Journal of Computational and Applied Mathematics 233 (2010) 1812–1820

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Error estimates of mixed finite element methods for quadratic optimal control problems[☆]

Xiaoqing Xing[a], Yanping Chen[a,*], Nianyu Yi[b]

[a] School of Mathematical Sciences, South China Normal University, Guangzhou 510631, PR China

[b] Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, PR China

ARTICLE INFO

Article history:
Received 3 August 2008

MSC:
65L10
65L12

Keywords:
Error estimates
Mixed finite element
Optimal control

ABSTRACT

In this paper, we investigate the error estimates for the solutions of optimal control problems by mixed finite element methods. The state and costate are approximated by Raviart–Thomas mixed finite element spaces of order k and the control is approximated by piecewise polynomials of order k. Under the special constraint set, we will show that the control variable can be smooth in the whole domain. We derive error estimates of optimal order both for the state variables and the control variable.

1. Introduction

Optimal control problems are playing an increasingly important role in modern life. They have various backgrounds in social, economic, scientific and engineering numerical simulation etc. Finite element approximation is widely applied though other methods are also used. A systematic introduction of finite element method for PDEs and optimal control can be found in, for example, [1–5].

For optimal control problem governed by elliptic equations, there are many works on error estimates and superconvergence, for a priori error estimate, see [6–12], for a posteriori error estimate, see [13–15], for the superconvergence results, see [16,17]. Note that all the above works aimed at standard finite element method.

In this work, we will use mixed finite element methods to deal with optimal control problem. When the objective functional contains the gradient of the state variable, mixed methods should be used for discretization of the state equation with which both the scalar variable and its flux variable can be approximated in the same accuracy. Though mixed finite element method has been widely used in engineering simulations, there does not seem to exist much work on theoretical analysis of mixed methods in computational optimal control.

For the constraint optimal control problem, the regularity of the optimal control is generally quite low. The goal of this paper is to investigate a priori error estimates for the elliptic optimal control problem with a special constraint set which will be specified later. We can see that with the special structure of this constraint set, the regularity of the optimal control can be raised to $H^2(\Omega)$.

☆ This work is supported by the Foundation for Talent Introduction of Guangdong Provincial University, Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2008), National Science Foundation of China, and the National Basic Research Program under the Grant 2005CB321703.

* Corresponding author.
E-mail addresses: xingxiaoqing@gmail.com (X. Xing), yanpingchen@scnu.edu.cn, ypchen@xtu.edu.cn (Y. Chen), yinianyu365109@126.com (N. Yi).

0377-0427/$ – see front matter Crown Copyright © 2009 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2009.09.018

In this paper, we consider the following distributed convex optimal control problem:

$$\min_{u\in K\subset L^2(\Omega)} \frac{1}{2}\{\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\} \tag{1.1}$$

$$-\mathrm{div}\,(A\,\mathbf{grad}\,y)=f+u,\quad x\in\ \Omega, \tag{1.2}$$

$$y=0,\quad x\in\partial\Omega, \tag{1.3}$$

where Ω is a bounded domain in R^2 and with a smooth boundary $\partial\Omega$. Here, K denotes the admissible set of the control variable, defined by

$$K=\left\{v\in L^2(\Omega):\int_\Omega v\geq 0\right\}. \tag{1.4}$$

The details will be specified later.

In this paper, we adopt the standard notation $W^{m,p}(\Omega)$ for Sobolev spaces on Ω with a norm $\|\cdot\|_{m,p}$ given by

$$\|\phi\|_{m,p}^p=\sum_{|\alpha|\leq m}\|D^\alpha\phi\|_{L^p(\Omega)}^p,$$

a semi-norm $|\cdot|_{m,p}$ given by

$$|\phi|_{m,p}^p=\sum_{|\alpha|=m}\|D^\alpha\phi\|_{L^p(\Omega)}^p.$$

We set $W_0^{m,p}(\Omega)=\{\phi\in W^{m,p}(\Omega):\phi|_{\partial\Omega}=0\}$. For $p=2$, we denote

$$H^m(\Omega)=W^{m,2}(\Omega),\qquad H_0^m(\Omega)=W_0^{m,2}(\Omega),$$

and

$$\|\cdot\|_m=\|\cdot\|_{m,2},\qquad \|\cdot\|=\|\cdot\|_{0,2}.$$

The paper is organized as follows: In Section 2, we study the mixed finite element approximation of the optimal control problems. In Section 3, we discuss the properties of the control variable and introduce some important projection operators. We give some intermediate error estimates in Section 4. The full error estimates for the state variables, costate variables and control variable are presented in Section 5.

2. Mixed formulation of optimal control problems

In this section, we study the mixed finite element approximation of the problem (1.1)–(1.3). First, the problem can be rewritten in the first order system as following:

$$\min_{u\in K\subset L^2(\Omega)} \frac{1}{2}\{\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\} \tag{2.1}$$

$$\mathrm{div}\,\boldsymbol{p}=f+u,\quad x\in\Omega, \tag{2.2}$$

$$\boldsymbol{p}=-A\,\mathbf{grad}\,y,\quad x\in\Omega, \tag{2.3}$$

with the boundary condition

$$y=0,\qquad x\in\partial\Omega. \tag{2.4}$$

Next, we introduce the costate elliptic equation

$$-\,\mathrm{div}\,(A(x)(\mathbf{grad}\,z+\boldsymbol{p}-\boldsymbol{p}_d))=y-y_d,\quad x\in\Omega, \tag{2.5}$$

with the boundary condition

$$z=0,\qquad x\in\partial\Omega. \tag{2.6}$$

For our purpose of this paper, we will make some assumptions:

