Unit 6: Eigenvectors (Characteristic Vectors)

3.6.1(L)

a. This is essentially a continuation of Unit 4. Knowing that

$$f(u_1) = -3u_1 + 2u_2$$

and

$$f(u_2) = 4u_1 - u_2$$

we have that the transpose* of the matrix of coefficients is given by

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

To find all $v = x_1u_1 + x_2u_2$ such that $f(v) = cv$, we let $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and then solve the equation

$$AX = cX \quad (= cIX)$$

for c.

This yields

$$\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix} . \qquad (1)$$

In this simple 2-dimensional case, we could solve (1) by direct computation but to prepare for the more general case, let us use the technique described in the lecture. Recall that since $AX = cX$ is automatically solved for any c if $X = 0$, we showed that if we want solutions other than $X = 0$ we must have that

*See the note at the end of this exercise for the procedure to be used if we wanted to use A to denote the matrix of coefficients.

3.6.1(L) continued

$$AX - cIX = 0$$

or

$$(A - cI)X = 0, \tag{2}$$

but if $A - cI$ is invertible, $X = 0$ is the only solution of (2). Hence, it is crucial that $A - cI$ be singular and this means that

$$|A - cI| = 0. \tag{3}$$

Since $A = \begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$ and $cI = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$, equation (3) yields

$$\begin{vmatrix} -3-c & 4 \\ 2 & -1-c \end{vmatrix} = 0,$$

or

$$(-3-c)(-1-c) - 8 = 0$$

or

$$c^2 + 4c - 5 = 0$$

or

$$(c + 5)(c - 1) = 0.$$

Hence,

$$|A - cI| = 0 \rightarrow c = -5 \text{ or } c = 1.$$

With $c = -5$, equation (1) becomes

$$\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -5x_1 \\ -5x_2 \end{bmatrix}$$

3.6.1(L) continued

or

$$\begin{bmatrix} -3x_1 + 4x_2 \\ 2x_1 - x_2 \end{bmatrix} = \begin{bmatrix} -5x_1 \\ -5x_2 \end{bmatrix}, \tag{4}$$

whereupon

$$\left.\begin{aligned} -3x_1 + 4x_2 &= -5x_1 \\ 2x_1 - x_2 &= -5x_2 \end{aligned}\right\}. \tag{5}$$

A quick check shows that (5) is equivalent to the single equation

$$2x_1 = -4x_2$$

or

$$x_1 = -2x_2.$$

In other words,

$$f(v) = f(x_1u_1 + x_2u_2) = 5(x_1u_1 + x_2u_2) \leftrightarrow$$

$$v = -2x_2u_1 + x_2u_2 = x_2(-2u_1 + u_2). \tag{6}$$

From (6), we see that $-2u_1 + u_2$ is a basis for the (1-dimensional) subspace of V defined by

$$V_1 = \{v : f(v) = -5v\}.$$

In a similar way, we may revisit (1) with c = 1 to obtain

$$\begin{bmatrix} -3x_1 + 4x_2 \\ 2x_1 - x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{7}$$

[Notice that the left sides of (4) and (7) are the same. This is no coincidence since c affects only the right side of (1).]

3.6.1(L) continued

From (7), we conclude that

$$\left.\begin{aligned} -3x_1 + 4x_2 &= x_1 \\ 2x_1 - x_2 &= x_2 \end{aligned}\right\}$$

or

$$x_1 = x_2.$$

In other words, if $V_2 = \{v\varepsilon V: f(v) = v\}$, then

$$V_2 = \{x_2u_1 + x_2u_2\} = x_2(u_1 + u_2)$$

so that V_2 is the 1-dimensional subspace of V which is spanned by $u_1 + u_2$.

To check our results so far, let $\alpha_1 = -2u_1 + u_2$ and $\alpha_2 = u_1 + u_2$. Then

$$\begin{aligned} f(\alpha_1) &= f(-2u_1 + u_2) \\ &= -2f(u_1) + f(u_2) \\ &= -2[-3u_1 + 2u_2] + [4u_1 - u_2] \\ &= 10u_1 - 5u_2 \\ &= -5(-2u_1 + u_2) \\ &= -5\alpha_1, \end{aligned}$$

while

3.6.1(L) continued

$$\begin{aligned} f(\alpha_2) &= f(u_1 + u_2) \\ &= f(u_1) + f(u_2) \\ &= [-3u_1 + 2u_2] + [4u_1 - u_2] \\ &= u_1 + u_2 \\ &= \alpha_2. \end{aligned}$$

In summary,

$$V = [\alpha_1] \oplus [\alpha_2]$$

where $f(\alpha_1) = -5\alpha_1$ and $f(\alpha_2) = \alpha_2$.

b. Relative to the basis $\{\alpha_1,\alpha_2\}$, we saw in (a) that

$$\left.\begin{aligned} f(\alpha_1) &= -5\alpha_1 = -5\alpha_1 + 0\alpha_2 \\ f(\alpha_2) &= \quad\alpha_2 = \ \ 0\alpha_1 + 1\alpha_2 \end{aligned}\right\}. \tag{8}$$

