III. Matrices

Definition III.1 An $m \times n$ matrix is a set of numbers arranged in a rectangular array having m rows and n columns. It is written

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

There are two important special cases. A $1 \times n$ matrix (that is, a matrix with 1 row) is called a row vector. An $m \times 1$ matrix (that is, a matrix with 1 column) is called a column vector. Our convention will be that row indices are always written before column indices. As a memory aid, I think of matrices as being RC (Roman Catholic or rows before columns).

§III.1. Matrix Operations

Definitions

1. Equality. For any two matrices A and B

$$A = B \iff \text{(a) } A \text{ and } B \text{ have the same number of rows and the same number of columns and} \\ \text{(b) } A_{ij} = B_{ij} \text{ for all } i, j$$

2. Addition. For any two $m \times n$ matrices A and B

$$(A + B)_{ij} = A_{ij} + B_{ij} \text{ for all } 1 \le i \le m,\ 1 \le j \le n$$

That is, the entry in row i, column j of the matrix $A + B$ is defined to be the sum of the corresponding entries in A and B. **Note:** The sum $A + B$ is only defined if A and B have the same number of rows and the same number of columns.

3. Scalar multiplication. For any number s and any $m \times n$ matrix A

$$(sA)_{ij} = sA_{ij} \text{ for all } 1 \le i \le m,\ 1 \le j \le n$$

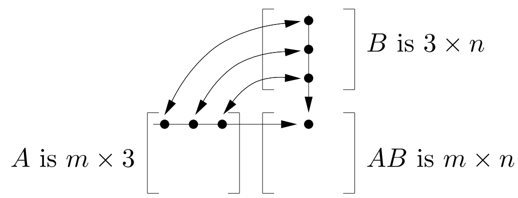
For example

$$2 \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 \times 1 + 0 & 2 \times 2 + 1 \\ 2 \times 0 + 1 & 2 \times 3 + 1 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix}$$

4. Matrix multiplication. For any $m \times p$ matrix A and any $p \times n$ matrix B

$$(AB)_{ik} = \sum_{j=1}^{p} A_{ij} B_{jk} \text{ for all } 1 \le i \le m,\ 1 \le k \le n$$

Note (a) AB is only defined if the number of columns of A is the same as the number of rows of B. If A is $m \times p$ and B is $p \times n$, then AB is $m \times n$.
(b) $(AB)_{ik}$ is the dot product of the $i^{\rm th}$ row of A (viewed as a row vector) and the $k^{\rm th}$ column of B
(c) Here is a memory aid. If you write the first factor of AB to the left of AB and the second factor above AB, then each entry $(AB)_{ik}$ of AB is built from the entries A_{ij} $j = 1, 2, \cdots$ of A that are



directly to its left and the entries B_{jk} $j = 1, 2, \cdots$ of B that are directly above it. These entries are multiplied in pairs, $A_{ij}B_{jk}$, $j = 1, 2, \cdots$, starting as far from AB as possible and then the products are added up to yield $(AB)_{ik} = \sum_j A_{ij}B_{jk}$.

(d) If A is a square matrix, then $AAA \cdots A$ (n factors) makes sense and is denoted A^n.

(e) At this stage we have no idea why it is useful to define matrix multiplication in this way. We'll get some first hints shortly.

Example III.2 Here is an example of the product of a 2×3 matrix with a 3×2 matrix, yielding a 2×2 matrix

$$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 0 \times 1 + 1 \times 2 + 2 \times 3 & 0 \times 3 + 1 \times 2 + 2 \times 1 \\ 3 \times 1 + 4 \times 2 + 5 \times 3 & 3 \times 3 + 4 \times 2 + 5 \times 1 \end{bmatrix} = \begin{bmatrix} 8 & 4 \\ 26 & 22 \end{bmatrix}$$

Example III.3 Here is an example of the product of a 3×3 matrix with a 3×1 matrix, yielding a 3×1 matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + 2x_2 + 3x_3 \\ 2x_1 + 3x_2 + x_3 \end{bmatrix}$$

Hence we may very compactly write the system of equations

$$\begin{aligned} x_1 + \ x_2 + \ x_3 &= 4 \\ x_1 + 2x_2 + 3x_3 &= 9 \\ 2x_1 + 3x_2 + \ x_3 &= 7 \end{aligned}$$

that we dealt with in Example II.2, as $A\vec{x} = \vec{b}$ where

$$A = \text{the matrix of coefficients} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \quad \vec{x} = \text{the column vector} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \vec{b} = \text{the column vector} \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}$$

Basic Properties of Matrix Operations.

Using A, B and C to denote matrices and s and t to denote numbers

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $s(A + B) = sA + sB$
4. $(s + t)A = sA + tA$
5. $(st)A = s(tA)$
6. $1A = A$
7. $A + 0 = A$ where 0 is the matrix all of whose entries are zero
8. $A + (-1)A = 0$ (The matrix $(-1)A$ is generally denoted $-A$.)
9. $A(B + C) = AB + AC$
10. $(A + B)C = AC + BC$
11. $A(BC) = (AB)C$
12. $s(AB) = (sA)B = A(sB)$

These properties are all almost immediate consequences of the definitions. For example to verify property 9, it suffices to write down the definitions of the two sides of property 9

$$\begin{aligned}
[A(B+C)]_{ik} &= \sum_j A_{ij}(B+C)_{jk} = \sum_j A_{ij}(B_{jk}+C_{jk}) \\
[AB+AC]_{ik} = [AB]_{ik}+[AC]_{ik} &= \sum_j A_{ij}B_{jk} + \sum_j A_{ij}C_{jk}
\end{aligned}$$

Counterintuitive Properties of Matrix Operations that Cause Numerous Errors

13. In general $AB \neq BA$. Here are three examples. In the first AB is defined and BA is not defined. In the second AB and BA are both defined, but are not of the same size. In the third AB and BA are both defined and are of the same size, but are different.

$$A = \begin{bmatrix} 1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 5 \\ 4 & 6 \end{bmatrix} \quad AB = \begin{bmatrix} 11 & 17 \end{bmatrix} \quad BA \text{ is not defined}$$

$$A = \begin{bmatrix} 1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad AB = \begin{bmatrix} 11 \end{bmatrix} \quad BA = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

A consequence of this "unproperty" is that $(A-B)(A+B)$ need not equal A^2-B^2. Multiplying out $(A-B)(A+B)$ carefully gives $AA+AB-BA-BB$. The middle two terms need not cancel. For example

$$\begin{aligned}
A &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} & B &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & A^2-B^2 &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
A-B &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} & A+B &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & (A-B)(A+B) &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

14. In general AB may be 0 even if A and B are both nonzero. For example

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \neq 0,\ B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \neq 0,\ AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0,$$

A consequence of this is that $AB = AC$ does not force $B = C$ even if every entry of A is nonzero. For example

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix},\ B = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \neq C = \begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix} \text{ and yet } AB = AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Example III.4 (The Form of the General Solution of Systems of Linear Equations Revisited) We have just seen that any system of linear equations can be written

$$A\vec{x} = \vec{b}$$

where A is the matrix of coefficients, $\vec{x}$ is the column vector of unknowns, $\vec{b}$ is the column vector of right hand sides and $A\vec{x}$ is the matrix product of A and $\vec{x}$. We also saw in §II.3 that, if the ranks of $[A]$ and $[A\,|\,\vec{b}]$ are the same, the general solution to this system is of the form

$$\vec{x} = \vec{u} + c_1\vec{v}_1 + \cdots + c_{n-\rho}\vec{v}_{n-\rho}$$

where n is the number of unknowns (that is, the number of components of $\vec{x}$, or equivalently, the number of columns of A), ρ is the rank of A and $c_1,\ \cdots,\ c_{n-\rho}$ are arbitrary constants. That $\vec{u}+c_1\vec{v}_1+\cdots+c_{n-\rho}\vec{v}_{n-\rho}$ is a solution of $A\vec{x}=\vec{b}$ for all values of $c_1,\ \cdots,\ c_{n-\rho}$ means that

$$A(\vec{u}+c_1\vec{v}_1+\cdots+c_{n-\rho}\vec{v}_{n-\rho})=\vec{b}$$

for all values of $c_1,\ \cdots,\ c_{n-\rho}$. By properties 9 and 12 of matrix operations, this implies that

$$A\vec{u}+c_1A\vec{v}_1+\cdots+c_{n-\rho}A\vec{v}_{n-\rho}=\vec{b}$$

But this is true for all $c_1,\ \cdots,\ c_{n-\rho}$ if and only if

$$\begin{aligned}
A\vec{u}&=\vec{b} && (\text{Set } c_1=\ \cdots\ =c_{n-\rho}=0)\\
A\vec{v}_1&=\vec{0} && (\text{Set } c_1=1,\ c_2=\ \cdots\ =c_{n-\rho}=0 \text{ and sub in } A\vec{u}=\vec{b}.)\\
&\ \vdots\\
A\vec{v}_{n-\rho}&=\vec{0} && (\text{Set } c_1=\ \cdots\ =c_{n-\rho-1}=0,\ c_{n-\rho}=1 \text{ and sub in } A\vec{u}=\vec{b}.)
\end{aligned}$$

In other words $\vec{u}$ is a solution of $A\vec{x}=\vec{b}$ and each of $\vec{v}_1,\ \cdots,\ \vec{v}_{n-\rho}$ is a solution of $A\vec{x}=\vec{0}$.

Exercises for §III.1

1) Define

$$A=\begin{bmatrix}1&2&3\\1&2&1\end{bmatrix}\quad B=\begin{bmatrix}-1&2\\-3&1\\-2&1\end{bmatrix}\quad C=[\,2\quad -2\quad 0\,]\quad D=\begin{bmatrix}2\\-11\\2\end{bmatrix}$$

Compute all pairwise products ($AA,\ AB,\ AC,\ AD, BA,\cdots$) that are defined.

2) Compute $A^2=AA$ and $A^3=AAA$ for

$$\text{a)}\quad A=\begin{bmatrix}0&a&b\\0&0&c\\0&0&0\end{bmatrix}\qquad\qquad \text{b)}\quad A=\begin{bmatrix}1&0&a\\0&1&0\\0&0&1\end{bmatrix}$$

3) Let $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$.

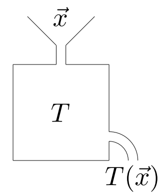
a) Find $A^2,\ A^3,\ A^4$.
b) Find A^k for all positive integers.
c) Find e^{At}. (Part of this problem is to invent a reasonable definition of e^{At}.)
d) Find a square root of A. That is, find a matrix B obeying $B^2=A$.
e) Find all square roots of A.

4) Compute A^k for $k=2,3,4$ when

$$A=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$$

§III.2 Matrices and Linear Transformations

Definition III.5 A **transformation** (a.k.a. **map**, a.k.a. **function**) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule which assigns to each vector $\vec{x}$ in $\mathbb{R}^n$ a vector $T(\vec{x})$ in $\mathbb{R}^m$. T should be thought of as a machine: if you put a



vector $\vec{x}$ into the input hopper, it spits out a new vector $T(\vec{x})$. A transformation is said to be **linear** if

$$T(s\vec{x}+t\vec{y}) = sT(\vec{x})+tT(\vec{y}) \qquad \text{for all numbers } s,\ t \text{ and vectors } \vec{x},\ \vec{y}.$$

Later in this section, we shall see that for each linear transformation $T(\vec{x})$ there is a matrix M_T such that $T(\vec{x})$ is the matrix product of M_T times the column vector $\vec{x}$. In other words, such that $T(\vec{x}) = M_T\vec{x}$. First, however, we look at a number of examples, both of transformations that are not linear and transformations that are linear.

The map

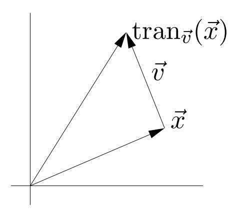
$$T([x_1,x_2]) = [0,x_2^2]$$

is not linear, because the two quantities

$$\begin{aligned} T(2\vec{x}) &= [0,(2x_2)^2] \\ 2T(\vec{x}) &= 2[0,(x_2)^2] \end{aligned}$$

are not equal whenever $x_2 \ne 0$. Another example of a map that is **not** linear is

Example III.6 (Translation) Define $\mathrm{tran}_{\vec{v}}(\vec{x})$ be the vector gotten by translating the head of the arrow $\vec{x}$ by $\vec{v}$ (while leaving the tail of the arrow fixed). In equations, $\mathrm{tran}_{\vec{v}}(\vec{x}) = \vec{x}+\vec{v}$. If translation were linear



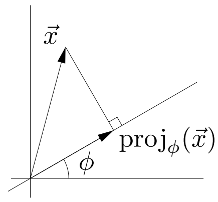
the two expressions

$$\begin{aligned} \mathrm{tran}_{\vec{v}}(s\vec{x}+t\vec{y}) &= s\vec{x}+t\vec{y}+\vec{v} \\ s\,\mathrm{tran}_{\vec{v}}(\vec{x})+t\,\mathrm{tran}_{\vec{v}}(\vec{y}) &= s(\vec{x}+\vec{v})+t(\vec{y}+\vec{v}) \\ &= s\vec{x}+t\vec{y}+(s+t)\vec{v} \end{aligned}$$

would be equal for all s and t. But if $\vec{v} \ne \vec{0}$ and $s+t \ne 1$, the two expressions are not equal.

We have just seen an example of a geometric operation that is not linear. Many other geometric operations are linear maps. As a result, linear maps play a big role in computer graphics. Here are some examples.

Example III.7 (Projection) Define $\mathrm{proj}_\phi(\vec{x})$ to be the projection of the vector $\vec{x}$ on the line in $\mathbb{R}^2$ that passes through the origin at an angle ϕ from the x–axis. The vector $\hat{b} = [\cos\phi, \sin\phi]$ is a unit vector



that lies on the line. So $\text{proj}_\phi(\vec{x})$ must have direction $[\cos\phi, \sin\phi]$ (or its negative) and must have length $\|\vec{x}\|\cos\theta = \vec{x}\cdot\hat{b}$, where θ is the angle between $\vec{x}$ and $\hat{b}$. The unique vector with the right direction and length is

$$\text{proj}_\phi(\vec{x}) = (\vec{x}\cdot\hat{b})\ \hat{b}$$

It is easy to verify that this really is a linear transformation:

$$\begin{aligned}\text{proj}_\phi(s\vec{x}+t\vec{y}) &= \big((s\vec{x}+t\vec{y})\cdot\hat{b}\big)\,\hat{b}\\ &= s(\vec{x}\cdot\hat{b})\,\hat{b} + t(\vec{y}\cdot\hat{b})\,\hat{b}\\ &= s\,\text{proj}_\phi(\vec{x})\ +\ t\,\text{proj}_\phi(\vec{y})\end{aligned}$$

Writing both $\text{proj}_\phi(\vec{x})$ and $\vec{x}$ as column vectors

$$\text{proj}_\phi(\vec{x}) = (\vec{x}\cdot\hat{b})\ \hat{b} = (x_1\cos\phi + x_2\sin\phi)\ \begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix} = \begin{bmatrix}x_1\cos^2\phi + x_2\sin\phi\cos\phi\\ x_1\sin\phi\cos\phi + x_2\sin^2\phi\end{bmatrix}$$

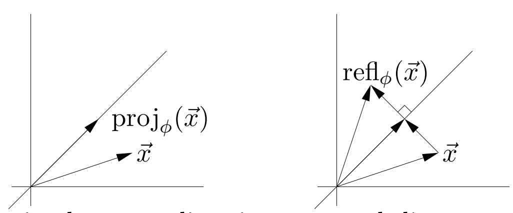
$$= \begin{bmatrix}\cos^2\phi & \sin\phi\cos\phi\\ \sin\phi\cos\phi & \sin^2\phi\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1+\cos 2\phi & \sin 2\phi\\ \sin 2\phi & 1-\cos 2\phi\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

where we have used the double angle trig identities

$$\begin{aligned}\sin(2\phi) &= 2\sin\phi\cos\phi\\ \cos(2\phi) &= \cos^2\phi - \sin^2\phi = 2\cos^2\phi - 1 = 1 - 2\sin^2\phi\end{aligned}$$

Notice that this is the matrix product of a matrix that depends only on ϕ (not on $\vec{x}$) times the column vector $\vec{x}$.

Example III.8 (Reflection) Define $\text{refl}_\phi(\vec{x})$ to be the reflection of the vector $\vec{x}$ on the line in $\mathbb{R}^2$ that passes through the origin at an angle ϕ from the x–axis. You can get from $\vec{x}$ to $\text{refl}_\phi(\vec{x})$ by first walking



from $\vec{x}$ to $\text{proj}_\phi(\vec{x})$ and continuing in the same direction an equal distance on the far side of the line. In terms of vectors, to get from $\vec{x}$ to $\text{proj}_\phi(\vec{x})$, you have to add the vector $\text{proj}_\phi(\vec{x}) - \vec{x}$ to $\vec{x}$. To continue an equal distance in the same direction, you have to add a second copy of $\text{proj}_\phi(\vec{x}) - \vec{x}$. So

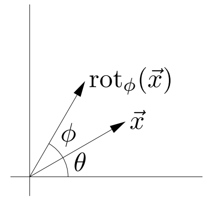
$$\text{refl}_\phi(\vec{x}) = \vec{x} + 2[\text{proj}_\phi(\vec{x}) - \vec{x}] = 2\text{proj}_\phi(\vec{x}) - \vec{x}$$

We may, once again, write this as the matrix product of a matrix that depends only on ϕ (not on $\vec{x}$) times the column vector $\vec{x}$.

$$\text{refl}_\phi(\vec{x}) = 2\frac{1}{2}\begin{bmatrix}1+\cos 2\phi & \sin 2\phi\\ \sin 2\phi & 1-\cos 2\phi\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} - \begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}1+\cos 2\phi & \sin 2\phi\\ \sin 2\phi & 1-\cos 2\phi\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} - \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

$$= \left(\begin{bmatrix}1+\cos 2\phi & \sin 2\phi\\ \sin 2\phi & 1-\cos 2\phi\end{bmatrix} - \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}\cos 2\phi & \sin 2\phi\\ \sin 2\phi & -\cos 2\phi\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

Example III.9 (Rotation) Define $\text{rot}_\phi(\vec{x})$ to be the result of rotating the vector $\vec{x}$ by an angle ϕ about the origin.



If we denote by r the length of $\vec{x}$ and by θ the angle between $\vec{x}$ and the x–axis, then

$$\vec{x} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}$$

To rotate this by ϕ we need only replace θ by $\theta + \phi$.

$$\begin{aligned}\text{rot}_\phi(\vec{x}) &= \begin{bmatrix} r\cos(\theta+\phi) \\ r\sin(\theta+\phi) \end{bmatrix} = \begin{bmatrix} r(\cos\theta\cos\phi - \sin\theta\sin\phi) \\ r(\sin\theta\cos\phi + \cos\theta\sin\phi) \end{bmatrix} = \begin{bmatrix} x_1\cos\phi - x_2\sin\phi \\ x_2\cos\phi + x_1\sin\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\end{aligned}$$

Note that in each of examples 7, 8 and 9 there was a matrix A such that the map could be written $T(\vec{x}) = A\vec{x}$, where $A\vec{x}$ is the matrix product of the matrix A times the column vector $\vec{x}$. Every map of the form $T(\vec{x}) = A\vec{x}$ is automatically linear, because $A(s\vec{x} + t\vec{y}) = s(A\vec{x}) + t(A\vec{y})$ by properties (9) and (12) of matrix operations. We shall now show that, conversely, for each linear map $T(\vec{x})$ there is a matrix A such that $T(\vec{x}) = A\vec{x}$.

