Recitation 8

Tuesday October 5, 2021

1 Recap

1.1 Orthogonality of Vectors

Let u and v be vectors of the same dimension. We say u and v are *orthogonal* iff their angle is $90°$, or equivalently $u \cdot v = u^\top v = v^\top u = 0$.
In addition, we say that a set of vectors $\{v_1, v_2, ..., v_n\}$ is *pairwise orthogonal* iff v_i and v_j are orthogonal for any $i \neq j \in \{1, 2, ..., n\}$. A set of pairwise orthogonal (nonzero) vectors is always linearly independent.
A set of vectors $\{u_1, u_2, ..., u_n\}$ is *pairwise orthonormal* if it is pairwise orthogonal, and each u_i is a unit vector.

1.2 Orthogonality of Subspaces

Two subspaces U and V of $\mathbb{R}^n$ are *orthogonal* if $u \cdot v = 0$ for all $u \in U$ and $v \in V$. In addition, it follows that $\dim U + \dim V \leq n$.

1.3 Orthogonal Complement of Subspaces

Given a subspace V, its *orthogonal complement* $V^\perp$ is defined as:

$$V^\perp = \{w : w \cdot v = 0 \text{ for any } v \in V\}.$$

Intuitively, $V^\perp$ is the largest subspace that is orthogonal to V.
Some important properties include

1. $\dim V + \dim V^\perp = n$
2. $\left(V^\perp\right)^\perp = V$

1.4 Decomposition

Theorem 1 *Let $V, W \subseteq \mathbb{R}^n$ are orthogonal complements – that is $V = W^\perp$ and $W = V^\perp$. Then every vector $x \in \mathbb{R}^n$ has a unique decomposition $x = v + w$ where $v \in V$ and $w \in W$. In addition, it follows that $v \cdot w = 0$.*

1.5 Some Familiar Orthogonal Complements

We have already seen and worked on orthogonal complements, but we just didn't realize that they are!

Theorem 2 *$N(A)$ and $C(A^\top)$ are orthogonal complements in $\mathbb{R}^n$. Similarly, $C(A)$ and $N(A^\top)$ are orthogonal complements in $\mathbb{R}^m$.*

In relation to Theorem 1, we can plug in $V = N(A)$ and $W = C(A^T)$ and derive the following result.

Theorem 3 *Suppose that we are given a matrix $A \in \mathbb{R}^{m\times n}$. Any vector $v \in \mathbb{R}^n$ can be written uniquely as $v = v_1 + v_2$ where $v_1 \in N(A)$ and $v_2 \in C(A^\top)$.*

1.6 Relationship to Projection

Suppose that we want to project a vector v onto a *unit* vector w, then the projection is

$$\text{proj}_w v = (v \cdot w)\, w.$$

We note that $v \cdot w$ is a scalar – which ensures that the projection is on w.
In general cases where w is not necessarily a unit vector, we have

$$\text{proj}_w v = \left(\frac{v \cdot w}{\|w\|^2}\right) w.$$

1.7 Gram-Schmidt

Let's suppose that we a set $\mathcal{V} = \{v_1, ..., v_k\}$ of linearly independent vectors. Our goal is to transform it into a set of orthonormal vectors $\mathcal{W}$.

Algorithm 1 GRAM-SCHMIDT
Input: a set $\mathcal{V} = \{v_1, ..., v_k\}$ of linearly independent vectors

$w_1 :=$ normalize(v_1)
$w_2 :=$ normalize$(v_2 - \text{proj}_{w_1} v_2)$
$w_3 :=$ normalize$(v_3 - \text{proj}_{w_1} v_3 - \text{proj}_{w_2} v_3)$
...
$w_k :=$ normalize $\left(v_k - \sum_{i=1}^{k-1} \text{proj}_{w_i} v_k\right)$
Output $\mathcal{W} = \{w_1, ..., w_k\}$

One crucial property is that $\mathcal{V}$ and $\mathcal{W}$ span the same subspace. In other words, if we are given a subspace $\mathcal{S}$ which is the span of basis $\mathcal{V}$, we can use Gram-Schmidt to derive its orthonormal basis $\mathcal{W}$ – meaning that $\mathcal{W}$ is a basis of $\mathcal{S}$ and is orthonormal.

