Unit 2F: Linear transformations

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

March 4, 2025

Summary

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc.

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.

Now, let's turn to the question:

What is this additional structure?

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.
Now, let's turn to the question:
What is this additional structure?
Relatedly, what is this matrix multiplication doing?

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.

Now, let's turn to the question:

What is this additional structure?

Relatedly, what is this matrix multiplication doing?

Relatedly, what does it "mean"?

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.

Now, let's turn to the question:

What is this additional structure?

Relatedly, what is this matrix multiplication doing?

Relatedly, what does it "mean"?

Relatedly, why is it so useful?

We know that the set of matrices forms a vector space, and so we can add them, multiply them by a scalar, etc. We also have the operation of matrix multiplication, which gives the set of matrices an additional structure.
Now, let's turn to the question:
What is this additional structure?
Relatedly, what is this matrix multiplication doing?
Relatedly, what does it "mean"?
Relatedly, why is it so useful?
We will consider the answers to these questions.

Link between the dot product and matrix multiplication:

Link between the dot product and matrix multiplication:

If $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$, then
$x \cdot y = x_1y_1 + ... + x_ny_n = \sum_{i=1}^{n} x_iy_i = y^T x.$

Link between the dot product and matrix multiplication:

If $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$, then $x \cdot y = x_1 y_1 + ... + x_n y_n = \sum_{i=1}^n x_i y_i = y^T x.$
Thus, we can consider the dot product to be a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x, y) = y^T x.$$

Link between the dot product and matrix multiplication:

If $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$, then $x \cdot y = x_1y_1 + ... + x_ny_n = \sum_{i=1}^{n} x_iy_i = y^T x$.
Thus, we can consider the dot product to be a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x, y) = y^T x.$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_j = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$.

Link between the dot product and matrix multiplication:

If $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$, then $x \cdot y = x_1 y_1 + ... + x_n y_n = \sum_{i=1}^{n} x_i y_i = y^T x$.
Thus, we can consider the dot product to be a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x, y) = y^T x.$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_j = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$. Then we can define a function $f : (\mathbb{R}^m \times \mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}^m$

Link between the dot product and matrix multiplication:

If $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}^n$, then $x \cdot y = x_1y_1 + ... + x_ny_n = \sum_{i=1}^{n} x_iy_i = y^Tx$.

Thus, we can consider the dot product to be a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x, y) = y^Tx.$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_j = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$. Then we can define a function $f : (\mathbb{R}^m \times \mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}^m$ given by

$$f(x, y) = \sum_{j=1}^{m}(\sum_{i=1}^{n} x_{ji}y_i) = y^Tx.$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_i = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$ and $y_1, ..., y_p \in \mathbb{R}^n$, where $y_k = (y_{1k}, ..., y_{nk})$ for $k = 1, 2, ..., p$.

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_i = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$ and $y_1, ..., y_p \in \mathbb{R}^n$, where $y_k = (y_{1k}, ..., y_{nk})$ for $k = 1, 2, ..., p$. Then we can define a function

$f : (\mathbb{R}^m \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^p) \to \mathbb{R}^m \times \mathbb{R}^p$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_i = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$ and $y_1, ..., y_p \in \mathbb{R}^n$, where $y_k = (y_{1k}, ..., y_{nk})$ for $k = 1, 2, ..., p$. Then we can define a function $f : (\mathbb{R}^m \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^p) \to \mathbb{R}^m \times \mathbb{R}^p$ given by

$$f(x, y) = \sum_{k=1}^{p} \sum_{j=1}^{m} (\sum_{i=1}^{n} x_{ji} y_{ik})$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_i = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$ and $y_1, ..., y_p \in \mathbb{R}^n$, where $y_k = (y_{1k}, ..., y_{nk})$ for $k = 1, 2, ..., p$. Then we can define a function $f : (\mathbb{R}^m \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^p) \to \mathbb{R}^m \times \mathbb{R}^p$ given by

$$\begin{aligned} f(x, y) &= \sum_{k=1}^{p} \sum_{j=1}^{m} (\sum_{i=1}^{n} x_{ji} y_{ik}) \\ &= \sum_{i=1}^{n} (\sum_{j=1}^{m} x_{ji})(\sum_{k=1}^{p} y_{ik}) = y^T x. \end{aligned}$$

