Lecture 6: Optimal Control Theory and Hamilton Jacobi Equations.

Goals:

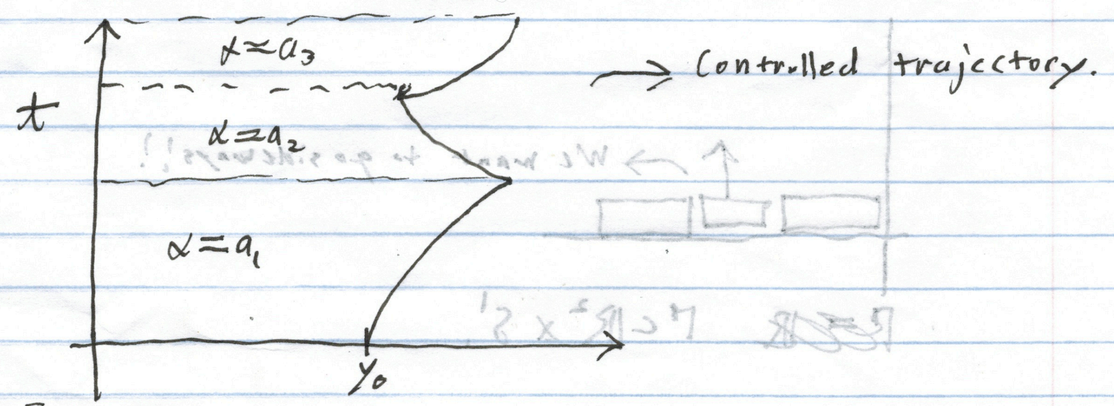
1. A scheme for "guessing" solutions to optimal control problems.
2. Make the connection between optimal control and Hamilton-Jacobi equations.
3. Deep connections between "Value function," H-J equations, and viscosity solutions.

1 Examples

A typical problem is of the form:

$$\boxed{\begin{gathered}\min_{\alpha} \int_0^T h(y(s), \alpha(s))\,ds + g(y(T)). \\ \dot{y}(s) = f(y(s), \alpha(s)),\ y(0) = y_0.\end{gathered}}$$

1. $\alpha : [0, \infty) \to A \subset \mathbb{R}^n$ is the control.
2. $P[\alpha] = \int_0^T h(y(s), \alpha(s))\,ds + g(y(T))$ is the payoff functional.
3. h = running payoff $\to$ fuel, consumption, time.
4. g = terminal payoff $\to$ parking penalty.



5. y = state of the system.

A. Bolza ~ Young

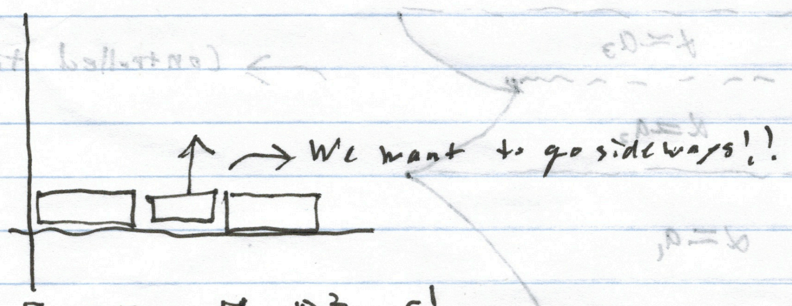
$$\min_{v(0)=0} \int_0^T (v_x^2 - 1)^2 + v^2\,dx$$

Let
$$\frac{dy}{dt} = \alpha, \quad y(0) = 0$$
be the state equation. The running cost is
$$h(y, \alpha) = (\alpha^2 - 1)^2 + y^2$$
$$g = 0.$$

B.) Minimal Arrival Time

$$\min \left\{ \text{time at which } y(s) \text{ reaches target set } \Gamma \right\}.$$
$$\dot{y}(s) = f(y(s), \alpha(s)), \quad y(0) = y_0$$
$$h = \begin{cases} 1 \text{ if } & y(t) \notin \Gamma, \text{ for all } t \leq s \\ 0 \text{ if } & y(t) \in \Gamma, \text{ for some } t \geq s. \end{cases}$$
$$T = \infty:$$

Parallel Parking:
$$\dot{x} = \alpha_1 \cos(\phi) + \alpha_2 \cos(\phi)$$
$$\dot{y} = \alpha_1 \sin(\phi) + \alpha_2 \sin(\phi)$$
$$\dot{\phi} = \alpha_1 b - \alpha_2 b$$
$$\bar{\alpha} = (1, 0) \text{ or } \bar{\alpha} = (0, 1).$$



$$\Gamma \subset \mathbb{R}^2 \times S^1.$$

C. Example Rocket Railroad

rocket engines

$x(t)$ = position

$\dot{x}(t) = v(t)$ = velocity

Goal: Park at the origin in minimal time

$\dot{x}(t) = v(t)$

$\dot{v}(t) = \alpha(t)$

$-1 \leq \alpha \leq 1$

$P[\alpha] = -\tau$, τ = first time when $x(\tau)=0, v(\tau)=0$.