(1) $A(x)=(a_{ij}(x))_{2\times 2}$ is a symmetric matrix with $a_{ij}(x)\in W^{1,\infty}(\Omega)$ and for any vector $\boldsymbol{X}\in R^2$, there is a constant $c>0$, such that

$$\boldsymbol{X}^tA\boldsymbol{X}\geq c\|\boldsymbol{X}\|_{R^2}^2. \tag{2.7}$$

(2) The given functions in the problem satisfy the following regularity:

$$y_d\in L^2(\Omega),\qquad \boldsymbol{p}_d\in(H^1(\Omega))^2. \tag{2.8}$$

Now, we give a weak formulation of the problem (2.1)–(2.4). Let

$$\boldsymbol{V}=H(\mathrm{div}\,;\Omega)=\{\boldsymbol{v}\in(L^2(\Omega))^2,\mathrm{div}\,\boldsymbol{v}\in L^2(\Omega)\},$$

endowed with the norm given by

$$\|\boldsymbol{v}\|_{\mathrm{div}}=(\|\boldsymbol{v}\|^2+\|\mathrm{div}\,\boldsymbol{v}\|^2)^{\frac{1}{2}},$$

and

$$W = L^2(\Omega).$$

Then, the mixed weak formulation for the problem (2.1)–(2.4) is to find $(\boldsymbol{p}, y, u) \in \boldsymbol{V} \times W \times K$ such that

$$\min_{u\in K} \frac{1}{2}\{\|\boldsymbol{p}-\boldsymbol{p}_d\|^2 + \|y-y_d\|^2 + \|u\|^2\} \tag{2.9}$$

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \operatorname{div}\boldsymbol{v}) = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.10}$$

$$(\operatorname{div}\boldsymbol{p}, w) = (f+u, w), \quad \forall w \in W, \tag{2.11}$$

where $(\cdot)$ denotes the inner product in $L^2(\Omega)$ or $(L^2(\Omega))^2$. It is well known that the convex control problem (2.9)–(2.11) has a unique solution $(\boldsymbol{p}, y, u)$, and that a triplet $(\boldsymbol{p}, y, u)$ is the solution of (2.9)–(2.11) if and only if there exists a costate $(\boldsymbol{q}, z) \in \boldsymbol{V} \times W$ such that $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \operatorname{div}\boldsymbol{v}) = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.12}$$

$$(\operatorname{div}\boldsymbol{p}, w) = (f+u, w), \quad \forall w \in W, \tag{2.13}$$

$$(A^{-1}\boldsymbol{q}, \boldsymbol{v}) - (z, \operatorname{div}\boldsymbol{v}) = -(\boldsymbol{p}-\boldsymbol{p}_d, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.14}$$

$$(\operatorname{div}\boldsymbol{q}, w) = (y-y_d, w), \quad \forall w \in W, \tag{2.15}$$

$$(z+u, \tilde{u}-u) \geq 0, \quad \forall \tilde{u} \in K. \tag{2.16}$$

In the following, we consider the mixed finite element approximation of the control problem. Let $\mathcal{T}^h$ denotes a quasi-uniform (in the sense of [12]) family of partition of Ω into triangles or rectangles, with boundary elements allowed to have one curved side. Here h is the maximum diameter of the element T in $\mathcal{T}^h$. Let $\boldsymbol{V}_h \times W_h \subset \boldsymbol{V} \times W$ denote the Raviart–Thomas space [18] of order k associated with the triangulations or rectangulations $\mathcal{T}^h$ of Ω. P_k denotes polynomials of total degree at most k, $Q_{m,n}$ indicates the space of polynomials of degree no more than m and n in x_1 and x_2 variables respectively, where $x = (x_1, x_2)$. If $T \in \mathcal{T}^h$ is a triangle, let

$$\boldsymbol{V}(T) = \boldsymbol{P}_k(T) \oplus \operatorname{span}(\boldsymbol{x}P_k(T)), \qquad W(T) = P_k(T).$$

Similarly, if $T \in \mathcal{T}^h$ is a rectangle, let

$$\boldsymbol{V}(T) = Q_{k+1,k}(T) \times Q_{k,k+1}(T), \qquad W(T) = P_k(T)$$

where $\boldsymbol{P}_k(T) = (P_k(T))^2$. Then we can define the finite dimensional spaces as follows

$$\boldsymbol{V}_h = \{\boldsymbol{v}_h \in \boldsymbol{V} : \boldsymbol{v}_h|_T \in \boldsymbol{V}(T),\ T \in \mathcal{T}^h\}, \tag{2.17}$$

$$W_h = \{w_h \in W : w_h|_T \in W(T),\ T \in \mathcal{T}^h\}, \tag{2.18}$$

$$K_h = \{\tilde{u}_h \in K : \tilde{u}_h|_T = P_k(T),\ T \in \mathcal{T}^h\}. \tag{2.19}$$

Then the finite element approximation of the problem (2.9)–(2.11) is to find $(\boldsymbol{p}_h, y_h, u_h) \in V_h \times W_h \times K_h$ such that

$$\min_{u_h\in K_h} \frac{1}{2}\{\|\boldsymbol{p}_h-\boldsymbol{p}_d\|^2 + \|y_h-y_d\|^2 + \|u_h\|^2\} \tag{2.20}$$

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \operatorname{div}\boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.21}$$

$$(\operatorname{div}\boldsymbol{p}_h, w_h) = (f+u_h, w_h), \qquad \forall w_h \in W_h. \tag{2.22}$$

