EE5110/EE6110: Special Topics in Automation and Control
Segment C: Control Optimization

1. Why optimal control?

Let's recall what we have learnt about how to design a control system in the past. If you are given a system without a mathematical model, what would you do?

- The PID controller should be tried first.
- If the performance of the PID controller is not good enough, then try to use physical laws and identification techniques to build a model.
- Once a model is available, there are many model-based approaches. For example, you can build a reference model which can meet all the performance requirements and then you can try to use pole-placement controller to make the transfer function of the closed-loop match the desired reference model.

In the pole-placement design, the attention is always focused upon the desired response of the closed-loop system. For instance, we build a state feedback controller in the form of

$$u(k) = -kx(k)$$

And then we just plug it into the equation of the system to analyze the closed-loop response. We never asked the question if the control input can be supplied by the actuator or what the cost of the control is. However, in reality, we must weigh the cost of undesirable performance of the system against the cost of control. This is one of the basic facts of control theory: nothing for nothing, or No Free Lunch.

How do we evaluate and compare the various consequences? There are many different approaches, none uniformly satisfactory. Each control process must be examined on its own merits. In what follows, we will consider one general approach that has certain advantages that will be indicated.

Consider the inverted pendulum example, whose linearized model is

$$\ddot{y} - y = u \qquad (1.1)$$

where y is the angle of the pendulum, and u is the torque (input) produced by the motor. If feedback control is used, then

$$u = g(y, \dot{y}) \qquad (1.2)$$

and

$$\ddot{y} - y = g(y, \dot{y}) \qquad (1.3)$$

and we would like to choose the function $g(y, \dot{y})$, the feedback control law, in such a way that y (the angle) and $\dot{y}$ (the rotation speed) approach zero reasonably rapidly without violating the condition that |y| and $|\dot{y}|$ must not be too large (otherwise, the linearization principle might not hold). This is clearly not a precise analytic formulation.

It can be made precise in several ways at the expense of introducing analytic difficulties of a higher order of difficulty than those we care to encounter at this stage.

Let us then compromise our original aims, the constant lot of the research scientist, and ask for a control law that makes y and $\dot{y}$ small on the average. For example, let us try to determine the function g so that the quantity

$$\int_0^T (y^2 + \dot{y}^2)dt \qquad (1.4)$$

is small. This formulation possesses certain merits, notably analytic simplicity.

We have, however, omitted to impose any cost of control. Different types of control laws require different amounts of effort to implement. How should we estimate this type of cost? One way is to use some average cost, such as

$$\int_0^T u^2 dt = \int_0^T g^2(y, \dot{y})dt \qquad (1.5)$$

If we add the two costs together, we have a total cost of

$$\int_0^T [y^2 + \dot{y}^2 + g^2(y, \dot{y})]dt \qquad (1.6)$$

How do we choose the function g so as to minimize this total cost?

We can now proceed to study questions of this nature in two ways, directly using dynamic programming, or be means of the calculus of variations using a simple artifice. We will discuss the basic ideas of both approaches in this class. Due to time constraint, only scalar case will be discussed in the class. Once the basic concepts are clear, the students can easily refer to advanced textbook to get more technical details to deal with higher dimensional systems.

Let us prepare the way for the calculus of variations by indicating how we can simplify the foregoing formulation. The desired function $g(y, \dot{y})$ is a function of time t, through the dependence of y and $\dot{y}$ upon t. Why not then write

$$\ddot{y} - y = u \qquad (1.7)$$

And determine the input u(t) by the condition that it minimizes the functional

$$\int_0^T [y^2 + \dot{y}^2 + u^2]dt \text{ ?} \qquad (1.8)$$

It is essential to include the initial conditions $y(0) = c_1, \dot{y}(0) = c_2$. Hopefully, once u(t) has been determined, we can interpret u(t) as a function of the state y(t) and $\dot{y}(t)$, and thus obtain the desired feedback law $u = g(y, \dot{y})$. This simplification possesses the overwhelming merit of allowing us to deal with linear differential equations such as (1.7) rather than nonlinear differential equations such as (1.3), in the description of the behavior of the system over time.

It turns out that it is wise initially to consider the simplest system, one described by a single state variable, a displacement y(t). Once we have carried through the analysis in this case, we can turn to the general case. It is always a good strategy by considering the simplest system first.

We consider then a system S described by a state variable y(t) satisfying the equation

$$\frac{dy}{dt} = f(y,u), y(0) = c, \qquad (1.9)$$

where the control variable u(t) is to be chosen so as to minimize the functional

$$J(y,u) = \int_0^T h(y,u)dt \qquad (1.10)$$

where h(y, u) is a prescribed function. This is still too complex a problem for a suitable introduction to the analytic aspects of control theory. Consequently, we use a linear approximation,

$$\frac{dy}{dt} = ay + u, y(0) = c, \qquad (1.11)$$

and a quadratic approximation of the cost function

$$J(y,u) = \int_0^T [y^2 + u^2]dt \qquad (1.12)$$

Finally, we begin by taking a = 0. We thus have the specific problem of minimizing the functional

$$J(y) = \int_0^T [y^2 + \dot{y}^2]dt \qquad (1.13)$$

where y(0) = c.

This will be our starting point. Once the basic ideas of the solution of this problem are made clear, it is not difficult to consider classes of more general variational problems. However, it is not to be anticipated that any all-purpose theory capable of handling completely general variational problems will ever be developed. The reason is quite simple. We face all the difficulties of treating general nonlinear differential equations.

What we can hope to do is to develop a toolkit of techniques that can be used singly, in unison, and together with an electronic computer to handle broader and broader classes of problems.

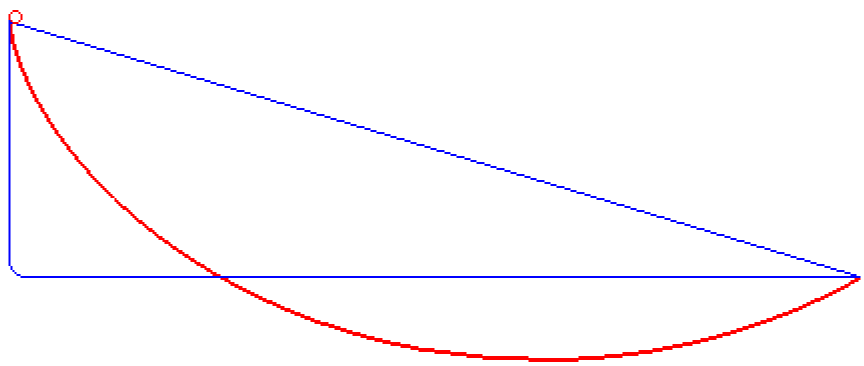
Before we move on to give the solution of optimal control using calculus of variations, let's give a brief introduction about the historical background of calculus of variations.

2. Historical Background of Calculus of Variations

The calculus of variations is a subject whose beginning can be precisely dated. It might be said to begin at the moment that Euler coined the name calculus of variations but this is, of course, not the true moment of inception of the subject. It would not have been unreasonable if I had gone back to the set of isoperimetric problems considered by Greek mathematicians such as Zenodorus (c. 200 B. C.) and preserved by Pappus (c. 300 A. D.).

I have not done this since these problems were solved by geometric means. Instead I have arbitrarily chosen to begin with Fermat's elegant principle of least time. He used this principle in 1662 to show how a light ray was refracted at the interface between two optical media of different densities. This analysis of Fermat seems to me especially appropriate as a starting point: He used the methods of the calculus to minimize the time of passage of a light ray through the two media, and his method was adapted by John Bernoulli to solve the brachistochrone curve problem in 1696.

In mathematics and physics, a **brachistochrone curve** (from Ancient Greek βράχιστος χρόνος *(brákhistos khrónos)*, meaning "shortest time"), or curve of fastest descent, is the one lying on plane between a point *A* and a lower point *B*, where *B* is not directly below *A*, on which a bead slides frictionlessly under the influence of a uniform gravitational field to a given end point in the shortest time. Incidentally, for a given starting point, the brachistochrone curve is the same as the tautochrone curve. More specifically, the solution to the brachistochrone and tautochrone problem are one and the same, the cycloid



It immediately occupied the attention of Jakob Bernoulli and the Marquis de l'Hôpital, but Leonhard Euler first elaborated the subject. His contributions began in 1733, and his Elementa Calculi Variationum gave the science its name. Lagrange contributed extensively to the theory, and Legendre (1786) laid down a method, not entirely satisfactory, for the discrimination of maxima and minima. Isaac Newton and Gottfried Leibniz also gave some early attention to the subject. To this discrimination Vincenzo Brunacci (1810), Carl Friedrich Gauss (1829), Siméon Poisson (1831), Mikhail Ostrogradsky (1834), and Carl Jacobi (1837) have been among the contributors. An important general work is that of Sarrus (1842) which was condensed and improved by Cauchy (1844). Other valuable treatises and memoirs have been written by Strauch (1849), Jellett (1850), Otto Hesse (1857), Alfred Clebsch (1858), and Carll (1885), but perhaps the most important work of the century is that of Weierstrass. His celebrated course on the theory is epoch-making, and it may be asserted that he was the first to place it on a firm and unquestionable foundation. The 20th and the 23rd Hilbert problem published in 1900 encouraged further development. In the 20th century David Hilbert, Emmy Noether, Leonida Tonelli, Henri Lebesgue and Jacques Hadamard among others made significant contributions. Marston Morse applied calculus of variations in what is now called Morse theory. Lev Pontryagin, Ralph Rockafellar and F. H. Clarke developed new mathematical tools for the calculus of variations in optimal control theory. The dynamic programming of Richard Bellman is an alternative to the calculus of variations.

