



最

东南大学
& 人 能学

@ . .



无



- ☐ 和
- ☐ 下
- ☐ 最 下
- ☐
- ☐ 自和
- ☐ 现





无



minimize $f_1(x)$



无



minimize $f_1(x)$



为 次 可微



无



minimize $f_1(x)$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有



无



minimize $f_1(x)$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有
- 无



无



minimize $f_1(x)$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有
- 无
- 生点 $x^{(k)} \in \mathbf{dom} f, k = 0, 1, \dots$ 中



无



minimize $f_1(x)$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有
- 无
- 生点 $x^{(k)} \in \text{dom } f, k = 0, 1, \dots$ 中
 $f(x^{(k)}) \rightarrow p^*$



无



minimize $f_1(x)$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有
- 无
- 生点 $x^{(k)} \in \text{dom } f, k = 0, 1, \dots$ 中
 $f(x^{(k)}) \rightarrow p^*$
- 过 最 的



无



$$\text{minimize } f(x)$$

- 为 次 可微
- 最 $p^* = \inf_x f(x)$ 在 为有
- 无
- 生点 $x^{(k)} \in \text{dom } f, k = 0, 1, \dots$ 中
 $f(x^{(k)}) \rightarrow p^*$
- 过 最 的
 $\nabla f(x^*) = 0$



点和下





点和下



要一个点 $x^{(0)}$



点和下



要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$



点和下



要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$



下 $S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为



点和下



要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$



下

$S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为



2

有的下 是



点和下



要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$



下 $S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为



2

有的下 是



$\mathbf{epi} f$ 是 的



$\mathbf{dom} f = \mathbf{R}^n$ 为真



点和下



- 要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$
- 下 $S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为
- 2 有的下 是
- $\mathbf{epi} f$ 是 的
- $\mathbf{dom} f = \mathbf{R}^n$ 为真
- $f(x) \rightarrow \infty$ as $x \rightarrow \mathbf{bd} \mathbf{dom} f$ 为真



点和下



- 要一个点 $x^{(0)}$
 $x^{(0)} \in \mathbf{dom} f$
- 下 $S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为
- 2 有的下 是
- $\mathbf{epi} f$ 是 的
- $\mathbf{dom} f = \mathbf{R}^n$ 为真
- $f(x) \rightarrow \infty$ as $x \rightarrow \mathbf{bd} \mathbf{dom} f$ 为真
- 不 有 的下 的可微



点和下



□ 要一个 点 $x^{(0)}$

$$x^{(0)} \in \mathbf{dom} f$$

□ 下 $S = \{x \mid f(x) \leq f(x^{(0)})\}$ 为

□ 2 有的下 是

□ $\mathbf{epi} f$ 是 的

□ $\mathbf{dom} f = \mathbf{R}^n$ 为真

□ $f(x) \rightarrow \infty$ as $x \rightarrow \mathbf{bd} \mathbf{dom} f$ 为真

□ 不 有 的下 的可微

$$f(x) = \log\left(\sum_{i=1}^m \exp(a_i^T x + b_i)\right), \quad f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$$





在 S 上 的 在 $m > 0$



在 S 上的 $\nabla^2 f(x)$ 在 $m > 0$
 $\nabla^2 f(x) \succeq mI$ for all $x \in S$



□ 在 S 上 的 在 $m > 0$

$$\nabla^2 f(x) \succeq mI \text{ for all } x \in S$$

□ 对 $x, y \in S$,



□ 在 S 上 的 在 $m > 0$

$$\nabla^2 f(x) \succeq mI \text{ for all } x \in S$$

□ 对 $x, y \in S$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|_2^2$$



□ 在 S 上 的 在 $m > 0$

$$\nabla^2 f(x) \succeq mI \quad \text{for all } x \in S$$

□ 对 $x, y \in S$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|_2^2$$

□ $p^* > -\infty$ 对 $x \in S$,



□ 在 S 上 的 在 $m > 0$

$$\nabla^2 f(x) \succeq mI \quad \text{for all } x \in S$$

□ 对 $x, y \in S$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|_2^2$$

□ $p^* > -\infty$ 对 $x \in S$,

$$f(x) - p^* \leq \frac{1}{2m} \|\nabla f(x)\|_2^2$$



□ 在 S 上 的 在 $m > 0$

$$\nabla^2 f(x) \succeq mI \quad \text{for all } x \in S$$

□ 对 $x, y \in S$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|_2^2$$

□ $p^* > -\infty$ 对 $x \in S$,

$$f(x) - p^* \leq \frac{1}{2m} \|\nabla f(x)\|_2^2$$

□ 为 分有用 m 知



下





下



$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$



下



$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$



他

$$x^+ = x + t \Delta x, \quad x := x + t \Delta x$$



下



$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$

□ 他 $x^+ = x + t\Delta x, x := x + t\Delta x$

□ Δx 为 t 为



下



$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$

□ 他 $x^+ = x + t\Delta x, x := x + t\Delta x$

□ Δx 为 t 为

□ $f(x^+) < f(x) \quad \nabla f(x)^T \Delta x < 0 \quad \Delta x$ 为下



下



$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$

□ 他 $x^+ = x + t\Delta x, x := x + t\Delta x$

□ Δx 为 t 为

□ $f(x^+) < f(x) \quad \nabla f(x)^T \Delta x < 0 \quad \Delta x$ 为下

算法 9.1 通用下降方法。

给定初始点 $x \in \text{dom } f$ 。

重复进行

1. 确定下降方向 Δx 。
2. 直线搜索。选择步长 $t > 0$ 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。





$$t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$$



$$t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$$



回

为 $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$



$$t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$$



回

为 $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$



$t = 1$ 开

$t := \beta t$ 到



$$t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$$



回

为 $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$



$t = 1$ 开

$t := \beta t$ 到

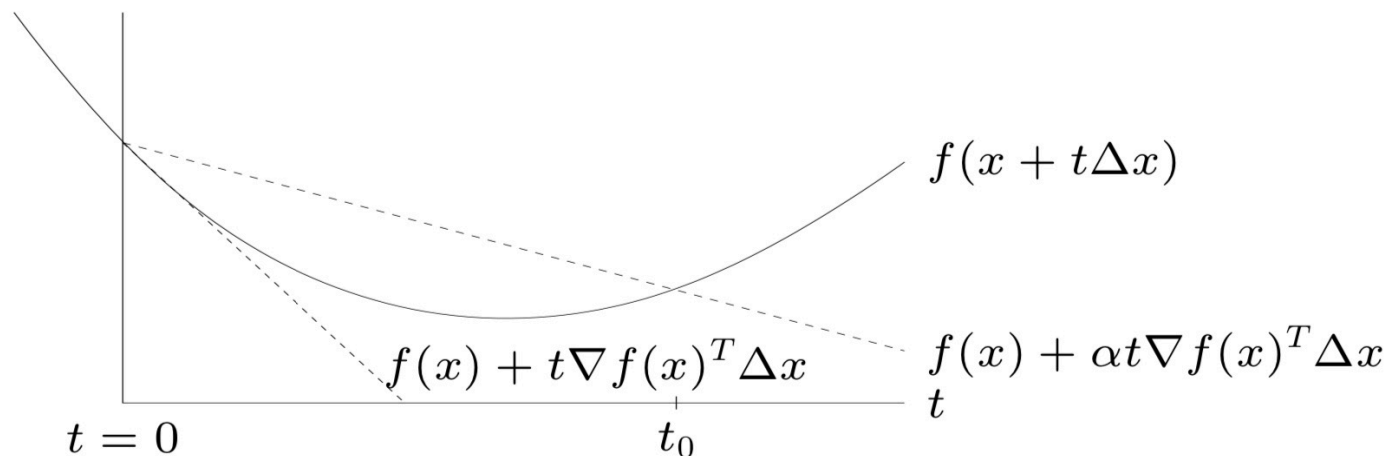
$$f(x + t\Delta x) < f(x) + \alpha t \nabla f(x)^T \Delta x$$



- $t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$
- 回 为 $\alpha \in (0, 1/2), \beta \in (0, 1)$
- $t = 1$ 开 $t := \beta t$ 到
 $f(x + t\Delta x) < f(x) + \alpha t \nabla f(x)^T \Delta x$
- 说 回 到 $t \leq t_0$



- $t = \operatorname{argmin}_{t>0} f(x + t\Delta x)$
- 回 为 $\alpha \in (0, 1/2), \beta \in (0, 1)$
- $t = 1$ 开 $t := \beta t$ 到
 $f(x + t\Delta x) < f(x) + \alpha t \nabla f(x)^T \Delta x$
- 说 回 到 $t \leq t_0$





下





下



下

中 $\Delta x = -\nabla f(x)$



下



下

中 $\Delta x = -\nabla f(x)$

算法 9.3 梯度下降方法。

给定 初始点 $x \in \text{dom } f$ 。

重复进行

1. $\Delta x := -\nabla f(x)$ 。
2. 直线搜索。通过精确或回溯直线搜索方法确定步长 t 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。



下



下

中 $\Delta x = -\nabla f(x)$

算法 9.3 梯度下降方法。

给定 初始点 $x \in \text{dom } f$ 。

重复进行

1. $\Delta x := -\nabla f(x)$ 。
2. 直线搜索。通过精确或回溯直线搜索方法确定步长 t 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。



如 $\|\nabla f(x)\|_2 \leq \epsilon$



下



□ 下 中 $\Delta x = -\nabla f(x)$

算法 9.3 梯度下降方法。

给定 初始点 $x \in \text{dom } f$ 。

重复进行

1. $\Delta x := -\nabla f(x)$ 。
2. 直线搜索。通过精确或回溯直线搜索方法确定步长 t 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。

□ 如 $\|\nabla f(x)\|_2 \leq \epsilon$

□ 对 $f(x^{(k)}) - p^* \leq c^k(f(x^{(0)}) - p^*)$



下



□ 下 中 $\Delta x = -\nabla f(x)$

算法 9.3 梯度下降方法。

给定 初始点 $x \in \text{dom } f$ 。

重复进行

1. $\Delta x := -\nabla f(x)$ 。
2. 直线搜索。通过精确或回溯直线搜索方法确定步长 t 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。

□ 如 $\|\nabla f(x)\|_2 \leq \epsilon$

□ 对 $f(x^{(k)}) - p^* \leq c^k(f(x^{(0)}) - p^*)$

□ $c \in (0, 1)$ $m, x^{(0)},$



下



下 中 $\Delta x = -\nabla f(x)$

算法 9.3 梯度下降方法。

给定 初始点 $x \in \text{dom } f$ 。

重复进行

1. $\Delta x := -\nabla f(x)$ 。
2. 直线搜索。通过精确或回溯直线搜索方法确定步长 t 。
3. 修改。 $x := x + t\Delta x$ 。

直到 满足停止准则。

如 $\|\nabla f(x)\|_2 \leq \epsilon$

对 $f(x^{(k)}) - p^* \leq c^k (f(x^{(0)}) - p^*)$

$c \in (0, 1)$ $m, x^{(0)},$

中很 用



R^2 中的 次





R^2 中的 次



$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$



R^2 中的 次



$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$



起 点为 $x^{(0)} = (\gamma, 1)$:



