Journal of Mathematical Research & Exposition
Vol.24, No.2, 257–266, May, 2004

Optimality Condition for Infinite-dimensional Programming Problem with Operator and Bound Constraints *

LI Chun-fa[1], *FENG En-min*[2]

(1. Sch. of Econ. & Man., Tianjin University of Technology, Tianjin 300191, China;
2. Dept. of Math., Dalian University of Technology, Liaoning 116024, China)

Abstract: A first-order necessary condition for an infinite-dimensional nonlinear optimization problem, which arises when the all-at-once method is employed to slove the optimal control problems, is formulated and analyzed. Operator constraint and simple bound on part of the variables are both considered. Based on this optimality condition, the trust-region subproblems are built, then the trust region method may be employed to deal with the optimization problem in infinite-dimensional space.

Key words: optimal control; operator constraint; trust-region method.

Classification: AMS(2000) 49K15, 49K27/CLC number: O232

Document code: A **Article ID:** 1000-341X(2004)02-0257-10

1. Introduction

Optimal control theory has been intensively applied to many fields. This kind of problem can be referred to as the infinite-dimensional optimization problem. It is well known that the necessary conditions are the basis for solving the nonlinear programming problems by using the gradient methods. The corresponding necessary conditions can be found in [1], for problems with only inequality operator constraint, and in [2] for problems with only equality operator constraint; Maurer[3] presented the general abstract optimality conditions for infinite dimensional programming problems with cone constraints. Based on the idea introduced by Coleman and Li[4], Ulbrich[5] gave a KKT necessary optimality condition of a special infinite dimensional optimization problem, which arises when the black-box approach is applied to optimal control problems with bound-constrained on the controls.

The present work is motivated by the application of all-at-once approach to optimal control problems with bounds on the controls, then it is possible to solve the above infinite-dimensional nonlinear programming problem using the trust-region algorithm based on the present necessary condition.

*Received date: 2002-12-30
Biography: LI Chun-fa (1968-), Ph.D., Associate Professor.

The paper is organized as follows. In the next section we describe constrained problem considered and the assumptions used. In Section 3, we present the classical necessary optimality conditions, and in terms of the operator theories and the introduction of the scaling functionals, we reformulate the necessary optimality condition in the infinite framework; In Section 4, the trust-region subproblems are built.

2. Problem formulation and assumptions

Optimal control problem can be formulated as the following abstract optimization problem with special structure in infinite dimensional space.

$$\begin{array}{ll} \min & f(y,u) \\ \text{s.t.} & e(y,u)=0 \\ & u \in \mathcal{U}_{ad}, \end{array} \tag{1}$$

where $\mathcal{U}_{ad} = \{u \mid a(x) \leq u(x) \leq b(x), x \in \Omega\}, \Omega \subset R^n$ is a domain with positive and finite Lebesgue measure $0 < \mu(\Omega) < \infty$. f is a real functional on $Z = Y \times U$, e is an operator from Z to W. Y, U and W are Hilbert spaces. $e(y,u) = 0$ is an operator equation corresponding to the state equation in the associated optimal control problem, the variables y, u are the state and the control respectively.

We adopt the following notations throughout: $z = (y,u) \in Y \times U = Z$; $\langle \cdot, \cdot \rangle$ stands for the inner product in a Hilbert space; $f'(z)$ and $e'(z)$ denote the Fréchet derivative of f and e respectively. ∇ denotes the gradient operator. Then for a fixed $z_0 \in Z, \nabla f(z_0)$ is a linear functional on Z, i.e. $\nabla f(z_0) \in Z^\star, Z^\star$ is the dual space of Z, but since Z is a Hilbert space, thus $Z^\star = Z; e'(z_0)$ is a linear operator from Z to W; $\langle e(z), e(z) \rangle$ is a real functional on Z, and for all $w \in W$, $\langle w, e'(z_0)(\cdot) \rangle$ is also a functional on Z, and when applied to some $z \in Z$, take the value $\langle w, e'(z_0)z \rangle$. For normed linear spaces E and F, we let $\mathcal{L}(E,F)$ denote the space of linear and bounded operators from E into F. For simplicity, $\mathcal{L}(\mathcal{E},\mathcal{E})$ will be abbreviated to $\mathcal{L}(\mathcal{E})$. For $A \in \mathcal{L}(\mathcal{E},\mathcal{F})$ we use the symbols $\mathcal{R}(A)$ and $\mathcal{N}(\mathcal{A})$ for the range space and the null space of A respectively, and the operator $A^\star \in \mathcal{L}(F^\star, E^\star)$ denotes the adjoint of A. Following the above notations, then $e'(z_0)z \in W$ for all $z \in Z$, and $e'(z_0) \in \mathcal{L}(Z,W), \nabla f(z_0) \in Z^\star$.

