Getting Control with Linear Algebra

James V. Herod
Professor Emeritus
School of Mathematics
Georgia Tech

In the Fall, 1998, MAA Southeastern Section Newsletter, Sylvia T. Bozeman pointed to "a possible evolution in the nature and role of linear algebra over the next two or three decades." In her Governor's report, she predicted that changes in linear algebra "would more likely be driven by the demands of applications and new designs in computer hardware than by faculty."

I am convinced that her prediction will prove to be accurate. It seems appropriate that we should try to anticipate some of those changes and prepare applications that lead toward these changes in linear algebra. Some of these, such as the one described here, will be new to the undergraduate curriculum, but familiar to engineers that seek methods of control for their machines. The ideas become available in the undergraduate curriculum because of the power of the computer hardware and software now available.

I have often had the impression that students thought that linear algebra was an interlude between the calculus and sophomore differential equations. Too often, linear algebra may seem like a study of matrix manipulations and abstract n-dimensional spaces.

Indeed, in these days when computer algebra systems, such as Maple, will perform more operations on a matrix than most students care to know about and will compute dot-products, cross products, and tensor products from a one-line command, what is there to consider?

It becomes a familiar theme for persons using the computer algebra systems to add emphasis to the applications for the linear algebra. One such application is A Problem of Control.

A Problem in Control. Suppose it is known where an object is to start and in which direction it is to be going. Suppose also that all the forces acting on this object as it moves through a field are known. Let a starting position be specified. The problem of control is to determine the forcing function which "drives" the system appropriately so that the object will arrive at a specified position and at a specified time.

One can imagine that for some situations, appropriate control is not possible. Think of driving a car which has a large minimum turning radius. In other situations, there might be several paths that could be followed through the field to achieve the specified position in the specified time. One would want a path that minimizes "energy used."

We will put a problem of this type into the context of linear algebra. In general, problems of control are hard and considerable investigative work is being done now in constructing controls. It is the goal of this discussion to present an application of linear algebra that may be different

from the ordinary and to introduce the concepts of control into sophomore differential equations.

Linear Algebra Background Information. We take E to be a vector space on which there is a dot product. We assume the structure of E to be rich enough so that E is complete -- that E is a Hilbert Space. We will illustrate the linear algebra ideas in R^3 and R^4. The purpose for stating the results in some generality is to see that the ideas are not dependent on the geometry of finite dimensional spaces. As traditional, we denote the dot product of x and y by < x, y>. Here are ideas that are used:

(1) A *linear functional* is a linear function from the space E to the scalars. For example, if E = R^3, then

$$L([x,y,z]) = x + 2y + 3z \qquad (1)$$

is a linear functional. It has domain R^3, it is linear, and its values are scalars -- are numbers.

(2) A *projection* is a function from E into E for which it is true that

$$P^2 = P.$$

For example, take P to be given by the matrix product

$$P([x,y,z]) = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \qquad (2)$$

This matrix defines a linear projection. To see this, simply compute P^2.

Ideas (1) and (2) were recollections of names of special kinds of linear transformations. Ideas (3) and (4) are results for such functions.

(3) If L is a linear functional on E, then there is exactly one member e of E such that, for all v in E,

$$L(v) = < v, e >.$$

To illustrate this result, note that the linear functional from equation (1) can be expressed as

$$L([x,y,z]) = < [x,y,z], [1,2,3] >.$$

(4) If P is a linear projection onto a subspace M, then these are equivalent:

(a) P is a *closest point projection* -- in the sense that if x is in E then

$$| x - Px | \le | x - m |$$

for all m in M, and

(b) P is an *orthogonal projection* onto M-- in the sense that if x is in E then

$$< x - Px, m > = 0$$

for all m in M.

Remark: Two other statements equivalent to (4.a) and (4.b) above are

(c) P is *non expansive* -- in the sense that $|Px| \leq |x|$ for all x in E, and

(d) P is *symmetric* (or self-adjoint) -- in the sense that the conjugate transpose of P is P.

We have given an example of a projection in equation (2). In fact, that projection was a closest point, or orthogonal, or non-expansive projection. To see that, just observe that the matrix is symmetric about the main diagonal. This observation is convincing if you believe the previous statements 4(a) – 4(d) are equivalent, and that symmetric linear transformation on R^n have matrix representations which are symmetric about the main diagonal.

