Vectors

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

What is vector?

- A vector is an ordered finite list of numbers. Written as:

$$a, X, p, \beta, E^{\text{aut}}, \mathbf{g}, \vec{a}$$

$$\begin{bmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{bmatrix} \quad \begin{pmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{pmatrix} \quad (-1.1, 0.0, 3.6, -7.2)$$

- Size (dimension or length): A vector of size n is called an n-vector ($x \in \mathcal{R}^n$)
- Elements (entries, coefficients, components) of a vector
- Two vectors a and b are equal, which we denote a = b, if they have the same size, and each of the corresponding entries is the same. If a and b are n-vectors, then a = b means a1 = b1, . . . , an = bn.
- Numbers are called scalars
- The set of all n-vectors is denoted $\mathbb{R}^n := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \middle| \, a_1, \ldots, a_n \in \mathbb{R} \right\}$

Block vectors

- Suppose b, c, and d are vectors with sizes m, n, p
- *stacked vector* or concatenation of b, c, and d. block vector with entries (blocks) b, c, d is:

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

- a has size m + n + p:
  - $a = (b_1, b_2, \ldots, b_m, c_1, c_2, \ldots, c_n, d_1, d_2, \ldots, d_p)$

Subvector

- $a_{r:s} = (a_r, \ldots, a_s)$ is a subvector of a. It is a vector with size (s-r+1).
- Colon notation is used to denote subvectors.
- The subscript r:s is called the index range
- In a block vector $a = \begin{bmatrix} b \\ c \\ d \end{bmatrix}$
  - b, c, and d are subvectors or slices of a, with sizes m, n, and p, respectively.
  - $b = a_{1:m}, \qquad c = a_{(m+1):(m+n)}, \qquad d = a_{(m+n+1):(m+n+p)}$

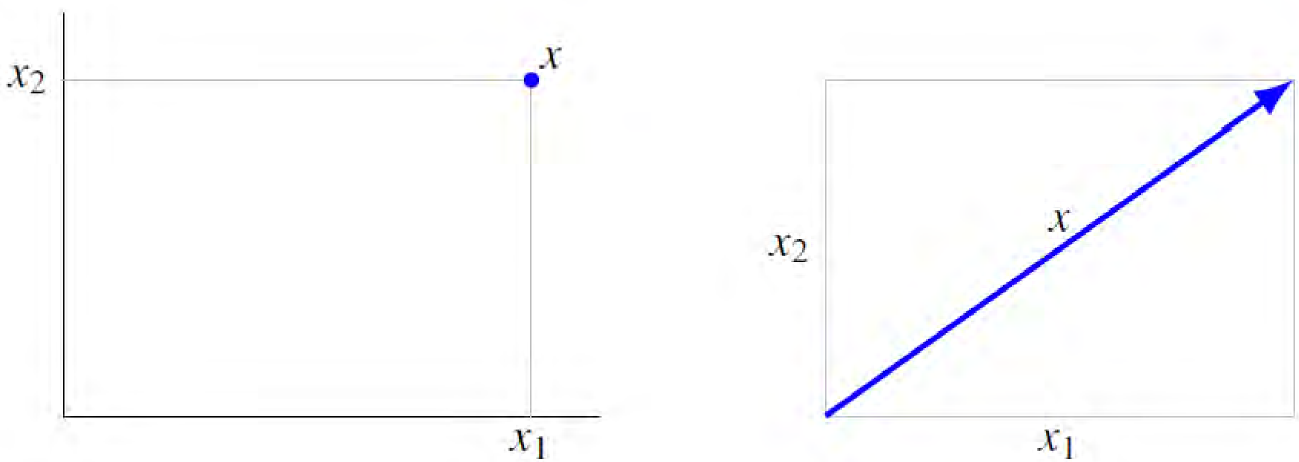
Famous vectors

- Zero vector: O_n
- Ones vector: I_n
- Unit vector: e_i (e_i is the entry with 1 value)
- Question: Write all unit vectors with length of 3?
- Sparse vector: a vector if many of its entries are 0
  - can be stored and manipulated efficiently on a computer
  - **nnz(x)** is number of entries that are nonzero
  - Question: What is the most sparsest vector?

Vectors examples

- Location or displacement in 2-D or 3-D

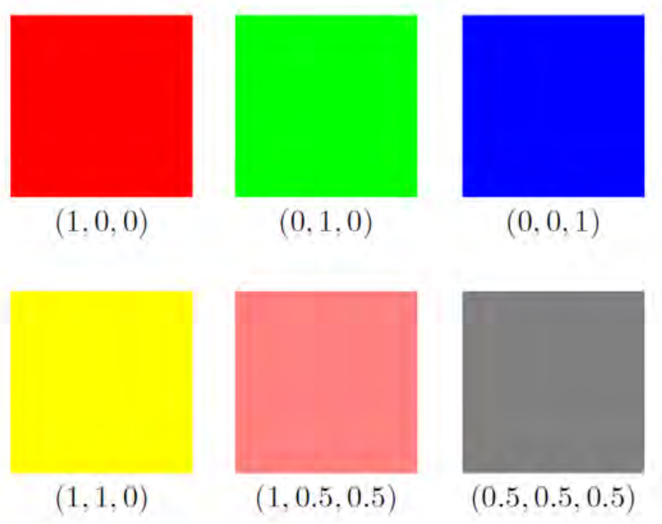
2-vector (x_1, x_2) can represent a location or a displacement in 2-D



- A vector can also be used to represent a displacement in a plane or 3-D space, in which case it is typically drawn as an arrow.
- A vector can also be used to represent the velocity or acceleration, at a given time, of a point that moves in a plane or 3-D space.

Vectors examples

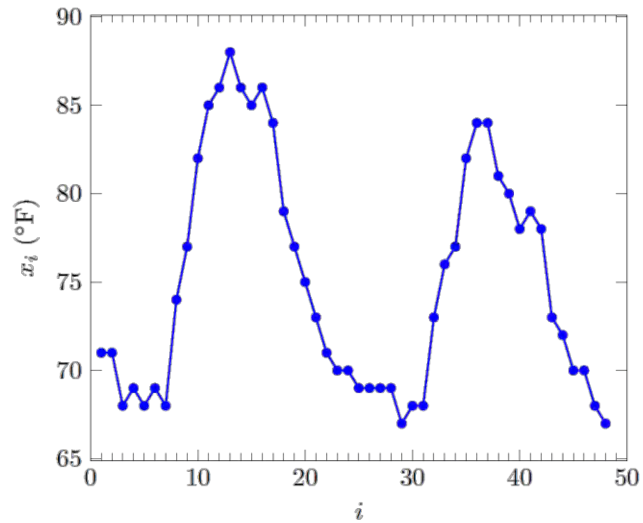
- Color (RGB)
  - A 3-vector can represent a color, with its entries giving the Red, Green, and Blue (RGB) intensity values (often between 0 and 1).



