

## 18.06 Problem Set 7 - Solutions

Due Wednesday, 07 November 2007 at 4 pm in 2-106.

**Problem 1:** (12=3+3+3+3) Consider the matrix  $A = \begin{pmatrix} -1 & 3 & -1 & 1 \\ -3 & 5 & 1 & -1 \\ 10 & -10 & -10 & 14 \\ 4 & -4 & -4 & 8 \end{pmatrix}$ .

(a) If one eigenvector is  $\mathbf{v}_1 = (1 \ 1 \ 0 \ 0)^T$ , find its eigenvalue  $\lambda_1$ .

**Solution**  $A\mathbf{v}_1 = (2 \ 2 \ 0 \ 0)^T = 2\mathbf{v}_1$ , thus  $\lambda_1 = 2$ .

(b) Show that  $\det(A) = 0$ . Give another eigenvalue  $\lambda_2$ , and find the corresponding eigenvector  $\mathbf{v}_2$ .

**Solution** Since  $\det(A) = 0$ , and the determinant is the product of all eigenvalues, we see that there must be a zero eigenvalue. So  $\lambda_2 = 0$ . To find  $\mathbf{v}_2$ , we need to solve the system  $A\mathbf{v}_2 = 0$ . By Gauss elimination, it is easy to see that one solution is given by  $\mathbf{v}_2 = (2 \ 1 \ 1 \ 0)^T$ .

(c) Given the eigenvalue  $\lambda_3 = 4$ , write down a linear system which can be solved to find the eigenvector  $\mathbf{v}_3$ .

**Solution** The system is  $A\mathbf{v}_3 = 4\mathbf{v}_3$ , or  $(A - 4I)\mathbf{v}_3 = 0$ :

$$\begin{pmatrix} -5 & 3 & -1 & 1 \\ -3 & 1 & 1 & -1 \\ 10 & -10 & -14 & 14 \\ 4 & -4 & -4 & 4 \end{pmatrix} \mathbf{v}_3 = 0.$$

The solution is  $\mathbf{v}_3 = (0 \ 0 \ 1 \ 1)^T$ .

(d) What is the trace of  $A$ ? Use this to find  $\lambda_4$ .

**Solution** The trace of  $A$  is  $tr(A) = -1 + 5 - 10 + 8 = 2$ . Since

$$tr(A) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4,$$

the previous parts tell us  $\lambda_4 = 2 - (2 + 0 + 4) = -4$ .

**Problem 2:** (12=3+3+3+3) (a) Suppose  $n \times n$  matrices  $A, B$  have the same eigenvalues  $\lambda_1, \dots, \lambda_n$ , with the same independent eigenvectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . Show that  $A = B$ .

**Solution** Let  $S$  be the eigenvector matrix,  $\Lambda$  be the diagonal matrix consists of the eigenvalues. Then we have  $A = S\Lambda S^{-1}$  and also  $B = S\Lambda S^{-1}$ . Thus  $A = B$ .

(b) Find the  $2 \times 2$  matrix  $A$  having eigenvalues  $\lambda_1 = 2, \lambda_2 = 5$  with corresponding eigenvectors  $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

**Solution** We have  $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  and  $\Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$ . Thus

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

(c) Find two different  $2 \times 2$  matrices  $A, B$ , both have the same eigenvalues  $\lambda_1 = \lambda_2 = 2$ , and both have the same eigenvector (only one)  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ .

**Solution** Let  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Since  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  is an eigenvector with eigenvalue 2, we see

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

which implies  $a = 2, c = 0$ . Since  $a + d = \text{tr}(A) = 2 + 2$ , we see  $d = 2$ . Thus  $A = \begin{pmatrix} 2 & b \\ 0 & 2 \end{pmatrix}$ . It is very easy to check any such matrix has two eigenvalues  $\lambda_1 = \lambda_2 = 2$ . Finally the condition that  $A$  has only one eigenvector implies  $b \neq 0$ .

