AMATH/MATH 516
Problem Set 6

(1) Use the conjugate gradient algorithm to determine the polynomial of degree 6 that best fits the data

$$y_i = \sin x_i \ , \text{ with } x_i := -1 + \frac{i}{5} \text{ for } i = 0.1. \cdots , 10,$$

in the least squares sense. Take $p^0 = 0$. Run the algorithm as an iterative procedure terminating when the magnitude of the gradient of the least squares objective is less than 10^{-12}. Explain what you observe. In particular, compare the theoretical finite convergence result with what you actually observe.

(2) In this problem set we revisit the nonlinear least squares problem of the previous problem set. That is, we are interested solving the nonlinear least squares problem

$$\mathcal{P} \qquad \min_{x \in B} f(x) := \frac{1}{2} \|g(x)\|_2^2$$

where matlab function files for the function $g : \mathbb{R}^7 \to \mathbb{R}^8$, its Jacobian, and its Hessians are available through the course webpage.

Each of the algorithms described below is to be initiated at the point $x^0 = zeros(7, 1)$ with a stopping criteria of $\|\nabla f(x)\| \leq 10^{-12}$. As part of your output, you need to include
(i) a graph of the primary stopping criteria, e.g. the norm of the gradient in the unconstrained case
(ii) a graph of the function values
(iii) a graph of the magnitude of the steps taken at each iteration
(iv) a table listing the total number of function, gradient, and Hessian evaluations.
Use a `maxit` in excess of 5000.

(a) Solve the problem using the safe-guarded Newton step described below and the non-monotone backtracking line search. (if you can't get the non-monotone line search working, just use Weak Wolf, see part (b) below).
Safeguard the Newton search direction by two checks:
(i) If $\left\|H_k^{-1}\nabla f(x_k)\right\| \geq 100 \|\nabla f(x_k)\|$, reset $d_k = -\nabla f(x_k)$.
(ii) If $-\nabla f(x_k)^T d_k \leq 10^{-4} \|\nabla f(x_k)\| \|d_k\|$, reset $d_k = -\nabla f(x_k)$.
(b) Solve the problem using the safe-guarded Newton step of the previous problem and the weak Wolfe line search.
(c) Solve the problem using (inverse) BFGS updating in conjunction with the weak Wolfe line search. Note that Weak Wolfe is essential here.

(3) We return to the nonlinear least squares problem, but now with box constraints:

$$\mathcal{P} \qquad \min_{x \in B} f(x) := \frac{1}{2} \|g(x)\|_2^2 \, ,$$

where matlab function files for the function $g : \mathbb{R}^7 \to \mathbb{R}^8$, its Jacobian, and its Hessians are available through the course webpage. We consider two cases. In the first $B = \mathbb{R}^7$ as before while in the second

$$B = \{x \in \mathbb{R}^7 \ : \ -5 \leq x_i \leq 5, \ i = 1, 2, \ldots , 7\}.$$

Each of the algorithms described below is to be initiated at the point $x^0 = zeros(7, 1)$ with a stopping criteria of $\|\nabla f(x)\| \leq 10^{-12}$. As part of your output, you need to include
(i) a graph of the primary stopping criteria, e.g. the norm of the gradient in the unconstrained case
(ii) a graph of the function values
(iii) a graph of the magnitude of the steps taken at each iteration
(iv) a table listing the total number of function, gradient, and Hessian evaluations.
Use a `maxit` in excess of 5000.

(a) Solve the constrained problem using the Gradient Projection Algorithm with the backtracking line search and the search direction given by

$$d^{k+1} := \begin{cases} P_B(x^{k+1} - \sigma_k \nabla f(x^{k+1})) - x^{k+1}, & \text{if } (s^k)^T y^k > 0.001, \\ P_B(x^{k+1} - \nabla f(x^{k+1})) - x^{k+1}, & \text{otherwise,} \end{cases}$$

where $\sigma_k = 1/\|\nabla f(x^{k+1})\|$ and $\sigma_k := (s^k)^T y^k/(y^k)^T y^k$. Does the Barzilai-Borwein scaling improve performance over vanilla steepest descent?

(b) Develop your own method to solve the constrained problem based on exterior penalization using the objective function

$$f_\alpha(x) := f(x) + \frac{\alpha}{2}\mathrm{dist}_2^2(x|\, B).$$

How should α be updated? You might try the method in Corollary 8.1.1.1 of the course notes.

(c) Develop your own method to solve the constrained problem based on a log-barrier method for the constraints:

$$f_t(x) := f(x) + t\sum_{j=1}^{7} \ln(5 + x_j)(5 - x_j)\ .$$

Use the ideas developed in previous homeworks if you wish.