年 月 日 第 次作业

Gradient Descent & Newton's Method

Guoruishen
guorui@outlook.com
May 3, 2020

GD:

① $f(x_{k+1}) \approx f(x_k) + \nabla f(x_k)^T (x_{k+1} - x_k)$

Let $x_{k+1} - x_k = -\lambda \nabla f(x_k)$, then we have

$f(x_{k+1}) \approx f(x_k) - \lambda \nabla f(x_k)^T \nabla f(x_k) \leq f(x_k)$

② Newton's Method: $\cancel{f(x_{k+1}) \approx f(x_k) + \nabla f(x_k)^T (x_{k+1} - x_k) + \frac{1}{2}}$

$f(x_k + \lambda \cdot v) \approx f(x_k) + \lambda \cdot \nabla f(x_k)^T v + \frac{1}{2} \lambda^2 v^T \nabla^2 f(x_k) v$

Therefore, $\dfrac{\partial f(x_k + \lambda v)}{\partial v} = \lambda \nabla f(x_k) + \lambda^2 \nabla^2 f(x_k) v = 0$

$\Rightarrow v^* = -\frac{1}{\lambda} [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

$\Rightarrow f(x_k + \lambda v^*) \approx f(x_k) \cancel{+} - \nabla f(x_k)^T [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) + \frac{1}{2} \nabla f(x_k)^T [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

$= f(x_k) - \frac{1}{2} \nabla f(x_k)^T [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

$\leq f(x_k)$