Advances in Applied Mathematics 应用数学进展, **2019, 8(8), 1398-1409**
Published Online August 2019 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2019.88164

Hans汉斯

A SQP Method for Linear Complementary Constraints Problem

Qingqun Huang

School of Mathematics and Statistics, Hechi University, Yizhou Guangxi
Email: 44584330@qq.com

Received: Jul. 25th, 2019; accepted: Aug. 12th, 2019; published: Aug. 19th, 2019

Abstract

In this paper, equilibrium problem with linear equilibrium constrains is studied. By using a smoothing complimentarily function which is differentiable everywhere, when smoothing parameter tends to zero, the original problem is equivalently transformed to a smoothing nonlinear optimization, then the smooth nonlinear optimization is solved by SQP algorithm. The algorithm converges globally under certain conditions. Finally, numerical experiments are given to prove the effectiveness of the algorithm.

Keywords

Complentarity Problem, Penalty Function, SQP Method, Global Convergence

求解线性互补约束问题的一个SQP方法

黄青群

河池学院数学与统计学院，广西 宜州
Email: 44584330@qq.com

收稿日期：2019年7月25日；录用日期：2019年8月12日；发布日期：2019年8月19日

摘　要

对线性互补约束优化问题，利用一个连续可微的光滑互补函数，光滑系数趋于零时，原问题转化为光滑非线性规划问题，再利用SQP算法来求解该光滑非线性规划问题。在适当的条件下，该算法具有全局收敛性，最后给出了数值实验，也证明了该算法是有效的。

文章引用: 黄青群. 求解线性互补约束问题的一个 SQP 方法[J]. 应用数学进展, 2019, 8(8): 1398-1409.
DOI: 10.12677/aam.2019.88164

关键词

均衡问题，罚函数，SQP法，全局收敛

1. 引言

近年来，带均衡约束的数学规划问题因为在经济、工程技术、对策决策等领域的广泛应用而引起关注[1] [2]。本文求解一个线性互补均衡约束问题，可以描写为以下形式：

$$\begin{aligned} &\min\ \ f(x,y) \\ &\text{s.t.}\ \ Ax\le b, \\ &\qquad w=Nx+My+q, \\ &\qquad 0\le w\perp y\ge 0, \end{aligned} \tag{1.1}$$

其中 $f:R^{n+m}\to R$，是连续可微的函数，$A\in R^{p\times n},N\in R^{m\times n},M\in R^{m\times m},b\in R^{p},q\in R^{m}$。

如果将互补条件 $w\perp y$ 写成内积的形式：$w^{\mathrm{T}}y=0$，那么均衡约束问题等价于一个光滑的非线性规划问题。但众所周知，对于此类非线性划归问题，即使较弱的 MFCQ 规格,在任何可行点处都不成立，故传统的优化算法一般不能直接用来求解此类问题。

对线性互补均衡约束问题，先介绍一些常见的光滑函数：

广义 Fischer-Burmeister 互补函数：

$$\phi_1(a,b)=a+b-\sqrt{a^2+b^2+\mu},$$

Zang 在文献[3]提出一个如下形式的光滑函数：

$$\phi_2(a,\mu)=\begin{cases}0, & a\le -\dfrac{\mu}{2}, \\ \dfrac{1}{2\mu}\left(a+\dfrac{\mu}{2}\right)^2, & -\dfrac{\mu}{2}<a<\dfrac{\mu}{2}, \\ a, & a\ge -\dfrac{\mu}{2}.\end{cases}$$

Song Xu 在文献[4]中提出一个如下形式的新的光滑函数：

$$\phi_3(a,\mu)=\mu\ln\sum_{i=1}^{m}\mathrm{e}^{\frac{a_i}{\mu}}.$$

本文尝试利用互补函数 $\phi(y_i,w_i,\mu)=-\mu\ln\left(1+\mathrm{e}^{-\frac{y_i}{\mu}}\right)-\mu\ln\left(1+\mathrm{e}^{-\frac{w_i}{\mu}}\right)$，将问题(1.1)中的互补约束条件 $w\perp y$ 光滑化，将问题(1.1)转化为一光滑非线性规划的一般约束优化问题，然后借助罚函数型 SQP 方法思想，通过求解后者逼近前者的解，并证明该算法是全局收敛的.

为了方便，本文采用如下记号：

$$z=(x,y,w),\ s=(x,y),\ t=(y,w),\ z^k=\left(x^k,y^k,w^k\right),\ s^k=\left(x^k,y^k\right),\ t^k=\left(y^k,w^k\right),$$
$$t_i=(y_i,w_i),\ \mathrm{d}z=(\mathrm{d}x,\mathrm{d}y,\mathrm{d}w),\ \mathrm{d}z^k=\left(\mathrm{d}x^k,\mathrm{d}y^k,\mathrm{d}w^k\right),\ X=\{z\in X_0\mid 0\le w\perp y\ge 0\},$$
$$X_0=\{z\mid Ax\le b, w=Nx+My+q\},\ A^{\mathrm{T}}=\left(a_1^{\mathrm{T}},\cdots,a_p^{\mathrm{T}}\right),\ b^{\mathrm{T}}=\left(b_1,\cdots,b_p\right).$$

本文主要符号解释如下：

R^n	表示一个实n维$Euclide$空间。
e	表示各分量均为1的列向量。
$\nabla f(x)$	表示函数 $f(x)$ 的一阶导数或梯度。
$\nabla^2 f(x)$	表示函数 $f(x)$ 的二阶导数或 $Hessian$ 矩阵。
$u>0$	表示向量u的每一个分量 $u_j>0$。
$u\ge 0$	表示向量u的每一个分量 $u_j\ge 0$。
x^{T} 和 A^{T}	分别表示向量x和矩阵A的转置。
$diag\left(a_j, j=1,2,\sim,m\right)$	表示以 a_j 为对角元素的 m 阶对角矩阵。