First we consider the case that $T(\vec{x})$ is a linear map from $\mathbb{R}^2$ (that is, two component vectors) to $\mathbb{R}^2$. Any vector $\vec{x}$ in $\mathbb{R}^2$ can be written

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\widehat{\mathbf{e}}_1 + x_2\widehat{\mathbf{e}}_2$$

where

$$\widehat{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \widehat{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Because T is linear

$$T(\vec{x}) = T(x_1\widehat{\mathbf{e}}_1 + x_2\widehat{\mathbf{e}}_2) = x_1T(\widehat{\mathbf{e}}_1) + x_2T(\widehat{\mathbf{e}}_2)$$

Define the numbers $a_{11},\ a_{12},\ a_{21},\ a_{22}$ by

$$T(\widehat{\mathbf{e}}_1) = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \qquad T(\widehat{\mathbf{e}}_2) = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

Then

$$T(\vec{x}) = x_1T(\widehat{\mathbf{e}}_1) + x_2T(\widehat{\mathbf{e}}_2) = x_1\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

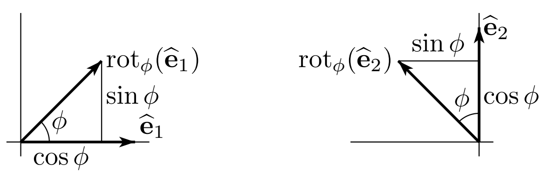
The same construction works for linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$. Define $\widehat{\mathbf{e}}_i$ to be the column vector all of whose entries are zero, except for the i^{th}, which is one. Note that every m component vector can be written

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + +x_m\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = x_1\widehat{\mathbf{e}}_1 + x_2\widehat{\mathbf{e}}_2 + \cdots x_m\widehat{\mathbf{e}}_m$$

For each i between 1 and m, $T(\hat{\mathbf{e}}_i)$ is an n–component vector. Think of this vector as a column vector. Then define the matrix $A = \big[T(\hat{\mathbf{e}}_1)\ T(\hat{\mathbf{e}}_2)\ \cdots\ T(\hat{\mathbf{e}}_m)\big]$. That is, A is the matrix whose $i^{\rm th}$ column is $T(\hat{\mathbf{e}}_i)$. This is the matrix we want, because

$$\begin{aligned}
A\vec{x} &= \big[T(\hat{\mathbf{e}}_1)\ T(\hat{\mathbf{e}}_2)\ \cdots\ T(\hat{\mathbf{e}}_m)\big]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m\end{bmatrix} \\
&= x_1T(\hat{\mathbf{e}}_1) + x_2T(\hat{\mathbf{e}}_2) + \cdots + x_mT(\hat{\mathbf{e}}_m) && \text{(by the definition of matrix multiplication)} \\
&= T(x_1\hat{\mathbf{e}}_1 + x_2\hat{\mathbf{e}}_2 + \cdots + x_m\hat{\mathbf{e}}_m) && \text{(by the linearity of } T) \\
&= T(\vec{x})
\end{aligned}$$

Example III.10 Let, as in Example III.9, rot_ϕ be the linear transformation which rotates vectors in the plane by ϕ. From the figure



we see that

$$\mathrm{rot}_\phi(\hat{\mathbf{e}}_1) = \begin{bmatrix}\cos\phi \\ \sin\phi\end{bmatrix} \qquad \mathrm{rot}_\phi(\hat{\mathbf{e}}_2) = \begin{bmatrix}-\sin\phi \\ \cos\phi\end{bmatrix}$$

Hence the matrix which implements rotatiom is

$$\big[\mathrm{rot}_\phi(\hat{\mathbf{e}}_1)\ \mathrm{rot}_\phi(\hat{\mathbf{e}}_2)\big] = \begin{bmatrix}\cos\phi & -\sin\phi \\ \sin\phi & \cos\phi\end{bmatrix}$$

This is exactly the matrix that we found in Example III.9. There is no need to memorize this matrix. This example has shown how to rederive it very quickly.

This formula $T(\vec{x}) = A\vec{x}$ is the reason we defined matrix multiplication the way we did. More generally, if S and T are two linear transformations with associated matrices M_S and M_T respectively (meaning that $T(\vec{x}) = M_T\vec{x}$ and $S(\vec{y}) = M_S\vec{y}$), then the map constructed by first applying T and then applying S obeys

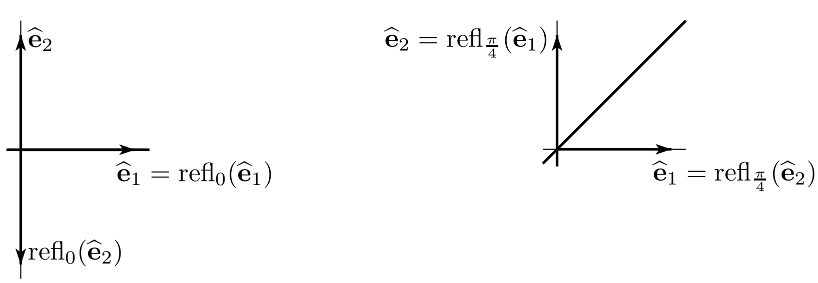
$$S\big(T(\vec{x})\big) = M_ST(\vec{x}) = M_SM_T\vec{x}$$

so that the matrix associated with the composite map $S\big(T(\vec{x})\big)$ is the matrix product M_SM_T of M_S and M_T. It is traditional to use the same symbol to stand for both a linear transformation and its associated matrix. For example, the matrix associated with the linear transformation $T(\vec{x})$ is traditionally denoted T as well, so that $T(\vec{x}) = T\vec{x}$.

Example III.11 Let, as in Example III.8, refl_ϕ be the linear transformation which reflects vectors in the line through the origin that makes an angle ϕ with respect to the x–axis. Define the linear transformation

$$T(\vec{x}) = \mathrm{refl}_{\frac{\pi}{4}}\big(\mathrm{refl}_0(\vec{x})\big)$$

From the figure



we see that

$$\mathrm{refl}_0(\hat{\mathbf{e}}_1) = \begin{bmatrix}1\\0\end{bmatrix} \qquad \mathrm{refl}_0(\hat{\mathbf{e}}_2) = \begin{bmatrix}0\\-1\end{bmatrix} \qquad \mathrm{refl}_{\frac{\pi}{4}}(\hat{\mathbf{e}}_1) = \begin{bmatrix}0\\1\end{bmatrix} \qquad \mathrm{refl}_{\frac{\pi}{4}}(\hat{\mathbf{e}}_2) = \begin{bmatrix}1\\0\end{bmatrix}$$

so that

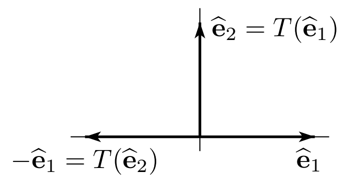
$$\mathrm{refl}_0(\vec{x}) = \big[\mathrm{refl}_0(\hat{\mathbf{e}}_1)\ \mathrm{refl}_0(\hat{\mathbf{e}}_2)\big]\vec{x} = \begin{bmatrix}1&0\\0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} \qquad \mathrm{refl}_{\frac{\pi}{4}}(\vec{y}) = \big[\mathrm{refl}_{\frac{\pi}{4}}(\hat{\mathbf{e}}_1)\ \mathrm{refl}_{\frac{\pi}{4}}(\hat{\mathbf{e}}_2)\big]\vec{y} = \begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix}$$

The linear transformation T is thus

$$T(\vec{x}) = \mathrm{refl}_{\frac{\pi}{4}}\big(\mathrm{refl}_0(\vec{x})\big) = \begin{bmatrix}0&1\\1&0\end{bmatrix}\mathrm{refl}_0(\vec{x}) = \begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&0\\0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

This tells us in particular that

$$T(\hat{\mathbf{e}}_1) = \begin{bmatrix}0\\1\end{bmatrix} = \hat{\mathbf{e}}_2 \qquad T(\hat{\mathbf{e}}_2) = \begin{bmatrix}-1\\0\end{bmatrix} = -\hat{\mathbf{e}}_1$$



The vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are each rotated by 90°. Since

$$\mathrm{refl}_{\frac{\pi}{4}}\big(\mathrm{refl}_0(\hat{\mathbf{e}}_i)\big) = T(\hat{\mathbf{e}}_i) = \mathrm{rot}_{\frac{\pi}{2}}(\hat{\mathbf{e}}_i)$$

for both $i = 1.2$, the matrix for T is the same as the matrix for $\mathrm{rot}_{\frac{\pi}{2}}$. So $\mathrm{refl}_{\frac{\pi}{4}}\big(\mathrm{refl}_0(\vec{x})\big) = \mathrm{rot}_{\frac{\pi}{2}}(\vec{x})$ for all $\vec{x}$. This is often written $\mathrm{refl}_{\frac{\pi}{4}} \circ \mathrm{refl}_0 = \mathrm{rot}_{\frac{\pi}{2}}$ where $S \circ T$ means "first apply T and then apply S". That is, $S \circ T(\vec{x}) = S\big(T(\vec{x})\big)$.

Exercises for §III.2

1) Find the matrices which project on the lines
   a) $x = y$ b) $3x + 4y = 0$
2) Find the matrices which reflect in the lines
   a) $x = y$ b) $3x + 4y = 0$
3) Find the matrices which project on the planes
   a) $x = y$ b) $x + 2y + 2z = 0$
4) Find the matrices which reflect in the planes
   a) $x = y$ b) $x + 2y + 2z = 0$
5) Find the matrices which rotate about the origin in $\mathbb{R}^2$ by
   a) $\pi/4$ b) $\pi/2$ c) π
6) Find the matrix which rotates about the z–axis in $\mathbb{R}^3$ by θ.
7) The matrix

$$\frac{1}{4}\begin{bmatrix}\sqrt{3}+2 & \sqrt{3}-2 & -\sqrt{2}\\ \sqrt{3}-2 & \sqrt{3}+2 & -\sqrt{2}\\ \sqrt{2} & \sqrt{2} & \sqrt{12}\end{bmatrix}$$

   is a rotation in $\mathbb{R}^3$. What axis does it rotate about and what is the angle of rotation?
8) Find the matrix which first reflects about the line in $\mathbb{R}^2$ that makes an angle ϕ with the x–axis and then reflects about the line that makes an angle θ with the x–axis. Give another geometric interpretation of this matrix.

§III.3 Another Application of Matrix Multiplication – Random Walks

"Random walks", or more precisely "discrete random walks", refer to a class of problems in which you are given the following information.

(H_1) The system of interest has a finite number of possible states, labelled 1, 2, 3, $\cdots$, S.

(H_2) We are interested in the system at times $t = 0,\ 1,\ 2,\ 3,\ \cdots$.

(H_3) If at some time n the system is in some state j, then, at time $n+1$ the system is in state i with probability $p_{i,j}$. This is the case for each $i = 1,\ 2,\ 3,\ \cdots,\ S$. The $p_{i,j}$'s are given numbers that obey $\sum_{i=1}^{S} p_{i,j} = 1$.

(Here H_p stands for "hypothesis number p".) Let the components of the column vector

$$\vec{x}_n = \begin{bmatrix} x_{n,1} \\ x_{n,2} \\ \vdots \\ x_{n,S} \end{bmatrix}$$

be the probabilities that, at time n, the system is in state 1, state 2, $\cdots$, state S, respectively. That is, $x_{n,j}$ denotes the probability that, at time n, the system is in state j. Rather than using the language of probability, you can imagine that the system consists of piles of sand located at sites 1 through S. There is a total of one ton of sand. At time n, the amount of sand at site 1 is $x_{n,1}$, the amount of sand at site 2 is $x_{n,2}$ and so on. According to (H_3), between time n and time $n+1$, the fraction $p_{i,j}$ of the sand at site j is moved to site i. So $p_{i,j}x_{n,j}$ tons of sand are moved from site j to site i by time $n+1$. The total amount of sand at site i at time $n+1$ is the sum, over j from 1 to S, of the amount $p_{i,j}x_{n,j}$ of sand moved to site i from site j. Hence

$$x_{n+1,i} = \sum_{j=1}^{S} p_{i,j} x_{n,j}$$

In vector and matrix notation

$$\vec{x}_{n+1} = P\vec{x}_n \tag{III.1}$$

where P is the $S \times S$ matrix whose entry in row i, column j is $p_{i,j}$. The $p_{i,j}$'s are called the transition probabilities for the random walk and P is called the transition matrix.

(H_4) In a random walk problem, you are also given the initial condition $\vec{x}_0$. Often, you will be told that at time 0, the system is in one specific state j_0. In this case, $x_{0,j_0} = 1$ and $x_{0,j} = 0$ for all $j \neq j_0$, so that

$$\vec{x}_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{row } j_0$$

We now have enough information to determine the time evolution of the system. Repeatedly applying (III.1),

$$\vec{x}_1 = P\vec{x}_0 \qquad \vec{x}_2 = P\vec{x}_1 = P^2\vec{x}_0 \qquad \vec{x}_3 = P\vec{x}_2 = P^3\vec{x}_0 \qquad \cdots \qquad \vec{x}_n = P^n\vec{x}_0 \qquad \cdots$$

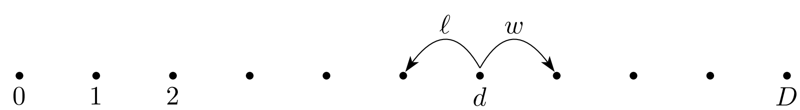
Example III.12 (Gambler's Ruin) I will give two descriptions of this random walk. The first description motivates the name "Gambler's Ruin".

Imagine that you are a gambler. At time zero you walk into a casino with a stake of \$$d$. At that time the house has \$$(D-d)$. (So the total amount of money in play is \$$D$.) At each time 1, 2, $\cdots$, you play a game of chance in which you win \$1 from the house with probability w and loose \$1 to the house with probability $\ell = 1 - w$. This continues until either you have \$0 (in which case you are broke) or you have \$$D$ (in which case the house is broke). That is, if at time n you have \$0, then at time $n+1$ you again have \$0 and if at time n you have \$$D$, then at time $n+1$ you again have \$$D$. The transition probabilities for Gambler's Ruin are

$$p_{i,j} = \begin{cases} 1 & \text{if } j=0 \text{ and } i=0 \\ 0 & \text{if } j=0 \text{ and } i\ne 0 \\ w & \text{if } 0<j<D \text{ and } i=j+1 \\ \ell = 1-w & \text{if } 0<j<D \text{ and } i=j-1 \\ 0 & \text{if } 0<j<D \text{ and } i\ne j-1, j+1 \\ 1 & \text{if } j=D \text{ and } i=D \\ 0 & \text{if } j=D \text{ and } i\ne D \end{cases}$$

and the transition matrix is

$$P = \begin{array}{c} \\ i=0 \\ i=1 \\ i=2 \\ i=3 \\ \vdots \\ i=D-3 \\ i=D-2 \\ i=D-1 \\ i=D \end{array} \begin{array}{c} \begin{array}{ccccccc} j=0 & j=1 & j=2 & \cdots & j=D-2 & j=D-1 & j=D \end{array} \\ \begin{pmatrix} 1 & \ell & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \ell & \cdots & 0 & 0 & 0 \\ 0 & w & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & w & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \ell & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \ell & 0 \\ 0 & 0 & 0 & \cdots & w & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & w & 1 \end{pmatrix} \end{array}$$



The second description, of the same mathematical system, motivates why it is called a "random walk".

At time zero, a drunk is at a location d. Once each unit of time, the drunk staggers to the right one unit with probability w and staggers to the left one unit with probability $\ell = 1-w$. This continues until the drunk reaches either the bar at 0 or the bar at D. Once the drunk reaches a bar, he remains there for ever.

Here is a table giving the time evolution of Gambler's Ruin, assuming that $d=2$, $D=8$, $w=0.49$ and $\ell = 0.51$. (Entries are rounded to three decimal places.)

$$\vec{x}_0 = \begin{bmatrix} 0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0 \end{bmatrix} \begin{array}{c} \$0\\ \$1\\ \$2\\ \$3\\ \$4\\ \$5\\ \$6\\ \$7\\ \$8\\ \$9\\ \$10 \end{array} \quad \vec{x}_1 = \begin{bmatrix} 0\\0.51\\0\\0.49\\0\\0\\0\\0\\0\\0\\0 \end{bmatrix} \quad \vec{x}_2 = \begin{bmatrix} 0.260\\0\\0.500\\0\\0.240\\0\\0\\0\\0\\0\\0 \end{bmatrix} \quad \vec{x}_3 = \begin{bmatrix} 0.260\\0.255\\0\\0.367\\0\\0.118\\0\\0\\0\\0\\0 \end{bmatrix} \quad \vec{x}_4 = \begin{bmatrix} 0.390\\0\\0.312\\0\\0.240\\0\\0.058\\0\\0\\0\\0 \end{bmatrix} \quad \vec{x}_5 = \begin{bmatrix} 0.390\\0.159\\0\\0.275\\0\\0.147\\0\\0.028\\0\\0\\0 \end{bmatrix} \quad \vec{x}_6 = \begin{bmatrix} 0.471\\0\\0.218\\0\\0.210\\0\\0.086\\0\\0.014\\0\\0 \end{bmatrix}$$

Here, for example, $\vec{x}_0$ says that the gambler started with \$2 at time 0. The vector $\vec{x}_1$ says that, at time 1, he has \$1 with probability 0.51 and \$3 with probability 0.49. The vector $\vec{x}_2$ says that, at time 2, he has \$0 with probability $0.51 \times 0.51 = 0.2601$, \$2 with probability $0.51 \times 0.49 + 0.49 \times 0.51 = 0.4998$ and \$4 with probability $0.49 \times 0.49 = 0.2401$.

§III.4 The Transpose of a Matrix

Definition III.13 The **transpose** of the $m \times n$ matrix A is the $n \times m$ matrix A^t with matrix elements

$$A^t_{i,j} = A_{j,i} \qquad \text{for all } 1 \le i \le n,\ 1 \le j \le m$$

The rows of A^t are the columns of A and the columns of A^t are the rows of A.

Example III.14 The transpose of the 2×3 matrix A on the left below is the 3×2 matrix A^t on the right below.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

There are two important properties of the transpose operation.

1. If A is any $m \times n$ matrix, $\vec{x}$ is any m component vector and y is any n component vector

$$\vec{x} \cdot (A\vec{y}) = (A^t\vec{x}) \cdot \vec{y}$$

To see this, we just compute the left and right hand sides

$$\begin{aligned} \vec{x} \cdot (A\vec{y}) &= \sum_{i=1}^m x_i (A\vec{y})_i \ = \sum_{i=1}^m x_i \Big(\sum_{j=1}^n A_{i,j} y_j \Big) = \sum_{\substack{1\le i\le m \\ 1\le j\le n}} x_i A_{i,j} y_j \\ (A^t\vec{x}) \cdot \vec{y} &= \sum_{j=1}^n (A^t\vec{x})_j y_j = \sum_{j=1}^n \Big(\sum_{i=1}^m A^t_{j,i} x_i \Big) y_j \ = \sum_{\substack{1\le i\le m \\ 1\le j\le n}} x_i A^t_{j,i} y_j \end{aligned}$$

and observe that they are the same, because of the definition of A^t.

