

最优化方法 第十二讲

Weiwen Wang(王伟文)

暨南大学

2025 年 6 月 9 日

目录

- 1 临近算子
- 2 临近点法 (Proximal Point Algorithm, PPA)
- 3 临近梯度法
- 4 加速临近梯度法-FISTA

定义 1

给定凸函数 f , f 在点 $\mathbf{x}$ 处点临近算子定义为

$$\text{prox}_f(\mathbf{x}) = \arg\min_{\mathbf{y}} \left\{ f(\mathbf{y}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 \right\}$$

- $\forall \mu > 0$,

$$\text{prox}_{\mu f}(\mathbf{x}) = \arg\min_{\mathbf{y}} \left\{ f(\mathbf{y}) + \frac{1}{2\mu} \|\mathbf{y} - \mathbf{x}\|_2^2 \right\}$$

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

例子

设 C 为凸集, 定义 C 上指示函数

$$f(\mathbf{x}) = \delta_C(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in C \\ +\infty, & \mathbf{x} \notin C \end{cases}$$

则

$$\text{prox}_f(\mathbf{x}) = \arg\min_{\mathbf{y}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 = \Pi_C(\mathbf{x})$$

例子

$$f(\mathbf{x}) = \lambda \|\mathbf{x}\|_2,$$

$$\text{prox}_f(\mathbf{x}) = \left(1 - \frac{\lambda}{\max(\|\mathbf{x}\|_2, \lambda)}\right) \mathbf{x}$$

$$f(\mathbf{x}) = \lambda \|\mathbf{x}\|_1,$$

$$\text{prox}_f(\mathbf{x}) = [|\mathbf{x}| - \lambda \mathbf{1}]_+ \odot \text{sgn}(\mathbf{x})$$

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

命题 2

设 f 为凸函数, 则有

- (a) $\mathbf{x}^*$ 为 $f(\mathbf{x})$ 的最小值点当且仅当 $\mathbf{x}^* = \text{prox}_f(\mathbf{x}^*)$;
- (b) $\|\text{prox}_f(\mathbf{x}) - \text{prox}_f(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$;
- (c) 若 f 为闭函数, $\forall \mathbf{x}, \mathbf{x} = \text{prox}_f(\mathbf{x}) + \text{prox}_{f^*}(\mathbf{x})$

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

证明.

(a)

 $\mathbf{x}^*$ 为 $f(\mathbf{x})$ 的最小值点 $\Leftrightarrow 0 \in \partial f(\mathbf{x}^*)$

$$\Leftrightarrow \mathbf{x}^* = \arg \min_{\mathbf{x}} \left\{ f(\mathbf{x}) + \frac{1}{2} \|\mathbf{x}^* - \mathbf{x}\|_2^2 \right\} = \text{prox}_f(\mathbf{x}^*)$$

(b) 记 $\tilde{\mathbf{x}} = \text{prox}_f(\mathbf{x})$, $\tilde{\mathbf{y}} = \text{prox}_f(\mathbf{y})$, 由最优性条件有,

$$\mathbf{0} \in \partial f(\tilde{\mathbf{x}}) + \tilde{\mathbf{x}} - \mathbf{x}, \quad \mathbf{0} \in \partial f(\tilde{\mathbf{y}}) + \tilde{\mathbf{y}} - \mathbf{y}.$$

故

$$f(\tilde{\mathbf{y}}) \geq f(\tilde{\mathbf{x}}) + \langle \mathbf{x} - \tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \tilde{\mathbf{x}} \rangle, \quad f(\tilde{\mathbf{x}}) \geq f(\tilde{\mathbf{y}}) + \langle \mathbf{y} - \tilde{\mathbf{y}}, \tilde{\mathbf{x}} - \tilde{\mathbf{y}} \rangle$$



临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

证明.

$$f(\tilde{\mathbf{y}}) \geq f(\tilde{\mathbf{x}}) + \langle \mathbf{x} - \tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \tilde{\mathbf{x}} \rangle, \quad f(\tilde{\mathbf{x}}) \geq f(\tilde{\mathbf{y}}) + \langle \mathbf{y} - \tilde{\mathbf{y}}, \tilde{\mathbf{x}} - \tilde{\mathbf{y}} \rangle$$

两不等式相加, 整理得到

$$\|\mathbf{x} - \mathbf{y}\| \cdot \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\| \geq \langle \mathbf{x} - \mathbf{y}, \tilde{\mathbf{x}} - \tilde{\mathbf{y}} \rangle \geq \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^2$$

故

$$\|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\| \leq \|\mathbf{x} - \mathbf{y}\|$$



证明.

