2.6 线性规划的非内部连续化算法

最优性条件

$$\begin{cases} Ax = b, \\ A^\top y + s = c, \\ x \geq 0, \quad s \geq 0, \quad x^\top s = 0, \end{cases} \tag{2.5.1}$$

注意到: 在最优性条件(2.5.1)中, $x \geq 0, s \geq 0, x^\top s = 0$当且仅当对任意的$i \in \{1, 2, \ldots, n\}$都有$x_i \geq 0, s_i \geq 0, x_i s_i = 0$. 因此, 最优性条件(2.5.1)等价地可表示为:

$$\begin{cases} Ax = b, \\ A^\top y + s = c, \\ x_i \geq 0, \quad s_i \geq 0, \quad (x_i)^\top s_i = 0, \quad \forall i \in \{1, 2, \ldots, n\}. \end{cases}$$

若存在函数$\phi : \mathbb{R}^2 \to \mathbb{R}$使得

$$\phi(a, b) = 0 \quad \Longleftrightarrow \quad a \geq 0, b \geq 0, ab = 0,$$

则最优性条件(2.5.1)可等价地转化为方程组.

2.6 线性规划的非内部连续化算法

定义函数

$$\phi(a,b) = a + b - \sqrt{(a-b)^2}, \quad \forall (a,b) \in \mathbb{R} \times \mathbb{R}.$$

则容易证明: $\phi(a,b) = 0$当且仅当$a \geq 0, b \geq 0, ab = 0$. 则最优性条件(2.5.1)等价地可表示为:

$$F(x,y,s) := \begin{pmatrix} Ax - b \\ A^\top y + s - c \\ \phi(x_1, s_1) \\ \vdots \\ \phi(x_n, s_n) \end{pmatrix} = 0. \tag{2.6.1}$$

这表明: 可通过求解方程组$F(x,y,s) = 0$来求解原线性规划问题(2.2.9)和对偶线性规划(2.2.10) 的最优解. 然而, 虽然函数$\phi(\cdot,\cdot)$是连续的, 但它不是可微的, 经典的牛顿法不能直接应用于求解方程组$F(x,y,s) = 0$. 为了能够应用牛顿法求解方程组$F(x,y,s) = 0$, 需要对函数$\phi(\cdot,\cdot)$进行光滑化. 下面将介绍一类光滑化技术.

2.6 线性规划的非内部连续化算法

任给参数$\mu > 0$, 定义函数

$$\phi_\mu(a,b) = a + b - \sqrt{(a-b)^2 + 4\mu^2}, \quad \forall (a,b) \in \mathbb{R} \times \mathbb{R}. \tag{2.6.2}$$

易知: 函数$\phi_\mu(\cdot,\cdot)$是无穷次连续可微的, 称$\phi_\mu(\cdot,\cdot)$为$\phi(\cdot,\cdot)$的光滑逼近函数, 简称$\phi_\mu(\cdot,\cdot)$为光滑函数. 定义函数

$$\Phi_\mu(x,s) = \begin{pmatrix} \phi_\mu(x_1,s_1) \\ \vdots \\ \phi_\mu(x_n,s_n) \end{pmatrix} \quad \text{和} \quad F_\mu(x,y,s) = \begin{pmatrix} Ax-b \\ A^\top y + s - c \\ \Phi_\mu(x,s) \end{pmatrix}. \tag{2.6.3}$$

则(x,y,s)是(2.2.9)和(2.2.10) 的最优解当且仅当

$$F_\mu(x,y,s) = 0 \quad \text{and} \quad \mu = 0.$$

2.6 线性规划的非内部连续化算法

容易得到: 函数$F_\mu(x,y,s)$连续可微, 并且它的Jacobi矩阵为

$$F'_\mu(x,y,s)=\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ \partial_x\Phi_\mu(x,s) & 0 & \partial_s\Phi_\mu(x,s) \end{pmatrix}, \tag{2.6.4}$$

其中, $I\in\mathbb{R}^{n\times n}$是单位矩阵, 且

$$\begin{aligned} \partial_x\Phi_\mu(x,s) &= \operatorname{diag}\left\{1-\frac{x_i-s_i}{\sqrt{(x_i-s_i)^2+4\mu^2}}\ \middle|\ i\in\{1,2,\ldots,n\}\right\}, \\ \partial_s\Phi_\mu(x,s) &= \operatorname{diag}\left\{1+\frac{x_i-s_i}{\sqrt{(x_i-s_i)^2+4\mu^2}}\ \middle|\ i\in\{1,2,\ldots,n\}\right\}. \end{aligned}$$

特别地, 类似于引理2.5.1的证明可得: 当系数矩阵A行满秩时, 则(2.6.4)式中的雅可比矩阵是可逆的. 因此, 人们可以用以下方法来求解原线性规划问题(2.2.9)和对偶线性规划(2.2.10)的最优解: 用牛顿型方法迭代求解方程组$F_\mu(x,y,s)=0$并且在迭代的过程中不断地降低μ值, 使其趋于零. 具体算法如下:

2.6 线性规划的非内部连续化算法

算法 2.6.1 (非内部连续化算法) 选择参数$\delta, \gamma, \sigma \in (0,1)$及$\mu_0 \in (0, \infty)$. 取$x^0 \in \mathbb{R}^n$并使得$Ax^0 = b$. 取$y^0 \in \mathbb{R}^m$, 并置$s^0 := c - A^\top y^0$且$w^0 := (x^0, y^0, s^0)$. 选择$\beta$使得$\beta \geq 2\sqrt{n}$且$\|F_{\mu_0}(w^0)\| \leq \beta\mu_0$. 置$k := 0$.