Similarly  $B$  has the same form. So we can take different values of  $b$  for  $A$  and  $B$ . For example,  $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$  and  $B = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$ .

(d) Find a matrix which has two different sets of independent eigenvectors.

**Solution** For the identity matrix, any set of independent vectors is a set of independent eigenvectors.

**Problem 3:** (12=3+3+3+3) Let  $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ .

(a) Find all eigenvalues and corresponding eigenvectors of  $A$ .

**Solution** The characteristic equation is

$$\lambda^2 - 4\lambda - 5 = 0,$$

thus the eigenvalues are  $\lambda_1 = 5, \lambda_2 = -1$ .

For  $\lambda_1 = 5$ , the eigenvector  $\mathbf{v}_1$  is a solution to

$$\begin{pmatrix} -4 & 4 \\ 2 & -2 \end{pmatrix} \mathbf{v}_1 = 0,$$

which implies  $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

For  $\lambda_1 = -1$ , the eigenvector  $\mathbf{v}_2$  is a solution to

$$\begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} \mathbf{v}_2 = 0,$$

which implies  $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

(b) Calculate  $A^{100}$  (not by multiplying  $A$  100 times!).

**Solution** We have

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}^{-1},$$

thus

$$\begin{aligned} A^{100} &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}^{100} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5^{100} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/3 & 2/3 \\ -1/3 & 1/3 \end{pmatrix} \\ &= \begin{pmatrix} (5^{100} + 2)/3 & (2 \cdot 5^{100} - 2)/3 \\ (5^{100} - 1)/3 & (2 \cdot 5^{100} + 1)/3 \end{pmatrix} \end{aligned}$$

(c) Find all eigenvalues and corresponding eigenvectors of  $A^3 - A + I$ .

**Solution** Use the formula above, we have

$$\begin{aligned} A^3 - A + I &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \left( \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}^3 - \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \left( \begin{pmatrix} 125 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 121 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \end{aligned}$$

Thus the eigenvalues of  $A^3 - A + I$  are

$$\lambda_1 = 121 = 5^3 - 5 + 1, \quad \lambda_2 = 1 = (-1)^3 - (-1) + 1.$$

The corresponding eigenvectors are still  $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

(d) For any polynomial function  $f$ , what are the eigenvalues and eigenvectors of  $f(A)$ ? Prove your statement.

**Solution** The eigenvalues of  $f(A)$  are  $f(5)$  and  $f(-1)$ , with the same eigenvector above. To prove this, we write

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Then

$$f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I.$$

If  $\lambda$  is an eigenvalue of  $A$  with eigenvector  $\mathbf{v}$ , i.e.  $A\mathbf{v} = \lambda\mathbf{v}$ , then we have

$$A^2\mathbf{v} = A(A\mathbf{v}) = A(\lambda\mathbf{v}) = \lambda A\mathbf{v} = \lambda^2\mathbf{v},$$

and in general by induction

$$A^k\mathbf{v} = \lambda^k\mathbf{v}.$$

This implies

$$f(A)\mathbf{v} = a_n \lambda^n \mathbf{v} + a_{n-1} \lambda^{n-1} \mathbf{v} + \cdots + a_1 \lambda \mathbf{v} + a_0 \mathbf{v} = f(\lambda)\mathbf{v},$$

which implies that  $\mathbf{v}$  is also an eigenvector of  $f(A)$ , which corresponding to the eigenvalue  $f(\lambda)$ .

(Thanks to Brian: In the special case that  $f(5) = f(-1)$ , any nonzero vector in the plane will be an eigenvector. In this case it is easy to see that  $f(A)$  is in fact a pure scalar matrix.)

---

**Problem 4:** (10=4+3+3) For simplicity, assume that  $A$  has  $n$  independent eigenvectors, thus diagonalizable. (However, the statement below also holds for general matrices.)

