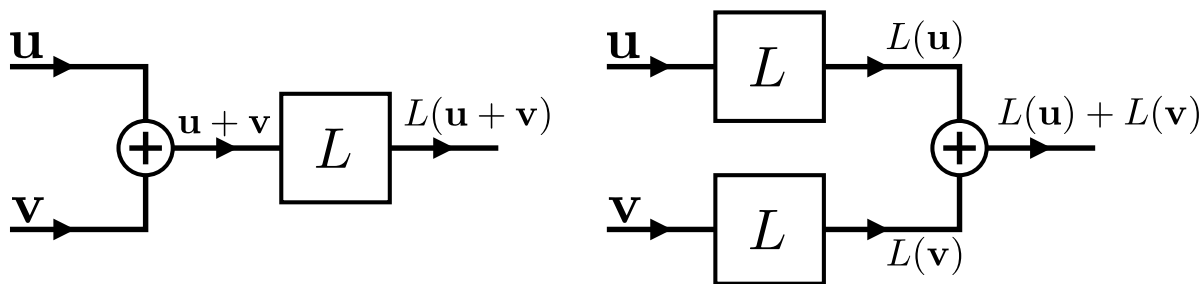


## Linear Mappings

Let  $L$  denote a function that accepts  $n$  component vectors ( $\mathbb{R}^n$ ) and returns  $m$  component vectors ( $\mathbb{R}^m$ ):  $L : \mathbb{R}^n \mapsto \mathbb{R}^m$ .

Function  $L$  will be assumed to be a **linear mapping**. A linear mapping has the property that applying the linear mapping to a sum is equal to applying the linear mapping to each summand and then summing the results after the mapping has been applied:

$$L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$$



Since adding  $\mathbf{0}$  to  $\mathbf{u}$  does not change the value of  $\mathbf{u}$ , adding  $L(\mathbf{0})$  to  $L(\mathbf{u})$  does not change the value of  $L(\mathbf{u})$ :

$$\mathbf{u} + \mathbf{0} = \mathbf{u} \quad \text{implies} \quad L(\mathbf{u}) + L(\mathbf{0}) = L(\mathbf{u})$$

Therefore  $L(\mathbf{0}) = \mathbf{0}$ .

Generalizing to adding the same vector to itself multiple times, if  $n$  is an arbitrary positive integer, then:

$$L(n \cdot \mathbf{u}) = L\left(\underbrace{\mathbf{u} + \mathbf{u} + \dots + \mathbf{u}}_n\right) = \underbrace{L(\mathbf{u}) + L(\mathbf{u}) + \dots + L(\mathbf{u})}_n = n \cdot L(\mathbf{u})$$

Generalizing  $n$  to an arbitrary real valued scalar  $c$ ,

$$L(c \cdot \mathbf{u}) = c \cdot L(\mathbf{u})$$

It is important to note that the common linear function  $f(x) = mx + c$  where  $m$  and  $c$  are constants is **not** a linear mapping if  $c \neq 0$ .

## Standard basis vectors

Given an arbitrary number of dimensions (entries)  $n$ , the  $i^{\text{th}}$  standard basis vector of  $\mathbb{R}^n$  is the vector where the  $i^{\text{th}}$  entry is 1, and all other entries are 0:

$$\mathbf{e}_1^n = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{e}_2^n = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{e}_3^n = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad \mathbf{e}_n^n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

In many cases the superscript  $n$ , which denotes the number of entries/dimensions that the basis vector is from is omitted.

Every possible  $n$  dimensional vector from  $\mathbb{R}^n$  can be expressed as a linear combination of the elementary basis vectors where the entries form the coefficients:

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + x_n \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \\
&= x_1 \mathbf{e}_1^n + x_2 \mathbf{e}_2^n + x_3 \mathbf{e}_3^n + \dots + x_n \mathbf{e}_n^n
\end{aligned}$$

If  $L$  is a linear mapping, then due to the linearity property,

$$\begin{aligned}
L \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) &= x_1 L \left( \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) + x_2 L \left( \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \right) + \dots + x_n L \left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right) \\
&= x_1 L(\mathbf{e}_1^n) + x_2 L(\mathbf{e}_2^n) + \dots + x_n L(\mathbf{e}_n^n)
\end{aligned}$$

If for each  $j = 1, 2, \dots, n$ ,  $L(\mathbf{e}_j^n) = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$ , then

$$L \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n \end{bmatrix}$$

## Coefficient matrices

From the above derivation, a linear mapping will always have the following form: given an arbitrary  $n$

component vector  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ ,

$$L(\mathbf{x}) = L \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n \end{bmatrix}$$

where the coefficients  $a_{1,1}, \dots, a_{m,n}$  define the mapping  $L$ . The coefficients that define  $L$  are quantified by a **matrix** with  $m$  rows and  $n$  columns.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

The  $j^{\text{th}}$  column is the vector that results from applying  $L$  to the  $j^{\text{th}}$  elementary basis vector.

The application of the linear mapping  $L$  to an arbitrary  $n$  component vector  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$  will be realized

by the **matrix vector product**  $A\mathbf{x}$ .  $A\mathbf{x} = L(\mathbf{x})$ :

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n \end{bmatrix}$$

Thanks to linearity, matrix-vector multiplication distributes over vector addition:

$$\begin{aligned} L(\mathbf{u} + \mathbf{v}) &= L(\mathbf{u}) + L(\mathbf{v}) \\ A(\mathbf{u} + \mathbf{v}) &= A\mathbf{u} + A\mathbf{v} \end{aligned}$$

- In general, when the dimensions of a matrix are described as being  $m \times n$ , the first quantity  $m$  is the number of rows, while the second quantity  $n$  is the number of columns.
- The “ $(i, j)$ ” entry is the entry at row  $i$  and column  $j$ .
- A matrix  $A$  with dimensions  $m \times 1$  has the same form as an  $m$  component vector. Let  $\mathbf{v}_A$  denote  $A$ , but interpreted as an  $m$  component vector rather than a matrix. The input of the linear mapping denoted by  $A$  is a 1 component vector, which is a scalar  $c$ , and whose output is the vector  $\mathbf{v}_A$  multiplied by the input scalar  $c$ :  $A[c] = \mathbf{v}_A \cdot c$ . In terms of matrix arithmetic, whether an  $m$  component vector is treated as a  $m \times 1$  matrix or vice versa, the math proceeds in exactly the same manner as will soon become apparent.
- An  $m \times n$  matrix where every entry is 0 is referred to as a zero matrix, and is denoted by  $0_{m \times n}$ . With a zero matrix, the corresponding linear mapping returns the zero vector no matter the input vector. If the size is unambiguous, the zero matrix is often denoted by simply 0.
- Any arbitrary linear system can be expressed as a matrix vector equation. The variables that are to be solved for are  $x_1, x_2, \dots, x_n$ .

