Unit 8

Linearity

Albert Sung

Outlines

- 8.1 Linearity
- 8.2 Matrix-Vector Multiplication
- 8.3 Geometric Transformations
- 8.4 Vector Spaces

Unit 8.1

Linearity

Mapping of a Vector

- In this unit, we consider f: $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

$$y = f(x),$$

  where x is an n-vector and y is an m-vector.

- In general, f is a *vector-valued* function.
  - i.e., the output y is a vector.
- In the special case when $m = 1$, f is a scalar-valued function.
  - i.e., the output y is a scalar.

Linearity: An Illustration

mori soba = 800 yen

sushi = 200 yen

- Exchange rate: 100 yen = HK$ 7.00
- How much are 2 plates of mori soba and 4 plates of sushi in HK dollars?

Linear Functions

- A function f is linear if it satisfies the following two conditions:

1. Additivity
$$f(x+y) = f(x) + f(y)$$

2. Scaling
$$f(cx) = cf(x)$$

- Putting the two properties together,

Superposition
$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

- A function f is linear if it satisfies the superposition property.

In our example, f is the function to convert from yen to HK dollars.

Zero-in Zero-out Property

- A linear function f must have
$$f(\mathbf{0}) = \mathbf{0}.$$
- Why?
  - Hint: Use scaling property.
- Zero-in zero-out is a necessary condition for f to be linear.
- Is it sufficient? Why?

Class Exercises

- Is each of the following scalar-valued functions linear?

a) $f(x) = \mathbf{avg}(x) = \frac{x_1 + x_2 \ldots + x_n}{n}$

b) $f(x) = \mathbf{min}(x) = \min\{x_1, x_2, \ldots, x_n\}$

c) $f(x) = x_1 - x_2 + 4$

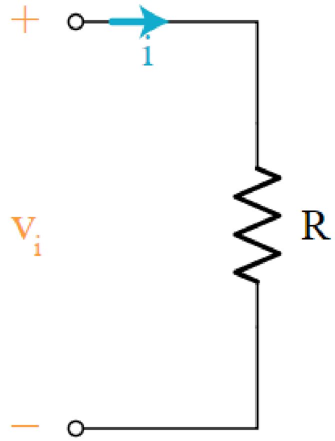
Application: Ohm's Law

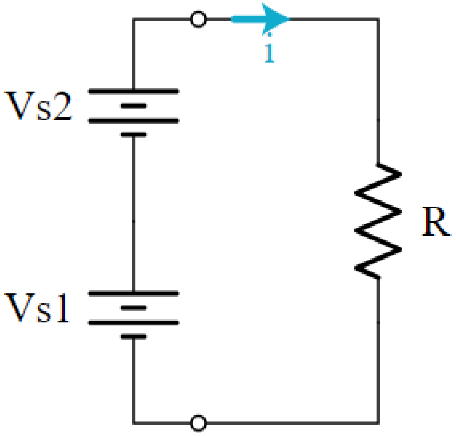
- The current is given by Ohm's Law:

$$i = f(v_i) = \frac{v_i}{R}.$$

  - The resistor can be viewed as a function that takes in a voltage and outputs a current.

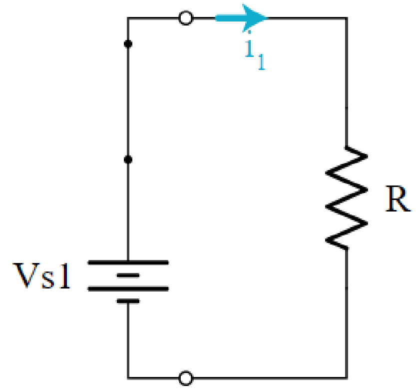
- It is a *linear* function.
- The figure on the right is a toy problem to demonstrate the use of *superposition principle*.

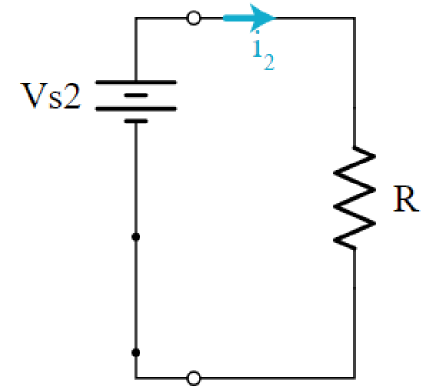




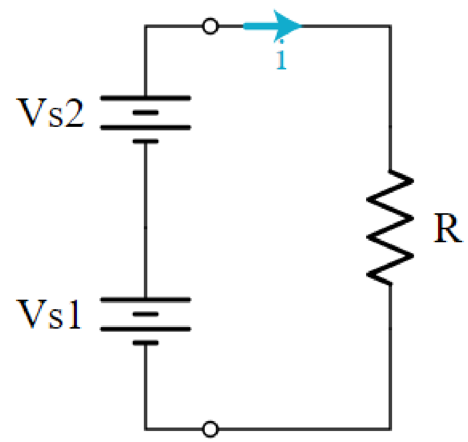
Using Superposition

- $f(V_{s1}) = \dfrac{V_{s1}}{R}$
- $f(V_{s2}) = \dfrac{V_{s2}}{R}$





- $$i = f(V_{s1} + V_{s2})$$
  $$= f(V_{s1}) + f(V_{s2})$$
  $$= \frac{V_{s1}}{R} + \frac{V_{s2}}{R}$$

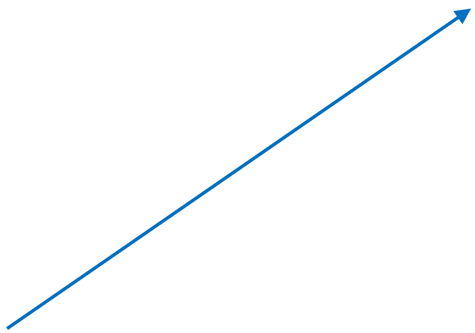


Although $f(V_{s1} + V_{s2})$ can be computed directly in this example, we demonstrate the use of superposition.

