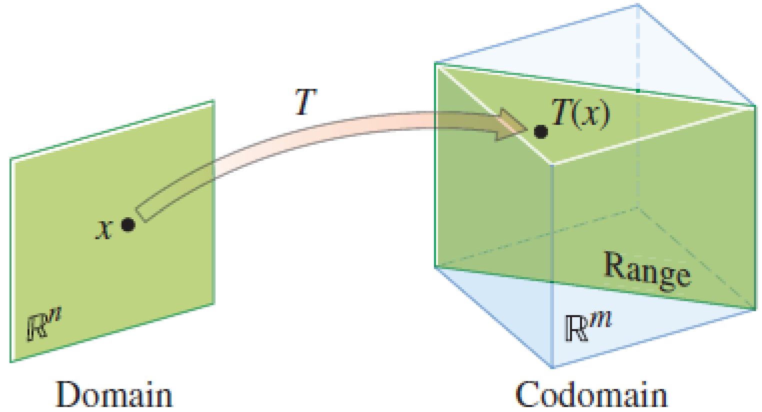
GTDM – 2019/20

LINEAR TRANSFORMATIONS

Based on Linear Algebra and Its Applications, David C. Lay, Steven R. Lay, and Judi J. McDonald, PEARSON 5^th ed.

A **transformation** (or **function** or **mapping**) T from $\mathbf{R}^n$ to $\mathbf{R}^m$ is a rule that assigns to each vector **x** in $\mathbf{R}^n$ a vector T(**x**) in $\mathbf{R}^m$



The set $\mathbf{R}^n$ is called the **domain** of T, and $\mathbf{R}^m$ is called the **codomain** of T. For **x** in $\mathbf{R}^n$, the vector T(**x**) in $\mathbf{R}^m$ is called the **image** of **x** (under the action of T). The set of all images T(**x**) is called the **range** of T.

Linear Transformations

Linear Transformations satisfy:

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

- **u** and **v** are vectors
- *a* and *b* are scalars

Linear mapping T from R^n to R^m can be expressed by using a *m x n* matrix A.

Example. The linear transformation T from R^3 to R^2 is defined as,

$$T\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 + 2\,u_2 \\ 3u_2 + 4u_3 \end{pmatrix}$$

Can be written as $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 4 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$

For each **x** in R^n, T(**x**) is computed as A**x**, where A is an m x n matrix. For simplicity, we sometimes denote such a *matrix transformation* by $\mathbf{x} \rightarrow A\mathbf{x}$.

Example.

Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, and

define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

Solve T (**x**) = **b** for **x**, that is, solve A**x** =**b**, means to find an **x** whose image under T is **b**.

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$$

Remark. The question of a *uniqueness* problem for a system of linear equations, translated here into the language of matrix transformations: Is **b** the image of a *unique* **x** in $\mathbf{R}^n$. Similarly, for the *existence* problem: does there *exist* an **x** whose image is **b?**

Identity matrix $I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$, for example $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Usually want a "formula" for T(**x**), Every linear transformation from R^n to R^m is actually a matrix transformation **x** → A**x** and that important properties of T are intimately related to familiar properties of A.
The key to finding A is to observe that T is completely determined by what it does to the columns of the n x n identity matrix I_n.

Example. The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Suppose T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ such that

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

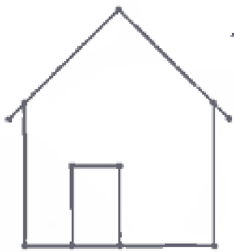
With no additional information, find a formula for the image of an arbitrary $\mathbf{x}$ in $\mathbb{R}^2$.

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

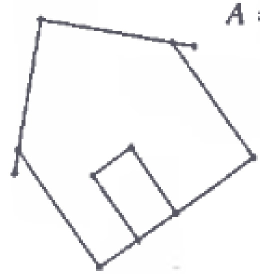
Since T is a *linear* transformation,

$$\begin{aligned} T(\mathbf{x}) &= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \\ &= x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix} \end{aligned}$$

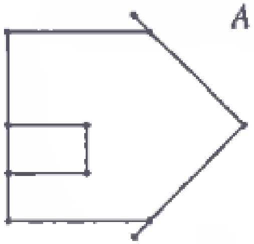
Examples of transformations $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ with their associated matrices.