(a) Since  $A$  is diagonalizable, we can write  $A = SAS^{-1}$  as in class. Substitute this into the characteristic polynomial

$$p(A) = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I)$$

to show that  $p(A) = 0$ . (This is called the **Cayley-Hamilton theorem**.)

Solution We have

$$\begin{aligned} p(A) &= (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I) \\ &= (S\Lambda S^{-1} - S\lambda_1 I S^{-1})(S\Lambda S^{-1} - S\lambda_2 I S^{-1}) \cdots (S\Lambda S^{-1} - S\lambda_n I S^{-1}) \\ &= S(\Lambda - \lambda_1 I)S^{-1}S(\Lambda - \lambda_2 I)S^{-1}S \cdots S^{-1}S(\Lambda - \lambda_n I)S \\ &= S \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 - \lambda_1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n - \lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1 - \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n - \lambda_2 \end{pmatrix} \\ &\quad \cdots \begin{pmatrix} \lambda_1 - \lambda_n & 0 & \cdots & 0 \\ 0 & \lambda_2 - \lambda_n & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} S^{-1} \\ &= S \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} S^{-1} \\ &= 0. \end{aligned}$$

(b) Test the Cayley-Hamilton theorem for the matrix  $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ . Then write  $A^{-1}$  as a polynomial function of  $A$ . [Hint: move the  $I$  term to one side of the equation  $p(A) = 0$  to write  $A \cdot (\text{something}) = I$ .]

Solution The characteristic polynomial of  $A$  is  $p(\lambda) = \lambda^2 - \lambda - 1$ . Thus we need

to check  $A^2 - A - I = 0$ :

$$\begin{aligned} A^2 - A - I &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 - \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

From  $A^2 - A - I = 0$  we have  $A^2 - A = I$ . In other words,  $A(A - I) = I$ . We conclude that  $A^{-1} = A - I$ .

(c) Use the Cayley-Hamilton theorem above to show that, for any invertible matrix  $A$ ,  $A^{-1}$  can always be written as a polynomial of  $A$ . (Inverting using elimination is usually much more practical, however!)

Solution Suppose  $A$  is invertible, then  $\det A \neq 0$ . From Cayley-Hamilton theorem we have

$$p(A) = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I) = 0.$$

Write

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

then

$$a_0 = \lambda_1 \lambda_2 \cdots \lambda_n = \det(A) \neq 0,$$

and we have

$$p(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I = 0.$$

Rewrite this as

$$-\frac{1}{a_0}A^n - \frac{a_{n-1}}{a_0}A^{n-1} - \cdots - \frac{a_1}{a_0}A = I,$$

or

$$A\left(-\frac{1}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \cdots - \frac{a_1}{a_0}I\right) = I,$$

which implies

$$A^{-1} = -\frac{1}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \cdots - \frac{a_1}{a_0}I.$$

This completes the proof.

**Problem 5:** (10=5+5) (a) If  $A$  (an  $n \times n$  matrix) has  $n$  nonnegative eigenvalues  $\lambda_k$  and independent eigenvectors  $\mathbf{x}_k$ , and if we define “ $\sqrt{A}$ ” as the matrix with eigenvalues  $\sqrt{\lambda_k}$  and the same eigenvectors, show that  $(\sqrt{A})^2 = A$ .

**Solution** We can decompose  $A$  into  $A = S\Lambda S^{-1}$ , where  $S$  is the matrix consists of eigenvectors of  $A$ , and

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

is the diagonal eigenvalue matrix. Then by definition,  $\sqrt{A}$  is the matrix with  $S$  as its eigenvector matrix, and  $\sqrt{\lambda_i}$  its eigenvalues. In other words, we have  $\sqrt{A} = S\sqrt{\Lambda}S^{-1}$ , where

$$\sqrt{\Lambda} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}.$$

Obviously, we have  $\sqrt{\Lambda}\sqrt{\Lambda} = \Lambda$ , so

$$(\sqrt{A})^2 = \sqrt{A}\sqrt{A} = S\sqrt{\Lambda}S^{-1}S\sqrt{\Lambda}S^{-1} = S\Lambda S^{-1} = A.$$

(b) Given  $\sqrt{A}$  from part (a), is the only other matrix whose square is  $A$  given by  $-\sqrt{A}$ ? Why or why not?