2. 问题光滑化

记F是问题(1.1)的可行域，定义在向量 $z\in F$ 处F的切锥如下：

$$L(z,F)=\left\{d=\lim_{k\to\infty}\frac{z^k-z}{\tau}: z^k\in F, \lim_{k\to\infty} z^k=z, \tau_k\downarrow 0\right\},$$

对问题(1.1)，需做如下基本假设：

H 2.1 问题(1.1)中的矩阵M是一个 P_0 矩阵，即所有的主子式均非负。

H 2.2 对任意的 $z\in X_0$，矩阵A的行向量组 $\left\{a_j \mid j\in L(z)\right\}$ 线性无关。

H 2.3 问题(1.1)的可行域非空。

H 2.4 函数 $f(x,y)$ 为二阶连续可微的。

定义 2.1 如果对于可行点 $z^*=\left(x^*,y^*,w^*\right)$ 满足下面的条件，那么称 z^* 是问题(1.1)的稳定点：

$$\mathrm{d}z=(\mathrm{d}x,\mathrm{d}y,\mathrm{d}w)\in\Gamma\left(X;z^*\right)\Rightarrow\nabla f\left(x^*,y^*\right)^{\mathrm{T}}\mathrm{d}s\ge 0,$$

其中 $\Gamma\left(X;z^*\right)$ 为X在 z^* 处的切锥。

问题(1.1)的稳定点与其最优解两者之间存在密切的联系，求解问题(1.1)的算法多半集中在稳定点的求解。

定理 2.1 [5] 假设 $z^*=\left(x^*,y^*,w^*\right)\in X$ 满足下层非退化条件：

$$\left(y_i,w_i\right)\ne(0,0),\quad \forall i=1,\cdots,m,$$

则 z^* 为问题(1.1)的一个稳定点的充分必要条件是存在乘子 $\left(\lambda^*,u^*,v^*\right)$，使得下面的方程组成立：

$$\begin{pmatrix}\nabla f_x\left(x^*,y^*\right)\\ \nabla f_y\left(x^*,y^*\right)\\ 0\end{pmatrix}+\begin{pmatrix}A^{\mathrm{T}}\\ 0\\ 0\end{pmatrix}\lambda^*+\begin{pmatrix}N^{\mathrm{T}}\\ M^{\mathrm{T}}\\ -I\end{pmatrix}u^*+\begin{pmatrix}0\\ W^*\\ Y^*\end{pmatrix}v^*=0,$$
$$\left(Ax^*-b\right)^{\mathrm{T}}\lambda^*=0,\quad Ax^*-b\ge 0,\quad \lambda^*\ge 0, \tag{2.1}$$

其中 $W^* = diag\left(w_i^*, i=1,\cdots,m\right), Y^* = diag\left(y_i^*, i=1,\cdots,m\right)$。

本文为问题(1.1)提出一个新的光滑函数.

令 $\varphi(y,w)=\min\{y,w\}$，其中 $\min\{y,w\}$ 定义为 $\min\{y_i,w_i\}e, i=1,\cdots,m$， $e=(1,\cdots,1)^{\mathrm{T}}\in R^m$。有

$$y^{\mathrm{T}}w=0, y\ge 0, w\ge 0 \Leftrightarrow \varphi(y,w)=\min\{y,w\}=0. \tag{2.2}$$

定义函数 $\Phi\in R^{2m}\times[0,+\infty)$：

$$\Phi(a,b,\mu)=\begin{pmatrix}\phi(a_1,b_1,\mu)\\ \vdots \\ \phi(a_m,b_m,\mu)\end{pmatrix}, \tag{2.3}$$

其中 $\phi\in R^{2m}\times[0,+\infty)$， μ 是一个非负的参数，

$$\phi(a,b,\mu)=-\mu\ln\left(1+\mathrm{e}^{-\frac{a}{\mu}}\right)-\mu\ln\left(1+\mathrm{e}^{-\frac{b}{\mu}}\right), \tag{2.4}$$

易知对于固定的 $\mu>0$，函数 $\phi(\cdot,\cdot,\mu)$ 处处可微.

$$\begin{aligned}&\nabla\phi(y,w,\mu)\\ &=\begin{pmatrix}\nabla_y\phi(y,w,\mu)\\ \nabla w\phi(y,w,\mu)\end{pmatrix}\\ &=\left(\frac{\partial\phi(y_1,w_1,\mu)}{\partial y_1},\cdots,\frac{\partial\phi(y_m,w_m,\mu)}{\partial y_m},\frac{\partial\phi(y_1,w_1,\mu)}{\partial w_1},\cdots,\frac{\partial\phi(y_m,w_m,\mu)}{\partial w_m}\right)^{\mathrm{T}}\\ &=\left(\frac{\mathrm{e}^{-\frac{y_1}{\mu}}}{1+\mathrm{e}^{-\frac{y_1}{\mu}}},\cdots,\frac{\mathrm{e}^{-\frac{y_m}{\mu}}}{1+\mathrm{e}^{-\frac{y_m}{\mu}}},\frac{\mathrm{e}^{-\frac{w_1}{\mu}}}{1+\mathrm{e}^{-\frac{w_1}{\mu}}},\cdots,\frac{\mathrm{e}^{-\frac{w_m}{\mu}}}{1+\mathrm{e}^{-\frac{w_m}{\mu}}}\right)^{\mathrm{T}}.\end{aligned} \tag{2.5}$$

