Pure Mathematics 理论数学, 2024, 14(12), 1-6
Published Online December 2024 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2024.1412399

二阶微分方程方法求解带不等式约束的优化问题

李思怡，姜　莹，宁文琪，任泓燃

沈阳航空航天大学理学院，辽宁　沈阳

收稿日期：2024年10月21日；录用日期：2024年11月22日；发布日期：2024年12月5日

摘　要

针对只含有不等式约束的优化问题，本文首先给出了其Karush-Kuhn-Tucker (KKT)条件，并利用光滑互补函数将KKT系统转化为一类光滑的方程组问题；其次，将光滑方程组问题转化为无约束优化问题；最后，本文提出一类二阶微分方程系统求解无约束优化问题，并讨论了二阶微分方程系统的解的稳定性及收敛速度。

关键词

二阶微分方程系统，不等式约束优化问题，KKT条件

Second-Order Differential Equation Method for Solving Optimization Problems with Inequality Constraints

Siyi Li, Ying Jiang, Wenqi Ning, Hongran Ren

School of Science, Shenyang Aerospace University, Shenyang Liaoning

Received: Oct. 21st, 2024; accepted: Nov. 22nd, 2024; published: Dec. 5th, 2024

Abstract

For optimization problems with only inequality constraints, this paper first presents their Karush-Kuhn-Tucker (KKT) conditions, and uses smooth complementarity functions to transform the KKT system into a class of smooth system of equations problems. Secondly, this article transforms the problem of smooth equation systems into an unconstrained optimization problem. Finally, this article proposes a class of second-order differential equation systems for solving unconstrained

文章引用: 李思怡, 姜莹, 宁文琪, 任泓燃. 二阶微分方程方法求解带不等式约束的优化问题[J]. 理论数学, 2024, 14(12): 1-6. DOI: 10.12677/pm.2024.1412399

optimization problems, and discusses the stability and convergence speed of the solutions of second-order differential equation systems.

Keywords

Second-Order Differential Equations, Optimization Problems with Inequality Constraints, KKT Conditions

1. 引言

在现实世界中，大多数实际问题都是包含约束条件的。例如：在建筑结构设计中，设计师需要在满足建筑安全标准、材料强度限制、建筑空间需求等多种约束条件下，优化结构的设计方案，以实现成本最低或结构性能最优[1] [2]；企业在制定生产计划时，需要考虑设备产能、原材料供应、人力成本、生产周期等约束条件，以实现生产效益的最大化；物流配送过程中，需要考虑车辆的载重限制、行驶路线的交通状况、配送时间要求等约束条件，以优化物流配送方案，降低运输成本等等[3] [4]。因此，约束优化问题在很多领域都有着广泛的应用，近年来备受学者们的关注[5]-[7]。

本文考虑带不等式约束的优化问题：

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x\in R^n \\ &\quad\ g(x)\le 0 \end{aligned} \tag{1}$$

其中 $f:R^n\to R$ 和 $g:R^n\to R^m$ 都是连续可微的，R^n 表示为 n 维实向量空间。

本文将应用一类二阶微分方程方法求解约束优化问题。早期工作 Arrow 与 Hurwicz [8]，应用微分方程研究了优化问题的最优性条件。Antipin [9]应用一阶微分方程方法研究了一类双约束变分不等式问题，并证明了算法的全局收敛性。Attouch [10] [11]应用二阶微分方程方法求解了凸优化问题。Evtushenko 与 Zhadan [12]-[14]对额二阶微分方程方法进行了系统的讨论和研究。微分方程方法的主要思想是，将原问题转化为无约束最小值问题，然后针对最小值问题构造相应的一阶或者二阶微分方程系统，讨论微分方程系统的解的稳定性。本文给出了一类带不等式约束的优化问题，应用一类光滑互补函数将带不等式约束的优化问题的 KKT 系统转化为最小值问题，从而构造出相应的二阶微分方程系统，为约束优化问题的方法研究提供思路。

2. 问题的转化

带不等式约束的优化问题(1)的 Karush-Kuhn-Tucker (KKT)条件为：

$$\begin{aligned} L(x,\lambda)=\nabla f(x)+\nabla g(x)\lambda=0 \\ -g(x)\perp\lambda \end{aligned} \tag{2}$$

