



最 方

东南大学
机&人 能学

songmf@seu.edu.cn



等



- ☐ 等
- ☐ 去等
- ☐ 有等 的**Newton**方
- ☐ 不可行 点的**Newton**方
- ☐ 现



等





等



$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$



f 为 次 可微



等



$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

□ f 为 次 可微

□ $A \in R^{p \times n}$ $\text{rank} A = p$



等



$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

- f 为 p 次可微
- $A \in R^{p \times n}$ $\text{rank} A = p$
- 最优值 p^* 有界



等



$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

□ f 为 次 可微

□ $A \in R^{p \times n}$ $\text{rank} A = p$

□ 最 p^* 有 在

□ 最 x^* 为最 在 v^*



等



$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

□ f 为 次 可微

□ $A \in R^{p \times n}$ $\text{rank} A = p$

□ 最 p^* 有 在

□ 最 x^* 为最 在 v^*

$$\nabla f(x^*) + A^T \nu^* = 0, \quad Ax^* = b$$



等

次





等

次



$$\begin{aligned} &\text{minimize} && (1/2)x^T P x + q^T x + r \\ &\text{subject to} && Ax = b \end{aligned}$$



等

次



$$\begin{aligned} &\text{minimize} && (1/2)x^T P x + q^T x + r \\ &\text{subject to} && Ax = b \end{aligned}$$

□ 最



等

次



$$\begin{aligned} &\text{minimize} && (1/2)x^T P x + q^T x + r \\ &\text{subject to} && Ax = b \end{aligned}$$

□ 最

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$



等

次



$$\begin{aligned} &\text{minimize} \quad (1/2)x^T P x + q^T x + r \\ &\text{subject to} \quad Ax = b \end{aligned}$$

□ 最

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

□

成为**KKT**



等

次



$$\begin{aligned} &\text{minimize} \quad (1/2)x^T P x + q^T x + r \\ &\text{subject to} \quad Ax = b \end{aligned}$$

□ 最

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

□

成为KKT

□

KKT 为



等

次



$$\begin{aligned} & \text{minimize} \quad (1/2)x^T P x + q^T x + r \\ & \text{subject to} \quad Ax = b \end{aligned}$$

□ 最

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

□ 成为KKT

□ KKT 为

$$Ax = 0, \quad x \neq 0 \quad \implies \quad x^T P x > 0$$



等

次



□ 最

□ 成为**KKT**

□ **KKT** 为



等

次



$$\begin{aligned} & \text{minimize} \quad (1/2)x^T P x + q^T x + r \\ & \text{subject to} \quad Ax = b \end{aligned}$$

□ 最

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

□ 成为KKT

□ KKT 为

$$Ax = 0, \quad x \neq 0 \quad \implies \quad x^T P x > 0$$

□ 的等 $P + A^T A \succ 0$



等





等



$\{x | Ax = b\}$ 的 为

$$\{x \mid Ax = b\} = \{Fz + \hat{x} \mid z \in \mathbf{R}^{n-p}\}$$



等



□ $\{x | Ax = b\}$ 的 为

$$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$

□ $\hat{x}$ 为 意一



等



- $\{x | Ax = b\}$ 的 为
- $$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$
- $\hat{x}$ 为 意一
- $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间



等



- $\{x | Ax = b\}$ 的 为
- $$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$
- $\hat{x}$ 为 意一
- $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间
- $\text{rank } F = n - p \quad AF = 0$



等



□ $\{x | Ax = b\}$ 的 为

$$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$

□ $\hat{x}$ 为 意一

□ $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间

□ $\text{rank } F = n - p \quad AF = 0$

□ 去 后的 minimize $f(Fz + \hat{x})$



等



□ $\{x | Ax = b\}$ 的 为

$$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$

□ $\hat{x}$ 为 意一

□ $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间

$$\square \text{rank } F = n - p \quad AF = 0$$

□ 去 后的 minimize $f(Fz + \hat{x})$

□ 为无 为 $z \in \mathbf{R}^{n-p}$



等



- $\{x \mid Ax = b\}$ 的 为
- $\{x \mid Ax = b\} = \{Fz + \hat{x} \mid z \in \mathbf{R}^{n-p}\}$
- $\hat{x}$ 为 意一
- $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间
- $\text{rank } F = n - p \quad AF = 0$
- 去 后的 $\text{minimize } f(Fz + \hat{x})$
- 为无 为 $z \in \mathbf{R}^{n-p}$
- z^* 可 x^* 和 v^*



