最优化方法

东南大学
计算机&人工智能学院
宋沫飞
songmf@seu.edu.cn

等式约束优化

- 等式约束优化问题
- 消去等式约束
- 带有等式约束的**Newton**方法
- 不可行初始点的**Newton**方法
- 实现

等式约束优化问题

等式约束优化问题

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

等式约束优化问题

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

- 函数f为凸函数，二次连续可微

等式约束优化问题

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

- 函数f为凸函数，二次连续可微
- $A \in R^{p\times n}$，秩$\mathrm{rank} A = p$

等式约束优化问题

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

- 函数f为凸函数，二次连续可微
- $A \in R^{p\times n}$，秩$\mathrm{rank}A = p$
- 假设最优值p^*有限且存在

等式约束优化问题

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

- 函数f为凸函数，二次连续可微
- $A \in R^{p\times n}$，秩$\mathrm{rank}A = p$
- 假设最优值p^*有限且存在
- 最优条件：x^*为最优解，当且仅当存在v^*满足

等式约束优化问题

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

- 函数f为凸函数，二次连续可微
- $A \in R^{p\times n}$，秩$\mathrm{rank}A = p$
- 假设最优值p^*有限且存在
- 最优条件：x^*为最优解，当且仅当存在v^*满足

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

等式约束凸二次规划

等式约束凸二次规划

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

等式约束凸二次规划

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

- 最优化条件：

等式约束凸二次规划

$$\begin{array}{ll} \text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Ax = b \end{array}$$

- 最优化条件：

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

等式约束凸二次规划

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

- 最优化条件：

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

- 系数矩阵成为**KKT**矩阵

等式约束凸二次规划

$$\begin{array}{ll}\text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Ax = b\end{array}$$

- 最优化条件：

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

- 系数矩阵成为**KKT**矩阵
- **KKT**矩阵为非奇异，当且仅当

等式约束凸二次规划

$$\begin{array}{ll}\text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Ax = b\end{array}$$

- 最优化条件：

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

- 系数矩阵成为**KKT**矩阵
- **KKT**矩阵为非奇异，当且仅当

$$Ax = 0, \quad x \neq 0 \qquad \Longrightarrow \qquad x^TPx > 0$$

等式约束凸二次规划

$$\begin{array}{ll}\text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Ax = b\end{array}$$

- 最优化条件：

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

- 系数矩阵成为**KKT**矩阵
- **KKT**矩阵为非奇异，当且仅当

$$Ax = 0, \quad x \neq 0 \qquad \Longrightarrow \qquad x^TPx > 0$$

- 非奇异的等价条件：

等式约束凸二次规划

$$\begin{array}{ll}\text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Ax = b\end{array}$$

- 最优化条件：

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

- 系数矩阵成为**KKT**矩阵
- **KKT**矩阵为非奇异，当且仅当

$$Ax = 0, \quad x \neq 0 \qquad \Longrightarrow \qquad x^TPx > 0$$

- 非奇异的等价条件：$P + A^TA \succ 0$

消除等式约束

消除等式约束

- 将$\{x \mid Ax = b\}$的求解表示为

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间
  - $\operatorname{rank} F = n - p$且$AF = 0$

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间
  - $\operatorname{rank} F = n - p$且$AF = 0$
- 消去约束后的问题 minimize $\quad f(Fz + \hat{x})$

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间
  - $\operatorname{rank} F = n - p$且$AF = 0$
- 消去约束后的问题 minimize $\quad f(Fz + \hat{x})$
- 为无约束问题，其优化变量为 $z \in \mathbf{R}^{n-p}$

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间
  - $\operatorname{rank} F = n - p$且$AF = 0$
- 消去约束后的问题 minimize $\quad f(Fz + \hat{x})$
- 为无约束问题，其优化变量为 $z \in \mathbf{R}^{n-p}$
- 根据解z^*，可获取x^*和 v^*

消除等式约束

- 将$\{x|Ax=b\}$的求解表示为

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

- $\hat{x}$为任意一特定解
- $F \in R^{n\times(n-p)}$的值域为矩阵A的零空间
  - $\operatorname{rank} F = n - p$且$AF = 0$
- 消去约束后的问题 minimize $\quad f(Fz + \hat{x})$
- 为无约束问题，其优化变量为 $z \in \mathbf{R}^{n-p}$
- 根据解z^*，可获取x^*和 ν^*

$$x^\star = Fz^\star + \hat{x}, \qquad \nu^\star = -(AA^T)^{-1}A\nabla f(x^\star)$$