2. If A is any $\ell \times m$ and B is any $m \times n$ matrix, then

$$(AB)^t = B^t A^t$$

Be careful about the order of matrix multiplication here. To see this, we also just compute the left and right hand sides

$$\begin{aligned} (AB)^t_{i,j} &= (AB)_{j,i} \qquad &= \sum_{k=1}^m A_{j,k} B_{k,i} \\ (B^t A^t)_{i,j} &= \sum_{k=1}^m B^t_{i,k} A^t_{k,j} &= \sum_{k=1}^m B_{k,i} A_{j,k} \end{aligned}$$

and observe that they are the same.

§III.5 Matrix Inverses

Suppose A is a matrix. What is A^{-1}? It's the thing you multiply A by to get 1. What is 1?

Definition III.15 The $m \times m$ identity matrix I_m (generally the subscript m is dropped from the notation) is the $m \times m$ matrix whose (i, j) matrix element is 1 if $i = j$ and 0 if $i \ne j$.

For example

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The reason we call this the identity matrix is that, for any $m \times n$ matrix A

$$I_m A = A I_n = A$$

It is easy to check that this is true. For example, fix any i and k with $1 \le i \le m$, $1 \le k \le n$. By the definition of I_{ij} the only nonzero term in the sum $(IA)_{ik} = \sum_{j=1}^{m} I_{ij} A_{jk}$ is that with $j = i$. Furthermore the I_{ii} that appears in that term takes the value one so $(IA)_{ik} = I_{ii} A_{ik} = A_{ik}$, as desired.

Definition III.16 A matrix B is called an **inverse** of the matrix A if $AB = I$ and $BA = I$. If A has an inverse, then A is said to be invertible or nonsingular. Otherwise, it is said to be singular. The inverse is generally denoted A^{-1}.

Example III.17 Let's consider the general 2×2 matrix

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

and see if we can find an inverse for it. Let's call the inverse

$$B = \begin{bmatrix} X & X' \\ Y & Y' \end{bmatrix}$$

The matrix A is to be treated as a given matrix. At this stage, B is unknown. To help keep straight what is known and what isn't, I've made all of the knowns lower case and all of the unknowns upper case. To be an inverse, B must obey

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X & X' \\ Y & Y' \end{bmatrix} = \begin{bmatrix} aX + bY & aX' + bY' \\ cX + dY & cX' + dY' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This condition consists of four equations

$$\begin{aligned} aX + bY &= 1 \qquad (1) \\ cX + dY &= 0 \qquad (2) \\ aX' + bY' &= 0 \qquad (1') \\ cX' + dY' &= 1 \qquad (2') \end{aligned}$$

in the four unknowns X, Y, X', Y'. Note that

a) The unknowns X, Y appear only in the first two equations.
b) The unknowns X', Y' appear only in the last two equations.
c) The coefficients on the left hand side of (1) are identical to the coefficients on the left hand side of (1').
d) The coefficients on the left hand side of (2) are identical to the coefficients on the left hand side of (2').

Consequently we can solve for X, Y and X', Y' at the same time:

$$\begin{aligned} &d(1) - b(2): \quad (ad - bc)X = d && d(1') - b(2'): \quad (ad - bc)X' = -b \\ &c(1) - a(2): \quad (bc - ad)Y = c && c(1') - a(2'): \quad (bc - ad)Y' = -a \end{aligned}$$

Dividing across gives

$$B = \begin{bmatrix} X & X' \\ Y & Y' \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Provided $ad - bc \ne 0$, this B exists and obeys equations (1) through (2') and hence $AB = I$. For B to be an inverse for A it must also obey

$$I = BA = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

If you go ahead and multiply out the matrices on the right, you see that this condition is indeed satisfied. We conclude that, if $\det A = ad - bc \neq 0$, the inverse of A exists and is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

On the other hand, if $\det A = ad - bc = 0$, equations d(1)−b(2), d(1')−b(2'), c(1)−a(2) and c(1')−a(2') force $d = -b = c = -a = 0$ and then the left hand side of equation (1) is zero for all values of X and Y so that equation (1) cannot be satisfied and A cannot have an inverse.

Properties of Inverses

1. If $AB = I$ and $CA = I$, then $B = C$. Consequently A has at most one inverse.

   Proof: If $AB = I$ and $CA = I$, then $B = IB = CAB = CI = C$. If B and C are both inverses of A, then, by definition, $AB = BA = I$ and $AC = CA = I$. In particular $AB = I$ and $CA = I$, so that $B = C$.

2. If A and B are both invertible, then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$.

   Proof: We have a guess for $(AB)^{-1}$. To check that the guess is correct, we merely need to check the requirements of the definition

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I$$

3. If A is invertible, then so is A^t and $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$.

   Proof: Let's use B to denote the inverse of A (so there won't be so many superscripts around.) By definition

$$AB = BA = I$$

   These three matrices are the same. So their transposes are the same. Since $(AB)^t = A^tB^t$, $(BA)^t = A^tB^t$ and $I^t = I$, we have

$$B^tA^t = A^tB^t = I^t = I$$

   which is exactly the definition of "the inverse of A^t is B^t".

4. Suppose that A is invertible. Then $A\vec{x} = \vec{b} \iff \vec{x} = A^{-1}\vec{b}$.

   Proof: Multiplying the first equation by A^{-1} gives the second and multiplying the second by A gives the first.

   WARNING: This property is conceptually important. But it is usually computationally much more efficient to solve $A\vec{x} = \vec{b}$ by Gaussian elimination than it is to find A^{-1} and then multiply $A^{-1}\vec{b}$.

5. Only square matrices can be invertible.

   Outline of Proof: Let A be an invertible $m \times n$ matrix. Then there exists an $n \times m$ matrix B such that $AB = I$, where I is the $m \times m$ identity matrix. We shall see in the next section, that the $j^{\rm th}$ column of B is solves the system of equations $A\vec{x} = \hat{\mathbf{e}}_j$, where $\hat{\mathbf{e}}_j$ is the $j^{\rm th}$ column of I. Because the identity matrix I has rank m, there must exist some $1 \le j \le m$ such that the augmented matrix $[A|\hat{\mathbf{e}}_j]$ also has rank m. By property 1 above, the corresponding system of equations must have a unique solution. Consequently, the number of unknowns, n, must equal the rank, m.

6. If A is an $n \times n$ matrix then the following are equivalent:
   (i) A is invertible.
   (ii) For each vector $\vec{b}$, the system of equations $A\vec{x} = \vec{b}$ has exactly one solution.

(iii) $\vec{x}=\vec{0}$ is the only solution of $A\vec{x}=\vec{0}$
(iv) A has rank n.
(v) There is a matrix B such that $AB=I$ (i.e. A has a right inverse).
(vi) There is a matrix B such that $BA=I$ (i.e. A has a left inverse).

In other words, if a given square matrix satisfies any one of properties (i)–(vi), it satisfies all of them.

Proof: The notation " (i)⇒(ii) " means " If statement (i) is true, then statement (ii) is also true".

(i)⇒(ii): If A is invertible, then $A\vec{x}=\vec{b}$ is true if and only if $\vec{x}=A^{-1}\vec{b}$ is true.

(ii)⇒(iii): $\vec{x}=0$ always solves $A\vec{x}=0$. So (iii) is the special case of (ii) with $\vec{b}=0$.

(iii)⇒(iv): $A\vec{x}=\vec{0}$ has precisely one solution $\Rightarrow$ $\text{rank}\,A=\#\text{unknowns}=n$.

(iv)⇒(v): Let A be an $n\times n$ matrix with rank n. We are to find another $n\times n$ matrix B that obeys $AB=I$. This is a system of linear equations for the unknown matrix B. We are more used to systems of equations with the unknowns being vectors. We can convert our "unknown matrix" problem into "unkown vector" problems just by giving names to the columns of B and I. Of course the j^{th} column of I is the standard unit vector $\widehat{\mathbf{e}}_j$, all of whose components are zero except for the j^{th}, which is 1. Let's use $\vec{B}_j$ to denote the j^{th} column of B. We are to solve $A\big[\vec{B}_1\ \vec{B}_2\ \cdots\ \vec{B}_n\big]=\big[\widehat{\mathbf{e}}_1\ \widehat{\mathbf{e}}_2\ \cdots\ \widehat{\mathbf{e}}_n\big]$. For the two matrices $A[\vec{B}_1\ \vec{B}_2\ \cdots\ \vec{B}_n]=[A\vec{B}_1\ A\vec{B}_2\ \cdots\ A\vec{B}_n]$ and $[\widehat{\mathbf{e}}_1\ \widehat{\mathbf{e}}_2\ \cdots\ \widehat{\mathbf{e}}_n]$ to be equal, all of their columns must agree, so the requirement $AB=I$ may be expressed as

$$A\vec{B}_i=\widehat{\mathbf{e}}_i \qquad \text{for } i=1,\ \cdots,\ n$$

(This argument is repeated, in more detail, with examples, in the next section.) Since A has rank n and $A\vec{B}_i=\widehat{\mathbf{e}}_i$ is a system of n linear equations in n unknowns, it has a unique solution.

(v)⇒(vi): Assume $AB=I$. Then $B\vec{x}=\vec{0}\ \Rightarrow\ AB\vec{x}=\vec{0}\ \Rightarrow\ \vec{x}=\vec{0}$. So condition (ii) is applicable to B, which implies that (iii) and subsequently (iv) are also applicable to B. This implies that there is a matrix C obeying $BC=I$. But $C=(AB)C=A(BC)=A$ so $BA=I$.

(vi)⇒(i): Assume that $BA=I$. Then $A\vec{x}=\vec{0}\ \Rightarrow\ BA\vec{x}=\vec{0}\ \Rightarrow\ \vec{x}=\vec{0}$. So (ii) applies to A. So (iv) and (v) apply to A.

Items (ii), (iii), (iv), (v) and (vi) of property 4 are all tests for invertibility. We shall get another test, once we have generalized the definition of determinant to matrices larger than 3×3: a square matrix A is invertible if and only if $\det A\neq 0$.

Example III.18 Let A be the matrix which implements reflection in the line $y=x$, let B be the matrix that implements reflection in the x axis and let C be the matrix for the linear transformation that first refelcts in the x axis and then reflects in the line $y=x$. We saw in Example III.11 that

$$A=\begin{bmatrix}0&1\\1&0\end{bmatrix} \qquad B=\begin{bmatrix}1&0\\0&-1\end{bmatrix} \qquad C=AB=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$$

Now the inverse of any reflection is itself. (That is, executing the same reflection twice returns every vector to its original location.) So

$$A^{-1}=A=\begin{bmatrix}0&1\\1&0\end{bmatrix} \qquad B^{-1}=B=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$$

(Go ahead and check for yourself that the matrix products AA and BB are both I.) So

$$C^{-1}=(AB)^{-1}=B^{-1}A^{-1}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

We saw in Example III.11 that C was rotation by $90°$. So C^{-1} should be rotation by $-90°$. Now rotation by $-90°$ maps $\widehat{\mathbf{e}}_1$ to $-\widehat{\mathbf{e}}_2$ and $\widehat{\mathbf{e}}_2$ to $\widehat{\mathbf{e}}_1$, which is exactly what $\begin{bmatrix}0&1\\-1&0\end{bmatrix}$ does.

Exercises for §III.5

1) Determine which of the following matrices are invertible and, for each that is, find its inverse. Use the method of Example II.10. Do not used the canned formula derived in Example II.10.

   a) $\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$ *b)* $\begin{bmatrix}1 & 3\\ -2 & 2\end{bmatrix}$ *c)* $\begin{bmatrix}2 & 4\\ 1 & 2\end{bmatrix}$

2) Either prove or find a counterexample to the statement "If $BA = I$ and $AC = I$ then $B = C$."

§III.6 Finding Inverses

Suppose that we are given an $n \times n$ matrix A and we wish to find a matrix B obeying $AB = I$. Denote by $\vec{B}_i$ the (as yet unknown) $i^{\rm th}$ column of B and by $\hat{\mathbf{e}}_i$ the column vector having all entries zero except for a one in the $i^{\rm th}$ row. For example, when $n = 2$

$$B = \begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix} \qquad \vec{B}_1 = \begin{bmatrix}b_{11}\\ b_{21}\end{bmatrix} \qquad \vec{B}_2 = \begin{bmatrix}b_{12}\\ b_{22}\end{bmatrix} \qquad \hat{\mathbf{e}}_1 = \begin{bmatrix}1\\ 0\end{bmatrix} \qquad \hat{\mathbf{e}}_2 = \begin{bmatrix}0\\ 1\end{bmatrix}$$

In this notation, the requirement $AB = I$ is

$$A[\vec{B}_1\ \vec{B}_2\ \cdots\ \vec{B}_n] = [\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\ \cdots\ \hat{\mathbf{e}}_n]$$

By the definition of matrix multiplication

$$A[\vec{B}_1\ \vec{B}_2\ \cdots\ \vec{B}_n] = [A\vec{B}_1\ A\vec{B}_2\ \cdots\ A\vec{B}_n]$$

That is, the first column of AB is $A\vec{B}_1$. For example, when $n = 2$, the first column of

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

is indeed identical to

$$A\vec{B}_1 = \begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}b_{11}\\ b_{21}\end{bmatrix} = \begin{bmatrix}a_{11}b_{11} + a_{12}b_{21}\\ a_{21}b_{11} + a_{22}b_{21}\end{bmatrix}$$

For the two matrices $A[\vec{B}_1\ \vec{B}_2\ \cdots\ \vec{B}_n] = [A\vec{B}_1\ A\vec{B}_2\ \cdots\ A\vec{B}_n]$ and $[\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\ \cdots\ \hat{\mathbf{e}}_n]$ to be equal, all of their columns must agree, so the requirement $AB = I$ may be expressed as

$$A\vec{B}_i = \hat{\mathbf{e}}_i \qquad\qquad \text{for } i = 1,\ \cdots,\ n$$

Recall that A and the $\hat{\mathbf{e}}_i$'s are all given matrices and that the $\vec{B}_i$'s are all unknown. We must solve n different systems of linear equations. The augmented matrix for system number i is $[A|\hat{\mathbf{e}}_i]$. We could apply Gauss reduction separately to the n systems. But because the left hand sides of all n systems are the same, we can solve the n systems simultaneously. We just form one big augmented matrix $[A|\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\ \cdots\ \hat{\mathbf{e}}_n]$. This augmented matrix is just short hand notation for the n systems of equations $A\vec{B}_1 = \hat{\mathbf{e}}_1$, $A\vec{B}_2 = \hat{\mathbf{e}}_2$, $\cdots$, $A\vec{B}_n = \hat{\mathbf{e}}_n$. Here are two examples of this technique.

Example III.19 Let

$$A = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}$$

Of course we have already derived, in Example III.17, a canned formula for the inverse of a 2×2 matrix. But we'll find the inverse of A using the Gaussian elimination technique anyway, just to provide a trivial example of the mechanics of the technique. In this example, the augmented matrix is

$$\big[A\,\big|\,\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\big] = \left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\ 1 & 2 & 0 & 1\end{array}\right]$$

It is important to always remember that if we were to erase all columns to the right of the vertical line except for the $i^{\rm th}$, we would have precisely the augmented matrix appropriate for the linear system $A\vec{B}_i = \widehat{\mathbf{e}}_i$. So any row operation applied to the big augmented matrix $[A|\widehat{\mathbf{e}}_1\ \widehat{\mathbf{e}}_2]$ really is a simultaneous application of the same row operation to 2 different augmented matrices $[A|e_i]$ at the same time. Row reduce in the usual way. The row echelon (upper triangular) form of this augment matrix is

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] \begin{array}{l} \\ (2)-(1) \end{array}$$

We could backsolve the two systems of equations separately. But it is easier to treat the two at the same time by further reducing the augmented matrix to reduced row echelon form.

$$\left[\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \end{array}\right] \begin{array}{l} (1)-(2) \\ \ \end{array}$$

What conclusion do we come to? Concentrate on, for example, the first column to the right of the vertical line. In fact, mentally erase the second column to the right of the vertical line in all of the above computations. Then the above row operations converted

$$[\,A\,|\,\widehat{\mathbf{e}}_1\,] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 2 & 0 \end{array}\right] \qquad \text{to} \qquad \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -1 \end{array}\right] = \left[\begin{array}{c|c} I & \begin{array}{c} 2 \\ -1 \end{array} \end{array}\right]$$

Because row operations have no effect on the set of solutions of a linear system we can conclude that

$$\vec{B}_1 \text{ obeys} \qquad A\vec{B}_1 = \widehat{\mathbf{e}}_1 \qquad \text{if and only if it obeys} \qquad I\vec{B}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Since $I\vec{B}_1 = \vec{B}_1$, we have that $\vec{B}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Similarly, $\vec{B}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Thus

$$A^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

which is exactly the matrix to the right of the vertical bar in the row reduced echelon form.

Example III.20 Now let's compute the inverse of a matrix for which we do not already have a canned formula. Let

$$A = \begin{bmatrix} 2 & -3 & 2 & 5 \\ 1 & -1 & 1 & 2 \\ 3 & 2 & 2 & 1 \\ 1 & 1 & -3 & 1 \end{bmatrix}$$

Form the augmented matrix

$$\left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 3 & 2 & 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & -3 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

Again note that if we were to erase all columns to the right of the vertical line except for the $i^{\rm th}$, we would have precisely the augmented matrix appropriate for the linear system $A\vec{B}_i = \widehat{\mathbf{e}}_i$, so that any row operation applied to the big augmented matrix $[A|\widehat{\mathbf{e}}_1\ \widehat{\mathbf{e}}_2\ \cdots\ \widehat{\mathbf{e}}_n]$ really is a simultaneous application of the same row operation to n different augmented matrices $[A|e_i]$ at the same time. Row reduce in the usual way.

$$\begin{array}{r} \\ (2)-0.5(1) \\ (3)-1.5(1) \\ (4)-0.5(1) \end{array} \left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & -0.5 & -0.5 & 1 & 0 & 0 \\ 0 & 6.5 & -1 & -6.5 & -1.5 & 0 & 1 & 0 \\ 0 & 2.5 & -4 & -1.5 & -0.5 & 0 & 0 & 1 \end{array}\right]$$

$$\begin{array}{r} \\ \\ (3)-13(2) \\ (4)-5(2) \end{array} \left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & -0.5 & -0.5 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 5 & -13 & 1 & 0 \\ 0 & 0 & -4 & 1 & 2 & -5 & 0 & 1 \end{array}\right]$$

$$\begin{array}{r} \\ \\ \\ (4)-4(3) \end{array} \left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & -0.5 & -0.5 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 5 & -13 & 1 & 0 \\ 0 & 0 & 0 & 1 & -18 & 47 & -4 & 1 \end{array}\right]$$

Again, rather than backsolving the four systems individually, it is easier to do all four at the same time by applying more row operations chosen to turn the left hand side into the identity matrix.

$$\begin{array}{r} \\ \\ -(3) \\ \\ \end{array} \left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & -0.5 & -0.5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -5 & 13 & -1 & 0 \\ 0 & 0 & 0 & 1 & -18 & 47 & -4 & 1 \end{array}\right]$$

$$\begin{array}{r} \\ 2(2)+(4) \\ \\ \\ \end{array} \left[\begin{array}{cccc|cccc} 2 & -3 & 2 & 5 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -19 & 49 & -4 & 1 \\ 0 & 0 & 1 & 0 & -5 & 13 & -1 & 0 \\ 0 & 0 & 0 & 1 & -18 & 47 & -4 & 1 \end{array}\right]$$

$$0.5[(1)+3(2)-2(3)-5(4)] \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 22 & -57 & 5 & -1 \\ 0 & 1 & 0 & 0 & -19 & 49 & -4 & 1 \\ 0 & 0 & 1 & 0 & -5 & 13 & -1 & 0 \\ 0 & 0 & 0 & 1 & -18 & 47 & -4 & 1 \end{array}\right]$$

By exactly the same argument as we used at the end of Example III.19, the inverse is the matrix to the right of the vertical bar in the row reduced echelon form. That is,

$$A^{-1} = \begin{bmatrix} 22 & -57 & 5 & -1 \\ -19 & 49 & -4 & 1 \\ -5 & 13 & -1 & 0 \\ -18 & 47 & -4 & 1 \end{bmatrix}$$

Exercises for §III.6

1) Determine which of the following matrices are invertible and, for each that is, find its inverse.

a) $\begin{bmatrix} 2 & 3 & -1 \\ 1 & 2 & 3 \\ -1 & -1 & 4 \end{bmatrix}$ *b*) $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & -1 \\ 2 & -1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$ *d*) $\begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 5 \\ 0 & -1 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ *f*) $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ where $ab \neq 0$

g) $\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *h*) $\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$

§III.7 Determinants – Definition

The determinant of a 1×1 matrix is defined by

$$\det [a_{11}] = a_{11}$$

The determinant of a 2×2 matrix is defined by

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

For $n > 2$, (in fact $n \geq 2$) the determinant of an $n \times n$ matrix A, whose ij entry is denoted a_{ij}, is defined by

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det M_{1j}$$

where M_{1j} is the $(n-1) \times (n-1)$ matrix formed by deleting, from the original matrix A, the row and column containing a_{1j}. This formula is called "expansion along the top row". There is one term in the formula for each entry in the top row. The term is a sign times the entry times the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the row and column that contains the entry. The sign alternates, starting with a $+$.