$\phi(a,b,\mu)$ 有一些很好的性质：

$$\lim_{\mu\to 0}\phi(a,b,\mu)=\lim_{\mu\to 0}-\mu\ln\left(1+\mathrm{e}^{-\frac{a}{\mu}}\right)-\mu\ln\left(1+\mathrm{e}^{-\frac{b}{\mu}}\right)=0,$$

且当 $\mu=0$ 时，

$$\lim_{\mu\to 0}\Phi(y,w,\mu)=\varphi(y,w)=0. \tag{2.6}$$

由上式，做如下定义：

$$\Phi(y,w,0)=\varphi(y,w), \tag{2.7}$$

从而有

$$\lim_{\mu\to 0}\phi(y,w,\mu)=\phi(y,w,0). \tag{2.8}$$

如将问题(1.1)的互补条件转化为线性方程组 $\phi(y,w,0)=0, i=1,\cdots,m$。则提出以下标准非线性规划问题：

$$\begin{aligned}&\min\ f(x,y)\\&\text{s.t.}\ \ Ax\le b,\\&\qquad w=Nx+My+q,\\&\qquad \Phi(y,w,\mu)=0,\end{aligned}\tag{2.9}$$

当 $\mu\ge 0$ 且充分逼近于0时，问题(2.9)是原问题(1.1)的一个近似逼近；当 $\mu=0$ 时，问题(2.9)是原问题(1.1)的等价；当 $\mu>0$ 时，问题(2.9)是一个一般的可微非线性规划问题。

对于 $z=(x,y,w)\in R^{n+2m}$，可以定义如下罚函数 $\varphi:R^{n+2m}\times(0,+\infty)\times(0,+\infty)\to R$：

$$\varphi(z,\varepsilon,\mu)=f(x,y)+\varepsilon\sum_{i=1}^{m}\left|\phi(y_i,w_i,\mu)\right|,\tag{2.10}$$

其中 $\varepsilon>0$ 是罚参数.

罚函数 φ 在z处沿方向 $\mathrm{d}z$ 的方向导数为 $\varphi'(z,\varepsilon,\mu,\mathrm{d}z)$：

$$\begin{aligned}&\varphi'(z,\varepsilon,\mu,\mathrm{d}z)\\&=\nabla f(s)^{\mathrm{T}}\mathrm{d}s+\varepsilon\sum_{i\in I_+}\nabla\phi(t_i,\mu)^{\mathrm{T}}\mathrm{d}t_i+\varepsilon\sum_{i\in I_0}\left|\nabla\phi(t_i,\mu)^{\mathrm{T}}\mathrm{d}t_i\right|-\varepsilon\sum_{i\in I_-}\left|\nabla\phi(t_i,\mu)^{\mathrm{T}}\mathrm{d}t_i\right|,\end{aligned}\tag{2.11}$$

其中下指标集定义如下：

$$\begin{aligned}&I_+\equiv I_+(y,w,\mu)=\{i\mid\phi(y_i,w_i,\mu)>0\},\\&I_0\equiv I_0(y,w,\mu)=\{i\mid\phi(y_i,w_i,\mu)=0\},\\&I_-\equiv I_-(y,w,\mu)=\{i\mid\phi(y_i,w_i,\mu)<0\}.\end{aligned}$$

3. 预备知识与算法

引理 3.1 对于任意的 $(a,b,\mu)\in R^2\times(0,+\infty)$，且 $(a,b,\mu)\ne(0,0,0)$，有

$$0<\partial_a\phi(a,b,\mu_k)<1,0<\partial_b\phi(a,b,\mu_k)<1,\ \ i=1,\cdots,m,\tag{3.1}$$

$$0<\left(\frac{\partial\phi(a,b,\mu_k)}{\partial a}\right)^2+\left(\frac{\partial\phi(a,b,\mu_k)}{\partial b}\right)^2\le 2\tag{3.2}$$

成立。

证明：由(2.5)，有(3.1)成立。

接下来证明(3.2)式如下：

$$\begin{aligned}&\left(\frac{\partial\phi(y,w,\mu_k)}{\partial y}\right)^2+\left(\frac{\partial\phi(y,w,\mu_k)}{\partial w}\right)^2\\&=\left(\frac{\mathrm{e}^{-\frac{y}{\mu}}}{1+\mathrm{e}^{-\frac{y}{\mu}}}\right)^2+\left(\frac{\mathrm{e}^{-\frac{w}{\mu}}}{1+\mathrm{e}^{-\frac{w}{\mu}}}\right)^2\\&\le\left(\frac{\mathrm{e}^{-\frac{y}{\mu}}}{\mathrm{e}^{-\frac{y}{\mu}}}\right)^2+\left(\frac{\mathrm{e}^{-\frac{w}{\mu}}}{\mathrm{e}^{-\frac{w}{\mu}}}\right)^2\\&=2.\end{aligned}$$

设 $z^k=(x^k,y^k,w^k)\in X_0$，是算法第 k 步的迭代点，$\mu>0$，B_k 是一个 $n+2m$ 阶的对称正定矩阵，在算法的第 k 步迭代中，求解以下二次序列子问题

$$\begin{aligned} \min\ & \nabla f\left(x^k\quad y^k\right)^{\mathrm{T}}\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\end{pmatrix}+\frac{1}{2}\left(\mathrm{d}x^{\mathrm{T}}\quad \mathrm{d}y^{\mathrm{T}}\quad \mathrm{d}w^{\mathrm{T}}\right)B_k\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\\ \mathrm{d}w\end{pmatrix},\\ \text{s.t.}\ & Ax^k+A\mathrm{d}x-b\le 0,\\ & \mathrm{d}w-N\mathrm{d}x-M\mathrm{d}y=0,\\ & \Phi\left(t^k,\mu^k\right)+\nabla_t\Phi\left(t^k,\mu\right)^{\mathrm{T}}\mathrm{d}t=0. \end{aligned} \tag{3.3}$$

