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$mI \le \nabla^2 f(x) \le MI\ (\forall x \in S)$
- ✓ 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;

$\nabla f(x^k)^T$

d_{sd}^k

$\|\nabla f(x^k)\|_*$

$f(x^k + td_{sd}^k)$

$f(x^{k+1})$

$f(x^k)$

γ

$\check{\gamma}$

已经如下公式：

$mI \le \nabla^2 f(x) \le MI ---①$

$\nabla f(x^k)^T d_{sd}^k = -\parallel \nabla f(x^k) \parallel_*^2 ---②$

存在$\gamma \in (0,1]$,使得$\parallel x \parallel \ge \gamma \parallel x \parallel_2 ---③$

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

$\alpha < \frac{1}{2} ---⑤$

$$\|d_{sd}^k\| = \left\| \ \|\nabla f(x^k)\|_* \times d_{nsd}^k \ \right\| = \parallel \nabla f(x^k) \parallel_* \parallel \nabla d_{nsd}^k \parallel = \parallel \nabla f(x^k) \parallel_* ---⑥$$

$f(x^k) - p^* \leq \frac{1}{2m} \parallel \nabla f(x^k) \parallel_2^2 --⑦$

回溯直线搜索

$f(x^{k+1}) = f(x^k + td_{sd}^k) \leq f(x^k) + t\nabla f(x^k)^T d_{sd}^k + \frac{M}{2} \parallel d_{sd}^k \parallel_2^2 t^2$ −−由泰勒展开和①

$\Rightarrow f(x^{k+1}) \leq f(x^k + td_{sd}^k) \leq f(x^k) + t\nabla f(x^k)^T d_{sd}^k + \frac{M}{2\gamma^2} \parallel d_{sd}^k \parallel^2 t^2$ −−由③

$\Rightarrow f(x^{k+1}) \leq f(x^k) - t \parallel \nabla f(x^k) \parallel_*^2 + \frac{Mt^2}{2\gamma^2} \parallel \nabla f(x^k) \parallel_*^2$ −− 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

$f(x^k + t_{min} d_{sd}^k) \leq f(x^k) - \frac{\gamma^2}{2M} \parallel \nabla f(x^k) \parallel_*^2 = f(x^k) - \frac{1}{2} t_{min} \parallel \nabla f(x^k) \parallel_*^2$ −− 由⑧

$\Rightarrow f(x^k + t_{min} d_{sd}^k) \leq f(x^k) - \alpha t_{min} \parallel \nabla f(x^k) \parallel_*^2$ −− 由⑤

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$t \geq \theta --⑨$

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

$\Rightarrow f(x^{k+1}) \leq f(x^k) - \alpha\theta \parallel \nabla f(x^k) \parallel_*^2$ −− 由⑨

$\Rightarrow f(x^{k+1}) \leq f(x^k) - \alpha\theta\check{\gamma}^2 \parallel \nabla f(x^k) \parallel_2^2$ −− 由④

$\Rightarrow f(x^{k+1}) - p^* \leq f(x^k) - p^* - \alpha\theta\check{\gamma}^2 \parallel \nabla f(x^k) \parallel_2^2$ −− 左右同时 $-p^*$

$\Rightarrow f(x^{k+1}) - p^* \leq \quad (1 - 2m\alpha\check{\gamma}^2\theta)(f(x^k) - p^*)$ −−由⑦

记$c = 1 - 2m\alpha\check{\gamma}^2\theta$

有$f(x^K) - p^* \leq \quad c^K(f(x^0) - p^*)$

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

$\Rightarrow K \geq \frac{log(f(x^0 - p^*)/\varepsilon)}{log(1/c)}$其中$c = 1 - 2m\alpha\check{\gamma}^2 min(1, \frac{\beta\gamma^2}{M})$

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)

$f(x^{k+1}) = f(x^k + td_{sd}^k) \leq f(x^k) + t\nabla f(x^k)^T d_{sd}^k + \frac{M}{2} \parallel d_{sd}^k \parallel_2^2 t^2$

$\Rightarrow f(x^{k+1}) \leq f(x^k) - t \parallel \nabla f(x^k) \parallel_*^2 + \frac{Mt^2}{2} \parallel \nabla f(x^k) \parallel_*^2$ −− 由②⑥

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

$f(x^{k+1}) \leq f(x^k) - \frac{1}{2M} \parallel \nabla f(x^k) \|_*^2$

$\Rightarrow f(x^k)-f(x^{k+1})\geq\frac{1}{2M}\parallel\nabla f(x^k)\|_*^2$

$\Rightarrow f(x^k)-f(x^{k+1})\geq\frac{\gamma^2}{2M}\parallel\nabla f(x^k)\parallel_2^2--$由③

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

记$c=1-\frac{m\gamma^2}{M}$

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

由迭代停止条件$f(x^K)-p^*\leq\varepsilon$

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