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= a_{11} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &= a_{11}\,(a_{22}a_{33} - a_{23}a_{32}) \quad - a_{12}\,(a_{21}a_{33} - a_{23}a_{31}) \quad + a_{13}\,(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

Example III.21

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} &= 1 \times \det \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix} - 2 \times \det \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} + 3 \times \det \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} \\ &= 1 \times (0-4) - 2(1-6) + 3(2-0) = 12 \end{aligned}$$

Example III.22 In this example we compute, using the definition of the determinant,

$$\det \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 3 & 3 & 3 \\ 7 & 8 & 9 & 12 \end{bmatrix} = \det \begin{bmatrix} 3 & 2 & 1 \\ 3 & 3 & 3 \\ 8 & 9 & 12 \end{bmatrix} - 2\det \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 3 \\ 7 & 9 & 12 \end{bmatrix} + 3\det \begin{bmatrix} 4 & 3 & 1 \\ 2 & 3 & 3 \\ 7 & 8 & 12 \end{bmatrix} - 4\det \begin{bmatrix} 4 & 3 & 2 \\ 2 & 3 & 3 \\ 7 & 8 & 9 \end{bmatrix}$$

As side computations, we evaluate the four 3×3 determinants

$$\begin{aligned} \det \begin{bmatrix} 3 & 2 & 1 \\ 3 & 3 & 3 \\ 8 & 9 & 12 \end{bmatrix} &= 3\det \begin{bmatrix} 3 & 3 \\ 9 & 12 \end{bmatrix} - 2\det \begin{bmatrix} 3 & 3 \\ 8 & 12 \end{bmatrix} + \det \begin{bmatrix} 3 & 3 \\ 8 & 9 \end{bmatrix} \\ &= 3(36-27) - 2(36-24) + (27-24) \\ &= 6 \\ \det \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 3 \\ 7 & 9 & 12 \end{bmatrix} &= 4\det \begin{bmatrix} 3 & 3 \\ 9 & 12 \end{bmatrix} - 2\det \begin{bmatrix} 2 & 3 \\ 7 & 12 \end{bmatrix} + \det \begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix} \\ &= 4(36-27) - 2(24-21) + (18-21) \\ &= 27 \end{aligned}$$

$$\det\begin{bmatrix}4&3&1\\2&3&3\\7&8&12\end{bmatrix} = 4\det\begin{bmatrix}3&3\\8&12\end{bmatrix} - 3\det\begin{bmatrix}2&3\\7&12\end{bmatrix} + \det\begin{bmatrix}2&3\\7&8\end{bmatrix}$$
$$= 4(36-24) - 3(24-21) + (16-21)$$
$$= 34$$
$$\det\begin{bmatrix}4&3&2\\2&3&3\\7&8&9\end{bmatrix} = 4\det\begin{bmatrix}3&3\\8&9\end{bmatrix} - 3\det\begin{bmatrix}2&3\\7&9\end{bmatrix} + 2\det\begin{bmatrix}2&3\\7&8\end{bmatrix}$$
$$= 4(27-24) - 3(18-21) + 2(16-21)$$
$$= 11$$

So

$$\det\begin{bmatrix}1&2&3&4\\4&3&2&1\\2&3&3&3\\7&8&9&12\end{bmatrix} = 6 - 2\times 27 + 3\times 34 - 4\times 11 = 10$$

This is clearly a very tedious procedure. We will develop a better one soon.

Exercises for §III.7

1) Evaluate the determinant of each of the following matrices

$$a)\ \begin{bmatrix}1&3\\2&4\end{bmatrix} \qquad b)\ \begin{bmatrix}0&-3\\2&-1\end{bmatrix} \qquad c)\ \begin{bmatrix}0&1\\1&0\end{bmatrix} \qquad d)\ \begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}$$

2) Evaluate the determinant of each of the following matrices by expanding along the top row

$$a)\ \begin{bmatrix}2&0&0\\1&-1&5\\2&3&-1\end{bmatrix} \qquad b)\ \begin{bmatrix}2&1&5\\1&0&3\\-1&2&0\end{bmatrix} \qquad c)\ \begin{bmatrix}7&-1&5\\3&4&-5\\2&3&0\end{bmatrix} \qquad d)\ \begin{bmatrix}1&a&b\\0&1&c\\0&0&1\end{bmatrix}$$

§III.8 Determinants – Properties

Property E

If two rows of an $n\times n$ matrix are exchanged, the determinant is multiplied by -1.

$$\det\begin{bmatrix}\vdots\\\vec a_i\\\vdots\\\vec a_j\\\vdots\end{bmatrix} = -\det\begin{bmatrix}\vdots\\\vec a_j\\\vdots\\\vec a_i\\\vdots\end{bmatrix}$$

where, for each $1\le k\le n$, $\vec a_k$ is the $k^{\rm th}$ row of the matrix and is an n component row vector.

Property M

Multiplying any **single** row of a matrix by t multiplies the determinant by t as well.

$$\det\begin{bmatrix}\vec a_1\\\vdots\\t\vec a_i\\\vdots\\\vec a_n\end{bmatrix} = t\det\begin{bmatrix}\vec a_1\\\vdots\\\vec a_i\\\vdots\\\vec a_n\end{bmatrix}$$

To multiply every entry in an $n \times n$ matrix by t, we have to apply Property M once for each row in the matrix, so we end up with a factor t^n. For example

$$\det \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} = t \times \det \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix} = t \times t \times \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = t^2$$

Property A

Adding any multiple of any row to any other row has no effect on the determinant.

$$\det \begin{bmatrix} \vdots \\ \vec{a}_i + t\vec{a}_m \\ \vdots \\ \vec{a}_m \\ \vdots \end{bmatrix} = \det \begin{bmatrix} \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_m \\ \vdots \end{bmatrix}$$

Property D

The determinant of any triangular matrix is the product of its diagonal entries.

$$\det \begin{bmatrix} a_{11} & * & \cdots & * & * \\ 0 & a_{22} & \cdots & * & * \\ & & \ddots & & \\ 0 & 0 & \cdots & a_{n-1\,n-1} & * \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix} = a_{11}a_{22}\cdots a_{nn}$$

Property P

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

Property T

$$\det A^t = \det A$$

Outline of Proof of Properties E, M, A, D

We will not give a complete proof. But we will give an outline that is sufficiently detailed that you should be able to fill in the gaps.

All of these properties follow fairly easily from another (in fact the standard) definition of determinant, which in turn follows fairly easily from our definition. For any $n \times n$ matrix A

$$\det A = \sum_{\sigma \in P_n} \operatorname{sgn} \sigma \; A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} \tag{III.2}$$

Here P_n is the set of all orderings of $(1, 2, \cdots, n)$. The symbol "P" stands for "permutation", which is the mathematical name for a reordering. So

$$P_3 = \big\{(1,2,3),\ (1,3,2),\ (2,1,3),\ (2,3,1),\ (3,1,2),\ (3,2,1)\big\}$$

We use $\sigma(i)$ to denote the i^{th} entry of the permutation σ. For example, when $\sigma = (2, 3, 1)$, $\sigma(2) = 3$. The sign of a permutation, here denoted $\operatorname{sgn}\sigma$ is $+1$ if $(1, 2, \cdots, n)$ can be transformed into σ using an even number of exchanges of pairs of numbers. Otherwise, the sign is -1. For example $(1, 2, 3)$ can be tranformed into $(1, 3, 2)$ by just exchanging 2 and 3, so the sign of $\sigma = (1, 3, 2)$ is -1. On the other hand, $(1, 2, 3)$ can be tranformed into $(3, 1, 2)$ by first exchanging 1 and 3 to yield $(3, 2, 1)$ and then exchanging 1 and 2 to yield $(3, 1, 2)$. So the sign of $\sigma = (3, 1, 2)$ is $+1$. Of course $(1, 2, 3)$ can also be transformed into $(3, 1, 2)$ by first exchanging 1 and 2 to yield $(2, 1, 3)$, then exchanging 1 and 3 to yield $(2, 3, 1)$, then exchanging 2 and 3 to yield $(3, 2, 1)$ and finally exchanging 1 and 2 to yield $(3, 1, 2)$. This used four exchanges, which is still even. It is possible to prove that the number of exchanges of pairs of numbers used to transform $(1, 2, 3)$ to $(3, 1, 2)$ is always even. It is also possible to prove the $\operatorname{sgn}\sigma$ is well–defined for all permutations σ.

If A is a 2×2 matrix, the above definition is

$$\det A = A_{11}A_{22} - A_{12}A_{21}$$

with the first term being the $\sigma = (1, 2)$ contribution and the second term being the $\sigma = (2, 1)$ contribution. If A is a 3×3 matrix, the above definition is

$$\det A = A_{11}A_{22}A_{33} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31}$$

with the terms being the contributions, in order, from $\sigma = (1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$, $(3, 2, 1)$.

To verify that the new and old definitions of determinant agree, it suffices to prove that they agree for 1×1 matrices, which is trivial, and that the new definition obeys the "expansion along the first row" formula.

Outline of Proof of Property P

Once again, we will not give a complete proof. First assume that A is invertible. Then there is a sequence of row operations, that when applied to A, convert it into the identity matrix. That was the basis of the algorithm for computing A^{-1} that we developed in §III.6. Any row operation can be implemented by multiplication by a matrix. For example, if A is a 4×4 matrix, all rows of

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A$$

with the exception of row (3) coincide with the corresponding row of A. Row (3) of the product is row (3) of A plus 4 times row (1) of A. The matrices that implement row operations are called elementary matrices. We shall denote them by Q_j here. We have

$$Q_h \cdots Q_1 A = I$$

for some sequence $Q_1, \ \cdots, \ Q_h$ of row operation implementing matrices.

The next step is to check that

$$\det QC = \det Q \det C$$

for any elementary matrix Q and any square matrix C. This is straight forward because we already know the effect that any row operation has on a determinant. Then

$$\begin{aligned} Q_h \cdots Q_1 A = I \quad &\implies \quad \det Q_h \cdots \det Q_1 \ \det A = 1 \\ Q_h \cdots Q_1 AB = B \quad &\implies \quad \det Q_h \cdots \det Q_1 \ \det AB = \det B \\ &\implies \quad \frac{1}{\det A} \det AB = \det B \end{aligned}$$

Now suppose that A is not invertible. Then there is a sequence of row operations, that when applied to A, convert it into a matrix $\tilde{A}$, whose last row is identically zero. Implementing these row operations using multiplication by elementary matrices, we have

$$\tilde{Q}_{\tilde{h}} \cdots \tilde{Q}_1 A = \tilde{A} \quad \Longrightarrow \quad \det \tilde{Q}_{\tilde{h}} \cdots \det \tilde{Q}_1 \det A = \det \tilde{A}$$

Any matrix that has at least one row identically zero, like $\tilde{A}$, has determinant zero. Applying $Q_h \cdots Q_1 A = I \implies \det Q_h \cdots \det Q_1 \ \det A = 1$ with $h = 1$ and A replaced by the inverse of Q_1, we see that every elementary matrix has nonzero determinant. So we conclude that if A is not invertible it necessarily has determinant zero. Finally, observe that if A fails to be invertible, the same is true for AB (otherwise $B(AB)^{-1}$ is an inverse for A) and so both $\det A = 0$ and $\det AB = 0$.

Outline of Proof of Property T

By (III.2) and the definition of "transpose"

$$\det A^t = \sum_{\sigma \in P_n} \operatorname{sgn} \sigma \ \prod_{i=1}^{n} A_{\sigma(i)\, i}$$

Concentrate on one term in this sum. By the definition of permutation, each of the integers 1, 2, $\cdots$, n appears exactly once in $\sigma(1)$, $\sigma(2)$, $\cdots$, $\sigma(n)$. Reorder the factors in the product $A_{\sigma(1)\,1} A_{\sigma(2)\,2} \cdots A_{\sigma(n)\,n}$ so that the first indices, rather than the second indices are in increasing order. This can be impemented by making the change of variables $i = \sigma^{-1}(j)$ in the product.

$$\det A^t = \sum_{\sigma \in P_n} \operatorname{sgn} \sigma \ \prod_{j=1}^{n} A_{j\, \sigma^{-1}(j)}$$

If we rename the permutation σ^{-1} to τ and use the facts that

- as σ runs over all permutations exacly once, σ^{-1} runs over all permutations exactly once and
- $\operatorname{sgn} \sigma = \operatorname{sgn} \sigma^{-1}$

we recover

$$\det A^t = \sum_{\tau \in P_n} \operatorname{sgn} \tau \ \prod_{j=1}^{n} A_{j\, \tau(j)} = \det A$$

Implications of Properties E, M, A ,D, P, T

1) If any two rows of a matrix A are the same, then $\det A = 0$ because

$$\det \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{b} \\ \vdots \\ \vec{b} \\ \vdots \\ \vec{a}_n \end{bmatrix} = - \det \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{b} \\ \vdots \\ \vec{b} \\ \vdots \\ \vec{a}_n \end{bmatrix}$$

On the right hand side, the two rows containing $\vec{b}$ have been interchanged.

2) Thanks to Property E, a determinant may be expanded along any row. That is, for any $1 \le i \le n$,

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det M_{ij}$$

where M_{ij} is the $(n-1)\times(n-1)$ matrix formed by deleting from the original matrix A the row and column containing a_{ij}. To get the signs $(-1)^{i+j}$ right, you just have to remember the checkerboard

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Example III.23 If we expand the matrix of Example III.21 along its second row, we get

$$\begin{aligned} \det\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} &= -1\times\det\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} + 0\times\det\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} - 2\times\det\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \\ &= -1\times(2-6)+0-2(2-6)=12 \end{aligned}$$

3) Thanks to Property T, a determinant may be expanded along any column too. That is, for any $1\le j\le n$,

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det M_{ij}$$

Example III.24 If we expand the matrix of Example III.21 along its second column, we get

$$\begin{aligned} \det\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} &= -2\times\det\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} + 0\times\det\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} - 2\times\det\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \\ &= -2\times(1-6)+0-2(2-3)=12 \end{aligned}$$

3) We can now use Gaussian elimination to evaluate determinants. Properties A, E, M say that if you take a matrix U and apply to it a row operation, the resulting matrix V obeys

$$\begin{array}{ll} \det U = \det V & \text{if the row operation is } (i)\to(i)+k(j) \text{ for some } j\ne i \\ \det U = -\det V & \text{if the row operation is } (i)\leftrightarrow(j) \text{ for some } j\ne i \\ \det U = \frac{1}{k}\det V & \text{if the row operation is } (i)\to k(i) \text{ for some } k\ne 0 \end{array}$$

These properties, combined with Gaussian elimination, allow us to relate the determinant of any given matrix to the determinant of a triangular matrix, that is trivially computed using Property D.

Example III.25

$$\det\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 4 & 0 & 3 \\ 1 & 3 & 5 & 6 \\ 1 & 3 & 3 & 9 \end{bmatrix} \overset{A}{=} \det\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & -2 & -1 \\ 0 & 1 & 4 & 4 \\ 0 & 1 & 2 & 7 \end{bmatrix}\begin{matrix} (1) \\ (2)-2(1) \\ (3)-(1) \\ (4)-(1) \end{matrix} \overset{E}{=} -\det\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & -2 & -1 \\ 0 & 1 & 2 & 7 \end{bmatrix}\begin{matrix} (1) \\ (3) \\ (2) \\ (4) \end{matrix}$$

$$\overset{A}{=} -\det\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & -2 & 3 \end{bmatrix}\begin{matrix} (1) \\ (2) \\ (3) \\ (4)-(2) \end{matrix} \overset{A}{=} -\det\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & 0 & 4 \end{bmatrix}\begin{matrix} (1) \\ (2) \\ (3) \\ (4)-(3) \end{matrix}$$

$$\overset{D}{=} -1\times1\times(-2)\times4 = 8$$

Example III.26 Let's redo the computation of the determinant in Example III.22, using row operations as well as the fact that we may expand along any row or column. Use Cj to denote expansion along column j and Rj to denote expansion along row j.

$$\det\begin{bmatrix}1&2&3&4\\4&3&2&1\\2&3&3&3\\7&8&9&12\end{bmatrix} \overset{A}{=} \det\begin{bmatrix}1&2&3&4\\0&-5&-10&-15\\0&-1&-3&-5\\0&-6&-12&-16\end{bmatrix}\begin{matrix}(1)\\(2)-4(1)\\(3)-2(1)\\(4)-7(1)\end{matrix} \overset{C1}{=} \det\begin{bmatrix}-5&-10&-15\\-1&-3&-5\\-6&-12&-16\end{bmatrix}$$

$$\overset{M}{=} -5\det\begin{bmatrix}1&2&3\\-1&-3&-5\\-6&-12&-16\end{bmatrix} \overset{A}{=} -5\det\begin{bmatrix}1&2&3\\0&-1&-2\\0&0&2\end{bmatrix}\begin{matrix}(1)\\(2)+(1)\\(3)+6(1)\end{matrix}$$

$$\overset{D}{=} (-5)\{1\times(-1)\times 2\} = 10$$

4) The determinant provides a test for invertibility: Let A be an $n\times n$ matrix. Then

$$\boxed{A \text{ is invertible} \qquad \Longleftrightarrow \qquad \det A \neq 0}$$

Proof: Let R be the triangular matrix that results from the application of Gaussian elimination to the matrix A. Then, by properties E, M and A,

$$\det A = (\text{nonzero number}) \det R$$

So

$$\begin{aligned}\det A\neq 0 &\iff \det R\neq 0\\ &\iff \text{the diagonal entries } R_{jj} \text{ of } R \text{ are all nonzero} && \text{(Property D)}\\ &\iff \operatorname{rank} R = n && \text{(See the definition of rank in §II.3)}\\ &\iff A \text{ is invertible} && \text{(Property 6 of §III.5)}\end{aligned}$$

5)

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

are easy consequences of property P.