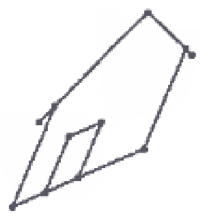
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



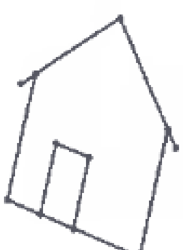
$$A = \begin{bmatrix} \cos 35^\circ & -\sin 35^\circ \\ \sin 35^\circ & \cos 35^\circ \end{bmatrix}$$



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



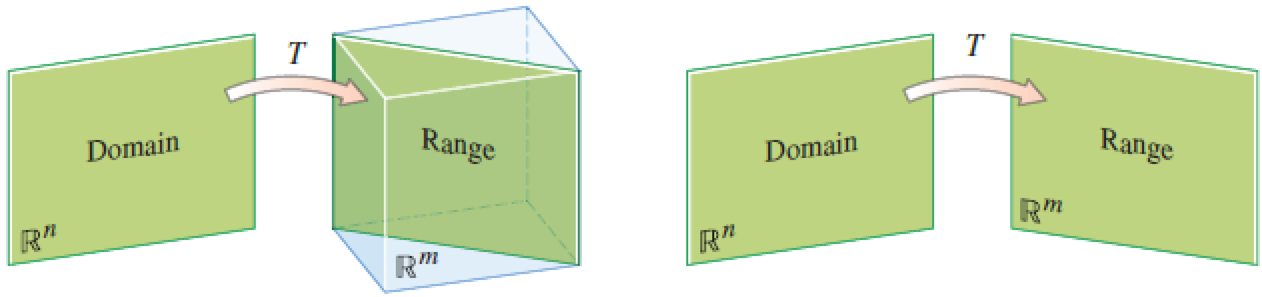
$$A = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$



$$A = \begin{bmatrix} 0.7 & 0.2 \\ -0.3 & 0.9 \end{bmatrix}$$

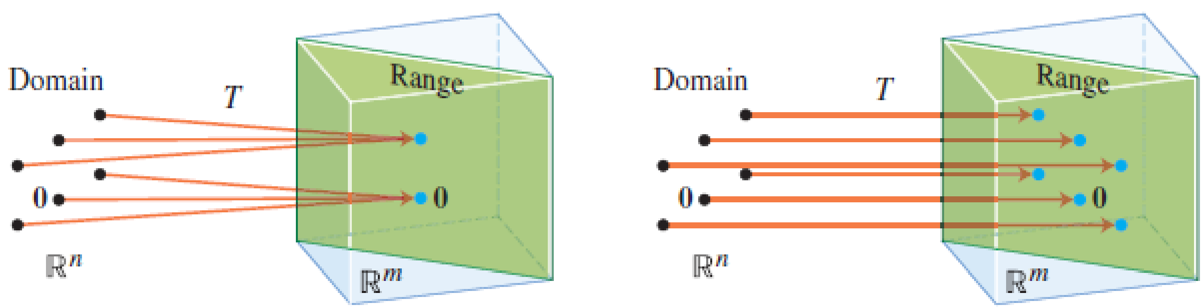
$$A = \begin{bmatrix} 0 & 1.1 \\ 0.1 & 0.3 \end{bmatrix}$$

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$.



T is *not* onto $\mathbb{R}^m$ T is onto $\mathbb{R}^m$

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most one* $\mathbf{x}$ in $\mathbb{R}^n$.



T is *not* one-to-one T is one-to-one

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.

