Homework (9)

1. Give the details of using the Sherman-Morrison formula to compute the inverse of

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\Delta \mathbf{g}^{(k)} \Delta \mathbf{g}^{(k)T}}{\Delta \mathbf{g}^{(k)T} \Delta \mathbf{x}^{(k)}} - \frac{\mathbf{B}_k \Delta \mathbf{x}^{(k)} \Delta \mathbf{x}^{(k)T} \mathbf{B}_k}{\Delta \mathbf{x}^{(k)T} \mathbf{B}_k \Delta \mathbf{x}^{(k)}}.$$

And further show that $\mathbf{H}_{k+1} = \mathbf{B}_{k+1}^{-1}$ can be rewritten as

$$\mathbf{H}_{k+1} = \mathbf{V}_k^T \mathbf{H}_k \mathbf{V}_k + \rho_k \Delta \mathbf{x}_k \Delta \mathbf{x}_k^T,$$

where $\rho_k = \frac{1}{\Delta \mathbf{g}_k^T \Delta \mathbf{x}_k}$, $\mathbf{V}_k = \mathbf{I} - \rho_k \Delta \mathbf{g}_k \Delta \mathbf{x}_k^T$.

2. Consider the problem to minimize $(3 - x_1)^2 + 7(x_2 - x_1^2)^2$. Starting from the point $(0, 0)^T$, solve the problem by the following procedures:

a. The method of Davidon-Fletcher-Powell (DFP).

b. The method of Broyden-Fletcher-Goldfarb-Shanno (BFGS).

Homework (9)

3. Use L-BFGS to solve extended Rosenbrock function

$$f(\mathbf{x}) = \sum_{i=1}^{n/2} \left[\alpha (x_{2i} - x_{2i-1}^2)^2 + (1 - x_{2i-1})^2 \right],$$

where α is a parameter that you can vary (for example, 1 or 100). The solution is $\mathbf{x}^* = (1, 1, \cdots, 1)^T$, $f^* = 0$. Choose the starting point as $(-1, -1, \cdots, -1)^T$. Observe the behavior of your program for various values of the memory parameter m.

4. Find the *symmetric* matrix $\mathbf{N}$ that minimizes the distance $\|\mathbf{N} - \mathbf{M}\|_F$ subject to the secant condition $\mathbf{N}\mathbf{d} = \mathbf{g}$, where $\mathbf{M}$ is a symmetric matrix.

Homework (9)

5. Solve $\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2+\lambda\|\mathbf{x}\|_1$ using the following majorant minimization methods:

   a. Use the Lipschitz gradient majorant function of $\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2$;

   b. Use the variational majorant function $\min_{\mathbf{d}>\mathbf{0}} \frac{1}{2}\left(\mathbf{x}^T\mathbf{D}\mathbf{x}+\mathbf{1}^T\mathbf{D}^{-1}\mathbf{1}\right)$ of $\|\mathbf{x}\|_1$, where $\mathbf{1}$ is an all-one vector and $\mathbf{D}$ is a diagonal matrix with $\mathbf{d}$ on the diagonal.

Compare their perfomances.

For coding problems, please write reports and hand in both codes and reports. Remember to restart your computer before running your codes.