When the black-box approach is applied to optimal control problem with bound constrained on u, problem (1) can be reformulated as the following reduced form:

$$\begin{array}{ll} \min & f(y(u),u) \\ \text{s.t.} & u \in \mathcal{U}_{ad}. \end{array} \tag{2}$$

In [5], Ulbrich investigated this problem, and the optimality conditions, the trust-region interior-point algorithm were presented to solve this problem based on [4].

Now we introduce the following functional

$$L(y,u,\lambda) = f(y,u) + \langle \lambda, e(y,u) \rangle_W, \tag{3}$$

where $L : Y \times U \times W^\star \to R^1, \lambda \in W^\star$ and $W^\star$ denotes the dual space of W. Since W is a Hilbert space, thus $W = W^\star$.

In order to establish the optimality conditions of problem (1), we introduce the following assumptions.

(A1) There exists local optimal solution $(y^\star, u^\star) \in Z$, and $f(y^\star, u^\star) \leq f(y, u)$, if $(y, u) \in Z, e(y, u) = 0, u \in \mathcal{U}_{ad}, \|y - y^\star\|_Y + \|u - u^\star\|_U \leq \varepsilon, \varepsilon > 0$.

(A2) f and e are twice continuously Fréchet-differentiable; $\nabla^2 f(y, u), e''(y, u)$ are Lipschitz continuous in a neighborhood of $(y, u) \in Z$;

(A3) $e'(y, u)$ is surjective and $e'_y(y, u)$ is bijective.

3. Optimality conditions

Completely analogous to those of finite-dimensional problem, we can obtain the first-order necessary optimality condition of problem (1) as follows:

Theorem 3.1 *Assume that (A1)-(A3) are valid, and that $(y^\star, u^\star)$ is a local optimal solution of problem (1), then there exist $\lambda^\star \in W^\star, \mu_a^\star \geq 0, \mu_b^\star \geq 0$, such that*

$$\begin{cases} \nabla_y f(y^\star, u^\star) + e'_y(y^\star, u^\star)^\star \lambda^\star = 0, \\ \nabla_u f(y^\star, u^\star) + e'_u(y^\star, u^\star)^\star \lambda^\star - \mu_a^\star + \mu_b^\star = 0, \\ e(y^\star, u^\star) = 0, \\ (u^\star(x) - a(x))\mu_a^\star + (b(x) - u^\star(x))\mu_b^\star = 0, \\ u^\star \in \mathcal{U}_{ad}, \end{cases} \tag{4}$$

where $e'_y(y^\star, u^\star)^\star$ and $e'_u(y^\star, u^\star)^\star$ denote the adjoint operators of $e'_y(y^\star, u^\star)$ and $e'_u(y^\star, u^\star)$, respectively.

Theorem 3.2 *Let the same assumptions as in the above theorem hold, and let $(y^\star, u^\star)$ be a local optimal solution of problem (1), then $(y^\star, u^\star)$ satisfies*

$$\begin{cases} e(y^\star, u^\star) = 0, \\ \nabla_u L(y^\star, u^\star, \lambda^\star) = \begin{cases} = 0 & a(x) < u^\star < b(x), \\ \geq 0 & u^\star = a(x), \\ \leq 0 & u^\star = b(x), \end{cases} \\ \lambda^\star = -(e'_y(y^\star, u^\star)^\star)^{-1} \nabla_y f(y^\star, u^\star), \\ u^\star \in \mathcal{U}_{ad}. \end{cases} \tag{5}$$

Proof Since $(y^\star, u^\star)$ is the local optimal solution of problem (1), obviously, $u^\star \in \mathcal{U}_{ad}$ and $e(y^\star, u^\star) = 0$, and for arbitrary $y \in Y$, the following equality holds.

