控制系统的状态空间分析

（第二册）

（日）绪方胜 廖新著 华东工学院译

武汉钢铁学院电气自动化教研室翻印

1979年1月

第三章　坐标变换

3—1　引　言

对于一给定的动态系统，我们可以用很多不同的方法来规定状态矢量。对于一个系统在受到一组对于 $t \geqslant t_0$ 时间中元素均为已知的输入中任何一个输入作用时，为了可能唯一地予测系统在 $t \geqslant t_0$ 任何时刻的行为而必须确定的至少必须的数的集合，就是状态矢量的一组元素，或者说一组状态变量。因此，无论状态矢量是如何规定的，其维数总是一样的。而确定状态变量的不同方法可以看作是系统的坐标变换。

为了深入了解系统的动态性能，我们往往通过坐标变换，使状态变量方程的系数矩阵成为在主对角线上直接显示它的特征值的规范形式，这样做是有好处的。这种规范式称为约当规范式。

约当规范式　如果

$$A = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s \end{bmatrix}$$

则称此 $n \times n$ 矩阵 A 为约当规范式

这里 A_i 是如下形式

$$A_i = \begin{bmatrix} \lambda & 1 & 0 \cdots\cdots 0 & 0 \\ 0 & \lambda & 1 \cdots\cdots 0 & 0 \\ \vdots & \vdots & \vdots \qquad \vdots & \vdots \\ 0 & 0 & 0 \cdots\cdots \lambda & 1 \\ 0 & 0 & 0 \cdots\cdots 0 & \lambda \end{bmatrix}$$

的 $m_i \times m_i$ 矩阵（A_i 中的 λ 和 A_j 中的可以一样也可以不同），并且 $m_1 + m_2 + \cdots\cdots + m_s = n$。矩阵 A 中每一个特征值 λ 在对角线上出现

的次数等于它的重数。例如，在 7 × 7 的矩阵 A 中，如果 $m_1=3$，$m_2=2$，$m_3=1$，$m_4=1$，且A的特征值为 λ_1，λ_1，λ_1，λ_1，λ_1，λ_6，λ_7，那么约当规范式就是

$$
\left[\begin{array}{ccc|cc|c|c}
\lambda_1 & 1 & 0 & & & & 0 \\
0 & \lambda_1 & 1 & & & & \\
0 & 0 & \lambda_1 & & & & \\
\hline
 & & & \lambda_1 & 1 & & \\
 & & & 0 & \lambda_1 & & \\
\hline
 & & & & & \lambda_6 & \\
\hline
 & & & & & & \lambda_7
\end{array}\right]
$$

注意一对角线矩阵就是约当规范式的一种特殊情况。

考虑一线性动态系统

$$\dot{X} = AX$$

其中X是状态矢量（n矢量）A是一 n × n 常数矩阵。对这个系统所定义的任意状态矢量Y与X都有关系式。

$$X = PY$$

其中 P＝ n × n非奇异矩阵。

在对多个对于这个系统所可能定义的状态矢量中最便于研究这个系统的动态性能的是一个特殊的状态矢量Y，它使系统微分方程具有下述的形式：

$$\dot{Y} = P^{-1}APY = JY$$

这里 $J = P^{-1}AP$ 是最简单的可能形式，例如对角线形式或者更一般地为约当规范形式，常常有可能找到一非奇异矩阵P，可以使任意的正方形矩阵A，通过简单的变换 $J = P^{-1}AP$ 变换成约当规范式。矩阵J具有和A相同的不变特性，但是运算更容易。因此将A转换成最简单

