Assignment 7 - Solutions

Due: November 20, in class
No late assignments accepted

Issued: November 13, 2013

Problem 1

For this problem we assume that eigenvalues and eigenvectors are all real valued.

(a) Let A be an $n \times n$ symmetric matrix. Let $\vec{q}_i$ and $\vec{q}_j$ be the eigenvectors of A corresponding to the eigenvalues λ_i and λ_j respectively. Show that if $\lambda_i \neq \lambda_j$, then $\vec{q}_i$ and $\vec{q}_j$ are orthogonal.

(b) Let A be an $n \times n$ matrix. We say that A is **positive definite** if for any non-zero vector $\vec{x}$, the following inequality holds

$$\vec{x}^T A\vec{x} > 0.$$

Show that the eigenvalues of a positive definite matrix A are all positive.

(c) $\star\star$ Let A be an $n \times n$ matrix. Show that

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i,$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A (λ_i's do not have to be all different).

[Hint 1: One way to prove this is to use the fact that any square matrix with real eigenvalues can be decomposed in the following way (called Schur decomposition)

$$A = QRQ^T,$$

where R is an upper triangular matrix and Q is an orthogonal matrix.]

[Hint 2: The following property of trace might be useful: given two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$, the trace of their product, $\mathrm{tr}(AB)$, is *invariant under cyclic permutations*, i.e. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
Note that this implies $\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$ for any matrices A, B, C with appropriately chosen dimension.]

Solution:

(a) By the definition of eigenvalues/eigenvectors, we have

$$A\vec{q}_i = \lambda_i\vec{q}_i, \quad A\vec{q}_j = \lambda_j\vec{q}_j$$

Then, since $A = A^T$,

$$\lambda_i\vec{q}_j^T\vec{q}_i = \vec{q}_j^T\lambda_i\vec{q}_i = \vec{q}_j^T A\vec{q}_i = \vec{q}_j^T A^T\vec{q}_i = (A\vec{q}_j)^T\vec{q}_i = (\lambda_j\vec{q}_j)^T\vec{q}_i = \lambda_j\vec{q}_j^T\vec{q}_i.$$

Rearranging, we have

$$(\lambda_i - \lambda_j)\vec{q}_j^T\vec{q}_i = 0.$$

When $\lambda_i \neq \lambda_j$ the term in the parentheses is non-zero, so it follows that

$$\vec{q}_j^T \vec{q}_i = 0,$$

which means that $\vec{q}_i$ and $\vec{q}_j$ are orthogonal.

(b) Let λ be an eigenvalue of A corresponding to an eigenvector $\vec{q}$, i.e.

$$A\vec{q} = \lambda\vec{q}.$$

Then, since A is positive definite, we have

$$0 < \vec{q}^T A\vec{q} = \vec{q}^T \lambda\vec{q} = \lambda\|\vec{q}\|_2^2.$$

Since $\|\vec{q}\|_2 > 0$ ($\vec{q} \neq \vec{0}$ for any eigenvector $\vec{q}$), it follows that $\lambda > 0$.

(c) Consider the Schur decomposition of A,

$$A = QRQ^T.$$

Using the property of trace that $\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$ (invariance under cyclic permutations), we have

$$\mathrm{tr}(A) = \mathrm{tr}(QRQ^T) = \mathrm{tr}(RQ^T Q) = \mathrm{tr}(R) = \sum_{i=1}^{n} R_{ii}.$$

We now prove that for any (upper) triangular matrix R the diagonal elements of R are its eigenvalues. Consider the characteristic polynomial of R,

$$p(\lambda) = \det(\lambda I - R) = (\lambda - R_{11})(\lambda - R_{22}) \cdots (\lambda - R_{nn}) = \prod_{i=1}^{n} (\lambda - R_{ii}),$$

where for the second equality we used the fact that the determinant of a triangular matrix is the product of its diagonal entries. Since the roots of a characteristic equation are the eigenvalues of R, we can immediately see that $\lambda_i = R_{ii}$ for $i = 1, \ldots, n$, so that

$$\mathrm{tr}(A) = \sum_{i=1}^{n} R_{ii} = \sum_{i=1}^{n} \lambda_i.$$

Now, it remains to show that the eigenvalues of R are equal to the eigenvalues of A. If α is an eigenvalue of A corresponding to eigenvector $\vec{v}$, then

$$\alpha\vec{v} = A\vec{v} = QRQ^T\vec{v}.$$

Pre-multiplying by Q^T, we have

$$Q^T\alpha\vec{v} = RQ^T\vec{v},$$
$$\alpha Q^T\vec{v} = RQ^T\vec{v},$$
$$\alpha\vec{u} = R\vec{u},$$