Let $x_1, ..., x_m \in \mathbb{R}^n$, where $x_i = (x_{j1}, ..., x_{jn})$ for $j = 1, 2, ..., m$ and $y_1, ..., y_p \in \mathbb{R}^n$, where $y_k = (y_{1k}, ..., y_{nk})$ for $k = 1, 2, ..., p$. Then we can define a function $f : (\mathbb{R}^m \times \mathbb{R}^n) \times (\mathbb{R}^n \times \mathbb{R}^p) \to \mathbb{R}^m \times \mathbb{R}^p$ given by

$$\begin{aligned} f(x, y) &= \sum_{k=1}^{p} \sum_{j=1}^{m} (\sum_{i=1}^{n} x_{ji} y_{ik}) \\ &= \sum_{i=1}^{n} (\sum_{j=1}^{m} x_{ji})(\sum_{k=1}^{p} y_{ik}) = y^T x. \end{aligned}$$

From this discussion, we can say that matrix multiplication is a generalization of dot product in some sense.

Link between a dot product and a linear function:

We have viewed a dot product

as a function that fixes y and takes x as an input. In this case,

$$f(x) = f_y(x) = y^T x,$$

and this is a function that takes as input a vector $x \in \mathbb{R}^n$ and returns as output a number that is an element of $\mathbb{R}$ that equals $\sum_{i=1}^{n} x_i y_i$, where recall y is assumed to be fixed.

Now in a dot product, if we replace the fixed vector y by a fixed matrix $A = \{A_{ij}\}$,

Now in a dot product, if we replace the fixed vector y by a fixed matrix $A = \{A_{ij}\}$, then

$$f(x) = f_A(x) = Ax = \sum_{j=1}^{n} A_{ij} x_j.$$

Now in a dot product, if we replace the fixed vector y by a fixed matrix $A = \{A_{ij}\}$, then

$$f(x) = f_A(x) = Ax = \sum_{j=1}^{n} A_{ij}x_j.$$

This function f takes as input a vector $x \in \mathbb{R}^n$ and returns a vector $y \in \mathbb{R}^n$, where

$$y = \sum_{j=1}^{n} A_{ij}x_j.$$

Now in a dot product, if we replace the fixed vector y by a fixed matrix $A = \{A_{ij}\}$, then

$$f(x) = f_A(x) = Ax = \sum_{j=1}^{n} A_{ij} x_j.$$

This function f takes as input a vector $x \in \mathbb{R}^n$ and returns a vector $y \in \mathbb{R}^n$, where

$$y = \sum_{j=1}^{n} A_{ij} x_j.$$

This enables us **to view a matrix as a linear transformation**.

For example, we have noted earlier that if we consider a **random walk transition matrix**

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

For example, we have noted earlier that if we consider a **random walk transition matrix**

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

and define

$$y = Wx,$$

For example, we have noted earlier that if we consider a **random walk transition matrix**

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

and define

$$y = Wx,$$

then we take one vector in $\mathbb{R}^4$

For example, we have noted earlier that if we consider a **random walk transition matrix**

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

and define

$$y = Wx,$$

then we take one vector in $\mathbb{R}^4$ and we apply W to get another vector in $\mathbb{R}^4$.

This perspective as viewing a matrix in terms of transformations or as a function is true more generally.

Linear transformations as matrices

An $m \times n$ matrix can be used to characterize a certain class of functions (**linear transformations**) from $\mathbb{R}^m$ to $\mathbb{R}^n$, and, conversely, every linear transformation can be written as a matrix.

Linear transformations as matrices

An $m \times n$ matrix can be used to characterize a certain class of functions (**linear transformations**) from $\mathbb{R}^m$ to $\mathbb{R}^n$, and, conversely, every linear transformation can be written as a matrix.

This representation also explains the definition of matrix multiplication:

Linear transformations as matrices

An $m \times n$ matrix can be used to characterize a certain class of functions (**linear transformations**) from $\mathbb{R}^m$ to $\mathbb{R}^n$, and, conversely, every linear transformation can be written as a matrix.

This representation also explains the definition of matrix multiplication:

Applying a function twice (as in iterating the random walk matrix) corresponds exactly to doing a multiplication of the two matrices.

As an example, consider $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$. It can be viewed as a transformation with domain $\mathbb{R}^3$ and range $\mathbb{R}^2$.

