Section 7: Orthogonal, Nonexpansive, & Self-Adjoint Projections

Some authors require that a projection should be nonexpansive, orthogonal, self adjoint ... even linear. These notes have made a point that a function with $P^2 = P$ need not have these properties. At least geometrically, the minimum requirement that $P^2 = P$ defines a "projection." So, one must ask, what does one gain with these other conditions? Theorem 16 in this section provides an answer.

Lemma 13 (Polarization Identity) If T is a linear transformation then

$$<Tx, y> = <T(\frac{x+y}{2}), \frac{x+y}{2}> - <T(\frac{x-y}{2}), \frac{x-y}{2}> +$$
$$i <T(\frac{x+iy}{2}), \frac{x+iy}{2}> - i <T(\frac{x-iy}{2}), \frac{x-iy}{2}>.$$

<u>Suggestion of Proof:</u> Just do it!

Lemma 14 If E is a Hilbert space over the complex field and T is a linear function then these are equivalent:
(a) $<Tx, y> = <x, Ty>$ for each x and y in E (or, T is self adjoint), and
(b) $<Tx, x>$ is real for each x in E.

<u>Suggestion of Proof:</u> Use the Polarization Identity.

Definition An *Orthogonal Projection* is a projection P for which the null space of P is perpendicular to the range of P.

Remark: Recall that if C of Theorem 11 is a linear *subspace* of E, then statement (a) of Theorem 11 can be replaced by
(a') z is in the subspace C and $< x-z, c-z> = 0$ for all c in C.
or
(a'') z is in the subspace C and $<x-z, m> = 0$ for all m in C.

Definition If M is a linear space then M   = {n: <m,n> = 0 for all m in M}, R(P) is the range of P, and and N(P) is the null space of P.

Lemma 15 If P is a linear, closest point projection onto a closed subspace M, then (R(P))   = N(P) -- that is, P is an orthogonal projection.

<u>Suggestion of Proof:</u> Suppose that n   N[P]. We want to show that <n,r> = 0 for all r in R(P). Since n   N[P] then P(n) = 0. By Corollary 12, since P is a closest point projection, we have that for all r in R(P),

$$0 = <n-Pn,r-Pn> = <n-0, r-0> = <n,r>$$

for all r in R(P). Hence, n   R(P) .
Now, suppose that n   R(P) . We want to show that Pn=0.

Since n R(P) , then <n,r> = 0 for all r R, or <n-0,r-0> = 0 for all r R. But, this is a characterization of closest point.

EXAMPLE : Here is an example of a projection that is not a closest point: P({x,y}) = {x-y,0}. What is the range and null space of this projection?

Theorem 16 Suppose that P is a linear projection. These are equivalent:
(a) P is an orthogonal projection.
(b) if x is in E then <(1-P)x, Px> = 0.
(c) if x and y are in E, then <Px, y> = <x, Py>, i.e., P is self adjoint.
(d) if x and y are in E then $|x-Px|^2 + |Px - Py|^2 = |x-Py|^2$, i.e., P is a right triangle projection.
(e) if x and y are in E then $|x-Px| \le |x-Py|$, i.e., P is a closest point projection.
(f) <x-Px, z-Px > = 0 for all x and for all z in P(E), i.e., P is a right angle projection.
(g) $|Px| \le |x|$ for each x in E, i.e., P is a non-expansive.

<u>Suggestion of Proofs:</u>
a b Px is in the range of P and (1−P) is in the nullspace of P.
b c Show <Px,x> is real: <Px,x> = <Px,x-Px+Px> = $|Px|^2$.
c a <n,r> = <n,Pr> = <Pn,r>.
a d Add to the left side <x-Px,Px-Py> + <Px-Py,x-Px> = 0+0.
d e Delete $|Px - Py|^2$ from the left side of d.
e f Use Corollary 12.
f b <x-Px, Px> = <x-Px, 0-Px> = 0 since 0 is in P(E)
c g $|Px|^2$ = <Px,Px> $\le |Px|\ |x|$.
g b Let y = Px + λ(x-Px). Note that Py = Px and $|y-Py| = |\lambda|\ |x-Px|$. By (g) $0 \le |y|^2 - |Py|^2 = 2 \,\mathrm{Re}\, \lambda^*$ <Px,x-Px> $+ |\lambda|^2\ |x-Px|^2$. Investigate this parabola with λ real or equal to -it, t real.

Remarks
(1) Some texts take all projections to be linear. It is useful not to do this so as to get orthogonal projections onto closed subsets.

(2) Some texts take a "resolution" of the identity to be a sum of orthogonal projections. We did not do that because this did not happen in our resolution of matrices.

(3) One might think that if $A = \sum_{p=1}^{\infty} \lambda_p P_p$ then these are equivalent

(a) A is self adjoint, and
(b) each λ_p is real.

