Name:	ID#:

CSE 4345 / CSE 5315 - *Computational Methods*

Practice Problems - 1- Solution - Fall 2011

Closed book, 2 pages of notes

Date:

Problems marked with * are required only for students of CSE 5315 but will be graded for extra credit for students of CSE 4345.

1. Briefly explain what the following types of errors measure and how they are affected by the choice of algorithm and its implementation.

   a) Data propagation error.

      Data propagation error measures the amount of error in the solution that is due to the imprecison (or error) in the input data. It assumes a perfect algorithm and perfect computations and is therefore not affected by the choice of algorithm or the implementation.

   b) Computation error.

      Computation error measures the error that is introduced by the algorithm and the imprecisions in the computing system and the implementation. It effectively measures the difference between what a perfect algorithm on a perfect computer would have computed with the input data and what the actual algorithm and implementation computes.

2. Briefly discuss the difference between stability and sensitivity. What do they measure and what are they influenced by ?

   Sensitivity is a measure of the problem to be solved. It measures how strongly data error would influence solution error for the perfect algorithm to solve the problem.

   Stability measures the algorithm and implementation. It evaluates how prone the algorithm and implementation is to introducing error in the solution of the problem beyond what sensitivity predicts would be unavoidable.

3. The equation $(2 + x^2) - (2x/sin(x))$ suffer from loss of significance (cancellation) for $x \to 0$. Provide a reformulation that avoids this problem.

   $\frac{(x^2-2x)+2}{sin(x)}$

4. For each of the following fixed point functions determine whether fixed point iteration would converge near the indicated solution.

   a) $g(x) = \frac{2}{x^2}$, Fixed point: $\sqrt[3]{2}$

   $g'(x) = -4\frac{1}{x^3}$ at root is $-4\frac{1}{2} = -2 \leq -1$ and will therefore not converge.

   b) $g(x) = 6 - x^2$, Fixed point: 2

   $g'(x) = -2x^3$ at root is $-16 \leq -1$ and will therefore not converge.

5. Illustrate the operation of the secant method for root finding for a single equation in one variable by showing two iterations on the function $f(x) = x^2 - 2$ starting with the two initial points $x_0 = 4, x_1 = 3$.

   $f(x_0) = 14$, $f(x_1) = 7$

   First iteration:

$x_2 = x_1 - f(x_1)\frac{x_1-x_0}{f(x_1-f(x_0)} = 3 - 7\frac{3-4}{7-14} = 3 - 1 = 2$
$f(x_2) = 2$

Second iteration:
$x_3 == x_2 - f(x_2)\frac{x_2-x_1}{f(x_2-f(x_1)} = 2 - 2\frac{2-3}{2-7} = 2 - \frac{2}{5} = 1.6$
$f(x_3) = 3.56 - 2 = 1.56$

6.* Briefly compare some of the key attributes of the interval bisection method and Newton's method for root finding, detailing their respective advantages and disadvantages. Also briefly discuss how this motivates hybrid methods that combine them.

Interval Bisection does not require any information about the function except for an initial bracket and that the function is continuous (in particular it does not require any information about derivatives of the function). In addition it is guaranteed to converge to a correct solution given that one exists within the initial bracket and that the function is continuous. However, the convergence rate is only linear, making it relatively slow despite the very low computational overhead (one function evaluation per iteration).

Newton's Method requires that the function for which the root is to be found is continuous and differentiable. When Newton's method converges its convergence rate is quadratic, making it a relatively fast iterative solution method. However, it will only converge when started from an initial point that is sufficiently close to the root. The computation of Newton's method requires a function evaluation and an evaluation of the derivative for every iteration (this additional overhead is more than made up for by the higher convergence rate).

The motivation of hybrid methods comes from the desire to combine the guaranteed convergence of methods like the interval bisection method to converge (making them safe solution techniques) with the much higher convergence rates of and thus faster solution finding of other methods such as Newton's method, Inverse Quadratic Interpolation, or the Secant method.

