CS532100 Numerical Optimization Homework 1

Student ID Name

Due Nov 11

1. (45%) Consider a function $f(x_1, x_2) = x_1^3 x_2 - 2x_1 x_2^2 + x_1 x_2^3$.
   (a) Compute the gradient and Hessian of f.
   (b) Is $(x_1, x_2) = (1, 1)$ a local minimizer? Justify your answer.
   (c) What is the steepest descent direction of f at $(x_1, x_2) = (1, 2)$?
   (d) What is the Newton's direction of f at $(x_1, x_2) = (1, 2)$?
   (e) Compute the LDL decomposition of the Hessian of f at $(x_1, x_2) = (1, 2)$. (No pivoting)
   (f) Is the Newton's direction of f at $(x_1, x_2) = (1, 2)$ a descent direction? Justify your answer.
   (g) Modify the LDL decomposition computed in (d) such that all diagonal elements of D is larger than or equal to 1, and use the modified LDL decomposition to compute a modified Newton's direction at $(x_1, x_2) = (1, 2)$.
   (h) Suppose $\vec{x_0} = (1, 1)$ and $\vec{x_1} = (1, 2)$ and $B_0 = I$, compute the quasi Newton direction p_1 using SR1.
   (i) Suppose $\vec{x_0} = (1, 1)$ and $\vec{x_1} = (1, 2)$ and $B_0 = I$, compute the quasi Newton direction p_1 using BFGS.

   Answers are put here.

   可以用中文作答。

2. (20%)
   (a) A set $S \subseteq R^n$ is a *convex* set if the straight line connecting any two points in S is also entirely in S. A function $f : S \to R$ is a *convex* function if S is a convex set. The following properties are equivalent:
      i. $S \subseteq R^n$ is a convex set, $f : S \to R$ is a convex function.
      ii. $f(\alpha\vec{x} + (1-\alpha)\vec{y}) \leq \alpha f(\vec{x}) + (1-\alpha) f(\vec{y})$ for all $\alpha \in [0, 1]$, $\vec{x}, \vec{y} \in S$.

      Prove that when f is a convex function, any local minimizer $\vec{x}^*$ is a global minimizer of f.
      (Hint: Suppose there is another point $\vec{z} \in S$ such that $f(\vec{z}) \leq f(\vec{x}^*)$, then $\vec{x}^*$ is not a local minimizer.)
   (b) Suppose $f(\vec{x}) = \vec{x}^T Q \vec{x}$, where Q is a symmetric positive semidefinite matrix, show that $f(\vec{x})$ is a convex function.
      (Hint: It might be easier to show $f(\vec{y} + \alpha(\vec{x} - \vec{y})) - \alpha f(\vec{x}) - (1 - \alpha) f(\vec{y}) \leq 0$.)

Answers are put here.

可以用中文作答。

3. (20%) (Line search method) Suppose $\phi(\alpha) = f(\vec{x_k} + \alpha\vec{p_k}) = (\alpha - 1)^2$.

   (a) The sufficient decrease condition asks $\phi(\alpha) \le \phi(0) + c_1\alpha\phi'(0)$, $\alpha \in [0, \infty)$. Suppose $c_1 = 0.1$, what is the feasible interval of α satisfying this condition ?

   (b) The curvature condition asks $\phi'(\alpha) \ge c_2\phi'(0)$. Suppose $c_2 = 0.9$, what is the feasible interval of α satisfying this condition ?

Answers are put here.

可以用中文作答。

4. (15%) The conjugate gradient method for solving $Ax = b$ is given in Figure 1, where z_k is the approximate solution. In class we only showed that $\alpha_k = (\vec{p_k}^T\vec{r_k})/(\vec{p_k}^T A\vec{p_k})$ and $\beta_k = -(\vec{p_k}^T A\vec{r_{k+1}})/(\vec{p_k}^T A\vec{p_k})$. Prove that the above formula of α_k and β_k are equivalent to the ones in step (3) and step (6). You may need the relations in step (4) and step (5), and the following properties:

   (a) $\vec{r_i}^T\vec{r_j} = 0$ for all $i \ne j$.

   (b) $\vec{p_i}^T A\vec{p_j} = 0$ for all $i \ne j$.

   (c) $\vec{p_k}$ is a linear combination of $\vec{r_0}, \ldots .\vec{r_k}$, $\vec{p_k} = \sum_{i=1}^{k}\gamma_i\vec{r_i}$.
   (which can be shown from step (7) by mathematical induction.)

(1)	Given $\vec{z}_0$. Let $\vec{p}_0 = \vec{b} - A\vec{z}_0$, and $\vec{r}_0 = \vec{p}_0$.
(2)	For $k = 0, 1, 2, \ldots$ until $\|\vec{r}_k\| \le \epsilon$
(3)	$\alpha_k = (\vec{r}_k^T\vec{r}_k)/(\vec{p}_k^T A\vec{p}_k)$
(4)	$\vec{z}_{k+1} = \vec{z}_k + \alpha_k\vec{p}_k$
(5)	$\vec{r}_{k+1} = \vec{r}_k - \alpha_k A\vec{p}_k$
(6)	$\beta_k = (\vec{r}_{k+1}^T\vec{r}_{k+1})/(\vec{r}_k^T\vec{r}_k)$
(7)	$\vec{p}_{k+1} = \vec{r}_{k+1} + \beta_k\vec{p}_k$

Figure 1: The CG algorithm.

Answers are put here.

可以用中文作答。