4.1. Linear Systems and Quadratic Cost.

A special case of a linear system:

$$x_{k+1} = A_k x_k + B_k u_k + w_k \quad . \quad k=0,1,\cdots,N-1.$$

Quadratic cost:

$$\mathop{E}_{\substack{w_k \\ k=0,1,\cdots,N-1}} \left\{ x_N' Q_N x_N + \sum_{k=0}^{N-1} (x_k' Q_k x_k + u_k' R_k u_k) \right\}.$$

x_k, u_k: vectors of dimension n and m.

$A_k. B_k. Q_k. R_k$: matrix with appropriate dimension.

$R_k \in \mathbb{R}_+^m$. $Q_k \in \mathbb{R}_+^n$. positive semidefinite. symmetric.

u_k: unconstrainted. w_k: ind of $x_k. u_k$, $E w_k = 0$. $E w_k^T w_k < \infty$

Applying DP Algorithm:

$$J_N(x_N) = x_N' Q_N x_N.$$

$$J_k(x_k) = \min_{u_k} \mathop{E}_{w_k} \left\{ x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}(A_k x_k + B_k u_k + w_k) \right\}.$$

* Cost-to-go functions J_k are quadratic

* Optimal control law is a linear function of the state.

Proved by. backward. induction:

when $k = N-1$

$$J_{N-1}(x_{N-1}) = \min_{u_{N-1}} \mathop{E}_{w_{N-1}} \left\{ x_{N-1}' Q_{N-1} x_{N-1} + u_{N-1}' R_{N-1} u_{N-1} + (A_{N-1} x_{N-1} + B_{N-1} u_{N-1} + w_{N-1})' Q_N (A_{N-1} x_{N-1} + B_{N-1} u_{N-1} + w_{N-1}) \right\}$$

$$= x_{N-1}' [Q_{N-1} + A_{N-1}' Q_N A_{N-1}] x_{N-1}' + \min_{u_{N-1}} \left\{ u_{N-1}' [R_{N-1} + B_{N-1}' Q_N B_{N-1}] u_{N-1} + 2 x_{N-1}' A_{N-1}' Q_N B_{N-1} u_{N-1} + \mathop{E}_{w_{N-1}} [w_{N-1}' Q_N w_{N-1}] \right\}.$$

上式可以看成最小化 u_{N-1} 的一个二项式. 而二次项系数为:

$R_{N-1} + B_{N-1}' Q_N B_{N-1}$ symmetric positive definte if $R.\ Q$ positive definite.

Thus, it is a minimize problem of a convex function of u_{N-1},

$$\Rightarrow u_{N-1}^* = -[R_{N-1} + B_{N-1}' Q_N B_{N-1}]^{-1} B_{N-1}' Q_N A_{N-1} x_{N-1}$$

(Matrix inverse exists since symmetric positive definite).

u_{N-1}^* is a linear function of state x_{N-1}.

$$J_{N-1}(x_{N-1}) = x_{N-1}' K_{N-1} x_{N-1} + \underset{w_{N-1}}{E}[w_{N-1}' Q_N w_{N-1}]$$

其中. $K_{N-1} = A_{N-1}'[Q_N - Q_N B_{N-1}(B_{N-1}' Q_N B_{N-1} + R_{N-1})^{-1} B_{N-1}' Q_N] A_{N-1} + Q_{N-1}$

$\forall\ x.$ $x' K_{N-1} x = \min_u [\underbrace{x' Q_{N-1} x}_{\geq 0} + \underbrace{u' R_{N-1} u}_{\geq 0} + \underbrace{(A_{N-1} x + B_{N-1} u)' Q_N (A_{N-1} x + B_{N-1} u)}_{\geq 0}] \geq 0$

K_{N-1} is symmetric, positive semidefinite.

cost-to-go function $J_{N-1}(x_{N-1})$ is quadratic.

when k proof is similar to $N-1$. Thus, we can obtain

$\mu_k^*(x_k) = L_k x_k$.

$L_k = -(B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1} A_k$.