3. Optimal Control using Calculus of Variations

3.1. Introduction

We wish to study the problem of minimizing the functional

$$J(y) = \int_0^T [y^2 + \dot{y}^2]dt \qquad (3.1.1)$$

using the methods of the calculus of variations. We will first consider the case where y(0)=c and there is no terminal condition, and then the cases where there are conditions at both boundaries, such as y(0)= c_1.,y(T)= c_2.

Following this, we shall consider the more general problem of minimizing the functional

$$J(y,u) = \int_0^T [y^2 + u^2]dt \qquad (3.1.2)$$

where the functions y and u are connected by the linear differential equation

$$\dot{y} = ay + u, y(0) = c. \qquad (3.1.3)$$

It is noted that (3.1.1) is the special case of (3.1.2) when a=0 in equation (3.1.3). We shall take T to be finite unless specifically stated otherwise, and all functions and parameters appearing are taken to be real.

3.2. Does a Minimum Exist?

The first question we must face, generally one of the thorniest in the calculus of variations, is that of determining the class of functions that we allow as candidates for the minimization of the functional

$$J(y) = \int_0^T [y^2 + \dot{y}^2]dt \qquad (3.2.1)$$

Do we wish to consider all possible functions y(t) defined over [0, T], satisfying the condition y(0)=c? Obviously the answer is negative. Our fundamental restriction is that y(t) is a function with the property that it possesses a derivative y'(t) whose square is integrable over [0, T]. This fact is occasionally written: $y' \in L^2[0,T]$. A consequence of this is that y(t), as the integral of y'(t), is also integrable over [0, T]: $y \in L^2[0,T]$.

In order to keep the presentation on the desired elementary path, we are going to follow a very circumscribed route. Rather than tackle the problem of existence of a minimizing function directly, we will first derive a necessary condition that any such function must satisfy. This necessary condition is the celebrated differential equation of Euler.

It is not difficult, as we show, to establish the fact that this equation possesses a unique solution. Next we demonstrate, by means of a simple direct calculation, that the function obtained in this fashion yields the absolute minimum of J(y) for all admissible y.

This circuitous approach enables us to provide a completely rigorous discussion of the variational problem with a minimum of analytical background. Furthermore, it provides a rigorous foundation for the dynamic programming approach. Finally, all the steps in the argument have been carefully constructed so as to generalize to the multidimensional case, granted a modicum of vector-matrix manipulation.

3.3. The Euler Equation

First of all, how do we solve the optimization problem for a scalar function f(x)?

A simple argument will go like this. Assume that x_o is the optimal point, then the values of its neighborhood can be expressed by Taylor series y_o

$$f(x_o+\Delta x)=f(x_o)+f'(x_o)\Delta x+\frac{f''(x_o)}{2}\Delta x^2+\cdots$$

If $f(x_o)$ is the optimal value, then its first order derivative must be zero. Otherwise, it is very easy to obtain values either bigger or smaller than $f(x_o)$.

From above argument, it is key to express the values of the cost function in the neighborhood of the optimal solution. We will follow the similar idea to find out the solution to our optimal control problem.

Let y_o denote a hypothetical solution to the minimization of the functional J(y),

$$J(y)=\int_0^T[y^2+\dot{y}^2]dt\,. \qquad (3.3.1)$$

Just as in the scalar case, we need to compute the variation of the optimal value $f(x_o)$ in the neighborhood of x_o, we need to compute the variations of J(y) in the neighborhood of the optimal solution y_o. But y_o is a function now in the infinite-dimensional function space, how to evaluate its neighborhood?

Since there are infinite directions around y_o, we just choose anyone of them, and try to compute the variations along the direction of the chosen function. Let z(t) denote any function of t with the property that J(z) exists.

The next step is critical to make the problem tractable. Instead of evaluating $\mathrm{J}(\mathrm{y}_o+\mathrm{z})$ directly, we put a constant parameter to control the size of the change in the direction of z. Take ε to be a scalar parameter and consider the expression

$$J(y_0+\varepsilon z)=\int_0^T[(y_0+\varepsilon z)^2+(\dot{y}_0+\varepsilon\dot{z})^2]dt \qquad (3.3.2)$$

as a function of the scalar variable ε for any fixed direction z(t). If y_o is indeed the function that yields the absolute minimum of J(y), then $J(y_0+\varepsilon z)$ as a function of the scalar parameter ε must have an absolute minimum at ε = 0. This is the key step in the argument of calculus of variations. The trick is to convert the problem of minimizing a functional into a problem of minimizing a normal function $f(\varepsilon)$.

Hence, using the usual result of calculus, a necessary condition is

$$\frac{d}{d\varepsilon} J(y_0 + \varepsilon z)\Big|_{\varepsilon=0} = 0 \qquad (3.3.3)$$

This actually does not require calculus since $J(y_0 + \varepsilon z)$ is a quadratic in ε,

$$J(y_0 + \varepsilon z) = \int_0^T (\dot{y}_0{}^2 + y_0{}^2)dt + 2\varepsilon \int_0^T (\dot{y}_0 \dot{z} + y_0 z)dt + \varepsilon^2 \int_0^T (\dot{z}^2 + z^2)dt$$

We see then that the variational condition derived from (3.3.3) is

$$\int_0^T (\dot{y}_0 \dot{z} + y_0 z)dt = 0 \qquad (3.3.4)$$

for all z such that J(z) exists.

If we knew that y had a second derivative, we would proceed in the following fashion. Integrating by parts, we obtain from (3.3.4) the relation

$$\begin{aligned} &\dot{y}_0 z\Big|_0^T + \int_0^T z(-\ddot{y}_0 + y_0)dt = 0 \\ &\dot{y}_0(T)z(T) - \dot{y}_0(0)z(0) + \int_0^T z(-\ddot{y}_0 + y_0)dt = 0 \end{aligned} \qquad (3.3.5)$$

This is not a legitimate step since we do not know that $\ddot{y}_0(t)$ exists. Let us, however, forge ahead boldly.

What is z(0)?

Since $y_0 + \varepsilon z$ as an admissible function satisfies the initial condition

$$y_0(0) + \varepsilon z(0) = c,$$

we see that z(0) = 0.

Since the left-hand side must be zero for all admissible z, we suspect that

$$-\ddot{y}_0 + y_0 = 0 \qquad (3.3.6)$$

and that $\dot{y}_0(T) = 0$. And we obtain no condition on $\dot{y}_0(0)$.

Equation (3.3.6) is the Euler equation associated with this variational problem.
Observe that the desired solution satisfies a two-point boundary condition

$$y_0(0) = \text{c}, \ \dot{y}_0(T) = 0. \qquad (3.3.7)$$

Consequently, it is not immediately obvious that there exists a function y(t) with the required properties.

The general solution of the differential equation is

$$y = c_1 e^t + c_2 e^{-t}$$

Using the boundary conditions, we have the two equations

$$c = c_1 + c_2$$

$$0 = c_1 e^T - c_2 e^{-T}$$

to determine the coefficients c_1 and c_2. Solving, we obtain the expression

$$y_o(t) = c\left(\frac{e^{t-T} + e^{-(t-T)}}{e^{-T} + e^T}\right) = c\left(\frac{\cosh(t-T)}{\cosh(T)}\right) \qquad (3.3.8)$$

where the hyperbolic cosine function cosh(t) is defined as

$$\cosh(t) = \frac{e^t + e^{-t}}{2} \qquad (3.3.9)$$

This is the unique solution of (3.3.6). It exists for all T>0 since the denominator is never zero. We have thus shown that there is one and only one function satisfying the necessary condition expressed by the Euler equation subject to the two-point boundary condition.

