Linear Algebra

Elementary Row Operations

Department of Computer Engineering

Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu
Maryam Ramezani maryam.ramezani@sharif.edu

Table of contents

01

Vector Operation

Dot Product

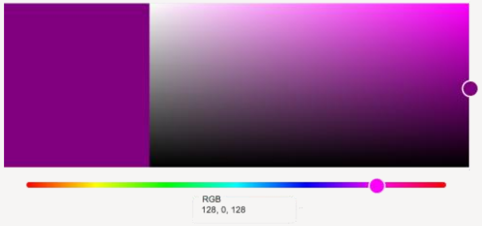
Review & **Geometric Interpretation**

Categorical (Non-numerical) Data

- Sometimes you work with categorical data in machine learning.
- It is common to encode categorical variables to make them easier to work with and learn by some techniques. A popular encoding for categorical variables is the one hot encoding.
- A one hot encoding is:

Example

$$\text{red}=\begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \text{green}=\begin{bmatrix}0\\1\\0\end{bmatrix} \qquad \text{blue}=\begin{bmatrix}0\\0\\1\end{bmatrix}$$



Categorical (Non-numerical) Data

- One-Hot Encodings (standard basis vector)
  - Assign to each word a vector with one 1 and 0s elsewhere.
  - Suppose our language only has four words:

$$\text{apple} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \quad \text{cat} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad \text{house} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad \text{tiger} = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

Drawbacks

- Very sparse vectors.
- Are never **similar**!

How to measure the similarity?

- **Dot Product**
  - The product of numbers is another number.
  - The dot product of vectors is not another vector! It is a number!!

$$2 \times 5 = 10$$

Numbers

VS

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} = 7$$

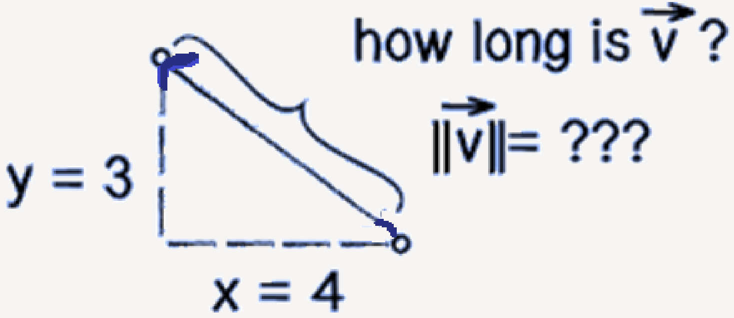
Vectors

A number

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

Length of vector

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



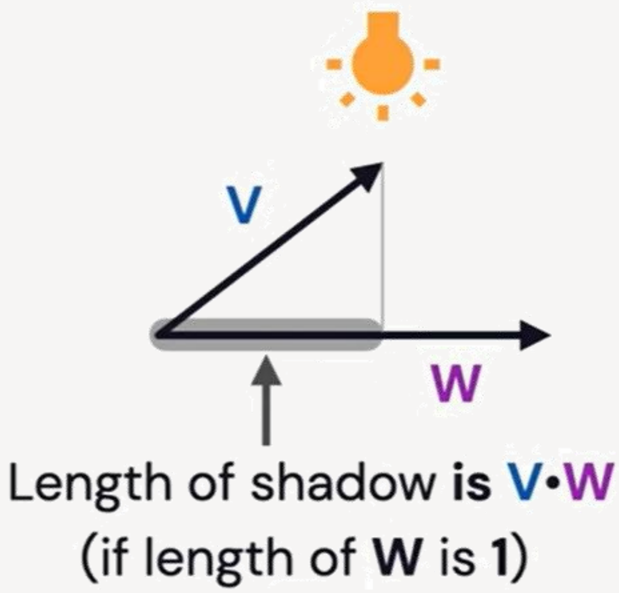
$\mathrm{v.v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} . \begin{bmatrix} 4 \\ 3 \end{bmatrix}$=16+9=25

Length(v)=5

$$a^T a = \|a\|^2 = \sum_{i=1}^{n} a_i a_i = \sum_{i=1}^{n} a_i^2$$

Dot Product (Geometric Interpretation and Intuition)

- Represents the length of the “shadow” of one vector along another.
- This indicates how similar the two vectors are.



