Proceedings of the World Congress on Engineering 2011 Vol I
WCE 2011, July 6 - 8, 2011, London, U.K.

First Order Method for Optimal Control using Parmetric Optimization

Amr Radwan and Andreas Griewank *

Abstract—**We consider the optimal control problem from view point of parametric aspects. We have examined the case of the parameterized problem. This case describes the situation when the objective functional contains time t as a parameter. We also show how to apply the parametric optimization techniques for finding a nominal optimal control path.**

Keywords: Optimal Control, Parametric Optimization, KKT conditions

1 Introduction

Parametric optimization offers very useful techniques for solving optimization problems in finite dimensional spaces whenever the objective function and the constraint function continuously depends on some unknown parameter. These techniques yield as a result a minimizing (or maximizing) curve that depends continuously on the original parameter.

On the other hand, traditional techniques for solving optimal control problems rely on finding the nominal control trajectory that minimizes the Hamiltonian at each instant along the time segment. Such trajectory should be continuously time-dependent as well. Basing on this argument, we can establish a bond between these two types of optimization and will finally show how to apply the parametric optimization techniques for solving optimal control problems. There are many works devoted to theory and methods of optimal control (see, e.g., [1, 3]).

The paper is organized as follows: The first section is devoted to basic problem of optimal control and traditional approaches for solving them. Second section examines application of parametric optimization to optimal control problems. Two numerical examples are discussed in the last section.

2 Basic optimal control problem

The basic problem of optimal control can be formulated as follows: find a control that minimizes the objective functional

$$\min \mathcal{J}(\mathbf{u}) = \phi(x(t_f)) + \int_{t_0}^{t_f} f_0(\mathbf{x}, \mathbf{u}, t)dt \qquad (1)$$

subject to

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \qquad \mathbf{x}(t_0) = \mathbf{x}^0 \qquad (2)$$

$$\mathbf{u}(t) \in U, \qquad t \in T = [t_0, t_f] \qquad (3)$$

where T is fixed. System (2) describes the connection between the state variable $x(t) \in R^n$ and the control variable $u(t) \in R^r$ at each $t \in T$ and also $u \in PC^r(T)$. Here $x^0 \in R^n$ is a given vector and U is a set in R^r that specifies the constraints imposed on all admissible control functions. The vector function f and scalar function f_0 are continuous together with their partial derivatives with respect to x for all admissible controls $u \in U$. The traditional approach for solving the problem (1)-(3) consists in applying the necessary condition of optimality in the form of Pontryagin's maximum principle, that is:

If $\mathbf{u}^*(t)$ is optimal in the problem (1)-(3), then it must satisfy the maximum condition

$$H(\psi^*, \mathbf{x}^*, \mathbf{u}^*, t) = \max_{\mathbf{v} \in U} H(\psi^*, \mathbf{x}^*, \mathbf{v}, t) \quad \text{almost for all} \quad t \in T \qquad (4)$$

where

$$H(\psi, x, u, t) = \langle \psi(t), f(x, u, t) \rangle - f_0(x, u, t)$$

denotes the Hamiltonian function and $\psi(t)$ is a solution of the conjugate system:

$$\begin{cases} \dot{\psi} = -\dfrac{\partial H(\psi, \mathbf{x}, \mathbf{u}, t)}{\partial \mathbf{x}} \\ \psi(t_f) = \phi_x(x(t_f)) \end{cases} \qquad (5)$$

while x^* and ψ^* are the solutions of (2) and (5) for $u = u^*(t)$, respectively. The maximum principle (4) is a key feature for many successive approximation algorithms. These algorithms generate a sequence of admissible control $\{u^k\}$ that is relaxational in the sense that

$$\mathcal{J}(u^{k+1}) < \mathcal{J}(u^k)$$

and a minimizing one, that is, $\|\mathbf{u}^k(t) - \mathbf{u}^*(t)\| \to 0$ when $k \to \infty$ almost for all $t \in T$.