步1 如果$F_0(w^k) = 0$, 算法终止.

步2 如果$F_{\mu_k}(w^k) = 0$, 那么置$w^{k+1} := w^k$及$\theta_k := 1$, 转步4; 否则, 令$\Delta w^k := (\Delta x^k, \Delta y^k, \Delta s^k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ 是下述方程组的解:

$$F'_{\mu_k}(w^k)\Delta w^k = -F_{\mu_k}(w^k). \tag{2.6.5}$$

步3 令θ_k是$\{1, \delta, \delta^2, \cdots\}$中使得下述不等式成立的最大值:

$$\|F_{\mu_k}(w^k + \theta_k \Delta w^k)\| \leq [1 - \sigma\theta_k]\beta\mu_k. \tag{2.6.6}$$

置$w^{k+1} := w^k + \theta_k \Delta w^k$. 若$F_0(w^{k+1}) = 0$, 则算法终止; 否则, 转步4.

步4 置

$$\hat{\mu}_k := \left(1 - \frac{1}{1 + \sqrt{2}}\sigma\theta_k\right)\mu_k. \tag{2.6.7}$$

令η_k是$\{1, \gamma, \gamma^2, \cdots\}$中使得下述不等式成立的最小值:

$$\|F_{\eta_k \hat{\mu}_k}(w^{k+1})\| \leq \beta\eta_k\hat{\mu}_k. \tag{2.6.8}$$

置$\mu_{k+1} := \eta_k\hat{\mu}_k$和$k := k + 1$, 转步1.

2.6 线性规划的非内部连续化算法

设$\mathcal{K}$表示迭代指标集. 构造两个集合:

$$\mathcal{K}_1 := \{k \in \mathcal{K} \mid F_{\mu_k}(w^k) = 0\} \quad 和 \quad \mathcal{K}_2 := \{k \in \mathcal{K} \mid F_{\mu_k}(w^k) \neq 0\} \tag{2.6.9}$$

那么, $\mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{K}$且$\mathcal{K}_1 \cap \mathcal{K}_2 = \emptyset$. 因此, 对任意的$k \in \mathcal{K}$, 有$k \in \mathcal{K}_1$或者$k \in \mathcal{K}_2$.

引理 2.6.1 令假设*2.5.1*中条件(b)成立. 那么算法*2.6.1*是适定的. 如果序列$\{(x^k, y^k, s^k)\}$是由算法*2.6.1*产生的迭代序列, 那么下面结论成立.

- 数列$\{\mu_k\}$是单调下降的且对任意的$k \in \mathcal{K}$有$\mu_k > 0$.
- 对任意的$k \in \mathcal{K}$, 有$\|F_{\mu_k}(w^k)\| \leq \beta\mu_k$.
- 对任意的$k \in \mathcal{K}_2$, 有$Ax^k = b$且$A^\top y^k + s^k = c$, 其中$\mathcal{K}_2$由*(2.6.9)*式定义.

2.6 线性规划的非内部连续化算法

首先考虑算法2.6.1的全局收敛性. 定义光滑路径的邻域:

$$\mathcal{N}(\beta,\mu) := \{w = (x, y, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mid \|F_\mu(w)\| \le \beta\mu\},$$

其中$\mu > 0$, 且$\beta > 0$是邻域的宽度. 那么, $\mu \in (0, \mu_0]$的邻域片被定义为

$$\mathcal{N}(\beta,(0,\mu_0)) := \{w \in \mathcal{N}(\beta,\mu) \mid \mu \in (0,\mu_0)\}. \tag{2.6.10}$$

本节使用下述假设条件.

假设 2.6.1 *假设(2.6.10)式定义的邻域片是有界的.*

这一假设在非内部连续化算法中是基本的条件. 事实上, 大多数在文献中使用的相关假设是这一假设的充分条件.

引理 2.6.2 *令假设2.5.1中的条件(b)和假设2.6.1条件成立, 且算法2.6.1产生的无穷迭代序列为$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$. 那么, $\lim_{k\to\infty}\mu_k = 0$.*

2.6 线性规划的非内部连续化算法

算法2.6.1的全局收敛性结论如下.

定理 2.6.1 令假设2.5.1中的条件(b)和假设2.6.1条件成立, 且算法(2.6.1)产生的无穷迭代序列为$\{\mu_k, w^k\}_{k\in\mathcal{K}}$. 那么, 序列$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$是有界的, 且序列$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$的每个聚点$(\mu_*, w^*) := (\mu_*, x^*, y^*, s^*) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n$满足$\mu_* = 0$并且$w^*$是最优性条件(2.5.1)的一个解.