3.4 Minimizing property of the Solution

Let us now show that the function obtained in this fashion yields the minimum value of J. Consider

$$\begin{aligned} J(y_0 + w) &= \int_0^T [(y_0 + w)^2 + (\dot{y}_0 + \dot{w})^2] dt \\ &= \int_0^T (\dot{y}_0^2 + y_0^2) dt + 2\int_0^T (\dot{y}_0 \dot{w} + y_0 w) dt + \int_0^T (w^2 + \dot{w}^2) dt \end{aligned} \qquad (3.4.1)$$

for any function w(t) such that w(0)=0 (to ensure that y(0)+w(0)=c), and such that $w' \in L^2[0,T]$. The middle term on the right disappears upon integration by parts,

$$\int_0^T (\dot{y}_0 \dot{w} + y_0 w) dt = \dot{y}_0 w\Big|_0^T + \int_0^T w(-\ddot{y}_0 + y_0) dt = 0 \qquad (3.4.2)$$

Hence

$$J(y_0 + w) = J(y_0) + J(w) > J(y_0) \qquad (3.4.3)$$

unless w is identically zero.

Thus we have established by a direct calculation that J has a minimum value assumed by the unique solution of the Euler equation determined by (3.3.8). The fact that the Euler equation was obtained by a formal procedure is no longer of any consequence.

3.5 Alternative Approach

In section 3.3 we used the fact that the Euler equation (3.3.6) could be solved explicitly. This enabled us to verify directly that a solution existed for all T>0 and that it was unique. Let us now employ a more powerful method which is applicable to the more common cases where we do not possess an explicit analytic representation for the solution.

Let v and w be two solutions of the Euler equation (3.3.6). Then u=v-w satisfies the differential equation and the conditions u(0)=0, u'(T)=0.
Suppose that the equation

$$\ddot{u}-u=0 \quad (3.5.1)$$

possesses a nontrivial solution satisfying the boundary conditions u(0)=0, u'(T)=0. Then, multiplying by u and integrating over [0,T],

$$\int_0^T u(\ddot{u}-u)dt=0 \qquad (3.5.2)$$

Integrating by parts, this yields

$$\int_0^T u(\ddot{u}-u)dt=u\dot{u}\,|_0^T-\int_0^T(\dot{u}^2+u^2)dt=0\ . \quad (3.5.3)$$

The integrated terms vanish due to the boundary conditions, and we are left with the relation

$$\int_0^T(\dot{u}^2+u^2)dt=0\,. \qquad (3.5.4)$$

This can hold only if u and u' are identically zero.

3.6 Asymptotic Control

In the calculus of variations, emphasis is customarily placed upon establishing the existence of a solution over some T-interval, however small. In control theory, it is natural to wish to examine the nature of the solution for all T>0. In particular, we are interested in the behavior of the control law as $T\to\infty$. Just as in the study of descriptive processes, we can expect a great deal of simplification as a steady-state regime takes over. As we shall see, this is indeed the case.

Let us examine the explicit form of the solution given in (3.3.8)

$$\begin{aligned} y_o(t)&=c\frac{\cosh(t-T)}{\cosh(T)}=c\Big(\frac{e^{t-T}+e^{-(t-T)}}{e^{-T}+e^{T}}\Big)\\ &=\frac{ce^{-t}}{1+e^{-2T}}+\frac{ce^{t-2T}}{1+e^{-2T}}\end{aligned} \qquad (3.6.1)$$

If $0\le t\le T$, the second term on the right is uniformly bounded by the quantity $|c|e^{-T}$. Hence,

$$\left| y_o(t) - \frac{ce^{-t}}{1+e^{-2T}} \right| \leq |c|\, e^{-T} \qquad (3.6.2)$$

And thus

$$\left| y_o(t) - ce^{-t} \right| \leq 2\,|c|\, e^{-T}$$

for $0 \leq t \leq T$. From the expression

$$\dot{y}_o(t) = c\,\frac{\sinh(t-T)}{\cosh(T)}, \qquad (3.6.3)$$

we obtain the same way

$$\begin{aligned} \left| \dot{y}_o(t) + \frac{ce^{-t}}{1+e^{-2T}} \right| &\leq |c|\, e^{-T}, \\ \left| \dot{y}_o(t) + ce^{-t} \right| &\leq 2\,|c|\, e^{-T}. \end{aligned} \qquad (3.6.4)$$

Thus, if T is large, we have the approximate relation

$$\dot{y}_0(t) \cong -y(t) \qquad (3.6.5)$$

Uniformly valid in $0 \leq t \leq T$. Since the control input u(t)=y'(t), it implies the simple feedback control law

$$u(t) = -y(t) \qquad (3.6.6)$$

This is exactly the type of simple control law we were seeking.

We have thus succeeded in converting a control process that was time-dependent, in the sense that we considered y'(t) as a function of time t, to a control process that is state-dependent, y'(t) depending on y. We may not always want to do this. What we do want is the flexibility to obtain whatever analytic representation of optimal control that is most convenient for the subsequent mathematical or engineering applications.

3.7 Infinite Control Process

Suppose we wish to consider a control process over an unbounded time interval. Take, for example, the problem of minimizing the functional

$$J(y) = \int_0^\infty [y^2 + \dot{y}^2]\,dt \qquad (3.7.1)$$

where again y(0)=c.

The first point to mention is that the class of admissible functions has changed. We must impose the additional restrictions that $\int_0^\infty y^2 dt$ and $\int_0^\infty \dot{y}^2 dt$ converge.

The second point to clarify is that of determining whether a solution to the foregoing minimization problem can be obtained as a limit of the solution to the finite control process as $T \to \infty$. In this case, it can be easily shown that the control law is exactly,

$$u(t) = \dot{y}(t) = -y(t) \quad (3.7.2)$$

3.8 The Minimum Value of J(u)

The minimum value of the cost function J(y) is then

$$\min_y J(y) = \int_0^T [y^2 + \dot{y}^2]dt = c^2 \tanh(T) = c^2 \frac{e^T - e^{-T}}{e^T + e^{-T}} \qquad (3.8.1)$$

We see that

$$\lim_{T\to\infty} \min_y J(y) = c^2 \qquad (3.8.2)$$

3.9 More General Quadratic Variational Problems

We have previously indicated that a more general form of the kind of variational problem that arises in control theory is that of minimizing the functional

$$J(y,u) = \int_0^T h(y,u)dt \qquad (3.9.1)$$

over all y and u connected by the differential equation

$$\frac{dy}{dt} = f(y,u), y(0) = c, \qquad (3.9.2)$$

In order to obtain some ideas as to how to proceed in the general case, we begin by considering the prototype problem

$$J(y,u) = \int_0^T [y^2 + u^2]dt \qquad (3.9.3)$$

where

$$\dot{y} = ay + u, y(0) = c. \qquad (3.9.4)$$

Previously, we considered the case a = 0.

There are several effective ways in which we can avoid the introduction of any new ideas and methods. For instance, we can solve for u in (3.16) and consider the more familiar problem of minimizing

$$J(y) = \int_0^T [y^2 + (\dot{y} - ay)^2]dt \qquad (3.9.5)$$

This can be reduced to the form

$$J(y) = \int_0^T [(1+a^2)y^2 + \dot{y}^2]dt - 2a\int_0^T \dot{y}y dt = \int_0^T [(1+a^2)y^2 + \dot{y}^2]dt - ay^2(T) + ac^2, \quad (3.9.6)$$

a type of functional already discussed.

Such method possesses certain advantages and should by no means be scorned. There is, however, an advantage in considering the problem in its original form, an advantage that becomes considerable when multidimensional, and still more general, control processes are studied.

3.10 Variational Procedure

Following the same general methods used in the discussion of the simpler question, let y_0, u_0 be a minimizing pair, assumed to exist. Write

$$y = y_0 + \varepsilon w,\ \ u = u_0 + \varepsilon z, \qquad (3.10.1)$$

where ε is an arbitrary real parameter, w(0) = 0, and w and z are otherwise arbitrary in the sense previously described.

Writing

$$J(y_0 + \varepsilon w, u_0 + \varepsilon z) = J(y_0, u_0) + 2\varepsilon \int_0^T (y_0 w + u_0 z)dt + \varepsilon^2 J(w, z) \qquad (3.10.2)$$

we see that the condition that y_0, u_0 furnish an absolute minimum yields the variational condition

$$\int_0^T (y_0 w + u_0 z)dt = 0 \qquad (3.10.3)$$

for all w and z. Since ($y_0 + \varepsilon w, u_0 + \varepsilon z$) must satisfy (3.9.4), we have

$$\frac{dw}{dt} = aw + z \qquad (3.10.4)$$

Hence, (3.10.3) yields

$$\int_0^T (y_0 w + u_0(\dot{w} - aw))dt = 0 \qquad (3.10.5)$$

Integrating by parts, this yields the relation

$$u_0 w\big|_0^T + \int_0^T w(y_0 - \dot{u}_0 - au_0)dt = 0$$

for all w. Hence, as above, we suspect that the variational equation is

$$\dot{u}_0 = -au_0 + y_0,\ \ u_0(T) = 0. \qquad (3.10.6)$$

Overall, the optimal solutions have to satisfy

$$\begin{aligned} &\dot{y} = ay + u, y(0) = c, \\ &\dot{u} = -au + y, u(T) = 0. \end{aligned} \qquad (3.10.7)$$

Note that it is a boundary value problem instead of an initial value problem, whose existence and uniqueness has yet to be established. The control signal u(t) obtained in this way will be open-loop control because it is a function of time. And we know that it is more desirable to obtain a feedback controller instead of open-loop controller. Then the question is:
Can we solve the optimal control using the idea of feedback?