**Solution** No. We can change some (but not all) of the  $\sqrt{\lambda_i}$  above to  $-\sqrt{\lambda_i}$ , then we will get a new matrix  $S\widetilde{\sqrt{\Lambda}}S^{-1}$ , whose square is  $A$ , but it is neither  $\sqrt{A}$ , nor  $-\sqrt{A}$ . So In general there are at least  $2^n$  possible matrices whose square is  $A$ .

As an example, we can take  $\widetilde{\sqrt{\Lambda}}$  to be

$$\begin{pmatrix} -\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}.$$

---

**Problem 6:** (10) If  $\lambda$  is an eigenvalue of  $A$ , is it also an eigenvalue of  $A^T$ ? What about the eigenvectors? Justify your answers.

**Solution** The eigenvalues of  $A$  must be the eigenvalues of  $A^T$ . In fact, the characteristic equation for  $A^T$  is  $\det(A^T - \lambda I) = 0$ . Since

$$\det(A^T - \lambda I) = \det(A^T - \lambda I)^T = \det(A - \lambda I),$$

we see that the characteristic equation for  $A^T$  coincides with the characteristic equation for  $A$ . This implies that the eigenvalues of  $A^T$  are the same as the eigenvalues of  $A$ .

However, the eigenvectors of  $A^T$  don't have to be the same as eigenvectors of  $A$ . In fact, the eigenvector of  $A$  corresponding to eigenvalue  $\lambda$  lies in the nullspace of  $A - \lambda I$ , while the eigenvector of  $A^T$  corresponding to the eigenvalue  $\lambda$  lies in the left nullspace of  $A - \lambda I$ , so they are different in general.

As an example, let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the both  $A$  and  $A^T$  has eigenvalues  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ , the eigenvectors of  $A$  are

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

while the eigenvectors of  $A^T$  are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Finally we point out that the number of independent eigenvectors of  $A^T$  is the same as the number of independent eigenvectors of  $A$ . In fact, for each eigenvalue  $\lambda$ , the number of independent eigenvectors of  $A^T$  or  $A$  both equal  $n - \text{rank}(A - \lambda I)$ , since  $\text{rank}(A - \lambda I) = \text{rank}(A^T - \lambda I)$ .

---



**Problem 7:** (10=2+2+2+2+2) This problem refers to *similar matrices* in the sense defined by section 6.6 (page 343) of the text.

(a)  $A$ ,  $B$ , and  $C$  are square matrices where  $A$  is similar to  $B$  and  $B$  is similar to  $C$ . Is  $A$  similar to  $C$ ? Why or why not?

**Solution** Yes,  $A$  is similar to  $C$ .

By definition,  $A$  is similar to  $B$  means  $A = PBP^{-1}$ , and  $B$  is similar to  $C$  means  $B = QCQ^{-1}$ . Thus we have  $A = PBP^{-1} = PQCQ^{-1}P^{-1} = (PQ)C(PQ)^{-1}$ . This implies that  $A$  is similar to  $C$ .

(b) If  $A$  is similar to  $Q$ , where  $Q$  is an orthogonal matrix ( $Q^T Q = I$ ), is  $A$  orthogonal too? Why or why not?

**Solution** No,  $A$  doesn't have to be orthogonal.

Counterexample: It is easy to check that  $Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  is orthogonal, while  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Q \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$  is not a orthogonal matrix.

(c) If  $A$  is similar to  $U$  where  $U$  is triangular, is  $A$  triangular too? Why or why not?

**Solution** No.

One can use the same example above:  $A$  is triangular,  $Q$  is similar to  $A$  but  $Q$  is not triangular.

