Linear Algebra Notes

Stanley Li

April 27, 2025

Contents

1 Vectors and Spaces **1**

1.1 Parametric representations of lines 1

1.2 Linear dependence 2

1.3 Linear subspaces 3

1.4 Dot product 5

1.4.1 Properties 5

1.4.2 Geometric representation 5

1.4.3 Interpretation 6

1.4.4 Cauchy-Schwarz Inequality 6

1.5 Cross product 8

1.5.1 Properties 8

1.5.2 Geometric representation 8

1 Vectors and Spaces

1.1 Parametric representations of lines

Example 1.1. Suppose that L_1 and L_2 are lines in the plane, that the x-intercepts of L_1 and L_2 are 5 and -1, respectively, and that the respective y-intercepts are 5 and 1. Find the point of intersection of L_1 and L_2.

Solution 1.1. Pick two points on L_1, i.e. $(5, 0)$ and $(0, 5)$. Let $\mathbf{a} = \langle 5,\ 0 \rangle$ and $\mathbf{b} = \langle 0,\ 5 \rangle$. Now, L_1 can be represented by the following:

$$
\begin{aligned}
L_1 &= \{(\mathbf{a}-\mathbf{b})t+\mathbf{b} \mid t \in \mathbb{R}\} \\
&= \left\{ \begin{bmatrix} 5t \\ -5t+5 \end{bmatrix} \middle| \, t \in \mathbb{R} \right\}
\end{aligned}
$$

The $(\mathbf{a}-\mathbf{b})t$ represents L_1 parameterized through the origin, so we apply a translation of $\mathbf{a}$ or $\mathbf{b}$. Similarly, let $\mathbf{c} = \langle -1, 0\rangle$ and $\mathbf{d} = \langle 0, 1\rangle$.

$$
\begin{aligned}
L_2 &= \{(\mathbf{c}-\mathbf{d})s+\mathbf{d} \mid s \in \mathbb{R}\} \\
&= \left\{ \begin{bmatrix} -s \\ -s+1 \end{bmatrix} \middle| \, s \in \mathbb{R} \right\}
\end{aligned}
$$

The point of intersection is where $L_{1_x} = L_{2_x}$ and $L_{1_y} = L_{2_y}$, allowing us to define a system of equations.

$$
5t = -s \qquad\qquad -5t+5 = -s+1
$$

$$
\begin{aligned}
s &= -2 \\
x &= -s = 2 \\
y &= -s+1 = 3
\end{aligned}
$$

The point of intersection is (2, 3). ■

1.2 Linear dependence

Definition 1.1 (Linear combination)**.** Let $V = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}\}$ where $\mathbf{v_i} \in \mathbb{R}^n$. A *linear combination* of V is defined to be $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k}$ where $c_i \in \mathbb{R}$.

Definition 1.2 (Linearly dependent)**.** Let $V = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}\}$ where $\mathbf{v_i} \in \mathbb{R}^n$. V is *linearly dependent* if and only if $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k} = \mathbf{0}$ where $c_i \in \mathbb{R}$ and there exists at least one c_i such that $c_i \neq 0$. In other words, V is *linearly independent* if and only if $c_1, c_2, \ldots, c_k = 0$ is the only solution.

Definition 1.3 (Span)**.** Given a set S of vectors, the *span* of S, denoted span(S), is the set of all linear combinations of S.

Example 1.2. Let $S = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}$. Is S linearly dependent?

Solution 1.2.

$$\begin{bmatrix}1\\-1\\2\end{bmatrix}c_1+\begin{bmatrix}2\\1\\3\end{bmatrix}c_2+\begin{bmatrix}-1\\0\\2\end{bmatrix}c_3=0$$

$$\begin{aligned}c_1+2c_2-c_3 &= 0\\ -c_1+c_2 &= 0\\ 2c_1+3c_2+2c_3 &= 0\end{aligned}$$

Solving this system, we find that the only solution is $c_1 = c_2 = c_3 = 0$. Thus, S is linearly independent. ■

We can go further by saying that the span of S is $\mathbb{R}^3$. This can be proven by setting the linear combination of S to an arbitrary 3-dimensional vector and isolating for the scalars c_1, c_2, and c_3, which tells us that any given vector in $\mathbb{R}^3$ can be represented in a specific linear combination of S. This can be thought of intuitively as well: if S is linearly independent, each vector of S introduces new directionality; there is no vector in S that can be defined as a linear combination of the other vectors.