*Humboldt University, Berlin, Germany Email: radwan@mathematik.hu-berlin.de, griewank@mathematik.hu-berlin.de

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The standard technique for construction of $\mathbf{u}^{k+1}$ out of $\mathbf{u}^k$ relies on the use of so-called nominal optimal control defined by

$$\hat{\mathbf{u}}^k(t) = \arg\min_{\mathbf{v}\in U} H(\psi^k, \mathbf{x}^k, \mathbf{v}, t) \quad \text{for almost all} \quad t \in T. \tag{6}$$

Several methods can now be used iteratively to construct u^{k+1} by maximizing the Hamiltonian subject to the boundary value problem. In each iteration step, the differential equation is numerically integrated forward in time while the adjoint differential equations are integrated backwards in time.

3 Application of parametric optimization to optimal control

Consider the optimal control problem

$$\min \mathcal{J}(\mathbf{u}) = \int_{t_0}^{t_f} f_0(\mathbf{x}, \mathbf{u}, t)dt \tag{7}$$

$$\dot{x_i} = f_i(\mathbf{x}, \mathbf{u}, t), i = 1, 2, \ldots, n \quad \mathbf{x}(t_0) = \mathbf{x}^0, t \in [t_0, t_f] \tag{8}$$

$$\mathbf{u} \in U = \{\mathbf{u}(t) \in R^r / g_i(\mathbf{u}) \leq 0, i = 1, \ldots, s\} \tag{9}$$

where t_0, t_f and $\mathbf{x}^0$ are given, the functions $f_i, i = 0, \ldots, n$ with partial derivatives $\frac{\partial f_i}{\partial x_k}, k = 1, \ldots n$ are continuous on $R^n \times U \times R$, and $g_i : R^r \to R, i = 1, \ldots, s$ are twice continuously differentiable convex functions. The Hamiltonian for problem (7)-(9) is written as follows:

$$H(\psi, x, u, t) = \sum_{i=1}^{n} \psi_i f_i(x, u, t) - f_0(x, u, t), \quad t \in [t_0, t_f]$$

$$\left\{ \begin{array}{rcl} \dot{x_i} & = & \dfrac{\partial H(\psi, \mathbf{x}, \mathbf{u}, t)}{\partial \psi_i} \\ \dot{\psi_i} & = & -\dfrac{\partial H(\psi, \mathbf{x}, \mathbf{u}, t)}{\partial x_i} \end{array} \right. \tag{10}$$

and

$$\left\{ \begin{array}{rcl} \mathbf{x}(t_0) & = & \mathbf{x}^0 \\ \psi(t_f) & = & \phi_{\mathbf{x}}(\mathbf{t_f}) \end{array} \right. \tag{11}$$

Furthermore, we assume that

[H1] The Hamiltonian is strictly concave with respect to $\mathbf{u}$;

[H2] $\tilde{\mathbf{u}}(t) = \arg\max_{\mathbf{u}\in U} H(\psi, \mathbf{x}, \mathbf{u}, t), t \in [t_0, t_f]$ is continuous on $[t_0, t_f]$

[H3] $\mathbf{u} \in C^r[t_0, t_f]$

Note that $\mathbf{u}(t)$ is determined in a unique way for each t since U is convex and H is concave. Now we consider the problem of maximizing the Hamiltonian with respect to $\mathbf{u}$:

$$\max_{u\in U} H(\psi, x, u, t) \quad \text{for each} \quad t \in [t_0, t_f]$$

Usually in literature, u is found explicitly as a function $u = u(\psi, x, t)$ and after substituting it into the system (10)-(11), the problem reduces to the boundary value problem.

Theorem 3.1 *Assume that the conditions [H1]-[H3] hold and problem (7)-(9) has an optimal solution* $(\mathbf{u}^*, \mathbf{x}^*)$. *Then for a given* ε *there exist a finite discretization*

$$t_0 = \tau_0 < \tau_1 < \ldots < \tau_i < \ldots < \tau_N = t_f$$

and approximate solution $\tilde{\mathbf{u}}(t), t \in [t_0, t_f]$ *such that*

$$\| \mathbf{u}^*(t_i) - \tilde{\mathbf{u}}(t_i) \| < \epsilon, i = 1, 2, \ldots, N$$