例：受资源约束的最优配置

例：受资源约束的最优配置

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

例：受资源约束的最优配置

minimize $\quad f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$

subject to $\quad x_1 + x_2 + \cdots + x_n = b$

- 消去 $x_n = b - x_1 - \cdots - x_{n-1}$ ，指选择

例：受资源约束的最优配置

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

- 消去 $x_n = b - x_1 - \cdots - x_{n-1}$ ，指选择

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

例：受资源约束的最优配置

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

- 消去 $x_n = b - x_1 - \cdots - x_{n-1}$，指选择

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

- 消去后的问题

例：受资源约束的最优配置

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

- 消去 $x_n = b - x_1 - \cdots - x_{n-1}$，指选择

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

- 消去后的问题

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

Newton方向

Newton方向

- 函数f在可行点x的**Newton**方向 $\Delta x_{\rm nt}$ 由解v给出

Newton方向

- 函数f在可行点x的**Newton**方向 Δx_{nt} 由解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方向

- 函数f在可行点x的**Newton**方向 Δx_{nt} 由解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}解决了二阶近似问题（优化变量为v）

Newton方向

- 函数f在可行点x的**Newton**方向 Δx_{nt} 由解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}解决了二阶近似问题（优化变量为v）

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

Newton方向

- 函数f在可行点x的Newton方向 Δx_{nt} 由解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}解决了二阶近似问题（优化变量为v）

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

- Δx_{nt} 等式来源于线性化最优条件

Newton方向

- 函数f在可行点x的**Newton**方向 Δx_{nt} 由解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}解决了二阶近似问题（优化变量为v）

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

- Δx_{nt} 等式来源于线性化最优条件

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

Newton减量

Newton减量

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

Newton减量

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

- 给出了$f(x) - p^\star$的一个估计，使用了二次近似：

Newton减量

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

- 给出了$f(x) - p^\star$的一个估计，使用了二次近似：

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

Newton减量

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

- 给出了$f(x) - p^\star$的一个估计，使用了二次近似：

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

- **Newton**方向的方向导数：

Newton减量

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

- 给出了$f(x) - p^\star$的一个估计，使用了二次近似：

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

- **Newton**方向的方向导数：

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

Newton减量

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

- 给出了$f(x) - p^\star$的一个估计，使用了二次近似：

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

- **Newton**方向的方向导数：

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

- 一般的，$\lambda(x) \neq \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)\right)^{1/2}$

带有等式约束的Newton方法

带有等式约束的Newton方法

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 $x \in \mathbf{dom}\, f$，误差阈值 $\epsilon > 0$。

重复进行

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

带有等式约束的Newton方法

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 $x \in \mathbf{dom}\, f$，误差阈值 $\epsilon > 0$。

重复进行

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

- 可行的下降方法：$x^{(k)}$ 可行且 $f(x^{(k+1)}) < f(x^{(k)})$
- 仿射不变性

Newton方法和消除法

Newton方法和消除法

- 面向简化问题的**Newton**方法

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; $\mathbf{rank}\, F = n - p$ and $AF = 0$

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; $\mathbf{rank}\, F = n - p$ and $AF = 0$
- 对函数执行**Newton**方法，起始点为 $z^{(0)}$ 进行迭代

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; $\mathbf{rank}\, F = n - p$ and $AF = 0$
- 对函数执行**Newton**方法，起始点为 $z^{(0)}$ 进行迭代
- 带有等式约束的**Newton**方法

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; $\mathbf{rank}\, F = n - p$ and $AF = 0$
- 对函数执行**Newton**方法，起始点为 $z^{(0)}$ 进行迭代
- 带有等式约束的**Newton**方法
- 当起始点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时，迭代为

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; $\mathbf{rank}\, F = n - p$ and $AF = 0$
- 对函数执行**Newton**方法，起始点为 $z^{(0)}$ 进行迭代
- 带有等式约束的**Newton**方法
- 当起始点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时，迭代为