$$\langle \nabla_y f(y^\star, u^\star), y \rangle + \langle e'_y(y^\star, u^\star) y, \lambda^\star \rangle = 0, \quad \forall y \in Y.$$

Since the choice of y is arbitrary in Y, then

$$\nabla_y f(y^\star, u^\star) + e'_y(y^\star, u^\star)^\star \lambda^\star = 0.$$

Based on the invertibility of $e'_y(y^\star, u^\star)$, we can obtain that

$$\lambda^\star = -(e'_y(y^\star, u^\star)^\star)^{-1} \nabla_y f(y^\star, u^\star).$$

If $a(x) < u^\star < b(x)$, then we can obtain that $\nabla_u L(y^\star, u^\star, \lambda^\star) = 0$; if $u^\star = a(x)$, then obey $\mu_a^\star \geq 0, \mu_b^\star = 0$. Based on the complementary condition, we can obtain $\nabla_u L(y^\star, u^\star, \lambda^\star) \geq 0$; Analogously, if $u^\star = b(x)$ we can obey $\mu_a^\star = 0, \mu_b^\star \geq 0$ respectively, it is obviously that $\nabla_u L(y^\star, u^\star, \lambda^\star) \leq 0$.

We introduce the following assumption, which is needed to reformulate the optimality conditions.

(A4) For each $z = (y,u) \in Y \times U$, there exists a linear bounded operator $T(z) : U \to Z$, such that

$$\mathcal{N}(e'(z)) = \{s \in Z \mid e'(z)(s) = 0\} = \{T(z)(u) \mid u \in U\} = \mathcal{R}(T(z)).$$

Let $R(z)$ be a linear bounded operator $R(z) : W \to Z$, i.e., $R(z) \in \mathcal{L}(W, Z)$, such that

$$e'(z)R(z) = I_W, \quad \forall z \in Z,$$

where I_W is an identity operator, $I_W \in \mathcal{L}(W)$.

Given $(y,u) \in Z$, then $(\xi, \nu) \in Z$ lies in the null space of $e'(y,u)$ if and only if

$$e'_y(y,u)\xi + e'_u(y,u)\nu = 0.$$

Since $e'_y(y,u)$ is bijective, we obtain

$$(\xi,\nu) \in \mathcal{N}(e'(y,u)) \Longleftrightarrow (\xi,\nu) = (-e'_y(y,u)^{-1}e'_u(y,u)\nu, \nu).$$

This leads to the following definition of operators

$$T(y,u) = (-e'_y(y,u)^{-1}e'_u(y,u), I_U)^T, \quad R(y,u) = (e'_y(y,u)^{-1}, 0)^T,$$

where I_U is identity operator, $I_U \in \mathcal{L}(U)$ and the second component in $R(y,u)$ denotes the null operator in $\mathcal{L}(W,U)$. According to the above expressions, we have

$$\begin{aligned}
\nabla_u L(y^\star, u^\star, \lambda^\star) &= \nabla_u f(y^\star, u^\star) + e'_u(y^\star, u^\star)^\star \lambda^\star \\
&= \nabla_u f(y^\star, u^\star) - e'_u(y^\star, u^\star)^\star (e'_y(y^\star, u^\star)^\star)^{-1} \nabla_y f(y^\star, u^\star)) \\
&= T(y^\star, u^\star)^\star \nabla f(y^\star, u^\star).
\end{aligned}$$

Then, if $(y^\star, u^\star) \in Y \times U$ is a local optimal solution of problem (1), the first-order necessary optimality condition (5) can be written as the following equivalent form

$$\begin{cases}
e(y^\star, u^\star) = 0, \\
T(y^\star, u^\star)^\star \nabla f(y^\star, u^\star) = \begin{cases} = 0 & a(x) < u^\star < b(x), \\ \geq 0 & u^\star = a(x), \\ \leq 0 & u^\star = b(x), \end{cases} \\
u^\star \in \mathcal{U}_{ad}.
\end{cases} \tag{6}$$

Based on the idea introduced by Coleman and Li[4], for arbitrary $(y,u) \in Y \times U, u \in \mathcal{U}_{ad}$, we introduce the following scaling functional which is assumed to satisfy

$$d(y,u)(x) = \begin{cases} = 0 & \text{if} \ \ u(x) = a(x) \ \text{ and } \ T(y,u)^\star \nabla f(y,u) \geq 0 \\ = 0 & \text{if} \ \ u(x) = b(x) \ \text{ and } \ T(y,u)^\star \nabla f(y,u) \leq 0 \\ > 0 & \text{else.} \end{cases} \tag{7}$$

