Numerical Optimization with applications: Homework 01

104021601 林俊傑
104021602 吳彥儒
104021615 黃翊軒

September 28, 2016

Exercise 1. *Compute the gradient* $\nabla f(x)$ *and Hessian* $\nabla^2 f(x)$ *of the Rosenbrock function*

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Show that $x^* = (1,1)^T$ *is the only local minimizer of this function, and that the Hessian matrix at that point is positive definite.*

Proof. Calculating the gradient of $f(x)$

$$\nabla f(x) = \begin{bmatrix} f_{x_1} \\ f_{x_2} \end{bmatrix} = \begin{bmatrix} 200(x_2 - x_1^2)(2x_1) + 2(1 - x_1)(-1) \\ 200(x_2 - x_1^2) \end{bmatrix} = \begin{bmatrix} -400x_1(x_2 - x_1^2) + 2(x_1 - 1) \\ 200(x_2 - x_1^2) \end{bmatrix}$$

Solving $\nabla f(x) = 0$, we obtain that $\nabla f(x) = 0$ if and only if x equals to $x^* = (1,1)^T$.

On the other hand, calculating the Hessian of $f(x)$

$$\nabla^2 f(x) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix} = \begin{bmatrix} 400(x_2 - x_1^2) - 400x_1(2x_2) + 4 & 400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

Observe that

$$\begin{aligned} P^T \nabla^2 f(x^*) P &= \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 804 & 400 \\ -400 & 200 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \\ &= 804p_1^2 + 400p_1p_2 - 400p_1p_2 + 200p_2^2 \\ &= 804p_1^2 + 200p_2^2 \end{aligned}$$

Therefore, $P^T \nabla^2 f(x^*) P > 0$ for all nonzero vector P. *i.e. the Hessian matrix at* x^* *is positive definite.* By Theorem2.4 (Second-Order Sufficient Conditions), $x^* = (1,1)^T$ is the only local minimizer of this function. □

Exercise 7. *Suppose that* $f(x) = x^T Q x$, *where* Q *is an* $n \times n$ *symmetric positive semidefinite matrix. Show using the definition (1.4) that* $f(x)$ *is convex on the domain* $\mathbb{R}^n$*.Hint: It may be convenient to prove the following equivalent inequality:*

$$f(y + \alpha(x - y)) - \alpha f(x) - (1 - \alpha) f(y) \leq 0$$

for all $\alpha \in [0,1]$ *and all* $x, y \in \mathbb{R}$*.*

Proof. By the definition of f and (1.4), for any $x, y \in \mathbb{R}^n$

$$\begin{aligned} & f(y + \alpha(x - y)) - \alpha f(x) - (1 - \alpha) f(y) \\ &= (y + \alpha(x - y))^T Q (y + \alpha(x - y)) - \alpha x^T Q x - (1 - \alpha) y^T Q y \\ &= y^T Q y + \alpha y^T Q (x - y) + \alpha (x - y)^T Q y + \alpha^2 (x - y)^T Q (x - y) - \alpha x^T Q x - (1 - \alpha) y^T Q y \\ &= \alpha [y^T Q y + y^T Q (x - y) - x^T Q x + (x - y)^T Q y] + \alpha^2 (x - y)^T Q (x - y) \\ &= \alpha [y^T Q x - x^T Q x + (x - y)^T Q y] + \alpha^2 (x - y)^T Q (x - y) \\ &= \alpha [-(x - y)^T Q x + (x - y)^T Q y] + \alpha^2 (x - y)^T Q (x - y) \\ &= -\alpha (x - y)^T Q (x - y) + \alpha^2 (x - y)^T Q (x - y) \\ &= (\alpha - \alpha^2)(x - y)^T Q (x - y) \leq 0 \end{aligned}$$

since $\alpha \in [0,1]$ and Q is positive semidefinite. This completes the proof. □

Exercise 8. *Suppose that f is a convex function. Show that the set of global minimizer of f is a convex set.*