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n = b_m \end{cases} \iff \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

**Examples:**

- Define the linear mapping  $L$  from  $\mathbb{R}^2$  (the set of 2 component vectors) to  $\mathbb{R}^3$  (the set of 3 component vectors) as

$$L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - 3x_2 \\ 4x_2 \\ -5x_1 + x_2 \end{bmatrix}$$

The coefficient matrix is the  $3 \times 2$  matrix:

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

Applying  $L$  to the vector  $\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$  gives:

$$L\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} (-1) - 3(2) \\ 4(2) \\ -5(-1) + 2 \end{bmatrix} = \begin{bmatrix} -7 \\ 8 \\ 7 \end{bmatrix}$$

The matrix vector product  $A\mathbf{x}$  is:

$$\begin{bmatrix} 1 & -3 \\ 0 & 4 \\ -5 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} (1)(-1) + (-3)(2) \\ (0)(-1) + (4)(2) \\ (-5)(-1) + (1)(2) \end{bmatrix} = \begin{bmatrix} -7 \\ 8 \\ 7 \end{bmatrix}$$

Applying the linear mapping gives the same result as the matrix vector product.

- Define the linear mapping  $L$  from  $\mathbb{R}^3$  (the set of 3 component vectors) to  $\mathbb{R}^3$  (the set of 3 component vectors) as

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_1 - x_3 \end{bmatrix}$$

The coefficient matrix is the  $3 \times 3$  matrix:

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

Applying  $L$  to the vector  $\mathbf{x} = \begin{bmatrix} 5 \\ 10 \\ -7 \end{bmatrix}$  gives:

$$L\left(\begin{bmatrix} 5 \\ 10 \\ -7 \end{bmatrix}\right) = \begin{bmatrix} 10 - 5 \\ (-7) - 10 \\ 5 - (-7) \end{bmatrix} = \begin{bmatrix} 5 \\ -17 \\ 12 \end{bmatrix}$$

The matrix vector product  $A\mathbf{x}$  is:

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 10 \\ -7 \end{bmatrix} = \begin{bmatrix} (-1)(5) + (1)(10) + (0)(-7) \\ (0)(5) + (-1)(10) + (1)(-7) \\ (1)(5) + (0)(10) + (-1)(-7) \end{bmatrix} = \begin{bmatrix} 5 \\ -17 \\ 12 \end{bmatrix}$$

Applying the linear mapping gives the same result as the matrix vector product.

## Adding linear mappings

Let  $L_A$  and  $L_B$  be linear mappings from  $\mathbb{R}^n$  (the set of  $n$  component vectors) to  $\mathbb{R}^m$  (the set of  $m$  component vectors). Let

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{bmatrix}$$

denote the respective coefficient matrices of  $L_A$  and  $L_B$ . Define  $L_C$  to be another linear mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^m$  whose return value is the sum of the return values of  $L_A$  and  $L_B$ :

$$\begin{aligned}
L_C(\mathbf{x}) &= L_A(\mathbf{x}) + L_B(\mathbf{x}) \\
&= \left( x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} \right) + \left( x_1 \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{m,n} \end{bmatrix} \right) \\
&= x_1 \left( \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{m,1} \end{bmatrix} \right) + x_2 \left( \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{m,2} \end{bmatrix} \right) + \dots + x_n \left( \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{m,n} \end{bmatrix} \right) \\
&= x_1 \begin{bmatrix} a_{1,1} + b_{1,1} \\ a_{2,1} + b_{2,1} \\ \vdots \\ a_{m,1} + b_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} + b_{1,2} \\ a_{2,2} + b_{2,2} \\ \vdots \\ a_{m,2} + b_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} + b_{1,n} \\ a_{2,n} + b_{2,n} \\ \vdots \\ a_{m,n} + b_{m,n} \end{bmatrix}
\end{aligned}$$

The coefficient matrix of  $L_C$  is therefore the **matrix sum** of the matrices  $A$  and  $B$ :

$$A+B = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix}$$

## Multiplication by a constant

Let  $L_A$  be a linear mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ . Let

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

denote the coefficient matrix of  $L_A$ . Define  $L_B$  to be another linear mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^m$  whose return value is the return value of  $L_A$  multiplied by a constant  $c$ :

$$\begin{aligned}
L_B(\mathbf{x}) &= c \cdot L_A(\mathbf{x}) = c \cdot \left( x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} \right) \\
&= x_1 \begin{bmatrix} ca_{1,1} \\ ca_{2,1} \\ \vdots \\ ca_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} ca_{1,2} \\ ca_{2,2} \\ \vdots \\ ca_{m,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} ca_{1,n} \\ ca_{2,n} \\ \vdots \\ ca_{m,n} \end{bmatrix}
\end{aligned}$$

The coefficient matrix of  $L_B$  is therefore the matrix  $A$  multiplied by a scalar  $c$ :

$$c \cdot A = A \cdot c = c \cdot \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} c \cdot a_{1,1} & c \cdot a_{1,2} & \cdots & c \cdot a_{1,n} \\ c \cdot a_{2,1} & c \cdot a_{2,2} & \cdots & c \cdot a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c \cdot a_{m,1} & c \cdot a_{m,2} & \cdots & c \cdot a_{m,n} \end{bmatrix}$$

The scalar can be multiplied on the left or right:  $c \cdot A = A \cdot c$ .

## Composition

Let  $L_A$  be a linear mapping from  $\mathbb{R}^r$  to  $\mathbb{R}^m$ , and let  $L_B$  be a linear mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^r$ . The coefficient matrix of  $L_A$  is the  $m \times r$  matrix  $A$ , and the coefficient matrix of  $L_B$  is the  $r \times n$  matrix  $B$ :

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,r} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r,1} & b_{r,2} & \cdots & b_{r,n} \end{bmatrix}$$

Define  $L_C$  to denote the composition of  $L_A$  after  $L_B$ :  $L_C = L_A \circ L_B$ .  $L_C$  takes an  $n$  component vector, first uses  $L_B$  to get an  $r$  component vector, and lastly uses  $L_A$  to get an  $m$  component vector.

$$\begin{aligned} L_A(L_B(\mathbf{x})) &= L_A \left( L_B \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \right) = L_A \left( x_1 \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{r,1} \end{bmatrix} + x_2 \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{r,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{r,n} \end{bmatrix} \right) \\ &= x_1 L_A \left( \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{r,1} \end{bmatrix} \right) + x_2 L_A \left( \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{r,2} \end{bmatrix} \right) + \dots + x_n L_A \left( \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{r,n} \end{bmatrix} \right) \\ &= x_1 \begin{bmatrix} \sum_{k=1}^r a_{1,k} b_{k,1} \\ \sum_{k=1}^r a_{2,k} b_{k,1} \\ \vdots \\ \sum_{k=1}^r a_{m,k} b_{k,1} \end{bmatrix} + x_2 \begin{bmatrix} \sum_{k=1}^r a_{1,k} b_{k,2} \\ \sum_{k=1}^r a_{2,k} b_{k,2} \\ \vdots \\ \sum_{k=1}^r a_{m,k} b_{k,2} \end{bmatrix} + \dots + x_n \begin{bmatrix} \sum_{k=1}^r a_{1,k} b_{k,n} \\ \sum_{k=1}^r a_{2,k} b_{k,n} \\ \vdots \\ \sum_{k=1}^r a_{m,k} b_{k,n} \end{bmatrix} \end{aligned}$$

