In this problem set we revisit the nonlinear least squares problem of Problem Set 3. That is, we are interested solving the nonlinear least squares problem

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

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 \( x^0 = \text{zeros}(7, 1) \) 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 \texttt{maxit} in excess of 5000.

(1) Solve the unconstrained problem using BFGS updating in conjunction with the bisection method for the weak Wolfe conditions.

(2) Solve the unconstrained problem using the Barzilai-Borwein scaling of the steepest descent direction at each iteration, i.e.

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

where \( s^k := x^{k+1} - x^k \), \( y^k := \nabla f(x^{k+1}) - \nabla f(x^k) \), and \( \sigma_k := (s^k)^T y^k / (y^k)^T y^k \). The step size should again be computed using the bisection method for the weak Wolfe conditions.

(3) 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 again \( \sigma_k := (s^k)^T y^k / (y^k)^T y^k \). Does the Barzilai-Borwein scaling improve performance over vanilla steepest descent?

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

\[ f_\alpha(x) := f(x) + \frac{\alpha}{2} \text{dist}_2^2(x \mid B). \]

How should \( \alpha \) be updated? You might try the method in Corollary 8.1.1.1 of the course notes.

(5) 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 problems 2 and 3 of Problem set 4 if you wish. Can central path ideas be used to tell you how fast the barrier parameter \( t \) can be taken to zero?

Compare your results with your previous work on this problem.