## MATH 21B, FEBRUARY 23: ORTHONORMAL BASES AND PROJECTION

Definition: Let $\vec{v}, \vec{w} \in \mathbb{R}^{n}$. Then we say $\vec{v}$ and $\vec{w}$ are orthogonal if the dot product

$$
\vec{v} \cdot \vec{w}=\left[\begin{array}{lll}
v_{1} & \cdots & v_{n}
\end{array}\right]\left[\begin{array}{c}
v_{1} \\
\vdots \\
v_{n}
\end{array}\right]=v_{1} w_{1}+\ldots+v_{n} w_{n}
$$

is equal to 0 . In general, the dot product of $\vec{v}$ and $\vec{w}$ is equal to $|\vec{v}||\vec{w}| \cos \theta$ where $\theta$ is the angle between $\vec{v}$ and $\vec{w}$. The number $\cos \theta$ is called the correlation coefficient: when it is positive, the vectors are positively correlated, when it is negative, the vectors are negatively correlated, and when it is zero, the vectors are orthogonal. Note that $\vec{v} \cdot \vec{v}=|\vec{v}|^{2}$ has correlation coefficient 1 .
(1) Suppose that $\vec{u}_{1}, \ldots, \vec{u}_{m}$ is a set of nonzero vectors in $\mathbb{R}^{n}$ such that each is orthogonal to all of the others. Must they be linearly independent? How do you know?

Solution: $\vec{u}_{1}, \ldots, \vec{u}_{m}$ must be linearly independent. Here's a nice trick to prove it. Suppose that you have some linear relation among these vectors, i.e.

$$
c_{1} \vec{u}_{1}+c_{2} \vec{u}_{2}+\ldots+c_{m} \vec{u}_{m}=\overrightarrow{0}
$$

Take the dot product of both sides with one of the $\vec{u}_{i}$ 's.

$$
c_{1}\left(\vec{u}_{1} \cdot \vec{u}_{i}\right)+c_{2}\left(\vec{u}_{2} \cdot \vec{u}_{i}\right)+\ldots+c_{m}\left(\vec{u}_{m} \cdot \vec{u}_{i}\right)=\overrightarrow{0} \cdot \vec{u}_{i}=0
$$

Because $\vec{u}_{1}, \ldots, \vec{u}_{m}$ are mutually orthogonal, every term except for one on the left hand side, is equal to zero

$$
c_{i}\left(\vec{u}_{i} \cdot \vec{u}_{i}\right)=0
$$

Since $\vec{u}_{i} \cdot \vec{u}_{i}=\left|\vec{u}_{i}\right|^{2}$ is nonzero, it follows that $c_{i}$ is equal to zero. We can use this argument for $i=1,2, \ldots, m$ to conclude that $c_{1}=c_{2}=\cdots=c_{m}=0$.

We have shown that for any linear relation, we can show $c_{1}=c_{2}=\cdots=c_{m}=0$. It follows that there are no nontrivial linear relations, and therefore, the vectors are independent.

Definition: A set of vectors $\vec{u}_{1}, \ldots, \vec{u}_{m}$ in $V$ is called an orthonormal basis for $V$ if they form a basis for $V$, each has length 1 , and each is orthogonal to all of the others. Expressing this condition another way,

$$
\vec{u}_{i} \cdot \vec{u}_{j}=\left\{\begin{array}{lll}
1 & \text { if } \quad i=j \\
0 & \text { if } \quad i \neq j
\end{array}\right.
$$

(2) Let $V$ be a subspace of $\mathbb{R}^{n}$. Any vector $\vec{x}$ in $\mathbb{R}^{n}$ can be written as $\vec{x}=\operatorname{proj}_{V}(\vec{x})+\vec{x}^{\perp}$, where $\operatorname{proj}_{V}(\vec{x})$ is the orthogonal projection of $\vec{x}$ onto $V$, and $\vec{x}^{\perp}$ is orthogonal to $V$. Suppose that we have an orthonormal basis $\left(\vec{u}_{1}, \ldots, \vec{u}_{m}\right)$ for $V$.
(a) Explain why $\operatorname{proj}_{V}(\vec{x})$ can be written as $c_{1} \vec{u}_{1}+\ldots+c_{m} \vec{u}_{m}$ for some scalars $c_{1}, \ldots, c_{m}$.