The coefficient matrix of  $L_C$  is therefore the **matrix product** of the matrices  $A$  and  $B$

$$\begin{aligned} AB &= \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,r} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r,1} & b_{r,2} & \cdots & b_{r,n} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^r a_{1,k} b_{k,1} & \sum_{k=1}^r a_{1,k} b_{k,2} & \cdots & \sum_{k=1}^r a_{1,k} b_{k,n} \\ \sum_{k=1}^r a_{2,k} b_{k,1} & \sum_{k=1}^r a_{2,k} b_{k,2} & \cdots & \sum_{k=1}^r a_{2,k} b_{k,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^r a_{m,k} b_{k,1} & \sum_{k=1}^r a_{m,k} b_{k,2} & \cdots & \sum_{k=1}^r a_{m,k} b_{k,n} \end{bmatrix} \end{aligned}$$

- $AB$  will have  $m$  rows and  $n$  columns.
- Matrix multiplication is the composition of two linear mappings. Since the output of  $L_B$  must match the input of  $L_A$ , the number of rows of  $B$  must match the number of columns of  $A$ .

- In summary, the  $(i, j)$  entry of  $AB$  is computed as follows: Take the  $i^{\text{th}}$  row of  $A$  and the  $j^{\text{th}}$  column of  $B$ . Both are a list of  $r$  quantities. Multiply the corresponding quantities together, and lastly sum the products in a manner similar to computing the dot product of two vectors.
- The process of matrix multiplication, which is computing the composition  $L_A \circ L_B$ , effectively applies the linear mapping  $L_A$  to each column of matrix  $B$  independent of the other columns:

$$A \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r,1} & b_{r,2} & \cdots & b_{r,n} \end{bmatrix} = \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{r,1} \end{bmatrix} A \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ \vdots \\ b_{r,2} \end{bmatrix} \cdots A \begin{bmatrix} b_{1,n} \\ b_{2,n} \\ \vdots \\ b_{r,n} \end{bmatrix}$$

This interpretation of matrix multiplication will allow the linear mapping  $L_A$  to be applied to multiple vectors simultaneously. This interpretation will also allow elementary row operations to be quantified as linear mappings, since the elementary row operations perform the same action on each column.

- With the matrix product, the matrices on the right correspond to mappings that are applied first, with matrices on the left being applied last. This notation is consistent with the notation  $L_A(L_B(\mathbf{x}))$ , but can be counterintuitive since instinct is to read the matrices from left to right as opposed to right to left.
- It is important to note that matrix multiplication is **not** always commutative:  $AB \neq BA$ . When two functions are composed, the order in which the functions are applied matters, and the same is true with matrix multiplication.
- If the matrix  $B$  has a single column and hence the dimensions  $r \times 1$ , then  $B$  is essentially an  $r$  component vector. In this case, let  $\mathbf{v}_B$  denote the vector interpretation of  $B$ . The matrix product  $AB$  is equal to the matrix vector product  $A\mathbf{v}_B$ , so whether  $B$  is interpreted as a matrix or a vector is irrelevant.
- If any matrix in a product is multiplied by a scalar, then the multiplication by a scalar can occur last: Let  $L_{cA}$  denote the linear mapping quantified by the matrix  $cA$ .

$$L_{cA}(L_B(\mathbf{x})) = cL_A(L_B(\mathbf{x})) \quad \text{and} \quad L_A(cL_B(\mathbf{x})) = cL_A(L_B(\mathbf{x}))$$

so

$$L_{cA}(L_B(\mathbf{x})) = L_A(cL_B(\mathbf{x})) = cL_A(L_B(\mathbf{x}))$$

so

$$(cA)B = A(cB) = c(AB)$$

- The distributive laws hold for matrix multiplication. If  $A$  is an  $m \times r$  matrix, and  $B$  and  $C$  are  $r \times n$  matrices, let  $L_A$ ,  $L_B$ , and  $L_C$  denote the corresponding linear mappings. Thanks to linearity,

$$L_A(L_B(\mathbf{x}) + L_C(\mathbf{x})) = L_A(L_B(\mathbf{x})) + L_A(L_C(\mathbf{x}))$$

so

$$A(B + C) = AB + AC$$

Now let  $A$  and  $B$  be  $m \times r$  matrices, and  $C$  be a  $r \times n$  matrix. Let  $L_A$ ,  $L_B$ , and  $L_C$  denote the corresponding linear mappings with coefficient matrices  $A$ ,  $B$ , and  $C$ . Let  $L_{A+B}$  denote the linear mapping quantified by the matrix  $A + B$ .

$$L_{A+B}(L_C(\mathbf{x})) = L_A(L_C(\mathbf{x})) + L_B(L_C(\mathbf{x}))$$

so

$$(A + B)C = AC + BC$$

- If  $k$  is an arbitrary positive integer, then  $A^k$  will denote the matrix  $A$  multiplied by itself  $k$  times:

$$A^k = \underbrace{A \cdot A \cdot \dots \cdot A}_k$$

For  $A^k$  to exist for any integer  $k$  that is  $\geq 2$ ,  $A$  must have an equal number of rows and columns, i.e.  $A$  must be square.

#### Examples:

- Let  $A = \begin{bmatrix} 3 & 7 & -1 & 0 & 5 \\ -8 & -1 & 2 & 2 & -6 \\ 0 & 1 & -1 & 7 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} -3 & 0 & 4 & 2 & -2 \\ 6 & 5 & -5 & 0 & 10 \\ -4 & -5 & -1 & 2 & -4 \end{bmatrix}$ . Both matrices  $A$  and  $B$  have the dimensions  $3 \times 5$ , so the matrix sum  $A + B$  can be computed:

$$A + B = \begin{bmatrix} 3 + (-3) & 7 + 0 & (-1) + 4 & 0 + 2 & 5 + (-2) \\ (-8) + 6 & (-1) + 5 & 2 + (-5) & 2 + 0 & (-6) + 10 \\ 0 + (-4) & 1 + (-5) & (-1) + (-1) & 7 + 2 & 4 + (-4) \end{bmatrix} = \begin{bmatrix} 0 & 7 & 3 & 2 & 3 \\ -2 & 4 & -3 & 2 & 4 \\ -4 & -4 & -2 & 9 & 0 \end{bmatrix}$$

- Let  $A = \begin{bmatrix} 4 & 5 & -1 \\ 2 & 3 & 0 \end{bmatrix}$  and  $B = \begin{bmatrix} 7 \\ -8 \\ 2 \end{bmatrix}$ . Matrix  $A$  has dimensions  $2 \times 3$ , while matrix  $B$  has dimensions  $3 \times 1$ . The dimensions do not match so  $A + B$  is not defined.