的可能形式就有可能用最少的计算工作量去研究动态系统的内在特性。

本章将讨论把系数矩阵A变成约当规范式（它包含了对角线形式作为一种特殊情况）的变换。在3—2节中，我们对矩阵的对角线化提供了初步的讨论。特别是我们将有不同的特征值的矢量矩阵微分方程的系数矩阵对角线化为对角线矩阵的形式。3—3节中我们可以获得一矩阵与一对角线矩阵相似的必要和充分条件。同时提出了与矩阵对角线有关的几条定理。在3—4节中我们将处理多重特征矢量的系统。我们将获得使变换后的方程式的系数矩阵成为约当规范式的变换矩阵。3—5节中简单的讨论了变换矩阵的计算机算法。3—6节给出了使二阶系统的系数矩阵成为约当规范式的变换矩阵的提要。3—7节中我们将处理用n阶标量微分方程描述的系统。我们将首先把它们变换成状态空间方程，然后获得将系统矩阵变换成约当规范式的变换矩阵。最后，在3—8节中我们将讨论线性自由动态系统的运动。

3—2　予先的讨论——有不同特征值的系统

考虑由

$$\dot{X}=AX \qquad X(0)=C$$

描述的系统，如果我们能找到一个合适的非奇异矩阵P，使变换式$X=PY$能够把给定的系统变换成

$$\dot{Y}=P^{-1}APY \qquad Y(0)=P^{-1}C=\xi$$

其中$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots\cdots,\lambda_n)$

那末n阶矢量矩阵微分方程（或者说对应的n个联立的一阶微分方程）可以变换成如下形式的n个一阶微分方程。

$$\dot{y}=\lambda_i y_i \quad y_i(0)=\xi_i \quad (i=1,2,\cdots\cdots,n)$$

这个方程组就可以互相独立的解出来，这样一个过程叫做系统矩阵A的对角线化。

3—4

在3—3节中，我们将证明对于一个有n个不等特征矢量的$n\times n$矩阵A这样的对角线化是可能的。如果我们取A的n个特征矢量作为新的坐标，并且把矢量X转化到矢量Y（如果A包多重特征矢量，那末A就不具有n个线性独立的特征矢量，A不可能与对角线矩阵相似

在这一节中，我们将用一个有n个不等特征值的系统提出矩阵对角线化的或多或少的初步讨论（这样一个系统有n个不等的特征矢量）。

考虑下面一组n个一阶微分方程式：

$$\left.\begin{aligned}\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1n}x_n \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots\cdots + a_{2n}x_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots\cdots + a_{nn}x_n\end{aligned}\right\} \quad (3-1)$$

$x_1(0)=C_1$，$x_2(0)=C_2$，……，$x_n(0)=C_n$

其中a_{ij}都是常数。方程式（3—1）可以写成矢量矩阵微分方程如下：

$$\dot{X}=AX \qquad X(0)=C \qquad (3-2)$$

$$X(t)=\begin{bmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{bmatrix} \qquad A=\begin{bmatrix}a_{11} & a_{12}\cdots\cdots & a_{1n}\\a_{21} & a_{22}\cdots\cdots & a_{2n}\\\vdots & \vdots & \vdots\\a_{n1} & a_{n2} & a_{nn}\end{bmatrix}$$

这里我们假定A是一个具有n个不等特征值λ_1，λ_2，……λ_n的$n\times$常数矩阵。

假定方程式（3—1）或（3—2）的解是

$$\left.\begin{aligned} x_1(t) &= C_1 e^{\lambda t} \\ x_2(t) &= C_2 e^{\lambda t} \\ &\cdots\cdots \\ x_n(t) &= C_n e^{\lambda t} \end{aligned}\right\} \qquad (3-3)$$

或

$$X(t) = Ce^{\lambda t} \qquad (3-4)$$

将方程式（3—3）代入到（3—1）得到

$$\left.\begin{aligned} (a_{11}-\lambda)C_1+a_{12}C_2+\cdots\cdots+a_{1n}C_n &= 0 \\ a_{21}C_1+(a_{22}-\lambda)C_2+\cdots\cdots+a_{2n}C_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}C_1+a_{n2}C_2+\cdots\cdots+(a_{nn}-\lambda)C_n &= 0 \end{aligned}\right\} \qquad (3-5)$$

