4

Elementary Matrix Theory

We will be using matrices:

1. For describing "change of basis" formulas.

2. For describing how to compute the result of any given linear operation acting on a "vector" (e.g., finding the components with respect to some basis of the force acting on an object having some velocity.)

These are two completely different things. Do not confuse them even though the same computational apparatus (i.e., matrices) is used for both. For example, if you confuse "rotating a vector" with "using a basis constructed by rotating the original basis", you are likely to discover that your computations have everything spinning backwards.

Throughout this set of notes, K , L , M and N are positive integers.

4.1 Basics

Our Basic Notation

A matrix $\mathbf{A}$ of size $M \times N$ is simply a rectangular array with M rows and N columns of "things",

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \cdots & a_{MN} \end{bmatrix} \quad .$$

As indicated, I will try to use "bold face, upper case letters" to denote matrices. We will use two notations for the $(i, j)^{\text{th}}$ entry of $\mathbf{A}$ (i.e., the "thing" in the i^{th} row and j^{th} column):

$$(i, j)^{\text{th}} \text{ entry of } \mathbf{A} = a_{ij} = [\mathbf{A}]_{ij}$$

Until further notice, assume the "thing" in each entry is a scalar (i.e., a real or complex number). Later, we'll use such "things" as functions, operators, and even other vectors and matrices.

The matrix $\mathbf{A}$ is a row matrix if and only if it consists of just one row, and $\mathbf{B}$ is a column matrix if and only if it consists of just one column. In such cases we will normally simplify the indices in the obvious way,

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_N \end{bmatrix} \quad .$$

Basic Algebra

Presumably, you are already acquainted with matrix equality, addition and multiplication. So all we'll do here is express those concepts using our $[\cdot]_{ij}$ notation:

Matrix Equality Two matrices $\mathbf{A}$ and $\mathbf{B}$ are equal (and we write $\mathbf{A} = \mathbf{B}$) if and only if both of the following hold:

(a) $\mathbf{A}$ is the same size as $\mathbf{B}$.

(b) Letting $M \times N$ be the size of $\mathbf{A}$ and $\mathbf{B}$,

$$[\mathbf{A}]_{jk} = [\mathbf{B}]_{jk} \qquad \text{for} \quad j = 1 \ldots M \quad \text{and} \quad k = 1 \ldots N \quad .$$

Matrix Addition and Scalar Multiplication Assuming $\mathbf{A}$ and $\mathbf{B}$ are both $M \times N$ matrices, and α and β are scalars, then $\alpha\mathbf{A} + \beta\mathbf{B}$ is the $M \times N$ matrix with entries

$$[\alpha\mathbf{A} + \beta\mathbf{B}]_{jk} = \alpha[\mathbf{A}]_{jk} + \beta[\mathbf{B}]_{jk} \qquad \text{for} \quad j = 1 \ldots M \quad \text{and} \quad k = 1 \ldots N \quad .$$

Matrix Multiplication Assuming $\mathbf{A}$ is a $L \times M$ matrix and $\mathbf{B}$ is a $M \times N$ matrix, their product $\mathbf{AB}$ is the $L \times N$ matrix with entries

$$\begin{aligned} [\mathbf{AB}]_{jk} &= \text{“} j^{\text{th}} \text{ row of } \mathbf{A} \text{ times } k^{\text{th}} \text{ column of } \mathbf{B}\text{”} \\ &= \begin{bmatrix} a_{j1} & a_{j2} & a_{j3} & \cdots & a_{jM} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ b_{3k} \\ \vdots \\ b_{Mk} \end{bmatrix} \\ &= a_{j1}b_{1k} + a_{j2}b_{2k} + a_{j3}b_{3k} + \cdots + a_{jM}b_{Mk} \\ &= \sum_{m=1}^{M} a_{jm}b_{mk} \\ &= \sum_{m=1}^{M} [\mathbf{A}]_{jm}[\mathbf{B}]_{mk} \quad . \end{aligned}$$