- Let  $A = \begin{bmatrix} 3 & -1 & 0 \\ 4 & 5 & 1 \\ 2 & 3 & 5 \\ -8 & 2 & 1 \end{bmatrix}$  and  $c = 5$ .

$$cA = \begin{bmatrix} (5)(3) & (5)(-1) & (5)(0) \\ (5)(4) & (5)(5) & (5)(1) \\ (5)(2) & (5)(3) & (5)(5) \\ (5)(-8) & (5)(2) & (5)(1) \end{bmatrix} = \begin{bmatrix} 15 & -5 & 0 \\ 20 & 25 & 5 \\ 10 & 15 & 25 \\ -40 & 10 & 5 \end{bmatrix}$$

- Given the two linear mappings

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -x_1 + 5x_2 \\ 3x_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_1 - 2y_2 \\ 7y_1 \end{bmatrix}$$

The composition of these mappings that expresses  $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$  as a function of  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  is:

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_1 - 2y_2 \\ 7y_1 \end{bmatrix} = \begin{bmatrix} (-x_1 + 5x_2) - 2(3x_2) \\ 7(-x_1 + 5x_2) \end{bmatrix} = \begin{bmatrix} (-x_1 + 5x_2) + (-6x_2) \\ (-7x_1 + 35x_2) \end{bmatrix} = \begin{bmatrix} -x_1 - x_2 \\ -7x_1 + 35x_2 \end{bmatrix}$$

Alternately, the linear mappings can be expressed using matrix vector products:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 5 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

The composition of these mappings that expresses  $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$  as a function of  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  is:

$$\begin{aligned} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= \begin{bmatrix} 1 & -2 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} -1 & 5 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (1)(-1) + (-2)(0) & (1)(5) + (-2)(3) \\ (7)(-1) + (0)(0) & (7)(5) + (0)(3) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} -1 + 0 & 5 + (-6) \\ (-7) + 0 & 35 + 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -7 & 35 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

As can be seen, the resultant mapping computed using matrices is identical to the mapping computed via direct substitution.



- Let  $A = \begin{bmatrix} 6 & 7 \\ -8 & 2 \\ 3 & 1 \end{bmatrix}$  and  $B = \begin{bmatrix} 4 & -1 & 0 & 3 \\ 2 & 0 & -6 & 1 \end{bmatrix}$ . Matrix  $A$  has the dimensions  $3 \times 2$ , while matrix  $B$  has the dimensions  $2 \times 4$ . Since the number of columns of  $A$  is equal to the number of rows of  $B$ , the matrix product  $AB$  can be computed:

$$\begin{aligned} AB &= \begin{bmatrix} (6)(4) + (7)(2) & (6)(-1) + (7)(0) & (6)(0) + (7)(-6) & (6)(3) + (7)(1) \\ (-8)(4) + (2)(2) & (-8)(-1) + (2)(0) & (-8)(0) + (2)(-6) & (-8)(3) + (2)(1) \\ (3)(4) + (1)(2) & (3)(-1) + (1)(0) & (3)(0) + (1)(-6) & (3)(3) + (1)(1) \end{bmatrix} \\ &= \begin{bmatrix} 24 + 14 & (-6) + 0 & 0 + (-42) & 18 + 7 \\ -32 + 4 & 8 + 0 & 0 + (-12) & (-24) + 2 \\ 12 + 2 & -3 + 0 & 0 + (-6) & 9 + 1 \end{bmatrix} = \begin{bmatrix} 38 & -6 & -42 & 25 \\ -28 & 8 & -12 & -22 \\ 14 & -3 & -6 & 10 \end{bmatrix} \end{aligned}$$

The matrix product  $BA$  cannot be computed since the number of columns of  $B$  does not equal the number of rows of  $A$ .

- Let  $A = \begin{bmatrix} 6 & 7 \\ -8 & 2 \end{bmatrix}$  and  $B = \begin{bmatrix} 4 & -1 \\ 2 & 0 \end{bmatrix}$ . Matrices  $A$  and  $B$  both have the dimensions  $2 \times 2$ . Since the number of columns of  $A$  is equal to the number of rows of  $B$ , the matrix product  $AB$  can be computed:

$$AB = \begin{bmatrix} (6)(4) + (7)(2) & (6)(-1) + (7)(0) \\ (-8)(4) + (2)(2) & (-8)(-1) + (2)(0) \end{bmatrix} = \begin{bmatrix} 24 + 14 & (-6) + 0 \\ (-32) + 4 & 8 + 0 \end{bmatrix} = \begin{bmatrix} 38 & -6 \\ -28 & 8 \end{bmatrix}$$

Since the number of columns of  $B$  is equal to the number of rows of  $A$ , the matrix product  $BA$  can be computed:

$$BA = \begin{bmatrix} (4)(6) + (-1)(-8) & (4)(7) + (-1)(2) \\ (2)(6) + (0)(-8) & (2)(7) + (0)(2) \end{bmatrix} = \begin{bmatrix} 24 + 8 & 28 + (-2) \\ 12 + 0 & 14 + 0 \end{bmatrix} = \begin{bmatrix} 32 & 26 \\ 12 & 14 \end{bmatrix}$$

Note that  $AB \neq BA$ .

## Square matrices and Inverse matrices

A **square matrix** is a matrix that has an equal number of rows and columns. The dimensions are  $n \times n$  for some positive integer  $n$ . With  $n \times n$  square matrices, the linear mapping takes as input and returns as output vectors from the same set  $\mathbb{R}^n$ . If  $A$  and  $B$  are both  $n \times n$  square matrices, then  $A + B$  and  $AB$  can always be computed.

Given an  $n \times n$  matrix  $A$ , the entries of  $A$  where the row and column indices are equal are referred to as the **diagonal entries**:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}$$

A square matrix  $A$  is **diagonal** if and only if all entries aside from the diagonal entries are 0:

$$A = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix}$$

The linear mapping encoded by a diagonal matrix multiplies each entry independently by a constant, where the entries have no influence on each other:

$$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \\ \vdots \\ d_n x_n \end{bmatrix}$$

A square matrix is **symmetric** if and only if it is symmetric about its diagonal. In other words, for all row column index pairs  $(i, j)$ , the entry at  $(j, i)$  is equal to the entry at  $(i, j)$ :  $a_{j,i} = a_{i,j}$ . An example symmetric matrix is:

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

A square matrix is **skew symmetric** if and only if for all row column index pairs  $(i, j)$ , the entry at  $(j, i)$  is equal to the negative of the entry at  $(i, j)$ :  $a_{j,i} = -a_{i,j}$ . An example skew symmetric matrix is:

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

Note that on the diagonal, since the condition  $a_{i,i} = -a_{i,i}$  is equivalent to  $a_{i,i} = 0$ , all entries along the diagonal are 0.

