Math 313 Lecture #28
§6.4: The Gram-Schmidt Process

We have seen that many computations in linear algebra can be accomplished by the inner product when we have an orthogonal basis for a vector space.

With the the use of the orthogonal projection onto a subspace, we can obtain an orthogonal basis from a given basis.

Example. Let W be the three dimensional subspace of $\mathbb{R}^4$ with the basis

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ \vec{x}_2 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix}, \ \vec{x}_3 = \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix}.$$

We seek an orthogonal basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ for W.

<u>Step 1</u>. We start by taking $\vec{v}_1 = \vec{x}_1$, for which $\operatorname{Span}\{\vec{v}_1\} = \operatorname{Span}\{\vec{x}_1\}$.

<u>Step 2</u>. We want to find $\vec{v}_2$ such that $\vec{v}_2 \perp \vec{v}_1$ and $\operatorname{Span}\{\vec{v}_1, \vec{v}_2\} = \operatorname{Span}\{\vec{x}_1, \vec{x}_2\}$.

The orthogonal projection of $\vec{x}_2$ onto the subspace $\operatorname{Span}\{\vec{v}_1\}$ is

$$\vec{p}_1 = \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 = \frac{6}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 3/2 \\ 3/2 \\ 3/2 \end{bmatrix}.$$

The vector

$$\vec{v}_2 = \vec{x}_2 - \vec{p}_1 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix} - \begin{bmatrix} 3/2 \\ 3/2 \\ 3/2 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{bmatrix}$$

belongs to $\operatorname{Span}\{\vec{x}_1, \vec{x}_2\}$ and is orthogonal to $\vec{v}_1$.

How do we check $\vec{v}_2$? It is orthogonal to $\vec{v}_1$? Yes it is.

<u>Step 3</u>. We want to find $\vec{v}_3$ such that $\vec{v}_3$ is orthogonal to $\operatorname{Span}\{\vec{v}_1, \vec{v}_2\}$ and $W = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

The orthogonal projection of $\vec{x}_3$ onto the subspace $\operatorname{Span}\{\vec{v}_1, \vec{v}_2\}$ is

$$\vec{p}_2 = \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 = \frac{4}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{10}{25}\begin{bmatrix} -5/2 \\ 5/2 \\ 5./2 \\ 5/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix}.$$

The vector

$$\vec{v}_3 = \vec{x}_3 - \vec{p}_2 = \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix}$$

belongs to Span$\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ and belongs to the orthogonal complement of Span$\{\vec{v}_1, \vec{v}_2\}$. How do we check $\vec{v}_3$? Is it orthogonal to both $\vec{v}_1$ and $\vec{v}_2$? Yes it is.

Thus an orthogonal basis for W is the set containing the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{bmatrix}, \ \vec{v}_3 = \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix}.$$

If there were more than 3 basis vectors given, we could continue this process to obtain an orthogonal basis.

Theorem 11 (The Gram-Schmidt Process). Let $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p\}$ be a basis for a nonzero subspace W of $\mathbb{R}^n$. The vectors

$$\begin{aligned}
\vec{v}_1 &= \vec{x}_1, \\
\vec{v}_2 &= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1, \\
\vec{v}_3 &= \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 \\
&\vdots \\
\vec{v}_p &= \vec{x}_p - \frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_p \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 - \cdots - \frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\vec{v}_{p-1},
\end{aligned}$$

form an orthogonal basis for W where for each $k = 1, \ldots, p$ we have

$$\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{Span}\{\vec{x}_1, \ldots, \vec{x}_k\}.$$

Even better than an orthogonal basis is an orthonormal basis.

A simple adjustment to the Gram-Schmidt process gives an orthonormal basis.

Example. We look at the same three vectors we started the first example with, but write them as the columns of

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix}.$$

The columns of A are linearly independent and therefore they form a basis for the column space of A.

We adapt the Gram-Schmidt process to the columns $\vec{a}_1, \vec{a}_2, \vec{a}_3$ of A to obtain an orthonormal basis $\vec{q}_1, \vec{q}_2, \vec{q}_3$ of the column space of A.

With $r_{11} = \|\vec{a}_1\| = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2$, we have

$$\vec{q}_1 = \frac{\vec{a}_1}{\|\vec{a}_1\|} = \frac{\vec{a}_1}{r_{11}} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}.$$

With $r_{12} = \vec{a}_2 \cdot \vec{q}_1 = 3$, the projection of $\vec{a}_2$ onto $\text{Span}\{\vec{q}_1\} = \text{Span}\{\vec{a}_1\}$ is

$$\vec{p}_1 = (\vec{a}_2 \cdot \vec{q}_1)\vec{q}_1 = r_{12}\vec{q}_1 = 3\vec{q}_1.$$

The vector orthogonal to $\text{Span}(\vec{q}_1)$ is

$$\vec{a}_2 - \vec{p}_1 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix} - \begin{bmatrix} 3/2 \\ 3/2 \\ 3/2 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{bmatrix}.$$

With $r_{22} = \|\vec{a}_2 - \vec{p}_1\| = 5$, the second vector in the orthonormal basis is

$$\vec{q}_2 = \frac{\vec{a}_2 - \vec{p}_1}{\|\vec{a}_2 - \vec{p}_1\|} = \frac{\vec{a}_2 - \vec{p}_1}{r_{22}} = \frac{1}{5}\begin{bmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{bmatrix}.$$

With $r_{13} = \vec{a}_3 \cdot \vec{q}_1 = 2$ and $r_{23} = \vec{a}_3 \cdot \vec{q}_2 = -2$, the projection of $\vec{a}_3$ onto $\text{Span}(\vec{q}_1, \vec{q}_2)$ is

$$\begin{aligned} \vec{p}_2 &= (\vec{a}_3 \cdot \vec{q}_1)\vec{q}_1 + (\vec{a}_3 \cdot \vec{q}_2)\vec{q}_2 \\ &= r_{13}\vec{q}_1 + r_{23}\vec{q}_2 \\ &= 2\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} - 2\begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix}. \end{aligned}$$

The vector orthogonal to $\text{Span}\{\vec{q}_1, \vec{q}_2\}$ is

$$\vec{a}_3 - \vec{p}_2 = \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix}.$$

With $r_{33} = \|\vec{a}_3 - \vec{p}_2\| = 4$, the third and last unit vector in the orthonormal basis is

$$\vec{q}_3 = \frac{\vec{a}_3 - \vec{p}_2}{\|\vec{a}_3 - \vec{p}_2\|} = \frac{\vec{a}_3 - \vec{p}_2}{r_{33}} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}.$$

You might not have noticed it, but we have been expressing A as a product of two matrices R and Q.

Define an upper triangular square matrix R by the inner products and norms r_{ij}:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{bmatrix}.$$

This square matrix is invertible (and this always happens – why?).

Define a matrix Q by the orthonormal basis vectors $\vec{q}_1, \vec{q}_1, \vec{q}_3$ as its the columns:

$$Q = (\vec{q}_1, \vec{q}_2, \vec{q}_3) = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \end{bmatrix}.$$

The columns of Q form an orthonormal basis for the column space of A.

You might not be surprised that the product of Q with R is the original matrix A:

$$QR = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix}.$$

This is called the QR factorization of A. $\square$

Theorem 11. If A is an $m \times n$ matrix with linearly independent columns, then there is an $m \times n$ matrix Q whose columns form an orthonormal set, and an invertible $n \times n$ upper triangular matrix R with positive numbers on its diagonal such that

$$A = QR.$$