方程式（3—5）有一个非平凡解，如果也只有如果

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots\cdots & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & & a_{nn}-\lambda \end{vmatrix} = 0 \qquad (3-6)$$

或

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

方程式（3—6）就是特征方程或者说特征值方程，它的根就是特征根或者特征值。

特征矢量　我们将要指出，如果选择一个新的坐标系，每一个坐标是一个特征矢量，那末原来的n个联立方程式中的变量就可以分离。

如果将第i个特征值λ_i代入到（3—5）式，那么就得到

$$(a_{11}-\lambda_i)C_1+a_{12}C_2+\cdots\cdots+a_{1n}C_n = 0$$

$$a_{21}C_1+(a_{22}-\lambda)C_2+\cdots\cdots+a_{2n}C_n = 0$$

$$\cdots\cdots\cdots\cdots\cdots \qquad (3-7)$$

$$a_{n1}C_1+a_{n2}C_2+\cdots\cdots+(a_{nn}-\lambda_i)C_n=0$$

特征值λ_i满足方程式（3—6）。对于这样一种情况，我们不可能决定象C_1，C_2，……C_n那样的常数而只能确定象C_j/C_1那样的比值，在下面方程式

$$\begin{vmatrix} a_{11}-\lambda_i & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22}-\lambda_i & \cdots\cdots & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & \cdots\cdots & a_{nn}-\lambda_i \end{vmatrix}=0 \qquad (3-8)$$

的行列式中，如果我们确定第一行的余子式如A_{i11}，A_{i12}，……A_{i1n}，并且如果把$\lambda_i=\lambda_1$代入到方程式（3—8）和把行列式以第一行的各项展开，那末就得到

$$a_{11}A_{111}+a_{12}A_{112}+\cdots\cdots+a_{1n}A_{11n}=\lambda_1A_{111}$$

如果方程式（3—8）中行列式的第一行用第m行来代替，这里m＝2，3，……，n（行列式的值仍旧等于零），并把结果的行列式以第一行的各项展开，我们将得到

$$a_{m1}A_{111}+a_{m2}A_{112}+\cdots\cdots+a_{mn}A_{11n}=\lambda_1A_{11m}$$

$$(m=2,3,\cdots\cdots,n)$$

这样，对于λ_1我们一共可以得到n个这样的关系。把同样的方法用到任意一个λ_i，我们就可以总共得到下面形式的n^2个方程式

$$a_{j1}A_{i11}+a_{j2}A_{i12}+\cdots\cdots+a_{jn}A_{i1n}=\lambda_iA_{i1j} \qquad (3-9)$$

$$(i,j=1,2,3,\cdots\cdots,n)$$

我们以后将应用方程式（3—9），从这个分析我们可以看出C_1，C_2，……，C_n能够决定如下：

$$\left.\begin{aligned} C_1 &= KA_{i11} \\ C_2 &= KA_{i12} \\ &\cdots\cdots\cdots \\ C_n &= KA_{i1n} \end{aligned}\right\} \quad (3-10)$$

其中K是一个常数。因此，由方程式（3—3）确定的x_k便成为

$$\left.\begin{aligned} x_1(t) &= KA_{i11}e^{\lambda_i t} \\ x_2(t) &= KA_{i12}e^{\lambda_i t} \\ &\cdots\cdots\cdots\cdots \\ x_n(t) &= KA_{i1n}e^{\lambda_i t} \end{aligned}\right\} \quad (3-11)$$

方程式（3—11）对每一个特征值λ_1，λ_2，……，λ_p都适用，这样方程式（3—1）的通解便是

$$\left.\begin{aligned} x_1(t) &= K_1A_{111}e^{\lambda_1 t}+K_2A_{211}e^{\lambda_2 t}+\cdots\cdots K_nA_{n11}e^{\lambda_n t} \\ x_2(t) &= K_1A_{112}e^{\lambda_1 t}+K_2A_{212}e^{\lambda_2 t}+\cdots\cdots+K_nA_{n12}e^{\lambda_n t} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n(t) &= K_1A_{11n}e^{\lambda_1 t}+K_2A_{21n}e^{\lambda_2 t}+\cdots\cdots+K_nA_{n1n}e^{\lambda_n t} \end{aligned}\right\} \quad (3-12)$$

