动手学深度学习 v2

李沐・Amazon资深首席科学家

线性代数

Scalars

- **Simple operations**

$$c = a + b$$

$$c = a \cdot b$$

$$c = \sin a$$

- **Length**

$$|a| = \begin{cases} a & \text{if} a > 0 \\ -a & \text{otherwise} \end{cases}$$

$$|a + b| \leq |a| + |b|$$

$$|a \cdot b| = |a| \cdot |b|$$

Vectors

- **Simple operations**

$$c = a + b \quad \text{where } c_i = a_i + b_i$$
$$c = \alpha \cdot b \quad \text{where } c_i = \alpha b_i$$
$$c = \sin a \quad \text{where } c_i = \sin a_i$$

Definition of a vector space

- **Length**

$$\|a\|_2 = \left[\sum_{i=1}^{m} a_i^2\right]^{\frac{1}{2}}$$
$$\|a\| \geq 0 \text{ for all } a$$
$$\|a + b\| \leq \|a\| + \|b\|$$
$$\|a \cdot b\| = |a| \cdot \|b\|$$

Definition of a norm

Vectors

$$c = a + b$$

$$c = \alpha \cdot b$$

Mathematician's **'parallel for all do'**

Vectors

- **Dot product**

$$a^\top b = \sum_i a_i b_i$$

- **Orthogonality**

$$a^\top b = \sum_i a_i b_i = 0$$

(e.g. if we have two vectors that are orthogonal with a third, their linear combination is it, too)

Matrices

- **Simple operations**

$$C = A + B \qquad \text{where } C_{ij} = A_{ij} + B_{ij}$$
$$C = \alpha \cdot B \qquad \text{where } C_{ij} = \alpha B_{ij}$$
$$C = \sin A \qquad \text{where } C_{ij} = \sin A_{ij}$$

- **Functional Analysis 101**

  vector = function, matrix = linear operator

  most theorems work sort-of in infinite dimensional spaces

Matrices

- **Multiplications (matrix vector)**

$$c = Ab \text{ where } c_i = \sum_j A_{ij} b_j$$

Matrices

https://courses.d2l.ai/zh-v2

Matrices

- **Multiplications (matrix matrix)**

$$C = AB \text{ where } C_{ik} = \sum_j A_{ij} B_{jk}$$

https://courses.d2l.ai/zh-v2

Matrices

- **Norms**

$$c = A \cdot b \text{ hence } \|c\| \leq \|A\| \cdot \|b\|$$

  - Choices depending on how to measure length of b and c

- **Popular norms**
  - Frobenius norm

$$\|A\|_{\text{Frob}} = \left[\sum_{ij} A_{ij}^2 \right]^{\frac{1}{2}}$$

  - H-infinity norm (or often just matrix norm) as smallest value for which condition holds for Euclidean norms on vectors.

Matrices

- **Eigenvectors and eigenvalue**
  - Vectors that aren’t changed by the matrix

$$Ax = \lambda x$$

  - For symmetric matrices we can always find this

https://courses.d2l.ai/zh-v2

Special Matrices

- **Symmetric, antisymmetric**

$$A_{ij} = A_{ji} \text{ and } A_{ij} = -A_{ji}$$

- **Positive definite**

$$\|x\|^2 = x^\top x \geq 0 \text{ generalizes to } x^\top A x \geq 0$$

(all positive eigenvalues)

https://courses.d2l.ai/zh-v2

Special Matrices

- **Orthogonal Matrices**
  - All rows of the matrix are orthogonal to each other
  - All rows of the matrix have unit length

$$U \text{ with } \sum_j U_{ij} U_{kj} = \delta_{ik}$$

Show that $U^\top U = \mathbf{1}$

  - Rewrite in matrix form

$$UU^\top = \mathbf{1}$$

Show that P is orthogonal

- **Permutation Matrices**

$$P \text{ where } P_{ij} = 1 \text{ if and only if } j = \pi(i)$$

GPUs love matrices and vectors
(they have many processors)