Proof: Let $\mathbf{u}^*$ be an optimal solution of problem (7)-(9). Then $(\mathbf{u}^*, \mathbf{x}^*)$ satisfies the conditions:

$$\begin{array}{rcl} \dot{x_i^*} & = & \dfrac{\partial H(\psi^*, \mathbf{x}^*, \mathbf{u}^*, t)}{\partial \psi_i} \\ \dot{\psi_i} & = & -\dfrac{\partial H(\psi^*, \mathbf{x}^*, \mathbf{u}^*, t)}{\partial x_i}, \quad i = 1, 2, \ldots, n \end{array}$$

where

$$H(\psi, x, u, t) = \sum_{i=1}^{n} \psi_i(t) f_i(x, u, t) - f_0(x, u, t), \quad t \in [t_0, t_f]$$

and $(\mathbf{u}^*, \mathbf{x}^*)$ satisfies the maximum principle:

$$H(\psi^*, x^*, u^*, t) = \max_{u\in U} H(\psi^*, x^*, u, t), \quad t \in [t_0, t_f] \tag{12}$$

Now we consider problem (12) as one parametric maximization problem. Since $H(\psi^*, x^*, u, t)$ is twice differentiable in u and assumptions [H1]-[H3] hold, we can apply the Theorem 3.4.1 from [4, p.78] to the problem (12). Then as a result, the method PATH1 [4] generates a discretization $t_0 = \tau_0 < \tau_1 < \ldots < \tau_i < \ldots < \tau_N = t_f$, and corresponding points $\tilde{u_i} = \tilde{u}(t_i)$ such that:

$$\| u^*(t_i) - \tilde{u}(t_i) \| < \epsilon, i = 1, 2, \ldots, N$$

which proves the assertion. Parametric optimization also can be applied in finding nominal optimal control [2] given in (6). It is easy to see that at each iteration k, the Hamiltonian function is a scalar function of $\mathbf{u} \in U \subset R^r$ and $t \in T = [t_0, t_f]$, that is

$$G_k(u, t) = -H(\psi^k(t), x^k(t), u, t)$$

The latter states that $\hat{u}^k(t)$ must be a minimizer of the following problem

$$\min_{u\in U} G_k(u, t), \quad t \in T \tag{13}$$

which is a problem of parametric optimization as formulated in various papers [4], where the independent variable t is now considered as unknown parameter $t \in T = [t_0, t_f]$. We can also consider a case when the set of admissible control is time-varying, i.e. $U = U(t), t \in T = [t_0, t_f]$. In this case, a general theory of parametric optimization is also applicable for finding the nominal optimal controls.

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Remark 3.2 *One can write (13) and add a penalty term as:*

$$\delta u^j(t) = \arg \min_{\delta u(t) \in U - u} G_k(u + \delta u, t) + \beta \| \delta u(t) \|^2$$

Then u^{k+1} *is constructed using the newton step:*

$$u^{k+1}(t) = u^k(t) + \delta u^k$$

with $\beta = 0$ *initially, If* U *is convex function and* H *is convex function w.r.t.* u *then* δu *is a descent direction and evaluating the objective for* $u^k(t) + \delta u^k(t)$ *must yield descent. In general we must still have descent so that* $u^k(t) + \delta u^k(t)$ *yields a reduction when* β *is chosen sufficiently large. When* $\beta = 0$ *does not work, we set* $\beta = 1$ *and then double it repeatedly until the objective is really reduced.*

Let us go back to the parametric optimization problem (13)

$$\min_{\mathbf{u} \in U} G(\mathbf{u}, t), \quad t \in [t_0, t_f], \tag{14}$$

$$U = \{\mathbf{u} \in R^r : g_i(\mathbf{u}) \leq 0, i \in J\}, \quad J = \{1, 2, \ldots, s\}.$$