Hence, from (8) we deduce that the matrix of (8) relative to $\{\alpha_1,\alpha_2\}$ is

$$B = \begin{bmatrix} -5 & 0 \\ 0 & 1 \end{bmatrix}. \tag{9}$$

<u>Note</u>

What this means is that with respect to the linear transformation f, the basis consisting of α_1 and α_2 is more convenient than the basis consisting of u_1 and u_2. In particular, we see from (9) how to compute f(v) relative to α_1 and α_2 coordinates; that is, if $v = y_1\alpha_1 + y_2\alpha_2$, then f(v) is given by

$$\begin{bmatrix} -5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

3.6.1(L) continued

or, without reference to matrix notation,

$$f(v) = f(y_1\alpha_1 + y_2\alpha_2)$$

$$= y_1 f(\alpha_1) + y_2 f(\alpha_2)$$

$$= -5y_1\alpha_1 + y_2\alpha_2.$$

c. We have

$$f:E^2 \to E^2$$

where

$$f(\vec{i}) = -3\vec{i} + 2\vec{j}$$

$$f(\vec{j}) = 4\vec{i} - \vec{j}.$$

We then let

$$\vec{\alpha}_1 = -2\vec{i} + \vec{j}$$

and

$$\vec{\alpha}_2 = \vec{i} + \vec{j}.$$

Then,

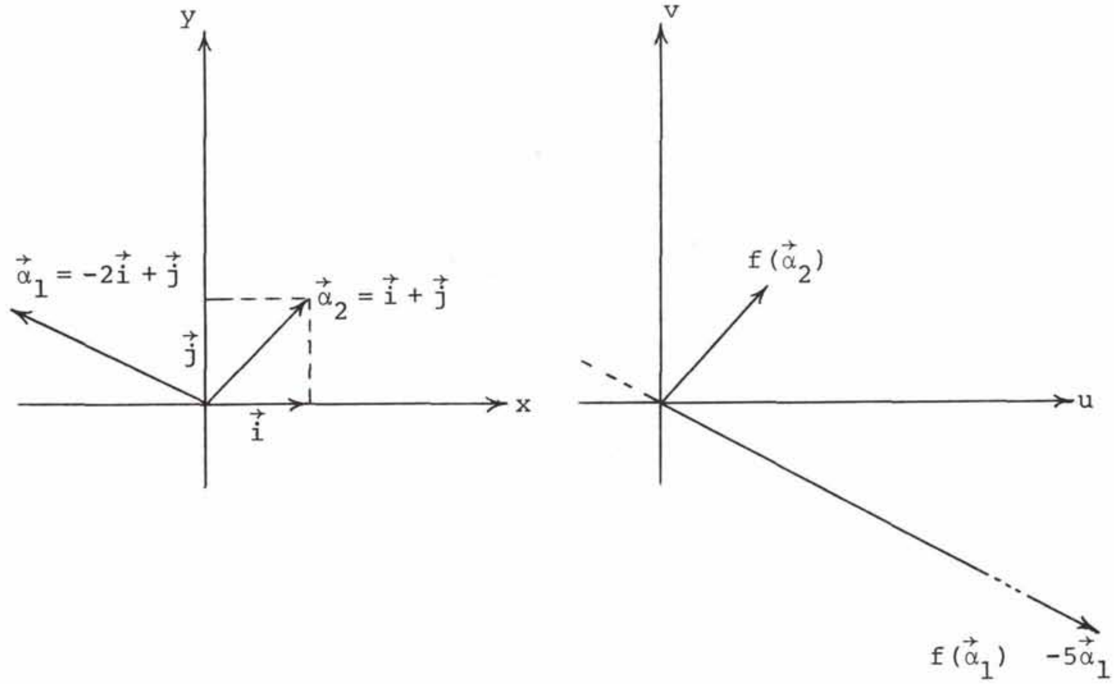
$$f(\vec{\alpha}_1) = -5\vec{\alpha}_1$$

and

$$f(\vec{\alpha}_2) = \vec{\alpha}_2.$$

Pictorially,

3.6.1(L) continued



In other words, relative to a coordinate system consisting of α_1 and α_2, f preserves the direction of the coordinate axes.

Note

Had we worked with the matrix of coefficients rather than its transpose, our matrix A (which is really A^T with A as in this exercise) would have been

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

in which case, our characteristic values would still have been $c = -5$ and $c = 1$.* The only difference now is that if we want to write, say, $f(v) = -5v$ in matrix form, where $v = x_1u_1 + x_2u_2$, we write

$\vec{X}A = -5\vec{X}$.

*Namely, $[A^T - cI]^T = (A^T)^T - c(I)^T = A - cI$. Hence, $|(A^T - cI)^T| = |A - cI|$.

3.6.1(L) continued

That is,

$$[x_1\ x_2]\begin{bmatrix}-3 & 2\\ 4 & -1\end{bmatrix} = [-5x_1\ -5x_2],$$

whereupon

$$[-3x_1 + 4x_2\quad 2x_1 - x_2] = [-5x_1\quad -5x_2],$$

or

$$\left.\begin{aligned}-3x_1 + 4x_2 &= -5x_1\\ 2x_1 - x_2 &= -5x_2\end{aligned}\right\}$$

which agrees with (5).

The point is that we use

$$A\vec{X} = c\vec{X}$$

when A is written as the transpose of the matrix of coefficients, but we write

$$\vec{X}A = c\vec{X}$$

when A denotes the matrix of coefficients itself.

It is not important which convention we use, but it is important to remember which is which. For example, if we use A as defined in this note and then solve

$$\begin{bmatrix}-3 & 2\\ 4 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}-5x_1\\ -5x_2\end{bmatrix},$$

we obtain

$$\begin{bmatrix}-3x_1 + 2x_2\\ 4x_1 - x_2\end{bmatrix} = \begin{bmatrix}-5x_1\\ -5x_2\end{bmatrix}$$

3.6.1(L) continued

or

$$\left.\begin{aligned} -3x_1 &= -7x_2 \\ 4x_1 &= -4x_2 \end{aligned}\right\}$$

which admits only $X = 0$ as a solution. In other words, once A is chosen to be the matrix of coefficients, X must be written as a row matrix which multiplies A on the left.

3.6.2

Method #1

We have $V = [u_1, u_2]$ and

$$\left.\begin{aligned} f(u_1) &= 8u_1 - 15u_2 \\ f(u_2) &= 2u_1 - 3u_2 \end{aligned}\right\}. \tag{1}$$

We let A denote the transpose of the matrix of coefficients in (1). This yields

$$A = \begin{bmatrix} 8 & 2 \\ -15 & -3 \end{bmatrix} \tag{2}$$

Then

$$|A - cI| = \begin{vmatrix} 8 - c & 2 \\ -15 & -3 - c \end{vmatrix} = (8 - c)(-3 - c) - 2(-15)$$

$$= -24 - 5c + c^2 + 30 = c^2 - 5c + 6$$

$$= (c - 2)(c - 3).$$

Hence,

3.6.2 continued

$|A - cI| = 0 \leftrightarrow c = 2$ or $c = 3$.

