Homework (12)

1. Given $f : \mathbb{R}^n \to \mathbb{R}$, consider the general iterative algorithm

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)},$$

where $\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \cdots$ are given vectors in $\mathbb{R}^n$ and α_k is chosen to minimize $f(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})$; that is,

$$\alpha_k = \underset{\alpha}{\operatorname{argmin}}\, f(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)}).$$

Show that for each k, the vector $\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$ is orthogonal to $\nabla f(\mathbf{x}^{(k+1)})$ (assuming that the gradient exists).

Homework (12)

2. Use the steepest descent method with the knowledge of Hessian, the Gauss-Newton method, Damped Newton method, and the Conjugate Gradient method with three update formulae for β_k to minimize Rosenbrock's function:

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Use an initial condition of $\mathbf{x}^{(0)} = [-2, 2]^T$. Terminate the algorithm when the norm of the gradient of f is less than 10^{-5}. Write a report to show the figures of $\log(f(\mathbf{x}^{(k)}) - p^*)$ vs. iteration number k and running time, respectively, and your choice of parameters, and what your observation is. p^* can be estimated by running the "best" algorithm with much more iterations. Remember to restart your computer before running the program. Hand in both your code and report.

Homework (12)

3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{b}$, where $\mathbf{b} \in \mathbb{R}^n$, and $\mathbf{Q}$ is a real symmetric positive definite $n \times n$ matrix. Show that in the conjugate gradient method for this f, $\mathbf{d}^{(k)T}\mathbf{Q}\mathbf{d}^{(k)} = -\mathbf{d}^{(k)T}\mathbf{Q}\mathbf{g}^{(k)}$.

4. Consider a conjugate gradient algorithm applied to a quadratic function.

   a) Show that the gradients associated with the algorithm are mutually orthogonal. Specifically show that $\mathbf{g}^{(k+1)T}\mathbf{g}^{(i)} = 0$ for all $0 \leq k \leq n-1$ and $0 \leq i \leq k$.

   b) Show that the gradients associated with the algorithm are $\mathbf{Q}$-conjugate if separated by at least two iterations. Specifically, show that $\mathbf{g}^{(k+1)T}\mathbf{Q}\mathbf{g}^{(i)} = 0$ for all $0 \leq k \leq n-1$ and $0 \leq i \leq k-1$.