As an example, consider $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$. It can be viewed as a transformation with domain $\mathbb{R}^3$ and range $\mathbb{R}^2$. For example, if $v = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{R}^3$, then

$$w = Av = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \in \mathbb{R}^2.$$

Thus, the $m \times n$ matrix A transforms vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$ according to the rule

$$w_i = \sum_{j=1}^{n} A_{ij} v_j \qquad \text{for } i = 1, 2, ..., m.$$

This is a linear function.

More abstractly, here is the definition of a linear transformation.

Linear transformation

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** if

1. $\forall\ x, y \in \mathbb{R}^n \quad T(x + y) = T(x) + T(y)$,
2. $\forall\ x \in \mathbb{R}^n,\ \forall\ \alpha \in \mathbb{R} \quad T(\alpha x) = \alpha T(x)$.

Observe that $T(0) = 0$. This indicates that a linear transformation goes through the origin.

Observe that $T(0) = 0$. This indicates that a linear transformation goes through the origin.

Affine transformation

An affine transformation is a transformation that is a linear transform plus a constant offset.

Observe that $T(0) = 0$. This indicates that a linear transformation goes through the origin.

Affine transformation

An affine transformation is a transformation that is a linear transform plus a constant offset.

In particular, this means that the function would be a linear transformation if it were modified by taking the image of the origin and transforming it back to the origin.

In $\mathbb{R}$, this is just another way of saying that

$$y_1 = ax_1 + b, \quad \text{for } b \neq 0,$$

is not a linear function, but $y_1 = ax_1$ is a linear function. This holds true more generally.

For example, a mapping

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix}$$

is affine.

For example, a mapping

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix}$$

is affine. This is because

$$\begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

For example, a mapping

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix}$$

is affine. This is because

$$\begin{aligned} \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix} &= \begin{pmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \end{aligned}$$

For example, a mapping

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix}$$

is affine. This is because

$$\begin{aligned} \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix} &= \begin{pmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \end{aligned}$$

i.e., $x \mapsto Ax + b$, where the first term is the linear transformation and the second term is the affine offset.

For example, a mapping

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix}$$

is affine. This is because

$$\begin{aligned} \begin{pmatrix} x_1 - x_2 + 2 \\ 2x_1 + x_2 + 1 \end{pmatrix} &= \begin{pmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \end{aligned}$$

i.e., $x \mapsto Ax + b$, where the first term is the linear transformation and the second term is the affine offset. Here the constant offset $b = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ is a vector offset and not a scalar offset.

Examples of nonlinear functions.

- A function given by the equation $y_1 = (x_1 - 2)^2 + 3$ is not linear.
- A function given by the equation $y_1 = \sin(x_1)$ is not a linear.

Example (linear)

1. A function given by the equation $y_1 = ax_1$ is a linear function.
2. The function that takes as input the vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and returns as output the vector
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$
Note that by writing the RHS of this equation as a matrix-vector product, we obtain
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Example

The composition of two (and thus more than two) linear functions is a linear function.

For example, if $f(x) = ax$ and $g(x) = bx$ (where $x, a, b, f(x), g(x)$ are all real numbers in $\mathbb{R}$), then

$$f(g(x)) = a(bx) = (ab)x,$$

which is a linear function.

Consider two linear functions that take as input vectors and returns as output vectors defined by

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Then we have

$$\begin{aligned} & \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ \Rightarrow \quad & \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

Clearly, the last equation defines a linear function.

Notice that a linear function can be represented as a matrix. Let us return to the equation considered above:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{1}$$

Put

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

$$f = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

We now observe that Equation (1) can be rewritten as follows:

$$y = f(x).$$

In general, if we have a linear function $f : \mathbb{R}^m \to \mathbb{R}^n$, then it can be fully described by an $m \times n$ matrix.

Recall that the multiplicative inverse of a number such as 5 is $1/5$ or 5^{-1}. This inverse satisfies the equations

$$5^{-1} \cdot 5 = 1 \text{ and } 5 \cdot 5^{-1} = 1.$$

The matrix generalization requires both equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative.
Furthermore, a full generalization is possible only if the matrices involved are square matrices.

We need the identity matrix in order to define and explain the inverse matrix. So, we begin with the identity matrix.

What does the term identity matrix mean?