(d) If  $A$  is similar to  $B$  where  $B$  is symmetric, is  $A$  symmetric too? Why or why not?

**Solution** No.

One can use the same example in part (b):  $Q$  is symmetric,  $A$  is similar to  $Q$  but  $A$  is not symmetric.

(e) For a given  $A$ , does the set of all matrices similar to  $A$  form a subspace of the set of all matrices (under ordinary matrix addition)? Why or why not?

**Solution** No.

The zero matrix will not lie in this set. In fact, if the zero matrix is similar to  $A$ , then  $A = P0P^{-1} = 0$ . In other words, the set is a subspace only if  $A$  is itself zero matrix.

---

**Problem 8:** (10=4+3+3) The *Pell numbers* are the sequence  $p_0 = 0$ ,  $p_1 = 1$ ,  $p_n = 2p_{n-1} + p_{n-2}$  ( $n > 1$ ).

(a) Analyzing the Pell numbers in the same way that we analyzed the Fibonacci sequence, via eigenvectors and eigenvalues of a  $2 \times 2$  matrix, find a closed-form expression for  $p_n$ .

Solution Rewrite the equation in matrix form, we have

$$\begin{pmatrix} p_{n+1} \\ p_n \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p_n \\ p_{n-1} \end{pmatrix}.$$

The characteristic equation of the matrix above is

$$\lambda^2 - 2\lambda - 1 = 0,$$

so the eigenvalues are

$$\lambda_1 = 1 + \sqrt{2}, \lambda_2 = 1 - \sqrt{2}.$$

The eigenvector corresponds to  $\lambda_1$  is given by

$$\begin{pmatrix} 1 - \sqrt{2} & 1 \\ 1 & -1 - \sqrt{2} \end{pmatrix} \mathbf{v}_1 = 0 \implies \mathbf{v}_1 = \begin{pmatrix} 1 \\ \sqrt{2} - 1 \end{pmatrix}.$$

Similarly the eigenvector corresponds to  $\lambda_2$  is given by

$$\begin{pmatrix} 1 + \sqrt{2} & 1 \\ 1 & -1 + \sqrt{2} \end{pmatrix} \mathbf{v}_2 = 0 \implies \mathbf{v}_2 = \begin{pmatrix} 1 - \sqrt{2} \\ 1 \end{pmatrix}.$$

So we have

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 - \sqrt{2} \\ \sqrt{2} - 1 & 1 \end{pmatrix} \begin{pmatrix} 1 + \sqrt{2} & 0 \\ 0 & 1 - \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 1 - \sqrt{2} \\ \sqrt{2} - 1 & 1 \end{pmatrix}^{-1}.$$

Thus

$$\begin{aligned} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}^n &= \begin{pmatrix} 1 & 1 - \sqrt{2} \\ \sqrt{2} - 1 & 1 \end{pmatrix} \begin{pmatrix} 1 + \sqrt{2} & 0 \\ 0 & 1 - \sqrt{2} \end{pmatrix}^n \begin{pmatrix} 1 & 1 - \sqrt{2} \\ \sqrt{2} - 1 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{4 - 2\sqrt{2}} \begin{pmatrix} (1 + \sqrt{2})^n + (\sqrt{2} - 1)^2(1 - \sqrt{2})^n & (\sqrt{2} - 1)((1 + \sqrt{2})^n - (1 - \sqrt{2})^n) \\ (\sqrt{2} - 1)((1 + \sqrt{2})^n - (1 - \sqrt{2})^n) & (\sqrt{2} - 1)^2(1 + \sqrt{2})^n + (1 - \sqrt{2})^n \end{pmatrix}. \end{aligned}$$

So  $p_n = \frac{1}{2\sqrt{2}}((1 + \sqrt{2})^n - (1 - \sqrt{2})^n)$ .

(b) Prove that the ratio  $(p_{n-1} + p_n)/p_n$  tends to  $\sqrt{2}$  as  $n$  grows.

