## MODULE 6

**Topics: Gram-Schmidt orthogonalization process**

We begin by observing that if the vectors $\{x_j\}_{j=1}^N$ are mutually orthogonal in an inner product space $V$ then they are necessarily linearly independent. For suppose that $\langle x_i, x_j\rangle = 0$ for $i \neq j$ and

$$\sum_{j=1}^N \alpha_j x_j = 0$$

then taking the inner product of this equation with $x_k$ shows that $\alpha_k\langle x_k, x_k\rangle = 0$ so that $\alpha_k = 0$. Hence only zero coefficients are possible. We shall now observe the simplification which arises when we compute the orthogonal projection of a vector $y$ in a subspace with an orthogonal basis. Hence assume that

$$M = \text{span}\{x_1, \ldots, x_N\} \subset V$$

and $\langle x_i, x_j\rangle = 0$ for $i \neq j$. Let $y$ be a given vector in $V$. According to Module 4 the orthogonal projection $Py$ in $M$ is given as

$$Py = \sum_{j=1}^N \alpha_j x_j$$

where

$$A\vec{\alpha} = b,$$

with

$$A_{ij} = \langle x_j, x_i\rangle$$

and

$$b_i = \langle y, x_i\rangle,$$

has to be solved to obtain the coefficients $\{\alpha_j\}$. But because the $\{x_j\}$ are all orthogonal the matrix $A$ is diagonal so that

$$\alpha_i = \frac{\langle y, x_i\rangle}{\langle x_i, x_i\rangle}$$

**Example.** Let $f \in L_2(-\pi, \pi)$ and let $M = \text{span}\{1, \cos t, \ldots, \cos Nt, \sin t, \ldots, \sin Nt\}$.

It is straightforward to verify that

$$\langle \cos mt, \sin nt \rangle = \int_{-\pi}^{\pi} \cos mt \sin nt \, dt = 0 \quad \text{for all } m, n$$

and that

$$\langle \cos mt, \cos nt \rangle = \langle \sin mt, \sin nt \rangle = 0 \qquad \text{for } m \neq n.$$

Hence all the elements spanning $M$ are mutually orthogonal and therefore linearly independent and thus a basis of $M$. The orthogonal projection of $f$ onto $M$ is then given by

$$Pf = \sum_{n=0}^{N} \alpha_n \cos nt + \sum_{n=1}^{N} \beta_n \sin nt$$

where the coefficients are found explicitly as

$$\alpha_n = \frac{\langle f, \cos nt \rangle}{\langle \cos nt, \cos nt \rangle} \qquad n = 0, 1, \ldots, N$$

and

$$\beta_n = \frac{\langle f, \sin nt \rangle}{\langle \sin nt, \sin nt \rangle} \qquad n = 1, \ldots, N.$$

These coefficients are known as the Fourier coefficients of $f$ in $L_2(-\pi, \pi)$ and $Pf$ is the $N$th partial sum of the Fourier series of $f$. This partial sum is the best approximation, in the mean square sense, over the interval $(-\pi, \pi)$ in terms of the given sine and cosine functions (i.e., in terms of a so-called trigonometric polynomial).

If in an application the linearly independent vectors $\{x_j\}$ spanning $M$ are not orthogonal then it may be advantageous to compute an equivalent basis $\{z_1, \ldots, z_N\}$ of $M$ of mutually orthogonal vectors and to express the projection as a linear combination of these new basis vectors. The process of finding the orthogonal basis $\{z_j\}$ equivalent to the basis $\{x_j\}$ is known as the Gram-Schmidt orthogonalization process and proceeds recursively as follows.

We set

$$z_1 = x_1.$$

Assume that for $j = 1, \ldots, k-1$ we have found orthogonal vectors $\{z_j\} \subset \text{span}\{x_j\}$. Then we set

$$z_k = x_k - \sum_{j=1}^{k-1} \alpha_j z_j$$

where the $\alpha_j$ are computed such that $\langle z_k, z_j \rangle = 0$ for $j = 1, \ldots, k-1$. Since $\langle z_i, z_j \rangle = 0$ for $i \neq j$ and $i, j < k$ this requires that

$$\alpha_j = \frac{\langle x_k, z_j \rangle}{\langle z_j, z_j \rangle}\,.$$

When $k = N$ we have generated $N$ mutually orthogonal vectors, each of which is obtained as a combination of the basis vectors $\{x_j\}$. Hence $\{z_j\}$ forms on orthogonal basis of $M$.

