Name:	ID#:

CSE 4345 / CSE 5315 - *Computational Methods*

Practice Exam 3- Sample Solution - Fall 2011

Closed book, 6 pages of notes

Date: Dec. 13 2011, 2:00 pm - 4:30 pm

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

Eigenvalues and Singular Values

1. Inverse iteration computes the smallest eigenvalue and the corresponding eigenvector by computing the largest eigenvalue of A^{-1}. Show one iteration of normalized inverse iteration for the following transformation matrix, A, and starting vector, x_0.

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Compute y_1:

$$\begin{aligned} & Ay_1 = x_0 \\ & \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} y_{1,1} \\ y_{1,2} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ & \Rightarrow \begin{pmatrix} 1 & 4 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} y_{1,1} \\ y_{1,2} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & \Rightarrow y_{1,2} = \tfrac{1}{5}\ ,\ y_{1,1} = \tfrac{1}{5} \end{aligned}$$

Compute x_1:

$$x_1 = \frac{y_1}{|y_1|} = \frac{y_1}{\frac{1}{5}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

2. When computing eigenvectors and eigenvalues, a number of transformations can be applied to the original matrix A without loosing the ability to compute its eigenvalues and eigenvectors.

   a) Shift (the subtracting of a constant from the diagonal terms, i.e. $B = A - \sigma I$) and Similarity Transformations (i.e. transformations of the form $B = T^{-1}AT$) are important problem transformations. Indicate how these transformations influence the eigenvalues and eigenvectors and discuss how they can be used to simplify the eigenvalue problem.

   Shift transfomation: The shift transformation maintains the eigenvectors but changes the corresponding eigenvalues. In particular, for every eigenvalue of the shifted system we have $(A - \sigma I)x_i = \lambda x_i$ and therefore we have $Ax_i = \lambda x_i + \sigma I x_i = (\lambda + \sigma)x_i$.

   The similarity transform maintains the eigenvalue but changes the eigenvector. In particular for an eigenvalue/eigenvector pair of the transformed system, $Bx_i = T^{-1}ATx_i = \lambda x_i$ and therefore for the original system, $A(Tx_i) = \lambda(Tx_i)$ and therefore the corresponding eigenvector of A is Tx_i

   Since they keep the eigenvalues and eigenvectors computable, they can be used to simplify the system, A into a simpler one or to accelerate the convergence of an iterative algorithm. In particular, the shift operation can be used to change the convergence rate of power iteration or inverse iteration since this convergence rate depends on the ratio of the largest and second largest, or on the ratio of the smallest and second smallest eigenvalue, respectively. Changing both of these eigenvalues by σ allows to change this ratio and thus the convergence rate of the iteration. Similarity transforms serve a different purpose. In particular they are efficient ways to convert the system matrix into diagonal (or block diagonal) or triangular form which makes solving for the eigenvalues and eigenvectors much simpler. Furthermore, they can be used for deflation to compute additional eigenvectors and eigenvalues.

   b) Eigenvalues and eigenvectors of diagonal and triangular matrices are significantly simpler to compute than for general matrices. Discuss why this is the case and how they can be determined.

   Diagonal or triangular form are simpler because in both cases the eigenvalues are the elements on the diagonal of the matrix. In the diagonal case, the eigenvectors are simply the elementary unit vectors e_i. In the case of a triangular matrix, eigenvectors still have to be computed but can easily be obtained using $(A - \lambda_i I)x = 0$.

3. Power iteration and inverse iteration are limited to computing the largest and the smallest eigenvalue (and corresponding eigenvector).

   a) Both, Deflation and Simultaneous Iteration can be used to address this limitation and determine more than just the largest and smallest eigenvectors. Discuss some of the differences of these methods and some of their advantages and disadvantages.

   Deflation computes the different eigenvalues (and corresponding eigenvectors) iteratively, starting with either the largest or the smallest eigenvalue. Once this is computed, the system is transformed using a similarity transformation and the component corresponding to the already computed eigenvector is removed, leaving a system that only contains the remaining eigenvalues and eigenvectors (and has dimension $n - 1$).

   In simultaneous iteration, multiple vectors are iterated simultaneously to derive the corresponding eigenvectors and eigenvalues simultaneously. In order to achieve this in a numerically stable way it is necessary that the different vectors are kept in a form that keeps them well conditioned.

   An advantage of simultaneous iteration is that it does not require an iterative deflation process and can operate solely on the original matrix A. As a result, it directly results in the eigenvalues (and potentially eigenvectors). However, it is numerically less stable if the eigenvectors are to be directly calculated and thus is usually performed in a way that does not preserve eigenvectors and thus requires them to be computed separately.

   b)* Simple simultaneous iteration becomes ill-conditioned relatively fast. Discuss why this is the case and how QR factorization in QR iteration addresses this problem.