Unit 8.2

Matrix-Vector Multiplication

What is a Vector?



An arrow

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

n elements of a field putting in a certain order

An n-vector can be used to represent n quantities or values in an application.

Special Vectors

- Zero Vector:
  - $\mathbf{0}_n$: An n-vector with all entries equal to 0.
    - Sometimes simply written as $\mathbf{0}$.
- Ones Vector:
  - $\mathbf{1}_n$: An n-vector with all entries equal to 1.
    - Sometimes simply written as $\mathbf{1}$.
- Standard Unit Vectors:
  - e_i: An n-vector with all entries equal to 0 except entry i equal to 1.
  - Example: In 3-dimensional space,

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

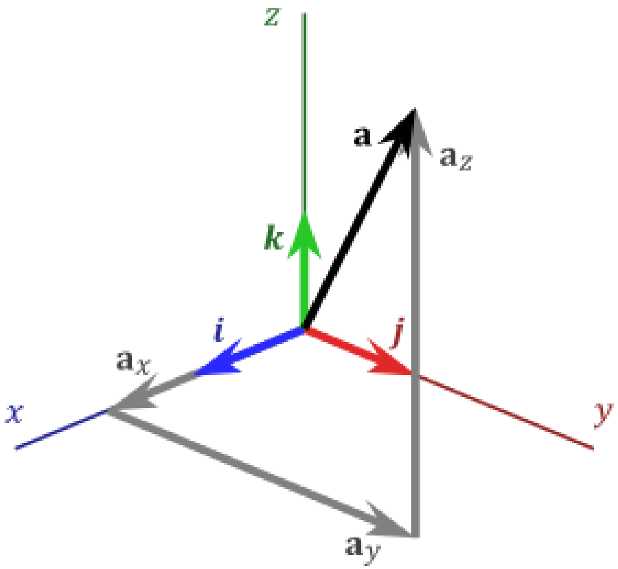
Standard Basis

- In the n-dimensional Euclidean space, $\mathbb{R}^n$, the set $\{e_1, e_2, \ldots, e_n\}$ is called standard basis.
- Any vector $x \in \mathbb{R}^n$ can be written as a linear combination of the standard basis vectors.
  - In 3-dimensional space,

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

  - Any vector a can be expressed as

$$a = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = a_x e_1 + a_y e_2 + a_z e_3.$$



Inner Product

- The inner product (or dot product) of two vectors $u = (u_1, u_{2,} \dots, u_n)$ and $v = (v_1, v_{2,} \dots, v_n)$ is defined as the scalar

$$u^T v = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$= u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Norm

- The Euclidean norm of a vector can be defined by the inner product:

$$\|x\| \triangleq \sqrt{x^T x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

which can be interpreted as the length of x.

- Is it a linear function of x?
  a) Yes
  b) No

Angle

- The angle between two vectors a and b is defined as

$$\theta \triangleq \cos^{-1}\left(\frac{a^T b}{\|a\|\|b\|}\right)$$

By Cauchy-Schwarz inequality, this ratio is always between -1 and $+1$.

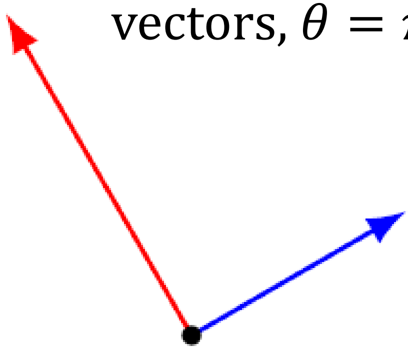
- In other words, θ is the unique number between 0 and π (in radians) that satisfies

$$\boxed{a^T b = \|a\|\|b\| \cos\theta}$$

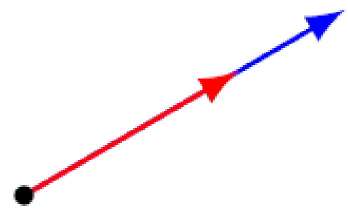
- a and b are said to be orthogonal if $\theta = \frac{\pi}{2}$.
  - Or equivalently, $a^T b = 0$.

Examples

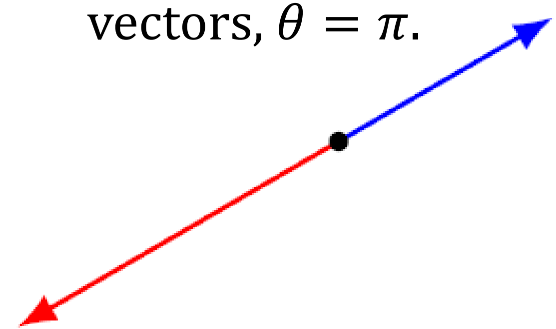
Orthogonal vectors, $\theta = \pi/2$.



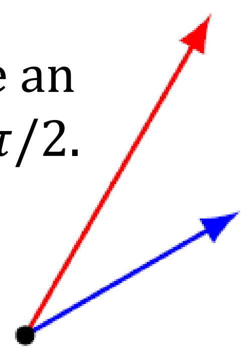
Aligned vectors, $\theta = 0$.



