Example Find the optimal trajectory for the problem of

$$\text{minimizing} \quad \int_0^3 (x^2 + \dot{x}^2)dt$$

$$\text{subject to} \quad x(0) = 0, x(3) = 3.$$

Solution Our example is a classic, see Bellman's text on optimisation. What we compute to begin with is actually only an approximation to the solution, the accuracy of which can be refined at will. We can label our stages using time t and label the states using x. The network is discretized using an appropriately small time step and x step. A simplified network is illustrated.

Figure 5. *Discretisation*

Note that the optimal path must be non-decreasing. The Bellman equation here reads

$$S(t,x) = \min_y \left\{ \int_t^{t+\Delta t} (y^2 + \dot{y}^2)d\tau \ + \ S(t + \delta t, x + \delta x) \right\}.$$

Using the approximations dictated by averaging we have

$$\delta x = \dot{y}\delta t, \qquad y = x + \frac{1}{2}\delta x$$

the transition cost is

$$C(t,x,y) = \delta t. \left\{ (x + \frac{1}{2}\delta x)^2 \ + \ (\delta x/\delta t)^2 \right\}$$

or if we take $\delta t = 1$

$$C(t,x,\delta x) = \left(x + \frac{1}{2}\delta x \right)^2 \ + \ (\delta x)^2.$$

As this is independent of time we denote the expression by $C(x, \delta x)$.

Our substitute problem is now to solve

$$S(t,x) \ = \min_{\delta x \in \{0,1,2,3\}} \{C(x,\delta x) + S(t+1, x+\delta x)\}$$

by the use of dynamic programming. We leave the actual calculation via, for instance Mathematica, to the reader (see section 7.6). The solution is sketched below.

Figure 6. *Optimal solution by discretisation.* Evidently a smaller step size on the t-axis will give a closer approximation.

7.3 Continuous-time dynamical programming In the next examples we study the effect on the Bellman Equation of passing to the limit as $\Delta t \to 0$. **Example 1** Minimize the cost

$$F(x) \equiv \int_0^T (x^2 \; + \; \dot{x}^2) dt$$

subject to $\quad x(0) = c.$

We regard the minimum cost as being a function of c and of T, thus

$$S(c, T) = \min_{x(t)} \left\{ \int_0^T (x^2 \; + \; \dot{x}^2) dt \right\}.$$

We think of the trajectory as consisting of a small first segment from $t = 0$ to $t = \triangle t$ and then the rest. Thus

$$S(c, T) = \int_0^{\triangle t} \; + \; \int_{\triangle t}^T = \int_o^{\triangle t} \; + \; S(c \; + \; \triangle x, T - \triangle t),$$

where the value of x increases from c to $c \; + \; \triangle x$.

We evaluate all these expressions in terms of Δt. Now if $v = \dot{x}(0)$ then for small enough $\triangle t$

$$\left| \frac{\triangle x}{\triangle t} - v \right| \; < \; \epsilon$$

i.e. $\qquad \triangle x \; = \; v \triangle t \; + \; o(|\triangle t|)$

Similarly

$$\int_0^{\triangle t} \{x^2 \; + \; \dot{x}^2\} dt = (c^2 \; + \; v^2) \triangle t \; + \; o(|\triangle t|)$$

$$S(c,T) \;=\; (v^2 \;+\; c^2)\triangle t \;+\; S(c \;+\; v\triangle t,\; T-\triangle t) \;+\; o(|\triangle t|)$$

Figure 7. *Breaking the curve into two.* We may regard v as the object we wish to compute. The traditional first-order argument hides behind some hand-waving and gives us a Bellman Equation where the error term is neglected:

$$S(c,T) \;=\; \min_v \left\{ (v^2+c^2)\triangle t \;+\; S(c+v\triangle t, T-\triangle t) \right\}.$$

