EN.553.761: Nonlinear Optimization I

Homework Assignment #4

Starred exercises require the use of MATLAB.

Exercise 4.1**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a μ-strongly convex function with $\mu > 0$. Show that f has a minimizer x^*.

Exercise 4.2**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex, L-smooth function that has a minimizer $x^\star$. Suppose we use random coordinate choice as a stochastic gradient oracle, with step lengths $\alpha_k = \frac{1}{nL}$. Let the (random) iterates be $x_0, x_1, x_2, \ldots$. Assume that the level set $\mathcal{L} := \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ is bounded and has diameter R, i.e., $\|x - y\| \leq R$ for all $x, y \in \mathcal{L}$. Show that for any $T \geq 1$,

$$\mathbb{E}[f(x_T) - f^\star] \leq \frac{2LnR^2}{T}.$$

[Hint: Try to adapt the deterministic analysis (Theorem 2.1 in the "Smooth Convex Optimization" lecture notes) and take expectations in an appropriate way.]

Exercise 4.3***.** This problem asks you to explore the performance of various coordinate minimization approaches for solving the very simple unconstrained optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}}\; f(x) := \tfrac{1}{2} x^T H x \tag{1}$$

for various choices of the symmetric, positive definite matrix H.

(i) Code the following coordinate minimization methods specifically written to solve problem (1).

  (a) *cyclic* coordinate minimization with *exact* linesearch minimization (closed form solution);
  (b) *cyclic* coordinate descent with "optimal" *fixed step size* given by $\alpha = 1/\|H\|_2$;
  (c) *random* coordinate descent with "optimal" *fixed step size* given by $\alpha = 1/\|H\|_2$; and
  (d) *Gauss-Southwell* coordinate descent with "optimal" *fixed step size* given by $\alpha = 1/\|H\|_2$.

(ii) Use your codes to *evaluate* the performance of methods (a)–(d) on problems of the form (1), where the matrix H is randomly created in MATLAB using the built in command SPRANDSYM. The command has options to constrain the matrix to be positive definite (by using the "kind" argument) and also the condition number.

  (a) How do they perform for $n = 10$ and $\text{cond}(H) \in \{10^1, 10^2, 10^3, 10^4\}$?
  (b) How do they perform for $n = 100$ and $\text{cond}(H) \in \{10^1, 10^2, 10^3, 10^4\}$?
  (c) How do they perform for $n = 1000$ and $\text{cond}(H) \in \{10^1, 10^2, 10^3, 10^4\}$?