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

Newton方法和消除法

- 面向简化问题的**Newton**方法

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

- 优化变量 $z \in \mathbf{R}^{n-p}$
- $\hat{x}$ 满足 $A\hat{x} = b$; **rank** $F = n - p$ and $AF = 0$
- 对函数执行**Newton**方法，起始点为 $z^{(0)}$ 进行迭代
- 带有等式约束的**Newton**方法
- 当起始点为 $x^{(0)} = Fz^{(0)} + \hat{x}$ 时，迭代为

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

- 因此，不需要分离的收敛分析

不可行初始点的**Newton**方法

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

不可行初始点的Newton方法

- 在不可行点*x*线性化最优条件

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

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y) = 0$ ，其中

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y) = 0$，其中

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

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y) = 0$，其中

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

- 线性化 $r(y) = 0$ 可得

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y) = 0$，其中

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

- 线性化 $r(y) = 0$ 可得 $r(y + \Delta y) \approx r(y) + Dr(y)\Delta y = 0$:

不可行初始点的Newton方法

- 在不可行点*x*线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y) = 0$，其中

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

- 线性化 $r(y) = 0$ 可得 $r(y + \Delta y) \approx r(y) + Dr(y)\Delta y = 0$:

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

不可行初始点的Newton方法

- 在不可行点x线性化最优条件

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

- 原对偶**Newton**方向的解释
- 最优条件为 $r(y)=0$，其中

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

- 线性化 $r(y)=0$ 可得 $r(y+\Delta y)\approx r(y)+Dr(y)\Delta y=0$:

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

不可行初始点Newton方法

不可行初始点Newton方法

算法 10.2　不可行初始点 Newton 方法。

给定 初始点 $x \in \mathbf{dom}\, f$，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** $\Delta x_{\rm nt}$，$\Delta \nu_{\rm nt}$。
2. **对** $\|r\|_2$ **进行回溯直线搜索。**

   $t := 1$。

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

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

不可行初始点Newton方法

算法 10.2　不可行初始点 Newton 方法。

给定 初始点 $x \in \mathbf{dom}\, f$，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}，$\Delta \nu_{\mathrm{nt}}$。
2. 对 $\|r\|_2$ 进行回溯直线搜索。

   $t := 1$。

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

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

- 不是下降法：可能存在 $f(x^{(k+1)}) > f(x^{(k)})$

不可行初始点Newton方法

算法 10.2　不可行初始点 Newton 方法。

给定 初始点 $x \in \mathbf{dom}\, f$，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** $\Delta x_{\rm nt}$，$\Delta \nu_{\rm nt}$。
2. **对 $\|r\|_2$ 进行回溯直线搜索。**

   $t := 1$。

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

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

- 不是下降法：可能存在 $f(x^{(k+1)}) > f(x^{(k)})$
- $\|r(y)\|_2$的有向导数在方向上 $\Delta y = (\Delta x_{\rm nt}, \Delta \nu_{\rm nt})$ 为

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

求解KKT系统

求解KKT系统

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

求解KKT系统

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

- 求解方法

求解KKT系统

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

- 求解方法
- **LDL^T**因式分解

求解KKT系统

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

- 求解方法
- **LDL$^{\mathbf{T}}$**因式分解
- 消去（若H非奇异）：

求解KKT系统

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

- 求解方法
- **LDL^T**因式分解
- 消去（若H非奇异）：

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

求解KKT系统

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

- 求解方法
- **LDL^T**因式分解
- 消去（若H非奇异）：

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

- 消去奇异矩阵H:

求解KKT系统

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

- 求解方法
- **LDL**$^{\mathbf{T}}$因式分解
- 消去（若H非奇异）：

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

- 消去奇异矩阵H:

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

求解KKT系统

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

- 求解方法
- **LDL^T**因式分解
- 消去（若H非奇异）：

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

- 消去奇异矩阵H:

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

- 其中，$Q \succeq 0$ ，满足 $H + A^TQA \succ 0$

等式约束的解析中心

等式约束的解析中心

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

等式约束的解析中心

- 原问题： minimize $-\sum_{i=1}^{n}\log x_i$ subject to $Ax=b$
- 对偶问题: maximize $-b^T\nu+\sum_{i=1}^{n}\log(A^T\nu)_i+n$