This did not happen in our examples. Not to worry! There is an inner-product with respect to which this is a theorem.

Assignment

(7.1) Give examples which contrast projections, linear projections and orthogonal, linear projections.

(7.2) Let C be the closed and bounded set {x: $|x| \le 1$ in R^2}. Let x = {1,1} and $P_c(1,1)$ be the closest point projection onto C. Find y_1 in C such that

$$<x-P_c(x), y_1 - P_c(x) > < 0.$$

Find y_2 in C such that

$$<x-P_c(x), y_2 - P_c(x) > = 0.$$

(7.3) Let S = $L^2[-1,1]$. Let *E* = {f: f   S, f(-x) = f(x)} and *O* = {f: f   S, f(-x) = - f(x)}. Show that *E* and *O* are orthogonal, linear subspaces. Find formulas for $P_E(f)$ and $[1-P_E](f)$. Show that P_E is a projection, is linear, and is orthogonal.

MAPLE remark: In this MAPLE exercise, we construct a non-orthogonal projection of R^3. To do this, choose u, v, and w linearly independent. We project onto the subspace spanned by multiples of u. Instead of being an orthogonal projection, however, it is a projection in the "direction of v and w." This will mean that

$$P(u) = u,\ P(v) = 0, \text{ and } P(w) = 0.$$

Such a projection can be accomplished as follows:
Choose x in R^3. Write

$$x = au + bv + cw.$$

Define P(x) = a u.

In a similar manner, we get the projection onto the subspace spanned by v and in the direction of u and w. And, we get the projection onto the supspace spanned by w in the direction of u and v.

Here's the technique: choose, for example, u = {1,0,1}, v = {1,1,0} and w = {0, 0, 1}. We seek the projection of R^3 onto the space spanned by u along v and w, the projection of R^3 onto the space spanned by v along u and w, and the projection of R^3 onto spanned by w along u and v.

Take {x, y, z} to be a point in R^3. we find the first matrix, P_u, which is the projection onto u. Thus we seek a, b, and c so that a u + b v + c w = {x,y,z} and we wish to write a, b, and c as a function of x, y, and z.

Make the matrix M such that the columns are u, v, and w.

```
> with(linalg):
> u:=vector([1,0,1]); v:=vector([1,1,0]); w:=vector([0,0,1]);
> M:=transpose(array([[u[1],u[2],u[3]],
                      [v[1],v[2],v[3]],[w[1],w[2],w[3]]]));
```

We know that {a, b, c} = M^{-1}\{x, y, z\} and that the projection onto u should be

$$< M^{-1}\{x,y,z\}, e_1>.$$

```
> preP:=evalm(inverse(M)&*vector([x,y,z]));
> coef1:=dotprod(preP,vector([1,0,0]));
> proj1:=evalm(coef1*u);
> col1:=vector([subs({x=1,y=0,z=0},proj1[1]),
                        subs({x=1,y=0,z=0},proj1[2]),
                        subs({x=1,y=0,z=0},proj1[3])]);
> col2:=vector([subs({x=0,y=1,z=0},proj1[1]),
                        subs({x=0,y=1,z=0},proj1[2]),
                        subs({x=0,y=1,z=0},proj1[3])]);
> col3:=vector([subs({x=0,y=0,z=1},proj1[1]),
                        subs({x=0,y=0,z=1},proj1[2]),
                        subs({x=0,y=0,z=1},proj1[3])]);
> Pu:=evalm(transpose(matrix([col1, col2, col3])));
```

Here is a different projection.

```
> coef2:=dotprod(preP,vector([0,1,0]));
> proj2:=evalm(coef2*v);
> col1:=vector([subs({x=1,y=0,z=0},proj2[1]),
                        subs({x=1,y=0,z=0},proj2[2]),
                        subs({x=1,y=0,z=0},proj2[3])]);
> col2:=vector([subs({x=0,y=1,z=0},proj2[1]),
                        subs({x=0,y=1,z=0},proj2[2]),
                        subs({x=0,y=1,z=0},proj2[3])]);
> col3:=vector([subs({x=0,y=0,z=1},proj2[1]),
                        subs({x=0,y=0,z=1},proj2[2]),
                        subs({x=0,y=0,z=1},proj2[3])]);
> Pv:=transpose(matrix([col1,col2,col3]));
```

And, here is the last projection.

```
> coef3:=dotprod(preP,vector([0,0,1]));
> proj3:=evalm(coef3*w);
> col1:=vector([subs({x=1,y=0,z=0},proj3[1]),
                        subs({x=1,y=0,z=0},proj3[2]),
                        subs({x=1,y=0,z=0},proj3[3])]);
> col2:=vector([subs({x=0,y=1,z=0},proj3[1]),
                        subs({x=0,y=1,z=0},proj3[2]),
                        subs({x=0,y=1,z=0},proj3[3])]);
> col3:=vector([subs({x=0,y=0,z=1},proj3[1]),
                        subs({x=0,y=0,z=1},proj3[2]),
                        subs({x=0,y=0,z=1},proj3[3])]);
> Pv:=transpose(matrix([col1,col2,col3]));
```

As, a check, it should be true that the sum of the projections is the identity matrix.

