VECTORS AND MATRICES

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. Linear algebra applies to both real and complex numbers. When you read this, 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, none appear at the level covered here. Scalars will be denoted by small Latin letters $a, b, c, \ldots$.

An n-tuple —pair, triple, quadruple, …—of scalars can be written 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

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

A column or row of length 1 is the same as a scalar.

You can *add* two vectors of the same length:

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

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

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

Every vector has a *negative*:

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

Clearly,

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

You can *multiply a vector by a scalar*:

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

Verify the following manipulative rules:

$$1\xi = \xi \qquad 0\xi = o \qquad (-a)\xi = -(a\xi) = a(-\xi)$$
$$(-1)\xi = -\xi \qquad ao = o \qquad a(b\xi) = (ab)\xi \qquad \text{—associative law}$$
$$(a + b)\xi = a\xi + b\xi \qquad \text{—distributive laws}$$
$$a(\xi + \eta) = a\xi + a\eta\,.$$

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

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

You can *multiply a row by a vector* of the same length:

$$\xi'\eta = [x_1\, x_2 \ldots x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n .$$

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

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

$$(\xi' + \eta')\zeta = \xi'\zeta + \eta'\zeta$$ —*distributive laws*
$$\xi'(\eta + \zeta) = \xi'\eta + \xi'\zeta.$$

An $m \times n$ *matrix* is an array of scalars with m rows and n columns. Matrices are denoted by large Latin letters $A, B, C, \ldots,$ and their entries, by the corresponding small letters with subscripts:

$$A = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{n \text{ columns.}} \left.\vphantom{\begin{bmatrix} a \\ a \\ \vdots \\ a \end{bmatrix}}\right] m \text{ rows}$$

A $1 \times n$ matrix is a row of length n, an $m \times 1$ matrix is a column of length m, and a 1×1 matrix is a scalar.

You can *add* $m \times n$ matrices in the obvious way. The $m \times n$ matrix O whose entries are all zeroes is called the *zero matrix*. You can also define the *negative* of a matrix, and the *product of a matrix by a scalar*. 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 an* $m \times n$ *matrix* A *by a vector* ξ *of length* n; the product $A\xi$ is the vector of length m whose entries are the products of the rows of A by ξ:

$$A\xi = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

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

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

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

The definition of the product of a matrix by a column was motivated by the appearance of a system of m linear equations in n unknowns $x_1, \ldots, x_n$: the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n \end{cases}$$

can be rewritten $A\xi = \beta$.

Similarly, you can *multiply a row* ξ' *of length* m *by an* $m \times n$ *matrix* A; the product $\xi'A$ is the row of length n whose entries are the products of ξ' by the columns of A:

$$\xi'A = [x_1\, x_2 \ldots x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$= [\, x_1 a_{11} + x_2 a_{21} + \cdots + x_m a_{m1} \quad x_1 a_{12} + x_2 a_{22} + \cdots + x_m a_{m2} \quad \cdots \quad x_1 a_{1n} + x_2 a_{2n} + \cdots + x_m a_{mn} \,].$$

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

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

You can *multiply an* $l \times m$ *matrix* A *by an* $m \times n$ *matrix* B. The product AB is an $l \times n$ 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:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + \cdots + a_{1m}b_{mn} & \cdots & a_{11}b_{1n} + \cdots + a_{1m}b_{mn} \\ \vdots & & \vdots \\ a_{l1}b_{11} + \cdots + a_{lm}b_{mn} & \cdots & a_{l1}b_{1n} + \cdots + a_{lm}b_{mn} \end{bmatrix}.$$

