DYNAMIC PROGRAMMING AND BELLMAN'S PRINCIPLE

Piermarco Cannarsa
Università di Roma Tor Vergata, Italy

Keywords: dynamic programming, Hamilton-Jacobi equation, optimal control, calculus of variations, linear quadratic regulator, optimal feedback control, optimal feedback synthesis, viscosity solutions, Riccati equation

Contents

Summary

Dynamic programming is a method that provides an optimal feedback synthesis for a control problem by solving a nonlinear partial differential equation, known as the Hamilton-Jacobi- Bellman equation. The main features of such a method are here described and applied to classical examples coming from calculus of variations and linear quadratic optimal control.

1. Introduction

Optimal control problems occupy a very special position in optimization theory. In fact, they represent natural examples of infinite dimensional optimization problems, even when referred to models with finitely many degrees of freedom.

In control theory one is given a system-usually described by differential equations-that can be influenced by an external action. In optimal control problems, such an action is to be exercised for minimizing a given cost functional. The cost functionals of interest may be of very different nature. In general, they may depend on the state of the system, on the control, and possibly on the system history during a given time interval. A control is *optimal* if the resulting evolution of the system minimizes the cost.

Constructing optimal controls and providing methods to compute them is the main goal of this theory that finds its motivations in applications. The areas of possible application can be very different, such as mechanical systems (see *Nonconvex Variational Problems*), space aircraft navigation (see *Optimization and Control of Distributed Processes*), population dynamics and economical models (see *Decision Analysis*).

Like in all optimization theory, one of the main tools for detecting minimum points–in this case, optimal controls–consists of necessary conditions of differential type. These conditions, that are usually referred to as the Pontryagin Maximum Principle, have been extensively studied and widely applied to all sort of different situations.

For optimal control problems, however, there is also an indirect method to derive necessary and sufficient optimality conditions. This method was called *dynamic programming* by its inventor R. Bellman. Developing the right theoretical framework to sort out the actual applicability of the dynamic programming method has taken several years, but nowadays the use of dynamic programming has become completely rigorous. At the same time, light has been spread on its drawbacks. In any case, dynamic programming remains a strikingly powerful bridge between two apparently unrelated branches of science, namely optimal control and partial differential equations.

2. Optimal Control

The basic object in the control theory of ordinary differential equations is the initial value problem

$$\begin{cases} \dot{y}(t) = b(t, y(t), \alpha(t)) \quad t \in [0,T] \ a.e. \\ y(s) = x \end{cases} \tag{1}$$

where $T > 0$ is given, $(s, x) \in [0, T] \times \mathbb{R}^n$, and where $\alpha(\cdot)$ is a measurable function with values in a given closed subset of a Euclidean space, A, that is called the *control space*. One usually refers to T as the time *horizon*, to (1) as the *state equation*, and to the elements $a \in A$ as the *controls*. To avoid any danger of confusion, the function $\alpha(\cdot)$ should be called a *control strategy* although the term ‘control’ is often used for $\alpha(\cdot)$ as well.

Mild conditions on the vector field $b : [0, T] \times \mathbb{R}^n \times A \to \mathbb{R}^n$ ensure that, for every initial condition $(s, x) \in [0, T] \times \mathbb{R}^n$ and every control strategy a, the solution of problem (1) is uniquely determined. Such a solution is called the *trajectory* starting at (s, x) with control α–hereafter abbreviated to $y_\alpha^{s,x}(\cdot)$–and may be subject to further constraints, but here this aspect of the theory will be put in the shade to unburden the presentation.

There are two main purposes for controlling system (1). One is to solve a positional problem: we want to steer the state from its initial configuration x to a given final target by the choice of $\alpha(\cdot)$; the other is to optimize the performance of the system. In *optimal control*, attention is focused on the latter task and so we shall follow this line neglecting the positional problem for the time being. Then, we shall consider the following problem: for any $(s, x) \in [0, T] \times \mathbb{R}^n$, minimize the *cost functional*

$$J^{s,x}(\alpha) = f(y_\alpha^{s,x}(T)) + \int_s^T L(t, y_\alpha^{s,x}(t), \alpha(t))dt \tag{2}$$

over all control strategies $\alpha(\cdot)$. Here, $f: \mathbb{R}^n \to \mathbb{R}$ and $L : [0, T] \times \mathbb{R}^n \times A \to \mathbb{R}$ are two given functions, the *final* and the *running cost* respectively. A control strategy $\alpha(\cdot)$ is said to be optimal if the above functional attains a minimum at α.

To account for the variety of problems that fit in the above model and better explain the techniques we are going to introduce, looking at a few examples is now in order.

Example 0.1 (Calculus of Variations)

A central role in classical mechanics is played by the minimization of the action integral

$$\int_0^T L(t, y(t), \dot{y}(t))dt \tag{3}$$

over all absolutely continuous arcs $y : [0, T] \to \mathbb{R}^n$ with fixed end-points. This problem has given rise to the fascinating theory of calculus of variations.

Let us consider the problem obtained replacing the initial time 0 with a generic time s and the terminal constraint on $y(\cdot)$ with a terminal cost, that is

$$\text{minimize} \quad f(y(T)) + \int_s^T L(t, y(t), \dot{y}(t))dt \qquad \text{subject to } y(s)=x. \tag{4}$$

This is the so-called simplest problem in calculus of variations. Notice that the above functional has the form (1)-(2) for $A = \mathbb{R}^n$ and $b(t, x, a) = a$.

