Homework 7

Iterative Methods

1. (Heath 4.42)

   (a) If a matrix A has a simple dominant eigenvalue λ_1, what quantity determines the convergence rate of the power method for computing λ_1?

   (b) How can the convergence rate of power iteration be improved?

   Solution:

   (a) The power method converges linearly with constant $|\lambda_2|/|\lambda_1|$.

   (b) The convergence can be improved by shifting the eigenvalues by a shift σ, so that $\lambda_i' = \lambda_i - \sigma$. The shift needs to be chosen such that $|\lambda_2'|/|\lambda_1'| < |\lambda_2|/|\lambda_1|$ and $|\lambda_1'| > |\lambda_2'| \geq |\lambda_k'|,\ k > 2$.

2. (Heath 4.24) Let A be an $n \times n$ real matrix of rank one.

   (a) Show that $A = \mathbf{u}\mathbf{v}^T$ for some nonzero real vectors $\mathbf{u}$ and $\mathbf{v}$.

   (b) Show that $\mathbf{u}^T\mathbf{v}$ is an eigenvalue of A.

   (c) If power iteration is applied to A, how many iterations are required for it to converge exactly to the eigenvector corresponding to the dominant eigenvalue?

   Solution:

   (a) If A has rank 1, then it has reduced SVD $A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$. Then $A = \mathbf{u}\mathbf{v}^T$, e.g., with $\mathbf{u} = \sigma_1 \mathbf{u}_1$ and $\mathbf{v} = \mathbf{v}_1$.

   (b) $A\mathbf{u} = \mathbf{u}\mathbf{v}^T\mathbf{u} = (\mathbf{u}^T\mathbf{v})\mathbf{u}$.

   (c) One iteration. This is because the other eigenvalues of A are all 0. Let those 0 eigenvalues have eigenvectors $\mathbf{v}_2 \ldots \mathbf{v}_n$. For any $\mathbf{x}_0 = \alpha_1\mathbf{u} + \alpha_2\mathbf{v}_2 + \ldots \alpha_n\mathbf{v}_n$ with $\alpha_1 \neq 0$,

   $$A\mathbf{x}_0 = (\mathbf{u}^T\mathbf{v})\alpha_1\mathbf{u}.$$

3. The asymptotic convergence rate of an iterative method is r if

   $$\lim_{k\to\infty} \|e_{k+1}\| = C\|e_k\|^r,$$

   where $e_k = x_k - x^*$ is the error at step k. Taking the log of both sides of the above equality

   $$\log\|e_{k+1}\| = \log C + r\log\|e_k\|.$$

   Thus the convergence rate can be seen to be the slope, r, of the log-log plot of error at $k+1$ vs error at k. Study the convergence rate of a few simple iterative methods in Matlab by implementing them and creating such plots. Include your code for each case.

   (a) Jacobi iteration to solve the system $A\mathbf{x} = \mathbf{b}$, with

   ```
   A = [ 3.5, 2, -1;  1 , 2.5, 0; 1, 2, -3.5]
   b = [ 1; 2; 3]
   ```

   and initial guess $\mathbf{x} = \mathbf{0}$.

   (b) The power method to find the dominant eigenvector of the symmetric matrix

```
rng(12)
A = rand(4,4);
A = A + A';
```

with initial guess

```
x0 = rand(4,1);
```

(c) Rayleigh quotient iteration to find an eigenvalue of the same matrix A from (b) starting with the same `x0`.

Nonlinear Equations

4. (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?

(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?

Solution:

(a)

$$\begin{aligned} f'(x) &= 2x \\ x_0 &= 1 \\ x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} = 1 + \frac{1}{2} = \frac{3}{2} \end{aligned}$$

(b)

$$\begin{aligned} x_0 &= 1, x_1 = 2 \\ x_2 &= x_1 - \frac{f(x_1)(x_1 - x_0)}{f(x_1) - f(x_0)} = 2 - \frac{2(2-1)}{2-(-1)} = 2 - \frac{2}{3} = \frac{4}{3}. \end{aligned}$$

5. (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

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

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

Solution:

Let $g(x) = x - \frac{f(x)}{d}$ so that $g'(x) = 1 - \frac{f'(x)}{d}$. x^* is a solution which satisfies $x^* = g(x^*)$. The scheme will be locally convergent when $|g'(x^*)| < 1$.

$$\begin{aligned} |g'(x^*)| &< 1 \\ |1 - \frac{f'(x^*)}{d}| &< 1 \\ 0 < \frac{f'(x^*)}{d} &< 2 \\ 0 < f'(x^*) &< 2d \end{aligned}$$

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

Solution:

If convergent, the convergence will generally be linear, since $g'(x^*) \neq 0$.

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

Solution:

If we choose for the constant $d = f'(x^*)$, then the convergence will be locally quadratic.

6. *Computer problem* (Heath 5.3) Implement the bisection, Newton, and secant methods for soving nonlinear 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) $x^3 - 3x^2 + 3x - 1 = 0$.

Solution:

Below is an outline of a possible solution.

- implement Newton's method, secant method, and bisection method and include your code - give one root for each method for each of parts a-d, e.g., a table

— NM — S — B

a

b

c

d

with the approximate roots.

- discuss termination criteria for each algorithm with some justification. For example, using a tolerance on the change in x in subsequent iterations, evaluating with how close $f(x)$ is to zero, or a combination of these two ideas.

- give a table for the convergence rates achieved for the roots in the first table, so another table of this form but this time with the rates:

— NM — S — B

a

b

c

d

- for one example, show the convergence, e.g., choosing NM and c, show

x0

x1

x2

x3

...

xn

and compute the convergence rate using the formula $\lim_{k->\infty} \frac{|x_{k+1}-L|}{x_k - L}$.

Note: The above is only an outline; we will be flexible in grading since the question left some details vague.

Systems of Nonlinear Equations

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

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

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

Solution:

$$
\begin{aligned}
\mathbf{f}(\mathbf{x}) &= \begin{pmatrix} x_1^2 - x_2^2 \\ 2x_1x_2 - 1 \end{pmatrix} \\
J_f(\mathbf{x}) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}
\end{aligned}
$$

To carry out one iteration we must solve the following linear system

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

We can then carry out the update as

$$
\begin{aligned}
\mathbf{x}_1 &= \mathbf{x}_0 + \mathbf{s}_0 \\
&= \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}
\end{aligned}
$$