Unit 7: The Dot (inner) Product

3.7.1

In our discussion of the game of mathematics at the start of this course, we indicated that one often tries to abstract the crucial aspects of a well-known system in order to see what inescapable consequences follow from the most "reasonable" set of axioms.

With this idea in mind, let us observe that as of this moment, we have not required our generalized vector space to have the equivalent of a "dot" product defined on it. That is, we have talked about bases of a vector space and about linear transformations, but we have not required that there be defined an operation which allows us to assign to each pair of vectors a number.

Thus, if we were interested in developing the anatomy of a vector space still further, it would next be reasonable to define that part of the structure which is characterized by what the dot product does for ordinary 2- and 3-dimensional vector spaces.

What mathematicians first noticed in this respect was the importance of linearity among the various attributes of the dot product. For this reason, they insisted that a rule that assigned to ordered pairs of vectors a real number not even be considered an "inner" (dot) product unless it possess the property of linearity. With this in mind, they defined a bilinear* function on a vector space to be any mapping

$f: V \times V^{**} \rightarrow R$

such that for $\alpha, \beta, \gamma \in V$ and $c \in R$

$$\left.\begin{array}{ll} 1. & f(\alpha + \beta, \gamma) = f(\alpha, \gamma) + f(\beta, \gamma) \\ 2. & f(\alpha, \beta + \gamma) = f(\alpha, \beta) + f(\alpha, \gamma) \\ 3. & f(\alpha, c\beta) = f(c\alpha, \beta) = cf(\alpha, \beta) \end{array}\right\} \quad (1)$$

*We say bilinear because the domain of f is ordered pairs.

**Recall that by A x B we mean $\{(a,b): a \in A, b \in B\}$. Hence S x S denotes the set of all ordered pairs of elements in S.

3.7.1 continued

Note #1

In keeping with the notation of the usual dot product, one often writes $\alpha\cdot\beta$ rather than $f(\alpha,\beta)$. With this in mind, (1) becomes rewritten as

$$\left.\begin{array}{l} 1.\quad (\alpha + \beta)\cdot\gamma = \alpha\cdot\gamma + \beta\cdot\gamma \\ 2.\quad \alpha\cdot(\beta + \gamma) = \alpha\cdot\gamma + \alpha\cdot\gamma \\ 3.\quad \alpha\cdot(c\beta) = (c\alpha)\cdot\beta = c(\alpha\cdot\beta) \end{array}\right\} \qquad (1')$$

Note #2

It is crucial to remember that $\alpha\cdot\beta$ is a number, not a member of V.

Note #3

Later we shall add the requirement that $\alpha\cdot\beta = \beta\cdot\alpha$, but for now this property is not imposed on our definition of a bilinear function. For this reason 1. and 2. are not redundant.

In other words, if we assume that $\alpha\cdot\beta = \beta\cdot\alpha$ for all $\alpha,\beta\in V$, then

$$\alpha\cdot(\beta + \gamma) = (\beta + \gamma)\cdot\alpha$$

so by (1),

$$\alpha\cdot(\beta + \gamma) = \beta\cdot\alpha + \gamma\cdot\alpha$$

or

$$\alpha\cdot(\beta + \gamma) = \alpha\cdot\beta + \alpha\cdot\gamma .$$

Since, however, we do not assume $\alpha\cdot\beta = \beta\cdot\alpha$, we cannot derive (2) from (1).

Note #4

There are, in general, many different ways (usually infinitely many ways) to define $f: V \times V \to R$ such that property (1) is obeyed. (We shall discuss this in more detail later). Thus, when we talk about $\alpha\cdot\beta$ we are assuming that we have chosen one particular bilinear function.

3.7.1 continued

With this as background, we now try to explore why the linear properties are so important. What we shall show is that many of the usual structural properties of a dot product follow just by linearity. In particular:

a. $\alpha \cdot \vec{0}^* = \alpha \cdot (\vec{0} + \vec{0})$
$= \alpha \cdot \vec{0} + \alpha \cdot \vec{0}.$ (2)

Since $\alpha \cdot \vec{0}$ is, in any event, a number (i.e., a member of R), let us denote it by b. We then conclude from (2) that

$b = b + b$ (3)

and since b is a number, we conclude by elementary arithmetic from (3) that b = 0. Thus, we have shown that

$\alpha \cdot \vec{0} = 0$** (4)

In summary, we have shown in part (a) that the usual property $\alpha \cdot \vec{0} = 0$ holds for <u>any</u> <u>bilinear</u> <u>function</u>.

b. Given,

$(a_1\alpha_1 + a_2\alpha_2) \cdot (a_3\alpha_3 + a_4\alpha_4)$, we may treat $a_3\alpha_3 + a_4\alpha_4$ as a single vector whereupon we may use 1. to conclude that

$(a_1\alpha_1 + a_2\alpha_2) \cdot (a_3\alpha_3 + a_4\alpha_4)$

$= (a_1\alpha_1 + a_2\alpha_2) \cdot \gamma$ [where $\gamma = a_3\alpha_3 + a_4\alpha_4$]

$= (a_1\alpha_1) \cdot \gamma + (a_2\alpha_2) \cdot \gamma$,

and this by 3. equals

$a_1(\alpha_1 \cdot \gamma) + a_2(\alpha_2 \cdot \gamma).$ (5)

*We have resorted to the notation $\vec{0}$ to reinforce the idea that $\alpha \cdot \beta$ in an operation on vectors not numbers.

**It is not an oversight that we have written 0 rather than $\vec{0}$ on the right side of (4). Namely, $\alpha \cdot \vec{0}$ is a number, not a vector.

3.7.1 continued

Replacing γ by $a_3\alpha_3 + a_4\alpha_4$, (5) becomes

$$a_1[\alpha_1\cdot(a_3\alpha_3 + a_4\alpha_4)] + a_2[\alpha_2\cdot(a_3\alpha_3 + a_4\alpha_4)]. \qquad (6)$$

By 2. we know that

$$\alpha_1\cdot(a_3\alpha_3 + a_4\alpha_4) = \alpha_1\cdot(a_3\alpha_3) + \alpha_1\cdot(a_4\alpha_4),$$

and this, in turn, by 3. is

$$a_3(\alpha_1 \cdot\alpha_3) + a_4(\alpha_1 \cdot\alpha_4). \qquad (7)$$

Similarly

$$\alpha_2 \cdot(a_3\alpha_3 + a_4\alpha_4) = a_3(\alpha_2 \cdot\alpha_3) + a_4(\alpha_2 \cdot\alpha_4). \qquad (8)$$

From (7) and (8) we see that (6) may be replaced by

$$a_1[a_3(\alpha_1\cdot\alpha_3) + a_4(\alpha_1 \cdot\alpha_4)] + a_2[a_3(\alpha_2 \cdot\alpha_3) + a_4(\alpha_2 \cdot\alpha_4)]. \qquad (9)$$

Since equation (9) involves only numbers*, we may use the rules of ordinary arithmetic to replace (9) by

$$a_1a_3(\alpha_1 \cdot\alpha_3) + a_1a_4(\alpha_1 \cdot\alpha_4) + a_2a_3(\alpha_2 \cdot\alpha_3) + a_2a_4(\alpha_2 \cdot\alpha_4). \qquad (10)$$

In summary, then, provided that we keep the order "straight" (i.e., $\alpha_1\cdot \alpha_4$ is not necessarily the same as $\alpha_4\cdot\alpha_1$), just as we had to do in our study of the cross-product, we see from (10) that the linearity of the dot product is what allows us to treat it as we treat the analogous expressions of numerical arithmetic.

*That is, while $\alpha_1,\alpha_2,\alpha_3$, and α_4 are vectors, $\alpha_1\cdot \alpha_2$, etc. are numbers. That is, why we have parenthesized $\alpha_1\cdot\alpha_2$, etc., namely to emphasize that $\alpha_1 \cdot \alpha_2$ is <u>a number</u>.

3.7.2

a. Once again we restrict our attention to $n = 2$ for computational simplicity, but our results hold for all n. More importantly, the technique we shall use here is the same one that is used in the higher dimensional cases.

Keep in mind that our main aim in this exercise is to show that a particular bilinear function is completely determined once we know its effect on each pair of basis elements.

So suppose that $V = [u_1, u_2]$ where $\{u_1, u_2\}$ is an arbitrarily chosen basis for V. Now let α and β be arbitrary elements of V. Say,

$$\left.\begin{aligned} \alpha &= x_1u_1 + x_2u_2 \\ \beta &= y_1u_1 + y_2u_2 \end{aligned}\right\} \qquad (1)$$

From (1) and part (b) of the previous exercise we conclude that

$$\alpha \cdot \beta = x_1y_1(u_1 \cdot u_1) + x_1y_2(u_1 \cdot u_2) + x_2y_1(u_2 \cdot u_1) + x_2y_2(u_2 \cdot u_2). \qquad (2)$$

Since x_1, x_2, y_1, y_2 are known once α and β are given, we see from (2) that $\alpha \cdot \beta$ is determined once the values of $u_1 \cdot u_1$, $u_1 \cdot u_2$, $u_2 \cdot u_1$, and $u_2 \cdot u_2$ are fixed.

Note #1

With a little insight, equation (2) gives us a hint as to why our matrix coding system is used in yet another context in vector spaces. Namely, if we let

$$A = \begin{bmatrix} u_1 \cdot u_1 & u_1 \cdot u_2 \\ u_2 \cdot u_1 & u_2 \cdot u_2 \end{bmatrix},$$

$\vec{X} = [x_1, x_2]$ and $\vec{Y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$;

we see that equation (2) says that

3.7.2 continued

$$\alpha\cdot\beta = \vec{X}\, A\, \overset{\downarrow}{Y} = [x_1 \quad x_2] \begin{bmatrix} u_1\cdot u_1 & u_1\cdot u_2 \\ u_2\cdot u_1 & u_2\cdot u_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad (2')$$

As a check on the equivalence of (2) and (2'), we have

$$\vec{X}\, A\, \overset{\downarrow}{Y} = [x_1(u_1\cdot u_1) + x_2(u_2\cdot u_1),\ x_1(u_1\cdot u_2) + x_2(u_2\cdot u_2)] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= [x_1(u_1\cdot u_1) + x_2(u_2\cdot u_1)]y_1 + [x_1(u_1\cdot u_2) + x_2(u_2\cdot u_2)]y_2$$

$$= x_1y_1(u_1\cdot u_1) + x_2y_1(u_2\cdot u_1) + x_1y_2(u_1\cdot u_2) + x_2y_2(u_2\cdot u_2) \qquad (3)$$

and it is easily seen that the number named by (3) is the same as that named by (2).

Note #2

The converse of Note #1 is also true. Namely, suppose we let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

denote any 2 by 2 matrix. We may now use A to induce a bilinear function f on V x V. Namely, if $V = [u_1,u_2]$, we define f as

$$f(u_1,u_1) = u_1\cdot u_1 = a_{11};\ f(u_1,u_2) = u_1\cdot u_2 = a_{12};\ f(u_2,u_1) = u_2\cdot u_1 = a_{21};\ \text{and } u_2\cdot u_2 = a_{22}.$$

We then use (2) to extend the definition to all of V.

This is the essence of part (b) wherein we illustrate our last remark more concretely.

b. Let $V = [u_1,u_2]$, then if $\alpha = (3,2)$ [i.e., $3u_1 + 2u_2$] and $\beta = (1,4)$, then

$$\alpha\cdot\beta = [3 \quad 2] \begin{bmatrix} 1 & -1 \\ -5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

3.7.2 continued

$$= [3 - 10, \ -3 + 12] \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$= [-7 \quad 9] \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$= -7 + 36$$

$$= 29. \tag{4}$$

In "long hand" our matrix A codes the bilinear function defined by

$u_1 \cdot u_1 = 1$, $u_1 \cdot u_2 = -1$, $u_2 \cdot u_1 = -5$, and $u_2 \cdot u_2 = 6$.