   Simultaneous iteration in its most basic form is ill-conditioned since all iterated vectors tend to converge to the dominant eigenvector, thus over time making the contributions of the other eigenvectors disappear within the rounding errors of the computer. To address this, it is necessary to explicitly maintain the individual eigenvectors in a separated form. Applying QR factorization in each iteration achieves this by decomposing the matrix of vectors into an orthogonal transform (which changes the eigenvectors but maintains the eigenvalues) and a triangular matrix which explicitly maintains the separate parts of the vectors. Only iterating on the latter part, the vectors are effectively kept separate and individually "rescaled" in each iteration to maintain k separate eigenvectors and thus ensuring that the system does not become ill-conditioned.

4. Singular values are important for a number of important problems related to the transformation A and the space it describes. List at least four of the characteristics of the transformation that can be answered using the result of singular value decomposition.

   Null space determination: The number of singular values equal to 0 represent the dimensionality of the null space of A and the corresponding right singular vectors span the null space.

   Span of the matrix: The number of non-zero singular values represents the dimensionality of the span of A and the corresponding left singular vectors span the corresponding space.

   Euclidean matrix norm determination: The maximum singular value is equal to the Euclidean matrix norm of A.

   Linear least squares solution: The least squares solution to the linear system represented by A can be computed directly from the singular values and the left and right singular vectors

Randomness and Monte Carlo Methods

5. Randomized algorithms can be effective ways to approximate solutions to highly complex problems. To apply them it is important to be able to generate random numbers. There are differnt types of random numbers based on their general properties. List 3 different types of random numbers and indicate one way to generate them for each of them.

   True random numbers: True random numbers have to be generated through a truly random process, such as radioactive decay, variations in crystal oscillations, etc.. The are not predictable, not repeatable, uncorrelated, uncompressable, and non-repeating.

   Pseudo-random numbers: Pseudo-random numbers are generated algorithmically. They appear random and unpredictable and generate number sequences that have appropriate statistical properties for random number sequences (low sequence correlation, independently distributed, etc). However, given the algorithm they are predictable and repeatable (which is very useful for debugging purposes).

   Quasi-random numbers: Quasi-random numbers are generated algorithmically and are aimed at applications that only require random distribution properties but no random sequence properties. Generally their overall distribution appears random but they have high sequence correlations. They are useful since they generate the overall distribution pattern even for small numbers of samples as they do not show any ”clumping” of samples early on.

6. Pseudo-random number generators can be very sensitive to the parameter settings.

   a) Briefly discuss why this is and what some of the effect of bad parameter choices can be.

   Most pseudo-random number generators are based on simple mathematical operations which will only lead to number sequences with the desired statistical properties if the parameters are chosen carefully. For example, in a congruential random number generator, choosing the base such that it has many common factors with the multiplier significantly reduces the maximum length of a non-repeating sequence that can be generated. Similarly, choice of a small multiplier will generally lead to sequences that show very high sequence correlations.

   b)* For the following, simple settings of a congruential random number generator, compute the first 5 pseudo-random numbers for a seed of 4. Base $m = 100$, Multiplier $a = 17$, Shift $c = 0$.

   $68, 56, 52, 84, 28, ...$

Errors

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

   Sensitivity is a function of the problem and independent of the algorithm used to solve it or the numerical precision of the computer used. It basically represents the degree to which the problem itself will magnify (or shrink) input errors.

   Stability is a function of the algorithm used and addresses how much error the algorithm introduces in light of the problem and the the numerical precision of the computer. A stable algorithm will always produce solutions that produce errors on the same order as indicated by the sensitivity (conditioning) of the problem. Unstable algorithms will either produce much larger errors or even fail on some problems and diverge.

8. For the following equations which suffer from a loss of significance (cancellation) for $x \to 0$, provide a reformulation that avoids this problem.

   a) $\frac{(x^2-4)^2-16}{x^4}$
   $1 - \frac{8}{x^2}$

   b) $\frac{\sqrt{x+9}-3}{2x}$
   $\frac{1}{2(\sqrt{x+9}+3)}$

Root Finding

9. Briefly explain the operation of the interval bisection method for equation solving with single equations in one variable. Illustrate the operation by performing the first two iterations of the bisection method on the function $f(x) = x^2 - 2x + 1$ starting with the bracket $[0, 2]$. Show the new points, their values, and the brackets for each of the four iterations.

   Bisection maintains a bracket and in each iteration halves the bracket by evaluating the middle point of the bracket and determining which half of the interval would yield a bracket (by evaluating whether the corresponding end point and the point in the middle of the bracket have opposite signs - or are equal to 0 in which case a solution has already been found.

