Lecture 37 : Orthonormal Bases

Math 3013
Oklahoma State University

January 31, 2022

Lecture 36

1. Orthonormal Bases
2. The Gram-Schmid Process

Orthogonality and Bases

One of the most useful properties of the standard basis $[\mathbf{e}_1, \ldots, \mathbf{e}_n]$ of $\mathbb{R}^n$ is the fact that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1}$$

For example, this property allows us to easily determine the component of a vector $\mathbf{v}$ along the i^{th} basis vector $\mathbf{e}_i$ by simply computing its inner product with $\mathbf{e}_i$:
For if

$$\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n$$

Then

$$\begin{aligned} \mathbf{e}_i \cdot \mathbf{v} &= \mathbf{e}_i \cdot (v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n) \\ &= v_1\mathbf{e}_i \cdot \mathbf{e}_1 + v_2\mathbf{e}_i \cdot \mathbf{e}_2 + \cdots + v_i\mathbf{e}_i \cdot \mathbf{e}_i + \cdots + v_n\mathbf{e}_i \cdot \mathbf{e}_n \\ &= 0 + 0 + \cdots + 0 + v_i + 0 + \cdots + 0 \\ &= v_i \end{aligned}$$

This circumstance is not true for a more general basis.

For a more general basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, in order to find the constants $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

you have to solve the linear system

$$\begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

which is a much more strenuous task.

On the other hand, we have lots and lots of choices of bases for $\mathbb{R}^n$ or for any subspace W of $\mathbb{R}^n$.

What we shall be developing in today's lecture is a way to contruct bases $B = \{\mathbf{n}_1, \ldots, \mathbf{n}_n\}$ that have orthogonality and normalization properties just like $\mathcal{E} = \{\mathbf{e}_1, , \ldots, \mathbf{e}_n\}$

$$\mathbf{n}_i \cdot \mathbf{n}_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

For such **orthonormal bases**, we will be able to rapidly determine the coefficients c_i such that

$$\mathbf{v} = c_1\mathbf{n}_1 + \cdots + c_n\mathbf{n}_n$$

by simply computing inner products

$$c_i = \mathbf{n}_i \cdot \mathbf{v}$$

Remark

The reason for the nomenclature **orthonormal basis** is because the condition

$$\mathbf{n}_i \cdot \mathbf{n}_j = \delta_{ij}$$

implies that the distinct basis vectors are not only orthogonal to one another

$$\mathbf{n}_i \cdot \mathbf{n}_j = 0 \quad \text{if } i \neq j$$

and they are "normalized" so that they have unit length

$$\|\mathbf{n}_i\| = \sqrt{\mathbf{n}_i \cdot \mathbf{n}_i} = \sqrt{1} = 1$$

Summary: Orthonormal Bases

A set of vectors $B = \{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$ of vectors in W is an **orthonormal basis** if

(i) $W = span\,(\mathbf{n}_1, \ldots, \mathbf{n}_k)$

(ii) $\{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$ are linearly independent

(iii) $\mathbf{n}_i \cdot \mathbf{n}_j = 0$ if $i \neq j$

(iv) $\mathbf{n}_i \cdot \mathbf{n}_i = 1$ for all i

When $B = \{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$ is an orthonormal basis for a subspace W then it's easy to find the coordinate vector $\mathbf{v}_B$ of a vector $\mathbf{v}$ with respect to the basis B,

$$\mathbf{v} = c_1\mathbf{n}_1 + \cdots + c_k\mathbf{n}_k \quad \Rightarrow \quad c_i = \mathbf{n}_i \cdot \mathbf{v} \quad , \quad i = 1, \ldots, k$$

Actually, these conditions are somewhat redundant, We'll see next that condition (iii) actually implies condition (ii).

Orthogonality and Linear Independence

Lemma

Any set of non-zero, mutually orthogonal vectors is linearly independent.

Proof. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors such that

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ if } i \neq j$$

From the definition of **linear independence**, we need to show that the equation

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \qquad (*)$$

can only be satisfied by setting each coefficient c_i equal to 0.

Proof of Lemma, Cont'd

So suppose

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \qquad (*)$$

Taking the dot product of both sides of (*) with $\mathbf{v}_1$ yields

$$\begin{aligned} 0 &= c_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + 0 + \cdots + 0 \\ \Rightarrow \quad & c_1 = 0 \quad \text{since } \mathbf{v}_1 \cdot \mathbf{v}_1 = \|\mathbf{v}_1\|^2 \neq 0 \end{aligned}$$

Likewise, taking the dot product of both sides with $\mathbf{v}_i$ leads to $c_i = 0$ for all i. Thus,

$$(*) \implies c_1 = 0, \ldots, c_k = 0$$

and so the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ are linearly independent. □

The Gram-Schmidt Process

The Gram-Schmidt Process is an algorithm by which one can create from **any** basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$, an orthogonal basis $B' = \{\mathbf{o}_1, \ldots, \mathbf{o}_k\}$ for which

$$\mathbf{o}_i \cdot \mathbf{o}_j = 0 \quad \text{wherever } i \neq j$$

This algorithm begins with setting

$$\mathbf{o}_1 = \mathbf{b}_1$$

Next, we want a second basis vector, $\mathbf{o}_2$ that's perpendicular to $\mathbf{o}_1$.