One-Hot Encodings Drawbacks

$$\text{apple} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{cat} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{house} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{tiger} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{apple} \cdot \text{cat} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \text{tiger} \cdot \text{cat}$$

Vector Operations

- Vector-Vector Addition
- Vector-Vector Subtraction
- Scalar-Vector Product
- Vector-Vector Products:
  - $\boldsymbol{x}.\boldsymbol{y}$ is called the inner product or dot product or scalar product of the vectors: $x^T y$ or $y^T x$
    - $< a, b >$ $< a|b >$ (a, b) $a.b$

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

  - Transpose of dot product:
    - $(a.b)^T = (a^T b)^T = (b^T a) = (b.a) = b^T a$
  - Length of vector

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?

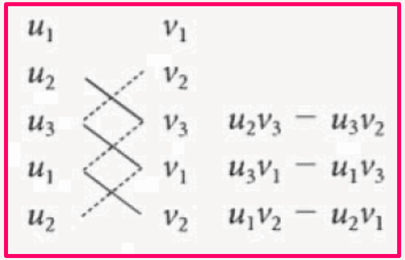
$$\underbrace{\mathbf{u}^T(\mathbf{v}^T\mathbf{w})}_{\substack{\text{vector–scalar product} \\ \text{row vector}}} \; ? \; \underbrace{(\mathbf{u}^T\mathbf{v})^T\mathbf{w}}_{\substack{\text{scalar–vector product} \\ \text{column vector}}}$$

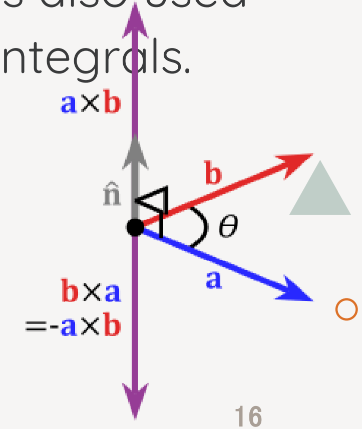
Cross product

- 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 (×).

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\| \sin(\theta_{ab}) \qquad \mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

- 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.





Vector Operations

- 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: $(xy^T)_{ij} = x_i y_j$

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

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 Product Properties

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

Vector Operations

- Vector-Vector Products:
  - Hadamard
  - 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 Product Properties

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

02

Matrix Multiplication

Basic Notation

- By $A \in \mathbb{R}^{m \times n}$ we denote a matrix with m rows and n columns, where the entries of A are real numbers.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

Definition

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

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

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

Matrix-Vector Multiplication

- If we write A by rows, then we can express Ax as,

$$A \in \mathbb{R}^{m \times n} \qquad y = Ax = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix}. \qquad a_i^T x = \sum_{j=1}^{n} a_{ij} x_j$$

- If we write A by columns, then we have:

$$y = Ax = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [a_1]x_1 + [a_2]x_2 + \cdots + [a_n]x_n \, .$$

  - y is a **linear combination** of the columns A.

We will learn in next lectures

columns of A are linearly independent if Ax = 0 implies x = 0

Matrix-Vector Multiplication

It is also possible to multiply on the left by a row vector.

- If we write A by columns, then we can express $x^T A$ as,

$$y^T = x^T A = x^T \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} = [x^T a_1 \quad x^T a_2 \quad \cdots \quad x^T a_n]$$

- Expressing A in terms of rows we have:

$$y^T = x^T A = [x_1 \quad x_2 \quad \cdots \quad x_m] \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

$$= x_1 [- \quad a_1^T \quad -] + x_2 [- \quad a_2^T \quad -] + \ldots + x_m [- \quad a_m^T \quad -]$$

  - y^T is a linear combination of the rows of A.

