Linear Algebra (Review)

Volker Tresp

Winter 2024-2025

Vectors

- k, M, N are scalars
- A order-1 array $\mathbf{c}$ is a column vector. Thus with two dimensions,

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

(more precisely, it is a representation of a vector in a specific coordinate system)

- c_i is the i-th component of $\mathbf{c}$

Transposed

- $\mathbf{c}^T = (c_1, c_2)$ is a row vector, the transposed of $\mathbf{c}$

Matrices

- An order-2 array $\mathbf{A}$ is a matrix, e.g.,

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}$$

- We also write: $\mathbf{A} = (a_{1,1}, a_{1,2}, a_{1,3}; a_{2,1}, a_{2,2}, a_{2,3})$; thus the semicolon ";" indicates row separations

- The colon ":" is sometimes used to select rows or columns; examples

$$\mathbf{A}_{:,1} = \begin{pmatrix} a_{1,1} \\ a_{2,1} \end{pmatrix} \qquad \mathbf{A}_{1,:} = (a_{1,1}, a_{1,2}, a_{1,3})$$

Transposed

- If $\mathbf{A}$ is an $N \times M$-dimensional matrix,
  - then the transposed $\mathbf{A}^T$ is an $M \times N$-dimensional matrix
  - the columns (rows) of $\mathbf{A}$ are the rows (columns) of $\mathbf{A}^T$ and vice versa

$$\mathbf{A}^T = \begin{bmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \\ a_{1,3} & a_{2,3} \end{bmatrix}$$

Addition of Two Vectors

- Let $\mathbf{c} = \mathbf{a} + \mathbf{d}$
- Then $c_j = a_j + d_j$

Multiplication of a Vector with a Scalar

- $\mathbf{c} = k\mathbf{a}$ is a vector with $c_j = ka_j$
- $\mathbf{C} = k\mathbf{A}$ is a matrix of the dimensionality of $\mathbf{A}$, with $c_{i,j} = ka_{i,j}$

Scalar Product of Two Vectors

- The **scalar product** (also called **dot product**) is defines as

$$\mathbf{a} \cdot \mathbf{c} = \mathbf{a}^T \mathbf{c} = \sum_{j=1}^{M} a_j c_j$$

  and is a scalar

- Remark: The **inner product** $\langle x, y \rangle$ is closely related and generalizes the dot product to abstract vector spaces

Scalar Product of Two Vectors (Cont'd)

- $\|\mathbf{a}\|$ is the length or Euclidean norm of the vector; then,

$$\|\mathbf{a}\|^2 = \mathbf{a}^T\mathbf{a} = \sum_{j=1}^{M} a_j^2$$

- $\|\mathbf{a}-\mathbf{b}\|$ is the Euclidean distance between both vectors; then,

$$\|\mathbf{a}-\mathbf{b}\|^2 = (\mathbf{a}-\mathbf{b})^T(\mathbf{a}-\mathbf{b}) = \sum_{j=1}^{M} (a_j - b_j)^2$$

Matrix-Vector Product

- A matrix consists of many row vectors. So a product of a matrix with a column vector consists of many scalar products of vectors
- If $\mathbf{A}$ is an $N \times M$-dimensional matrix and $\mathbf{c}$ is an M-dimensional column vector
- Then $\mathbf{d} = \mathbf{A}\mathbf{c}$ is an N-dimensional column vector $\mathbf{d}$ with

$$d_i = \sum_{j=1}^{M} a_{i,j} c_j$$

Matrix-Matrix Product

- A matrix also consists of many column vectors. So a product of matrix with a matrix consists of many matrix-vector products
- If $\mathbf{A}$ is an $N \times M$-dimensional matrix and $\mathbf{C}$ an $M \times K$-dimensional matrix
- Then $\mathbf{D} = \mathbf{AC}$ is an $N \times K$-dimensional matrix with

$$d_{i,k} = \sum_{j=1}^{M} a_{i,j} c_{j,k}$$

Multiplication of a Row-Vector with a Matrix

- **Multiplication of a row vector with a matrix is a row vector**. If A is a $N \times M$-dimensional matrix and $\mathbf{d}$ a N-dimensional vector and if

$$\mathbf{c}^T = \mathbf{d}^T A$$

Then $\mathbf{c}$ is a M-dimensional vector with $c_j = \sum_{i=1}^{N} d_i a_{i,j}$

Outer Product

- Special case: **Multiplication of a column vector with a row vector is a matrix**.

  Let $\mathbf{d}$ be a N-dimensional vector and $\mathbf{c}$ be a M-dimensional vector, then

  $$\mathbf{A} = \mathbf{d}\mathbf{c}^T$$

  is an $N \times M$ matrix with $a_{i,j} = d_i c_j$

  Example:

  $$\begin{bmatrix} d_1c_1 & d_1c_2 & d_1c_3 \\ d_2c_1 & d_2c_2 & d_2c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}$$

Matrix Transposed

- The transposed $\mathbf{A}^T$ changes rows and columns
- We have

$$\left(\mathbf{A}^T\right)^T = \mathbf{A}$$

$$(\mathbf{AC})^T = \mathbf{C}^T\mathbf{A}^T$$

Unit Matrix

-

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ldots & \\ 0 & \ldots & 0 & 1 \end{pmatrix}$$

Diagonal Matrix

- $N \times N$ diagonal matrix:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 \\ & & \dots & \\ 0 & \dots & 0 & a_{N,N} \end{pmatrix}$$

Matrix Inverse

- Let $\mathbf{A}$ be an $N \times N$ square matrix
- If there is a unique inverse matrix $\mathbf{A}^{-1}$, then we have

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

- If the corresponding inverse exist, $(\mathbf{AC})^{-1} = \mathbf{C}^{-1}\mathbf{A}^{-1}$
- and $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$

Orthogonal Matrices

- **Orthogonal Matrix** (more precisely: orthonormal matrix): $\mathbf{R}$ is a (quadratic) orthogonal matrix, if all columns are orthonormal. It follows (non-trivially) that all rows are orthonormal as well and

$$\mathbf{R}^T\mathbf{R} = \mathbf{I} \quad \mathbf{R}\mathbf{R}^T = \mathbf{I} \quad \mathbf{R}^{-1} = \mathbf{R}^T \tag{1}$$

A Function as a Vector in Infinite Dimensions

- Inner product between two functions

$$\mathbf{a}^T\mathbf{b} = \sum_{j=1}^{M} a_j b_j \rightarrow \int_{x_1}^{x_2} a(x)b(x)dx$$

- Distance between two functions

$$\|\mathbf{a} - \mathbf{b}\|^T = \sum_{j=1}^{M} (a_j - b_j)^2 \rightarrow \int_{x_1}^{x_2} (a(x) - b(x))^2 dx$$

- Average squared distance between two functions

$$\|\mathbf{a} - \mathbf{b}\|^T / M = 1/M \sum_{j=1}^{M} (a_j - b_j)^2 \rightarrow \frac{1}{|x_2 - x_1|} \int_{x_1}^{x_2} (a(x) - b(x))^2 dx$$

Functions (cont'd)

- Expected squared distance (averaged over a random input) between two functions:

$$\sum_{j=1}^{M} P(j)(a_j - b_j)^2 \rightarrow \int P(x)(a(x) - b(x))^2 dx$$

where $P(j)$ is a probability and $P(x)$ is a probability density