等式约束的解析中心

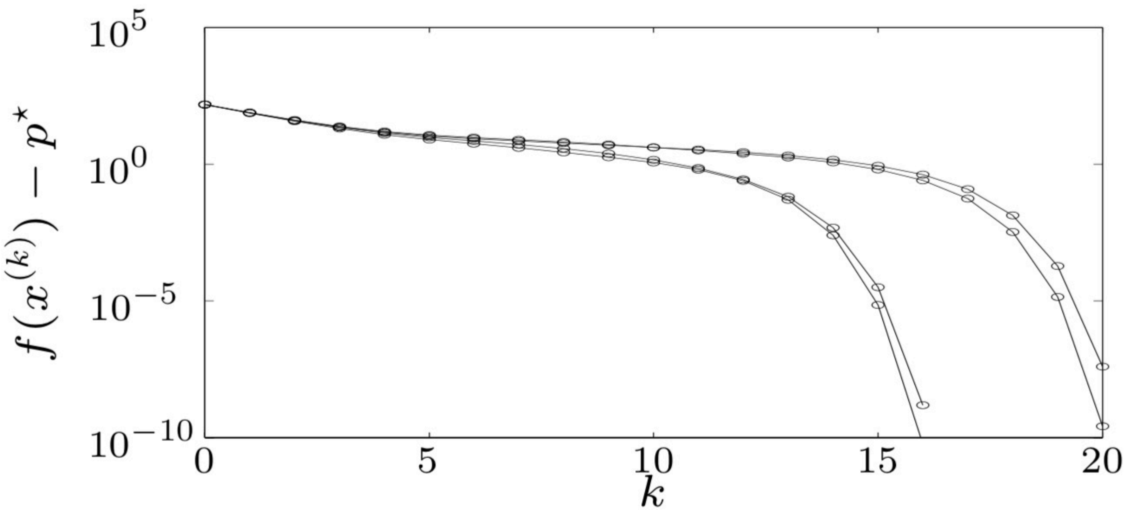
- 原问题: minimize $-\sum_{i=1}^{n}\log x_i$ subject to $Ax=b$
- 对偶问题: maximize $-b^T\nu+\sum_{i=1}^{n}\log(A^T\nu)_i+n$
- 使用三种方法求解 $A\in\mathbf{R}^{100\times 500}$ 中的例子，三个不同起始点

等式约束的解析中心

- 原问题： minimize $-\sum_{i=1}^{n}\log x_i$ subject to $Ax=b$
- 对偶问题: maximize $-b^T\nu+\sum_{i=1}^{n}\log(A^T\nu)_i+n$
- 使用三种方法求解 $A\in\mathbf{R}^{100\times 500}$ 中的例子，三个不同起始点
- 带有等式约束的**Newton**方法，需 $x^{(0)}\succ 0$, $Ax^{(0)}=b$

等式约束的解析中心

- 原问题： minimize $-\sum_{i=1}^{n}\log x_i$ subject to $Ax=b$
- 对偶问题: maximize $-b^T\nu+\sum_{i=1}^{n}\log(A^T\nu)_i+n$
- 使用三种方法求解 $A\in\mathbf{R}^{100\times 500}$ 中的例子，三个不同起始点
- 带有等式约束的**Newton**方法，需 $x^{(0)}\succ 0$, $Ax^{(0)}=b$