等



□ $\{x | Ax = b\}$ 的 为

$$\{x | Ax = b\} = \{Fz + \hat{x} | z \in \mathbf{R}^{n-p}\}$$

□ $\hat{x}$ 为 意一

□ $F \in \mathbf{R}^{n \times (n-p)}$ 的 为 A 的 间

□ $\text{rank } F = n - p \quad AF = 0$

□ 去 后的 $\text{minimize } f(Fz + \hat{x})$

□ 为无 为 $z \in \mathbf{R}^{n-p}$

□ z^* 可 x^* 和 v^*

$$x^* = Fz^* + \hat{x}, \quad v^* = -(AA^T)^{-1}A\nabla f(x^*)$$



的最





的最



$$\begin{array}{ll} \text{minimize} & f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ \text{subject to} & x_1 + x_2 + \cdots + x_n = b \end{array}$$



的最



$$\begin{aligned} &\text{minimize} && f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ &\text{subject to} && x_1 + x_2 + \cdots + x_n = b \end{aligned}$$

□ 去 $x_n = b - x_1 - \cdots - x_{n-1}$



的最



$$\begin{aligned} &\text{minimize} && f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ &\text{subject to} && x_1 + x_2 + \cdots + x_n = b \end{aligned}$$

□ 去 $x_n = b - x_1 - \cdots - x_{n-1}$

$$\hat{x} = be_n, \quad F = \begin{bmatrix} I \\ -\mathbf{1}^T \end{bmatrix} \in \mathbf{R}^{n \times (n-1)}$$



的最



$$\begin{aligned} &\text{minimize} && f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ &\text{subject to} && x_1 + x_2 + \cdots + x_n = b \end{aligned}$$

□ 去 $x_n = b - x_1 - \cdots - x_{n-1}$

$$\hat{x} = be_n, \quad F = \begin{bmatrix} I \\ -\mathbf{1}^T \end{bmatrix} \in \mathbf{R}^{n \times (n-1)}$$

□ 去后的



的最



$$\begin{aligned} &\text{minimize} && f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n) \\ &\text{subject to} && x_1 + x_2 + \cdots + x_n = b \end{aligned}$$

□ 去 $x_n = b - x_1 - \cdots - x_{n-1}$

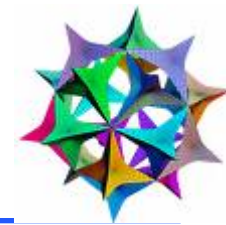
$$\hat{x} = be_n, \quad F = \begin{bmatrix} I \\ -\mathbf{1}^T \end{bmatrix} \in \mathbf{R}^{n \times (n-1)}$$

□ 去后的

$$\text{minimize} \quad f_1(x_1) + \cdots + f_{n-1}(x_{n-1}) + f_n(b - x_1 - \cdots - x_{n-1})$$



Newton方





Newton方



f 在可行点 x 的**Newton**方 Δx_{nt} v 给出



Newton方



f 在可行点 x 的**Newton**方 Δx_{nt} v 给出

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$



Newton方



□ f 在可行点 x 的**Newton**方 Δx_{nt} v 给出

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

□ Δx_{nt} 了 为 v



Newton方



□ f 在可行点 x 的**Newton方** Δx_{nt} v 给出

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

□ Δx_{nt} 了 为 v

$$\begin{array}{ll} \text{minimize} & \hat{f}(x+v) = f(x) + \nabla f(x)^T v + (1/2)v^T \nabla^2 f(x)v \\ \text{subject to} & A(x+v) = b \end{array}$$



Newton方



□ f 在可行点 x 的**Newton方** Δx_{nt} v 给出

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

□ Δx_{nt} 了 为 v

$$\begin{array}{ll} \text{minimize} & \hat{f}(x+v) = f(x) + \nabla f(x)^T v + (1/2)v^T \nabla^2 f(x)v \\ \text{subject to} & A(x+v) = b \end{array}$$