Proof. Let E denotes the set of global minimizer of f.
For any $x^*, y^* \in E$, $f(x^*) \leq f(y^*)$ and $f(y^*) \leq f(x^*)$. i.e. $f(x^*) = f(y^*)$
Then for any $\alpha \in [0, 1]$

$$f(\alpha x^* + (1-\alpha)y^*) \leq \alpha f(x^*) + (1-\alpha)f(y^*) = f(x^*) \leq f(x) \qquad \forall x \in \mathbb{R}^n$$

since x^* is a global minimizer.
By above, $f(\alpha x^* + (1-\alpha)y^*) \leq f(x), \forall x \in \mathbb{R}, \alpha \in [0, 1]$.
E is a convex set. □

Exercise 16. *Consider the sequence x_k defined by*

$$x_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k}, & k \text{ even}, \\ (x_{k-1})/k, & k \text{ odd}. \end{cases}$$

Is this sequence Q-superlinearly convergent? Q-quadratically convergent? R-quadratically convergent?

Proof. Clearly, x_k converges to $x^* = 0$

(i) Q-superlinearly:
If k is even: $\lim\limits_{k\to\infty} \dfrac{|x_{k+1} - x^*|}{|x_k - x*|} = \lim\limits_{k\to\infty} \dfrac{\frac{x_k}{k+1}}{x_k} = \lim\limits_{k\to\infty} \dfrac{1}{k+1} = 0$
If k is odd: $\lim\limits_{k\to\infty} \dfrac{|x_{k+1} - x^*|}{|x_k - x*|} = \lim\limits_{k\to\infty} \dfrac{x_{k+1}}{\frac{x_{k-1}}{k}} = \lim\limits_{k\to\infty} \dfrac{k(\frac{1}{4})^{2^k}}{(\frac{1}{4})^{2^{k-1}}} = \lim\limits_{k\to\infty} k(\dfrac{1}{4})^{2^{k-1}} = 0$
This implies x_k is Q-superlinearly convergence.

(ii) Q-quadratically:
If k is even: $\lim\limits_{k\to\infty} \dfrac{|x_{k+1} - x^*|}{|x_k - x*|^2} = \lim\limits_{k\to\infty} \dfrac{\frac{x_k}{k+1}}{x_k^2} = \lim\limits_{k\to\infty} \dfrac{1}{kx_k} = \lim\limits_{k\to\infty} \dfrac{1}{k(\frac{1}{4})^{2^k}} = +\infty$ This implies x_k is not Q-quadratically convergent.

(iii) R-quadratically:
Let $\epsilon_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k}, & \text{k even}, \\ \left(\frac{1}{4}\right)^{2^{k-1}}, & \text{k odd}. \end{cases}$
When k is even, $|x_k - x^*| = |x_k| = (\frac{1}{4})^{2^k} \leq \epsilon_k$
When k is odd, $|x_k - x^*| = |x_k| = x_{k-1}/k = (\frac{1}{4})^{2^{k-1}}(\frac{1}{k}) \leq (\frac{1}{4})^{2^{k-1}} = \epsilon_k$
By the above, $|x_k - x^*| \leq \epsilon_k \quad \forall k$
If k is even: $\lim\limits_{k\to\infty} \dfrac{|\epsilon_{k+1} - 0|}{|\epsilon_k - 0|^2} = \lim\limits_{k\to\infty} \dfrac{(\frac{1}{4})^{2^k}}{(\frac{1}{4})^{2^k}} = 1$
If k is odd: $\lim\limits_{k\to\infty} \dfrac{|\epsilon_{k+1} - 0|}{|\epsilon_k - 0|^2} = \lim\limits_{k\to\infty} \dfrac{(\frac{1}{4})^{2^{k+1}}}{(\frac{1}{4})^{2^{k-1}}} = \lim\limits_{k\to\infty} [(\dfrac{1}{4})^{2^{k-1}}]^3 = 0$
This implies ϵ_k is Q-quadratically convergent.
And therefore, x_k is R-quadratically convergent.

□