清華大学

Tsinghua University

自93 姜永鹏 2019010465 凸优化 2022/11/07

1. 假设 $\exists m, M \in \mathbb{R}_{++}$ s.t.

$$mI \preceq \nabla^2 f(x) \preceq MI$$

$$x_{k+1} = T(x_k) = x_k - t\nabla f(x_k).$$

$$T(x) - T(y) = x - t\nabla f(x) - (y - t\nabla f(y))$$
$$= x - y - t(\nabla f(x) - \nabla f(y))$$

由中值定理，化简上式

$$T(x) - T(y) = x - y - t \cdot \nabla^2 f(\xi)(x-y)$$
$$= (I - t\nabla^2 f(\xi))(x-y)$$

其中ξ在线段$[a,b]$上.

特别地，取 $x = x_k$，$y = x^*$(最优解)

$$\|T(x_k) - T(x^*)\| = \|x_{k+1} - x^*\|$$
$$\le \|I - t\nabla^2 f(\xi)\| \|x_k - x^*\|$$

其中$\|x_k - x^*\|$用2-范数度量，与之相容的矩阵范数是谱范数. 上式可写为 $\cancel{\|T(x_k) - T(x^*)\|}$

$$\|x_{k+1} - x^*\| \le \sqrt{\lambda_{max}(B^TB)} \|x_k - x^*\|$$

其中 $B = I - t\nabla^2 f(\xi) = \begin{pmatrix} 1-t & 0 \\ 0 & 1-100t \end{pmatrix}$.

$$\sqrt{\lambda_{max}(B^TB)} = \max\{|1-t|, |1-100t|\}.$$

令 $\frac{(1-t)+(1-100t)}{2} = 0$ 即 $t^* = 2/101$.

满足. $t^* = \underset{t}{\operatorname{argmin}} \sqrt{\lambda_{max}(B^TB)}$

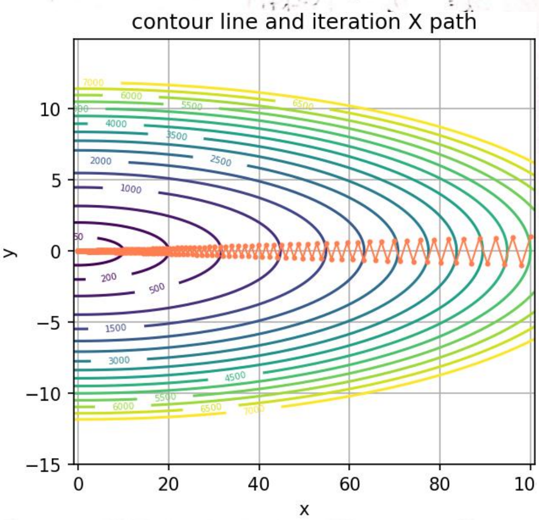
$$\|x_{k+1} - x^*\| \le \frac{99}{101} \|x_k - x^*\|.$$

压缩系数 $\frac{99}{101}$ 最小

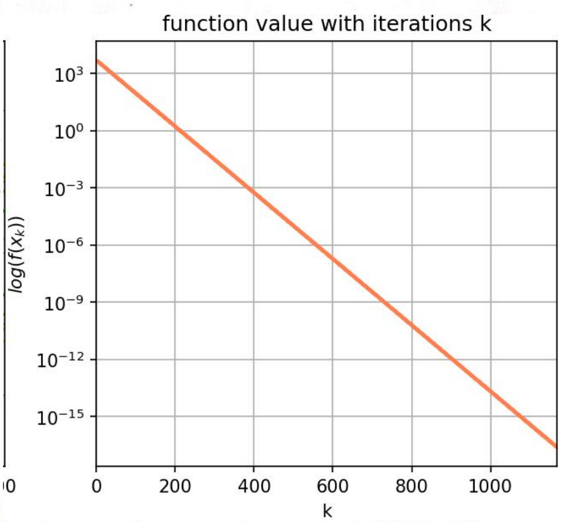
即 $t^* = \frac{99}{101}$ 下收敛最快.

图像:

解的收敛轨迹



函数值的半对数坐标图



2\. 精确直线搜索，记 $f(x)=\frac{1}{2}x^TAx$

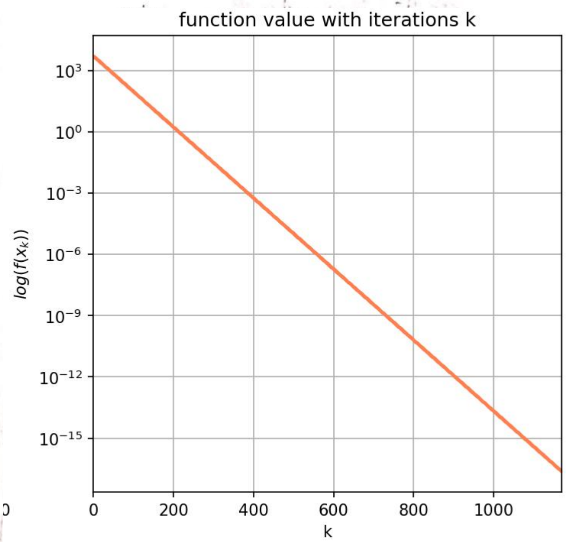
$x_{k+1}=x_k-t_k\nabla f(x_k)$，其中

$$
\begin{aligned}
t_k &= \underset{t}{\operatorname{argmin}}\ f(x_{k+1}) \\
&= \underset{t}{\operatorname{argmin}}\ g(t) \\
&= \underset{t}{\operatorname{argmin}}\ \frac{1}{2}(x_k-t_k\nabla f(x_k))^TA(x_k-t_k\nabla f(x_k))
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial g(t)}{\partial t} &= -\nabla^T f(x_k)A(x_k-t_k\nabla f(x_k)) \\
&= -(Ax_k)^TA(x_k-t_kAx_k) \\
&= 0
\end{aligned}
$$

解得 $t_k=\dfrac{x_k^TA^2x_k}{x_k^TA^3x_k}$

函数值的半对数坐标图



解的收敛轨迹