□ Δx_{nt} 等 来 于 最



Newton方



□ f 在可行点 x 的**Newton方** Δx_{nt} v 给出

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

□ Δx_{nt} 了 为 v

$$\begin{array}{ll} \text{minimize} & \hat{f}(x+v) = f(x) + \nabla f(x)^T v + (1/2)v^T \nabla^2 f(x)v \\ \text{subject to} & A(x+v) = b \end{array}$$

□ Δx_{nt} 等 来 于 最

$$\nabla f(x+v) + A^T w \approx \nabla f(x) + \nabla^2 f(x)v + A^T w = 0, \quad A(x+v) = b$$



Newton





Newton



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



Newton



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

□ 给出了 $f(x) - p^*$ 的一个 用了 次



Newton





Newton



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

□ 给出了 $f(x) - p^*$ 的一个 用了 次

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

□ **Newton** 方 的方



Newton



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

□ 给出了 $f(x) - p^*$ 的一个 用了 次

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

□ **Newton** 方 的方

$$\left. \frac{d}{dt} f(x + t \Delta x_{\text{nt}}) \right|_{t=0} = -\lambda(x)^2$$



Newton



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

□ 给出了 $f(x) - p^*$ 的一个 用了 次

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

□ **Newton** 方 的方

$$\left. \frac{d}{dt} f(x + t \Delta x_{\text{nt}}) \right|_{t=0} = -\lambda(x)^2$$

□ 一 的 $\lambda(x) \neq (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$



有等 的Newton方





有等 的Newton方



算法 10.1 等式约束优化问题的 Newton 方法。

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

重复进行

1. 计算 Newton 方向和 Newton 减量 $\Delta x_{\text{nt}}, \lambda(x)$ 。
2. 停止准则。如果 $\lambda^2/2 \leq \epsilon$ 则退出。
3. 直线搜索。通过回溯直线搜索确定步长 t 。
4. 修改。 $x := x + t\Delta x_{\text{nt}}$ 。



有等式的Newton方



算法 10.1 等式约束优化问题的 Newton 方法。

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

重复进行

1. 计算 Newton 方向和 Newton 减量 $\Delta x_{\text{nt}}, \lambda(x)$ 。
2. 停止准则。如果 $\lambda^2/2 \leq \epsilon$ 则退出。
3. 直线搜索。通过回溯直线搜索确定步长 t 。
4. 修改。 $x := x + t\Delta x_{\text{nt}}$ 。

☐ 可行的下方 $x^{(k)}$ 可行 $f(x^{(k+1)}) < f(x^{(k)})$

☐ 不

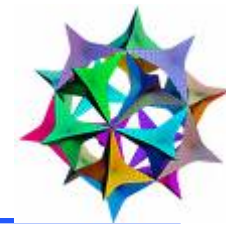


Newton 方 和





Newton方 和



□ 面 的Newton方



Newton方 和



□ 面

的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x} \quad A\hat{x} = b; \text{rank } F = n - p \text{ and } AF = 0$



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x}$ $A\hat{x} = b; \text{rank } F = n - p \text{ and } AF = 0$

□ 对 行Newton方 起 点为 $z^{(0)}$ 行



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x}$ $A\hat{x} = b$; $\text{rank } F = n - p$ and $AF = 0$

□ 对 行Newton方 起 点为 $z^{(0)}$ 行

□ 有等 的Newton方



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x}$ $A\hat{x} = b$; $\text{rank } F = n - p$ and $AF = 0$

□ 对 行Newton方 起 点为 $z^{(0)}$ 行

□ 有等 的Newton方

□ 起 点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时 为



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x}$ $A\hat{x} = b$; $\text{rank } F = n - p$ and $AF = 0$