求解(3.3)式，得解 $\mathrm{d}z^k=\left(\mathrm{d}x^k\quad \mathrm{d}y^k\quad \mathrm{d}w^k\right)$，及相应的 KKT 乘子 $\left(\lambda^k,l^k,m^k\right)$，定义 $\nabla\Phi\left(y^k,w^k,\mu^k\right)\in R^{2m}\times R^m\times R\to R^m$。

$$\nabla\Phi\left(y^k,w^k,\mu^k\right)=\begin{pmatrix}D_y^k\\ D_w^k\end{pmatrix}, \tag{3.4}$$

其中

$$D_y^k=diag\left(\partial_{y_i}\phi\left(y_i^k,w_i^k,\mu_k\right)\right),\ i=1,\cdots,m$$
$$D_w^k=diag\left(\partial_{w_i}\phi\left(y_i^k,w_i^k,\mu_k\right)\right),\ i=1,\cdots,m$$

从而 QP 子问题(3.3)也可以写成如下形式：

$$\begin{aligned} \min\ & \nabla f\left(x^k\quad y^k\right)^{\mathrm{T}}\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\end{pmatrix}+\frac{1}{2}\left(\mathrm{d}x^{\mathrm{T}}\quad \mathrm{d}y^{\mathrm{T}}\quad \mathrm{d}w^{\mathrm{T}}\right)B_k\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\\ \mathrm{d}w\end{pmatrix},\\ \text{s.t.}\ & Ax^k+A\mathrm{d}x-b\le 0,\\ & \begin{pmatrix}N & M & -I\\ 0 & D_y^k & D_w^k\end{pmatrix}\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\\ \mathrm{d}w\end{pmatrix}+\begin{pmatrix}0\\ \Phi\left(y_i,w_i,\mu\right)\end{pmatrix}=0. \end{aligned} \tag{3.5}$$

QP 子问题(3.3)是一个带线性约束的凸规划问题，所以问题(3.5)的最优解与 KKT 点等价，进而其 KKT 系统可以写成：

$$\begin{pmatrix}\nabla f\left(x^k\right)\\ \nabla f\left(y^k\right)\\ 0\end{pmatrix}+B_k\begin{pmatrix}\mathrm{d}x\\ \mathrm{d}y\\ \mathrm{d}w\end{pmatrix}+\begin{pmatrix}N^{\mathrm{T}} & 0\\ M^{\mathrm{T}} & D_y^k\end{pmatrix}\begin{pmatrix}l\\ m\end{pmatrix}+\begin{pmatrix}A^{\mathrm{T}}\\ 0\\ 0\end{pmatrix}\lambda=0, \tag{3.6}$$

$$0\le\left(b-Ax^k-A\mathrm{d}x\right)\perp\lambda\ge 0, \tag{3.7}$$

$$\begin{pmatrix}N & M & -I\\ 0 & D_y^k & D_w^k\end{pmatrix}\mathrm{d}z+\begin{pmatrix}0\\ \Phi\left(y_i,w_i,\mu\right)\end{pmatrix}=0, \tag{3.8}$$

其中 $(\lambda,l,m)\in R^p\times R^m\times R^m$ 是 KKT 乘子。

命题 3.1 [6] 设假设 H 2.1 成立，$z^k=\left(x^k,y^k,w^k\right)\in X$，罚参数 μ_k 是个给定的正数，B_k 是一个对称正定矩阵，那么 QP 子问题(3.3)有唯一最优解。

现假设问题(3.3)的唯一最优解为 $dz^k=\left(dx^k,dy^k,dw^k\right)$，通过对罚函数 φ 作线搜索，得到下一个迭代点 $z^{k+1}=\left(x^{k+1},y^{k+1},w^{k+1}\right)$，也就是说对于给定的参数 ρ,σ，满足 $0<\rho<1,0<\sigma<1$，第 $k+1$ 次迭代点 $z^{k+1}=z^k+\alpha_k dz^k$，其中 α_k 是步长因子，令 $\alpha_k=\delta^{m_k}$，m_k 是使得下列不等式成立的最小的非负整数：

$$\varphi\left(z^k+\delta^{m_k}dz^k,\varepsilon,\mu_k\right)\le\varphi\left(z^k,\varepsilon,\mu_k\right)+\sigma\delta^{m_k}\varphi'\left(z^k,\varepsilon,\mu_k;dz^k\right). \tag{3.9}$$

下面给出求解问题(1.1)的 SQP 算法。

算法 3.1

步骤 0. 选取初始点 $z^0=\left(x^0,y^0,z^0\right)\in X_0$，参数 $\xi,\mu_0,\varepsilon_{-1}\in(0,+\infty)$，$B_0\in R^{n+2m}\times R^{n+2m}$。选取序列 $\{\mu_k\}_{k=0}^{\infty}$ 满足 $\mu_k>0,\mu_{k+1}<\mu_k,\lim\limits_{k\to\infty}\dfrac{\mu_{k+1}}{\mu_k^{\zeta}}=\eta$，其中 $0<\eta<1,1<\zeta<2$，令 $k=0$。

步骤 1. 求解 QP 子问题(3.3)得唯一最优解 $dz^k=\left(dx^k,dy^k,dw^k\right)$ 以及相应的乘子 $\left(\lambda^k,l^k,m^k\right)$，如果 $dz^k=0$，则令 $z^{k+1}=z^k$，$\varepsilon_k=\varepsilon_{k-1}$，转到步骤 4。