Theorem

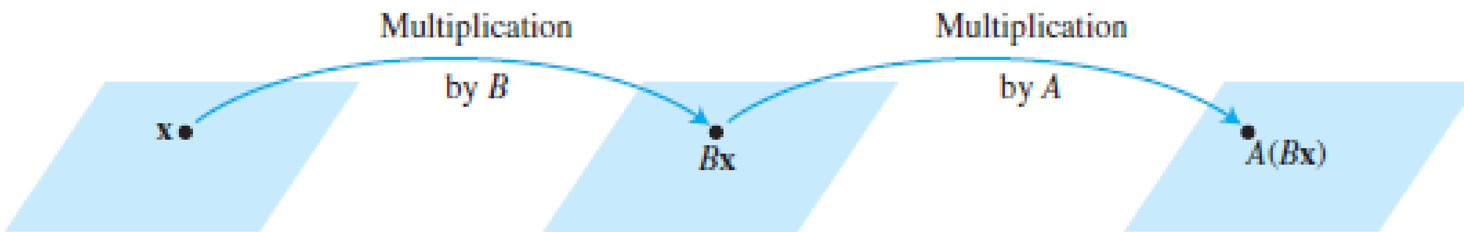
Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;

b. T is one-to-one if and only if the columns of A are linearly independent.

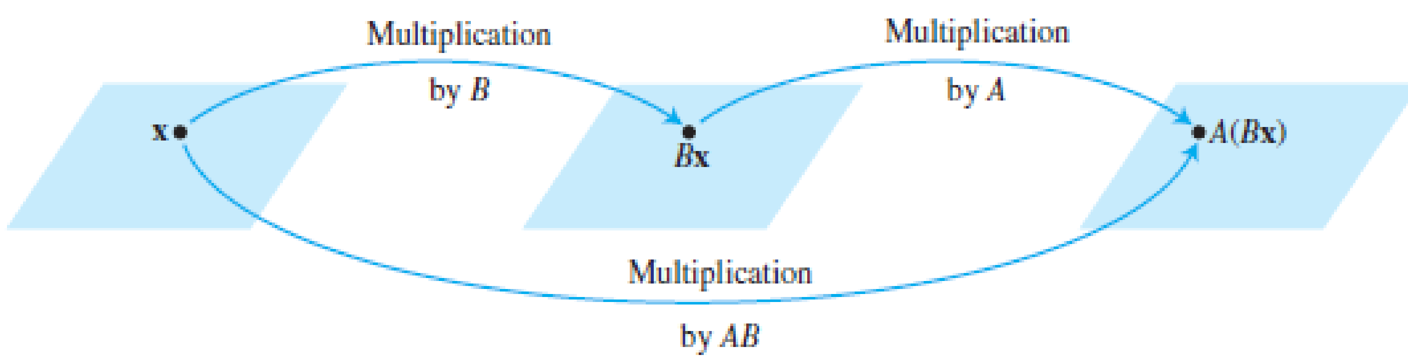
Matrix Multiplication

When a matrix B multiplies a vector **x**, it transforms **x** into the vector B**x**. If this vector is then multiplied in turn by a matrix A, the resulting vector is A(B**x**)



Thus A(B**x**) is produced from **x** by a *composition* of mappings. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by *AB*, so that

$$A(B\mathbf{x}) = (AB)\,\mathbf{x}$$



Example. Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

Write $B = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,]$, and compute:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$AB = A[\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$A\mathbf{b}_1$ $A\mathbf{b}_2$ $A\mathbf{b}_3$

ROW–COLUMN RULE FOR COMPUTING AB

If the product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Example.

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

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square \\ -4+21-12 & 6+3-8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -27 & -17 \\ 5 & 1 \\ 15 & 36 \end{bmatrix}$$

Powers of a Matrix

If A is an nxn matrix and if k is a positive integer, then A^k denotes the product of k copies of A

$$A^k = \underbrace{A \cdots A}_{k}$$

The Transpose of a Matrix

Given an m x n matrix A, the **transpose** of A is the n x m matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

THE INVERSE OF A MATRIX

An n x n matrix A is said to be **invertible** if there is an n x n matrix C such that

CA =I and AC = I

where I is the n x n identity matrix. In this case, C is an **inverse** of A. In fact, C is uniquely determined by A, because if B were another inverse of A, then B =C : this unique inverse is denoted by A^{-1}

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

Example.