(c) 记 $\tilde{\mathbf{x}} = \text{prox}_f(\mathbf{x})$, 由最优性条件

$$\tilde{\mathbf{x}} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \tilde{\mathbf{x}} \in \partial f(\tilde{\mathbf{x}}).$$

又因为 f 为闭凸函数, 故

$$\mathbf{x} - \tilde{\mathbf{x}} \in \partial f(\tilde{\mathbf{x}}) \Leftrightarrow \tilde{\mathbf{x}} \in \partial f^*(\mathbf{x} - \tilde{\mathbf{x}})$$

记 $\mathbf{u} = \mathbf{x} - \tilde{\mathbf{x}}$, 即

$$\mathbf{x} - \mathbf{u} \in \partial f^*(\mathbf{u}) \Leftrightarrow \mathbf{u} = \text{prox}_{f^*}(\mathbf{x})$$

因此

$$\mathbf{x} = \tilde{\mathbf{x}} + (\mathbf{x} - \tilde{\mathbf{x}}) = \text{prox}_f(\mathbf{x}) + \text{prox}_{f^*}(\mathbf{x}).$$



定义 3 (Moreau envelop)

设 f 为闭凸函数, $\mu > 0$, f 的 Moreau envelop 定义为

$$M_f^\mu(\mathbf{x}) = \min_{\mathbf{y}} \left\{ f(\mathbf{y}) + \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 \right\}$$

- $M_f^\mu(\mathbf{x})$ 为 $\frac{1}{\mu}$ -smooth 且

$$\nabla M_f^\mu(\mathbf{x}) = \frac{1}{\mu} (\mathbf{x} - \text{prox}_{\mu f}(\mathbf{x}))$$

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

例子

设 $f(\mathbf{x}) = \|\mathbf{x}\|_2$, $\mu > 0$, 则

$$\text{prox}_{\mu f}(\mathbf{x}) = \left(1 - \frac{\mu}{\max(\|\mathbf{x}\|_2, \mu)}\right) \mathbf{x}$$

$$M_f^\mu(\mathbf{x}) = \|\text{prox}_{\mu f}(\mathbf{x})\|_2 + \frac{1}{2\mu} \|\mathbf{x} - \text{prox}_{\mu f}(\mathbf{x})\|_2^2 = \begin{cases} \frac{1}{2\mu} \|\mathbf{x}\|_2^2, & \|\mathbf{x}\|_2 \leq \mu \\ \|\mathbf{x}\|_2 - \frac{\mu}{2}, & \|\mathbf{x}\|_2 > \mu. \end{cases}$$

梯度下降法求解 M_f^μ

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mu \nabla M_f^\mu(\mathbf{x}_k) = \text{prox}_{\mu f}(\mathbf{x}_k)$$

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

PPA

设 f 为闭凸函数, 求解优化问题

$$\min_{\mathbf{x}} f(\mathbf{x})$$

的 PPA 迭代格式为

$$\mathbf{x}_{k+1} = \text{prox}_{\gamma_k f}(\mathbf{x}_k).$$

即对目标函数使用 Moreau envelope 光滑化后的梯度下降法

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

收敛性分析

定理 4

设 f 为闭凸函数, $\mathbf{x}^*$ 为 f 的最小值点, 相应的最小值记为 f^* , 则临近点法 (PPA) 满足

$$f(\mathbf{x}_K) - f^* \leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|_2^2}{2 \sum_{k=0}^{K-1} \gamma_k} \quad \forall K \geq 1.$$

证明.