If $\mathbf{A}$ is a row matrix, then so is $\mathbf{AB}$ and the above formula reduces to

$$[\mathbf{AB}]_k = \sum_{m=1}^{M} [\mathbf{A}]_m [\mathbf{B}]_{mk} \quad .$$

If, instead, $\mathbf{B}$ is a column matrix, then so is $\mathbf{AB}$ and the above formula reduces to

$$[\mathbf{AB}]_j = \sum_{m=1}^{M} [\mathbf{A}]_{jm} [\mathbf{B}]_m \quad .$$

Do remember that matrix multiplication is not commutative in general, but is associative. That is, except for a few special cases,

$$\mathbf{AB} \neq \mathbf{BA} \quad ,$$

but, as long as the sizes are appropriate, we can assume

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \quad .$$

Also, you can*not* assume that if the product $\mathbf{AB}$ is a matrix of just zeros, then either $\mathbf{A}$ or $\mathbf{B}$ must all be zeros.

?►*Exercise 4.1* ***a:*** *Give an example where* $\mathbf{AB} \neq \mathbf{BA}$.

b: *Give an example of two nonzero matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *for which the product* $\mathbf{AB}$ *is a matrix of just zeros. (*$\mathbf{A}$ *and* $\mathbf{B}$ *can have some zeros, just make sure that at least one entry in each is nonzero.)*

Inner Products of Matrices

Using the above definitions for addition and scalar multiplication, the set of all $M \times N$ matrices form a vector space of dimension MN. The default inner product for this vector space is the natural extension of the inner product for $\mathbb{C}^N$,

$$\langle \mathbf{A} \mid \mathbf{B} \rangle = \sum_{j=1}^{M} \sum_{k=1}^{N} a_{jk}{}^* b_{jk} = \sum_{j=1}^{M} \sum_{k=1}^{N} [\mathbf{A}]_{jk}{}^* [\mathbf{B}]_{jk} \quad .$$

Note that, if $\mathbf{A}$ and $\mathbf{B}$ are both row matrices

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_N \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_N \end{bmatrix}$$

or are both column matrices

$$\mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_M \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_M \end{bmatrix}$$

then

$$\langle \mathbf{A} \mid \mathbf{B} \rangle = \sum_j a_j{}^* b_j = \sum_j [\mathbf{A}]_j{}^* [\mathbf{B}]_j \quad .$$

Conjugates, Transposes and Adjoints

In the following, assume $\mathbf{A}$ is a $M \times N$ matrix and $\mathbf{B}$ is a $N \times L$ matrix.

Complex Conjugates

The complex conjugate of $\mathbf{A}$, denoted by $\mathbf{A}^*$ is simply the matrix obtained taking the complex conjugate of each entry,

$$\left[\mathbf{A}^*\right]_{mn} = \left([\mathbf{A}]_{mn}\right)^* \quad .$$

!►*Example 4.1:*

$$\textit{If} \quad \mathbf{A} = \begin{bmatrix} 2+3i & 5 & 7-2i \\ 6i & 4+4i & 8 \end{bmatrix} \quad \textit{then} \quad \mathbf{A}^* = \begin{bmatrix} 2-3i & 5 & 7+2i \\ -6i & 4-4i & 8 \end{bmatrix} \quad .$$

We will say $\mathbf{A}$ is *real* if and only if all the entries of $\mathbf{A}$ are real, and we will say $\mathbf{A}$ is *imaginary* if and only if all the entries of $\mathbf{A}$ are imaginary.

The following are easily verified, if not obvious:

1. $(\mathbf{AB})^* = \mathbf{A}^*\mathbf{B}^*$.
2. $(\mathbf{A}^*)^* = \mathbf{A}$.
3. $\mathbf{A}$ is real $\iff$ $\mathbf{A}^* = \mathbf{A}$.
4. $\mathbf{A}$ is imaginary $\iff$ $\mathbf{A}^* = -\mathbf{A}$.