1.3 Linear subspaces

Definition 1.4 (Linear subspace)**.** A set of vectors $V \subseteq \mathbb{R}^n$ is defined to be a linear/vector *subspace* of $\mathbb{R}^n$ if and only if it contains $\mathbf{0}$, it is closed under scalar multiplication, and it is closed under addition:

$$\begin{aligned}\mathbf{0} &\in V\\ \forall\ c\in\mathbb{R},\ c\mathbf{v} &\in V\\ \forall\ \mathbf{v_1},\mathbf{v_2}\in V, \mathbf{v_1}+\mathbf{v_2} &\in V\end{aligned}$$

Theorem 1.1. *Let $V = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$. It holds true that* $\operatorname{span}(V)$ *is a subspace of* $\mathbb{R}^n$.

Proof. If $\operatorname{span}(V)$ is a valid linear subspace, it must contain the zero vector, it must be closed under scalar multiplication, and it must be closed under addition.

By definition, the span of V is the set of all linear combinations of V:

$$\operatorname{span}(V) = \{c_1\mathbf{v_1}+c_2\mathbf{v_2}+\cdots+c_n\mathbf{v_n}\ \forall\ c_i\in\mathbb{R}\}$$

1. Inclusion of zero vector:

$$\text{Let } c_1, c_2, \dots, c_n = 0.$$

$$\implies c_1\mathbf{v_1}+c_2\mathbf{v_2}+\cdots+c_n\mathbf{v_n}=\mathbf{0}$$

$$\implies \mathbf{0}\in\operatorname{span}(V)$$

2. Closure under scalar multiplication:

$$\text{Let } \mathbf{a} \in \text{span}(V).$$

$$\iff c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n} = \mathbf{a}$$

$$d(c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}) = d\mathbf{a} \text{ where } d \in \mathbb{R}$$
$$\implies dc_1\mathbf{v_1} + dc_2\mathbf{v_2} + \cdots + dc_n\mathbf{v_n} = d\mathbf{a}$$

dc_i is just a scalar, meaning $d\mathbf{a}$ is another linear combination of V:

$$d\mathbf{a} \in \text{span}(V)$$

3. Closure under addition:

Let $\mathbf{a}, \mathbf{b} \in \text{span}(V)$.

$$\begin{aligned} c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n} &= \mathbf{a} \\ d_1\mathbf{v_1} + d_2\mathbf{v_2} + \cdots + d_n\mathbf{v_n} &= \mathbf{b} \\ (c_1 + d_1)\mathbf{v_1} + (c_2 + d_2)\mathbf{v_2} + \cdots + (c_n + d_n)\mathbf{v_n} &= \mathbf{a} + \mathbf{b} \end{aligned}$$

Again, $(c_i + d_i)$ is just a scalar, meaning $\mathbf{a} + \mathbf{b}$ is another linear combination of V:

$$\mathbf{a} + \mathbf{b} \in \text{span}(V)$$

■

Definition 1.5 (Basis)**.** Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be linearly independent. It follows that S is a *basis* for the subspace $V = \text{span}(S)$.

Lemma 1.1. *Any vector in a subspace V is the result of a unique linear combination of some basis for V.*

Proof. Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be a basis for the subspace V. Suppose $\mathbf{a} \in V$.

$$\mathbf{a} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n} \text{ where } c_i \in \mathbb{R}$$

Assume that $\mathbf{a}$ can be represented by another linear combination of S.

$$\begin{aligned} \mathbf{a} &= d_1\mathbf{v_1} + d_2\mathbf{v_2} + \cdots + d_n\mathbf{v_n} \text{ for some } d_j \neq c_j \\ \mathbf{0} &= (c_1 - d_1)\mathbf{v_1} + (c_2 - d_2)\mathbf{v_2} + \cdots + (c_n - d_n)\mathbf{v_n} \end{aligned}$$

S is linearly independent so the scalar $(c_i - d_i)$ must be zero for $1 \leq i \leq n$. This implies that $c_i = d_i$ which contradicts the statement that some $d_j \neq c_j$. This further implies that $\mathbf{a}$ cannot be represented by more than one linear combination of S. ■

1.4 Dot product

Definition 1.6. The *dot product* of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, denoted $\mathbf{a} \cdot \mathbf{b}$, is defined to be $\mathbf{a} \cdot \mathbf{b} = \mathbf{a_1 b_1} + \mathbf{a_2 b_2} + \cdots + \mathbf{a_n b_n}$.