The control problem (2.20)–(2.22) again has a unique solution $(\boldsymbol{p}_h, y_h, u_h)$ and that a triplet $(\boldsymbol{p}_h, y_h, u_h)$ is the solution of (2.20)–(2.22) if and only if there exists a costate $(\boldsymbol{q}_h, z_h) \in \boldsymbol{V}_h \times W_h$ such that $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \operatorname{div}\boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.23}$$

$$(\operatorname{div}\boldsymbol{p}_h, w_h) = (f+u_h, w_h), \quad \forall w_h \in W_h, \tag{2.24}$$

$$(A^{-1}\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \operatorname{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h-\boldsymbol{p}_d, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.25}$$

$$(\operatorname{div}\boldsymbol{q}_h, w_h) = (y_h-y_d, w_h), \quad \forall w_h \in W_h, \tag{2.26}$$

$$(z_h+u_h, \tilde{u}_h-u_h) \geq 0, \quad \forall \tilde{u}_h \in K_h. \tag{2.27}$$

In the rest of the paper, we shall use some intermediate variables. For any control function $\tilde{u} \in K$, we first define the state solution $(\boldsymbol{p}(\tilde{u}), y(\tilde{u}), \boldsymbol{q}(\tilde{u}), z(\tilde{u}))$ associated with $\tilde{u}$ that satisfies

$$(A^{-1}\boldsymbol{p}(\tilde{u}), \boldsymbol{v}) - (y(\tilde{u}), \operatorname{div}\boldsymbol{v}) = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.28}$$

$$(\operatorname{div}\boldsymbol{p}(\tilde{u}), w) = (f+\tilde{u}, w), \quad \forall w \in W, \tag{2.29}$$

$$(A^{-1}\boldsymbol{q}(\tilde{u}), \boldsymbol{v}) - (z(\tilde{u}), \operatorname{div}\boldsymbol{v}) = -(\boldsymbol{p}(\tilde{u})-\boldsymbol{p}_d, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.30}$$

$$(\operatorname{div}\boldsymbol{q}(\tilde{u}), w) = (y(\tilde{u})-y_d, w), \quad \forall w \in W. \tag{2.31}$$

Correspondingly, we define the discrete state solution $(\boldsymbol{p}_h(\tilde{u}), y_h(\tilde{u}), \boldsymbol{q}_h(\tilde{u}), z_h(\tilde{u}))$ associated with $\tilde{u} \in K$ that satisfies

$$(A^{-1}\boldsymbol{p}_h(\tilde{u}), \boldsymbol{v}_h) - (y_h(\tilde{u}), \operatorname{div} \boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.32}$$

$$(\operatorname{div} \boldsymbol{p}_h(\tilde{u}), w_h) = (f + \tilde{u}, w_h), \quad \forall w_h \in W_h, \tag{2.33}$$

$$(A^{-1}\boldsymbol{q}_h(\tilde{u}), \boldsymbol{v}_h) - (z_h(\tilde{u}), \operatorname{div} \boldsymbol{v}_h) = -(\boldsymbol{p}_h(\tilde{u}) - \boldsymbol{p}_d, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.34}$$

$$(\operatorname{div} \boldsymbol{q}_h(\tilde{u}), w_h) = (y_h(\tilde{u}) - y_d, w_h), \quad \forall w_h \in W_h. \tag{2.35}$$

Thus, for the sake of simplicity, we write the exact solution and its approximation in the following way:

$$(\boldsymbol{p}, y, \boldsymbol{q}, z) = (\boldsymbol{p}(u), y(u), \boldsymbol{q}(u), z(u)),$$
$$(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h) = (\boldsymbol{p}_h(u_h), y_h(u_h), \boldsymbol{q}_h(u_h), z_h(u_h)).$$

3. Properties of the control variable

Generally, the regularity of the optimal control for a constrained optimal control problem is quite low, say only in $H^1(\Omega)$, see [19]. For example, in our priori work, we have considered the obstacle type constraint set: $K = \{v : a \le v \le b\}$, where a and b are real numbers. For optimal control problem with this constraint set, we have the following relationship between the control variable u and the costate variable z:

$$u = \max(a, \min(b, -z(x))).$$

Thus, the gradient of u jumps along the boundary of the set of z where $z = a$ or $z = b$. Moreover, the location is generally unknown.

In this paper, we consider the special constraint set K defined as (1.4). We will show that the optimal control of the optimal control problem (2.12)–(2.16) can be infinitely smooth if the initial data are so. To this end, we first derive an important relationship between the optimal control and the optimal costate z. The following Theorem has been proved in Chen–Yi–Liu [20].

Theorem 3.1. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times K$ be the solution of (2.12)–(2.16). Then we have*

$$u = \max(0, \bar{z}) - z,$$

where $\bar{z} = \int_\Omega z/|\Omega|$ denotes the integral average on Ω of the function z.

Proof. If $\bar{z} > 0$, then $u = \bar{z} - z$ and for any $v \in K$

$$\begin{aligned}(u + z, v - u) &= \int_\Omega (u + z)(v - u) \\ &= \int_\Omega \bar{z}(v - \bar{z} + z) \\ &= \bar{z} \int_\Omega v \ge 0.\end{aligned}$$

If $\bar{z} \le 0$, then $u = -z$ and

$$(u + z, v - u) = 0.$$

Now, note that for the costate solution z the solution of (2.16) is unique. Then we have proved the theorem. □

Due to this theorem, we can obtain the following regularity result for the control variable.

