7

Vectors, Matrices, and Tensors

For both state space control theory and kinematics, we can take advantage of matrix methods.

There is a tendency among mathematicians to regard matrices as arcane and mystic entities, with cryptic properties that reward a lifetime of study. Engineers can be duped into this point of view if they are not careful.

7.1 MEET THE MATRIX

Matrices are, in fact, just a form of shorthand that can come in very useful when a lot of calculating operations are involved. There are strict rules to observe, but when used properly matrices, vectors, and tensors are mere tools that are the servant of the engineer.

You will probably have first encountered matrices in the solution of simultaneous equations. To take a simple example, the equations

$$5x + 7y = 2$$

$$2x + 3y = 1$$

can be "tidied up" by separating the coefficients from the variables in the form

Essentials of Mechatronics, by John Billingsley

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

where the variables x and y are now conveniently grouped as a vector. Now the multiplication rule has defined itself.

We move across the top row of the matrix, multiplying each element by the corresponding component as we move down the vector to its right, adding up these products as we go. We put the resulting total in the top element, here $5x + 7y$.

Then we do the same for the next row, and so on.

7.2 MORE ON VECTORS

What does a *vector* actually "mean"? The answer has to be "anything you like." Anything, that is, that cannot be represented by a single number but requires a string of numbers to define it. It could even be a shopping list:

$$5 \text{ oranges} + 3 \text{ lemons} + 2 \text{ grapefruit}$$

can be written in matrix format as

$$[\text{orange} \quad \text{lemon} \quad \text{grapefruit}]\begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$$

which we might write in a line of text as (orange, lemon, grapefruit) (5,3,2)′ or else place the dot between them that we use for *scalar product*. The numbers on the right have defined a "mixture" of the items on the left.

Rather than fruit, we are more likely to apply vectors to coordinate systems—but we are still just picking from a list.

We might define $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ to be *unit vectors* all at right angles, say, east, north, and up. We can call them *basis vectors.*

When we say that point P has coordinates (2,3,4)′, we mean that to get there, you start at the origin and go 2 m east, then 3 m north, and 4 m up.

We could write this as

$$2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}$$

which is a mixture of the basis vectors defined by a matrix multiplication—vectors are just skinny matrices.

Now, when we turn our minds to applications, we can see many uses for vector operations. When a force $\boldsymbol{F}$ moves a load a distance $\boldsymbol{x}$, the work done is given by their scalar product $\boldsymbol{F} \cdot \boldsymbol{x}$.

As before, we take products of corresponding elements and add them up, to get a scalar number.

We usually think in terms of “the matrix multiplies the vector.” But how about thinking of the vector multiplying the matrix? What does it do to it? Consider the following matrix:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1x + 2y + 3z \\ 4x + 5y + 6z \\ 7x + 8y + 9z \end{bmatrix}$$

From one perspective, the top element is equal to the scalar product of the top row of the matrix with the vector $(x,y,z)'$. Similarly, the other elements are the scalar products of the vector with the middle and bottom rows of the matrix, respectively.

So we have

> The product of a matrix and a (column) vector is made up of the scalar products of the vector with each of the rows of the matrix.

But there is another way of seeing it. The answer is the same as

$$\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} x + \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} y + \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} z$$

So we also have

> The product of a matrix and a column vector is a mixture of the vectors that make up the columns of the matrix.

Suppose that point P is defined in terms of a second set of basis vectors, $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$, so that its coordinates $(x,y,z)'$ mean $x\boldsymbol{u} + y\boldsymbol{v} + z\boldsymbol{w}$. To find the coordinates in terms of $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$, we simply multiply and add up the contributions from $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$.

We can “transform the coordinates” by multiplying $(x,y,z)'$ by a matrix made up of columns representing vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$, to end up with a vector for P as a mixture of $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$.

7.3 MATRIX MULTIPLICATION

Often we will find a need to multiply one matrix by another. To see this in action, let us look at another simple "mixing" example.

