



# 清华大学

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.

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

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

$$\begin{aligned} T(x) - T(y) &= x - t \nabla f(x) - (y - t \nabla f(y)) \\ &= x - y - t(\nabla f(x) - \nabla f(y)) \end{aligned}$$

由中值定理, 化简上式

$$\begin{aligned} T(x) - T(y) &= x - y - t \nabla^2 f(z)(x - y) \\ &= (I - t \nabla^2 f(z))(x - y) \end{aligned}$$

其中  $z$  在线段  $[a, b]$  上.

特别地, 取  $x = x_k, y = x^*$  (最优解)

$$\begin{aligned} \|T(x_k) - T(x^*)\| &= \|x_{k+1} - x^*\| \\ &\leq \|I - t \nabla^2 f(z)\| \|x_k - x^*\| \end{aligned}$$

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  $\|T(x_k) - T(x^*)\|$

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

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

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

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

$$\text{满足, } t^* = \arg\min_t \sqrt{\lambda_{\max}(B^T B)}$$

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

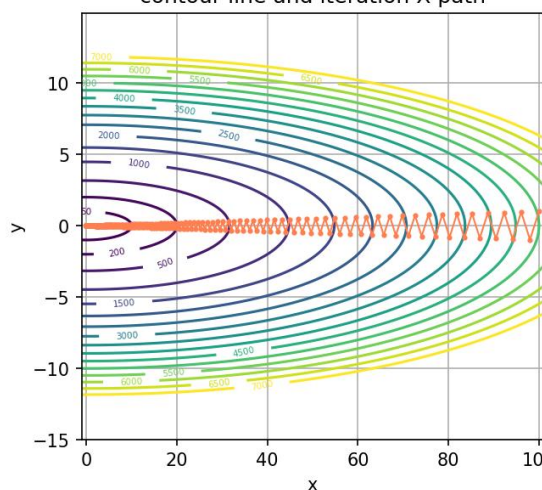
压缩系数  $\frac{99}{101}$  最小

即  $t^* = \frac{99}{101}$  下收敛最快.

图像:

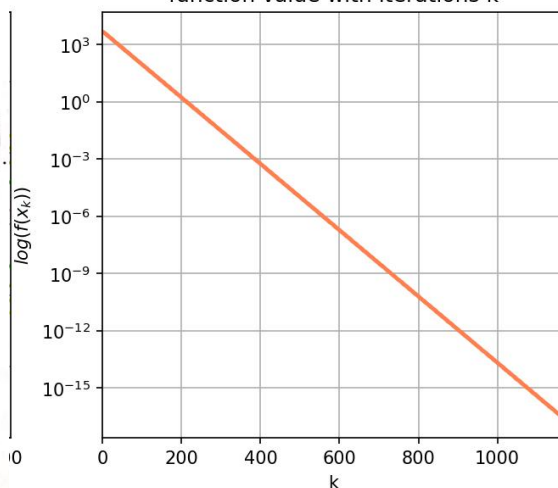
## 解的收敛轨迹

contour line and iteration X path



## 函数值的半对数坐标图

function value with iterations k



2. 精确直线搜索, 记  $f(x) = \frac{1}{2}x^T A x$

$x_{k+1} = x_k - t_k \nabla f(x_k)$ , 其中

$$t_k = \operatorname{argmin}_t f(x_{k+1})$$

$$= \operatorname{argmin}_t g(t)$$

$$= \operatorname{argmin}_t \frac{1}{2} (x_k - t_k \nabla f(x_k))^T A (x_k - t_k \nabla f(x_k))$$

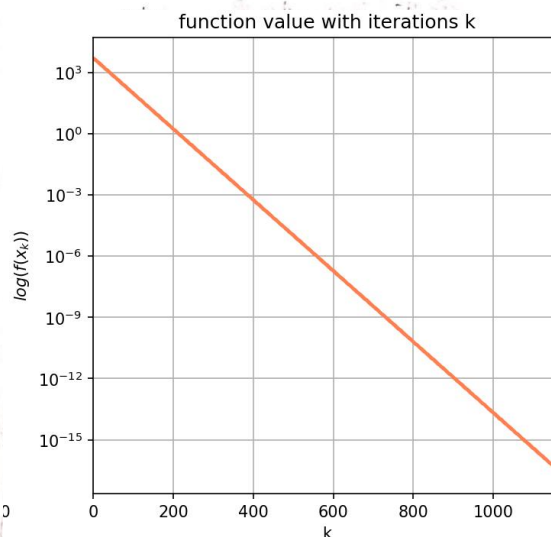
$$\frac{\partial g(t)}{\partial t} = -\nabla f(x_k)^T A (x_k - t_k \nabla f(x_k))$$

$$= -(Ax_k)^T A (x_k - t_k \nabla f(x_k))$$

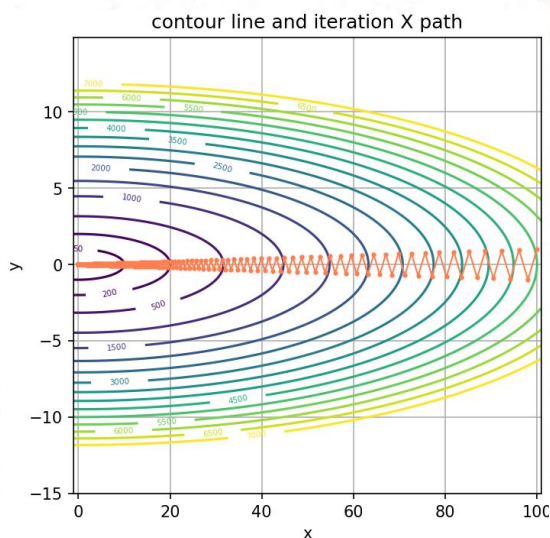
$$= 0$$

$$\text{解得 } t_k = \frac{x_k^T A^2 x_k}{x_k^T A^3 x_k}$$

函数值的半对数坐标图



解的收敛轨迹



由上可知, 该算法在每一步迭代中, 均沿当前点处的负梯度方向进行搜索, 且搜索步长  $t_k$  为使得  $f(x_{k+1})$  达到最小值的步长。因此, 该算法在每一步迭代中, 均沿当前点处的负梯度方向进行搜索, 且搜索步长  $t_k$  为使得  $f(x_{k+1})$  达到最小值的步长。因此, 该算法在每一步迭代中, 均沿当前点处的负梯度方向进行搜索, 且搜索步长  $t_k$  为使得  $f(x_{k+1})$  达到最小值的步长。