Characteristic Polynomial

CALIFORNIA STATE UNIVERSITY EAST BAY

Massoud Malek

♣ **Preliminary Results**. In all that follows, we denote the $n \times n$ identity matrix by I_n and the $n \times n$ zero matrix by Z_n.

Let $A = (a_{ij})$ be an $n \times n$ matrix. If $A\,u = \lambda\,u$, then λ and u are called the *eigenvalue* and *eigenvector* of A, respectively. The eigenvalues of A are the roots of the *characteristic polynomial*

$$K_A(\lambda) = \det(\lambda\,I_n - A).$$

The eigenvectors are the solutions to the *Homogeneous system* $(\lambda\,I_n - A)\,X = \theta$.

Note that $K_A(\lambda)$ is a *monic polynomial* (i.e., the leading coefficient is one).

Cayley-Hamilton Theorem. *If* $K_A(\lambda) = \lambda^n + p_1\,\lambda^{n-1} + \cdots + p_{n-1}\,\lambda + p_n$ *is the characteristic polynomial of the* $n \times n$ *matrix* A*, then*

$$K_A(A) = A^n + p_1 A^{n-1} + \cdots + p_{n-1}A + p_n I_n = Z_n,$$

Corollary. *Let* $K_A(\lambda) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$ *be the characteristic polynomial of the* $n \times n$ *invertible matrix* A*. Then*

$$A^{-1} = \frac{1}{-p_n}\left[A^{n-1} + p_1 A^{n-2} + \cdots + p_{n-2}A + p_{n-1}I_n\right].$$

Proof. *According to the Cayley Hamilton's theorem we have*

$$A\,\left[\,A^{n-1} + p_1\,A^{n-2} + \cdots + p_{n-1}\,I_n\,\right] = -p_n\,I_n\,.$$

Since A *is nonsingular,* $p_n = (-1)^n \det(A) \neq 0$*; thus the result follows.* □

Newton's Identity. *Let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be the roots of the polynomial*

$$K(\lambda) = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \cdots\ \cdots + p_{n-1}\lambda + p_n.$$

If $s_k = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k$*, then*

$$p_k = -\frac{1}{k}\left(s_k + s_{k-1}\,p_1 + s_{k-2}\,p_2 + \cdots + s_2\,p_{k-2}p_1 + s_1\,p_{k-1}\right).$$

Proof. *From* $K(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots\ \ldots(\lambda - \lambda_{n-1})(\lambda - \lambda_n)$ *and the use of logarithmic differentiation, we obtain*

$$\frac{K'(\lambda)}{K(\lambda)} = \frac{n\,\lambda^{n-1} + (n-1)\,c_1\lambda^{n-2} + \cdots + 2\,c_{n-2}\lambda + c_{n-1}}{\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n} = \sum_{i=1}^{n}\frac{1}{(\lambda - \lambda_i)}\,.$$

By using the geometric series for $\dfrac{1}{(\lambda - \lambda_i)}$ *and choosing* $|\lambda| > \max\limits_{1 \le i \le n} |\lambda_i|$*, we obtain*

$$\sum_{i=1}^{n}\frac{1}{(\lambda - \lambda_i)} = \frac{n}{\lambda} + \frac{s_1}{\lambda^2} + \frac{s_2}{\lambda^3} + \cdots\cdots$$

Hence $n\,\lambda^{n-1} + (n-1)\,p_1\,\lambda^{n-2} + \cdots + p_{n-1} = \left(\lambda^n + p_1\,\lambda^{n-1} + p_2\,\lambda^{n-2} + \cdots + p_n\right)\left(\dfrac{n}{\lambda} + \dfrac{s_1}{\lambda^2} + \dfrac{s_2}{\lambda^3} + \cdots\right).$

By equating both sides of the above equality we may obtain the Newton's identities. □

♣ **The Method of Direct Expansion**. *The characteristic polynomial of an* $n \times n$ *matrix* $A = (a_{ij})$ *is defined as:*

$$K_A(\lambda) = \det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix} = \lambda^n - \sigma_1 \lambda^{n-1} + \sigma_2 \lambda^{n-2} - \dots + (-1)^n \sigma_n,$$

where

$$\sigma_1 = \sum_{i=1}^{n} a_{ii} = trace(A)$$

is the sum of all first-order diagonal minors of A*,*

$$\sigma_2 = \sum_{i<j} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$$

is the sum of all second-order diagonal minors of A*,*

$$\sigma_3 = \sum_{i<j<k} \begin{vmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{vmatrix}$$

is the sum of all third-order diagonal minors of A*, and so forth. Finally,*

$$\sigma_n = det(A)$$

There are $\binom{n}{k}$ *diagonal minors of order* k *in* A*. From this we find that the direct computation of the coefficients of the characteristic polynomial of an* $n \times n$ *matrix is equivalent to computing*

$$\binom{n}{1} + \binom{n}{2} + \cdots \binom{n}{n} = 2^n - 1$$

determinants of various orders, which, generally speaking, is a major task. This has given rise to special methods for expanding characteristic polynomial. We shall explain some of these methods.

Example. *Compute the characteristic polynomial of* $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & -4 \\ 1 & 0 & 2 \end{pmatrix}$.