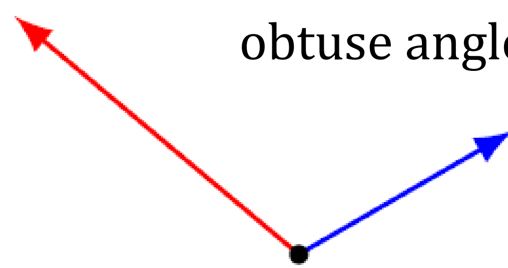
Anti-aligned vectors, $\theta = \pi$.



Vectors that make an acute angle, $\theta < \pi/2$.



Vectors that make an obtuse angle, $\theta > \pi/2$.



Matrix-Vector Multiplication

- Let A be an $m \times n$ matrix and x be an n-vector.
- Then

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix},$$

where a_i^T is the i-th row of A.

Alternative Representation

- $$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Matrix Multiplication is Linear

- Consider a function
$$f(x) = Ax,$$
where A is a matrix and x is a vector.
- It is easy to check that superposition holds:
$$\begin{aligned} f(\alpha x + \beta y) &= A(\alpha x + \beta y) \\ &= \alpha Ax + \beta Ay \qquad \text{(matrix algebra)} \\ &= \alpha f(x) + \beta f(y). \end{aligned}$$
- Hence, any matrix leads to a linear transformation.

Is the converse true?

Does every linear transformation lead to a matrix?

Representation of Linear Functions

If $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear function, then it can be represented as

$$f(x) = Ax,$$

where A is an $m \times n$ matrix, and x is an n-vector.

- To see this, we write x as $x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$.
- By superposition,

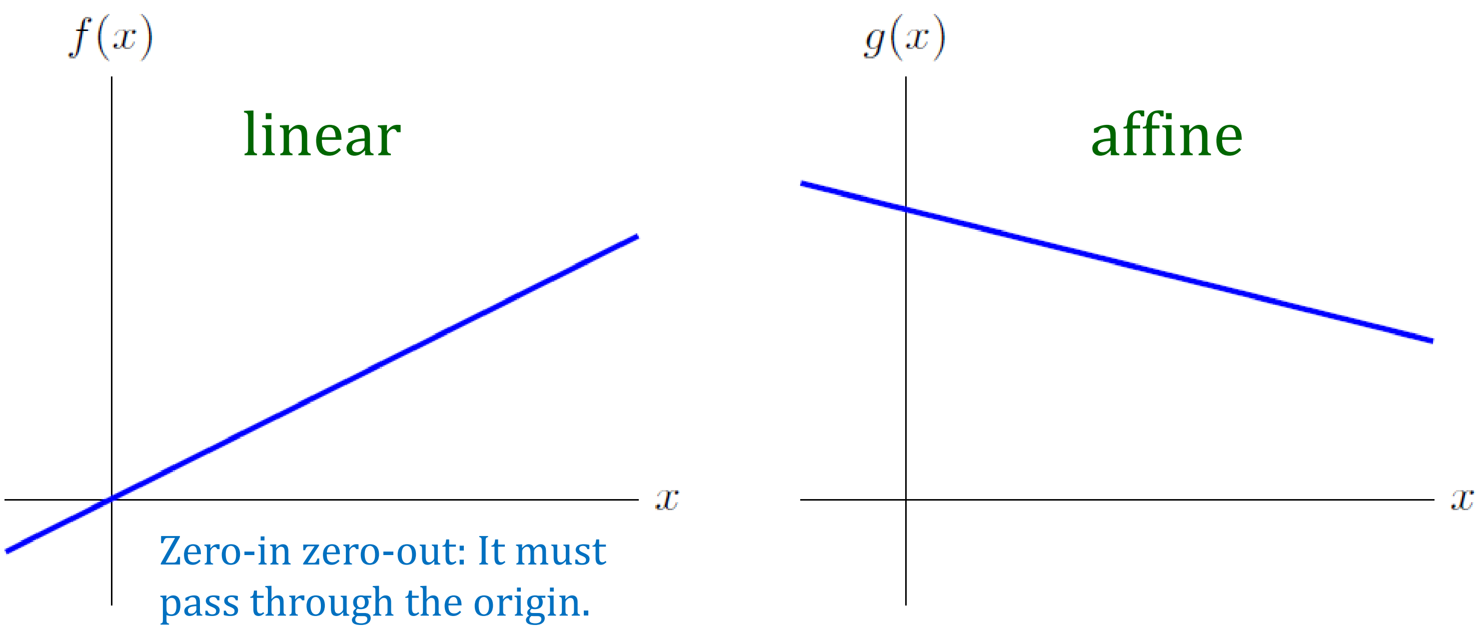
$$f(x) = x_1 f(e_1) + x_2 f(e_2) + \cdots + x_n f(e_n)$$

$$= \begin{bmatrix} | & | & \cdots & | \\ f(e_1) & f(e_2) & \ddots & f(e_n) \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