其中K_i都是从初始条件确定的常数。如果初始条件已给出，如

$$\left.\begin{aligned} x_1(0) &= K_1A_{111} \\ x_2(0) &= K_1A_{112} \\ &\cdots\cdots\cdots \\ x_n(0) &= K_1A_{11n} \end{aligned}\right\} \quad (3-13)$$

把方程式（3—13）代入到下面公式，该公式是在方程式（3—12）中令$t=0$得到的：

$$x_1(0) = K_1A_{111}+K_2A_{211}+\cdots\cdots+K_nA_{n11}$$

$$x_2(0) = K_1A_{112}+K_2A_{212}+\cdots\cdots+K_nA_{n12}$$

……………………………………………………

$$x_n(0)=K_1A_{11n}+K_2A_{21n}+\cdots\cdots+K_nA_{n1n}$$

结果得到

$$K_2=K_3=\cdots\cdots=K_n=0$$

这意味着在这样一个初始条件时方程式的解是

$$\left.\begin{aligned}
x_1(t)&=K_1A_{111}e^{\lambda_1 t}\\
x_2(t)&=K_1A_{112}e^{\lambda_1 t}\\
&\cdots\cdots\cdots\cdots\\
x_n(t)&=K_1A_{11n}e^{\lambda_1 t}
\end{aligned}\right\}\qquad(3—14)$$

很清楚，方程式（3—14）决定了下列超直线（在n维状态空间）：

$$\frac{x_1}{A_{111}}=\frac{x_2}{A_{112}}=\cdots\cdots=\frac{x_n}{A_{11n}}\qquad(3—15)$$

（注意，如果$A_{11j}=0$，那末相应的x_j也为零）。方程式（3—15）是对于特征值λ_1获得的，对于n个不等的特征值λ_i（$i=1,2,\cdots n$）一共有n条这样的超直线，这样我们得到

$$\frac{x_1}{A_{i11}}=\frac{x_2}{A_{i12}}=\cdots\cdots=\frac{X_n}{A_{i1n}}\qquad(3—16)$$

$$(i=1,2,\cdots\cdots n)$$

这样的超直线都是特征矢量。特征矢量也叫做主坐标，这里A_{k1j}是方程式（3—8）中的行列式当以λ_k代λ_i时a_{1j}的余子式（有时候我们会遇到$A_{i11}=A_{i12}=\cdots\cdots=A_{i1n}=0$。后面我们将要讨论这种特殊情况）。

注意，对于实数λ_i（$i=1,2,\cdots\cdots,n$）如果初始条件是在对应于λ_i的特征矢量上，那么表示$t>0$时，系统状态的点常常

是在同一特征矢量上（如果λ_i和λ_{i+1}是共轭复数，那末运动$e^{\lambda_i t}$和$e^{\lambda_{i+1} t}$就不能分离，合成的运动就发生在与λ_i和λ_{i+1}相对应的二个特征矢量所决定的平面上。这个运动是振荡的，详细参阅3—8节）。

坐标变换　n维状态空间中的一个点可以用一个矢量X来表示。如果A的n个特征矢量被选择用来作为一组新的坐标，那末在n维状态空间中的同一个点就可以用一个矢量Y来表示。由于特征矢量由方程式（3—16）所决定，只要我们对坐标上选择适当的比例尺，坐标X和坐标Y间的关系可以用下面的方程式来给出：

$$x_1 = A_{11}Y_1 + A_{21}Y_2 + \cdots\cdots A_{n1}Y_n$$

$$x_2 = A_{12}Y_1 + A_{22}Y_2 + \cdots\cdots A_{n2}Y_n$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (3—17)$$

$$x_n = A_{1n}Y_1 + A_{2n}Y_2 + \cdots\cdots + A_{nn}Y_n$$

方制式（3—17）可以重写成下面的形式

$$X = PY$$

这里

$$P = \begin{bmatrix} A_{11} & A_{21} & \cdots\cdots & A_{n1} \\ A_{12} & A_{22} & \cdots\cdots & A_{n2} \\ & & & \\ A_{1n} & A_{2n} & \cdots\cdots & A_{nn} \end{bmatrix} \qquad (3—18)$$