Transposes

The transpose of the $M \times N$ matrix $\mathbf{A}$, denoted in these notes[1] by $\mathbf{A}^{\mathsf{T}}$, is the $N \times M$ matrix whose rows are the columns of $\mathbf{A}$ (or, equivalently, whose columns are the rows of $\mathbf{A}$),

$$\left[\mathbf{A}^{\mathsf{T}}\right]_{mn} = [\mathbf{A}]_{nm} \quad .$$

!►*Example 4.2:*

$$\textit{If} \quad \mathbf{A} = \begin{bmatrix} 2+3i & 5 & 7-2i \\ 6i & 4+4i & 8 \end{bmatrix} \quad \textit{then} \quad \mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 2+3i & 6i \\ 5 & 4+4i \\ 7-2i & 8 \end{bmatrix} \quad .$$

Also,

$$\textit{If} \quad |\mathbf{a}\rangle = \begin{bmatrix} 1+2i \\ 3i \\ 5 \end{bmatrix} \quad \textit{then} \quad |\mathbf{a}\rangle^{\mathsf{T}} = \begin{bmatrix} 1+2i & 3i & 5 \end{bmatrix} \quad .$$

---

[1] but Arfken, Weber & Harris use $\widetilde{\mathbf{A}}$

If you just think about what "transposing a transpose" means, it should be pretty obvious that

$$\left(\mathbf{A}^{\mathsf{T}}\right)^{\mathsf{T}} = \mathbf{A} \quad .$$

If you think a little bit about the role of rows and columns in matrix multiplication, then you may not be surprised by the fact that

$$(\mathbf{AB})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} \quad .$$

This is easily proven using our $[\cdot]_{jk}$ notation:

$$\begin{aligned}\left[(\mathbf{AB})^{\mathsf{T}}\right]_{jk} = [\mathbf{AB}]_{kj} &= \sum_{m=1}^{M} [\mathbf{A}]_{km}[\mathbf{B}]_{mj} \\ &= \sum_{m=1}^{M} \left[\mathbf{A}^{\mathsf{T}}\right]_{mk} \left[\mathbf{B}^{\mathsf{T}}\right]_{jm} \\ &= \sum_{m=1}^{M} \left[\mathbf{B}^{\mathsf{T}}\right]_{jm} \left[\mathbf{A}^{\mathsf{T}}\right]_{mk} = \left[\mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\right]_{jk} \quad ,\end{aligned}$$

showing that

$$\left[(\mathbf{AB})^{\mathsf{T}}\right]_{jk} = \left[\mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\right]_{jk} \qquad \text{for every} \quad (j,k) \quad .$$

Hence, $(\mathbf{AB})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}$ as claimed above.

Two more terms you may recall:

$$\mathbf{A} \text{ is } \textit{symmetric} \quad \Longleftrightarrow \quad \mathbf{A}^{\mathsf{T}} = \mathbf{A} \quad \Longleftrightarrow \quad [\mathbf{A}]_{jk} = [\mathbf{A}]_{kj} \quad .$$

And

$$\mathbf{A} \text{ is } \textit{antisymmetric} \quad \Longleftrightarrow \quad \mathbf{A}^{\mathsf{T}} = -\mathbf{A} \quad \Longleftrightarrow \quad [\mathbf{A}]_{jk} = -[\mathbf{A}]_{kj} \quad .$$

Note that, to be symmetric or antisymmetric, the matrix must have the same number of rows as it has columns (i.e., it must be "square").

You may recall from your undergraduate linear algebra days that there is a nice theory concerning the eigenvalues and eigenvectors of real symmetric matrices. We'll extend that theory later.

Adjoints

Combining the transpose with complex conjugation yields the "adjoint".[2]

The *adjoint* of the $M \times N$ matrix $\mathbf{A}$, denoted by either $\operatorname{adj}(\mathbf{A})$ or $\mathbf{A}^{\dagger}$, is the $N \times M$ matrix

$$\mathbf{A}^{\dagger} = \left(\mathbf{A}^{\mathsf{T}}\right)^{*} = \left(\mathbf{A}^{*}\right)^{\mathsf{T}} \quad .$$

Using the $[\cdot]_{jk}$ notation,

$$\left[\mathbf{A}^{\dagger}\right]_{jk} = \left([\mathbf{A}]_{kj}\right)^{*} \quad .$$

[2] The adjoint discussed here is sometimes called the *operator adjoint*. You may also find reference to the "classical" adjoint, which is something totally different, and is of no interest to us at all.

