Application of Diagonalization of Coefficient Matrices to Differential Equations

First-order Differential Equations

Consider the following three by three system:

\[
\begin{align*}
\dot{y}_1 &= y_1 - y_2 + 2y_3 \\
\dot{y}_2 &= 3y_1 + 4y_3 \\
\dot{y}_3 &= 2y_1 + y_2
\end{align*}
\]

We can write this set of equations as \( \dot{Y} = [A][Y] \) where

\[
[A] = \begin{bmatrix}
1 & -1 & 2 \\
3 & 0 & 4 \\
2 & 1 & 0
\end{bmatrix}
\]

The eigenvalues of \( [A] \) are 1.303, -2.303 and 2. The corresponding eigenvectors are:

\[
\begin{bmatrix}
-0.172 \\
0.891 \\
0.42
\end{bmatrix}, \quad
\begin{bmatrix}
-0.557 \\
-0.467 \\
0.687
\end{bmatrix}, \quad \text{and} \quad
\begin{bmatrix}
0 \\
0.894 \\
0.447
\end{bmatrix}
\]

The modal transformation matrix is:

\[
[M] = \begin{bmatrix}
-0.172 & -0.557 & 0 \\
0.891 & -0.467 & 0.894 \\
0.42 & 0.687 & 0.447
\end{bmatrix}
\]

\( [M] \) diagonalizes \( [A] \) and setting \( [Y] = [M][Z] \) transforms \( \dot{Y} = [A][Y] \) into \( \dot{Z} = [\Lambda][Z] \) where

\[
[\Lambda] = [M]^{-1}[A][M]
\]

\[
\begin{bmatrix}
1.303 & 0 & 0 \\
0 & -2.303 & 0 \\
0 & 0 & 2
\end{bmatrix}
\]

Therefore, \( \dot{Z} = [\Lambda][Z] \) can be written as:
\[ \dot{z}_1 = 1.303z_1 \]
\[ \dot{z}_2 = -2.303z_2 \]
\[ \dot{z}_3 = 2z_3 \]

These equations can be solved easily.

\[ z_1 = ae^{1.303t} \]
\[ z_2 = be^{-2.303t} \]
\[ z_3 = ce^{2t} \]

Therefore,

\[
[Z] = \begin{bmatrix}
ae^{1.303t} \\
be^{-2.303t} \\
ce^{2t}
\end{bmatrix}
\]

Since \([Y]=[M][Z]\)

\[
[Y] = \begin{bmatrix}
-0.172 & -0.557 & 0 \\
0.891 & -0.467 & 0.894 \\
0.42 & 0.687 & 0.447
\end{bmatrix}
\begin{bmatrix}
ae^{1.303t} \\
be^{-2.303t} \\
ce^{2t}
\end{bmatrix}
\]

In terms of component, the general solution is

\[ y_1 = -0.172ae^{1.303t} - 0.557be^{-2.303t} \]
\[ y_2 = 0.891ae^{1.303t} - 0.467be^{-2.303t} + 0.894ce^{2t} \]
\[ y_3 = 0.42ae^{1.303t} + 0.687be^{-2.303t} + 0.447ce^{2t} \]

The arbitrary constants \(a, b\) and \(c\) can be found if initial condition is given. Given an initial condition, \([Y(0)] = [0 \quad 0 \quad 1]^T\) we have the following.

\[-0.172a - 0.557b = 0 \]
\[0.891a - 0.467b + 0.894c = 0 \]
\[0.42a + 0.687b + 0.447c = 1 \]
Therefore, constants \( a, b \) and \( c \) are -3.228, 0.997 and 3.738 respectively.

