CLOSE APPROACH TO CONSTRAINT SET: OPTIMAL CONTROL PROBLEM

PAUL J. SCOTT

ABSTRACT. A system of ordinary differential equations subject to additional equality constraints is considered. An optimal control problem and necessary condition are posed for close approach to the constraint set. This requires constraint set proximity related control constraints.

1. INTRODUCTION

There is interest in systems of ordinary differential equations acting on a set constrained by a system of equality constraints. An initial point satisfying the constraints may not be available. In this case trajectories that approach the constrained set must be found. Such trajectories are extremals of a suitable optimal control problem. This problem includes constraints of the control variables related to proximity of the constraint set.

Let n and m be positive integers. Let y denote a vector in $\mathbb{R}^n$ and z a vector in $\mathbb{R}^m$. The pair (y, z) is called a state vector. Also f is a smooth function mapping $\mathbb{R}^{n+m} \longrightarrow \mathbb{R}^n$ and g is a smooth function mapping $\mathbb{R}^{n+m} \longrightarrow \mathbb{R}^m$.

Let t denote a real number, and for vector valued functions of t, $(\ )'$ denotes the derivative with respect to, wrt, t. State trajectories $(y(t), z(t))$ satisfying the differential equation

$$y'(t) = f(y(t), z(t)) \tag{1.1}$$

and the constraint

$$g(y(t), z(t)) = 0 \tag{1.2}$$

are of interest.

Starting from some initial state (y_0, z_0) that does not satisfy constraint (1.2), a trajectory approaching the constraint is desired. To find such trajectories an optimal control problem is defined in section 2. A necessary condition for solution of this problem is formulated in section 3. First some definitions are required.

The transpose is denoted by $(\ )^\tau$. The Jacobian of g wrt z is defined by

$$g_z = \begin{pmatrix} g_{1,z_1} & \cdots & g_{1,z_m} \\ \vdots & \ddots & \vdots \\ g_{m,z_1} & \cdots & g_{m,z_m} \end{pmatrix}$$

with g_{i,z_j} denoting the partial derivative of g_i wrt z_j. Similar definitions for g_y, f_z, and f_y. In the following assume the inverse of g_z, denoted g_z^{-1} exists.

Date: November 3, 1996.

Key words and phrases. ODE's, optimal control, constraints.

Let u denote a vector in $\mathbb{R}^m$. u is called a control and for some control trajectory $u(t)$ the following differential equation is introduced.

$$z'(t) = u(t) \tag{1.3}$$

The control trajectory $u(t)$ will be adjusted so solutions of state equations (1.1) and (1.3) approach the constrained set. A measure of proximity to the constrained set must be included in this adjustment. First note that along a state trajectory the derivative of g wrt t is given by

$$g'(y(t), z(t)) = g_y(y(t), z(t)) \cdot f(y(t), z(t)) + g_z(y(t), z(t)) \cdot u(t) \tag{1.4}$$

Next define $g^2 : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^m$ by

$$g_i^2(y, z) = (g_i(y, z))^2 \text{ for } i = 1, \ldots, m \tag{1.5}$$

The following control constraints include a measure of proximity to the constraint set

$$-g^2(y, z) \leq g_y(y, z) \cdot f(y, z) + g_z(y, z) \cdot u \leq g^2(y, z) \tag{1.6}$$

Express these constraints, in terms of two functions mapping $\mathbb{R}^{n+m+m} \longrightarrow \mathbb{R}^m$ defined by

$$c_+(y, z, u) = \quad g_y(y, z) \cdot f(y, z) + g_z(y, z) \cdot u - g^2(y, z) \tag{1.7}$$

$$c_-(y, z, u) = -g_y(y, z) \cdot f(y, z) - g_z(y, z) \cdot u - g^2(y, z) \tag{1.8}$$

as

$$c_+(y, z, u) \leq 0 \tag{1.9}$$

$$c_-(y, z, u) \leq 0 \tag{1.10}$$

At a point (y, z, u) a component of a constraint is called active if it is zero, and inactive if it is negative.

2. Optimal Control Problem

A control trajectory is sought for generating a state trajectory starting from the initial state that approachs the constraint set. This can be possed as the following optimal control problem.

