Fall 2015:
Computational and Variational Methods for Inverse Problems
CSE 397/GEO 391/ME 397/ORI 397
Assignment 2 (due Oct. 14, 2015)

1. Consider the unconstrained optimization problem

$$\min f(x, y) \equiv -\cos x \cos(y/10).$$

   (a) Find and classify all stationary points in the region $-\pi/2 \leq x \leq \pi/2, -10\pi/2 \leq y \leq 10\pi/2$

   (b) There is a portion of the problem region within which the Hessian matrix of $f(x, y)$ is positive definite. Give expressions for this portion. You should be able to do this analytically.

   (c) Derive expressions for the search directions associated with the steepest descent and Newton methods.

   (d) Write a program that performs iterations for both methods, both without a line search and with an exact line search. Note that you will not be able to find the value of the optimal step length analytically; instead, determine it numerically[1].

   (e) Run your program for various initial guesses within the region. Verify through numerical experiments the following:

      i. Steepest descent converges to the minimum x^* for any starting point within the region.

      ii. Newton's method with line search converges to the minimum only for initial points for which the Hessian matrix is positive definite.

      iii. Newton's method without line search has an even more restricted radius of convergence. *Extra credit:* Determine this radius of convergence.

   (f) What do you observe about the convergence rate in these cases? *Extra credit:* Verify the observed convergence rate analytically.

2. Consider the minimization problems

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} f_1(\boldsymbol{x}) \qquad \text{and} \qquad \min_{\boldsymbol{x} \in \mathbb{R}^n} f_2(\boldsymbol{x}),$$

   where $f_2(\boldsymbol{x}) = \beta f_1(\boldsymbol{x})$ with an $\beta > 0$.

   (a) Show that these two problems have the same minimizers and compare the steepest descent and the Newton directions at $\boldsymbol{x}_0 \in \mathbb{R}^n$. In class we showed that (locally) a good step length for Newton's methods is $\alpha = 1$, and thus we initialize a backtracking line search with that value. Is is possible to give a good initial step length for steepest descent?

[1] You may use the built-in one-dimensional minimization function `fzero` in MATLAB, for example

(b) Newton's method for optimization problems can also be seen as a method to find stationary points $\boldsymbol{x}$ of the gradient, i.e., points where $\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$. Show that the Newton step for $\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$ coincides with the Newton step for the modified problem

$$\boldsymbol{B}\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}, \tag{1}$$

where $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ is a non-singular matrix[2].

3. Write a program that implements the inexact Newton-conjugate gradient method as described in class[3] and use it to solve the following problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^4} f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T(\boldsymbol{I} + \mu\boldsymbol{A})\boldsymbol{x} + \frac{\sigma}{4}\left(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\right)^2$$

with parameters $\sigma > 0, \mu \geq 0$, the identity matrix $\boldsymbol{I}$, and the matrix $\boldsymbol{A}$ given by

$$\boldsymbol{A} = \begin{pmatrix} 5 & 1 & 0 & 0.5 \\ 1 & 4 & 0.5 & 0 \\ 0 & 0.5 & 3 & 0 \\ 0.5 & 0 & 0 & 2 \end{pmatrix}.$$

Your implementation should have the following features:

- Terminate the CG iterations when $\|\boldsymbol{H}_k\boldsymbol{p}_k + \boldsymbol{g}_k\| \leq \eta_k\|\boldsymbol{g}_k\|$, and implement the following three options for η:
  - $\eta_k = 0.5$
  - $\eta_k = \min(0.5, \sqrt{\|\boldsymbol{g}_k\|})$
  - $\eta_k = \min(0.5, \|\boldsymbol{g}_k\|)$
- Also terminate the CG iterations when a direction of negative curvature is detected within the CG iteration.
- For a line search, implement a backtracking Armijo line search as described in class.

(a) Please turn in code listings of your implementation.

(b) Compare the performance of Newton and steepest descent for $\sigma = 1, \mu = 0$, as well as for $\sigma = 1, \mu = 10$. Use the starting point $\boldsymbol{x} = (\cos 70°, \sin 70°, \cos 70°, \sin 70°)^T$. Can you explain the different behavior?

(c) Experiment with the three different choices of η for the case $\sigma = 1, \mu = 10$. Verify the theoretical convergence rates for the different choices of η.

[2]This property is called *affine invariance* of Newton's method, and it is one of the reasons why Newton's method is so efficient. A comprehensive reference for Newton's method is the book by P. Deuflhard, *Newton Methods for Nonlinear Problems*, Springer 2006.

[3]Reference: S.C. Eisenstat and H.F Walker, *Globally convergent inexact Newton's method*, SIAM Journal on Optimization, Vol. 4, p.393–422, 1994.