3.6 Other applications using calculus of variations

Minimization problems that can be analyzed by the calculus of variations serve to characterize the equilibrium configurations of almost all continuous physical systems, ranging through elasticity, solid and fluid mechanics, electro-magnetism, gravitation, quantum mechanics, string theory, and many, many others. Many geometrical configurations, such as minimal surfaces, can be conveniently formulated as optimization problems. Moreover, numerical approximations to the equilibrium solutions of such boundary value problems are based on a nonlinear finite element approach that reduces the infinite-dimensional minimization problem to a finite-dimensional problem.

Just as the vanishing of the gradient of a function of several variables singles out the critical points, among which are the minima, both local and global, so a similar "functional gradient" will distinguish the candidate functions that might be minimizers of the functional. The finite-dimensional calculus leads to a system of algebraic equations for the critical points; the infinite-dimensional functional analog results a boundary value problem for a nonlinear ordinary or partial differential equation whose solutions are the critical functions for the variational problem. So, the passage from finite to infinite dimensional nonlinear systems mirrors the transition from linear algebraic systems to boundary value problems.

The best way to appreciate the calculus of variations is by introducing a few concrete examples of both mathematical and practical importance. Some of these minimization problems played a key role in the historical development of the subject. And they still serve as an excellent means of learning its basic constructions.

The minimal curve problem is to find the shortest path between two specified locations. In its simplest manifestation, we are given two distinct points, A and B,

$A=(x_a, y_a)$ and $B=(x_b, y_b)$ in plane R^2 (3.6.1)

and our task is to find the curve of shortest length connecting them. "Obviously", as you learn in childhood, the shortest route between two points is a straight line. Mathematically, then, the minimizing curve should be the graph of the particular affine function (assuming $x_a \neq x_b$, i.e. the points A, B do not lie on a common vertical line)

$$y = cx + d = \frac{(y_b - y_a)}{(x_b - x_a)}(x - x_a) + y_a \qquad (3.6.2)$$

that passes through or interpolates the two points. However, this commonly accepted "fact" — that (3.6.2) is the solution to the minimization problem — is, upon closer inspection, perhaps not so immediately obvious from a rigorous mathematical standpoint.

Let us see how we might formulate the minimal curve problem in a mathematically precise way. For simplicity, we assume that the minimal curve is given as the graph of a smooth function y(x). Then, the length of the curve is given by the standard arc length integral

$$J(y) = \int_{x_a}^{x_b} \sqrt{1 + y'(x)^2}\, dx \qquad (3.6.3)$$

where we abbreviate y'=dy/dx. The function y(x) is required to satisfy the boundary conditions

$$y(x_a) = y_a, y(x_b) = y_b \qquad (3.6.4)$$

in order that its graph pass through the two prescribed points (3.6.1). The minimal curve problem asks us to find the function y(x) that minimizes the arc length functional (3.6.3) among all "reasonable" functions satisfying the prescribed boundary conditions. The reader might pause to meditate on whether it is analytically obvious that the affine function (3.6.2) is the one that minimizes the arc length integral (3.6.3) subject to the given boundary conditions. One of the motivating tasks of the calculus of variations, then, is to rigorously prove that our everyday intuition is indeed correct.

Let us now discuss the most basic analytical techniques for solving such minimization problem. Let us concentrate on the simplest class of variational problems, in which the unknown is a continuously differentiable scalar function, and the functional to be minimized depends upon at most its first derivative. The basic minimization problem, then, is to determine a suitable function y(x) that minimizes the objective functional

$$J(y) = \int_a^b L(x, y, y')\, dx \qquad (3.6.5)$$

The integrand is known as the *Lagrangian* for the variational problem, in honor of Lagrange. We usually assume that the Lagrangian L(x,y,p) is a reasonably smooth function of all three of its (scalar) arguments x, y and p, which represents the derivative dy/dx. For example, the arc length functional (3.6.3) has Lagrangian function

$$L(x, y, p) = \sqrt{1 + p^2} \qquad (3.6.6)$$

In order to uniquely specify a minimizing function, we must impose suitable boundary conditions. All of the usual suspects — Dirichlet (fixed), Neumann (free), as well as mixed and periodic boundary conditions — are also relevant here. In the interests of brevity, we shall concentrate on the Dirichlet boundary conditions

$$y(a) = y_a, y(b) = y_b \qquad (3.6.7)$$

The (local) minimizers of a (sufficiently nice) objective function defined on a finite dimensional vector space are initially characterized as critical points, where the objective function's gradient vanishes. An analogous construction applies in the infinite-dimensional context treated by the calculus of variations. Every sufficiently nice minimizer of a sufficiently nice functional J(y) is a "critical function".

Given a path y(x), we consider a variation of this path in the direction of v(x) by $y + \varepsilon v$. Follow the same argument as in the past. Suppose y is a local minimum, then for any such variation v, $J(y + \varepsilon v)$ takes minimum at $\varepsilon = 0$. This leads to a necessary condition:

$$\left.\frac{dJ(y+\varepsilon v)}{d\varepsilon}\right|_{\varepsilon=0}=0 \qquad (3.6.8)$$

We have

$$J(y+\varepsilon v)=J(x,y+\varepsilon v,y'+\varepsilon v') \qquad (3.6.9)$$

Take derivative with respect to ε, we have

$$\frac{dJ(y+\varepsilon v)}{d\varepsilon}=\int_a^b \frac{dL(x,y+\varepsilon v,y'+\varepsilon v')}{d\varepsilon}dx \qquad (3.6.10)$$

Where

$$\frac{dL(x,y+\varepsilon v,y'+\varepsilon v')}{d\varepsilon}=\frac{\partial L}{\partial y}(x,y+\varepsilon v,y'+\varepsilon v')v+\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v')v' \qquad (3.6.11)$$

So we get

$$\frac{dJ(y+\varepsilon v)}{d\varepsilon}=\int_a^b [\frac{\partial L}{\partial y}(x,y+\varepsilon v,y'+\varepsilon v')v+\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v')v']dx \qquad (3.6.12)$$

Integrate by parts we have

$$\begin{aligned}\frac{dJ(y+\varepsilon v)}{d\varepsilon}&=\int_a^b [\frac{\partial L}{\partial y}(x,y+\varepsilon v,y'+\varepsilon v')v+\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v')v']dx\\ &=\left.v\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v')\right|_a^b+\int_a^b v[\frac{\partial L}{\partial y}(x,y+\varepsilon v,y'+\varepsilon v')-\frac{d}{dx}(\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v'))]dx\end{aligned} \qquad (3.6.13)$$

Since $y+\varepsilon v$ satisfies the boundary condition, we have

v(a)=0, and v(b)=0 (3.6.14)

So we obtain

$$\frac{dJ(y+\varepsilon v)}{d\varepsilon}=\int_a^b v[\frac{\partial L}{\partial y}(x,y+\varepsilon v,y'+\varepsilon v')-\frac{d}{dx}(\frac{\partial L}{\partial p}(x,y+\varepsilon v,y'+\varepsilon v'))]dx$$

Set $\varepsilon=0$, we have

$$\int_a^b v[\frac{\partial L}{\partial y}(x,y,y')-\frac{d}{dx}(\frac{\partial L}{\partial p}(x,y,y'))]dx=0 \qquad (3.6.15)$$

Since v(x) is arbitrary function satisfying v(a)=0 and v(b)=0, we have

$$\frac{\partial L}{\partial y}(x,y,y')-\frac{d}{dx}(\frac{\partial L}{\partial p}(x,y,y'))=0 \qquad (3.6.16)$$

This critical equation (3.6.16) is, in fact, a second order ordinary differential equation,

$$\frac{\partial L}{\partial y}(x,y,y')-\frac{\partial^2 L}{\partial x\partial p}(x,y,y')-y'\frac{\partial^2 L}{\partial y\partial p}(x,y,y')-y''\frac{\partial^2 L}{\partial p^2}(x,y,y')=0 \qquad (3.6.17)$$

known as the Euler–Lagrange equation associated with the variational problem (3.6.5), in honor of two of the most important contributors to the subject. Any solution to the Euler-Lagrange equation that is subject to the assumed boundary conditions forms a critical point for the functional, and hence is a potential candidate for the desired minimizing function. And, in many cases, the Euler–Lagrange equation suffices to characterize the minimizer without further ado.