We would then obtain

$$\begin{aligned}(3u_1 + 2u_2)\cdot(u_1 + 4u_2) &= 3(u_1\cdot u_1) + 12(u_1\cdot u_2) + 2(u_2\cdot u_1) \\ &\quad + 8(u_2\cdot u_2) \\ &= 3(1) + 12(-1) + 2(-5) + 8(6) = 29,\end{aligned}$$

which checks with (4).

<u>Note #3</u>

What we have essentially shown is that if dim V = n (although we used n = 2) then there are as many bilinear functions which can be defined on V (relative to a fixed basic $\{u_1, \ldots, u_n\}$) as there are n by n matrices. Namely, if $A = [a_{ij}]$ is any n by n matrix, we define a bilinear function by $u_i \cdot u_j = a_{ij}$ whereupon

$$(x_1, \ldots, x_n)\cdot(y_1, \ldots, y_n) = [x_1 \quad x_n] \begin{bmatrix} a_1, \ldots a_{1n} \\ \vdots \qquad \vdots \\ a_n, \ldots a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ \\ y_n \end{bmatrix}$$

is <u>the</u> bilinear function induced by A.

3.7.2 continued

Note #4

Since there are many different bases for the same vector, two different n by n matrices may code the same bilinear function but with respect to different bases. When we become concerned with this problem we have yet another situation in which we must define what we mean for two matrices to be equivalent. That is, in the same way that we have identified different matrices as being equivalent (in this case, called similar) if they denote the same linear transformation of V but with respect to (possibly) different bases, we might want to identify different matrices as being equivalent if they define the same dot product. Certainly these two definitions of "equivalent" need not be the same. That is, two matrices may code the same linear transformation but not the same dot product.

3.7.3

Hopefully, our treatment in the first two exercises makes it clear that one does not need the bilinear function to be commutative in order for it to have a nice structure. We observe, however, that the "usual" dot product does obey the commutative rule. That is, for any "arrows" $\vec{b}$ and $\vec{c}$, $\vec{b} \cdot \vec{c} = \vec{c} \cdot \vec{b}$.

Thus, to capture, even more, the flavor of the ordinary dot product we add another property to the bilinear function f, namely that $f(\alpha, \beta) = f(\beta, \alpha)$; or in the language of the dot product, $\alpha \cdot \beta = \beta \cdot \alpha$.

We call the resulting bilinear form a symmetric bilinear form, and we notice that nothing new happens that didn't already occur in our previous treatment except now all of our matrices which represent the dot product are also symmetric. In other words, for a symmetric bilinear function we know that $u_1 \cdot u_2 = u_2 \cdot u_1$, and even more generally that $u_i \cdot u_j = u_j \cdot u_i$. Hence the matrix

$$[u_i \cdot u_j] = \begin{bmatrix} u_1 \cdot u_1, \ldots, u_1 \cdot u_n \\ \vdots \qquad \vdots \\ u_n \cdot u_1, \ldots, u_n \cdot u_n \end{bmatrix}$$

is a symmetric matrix (that is, we get the same matrix if we interchange rows and columns). In any event, it is the study

3.7.3 continued

of symmetric bilinear functions that makes the study of symmetric matrices especially important.

a. If we assume that $V = [u_1, u_2]$ then the symmetric matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \tag{1}$$

codes the symmetric bilinear function defined by

$$\left.\begin{matrix} u_1 \cdot u_1 = 1 & u_1 \cdot u_2 = 2 \\ u_2 \cdot u_1 = 2 & u_2 \cdot u_2 = 3 \end{matrix}\right\} \tag{2}$$

We then have that

$$(x_1u_1 + x_2u_2) \cdot (y_1u_1 + y_2u_2)$$

$$= x_1y_1(u_1 \cdot u_1) + x_1y_2(u_1 \cdot u_2) + x_2y_1(u_2 \cdot u_1) + x_2y_2(u_2 \cdot u_2)$$

$$= x_1y_1 + 2(x_1y_2 + x_2y_1) + 3x_2y_2. \tag{3}$$

Had we wished to arrive at (3) using matrix notation, we would have

$$(x_1u_1 + x_2u_2) \cdot (y_1u_1 + y_2u_2) = [x_1 \quad x_2] \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= (x_1 + 2x_2, \; 2x_1 + 3x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= (x_1 + 2x_2)y_1 + (2x_1 + 3x_2)y_2$$

$$= x_1y_1 + 2(x_1y_2 + x_2y_1) + 3x_2y_2,$$

which checks with (3).

3.7.3 continued

b. We know from our geometric treatment of the dot product earlier in our course that the length of the projection of v_2 on v_1 is given by

$$v_2 \cdot \frac{v_1}{|v_1|}$$

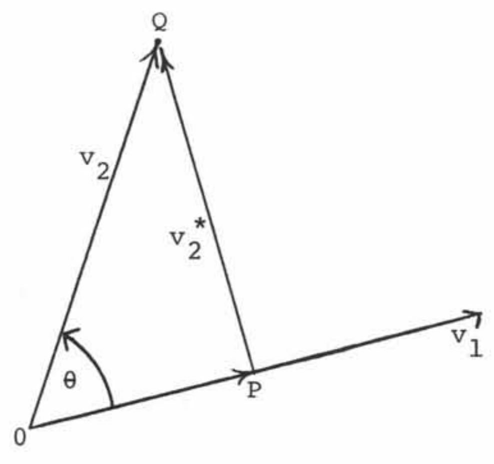
so that the vector projection of v_2 onto v_1 is given by

$$\left(v_2 \cdot \frac{v_1}{|v_1|}\right) \frac{v_1}{|v_1|}$$

$$= \frac{v_2 \cdot v_1}{|v_1|^2} v_1$$

$$= \left(\frac{v_2 \cdot v_1}{v_1 \cdot v_1}\right) v_1.$$

Pictorially,



$$v_2 = \vec{OQ}$$

$$\left(\frac{v_2 \cdot v_1}{v_1 \cdot v_1}\right) v_1 = \vec{OP}$$

$$\vec{OQ} - \vec{OP} = v_2^*$$

$$= v_2 - \left(\frac{v_2 \cdot v_1}{v_1 \cdot v_1}\right) v_1$$

(Figure 1)

3.7.3 continued

We may now divide v_1 and v_2* by their magnitudes to determine an orthonormal basis for the space spanned by v_1 and v_2.

Before proceeding to the actual illustration of this exercise, let us observe that while it was convenient to use geometry to construct v_2*, it was not necessary to do so. We could have done the same thing algebraically (axiomatically). More importantly, the algebraic approach remains valid for $n \geq 3$ while the geometric approach doesn't.

The algebraic determination of v_2* comes from observing that $\vec{OP} = xv_1$ (where x is to be determined), that $xv_1 + v_2* = v_2$ (or $v_2* = v_2 - xv_1$) and that $v_2* \cdot v_1 = 0$. Namely,

$$v_2* \cdot v_1 = 0 \rightarrow$$

$$(v_2 - xv_1) \cdot v_1 = 0 \rightarrow$$

$$(v_2 \cdot v_1) - x(v_1 \cdot v_1) = 0 \rightarrow$$

$$x = \frac{v_2 \cdot v_1}{v_1 \cdot v_1}, \qquad (4)$$

so that v_2* must be given by

$$v_2* = v_2 - \left(\frac{v_2 \cdot v_1}{v_1 \cdot v_1} \right) v_1 \qquad (5)$$

which agrees with the geometric situation.

Turning now to our specific exercise, we have

$$\left. \begin{aligned} v_1 &= \vec{i} + 2\vec{j} + 3\vec{k} \\ v_2 &= 2\vec{i} + 5\vec{j} - 2\vec{k} \end{aligned} \right\} \qquad (6)$$

Hence,

$$\left. \begin{aligned} v_1 \cdot v_1 &= 1 + 4 + 9 = 14 \\ v_2 \cdot v_2 &= 4 + 25 + 4 = 33 \\ v_2 \cdot v_1 &= 2 + 10 - 6 = 6 \end{aligned} \right\} . \qquad (7)$$

3.7.3 continued

Then,

$$v_2^* = v_2 - \left(\frac{v_2 \cdot v_1}{v_1 \cdot v_1}\right) v_1$$

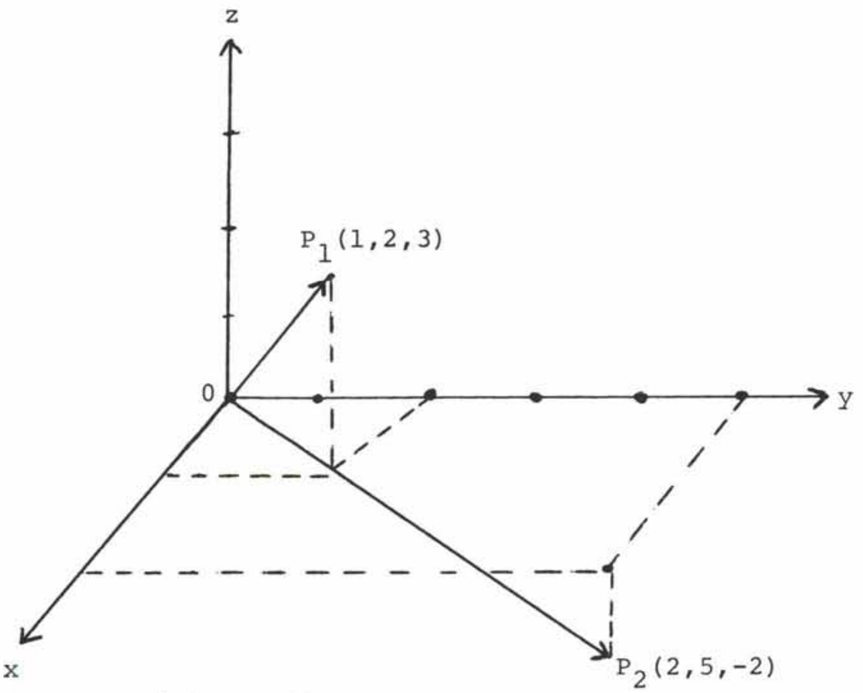
$$= v_2 - \frac{6}{14} v_1$$

$$= 2\vec{i} + 5\vec{j} - 2\vec{k} - \frac{3}{7}(\vec{i} + 2\vec{j} + 3\vec{k})$$

$$= \frac{11}{7}\vec{i} + \frac{29}{7}\vec{j} - \frac{23}{7}\vec{k}$$

$$= \frac{1}{7}(11\vec{i} + 29\vec{j} - 23\vec{k}). \qquad (8)$$

Pictorially,



(Figure 2)

v_2^* is the vector projection of $\vec{OP}_2$ onto $\vec{OP}_1$.

$v_1 = \vec{OP}_1$
$v_2 = \vec{OP}_2$.

The space spanned by v_1 and v_2 is the plane determined by 0, P_1, and P_2.

3.7.3 continued

As a check that v_1 and $v_2{}^*$ is an orthogonal basis for the space spanned by v_1 and v_2, we need only check that $v_2{}^*$ belongs to this space, and $v_1 \cdot v_2{}^* = 0$.