叫做对角线化变换矩阵，或者简单地叫变换矩阵。如果完成了这个坐标

坐标变换，我们就能够证明方程式（3—2）可以写成：

$$\dot{X} = P\dot{Y} = APY$$

或

$$\dot{Y} = P^{-1}APY = DY$$

这里D是一个对角线矩阵，初始条件各项变换成如下形式：

$$Y(0)=P^{-1}X(0)$$

现在我们将说明系数矩阵的对角线化

让我们计算AP·利用方程式（3—9）可获得AP如下：

$$AP=\begin{bmatrix} \lambda_1 A_{111} & \lambda_2 A_{211} \cdots\cdots & \lambda_n A_{n11} \\ \lambda_1 A_{112} & \lambda_2 A_{212} \cdots\cdots & \lambda_n A_{n12} \\ & & \\ \lambda_1 A_{11n} & \lambda_2 A_{21n} \cdots\cdots & \lambda_n A_{n1n} \end{bmatrix}$$

$$=\begin{bmatrix} A_{111} & A_{211} \cdots\cdots & A_{n11} \\ A_{112} & A_{212} \cdots\cdots & A_{n12} \\ & & \\ A_{11n} & A_{21n} \cdots\cdots & A_{n1n} \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_2 & \cdots\cdots & 0 \\ & & & \\ 0 & 0 & \cdots\cdots & \lambda_n \end{bmatrix}=PD$$

（3—19）

这里

$$D=\begin{bmatrix} \lambda_1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_2 & \cdots\cdots & 0 \\ & & & \\ 0 & 0 & & \lambda_n \end{bmatrix}$$

因为 P的n列是线性独立的，P是非奇异的P^{-1}存在，所以

$$P^{-1}AP=D=\mathrm{diag}(\lambda_1, \lambda_2, \cdots\cdots, \lambda_n)$$

利用方程式（3—18）所给出的变换矩阵P，特征值是不等的（这也说明特征矢量是不等的）方程式（3—2）可以转换成

$$\dot{Y}=DY$$

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \vdots \\ \dot{y}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad (3-20)$$

这方程式可以写成下面一组n个一阶微分方程：

$$\dot{y}_1 = \lambda_1 y_1$$

$$\dot{y}_2 = \lambda_2 y_2$$

$$\cdots\cdots\cdots$$

$$\dot{y}_n = \lambda_n y_n$$

初始条件决定于：

$$Y(0) = P^{-1}X(0)$$

方程式（3—20）指出变量 $y_1, y_2, \cdots\cdots, y_n$ 是互相分离的。

方程式（3—20）所代表的系统的特征方程或特征值式程是

$$(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots\cdots(\lambda_n-\lambda) = 0$$

这里特征值没有变化。事实上，对于任何线性非奇异变换特征值都保持不变。

应该指出，如果P是一个对角线化变换矩阵，那末矩阵PC（这里C是一奇非异对角线矩阵）也是一个对角化变换矩阵，因为如果

$$P^{-1}AP = D$$

那末

$$C^{-1}P^{-1}APC = C^{-1}DC = D$$

因此，方程式（3—18）所决定的P并不是唯一的对角线化变换矩阵。P的每一列可以乘一不为零的常数。一般的说，变换矩阵可以写成

$$\begin{bmatrix} K_1A_{111} & K_2A_{211} & \cdots\cdots & K_nA_{n11} \\ K_1A_{112} & K_2A_{212} & \cdots\cdots & K_nA_{n12} \\ \\ K_1A_{11n} & K_2A_{21n} & \cdots\cdots & K_nA_{n1n} \end{bmatrix} \qquad (3-21)$$