Now we are ready to present the following theorem:

Theorem 3.1. Suppose the Lagrangian function is at least twice continuously differentiable: $L(x,y,p)\in C^2$. Then any C^2 minimizer y(x) to the corresponding functional

$J(y)=\int_a^b L(x,y,y')dx$ subject to the selected boundary conditions, must satisfy the associated Euler–Lagrange equation (3.6.16).

Let us return to the most elementary problem in the calculus of variations: finding the curve of shortest length connecting two points in the plane. As we derived earlier, such planar geodesics minimize the arc length integral

$J(y)=\int_{x_a}^{x_b}\sqrt{1+y'(x)^2}\,dx$ with Lagrangian $L(x,y,p)=\sqrt{1+p^2}$,

subject to the boundary conditions

$$y(a)=y_a, y(b)=y_b$$

Since

$\frac{\partial L}{\partial y}(x,y,y')=0$ and $\frac{\partial L}{\partial p}(x,y,y'))=\frac{p}{\sqrt{1+p^2}}$

the Euler–Lagrange equation (3.6.16) in this case takes the form

$$-\frac{d}{dx}\frac{y'}{\sqrt{1+(y')^2}}=-\frac{y''}{(1+(y')^2)^{\frac{3}{2}}}=0 \qquad (3.6.18)$$

Since the denominator does not vanish, this is the same as the simplest second order ordinary differential equation

$$y''=0 \qquad (3.\ 6.19)$$

We deduce that the solutions to the Euler–Lagrange equation are all affine functions, y=cx+d, whose graphs are straight lines. Since our solution must also satisfy the boundary conditions, the only critical function — and hence the sole candidate for a minimizer — is the straight line

$$y=\frac{(y_b-y_a)}{(x_b-x_a)}(x-x_a)+y_a \qquad (3.6.20)$$

passing through the two points. Thus, the Euler–Lagrange equation helps to reconfirm our intuition that straight lines minimize distance.

Be that as it may, the fact that a function satisfies the Euler–Lagrange equation and the boundary conditions merely confirms its status as a critical function, and does not guarantee that it is the minimizer. Indeed, any critical function is also a candidate for maximizing the variational problem, too. The nature of a critical function will be elucidated by the second derivative test, and requires some further work. Of course, for the minimum distance problem, we "know" that a straight line cannot maximize distance, and must be the minimizer. Nevertheless, the reader should have a small nagging doubt that we may not have completely solved the problem at hand . . .

You may wish to explore more technical details on the second derivatives for functional, which is beyond the scope of this introduction. It is enough for you to understand the basic idea for calculus of variations which is to consider the variation of the optimal solution in the form of $y+\varepsilon v$, and use the condition for optimal value for normal functions.

4. Dynamic Programming

4.1 Introduction

Following we will introduce an entirely different approach to the treatment of control processes. It will be based on the theory of dynamic programming, a mathematical abstraction and extension of the fundamental engineering concept of feedback control.

4.2. Control as a Multistage Decision Process

Consider the problem of minimizing the functional

$$J(y)=\int_0^T [y^2+\dot{y}^2]dt \qquad (4.2.1)$$

over all y(t) subject to the initial condition y(0) = c, a problem we completely resolved in the preceding discussion. Our approach was a straightforward extension of the ideas of calculus. The existence of a minimizing function y was assumed and the variation of J(y) in the neighborhood of this extremum provided a necessary condition, the Euler equation, which we showed was also sufficient.

We now want to employ some different techniques generated by the idea that the minimization problem can be considered to arise from a control process. In place of asking for the solution y as a function of time we ask for the most efficient control at each point in the state space. At any particular point in the history of the process, we want instructions as to what to do next.

In terms of the foregoing minimization question, at the point P = P(y, t), we want to determine the control input $u=\dot{y}$ as a function of y and t. Intuitively, we conceive of this as a guidance process in which it is required to furnish steering directions continuously.

A control process is thus considered to be a multistage decision process consisting of the following operations:

(a) At time t, the state of the system is observed.
(b) Based on this information and a control law, a decision is made.
(c) This decision produces an effect upon the system.

In order to keep the presentation simple, we assume, quite unrealistically of course, that it is always possible to observe the system and obtain complete information about its state, that no time or effort is required for this, that the correct decision is always made, that we know exactly what the effect of every decision is, and finally, that we know what the purpose of the control process is.

The mention of these more realistic aspects of control is made to give you some brief idea of the scope of control theory and of the vast amount of research that remains to be done. In particular, we wish to emphasize the grave difficulties entailed in fitting real control process into simplified molds. Nevertheless, there is considerable merit in considering the simpler versions in detail, and remarkably, on occasion the case even occurs that these naïve formulations are useful.

4.3. Preliminary Concepts

As indicated above, the direction of the path at a particular point P is to be considered dependent upon the coordinates of the point, the state y and the time, as measured in terms of time transpired or time remaining in the process. This point of view combined with the additive nature of the quadratic functional used to evaluate the control process enables us to regard each decision that is made as the initial decision of a new control process. Consequently, we can constrain our attention to the initial slope, the value $\dot{y}(0)$, and regard this as a function of c, the initial state, and T, the duration of the process. This basic function, which we denote by u(c, T), is called a <u>policy</u>.

A policy is a rule for prescribing the action to take at every possible position of the system in state space. The term is used deliberately because it agrees so closely with the intuitive concept of "policy."

The analytic structure of the policy function depends upon the way in which the system has been described in mathematical terms. There are many ways of constructing a mathematical formulation and it is not always clear what is irreducible amount of information concerning the system that is required for either optimal or effective control. Careful analysis of this type is essential in the study of large complex systems occurring in the economic, industrial, and biomedical area. It is important to emphasize that the concepts of “state” and “policy” are elusive and require continuing examination.

A policy that minimizes the functional representing the overall cost of the control process, or, generally, that most effectively achieves the desired goal, is called an optimal policy. Even though there is a unique minimum or maximum value, there need not be a unique optimal policy that attains this value. There may be several ways of carrying this out.

As we shall see, there is a very simple and intuitive characterization of optimal policies, the principle of optimality. This principle can be used to obtain functional equations that simultaneously determine the required minimum or maximum and all optimal policies.

It is of some interest to point out that the concept of policy is more fundamental by far than that of minimizing function and thus that dynamic programming can profitably be applied in many situations where there is either no criterion function or a vector-valued criterion function. This is particularly the case in simulation.

4.4 Formalism

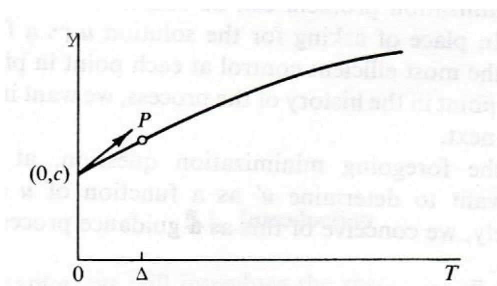
We begin with the observation that the minimum value of J(y) depends on the two quantities singled out above: the initial state c and the duration of the process T. Hence, following the usual mathematical translation of “depends on," we may write

$$\mathrm{V}(c,T) = \min_{y} J(y). \qquad (4.4.1)$$

This value function (minimal cost) is defined for $T \geq 0$ and $-\infty < c < \infty$. The original minimization problem is thus imbedded within a family of variational problems of the same type. It remains to show that this is a useful step to take in our pursuit of the solution of the original problem.

This is a critical step in the approach of dynamic programming. Instead of finding the control policy u(c, T) directly, we try to find out the value function V(c,T), which can in turn leads to the control policy.

This dependence is intuitively clear, and, of course, analytically clear by virtue of the complete solution of the minimization problem obtained in the previous section. To obtain an equation for V(c, T), which, as it turns out, will provide the optimal policy for us, we argue as follows.



Let u = u(c, T) denote the initial slope at t= 0, the control action at time t=0. We need to evaluate the consequence of this control action upon the total cost. For that purpose, we need to consider the immediate effect of this action and the long-term effect. This is also the way we make decisions in our daily life. We need to consider both the long term and short term benefits to figure out the best action to take.

Let Δ be an infinitesimal, then at time Δ, the new state will be $c + \Delta u$.
We have the total cost expressed as the summation of immediate cost and the long-term cost

$$J(y) = \int_0^\Delta + \int_\Delta^T . \qquad (4.4.2)$$

It is this additivity of the functional representing the total cost, together with our assumptions of instantaneous observation, decision, and action, that enables us to regard the problem of determining the form of the minimizing function over (Δ, T) as a new control process starting in state c + uΔ with duration T-Δ.