R^2 中的 次



$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$



起 点为 $x^{(0)} = (\gamma, 1)$:

$$x_1^{(k)} = \gamma \left(\frac{\gamma - 1}{\gamma + 1} \right)^k, \quad x_2^{(k)} = \left(-\frac{\gamma - 1}{\gamma + 1} \right)^k$$



R^2 中的 次



$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$



起 点为 $x^{(0)} = (\gamma, 1)$:

$$x_1^{(k)} = \gamma \left(\frac{\gamma - 1}{\gamma + 1} \right)^k, \quad x_2^{(k)} = \left(-\frac{\gamma - 1}{\gamma + 1} \right)^k$$



只 $\gamma \gg 1$ or $\gamma \ll 1$



R^2 中的 次



$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$



起 点为 $x^{(0)} = (\gamma, 1)$:

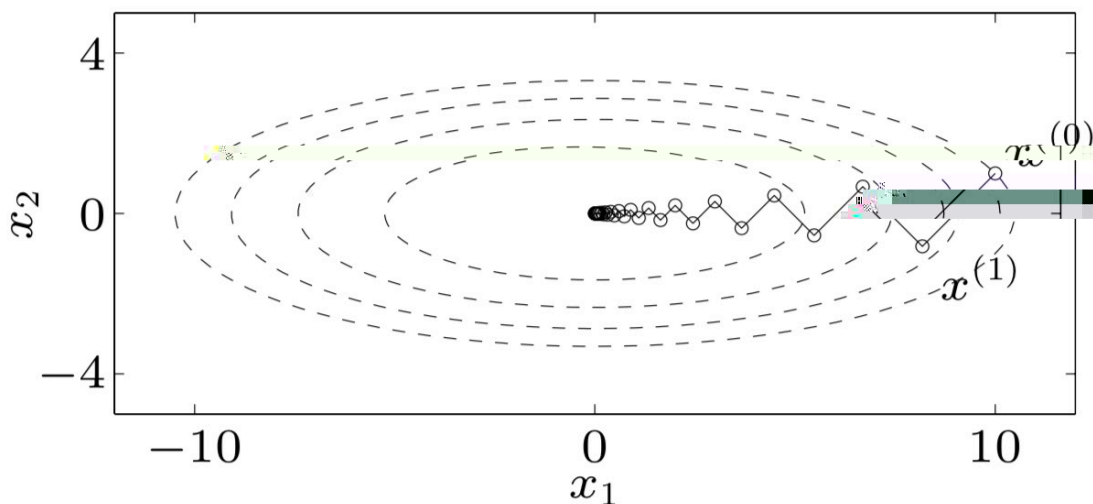
$$x_1^{(k)} = \gamma \left(\frac{\gamma - 1}{\gamma + 1} \right)^k, \quad x_2^{(k)} = \left(-\frac{\gamma - 1}{\gamma + 1} \right)^k$$



只 $\gamma \gg 1$ or $\gamma \ll 1$



$\gamma = 10$

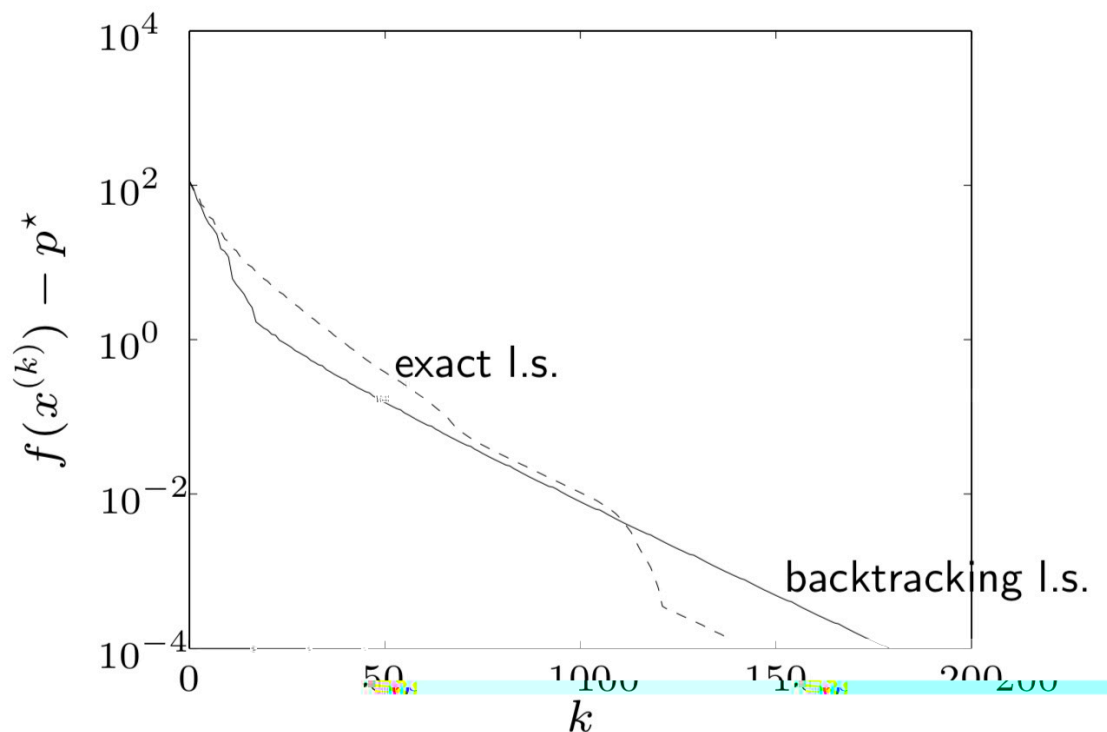




R^{100} 中的



$$f(x) = c^T x - \sum_{i=1}^{500} \log(b_i - a_i^T x)$$





最 下





最 下



的最 下

对

$\parallel \cdot \parallel$



最 下



的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$



最 下



- 的最 下 对 $\|\cdot\|$
 $\Delta x_{\text{nsd}} = \operatorname{argmin}\{\nabla f(x)^T v \mid \|v\| = 1\}$
- 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$



最 下



- 的最 下 对 $\|\cdot\|$
 $\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$
- 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$
- Δx_{nsd} 是 有最大 下 的 的



最 下



- 的最 下 对 $\|\cdot\|$
 $\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$
- 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$
- Δx_{nsd} 是 有最大 下 的 的
- 的最 下



最 下



□ 的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$

□ 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$

□ Δx_{nsd} 是 有最大 下 的 的

□ 的最 下

$$\Delta x_{\text{sd}} = \|\nabla f(x)\|_* \Delta x_{\text{nsd}}$$



最 下



□ 的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$

□ 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$

□ Δx_{nsd} 是 有最大 下 的 的

□ 的最 下

$$\Delta x_{\text{sd}} = \|\nabla f(x)\|_* \Delta x_{\text{nsd}}$$

□
$$\nabla f(x)^T \Delta x_{\text{sd}} = -\|\nabla f(x)\|_*^2$$



最 下



□ 的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$

□ 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$

□ Δx_{nsd} 是 有最大 下 的 的

□ 的最 下

$$\Delta x_{\text{sd}} = \|\nabla f(x)\|_* \Delta x_{\text{nsd}}$$

$$\nabla f(x)^T \Delta x_{\text{sd}} = -\|\nabla f(x)\|_*^2$$

□ 最 下



最 下



□ 的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$

□ 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$

□ Δx_{nsd} 是 有最大 下 的 的

□ 的最 下

$$\Delta x_{\text{sd}} = \|\nabla f(x)\|_* \Delta x_{\text{nsd}}$$

$$\nabla f(x)^T \Delta x_{\text{sd}} = -\|\nabla f(x)\|_*^2$$

□ 最 下

□ 用下 $\Delta x = \Delta x_{\text{sd}}$



最 下



□ 的最 下 对 $\|\cdot\|$

$$\Delta x_{\text{nsd}} = \operatorname{argmin}\{\|\nabla f(x)^T v\| \mid \|v\| = 1\}$$

□ 对 小的 v 有 $f(x+v) \approx f(x) + \nabla f(x)^T v$

□ Δx_{nsd} 是 有最大 下 的 的

□ 的最 下

$$\Delta x_{\text{sd}} = \|\nabla f(x)\|_* \Delta x_{\text{nsd}}$$

□ $\nabla f(x)^T \Delta x_{\text{sd}} = -\|\nabla f(x)\|_*^2$

□ 最 下

□ 用下 $\Delta x = \Delta x_{\text{sd}}$

□ 和 下 一





$$\Delta x_{\text{sd}} = -\nabla f(x)$$



$$\Delta x_{\text{sd}} = -\nabla f(x)$$



次 $\|x\|_P = (x^T P x)^{1/2} \ (P \in \mathbf{S}_{++}^n): \Delta x_{\text{sd}} = -P^{-1} \nabla f(x)$



$$\Delta x_{\text{sd}} = -\nabla f(x)$$



次

$$\|x\|_P = (x^T P x)^{1/2} \quad (P \in \mathbf{S}_{++}^n): \quad \Delta x_{\text{sd}} = -P^{-1} \nabla f(x)$$



ℓ_1 -

$$\Delta x_{\text{sd}} = -(\partial f(x)/\partial x_i)e_i, \text{ where } |\partial f(x)/\partial x_i| = \|\nabla f(x)\|_\infty$$



$$\Delta x_{\text{sd}} = -\nabla f(x)$$



次 $\|x\|_P = (x^T P x)^{1/2} \ (P \in \mathbf{S}_{++}^n): \Delta x_{\text{sd}} = -P^{-1} \nabla f(x)$



ℓ_1 - $\Delta x_{\text{sd}} = -(\partial f(x)/\partial x_i)e_i$, where $|\partial f(x)/\partial x_i| = \|\nabla f(x)\|_\infty$



次 和 ℓ_1 - 的 和 最 下



$$\Delta x_{\text{sd}} = -\nabla f(x)$$



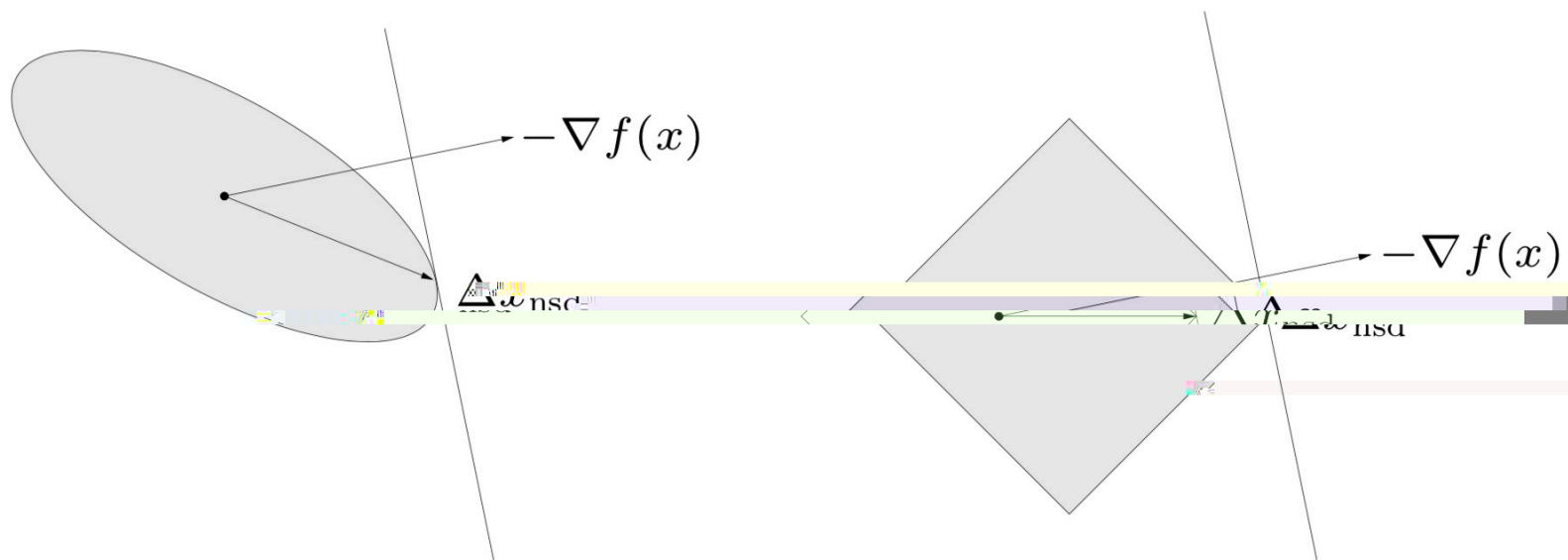
次 $\|x\|_P = (x^T P x)^{1/2} \ (P \in \mathbf{S}_{++}^n): \Delta x_{\text{sd}} = -P^{-1} \nabla f(x)$