It seems well to have an illustration of a projection that in not a closest point projection. The following is a projection and, in view of the equivalent statement (4.d), it is not an orthogonal projection:

$$P = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

Visualize this last projection; see that the range consist of points of the form [x, 0], that P([x, y]) is not the closest point in the range to [x, y], and that some points are closer to the origin than their images. For example,

$$P([1/2, -1/2]) = [1, 0].$$

Also, the projection does not project orthogonally onto its range.

It is of value to consider non-linear projections in analyzing the problem of control. For example, suppose that M is an affine set - - in the sense that there is a vector e in E such that

$$e - M$$

is a linear subspace. A projection onto M will be a function from E to M for which $P^2 = P$. Note that P is not linear if M is affine, but not a subspace. If M is any closed, convex subset of E, there is a unique closest point projection function, albeit not necessarily linear. Our interest will be these closest point projections.

We now generalize the notion of orthogonal projection in statements (4) above. The equivalent statements take a new form. Here are the first two.

5. If P is a projection onto an affine set M, then these are equivalent:

(a) P_m is a closest point projection -- in the sense that

$$|u - P_m(u)| \leq |u - m| \text{ for all m in M, and}$$

(b) P_m is an orthogonal projection -- in the sense that

$$\langle u - P_m(u), m - P_m(u) \rangle = 0 \text{ for all m in M.}$$

A moment's thought leads to an understanding that this last result contains the previous equivalence of (4.a) and (4.b) in case M is a linear subspace.

It is not a surprise that a closest point projection should arise in a problem of control. Think of the reason why this is true: if the problem is a linear problem, then the family of all forcing functions that *drive* the object to a specific point in time T will be a *convex set*.

The control problem has two parts, find the convex set and find the driver in the convex set that is closest to zero with some measure of closeness. To do this latter, we simply take the closest point projection of zero onto the convex set.

Example 1: The set

$$M = \{ [u,v]: v = u + 1 \}$$

is an affine subspace for if e = [− 1/2, 1/2] , then

$$e - M = \{ [u,v]: v = u\}$$

is a subspace of R^2. The nonlinear, closest point projection onto M is

$$P(x,y) = [(x+y)/2, (x+y)/2] + e .$$

Illustration of the linear algebra tools. We provide examples and illustrations of the linear algebra tools that are to be used. The illustrations are different from the examples in that they have a general context. The examples will be specific cases. These examples and illustrations will seem to be a detour from the problem in control, but they use exactly the ideas from linear algebra that are needed for the problem in control.

<u>Example 2</u> Let M be the line in the intersection of the planes

P_1: x+y+z=0 and P_2: x-y+z=0.

Also, consider the point [1,2,3] which is not in either plane. The problem is to find the closest point in M to [1,2,3]. Here is a solution using the techniques of calculus:

P_1 is the plane consisting of points {x, y, z} such that

$$\langle \{x,y,z\}, \{1,1,1\} \rangle = 0.$$

P_2 is the plane consisting of points {x, y, z} such that

$$<\{x,y,z\}, \{1,-1,1\}> = 0.$$

The two vectors [1,1,1] and [1, -1, 1] are perpendicular to all points in P_1 and P_2, respectively. The line M has the same direction as a vector which is perpendicular to both these vectors. That direction can be computed as the cross product of [1, 1, 1] and [1, -1, 1]:

$$[1,1,1] \times [1,-1,1] = [\ 2, 0, -2]\ .$$

The line M also contains {0,0,0} since this point is in P_1 and P_2. Hence, an equation for this line is $M(t) = [\ 2\ t, 0, -\ 2\ t\ ]$. We want to choose t in order to minimize the distance between [2 t, 0, −2 t] and [1, 2, 3]:

$$|[2\ t,0,-2\ t] - [1,2,3]|^2$$

$$= (2\ t-1)^2 + 2^2 + (-2\ t-3)^2 = D(t).$$

Well, $\qquad D'(t) = 2\ (2\ t-1)\ 2 + 2\ (\ (-2\ t-3)\ (-2) = 16\ t + 8.$

And, $D'(t) = 0$ provided that $t = -1/2$. Hence, the closest point in M to [1,2,3] is [−1,0,1].

Remark: The method of this example works only in 3-D, because the cross product is unique to R^3. We make the first illustration to do the same problem as Example 2, but to be independent of dimension.

