2.2 Transformations, canonical forms and eigen-values

Introduction

We have seen how the dynamics of a linear, time invariant network may be represented by a set of linear state-space equations, describing the behaviour of the system as well as its output.

Consider the following set of linear equations:

$$d/dt\,[X] = AX + BU$$
$$d/dt\,[Y] = CX + DU$$

where X represents the state vector and Y is the output. Matrix A is the system matrix. Matrices B, C and D represent the influence of the input on the state, the influence of the state on the output and of the input on the output, respectively.

[We have referred to this as a standard form of representation. A standard form is technically known as a canonical form.]

The state vector X is not unique.

An infinite number of linear transformations of X exist, which are equally valid representations of the system.

Let us consider a simple example.

$$dx_1/dt = x_1 + x_2$$
$$dx_2/dt = x_1 - x_2$$

This is a simple second order system, with no excitation.

Now consider a new set of variables z_1 and z_2 such that:

$$z_1 = x_1 + 0.4142\, x_2$$
$$z_2 = x_1 - 2.4142\, x_2$$

Then, we have:

$$x_1 = 0.8536\, z_1 + 0.1464\, z_2$$
$$x_2 = 0.3536\, z_1 - 0.3536\, z_2$$

Substituting, we get:

$dz_1/dt = 1.414\, z_1$
$dz_2/dt = -1.414\, z_2$

What we have found is that the two systems

$dx_1/dt = x_1 + x_2$
$dx_2/dt = x_1 – x_2$

and

$dz_1/dt = 1.414\, z_1$
$dz_2/dt = -1.414\, z_2$

are equivalent.

We have used a somewhat funny-looking transformation from X to Z, and ended up with an unusual set of equations in Z. It is unusual because the two equations in Z are uncoupled from each other. You will suspect that the transformation we used is not an arbitrary one, but a very special one to yield such a result. On the other hand, it also illustrates that there can be any number of transformations (as long as the two relations are independent of each other) yielding new sets of state variable, which are equally valid descriptions of the system.

The particular form of the equation that we obtained by this unusual transformation is known as a diagonal matrix, and is of great significance in the study of dynamic systems. This form of the system equations was made possible because of a particular property of the system we had chosen, that it has distinct characteristic roots or in other words, distinct eigen-values. Remember that we found that the roots of system functions of RLC networks are simple, so that we will not come across systems with multiple roots in the study of such systems.

We will now look at a more systematic method for the transformation of system matrices with distinct eigen-values to diagonal form. You will find that it is not possible to transform matrices with multiple eigen-values to diagonal form.

The Jordan canonical form is a more inclusive form (that includes the diagonal form for matrices with distinct eigen-values) that allows us to have a common standard or canonical form, similar to any given matrix. Matrices are said to be similar if they have the same eigen-values.

2.2.1 Eigen-values and eigen-vectors

Eigen-value is a hybrid word, made up of the German 'eigen' (meaning 'characteristic') and the English 'value'.

As its name implies, its value is characteristic of the system that it represents. We also have other related concepts such as the characteristic equation, characteristic roots etc.

Eigen-values are characteristic of a matrix. In the case of the system equations that we studied, the matrix is the system matrix A.

The eigen-values are given by the solution of the equation:

$$|A - \lambda I| = 0$$

We could of course look at this as the solution to the problem of finding a (non-trivial) solution to:

$$Ax = \lambda x$$

Let us consider the example that we already have:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) - 1$$

$$\therefore \lambda^2 - 2 = 0$$

$$\lambda = \pm\sqrt{2}$$

The eigen-values are +1,414 and –1,414

These correspond to the transformed equation that we obtained, namely,

$$\dot{z}_1 = 1{,}414 z_1$$

$$\dot{z}_2 = -1.414 z_2$$

or, in matrix form:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 1.414 & 0 \\ 0 & -1.414 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

