CS 4220: Project Assignment

Due: Tuesday, May 18, 2010 .

The project assignment is worth 20% of the course grade and will be graded on a 0-to-25 scale. Clarity of code (5pts), clarity of write-up (5pts), information content of code output (5pts), algorithmic inventiveness (5pts), and integration of course concepts (5pts).

General Requirements

The idea behind the project is for you to choose a matrix problem or a nonlinear problem and solve it. The submission is to have two parts:

Code. You must write a function that solves the chosen problem and a script that showcases the function on a small number of illustrative examples. This is to be submitted electronically in a single function mfile `YourName.m`. Commenting must be good enough so that I can appreciate what you have done. The execution of `YourName()` should take less than 3 minutes and be informative. You have seen enough test scripts over the term to design a good one yourself.

Write-up. A physically-submitted two-pager that clearly describes the problem you solved and/or what you discovered. Describe anything that might be special about the the matrices, objective functions, gradients, Jacobians, or Hessians that arise. OK if you need a little more than 2-pages, but it better be good if you go past that limit! Do not forget to cite any references that you used.

I am happy to offer suggestions as you develop your solutions. But that won't be too worthwhile unless you begin early enough.

The options...

0. (Your Choice) Must be pre-approved before the last day of classes. A typical situation might be that you solve a sparse linear equation, least squares, or eigenvalue problem using MATLAB's sparse matrix tools or a nonlinear problem of special structure. The demo script could shed light on problem sensitivity and/or time complexity.

1. (Ellipse Fitting) Write a function `E = BestEllipse(x,y)` that returns the ellipse that best fits the points defined by the column n-vectors `x` and `y`. Define an appropriate structure for `E`. For an objective function use

$$\phi(E) \;=\; \sum_{i=1}^{n} \mathrm{dist}(E, x_i, y_i)$$

where $\mathrm{dist}(E, a, b)$ is the distance from (a, b) to the nearest point on E. Focus on the inner/outer iteration issue by noting that the evaluation of dist is itself iterative. Approximate evaluations may be OK at the start of the process. Towards the end, the nearest ellipse point for each (x_i, y_i) probably doesn't move much.

2. (Broyden on Linear Systems) What happens when Broyden's method is applied to find a zero of $F(x) = Ax - b$ where $A \in \mathbb{R}^{n\times n}$ is nonsingular and $b \in \mathbb{R}^n$? Set the initial guess to be $x_0 = 0$ and the initial approximate Jacobian to be $B_0 = I_n$. Does the condition of A have a bearing on what happens? What does $\| x - x_k \|$ look like? What if $A = I + YZ^T$ when $Y, Z \in \mathbb{R}^{n\times r}$ with $r < n$? Not necessary to implement the $O(n^2)$ QR update strategy that takes you from $B_k = Q_k R_k$ to $B_{k+1} = Q_{k+1} R_{k+1}$.

3. (Newton Fractals) Look up Newton Fractals in Wikipedia. Write a script that graphically demos the idea. Might want to vectorize the Newton iterations for efficiency. Write about termination criteria and special features of the fractal set associated with $f(z) = z^r - 1$. What if the secant method is used instead?

4. (Golden Section Search) What happens if Golden Section search is applied to a function that is not unimodal on the initial interval? Does it "usually" converge to a local minimum? Is there a way to a apply it multiple times so as to emerge with something meaningful? Develop a point-and-click environment that displays a function across an interval and uses GS search to get all the local minima.

5. (Polygon Smoothing in 3D) Suppose $(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)$ are given points with $\| x \|_2 = 1$, $\| y \|_2 = 1$, and $\| z \|_2 = 1$. Play with the following script and offer a set of explanations.

```
n = input('Enter n');
x = randn(n,1); x = x - sum(x)/n; x = x/norm(x);
y = randn(n,1); y = y - sum(y)/n; y = y/norm(y);
z = randn(n,1); z = z - sum(z)/n; z = z/norm(z);
for k = 1:20*n
  x = (x + [x(n);x(1:n-1)])/2; x = x/norm(x);
  y = (y + [x(n);x(1:n-1)])/2; y = y/norm(y);
  z = (z + [z(n);z(1:n-1)])/2; z = z/norm(z);
  plot3(x,y,z)
end
```

For background, read `http://www.cs.cornell.edu/cv/ResearchPDF/EllipsePoly.pdf`. Embellish the above script so that the graphics is more illuminating.

6. (Exponential of an Intensity Matrix) The (m, m) Pade approximation to the matrix exponential is given by

$$e^A \;=\; \sum_{k=0}^{\infty} \frac{A^k}{k!} \;\approx\; \left(\sum_{k=0}^{m} c_k A^k \right)^{-1} \left(\sum_{k=0}^{m} c_k (-A)^k \right) \;=\; R_{mm}(A)$$

where $c_k \;=\; ((2m-k)!m!)/((2m)!k!(m-k)!)$. The quality of the approximation deteriorates if $\| A \|$ is too large so in practice one uses the approximation

$$e^A \;\approx\; \left(R_{mm}(A/2^j)\right)^{2^j} \;=\; F(A, m, j)$$

where j is the smallest positive integer so $\| A \|_1 / 2^j \leq 1$.

In this problem you are to explore the use of this approximation when A is an intensity matrix. We say that $A \in \mathbb{R}^{n \times n}$ is an intensity matrix if its diagonal entries are strictly negative, its off-diagonal entries are strictly positive, and its column sums are zero. It can be shown that e^A is stochastic, i.e., entries are all positive and column sums are 1.

Write a function that computes $F(A, m, j)$. The demo script should shed light on the accuracy of $F(A, m, j)$ as a function of m. Other questions. How does e^A change if the k-th column of A is multiplied by a scalar μ? How close is $F(A, m, j)$ to being stochastic? Can you prove that its entries are nonnegative?

7. (Eigenvalues of a Big Matrix)

The Kronecker product of two matrices $B \in \mathbb{R}^{m_1 \times m_2}$ and $C \in \mathbb{R}^{n_1 \times n_2}$ is denoted by $B \otimes C$ and is defined by

$$B \otimes C \;=\; \begin{bmatrix} b_{11}C & \cdots & b_{1,m_2}C \\ \vdots & \ddots & \vdots \\ b_{m_1,1}C & \cdots & b_{m_1,m_2}C \end{bmatrix}$$

It can be shown that if $x \in \mathbb{R}^{m_2 n_2}$ and $y = (B \otimes C)x$, then $Y = CXB^T$ where $Y = \text{reshape}(y, n_1, m_1)$ and $X = \text{reshape}(x, n_2, m_2)$. Write a function that can be used to compute the k largest eigenvalues of

$$A \;=\; B_1 \otimes C_1 \;+ \cdots +\; B_d \otimes C_d$$

where all the B_i are m-by-m and all the C_i are n-by-n. Make effective use of `eigs`.

8. (Big Vector Approximation) Given a unit 2-norm vector $x \in \mathbb{R}^{2^d}$, compute $\theta_1, \ldots, \theta_d$ so that $\| x - z_1 \otimes z_2 \otimes \cdots \otimes z_d \|_2$ is minimum where $z_k \;=\; [\cos(\theta_k); \sin(\theta_k)]$. (The $\otimes$ operator is defined in the previous problem.) Choose your poison: `lsqnonlin` or `fminsearch`. Supply as much gradient/jacobian/Hessian info as possible.