In a candy store, "scrunches", "munches," and "chews" are on sale.

Also on sale are "Jumbo" bags each containing 2 scrunches, 3 munches, and 4 chews, and "Giant" bags containing 5 scrunches, 6 munches and only one chew. If I purchase 7 Jumbo bags and 8 Giant bags, how many of each sweet have I bought?

The bag contents can be expressed algebraically as

$$J = 2s + 3m + 4c$$

and

$$G = 5s + 6m + 1c$$

or in matrix form as

$$\begin{bmatrix} J \\ G \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 1 \end{bmatrix} \begin{bmatrix} s \\ m \\ c \end{bmatrix}$$

Note that matrices do not have to be square, as long as the terms to be multiplied correspond in number.

Now my purchase of 7 Jumbo bags and 8 Giant bags can be written as

$$7J + 8G$$

or in grander form as the product of a row vector with a column vector:

$$\begin{bmatrix} 7 & 8 \end{bmatrix} \begin{bmatrix} J \\ G \end{bmatrix}$$

But I can substitute for the J,G vector to obtain

$$\begin{bmatrix} 7 & 8 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 1 \end{bmatrix} \begin{bmatrix} s \\ m \\ c \end{bmatrix}$$

To get numerical counts of scrunches, munches, and chews we have to calculate the product of a numerical row vector with a numerical matrix. As before, we march across the row(s) of the one on the left, taking the scalar product with the columns on the right.

The answer is what common sense would give.

From 7 Jumbo bags, with scrunches at 2 to a bag, we find 7 times 2 scrunches.
From 8 Giant bags, we find 8 times 5 more, giving a grand total of 54.

The final answer is

$$[54 \quad 69 \quad 36]\begin{bmatrix} s \\ m \\ c \end{bmatrix}$$

that is 54 scrunches, 69 munches, and 36 chews.

Now the shop is selling an Easter bundle of 3 Jumbo bags and a Giant bag, and still has in stock Christmas bundles of 2 Jumbo bags and 4 Giant bags. If I buy five Easter packs and one Christmas pack, how many scrunches, munches, and chews will I have?

As an exercise, write down the matrices involved and multiply them out by the rules that we have found. (Your answer should be 89 scrunches + 105 munches + 77 chews.)

The mathematician will still worry about the order in which the matrix multiplication is carried out. We must not alter the order of the matrices, but we can group the pairs for calculation in either of two ways.

The Christmas and Easter bags can first be opened to reveal a total of Jumbo and Giant bags, then these can be expanded into individual sweets. Alternatively, work out the total of each sweet for a Christmas bag and for an Easter bag first. The result must be the same. (Check it.)

Mathematicians would say that "multiplication of matrices is associative:"

$$ABC = (AB)C = A(BC)$$

7.4 TRANSPOSITION OF MATRICES

Our mixed fruit multiplication can be written as

$$[\text{orange} \quad \text{lemon} \quad \text{grapefruit}]\begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$$

or equally well as

$$[5 \quad 3 \quad 2]\begin{bmatrix} \text{orange} \\ \text{lemon} \\ \text{grapefruit} \end{bmatrix}$$

giving 5 oranges + 3 lemons + 2 grapefruit in both cases—this result is in the form of a scalar. But note that in reversing the order in which we multiply the vectors, we have had to transpose them.

Transposing a scalar is not very spectacular—but when two matrices are multiplied together to give another matrix, $C = AB$, then, if we wish to find out the transpose of C, we must both transpose A and B and reverse the order in which we multiply them:

$$C' = B'A'$$

7.5 THE UNIT MATRIX

One last point to note before moving on is that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The matrix with 1s down its diagonal and 0s elsewhere has the special property that its product with any vector or matrix leaves that vector or matrix unchanged. Of course, there is not just one unit matrix; they come in all sizes to fit the rows of the matrix that they have to multiply. This one is the 3×3 version.