For some positive value of t, denoted by t_1, find the control trajectory $u(t)$ defined for $0 \leq t \leq t_1$ that minimizes the functional $J(u(t))$ defined by

$$J(u(t)) = \int_0^{t_1} g^\tau(y(t), z(t)) \cdot g(y(t), z(t)) dt \tag{2.1}$$

subject to the initial conditions

$$y(0) = y_0 \tag{2.2}$$

$$z(0) = z_0$$

state equations (1.1) and (1.3), and such that for each t, $(y(t), z(t), u(t))$ satisfies control constraints (1.9) and (1.10).

3. Necessary Conditions

Let λ denote a vector in $\mathbb{R}^n$ and ξ a vector in $\mathbb{R}^m$. The pair (λ, ξ) is called a costate vector. Let l_+ and l_- denote vectors in $\mathbb{R}^m$ which will be called K-T multipliers. A Hamiltonian function H is defined by

$$H(y,z,u,\lambda,\xi,l_+,l_-) = g^\tau(y,z)\cdot g(y,z) + \lambda^\tau \cdot f(y,z) + \xi^\tau \cdot u + l_+^\tau \cdot c_+(y,z,u) + l_-^\tau \cdot c_-(y,z,u) \tag{3.1}$$

If control trajectory $u(t)$ is a solution of the optimal control problem it is necessary that a costate trajectory $(\lambda(t), \xi(t))$ exist satisfying the differential equations

$$\lambda'(t) = -H_y^\tau(y(t),z(t),u(t),\lambda(t),\xi(t)) \tag{3.2}$$

$$\xi'(t) = -H_z^\tau(y(t),z(t),u(t),\lambda(t),\xi(t)) \tag{3.3}$$

and end conditions

$$\lambda(t_1) = 0 \qquad \xi(t_1) = 0 \tag{3.4}$$

the control $u(t)$ gives a minimum of $H(y(t), z(t), u, \lambda(t), \xi(t), l_+(t), l_-(t))$ among all controls u satisfying

$$c_+(y(t),z(t),u) \le 0 \tag{3.5}$$

$$c_-(y(t),z(t),u) \le 0 \tag{3.6}$$

and the K-T mutilpliers $(l_+(t), l_-(t))$ satisfy the complimentary slackness conditions

$$l_+(t) \ge 0 \qquad l_+^\tau(t)\cdot c_+(y(t),z(t),u(t)) = 0 \tag{3.7}$$

$$l_-(t) \ge 0 \qquad l_-^\tau(t)\cdot c_-(y(t),z(t),u(t)) = 0 \tag{3.8}$$

and

$$H_u(y(t),z(t),u(t),\lambda(t),\xi(t),l_+(t),l_-(t)) = 0 \tag{3.9}$$

The ordered set consisting of a state trajectory, costate trajectory, control trajectory, and K-T trajectory satisfying the constraints of the optimal control problem and its necessary conditions is called an extremal.

4. Expansion of Necessary Conditions

Let L_- be a vector in $\mathbb{R}^m$ defined by

$$L_- = l_+ - l_- \tag{4.1}$$

and let L_+ be an $m \times m$ diagonal matrix defined by

$$L_+ = \begin{pmatrix} l_{+,1} + l_{-,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & l_{+,m} + l_{-,m} \end{pmatrix} \tag{4.2}$$

Using the constraint definitions (1.7), the definition of g^2 (1.5), and the definitions of L_-, and L_+ the Hamiltonian H can be expressed as

$$H(y,z,u,\lambda,\xi,L_+,L_-) = g^\tau(y,z)\cdot(I - L_+)\cdot g(y,z) + \left(\lambda + g_y^\tau(y,z)\cdot L_-\right)^\tau \cdot f(y,z) + (\xi + g_z^\tau(y,z)\cdot L_-)^\tau \cdot u \tag{4.3}$$