Theorem 3.2. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ satisfy the optimality conditions (2.12)–(2.16). Assume that the data functions $\boldsymbol{p}_d, y_d$ and the domain Ω are infinitely smooth. Then,*

$$u \in C^\infty(\bar{\Omega}).$$

Proof. Applying the regularity argument of elliptic problems, it is clear that $y \in H^2(\Omega)$, so that $\boldsymbol{p} \in H^1(\Omega)$. It follows from the costate equation (2.5) and the assumption of $\boldsymbol{p}_d, y_d$, we can obtain that $z \in H^2(\Omega)$. With the relationship between the control and the costate

$$u = \max(0, \bar{z}) - z,$$

that $u \in H^2(\Omega)$. Thus $y \in H^4(\Omega)$, $\boldsymbol{p} \in H^3(\Omega)$. By repeating the above process, we can conclude that $u \in C^\infty(\bar{\Omega})$. □

Remark 3.1. If Ω is a convex open domain with a Lipschitz boundary $\partial\Omega$, and $\boldsymbol{p}_d \in (H^1(\Omega))^2$, $y_d \in L^2(\Omega)$, then we have $u \in H^2(\Omega)$.

4. Error estimates for the intermediate error

In this section, we will give some error estimates for the intermediate error. First of all, we define the standard $L^2(\Omega)$-orthogonal projection $P_h : W \to W_h$ which satisfies: for any $w \in W$

$$(w - P_h w, w_h) = 0, \quad \forall w_h \in W_h. \tag{4.1}$$

We also consider the Fortin projection [21,22] $\Pi_h : \boldsymbol{V} \to \boldsymbol{V}_h$, which satisfies: for any $\boldsymbol{q} \in \boldsymbol{V}$,

$$(\operatorname{div}(\boldsymbol{q} - \Pi_h \boldsymbol{q}), w_h) = 0, \quad \forall w_h \in W_h. \tag{4.2}$$

For the projection defined above, we have the following relations (see [21–23]):

$$\operatorname{div} \circ \Pi_h = P_h \circ \operatorname{div}, \tag{4.3}$$

$$\|\boldsymbol{q} - \Pi_h \boldsymbol{q}\|_0 \le Ch|\boldsymbol{q}|_1, \quad \text{for } \boldsymbol{q} \in (H^1(\Omega))^2, \tag{4.4}$$

$$\|\operatorname{div}(\boldsymbol{q} - \Pi_h \boldsymbol{q})\|_{-s} \le Ch^{1+s}|\operatorname{div}\boldsymbol{q}|_1, \quad s = 0, 1, \text{ for all } \operatorname{div}\boldsymbol{q} \in H^1(\Omega), \tag{4.5}$$

$$\|\phi - P_h\phi\|_{-s} \le Ch^{1+s}|\phi|_1, \quad s = 0, 1, \text{ for } \phi \in H^1(\Omega). \tag{4.6}$$

Now, if we choose $\tilde{u} = u$ in (2.32)–(2.35), we can obtain the intermediate solution $(\boldsymbol{p}_h(u), y_h(u), \boldsymbol{q}_h(u), z_h(u))$ following equations:

$$(A^{-1}\boldsymbol{p}_h(u), \boldsymbol{v}_h) - (y_h(u), \operatorname{div}\boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.7}$$

$$(\operatorname{div}\boldsymbol{p}_h(u), w_h) = (f + u, w_h), \quad \forall w_h \in W_h, \tag{4.8}$$

$$(A^{-1}\boldsymbol{q}_h(u), \boldsymbol{v}_h) - (z_h(u), \operatorname{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(u) - \boldsymbol{p}_d, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.9}$$

$$(\operatorname{div}\boldsymbol{q}_h(u), w_h) = (y_h(u) - y_d, w_h), \quad \forall w_h \in W_h. \tag{4.10}$$

Remark 4.1. Obviously, we can see that $(\boldsymbol{p}_h(u), y_h(u))$ is the mixed finite element approximation of the elliptic problem (2.12) and (2.13).

Now, we define some intermediate errors:

$$\eta_1 = \boldsymbol{p} - \boldsymbol{p}_h(u), \qquad \lambda_1 = y - y_h(u),$$
$$\eta_2 = \boldsymbol{q} - \boldsymbol{q}_h(u), \qquad \lambda_2 = z - z_h(u).$$

Then, from (2.12), (2.13), (4.7) and (4.8), we can obtain the following error equations:

$$(A^{-1}\eta_1, \boldsymbol{v}_h) - (\lambda_1, \operatorname{div}\boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h,$$
$$(\operatorname{div}\eta_1, w_h) = 0, \quad \forall w_h \in W_h.$$

From the classical mixed finite element error estimates in [22], we can establish the following results:

Lemma 4.1. *For h sufficiently small, there exists a constant C which only dependent on A and Ω, such that*

$$\|y - y_h(u)\| \le Ch^{k+1}\|y\|_{k+1+\delta_{k0}},$$
$$\|\boldsymbol{p} - \boldsymbol{p}_h(u)\| \le Ch^{k+1}\|y\|_{k+2},$$
$$\|\operatorname{div}(\boldsymbol{p} - \boldsymbol{p}_h(u))\| \le Ch^{k+1}\|y\|_{k+3},$$

where δ_{k0} is Dirac function.