ℓ_1 - $\Delta x_{\text{sd}} = -(\partial f(x)/\partial x_i)e_i$, where $|\partial f(x)/\partial x_i| = \|\nabla f(x)\|_\infty$



次 和 ℓ_1 - 的 和 最 下



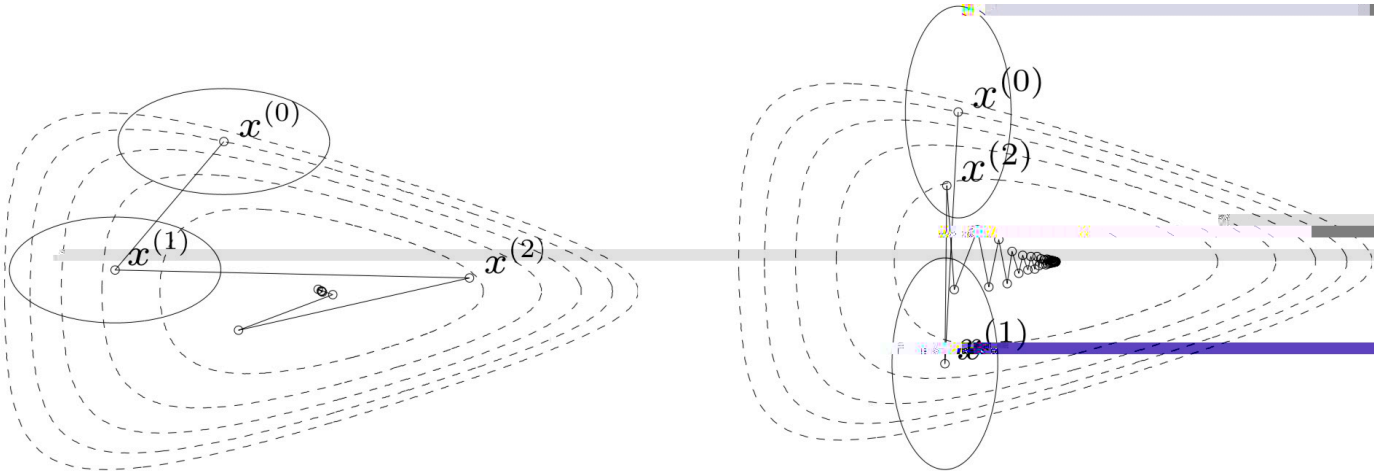


最 下 的



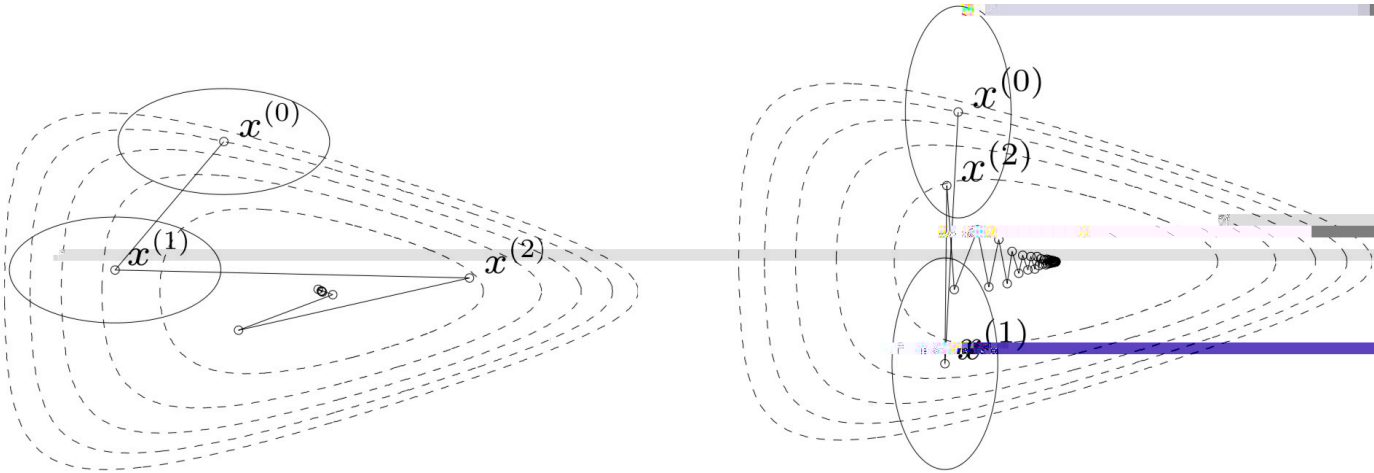


最 下 的





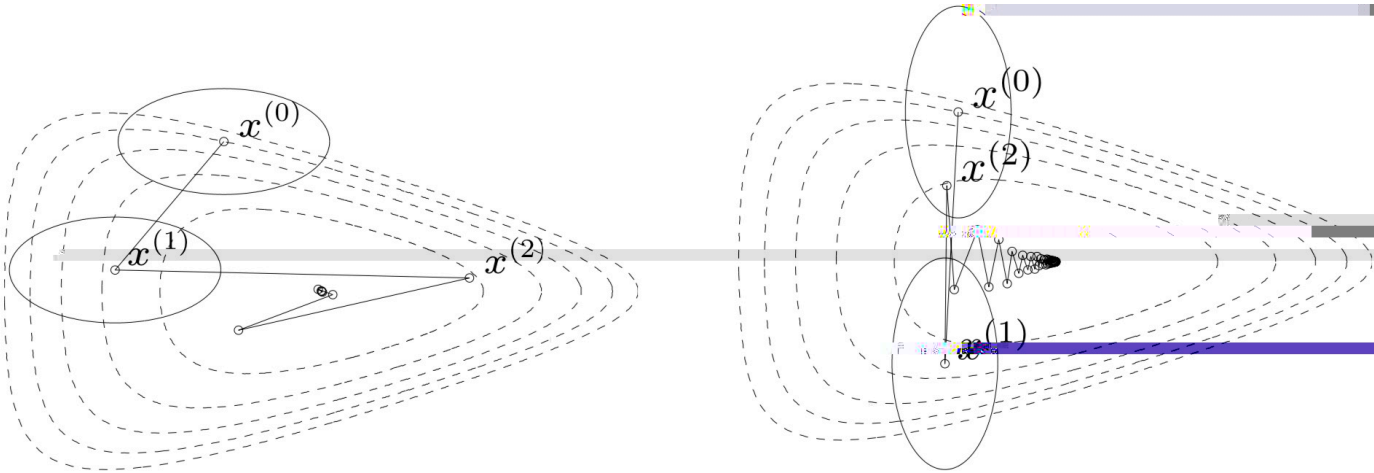
最 下 的



□ 分 用两个 次 和回 的最 下



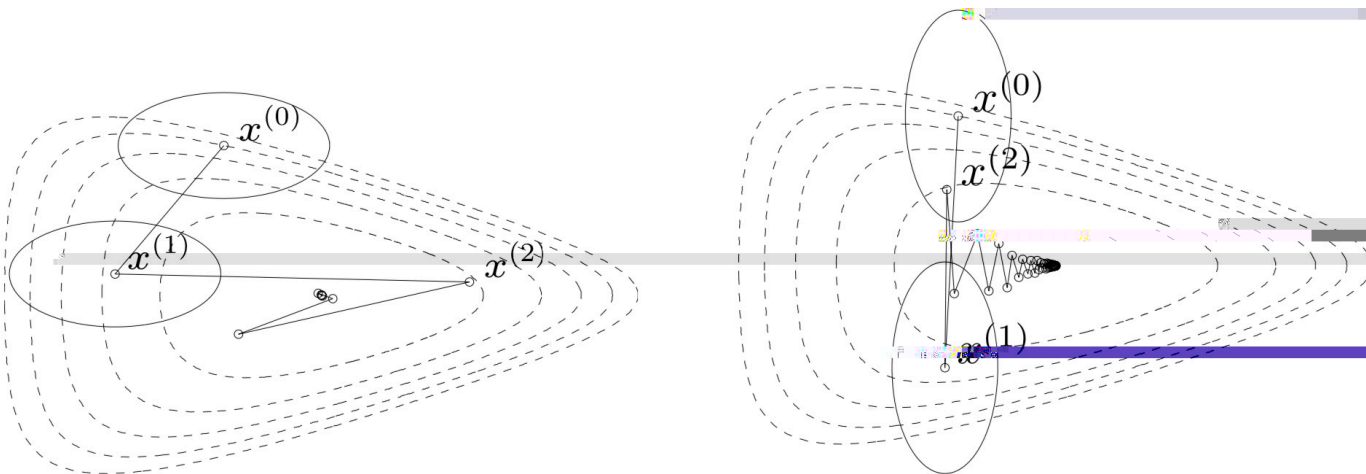
最 下 的



$$\{x \mid \|x - x^{(k)}\|_P = 1\}$$



最 下 的



分 用两个 次 和回 的最 下

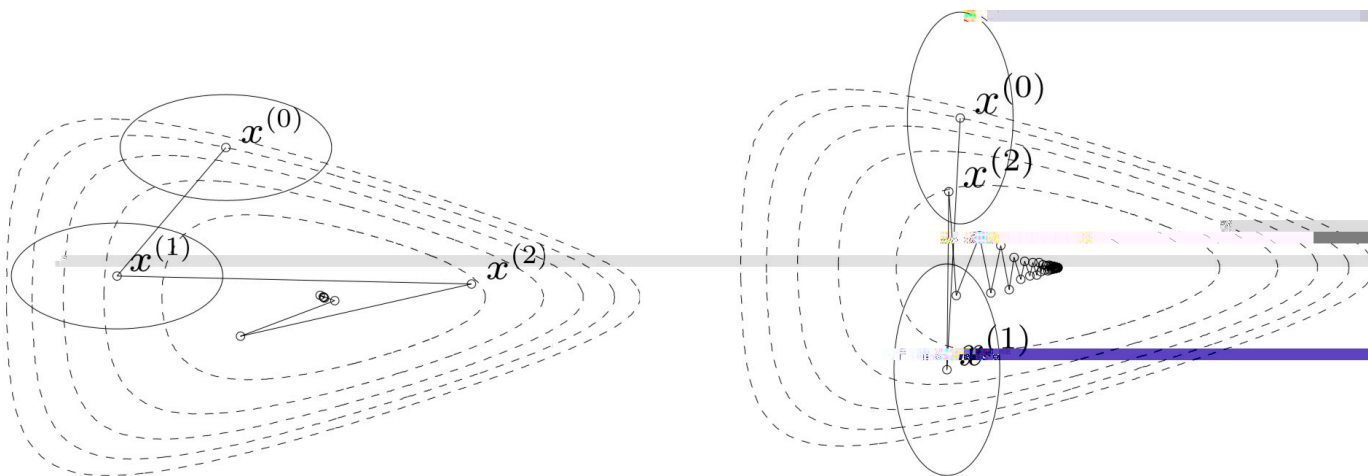
$\{x \mid \|x - x^{(k)}\|_P = 1\}$

次 最 下 的
的 下

$$\bar{x} = P^{1/2}r$$



最 下 的



分 用两个 次 和回 的最 下

$\{x \mid \|x - x^{(k)}\|_P = 1\}$

次 最 下 的
的 下

$$\bar{x} = P^{1/2}x$$

对 有很大





$$\Delta x_{\text{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$



$$\Delta x_{\text{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

□ $x + \Delta x_{\text{nt}}$ 最小



$$\Delta x_{\text{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

□ $x + \Delta x_{\text{nt}}$ 最小

$$\hat{f}(x + v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v,$$



$$\Delta x_{\text{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

□ $x + \Delta x_{\text{nt}}$ 最小

$$\hat{f}(x + v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v,$$

□ $x + \Delta x_{\text{nt}}$ 最



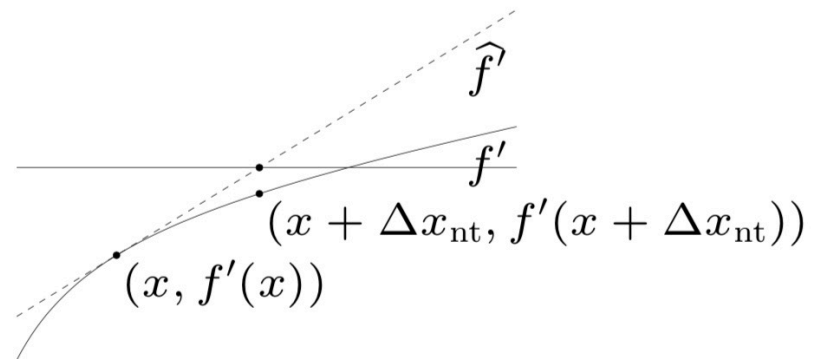
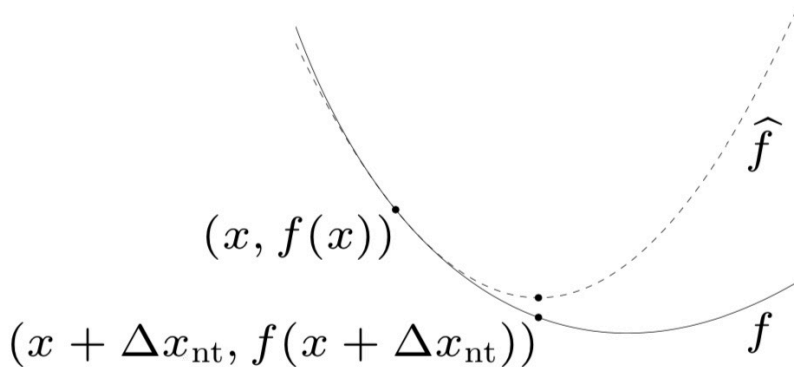
$$\Delta x_{\text{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

□ $x + \Delta x_{\text{nt}}$ 最小

$$\hat{f}(x + v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v,$$

□ $x + \Delta x_{\text{nt}}$ 最