步骤 2. 调整罚参数，计算

$$\varepsilon_k=\begin{cases}\varepsilon_{k-1}, & \varepsilon_{k-1}\ge\left\|c^k\right\|_\infty+\xi,\\ \max\left\{\left\|c^k\right\|_\infty+\xi,\varepsilon_{k-1}+2\xi\right\}, & \varepsilon_{k-1}<\left\|c^k\right\|_\infty+\xi,\end{cases} \tag{3.10}$$

其中 $\left\|c^k\right\|_\infty=\max\limits_{1\le i\le m}\left|c_i^k\right|$。

步骤 3. 对函数 $\varphi\left(\cdot,\varepsilon_k,\mu_k\right)$ 按(3.10)做 Armijo 先搜索求得步长 α_k,令 $z^{k+1}=z^k+\alpha_k dz^k$。

步骤 4. 如果 $dz^k=0$,停止;否则,令 $\mu_{k+1}=\beta\mu_k$,更新修正 B_k,得新的对称正定阵,$B_{k+1}\in R^{(n+2m)\times(n+2m)}$,令 $k=k+1$，返回步骤 1。

引理 3.2 若假设 H2.1 成立，$\left(x^k,y^k,w^k\right)\in X_0$，$\mu_k>0$，$B_k$ 是对称正定阵，假设 $dz^k=\left(dx^k,dy^k,dw^k\right)$ 是子问题(3.3)的唯一最优解，相应的 KKT 乘子为 $\left(\lambda^k,l^k,m^k\right)$，若 $\varepsilon_k\ge\left\|c^k\right\|_\infty$，有

$$\varphi'\left(z^k,\varepsilon^k,\mu_k;dz^k\right)\le\left(-dz^k\right)^{\mathrm T}B_k dz^k<0 \tag{3.11}$$

成立，且对任意的 $\rho\in(0,1)$，存在 $\bar\rho>0$，使得对于所有的 $\rho\in(0,\bar\rho)$，有

$$\varphi\left(z^k+\rho dz,\varepsilon,\mu_k\right)\le\varphi\left(z^k,\varepsilon,\mu_k\right)+\sigma\rho\varphi'\left(z^k,\varepsilon,\mu_k;dz^k\right). \tag{3.12}$$

证明：由(3.5)有

$$D_y^k dy^k+D_w^k dw^k+\Phi\left(y^k,w^k,\mu_k\right)=0,$$

$$\begin{pmatrix}\nabla_y\phi\left(y_i,w_i,\mu\right)\\ \nabla_w\phi\left(y_i,w_i,\mu\right)\end{pmatrix}^{\mathrm T}\begin{pmatrix}dy_i^k\\ dw_i^k\end{pmatrix}+\Phi\left(y_i^k,w_i^k,\mu\right)=0. \tag{3.13}$$

由(2.10)和(2.11)有

$$\varphi'\left(z^k,\varepsilon_k,\mu_k;dz^k\right)=\nabla f\left(x^k\quad y^k\right)^{\mathrm T}\begin{pmatrix}dx^k\\ dy^k\end{pmatrix}-\varepsilon_k\sum\left|\phi\left(y^k,w^k,\mu_k\right)\right|. \tag{3.14}$$

由(3.6)得

$$\begin{aligned}&\nabla f\left(x^k\quad y^k\right)^{\mathrm T}\begin{pmatrix}dx^k\\ dy^k\end{pmatrix}+\left(\lambda^k\right)^{\mathrm T}Adx^k+\left(dz^k\right)^{\mathrm T}B_k dz^k\\ &+\left(l^k\right)^{\mathrm T}\left(Ndx+Mdy-dw^k\right)+\left(m^k\right)^{\mathrm T}\left(D_y^k dy^k+D_w^k dw^k\right)=0,\end{aligned}$$

故

$$\begin{aligned}&\nabla f\left(x^k \quad y^k\right)^{\mathrm{T}}\begin{pmatrix}\mathrm{d}x^k\\ \mathrm{d}y^k\end{pmatrix}+\left(\mathrm{d}z^k\right)^{\mathrm{T}}B_k\mathrm{d}z^k-\left(m^k\right)^{\mathrm{T}}\Phi\left(y^k,w^k,\mu_k\right)\\&=\left(\lambda^k\right)^{\mathrm{T}}\left(Ax-b\right)\le 0.\end{aligned}\tag{3.15}$$

由(3.14)，(3.15)有

$$\begin{aligned}&\varphi'\left(z^k,\varepsilon_k,\mu_k;\mathrm{d}z^k\right)\\&\le-\left(\mathrm{d}z^k\right)^{\mathrm{T}}B_k\mathrm{d}z^k+\sum_{i\in I_+^k}\left(c_i^k-\varepsilon_k\right)\phi\left(y_i^k,w_i^k,\mu_k\right)+\sum_{i\in I_-^k}\left(c_i^k+\varepsilon_k\right)\phi\left(y_i^k,w_i^k,\mu_k\right),\end{aligned}$$

其中 I_+^k,I_-^k 分别定义如下：

$$I_+^k=\left\{i:\phi\left(y_i^k,w_i^k,\mu_k\right)>0\right\},$$

$$I_-^k=\left\{i:\phi\left(y_i^k,w_i^k,\mu_k\right)<0\right\}.$$

又由(3.10)及 $\varepsilon_k\ge\left\|c^k\right\|_\infty$，有

$$\varphi'\left(z^k,\varepsilon^k,\mu_k;\mathrm{d}z^k\right)\le\left(-\mathrm{d}z^k\right)^{\mathrm{T}}B_k\mathrm{d}z^k<0.$$

由(2.10)，(3.10) (3.11)和(3.14)易得(3.13)成立.

