数值优化（Numerical Optimization）学习系列-线性规划（Linear Programming）

下一步 于 2015-12-27 18:54:46 发布

概述

> 线性规划问题是指目标和约束函数都是线性的最简单的约束最优化问题，也是在实际中最长使用的模型之一。其求解算法也是相对成熟，各个代数软件中都会有求解该问题的工具，本节主要介绍：
> 1. 线性规划的基本形式已经对偶
> 2. 线性规划两类求解算法：单纯形和 内点法
> 3. 总结

线性规划问题

标准形式

线性规划的标准形式通常表示为

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

其中矩阵A是一个m*n的矩阵，通常情况下是 **行满秩矩阵**，否则可能解不存在或者唯一解。
对于其他形式都可以转化为标准形式，例如

$$min\ c^Tx,\ s.t\ Ax\leq b$$

可以采用松弛的方法进行转换

$$min\ c^Tx,\ s.t\ Ax+z=b; z\geq 0$$

此时目标函数中的变量还不满足非负数约束，可以采用 $x=x^+-x^-$，其中 $x^+=max(x,0); x^-=max(-x,0)$
同理对于其他形式，均可以转换

$$x\leq u <=> x+w=u, w\geq 0$$

$$Ax\geq b <=> Ax-y=b; y\geq 0$$

> 线性规划解的形式，根据约束的不同，该问题可能有唯一解、一条线、一个平面，无解（Infeasible）或者无界（Unbounded）
> 最后两种情况都可以认为无界，一个可行解为空，一个可行解不可达
> 线性规划主要考虑，即行满秩，此时可能有多个解，其他都可能唯一解或者无界。

对偶问题

由于线性的关系，该问题满足LICQ时，局部最优解就是全局最优解，因此原始问题和对偶问题的解一致。拉格朗日函数为

$$\mathcal{L}(x,\lambda,s)=c^Tx-\lambda^T(Ax-b)-s^Tx$$

其中拉格朗日因子 λ和s分别对应等式约束和非负约束，根据KKT条件有

$$\begin{gathered}A^T\lambda+s=c\\Ax=b\\x\geq 0\\s\geq 0\\x^Ts=0\end{gathered}$$

> 1.对于该KKT条件对应的最优解有如下一个性质，
>
> $$c^Tx^*=(A^T\lambda^*+s^*)x^*=(Ax^*)^T\lambda^*=b^T\lambda^*$$
>
> 即目标函数转换为 $b^T\lambda$，即对偶问题的目标
> 2.可以证明满足该KKT条件的可行解，就是全局最优解。

对偶问题

$$max\ b^T\lambda\ \ s.t\ A^T\lambda\leq c$$

转换为标准形式为

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

其中 (λ,s)叫做对偶变量。

性质

原始问题和对偶问题是对称的，并且满足强对偶条件，因此

1. 如果原始或者对偶问题有解，则对应问题也有解
2. 如果颜色或者对偶问题是无界的，则对应问题则为无解。

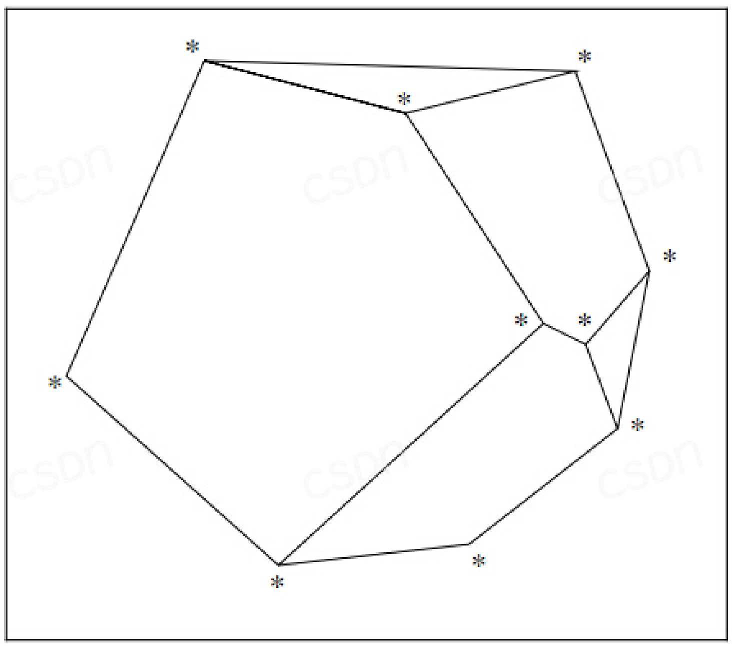
线性规划求解

概念

在介绍算法之前，有一个比较重要的概念（基矩阵和基础可行解）介绍一下，在后续的介绍中会频繁应用到。

> 基础可行点，需要满足如下条件
> 1. 指示集合包含m个下标
> 2. 当
> 3. 基矩阵B定义为

简单示意如下图



单纯形算法

单纯形算法的基本策略遍历线性规划问题所有的基础可行解，直到收敛到一个基础最优解上，并且有定理证明**当线性规划有界并且有可行解时，肯定会收敛到基础最优解上**

定义，则单纯形算法其中一步如下

Procedure 13.1 (One Step of Simplex).