We can obtain the eigen-values of a matrix using the function *eig* in MATLAB. The following is a transcript of a MATLAB session, which calculates the eigen-values of the above matrix:

```
A=[1 1; 1 –1]

A=
        1       1
        1      -1

lamda=eig(A)

lamda=
       -1.4142
        1.4142
```

Now let us try some slightly more complex examples:

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

Using MATLAB, let us evaluate their eigen-values:

```
B={1 1 0; 0 1 1; 0 0 1]

B=
        1       1       0
        0       1       1
        0       0       1
eig(B)

ans =

        1
        1
        1
C=[2 2 1; 1 2 2; 1 1 2]
```

```
C=
        2       2       1
        1       2       2
        1       1       2

eig(C)

ans =
        4.6274
        0.6863 + 0.4211i
        0.6863 - 0.4211i
```

These examples illustrate two other phenomena. Matrix B has three coincident eigen-values, while C has one real eigen-value and a pair of complex eigen-values, which are conjugates of each other.

We are now ready to look at another concept, that of eigen-vectors.

Associated with each distinct characteristic value (eigen-value) λ, there is a characteristic vector (eigen-vector), determined up to a scalar multiple. Earlier, we used the relationship

$$Ax = \lambda x$$

to find the eigen-values λ, so that x is non-trivial. We will now find these values of x, corresponding to each eigen-value.

$$(A - \lambda I)\, x = 0$$

for the example considered.

With $\lambda_1 = \sqrt{2}$, the equations are:

$$-0.4142x_1 + x_2 = 0$$
$$x_1 - 2.4142x_2 = 0$$

From either of these, we can obtain

$$x_1 = 2.4142x_2$$

[1/0.4142 is equal to 2.4142, so that both equations yield the same relationship]

If we select $x_2 = 1$, then $x_1 = 2.4142$.

If we normalise this vector, such that the sum of their squares is equal to 1,

$$x_1 = \frac{2.4142}{\sqrt{\left(2.4142^2 + 1^2\right)}} = \frac{2.4142}{2.6131} = 0.9239$$

$$x_2 = \frac{1}{\sqrt{\left(2.4142^2 + 1^2\right)}} = 0.3827$$

If we now start with the other eigen-value
$\lambda_2 = -\sqrt{2}$, we get:

$$2.4142\,x_1 + x_2 = 0$$
$$x_1 + 0.4142\,x_2 = 0$$

Either of these will yield the relationship

$$x_1 = -0.4142\,x_2$$

We can, up to a scalar multiple, assume

$$x_1 = -0.4142$$
$$x_2 = 1$$

After normalisation, we get:

$$x_1 = -0.3827$$
$$x_2 = 0.9239$$

We have now calculated the two eigen-vectors of the matrix A, up to (a) scalar multiple. They are:

$$\begin{bmatrix} 0.9239 \\ 0.3827 \end{bmatrix}, \begin{bmatrix} -0.3827 \\ 0.9239 \end{bmatrix}$$

We can use MATLAB to obtain the eigen-vectors of a matrix. Let us illustrate this using the example used:

```
A=[1 1; 1 –1]

A =
        1        1
        1        -1

[V,D] = eig(A)

V =
        0.3827          -0.9239
        -0.9239         -0.3827
```

```
D =
          -1.4142                   0
                0            1.4142
```

[Note: When the function *eig* is used with two expected matrix responses as here, the first answer returned is an array of eigen vectors and the second is a diagonal matrix with eigen-values on the diagonal.]