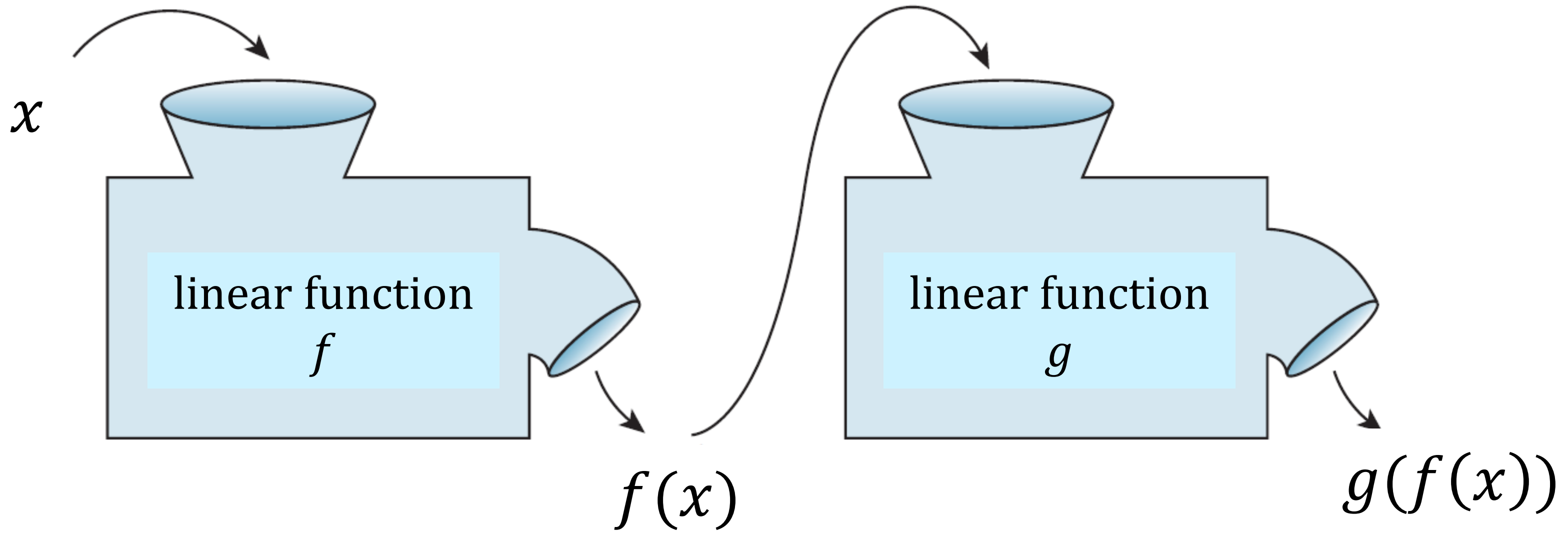
Affine Functions

- A function g is affine if it is the sum of a linear function and a constant, i.e.,

$$g(x) = Ax + c.$$



Composition of Linear Functions



- Is the composite function $g \circ f$ linear?

Linearity Preserves under Composition

- Since f and g are linear, they can be expressed as

$$f(x) = Ax \quad \text{and} \quad g(y) = By.$$

- Consider the composition:

$$g\big(f(x)\big) = B(Ax) = BAx = Cx,$$

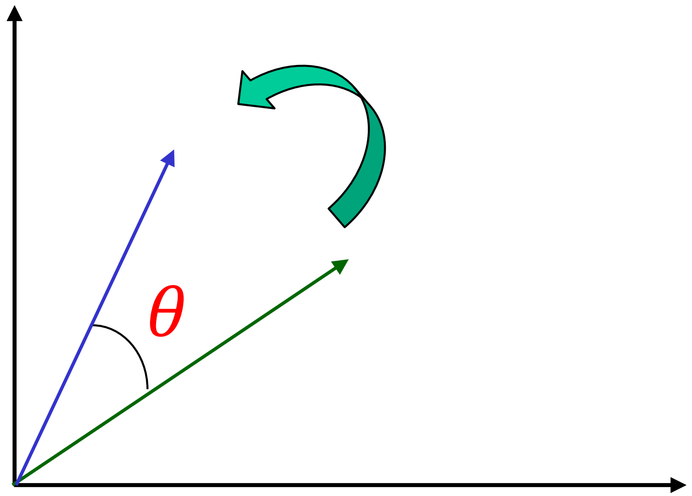
where $C = BA$ is a matrix.

- Hence, $g\big(f(x)\big)$ is a linear function of x.

Unit 8.3

Geometric Transformations

Rotation

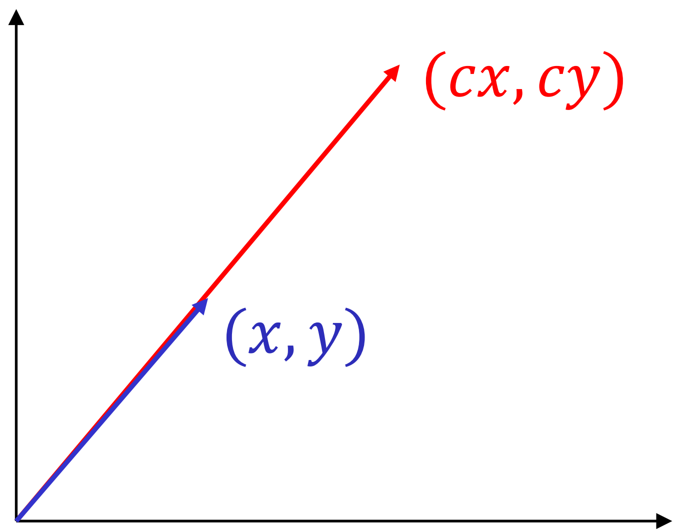


- Useful operation in computer graphics.
- Is it a linear transformation?
    a) Yes
    b) No