Given $\mathcal{B}, \mathcal{N}$, $x_B = B^{-1}b \geq 0$, $x_N = 0$;
Solve $B^T\lambda = c_B$ for λ,
Compute $s_N = c_N - N^T\lambda$; (* pricing *)
if $s_N \geq 0$
  stop; (* optimal point found *)
Select $q \in \mathcal{N}$ with $s_q < 0$ as the entering index;
Solve $Bd = A_q$ for d;
if $d \leq 0$
  stop; (* problem is unbounded *)
Calculate $x_q^+ = \min_{i \mid d_i > 0} (x_B)_i / d_i$, and use p to denote the minimizing i;
Update $x_B^+ = x_B - dx_q^+$, $x_N^+ = (0, \ldots, 0, x_q^+, 0, \ldots, 0)^T$;
Change $\mathcal{B}$ by adding q and removing the basic variable corresponding to column p of B

单纯形算法需要解决如下几个关键点

> 1. 如何选取初始点，主要有两类算法可以求解得到，都是转换线性规划本身进行求解，一是PhaseI 转换为
>
> $$min\ e^T z,\ \ s.t.\ \ Ax + Ez = b;\ \ (x, z) \geq 0$$
>
> 二是PhaseII 转换为
>
> $$min\ c^T x\ \ s.t\ Ax + z = b, x \geq 0,\ 0 \geq z \geq 0$$
>
> 2.如何有效表示x_N
>
> 2. 如何计算离开下标
> **详细的单纯形算法参考各类参考文献。**

内点法

对于小规模线性规划问题，单纯形算法能很好的解决，但是大规模问题效率比较低，而内点法能够很好的解决该问题。
内点法每一步骤计算比较昂贵但是能够有效的解决最优解。
内点法的主要思路，逐渐寻找满足KKT条件并且最小化某价值函数的解，根据KKT条件有

$$
\begin{gathered}
A^T\lambda + s = c \\
Ax = b \\
(x, s) \geq 0 \\
x^T s = 0
\end{gathered}
$$

转换为矩阵表示为

$$
F(x,\lambda,s) = \begin{bmatrix} A^T + s - c \\ AX - b \\ XSe \end{bmatrix} = 0
$$

问题转换为求解该方程，根据之前的学习，此时通过牛顿法进行求解，一般情况下还需要一个解的度量函数，一般定义为 $u = \frac{x^T s}{n}$
根据牛顿方法有

$$
J(x,\lambda,s) \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = -F(x,\lambda,s)
$$

根据该方程求解搜索方向，一般情况下还要满足非负约束 $(x, s) \geq 0$，因此要求解一定步长，即下一个搜索点为 $(x,\lambda,s) + \alpha(\Delta x, \Delta\lambda, \Delta s)$
但是大多数内点法不会直接求解上述方程，一般会进行一定缩放，即$x^T s = \delta u$，即求解

$$
F(x,\lambda,s) = \begin{bmatrix} A^T + s - c \\ AX - b \\ XSe + \delta ue \end{bmatrix} = 0
$$

因此一个相对完整的原始对偶内点法步骤为

Framework 14.1 (Primal-Dual Path-Following).

Given (x^0, λ^0, s^0) with $(x^0, s^0) > 0$;

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

Choose $\sigma_k \in [0, 1]$ and solve

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

where $\mu_k = (x^k)^T s^k / n$;

Set

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

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

end (for).

中心路径

中心路径在内点法中是一个非常重要的概念，因为每次迭代点都是中心路径的点，中心路径的选取非常关键，参数选择太大或者太小都会影响计算复杂度。某个点属于中心路径，即则满足

$$\begin{gathered} A^T \lambda + s = c \\ Ax = b \\ (x, s) \geq 0 \\ x^T s = \tau \end{gathered}$$

该KKT条件也对应某个最优化问题（对数障碍问题），也是内点法求解其他非线性约束最优化问题的基础。

$$min\ c^T x - \tau \sum_{i=1 \ldots n} ln x_i, \quad Ax = b$$

此时只要逐渐改变 $\tau \rightarrow 0$的值，就能找到最优解，因此 τ如何改变非常重要，也是很多算法改进的关键点。

总结

通过本小节的学习，能够了解到

1. 线性规划的基本形式已经原始问题和对偶问题的关系
2. 单纯形和内点法如何求解线性规划问题

详细的单纯形算法网上介绍的非常多，经典的可以参考wiki