Identity matrix

The **identity matrix** is a matrix denoted by I such that

$$AI = A \text{ for any matrix } A$$

What does the identity matrix look like?

An identity matrix is a square matrix defined by

$$I = (I_{kj}) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}$$

This means, all the diagonal elements of a matrix I are 1:

$$i_{11} = i_{22} = i_{33} = ... = 1$$

and all the other entries are zero.

Example.

For a 2×2 matrix we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

where a, b, c and d are real numbers.

Invertible

A square matrix A is said to be **invertible** or **non-singular** if there is a matrix B of the same size such that

$$AB = BA = I.$$

Matrix B is called the (multiplicative) **inverse** of A and is denoted by A^{-1}.

Here is a basic result about inverses.

Theorem

If A and B are invertible, then AB is invertible, and

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

Proof.

… $\square$

From a linear equation perspective, inverse matrices are important. For example, a system of linear equations can generally be written as $Ax = b$ where x is the vector of unknowns that we need to find. If we multiply both sides of this $Ax = b$ by the inverse matrix A^{-1} we obtain:

$$\begin{aligned} & A^{-1}Ax = A^{-1}b \\ \Rightarrow \quad & x = A^{-1}b. \end{aligned}$$

Hence we can find the unknowns by finding the inverse matrix.

Question. When does such an inverse matrix exist, i.e., when is a matrix invertible?

Answer. Here is the partial answer.

Case $m = n = 1$: We can write a 1×1 matrix as $A = (a)$. In this case, $A^{-1} = 1/a$, which is defined for all $a \neq 0$. So, for 1×1 matrices, i.e., real numbers, such a number exists for every $a \in \mathbb{R}$ such that $a \neq 0$. This is particularly simple, and the situation is considerably more subtle for matrices of larger size.

Case $m = n = 2$: We can write a 2×2 matrix as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

in which case

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

which is defined only when $ad - bc \neq 0$, and otherwise the inverse is not defined.

Check that the above matrix A^{-1} is the inverse of the matrix A, and vice versa

A 2×2 matrix A doesn't have an inverse in two cases:

A 2×2 matrix A doesn't have an inverse in two cases:

1. When it is the all-zeros matrix, in which case it doesn't have any non-trivial information and it sends all input vectors to the zero vector; and
2. when it's two columns are the same, up to scaling, in which case you might imagine that it is missing some information. Note that this degeneracy corresponds to multiplication by a scalar, i.e., one column is a scalar multiple of the other column. Note also that if $\alpha = 0$, then the second column is the all-zeros column, and it still holds that the matrix is not invertible.

The quantity $ad - bc$ is called the **determinant** of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and we write

$$\det A = ad - bc.$$

We also write

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

For the case of 3×3 matrices, if

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix},$$

the determinant of this 3×3 matrix is defined by

$$\det A = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

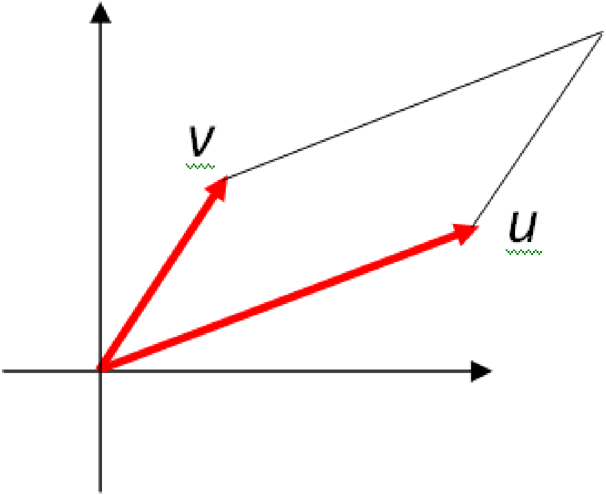
In this way, we can also define the determinant of an $n \times n$ matrix with $n \geq 4$.