7.6 COORDINATE TRANSFORMATIONS

It has been mentioned that vector geometry is usually introduced with the aid of three orthogonal unit vectors: $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$.

For now, let us keep to two dimensions and consider just $(x,y)'$, meaning $x\boldsymbol{i} + y\boldsymbol{j}$.

Now suppose that there are two sets of axes in action. With respect to our first set the point is $(x,y)'$ but with respect to a second set it is $(u,v)'$. Just how can these two vectors be related?

What we have in effect is one pair of unit vectors $\boldsymbol{i}, \boldsymbol{j}$, and another pair, $\boldsymbol{l}$, $\boldsymbol{m}$, say. Since both sets of coordinates represent the same vector, we have

$$x\boldsymbol{i} + y\boldsymbol{j} = u\boldsymbol{l} + v\boldsymbol{m}$$

Now each of the vectors $\boldsymbol{l}$ and $\boldsymbol{m}$ must be expressible in terms of $\boldsymbol{i}$ and $\boldsymbol{j}$. Suppose that

$$\boldsymbol{l} = a\boldsymbol{i} + b\boldsymbol{j}$$

$$\boldsymbol{m} = c\boldsymbol{i} + d\boldsymbol{j}$$

or in matrix form

$$[\boldsymbol{l} \quad \boldsymbol{m}] = [\boldsymbol{i} \quad \boldsymbol{j}]\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

We want the relationship in this slightly twisted form, because we want to substitute into

$$[\boldsymbol{l} \quad \boldsymbol{m}]\begin{bmatrix} u \\ v \end{bmatrix}$$

to eliminate vectors $\boldsymbol{l}$ and $\boldsymbol{m}$ to get

$$[\boldsymbol{i} \quad \boldsymbol{j}]\begin{bmatrix} a & c \\ b & c \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}$$

Now the ingredients must match:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}$$

Although this exercise is now graced with the name "vector geometry," we are merely adding up mixtures in just the same form as the antics in the candy store.

To convert our $(u,v)'$ coordinates into the $(x,y)'$ frame, we simply multiply the coordinates by an appropriate matrix that defines the mixture.

Suppose, however, that we are presented with the values of x and y and are asked to find $(u,v)'$. We are left trying to solve two simultaneous equations:

$$x = au + cv$$

$$y = bu + dv$$

In traditional style, we multiply the top equation by d and subtract c times the second equation to obtain

$$dx - cy = (ad - bc)u$$

and in a similar way, we find

$$-bx + ay = (ad - bc)v$$

which we can rearrange as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{(ad - bc)} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where the constant $1/(ad - bc)$ multiplies each of the coefficients inside the matrix.

If the original relationship between $(x,y)'$ and $(u,v)'$ was

$$\begin{bmatrix} x \\ y \end{bmatrix} = T \begin{bmatrix} u \\ v \end{bmatrix}$$

then we have found an "inverse matrix" such that

$$\begin{bmatrix} u \\ v \end{bmatrix} = T^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

The value of $(ad - bc)$ obviously has special importance—we will have great trouble in finding an inverse if $(ad - bc) = 0$. Its value is the "determinant" of the matrix T.

7.7 MATRICES, NOTATION, AND COMPUTING

In a computer program, rather than using separate variables x, y, u, v, and so on, it is more convenient mathematically to use "subscripted variables" as the elements of a vector.

The entire vector is then represented by the single symbol $\boldsymbol{x}$, which is made up of several elements x_1, x_2, and so on.

Matrices are now made up of elements with two suffices:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

In a computer program, the subscripts appear in brackets, so that a vector could be represented by the elements **`X(1)`**, **`X(2)`**, and **`X(3)`**, while the elements of the matrix are **`A(1,1)`**, **`A(1,2)`**, and so on.

