MS-E2148 Dynamic optimization

Lecture 8

Contents

- The role of perfect state information
- Inperfect state information
- Application: inventory control with piecewise continuous cost
- Application: portfolio problem
- How to model utility mathematically
- Material Bertsekas 4.2, 4.3 and 5.1

Recap

- Shortest path problem was one special case of problem suitable for DP algorithm

Perfect state information in basic problem

- Special case; it is more common that the decision maker does **not** have perfect information about the state
  - Sensors can be imprecise (e.g. how to measure the wind speed for jumbo jet flying in the air for 900 kph)
  - Getting perfect information can be costly or technically impossible (e.g. in gene mapping)
- Imperfect state information can be imprecise *observations* and the control laws may depend on all possible realizable states
- Note: Kalman filter

Imperfect state information

- Let us modify the basic problem the following way: instead of knowing x_k, we observe

$$z_0 = h_0(x_0, v_0), \quad z_k = h_k(x_k, u_{k-1}, v_k), \quad k \geq 1.$$

- Observation $z_k \in Z_k$
- The noise in the observation may depend on the state and past states, controls and noise (w and v)
- The noise w may depend on state and control but not past noises (w and v)

Information vector

- The available information is given in vector I_k, $I_0 = z_0$,

$$I_k = (z_0, z_1, \ldots, z_k, u_0, u_1, \ldots, u_{k-1}), \quad k \geq 1$$

- The control is chosen for each information vector
- We get the basic problem with small modifications:

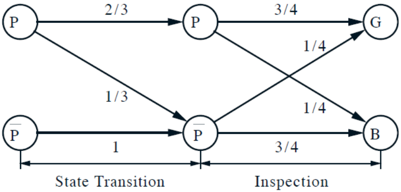
$$J_\pi = E(g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(I_k), w_k))$$

$$x_{k+1} = f_k(x_k, \mu_k(I_k), w_k), \quad k \geq 0$$

$$z_0 = h_0(x_0, v_0), \; z_k = h_k(x_k, \mu_{k-1}(I_{k-1}), v_k), \; k \geq 1$$

Application: machine repair

- The machine can be in good state P or in bad state $\bar{P}$
- In every state the machine is examined (three period problem)
- Two possibilities: it looks good G or bad B



- Control: continue C or stop S and check the state and fix
- Costs: $g(P, C) = 0, g(P, S) = 1, g(\bar{P}, C) = 2, g(\bar{P}, S) = 1$

Application: machine repair

- Information at stage 0 is $I_0 = z_0$ and at stage 1 $I_1 = (z_0, z_1, u_0)$
- Find $\mu_0(I_0)$ and $\mu_1(I_1)$

$$E(g(x_0, \mu_0(z_0)) + g(x_1, \mu_1(z_0, z_1, \mu_0(z_0))))$$

- DP: Start from $J_2(I_2) = 0$

$$\begin{aligned} J_k(I_k) = \min \big[& P(x_k = P|I_k) g(P, C) + P(x_k = \bar{P}|I_k) g(\bar{P}, C) \\ & + E(J_{k+1}(I_k, C, z_{k+1})|I_k, C), P(x_k = P|I_k) g(P, S) \\ & + P(x_k = \bar{P}|I_k) g(\bar{P}, S) + E(J_{k+1}(I_k, S, z_{k+1})|I_k, S) \big] \end{aligned}$$

Application: machine repair

- Last stage: Compute $J_1(I_1)$ for each of the eight information vectors
- Cost for S is 1 and for C $2P(x_1 = \bar{P}|I_1)$
- thus $J_1(I_1)$ is minimum of these
- $P(x_1 = \bar{P}|I_1)$ is computed with Bayes rule
- $I_1 = (G, G, S)$:
  $P(x_1 = \bar{P}|G, G, S) = P(x_1 = \bar{P}, G, G, S)/P(G, G, S)$

$$= \frac{1/3 \cdot 1/4(2/3 \cdot 3/4 + 1/3 \cdot 1/4)}{(2/3 \cdot 3/4 + 1/3 \cdot 1/4)^2} = 1/7$$

- $J_1(G, G, S) = 2/7$ so it is better to continue C. Similarly for $I_1(B, G, S)$

Application: machine repair

- $I_1 = (G, B, S)$:
  $P(x_1 = \bar{P} | G, B, S) = P(x_1 = \bar{P}, G, B, S) / P(G, B, S)$
  $$= \frac{1/3 \cdot 3/4 (2/3 \cdot 3/4 + 1/3 \cdot 1/4)}{(2/3 \cdot 3/4 + 1/3 \cdot 1/4)(2/3 \cdot 1/4 + 1/3 \cdot 3/4)} = 3/5$$
- $J_1(G, B, S) = 1$ so it is better to stop S. Similarly, for $I_1(B, B, S)$
- It is possible to compute the other information states the same way

