5.6 Gram- Schmidt

Let $(V, \langle , \rangle)$ be an n-dim inner product space, and suppose $\{\underline{x}_1, \ldots, \underline{x}_n\}$ is any basis of the vector space V.

We want to use $\{\underline{x}_i\}$ to construct a new basis of V that respects the inner product $\langle , \rangle$ in the sense that it is orthonormal.

This new basis $\{\underline{u}_1, \ldots, \underline{u}_n\}$ is obtained by the following Gram-Schmidt Orthogonalization Process.

Idea: we want all of the $\underline{u}_i$ to be unit vectors, and we also want them to be orthogonal to each other:

1. $\|\underline{u}_i\| = 1$ for all i
2. $\langle \underline{u}_i, \underline{u}_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$.

property 1. is actually encoded in property 2. if you recall what the definition of $\|\cdot\|$ is.

Here we go:

Step 1. Put $\underline{u}_1 = \dfrac{\underline{x}_1}{\|\underline{x}_1\|}$

Step 2. Put $\underline{p}_1 = \mathrm{proj}_{\underline{u}_1} \underline{x}_2 = \langle \underline{x}_2, \underline{u}_1 \rangle \underline{u}_1$

Then $\mathrm{orth}_{\underline{u}_1} \underline{x}_2 = (\underline{x}_2 - \underline{p}_1)$ is orthogonal to $\underline{x}_1$ and $\underline{u}_1$.

Further,

$\underline{x}_2 - \underline{p}_1 \neq \underline{0}$ since

$$\underline{x}_2 - \underline{p}_1 = \frac{-\langle \underline{x}_2, \underline{u}_1 \rangle}{\|\underline{x}_1\|} \underline{x}_1 + \underline{x}_2 . \qquad (*)$$

and $\underline{x}_2 \neq 0$ since it is a basis vector, and $\underline{x}_1, \underline{x}_2$ are linearly independent.

Step 3. Put $\underline{u}_2 = \dfrac{(\underline{x}_2 - \underline{p}_1)}{\|\underline{x}_2 - \underline{p}_1\|}$

Then $\underline{u}_2$ is a unit vector orthogonal to $\underline{u}_1$. Notice that span $\{\underline{x}_1, \underline{x}_2\}$ = span $\{\underline{u}_1, \underline{u}_2\}$ by eq'n $(*)$. Thus we have produced an orthonormal basis of span $\{\underline{x}_1, \underline{x}_2\}$, constructed from $\{\underline{x}_1, \underline{x}_2\}$.

Step 4. Repeat this process to produce the rest of the basis.

Define $\underline{u}_2, \dots, \underline{u}_n$ recursively by

$$\underline{u}_{k+1} = \frac{(\underline{x}_{k+1} - \underline{p}_k)}{\|\underline{x}_{k+1} - \underline{p}_k\|} \quad \text{where}$$

$$\underline{p}_k = \langle \underline{x}_{k+1}, \underline{u}_1 \rangle \underline{u}_1 + \langle \underline{x}_{k+1}, \underline{u}_2 \rangle \underline{u}_2 + \dots + \langle \underline{x}_{k+1}, \underline{u}_k \rangle \underline{u}_k$$

the projection of $\underline{x}_{k+1}$ onto the subspace span $\{\underline{u}_1, \dots, \underline{u}_k\}$ = span $\{\underline{x}_1, \dots, \underline{x}_k\}$.

Ex. Find an orthormal basis for P_3 w/ inner product defined by

$$\langle p, q\rangle = \sum_{i=1}^{3} p(x_i)\, q(x_i)$$

where $x_1 = -1$, $x_2 = 0$, $x_3 = 1$.