**Exercise problems**

Solve the systems of differential equations using diagonalization

\[
\begin{align*}
1. \quad \dot{y}_1 &= 5y_1 - y_2, \quad Y(0) = 1 \\
\dot{y}_2 &= -2y_2, \quad Y(0) = 1 \\
\dot{y}_3 &= 8y_1 + 7y_2, \quad Y(0) = 1 \\
2. \quad \dot{y}_1 &= 5y_1 - 4y_2 + 2y_3, \quad Y(0) = 0 \\
\dot{y}_2 &= y_1 - 6y_2 + 3y_3, \quad Y(0) = 0 \\
\dot{y}_3 &= 2y_3, \quad Y(0) = 1 \\
3. \quad \dot{y}_1 &= 3y_1 - y_2 - y_3, \quad Y(0) = 1 \\
\dot{y}_2 &= y_1 + 2y_2 - y_3, \quad Y(0) = 0 \\
\dot{y}_3 &= 4y_2 - y_3, \quad Y(0) = 0 \\
4. \quad \dot{y}_1 &= 4y_1 - 6y_2 + y_3, \quad Y(0) = 1 \\
\dot{y}_2 &= y_1 + 3y_2 - 2y_3, \quad Y(0) = 0 \\
\dot{y}_3 &= y_2 + 4y_3, \quad Y(0) = 2
\end{align*}
\]

**Second-order Differential Equations**

Consider the mass-spring system shown in Figure 1 where \( y_1 \) and \( y_2 \) are displacements of masses \( m_1 \) and \( m_2 \) respectively from equilibrium positions. The spring constants are \( k_1 \) and \( k_2 \) as shown and both masses are chosen as 1.

![Mass-spring System](image)

**Figure 1. Mass-spring System**
Assume that there are no damping and external forces acting on the system. The motions of the masses are governed by the following differential equations.

\[
\begin{align*}
\dddot{y}_1 &= -(k_1 + k_2)y_1 + k_2 y_2 \\
\dddot{y}_2 &= -k_2 (y_2 - y_1) = k_2 y_1 - k_2 y_2
\end{align*}
\]

where \( y_1 \) and \( y_2 \) are functions of time \( t \) and the dots denotes derivatives with respect to time. In matrix form the equations can be written as

\[
\begin{bmatrix}
\dddot{y}_1 \\
\dddot{y}_2
\end{bmatrix} =
\begin{bmatrix}
-(k_1 + k_2) & k_2 \\
-2k_2 & -k_2
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]

For the sake of this analysis we assume that \( k_1 = 5 \) and \( k_2 = 6 \). Then the equation in matrix form becomes

\[
\begin{bmatrix}
\dddot{y}_1 \\
\dddot{y}_2
\end{bmatrix} = [A]\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]

where \( [A] = \begin{bmatrix} -11 & 6 \\ 6 & -6 \end{bmatrix} \).

Let us try \( [Y] = e^{wt}[C] \) as a solution where \( w \) will be determined and \( [C] \) is an unknown constant matrix

\[
[C] = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}
\]

Substituting \( [Y] = e^{wt}[C] \) into the differential equation \( \dddot{Y} = [A][Y] \) gives us

\[
w^2 e^{wt}[C] = [A]e^{wt}[C] \tag{1}
\]

Dividing Eqn. (1) by \( e^{wt} \) gives us

\[
[A][C] = w^2 [C]
\]

or, \( [A][C] = \lambda[C] \) \tag{2}

where \( \lambda = w^2 \)

Eqn. (2) shows that \( [C] \) must be an eigenvector corresponding to eigenvalue \( \lambda \) for \( [Y] = e^{wt}[C] \) to satisfy the differential equation \( \dddot{Y} = [A][Y] \). Therefore, the next step would be to solve for the eigenvalues of \( [A] \). The eigenvalues are \( \lambda_1 = -2 \) and \( \lambda_2 = -15 \). Corresponding to \( \lambda_1 = -2 \) we get the eigenvector \( \begin{bmatrix} 0.555 \\ 0.832 \end{bmatrix} \).

Therefore \( w_1 = \pm \sqrt{2i} \) and we get solution of the form

\[
[Y_1] = \alpha e^{2i\beta} \begin{bmatrix} 0.555 \\ 0.832 \end{bmatrix} + \beta e^{-2i\beta} \begin{bmatrix} 0.555 \\ 0.832 \end{bmatrix} \tag{3}
\]
Corresponding to \( \lambda_2 = -15 \) we get the eigenvector \( \begin{bmatrix} -0.832 \\ 0.555 \end{bmatrix} \).