Examples

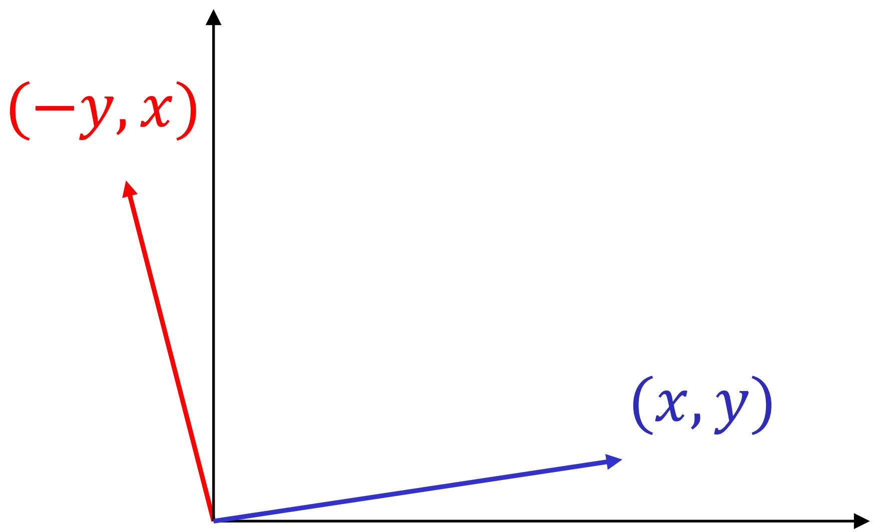
Stretching by the factor *c*

- $A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$ stretches every vector x.



Rotation by 90^o

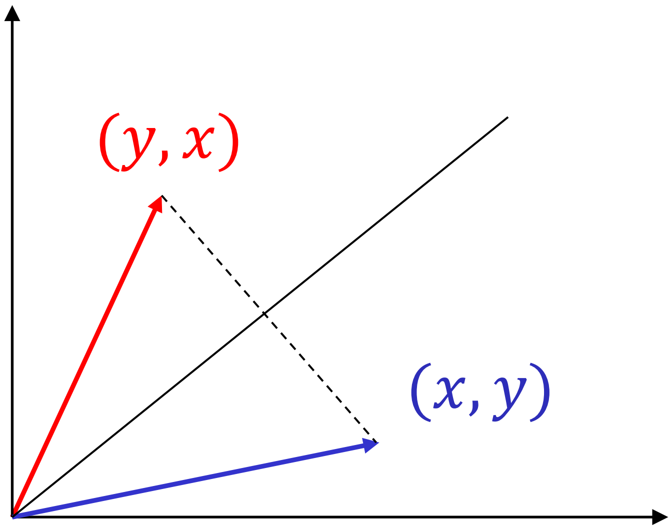
- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ turns every point (x, y) to $(-y, x)$.



Examples

Reflection across 45^o line

- $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ flips every vector x to its mirror image.



Projection onto x-axis

- $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ projects every point (x, y) to the x-axis.

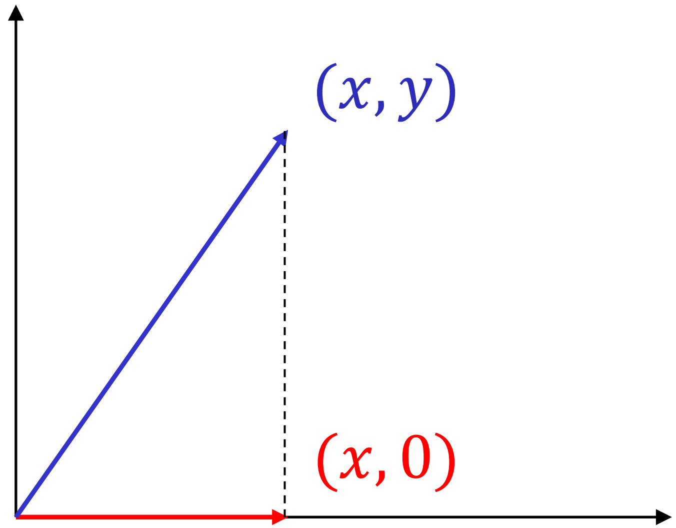
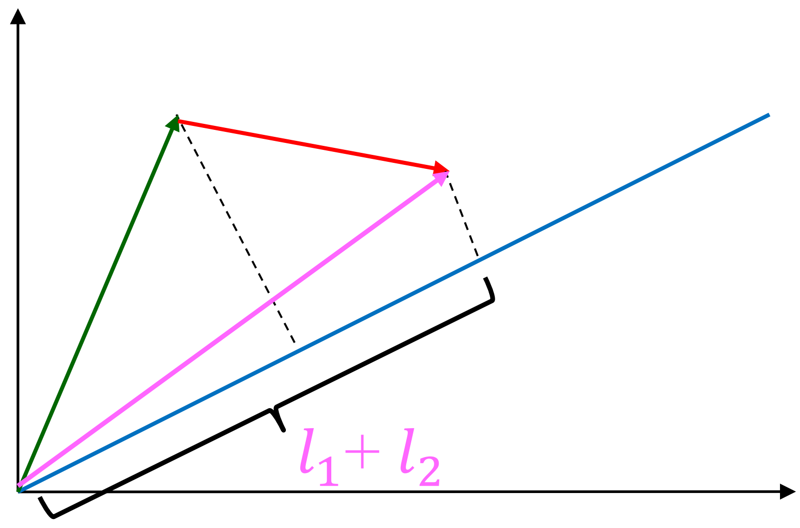
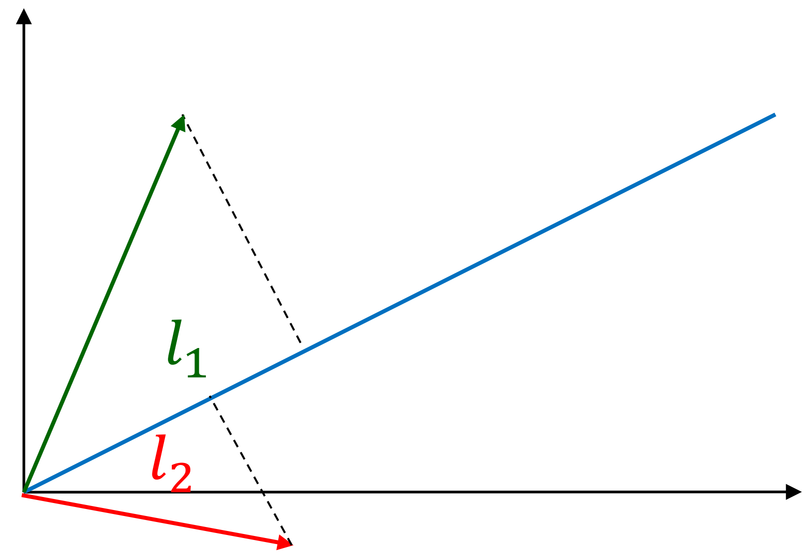