Using $c = 2$ and letting $\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $v = x_1u_1 + x_2u_2$, we have

$$\begin{bmatrix} 8 & 2 \\ -15 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \tag{3}$$

and this says that

$$\begin{bmatrix} 8x_1 + 2x_2 \\ -15x_1 - 3x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}.$$

Consequently,

$$\left.\begin{aligned} 8x_1 + 2x_2 &= 2x_1 \\ -15x_1 - 3x_2 &= 2x_2 \end{aligned}\right\}$$

or

$x_2 = -3x_1$.

In other words,

$f(x_1u_1 + x_2u_2) = 2(x_1u_1 + x_2u_2) \leftrightarrow x_2 = -3x_1$.

That is,

$f(v) = 2v \leftrightarrow v = x_1u_1 - 3x_1u_2 = x_1(u_1 - 3u_2)$.

Hence, if we let

$V_1 = \{v \varepsilon V : f(v) = 2v\}$,

then

3.6.2 continued

$V_1 = [u_1 - 3u_2].$ (4)

Check

Let $\alpha_1 = u_1 - 3u_2$. Then

$$\begin{aligned} f(\alpha_1) &= f(u_1) - 3f(u_2) \\ &= (8u_1 - 15u_2) - 3(2u_1 - 3u_2) \\ &= 2u_1 - 6u_2 \\ &= 2(u_1 - 3u_2) \\ &= 2\alpha_1. \end{aligned}$$

If we next let c = 3, equation (3) is replaced by

$$\begin{bmatrix} 8 & 2 \\ -15 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ 3x_2 \end{bmatrix}$$

or

$$\begin{bmatrix} 8x_1 + 2x_2 \\ -15x_1 - 3x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ 3x_2 \end{bmatrix}.$$

Hence,

$$\left.\begin{aligned} 8x_1 + 2x_2 &= 3x_1 \\ -15x_1 - 3x_2 &= 3x_2 \end{aligned}\right\}$$

or

$$2x_2 = -5x_1.$$

In particular, $f(2u_1 - 5u_2) = 3(2u_1 - 5u_2)$.

3.6.2 continued

Check

$$\begin{aligned} f(2u_1 - 5u_2) &= 2f(u_1) - 5f(u_2) \\ &= 2(8u_1 - 15u_2) - 5(2u_1 - 3u_2) \\ &= 6u_1 - 15u_2 \\ &= 3(2u_1 - 5u_2). \end{aligned}$$

Consequently, if $V_2 = \{v \varepsilon V : f(v) = 3v\}$, then

$$V_2 = [2u_1 - 5u_2]. \tag{5}$$

Looking at (4) and (5), and letting

$$\alpha_1 = u_1 - 3u_2$$

$$\alpha_2 = 2u_1 - 5u_2$$

we see that

$$V = [\alpha_1] + [\alpha_2]$$

and relative to $\{\alpha_1, \alpha_2\}$, the matrix of f is given by

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

That is,

$$f(\alpha_1) = 2\alpha_1 = 2\alpha_1 + 0\alpha_2$$

$$f(\alpha_2) = 3\alpha_2 = 0\alpha_1 + 3\alpha_2.$$

Method #2

This is the same as our first method except that we now let A denote the matrix of coefficients in (1).

3.6.2 continued

That is

$$A = \begin{bmatrix} 8 & -15 \\ 2 & -3 \end{bmatrix}$$

We again obtain that our eigenvalues are c = 2 and c = 3. The only subtlety is that we now find α_1 and α_2 by solving

$$[x_1 \quad x_2] \begin{bmatrix} 8 & -15 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix} \text{ where } c = 2 \text{ or } c = 3.$$

A quick check shows that everything works out as in Method #1.

3.6.3(L)

Our main aim here is to show still another invariant of the matrix which represents a given linear transformation. We know that if A and B are two matrices that represent the same transformation but with respect to different bases, then there exists a non-singular matrix P such that

$$B = PAP^{-1}.$$

What we already know here is that A and B are characterized by having the same determinant. That is,

$$\begin{aligned} |B| &= |PAP^{-1}| \\ &= |P|\ |A|\ |P^{-1}| \\ &= |P|\ |A|\ \frac{1}{|P|}, \quad \left(|P| \neq 0\right) \\ &= |A|. \end{aligned}$$

What we shall show in this exercise is that the characteristic equation

3.6.3(L) continued

$|A - cI| = 0$

is the same equation as

$|B - cI| = 0.$

More specifically, we shall show that

$|A - cI| = |PAP^{-1} - cI|.$

To this end, we simply begin by rewriting I as $P^{-1}IP$, and this leads us to the sequence of steps

$$\begin{aligned} P^{-1}BP - cI &= P^{-1}BP - cP^{-1}IP \\ &= P^{-1}(BP - cIP) \\ &= P^{-1}[(B - cI)P], \qquad (1) \end{aligned}$$

and since the determinant of a product is the product of the determinants, we conclude from (1) that

$$\begin{aligned} |P^{-1}BP - cI| &= |P^{-1}(B - cI)P| \\ &= |P^{-1}|\ |B - cI|\ |P| \\ &= |B - cI|, \end{aligned}$$

or, since $A = P^{-1}BP$,

$|A - cI| = |B - cI|.$

In summary, then, what we have shown is that if we want to find the characteristic equation for a linear transformation and if A is any matrix which represents the given transformation, then the characteristic equation is given by

$|A - cI| = 0$

3.6.3(L) continued

regardless of the choice of A.

In particular, if the matrix of the transformation is chosen to be of triangular form (and this is always possible), we may write down the characteristic equation very easily simply by equating the product of the diagonal elements to 0.