If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus $C = A^{-1}$.

Matrices

1 $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ is a 3 by 2 matrix : $m = 3$ rows and $n = 2$ columns.

2 $Ax = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a **combination of the columns** $Ax = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$.

3 The 3 components of Ax are dot products of the 3 rows of A with the vector x :

Row at a time $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 2 \cdot 8 \\ 3 \cdot 7 + 4 \cdot 8 \\ 5 \cdot 7 + 6 \cdot 8 \end{bmatrix} = \begin{bmatrix} 23 \\ 53 \\ 83 \end{bmatrix}$.

4 Equations in matrix form $Ax = b$: $\begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ replaces $\begin{matrix} 2x_1 + 5x_2 = b_1 \\ 3x_1 + 7x_2 = b_2 \end{matrix}$.

5 The solution to $Ax = b$ can be written as $x = A^{-1}b$. But some matrices don't allow A^{-1}.

Inner Product of the Euclidean n-space

R^n was defined to be **the *set* of all ordered n-tuples of real numbers**. When *R^n* is combined with the standard operations of vector addition, scalar multiplication, **vector length**, and the **dot product**, the resulting vector space is called **Euclidean *n*-space**.

The dot product of two vectors is defined to be

$$\mathbf{u} \cdot \mathbf{v} = \mathrm{u}_1\mathrm{v}_1 + \mathrm{u}_2\mathrm{v}_2 + \ldots + \mathrm{u}_n\mathrm{v}_n$$

The definitions of the vector length and the dot product are needed to provide **the metric concept** for the vector space.

The length of a vector **v**$= (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n)$ in the Euclidean space **R*^n*** is given by

$$\|\mathbf{v}\| = \sqrt{\mathbf{u} \cdot \mathbf{v}} = \sqrt{\mathrm{v}_1^2 + \mathrm{v}_2^2 + \ldots + \mathrm{v}_n^2}$$

Example

(a) In $\mathbf{R}^5$, the length of $\mathbf{v}=(0\,,-2\,,1\,,4\,,-2)$ is given by

$$\|\mathbf{v}\|=\sqrt{0^2+(-2)^2+1^2+4^2+(-2)^2}=\sqrt{25}=5$$

(b) In $\mathbf{R}^3$ the length of $\mathbf{v}=(\frac{2}{\sqrt{17}}\,,\frac{-2}{\sqrt{17}}\,,\frac{3}{\sqrt{17}})$ is given by

$$\|\mathbf{v}\|=\sqrt{\left(\frac{2}{\sqrt{17}}\right)^2+\left(\frac{-2}{\sqrt{17}}\right)^2+\left(\frac{3}{\sqrt{17}}\right)^2}=\sqrt{\frac{17}{17}}=1$$

(**v** is a unit vector)

Lemma (Unit vector in the direction of **v)** If **v** is a nonzero vector in $\mathbf{R}^n$, then the vector

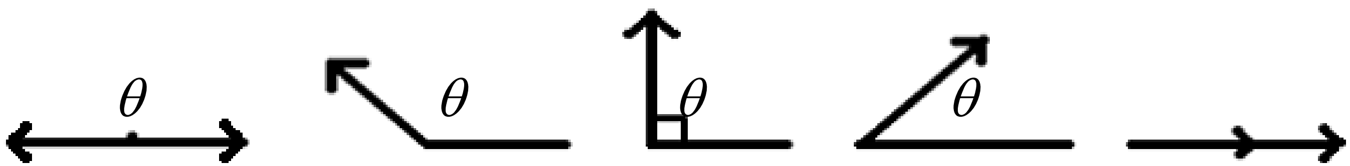
$$\mathbf{u}=\frac{\mathbf{v}}{\|\mathbf{v}\|}$$ (normalization)

has length 1 and has the same direction as **v**. This vector **u** is called the **unit vector in the direction of v.**

The angle between two vectors in $\mathbf{R}^n$:

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|},\ 0\le\theta\le\pi$$

Opposite direction	$\mathbf{u}\cdot\mathbf{v}<\mathbf{0}$	$\mathbf{u}\cdot\mathbf{v}=\mathbf{0}$	$\mathbf{u}\cdot\mathbf{v}>\mathbf{0}$	Same direction



$\theta=\pi$	$\frac{\pi}{2}<\theta<\pi$	$\theta=\frac{\pi}{2}$	$0<\theta<\frac{\pi}{2}$	$\theta=0$
$\cos=-1$	$\cos<0$	$\cos=0$	$\cos>0$	$\cos=1$

- Note:

  The angle between the zero vector and another vector is not defined.

Axioms of inner product (more general than dot product)

Let **u**, **v**, and **w** be vectors in a vector space *V*, and let *c* be any scalar. **An inner product on *V* is a function that associates a real number <u, v> with each pair of vectors u and v** and satisfies the following axioms.

(1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(2) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(3) $c \langle \mathbf{u}, \mathbf{v} \rangle = \langle c\mathbf{u}, \mathbf{v} \rangle$

(4) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = 0$

Note:

$$\mathbf{u} \cdot \mathbf{v} = \text{dot product (Euclidean inner product for } R^n\text{)}$$

$$< \mathbf{u}, \mathbf{v} > = \text{general inner product for vector space } V$$

- Ex: (A different inner product for $\mathbf{R}^n$)

Show that the function defines an inner product on R^2, where $\mathbf{u}=(u_1, u_2)$ and $\mathbf{v}=(v_1, v_2)$.

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2$$

Sol:

$$(a)\quad \langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2 = v_1 u_1 + 2v_2 u_2 = \langle \mathbf{v}, \mathbf{u} \rangle$$

$$\begin{aligned}(b)\quad \mathbf{w}=(w_1, w_2) \Rightarrow \langle \mathbf{u}, \mathbf{v}+\mathbf{w} \rangle &= u_1(v_1+w_1) + 2u_2(v_2+w_2) \\ &= u_1v_1 + u_1w_1 + 2u_2v_2 + 2u_2w_2 \\ &= (u_1v_1 + 2u_2v_2) + (u_1w_1 + 2u_2w_2) \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle\end{aligned}$$

$$(c)\quad c\langle \mathbf{u}, \mathbf{v} \rangle = c(u_1v_1 + 2u_2v_2) = (cu_1)v_1 + 2(cu_2)v_2 = \langle c\mathbf{u}, \mathbf{v} \rangle$$

$$(d)\quad \langle \mathbf{v}, \mathbf{v} \rangle = {v_1}^2 + 2{v_2}^2 \geq 0$$

$$\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow {v_1}^2 + 2{v_2}^2 = 0 \quad \Rightarrow \quad v_1 = v_2 = 0 \quad (\mathbf{v}=0)$$

Norm (length) of u from the inner product:

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} \qquad \|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle$$

Distance between **u** and **v**:

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

Angle between two nonzero vectors **u** and **v**:

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}, \ 0 \le \theta \le \pi$$

Orthogonal nonzero vectors : $(\mathbf{u} \perp \mathbf{v})$

$\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Theorem

Let **u** and **v** be vectors in an inner product space *V*.

(1) **Cauchy-Schwarz inequality**:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \le \|\mathbf{u}\| \, \|\mathbf{v}\|$$

(2) **Triangle inequality**:

$$\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$$

(3) **Pythagorean theorem** :

u and **v** are orthogonal if and only if

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

For a norm, and a distance, there are many possibilities!

Properties of norm:

(1) $\| \mathbf{u} \| \geq 0$ and $\| \mathbf{u} \| = 0$ if and only if $\mathbf{u} = \mathbf{0}$

(2) $\| c\mathbf{u} \| = |c| \, \| \mathbf{u} \|$ For any scalar c and vector **u**

(3) $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ For any vectors **u** and **v**

Examples.