Six colors and their RGB vectors.

Vectors examples

- Time series
  - An n-vector can represent a time series or signal, that is, the value of some quantity at different times.
  - The entries in a vector that represents a time series are sometimes called samples, especially when the quantity is something measured.
  - An audio (sound) signal can be represented as a vector whose entries
  - give the value of acoustic pressure at equally spaced times (typically 48000 or 44100 per second).
  - A vector might give the hourly rainfall (or temperature, or barometric pressure) at some location, over some time period.
  - These lines carry no information; they are added only to make the plot
  - easier to understand visually.



Hourly temperature in downtown Los Angeles on August 5 and 6, 2015 (starting at 12:47AM, ending at 11:47PM).

Vectors examples

Word count vectors

- a short document:

  > **Word** count vectors are used **in** computer based **document** analysis. Each entry of the **word** count vector is the **number** of times the associated dictionary **word** appears **in** the **document**.

- a small dictionary (left) and word count vector (right)

$$
\begin{array}{l}
\text{word} \\
\text{in} \\
\text{number} \\
\text{horse} \\
\text{the} \\
\text{document}
\end{array}
\qquad\qquad
\begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 4 \\ 2 \end{bmatrix}
$$

- dictionaries used in practice are much larger

Basic Notation

- Column vector $x \in R^n$
- Transpose:

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

$$\begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 4 & 3 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 3 & 0 \end{bmatrix}^{\mathrm{TT}} = \begin{bmatrix} 4 & 3 & 0 \end{bmatrix}$$

$$4^{\mathrm{T}} = 4,$$

- Row vector $x^T \in R^{1\times n}$
- ith element of x is: x_i

Vector Addition

- n-vectors a and b

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{ and } b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \qquad a + b = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

- Can be added, with sum denoted: a + b
- Subtraction is similar: (a-b)
- The result of vector subtraction is called the difference of the two vectors.

Vector Addition and Subtraction

The Head-to-Tail Rule

Given vectors **u** and **v** in $\mathbb{R}^2$, translate **v** so that its tail coincides with the head of **u**. The *sum* **u** + **v** of **u** and **v** is the vector from the tail of **u** to the head of **v**. (See Figure 1.7.)

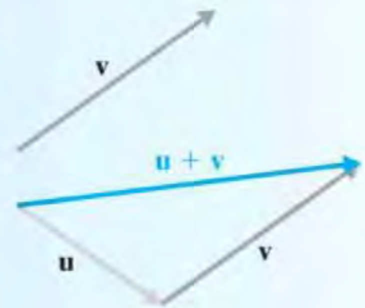
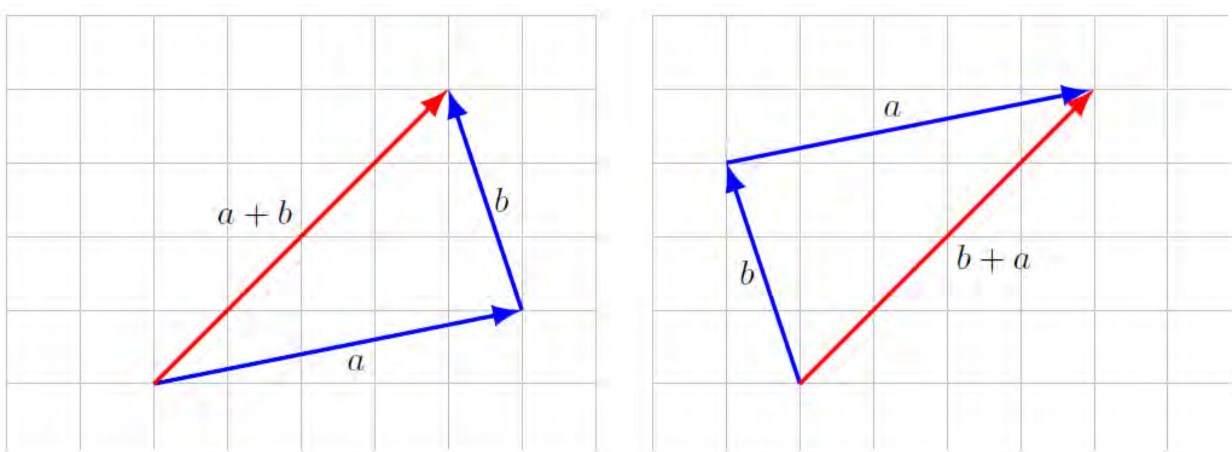


Figure 1.7
The head-to-tail rule



Vector Addition and Subtraction

The Parallelogram Rule

Given vectors **u** and **v** in $\mathbb{R}^2$ (in standard position), their ***sum*** $\mathbf{u} + \mathbf{v}$ is the vector in standard position along the diagonal of the parallelogram determined by **u** and **v**. (See Figure 1.9.)

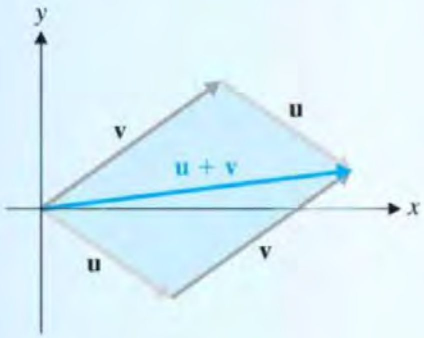
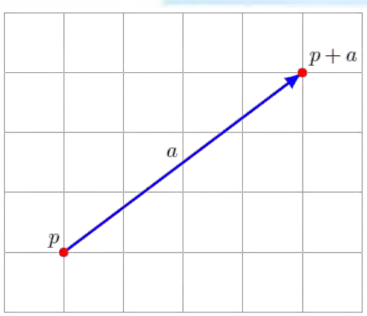
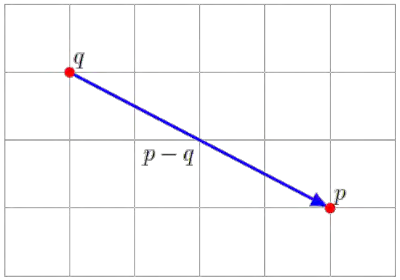


Figure 1.9
The parallelogram rule



The vector $p+a$ is the position of the point represented by p displaced by the displacement represented by a.



The vector $p-q$ represents the displacement from the point represented by q to the point represented by p.

Vector Addition Properties

- Commutative $a + b = b + a$
- Associative
  - Note: the associative law is that parentheses can be moved around, e.g., (x+y)+z = x+(y+z) and x(yz) = (xy)z

$$(a + b) + c = a + (b + c) = a + b + c$$

- Adding the zero vector to a vector has no effect

$$a + 0 = 0 + a = a$$

  - What constraints should you have?
- Subtracting a vector from itself yields the zero vector

$$a - a = 0$$

  - What is size of 0 here?

