Recall model from last class

$$x_{k+1} = f_k(x_k, u_k) \quad , \quad k \in \{0, 1, \dots, N-1\}$$

k - stage or time u_k controls $\in U_k(x_k)$

x_k - state of system

f - rule governing dynamics

cost ~~given by~~ functions : $g_0, g_1, \dots, g_N$

For initial condition x_0 and control sequence $\vec{u} = u_0, u_1, \dots, u_{N-1}$

cost is given by

$$J(x_0; \vec{u}) = g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k)$$

Optimal control problem :

solve

$$J^*(x_0) = \min\{J(x_0; \vec{u}) ; u_k \in U_k(x_k), x_{k+1} = f(x_k, u_k)\}$$

Introduced tail problems for $k = 0, 1, \dots, N$

$$\cancel{J_k(x_0; \vec{u})} \quad J_k(x_k; u_k, \dots, u_{N-1}) = \sum_{m=k}^{N-1} g_m(x_m, u_m) + g_N(x_N)$$

$$J_k^*(x_k; u_k, \dots, u_{N-1}) = \min\{J_k(x_k; u_k, \dots, u_{N-1}) ; u_m \in U(x_m), x_{m+1} = f(x_m, u_m)\}$$

Overview of DP Algorithm

1) Solve tail problem with only last stage

1) Solve all tail problems with length 1.

2) Solve all ~~tail~~ tail problems of length 2

⋮

N-k) Solve all tail problems of length k

DP Algorithm

a) ~~Set~~ Tail ~~su~~ problem of length 0 : $J_N^*(x_N) = g_N(x_N)$

b) For $k = N-1, N-2, \ldots, 1, 0$ ~~$k=0,1,2,\ldots,N-1$~~ solve

$$J_k^*(x) = \min_{u \in U(x)} \left[g_k(x,u) + J_{k+1}^*(f(x,u)) \right] \qquad \text{(DPE) (B.E.)}$$

for all x.

→ when $k = N-1$ this is ~~subprob~~ tail problem of length 1.

At end of for loop

$$J_0^*(x_0) = J^*(x_0) = \min \left(J(x_0; u_0, u_1, \ldots, u_{N-1});\ u_k \in U(x_k),\ x_{k+1} = f(x_k,u_k) \right)$$

will prove in HW.

J_k^* is ~~defined~~ often called the 'value function' or 'optimal cost to go function'

- After obtaining value functions $J_0^*, J_1^*, \ldots, J_N^*$ we can use these functions to get optimal controls

$$u_0^* \in \operatorname*{argmin}_{u \in U(x_0)} \left[g_0(x_0,u) + J_1^*(f(x_0,u)) \right]$$

set $x_1^* = f_0(x_0, u_0^*)$ and for $k = 1, 2, \ldots, N-1$

$$u_k^* \in \operatorname*{argmin}_{u \in U(x_k^*)} \left[g_k(x_k^*, u) + J_{k+1}^*(f(x_k^*, u)) \right]$$

and $x_{k+1}^* = f(x_k^*, u_k^*)$

IE 5571 Lecture 2

Recall ~~problem~~ Model from last class

$$x_{k+1} = f_k(x_k, u_k), \quad k \in \{0,1,\dots,N-1\}$$

k - stage

x_k - state, f_k - dynamics

u_k - controls $\in U_k(x_k)$ ← feasible controls

cost $g_0, g_1, \dots, g_N$

For initial condition x_0, ~~cost~~ and control sequence $u_0, u_1, \dots, u_{N-1}$

cost is given by

$$J(x_0; u_0, u_1, \dots, u_{N-1}) = g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k)$$

Optimal control problem:

Find $J^*(x_0) = \min\{J(x_0; u_0, \dots, u_{N-1}); u_k \in U_k(x_k), x_{k+1} = f(x_k, u_k)\}$

↑ solve with D.P. important tool is Value function and DPE

$V_N(x) = g_N(x)$, $V_n(x) = \min\limits_{u_n}\left(g_n(x, u_n) + V_{n+1}^{(f_n(x))}\right)$

Agent start with wealth X_0.

At each ~~ti~~ stage $n \in \{0,1,\dots,N-1\}$, agent consumes u_n proportion of their wealth, X_n.

Remainder goes into bank ~~and~~ at interest rate r.

