计算机91刘青帅第四次作业



作业1

 $\min f(x)$ **迭代求解**最优性条件 $\nabla f(x^*) = 0$

假设:

- ✓ 存在m, M满足 $mI \leq \nabla^2 f(x) \leq MI \ (\forall x \in S)$
- ✓ 存在常数 $\gamma \in (0,1]$, 使得 $\|x\|_* \ge \gamma \|x\|_2$

证明:

- 1. 精确直线搜索时, 非规范化最速下降方法的收敛性;
- 2. 回溯直线搜索时,非规范化最速下降方法的收敛性;

 $egin{aligned} &
abla f(x^k)^T \ & d^k_{sd} \ & \|
abla f(x^k)\|_* \ & f(x^k+td^k_{sd}) \ & f(x^{k+1}) \ & f(x^k) \ & \gamma \end{aligned}$

已经如下公式:

 $\check{\gamma}$

$$mI \leq
abla^2 f(x) \leq MI - --$$
 ① $abla f(x^k)^T d^k_{sd} = - \parallel
abla f(x^k) \parallel^2_* - --$ ② $abla f(x^k)^T d^k_{sd} = - \parallel
abla f(x^k) \parallel^2_* - --$ ②

存在
$$\check{\gamma}\in(0,1]$$
,使得 $\parallel x\parallel_*\geq\check{\gamma}\parallel x\parallel_2---$ ④ $lpha<rac{1}{2}---$ ⑤

$$egin{aligned} \|d_{sd}^k\parallel=igg| &\|
abla f(x^k)\parallel_* imes d_{nsd}^k &\| &=\|
abla f(x^k)\parallel_*\|
abla f(x^k)-p^* &\leq rac{1}{2m}\|
abla f(x^k)\parallel_2^2--rac{1}{2m}\|
abla f(x^k)\parallel_2^2-abla f(x^k)\parallel_2^2-abla f(x^k)\parallel_2^2-land f(x^k)\parallel_2^2-lan$$

回溯直线搜索

$$f(x^{k+1}) = f(x^k + td_{sd}^k) \le f(x^k) + t\nabla f(x^k)^T d_{sd}^k + \frac{M}{2} \parallel d_{sd}^k \parallel_2^2 t^2 - -$$
由泰勒展开和① $\Rightarrow f(x^{k+1}) \le f(x^k + td_{sd}^k) \le f(x^k) + t\nabla f(x^k)^T d_{sd}^k + \frac{M}{2\gamma^2} \parallel d_{sd}^k \parallel^2 t^2 - -$ 由③ $\Rightarrow f(x^{k+1}) \le f(x^k) - t \parallel \nabla f(x^k) \parallel_*^2 + \frac{Mt^2}{2\gamma^2} \parallel \nabla f(x^k) \parallel_*^2 - -$ 由②⑥

上式右侧可看成关于t的一元二次函数,取极小时, $t_{min}=\frac{\gamma^2}{M}---8$ 代入得

$$egin{aligned} f(x^k + t_{min} d_{sd}^k) & \leq f(x^k) - rac{\gamma^2}{2M} \parallel
abla f(x^k) \parallel_*^2 = f(x^k) - rac{1}{2} t_{min} \parallel
abla f(x^k) \parallel_*^2 - - eta \otimes f(x^k + t_{min} d_{sd}^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) \parallel_*^2 - - eta \otimes f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) \parallel_*^2 - - eta \otimes f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) - lpha t_{min} \parallel
abla f(x^k) & \leq f(x^k) -$$

由上述证明过程可知 $t \geq min(1, \frac{\beta\gamma^2}{M}),$ 记 $\theta = min(1, \frac{\beta\gamma^2}{M}),$ 则有 $t \geq \theta - -9$

$$f(x^{k+1}) \le f(x^k) - \alpha t_{min} \| \nabla f(x^k) \|_*^2$$

$$\Rightarrow f(x^{k+1}) \leq f(x^k) - lpha heta \parallel
abla f(x^k) \parallel_*^2 - - heta$$

$$\Rightarrow f(x^{k+1}) \leq f(x^k) - lpha heta ilde{\gamma}^2 \parallel
abla f(x^k) \parallel_2^2 - -$$
 由④

$$0 \Rightarrow f(x^{k+1}) - p^* \leq f(x^k) - p^* - lpha heta \dot{\gamma}^2 \parallel
abla f(x^k) \parallel_2^2 - -$$
 左右同时 $-p^*$

$$\phi \Rightarrow f(x^{k+1}) - p^* \leq (1 - 2mlpha \check{\gamma}^2 heta) (f(x^k) - p^*) - -$$
由⑦

ப்
$$c=1-2mlpha\check{\gamma}^2 heta$$

有
$$f(x^{\mathsf{K}}) - p^* \leq c^K (f(x^0) - p^*)$$

由迭代停止条件 $f(x^K) - p^* \le \varepsilon$

$$extrm{ }\Rightarrow K \geq rac{log(f(x^0-p^*)/arepsilon)}{log(1/c)}$$
其中 $c=1-2mlpha\check{\gamma}^2min(1,rac{eta\gamma^2}{M})$

精确直线搜索(由于证明过程很多与上述类似,故不在详细给出每一步证明,只给出重要步骤)

$$egin{aligned} f(x^{k+1}) &= f(x^k + t d_{sd}^k) \leq f(x^k) + t
abla f(x^k)^T d_{sd}^k + rac{M}{2} \parallel d_{sd}^k \parallel_2^2 t^2 \end{aligned}$$
 $\Rightarrow f(x^{k+1}) \leq f(x^k) - t \parallel
abla f(x^k) \parallel_*^2 + rac{Mt^2}{2} \parallel
abla f(x^k) \parallel_*^2 - - 曲②⑥$

上式右侧可看成关于t的一元二次函数,取极小时, $t=\frac{1}{M}$ 代入有

$$f(x^{k+1}) \leq f(x^k) - rac{1}{2M} \parallel
abla f(x^k) \parallel_*^2$$

$$\Rightarrow f(x^k) - f(x^{k+1}) \geq rac{1}{2M} \parallel
abla f(x^k) \parallel_*^2$$

$$\Rightarrow f(x^k) - f(x^{k+1}) \geq rac{\gamma^2}{2M} \parallel
abla f(x^k) \parallel_2^2 - -$$
 曲③

$$\Rightarrow f(x^{k+1}) - p^* \leq (1 - rac{m\gamma^2}{M})(f(x^k) - p^*) - -$$
由⑦

记
$$c=1-rac{m\gamma^2}{M}$$

有
$$f(x^{\mathsf{K}}) - p^* \leq c^{\mathsf{K}}(f(x^0) - p^*)$$

由迭代停止条件
$$f(x^K) - p^* \leq \varepsilon$$

$$\Rightarrow K \geq rac{log(f(x^0-p^*)/arepsilon)}{log(1/c)}$$
其中 $c = 1 - rac{m\gamma^2}{M}$