Matrix-Vector Multiplication

- $A(u+v) = Au + Av$
- $(A+B)u = Au + Bu$
- $(\alpha A)u = \alpha(Au) = A(\alpha u) = \alpha Au$
- $0u = 0$
- $A0 = 0$
- $Iu = u$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

Example: Write in matrix-vector multiplication

- Column j: $a_j =$
- Row i: $a_i^T =$
- Vector sum of rows of $A =$
- Vector sum of columns of $A =$

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

Matrix-Matrix Multiplication

Definition

Let A be an $m \times n$ matrix over the field F and let B be an $n \times p$ matrix over F. The product AB is the $m \times p$ matrix C whose i, j entry is:

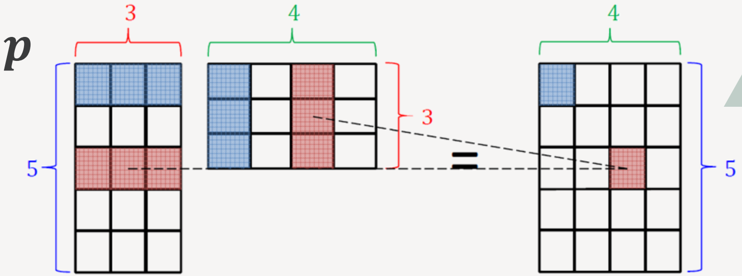
$$C_{ij} = \sum_{r=1}^{n} A_{ir} B_{rj}$$

- $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p} \rightarrow \mathbb{R}^{m \times p}$
  - a_i rows of A, b_j cols of B

$$\boldsymbol{C = AB} \qquad \boldsymbol{for\ 1 \le i \le m\ , 1 \le j \le p}$$

$$\boldsymbol{dot\ product(a_i . b_j)}$$

$$\boldsymbol{C_{ij} = a_i^T b_j}$$



Matrix-Matrix Multiplication (different views)

1. As a set of vector-vector products

$$C = AB = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_p \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_p \end{bmatrix}$$

2. As a sum of outer products

$$C = AB = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} - & b_1^T & - \\ - & b_2^T & - \\ & \vdots & \\ - & b_n^T & - \end{bmatrix} = \sum_{i=1}^{n} a_i b_i^T$$

Matrix-Matrix Multiplication (different views)

3. As a set of matrix-vector products.

$$C = AB = A\begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ Ab_1 & Ab_2 & \cdots & Ab_p \\ | & | & & | \end{bmatrix}$$

Here the ith column of C is given by the matrix-vector product with the vector on the right, $c_i = Ab_i$. These matrix-vector products can in turn be interpreted using both viewpoints given in the previous subsection.

4. As a set of vector-matrix products.

$$C = AB = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} B = \begin{bmatrix} - & a_1^T B & - \\ - & a_2^T B & - \\ & \vdots & \\ - & a_m^T B & - \end{bmatrix}$$

Matrix-Matrix Multiplication

- Properties:
  - Associative

$$(AB)C = A(BC)$$

  - Distributive

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

  - <u>NOT commutative</u>

$$AB \neq BA$$

  - Dimensions may not even be conformable

03

Elementary Row Operations

Gaussian Elimination: Elementary Row Operations

- Elementary Row Operations
  1. Scaling: Multiply all entries in a row by a nonzero scalar.
  2. Replacement: Replace one row by the sum of itself and a multiple of another row.
  3. Interchange: Interchange two rows.

- Elementary Row Operation is a special type of function e on $m \times n$ matrix A and gives an $m \times n$ matrix $e(A)$ where $c \neq 0$.
  1. Scaling : $e(A)_{\mathrm{ij}} = \mathrm{cA_{ij}}$
  2. Replacement: $e(A)_{\mathrm{ij}} = A_{ij} + \mathrm{cA_{kj}}$
  3. Interchange: $e(A)_{\mathrm{ij}} = \mathrm{A_{kj}}$, $e(A)_{\mathrm{kj}} = \mathrm{A_{ij}}$

In defining $e(A)$, it is not really important how many columns A has, but the number of rows of A is crucial.

Inverse of Elementary Row Operation

Theorem

The inverse operation (function) of an elementary row operation exists and is an elementary row operation of the same type.