<u>Illustration 1.</u> Let L_1 and L_2 be linear functions from E to the scalars . Let M be the intersection of the null space of L_1 and of L_2. The problem is: given u in E, find the closest point in E to the null space of L_1 and L_2. (It is important to note that the previous example is a special case of this setting.) By statement (1), there is y_1 and y_2 in E such that

$$L_1(v) = <v, y_1> \text{ and } L_2(v) = <v, y_2>.$$

It follows that

$$M = \{x: <x, y_1> = 0 = <x, y_2> \}. \qquad (3)$$

Then

$$M = \{x: <x, \alpha y_1 + \beta y_2> = 0 \text{ for all } \alpha \text{ and } \beta\}.$$

Also, the collection of points perpendicular to M can be characterized:

$$M \quad = \text{span}\{y_1, y_2\}. \qquad (4)$$

Now, suppose that u is a point of E and that P_m is the orthogonal projection onto M. The point $P_m(u)$ provides the closest point in M to u. We use this fact. First, the point $P_m(u)$ is in M. From (3), this implies that

$$<P_m(u), y_1> = 0 = <P_m(u), y_2>. \qquad (5)$$

Using that P_m is a closest point projection, we have from 4(b) that u - $P_m(u)$ M . Thus, from (4),

$$u - P_m(u) = \alpha y_1 + \beta y_2 \text{ for some } \alpha \text{ and } \beta \qquad (6)$$

Take the inner product of equation (6) with y_1 and y_2 and use (5):

$$<u, y_1> = \alpha <y_1, y_1> + \beta <y_2, y_1> \qquad (7)$$

and

$$<u, y_2> = \alpha <y_1, y_2> + \beta <y_2, y_2>,$$

or

$$\begin{pmatrix} <u,y_1> & <u,y_2> \end{pmatrix} = \begin{pmatrix} <y_1,y_1> & <y_2,y_1> \\ <y_1,y_2> & <y_2,y_2> \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \qquad (8)$$

We can solve the system (7), or (8), uniquely provided the matrix in equation (8) has an inverse. From information about determinants, we want

$$|y_1|^2 \, |y_2|^2 - |<y_1,y_2>|^2 \neq 0.$$

This inequality holds if y_1 and y_2 are linearly independent.

These then are the ideas: $P_m(u)$ is in M gives equation (5), P_m is a closest point projection gives equation (6). Applying these two ideas from linear algebra solves the problem in control ... almost.

We provide another example where we techniques use the two ideas.

<u>Example 3.</u> Let L_1 be the linear function

$$L_1(w, x,y,z) = < [w, x,y,z], [1, 1, 1, 1] >$$

and L_2 be the linear function

$$L_2(w, x, y, z) = < [w, x, y, z], [1, -1, 1, -1\} >.$$

Let M be the intersection of the null space of L_1 and L_2. Take u to be given as

$$u = [0, 1, 2, 3].$$

We find the closest point in M to u. From equation (6), we have that

$$P_m(u) = u - \alpha\, y_1 - \beta\, y_2 \,.$$

From equations (7)

$$6 = 4\,\alpha + 0\,\beta$$

and $$-2 = 4\alpha + 0\,\beta.$$

Solve these equations for α and β to find that

$$P_m(u) = [-1, -1, 1, 1].$$

CHECK: As a check of this result, we use ideas available with finite dimensional problems. The null space of L_1 and L_2 is the same as the nullspace of the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}.$$

Using tools taught in linear algebra, this is computed to be

$$\{ v\!: v = s \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \text{ for all } s \text{ and } t \}.$$

To compute the distance from u to any of such points v in the null space of the two, we compute the norm of $u - v$. We find this is smallest at s = 1, t = 1, or at $P_m(u)$ as found above.

As already suggested, we need to make one further variation in the problem. The set M needs to be an affine subset, not a subspace. This occurs when the two linear equations are non-homogeneous equations. Consider the following.

<u>Illustration 2.</u> Let L_1 and L_2 be the linear functions from E to the scalars. Let

$$M = \{x\!: L_1(x) = A_1 \text{ and } L_2(x) = A_2\}. \qquad (9)$$

By Statement (1), L_1 and L_2 have a special form: there is y_1 and y_2 such that

$$L_1(v) = \langle v, y_1\rangle \text{ and } L_2(v) = \langle v, y_2\rangle.$$

Now, M is a closed, convex set. If u   E and if we seek the closest point in M to u then $P_m(u)$ again provides this point. While $P_m(u)$ is a projection, it is not necessarily linear. We want a formula for $P_m(u)$. To get this formula, see from Statement 5(b) that u - $P_m(u)$ is perpendicular to the intersection of the nullspace of L_1 and of L_2.

