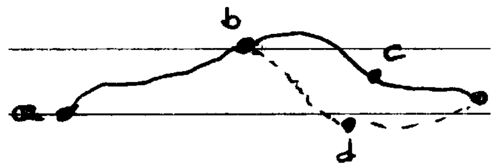
Dynamic Programming

Basic Idea



Suppose start at point a

IF The path a-b-c-F is The minimum cost path From Then b-c-F is The " " b-F

PF By contradiction

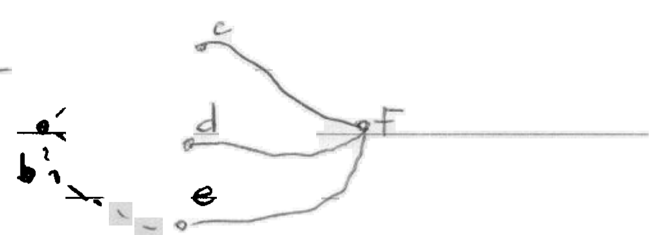
$J_{aF} = J_{ab}\ J_{bF}$

$= J_{ab} + J_{bc}\ J_{cF}$

if $J_{bdF} < J_{bcF}$ Then $J_{abdF} < J_{abcF}$

Bellman: An optimal policy has The property That whatever The initial state & initial decision are, The remaining decisions must constitute an optimal policy with regard To The state resulting from The first decision.

Approach



If we know the optimal cost from c-F, J^*_{cF}, J^*_{dF} J^*_{eF}, Then The optimal cost From point b is The minimum

$J_{bc} + J^*_{cF}$, $J_{bd} + J^*_{dF}$, $J_{be} + J^*_{eF}$

Bellman Equation (discrete-Time)

Suppose we have a decison (aka control) problem

$$\min_{u_k \in \Omega} \sum_{k=0}^{N} L(x_k, u_k)$$

decisions; may come from ~~set~~ of Finite choices.

with constraints

$$x_{k+1} = f(x_k, u_k) \quad \forall k = 0, 1, \ldots N-1$$

$$x_0 = x_0$$

Then This equivalent To

$$\min_{u_0} \left\{ L(x_0, u_0) + \min_{\substack{u_k, \\ k=1\ldots N}} \sum_{k=1}^{N} L(x_k, u_k), \quad \text{with } x_1 = f(x_0, a_0) \right\}$$

Define $V(x_n) = \min_{u_k} \sum_{k=n}^{N} L(x_k, u_k)$

Then above can be written as

$$V(x_n) \quad \min_{u_0} \left[L(x_0, u_0) + V(x_1) \right]$$

Or more generally:

$$\boxed{V(x) = \min_{u \in \Omega} \left[L(x, u) + V(f(x, u)) \right]}$$

Bellman Eqn

(e.g. 1957, Thesis 1946)

Aside: I could include a final cost

$$\min_{u} \sum_{k=0}^{N} L(x_k, u_k) + V(x_{N+1})$$

This becomes The boundary condition on The value function.

Remark: This is necessary and sufficient

Hamilton - Jacobi - Bellman (HJB) (continuous-Time)

Consider:

$$\min_{u} \left\{ \int_0^T L(x,u)\,dt + V_1(x(T)) \right\} \qquad x(0) = x_0$$

$$\dot{x} = F(x,u) \quad \forall t \in [0,t]$$

L: cost rate (scalar)

And define $V(x,t)$ as The value Function

Then

HJB:

$$\boxed{\begin{aligned} &\dot{V} + \min_{u}\left\{ \frac{\partial V}{\partial x} F(x,u) + L(x,u) \right\} = 0 \\ &V(x,T) = V_1(x(T)) \end{aligned}}$$

- Differential equation for value function

Derivation From Bellman: consider interval from t to $t+dt$

$$V(x(t),t) = \min_{u(t)} \left\{ L(x,u)\,dt + \underbrace{V(x(t+dt),\, t+dt)}_{= V(x(t),t) + \dot{V}(x,t)\,dt + \frac{\partial V}{\partial x}\dot{x}\,dt + O(dt^2)} \right\}$$

$$\Rightarrow \quad \min_{u} \left\{ \left[L(x,u) + \dot{V}(x,t) + \frac{\partial V}{\partial x} F(x,u) \right] dt \right\} = 0 \qquad (\text{cancel } V)$$

$$\Rightarrow \quad \dot{V}(x,t) + \min_{u} \left\{ L(x,u) + \frac{\partial V}{\partial x} F(x,u) \right\} = 0 \qquad \left(V \text{ and } \dot{V} \text{ don't depend on } u \right)$$

(Note: $\frac{\partial V}{\partial t}(x,t)$ here is The explicit dependence on t; The dependence on $\frac{dx}{dt}$ is already accounted for)

LQR (again)

Linear system $\dot{x} = f(x, u)$

Quadratic cost $J = \frac{1}{2} \int_0^T x^T Q x + u^T R u \, dt + x(T)^T P_1 \, x(T)$

HJB: $\dot{V} \quad \min_u \left\{ \frac{1}{2}(x^T Q x + u^T R u) + \frac{\partial V}{\partial x}(Ax + Bu) \right\} = 0$

$$\left(V = \int_t^T \frac{1}{2}(x^T Q x + u^T R u)\, dt + x(T)^T P_1 \, x(T) \right)$$

Assume $V(x) = \frac{1}{2} x^T P x$

So $\frac{\partial V}{\partial x} = x^T P$

Then i) $\min_u \left\{ \frac{1}{2}(x^T Q x + u^T R u) + x^T P (Ax + Bu) \right\}$

$\Rightarrow u^T R + x^T P B = 0$

$\Rightarrow u = -R^{-1} B^T P x$

So $\min_u \{ \quad \} = \frac{1}{2} x \, (Q + PBR^{-1}B^T P) x$

$+ \; x^T (PA - PBR^{-1}B^T P) x$

and note $x^T(PA)x = \frac{1}{2}\left(x^T (PA + A^T P) x \right)$

ii) $\dot{P} \quad PA + A^T P + Q - PBR^{-1}B^T P$ from HJB

$P(T) = P_1$

and if $T \to \infty$ then $\dot{V} = 0 \Rightarrow \dot{P} = 0$

Discrete-Time LQR

System $x_{k+1} = A x_k + B u_k$ $\qquad Q \geq 0$, $R \geq 0$

Cost $J = \frac{1}{2}\sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k \quad + \quad x_N^T Q_N x_N$

Define value function $V_k(x) = \frac{1}{2}\sum_{i=k}^{N-1} \cdots$

and assume

$$V_k(x) = \frac{1}{2} x^T P_k x \qquad (\text{so } P_N = Q_N)$$

Bellman:

$$V_k = \min_u \left[\frac{1}{2}\left(x^T Q x + u^T R u\right) + \frac{1}{2} x_{k+1}^T P_{k+1} x_{k+1} \right]$$

$$\vdots$$

$$u_k = -\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A x_k$$

and $\boxed{P_k = A^T P_{k+1} A + Q - A^T P_{k+1} B \left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A}$

("Discrete algebraic Riccati equation" or DARE)

<u>Note</u>: Also useful to consider the best single-step control (i.e. no dynamics)

$$z = z_0 + H u$$

$$J = z^T Q z + u^T R u \qquad R \geq 0,\ Q \geq 0$$

Choose u to minimize J:

$$\frac{dJ}{du} = 0 \ : \quad u^* = -\left(R + H^T Q H\right)^{-1} H^T Q z_0$$

- Weighted least-squares
- If $Q = I$, $\lim_{R \to 0}$ is Moore-Penrose pseudo-inverse