内点法求解线性规划问题

原创 余孙婕 LOA算法学习笔记 2021-01-19 21:11

这里主要介绍两种内点法，即障碍函数法和原始对偶内点法。另外还有投影尺度法、仿射尺度法等经典内点算法，这里不做介绍。

01 障碍函数法

障碍函数法的思想是在可行域的边界处人为的架起边界“墙壁”，使得目标函数在迭代值到达边界使迅速增大作为惩罚，以阻止迭代值穿过边界，这样就可以保证最优解被拦在可行域范围之内了。
换句话说，是通过引入效用函数的方法将约束优化问题转换成无约束问题，再利用优化迭代过程不断地更新效用函数，以使得算法收敛。
对于一般的线性规划问题

$$\min c^T x$$

$$s.t. Ax \le b$$

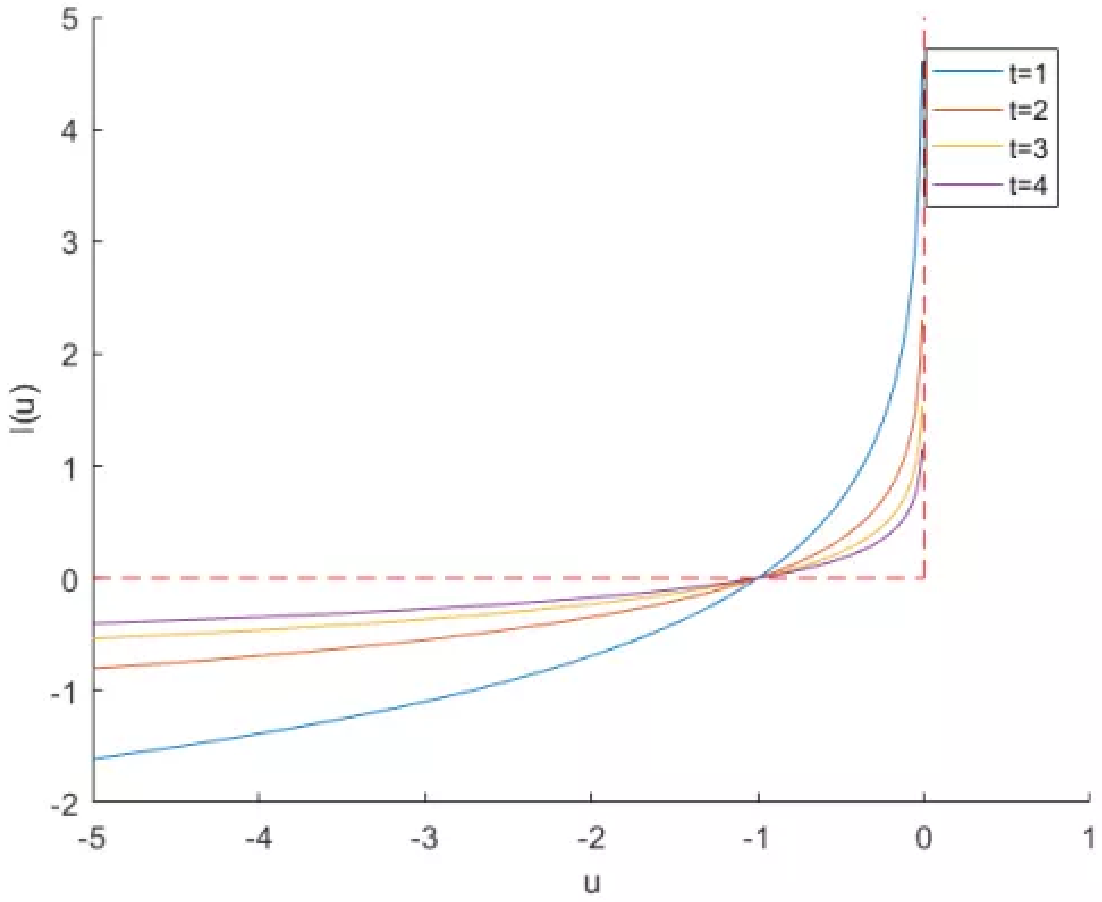
我们希望可以把它转化为无约束问题：

$$f(x) = c^T x + \sum_{i=1}^{m} I(A_{ij} x_j - b_i)$$

这里我们引入了I(u)(也就是惩罚/障碍函数)，目的是为了将约束条件写入目标函数中，我们只需要对新的无约束目标函数迭代就行了。另外，我们希望 I(u) 函数具有如下的功能：在没有违反约束时，函数值为 0，而在违反约束时，函数值为正无穷，从而达到将迭代值限制在可行域的目标。把I(u)转换成公式就是：

$$I(u) = \begin{cases} 0, x \le 0 \\ \infty, x > 0 \end{cases}$$

也就是下图中的红色虚线部分。



但是很可惜的是，红色的虚线很明显地在0点处不可导，为此我们用实线对其进行近似，其中一个用的比较多的近似函数如下：

$$I_t(u) = -\frac{1}{t}\log(-u)$$

该式在u<0 时才有定义，我们人为规定 u>0 时为正无穷，且参数 t>0 越大，$I_t(u)$ 就越接近我们的理想障碍函数，我们可以通过调节 t 来调节函数的近似程度。因此，新的目标函数可以写作：