It is in matrix operations that this notation really earns its keep. Suppose that we have a relationship

$$\boldsymbol{x} = T\boldsymbol{u}$$

where the vectors have three elements and the matrix is 3×3. Instead of a massive block of arithmetic, the entire product is expressed in just five lines of Basic program:

```
FOR I=1 TO 3
   X(I)=0
   FOR J=1 TO 3
      X(I)=X(I)+T(I, J)*U(J)
   NEXT J
NEXT I
```

For the matrix product $C = AB$, the program is hardly any more complex:

```
FOR I=1 TO 3
   FOR J=1 TO 3
      C(I,J)=0
      FOR K=1 TO 3
         C(I,J)=C(I,J)+A(I,K)*B(K,J)
      NEXT K
   NEXT J
NEXT I
```

Or in Java or C it becomes

```
for(i = 1; i<=3; i++){
   for(j = 1; j<=3; j++) {
      c[i][j] = 0;
      for(k = 1; k<=3;k++) {
            c[i][j] += a[i][k]*b[k][j];
      }
   }
}
```

These examples would look almost identical in a variety of languages and would show the same economy of programming effort.

In Matlab the shorthand of matrix operations goes even farther—but there is a danger that the engineroom will be lost to view behind the paintwork.

> Clearly, if we are to try to analyze any except the simplest of systems by computer, we should first represent the problem in a matrix form.

But beware!!

If you have no computer to hand, it will almost certainly be quicker, easier, and less prone to errors to use non-matrix methods to solve the problem.

7.8 EIGENVECTORS

If we multiply a vector and a matrix, what do we get?

We get another vector. For example

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

From the vector (1,0)′, we get (1,–1)′. This new vector is not only a different "size"; it represents a different direction. Another example is:

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

So, from the vector (0,1)′, we get (2,4)′—again in a new direction.

Are there any vectors that can be multiplied by the matrix

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

to give another vector in the same direction?

If we start with $(x,y)'$, another vector in the same direction will be $(\lambda x,\lambda y)'$—where λ is some constant.

We are looking for a vector $\boldsymbol{x}$ for which

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

or

$$A\boldsymbol{x} = \lambda I\boldsymbol{x}$$

where I is the unit matrix. We can move both terms to the lefthand side to get

$$A\boldsymbol{x} - \lambda I\boldsymbol{x} = \mathbf{0}$$

or

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0}$$

where the **0** is a vector with all components zero.

You will recall that we could consider the matrix–vector product as a mixture of the columns of the matrix.

So here, if the vector $\boldsymbol{x}$ is not **0**, we have a combination of the columns of $(A - \lambda I)$ that will give (0,0)′.

Remember also that to evaluate a determinant of a matrix, you can first add multiples of columns to other columns of the matrix without changing the determinant's value.

Thus we have a way to reduce a column of $(A - \lambda I)$ to all zeros, and so its determinant must be zero.

Now, when we construct $A - \lambda I$ and take its determinant, we get

$$\det\begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = 0$$

which we can expand as

$$(1-\lambda)(4-\lambda) - (-1)2 = 0$$

or

$$\lambda^2 - 5\lambda + 6 = 0$$

So, we have not just one value for λ, but two: 2 and 3.

If we substitute the value 2, we get

$$\begin{bmatrix} 1-2 & 2 \\ -1 & 4-2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which is satisfied if $\boldsymbol{x} = (2,1)'$.

Let us try it out:

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

So $A\boldsymbol{x} = 2\boldsymbol{x}$, just as we hoped to find, and $\boldsymbol{x}$ is an *eigenvector* of A. The value of λ is called an *eigenvalue*.

As an exercise, find the other eigenvector, corresponding to eigenvalue $\lambda = 3$.

If the matrix A is $n \times n$, the equation for λ will be nth order and there will be n roots. But the method is just the same:

1. Write down $(A - \lambda I)$ and take its determinant.
2. Equate the determinant to 0, giving a polynomial for λ.
3. Solve this, to get a set of n eigenvalues.
4. For each eigenvalue, substitute that value back into $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$, getting a set of simultaneous equations for the elements of $\boldsymbol{x}$.
5. Solve these equations, and you have each corresponding eigenvector.