Exercises for §III.8

1) Evaluate the determinant of each of the following matrices by expanding along some row other than the first

$a)$ $\begin{bmatrix}2&1&5\\1&0&3\\-1&2&0\end{bmatrix}$ $\quad b)$ $\begin{bmatrix}7&-1&5\\3&4&-5\\2&3&0\end{bmatrix}$ $\quad c)$ $\begin{bmatrix}1&a&b\\0&1&c\\0&0&1\end{bmatrix}$

2) Evaluate the determinant of each of the following matrices using row reduction

$a)$ $\begin{bmatrix}1&-1&1&-1\\1&2&4&8\\1&-2&4&-8\\1&1&1&1\end{bmatrix}$ $\quad b)$ $\begin{bmatrix}0&1&-2&3\\-1&0&1&2\\2&-1&0&1\\-3&-2&-1&0\end{bmatrix}$

$c)$ $\begin{bmatrix}3&1&2&0\\-2&-1&5&-2\\1&-3&1&1\\4&1&2&-3\end{bmatrix}$ $\quad d)$ $\begin{bmatrix}1&1&1&1\\1&2&4&8\\1&3&9&27\\1&4&16&64\end{bmatrix}$

3) For which values of the parameter λ does the matrix

$$\begin{bmatrix} -\frac{1}{2}-\lambda & \frac{1}{2} \\ -\frac{3}{2} & -\frac{5}{2}-\lambda \end{bmatrix}$$

have an inverse?

4) For which values of the parameter λ does the system of linear equations

$$\begin{aligned} 2x_1 - x_2 &= \lambda x_1 \\ 2x_1 + 5x_2 &= \lambda x_2 \end{aligned}$$

have a solution other than $x_1 = x_2 = 0$?

§III.9 Determinants – Applications

Testing for Invertibility

One of the main uses for determinants is testing for invertibility. A square matrix A is invertible if and only if $\det A \neq 0$. Equivalently, (see property 4 of inverses given in §III.5), the system of equations $A\vec{x} = \vec{0}$ has a nonzero solution $\vec{x}$ if and only if $\det A = 0$. This is heavily used in the eigenvalue/eigenvector problem in the next chapter.

Concise Formulae

a) If A is a square matrix with $\det A \neq 0$,

$$\left(A^{-1}\right)_{ij} = (-1)^{i+j}\frac{\det M_{ji}}{\det A}$$

where M_{ji} is the matrix gotten by deleting from A the row and column containing A_{ji}.

WARNING: This formula is useful, for example, for studying the dependence of A^{-1} on matrix elements of A. But it is usually computationally much more efficient to solve $AB = I$ by Gaussian elimination than it is to apply this formula.

Proof: The foundation for this formula is the expansion formula

$$\det A = \sum_i (-1)^{i+j} A_{ji} \det M_{ji}$$

It is true for any j. This was implication number 2 in the last section. Let B be the matrix with

$$B_{ij} = (-1)^{i+j} \det M_{ji}$$

We have to show that $(AB)_{kj}$ is $\det A$ if $j = k$ and 0 if $j \neq k$. But

$$(AB)_{kj} = \sum_i A_{ki} B_{ij} = \sum_i (-1)^{i+j} A_{ki} \det M_{ji}$$

If $k = j$ this is precisely the formula for $\det A$ given above. If $k \neq j$, this is the expansion along row j for the determinant of another matrix $\tilde{A}$. This other matrix is constructed from A by replacing row j of A by row k of A. Row numbers j and k of $\tilde{A}$ are identical so that $\det \tilde{A} = 0$ by implication number 1 of Property E.

Example III.27 Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

Then

$$\begin{array}{lll} \det M_{11} = \det \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = -7 & \det M_{12} = \det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = -5 & \det M_{13} = \det \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = -1 \\ \det M_{21} = \det \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} = -2 & \det M_{22} = \det \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = -1 & \det M_{23} = \det \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = 1 \\ \det M_{31} = \det \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = 1 & \det M_{32} = \det \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} = 2 & \det M_{33} = \det \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = 1 \end{array}$$

so that

$$\det A = 1 \times M_{11} - 1 \times M_{12} + 1 \times M_{13} = -7 + 5 - 1 = -3$$

and

$$A^{-1} = \frac{1}{-3} \begin{bmatrix} -7 & 2 & 1 \\ 5 & -1 & -2 \\ -1 & -1 & 1 \end{bmatrix}$$

b) (Cramer's rule) If A is a square matrix with $\det A \neq 0$, the solution to $A\vec{x} = \vec{b}$ is

$$x_i = \frac{\det \big(A \text{ with } i^{\text{th}} \text{ column replaced by } \vec{b} \big)}{\det A}$$

WARNING: This formula is useful, for example, for studying the dependence of $A^{-1}\vec{b}$ on matrix elements of A and $\vec{b}$. But it is usually computationally much more efficient to solve $A\vec{x} = \vec{b}$ by Gaussian elimination than it is to apply this formula.

Proof: Since $\vec{x} = A^{-1}\vec{b}$,

$$\det A \times x_i = \det A \times \big(A^{-1}\vec{b} \big)_i = \det A \sum_j A^{-1}_{ij} \vec{b}_j = \sum_j (-1)^{i+j} b_j \det M_{ji}$$

The right hand side is the expansion along column i of the determinant of the matrix constructed from A by replacing column number i with $\vec{b}$.

Example III.28 Consider the system of equations of Example II.2.

$$\begin{aligned} x_1 + \ x_2 + \ x_3 &= 4 \\ x_1 + 2x_2 + 3x_3 &= 9 \\ 2x_1 + 3x_2 + \ x_3 &= 7 \end{aligned}$$

This system can be written in the form $A\vec{x} = \vec{b}$ with

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix}$$

According to Cramer's rule

$$x_1 = \frac{1}{-3}\det\begin{bmatrix}4&1&1\\9&2&3\\7&3&1\end{bmatrix} = -\frac{1}{3}\det\begin{bmatrix}4&1&1\\-3&-1&0\\3&2&0\end{bmatrix}\begin{matrix}(1)\\(2)-3(1)\\(3)-(1)\end{matrix} = -\frac{-3}{3} = 1$$

$$x_2 = \frac{1}{-3}\det\begin{bmatrix}1&4&1\\1&9&3\\2&7&1\end{bmatrix} = -\frac{1}{3}\det\begin{bmatrix}1&4&1\\0&5&2\\0&-1&-1\end{bmatrix}\begin{matrix}(1)\\(2)-(1)\\(3)-2(1)\end{matrix} = -\frac{-3}{3} = 1$$

$$x_3 = \frac{1}{-3}\det\begin{bmatrix}1&1&4\\1&2&9\\2&3&7\end{bmatrix} = -\frac{1}{3}\det\begin{bmatrix}1&1&4\\0&1&5\\0&1&-1\end{bmatrix}\begin{matrix}(1)\\(2)-(1)\\(3)-2(1)\end{matrix} = -\frac{-6}{3} = 2$$

c) The cross product

$$\vec{a}\times\vec{b} = \begin{bmatrix}\hat{\imath}&\hat{\jmath}&\hat{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{bmatrix}$$

Areas and Volumes

We have aready seen in §I.5 that areas of parallelograms and volumes of parallelopipeds are given by determinants.

$$\text{volume of parallelogram with sides } \vec{a},\ \vec{b} \ = \ \left|\det\begin{bmatrix}a_1&a_2\\b_1&b_2\end{bmatrix}\right|$$

$$\text{volume of parallelopiped with edges } \vec{a},\ \vec{b},\ \vec{c} \ = \ \left|\det\begin{bmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{bmatrix}\right|$$

As a result, the standard change of variables formula for integrals in more than one dimension involves a determinant.

§III.10 Worked Problems

Questions

1) Compute the following matrix products:

(a) $\begin{bmatrix}1&0\\3&5\end{bmatrix}\begin{bmatrix}0&3\\1&4\end{bmatrix}$ (b) $\begin{bmatrix}1&2&0\\0&-1&1\end{bmatrix}\begin{bmatrix}2&1\\0&1\\3&0\end{bmatrix}$ (c) $\begin{bmatrix}1&2&3\\0&0&1\\1&0&0\end{bmatrix}\begin{bmatrix}0&4&3\\2&1&1\\8&0&4\end{bmatrix}$

(d) $\begin{bmatrix}1&0\\2&3\\5&0\end{bmatrix}\begin{bmatrix}1&0&3&5\\0&1&1&2\end{bmatrix}$ (e) $\begin{bmatrix}3&1&0\end{bmatrix}\begin{bmatrix}2&3\\3&0\\0&5\end{bmatrix}$ (f) $\begin{bmatrix}2&3\end{bmatrix}\begin{bmatrix}4\\5\end{bmatrix}$

(g) $\begin{bmatrix}2&1\\0&4\end{bmatrix}\begin{bmatrix}0&1\\1&2\end{bmatrix}\begin{bmatrix}-1&0&1\\2&1&0\end{bmatrix}$ (h) $\begin{bmatrix}2&-1&0\\0&0&3\end{bmatrix}\begin{bmatrix}3&3&5\\1&0&0\\0&2&0\end{bmatrix}\begin{bmatrix}3&2\\1&0\\2&0\end{bmatrix}$ (i) $\begin{bmatrix}1&3\\5&4\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}$

(j) $\begin{bmatrix}x&y\\2&3\end{bmatrix}\begin{bmatrix}a&b\\-1&4\end{bmatrix}$

2) Let

$$A = \begin{bmatrix}1&2&3\\4&5&6\end{bmatrix} \qquad B = \begin{bmatrix}-3&1&2\\-3&2&0\end{bmatrix}$$

(a) Compute $2A$, $3B$, $2A+3B$ and $3(2A+3B)$.
(b) Compute $6A$, $9B$ and $6A+9B$.
(c) Why are the last results in parts (a) and (b) the same?

3) Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 6 & 1 \\ 1 & -1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 4 & 1 \end{bmatrix}$$

(a) Compute AB, $(AB)C$, BC and $A(BC)$ and verify that $(AB)C = A(BC)$. So it is not necessary to bracket ABC.
(b) Can the order of the factors in the product ABC be changed?

4) Let

$$A = \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}$$

(a) Compute $A^2 + 2AB + B^2$.
(b) Compute $(A+B)^2$.
(c) Account for the difference between the answers to parts a and b.

5) Define

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(a) Find all matrices that commute with A. That is, find all matrices M that obey $AM = MA$.
(b) Find all matrices that commute with B. That is, find all matrices M that obey $BM = MB$.
(c) Find all matrices that commute with both A and B. That is, find all matrices M that obey $AM = MA$ and $BM = MB$.

6) Suppose the matrix B obeys $B^7 - 3B + I = 0$. Is B invertible? If so, what is its inverse?

7) State whether each of the following statments is true or false. In each case give a *brief* reason.
(a) No square matrix with real entries can obey $A^2 = -I$.
(b) The only 2×2 matrix that obeys $A^2 = 0$ is $A = 0$.
(c) The only 2×2 matrices that obey $A^2 = A$ are $A = 0$ and $A = I$.

8) Suppose that a taxicab company uses the following strategy to maintain a fleet of fixed size: on December 31 of each year it sells 25% of the cars that are one year old (the lemons), 50% of all two year cars and all three year old cars; the next day, January 1, it buys enough new cars to replace those sold the previous day.
(a) Let $\vec{x}(n)$ be the vector which gives the number of cars that are in each age category during year n. Find a 3×3 matrix B such that $\vec{x}(n+1) = B\vec{x}(n)$.
(b) Suppose the company starts with a fleet consisting of entirely third year cars. What proportion of the fleet will consist of third year cars in each of the following four years?
(c) Find all equilibrium fleet vectors. That is, all vectors $\vec{x}(n)$ obeying $\vec{x}(n+1) = \vec{x}(n)$.
(d) Find a matrix C that moves the fleet ahead two years. That is $\vec{x}(n+2) = C\vec{x}(n)$.
(e) Are there any fleet vectors which repeat themselves every two years, but not every year?

9) Determine whether or not each of the following functions is a linear transformation.
(a) $f(x,y,z) = 3x - 2y + 5z$ (b) $f(x,y,z) = -2y + 9z - 12$ (c) $f(x,y,z) = x^2 + y^2 + z^2$
(d) $f(x,y,z) = [x+y, x-y, 0]$ (e) $f(x,y,z) = [x+y, x-y, 1]$ (f) $f(x,y,z) = [x, y, z^2]$

10) Suppose that a linear transformation maps $[1,1]$ to $[4,7]$ and $[1,-1]$ to $[8,3]$. What vector does it map $[5,14]$ to?

11) Is it possible for a linear transformation to map $[1,2]$ to $[1,0,-1]$, $[3,4]$ to $[1,2,3]$ and $[5,8]$ to $[3,1,6]$?

12) Find the matrices representing the linear transformations

(a) $f(x,y) = [x+2y, y-x, x+y]$ (b) $g(x,y,z) = [x-y-z, 2y+5z-3x]$
(c) $f\big(g(x,y,z)\big)$ (d) $g\big(f(x,y)\big)$

13) A linear transformation maps $\hat\imath$ to $-\hat\jmath$, $\hat\jmath$ to $\hat k$ and $\hat k$ to $-\hat\imath$.

(a) Find the matrix of the linear transformation.
(b) The linear transformation is a rotation. Find the axis of rotation.
(c) Find the angle of rotation.
(d) Show that this linear transformation really is a rotation.

14) Determine the matrices of the reflections in the following planes in $\mathbb{R}^3$

(a) $x+y+z=0$ (b) $2x-2y-z=0$

15) A solid body is rotating about an axis which passes through the origin and has direction $\vec\Omega = \Omega_1\hat\imath + \Omega_2\hat\jmath + \Omega_3\hat k$. The rate of rotation is $\|\vec\Omega\|$ radians per second. Denote by $\vec x$ the coordinates, at some fixed time, of a point fixed to the body and by $\vec v$ the velocity vector of the point at that time. Find a matrix A such that $\vec v = A\vec x$.

16) Find the matrix of the linear transformation on $\mathbb{R}^2$ which

(a) rotates each vector $\vec x$ clockwise by 45°.
(b) reflects each vector $\vec x$ in the x–axis and then rotates the result counterclockwise by 90°.
(c) reflects each vector $\vec x$ about the line $x=y$ and then projects the result onto the x–axis.

17) The matrix on the left below is the inverse of the matrix on the right below. Fill in the missing entries.

$$\begin{bmatrix} 4 & 3 & 2 & \bullet \\ 5 & 4 & 3 & 1 \\ -2 & -2 & -1 & -1 \\ 11 & 6 & 4 & 3 \end{bmatrix} \qquad \begin{bmatrix} \bullet & -1 & 1 & 0 \\ 7 & -3 & 1 & -1 \\ -10 & 5 & \bullet & \bullet \\ -8 & 3 & \bullet & 1 \end{bmatrix}$$

18) Find, if possible, the inverses of each of the following matrices.

$$(a)\ \begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix} \quad (b)\ \begin{bmatrix} 2 & 5 \\ 4 & 8 \end{bmatrix} \quad (c)\ \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \quad (d)\ \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ 1 & 3 & 9 \end{bmatrix} \quad (e)\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{bmatrix}$$

$$(f)\ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad (g)\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}$$

19) Let

$$B = \begin{bmatrix} -1 & 2 & p \\ 0 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

(a) For which values of p does B have an inverse?
(b) Find B^{-1}, for those values of p.

20) Suppose that, for some square matrix A, the series $\sum_{n=0}^\infty A^n = I + A + A^2 + A^3 + \cdots$ converges. (In series notation, A^0 is defined to be I.) Show that

$$(I-A)^{-1} = \sum_{n=0}^\infty A^n$$

21) Suppose that some square matrix obeys $A^n = 0$ for some positive integer.
(a) Find the inverse of A.
(b) Find the inverse of $I - A$.

22) Suppose that L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Prove that its inverse, if it exists, is also linear.

23) Suppose that $\det\begin{bmatrix}3 & 2 & p\\ 0 & p & 1\\ 1 & 0 & 2\end{bmatrix} = 10$. What are the possible values of p?

24) Let

$$A = \begin{bmatrix}1 & 3 & 5 & *\\ 0 & 4 & 0 & 6\\ 0 & 1 & 0 & 2\\ 3 & * & 7 & 8\end{bmatrix}$$

where the $*$'s denote unknown entries. Find all possible values of $\det A$.

25) Suppose that the 3×3 matrix A obeys $\det A = 5$. Compute (a) $\det(4A)$ (b) $\det(A^2)$ (c) $\det(4A^2)$

26) Suppose that the 6×6 matrix A obeys $A^4 = 2A$. Find all possible values of $\det A$.