(1) $\|(a_1, \ldots, a_n)\| = \sqrt{\sum_{k=1}^{n} |a_k|^2}$

(2) $\|(a_1, \ldots, a_n)\| = \sum_{k=1}^{n} |a_k|$

(3) $\|(a_1, \ldots, a_n)\| = \max_{k=1}^{n} |a_k|$

(4) $\|(a_1, \ldots, a_n)\| = \max_{k=1}^{n} k|a_k|$.

Properties of distance:

(1) $d(\mathbf{u}, \mathbf{v}) \geq 0$, and $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$

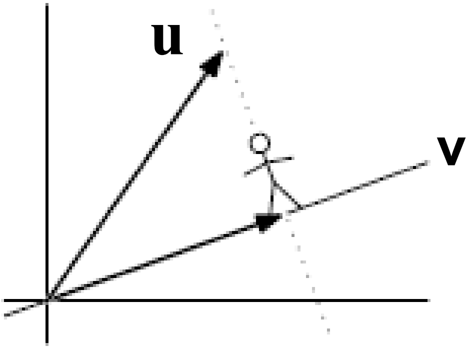
(2) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

(3) $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$

- Orthogonal projections in inner product spaces:

Let **u** and **v** be two vectors in an inner product space V, such that $\mathbf{v} \neq \mathbf{0}$. Then the **orthogonal projection of u onto v** is given by

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \frac{\langle \mathbf{u}\,,\, \mathbf{v} \rangle}{\langle \mathbf{v}\,,\, \mathbf{v} \rangle}\mathbf{v}$$



- Note:

If **v** is a init vector, then $\langle \mathbf{v}\,,\, \mathbf{v} \rangle = \|\mathbf{v}\|^2 = 1$.

The formula for the orthogonal projection of **u** onto **v** takes the following simpler form.

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \langle \mathbf{u}\,,\, \mathbf{v} \rangle \mathbf{v}$$

Example (Finding an orthogonal projection in R^3) Use the Euclidean inner product in R^3 to find the orthogonal projection of u=(6, 2, 4) onto v=(1, 2, 0).

Sol:

$$\langle \mathbf{u}, \mathbf{v} \rangle = (6)(1) + (2)(2) + (4)(0) = 10$$

$$\langle \mathbf{v}, \mathbf{v} \rangle = 1^2 + 2^2 + 0^2 = 5$$

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v} = \tfrac{10}{5}(1, 2, 0) = (2, 4, 0)$$

Note:

$\boldsymbol{u} - \text{proj}_v \boldsymbol{u} = (6, 2, 4) - (2, 4, 0) = (4, -2, 4)$ is orthogonal to $v = (1,2,0)$.

Orthonormal Bases

Orthogonal:

A set S of nonzero vectors in an inner product space V is called an **orthogonal set** if every pair of vectors in the set is orthogonal.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\} \subseteq V,$$

$$< \mathbf{v}_i \ \mathbf{v}_j > = 0 \quad i \neq j$$

Orthonormal:

An orthogonal set in which each vector is a unit vector is called **orthonormal**.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\} \subseteq V$$

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Note:

If S is a basis, then it is called an orthogonal basis or an orthonormal basis.

Theorem: If $S=\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ in $\mathbf{R}^n$ is an orthogonal set of nonzero vectors, then *S* is linearly independent and hence is a basis for the subspace spanned by *S*.

Definition: An **orthogonal basis** for a subspace *W* of $\mathbf{R}^n$ is a basis for *W* that is also an orthogonal set.

Theorem: Let $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ be an orthogonal basis for a subspace *W* of $\mathbf{R}^n$. For each **y** in *W*, the weights in the linear combination

$$\mathbf{y} = c_1\,\mathbf{u}_1 + \ldots + c_m\,\mathbf{u}_m$$

are (uniquely) given by

$$c_j = \frac{\langle \mathbf{y},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle} \qquad j=1,\ldots,m$$

Remark.
From a bases to an orthonormal Bases: the Gram-Schmidt Algorithm.

Example (A nonstandard orthonormal basis for R^3). Show that the following set is an orthonormal basis.

$$S = \left\{ \overset{\mathbf{v}_1}{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)}, \quad \overset{\mathbf{v}_2}{\left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right)}, \quad \overset{\mathbf{v}_3}{\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)} \right\}$$

Solution,

Show that the three vectors are mutually orthogonal.