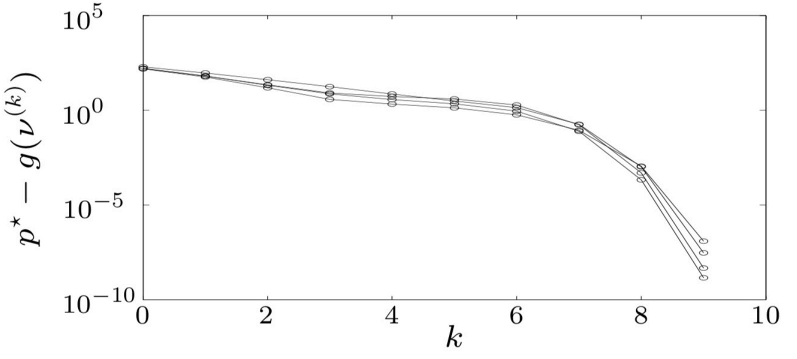
其他两个方法

其他两个方法

- **Newton**方法应用于对偶问题，需 $A^T\nu^{(0)} \succ 0$

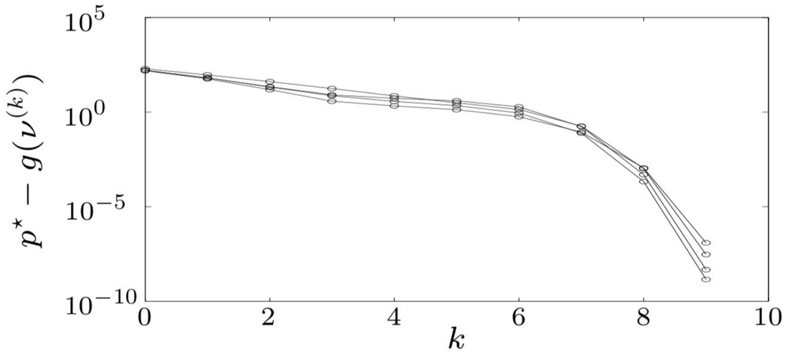
其他两个方法

- **Newton**方法应用于对偶问题，需 $A^T\nu^{(0)} \succ 0$



其他两个方法

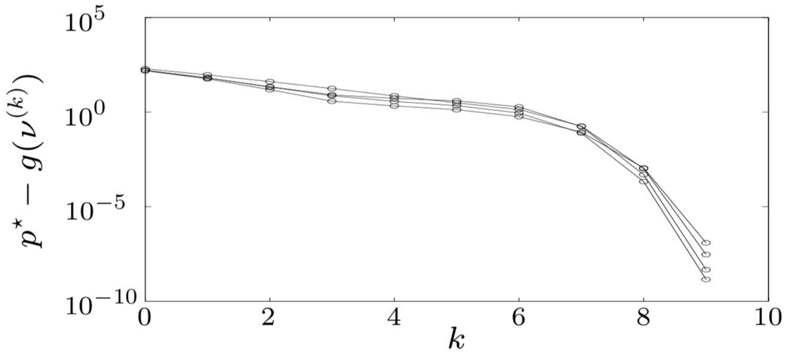
- **Newton**方法应用于对偶问题，需 $A^T\nu^{(0)} \succ 0$



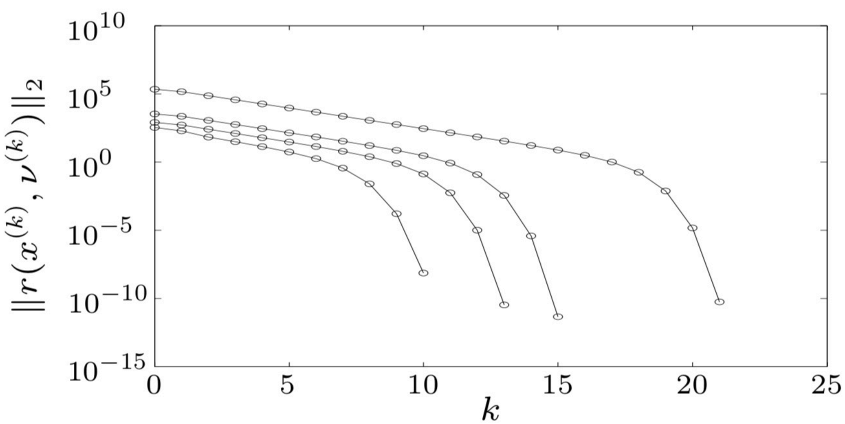
- 不可行初始点的**Newton**方法，需 $x^{(0)} \succ 0$

其他两个方法

- **Newton**方法应用于对偶问题，需 $A^T\nu^{(0)} \succ 0$



- 不可行初始点的**Newton**方法，需 $x^{(0)} \succ 0$



三个方法每次迭代的复杂度

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$
- 求解**Newton**系统 $A\,\mathbf{diag}(A^T\nu)^{-2}A^T\Delta\nu = -b + A\,\mathbf{diag}(A^T\nu)^{-1}\mathbf{1}$

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$
- 求解**Newton**系统 $A\,\mathbf{diag}(A^T\nu)^{-2}A^T\Delta\nu = -b + A\,\mathbf{diag}(A^T\nu)^{-1}\mathbf{1}$
- 使用分块消元求解**KKT**系统

