About this chapter

• Not a comprehensive survey of all of linear algebra

• Focused on the subset most relevant to deep learning

• Larger subset: e.g., *Linear Algebra* by Georgi Shilov
 Scalars

• A scalar is a single number

• Integers, real numbers, rational numbers, etc.

• We denote it with italic font:

\[ a, n, x \]
Vectors

- A vector is a 1-D array of numbers:
  \[
  \mathbf{x} = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
  \end{bmatrix}.
  \]  
  \hspace{1cm} (2.1)

- Can be real, binary, integer, etc.

- Example notation for type and size:
  \[ \mathbb{R}^n \]
Matrices

• A matrix is a 2-D array of numbers:

\[
\begin{bmatrix}
A_{1,1} & A_{1,2} \\
A_{2,1} & A_{2,2}
\end{bmatrix}
\]  

(2.2)

• Example notation for type and shape:

\[ A \in \mathbb{R}^{m \times n} \]
Tensors

• A tensor is an array of numbers, that may have
  • zero dimensions, and be a scalar
  • one dimension, and be a vector
  • two dimensions, and be a matrix
  • or more dimensions.
Matrix Transpose

\[(A^\top)_{i,j} = A_{j,i}.\]  \hspace{1cm} (2.3)

\[
\begin{pmatrix}
A_{1,1} & A_{1,2} \\
A_{2,1} & A_{2,2} \\
A_{3,1} & A_{3,2}
\end{pmatrix}
\Rightarrow
\begin{pmatrix}
A_{1,1} & A_{2,1} & A_{3,1} \\
A_{1,2} & A_{2,2} & A_{3,2}
\end{pmatrix}
\]

Figure 2.1: The transpose of the matrix can be thought of as a mirror image across the main diagonal.

\[(AB)^\top = B^\top A^\top.\]  \hspace{1cm} (2.9)

(Goodfellow 2016)
Matrix (Dot) Product

\[ C = AB. \]  \hspace{2cm} (2.4)

\[ C_{i,j} = \sum_k A_{i,k} B_{k,j}. \]  \hspace{2cm} (2.5)

(Goodfellow 2016)
Identity Matrix

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

Figure 2.2: Example identity matrix: This is \(I_3\).

\[\forall \mathbf{x} \in \mathbb{R}^n, I_n \mathbf{x} = \mathbf{x}.\]  \hspace{1cm} (2.20)
Systems of Equations

\[ Ax = b \] \hspace{1cm} (2.11)

expands to

\[ A_1 : x = b_1 \] \hspace{1cm} (2.12)
\[ A_2 : x = b_2 \] \hspace{1cm} (2.13)
\[ \ldots \] \hspace{1cm} (2.14)
\[ A_m : x = b_m \] \hspace{1cm} (2.15)
Solving Systems of Equations

• A linear system of equations can have:
  • No solution
  • Many solutions
  • Exactly one solution: this means multiplication by the matrix is an invertible function
Matrix Inversion

- Matrix inverse:
  \[ A^{-1}A = I_n. \] (2.21)

- Solving a system using an inverse:
  \[ Ax = b \] (2.22)
  \[ A^{-1}Ax = A^{-1}b \] (2.23)
  \[ I_nx = A^{-1}b \] (2.24)

- Numerically unstable, but useful for abstract analysis
Invertibility

• Matrix can’t be inverted if…
  • More rows than columns
  • More columns than rows
  • Redundant rows/columns ("linearly dependent", "low rank")
Norms

- Functions that measure how “large” a vector is
- Similar to a distance between zero and the point represented by the vector

\[
f(x) = 0 \Rightarrow x = 0
\]

\[
f(x + y) \leq f(x) + f(y) \text{ (the triangle inequality)}
\]

\[
\forall \alpha \in \mathbb{R}, f(\alpha x) = |\alpha| f(x)
\]
Norms

- **$L^p$ norm**

\[
\|x\|_p = \left( \sum_{i} x_i^p \right)^{\frac{1}{p}}
\]

- Most popular norm: L2 norm, $p=2$

- L1 norm, $p=1$: $\|x\|_1 = \sum_{i} |x_i|$.

- Max norm, infinite $p$: $\|x\|_\infty = \max_{i} |x_i|$.
Special Matrices and Vectors

- Unit vector:
  \[ \|x\|_2 = 1. \]  
  \[ (2.36) \]

- Symmetric Matrix:
  \[ A = A^\top. \]  
  \[ (2.35) \]

- Orthogonal matrix:
  \[ A^\top A = AA^\top = I. \]
  \[ A^{-1} = A^\top \]  
  \[ (2.37) \]
Trace

\[ \text{Tr}(A) = \sum_i A_{i,i} . \quad (2.48) \]

\[ \text{Tr}(ABC) = \text{Tr}(CAB) = \text{Tr}(BCA) \quad (2.51) \]
Learning linear algebra

• Do a lot of practice problems

• Start out with lots of summation signs and indexing into individual entries

• Eventually you will be able to mostly use matrix and vector product notation quickly and easily