# Numerical Methods I

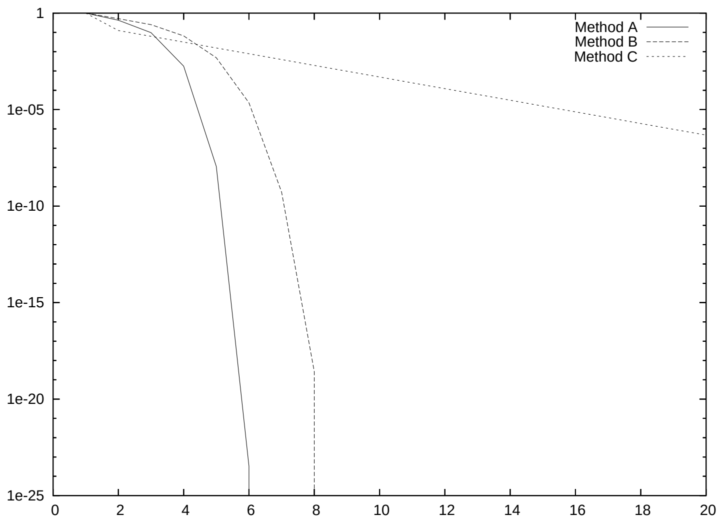
Midterm Exam

October 18, 2005

1. The following graphs compare the performance of the bisection, Newton, and secant root finding method for two different functions.

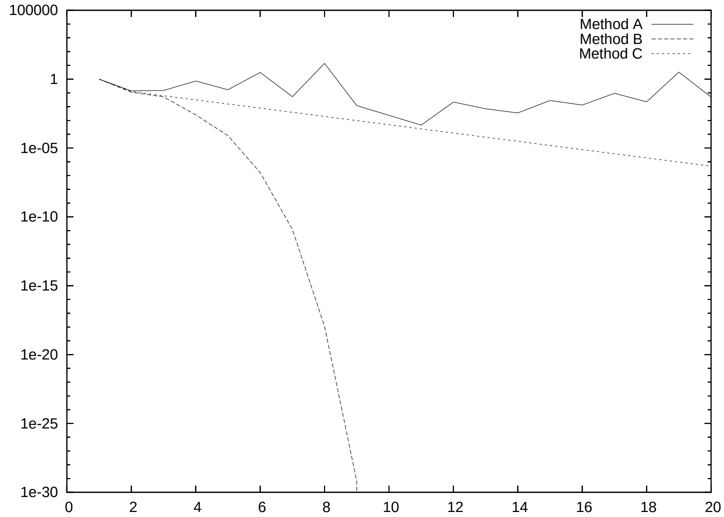
   The first test is for finding the root at $x = 0$ of the function

   $$f(x) = x + x^2 \sin x\,.$$




   The second test is for finding the root at $x = 0$ of the function

   $$g(x) = \begin{cases} x + x^2 \sin \frac{1}{x} & \text{for } x \neq 0\,, \\ 0 & \text{for } x = 0\,. \end{cases}$$




(a) Identify methods A, B, and C. (The labeling in the two graphs is the same.) Comment briefly on each match.

(b) **Extra credit.** Analyze why one of the methods fails so badly for finding the root of $g$ while the others do just fine.

(10+*10*)

2. The secant method, which can be written as

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

involves subtraction of almost equal numbers as the sequence approaches a limit. Yet, the method proves very robust in practice. We shall thus analyze the propagation of rounding errors for the secant method.

(a) Assume that all operations are exact *except* for the evaluation of $f$. Why can we expect that this simplified analysis still captures the essential error behavior of the secant method?

(b) Show that one step of the secant method in floating point is approximately

$$x_{k+1} \approx x_k - (x_k - \xi)\,\frac{1}{1 - \delta\,\dfrac{x_{k-1} - \xi}{x_k - x_{k-1}}}\,, \qquad (*)$$

where $\xi$ denotes a simple root of $f$, for some $\delta \ll 1$.
Hint: It is easiest to write

$$\mathrm{fl}\left(\frac{f(x_{k-1})}{f(x_k)}\right) = \frac{f(x_{k-1})}{f(x_k)}\,(1+\delta)\,,$$

then use truncated Taylor series to estimate the quotient.

(c) How does (*) imply that the secant method is robust to rounding errors?

(5+10+5)

3. Compute the $LU$ factorization without pivoting of the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 8 & 14 \end{pmatrix} .$$

(15)

4. (a) Prove that if $A \in M(n \times n)$ is positive definite, then all diagonal entries must be positive.

   (b) Let $A \in M(n \times n)$ be symmetric and positive definite, with $A = L + D + R$ denoting the splitting of $A$ into its lower left triangular, diagonal, and upper right triangular parts (as in the derivation of the Jacobi and Gauss–Seidel algorithms). Show that the so-called *symmetric Gauss–Seidel preconditioner*

$$B_{\mathrm{SGS}} = (D + R)\, D^{-1}\, (D + L)$$

   is indeed symmetric and positive definite.

(10+10)

5. Let

$$A = \begin{pmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{pmatrix}$$

   with $|\varepsilon| < 1$.

   (a) Show that the Jacobi method for solving $A\boldsymbol{x} = \boldsymbol{b}$ does not converge for any choice of right hand side $\boldsymbol{b}$ and starting vector $\boldsymbol{x}_0$.

   (b) Suggest a "Jacobi method with pivoting" so that the modified Jacobi iteration converges for the given matrix $A$.

(10+10)

6. (a) State the definition of the norm of a matrix induced by a vector norm; state the definition of the condition number of a matrix.

   (b) Let $A \in M(n \times n)$ be invertible and $\boldsymbol{b} \in \mathbb{R}^n$. Let $\boldsymbol{x}^*$ denote the exact solution to the linear system $A\boldsymbol{x}^* = \boldsymbol{b}$ and let $\boldsymbol{x}$ denote an approximation to $\boldsymbol{x}^*$. Show that

$$\frac{\|\boldsymbol{x}^* - \boldsymbol{x}\|}{\|\boldsymbol{x}^*\|} \leq \kappa(A)\, \frac{\|\boldsymbol{b} - A\boldsymbol{x}\|}{\|\boldsymbol{b}\|}$$

   where the vector norm $\|\cdot\|$ is arbitrary, and the condition number is defined with respect to the induced matrix norm.

(5+10)