Vector Addition Properties

- Transpose: For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$, $(\boldsymbol{u}+\boldsymbol{v})^T = \boldsymbol{u}^T + \boldsymbol{v}^T$
  - Proof?

- Can scalar and vector be added?

$$4 + \begin{bmatrix} 1 \\ 2 \\ -10 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \\ -10 \end{bmatrix} + 4$$

Scalar-Vector Product

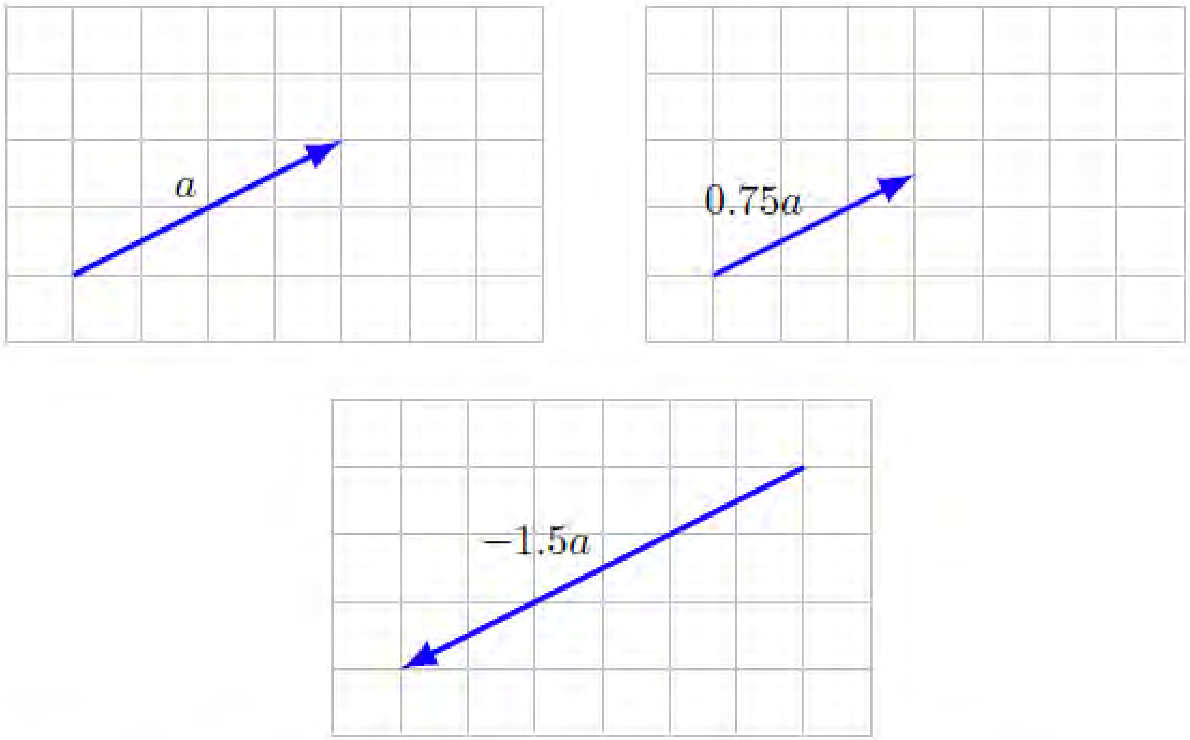
- Scalar multiplication or scalar-vector multiplication: a vector is multiplied by a scalar (i.e., number), which is done by multiplying every element of the vector by the scalar.
  - scalar on the left or scalar on the right

$$(-2)\begin{bmatrix} 1 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ -18 \\ -12 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 9 \\ 6 \end{bmatrix}(1.5) = \begin{bmatrix} 1.5 \\ 13.5 \\ 9 \end{bmatrix}$$

  - Some notations:
    - $\boldsymbol{a}/2$ is a vector means $\left(\frac{1}{2}\right)\boldsymbol{a}$
    - $-\boldsymbol{a}$ is a vector means $(-1)\boldsymbol{a}$
    - $0\boldsymbol{a} = \boldsymbol{0}$ vector

scalar

Scalar-Vector Product



The vector $0.75a$ represents the displacement in the direction of the displacement a, with magnitude scaled by 0.75; $(-1.5)a$ represents the displacement in the opposite direction, with magnitude scaled by 1.5.

Scalar-Vector Product Properties

- Commutative $\beta \boldsymbol{a} = \boldsymbol{a}\beta$
- Associative

$$(\beta\gamma)\boldsymbol{a} = \beta(\gamma\boldsymbol{a}) = (\beta\boldsymbol{a})\gamma = \beta\boldsymbol{a}\gamma = \beta\gamma\boldsymbol{a}$$

- Left-Distributive

$$(\beta + \gamma)\boldsymbol{a} = \beta\boldsymbol{a} + \gamma\boldsymbol{a}$$

- Right-Distributive

$$\boldsymbol{a}(\beta + \gamma) = \boldsymbol{a}\beta + \boldsymbol{a}\gamma$$

$$\beta(\boldsymbol{a} + \boldsymbol{b}) = \beta\boldsymbol{a} + \beta\boldsymbol{b}$$

Addition of n-vectors

Vector-Vector Products

- Given two vectors $x, y \in R^n$: (should have same size)
  - $\boldsymbol{x.y}$ is called the inner product or dot product or scalar product of the vectors: $x^T y$ $(y^T x)$

$$x^T y \in \mathbb{R} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i .$$

- Dot product is a single number that provides information about the relationship between two vectors
- It is the basic computational building-block from which many operations and algorithms are built, including convolution, correlation, the Fourier transform, matrix multiplication, signal filtering, and so on.
- The term "inner product" is used when the two vectors are continuous functions.
- Why is named scalar product, too?
- Notations: $< a, b >$ $\quad < a|b >$ $\quad (a, b)$ $\quad a.b$

Vector-Vector Products

- Dot product between a vector and itself: magnitude-squared, the length squared, or the squared-norm, of the vector.

$$\mathbf{a}^{\mathrm{T}}\mathbf{a} = \|\mathbf{a}\|^2 = \sum_{i=1}^{n} \mathrm{a}_i \mathrm{a}_i = \sum_{i=1}^{n} \mathrm{a}_i^2$$

- If the vector is mean-centered—the average of all vector elements is subtracted from each element—then the dot product of a vector with itself is call *variance* in statistics lingo.
- When n = 1, the inner product reduces to the usual product of two numbers.

Vector-Vector Products

- The scalar product can be viewed as function taking two vectors as arguments and producing a single scalar as a result. The usual notation in this case is

$$\langle\, ,\rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R},\ \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^T \boldsymbol{v} = \sum_{i=1}^{m} u_i v_i$$

  with $\mathcal{V} = \mathbb{R}^m$.