We have:

$$\sigma_1 = 1 + 1 + 2 = 4,$$
$$\sigma_2 = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} + \begin{vmatrix} 1 & -4 \\ 0 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = (-3) + (2) + (-1) = -2,$$
$$\sigma_3 = \det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & -4 \\ 1 & 0 & 2 \end{vmatrix} = -17.$$

Thus

$$K_A(\lambda) = \det(\lambda I_3 - A) = \lambda^3 - \sigma_1 \lambda^2 + \sigma_2 \lambda - \sigma_3 = \lambda^3 - 4\lambda - 2\lambda + 17.$$

♣ **Leverrier's Algorithm**. *This method allows us to find the characteristic polynomial of any $n \times n$ matrix A using the trace of the matrix A^k, where $k = 1, 2, \cdots n$. Let*

$$\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$$

be the set of all eigenvalues of A which is also called the spectrum of A. Note that

$$s_k = trace(A^k) = \sum_{i=1}^{n} \lambda_i^k \ , \ \textit{for all } k = 1, 2, \cdots, n.$$

Let

$$K_A(\lambda) = \det(\lambda I_n - A) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

be the characteristic polynomial of the matrix A, then for $k \leq n$, the Newton's identities hold true:

$$p_k = -\frac{1}{k}\left[s_k + p_1 s_{k-1} + \cdots + p_{k-1}s_1\right] \quad (k = 1, 2, \cdots, n)$$

Example. *Let* $A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix}$. *Then*

$$A^2 = \begin{pmatrix} 1 & 8 & 4 & 0 \\ 9 & -1 & -1 & 9 \\ 13 & -12 & 5 & 8 \\ 15 & -12 & -6 & 7 \end{pmatrix} \quad A^3 = \begin{pmatrix} 17 & 6 & 13 & 19 \\ 42 & -28 & 8 & 23 \\ 43 & -9 & -16 & 22 \\ 19 & -11 & -3 & -17 \end{pmatrix} \quad A^4 = \begin{pmatrix} 125 & -48 & 16 & 104 \\ 122 & -23 & -22 & 46 \\ 90 & -40 & 41 & -12 \\ -66 & 120 & 0 & -107 \end{pmatrix}.$$

So $s_1 = 4$, $s_2 = 12$, $s_3 = -44$, *and* $s_4 = 36$. *Hence*

$$\begin{aligned}
p_1 &= -s_1 = -4, \\
p_2 &= -\frac{1}{2}(s_2 + p_1 s_1) = -\frac{1}{2}(12 + (-4)4) = 2, \\
p_3 &= -\frac{1}{3}(s_3 + p_1 s_2 + p_2 s_1) = -\frac{1}{3}(-44 + (-4)12 + 2(4)) = 28, \\
p_4 &= -\frac{1}{4}(s_4 + p_1 s_3 + p_2 s_2 + p_3 s_1) = -\frac{1}{4}(36 + (-4)(-44) + 2(12) + 28(4)) = -87.
\end{aligned}$$

Therefore

$$K_A(\lambda) = \lambda^4 - 4\lambda^3 + 2\lambda^2 + 28\lambda - 87$$

and

$$\begin{aligned}
A^{-1} &= \frac{1}{87}\left[A^3 - 4\,A^2 + 2\,A + 28\,I_4\right] \\
&= \frac{1}{87}\left[\begin{pmatrix} 17 & 6 & 13 & 19 \\ 42 & -28 & 8 & 23 \\ 43 & -9 & -16 & 22 \\ 19 & -11 & -3 & -17 \end{pmatrix} - 4\begin{pmatrix} 1 & 8 & 4 & 0 \\ 9 & -1 & -1 & 9 \\ 13 & -12 & 5 & 8 \\ 15 & -12 & -6 & 7 \end{pmatrix} + 2\begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix} + 28\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right] \\
&= \frac{1}{87}\begin{pmatrix} 43 & -22 & -1 & 17 \\ 8 & 4 & 16 & -11 \\ -5 & 41 & -10 & -4 \\ -33 & 27 & 21 & -9 \end{pmatrix}.
\end{aligned}$$

♣ **The Method of Souriau (or Fadeev and Frame)**. *This is an elegant modification of the Leverrier's method.*
Let A be an $n \times n$ matrix, then define

$$\begin{array}{rcl} A_1 = A, & q_1 = -trace(A_1), & B_1 = A_1 + q_1 I_n; \\ A_2 = A\,B_1, & q_2 = -\frac{1}{2} trace(A_2), & B_2 = A_2 + q_2 I_n; \\ \vdots \ \vdots \ \vdots & \vdots \ \vdots \ \vdots \ \vdots & \vdots \ \vdots \ \vdots \\ A_{n-1} = A\,B_{n-2}, & q_{n-1} = -\frac{1}{n-1} trace(A_{n-1}), & B_{n-1} = A_{n-1} + q_{n-1} I_n; \\ A_n = A\,B_{n-1}, & q_n = -\frac{1}{n} trace(A_n), & B_n = A_n + q_n I_n \end{array} .$$

Theorem. *$B_n = Z_n$, and*

$$K_A(\lambda) = \det(\lambda I_n - A) = \lambda^n + q_1 \lambda^{n-1} + \cdots + q_{n-1}\lambda + q_n\,,$$

where $Adj\,(A) = B_{n-1}$ and if A is nonsingular, then

$$A^{-1} = -\frac{1}{q_n} B_{n-1}\,.$$

Proof. Suppose the characteristic polynomial of A is

$$K_A(\lambda) = \det(\lambda I_n - A) = \lambda^n + p_1 \lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n\,,$$

where $p_k's$ are defined in the Leverrier's method.
Clearly $p_1 = -trace(A) = -trace(A_1) = q_1$, and now suppose that we have proved that

$$q_1 = p_1, q_2 = p_2, \ \ldots \ , q_{k-1} = p_{k-1}\,.$$

Then by the hypothesis we have

$$\begin{aligned} A_k = AB_{k-1} &= A(A_{k-1} + q_{k-1} I_n) = AA_{k-1} + q_{k-1} A \\ &= A[A(A_{k-2} + q_{k-2} I_n)] + q_{k-1} A \\ &= A^2 A_{k-1} + q_{k-2} A^2 + q_{k-1} A \\ &= \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ &= A^k + q_1 A^{k-1} + \ \cdots \ + q_{k-1} A\,. \end{aligned}$$