We can define an affine scaling function as follows

$$d_I(y,u)(x) = \begin{cases} u(x)-a(x) & \text{if } T(y,u)^\star\nabla f(y,u)>0 \text{ or } T(y,u)^\star\nabla f(y,u)=0 \\ & \text{and } u(x)\le \frac{1}{2}(a(x)+b(x)), \\ b(x)-u(x) & \text{if } T(y,u)^\star\nabla f(y,u)<0 \text{ or } T(y,u)^\star\nabla f(y,u)=0 \\ & \text{and } u(x)> \frac{1}{2}(a(x)+b(x)). \end{cases} \tag{8}$$

It is easily to verify that $d_I(y,u)$ satisfies the condition (7).

Following the above discussion, we can obtain the conclusion as follows:

Theorem 3.3 *The point $(y^\star,u^\star)$ satisfies the first-order necessary conditions if and only if the following conditions are valid*

$$\begin{cases} e(y^\star,u^\star)=0, \\ d(y^\star,u^\star)T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)=0. \end{cases} \tag{9}$$

Proof It is obviously that we need only to prove (9) is equivalent to (6). Firstly, we assume (6) is valid. If $a(x)<u^\star<b(x)$, then $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)=0$; if $u^\star=a(x)$, then $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)\ge 0$, in terms of the definition of $d(y,u)(x)$, we can obtain $d(y^\star,u^\star)(x)=0$ for arbitrary $x\in\Omega$; Analogously, if $u^\star=b(x)$, then $d(y^\star,u^\star)(x)=0$ for any $x\in\Omega$, hence (9) is valid. On the other hand, let $d(y^\star,u^\star)T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)=0$ hold. For all $x\in\Omega$ with $a(x)<u^\star<b(x)$, we have $d(y^\star,u^\star)(x)>0$, which implies $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)=0$. For all $x\in\Omega$ with $u^\star=a(x)$, we obtain $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)\ge 0$, since $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)>0$ would yield the contradiction $d(y^\star,u^\star)(x)>0$; Analogously, we can show that $T(y^\star,u^\star)^\star\nabla f(y^\star,u^\star)\le 0$ for all $x\in\Omega$ with $u^\star=b(x)$. Thus, (6) holds.

4. Trust-region subproblem

One way to motivate the algorithms is to apply Newton's method augmented by trust-region globalization to the system of nonlinear equation (9). In this section we focus on constructing the trust-region subproblems based on (9). Firstly, we need a substitution for the derivative of $d(y,u)T(y,u)^\star\nabla f(y,u)$. Formal application of the product rule suggests choosing an approximate derivative of the form[5].

$$D(y,u)\nabla\left[T(y,u)^\star\nabla f(y,u)\right]+D'(y,u)\left[T(y,u)^\star\nabla f(y,u)\right],\quad u\in\mathcal{U}_{ad}. \tag{10}$$

Here $D(y,u)$ denotes the pointwise multiplication operator associated with $d(y,u)$. It is clearly when $d(y,u)=d_I(y,u)$, the choice $D_u(y,u)\omega=d_{Iu}\omega, D_y(y,u)\eta=d_{Iy}\eta$, for all $\eta\in Y,\omega\in U$, with

$$d_{Iu}(y,u)(x) = \begin{cases} 1 & \text{if } T(y,u)^\star\nabla f(y,u)>0 \text{ or } T(y,u)^\star\nabla f(y,u)=0 \\ & \text{and } u(x)\le \frac{1}{2}(a(x)+b(x)), \\ -1 & \text{if } T(y,u)^\star\nabla f(y,u)<0 \text{ or } T(y,u)^\star\nabla f(y,u)=0 \\ & \text{and } u(x)> \frac{1}{2}(a(x)+b(x)) \end{cases} \tag{11}$$

and

$$d_{Iy}(y,u)(x) = 0. \tag{12}$$

Now we introduce the derivative of $T(y,u)^\star \nabla f(y,u)$. Since

$$T(y,u)^\star \nabla f(y,u) = \nabla_u f(y,u) + e'_u(y,u)^\star \lambda, \tag{13}$$