From the above discussion, we observe that

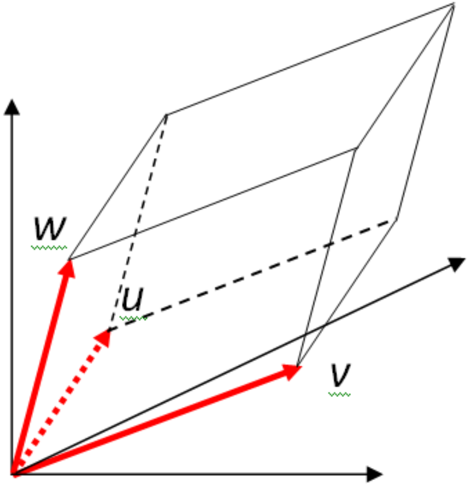
Determinant as a function

$\det(\cdot)$ is a real-valued function defined on the set of all square matrices.

Geometrical interpretation of determinants.
For a 2×2 matrix, if we consider the parallelogram defined by the two columns (or rows) of the matrix, for example, $x = \begin{pmatrix} a \\ c \end{pmatrix}$ and $y = \begin{pmatrix} b \\ d \end{pmatrix}$, then the determinant equals the area of the parallelogram.



Clearly, this equals zero if the two columns are linearly dependent, i.e., if one is a scalar multiple of the other, and thus they point in the same direction, For a 3×3 matrix, if we consider the parallelepiped defined by the three columns (or rows) of the matrix, then the determinant equals the volume of the parallelepiped.



Clearly, this equals zero if the three columns are linearly dependent, i.e., if one is a scalar multiple of another or if one can be obtained by a linear

Some important theorems

Here are two important theorems that we won't prove but that the above discussion suggests.

Theorem.

Let A be an $m \times n$ matrix. Then, A defines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ by matrix multiplication: $T(v) = Av$, where v is a column vector.

Theorem.

Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is given by an $m \times n$ matrix, call it A. The functional form is given by $T(v) = Av$, i.e., a matrix -vector multiplication, where th jth column of A is $T(e_j)$..

The two theorems given above relate matrices and linear transformations more precisely.

The two theorems given above relate matrices and linear transformations more precisely. Due to them, we will sometimes, but not always, distinguish between a linear transformation T and its associated matrix A.

The two theorems given above relate matrices and linear transformations more precisely. Due to them, we will sometimes, but not always, distinguish between a linear transformation T and its associated matrix A.

The second theorem is powerful and surprising.

The two theorems given above relate matrices and linear transformations more precisely. Due to them, we will sometimes, but not always, distinguish between a linear transformation T and its associated matrix A.

The second theorem is powerful and surprising. It says not just that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is given by a matrix; it also says that one can construct the matrix by seeing how the transformation acts on the standard basis vectors. This is rather remarkable.

For completeness, we note the following results, which states that if we have a linear transformation corresponding to a matrix, then the inverse linear transformation corresponds to the inverse matrix.

Theorem.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is invertible iff the $m \times n$ matrix A associated with it is invertible, and $T^{-1} = A^{-1}$.

Note that only square matrices can have inverses. Thus for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ to be invertible, we must have $m = n$.

Standard basis vectors

Standard basis vectors

The **standard basis vectors** (or **standard unit vectors**) in $\mathbb{R}^n$, denoted e_k, have n entries, with a 1 in the kth position and a 0 in all the other positions.

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad k\text{th position}$$

Figure 1: kth standard basis vector

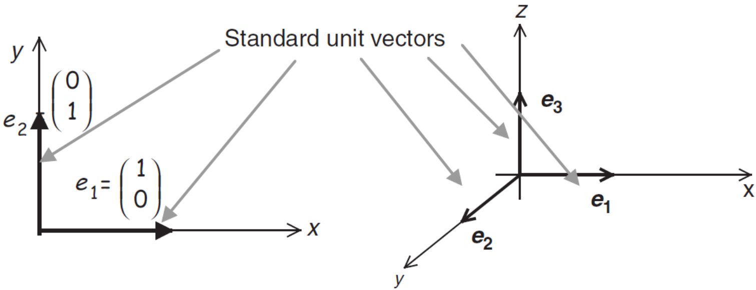


Figure 2: Standard basis vectors in the plane and in the space.

Question: Why are these standard basis vectors important?

Question: Why are these standard basis vectors important?
Ans. Every vector in the space can be expressed as a linear combination of the basis vectors of the space. It is easier to work with linear combinations.

Consider a vector $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \in \mathbb{R}^2$.

Consider a vector $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \in \mathbb{R}^2$. It can be written as

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Consider a vector $x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \in \mathbb{R}^2$. It can be written as

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This gives

$$x = x_1 e_1 + x_2 e_2,$$

where $x_1 = 2$ and $x_2 = 3$.