Therefore \( w_1 = \pm \sqrt{15}i \) and we get solution of the form

\[
\begin{bmatrix} Y_2 \end{bmatrix} = \gamma e^{\sqrt{15}it} \begin{bmatrix} -0.832 \\ 0.555 \end{bmatrix} + \delta e^{-\sqrt{15}it} \begin{bmatrix} -0.832 \\ 0.555 \end{bmatrix}
\]

(4)

\( \alpha, \beta, \gamma, \) and \( \delta \) are arbitrary scalars. These solutions can also be written as

\[
\begin{bmatrix} Y_1 \end{bmatrix} = \left[ a \cos(\sqrt{2}t) + b \sin(\sqrt{2}t) \right] \begin{bmatrix} 0.555 \\ 0.832 \end{bmatrix}
\]

(5)

\[
\begin{bmatrix} Y_2 \end{bmatrix} = \left[ c \cos(\sqrt{15}t) + d \sin(\sqrt{15}t) \right] \begin{bmatrix} -0.832 \\ 0.555 \end{bmatrix}
\]

(6)

The general solution is:

\[
\begin{bmatrix} Y \end{bmatrix} = [Y_1] + [Y_2]
\]

In terms of individual displacement

\[
y_1 = 0.555a \cos(\sqrt{2}t) + 0.555b \sin(\sqrt{2}t) - 0.832c \cos(\sqrt{15}t) - 0.832d \sin(\sqrt{15}t)
\]

\[
y_2 = 0.832a \cos(\sqrt{2}t) + 0.832b \sin(\sqrt{2}t) + 0.555c \cos(\sqrt{15}t) + 0.555d \sin(\sqrt{15}t)
\]

We can determine the values of \( a, b, c, \) and \( d \) if we know the initial position and initial velocity.

**Exercise problems**

Solve the system \( [\ddot{Y}] = [A][Y] \), with \([A]\) as given below.

1. \( [A] = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \), \( [Y(0)] = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \), \( [\dot{Y}(0)] = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \)

2. \( [A] = \begin{bmatrix} 6 & -3 \\ 2 & 0 \end{bmatrix} \), \( [Y(0)] = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \), \( [\dot{Y}(0)] = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \)

3. \( [A] = \begin{bmatrix} 5 & 1 \\ 0 & -1 \end{bmatrix} \), \( [Y(0)] = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \), \( [\dot{Y}(0)] = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \)

4. \( [A] = \begin{bmatrix} -2 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix} \)
Definitions:

**Symetric Matrix**: A symmetric matrix is a square matrix that satisfies $[A]^T = [A]$

The eigenvalues of a real symmetric matrix are real.


An orthogonal matrix has the following properties.

1. The rows (columns) of an orthogonal matrix are orthogonal in pairs.
   \[
   \sum_{k=1}^{n} a_{ik} a_{jk} = 0 \quad \text{for} \quad i \neq j \quad \text{and} \quad \sum_{k=1}^{n} a_{ki} a_{kj} = 0 \quad \text{for} \quad i \neq j
   \]

2. The sum of the squares of the elements of each row (column) of an orthogonal matrix is equal to unity.
   \[
   \sum_{k=1}^{n} a_{ik}^2 = \sum_{k=1}^{n} a_{ki}^2 = 1
   \]

3. The determinant of an orthogonal matrix is equal to $\pm 1$.

4. The transpose and the inverse of an orthogonal matrix are also orthogonal matrices.

**Orthogonal Matrices and Real Quadratic Forms**

A real quadratic form in $x_1, x_2, \cdots, x_n$ can be written as
\[
\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j
\]
where $a_{ij}$'s are all real numbers. This is a polynomial expression similar to $6x_1^2 + 4x_1 x_2 - x_2^2$ and may arise in many areas including physics and analytic geometry. In many cases we want to write this polynomial expression in a more compact form, specifically in self terms only. This requires a coordinate transformation.