□ 对 行Newton方 起 点为 $z^{(0)}$ 行

□ 有等 的Newton方

□ 起 点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时 为

$$x^{(k+1)} = Fz^{(k)} + \hat{x}$$



Newton方 和



□ 面 的Newton方

$$\text{minimize } \tilde{f}(z) = f(Fz + \hat{x})$$

□ $z \in \mathbf{R}^{n-p}$

□ $\hat{x} \quad A\hat{x} = b; \text{rank } F = n - p \text{ and } AF = 0$

□ 对 行Newton方 起 点为 $z^{(0)}$ 行

□ 有等 的Newton方

□ 起 点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时 为

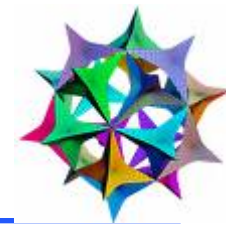
$$x^{(k+1)} = Fz^{(k)} + \hat{x}$$

□ 不 要分 的 分



不可行

点的**Newton**方





不可行

点的Newton方



□ 在不可行点 x 最



不可行

点的Newton方



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$



不可行

点的Newton方



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton**方 的



不可行

点的Newton方



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton**方 的

□ 最 为 $r(y) = 0$ 中



不可行

点的Newton方



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton**方 的

□ 最 为 $r(y) = 0$ 中

$$y = (x, \nu), \quad r(y) = (\nabla f(x) + A^T \nu, Ax - b)$$



不可行

点的Newton方



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton方** 的

□ 最 为 $r(y) = 0$ 中

$$y = (x, \nu), \quad r(y) = (\nabla f(x) + A^T \nu, Ax - b)$$

□ $r(y) = 0$ 可得



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton方** 的

□ 最 为 $r(y) = 0$ 中

$$y = (x, \nu), \quad r(y) = (\nabla f(x) + A^T \nu, Ax - b)$$

□ $r(y) = 0$ 可得 $r(y + \Delta y) \approx r(y) + Dr(y)\Delta y = 0$:



不可行

点的Newton方



□ 在不可行点 x



□ 在不可行点 x 最

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ w \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

□ 对 **Newton**方 的

□ 最 为 $r(y) = 0$ 中

$$y = (x, \nu), \quad r(y) = (\nabla f(x) + A^T \nu, Ax - b)$$

□ $r(y) = 0$ 可得 $r(y + \Delta y) \approx r(y) + Dr(y)\Delta y = 0$:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ \Delta \nu_{\text{nt}} \end{bmatrix} = - \begin{bmatrix} \nabla f(x) + A^T \nu \\ Ax - b \end{bmatrix} \quad w = \nu + \Delta \nu_{\text{nt}}$$



不可行

点Newton方







不可行点Newton方



算法 10.2 不可行初始点 Newton 方法。

给定 初始点 $x \in \text{dom } f$, ν , 误差阈值 $\epsilon > 0$, $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$ 。

重复进行

1. 计算原对偶 Newton 方向 Δx_{nt} , $\Delta \nu_{\text{nt}}$ 。

2. 对 $\|r\|_2$ 进行回溯直线搜索。

$t := 1$ 。

只要 $\|r(x + t\Delta x_{\text{nt}}, \nu + t\Delta \nu_{\text{nt}})\|_2 > (1 - \alpha t)\|r(x, \nu)\|_2$, $t := \beta t$ 。

3. 改进。 $x := x + t\Delta x_{\text{nt}}$, $\nu := \nu + t\Delta \nu_{\text{nt}}$ 。

直到 $Ax = b$ 并且 $\|r(x, \nu)\|_2 \leq \epsilon$ 。

□ 不是下 可能 在 $f(x^{(k+1)}) > f(x^{(k)})$



算法 10.2 不可行初始点 Newton 方法。

给定 初始点 $x \in \text{dom } f$, ν , 误差阈值 $\epsilon > 0$, $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$ 。

重复进行

1. 计算原对偶 Newton 方向 $\Delta x_{\text{nt}}, \Delta \nu_{\text{nt}}$ 。

2. 对 $\|r\|_2$ 进行回溯直线搜索。

$t := 1$ 。

只要 $\|r(x + t\Delta x_{\text{nt}}, \nu + t\Delta \nu_{\text{nt}})\|_2 > (1 - \alpha t)\|r(x, \nu)\|_2$, $t := \beta t$ 。