Thus, we have expressed x as a *linear combination* of the basis vectors e_1 and e_2, Thus, we have expressed x as a *linear combination* of the basis vectors e_1 and e_2, and the coefficients involved in the linear combination are the components of the vector x.

Further, consider a linear combination:

$$x_1 e_1 + x_2 e_2,$$

Further, consider a linear combination:

$$x_1 e_1 + x_2 e_2,$$

Now the question is:

Under what conditions is this linear combination zero?

that is,

$$\textit{When does the equality } x_1 e_1 + x_2 e_2 = 0 \textit{ hold?}$$

To answer to this question, put

$$x = x_1e_1 + x_2e_2.$$

To answer to this question, put

$$x = x_1 e_1 + x_2 e_2.$$

Then

$$x = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

To answer to this question, put

$$x = x_1 e_1 + x_2 e_2.$$

Then

$$\begin{aligned} x &= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix} \end{aligned}$$

To answer to this question, put

$$x = x_1 e_1 + x_2 e_2.$$

Then

$$\begin{aligned} x &= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

If $x_1e_1 + x_2e_2 = 0$, then

$$x = 0 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Rightarrow x_1 = x_2 = 0.$$

If $x_1e_1 + x_2e_2 = 0$, then

$$x = 0 \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Rightarrow x_1 = x_2 = 0.$$

On the otherhand,

$$x_1 = x_2 = 0 \Rightarrow x_1e_1 + x_2e_2 = 0.$$

Therefore,

$$x_1e_1 + x_2e_2 = 0 \Leftrightarrow x_1 = x_2 = 0.$$

Problem

1. Show that any vector in $\mathbb{R}^3$ can be expressed as a linear combination of the three unit basis vectors in $\mathbb{R}^3$. Also, show that a linear combination of the three unit basis vectors in $\mathbb{R}^3$ equals to 0 if and only if all coefficients in the linear combination are zeros.
2. Do the above problem for $\mathbb{R}^n$.

Multiplying a matrix by a standard basis vector

Multiplying a matrix A by the standard basis vector e_i selects out the ith column of A.

This is shown in the following example.

Example

Show that the second column of $A = \begin{pmatrix} 3 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 4 & 4 \\ 1 & 0 & 2 \end{pmatrix}$ is Ae_2.

Example

Show that the second column of $A = \begin{pmatrix} 3 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 4 & 4 \\ 1 & 0 & 2 \end{pmatrix}$ is Ae_2.

Solution. We have

$$Ae_2 = \begin{pmatrix} 3 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 4 & 4 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 4 \\ 0 \end{pmatrix} = A_{:2}.$$

$\square$

Generalizing it we get

jth column of a matrix

The jth column of a matrix A is Ae_j, that is,

$$Ae_j = A_{:j}.$$

jth column of the product AB

The jth column of the product AB is the product of A and the jth column of B.

Example

The second column of the product AB is the product of A and the second column of B.

Example

The second column of the product AB is the product of A and the second column of B.

To show it, let

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 4 & -2 \\ 3 & 0 & 2 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 4 & -2 \\ 3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 8 & -6 \\ 9 & 12 & -6 \end{pmatrix}.$$

Now, we find the product $AB_{:2}$.

$$AB_{:2} = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = (AB)_{:2}.$$

The next theorem is a key result for application of matrix operations in data science.

Theorem.

Let $S : \mathbb{R}^\ell \to \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^\ell$ be linear transformations, given by matrices A and B, respectively. Then, the composition $S \circ T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and is given by AB.

Scheme of proof:

- A composition of linear transformations is linear.
- Each column of the matrix representing the compositon coincides with corresponding column of AB.

Proof.

Let $v, w \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$. Then

$$\begin{aligned}(S \circ T)(av + bw) &= S(T(av + bw)) \\ &= S(aT(v) + bT(w)) \\ &= aS(T(v)) + bS(T(w)) \\ &= a(S \circ T)(v) + b(S \circ T)(w).\end{aligned}$$

This shows that $S \circ T$ is linear. Then it can be represented by a matrix, say, C.

Proof ...