Using this form equation (3.9) becomes

$$\xi(t) + g_z^{\tau}(y(t), z(t)) \cdot L_-(t) = 0 \tag{4.4}$$

and solving for $L_-(t)$ gives

$$L_-(t) = -g_z^{-\tau}(y(t), z(t)) \cdot \xi(t) \tag{4.5}$$

The Hamiltonian H, (3.1), is linear wrt u. For a minimizing u, m constraint components must be active, and the active components of c_+ and c_- are mutualy exclusive. So, by the complimentary slackness conditions (3.7), $L_+(t)$ can be found as follows

$$L_{-,i}(t) > 0 \Longrightarrow l_{-,i}(t) = 0 \text{ and } l_{+,i}(t) > 0 \Longrightarrow L_{+,[i,i]}(t) = l_{+,i}(t) = L_{-,i}(t) \tag{4.6}$$

$$L_{-,i}(t) < 0 \Longrightarrow l_{+,i}(t) = 0 \text{ and } l_{-,i}(t) > 0 \Longrightarrow L_{+,[i,i]}(t) = l_{-,i}(t) = -L_{-,i}(t) \tag{4.7}$$

so $L_+(t)$ is given by

$$L_+(t) = \begin{pmatrix} sign(L_{-,1}(t))L_{-,1}(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & sign(L_{-,m}(t))L_{-,m}(t) \end{pmatrix} \tag{4.8}$$

For $u(t)$ notice that

$$L_{-,i}(t) > 0 \Longrightarrow c_{+,i}(y(t), z(t), u(t)) \text{ is active } \Longrightarrow \tag{4.9}$$
$$g_{i,y}(y(t), z(t)) \cdot f(y(t), z(t)) + g_{i,z}(y(t), z(t)) \cdot u(t) = g_i^2(y(t), z(t))$$

$$L_{-,i}(t) < 0 \Longrightarrow c_{-,i}(y(t), z(t), u(t)) \text{ is active } \Longrightarrow \tag{4.10}$$
$$g_{i,y}(y(t), z(t)) \cdot f(y(t), z(t)) + g_{i,z}(y(t), z(t)) \cdot u(t) = -g_i^2(y(t), z(t))$$

which can be expressed as the following linear system

$$g_y(y(t), z(t)) \cdot f(y(t), z(t)) + g_z(y(t), z(t)) \cdot u(t) = \text{sign}\,(L_-(t))\, g^2(y(t), z(t)) \tag{4.11}$$

and solving for $u(t)$ gives

$$u(t) = g_z^{-1}(y(t), z(t)) \cdot \left(\text{sign}\,(L_-(t))\, g^2(y(t), z(t)) - g_y(y(t), z(t)) \cdot f(y(t), z(t))\right) \tag{4.12}$$

Using (4.12) the derivatives of H wrt y and z are

$$H_y = 2g^{\tau}(y, z) \cdot (I - L_+) \cdot g_y(y, z) + \left(\lambda + g_y^{\tau}(y, z) \cdot L_-\right)^{\tau} \cdot \tag{4.13}$$
$$f_y(y, z) + L_-^{\tau} \cdot \bigotimes_{i=1}^{n} \left(g_{y,y_i}(y, z) - g_{z,y_i}(y, z)\right) \cdot g_z^{-1}(y, z) \cdot g_y(y, z) \cdot f(y, z) +$$
$$L_-^{\tau} \cdot \bigotimes_{i=1}^{n} g_{z,y_i}(y, z) \cdot g_y^{-1}(y, z) \cdot \text{sign}\,(L_-)\, g^2(y, z)$$

and

$$H_z = 2g^{\tau}(y,z)\cdot(I-L_+)\cdot g_z(y,z) + \left(\lambda + g_y^{\tau}(y,z)\cdot L_-\right)^{\tau}\cdot f_z(y,z) + L_-^{\tau}\cdot\bigotimes_{i=1}^{m}\left(g_{y,z_i}(y,z) - g_{z,z_i}(y,z)\right)\cdot g_z^{-1}(y,z)\cdot g_y(y,z)\cdot f(y,z) + L_-^{\tau}\cdot\bigotimes_{i=1}^{m} g_{z,z_i}(y,z)\cdot g_y^{-1}(y,z)\cdot \operatorname{sign}(L_-)\, g^2(y,z) \tag{4.14}$$

P. O. Box 369, St. Andrews, New Brunswick, Canada. E0G 2X0
E-mail address: pauls@nb.sympatico.ca