A square matrix is **upper triangular** if and only if the only nonzero entries are above or on the diagonal: for all row column indices  $(i, j)$ , if  $j < i$  then  $a_{i,j} = 0$ :

$$\begin{bmatrix} * & * & * & * & \cdots & * \\ 0 & * & * & * & \cdots & * \\ 0 & 0 & * & * & \cdots & * \\ 0 & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & * \end{bmatrix}$$

A square matrix is **lower triangular** if and only if the only nonzero entries are below or on the diagonal: for all row column indices  $(i, j)$ , if  $j > i$  then  $a_{i,j} = 0$ :

$$\begin{bmatrix} * & 0 & 0 & 0 & \cdots & 0 \\ * & * & 0 & 0 & \cdots & 0 \\ * & * & * & 0 & \cdots & 0 \\ * & * & * & * & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & * & \cdots & * \end{bmatrix}$$

The **trace** of a square matrix is the sum of the diagonal entries:

$$\text{tr} \left( \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \right) = a_{1,1} + a_{2,2} + \cdots + a_{n,n}$$

## The identity matrix

Of special interest is the “identity matrix”  $I_n$ . Consider the linear mapping  $L_I$  from  $\mathbb{R}^n$  to itself that has no impact on the value of its input vector:

$$L_I \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

The coefficient matrix of  $L_I$  is the  $n \times n$  identity matrix  $I_n$ . The identity matrix has 1’s along the diagonal and 0’s everywhere else:

$$I_n = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

- The subscript of  $n$  is often excluded when the size is unambiguous.
- The identity matrix quantifies a linear mapping that does nothing to its input vector.
- Introducing the identity matrix (i.e. the do-nothing function) into a matrix product (i.e. a composition of functions) has no impact of the final matrix/linear mapping. Given an arbitrary  $m \times n$  matrix  $A$ ,

$$I_m A = A I_n = A$$

- Given any  $n \times n$  square matrix  $A$ , then  $A^0 = I_n$ . In other words, the empty matrix product is the do-nothing matrix, the identity matrix.

## Inverse matrices

In a similar manner that there are inverse functions, if a linear mapping  $L$  from  $\mathbb{R}^n$  to  $\mathbb{R}^n$  is one-to-one, which means that no two different input vectors generate the same output vector, then there exists an inverse linear mapping  $L^{-1}$  that reverses  $L$ :

$$L^{-1}(L(\mathbf{x})) = \mathbf{x}$$

If  $A$  is the  $n \times n$  coefficient matrix of linear mapping  $L$ , then the  $n \times n$  coefficient matrix of the inverse linear mapping  $L^{-1}$  is the “inverse matrix” of  $A$ , denoted by  $A^{-1}$ . An inverse matrix only exists for **square** matrices where the linear mapping is **one-to-one**. If an inverse matrix exists for matrix  $A$ , then  $A$  is referred to as being **non-singular**. If  $A$  does not have an inverse, then  $A$  is **singular**.

- Given a linear mapping  $L$  that has an inverse  $L^{-1}$ , the function  $L^{-1}(L(\mathbf{x}))$  has no effect on  $\mathbf{x}$ :  $L^{-1}(L(\mathbf{x})) = \mathbf{x}$ . If  $A$  is the coefficient matrix of  $L$ , then the fact that the function  $L^{-1}(L(\mathbf{x}))$  does nothing means that  $A^{-1}A = I_n$ . Conversely, the function  $L(L^{-1}(\mathbf{x})) = \mathbf{x}$  also does nothing, so  $AA^{-1} = I_n$ . Therefore:

$$L^{-1}(L(\mathbf{x})) = \mathbf{x} \quad \text{and} \quad L(L^{-1}(\mathbf{x})) = \mathbf{x}$$

so

$$A^{-1}A = I_n \quad \text{and} \quad AA^{-1} = I_n$$

- For any positive integer  $k$ ,  $A^{-k}$  is the result of multiplying  $k$  copies of  $A^{-1}$ . In other words, reversing  $A$   $k$  times:

$$A^{-k} = \underbrace{A^{-1} \cdot A^{-1} \cdot \dots \cdot A^{-1}}_k$$

In summary, if  $k$  is any integer, then

$$A^k = \begin{cases} \underbrace{A \cdot A \cdot \dots \cdot A}_k & (k > 0) \\ I_n & (k = 0) \\ \underbrace{A^{-1} \cdot A^{-1} \cdot \dots \cdot A^{-1}}_{|k|} & (k < 0) \end{cases}$$

It is then easy to establish that for any integers  $k_1$  and  $k_2$  that:

$$A^{k_1+k_2} = A^{k_1} A^{k_2} \quad \text{and} \quad A^{k_1 k_2} = (A^{k_1})^{k_2}$$

- Let  $A$  and  $B$  be both nonsingular  $n \times n$  matrices. The matrix  $AB$  denotes the mapping that consists of applying  $B$  followed by  $A$ . To reverse this mapping, the more recent mapping  $A$  must be reversed first, followed by reversing the earlier mapping  $B$ :

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

Now will be described an algorithm for computing the inverse of a square matrix. Inverting a function typically involves reversing the steps that comprise the function. For example, consider the function  $f(X) = (5X + 2)^3$ . Given an arbitrary value of  $Y$ , to find the value of  $X$  for which  $f(X) = Y$ , we reverse the steps that were applied to  $X$ , and apply the reversed steps to  $Y$ :

$$f(X) = Y \iff (5X + 2)^3 = Y \iff 5X + 2 = \sqrt[3]{Y} \iff 5X = \sqrt[3]{Y} - 2 \iff X = \frac{\sqrt[3]{Y} - 2}{5}$$

so the inverse function is  $f^{-1}(Y) = \frac{\sqrt[3]{Y} - 2}{5}$ .

A linear mapping has the form  $L(\mathbf{x}) = A\mathbf{x}$ . Given an arbitrary output vector of  $\mathbf{y}$ , to find the value of  $\mathbf{x}$  for which  $L(\mathbf{x}) = \mathbf{y}$ , the equation  $L(\mathbf{x}) = \mathbf{y}$  must be solved for  $\mathbf{x}$  with  $\mathbf{y}$  held as an arbitrary constant, just like how the function  $f(X) = (5X + 2)^3$  was inverted.

With  $\mathbf{y}$  as an arbitrary vector, the equation  $A\mathbf{x} = \mathbf{y}$  must be solved for  $\mathbf{x}$ , and the expression for  $\mathbf{x}$  will be  $\mathbf{x} = A^{-1}\mathbf{y}$ .