Application: machine repair

- First stage: Compute $J_0(I_0)$: for $I_0 = G$ and for $I_0 = B$
- Cost for C: $2P(x_0 = \bar{P}|I_0) + E(J_1(I_0, z_1, C)|I_0, C)$

  $= \ldots + P(z_1 = G|I_0, C)J_1(I_0, G, C) + P(z_1 = B|I_0, C)J_1(I_0, B, C)$
- Cost for S: $1 + E(J_1(I_0, z_1, S)|I_0, S)$

  $= 1 + P(z_1 = G|I_0, S)J_1(I_0, G, S) + P(z_1 = B|I_0, S)J_1(I_0, B, S)$
- $J_0(I_0)$ is the minimum of these and
  $J^* = P(G)J_0(G) + P(B)J_0(B)$

Application: inventory control (continued)

- Inventory, independent and restricted demands:

$$x_{k+1} = x_k + u_k - w_k, \qquad k = 0, 1, ..., N-1$$

- Let us assume that the inventory/shortage cost of amount z is

$$r(z) = p \max(0, -z) + h \max(0, z), \qquad p, h \geq 0$$

- The total cost of inventory is

$$E\left\{ \sum_{k=0}^{N-1} \left(cu_k + p \max(0, w_k - x_k - u_k) + h \max(0, x_k + u_k - w_k) \right) \right\}$$

- Let us assume that $p > c$.

Application: inventory control

- With DP-algorithm we get for all $k = 0, ..., N-1$

$$\begin{aligned} J_N(x_N) &= 0, \\ J_k(x_k) &= \min_{u_k \geq 0} \Big[cu_k + H(x_k + u_k) \\ &\qquad + E\{J_{k+1}(x_k + u_k - w_k)\} \Big] \end{aligned} \tag{1}$$

where

$$H(y) = pE\{\max(0, w_k - y)\} + hE\{\max(0, y - w_k)\}$$

Application: inventory control

- Let us change the variables $y_k = x_k + u_k$, and we get (1) into the following form

$$\min_{y_k \geq x_k} [cy_k + H(y_k) + E_{w_k}\{J_{k+1}(y_k - w_k)\}] - cx_k \qquad (2)$$

- Does the minimum exist? We examine the convexity:
  - Function $H(\cdot)$ is convex since max-functions $\max(0, w_k - y_k)$ and $\max(0, y_k - w_k)$ are convex in y_k for each w_k
  - The expectation preserves convexity, so the remaining requirement is that J_{k+1} is convex

Application: inventory control

- *Assume* that the minimum (2) with respect to y_k exists and denote it by constant S_k; then the minimizing y_k is equal to S_k if $x_k < S_k$ and otherwise equal to x_k
- Thus, the optimal control law for inventory is based on the sequence $\{S_0, S_1, ..., S_{N-1}\}$ and is

$$\mu_k^*(x_k) = \begin{cases} S_k - x_k, & x_k < S_k, \\ 0, & x_k \geq S_k \end{cases} \tag{3}$$

- For each k, the constant S_k minimizes function

$$G_k(y) = cy + H(y) + E\{J_{k+1}(y - w)\} \tag{4}$$

Application: inventory control

- The optimality of control law (3) can be proven if can be shown that the cost-to-go J_k (and functions G_k) are convex and $\lim_{|y|\to\infty} G_k(y) = \infty$ so that the minimizing constants S_k exist
- The cost-to-go with the control law (3) can be derived with the DP algorithm (1) at stage $N-1$

$$J_{N-1}(x_{N-1}) = \begin{cases} c(S_{N-1} - x_{N-1}) + H(S_{N-1}), & x_{N-1} < S_{N-1} \\ H(x_{N-1}), & x_{N-1} \geq S_{N-1} \end{cases} \tag{5}$$

- This is a convex function since H is convex
- This argument can be repeated inductively backwards

Application: portfolio problem

- How to invest a given amount of wealth in different sources (stocks, bonds, buildings, startups etc.)
- How to model the problem if the investment decisions are made consecutively in many periods and the aim is to maximize some final wealth?