For example, relative to the previous exercise, we found that the matrix

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

represented the same transformation as did the matrix

$$\begin{bmatrix} 8 & 2 \\ -15 & 3 \end{bmatrix}.$$

Notice, however, that our first matrix yields the characteristic equation $(2 - c)(3 - c) = 0$ almost instantly.

3.6.4(L)

From a mechanical point of view, this exercise is not too difficult. It is designed to show that if f is a linear transformation of the n-dimensional vector space V onto itself, there need not be n linearly independent vectors which are mapped into scalar multiples of themselves.

In order that you see this from a more intuitive point of view, we have chosen a problem that has a simple geometric interpretation. Namely, we have chosen a rotation of the plane. In particular, if

$$\left.\begin{aligned} u &= x \cos \alpha - y \sin \alpha \\ v &= x \sin \alpha + y \cos \alpha \end{aligned}\right\} \qquad (1)$$

3.6.4(L) continued

then the uv-plane is just the xy-plane rotated about the origin through an angle of measure α.* It should be clear that unless α is an integral multiple of 180°, no line through the origin has its direction preserved. For example, if we rotate the xy-plane through 45°, every line which passes through the origin is rotated by 45°, and, consequently, cannot have the same direction it had prior to the rotation.

Note

Equation (1) may be seen to be equivalent to the given transformation by letting $x = 1$ and $y = 0$, and then letting $x = 0$ and $y = 1$. Namely, we see from (1) that the mapping defined by

$$f(x,y) = (u,v) = (x \cos \alpha - y \sin \alpha,\ x \sin \alpha + y \cos \alpha)$$

implies that

$$\left.\begin{aligned} f(1,0) &= (\cos \alpha,\ \sin \alpha) \\ f(0,1) &= (-\sin \alpha,\ \cos \alpha) \end{aligned}\right\} . \qquad (2)$$

Identifying (a,b) with $au_1 + bu_2$, equation (2) becomes

$$\left.\begin{aligned} f(u_1) &= (\cos \alpha)u_1 + (\sin \alpha)u_2 \\ f(u_2) &= (-\sin \alpha)u_1 + (\cos \alpha)u_2 \end{aligned}\right\} \qquad (3)$$

which is the given transformation.

The matrix of coefficients in (3) is given by

---

*By way of review, here is a good application of the arithmetic of complex numbers. $c = \cos \alpha + i \sin \alpha$ has magnitude 1 and argument α. Hence, cz rotates z through an angle α. In other words,

$$w = cz \rightarrow$$

$$u + iv = (\cos \alpha + i \sin \alpha)(x + iy)$$

$$= x \cos \alpha - y \sin \alpha + i(x \sin \alpha + y \cos \alpha).$$

Hence, $u = x \cos \alpha - y \sin \alpha$ and $v = x \sin \alpha + y \cos \alpha$.

3.6.4(L) continued

$$A = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

so that

$$|A - cI| = \begin{vmatrix} \cos\alpha - c & \sin\alpha \\ -\sin\alpha & \cos\alpha - c \end{vmatrix} = (\cos\alpha - c)^2 + \sin^2\alpha$$

or

$$|A - cI| = \cos^2\alpha - 2c\cos\alpha + c^2 + \sin^2\alpha$$

$$= c^2 - 2c\cos\alpha + 1.$$

Hence,

$$|A - cI| = 0 \leftrightarrow c^2 - 2c\cos\alpha + 1 = 0$$

$$\leftrightarrow c = \frac{2\cos\alpha \pm \sqrt{4\cos^2\alpha - 4}}{2}$$

$$\leftrightarrow c = \cos\alpha \pm \sqrt{\cos^2\alpha - 1}.$$

But, $\cos^2\alpha - 1 < 0$ unless $\cos^2\alpha = 1$ or $\cos\alpha = \pm 1$. That is, $\sqrt{\cos^2\alpha - 1}$ is real $\leftrightarrow \alpha = k\pi$ (radians) where k is any integer.

3.6.5

a. The matrix of coefficients is

$$A = \begin{bmatrix} 8 & 0 & 4 & 0 \\ 9 & 2 & 6 & 0 \\ -9 & 0 & -4 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix}$$

Hence, the characteristic equation of our transformation is

3.6.5 continued

$$\begin{vmatrix} 8-c & 0 & 4 & 0 \\ 9 & 2-c & 6 & 0 \\ -9 & 0 & -4-c & 0 \\ 0 & 2 & 0 & 3-c \end{vmatrix} = 0. \tag{1}$$

Since the determinant in (1) has its last column with all but one zero entry, we may expand it along the last column to obtain from (1) that

$$(3-c)\begin{vmatrix} 8-c & 0 & 4 \\ 9 & 2-c & 6 \\ -9 & 0 & -4-c \end{vmatrix} = 0. \tag{2}$$

The determinant in (2) has two zero entries in its second column so if we expand it along the second column, we obtain from (2) that

$$(3-c)(2-c)\begin{vmatrix} 8-c & 4 \\ -9 & -4-c \end{vmatrix} = 0$$

or

$$(3-c)(2-c)[(8-c)(-4-c)+36] = 0,$$

or

$$(3-c)(2-c)(-32-4c+c^2+36) = 0,$$

or

$$(3-c)(2-c)(c^2-4c+4) = 0.$$

Hence,

$$(3-c)(2-c)(c-2)^2 = 0,$$

3.6.5 continued

or, since $(c - 2)^2 = (2 - c)^2$,

$$(3 - c)(2 - c)^3 = 0. \tag{3}$$

Thus, we conclude from (3) that the characteristic values of our transformation are given by $c = 2$ and $c = 3$.

b. Letting $v = x_1u_1 + x_2u_2 + x_3u_3 + x_4u_4$, the required matrix equation to solve is

$$\vec{X}A = 2\vec{X}, \text{ where } \vec{X} = [x_1\ x_2\ x_3\ x_4].$$

[Recall that we must write $\vec{X}A$ rather than $A\overset{\downarrow}{X}$ since we have chosen A to be the matrix of coefficients rather than its transpose.]