Note that the eigen-vectors computed by MATLAB are different from the ones we obtained, by a factor of (-1). This is because they are determined only up to a scalar multiple, and even normalisation leaves us with this ambiguity.

```
>> B=[1 1 0;0 1 1; 0 0 1]

B =

   1   1   0
   0   1   1
   0   0   1

>> C=[2 2 1; 1 2 2; 1 1 2]

C =

   2   2   1
   1   2   2
   1   1   2

>> [P,Q]=eig(B)

P =

  1.0000          -1.0000           1.0000
       0           0.0000          -0.0000
       0           0                0.0000

Q =

   1   0   0
   0   1   0
   0   0   1

>> [R,S]=eig(C)

R =

  0.6404            0.7792              0.7792
  0.6044           -0.4233 + 0.3106i   -0.4233 - 0.3106i
  0.4738           -0.1769 - 0.2931i   -0.1769 + 0.2931i

S =

  4.6274           0                    0
       0           0.6863 + 0.4211i     0
       0           0                    0.6863 - 0.4211i
```

[Note that this algorithm has not been able to compute the eigen-vectors of the matrix B, as it has multiple eigen-values. An indication of this may be obtained by invoking the function *condeig,* which returns a vector of condition numbers for the evaluation of eigen-vectors. For a well-conditioned matrix, it should have values close to unity.

```
>> condeig(B)

Warning: Matrix is close to singular or badly scaled.
Results may be inaccurate. RCOND = 2.465190e-032.
> In D:\matlabR12\toolbox\matlab\matfun\condeig.m at line 30

ans =

  1.0e+031 *

    0.0000
    2.0282
    2.0282
```

]

2.2.2 Diagonal matrices

Let us assume that the matrix A has distinct characteristic roots or eigen-values $\lambda_1, \lambda_2, \lambda_3. \ . \ . \ , \lambda_n$. Let us further assume that the eigen-vectors associated with these eigen-values are $x^1, x^2, x^3, \ . \ . \ . \ . \ , x^n$. We will further assume that the eigen-vectors have been normalised, so that inner product of each eigen-vector and itself is unity:

$$(x^i, x^i) = \sum_{j=1}^{n} (x_j^i)^2 = 1, for \ \ i = 1, n$$

Consider the matrix T formed by assembling the vectors x^i as columns, as follows:

$$T = \left[x^1 \,|x^2 \,|x^3 \ldots |\, x^n \right]$$

$$= \begin{bmatrix} x_1^1 & x_1^2 & . & . & x_1^n \\ x_2^1 & x_2^2 & . & . & x_2^n \\ . & . & . & . & . \\ x_n^1 & x_n^2 & . & . & x_n^n \end{bmatrix}$$

T', the transpose of T, would then be the matrix obtained by arranging the eigen-vectors as rows:

$$T' = \begin{bmatrix} x^1 \\ x^2 \\ . \\ . \\ . \\ x^n \end{bmatrix} = \begin{bmatrix} x_1^1 & x_2^1 & . & . & x_n^1 \\ x_1^2 & x_2^2 & . & . & x_n^2 \\ . & . & & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_1^n & x_2^n & . & . & x_n^n \end{bmatrix}$$

Since the eigen-vectors are orthogonal, we have:

$$T'T = ((x^i x^j)) = (\delta_{ij})$$

[T is thus an orthogonal matrix.]

As each of the x^i are eigen-vectors, we have:

$$Ax^i = \lambda_i x^i$$

so that:

$$AT = [\lambda_1 x^1 \mid \lambda_2 x^2 \mid . \quad \mid \lambda_n x^n]$$

$$\therefore T'AT = (\lambda_i (x^i x^j)) = (\lambda_i \delta_{ij})$$

$$\therefore T'AT = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & 0 & \\ & & \lambda_3 & & \\ . & & & & \\ & 0 & & & \\ & & & & \lambda_n \end{bmatrix}$$

T'AT has eigen-values on the diagonal, and zero everywhere else. It is a diagonal matrix.

Let us define Λ (lamda) = T'AT

Pre-multiplying by T and post-multiplying by T', we have:

$T\Lambda T' = TT'\ A\ TT' = A$

For the examples we considered earlier, from the MATLAB simulation, we had:

A =

```
     1     1
     1    -1
```

T=

```
0.3827      -0.9239
-0.9239     -0.3827
```

Since T is symmetric, T' = T. and we have:

lamda =

```
  -1.4142   -0.0000
  -0.0000    1.4142
```

This, as expected, is a diagonal matrix with the eigen-values on the diagonal.
We can now get back A as $T\Lambda T'$:

ans =

```
    1.0000    1.0000
     1.00      -1.0000
```

2.2.3 The Jordan canonical form

We saw how symmetrical matrices with real eigen-values may be transformed to diagonal form.