其中，$L(x,\lambda)$ 是约束优化问题的 Lagrange 函数，$\nabla f(x)$ 和 $\nabla g(x)$ 分别是函数 f 和 g 的梯度，$\lambda\in R^m$ 是拉格朗日乘子。

为了将 KKT 条件(2)转化为光滑方程组问题，本文借助文献[15]中的光滑互补函数

$$\phi_{NR}^{\varepsilon}(a,b)=a-\varphi(\varepsilon,a-b)$$

其中

$$\varphi(\varepsilon,a-b)=\frac{1}{2}\left[a-b+\sqrt{\varepsilon^2+(a-b)^2}\right]$$

且 $\varepsilon\in R_+$， $a,b\in R_+$， $\phi:R\times R\to R$ 满足 $\phi(a,b)=0$ 当且仅当 $ab=0$。

由此，带不等式约束的优化问题(1)的 KKT 条件(2)可以转化下面光滑方程组问题

$$H(\varepsilon,x,\lambda)=\begin{pmatrix}\varepsilon\\ L(x,\lambda)\\ \phi_{NR}^{\varepsilon}(-g_1(x),\lambda_1)\\ \vdots\\ \phi_{NR}^{\varepsilon}(-g_m(x),\lambda_m)\end{pmatrix}=0 \tag{3}$$

其中， $g_i(x),\lambda_i\in R_+$ 分别是 $g(x),\lambda$ 的分量。

令

$$\begin{aligned}\theta(\varepsilon,x,\lambda)&=\frac{1}{2}H^{\mathrm{T}}(\varepsilon,x,\lambda)H(\varepsilon,x,\lambda)\\&=\frac{1}{2}\left\|H(\varepsilon,x,\lambda)\right\|^2\end{aligned}$$

则光滑方程组(3)与下面无约束优化问题(4)同解。

$$\min\theta(z)=\min\frac{1}{2}\left\|H(z)\right\|^2 \tag{4}$$

其中 $z=(\varepsilon,x,\lambda)$。

3. 二阶微分方程系统

本文应用具有阻尼惯性系数 $\gamma(t)$ 和时间缩放参数 $\beta(t)$ 的二阶微分方程系统(DIN-AVD) [11]

$$\ddot{z}(t)+\frac{\alpha}{t}\dot{z}(t)+\beta\nabla^2\theta(z(t))\dot{z}(t)+\nabla\theta(z(t))=0 \tag{5}$$

来求解无约束优化问题(4)，其中 $\gamma(t)$ 和 $\beta(t)$ 是 $[t_0,+\infty)$ 上非负连续函数。

从而，建立与无约束优化问题所对应的二阶微分方程系统

$$\begin{pmatrix}\ddot{\varepsilon}(t)\\ \ddot{x}(t)\\ \ddot{\lambda}(t)\end{pmatrix}+\frac{\alpha}{t}\begin{pmatrix}\dot{\varepsilon}(t)\\ \dot{x}(t)\\ \dot{\lambda}(t)\end{pmatrix}+\beta\nabla^2\theta\begin{pmatrix}\varepsilon(t)\\ x(t)\\ \lambda(t)\end{pmatrix}\begin{pmatrix}\dot{\varepsilon}(t)\\ \dot{x}(t)\\ \dot{\lambda}(t)\end{pmatrix}+\begin{pmatrix}\nabla_{\varepsilon}H^{\mathrm{T}}(z)H(z)\\ \nabla_{x}H^{\mathrm{T}}(z)H(z)\\ \nabla_{\lambda}H^{\mathrm{T}}(z)H(z)\end{pmatrix}=0 \tag{6}$$

下面，分析二阶微分方程系统(6)沿轨迹值的收敛速度。假设 $\alpha\ge 3$ 且 $z^*\in\arg\min\theta$。设 z 是(6)的解，则 $z(t)\to z^*(t\to\infty)$。对于 $\eta\in[2,\alpha-1]$ 我们定义函数 $\mathrm{E}_{\eta}(t):[t_0,+\infty)\to R$ 即

$$\begin{aligned}\mathrm{E}_{\eta}(t)=&\,t\left(t-\beta(\eta+2-\alpha)\right)\left(\theta(z(t))-\min\theta\right)+\frac{1}{2}\left\|\eta\left(z(t)-z^*\right)+t\dot{u}_{\beta}(t)\right\|^2\\&+\eta(\alpha-\eta-1)\frac{1}{2}\left\|z(t)-z^*\right\|^2\end{aligned} \tag{7}$$