2 Exercises

1. Among the following six 3-dimensional vectors, which pairs are orthogonal?

$$a = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ -6 \\ -3 \end{bmatrix}, c = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}, d = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, e = \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix}, f = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

2. Denote a subspace $V = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} : 2v_1 + 3v_3 + 5v_5 = 0 \right\}$. Find $V^{\perp}$.

3. Suppose that we have a subspace $\mathcal{S}$ with an orthongonal basis $\{v_1, ..., v_k\}$. By the definition of basis, any vector $v \in \mathcal{S}$ can be expressed as

$$v = \sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + ...\alpha_k v_k$$

for some constants $\alpha_1, ..., \alpha_k$. Determine each α_j in terms of $v_1, ..., v_k$ and v. Will the same derivation work if it not for the orthogonality of $\{v_1, ..., v_k\}$?

4. Suppose that a set of vectors $\{v_1, ..., v_n\}$ generates a subspace $\mathcal{S}$. In other words, $\mathcal{S} = \text{Span}\{v_1, ..., v_n\}$. Describe a procedure to derive an orthonormal basis of $\mathcal{S}$.

5. In this problem, we will explore the effect of ordering on the Gram-Schmidt algorithm. Denote

$$u_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \; u_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \; u_3 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix}$$

and

$$v_1 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \; v_2 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix}, \; v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

for which each $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ is a set of three linearly independent vectors. Moreover, the two sets $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ are identical, but are in different orders. This means both sets are bases of the same subspace $\mathcal{S}$.

   (a) Perform Gram-Schmidt on $\{u_1, u_2, u_3\}$ to derive an orthonormal basis of $\mathcal{S}$.

   (b) Perform Gram-Schmidt on $\{v_1, v_2, v_3\}$ to derive an orthonormal basis of $\mathcal{S}$.

   (c) Each of the answer to the previous parts is an orthonormal basis of $\mathcal{S}$. Are they identical? What can we conclude about the effect of order to the Gram-Schmidt?

3 Solutions

1. Two vectors are orthogonal iff their inner product is 0. Those pairs are $(a, c), (a, d), (a, f), (b, e), (b, f), (c, d), (c, f), (d, e), (d, f), (e, f)$.

2. There are several ways to solve this problem. First, we realize that V resides in an ambient space $\mathbb{R}^3$ and has dimension 2. This means $V^\perp$ has dimension $3 - 2 = 1$, which means that if we can find a non-zero vector $w \in V^\perp$, we can write $V^\perp = \text{Span}\{w\}$.

   Any $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in V$ must satisfy $2v_1 + 3v_2 + 5v_3 = 0$, i.e. $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$. By definition, $V^\perp = \{w : w \cdot v = 0 \text{ for any } v \in V\}$. This implies that $w = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$ works, which means $V^\perp = \text{Span}\left\{\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}\right\}$.

   Another way to solve this problem is to notice that $V = N(A)$ when $A = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix}$. This means $V^\perp = C(A^T) = \text{Span}\left\{\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}\right\}$.

3. To derive α_j, we take inner product, for both sides of the equation, with v_j.

$$\langle v, v_j \rangle = \sum_{i=1}^{k} \langle \alpha_i v_i, v_j \rangle = \alpha_j \langle v_j, v_j \rangle + \sum_{\substack{1 \le i \le k \\ i \ne j}} \alpha_i \langle v_i, v_j \rangle = \alpha_j \|v_j\|^2.$$

   This implies $\alpha_j = \frac{\langle v, v_j \rangle}{\|v_j\|^2}$. We make a crucial note that this derivation only works when $\{v_1, ..., v_k\}$ are orthogonal as we use the fact that $\langle v_i, v_j \rangle = 0$ for any $i \ne j$.

4. At the very first glance, it seems like Gram-Schmidt would do the job for us; however, it is not always the case. Recall that the input to Gram-Schmidt must be a set of linearly independent vectors, while such property is not guaranteed for $\{v_1, ..., v_n\}$.