But how to evaluate the long-term cost without any knowledge on the rest of the control actions except the first one?

The next step is the critical one, or "signature" , in the approach of dynamic programming. Regardless of how u has been chosen at t=0, we will continue from P at time Δ in such a way as to minimize the cost incurred,

$$\int_\Delta^T [y^2 + \dot{y}^2]dt, \qquad (4.4.3)$$

over the remaining time interval (Δ, T). In other words, the rest of the actions follow the optimal policy which is the best we can do. The key feature of dynamic programming is that it does not need to specify the complete control actions at every step by assuming taking optimal actions for the rest of the process. In doing so, the long-term cost can be easily obtained, and the attention is paid only to the current control action to be taken! As will be shown later, this is also the main advantage of dynamic programming over brutal force approach (like exhaustive search).

Since the optimal control actions will be followed after Δ, by definition of the value function V(c, T), this minimum value is V(c + uΔ, T-Δ).

We are considering Δ to be an infinitesimal and systematically neglecting terms of order Δ^2. Hence, we can write

$$\int_0^\Delta [y^2 + \dot{y}^2]dt = (c^2 + u^2)\Delta + O(\Delta^2), \qquad (4.4.4)$$

as an estimate of the cost of control over the initial interval [0, Δ]. Thus, (4.4.2) takes the form

$$J(y) = (c^2 + u^2)\Delta + V(c + u\Delta, T - \Delta) + O(\Delta^2), \qquad (4.4.5)$$

How should u be chosen? It is reasonable to suppose that we choose u to minimize the right-hand side. We must balance the cost of control over [0,Δ], the short term cost, against the cost of optimal control over [Δ, T], the long term cost. Hence, we have

$$\mathrm{V}(c,T) = \min_u [(c^2 + u^2)\Delta + V(c + u\Delta, T - \Delta) + O(\Delta^2)]. \qquad (4.4.6)$$

Expanding in a Taylor series, we have

$$V(c+u\Delta, T-\Delta) = V(c,T) + \frac{\partial V}{\partial c}u\Delta - \frac{\partial V}{\partial T}\Delta + O(\Delta^2). \qquad (4.4.7)$$

Hence, (4.4.6) becomes

$$V(c,T) = \min_u [(c^2+u^2)\Delta + V(c,T) + \frac{\partial V}{\partial c}u\Delta - \frac{\partial V}{\partial T}\Delta + O(\Delta^2)]. \qquad (4.4.8)$$

Cancelling the common terms f(c, T) on both sides, dividing through by Δ, and letting $\Delta \to 0,$ we obtain the nonlinear partial differential equation

$$\frac{\partial V}{\partial T} = \min_u [(c^2+u^2) + \frac{\partial V}{\partial c}u]. \qquad (4.4.9)$$

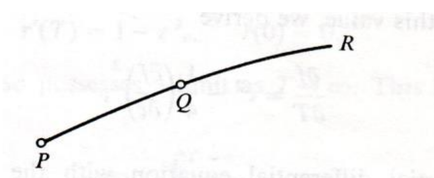
The initial condition is clearly V(c, 0) = 0. As we shall see, this equation will provide us with both the minimum value V(c, T) and the optimal policy u(c, T).

4.5. Principle of Optimality

The argument we used to obtain (4.4.6) is a particular case of the following principle.

Principle of Optimality. An optimal policy has the property that what ever the initial state and initial decision are the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

If an optimal policy is conceived of as a means of tracing out an optimal trajectory in state space, say a geodesic as far as time or distance are concerned, then the principle becomes quite clear:



If PR is to be a path of minimum time, then clearly the last part, QR must also be a path of minimum time.

Once Q has been determined, PQ must also be a geodesic. The point is, of course, that starting from P we don't always know the point Q in advance. This means that we can only assert that PQ + QR must constitute an optimal trajectory.

Another way of putting this is to state that in a multistage decision process we must balance the gain or loss resulting from the first decision against the return from the optimal continuation from the remaining decisions.

In the remaining part of the lecture we will show how easy it is to apply the foregoing principle to treat a number of optimal control problems.

4.6. Discussion

So far we do not seem to have gained much. Whereas the formalism of the calculus of variations yielded an ordinary differential equation, the new formalism confronts us with a partial differential equation of unfamiliar type. There is, however, one point to emphasize. The differential equation was subject to a two-point boundary condition, whereas the partial differential equation has its solution determined by an initial condition. These are very important points as far as analytic and computational solutions are concerned.

We can remove some of the strangeness by carrying out the minimization with respect to u. We have, differentiating in (4.4.9),

$$u = -\frac{1}{2}\frac{\partial V}{\partial c} \tag{4.6.1}$$

and thus, using this value, we derive

$$\frac{\partial V}{\partial T} = c^2 - \frac{1}{4}(\frac{\partial V}{\partial c})^2, \tag{4.6.2}$$

a nonlinear partial differential equation with the initial condition V(c, 0) = 0.

4.7 Simplification

We start with the observation that

$$\mathrm{V}(c,T) = c^2 r(T) \tag{4.7.1}$$

We know this using the explicit form provided by the results of section (3.8). We can also derive this by using the common technique of separation of variables for solving PDE.

Assume that the solution of to the PDE takes the form

$$\mathrm{V}(c,T) = f(c)r(T) \tag{4.7.2}$$

Plug it into (4.6.2), we have

$$\mathrm{f(c)}r' = c^2 - \frac{1}{4}(f'(c))^2\,\mathrm{r}^2 \tag{4.7.3}$$

For this equation to hold, we have the following condition on f(c),

$$\mathrm{f(c)} = c^2 = \frac{1}{4}(f'(c))^2 \tag{4.7.4}$$

which confirms the validity of (4.7.1).

Using (4.7.1) in (4.6.2), we see that r(T) satisfies the ordinary differential equation

$$\dot{r}(T) = 1 - r^2(T),\ \mathrm{r(0)=0}, \tag{4.7.5}$$

a <u>Riccati equation</u>.

4.8. Validation

We have already established in section (3.8) that

$$\mathrm{V}(c,T) = c^2 \tanh(T) \tag{4.8.1}$$

thus

$$r(T) = \tanh(T) \tag{4.8.2}$$

does indeed satisfy (4.7.5).

The point to emphasize is that we have once again invoked a familiar technique of analysis. A formal procedure based upon intuitive concepts is first used to obtain a relation. Once the form of the relation is observed, we can employ either an entirely differential method to establish it rigorously, or, guided by our knowledge of the desired result, return to the original method and smooth the edges.

After all, we have a choice of our techniques. No holy vow has been sworn to adhere rigidly to one or the other general theories. Successful mathematical research requires eclecticism.

4.9. Infinite Process

The formalism is particularly simple in the case where $T = \infty$. Write

$$\mathrm{V}(c) = \min_y \int_0^\infty (\dot{y}^2 + y^2)dt. \tag{4.9.1}$$

Then the same formalism as above yields the relation

$$\mathrm{V}(c) = \min_u [(u^2 + c^2)\Delta + V(c + u\Delta)] + O(\Delta^2), \tag{4.9.2}$$

hence

$$0 = \min_u [(u^2 + c^2) + u\dot{V}(c)] \tag{4.9.3}$$

and

$$u = -\frac{\dot{\mathrm{V}}(c)}{2}, \qquad \dot{V}^2 = 4c^2 \tag{4.9.4}$$

Since $\mathrm{V}(c) \geq 0$ for $c \geq 0$, and is increasing as c increases, we see that

$$\dot{V} = 2c$$

hence

$$\mathrm{V}(c) = c^2, u = -c. \tag{4.9.5}$$

This is the same result we obtained previously as the limit of V(c, T) as $T \to \infty$

4.10. Discrete Control Processes

In many situations, we cannot control a system in a continuous fashion. It may not be possible to observe it continuously, nor to make an unbroken succession of decisions. Furthermore, it is of considerable interest to determine whether such close scrutiny of a system is necessary. It may be the case that intermittent control will be just as effective, and much cheaper and simpler to implement. Esthetically, there is a great advantage to formulating a mathematical model directly in discrete terms if we expect to use a digital computer to obtain a numerical solution of the resulting equations. From the mathematical point of view, there is the advantage that a direct rigorous approach can be made without any worry about the subtleties of the calculus of variations.

Let us then consider a system whose state at time n, n=0,1,2,..., is described by the scalar variable y(n) . Let the equation determining the history of the system be

$$y(n+1) = ay(n) + u(n), y(0) = c \qquad (4.10.1)$$

and suppose that the u(n), the control variables are chosen so that the quadratic form

$$J_N(y,u) = \sum_{n=0}^{N} (y^2(n) + u^2(n)) \qquad (4.10.2)$$

is minimized.