That $v_1 \cdot v_2{}^* = 0$ may be checked at once from (6). In particular,

$$v_1 \cdot v_2{}^* = (1,2,3) \cdot \frac{1}{7}(11,29,-23)$$

$$= \frac{1}{7}(11 + 58 - 69)$$

$$= 0.$$

As to whether $v_2{}^*$ belongs to the space spanned by v_1 and v_2, we need simply row-reduce the matrix

$$\begin{array}{ccccc} \vec{i} & \vec{j} & \vec{k} & v_1 & v_2 \end{array}$$
$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 \\ 2 & 5 & -2 & 0 & 1 \end{bmatrix}$$

to obtain

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 \\ 0 & 1 & -8 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 19 & 5 & 1 \\ 0 & 1 & -8 & -2 & 1 \end{bmatrix}$$

from which we conclude that $S(v_1,v_2) = \{x_1(1,0,19) + x_2(0,1,-8)\}$ or,

$$\{(x_1,x_2, 19x_1 - 8x_2)\}.$$

In particular,

$$(11,29,z) \varepsilon S(v_1,v_2) \leftrightarrow z = 19(11) - 8(29)$$

$$\leftrightarrow z = 209 - 232$$

$$\leftrightarrow z = -23.$$

Hence, $(11,29,-23) \varepsilon S(v_1,v_2)$. Therefore, $v_2{}^* = \frac{1}{7}(11,29,-23) \varepsilon S(v_1,v_2)$.

3.7.3 continued

Note:
Do not let the fact that we are now dealing with dot products make you forget that one uses the row-reduced matrix technique quite apart from any knowledge of a dot product. In particular, we analyze the space $S(v_1, v_2)$ in the same way as we would have in the first three units of this Block. What we should be beginning to feel now is the number of different ways in which matrix notation is used to code completely different aspects of vector spaces.

To conclude this part of the exercise we need only divide both v_1 and $v_2{}^*$ by their magnitudes and we have obtained the desired orthonormal basis. To this end we have:

$$|v_1| = \sqrt{v_1 \cdot v_1} = \sqrt{14}$$

$$v_2{}^* = \sqrt{v_2{}^* \cdot v_2{}^*} = \frac{1}{7}\sqrt{121 + 841 + 529} = \frac{1}{7}\sqrt{1491}\ .$$

Hence, we let

$$u_1 = \frac{1}{\sqrt{14}} v_1 = \frac{1}{\sqrt{14}} (\vec{i} + 2\vec{j} + 3\vec{k})$$

$$u_2 = \frac{7}{\sqrt{1491}} v_2{}^* = \frac{1}{\sqrt{1491}} (11\vec{i} + 29\vec{j} - 23\vec{k}).$$

Then,

$$S(v_1, v_2) = [u_1, u_2] \text{ where } u_1 \cdot u_1 = u_2 \cdot u_2 = 1 \text{ and } u_1 \cdot u_2 = 0.$$

c. At first glance it might seem that part (b) was an interruption of our train of thought in going from part (a) to part (c), but this is hardly the case. Rather what we wanted to show is that the technique used in part (b) which is so obvious from a geometric point of view works, word for word, in the more abstract situations.

In particular, in this exercise we want to express the space $V = [u_1, u_2]$ described in part (a) by replacing u_2 by an appropriate $u_2{}^*$, where $u_1 \cdot u_2{}^* = 0$.

3.7.3 continued

To this end, we need simply take

$$u_2{}^* = u_2 - \left[\frac{u_2 \cdot u_1}{u_1 \cdot u_1} \right] u_1$$

$$= u_2 - \left[\frac{2}{1} \right] u_1$$

$$= -2u_1 + u_2. \qquad (9)$$

That is,

$V = [u_1{}^*, u_2{}^*]$

where

$$\begin{cases} u_1{}^* = u_1 \\ u_2{}^* = -2u_1 + u_2. \end{cases} \qquad (10)$$

Moreover, we see from (10) that

$$u_1{}^* \cdot u_1{}^* = u_1 \cdot u_1 = 1$$

$$u_1{}^* \cdot u_2{}^* = u_1 \cdot (-2u_1 + u_2) = -2u_1 \cdot u_1 + u_1 \cdot u_2 = -2 + 2 = 0$$

$$u_2{}^* \cdot u_2{}^* = (-2u_1 + u_2) \cdot (-2u_1 + u_2) = 4u_1 \cdot u_1 - 4u_1 \cdot u_2 + u_2 \cdot u_2$$

$$= 4 - 8 + 3 = -1. \qquad (11)$$

Note #1:

As we may see from (11), the matrix of our symmetric bilinear function relative to the basis $\{u_1{}^*, u_2{}^*\}$ is given by

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \qquad (12)$$

That is, the matrix A of part (a) and the matrix B of (12), code the same symmetric bilinear function but with respect to a different basis.

The matrix B shows us why $\{u_1{}^*, u_2{}^*\}$ is a "better" basis than $\{u_1, u_2\}$ at least with respect to the given bilinear function.

3.7.3 continued

Namely from (12) we see that

$$(a_1u_1^* + a_2u_2^*) \cdot (b_1u_1^* + b_2u_2^*)$$

$$= a_1b_1\underbrace{u_1^* \cdot u_1^*}_{=1} + a_1b_2\underbrace{u_1^* \cdot u_2^*}_{=0} + a_2b_1\underbrace{u_2^* \cdot u_1}_{=0} + a_2b_2\underbrace{u_2^* \cdot u_2^*}_{=-1}$$

$$= a_1b_1 - a_2b_2. \qquad (13)$$

In matrix form, (13) may be derived from

$$[a_1 \quad a_2]\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}b_1\\ b_2\end{bmatrix}$$

$$= [a_1 \quad -a_2]\begin{bmatrix}b_1\\ b_2\end{bmatrix}$$

$$= a_1b_1 - a_2b_2.$$

<u>Note #2</u>:
We may use (11) to find a matrix P such that $PAP^T = B$. That is, we may diagonalize a symmetric matrix A by appropriately choosing P. This choice of P comes from the following.

1. We know that $u_1^* \cdot u_1^* = u_1 \cdot u_1$, so relative to $\{u_1, u_2\}$ as a basis we may code this piece of information by writing

$$[1 \quad 0]\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} \left(= [1 \quad 2]\begin{bmatrix}1\\ 0\end{bmatrix} = 1\right). \qquad (14)$$

Since $u_2^* = -2u_1 + u_2$, we may represent $u_2^* \cdot u_2^*$ by

$$(-2 \quad 1]\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}-2\\ 1\end{bmatrix} \left(= [0 \quad -1]\begin{bmatrix}-2\\ 1\end{bmatrix} = -1\right). \qquad (15)$$

Putting (14) and (15) into a single form we have

3.7.3 continued

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

where

$$P = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

and

$$P^T = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Note #3:
The procedure used in obtaining the diagonal matrix B from the symmetric matrix A generalizes to n-dimensional space. The strange part, at least in terms of our usual experience, is that the diagonalized matrix may include negative entries as well as zero among its diagonal entries. For example, with respect to (12), notice that -1 appears on the diagonal.

What this means is that if we still want to identify $v \cdot v$ with the "length" of v, then some symmetric bilinear functions produce the "disturbing" results that a vector can have an immaginary length and that a non-zero vector can have a zero length. While it is understandable that such a result may seem disturbing, it is important that we try to understand the structure of mathematics to an extent which allows us to accept such results in a natural way. In essence when we say that we want to think of $v \cdot v$ as the square of the magnitude of v, we have to realize that unless our axiomatic approach has captured every pertinent ingredient of the system we are trying to characterize, it is possible that certain concepts may possess strange properties in the generalized, abstract system. This happened to us in Block 1 when we introduced modular arithmetic as an example that obeyed much of the structure of ordinary arithmetic, yet was different enough to cause us some startling results.

3.7.3 continued

In any event, let us observe that this exercise shows us that relative to our present axioms, a symmetric bilinear function may define a dot product for which $v \cdot v < 0$ or $v \cdot v = 0$ even if $v \neq 0$.

Note #4:
Pursuing the previous note in somewhat more detail, we see from (13) that relative to the basis $\{u_1^*, u_2^*\}$

$$(x,y) \cdot (x,y) = x^2 - y^2. \tag{16}$$

That is,

$$(x,y) = (xu_1^* + yu_2^*),$$

hence

$$\begin{aligned}(x,y) \cdot (x,y) &= (xu_1^* + yu_2^*) \cdot (xu_1^* + yu_2^*)\\ &= x^2\underbrace{(u_1^* \cdot u_1^*)}_{=1} + 2xy\underbrace{(u_1^* \cdot u_2^*)}_{=0} + y^2\underbrace{(u_2^* \cdot u_2^*)}_{=-1}\\ &= x^2 - y^2.\end{aligned}$$

Thus, using (16), we see that if $v = xu_1^* + yu_2^*$, then

$$\begin{aligned}v \cdot v = 0 &\leftrightarrow x^2 - y^2 = 0\\ &\leftrightarrow x = \pm y.\end{aligned}$$

In other words, the vectors v, which are non-zero scalar multiples of $u_1^* \pm u_2^*$, are distinguished by the fact that they are null vectors (v is called a null vector if $v \neq 0$ but $v \cdot v = 0$).

We may use (10) to investigate how null vectors look relative to the basis $\{u_1, u_2\}$. Namely,

$$\begin{aligned}u_1^* + u_2^* &= u_1 + (-2u_1 + u_2)\\ &= -u_1 + u_2\end{aligned}$$

3.7.3 continued

while

$$u_1^* - u_2^* = u_1 - (-2u_1 + u_2)$$

$$= 3u_1 - u_2.$$

Check:

$$(-u_1 + u_2) \cdot (-u_1 + u_2) = u_1 \cdot u_1 - 2u_1 \cdot u_2 + u_2 \cdot u_2$$

$$= 1 - 4 + 3 = 0$$

$$(3u_1 - u_2) \cdot (3u_1 - u_2) = 9u_1 \cdot u_1 - 6u_1 \cdot u_2 + u_2 \cdot u_2$$

$$= 9 - 12 + 3 = 0.$$

Note #5:

The material discussed in Notes 3 and 4 leads to an application of symmetric bilinear functions in the study of the algebraic topic known as quadratic forms. In n variables a quadratic form is any expression of the type

$$a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 + \sum_{i<j} b_{ij}x_ix_j. \tag{17}$$

In the special case n = 2, (17) becomes

$$a_{11}x_1^2 + b_{12}x_1x_2 + a_{22}x_2^2. \tag{18}$$

The connection between quadratic forms and symmetric bilinear functions should become clear rather quickly once we realize that if we let $a_{12} = \frac{1}{2} b_{12}$, then

$$[x_1 \quad x_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

$$= a_{11}x_1^2 + b_{12}x_1x_2 + a_{22}x_2^2.$$

3.7.3 continued

In other words, any equation of the form

$$a_{11}x_1^2 + b_{12}x_1x_2 + a_{22}x_2^2 = m$$

may be viewed as the vector equation

$$v \cdot v = m$$

where $v \cdot v$ is defined by the symmetric matrix

$$\begin{bmatrix} a_{11} & \frac{1}{2}b_{12} \\ \frac{1}{2}b_{12} & a_{22} \end{bmatrix}.$$

Relative to our present exercise, we have that

$$x_1^2 + 4x_1x_2 + 3x_2^2 = m \qquad (19)$$

is equivalent to solving $v \cdot v = m$ where $v = x_1u_1 + x_2u_2$. That is,

$$v \cdot v = [x_1 \quad x_2] \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

More concretely if we let $m = 0$ in (19), then we want to solve the equation

$$x_1^2 + 4x_1x_2 + 3x_2^2 = 0. \qquad (20)$$

This can be done rather simply here (since there are only two variables) by completing the square to obtain

$$x_1^2 + 4x_1x_2 + 4x_2^2 - x_2^2 = 0$$

or

$$(x_1 + 2x_2)^2 - x_2^2 = 0$$

3.7.3 continued

or

$x_2 = \pm (x_1 + 2x_2)$.

