Common Background(?)

Here are some background facts you may already know, and will want to be familiar with as we enter Chapter 2. (Some of these, at least, are discussed in Chapter 1.)

- $\mathbb{R}$ is a way of referring to the set of *real* numbers. $\mathbb{R}^2$ is the collection of all 2-tuples (x_1, x_2), where both x_1, x_2 are *real* numbers. $\mathbb{R}^3$ is the collection of all 3-tuples (x_1, x_2, x_3), where each of x_1, x_2, x_3 are *real* numbers. In the same way, $\mathbb{R}^n$, with n unspecified, is the collection of all n-tuples $(x_1, x_2, \ldots, x_n)$, where each of x_1, x_2, $\ldots$, x_n are *real* numbers. When we talk of **Euclidean** space, we are talking of $\mathbb{R}^n$ for some n.

- The dot product of vectors $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$ of $\mathbb{R}^2$ is defined as

  $$\mathbf{x} \cdot \mathbf{y} \;=\; (x_1, x_2) \cdot (y_1, y_2) \;=\; x_1 y_1 + x_2 y_2.$$

  The dot product of vectors $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ of $\mathbb{R}^3$ is defined as

  $$\mathbf{x} \cdot \mathbf{y} \;=\; (x_1, x_2, x_3) \cdot (y_1, y_2, y_3) \;=\; x_1 y_1 + x_2 y_2 + x_3 y_3.$$

  We generalize to the dot product of vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ of $\mathbb{R}^n$ is defined as

  $$\mathbf{x} \cdot \mathbf{y} \;=\; (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) \;=\; \sum_{j=1}^{n} x_j y_j \;=\; x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

- An equation in n (implied) variables $x_1, \ldots, x_n$ of the form

  $$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b, \tag{1}$$

  where the $a_1, \ldots, a_n$ and b are constants, is said to be **linear**, or **linear in $x_1, \ldots, x_n$**. When we graph such an equation, and the number of variables is

  - *two*, the graph is a *line*, for example, $3x + 2y = 5,\ \ y = 2x - 1,\ \ x = 3$
  - *three*, the graph is a *plane*. $3x + 2y - z = 5,\ \ 2x + z = 5,\ \ y = 8$
  - *four*, the graph called a *hyperplane*.
  - n, we continue to call the graph a *hyperplane*.

  In all these cases, the graph is said to have **codimension** 1.

- The length of a vector $\mathbf{x} = (x_1, x_2)$ in $\mathbb{R}^2$ is $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}$. The length of a vector $\mathbf{x} = (x_1, x_2, x_3)$ in $\mathbb{R}^3$ is $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Generalizing this, the length of a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is

  $$\|\mathbf{x}\| \;=\; \sqrt{x_1^2 + \cdots + x_n^2} \;=\; \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

- The angle $\theta \in [0, \pi]$ between two nonzero vectors $\mathbf{x}$, $\mathbf{y}$ satisfies the relationship

$$\cos\theta \;=\; \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}. \tag{2}$$

  Multiplying through by the denominator, we get

$$\mathbf{x}\cdot\mathbf{y} \;=\; \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$

  Taking absolute values of both sides, and noting that $|\cos\theta| \leqslant 1$, we obtain the **Cauchy-Schwarz Inequality**

$$|\mathbf{x}\cdot\mathbf{y}| \;\leqslant\; \|\mathbf{x}\|\|\mathbf{y}\|. \tag{3}$$

  The two sides are equal only when $|\cos\theta| = 1$—that is, when the angle between $\mathbf{x}$ and $\mathbf{y}$ is 0 or π, making $\mathbf{y}$ a scalar multiple of $\mathbf{x}$ (or vice-versa).

- The vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x}+\mathbf{y}$ can be viewed as the three sides of a triangle. The triangle inequality says that

$$\|\mathbf{x}+\mathbf{y}\| \;\leqslant\; \|\mathbf{x}\| + \|\mathbf{y}\|.$$

- A single linear equation (1) can be written as a statement about the product of a 1-by-n matrix with the vector $\mathbf{x}$ of unknowns:

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}\mathbf{x} \;=\; b.$$

  A system of m such equations in the same n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\;\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned},$$

  each with a graph that is a hyperplane, can be arranged as the matrix equation

$$\mathbf{A}\mathbf{x} \;=\; \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \;=\; \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

  Any solutions could be thought of as points which lie simultaneously on all m hyperplanes.