Vector Geometry

1. Norm and Inner Product

- **Norm (||x||):** Measures the "length" of a vector.

  For a vector `x = [x₁, x₂, ..., xₙ]ᵗ`,

  $||x|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$

- **Inner Product (Dot Product):** Measures how much two vectors "point in the same direction".

  $\langle x, y \rangle = x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$

Important properties:

- $\langle x, x \rangle = ||x||^2$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle Ax, y \rangle = \langle x, A^T y \rangle$

---

2. Angle and Orthogonality

- Using the cosine rule:

  $||y - x||^2 = ||x||^2 + ||y||^2 - 2||x||\,||y|| \cos(\theta)$

- **Orthogonal Vectors:**

  $\langle x, y \rangle = 0 \Rightarrow x \perp y$

- **Orthonormal Set:**

  Vectors are orthogonal and of unit length:

  $\langle q_i, q_j \rangle = \delta_{ij}$

  (Kronecker delta = 1 if i = j, 0 otherwise)

What Are Orthonormal Vectors?

A set of **orthonormal vectors** is a set of vectors that are both:

1. **Orthogonal** to each other (i.e., at 90° to each other, or their dot product is 0), and
2. **Normalized** (i.e., each vector has a length or norm of 1).

In More Precise Terms:

Let's say we have two vectors $\vec{u}$ and $\vec{v}$. They are **orthonormal** if:

- $\vec{u} \cdot \vec{v} = 0$ (orthogonality)
- $\|\vec{u}\| = 1$ and $\|\vec{v}\| = 1$ (normalization)

---

3. Gram-Schmidt Process (to get an orthonormal set)

Why use Gram-Schmidt?

If you have a bunch of vectors that are linearly independent, but **not orthogonal or normalized**, Gram-Schmidt gives you a systematic way to make them **orthonormal**, which makes things like solving systems, projections, and factorizations much easier.

Step-by-Step Gram-Schmidt Process

Given a linearly independent set $\{a_1, a_2, ..., a_n\}$:

- Step 1: $q_1 = \frac{a_1}{||a_1||}$
- Step 2: Subtract projection of a_2 on q_1, normalize:

$$v_2 = a_2 - \langle q_1, a_2 \rangle q_1, \quad q_2 = \frac{v_2}{||v_2||}$$

- Repeat for others:

$$v_i = a_i - \sum_{k=1}^{i-1} \langle q_k, a_i \rangle q_k, \quad q_i = \frac{v_i}{||v_i||}$$

---

4. Orthogonal and QR Matrices

- **Orthogonal Matrix Q:** $Q^T Q = I \Rightarrow Q^{-1} = Q^T$

QR Factorization (or **QR decomposition**) is a method in linear algebra where you take a **matrix A** and break it into the product of two matrices:

A = QR

Where:

- **Q** is an **orthogonal (or orthonormal)** matrix — its columns are orthonormal vectors (i.e., they're at right angles to each other and have length 1).
- **R** is an **upper triangular matrix** — all entries below the diagonal are zero.

🔍 Why is it useful?

QR factorization is super useful in:

- **Solving linear systems**, especially least squares problems.
- **Numerical stability** in computations (e.g., better than solving $A^TAx = A^Tb$ directly).
- **Eigenvalue algorithms**
- Under the hood of **machine learning** and **optimization** techniques (e.g., in linear regression).

📐 Geometric Intuition

Think of matrix A as having columns that are not orthogonal. QR factorization says:

> "Let's find a new basis (the Q matrix) that is orthonormal but still spans the same space as the columns of A, and then express A in terms of those basis vectors using R."

That's where **Gram-Schmidt** comes in — it's one way to find the orthonormal vectors (columns of Q).

🧠 Simple Example (Conceptually)

Suppose:

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

You could apply the **Gram-Schmidt process** to get orthonormal vectors:

$$Q = \left[\text{orthonormal columns}\right], \quad R = \text{upper triangular}$$

Then A = QR reconstructs the original matrix.

✅ Quick Facts:

- $Q^T Q = I$, meaning Q is orthonormal.
- QR is often used **instead of matrix inversion** in practice.
- Faster and more stable than solving equations using A^{-1}.

---

5. Least Squares Fit

Problem: Given data points (x_i, y_i), find the best-fit line y=ax+b

- Form as matrix: Ax=b
- Find x that minimizes ||Ax - b||
- Solution:

  $$x = (A^T A)^{-1} A^T b$$
- These are called **normal equations**.

---

6. Projection Matrix

- For matrix A, the projection matrix is:

  $$P = A(A^T A)^{-1} A^T$$
- **Properties:**
  - $P^2 = P$
  - P is symmetric
  - Projects any vector onto the column space of A

---

7. Eigenvalues and Eigenvectors

- Problem: $Ax = \lambda x$
  - λ: eigenvalue

- x: eigenvector

- To find: solve (Determinant of A - lambda I) $\det(A - \lambda I) = 0$→ characteristic polynomial
- Properties:
  - |A|=product of eigenvalues
  - (The **Trace** of A) $\text{tr}(A) = \sum \text{eigenvalues}$
  - If eigenvalues are distinct, eigenvectors are linearly independent.

---

8. Diagonalization

Diagonalization is the process of transforming a matrix into a **diagonal matrix** — a matrix with nonzero entries only on the diagonal — using its **eigenvalues** and **eigenvectors**.

If A has n independent eigenvectors:

- Define $S = [x_1, x_2, ..., x_n]$ (matrix of eigenvectors)
- $S^{-1}AS = \Lambda$ (diagonal matrix with eigenvalues)

  where:

  - A = original matrix
  - S = matrix whose **columns are eigenvectors** of A
  - Λ = **diagonal matrix** whose **diagonal entries are eigenvalues** of A

This simplifies computations like A^n using:

$A^n = S\Lambda^n S^{-1}$

---

9. Quadratic Forms

- General form:

  $Q = x^T Ax$

- If A is symmetric, we can simplify Q by diagonalizing

  $Q = x^T Ax = x^{*\top} \Lambda x^* \quad (\text{change of variables: } x = Mx^*)$

- Useful for identifying conics (ellipse, hyperbola, etc.)

Matrix Representation of Quadratic Forms

You can **rewrite** any quadratic form using matrices:

$$Q(x) = x^T A x$$

where:

- $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- A is a **symmetric matrix**.

Example:

$$Q(x) = ax_1^2 + bx_1x_2 + cx_2^2 \quad \text{can be written as} \quad x^T \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} x$$

Diagonalizing a Quadratic Form

Sometimes you want to **simplify** the quadratic form **— remove the cross terms**.

You can do that by **diagonalizing the matrix A**:

$$M^T A M = \Lambda$$

where:

- M = matrix of eigenvectors (orthogonal matrix if A is symmetric)
- Λ = diagonal matrix (eigenvalues on diagonal)

After a **change of variables**:

$$x = Mx^*$$

the quadratic form becomes:

$$Q(x) = {x^*}^T \Lambda x^*$$

which is **purely sums of squares**:

$$Q(x^*) = \lambda_1 (x_1^*)^2 + \lambda_2 (x_2^*)^2 + \cdots$$
