CS 357 - 08 Vectors, Matrices and Norms

Boyang Li (boyangl3)

Vector: An array of numbers that represents both a magnitude and a direction. A n-dimensional vector has n elements. A Vector is an element if a vector space.

Vector Space: A Vector Space is a set $\boldsymbol{V}$ of vectors with elements in the field $\boldsymbol{F}$. The elements in the field $\boldsymbol{F}$ are called scalars. There are 2 operation defined:

1. Closed under vector addition: $\forall \mathbf{v}, \mathbf{w} \in V$, $\mathbf{v} + \mathbf{w} \in V$.
2. Closed under scalar multiplication: $\forall \alpha \in F$, $\mathbf{v} \in V$, $\alpha\mathbf{v} \in V$.

Moreover, the following properties hold:

1. Associativity (Vector): $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
2. "Zero" vector: There exists a zero vector $\vec{0}$, such that $\forall \mathbf{u} \in V$, $\vec{0} + \mathbf{u} = \mathbf{u}$.
3. Additive inverse: $\forall \mathbf{u} \in V$, there exists $-\mathbf{u} \in V$, such that $\mathbf{u} + (-\mathbf{u}) = \vec{0}$.
4. Associativity (scalar): $\forall \alpha, \beta \in F, \mathbf{u} \in V$, $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u})$.
5. Distributivity: $\forall \alpha, \beta \in F, \mathbf{u} \in V$, $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$.
6. Unitarity: $\forall \mathbf{u} \in V$.

Linear combination and liearly independence:

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Where $\mathbf{v}_1$, $\mathbf{v}_2$, ... $\mathbf{v}_n$ are all vectors in the vector space V, and c_1, c_2, ... c_n are constants, then we call $\mathbf{x}$ is a linear combination of $\mathbf{v_1}$, $\mathbf{v}_2$, ... $\mathbf{v}_n$.

If any vector $\mathbf{x} \in V$ can be written in the form of the linear combination with uniquely defined scalar c_1, c_2, ... c_n, then we call $\mathbf{v}_1$, $\mathbf{v}_2$, ... $\mathbf{v}_n$ a **basis** of V. The size n of the basis is called the **dimension** of V. The dimension of $\mathbb{R}^n$ is n.

If $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \vec{0}$ only has one solution $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, then $\mathbf{v}_1$, $\mathbf{v}_2$, ... $\mathbf{v}_n$ are called **linearly independent**. Otherwise the vectors are linearly dependent, and at least one of the vectors can be written as a linear combination of the other vectors in the set. A basis is always linearly independent.

Inner product: An inner product is a function that takes 2 vectors from the same real vector space and returns a real number. There are 4 properties hold ($\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\alpha, \beta \in \mathbb{R}$):

1. Positivity: $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$
2. Definiteness: $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \vec{0}$
3. Symmetric: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
4. Linearity: $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{u}, w \rangle + \beta\langle \mathbf{v}, \mathbf{w} \rangle$.

If $\mathbf{U}, \mathbf{V} \in V$ and $\langle \mathbf{U}, \mathbf{V} \rangle = \vec{0}$, then $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal**. The standard inner product is the **dot product**: $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i$.

Linear transformations: A function $f : V \to W$ between 2 vectors v and w is considered to be **linear** if the following 2 properties hold:

1. $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$, where $\mathbf{u}, \mathbf{v} \in V$.
2. $f(c\mathbf{v}) = cf(\mathbf{v})$, for all $\mathbf{v} \in V$ and scalars c.

Where f is called a **linear transformation**. If n and m are the dimensions of V and W, then f can be represented as a $m \times n$ matrix $\mathbf{A}$.

Matrices:

- Special matrices:
  - **Zero Matrix:** A matrix with all entries to be zero.
  - **Identity Matrix:** An $n \times n$ matrix with 1 in the diagonal entries and 0 elsewhere. Can be considered as "1" because $\mathbf{AI}_n = \mathbf{A}$.
  - **Diagonal Matrix:** An $n \times n$ diagonal matrix has all zeroes at non-diagonal entries.
  - **Triangular Matrix:** A lower-triangular matrix ($\mathbf{L}$) is a square matrix that is entirely zero above the diagonal. An upper triangular matrix ($\mathbf{U}$) is a square matrix that is entirely zero below the diagonal. Here are the properties:
    * A $n \times n$ triangular matrix has $n(n-1)/2$ entries that must be zero, and $n(n+1)/2$ entries that are allowed to be non-zero.
    * Zero matrices, identity matrices, and diagonal matrices are all both lower triangular and upper triangular.
  - **Permutation Matrix:** A square matrix that is all zero, except for a single entry in each row and each column which is 1. Here are the properties:
    * Exactly n entries are zero.
    * Multiplying a vector with a permutation matrix permutes (rearranges) the order of the entries in the vector.
    * If $\mathbf{P}_{ij} = 1$, the ($\mathbf{Px}_i = \mathbf{x}_j$)
    * The inverse of a permutation matrix is its transpose, so $\mathbf{PP}^T = \mathbf{P}^T\mathbf{P} = \mathbf{I}$.
- Block form: A matrix in block form is a matrix partitioned into blocks. For example: $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$.
- Matrix rank: The rank of a matrix is the number of linearly independent columns of the matrix. It can also be shown that the matrix has the same number of linearly independent rows, as well. If $\mathbf{A}$ is an $m \times n$ matrix, then
  1. $\text{rank}(\mathbf{A}) \leq \min(m, n)$
  2. If $\text{rank}(\mathbf{A}) = \min(m, n)$, then $\mathbf{A}$ is full rank. Otherwise, A is rank deficient.