Proof:

Proof. (1) Suppose e is the operation which multiplies the rth row of a matrix by the non-zero scalar c. Let e_1 be the operation which multiplies row r by c^{-1}. (2) Suppose e is the operation which replaces row r by row r plus c times row s, $r \neq s$. Let e_1 be the operation which replaces row r by row r plus $(-c)$ times row s. (3) If e interchanges rows r and s, let $e_1 = e$. In each of these three cases we clearly have $e_1(e(A)) = e(e_1(A)) = A$ for each A. ∎

Row-Equivalent

Definition

If A and B are $m \times n$ matrices over the field F, we say that B is **row-equivalent** to A if B can be obtained from A by a finite sequence of elementary row operations.

Note (from pervious theorem and this definition)

- Each matrix is row-equivalent to itself
- If B is row-equivalent to A, then A is row-equivalent to B.
- If B is row-equivalent to A, C is row-equivalent to B, then C is row-equivalent to A

04

Elementary Matrices

Elementary Matrices

Definition

A $m \times m$ matrix is an elementary matrix if it can be obtained from the $m \times m$ identity matrix by means of a single elementary row operation.

Example

Find all 2×2 elementary matrices.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$$

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}, \quad c \neq 0, \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, \quad c \neq 0.$$

Elementary Matrices and Elementary Row Operation

Theorem

Let e be an elementary row operation and let E be the $m \times m$ elementary matrix $E = e(I)$. Then, for every $m \times n$ matrix A:

$$e(A) = EA$$

Proof:

Proof. The point of the proof is that the entry in the ith row and jth column of the product matrix EA is obtained from the ith row of E and the jth column of A. The three types of elementary row operations should be taken up separately. We shall give a detailed proof for an operation of type (ii). The other two cases are even easier to handle than this one and will be left as exercises. Suppose $r \neq s$ and e is the operation 'replacement of row r by row r plus c times row s.' Then

$$E_{ik} = \begin{cases} \delta_{ik}, & i \neq r \\ \delta_{rk} + c\delta_{sk}, & i = r. \end{cases}$$

Therefore,

$$(EA)_{ij} = \sum_{k=1}^{m} E_{ik}A_{kj} = \begin{cases} A_{ij}, & i \neq r \\ A_{rj} + cA_{sj}, & i = r. \end{cases}$$

In other words $EA = e(A)$. ∎

Multiplication of a matrix on the left by a square matrix performs row operations.

Elementary Matrices

Example

Matrix	Elementary row operation	Elementary matrix
$\begin{bmatrix} 1 & 0 & 2 \\ -2 & 0 & -3 \\ 0 & 2 & 0 \end{bmatrix}$	$R_2 \leftarrow R_2 + 2R_1$	$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$	$R_2 \leftrightarrow R_3$	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$R_2 \leftarrow \frac{1}{2} R_2$	$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$R_1 \leftarrow R_1 + (-2)R_3$	$E_4 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$		

(From property $(AB)C = A(BC)$)

$$E_4(E_3(E_2(E_1A))) = \left(E_4(E_3(E_2E_1))\right)A$$

Row-Equivalent and Elementary Matrices

Theorem

Let A and B be $m \times n$ matrices over the field F. Then B is row-equivalent to A if and only if $B = PA$, where P is a product of $m \times m$ elementary matrices.

Proof:

Corollary. *Let* A *and* B *be* $m \times n$ *matrices over the field* F. *Then* B *is row-equivalent to* A *if and only if* $B = PA$, *where* P *is a product of* $m \times m$ *elementary matrices.*

Proof. Suppose $B = PA$ where $P = E_s \cdots E_2E_1$ and the E_i are $m \times m$ elementary matrices. Then E_1A is row-equivalent to A, and $E_2(E_1A)$ is row-equivalent to E_1A. So E_2E_1A is row-equivalent to A; and continuing in this way we see that $(E_s \cdots E_1)A$ is row-equivalent to A.

Now suppose that B is row-equivalent to A. Let $E_1, E_2, \ldots, E_s$ be the elementary matrices corresponding to some sequence of elementary row operations which carries A into B. Then $B = (E_s \cdots E_1)A$. ∎

05

Linear Equations

Systems of Linear Equations

Definition

A system of m linear equations with n unknowns:

- F is a field, we want to find n scalars (elements of F) $x_1, \dots, x_n$ which satisfy the conditions: (A_{ij}, y_k are elements of F)

$$\begin{gathered} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n = y_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n = y_2 \\ \dots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n = y_m \end{gathered}$$

If $y_1 = y_2 = \cdots = y_m = 0$, we say that the system is homogeneous.