!►*Example 4.3:*

$$\textit{If}\quad \mathbf{A} = \begin{bmatrix} 2+3i & 5 & 7-2i \\ 6i & 4+4i & 8 \end{bmatrix} \quad \textit{then}\quad \mathbf{A}^\dagger = \begin{bmatrix} 2-3i & -6i \\ 5 & 4-4i \\ 7+2i & 8 \end{bmatrix} \quad .$$

Also,

$$\textit{If}\quad |\mathbf{a}\rangle = \begin{bmatrix} 1+2i \\ 3i \\ 5 \end{bmatrix} \quad \textit{then}\quad |\mathbf{a}\rangle^\dagger = \begin{bmatrix} 1-2i & -3i & 5 \end{bmatrix} \quad .$$

From what we've already discussed regarding complex conjugates and transposes, we immediately get the following facts:

1. $\left(\mathbf{A}^\dagger\right)^\dagger = \mathbf{A}$.

2. $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$.

3. $\mathbf{A}^\dagger = \mathbf{A}^\mathsf{T} \iff \mathbf{A}$ is real .

4. $\mathbf{A}^\dagger = -\mathbf{A}^\mathsf{T} \iff \mathbf{A}$ is imaginary .

Any matrix that satisfies $\mathbf{A}^\dagger = \mathbf{A}$ is said to be either *self adjoint* or *Hermitian*, depending on the mood of the speaker. Such matrices are the complex analogs of symmetric real matrices and will be of great interest later.

Along the same lines, any matrix that satisfies $\mathbf{A}^\dagger = -\mathbf{A}$ is said to be *antiHermitian*. I suppose you could also call them "anti-self adjoint" ("self anti-adjoint"?) though that is not commonly done.

4.2 Extending the "Bra – Ket" Notation Using an Orthonormal Basis

Assume we have an N-dimensional vector space $\mathcal{V}$ with orthonormal basis

$$\mathcal{B} = \{\, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N \,\} \quad .$$

Remember, for any vector $\mathbf{v}$ in this space,

$$|\mathbf{v}\rangle = |\mathbf{v}\rangle_{\mathcal{B}} = \text{column matrix of components of } \mathbf{v} \text{ w.r.t. } \mathcal{B}$$

$$= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad \text{where}\quad \mathbf{v} = \sum_{j=1}^{N} v_j \mathbf{e}_j \quad .$$

We now define $\langle\mathbf{v}|$ to be the adjoint of $|\mathbf{v}\rangle$,

$$\langle\mathbf{v}| = \langle\mathbf{v}|_{\mathcal{B}} = |\mathbf{v}\rangle^\dagger = \begin{bmatrix} {v_1}^* & {v_2}^* & \cdots & {v_N}^* \end{bmatrix} \quad .$$

Observe that

$$\big(\mathbf{A}\,|\mathbf{v}\rangle\big)^{\dagger} \;=\; \langle\mathbf{v}|\,\mathbf{A}^{\dagger} \quad .$$

Also note that

$$\langle\mathbf{v}|\,|\mathbf{w}\rangle \;=\; \begin{bmatrix} {v_1}^* & {v_2}^* & \cdots & {v_N}^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \;=\; \sum_{j=1}^{N} {v_j}^* w_j$$

which, since we are assuming an orthonormal basis, reduces to

$$\underbrace{\langle\mathbf{v}|\,|\mathbf{w}\rangle}_{\substack{\text{matrix}\\ \text{product}}} \;=\; \underbrace{\langle\, \mathbf{v} \mid \mathbf{w} \,\rangle}_{\substack{\text{vector}\\ \text{inner}\\ \text{product}}} \quad .$$

The " $\langle\cdot|$ " is the "bra" in the (Dirac) "bra-ket" notation.

Using an Arbitrary Basis*

Now let

$$\mathcal{B} \;=\; \{\, \mathbf{b}_1,\, \mathbf{b}_2,\, \dots,\, \mathbf{b}_N \,\} \quad .$$

be any basis for our vector space $\mathcal{V}$, and let

$$\mathcal{D} \;=\; \{\, \mathbf{d}_1,\, \mathbf{d}_2,\, \dots,\, \mathbf{d}_N \,\} \quad .$$

be the corresponding reciprocal basis. Recall from section 2.6 that

$$\langle\, \mathbf{v} \mid \mathbf{w} \,\rangle \;=\; \sum_{k=1}^{N} \bar{v}_k^{\,*} w_k$$

where

$$\begin{bmatrix} \bar{v}_1 \\ \bar{v}_2 \\ \vdots \\ \bar{v}_N \end{bmatrix} \;=\; |\mathbf{v}\rangle_{\mathcal{D}} \qquad \text{and} \qquad \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \;=\; |\mathbf{w}\rangle_{\mathcal{B}} \quad .$$

