Homework 3 – due Friday, November 4th.

From Bindel & Goodman, Chapter 6: #'s 1 & 3 (typo: $x_{k+1} = -x_k^3$).

This problem first concerns a more sensible derivation of Broyden's method than that given in class. The question is where does that rank-one update come from. Recall that we had the secant equation:

$$\mathbf{A}_+ \mathbf{s}_c = \mathbf{y}_c$$

and wanted to minimize the change in $\mathbf{A}_+$ relative to the Jacobian "approximation", $\mathbf{A}_c$, of the previous iteration. That is, let's minimize $\|\mathbf{A}_+ - \mathbf{A}_c\|$ under the constraint of the secant equation. To get Broyden's method we choose the matrix norm to be the Frobenius norm: $\|\mathbf{A}\|_F = \left(\frac{1}{2}\sum_{i=1}^N \sum_{j=1}^N A_{i,j}^2\right)^{1/2}$ for $\mathbf{A} \in \mathbb{R}^{N\times N}$(note: this is not a norm induced by a vector norm). We use the method of Lagrange multipliers, and so seek to minimize in $\mathbf{A}_+$ and $\boldsymbol{\lambda}$ the function

$$F(\mathbf{A}_+;\boldsymbol{\lambda}) = \left[\, \|\mathbf{A}_+ - \mathbf{A}_c\|_F^2 + \boldsymbol{\lambda}^T(\mathbf{A}_+\mathbf{s}_c - \mathbf{y}_c)\,\right]$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ is a vector of Lagrange multipliers that enforces the secant equation.

(**i**) Show that the stationary (critical) points of F in the components of $\mathbf{A}_+$ and $\boldsymbol{\lambda}$ satisfy

$$\boldsymbol{\lambda} = \frac{1}{\mathbf{s}_c^T\mathbf{s}_c}(\mathbf{A}_c\mathbf{s}_c - \mathbf{y}_c)$$

$$\mathbf{A}_+ = \mathbf{A}_c + \frac{1}{\mathbf{s}_c^T\mathbf{s}_c}(\mathbf{y}_c - \mathbf{A}_c\mathbf{s}_c)\mathbf{s}_c^T$$

(**ii**) Prove the **Sherman-Morrison formula** for calculating the inverse of a rank-one change to a matrix:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$$

This gives a method for directly updating the inverse of $\mathbf{A}_c$. Comment on the structure of the inverse.

(**iii**) There are now two ways to implement Broyden's method. In the first we write:

$$\begin{aligned}
\mathbf{A}_k\mathbf{s}_k &= -\mathbf{f}(\mathbf{x}_k) \\
\mathbf{x}_{k+1} &= \mathbf{x}_k + \mathbf{s}_k \\
\mathbf{y}_k &= \mathbf{f}(\mathbf{x}_{k+1}) - \mathbf{f}(\mathbf{x}_k) \\
\mathbf{A}_{k+1} &= \mathbf{A}_k + \frac{1}{\mathbf{s}_k^T\mathbf{s}_k}(\mathbf{y}_k - \mathbf{A}_k\mathbf{s}_k)\mathbf{s}_k^T
\end{aligned}$$

and in the second:

$$
\begin{aligned}
\mathbf{s}_k &= -\mathbf{A}_k^{-1}\mathbf{f}(\mathbf{x}_k) \\
\mathbf{x}_{k+1} &= \mathbf{x}_k + \mathbf{s}_k \\
\mathbf{y}_k &= \mathbf{f}(\mathbf{x}_{k+1}) - \mathbf{f}(\mathbf{x}_k) \\
\mathbf{A}_{k+1}^{-1} &= \mathbf{A}_k^{-1} + \frac{(\mathbf{s}_k - \mathbf{A}_k^{-1}\mathbf{y}_k)}{\mathbf{s}_k^T\mathbf{A}_k^{-1}\mathbf{y}_k}\mathbf{s}_k^T\mathbf{A}_k^{-1}
\end{aligned}
$$

Compare in terms of rough operation count (the scaling of the number floating point ops with N) the two different ways of implementing Broyden's method.

(**iv**) Consider the Lorenz equations [E. N. Lorentz, 1963, *Deterministic nonperiodic flow,* J. Atmospheric Science]

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= \rho x - y - xz \\
\dot{z} &= -\beta z + xy
\end{aligned}
$$

where $\sigma,\ \rho,\ \beta$ are parameters. This systems has 3 fixed points: $(x,y,z)_0 = (0,0,0)$ and $(x,y,z)_\pm = \left(\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho - 1\right)$. Implement both Newton's method and Broyden's method for finding these fixed points (use the exact Jacobian to start the Broyden method). Fixing $\mu = 1$, $\rho = 2$ and $\beta = 1$, demonstrate that your implementations shows convergence to $(x,y,z)_+$ if the initial guess is sufficiently close (but don't start on the solution. That's cheating). Demonstrate the quadratic convergence of Newton's method, and try and extract from your results a convergence rate for the Broyden method (i.e. try to find γ such that $\|\mathbf{x}_{k+1} - \mathbf{x}_+\| \sim C\|\mathbf{x}_k - \mathbf{x}_+\|^\gamma$).

(**v**; **extra credit**) Lastly, find numerically the solution branch for $\beta \in [0,1]$ (even though we know it analytically). Given our numerically determined solution at $\beta = 1$, take advantage of the local convergence properties of Newton's method by slightly decreasing β (say by $\Delta\beta = 0.1$) and restarting Newton's method (now with $\beta = 0.9$) using as initial guess the solution determined for $\beta = 1$. It will converge very quickly as the two solutions are close. Now decrease β again, and use the $\beta = 0.9$ solution as the initial guess, and so on, and decrease β towards zero (sounds like a for-loop). This is called a *continuation method*. What happens to the convergence rate of Newton's method as $\beta = 0$ is approached? What happens to the determinant of the Jacobian? Do the same study, starting from $\rho = 2$, $\beta = 1$, but decreasing ρ towards 1.