Introduction — Dot Product, Length and Orthogonality

Matteo Paz, Dylan Rupel

March 20th, 2024

Recall the dot product between two vectors. For this unit, we will primarily consider vectors to be column vectors, or matrices of size $N \times 1$.

Definition 1 *Given two column vectors* $\mathbf{v}, \mathbf{w}$*, their dot product denoted* $\mathbf{v} \cdot \mathbf{w}$ *is given by the form* $\mathbf{v}_1\mathbf{w}_1 + \cdots + \mathbf{v}_n\mathbf{w}_n$*, but defined by the notation*

$$\mathbf{v}^T\mathbf{w}$$

Proposition 1 *The dot product gives a positive definite symmetric bilinear form.*

First, we need to define these terms.

Definition 2 *A bilinear form is a function* $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *which is linear in each component separately:*

- $T(\lambda\mathbf{v}, \mathbf{w}) = \lambda T(\mathbf{v}, \mathbf{w}) = T(\mathbf{v}, \lambda\mathbf{w})$
- $T(\mathbf{u}, \mathbf{v} + \mathbf{w}) = \mathbf{T}(\mathbf{u}, \mathbf{v}) + \mathbf{T}(\mathbf{u}, \mathbf{w})$
- $T(\mathbf{u} + \mathbf{v}, \mathbf{w}) = \mathbf{T}(\mathbf{u}, \mathbf{w}) + \mathbf{T}(\mathbf{v}, \mathbf{w})$

It is symmetric if $T(\mathbf{v}, \mathbf{w}) = T(\mathbf{w}, \mathbf{v})$.
It is positive semidefinite if $T(\mathbf{v}, \mathbf{v}) \geq 0$.
It is positive definite if $T(\mathbf{v}, \mathbf{v}) \geq 0$ *with equality iff* $\mathbf{v} = \mathbf{0}$

We typically denote a bilinear form via $\langle \mathbf{v}, \mathbf{w} \rangle$*. It can always be formulated as a square matrix* $A \in \mathbb{R}^{n \times n}$*, a way of multiplying two vectors in the space:*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T A \mathbf{w}$$

Now, we can notice that the dot product is simply a bilinear form with $A = \mathrm{Id}$. It is clearly symmetric, and since we are summing squares ($\mathbf{v} \cdot \mathbf{v} = \mathbf{v}_1^2 + \cdots$) it is obviously always greater than or equal to zero, with equality when $\mathbf{v} = 0$.

Definition 3 *The length or norm of a vector* $\mathbf{v}$ *is defined to be*

$$\sqrt{\mathbf{v} \cdot \mathbf{v}}$$

How can we define the angle of vectors in this space? We only need an inner product, here the dot product.

Consider vectors $\mathbf{v}, \mathbf{w}$. Then $(\mathbf{v}, \mathbf{w}, \mathbf{v}-\mathbf{w})$ forms a triangle. By the law of cosines, having $\theta \sim \mathbf{v}, \mathbf{w}$:

$$|\mathbf{v}-\mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}||\mathbf{w}|\cos(\theta)$$

We know that $|\mathbf{v}-\mathbf{w}|^2 = (\mathbf{v}-\mathbf{w}) \cdot (\mathbf{v}-\mathbf{w})$. By linearity we have then $|\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}||\mathbf{w}|$. Substitution:

$$|\mathbf{v}|^2 + |\mathbf{w}|^2 - 2(\mathbf{v} \cdot \mathbf{w}) = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}||\mathbf{w}|\cos(\theta)$$

Finally we then have

$$\cos(\theta) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}||\mathbf{w}|}$$

From this we arrive at our notion of angle. Notice that this angle is only $\pi/2$ when $\mathbf{v} \cdot \mathbf{w} = 0$. Thus this is our notion of **orthogonality**.

Proposition 2 *Let $U = \{\mathbf{w}_1, \cdots, \mathbf{w}_k\}$ be an orthogonal basis for $W \subseteq \mathbb{R}^n$ and $V = \{\mathbf{v}_1, \cdots, \mathbf{v}_l\}$ be an orthogonal basis for $W^\perp$. Then $U \cup V$ is an orthogonal basis for $\mathbb{R}^n$.*

Proof.
Let $\mathbf{x} \in \mathbb{R}^\mathbf{n}$. If it can be written as $\mathbf{x} = \mathbf{c_1 w_1} + \cdots + \mathbf{c_k w_k} + \mathbf{d_1 v_1} + \cdots + \mathbf{d_l v_l}$, then $\mathbf{x} \cdot \mathbf{w_i} = \mathbf{c_i w_i} \cdot \mathbf{w_i}$. Thus $c_i = \frac{\mathbf{x} \cdot \mathbf{w_i}}{\mathbf{w}_i \cdot \mathbf{w}_i}$. This is useful to find the component in W.

Now, consider the vector $\mathbf{x} - \sum_i^k \frac{x \cdot \mathbf{w}_i}{\mathbf{w}_i \cdot \mathbf{w}_i}$. We want this to be in $W^\perp$. Thus, dot with arbitrary $\mathbf{w}_j$:

$$\left(\mathbf{x} - \sum_i^k \frac{\mathbf{x} \cdot \mathbf{w}_i}{\mathbf{w}_i \cdot \mathbf{w}_i}\mathbf{w}_i\right) \cdot \mathbf{w}_j = \mathbf{x} \cdot \mathbf{w}_j - \sum_i^k \frac{\mathbf{x} \cdot \mathbf{w}_i}{\mathbf{w}_i \cdot \mathbf{w}_i}\mathbf{w}_i \cdot \mathbf{w}_j$$

$$=^{i=j} \mathbf{x} \cdot \mathbf{w}_j - \frac{\mathbf{x} \cdot \mathbf{w}_j}{\mathbf{w}_j \cdot \mathbf{w}_j}\mathbf{w}_j \cdot \mathbf{w}_j$$

$$= 0$$

Thus $\mathbf{x} - \sum_i^k \frac{x \cdot \mathbf{w}_i}{\mathbf{w}_i \cdot \mathbf{w}_i} \in W^\perp$. Thus it can be written as a linear combination d_i of vectors in V. Thus

$$\mathbf{x} \in \text{Span}\{U \cup V\}$$

Since all vectors in $U \cup V$ are orthogonal, that set is linearly independent. Thus $U \cup V$ is a basis of $\mathbb{R}^n$. Then $k + l = n$. $\square$

1 The Gram-Schmidt Process

The Gram-Schmidt Process allows us to get an orthonormal basis from any basis. It works in steps. First, consider this basis of $\mathbb{R}^2$:

$$\left\{ \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$$

The steps are as follows

1. For the first vector, normalize it: $\begin{bmatrix} 2 \\ 2 \end{bmatrix} \to \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$
2. Remove the component of this first vector $\mathbf{v}_1$ from all other vectors ahead of it. $\mathbf{w} - (\mathbf{w} \cdot \mathbf{v})\mathbf{v}$: $\begin{bmatrix} 1 \\ 3 \end{bmatrix} - 2\sqrt{2} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
3. Repeat (1), (2) going down the list of vectors, until all have been normalized: $\begin{bmatrix} -1 \\ 1 \end{bmatrix} \to \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Now we are left with $\left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$, an orthonormal basis.