The key computational point is that relative to the basis $\{u_1^*, u_2^*\}$ equation (20) may be written as

$$y_1^2 - y_2^2 = 0$$

where now $v = y_1u_1^* + y_2u_2^*$.

In summary, the process of replacing the symmetric matrix A by the diagonal matrix B allows us to replace a quadratic form by an equivalent form in which only the square terms are present.

3.7.4

a. Since the matrix

$$A = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \qquad (1)$$

is symmetric, it codes a symmetric bilinear function. In particular, if $V = [u_1, u_2]$, then the matrix A codes the symmetric bilinear function defined by

$$\left.\begin{matrix} u_1 \cdot u_1 = 3 & u_1 \cdot u_2 = 4 \\ u_2 \cdot u_1 = 4 & u_2 \cdot u_2 = 5 \end{matrix}\right\} \qquad (2)$$

where we have written (2) in a form which we hope suggests the matrix A given in (1). We then have

$$\begin{aligned}(x_1, x_2) \cdot (y_1, y_2) &= [x_1 \quad x_2] \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= [3x_1 + 4x_2, \quad 4x_1 + 5x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= (3x_1y_1 + 4x_2y_1) + (4x_1y_2 + 5x_2y_2) \\ &= 3x_1y_1 + 4x_1y_2 + 4x_2y_1 + 5x_2y_2. \qquad (3)\end{aligned}$$

3.7.4 continued

In particular, we see from (3) that if $(x_1,x_2) = (y_1,y_2)$ then

$$(x_1,x_2) \cdot (x_1,x_2) = 3x_1^2 + 8x_1x_2 + 5x_2^2 \tag{4}$$

By way of review, the right side of (4) is known as a quadratic form. We may always view

$$ax_1^2 + 2bx_1x_2 + cx_2^2 \tag{5}$$

as $v \cdot v$ where the dot product is defined by the symmetric matrix

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

with a, b, and c as in (5).

b. So far we have that

$$v = [u_1, u_2]$$

and

$$\left.\begin{array}{l} u_1 \cdot u_1 = 3 \\ u_1 \cdot u_2 = u_2 \cdot u_1 = 4 \\ u_2 \cdot u_2 = 5 \end{array}\right\} \tag{6}$$

We also know that if $\{u_1, u_2\}$ is a basis for V so also is $\{u_1 + xu_2, u_2\}$ where x is any real number. So what we try to do is choose x so that

$$(u_1 + xu_2) \cdot u_2 = 0. \tag{7}$$

To solve (7) we see that

$$u_1 \cdot u_2 + x(u_2 \cdot u_2) = 0$$

so that

3.7.4 continued

$$x = \frac{-u_1 \cdot u_2}{u_2 \cdot u_2} . \quad (8)$$

Using (6), we see that (8) becomes

$$x = -\frac{4}{5} .$$

We now replace u_1 by

$$u_1{}^* = u_1 - \frac{4}{5} u_2 \quad (9)$$

and we choose as our new basis for V, $\{u_1{}^*, u_2\}$.

From (9)

$$u_1{}^* \cdot u_1{}^* = (u_1 - \frac{4}{5} u_2) \cdot (u_1 - \frac{4}{5} u_2)$$

$$= u_1 \cdot u_1 - \frac{4}{5} u_2 \cdot u_1 - \frac{4}{5} u_1 \cdot u_2 + \frac{16}{25} u_2 \cdot u_2,$$

so by (6),

$$u_1{}^* \cdot u_1{}^* = 3 - \frac{4}{5} (4 + 4) + \frac{16}{25} (5)$$

$$= 3 - \frac{32}{5} + \frac{16}{5}$$

$$= -\frac{1}{5} . \quad (10)$$

We also know that

$$u_1{}^* \cdot u_2 = 0 \quad (11)$$

since this is how $u_1{}^*$ was chosen.

<u>Check</u>:

$$u_1{}^* \cdot u_2 = (u_1 - \frac{4}{5} u_2) \cdot u_2$$

$$= u_1 \cdot u_2 - \frac{4}{5} u_2 \cdot u_2$$

$$= 4 - \frac{4}{5} (5) = 0.$$

3.7.4 continued

Finally

$$u_2 \cdot u_2 = 5. \tag{12}$$

Hence, from (10), (11), and (12) we conclude that the matrix of the given dot product relative to the basis $\{u_1^*, u_2\}$ is

$$\begin{bmatrix} -\frac{1}{5} & 0 \\ 0 & 5 \end{bmatrix}. \tag{13}$$

In other words, if $v \varepsilon V$ is given by $v = y_1u_1^* + y_2u_2$, then $v \cdot v = -\frac{1}{5}y_1{}^2 + 5y_2{}^2$ and the y_1y_2 term is missing. Again, by way of review, (13) may be obtained from (1) as follows. We know that

$$u_1' \cdot u_1' = (u_1 - \frac{4}{5}u_2) \cdot (u_1 - \frac{4}{5}u_2)$$

$$= [1, -\frac{4}{5}] \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{4}{5} \end{bmatrix} \tag{14}$$

and that

$$u_2 \cdot u_2 = [0 \quad 1] \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{15}$$

(14) and (15) may be combined into the single form

$$\begin{bmatrix} 1 & -\frac{4}{5} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{4}{5} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{5} & 0 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{4}{5} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{5} & 0 \\ 0 & 5 \end{bmatrix}.$$

3.7.4 continued

c. What (b) shows is that the equation $3x_1^2 + 8x_1x_2 + 5x_2^2 = m$ is equivalent to $-\frac{1}{5}y_1^2 + 5y_2^2 = m$ where the x's and y's are related as follows.

If $v = y_1u_1^* + y_2u_2^*$, then

$$v = y_1(u_1 - \tfrac{4}{5}u_2) + y_2u_2$$

$$= y_1u_1 + (y_2 - \tfrac{4}{5}y_1)u_2$$

$$= x_1u_1 + x_2u_2$$

where $x_1 = y_1$ and $x_2 = y_2 - \frac{4}{5}y_1$. In particular,

$$-\tfrac{1}{5}y_1^2 + 5y_2^2 = 0 \leftrightarrow$$

$$25y_2^2 = y_1^2 \leftrightarrow$$

$$y_2 = \pm\tfrac{1}{5}y_1 .$$

Hence,

$v = y_1u_1^* + y_2u_2^*$ is a null vector $\leftrightarrow y_2 = \pm\frac{1}{5}y_1$

$\leftrightarrow$ v is a non zero scalar multiple of $u_1^* \pm \frac{1}{5}u_2^*$

$\leftrightarrow$ v is a non zero scalar multiple of $(u_1 - \frac{4}{5}u_2) \pm \frac{1}{5}u_2$

$\leftrightarrow$ $v = k(u_1 - u_2)$ or $k(u_1 - \frac{3}{5}u_2)$, k an arbitrary non-zero constant.

<u>Check:</u>

$$(u_1 - u_2)\cdot(u_1 - u_2) = \underbrace{u_1\cdot u_1}_{=3} - 2\underbrace{u_1\cdot u_2}_{=4} + \underbrace{u_2\cdot u_2}_{=5}$$

$$= 3 - 8 + 5 = 0$$

3.7.4 continued

$$(u_1 - \frac{3}{5}u_2)\cdot(u_1 - \frac{3}{5}u_2) = \underbrace{u_1\cdot u_1}_{=3} - \frac{6}{5}\underbrace{u_1\cdot u_2}_{=4} + \frac{9}{25}\underbrace{u_2\cdot u_2}_{=5}$$

$$= 3 - \frac{24}{5} + \frac{9}{5}$$

$$= 0.$$

3.7.5 (Review of the Lecture)

As described in the lecture, the process of finding an orthogonal basis may be extended to an n-dimensional space. The procedure is inductive. Namely, we have already seen that if $V = [u_1, u_2, \ldots, u_n]$ then we may replace u_2 by

$$u_2^* = u_2 - \left[\frac{u_2\cdot u_1}{u_1\cdot u_1}\right]u_1, \qquad (1)$$

whereupon, $V = [u_1, u_2^*, \ldots, u_n]$, but now $u_1 \cdot u_2^* = 0$. Once we know that u_1 and u_2 are orthogonal we use the fact that if u_3 is replaced by u_3^* where u_3^* had the form:

$$u_3 - x_1u_1 - x_2u_2$$

then u_1, u_2, and u_3^* span the same space as to u_1, u_2, and u_3. We then try to determine x_1 and x_2 in such a way that $u_1 \cdot u_3^* = u_2 \cdot u_3^* = 0$.

As long as neither $u_1 \cdot u_1$ nor $u_2 \cdot u_2 = 0$, this can always be done. Namely, given that

$$u_3^* = u_3 - x_1u_1 - x_2u_2 \qquad (2)$$

and that $u_1 \cdot u_2 = 0$, we may dot both sides of (2) by u_1 to obtain

$$u_1\cdot u_3^* = u_1\cdot(u_3 - x_1u_1 - x_2u_2)$$

$$= u_1\cdot u_3 - x_1(u_1\cdot u_1) - x_2\underbrace{(u_1\cdot u_2)}_{=0}, \qquad (3)$$

3.7.5 continued

whereupon the condition that $u_1 \cdot u_3^*$ is to equal 0 allows us to conclude from (3) that

$$x_1 = \frac{u_1 \cdot u_3}{u_1 \cdot u_1}. \tag{4}$$

Similarly, if we had dotted both sides of (3) with u_2 and set $u_2 \cdot u_3^*$ equal to 0, we would have obtained

$$x_2 = \frac{u_2 \cdot u_3}{u_2 \cdot u_2}. \tag{5}$$

So, putting (4) and (5) into (2) would yield

$$u_3^* = u_3 - \left(\frac{u_1 \cdot u_3}{u_1 \cdot u_1} \right) u_1 - \left(\frac{u_2 \cdot u_3}{u_2 \cdot u_2} \right) u_2. \tag{6}$$

Equation (6) represents a generalization of our approach in Exercise 3.7.3. Notice that

$$\left(\frac{u_1 \cdot u_3}{u_1 \cdot u_1} \right) u_1$$

is the vector projection of u_3 onto u_1, while

$$\left(\frac{u_2 \cdot u_3}{u_2 \cdot u_2} \right) u_2$$

is the vector projection of u_3 onto u_2. The procedure indicated in (6) is known as the Gram-Schmidt Orthogonalization Process.