Section 8: Orthonormal Vectors

Surely, any view of a Hilbert space will have the notion of orthogonality and orthogonal vectors at its core. The very notion of orthogonality is embedded in the concept of the dot product. The classical applications of orthogonal functions pervade applied mathematics. Indeed, there are books on orthogonal functions and their application to applied mathematics.

We have seen that orthogonal vectors arise naturally as eigenvectors of self adjoint transformations. This section provides a process for generating an orthogonal set of vectors. It also examines the implication of having a *maximal* family of orthogonal vectors.

Remark One should review the notion of linearly independent vectors and verify that any orthogonal collection is linearly independent.

Definition The collection $\{x_p\}_{p=1}^{\infty}$ is a *maximal* orthonormal family if the only vector y satisfying $< x_p, y > = 0$, for all p, is $y = 0$.

Theorem 17 Let $\{x_p\}_{p=1}^{\infty}$ be an orthonormal set in $\{E, < \;, >\}$.

$$\sum |<y, x_p>|^2 \leq |y|^2,$$

with equality holding only in case $y = \sum <y, x_p> x_p$.

<u>Suggestion of Proof:</u> $0 \leq |y - \sum <y, x_p> x_p|^2 = |y|^2 - \sum |<y, x_p>|^2$.

Theorem 18 Suppose that $\{x_p\}_{p=1}^{\infty}$ is an orthonormal sequence in $\{E,< \;, >\}$. These are equivalent.

(a) $\{x_p\}_{p=1}^{\infty}$ is maximal - in the sense that if y is in E and $<y,x_p> = 0$ for all p then $y = 0$.

(b) $\{x_p\}_{p=1}^{\infty}$ is an orthonormal basis - in the sense that if y is in E then

$$y = \sum <y,x_p> x_p .$$

(c) Parseval's equality holds - if u and v are in E then

$$< u,v> = \sum_p <u, x_p> <x_p, v>.$$

(d) Bessel's equality holds - if y is in E then

$$|y|^2 = \sum |<y, x_p>|^2.$$

(e) The span S of $\{x_p\}_{p=1}^{\infty}$ is dense in E - in the sense that if y is in E then there is a sequence $\{u_p\}_{p=1}^{\infty}$ in S such that $\lim_p u_p = y$.

<u>Suggestion of Proof.</u>

a b Suppose that $\{x_p\}$ is maximal. Let M be all y that can be written as

$\sum <y,x_p> x_p$. Suppose M is not E. Let z be in M-E. Consider

$$z_0 = z - \sum <z,x_p> x_p.$$

Then $<z_0,x_p> = 0$ for all p.

b c $<u,v> = \Sigma_p\Sigma_q <u,x_p> <x_q,v> <x_p,x_q>$.

c d $<y,y> = \Sigma_p <y,x_p> <x_p,y>$.

d a If $<y,x_p> = 0$ for all p then $<y,y> = \Sigma_p <y,x_p> <x_p,y> = 0$.

b e If $u_n = \sum_{p=1}^{n} <y,x_p>x_p$ then $\lim_n u_n = y$.

e a Suppose that $y \neq 0$ and $<y,x_p> = 0$ for all p. Let $\{a_p\}$ be any number sequence.

$$|y - \sum a_p x_p|^2 = |y|^2 + \sum |a_p|^2 > |y|^2.$$

Hence there is no such $\{u_p\}$ for this y.

Definition If $\{x_p\}_{p=1}^{\infty}$ is a linearly independent sequence of vectors then the *Gramm-Schmidt process* generates an orthonormal sequence as follows:

$$u_1 = x_1, \qquad v_1 = x_1/|x_1|$$

$$u_n = x_n - \sum_{p=1}^{n-1} <x_n, v_p> v_p \qquad v_n = u_n/|u_n|.$$

Remark This process generates an orthogonal sequence: if n > k,

$$<u_n, v_k> = <x_n - \sum_{p=1}^{n-1} <x_n, v_p>v_p, v_k> = 0.$$

If $1 \le J \le n$ then the span of $\{x_p\}_{p=1}^{J}$ is the same as the span of $\{v_p\}_{p=1}^{J}$.