4. 收敛性分析

本节分析算法的收敛性，为此，我们对对称正定矩阵 B_k 以及算法产生的无穷点列 $\{z_k\}$ 作如下基本假设：

H4.1 无穷点列 $\{z_k\}$ 有界。

H4.2 若 H4.1 成立，序列 $\{z_k\}$ 的每个极限点 $z^*=\left(x^*,y^*,w^*\right)$，满足下层非退化条件：$\left(y^*,w^*\right)\ne(0,0),i=1,\cdots,m.$

H4.3 存在常数 $a,b>0$，有

$$a\|z\|^2\le z^{\mathrm{T}}B_kz\le b\|z\|^2,\ \forall z\in R^{n+2m},k=1,2,\cdots.$$

根据上面的几个假设，有以下引理成立.

引理 4.1 [5] 如果假设 H4.1~H4.3 成立，则

1) 存在一个常数 $C>0$，有

$$\left\|\begin{pmatrix}M & -I\\ D_y^k & D_W^k\end{pmatrix}^{-1}\right\|\le C,\ \forall k\in K.$$

2) 序列 $\left\{\left(\lambda^k,l^k,m^k\right)\right\}$，$\left\{\mathrm{d}z^k\right\}$ 有界。

3) 存在一个正整数 k_0，使得 $\varepsilon_k=\varepsilon_{k_0}=\varepsilon,\forall k\ge k_0$。

命题 4.2 [6] 若数列 $\{v_k\},\{\gamma_k\}$ 满足

$$v_k\ge 0,\ \sum_{k=1}^{\infty}v_k<\infty,\ \gamma_{k+1}\le\gamma_k+v_k,\ k=1,2,\cdots.\tag{4.1}$$

则

1) 数列 $\{\gamma_k\}$ 有上界。即

$$\overline{\lim_{k\to\infty}}\gamma_k < +\infty;$$

2) 数列 $\{\gamma_k\}$ 整列收敛。

命题 4.3 若 H4.1 成立，则对任意的 $\delta_1 > \delta_2 > 0$，点列 $\left(y^k, w^k\right) \in R^2$，有不等式

$$\left|\phi\left(y^k, w^k, \delta_2\right)\right| \le \left|\phi\left(y^k, w^k, \delta_1\right)\right| + 2\left(\ln 2 + \frac{\tilde{M}}{\delta_2}\right)\delta_1$$

成立。

证明：存在一个 $\bar{\delta} > 0$，且 $\delta_1 < \bar{\delta} < \delta_2$，由中值定理，有

$$\begin{aligned}
&\left|\phi\left(y^k, w^k, \delta_2\right)\right| \\
&\le \left|\phi\left(y^k, w^k, \delta_1\right) + \phi'_\mu\left(y^k, w^k, \bar{\delta}\right)\left(\delta_2 - \delta_1\right)\right| \\
&\le \left|\phi\left(y, w, \delta_1\right)\right| + \left[\ln\left(1 + e^{-\frac{y}{\mu}}\right) + \frac{\frac{y}{\mu}\cdot e^{-\frac{y}{\mu}}}{1 + e^{-\frac{y}{\mu}}} + \ln\left(1 + e^{-\frac{w}{\mu}}\right) + \frac{\frac{w}{\mu}\cdot e^{-\frac{w}{\mu}}}{1 + e^{-\frac{w}{\mu}}}\right]\left(\delta_2 - \delta_1\right) \\
&\le \left|\phi\left(y, w, \delta_1\right)\right| + \left[2\ln 2 + \frac{1}{\mu}\left(\frac{y}{\mu} + \frac{w}{\mu}\right)\right]\left(\delta_2 - \delta_1\right) \\
&\le \left|\phi\left(y, w, \delta_1\right)\right| + \left[2\ln 2 + \frac{2}{\mu^2}\max\{y, w\}\right]\left(\delta_2 - \delta_1\right) \\
&\le \left|\phi\left(y, w, \delta_1\right)\right| + 2\left[\ln 2 + \frac{1}{\delta}\max\{y, w\}\right]\left(\delta_2 - \delta_1\right) \\
&\le \left|\phi\left(y^k, w^k, \delta_1\right)\right| + 2\left(\ln 2 + \frac{\tilde{M}}{\delta_2}\right)\delta_1.
\end{aligned}$$

由引理 3.2，命题 4.2 及命题 4.3 易知有以下结论成立.

引理 4.2 若 H4.1 成立，则

$$\lim_{k\to\infty}\varphi\left(z^k, \varepsilon^k, \mu_k\right) = \lim_{k\to\infty}\varphi\left(z^{k+1}, \varepsilon^k, \mu_k\right) = \varphi\left(z^*, \varepsilon, 0\right).$$

由引理 4.1(2)，不妨做如下假设：

$$\mathrm{d}z^k = \mathrm{d}z^* = \left(\mathrm{d}x^*, \mathrm{d}y^*, \mathrm{d}w^*\right), x^k \to x^*, B_k \to B_*, l^k \to l^*, m^k \to m^*, k \in K. \tag{4.2}$$

利用上述引理及假设,来证明本文的算法具有全局收敛性。

引理 4.3 若假设 H 4.1-H4.3 成立，则序列 $\left\{\mathrm{d}z^k\right\}$ 收敛到 0，即

$$\lim_{k\in K}\mathrm{d}z^k = \mathrm{d}z^* = 0. \tag{4.3}$$

证明：我们用反证法来证明，假设 $\mathrm{d}z^* \ne 0$。

由(3.12)有

$$\varphi'\left(z^*, \varepsilon, 0; \mathrm{d}z^k\right) = -\left(\mathrm{d}z^*\right)^{\mathrm{T}} B_* \mathrm{d}z^* < 0.$$