One might justifiably regard this as a “substitute problem” for that posed originally. Note that the substitute problem mysteriously replaces minimisation over all curves $x(t)$ by minimisation over all velocities v. We shall return to a discussion of the neglected term in Section 7.5. Suffice it to say that a better, but somewhat longer argument can be presented through the use of the Mean Value Theorem. We may take the calculation yet one step further by writing

$$S(c+v\triangle t, T-\triangle t) \;=\; S(c,T) \;+\; v\frac{\partial S}{\partial c}\triangle t \;-\; \frac{\partial S}{\partial T}\triangle t \;+\; o(|\triangle t|).$$

Our substitute problem becomes

$$S(c,T) \;=\; \min_v \left\{ (v^2+c^2)\triangle t \;+\; S(c,T) \;+\; v\frac{\partial S}{\partial c}\triangle t \;-\; \frac{\partial S}{\partial T}\triangle t \right\}.$$

This reduces, after suppression of the common term on both sides and after cancellation by the *positive* quantity Δt (which preserves the minimisation), to

$$0 = \min_v \left\{ (v^2+c^2) \;+\; v\frac{\partial S}{\partial c} \;-\; \frac{\partial S}{\partial T} \right\}$$

or equivalently

$$\frac{\partial S}{\partial T} \;=\; \min_v \left\{ (c^2+v^2) \;+\; v\frac{\partial S}{\partial c} \right\}.$$

with $S(c, 0) = 0$. If we can solve for $v(= \dot{x})$ we shall in effect have the optimal trajectory.

Though Bellman's Equation looks forbidding, luckily, under certain circumstances it can be solved explicitly. One such circumstance is when its variables can be **separated**, and our example is a case in point. We can easily show that

$$S(c, T) = c^2 \cdot S(1, T).$$

Indeed if $z(t) = x(t)/c$ then $z(0) = 1$. Moreover

$$F(z) = c^{-2} \int_0^T (x^2 + \dot{x}^2) dt = c^{-2} \cdot F(x),$$

so that if $x(t)$ is the optimising trajectory for the value c, then obviously (why?), $z(t)$ is the optimising trajectory, for the special case when $c = 1$. Thus $S(c, T) = F(x) = c^2 F(z) = c^2 S(1, T)$. The above argument is, evidently, based on the fact that the objective function is homogeneous (of degree 2), whereas the initial constraint is linear in the parameter c. We stop to note another example of the same sort. **Example 2** Maximize for any t with $0 < t < 1$ the þrevenueþ function:

$$F(x) \equiv \int_t^1 f(\tau) \sqrt{x(\tau)}\, d\tau$$

subject to $\int_t^1 x(\tau)\ d\tau = c$. Here the maximum revenue $S(c, t) = \sqrt{c} \cdot S(1, t)$, since the objective function is trivially homogeneous of degree one-half. We now solve the Bellman Equation of Example 1. Writing $S = c^2 \cdot G(T)$, we obtain the equation:

$$c^2 G'(T) = \min_v \left\{ (c^2 + v^2) + v \cdot 2c \cdot G(T) \right\}.$$

with $G(0) = 0$. For this the first order condition reads:

$$2v + 2cG(T) = 0,$$

i.e. $v = -c \cdot G(T)$. On substitution we obtain:

$$c^2 G'^2 + c^2 G(T)^2 - 2c^2 G(T)^2,$$

that is

$$\frac{G'}{(1 - G^2)} = 1,$$

whence, treating G as a function of a parameter τ, and integrating from 0 to T, we have:

$$\int_0^T \left\{ \frac{G'}{2(1 - G)} + \frac{G'}{2(1 + G)} \right\} d\tau = \int_0^T 1\, d\tau = T.$$

Thus

$$\frac{1+G}{1-G} = e^{2T},$$

or

$$G(T) = \frac{e^{2T}-1}{e^{2T}+1}.$$

This result may, of course, be verified by solving the Euler-Lagrange Equation of the problem and computing the cost along the optimal trajectory. A second case where the Bellman Equation can be solved explicitly is where the cost function depends only one variable. This is the case in Example 1 when we put $T = \infty$, so that the dependence on T disappears. The case is dealt with in Example 3 below. **Remark** It is not very difficult to prove that the Bellman Equation may always be solved explicitly in the above manner (i.e. analytically), if the constraint is linear in a parameter and the objective function $F(x) = \int_0^T f_0(x, \dot{x})\, dt$ is homogeneous, while the derivative, which is then necessarily also homogeneous, $\frac{\partial f}{\partial \dot{x}}(1, v)$ is an increasing function of v.