For any basis vector e_i of $\mathbb{R}^n$ we then have

$$\begin{aligned} Ce_i &= (S \circ T)(e_i) = S(T(e_i)) \\ &= S(Be_i) = A(Be_i) \\ &= (AB)e_i. \end{aligned}$$

This implies that each column of C is equal to the corresponding column of AB and so,

$$C = AB.$$

Therefore, $S \circ T$ is given by AB. $\square$

Some special matrices

Symmetric matrices are important in general but are very important in data science.

Symmetric matrices are important in general but are very important in data science.

Symmetric matrix

A symmetric matrix is a matrix that equals its transpose.

This means, a matrix $A = (A_{ij})$ is symmetric iff
$$A_{ij} = A_{ji}.$$

For example, the matrix

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

is a symmetric matrix, while the matrix

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

is not a symmetric matrix.

Triangular matrices

An **upper triangular matrix** is a square matrix with non-zero entries only on or above the diagonal.

An **lower triangular matrix** is a square matrix with non-zero entries only on or below the diagonal.

For example, the matrix

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

is neither upper-triangular nor lower-triangular, the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{pmatrix}$$

is upper triangular but not lower triangular.

The matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

is both upper triangular and lower triangular.

Diagonal matrices.

A **diagonal matrix** is a square matrix with nonzero entries (if any) only on the main diagonal.

Diagonal matrices.

A **diagonal matrix** is a square matrix with nonzero entries (if any) only on the main diagonal.

For example, the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

is a diagonal matrix.

<u>Note that</u> a diagonal matrix is both upper triangular and lower triangular.

Diagonal matrices have many nice properties. For example,

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}^k = \begin{pmatrix} A_{11}^k & 0 \\ 0 & A_{22}^k \end{pmatrix}.$$

Diagonal matrices have many nice properties. For example,

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}^k = \begin{pmatrix} A_{11}^k & 0 \\ 0 & A_{22}^k \end{pmatrix}.$$

Due to this property, if $0 < A_{22} < 1$, then we obtain

$$\begin{pmatrix} 1 & 0 \\ 0 & A_{22} \end{pmatrix}^k = \begin{pmatrix} 1 & 0 \\ 0 & A_{22}^k \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{as } k \to \infty.$$

Examples of matrices as transformations

We saw that every matrix represents a linear transformation, and vice versa. In spite of that, it can be difficult to tell what exactly an arbitrary matrix is "doing" when it is just presented as an array of numbers.

- Fortunately, there are some basic examples of matrices that are relatively simple to think about, in the sense that their associated transformations are relatively simple to understand.

- Fortunately, there are some basic examples of matrices that are relatively simple to think about, in the sense that their associated transformations are relatively simple to understand.
- In some cases, these basic examples form the "building blocks" of arbitrary matrices.

- Fortunately, there are some basic examples of matrices that are relatively simple to think about, in the sense that their associated transformations are relatively simple to understand.
- In some cases, these basic examples form the "building blocks" of arbitrary matrices.
- When that happens, one can express the arbitrary matrix in terms of those building blocks; these are often called **matrix decompositions**.

Matrix decompositions are useful for two things: They can help

- To understand structure in data (by describing what a matrix is "doing," and thus what is happening in data that are being represented by the matrix, in terms of simpler operations).
- To perform computations faster (by describing a difficult-to-work with matrix in terms of simpler easier-to-implement parts).

Now, we will give several examples.

1. Identity: The identity transformation $I : \mathbb{R}^n \to \mathbb{R}^n$ is linear and is given by the matrix

$$I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

This is the trivial transformation that doesn't do anything.

2. Scaling: The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that takes any input vector and multiplies it by $a \in \mathbb{R}$ is given by

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

2. Scaling: The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that takes any input vector and multiplies it by $a \in \mathbb{R}$ is given by

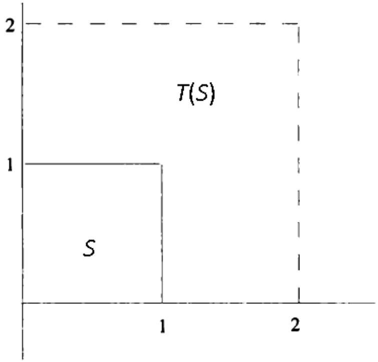
$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

For example,

$$Ae_1 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix},$$

$$Ae_2 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ a \end{pmatrix}.$$

The following figure shows the result of applying T represented by the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ to the unit square S.