In terms of matrix operations, the above is

$$\langle\, \mathbf{v} \mid \mathbf{w} \,\rangle \;=\; \sum_{k=1}^{N} \bar{v}_k^{\,*} w_k \;=\; \begin{bmatrix} \bar{v}_1^{\,*} & \bar{v}_2^{\,*} & \cdots & \bar{v}_N^{\,*} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \;=\; |\mathbf{v}\rangle_{\mathcal{D}}{}^{\dagger}\; |\mathbf{w}\rangle_{\mathcal{B}} \quad .$$

Because of this, the appropriate general definition of "bra **v**" is

$$\langle\mathbf{v}| \;=\; \langle\mathbf{v}|_{\mathcal{B}} \;=\; |\mathbf{v}\rangle_{\mathcal{D}}{}^{\dagger}$$

where $\mathcal{D}$ is the reciprocal basis corresponding to $\mathcal{B}$. That way, we still have

$$\underbrace{\langle\mathbf{v}|\,|\mathbf{w}\rangle}_{\substack{\text{matrix}\\ \text{product}}} \;=\; \underbrace{\langle\, \mathbf{v} \mid \mathbf{w} \,\rangle}_{\substack{\text{vector}\\ \text{inner}\\ \text{product}}} \quad .$$

* This material is optional. It requires the "reciprocal basis" from section 2.6.

4.3 Square Matrices

For the most part, the only matrices we'll have much to do with (other than row or column matrices) are square matrices.

The Zero, Identity and Inverse Matrices

A *square matrix* is any matrix having the same number of rows as columns. Two important $N \times N$ matrices are

The *zero matrix*, denoted by $\mathbf{O}$ or $\mathbf{O}_N$, which is simply the $N \times N$ matrix whose entries are all zero,

$$[\mathbf{O}]_{jk} = 0 \qquad \text{for all} \quad (j,k) \quad .$$

The *identity matrix*, denoted by $\mathbf{I}$ or $\mathbf{I}_N$ or $\mathbf{1}$ (Afkin, Weber & Harris's notation), which is the matrix whose entries are all zero except for 1's on the major diagonal,

$$[\mathbf{I}]_{jk} = \delta_{jk} \quad .$$

Recall that, if $\mathbf{A}$ is any $N \times N$ matrix, then

$$\mathbf{A} + \mathbf{O} = \mathbf{A} \quad , \qquad \mathbf{AO} = \mathbf{OA} = \mathbf{O} \qquad \text{and} \qquad \mathbf{AI} = \mathbf{IA} = \mathbf{A} \quad .$$

!►*Example 4.4:* *The* 3×3 *zero and identity matrices are*

$$\mathbf{O} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \textit{and} \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad .$$

The *inverse* of a (square) matrix $\mathbf{A}$ is the matrix, denoted by $\mathbf{A}^{-1}$, such that

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

Not all square matrices have inverses. If $\mathbf{A}$ has an inverse, it is said to be *invertible* or *nonsingular*. If $\mathbf{A}$ does not have an inverse, we call it *noninvertible* or *singular*.

It may later be worth noting that, while the definition of the inverse requires that

$$\mathbf{AA}^{-1} = \mathbf{I} \qquad \textit{and} \qquad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad ,$$

it can easily be shown that, if $\mathbf{B}$ is any square matrix such that

$$\mathbf{AB} = \mathbf{I} \qquad \textit{or} \qquad \mathbf{BA} = \mathbf{I} \quad ,$$

then

$$\mathbf{AB} = \mathbf{I} \qquad \textit{and} \qquad \mathbf{BA} = \mathbf{I} \quad ,$$

and thus, $\mathbf{B} = \mathbf{A}^{-1}$.