这里 K_1，K_2，……，K_n 都是不为零的常数（注意每一列表示一个特征矢量。在 n 维状态空间中特征矢量只决定坐标的方向）。

特殊情况 $A_{i11}=A_{i12}=\cdots\cdots=A_{i1n}=0$，有时候遇到 $A_{i11}=A_{i12}=\cdots\cdots=A_{i1n}=0$，在这种情况下，我们可以取 A_{ik1}，A_{ik2}，……，A_{ikn}（$k\neq1$）组成第 i 个特征矢量（注意在获得方程式（3—9）时，我们是把方程式（3—8）中的行列式对第一行展开。这并不是必须如此的，我们同样可以将行列式对第 k 行（$k\neq1$）展开）考虑下面的例子：

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -1 & 0 \\ 5 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

系数矩阵的三个特征值是 $\lambda_1=-2$，$\lambda_2=-1$，$\lambda_3=1$。让我们对这个例题计算由方程式（3—18）所决定的矩阵：

$$\begin{bmatrix} 1 & 0 & 4 \\ -2 & 0 & 4 \\ -5 & 0 & 10 \end{bmatrix}$$

第二列的元素为零，因此这个矩阵的倒数不存在代替 A_{211}，A_{212}，A_{213}，让我们取 A_{221}，A_{222}，A_{223} 作为第二列的元素（第一和第三列的元素保持不变），那末

$$P=\begin{bmatrix}1 & 0 & 4\\ -2 & 0 & 4\\ -5 & 2 & 10\end{bmatrix}$$

和

$$P^{-1}=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & 0\\ 0 & \frac{5}{4} & \frac{1}{2}\\ \frac{1}{6} & \frac{1}{12} & 0\end{bmatrix}$$

因为这个方法得到的矩阵P可以将A转换到对角线形式·为了检查让我们计算$P^{-1}AP$

$$P^{-1}AP=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & 0\\ 0 & \frac{5}{4} & \frac{1}{2}\\ \frac{1}{6} & \frac{1}{12} & 0\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\ 2 & -1 & 0\\ 5 & 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 4\\ -2 & 0 & 4\\ -5 & 2 & 10\end{bmatrix}=\begin{bmatrix}-2 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

对角线化变换矩阵P的计算（应用于当A有不等的特征值时）应该指出，在有不等特征值的自由动态系统中，只要系统的特征值已知时，为了获得对角线形式，没有必要再去寻找矩阵P·从所知的特征值就可以把对角线形式直接写出·在强迫的动态系统中，我们需要P^{-1}（参看第四章）·下面我们将讨论二种寻找P的方法·

应该指出，如果A有下面二种形式中的一种：

$$A=\begin{bmatrix}0 & 1 & 0 & \cdots\cdots & 0 & 0\\ 0 & 0 & 1 & \cdots\cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots\cdots & 0 & 1\\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_2 & -a_1\end{bmatrix}$$

3—14 .

或

$$A=\begin{bmatrix} -a_1 & 1 & 0\cdots\cdots 0 & 0 \\ -a_2 & 0 & 1\cdots\cdots 0 & 0 \\ \vdots & \vdots & \vdots \quad\quad \vdots & \vdots \\ -a_{n-1} & 0 & 0\cdots\cdots 0 & 1 \\ -a_n & 0 & 0\cdots\cdots 0 & 0 \end{bmatrix}$$