Let $s_i = trace(A^i)$ $(i = 1, 2, \ldots, k)$, then by *Newton's identities*

$$\begin{aligned} -kq_k = trace(A_k) &= trace(A^k) + q_1\ trace(A^{k-1}) + \cdots + q_{k-1}\ trace(A) \\ &= s_k + q_1 s_{k-1} + \cdots + q_{k-1} s_1 \\ &= s_k + p_1 s_{k-1} + \cdots + p_{k-1} s_1 \\ &= -kp_k\,. \end{aligned}$$

showing that $p_k = q_k$. Hence this relation holds for all k.
By the Cayley-Hamilton theorem,

$$B_n = A^n + q_1 A^{n-1} + \cdots + q_{n-1} A + q_n I_n = Z_n\,.$$

and so

$$B_n = A_n + q_n I_n = Z_n; \qquad A_n = AB_{n-1} = -q_n I_n.$$

If A is nonsingular, then $\det(A) = (-1)^n K_A(0) = (-1)^n q_n \neq 0$, and thus

$$A^{-1} = -\frac{1}{q_n} B_{n-1}\,. \qquad \square$$

Example. Find the characteristic polynomial and if possible the inverse of the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix}.$$

For $k = 1, 2, 3, 4$, compute

$$A_k = A\,B_{k-1} \qquad q_k = \frac{-1}{k} trace(A_k), \quad B_k = A_k + q_k I_4\,.$$

$$A_1 = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix}, \qquad q_1 = -4, \qquad B_1 = \begin{pmatrix} -3 & 2 & 1 & -1 \\ 1 & -4 & 2 & 1 \\ 2 & 1 & -5 & 3 \\ 4 & -5 & 0 & 0 \end{pmatrix};$$

$$A_2 = \begin{pmatrix} -3 & 0 & 0 & 4 \\ 5 & -1 & -9 & 5 \\ 5 & -16 & 9 & -4 \\ -1 & 8 & -6 & -9 \end{pmatrix}, \qquad q_2 = 2, \qquad B_2 = \begin{pmatrix} -1 & 0 & 0 & 4 \\ 5 & 1 & -9 & 5 \\ 5 & -16 & 11 & -4 \\ -1 & 8 & -6 & -7 \end{pmatrix};$$

$$A_3 = \begin{pmatrix} 15 & -22 & -1 & 17 \\ 8 & -24 & 16 & -11 \\ -5 & 41 & -38 & -4 \\ -33 & 27 & 21 & -37 \end{pmatrix}, \quad q_3 = 28, \quad B_3 = \begin{pmatrix} 43 & -22 & -1 & 17 \\ 8 & 4 & 16 & -11 \\ -5 & 41 & -10 & -4 \\ -33 & 27 & 21 & -9 \end{pmatrix};$$

$$A_4 = \begin{pmatrix} 87 & 0 & 0 & 0 \\ 0 & 87 & 0 & 0 \\ 0 & 0 & 87 & 0 \\ 0 & 0 & 0 & 87 \end{pmatrix}, \qquad q_4 = -87, \qquad B_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore the characteristic polynomial of A is:

$$K_A(\lambda) = \lambda^4 - 4\lambda^3 + 2\lambda^2 + 28\lambda - 87.$$

Note that A_4 is a diagonal matrix, so we only need to multiply the first row of A by the first column of B_3 to obtain 87. Since $q_4 = -87$, the matrix A has an inverse.

$$A^{-1} = \frac{-1}{q_4} B_3 = \frac{1}{87} \begin{pmatrix} 43 & -22 & -1 & 17 \\ 8 & 4 & 16 & -11 \\ -5 & 41 & -10 & -4 \\ -33 & 27 & 21 & -9 \end{pmatrix}. \qquad \square$$

Matlab Program

$A = input('Enter\ a\ square\ matrix\ :\ ')$
$m = size(A);\ n = m(1);\ q = zeros(1, n);\ B = A;\ AB = A;\ In = eye(n);$
$for\ k = 1 : n-1,\ q(k) = -(1/k) * trace(AB)B = AB + q(k) * In;\ AB = A * B;\ end$
$C = B;\ q(n) = -(1/n) * trace(AB);\ Q = [1\ q];$
$disp('The\ Characteristic\ polynomial\ looks\ like\ :\ ')$
$disp(\ 'K_A(x) = x \wedge n + q(1)x \wedge (n-1) + ... + q(n-1)x + q(n)'\ ),\ disp('\ '),$
$disp('The\ coefficients\ list\ c(k)\ is\ :\ '),\ disp('\ '),$
$disp(Q),\ disp('\ ')$
$if\ q(n) == 0,\ disp('The\ matrix\ is\ singular\ ');$
$else,\ disp('The\ matrix\ has\ an\ inverse.\ '),\ disp('\ ')$
  $C = -(1/q(n)) * B;$
  $disp('The\ inverse\ of\ A\ is\ :\ '), disp('\ '),$
  $disp(C)$
end $\square$

♣ **<u>The Method of Undetermined Coefficients</u>**. If one has to expand large numbers of characteristic polynomials of matrices of the same order, then the method of undetermined coefficients may be used to produce characteristic polynomials of those matrices.