$$f_t(x) = tc^T x - \sum_{i=1}^{m} \log(-A_{ij}x_j + b_i)$$

内点法不仅可以解决线性规划问题，也可以有效解决非线性规划问题，表示如下：

$$\min f(x)$$

$$s.t. g_i(x) \le 0$$

构造的无约束目标函数一般如下：

$$\varphi(x,r) = f(x) - r\sum_{i=1}^{m} \ln[-g_i(x)]$$

由于 $\varphi(x,r)$ 是凸函数，我们可以用凸优化的经典方法得到该函数的极小值。

算法步骤如下：

(1) 取初始惩罚因子 $r^{(0)}>0$，误差 $\varepsilon>0$

(2) 在可行域 D 内选取初始点 x_0，令 k = 1

(3) 构造惩罚函数 $\varphi(x,r_k)$，从 x_{k-1} 出发用无约束优化方法求惩罚函数 $\varphi(x,r_k)$ 的极值点 $\varphi(x^*,r_k)$

(4) 如果满足 $||x^*(r_k)-x^*(r_{k-1})||\le\varepsilon$ 则停止迭代，并将 $x^*(r_k)$ 作为原目标函数f(x)的约束最优解，否则取 $r_{k-1}=cr_k$ (其中c常取0.1)，$x_0=x^*(r_k),k=k+1$，重复(3)-(4)

举例分析：

$$\min f(x)=x_1^2+x_2^2-x_1x_2-10x_1-4x_2+60$$
$$s.t. x_1+x_2-8\le 0$$

令

$$x_0=(0,0)^T,r_0=1,\varepsilon=0.01$$

构造无约束目标函数

$$\varphi(x,r)=x_1^2+x_2^2-x_1x_2-10x_1-4x_2+60-r\ln(x_1+x_2-8)$$

求梯度

$$\nabla\varphi(x,r)=\begin{pmatrix}2x_1-x_2-10-\dfrac{r}{x_1+x_2-8} & 2x_2-x_1-4-\dfrac{r}{x_1+x_2-8}\end{pmatrix}=0$$

则有

$$x_1^*(r)=\begin{pmatrix}\dfrac{13+\sqrt{9+2r}}{2} & \dfrac{9+\sqrt{9+2r}}{2}\end{pmatrix}^T$$

由于 $g(x_1^*(r))>0$，故舍去；

$$x^*(r)=x_2^*(r)=\begin{pmatrix}\dfrac{13-\sqrt{9+2r}}{2} & \dfrac{9-\sqrt{9+2r}}{2}\end{pmatrix}^T$$

为当前最优解。

$$r_0=1,x^*(r_0)=(4.8417\quad 2.8417)^T,\left\|x^*(r_0)-x_0\right\|=5.6140>\varepsilon$$

后续迭代过程同上，直至 $x^*(r_k)<\varepsilon$

02 原始对偶内点法

原始对偶内点法始终保持原始问题和对偶问题的解在每次迭代中严格可行，并在搜索过程中系统性的减少互补松弛性的违背程度。

考虑标准 LP 问题

$$\min c^T x$$
$$s.t. Ax = b, x \geq 0$$

以及它的对偶问题

$$\max b^T \lambda$$
$$s.t. A^T \lambda + s = c, s \geq 0$$

我们找到上述两个问题的 KKT 条件如下：

$$A^T \lambda + s = c$$
$$Ax = b$$
$$x_i s_i = 0$$
$$(x, s) \geq 0$$

即

$$F(x,\lambda,s) = \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe \end{pmatrix} = 0, 其中 (x,s) \geq 0, X = diag(x), S = diag(s)$$

求解原目标问题，其实本质是求解上述方程组，显然这是个非线性方程组(XSe非线性)，我们同样适用牛顿法进行求解，即确定当前点的一个前进方向，并通过迭代逼近非线性方程组的解。

这里步长的选取也十分的有意思，但本文不做过多介绍，其结果是导致 KKT 条件中， $x_i s_i = \tau = \sigma\mu, \mu = \frac{x^T s}{n}$ ，衡量迭代点与最优点的差距，替换了原 $x_i s_i = 0$ 。给出以下算法步骤：

(1) 选取初始点满足

$$\|f_0 - \mu_0 e\| \leq \theta\mu_0, 其中 f_0 = (x_1^0 s_1^0, \ldots, x_n^0 s_n^0)^T, \mu_0 = \frac{(x_0)^T s_0}{n}, e = (1 \ldots 1)^T$$

给定较小的 θ, δ ，给定终止误差 $\varepsilon > 0$ ，令k=0

(2) 计算

$$\left(\left\| A^T \lambda_k + s_k - c \right\|_2^2 + \left\| A^T x_k - b \right\|_2^2 \right)^{1/2} < \varepsilon$$

和对偶间隙 $x_k^T s_k < \varepsilon$ ，如果满足，则 x_k 为近似解，反之则继续(3)

(3) 计算 $\mu_{k+1} = \mu_k(1 - \delta/\sqrt{n})$ ，利用

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_k & 0 & X_k \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{pmatrix} = J(x_k, \lambda_k, s_k)\begin{pmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{pmatrix} = -F(x_k, \lambda_k, s_k) = \begin{pmatrix} -(A^T\lambda_k + s_k - c) \\ -(Ax_k - b) \\ -X_k S_k e + \sigma\mu_k e \end{pmatrix}$$

求解 $\Delta F_k = (\Delta x, \Delta\lambda, \Delta s)$

(4) $F_{k+1} = F_k + \Delta F_k, k = k+1$，重复步骤(2)-(4)