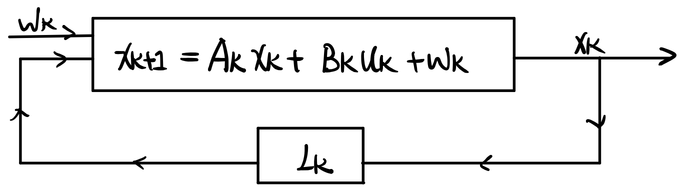
K_k is given recursively:

$K_N = Q_N$.

$K_k = A_k'[K_{k+1} - K_{k+1} B_k (B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1}] A_k + Q_k$. (1)

K_k : symmetric positive semidefinite.

$$\Rightarrow J_0(x_0) = x_0' K_0 x_0 + \sum_{k=0}^{N-1} E\{w_k' K_{k+1} w_k\}.$$



Automatic control by computation of L_k.

* The Riccati Equation and Its Asymptotic Behavior.

Eq (1) is called discrete time Riccati equation.

Require $A_k = A$. $B_k = B$. $Q_k = Q$. $R_k = R$. then K_k by Eq (1) converge to K as $k \to \infty$ (under mild conditions).

$$K = A'(K - KB(B'KB + R)^{-1} B'K)A + Q.$$

which implies that for system.

$$x_{k+1} = Ax_k + Bu_k + w_k. \quad , k = 0, \cdots, N-1.$$

when N is large enough. we can approximate optimal control law as:

$$\{\mu_0^*, \cdots, \mu_{N-1}^*\} = \{\mu^*, \mu^*, \cdots, \mu^*\}.$$

$$\mu^*(x) = Lx. \qquad L = -[B'KB + R]^{-1} B'KA.$$

4.2. Inventory Control.

Assumption: excessed demand at each period is backlogged and is fulfilled when additional inventory becomes available

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

x_k: on hand inventory at the beginning of period k.

u_k: order quantity in period k.

w_k: demand in period k. (value within some bounded interval and ind).

$r(x) = p \max(0, -x) + h \max(0, x)$: holding / shortage cost.

$p \geq 0$. $h \geq 0$

Hence, Total expected cost:

$E\{\sum_{k=0}^{N-1}(cu_k + r(x_k+u_k-w_k))\}$ $\quad$ c: purchase cost per unit.

r: convex function and $r(-\infty)=r(\infty)=\infty$.

Apply DP Algorithm:

$$J_N(x_N)=0$$

$$J_k(x_k)=\min_{u_k}[cu_k + H(x_k+u_k) + E\{J_{k+1}(x_k+u_k-w_k)\}]$$

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

H independ of k (if H depends on k. results still hold).

H is convex. since $r(y-w_k)$ is convex in y for each fixed w_k. Taking expectation over w_k keeps convexity.

Now. denote $y_k \triangleq x_k+u_k$. and rewrite DP:

$$J_k(x_k)=\min_{y_k\geq x_k} G_k(y_k) - cx_k.$$

where. $G_k(y)=cy + H(y) + E\{J_{k+1}(y-w_k)\}$.

$G_k(y)$ convex. denote $S_k \triangleq \arg\min_{y\in R} G_k(y)$.

$\Rightarrow$ if $x_k \leq S_k$, then $y_k=S_k$.

if $x_k \geq S_k$. then $y_k=x_k$.

$\Rightarrow$ minimum is attained at

$$\mu_k^*(x_k)=u_k^*\begin{cases} S_k-x_k & \text{if } x_k\leq S_k. \\ 0 & \text{otherwise.}\end{cases} \Rightarrow \text{optimal policy.}$$

(optimal when J_k is convex and $\lim_{N\to\infty} G_{N-1}(y)=\infty$)

proof. By Backward Induction:

$J_N = 0$ convex since $c < p$ and $\mathcal{H}'(y) \to p$ as $y \to -\infty$

$\mathcal{H}'(y) \to h$ as $y \to \infty$, $G(y) \to \infty$ as $|y| \to \infty$

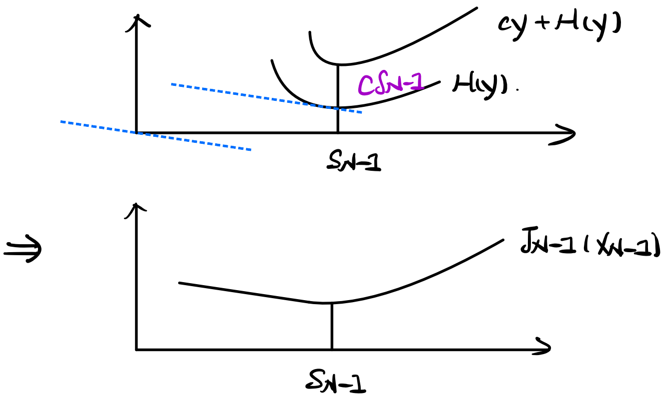
when $N-1$. optimal policy:

$$\mu^*_{N-1}(x_{N-1}) = \begin{cases} S_{N-1} - x_{N-1} & \text{if } x_{N-1} < S_{N-1} \\ 0 & \text{otherwise.} \end{cases}$$

By DP Algorithm:

$$J_{N-1}(x_{N-1}) = \begin{cases} c(S_{N-1} - x_{N-1}) + \mathcal{H}(S_{N-1}) & \text{if } x_{N-1} < S_{N-1} \\ \mathcal{H}(x_{N-1}). & \text{otherwise} \end{cases}$$

is convex. Since $\mathcal{H}$ is convex, S_{N-1} minimize $cy + H(y)$.



$$\lim_{|y| \to \infty} J_{N-1}(y) = \infty$$

we can repeat this argument to show for all $k = N-2, \cdots, 0$

If J_{k+1} is convex, $\lim_{|y| \to \infty} J_{k+1}(y) = \infty$, $\lim_{|y| \to \infty} G_k(y) = \infty$.

then $$J_k(x_k) = \begin{cases} c(S_k - x_k) + H(S_k) + E\, J_{k+1}(S_k - w_k) & x_k < S_k. \\ H(x_k) + E\, J_{k+1}(x_k - w_k) & x_k \geq S_k. \end{cases}$$

S_k is an unconstrained minimum of $G_k(y)$

J_k is convex, $\lim_{|y| \to \infty} J_k(y) = \infty$, $\lim_{|y| \to \infty} G_{k-1}(y) = \infty$

* Positive Fixed cost and (s, S) policies.

Assumption: a fixed cost K associated with a positive inventory cost.

cost of ordering u units: $C(u)$

$$C(u) = \begin{cases} K + cu & \text{if } u>0 \\ 0 & \text{if } u=0 \end{cases}$$

Apply DP Algorithm:

$$J_N(x_N) = 0$$

$$J_k(x_k) = \min_{u_k \geq 0} \left\{ C(u_k) + H(x_k+u_k) + E\left\{ J_{k+1}(x_k+u_k-w_k) \right\} \right\}.$$

H is defined as.

$$H(y) = E\, r(y-w) = p\, E\{\max(0, w-y)\} + h\, E\{\max(0, y-w)\}.$$

Denote $y_k = x_k + u_k$. $\quad G_k(y) = cy + H(y) + E\{J_{k+1}(y-w_k)\}.$

rewrite DP.

$$J_k(x_k) = \min\left[G_k(x_k), \min_{y_k > x_k} \left[K + G_k(y_k) \right] \right] - cx_k.$$

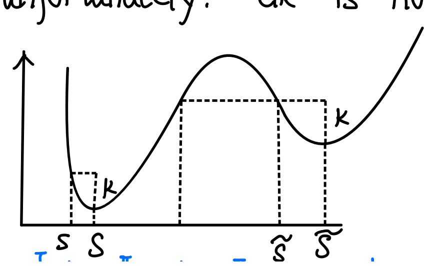
If $K=0$, return to the previous situation.

$$\mu_k^*(x_k) = \begin{cases} S_k - x_k & \text{if } x_k < S_k \\ 0 & \text{otherwise.} \end{cases}$$

$S_k = \arg\min G_k(y)$. $\quad s_k = \min\{ y \mid G_k(y) = K + G_k(S_k) \}$.

so-called. (s, S) policy.

If $K>0$, unfortunately. G_k is not convex, we may have.