For the costate variables, we derive the following error equations from (2.14), (2.15), (4.9) and (4.10):

$$(A^{-1}\eta_2, \boldsymbol{v}_h) - (\lambda_2, \operatorname{div}\boldsymbol{v}_h) = -(\eta_1, \boldsymbol{v}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h,$$
$$(\operatorname{div}\eta_2, w_h) = (\lambda_1, w_h), \quad \forall w_h \in W_h.$$

Using the standard stability result [21] of mixed finite element, we can easily deduce that

$$\|\boldsymbol{q} - \boldsymbol{q}_h(u)\|_{\operatorname{div}} + \|z - z_h(u)\| \le C(\|\boldsymbol{p} - \boldsymbol{p}_h(u)\|_{\operatorname{div}} + \|y - y_h(u)\|) \le Ch^{k+1}\|y\|_{k+3}, \tag{4.11}$$

where we have used Lemma 4.1.

5. Error estimates for optimal control problems

In this section, we consider the constraint optimal control problem where the convex set K is defined in (1.4). First, set the following errors:

$$r_1 = \boldsymbol{p}_h(u) - \boldsymbol{p}_h, \qquad e_1 = y_h(u) - y_h,$$
$$r_2 = \boldsymbol{q}_h(u) - \boldsymbol{q}_h, \qquad e_2 = z_h(u) - z_h.$$

Thus, we have

$$\boldsymbol{p} - \boldsymbol{p}_h = r_1 + \eta_1, \qquad y - y_h = e_1 + \lambda_1, \tag{5.1}$$
$$\boldsymbol{q} - \boldsymbol{q}_h = r_2 + \eta_2, \qquad z - z_h = e_2 + \lambda_2. \tag{5.2}$$

In order to estimate $\|\boldsymbol{p} - \boldsymbol{p}_h\|$, $\|y - y_h\|$, $\|\boldsymbol{q} - \boldsymbol{q}_h\|$, $\|z - z_h\|$, along with the result given in Section 4, we only need to estimate $\|r_1\|$, $\|e_1\|$, $\|r_2\|$, $\|e_2\|$.

From (2.23)–(2.26) and (4.7)–(4.10), we have that

$$(A^{-1}r_1, \boldsymbol{v}_h) - (e_1, \operatorname{div} \boldsymbol{v}_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{5.3}$$
$$(\operatorname{div} r_1, w_h) = (u - u_h, w_h), \quad \forall w_h \in W_h, \tag{5.4}$$
$$(A^{-1}r_2, \boldsymbol{v}_h) - (e_2, \operatorname{div} \boldsymbol{v}_h) = -(\boldsymbol{p}_h(u) - \boldsymbol{p}_h, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{5.5}$$
$$(\operatorname{div} r_2, w_h) = (y_h(u) - y_h, w_h), \quad \forall w_h \in W_h. \tag{5.6}$$

From the stability property of the saddle point problem (5.3)–(5.6), we have

$$\|r_1\|_{\text{div}} + \|e_1\| \le C\|u - u_h\|, \tag{5.7}$$
$$\|r_2\|_{\text{div}} + \|e_2\| \le C(\|\boldsymbol{p}_h(u) - \boldsymbol{p}_h\| + \|y_h(u) - y_h\|) \le C\|u - u_h\|. \tag{5.8}$$

So, we need to give the estimate for $\|u - u_h\|$.

Theorem 5.1. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times K$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h \times W_h)^2 \times K_h$ be the solution of (2.12)–(2.16) and (2.23)–(2.27) respectively. Then, we have*

$$\|u - u_h\| \le Ch^{k+1}(\|y\|_{k+3} + \|u\|_{k+1} + \|z\|_{k+1}), \tag{5.9}$$

and

$$\|y - y_h\| + \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\text{div}} \le Ch^{k+1}(\|y\|_{k+3} + \|u\|_{k+1} + \|z\|_{k+1}),$$
$$\|z - z_h\| + \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\text{div}} \le Ch^{k+1}(\|y\|_{k+3} + \|u\|_{k+1} + \|z\|_{k+1}).$$

Proof. From (2.27) and (2.32)–(2.35), for every $\tilde{u}_h \in K_h$, we have

$$\begin{aligned} 0 \ge\ & (u_h + z_h, u_h - \tilde{u}_h) = (u_h, u_h - \tilde{u}_h) + (z_h, u_h - \tilde{u}_h) \\ =\ & (u_h, u_h - \tilde{u}_h) + (\operatorname{div}(\boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)), z_h) \\ =\ & (u_h, u_h - \tilde{u}_h) + (A^{-1}\boldsymbol{q}_h, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)) \\ =\ & (u_h, u_h - \tilde{u}_h) + (A^{-1}(\boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)), \boldsymbol{q}_h) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)) \\ =\ & (u_h, u_h - \tilde{u}_h) + (y_h - y_h(\tilde{u}_h), \operatorname{div} \boldsymbol{q}_h) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)) \\ =\ & (u_h, u_h - \tilde{u}_h) + (y_h - y_d, y_h - y_h(\tilde{u}_h)) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)), \end{aligned} \tag{5.10}$$

that is

$$(u_h, u_h - \tilde{u}_h) + (y_h - y_d, y_h - y_h(\tilde{u}_h)) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(\tilde{u}_h)) \le 0, \quad \forall \tilde{u}_h \in K_h. \tag{5.11}$$