因为

$$\mathbf{x}_{k+1} = \text{prox}_{\gamma_k f}(\mathbf{x}_k) = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ f(\mathbf{x}) + \frac{1}{2\gamma_k} \|\mathbf{x} - \mathbf{x}_k\|_2^2 \right\}$$

故

$$\frac{1}{\gamma_k} (\mathbf{x}_k - \mathbf{x}_{k+1}) \in \partial f(\mathbf{x}_{k+1})$$

且

$$f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_{k+1}) + \frac{1}{2\gamma_k} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2^2 \leq f(\mathbf{x}_k)$$

由次梯度定义,

$$f(\mathbf{x}^*) \geq f(\mathbf{x}_{k+1}) + \left\langle \frac{1}{\gamma_k} (\mathbf{x}_k - \mathbf{x}_{k+1}), \mathbf{x}^* - \mathbf{x}_{k+1} \right\rangle$$



临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

续.

$$f(\mathbf{x}^*) \geq f(\mathbf{x}_{k+1}) + \left\langle \frac{1}{\gamma_k}(\mathbf{x}_k - \mathbf{x}_{k+1}), \mathbf{x}^* - \mathbf{x}_{k+1} \right\rangle$$

移项得到

$$\begin{aligned} \gamma_k(f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*)) &\leq \langle \mathbf{x}_{k+1} - \mathbf{x}_k, \mathbf{x}^* - \mathbf{x}_{k+1} \rangle \\ &= \frac{1}{2} \|\mathbf{x}_k - \mathbf{x}^*\|_2^2 - \frac{1}{2} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2^2 - \frac{1}{2} \|\mathbf{x}_{k+1} - \mathbf{x}^*\|_2^2 \\ &\leq \frac{1}{2} \|\mathbf{x}_k - \mathbf{x}^*\|_2^2 - \frac{1}{2} \|\mathbf{x}_{k+1} - \mathbf{x}^*\|_2^2. \end{aligned}$$

余下证明与以往的相同.



临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

考虑复合模型

$$\min_{\mathbf{x}} F(\mathbf{x}) \equiv f(\mathbf{x}) + g(\mathbf{x})$$

假设

- g 为闭凸函数;
- f 的定义域 $\text{dom } f$ 为凸集且 f 是 L_f -smooth;
- 最优解集 X^* 非空.

投影梯度法回顾

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

$$\begin{aligned}\mathbf{x}_{k+1} &= \Pi_C(\mathbf{x}_k - \gamma_k \nabla f(\mathbf{x}_k)) \\ &= \arg \min_{\mathbf{x} \in C} \left\{ f(\mathbf{x}_k) + \langle \nabla f(\mathbf{x}_k), \mathbf{x} - \mathbf{x}_k \rangle + \frac{1}{2\gamma_k} \|\mathbf{x} - \mathbf{x}_k\|_2^2 \right\}\end{aligned}$$

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

临近梯度法

$$\begin{aligned}
 \mathbf{x}_{k+1} &= \arg\min_{\mathbf{x}} \left\{ f(\mathbf{x}_k) + \langle \nabla f(\mathbf{x}_k), \mathbf{x} - \mathbf{x}_k \rangle + \frac{1}{2\gamma_k} \|\mathbf{x} - \mathbf{x}_k\|_2^2 + g(\mathbf{x}) \right\} \\
 &= \arg\min_{\mathbf{x}} \left\{ \gamma_k g(\mathbf{x}) + \frac{1}{2} \|\mathbf{x} - (\mathbf{x}_k - \gamma_k \nabla f(\mathbf{x}_k))\|_2^2 \right\} \\
 &= \text{prox}_{\gamma_k g}(\mathbf{x}_k - \gamma_k \nabla f(\mathbf{x}_k))
 \end{aligned}$$

记

$$\mathcal{T}_L^{f,g}(\mathbf{x}) = \text{prox}_{\frac{1}{L}g}\left(\mathbf{x} - \frac{1}{L}\nabla f(\mathbf{x})\right)$$

简记为 $\mathcal{T}_L(\mathbf{x})$.