27) Evaluate

$$\det\begin{bmatrix}1 & a & a^2 & a^3\\ a & a^2 & a^3 & 1\\ a^2 & a^3 & 1 & a\\ a^3 & 1 & a & a^2\end{bmatrix}$$

Solutions

1) Compute the following matrix products:

(a) $\begin{bmatrix}1 & 0\\ 3 & 5\end{bmatrix}\begin{bmatrix}0 & 3\\ 1 & 4\end{bmatrix}$ (b) $\begin{bmatrix}1 & 2 & 0\\ 0 & -1 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\ 0 & 1\\ 3 & 0\end{bmatrix}$ (c) $\begin{bmatrix}1 & 2 & 3\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}\begin{bmatrix}0 & 4 & 3\\ 2 & 1 & 1\\ 8 & 0 & 4\end{bmatrix}$

(d) $\begin{bmatrix}1 & 0\\ 2 & 3\\ 5 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & 3 & 5\\ 0 & 1 & 1 & 2\end{bmatrix}$ (e) $\begin{bmatrix}3 & 1 & 0\end{bmatrix}\begin{bmatrix}2 & 3\\ 3 & 0\\ 0 & 5\end{bmatrix}$ (f) $\begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}4\\ 5\end{bmatrix}$

(g) $\begin{bmatrix}2 & 1\\ 0 & 4\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}-1 & 0 & 1\\ 2 & 1 & 0\end{bmatrix}$ (h) $\begin{bmatrix}2 & -1 & 0\\ 0 & 0 & 3\end{bmatrix}\begin{bmatrix}3 & 3 & 5\\ 1 & 0 & 0\\ 0 & 2 & 0\end{bmatrix}\begin{bmatrix}3 & 2\\ 1 & 0\\ 2 & 0\end{bmatrix}$ (i) $\begin{bmatrix}1 & 3\\ 5 & 4\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}$

(j) $\begin{bmatrix}x & y\\ 2 & 3\end{bmatrix}\begin{bmatrix}a & b\\ -1 & 4\end{bmatrix}$

Solution.
(a)

$$\begin{bmatrix}1 & 0\\ 3 & 5\end{bmatrix}\begin{bmatrix}0 & 3\\ 1 & 4\end{bmatrix} = \begin{bmatrix}1\times 0+0\times 1 & 1\times 3+0\times 4\\ 3\times 0+5\times 1 & 3\times 3+5\times 4\end{bmatrix} = \boxed{\begin{bmatrix}0 & 3\\ 5 & 29\end{bmatrix}}$$

(b)

$$\begin{bmatrix}1 & 2 & 0\\ 0 & -1 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\ 0 & 1\\ 3 & 0\end{bmatrix} = \begin{bmatrix}1\times 2+2\times 0+0\times 3 & 1\times 1+2\times 1+0\times 0\\ 0\times 2-1\times 0+1\times 3 & 0\times 1-1\times 1+1\times 0\end{bmatrix} = \boxed{\begin{bmatrix}2 & 3\\ 3 & -1\end{bmatrix}}$$

(c)

$$\begin{bmatrix}1&2&3\\0&0&1\\1&0&0\end{bmatrix}\begin{bmatrix}0&4&3\\2&1&1\\8&0&4\end{bmatrix}=\begin{bmatrix}1\times0+2\times2+3\times8 & 1\times4+2\times1+3\times0 & 1\times3+2\times1+3\times4\\0\times0+0\times2+1\times8 & 0\times4+0\times1+1\times0 & 0\times3+0\times1+1\times4\\1\times0+0\times2+0\times8 & 1\times4+0\times1+0\times0 & 1\times3+0\times1+0\times4\end{bmatrix}$$
$$=\boxed{\begin{bmatrix}28&6&17\\8&0&4\\0&4&3\end{bmatrix}}$$

(d)

$$\begin{bmatrix}1&0\\2&3\\5&0\end{bmatrix}\begin{bmatrix}1&0&3&5\\0&1&1&2\end{bmatrix}=\begin{bmatrix}1\times1+0\times0 & 1\times0+0\times1 & 1\times3+0\times1 & 1\times5+0\times2\\2\times1+3\times0 & 2\times0+3\times1 & 2\times3+3\times1 & 2\times5+3\times2\\5\times1+0\times0 & 5\times0+0\times1 & 5\times3+0\times1 & 5\times5+0\times2\end{bmatrix}$$
$$=\boxed{\begin{bmatrix}1&0&3&5\\2&3&9&16\\5&0&15&25\end{bmatrix}}$$

(e)

$$\begin{bmatrix}3&1&0\end{bmatrix}\begin{bmatrix}2&3\\3&0\\0&5\end{bmatrix}=\begin{bmatrix}3\times2+1\times3+0\times0 & 3\times3+1\times0+0\times5\end{bmatrix}=\boxed{\begin{bmatrix}9&9\end{bmatrix}}$$

(f)

$$\begin{bmatrix}2&3\end{bmatrix}\begin{bmatrix}4\\5\end{bmatrix}=\begin{bmatrix}2\times4+3\times5\end{bmatrix}=\boxed{\begin{bmatrix}23\end{bmatrix}}$$

(g)

$$\begin{bmatrix}2&1\\0&4\end{bmatrix}\begin{bmatrix}0&1\\1&2\end{bmatrix}\begin{bmatrix}-1&0&1\\2&1&0\end{bmatrix}=\begin{bmatrix}1&4\\4&8\end{bmatrix}\begin{bmatrix}-1&0&1\\2&1&0\end{bmatrix}=\boxed{\begin{bmatrix}7&4&1\\12&8&4\end{bmatrix}}$$

(h)

$$\begin{bmatrix}2&-1&0\\0&0&3\end{bmatrix}\begin{bmatrix}3&3&5\\1&0&0\\0&2&0\end{bmatrix}\begin{bmatrix}3&2\\1&0\\2&0\end{bmatrix}=\begin{bmatrix}5&6&10\\0&6&0\end{bmatrix}\begin{bmatrix}3&2\\1&0\\2&0\end{bmatrix}=\boxed{\begin{bmatrix}41&10\\6&0\end{bmatrix}}$$

(i)

$$\begin{bmatrix}1&3\\5&4\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\boxed{\begin{bmatrix}a+3c&b+3d\\5a+4c&5b+4d\end{bmatrix}}$$

(j)

$$\begin{bmatrix}x&y\\2&3\end{bmatrix}\begin{bmatrix}a&b\\-1&4\end{bmatrix}=\boxed{\begin{bmatrix}ax-y&bx+4y\\2a-3&2b+12\end{bmatrix}}$$

2) Let

$$A=\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\qquad B=\begin{bmatrix}-3&1&2\\-3&2&0\end{bmatrix}$$

(a) Compute $2A$, $3B$, $2A+3B$ and $3(2A+3B)$.
(b) Compute $6A$, $9B$ and $6A+9B$.
(c) Why are the last results in parts (a) and (b) the same?

Solution. (a)

$$2A = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix} \quad 3B = \begin{bmatrix} -9 & 3 & 6 \\ -9 & 6 & 0 \end{bmatrix} \quad 2A + 3B = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix} + \begin{bmatrix} -9 & 3 & 6 \\ -9 & 6 & 0 \end{bmatrix} = \begin{bmatrix} -7 & 7 & 12 \\ -1 & 16 & 12 \end{bmatrix}$$

$$3(2A + 3B) = 3 \begin{bmatrix} -7 & 7 & 12 \\ -1 & 16 & 12 \end{bmatrix} = \boxed{\begin{bmatrix} -21 & 21 & 36 \\ -3 & 48 & 36 \end{bmatrix}}$$

(b)

$$6A = \begin{bmatrix} 6 & 12 & 18 \\ 24 & 30 & 36 \end{bmatrix} \quad 9B = \begin{bmatrix} -27 & 9 & 18 \\ -27 & 18 & 0 \end{bmatrix}$$

$$6A + 9B = \begin{bmatrix} 6 & 12 & 18 \\ 24 & 30 & 36 \end{bmatrix} + \begin{bmatrix} -27 & 9 & 18 \\ -27 & 18 & 0 \end{bmatrix} = \boxed{\begin{bmatrix} -21 & 21 & 36 \\ -3 & 48 & 36 \end{bmatrix}}$$

(c) Applying properties 3 and 5 of the "Basic Properties of Matrix Operations" given in §III.1, we have,

$\boxed{3(2A + 3B) \overset{3}{=} 3(2A) + 3(3B) \overset{5}{=} (3 \times 2)A + (3 \times 3)B = 6A + 9B}$

3) Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 6 & 1 \\ 1 & -1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 4 & 1 \end{bmatrix}$$

(a) Compute AB, $(AB)C$, BC and $A(BC)$ and verify that $(AB)C = A(BC)$. So it is not necessary to bracket ABC.
(b) Can the order of the factors in the product ABC be changed?

Solution. (a)

$$AB = \begin{bmatrix} 4 & 4 & 1 \\ 5 & 19 & 3 \end{bmatrix} \quad (AB)C = \begin{bmatrix} 4 & 4 & 1 \\ 5 & 19 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 2 \\ 4 & 1 \end{bmatrix} = \boxed{\begin{bmatrix} 20 & 17 \\ 46 & 51 \end{bmatrix}}$$

$$BC = \begin{bmatrix} 16 & 17 \\ 2 & 0 \end{bmatrix} \qquad A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 16 & 17 \\ 2 & 0 \end{bmatrix} = \boxed{\begin{bmatrix} 20 & 17 \\ 46 & 51 \end{bmatrix}}$$

Yup. They're the same.
(b) The matrix product $\boxed{BAC \text{ is not defined}}$ because the matrix product BA is not defined owing to a matrix size mismatch.

4) Let

$$A = \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}$$

(a) Compute $A^2 + 2AB + B^2$.
(b) Compute $(A + B)^2$.
(c) Account for the difference between the answers to parts a and b.

Solution. (a)

$$A^2 = \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} 7 & -10 \\ -15 & 22 \end{bmatrix}$$

$$2AB = 2 \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} = 2 \begin{bmatrix} -1 & 7 \\ 3 & -15 \end{bmatrix} = \begin{bmatrix} -2 & 14 \\ 6 & -30 \end{bmatrix}$$

$$B^2 = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ 0 & 9 \end{bmatrix}$$

$$A^2 + 2AB + B^2 = \begin{bmatrix} 7 & -10 \\ -15 & 22 \end{bmatrix} + \begin{bmatrix} -2 & 14 \\ 6 & -30 \end{bmatrix} + \begin{bmatrix} 1 & -4 \\ 0 & 9 \end{bmatrix} = \boxed{\begin{bmatrix} 6 & 0 \\ -9 & 1 \end{bmatrix}}$$

(b)

$$A+B=\begin{bmatrix}-1&2\\3&-4\end{bmatrix}+\begin{bmatrix}1&-1\\0&3\end{bmatrix}=\begin{bmatrix}0&1\\3&-1\end{bmatrix}$$

$$(A+B)^2=\begin{bmatrix}0&1\\3&-1\end{bmatrix}\begin{bmatrix}0&1\\3&-1\end{bmatrix}=\boxed{\begin{bmatrix}3&-1\\-3&4\end{bmatrix}}$$

(c) $(A+B)^2 = A(A+B)+B(A+B) = A^2+AB+BA+B^2$, so the answer to part a minus the answer to part b ought to be

$$\begin{aligned}(A^2+2AB+B^2)-(A^2+AB+BA+B^2) &= AB-BA\\ &= \begin{bmatrix}-1&2\\3&-4\end{bmatrix}\begin{bmatrix}1&-1\\0&3\end{bmatrix}-\begin{bmatrix}1&-1\\0&3\end{bmatrix}\begin{bmatrix}-1&2\\3&-4\end{bmatrix}\\ &= \begin{bmatrix}-1&7\\3&-15\end{bmatrix}-\begin{bmatrix}-4&6\\9&-12\end{bmatrix}=\boxed{\begin{bmatrix}3&1\\-6&-3\end{bmatrix}}\end{aligned}$$

This is indeed the difference.

5) Define

$$A=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\qquad B=\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

(a) Find all matrices that commute with A. That is, find all matrices M that obey $AM = MA$.
(b) Find all matrices that commute with B. That is, find all matrices M that obey $BM = MB$.
(c) Find all matrices that commute with both A and B. That is, find all matrices M that obey $AM = MA$ and $BM = MB$.

Solution. (a) In order for both AM and MA to be defined, M must be a 2×2 matrix. Let

$$M=\begin{bmatrix}a&b\\c&d\end{bmatrix}$$

Then

$$AM=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}a&b\\-c&-d\end{bmatrix}$$

$$MA=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}1&0\\0&-1\end{bmatrix}=\begin{bmatrix}a&-b\\c&-d\end{bmatrix}$$

These are the same if and only if $b = -b$ and $c = -c$ which in turn is the case if and only if $b = c = 0$. So we need M to be of the form $\boxed{\begin{bmatrix}a&0\\0&d\end{bmatrix}}$ for some numbers a and d.

(b) In order for both BM and MB to be defined, M must be a 2×2 matrix. Let

$$M=\begin{bmatrix}a&b\\c&d\end{bmatrix}$$

Then

$$BM=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}c&d\\a&b\end{bmatrix}$$

$$MB=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}b&a\\d&c\end{bmatrix}$$

These are the same if and only if $c = b$ and $a = d$. So we need M to be of the form $\boxed{\begin{bmatrix}a&b\\b&a\end{bmatrix}}$ for some numbers a and b.

(c) In order to satisfy the conditions of both part a and part b, we need $b = c = 0$ and $a = d$, so M must be of the form $\boxed{a\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}}$ for some number a.

6) Suppose the matrix B obeys $B^7 - 3B + I = 0$. Is B invertible? If so, what is its inverse?

Solution. Because $B(3I - B^6) = 3B - B^7 = I$, the matrix B $\boxed{\text{is invertible}}$ with inverse $\boxed{3I - B^6}$.

7) State whether each of the following statments is true or false. In each case give a *brief* reason.
(a) No square matrix with real entries can obey $A^2 = -I$.
(b) The only 2×2 matrix that obeys $A^2 = 0$ is $A = 0$.
(c) The only 2×2 matrices that obey $A^2 = A$ are $A = 0$ and $A = I$.

Solution. (a) This statement is $\boxed{\text{false}}$. For example

$$A = \begin{bmatrix}0 & a\\ b & 0\end{bmatrix} \quad\Longrightarrow\quad A^2 = \begin{bmatrix}ab & 0\\ 0 & ab\end{bmatrix}$$

So if $ab = -1$ (e.g. $a = 1,\ b = -1$), $A^2 = -I$.
(b) This statement is $\boxed{\text{false}}$. For example

$$A = \begin{bmatrix}0 & a\\ b & 0\end{bmatrix} \quad\Longrightarrow\quad A^2 = \begin{bmatrix}ab & 0\\ 0 & ab\end{bmatrix}$$

So if $ab = 0$ (e.g. $a = 1,\ b = 0$), $A^2 = 0$.
(c) This statement is $\boxed{\text{false}}$. For example

$$A = \begin{bmatrix}a & 0\\ 0 & b\end{bmatrix} \quad\Longrightarrow\quad A^2 = \begin{bmatrix}a^2 & 0\\ 0 & b^2\end{bmatrix}$$

So if $a^2 = a$ and $b^2 = b$ (e.g. $a = 1,\ b = 0$), $A^2 = A$.

8) Suppose that a taxicab company uses the following strategy to maintain a fleet of fixed size: on December 31 of each year it sells 25% of the cars that are one year old (the lemons), 50% of all two year cars and all three year old cars; the next day, January 1, it buys enough new cars to replace those sold the previous day.
(a) Let $\vec{x}(n)$ be the vector which gives the number of cars that are in each age category during year n. Find a 3×3 matrix B such that $\vec{x}(n+1) = B\vec{x}(n)$.
(b) Suppose the company starts with a fleet consisting of entirely third year cars. What proportion of the fleet will consist of third year cars in each of the following four years?
(c) Find all equilibrium fleet vectors. That is, all vectors $\vec{x}(n)$ obeying $\vec{x}(n+1) = \vec{x}(n)$.
(d) Find a matrix C that moves the fleet ahead two years. That is $\vec{x}(n+2) = C\vec{x}(n)$.
(e) Are there any fleet vectors which repeat themselves every two years, but not every year?

Solution. (a) Pretend that it is now December 31 of year n. We have $x(n)_1$ cars that are (almost) one year old, $x(n)_2$ cars that are two years old and $x(n)_3$ cars that are three years old. So, on December 31 of year n, we sell $.25x_1(n) + .50x_2(n) + 1.0x_3(n)$ cars. The next day, January 1 of year $n+1$ we buy the same number, $.25x_1(n) + .50x_2(n) + 1.0x_3(n)$, of replacement cars. During year $n+1$, these cars are all between zero and one year old. So

$$x_1(n+1) = .25x_1(n) + .50x_2(n) + 1.0x_3(n)$$

Of the $x_1(n)$ cars that had their first birthday on January 1, $.75x_1(n)$ remain with the company. During year $n+1$, these cars are all between one and two years old. So

$$x_2(n+1) = .75x_1(n)$$

Similarly,

$$x_3(n+1) = .50x_2(n)$$

In summary,

$$\begin{aligned}
\big(B\vec{x}(n)\big)_1 &= x_1(n+1) = .25x_1(n) + .50x_2(n) + 1.0x_3(n)\\
\big(B\vec{x}(n)\big)_2 &= x_2(n+1) = .75x_1(n)\\
\big(B\vec{x}(n)\big)_3 &= x_3(n+1) = \qquad\qquad .50x_2(n)
\end{aligned}$$

So

$$\boxed{B = \begin{bmatrix} .25 & .5 & 1 \\ .75 & 0 & 0 \\ 0 & .5 & 0 \end{bmatrix}}$$

does the job.
(b) If there are q cars in the fleet, we start with

$$\vec{x}(1) = \begin{bmatrix} 0 \\ 0 \\ q \end{bmatrix}$$

so that

$$\vec{x}(2) = B\vec{x}(1) = \begin{bmatrix} .25 & .5 & 1 \\ .75 & 0 & 0 \\ 0 & .5 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ q \end{bmatrix} = \begin{bmatrix} q \\ 0 \\ 0 \end{bmatrix} \qquad \vec{x}(3) = B\vec{x}(2) = \begin{bmatrix} .25 & .5 & 1 \\ .75 & 0 & 0 \\ 0 & .5 & 0 \end{bmatrix}\begin{bmatrix} q \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}q \\ \frac{3}{4}q \\ 0 \end{bmatrix}$$

$$\vec{x}(4) = B\vec{x}(3) = \begin{bmatrix} .25 & .5 & 1 \\ .75 & 0 & 0 \\ 0 & .5 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{4}q \\ \frac{3}{4}q \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{7}{16}q \\ \frac{3}{16}q \\ \frac{3}{8}q \end{bmatrix} \quad \vec{x}(5) = B\vec{x}(4) = \begin{bmatrix} .25 & .5 & 1 \\ .75 & 0 & 0 \\ 0 & .5 & 0 \end{bmatrix}\begin{bmatrix} \frac{7}{16}q \\ \frac{3}{16}q \\ \frac{3}{8}q \end{bmatrix} = \begin{bmatrix} \frac{37}{64}q \\ \frac{21}{64}q \\ \frac{6}{64}q \end{bmatrix}$$

The proportions are $\boxed{0,\ 0,\ \frac{3}{8},\ \frac{3}{32}}$.
(c) In equilibrium (using I to denote the identity matrix)

$$\begin{aligned}
\vec{x}(n+1) &= \vec{x}(n)\\
B\vec{x}(n) &= I\vec{x}(n)\\
(B-I)\vec{x}(n) &= 0
\end{aligned}$$

If we call the components of $\vec{x}(n)$, $\alpha,\ \beta,\ \gamma$ respectively

$$(B-I)\vec{x}(n) = \begin{bmatrix} -.75 & .5 & 1 \\ .75 & -1 & 0 \\ 0 & .5 & -1 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The last equation forces $\beta = 2\gamma$ and the second equation forces $\alpha = \frac{4}{3}\beta = \frac{8}{3}\gamma$. In order to have an integer numbers of cars, γ has to be a positive multiple of 3. Let $\gamma = 3p$. The general solution is

$$\boxed{\vec{x}(n) = p\begin{bmatrix} 8 \\ 6 \\ 3 \end{bmatrix},\ p = 0,\ 1,\ 2,\ 3\ \cdots}$$

(d) We wish to find a matrix C obeying $\vec{x}(n+2) = C\vec{x}(n)$. As $\vec{x}(n+1) = B\vec{x}(n)$ and $\vec{x}(n+2) = B\vec{x}(n+1)$, we have $\vec{x}(n+2) = B\vec{x}(n+1) = B^2\vec{x}(n)$. So

$$C = B^2 = \begin{bmatrix} 1/4 & 2/4 & 1 \\ 3/4 & 0 & 0 \\ 0 & 2/4 & 0 \end{bmatrix} \begin{bmatrix} 1/4 & 2/4 & 1 \\ 3/4 & 0 & 0 \\ 0 & 2/4 & 0 \end{bmatrix} = \boxed{\frac{1}{16}\begin{bmatrix} 7 & 10 & 4 \\ 3 & 6 & 12 \\ 6 & 0 & 0 \end{bmatrix}}$$

(e) If $C\vec{x}(n) = \vec{x}(n)$, then

$$(C - I)\vec{x}(n) = \frac{1}{16}\begin{bmatrix} -9 & 10 & 4 \\ 3 & -10 & 12 \\ 6 & 0 & -16 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The last equation forces $\alpha = \frac{8}{3}\gamma$, just as in part c. The last equation minus twice the second equation forces $20\beta - 40\gamma = 0$ or $\beta = 2\gamma$ just as in part c. The general solution is once again

$$\vec{x}(n) = p\begin{bmatrix} 8 \\ 6 \\ 3 \end{bmatrix}, \ p = 0,\ 1,\ 2,\ 3\ \cdots$$

$\boxed{\text{No fleet vector repeats itself every two years but not every year.}}$

9) Determine whether or not each of the following functions is a linear transformation.
(a) $f(x, y, z) = 3x - 2y + 5z$ (b) $f(x, y, z) = -2y + 9z - 12$ (c) $f(x, y, z) = x^2 + y^2 + z^2$
(d) $f(x, y, z) = [x + y, x - y, 0]$ (e) $f(x, y, z) = [x + y, x - y, 1]$ (f) $f(x, y, z) = [x, y, z^2]$

Solution. We use the notations $\vec{x}$ to stand for $[x, y, z]$, $\vec{x}'$ to stand for $[x', y', z']$ and $f(\vec{x})$ to stand for $f(x, y, z)$. Observe that $s\vec{x} + t\vec{x}' = [sx + tx', sy + ty', sz + tz']$.
(a) $f(\vec{x}) = 3x - 2y + 5z$ is a function from $\mathbb{R}^3$ to $\mathbb{R}$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x} + t\vec{x}') &= 3(sx + tx') - 2(sy + ty') + 5(sz + tz') \\ sf(\vec{x}) + tf(\vec{x'}) &= s(3x - 2y + 5z) + t(3x' - 2y' + 5z') \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. They are, so this f $\boxed{\text{is linear}}$.
(b) $f(\vec{x}) = -2y + 9z - 12$ is a function from $\mathbb{R}^3$ to $\mathbb{R}$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x} + t\vec{x}') &= -2(sy + ty') + 9(sz + tz') - 12 \\ sf(\vec{x}) + tf(\vec{x'}) &= s(-2y + 9z - 12) + t(-2y' + 9z' - 12) \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. They aren't, because when $x = y = z = x' = y' = z' = 0$ the first expression reduces to -12 and the second reduces to $-12(s + t)$. These are equal only when $s + t = 1$. So this f $\boxed{\text{is not linear}}$.
(c) $f(\vec{x}) = x^2 + y^2 + z^2$ is a function from $\mathbb{R}^3$ to $\mathbb{R}$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x} + t\vec{x}') &= (sx + tx')^2 + (sy + ty')^2 + (sz + tz')^2 \\ sf(\vec{x}) + tf(\vec{x'}) &= s(x^2 + y^2 + z^2) + t(x'^2 + y'^2 + z'^2) \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. They aren't, because when $y = z = x' = y' = z' = 0$ the first expression reduces to s^2x^2 and the second reduces to sx^2. These are equal only when $s^2 = s$ or when $x = 0$. So this f $\boxed{\text{is not linear}}$.