Application: portfolio problem

One period problem

- x_0 is the initial wealth
- Let there be n uncertain sources that have *stochastic* revenue $e_1, \ldots, e_n$
- Beside the uncertain ones, it is possible to invest in certain account that gives *deterministic* revenue s

- Let the investment in the uncertain sources be $u_1, \ldots, u_n$ and the certain one is then $x_0 - u_1 - \ldots - u_n$

Application: portfolio problem

One period problem

- Wealth after one period is

$$\begin{aligned} x_1 &= s(x_0 - u_1 - \ldots - u_n) + \sum_{i=1}^{n} e_i u_i \\ &= s x_0 + \sum_{i=1}^{n} (e_i - s) u_i \end{aligned}$$

- What are we maximizing? Let us assume that the **expected utility** of the wealth is

$$E\{U(x_1)\}$$

where $U(\cdot)$ is *concave* and twice differentiable function

Application: portfolio problem

One period problem

- Let us assume that the expectation is well-defined and finite for all x_0 and u_i
- No other assumptions are needed, not even $x_0 \geq \sum u_i$ (which allows to invest with loan money)

- Let us try to find an optimal control law of the form

$$\mu^{i*}(x_0) = \beta^i(x_0)(a + bsx_0) \tag{6}$$

where $\beta^i(x_0)$, $i = 1, ..., n$, are some differentiable functions and a, b are parameters that define the shape of utility function

Application: portfolio problem

One period problem

- We will find out if our assumed form (6) is optimal and can we simplify it

- First-order condition for utility maximization is

$$\frac{\partial E\{U(x_1)\}}{\partial u_i} = 0 \qquad \forall x_0, \quad i = 1, ..., n \tag{7}$$

- Let us substitute the system equation (6) with the assumed form (6) into (7), which gives us n equations, $\forall i = 1, ..., n$:

$$E\Big\{U'\Big(sx_0 + \sum_{j=1}^{n}(e_j - s)\beta^j(x_0)(a + bsx_0)\Big)(e_i - s)\Big\} = 0 \tag{8}$$

Application: portfolio problem

One period problem

- We can solve the equations (8), but we can examine what they imply based on the control law (6) when x_0 changes. By taking the partial differential with respect to x_0:

$$\Rightarrow E\{M \cdot U''(x_1)(a + bsx_0)\} \cdot \begin{bmatrix} \frac{\partial \beta^1(x_0)}{\partial x_0} \\ \vdots \\ \frac{\partial \beta^n(x_0)}{\partial x_0} \end{bmatrix} = -K \qquad (9)$$

where the $n \times n$ matrix is

$$M = \begin{bmatrix} (e_1 - s)^2 & \cdots & (e_1 - s)(e_n - s) \\ \vdots & & \vdots \\ (e_n - s)(e_1 - s) & \cdots & (e_n - s)^2 \end{bmatrix}$$

Application: portfolio problem

One period problem

- and n vector

$$K = \begin{bmatrix} E\{U''(x_1)(e_1 - s)s\big(1 + \sum_{i=1}^{n}(e_i - s)\beta^i(x_0)b\big)\} \\ \vdots \\ E\{U''(x_1)(e_n - s)s\big(1 + \sum_{i=1}^{n}(e_i - s)\beta^i(x_0)b\big)\} \end{bmatrix}$$

- Now, we should find out under what conditions $\beta^i(x_0)$ satisfy equation (9). It contains the utility function $U(\cdot)$ on both sides, so we need to make some assumptions about it

Application: portfolio problem
One period problem

- Let us assume that the utility function satisfies

$$-\frac{U'(y)}{U''(y)} = a + by, \qquad \forall y \tag{10}$$

  where a, b are parameters of the utility function
- From this, we get the terms for (9):

$$\begin{aligned} U''(x_1) &= -\frac{U'(x_1)}{a+b(sx_0+\sum_i(e_i-s)\beta^i(x_0)(a+bsx_0))} \\ &= -\frac{U'(x_1)}{(a+bsx_0)(1+\sum_i(e_i-s)\beta^i(x_0)b)} \end{aligned} \tag{11}$$

Application: portfolio problem

One period problem

- Let us substitute (11) into (9)
- Vector K is null vector since by the first-order condition (8) we get $E\{U'(x_1)(e_i - s)\} = 0$
- We can assume that matrix M in equation (9) is non-singular. This holds in many cases, so we can multiply by M on both sides of (9)

$$\frac{\partial \beta^i(x_0)}{\partial x_0} = 0 \qquad \forall i = 1, ..., n$$

Application: portfolio problem

One period problem

- **Result**: The expected control law is linear of the form

$$\mu^{i*}(x_0) = \alpha^i(a + bsx_0), \qquad \forall i = 1, ..., n \tag{12}$$

where α^i are constants

- We can also show that it holds for the portfolio value $J(x_0) = \max_{u_i} E\{U(x_1)\}$ that

$$-\frac{J'(x_0)}{J''(x_0)} = \frac{a}{s} + bx_0, \qquad \forall x_0 \tag{13}$$

- This we get by differentiating $J(x_0)$ and reordering the terms; see Bertsekas Sect. 4.3

