第 28 卷第 1 期 四川理工学院学报(自然科学版) Vol. 28 No. 1
2015 年 2 月 Journal of Sichuan University of Science & Engineering(Natural Science Edition) Feb. 2015
文章编号:1673-1549(2015)01-0076-04 DOI:10.11863/j.suse.2015.01.18

椭圆型最优控制问题中的 $L^{1,2}$ - 方向稀疏

严春梅, 张 维

(成都理工大学管理科学学院, 成都 610059)

摘　要:介绍一种带有 $L^{1,2}$ - 方向稀疏项的椭圆型最优控制问题,分析条纹稀疏模式,从理论角度研究该问题的一阶最优性条件。为解决不可微控制问题,基于广义微分,提出一个半光滑牛顿方法,将问题在泛函空间中进行表示和分析,并具有局部超线性收敛率。

关键词:方向稀疏;非光滑正则化;半光滑牛顿

中图分类号:O29　　**文献标志码**:A

引　言

罚项是 L^1 类型的最优控制问题能产生稀疏解[1]。即控制函数在区域的某些部分为 0,因为在这些点,不必要应用控制。分析一个带有稀疏测度的椭圆型问题,稀疏测度能够促进条纹稀疏模式。考虑带有条纹稀疏模式的椭圆型的最优控制 (p) 问题:

$$\begin{cases} \min J(y,u):=\dfrac{1}{2}\|y-y_d\|_{L^2}^2+\dfrac{\alpha}{2}\|u\|_{L^2}^2+\beta\|u\|_{1,2} \\ (y,u)\in H_0^1(\Omega)\times L^2(\Omega) \\ s.t: Ay=u+f\in\Omega \\ a\leqslant u\leqslant b \end{cases}$$

$\Omega\in R^n$ 是一个有界区域,带有足够光滑的边界 $\Gamma=\partial\Omega$, $y_d,a,b\in L^2(\Omega)$,且在 Ω 上几乎处处有 $a<0<b$, α, $\beta\geqslant 0$。$A: H_0^1(\Omega)\to H^{-1}(\Omega)$ 是二阶线性椭圆型微分算子,并且 $\|\cdot\|_{L^2}$ 和 $\|\cdot\|_{1,2}$ 分别代表 $L^2(\Omega)$ 和 $L^1(L^2(\Omega))$ 范数,y 是状态,u 是控制变量,y_d 是理想状态。

非光滑项在偏微分方程(PDE)的最优问题中经常被用于反问题的正则化,如处理图像过程[2]。文献[3],L^1 - 正则化被使用在 PDE 的最优控制背景下。文献[4]分析了带有非方向稀疏项 $\beta\|u\|_{L^1(\Omega)}$ 的椭圆型最优控制问题。除稀疏项外,L^2 范数的平方是成本函数的一部分,它允许问题在 Hilbert 空间结构下被分析,用牛顿类型算法解决[5]。方向稀疏项在 2012 年被 Heraog R[6]第一次提出,从理论到数值的实现都具有很大的难度,本文在此基础上研究椭圆型问题的方向稀疏项,着重研究其理论和计算方法。

1　预备知识

空间 $\Omega\subset R^N$, $N\geqslant 2$ 是一个有界可测集,可以 $R^N=R^n\times R^{N-n}(1\leqslant n<N)$ 将其按坐标分割。故有集合:

$$\Omega_1=\{x_1\in R^N:\exists x_2\in R^{N-n},(x_1,x_2)\in\Omega\}$$

$$\Omega_2(x_1)=\{x_2\in R^{N-n}:(x_1,x_2)\in\Omega\}$$

Ω_1 可以看作 Ω 在 R^N 上的射影,$\Omega_2(x_1)$ 是 Ω 在 $x_1\in R^n$ 方向的横截面。方向稀疏项 $\|u\|_{1,2}$ 的一般形式为:

$$\|u\|_{1,2}:=\int_{\Omega_1}\left(\int_{\Omega_2(x_1)}u^2(x_1,x_2)\,\mathrm{d}x_2\right)^{1/2}\mathrm{d}x_1$$

引理 1[6]　$L^{1,2}$ 的对偶空间是 Bochner 类型空间 $L^{\infty,2}$。$L^{\infty,2}$ 是定义在 Ω 上的所有测度函数 φ 且

$$\|\varphi\|_{\infty,2}=\operatorname*{esssup}_{x_1\in\Omega}\left(\int_{\Omega_2(x_1)}\varphi^2(x_1,x_2)\,\mathrm{d}x_2\right)^{1/2}$$