A solution of this system of linear equations is vector $\begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$ whose components satisfy

$$x_1 = s_1, \dots, x_n = s_n$$

Linear Equation (Geometric Interpretation and Intuition)

- Consider this simple system of equations,

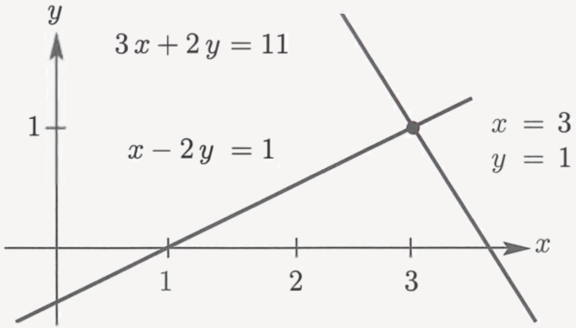
$$x - 2y = 1$$
$$3x + 2y = 11$$

- Can be expressed as a matrix-vector multiplication
- Matrix Equation: $Ax=b$

$$\underbrace{\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 11 \end{bmatrix}}_{b}$$

- A is often called coefficient matrix $\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}$

- Ab is an Augmented matrix: $\begin{bmatrix} 1 & -2 & 1 \\ 3 & 2 & 11 \end{bmatrix}$



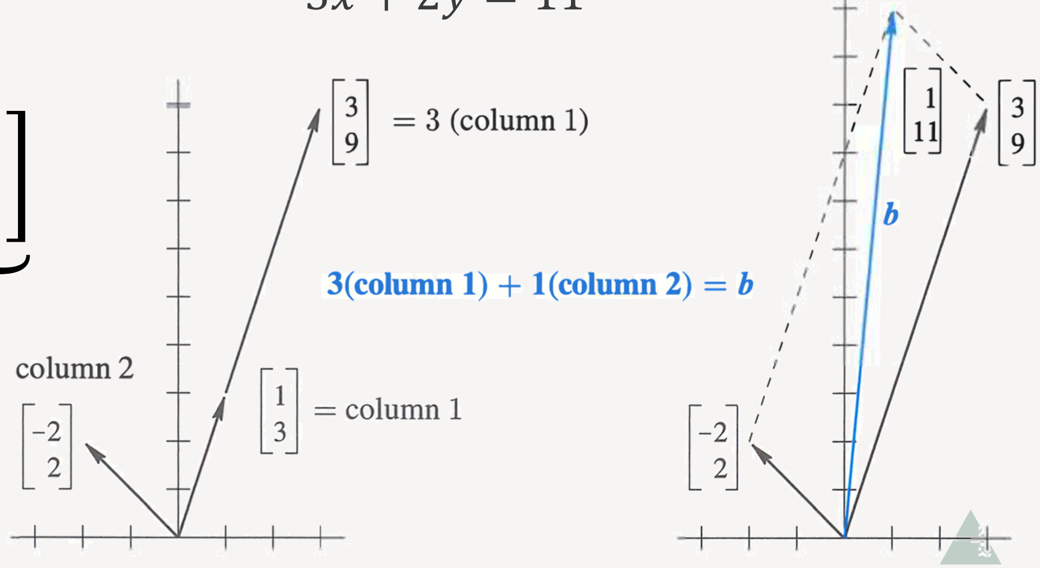
Vectors & Linear Equation

- Also, Can be expressed as linear combination of cols:

$$x - 2y = 1$$
$$3x + 2y = 11$$

$$\underbrace{\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}x\\y\end{bmatrix}}_{x} = \underbrace{\begin{bmatrix}1\\11\end{bmatrix}}_{b}$$

$$x\begin{bmatrix}1\\3\end{bmatrix} + y\begin{bmatrix}-2\\2\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix} = b$$



- Same for n equation, n variable

Idea Of Elimination

- Subtract a multiple of equation (1) from (2) to eliminate a variable

$$\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned}$$

multiply equation 1 by 3

Subtract to eliminate $3x$

$$\begin{aligned} x - 2y &= 1 \\ 8y &= 8 \end{aligned}$$

$$\underbrace{\begin{bmatrix} 1 & -2 \\ 0 & 8 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 8 \end{bmatrix}}_{c}$$

A has become a upper triangle matrix U

Idea Of Elimination (Row Reduction Algorithm)

Definition

A leading entry of a row refers to **the left most** nonzero entry in a nonzero row.

