Important Mathematics for Machine Learning

Type	Data science masterclass

I. Vectors & Matrices

Scalars, Vectors, Matrices, and Tensors

Scalars

- A scalar is a single numerical value.
- It has only magnitude and no direction.
- Example: 5, -3.2, 100, π

Vectors

- A vector is an ordered list of numbers.
- Represented as a column or row:

  Column Vector:

  $$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$

  Row Vector:

  $$\mathbf{w} = \begin{bmatrix} 4 & -1 & 6 \end{bmatrix}$$

- **Operations on Vectors:**
  - **Addition/Subtraction:** Component-wise operation.
  - **Scalar Multiplication:** Each element is multiplied by a scalar.
  - **Dot Product:**

$$\mathbf{a} \cdot \mathbf{b} = \sum a_i b_i$$

- **Cross Product:** Only defined in 3D, results in another vector perpendicular to both.

Matrices

- A matrix is a two-dimensional array of numbers arranged in rows and columns.
- Denoted as:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

- **Matrix Properties:**
  - **Square Matrix:** Equal rows and columns.
  - **Diagonal Matrix:** Non-diagonal elements are zero.
  - **Symmetric Matrix:**

    $$A^T = A$$

Tensors

- A tensor is a generalization of scalars, vectors, and matrices to higher dimensions.
- Example: A 3D tensor has shape (3,3,3), like a cube of numbers.

II. Matrix Operations

Matrix Addition & Subtraction

- Only possible if both matrices have the same dimensions.
- Performed element-wise:

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Matrix Multiplication

Scalar Multiplication

- Multiply each element by a scalar :

  λ

  $$\lambda A = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{bmatrix}$$

Matrix-Matrix Multiplication

- If A is m×n and B is n×p , the result is m×p.
- Formula:

  $(AB)ij = \sum_{k=1}^{n} A_{ik}B_{kj}$

Transpose of a Matrix

- Swap rows and columns:

  $$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

---

3. Special Matrices

Identity Matrix (I)

- A square matrix with 1s on the diagonal and 0s elsewhere.
- Example:

  $$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- Property: ,

  AI =A

  IA = A

Inverse Matrix (A^{-1})

- A matrix A is invertible if .

  $AA^{-1} = I$

- For a 2×2 matrix:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where $\det(A) = ad - bc$

- A matrix is non-invertible (singular) if .

  $det(A) = 0$

III. Linear Algebra: Vector Spaces & Transformations

1. Vector Spaces

A **vector space** is a set of vectors where vector addition and scalar multiplication are defined and satisfy the following **axioms**:

1. **Closure under Addition**: If u,v are in V, then u+v is also in V.
2. **Closure under Scalar Multiplication**: If v is in V and c is a scalar, then cv is also in V.
3. **Commutativity**: .

   u+v = v + u
4. **Associativity of Addition**: .

   (u+v)+w = u + (v + w)
5. **Existence of Zero Vector**: There exists a vector 0 such that v+0=v for all v∈V.
6. **Existence of Additive Inverse**: For every v, there exists -v such that v+(−v)=0.
7. **Associativity of Scalar Multiplication**: a(bv)=(ab)v
8. **Distributive Property (Scalars and Vectors)**: a(u+v) = a u + a v

9. **Distributive Property (Scalars)**: (a+b)v = a v + b v
10. **Multiplicative Identity**: for all v: 1v= v

Examples of vector spaces:

- The set of all n-dimensional column vectors $\mathbb{R}^n$.
- The set of all continuous functions f(x).
- The set of all polynomials of degree n.

---

2. Basis & Dimension

Basis of a Vector Space

- A **basis** of a vector space is a **set of linearly independent vectors** that **span** the entire space.
- If $\{v_1, v_2, \ldots, v_n\}$ is a basis for , then every vector in can be written as a **linear combination** of basis vectors:

  $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$
- A basis must satisfy two conditions:
  1. **Linear Independence**: No vector in the basis can be expressed as a linear combination of others.
  2. **Spanning**: Any vector in the space can be written as a combination of the basis vectors.

Dimension of a Vector Space

- The **dimension** of a vector space, denoted as **V**, is the number of vectors in any basis for V.
- Examples:
  - $\mathbb{R}^2$ has a standard basis $\{(1,0),(0,1)\}$ and **dimension = 2**.
  - The space of all polynomials of degree at most has dimension n+1.
  - The **zero vector space** {0} has **dimension = 0**.

---

3. Orthogonality & Dot Product

Orthogonality

- Two vectors u,v in R^n are **orthogonal** if their dot product is **zero**:

  $$u \cdot v = 0$$

- Example: In $\mathbb{R}^2$, vectors (1,2) and (−2,1) are orthogonal because:

  $$(1,2) \cdot (-2,1) = (1 \times -2) + (2 \times 1) = -2 + 2 = 0$$

Dot Product

- The **dot product** of two vectors u,v in $\mathbb{R}^n$ is:

  $$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- It measures the **projection** of one vector onto another and relates to the angle θ between them:

  $$u \cdot v = ||u||||v|| \cos\theta$$

- Properties:
  - **Commutative**:

    $$u \cdot v = v \cdot u$$

  - **Distributive**:

    $$u \cdot (v + w) = u \cdot v + u \cdot w$$

  - **Scalar Multiplication**:

    $$(cu) \cdot v = c(u \cdot v)$$

---

4. Eigenvalues & Eigenvectors

Definition

- Given a square matrix A, an **eigenvector** v is a nonzero vector that satisfies:

  $$Av = \lambda v$$

where λ is a
scalar called an **eigenvalue**.

- This means that applying A to v only **scales** it without changing its direction.

Finding Eigenvalues

- Eigenvalues are solutions to the **characteristic equation**:

  $$\det(A-\lambda I)=0$$

- Example: If

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

  then the characteristic equation is:

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

which simplifies to
$(2-\lambda)^2-1=0$, giving eigenvalues **λ=3,1.**

Finding Eigenvectors

- For each eigenvalue λ, solve (A−λI)v=0 to find the corresponding eigenvectors.

---

5. Singular Value Decomposition (SVD)

Definition

- Any m×n matrix A can be decomposed as:
- $A=U\Sigma V^T$
  where:
  - U is an m×m **orthogonal matrix**.

- Σ is an m×n **diagonal matrix** with **singular values**.
- V is an n×n **orthogonal matrix**.

Why SVD is Important?

- Used in **dimensionality reduction** (e.g., PCA).
- Helps in solving **ill-conditioned linear systems**.
- Provides insights into the **geometry of transformations**.

Steps to Compute SVD

1. Compute **eigenvalues and eigenvectors** of $A^T A$ (columns of V).
2. Compute **eigenvalues and eigenvectors** of AA^T (columns of U).
3. Compute **singular values** as square roots of eigenvalues of $A^T A$.

Example

For

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

1. Compute $A^T A$, find eigenvalues/eigenvectors.
2. Compute AA^T, find eigenvalues/eigenvectors.
3. Construct U, Σ, V^T.