It works as follows. Suppose $\{u_1, \ldots, u_n\}$ is an orthogonal set and that $u_1 \cdot u_1, \ldots,$ and $u_{n-1} \cdot u_{n-1}$ are all different from zero. Given u_n, we let

$$u_n^* = u_n - x_1 u_1 - \cdots - x_{n-1} u_{n-1}. \tag{7}$$

We then dot both sides with u_k where $k = 1, \ldots,$ or $n - 1$ and this yields

$$u_k \cdot u_n^* = u_k \cdot u_n - x_1 u_k \cdot u_1 \ldots - x_k u_k \cdot u_k - \ldots - x_{n-1} u_k \cdot u_{n-1}. \tag{8}$$

3.7.5 continued

Since $u_k \cdot u_k \neq 0$ but $u_k \cdot u_i = 0$ for $i \neq k$, we may set $u_k \cdot u_n^*$ equal to 0 in (8) to obtain

$$x_k = \frac{u_k \cdot u_n}{u_k \cdot u_k}, \ k = 1,\ldots,n - 1. \qquad (9)$$

Putting the result of (9) into (7) we have that if $u_i \cdot u_j = 0$ for $i \neq j$; $i, j = 1,\ldots,n - 1$ then if

$$u_n^* = u_n - \sum_{k=1}^{n-1} \left(\frac{u_k \cdot u_n}{u_k \cdot u_k}\right) u_k, \qquad (10)$$

$u_k \cdot u_n^* = 0.$

Looking at (10) one dimension at a time, notice that the coefficient of u_k is the vector projection of u_n onto u_k. In other words, the coefficient of u_k has its numerator equal to (minus) $u_k \cdot u_n$ and its denominator equal to $u_k \cdot u_k$. With this as background, we proceed with the present exercise.

a. With $V = [u_1,u_2,u_3]$, the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix} \qquad (11)$$

defines the symmetric bilinear function given by

$$\left.\begin{matrix} u_1 \cdot u_1 = 1 & u_1 \cdot u_2 = 1 & u_1 \cdot u_3 = 1 \\ u_2 \cdot u_1 = 1 & u_2 \cdot u_2 = 2 & u_2 \cdot u_3 = 3 \\ u_3 \cdot u_1 = 1 & u_3 \cdot u_2 = 3 & u_3 \cdot u_3 = 4 \end{matrix}\right\} \qquad (12)$$

Hence to find an orthogonal basis for V we begin by letting

$$u_1^* = u_1 \qquad (13)$$

3.7.5 continued

and

$$u_2^* = u_2 - \left(\frac{u_1 \cdot u_2}{u_1 \cdot u_1}\right) u_1;$$

or by (12),

$$u_2^* = u_2 - \left(\frac{1}{1}\right) u_1$$

$$= u_2 - u_1. \qquad (14)$$

Hence,

$$V = [u_1^*, u_2^*, u_3] \text{ and } u_1^* \cdot u_2^* = 0. \qquad (15)$$

Since u_1^* and u_2^* are orthogonal, we use the Gram-Schmidt process to replace u_3 by

$$u_3^* = u_3 - \left(\frac{u_1 \cdot u_3}{u_1^* \cdot u_1^*}\right) u_1^* - \left(\frac{u_2^* \cdot u_3}{u_2^* \cdot u_2^*}\right) u_2^*. \qquad (16)$$

Using (11), (13), and (14) we have

$$u_1^* \cdot u_3 = u_1 \cdot u_3 = 1$$

$$u_1^* \cdot u_1^* = u_1 \cdot u_1 = 1$$

$$u_2^* \cdot u_3 = (u_2 - u_1) \cdot u_3 = u_2 \cdot u_3 - u_1 \cdot u_3 = 3 - 1 = 2$$

$$u_2^* \cdot u_2^* = (u_2 - u_1) \cdot (u_2 - u_1) = u_2 \cdot u_2 - 2u_1 \cdot u_2 + u_1 \cdot u_1$$
$$= 2 - 2 + 1 = 1. \qquad (17)$$

Thus, using (17), we see from (16) that

$$u_3^* = u_3 - u_1 - 2u_2^* = u_3 - u_1 - 2(u_2 - u_1)$$

$$= u_1 - 2u_2 + u_3. \qquad (18)$$

From (13), (14), and (18) we conclude that an orthogonal basis for V is given by

3.7.5 continued

$$\{u_1, -u_1 + u_2, u_1 - 2u_2 + u_3\} \tag{19}$$

<u>Check</u>:

$$u_1 \cdot (-u_1 + u_2) = -(u_1 \cdot u_1) + u_1 \cdot u_2 = -1 + 1 = 0$$

$$u_1 \cdot (u_1 - 2u_2 + u_3) = (u_1 \cdot u_1) - 2(u_1 \cdot u_2) + u_1 \cdot u_3$$

$$= 1 - 2 + 1 = 0$$

$$(-u_1 + u_2) \cdot (u_1 - 2u_2 + u_3) = -(u_1 . u_1) + 2(u_1 \cdot u_2)$$

$$-(u_1 \cdot u_3) + (u_1 \cdot u_2)$$

$$- 2(u_2 \cdot u_2) + u_2 \cdot u_3$$

$$= -1 + 2 - 1 + 1 - 4 + 3 = 0.$$

b. From (17) we already know that $u_1^* \cdot u_1^* = 1$ and $u_2^* \cdot u_2^* = 1$. From (18),

$$u_3^* \cdot u_3^* = (u_1 - 2u_2 + u_3) \cdot (u_1 - 2u_2 + u_3)$$

$$= (u_1 \cdot u_1) + 4(u_2 \cdot u_2) + (u_3 \cdot u_3)$$

$$- 4(u_1 \cdot u_2) + 2(u_1 \cdot u_3) - 4(u_2 \cdot u_3)$$

$$= 1 + 8 + 4 - 4 + 2 - 12$$

$$= -1.$$

Hence, relative to $\{u_1^*, u_2^*, u_3^*\}$, $v = x_1u_1^* + x_2u_2^* + x_3u_3^* \rightarrow$

$$v \cdot v = x_1^2 \underbrace{u_1^* \cdot u_1^*}_{=1} + x_2^2 \underbrace{u_2^* \cdot u_2^*}_{=1} + x_3^2 \underbrace{u_3^* \cdot u_3^*}_{=-1}$$

$$= x_1^2 + x_2^2 - x_3^2.$$

Therefore,

3.7.5 continued

$$v \cdot v = 0 \leftrightarrow x_1^2 + x_2^2 - x_3^2 = 0$$

$$\leftrightarrow x_3^2 = x_1^2 + x_2^2. \qquad (20)$$

c. Given that $v = 3u_1^* + 4u_2^* + 5u_3^*$ we see that equation (20) is obeyed. Hence $v \cdot v = 0$, and, accordingly, v is a null vector. Again using (13), (14), and (18) we have

$$v = 3u_1 + 4(u_2 - u_1) + 5(u_1 - 2u_2 + u_3)$$

$$= 4u_1 - 6u_2 + 5u_3.$$

Check:

$$v \cdot v = 16(u_1 \cdot u_1) + 36(u_2 \cdot u_2) + 25(u_3 \cdot u_3) - 48(u_1 \cdot u_2)$$

$$+ 40(u_1 \cdot u_3) - 60(u_2 \cdot u_3)$$

$$= 16 + 72 + 100 - 48 + 40 - 180$$

$$= 0.$$

d.

$$u_1^* \cdot u_1^* = [1 \quad 0 \quad 0]\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1$$

$$u_2^* \cdot u_2^* = [-1 \quad 1 \quad 0]\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = 1$$

$$u_3^* \cdot u_3^* = [1 \quad -2 \quad 1]\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = -1.$$

Hence, putting these three results into a single matrix equation yields

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

3.7.5 continued

e. The equation

$$x_1^2 + 2x_2^2 + 4x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3 = m \quad (21)$$

is equivalent to solving

$$y_1^2 + y_2^2 - y_3^2 = m. \quad (22)$$

Namely (21) evolves from computing $v \cdot v = m$ relative to $\{u_1, u_2, u_3\}$ and obtaining

$$[x_1 \quad x_2 \quad x_3] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = m$$

(i.e., $y = x_1u_1 + x_2u_2 + x_3u_3$) and (22) evolves from solving $v \cdot v = m$ relative to $\{u_1^*, u_2^*, u_3^*\}$; namely,

$$v \cdot v = m \rightarrow [y_1, y_2, y_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = m$$

where $v = y_1u_1^* + y_2u_2^* + y_3u_3^*$.

3.7.6

We now define an <u>inner</u>, or a <u>dot</u>, product on a vector space V to be any mapping of ordered pairs of elements of V into the real numbers, say, $f: V \times V \rightarrow R$ which has the following properties.

For $v_1, v_2, v_3 \varepsilon V$ and $c, \varepsilon R$

(i) $f(v_1, v_2) = f(v_1, v_2)$
(ii) $f(v_1, v_2 + v_3) = f(v_1, v_2) + f(v_1, v_3)$
(iii) $f(v_1, cv_2) = cf(v_1, v_2)$
(iv) $f(v_1, v_1) \geq 0$; and $f(v_1, v_1) = 0 \leftrightarrow v_1 = 0$.

3.7.6 continued

Rewritten in terms of the dot notation we have:

(i) $v_1 \cdot v_2 = v_2 \cdot v_1$
(ii) $v_1 \cdot (v_2 + v_3) = v_1 \cdot v_2 + v_1 \cdot v_3$
(iii) $v_1 \cdot (cv_2) = c(v_1 \cdot v_2)$
(iv) $v_1 \cdot v_1 \geq 0$; $v_1 \cdot v_1 = 0 \leftrightarrow v_1 = 0$.

When a symmetric bilinear function satisfies property (iv), the function is called positive definite. From another point of view, what positive definite means is that when we replace the symmetric matrix by the equivalent diagonal matrix, every entry on the diagonal is a positive (real) number.

The crucial point is that once we have a positive definite, symmetric, bilinear function, then we can identify the traditional definition of a dot product [namely the definition which says that $(a_1,\ldots,a_n) \cdot (b_1,\ldots,b_n) = a_1b_1 + \ldots + a_nb_n$] with our "new" definition.

More specifically, and we shall illustrate this more concretely in the present exercise, if we have a positive definite, symmetric, bilinear function defined on V (and this usually is abbreviated by saying that we have an inner product defined on V), then we may use the Gram-Schmidt Orthogonalization process to find an orthogonal basis for V. Of course, that much we could do before; but now the fact that $u_k^* \cdot u_k^*$ is positive for each member of our orthogonal basis, means that

$$\sqrt{u_k^* \cdot u_k^*}$$

is real. In this case we can let

$$w_k = \frac{u_k^*}{\sqrt{u_k^* \cdot u_k^*}}$$

whereupon $V = [w_1,\ldots,w_n]$ and $\{w_1,\ldots,w_n\}$ is not only orthogonal but also orthonormal, meaning in addition to $w_i \cdot w_j = 0$ if $i \neq j$ that $w_k \cdot w_k = 1$ for each $k = 1,\ldots,n$. This shall be summarized as a note at the end of this exercise.

3.7.6 continued

Turning to this exercise we have that

$$A = \begin{bmatrix} 3 & 4 & 4 \\ 4 & 6 & 5 \\ 4 & 5 & 6 \end{bmatrix} \qquad (1)$$

codes the symmetric bilinear function, defined on $V = [u_1, u_2, u_3]$, by

$$\left.\begin{matrix} u_1 \cdot u_1 = 3 & u_1 \cdot u_2 = 4 & u_1 \cdot u_3 = 4 \\ u_2 \cdot u_1 = 4 & u_2 \cdot u_2 = 6 & u_2 \cdot u_3 = 5 \\ u_3 \cdot u_1 = 4 & u_3 \cdot u_2 = 5 & u_3 \cdot u_3 = 6 \end{matrix}\right\} \qquad (2)$$

Now just as before we let

$$u_1^* = u_1 \qquad (3)$$

and

$$\begin{aligned} u_2^* &= u_2 - \left(\frac{u_1 \cdot u_2}{u_1 \cdot u_1}\right) u_1 \\ &= u_2 - \frac{4}{3} u_1 \\ &= -\frac{4}{3} u_1 + u_2. \end{aligned} \qquad (4)$$

Then $[u_1, u_2] = [u_1^*, u_2^*]$, and $u_1^* \cdot u_2^* = 0$.