Lemma 1.2. $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$.

Proof. Let $\mathbf{a} \in \mathbb{R}^n$.

$$\begin{aligned} \|\mathbf{a}\| &= \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \\ \|\mathbf{a}\|^2 &= |a_1^2 + a_2^2 + \cdots + a_n^2| \\ &= a_1^2 + a_2^2 + \cdots + a_n^2 \\ &= \mathbf{a} \cdot \mathbf{a} \end{aligned}$$

■

1.4.1 Properties

Dot products are commutative ($\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$), distributive ($\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$), and associative with scalars ($c(\mathbf{a} \cdot \mathbf{b}) = (c\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (c\mathbf{b})$). These properties can be easily proven with the definition of the dot product.

1.4.2 Geometric representation

The dot product of two vectors can also be represented in relation to the angle between them, θ, by the following identity: $\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta)$.

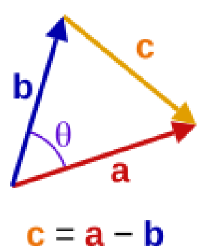

Figure 1: The angle θ in the triangle constructed by $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a} - \mathbf{b}$

Proof. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ such that $\mathbf{a}, \mathbf{b} \neq 0$. Using the Law of Cosines,

$$\begin{aligned} \|\mathbf{a} - \mathbf{b}\|^2 &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) \\ \|\mathbf{a}\|^2 - 2(\mathbf{a} \cdot \mathbf{b}) + \|\mathbf{b}\|^2 &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) \\ -2(\mathbf{a} \cdot \mathbf{b}) &= -2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) \\ \mathbf{a} \cdot \mathbf{b} &= \|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) \end{aligned}$$

■

1.4.3 Interpretation

The geometric representation makes it easy to recognize that $\mathbf{a} \cdot \mathbf{b}$ is maximized when $\theta = 0°$, which is when $\mathbf{a}$ and $\mathbf{b}$ are collinear in the same direction. On the other hand, $\mathbf{a} \cdot \mathbf{b}$ is minimized when $\theta = 180°$, which is when $\mathbf{a}$ and $\mathbf{b}$ are collinear in opposite directions. When $\mathbf{a}$ and $\mathbf{b}$ are perpendicular to each other (i.e. $\theta = 90°$), $\mathbf{a} \cdot \mathbf{b} = 0$. This means that we can interpret the dot product as a measure of collinearity.

Definition 1.7 (Orthogonal)**.** Vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ are *orthogonal* to each other if and only if $\mathbf{a} \cdot \mathbf{b} = 0$.

1.4.4 Cauchy-Schwarz Inequality

Lemma 1.3. *For some scalar $c \geq 0$ and some vector $\mathbf{a}$, $c\|\mathbf{a}\| = \|c\mathbf{a}\|$.*

Proof. Let $c \geq 0, c \in \mathbb{R}$. Let $\mathbf{a} \in \mathbb{R}^n$.

$$\begin{aligned} c\|\mathbf{a}\| &= c\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \\ &= \sqrt{c^2(a_1^2 + a_2^2 + \cdots + a_n^2)} \\ &= \sqrt{(ca_1)^2 + (ca_2)^2 + \cdots + (ca_n)^2} \\ &= \|c\mathbf{a}\| \end{aligned}$$

■

Theorem 1.2 (Cauchy-Schwarz inequality)**.** *If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\|\mathbf{x}\|\|\mathbf{y}\| \geq |\mathbf{x} \cdot \mathbf{y}|$. Furthermore, $\|\mathbf{x}\|\|\mathbf{y}\| = |\mathbf{x} \cdot \mathbf{y}|$ if and only if $\mathbf{y} = c\mathbf{x}$ for some $c \in \mathbb{R}$.*