We obtain

$$[x_1\ x_2\ x_3\ x_4]\begin{bmatrix} 8 & 0 & 4 & 0 \\ 9 & 2 & 6 & 0 \\ -9 & 0 & -4 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix} = [2x_1\ 2x_2\ 2x_3\ 2x_4]$$

or

$$\left.\begin{array}{ll} \text{(i)} & 8x_1 + 9x_2 - 9x_3 = 2x_1 \\ \text{(ii)} & 2x_2 + 2x_4 = 2x_2 \\ \text{(iii)} & 4x_1 + 6x_2 - 4x_3 = 2x_3 \\ \text{(iv)} & 3x_4 = 2x_4 \end{array}\right\} . \tag{4}$$

From (iv), we see that $x_4 = 0$ and letting $x_4 = 0$ in (ii) tells us that x_2 is an arbitrary constant (i.e. $x_2 = x_2$ is true for any constant). If we now pick x_1 at random, we see from (i) that

$$9x_3 = 6x_1 + 9x_2$$

or

3.6.5 continued

$x_3 = \frac{2}{3} x_1 + x_2.$

This checks with (iii) which says that

$4x_1 + 6x_2 = 6x_3,$

or

$x_3 = \frac{2}{3} x_1 + x_2.$

We have, therefore, shown that if

$v = (x_1,x_2,x_3,x_4) = x_1u_1 + x_2u_2 + x_3u_3 + x_4u_4$

then

$f(v) = 2v \leftrightarrow$

$$v = (x_1,x_2,\tfrac{2}{3} x_1 + x_2,0). \qquad (5)$$

To pick a convenient basis for

$V_2 = \{v \varepsilon V: f(v) = 2v\},$

we may let $x_1 = 0$ and $x_2 = 1$ in (5) to obtain

$\alpha_1 = (0,1,1,0) = u_2 + u_3$

and we may let $x_1 = 3$* and $x_2 = 0$ in (5) to obtain

$\alpha_2 = (3,0,2,0) = 3u_1 + 2u_3.$

Hence, dim $V_2 = 2$, and

$V_2 = [\alpha_1] \oplus [\alpha_2].$

*We picked $x_1 = 3$ rather than 1 simply to avoid having $\frac{2}{3} x_1$ be a non-integer.

3.6.5 continued

<u>Check</u>

$$f(\alpha_1) = f(0,1,1,0) = f(u_2 + u_3)$$

$$= f(u_2) + f(u_3)$$

$$= (9u_1 + 2u_2 + 6u_3) + (-9u_1 - 4u_3)$$

$$= 2u_2 + 2u_3$$

$$= 2(u_2 + u_3)$$

$$= 2\alpha_1$$

$$f(\alpha_2) = f(3u_1 + 2u_3)$$

$$= 3f(u_1) + 2f(u_3)$$

$$= 3(8u_1 + 4u_3) + 2(-9u_1 - 4u_3)$$

$$= 6u_1 + 4u_3$$

$$= 2(3u_1 + 2u_3)$$

$$= 2\alpha_2$$

3.6.6(L)

a. Suppose $v_1 \neq 0$ and that

$$f(v_1) = c_1v_1. \tag{1}$$

Suppose also that $v_2 \neq 0$ and that

$$f(v_2) = c_2v_2 \tag{2}$$

where

$$c_1 \neq c_2. \tag{3}$$

3.6.6(L) continued

Now if there exist scalars x_1 and x_2 such that

$$x_1v_1 + x_2v_2 = 0 \tag{4}$$

then

$$f(x_1v_1 + x_2v_2) = 0,$$

or

$$x_1f(v_1) + x_2f(v_2) = 0.$$

Hence,

$$x_1c_1v_1 + x_2c_2v_2 = 0. \tag{5}$$

If we multiply both sides of (4) by c_1 and subtract this result from (5), we obtain

$$x_2c_2v_2 - x_2c_1v_2 = 0$$

or

$$x_2(c_2 - c_1)v_2 = 0. \tag{6}$$

Since $v_2 \neq 0$, we see from (6) that

$$x_2(c_2 - c_1) = 0, \tag{7}$$

and since $c_2 - c_1 \neq 0$, we see from (7) that

$$x_2 = 0.$$

Knowing that $x_2 = 0$, we return to (4) to conclude that

$$x_1v_1 = 0, \tag{8}$$

and since $v_1 \neq 0$, we see from (8) that

3.6.6(L) continued

$x_1 = 0.$

Since x_1 and x_2 must both equal 0, $\{v_1, v_2\}$ is a linearly independent set.

b. We continue inductively along the lines of (a). We have that $v_1 \neq 0$, $v_2 \neq 0$, $v_3 \neq 0$ and that

$$f(v_1) = c_1v_1,\ f(v_2) = c_2v_2,\ f(v_3) = c_3v_3$$

where $c_i \neq c_j$ if $i \neq j$.

If we assume that

$$x_1v_1 + x_2v_2 + x_3v_3 = 0, \tag{9}$$

then

$$f(x_1v_1 + x_2v_2 + x_3v_3) = 0,$$

or

$$x_1f(v_1) + x_2f(v_2) + x_3f(v_3) = 0.$$

Hence,

$$x_1c_1v_1 + x_2c_2v_2 + x_3c_3v_3 = 0. \tag{10}$$

We now multiply (9) by c_3, say, and subtract the result from (10), we obtain

$$x_1c_1v_1 + x_2c_2v_2 - x_1c_3v_1 - x_2c_3v_2 = 0$$

or

$$x_1(c_1 - c_3)v_1 + x_2(c_2 - c_3)v_2 = 0. \tag{11}$$

By the result of part (a), we may conclude from (11) that

3.6.6(L) continued

$$x_1(c_1 - c_3) = 0 \text{ and } x_2(c_2 - c_3) = 0. \tag{12}$$

But since $c_1 \neq c_3$ and $c_2 \neq c_3$, we conclude from (12) that

$x_1 = 0$ and $x_2 = 0$.