Note that

$$Ax = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \end{pmatrix} = a \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax.$$

3. Stretching: The transformation T given by

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{with } a \neq b$$

is like scaling, except that the scaling is by a different amount in each direction.

3. Stretching: The transformation T given by

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{with } a \neq b$$

is like scaling, except that the scaling is by a different amount in each direction. For example, since

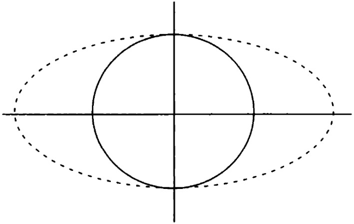
$$Ae_1 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix},$$

$$Ae_2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix},$$

we have that

$$Ax = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 \\ bx_2 \end{pmatrix}.$$

The linear transformation given by the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ stretches the unit circle into an ellipse.



Note that although we are calling this stretching, if one of the entries is negative, it might be a reflection.

Note that although we are calling this stretching, if one of the entries is negative, it might be a reflection. For example, if $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then

$$Ax = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

This is a reflection about the line $x_1 = 0$.

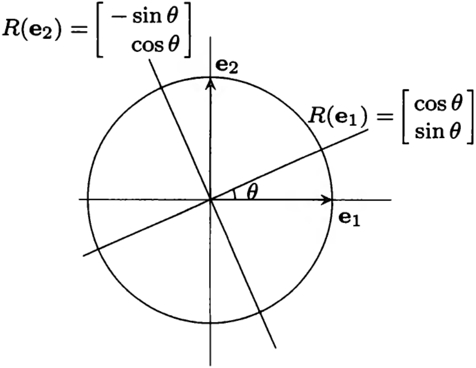
4. Rotation: The transformation given by the matrix $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ involves rotating by an angle of θ.

4. Rotation: The transformation given by the matrix $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ involves rotating by an angle of θ. In particular,

$$Re_1 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix},$$

$$Re_2 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

See the figure given below.


$R(\mathbf{e}_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$
$\mathbf{e}_2$
$R(\mathbf{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$
θ
$\mathbf{e}_1$

Consider an adjacency matrix:

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Divide each (vertical) column by the sum of the entries in that column.

$$\begin{pmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 1 & 0 & 0 \end{pmatrix}.$$

The resulting matrix is called the **random walk matrix**.

In this example, we look for the **meaning of the matrix product from a transformation perspective**. Let S and T be matrix transformations defined by

$$S(y) = Ay \text{ and } T(x) = Bx,$$

where

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{pmatrix}.$$

In this example, we look for the **meaning of the matrix product from a transformation perspective**. Let S and T be matrix transformations defined by

$$S(y) = Ay \text{ and } T(x) = Bx,$$

where

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{pmatrix}.$$

(a) What are the domains and codomains of S and T?

In this example, we look for the **meaning of the matrix product from a transformation perspective**. Let S and T be matrix transformations defined by

$$S(y) = Ay \text{ and } T(x) = Bx,$$

where

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{pmatrix}.$$

(a) What are the domains and codomains of S and T? Why is the composite transformation $S \circ T$ defined?

In this example, we look for the **meaning of the matrix product from a transformation perspective**. Let S and T be matrix transformations defined by

$$S(y) = Ay \text{ and } T(x) = Bx,$$

where

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{pmatrix}.$$

(a) What are the domains and codomains of S and T? Why is the composite transformation $S \circ T$ defined? What are the domain and the codomain of $S \circ T$?

(b) Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $T(x)$.

(b) Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $T(x)$.

(c) Find $(S \circ T)(x)$.

(b) Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $T(x)$.

(c) Find $(S \circ T)(x)$.

(d) Find a matrix C so that $(S \circ T)(x) = Cx$.

(b) Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $T(x)$.

(c) Find $(S \circ T)(x)$.

(d) Find a matrix C so that $(S \circ T)(x) = Cx$.

(e) Why is it reasonable to define AB to be the matrix C. Does the matrix C agree with the AB.

(b) Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Determine $T(x)$.

(c) Find $(S \circ T)(x)$.

(d) Find a matrix C so that $(S \circ T)(x) = Cx$.

(e) Why is it reasonable to define AB to be the matrix C. Does the matrix C agree with the AB.

(f) Show that S, T and $S \circ T$ are linear.