Example 3 Solve Example 1 with $T = \infty$.

Let $\quad S(c) = \lim_{T\to\infty} S(c, T).$

Then we have from the earlier calculation

$$0 = \min_v \left\{ (c^2 + v^2) + v\ S'(c) \right\},$$

where all reference to T has been obliterated. Let us see why this is. We return to the standard invocation of the Optimality Principle.

$$S(c) = \min_{\substack{x(\tau)\\ x(0)=c}} \left(\int_0^{\Delta t} (x^2 + \dot{x}^2)\, d\tau + \int_{\Delta t}^{\infty} (x^2 + \dot{x}^2)\, d\tau \right).$$

Now the second term may be rewritten by using the substitution:

$$\theta = \tau - \Delta t$$

to yield:

$$\int_0^{\infty} \left(x(\theta + \Delta t)^2 + \dot{x}(\theta + \Delta t)^2 \right)\ d\theta.$$

Setting $z(\theta) = x(\theta + \Delta t)$ we obtain a function z such that $z(0) = x(\Delta t)$ with associated cost $F(z)$ where

$$F(z) = \int_0^{\infty} (z^2 + \dot{z}^2)\, dt.$$

Thus, if x is optimal for the given problem and z is optimal for the same problem but with c replaced by $c + \Delta c = x(\Delta t)$, then , since $F(z) = S(c + \Delta c)$, we have

$$S(c) = \min_{\substack{x,z\\ x(0)=c,\quad z(0)=x(\Delta t)}} \left(\int_0^{\Delta t} (x^2 + \dot{x}^2)\, d\tau + F(z) \right), = \min_x \left(\int_0^{\Delta t} (x^2 + \dot{x}^2)\, d\tau + S(c + \Delta c) \right).$$

From here, the old argument from Taylor's Theorem leads us to the current version of Bellman's Equation.

Minimizing by the use of calculus gives the first order condition:

$$0 = 2v \; + \; S'(c)$$

i.e. $\frac{1}{2}S'(c) = -v.$

Substituting back gives

$$0 = c^2 \; + \; \frac{1}{4}S'^2 \; - \; \frac{1}{2}S'(c).S'(c).$$

or $S'^2 = 4c^2.$

As in Example 1, notice that, if $c \neq 0$, the substitution $c.y = x$ gives us

$$F(cy) = c^2 F(y) \quad \text{and} \quad y(0) = 1,$$

where $F(x)$ denotes the cost of the trajectory x. Hence $S(c) = c^2 S(1)$, and it is clear that $S(1) > 0$. Thus $S'(c) = 2c$ unless $c = 0$ and consequently $S(1) = 1$. (It is also clear that $S(0) = 0$.) Furthermore, $c = -v = -\dot{x}(0)$. It is interesting to note that the condition $x(t) = -\dot{x}(t)$ may be established not just at $t = 0$, but for any $t > 0$, by the same argument (essentially, since the integrand does not depend on t explicitly). We thus see that the optimal path obeys the differential equation:

$$\frac{dx}{dt} = -x,$$

so that $x(t) = x(0)e^{-t}$. A classic application of the Optimality Principle is concerned with evaluating the maximum height reached by a particle projected vertically with initial velocity v moving against gravity. This is a simple problem in one dimension so that the trajectory is linear.

Example 4: The maximum height problem A mass m is projected upwards with initial velocity v being subjected only to gravity. What is its maximum height H above ground?

