Lecture 23: 6.2 Orthogonal Projections and Gram-Schmidt.

Ex Suppose that W is a vector subspace of $\mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$. If $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$, then we say that $\hat{\mathbf{y}}$ is the **orthogonal projection** of $\mathbf{y}$ onto W.
Example: Suppose $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbf{R}^3$ and let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Suppose $\mathbf{y} \in \mathbb{R}^3$. By a previous theorem we can write

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3.$$

Let $\widehat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \dfrac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$. Then $\widehat{\mathbf{y}} \in W$ and $\mathbf{z} = \dfrac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3$ is orthogonal to W, since $\mathbf{z}\cdot\mathbf{u}_1 = \mathbf{z}\cdot\mathbf{u}_2 = 0$.
$\widehat{\mathbf{y}}$ is called the **projection of y onto** W

Ex Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and let $\mathbf{y} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$.
Write $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$, where $\widehat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$.
Sol Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis it follows that

$$\widehat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1} + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2} = \frac{10}{10}\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \frac{3}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix}.$$

$$\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} - \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

The orthogonal decomposition theorem Let W be a subspace of $\mathbf{R}^n$ and suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W. Any $\mathbf{y} \in \mathbf{R}^n$ can be written uniquely as

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z},$$

where

$$\widehat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p$$

and $\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}} \in W^\perp$, the orthogonal complement $W^\perp = \{\mathbf{z} \in \mathbf{R}^n; \mathbf{z}\cdot\mathbf{u}_1 = 0, \ldots, \mathbf{z}\cdot\mathbf{u}_p = 0\}$. $\widehat{\mathbf{y}} = \text{proj}_W\mathbf{y}$ is called the **orthogonal projection of y onto** W.

The best approximation theorem Let W be a subspace of $\mathbf{R}^n$, $\mathbf{y}$ a vector and $\widehat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\widehat{\mathbf{y}}$ is the point in W closest to $\mathbf{y}$:

$$\|\mathbf{y} - \widehat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|, \qquad \mathbf{v} \in W, \quad \mathbf{v} \neq \widehat{\mathbf{y}}.$$

Pf We can write

$$\mathbf{y} - \mathbf{v} = \mathbf{y} - \widehat{\mathbf{y}} + \widehat{\mathbf{y}} - \mathbf{v}$$

where $\mathbf{y}-\widehat{\mathbf{y}} \in W^{\perp}$ and $\widehat{\mathbf{y}}-\mathbf{v} \in W$ are orthogonal and hence by the Pythagorean theorem:

$$||\mathbf{y}-\mathbf{v}\|^2 = \|\mathbf{y}-\widehat{\mathbf{y}}\|^2 + \|\widehat{\mathbf{y}}-\mathbf{v}\|^2 > \|\mathbf{y}-\widehat{\mathbf{y}}\|^2.$$

Ex Find the closest point to $\mathbf{y}=\begin{bmatrix} 2 \\ 4 \\ 0 \\ -2 \end{bmatrix}$ to $W=\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, $\mathbf{u}_1=\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2=\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$.

Sol $\widehat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \dfrac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = 3\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ -1 \\ -1 \end{bmatrix}$.

Th Suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for W, i.e. $\mathbf{u}.\mathbf{u}_j = \delta_{ij}$. Then

$$\mathrm{proj}_W\mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p$$

6.4 The Gram-Smith Orthogonalization Process.

In this section we will learn a process for constructing an orthonormal basis for subspace W of $\mathbf{R}^m$. We start with any basis $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ for W and from it we will use projections to construct an orthonormal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ for W.
We will construct the $\mathbf{u}_i$'s inductively so that $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ are orthonormal and

$$\mathrm{Span}(\mathbf{u}_1, \cdots, \mathbf{u}_k) = \mathrm{Span}(\mathbf{x}_1, \cdots, \mathbf{x}_k) = W_k$$

for $k = 1, ..., n$. To begin the process, let

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\, \mathbf{x}_1$$

Then $\mathrm{Span}(\mathbf{u}_1) = \mathrm{Span}(\mathbf{x}_1)$, since $\mathbf{u}_1$ is a multiple of $\mathbf{x}_1$ and $\|\mathbf{u}_1\| = 1$.
Let $\mathbf{p}_1$ be the projection of $\mathbf{x}_2$ onto $\mathrm{Span}(\mathbf{x}_1) = \mathrm{Span}(\mathbf{u}_1)$, i.e. by section 5.5:

$$\mathbf{p}_1 = \langle \mathbf{x}_2, \mathbf{u}_1 \rangle\, \mathbf{u}_1, \qquad \mathbf{x}_2 - \mathbf{p}_1 \in \mathrm{Span}(\mathbf{u}_1)^{\perp}$$