定理 5

设 $L > 0$, 满足

$$f(\mathcal{T}_L(\mathbf{y})) \leq f(\mathbf{y}) + \langle \nabla f(\mathbf{y}), \mathcal{T}_L(\mathbf{y}) - \mathbf{y} \rangle + \frac{L}{2} \|\mathcal{T}_L(\mathbf{y}) - \mathbf{y}\|_2^2 \quad \forall \mathbf{x}, \mathbf{y},$$

则有

$$F(\mathbf{x}) - F(\mathcal{T}_L(\mathbf{y})) \geq \frac{L}{2} \|\mathbf{x} - \mathcal{T}_L(\mathbf{y})\|_2^2 - \frac{L}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \ell_f(\mathbf{x}, \mathbf{y}),$$

其中

$$\ell_f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle.$$

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

证明.

$$\mathcal{T}_L(\mathbf{y}) = \arg \min_{\mathbf{x}} \left\{ \frac{1}{L}g(\mathbf{x}) + \frac{1}{2} \left\| \mathbf{x} - \left(\mathbf{y} - \frac{1}{L} \nabla f(\mathbf{y}) \right) \right\|^2 \right\}$$

由最优性条件

$$L(\mathbf{y} - \frac{1}{L} \nabla f(\mathbf{y}) - \mathcal{T}_L(\mathbf{y})) \in \partial g(\mathcal{T}_L(\mathbf{y}))$$

因此

$$g(\mathbf{x}) \geq g(\mathcal{T}_L(\mathbf{y})) + \left\langle L(\mathbf{y} - \frac{1}{L} \nabla f(\mathbf{y}) - \mathcal{T}_L(\mathbf{y})), \mathbf{x} - \mathcal{T}_L(\mathbf{y}) \right\rangle.$$



临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

证明.

由定理条件得到

$$\begin{aligned}
 f(\mathbf{x}) &\geq f(\mathcal{T}_L(\mathbf{y})) - \langle \nabla f(\mathbf{y}), \mathcal{T}_L(\mathbf{y}) - \mathbf{y} \rangle - \frac{L}{2} \|\mathcal{T}_L(\mathbf{y}) - \mathbf{y}\|_2^2 + f(\mathbf{x}) - f(\mathbf{y}) \\
 &= f(\mathcal{T}_L(\mathbf{y})) + \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathcal{T}_L(\mathbf{y}) \rangle - \frac{L}{2} \|\mathcal{T}_L(\mathbf{y}) - \mathbf{y}\|_2^2 + \ell_f(\mathbf{x}, \mathbf{y})
 \end{aligned}$$

两不等式左右端相加得到

$$\begin{aligned}
 F(\mathbf{x}) &\geq F(\mathcal{T}_L(\mathbf{y})) + L \langle \mathbf{y} - \mathcal{T}_L(\mathbf{y}), \mathbf{x} - \mathcal{T}_L(\mathbf{y}) \rangle - \frac{L}{2} \|\mathcal{T}_L(\mathbf{y}) - \mathbf{y}\|_2^2 + \ell_f(\mathbf{x}, \mathbf{y}) \\
 &= F(\mathcal{T}_L(\mathbf{y})) + \frac{L}{2} \|\mathbf{x} - \mathcal{T}_L(\mathbf{y})\|_2^2 - \frac{L}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \ell_f(\mathbf{x}, \mathbf{y}).
 \end{aligned}$$



临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

凸函数收敛性

定理 6

若 f 为凸函数, 设 $\mathbf{x}_*$ 为复合优化问题的最优解, 取步长 $\gamma_k = \frac{1}{L_f}$, 则临近梯度法的迭代序列 $\{\mathbf{x}_k\}_{k \geq 0}$ 满足

$$F(\mathbf{x}_k) - F(\mathbf{x}^*) \leq \frac{L_f \|\mathbf{x}_0 - \mathbf{x}^*\|_2^2}{2k} \quad \forall k \geq 1.$$

证明.

在定理 5 中, 取 $\mathbf{x} = \mathbf{y} = \mathbf{x}_k$, $\mathcal{T}_{L_f}(\mathbf{x}_k) = \mathbf{x}_{k+1}$, 此时有

$$F(\mathbf{x}_k) - F(\mathbf{x}_{k+1}) \geq \frac{L_f}{2} \|\mathbf{x}_k - \mathbf{x}_{k+1}\|_2^2 \geq 0.$$

故序列 $\{F(\mathbf{x})_k\}_{k \geq 0}$ 单调非增.