(d) $f(\vec{x}) = [x+y, x-y, 0]$ is a function from $\mathbb{R}^3$ to $\mathbb{R}^3$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x}+t\vec{x}') &= [(sx+tx')+(sy+ty'), (sx+tx')-(sy+ty'), 0] \\ sf(\vec{x})+tf(\vec{x'}) &= s[x+y, x-y, 0]+t[x'+y', x'-y', 0] \\ &= [sx+sy+tx'+ty', sx-sy+tx'-ty', 0] \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. They are, so this f $\boxed{\text{is linear}}$.

(e) $f(\vec{x}) = [x+y, x-y, 1]$ is a function from $\mathbb{R}^3$ to $\mathbb{R}^3$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x}+t\vec{x}') &= [(sx+tx')+(sy+ty'), (sx+tx')-(sy+ty'), 1] \\ sf(\vec{x})+tf(\vec{x'}) &= s[x+y, x-y, 1]+t[x'+y', x'-y', 1] \\ &= [sx+sy+tx'+ty', sx-sy+tx'-ty', s+t] \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. The third components are equal only for $s+t=1$, so this f $\boxed{\text{is not linear}}$.

(f) $f(\vec{x}) = [x, y, z^2]$ is a function from $\mathbb{R}^3$ to $\mathbb{R}^3$. For it to be linear, the two expressions

$$\begin{aligned} f(s\vec{x}+t\vec{x}') &= [sx+tx', sy+ty', (sz+tz')^2] \\ sf(\vec{x})+tf(\vec{x'}) &= s[x, y, z^2]+t[x', y', z'^2] \\ &= [sx+tx', sy+ty', sz^2+tz'^2] \end{aligned}$$

have to be equal for all $\vec{x},\ \vec{x}'$ in $\mathbb{R}^3$ and all $s,\ t$ in $\mathbb{R}$. If $z'=0$ the third component of the first expression reduces to s^2z^2 while that of the second expression reduces to sz^2. These agree only for $z=0$ or $s=0,1$. So this f $\boxed{\text{is not linear}}$.

10) Suppose that a linear transformation maps $[1,1]$ to $[4,7]$ and $[1,-1]$ to $[8,3]$. What vector does it map $[5,14]$ to?

Solution 1. The vector

$$[5,14] = 5[1,0]+14[0,1] = \tfrac{5}{2}([1,1]+[1,-1]) + \tfrac{14}{2}([1,1]-[1,-1]) = \tfrac{19}{2}[1,1] - \tfrac{9}{2}[1,-1]$$

is mapped to

$$\tfrac{19}{2}[4,7] - \tfrac{9}{2}[8,3] = \left[\tfrac{4}{2}, \tfrac{106}{2}\right] = \boxed{[2,53]}$$

Solution 2. Write the vectors as column, rather than row vectors. Let

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be the matrix of the linear transformation. This matrix must obey

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ c+d \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} a-b \\ c-d \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

This consists of a system of two equations in the two unknowns a and b as well as a system of two equations in the two unknowns c and d. Both systems are easy to solve

$$\begin{aligned} a+b=4 \quad a-b=8 \quad &\implies \quad a=6 \quad b=-2 \\ c+d=7 \quad c-d=3 \quad &\implies \quad c=5 \quad d=2 \end{aligned}$$

Now that we know the matrix of the linear transformation, we just have to apply it to the specified input vector.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 5 \\ 14 \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 14 \end{bmatrix} = \begin{bmatrix} 30-28 \\ 25+28 \end{bmatrix} = \boxed{\begin{bmatrix} 2 \\ 53 \end{bmatrix}}$$

11) Is it possible for a linear transformation to map $[1,2]$ to $[1,0,-1]$, $[3,4]$ to $[1,2,3]$ and $[5,8]$ to $[3,1,6]$?

Solution. $[5,8] = 2[1,2] + [3,4]$. So for the map to be linear, it is necessary that $[3,1,6] = 2[1,0,-1] + [1,2,3] = [3,2,1]$, which is false. So, $\boxed{\text{it is not possible}}$.

12) Find the matrices representing the linear transformations

(a) $f(x,y) = [x+2y, y-x, x+y]$ (b) $g(x,y,z) = [x-y-z, 2y+5z-3x]$

(c) $f\big(g(x,y,z)\big)$ (d) $g\big(f(x,y)\big)$

Solution. Writing column, rather than row vectors.

(a)

$$\begin{bmatrix} x+2y \\ y-x \\ x+y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \\ ex+fy \end{bmatrix}$$

provided we choose $a=1,\ b=2,\ c=-1,\ d=1,\ e=1,\ f=1$. So the matrix is

$$\boxed{\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}}$$

(b)

$$\begin{bmatrix} x-y-z \\ 2y+5z-3x \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax+by+cz \\ dx+ey+fz \end{bmatrix}$$

provided we choose $a=1,\ b=-1,\ c=-1,\ d=-3,\ e=2,\ f=5$. So the matrix is

$$\boxed{\begin{bmatrix} 1 & -1 & -1 \\ -3 & 2 & 5 \end{bmatrix}}$$

(c) Substituting

$$\begin{bmatrix} u \\ v \end{bmatrix} = g\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & -1 \\ -3 & 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

into

$$f\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

gives

$$f\left(g\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)\right) = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} g\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -3 & 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \boxed{\begin{bmatrix} -5 & 3 & 9 \\ -4 & 3 & 6 \\ -2 & 1 & 4 \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(d) Substituting

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

into

$$g\left(\begin{bmatrix} u\\ v\\ w\end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & -1\\ -3 & 2 & 5\end{bmatrix}\begin{bmatrix} u\\ v\\ w\end{bmatrix}$$

gives

$$g\left(f\left(\begin{bmatrix} x\\ y\end{bmatrix}\right)\right) = \begin{bmatrix} 1 & -1 & -1\\ -3 & 2 & 5\end{bmatrix} f\left(\begin{bmatrix} x\\ y\end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & -1\\ -3 & 2 & 5\end{bmatrix}\begin{bmatrix} 1 & 2\\ -1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix} x\\ y\end{bmatrix} = \boxed{\begin{bmatrix} 1 & 0\\ 0 & 1\end{bmatrix}}\begin{bmatrix} x\\ y\end{bmatrix}$$

13) A linear transformation maps $\hat{\imath}$ to $-\hat{\jmath}$, $\hat{\jmath}$ to $\hat{k}$ and $\hat{k}$ to $-\hat{\imath}$.
(a) Find the matrix of the linear transformation.
(b) The linear transformation is a rotation. Find the axis of rotation.
(c) Find the angle of rotation.
(d) Show that this linear transformation really is a rotation.
Solution. (a) The matrix must obey

$$\begin{bmatrix} a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}\begin{bmatrix} 1\\ 0\\ 0\end{bmatrix} = \begin{bmatrix} 0\\ -1\\ 0\end{bmatrix} \qquad \begin{bmatrix} a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}\begin{bmatrix} 0\\ 1\\ 0\end{bmatrix} = \begin{bmatrix} 0\\ 0\\ 1\end{bmatrix} \qquad \begin{bmatrix} a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}\begin{bmatrix} 0\\ 0\\ 1\end{bmatrix} = \begin{bmatrix} -1\\ 0\\ 0\end{bmatrix}$$

The desired matrix is

$$T = \boxed{\begin{bmatrix} 0 & 0 & -1\\ -1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}}$$

(b) Perform the following little experiment. Take a book. Rotate the book, about its spine, by 45°. Observe that the spine of the book does not move at all. Vectors lying on the axis of rotation, do not change when the rotation is executed. To find the axis of rotation of T, we just need to find a nonzero vector $\vec{n}$ obeying $T\vec{n} = \vec{n}$.

$$\begin{bmatrix} 0 & 0 & -1\\ -1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix} n_1\\ n_2\\ n_3\end{bmatrix} = \begin{bmatrix} n_1\\ n_2\\ n_3\end{bmatrix} \implies \begin{bmatrix} -n_3\\ -n_1\\ n_2\end{bmatrix} = \begin{bmatrix} n_1\\ n_2\\ n_3\end{bmatrix} \implies \begin{bmatrix} n_1\\ n_2\\ n_3\end{bmatrix} = \boxed{c\begin{bmatrix} 1\\ -1\\ -1\end{bmatrix}} \text{ for any } c$$

(c) Rotate the book, about its spine, by 45° again. Observe that the bottom and top edges of the book rotate by 45°. Under a rotation of $\theta°$, vectors perpendicular to the axis of rotation rotate by $\theta°$. Vectors that are neither perpendicular to nor parallel to the axis of rotation, rotate by angles that are strictly between 0° and $\theta°$. (Repeat the book experiment a few times, concentrating on vectors that are almost parallel to the spine and then on vectors that are almost perpendicular to the spine, to convince yourself that this is true.) So to determine the angle of rotation, we select a vector, $\vec{v}$, perpendicular to the axis of rotation, $[1, -1, -1]$, and compute the angle between $\vec{v}$ and $T\vec{v}$. The vector $\vec{v} = \hat{\imath} + \hat{\jmath}$ is perpendicular to the axis of rotation, because $[1, 1, 0] \cdot [1, -1, -1] = 0$. It gets mapped to

$$\begin{bmatrix} 0 & 0 & -1\\ -1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix} 1\\ 1\\ 0\end{bmatrix} = \begin{bmatrix} 0\\ -1\\ 1\end{bmatrix}$$

The angle, θ, between $\hat{\imath} + \hat{\jmath}$ and the vector $-\hat{\jmath} + \hat{k}$ that it is mapped to obeys

$$(\hat{\imath} + \hat{\jmath}) \cdot (-\hat{\jmath} + \hat{k}) = \|\hat{\imath} + \hat{\jmath}\|\,\| - \hat{\jmath} + \hat{k}\|\,\cos\theta \implies 2\cos\theta = -1 \implies \boxed{\theta = 120°}$$

Remark. We have already seen, in part b, that every vector parallel to $\hat\imath - \hat\jmath - \hat k$ gets mapped to itself. That is, the linear transformation does not move it at all. If the linear transformation really is a rotation then every vector perpendicular to $\hat\imath-\hat\jmath-\hat k$, i.e. every vector $c\hat\imath+a\hat\jmath+b\hat k$ obeying $(c\hat\imath+a\hat\jmath+b\hat k)\cdot(\hat\imath-\hat\jmath-\hat k) = c-a-b = 0$, i.e. every vector of the form $(a+b)\hat\imath + a\hat\jmath + b\hat k$, should get mapped to a vector which is perpendicular to $\hat\imath-\hat\jmath-\hat k$, has the same length as $(a+b)\hat\imath+a\hat\jmath+b\hat k$ and makes an angle 120° with respect to $(a+b)\hat\imath + a\hat\jmath + b\hat k$. The vector $(a+b)\hat\imath + a\hat\jmath + b\hat k$ gets mapped to

$$\begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a+b \\ a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ -a-b \\ a \end{bmatrix}$$

This vector is again perpendicular to $\hat\imath-\hat\jmath-\hat k$, by the dot product test. The lengths $\|(a+b)\hat\imath+a\hat\jmath+b\hat k\| = \sqrt{(a+b)^2+a^2+b^2}$ and $\|-b\hat\imath-(a+b)\hat\jmath+a\hat k\| = \sqrt{b^2+(a+b)^2+a^2}$ both equal $\sqrt{2(a^2+b^2+ab)}$. The dot product

$$\begin{aligned} [(a+b)\hat\imath + a\hat\jmath + b\hat k]\cdot[-b\hat\imath-(a+b)\hat\jmath+a\hat k] &= -b(a+b) - a(a+b) + ab \\ &= -a^2 - b^2 - ab \\ &= -\tfrac{1}{2}\|(a+b)\hat\imath + a\hat\jmath + b\hat k\|\|-b\hat\imath-(a+b)\hat\jmath+a\hat k\| \end{aligned}$$

shows that the angle between the two vectors is 120°.

14) Determine the matrices of the reflections in the following planes in $\mathbb{R}^3$

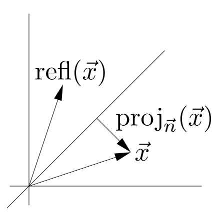
(a) $x+y+z=0$

(b) $2x-2y-z=0$

Solution. (a) The vector $\vec n = \hat\imath+\hat\jmath+\hat k$ is perpendicular to the given plane. The projection of any vector $\vec x$ on $\vec n$ is

$$\text{proj}_{\vec n}(\vec x) = \frac{\vec n\cdot\vec x}{\|\vec n\|^2}\vec n = \frac{x+y+z}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

The reflection is



$$\vec x - 2\text{proj}_{\vec n}(\vec x) = \begin{bmatrix}x\\y\\z\end{bmatrix} - \frac{x+y+z}{3}\begin{bmatrix}2\\2\\2\end{bmatrix} = \frac{1}{3}\begin{bmatrix} x-2y-2z \\ y-2x-2z \\ z-2x-2y\end{bmatrix}$$

$$= \boxed{\frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1\end{bmatrix}}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

Check: The vector $\vec n$, which is perpendicular to the plane, should be mapped to $-\vec n$. On the other hand, the vectors $\hat\imath-\hat\jmath$ and $\hat\jmath-\hat k$, both of which are parallel to the plane, should be mapped to themselves.

$$\frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = -\begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$\frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1\end{bmatrix}\begin{bmatrix}1\\-1\\0\end{bmatrix} = \begin{bmatrix}1\\-1\\0\end{bmatrix}$$

$$\frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1\end{bmatrix}\begin{bmatrix}0\\1\\-1\end{bmatrix} = \begin{bmatrix}0\\1\\-1\end{bmatrix}$$

(b) The vector $\vec{n} = 2\hat{\imath} - 2\hat{\jmath} - \hat{k}$ is perpendicular to the given plane. The projection of any vector $\vec{x}$ on $\vec{n}$ is

$$\text{proj}_{\vec{n}}(\vec{x}) = \frac{\vec{n}\cdot\vec{x}}{\|\vec{n}\|^2}\vec{n} = \frac{2x-2y-z}{9}\begin{bmatrix}2\\-2\\-1\end{bmatrix}$$

The reflection is

$$\vec{x} - 2\text{proj}_{\vec{n}}(\vec{x}) = \begin{bmatrix}x\\y\\z\end{bmatrix} - 2\frac{2x-2y-z}{9}\begin{bmatrix}2\\-2\\-1\end{bmatrix} = \frac{1}{9}\begin{bmatrix}x+8y+4z\\y+8x-4z\\7z+4x-4y\end{bmatrix} = \boxed{\frac{1}{9}\begin{bmatrix}1&8&4\\8&1&-4\\4&-4&7\end{bmatrix}}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

Check: The vector $\vec{n}$, which is perpendicular to the plane, should be mapped to $-\vec{n}$. On the other hand, the vectors $\hat{\imath} + \hat{\jmath}$ and $\hat{\imath} + 2\hat{k}$, both of which are parallel to the plane, should be mapped to themselves.

$$\frac{1}{9}\begin{bmatrix}1&8&4\\8&1&-4\\4&-4&7\end{bmatrix}\begin{bmatrix}2\\-2\\-1\end{bmatrix} = -\begin{bmatrix}2\\-2\\-1\end{bmatrix}$$

$$\frac{1}{9}\begin{bmatrix}1&8&4\\8&1&-4\\4&-4&7\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix}$$

$$\frac{1}{9}\begin{bmatrix}1&8&4\\8&1&-4\\4&-4&7\end{bmatrix}\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}1\\0\\2\end{bmatrix}$$

15) A solid body is rotating about an axis which passes through the origin and has direction $\vec{\Omega} = \Omega_1\hat{\imath} + \Omega_2\hat{\jmath} + \Omega_3\hat{k}$. The rate of rotation is $\|\vec{\Omega}\|$ radians per second. Denote by $\vec{x}$ the coordinates, at some fixed time, of a point fixed to the body and by $\vec{v}$ the velocity vector of the point at that time. Find a matrix A such that $\vec{v} = A\vec{x}$.

Solution. We saw in §I.7 that the velocity vector is

$$\vec{v} = \vec{\Omega}\times\vec{x} = \begin{bmatrix}\hat{\imath}&\hat{\jmath}&\hat{k}\\\Omega_1&\Omega_2&\Omega_3\\x&y&z\end{bmatrix} = \begin{bmatrix}\Omega_2 z - \Omega_3 y\\-\Omega_1 z + \Omega_3 x\\\Omega_1 y - \Omega_2 x\end{bmatrix} = \boxed{\begin{bmatrix}0&-\Omega_3&\Omega_2\\\Omega_3&0&-\Omega_1\\-\Omega_2&\Omega_1&0\end{bmatrix}}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

16) Find the matrix of the linear transformation on $\mathbb{R}^2$ which
(a) rotates each vector $\vec{x}$ clockwise by 45°.
(b) reflects each vector $\vec{x}$ in the x–axis and then rotates the result counterclockwise by 90°.
(c) reflects each vector $\vec{x}$ about the line $x = y$ and then projects the result onto the x–axis.