An $n \times n$ matrix $\mathbf{A}$ is invertable if and only if there exists another matrix $\mathbf{B}$ such at $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$. Then, $\mathbf{B}$ is called the inverse of A, or write as $\mathbf{A}^{-1}$. A square matrix is invertible if and only if it has full rank. A square matrix that is not invertible is called a singular matrix.

- Matrix-vector multiplication: There are 2 forms to represent $\mathbf{Ax} = \mathbf{b}$:

  1. Writing a matrix-vector multiplication as inner products of the rows $\mathbf{A}$:

  $$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1^T \cdot \mathbf{x} \\ \mathbf{a}_2^T \cdot \mathbf{x} \\ \vdots \\ \mathbf{a}_m^T \cdot \mathbf{x} \end{bmatrix}$$

  2. Writing a matrix-vector multiplication as linear combination of the columns of $\mathbf{A}$:

  $$\mathbf{Ax} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots x_n\mathbf{a}_n = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

- Matrices as operator:

  - Rotation (counterclockwise): $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
  - Scale (factor a in x-direction and b in y-direction): $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
  - Reflection (both x and y axis): $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
  - Transition (this is **not** linear): $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$

Vector and matrix norm: A vector norm is a function $\|\mathbf{u}\| : V \to \mathbb{R}_0^+$ (i.e., it takes a vector and returns a nonnegative real number) that satisfies the following properties, where $\mathbf{u}, \mathbf{v} \in V$ and $\alpha \in \mathbb{R}$:

1. Positivity: $\|\mathbf{u}\| \geq 0$
2. Definiteness: $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \vec{0}$
3. Homogeneity: $\|\alpha\mathbf{u}\| = |\alpha| \, \|\mathbf{u}\|$
4. Triangle inequality: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

The p-norm of a vector is defined as $\|\mathbf{w}\|_p = (|\mathbf{w}_1|^p + |\mathbf{w}_2|^p + \cdots + |\mathbf{w}_n|^p)^{\frac{1}{p}} = \left(\sum |\mathbf{w}_i|^p\right)^{\frac{1}{p}}$. Especially, When $p = 2$ (2-norm), this is called the Euclidean norm and it corresponds to the length of the vector.

The matrix norm has the same properties that a vector norm has. Here are some common matrix norms:

- **Induced (or operator) norm:** $\|\mathbf{A}\| := \max\limits_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|$. An induced matrix norm is a particular type of a general matrix norm. Induced matrix norms tell us the maximum amplification of the norm of any vector when multiplied by the matrix. Note that the definition above is equivalent to $\|\mathbf{A}\| = \max\limits_{\|\mathbf{x}\| \neq 0} \dfrac{\|\mathbf{A}\mathbf{x}\|}{\|x\|}$. Induced matrix norms also satisfy the submultiplicative conditions: $\|AB\| \leq \|A\| \, \|B\|$.
- **Frobenius norm:** Square root of the sum of every squared element of the matrix. $\|\mathbf{A}\|_F = \sqrt{\sum\limits_{i,j} a_{ij}^2}$.
- ***p*-norm of a matrix:** The matrix p-norm is induced by the p-norm of a vector. It is $\|\mathbf{A}\|_p := \max\limits_{\|\mathbf{x}\|_p = 1} \|\mathbf{A}\mathbf{x}\|_p$.
  - 1-norm: The maximum absolute column sum of the matrix. $\|\mathbf{A}\|_1 = \max\limits_j \sum\limits_{i=1}^n |a_{ij}|$.
  - 2-norm: The maximum singular value of the matrix. $\|A\|_2 = \max\limits_k \sigma_k$
  - ∞-norm: The maximum absolute row sum of the matrix. $\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$.

Python for vectors and matrices:

```python
import numpy as np
import numpy.linalg as la

a = np.array([[8, 9]])
b = np.array([[1, 2, 3], [4, 5, 6]])

# The most important operation you'll do is matrix multiplication
# we can do this easily in 2 ways (both are the same)
c = np.dot(a, b)
c = a @ b

A = np.array([1, 2], [3, 4])
b = np.array([5, 6])

matrix_inv = la.inv(A)          # inverse matrix

vec_norm = la.norm(A)           # calculate norm of A
vec_norm_3 = la.norm(A, 3)      # calculate 3-norm of A
```