CS5321 Numerical Optimization Midterm

April 19, 15:20-17:20

Note: Since many problems are in sets, to avoid your wrong answers in the previous problems affecting the latter ones, write down your derivations. Also, if you have any questions about the problems, write them down.

1. Consider a function $f(x_1, x_2) = x_1^3 x_2 - 2x_1^2 x_2^2 + x_1 x_2^3$.

   (a) (5pt) Compute the gradient of f.

   (b) (5pt) Compute the Hessian of f.

   (c) (5pt) Is $(x_1, x_2) = (1, 1)$ a local minimizer? Justify your answer?

   (d) (5pt) What is the steepest descent direction of f at $(x_1, x_2) = (1, 2)$?

   (e) (5pt) Compute the LDL decomposition of the Hessian of f at $(x_1, x_2) = (1, 2)$. (No pivoting)

   (f) (5pt) What is the Newton's direction of f at $(x_1, x_2) = (1, 2)$?

   (g) (5pt) Is the Newton's direction of f at $(x_1, x_2) = (1, 2)$ a *descent direction*? Justify your answer.

   (h) (5pt) 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 $(x_1, x_2) = (1, 2)$.

2. (Line search method) Suppose $\phi(\alpha) = f(\vec{x}_k + \alpha \vec{p}_k) = (\alpha - 1)^2$.

   (a) (10pt) The sufficient decrease condition asks $\phi(\alpha) \le \phi(0) + c_1 \alpha \phi'(0)$. Suppose $c_1 = 0.1$. What is the feasible interval of α satisfying this condition? Note that $\alpha \in [0, \infty)$.

   (b) (10pt) 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?

3. (Quasi-Newton method) The BSGS update formula for approximating the inverse of Hessian matrix is

$$B_{k+1} = (I - \rho_k \vec{s}_k \vec{y}_k^T) B_k (I - \rho_k \vec{y}_k \vec{s}_k^T) + \rho_k \vec{s}_k \vec{s}_k^T,$$

where $\rho = 1/(\vec{y}_k^T s_k)$.

(a) (10pt) Prove that B_{k+1} satisfies the secant equation $B_{k+1}\vec{y}_k = \vec{s}_k$.

(b) (10pt) Prove that B_{k+1} is positive definite if B_k is positive definite and $\vec{y}_k^T B_{k+1} \vec{y}_k > 0$.

4. (CG method)

(a) (10pt) What are the restrictions of using the conjugate gradient method to solve $A\vec{x} = \vec{b}$?

(b) (10pt) Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}, \vec{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \vec{y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Find the scalar β that makes two vectors, $\vec{x}$ and $(\vec{x} - \beta\vec{y})$, A-conjugate.

5. (10pt) In many proofs, we need the function f or its derivative to be continuous. Give an example to explain why this property is important.