Linear Algebra

AI ToolKit

Fall Semester, 2023

Scalars and Vectors

- **Scalar,** $x \in \mathbb{R}$
  - A single number.
  - Usually written with x, y (normal lower-case letters).

- **Vector,** $\mathbf{x} \in \mathbb{R}^d$
  - A fixed-length arrays of scalars. $\mathbf{x} = [x_1, \ldots, x_d]^\top$
    - Each scalar x_i is called as “element”, “entries”, “components”.
  - Usually written with $\mathbf{x}$ (bold lower-case letters).
  - Can be represented with arrows.

Matrices and Tensors

- **Matrix,** $\mathbf{A} \in \mathbb{R}^{m \times n}$ (m rows and n columns)
  - Usually written with $\mathbf{A}$ (bold upper-case letters).

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, a_{ij} \in \mathbb{R}$$

    - Can be represented as $\mathbf{a} \in \mathbb{R}^{mn}$, by stacking all n columns vertically, or concatenating m rows horizontally.
- **Tensor**
  - Stack k number of $\mathbf{X} \in \mathbb{R}^{m \times n}$.
  - The results can be called as 3th-order array, tensor $\mathsf{X} \in \mathbb{R}^{k \times m \times n}$
  - Can be written with X (capital letters with a special font face).

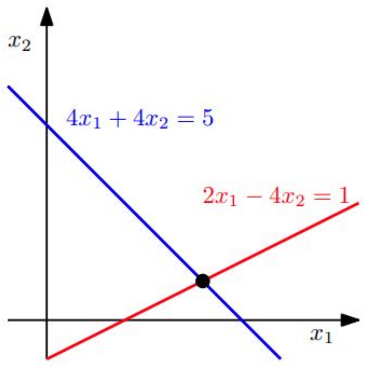
Linear Algebra

- Gives us the tools for solving linear equations.

$$a_{11}x_1 + \ldots + a_{1n}x_n = b_1$$
$$\vdots$$
$$a_{m1}x_1 + \ldots + a_{mn}x_n = b_n$$

- $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbb{R}^n$: the unknown variable of this system.
- Every $\mathbf{x}$ that satisfies above equation is a **solution** of the linear equation system.
- There can be *one solution*, *no solution* or *infinite solutions*.

Linear Equation: Example



Each linear equation defines a **line** on $x_1 x_2$ - plane.

- A solution set is the intersection of these lines.
- *Infinite solution* : If two equations refers to the same line.
- *No solution* : If two lines are parallel.

- When there are three variables,
  - Each linear equation defines a **plane** in three-dimensional space.
  - A solution set is the intersection between planes.
    - Point
    - Line
    - Plane
    - Empty

Two ways of representing linear equations

1. $$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \ldots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

- Finding a proper weight vector $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbb{R}^n$ for this linear combination between column vectors, which are $\mathbf{a}_j = [a_{1j}, \ldots, a_{mj}]^T$.

2. $$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

- $\mathbf{Ax} = \mathbf{b}$: How do we solve this?

Matrix Addition and Multiplication

- **Addition** : an element-wise sum between two same-sized matrices.

- **Multiplication**
  - For $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times l}$, the element c_{ij} of the product $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m\times l}$ is :
    - $c_{ij} = \sum_{l=1}^{k} a_{ik}b_{kj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k.$
      - An inner product between i-th row of $\mathbf{A}$ with the j-th column of $\mathbf{B}$.
    - # columns of $\mathbf{A}$ (dim. of row vector) == # rows of $\mathbf{B}$ (dim. of column vector).
  - $\mathbf{BA}$ is not defined if $l \neq n$.
  - $\mathbf{AB} \neq \mathbf{BA}$
  - Element-wise multiplication is called as a Hadamard product.

Associativity and Distributivity of Multiplication

- **Associativity**
  - $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
  - For scalar λ and ϕ,
    - $(\lambda\phi)\mathbf{C} = \lambda(\phi\mathbf{C})$
    - $\lambda(\mathbf{BC}) = (\lambda\mathbf{B})\mathbf{C} = \mathbf{B}(\lambda\mathbf{C}) = (\mathbf{BC})\lambda$
    - $(\lambda\mathbf{C})^\top = \mathbf{C}^\top\lambda^\top = \mathbf{C}^\top\lambda = \lambda\mathbf{C}^\top$
- **Distributivity**
  - $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
  - $\mathbf{A}(\mathbf{C} + \mathbf{D}) = \mathbf{AC} + \mathbf{DC}$
  - For scalar λ and ϕ,
    - $(\lambda + \phi)\mathbf{C} = \lambda\mathbf{C} + \phi\mathbf{C}$
    - $\lambda(\mathbf{B} + \mathbf{C}) = \lambda\mathbf{B} + \lambda\mathbf{C}$

