Dynamic Programming in discrete time

[L20912] Ch. 5, pp.156-179

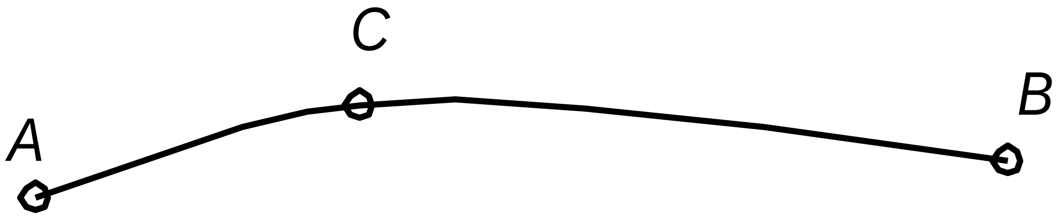
Bellman, years 1950
Actually older roots, starting from the XVII century (Jacob Bernouilli solution of the Brachistochrone problem).

Bellman's Optimum Principle

Traveling in state space from A to B.

Find the optimal path.

If you start at an intermediate point C on the optimal path, the optimal path between C and B is the same as the section of the optimal path between A and B that lies between C and B.



Dynamic Programming algorithm

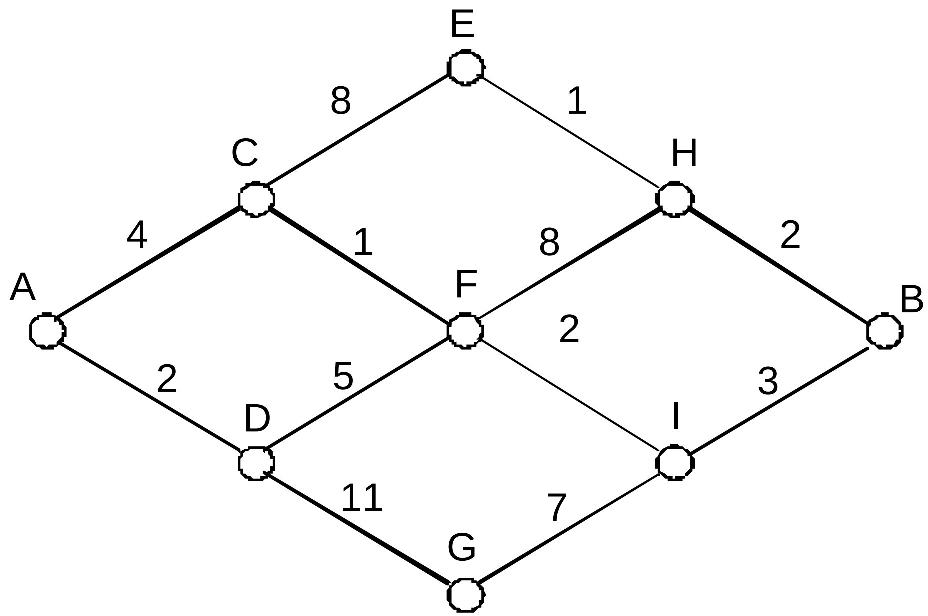
1. Characterize each state by:

   - Its value: The value of the objective function when you follow the optimal path from it to the end.
   - The optimal decision to take when you are on it.

   Do this characterization by progressing backwards, from the last state to the others that precede it.