当 $k \to \infty$，序列 $\left\{z^k\right\}$ 收敛于 z^*，记

$$\lim_{k\in K}\varphi'\left(z^k,\varepsilon,\mu_k;\mathrm{d}z^k\right)=\varphi'\left(z^*,\varepsilon,0;\mathrm{d}z^*\right).$$

考虑算法步骤 3 的线搜索，步长 α 在 K 上有界，且远大于 0，设

$$\alpha_k \ge \alpha_* = \inf\left(\alpha_k, k\in K\right) > 0,\ k\in K.$$

又由序列 $\left\{\varphi\left(z^k,\varepsilon,\mu_k\right)\right\}$ 为单调递减序列，当 μ_k 固定时，由序列 $\left\{z^k\right\}$ 是收敛到 z^* 的,及函数 φ 的连续性，当 $k\to\infty$ 时，有 $\varphi\left(z^k,\varepsilon,\mu_k\right)\to\varphi\left(z^*,\varepsilon,0\right)$。

由(3.9)及引理 2.4.1(3)，有

$$0=\lim_{k\in K}\left(\varphi\left(z^{k+1},\varepsilon,\mu_k\right)-\varphi\left(z^k,\varepsilon,\mu_k\right)\right)\le\lim_{k\in K}\delta\alpha\varphi'\left(z^k,\varepsilon,\mu_k;\mathrm{d}z^k\right)\le\frac{1}{2}\delta\alpha_*\varphi'\left(z^k,\varepsilon,0;\mathrm{d}z^k\right),$$

当 $k\in K,k\to\infty$ 时。这与(3.12)矛盾，从而假设 $\mathrm{d}z^*\ne 0$ 不成立，即(4.3)成立.

定理 4.1 H4.1~H4.3 成立，则算法 3.1 产生的点列 $\left\{z^k\right\}$，它的每一个聚点 $z^*=\left(x^*,y^*,w^*\right)$ 都是问题(1.1)的稳定点，即本文的算法是全局收敛的。

证明：由已知，序列 $\left\{z^k\right\}$ 是个无穷序列， $z^k\to z^*$，有 $z^*\in X_0$,即

$$Ax^*\le b,\ w^*=Nx^*+My^*+q,$$

由(2.7)和(3.3)有

$$\begin{aligned}
&\nabla f\left(x^*,y^*\right)+A^{\mathrm{T}}\lambda^*+N^{\mathrm{T}}l^*=0,\\
&M^{\mathrm{T}}l^*+D_y^*m^*=0,\\
&l^*=D_w^*m^*,\\
&0\le\lambda^*\perp\left(b-Ax^*\right)\ge 0,\\
&\lim_{k\to\infty}\Phi\left(a,b,\mu_k\right)=\min\left(a,b\right)=0,
\end{aligned}\tag{4.4}$$

其中 D_y^*， D_w^* 定义为

$$D_y^*=diag\left(\nabla_{y_i}\phi\left(y_i^*,w_i^*,0\right),i=1,\cdots,m\right),$$

$$D_w^*=diag\left(\nabla_{w_i}\phi\left(y_i^*,w_i^*,0\right),i=1,\cdots,m\right).$$

由(2.2)和(3.11)有

$$0\le w^*\perp y^*\ge 0,$$

又 $z^*\in X_0$， 从而有 $z^*\in X$ 。

综上所述， z^* 是问题(1.1)的稳定点。

5. 数值实验

本文提出了求解均衡约束优化问题的算法，在这一节我们将对这个算法进行数值试验，利用 MATLAB 软件算法进行了数值实验，通过实验说明这个算法具有有效性和可行性.

算法中的矩阵 B_k 是 Lagrange 函数 $L(z,\mu)=f(z)+\sum_{j=1}^{m}\mu_j g_j(z,\mu)$ 的二阶 Hessian 阵的近似估计。矩阵 B_k 采用 BFGS 公式修正：

$$B_{k+1}=B_k-\frac{B_k\hat{s}_k\left(B_k\hat{s}_k\right)^{\mathrm{T}}}{\hat{s}_k^{\mathrm{T}}B_k\hat{s}_k}+\frac{\hat{\eta}_k\hat{\eta}_k^{\mathrm{T}}}{\hat{s}_k^{\mathrm{T}}\hat{\eta}_k},$$

其中 $\hat{s}_k = z^{k+1} - z^k, \hat{\eta}_k = \theta_k \hat{\gamma}_k + (1-\theta) B_k \hat{s}_k$，这里

$$\hat{\gamma}_k = \nabla L_z\left(z^k + \hat{s}_k, \mu_k\right) - \nabla L_z\left(z^k, \mu_k\right),$$

$$\theta = \begin{cases} 1, & \hat{s}_k \hat{\gamma}_k \geq 0.2 \hat{s}_k^{\mathrm{T}} B_k \hat{s}_k \\ \dfrac{0.8 \hat{s}_k^{\mathrm{T}} B_k \hat{s}_k}{\hat{s}_k^{\mathrm{T}} B_k \hat{s}_k - \hat{s}_k \hat{\gamma}_k}, & \text{否则}. \end{cases}$$

分别选取参数 $\xi = 10$，$\varepsilon_0 = 10$，B_0 为 $(n+2m)\times(n+2m)$ 的单位阵.