Solution: $\operatorname{proj}_{V}(\vec{x})$ is the orthogonal projection of $\vec{x}$ onto $V$, and therefore lies in $V$. Since $\vec{u}_{1}, \ldots, \vec{u}_{m}$ form a basis for $V, \operatorname{proj}_{V}(\vec{x})$ can be written as a linear combination $c_{1} \vec{u}_{1}+\ldots+c_{m} \vec{u}_{m}$ of these vectors.
(b) How do you calculate these coefficients $c_{i}$ in terms of $\vec{x}$ and $\vec{u}_{1}, \ldots, \vec{u}_{m}$ ? (Hint: Take the equation $\vec{x}=\left(c_{1} \vec{u}_{1}+\ldots+c_{m} \vec{u}_{m}\right)+\vec{x}^{\perp}$ and take the dot product of both sides with $\vec{u}_{i}$ 's.)

Solution: As suggested, we take the equation

$$
\vec{x}=\operatorname{proj}_{V}(\vec{x})+\vec{x}^{\perp}=\left(c_{1} \vec{u}_{1}+\ldots+c_{m} \vec{u}_{m}\right)+\vec{x}^{\perp}
$$

and dot both sides with $\vec{u}_{i}$.

$$
\left.\vec{u}_{i} \cdot \vec{x}=c_{1}\left(\vec{u}_{i} \cdot \vec{u}_{1}\right)+c_{2}\left(\vec{u}_{i} \cdot \vec{u}_{2}\right)+\ldots+c_{m}\left(\vec{u}_{i} \cdot \vec{u}_{m}\right)\right)+\vec{u}_{i} \cdot \vec{x}^{\perp}
$$

Since $\vec{x}^{\perp}$ is orthogonal to any vector in $V, \vec{u}_{i} \cdot \vec{x}^{\perp}=0$. And since the basis is orthonormal, we get the simplification

$$
\vec{u}_{i} \cdot \vec{x}=c_{1}(0)+c_{2}(0)+\ldots+c_{i}(1)+\ldots+c_{m}(0)+0=c_{i}
$$

Therefore, for each $i=1,2, \ldots, m, \overrightarrow{u_{i} \cdot \vec{x}=c_{i}}$.
(c) Use this to write a formula for $\operatorname{proj}_{V}(\vec{x})$ in terms of $\vec{x}$ and our basis $\vec{u}_{1}, \ldots, \vec{u}_{m}$. At what step did you require the basis to be orthonormal?

Solution: Since $\operatorname{proj}_{V}(\vec{x})=c_{1} \vec{u}_{1}+\ldots+c_{m} \vec{u}_{m}$, we have the formula

$$
\operatorname{proj}_{V}(\vec{x})=\left(\vec{u}_{1} \cdot \vec{x}\right) \vec{u}_{1}+\left(\vec{u}_{2} \cdot \vec{x}\right) \vec{u}_{2}+\ldots+\left(\vec{u}_{m} \cdot \vec{x}\right) \vec{u}_{m}
$$

(3) Let $V$ be the plane $2 x+2 y+z=0, \vec{u}_{1}=\left[\begin{array}{c}1 / 3 \\ -2 / 3 \\ 2 / 3\end{array}\right]$, and $\vec{u}_{2}=\left[\begin{array}{c}-2 / 3 \\ 1 / 3 \\ 2 / 3\end{array}\right]$. Let $\vec{x}=\left[\begin{array}{l}1 \\ 4 \\ 8\end{array}\right]$.
(a) Verify that $\left(\vec{u}_{1}, \vec{u}_{2}\right)$ is an orthonormal basis of $V$.