是有限的。

引理 2[6]　$\|\cdot\|_{1,2}: L^{1,2}\to R$ 在 $u\in L^{1,2}(\Omega)$ 的次微分为:

$$\partial\|\cdot\|_{1,2}(u)=\begin{Bmatrix} v\in L^{\infty,2}(\Omega): v(x_1,\cdot)\in \\ \partial\|\cdot\|_2(u(x_1,\cdot)),\text{对所有的 } x_1\in\Omega \end{Bmatrix}$$

收稿日期:2014-08-09

作者简介:严春梅(1992-),女,四川宜宾人,硕士生,主要从事应用泛函分析方面的研究,(E-mail)yanchunmei199202@163.com

$L^2(\Omega_2(x_1))$ 范数的次微分为:

$$s \in \partial\|\cdot\|_2(w) \Leftrightarrow \begin{cases} \|s\|_2 \leqslant 1 \text{ 若 } w = 0, \text{在 } \Omega_2(x_1) \text{ 上} \\ s = \dfrac{w}{\|w\|^2} \text{ 其他} \end{cases}$$

2　一阶最优系统

用一个减少公式的问题替换 (p)。这个问题的减少仅仅涉及到控制变量 u,因为微分算子 A 存在逆映射 $A^{-1}: H^{-1}(\Omega) \to H_0^1(\Omega)$,则减少后的问题是$(\hat{p})$:

$$\begin{cases} \min \hat{J}(u): = \dfrac{1}{2}\|A^{-1}u + A^{-1}f - y_d\|_{L^2}^2 + \dfrac{\alpha}{2}\|u\|_{L_2}^2 + \beta\|u\|_{1,2} \\ u \in U_{ad}: = \{u \in L^2(\Omega): a \leqslant u \leqslant b\} \end{cases}$$

这是一个在 Hilbert 空间上凸的最优性问题。

引理 3　当 $\beta \geqslant 0, \alpha \geqslant 0$ 时,问题$(\hat{p})$ 有唯一解。

证明　当 $\beta \geqslant 0, \alpha > 0$ 时,$J(\hat{u})$ 是严格凸的,连续的。当 $\beta \geqslant 0, \alpha = 0$ 时,由于算子 A 是单射,则$J(\hat{u})$ 同样是严格凸的,而且集合 $U_{ad} \subset L^2_{(\Omega)}$ 是凸的,弱紧的,则问题 $\hat{p}$ 遵循标准参数[2],故 $\hat{p}$ 问题的解存在且唯一。

问题 $\hat{p}$ 的解 $\bar{u} \in U_{ad}$ 满足变分不等式:

$$(A^{-*}(A^{-1}\bar{u} + A^{-1}f - y_d) + \alpha\bar{u}, u - \bar{u}) + \varphi(u) - \varphi(\bar{u}) \geqslant 0 \tag{1}$$

对所有的 $u \in U_{ad}$ 都成立。其中,A^{-*} 代表A 的转置的逆,即 $A^{-*} = (A^*)^{-1}$,而且 $\varphi(v): = \beta\|v\|_{1,2}$。用 $\partial\varphi(\bar{u})$ 表示 φ 在 $\bar{u}$ 处的次微分。

引理 4　$\bar{u} \in L^2(\Omega)$ 是$(\hat{p})$ 的解,若满足

$$(u - \bar{u}, -\bar{p} + \alpha\bar{u} + \bar{\lambda}) \geqslant 0 \tag{2}$$

对所有的 $u \in U_{ad}$ 都成立,其中,伴随状态 $\bar{p}: = -A^{-*}(A^{-1}\bar{u} + A^{-1}f - y_d)$,次微分 $\bar{\lambda} \in \partial\varphi(\bar{u})$。

变分不等式(2)式几乎需要在所有的地方进行逐点讨论。但在文献[7]允许存在非负函数 $\bar{\lambda}_a$ 和 $\bar{\lambda}_b$ 在 U_{ad} 中起到不等式约束的拉格朗日乘数,而且估计微分 $\bar{\lambda} \in \partial\varphi(\bar{u})$ 涉及到 $\bar{u}$ 的符号,由此引进拉格朗日乘数 $\bar{\lambda}_a$ 和 $\bar{\lambda}_b$,则变分不等式(2)式变为:

$$-\bar{p} + \alpha\bar{u} + \bar{\lambda} + \bar{\lambda}_a - \bar{\lambda}_b = 0 \tag{3}$$