- The **pivots** are on the diagonal of the triangle after elimination. **The first non zero element in each row** (boldface 2 below is the first pivot)

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 10 \end{aligned}$$

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 1y + 1z &= 4 \\ 4z &= 8 \end{aligned}$$

  - Step 1: subtract $2 * (1)$ from (2) to eliminate x's in (2) $\Rightarrow 1y + 1z = 4$
  - Step 2: add (1) to (3) to totally eliminate x $\Rightarrow 1y + 5z = 12$
  - Step 3: subtract new (2) from new (3) $\Rightarrow 4z = 8$

Definition

The variables corresponding to pivot columns in the matrix are called basic variables.

The other variables are called a free variable.

$$\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix} \quad \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & x \end{bmatrix}$$

Homogenous system

Theorem

If A and B are row-equivalent $m \times n$ matrices, the homogenous systems of linear equations $Ax = 0$ and $Bx = 0$ have exactly the same solutions.

Proof:

Proof. Suppose we pass from A to B by a finite sequence of elementary row operations:

$$A = A_0 \to A_1 \to \cdots \to A_k = B.$$

It is enough to prove that the systems $A_jX = 0$ and $A_{j+1}X = 0$ have the same solutions, i.e., that one elementary row operation does not disturb the set of solutions.

So suppose that B is obtained from A by a single elementary row operation. No matter which of the three types the operation is, (1), (2), or (3), each equation in the system $BX = 0$ will be a linear combination of the equations in the system $AX = 0$. Since the inverse of an elementary row operation is an elementary row operation, each equation in $AX = 0$ will also be a linear combination of the equations in $BX = 0$. Hence these two systems are equivalent, and by Theorem 1 they have the same solutions. ∎

45

Homogenous system

Example

Find the solution for this system.
Suppose F is the field of complex number and the coefficient matrix is:

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

In performing row operations it is often convenient to combine several operations of type (2). With this in mind

$$\begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 2+i \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & 1 \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Thus the system of equations

$$\begin{aligned} -x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 0 \end{aligned}$$

has only the trivial solution $x_1 = x_2 = 0$.

Solution of system of linear equations

Definition

The two systems of linear equations are equivalent if each equation in each system is a linear combination of the equations in other system.

Theorem

Equivalent systems of linear equations have exactly the same solutions.

Proof:

Note

- It is important to note that row operations are reversible. If two rows are interchanged, they can be returned to their original positions by another interchange.
- If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Existence and Uniqueness Questions

- A system of linear equations has:
  - No solution → inconsistent
  - Exactly one solution
  - Infinitely many solutions } → consistent

Next session:

1. Is the system consistent? That is, does at least one solution exist?
2. If a solution exists, is it the only one? That is, is the solution unique?

Conclusion

- Different view of matrix multiplication
- Linear combination and matrix multiplication
- Associativity of three matrices multiplication
- Gaussian Elimination
- Row-equivalent of two matrices
- Elementary matrices
- System of linear equations
- Equivalent systems of linear equations have exactly the same solutions.

Resources

- ❑ Chapter 1: Kenneth Hoffman and Ray A. Kunze. Linear Algebra. PHI Learning, 2004.
- ❑ Chapter 1: David C. Lay, Steven R. Lay, and Judi J. McDonald. Linear Algebra and Its Applications. Pearson, 2016.
- ❑ Chapter 2: David Poole, Linear Algebra: A Modern Introduction. Cengage Learning, 2014.
- ❑ Chaper1: Gilbert Strang. Introduction to Linear Algebra. Wellesley-Cambridge Press, 2016.