When solving an equation, the same operations must be applied to both sides of the equation in order to preserve the truth of the equation. Moreover, the steps must be reversible. Multiplying both sides by 0 is not allowed, as multiplication by 0 cannot be reversed and information is hence destroyed by multiplication by 0. The equation  $A\mathbf{x} = \mathbf{y}$  will be solved for  $\mathbf{x}$  by applying equivalent linear mappings to both sides. These linear mappings will be easily invertible.

What linear mappings are easily invertible? The linear mappings that correspond to the elementary row operations are easy to invert.

First is row replacement. Choose  $i$  and  $j$  to be arbitrary row indices, and let  $k$  be a nonzero constant.

Consider an arbitrary vector  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ . The mapping  $L_{R_j \rightarrow R_j + kR_i}$  will be the linear mapping that adds  $k$  times row  $i$  to row  $j$ :

$$L_{R_j \rightarrow R_j + kR_i} \left( \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \\ x_j \\ \vdots \end{bmatrix} \right) = \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \\ x_j + kx_i \\ \vdots \end{bmatrix}$$

The coefficient matrix  $E_{R_j \rightarrow R_j + kR_i}$  for this simple linear mapping is referred to as an elementary matrix, and this matrix is simply the identity matrix with entry  $(j, i)$  set to  $k$ . For example, given the elementary operation  $L_{R_4 \rightarrow R_4 - 7R_2}$  where  $n = 5$ .

$$E_{R_4 \rightarrow R_4 - 7R_2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -7 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

When a matrix is multiplied **on the left** by an elementary matrix, the rows are modified according to the corresponding elementary row operation. For example, given an arbitrary  $5 \times 5$  matrix  $A$ ,

$$\begin{aligned} E_{R_4 \rightarrow R_4 - 7R_2} A &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -7 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ a_{4,1} - 7a_{2,1} & a_{4,2} - 7a_{2,2} & a_{4,3} - 7a_{2,3} & a_{4,4} - 7a_{2,4} & a_{4,5} - 7a_{2,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \end{aligned}$$

The linear mapping  $L_{R_j \rightarrow R_j + kR_i}$  is very easy to invert. Just subtract  $k$  times row  $i$  from row  $j$ :

$$L_{R_j \rightarrow R_j + kR_i}^{-1} \left( \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \\ x_j \\ \vdots \end{bmatrix} \right) = \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \\ x_j - kx_i \\ \vdots \end{bmatrix}$$

The inverse of the elementary matrix  $E_{R_j \rightarrow R_j + kR_i}$  is simply the identity matrix with entry  $(j, i)$  set to  $-k$ .

The next type of elementary row operation is to multiply a row by a fixed constant. Choose  $i$  to be an arbitrary row index, and let  $k$  be a nonzero constant. Consider an arbitrary vector  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ . The mapping

$L_{R_i \rightarrow kR_i}$  will be the linear mapping that multiplies row  $i$  by  $k$ :

$$L_{R_i \rightarrow kR_i} \left( \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix} \right) = \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \end{array} \begin{bmatrix} \vdots \\ kx_i \\ \vdots \end{bmatrix}$$

The coefficient matrix  $E_{R_i \rightarrow kR_i}$  for this simple linear mapping is simply the identity matrix with entry  $(i, i)$  set to  $k$ . For example, given the elementary operation  $L_{R_4 \rightarrow -2R_4}$  where  $n = 5$ .

$$E_{R_4 \rightarrow -2R_4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

When a matrix is multiplied **on the left** by an elementary matrix, the rows are modified according to the corresponding elementary row operation. For example, given an arbitrary  $5 \times 5$  matrix  $A$ ,

$$\begin{aligned} E_{R_4 \rightarrow -2R_4} A &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ -2a_{4,1} & -2a_{4,2} & -2a_{4,3} & -2a_{4,4} & -2a_{4,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \end{aligned}$$

The linear mapping  $L_{R_i \rightarrow kR_i}$  is very easy to invert. Just divide row  $i$  by  $k$ :

$$L_{R_i \rightarrow kR_i}^{-1} \left( \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix} \right) = \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i/k \\ \vdots \end{bmatrix}$$

The inverse of the elementary matrix  $E_{R_i \rightarrow kR_i}$  is simply the identity matrix with entry  $(i, i)$  set to  $1/k$ .

The last type of elementary row operation is to swap two rows. Choose  $i$  and  $j$  to be arbitrary row indices. Consider an arbitrary vector  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ . The mapping  $L_{R_i \leftrightarrow R_j}$  will be the linear mapping that swaps row  $i$  with row  $j$ :

$$L_{R_i \leftrightarrow R_j} \left( \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_i \\ \vdots \\ x_j \\ \vdots \end{bmatrix} \right) = \begin{array}{c} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{array} \begin{bmatrix} \vdots \\ x_j \\ \vdots \\ x_i \\ \vdots \end{bmatrix}$$

The coefficient matrix  $E_{R_i \leftrightarrow R_j}$  for this simple linear mapping is simply the identity matrix with entries  $(i, i)$  and  $(j, j)$  set to 0, and the entries  $(i, j)$  and  $(j, i)$  set to 1. For example, given the elementary operation  $L_{R_2 \leftrightarrow R_4}$  where  $n = 5$ .

$$E_{R_2 \leftrightarrow R_4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

When a matrix is multiplied **on the left** by an elementary matrix, the rows are modified according to the corresponding elementary row operation. For example, given an arbitrary  $5 \times 5$  matrix  $A$ ,

$$\begin{aligned} E_{R_2 \leftrightarrow R_4} A &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} \end{bmatrix} \end{aligned}$$

The linear mapping  $L_{R_i \leftrightarrow R_j}$  is very easy to invert. Just swap row  $i$  with row  $j$  again:

$$L_{R_i \leftrightarrow R_j}^{-1} \left( \begin{array}{c|c} \vdots & \begin{bmatrix} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{bmatrix} \\ \hline & \begin{bmatrix} x_i \\ \vdots \\ x_j \\ \vdots \end{bmatrix} \end{array} \right) = \begin{array}{c|c} \vdots & \begin{bmatrix} \vdots \\ \text{row } i \\ \vdots \\ \text{row } j \\ \vdots \end{bmatrix} \\ \hline & \begin{bmatrix} x_j \\ \vdots \\ x_i \\ \vdots \end{bmatrix} \end{array}$$

The inverse of the elementary matrix  $E_{R_i \leftrightarrow R_j}$  is itself:  $E_{R_i \leftrightarrow R_j}^{-1} = E_{R_i \leftrightarrow R_j}$ .