We next let

$$u_3^* = u_3 - \left(\frac{u_1^* \cdot u_3}{u_1^* \cdot u_1^*}\right) u_1^* - \left(\frac{u_2^* \cdot u_3}{u_2^* \cdot u_2^*}\right) u_2^*. \qquad (5)$$

Now,

$$u_1^* \cdot u_3 = u_1 \cdot u_3 = 4,$$

$$u_1^* \cdot u_1^* = u_1 \cdot u_1 = 3,$$

3.7.6 continued

$$u_2{}^* \cdot u_2{}^* = (-\tfrac{4}{3}u_1 + u_2) \cdot (-\tfrac{4}{3}u_1 + u_2)$$

$$= \tfrac{16}{9}(u_1 \cdot u_1) - \tfrac{8}{3}(u_1 \cdot u_2) + u_2 \cdot u_2$$

$$= \tfrac{16}{9}(3) - \tfrac{8}{3}(4) + 6$$

$$= \tfrac{16}{3} - \tfrac{32}{3} + 6 = 6 - \tfrac{16}{3}$$

$$= \tfrac{2}{3},$$

and

$$u_2{}^* \cdot u_3 = (-\tfrac{4}{3}u_1 + u_2) \cdot u_3 = -\tfrac{4}{3}(u_1 \cdot u_3) + u_2 \cdot u_3$$

$$= -\tfrac{4}{3}(4) + 5$$

$$= -\tfrac{1}{3}.$$

Consequently (5) may be rewritten as

$$u_3{}^* = u_3 - \tfrac{4}{3}u_1{}^* - \left[\frac{-\frac{1}{3}}{\frac{2}{3}}\right] u_2{}^*$$

$$= u_3 - \tfrac{4}{3}u_1{}^* + \tfrac{1}{2}u_2{}^*$$

$$= u_3 - \tfrac{4}{3}u_1 + \tfrac{1}{2}(-\tfrac{4}{3}u_1 + u_2)$$

$$= -2u_1 + \tfrac{1}{2}u_2 + u_3.$$

<u>Matrix Check</u>

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ -2 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 4 \\ 4 & 6 & 5 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & -\frac{4}{3} & -2 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 4 & 4 \\ 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{4}{3} & -2 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}. \qquad (6)$$

3.7.6 continued

From (6) we conclude that

$$V = [u_1^*, u_2^*, u_3^*]$$

where

$$u_1^* \cdot u_1^* = 3 \quad ; \sqrt{u_1^* \cdot u_1^*} = \sqrt{3}$$

$$u_2^* \cdot u_2^* = \frac{2}{3} \quad ; \sqrt{u_2^* \cdot u_2^*} = \sqrt{\frac{2}{3}} = \frac{1}{3}\sqrt{6}$$

$$u_3^* \cdot u_3^* = \frac{1}{2} \quad ; \sqrt{u_3^* \cdot u_3^*} = \sqrt{\frac{1}{2}} = \frac{1}{2}\sqrt{2}$$

In particular, (6) tells us that our bilinear form is positive definite since each $v = V$ has the form

$$v = x_1u_1^* + x_2u_2^* + x_3u_3^*$$

whereupon

$$v \cdot v = x_1^2(u_1^* \cdot u_1^*) + x_2^2(u_2^* \cdot u_2^*) + x_3^2(u_3^* \cdot u_3^*)$$

$$= 3x_1^2 + \frac{2}{3}x_2^2 + \frac{1}{2}x_3^2 \geq 0;$$

$$= 0 \leftrightarrow x_1 = x_2 = x_3 = 0 \leftrightarrow v = 0.$$

Finally, if we let

$$w_1 = \frac{u_1^*}{\sqrt{3}}$$

$$w_2 = \frac{u_2^*}{\frac{1}{3}\sqrt{6}} = \frac{3}{\sqrt{6}}u_2^* = \frac{1}{2}\sqrt{6}\left(-\frac{4}{3}u_1 + u_2\right)$$

$$= -\frac{2}{3}\sqrt{6}\,u_1 + \frac{1}{2}\sqrt{6}\,u_2$$

$$w_3 = \frac{u_3^*}{\frac{1}{2}\sqrt{2}} = \frac{2u_3^*}{\sqrt{2}} = \sqrt{2}\,u_3^* = \sqrt{2}\left(-2u_1 + \frac{1}{2}u_2 + u_3\right)$$

$$= -2\sqrt{2}\,u_1 + \frac{1}{2}\sqrt{2}\,u_2 + \sqrt{2}\,u_3,$$

3.7.6 continued

we may verify that $\{w_1, w_2, w_3\}$ is an orthonormal bases for V. In particular, relative to $\{w_1, w_2, w_3\}$, $(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = x_1y_1 + x_2y_2 + x_3y_3$.

Check:

$$w_1 \cdot w_1 = \frac{u_1^*}{\sqrt{3}} \cdot \frac{u_1^*}{\sqrt{3}} = \frac{u_1 \cdot u_1}{3} = \frac{3}{3} = 1$$

$$w_1 \cdot w_2 = \sqrt{6}\,(-\tfrac{2}{3}u_1 + \tfrac{1}{2}u_2) \cdot \sqrt{6}\,(-\tfrac{2}{3}u_1 + \tfrac{1}{2}u_2)$$

$$= 6[\tfrac{4}{9}u_1 \cdot u_1 - \tfrac{2}{3}u_1 \cdot u_2 + \tfrac{1}{4}u_2 \cdot u_2]$$

$$= 6[\tfrac{4}{9}(3) - \tfrac{2}{3}(4) + \tfrac{1}{4}(6)]$$

$$= 6[\tfrac{4}{3} - \tfrac{8}{3} + \tfrac{3}{2}] = 6[\tfrac{3}{2} - \tfrac{4}{3}] = 6[\tfrac{9-8}{6}] = 1;$$

$$w_3 \cdot w_3 = \tfrac{1}{2}\sqrt{2}\,(-4u_1 + u_2 + 2u_3) \cdot \tfrac{1}{2}\sqrt{2}\,(-4u_1 + u_2 + 2u_3)$$

$$= \tfrac{1}{2}[16u_1 \cdot u_1 + u_2 \cdot u_2 + 4u_3 \cdot u_3 - 8u_1 \cdot u_2 - 16u_1 \cdot u_3 + 4u_2 \cdot u_3]$$

$$= \tfrac{1}{2}[48 + 6 + 24 - 32 - 64 + 20]$$

$$= 1.$$

Note:

This is the crucial exercise if we are to identify our definition of a dot product with that given in the text (for that matter, with that given in most texts). Notice that in the text, one identifies vectors with n-tuples, whereupon the dot product is defined by

$$(x_1,\ldots,x_n) \cdot (y_1,\ldots,y_n) = x_1y_1 + \ldots + x_ny_n. \qquad (7)$$

The key point is that there are as many different ways of representing an n-dimensional space V in terms of n-tuples as there are ways of choosing a basis for V. What this exercise shows is that if our symmetric bilinear function is positive definite (i.e., a dot product), then we can always construct an orthonormal basis for V. If $u_1,\ldots,u_n$ denotes this orthonormal basis,

3.7.6 continued

then relative to this basis, our dot product is as given by (7). In other words;

$(x_1u_1 + \dots + x_nu_n) \cdot (y_1u_1 + \dots + y_nu_n)$

$= x_1y_1 + \dots + x_ny_n,$

since $u_i \cdot u_i = 1$ and $u_i \cdot u_j = 0$ if $i \neq j$.

Notice, however, that if all we know is that $u_1, \dots, u_n$ is an arbitrary basis for V, then we are not allowed to conclude that relative to this basis $(x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = x_1y_1 + \dots + x_ny_n$ since we do not know that $u_i \cdot u_i = 1$ or that $u_i \cdot u_j$ for $i \neq j$.

The key point is that once we know that we have a positive definite, symmetric, bilinear function then relative to this there is an orthonormal basis (which may be found by the Gram-Schmidt ortogonalization process). If we now agree to use this basis, thenwe may think of the dot product in terms of the usual n-tuple definition.

3.7.7

By the usual use of the cross product we know that $u_1 \times u_2$ is perpendicular to the plane determined by u_1 and u_2. We also know that

$$u_1 \times u_2 = \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$= \vec{i}(0) - \vec{j}(-1) + \vec{k}(-1)$$

$$= \vec{j} - \vec{k}. \qquad (1)$$

On the other hand,

$$-u_1 - \frac{1}{2}u_2 + u_3 = -(\vec{i} + \vec{j} + \vec{k}) - \frac{1}{2}(2\vec{i} + \vec{j} + \vec{k}) + (2\vec{i} + \vec{j} + 2\vec{k})$$

$$= -\vec{i} - \vec{j} - \vec{k} - \vec{i} - \frac{1}{2}\vec{j} - \frac{1}{2}\vec{k} + 2\vec{i} + \vec{j} + 2\vec{k}$$

3.7.7 continued

$= -\frac{1}{2}\vec{j} + \frac{1}{2}\vec{k}$

$= -\frac{1}{2}(\vec{j} - \vec{k}).$ (2)

Comparing (1) and (2) we have that $-u_1 - \frac{1}{2}u_2 + u_3$ is parallel to (i.e., is a scalar multiple of) $u_1 \times u_2$. Hence $-u_1 - \frac{1}{2}u_2 + u_3$ is perpendicular to the plane determined by u_1 and u_2. In fact, if we want to make the geometric interpretation complete, our construction of u_2^* in the lecture was equivalent to taking the component of u_2 which was perpendicular to u_1. Clearly u_1 and u_2^* determined the same plane as did u_1 and u_2.

It is worth noting that the usual geometric procedure seems much easier to handle than the more abstract Gram-Schmidt Orthogonalization Process, but the fact remains that the latter method applies to all vector spaces in which an inner product is defined, while the more intuitive geometric technique applies only to 2- and 3-dimensional space.

3.7.8

You may recall in our first Unit in this Block, we emphasized the fact that there were infinite dimensional vector spaces and that one such space was the set of continuous functions on an interval [a,b]. In this exercise we want to point out that the idea of a dot product and the associated concept of "distance" make sense in this rather abstract model of a vector space. In fact, the study of orthogonal functions (of which the special case of Fourier Series will be discussed in the next and final unit) makes excellent sense in terms of the generalized notion of "distance".

More specifically, we know that since f(x) is continuous on [a,b], $\int_a^b f(x)\,dx$ exists; and we also know that the product of two integrable functions is an integrable function. Consequently, since both f and g are continuous on [a,b], $\int_a^b f(x)g(x)\,dx$ is a well-defined number. Thus, the definition

$$f \cdot g = \int_a^b f(x)g(x)\,dx \qquad (1)$$

3.7.8 continued

is at least a function which maps ordered pairs of continuous functions, defined on [a,b], into the real numbers; and accordingly the definition given above is at least eligible to be considered as a possible dot product. Our aim in this exercise is to show that indeed (1) does define a dot product; and we do this by showing that the definition given in (1) satisfies properties (i), (ii), (iii), and (iv) required of any dot product.

Given that

$$f \cdot g = \int_a^b f(x)g(x)dx$$

then

$$\text{(i)}\ f \cdot (g + h) = \int_a^b f(x)[g(x) + h(x)]dx$$

$$= \int_a^b [f(x)g(x) + f(x)h(x)]dx$$

$$= \int_a^b f(x)g(x)dx + \int_a^b f(x)h(x)dx$$

$$= f \cdot g + f \cdot h.$$

$$\text{(ii)}\ (cf) \cdot g = \int_a^b [cf(x)]g(x)dx$$

$$= c\int_a^b f(x)g(x)dx$$

$$= c(f \cdot g).$$

$$\text{(iii)}\ f \cdot g = \int_a^b f(x)g(x)dx$$

$$= \int_a^b g(x)f(x)dx$$

$$= g \cdot f.$$

3.7.8 continued

$$\text{(iv)}\quad f \cdot f = \int_a^b f(x)f(x)\,dx = \int_a^b [f(x)]^2 dx,$$

and since $[f(x)]^2 \geq 0$ for all x, $\int_a^b [f(x)]^2 dx \geq 0$ for all x and $\int_a^b [f(x)]^2 dx = 0 \leftrightarrow f(x) \equiv 0^* \leftrightarrow f = 0$; $f \cdot f \geq 0$ and $f \cdot f = 0 \leftrightarrow f = 0$. Hence, since (i), (ii), (iii) and (iv) are all obeyed $f \cdot g$, defined by $f \cdot g = \int_a^b f(x)g(x)dx$ is a positive definite, symmetric, bilinear function. In other words, it is an inner product.

