The converse of this last proposition is also true: A is invertible if the system $A\xi = \beta$ has a solution for every vector β. The proof is postponed until the end of these notes.

The *determinant* of a matrix A is

$$\det A = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Clearly,

$$\det O = 0 \qquad \det(cA) = c^2 \det A \qquad \det A' = \det A$$
$$\det I = 1 \qquad \det(-A) = \det A.$$

Another such rule, critically important, is very tedious to check:

$$\det AB = \det A \det B.$$

Proof:

$$\begin{aligned}
\det AB &= \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \\
&= \det \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} = \\
&= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) \\
&= \quad (a_{11}a_{21}b_{11}b_{12} + a_{11}a_{22}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21} + a_{12}a_{22}b_{21}b_{22}) \\
&\quad - (a_{11}a_{21}b_{11}b_{12} + a_{11}a_{22}b_{12}b_{21} + a_{12}a_{21}b_{11}b_{22} + a_{12}a_{22}b_{21}b_{22})
\end{aligned}$$

$$
\begin{aligned}
&= (a_{11}a_{22}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21}) - (a_{11}a_{22}b_{12}b_{21} + a_{12}a_{21}b_{11}b_{22}) \\
&= (a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21}) \\
&= (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}) \\
&= \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \det\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \det A \det B.
\end{aligned}
$$

From this *product rule* you can deduce that if A is invertible, then $\det A \neq 0$ and

$$\det A^{-1} = (\det A)^{-1}.$$

Proof: $1 = \det I = \det AA^{-1} = \det A \det A^{-1}$.

Some strands in the theory of inverses, linear systems, and determinants still need to be tied together. First, verify the equation

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (\det A)I.$$

This shows that if $\det A \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

It follows that the single equation $AB = I$ implies that A is invertible and $B = A^{-1}$. *Proof*: $AB = I$ implies $\det A \det B = \det AB = \det I = 1$, hence $\det A \neq 0$ and A is invertible, hence $A^{-1} = A^{-1}I = A^{-1}(AB) = (A^{-1}A)B = IB = B$.

The single equation $BA = I$ also implies that A is invertible and $A^{-1} = B$. *Proof*: $A'B' = (BA)' = I' = I$, hence A' is invertible by the previous paragraph, and $(A')^{-1} = B'$. But then $A = A''$ is invertible and $A^{-1} = (A'')^{-1} = ((A')^{-1})' = B'' = B$.

Now it's possible to give the proof about linear systems that was postponed: if the system $A\xi = \beta$ has a solution for every vector β, then A is invertible. *Proof*: Construct a matrix B from columns obtained by solving linear systems as follows:

$$A\begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad A\begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then

$$AB = A\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

and the result follows. Note that this proof gives a method for computing the inverse that doesn't depend on the determinant.

These notes have given a fairly complete treatment of elementary linear algebra in two dimensions. It should be apparent to you that the notions of column vector, row, and matrix can be extended to higher dimensions. The definitions of vector and matrix addition and multiplication, and matrix inverse hold in that more general context, as well as the connection between matrix inversion and linear systems. There's one major obstacle in extending the entire theory to higher dimensions, however: the determinant. The determinant formula used in these notes holds only for two dimensions. A more complicated formula for 3×3 determinants is introduced in elementary algebra courses. From that it's possible, but extremely tedious, to verify the product rule; you can also devise a formula for the matrix inverse in terms of the determinant and complete the theory exactly as in the two dimensional case.

The theory of linear systems of arbitrary size has applications in nearly every area of mathematics. Thus there's considerable reason to generalize linear algebra beyond three dimensions. To follow exactly the route taken in these notes, you'd have to develop the theory of the determinant in higher dimensions—a major task. While that's an important mathematical concept, it seems inappropriate to spend such an effort on a tool that's only used once in the theory of linear systems. It's possible to avoid this inefficiency by analyzing in extreme detail the standard method for computing solutions of large linear systems: Gauss elimination. You can develop the connection between linear systems and matrix inversion completely, as in the two dimensional case, but without mentioning determinants. When you do study determinant theory, you find that the standard method for calculating higher dimensional determinants also uses Gauss elimination.

The theories of determinants and linear systems were developed in considerable detail during the period 1650–1850, but in an *ad hoc* fashion, only as required for applications in other areas of mathematics. The unification in terms of matrix algebra first appeared in papers of Arthur Cayley in the late 1850s.