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -\tfrac{1}{6} + \tfrac{1}{6} + 0 = 0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = \frac{2}{3\sqrt{2}} - \frac{2}{3\sqrt{2}} + 0 = 0$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = -\frac{\sqrt{2}}{9} - \frac{\sqrt{2}}{9} + \frac{2\sqrt{2}}{9} = 0$$

Show that each vector is of length 1: thus it is an orthonormal set

$$\|\mathbf{v}_1\| = \sqrt{\mathbf{v}_1 \cdot \mathbf{v}_1} = \sqrt{\tfrac{1}{2} + \tfrac{1}{2} + 0} = 1$$

$$\|\mathbf{v}_2\| = \sqrt{\mathbf{v}_2 \cdot \mathbf{v}_2} = \sqrt{\tfrac{2}{36} + \tfrac{2}{36} + \tfrac{8}{9}} = 1$$

$$\|\mathbf{v}_3\| = \sqrt{\mathbf{v}_3 \cdot \mathbf{v}_3} = \sqrt{\tfrac{4}{9} + \tfrac{4}{9} + \tfrac{1}{9}} = 1$$

Example. (Representing vectors relative to an orthonormal basis) Find the coordinates of **w** = (5, -5, 2) relative to the following orthonormal basis:

$$B = \{(\tfrac{3}{5}, \tfrac{4}{5}, 0), (-\tfrac{4}{5}, \tfrac{3}{5}, 0), (0, 0, 1)\}$$

$$\langle \mathbf{w}, \mathbf{v}_1 \rangle = \mathbf{w} \cdot \mathbf{v}_1 = (5, -5, 2) \cdot (\tfrac{3}{5}, \tfrac{4}{5}, 0) = -1$$

$$\langle \mathbf{w}, \mathbf{v}_2 \rangle = \mathbf{w} \cdot \mathbf{v}_2 = (5, -5, 2) \cdot (-\tfrac{4}{5}, \tfrac{3}{5}, 0) = -7$$

$$\langle \mathbf{w}, \mathbf{v}_3 \rangle = \mathbf{w} \cdot \mathbf{v}_3 = (5, -5, 2) \cdot (0, 0, 1) = 2$$

$$\Rightarrow [\mathbf{w}]_B = \begin{bmatrix} -1 \\ -7 \\ 2 \end{bmatrix}$$

Orthogonal Matrix

- *A* square matrix *A* with the property

$$A^{-1} = A^T$$

  is said to be an orthogonal matrix.

Remark

– A square matrix *A* is orthogonal if and only if $AA^T = I$ or $A^TA = I$.

- Example 1

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} \qquad A^T A = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Example 2

  Rotation and reflection matrices is orthogonal.

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Theorem

The following are equivalent for an $n \times n$ matrix A.

- A is orthogonal.
- The row vectors of A form an orthonormal set in R^n with the Euclidean inner product.
- The column vectors of A form an orthonormal set in R^n with the Euclidean inner product.

Moreover:

- The inverse of an orthogonal matrix is orthogonal.

- A product of orthogonal matrices is orthogonal.

Theorem (Orthogonal Matrices as Linear Operators)

- If A is an $n \times n$ matrix, then the following are equivalent.
  - A is orthogonal.
  - $\| A\mathbf{x} \| = \| \mathbf{x} \|$ for all $\mathbf{x}$ in R^n.
  - $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in R^n.

- Remark:
  - If $T: R^n \to R^n$ is multiplication by an orthogonal matrix A, then T is called an orthogonal operator on R^n.
  - It follows from the preceding theorem that the orthogonal operator on R^n are precisely those operators that leave the length of all vectors unchanged.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

If $ad - bc = 0$, then A is not invertible.

The quantity ad-bc is called the **determinant** of A, and we write
det A = ad-bc

Example. $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$. $\Longrightarrow$ $A^{-1} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$

If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.