**ML Math Core — Formal Definitions & Examples (C#)**

# Notation

Vectors are bold lowercase (a, b) with elements ai. Matrices are uppercase (A, B) with entries Aij. Code uses 0-based indices; math uses 1-based.

# Vector Operations

## Addition

**(a + b)**i = ai + bi

Definition: The i-th element equals the sum of the i-th elements of a and b.

Example: a = (1, 2, 3), b = (4, 5, 6) ⇒ a + b = (5, 7, 9).

## Subtraction

**(a − b)**i = ai − bi

Definition: The i-th element equals the difference of the i-th elements of a and b.

Example: a = (1, 2, 3), b = (4, 5, 6) ⇒ a − b = (−3, −3, −3).

## Scalar Multiplication

**(s·a)**i = s · ai

Definition: Multiply each element of a by the scalar s.

Example: s = 2, a = (1, 2, 3) ⇒ 2·a = (2, 4, 6).

## Dot Product

a · b = ∑i ai bi

Definition: Multiply corresponding elements of a and b and sum to produce a scalar.

Example: a = (1, 2, 3), b = (4, 5, 6) ⇒ a · b = 32.

## L2 Norm (Euclidean)

**||a||**2 = √(∑i ai²)

Definition: Square each element, sum them, and take the square root.

Example: a = (3, 4) ⇒ ||a||₂ = 5.

## L1 Norm (Manhattan)

**||a||**1 = ∑i |ai|

Definition: Sum of the absolute values of the elements.

Example: a = (3, −4, 5) ⇒ ||a||₁ = 12.

## Normalize

Normalize(a, ε) = a / max(**||a||**2, ε)

Definition: Divide each element by the L2 norm (or ε if the norm is too small).

Example: a = (3, 4), ε = 10⁻¹² ⇒ Normalize(a) = (0.6, 0.8).

## Cosine Similarity

cos(a,b) = (a · b) / (**||a||**2 **||b||**2)

Definition: The dot product divided by the product of the L2 norms.

Example: a = (1, 0), b = (0, 1) ⇒ cos(a,b) = 0.

# Matrix Operations

## Addition

**(A + B)**ij = Aij + Bij

Definition: Add corresponding elements (shapes must match).

Example: A = [[1,2],[3,4]], B = [[5,6],[7,8]] ⇒ A + B = [[6,8],[10,12]].

## Subtraction

**(A − B)**ij = Aij − Bij

Definition: Subtract corresponding elements (shapes must match).

Example: A = [[1,2],[3,4]], B = [[5,6],[7,8]] ⇒ A − B = [[−4,−4],[−4,−4]].

## Scalar Multiplication

Definition: Multiply every element of A by the scalar s.

Example: s = 2, A = [[1,2],[3,4]] ⇒ 2·A = [[2,4],[6,8]].

## Transpose

**(Aᵀ)**ij = Aji

Definition: Swap rows and columns.

Example: A = [[1,2,3],[4,5,6]] ⇒ Aᵀ = [[1,4],[2,5],[3,6]].

## Matrix–Vector Multiplication (MatVec)

**(A x)**i = ∑j Aij xj

Definition: Each output entry is the dot product of row i of A with x (requires x length equals columns).

Example: A = [[1,2,3],[4,5,6]], x = [7,8,9] ⇒ A x = [50, 122].

## Matrix–Matrix Multiplication (MatMul)

**(A B)**ik = ∑j Aij Bjk

Definition: The (i,k) entry of A B is the sum over j of Aij times Bjk (inner dimensions must match).

Example: A = [[1,2,3],[4,5,6]] (2×3), B = [[7,8],[9,10],[11,12]] (3×2) ⇒ A B = [[58,64],[139,154]].

## Identity Matrix

Definition: In has ones on the main diagonal and zeros elsewhere; acts as multiplicative identity.

Examples: I₂ = [[1,0],[0,1]]. For A (m×n): Im A = A and A In = A (when dimensions match).

## Trace

Definition: tr(A) = sum of diagonal elements (square matrices). Example: A = [[1,2],[3,4]] ⇒ tr(A) = 5.

## Frobenius Norm

Definition: ||A||F = √(sum of squares of all entries). Example: A = [[1,2],[3,4]] ⇒ ||A||F = √30.

# Descriptive Statistics

## Mean

Definition: μ = (1/n) ∑ xi. Example: x = [1,2,3] ⇒ μ = 2.

## Variance

Definition: Population σ² = (1/n) ∑ (xi − μ)²; Sample (unbiased) s² = (1/(n−1)) ∑ (xi − μ)².

Example: x = [1,2,3], μ = 2 ⇒ σ² = 2/3, s² = 1.

## Standard Deviation

Definition: σ = √(variance). Example: for x = [1,2,3], population σ = √(2/3), sample s = 1.