$$\nabla f(x + v) \approx \nabla \hat{f}(x + v) = \nabla f(x) + \nabla^2 f(x) v = 0$$





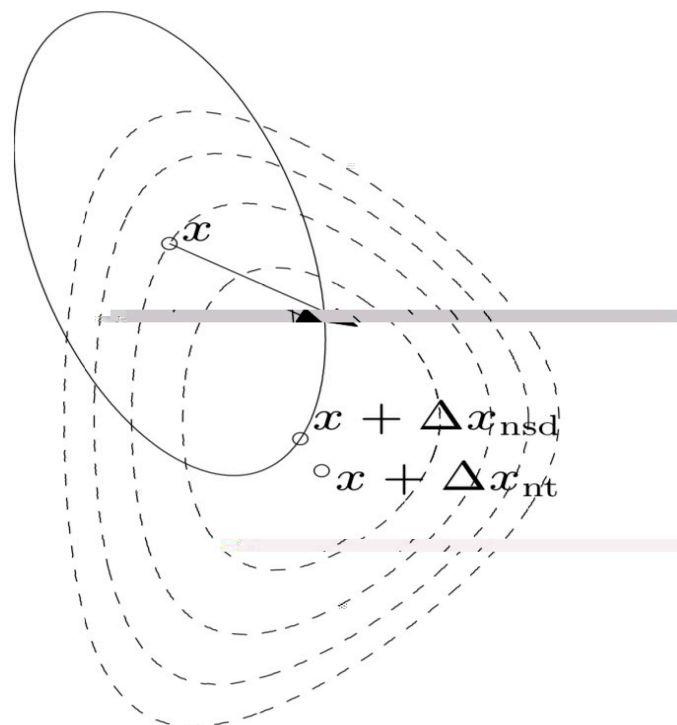


□ Δx_{nt} 为 $\mathbf{H}$ 的最 下



□ Δx_{nt} 为 $\mathbf{H}$ 的最 下

$$\|u\|_{\nabla^2 f(x)} = (u^T \nabla^2 f(x) u)^{1/2}$$

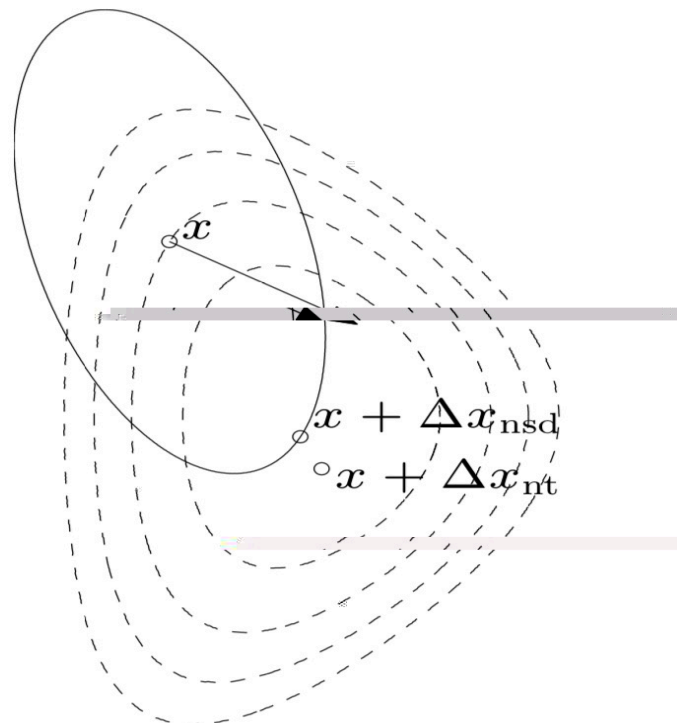




□ Δx_{nt} 为 $\mathbf{H}$ 的最 下

$$\|u\|_{\nabla^2 f(x)} = (u^T \nabla^2 f(x) u)^{1/2}$$

□ 为 的 平



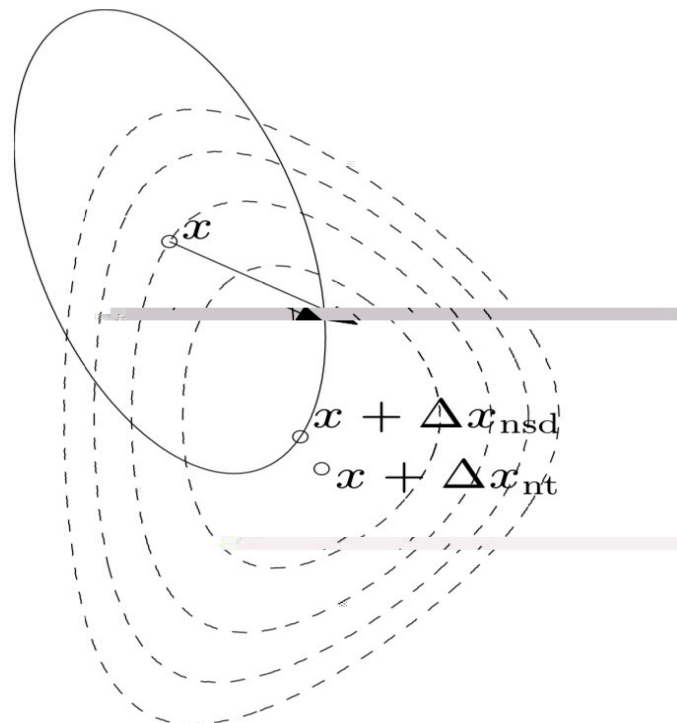


□ Δx_{nt} 为 $\mathbf{H}$ 的最 下

$$\|u\|_{\nabla^2 f(x)} = (u^T \nabla^2 f(x) u)^{1/2}$$

□ 为 的 ∇^2

□ 为 $\{x + v \mid v^T \nabla^2 f(x) v = 1\}$





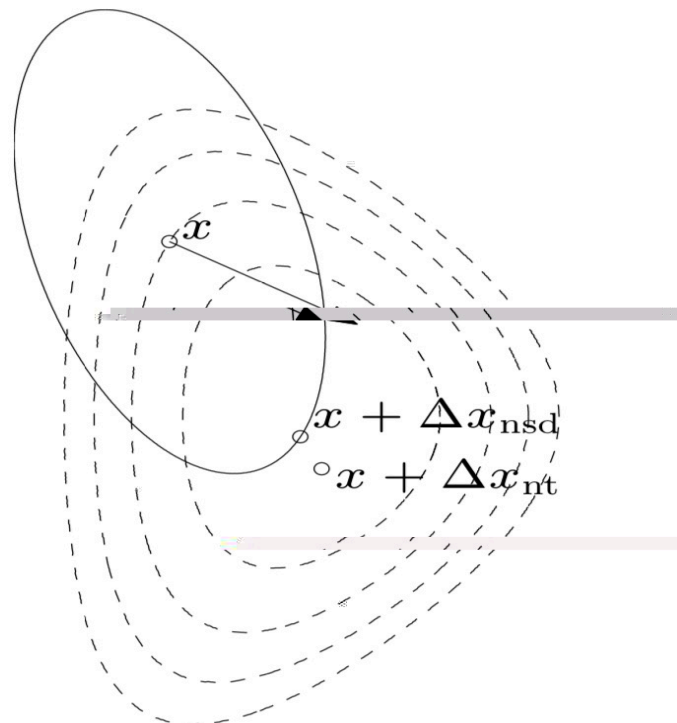
□ Δx_{nt} 为 $\mathbf{H}$ 的最 下

$$\|u\|_{\nabla^2 f(x)} = (u^T \nabla^2 f(x) u)^{1/2}$$

□ 为 的 ∇

□ 为 $\{x + v \mid v^T \nabla^2 f(x) v = 1\}$

□ $-\nabla f(x)$







$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{1/2}$$



$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{1/2}$$

□ 对 x 和 x^* 的



$$\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$$

□ 对 x 和 x^* 的





$$\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$$

□ 对 x 和 x^* 的



□ 给出了 $f(x) - p^*$ 的一个 用 次





$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{1/2}$$

□ 对 x 和 x^* 的



□ 给出了 $f(x) - p^*$ 的一个 用 次

$$f(x) - \inf_y \hat{f}(y) = \frac{1}{2} \lambda(x)^2$$

□ 为 的 次 有 次 $\mathbf{H}$



$$\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$$

□ 对 x 和 x^* 的



□ 给出了 $f(x) - p^*$ 的一个 用 次

$$f(x) - \inf_y \hat{f}(y) = \frac{1}{2} \lambda(x)^2$$

□ 为 的 次 有 次 $\mathbf{H}$

$$\lambda(x) = (\Delta x_{\text{nt}}^T \nabla^2 f(x) \Delta x_{\text{nt}})^{1/2}$$