where $\lambda \in Z^\star$ is determined by

$$e(y,u)^\star \lambda = -\nabla_y f(y,u), \tag{14}$$

and the following equalities are fulfilled

$$\frac{\partial}{\partial y}\left(T(y,u)^\star \nabla f(y,u)\right) = T(y,u)^\star \begin{pmatrix} \nabla^2_{yy} L(y,u,\lambda) \\ \nabla^2_{uy} L(y,u,\lambda) \end{pmatrix}, \tag{15}$$

$$\frac{\partial}{\partial u}\left(T(y,u)^\star \nabla f(y,u)\right) = T(y,u)^\star \begin{pmatrix} \nabla^2_{yu} L(y,u,\lambda) \\ \nabla^2_{uu} L(y,u,\lambda) \end{pmatrix}, \tag{16}$$

where $\lambda = (e'_y(y,u)^\star)^{-1} \nabla_y f(y,u)$.

It can be shown that a Newton's step in (y,u) on nonlinear system (9) is given by

$$\left\{ \begin{array}{l} e'_y(y,u)s_y + e'_u(y,u)s_u = -e(y,u), \\ \left(D(y,u)T(y,u)^\star \nabla^2 L(y,u,\lambda) + (0, D'_u(y,u)\right) \begin{pmatrix} s_y \\ s_u \end{pmatrix} \\ \qquad = -D(y,u)T(y,u)^\star \nabla f(y,u), \end{array} \right. \tag{17}$$

where

$$\nabla^2 L(y,u,\lambda) = \begin{pmatrix} \nabla^2_{yy} L(y,u,\lambda) & \nabla^2_{yu} L(y,u,\lambda \\ \nabla^2_{uy} L(y,u,\lambda) & \nabla^2_{uu} L(y,u,\lambda \end{pmatrix}.$$

Following (A3), the solutions of the linearized state equation in (17) are of the form

$$s = \begin{pmatrix} s_y \\ s_u \end{pmatrix} = s^n + T(y,u)s_u, \tag{18}$$

where

$$s^n = -R(y,u)e(y,u) = -\begin{pmatrix} e'_y(y,u)^{-1} e(y,u) \\ 0 \end{pmatrix}. \tag{19}$$

The above system (17) can be equivalently written as

$$\left\{ \begin{array}{l} s = s^n + T(y,u)s_u, \\ \left(D(y,u)T(y,u)^\star \nabla^2 L(y,u,\lambda)T(y,u) + D'_u(y,u)\right) s_u \\ \qquad = -D(y,u)T(y,u)^\star \left(\nabla^2 L(y,u,\lambda)s^n + \nabla f(y,u)\right). \end{array} \right. \tag{20}$$

Then above system (20) can be reformulated as

$$\left\{ \begin{array}{l} s = s^n + T(y,u)s_u, \\ \left(T(y,u)^\star \nabla^2 L(y,u,\lambda)T(y,u) + D^{-1}(y,u)D'_u(y,u)\right) s_u \\ \qquad = -T(y,u)^\star \left(\nabla^2 L(y,u,\lambda)s^n + \nabla f(y,u)\right). \end{array} \right. \tag{21}$$

Introducing the scaled step $\hat{s}_u = D^{-\frac{1}{2}}(y,u)s_u$, then (21) takes the form

$$\begin{cases} \hat{s} = s^n + D^{\frac{1}{2}}(y,u)T(y,u)\hat{s}_u, \\ \left(D^{\frac{1}{2}}(y,u)T(y,u)^\star\nabla^2 L(y,u,\lambda)T(y,u)D^{\frac{1}{2}}(y,u) + D'_u(y,u)\right)\hat{s}_u \\ \quad = -D^{\frac{1}{2}}(y,u)T(y,u)^\star\left(\nabla^2 L(y,u,\lambda)s^n + \nabla f(y,u)\right). \end{cases} \tag{22}$$

Following the idea introduced by Dennis and Vicente[6], we can structure two trust-region subproblems corresponding to the system of equation of (22) respectively.