Let A be an $n \times n$ matrix and

$$K_A(\lambda) = \det(\lambda I_n - A) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n\,.$$

be its characteristic polynomial. In order to find the coefficients p_i's of $K_A(\lambda)$ we evaluate

$$D_j = K_A(j) = \det(jI_n - A) \quad j = 0, 1, 2, \ldots, n-1$$

and obtain the following system of linear equations:

$$\begin{aligned} p_n &= D_0 \\ 1^n + p_1.1^{n-1} + \cdots\cdots\cdots\cdots\cdots + p_n &= D_1 \\ 2^n + p_1.2^{n-1} + \cdots\cdots\cdots\cdots\cdots + p_n &= D_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (n-1)^n + p_1.(n-1)^{n-1} + \cdots + p_n &= D_{n-1} \end{aligned}$$

Which can be changed into:

$$S_{n-1}P = \begin{bmatrix} 1^{n-1} & 1^{n-2} & \ldots & 1 \\ 2^{n-1} & 2^{n-2} & \ldots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ (n-1)^{n-1} & (n-1)^{n-2} & \ldots & n-1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \end{bmatrix} = \begin{bmatrix} D_1 - D_0 - 1^n \\ D_2 - D_0 - 2^n \\ \vdots \; \vdots \; \vdots \; \vdots \\ D_{n-1} - D_0 - (n-1)^n \end{bmatrix} = D\,.$$

The system may be solved as follows:

$$P = S_n^{-1} D\,.$$

Since the $(n-1)\times(n-1)$ matrix S_n depends only on the order of A, we may store R_n, the inverse of S_{n-1} beforehand and use it to find the coefficients of characteristic polynomial of various $n \times n$ matrices.

Examples. Compute the characteristic polynomials of the 4×4 matrices

$$A = \begin{pmatrix} 1 & 3 & 0 & 4 \\ 2 & -3 & 1 & 3 \\ 1 & 2 & 1 & 2 \\ -1 & 3 & 2 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix}.$$

First we find

$$S = \begin{pmatrix} 1 & 1 & 1 \\ 8 & 4 & 2 \\ 27 & 9 & 3 \end{pmatrix} \quad \text{and} \quad R = S^{-1} = -\frac{1}{12}\begin{pmatrix} -6 & 6 & -2 \\ 30 & -24 & 6 \\ -36 & 18 & -4 \end{pmatrix}.$$

Then for the matrix A we obtain

$$\begin{aligned} &D_0 = \det(-A) = -48, \quad D_1 = \det(I_4 - A) = -72, \\ &D_2 = \det(2I_4 - A) = -128, \quad \text{and} \quad D_3 = \det(3I_4 - A) = -180\,. \end{aligned}$$

$$D = \begin{pmatrix} D_1 - D_0 - 1^4 \\ D_2 - D_0 - 2^4 \\ D_3 - D_0 - 3^4 \end{pmatrix} = \begin{pmatrix} -25 \\ -96 \\ -213 \end{pmatrix}.$$

Hence

$$P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = -\frac{1}{12}\begin{pmatrix} -6 & 6 & -2 \\ 30 & -24 & 6 \\ -36 & 18 & -4 \end{pmatrix}\begin{pmatrix} -25 \\ -96 \\ -213 \end{pmatrix} = \begin{pmatrix} 0 \\ -23 \\ -2 \end{pmatrix}.$$

Thus

$$K_A(\lambda) = \lambda^4 - 23\lambda^2 - 2\lambda - 48.$$

For the matrix B we have

$$D_0 = \det(-B) = -87, \quad D_1 = \det(I_4 - B) = -60,$$
$$D_2 = \det(2I_4 - B) = -39, \quad \text{and} \quad D_3 = \det(3I_4 - B) = -12.$$

$$D = \begin{pmatrix} D_1 - D_0 - 1^4 \\ D_2 - D_0 - 2^4 \\ D_3 - D_0 - 3^4 \end{pmatrix} = \begin{pmatrix} 26 \\ 32 \\ -6 \end{pmatrix}.$$

Hence

$$P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = -\frac{1}{12} \begin{pmatrix} -6 & 6 & -2 \\ 30 & -24 & 6 \\ -36 & 18 & -4 \end{pmatrix} \begin{pmatrix} 26 \\ 32 \\ -6 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 28 \end{pmatrix}.$$

Thus

$$K_B(\lambda) = \lambda^4 - 4\lambda^3 + 2\lambda^2 + 28\lambda - 87.$$

Matlab Program

```
N = input('Enter the size of your square matrix : ');
n = N − 1; In = eye(N); S = zeros(n); R = zeros(n); D = zeros(1, n);
DSP1 = [' For any ', int2str(N),' −square matrix, you need S =' ];
DSP2 = ['Do you want to try with another ', int2str(N),' −square matrix? (Yes = 1/No = 0)' ];
%DEFINING S
for i = 1 : n, for j = 1 : n, S(i, j) = i ∧ (N − j); end; end;
disp(' '), disp(DSP1), disp(' '), disp(S),
R = inv(S);
ok = 1;
while ok == 1;
  A = input(['Enteran ', int2str(N),' x ', int2str(N),' matrix A : ');   disp(' ')
  D0 = det(A);
  for k = 1 : n; D(k) = det(k ∗ In − A); end;
  for i = 1 : n; DD(i) = D(i) − D0 − i ∧ N; end;
  P = R ∗ DD';
  disp('The Characteristic polynomial looks like : ')
  disp( 'K_A(x) = x ∧ n + p(1)x ∧ (n − 1) + ... + p(n − 1)x + p(n)' ), disp(' '),
  disp('The coefficients list p(k) is : '), disp(' '),
  disp([1 P' D0]), disp(' '),
  disp(DSP2), disp(' '),
  ok = input(DSP2);
end
```

□

♣ **The Method of Danilevsky**. Consider an $n \times n$ matrix A and let

$$K_A(\lambda) = det(\lambda I_n - A) = \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

be its characteristic polynomial. Then the **companion matrix** of $K_A(\lambda)$

$$F[A] = \begin{pmatrix} -p_1 & -p_2 & -p_3 & \dots & -p_{n-1} & -p_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

is similar to A and is called the *Frobenius form* of A.

The method of Danilevsky (1937) applies the Gauss-Jordan method to obtain the Frobenius form of an $n \times n$ matrix. According to this method the transition from the matrix A to $F[A]$ is done by means of $n-1$ similarity transformations which successively transform the rows of A, beginning with the last, into corresponding rows of $F[A]$.