$$\bar{\lambda}_b \geqslant 0, b - \bar{u} \geqslant 0, \bar{\lambda}_b(b - \bar{u}) = 0 \tag{4}$$

$$\bar{\lambda}_a \geqslant 0, \bar{u} - a \geqslant 0, \bar{\lambda}_a(\bar{u} - a) = 0 \tag{5}$$

$$\bar{\lambda} = \beta \text{ 当} \{x \in \Omega: \bar{u} > 0\} \text{ 时} \tag{6}$$

$$|\bar{\lambda}| \leqslant \beta \text{ 当} \{x \in \Omega: \bar{u} = 0\} \text{ 时} \tag{7}$$

$$\bar{\lambda} = -\beta \text{ 当} \{x \in \Omega: \bar{u} < 0\} \text{ 时} \tag{8}$$

将拉格朗日乘子 $\bar{\lambda}, \bar{\lambda}_a, \bar{\lambda}_b$ 合成一个拉格朗日乘子 $\bar{\mu}$:

$$\bar{\mu}: = \bar{\lambda} - \bar{\lambda}_a + \bar{\lambda}_b \tag{9}$$

又因为 $a < 0 < b$,则 $\bar{\mu}, \bar{\lambda}, \bar{\lambda}_a, \bar{\lambda}_b$ 有关系式:

$$\begin{cases} \bar{\lambda} = \min(\beta, \max(-\beta, \bar{\mu})) \\ \bar{\lambda}_a = -\min(0, \bar{\mu} + \beta) \\ \bar{\lambda}_b = \max(0, \bar{\mu} - \beta) \end{cases} \tag{10}$$

用(9)式和(10)式替换(4) - (8)式,对于 $c > 0$,则替换后的系统为非光滑方程:

$$C(\bar{u}, \bar{\mu}): = \bar{u} - \max(0, \bar{u} + c(\bar{\mu} - \beta)) - \min(0, \bar{u} + c(\bar{\mu} + \beta)) + \max(0, (\bar{u} - b) + c(\bar{\mu} - \beta)) + \min(0, (\bar{u} - a) + c(\bar{\mu} + \beta)) = 0 \tag{11}$$

因此,最大、最小函数就可以被理解为逐点。

定理 1　$(\bar{y}, \bar{u}) \in H_0^1(\Omega) \times L^2(\Omega)$ 是(p) 的解,若$(\bar{p}, \bar{\mu}) \in H_0^1(\Omega) \times L^2(\Omega)$ 满足:

$$A\bar{y} - \bar{u} - \bar{f} = 0 \tag{12}$$

$$A^* p + \bar{y} - y_d = 0 \tag{13}$$

$$-\bar{p} + \alpha\bar{u} + \bar{\mu} = 0 \tag{14}$$

$$\bar{u} - \max(0, \bar{u} + c(\bar{\mu} - \beta) - \min(0, \bar{u} + c(\bar{\mu} + \beta)) + \max(0, (\bar{u} - b) + c(\bar{\mu} - \beta)) + \min(0, (\bar{u} - a) + c(\bar{\mu} + \beta)) = 0 \tag{15}$$

其中,$c > 0$,可以得到无方向稀疏项的(p) 的最优性条件,就是令 $\beta = 0$,则(12) - (14)式仍然保持不变,(15)式变为:

$$\bar{\mu} + \max(0, \bar{\mu} + c^{-1}(\bar{u} - b)) + \min(0, \bar{\mu} + c^{-1}(\bar{u} - a)) = 0$$

已经用于构建一种算法,应用于双边控制约束最优控制问题。

3　半光滑牛顿法

半光滑牛顿方法研究有限维函数空间,常用于最优控制问题。在确定条件下,局部超线性收敛甚至全局收敛性都可以被证明。

X, Y 是 Banach 空间,$D \subset X$ 是开的,$F: D \to Y$ 是非线性映射。映射F 是在开集 $U \subset D$ 中是广义可微的,若存在一个映射 $g: U \to L(X, Y)$,对每个 $x \in U$,都有:

$$\lim_{h \to 0} \frac{1}{\|h\|_X}\|F(x + h) - F(x) - g(x + h)h\|_Y = 0 \tag{16}$$

假设用牛顿迭代法发现半光滑映射 $F(x) = 0$ 的一个根 $\bar{x}$。则下面局部收敛性结果成立。

定理 2[8]　假设 $\bar{x} \in D$ 是 $F(x) = 0$ 的一个解,F 在 $\bar{x}$ 的一个开领域 U 内是半光滑的,且有广义微分 g。若对所有的 $x \in U$,$g(x)^{-1}$ 都存在 $\{\|g(x)^{-1}\|: x \in U\}$ 是有界的,初始值 $x_0 \in U$ 已知,则牛顿迭代 $x^{k+1} = x^k -$