$$\begin{array}{ll} \min & \frac{1}{2}\|(e'_y(y,u), e'_u(y,u))s^n + e(y,u)\|_W^2 \\ \text{s.t.} & \|s^n\|_Z \le \Delta \end{array} \tag{23}$$

and

$$\begin{array}{ll} \min & \frac{1}{2}\langle \hat{s}_u, (D^{\frac{1}{2}}(y,u)T(y,u)^\star\nabla^2 L(y,u,\lambda)T(y,u)D^{\frac{1}{2}}(y,u) + D'_u(y,u))\hat{s}_u\rangle + \\ & \langle \hat{s}_u, D^{\frac{1}{2}}(y,u)T(y,u)^\star\left(\nabla^2 L(y,u,\lambda)s^n + \nabla f(y,u)\right)\rangle \\ \text{s.t.} & \|\hat{s}_u\|_U \le \Delta. \end{array} \tag{24}$$

Let B denote a symmetric approximation of $\nabla^2 L(y,u,\lambda)$, and assume that the norms $\|B\|_{\mathcal{L}(Z,Z^\star)}$ are uniformly bounded by a constant $c_1 > 0$; In the subproblem (23), since s^n is required to have the form $\binom{s_y^n}{0}$, the displacement along s^n is made only in the y variable. As a consequence, z and $z + s^n$ have the same u component, then the previous two trust-region subproblems can be rewritten as

$$\begin{array}{ll} \min & \frac{1}{2}\|e'_y(y,u)s_y^n + e(y,u)\|_W^2 \\ \text{s.t.} & \|s_y^n\|_Y \le \Delta, \end{array} \tag{25}$$

$$\begin{array}{ll} \min & \frac{1}{2}\langle \hat{s}_u, (D^{\frac{1}{2}}(y,u)T(y,u)^\star BT(y,u)D^{\frac{1}{2}}(y,u) + D'_u(y,u))\hat{s}_u\rangle + \\ & \langle \hat{s}_u, D^{\frac{1}{2}}(y,u)T(y,u)^\star\left(Bs^n + \nabla f(y,u)\right)\rangle \\ \text{s.t.} & \|\hat{s}_u\|_U \le \Delta. \end{array} \tag{26}$$

If we work with the original variable, the subproblem (26) reads as follows

$$\begin{array}{ll} \min & \frac{1}{2}\langle s_u, (T(y,u)^\star BT(y,u) + D^{-1}(y,u)D'_u(y,u))s_u\rangle + \\ & \langle s_u, T(y,u)^\star\left(Bs^n + \nabla f(y,u)\right)\rangle \\ \text{s.t.} & \|D^{-1}(y,u)s_u\|_U \le \Delta \\ & u + s_u \in \mathcal{U}_{ad}. \end{array} \tag{27}$$

In the above subproblem, we require that the new iterate is in the interior of the box constraint. However, it is important to remark that the bound constraint do not need to be strictly enforced when computing s_u[6]. An approximate of

$$\begin{array}{ll} \min & \frac{1}{2}\langle s_u, (T(y,u)^\star BT(y,u) + D^{-1}(y,u)D'_u(y,u)s_u\rangle + \\ & \langle s_u, T(y,u)^\star\left(Bs^n + \nabla f(y,u)\right)\rangle \\ \text{s.t.} & \|D^{-1}(y,u)s_u\|_U \le \Delta \end{array} \tag{28}$$

is computed and then scaled by a constant $\gamma > 0$ so that $u + \gamma s_u \in \mathcal{U}_{ad}$.

Let $z_k = (y_k, u_k) \in Z$ be given with $u_k \in \mathcal{U}_{ad}$, we introduce a quadratic model

$$q_k(s) = L_k + \langle s, \nabla L_k \rangle + \frac{1}{2}\langle s, B_k s \rangle$$

of $L(z_k + s, \lambda_k)$ defined by (3) about (z_k, λ_k). Obviously, we can derive that

$$q_k(s_k^n + T(y_k, u_k)s_u) = q_k(s_k^n) + \langle s_u, \bar{g}_k \rangle + \frac{1}{2}\langle s_u, T(y_k, u_k)^\star B_k T(y_k, u_k)s_u \rangle, \tag{29}$$

where

$$\bar{g}_k = T(y_k, u_k)^\star \nabla q_k(s_k) = T(y_k, u_k)^\star (B_k s_k^n + \nabla f(y_k, u_k)),$$

and B_k is the symmetric approximation of $\nabla^2 L(y_k, u_k, \lambda_k)$, $B_k \in \mathcal{L}(Z, \mathcal{L}(Z, R^1)) = \mathcal{L}(Z, Z^\star)$. For convenience, we define

$$\varphi_k(s_k) = q_k(s_k^n + T(y_k, u_k)s_u) + \frac{1}{2}\langle s_u, D_k^{-1} D_{ku} s_u \rangle. \tag{30}$$