2. Start from the beginning and follow the optimal decision at each state.

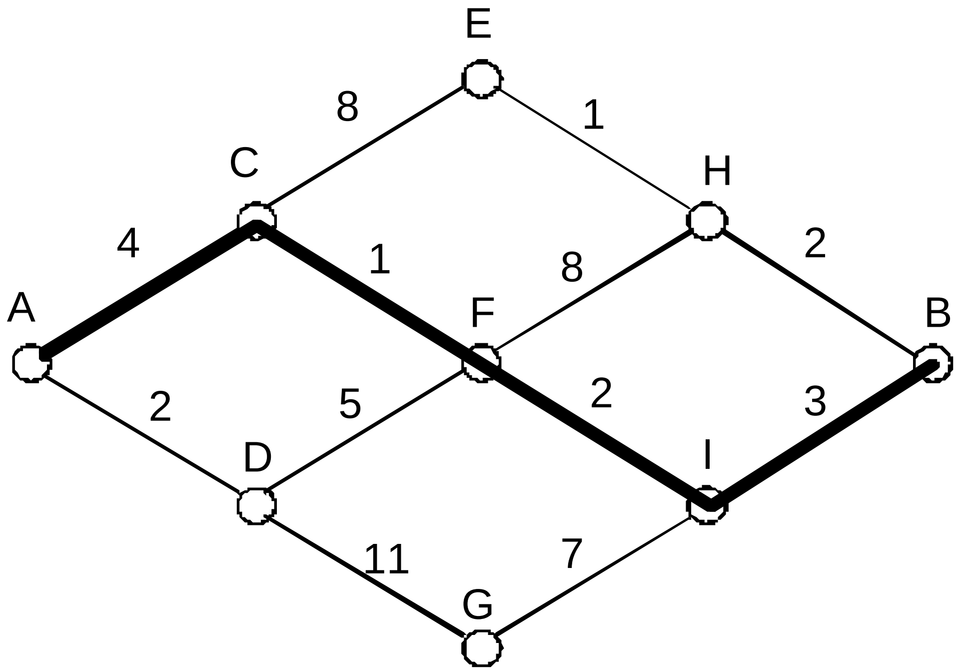
Example: Find the optimal path between A and B



Use the index of the levels of the network as “time”.

Start from B and go backwards, labeling each node with its value and the optimal decision to take whenm you are on it.

Solution for the optimal path:



Dynamic optimization problem

Discrete time nonlinear state model

$$x(k+1) = f(x(k, u(k)), \qquad x(0) = x_0$$

Objective function (*e. g*., a cost to minimize)

$$J\left(u|_{[0,N-1]}\right) = \psi\left(x(N)\right) + \sum_{k=0}^{N-1} L(x(k), u(k))$$

Find the sequence of control values that optimizes the objective function.

Value function

The value function represents the minimum cost that can be obtained when starting from state x, at instant k, up to the final instant N:

$$V(x,k) = \min_{u|_{[k,N-1]}} \left(\psi(x(N) + \sum_{i=k}^{N-1} L(x(i),u(i)) \right)$$

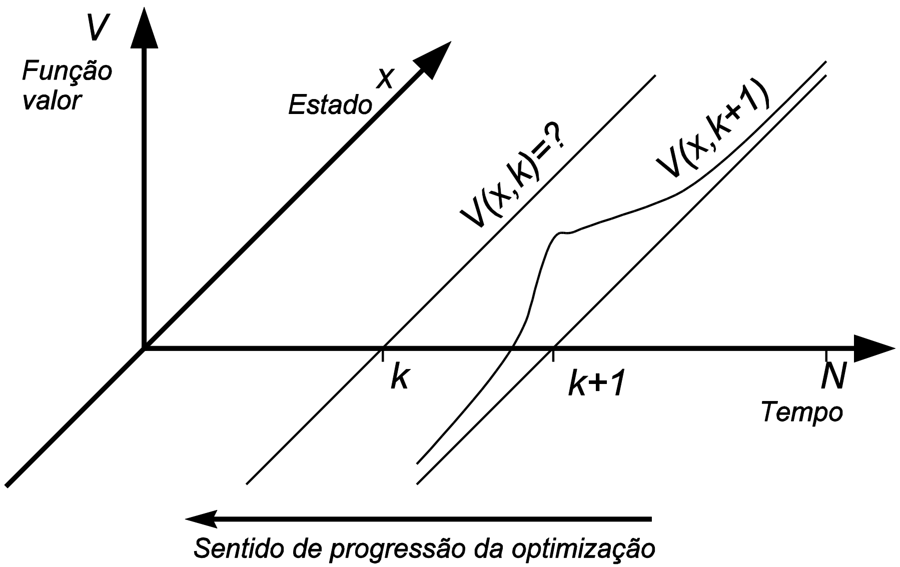
$u|_{[k,N-1]}$ represents the optimal control function restricted to the discrete time span from k to $N-1$.