$$\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$$

□ 对 x 和 x^* 的



□ 给出了 $f(x) - p^*$ 的一个 用 次

$$f(x) - \inf_y \hat{f}(y) = \frac{1}{2} \lambda(x)^2$$

□ 为 的 次 有 次 $\mathbf{H}$

$$\lambda(x) = (\Delta x_{\text{nt}}^T \nabla^2 f(x) \Delta x_{\text{nt}})^{1/2}$$



不 和 $\|\nabla f(x)\|_2$ 不一样





算法 9.5 Newton 方法。

给定 初始点 $x \in \text{dom } f$, 误差阈值 $\epsilon > 0$ 。

重复进行

1. 计算 Newton 步径和减量。

$$\Delta x_{\text{nt}} := -\nabla^2 f(x)^{-1} \nabla f(x); \quad \lambda^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)。$$

2. 停止准则。如果 $\lambda^2/2 \leq \epsilon$, 退出。
3. 直线搜索。通过回溯直线搜索确定步长 t 。
4. 改进。 $x := x + t \Delta x_{\text{nt}}$ 。



算法 9.5 Newton 方法。

给定 初始点 $x \in \text{dom } f$, 误差阈值 $\epsilon > 0$ 。

重复进行

1. 计算 Newton 步径和减量。

$$\Delta x_{\text{nt}} := -\nabla^2 f(x)^{-1} \nabla f(x); \quad \lambda^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)。$$

2. 停止准则。如果 $\lambda^2/2 \leq \epsilon$, 退出。
3. 直线搜索。通过回溯直线搜索确定步长 t 。
4. 改进。 $x := x + t \Delta x_{\text{nt}}。$



不 和 的 无







算法 9.5 Newton 方法。

给定 初始点 $x \in \text{dom } f$, 误差阈值 $\epsilon > 0$ 。

重复进行

1. 计算 Newton 步径和减量。

$$\Delta x_{\text{nt}} := -\nabla^2 f(x)^{-1} \nabla f(x); \quad \lambda^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)。$$

2. 停止准则。如果 $\lambda^2/2 \leq \epsilon$, 退出。
3. 直线搜索。通过回溯直线搜索确定步长 t 。
4. 改进。 $x := x + t \Delta x_{\text{nt}}。$

□ 不 和 的 无

□ $\tilde{f}(y) = f(Ty)$

□ 起 点为 $y^{(0)} = T^{-1}x^{(0)}$

$$y^{(k)} = T^{-1}x^{(k)}$$



分





分



f 在 S 上 有 为 m 的



分



- f 在 S 上 有 为 m 的
- $\nabla^2 f$ 在 S 上 有 以 L 为 的



分



□ f 在 S 上 有 为 m 的

□ $\nabla^2 f$ 在 S 上 有以 L 为 的

$$\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2$$



分



□ f 在 S 上 有 为 m 的

□ $\nabla^2 f$ 在 S 上 有以 L 为 的

$$\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2$$

□ L 了 用 次 f 的



分



□ f 在 S 上 有 为 m 的

□ $\nabla^2 f$ 在 S 上 有以 L 为 的

$$\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2$$

□ L 了 用 次 f 的

□ 在 $\eta \in (0, m^2/L), \gamma \geq 3$



分



□ f 在 S 上 有 为 m 的

□ $\nabla^2 f$ 在 S 上 有以 L 为 的

$$\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2$$

□ L 了 用 次 f 的

□ 在 $\eta \in (0, m^2/L), \gamma > 0$

if $\|\nabla f(x)\|_2 \geq \eta$, then $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$

if $\|\nabla f(x)\|_2 < \eta$, then

$$\frac{L}{2m^2} \|\nabla f(x^{(k+1)})\|_2 \leq \left(\frac{L}{2m^2} \|\nabla f(x^{(k)})\|_2 \right)^2$$





$$(\|\nabla f(x)\|_2 \geq \eta)$$



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回
□ 为 γ



$$(\|\nabla f(x)\|_2 \geq \eta)$$

- 大多 过 要回
- 为 γ
- $p^* > -\infty$ 最多 要的 次 是



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回

□ 为 γ

□ $p^* > -\infty$ 最多 要的 次 是

$$(f(x^{(0)}) - p^*)/\gamma$$



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回

□ 为 γ

□ $p^* > -\infty$ 最多 要的 次 是

$$(f(x^{(0)}) - p^*)/\gamma$$

□ 次 $(\|\nabla f(x)\|_2 < \eta)$



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回

□ 为 γ

□ $p^* > -\infty$ 最多 要的 次 是

$$(f(x^{(0)}) - p^*)/\gamma$$

□ 次 $(\|\nabla f(x)\|_2 < \eta)$

□ 有 用 $t = 1$



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回

□ 为 γ

□ $p^* > -\infty$ 最多 要的 次 是

$$(f(x^{(0)}) - p^*)/\gamma$$

□ 次 $(\|\nabla f(x)\|_2 < \eta)$

□ 有 用 $t = 1$

□ $\|\nabla f(x)\|_2$ 次 到 **0** $\|\nabla f(x^{(k)})\|_2 < \eta$



$$(\|\nabla f(x)\|_2 \geq \eta)$$

□ 大多 过 要回

□ 为 γ

□ $p^* > -\infty$ 最多 要的 次 是

$$(f(x^{(0)}) - p^*)/\gamma$$

□ 次 $(\|\nabla f(x)\|_2 < \eta)$

□ 有 用 $t = 1$

□ $\|\nabla f(x)\|_2$ 次 到 0 $\|\nabla f(x^{(k)})\|_2 < \eta$

$$\frac{L}{2m^2} \|\nabla f(x^l)\|_2 \leq \left(\frac{L}{2m^2} \|\nabla f(x^k)\|_2 \right)^{2^{l-k}} \leq \left(\frac{1}{2} \right)^{2^{l-k}}, \quad l \geq k$$





□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为



□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon)$$



□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon)$$

□ γ, ϵ_0 为 $m, L, x^{(0)}$



□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon)$$

□ γ, ϵ_0 为 $m, L, x^{(0)}$

□ 小 中 为



□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon)$$

□ γ, ϵ_0 为 $m, L, x^{(0)}$

□ 小 中 为

□ 上 m, L 知 γ, ϵ_0 也 知



□ 到 $f(x) - p^* \leq \epsilon$ 的 次 上 为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon)$$

□ γ, ϵ_0 为 $m, L, x^{(0)}$

□ 小 中 为

□ 上 m, L 知 γ, ϵ_0 也 知

□ 为 了 了两个

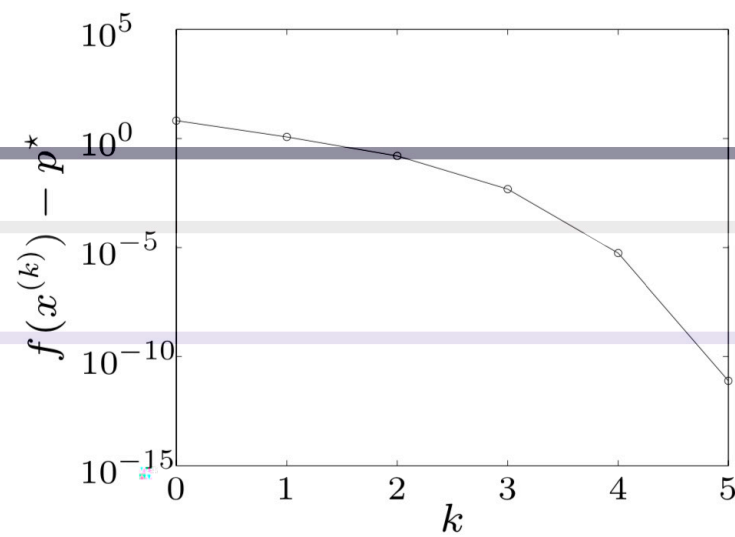
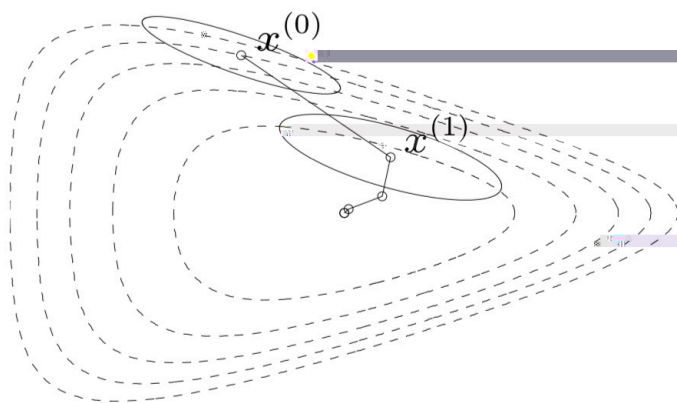