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be nonzero. Let $p(t) = \|t\mathbf{y} - \mathbf{x}\|^2$. Note that $p(t) \geq 0$ for all $t \in \mathbb{R}$.

$$\begin{aligned} p(t) &= (t\mathbf{y} - \mathbf{x}) \cdot (t\mathbf{y} - \mathbf{x}) \\ &= t^2\|\mathbf{y}\|^2 - 2t(\mathbf{x} \cdot \mathbf{y}) + \|\mathbf{x}\|^2 \end{aligned}$$

Let $a = \|\mathbf{y}\|^2$. Let $b = 2(\mathbf{x} \cdot \mathbf{y})$. Let $c = \|\mathbf{x}\|^2$. Note that $a \neq 0$ because $\mathbf{y} \neq \mathbf{0}$.

$$\begin{aligned} p(t) &= at^2 - bt + c \geq 0 \\ p\left(\frac{b}{2a}\right) &= a\left(\frac{b}{2a}\right)^2 - b\left(\frac{b}{2a}\right) + c \\ &= \frac{b^2}{4a} - \frac{b^2}{2a} + c \\ &= -\frac{b^2}{4a} + c \geq 0 \\ \implies c &\geq \frac{b^2}{4a} \end{aligned}$$

Substituting back for a, b, and c:

$$\begin{aligned}
\|\mathbf{x}\|^2 &\geq \frac{4(\mathbf{x}\cdot\mathbf{y})^2}{4\|\mathbf{y}\|^2} \\
\|\mathbf{x}\|^2\|\mathbf{y}\|^2 &\geq (\mathbf{x}\cdot\mathbf{y})^2 \\
\|\mathbf{x}\|\|\mathbf{y}\| &\geq |\mathbf{x}\cdot\mathbf{y}|
\end{aligned}$$

Consider the case of the equality:

$$\begin{aligned}
\|\mathbf{x}\|\|\mathbf{y}\| &= |\mathbf{x}\cdot\mathbf{y}| \\
\|\mathbf{x}\|\|\mathbf{y}\| &= |\|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta)|
\end{aligned}$$

This can be true if and only if $\theta = 0°$ or $\theta = 180°$, meaning $\mathbf{x}$ and $\mathbf{y}$ must be collinear with each other.

While we initially assumed that $\mathbf{x}$ and $\mathbf{y}$ were nonzero, it is easy to see that the inequality still holds when $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$. ■

Theorem 1.3 (Triangle inequality)**.** *If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. This inequality becomes an equality if and only if $\mathbf{y} = c\mathbf{x}$ for some $c \in \mathbb{R}$ such that $c \geq 0$. In other words, the sum of the lengths of two sides of a triangle is always greater than the length of its third side.*

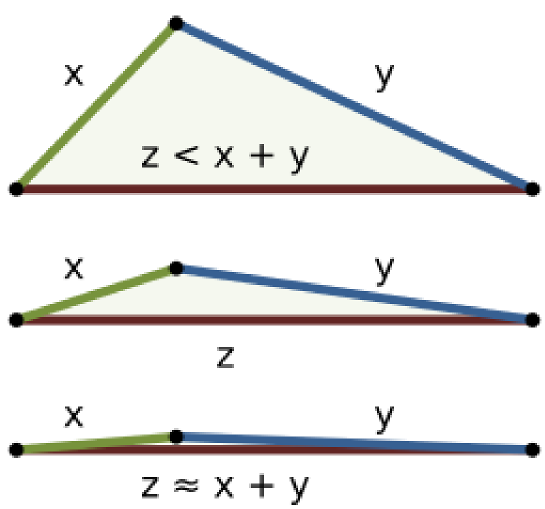

Figure 2: As two sides of a triangle get closer to being collinear, the sum of their lengths gets closer to the length of the third side.

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. By the Cauchy-Schwarz inequality (Theorem 1.2),

$$\|\mathbf{x}\|\|\mathbf{y}\| \geq |\mathbf{x}\cdot\mathbf{y}| \geq \mathbf{x}\cdot\mathbf{y}$$

Note that $|\mathbf{x}\cdot\mathbf{y}| = \mathbf{x}\cdot\mathbf{y}$ if and only if $\mathbf{x}\cdot\mathbf{y} \geq 0$. Else, $|\mathbf{x}\cdot\mathbf{y}| > \mathbf{x}\cdot\mathbf{y}$. Additionally, $|\mathbf{x}\cdot\mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$ if and only if $\mathbf{y} = c\mathbf{x}$ for some $c \in \mathbb{R}$, as stated in the Cauchy-Schwarz inequality. So, if $c \geq 0$ and $\mathbf{y} = c\mathbf{x}$, then $\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}\cdot\mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$.

