Linear Algebra: Review

CS115 - Math for Computer Science

TS. Lương Ngọc Hoàng
TS. Dương Việt Hằng

September 9, 2023

Roadmap

- Systems of Linear Equations
- Matrices
- Solving Systems of Linear Equations
- Linear Independence
- Basis and Rank
- Linear Mappings

System of Linear Equations

- For unknown variables $(x_1, \cdots, x_n) \in \mathbb{R}^n$,

$$
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
&\vdots \\
a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

- Matrix representations:

$$
\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \Longleftrightarrow \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\boldsymbol{A}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{\boldsymbol{b}}
$$

Identity Matrix and Matrix Properties

- A square matrix[1] $\boldsymbol{I}_n$ with $I_{ii} = 1$ and $I_{ij=0}$ for $i \neq j$, where n is the number of rows and columns. For example,

$$\boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Associativity: For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n \times p}$, $\boldsymbol{C} \in \mathbb{R}^{p \times q}$, $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$
- Distributivity: For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$, and $\boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{n \times p}$,
  (i) $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$ and (ii) $\boldsymbol{A}(\boldsymbol{C} + \boldsymbol{D}) = \boldsymbol{AC} + \boldsymbol{AD}$
- Multiplication with the identity matrix: For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$

[1] $\#$ of rows $= \#$ of cols

Matrix: Addition and Multiplication

- For two matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{B} \in \mathbb{R}^{m \times n}$,

$$\boldsymbol{A} + \boldsymbol{B} := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

- For two matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times k}$, the elements c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m \times k}$ is:

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k.$$

Inverse Matrix

- For a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B}$ is the inverse of A, denoted by $\boldsymbol{A}^{-1}$, if

$$\boldsymbol{AB} = \boldsymbol{I}_n = \boldsymbol{BA}.$$

- Called regular/invertible/nonsingular, if it exists.
- If it exists, it is unique.

Transpose Matrix

- For a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ with $b_{ij} = a_{ji}$ is the transpose of $\boldsymbol{A}$, which we denote by $\boldsymbol{A}^\top$.

- Example. For $\boldsymbol{A} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$,

$$\boldsymbol{A}^\top = \begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix}$$

- If $\boldsymbol{A} = \boldsymbol{A}^\top$, $\boldsymbol{A}$ is called symmetric.

Inverse and Transpose: More Properties

- $AA^{-1} = I = A^{-1}A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A+B)^{-1} \neq A^{-1} + B^{-1}$
- $(A^{\top})^{\top} = A$
- $(A+B)^{\top} = A^{\top} + B^{\top}$
- $(AB)^{\top} = B^{\top}A^{\top}$

Scalar Multiplication

- Multiplication by a scalar $\lambda \in \mathbb{R}$ to $\boldsymbol{A} \in \mathbb{R}^{m\times n}$
- Example. For $\boldsymbol{A} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$, $3 \times \boldsymbol{A} = \begin{pmatrix} 0 & 6 \\ 3 & -3 \\ 0 & 3 \end{pmatrix}$
- Associativity
  - $(\lambda\psi)\boldsymbol{C} = \lambda(\psi\boldsymbol{C})$
  - $\lambda(\boldsymbol{BC}) = (\lambda\boldsymbol{B})\boldsymbol{C} = \boldsymbol{B}(\lambda\boldsymbol{C}) = (\boldsymbol{BC})\lambda$
  - $(\lambda\boldsymbol{C})^{\top} = \boldsymbol{C}^{\top}\lambda^{\top} = \boldsymbol{C}^{\top}\lambda = \lambda\boldsymbol{C}^{\top}$
- Distributivity
  - $(\lambda + \psi)\boldsymbol{C} = \lambda\boldsymbol{C} + \psi\boldsymbol{C}$
  - $\lambda(\boldsymbol{B} + \boldsymbol{C}) = \lambda\boldsymbol{B} + \lambda\boldsymbol{C}$

Solving Systems of Linear Equations

- Start as usual by getting echelon form.

$$\begin{array}{rcl} x + y - z &=& 2 \\ 2x - y \phantom{{}+2z} &=& -1 \\ x - 2y + 2z &=& -1 \end{array} \xrightarrow[-1\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{array}{rcl} x + y - z &=& 2 \\ -3y + 2z &=& -5 \\ -3y + 3z &=& -3 \end{array} \xrightarrow{-1\rho_2+\rho_3} \begin{array}{rcl} x + y - z &=& 2 \\ -3y + 2z &=& -5 \\ z &=& 2 \end{array}$$