Next, the relation (2.25), (2.26), (2.32) and (2.33) imply that, for any $\tilde{u}_h \in K_h$,

$$\begin{aligned} (y_h - y_d, y_h(\tilde{u}_h)) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h(\tilde{u}_h)) &= (\operatorname{div} \boldsymbol{q}_h, y_h(\tilde{u}_h)) - (A^{-1}\boldsymbol{q}_h, \boldsymbol{p}_h(\tilde{u}_h)) + (z_h, \operatorname{div} \boldsymbol{p}_h(\tilde{u}_h)) \\ &= (\operatorname{div} \boldsymbol{q}_h, y_h(\tilde{u}_h)) - (A^{-1}\boldsymbol{p}_h(\tilde{u}_h), \boldsymbol{q}_h) + (z_h, \operatorname{div} \boldsymbol{p}_h(\tilde{u}_h)) \\ &= (\tilde{u}_h, z_h), \end{aligned}$$

thus,

$$(y_h - y_d, y_h(\tilde{u}_h) - y_h(u)) + (\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{p}_h(\tilde{u}_h) - \boldsymbol{p}_h(u)) = (z_h, \tilde{u}_h - u). \tag{5.12}$$

Similarly,

$$(y_h(u) - y_d, y_h - y_h(u)) + (\boldsymbol{p}_h(u) - \boldsymbol{p}_d, \boldsymbol{p}_h - \boldsymbol{p}_h(u)) = (z_h(u), u_h - u). \tag{5.13}$$

Finally, we observe that

$$\begin{aligned}\|u-u_h\|^2 &\le \|u-u_h\|^2+\|\boldsymbol{p}_h-\boldsymbol{p}_h(u)\|^2+\|y_h-y_h(u)\|^2\\ &= (u,u-u_h)+(\boldsymbol{p}_h-\boldsymbol{p}_d,\boldsymbol{p}_h-\boldsymbol{p}_h(u))+(y_h-y_d,y_h-y_h(u))-(u_h,u-u_h)\\ &\quad -(\boldsymbol{p}_h(u)-\boldsymbol{p}_d,\boldsymbol{p}_h-\boldsymbol{p}_h(u))-(y_h(u)-y_d,y_h-y_h(u))\\ &= (u_h,u_h-\tilde{u}_h)+(\boldsymbol{p}_h-\boldsymbol{p}_d,\boldsymbol{p}_h-\boldsymbol{p}_h(\tilde{u}_h))+(y_h-y_d,y_h-y_h(\tilde{u}_h))\\ &\quad +(u_h,\tilde{u}_h-u)+(\boldsymbol{p}_h-\boldsymbol{p}_d,\boldsymbol{p}_h(\tilde{u}_h)-\boldsymbol{p}_h(u))+(y_h-y_d,y_h(\tilde{u}_h)-y_h(u))-(u,u_h-u)\\ &\quad -(\boldsymbol{p}_h(u)-\boldsymbol{p}_d,\boldsymbol{p}_h-\boldsymbol{p}_h(u))-(y_h(u)-y_d,y_h-y_h(u)),\end{aligned}\tag{5.14}$$

using (5.11)–(5.13),

$$\begin{aligned}\|u-u_h\|^2 &\le (u_h,\tilde{u}_h-u)+(\boldsymbol{p}_h-\boldsymbol{p}_d,\boldsymbol{p}_h(\tilde{u}_h)-\boldsymbol{p}_h(u))+(y_h-y_d,y_h(\tilde{u}_h)-y_h(u))\\ &\quad -(u,u_h-u)-(\boldsymbol{p}_h(u)-\boldsymbol{p}_d,\boldsymbol{p}_h-\boldsymbol{p}_h(u))-(y_h(u)-y_d,y_h-y_h(u))\\ &= (u_h+z_h,\tilde{u}_h-u)-(u+z_h(u),u_h-u).\end{aligned}$$

Let $\tilde{u}_h=u_h$, then for every $w_h\in K_h$, using (2.16) and (2.27),

$$\begin{aligned}\|u-u_h\|^2 &\le (u_h+z_h,u_h-u)-(u+z_h(u),u_h-u)\\ &= (u_h+z_h,u_h-w_h)+(u_h+z_h,w_h-u)+(u+z,u-u_h)+(z-z_h(u),u_h-u)\\ &\le (u_h+z_h,w_h-u)+(z-z_h(u),u_h-u).\end{aligned}\tag{5.15}$$

Let $w_h=P_hu$, we will prove that $P_hu\in K_h$, where P_h is the L^2-orthogonal projection operator defined in (4.1). Note that

$$(u-P_hu,e_h)=0,\quad \forall e_h\in W_h,$$

especially letting $e_h=1\in W_h$ we have

$$(u-P_hu,e_h)=\int_\Omega (u-P_hu)=0,$$

thus

$$\int_\Omega P_hu=\int_\Omega u\ge 0.$$

Then $P_hu\in K_h$. So, if we choose $w_h=P_hu$ in (5.15), we can obtain that

$$\begin{aligned}(u_h+z_h,P_hu-u) &= (u_h-u,P_hu-u)+(z_h-z_h(u),P_hu-u)+(z_h(u)-z,P_hu-u)+(u+z,P_hu-u)\\ &\le \delta(\|u_h-u\|^2+\|z_h-z_h(u)\|^2+\|z_h(u)-z\|^2)C\|u-P_hu\|^2+(u+z,P_hu-u),\end{aligned}\tag{5.16}$$

where δ is small enough. For the last term of (5.16), we have

$$\begin{aligned}(u+z,P_hu-u) &= (u+z-P_h(u+z),P_hu-u)\\ &\le Ch^{2(k+1)}(\|u\|_{k+1}^2+\|z\|_{k+1}^2),\end{aligned}\tag{5.17}$$

where we use the estimate

$$\|u-P_hu\|\le Ch^{k+1}\|u\|_{k+1}.\tag{5.18}$$

From (4.11), (5.8) and (5.15)–(5.18) we have

$$\|u-u_h\|^2\le \delta\|u-u_h\|^2+Ch^{2(k+1)}\|y\|_{k+3}^2+Ch^{2(k+1)}(\|u\|_{k+1}^2+\|z\|_{k+1}^2).\tag{5.19}$$