Set $R = 1 + r$ then

$$X_{n+1} = R X_n (1 - u_n), \quad n = 0, 1, \dots, N-1$$

Consuming wealth x at $n \in \{0,1,\dots,N-1\}$ agent receives utility $g(x)$ and ~~$g_N(x)$ at~~ and at $n = N$ utility is $g_N(x)$

~~So total utility over time period $n \in \{0,1,\dots,N-1\}$ is~~

Goal of agent is to come up with a sequence of consumptions to minimize

$$J^*(x_0; u_1,\dots,u_n) = g_N(x_N) + \sum_{n=0}^{N-1} g(u_n x_n)$$

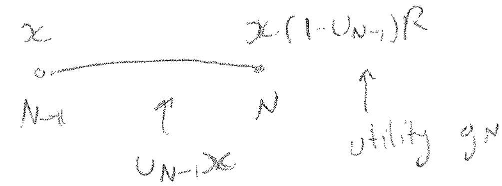
~~We solve this problem using dynamic programming:~~

i.e. find $J^*(x_0) = \min_{u_0,u_1,\dots,u_{N-1}} J^*(x_0; u_0,\dots,u_{N-1})$

Solve using DP. (solve tail subproblems)

Start at $n=N-1$, with wealth x.

Maximize $g(u_{N-1}x) + g_N(Rx(1-u_{N-1}))$ over u_{N-1}



optimal $u^*_{N-1}(x)$ tells us how much to consume ~~at time~~

Maximum possible utility we can have is

optimal value of wealth x at $n=N-1$

$$V_{N-1}(x) = \max_{0 \le u_{N-1} \le 1} \left[g(u_{N-1}x) + g_N(Rx(1-u_{N-1})) \right]$$

For given $g + g_N$ we can solve this by basic calculus.

Suppose $n=N-2$ and wealth is x.

Agents goal is to maximize

V_{N-1} - value of weal

$$g(u_{N-2}x) + V_{N-1}(R(1-u_{N-2})x)$$

then maximal value of wealth at $n=N-2$ is

$$V_{N-2}(x) = \max_{0 \le u_{N-2} \le 1} \left[g(u_{N-2}x) + V_{N-1}(R(1-u_{N-2})x) \right]$$

Example

~~II~~ A salesperson based in NYC needs to visit Miami, Dallas & Chicago and then return to NYC.

Given the following distance matrix

	NYC	Miami	Dallas	Chicago
NYC	—	1334	1559	809
Miami	1334	—	1343	1397
Dallas	1559	1343	—	921
Chicago	809	1397	921	—

What order should they visit cities to minimize distance traveled?

Solution: Would like to use DP to solve problem, but not clear how.

Stage j - j^{th} stop of trip, $j=0$ starting and $j=4$ final stop.

State a bit subtle: $x_j = \{c_1, .., c_j\}$, ~~c_j is~~ $c_1, .., c_j$ is list of cities visited, c_j - last city visited. $\{c_1, .., c_{j-1}\} = L$

~~$f_j(L, c_j)$~~ $f_j(c_j, L)$ - minimal distance that must be travelled to complete the tour if salesperson at j^{th} stop in city c_j and L is all cities visited.

~~When $j=3$, all we have to do is return to~~

For ease of notation, let list cities as $\{N, M, D, C\}$ and ~~d_{ij}~~ ~~d_{IJ}~~ = distance from city $I \to$ city J.

$d(I, J)$

(3) Using (DPE) we have

$$f_1(c_1, L) = \min_{J \notin L} \{ d(c_1, J) + f_2(J, L \cup \{J\}) \}$$

Therefore

i) $$f_1(M, \{M\}) = \min_{J \in \{D, C\}} [d(M, J) + f_2(J, \{M, J\})]$$

$$= \min \{ d(M, D) + f_2(D, \{M, D\}), d(M, C) + f_2(C, \{M, C\}) \}$$

$$= \min \{ 1343 + 1730, 1397 + 2480 \} = 3073$$

min achieved at $J = D$

ii) $$f_1(D, \{D\}) = \min_{J \in \{M, C\}} [d(D, J) + f_2(J, \{D, J\})]$$

$$= \min (d(D, M) + f_2(M, \{D, M\}), d(D, C) + f_2(C, \{D, C\}))$$

$$= \min (1343 + 2206, 921 + 2731) = 3549, \quad J^* = M$$

iii) $$f_1(C, \{C\}) = \min_{J \in \{M, D\}} (d(C, J) + f_2(J, \{C, J\})) = 3598$$

$$J^* = D$$

Finally $$f_0(N, \{N\}) = \min_{J \in \{M, D, C\}} [d_{NJ} + f_1(J, \{J\})]$$

$$= \min [d(N, M) + f_1(M, \{M\}), d(N, D) + f_1(D, \{D\}), d(N, C) + f_1(C, \{C\})]$$

$$= \min [1334 + 3073, 1559 + 3549, 809 + 3598] = 4407, \quad J^* \in \{M, C\}$$

Optimal Tour: $N \to M \to D \to C \to N$ or $N \to C \to D \to M \to N$

~~When $j=3$~~ Can write DPE for this problem as

$$f_j(c_j, L) = \min_{J \notin L} \{ d(c_j, J) + f_{j+1}(J, L \cup \{J\}) \} \qquad \text{(DPE)}$$

with $f_3(c_3, \{M, D, C\}) = d(c_2, N)$ ~~$d_{c_3 N}$~~ since at stage 3

we have to travel from current city to NYC.