**Solution** Since  $-1 < 1 - \sqrt{2} < 0$ , from the above expression we see that, for large  $n$ ,

$$p_n = \frac{1}{2\sqrt{2}}((1 + \sqrt{2})^n - (1 - \sqrt{2})^n)$$

approaches  $(1 + \sqrt{2})^n/2\sqrt{2}$ . Thus  $p_{n-1}/p_n$  tends to  $1/(1 + \sqrt{2}) = \sqrt{2} - 1$ , after normalizing the denominator. This implies that  $(p_{n-1} + p_n)/p_n$  tends to  $\sqrt{2}$  as  $n$  grows to  $\infty$ .

(c) Using MATLAB or a calculator, evaluate the ratio from (b) up to  $n = 10$  or so, and subtract  $\sqrt{2}$  to find the difference for each  $n$ . Compare this method of computing  $\sqrt{2}$  to Newton's method from 18.01:<sup>1</sup> starting with  $x = 1$ , repeatedly replace  $x$  by  $(x + 2/x)/2$ . Which technique goes faster to  $\sqrt{2}$ ?

**Solution**

We use MATLAB. The ratios from (b) are given by

```
>> x=0;y=1;z=y;
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
    0.0858
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
   -0.0142
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
    0.0025
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
  -4.2046e-04
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
   7.2152e-05
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
```

---

<sup>1</sup>The square root  $\sqrt{y}$  is the root of  $f(x) = x^2 - y$ , and Newton's method replaces  $x$  by  $x - f(x)/f'(x) = (x + y/x)/2$ . Actually, this technique for square roots was known to the Babylonians 3000 years ago.

```

-1.2379e-05
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
  2.1239e-06
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
 -3.6440e-07
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
  6.2522e-08
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
 -1.0727e-08
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
  1.8405e-09
>> y=2*y+x;x=z;z=y;(x+y)/y-sqrt(2)
ans =
 -3.1577e-10

```

while Euler method gives

```

>> x=1;
>> x=(x+2/x)/2;x-sqrt(2)
ans =
  0.0858
>> x=(x+2/x)/2;x-sqrt(2)
ans =
  0.0025
>> x=(x+2/x)/2;x-sqrt(2)
ans =
  2.1239e-06
>> x=(x+2/x)/2;x-sqrt(2)
ans =
  1.5947e-12
>> x=(x+2/x)/2;x-sqrt(2)
ans =

```

```

-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16
>> x=(x+2/x)/2;x-sqrt(2)
ans =
-2.2204e-16

```

The Newton's method is faster. In fact, the difference between the Pell ratios and  $\sqrt{2}$  decreases exponentially (as the other eigenvector disappears), and as a consequence the number of digits of accuracy (the log of the error) increases linearly with the number of iterations ("linear convergence"). In contrast, Newton's method roughly squares the error or doubles the number of significant digits with each iteration ("quadratic convergence"), which is far faster. (Of course, the error stops improving once the machine precision is reached.)

---

**Problem 9:** (14=2+2+2+2+2+2+2) This is a MATLAB problem on symmetric matrices.

(a) Use MATLAB to construct a random  $4 \times 4$  symmetric matrix ( $A = \text{rand}(4,4); A = A' * A$ ), and find its eigenvalues via the command `eig(A)`.

Solution The codes

```
>> A=rand(4,4);A=A'*A
```

```
A =
```

```
    2.3346    1.1384    2.5606    1.4507
    1.1384    0.7860    1.2743    0.9531
    2.5606    1.2743    2.8147    1.6487
    1.4507    0.9531    1.6487    1.8123
```

```
>> eig(A)
```

```
ans =
```

```
    0.0009
    0.1441
    0.7379
    6.8648
```

(b) Construct 1000 random vectors as the columns of  $X$  via  $X = \text{rand}(4,1000) - 0.5$ ; and for each vector  $\mathbf{x}_k$  compute  $d_k = \mathbf{x}_k^T A \mathbf{x}_k / \mathbf{x}_k^T \mathbf{x}_k$  via the command `d = diag(X'*A*X) ./ diag(X'*X)`; (which computes a vector  $\mathbf{d}$  of these 1000 ratios).