where $\vec{u} = Q^T\vec{v}$ is an eigenvector of R corresponding to eigenvalue α.
Similarly, if β is an eigenvalue of R corresponding to eigenvector $\vec{u}$, then

$$\beta\vec{u} = R\vec{u} = Q^T AQ\vec{u}.$$

Pre-multiplying by Q, we have

$$\beta Q\vec{u} = AQ\vec{u},$$
$$\beta\vec{v} = A\vec{v},$$

where $\vec{v} = Q\vec{u}$ is an eigenvector of A corresponding to eigenvalue β.
This proves the claim. Therefore,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i,$$

where $\lambda_i,\ i = 1, \dots, n$ are the eigenvalues of A (and R).

[**Note.** The fact that eigenvalues of A and R were the same holds in a more general setting. We say that two matrices A and B are **similar** if there exists and invertible matrix P, such that

$$B = P^{-1}AP.$$

The proof above shows that all similar matrices A and B have the same set of eigenvalues. This property is referred to as the **invariance of eigenvalues under similarity transformations**.]

Problem 2

We are interested in finding the fixed points (the points at which the time derivatives are zero) of the following system of equations:

$$\begin{aligned} \frac{dx_1}{dt} &= x_1(a - bx_2) \\ \frac{dx_2}{dt} &= -x_2(c - dx_1) \end{aligned}$$

for $a = 3, b = 1, c = 2, d = 1$. We can use the Newton-Raphson method to find these fixed points, simply by setting the derivatives zero in the given system of equations.

(a) In the scalar case, Newton-Raphson breaks down at points at which the derivative of the nonlinear function is zero. In general, where can it break down for systems of nonlinear equations? For the system given above, find the troublesome points.

(b) Find the fixed points of the above system analytically.

(c) Find all fixed points using repeated application of the Newton-Raphson method. You will have to judiciously choose your starting points (but of course, you are not allowed to use the known roots as starting points!). You may use MATLAB to program the method if you like.

Solution:

1. The Newton-Raphson method can break down for systems of nonlinear equations when the Jacobian matrix is singular. Lets examine the given system:

$$\begin{aligned} f_1(x_1, x_2) &= x_1(3 - x_2) \\ f_2(x_1, x_2) &= -x_2(2 - x_1) \\ J(x_1, x_2) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} \\ &= \begin{pmatrix} 3 - x_2 & -x_1 \\ x_2 & x_1 - 2 \end{pmatrix} \end{aligned}$$

The troublesome points can be found by setting the determinant of $J = 0$

$$\begin{aligned} \det(J) &= (3 - x_2)(x_1 - 2) + x_1x_2 \\ &= 0 \\ 3x_1 + 2x_2 &= 0 \end{aligned}$$

Hence, the troublesome points in this case form a line in two dimensional space.

2. The fixed points of the above system occur when the derivatives of x w.r.t. t are zero. Therefore

$$f_1(x_1, x_2) = x_1(3 - x_2) = 0$$
$$f_2(x_1, x_2) = -x_2(2 - x_1) = 0$$

This is easily solved to be (0,0) and (2,3).

3. The equation to solve at every iteration is

$$J(\vec{x}^{(k)})(\vec{x}^{(k+1)} - \vec{x}^{(k)}) = -f(\vec{x}^{(k)})$$

$$\begin{pmatrix} {x_1}^{(k+1)} \\ {x_2}^{(k+1)} \end{pmatrix} = -\begin{pmatrix} 3 - {x_2}^{(k)} & {x_1}^{(k)} \\ {x_2}^{(k)} & {x_1}^{(k)} - 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1^{(k)}(3 - x_2^{(k)}) \\ x_2^{(k)}(2 - x_1^{(k)}) \end{pmatrix} + \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix}$$

This iteration can be computed in Matlab or by hand. Of course, we should not come across points that make the Jacobian singular and therefore, we choose the starting points to be not close to the line of troublesome points calculated in the first part. Interestingly, travelling across the line when choosing the starting point, changes the final solution. So, we choose one starting point which lies above the line(e.g. (4,2)) to get (2,3) and another which is below the line(e.g (-1,1)) to get the other solution(0,0). This takes about 5 iterations to converge.