The law of motion is $\ddot{x} = -g$ with $\dot{x}(0) = v$. Taking $z = \dot{x}$ as the state variable we have in fact

$$\dot{z} = -g \quad \text{and} \quad z(0) = v.$$

Before offering a solution in the spirit of Example 1, we look at the traditional solution of the problem.

$$z(t) - z(0) = -gt, \quad \text{i.e.} \quad z(t) = v - gt. \tag{1}$$

At the highest point $z = 0$. Of this occurs when $t = \tau$ where we have $0 - v = -gt$ i.e. $\tau = v/g$.
Evidently $x(t) - x(0) = \int_0^t \dot{x} dt = \int_0^\tau z(t) dt$, or since $x(0) = 0$, $x(t) = vt - \frac{1}{2}gt^2$.
Thus $H = v^2/g - \dfrac{1}{2} g \dfrac{v^2}{g^2} = \dfrac{v^2}{2g}$.

We shall solve the problem now by **appeal to the principle of optimality**. Taking vertical distance upwards as $x(t)$ the underlying problem is

$$\text{maximize} \quad x(T) \ = \ \int_0^T \dot{x}(t) dt$$

subject to

$$\left. \ddot{x} = -g \dot{x}(0) = v \right\}$$

Thus the natural state variable is $z(t) = \dot{x}(t)$ leaving us with $\dot{z} = -g, z(0) = v$ as the state equations with maximization of $\displaystyle\int_0^T z(t)dt$, where T is not specified. We begin by assuming that the maximum height reached is a function of only v. This may be justified as follows. There is a theorem in the theory of differential equations which asserts that the solution of the differential equation $\dot{z} = f(z)$ subject to $z(0) = v$ is a continuous function of the initial conditions v (under assumptions on f, such as f obeys the Lipschitz condition). Thus $z(t) = F(t, v)$ just as in (1). Hence the time taken to reach the maximum height, τ, is a function of v given implicitly by $F(\tau, v) = 0$. Finally, as

$$H = \int_0^\tau z(t) dt,$$

we see that H is a function of v, indeed.

Figure 8. *The maximum height left after elapse of time Δt is the original maximum height less the amount the particle has risen* Now we realize that after an elapse of time

$\triangle t$ the particle has moved up a little and lost some of its initial velocity. The remaining trajectory, however, has an associated maximum height and we seek to relate that maximum height to the original maximum height.

We analyze this formally.

The **loss in speed** is $\triangle v$ and we have

$$\triangle v \;=\; g\triangle t \;+\; o(|\triangle t|). \tag{1}$$

This is because for any $\epsilon > 0$ and $\triangle t$ small enough

$$\left| g - \frac{\triangle v}{\triangle t} \right| \;<\; \epsilon \qquad \text{(since } \dot{z} = -g).$$

The **loss in height** $\triangle x$ may similarly be computed, for since $z \;=\; \dot{x}$ and initially $\dot{x} = v$ we have

$$\triangle x \;=\; v\triangle t \;+\; o(|\triangle t|). \tag{2}$$

A **Bellman equation** now **relates** the two trajectory "costs" ("returns", "performance-indicators")

$$H(v) \;=\; H(v - \triangle v) \;+\; \triangle x. \tag{3}$$

Finally, we assume that H is differentiable so that

$$H(v) \;-\; H(v - \triangle v) \;=\; \frac{dH}{dv}.\triangle v \;+\; o(|\triangle v|). \tag{4}$$

Finally, then

$$v\triangle t \;+\; o(|\triangle t|) \;=\; \frac{dH}{dv}.(g\triangle t \;+\; o(|\triangle t|)) \;+\; o(|\triangle v|).$$

But by (1), if $\epsilon < 1$, then (Exercise!),

$$o(|\triangle v|) \;=\; o((g+1)|\triangle t|) \;=\; o(|\triangle t|).$$

Hence,

$$v\triangle t = \frac{dH}{dv}.g\triangle t \;+\; o(|\triangle t|).$$

Consequently, (why?)