Note:

Once we have a dot product defined on a vector space, it makes sense to talk about orthogonality. In particular, suppose that $f_1, \ldots, f_n$ are any linearly independent functions which are continuous on [a,b] and that for $i \neq j$, $f_i \cdot f_j = 0$. That is, for $i \neq j$,

$$\int_a^b f_i(x)f_j(x)\,dx = 0. \qquad (2)$$

We then have a simple way of computing the coefficients $c_1, \ldots,$ and c_n in an expression such as

$$g(x) = c_1 f_1(x) + \ldots + c_n f_n(x). \qquad (3)$$

Indeed we need only employ the usual technique of dotting both sides of (3) with $f_k(x)$ (for any $k = 1,2,3,\ldots,n$) in order to find each c_k.

In particular, this leads to

$$g \cdot f_k = c_1 f_1 \cdot f_k + \ldots + c_k f_k \cdot f_k + \ldots + c_n f_n \cdot f_k. \qquad (4)$$

Each term on the right side of (4) is zero except for $c_k f_k \cdot f_k$, since $f_i \cdot f_j = 0$ whenever $i \neq j$.

If we rewrite (4) in terms of definition (1), we obtain

*Here we use the fact that f is continuous for otherwise we could allow f(x) to be unequal to 0 at, for example, a finite number of points in [a,b] and yet $\int_a^b [f(x)]^2 dx = 0$.

3.7.8 continued

$$\int_a^b g(x) f_k(x)\,dx = \int_a^b c_k [f_k(x)]^2 dx \qquad (5)$$

from which we conclude that

$$c_k = \frac{\int_a^b g(x) f_k(x)\,dx}{\int_a^b [f_k(x)]^2 dx}$$

[Notice that $f_k \neq 0$ since $\{f_1,\ldots,f_n\}$ is linearly independent, hence by (iv), $\int_a^b [f_k(x)]^2 dx \neq 0$ and, consequently (6) is well-defined.]

If also we assume that the f's are normalized, i.e., $f \cdot f = \int_a^b [f(x)]^2 dx = 1$*, then (6) may be further simplified to read

$$c_k = \int_a^b g(x) f_k(x)\,dx. \qquad (7)$$

Using the result of (7) in (3) we obtain

$$g(x) = \left[\int_a^b g(x) f_1(x)\,dx\right] f_1(x) + \ldots + \left[\int_a^b g(x) f_n(x)\,dx\right] f_n(x).$$

As is usually the case, very few serious problems arise as long as we look at finite linear combinations. The deep study begins when we assume that we are dealing with an infinite sets of functions defined and continuous on an interval [a,b]. For now, just as in our study of power series in Part 1 of this course, we must ask the question of whether every continuous function defined on [a,b] can be represented as a convergent series in

*If $f \cdot f \neq 1$, we are in no great trouble. Namely, if $f \cdot f = k \neq 1$ $(k \neq 0)$, then since $f \cdot f \geq 0$, $k \geq 0$, $\sqrt{k}$ is real. We then replace f by $f/\sqrt{k}$ whereupon

$$\frac{f}{\sqrt{k}} \cdot \frac{f}{\sqrt{k}} = \frac{f \cdot f}{k} = \frac{k}{k} = 1.$$

$1/\sqrt{k}$ is then called a weighting factor for f.

3.7.8 continued

which the series is an infinite linear combination of elements of our set S. Once we have answered this question (hopefully, in the affirmative) we must then ask whether the process of dotting term by term as we did in this exercise is valid. In other words, as usual, such statements as "the integral of a sum is the sum of the integrals" depends on the fact that our sum involves only a finite number of terms.

The discussion hinted at here forms the foundation of such topics as representations by Fourier series. We shall not go into the theory of Fourier series in this course, but we shall, in the next unit, talk about a few properties of Fourier series.

3.7.9 (optional)

a. Suppose

$$c_1u_1 + \dots + c_nu_n = 0 , \tag{1}$$

then

$$u_1 \cdot (c_1u_1 + \dots + c_nu_n) = u_1 \cdot 0. \tag{2}$$

Since $\{u_1,\dots,u_n\}$ are orthogonal we have that $u_1 \cdot u_j = 0$ for $j = 2,\dots,n$. This, coupled with the fact that $u_1 \cdot 0 = 0$, reduces (2) to

$$c_1(u_1 \cdot u_1) = 0. \tag{3}$$

Now, since c_1 and $u_1 \cdot u_1$ are real numbers, we conclude from (3) that either $c_1 = 0$ or $u_1 \cdot u_1 = 0$.

But $u_1 \cdot u_1 = 0 \leftrightarrow u_1 = 0$ (i.e., a dot product is positive definite). Hence, if $u_1 \neq 0$, then $c_1 = 0$. We may repeat this procedure by dotting both sides of (1) with $u_2, u_3,\dots,$ or u_n; and we conclude that if $\{u_1,\dots,u_n\}$ is any set of non-zero orthogonal vectors, then $c_1u_1 + \dots + c_nu_n = 0 \leftrightarrow c_1 = \dots = c_n = 0$.

In other words, any set of non-zero orthogonal vectors is linearly independent. This does not mean that one needs the concept of a dot product to study linear independence but rather

3.7.9 continued

that once one knows about a dot product, there are easier ways of testing for linear independence.

Along these same lines, notice how we used this idea when we dealt with $\vec{i}$, $\vec{j}$, and $\vec{k}$ components in E^3. Quite generally, if

$$v = c_1u_1 + \ldots + c_nu_n \tag{4}$$

where $u_1, \ldots,$ and u_n are orthogonal, we may compute c_1 by dotting both sides of (4) with u_1. Since

$$u_1 \cdot u_2 = \ldots = u_1 \cdot u_n = 0$$

we deduce from (4) that

$$u_1 \cdot v = c_1(u_1 \cdot u_1). \tag{5}$$

If we now assume, in addition, that $\{u_1, \ldots, u_n\}$ is orthonormal, then $u_1 \cdot u_1 = 1$, whereupon (5) yields

$$c_1 = u_1 \cdot v. \tag{6}$$

What equation (6) corroborates is our usual property of 3-dimensional orthonormal coordinate systems. Namely, if $\{u_1, \ldots, u_n\}$ is orthonormal and if v is any linear combination of $u_1, \ldots,$ and u_n; then we find the u_i-component of v merely by dotting v with u_i.

b. In this part of the exercise we are trying to show how the concept of an orthogonal complement extends the idea of perpendicularity as studied in the lower dimensional cases. For example, in 2-space we know that the 1-dimensional subspaces are lines. Any two non-parallel lines span the plane, but we can always pick our lines to be perpendicular. In a similar way, given a plane W in 3-space then any vector in 3-space is the sum of two vectors, one of which lies in W and the other of which doesn't. We can always choose the second vector to be perpendicular to W.

Our point is that this idea extends to any vector space on which a dot product is defined. The basic idea is as follows:

3.7.9 continued

(i) Suppose that W is any subspace of V. We then define a new set W_p by

$$W_p = \{v \varepsilon V : v \cdot w = 0 \text{ for \underline{each} } w \varepsilon W\} \qquad (7)$$

Clearly, by its very definition, W_p is a subset of V, but as we ask you to show in this part of the exercise, W_p is also a subspace of V. To show this, all we have to do is prove that the sum of two members of W_p is also a member of W_p, and that every scalar multiple of an element of W_p belongs to W_p.

Computationally, this is certainly easy enough to do. Namely:

$$\begin{aligned} v_1, v_2 \varepsilon W_p &\rightarrow v_1 \cdot w = v_2 \cdot w = 0 \text{ for each } w \varepsilon W \\ &\rightarrow v_1 \cdot w + v_2 \cdot w = 0 \text{ for each } w \varepsilon W \\ &\rightarrow (v_1 + v_2) \cdot w = 0 \quad \text{for each } w \varepsilon W \\ &\rightarrow v_1 + v_2 \varepsilon W_p \end{aligned}$$

$$\begin{aligned} v \varepsilon W_p &\rightarrow v \cdot w = 0 \text{ for each } w \varepsilon W \\ &\rightarrow c(v \cdot w) = 0 \text{ for each } w \varepsilon W\text{, where c is any scalar} \\ &\rightarrow (cv) \cdot w = 0 \text{ for each } w \varepsilon W \\ &\rightarrow cv \varepsilon W_p. \end{aligned}$$

(ii) If $v \varepsilon W_p$ then $v \cdot w = 0$ for each $w \varepsilon W$. In particular, then, if v also belongs to W, then by virtue of it being a member of W_p, $v \cdot v = 0$. Since our dot product is, by definition, positive definite, $v \cdot v = 0 \rightarrow v = 0$. Hence, $v \varepsilon W \cap W_p \rightarrow v = 0$, or $W \cap W_p = \{0\}$

(iii) We now show that $V = W \oplus W_p$. That is, each $v \varepsilon V$ may be expressed in one and only one way as a sum of two vectors, one of which is in W and the other of which is perpendicular to W (i.e., in W_p).

3.7.9 continued

Since $W \cap W_p = \{0\}$, it is sufficient to show that each $v \varepsilon V$ may be written in at least one way in the required form.*

We introduce the Gram Schmidt process to solve our problem. Namely, we begin by constructing an orthonormal basis for W, say $\{w_1, \ldots, w_r\}$. We may now use the Gram-Schmidt Orthogonalization Process to augment this basis to become a basis for V. That is

$$V = [w_1, \ldots, w_r, v_1, \ldots, v_{n-r}], \tag{8}$$

where dim V = n.

The key lies in the fact that $W_p = [v_1, \ldots, v_{n-r}]$ where $v_1, \ldots,$ and v_{n-r} are as given by (8). To prove this assertion we first notice that it is trivial to show that $v_1, \ldots,$ and v_{n-r} all belong to W_p. Namely each of these vectors is by construction orthogonal to each of the vectors $w_1, \ldots,$ and w_r; hence to the space spanned by $w_1, \ldots,$ and w_r. As a computational review, we have that for each $w \varepsilon W$,

$$w = c_1 w_1 + \ldots + c_r w_r \rightarrow$$

$$v_i \cdot w = v_i \cdot (c_1 w_1 + \ldots + c_r w_r)$$

$$= [v_i \cdot (c_1 w_1) + \ldots + v_i \cdot (c_r w_r)]$$

$$= c_1 \underbrace{(v_i \cdot w_1)}_{0} + \ldots + c_r \underbrace{(v_i \cdot w_r)}_{0}$$

$$= 0.$$

Hence, $v_i \varepsilon W_p$.