The ikth entry of AB is $a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{im}b_{mk}$, the product of the ith row of A by the kth column of B. You can check these manipulation rules:

$$AO = O = OB \qquad (aA)C = a(AC) = A(aC)$$
$$(-A)C = -(AC) = A(-C)$$

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

The definition of the product of two matrices was motivated by the formulas for linear substitution: from

$$\begin{cases} z_1 = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1m}y_m \\ z_2 = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2m}y_m \\ \quad\vdots \\ z_l = a_{l1}\,y_1 + a_{l2}y_2 + \cdots + a_{lm}x_m \end{cases} \qquad \begin{cases} y_1 = b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ y_2 = b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \quad\vdots \\ y_m = b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{cases}$$

you can derive

$$\begin{cases} z_1 = (a_{11}b_{11} + \cdots + a_{1m}b_{mn})x_1 + \cdots + (a_{11}b_{1n} + \cdots + a_{1m}b_{mn})x_n \\ \quad\vdots \\ z_l = (a_{l1}b_{11} + \cdots + a_{lm}b_{mn})\,x_1 + \cdots + (a_{l1}b_{1n} + \cdots + a_{lm}b_{mn})x_n \end{cases}$$

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

$$A(B\xi) = (AB)\xi$$ —*associative law.*

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

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

Proof. jth column of $A(BC)$ $= A\,(j\text{th column of } BC)$
$= A\,(B \cdot j\text{th column of } C)$
$= AB\,(j\text{th column of } C)$
$= j\text{th column of } (AB)C.$

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 $m \times n$ matrix A has a *transpose* A', the $n \times m$ matrix whose jith entry is the ijth entry of A:

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

The following manipulation rules hold:

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

The transpose of a vector is a row, and vice-versa. If A is an $l \times m$ matrix and B is an $m \times n$ matrix, then

$$(AB)' = B'A'.$$

Proof. jith entry of $(AB)'$ $=$ ijth entry of AB
$= (i\text{th row of } A)(j\text{th column of } B)$
$= (j\text{th column of } B)'(i\text{th row of } A)'$
$= (j\text{th row of } B)(i\text{th column of } A)$
$= ji\text{th entry of } B'A'.$

Consider vectors of a given length n. The jth *unit* vector is the vector u^j whose entries are all 0 except the jth, which is 1. For any row ξ', $\xi' u^j$ is the jth entry of ξ'. For any $m \times n$ matrix A, Au^j is the jth column of A. For example,

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

$$Au^1 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

The $n \times n$ matrix I whose jth column is the jth unit vector is called the *identity* matrix. Its only nonzero entries are the diagonal entries 1. Clearly, $I' = I$. For any $m \times n$ matrix A, $AI = A$. *Proof.* jth column of $AI = A$(jth column of I) $= Au^j =$ jth column of A. In particular, for any row ξ' of length n, $\xi' I = \xi'$.

Similarly, you may consider rows of a given length m. The *unit* rows $u^{i\prime}$ are the rows of the $m \times m$ identity matrix I. Verify that for any column ξ, $u^{i\prime}\xi$ is the ith entry of ξ. For any $m \times n$ matrix A, $u^{i\prime}A$ is the ith row of A. This yields $IA = A$ for any $m \times n$ matrix A. In particular, $I\xi = \xi$ for any column ξ of length m.

A matrix A is called *invertible* if there's a matrix B such that $AB = I = BA$. Clearly, invertible matrices must be square. O isn't invertible because $OB = O \neq I$ for every B. Also, some nonzero square 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 exists 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 $AA^{-1} = I = A^{-1}A$; 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 $\mathbb{R}^n$ to $\mathbb{R}^n$ is bijective. *Proof.* $A^{-1}\beta$ is a solution because $A(A^{-1}\beta) = (AA^{-1})\beta = I\beta = \beta$. 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. To see that, construct matrix B from columns obtained by solving the following linear systems:

$$A\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = u^j.$$

Then the jth column of AB is the product u^j of A and the jth column of B. That is, $AB = I$. If also $BA = I$, then A is invertible and $A^{-1} = B$. This argument gives a method for computing the inverse.

The gap in the previous argument—showing that the equation $AB = I$ implies $BA = I$ —isn't easy to fill. The most direct approach is to analyze in extreme detail the standard method for computing solutions to large linear systems: Gauss elimination. That requires another set of notes.

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

The theory of linear systems was 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.