Since the minimum value of J(y, u) depends on c, the initial value, and N, the number of stages, we write

$$V_N(c) = \min_{\{u(n)\}} J_N(y,u) \qquad (4.10.3)$$

Does the minimum actually exist? Suppose that we take u(n) = 0, n= 0, 1, 2,..., N. The corresponding value of $y(n) = a^n c$, and the cost function $J_N(y,0)$ is

$$J_N(y,0) = \sum_{n=0}^{N} (a^n c)^2 = c^2 \sum_{n=0}^{N} a^{2n}$$

Hence, we know that

$$\min_{\{u(n)\}} J_N(y,u) \le c^2 \sum_{n=0}^{N} a^{2n} \qquad (4.10.4)$$

Thus, since

$$\sum_{n=0}^{N} u^2(n) \le \sum_{n=0}^{N} (y^2(n) + u^2(n)),$$

it is sufficient to ask for the minimum of $J_N(y,u)$ over the finite region

$$\sum_{n=0}^{N} u^2(n) \le c^2 \sum_{n=0}^{N} a^{2n}. \qquad (4.10.5)$$

The minimum of $J_N(y,u)$ is assumed over this closed region and thus $V_N(c)$ exists as a function of c for all c and N = 0, 1, 2,… As before it can be shown directly that

$$V_N(c) = c^2 r_N \qquad (4.10.6)$$

where r_N is independent of c.

4.12. Recurrence Relation

Let us now apply the principle of optimality to obtain a recurrence relation, a nonlinear difference equation, satisfied by $V_N(c)$.

After u(0) is chosen, the new state of the system is y(1)=ac+u(0) . The cost function can be broken down into the summation of short term cost and long term cost as

$$c^2 + u^2(0) + \sum_{n=1}^{N} (y^2(n) + u^2(n)). \qquad (4.12.1)$$

The additivity of the integrand shows that regardless of the choice of u(0), we proceed from the new state to minimize the remaining sum. Hence, we know that for any choice of u(0), the cost function takes the form

$$c^2 + u^2(0) + V_{N-1}(ac + u(0)). \quad (4.12.2)$$

The quantity u(0) is now to be chosen to minimize the expression. Thus, we obtain the nonlinear recurrence relation

$$\begin{aligned} &V_N(c) = \min_{u(0)} [c^2 + u^2(0) + V_{N-1}(ac + u(0))], \\ &V_0(c) = c^2, N \geq 1. \end{aligned} \qquad (4.12.3)$$

Using the fact that $V_N(c) = c^2 r_N$, (4.12.3) yields the relation

$$c^2 r_N = \min_{u(0)} [c^2 + u^2(0) + r_{N-1}(ac + u(0))^2], \qquad (4.12.4)$$

The value of u(0) that minimizes is readily obtained by differentiation,

$$\begin{aligned} &2u(0) + 2r_{N-1}(ac + u(0)) = 0, \\ &u(0) = \frac{-r_{N-1}ac}{1 + r_{N-1}} \end{aligned} \qquad (4.12.5)$$

Using this value and a small amount of algebra in (4.12.4), we obtain the relation

$$\begin{aligned} &r_N = (1 + a^2) - \frac{a^2}{(1 + r_{N-1})} \\ &r_0 = 1, N \geq 1 \end{aligned} \qquad (4.12.6)$$

This is a discrete analogue of the Riccati equation obtained in the continuous case.

Note that the controller is in a feedback form since c=y(0). Overall, we have obtained the following optimal controller in a feedback fashion.

At each time t=k, measure the state variable y(k), and calculate

$$u(k) = \frac{-r_{N-k-1}ay(k)}{1 + r_{N-k-1}} \qquad (4.12.7)$$

Comparing to the pole-placement controller, the optimal controller is simply a state-feedback controller, but with a time-varying feedback gain, which depends upon the time duration left for the control process. If the time duration is infinite, then we can expect that the feedback gain becomes a constant.

4.13. Infinite Process

It is clear from (4.12.6) that r_N is uniformly bounded by the quantity $(1 + a^2)$, and that it is monotone increasing. We have

$$r_1 = (1+a^2) - \frac{a^2}{(1+1)} = 1 + \frac{a^2}{2} > 1 = r_0,$$

$$r_2 = (1+a^2) - \frac{a^2}{(1+r_1)} > (1+a^2) - \frac{a^2}{(1+r_0)} = r_1$$

and thus, inductively, $r_N > r_{N-1}$. This monotone behavior in N follows, of course, directly from the definition of $V_N(c)$.

Let $r = \lim_{N\to\infty} r_N$. Then r is the positive root of the quadratic equation

$$r = (1+a^2) - \frac{a^2}{(1+r)} \tag{4.13.1}$$

Observe that u(0), the initial decision, also converges as $N \to \infty$, and

$$\lim_{N\to\infty} u(0) = -\frac{rac}{1+r} \tag{4.13.2}$$

If we formally consider the infinite process,

$$\mathrm{V}(c) = \min_{\{u(n)\}} \sum_{n=0}^{\infty} (y^2(n) + u^2(n)), \tag{4.13.3}$$

we see that

$$\mathrm{V}(c) = \min_{u(0)} [c^2 + u^2(0) + V(ac + u(0))] \tag{4.13.4}$$

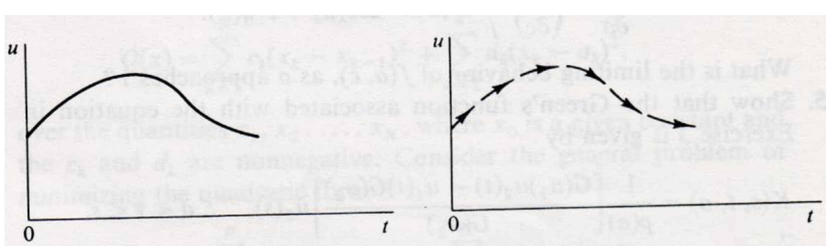
One solution of this equation is certainly $\mathrm{V}(c) = rc^2$.

Therefore, for the infinite time process, the optimal feedback controller is simply,

$$u(k) = -\frac{ray(k)}{1+r} \tag{4.13.5}$$

4.14 Comparison between calculus of Variations and Dynamic Programming

In the calculus of variations, we wish to minimize a functional J(u). The minimizing function is regarded as a point in a function space. In the dynamic programming approach, the tangent at a point is determined by the policy.



Thus, in the calculus of variations we regard a curve as a locus of points, and the resulting controller is a function of time; while in dynamic programming we regard a curve as an envelope of tangents, and the resulting controller depends upon the state variables rather on time. The basic duality theorem of Euclidean space is that these are equivalent formulations.

4.15. Other Applications using Dynamic Programming

The basic idea of dynamic programming is to use *Principle of Optimality* to figure out the optimal policy. Given the state, and at time k, you wish to find out the optimal action. At the beginning, you have no idea which action to take. Then you just pick any one, and find out the total cost of this action, which can be broken down into summation of short term and long term costs. In order to evaluate the long-term cost properly, we simply assume the optimal policy will be implemented for the rest of the process, which is the key step in dynamic programming. Then you try to find out the best action to minimize the instant cost and long term cost. In this way, you can get an equation for the optimal cost, from which you can then find the optimal policy.

In summary, there are three steps.

1. Define the optimal value function, V(x), where x represent all the information affecting the outcome. The value could be cost or utility, and the problem could be minimization (for cost) or maximization (for utility)
2. At time t, take any action u(t), and find out the immediate consequence associated with this action. The greedy policy would be just to choose the action minimize the immediate cost. The optimal policy is to balance the immediate cost and the long term cost. Then comes the critical step for dynamic programming—determine the long term cost by assuming the optimal actions for the rest!
3. Assume that the rest of the actions will follow the optimal policy, which of course will depend upon the state driven by the current action. Then try to minimize (or maximize) the total value by choosing the best action, which will lead to the optimal cost. In this way, a recurrence relation will be derived to relate the current optimal value and the future value (immediately after the current action).

Corresponding to above three steps, the basic functional equation of dynamic programming has the form

$$\mathrm{V(x)} = \max_u [H(x,u,\mathrm{V(T(x,u))})] \quad (4.15.1)$$

Where the function T(x,u) describes the consequence resulting from a particular decision u, and H is the total value associated with this decision combined with the rest. In many case the total cost can be expressed as summation of short term cost and long term cost.

$$H(x,u,\mathrm{V(T(x,u))}) = \mathrm{C(x,u)} + \mathrm{V(T(x,u))}$$

There are many applications which can be solved by dynamic programming. In the following, we will show one example.