Suppose we only need $\alpha = \pm 1$.

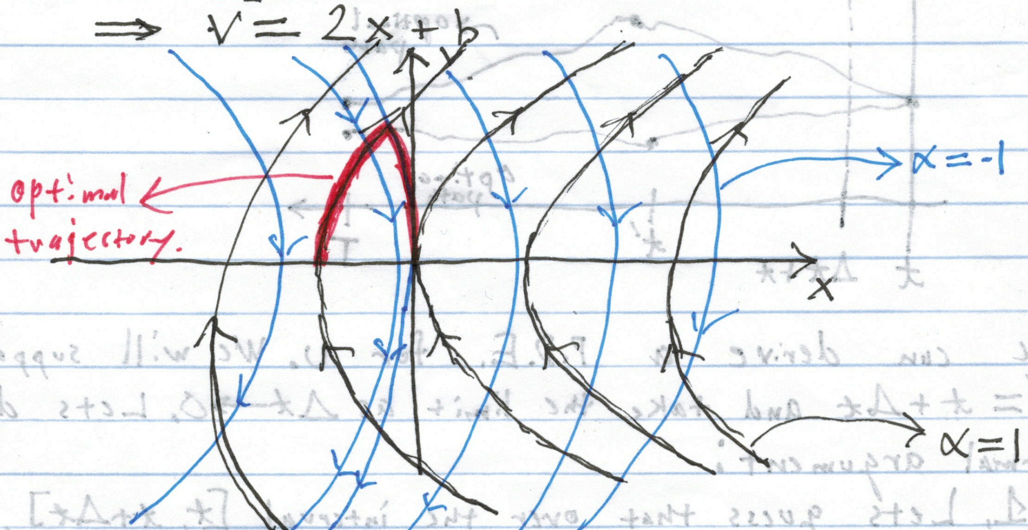
Case 1:

If $\alpha = 1$ then, we also know:

$v(t) = t + b$, $\frac{dx}{dv} = v \Rightarrow \frac{v^2}{2} - \frac{v_0^2}{2} = x - x_0$

$x(t) = \frac{t^2}{2} + bt + c$,

$\Rightarrow v^2 = 2x + b$



Case 2:

If $\alpha = 2$ then

$\frac{dx}{dv} = -v \Rightarrow v^2 = -2x + c$

* Optimal trajectory is a "bang-bang" control. Fully acclerate and then jam on brakes.

2 Dynamic Programming

A typical trick in applied math is to embed the problem in a larger class of problems and solve all of them at once. Lets define the value function:

$$\boxed{u(x,t) = \min_{\alpha(s)\in A} \int_t^T h(y(s),\alpha(s))\,ds + g(y(T)).}$$

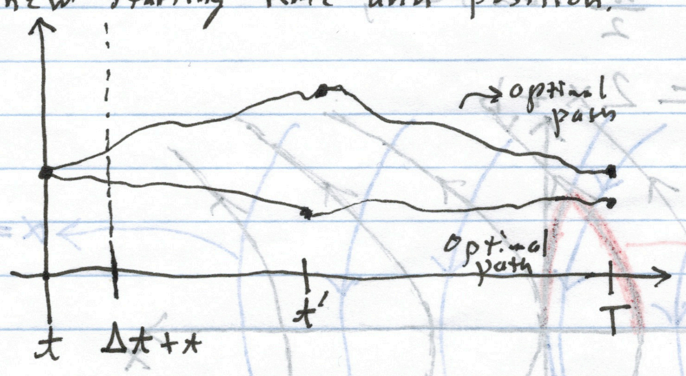
$$\dot{y}(s) = f(s,\alpha(s)) \quad \text{for } t<s<T$$
$$y(t) = x.$$

The dynamic programming principle is that:

$$\boxed{u(x,t) = \min_{\substack{\alpha(s)\in A,\\ t<s<t'}} \left\{ \int_t^{t'} h(y(s),\alpha(s))\,ds + u(y_{x,t,\alpha}(t'),t') \right\}}$$

Interpretation:

It is better to be smart at the beginning then wait until the end. The optimal strategy should do something between t and t'; Starting from t' it should solve the same problem with new starting time and position.