The **KKT** conditions for the problem (14) state that

$$D_{\mathbf{u}} G(\mathbf{u}, t) + \sum_{j \in J} \mu_i D_{\mathbf{u}} g_i(\mathbf{u}, t) = 0$$

$$g_i(\mathbf{u}, t) \leq 0, \quad \mu_i \geq 0, \quad i \in J$$

$$\mu_i g_i(\mathbf{u}, t) = 0, \quad i \in J, \quad t \in [t_0, t_f]$$

where

$$D_{\mathbf{u}} f(\mathbf{u}, t) = \left(\frac{\partial f(\mathbf{u}, t)}{\partial u_1}, \ldots, \frac{\partial f(\mathbf{u}, t)}{\partial u_r} \right).$$

Consider the auxiliary parametric optimization problem

$$\min G(\mathbf{u}, t), \quad t \in [t_0, t_f] \tag{15}$$

subject to

$$g_i(\mathbf{u}) = 0, i \in \tilde{J} \subset J. \tag{16}$$

Let $\mathbf{v}^0 = (\mathbf{u}^0(t), \mu^0(t))$ satisfy the **KKT** conditions for problem (15)-(16) with $\tilde{J} = J_0$. This system can be written in the following compact notation

$$F(\mathbf{u}, t) = 0, \quad t \in [t_0, t_f] \tag{17}$$

where, $\mathbf{v} = (\mathbf{u}^0(t), \mu^0(t))$. In order to apply Newton's method to system (17), we have to solve a linear system with $D_{\mathbf{u}} F(\mathbf{v}(t), t)$ as matrix.

Algorithm:
Choose intial control trajectory $u^0, t \in [t_0, t_f], \beta^0, k = 0$.
Do:

Original initialisation $x^k(0) = x_0$

Original sweep $t : t_0 \to t_f$
Integrate forward $\dot{x}^k = f(x^k(t), u^k(t), t)$.

Adjoint initilisation Set $\lambda^k(t_f) = \phi_x^\top(t_f)$.

Adjoint sweep $t : t_f \to t_0$
Integrate backward $\dot{\lambda}^\top = -\frac{\partial H}{\partial x}$

Final sweep Solve the following parametric optimization problem to get $\delta u, t \in [t_0, t_f]$
$\delta u^k(t) = \arg\min_{\delta u(t) \in U - u} G_k(u + \delta u, t) + \beta \| \delta u(t) \|^2$

$u^{k+1}(t) = u^k(t) + \delta u^k(t)$.

$k = k + 1$

$\beta^{k+1} = 2\beta^k$

while: $\| \delta u^k(t) \| \geq Tol$ *and* $k < MAXITER$

Example 1 *([5])*

Determine the optimal mixing policy of two catalysts along the length of a tubular plug flow reactor involving several reactions. The nonlinear model that describes the reactions is:

$$\dot{x}_1(t) = u(t)(10x_2(t) - x_1(t)) \tag{18}$$

$$\dot{x}_2(t) = u(t)(x_1(t) - 10x_2(t)) - (1 - u(t))x_2(t) \tag{19}$$

Initial conditions for (18,19) are $x_1(0) = 1$ *and* $x_2(0) = 0$. *The control variable* u *represents the mixing ratio of the catalysts and must satisfy the bounds.*

$$0 \leq u(t) \leq 1$$

The problem is to minimize

$$\mathcal{J} = -1 + x_1(t_f) + x_2(t_f), \quad t_f = 1$$

The adjoint equations are determined from the Hamiltonian function,

$H(x(t),u(t),\psi(t)) = \psi_0(u(t)(10x_2(t) - x_1(t)) + \psi_1(u(t)(x_1(t) - 10x_2(t)) - (1 - u(t))x_2(t))$

as

$$\dot{\psi}_0(t) = \psi_0 - \psi_1 u(t)$$

$$\dot{\psi}_1(t) = u(t)(-10\psi_0 + 9\psi_1) + \psi_1$$

With $u^{(0)}(t) = 1.0, t \in [0, 1]$, *the value of the optimal cost for is equal* $\mathcal{J} = -0.0482226380432$. *The control function is shown in figure (1).*