Let us illustrate the beginning of the process. Our purpose is to carry the *nth* row of

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2,n-1} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}$$

into the row $\begin{pmatrix} 0 & 0 & \dots & 1 & 0 \end{pmatrix}$. Assuming that $a_{n,n-1} \neq 0$, we replace the $(n-1)th$ row of the $n \times n$ identity matrix with the *nth* row of A and obtained the matrix

$$U_{n-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} & a_{nn} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

The inverse of U_{n-1} is

$$V_{n-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots & \vdots \\ v_{n-1,1} & v_{n-1,2} & v_{n-1,3} & \dots & v_{n-1,n-1} & v_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

where

$$v_{n-1,i} = -\frac{a_{ni}}{a_{n,n-1}} \qquad \text{for} \quad i \neq n-1$$

and

$$v_{n-1,n-1} = \frac{1}{a_{n,n-1}} \ .$$

Multiplying the right side of A by V_{n-1}, we obtain

$$AV_{n-1} = B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1,n-1} & b_{1n} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2,n-1} & b_{2n} \\ \vdots & \vdots & \ddots & \dots & \vdots & \vdots \\ b_{n-1,1} & b_{n-1,2} & b_{n-1,3} & \dots & b_{n-1,n-1} & b_{n-1,n-1} \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

However the matrix $B = A\,M_{n-1}$ is not similar to A. To have a similarity transformation, it is necessary to multiply the left side of B by $U_{n-1} = V_{n-1}^{-1}$. Let $C = U_{n-1}AV_{n-1}$, then C is similar to A and is of the form

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1,n-1} & c_{1n} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2,n-1} & b_{2n} \\ \vdots & \vdots & \ddots & \dots & \vdots & \vdots \\ c_{n-1,1} & c_{n-1,2} & c_{n-1,3} & \dots & c_{n-1,n-1} & c_{n-1,n-1} \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Now, if $c_{n-1,n-1} \neq 0$, then similar operations are performed on matrix C by taking its $(n-2)th$ row as the principal one. We then obtain the matrix

$$D = U_{n-2}CV_{n-2} = U_{n-2}U_{n-1}AV_{n-1}V_{n-2}$$

with two reduced rows. We continue the same way until we finally obtain the Frobenius form

$$F[A] = U_1U_2\cdots U_{n-2}U_{n-1}AV_{n-1}V_{n-2}\cdots V_2V_1\,.$$

if, of course, all the $n-1$ intermediate transformations are possible.

Exceptional case in the Danilevsky method. Suppose that in the transformation of the matrix A into its Frobenius form $F[A]$ we arrived, after a few steps, at a matrix of the form

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} & \dots & r_{1,n-1} & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2k} & \dots & r_{2,n-1} & r_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{k1} & r_{k2} & \dots & r_{kk} & \dots & r_{k,n-1} & r_{kn} \\ 0 & 0 & \dots & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & 0 \end{pmatrix}$$

and it was found that $r_{k,k-1} = 0$ or $|r_{k,k-1}|$ is very small. It is then possible to continue the transformation by the Danilevsky method.

Two cases are possible here.

Case 1. Suppose for some $j = 1, 2, \dots, k-2$, $r_{kj} \neq 0$. Then by permuting the jth row and $(k-1)th$ row and the jth column and $(k-1)th$ column of R we obtain a matrix $R' = (r'_{ij})$ similar to R with $r'_{k,k-1} \neq 0$.

Case 2. Suppose now that $r_{kj} = 0$ for all $j = 1, 2, \dots, k-2$. Then R is in the form

$$R = \begin{bmatrix} R_1 & R_2 \\ 0 & R_3 \end{bmatrix} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1,k-1} & r_{1,k} & r_{1,k+1} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2,k-1} & r_{2,k} & r_{2,k+1} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{k-1,1} & r_{k-1,2} & \dots & r_{k-1,k-1} & r_{k-1,k} & r_{k-1,k+1} & \dots & r_{k-1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & r_{kk} & r_{k,k+1} & \dots & r_{kn} \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

In this case the characteristic polynomial of R breaks up into two determinants:

$$\det(\lambda I_n - R) = \det(\lambda I_{k-1} - R_1)\det(\lambda I_{n-k+1} - R_3).$$

Here, the matrix R_3 is already reduced to the Frobenius form. It remains to apply the Danilevsky's method to the matrix R_1.

Note. Since $U_k\, A_{k-1}$ only changes the kth row of A_{k-1}, it is more efficient to multiply first A_{k-1} by its $(k+1)th$ row and then multiply on the right side the resulting matrix by V_k.

The next result shows that once we transform A into its Frobenius form ; we may obtain the eigenvectors with the help of the matrices $V_i' s$.

Theorem. *Let A be an $n \times n$ matrix and let $F[A]$ be its Frobenius form. If λ is an eigenvalue of A, then*

$$v = \begin{pmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \\ 1 \end{pmatrix} \qquad \text{and} \qquad w = V_{n-1} V_{n-2} \cdots V_2 V_1 v$$

are the eigenvectors of $F[A]$ and A respectively.

Proof. Since

$$\det(\lambda I_n - A) = \det(\lambda I_n - F[A]\,) = \lambda^n + p_1 \lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n\,,$$

we have

$$(\lambda I_n - F[A])v = \begin{pmatrix} \lambda - p_1 & -p_2 & -p_3 & \ldots & -p_{n-1} & -p_n \\ 1 & \lambda & 0 & \ldots & 0 & 0 \\ \lambda & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & \lambda \end{pmatrix} \begin{pmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Since $F[A] = V_1^{-1} V_2^{-1} \cdots V_{n-2}^{-1} V_{n-1}^{-1} A V_{n-1} V_{n-2} \cdots V_2 V_1$ and $F[A]v = \lambda v$, we conclude that

$$\lambda w = V_{n-2} \cdots V_2\, V_1(\lambda v) = (V_{n-2} \cdots V_2\, V_1)\ F[A]v = A\,(V_{n-1} V_{n-2} \cdots V_2\, V_1 v) = Aw \qquad \square$$

Note. For expanding characteristic polynomials of matrices of order higher than fifth, the method of Danilevsky requires less multiplications and additions than other methods.