As in Illustration 1, equation (6), this implies that

$$u - P_m(u) = \alpha\, y_1 + \beta\, y_2. \qquad (10)$$

We also have from the equations (5)

$$A_1 = L_1(P_m(u)) = \langle P_m(u), y_1\rangle$$

and $$A_2 = L_2(P_m(u)) = \langle P_m(u), y_2\rangle.$$

Suppose we know u and seek $P_m(u)$. From (10), we see that

$$u - \alpha y_1 - \beta y_2 = P_m(u).$$

Hence,

$$\langle u, y_1\rangle - \alpha \langle y_1, y_1\rangle - \beta \langle y_2, y_1\rangle = L_1(P_m(u)) = A_1,$$

$$\langle u, y_2\rangle - \alpha \langle y_1, y_2\rangle - \beta \langle y_2, y_2\rangle = L_2(P_m(u)) = A_2. \qquad (11)$$

Or, $$\begin{pmatrix} \langle u,y_1\rangle - A_1 \\ \langle u,y_2\rangle - A_2 \end{pmatrix} = \begin{pmatrix} \langle y_1,y_1\rangle & \langle y_2,y_1\rangle \\ \langle y_1,y_2\rangle & \langle y_2,y_2\rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \qquad (12)$$

Example 4. Take M = { [x,y,z]: x + y + z = 1 and x − y + z = 2 }.
Take

$$u = [1, 2, 3].$$

The vectors y_1 and y_2 are given by

$$y_1 = [1, 1, 1] \quad \text{and} \quad y_2 = [1, -1, 1].$$

From (11) we must solve a system of equations:

$$5 = 2\alpha + \beta$$

and $$0 = \alpha + 3\beta.$$

This system has solution

$$\alpha = 15/8, \qquad \beta = -5/8.$$

Thus, $P_m([1, 2, 3]) = [-1/4. -1/2, 7/4]$. This has distance $5/4\sqrt{6}$ from u.

Example 5: Because the previous example was in R^3, it could have been worked using the cross product. If the equations were

$$w + x + y + z = 3 \text{ and } w - x + y - z = 5$$

and u were [1, 2, 3, 4], cross productions could not be used. In this case

$$P_m(u) = [1, -3/2, 3, 1/2].$$

Differential Equations Background Information. The control problem will be stated as a differential equation. There will be two ideas that are important. First, we will re-write a constant coefficient, second order differential equation as a two dimensional system. Second, we will change the differential equation into an integral equation. Here are the ideas:

(6) Suppose that each of a and b is a number and that f(t) is a continuous function. These are equivalent:

(a) The function y(t) satisfies the differential equation

$$y'' + a\, y' + b\, y = f(t),\ y(0) = \alpha,\ y'(0) = \beta, \text{ and}$$

(b) The functions Z and F are defined by

$$Z(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix},\quad F(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}$$

and Z satisfies the differential equation

$$Z' = M\, Z + F(t), \text{ with } Z(0) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

where $M = \begin{pmatrix} 0 & 1 \\ -\beta & -\alpha \end{pmatrix}$.

It is common to change, for example, constant coefficient differential equations to this matrix form, for then the ideas from linear algebra about the location of eigenvalues for the matrix come to play.

The second idea provides a solution for the differential equation in terms of an integral. In this idea, the notion of an exponential for a matrix is used.

(7) Suppose that Z is any continuously differentiable function and M is a matrix. These are equivalent