3. 改进。 $x := x + t\Delta x_{\text{nt}}, \nu := \nu + t\Delta \nu_{\text{nt}}$ 。

直到 $Ax = b$ 并且 $\|r(x, \nu)\|_2 \leq \epsilon$ 。

□ 不是下 可能 在 $f(x^{(k+1)}) > f(x^{(k)})$

□ $\|r(y)\|_2$ 的有 在方 上 $\Delta y = (\Delta x_{\text{nt}}, \Delta \nu_{\text{nt}})$ 为

$$\left. \frac{d}{dt} \|r(y + t\Delta y)\|_2 \right|_{t=0} = -\|r(y)\|_2$$



KKT





KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$



方



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分

□ 去 H



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分

□ 去 H

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分

□ 去 H

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$

□ 去 H



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分

□ 去 H

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$

□ 去 H

$$\begin{bmatrix} H + A^TQA & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g + A^TQh \\ h \end{bmatrix}$$



KKT



$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ 方

□ LDL^T 分

□ 去 H

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$

□ 去 H

$$\begin{bmatrix} H + A^TQA & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g + A^TQh \\ h \end{bmatrix}$$

□ 中 $Q \succeq 0 \quad H + A^TQA \succ 0$



等的中心





等的中心



$$\text{minimize } -\sum_{i=1}^n \log x_i \text{ subject to } Ax = b$$



等 的 中心



minimize $-\sum_{i=1}^n \log x_i$ subject to $Ax = b$



对

maximize $-b^T \nu + \sum_{i=1}^n \log(A^T \nu)_i + n$



等的中心



- minimize $-\sum_{i=1}^n \log x_i$ subject to $Ax = b$
- 对 maximize $-b^T \nu + \sum_{i=1}^n \log(A^T \nu)_i + n$
- 用三种方 $A \in \mathbf{R}^{100 \times 500}$ 中的 子 三个不
起 点



等 的 中心



- minimize $-\sum_{i=1}^n \log x_i$ subject to $Ax = b$
- 对 maximize $-b^T \nu + \sum_{i=1}^n \log(A^T \nu)_i + n$
- 用三种方 $A \in \mathbf{R}^{100 \times 500}$ 中的 子 三个不
起 点
- 有等 的 **Newton** 方 $x^{(0)} \succ 0, Ax^{(0)} = b$



