Homework #4
Math 465/565

A few notes about this homework.

- You will be graded on your presentation as well as the content of the work. This means that all of your plots should have axis labels, a title and legends, and any other annotation that makes it clear what you are demonstrating.

- You may submit your homework to BlackBoard. Please submit a PDF of your homework.

- Students in 465 may do the problems for 565 for extra credit.

1. (**Fixed point algorithm.**) Consider the fixed-point iteration

$$x_{k+1} = g(x_k). \tag{1}$$

   applied to solving the problem $g(x) = ax + b$, for $|a| < 1$.

   (a) Show analytically that iteration (1) converges to the solution $\bar{x} = b/(1 - a)$.

   (b) Show that error $e_{k+1} \equiv x_k - \bar{x}$ satisfies $e_{k+1} = a^k e_0$.

   (c) The result in the Problem 1b is not very useful, since we don't know the initial error. However, show that we can approximate the error using

$$e_{k+1} \approx \frac{a}{a-1}(x_{k+1} - x_k)$$

   Hint: Start with $e_{k+1} = x_{k+1} - \bar{x}$. Add and subtract x_k to the right hand side to get an expression involving e_k. Then use the fact that $e_{k+1} = ae_k$ to solve for e_{k+1}.

   (d) Show that for the linear problem, the error estimate in Problem 1c is exactly equal to the true error you found in Problem 1b.

   (e) Provide a formula for how many iterations the fixed point algorithm requires to solve $g(x) = \frac{1}{10}x + 1 = x$ to a tolerance of 10^{-8}. For this problem, you do not need to code anything. Your formula will involve the initial error e_0.

2. (**Steffensen's Method**) From Problem 1, you might conclude that the fixed point method isn't very useful if it takes so many iterations to solve a simple linear problem. Steffensen's Method, given by

$$x_{k+1} = x_k - \frac{(g(x_k) - x_k)^2}{g(g(x_k)) - 2g(x_k) + x_k} \tag{2}$$

   improves on the fixed point algorithm by solving the linear problem in one step.

   (a) Show analytically that for any $g(x)$, the iteration used in Steffensen's Method satisfies

$$\lim_{x \to \bar{x}} \left(x - \frac{(g(x) - x)^2}{g(g(x)) - 2g(x) + x} \right) = \bar{x}$$

   where $\bar{x}$ satisfies $g(\bar{x}) = \bar{x}$.

   (b) Show analytically that Steffensen's iteration converges in one step for the fixed point problem $g(x) = ax + b = x$, for $|a| < 1$.

   (c) Code Steffensen's Method and solve the fixed point problem $g(x) = \frac{1}{10}x + 1 = x$. Show that you get convergence in one step. Compare this to the number of iterations required in Problem 1d.

(d) Use Steffensen's method to accelerate the convergence of a fixed-point algorithm used to solve the root-finding problem $f(x) = x^3 + x^2 - 3x - 3$. Use the function $g(x) = \sqrt[3]{3 + 3x - x^2}$. See lecture notes on the fixed-point iteration for a suitable starting guess. Use a tolerance of 10^{-8}.

- Create a loglog plot showing the error e_{n+1} verses e_n for both the fixed-point method and the accelerated fixed-point method (Steffensen's algorithm). Compute the slope of each plot and show that the fixed-point algorithm is linearly convergent, whereas Steffensen's method is quadratically convergent.

Note: In Homework #3 (Problem 3), we showed that for a superlinear scheme, the error can be approximated as

$$|e_{k+1}| = |x_{k+1} - \bar{x}| \approx |x_{k+1} - x_k|$$

3. (**LU Decomposition**). In class, we used "elementary matrices" to derive the LU decomposition. An example of a 4×4 elementary matrix is

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\ell_{21} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where ℓ_{21} is a multiplier computed during the elimination process.

Consider the matrix

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} \tag{4}$$

(a) Show that

$$E_{41} E_{31} E_{21} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\ell_{21} & 1 & 0 & 0 \\ -\ell_{31} & 0 & 1 & 0 \\ -\ell_{41} & 0 & 0 & 1 \end{bmatrix}$$

(b) Choose multipliers ℓ_{j1} so that multiplying A on the left by the product $E_{41} E_{31} E_{21}$ zeros out the entries below a_{11}.

(c) Show that the inverse of $E_{41} E_{31} E_{21}$ is

$$(E_{41} E_{31} E_{21})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \ell_{21} & 1 & 0 & 0 \\ \ell_{31} & 0 & 1 & 0 \\ \ell_{41} & 0 & 0 & 1 \end{bmatrix}$$

Hint : Use the formula $B^{-1} = C^T / \det(B)$ where C is the co-factor matrix of B.

(d) From the above steps, derive the LU decomposition of a general 4×4 matrix. Describe how the entries of L are found and describe how the matrix U is formed.

(e) Find the LU decomposition of the matrix in (4). Report your results for both L and U and show that $LU = A$.

