MATH 689: Numerical Optimization

Prof. Wolfgang Bangerth // Blocker 507D // bangerth@math.tamu.edu

Homework assignment 5 – due Tuesday 2/25/2014

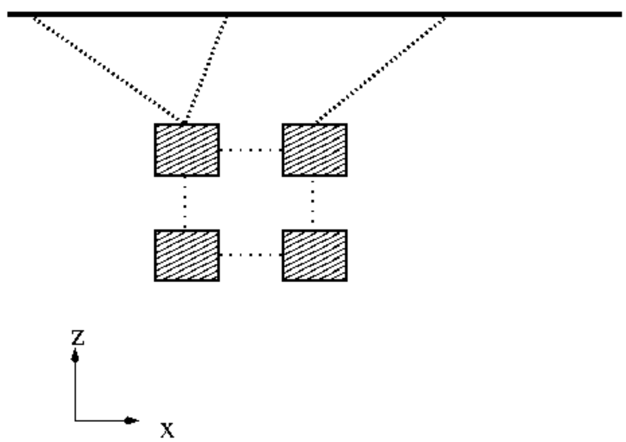
Problem 1 (BFGS for quadratic functions?). Consider $f : \mathbb{R}^4 \to \mathbb{R}$ given by $f(x) = \frac{1}{2}x^T Ax$ with

$$A = \begin{pmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{pmatrix}$$

Its minimum lies at $x = 0$. Implement the BFGS algorithm for this problem, using full step length $\alpha_k = 1$ in each step, starting at $x_0 = (1, 2, 3, 4)^T$. How many steps do you need to converge to the minimum? How many steps would you need if you used the exact Newton matrix instead of the BFGS approximation?

(4 points)

Problem 2 (Not just three springs). Let us revisit one of last week's problems. Consider this system here:



The three springs at the top are suspended from the ceiling at positions $(x, z) = (-20\text{cm}, 0\text{cm})$, $(-5\text{cm}, 0\text{cm})$, and $(7\text{cm}, 0\text{cm})$ and have rest lengths of $L_0^{\text{top}} = 20\text{cm}$. The 4 bodies are connected with 4 springs of rest lengths $L_0^{\text{between}} = 5\text{cm}$. All springs have a spring constant of $D = 300\frac{N}{m}$ and are attached to the centers of the bodies (which for the purpose of this problem we will consider as point-like). All bodies have a mass of $m = 500g$.

Express the total energy in the system (spring energies plus potential energies) as a function of the 4 bodies' positions $(x_i, z_i), i = 1, \ldots, 4$. As before, nature likes to do things so that the energy is minimal, so use a line search quasi-Newton method to find the location at which this energy is minimal. Use the BFGS formula to approximate the inverse of the Hessian matrix and compute quasi-Newton updates with it. (Note that this saves you the trouble of computing awkward looking second derivatives of the objective function, and you also don't have to worry whether the Hessian may have negative eigenvalues.)

In your answer, give the location of the centers of the 4 bodies as well as the value of the total energy function at the minimum. **(6 points)**

Problem 3 (Is the BFGS matrix always positive definite?). Given a previous BFGS matrix B_k, and given previous steps $s_k = x_{k+1} - x_k$ and changes in the gradient $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$, the BFGS method computes the next approximation to the inverse of the Hessian using

$$B_{k+1}^{-1} = (\mathbf{I} - \rho_k s_k y_k^T) B_k^{-1} (\mathbf{I} - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$$

where $\rho_k = \frac{1}{y_k^T s_k}$. This matrix is certainly positive definite if B_k was already positive definite and if $\rho_k > 0$. Indeed, it can be shown that $\rho_k > 0$ if the step length is chosen according to the Wolfe conditions (think about why that might be so).

As an alternative to the previous statement, show that $\rho_k > 0$ for any pair of points x_{k+1}, x_k if $f : \mathbb{R}^n \to \mathbb{R}$ is a strictly convex function. You can assume that $f \in C^2$.

(2 points)

The same can be shown with a slightly more clever trick for strictly convex functions f that are only C^1. The result then obviously also holds if $f \in C^2$. See if you can show that $\rho_k > 0$ under this weaker condition.

(1 bonus points)

One can make things even more general by just assuming that f only have a non-empty set of subdifferentials at all points. That is certainly true if $f \in C^1$, but also for a larger set of strictly convex functions. Try to repeat the calculation in this case. **(1 more bonus points)**

If you have comments on the way I teach – in particular suggestions how I can do things better, if I should do more or less examples, powerpoint slides vs whiteboard, etc – or on other things you would like to critique, feel free to hand those in with your homework as well. I want to make this as good a class as possible, and all comments are certainly much appreciated!