**Examples:**

1) Let us find an orthogonal basis of $M = \text{span}\{(1,2,1,2),(0,1,0,1),(1,0,0,-1)\} \subset E_4$. The notation $E_4$ implies that the inner product is the dot product. We set

$$z_1 = (1,2,1,2)$$

and compute

$$z_2 = (0,1,0,1) - \alpha_1(1,2,1,2)$$

where

$$\alpha_1 = \frac{(0,1,0,1) \bullet (1,2,1,2)}{(1,2,1,2) \bullet (1,2,1,2)} = \frac{4}{10}$$

so that

$$z_2 = (-2/5, 1/5, -2/5, 1/5).$$

Since the span of a set remains unchanged if the vectors are scaled we can simplify the notation in our long-hand calculation by setting

$$z_2 = (-2, 1, -2, 1).$$

Then

$$z_3 = (1,0,0,-1) - \alpha_1(1,2,1,2) - \alpha_2(-2,1,-2,1)$$

where

$$\alpha_1 = \frac{(1,0,0,-1) \bullet (1,2,1,2)}{10} = \frac{-1}{10}$$

$$\alpha_2 = \frac{(1,0,0,-1) \bullet (-2,1,-2,1)}{(-2,1,-2,1) \bullet (-2,1,-2,1)} = \frac{-3}{10}$$

so that

$$z_3 = (1/2, 1/2, -1/2, -1/2).$$

Hence as an orthogonal basis of $M$ we may take

$$\{(1,2,1,2), (-2,1,-2,1), (1,1,-1,-1)\}.$$

The orthogonal projection of an arbitrary vector $y \in E_4$ onto $M$ is then given by

$$Py = \sum_{i=1}^{3} \frac{\langle y, z_i \rangle}{\langle z_i, z_i \rangle} z_i.$$

On occasion the vectors $z_j$ are scaled so that they are unit vectors in the norm $\|z\| = \langle z, z \rangle^{1/2}$. The vectors $\{z_j\}$ are said to be orthonormal in this case.

2) Find an orthogonal basis of span$\{1, t, t^2\} \subset L_2(-1,1)$. Scale the functions such that they assume a value of 1 at $t = 1$. The notation indicates that the inner product is

$$\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)dt.$$

We now apply the Gram-Schmidt process to generate orthogonal functions $\{\phi_0, \phi_1, \phi_2\}$. We set

$$\phi_0(t) \equiv 1 \qquad \text{(which already satisfies } \phi(1) = 1\text{)}$$

and compute

$$\phi_1(t) = t - \alpha_0 \, 1$$

where

$$\alpha_0 = \frac{\langle t, 1 \rangle}{\langle 1, 1 \rangle} = 0.$$

Hence

$$\phi_1(t) = t \qquad \text{(which also is already properly scaled).}$$

Then

$$\phi_2(t) = t^2 - \alpha_0 \phi_0(t) - \alpha_1 \phi_1(t)$$

with

$$\alpha_0 = \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{3}$$

and

$$\alpha_1 = \frac{\langle t^2, t\rangle}{\langle t, t\rangle} = 0.$$

Hence a function orthogonal to $\phi_0(t)$ and $\phi_1(t)$ is $c(t^2 - {}^1\!/_3)$ for any constant $c$. To insure that $\phi_2(1) = 1$ we choose

$$\phi_2(t) = {}^1\!/_2\,(3t^2 - 1).$$

These three orthogonal polynomials are known as the first three Legendre polynomials which arise, for example, in the solution of the heat equation in spherical coordinates. Legendre polynomials of order up to $N$ can be found by applying the Gram-Schmidt process to the linearly independent functions $\{1, t, \ldots, t^N\}$ with the $L_2$ inner product.

We shall conclude our discussion of projections and best approximations with a problem which, strictly speaking, leads neither to a projection nor a best approximation, but which has much the same flavor as the material presented above.

Suppose we are in a vector space $V$ with inner product $\langle\ ,\ \rangle$ and its associated norm. Suppose further that we wish to find an element $u \in V$ which satisfies the following $N$ constraint equations

$$\langle u, x_i\rangle = b_i, \qquad i = 1, \ldots, N$$

where the vectors $\{x_1, .., x_N\}$ are assumed to be linearly independent in $V$. In general, the solution to this problem will not be unique. The solution does exist and is unique if we restrict it to lie in

$$M = \text{span}\{x_1, \ldots, x_n\}.$$

In this case $u$ has to have the form

$$u = \sum_{j=1}^{N} \alpha_j x_j.$$

Substitution into the $N$ equations shows that $\{\alpha_j\}$ is a solution of the system

$$A\vec{\alpha} = b$$

where as before

$$A_{ij} = \langle x_j, x_i\rangle.$$

Linear independence of the $\{x_j\}$ guarantees that $A$ is invertible and that therefore the $\{\alpha_j\}$ are uniquely defined.

**Theorem:** *Let $y$ be any vector in $V$ which satisfies the $N$ constraint equations and let $u$ be the specific solution which belongs to $M$ then*

$$\|u\| \leq \|y\|.$$

**Proof:** For any $y$ we can write $y = y - u + u$. Then

$$\|y\|^2 = \langle y - u + u, y - u + u\rangle = \langle y - u, y - u\rangle + 2\text{Re}\langle y - u, u\rangle + \langle u, u\rangle.$$

But

$$\langle y - u, u\rangle = \sum_{j=1}^{N} \alpha_j \langle y - u, x\rangle = 0 \text{ because } \langle y, x_j\rangle = b_j = \langle u, x_j\rangle.$$

Hence

$$\|y\|^2 = \|y - u\|^2 + \|u\|^2 > \|u\|^2 \qquad \text{for } y \neq u.$$

Problems of this type arise in the theory of optimal controls for linear state equations and quadratic cost functionals. We shall not pursue this subject here but instead consider the following simpler geometric problem.