Then with DPE we have

$$f_2(c_2, L) = \min_{J \notin L} (d(c_2, J) + f_3(J, L \cup \{J\}))$$

$$= \min_{J \notin L} (d(c_2, J) + d(J, N))$$

Note that to use (DPE) when $j=1$ we will need to evaluate ~~$f_2(c_2, L)$~~ $f_2(\cdot, \cdot)$ at all possible states

(c_2, L):

$f_2(M, \{M, D\}) = d_{MC} + d_{CN} = 2206$

$f_2(D, \{M, D\}) = d(D, C) + d(C, N) = 1730$

$f_2(M, \{M, C\}) = d(M, D) + d(D, N) = 2902$

$f_2(C, \{M, C\}) = 2480$

$f_2(D, \{D, C\}) = 2677$

$f_2(C, \{D, C\}) = 2731$

In general if the agent has wealth x at time n then agent should solve $\swarrow$ $\boxed{\text{DPE.}}$

$$V_n(x) = \max_{0 \le u_n \le 1} \left[g(u_n x) + V_{n+1}(R(1-u_n)x) \right]$$

Intuitively $V_n(x)$ = maximal total utility possible from time n on starting from wealth x at time n.

$\boxed{\text{Do TSP FIRST}}$

$\boxed{\text{Stochastic Dynamic Programming}}$

Instead of assuming ~~X_n~~ $X_{n+1} = f(X_n, u_n)$ is exact we assume $X_{n+1} = f(X_n, u_n; \xi_{n+1})$ where $\{\xi_{n+1}\}_{n \ge 1}$ are a sequence i.i.d R.V.'s.

~~Co~~ Finite stage problem: ~~Re~~ $n \in \{0, 1, \dots, N-1\}$

Cost: $g_0, g_1, \dots, g_N$.

Since $\{X_n\}_{n=0}^{N}$ is a stochastic process we are interested in expected cost:

$$J(x_0; \{u_n\}_{n=0}^{N-1}) = E_{x_0}\left[g_N(X_N) + \sum_{n=0}^{N-1} g_n(X_n, u_n) \right]$$

E_{x_0} $\leftarrow$ expected value conditioned on $X_0 = x_0$

Two differences between stochastic and deterministic dynamic programming

1) In deterministic d.p. ~~ea~~ system dynamics are given by once we choose control sequence. For stochastic system choosing control sequence doesn't determine dynamics. Therefore we need more than optimal control sequence we need 'optimal policies', i.e. for stage k optimal policy $\mu_k^*(x)$ tells us optimal action if we are in ~~stage k~~ state x on stage k.

2) Evaluating our objective function J requires evaluating an expected value.

If we define the minimal possible cost as

$$J^*(x_0; \{U_n\}_{n=0}^{N-1}) = \min_{U_0,\dots,U_{N-1}} J(x_0; \{U_n\}_{n=0}^{N-1})$$

then we can set up a D.P.E.

$$V_N(x) = g_N(x)$$

$$V_n(x) = \min_{u \in U(x)} \left(g_n(x,u) + E[V_{n+1}(f(x,u,\xi))] \right), \quad n \in \{0,1,\dots,N-1\}$$

can prove that $V_0(x) = J^*(x_0; \{U_n\}_{n=0}^{N-1})$

Example: Agent starts with wealth X_0

- At each stage $n \in \{0,1,\dots,N-1\}$ agent consumes proportion u_n of their wealth X_n.
- Remainder goes into bank and earns interest at rate r_n. r_n a RV with CDF F.

Assume $\{r_n\}_{n\geq 1}$ are i.i.d RV's. (5)

Define $R_n = 1 + r_n$ then

$$X_{n+1} = R_n X_n (1 - u_n)$$

Consuming wealth x at stage $n \in \{0, 1, \dots, N-1\}$ gives utility $g(x)$ and $g_N(x)$ at stage N.

~~Discuss~~ $\boxed{\text{Solve with DP}}$

$\underline{n = N-1}$ wealth $= x$

Maximize $g(u_{N-1}x) + E[g_N(R_N x(1-u_{N-1}))]$ over u_{N-1}

↳ we now have to evaluate expected reward in next stage because we don't know interest rates yet.

Maximum possible utility for wealth x at stage $N-1$:

$$V_{N-1}(x) = \max_{0 \leq u_{N-1} \leq 1} [g(u_{N-1}x) + g_N(R_N x(1-u_{N-1}))]$$

~~optimal $u^*_{N-1}(x)$~~ – minimizing u_{N-1} $u^*_{N-1}(x)$ will give us optimal policy at stage $N-1$.

For $n = 0, 1, \dots, N-1$

$$V_n(x) = \max_{0 \leq u_n \leq 1} \left(g(u_n x) + \underline{E}[V_{n+1}(R_n(1-u_n)x)] \right)$$

$V_0(x)$ – optimal possible utility over total horizon and $u^*_n(x)$ sequence of optimal policies.