The first step towards a coordinate transformation is to arrange the polynomial expression into a matrix expression. The polynomial expression, $\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$ will include square terms $a_{11} x_1^2, a_{22} x_2^2, \cdots, a_{nn} x_n^2$ and product (mixed) terms.
\((a_{ij} + a_{ji})x_i x_j, \quad i \neq j\) and can be written as a matrix product \([X]^T [A] [X]\), where \([A]\) is a real symmetric matrix.

\[
[X] = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

and the elements of \([A]\) can be found in the following manner.

\[
A_{ii} = a_{ii} \quad \text{for} \quad i = 1, 2, \cdots, n \quad \text{and}
\]

\[
A_{ij} = (a_{ij} + a_{ji})/2, \quad i \neq j.
\]

Let us consider a quadratic expression \(-12x_1^2 + 6x_1x_2 + 2x_2^2 + 5x_1x_3 - 4x_2x_3 + 7x_3^2\).

This expression can be written as:

\[
\begin{bmatrix}
  x_1 & x_2 & x_3
\end{bmatrix}
\begin{bmatrix}
  -12 & 3 & 2.5 \\
  3 & 2 & -2 \\
  2.5 & -2 & 7
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix}
\]

The second step is to find the coordinate system where this quadratic expression can be expressed in self terms only.

**Theorem:** If \([A]\) is a real symmetric matrix then its modal transformation matrix is orthogonal.

Start with the quadratic expression in matrix form, \([X]^T [A] [X]\). Let \([Q]\) be the modal transformation matrix that diagonalizes \([A]\). The change of coordinate is achieved with the transformation, \([X] = [Q] [Y]\). Therefore,

\[
[X]^T [A] [X] = ([Q] [Y])^T [A] ([Q] [Y])
\]

\[
= [Y]^T ([Q]^T [A] [Q]) [Y]
\]

\[
= [Y]^T ([Q]^{-1} [A] [Q]) [Y]
\]

\[
= [Y]^T \begin{bmatrix}
  \lambda_1 & 0 & 0 \\
  0 & \ddots & 0 \\
  0 & 0 & \lambda_n
\end{bmatrix} [Y]
\]

\[
= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2
\]

The result, termed as the *principal axis theorem*, illustrates the application of a new axes with respect to which the quadratic form has a particularly simple appearance. The resulting expression with self terms is termed as the standard form.

Let us continue with our previous quadratic expression which was written in the following matrix form.
\[
\begin{bmatrix}
    x_1^T \\
    x_2^T \\
    x_3^T
\end{bmatrix}
\begin{bmatrix}
    -12 & 3 & 2.5 \\
    3 & 2 & -2 \\
    2.5 & -2 & 7 \\
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{bmatrix}
\]

Note that the coefficient matrix \([A]\) is symmetric.

\[
[A] =
\begin{bmatrix}
    -12 & 3 & 2.5 \\
    3 & 2 & -2 \\
    2.5 & -2 & 7 \\
\end{bmatrix}
\]

The eigenvalues are 2.218, 7.807 and -13.024 and the corresponding modal transformation matrix is:

\[
[Q] =
\begin{bmatrix}
    0.243 & 0.077 & -0.967 \\
    0.934 & -0.289 & 0.212 \\
    0.263 & 0.954 & 0.142 \\
\end{bmatrix}
\]

Note that the transformation matrix is a normalized one.

The resulting standard quadratic form is: \(2.218y_1^2 + 7.807y_2^2 - 13.024y_3^2\).

The coordinate transformation is:

\[
\begin{align*}
    x_1 &= 0.243y_1 + 0.077y_2 - 0.967y_3 \\
    x_2 &= 0.934y_1 - 0.289y_2 + 0.212y_3 \\
    x_3 &= 0.263y_1 + 0.954y_2 + 0.142y_3 \\
\end{align*}
\]

The new axes system is related to the old axes system in the following manner.

\[
\begin{align*}
    y_1 &= 0.243x_1 + 0.934x_2 + 0.263x_3 \\
    y_2 &= 0.077x_1 - 0.289x_2 + 0.954x_3 \\
    y_3 &= -0.967x_1 + 0.212x_2 + 0.142x_3 \\
\end{align*}
\]

In the following two examples we will consider two similar quadratic forms and yet they represent two different conics. The only way we know that is by converting these expressions into their standard forms.