Application: portfolio problem

Multiperiod problem

- The investments are made in multiple periods
- x_k is the wealth at the beginning of stage k
- u_i^k is the amount invested at the beginning of stage k to uncertain source i
- e_i^k is the revenue of i at stage k
- s_k is the certain revenue at stage k
- The system is

$$x_{k+1} = s_k x_k + \sum_{i=1}^{n} (e_i^k - s_k) u_i^k, \quad k = 0, 1, ..., N-1$$

Application: portfolio problem

Multiperiod problem

- The revenues $e^k = (e_1^k, ..., e_n^k)$, $k = 0, ..., N-1$ are vectors; assume that they are independent $\forall k$ and yield finite expected values
- The objective is to maximize the final wealth $E\{U(x_N)\}$, assuming that the utility function satisfies (10)
- DP-algorithm:

$$J_N(x_N) = U(x_N)$$

$$J_k(x_k) = \max_{u_1^k,...,u_n^k} E\Big\{J_{k+1}\Big(s_k x_k + \sum_i (e_i^k - s_k) u_i^k\Big)\Big\}$$

Application: portfolio problem

Multiperiod problem

- By the principle of optimality, it is optimal to act like in the single period problem, and the optimal control law (vector) is

$$\mu^*_{N-1}(x_{N-1}) = \alpha_{N-1}(a + bs_{N-1}x_{N-1}) \tag{14}$$

  where α_{N-1} is a vector of length n

- Cost-to-go at stage $N-1$ is

$$J_{N-1}(x_{N-1}) = E\{U(s_{N-1}x_{N-1} + \sum_i (e_i^{N-1} - s_{N-1})\alpha_{N-1}(a + bs_{N-1}x_{N-1}))\}$$

- By differentiating the cost-to-go:

$$J'_{N-1}(x_{N-1}) = E\left\{U'(x_N)\cdot\left(s_{N-1}+\sum_i (e_i^{N-1}-s_{N-1})\alpha_{N-1} b s_{N-1}\right)\right\}$$

$$J''_{N-1}(x_{N-1}) = E\left\{U''(x_N)\cdot\left(s_{N-1}+\sum_i (e_i^{N-1}-s_{N-1})\alpha_{N-1} b s_{N-1}\right)^2\right\}$$

- Let us use again the term (11) for $U''(x_N)$, which gives:

$$-\frac{J'_{N-1}(x)}{J''_{N-1}(x)} = \frac{a}{s_{N-1}} + bx \qquad (15)$$

Application: portfolio problem

Multiperiod problem

- By similar differentiation, we get the optimal control law for stage $N-2$ by DP algorithm

$$\mu^*_{N-2}(x_{N-2}) = \alpha_{N-2}\left(\frac{a}{s_{N-1}} + bs_{N-2}x_{N-2}\right) \quad (16)$$

- Similarly, proceeding backwards we get the optimal control law for stage k

$$\mu^*_{k}(x_{k}) = \alpha_{k}\left(\frac{a}{s_{N-1}\cdots s_{k+1}} + bs_{k}x_{k}\right) \quad (17)$$

Application: portfolio problem

Multiperiod problem

- **Result**: the investment in multiperiod problem is similar to the single period problem
- Multiple stages are seen in the control law (17) by dependency of the certain revenues $s_{k+1}, ..., s_{N-1}$ at stages $k+1, ..., N-1$

Utility

- Decision makers are usually risk averse; the more concave the utility function is, the more risk averse the person
- The utility function is unique up to affine transormations; affine transformed utility function describes the decision maker *as good*:

$$U_2(\cdot) = k_1 + k_2 U_1(\cdot) \tag{18}$$

- Decision maker who maximizes U_1 acts exactly like decision maker who maximizes U_2 and vice versa

Utility

- The second derivative $U''(y)$ tells how concave the utility function is, but this factor does not determine the utility function uniquely and is not enough to measure behavior towards risk
- By normalizing the second derivative with the first derivative, we get *Arrow-Pratt measure of absolute risk-aversion* (ARA)

$$-\frac{U'(y)}{U''(y)} = f(y), \qquad \forall y \tag{19}$$

Utility

- HARA (*hyperbolic absolute risk aversion*):

$$f(y) = a + by, \qquad \forall y$$

  $\Rightarrow$ risky behavior increases with wealth and vice versa

- CARA (*constant absolute risk aversion*):

$$f(y) = \text{vakio} \qquad \forall y$$

  $\Rightarrow$ risk behavior is independent of wealth