   For the problem above the initial interval is not actually a bracket (as both endpoint function values are larger than 0). Given a modified function, $f(x) = x^2 - 3x + 1$, $[0, 2]$ becomes a bracket so this will be used here.

   $f(0) = 1, f(2) = -1$
   Middle point: $c = 1, f(1) = -1$
   Since $f(1) * f(0) < 0$, the new bracket is $[0, 1]$

   Middle point: $c = 0.5, f(0.5) = -0.25$
   Since $f(0.5) * f(0) < 0$, the new bracket is $[0, 0.5]$

   Middle point: $c = 0.25, f(0.25) = \frac{5}{16}$
   Since $f(0.25) * f(0.5) < 0$, the new bracket is $[0.25, 0.5]$

10. 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.

    For the multivariate Newton method the Jacobian, $J_f(x)$, of the system of equations has to be calculated. Using this, the step for the Newton method can be computed as the solution to the linear system, $J_f(x)s = -f(x)$. This step vector, s is the added to the previous estimate, x. Effectively, multiplication with the inverse of the Jacobian takes the place of the division by the derivative in the one-dimensional Newton method.

Interpolation

11. There are generally an infinite number of possible piecewise quadratic or cubic interpolation functions for a set of data points. Provide a piecewise cubic interpolation for the data points $(0,1),(1,1),(2,-2)$. (Note that it does not have to be differentiable or smooth.)

    Since piecewise interpolation only uses 2 data points for each piece, the simplest cubic interpolation uses coefficients of o for the second and third base monomials, resulting in the following piecewise interpolation:

    For $0 \leq x < 1$*:* $p(x) = 1$
    For $1 \leq x \leq 2$*:* $p(x) = -3(x-1)+1$

12.* Interpolate the data points of problem 17. using polynomial interpolation with monomial basis functions.

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}$$
$$\Rightarrow a_3 = -\tfrac{3}{2}, a_2 = \tfrac{3}{2}, a_1 = 1$$
$$\Rightarrow p(x) = 1 + \tfrac{3}{2}x - \tfrac{3}{2}x^2$$

Optimization

13. First and second-order optimality conditions are necessary conditions that allow to characterize optima. In particular, the first order optimality condition states that in a continuous differentiable function, the first derivative of the function at an optimum has to be 0.

    a) Fulfilling the first-order optimality condition is not sufficient for a point to be identified as an optimum. Discuss why this is the case.

    If the first derivative is 0, the point can still be a saddle point and thus not be a local minimum (there is a neighboring point that has a lower function value).

    b) How does the second order optimality condition (i.e. information about the Hessian - the second derivatives of the function) address this ?

    The second order condition measures the second derivative (or Hessian) and indicates the curvature. at a saddle, the curvature is 0 while it is positive at a local minimum.

14. In multi-dimensional optimization, a variety of optimization methods can be used that are variations on Newton's method. Briefly discuss the differences between Newton's method and quasi-Newton methods and list some of the advantages of quasi-Newton methods ?

    Newton's method uses the actual Hessian of the function to determine the step size while Quasi-Newton methods use an approximation of it that is updated in each iteration. As a result, quasi-newton methods are significantly less complex as the determination of the Hessian is a very complex operation. In addition, computation of the hessian is a relatively sensitive operation, allowing Quasi-Newton methods also to be often more stable than Newton's method.

15. Linear programs are a special type of constrained optimization problem. Briefly discuss what makes a linear programming problem easier to solve than other, more general constrained optimization problems.

    Linear programs are constrained optimization problems in which the objective function as well as the constraints are linear. This makes them easier to solve because the constraints describe a convex feasible region. In addition, the linear objective function means that all points with the same objective function value lie on a hyperplane with two differnt values forming parallel hyperplanes. As a result, the objective function changes linearly with the distance from the origin of the system, meaning that if the unconstraint minimum of the objective function is not inside the feasible region, it has to lie on the intersection of the value hyperplane and the boundary of the feasible region that is closes to the unconstrained minimum. As a result, it has to occur at an intersection point of multiple constraints (i.e. a vertex of the constraint surface). This allows methods like the simplex method to be used that only evaluate vertices in constraint space.

16. Define the Lagrange function for the following constrained optimization problem with equality constraints.
    Objective function: $f(x, y) = x^2 + y^2 - 2xy + 7$, Constraints: $g_1(x, y) = 7x + 2y = 0$, $g_2(x, y) = x^2 - xy = 0$
    $L(x, y, \lambda_1, \lambda_2) = f(x, y) + \lambda_1 g_1(x, y) + \lambda_2 g_2(x, y)$