Start w/ the standard basis $\{1, x, x^2\}$ and use G-S to produce the orthonormal basis $\{\underline{u}_1, \underline{u}_2, u_3\}$

$$\underline{u}_1 = \frac{1}{\sqrt{3}}$$

$$\underline{u}_2 = \frac{1}{\sqrt{2}} x$$

$$\underline{u}_3 = \frac{\sqrt{6}}{2}\left(x^2 - \frac{2}{3}\right)$$

Ex. Let $A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{pmatrix}$ Find an orthonormal basis for the column space.

~~But~~ Let $r_{11} = \|a_1\| = 2$

put $\underline{q}_1 = \frac{a_1}{r_{11}} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)^T$

let $r_{12} = \langle \underline{q}_1, \underline{a}_2\rangle = \underline{q}_1^T \underline{a}_2 = 3$

$$\underline{p}_1 = r_{12}\, q_1 = 3 q_1$$

$$\underline{a}_2 - \underline{p}_1 = \left(\frac{-5}{2}, \frac{5}{2}, \frac{5}{2}, \frac{-5}{2}\right)$$

$$r_{22} = \|\underline{a}_2 - \underline{p}_1\| = 5$$

$$\underline{q}_2 = \frac{\underline{a}_2 - \underline{p}_1}{r_{22}} = \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)^T$$

Similarly $\quad r_{13} = \langle \underline{q}_1, \underline{a}_3 \rangle = 2$

$$r_{23} = \langle \underline{q}_2, \underline{a}_3 \rangle = -2$$

$$\underline{p}_2 = r_{13}\underline{q}_1 + r_{23}\underline{q}_2 = (2, 0, 0, 2)^T$$

$$\underline{a}_3 - \underline{p}_2 = (2, -2, 2, -2)^T$$

$$r_{33} = \|\underline{a}_3 - \underline{p}_2\| = 4$$

$$\underline{q}_3 = \frac{\underline{a}_3 - \underline{p}_2}{r_{33}} = \left(\tfrac{1}{2}, \tfrac{-1}{2}, \tfrac{1}{2}, \tfrac{-1}{2}\right)^T$$

so $Q = \begin{pmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \end{pmatrix}$ has the same column space as A, but is an orthogonal matrix.

If we collect all of the r_{ij}'s into a matrix, we get

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

<u>RE</u>. Verify that $A = QR$

This is called the Q-R factorization of A.

In general, we can always decompose an $m \times n$ matrix A w/ rank n into such a product where $Q \in \mathbb{R}^{m \times n}$ is orthogonal (its columns are an orthonormal basis of the column space) and $R \in \mathbb{R}^{n \times n}$ is nonsingular, and whose diagonal entries are all positive.

Think of R as an orientation-preserving change of basis matrix.

Least Squares Problem

Let $A \in \mathbb{R}^{m \times n}$ w/ $\text{rank}(A) = n$, and let $\underline{b} \in \mathbb{R}^m$. The *least squares problem* is to find a vector $\underline{\hat{x}} \in \text{Row}(A)$ such that $A\underline{\hat{x}}$ is "as close as possible" to $\underline{b}$. (whoops ☹)

To find $\underline{\hat{x}}$, one solves the *normal equation(s)*

$$A^T A \underline{x} = A^T \underline{b}$$

This has a unique sol'n since $A^T \in \mathbb{R}^{n \times m}$ w/ $\text{rank}(A^T) = \text{rank}(A) = n$; i.e., the column space (image) of A^T is all of $\mathbb{R}^n$.

When the columns of A form an orthonormal set of vectors in $\mathbb{R}^m$, then $A^T A = I \in \mathbb{R}^{n \times n}$. Then the solution of the least squares problem is

$$\underline{\hat{x}} = A^T \underline{b}.$$

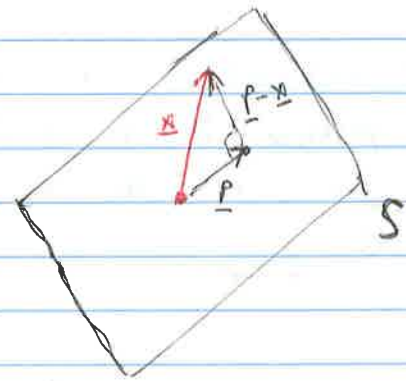
Thm. Let S be a subspace of an inner product space $(V, \langle , \rangle)$ and let $\underline{x} \in V$. Let $\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n\}$ be an orthonormal basis for S.