We then see from (9) that $x_3 = 0$, thus, establishing that $\{v_1,v_2,v_3\}$ is linearly independent.

Note #1

The result established in this exercise gives us another "excuse" for excluding the zero vector from the ranks of the eigenvectors. In particular, the result stated in this exercise requires that v_1, v_2, and v_3 not be zero vectors (since otherwise the coefficients need not equal 0).

Note #2

If we extend the results of this exercise inductively, we may conclude that any set of eigenvectors which have distinct eigenvalues is linearly independent. In particular, then, if dim V = n, there cannot be more than n such eigenvectors (although as we shall show in the next exercise, there can be fewer) since the number of linearly independent vectors in V cannot exceed the dimension of V.

Moreover, if dim V = n and we find n such distinct eigenvalues, then the set of distinct eigenvectors which correspond to each of these eigenvalues form a basis for V.

Finally, as we mentioned earlier, if V has n such eigenvectors, we can find a matrix to represent the transformation which is diagonal and whose diagonal elements are the eigenvalues.

In summary, we have the following important theorem.

Suppose $f:V\to V$ is a linear transformation of the n-dimensional vector space V into itself. Then f has at most n distinct eigenvalues. Moreover, when the number of distinct eigenvalues is equal to n, then any complete set of eigenvectors (one for each eigenvalue) is a basis for V, and the matrix of f relative to such a basis is

3.6.6(L) continued

$$\begin{bmatrix} c_1 & & & & 0 \\ & c_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & c_n \end{bmatrix} .$$

3.6.7(L)

a. Using the transpose of the matrix of coefficients, we have

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

Hence,

$$|A - cI| = \begin{vmatrix} 2 - c & -1 & 1 \\ 0 & 2 - c & 0 \\ 1 & 3 & 2 - c \end{vmatrix},$$

and this in turn, expanding along the second row, yields

$$|A - cI| = (2 - c) \begin{vmatrix} 2 - c & 1 \\ 1 & 2 - c \end{vmatrix}$$

$$= (2 - c)[(2 - c)^2 - 1]$$

$$= (2 - c)(c^2 - 4c + 3)$$

$$= (2 - c)(c - 3)(c - 1). \qquad (1)$$

From (1), we conclude that

$$|A - cI| = 0 \leftrightarrow c = 1 \text{ or } c = 2 \text{ or } c = 3.$$

b. An eigenvector, $v = x_1u_1 + x_2u_2 + x_3u_3$, corresponding to $c = 1$ is found by solving the matrix equation

3.6.7(L) continued

$$\begin{bmatrix} 2 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{2}$$

This yields

$$\begin{bmatrix} 2x_1 - x_2 + x_3 \\ 2x_2 \\ x_1 + 3x_2 + 2x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{2'}$$

or

$$\left.\begin{array}{ll} \text{(i)} & 2x_1 - x_2 + x_3 = x_1 \\ \text{(ii)} & 2x_2 = x_2 \\ \text{(iii)} & x_1 + 3x_2 + 2x_3 = x_3 \end{array}\right\} . \tag{3}$$

From (ii), we conclude that $x_2 = 0$, whereupon (1) implies that $2x_1 + x_3 = x_1$, or $x_3 = -x_1$. This condition is reaffirmed by (iii). In other words, (3) is equivalent to

$$\left.\begin{array}{l} x_3 = -x_1 \\ x_2 = 0 \end{array}\right\} ,$$

from which we conclude that

$$f(v) = v \leftrightarrow v = x_1u_1 + 0u_2 - x_1u_3 = x_1(u_1 - u_3). \tag{4}$$

From (4), we see that

$$\alpha_1 = u_1 - u_3 \tag{5}$$

is a basis for $V_1 = \{v\varepsilon V: f(v) = v\}$.

3.6.7(L) continued

<u>Check</u>

$$\begin{aligned} f(\alpha_1) &= f(u_1 - u_3) \\ &= f(u_1) - f(u_3) \\ &= (2u_1 + u_3) - (u_1 + 2u_3) \\ &= u_1 - u_3 \\ &= \alpha_1. \end{aligned}$$

We now find an eigenvector corresponding to $c = 2$ by solving

$$\begin{bmatrix} 2x_1 - x_2 + x_3 \\ 2x_2 \\ x_1 + 3x_2 + 2x_3 \end{bmatrix}^* = \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix}$$

or,

$$\left.\begin{array}{ll} \text{(i)} & 2x_1 - x_2 + x_3 = 2x_1 \\ \text{(ii)} & 2x_2 = 2x_2 \\ \text{(iii)} & x_1 + 3x_2 + 2x_3 = 2x_3 \end{array}\right\} . \qquad (6)$$

From (1) we see that $-x_2 + x_3 = 0$ or $x_2 = x_3$. From (iii) we see that $x_1 + 3x_2 = 0$, or $x_1 = -3x_2$; and from (ii) we see that no restriction is placed on x_2. Thus, we may view x_2 as an arbitrary number, whereupon $x_3 = x_2$ and $x_1 = -3x_2$. This tells us that

$$f(v) = 2v \leftrightarrow v = -3x_2u_1 + x_2u_2 + x_2u_3 = x_2(-3u_1 + u_2 + u_3). \qquad (7)$$

From (7), we see that

---

*Notice that this matrix is $\vec{AX}$ which does not depend on c. That is, $\vec{AX} = c\vec{X}$ implies that in (2) and (2') the left side remains the same no matter what the value of c is.

3.6.7(L) continued

$$\alpha_2 = -3u_1 + u_2 + u_3 \tag{8}$$

is an eigenvector corresponding to the eigenvalue $c = 2$. That is,

$[\alpha_2] = V_2 = \{v \varepsilon V : f(v) = 2v\}$.