Illustration: Projection is Linear





Add two vectors and then project onto the line. $\equiv$ Project two vectors onto the line and then add them.

Linear Mapping via Standard Basis

- Every vector $x \in \mathbb{R}^n$ can be written as
$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$
- By superposition, if we know $v_i = Ae_i$ for all i, then we can determine
$$\begin{aligned} Ax &= x_1 Ae_1 + x_2 Ae_2 + \cdots + x_n Ae_n \\ &= x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \\ &= \begin{bmatrix} | & | & \cdots & | \\ v_1 & v_2 & \ddots & v_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{aligned}$$
- Note: v_i is obtained by transforming e_i.

Example

- Let e_1 and e_2 be the standard basis of $\mathbb{R}^2$.
- Consider a linear function $f: \mathbb{R}^2 \to \mathbb{R}^3$.
- Suppose we know that

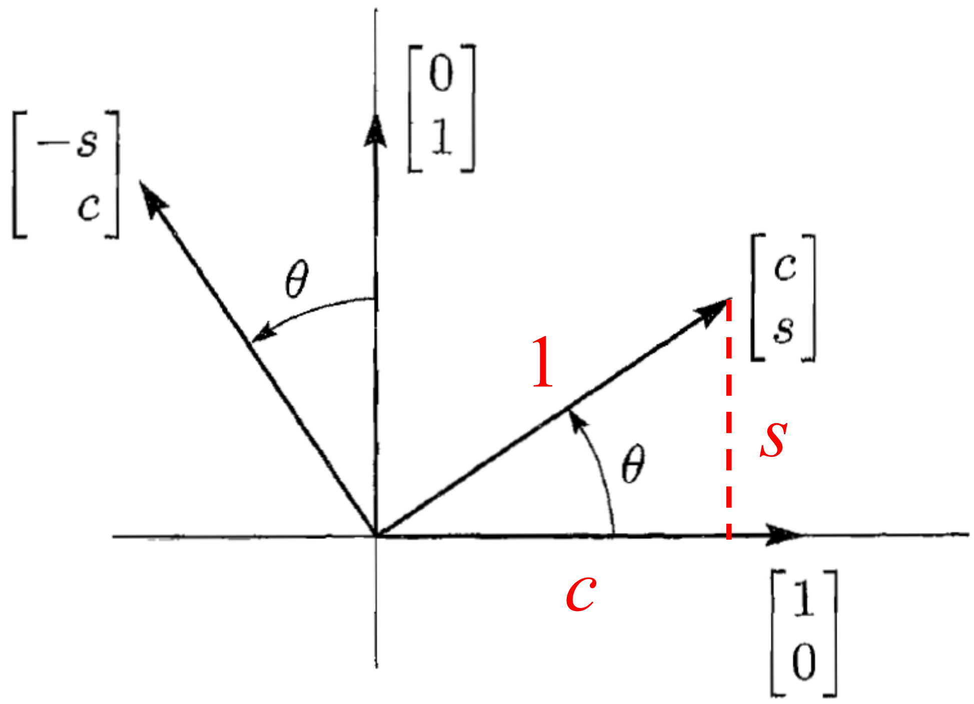
$$f(e_1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ and } f(e_2) = \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix}.$$

- A vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ can be written as $x = x_1 e_1 + x_2 e_2$.
- Hence, $f(x) = x_1 f(e_1) + x_2 f(e_2) = \begin{bmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Ax$,

  where $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 4 \end{bmatrix}$.