It is easy to obtain that

$$\|u-u_h\|\le Ch^{k+1}(\|y\|_{k+3}+\|u\|_{k+1}+\|z\|_{k+1}).\tag{5.20}$$

Next, we will give the error estimates for the state variables and costate variables. By using Lemma 4.1, (4.11), (5.7), (5.8) and (5.20), we get that

$$\begin{aligned}\|y-y_h\|+\|\boldsymbol{p}-\boldsymbol{p}_h\|_{\mathrm{div}} &\le \|y-y_h(u)\|+\|y_h(u)-y_h\|+\|\boldsymbol{p}-\boldsymbol{p}_h(u)\|_{\mathrm{div}}+\|\boldsymbol{p}_h(u)-\boldsymbol{p}_h\|_{\mathrm{div}}\\ &\le Ch^{k+1}\|y\|_{k+1+\delta_{k0}}+Ch^{k+1}\|y\|_{k+2}+C\|u-u_h\|\\ &\le Ch^{k+1}(\|y\|_{k+3}+\|u\|_{k+1}+\|z\|_{k+1}),\\ \|z-z_h\|+\|\boldsymbol{q}-\boldsymbol{q}_h\|_{\mathrm{div}} &\le \|z-z_h(u)\|+\|z_h(u)-z_h\|+\|\boldsymbol{q}-\boldsymbol{q}_h(u)\|_{\mathrm{div}}+\|\boldsymbol{q}_h(u)-\boldsymbol{q}_h\|_{\mathrm{div}}\\ &\le Ch^{k+1}\|y\|_{k+3}+C\|u-u_h\|\\ &\le Ch^{k+1}(\|y\|_{k+3}+\|u\|_{k+1}+\|z\|_{k+1}).\end{aligned}$$

Thus, the theorem has been proved. □

Table 1
L^2 error estimates between the exact solutions and the approximations.

Resolution	$\|u-u_h\|$	$\|y-y_h\|$	$\|z-z_h\|$	$\|\boldsymbol{p}-\boldsymbol{p}_h\|$	$\|\boldsymbol{q}-\boldsymbol{q}_h\|$
16×16	2.454e−02	1.243e−03	2.454e−02	3.512e−03	3.466e−02
32×32	6.139e−03	3.110e−04	6.139e−03	8.800e−04	8.685e−03
64×64	1.535e−03	7.776e−05	1.535e−03	2.203e−04	2.174e−03
128×128	3.838e−04	1.944e−05	3.838e−04	5.510e−05	5.438e−04

6. Numerical tests

In this section, we will present an example to test our theoretical results. The problem we will consider is:

$$\min_{u\in K}\frac{1}{2}\{\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\} \tag{6.1}$$

$$\operatorname{div}\boldsymbol{p}+a_0y=u+f,\qquad \boldsymbol{p}=-\mathbf{grad}\,y,\quad x\in\Omega, \tag{6.2}$$

$$y=0,\quad x\in\partial\Omega \tag{6.3}$$

and the admissible set is:

$$K=\left\{u\in L^2(\Omega):\int_\Omega u(x)\mathrm{d}x\geqslant 0\right\}. \tag{6.4}$$

We make $\Omega=[0,1]\times[0,1]$ and $a_0=0$, then the state equation can be written as:

$$\operatorname{div}\boldsymbol{p}=u+f,\qquad \boldsymbol{p}=-\mathbf{grad}\,y,\quad x\in\Omega, \tag{6.5}$$

and the costate elliptic equation is:

$$\operatorname{div}\boldsymbol{q}=y-y_d,\qquad \boldsymbol{q}=-(\mathbf{grad}\,z+\boldsymbol{p}-\boldsymbol{p}_d),\quad x\in\Omega, \tag{6.6}$$

with the boundary condition

$$z=0,\quad x\in\partial\Omega. \tag{6.7}$$

Now, we choose the state function

$$y(x_1,x_2)=\sin(\pi x_1)\sin(\pi x_2)$$

and

$$\boldsymbol{p}_d=(-\pi(1+\pi^2)\cos(\pi x_1)\sin(\pi x_2),-\pi(1+\pi^2)\sin(\pi x_1)\cos(\pi x_2)).$$

For simplicity, we can let $f=0$. The costate function can be chosen as:

$$z(x_1,x_2)=-2\pi^2\sin(\pi x_1)\sin(\pi x_2).$$

Now, we can easily obtain that $\int_\Omega z=-8$. From the projection

$$u=\max(0,\bar{z})-z,$$

we have $\int_\Omega u=8\geq 0$.

From (6.6), we can obtain that

$$-\operatorname{div}(\mathbf{grad}\,z+\boldsymbol{p}-\boldsymbol{p}_d)=-\operatorname{div}(\mathbf{grad}\,z)-\operatorname{div}\boldsymbol{p}-\operatorname{div}\boldsymbol{p}_d=y-y_d. \tag{6.8}$$

From the known functions y, z, and $\boldsymbol{p}_d$, we can get that

$$y_d(x_1,x_2)=(1+2\pi^4)\sin(\pi x_1)\sin(\pi x_2).$$

In the following, without loss of generality, we use the order $k=1$ Raviart–Thomas mixed finite element spaces to approximate the state and costate variables, use piecewise linear function to approximate the control variable.