Now, consider the following:

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2$$

Since $\|\mathbf{x}\|\|\mathbf{y}\| \geq \mathbf{x}\cdot\mathbf{y}$, we can create the inequality $\|\mathbf{x}+\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2$. For the reason previously mentioned, this inequality becomes an equality if and only if $\mathbf{y} = c\mathbf{x}$ where $c \geq 0$.

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &\leq \ |\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2\\ \|\mathbf{x}+\mathbf{y}\|^2 &\leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2\\ \|\mathbf{x}+\mathbf{y}\| &\leq \|\mathbf{x}\| + \|\mathbf{y}\|\end{aligned}$$

■

1.5 Cross product

Definition 1.8. The *cross product* of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, denoted $\mathbf{a}\times\mathbf{b}$, is defined to be

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix} \times \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}\\ &= \begin{bmatrix}a_2b_3 - a_3b_2\\a_3b_1 - a_1b_3\\a_1b_2 - a_2b_1\end{bmatrix}\end{aligned}$$

Remark. A way to easily remember how to calculate the cross product is by writing it as a determinant (covered in Section ??):

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix}\hat{\imath} & \hat{\jmath} & \hat{k}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix}$$

1.5.1 Properties

Cross products are anticommutative ($\mathbf{a}\times\mathbf{b} = -(\mathbf{b}\times\mathbf{a})$), distributive ($\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = (\mathbf{a}\times\mathbf{b})+(\mathbf{a}\times\mathbf{c})$), and associative with scalars ($c(\mathbf{a}\times\mathbf{b}) = (c\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(c\mathbf{b})$). They are not associative themselves. These properties can be proven from the definition of the cross product.

1.5.2 Geometric representation

Theorem 1.4. *If* $\mathbf{a}\times\mathbf{b} = \mathbf{c}$, *then* $\mathbf{c}$ *is orthogonal to both* $\mathbf{a}$ *and* $\mathbf{b}$.

Proof. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

$$\mathbf{a}\times\mathbf{b} = \begin{bmatrix}a_2b_3 - a_3b_2\\a_3b_1 - a_1b_3\\a_1b_2 - a_2b_1\end{bmatrix} = \mathbf{c}$$

$$
\begin{aligned}
\mathbf{a} \cdot \mathbf{c} &= a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1) \\
&= a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_1a_2b_3 + a_1a_3b_2 - a_2a_3b_1 \\
&= 0 \\
\mathbf{b} \cdot \mathbf{c} &= b_1(a_2b_3 - a_3b_2) + b_2(a_3b_1 - a_1b_3) + b_3(a_1b_2 - a_2b_1) \\
&= a_2b_1b_3 - a_3b_1b_2 + a_3b_1b_2 - a_1b_2b_3 + a_1b_2b_3 - a_2b_1b_3 \\
&= 0
\end{aligned}
$$

■

The direction of the normal vector obtained from the cross product $\mathbf{a} \times \mathbf{b}$ can be determined using the right-hand rule: point your index finger in the direction of $\mathbf{a}$, then point your middle finger in the direction of $\mathbf{b}$, then stick out your thumb. Your thumb represents the direction of the normal vector.

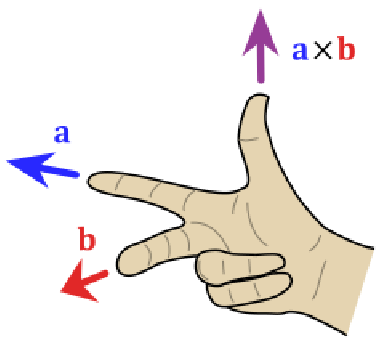

Figure 3: The right-hand rule

Similar to the dot product, the cross product of two vectors can also be represented in relation to the angle between them, θ, by the following identity: $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\| \sin(\theta)$.

Proof. ■