Here are few other things you should recall about inverses (assume $\mathbf{A}$ is an invertible $N \times N$ matrix):

1. $\mathbf{AB} = \mathbf{C} \iff \mathbf{B} = \mathbf{A}^{-1}\mathbf{C}$.
2. $\mathbf{BA} = \mathbf{C} \iff \mathbf{B} = \mathbf{C}\mathbf{A}^{-1}$.
3. $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$.
4. If $\mathbf{B}$ is also an invertible $N \times N$ matrix, then

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

?►***Exercise 4.2:*** *Verify anything above that is not immediately obvious to you.*

4.4 Determinants
Definition, Formulas and Cramer's Rule

Recall that each $N \times N$ matrix has an associated scalar called the *determinant* of that matrix. We will denote the determinant of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$$

by

$$\det(\mathbf{A}) \qquad \text{or} \qquad \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix} \qquad \text{or} \qquad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix}$$

as seems convenient.

The determinant naturally arises when solving a system of N linear equations for N unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N &= b_2 \\ &\vdots \\ a_{N1}x_1 + a_{N2}x_2 + \cdots + a_{NN}x_N &= b_N \end{aligned} \tag{4.1a}$$

where the a_{jk}'s and b_k's are known values and the x_j's are the unknowns. Do observe that we can write this system more concisely in "matrix form",

$$\mathbf{Ax} = \mathbf{b} \tag{4.1b}$$

where $\mathbf{A}$ is the $N \times N$ matrix of coefficients, and $\mathbf{x}$ and $\mathbf{b}$ are the column matrices

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \quad .$$

In solving this, you get (trust me)

$$\det(\mathbf{A})\, x_k = \det(\mathbf{B}_k) \qquad \text{for} \quad k = 1, 2, \ldots N \tag{4.2}$$

where $\mathbf{B}_k$ is the matrix obtained by replacing the k^{th} column in $\mathbf{A}$ with $\mathbf{b}$. We can then find each x_k by dividing through by $\det(\mathbf{A})$, provided $\det(\mathbf{A}) \neq 0$. (Equation (4.2) is *Cramer's rule* for solving system (4.1a) or (4.1b), written the way it should be — but almost never is — written.)

?►*Exercise 4.3:* *What can you say about the possible values for the* x_k*'s if* $\det(\mathbf{A}) = 0$ *in the above?*

In theory, the definition of and formula for the determinant of any square matrix can be based on equation (4.2) as a solution to system (4.1a) or (4.1b). (Clever choices for the $\mathbf{b}$ vectors may help.) Instead, we will follow standard practice: I'll give you the formulas and tell you to trust me. (And if you remember how to compute determinants — as I rather expect — just skip ahead to **Properties and Applications of Determinants**.)

For $N = 1$ and $N = 2$, we have the well-known formulas

$$\det[a_{11}] = a_{11} \qquad \text{and} \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad .$$

For $N > 2$, the formula for the determinant of an $N \times N$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$$

can be given in terms of the determinants of $(N-1) \times (N-1)$ "submatrices" via either the "expansion by minors of the first row"

$$\det \mathbf{A} = \sum_{k=1}^{N} (-1)^{k-1} a_{1k} \det(\mathbf{A}_{1k})$$

or the "expansion by minors of the first column"

$$\det \mathbf{A} = \sum_{j=1}^{N} (-1)^{j-1} a_{j1} \det\big(\mathbf{A}_{j1}\big)$$

where $\mathbf{A}_{jk}$ is the $(N-1) \times (N-1)$ matrix obtained from $\mathbf{A}$ by deleting the j^{th} row and k^{th} column.

Either of the above formulas gives you the determinant of $\mathbf{A}$. If you are careful with your signs, you can find the determinant via an expansion using any row or column. If you are interested[3], here is a "most general" formula for the determinant for the above $N \times N$ matrix $\mathbf{A}$:

$$\det(\mathbf{A}) = \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \cdots \sum_{i_N=1}^{N} \varepsilon(i_1, i_2, \ldots, i_N)\, a_{i_1 1} a_{i_2 2} \cdots a_{i_N N}$$

where ε is a function of all N-tuples of the integers from 1 to N satisfying all the following:

1. If any two of the entries in $(i_1, i_2, \ldots, i_N)$ are the same, then

$$\varepsilon(i_1, i_2, \ldots, i_N) = 0 \quad .$$

2. If $i_1 = 1$, $i_2 = 2$, $i_3 = 3$, etc., then

$$\varepsilon(i_1, i_2, i_3 \ldots, i_N) = \varepsilon(1, 2, 3, \ldots, N) = 1 \quad .$$

3. Suppose $(j_1, j_2, \ldots, j_N)$ is obtained from $(i_1, i_2, \ldots, i_N)$ by simply switching just two adjacent entries in $(i_1, i_2, \ldots, i_N)$. In other words, for some integer K with $1 \leq K < N$,

$$j_k = \begin{cases} i_k & \text{if} \quad k \neq K \text{ or } k \neq K+1 \\ i_{K+1} & \text{if} \quad k = K \\ i_K & \text{if} \quad k = K+1 \end{cases} \quad .$$

Then

$$\varepsilon(j_1, j_2, j_3 \ldots, j_N) = -\varepsilon(i_1, i_2, i_3 \ldots, i_N) \quad .$$

Properties and Applications of Determinants

Let me remind you of many of the properties and a few of the applications of determinants.

In many ways, you can consider Cramer's rule (equation (4.2)) for solving system (4.1a) or, equivalently, matrix equation (4.1b) as a basic defining equation for the determinant. Cramer's rule tells us that system (4.1a) can be solved for every choice of the b_j's if and only if $\det(\mathbf{A}) \neq 0$. But the ability to solve that system is equivalent to the ability to solve the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

for every choice of $\mathbf{b}$, and since this matrix equation is solvable for every choice of $\mathbf{b}$ if and only if $\mathbf{A}$ is invertible, we must have that:

$$\text{An } N \times N \text{ matrix } \mathbf{A} \text{ is invertible} \quad \Longleftrightarrow \quad \det(\mathbf{A}) \neq 0 \quad .$$

This is doubtlessly our favorite test for determining if $\mathbf{A}^{-1}$ exists. It will also probably be the most important fact about determinants that we will use.

Another property that can be derived from Cramer's rule is that, if $\mathbf{A}$ and $\mathbf{B}$ are both $N \times N$ matrices, then

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}) \quad . \tag{4.3}$$

[3] If you are not interested, skip to **Properties and Applications of Determinants** below.

(Just consider using Cramer's rule to solve $(\mathbf{AB})\mathbf{x} = \mathbf{b}$, rewritten as $\mathbf{A}(\mathbf{Bx}) = \mathbf{b}$.)

Now it is real easy to use the formulas (or even Cramer's rule) to verify that

$$\det(\mathbf{I}) = \det \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = 1 \quad .$$

From this and the observation regarding determinants of products you can easily do the next exercise.

?►*Exercise 4.4:* *Assume* $\mathbf{A}$ *is invertible, and show that*

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})} \quad .$$

A few other properties of the determinant are listed here. Assume $\mathbf{A}$ is an $N \times N$ matrix and α is some scalar.

1. $\det(\alpha\mathbf{A}) = \alpha^N \det(\mathbf{A})$.
2. $\det\left(\mathbf{A}^*\right) = \det(\mathbf{A})^*$.
3. $\det\left(\mathbf{A}^\mathsf{T}\right) = \det(\mathbf{A})$.
4. $\det\left(\mathbf{A}^\dagger\right) = \det(\mathbf{A})^*$.

?►*Exercise 4.5:* *Convince yourself of the validity of each of the above statements. (For the first,* $\det(\alpha\mathbf{A}) = \alpha^N \det(\mathbf{A})$*, you might first show that* $\det(\alpha\mathbf{I}) = \alpha^N$*, and then consider rewriting* $\alpha\mathbf{A}$ *as* $\alpha\mathbf{IA}$ *and taking the determinant of that.)*

Also remind yourself of the computational facts regarding determinants listed on page 87 of Arfken, Weber & Harris.

Finally, let me comment on the importance of Cramer's rule: It can be viewed as being very important, theoretically, for the way it relates determinants to to solving of linear systems. However, as a practical tool for solving these systems, it is quite *un*important. The number of computations required to find all the determinants is horrendous, especially if the system is large. Other methods, such as the Gauss elimination/row reduction learned in your undergraduate linear algebra course are much faster and easier to use.