7. Perform Gaussian Elimination on the following system of equations.

$$2x + y + 3z = 0$$
$$6x + 4y + 7z = 2$$
$$4x + 4y + 7z = 14$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 7 \\ 4 & 4 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 14 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 7 \\ 4 & 4 & 7 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 2 \\ 14 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 2 \\ 14 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix}$$

$z = 5$, $y = (2 - (-2 * 5))/1 = 12$, $x = (0 - 3 * 5 - 12)/2 = -13.5$

8. The Multivariate Newton method solves a system of nonlinear equations by iteratively solving a sequence of linear systems of equations. List the basic operation steps of the Multivariate Newton method.

   Multivariate Newton first computes the Jacobian of the system at the current point, x. Then it solves the system $J(x)s = -f(x)$ to effectively determine the step size $s = -J(x)^{-1}f(x)$ while avoiding the matrix inversion. It then updates the point for the next iteration to $x + s$.

9. Solving systems of linear equations, $Ax = b$, only has a solution if the rank of the matrix A is less than or equal to n (the number of variables, x), and

has a unique solution only if the rank is equal to n. Under what conditions does the linear least squares problem, $Ax \cong b$, have a solution and under which conditions is the solution unique ?

As opposed to the root finding problem for systems of equations, the linear least squares problem always has a solution.

The linear least squares solution is unique if the rank of matrix A is equal to n.

10. Consider the linear least squares data fitting problem where a linear function, $f(x) = \alpha x$ is to be fit to the data points, $\{(x, y)_i\} = \{(1, 2), (2, 3), (3, 4)\}$, such that the square difference (residual) between $f(x)$ and y is minimized.

    a) Transform the above problem (and data) into the Normal Equations for the linear least squares problem.

       First we have to formulate the problem. Given that we are fitting αx we get:

$$A = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$y = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

$$A\alpha \cong y$$

       From this we get the normal equations:

$$A^T A\alpha = A^T y$$

$$(1, 2, 3)\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\alpha = (1 + 4 + 9)\alpha = (1, 2, 3)\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 2 + 6 + 12$$

    b) Derive the solution using Normal Equations.

       To solve the system we have to solve the linear equations above:

$$14\alpha = 20$$
$$\alpha = \frac{20}{14} = \frac{10}{7}$$

11. Consider the following linear least squares approximation problem where a function, $f(\alpha, x)$ that is linear in the parameters α is to be fitted to a set of data points such that the square residual (difference between $f(\alpha, x)$ and y of the data) is minimized.

$$f(\alpha, x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 \quad , \quad Data\ (x, y) : \ \{(0,1), (1,3), (2,4), (3,5)\}$$

a) Provide the augmented system formulation for the problem.
We can first rewrite the problem in matrix form as $A\alpha \cong b$:

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

$$b = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \end{pmatrix}$$

Now we can derive the augmented system formulation $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} r \\ \alpha \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$ as:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 1 & 1 & 3 & 9 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 1 & 4 & 9 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

b) Discuss some of the differences (in particular in terms of sensitivity and complexity) between the normal equation solution and the solution using the augmented formulation.

The normal equation solution to the linear least squares problem consists of solving the system $A^T Ax = A^T b$. As a result, it reduces the complexity of the solution to (besides the matrix multiplication) solving a $n \times n$ linear equation solving problem. The augmented system method, on the other hand, requires the solution of an $(m+n)\times(m+n)$ linear equation solving problem. As a result, the computational complexity of the normal equation solution is generally substantially lower than the one of the augmented system method.

On the other hand, the sensitivity of the solution to the normal equations is determined by the condition number of the matrix $A^T A$ and is thus the square of the condition number of the original system matrix A. The augmented system model does not require the multiplication of the system matrix and thus results in a significantly better conditioned system. As a result, the augmented system model is a significantly more stable solution algorithm in particular for large linear least squares problems. This is further increased by the better selection method for pivots when solving the resulting system of equations (e.g. using LU factorization).

Overall, the normal equation solution is significantly less computationally complex while the augmented system method results in a better conditioned solution system and thus represents a more stable solution approach for the linear least squares problem.