In order to compute the new iterate $z_{k+1} = z_k + s_k$, with $s_k = s_k^n + T_k(s_u)_k, u_k + (s_u)_k \in \mathcal{U}_{ad}$, we should solve the following trust-region subproblems in order.

$$\begin{array}{ll} \min & \|e'_y(y_k, u_k)(s_y^n)_k + e(y_k, u_k)\|_W^2 \\ \text{s.t.} & \|(s_y^n)_k\|_Y \le \Delta_k \end{array} \tag{31}$$

and

$$\begin{array}{ll} \min & \varphi_k(s_k) \\ \text{s.t.} & \|D_k^{-1}(s_u)_k\|_U \le \Delta_k \\ & u_k + (s_u)_k \in \mathcal{U}_{ad}. \end{array} \tag{32}$$

Due to the above discussion, the problem (1) is transferred into two trust-region subproblems which are similar to the trust-region subproblem for the unconstrained case. In unconstrained optimization, the use of a trust-region has made it possible to make strong guarantees of convergence, in order to ensure global convergence, the step is required to satisfy the FCD condition; namely, the step must product at least a fraction of decrease obtained by Cauchy step[6]. Here the Cauchy decrease denotes the maximum possible decrease along the steepest descent direction of $\|e'_y(y_k, u_k)s_y^n + e(y_k, u_k)\|_W$ at $s_y^n = 0$ and φ_k at $s_u = 0$ inside the feasible region of the subproblem (31) and (32) respectively.

In order to guarantee global convergence we require $(s_y^n)_k$ to satisfy

$$\|(s_y^n)_k\|_Y \le K\|e(y_k, u_k)\|_W, \tag{33}$$

where K is a positive constant independent of the iterate k of the algorithm; and satisfy the following fraction of Cauchy decrease condition

$$\left\{ \begin{array}{l} \|e(y_k, u_k)\|_W^2 - \|e'_y(y_k, u_k)(s_y^n)_k + e(y_k, u_k)\|_W^2 \\ \qquad \ge \kappa_1 \left(\|e(y_k, u_k)\|_w^2 - \|e'_y(y_k, u_k)(s_y^c)_k + e(y_k, u_k)\|_W^2 \right), \\ \|(s_y^n)_k\|_Y \le \Delta_k, \end{array} \right. \tag{34}$$

where $\kappa_k > 0$ does not depend on the iterate k of the algorithm and $(s_y^c)_k$ is the so-called Cauchy step for this trust-region subproblem, i.e. $(s_y^c)_k$ is the optimal solution of the following one-dimensional optimization problem

$$\begin{array}{ll} \min & \frac{1}{2}\|e_y'(y_k,u_k)s_y + e(y_k,u_k)\|_W^2 \\ \text{s.t.} & \|s_y\|_Y \le \Delta_k \\ & s_y = -te_y'(y_k,u_k)^\star e(y_k,u_k), \quad t \ge 0. \end{array}$$

Now following the powerful lemma presented by Powell[8], we can deduce that the decrease given by $(s_k^n)_k$ is such that

$$\begin{aligned} &\|e(y_k,u_k)\|_W^2 - \|e_y'(y_k,u_k)(s_y^n)_k + e(y_k,u_k)\|_W^2 \\ &\ge \frac{1}{2}\kappa_1\|e_y'(y_k,u_k)^\star e(y_k,u_k)\|_Y \min\{\frac{\|e_y'(y_k,u_k)^\star e(y_k,u_k)\|_Y}{e_y'(y_k,u_k)^\star e_y'(y_k,u_k)\|_Y}, \Delta_k\}. \end{aligned} \tag{35}$$

However, we do not have to solve (31) exactly. For example, in the finite-dimensional setting, it is only to assume that $(s_y^n)_k$ satisfies (33) and

$$\begin{aligned} &\|e(y_k,u_k)\|_W^2 - \|e_y'(y_k,u_k)(s_y^n)_k + e(y_k,u_k)\|_W^2 \\ &\quad \ge \kappa_1\|e(y_k,u_k)\|_W \min\{\kappa_2\|e(y_k,u_k)\|_W, \Delta_k\}, \end{aligned} \tag{36}$$

where K, κ_2, κ_3 are positive constants independent of k. Similar techniques can be applied in the infinite-dimensional framework. In fact, the condition (34) is just a weaker form of Cauchy decrease for the trust-region subproblem (31).