The following are two important properties of real symmetric matrices:

- The characteristic roots (eigen-values) of a real symmetric matrix are real.
- The characteristic vectors (eigen-vectors) associated with distinct characteristic roots of a real symmetric matrix are orthogonal.

In practice, we do come across matrices with complex eigen-values. What difference does it make?

Before we proceed, we will need to agree on certain notations.

We have already used the terms *symmetric* matrix and *transpose* of a matrix.

1. A *symmetric matrix* is one where

$$a_{ij} = a_{ji}$$

2. The *transpose* A' of a matrix A is given by:

$$A = (a_{ij})$$
$$A' = (a_{ji})$$

[For a symmetric matrix, A=A']

3. When the elements a_{ij} of a matrix are real, we call such a matrix a *real matrix*.

4. The *inner product* of two vectors x and y is written as (xy), and is an important scalar function of x and y. It is defined as

$$(xy) = \sum_{1}^{n} x_i y_i$$

5. The complex conjugate of a complex variable x is denoted by placing a ¯ over x.

$$x = \alpha + j\beta$$
$$\bar{x} = \alpha - j\beta$$

[For complex vectors, the product $(x, \bar{y})$ is of greater significance than the usual inner product (x,y)]

6. Corresponding to the symmetric real matrices where A = A', the significant form for complex matrices is called *Hermitian matrices*, where:

$$A = \bar{A}' = A^*$$

[Parallel with the orthogonal matrices in the study of symmetric matrices, we have the concept of unitary matrices in the study of Hermitian matrices, where $T^*T = I$]

Let us now consider an example of a non-symmetric matrix with real coefficients:

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

As before, we will look at a MATLAB transcript.

```
» A=[1 -2 ; 2 1]

A =

     1    -2
     2     1

» [T,D]=eig(A)

T =

  -0.7071              -0.7071
       0 + 0.7071i          0 - 0.7071i


D =

   1.0000 + 2.0000i        0
        0              1.0000 - 2.0000i

» TINV= inv(T)

TINV =

  -0.7071                   0 - 0.7071i
  -0.7071                   0 + 0.7071i

» LAMDA=TINV*A*T

LAMDA =

   1.0000 + 2.0000i        0
        0              1.0000 - 2.0000i

» T*LAMDA*TINV

ans =

     1    -2
     2     1

»
```

We first compute the matrix formed by the eigen-vectors (T) and one with the eigen-values as diagonal elements (D). Note that the eigen-values are a complex conjugate pair. We then find the inverse TINV of T. We then note that

$$\Lambda = T^{-1}AT$$

is the diagonal matrix with the eigen-values on the diagonal, the same as D. Finally, we can get back A as:

$$A = T\Lambda T^{-1}$$

[In this particular case, the inverse of T is equal to the transpose of its complex conjugate, but this is not so in general, unless there are only complex roots.]

```
» CTRANS=ctranspose (T)

CTRANS =

  -0.7071              0 - 0.7071i
  -0.7071              0 + 0.7071i

» TINV=inv(T)

TINV =

  -0.7071              0 - 0.7071i
  -0.7071              0 + 0.7071i
```

Here, the transpose and the inverse are equal.

