Week 2: Dot Product and Hyperplanes in $\mathbb{R}^n$

Definition 2.1. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, write $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$. The *dot product* of $\mathbf{x}$ and $\mathbf{y}$ is defined to be the real number

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + ... + x_n y_n.$$

Theorem 2.2. The dot product has the following properties.
(1) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
(2) $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2 \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Moreover $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
(3) $(c\mathbf{x}) \cdot \mathbf{y} = c(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (c\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.
(4) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.

Proof. We prove (2) and (4). Write $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$ and $\mathbf{z} = (z_1, z_2, ..., z_n)$.
(2) Since $r^2 \geq 0$ for all $r \in \mathbb{R}$

$$\begin{aligned} & \mathbf{x} \cdot \mathbf{x} \\ = \;& x_1^2 + x_2^2 + ... + x_n^2 \\ = \;& \sqrt{x_1^2 + x_2^2 + ... + x_n^2}^{\,2} \\ = \;& \|\mathbf{x}\|^2 \\ \geq \;& 0 \,. \end{aligned}$$

Similarly, if $\mathbf{x} \cdot \mathbf{x} = 0$, then $x_1^2 + x_2^2 + ... + x_n^2 = 0$. But if the sum of nonnegative real numbers is 0, then each of the summands must be 0. That is, $x_i^2 = 0$ for each $i \leq n$ and so $x_i = 0$ for each $i \leq n$. It follows that $(x_1, x_2, ..., x_n) = (0, 0, ..., 0)$; in other words $\mathbf{x} = \mathbf{0}$.
(4) We have

$$\begin{aligned} & \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) \\ = \;& (x_1, x_2, ..., x_n) \cdot ((y_1, y_2, ..., y_n) + (z_1, z_2, ..., z_n)) \\ = \;& (x_1, x_2, ..., x_n) \cdot (y_1 + z_1, y_2 + z_2, ..., y_n + z_n) \quad \text{(Defn of } + \text{ in } \mathbb{R}^n) \\ = \;& x_1(y_1 + z_1) + x_2(y_2 + z_2) + ... + x_n(y_n + z_n) \quad \text{(Defn of } \cdot \text{ in } \mathbb{R}^n) \\ = \;& x_1y_1 + x_1z_1 + x_2y_2 + x_2z_2 + ... + x_ny_n + x_nz_n \quad \text{(Distributive Property in } \mathbb{R}) \\ = \;& x_1y_1 + x_2y_2 ... + x_ny_n + x_1z_1 + x_2z_2 + ... + x_nz_n \quad \text{(Commutative Property of } + \text{ in } \mathbb{R}) \\ = \;& (x_1, x_2, ..., x_n) \cdot (y_1, y_2, ..., y_n) + (x_1, x_2, ..., x_n) \cdot (z_1, z_2, ..., z_n) \quad \text{(Defn of } \cdot \text{ in } \mathbb{R}^n) \\ = \;& \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}. \end{aligned}$$

Corollary 2.3. $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Proof. We no longer need to write vectors in coordinate form to prove things about the dot product. We have

$$\begin{aligned}
& \|\mathbf{x}+\mathbf{y}\|^2 \\
= \; & (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) \text{ (Theorem 2.2(2))} \\
= \; & (\mathbf{x}+\mathbf{y})\cdot\mathbf{x}+(\mathbf{x}+\mathbf{y})\cdot\mathbf{y} \text{ (Theorem 2.2(4))} \\
= \; & \mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{y} \text{ (Theorem 2.2(4))} \\
= \; & \|\mathbf{x}\|^2+\mathbf{y}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{y}+\|\mathbf{y}\|^2 \text{ (Theorem 2.2(2))} \\
= \; & \|\mathbf{x}\|^2+\mathbf{x}\cdot\mathbf{y}+\mathbf{x}\cdot\mathbf{y}+\|\mathbf{y}\|^2 \text{ (Theorem 2.2(1))} \\
= \; & \|\mathbf{x}\|^2+2(\mathbf{x}\cdot\mathbf{y})+\|\mathbf{y}\|^2 .
\end{aligned}$$

Definition 2.4. Two vectors $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ are called orthogonal if $\mathbf{x}\cdot\mathbf{y}=0$.