The elementary row operations and their corresponding elementary matrices are what will be used to solve the equation  $A\mathbf{x} = \mathbf{y}$ . Both sides will be multiplied by elementary matrices until the matrix  $A$  is undone. As both sides are multiplied by the same elementary matrices, the equation  $A\mathbf{x} = \mathbf{y}$  will take the form  $B\mathbf{x} = C\mathbf{y}$ . Omitting the terms  $\mathbf{x}$  and  $\mathbf{y}$ , the equation can be shorthand via the  $n \times 2n$  augmented matrix  $[B|C]$ , where the first  $n$  columns consist of  $B$ , and the latter  $n$  columns consist of  $C$ . Applying an elementary row operation to the entire augmented matrix effectively applies the elementary row operation to both matrices. The augmented matrix starts with the form  $[A|I]$ , and the aim is to apply elementary row operations until the form  $[I|A^{-1}]$  is achieved. The matrix  $[I|A^{-1}]$  is in **row reduced echelon form**, since the identity matrix on the left hand side manifests a pivot in every row and column on the left hand side. If one of the columns on the left hand side does not have a pivot, then the form  $[I|A^{-1}]$  is unachievable, and the matrix  $A$  does not have an inverse and is hence **singular**. This also means that the linear mapping  $\mathbf{y} = L(\mathbf{x})$  maps different input vectors onto the same output vectors, and is not one-to-one.

**Examples:\***

- Find the inverse of

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

Creating and row reducing the augmented matrix proceeds as follows:

$$\begin{aligned}
 & \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - R_1}} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array} \right] \\
 & \xrightarrow{R_3 \rightarrow R_3 + 2R_2} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array} \right] \xrightarrow{R_3 \rightarrow -R_3} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right] \\
 & \xrightarrow{\substack{R_1 \rightarrow R_1 - 3R_3 \\ R_2 \rightarrow R_2 + 3R_3}} \left[ \begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right] \xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array} \right]
 \end{aligned}$$

Therefore  $A$  is non-singular and has the inverse matrix:

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

- Find the inverse of

$$A = \begin{bmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$$

Creating and row reducing the augmented matrix proceeds as follows:

$$\begin{aligned}
 & \left[ \begin{array}{ccc|ccc} -1 & 3 & -4 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 & 1 & 0 \\ -4 & 2 & -9 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{R_2 \rightarrow R_2 + 2R_1 \\ R_3 \rightarrow R_3 - 4R_1}} \left[ \begin{array}{ccc|ccc} -1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 10 & -7 & 2 & 1 & 0 \\ 0 & -10 & 7 & -4 & 0 & 1 \end{array} \right] \\
 & \xrightarrow{R_3 \rightarrow R_3 + R_2} \left[ \begin{array}{ccc|ccc} -1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 10 & -7 & 2 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{array} \right]
 \end{aligned}$$

The 3<sup>rd</sup> column does not contain a pivot so matrix  $A$  is singular (does not have an inverse).

- Find the inverse of

$$A = \begin{bmatrix} 1/5 & 1/5 & -2/5 \\ 2/5 & -3/5 & -3/10 \\ 1/5 & -4/5 & 1/10 \end{bmatrix}$$

Creating and row reducing the augmented matrix proceeds as follows:

$$\begin{aligned}
 & \left[ \begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 2/5 & -3/5 & -3/10 & 0 & 1 & 0 \\ 1/5 & -4/5 & 1/10 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - R_1}} \left[ \begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & -1 & 1/2 & -2 & 1 & 0 \\ 0 & -1 & 1/2 & -1 & 0 & 1 \end{array} \right] \\
 & \xrightarrow{R_3 \rightarrow R_3 - R_2} \left[ \begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & -1 & 1/2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]
 \end{aligned}$$

The 3<sup>rd</sup> column does not contain a pivot so matrix  $A$  is singular (does not have an inverse).



- Find the inverse of

$$A = \begin{bmatrix} 2 & -4 & 0 & 0 \\ 1 & 2 & 12 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & -1 & -4 & -5 \end{bmatrix}$$

Creating and row reducing the augmented matrix proceeds as follows:

$$\begin{aligned} & \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 12 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -4 & -5 & 0 & 0 & 0 & 1 \end{array} \right] \xrightarrow{R_2 \rightarrow R_2 - (1/2)R_1} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 12 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -4 & -5 & 0 & 0 & 0 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow R_4 + (1/4)R_2} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 12 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -5 & -1/8 & 1/4 & 0 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow R_4 + (1/2)R_3} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 12 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -5 & -1/8 & 1/4 & 1/2 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow -(1/5)R_4} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 12 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \\ & \xrightarrow{R_3 \rightarrow (1/2)R_3} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 12 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \\ & \xrightarrow{R_2 \rightarrow R_2 - 12R_3} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & -1/2 & 1 & -6 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \\ & \xrightarrow{R_2 \rightarrow (1/4)R_2} \left[ \begin{array}{cccc|cccc} 2 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/8 & 1/4 & -3/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \\ & \xrightarrow{R_1 \rightarrow R_1 + 4R_2} \left[ \begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1/2 & 1 & -6 & 0 \\ 0 & 1 & 0 & 0 & -1/8 & 1/4 & -3/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \\ & \xrightarrow{R_1 \rightarrow (1/2)R_1} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/4 & 1/2 & -3 & 0 \\ 0 & 1 & 0 & 0 & -1/8 & 1/4 & -3/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 & 1/40 & -1/20 & -1/10 & -1/5 \end{array} \right] \end{aligned}$$

Therefore  $A$  is nonsingular and has the inverse matrix:

$$A^{-1} = \begin{bmatrix} 1/4 & 1/2 & -3 & 0 \\ -1/8 & 1/4 & -3/2 & 0 \\ 0 & 0 & 1/2 & 0 \\ 1/40 & -1/20 & -1/10 & -1/5 \end{bmatrix}$$

- Find the inverse of

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

Creating and row reducing the augmented matrix proceeds as follows:

$$\begin{aligned} & \left[ \begin{array}{cccc|cccc} 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 5 & -3 & 0 & 0 & 0 & 1 \end{array} \right] \xrightarrow{R_1 \leftrightarrow R_2} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 5 & -3 & 0 & 0 & 0 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow R_4 - 2R_1} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 5 & -5 & 0 & -2 & 0 & 1 \end{array} \right] \\ & \xrightarrow{R_2 \leftrightarrow R_3} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & -5 & 0 & -2 & 0 & 1 \end{array} \right] \xrightarrow{R_4 \rightarrow R_4 + R_2} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 8 & -5 & 0 & -2 & 1 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow R_4 - 4R_3} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & -4 & -2 & 1 & 1 \end{array} \right] \\ & \xrightarrow{R_4 \rightarrow -(1/5)R_4} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array} \right] \\ & \xrightarrow{R_1 \rightarrow R_1 - R_4} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -4/5 & 3/5 & 1/5 & 1/5 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array} \right] \\ & \xrightarrow{R_3 \rightarrow (1/2)R_3} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -4/5 & 3/5 & 1/5 & 1/5 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array} \right] \\ & \xrightarrow{R_2 \rightarrow R_2 - 3R_3} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -4/5 & 3/5 & 1/5 & 1/5 \\ 0 & -1 & 0 & 0 & -3/2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array} \right] \\ & \xrightarrow{R_2 \rightarrow -R_2} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -4/5 & 3/5 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 3/2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array} \right] \end{aligned}$$