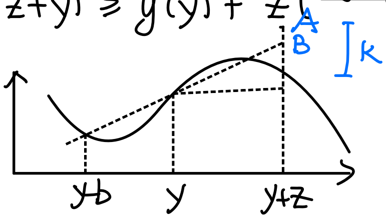
I: opt policy: $S - x_k$. II: opt policy: x_k.

III: opt policy: $\tilde{S} - x_k$. IV: opt policy: x_k.

Def: K-convex:

A Real-valued function g is K-convex, $K \geq 0$ if

$$K + g(z+y) \geq g(y) + z\left(\frac{g(y)-g(y-b)}{b}\right). \quad \forall\ z \geq 0,\ b > 0,\ y.$$



$A \geq B$

$K=0$, convex 定义。 K越大，容忍的振荡越大。
(越不 convex)

Lemma. 4.2.1.

a) A Real-valued convex function is also 0-convex and hence K-convex for all $K \geq 0$

b) $g_1(y)$ and $g_2(y)$ K-convex and L-convex ($K \geq 0$, $L \geq 0$)
$\alpha g_1(y) + \beta g_2(y)$ is $(\alpha K + \beta L)$-convex for all $\alpha > 0$, $\beta > 0$

c) If $g(y)$ is K-convex and w is a r.v., then $E_w\{g(y-w)\}$ is also K-convex. [$E_w\{g(y-w)\} < \infty$ for all y).

d) If g is a continuous K-convex function, $g(y) \to \infty$ as $|y| \to \infty$, there exist scalars s and S with $s < S$ such that:

(i) $g(S) \leq g(y)$ $\forall\ y$.

(ii) $g(S) + K = g(s) < g(y)$ $\forall\ y < s$

(iii) $g(y)$ is a decreasing function on $(-\infty, s)$

(iv) $g(y) \le g(z) + K$ for all y, z with $s \le y \le z$.

(lemma α 指出即使 G_K not convex、但不影响 (s, S) policy, 上图中 III 不会出现).

proof. Consider G_{N-1}.

$G_{N-1}(y) = c(y) + h(y) \Rightarrow$ convex $\Rightarrow$ K-convex.

$J_{N-1} = \min \{ G_{N-1}(x), \min\limits_{y \ge x} \{ K + G_{N-1}(y) \} \} - cx$

$$= \begin{cases} K + G_{N-1}(S_{N-1}) - cx & \text{for } x \le s_{N-1} \\ G_{N-1}(x) - cx & \text{for } x > s_{N-1} \end{cases}$$

$S_{N-1} = \arg\min\limits_{x} G_{N-1}(x)$.

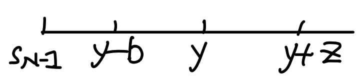
s_{N-1} : smallest value of y for which $G_{N-1}(y) = K + G_{N-1}(S_{N-1})$

$\forall$ $z \ge 0$, $b > 0$. we assert that

$$K + J_{N-1}(y+z) \ge J_{N-1}(y) + z \left(\frac{J_{N-1}(y) - J_{N-1}(y-b)}{b} \right) \quad \text{(Eq 2)}$$

Now we prove it by 3 cases:

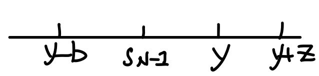
Case I. $y \ge s_{N-1}$



* $y-b \ge s_{N-1}$

$J_{N-1}(y-b)$, $J_{N-1}(y)$, $J_{N-1}(y+z)$ takes value from $G_{N-1}(x) - cx$, which is convex.

* $y-b \le s_{N-1}$



Eq (2) $\Leftrightarrow$ $K + G_{N-1}(y+z) - c(y+z) \ge G_{N-1}(y)$

$-c(y) + z \left(\frac{G_{N-1}(y) - cy - G_{N-1}(S_{N-1}) + c(y-b)}{b} \right)$

$\Leftrightarrow K + G_{N-1}(y+z) \ge G_{N-1}(y) + z \left(\frac{G_{N-1}(y) - G_{N-1}(S_{N-1})}{b} \right)$

If y is such that $G_{N-1}(y) \ge G_{N-1}(S_{N-1})$

then by k-convexity of G_{N-1}

$$k + G_{N-1}(y+z) \geq G_{N-1}(y) + z\left(\frac{G_{N-1}(y) - G_{N-1}(S_{N-1})}{y - S_{N-1}}\right) \quad [y - S_{N-1} < b]$$

$$\geq G_{N-1}(y) + z\left(\frac{G_{N-1}(y) - G_{N-1}(S_{N-1})}{b}\right)$$

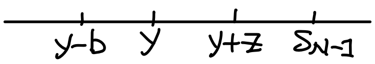
G convex.

