

线性规划中的原始对偶内点法

主讲人 邱欣欣
幻灯片制作 邱欣欣

中国海洋大学 信息科学与工程学院

2014 年 6 月 6 日

线性规划的标准型



$$\begin{aligned} \min \quad & c^T x, \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0 \end{aligned} \tag{1}$$

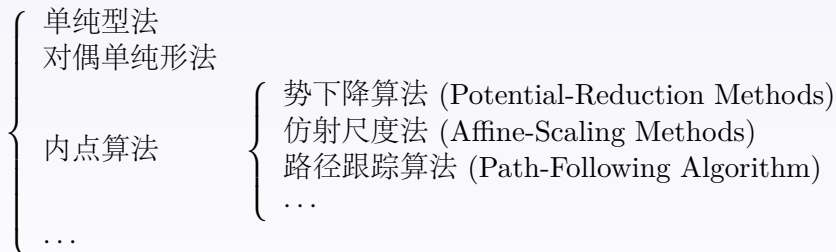
$c = (c_1, c_2, \dots, c_n)^T$ 称为价值系数向量, $b = (b_1, b_2, \dots, b_m) \in R^m$ 为约束的右端向量并要求 $b \geq 0$, A 为约束的 $m \times n$ 阶系数矩阵, 其中我们假定 $m < n$ 且行满秩。

线性规划的基本概念和特性

- 满足所有约束条件的解被称为可行解。
- 使目标函数值取得最优的可行解被称为最优解。
- 满足非负性约束的解称为基本解。
- 若线性规划有最优解，则最优解一定可在某个基本可行解上取得，也即在可行域的某个顶点（极点）上取得。

求解线性规划的最优化

•



线性规划的对偶

- 对一个线性规划问题，可以构造另一个与之相应且密切相关的线性规划问题，如果称前者为原始问题，后者就称为对偶问题。

-

$$\begin{aligned} \max \quad & b^T \lambda, \\ \text{s.t.} \quad & A^T \lambda + s = c, \\ & s \geq 0 \end{aligned} \tag{2}$$

其中 λ 是 R^m 中的向量， s 是 R^n 中的向量， s 称为松弛变量。

线性规划的原始问题与对偶问题

- 线性规划的弱对偶定理和强对偶定理描述了线性规划的原始问题的最优值和对偶问题的最优值之间的关系。
- 如果两个问题各有一个可行解，且相应的目标函数值相同，则这两个可行解各是相应问题的最优解，即设 x^* 和 y^* 分别是原始问题和对偶问题的可行解， $b^T \lambda^* = c^T x^*$ ，此时 x^* 是 (1) 的最优解， (λ^*, s^*) 是 (2) 的最优解。
- 由 $b^T \lambda = c^T x$ ，可得 $x^T s = 0$ 。

Karush-Kuhn-Tucker(KKT) condition

THEOREM 1

The vector $x^ \in R^n$ is a solution of (1) if and only if there exist vectors $s^* \in R^n$ and $\lambda^* \in R^m$ for which the following conditions hold for $(x, \lambda, s) = (x^*, \lambda^*, s^*)$:*

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= 0, i = 1, 2, \dots, n, \\ (x, s) &\geq 0 \end{aligned} \tag{3}$$

原始-对偶内点法

- 一个向量 (x^*, λ^*, s^*) 能够求解方程 (3) 当且仅当 x^* 求解原始问题 (1) and (λ^*, s^*) 求解对偶问题 (2). 向量 (x^*, λ^*, s^*) 被称作原始-对偶解.
- 原始对偶内点算法就是通过应用牛顿法的变体求上述方程组的三个等式, 修正搜索方向和步长使不等式 $(x, s) \geq 0$ 在每次迭代严格满足从而求解 (x^*, λ^*, s^*) 。

原始-对偶内点法

•

$$F(x, \lambda, s) = \left\{ \begin{array}{c} A^T \lambda + s - c \\ Ax = b \\ XSe \end{array} \right\} = 0, \\ (x, s) \geq 0,$$

其中 $X = \text{diag}(x_1, x_2, \dots, x_n)$, $S = \text{diag}(s_1, s_2, \dots, s_n)$, 第一二项是线性的, 第三项是非线性的。

- 每次迭代中, $x^k > 0$, $s^k > 0$, 这个性质是内点的性质。
- 牛顿步长方程为

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe \end{bmatrix}$$

原始-对偶内点法

原始-对偶方法在两个方面修正了基本的牛顿过程：

- 它使搜索方向沿着非负区域 $(x, s) \geq 0$ 的内部前进，因此我们可以在 (x, s) 的分量为负前沿着方向走的更远。
- 它避免迭代点过于接近非负区域的边界，因为接近于边界的点计算的搜索方向对于到达最优解没有太大帮助。
- 我们沿着牛顿方向采用线性搜索使新的迭代满足

$$(x, \lambda, s) + \alpha * (\Delta x, \Delta \lambda, \Delta s)$$

线性搜索参数 $\alpha \in (0, 1]$ 。

•

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= \tau, i = 1, 2, \dots, n, \\ (x, s) &> 0 \end{aligned} \tag{4}$$