Solution: First, we need to verify that $\vec{u}_{1}, \vec{u}_{2}$ lie in the plane $V$.

$$
2(1 / 3)+2(-2 / 3)+1(2 / 3)=0 \quad 2(-2 / 3)+2(1 / 3)+1(2 / 3)=0
$$

Then we need to verify that $\vec{u}_{1}, \vec{u}_{2}$ both have length 1 and are orthogonal to each other.

$$
\begin{gathered}
\vec{u}_{1} \cdot \vec{u}_{1}=(1 / 3)^{2}+(-2 / 3)^{2}+(2 / 3)^{2}=1 / 9+4 / 9+4 / 9=1 \\
\vec{u}_{2} \cdot \vec{u}_{2}=(-2 / 3)^{2}+(1 / 3)^{2}+(2 / 3)^{2}=4 / 9+1 / 9+4 / 9=1 \\
\vec{u}_{1} \cdot \vec{u}_{2}=(1 / 3)(-2 / 3)+(-2 / 3)(1 / 3)+(2 / 3)(2 / 3)=-2 / 9-2 / 9+4 / 9=0
\end{gathered}
$$

(b) Find $\operatorname{proj}_{V}(\vec{x})$. (Check that your answer is reasonable by computing the difference $\vec{x}-\operatorname{proj}_{V}(\vec{x})$. What should be true about this vector?)

Solution: We use the formula from the previous question.

$$
\operatorname{proj}_{V}(\vec{x})=\left(\vec{u}_{1} \cdot \vec{x}\right) \vec{u}_{1}+\left(\vec{u}_{2} \cdot \vec{x}\right) \vec{u}_{2}=\left(\left[\begin{array}{c}
1 / 3 \\
-2 / 3 \\
2 / 3
\end{array}\right] \cdot\left[\begin{array}{l}
1 \\
4 \\
8
\end{array}\right]\right)\left[\begin{array}{c}
1 / 3 \\
-2 / 3 \\
2 / 3
\end{array}\right]+\left(\left[\begin{array}{c}
-2 / 3 \\
1 / 3 \\
2 / 3
\end{array}\right] \cdot\left[\begin{array}{l}
1 \\
4 \\
8
\end{array}\right]\right)\left[\begin{array}{c}
-2 / 3 \\
1 / 3 \\
2 / 3
\end{array}\right]
$$

Straightforward computation gives $\vec{u}_{1} \cdot \vec{x}=3$ and $\vec{u}_{2} \cdot \vec{x}=6$. So

$$
\operatorname{proj}_{V}(\vec{x})=3\left[\begin{array}{c}
1 / 3 \\
-2 / 3 \\
2 / 3
\end{array}\right]+6\left[\begin{array}{c}
-2 / 3 \\
1 / 3 \\
2 / 3
\end{array}\right]=\left[\begin{array}{c}
-3 \\
0 \\
6
\end{array}\right]
$$

We check that this is reasonable by computing $\vec{x}^{\perp}=\vec{x}-\operatorname{proj}_{V}(\vec{x})=\left[\begin{array}{l}4 \\ 4 \\ 2\end{array}\right]$. This vector is supposed to be orthogonal to $V$, and indeed, it's a multiple of the normal vector.
(4) Let $V$ be an $m$-dimensional subspace of $\mathbb{R}^{n}$. Consider the linear transformation proj $_{V}$ : $\mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$.
(a) What is $\operatorname{im}\left(\operatorname{proj}_{V}\right)$ ? What is its dimension?

Solution: The image of projection onto $V$ is the subspace $V$ (in $\mathbb{R}^{n}$ ). By definition, it has dimension $m$.
(b) What is $\operatorname{ker}\left(\operatorname{proj}_{V}\right)$ ? What is its dimension?

Solution: The kernel is the space of vectors whose projection onto $V$ is zero. This is the set of vectors orthogonal to $V$, called the orthogonal complement (denoted $V^{\perp}$ ). It has dimension $n-m$.
(c) If you have a basis $\vec{u}_{1}, \ldots, \vec{u}_{m}$ for $V$, how would you calculate $\operatorname{ker}\left(\operatorname{proj}_{V}\right)$ ?