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$

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

- 使用分块消元求解**KKT**系统

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

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$

- 求解**Newton**系统 $A\,\mathbf{diag}(A^T\nu)^{-2}A^T\Delta\nu = -b + A\,\mathbf{diag}(A^T\nu)^{-1}\mathbf{1}$
- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = 2Ax - b$

三个方法每次迭代的复杂度

- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = b$
- 求解**Newton**系统 $A\,\mathbf{diag}(A^T\nu)^{-2}A^T\Delta\nu = -b + A\,\mathbf{diag}(A^T\nu)^{-1}\mathbf{1}$
- 使用分块消元求解**KKT**系统

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

  - 简化为求解 $A\,\mathbf{diag}(x)^2 A^T w = 2Ax - b$
- 结论：求解带有正对角矩阵*D*的 $ADA^T w = h$

最优网络流

最优网络流

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

最优网络流

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

- 有向图带有n条边和p+**1**个结点

最优网络流

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

- 有向图带有n条边和p+**1**个结点
- x_i： 通过边i的流量；ϕ_i: 边i的流程成本函数（ $\phi_i''(x) > 0$ ）

最优网络流

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

- 有向图带有n条边和p+**1**个结点
- x_i： 通过边i的流量; ϕ_i: 边i的流程成本函数（ $\phi_i''(x) > 0$ ）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

最优网络流

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

- 有向图带有n条边和p+1个结点
- x_i：通过边i的流量；ϕ_i: 边i的流程成本函数（$\phi_i''(x) > 0$）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

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

最优网络流

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

- 有向图带有*n*条边和*p*+**1**个结点
- x_i： 通过边*i*的流量；ϕ_i: 边*i*的流程成本函数（ $\phi_i''(x) > 0$ ）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

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

- 简化结点进入矩阵 $A \in \mathbf{R}^{p\times n}$， 去除最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

最优网络流

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

- 有向图带有n条边和p+1个结点
- x_i： 通过边i的流量；ϕ_i: 边i的流程成本函数（ $\phi_i''(x) > 0$ ）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

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

- 简化结点进入矩阵 $A \in \mathbf{R}^{p\times n}$，去除最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$
- （简化的）资源向量 $b \in \mathbf{R}^{p}$

最优网络流

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

- 有向图带有*n*条边和***p*+1**个结点
- x_i： 通过边*i*的流量；ϕ_i: 边*i*的流程成本函数（ $\phi_i''(x) > 0$ ）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

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

- 简化结点进入矩阵 $A \in \mathbf{R}^{p\times n}$， 去除最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$
- （简化的）资源向量 $b \in \mathbf{R}^p$
- 若图是连接的，$\mathbf{rank}\, A = p$

最优网络流

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

- 有向图带有n条边和p+**1**个结点
- x_i：通过边i的流量；ϕ_i: 边i的流程成本函数（$\phi_i''(x) > 0$）
- 结点进入矩阵 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$

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

- 简化结点进入矩阵 $A \in \mathbf{R}^{p\times n}$，去除最后一行的 $\tilde{A} \in \mathbf{R}^{(p+1)\times n}$
- （简化的）资源向量 $b \in \mathbf{R}^p$
- 若图是连接的，$\mathbf{rank}\, A = p$

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

最优控制

- **KKT**系统 $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = -\begin{bmatrix} g \\ h \end{bmatrix}$
- $H = \mathbf{diag}(\phi_1''(x_1), \ldots, \phi_n''(x_n))$，正对角矩阵

最优控制

- **KKT**系统 $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$
- $H = \mathbf{diag}(\phi_1''(x_1), \ldots, \phi_n''(x_n))$，正对角矩阵
- 通过消元法求解：

最优控制

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

- $H = \mathbf{diag}(\phi_1''(x_1), \ldots, \phi_n''(x_n))$，正对角矩阵
- 通过消元法求解：

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

最优控制

- **KKT**系统 $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = - \begin{bmatrix} g \\ h \end{bmatrix}$
- $H = \mathbf{diag}(\phi_1''(x_1), \ldots, \phi_n''(x_n))$，正对角矩阵
- 通过消元法求解：

  $$AH^{-1}A^T w = h - AH^{-1}g, \qquad Hv = -(g + A^T w)$$
- 系数矩阵的稀疏模式通过图的连接性获取

最优控制

- **KKT**系统 $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = -\begin{bmatrix} g \\ h \end{bmatrix}$
- $H = \mathbf{diag}(\phi_1''(x_1), \ldots, \phi_n''(x_n))$，正对角矩阵
- 通过消元法求解：

