Cheat sheet linear algebra

Michael Franke

Notation & conventions

- vectors are written in bold lower-case $\mathbf{x}$
- matrices are written in bold upper-case $\mathbf{A}$
- use square brackets for vectors and matrices, e.g., $\mathbf{x} = [x_1, \dots, x_n]$ or

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

  - an $n \times m$ matrix $\mathbf{A}$ can be written as $A = [a_{ij}] \in \mathbb{R}^{n \times m}$, or shorter as: $A \in \mathbb{R}^{n \times m}$
- the **transpose** of a matrix or vector is written as $\mathbf{A}^\top$ and $\mathbf{x}^\top$
- we use **row-first**
  - first index of a 2-D matrix is the row
  - indices for vectors give the *row*, so that vectors are column vectors
    - by convention, if the context is clear, we can write $\mathbf{x} = [x_1, \dots, x_n]$ to denote a column vector
    - it would be more precise to write $\mathbf{x} = [x_1, \dots, x_n]^T$ or

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Vector & matrix operations

- if $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{B} \in \mathbb{R}^{m \times o}$, the **matrix product** $AB = C \in \mathbb{R}^{n \times o}$ is defined via:

$$c_{ik} = \sum_j a_{ij} b_{jk}$$

- the **dot product** between equal-length vectors $\mathbf{x}$ and $\mathbf{y}$ is:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\intercal \mathbf{y} = \sum_i x_i y_i$$

- **vector concatenation** of $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^m$ is written as $\mathbf{x} \oplus \mathbf{y} = [x_1, \ldots, x_n, y_1, \ldots, y_m]$
  - We can generalize this notation to many vectors: $\bigoplus_{i=1}^{n} \mathbf{x}_i = \mathbf{x}_1 \oplus \cdots \oplus \mathbf{x}_n$

Interpretation

- think of matrix $\mathbf{A}$ with dimensions (n, m) as a **linear mapping** $f_{\mathbf{A}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ from vectors of length m to vectors for length n, so that with $\mathbf{x} = [x_1, \ldots, x_m]$:

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

Vector similarity

- the **dot-product similarity** of vectors $\mathbf{x}$ and $\mathbf{y}$ is the dot product between the vectors:

$$\text{DotSim}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$$

- the **cosine similarity** of vectors $\mathbf{x}$ and $\mathbf{y}$ is the dot-product similarity adjusted for vector length:

$$\text{CosineSim}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||\ ||\mathbf{y}||}$$

  where $||\mathbf{x}|| = \sqrt{\sum_i x_i^2}$ is a measure of the length of vector $\mathbf{x}$