If y is such that $G_{N-1}(y) < G_{N-1}(S_{N-1})$

$$\begin{aligned} k + G_{N-1}(y+z) &\geq k + G_{N-1}(S_{N-1}) \\ &= G_{N-1}(s_{N-1}) \\ &\geq G_{N-1}(y) \\ &\geq G_{N-1}(y) + z\left(\frac{G_{N-1}(y) - G_{N-1}(S_{N-1})}{b}\right) \end{aligned}$$

Hence, k-convexity Equation holds.

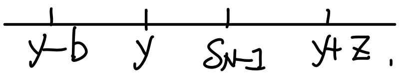
Case II. $y \leq y + z \leq S_{N-1}$



J_{N-1} is linear in x,

result is obvious.

Case III. $y < S_{N-1} \leq y + z$



Eq 2 $\Leftrightarrow$

$$k + G_{N-1}(y+z) - c(y+z) \geq G_{N-1}(S_{N-1}) - cy + z\left(\frac{G_{N-1}(S_{N-1}) - cy - G_{N-1}(S_{N-1}) + c(y-b)}{b}\right)$$

$$\Leftrightarrow \quad k + G_{N-1}(y+z) \geq G_{N-1}(S_{N-1}) \quad (= k + G_{N-1}(S_{N-1}))$$

the Equation holds by definition of S_{N-1}.

thus, by three cases, $G_{N-1}(x)$ is k-convex and $G_{N-1}(y) \to \infty$ as $|y| \to \infty$ implies $J_{N-1}(x)$ is k-convex.

Since J_{N-1} is continuous, Lemma 4.2.1 (c) $\Rightarrow$ $E\, J_{N-1}(y-w)$ is also k-convex $\Rightarrow$ $G_{N-2}(y) = cy + H(y) + E\, J_{N-1}(y-w)$ is k-convex

Since $G_{N-2}(y)$ is continuous, $G_{N-2}(y) \to \infty$ as $|y| \to \infty$

$\Rightarrow$ J_{N-2} is also K-convex.

By lemma 4.2.1 (d). we can get (s, S) policy is optimal

4. Optimal Stopping Problems.

* Asset selling.

Problem: * A person has an asset (such as land)

* $w_0, w_1, \cdots, w_{N-1}$ offers money to buy the asset.
* w_k takes nonnegative bounded value, $w_k = 0$ implies no offers in period k.
* If the person accepts the offer, then he invents the money at fixed rate $r > 0$
* If he rejects the offer, then he waits until next period to consider the next offer
* Rejected offers are not renewed, and last offer w_{N-1} must be accepted if every prior offer has been rejected

Objective: find a policy for accepting and rejecting offers that maximize the revenue of the person in the Nth. period.

DP:

State space: Real + termination state "T"

$x_k = T$, $k \le N-1$: the asset has already been sold

$x_k \neq T$, $k \le N-1$: the asset has not been sold and offer under consideration is x_k

We assume $x_0 = 0$

Control Space: two elements $u^1 = \text{sell}$, $u^2 = \text{do not sell}$

$u_k = u^1$: sell at time k and accepts offer w_{k-1}

w^k: disturbance at time k.

System Equation: $x_{k+1} = f_k(x_k, u_k, w_k)$.

$$= \begin{cases} T & \text{if } x_k = T \text{ or } x_k \neq T,\ u_k = u^1 \\ w_k & \text{otherwise.} \end{cases}$$

Reward Function:

$$\underset{\substack{w_k \\ k=0,1,\cdots,N-1}}{E} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\}.$$

其中: $g_N(x_N) = \begin{cases} x_N & \text{if } x_N \neq T \\ 0 & \text{if } x_N = T \end{cases}$

$$g_k(x_k, u_k, w_k) = \begin{cases} (1+r)^{N-k} x_k & \text{if } x_k \neq T \text{ and } u_k = u^1 \\ 0 & \text{otherwise} \end{cases}$$