Example. Reduce the matrix

$$A = \begin{pmatrix} 1 & 1 & 3 & 4 \\ 2 & 0 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & -1 & -1 \end{pmatrix}$$

to its Frobenius form.

The matrix $B = A_3 = U_3 A V_3$ is as follows:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 & 4 \\ 2 & 0 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -3 & 1 \\ 2 & 0 & -2 & -1 \\ -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Since $b_{32} = 0$, we need the permutation matrix $J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$; thus

$$C = J\,B\,J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -3 & 1 \\ 2 & 0 & -2 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -2 & -1 \\ 1 & 1 & -3 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Next we obtain the matrix $D = A_2 = U_2 C\, V_2$

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 & -1 \\ 1 & 1 & -3 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -2 & -3 \\ -1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Finally the Frobenius form $F[A] = A_1 = U_1 D V_1$,

$$F[A] = \begin{pmatrix} -1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -2 & -2 & -3 \\ -1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Thus the Characteristic polynomial of A is $K_A(\lambda) = x^4 - x^3 - 4\,x^2 - 2\,x - 3$ □

Using MATLAB for the Danilevsky Method

Consider the matrix A:

$>>\ A = [\ 1\ 2\ 4\ 3\ ;\ \ 2\ 4\ 5\ 1\ ;\ \ 3\ 2\ 1\ 4\ ;\ \ 5\ 1\ 2\ 3\ ]\ ,\ M = A\ ;\ \ I = eye(4);$

$$A = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 4 & 5 & 1 \\ 3 & 2 & 1 & 4 \\ 5 & 1 & 2 & 3 \end{pmatrix}$$

Use the fourth row of A to define U and its inverse V:

$>>\ U = I\ ;\ U\,(3,:) = A\,(4,:)\ ,\ \ V = I\ ;\ \ V\,(3,:) = -A\,(4,:)\,/\,A\,(4,3)\ ;\ V\,(3,3) = 1/\,A\,(4,3)$

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 5 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad V = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ -2.5000 & -0.5000 & 0.5000 & -1.5000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial.

$>>\ B = U * A * V$

$$B = \begin{pmatrix} -9.0000 & 0.0000 & 2.0000 & -3.0000 \\ -10.5000 & 1.5000 & 2.5000 & -6.5000 \\ -54.5000 & 4.5000 & 16.5000 & -16.5000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{pmatrix}$$

Change B into A and find a new U and its inverse V, by using the third row of the new A:

$>>\ A = B\,;\ U = I\,;\ U\,(2,:) = A\,(3,:)\,,\ V = I\,;\ V(2,:) = -A\,(3,:)\,/\,A\,(3,2)\,;V\,(2,2) = 1/\,A\,(3,2)$

$$U = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ -54.5000 & 4.5000 & 16.5000 & -16.5000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix} \qquad V = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 12.1111 & 0.2222 & -3.6667 & 3.6667 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial.

>> $B = U * A * V$

$$B = \begin{pmatrix} -9 & 0 & 2 & -3 \\ 525 & 18 & -139 & 159 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Change B into A and find a new U and its inverse V, using the second row of the new A:

>> $A = B$; $U = I$; $U(1,:) = A(2,:)$, $V(1,:) = -A(2,:)/A(2,1)$; $V(1,1) = 1/A(2,1)$

$$U = \begin{pmatrix} -10.5000 & 1.5000 & 2.5000 & -6.5000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix} \qquad V = \begin{pmatrix} 0.0019 & -0.0343 & 0.2648 & -0.3029 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

>> $B = U * A * V$

$$B = \begin{pmatrix} 9 & 23 & -42 & -144 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

B is the companion matrix of our original matrix A. Here is the characteristic polynomial of A:

$$1.0000 \quad -9.0000 \quad -23.0000 \quad 42.0000 \quad 144.0000$$

Exceptional Case.

>> $A = [\ 1\ 2\ 4\ 3\ ;\ 2\ 4\ 5\ 1\ ;\ 3\ 2\ 1\ 4\ ;\ 5\ 1\ 0\ 3\ ]$

$$A = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 4 & 5 & 1 \\ 3 & 2 & 1 & 4 \\ 5 & 1 & 0 & 3 \end{pmatrix}$$

Since $A(4,3) = 0$, we need a permutation matrix P which moves $A(4,3) = 0$ into another position:

>> $P = I$; $Q = P$ $P(4,:) = P(3,:)$ $P(3,:) = Q(4,:)$

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial, but $B(4,3) \neq 0$.

Note that every time there is a zero in $A(k, k+1)$ entry, use a permutation matrix, to obtain a new matrix similar the matrix A, but with a non-zero value at that position.

>> $B = P * A * P'$

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 5 \\ 5 & 1 & 3 & 0 \\ 3 & 2 & 4 & 1 \end{pmatrix}$$

We set A to B and continue the same way as the previous example.

$>>\ U = I\ ;\ U\,(3,:) = A\,(4,:)\ ,\quad V = I\ ;\quad V\,(3,:) = -A\,(4,:)\,/\,A\,(4,3)\ ;\ V\,(3,3) = 1/\,A\,(4,3)$

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 2 & 4 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad V = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ -0.7500 & -0.5000 & 0.2500 & -0.2500 \\ 1.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial.