We can derive a P.D.E. for u. We will suppose that $t' = t + \Delta t$ and take the limit as $\Delta t \to 0$. Lets do the formal argument:

1. Lets guess that over the interval $[t, t+\Delta t]$, that $\alpha(t) = a$ a constant.
2. $u(x,t) \approx \min_{a\in A} \left\{ h(x,a)\Delta t + u(x + f(x,a)\Delta t, t+\Delta t) \right\}$

where we have applied

$$\dot{y} = f(y(s),a) \Rightarrow y \approx x + f(x,a)s.$$
$$y(t) = x$$

and dropped terms of order Δt^2.

3. Now, if we assume that v is differentiable then
$$v(x + f(y(t),a)\Delta t, t+\Delta t) \approx v(x,t) + \nabla v \cdot f \cdot \Delta t + v_t \Delta t.$$
4. $$\cancel{v(x,t)} \approx \min_{a\in A}\{h(x,a)\Delta t + \cancel{v(x,t)} + \nabla v \cdot f \cdot \Delta t + v_t \Delta t\}$$
$$\Rightarrow \cancel{v(t)}\ \boxed{v_t + \min_{a\in A}\{h(x,a) + \nabla v \cdot f(t,a)\} = 0.}$$
This is the Hamilton-Jacobi-Bellman equation.
We obtain initial conditions by noticing that:
$$\boxed{v(x,T) = g(x).}$$
Notice, if we define
$$H(\cancel{t},x, \nabla v) = \min_{a\in A}\{h(x,a) + \nabla v \cdot f(t,a)\}$$
we get the ~~?~~ P.D.E
$$\boxed{v_t + H(t,x,\nabla v) = 0.}$$

3. Hopf-Lax Formula

Recall we can obtain "Calculus of variations" problems if we take:
$$\dot{y} = \alpha(t) = f(y(s), \alpha(s)).$$
$$y(t) = x$$
~~Assume that $h(x,a) = h(a)$~~
$$v_t + \max_{a\in A}\{-h(x,a) + \nabla v \cdot a\} = 0$$
$$\cancel{v_t + H(t,x,\nabla v) = 0}$$
$$\boxed{\begin{aligned} &v_t + H(x,\nabla v) = 0\\ &H = \max_{a\in A}\{-h(x,a) + \nabla v \cdot a\} = 0\\ &-h = H^* \to \text{the Frenchel transform.}\end{aligned}}$$
Legendre.

* Hamilton-Jacobi ~~functions generate~~ equations generate calculus of variations problems
* Calculus of variations generate H-J equations for the value function
* The Hamiltonian is the Legendre transform of the Lagrangian

Now, lets assume that h is independant of y.

$$\Rightarrow u(x,t) = \min \int_t^T h(\alpha(s))ds + g(y(T))$$

$$\dot{y} = \alpha, \quad y(t) = x.$$

and we assume that h is convex. Then,

$$-h\left(\int_t^T \frac{\alpha(s)}{T-t}ds\right) = -h\left(\int_t^T \frac{\dot{y}(s)}{T-t}ds\right)$$

$$\geq \frac{1}{T-t}\left(\int_t^T h(\dot{y})\,ds\right)$$

Jensen's Inequality.

With equality if and only if $\dot{y}(s) = \text{constant}$.

Therefore, the optimal trajectory must have α a constant!

This gives us the Hopf-Lax formula:

~~$u(x,t) = \min$~~

$$\boxed{u(x,t) = \max_z \left\{(T-t)\,h\left(\frac{z-x}{T-t}\right) + g(z)\right\}}$$

~~Summary~~

Summary:

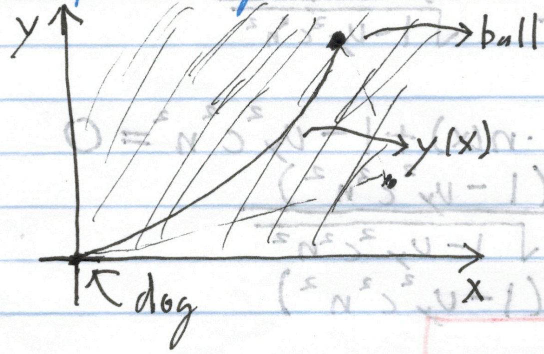
1. If we want to solve
   $$u_t + H(\nabla u) = 0$$
   $$u(x,0) = g(x)$$
2. Let $h = -H^*$
3. We just need to solve the 1-D optimization problem:

~~$u(x,t) = \max_z \{$~~

~~$u(x,t) = \min_z \{ t\,h(z-x$~~

$$\boxed{u(x,t) = \min_z \left\{ t\,h\left(\frac{x-z}{t}\right)\right\}}$$

4. Example – Dog in the sand



$$I[f] = \int_0^1 \frac{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}{c\,n(x)}\,dx, \quad A = \{y \in W^{1,1} : y(0)=0,\ y(1)=1\}.$$

- c top speed of dog
- $n(x)$ the effect of the sand on the dog's speed.

Convert to optimal control:

$$P[\alpha] = \int_0^1 \frac{\sqrt{1+\alpha^2}}{c\,n(x)}\,dx$$

$$\frac{dy}{dx} = \alpha(x), \quad y(0)=0, \quad A = \{\alpha : y(1)=1\}.$$

Define a value function:

$$u(x,y) = \inf_{\alpha \in A} \int_x^1 \frac{\sqrt{1+\alpha^2}}{c\,n(s)}\,ds$$

$$\frac{d\bar{y}}{ds} = \alpha(s), \quad \cancel{y(x)}\ \bar{y}(0) = y.$$

$$u_x + \min_a \left\{ \frac{\sqrt{1+a^2}}{c\,n(x)} + u_y \cdot a \right\} = 0$$

$$\Rightarrow \boxed{a = \frac{\pm u_y\, c \cdot n(x)}{\sqrt{1 - u_y^2 c^2 n^2(x)}}}$$

$\Rightarrow$ Getting the sign correct is not so trivial. We clearly want $a > 0$ and we expect $u_y < 0$ so we must take $-$ root!

$$\Rightarrow U_x + \frac{\sqrt{1+\frac{U_y^2c^2n^2}{1-U_y^2c^2n^2}}}{cn} - \frac{U_y^2cn}{\sqrt{1-U_y^2c^2n^2}} = 0$$

$$\Rightarrow U_x\left(\sqrt{1-U_y^2c^2n^2}\right)c\cdot n(x) + 1 - U_y^2c^2n^2 = 0$$

$$\Rightarrow U_x \cdot c \cdot n(x) = -\frac{(1-U_y^2c^2n^2)}{\sqrt{1-U_y^2c^2n^2}}$$

$$\Rightarrow U_x^2 \cdot c^2 n(x)^2 = (1-U_y^2c^2n^2)$$

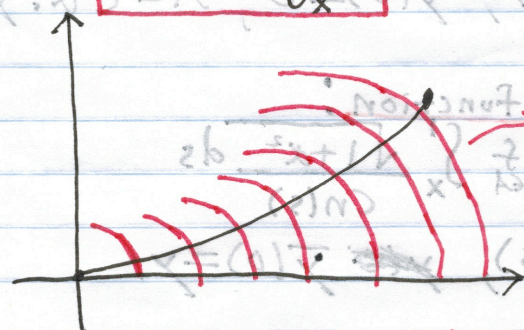
$$\Rightarrow \boxed{|\nabla U|^2 = \frac{1}{c^2n^2(x)}}$$

This is an eikonal equation for the value function!

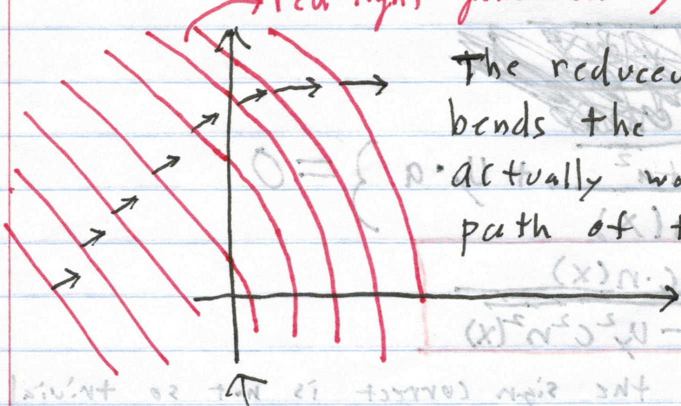
Interpretation:

The contours of the value function are wave fronts, and the outward normal is the tangent to the optimal curve. We see this by noticing that

$$\boxed{a = \pm\frac{U_y}{U_x}}$$



red light generated by E-M radiation



The reduced speed of light in silica bends the light. This is why Snell's law actually works i.e. light takes the shortest path of time.

silica interface

5. Issues

1. $u(x,t)$ may not be C^1 bringing into question our derivation
2. The P.D.E. may have many solutions; for example Bolza-Young:
   $$|u_x| = 1 \quad \text{on } [-1,1]$$
3. How can we use the P.D.E. to actually find the optimal control?

* Answering 1 and 2 requires the viscosity method to select distinguished solutions to the P.D.E. The viscosity solution is the value function.

* Lets answer some questions about rigor.

Theorem:

$$\boxed{\begin{aligned} u(t,x) &= \inf_{\substack{\alpha\in A\\ t<s<t'}} \int_t^{t'} h(y(s),\alpha(s))\,ds + u(t', y_{\alpha,t',x}(t')) \\ y(t) &= x, \quad \dot y(s) = f(y(s),\alpha(s)). \end{aligned}}$$

~~proof~~:

Let $\bar v(t,x) = $ R.H.S. of the above equation. From definition of $u(t,x)$ we have $\forall \varepsilon > 0, \exists \alpha_\varepsilon \in A$ such that

$$\begin{aligned} u(t,x) + \varepsilon &\geq P[\alpha_\varepsilon] \\ &= \int_t^{t'} h(y(s),\alpha(s))\,ds + P[\alpha_\varepsilon]\Big|_{t'}, \end{aligned}$$

with $P[\alpha_\varepsilon]_{t'}$ denoting the payoff functional for α_ε with initial ~~condition~~ value problem:

$$\dot y(s) = f(\alpha_\varepsilon(s), y(s)), \quad y(t') = y_{\alpha_\varepsilon,x,t}(t').$$

$$\Rightarrow u(t,x) + \varepsilon \geq \int_t^{t'} h(y(s),\alpha_\varepsilon(s))\,ds + u(t', y_{\alpha,x,t}(t'))$$

$$\Rightarrow u(t,x) + \varepsilon \geq \bar v(t,x).$$

The other inequality is easy:

$$\bar v(t,x) \leq \inf_{\alpha\in A}\left\{\int_t^T h(y(s),\alpha(s))\,ds + u(T, y_{\alpha,T,x}(T))\right\} = u(x,t).$$

* From this rigorous proof of the dynamic programming ~~princ~~ principle we ~~say~~ see that C^1 continuity/regularity is enough to derive HJB equations.

* Now, to design an optimal control we select $\alpha(s)$ to be the value where the minimum of the HJB equation is obtained.

Steps:

1. Solve HJB equations and compute $u(x,t)$.
2. Use $u(x,t)$ to define an optimal feedback control: For each $x \in \mathbb{R}^n$ and $t \in [0,T]$ set
$$\alpha(x,t) = a \in A$$
to be where the minimum is obtained
$$\Rightarrow u_t + f(x, \alpha(x,t)) \cdot \nabla_x u(x,t) + h(x, \alpha(x,t)) = 0.$$
3. Next solve the O.D.E.
$$\boxed{\begin{aligned} \dot{x}^*(s) &= f(x^*(s), \alpha(x^*(s), s)) \\ x(t) &= x \end{aligned}}$$
4. Define the feedback control
$$\alpha^*(s) = \alpha(x^*(s), s).$$

Theorem - The control $\alpha^*(\cdot)$ defined by this construction is optimal.

proof:
$$P[\alpha^*]\big|_{x,t} = \int_t^T h(x^*(s), \alpha^*(s))\,ds + g(x^*(T)).$$
$$= \int_t^T \left(-u_t(x^*(s), s) - f(x^*(s), \alpha^*(s)) \cdot \nabla_x u(x^*(s), s)\right) ds + g(x^*(T))$$
$$= -\int_t^T \left(u_t(x^*(s), s) + \dot{x}^*(s) \cdot \nabla_x u(x^*(s), s)\right) ds + g(x^*(T))$$
$$= -\int_t^T \frac{d}{ds}(u(x^*(s), s)) + g(x^*(T)).$$
$$= u(x^*(t), s)$$
$$= u(x,t).$$