$g(x^k)^{-1}F(x^k)$ 是定义明确的,如果 x^0 足够的接近 $\bar{x}$,则具有超线性收敛率。

用牛顿法求解(12)-(15)式的解,需要半光滑的逐点的最大和最小算子。它们的定义为

$F_{\max}, F_{\min}: L^r \to L^s(\Omega)$,

$F_{\max}(v) = \max(0,v)$ 和 $F_{\min}(v) = \min(0,v)$, $v \in L^r(\Omega)$, $1 \leqslant s \leqslant r \leqslant \infty$。

映射:

$$g_{\max}(v)(x) = \begin{cases} 1 \text{ 若 } v(x) \geqslant 0 \\ 0 \text{ 若 } v(x) < 0 \end{cases}$$

$$g_{\min}(v)(x) = \begin{cases} 1 \text{ 若 } v(x) \leqslant 0 \\ 0 \text{ 若 } v(x) > 0 \end{cases} \quad (17)$$

可以作为 $F_{\max}, F_{\min}$ 在 v 处的广义导数。

由(14)式可以推出 $\bar{\mu} = \bar{p} - \alpha\bar{u}$,代入(15)式中,并令 $c := \alpha^{-1}$,则有

$$\bar{u} - \alpha^{-1}\max(0,\bar{p}-\beta) - \alpha^{-1}\min(0,\bar{p}+\beta) + \alpha^{-1}\max(0,\bar{p}-\beta-\alpha b) + \alpha^{-1}\min(0,\bar{p}+\beta-\alpha a) = 0 \quad (18)$$

由于 c 的选择[9],只有 $\bar{p}$ 出现在逐点的最大和最小算子中,且 $\bar{p}$ 具有更多的规律性。为了更加明确,引进算子:

$S: = -A^{-*}A^{-1}: L^2(\Omega) \to H_0^1(\Omega)$

$h: = -A^{-*}(A^{-1}f - y_d) \in H_0^1(\Omega)$

则 $\bar{p}: = -A^{-*}(A^{-1}\bar{u} + A^{-1}f - y_d)$ 可以写成 $\bar{p} = S\bar{u} + h$。考虑映射 $T: L^{6,2}(\Omega) \to L^s(\Omega)$,

$$s \in \begin{cases} (2,\infty] \text{ 当 } n = 1 \text{ 时} \\ (2,\infty) \text{ 当 } n = 2 \text{ 时} \\ \left(2, \dfrac{2n}{(n-2)}\right] \text{当 } n \geqslant 3 \text{ 时} \end{cases} \quad (19)$$

定义 $Tu = p = Su + h$。严格的说,$Tu = I(Su + h)$,I 表示 $H_0^1(\Omega)$ 嵌入到 $L^s(\Omega)$ 中。可以知道 T 是定义明确的,且是连续的。由于它是仿射,所以它也是 *Frechet* 可微的。用 $T\bar{u}$ 替换(18)式中的 $\bar{p}$,定义

$F: L^2(\Omega) \to L^2(\Omega)$

$$F(u): = u - \alpha^{-1}\max(0,Tu-\beta) - \alpha^{-1}\min(0,Tu+\beta) + \alpha^{-1}\max(0,Tu-\beta-\alpha b) + \alpha^{-1}\min(0,Tu+\beta-\alpha a) \quad (20)$$

紧凑的形式 $F(u) = 0$ 表示最优系统(12)-(15)式。讨论函数 F 的广义可微性,并用牛顿迭代法求解 $F(u) = 0$,即为(p)问题的解。

定理3 如(20)式定义的函数 F 是广义可微的,广义导数为:

$$g(u)(v) = v - \alpha^{-1}\chi(J_- \cup J_+)(Sv) \quad (21)$$

其中,J_-, J_+ 是互不相交的集合:

$$J_- = \{x \in \Omega: \alpha a < Tu + \beta \leqslant 0 \text{ 在 } \Omega \text{ 上}\}$$

$$J_+ = \{x \in \Omega: 0 \leqslant Tu - \beta < \alpha b \text{ 在 } \Omega \text{ 上}\} \quad (22)$$

证明 因为放射算子 T 具有光滑性性质,则对于每一个 $u \in L^2(\Omega)$ 都有 $Tu \in L^s(\Omega)(s > 2)$。这表明映射