Solution: A vector $\vec{x}$ is in $V^{\perp}$ if and only if $\vec{u}_{1} \cdot \vec{x}=\cdots=\vec{u}_{m} \cdot \vec{x}$. That is, $\vec{x}$ is in $V^{\perp}$ if and only if

$$
\left[\begin{array}{c}
\vec{u}_{1}^{T} \\
\vdots \\
\vec{u}_{m}^{T}
\end{array}\right] \vec{x}=\left[\begin{array}{c}
0 \\
\vdots \\
0
\end{array}\right]
$$

where $\vec{u}_{i}^{T}$ just means the vector $\vec{u}_{i}$ written as a row vector (so the above is a matrix times vector equals vector equation). Therefore, $V^{\perp}$ is the kernel of the matrix $\left[\begin{array}{c}\vec{u}_{1}^{T} \\ \vdots \\ \vec{u}_{m}^{T}\end{array}\right]$ - we know how to compute kernels.
(5) Suppose $\mathfrak{B}=\left(\vec{u}_{1}, \vec{u}_{2}, \vec{u}_{3}\right)$ is an orthonormal basis of $\mathbb{R}^{3}$. In each part of this problem, you are given the $\mathfrak{B}$-matrix of a linear transformation $T: \mathbb{R}^{3} \rightarrow \mathbb{R}^{3}$. Describe the linear transformation geometrically.
(a) $\left[\begin{array}{ccc}0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1\end{array}\right]$
(b) $\left[\begin{array}{lll}1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1\end{array}\right]$

Solution: This matrix sends $\vec{u}_{1} \mapsto \vec{u}_{2}$, $\vec{u}_{2} \mapsto-\vec{u}_{1}$, and $\vec{u}_{3} \mapsto \vec{u}_{3}$. So in the plane

Solution: This matrix sends $\vec{u}_{1} \mapsto \vec{u}_{1}$, $\vec{u}_{2} \mapsto \overrightarrow{0}$, and $\vec{u}_{3} \mapsto \vec{u}_{3}$. Thus, it is projecdefined by $\vec{u}_{1}, \vec{u}_{2}$, it is a rotation by $\pi / 2$, tion onto the plane spanned by $\vec{u}_{1}$ and $\vec{u}_{3}$. and it leaves the vector $\vec{u}_{3}$ fixed. Therefore, it is rotation by $\pi / 2$ around the line $\left\langle\vec{u}_{3}\right\rangle$.

If $\mathfrak{B}$ were not orthonormal, how would your answers change?
Solution: A rotation has to preserve all lengths and angles, but if $\vec{u}_{1}, \vec{u}_{2}$ have different lengths, then (a) doesn't preserve lengths, and if $\vec{u}_{1}, \vec{u}_{2}$ are not orthogonal to each other, then (a) doesn't preserve angles. Therefore, the description for (a) heavily depends on orthonormality. For (b), it can still be described as projection onto the plane $\left\langle\vec{u}_{1}, \vec{u}_{3}\right\rangle$ as long as $\vec{u}_{2}$ is orthogonal to $\left\langle\vec{u}_{1}, \vec{u}_{3}\right\rangle$.
(6) Let $\vec{v}_{1}=\left[\begin{array}{l}1 \\ 1 \\ 1 \\ 1\end{array}\right], \vec{v}_{2}=\left[\begin{array}{c}3 \\ -1 \\ -1 \\ 3\end{array}\right], \vec{v}_{3}=\left[\begin{array}{c}1 \\ 3 \\ 1 \\ -1\end{array}\right]$; these three vectors are linearly independent. Let $V$ be the subspace of $\mathbb{R}^{4}$ spanned by $\vec{v}_{1}, \vec{v}_{2}, \vec{v}_{3}$. Find an orthonormal basis of $V$.

NOTE: this problem involves the Gram-Schmidt orthonormalization procedure, which you do NOT need to know for the exam. We will talk about this in detail on Thursday March 2.