If $$\underline{p} = \sum_{i=1}^{n} c_i \underline{u}_i$$

where $c_i = \langle \underline{u}_i, \underline{x} \rangle \quad \forall i$,

then $\underline{p} - \underline{x} \in S^\perp$.

The picture looks as follows.



We can then decompose $\underline{x}$ into a sum of a vector in S and one in $S^{\perp}$.

Namely,

$$\underline{x} = \underline{p} + (\underline{p} - \underline{x}) = [\underline{x}]_S + [\underline{x}]_{S^{\perp}}$$

These are components of $\underline{x}$ that respect the subspace S.

We call $[\underline{x}]_S = \underline{p}$ the (orthogonal) projection of $\underline{x}$ onto S.

Thm. Let S be a nonzero subspace of $\mathbb{R}^m$ and let $\underline{b} \in \mathbb{R}^m$. If $\{\underline{u}_1, \ldots, \underline{u}_k\}$ is an orthonormal basis for S and $U = (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_k)$, then the projection $\underline{p}$ of $\underline{b}$ onto S is given by

$$\underline{p} = UU^T \underline{b}$$

Ex. Let S be the set of all vectors in $\mathbb{R}^3$ of the form $(x, y, 0)^T$. Find the vector $\underline{p} \in S$ closest to $\underline{w} = (5, 3, 4)^T$.

Sol'n. $\underline{e}_1$ and $\underline{e}_2$ form an o-n basis for S.

$$\text{Then } U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad U^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$UU^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{and} \quad \underline{p} = UU^T \underline{w} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix}$$

This is exactly what one would expect!

Ex. Find the best linear approximation of e^x on the interval $[0,1]$ by a linear function.

Let $S = P_2 = \{ p \in C[0,1] \mid p(x) = ax + b. \}$

Although $\{1, x\}$ span P_2, they are not orthogonal wrt the inner product

$$\langle p, q \rangle = \int_0^1 p(x)\, q(x)\, dx$$

By Gram-Schmidt, $\{1, \sqrt{12}(x - \tfrac{1}{2})\}$ form an orthonormal basis.

Let $c_1 = \langle \underline{u}_1, e^x \rangle = \int_0^1 e^x \, dx = e - 1$

$c_2 = \langle \underline{u}_2, e^x \rangle = \int_0^1 \sqrt{12}\left(x - \frac{1}{2}\right) e^x \, dx = \sqrt{3}\,(3 - e)$

Then $p(x) = c_1 \underline{u}_1(x) + c_2 \underline{u}_2(x)$

$= (e-1) \cdot 1 + \sqrt{3}(3-e)\left[\sqrt{12}\left(x - \frac{1}{2}\right)\right]$

$= (4e - 10) + 6(3-e)\,x \quad \approx e^x$

This is the best approximation of e^x on this interval!

linear

Pretty cool, right?!

<u>Ex.</u> Show that the functions $\sin x, \sin 2x, \ldots, \sin nx$ are orthonormal wrt the inner product

$$\langle f, g \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\, g(x)\, dx$$

in $C[-\pi, \pi]$.

Let $f \in C[-\pi, \pi]$ be 2π-periodic. Then we can approximate ~~[illegible]~~ f by a <u>Fourier series</u> or <u>trigonometric polynomial</u>.

$$f(x) \approx \sum_{k=1}^{n} a_k(x) \sin(kx)$$

where $a_k(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\, dx = \langle f, \sin(kx) \rangle$.

This can also be done w/ cosine. See pp. 252-3.