Solution The codes

```
>> X=rand(4,1000)-0.5;
>> d=diag(X'*A*X)./diag(X'*X);
```

(c) Find the minimum ratio  $d_k$  via `min(d)` and the maximum via `max(d)`. How do these compare to the eigenvalues? Try increasing from 1000 to 10000 above to see if your hypothesis holds up.

Solution The codes

```
>> min(d),max(d)
```

```
ans =
```

```
0.0122
```

```
ans =
```

```
6.8117
```

It seems that the minimum  $d_k$  is close to the smallest eigenvalue of  $A$ , and the maximum  $d_k$  is close to the largest eigenvalue of  $A$ . Moreover, the values of  $d_k$  are also bounded below/above by the smallest/largest eigenvalue.

We increase 1000 to 10000, which also implies the above judgement:

```
>> X=rand(4,10000)-0.5;
```

```
>> d=diag(X'*A*X)./diag(X'*X);
```

```
>> min(d),max(d)
```

```
ans =
```

```
0.0050
```

```
ans =
```

```
6.8451
```

(d) Find the eigenvectors of  $A$  via  $[S,L] = \text{eig}(A)$ :  $L$  is the diagonal matrix  $\Lambda$  of eigenvalues, and  $S$  is the matrix whose columns are the eigenvectors, so that  $AS = SL$ . Does  $S$  have any special properties? (Hint: try looking at  $\det(S)$  and  $\text{inv}(S)$  compared to  $S$ .) (We will see in class, eventually, that these and other nice properties come from  $A$  being symmetric.)

Solution The codes:

```
>> [S,L]=eig(A)
```

```
S =
```

```
    0.7196    0.1621    0.3613    0.5705
    0.0416   -0.9401   -0.1444    0.3061
   -0.6924    0.1247    0.3268    0.6310
    0.0321    0.2729   -0.8613    0.4274
```

```
L =
```

```
    0.0009         0         0         0
         0    0.1441         0         0
         0         0    0.7379         0
         0         0         0    6.8648
```

```
>> det(S)
```

```
ans =
```

```
    -1
```

```
>> inv(S)
```

```
ans =
```

```
    0.7196    0.0416   -0.6924    0.0321
    0.1621   -0.9401    0.1247    0.2729
    0.3613   -0.1444    0.3268   -0.8613
    0.5705    0.3061    0.6310    0.4274
```



We can see that  $S^{-1} = S^T$ , which means that  $S$  is an orthogonal matrix. In other words, for symmetric matrices, we can choose the eigenvectors such that they form an orthonormal basis!

(e) If you wanted to pick a vector  $\mathbf{x}$  to get the maximum possible  $d$ , what would  $\mathbf{x}$  and the resulting  $d$  be? What about the minimum?

**Solution** We have seen from part (c) that the maximum possible  $d$  is the maximum eigenvalue. So we should take the vector  $\mathbf{x}$  to be an eigenvector corresponding to the maximum eigenvalue. Similarly, if we want get the minimum possible  $d$ , we need to take  $\mathbf{x}$  to be an eigenvector of the minimum eigenvalue.

(f) Repeat the process for a  $2 \times 2$  complex *asymmetric*  $A$  and a *complex*  $X$ . This time, plot the resulting (complex)  $d_k$  values in the complex plane as black dots, and the (complex) eigenvalues as red circles:

```
A = rand(2,2) + i*rand(2,2);
X = rand(2,1000)-0.5+i*(rand(2,1000)-0.5);
d = diag(X'*A*X) ./ diag(X'*X);
v = eig(A)
plot(real(d), imag(d), 'k.', real(v), imag(v), 'ro')
```

You should get an elliptical region with the eigenvalues as foci.