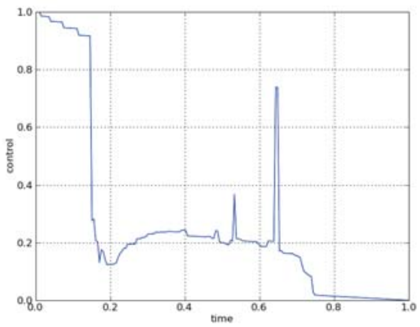

Figure 1: control function

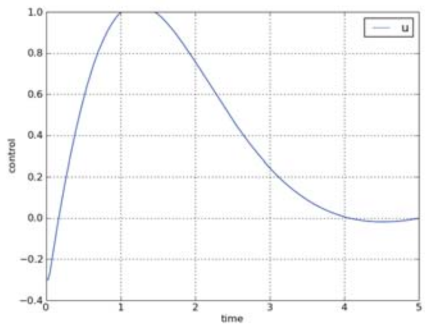

Figure 2: control function

Example 2 *This is a well-known problem which has been studied by several authors([5]). The dynamic optimization problem is to minimize*

$$\mathcal{J} = x_3(t_f)$$

subject to

$$\dot{x}_1(t) = (1 - x_2^2)x_1 - x_2 + u \tag{20}$$

$$\dot{x}_2(t) = x_1 \tag{21}$$

$$\dot{x}_3(t) = x_1^2 + x_2^2 + u^2 \tag{22}$$

The adjoint equations are determined from the Hamiltonian function,

$H(x(t),u(t),\psi(t)) = \psi_0((1 - x_2^2)x_1 - x_2 + u) + \psi_1(x_1) + \psi_1(x_1^2 + x_2^2 + u^2)$ *as*

$$\dot{\psi}_0(t) = -\psi_0((1 - x_2^2)) - \psi_1 - 2x_1\psi_2$$

$$\dot{\psi}_1(t) = \psi_0(2x_1x_2 + 1) - 2x_2\psi_2$$

$$\dot{\psi}_2(t) = 0$$

and the algebraic relation that must be satisfied is

$$\frac{\partial H}{\partial u} = \psi_0(t) + 2\psi_2(t)u = 0.$$

with $-3.0 \leq u \leq 1.0$ *and initial and terminal conditions* $x(0) = [0\ 1\ 0]^\top, u^{(0)}(t) = -0.01, t \in [0, 5]$, $\psi_0(5) = 0, \psi1(5) = 0, \psi_2(5) = 1$ *the value of the optimal cost for is equal* $\mathcal{J} = 2.92015428748$. *The control and state functions are shown in figure (2, 3).*

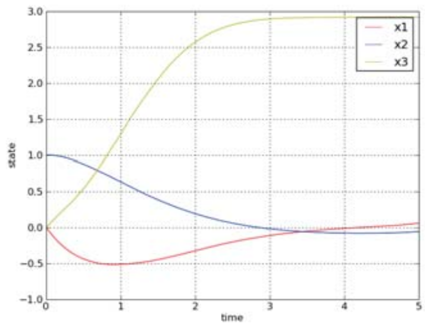

Figure 3: state function

4 Conclusions and Future Work

We have examined parametric optimal control problems. We have shown that the parametric optimization technique can be applied in maximizing the Hamiltonian under some assumptions. Two numerical examples have been introduced as well. We aim in the future to use the second order method as the next step to handle the same problem.

References

[1] A.E. Bryson and Y.C. Ho, Applied Optimal Control: Optimization, *Washington, DC: Hemisphere* (1975).

[2] A. Radwan, O. Vasilieva, R. Enkhbat, A. Griewank and J. Guddat, Parametric Approach to Optimal Control, *to appear in Journal of Optimization Letter.*

[3] H.Th. Jongen, P. Jonker, F. Twilt, Critical Sets in Parametric Optimization, *Math. Program.* 34 (1986) 333-353.

[4] J. Guddat, F. Guerra and H.Th. Jongen, Parametric Optimization: Singularities *Pathfollowing and Jumps. E.G. Teubner. Stuttgard and John Wiley, Chichester* (1990).

[5] V. S. Vassiliadis, J. R. Banga, and E. Balsa-Canto, "Second-order sensitivities of general dynamic systems with application to optimal control problems," *Chemical Engineering Science*, vol. 54, pp. 3851-3860, 1999.

ISBN: 978-988-18210-6-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)