其中， $\dot{u}_{\beta}(t)=\dot{z}(t)+\beta\nabla\theta(z(t))$。下面，讨论 $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_{\eta}(t)$ 的表达式。

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[t\big(t-\beta(\eta+2-\alpha)\big)\big(\theta(z(t))-\min\theta\big)\Big]=\big(2t-\beta(\eta+2-\alpha)\big)\big(\theta(z(t))-\min\theta\big)$$
$$+t\big(t-\beta(\eta+2-\alpha)\big)\big\langle\dot z(t),\nabla\theta(z(t))\big\rangle$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\big\|\eta(z(t)-z^*)+t\dot u_\beta(t)\big\|^2&=\big\langle\eta(z(t)-z^*)+t\dot u_\beta(t),\eta\dot z(t)+\dot u_\beta(t)+t\ddot u_\beta(t)\big\rangle\\&=\big\langle\eta(z(t)-z^*)+t\dot u_\beta(t),(\eta+1-\alpha)\dot z(t)-(t-\beta)\nabla\theta(z(t))\big\rangle\\&=\eta(\eta+1-\alpha)\big\langle z(t)-z^*,\dot z(t)\big\rangle-t(\alpha-\eta-1)\|\dot z(t)\|^2\\&\quad-\beta t(t-\beta)\|\nabla\theta(z(t))\|^2-\eta(t-\beta)\big\langle z(t)-z^*,\nabla\theta(z(t))\big\rangle\\&\quad-t\big(t-\beta(\eta+2-\alpha)\big)\big\langle\dot z(t),\nabla\theta(z(t))\big\rangle\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta(\alpha-\eta-1)\frac{1}{2}\|z(t)-z^*\|^2=\eta(\alpha-\eta-1)\big\langle z(t)-z^*,\dot z(t)\big\rangle$$

因此，

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\eta(t)&=\big(2t-\beta(\eta+2-\alpha)\big)\big(\theta(z(t))-\min\theta\big)-\eta(t-\beta)\big\langle z(t)-z^*,\nabla\theta(z(t))\big\rangle\\&\quad-t(\alpha-\eta-1)\|\dot z(t)\|^2-\beta t(t-\beta)\|\nabla\theta(z(t))\|^2\end{aligned}$$

由于 $\langle z(t)-z^*,\nabla\Psi(z(t))\rangle\geq\Psi(z(t))-\Psi(z^*)$，当 $t\geq\max\{t_0,\beta\}$，我们得到

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\eta(t)&\leq-\big((\eta-2)t-\beta(\alpha-2)\big)\big(\theta(z(t))-\min\theta\big)-t(\alpha-\eta-1)\|\dot z(t)\|^2\\&\quad-\beta t(t-\beta)\|\nabla\theta(z(t))\|^2\end{aligned}\tag{8}$$

由定义可知函数 E_η 是非负的。对于不等式(8)：如果 $\eta>2$，当 $t\geq t_1=\max\left\{t_0,\beta,\frac{\beta(\alpha-2)}{\eta-2}\right\}$ 时，有 $(\eta-2)t-\beta(\alpha-2)\geq0$。由 $\alpha-\eta-1\geq0$ 得 $\eta\leq\alpha-1$。因此有 $2<\eta\leq\alpha-1$ 即 $\alpha>3$。其极限情况 $\eta=2$ (结果为 $\alpha=3$)定理 3.1 也包含。当 $t\geq\beta$ 时有 $\beta t(t-\beta)\geq0$。综上可得，若 $\eta\in(2,\alpha-1]$，可推导出 E_η 在 $[t_1,+\infty)$ 上是非递增的且 $\lim_{t\to+\infty}\mathrm{E}_\eta(t)$ 存在。

定理 3.1 令 $\alpha\geq3$ 且 $\arg\min\theta\neq\varphi$，设 $z:[t_0,+\infty)\to R$ 是(6)的解，若 $\eta\in[2,\alpha-1]$，则函数

$$t\mapsto\left(\frac{t}{t-\beta}\right)^{\alpha-2}\mathrm{E}_\eta(t)$$

是非递增的且 $\lim_{t\to+\infty}\mathrm{E}_\eta(t)$ 存在。

证明：设 $t>\max\{t_0,\beta\}$，由(8)可得

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\eta(t)\leq\beta(\alpha-2)\big(\theta(z(t))-\min\theta\big)\tag{9}$$