那么可以用简单的公式获得对角线化变换矩阵P（参看方程式（3—37），方程式（3—50）和方程式（3—52）；和习题A—3—10和习题A—3—11）

方法1 一个对角线化变换矩阵，可以通过计算A_{1ij}直接得到。作为例子，考虑下面矩阵：

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

A的特征值是1和2，矩阵P由

$$P=\begin{bmatrix} A_{111} & A_{211} \\ A_{112} & A_{212} \end{bmatrix}$$

给出。为了找出A_{111}和A_{112}让我们将$\lambda_1=1$代入到特征多项式，我们得到

$$|A-\lambda_1 I|=\begin{vmatrix} 3 & 2 \\ -3 & -2 \end{vmatrix}$$

因此，$A_{111}=-2$，$A_{112}=-(-3)$

为了找出A_{211}和A_{212}，让我们将$\lambda_2=2$代入到特征多项式

$$|A-\lambda_2 I|=\begin{vmatrix} 2 & 2 \\ -3 & -3 \end{vmatrix}$$

我们得到$A_{211}=-3$，　　$A_{212}=-(-3)$

所以我们找出P为：

$$P=\begin{pmatrix}-2 & -3\\ 3 & 3\end{pmatrix}$$

然后我们得到

$$P^{-1}AP=\begin{pmatrix}1 & 1\\ -1 & -\frac{2}{3}\end{pmatrix}\begin{pmatrix}4 & 2\\ -3 & -1\end{pmatrix}\begin{pmatrix}-2 & -3\\ 3 & 3\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

求取使2×2常数矩阵和3×3常数矩阵转换成对角线形式的对角线化变换矩阵P的公式分别给出如下：

(1)　对于如下的2×2常数矩阵A

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

其中　$|A-\lambda I|=(\lambda-\lambda_1)(\lambda-\lambda_2)$　　$\lambda_1\neq\lambda_2$

将A转换成对角线矩阵的对角线化变换矩阵P为

$$P=\begin{pmatrix}A_{111} & A_{211}\\ A_{112} & A_{212}\end{pmatrix}=\begin{pmatrix}d-\lambda_1 & d-\lambda_2\\ -C & -C\end{pmatrix}$$

(2)对于如下的3×3常数矩阵A：

$$A=\begin{pmatrix}a & b & c\\ d & e & f\\ g & h & i\end{pmatrix}$$

其中　$|A-\lambda I|=-(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$

并且λ_1，λ_2和λ_3互不相等，将A转换成对角线矩阵的对角线化变换矩阵P为：

$$P=\begin{bmatrix} A_{111} & A_{211} & A_{311} \\ A_{112} & A_{212} & A_{312} \\ A_{113} & A_{213} & A_{313} \end{bmatrix}$$

$$=\begin{bmatrix} \begin{vmatrix} e-\lambda_1 & f \\ h & i-\lambda_1 \end{vmatrix} & \begin{vmatrix} e-\lambda_2 & f \\ h & i-\lambda_2 \end{vmatrix} & \begin{vmatrix} e-\lambda_3 & f \\ h & i-\lambda_3 \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i-\lambda_1 \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i-\lambda_2 \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i-\lambda_3 \end{vmatrix} \\ \begin{vmatrix} d & e-\lambda_1 \\ g & h \end{vmatrix} & \begin{vmatrix} d & e-\lambda_2 \\ g & h \end{vmatrix} & \begin{vmatrix} d & e-\lambda_3 \\ g & h \end{vmatrix} \end{bmatrix}$$

方法2，让我们写 P

$$P=\begin{bmatrix} P_{11} & P_{12} & \cdots\cdots & P_{1n} \\ P_{21} & P_{22} & \cdots\cdots & P_{2n} \\ & & & \\ P_{n1} & P_{n2} & \cdots\cdots & P_{nn} \end{bmatrix}$$

由于 P 的每一列是 A 的一个特征矢量，我们就有

$$AP_i=\lambda_i P_i$$

这里 P_i 是矩阵 P 的第 i 列。为了计算列矩阵 P_i 的元素，我们从方程式

$$(A-\lambda_i I)P_i=0 \qquad (i=1, 2, \cdots\cdots, n)$$

解出 P_{1i}，P_{2i}，……，P_{ni}

例如，考虑如方法1的例子中的同一矩阵 A 或者说

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

A 的特征值是1和2，为了获得对角线化变换矩阵 P，我们解一个方