□ R^2 中的子

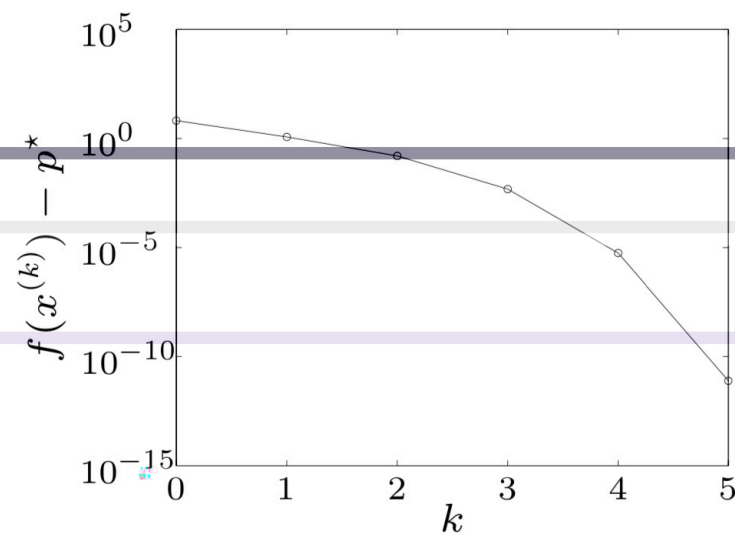
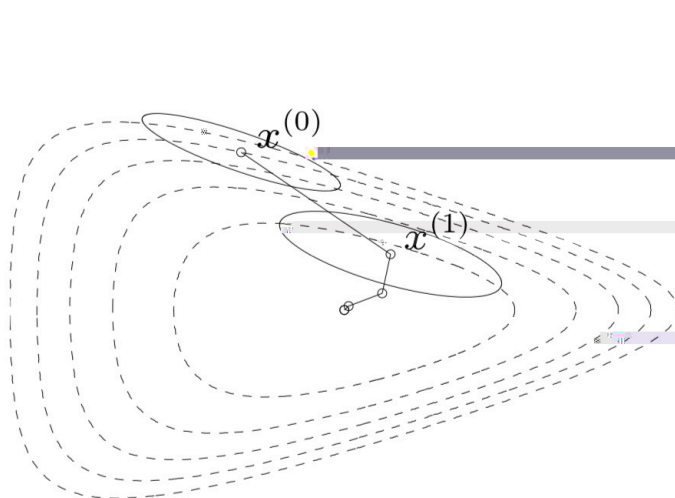


$\square R^2$ 中的子





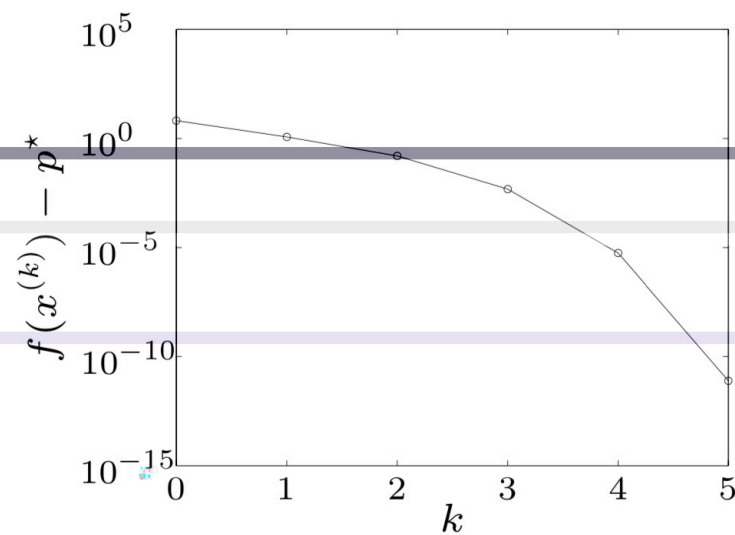
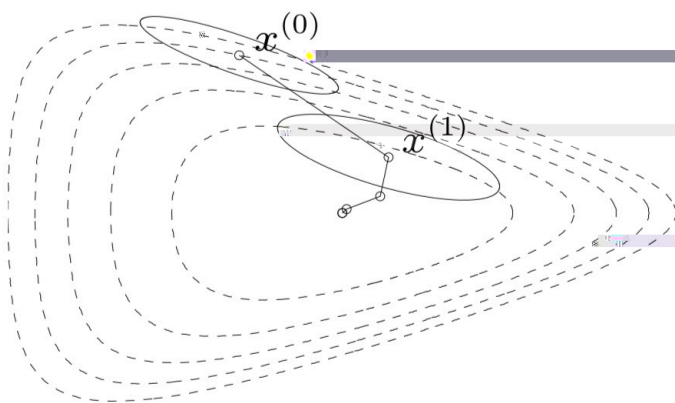
□ R^2 中的子



□ 回 $\alpha = 0.1, \beta = 0.7$



□ R^2 中的子

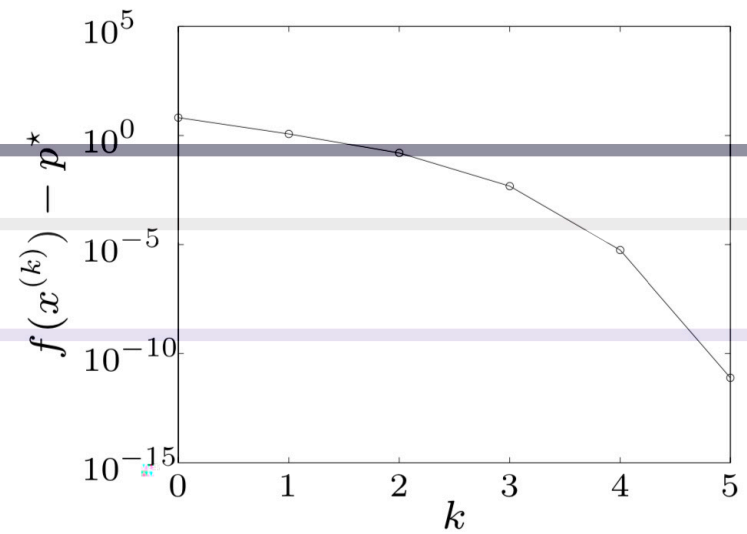
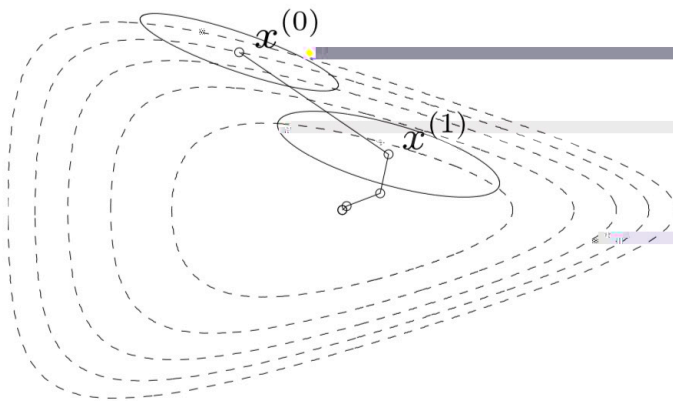


□ 回 $\alpha = 0.1, \beta = 0.7$

□ 用5



□ R^2 中的子



□ 回 $\alpha = 0.1, \beta = 0.7$

□ 用5

□ 次

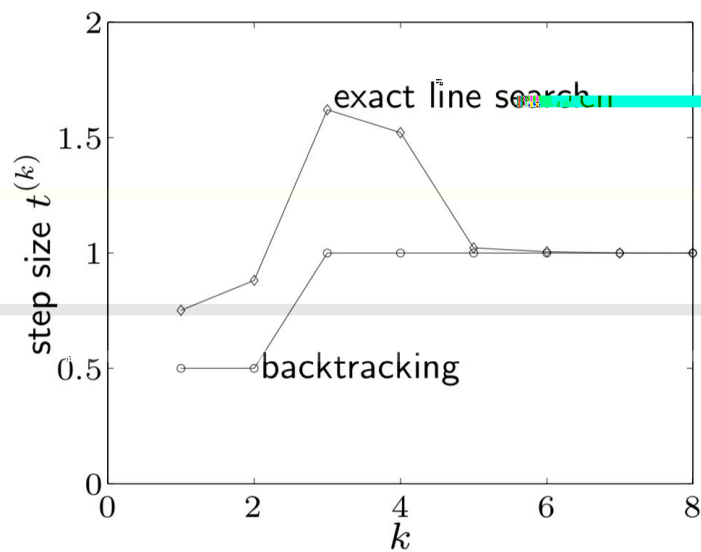
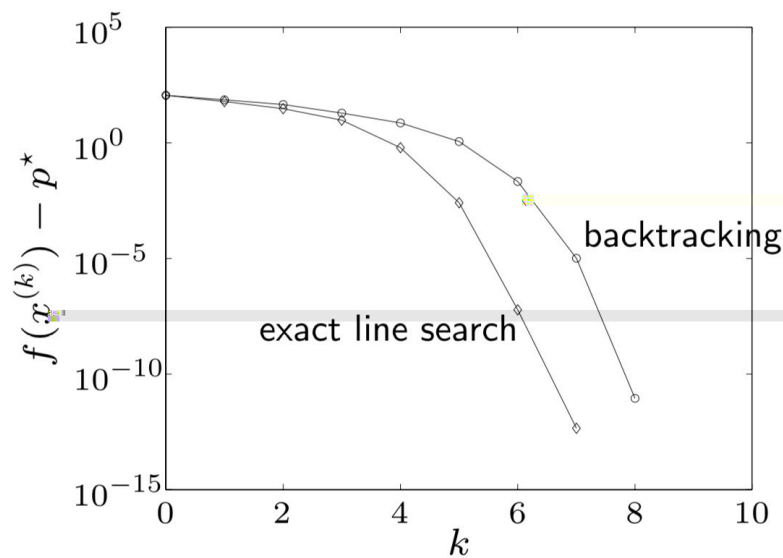




□ R^{100} 中的子

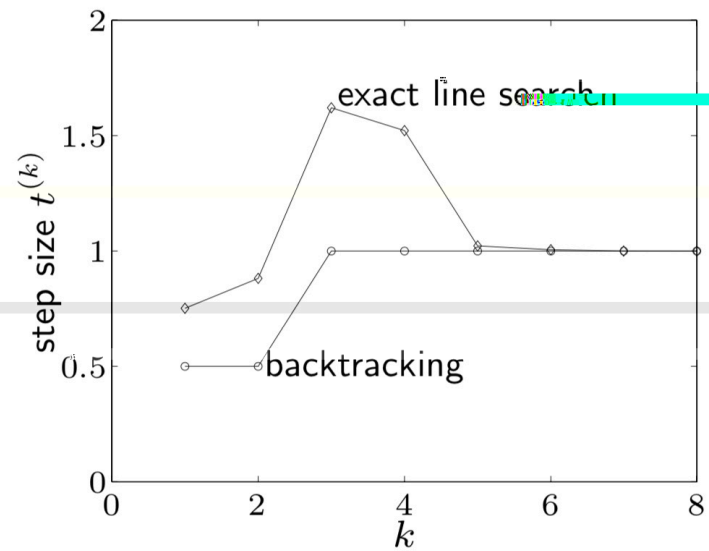
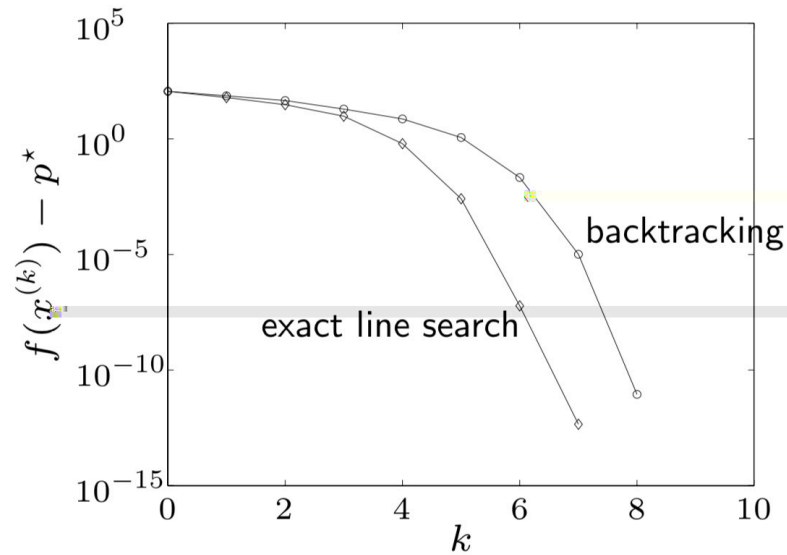