Theorem 2.5. Given $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ with $\mathbf{y}\neq\mathbf{0}$. There exists vectors $\mathbf{x}^{\|},\mathbf{x}^{\perp}\in\mathbb{R}^n$ such that

(i) $\mathbf{x}^{\|}$ is parallel to $\mathbf{y}$
(ii) $\mathbf{x}^{\perp}$ is orthogonal to $\mathbf{y}$
(iii) $\mathbf{x}=\mathbf{x}^{\|}+\mathbf{x}^{\perp}$.

Proof. Suppose that $\mathbf{x}^{\|},\mathbf{x}^{\perp}\in\mathbb{R}^n$ satisfy conditions (i), (ii), and (iii). Then

$$\mathbf{x}\cdot\mathbf{y}=(\mathbf{x}^{\|}+\mathbf{x}^{\perp})\cdot\mathbf{y}=\mathbf{x}^{\|}\cdot\mathbf{y}+\mathbf{x}^{\perp}\cdot\mathbf{y}=\mathbf{x}^{\|}\cdot\mathbf{y}+\mathbf{0}=\mathbf{x}^{\|}\cdot\mathbf{y}=c\mathbf{y}\cdot\mathbf{y}=c\,\|\mathbf{y}\|^2 .$$

It follows that

$$c=\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}.$$

Take

$$\mathbf{x}^{\|}=\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}$$

and so

$$\mathbf{x}^{\perp}=\mathbf{x}-\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}.$$

Since

$$\mathbf{x}^{\perp}\cdot\mathbf{y}=(\mathbf{x}-\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y})\cdot\mathbf{y}=\mathbf{x}\cdot\mathbf{y}-\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}\cdot\mathbf{y}=\mathbf{x}\cdot\mathbf{y}-\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\|\mathbf{y}\|^2=\mathbf{x}\cdot\mathbf{y}-\mathbf{x}\cdot\mathbf{y}=\mathbf{0}$$

we have verified that $\mathbf{x}^{\perp}$ is orthogonal to $\mathbf{y}$.

Definition 2.6. Given $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ with $\mathbf{y}\neq\mathbf{0}$. We call the vector $\mathbf{x}^{\|}$ the projection of $\mathbf{x}$ onto $\mathbf{y}$. We denote this projection by

$$\mathrm{proj}_{\mathbf{y}}(\mathbf{x})=\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}.$$

Theorem 2.7. (Cauchy-Schwarz Inequality) If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$. Equality holds ifand only if $\mathbf{x}$ and $\mathbf{y}$ are parallel.

Proof. We give an alternative proof than the one from the text. Let

$$f(t) = \|\mathbf{x} - t\mathbf{y}\|^2 .$$

Notice that f is the square of a real number so that $\|\mathbf{x} - t\mathbf{y}\|^2 \geq 0$. Now,

$$f(t) = \|\mathbf{x} - t\mathbf{y}\|^2 = (\mathbf{x} - t\mathbf{y}) \cdot (\mathbf{x} - t\mathbf{y}) = \|\mathbf{x}\|^2 + 2t\mathbf{x} \cdot \mathbf{y} + t^2 \|\mathbf{y}\|^2 .$$

By inspection, we see that $f(t)$ is a quadratic function in the variable t. Indeed,

$$f(t) = c + bt + at^2 \text{ where } c = \|\mathbf{x}\|^2 \text{ and } b = 2\mathbf{x} \cdot \mathbf{y} \text{ and } a = \|\mathbf{y}\|^2 .$$