在式(9)两边同时乘以 $t(t-\beta)$，由 $\lambda+2-\alpha\leq1$，得

$$\begin{aligned}t(t-\beta)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\eta(t)&\leq\beta(\alpha-2)t(t-\beta)\big(\theta(z(t))-\min\theta\big)\\&\leq\beta(\alpha-2)t\big(t-\beta(\eta+2-\alpha)\big)\big(\theta(z(t))-\min\theta\big)\\&\leq\beta(\alpha-2)\mathrm{E}_\eta(t)\end{aligned}\tag{10}$$

将(10)两边同时乘以 $t^{\alpha-3}(t-\beta)^{1-\alpha}$，得

$$\left(\frac{t}{t-\beta}\right)^{\alpha-2}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\eta(t)\le\beta(\alpha-2)\frac{t^{\alpha-3}}{(t-\beta)^{\alpha-1}}\mathrm{E}_\eta(t)$$

进一步得

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(\frac{t}{t-\beta}\right)^{\alpha-2}\mathrm{E}_\lambda(t)\right]\le 0$$

因此，$\left(\frac{t}{t-\beta}\right)^{\alpha-2}\mathrm{E}_\eta(t)$ 是非递增的。由 E_η 的非负性，当 $t\to+\infty$ 时，其极限存在。

定理 3.2 令 $\alpha>3$ 且 $\arg\min\theta\ne\phi$，设 $z:[t_0,+\infty)\to R$ 是(6)的解，则 z 有界。同时，设 $\eta\in[2,\alpha-1]$，$t_1=\max\{t_0,\beta\}$，对于所有的 $t\ge s>t_1$，我们有

$$\theta(z(t))-\min\theta\le\frac{1}{t^2}\left(\frac{s}{s-\beta}\right)^{\alpha-2}\mathrm{E}_\eta(s)=\mathrm{O}(t^{-2})$$

证明：取 $\eta\in[2,\alpha-1]$，由 E_η 的定义，有

$$\frac{1}{2}\left\|\lambda(z(t)-z^*)+t(\dot z(t)+\beta\nabla\theta(z(t)))\right\|^2\le\mathrm{E}_\eta(t)\tag{11}$$

根据定理 3.1，有 $\lim\limits_{t\to+\infty}\mathrm{E}_\eta(t)$ 存在，令 M 为 E_η 的上界。将(11)左侧的平方项展开，有

$$\begin{aligned}&\lambda^2\frac{1}{2}\left\|z(t)-z^*\right\|^2+\lambda t\left\langle z(t)-z^*,\dot z(t)\right\rangle\\&+\lambda t\left\langle z(t)-z^*,\beta\nabla\Psi(z(t))\right\rangle+t^2\frac{1}{2}\left\|\dot z(t)+\beta\nabla\Psi(z(t))\right\|^2\le M\end{aligned}\tag{12}$$

由此可以推导出

$$\eta\frac{1}{2}\left\|z(t)-z^*\right\|^2+t\left\langle z(t)-z^*,\dot z(t)\right\rangle\le\frac{M}{\eta}\tag{13}$$

令 $h(t)=\frac{1}{2}\left\|z(t)-z^*\right\|^2$，式(13)两边同时乘 $t^{\lambda-1}$，可以得到

$$\eta t^{\eta-1}h(t)+t^\eta\dot h(t)=\frac{\mathrm{d}}{\mathrm{d}t}\left(t^\eta h(t)\right)\le\frac{M}{\eta}t^{\eta-1}\tag{14}$$

将(14)两边从 t_1 到 $t>t_1$ 进行求定积分，得到

$$t^\eta h(t)-t_1^\eta h(t_1)\le\frac{M}{\eta^2}\left(t^\eta-t_1^\eta\right)$$

因此

$$h(t)\le h(t_1)+\frac{M}{\eta^2}$$

因此 h 是有界的。有 E_η 的定义，有

$$\theta(z(t))-\min\theta\le\frac{\mathrm{E}_\eta(t)}{t(t-\beta(\eta+2-\alpha))}\le\frac{\mathrm{E}_\eta(t)}{t(t-\beta)}$$