Solution.
(a) The vector $\hat{\imath}$ should be mapped to $\frac{1}{\sqrt{2}}(\hat{\imath} - \hat{\jmath})$ and the vector $\hat{\jmath}$ should be mapped to $\frac{1}{\sqrt{2}}(\hat{\imath} + \hat{\jmath})$. The matrix

$$\boxed{\tfrac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-1&1\end{bmatrix}}$$

does the job.
(b) The vector $\hat{\imath}$ is reflected to $\hat{\imath}$ and then rotated to $\hat{\jmath}$ and the vector $\hat{\jmath}$ is reflected to $-\hat{\jmath}$ and then rotated to $\hat{\imath}$. So, in the end, $\hat{\imath}$ is mapped to $\hat{\jmath}$ and $\hat{\jmath}$ is mapped to $\hat{\imath}$. The matrix

$$\boxed{\begin{bmatrix}0&1\\1&0\end{bmatrix}}$$

does the job.
(c) The vector $\hat{\imath}$ is reflected to $\hat{\jmath}$ and then projected to $\vec{0}$ and the vector $\hat{\jmath}$ is reflected to $\hat{\imath}$ and then projected to $\hat{\imath}$. So, in the end, $\hat{\imath}$ is mapped to $\vec{0}$ and $\hat{\jmath}$ is mapped to $\hat{\imath}$. The matrix

$$\boxed{\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}}$$

does the job.

17) The matrix on the left below is the inverse of the matrix on the right below. Fill in the missing entries.

$$\begin{bmatrix}4 & 3 & 2 & \bullet\\ 5 & 4 & 3 & 1\\ -2 & -2 & -1 & -1\\ 11 & 6 & 4 & 3\end{bmatrix} \qquad \begin{bmatrix}\bullet & -1 & 1 & 0\\ 7 & -3 & 1 & -1\\ -10 & 5 & \bullet & \bullet\\ -8 & 3 & \bullet & 1\end{bmatrix}$$

Solution. The matrices must obey

$$\begin{bmatrix}4 & 3 & 2 & a\\ 5 & 4 & 3 & 1\\ -2 & -2 & -1 & -1\\ 11 & 6 & 4 & 3\end{bmatrix}\begin{bmatrix}b & -1 & 1 & 0\\ 7 & -3 & 1 & -1\\ -10 & 5 & c & d\\ -8 & 3 & e & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

In particular, using (m, n) to denote the equation that matches the matrix element in row m and column n of the product of the two matrices on the left with the corresponding matrix element of the identity matrix on the right

$$\begin{array}{llcr}
(3,4): & (-2)(0) + (-2)(-1) + (-1)d + (-1)(1) = 0 & \Longrightarrow & d = 1\\
(1,4): & (4)(0) + (3)(-1) + (2)d + (a)(1) = 0 & \Longrightarrow & a = 1\\
(3,1): & (-2)(b) + (-2)(7) + (-1)(-10) + (-1)(-8) = 0 & \Longrightarrow & b = 2\\
(2,3): & (5)(1) + (4)(1) + (3)(c) + (1)(e) = 0 & \Longrightarrow & 3c + e = -9\\
(3,3): & (-2)(1) + (-2)(1) + (-1)(c) + (-1)(e) = 1 & \Longrightarrow & c + e = -5
\end{array}$$

The solution is $\boxed{a = 1,\ b = 2,\ c = -2,\ d = 1,\ e = -3}$.

18) Find, if possible, the inverses of each of the following matrices.

$$(a)\ \begin{bmatrix}1 & 4\\ 2 & 7\end{bmatrix} \qquad (b)\ \begin{bmatrix}2 & 5\\ 4 & 8\end{bmatrix} \qquad (c)\ \begin{bmatrix}3 & 4\\ 4 & -3\end{bmatrix} \qquad (d)\ \begin{bmatrix}1 & 2 & -1\\ 2 & 5 & 3\\ 1 & 3 & 9\end{bmatrix} \qquad (e)\ \begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 4\\ 3 & 4 & 6\end{bmatrix}$$

$$(f)\ \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix} \qquad (g)\ \begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 2 & 0 & 0\\ 3 & 0 & 0 & 0\end{bmatrix}$$

Solution. We have a canned formula for the inverse of 2×2 matrices. But I'll use row reduction, just for practice.
(a)

$$\left[\begin{array}{cc|cc}1 & 4 & 1 & 0\\ 2 & 7 & 0 & 1\end{array}\right] \quad \begin{array}{c}(1)\\ (2) - 2(1)\end{array}\left[\begin{array}{cc|cc}1 & 4 & 1 & 0\\ 0 & -1 & -2 & 1\end{array}\right] \quad \begin{array}{c}(1) + 4(2)\\ -(2)\end{array}\left[\begin{array}{cc|c}1 & 0 & \boxed{\begin{matrix}-7 & 4\\ 2 & -1\end{matrix}}\\ 0 & 1 & \end{array}\right]$$

Check:

$$\begin{bmatrix}1 & 4\\ 2 & 7\end{bmatrix}\begin{bmatrix}-7 & 4\\ 2 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

(b)

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 4 & 8 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2)-2(1) \end{array} \left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 0 & -2 & -2 & 1 \end{array}\right] \quad \begin{array}{c} (1)/2+5(2)/4 \\ -(2)/2 \end{array} \left[\begin{array}{cc|c} 1 & 0 & \boxed{\begin{matrix} -2 & 5/4 \\ 1 & -1/2 \end{matrix}} \\ 0 & 1 & \end{array}\right]$$

Check:

$$\begin{bmatrix} 2 & 5 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} -2 & 5/4 \\ 1 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(c)

$$\left[\begin{array}{cc|cc} 3 & 4 & 1 & 0 \\ 4 & -3 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1)/3 \\ (2)-4(1)/3 \end{array} \left[\begin{array}{cc|cc} 1 & 4/3 & 1/3 & 0 \\ 0 & -25/3 & -4/3 & 1 \end{array}\right] \quad \begin{array}{c} (1)+4(2)/25 \\ -3(2)/25 \end{array} \left[\begin{array}{cc|c} 1 & 0 & \boxed{\begin{matrix} 3/25 & 4/25 \\ 4/25 & -3/25 \end{matrix}} \\ 0 & 1 & \end{array}\right]$$

Check:

$$\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 3/25 & 4/25 \\ 4/25 & -3/25 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(d)

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 3 & 9 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2)-2(1) \\ (3)-(1) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 1 & 10 & -1 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2) \\ (3)-(2) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 0 & 5 & 1 & -1 & 1 \end{array}\right]$$

$$\begin{array}{c} (1)+(3)/5 \\ (2)-(3) \\ (3)/5 \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 6/5 & -1/5 & 1/5 \\ 0 & 1 & 0 & -3 & 2 & -1 \\ 0 & 0 & 1 & 1/5 & -1/5 & 1/5 \end{array}\right] \quad \begin{array}{c} (1)-2(2) \\ (2) \\ (3) \end{array} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & \boxed{\begin{matrix} 36/5 & -21/5 & 11/5 \\ -3 & 2 & -1 \\ 1/5 & -1/5 & 1/5 \end{matrix}} \\ 0 & 0 & 1 & \end{array}\right]$$

Check:

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} 36/5 & -21/5 & 11/5 \\ -15/5 & 10/5 & -5/5 \\ 1/5 & -1/5 & 1/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(e)

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2)-2(1) \\ (3)-3(1) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -3 & -3 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2) \\ (3)-2(2) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

$$\begin{array}{c} (1)-3(3) \\ -(2)-2(3) \\ (3) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right] \quad \begin{array}{c} (1)-2(2) \\ (2) \\ (3) \end{array} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & \boxed{\begin{matrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{matrix}} \\ 0 & 0 & 1 & \end{array}\right]$$

Check:

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

(f)

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2)-4(1) \\ (3)-7(1) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -12 & -7 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (1) \\ (2) \\ (3)-2(2) \end{array} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$$

The last equation has a zero left hand side and nonzero right hand side and so cannot be satisfied. There is $\boxed{\text{no inverse}}$. As a check, observe that

$$\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}1\\-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

No invertible matrix can map a nonzero vector to the zero vector.

(g)

$$\left[\begin{array}{cccc|cccc}0&0&1&0&1&0&0&0\\0&0&0&1&0&1&0&0\\0&2&0&0&0&0&1&0\\3&0&0&0&0&0&0&1\end{array}\right]\quad\begin{array}{c}(4)/3\\(3)/2\\(1)\\(2)\end{array}\left[\begin{array}{cccc|}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\boxed{\begin{array}{cccc}0&0&0&1/3\\0&0&1/2&0\\1&0&0&0\\0&1&0&0\end{array}}\right]$$

Check:

$$\begin{bmatrix}0&0&1&0\\0&0&0&1\\0&2&0&0\\3&0&0&0\end{bmatrix}\begin{bmatrix}0&0&0&1/3\\0&0&1/2&0\\1&0&0&0\\0&1&0&0\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

19) Let

$$B=\begin{bmatrix}-1&2&p\\0&-1&1\\2&1&0\end{bmatrix}$$

(a) For which values of p does B have an inverse?
(b) Find B^{-1}, for those values of p.

Solution.

$$\left[\begin{array}{ccc|ccc}-1&2&p&1&0&0\\0&-1&1&0&1&0\\2&1&0&0&0&1\end{array}\right]\qquad\begin{array}{c}(1)\\(2)\\(3)+2(1)\end{array}\left[\begin{array}{ccc|ccc}-1&2&p&1&0&0\\0&-1&1&0&1&0\\0&5&2p&2&0&1\end{array}\right]$$

$$\begin{array}{c}(1)\\(2)\\(3)+5(2)\end{array}\left[\begin{array}{ccc|ccc}-1&2&p&1&0&0\\0&-1&1&0&1&0\\0&0&2p+5&2&5&1\end{array}\right]$$

There is an inverse if and only if $\boxed{2p+5\neq 0}$. In this case, we can continue

$$\begin{array}{c}-(1)\\-(2)\\(3)/(2p+5)\end{array}\left[\begin{array}{ccc|ccc}1&-2&-p&-1&0&0\\0&1&-1&0&-1&0\\0&0&1&\frac{2}{2p+5}&\frac{5}{2p+5}&\frac{1}{2p+5}\end{array}\right]\qquad\begin{array}{c}(1)+p(3)\\(2)+(3)\\(3)\end{array}\left[\begin{array}{ccc|ccc}1&-2&0&\frac{-5}{2p+5}&\frac{5p}{2p+5}&\frac{p}{2p+5}\\0&1&0&\frac{2}{2p+5}&\frac{-2p}{2p+5}&\frac{1}{2p+5}\\0&0&1&\frac{2}{2p+5}&\frac{5}{2p+5}&\frac{1}{2p+5}\end{array}\right]$$

$$\begin{array}{c}(1)+2(2)\\(2)\\(3)\end{array}\left[\begin{array}{ccc|}1&0&0\\0&1&0\\0&0&1\end{array}\boxed{\begin{array}{ccc}\frac{-1}{2p+5}&\frac{p}{2p+5}&\frac{p+2}{2p+5}\\\frac{2}{2p+5}&\frac{-2p}{2p+5}&\frac{1}{2p+5}\\\frac{2}{2p+5}&\frac{5}{2p+5}&\frac{1}{2p+5}\end{array}}\right]$$

20) Suppose that, for some square matrix A, the series $\sum_{n=0}^{\infty}A^n=I+A+A^2+A^3+\cdots$ converges. (In series notation, A^0 is defined to be I.) Show that

$$(I-A)^{-1}=\sum_{n=0}^{\infty}A^n$$

Solution. We have to verify that

$$(I-A)\Big(\sum_{n=0}^{\infty} A^n\Big) = \Big(\sum_{n=0}^{\infty} A^n\Big)(I-A) = I$$

For the left hand side

$$(I-A)\Big(\sum_{n=0}^{\infty} A^n\Big) = \sum_{n=0}^{\infty} A^n - A\sum_{n=0}^{\infty} A^n = \sum_{n=0}^{\infty} A^n - \sum_{n=0}^{\infty} A^{n+1}$$

The second sum

$$\sum_{n=0}^{\infty} A^{n+1} = A + A^2 + A^3 + \cdots = \sum_{n=1}^{\infty} A^n$$

so

$$(I-A)\Big(\sum_{n=0}^{\infty} A^n\Big) = \sum_{n=0}^{\infty} A^n - \sum_{n=1}^{\infty} A^n = I$$

The argument for $\Big(\sum_{n=0}^{\infty} A^n\Big)(I-A)$ is similar.

21) Suppose that some square matrix obeys $A^n = 0$ for some positive integer.
(a) Find the inverse of A.
(b) Find the inverse of $I-A$.

Solution. (a) Trick question!! $\boxed{A \text{ has no inverse.}}$ If A had an inverse then multiplying both sides of $A^n = 0$ by $(A^{-1})^n$ would give $A^{-1}\cdots A^{-1}A\cdots A = 0$ (with n A^{-1}'s and n A's) and then $I = 0$.
(b) From problem 20, we would guess $\boxed{(I-A)^{-1} = I + A + A^2 + \cdots + A^{n-1}}$. All the remaining terms in the series $I + A + A^2 + \cdots$ vanish because of $A^n = 0$. To verify that the guess is correct we multiply out

$$\begin{aligned}(I-A)(I + A + A^2 + \cdots + A^{n-1}) &= (I + A + A^2 + \cdots + A^{n-1}) - A(I + A + A^2 + \cdots + A^{n-1})\\ &= (I + A + A^2 + \cdots + A^{n-1}) - (A + A^2 + A^3 + \cdots + A^n)\\ &= I - A^n = I\end{aligned}$$

The same argument also gives $(I + A + A^2 + \cdots + A^{n-1})(I-A) = I$

22) Suppose that L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Prove that its inverse, if it exists, is also linear.

Solution. We are to assume that there exists a map M such that $L(\vec{x}) = \vec{y}$ if and only if $M(\vec{y}) = \vec{x}$. Let $\vec{x}$ and $\vec{x}'$ be two arbitrary vectors in $\mathbb{R}^n$ and let s and t be two arbitrary numbers. Define $\vec{u} = M(\vec{x})$ and $\vec{u}' = M(\vec{x}')$. By hypothesis, $\vec{u} = M(\vec{x}) \Rightarrow \vec{x} = L(\vec{u})$ and $\vec{u}' = M(\vec{x}') \Rightarrow \vec{x}' = L(\vec{u}')$. As L is linear $L(s\vec{u} + t\vec{u}') = s\vec{x} + t\vec{x}'$, which in turn implies $M(s\vec{x} + t\vec{x}') = s\vec{u} + t\vec{u}' = sM(\vec{x}) + tM(\vec{x}')$. Thus, $\boxed{M \text{ is linear}}$.

23) Suppose that $\det\begin{bmatrix} 3 & 2 & p \\ 0 & p & 1 \\ 1 & 0 & 2 \end{bmatrix} = 10$. What are the possible values of p?

Solution. Expanding along the last row

$$\begin{aligned}\det\begin{bmatrix} 3 & 2 & p \\ 0 & p & 1 \\ 1 & 0 & 2 \end{bmatrix} &= \det\begin{bmatrix} 2 & p \\ p & 1 \end{bmatrix} + 2\det\begin{bmatrix} 3 & 2 \\ 0 & p \end{bmatrix}\\ &= (2-p^2) + 2(3p-0) = -p^2 + 6p + 2\end{aligned}$$

For this to be 10

$$-p^2 + 6p + 2 = 10 \iff p^2 - 6p + 8 = 0 \iff \boxed{p = 2, 4}$$

24) Let

$$A = \begin{bmatrix} 1 & 3 & 5 & * \\ 0 & 4 & 0 & 6 \\ 0 & 1 & 0 & 2 \\ 3 & * & 7 & 8 \end{bmatrix}$$

where the $*$'s denote unknown entries. Find all possible values of $\det A$.

Solution.

$$\det \begin{bmatrix} 1 & 3 & 5 & * \\ 0 & 4 & 0 & 6 \\ 0 & 1 & 0 & 2 \\ 3 & * & 7 & 8 \end{bmatrix} = \det \begin{bmatrix} 1 & 3 & 5 & * \\ 0 & 4 & 0 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & * & -8 & * \end{bmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4)-3(1) \end{matrix}$$

$$= \det \begin{bmatrix} 4 & 0 & 6 \\ 1 & 0 & 2 \\ * & -8 & * \end{bmatrix} = \det \begin{bmatrix} 0 & 0 & -2 \\ 1 & 0 & 2 \\ * & -8 & * \end{bmatrix} \begin{matrix} (1)-4(2) \\ (2) \\ (3) \end{matrix}$$

$$= -2 \det \begin{bmatrix} 1 & 0 \\ * & -8 \end{bmatrix} = \boxed{16}$$

for all values of the $*$'s.

25) Suppose that the 3×3 matrix A obeys $\det A = 5$. Compute (a) $\det(4A)$ (b) $\det(A^2)$ (c) $\det(4A^2)$

Solution.

$$\det(4A) = \det \left(\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} A \right) = \det \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \det A = 4^3 \times 5 = \boxed{320}$$

$$\det(A^2) = \det A \det A = 5 \times 5 = \boxed{25}$$

$$\det(4A^2) = \det(4A) \det A = 320 \times 5 = \boxed{1600}$$

26) Suppose that the 6×6 matrix A obeys $A^4 = 2A$. Find all possible values of $\det A$.

Solution.

$$A^4 = 2A \implies \det(A^4) = \det(2A) \implies (\det A)^4 = \det(2I) \det A$$

As in the last question $\det(2I) = 2^6 = 64$. So

$$0 = (\det A^4) - 64 \det A = \det A \big[(\det A)^3 - 64 \big]$$

Consequently $\det A = 0$ or $(\det A)^3 = 64$ or $\boxed{\det A = 0, 4}$, assuming that A has real matrix entries.

27) Evaluate

$$\det \begin{bmatrix} 1 & a & a^2 & a^3 \\ a & a^2 & a^3 & 1 \\ a^2 & a^3 & 1 & a \\ a^3 & 1 & a & a^2 \end{bmatrix}$$

Solution.

$$\det \begin{bmatrix} 1 & a & a^2 & a^3 \\ a & a^2 & a^3 & 1 \\ a^2 & a^3 & 1 & a \\ a^3 & 1 & a & a^2 \end{bmatrix} = \det \begin{bmatrix} 1 & a & a^2 & a^3 \\ 0 & 0 & 0 & 1-a^4 \\ 0 & 0 & 1-a^4 & a-a^5 \\ 0 & 1-a^4 & a-a^5 & a^2-a^6 \end{bmatrix} \begin{matrix} (1) \\ (2)-a(1) \\ (3)-a^2(1) \\ (4)-a^3(1) \end{matrix}$$

$$= \det \begin{bmatrix} 0 & 0 & 1-a^4 \\ 0 & 1-a^4 & a-a^5 \\ 1-a^4 & a-a^5 & a^2-a^6 \end{bmatrix} = (1-a^4) \det \begin{bmatrix} 0 & 1-a^4 \\ 1-a^4 & a-a^5 \end{bmatrix}$$

$$= (1-a^4)\big[-(1-a^4)(1-a^4) \big] \qquad = -(1-a^4)^3 = \boxed{(a^4-1)^3}$$

We expanded along the first row to achieve each of the the second, third and fourth equalities.