- Transpose of dot product:
  - $(a.b)^T = \left(a^T b\right)^T = \left(b^T a\right) = (b.a) = b^T a$

Dot product properties

- Commutativity
  - The order of the two vector arguments in the inner product does not matter.

$$a^T b = b^T a$$

- Distributivity with vector addition
  - The inner product can be distributed across vector addition.

$$(a+b)^T c = a^T c + b^T c$$
$$a^T (b+c) = a^T b + a^T c$$

Dot product properties

- Bilinear (linear in both a and b)

$$a^T(\lambda b + \beta c) = \lambda a^T b + \beta a^T c$$

- Positive Definite:

$$(a.a) = a^T a \geq 0$$

  - 0 only if a itself is a zero vector $a = \mathbf{0}$

Dot product properties

- Associative
  - Note: the associative law is that parentheses can be moved around, e.g., (x+y)+z = x+(y+z) and x(yz) = (xy)z
- 1) Associative property of the vector dot product with a scalar (scalar-vector multiplication embedded inside the dot product)

scalar

$$\gamma(\mathbf{u}^{\mathrm{T}}\mathbf{v}) = (\gamma\mathbf{u}^{\mathrm{T}})\mathbf{v} = \mathbf{u}^{\mathrm{T}}(\gamma\mathbf{v}) = (\mathbf{u}^{\mathrm{T}}\mathbf{v})\gamma$$

$$= (\gamma \boldsymbol{u})^{T}\boldsymbol{v} = \gamma \boldsymbol{u}^{T}\boldsymbol{v}$$

Dot product properties

- Associative
  - 2) Does vector dot product obey the associative property?

$$\underbrace{\mathbf{u}^{\mathrm{T}}(\mathbf{v}^{\mathrm{T}}\mathbf{w})}_{\substack{\text{vector-scalar product} \\ \text{row vector}}} = \underbrace{(\mathbf{u}^{\mathrm{T}}\mathbf{v})^{\mathrm{T}}\mathbf{w}}_{\substack{\text{scalar-vector product} \\ \text{column vector}}}$$

General Examples

- The inner product of a vector with the ith standard unit vector gives (or `picks out') the ith element of a.

$$e_i^T a = a_i$$

- The inner product of a vector with the vector of ones gives the sum of the elements of the vector.

$$\mathbf{1}^T a = a_1 + \cdots + a_n$$

- The inner product of an n-vector with the vector $\mathbf{1}/n$ gives the average or mean of the elements of the vector.

$$\boldsymbol{aveg}(a) = \boldsymbol{\mu}_a = (\mathbf{1}/n)^T a = (a_1 + \cdots + a_n)/n$$

General Examples

- The inner product of a vector with itself gives the sum of the squares of the elements of the vector.

$$a^T a = a_1^2 + \cdots + a_n^2$$

- Selective sum: Let b be a vector all of whose entries are either 0 or 1. Then $b^T a$ is the sum of the elements in a for which $b_i = 1$.

Inner product of block vectors

- If two block vectors conform, then the inner product of them is the sum of inner products of the blocks:
  - Proof?

Dot product properties

- Example
  - For any vectors a, b, c, d with the same size:

$$(a+b)^T(c+d) = a^Tc + a^Td + b^Tc + b^Td$$

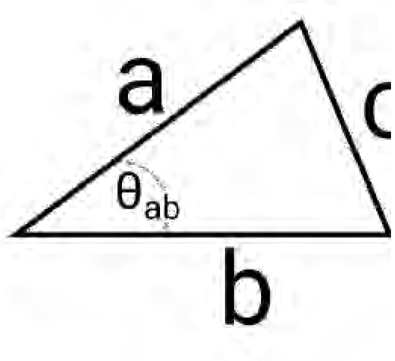
  - Specify the vector and scalar additions?
  - Applying the distributive property to the dot product between a vector and itself?

$$(\mathbf{u}+\mathbf{v})^{\mathrm{T}}(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}+\mathbf{v}\|^2 = \mathbf{u}^{\mathrm{T}}\mathbf{u} + 2\mathbf{u}^{\mathrm{T}}\mathbf{v} + \mathbf{v}^{\mathrm{T}}\mathbf{v}$$

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

Vector dot product: Geometry

- Dot Product: the cosine of the angle between the two vectors, times the lengths of the two vectors.

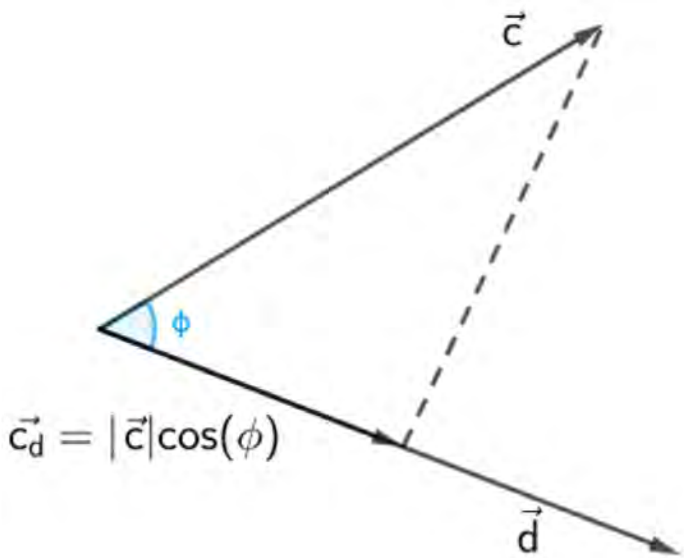


$$a^T b = \|a\| \, \|b\| \cos(\theta_{ab})$$

  - proof
- In statistics, cos() with suitable normalization is called the Pearson correlation coefficient.

Vector dot product: Geometry

- This is called scalar projection. To find the vector projection of vector c on vector d we have to multiply scalar projection with unit vector d.



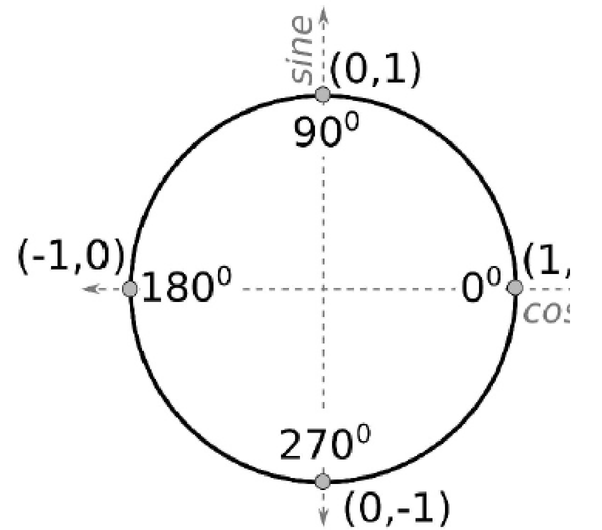
$$\vec{c}_d = |\vec{c}|\cos(\phi)\cdot\frac{\vec{d}}{|\vec{d}|}$$