取 $\mathbf{x} = \mathbf{x}^*$, $\mathbf{y} = \mathbf{x}_k$, $\mathbf{x}_{k+1} = \mathcal{T}_{L_f}(\mathbf{x}_k)$, 此时

$$F(\mathbf{x}^*) - F(\mathbf{x}_{k+1}) \geq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_{k+1}\|_2^2 - \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_k\|_2^2 + \ell_f(\mathbf{x}^*, \mathbf{x}_k)$$

移项得到

$$F(\mathbf{x}_{k+1}) - F(\mathbf{x}^*) \leq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_k\|_2^2 - \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_{k+1}\|_2^2 - \ell_f(\mathbf{x}^*, \mathbf{x}_k)$$

□

临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

续.

$$\begin{aligned}
 F(\mathbf{x}_{k+1}) - F(\mathbf{x}^*) &\leq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_k\|_2^2 - \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_{k+1}\|_2^2 - \ell_f(\mathbf{x}^*, \mathbf{x}_k) \\
 &\leq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_k\|_2^2 - \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_{k+1}\|_2^2 \quad (\text{convexity of } f)
 \end{aligned}$$

取 $k = 0, 1, \dots, K$, 对 $K+1$ 组不等式求和得到

$$\begin{aligned}
 (K+1)(F(\mathbf{x}_{K+1}) - F(\mathbf{x}^*)) &\leq \sum_{k=0}^K (F(\mathbf{x}_{k+1}) - f(\mathbf{x}^*)) \\
 &\leq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_0\|_2^2 - \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_{K+1}\|_2^2 \\
 &\leq \frac{L_f}{2} \|\mathbf{x}^* - \mathbf{x}_0\|_2^2
 \end{aligned}$$



临近算子

临近点法 (Proximal Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

强凸函数收敛性

定理 7

若 f 为 σ -强凸函数, 设 $\mathbf{x}_*$ 为复合优化问题的最优解, 取步长 $\gamma_k = \frac{1}{L_f}$, 则临近梯度法的迭代序列 $\{\mathbf{x}_k\}_{k \geq 0}$ 满足, $\forall k \geq 0$

- (a) $\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_2^2 \leq \left(1 - \frac{\sigma}{L_f}\right) \|\mathbf{x}_k - \mathbf{x}_*\|_2^2;$
- (b) $\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_2^2 \leq \left(1 - \frac{\sigma}{L_f}\right)^k \|\mathbf{x}_0 - \mathbf{x}_*\|_2^2;$
- (c) $F(\mathbf{x}_{k+1}) - F(\mathbf{x}^*) \leq \frac{L_f}{2} \left(1 - \frac{\sigma}{L_f}\right)^{k+1} \|\mathbf{x}_0 - \mathbf{x}^*\|_2^2.$

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

证明.

定理 5 的直接结果.



临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

在复合模型中, 假设

- g 为闭凸函数;
- f 是 L_f -smooth 且为凸函数;
- 最优解集 X^* 非空.

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

FISTA 迭代格式

$$\begin{aligned}\mathbf{x}_{k+1} &= \text{prox}_{\gamma_k g}(\mathbf{y}_k - \gamma_k \nabla f(\mathbf{y}_k)), \\ t_{k+1} &= \frac{1 + \sqrt{1 + 4t_k^2}}{2}, \\ \mathbf{y}_{k+1} &= \mathbf{x}_{k+1} + \frac{t_k - 1}{t_{k+1}}(\mathbf{x}_{k+1} - \mathbf{x}_k).\end{aligned}$$

其中 $\mathbf{x}_0 = \mathbf{y}_0$, $t_0 = 1$.

临近算子

临近点法 (Proximal
Point Algorithm, PPA)

临近梯度法

加速临近梯度法-FISTA

FISTA 收敛性

定理 8

设 $\mathbf{x}_*$ 为复合优化问题的最优解, 取步长 $\gamma_k = \frac{1}{L_f}$, 则 *FISTA* 的迭代序列 $\{\mathbf{x}_k\}_{k \geq 0}$ 满足

$$F(\mathbf{x}_k) - F(\mathbf{x}^*) \leq \frac{2L_f \|\mathbf{x}_0 - \mathbf{x}^*\|_2^2}{(k+1)^2} \quad \forall k \geq 1.$$