问题 1

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}x^2 + \frac{1}{2}xy - 95x \\ \text{s.t.} \quad & 0 \leq x \leq 200, \\ & 0 \leq w \perp y \geq 0. \end{aligned}$$

问题 2

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}\left[\left(x_1 + x_2 + y_1 - 15\right)^2 + \left(x_1 + x_2 + y_2 - 15\right)^2\right] \\ \text{s.t.} \quad & 0 \leq x \leq 10, \\ & w = Nx + My + q, \\ & 0 \leq w \perp y \geq 0, \\ & (x, y, w) \in R^2 \times R^2 \times R^2. \end{aligned}$$

其中

$$q = \begin{pmatrix} -36 \\ -25 \end{pmatrix}, M = \begin{pmatrix} 2 & 8/3 \\ 5/4 & 2 \end{pmatrix}, N = \begin{pmatrix} 8/3 & 2 \\ 2 & 5/4 \end{pmatrix}.$$

问题 3

$$\begin{aligned} \min \quad & f(x,y) = \frac{1}{2}\left(x_1 + 3y_1 - 4y_2 - 5\right)^2 + \left(x_2 - \frac{15}{8}y_1 + 3y_2 - 9\right)^2 \\ \text{s.t.} \quad & 0 \leq x_1 \leq 10, \\ & 0 \leq x_2 \leq 10, \\ & 0 \leq y \perp \left[\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^{\mathrm{T}} x + \begin{pmatrix} 3 & -\frac{15}{8} \\ -4 & 3 \end{pmatrix}^{\mathrm{T}} y + \begin{pmatrix} 10 \\ 6 \end{pmatrix}\right] \geq 0. \end{aligned}$$

问题 4

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}x^{\mathrm{T}}x + e^{\mathrm{T}}y \\ \text{s.t.} \quad & Ax \leq b, \\ & w = Nx + My + q, \\ & 0 \leq w \perp y \geq 0. \end{aligned}$$

其中，$M \in R^{m\times m}, N \in R^{m\times n}, A \in R^{p\times n}$ 分别为随机产生的矩阵，矩阵 M 严格对角占优的，是 P 矩阵，其非对角线上的非负向量 $q \in R^m, b \in R^p$ 是随机产生的，对于单位向量 $e \in R^n, A, b$ 满足 $Ae < b$。

表 1 和表 2 是算法 3.1 的数值实验结果。

Table 1. The numerical experiment results of problem 1 - problem 3
表 1. 问题 1~问题 3 的数值实验结果

NO	(x^0, y^0)	k	(x^*, y^*)	$f(x^*, y^*)$	$\|dz^k\|$
1	(0, 0)	13	(93.33333333333336, 26.66666666666664)	-3.266666666666669e+003	9.300000000000003e-012
2	(0, 0)	7	(5.00000000000006 9.99999999999994, -0.00000000000000 0.00000000000000)	1.262177448353619e-029	1.010000000000000e-008
3	(0, 0)	3	(5.00000000000000, 9.00000000000000, 0.00000000000000, 0.00000000000000)	0	2.500000000000000e-009

Table 2. The numerical experiment results of problem 4
表 2. 问题 4 的数值实验结果

NO.	(p, m, n)	初始点	迭代次数	最优值	$\|dz^k\|$
4	(30, 30, 50)	(*ones*, *zeros*)	48	6.376485162e-016	4.207400000e-8
4	(50, 60, 50)	(*ones*, *zeros*)	56	4.721785683e-014	1.621502153e-6
4	(100, 60, 60)	(*ones*, *zeros*)	72	8.190904715e-016	3.100586202e-6

在表 1 和表 2 中，NO.表示问题的编号，(p, m, n) 分别表示 x 的维数，m 表示 y 的维数，n 表示约束的个数。$\|dz^k\|$ 表示终止时，$\|dz^k\|$ 的对应值。数值实验表明本文的算法 3.1 是有效的。

基金项目

河池学院 2019 年高层次人才科研启动费项目“比例时滞神经网络的周期解与稳定状态研究”(2019GCC005)。

参考文献

[1] Yu, B., Mitchell, J.E. and Pang, J.-S. (2019) Solving Linear Programs with Complementarity Constraints Using Branch-and-Cut. *Mathematical Programming Computation*, **11**, 267-310. https://doi.org/10.1007/s12532-018-0149-2

[2] 吴学谦, 李声杰. 随机线性互补问题的无约束优化再定式[J]. 数学年刊 A 辑(中文版), 2019, 40(1): 43-54.

[3] Zang, I. (1980) A Smoothing-Out Technique for Min-Max Optimization. *Mathematical Programming*, **19**, 61-71. https://doi.org/10.1007/BF01581628

[4] Song, X. (2001) Smoothing Method for Minimax Problems. *Computational Optimization and Applications*, **20**, 267-279. https://doi.org/10.1023/A:1011211101714

[5] Luo, Z.Q., Pang, J.S. and Ralph, D. (1996) Mathematical Program with Equilibrium Constraints. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511983658

[6] 简金宝. 光滑约束优化快速算法[M]. 北京: 科学出版社, 2010.

知网检索的两种方式：

1. 打开知网首页：http://cnki.net/，点击页面中“外文资源总库 CNKI SCHOLAR”，跳转至：http://scholar.cnki.net/new，搜索框内直接输入文章标题，即可查询；
   或点击“高级检索”，下拉列表框选择：[ISSN]，输入期刊 ISSN：2324-7991，即可查询。
2. 通过知网首页 http://cnki.net/顶部“旧版入口”进入知网旧版：http://www.cnki.net/old/，左侧选择“国际文献总库”进入，搜索框直接输入文章标题，即可查询。

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：aam@hanspub.org