$F_1: L^2(\Omega) \to L^2(\Omega)$

$u \to \max(0,Tu-\beta) = \max(0,Su+h-\beta)$

是光滑的,此外,F_1 的广义导数为 $g_1(u)(v) = \chi_A(Sv)$,χ_A 代表集合 $A = \{x \in \Omega: Tu - \beta \geqslant 0 \text{ 在 } \Omega \text{ 上}\}$ 的特征函数。在(20)式中,最大和最小函数可以进行类似的讨论。可以得出整个函数 F 是广义可微的合并特征函数,可以得到 F 的广义导数为(21)式。证毕。

求解(p)问题具体算法(半光滑牛顿算法)。

(1)给出初始值 $u^0 \in L^2(\Omega)$ 并令 $k: = 0$。

(2)除非某些停止准则是满意的,否则计算广义导数 $g(u^k)$,导出 δu^k,

$$g(u^k)\delta u^k = -F(u^k) \quad (23)$$

更新 $u^{k+1}: = u^k + \delta u^k$,令 $k: = k + 1$,并回到第一步。

应用定理2的半光滑牛顿算法的结果是收敛的。

定理4 初始值 u^0 足够靠近(p)的解 $\bar{u}$,则半光滑牛顿法迭代的 u^k 在 $L^2(\Omega)$ 中超线性收敛于 $\bar{u}$。而且,相应的状态 y^k 在 $H_0^1(\Omega)$ 中超线性收敛于 $\bar{y}$。

使用的半光滑牛顿方法服从有效集的设置。与半光滑牛顿方法有关的"双重有效集"[10]和"非有效集策略"[11]已经被广泛的讨论和应用。

4 结束语

本文从理论上分析了方向稀疏椭圆型最优控制问题,该问题可以被基于广义微分的半光滑牛顿法解决,且具有超线性收敛率。进一步的研究可以用双重有效集法。方向稀疏项为 $L^{0,2}$ 类型的偏微分最优控制问题是非凸且高度非线性的,至今还有待解决。

参 考 文 献:

[1] Nikolova M.Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares[J].Multiscale Model,2005,4(3):960-991.

[2] Troltzsch F. Optimal Control of Partial Differential Equations:Theory, Methods and Application[M].Providence:American Mathematical Society,RI,2010.

[3] Clason C,Kunisch K.A duality-based approach to elliptic control problems in nonreflexive Banach space[J]. ESAIM Control Optimal,2011(17):243-266.

[4] Wachsmuth G,wachsmuth D.Cocergence and regularization results for optimal control problems with sparsity functional[J].ESAIM Control optimal,2011(17):858-886.

[5] Casas E,Herzog R,Wachsmuth G.Optimality conditions and error analysis of semilinear elliptic control problems with L1 cost functional[J].SIAM J,Optimal,2012,22(3):795-820.

[6] Herzog R,Stadler G,Wachsmuth G.Directional sparsity in optimal control of partial differential equations[J]. SIAM Journal on Control and Optimization, 2012,50(2):943-963.

[7] Troltzsch V F,Wiesbaden.Optimal Steuerung partieller Differential gleichungen[M].Wiesbaden:Vieweg,2005.

[8] Sun D,Han J.Newton and quasi-Newton methods for a class of nonsmooth equations and related problems[J]. SIAM J.Optimal,1997(7):463-480.

[9] Hintermüller M I,Kunisch K.The primal-dual active set strategy as a semi-smooth Newton method[J].SIAM J. Optimal,2003,13(3):865-888.

[10] Davis T A,Hager W W.A sparse proximal implementation of the LP dual active set algorithm[J].Math.Program-Mathematical Programming, 2008, 112 (2): 275-301.

[11] Ito K,Kunisch K.The primal-dual active set method for nonlinear optimal control problems with bilateral constraints[J].SIAM.Control.Optimal,2004,43(1):357-376.

$L^{1,2}$ -Directional Sparsity of Elliptic Optimal Control Problems

YAN Chunmei, *ZHANG Wei*

(School of Administrative Science, Chengdu University of Technology, Chengdu 610059, China)

Abstract: The elliptic optimal control problems with $L^{1,2}$- directional sparsity are introduced and the stripe sparse mode is analyzed. Emphatically, the first-order optimality conditions of the problem are studied from the theory angle. For solving the non-differentiable control problem, a semi-smooth Newton method based on the generalized differential is proposed, with which the problem can be stated and analyzed in the functional space and has local superlinear convergence rate.

Key words: directional sparsity; non-smooth regularization; semi-smooth Newton