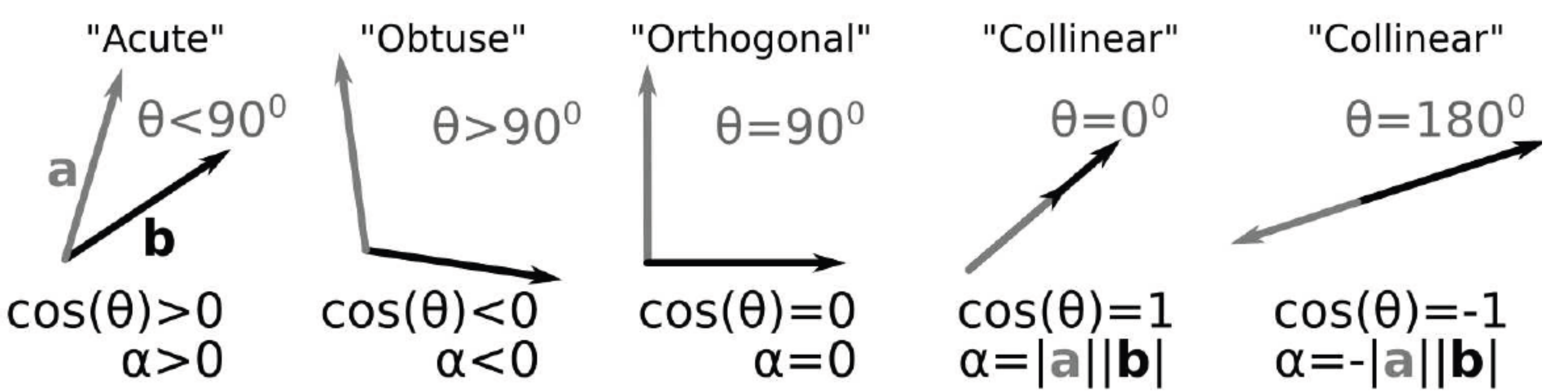
$$\vec{c}\cdot\vec{d} = \vec{c}_d|\vec{d}| = \vec{d}_c|\vec{c}|$$

- Projections have wide use in linear algebra and machine learning (Support Vector Machine(SVM) is a machine learning algorithm, used for classification of data).

Vector dot product: Geometry

- $\theta < 90^o$
- $\theta > 90^o$
- $\theta = 90^o$: vectors are orthogonal
- $\theta = 0^o$: collinear
- $\theta = 180^o$: collinear





Vector-Vector Products

- Given two vectors $x \in R^m, y \in R^n$:
  - $x \otimes y = xy^T \in R^{m \times n}$ is called the outer product of the vectors: $(xy^T)_{ij} = x_i y_j$

$$xy^T \in \mathbb{R}^{m \times n} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_my_1 & x_my_2 & \cdots & x_my_n \end{bmatrix}$$

  - Is it symmetric?
  - Example: Represent $A \in R^{m \times n}$ with outer product of two vectors:

$$A = \begin{bmatrix} | & | & & | \\ x & x & \cdots & x \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x_1 & x_1 & \cdots & x_1 \\ x_2 & x_2 & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_m & x_m & \cdots & x_m \end{bmatrix}$$

Outer Products

- Properties:
  - $(u \otimes v)^T = (v \otimes \mathrm{u})$
  - $(v + w) \otimes u = v \otimes u + w \otimes u$
  - $u \otimes (v + w) = u \otimes v + u \otimes w$
  - $c(v \otimes u) = (cv) \otimes u = v \otimes (cu)$
  - $(u.v) = trace(u \otimes v) \ \ (u, v \epsilon R^n)$
  - $(u \otimes v)w = (v.w)u$

Hadamard vector product

- Element-wise product

$$c = a \odot b = \begin{bmatrix} a_1 b_1 \\ a_2 b_2 \\ \cdot \\ \cdot \\ a_n b_n \end{bmatrix}$$

- Hadamard product is used in image compression techniques such as JPEG. It is also known as Schur product

- Hadamard Product is used in LSTM (Long Short-Term Memory) cells of Recurrent Neural Networks (RNNs).

Hadamard vector product

- Properties:
  - $a \odot b = b \odot a$
  - $a \odot (b \odot c) = (a \odot b) \odot c$
  - $a \odot (b + c) = a \odot b + a \odot c$
  - $(\theta a) \odot b = a \odot (\theta b) = \theta (a \odot b)$
  - $a \odot \mathbf{0} = \mathbf{0} \odot a = \mathbf{0}$

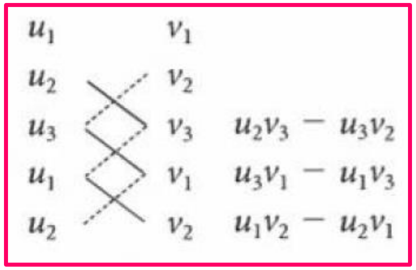
Cross product

- The cross product is defined only for two 3-element vectors, and the result is another 3-element vector. It is commonly indicated using a multiplication symbol (×).

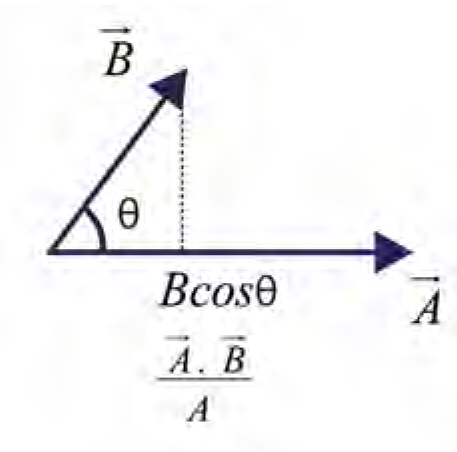
$$\|\mathbf{a}\times\mathbf{b}\| = \|\mathbf{a}\|\,\|\mathbf{b}\|\sin(\theta_{ab})$$

$$\mathbf{a}\times\mathbf{b} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

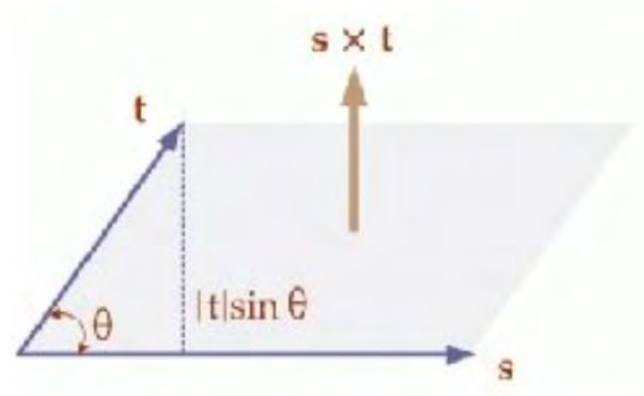
- It used often in geometry, for example to create a vector c that is orthogonal to the plane spanned by vectors a and b. It is also used in vector and multivariate calculus to compute surface integrals.