   This means in order to derive an orthonormal basis for $\mathcal{S}$, we can do as follows.

   I. Find a basis $\mathcal{B}$ of $\mathcal{S}$ from $\{v_1, ..., v_n\}$. This can be done in several ways. We already proposed two ways of doing it in Recitation 6 – by an algorithm or by Gaussian elimination.

   II. Use $\mathcal{B}$ as an input to Gram-Schmidt which will output an orthonormal basis to $\mathcal{S}$.

5. (a) $\|u_1\| = 1$ so $w_1 = \text{normalize}(u_1) = \frac{u_1}{\|u_1\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Next, $\text{proj}_{w_1} u_2 = (u_2 \cdot w_1)w_1 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$, so $u_2 - \text{proj}_{w_1} u_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$. Therefore,

$w_2 = \text{normalize}\left(\begin{bmatrix}2\\2\\0\end{bmatrix}\right) = \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}$. Finally, $\text{proj}_{w_1}u_3 = (u_3 \cdot w_1)w_1 = \begin{bmatrix}0\\0\\1\end{bmatrix}$ and $\text{proj}_{w_2}u_3 = (u_3 \cdot w_2)w_2 = \begin{bmatrix}2\\2\\0\end{bmatrix}$, so $u_3 - \text{proj}_{w_1}u_3 - \text{proj}_{w_2}u_3 = \begin{bmatrix}4\\0\\-1\end{bmatrix} - \begin{bmatrix}0\\0\\-1\end{bmatrix} - \begin{bmatrix}2\\2\\0\end{bmatrix} = \begin{bmatrix}2\\-2\\0\end{bmatrix}$. Therefore, $w_3 = \text{normalize}\left(\begin{bmatrix}2\\-2\\0\end{bmatrix}\right) = \begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\\0\end{bmatrix}$.

Thus, Gran-Schmidt yields an orthogonal basis $\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\\0\end{bmatrix}\right\}$.

(b) $\|v_1\| = 3$ so $w_1 = \text{normalize}(v_1) = \frac{v_1}{\|v_1\|} = \begin{bmatrix}2/3\\2/3\\-1/3\end{bmatrix}$. Next, $\text{proj}_{w_1}v_2 = (v_2 \cdot w_1)w_1 = \begin{bmatrix}2\\2\\-1\end{bmatrix}$, so $v_2 - \text{proj}_{w_1}v_2 = \begin{bmatrix}4\\0\\-1\end{bmatrix} - \begin{bmatrix}2\\2\\-1\end{bmatrix} = \begin{bmatrix}2\\-2\\0\end{bmatrix}$. Therefore, $w_2 = \text{normalize}\left(\begin{bmatrix}2\\-2\\0\end{bmatrix}\right) = \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\\0\end{bmatrix}$. Finally, $\text{proj}_{w_1}v_3 = (v_3 \cdot w_1)w_1 = \begin{bmatrix}-2/9\\-2/9\\1/9\end{bmatrix}$ and $\text{proj}_{w_2}v_3 = (v_3 \cdot w_2)w_2 = \begin{bmatrix}0\\0\\0\end{bmatrix}$, so $v_3 - \text{proj}_{w_1}v_3 - \text{proj}_{w_2}v_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} - \begin{bmatrix}-2/9\\-2/9\\1/9\end{bmatrix} - \begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix}2/9\\2/9\\8/9\end{bmatrix}$. Therefore, $w_3 = \text{normalize}\left(\begin{bmatrix}2/9\\2/9\\8/9\end{bmatrix}\right) = \begin{bmatrix}1/\sqrt{18}\\1/\sqrt{18}\\4/\sqrt{18}\end{bmatrix}$.

Thus, Gran-Schmidt yields an orthogonal basis $\left\{\begin{bmatrix}2/3\\2/3\\-1/3\end{bmatrix}, \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\\0\end{bmatrix}, \begin{bmatrix}1/\sqrt{18}\\1/\sqrt{18}\\4/\sqrt{18}\end{bmatrix}\right\}$.

(c) As we can see, the answers from two parts are not identical. This means the ordering of vectors affects the output orthogonal basis of the subspace.