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

- 系数矩阵的稀疏模式通过图的连接性获取

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

$\Longleftrightarrow$ 结点*i*和*j*存在边连接

线性矩阵不等式的解析中心

线性矩阵不等式的解析中心

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

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$
- 优化条件

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$
- 优化条件

$$X^\star \succ 0, \qquad -(X^\star)^{-1} + \sum_{j=1}^{p} \nu_j^\star A_i = 0, \qquad \mathbf{tr}(A_iX^\star) = b_i, \quad i = 1,\ldots,p$$

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$
- 优化条件

$$X^\star \succ 0, \qquad -(X^\star)^{-1} + \sum_{j=1}^{p} \nu_j^\star A_i = 0, \qquad \mathbf{tr}(A_i X^\star) = b_i, \quad i = 1, \ldots, p$$

- 在可行点X的**Newton**等式

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$
- 优化条件

$$
X^\star \succ 0, \qquad -(X^\star)^{-1} + \sum_{j=1}^{p} \nu_j^\star A_i = 0, \qquad \mathbf{tr}(A_i X^\star) = b_i, \quad i = 1, \ldots, p
$$

- 在可行点*X*的**Newton**等式

$$
X^{-1}\Delta X X^{-1} + \sum_{j=1}^{p} w_j A_i = X^{-1}, \qquad \mathbf{tr}(A_i \Delta X) = 0, \quad i = 1, \ldots, p
$$

线性矩阵不等式的解析中心

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

- 优化变量 $X \in \mathbf{S}^n$
- 优化条件

$$X^\star \succ 0, \qquad -(X^\star)^{-1}+\sum_{j=1}^{p}\nu_j^\star A_i=0, \qquad \mathbf{tr}(A_iX^\star)=b_i, \quad i=1,\dots,p$$

- 在可行点X的**Newton**等式

$$X^{-1}\Delta X X^{-1}+\sum_{j=1}^{p}w_jA_i=X^{-1}, \qquad \mathbf{tr}(A_i\Delta X)=0, \quad i=1,\dots,p$$

- 根据线性近似 $(X+\Delta X)^{-1}\approx X^{-1}-X^{-1}\Delta X X^{-1}$
  得到 $n(n+1)/2+p$ 个变量 ΔX, w

分块消元求解

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

- 为一个密集正定的线性方程组，变量为 $w \in \mathbf{R}^p$

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

- 为一个密集正定的线性方程组，变量为 $w \in \mathbf{R}^p$
- 使用**Cholesky**因式分解$X=LL^T$的浮点运算次数：

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

- 为一个密集正定的线性方程组，变量为 $w \in \mathbf{R}^p$
- 使用**Cholesky**因式分解X=LL^T的浮点运算次数：
- 计算p个乘积 $L^T A_j L$: $(3/2)pn^3$

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

- 为一个密集正定的线性方程组，变量为 $w \in \mathbf{R}^p$
- 使用**Cholesky**因式分解$X=LL^T$的浮点运算次数：
- 计算p个乘积 $L^T A_j L$: $(3/2)pn^3$
- 计算 $p(p+1)/2$个内积：$\mathbf{tr}((L^T A_i L)(L^T A_j L))$: $(1/2)p^2 n^2$

分块消元求解

- 从第一个方程中消去 ΔX: $\Delta X = X - \sum_{j=1}^{p} w_j X A_j X$
- 将 ΔX 带入第二个方程

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

- 为一个密集正定的线性方程组，变量为 $w \in \mathbf{R}^p$
- 使用**Cholesky**因式分解X=LL^T的浮点运算次数:
- 计算p个乘积 $L^T A_j L$: $(3/2)pn^3$
- 计算 $p(p+1)/2$个内积: $\mathbf{tr}((L^T A_i L)(L^T A_j L))$: $(1/2)p^2n^2$
- 通过**Cholesky**因式分解求解: $(1/3)p^3$