等的中心





他两个方





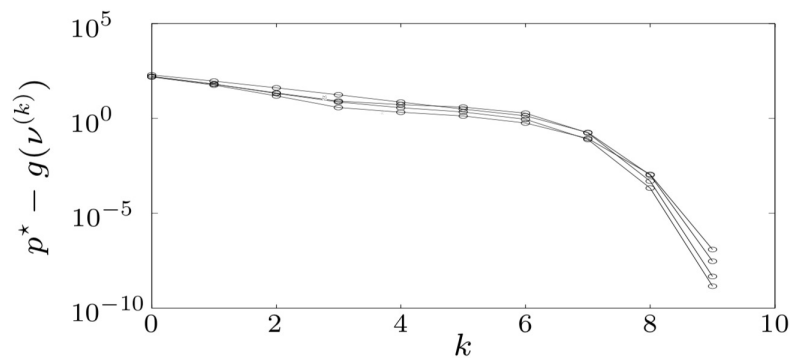


他两个方



□ **Newton方** 用于对

$$A^T \nu^{(0)} \succ 0$$



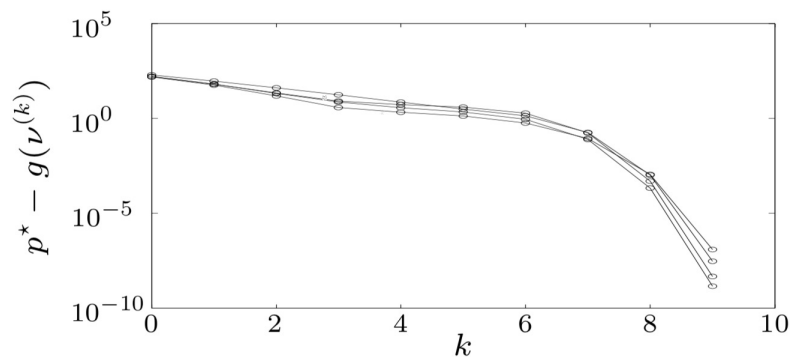


他两个方



□ **Newton方** 用于对

$$A^T \nu^{(0)} \succ 0$$



□ 不可行 点的**Newton方**

$$x^{(0)} \succ 0$$

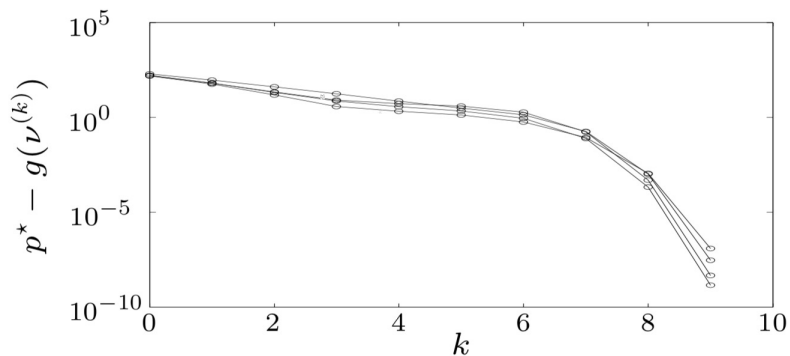


他两个方



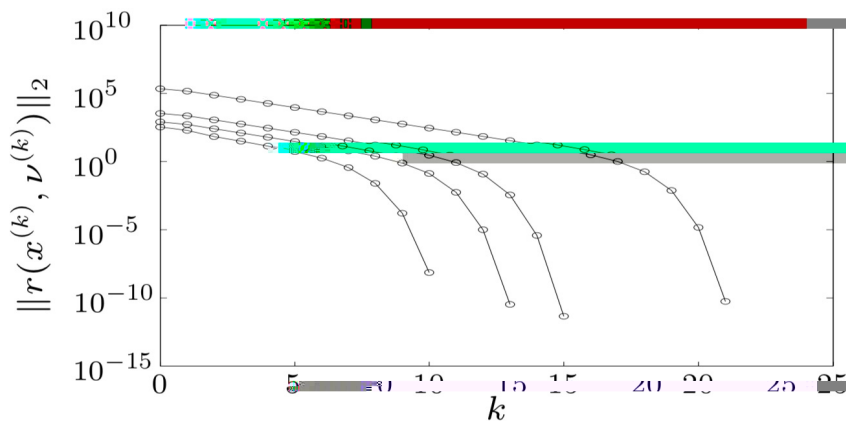
□ **Newton方** 用于对

$$A^T \nu^{(0)} \succ 0$$



□ **不可行** 点的**Newton方**

$$x^{(0)} \succ 0$$





三个方 每次 的





三个方 每次 的



□ 用分 **KKT**



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(w)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(w)^{-1} \\ 0 \end{bmatrix}$$



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ \lambda \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} h \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = h$



三个方 每次 的



□ 用分 **KKT**

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = b$

□ **Newton** $A \text{diag}(A^T \nu)^{-2} A^T \Delta \nu = -b + A \text{diag}(A^T \nu)^{-1} \mathbf{1}$



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = b$

□ Newton $A \text{diag}(A^T \nu)^{-2} A^T \Delta \nu = -b + A \text{diag}(A^T \nu)^{-1} \mathbf{1}$

□ 用分 KKT



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = b$

□ Newton $A \text{diag}(A^T \nu)^{-2} A^T \Delta \nu = -b + A \text{diag}(A^T \nu)^{-1} \mathbf{1}$

□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} - A^T \nu \\ b - Ax \end{bmatrix}$$



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = b$

□ Newton $A \text{diag}(A^T \nu)^{-2} A^T \Delta \nu = -b + A \text{diag}(A^T \nu)^{-1} \mathbf{1}$

□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} - A^T \nu \\ b - Ax \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = 2Ax - b$



三个方 每次 的



□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} \\ 0 \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = h$

□ Newton $A \text{diag}(A^T \nu)^{-2} A^T \Delta \nu = -b + A \text{diag}(A^T \nu)^{-1} \mathbf{1}$

□ 用分 KKT

$$\begin{bmatrix} \text{diag}(x)^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = \begin{bmatrix} \text{diag}(x)^{-1} \mathbf{1} - A^T \nu \\ b - Ax \end{bmatrix}$$