We will consider one more example.

```
» B=[1 2 3; 2 0 1; 1 2 0]

B =

     1     2     3
     2     0     1
     1     2     0

» [T,D]=eig(B)

T =

  -0.7581          -0.3589 - 0.4523i  -0.3589 + 0.4523i
  -0.4874          -0.2122 + 0.5369i  -0.2122 - 0.5369i
  -0.4332           0.5711 - 0.0847i   0.5711 + 0.0847i


D =

   4.0000                0                 0
        0          -1.5000 + 0.8660i       0
        0                0           -1.5000 - 0.8660i

» CTRANS=ctranspose(T)

CTRANS =

  -0.7581           -0.4874           -0.4332
  -0.3589 + 0.4523i  -0.2122 - 0.5369i   0.5711 + 0.0847i
  -0.3589 - 0.4523i  -0.2122 + 0.5369i   0.5711 - 0.0847i

» TINV=inv(T)

TINV =
```

```
  -0.5957 + 0.0000i  -0.5957 - 0.0000i  -0.5957
  -0.2826 + 0.3820i  -0.1359 - 0.6071i   0.6474 + 0.0145i
  -0.2826 - 0.3820i  -0.1359 + 0.6071i   0.6474 - 0.0145i
```

The transpose and the inverse are not the same.

```
» LAMDA=TINV*B*T

LAMDA =

   4.0000 - 0.0000i  -0.0000 - 0.0000i  -0.0000 + 0.0000i
   0.0000 - 0.0000i  -1.5000 + 0.8660i  -0.0000 - 0.0000i
   0.0000 + 0.0000i  -0.0000 + 0.0000i  -1.5000 - 0.8660i

» T*LAMDA*TINV

ans =

   1.0000 - 0.0000i   2.0000 + 0.0000i   3.0000 - 0.0000i
   2.0000 + 0.0000i        0 + 0.0000i   1.0000 + 0.0000i
   1.0000 - 0.0000i   2.0000 + 0.0000i   0.0000 - 0.0000i
```

Note that transpose of the complex conjugate of T and the inverse of T are quite different. However, we do have a systematic method for transforming a matrix to diagonal form. There is still one assumption we have made, that the eigen-values are distinct.

Let us now look at the case where you get coincident eigen-values.

We will re-examine the example we considered earlier:

$$C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

```
» C=[1 1 0; 0 1 1; 0 0 1]

C =

    1    1    0
    0    1    1
    0    0    1

» eig(C)

ans =

    1
    1
    1
```

There are three coincident eigen-values, each equal to 1. If you attempt to find the eigen-vectors, you will find that x_1 may take any value, but that the other two components are identically zero. MATLAB also gives the same result.

```
» [T,D]=eig(C)

T =

   1.0000   -1.0000    1.0000
        0    0.0000   -0.0000
        0         0    0.0000


D =

     1     0     0
     0     1     0
     0     0     1
```

T cannot be inverted, as its rank is only 1.

We are now ready to introduce the *Jordan canonical form*.

Definition

Let us denote by $L_k(\lambda)$ a k x k matrix of the form:

$$L_h(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & & & & & 0. \\ 0 & \lambda & 1 & 0 & & . & . & 0 \\ . & . & . & & & & & . \\ . & . & . & . & . & . & . & . \\ & & & & & & \lambda & 1 \\ 0 & 0 & & & & & & \lambda \end{bmatrix}$$

[L_1 would be equal to λ]

It can be shown that there exists a matrix T such that

$$T^{-1}AT = \begin{bmatrix} L_{k_1}(\lambda_1) & & & \\ & L_{k_2}(\lambda_2) & & 0 \\ & & & \\ & 0 & & \\ & & & L_{k_r}(\lambda_r) \end{bmatrix}$$

$$k_1 + k_2 + . . . + k_r = n$$

λ_1,λ_2, . . .,λ_r are eigen-values with multiplicity k_1, k_2, . . . k_r.

This representation (that is $T^{-1}AT$) is called the **Jordan Canonical Form.**

The diagonal form we had for the case of distinct eigen-values obviously satisfies this condition.

The matrix

$$C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

which we picked for study earlier on is already in this form.

We will now consider another example.

Consider

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

The eigen-values of A are 3 and 3. We can then write down the Jordan canonical form of the matrix as:

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

This does not, however, tell us anything of the transformation T that will yield J from A, other than that:

$$J = T^{-1}AT$$