Math 4500/6500 Homework #11

This homework assignment covers our notes on Direction Set methods in multivariable minimization.

1. Consider the function
$$f(x_1, x_2, x_3) = 3e^{x_1 x_2} - x_3 \cos x_1 + x_2 \ln x_3.$$
Now do the following problems:
   (1) Find the gradient vector and the Hessian matrix of f.
   (2) Find the first three terms in the Taylor expansion of f around $(0, 1, 1)$.
   (3) Set up and solve a linear system to guess where this function has a minimum. Actually evaluate the function at this point. Is this near the minimum of the function?

2. An inequality of I. Schur states that
$$x^2(x-y)(x-z) + y^2(y-z)(y-x) + z^2(z-x)(z-y) \geq 0$$
for all x, y, and z, with equality if and only if $x = y = z$ or two of x, y, z are equal and the other is zero. Suppose that we restrict our attention to the sphere $x^2 + y^2 + z^2 = 1$. Use gradient descent to find the maximum value of the function above on the sphere.
   (1) First, translate the function into spherical polar coordinates θ and ϕ (using *Mathematica*).
   (2) Second, compute the gradient of the new function $f(\theta, \phi)$ using *Mathematica*.
   (3) Now pick a starting point and describe why you picked the given starting point. Hint: I'd graph the function.
   (4) Third, find the maximum of the function using your method.

3. Consider the data

x	y	x	y
1.0	-1.945	3.2	0.764
1.2	-1.253	3.4	0.532
1.4	-1.140	3.6	1.073
1.6	-1.087	3.8	1.286
1.8	-0.760	4.0	1.502
2.0	-0.682	4.2	1.582
2.2	-0.424	4.4	1.993
2.4	-0.012	4.6	2.473
2.6	0.190	4.8	2.503
2.8	0.452	5.0	2.322
3.0	0.337		

We believe that this data comes from a linear function of the form $y = mx + b$, and we wish to discover m and b using a numerical method. The method of least-squares minimizes the *root mean square error* (or RMS error) in the model $y = f(x) = mx + b$, which is given by the formula

$$\mathrm{RMS}(m,b) = \left(\frac{1}{n} \sum_{i=1}^{n} [(mx_i + b) - y_i]^2 \right)^{1/2}.$$

This is the square root of the mean of the squares of the differences between the predicted data value $mx_i + b$ and the actual data value y_i for each observation (x_i, y_i). Minimizing RMS error is the same as minimizing $\mathrm{RMS}^2(m,b)$.

(1) Using *Mathematica*, write out and algebraically `Simplify` the function $\mathrm{RMS}^2(m,b)$ for the data above.
(2) Minimize the resulting function using our conjugate gradient code from the class demonstration notebook. How many iterations were required? (Are you not entertained?)
(3) Plot the data, along with the linear model. Does it look like a good fit?

4. There are theoretical results and special-purpose numerical methods available for the least-squares problem in the previous exercise. But one of the virtues of generic numerical methods is that they can handle cases which are much harder to analyze theoretically. So suppose we were interested in minimizing the maximum error of approximation for a linear model for the data above:

$$\mathrm{MaxError}(m,b) = \max_i |(mx_i + b) - y_i|.$$

This function is harder to analyze.

(1) This function is not differentiable. Produce a surface plot for $\mathrm{MaxError}(m,b)$ as a function of m and b using *Mathematica* which clearly shows "corners" and "ridges" in the surface.
(2) Minimize the function anyway using Nelder-Mead.
(3) Plot the data, along both linear models. Are the linear models very different?
(4) Change the last data point to $(5,3)$ and repeat the calculations. Now are the linear models very different?