Check

$$\begin{aligned} f(\alpha_2) &= f(-3u_1 + u_2 + u_3) \\ &= -3f(u_1) + f(u_2) + f(u_3) \\ &= -3(2u_1 + u_3) + (-u_1 + 2u_2 + 3u_3) + (u_1 + 2u_3) \\ &= -6u_1 + 2u_2 + 2u_3 \\ &= 2(-3u_1 + u_2 + u_3) \\ &= 2\alpha_2. \end{aligned}$$

Finally, to find an eigenvector corresponding to $c = 3$, we have

$$\begin{bmatrix} 2x_1 - x_2 + x_3 \\ 2x_2 \\ x_1 + 3x_2 + 2x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ 3x_2 \\ 3x_3 \end{bmatrix}$$

or

$$\left.\begin{array}{ll} \text{(i)} & 2x_1 - x_2 + x_3 = 3x_1 \\ \text{(ii)} & 2x_2 = 3x_2 \\ \text{(iii)} & x_1 + 3x_2 + 2x_3 = 3x_3 \end{array}\right\} . \tag{9}$$

From (ii), $x_2 = 0$ whereupon both (i) and (iii) yield that $x_1 = x_3$. Hence,

3.6.7(L) continued

$f(v) = 3v \leftrightarrow v = x_1u_1 + 0u_2 + x_1u_3 = x_1(u_1 + u_3)$.

Consequently,

$$\alpha_3 = u_1 + u_3 \tag{10}$$

spans $V_3 = \{v\varepsilon V: f(v) = 3v\}$.

<u>Check</u>

$$\begin{aligned} f(\alpha_3) &= f(u_1 + u_3) \\ &= f(u_1) + f(u_3) \\ &= (2u_1 + u_3) + (u_1 + 2u_3) \\ &= 3u_1 + 3u_3 \\ &= 3(u_1 + u_3) \\ &= 3\alpha_3. \end{aligned}$$

c. By the theorem in Note #2 of the previous exercise, we have that

$$V = [\alpha_1] \oplus [\alpha_2] \oplus [\alpha_3]$$

and, in particular, the matrix of f relative to the (ordered) basis $\{\alpha_1,\alpha_2,\alpha_3\}$ is

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

That is,

3.6.7(L) continued

$$f(\alpha_1) = \alpha_1 = 1\alpha_1 + 0\alpha_2 + 0\alpha_3$$

$$f(\alpha_2) = 2\alpha_2 = 0\alpha_1 + 2\alpha_2 + 0\alpha_3$$

$$f(\alpha_3) = 3\alpha_3 = 0\alpha_1 + 0\alpha_2 + 3\alpha_3$$

d. We want to find ξ such that $f(\xi) = v_o$ for a <u>given</u> $v_o \varepsilon V$. To use the eigenvector method, we choose the basis (or one like it) $\{\alpha_1,\alpha_2,\alpha_3\}$ where $\alpha_1,\alpha_2,\alpha_3$ are as in the previous part.

Say

$$\xi = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$$

and

$$v_o = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3.$$

Then

$$\begin{aligned} f(\xi) &= f(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3) \\ &= x_1 f(\alpha_1) + x_2 f(\alpha_2) + x_3 f(\alpha_3) \\ &= x_1\alpha_1 + 2x_2\alpha_2 + 3x_3\alpha_3. \end{aligned}$$

Hence,

$$f(\xi) = v_o \rightarrow$$

$$x_1\alpha_1 + 2x_2\alpha_2 + 3x_3\alpha_3 = a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 \tag{11}$$

and since $\{\alpha_1,\alpha_2,\alpha_3\}$ is a basis for V, we conclude from (11) that

$$\begin{cases} x_1 = a_1 \\ 2x_2 = a_2 \\ 3x_3 = a_3 \end{cases}$$

3.6.7(L) continued

or

$x_1 = a_1$

$x_2 = \frac{1}{2} a_2$

$x_3 = \frac{1}{3} a_3$

What we hope is clear is that the basis consisting of eigenvectors (when such a basis exists) is particularly convenient for computational purposes.

e. Letting $v = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$, we have

$f(v) = \alpha_1 + 4\alpha_2 + 12\alpha_3 \rightarrow$

$x_1 f(\alpha_1) + x_2 f(\alpha_2) + x_3 f(\alpha_3) = \alpha_1 + 4\alpha_2 + 12\alpha_3 \rightarrow$

$x_1\alpha_1 + 2x_2\alpha_2 + 3x_3\alpha_3 = \alpha_1 + 4\alpha_2 + 12\alpha_3 \rightarrow$

$x_1 = 1,\ x_2 = 2,\ x_3 = 4.$

Hence,

$$v = \alpha_1 + 2\alpha_2 + 4\alpha_3. \tag{12}$$

If we wish to express v in terms of u_1, u_2, and u_3, we have from (5), (8), and (10) that

$$\begin{aligned} v &= (u_1 - u_3) + 2(-3u_1 + u_2 + u_3) + 4(u_1 + u_3) \\ &= -u_1 + 2u_2 + 5u_3. \end{aligned} \tag{13}$$

<u>Note</u>

Since the standard basis for V was given as $\{u_1, u_2, u_3\}$, there is a good chance that all vectors will be given in terms of u_1, u_2, and u_3 coordinates. Hence, we must expect to set up the technique

3.6.7(L) continued

used in (d) and (e) by converting vectors from u_1, u_2, u_3 coordinates into $\alpha_1, \alpha_2, \alpha_3$ coordinates. This is done in the usual way, by inverting the system of equations

$$\left.\begin{array}{ll} \text{(i)} & \alpha_1 = u_1 - u_3 \\ \text{(ii)} & \alpha_2 = -3u_1 + u_2 + u_3 \\ \text{(iii)} & \alpha_3 = u_1 + u_3 \end{array}\right\} . \tag{14}$$

In this particular case, we may avoid the row-reduced matrix technique (if we must), by observing that the sum of (i) and (iii) yields

$$u_1 = \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_3 \tag{15}$$

while the difference of (i) and (iii) yields

$$u_3 = -\frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_3. \tag{16}$$

Putting (15) and (16) into (ii) yields

$$\alpha_2 = -\frac{3}{2}\alpha_1 - \frac{3}{2}\alpha_3 + u_2 - \frac{1}{2}\alpha_1 + \frac{1}{2}\alpha_3$$

or

$$u_2 = 2\alpha_1 + \alpha_2 + \alpha_3. \tag{17}$$

From (15), (16), and (17) we can now convert from u-coordinates to α-coordinates quite readily.