$$v = \frac{dH}{dv}.g$$

Integrating up we get, since $v = 0 \Rightarrow H = 0$, that

$$\frac{1}{2}v^2 \;=\; H(v).g, \quad \text{or} \quad H \;=\; \frac{v^2}{2g}.$$

7.4 Repairing the first-order argument By an exercise we obtain that if $H(s) = G(s\Delta t)$, where

$$G(s) = \int_0^s (x^2 + \dot{x}^2)\, dt + S(x(s), T - s),$$

then since

$$G'(s) = (x^2 + \dot{x}^2)|_{t=s} + \frac{\partial S}{\partial x}(x(s), T - s)\dot{x}(s) - \frac{\partial S}{\partial T}(x(s), T - s)$$

and $H'(s) = G'(s\Delta t) \cdot \Delta t$, we have by the Mean Value Theorem that for some $0 < \theta < 1$

$$H(1) - H(0) = H'(\theta) = G'(\theta\Delta t)\Delta t.$$

In the present context, since $H(0) = S(x(0), T) = S(c, T)$ and $H(1) = S(c + \Delta x, T - \Delta t)$, we have

$$H(1) = H(0) + H'(\theta)$$

which yields

$$S(c + \Delta x, T - \Delta t) =$$

$$S(c, T) + \Delta t\{(x^2 + \dot{x}^2)|_{t=\theta\Delta t} + \frac{\partial S}{\partial x}(x(\theta\Delta t), T - \theta\Delta t)\dot{x}(\theta\Delta t) - \frac{\partial S}{\partial T}(x(\theta\Delta t), T - \theta\Delta t)\}.$$

The Bellman equation reads:

$$S(c, T) = \min_{x(t)} \Bigg(S(c, T) +$$

$$\Delta t\{(x^2 + \dot{x}^2)|_{t=\theta\Delta t} + \frac{\partial S}{\partial x}(x(\theta\Delta t), T - \theta\Delta t)\dot{x}(\theta\Delta t) - \frac{\partial S}{\partial T}(x(\theta\Delta t), T - \theta\Delta t)\} \Bigg),$$

where the minimisation is over all $x(t)$ subject to $x(0) = c$. Suppressing the common term and dividing by the positive quantity Δt (which preserves the minimum operation):

$$0 = \min_{x(t)} \left(\{(x^2 + \dot{x}^2)|_{t=\theta\Delta t} + \frac{\partial S}{\partial x}(x(\theta\Delta t), T - \theta\Delta t)\dot{x}(\theta\Delta t) - \frac{\partial S}{\partial T}(x(\theta\Delta t), T - \theta\Delta t)\} \right).$$

This equation asserts two facts. The first is that for the optimal trajectory $x(t)$ we have:

$$0 = \left(\{(x^2 + \dot{x}^2)|_{t=\theta\Delta t} + \frac{\partial S}{\partial x}(x(\theta\Delta t), T - \theta\Delta t)\dot{x}(\theta\Delta t) - \frac{\partial S}{\partial T}(x(\theta\Delta t), T - \theta\Delta t)\} \right).$$

So if we *fix* $x(t)$ and allow $\Delta t \to 0$, we obtain under assumptions of continuity that:

$$0 = \left((c^2 + v^2) + \frac{\partial S}{\partial x}(c, T)v - \frac{\partial S}{\partial T}(c, T) \right), \qquad (*)$$

where v denotes $\dot{x}(0)$ for the optimal trajectory.

The second fact is that if $x(t)$ is *any* trajectory with $x(0) = c$ then

$$0 \leq \left((x^2 + \dot{x}^2)|_{t=\theta\Delta t} + \frac{\partial S}{\partial x}(x(\theta\Delta t), T - \theta\Delta t)\dot{x}(\theta\Delta t) - \frac{\partial S}{\partial T}(x(\theta\Delta t), T - \theta\Delta t) \right).$$

So if we *fix* any such $x(t)$ and allow $\Delta t \to 0$, we obtain under assumptions of continuity that:

$$0 \leq \left((c^2 + v^2) + \frac{\partial S}{\partial x}(c, T)v - \frac{\partial S}{\partial T}(c, T) \right).$$

Here v again denotes $\dot{x}(0)$ but this time v is arbitrary.

This last inequality and equation (*) for the special value of v prove that,

$$0 = \min_v \left((c^2 + v^2) + \frac{\partial S}{\partial x}(c, T)v - \frac{\partial S}{\partial T}(c, T) \right).$$

Note how in the last equation the earlier minimisation over $x(t)$ has been replaced by minimisation over v.