Example 0.2 (The Linear Quadratic Regulator Problem)

A large number of control systems of interest to engineering are of the form

$$\begin{cases} \dot{y}(t) = B(t)y(t) + C(t)\alpha(t) & t \in [s,T] \quad a.e. \\ y(s) = x \end{cases} \tag{5}$$

with given matrices $B(t)$ and $C(t)$ of dimensions $n \times n$ and $n \times m$ respectively. This is a typical example of a state equation of type (1) with $\mathbb{R}^m$ as the control space. The Linear Quadratic Regulator (LQR) problem consists in minimizing the quadratic cost functional

$$J^{s,x}(\alpha) = y_\alpha^{s,x}(T) \cdot D y_\alpha^{s,x}(T) + \int_s^T \left[y_\alpha^{s,x}(t) \cdot M(t) y_\alpha^{s,x}(t) + a(t) \cdot N(t) a(t) \right] dt. \tag{6}$$

Notice that the above functional is of type (2) for the following choice of coefficients:

$$f(x) = x \cdot Dx\,, \qquad L(t, x, a) = x \cdot M(t)x + a \cdot N(t)a\,, \tag{7}$$

where we have denoted by $p \cdot q$ the scalar product of any two vectors $p, q \in \mathbb{R}^n$. Typical nondegeneracy assumptions on J require D and $M(t)$ to be nonnegative definite, symmetric $n \times n$ matrices and $N(t)$ to be a symmetric, positive definite $m \times m$ matrix.

Three are the basic theoretical issues in optimal control:

- Prove the existence of optimal controls;
- Derive necessary conditions that must be satisfied by any optimal control;
- Provide sufficient conditions for a control-trajectory pair $\{a, y\}$ to be optimal.

Extensively though each of these issues may have been investigated, there is still plenty of room for open questions and research work to be done before the subject can be fully understood. As anticipated in the introduction, the dynamic programming method aims at providing sufficient optimality conditions, thus addressing the third issue above.

-
-
-

Bibliography

Bardi M. and Capuzzo Dolcetta I. (1997). *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, 570pp. Boston, MA, USA: Birkhäuser. [This presents a comprehensive description of the viscosity solution approach to deterministic optimal control problems and differential games.]

Bellman R. (1957). *Dynamic Programming*, 342 pp. Princeton, NJ, USA: Princeton University Press. [This represents a pioneering text discussing the dynamic programming method.]

Clarke F.H., Ledyaev Yu.S., Stern R.J. and Wolenski P.R. (1998). *Nonsmooth Analysis and Control Theory*, 276 pp. New York, NY, USA: Springer. [This presents a handy introduction to the nonsmooth analysis approach to optimal control problems.]

Crandall M.G., Ishii H. and Lions P.L. (1992). User's Guide to Viscosity Solutions of Second Order Partial Differential Equations. *Bull. Amer. Math. Soc.* **27**, 1-67. [This is an updated exposition of the viscosity solution approach to Hamilton-Jacobi equations.]

Fleming W.H. and Rishel W.R. (1975). *Deterministic and Stochastic Optimal Control*, 222 pp. Berlin, Germany: Springer. [This presents a comprehensive yet simple introduction to optimal control problems discussing some aspects of dynamic programming as they were perceived before the introduction of viscosity solutions.]

Fleming W.H. and Soner H.M. (1993). *Controlled Markov Processes and Viscosity Solutions*, 428 pp. Berlin, Germany: Springer. [This presents a comprehensive description of the viscosity solution approach to stochastic optimal control problems and viscosity solutions to second order Hamilton-Jacobi equations.]

Li X. and Yong J. (1995). *Optimal Control Theory for Infinite Dimensional Systems*, 448 pp. Boston, MA, USA: Birkhäuser. [This contains a detailed treatment of infinite dimensional problems with plenty of examples coming from partial differential equations, delay equations and applications.]

Lions P.L. (1982). *Generalized Solutions of Hamilton-Jacobi Equations*, 317 pp. Boston, MA, USA: Pitman. [This presents the state-of-the art of Hamilton-Jacobi theory before viscosity solutions were developed.]

Subbotin A.I. (1995). *Generalized Solutions of First Order PDEs: The Dynamic Optimization Perspective*, 312 pp. Boston, MA, USA: Birkäuser. [This represents an alternative approach to deterministic optimal control problems and differential games.]

Young L.C. (1980). *Calculus of Variations and Optimal Control Theory* (second edition), 337 pp. New York, NY, USA: Chelsea Publishing Company. [This is a classic on calculus of variations; various interesting examples are here discussed.]

Biographical Sketch

Piermarco Cannarsa, who received his education at the Scuola Normale Superiore di Pisa, is Professor in Mathematical Analysis at the University of Rome Tor Vergata since 1990. Cannarsa's scientific interests include partial differential equations, functional analysis and control theory. His current research is concerned with the regularity of viscosity solutions to Hamilton-Jacobi-Bellman equations, singularities of semiconcave functions in finite and infinite dimensions, and controllability of nonlinear evolution equations. He authored more than 50 papers, which are published in international journals, and the monograph entitled *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Cannarsa served as an Associate Editor of the *Journal of Mathematical Systems, Estimation and Control* and of the *SIAM Journal on Control and Optimization*. Cannarsa gave lectures at many Italian and foreign universities, including Scuola Normale Superiore di Pisa, Ecole Normale Superieure de Paris and Princeton University. He was an invited speaker at several international conferences, including the *1998 Fourth SIAM Conference on Control and Its Applications* and the *2003 Congress of the Italian Mathematical Society*. Piermarco Cannarsa is vice-president of the Istituto Nazionale di Alta Matematica "Francesco Severi".