Then $\mathbf{x}_2 - \mathbf{p}_1 \neq \mathbf{0}$ since $\mathbf{x}_2 \notin \mathrm{Span}(\mathbf{u}_1)$. If we set

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}_1\|}\, (\mathbf{x}_2 - \mathbf{p}_1)$$

then $\mathbf{u}_2$ is a unit vector orthogonal to $\mathrm{Span}(\mathbf{u}_1)$ and $\mathrm{Span}(\mathbf{u}_1, \mathbf{u}_2) = \mathrm{Span}(\mathbf{x}_1, \mathbf{x}_2)$.
To construct $\mathbf{u}_3$ let $\mathbf{p}_3$ be the projection of $\mathbf{x}_3$ into $\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$:

$$\mathbf{p}_3 = \langle \mathbf{x}_3, \mathbf{u}_1 \rangle\, \mathbf{u}_1 + \langle \mathbf{x}_3, \mathbf{u}_2 \rangle\, \mathbf{u}_2$$

and set

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3 - \mathbf{p}_3\|}\, (\mathbf{x}_3 - \mathbf{p}_3)$$

In general we define $\mathbf{u}_k$ recursively by

$$\mathbf{u}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1} - \mathbf{p}_{k+1}\|}\, (\mathbf{x}_{k+1} - \mathbf{p}_{k+1})$$

where

$$\mathbf{p}_{k+1} = \langle \mathbf{x}_{k+1}, \mathbf{u}_1 \rangle\, \mathbf{u}_1 + \cdots + \langle \mathbf{x}_{k+1}, \mathbf{u}_k \rangle\, \mathbf{u}_k$$

is the projection of $\mathbf{x}_{k+1}$ onto $\mathrm{Span}(\mathbf{u}_1, \cdots, \mathbf{u}_k)$.
This procedure, called the **Gram-Smith orthogonalization process** yields an orthonormal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ for W.

Ex Find an orthonormal basis for the plane $F = \{\mathbf{x} \in \mathbf{R}^3;\, x_1 + x_2 + x_3 = 0\}$.
Sol $\mathbf{x}_1 = (1, -1, 0)^T$ and $\mathbf{x}_2 = (1, 0, -1)^T$ are two vectors in the plane. First let

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\, \mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Then let

$$\mathbf{p}_1 = \langle \mathbf{x}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Since

$$\mathbf{x}_2 - \mathbf{p}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

we get

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}_1\|} \big(\mathbf{x}_2 - \mathbf{p}_1\big) = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

One can also use the Gram-Smith process to obtain the so called QR factorization of a matrix $A = QR$, where the column vectors of Q are orthonormal and R is upper triangular. In fact if A is an $m \times n$ matrix such that the n column vectors of $A = \big[\mathbf{x}_1 \cdots \mathbf{x}_n\big]$ form a basis we can perform the Gram-Smith process on these to obtain an orthonormal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ such that $\mathrm{Span}\big(\mathbf{u}_1, \cdots, \mathbf{u}_k\big) = \mathrm{Span}\big(\mathbf{x}_1, \cdots, \mathbf{x}_k\big)$, for $k = 1, ..., n$. Hence for some constants r_{ij}

$$\mathbf{x}_k = r_{1k}\mathbf{u}_1 + \cdots + \mathbf{r}_{kk}\mathbf{u}_k + 0\mathbf{u}_{k+1} + \cdots 0\mathbf{u}_n, \qquad k = 1, ..., n.$$

Let R be the upper triangular matrix with column vectors defined by

$$R = \big[\mathbf{r}_1 \cdots \mathbf{r}_n\big], \qquad \text{where} \qquad \mathbf{r}_k = \begin{bmatrix} r_{1k} \\ \cdot \\ r_{kk} \\ 0 \\ \cdot \\ 0 \end{bmatrix},$$

and let $Q = \big[\mathbf{u}_1 \cdots \mathbf{u}_n\big]$. Then

$$Q\mathbf{r}_k = r_{1k}\mathbf{u}_1 + \cdots + r_{kk}\mathbf{u}_k = \mathbf{x}_k$$

and hence

$$QR = \big[Q\mathbf{r}_1 \cdots Q\mathbf{r}_n\big] = \big[\mathbf{x}_1 \cdots \mathbf{x}_n\big] = A.$$

Note in principle one can calculate what R from the Gram-Smith process, but it is simpler to get R just from using that $A = QR$ so since $Q^T Q = I$ and

$$R = Q^I QR = Q^T A.$$

Ex Find the QR factorization of $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 3 \end{bmatrix}$.

Sol Use Gram Schmidt on the columns of A to find and orthonormal basis and from construct $Q = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{bmatrix}$. From it let $R = Q^T A = \cdots = \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 0 & 3 \end{bmatrix}$.