As for subproblem (32), it is analogous to the situation in [7]. We may take

$$-D_k\bar{g}_k = -D_k(T(y_k,u_k)^\star B_k s_k^n + \nabla f(y_k,u_k))$$

as the Cauchy decrease direction of φ_k and therefore define the following fraction of Cauchy decrease condition: There exists $\kappa_4 > 0$ (fixed for all k) such that $(s_u)_k$ is an approximate solution of

$$(T(y_k,u_k)^\star B_k T(y_k,u_k) + D_k^{-1}D_{ku}')s_u = -T(y_k,u_k)^\star(B_k s_k^n + \nabla f(y_k,u_k)) \tag{37}$$

satisfying

$$\begin{cases} \|D^{-1}(s_u)_k\|_U \le \Delta_k, \\ u_k + (s_u)_k \in \mathcal{U}_{ad}, \\ \varphi_k((s_u)_k) < \kappa_4\varphi((s_u^c)_k), \end{cases} \tag{38}$$

where $(s_u^c)_k$ is a solution of the following one-dimensional optimization problem.

$$\begin{array}{ll} \min & \varphi_k(s_u) \\ \text{s.t.} & s_u = -tD_kT(y_k,u_k)^\star(B_k s_k^n + \nabla f(y_k,u_k)), \ t \ge 0 \\ & \|D_k^{-1}s_u\|_U \le \Delta_k \\ & \sigma_k(a(x) - u_k) \le s_k \le \sigma_k(b(x) - u_k), \end{array} \tag{39}$$

where $\sigma_k \in [\sigma, 1]$ ensure that the Cauchy step $(s_u^c)_k$ remains strictly feasible with respect to the box constraint, and the parameter $\sigma \in (0,1)$ is fixed for all iterate k.

References:

[1] LUENBERGER D G. *Optimization by Vector Space Methods* [M]. John Wiley & Sons Inc., New York, 1969.

[2] IFFE A D. *Theory of Extremal Problems* [M]. North-Holland, Amsterdam, 1979.

[3] MAURER H. *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems* [J]. Math. Program., 1979, **16**(2): 98–110.

[4] COLEMAN T F, LI Y. *An interior point trust-region approach for nonlinear minimization subject to bounds* [J]. SIAM. J. Optimization, 1996, **56**(6): 418–445.

[5] ULBRICH M, ULBRICH S. *Global convergence of trust region interior point algorithms for infinite-dimensional non-convex minimization subject to pointwise bounds* [J]. SIAM J. Control Optimization, 1999, **37**(3): 731–764.

[6] DENNIS J E, VICENTE L N. *Trust-region interior-point algorithms for a class of nonlinear programming problems* [J]. SIAM J. Control Optimization, 1998, **36**(4): 1750–1794.

[7] ULBRICH M, ULBRICH S. *Superlinear convergence of affine-scaling interior-point newton methods for infinite-dimensional nonlinear problems with pointwise bounds* [J]. SIAM J. Control and Optimization, 2000, **38**(5): 1938–1984.

[8] POWELL J D. *A New Algorithm for Unconstrained Optimization in Nonlinear Programming* [M]. Academic Press, New York, 1970.

具有算子与有界约束的无穷维规划问题的最优性条件

李春发[1]，冯恩民[2]

(1. 天津理工学院经济与管理学院，天津 300191; 2. 大连理工大学应用数学系，辽宁 大连 116024)

摘 要: 本文给出了当 all-at-once 方法用于求解最优控制问题而产生的一类同时具有算子和对部分变量具有简单界约束的无穷维最优化问题的一个一阶必要条件，构造了相应的信赖域子问题，据此，信赖域法可以用于求解无穷空间中的最优化问题.

关键词: 最优控制；算子约束；信赖域法.