*Recall that if $W_1 \cap W_2 = \{0\}$ and $w_1 + w_2 = w_1' + w_2'$, where $w_1, w_1' \varepsilon W_1$ and $w_2, w_2' \varepsilon W_2$, then $w_1 = w_1'$ and $w_2 = w_2'$. Namely $w_1 + w_2 = w_1' + w_2' \rightarrow w_1 - w_1' = w_2' - w_2$. Hence, $w_1 - w_1' \varepsilon W_2$ since $w_2' - w_2$ is in W_2. Since $w_1 - w_1' W_1$ we have that $w_1 - w_1' = W_1 \cap W_2$; and since $W_1 \, W_2 = 0$, $w_1 - w_1' = 0$ from which we conclude that $w_1 = w_1'$. A similar argument shows that $w_2 = w_2'$.

3.7.9 continued

Conversely if w_p is any element of W_p then since $W_p \subset V$, we see from (8) that

$$w_p = c_1w_1 + \dots + c_rw_r + a_1v_1 + \dots + a_{n-r}v_{n-r}. \qquad (9)$$

Now, since $w_p \cdot w = 0$ for each $w \epsilon W$, we know that in particular $w_p \cdot w_1 = \dots = w_p \cdot w_r = 0$. Thus, for example, if we dot both sides of (9) with w_1 we obtain

$$w_p \cdot w_1 = (c_1w_1 + \dots + c_rw_r + a_1v_1 + \dots + a_{n-r}v_{n-r}) \cdot w_1$$

or

$$0 = c_1w_1 \cdot w_1 + c_2w_2 \cdot w_1 + \dots + c_rw_r \cdot w_1 + a_1v_1 \cdot w_1 + \dots + a_{n-r}v_{n-r} \cdot w_1 \qquad (10)$$

and since $\{w_1,\dots,w_r, v_1,\dots, v_{n-r}\}$ is orthonormal, we conclude from (10) that $0 = c_1$. Similarly, we may show that $c_2 = \dots = c_r = 0$, so that by (9),

$$w_p = a_1v_1 + \dots + a_{n-r}v_{n-r},$$

and this shows that $\{v_1,\dots,v_{n-r}\}$ span W_p. This completes our proof that

$$V = W \oplus W_p.$$

3.7.10 (optional)

We have

$$\left.\begin{array}{l} u_1 \cdot u_1 = 4,\ u_1 \cdot u_2 = 1,\ u_1 \cdot u_3 = 2,\ u_1 \cdot u_4 = 1 \\ u_2 \cdot u_1 = 1,\ u_2 \cdot u_2 = 7,\ u_2 \cdot u_3 = 3,\ u_2 \cdot u_4 = 2 \\ u_3 \cdot u_1 = 2,\ u_3 \cdot u_2 = 3,\ u_3 \cdot u_3 = 2,\ u_3 \cdot u_4 = 1 \\ u_4 \cdot u_1 = 1,\ u_4 \cdot u_2 = 2,\ u_4 \cdot u_3 = 1,\ u_4 \cdot u_4 = 9 \end{array}\right\} \qquad (1)$$

We first let

$$\underline{u_1{}^* = u_1.} \qquad (2)$$

3.7.10 continued

Then by the Gram-Schmidt process we have

$$u_2^* = u_2 - \frac{u_2 \cdot u_1^*}{u_1^* \cdot u_1^*} u_1^*$$

$$= u_2 - \frac{u_2 \cdot u_1}{u_1 \cdot u_1} u_1,$$

or from (1),

$$\underline{u_2^* = u_2 - \frac{1}{4} u_1.} \qquad (3)$$

[Check: $u_1^* \cdot u_2^* = u_1 \cdot (u_2 - \frac{1}{4} u_1) = u_1 \cdot u_2 - \frac{1}{4} u_1 \cdot u_1$
$= 1 - \frac{1}{4}(4) = 1 - 1 = 0.$]

With u_1^* as in (2) and u_2^* as in (3), we next define u_3^* by

$$u_3^* = u_3 - \frac{u_3 \cdot u_1^*}{u_1^* \cdot u_1^*} u_1^* - \frac{u_3 \cdot u_2^*}{u_2^* \cdot u_2^*} u_2^*. \qquad (4)$$

Now,

$$u_1^* \cdot u_1^* = u_1 \cdot u_1 = 4 \qquad (5)$$

$$u_3 \cdot u_1^* = u_3 \cdot u_1 = 2 \qquad (6)$$

$$u_3 \cdot u_2^* = u_3 \cdot (u_2 - \frac{1}{4} u_1) = u_3 \cdot u_2 - \frac{1}{4} u_3 \cdot u_1$$

$$= 3 - \frac{1}{4}(2)$$

$$= \frac{5}{2}. \qquad (7)$$

$$u_2^* \cdot u_2^* = (u_2 - \frac{1}{4} u_1) \cdot (u_2 - \frac{1}{4} u_1)$$

$$= u_2 \cdot u_2 - \frac{1}{2} u_1 \cdot u_2 + \frac{1}{16} u_1 \cdot u_1$$

$$= 7 - \frac{1}{2} + \frac{1}{16}(4)$$

$$= 7 - \frac{1}{2} + \frac{1}{4}$$

$$= \frac{27}{4}. \qquad (8)$$

3.7.10 continued

Putting the results of (5), (6), (7) and (8) into (4) yields

$$u_3^* = u_3 - \frac{2}{4}u_1 - \frac{\frac{5}{2}}{\frac{27}{4}}\left(u_2 - \frac{1}{4}u_1\right);$$

or

$$u_3^* = u_3 - \frac{1}{2}u_1 - \frac{10}{27}\left(u_2 - \frac{1}{4}u_1\right)$$

$$= u_3 - \frac{11}{27}u_1 - \frac{10}{27}u_2. \qquad (9)$$

Finally,

$$u_4^* = u_4 - \frac{u_4 \cdot u_1^*}{u_1^* \cdot u_1^*}u_1^* - \frac{u_4 \cdot u_2^*}{u_2^* \cdot u_2^*}u_2^* - \frac{u_4 \cdot u_3^*}{u_3^* \cdot u_3^*}u_3^* . \qquad (10)$$

Now,

$$u_1^* \cdot u_1^* = u_1 \cdot u_1 = 4$$

$$u_4 \cdot u_1^* = u_4 \cdot u_1 = 1$$

$$u_4 \cdot u_2^* = u_4 \cdot \left(u_2 - \frac{1}{4}u_1\right) = u_4 \cdot u_2 - \frac{1}{4}u_4 \cdot u_1$$

$$= 2 - \frac{1}{4}(1) = \frac{7}{4} .$$

$$u_2^* \cdot u_2^* \text{ [by (8)]} = \frac{27}{4}$$

$$u_4 \cdot u_3^* = \text{[by (9)]}\ u_4 \cdot \left(u_3 - \frac{11}{27}u_1 - \frac{10}{27}u_2\right)$$

$$= u_4 \cdot u_3 - \frac{11}{27}u_4 \cdot u_1 - \frac{10}{27}u_4 \cdot u_2$$

$$= 1 - \frac{11}{27}(1) - \frac{10}{27}(2)$$

$$= 1 - \frac{11}{27} - \frac{20}{27}$$

$$= -\frac{4}{27} .$$

3.7.10 continued

$$u_3^* \cdot u_3^* = (u_3 - \tfrac{11}{27} u_1 - \tfrac{10}{27} u_2) \cdot (u_3 - \tfrac{11}{27} u_1 - \tfrac{10}{27} u_2)$$

$$= u_3 \cdot u_3 + \tfrac{121}{729} u_1 \cdot u_1 + \tfrac{100}{729} u_2 \cdot u_2 - \tfrac{22}{27} u_3 \cdot u_1 - \tfrac{20}{27} u_3 \cdot u_2 + \tfrac{220}{729} u_1 \cdot u_2$$

$$= 2 + \tfrac{484}{729} + \tfrac{700}{729} - \tfrac{44}{27} - \tfrac{60}{27} + \tfrac{220}{729}$$

$$= \frac{1458 + 484 + 700 - 1188 - 1620 + 220}{729}$$

$$= \frac{54}{729}$$

$$= \frac{2}{27}.$$

Putting these results into (10) yields

$$u_4^* = u_4 - \tfrac{1}{4} u_1 - \frac{\frac{7}{4}}{\frac{27}{4}} (u_2 - \tfrac{1}{4} u_1) - \frac{(-\frac{4}{27})}{\frac{2}{27}} (u_3 - \tfrac{11}{27} u_1 - \tfrac{10}{27} u_2)$$

$$= u_4 - \tfrac{1}{4} u_1 - \tfrac{7}{27} (u_2 - \tfrac{1}{4} u_1) + 2(u_3 - \tfrac{11}{27} u_1 - \tfrac{10}{27} u_2)$$

$$= u_4 + (-\tfrac{1}{4} + \tfrac{7}{108} - \tfrac{22}{27}) u_1 + (-\tfrac{7}{27} - \tfrac{20}{27}) u_2 + 2u_3$$

$$= u_4 - u_1 - u_2 + 2u_3.$$

Summarizing, we have from (2), (3), (9), and (11)

$$\left.\begin{aligned} u_1^* &= u_1 \\ u_2^* &= -\tfrac{1}{4} u_1 + u_2 \\ u_3^* &= -\tfrac{11}{27} u_1 - \tfrac{10}{27} u_2 + u_3 \\ u_4^* &= -u_1 - u_2 + 2u_3 + u_4 \end{aligned}\right\} \qquad (12)$$

3.7.10 continued

<u>Note</u>:

If we write (12) in matrix form we have

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{4} & 1 & 0 & 0 \\ -\frac{11}{27} & -\frac{10}{27} & 1 & 0 \\ -1 & -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 & 1 \\ 1 & 7 & 3 & 2 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 1 & 9 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{4} & -\frac{11}{27} & -1 \\ 0 & 1 & -\frac{10}{27} & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 1 & 2 & 1 \\ 0 & \frac{27}{4} & \frac{5}{2} & \frac{7}{4} \\ 0 & 0 & \frac{2}{27} & -\frac{4}{27} \\ 0 & 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{4} & -\frac{11}{27} & -1 \\ 0 & 1 & -\frac{10}{27} & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & \frac{27}{4} & 0 & 0 \\ 0 & 0 & \frac{2}{27} & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} . \qquad (13)$$

Matrix (13) tells us that

$u_1^* \cdot u_1^* = 4,\ u_1^* \cdot u_2^* = 0,\ u_1^* \cdot u_3^* = 0,\ u_1^* \cdot u_4^* = 0$

$u_2^* \cdot u_1^* = 0,\ u_2^* \cdot u_2^* = \frac{27}{4},\ u_2^* \cdot u_3^* = 0,\ u_2^* \cdot u_4^* = 0$

$u_3^* \cdot u_1^* = 0,\ u_3^* \cdot u_2^* = 0,\ u_3^* \cdot u_3^* = \frac{2}{27},\ u_3^* \cdot u_4^* = 0$

$u_4^* \cdot u_1^* = 0,\ u_4^* \cdot u_2^* = 0,\ u_4^* \cdot u_3^* = 0,\ u_4^* \cdot u_4^* = 8.$

3.7.10 continued

Hence, $u_1^*, u_2^*, u_3^*, u_4^*$ is an orthogonal basis for V; and

$$(x_1u_1^* + x_2u_2^* + x_3u_3^* + x_4u_4^*) \cdot (y_1u_1^* + y_2u_2^* + y_3u_3^* + y_4u_4^*)$$

$$= 4x_1y_1 + \frac{27}{4}x_2y_2 + \frac{2}{27}x_3y_3 + 8x_4y_4.$$

This is much neater than the corresponding result using u_1, u_2, u_3, and u_4 as a basis - in which case we would also have had to worry about the terms involving x_1y_2, x_1y_3, x_1y_4, x_2y_1, x_2y_3, x_2y_4, x_3y_1, x_3y_2, x_3y_4, x_4y_1, x_4y_2, and x_4y_3.