**Example:** Consider the quadratic expression \(x_1^2 - 2x_1x_2 + x_2^2\). We want to transform the expression to its standard form. First write the quadratic form

\[
\begin{align*}
    [x]^{T} [A] [x] &=
    \begin{bmatrix}
        x_1 \\
        x_2
    \end{bmatrix}
    \begin{bmatrix}
        1 & -1 \\
        -1 & 1
    \end{bmatrix}
    \begin{bmatrix}
        x_1 \\
        x_2
    \end{bmatrix}
    \\
    [A] &=
    \begin{bmatrix}
        1 & -1 \\
        -1 & 1
    \end{bmatrix}
\end{align*}
\]
The eigenvalues are 2 and 0. The modal transformation matrix is 
\[
\begin{bmatrix}
0.707 & 0.707 \\
-0.707 & 0.707
\end{bmatrix}
\].

Using \([x] = [Q][y]\), the quadratic expression \([X]^T [A] [X]\) is transformed to its standard form, \(2y_1^2\).

The required transformation in coordinate form is
\[
x_1 = 0.707y_1 + 0.707y_2
\]
\[
x_2 = -0.707y_1 + 0.707y_2
\]

The new axes system with respect to the old one is
\[
y_1 = 0.707x_1 - 0.707x_2
\]
\[
y_2 = 0.707x_1 + 0.707x_2
\]

The equation \(x_1^2 - 2x_1x_2 + x_2^2 = K\) (for \(K > 0\)) represents a conic in the plane. However, such a conic is easier to visualize when put into standard form. So the standard form, \(2y_1^2 = K\) is identified as two straight lines \(y_1 = \sqrt{K}/2\) and \(y_1 = -\sqrt{K}/2\).

![Figure 1. Geometrical interpretation of \(x_1^2 - 2x_1x_2 + x_2^2 = K\)](image)

**Example:** Analyze the conic \(4x_1^2 - 3x_1x_2 + 2x_2^2 = 8\)
This expression is \[ [X]^T [A][X] = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & -1.5 \\ -1.5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 8 \]

\[ [A] = \begin{bmatrix} 4 & -1.5 \\ -1.5 & 2 \end{bmatrix} \]

The eigenvalues are 4.803 and 1.197. Using \([X] = [Q][Y]\), the quadratic expression \([X]^T [A][X]\) is transformed to its standard form. \(4.803y_1^2 + 1.197y_2^2 = 8\).

\[ \frac{y_1^2}{1.666} + \frac{y_2^2}{6.683} = 1 \]

\[ \frac{y_1^2}{(1.291)^2} + \frac{y_2^2}{(2.585)^2} = 1 \]

In its standard form we can easily recognize the conic to be an ellipse. The intercepts with \(y_1\) and \(y_2\) are \pm 1.291 and \pm 2.585 respectively.

The modal transformation matrix (which is orthogonal) is

\[ [Q] = \begin{bmatrix} 0.882 & 0.472 \\ -0.472 & 0.882 \end{bmatrix} \]

Every 2 by 2 orthogonal matrix is a rotation of the plane and \([Q] = \begin{bmatrix} 0.882 & 0.472 \\ -0.472 & 0.882 \end{bmatrix}\) represents a clockwise rotation of about 28.155.

The required transformation in coordinate form is

\[ x_1 = 0.882y_1 + 0.472y_2 \]
\[ x_2 = -0.472y_1 + 0.882y_2 \]

The new axes system with respect to the old one is

\[ y_1 = 0.882x_1 - 0.472x_2 \]
\[ y_2 = 0.472x_1 + 0.882x_2 \]
Exercise problems

Write the following quadratic expressions as $\mathbf{X}^\top \mathbf{A} \mathbf{X}$ and then write them into their standard forms.

1. $x_1^2 + 2x_1x_2 + 6x_2^2$
2. $x_1^2 - 4x_1x_2 + x_2^2$
3. $3x_1^2 - 4x_1x_2 + 3x_2^2 - 3x_1x_3 + 2x_2x_3 + x_3^2$
4. $x_1^2 - x_2^2 - x_1x_3 + 4x_2x_3 + x_3^2$