Hamilton-Jocobi-Bellman (HJB) equation in discrete time

Relates the value function in successive instants of time

$$V(x,k) = \min_{u \in U} \left(L(x,u) + V(f(x,u), k+1) \right)$$

Terminal condition $V(x,N) = \psi(x)$



Linear Quadratic Problem

Linear dynamics

$$x(k+1) = \Phi x(k) + \Gamma u(k), \qquad x(0) = x_0$$

Quadratic cost

$$J = x^T(N) Q_0 x(N) + \sum_{k=0}^{N-1} x^T(k) Q x(k) + u^T(k) R u(k)$$

HJB equation

$$V(x,k) = \min_{u} \left[x^T Q x + u^T R u + V(\Phi x + \Gamma u, k+1) \right]$$

$$V(x,N) = x^T(N) Q_0 x(N)$$

$$V(x,k) = x^T Q x + \min_u \left[u^T R u + V(\Phi x + \Gamma u, k+1)\right]$$

Assume a solution of the form

$$V(x,k) = x^T P(k) x$$

The HJB equation reads

$$x^T P(k) x = x^T Q x + \min_u \left[u^T (R + \Gamma^T P(k+1)\Gamma) u + 2x^T \Phi^T P(k+1)\Gamma u\right] +$$

$$+ x^T \Phi^T P(k+1) \Phi x$$

$$x^T P(k) = x^T Q x + \min_u \left[u^T (R + \Gamma^T P(k+1)\Gamma) u + 2x^T \Phi^T P(k+1)\Gamma u\right] +$$

$$+ x^T \Phi^T P(k+1)\Phi x$$

Comparing with the quadratic form

$$(u - u^*)^T M (u - u^*) = u^* M u - 2u^{*T} M u + u^{*T} M u^*$$

yields

$$M = R + \Gamma^T P(k+1)\Gamma \qquad -u^{*T} M = x^T \Phi^T P(k+1)\Gamma$$

$$u^* = F(k) x \qquad F(k) = -[R + \Gamma^T P(k+1)\Gamma]^{-1} \Gamma^T P(k+1)\Phi$$

Riccati equation

$$P(k) = Q - \Phi^T P(k+1)\Gamma[R + \Gamma^T P(k+1)\Gamma]^{-1}\Gamma^T P(k+1)\Phi + \Phi^T P(k+1)\Phi$$

Dynamic Programming in Continuous time

Basic problem

$$\dot{x} = f(x,u) \qquad x(0) = x_0$$

Functional to minimize

$$J(u) = \Psi\big(x(T)\big) + \int_0^T L\big(x(t),u(t)\big)dt$$

Bellman's Optimal Principle

Let u^* be optimal, with an optimal state x^*.

Take $t_1 \in [0, T]$and let $x_1 = x^*(t_1)$.

Then, $u^*|_{[t_1, T]}$ must be optimal for the problem

$$\text{Minimize } \Psi\big(x(T)\big) + \int_{t_1}^{T} L\big(x(t), u(t)\big) dt$$

$$\text{Subject to } \dot{x} = f(x, u) \qquad x(t_1) = x_1$$

The value function

$$V(t,x) = \min_{u|_{[t,T]}} \Psi\big(x(T)\big) + \int_t^T L\big(x(t),u(t)\big)dt$$

Subject to $\dot{x} = f(x,u) \quad x(t) = x$

The value function represents the optimal cost to from a given state x , at a given time t to the terminal instant T.

Hamilton-Jacobi-Bellman (HJB) equation

$$\frac{\partial V(t,x)}{\partial t} + \min_{u \in U}\left(\nabla_x V(t,x) f(x,u) + L(x,u)\right)$$

Verification theorem

If u satisfies the HJB and $\nabla_x V(t,x) f(x,u) + L(x,u) \leq \nabla_x V(t,x) f(x,v) + L(x,v)$ for any $v \in U$, then u is optimal.