Solution: In the Gram-Schmidt process, we first 'turn $\vec{v}_{1}$ into a unit vector' by dividing it by its length. That is, we let $\vec{u}_{1}=\frac{1}{\left\|\vec{v}_{2}\right\|} \vec{v}_{1}$. Here, $\left\|\vec{v}_{1}\right\|=\sqrt{\vec{v}_{1} \cdot \vec{v}_{1}}=2$, so

$$
\vec{u}_{1}=\left[\begin{array}{l}
1 / 2 \\
1 / 2 \\
1 / 2 \\
1 / 2
\end{array}\right]
$$

Next, we want to get a vector $\vec{v}_{2}^{\perp}$ in $\left\langle\vec{v}_{1}, \vec{v}_{2}\right\rangle$ which is perpendicular to $\vec{u}_{1}$. We do this by letting $\vec{v}_{2}^{\perp}=\vec{v}_{2}-\operatorname{proj}_{\left\langle\vec{u}_{1}\right\rangle}\left(\vec{v}_{2}\right)$, or

$$
\vec{v}_{2}^{\perp}=\vec{v}_{2}-\left(\vec{u}_{1} \cdot \vec{v}_{2}\right) \vec{u}_{1}=\left[\begin{array}{c}
3 \\
-1 \\
-1 \\
3
\end{array}\right]-2\left[\begin{array}{c}
1 / 2 \\
1 / 2 \\
1 / 2 \\
1 / 2
\end{array}\right]=\left[\begin{array}{c}
2 \\
-2 \\
-2 \\
2
\end{array}\right]
$$

(Note that $\vec{v}_{2}^{\perp}$ is indeed orthogonal to $\vec{u}_{1}$.) We turn this into a unit vector by dividing by its length.

$$
\vec{u}_{2}=\frac{1}{\left\|\vec{v}_{2}^{\perp}\right\|} \vec{v}_{2}^{\perp}=\left[\begin{array}{c}
1 / 2 \\
-1 / 2 \\
-1 / 2 \\
1 / 2
\end{array}\right]
$$

Next, we want a vector $\vec{v}_{3}^{\perp}$ in $\left\langle\vec{v}_{1}, \vec{v}_{2}, \vec{v}_{3}\right\rangle$ which is orthogonal to $\vec{u}_{1}, \vec{u}_{2}$. We get this by letting $\vec{v}_{3}^{\perp}=\vec{v}_{3}-\operatorname{proj}_{\left\langle\vec{v}_{1}, \vec{v}_{2}\right\rangle}\left(\vec{v}_{3}\right)$. To compute this projection, we need an orthonormal basis of $\left\langle\vec{v}_{1}, \vec{v}_{2}\right\rangle$, but we have one: $\vec{u}_{1}, \vec{u}_{2}$ ! So

$$
\vec{v}_{3}^{\perp}=\vec{v}_{3}-\left(\vec{u}_{1} \cdot \vec{v}_{3}\right) \vec{u}_{1}-\left(\vec{u}_{2} \cdot \vec{v}_{3}\right) \vec{u}_{2}=\left[\begin{array}{c}
1 \\
1 \\
-1 \\
-1
\end{array}\right]
$$

(It's easy to check at this point that $\vec{v}_{3}^{\perp}$ is really orthogonal to $\vec{u}_{1}$ and $\vec{u}_{2}$.) Finally, we turn $\vec{v}_{3}^{\perp}$ into a unit vector by dividing by its length

$$
\vec{u}_{3}=\frac{1}{\left\|\vec{v}_{3}^{\perp}\right\|} \vec{v}_{3}^{\perp}=\left[\begin{array}{c}
1 / 2 \\
1 / 2 \\
-1 / 2 \\
-1 / 2
\end{array}\right]
$$

Thus, an orthonormal basis of $V$ is

$$
\vec{u}_{1}=\left[\begin{array}{l}
1 / 2 \\
1 / 2 \\
1 / 2 \\
1 / 2
\end{array}\right], \vec{u}_{2}=\left[\begin{array}{c}
1 / 2 \\
-1 / 2 \\
-1 / 2 \\
1 / 2
\end{array}\right], \vec{u}_{3}=\left[\begin{array}{c}
1 / 2 \\
1 / 2 \\
-1 / 2 \\
-1 / 2
\end{array}\right]
$$

Check for yourself that this basis is orthonormal!