Identity, Inverse and Transpose

- **Identity Matrix**
  - $\mathbf{I} \in \mathbb{R}^{n \times n}$ is an identity matrix if 1 on the diagonal and 0 everywhere else.
  - ex) $\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 3 x 3 identity matrix
  - $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$
- **Inverse Matrix $\mathbf{A}^{-1}$**
  - For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, another square matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ is $\mathbf{A}^{-1}$ if $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$.
  - "Singular, noninvertible" $\mathbf{A}$: $\mathbf{A}^{-1}$ does not exists.
- **Transpose:** For $\mathbf{A}, \mathbf{B}$ with $b_{ij} = a_{ji}$ is called the transpose of $\mathbf{A}$, written as $\mathbf{A}^\top$.
  - A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric if $\mathbf{A} = \mathbf{A}^\top$.

Vector Space

- A set of elements (vectors) with rules for **“linear combinations”**
    1. Addition,
    2. Multiplication by real numbers (scalars).
- $\mathbb{R}^n$: a space consists of all column vectors with n components.
- Addition and Scalar Multiplications are required to satisfy below rules:
    - $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$
    - $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$
    - $\mathbf{x}+\mathbf{0}=\mathbf{x}$
    - There is a unique “zero vector” such that $\mathbf{x}+\mathbf{0}=\mathbf{x}$ for all $\mathbf{x}$.
    - $1\mathbf{x}=\mathbf{x}$
    - $(c_1c_2)\mathbf{x}=c_1(c_2\mathbf{x})$
    - $c(\mathbf{x}+\mathbf{y})=c\mathbf{x}+c\mathbf{y}$
    - $(c_1+c_2)\mathbf{x}=c_1\mathbf{x}+c_2\mathbf{x}$

Vector Subspace

- A non-empty subset of a vector space:
  1. If $\mathbf{x}$ and $\mathbf{y}$ are in the subspace, $\mathbf{x}+\mathbf{y}$ is in the subspace.
  2. If multiply $\mathbf{x}$ in the subspace by any scalar c, $c\mathbf{x}$ is in the subspace.
     - Linear combinations stays in the subspace.
     - “Closed” under addition and scalar multiplication.

- A zero vector belongs to every subspace, due to zero scalar.

- **Example**
  - For a vector space by 3 by 3 matrices,
    - A set of symmetric matrices. (?)
    - A set of lower triangular matrices is its subspace. (?)
    - A set of matrices with all positive elements. (?)

Span, Rank, Linear Independence

- **Span**
  - If a vector space $\mathbf{V}$ consists of all linear combinations of $\mathbf{w}_1, \ldots \mathbf{w}_l$, then these vectors **span the space.**
    - Every vector in $\mathbf{V}$ is some combination of the $\mathbf{w}$'s.
      - $\mathbf{v} = c_1\mathbf{w}_1 + \ldots + c_l\mathbf{w}_l$
  - The column space of $\mathbf{A}$ is "spanned" by its columns.
- **Rank** r
  - The number of genuinely independent rows/columns in the matrix.
  - When $r = m = n$, the matrix has an inverse.
- **Linear Independence**
  - Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent,
    If $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$ only happens when $c_1 = \ldots = c_k = 0$.
  - If any c's are non-zero, the $\mathbf{v}$'s are linearly dependent
    - One vector is a linear combination of the others.

Linear Transformations

- $\mathbf{T}(\mathbf{x})$ is a linear transformation if $\mathbf{T}(c\mathbf{x} + d\mathbf{y}) = c(\mathbf{T}(\mathbf{x})) + d(\mathbf{T}(\mathbf{y}))$
- Consider transformation by matrix $\mathbf{A}$: $\mathbf{A}\mathbf{x}$.
- If basis : $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{x} = c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n$,
  - $\mathbf{A}\mathbf{x} = c_1(\mathbf{A}\mathbf{x_1}) + \ldots + c_n(\mathbf{A}\mathbf{x_n})$
  - If we know $\mathbf{A}\mathbf{x}$ for each vector in a basis, then we know $\mathbf{A}\mathbf{x}$ for each vector in the entire space.
- Is composition of two linear transformations is also linear transform?

  $\Rightarrow$

Basis for Vector Space

- A sequence of vectors having below two properties:
  1. The vectors are linearly independent (not too many vectors).
  2. The vectors span the vector space (not too few vectors).
- There is one and only one way to write a vector as a linear combination of basis vectors.
  $\Rightarrow$
- If columns of a matrix is a basis for $\mathbb{R}^n$, then the matrix must be square and invertible.
  $\Rightarrow$
- "Dimension" of a vector space is "the number of basis vectors".

Any Questions?