Since f has at most one real root, the discriminant $D = b^2 - 4ac$ satisfies $D \leq 0$. That is,

$$\begin{aligned} (2\mathbf{x} \cdot \mathbf{y})^2 - 4\|\mathbf{y}\|^2 \|\mathbf{x}\|^2 &\leq 0 \\ 4(\mathbf{x} \cdot \mathbf{y})^2 &\leq 4\|\mathbf{y}\|^2 \|\mathbf{x}\|^2 \\ (\mathbf{x} \cdot \mathbf{y})^2 &\leq \|\mathbf{y}\|^2 \|\mathbf{x}\|^2 \\ |\mathbf{x} \cdot \mathbf{y}| &\leq \|\mathbf{x}\| \|\mathbf{y}\| . \end{aligned}$$

Referring back to our quadratic, Notice that $f(t) = \|\mathbf{x} - t\mathbf{y}\|^2 = 0$ (i.e. has one real root) exactly when $D = 0$. In other words,

$$\begin{aligned} |\mathbf{x} \cdot \mathbf{y}| &= \|\mathbf{x}\| \|\mathbf{y}\| \\ \text{iff } \|\mathbf{x} - t\mathbf{y}\|^2 &= 0 \\ \text{iff } \|\mathbf{x} - t\mathbf{y}\| &= 0 \\ \text{iff } \mathbf{x} - t\mathbf{y} &= 0 \\ \text{iff } \mathbf{x} &= t\mathbf{y}. \end{aligned}$$

Remark 2.8. Notice that $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ if and only if $\frac{|\mathbf{x} \cdot \mathbf{y}|}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$ if and only if $-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$. This allows us to make the following definition.

Definition 2.9. We define the angle θ between the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ to be the real number

$$\theta = \cos^{-1}\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}\right).$$

Exercises. Section 1.2: 1(d), 2(d), 7, 10, 11, 13, 14, 17, 18.

Hyperplanes in $\mathbb{R}^n$

Definition 2.10. Let $\mathbf{a} \in \mathbb{R}^n$ be a nonzero vector. The *hyperplane* in $\mathbb{R}^n$ with normal vector $\mathbf{a}$ through the point $\mathbf{x}_0$ is the set

$$H(\mathbf{x}_0, \mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a} \cdot \mathbf{x} = \mathbf{a} \cdot \mathbf{x}_0\}.$$

Example 2.11. If $n = 2$, then $H(\mathbf{x}_0, \mathbf{a})$ is a line through $\mathbf{x}_0$. If $n = 3$, then $H(\mathbf{x}_0, \mathbf{a})$ is a plane through $\mathbf{x}_0$. Generalize. Let us prove the second statement. Write $\mathbf{a} = (a, b, c)$ and $\mathbf{x} = (x, y, z)$ and set $\mathbf{a} \cdot \mathbf{x}_0 = d$.

$$\begin{aligned} \mathbf{x} &\in H(\mathbf{x}_0, \mathbf{a}) \\ \mathbf{a} \cdot \mathbf{x} &= \mathbf{a} \cdot \mathbf{x}_0 \\ (a, b, c) \cdot (x, y, z) &= c \\ ax + by + cz &= d. \end{aligned}$$

We can assume that $a \neq 0$ (one of a, b, c must be nonzero). Then $x = d - \frac{b}{a}y - \frac{c}{a}z$ and

$$\begin{aligned} \mathbf{x} &= (x, y, z) \\ &= (d - \frac{b}{a}y - \frac{c}{a}z, y, z) \\ &= (d, 0, 0) + (-\frac{b}{a}y, y, 0) + (\frac{c}{a}z, 0, z) \\ &= (d, 0, 0) + y(-\frac{b}{a}, 1, 0) + z(\frac{c}{a}, 0, 1). \end{aligned}$$

Take $\mathbf{y}_0 = (d, 0, 0)$ and $\mathbf{u} = (-\frac{b}{a}, 1, 0)$ and $\mathbf{v} = (\frac{c}{a}, 0, 1)$. The computations above show that $H(\mathbf{x}_0, \mathbf{a}) = P(\mathbf{y}_0, \mathbf{u}, \mathbf{v})$.

Exercises. Section 1.3: 4, 6, 7, 9-13