R^{100} 中的子





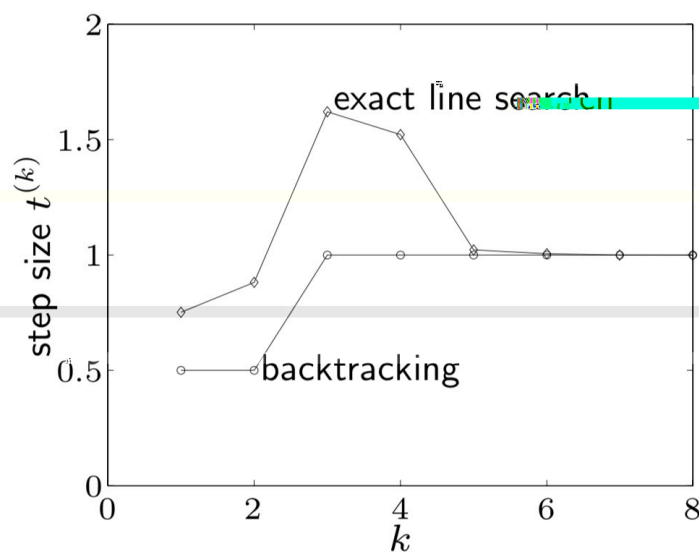
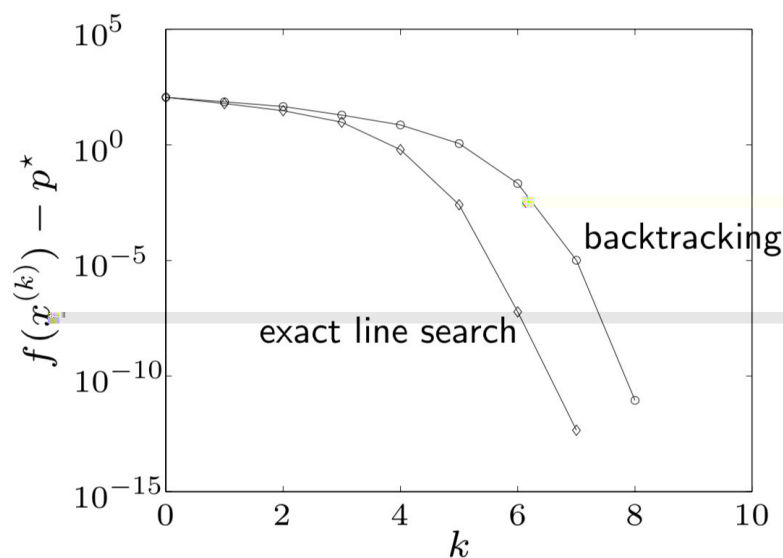
□ R^{100} 中的子



□ 回 $\alpha = 0.01, \beta = 0.5$



□ R^{100} 中的子

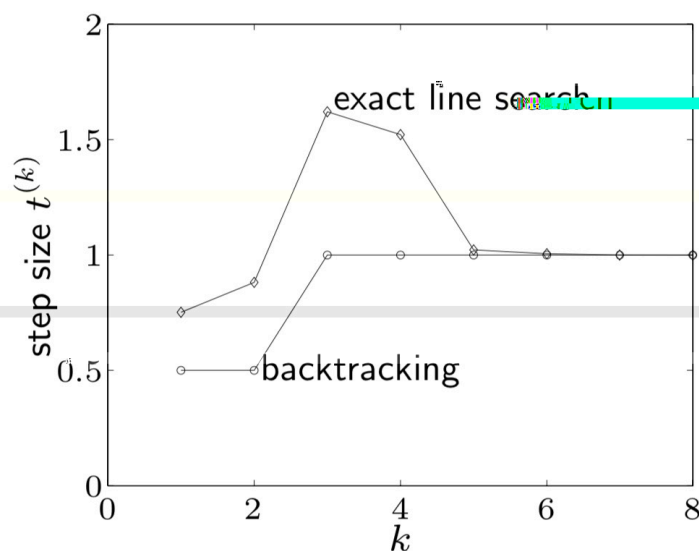
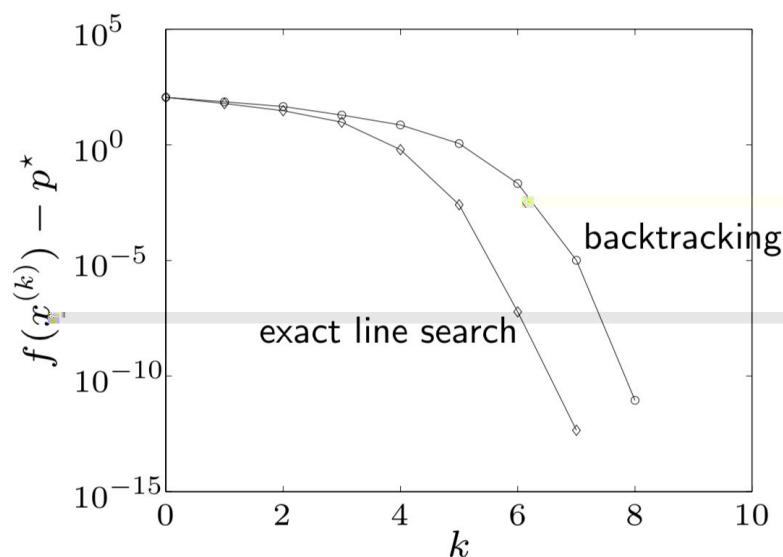


□ 回 $\alpha = 0.01, \beta = 0.5$

□ 回 和 一样快 更为



□ R^{100} 中的子



□ 回 $\alpha = 0.01, \beta = 0.5$

□ 回 和 一样快 更为

□ 的 了 的两个



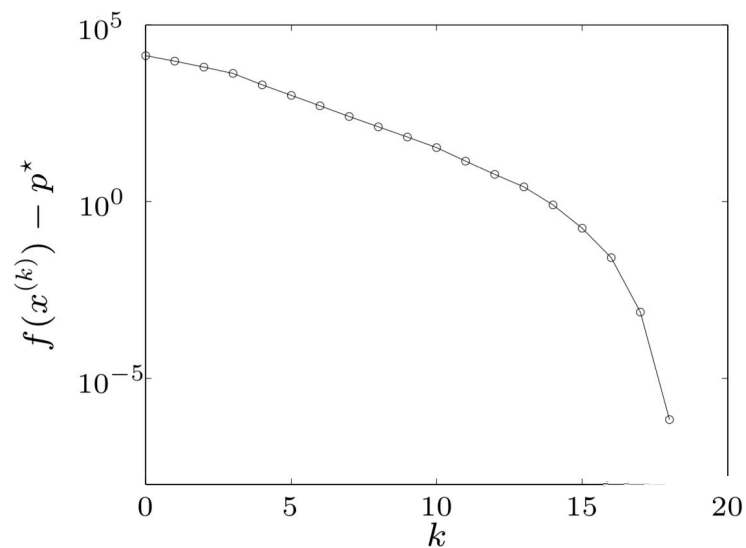


□ R^{10000} 的子 有 的 a_i



□ R^{10000} 的子 有 的 a_i

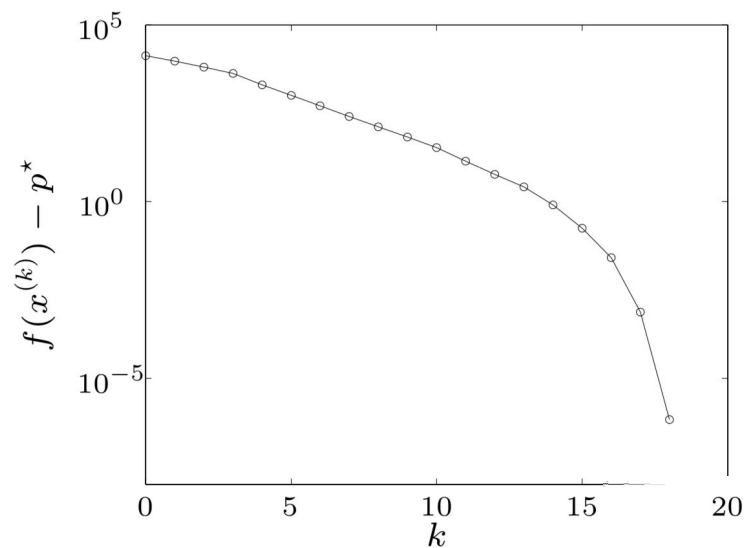
$$f(x) = - \sum_{i=1}^{10000} \log(1 - x_i^2) - \sum_{i=1}^{100000} \log(b_i - a_i^T x)$$





□ R^{10000} 的子 有 的 a_i

$$f(x) = - \sum_{i=1}^{10000} \log(1 - x_i^2) - \sum_{i=1}^{100000} \log(b_i - a_i^T x)$$

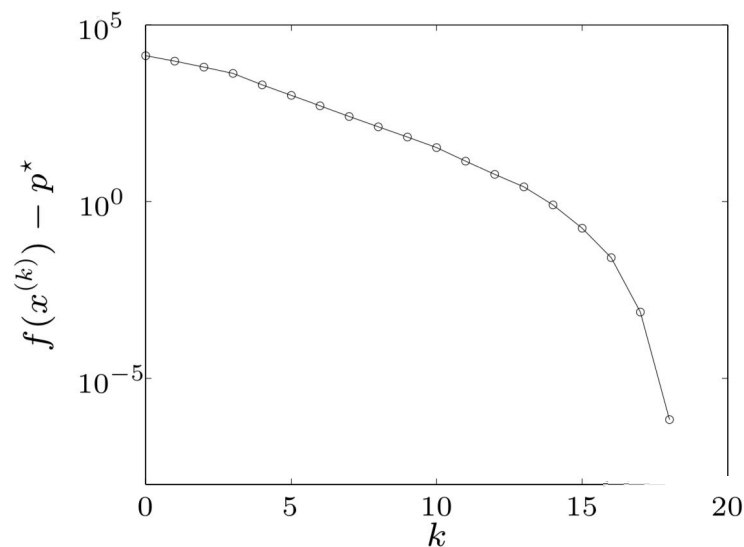


□ 回 $\alpha = 0.01, \beta = 0.5$



□ R^{10000} 的子 有 的 a_i

$$f(x) = - \sum_{i=1}^{10000} \log(1 - x_i^2) - \sum_{i=1}^{100000} \log(b_i - a_i^T x)$$