$$\vec{x}^{(1)} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \to \vec{x}^{(2)} = \begin{pmatrix} 1.6 \\ 2.4 \end{pmatrix} \to \vec{x}^{(3)} = \begin{pmatrix} 2.1333 \\ 3.2 \end{pmatrix} \to \vec{x}^{(4)} = \begin{pmatrix} 2.0078 \\ 3.2 \end{pmatrix} \to \vec{x}^{(5)} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\vec{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \to \vec{x}^{(2)} = \begin{pmatrix} 0.2857 \\ 0.4286 \end{pmatrix} \to \vec{x}^{(3)} = \begin{pmatrix} -0.0015 \\ -0.0023 \end{pmatrix} \to \vec{x}^{(4)} = 10^{-5}\begin{pmatrix} 0.1191 \\ 0.1786 \end{pmatrix} \to \vec{x}^{(5)} = 10^{-11}\begin{pmatrix} -0.0709 \\ -0.1063 \end{pmatrix}$$

Problem 3

(a) If $P^2 = P$, show that

$$e^P \approx I + 1.718P$$

(b) Convert the equation below to a matrix equation and then by using the exponential matrix find the solution in terms of $y(0)$ and $y'(0)$:

$$y'' = 0$$

(c) Show that $e^{A+B} = e^A e^B$ is not generally true for matrices.

Solution:

(a) From the definition of exponential matrices we have:

$$e^P = I + P + \frac{P^2}{2!} + \frac{P^3}{3!} + \cdots$$

In addition we know that:

$$P^2 = P$$
$$\Rightarrow P^n = P \quad \forall n \text{ by simple induction}$$

Therefore

$$\begin{aligned}
e^P &= I + P + \frac{P^2}{2!} + \frac{P^3}{3!} + \cdots \\
&= I + P + \frac{P}{2!} + \frac{P}{3!} + \cdots \\
&= I + P(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots)
\end{aligned}$$

From Taylor expansion of exponential function we have

$$\begin{aligned}
e^x &= 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots \\
e &= 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} \cdots \\
e - 1 &= 1 + \frac{1}{2!} + \frac{1}{3!} \cdots \\
&= 1.718
\end{aligned}$$

Therefore

$$\begin{aligned}
e^P &= I + P(1 + \frac{1}{2!} + \frac{1}{3!} \cdots) \\
&\approx I + 1.718P
\end{aligned}$$

(b) In order to transform this ODE to a matrix equation, we use:

$$\begin{aligned}
\frac{d}{dt}\begin{pmatrix} y \\ y' \end{pmatrix} &= \begin{pmatrix} y' \\ y'' \end{pmatrix} \\
&= \begin{pmatrix} y' \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} y \\ y' \end{pmatrix}
\end{aligned}$$

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

$$\begin{aligned}
\frac{d}{dt}\begin{pmatrix} y \\ y' \end{pmatrix} &= A\begin{pmatrix} y \\ y' \end{pmatrix} \\
\Rightarrow \begin{pmatrix} y \\ y' \end{pmatrix} &= e^{At}\begin{pmatrix} y(0) \\ y'(0) \end{pmatrix}
\end{aligned}$$

Now we look at

$$\begin{aligned}
e^{At} &= I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} \cdots \\
&= I + At + \frac{A^2}{2!}t^2 + \frac{A^3}{3!}t^3 \cdots
\end{aligned}$$

But A is nilpotent, i.e.

$$\begin{aligned} A^2 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &= A^n \quad \forall n \text{ by simple induction} \end{aligned}$$

Therefore

$$\begin{aligned} e^{At} &= I + At \\ &= \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} y \\ y' \end{pmatrix} &= \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} e^{At} \\ &= \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} \\ &= \begin{pmatrix} y(0) + y'(0)t \\ y'(0) \end{pmatrix} \end{aligned}$$

Therefore we have $y = y(0) + y'(0)t$.

(c) Let

$$\begin{aligned} A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B &= \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \\ \Rightarrow A^2 &= B^2 = 0 \\ \Rightarrow e^A &= I + A \\ e^B &= I + B \end{aligned}$$

Therefore

$$\begin{aligned} e^A e^B &= I + A + B + AB \\ e^B e^A &= I + B + A + BA \end{aligned}$$

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, BA = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$$

Therefore $e^A e^B \neq e^B e^A$ for these two matrices. Now suppose for contradiction that $e^{A+B} = e^A e^B$. Then

$$\begin{aligned} e^A e^B &= e^{A+B} \\ &= e^{B+A} \quad \text{addition is always commutative} \\ &= e^B e^A \end{aligned}$$

This is clearly not true because we showed earlier that $e^A e^B \neq e^B e^A$. Therefore $e^A e^B = e^{A+B}$ is generally not true.