The **knapsack problem** is a problem in combinatorial optimization: Given a set of items, each with a weight and a value, determine the number of each item to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible. It derives its name from the problem faced by someone who is constrained by a fixed-size knapsack and must fill it with the most valuable items.

The problem often arises in resource allocation where there are financial constraints and is studied in fields such as combinatorics, computer science, complexity theory, cryptography, applied mathematics, and daily fantasy sports.

The knapsack problem has been studied for more than a century, with early works dating as far back as 1897. The name "knapsack problem" dates back to the early works of mathematician Tobias Dantzig (1884–1956), and refers to the commonplace problem of packing your most valuable or useful items without overloading your luggage.

Knapsack problems appear in real-world decision-making processes in a wide variety of fields, such as finding the least wasteful way to cut raw materials, selection of investments and portfolios, selection of assets for asset-backed securitization, and generating keys for the Merkle–Hellman and other knapsack cryptosystems.

One early application of knapsack algorithms was in the construction and scoring of tests in which the test-takers have a choice as to which questions they answer. For small examples it is a fairly simple process to provide the test-takers with such a choice. For example, if an exam contains 12 questions each worth 10 points, the test-taker need only answer 10 questions to achieve a maximum possible score of 100 points. However, on tests with a heterogeneous distribution of point values—i.e. different questions are worth different point values— it is more difficult to provide choices.

George Dantzig proposed a greedy approximation algorithm to solve the unbounded knapsack problem. His version sorts the items in decreasing order of value per unit of weight. It then proceeds to insert them into the sack, starting with as many copies as possible of the first kind of item until there is no longer space in the sack for more. Provided that there is an unlimited supply of each kind of item, then the greedy algorithm is guaranteed to achieve the optimal solution. However, for the bounded problem, where the supply of each kind of item is limited, the algorithm may be far from optimal.

The most common problem being solved is the **0-1 knapsack problem**, which restricts the number x_i of copies of each kind of item to zero or one (either selected or not selected). Given a set of n items numbered from 1 up to n, each with a weight w_i and a value v_i, along with a maximum weight capacity W, we wish to determine a subset

$$S \subseteq \{1, 2, \cdots n\} \qquad (4.15.2)$$

$$\text{that maximizes} \quad \sum_{i \in S} v_i \qquad (4.15.3)$$

$$\text{subject to} \sum_{i \in S} w_i \le W \qquad (4.15.4)$$

Of course, the exhaustive search (brutal force) is to try all 2^n possible subsets. The greedy algorithm is to rank the items by the values, and place the items one by one from the top value to the lower ones until the bag is full. It is fast and very often gives a good solution, but there is no guarantee to get the optimal solution.

Question: Any better solution?
Yes–Dynamic programming (DP)!

First, define the optimal value function V. What are the factors affecting the optimal value? The set of items and the total capacity.

Therefore, let's define

$$V(i,w), \quad 1 \le i \le n \quad 0 \le w \le W \qquad (4.15.5)$$

as the maximal value of the bag with the available set of {1,2,3,…,i}, and the capacity of w. We need to find out this function for all possible i and w. For many applications, the arguments are discrete type, like the **knapsack problem.** Then we need to use a table for storing all the outcomes, rather than a function with continuous variable.

If we can compute all the entries of this array, then the array entry V(n,W) will contain the maximum value of the items that can fit into the knapsack, that is, the solution to our problem. Of course, for each optimal value, we can also find out the choices of the items to be placed, and store them. For the time being, let's just focus upon how to find out V(n,W) first.

The key is to use dynamic programming to form the recurrence relation such that all the entries can be computed.

When i=1, we can easily find out V(1,w) for $0 \le w \le W$ as follows

$$V(1,w) = \begin{cases} 0, & w_1 > w \\ v_1, & w_1 \le w \end{cases} \qquad (4.15.6)$$

How do we find out V(2,w) for all $0 \le w \le W$?

Now there are two items available, it appears that we need to make a decision on each of them. However, since we already know V(1,w), it turns out that we only need to make decision on item 2, which significantly reduces the computing time.

Since it is just 1 or 0, there are only two choices for item 2, select it or not. So just need to compute the value for each of them and pick the bigger one. Compared to the control problem, this step is much easier as only two actions to be chosen rather than a continuous variable.

If item 2 is confirmed to be selected, then the totally capacity of the bag reduces from w to $w - w_2$, and the available set for selection is the set of items {1}. The short term value for this action is the value associated with the item 2, v_2 . But how about the total value? Following the earlier idea illustrated for optimal control, the rest of the items should be chosen in an optimal way, for which we already have the optimal value of $V(1, w - w_2)$, therefore the total value is $V(1, w - w_2) + \mathrm{v}_2$.

If item 2 is not selected, then the optimal value is simply V(1,w). So we just select the bigger one as

$$\max(V(1,w), V(1, w - w_2) + \mathrm{v}_2)$$

Overall, we have

$$V(2,w) = \begin{cases} V(1,w) & if \quad 0 \le w < w_2 \\ \max(V(1,w), V(1, w - w_2) + \mathrm{v}_2) & if \quad w_2 \le w < W \end{cases} \qquad (4.15.7)$$

Following the same argument, we can easily get the recurrence relation between V(i-1,w) and V(i,w).

Assuming we have the complete information about the subset {1,2,…, i-1} with the optimal value V(i-1,w), how to compute V(i,w)?

Again, there are only two choices for item i, select it or not. If not, then the value is still V(i-1,w). If selected, then the totally capacity of the bag reduces from w to $w - w_i$, and the available set for selection is the subset of items {1,2,..,i-1}, for which we already have the optimal value of $V(\mathrm{i}-1, w - w_i)$, therefore the total value is $V(\mathrm{i}-1, w - w_i) + \mathrm{v}_i$.

Overall, we have the recurrence equation to compute $V(\mathrm{i}, w)$ from $V(\mathrm{i}-1, w)$,

$$V(\mathrm{i},w) = \begin{cases} V(\mathrm{i}-1,w) & if \quad 0 \le w < w_i \\ \max(V(\mathrm{i}-1,w), V(\mathrm{i}-1, w - w_i) + \mathrm{v}_i) & if \quad w_i \le w \le W \end{cases} \qquad (4.15.8)$$

Tabulating the results from $V(\mathrm{i},1)$ up through $V(\mathrm{i},\mathrm{W})$ gives the solution. This solution will therefore run in $O(n\mathrm{W})$ time and $O(n\mathrm{W})$ space. Compare to the brutal force in the order of 2^n in time, the computational load is significantly reduced for large n.

Once the optimal value function V(i,w) is found, it is very easy to find out whether item i should be chosen or not by working backwards from item n to item 1. For instance, if you want to decide whether item n should be included or not, you just compare V(n,W) and V(n-1,W) and use the following decision rule:

Item n should be selected if $V(\mathrm{n}, W) > V(\mathrm{n}-1, W)$ (4.15.9)

For the next item n-1, it will depend upon the choice on item n. If item n is confirmed to be selected, then the total packing space reduces from W to $W - w_n$. Then using similar argument, you just compare $V(n-1, W - w_n)$ and $V(n-2, W - w_n)$, and use the decision rule that

$$\text{Item n-1 should be selected if } V(n-1, W - w_n) > V(n-2, W - w_n) \quad (4.15.10)$$

If item n is confirmed not to be selected, the total packing space remains as W. Then use the decision rule that

$$\text{Item n-1 should be selected if } V(n-1, W) > V(n-2, W) \quad (4.15.11)$$

Overall, to check whether item i should be selected, you need to know the choices of all the items from i+1 to n, and obtain the total packing space after packing the items i+1 to n as W_i, and then decide

$$\text{Item i should be selected if } V(\mathrm{i}, W_i) > V(i-1, W_i) \quad (4.15.12)$$

In summary, both calculus of variations and dynamic programming can be applied for solving optimization problems in continuous-time space, while dynamic programming is more suitable for discrete-time systems.

It often happens that one type of mathematical model is well suited to one type of analytic approach and not to another. Generally, the three principal parts of a mathematical model, the conceptual, analytic, and computational aspects, must be considered simultaneously and not separately. Consequently, a certain amount of effort is involved in translating verbal problems posed in such vague terms of efficiency, feasibility, cost and so on, into precise analytic problems requiring the solution of equations and determination of extrema.

Remember: if the optimization problem can be cast as a multistage decision process, dynamic programming should be the first choice of your weapon to attack the problem. The key idea of dynamic programming is to focus upon the decision at present while assuming the rest will be optimal, which simplifies the calculation of long term and short term cost for each action.

Reference:

The material presented in the lecture notes is mostly exacted from the following classical textbook on control theory:

R. Bellman, Introduction to the Mathematical Theory of Control Processes, Vol. 1, Academic Press, 1967.