(a) The function Z satisfies the differential equation

$$Z' = M\, Z + F,\ Z(0) = C.$$

(b) The function Z(t) is defined by

$$Z(t) = \exp(t\, M)\, Z(0) + \int_0^t \exp((t-s)\, M)\, F(s)\, ds$$

A Problem of Control. We now have linear algebra tools and differential equations tools to discuss the control problem. Suppose that $Q_0 \in R^2$ and A is a 2x2 matrix. Suppose that $b \in R^2$ and $Q_1 \in R^2$. We seek v with minimum norm such that if

$$Z' = M\, Z + b\, v,\ Z(0) = Q_0, \qquad (13)$$

then also, $Z(1) = Q_1$. We know

$$Z(t) = \exp(t\, M)\, Q_0 + \int_0^t \exp((t-s)\, M)\, b\, v(s)\, ds.$$

Since $Q_1 = Z(1)$, We have

$$Q_1 = \exp(M)\, Q_0 + \int_0^1 \exp((1-s)\, M)\, b\, v(s)\, ds$$

or,
$$Q_1 - \exp(M)\, Q_0 = \int_0^1 \exp((1-s)\, M)\, b\, v(s)\, ds.$$

Let $\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = Q_1 - \exp(M)\, Q_0$. **Let** L_1 **and** L_2 **be defined so that**

$$\begin{pmatrix} L_1(v) \\ L_2(v) \end{pmatrix} = \int_0^1 \exp(\,(1-s)\, M)\, b\, v(s)\, ds$$

To identify y_1 **and** y_2**, take the first and second components, respectively, of**

$$\exp(\,(\,1-s)\, M)\, b.$$

That is, $y_1(s) = \langle \exp((1-s)\, M)\, b, [1,0] \rangle$

and $y_2(s) = \langle \exp((1-s)\, M)\, b, [0,1] \rangle$.

To ask for the point v with minimum norm which satisfies $A_1=L_1(v)$ **and** $A_2= L_2(v)$**, we choose u from Illustration 2 to be 0. It follows that v will be** $P_m(0)$ **where** P_m **is the nonlinear projection onto**

$$\{x: L_1(x) = A_1 \text{ and } L_2(x) = A_2\}.$$

From the equation (11) of Illustration (2)

$$v(s) = 0 - \alpha \langle \exp((1-s)\, M)\, b, [1,0]\rangle - \beta \langle \exp((1-s)\, M)\, b, [0,1]\rangle \qquad (14)$$
$$= -\alpha\, y_1(s) - \beta\, y_2(s)$$

where α **and** β **satisfy an equation (12) in Illustration 2.**

<u>Example 6.</u> We find v(s): $[0, \pi] \to R$ **such that**

$$y'' + y = v, \quad y(0) = 0, \quad y'(0) = 1, \qquad (15)$$

and
$$y(\pi)=1,\ y'(0) = 0. \qquad (16)$$

First we rewrite the problem in the form of (13). Take

$$M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

with
$$Z(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix},$$

$$Q_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad Q_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In this case,

$$\exp(t\,M) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix},$$

$$b = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and v(t) is to be found. The matrix $\exp(\pi - s)$ is given as

$$\exp(\pi - s) = \begin{pmatrix} -\cos(s) & \sin(s) \\ -\sin(s) & -\cos(s) \end{pmatrix}$$

The functions y_1 and y_2 are sin(s) and − cos(s) respectively. Carrying out the calculations from (10) and (11) we have

$$v(t) = 2\,(\sin(t) - \cos(t)\,)\,/\,\pi.$$

Using this function v(t) on the right side of equation (15) with the indicated initial conditions produces the solution

$$y(t) = \sin(t)\,\frac{1+\pi}{\pi} \;-\; t\;(\sin(t) + \cos(t)\,).$$

This function, y(t), satisfies the end conditions (16) at $t = \pi$.

Accessibility. The following example shows that it may be that not every point is accessible in a control model. Suppose that P_1 and P_2 are orthogonal projections, $M = \alpha\, P_1 + \beta\, P_2$, and that $P_1(b) = 0$. If x is in E and u is in $L^2[0,1]$ then

$$|P_1(x) - \int_0^1 \exp((1-s)\,M\,)b\,u(s)\,ds| \;\geq\; |P_1(x)|.$$

Here's why:

$$|P_1(x) - (\int_0^1 \exp((1-s)\,\alpha\,)\,P_1\,b\,u(s)\,ds + \int_0^1 \exp((1-s)\,\beta\,)\,P_2\,b\,u(s)\,ds)|^2 =$$

$$|\,P_1(x) - \int_0^1 \exp((1-s)\,\beta\,)\,u(s)\,ds\,P_2 b\,|^2$$

$$= |P_1(x)|^2 + \left| \int_0^1 \exp((1-s)\beta)\, u(s)\, ds \right|^2 \; |P_2(b)|^2$$

$$\geq |P_1(x)|^2.$$