Rotation through θ

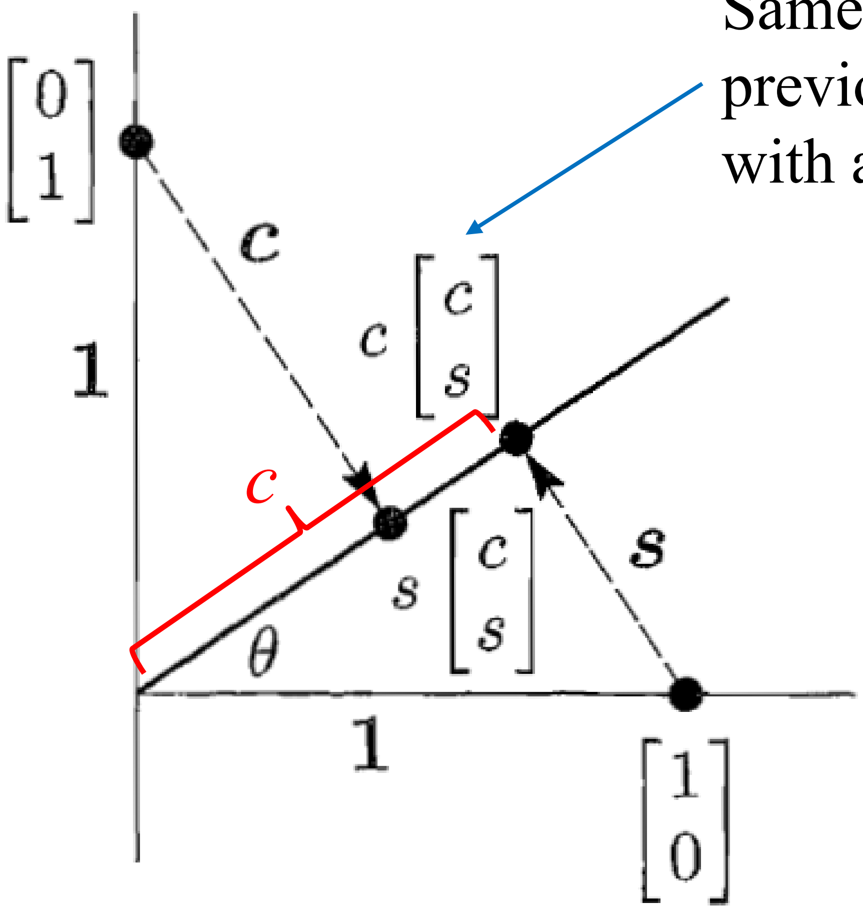
- Let $c = \cos\theta$ and $s = \sin\theta$.



Hence, the rotation matrix is given by

$$R = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

Projection onto the θ-line



Same as in a previous slide, except with a factor of c.

Hence, the projection matrix is given by

$$P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$$

Unit 8.4

Vector Spaces

What is a Field?

- Roughly speaking, a field is a set of elements which you can add, subtract, multiply, and divide.
- Examples:
  - Rational numbers $\mathbb{Q}$
  - Real numbers $\mathbb{R}$
  - Complex numbers $\mathbb{C}$
- Non-examples:
  - Integers $\mathbb{Z}$

Field Definition (optional)

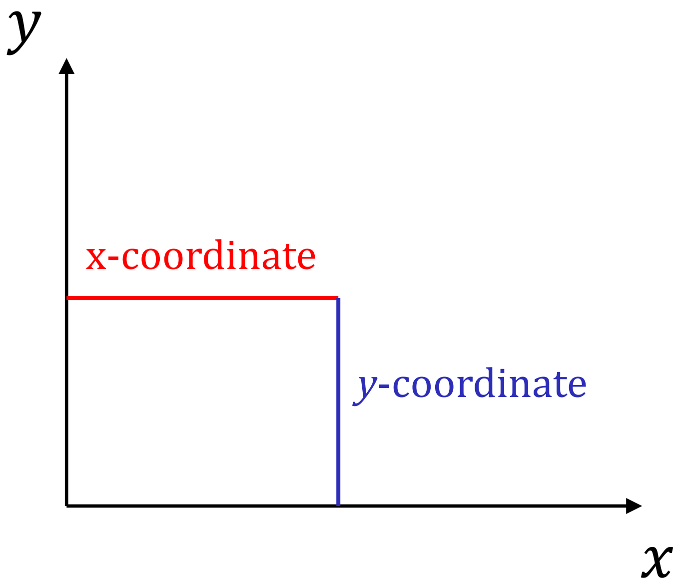
- A set F of elements
- Two operations often denoted by $+$, $\times$
- F forms a commutative group under $+$
  - Additive identity is denoted by 0.
- $F \backslash \{0\}$ forms a commutative group under $\times$
  - Multiplicative identity is denoted by 1.
  - Zero is excluded because division by 0 is not allowed, meaning that multiplicative inverse of 0 does not exist.
- Distributive Property:
  - $a \times (b + c) = a \times b + a \times c$
  - $(b + c) \times a = b \times a + c \times a$

We will study commutative group in Unit 10.

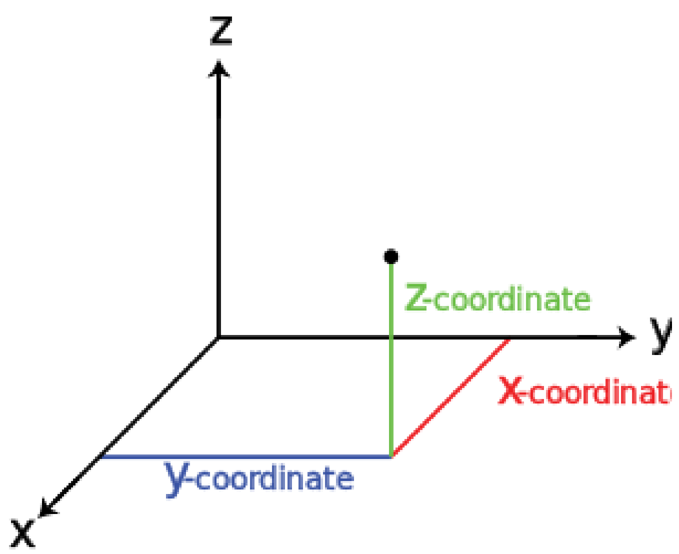
Vector Spaces: Examples

A vector is an element of a vector space.

a set with some special properties



2-dimensional
Euclidean space, $\mathbb{R}^2$



3-dimensional
Euclidean space, $\mathbb{R}^3$

What is a Vector Space?

