

Homework 4

CS 210

Andres Calderon
SID 861243796

June 3, 2016

| Question | Points | Score |
|----------|--------|-------|
| 1 | 5 | |
| 2 | 10 | |
| 3 | 10 | |
| 4 | 10 | |
| 5 | 10 | |
| 6 | 5 | |
| 7 | 5 | |
| 8 | 10 | |
| 9 | 5 | |
| 10 | 15 | |
| Total | 85 | |

Eigenvalue problems

1. Consider the following statements about eigenvalue problems. Mark each statement as true or false.

T / ☒ F A defective eigenvalue is one where the geometric multiplicity is greater the algebraic multiplicity.

T / ☒ F A good way to compute eigenvalues is by finding roots of the associated characteristic polynomial.

T / ☒ F An orthogonal projection matrix has one eigenvalue equal to 0 and the other eigenvalues equal to 1.

☒ T / F Symmetric matrices have orthogonal set of eigenvectors.

T / ☒ F A projection matrix must have at least one eigenvalue equal to 0.

☒ T / F A matrix that has an orthogonal set of eigenvectors can be decomposed as $A = U\Lambda U^T$ where U is orthogonal and Λ is diagonal.

2. Given a symmetric $n \times n$ matrix A , show that two eigenvectors corresponding to two distinct eigenvalues must be orthogonal.

Answer:

Let's have two eigenvectors \mathbf{x}_1 and \mathbf{x}_2 with different eigenvalues λ_1 and λ_2 respectively. Given $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ we can multiply both sides by \mathbf{x}_2^T ,

$$\mathbf{x}_2^T A\mathbf{x}_1 = \lambda_1 \mathbf{x}_2^T \mathbf{x}_1$$

Now, let's transpose both sides,

$$\mathbf{x}_1^T \mathbf{x}_2 \lambda_1 = \mathbf{x}_1^T A^T \mathbf{x}_2$$

As λ_1 is a scalar and A is symmetric ($A = A^T$) we can reorganized as,

$$\lambda_1 \mathbf{x}_1^T \mathbf{x}_2 = \mathbf{x}_1^T A \mathbf{x}_2$$

Similarly, given $A \mathbf{x}_2 = \lambda_2 \mathbf{x}_2$, this time we will multiply both sides by \mathbf{x}_1^T ,

$$\mathbf{x}_1^T A \mathbf{x}_2 = \lambda_2 \mathbf{x}_1^T \mathbf{x}_2$$

So we can compare previous expressions as follows,

$$\lambda_1 \mathbf{x}_1^T \mathbf{x}_2 = \lambda_2 \mathbf{x}_1^T \mathbf{x}_2$$

Then,

$$\lambda_1 \mathbf{x}_1^T \mathbf{x}_2 - \lambda_2 \mathbf{x}_1^T \mathbf{x}_2 = 0$$

$$(\lambda_1 - \lambda_2) \mathbf{x}_1^T \mathbf{x}_2 = 0$$

As $\lambda_1 \neq \lambda_2$, $(\lambda_1 - \lambda_2)$ cannot be 0. We follow that $\mathbf{x}_1^T \mathbf{x}_2 = 0$ and conclude that $\mathbf{x}_1 \perp \mathbf{x}_2$.

Matrix deflation

Suppose we have found an eigenvalue, eigenvector pair $(\lambda_1, \mathbf{x}_1)$ of A , i.e., $A \mathbf{x}_1 = \lambda_1 \mathbf{x}_1$. *Matrix deflation* effectively removes the known eigenvalue from the matrix so that another eigenvalue, eigenvector pair can be computed. Below we look at two approaches to matrix deflation. See Heath 2002 Section 4.5.4, or Heath 1997 Section 4.3.5.

3. Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Suppose we have found an eigenvalue, eigenvector pair $(\lambda_1, \mathbf{x}_1)$ of A , i.e., $A \mathbf{x}_1 = \lambda_1 \mathbf{x}_1$. Let H be a Householder matrix such that $H \mathbf{x}_1 = \alpha \mathbf{e}_1$, a scalar multiple of the first column of the identity matrix.