The Gram-Schmidt Process, Cont'd

In Lecture 34, we learned that if **a** and **b** are two vectors we can decompose **a** into two components:

$$\mathbf{a} = \mathbf{a_b} + \mathbf{a}_\perp$$

where

$$\mathbf{a_b} = \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\mathbf{b} \quad , \quad \mathbf{a}_\perp = \mathbf{a} - \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\mathbf{b}$$

Here $\mathbf{a}_b$ has the same direction as the vector **b** and the direction of $\mathbf{a}_\perp$ is perpendicular to that of **b**.

The Gram-Schmidt Process, Cont'd

Using a similar orthogonal decomposition, with $\mathbf{o}_1$ playing the role of $\mathbf{b}$ and $\mathbf{b}_2$ playing the role of $\mathbf{a}$, we can split $\mathbf{b}_2$ into a component $(\mathbf{b}_2)_{||\mathbf{o}_1}$ that runs parallel to $\mathbf{o}_1$, and a component $(\mathbf{b}_2)_{\perp\mathbf{o}_1}$ that is runs perpendicular to $\mathbf{o}_1$:

$$\mathbf{b}_2 = (\mathbf{b}_2)_{||\mathbf{o}_1} + (\mathbf{b}_2)_{\perp\mathbf{o}_1}$$

We then set

$$\mathbf{o}_2 = (\mathbf{b}_2)_{\perp\mathbf{o}_1} = \mathbf{b}_2 - (\mathbf{b}_2)_{||\mathbf{o}_1} = \mathbf{b}_2 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_2}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1$$

Then $\mathbf{o}_2$ is automatically perpendicular to $\mathbf{o}_1$. The vector $\mathbf{o}_2$ constructed in this way will be our second orthogonal basis vector. Note that, by the Lemma, $\{\mathbf{o}_1, \mathbf{o}_2\}$ are also automatically linearly independent, since $\mathbf{o}_1$ and $\mathbf{o}_2$ are non-zero orthogonal vectors.

It's helpful to understand this construction of $\mathbf{o}_2$ as simply removing from $\mathbf{b}_2$ the component that runs parallel to $\mathbf{o}_1$. Because that's essentially how we'll get the other orthogonal basis vectors.

For instance, to get a third basis vector that orthogonal to both $\mathbf{o}_1$ and $\mathbf{o}_2$, we start with $\mathbf{b}_3$, and remove from it, both the part that runs parallel to $\mathbf{o}_1$ and the part of $\mathbf{b}_3$ that runs parallel to $\mathbf{o}_2$:

$$\mathbf{o}_3 = \mathbf{b}_3 - (\mathbf{b}_3)_{||\mathbf{o}_1} - (\mathbf{b}_3)_{||\mathbf{o}_2} = \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2}\mathbf{o}_2$$

and then $\{\mathbf{o}_1, \mathbf{o}_2, \mathbf{o}_3\}$ will be an orthogonal basis for $\mathit{span}\,(\mathbf{o}_1, \mathbf{o}_2, \mathbf{o}_3)$
And we can similarly construct more orthogonal vectors $\mathbf{o}_4, \mathbf{o}_5 \cdots$, until we reach

$$\mathbf{o}_k = \mathbf{b}_k - \frac{\mathbf{o}_1 \cdot \mathbf{b}_k}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_k}{\mathbf{o}_2 \cdot \mathbf{o}_2}\mathbf{o}_2 - \cdots - \frac{\mathbf{o}_{k-1} \cdot \mathbf{b}_3}{\mathbf{o}_{k-1} \cdot \mathbf{o}_{k-1}}\mathbf{o}_{k-1}$$

In this way, we end up with a set of k mutually orthogonal, and so linear independent, vectors in W.
But any set of k linearly independent vectors in a k-dimensional space W will be a basis for W.
So $B' = \{\mathbf{o}_1, \ldots, \mathbf{o}_k\}$, so constructed, will be an **orthogonal basis** for $W = span(\mathbf{b}_1, \ldots, \mathbf{b}_k)$.

The Gram-Schmidt Process

The Gram-Schmidt Process is the following algorithm.
Given any basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ for a subspace W, we can systematically construct an orthogonal basis $B' = \{\mathbf{o}_1, \ldots, \mathbf{o}_k\}$ for W, by setting

$$\begin{aligned}
\mathbf{o}_1 &= \mathbf{b}_1 \\
\mathbf{o}_2 &= \mathbf{b}_2 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_2}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 \\
\mathbf{o}_3 &= \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2}\mathbf{o}_2 \\
&\vdots \\
\mathbf{o}_k &= \mathbf{b}_k - \frac{\mathbf{o}_1 \cdot \mathbf{b}_k}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_k}{\mathbf{o}_2 \cdot \mathbf{o}_2}\mathbf{o}_2 - \cdots - \frac{\mathbf{o}_{k-1} \cdot \mathbf{b}_3}{\mathbf{o}_{k-1} \cdot \mathbf{o}_{k-1}}\mathbf{o}_{k-1}
\end{aligned}$$

The basis $\{\mathbf{o}_1, \cdots, \mathbf{o}_k\}$ obtained by the above algorithm, however, is an **orthogonal basis**, but not yet an **orthonormal basis**. For the vectors $\mathbf{o}_i$ while mutually orthogonal do not necessarily have the length 1.