■

6. Example:

$$\boxed{\begin{aligned} u(x,t) &= \max_{\alpha} \int_t^T e^{-\rho(s-t)} \alpha^q(s)\,ds \\ \frac{dy}{dt} &= ry - \alpha(s), \quad y(t) = x \end{aligned}}$$

with the constraint:

$$\alpha(s) \ge 0, \quad y(s) \ge 0.$$

- $r > 0$ is an interest rate
- $\alpha(s)$ consumption
- $u(x,t)$ is utility.
- $y(s)$ is money $

Step #1

Lets find the HJB equation

$$u(x,t) = \max_{\substack{\alpha(s) \\ t<s<t'}} \left\{ \int_t^{t'} e^{-\rho(s-t)} \alpha^q(s)\,ds + e^{\rho(t-t')} \cdot u(y_{\alpha,x,t'}(t'), t') \right.$$

take $t' = t + \Delta t$

$$u(x,t) \approx \max_{a \ge 0} \left\{ a^q \Delta t + e^{-\rho \Delta t} u(x + (rx-a)\Delta t, t + \Delta t) \right.$$

→ This term is necessary.

$$\approx \max_{a \ge 0} \left\{ a^q \Delta t + (1 - \rho \Delta t)(u(x + (rx-a)\Delta t, t + \Delta t) \right\}$$

$$\cancel{\approx \max_{a \ge 0} \left\{ a^q \Delta t + (1 - \rho \Delta t)(u(x,t) + u_t \Delta t + x + (rx-a) \right.}$$

$$\approx \max_{a \ge 0} \left\{ a^q \Delta t + (1 - \rho \Delta t)(u(x,t) + (rx-a) u_x \Delta t + u_t \Delta t \right\}$$

$$\Rightarrow \boxed{u_t + \max_a \left\{ a^q - \rho u + (rx-a) u_x \right\}}$$

Step #2

Lets try to find the optimal policy. Lets show that

$$u(x,t) = g(t) x^q.$$ → I.e. a seperable solution

It is sufficient to show that

$$u(\lambda x, t) = \lambda^q u(x,t) \Rightarrow g(t) = u(1,t).$$

Let $\lambda\alpha(s)$ be a control for problem starting at λx, where $\alpha(s)$ is the optimal choice starting from x.

$\cancel{\Rightarrow y_{\lambda x} = \lambda y_x(s)}$

$$\frac{dy_\lambda}{dx} = ry_\lambda - \lambda\alpha(s)$$

$$y_\lambda(0) = \lambda x$$

Simple change of variables in O.D.E.

It is clear then that $y_\lambda(x) = \lambda y(x)$. Therefore,

$$v(\lambda x, t) \geq \lambda^q v(x,t).$$

Replacing λ with λ^{-1} gives $\longrightarrow v(\lambda^{-1}x, t) \geq \lambda^{-q} v(x,t)$

$\cancel{v(\lambda x,t)\ v(\lambda x,t) \leq \lambda^{-q} v(x,t)}$ $\Longrightarrow v(x,t) \geq \lambda^{-q} v(\lambda x, t).$

$$v(x,t) \geq \lambda^{-q} v(\lambda x, t)$$

This gives us the result.

$$\boxed{v(\lambda x, t) = \lambda^q v(x,t)}$$

Step 3:

Now lets solve the HJB equation. From step #2 we have that $v_x > 0$, i.e. it is increasing in x.

Maximizing we get that:

$$\boxed{a = \left(\frac{1}{q} v_x\right)^{\frac{1}{q-1}}}$$

$$\Rightarrow a(t) = g(t)^{\frac{1}{q-1}} x \quad \text{since } v(x,t) = x^q g(t).$$

Substituting back into the P.D.E. we get that

$$g_t x^q - \varsigma g x^q + \left(g^{\frac{q}{q-1}}(1-q) + rqg\right)x^q = 0$$

Letting $H(t) = g(t)^{\frac{1}{1-q}}$ we have the linear O.D.E.

$$\boxed{H_t - \nu H = 0, \quad \text{with } \nu = \frac{\varsigma - rq}{1-q}}$$

The solution satisfying $v(x,T) = 0$ gives

$$\boxed{H(t) = \nu^{-1}\left(1 - e^{-\nu(T-t)}\right)}$$

The optimal control is then,

$$\boxed{\alpha(t) = \frac{\nu x}{1 - e^{-\nu(T-t)}}}$$