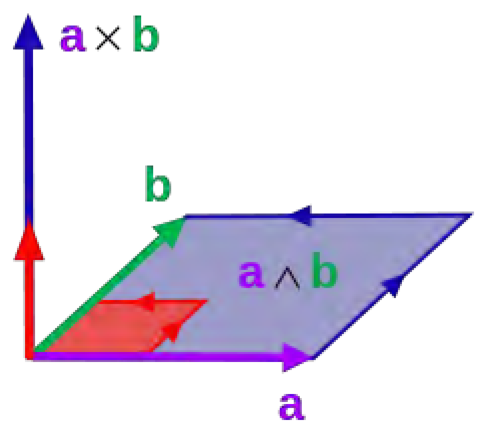
Products

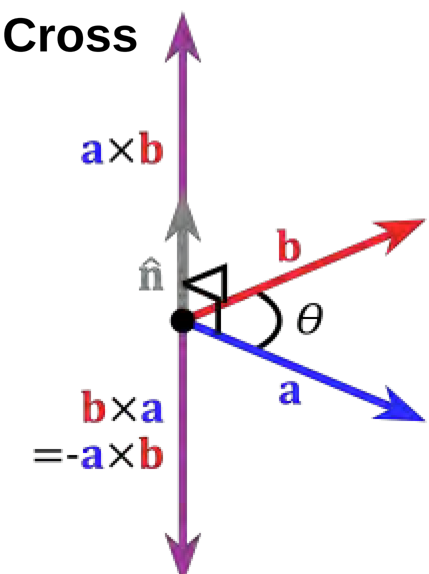


Dot



Cross





Wedge and Cross

Linear Combinations

- The linear combinations of m vectors $a_1, \dots a_m$, each with size n is:

$$\beta_1 a_1 + \cdots + \beta_m a_m$$

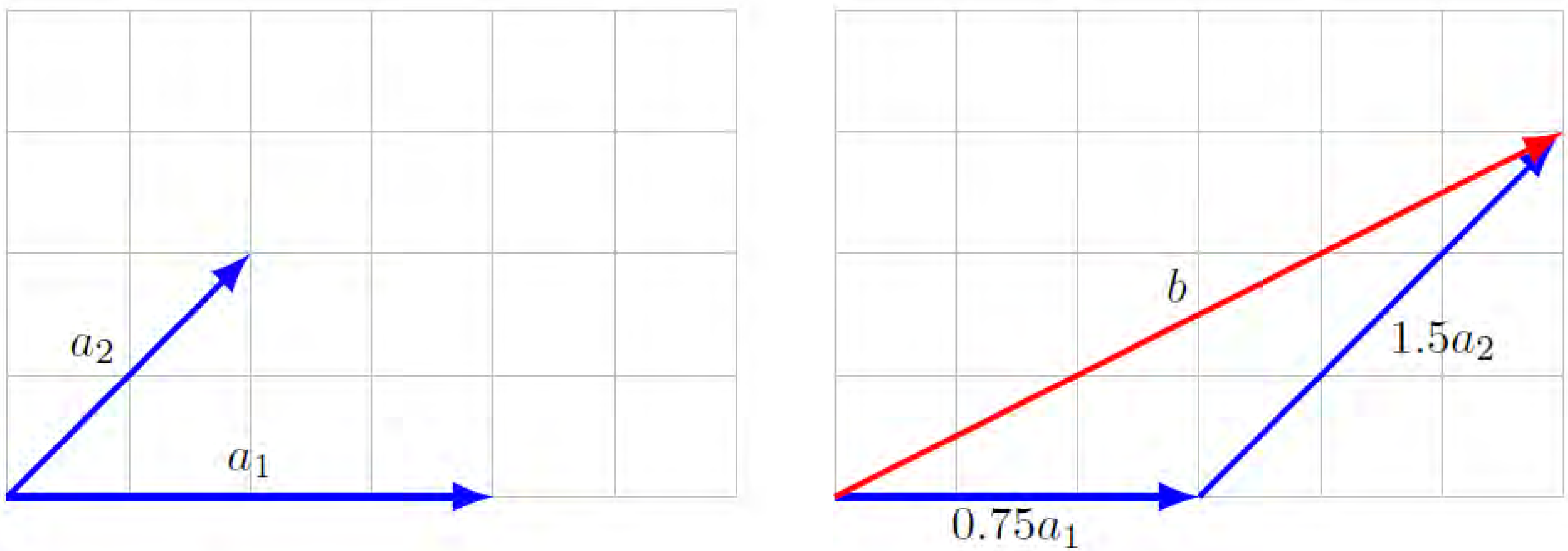
where $\beta_1, \dots, \beta_m$ are scalars and called the coefficients of the linear combination

- Coordinates: We can write any n-vector b as a linear combination of the standard unit vectors, as:

$$b = b_1 e_1 + \cdots + b_n e_n$$

  - Example: What are the coefficients and combination for this vector? $\begin{bmatrix} -1 \\ 3 \\ 5 \end{bmatrix}$

Linear Combinations



Left. Two 2-vectors a_1 and a_2. *Right.* The linear combination $b = 0.75a_1 + 1.5a_2$