- A vector space is
  - a set of elements, and
  - two operations within the space.

- The two operations are
  i. (Vector Addition) Adding two vectors, and
  ii. (Scalar Multiplication) Multiplying a vector by a scalar.
  - These operations need to satisfy eight properties to be defined in the next slide.

Eight Properties

No need to memorize them.

1. (Commutative) $x + y = y + x$.
2. (Associative) $x + (y + z) = (x + y) + z$.
3. (Zero) There exists an element $\mathbf{0}$, called zero vector, such that $x + \mathbf{0} = x$ for all x.
4. (Inverse) For each x, there exists a unique vector $-x$ such that $x + (-x) = \mathbf{0}$.
5. (Associative) $(c_1 c_2)x = c_1(c_2 x)$.
6. (Unitarity) $1x = x$.
7. (Distributive I) $c(x + y) = cx + cy$.
8. (Distributive II) $(c_1 + c_2)x = c_1 x + c_2 x$.

Example 1: Matrices

- Consider the set of all 2×2 matrices with real entries, denoted by $\mathbb{R}^{2\times 2}$.
- The two operations are defined as follows:
  - Addition: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a+w & b+x \\ c+y & d+z \end{bmatrix}$.
  - Scalar Multiplication: $\alpha \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{bmatrix}$.
- Do these two operations satisfy the eight conditions?
  - Yes. (The checking is tedious, thus omitted).
  - Details can be found here: https://www.youtube.com/watch?v=ug3FpapN8Ng (start from 7:38)

Example 2: Real Functions

- Consider the set of all functions $f: \mathbb{R} \to \mathbb{R}$.
- The two operations are defined as follows:
  - Addition: $(f + g)(x) = f(x) + g(x)$.
  - Scalar Multiplication: $(\alpha f)(x) = \alpha f(x)$.
- Next, check the eight conditions.
- Yes, they are satisfied.
- Zero element is the constant function $\mathbf{0}(x) = 0$.

Example 3: Polynomials

- A real polynomial p is of the form
$$p = a_0 + a_1 x + \cdots + a_n x^n,$$
where the coefficients are real numbers.

- The two operations are defined in the usual way.
- It can be checked that the set of all polynomials is a vector space.

Subspace

- A vector space is a set with two operations that satisfy eight conditions.
- Its subset is called a subspace if the subset is also a vector space.
- We only need to check:

> Is it closed under addition and scalar multiplication?

"Closed" means that the result remains in the subset.

- The eight conditions will automatically be satisfied, since it is a subset of a vector space.

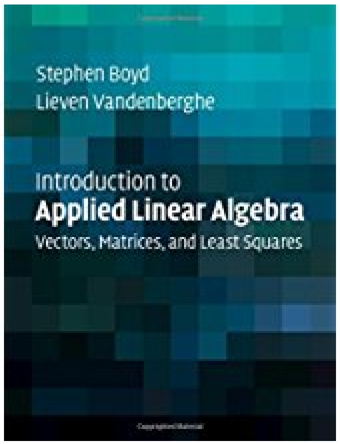
Example 4: The x-y plane in $\mathbb{R}^3$

- Is the x-y plane a subspace of $\mathbb{R}^3$?
- The x-y plane consists of all vectors in the form of $(x, y, 0)$.
- Closed under addition:
  - $(x_1, y_1, 0) + (x_2, y_2, 0) = (x_1 + x_2, y_1 + y_2, 0)$ is still in the x-y plane.
- Closed under scalar multiplication:
  - $c(x, y, 0) = (cx, cy, 0)$ is still in the x-y plane.
- Therefore, it is a subspace.

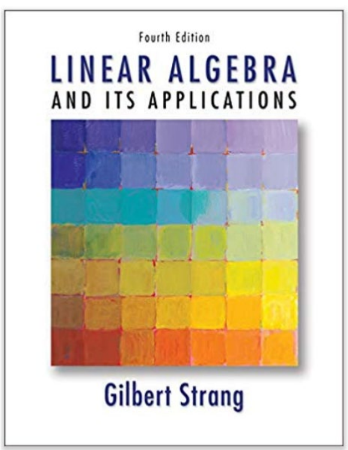
Example 5: A non-example

- What if we lift the x-y plane by 1 unit along the z-axis? Is it a subspace of $\mathbb{R}^3$?
- No.
  - *Not* closed under addition:
    $$(x_1, y_1, 1) + (x_2, y_2, 1) = (x_1 + x_2, y_1 + y_2, 2)$$
  - *Not* closed under scalar multiplication if $c \neq 1$:
    $$c(x, y, 1) = (cx, cy, c)$$
- Recall that condition 3 says that there must be a zero element. In this case, the zero vector is not in the lifted x-y plane, so it must not be a vector space.

Recommended Reading



- Chapters 1 and 2, S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*, Cambridge University Press, 2018.
  - Available on the web, http://web.stanford.edu/~boyd/vmls/



- Sections 2.6 and 3.2, G. Strang, *Linear Algebra and its Applications*, 4^{th} ed., Thomson Learning, 2006.
  - This book is more advanced.