□ 为 $A \text{diag}(x)^2 A^T w = 2Ax - b$

□ 有 对 D 的 $ADA^T w = h$



最





最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$



最



$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b\end{array}$$

□ 有 n 和 $p+1$ 个 点



最



$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b\end{array}$$

□ 有 n 和 $p+1$ 个 点

□ x_i 过 i 的 ϕ_i : i 的 成本 $\phi_i''(x) > 0$



最



□ 有 n 和 $p+1$ 个 点

□ x_i 过 i 的 i 的 成本



最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$

□ 有 n 和 $p+1$ 个 点

□ x_i 过 i 的 ϕ_i : i 的 成本 $\phi_i''(x) > 0$

□ 点 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

$$\tilde{A}_{ij} = \begin{cases} 1 & \text{arc } j \text{ leaves node } i \\ -1 & \text{arc } j \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$



最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$

□ 有 n 个点和 $p+1$ 个约束

□ x_i 过 i 的弧 ϕ_i : i 的成本 $\phi_i''(x) > 0$

□ 点 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

$$\tilde{A}_{ij} = \begin{cases} 1 & \text{arc } j \text{ leaves node } i \\ -1 & \text{arc } j \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$

□ 点 $A \in \mathbf{R}^{p \times n}$ 去 最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$



最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$

□ 有 n 个点和 $p+1$ 个弧

□ x_i 是节点 i 的成本 $\phi_i: \phi_i''(x) > 0$

□ 弧 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

$$\tilde{A}_{ij} = \begin{cases} 1 & \text{arc } j \text{ leaves node } i \\ -1 & \text{arc } j \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$

□ 弧 $A \in \mathbf{R}^{p \times n}$ 去最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

□ 的 $b \in \mathbf{R}^p$



最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$

□ 有 n 个点和 $p+1$ 个

□ x_i 过 i 的 ϕ_i : i 的成本 $\phi_i''(x) > 0$

□ 点 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

$$\tilde{A}_{ij} = \begin{cases} 1 & \text{arc } j \text{ leaves node } i \\ -1 & \text{arc } j \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$

□ 点 $A \in \mathbf{R}^{p \times n}$ 去 最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

□ 的 $b \in \mathbf{R}^p$

□ 是 的 $\text{rank } A = p$



最



$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i) \\ \text{subject to} & Ax = b \end{array}$$

□ 有 n 个点和 $p+1$ 个

□ x_i 过 i 的 ϕ_i : i 的成本 $\phi_i''(x) > 0$

□ 点 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

$$\tilde{A}_{ij} = \begin{cases} 1 & \text{arc } j \text{ leaves node } i \\ -1 & \text{arc } j \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$

□ 点 $A \in \mathbf{R}^{p \times n}$ 去 最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1) \times n}$

□ 的 $b \in \mathbf{R}^p$

□ 是 的 $\text{rank } A = p$



最





最



□ KKT

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$



最



□ KKT
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ $H = \text{diag}(\phi_1''(x_1), \dots, \phi_n''(x_n))$ 对



最



□ KKT
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ $H = \text{diag}(\phi_1''(x_1), \dots, \phi_n''(x_n))$ 对

□ 过



最



□ **KKT**
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ $H = \mathbf{diag}(\phi_1''(x_1), \dots, \phi_n''(x_n))$ 对

□ 过

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$



最



□ KKT
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ $H = \text{diag}(\phi_1''(x_1), \dots, \phi_n''(x_n))$ 对

□ 过

$$AH^{-1}A^Tw = h - AH^{-1}g, \quad Hv = -(g + A^Tw)$$

□ 的 过 的



最



□ KKT
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$$

□ $H = \text{diag}(\phi_1''(x_1), \dots, \phi_n''(x_n))$ 对

□ 过

$$AH^{-1}A^T w = h - AH^{-1}g, \quad Hv = -(g + A^T w)$$

□ 的 过 的

$$(AH^{-1}A^T)_{ij} \neq 0 \iff (AA^T)_{ij} \neq 0$$

$\iff$ 点*i*和*j* 在



不等的中心







不等 的 中心



$$\begin{aligned} & \text{minimize} && -\log \det X \\ & \text{subject to} && \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{aligned}$$



$$X \in \mathbf{S}^n$$



不等 的 中心



$$\begin{aligned} &\text{minimize} && -\log \det X \\ &\text{subject to} && \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{aligned}$$



$$X \in \mathbf{S}^n$$





不等 的 中心



$$\begin{array}{ll} \text{minimize} & -\log \det X \\ \text{subject to} & \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{array}$$



$$X \in \mathbf{S}^n$$



$$X^* \succ 0, \quad -(X^*)^{-1} + \sum_{j=1}^p \nu_j^* A_j = 0, \quad \text{tr}(A_i X^*) = b_i, \quad i = 1, \dots, p$$