其中 $\tau > 0$ ，从而得中心路径

$$C = \{(x_\tau, \lambda_\tau, s_\tau) | \tau > 0\}$$

- 当 $\tau \rightarrow 0$ 时， $x_i s_i \rightarrow 0$ ，因此对逐步减小并收敛于零的正数序列 τ ，求方程组 (4) 的解，就可保持解的严格可行性，并在 $\tau \rightarrow 0$ 时最终确定原问题的最优解 x^* 和相应对偶问题的最优解 (λ^*, s^*) 。

原始-对偶内点法

- 任取一点 (x, λ, s) , 其中 $x > 0, s > 0$, 求取一个方向 $(\Delta x, \Delta \lambda, \Delta s)$, 使得新的迭代点 $(x + \Delta x, \lambda + \Delta \lambda, s + \Delta s)$ 位于原始-对偶中心路径上, 即满足方程 (4)。

$$\begin{aligned} A^T(\lambda + \Delta \lambda) + s + \Delta s &= c, \\ A(x + \Delta x) &= b, \\ (x_i + \Delta x_i)(s_i + \Delta s_i) &= \tau^+, i = 1, 2, \dots, n, \\ x + \Delta x > 0, s + \Delta s &> 0, \end{aligned} \tag{5}$$

其中 $\tau^+ < \tau$, 展开三个方程并略去 Δx_i 和 Δs_i 的高次项。



$$\begin{aligned} A^T \Delta \lambda + \Delta s &= 0, \\ A \Delta x &= 0, \\ x_i \Delta s_i + s_i \Delta x_i &= \tau^+ - x_i s_i, i = 1, 2, \dots, n, \\ x + \Delta x &> 0, s + \Delta s > 0, \end{aligned} \tag{6}$$

我们引入参数 $\sigma \in [0, 1]$, 使 $\tau^+ = \sigma * \mu$ 。 σ 的取值应能使 τ^+ 较快下降收敛于 0。 μ 定义为 $\mu = \frac{1}{n} \sum_{i=1}^n x_i s_i = x^T s / n$ 。

- 步长方程变为

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \sigma \mu e \end{bmatrix} \quad (7)$$

搜索方向是朝着点 $(x_{\sigma\mu}, \lambda_{\sigma\mu}, s_{\sigma\mu})$ 的牛顿步，如果 $\sigma = 1$ ，方程 (7) 定义了中心方向，其中所有的 $x_i s_i$ 等于 μ 。如果 $\sigma = 0$ ，是标准牛顿步。

- 当 μ 的值以较快的速度下降时， $x + \Delta x > 0, s + \Delta s > 0$ 不再得到保证，为确保迭代点的严格可行性，取 $x^+ = x + \alpha * \Delta x$, $\lambda^+ = \lambda + \alpha * \Delta \lambda, s^+ = s + \alpha * \Delta s$ ，步长 α 应在确保 $x > 0, s > 0$ 的前提下尽可能大。

原始-对偶内点法

- 原始-对偶方法的框架如下:

Given $(x^0, \lambda^0, s^0) \in F^0$

for $k=0,1,2,\dots$

 solve

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -X^k S^k e + \sigma_k \mu_k e \end{bmatrix}$$

 where $\sigma_k \in [0, 1]$ and $\mu_k = (x^k)^T s^k / n$;

 set

$(x^{k+1}, \lambda^{k+1}, s^{k+1}) \leftarrow (x^k, \lambda^k, s^k) + \alpha_k (\Delta x^k, \Delta \lambda^k, \Delta s^k)$,

 choosing α_k so that $(x^{k+1}, s^{k+1}) > 0$.

end (for)

- $F^0 = \{(x, \lambda, s) | Ax = b, A^T \lambda + s = c, (x, s) > 0\}$

不可行的原始对偶内点算法

- 上述算法要求用户提供一个可行的初始内点 (x^0, λ^0, s^0) ，对大多数实际问题这是一个比较困难的问题，其中一种解决的方法是采用不可行的原始对偶内点算法，该算法不要求初始点可行，而只要求是严格内点，也就是可以取任 $x^0 > 0, s^0 > 0$ 作为初始点。
- 我们定义两个线性方程的残差

$$r_b = Ax - b, \quad r_c = A^T \lambda + s - c$$

步长方程变为

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma \mu e \end{bmatrix}$$

如果点 x 可行，则 r_b 为零；同样如果点 (λ, s) 可行，则 r_c 为零。

- 线性规划问题的最优解必定在可行域的边界。而上述介绍的内点算法，不管求解的精度多高，所确定的只是一个接近最优解的严格内点。对此可利用最后所得迭代点提供的信息确定问题的一个最优基本可行解。

- 大多数线性规划的内点代码是依据于 Mehrotra 的预测校正算法 (Predictor-Corrector Algorithm), 这个算法的两个要点是:
在上述的原始-对偶框架中的搜索方向添加修正步, 因此使算法沿着解集轨道前进。
中心参数 σ 的适应选择。
- 路径-跟踪算法要求每步产生的迭代点均落在一个包含中心路径 C 的邻域内, 且沿着这条中心路径可搜索至原问题的最优解, 其中 C 的邻域是 2-范数邻域和 ∞ -范数。