CO 367 Fall 2018: Homework 2

Due: November 12th, 1:30pm

Instructions For every nontrivial step you perform, you must justify why the step is valid and what assumptions it exploits. In other words, you do not need to justify basic algebraic operations (rearranging or distributing terms, multiplying both sides of an equation by a constant, etc.), but you *do* need to explain all steps that exploit hypotheses and assumptions (positive semidefiniteness of a matrix, continuity or convexity of a function, taking a limit that must exist, etc.). If you exploit a result seen in class, or an elementary theorem, clearly state which one.

Question 1 [8 marks] Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1-smooth function over $\mathbb{R}^n$. Let us fix a point $x^k \in \mathbb{R}^n$ and a descent direction $p^k \in \mathbb{R}^n$. We define a function $\psi : \mathbb{R} \to \mathbb{R}$ for the line search problem, i.e. $\psi(\alpha) = f(x^k + \alpha p^k)$.
(i) Prove that if ∇f is Lipschitz continuous with constant L, then $\frac{d}{d\alpha}\psi$ is also Lipschitz continous. Give a Lipschitz constant K for $\frac{d}{d\alpha}\psi$.
(ii) Assuming that K is known, prove that setting $\alpha = \frac{2}{3K}\left(-\frac{d}{d\alpha}\psi(0)\right)$ will satisfy the sufficient decrease condition for line search with $\sigma = \frac{1}{3}$.
(iii) If f is strongly convex, give a constant γ such that the curvature condition is satisfied for all $\alpha > \gamma$. The expression of γ may involve $\psi(0)$ and p.

Question 2 [4 marks] In class, we did not use line search for Newton's method: we simply set $x^{k+1} = x^k + p^k$, where $p^k \neq 0$ is given by

$$p^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k).$$

Instead, consider the possibility of doing line search along the direction p^k, i.e. we set $x^{k+1} = x^k + \alpha p^k$, for some $\alpha > 0$. We consider the special case where f is a strictly convex quadratic function.
(i) Prove that $\psi(\alpha) = f(x^k + \alpha p^k)$ is also a strictly convex quadratic function.
(ii) Briefly explain why the unique global minimizer of ψ is $\alpha^* = 1$.
(iii) Prove that setting $\alpha = 1$ satisfies the sufficient decrease condition and the curvature condition for $\sigma \leq \frac{1}{2}$.

Question 3 [3 marks] Assume that we can solve the trust region subproblem. In other words, we are given a function $s(A, b) = \text{argmin}\{x^T Ax + b^T x \ : \ ||x||_2 \leq 1\}$ that is defined for any $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$. Consider the problem

$$\min\{x^T Cx + d^T x \ : \ (x - z)^T R(x - z) \leq 1\}, \tag{E}$$

where we are given constants $C \in \mathbb{R}^{n\times n}$, $d \in \mathbb{R}^n, z \in \mathbb{R}^n$, $R \in \mathbb{R}^{n\times n}$, and where R is symmetric and positive definite. Give an expression for an optimal solution x^* to (E). This expression may use the function s.