不等 的 中心



$$\begin{aligned} & \text{minimize} && -\log \det X \\ & \text{subject to} && \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{aligned}$$



$$X \in \mathbf{S}^n$$



$$X^* \succ 0, \quad -(X^*)^{-1} + \sum_{j=1}^p \nu_j^* A_j = 0, \quad \text{tr}(A_i X^*) = b_i, \quad i = 1, \dots, p$$



在可行点 X 的**Newton**等



不等 的 中心



$$\begin{aligned} & \text{minimize} && -\log \det X \\ & \text{subject to} && \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{aligned}$$



$$X \in \mathbf{S}^n$$



$$X^* \succ 0, \quad -(X^*)^{-1} + \sum_{j=1}^p \nu_j^* A_j = 0, \quad \text{tr}(A_i X^*) = b_i, \quad i = 1, \dots, p$$



在可行点 X 的Newton等

$$X^{-1} \Delta X X^{-1} + \sum_{j=1}^p w_j A_j = X^{-1}, \quad \text{tr}(A_i \Delta X) = 0, \quad i = 1, \dots, p$$



不等 的 中心



$$\begin{aligned} & \text{minimize} && -\log \det X \\ & \text{subject to} && \text{tr}(A_i X) = b_i, \quad i = 1, \dots, p \end{aligned}$$



$$X \in \mathbf{S}^n$$



$$X^* \succ 0, \quad -(X^*)^{-1} + \sum_{j=1}^p \nu_j^* A_j = 0, \quad \text{tr}(A_i X^*) = b_i, \quad i = 1, \dots, p$$



在可行点 X 的**Newton**等

$$X^{-1} \Delta X X^{-1} + \sum_{j=1}^p w_j A_j = X^{-1}, \quad \text{tr}(A_i \Delta X) = 0, \quad i = 1, \dots, p$$



$$(X + \Delta X)^{-1} \approx X^{-1} - X^{-1} \Delta X X^{-1}$$

得到 $n(n+1)/2 + p$ 个 $\Delta X, w$



分





分



□ 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$



分



- 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$
- ΔX 第 个方



分



□ 第一个方程中去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 i 个方程

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$



分



□ 第一个方程中去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 i 个方程

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$

□ 为一个 $p \times p$ 的方程组



分



□ 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 个方

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$

□ 为一个 的 方 为 $w \in \mathbf{R}^p$

□ 用 **Cholesky** 分 $X=LL^T$ 的 点 次



分



□ 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 个方

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$

□ 为一个 的 方 为 $w \in \mathbf{R}^p$

□ 用 **Cholesky** 分 $X = LL^T$ 的 点 次

□ p 个 $L^T A_j L$: $(3/2)pn^3$



分



□ 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 个方

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$

□ 为一个 的 方 为 $w \in \mathbf{R}^p$

□ 用 **Cholesky** 分 $X = LL^T$ 的 点 次

□ p 个 $L^T A_j L$: $(3/2)pn^3$

□ $p(p+1)/2$ 个 $\text{tr}((L^T A_i L)(L^T A_j L))$: $(1/2)p^2 n^2$



分



□ 第一个方 中 去 ΔX $\Delta X = X - \sum_{j=1}^p w_j X A_j X$

□ ΔX 第 个方

$$\sum_{j=1}^p \text{tr}(A_i X A_j X) w_j = b_i, \quad i = 1, \dots, p$$

□ 为一个 的 方 为 $w \in \mathbf{R}^p$

□ 用 **Cholesky** 分 $X = LL^T$ 的 点 次

□ p 个 $L^T A_j L$: $(3/2)pn^3$

□ $p(p+1)/2$ 个 $\text{tr}((L^T A_i L)(L^T A_j L))$: $(1/2)p^2 n^2$

□ 过 **Cholesky** 分 $(1/3)p^3$