由定理 3.1，函数$\left(\frac{t}{t-\beta}\right)^{\alpha-2}\mathrm{E}_{\lambda}(t)$是非递增的。因此，当$t \geq s > t_1$时，可以推出

$$\begin{aligned}\theta(z(t))-\min\theta &\leq \frac{1}{t(t-\beta)}\left(\frac{t-\beta}{t}\right)^{\alpha-2}\left(\frac{s}{s-\beta}\right)^{\alpha-2}\mathrm{E}_{\eta}(s)\\ &\leq \frac{1}{t^2}\left(\frac{t-\beta}{t}\right)^{\alpha-3}\left(\frac{s}{s-\beta}\right)^{\alpha-2}\mathrm{E}_{\eta}(s)\\ &\leq \frac{1}{t^2}\left(\frac{s}{s-\beta}\right)^{\alpha-2}\mathrm{E}_{\eta}(s)\end{aligned}$$

由此定理 3.2 得证。

4. 结论

本文应用一类光滑互补函数(NR 函数)，将带不等式约束的优化问题的 KKT 条件转化为一个无约束优化问题。基于无约束优化问题，本文构造了一个二阶微分方程系统，说明了二阶微分方程系统的解收敛到无约束最优化问题，并通过讨论二阶微分方程系统的解稳定性理论，说明了其解的收敛速度。本文基于约束优化问题的理论研究，为约束优化问题的求解方法提供思路。

参考文献

[1] 蓝倜恩. 优化技术在土建结构设计中的应用[J]. 建筑结构学报, 1981, 2(6): 12-19.

[2] 孙芳垂. 建筑结构优化设计案例分析[M]. 北京: 中国建筑工业出版社, 2011.

[3] 常剑峰, 钟约先, 韩赞东. 制造系统中能力约束下的生产批量计划优化方法[J]. 清华大学学报: 自然科学版, 2004, 44(5): 605-608.

[4] 肖依永, 常文兵, 张人千. 能力与资源双重约束下的启发式组合生产计划研究[J]. 中国管理科学, 2008, 16(6): 33-40.

[5] 周珩. RICU 异常行为患者身体约束优化流程应用效果观察[J]. 护理学, 2022, 11(2): 213-217.

[6] 钱立泽. 树种算法改进及其求解实际约束优化问题应用[D]: [硕士学位论文]. 长春: 吉林财经大学, 2022.

[7] 黄文, 陈建恒, 郭媛媛, 等. 黎曼流形上的无约束优化及其应用[J]. 厦门大学学报(自然科学版), 2023, 62(6): 1012-1038.

[8] Arrow, K.J. and Hurwicz, L. (1956) Reduction of Constrained Maxima to Saddle-Point Problems. *Berkeley Symposium on Mathematical Statistics and Probability*, **3**, 1-20.

[9] Antipin, A.S. (2000) Solving Variational Inequalities with Coupling Constraints with the Use of Differential Equations. *Differential Equations*, **36**, 1587-1596. https://doi.org/10.1007/bf02757358

[10] Attouch, H., Chbani, Z. and Riahi, H. (2019) Fast Proximal Methods via Time Scaling of Damped Inertial Dynamics. *SIAM Journal on Optimization*, **29**, 2227-2256. https://doi.org/10.1137/18m1230207

[11] Attouch, H. and Cabot, A. (2017) Asymptotic Stabilization of Inertial Gradient Dynamics with Time-Dependent Viscosity. *Journal of Differential Equations*, **263**, 5412-5458. https://doi.org/10.1016/j.jde.2017.06.024

[12] Extushenko, Y. (1974) Two Numerical Methods of Solving Nonlinear Programming. *Doklady Mathematics*, **15**, 420-423.

[13] Evtushenko, Y.G. and Zhadan, V.G. (1973) Numerical Methods of Solving Some Operational Research Problems. *USSR Computational Mathematics and Mathematical Physics*, **13**, 56-77. https://doi.org/10.1016/0041-5553(73)90100-6

[14] Evtushenko, Y.G. and Zhadan, V.G. (1977) A Relaxation Method for Solving Problems of Non-Linear Programming. *USSR Computational Mathematics and Mathematical Physics*, **17**, 73-87. https://doi.org/10.1016/0041-5553(77)90105-7

[15] Sun, J., Fu, W., Alcantara, J.H. and Chen, J. (2021) A Neural Network Based on the Metric Projector for Solving SOCCVI Problem. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 2886-2900. https://doi.org/10.1109/tnnls.2020.3008661