**Problem:** Let $\vec{x} = (x, y, z)$ be a point in $\mathbb{E}_3$. Find the point in the intersection of the planes

$$x + 2y + 3z = 1$$

$$3x + 2y + z = 5$$

which is closest to the point $(0, 1, 0)$.

**Answer:** The geometry is clear. These are two planes which are not parallel to each other. Hence they intersect in a line. The problem then is to find the point on the line which is closest to $(0, 1, 0)$. Since the setting is $\mathbb{E}_3$, closest means closest in Euclidean distance.

We shall examine three different approaches to solving this problem.

i) If we define $u = \vec{x} - (0, 1, 0)$ then we want the minimum norm $u$ which satisfies

$$\langle u, (1, 2, 3)\rangle = \langle \vec{x}, (1, 2, 3)\rangle - \langle (0, 1, 0), (1, 2, 3)\rangle = -1$$

$$\langle u, (3, 2, 1)\rangle = \langle \vec{x}, (3, 2, 1)\rangle - \langle (0, 1, 0), (3, 2, 1)\rangle = 3.$$

According to the last theorem the minimum norm solution is the uniquely defined solution belonging to $M = \text{span}\{(1,2,3),(3,2,1)\}$. It is computed as

$$u = \alpha_1(1,2,3) + \alpha_2(3,2,1)$$

where $\{\alpha_1, \alpha_2\}$ are found from the linear system

$$\begin{pmatrix} \langle(1,2,3),(1,2,3)\rangle & \langle(3,2,1),(1,2,3)\rangle \\ \langle(1,2,3),(3,2,1)\rangle & \langle(3,2,1),(3,2,1)\rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 14 & 10 \\ 10 & 14 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Doing the arithmetic we find that

$$(x,y,z) = (7/6, 7/6, -5/6).$$

ii) We shall find the equation of the line of intersection. Subtracting the equation of one plane from the other we find that coordinates of points in the intersection must satisfy

$$x - z = 2.$$

We set $z = s$. Then $x = 2 + s$ and $y = -1/2 - 2s$. Hence the line is

$$(x,y,z) = (2, -1/2, 0) + s(1,-2,1)$$

i.e., the line through $(2, -1/2, 0)$ with direction $(1,-2,1)$. The square of the distance from $(0,1,0)$ to a point on the line is

$$g(s) = (2+s)^2 + (3/2 + 2s)^2 + s^2.$$

This function is minimized for $s = -5/6$ yielding

$$(x,y,z) = (7/6, 7/6, -5/6).$$

iii) The problem can be solved with Lagrange multipliers. We want to minimize

$$g(\vec{x}) = g(x,y,z) = \langle \vec{x} - (0,1,0), \vec{x} - (0,1,0)\rangle.$$

subject to the constraints

$$\langle \vec{x}, (1,2,3)\rangle - 1 = 0$$

$$\langle \vec{x}, (3,2,1) \rangle - 5 = 0.$$

The Lagrangian is

$$\mathcal{L} = x^2 + (y-1)^2 + z^2 + \lambda_1(x + 2y + 3z - 1) + \lambda_2(3x + 2y + z - 5).$$

The minimizer has to satisfy

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial \mathcal{L}}{\partial z} = 0$$

as well as the constraint equations. This leads to the linear system

$$\begin{pmatrix} 2 & 0 & 0 & 1 & 3 \\ 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 2 & 3 & 1 \\ 1 & 2 & 3 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \\ 5 \end{pmatrix}$$

which again has the solution

$$(x, y, z) = ({}^7\!/_6, {}^7\!/_6, -{}^5\!/_6).$$

**Module 6 - Homework**

1) Consider the vectors $\{(1,1,0),(1,0,1),(0,1,1),(1,1,1)\}$.

   i) Why can these vectors not be linearly independent?

   ii) Carry out the Gram-Schmidt process. How does the dependence of the vectors $\{x_j\}$ affect the calculation of the $\{z_j\}$?

2) Let $y = (1,2,3) \in \mathbb{E}_3$. Find three orthonormal vectors $\{u_1, u_2, u_3\}$ such that $u_1$ is parallel to $y$.

3) Let $V$ be the set of all continuous real valued functions which are square integrable over $(0,\infty)$ with respect to the weight function

$$w(t) = e^{-t},$$

   i.e.,

$$V = \left\{ f : f \in C^0(0,\infty) \text{ and } \int_0^\infty f(t)^2 e^{-t} dt < \infty \right\}$$

   i) Show that $V$ is a vector space over $\mathbb{R}$.

   ii) Show that

$$\langle f, g \rangle = \int_0^\infty f(t) g(t) e^{-t} dt$$

   defines an inner product on $V$.

   iii) Show that $M = \text{span}\{1, t, t^2\} \subset V$.

   iv) Find an orthogonal basis of $M$. Scale the vectors such that they assume the value 1 at $t = 0$ (if you solve this problem correctly you will find the first three so-called Laguerre polynomials).