下面考虑算法的全局线性收敛性. 需要使用如下假设.

假设 2.6.2 令序列$\{\mu_k\}_{k\in\mathcal{K}}$和$\{\Delta w^k\}_{k\in\mathcal{K}_2}$由算法2.6.1产生, 其中$\mathcal{K}_2$由(2.6.9)式定义. 那么, 存在一个常数$C > 0$使得对所有的$k \in \mathcal{K}_2$, 有$\|\Delta w^k\| \le C\mu_k$.

这一假设弱于文献中相关的假设条件(见文献[16]).

对任意的$\mu \in \mathbb{R}$和$c := (a, b) \in \mathbb{R}^2$, 令$\nabla_c\phi(\mu, c) := \nabla_{(a,b)}\phi(\mu, a, b)$, 它表示$\phi$相对于变量$a$和$b$的梯度. 令$\{\Delta w^k\}_{k\in\mathcal{K}_2}$由算法2.6.1产生, 其中$\mathcal{K}_2$由(2.6.9)式定义. 记

$$\widetilde{\Delta w}_i^k := ((\Delta x^k)_i, (\Delta s^k)_i)^\top, \quad \forall i \in \{1, 2, \ldots, n\}, \quad \forall k \in \mathcal{K}_2.$$

2.6 线性规划的非内部连续化算法

引理 2.6.3 令函数ϕ由(2.6.2)式定义, 序列$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$和$\{\Delta w^k\}_{k\in\mathcal{K}_2}$由算法2.6.1产生, 其中指标集$\mathcal{K}_2$由(2.6.9)式定义. 那么对任意的$i \in \{1,2,\ldots,n\}$, $k \in \mathcal{K}_2$和$\alpha \in [0,1)$, 有

$$\left|\phi(\mu_k, \tilde{w}_i^k + \alpha\widetilde{\Delta w}_i^k) - \phi(\mu_k, \tilde{w}_i^k) - \alpha\nabla_{\tilde{w}_i^k}\phi(\mu_k, \tilde{w}_i^k)^\top \widetilde{\Delta w}_i^k\right| \le \frac{\alpha^2}{\mu_k}\|\widetilde{\Delta w}_i^k\|^2.$$

算法2.6.1的全局线性收敛性结论如下.

定理 2.6.2 令假设2.5.1中的条件(b), 假设2.6.1以及假设2.6.2条件成立, 且$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$为由算法2.6.1产生的无穷迭代序列. 那么, 存在常数$\breve{C} \in (0,1)$使得对任意的$k \in \mathcal{K}$, 有$\mu_{k+1} \le \breve{C}\mu_k$.

2.6 线性规划的非内部连续化算法

下面考虑算法的局部二次收敛性. 令$(\mu_*, w^*) := (\mu_*, x^*, y^*, s^*)$为算法2.6.1所产生迭代序列的聚点, 那么$\mu_* = 0$且$w^*$是线性规划问题(2.2.9)和对偶线性规划(2.2.10)的最优解, 做如下假设.

假设 2.6.3 w^*是线性规划问题*(2.2.9)*和对偶线性规划*(2.2.10)*的严格互补最优解, 即: $x^* + s^* > 0$.

如果假设2.6.3成立, 那么假设2.6.2成立(见文献[16]).

引理 2.6.4 令无穷序列$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$和$\{\triangle w^k\}_{k\in\mathcal{K}_2}$由算法*2.6.1*产生, 其中$\mathcal{K}_2$由*(2.6.9)*式定义. 如果假设*2.5.1*中的条件*(b)*, 假设*2.6.1*和假设*2.6.3*条件成立, 那么对所有充分大的$k \in \mathcal{K}_2$, 有$\theta^k = 1$, 其中θ_k由算法*2.6.1*的第*3*步定义.

定理 2.6.3 令假设*2.5.1*中的条件(*b*), 假设*2.6.1*以及假设*2.6.3*条件成立, 且$\{(\mu_k, w^k)\}_{k\in\mathcal{K}}$为由算法*2.6.1* 产生的无穷迭代序列. 那么,

$$\mu_{k+1} = O(\mu_k^2).$$

2.6线性规划的非内部连续化算法— 作业

2.20 证明以下两个命题:

(1) 如果定义函数

$$\phi(a,b) = a + b - \sqrt{a^2 + b^2}, \quad \forall (a,b) \in \mathbb{R} \times \mathbb{R}.$$

那么$\phi(a,b) = 0$当且仅当$a \geq 0, b \geq 0, ab = 0.$

(2) 如果定义函数

$$\phi(\mu,a,b) = a + b - \sqrt{a^2 + b^2 + 2\mu^2}, \quad \forall (\mu,a,b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

那么$\phi(\mu,a,b) = 0$当且仅当$a \geq 0, b \geq 0, ab = \mu.$