**Solution** The codes

```
>> A=rand(2,2)+i*rand(2,2);
>> X=rand(2,1000)-0.5+i*(rand(2,1000)-0.5);
>> d=diag(X'*A*X)./diag(X'*X);
>> v=eig(A)

v =

    0.4377 + 0.2842i
    1.2904 + 0.8951i

>> plot(real(d),imag(d),'k.',real(v),imag(v),'ro')
```

```
>> axis equal
```

The output graph is

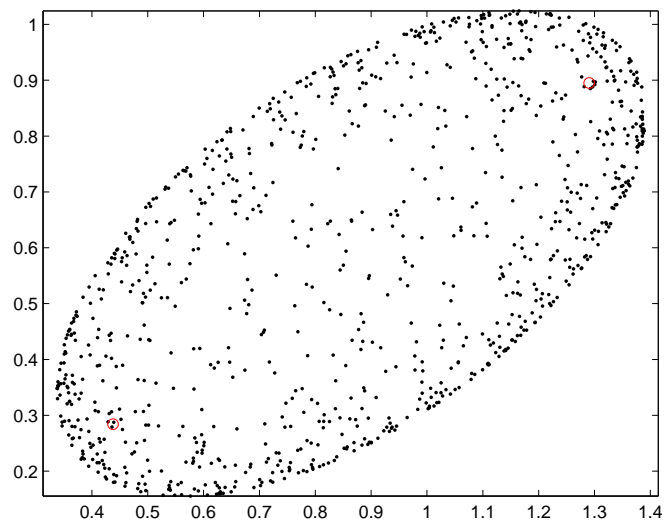


Figure 1: numerical range

(g) It is hard to see in part (f) that the eigenvalues are foci unless they are close together, which is unlikely. Construct a random non-symmetric  $A$  with eigenvalues  $1 + 1i$  and  $1.05 + 0.9i$  by performing a random  $2 \times 2$  similarity transformation on  $D = \text{diag}([1+1i, 1.05+0.9i])$  (i.e. construct a random  $S$  and multiply  $A = \text{inv}(S) * D * S$ ), and then repeat (f) with this  $A$ .

Solution The codes

```
>> S=rand(2,2);A=inv(S)*diag([1+1*i,1.05+0.9*i])*S;
>> X=rand(2,1000)-0.5+i*(rand(2,1000)-0.5);
>> d=diag(X'*A*X)./diag(X'*X);
>> v=eig(A)
```

v =

```
1.0000 + 1.0000i  
1.0500 + 0.9000i
```

```
>> plot(real(d),imag(d),'k.',real(v),imag(v),'ro')  
>> axis equal
```

The output graph is

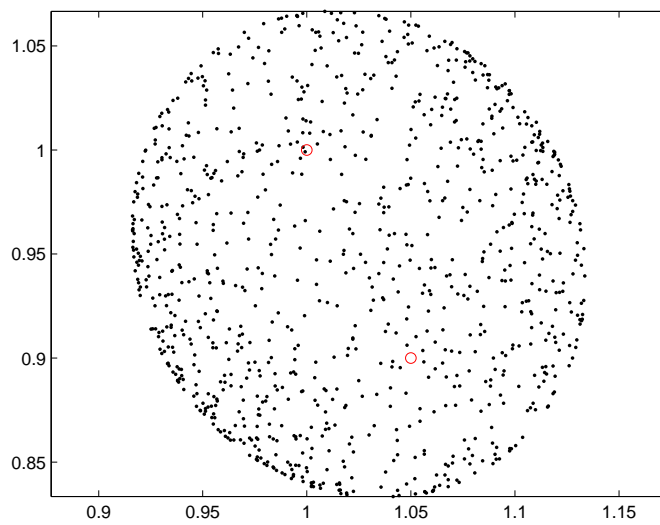


Figure 2: numerical range