But there is an easy fix for this. All we have to do is divide each of the orthogonal basis vectors $\mathbf{o}_i$ by their lengths $\|\mathbf{o}_i\| = \sqrt{\mathbf{o}_i \cdot \mathbf{o}_i}$ to get a set of k, mutually orthogonal, linearly independent vectors, all of length 1 :

$$\begin{array}{ccc} \mathbf{o}_1 & \longrightarrow & \mathbf{n}_1 = \dfrac{1}{\sqrt{\mathbf{o}_1 \cdot \mathbf{o}_1}}\mathbf{o}_1 \\ & \vdots & \\ \mathbf{o}_k & \longrightarrow & \mathbf{n}_k = \dfrac{1}{\sqrt{\mathbf{o}_k \cdot \mathbf{o}_k}}\mathbf{o}_k \end{array}$$

The vectors $\{\mathbf{n}_1, \cdots, \mathbf{n}_k\}$ so obtained will be an **orthonormal basis** for W.

Example

Find a orthonormal basis for the subspace

$$W = span\left([1,-1,1,0],[-1,0,0,1],[0,0,1,1]\right)$$

of $\mathbb{R}^5$.

First we need a basis for W.

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

This last matrix is in row echelon form with no non-zero rows. From this calculation we can conclude that the original three vectors are linearly independent and so will constitute a basis for W.

We can thus begin with the basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ with

$$\begin{aligned} \mathbf{b}_1 &= [1,-1,1,0] \\ \mathbf{b}_2 &= [-1,0,0,1] \\ \mathbf{b}_3 &= [0,0,1,1] \end{aligned}$$

as an initial basis to start the Gram-Schmidt Process.

Thus, we set

$$\begin{aligned} \mathbf{o}_1 &= \mathbf{b}_1 = [1, -1, 1, 0] \\ \Longrightarrow & \quad \|\mathbf{o}_1\|^2 = 3 \\ \Longrightarrow & \quad \mathbf{n}_1 = \left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right] \end{aligned}$$

Next we compute $\mathbf{o}_2$,

$$\begin{aligned} \mathbf{o}_2 &= \mathbf{b}_2 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_2}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 \\ &= [-1, 0, 0, 1] - \frac{(-1)}{3}[1, -1, 1, 0] \\ &= \left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 1\right] \end{aligned}$$

We have

$$\begin{aligned} \|\mathbf{o}_2\|^2 &= \frac{4}{9} + \frac{1}{9} + \frac{1}{9} + 1 = \frac{5}{3} \\ \Longrightarrow \quad \mathbf{n}_2 &= \frac{\mathbf{o}_2}{\|\mathbf{o}_2\|} = \sqrt{\frac{3}{5}}\left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 1\right] \end{aligned}$$

And, finally,

$$\begin{aligned}
\mathbf{o}_3 &= \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1}\mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2}\mathbf{o}_2 \\
&= [0, 0, 1, 1] - \frac{(1)}{3}[1, -1, 1, 0] - \frac{\frac{4}{3}}{\frac{5}{3}}\left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 1\right] \\
&= \left[\frac{1}{5}, \frac{3}{5}, \frac{2}{5}, \frac{1}{5}\right]
\end{aligned}$$

and

$$\|\mathbf{o}_3\|^2 = \frac{1 + 9 + 4 + 1}{25} = \frac{3}{5}$$

so

$$\mathbf{n}_3 = \sqrt{\frac{5}{3}}\left[\frac{1}{5}, \frac{3}{5}, \frac{2}{5}, \frac{1}{5}\right]$$

Thus,

$$B' = \left\{ \left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0 \right], \sqrt{\frac{3}{5}} \left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 1 \right], \sqrt{\frac{5}{3}} \left[\frac{1}{5}, \frac{3}{5}, \frac{2}{5}, \frac{1}{5} \right] \right\}$$

will be an orthonormal basis for W.
So now if one wants to determine the coefficients c_1, c_2, c_3 in the expansion

$$\mathbf{v} = c_1\mathbf{n}_1 + c_2\mathbf{n}_2 + c_3\mathbf{n}_3$$

of a vector $\mathbf{v}$ w.r.t. the orthonormal basis B', we just need to compute some dot products

$$\begin{aligned} c_1 &= \mathbf{n}_1 \cdot \mathbf{v} \\ c_2 &= \mathbf{n}_2 \cdot \mathbf{v} \\ c_3 &= \mathbf{n}_3 \cdot \mathbf{v} \end{aligned}$$