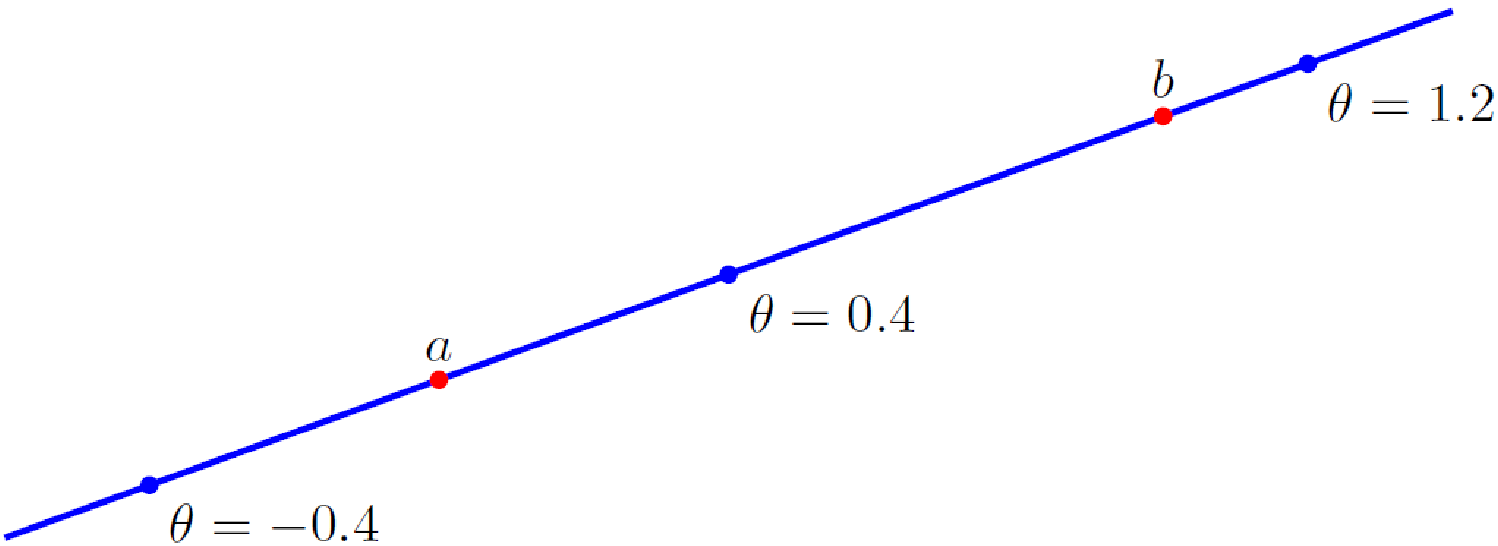
Special Linear Combinations

- Sum of vectors
- Average of vectors
- Affine combination

$$\beta_1 + \cdots + \beta_m = 1$$

- Convex combination, mixture average, weighted average: When the coefficients in an affine combination are nonnegative
  - Note: The coefficients in an affine or convex combination are sometimes given as percentages, which add up to 100%.

Linear Combinations Example



The affine combination $(1-\theta)a + \theta b$ for different values of θ.
These points are on the line passing through a and b; for θ between 0 and 1, the points are on the line segment between a and b.

Linear Combinations

- For vectors $x_1, x_2, \ldots, x_k$:any point y is a linear combination of them iff:

$$y = \alpha_1 x_1 + \alpha_2 x_2 \cdots + \alpha_k x_k \quad \forall i,\ \alpha_i \in \mathbb{R}$$

- If we restrict α_i's to be positive then we get a conic combination.

$$y = \alpha_1 x_1 + \alpha_2 x_2 \cdots + \alpha_k x_k \quad \forall i,\ \alpha_i \geq 0 \in \mathbb{R}$$

- Instead of being positive, if we put the restriction that α_i's sum up to 1, it is called an affine combination

$$y = \alpha_1 x_1 + \alpha_2 x_2 \cdots + \alpha_k x_k \quad \forall i,\ \alpha_i \in \mathbb{R}, \quad \sum \alpha_i = 1$$

- When a combination is affine as well as conic, it is called a convex combination

$$y = \alpha_1 x_1 + \alpha_2 x_2 \cdots + \alpha_k x_k \quad \forall i,\ \alpha_i \geq 0 \in \mathbb{R}, \quad \sum_i \alpha_i = 1$$

Complexity of vector computations

- Computers store (real) numbers in floating-point format
- Floating point= 64 bits or 8 bytes
  - How many possible sequences of bits?
  - How many bytes to store n-vector?
- Current memory and storage devices, with capacities measured in many gigabytes (109 bytes), can easily store vectors with dimensions in the millions or billions.
- Sparse vectors are stored in a more efficient way that keeps track of indices and values of the nonzero entries.
- Note about floating point operations and round-off error.

Complexity of vector computations

- How quickly the vector operations can be carried out by a computer depends very much on the computer hardware and software, and the size of the vector.
- Basic arithmetic operations (addition, multiplication, . . .) are called Floating Point Operations (FLOP)s.
- Estimate the time of computation= counting the total number of Floating Point Operations (FLOP)s.
- The complexity of an operation is the number of flops required to carry it out, as a function of the size or sizes of the input to the operation.
- Crude approximation of time to execute: (flopsneeded)/(computer speed)
- current computers are around 1Gflop/sec (10^9 flops/sec)

Complexity of vector computations

Floating point operation

Floating point operation (flop)

- the unit of complexity when comparing vector and matrix algorithms
- 1 flop = one basic arithmetic operation ($+, -, *, /, \surd, \ldots$) in $\mathbf{R}$ or $\mathbf{C}$

Comments: this is a very simplified model of complexity of algorithms

- we don't distinguish between the different types of arithmetic operations
- we don't distinguish between real and complex arithmetic
- we ignore integer operations (indexing, loop counters, $\ldots$)
- we ignore cost of memory access

Complexity of vector computations

Complexity

Operation count (flop count)

- total number of operations in an algorithm
- in linear algebra, typically a polynomial of the dimensions in the problem
- a crude predictor of run time of the algorithm:

$$\text{run time} \approx \frac{\text{number of operations (flops)}}{\text{computer speed (flops per second)}}$$

Dominant term: the highest-order term in the flop count

$$\frac{1}{3}n^3 + 100n^2 + 10n + 5 \approx \frac{1}{3}n^3$$

Order: the power in the dominant term

$$\frac{1}{3}n^3 + 10n^2 + 100 = \text{order } n^3$$

Complexity of vector computations

Examples

complexity of vector operations in this lecture (for vectors of size n)

- addition, subtraction: n flops
- scalar multiplication: n flops
- componentwise multiplication: n flops
- inner product: $2n - 1 \approx 2n$ flops

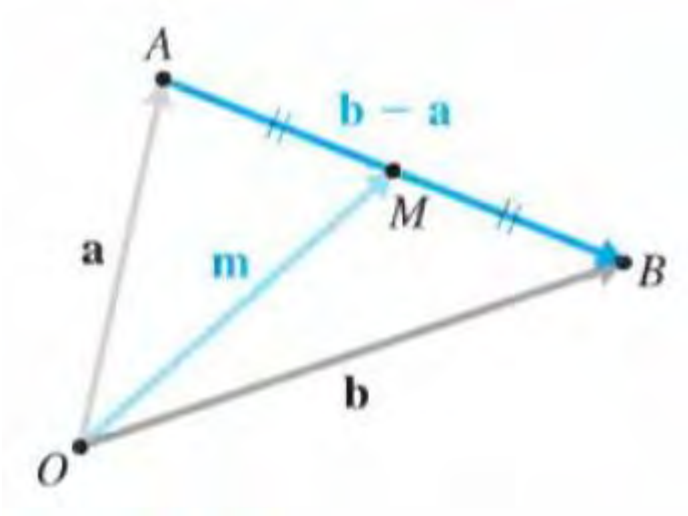
these operations are all order n

Complexity of vector computations

		#FLOPS		Complexity	
Operation		General	Sparse	General	Sparse
Scalar-Vector product					
Vector-Vector sum					
Inner product					
Outer product (vectors with sizes “n” and “m”					
Hadamard product					

Vectors and Geometry

- Give a vector description of the midpoint M of a line segment $\overline{AB}$.