$>>\ B = U * A * V$

$$B = \begin{pmatrix} -1.2500 & 0.5000 & 0.7500 & 3.2500 \\ 1.2500 & 3.5000 & 0.2500 & 4.7500 \\ 9.7500 & 6.5000 & 6.7500 & 16.2500 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{pmatrix}$$

Change B into A and find a new U and its inverse V, using the second row of the new A:

$>>\ A = B\,;\ U = I\,;\ U\,(2,:) = A\,(3,:)\ ,\ V = I\,;\ V\,(2,:) = -A\,(3,:)\,/\,A\,(3,2)\,;\ V\,(2,2) = 1/\,A\,(3,2)$

$$U = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ -45.5000 & 11.0000 & -3.5000 & -6.5000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix} \qquad V = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 4.1364 & 0.0909 & 0.3182 & 0.5909 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial.

$>>\ B = U * A * V$

$$B = \begin{pmatrix} -1.6818 & 0.0070 & 0.2552 & 2.0455 \\ 57.0227 & -3.2797 & -53.4598 & -133.5682 \\ 0.0000 & 1.0000 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 \end{pmatrix}$$

Note that $A\,(2,1) \neq 0$, so we don't need any permutation matrix.

Change B into A and find a new U and its inverse V, using the second row of the new A:

$>>\ A = B\,;\ U = I\,;\ U\,(1,:) = A\,(2,:)\ ,\ V = I\,;\ V\,(1,:) = -A\,(2,:)\,/\,A\,(2,1)\,;\ V\,(1,1) = 1/\,A\,(2,1)$

$$U = \begin{pmatrix} 57.0227 & -3.2797 & -53.4598 & -133.5682 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix} \qquad V = \begin{pmatrix} 0.0175 & 0.0575 & 0.9375 & 2.3424 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Define the matrix B similar to A which has the same characteristic polynomial.

$>>\ B = U * A * V$

$$B = \begin{pmatrix} -4.9615 & -58.5769 & -208.9231 & -108.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{pmatrix}$$

B is the companion matrix of our original matrix A. Here is the characteristic polynomial of A:

$$1.0000 \quad 4.9615 \quad 58.5769 \quad 208.9231 \quad 108.0000$$

Matlab Program

```
A = input('Enter the square matrix A  : ');
m = size(A);  N = m(1);  b = [1];  B = zeros(N);  i = 1;
while i < N,
   J = eye(N); h = A(N − i + 1, N − i)
     while h == 0;
     c = A(N − i + 1, 1 : N − i);  z = norm(c, inf);
     if z  = 0;
       k = 1;  r = 0;
       while r == 0 & k < N − i;
           r = r + c(N − i − k);  k = k + 1;
           J(N − i, N − i) = 0;  J(N − i, N − i − k + 1) = 1;
           J(N − i − k + 1, N − i − k + 1) = 0;  J(N − i − k + 1, N − i) = 1,
           A = J ∗ A ∗ J;  k = k + 1;
       end
     else
       b = conv(b, [1 − A(N − i + 1, N − i + 1 : N)]);
       B = A(1 : N − i, 1 : N − i); A = B, N = N − i;
     end
      h = A(N − i + 1, N − i);
end
U = eye(N);  V = eye(N);
U(N − i, :) = A(N − i + 1, :);
V(N − i, :) = −A(N − i + 1, :)/A(N − i + 1, N − i);  V(N − i, N − i) = 1/A(N − i + 1, N − i);
A = U ∗ A ∗ V;
i = i + 1;
end;
b = conv(b,  [1 − A(N − i + 1, N − i + 1 : N)]);  disp(' '),
disp('The Characteristic polynomial looks like  :  '),  disp(' '),
disp(['K_A(x) = x ∧ n + c(1)x ∧ (n − 1), +... + c(n − 1)x + c(n)']), disp(' '),
disp('The coefficients list c(k) is  :  '), disp(' '),
disp(b),  disp(' ')
```

□

♣ **The Method of Krylov**. Let A be an $n \times n$ matrix. For any n-dimensional nonzero column vector v we associate its successive transforms

$$v_k = A^k v \qquad (k = 0, 1, 2, \ldots);$$

this sequence of vectors is called the *Krylov sequence* associated to the matrix A and the vector v.

At most n vectors of the sequence $v_0, v_1, v_2, \ldots$ will be linearly independent.

Suppose for some $r = r(v) \leq n$, the vectors $v_0 , v_1 , v_2 , \ldots , v_r$ are linearly independent and the vector v_{r+1} is a linear combination of the preceding ones. Hence there exists a monic polynomial

$$\phi(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{r-1}\lambda^{r-1} + \lambda^r$$

such that

$$\phi(A)\, v = (c_0\, I_n + c_1\, A + \cdots + c_{r-1}\, A^{r-1} + A^r)\, v = c_0\, v_0 + c_1\, v_1 + \cdots + c_{r-1}\, v_r + c_r\, v_{r+1} = \theta\,.$$

The polynomial $\phi(\lambda)$ is said to *annihilate* v and to be *minimal* for v. If $\omega\,(\lambda)$ is another monic polynomial which annihilates v,

$$\omega(A)\, v = \theta\,,$$

then $\phi\,(\lambda)$ divides $\omega\,(\lambda)$.

To show that; suppose

$$\omega\,(\lambda) = \gamma\,(\lambda)\,\phi\,(\lambda) + \rho\,(\lambda),$$

where $\rho\,(\lambda)$ is the remainder after dividing ω by ϕ, hence of degree strictly less than ρ, it follows that

$$\rho(A)\, v = \theta\,.$$

But $\phi\,(\lambda)$ is minimal for v, hence $\rho\,(\lambda) = 0$.