- Make all the leading entries one.

$$\xrightarrow{(-1/3)\rho_2} \begin{array}{rcl} x + y - z &=& 2 \\ y - (2/3)z &=& 5/3 \\ z &=& 2 \end{array}$$

- Finish by using the leading entries to eliminate upwards, until we can read off the solution.

$$\begin{array}{rcl} x + y - z &=& 2 \\ y - (2/3)z &=& 5/3 \\ z &=& 2 \end{array} \xrightarrow[(2/3)\rho_3+\rho_2]{\rho_3+\rho_1} \begin{array}{rcl} x + y &=& 4 \\ y &=& 3 \\ z &=& 2 \end{array} \xrightarrow{-\rho_2+\rho_1} \begin{array}{rcl} x &=& 1 \\ y &=& 3 \\ z &=& 2 \end{array}$$

Algorithms for Solving System of Linear Equations

1. Pseudo-inverse

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{b} \Longleftrightarrow \boldsymbol{x} = (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{b}$$

   - $(\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top$: *Moore-Penrose pseudo-inverse*
   - many computations: matrix product, inverse, etc
2. Gaussian elimination
   - intuitive and constructive way
   - cubic complexity (in terms of # of simultaneous equations)
3. Iterative methods
   - practical ways to solve indirectly

   (a) stationary iterative methods: Richardson method, Jacobi method, Gaus-Seidel method, successive over-relaxation method

   (b) Krylov subspace methods: conjugate gradients, generalized minimal residual, biconjugate gradients

Linear Combinations of Linearly Independent Vectors

- Vector space V with k linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k$
- m linear combinations $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m$. (Q) Are they linearly independent?

$$\boldsymbol{x}_1 = \lambda_{11}\boldsymbol{b}_1 + \lambda_{21}\boldsymbol{b}_2 + \cdots + \lambda_{k1}\boldsymbol{b}_k$$

$$\vdots$$

$$\boldsymbol{x}_m = \lambda_{1m}\boldsymbol{b}_1 + \lambda_{2m}\boldsymbol{b}_2 + \cdots + \lambda_{km}\boldsymbol{b}_k$$

$$\boldsymbol{x}_j = \overbrace{\begin{pmatrix}\boldsymbol{b}_1, & \cdots, & \boldsymbol{b}_k\end{pmatrix}}^{\boldsymbol{B}} \overbrace{\begin{pmatrix}\lambda_{1j} \\ \vdots \\ \lambda_{kj}\end{pmatrix}}^{\boldsymbol{\lambda}_j}, \quad \boldsymbol{x}_j = \boldsymbol{B}\boldsymbol{\lambda}_j$$

- $\sum_{j=1}^{m} \psi_j \boldsymbol{x}_j = \sum_{j=1}^{m} \psi_j \boldsymbol{B}\boldsymbol{\lambda}_j = \boldsymbol{B}\sum_{j=1}^{m} \psi_j \boldsymbol{\lambda}_j$
- $\{\boldsymbol{x}\}$ linearly independent $\Longleftrightarrow$ $\{\boldsymbol{\lambda}\}$ linearly independent

Generating Set and Basic

- Definition. A vector space $V = (\mathcal{V}, +, \cdot)$ and a set of vectors $\mathcal{A} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subset \mathcal{V}$.
  - If every $v \in \mathcal{V}$ can be expressed as a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_k$, $\mathcal{A}$ is called a generating set of V.
  - The set of all linear combinations of $\mathcal{A}$ is called the span of $\mathcal{A}$.
  - If $\mathcal{A}$ spans the vector space V, we use $V = \text{span}[\mathcal{A}]$ or $V = \text{span}[\mathbf{x}_1, \ldots, \mathbf{x}_k]$
- Definition. The minimal generating set $\mathcal{B}$ of V is called basis of V. We call each element of $\mathcal{B}$ basis vector. The number of basis vectors is called dimension of V.
- Properties
  - $\mathcal{B}$ is a maximally[2] linearly independent set of vectors in V.
  - Every vector $x \in V$ is a linear combination of $\mathcal{B}$, which is unique.

[2] Adding any other vector to this set will make it linearly dependent.