4. (**Partial Pivoting**) The basic LU decomposition algorithm may fail if it hits a zero pivot, or can lead to significant errors for very small pivots.

(a) The code `lu_bug_pp.m` on the course website implements the partial pivoting algorithm. Find the three bugs in the code to implement partial pivoting correctly.

(b) The code `lu_solve.m` on the course website carries out forward and backward substitutions needed to solve $A\mathbf{x} = \mathbf{b}$ using LU decomposition. Modify this to use L and U obtained from the partial pivoting algorithm.

(c) Solve the linear system $A\mathbf{x} = \mathbf{b}$ for A and $\mathbf{b}$ given by

$$A = \begin{bmatrix} 10^{-10} & 4 \\ 2 & 1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

using LU decomposition both without and with partial pivoting. Compute the true solution and use this to compute the errors using the LU decomposition both without and with partial pivoting. Show that the results using partial pivoting are significantly smaller than the errors obtained without partial pivoting.

(d) (**Math 565**). At what step in the LU decomposition process without partial pivoting do errors (due to catastrophic cancellation) enter the calculations?

5. (**Jacobi Method**). Assume that the following 2×2 matrix A is strictly diagonally dominant.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

(a) Show that $\rho(I - D^{-1}A) < 1$. From this conclude that the Jacobi iteration will converge.

6. (**Gauss-Seidel Method**). Write a code that implements the Gauss-Seidel method to solve a 4×4 matrix equation $A\mathbf{x} = \mathbf{b}$ for a symmetric positive definite matrix A and right hand side vector $\mathbf{b}$ of your choice.

(a) Solve the system to a tolerance of 10^{-8}. Report on the number of iterations required to solve your system, as well as the final solution.

To generate your matrix A, use the `rand` function in Matlab (or the equivalent in Python) to create a matrix B with random entries. Then set $A = B^T B$. Show that A is symmetric positive definite. **Hint:** Show that $\mathbf{x}^T A\mathbf{x} > 0$ for any vector x.

7. (**Math 565**). The iteration step in a general stationery iterative method for solving the linear system $A\mathbf{x} = \mathbf{b}$ can be written as

$$\mathbf{x}_{k+1} = (I - M^{-1}A)\mathbf{x}_k + M^{-1}\mathbf{b}$$

(a) Show that if $M = A$, the iteration converges in one step.

(b) Show that $A\mathbf{e}_k = -\mathbf{r}_k$, where $\mathbf{r}_k = b - A\mathbf{x}_k$, $\mathbf{e}_k = \mathbf{x}_k - \bar{x}$, and $\bar{x}$ solves $A\mathbf{x} = \mathbf{b}$ exactly.

(c) Show that the iteration for the error $\mathbf{e}_k$ is given by

$$\mathbf{e}_{k+1} = (I - M^{-1}A)\mathbf{e}_k$$

(d) Show that

$$\| \mathbf{e}_k \| \leq \| I - M^{-1}A \|^k \| \mathbf{e}_0 \|$$

where $\| \cdot \|$ is a matrix norm. **Hint:** Use the following properties of matrix norms:

   i. $\| AB \| \leq \| A \| \| B \|$ for matrices A and B.

   ii. $\| A\mathbf{x} \| \leq \| A \| \| \mathbf{x} \|$ for matrix A and vector $\mathbf{x}$.

(e) Show that we can estimate the number of iterations needed for a stationery iterative method to achieve a desired error tolerance ε using the formula

$$k \geq \frac{\log(\varepsilon)}{\log \| I - M^{-1}A \|}.$$

(gs_iterations.png)

Figure 1: Plot of iterations required by a Gauss-Seidel method for random, 4×4 SPD matrices.

(f) The plot in Figure 7 shows the number of iterations required to solve a series of $4x4$ symmetric positive definite systems using the Gauss-Seidel method. From the plot, you see that generally, the number of actual iterations required increases as the spectral radius of the matrix $I - M^{-1}A$ approaches 1. But the estimated number of iterations seems to underestimate the actual number for many larger values of $\rho(I - M^{-1}A)$. What is the reason for this?

8. (**Math 565**). The conjugate gradient algorithm improves on the Steepest Descent Algorithm by choosing *A-conjugate* directions rather than just orthogonal directions.

(a) Show analytically that the update to the residual $\mathbf{r}_{k+1}$ in the Conjugate Gradient algorithm is equal to $b - A\mathbf{x}_k$.

(b) Show that the directions $\mathbf{p}_k$ satisfy $\mathbf{p}_{k+1}^T A \mathbf{p}_k = 0$. That is, the directions are *A-conjugate*.

(c) Code the conjugate gradient algorithm and use it to solve a 100×100 linear positive definite system $A\mathbf{x} = \mathbf{b}$. Construct A and $\mathbf{b}$ using random entries.

(d) Compare the number of iterations needed using CG to the number required by Gauss-Seidel.