Now of all vectors v there is at least one vector for which the degree v is maximal, since for any vector v, $r(v) \leq n$. We call such vector a *maximal vector.*

Remark. A monic polynomial $\mu_A\,(\lambda)$ of an $n \times n$ matrix A is said to be its *minimal polynomial*, if $\mu_A\,(\lambda)$ is of minimum degree satisfying

$$\mu\,(A) = Z_n.$$

If the minimal polynomial of a matrix is equal to its characteristic polynomial, then we may use Krylov method to find the characteristic polynomial. For example, if the matrix has n distinct eigenvalues, then its minimal polynomial is equal to its characteristic polynomial. Hence the Krylov method would be successful.

Algorithm. To produce the characteristic polynomial of the matrix A by Krylov method, we follow the following steps:

Step 1. Choose an arbitrary n-dimensional nonzero column vector v such as e_1, then use the Krylov sequence to define the matrix

$$V = \left[\, v\,,\, A\, v\,,\, A^2\, v\,,\, \ldots\,,\, A^{n-2}\, v\,,\, A^{n-1}\, v \,\right] = [\, v_0\,,\, v_1\,,\, v-2\,,\, \ldots v_{n-2}\,,\, v_{n-1}]\,.$$

Step 2. If the matrix V has rank n, then the system $V\, c = -v_n$ has a unique solution

$$c^t = (c_0\,,\, c_1\,,\, c_2\,,\, \cdots\,,\, c_{n-1}).$$

The monic polynomial

$$\phi(\lambda) = c_0 + c_1\,\lambda + c_2\,\lambda^2 + \cdots + c_{n-1}\,\lambda^{n-1} + \lambda^n$$

which annihilates the vector v is the characteristic polynomial of A.

Step 3. If the system $V\, c = -v_n$ does not have a unique solution, then change the initial vector and try for example with e_2. □

Step 4. If again, the system $V\,c = -v_n$ does not have a unique solution, then either choose another initial vector v or suspect that the minimal polynomial and characteristic polynomial of A are different, in this case abandon this method and use another method.

Examples. Compute the characteristic polynomials of the following matrices:

$$A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 3 \\ 4 & -5 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & -3 & 1 \\ 1 & 0 & -2 & 1 \\ 1 & -3 & -1 & 3 \\ 1 & 0 & 1 & -2 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Choosing the initial vector $v = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ for the matrices A and B, we obtain

$$V_A = \left[\, v\,,\, A\,v\,,\, A^2\,v\,,\, A^3\,v \,\right] = \begin{pmatrix} 1 & 1 & 1 & 17 \\ 0 & 1 & 9 & 42 \\ 0 & 2 & 13 & 43 \\ 0 & 4 & 15 & 19 \end{pmatrix} \quad \text{and} \quad V_B = \left[\, v\,,\, B\,v\,,\, B^2\,v\,,\, B^3\,v \,\right] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

The matrix V_A is nonsingular, hence $c_A = -V_A^{-1} A^4 v = \begin{pmatrix} -87 \\ 28 \\ 2 \\ -4 \end{pmatrix}$. From the vector c_A we obtain the characteristic polynomial of A which is

$$K_A\,(\lambda) = \left(1\,,\, \lambda\,,\, \lambda^2\,,\, \lambda^4\,\right) \cdot \begin{pmatrix} -87 \\ 28 \\ 2 \\ -4 \end{pmatrix} + \lambda^4 = -87 + 28\,\lambda + 2\,\lambda^2 - 4\,\lambda^3 + \lambda^4.$$

The matrix V_B is singular, so we need another initial vector such as $v = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. The new matrix

$V_B = \begin{pmatrix} 0 & 2 & 11 & 11 \\ 1 & 0 & 8 & 0 \\ 0 & -3 & 5 & -21 \\ 0 & 0 & -1 & 18 \end{pmatrix}$ is invertible, so $c_B = -V_B^{-1}\, B^4\, v = \begin{pmatrix} 9 \\ -2 \\ -10 \\ 2 \end{pmatrix}$. From the vector c_B we obtain the characteristic polynomial of B which is

$$K_B(\lambda) = \left(1\,,\, \lambda\,,\, \lambda^2\,,\, \lambda^4\,\right) \cdot \begin{pmatrix} 9 \\ -2 \\ -10 \\ 2 \end{pmatrix} + \lambda^4 = 9 - 2\,\lambda - 10\,\lambda^2 + 2\,\lambda^3 + \lambda^4.$$

The minimal and characteristic polynomials of the matrix C are

$$m_C\,(\lambda) = \lambda^3 - 3\lambda^2 - 7\lambda \quad \text{and} \quad K_C\,(\lambda) = \lambda^4 - 3\lambda^3 - 7\lambda^2,$$

respectively. Therefore by choosing any initial vector v, the matrix $V_C = \left[\, v\,,\, C\,v\,,\, C^2\,v\,,\, C^3\,v \,\right]$ will always be singular. This means that the Krylov sequence will never produce the characteristic polynomial $K_C\,(\lambda)$.

Matlab Program

```
A = input('Enter a square matrix A : ');
m = size(A); n = m(1); V = zeros(n, n);
DL1 = ['Enter an initial ', int2str(n) , ' - dimensional row vector v0 = '];
v0 = input(DL1);
z = 0; k = 1;
while z == 0 & k < 5
    w = v0; V(:, 1) = w;
    for i = 2 : n,   w = A * w;   V(:, i) = w;   end,
    if det(V) ~= 0;   k = 8;   c = -inv(V) * A * w;
    else
      while k < 5
        v0 = input('The matrix V is singular, please enter another initial row vector v0 : ');
        k = k + 1;
      end;
    end;
   z = det(V);
end;
if k == 5;
  disp('Sorry, the Krylov method is not suited for this matrix. '), disp(' '),
else;
  disp('The Characteristic polynomial looks like : '), disp(' '),
  disp(['K_A(x) = c(0) + C(1)x + c(2)x ^ 2 + ··· + c(n - 1)x ^ (n - 1) + x ^ n']), disp(' '),
  disp('The coefficients list c(k) is : '), disp(' '),
  disp([c', 1])
end;
```

□