- (a) Show that the similarity transformation determined by H transforms A into the block triangular form

$$H A H^{-1} = \begin{pmatrix} \lambda_1 & \mathbf{b}^T \\ \mathbf{0} & B \end{pmatrix},$$

where B is an $(n-1) \times (n-1)$ matrix.

- (b) Show that B has eigenvalues $\lambda_2, \dots, \lambda_n$, and that if $B \mathbf{y}_k = \lambda_k \mathbf{y}_k$, then $A \mathbf{x}_k = \lambda_k \mathbf{x}_k$ where $\mathbf{x}_k = H^{-1} \begin{pmatrix} \gamma \\ \mathbf{y}_k \end{pmatrix}$, $\gamma = \frac{\mathbf{b}^T \mathbf{y}_k}{\lambda_k - \lambda_1}$.

4. Another approach to matrix deflation is to let \mathbf{u}_1 be any vector such that $\mathbf{u}_1^T \mathbf{x}_1 = \lambda_1$, and define $B = A - \mathbf{x}_1 \mathbf{u}_1^T$.

- (a) Show that B has eigenvalues $0, \lambda_2, \dots, \lambda_n$.

(b) What are the eigenvectors of B ?

Answer:

To answer this question we have to make two assumptions:

- i. The eigenvalues of A^T are the same as those of A .
- ii. The eigenvectors from A and A^T that are associated with different eigenvalues are orthogonal.

Based on that, we start by multiplying $B = A - \mathbf{x}_1 \mathbf{u}_1^T$ by \mathbf{x}_1 ,

$$\begin{aligned} B\mathbf{x}_1 &= (A - \mathbf{x}_1 \mathbf{u}_1^T)\mathbf{x}_1 \\ B\mathbf{x}_1 &= A\mathbf{x}_1 - \mathbf{x}_1(\mathbf{u}_1^T \mathbf{x}_1) \end{aligned}$$

As $\mathbf{u}_1^T \mathbf{x}_1 = \lambda_1$,

$$\begin{aligned} B\mathbf{x}_1 &= A\mathbf{x}_1 - \mathbf{x}_1 \lambda_1 \\ B\mathbf{x}_1 &= \lambda_1 \mathbf{x}_1 - \lambda_1 \mathbf{x}_1 \\ B\mathbf{x}_1 &= 0 \end{aligned}$$

It shows that the first eigenvalue of B is 0. Now, for the remaining eigenvalues λ_i we transpose B before multiplying by the corresponding eigenvector \mathbf{x}_i ,

$$\begin{aligned} B^T &= A^T - \mathbf{u}_1 \mathbf{x}_1^T \\ B^T \mathbf{x}_i &= (A^T - \mathbf{u}_1 \mathbf{x}_1^T)\mathbf{x}_i \\ B^T \mathbf{x}_i &= A^T \mathbf{x}_i - \mathbf{u}_1(\mathbf{x}_1^T \mathbf{x}_i) \end{aligned}$$

Based on our previous assumptions we got,

$$\begin{aligned} B^T \mathbf{x}_i &= \lambda_i \mathbf{x}_i - \mathbf{u}_1(0) \\ B^T \mathbf{x}_i &= \lambda_i \mathbf{x}_i \end{aligned}$$

Showing that A and B share the same remaining eigenvalues $(\lambda_2, \dots, \lambda_n)$.

Nonlinear Equations

5. Consider the following statements about nonlinear equation solving. Mark each statement as true or false.

T / (F) A small residual $\|\mathbf{f}(\mathbf{x})\|$ guarantees an accurate solution of a system of nonlinear equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

(T) / F Newton's method is an example of a fixed point iteration.

T / (F) If an iterative method for solving a nonlinear equation gains more than one bit of accuracy per iteration, then it is said to have a superlinear convergence rate.

T / (F) Newton's method always converges quadratically.

T / (F) A *fixed point* of a function $f(x)$ is a point x^* such that $f(x^*) = 0$.

6. Compare Newton's method and the Secant Method for solving a scalar nonlinear equation. What are the advantages and disadvantages of each?

Answer:

As it is stated at the end of section 5.5.4 of the textbook, in general, Newton's method has a quadratic convergence rate faster than the superlinear convergence of the secant method. However, an important advantage of the secant method is that it just requires one new function evaluation per iteration instead of a derivative computation. Another drawback of the secant method is the requirement of two initial points instead of one of the Newton's method.