Theorem 19: Suppose that $\{P_n\}_{n=1}^{\infty}$ is a sequence of orthonormal polynomials, each having degree n and that that the generating dot product has the property that

$$<x\, f(x), g(x)> = <f(x), x\, g(x)>.$$

It follows that for each n, there are numbers α_n, β_n, and γ_n such that

$$P_{n+1}(x) = (\alpha_n x + \beta_n)\, P_n(x) + \gamma_n\, P_{n-1}(x).$$

Warning: In this discusion, P_n is a polynomial, not a projection.

<u>Suggestion for proof.</u> Here's why there is the recursion formula for orthonormal polynomials: We suppose that P_n is a polynomial of degree n, that $<P_n, P_m> = 0$ if $n \ne m$ and $= 1$ if $n = m$, and that the dot product has the property that $<x\, f(x), g(x)> = <f(x), x\, g(x)>$.

Choose β_n such that the function $P_{n+1}(x) - \beta_n x\, P_n(x)$ is a polynomial of degree n. Then there is a sequence $\{\alpha_p\}$ of numbers such that

$$P_{n+1}(x) - \beta_n x\, P_n(x) = \sum_{p=0}^{n} \alpha_p\, P_p(x).$$

Suppose that $k \le n-2$.

$$0 = <P_{n+1}, P_k> = <\beta_n x\, P_n(x), P_k(x)> + < \sum_{p=0}^{n} \alpha_p\, P_p(x), P_k(x)>$$

$$= \beta_n <P_n(x), xP_k(x)> + \alpha_k$$

$$= \beta_n <P_n(x), \sum_{i=0}^{k+1} \gamma_i\, P_i(x)> + \alpha_k$$

$$= \beta_n \sum_{i=0}^{k+1} \gamma_i <P_n(x), P_i(x)> + \alpha_k = 0 + \alpha_k.$$

Hence, $P_{n+1} = \beta_n x\, P_n + \alpha_n\, P_n(x) + \alpha_{n-1}\, P_{n-1}(x)$

Examples From the CRC:
Legendre Polynomials: $(n+1)\, P_{n+1}(x) = (2n+1)\, x\, P_n(x) - n\, P_{n-1}(x)$.

Here, $$<f, g> = \int_{-1}^{1} f(x)\, g(x)\, dx.$$

Tschebysheff Polynomials: $T_{n+1}(x) = 2\, x\, T_n(x) - T_{n-1}(x)$.

Here, $$<f, g> = \int_{-1}^{1} \sqrt{1-x^2}\ f(x)\, g(x)\, dx.$$

Laguerre Polynomials: $(n+1)\, L_{n+1}(x) = [2n+1-x]\, L_n(x) - n\, L_{n-1}(x)$.

Here, $$<f, g> = \int_{0}^{\infty} e^{-x} f(x)\, g(x)\, dx.$$

Assignment

(8.1) Let $\{x_p\}_{p=1}^{\infty}$ be an orthonormal sequence and $\{\alpha_p\}_{p=1}^{\infty}$ be a number sequence. These are equivalent:

(a) $\sum_{p=1}^{\infty} |\alpha_p|^2 < \infty$, and

(b) $\sum_{p=1}^{\infty} \alpha_p x_p$ converges in $\{E, <\ ,\ >\}$.

(8.2) Let $\{x_p\}_{p=1}^{N}$ be an orthonormal sequence and M be the span of $\{x_p\}_{p=1}^{N}$. Give a formula for the closest point projection P_M onto M.

(8.3) If y is in E and $\{x_p\}_{p=1}^{\infty}$ is an orthonormal sequence then the series

$$\sum_{p=1}^{\infty} <y, x_p>\, x_p \text{ converges.}$$

(8.4) Gramm-Schmidt $\{e_1, e_2, e_1+e_2+e_3\}$ in R^3.

(8.5) Gramm-Schmidt $\{1,x,x^2\}$ in $L^2([-1,1])$.

(8.6) Gramm-Schmidt $\{e^{-x}, xe^{-x}, x^2e^{-x}\}$ in $L^2([0,\infty),e^x)$.

MAPLE Remark. The procedure to perform the Gram-Schmidt process on a sequence of vectors is a part of the MAPLE linear algebra package. That procedure uses the standard dot-product. When using this procedure, it should be noted that the vectors returned are not normalized.

```
> with(linalg):
> u:=vector([1,0,1]); v:=vector([1,1,0]); w:=vector([0,0,1]);
> GramSchmidt({u,v,w});
```

It is not so hard to write a Gram-Schmidt procedure for dot-products other than the standard one, and even in a function space.

MAPLE also contains the standard orthogonal functions. For example, here are the Legendre polynomials.

```
> with(orthopoly);
> P(0,x); P(1,x); P(2,x);P(3,x);
> int(P(2,x)*P(3,x),x=-1..1);
```