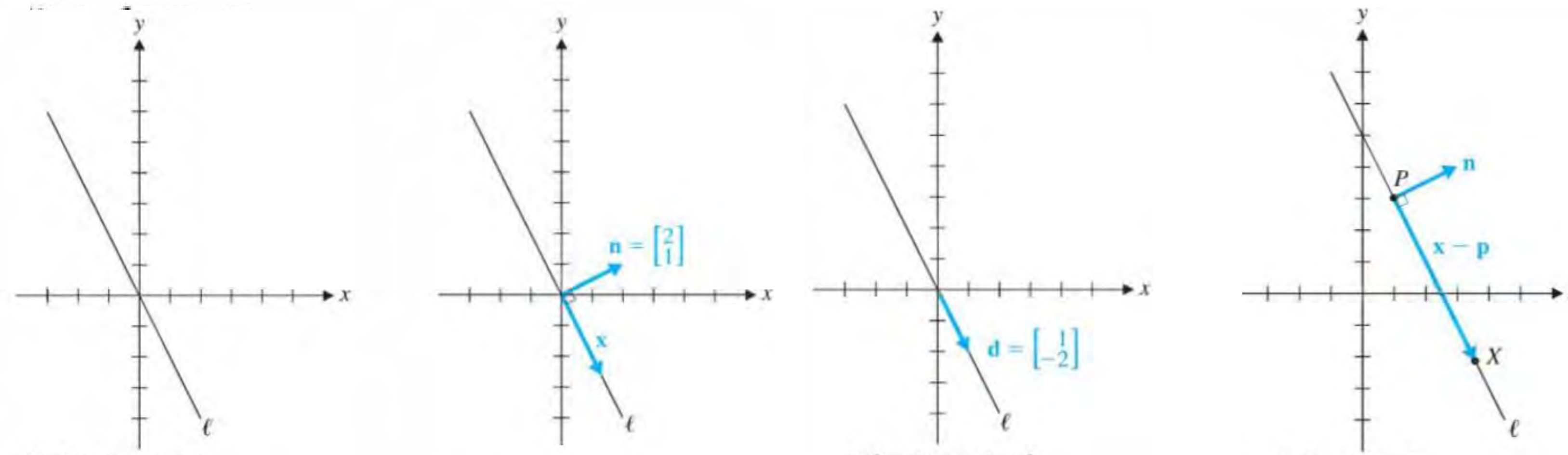


$$\mathbf{m} - \mathbf{a} = \overrightarrow{AM} = \tfrac{1}{2}\overrightarrow{AB} = \tfrac{1}{2}(\mathbf{b} - \mathbf{a})$$

$$\mathbf{m} = \mathbf{a} + \tfrac{1}{2}(\mathbf{b} - \mathbf{a}) = \tfrac{1}{2}(\mathbf{a} + \mathbf{b})$$

Line (R^2)

- The line ℓ with equation $2x + y = 0$
- $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, then the equation becomes $\mathbf{n} \cdot \mathbf{x} = 0$.
- ℓ as $\mathbf{x} = t\mathbf{d}$.

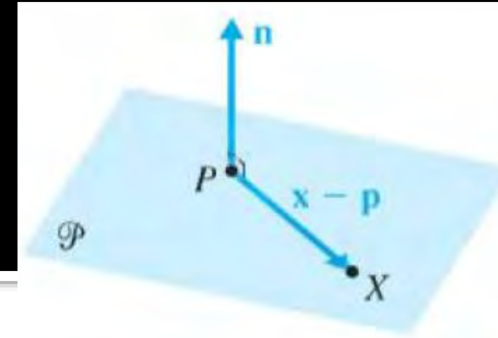


The line $2x + y = 0$ | A normal vector $\mathbf{n}$ | A direction vector $\mathbf{d}$ | $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$

Equations of Lines in $\mathbb{R}^2$

Normal Form	General Form	Vector Form	Parametric Form
$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$	$ax + by = c$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$\begin{cases} x = p_1 + td_1 \\ y = p_2 + td_2 \end{cases}$

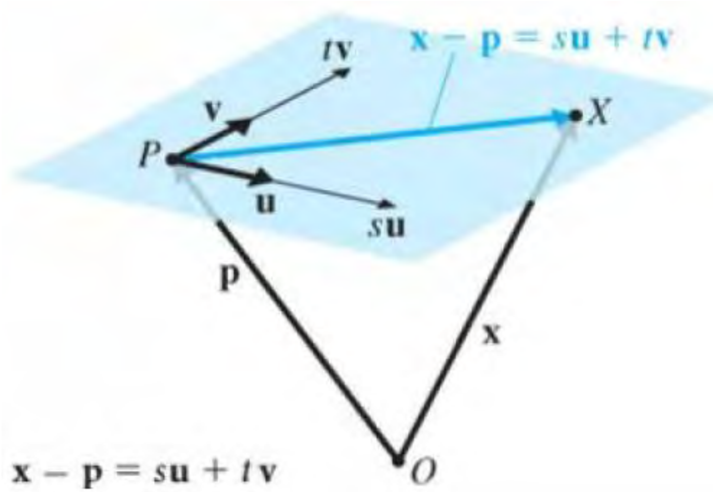
Plan (R^3)



$$\mathbf{n}\cdot(\mathbf{x}-\mathbf{p}) = 0$$

$$\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

$$ax + by + cz = d \,(\text{where } d = \mathbf{n}\cdot\mathbf{p})$$



$$\mathbf{x}-\mathbf{p} = s\mathbf{u} + t\mathbf{v}$$

Table 1.3 Lines and Planes in $\mathbb{R}^3$

	Normal Form	General Form	Vector Form	Parametric Form
Lines	$\begin{cases}\mathbf{n}_1\cdot\mathbf{x} = \mathbf{n}_1\cdot\mathbf{p}_1 \\ \mathbf{n}_2\cdot\mathbf{x} = \mathbf{n}_2\cdot\mathbf{p}_2\end{cases}$	$\begin{cases}a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2\end{cases}$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$\begin{cases}x = p_1 + td_1 \\ y = p_2 + td_2 \\ z = p_3 + td_3\end{cases}$
Planes	$\mathbf{n}\cdot\mathbf{x} = \mathbf{n}\cdot\mathbf{p}$	$ax + by + cz = d$	$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$	$\begin{cases}x = p_1 + su_1 + tv_1 \\ y = p_2 + su_2 + tv_2 \\ z = p_3 + su_3 + tv_3\end{cases}$

Reference

- Chapter 2,3,4: LINEAR ALGEBRA: Theory, Intuition, Code
- Chapter 1: Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares
- Chapter 8: Linear Algebra and its applications
- Chapter 2: Linear Algebra Jim Hefferon
- Chapter 4: Linear Algebra Devid Cherney