Table 1 shows the error data of L^2 norm both for state and costate variables and control variable with triangulation. Fig. 1 shows the numerical solution of control function u in 64×64 mesh grid while the state and costate were approximated by RT1.

7. Conclusion and future works

In this paper, we give a complete estimate for control variable, state variables and the adjoint variables of optimal control problem (1.1), (1.3) and (1.4) using mixed finite element methods. Our L^2-error estimates for the special control constraint set by mixed methods seem to be new. We have used piecewise constant functions to approximate the control variable. In our future work, we shall use the standard linear element space to approximate the control function. Furthermore, we shall consider the optimal boundary control problem and parabolic optimal control problem.

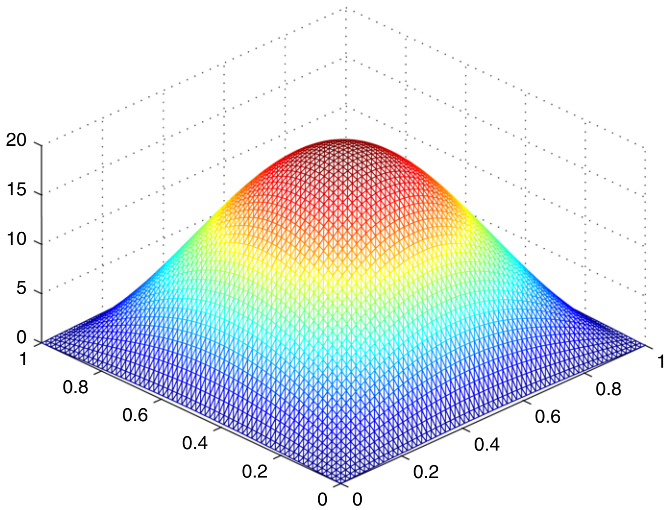

Fig. 1. The profile of numerical approximation of the control function with triangulation.

Acknowledgment

The author would like to thank Yongquan Dai for assistance in the numerical work.

References

[1] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[2] J. Haslinger, P. Neittaanmaki, Finite Element Approximation for Optimal Shape Design, John Wiley and Sons, Chichester, UK, 1989.
[3] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.
[4] P. Neittaanmaki, D. Tiba, Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications, M. Dekker, New York, 1994.
[5] D. Tiba, Lectures on the Optimal Control of Elliptic Equations, University of Jyvaskyla Press, Jyvaskyla, Finland, 1995.
[6] W. Alt, On the approximation of infinite optimization problems with an application to optimal control problems, Appl. Math. Optim. 12 (1984) 15–27.
[7] W. Alt, U. Machenroth, Convergence of finite element approximations to state constrained convex parabolic boundary control problems, SIAM J. Control Optim. 27 (1989) 718–736.
[8] N. Arada, E. Casas, F. Tröltzsch, Error estimates for a semilinear elliptic control problem, Comput. Optim. Appl. 23 (2002) 201–229.
[9] E. Casas, Optimal control in coefficients of elliptic equations with state constraints, Appl. Math. Optim. 26 (1992) 21–37.
[10] F.S. Falk, Approximation of a class of optimal control problems with order of convergence estimates, J. Math. Anal. Appl. 44 (1973) 28–47.
[11] T. Geveci, On the approximation of the solution of an optimal control problem governed by an elliptic equation, RAIRO Anal. Numer. 13 (1979) 313–328.
[12] V. Girault, P.A. Raviart, Finite Element Methods for Navier–Stokes equations: Theory and Algorithms, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986.
[13] W.B. Liu, N.N. Yan, A posteriori error estimates for optimal boundary control, SIAM J. Numer. Anal. 39 (2001) 73–99.
[14] W.B. Liu, N.N. Yan, A posteriori error estimates for control problems governed by Stokes' equations, SIAM J. Numer. Anal. 40 (2003) 1850–1869.
[15] W.B. Liu, N.N. Yan, A posteriori error estimates for distributed convex optimal control problems, Adv. Comput. Math. 15 (2001) 285–309.
[16] C. Meyer, A. Rösch, Superconvergence properties of optimal control problems, SIAM J. Control Optim. 43 (3) (2004) 970–985.
[17] N. Yan, Superconvergence and recovery type a posteriori error estimate for constrained convex optimal control problems, in: Y. Lu, W. Sun, T. Tang (Eds.), Advances in Scientific Computing and Applications, Science Press, Beijing, New York, 2004, pp. 408–419.
[18] P.A. Raviart, J.M. Thomas, A mixed finite element method for 2nd order elliptic problems, in: Math. Aspecs of the Finite Element Method, in: Lecture Notes in Math., vol. 606, Springer, Berlin, 1977, pp. 292–315.
[19] R. Li, W.B. Liu, H. Ma, T. Tang, Adaptive finite element approximation for distributed elliptic optimal control problems, SIAM J. Control Optim. 41 (2002) 1321–1349.
[20] Y. Chen, N. Yi, W.B. Liu, A legendre galerkin spectral method for optimal control problems governed by elliptic equations, SIAM J. Numer. Anal. 46 (5) (2008) 2254–2275.
[21] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, 1991.
[22] J. Douglas Jr, J.E. Roberts, Global estimates for mixed finite element methods for second order elliptic equations, Math. Comp. 44 (1985) 39–52.
[23] Y. Kwon, F.A. Milner, L^∞-error estimates for mixed methods for semilinear second-order elliptic equations, SIAM J. Numer. Anal. 25 (1988) 46–53.