Therefore  $A$  is nonsingular and has the inverse matrix:

$$A^{-1} = \begin{bmatrix} -4/5 & 3/5 & 1/5 & 1/5 \\ 3/2 & 0 & -1 & 0 \\ 1/2 & 0 & 0 & 0 \\ 4/5 & 2/5 & -1/5 & -1/5 \end{bmatrix}$$

- If  $A$ ,  $B$ , and  $C$  are all nonsingular  $n \times n$  matrices, then solve the following for matrix  $D$ :

$$BADCB^{-1} = BAB^{-1}$$

Since  $A$ ,  $B$ , and  $C$  all have inverses, multiplying both sides by one of these matrices causes no loss of information. Start by multiplying both sides **on the left** by  $(BA)^{-1}$ :

$$BADCB^{-1} = BAB^{-1} \iff (BA)^{-1}BADCB^{-1} = (BA)^{-1}BAB^{-1} \iff DCB^{-1} = B^{-1}$$

Next multiply both sides **on the right** by  $(CB^{-1})^{-1}$ :

$$DCB^{-1} = B^{-1} \iff DCB^{-1}(CB^{-1})^{-1} = B^{-1}(CB^{-1})^{-1} \iff D = B^{-1}(BC^{-1}) \iff D = C^{-1}$$

Therefore:  $D = C^{-1}$

\*Examples from the problem set of chapter 1.5 of the textbook:

Anton, Howard; Rorres, Chris, *Elementary Linear Algebra 11th edition, Applications Version*, Wiley, 2014.

## The matrix transpose

Lastly, the **transpose** of a matrix is computed by converting the rows into columns and the columns into rows. If  $A$  is an  $m \times n$  matrix, then the transpose  $A^T$  is an  $n \times m$  matrix where the  $(i, j)$  entry of  $A$  is the  $(j, i)$  entry of  $A^T$ :

$$A^T = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix}$$

For example,

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \\ 7 & 0 \end{bmatrix}^T = \begin{bmatrix} 3 & -1 & 7 \\ 4 & 2 & 0 \end{bmatrix}$$

The following properties hold with the transpose, and are easy to verify:

- If  $A$  and  $B$  are both  $m \times n$  matrices, then  $(A + B)^T = A^T + B^T$
- If  $A$  is a matrix and  $c$  is a scalar, then  $(c \cdot A)^T = c \cdot A^T$
- If  $A$  is an  $m \times r$  matrix and  $B$  is an  $r \times n$  matrix, then  $(AB)^T = B^T A^T$ . The order is reversed thanks to the reversal of the rows and columns.
- $I_n^T = I_n$
- If  $A$  is a nonsingular matrix, then  $(A^{-1})^T = (A^T)^{-1}$

## Matrix inverse properties

A linear system with  $n$  variables and  $n$  equations has the form:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,n}x_n = b_n \end{cases} \iff \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Let  $\mathbf{x}$  be the vector that consists of the sought after variables,  $A$  is an  $n \times n$  matrix that consists of the coefficients, and  $\mathbf{b}$  are the constants on the right hand side of the equations. The linear system has the form:

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

This is referred to as matrix-vector form. If  $A$  is nonsingular and the inverse  $A^{-1}$  is known, then this system has the easy solution:

$$A\mathbf{x} = \mathbf{b} \iff \mathbf{x} = A^{-1}\mathbf{b}$$

When matrix  $A$  is singular (non-invertible),

- Multiple input vectors give rise to the same output vector.
- There are vectors from the codomain  $\mathbb{R}^n$  that are not the output for any input vector.

In the case of square matrices where the domain and codomain have the same number of dimensions, these two conditions are synonymous.

If  $\mathbf{u}_1$  and  $\mathbf{u}_2$  give the same output value  $\mathbf{v}$ , more specifically  $A\mathbf{u}_1 = A\mathbf{u}_2 = \mathbf{v}$ , then  $A(\mathbf{u}_1 - \mathbf{u}_2) = A\mathbf{u}_1 - A\mathbf{u}_2 = \mathbf{v} - \mathbf{v} = \mathbf{0}$ , and then  $\mathbf{x} = \mathbf{u}_2 - \mathbf{u}_1$  is a solution to the **homogeneous** linear system,  $A\mathbf{x} = \mathbf{0}$ .

If  $A$  is an arbitrary  $m \times n$  matrix, the **homogeneous** linear system is  $A\mathbf{x} = \mathbf{0}$  where  $\mathbf{x}$  is the vector of variables that are to be solved for. Right now however,  $A$  is a singular (non-invertible)  $n \times n$  matrix.

If  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  are all solutions to the homogeneous linear system, then any linear combination of these vectors,  $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k$ , is also a solution to the homogeneous linear system:

$$A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_k\mathbf{x}_k) = c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 + \dots + c_kA\mathbf{x}_k = c_1\mathbf{0} + c_2\mathbf{0} + \dots + c_k\mathbf{0} = \mathbf{0}$$

This means that if there is a nonzero solution to the homogeneous system of equations, which is the case when  $A$  is singular, then there are infinitely many solutions to the homogeneous system of equations.

If  $A$  is nonsingular, then there is only one solution to the homogeneous linear system:

$$A\mathbf{x} = \mathbf{0} \iff \mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$$

namely the zero vector, also referred to as the trivial solution.

The set of solutions to the homogeneous system of equations, also referred to as the **null space** of matrix  $A$ , can be used to characterize the set of solutions for the nonhomogeneous system  $A\mathbf{x} = \mathbf{b}$ . Let  $\mathbf{x}_0$  be any solution to the nonhomogeneous system  $A\mathbf{x} = \mathbf{b}$ . Given any nonzero (nontrivial) solution  $\mathbf{x} = \mathbf{u}$  to the homogeneous system  $A\mathbf{x} = \mathbf{0}$ , then  $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$  is also a solution to the nonhomogeneous system:

$$A(\mathbf{x}_0 + \mathbf{u}) = A\mathbf{x}_0 + A\mathbf{u} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

In summary,

- If the square matrix  $A$  is non singular, then the system  $A\mathbf{x} = \mathbf{b}$  has exactly 1 solution, namely  $\mathbf{x} = A^{-1}\mathbf{b}$ .
- If the square matrix  $A$  is singular, then the system  $A\mathbf{x} = \mathbf{b}$  has either 0 or infinitely many solutions. The homogeneous system  $A\mathbf{x} = \mathbf{0}$  has infinitely many solutions.