7. (Heath 5.1) Consider the nonlinear equation

$$f(x) = x^2 - 2 = 0.$$

- (a) With $x_0 = 1$, as a starting point, what is the value of x_1 if you use Newton's method for solving this problem?

Answer:

For Newton's method we will need the derivative of $f(x)$ which is given by,

$$f'(x) = 2x$$

So the iteration function is given by,

$$x_{k+1} = x_k - \frac{x_k^2 - 2}{2x_k}$$

For $x_0 = 1$ the value of x_1 will be,

$$\begin{aligned} x_1 &= x_0 - \frac{x_0^2 - 2}{2x_0} \\ x_1 &= 1 - \frac{1^2 - 2}{2(1)} \\ x_1 &= 1.5 \end{aligned}$$

- (b) With $x_0 = 1$ and $x_1 = 2$ as a starting points, what is the value of x_2 if you use the secant method for the same problem?

Answer:

For the secant method, the iteration function is given by,

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

Table 1 summarize the computation of x_2 given $x_0 = 1$ and $x_1 = 2$ as a starting points.

So, $x_2 = 1.33$

8. (Heath 5.12) Newton's method for solving a scalar nonlinear equation $f(x) = 0$ requires computation of the derivative of f at each iteration. Suppose that we instead replace the true derivative with a constant value d , that is, we use the iteration scheme

$$x_{k+1} = x_k - \frac{f(x_k)}{d}.$$

| k | x_k | $f(x_k)$ |
|---|-------|----------|
| 0 | 1.00 | -1.00 |
| 1 | 2.00 | 2.00 |
| 2 | 1.33 | |

Table 1: Computation of x_2 using the secant method.

- (a) Under what condition on the value of d will this scheme be locally convergent?

Answer:

Given the iteration function $g(x_k) = x_k - \frac{f(x_k)}{d}$, to reach convergence we will need find the derivative,

$$g'(x_k) = 1 - \frac{f'(x_k)}{d}$$

So the bounded condition for local convergence will be,

$$\left| 1 - \frac{f'(x_k)}{d} \right| < 1$$

$$-1 < 1 - \frac{f'(x_k)}{d} < 1$$

- (b) What will be the convergence rate, in general?

Answer:

Overall, the convergence rate should be linear with constant $C = \left| 1 - \frac{f'(x_k)}{d} \right|$.

- (c) Is there any value of d that would still yield quadratic convergence?

Answer:

To achieve quadratic convergence, $1 - \frac{f'(x_k)}{d} = 0$ should hold. It implies that for $d = f'(x_k)$ the convergence would be quadratic.

9. (Heath 5.10) Carry out one iteration of Newton's method applied to the system of nonlinear equations

$$\begin{aligned} x_1^2 - x_2^2 &= 0 \\ 2x_1x_2 &= 1 \end{aligned}$$

with starting value $\mathbf{x}_0 = (0, 1)^T$.

Answer:

Given,

$$f(x) = \begin{bmatrix} x_1^2 - x_2^2 \\ 2x_1x_2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

the Jacobian matrix is given by,

$$J_f(x) = \begin{bmatrix} 2x_1 & 2x_2 \\ 2 & 2 \end{bmatrix}$$

for $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ then,

$$f(\mathbf{x}_0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

and,

$$J_f(\mathbf{x}_0) = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}$$

solving the system,

$$J_f(\mathbf{x}_0)\mathbf{s}_0 = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix} \mathbf{s}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -f(\mathbf{x}_0)$$

gives $\mathbf{s}_0 = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}$ and hence,

$$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{s}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix}$$

10. *Computer problem* (Heath 5.3) Implement the bisection, Newton, and secant methods for solving non-linear equations in one dimension, and test your implementation by finding at least one root for each of the following equations. What termination criterion should you use? What convergence rate is achieved in each case?

(a) $x^3 - 2x - 5 = 0$.

(b) $e^{-x} = x$.

(c) $x \sin(x) = 1$.

(d) $x^3 - 3x^2 + 3x - 1 = 0$.