□ 回 $\alpha = 0.01, \beta = 0.5$

□ 和 小的 子 有 样的 能



自和





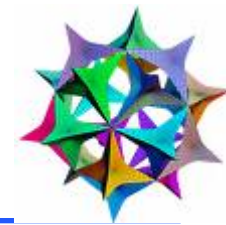
自和



分 的不



自和



分 的不



知 m L



自和



分 的不



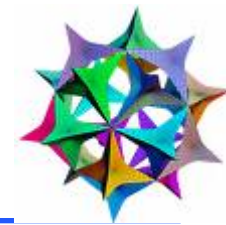
知 m L



不 不



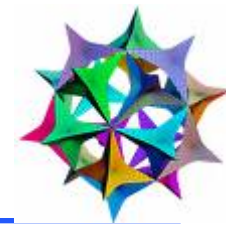
自和



- 分 的不
- 知 m L
- 不 不
- 用自和 的 分 和



自和



- 分 的不
- 知 m L
- 不 不
- 用自和 的 分 和
- 不 知



自和



- 分 的不
- 知 m L
- 不 不
- 用自和 的 分 和
- 不 知
- 给出了 有 不 的



自和



- 分 的不
- 知 m L
- 不 不
- 用自和 的 分 和
- 不 知
- 给出了 有 不 的
- 可用 的 自和



自和



- 分 的不
- 知 m L
- 不 不
- 用自和 的 分 和
- 不 知
- 给出了 有 不 的
- 可用 的 自和
- 可用来分 的多 时 点



自和





自和





自和



$f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和



自和



对 $f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和
有 $x \in \text{dom } f$



自和



对

$f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和

有 $x \in \mathbf{dom} f$

$$|f'''(x)| \leq 2f''(x)^{3/2}$$



自和



对 $f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和
有 $x \in \text{dom } f$
 $|f'''(x)| \leq 2f''(x)^{3/2}$



$f : \mathbf{R}^n \rightarrow \mathbf{R}$ 为自和



自和



对 $f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和
有 $x \in \mathbf{dom} f$
 $|f'''(x)| \leq 2f''(x)^{3/2}$



$f : \mathbf{R}^n \rightarrow \mathbf{R}$ 为自和
对 有 $x \in \mathbf{dom} f, v \in \mathbf{R}^n$
为自和

$$g(t) = f(x + tv)$$



自和





自和



□ R



自和



□ R



和 次



自和



□ R

□

和 次

□

对

$$f(x) = -\log x$$



自和



□ R

□

和 次

□

对

$$f(x) = -\log x$$

□

对

$$f(x) = x \log x - \log x$$



自和



□ R

□

和 次

□

对

$$f(x) = -\log x$$

□

对

$$f(x) = x \log x - \log x$$

□

不



自和



□ $\mathbf{R}$

□

和 次

□

对

$$f(x) = -\log x$$

□

对

$$f(x) = x \log x - \log x$$

□

不

□

$f: \mathbf{R} \rightarrow \mathbf{R}$ 为自和
为自和

$$\tilde{f}(y) = f(ay + b)$$



自和



□ $\mathbf{R}$

□

和 次

□

对

$$f(x) = -\log x$$

□

对

$$f(x) = x \log x - \log x$$

□

不

□

$f : \mathbf{R} \rightarrow \mathbf{R}$ 为自和
为自和

$$\tilde{f}(y) = f(ay + b)$$

$$\tilde{f}'''(y) = a^3 f'''(ay + b), \quad \tilde{f}''(y) = a^2 f''(ay + b)$$



自和的 分





自和的分



□ 在 $\eta \in (0, 1/4], \gamma > 0$



自和的分



□ 在 $\eta \in (0, 1/4]$, $\gamma > 0$

□ $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$



自和的分



- 在 $\eta \in (0, 1/4], \gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$



自和的分



- 在 $\eta \in (0, 1/4]$, $\gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$
- η 和 γ 只回 α, β



自和的分



- 在 $\eta \in (0, 1/4], \gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$
- η 和 γ 只回 α, β
- 次上为



自和的分



- 在 $\eta \in (0, 1/4], \gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$
- η 和 γ 只回 α, β
- 次上为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(1/\epsilon)$$



自和的分



- 在 $\eta \in (0, 1/4], \gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$
- η 和 γ 只回 α, β
- 次上为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(1/\epsilon)$$

- $\alpha = 0.1, \beta = 0.8, \epsilon = 10^{-10}$, 上为



自和的分



- 在 $\eta \in (0, 1/4], \gamma > 0$
- $\lambda(x) > \eta$ 有 $f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$
- $\lambda(x) \leq \eta$ 有 $2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$
- η 和 γ 只回 α, β
- 次上为

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(1/\epsilon)$$

- $\alpha = 0.1, \beta = 0.8, \epsilon = 10^{-10}$, 上为
- $375(f(x^{(0)}) - p^*)$



子





子



□ 150个 生成的



□ 150个 生成的

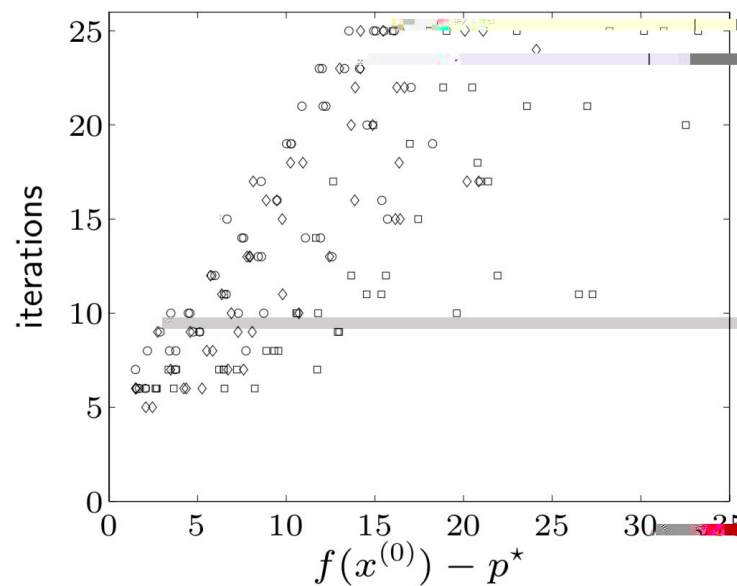
$$\text{minimize } f(x) = -\sum_{i=1}^m \log(b_i - \alpha_i^T x)$$



150个 生成的

$$\text{minimize } f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$$

- : $m = 100, n = 50$
- : $m = 1000, n = 500$
- ◇: $m = 1000, n = 50$





子



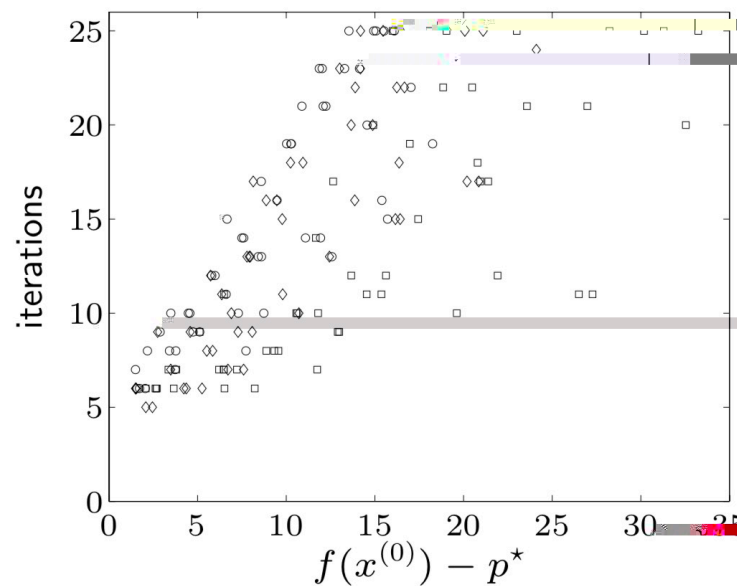
150个 生成的

$$\text{minimize } f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$$

○: $m = 100, n = 50$

□: $m = 1000, n = 500$

◇: $m = 1000, n = 50$



次 大大 $375(f(x^{(0)}) - p^*) + 6$



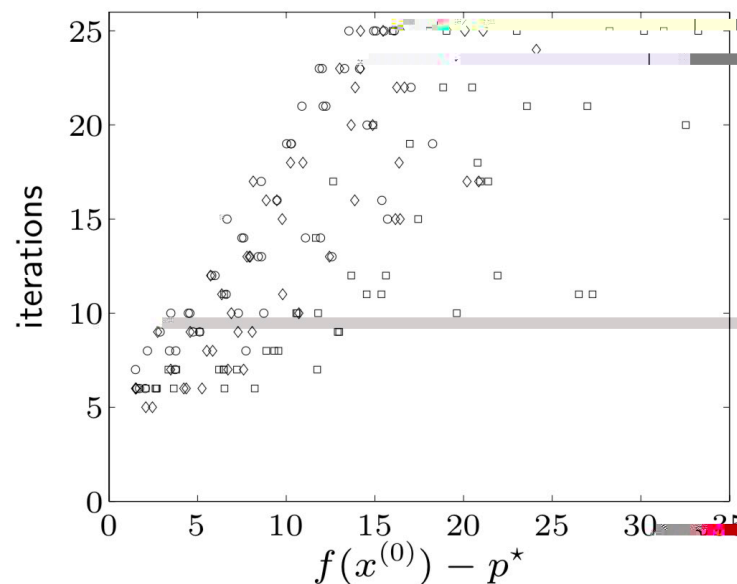
子



150个 生成的

$$\text{minimize } f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$$

○: $m = 100, n = 50$
 □: $m = 1000, n = 500$
 ◇: $m = 1000, n = 50$



次 大大 $375(f(x^{(0)}) - p^*) + 6$

$c(f(x^{(0)}) - p^*) + 6$ 是 次 的一个不 的 c 小



现





现



□ 次 的 要

$$H\Delta x = -g$$



现



□ 次 的 要

$$H\Delta x = -g$$

□ 中 $H = \nabla^2 f(x)$, $g = \nabla f(x)$



现



□ 次 的 要

$$H\Delta x = -g$$

□ 中 $H = \nabla^2 f(x)$, $g = \nabla f(x)$

□ 过 分



现



□ 次 的 要

$$H\Delta x = -g$$

□ 中 $H = \nabla^2 f(x)$, $g = \nabla f(x)$

□ 过 分

$$H = LL^T, \quad \Delta x_{\text{nt}} = -L^{-T}L^{-1}g, \quad \lambda(x) = \|L^{-1}g\|_2$$



现



□ 次 的 要

$$H\Delta x = -g$$

□ 中 $H = \nabla^2 f(x)$, $g = \nabla f(x)$

□ 过 分

$$H = LL^T, \quad \Delta x_{\text{nt}} = -L^{-T}L^{-1}g, \quad \lambda(x) = \|L^{-1}g\|_2$$

□ 对 要 $(1/3)n^3$ 次 点



现



□ 次 的 要

$$H\Delta x = -g$$

□ 中 $H = \nabla^2 f(x)$, $g = \nabla f(x)$

□ 过 分

$$H = LL^T, \quad \Delta x_{\text{nt}} = -L^{-T}L^{-1}g, \quad \lambda(x) = \|L^{-1}g\|_2$$

□ 对 要 $(1/3)n^3$ 次 点

□ H 为 的 $\ll (1/3)n^3$