f. Since we are using transposes our matrix P is the transpose of the matrix of coefficients given by

$$\alpha_1 = u_1 - u_3$$

$$\alpha_2 = -3u_1 + u_2 + u_3$$

$$\alpha_3 = u_1 + u_3$$

3.6.7(L) continued

That is

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

Then, as we discussed in Unit 4, the matrix of f relative to $\{\alpha_1, \alpha_2, \alpha_3\}$ is given by

$P^{-1}AP$.*

Now equations (15), (16), and (17) imply that

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

is the inverse of

$$\begin{bmatrix} 1 & 0 & -1 \\ -3 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = P^T.$$

Hence,

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

We then have

*Again, notice the order. We would write PAP^{-1} if we were using the matrices of coefficients rather than their inverses.

3.6.7(L) continued

$$P^{-1}AP = \begin{bmatrix} \frac{1}{2} & 2 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

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

$$\begin{matrix} \downarrow & \downarrow & \downarrow \\ f(\alpha_1) & f(\alpha_2) & f(\alpha_3) \end{matrix}$$

with respect to u_1, u_2, u_3

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

$$= D. \tag{18}$$

<u>Note</u>

From (18), we conclude that

$$(P^{-1}AP)^T = D^T = D$$

or

$$P^T A^T \left(P^T\right)^{-1} = D. \tag{19}$$

Hence, if we let B be the transpose, A^T, of the matrix of coefficients, we see from (19) that

$$QBQ^{-1} = D \text{ where } Q = P^T.$$

In other words,

3.6.7(L) continued

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

Again, it is not important whether we work with A or A^T but it is important that we keep the proper order. For example, with reference to (18), PAP^{-1} would not yield D. In particular,

$$PAP^{-1} = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 2 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}$$

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

$$= \begin{bmatrix} 3 & 5 & 0 \\ 0 & 2 & 0 \\ 0 & 5 & 1 \end{bmatrix}.$$

3.6.8 (Optional)

a. We have that the characteristic equation for f is given by

$$(c - 1)(c - 2)(c - 3) = 0$$

or

$$c^3 - 6c^2 + 11c - 6 = 0 \tag{1}$$

and that this equation does not depend on the matrix which represents f.

The more amazing fact (the proof of which is beyond our scope) is that if we replace c in (1) by the matrix of f (relative to any

3.6.8 continued

basis we choose) the equation is satisfied by the matrix.

As a check, with

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

we have

$$A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

and

$$A^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 27 \end{bmatrix},$$

whence

$$A^3 - 6A^2 + 11A - 6I =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 27 \end{bmatrix} + \begin{bmatrix} -6 & 0 & 0 \\ 0 & -24 & 0 \\ 0 & 0 & -54 \end{bmatrix} + \begin{bmatrix} 11 & 0 & 0 \\ 0 & 22 & 0 \\ 0 & 0 & 33 \end{bmatrix} + \begin{bmatrix} -6 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -6 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now, since $|PAP^{-1} - \lambda I| = |A - \lambda I|$, it would seem that the identity

$$A^3 - 6A^2 + 11A - 6I = 0 \qquad (2)$$

would hold for <u>each</u> matrix A which represents f. This turns out to be true. More emphatically, what this means is that the linear

3.6.8 continued

transformation defined by

$$f \circ f \circ f - 6 f \circ f + 11 f - 6$$

maps V into the 0-space.

b. At any rate, we have that

$$
\begin{array}{r|l}
 & \overline{A^4 + 6A^3 + 25A^2 + 90A + 301} \\
A^3 - 6A^2 + 11A - 6I & A^7 \\
 & \underline{A^7 - 6A^6 + 11A^5 - 6A^4} \\
 & \quad 6A^6 - 11A^5 + 6A^4 \\
 & \quad \underline{6A^6 - 36A^5 + 66A^4 - 36A^3} \\
 & \qquad 25A^5 - 60A^4 + 36A^3 \\
 & \qquad \underline{25A^5 - 150A^4 + 275A^3 - 150A^2} \\
 & \qquad\quad 90A^4 - 239A^3 + 150A^2 \\
 & \qquad\quad \underline{90A^4 - 540A^3 + 990A^2 - 540A} \\
 & \qquad\qquad 301A^3 - 840A^2 + 540A \\
 & \qquad\qquad \underline{301A^3 - 1806A^2 + 3311A - 1806I} \\
 & \qquad\qquad\qquad 966A^2 - 2771A + 1806I.
\end{array}
$$

In other words,

$$A^7 = (A^4 + 6A^3 + 25A^2 + 90A + 311)\underbrace{(A^3 - 6A^2 + 11A - 6I)}_{= 0} +$$

$$+ 966A^2 - 2771A + 1806I.$$

Hence,

$$A^7 = 1026A^2 - 2881A + 1866I.$$

In other words, then, every polynomial in A may be reduced to an equivalent linear combination of I, A, and A^2 by virtue of the fact that

$$A^3 - 6A^2 + 11A - 6I = 0.$$

3.6.8 continued

Note

While we do not intend to pursue this point at this level of our course, it should be noted that what we are showing here is that since each matrix satisfies its characteristic equation, we may always reduce a power of an n by n matrix A to a linear combination of I, A, ..., and A^{n-1}. This allows us, for example, to talk about power series of a matrix and to define such "weird" things as e^A, cos A, etc., where A is a matrix rather than a number. For example, we define e^A, analogous to the corresponding numerical result to be

$$I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots + \frac{A^n}{n!} + \ldots.$$

and if A happens to be a k by k matrix, we can replace A^n, whenever n is k or greater, by an appropriate linear combination of I, A, ..., and A^{k-1}.