DP Algorithm:

$$J_N(x_N) = \begin{cases} x_N, & \text{if } x_N \neq T \\ 0 & \text{if } x_N = T \end{cases}$$

$$J_k(x_k) = \begin{cases} \max\left\{ (1+r)^{N-k} x_k,\ E[J_{k+1}(w_k)] \right\}. & \text{if } x_k \neq T \\ 0 & \text{if } x_k = T \end{cases}$$

$(1+r)^{N-k} x_k$: revenue when sell at period k.

$E[J_{k+1}(w_k)]$: revenue not sell at period k.

Optimal Policy: denote $\alpha_k = \dfrac{E[J_{k+1}(w_k)]}{(1+r)^{N-k}}$

accept if $x_k \geq \alpha_k$.

reject if $x_k < \alpha_k$.

Properties of the optimal policy:

Assume w_k identically distributed : drop k, $w_k \triangleq w$ and denote by $E_w\{\cdot\}$.

Now we will show that $\alpha_k \geq \alpha_{k+1}$ for all k. and $\alpha_k \to \alpha$ as $k \to -\infty$

*proof : Introduce the functions:

$$V_N(x_N) = x_N.$$

$$V_k(x_k) = \frac{J_k(x_k)}{(1+r)^{N-k}}, \quad x_k \neq T$$

$$= \max\{x_k . (1+r)^{-1} E_w\{V_{k+1}(w)\}\}$$

$$\alpha_k = \frac{E[J_{k+1}(x_k)]}{(1+r)^{N-k}} = (1+r)^{-1} E_w[V_{k+1}(w)]$$

To prove $\alpha_k \geq \alpha_{k+1}$. ($V_{k+1} \leq V_k$). Induction.

$$V_{N-1}(x) \geq V_N(x), \quad \text{for all } x \geq 0$$

$$V_{N-2}(x) = \max\{x . (1+r)^{-1} E_w[V_{N-1}(w)]\}$$

$$\geq \max\{x . (1+r)^{-1} E_w[V_N(w)]\} . = V_{N-1}(x).$$

repeat : $V_k(x) \geq V_{k+1}(x)$. $\forall\ x, k$

$$\Rightarrow \alpha_k \geq \alpha_{k+1}$$

when N is large:

$$V_k(x_k) = \max\{x_k, \alpha_k\}.$$

$$\alpha_k = (1+r)^{-1} E_w[V_{k+1}(w)]$$

$$= \frac{1}{1+r} E[\max\{x_{k+1}, \alpha_{k+1}\}]$$

$$= \frac{1}{1+r}\left[\int_0^{\alpha_{k+1}} \alpha_{k+1}\, dP(w) + \int_{\alpha_{k+1}}^{\infty} w\, dP(w)\right]$$

$$= \frac{\alpha_{k+1}}{1+r} P(\alpha_{k+1}) + \frac{1}{1+r}\int_{\alpha_{k+1}}^{\infty} w\, dP(w)$$

where. $P(\alpha_{k+1}) = \text{Prob}(w \leq \alpha_{k+1})$, $\alpha_N = 0$

Because $0 \leq \frac{P(\alpha)}{1+r} \leq \frac{1}{1+r} \leq 1$.

$0 \leq \frac{1}{1+r} \int_{\alpha_{k+1}}^{\infty} w \, dP(w) \leq \frac{Ew}{1+r} \quad \forall k$.

$\alpha_k \geq \alpha_{k+1}$, $\alpha_k \geq 0$. $\alpha_k \to \alpha$ as $k \to -\infty$

$\Rightarrow$ $(1+r)\alpha = P(\alpha)\alpha + \int_{\alpha}^{\infty} w \, dP(w)$ $\quad k \to -\infty$

$\Rightarrow$ when N is Large. policy become steady.

accept if $x_k > \alpha$

reject if $x_k < \alpha$.

* Purchasing with a deadline.

Problem background: Buy raw material by a certain time and price flunctuates. We need to minimize purchase price.

Assumptions: * successive price $w_k \overset{iid}{\sim} P(w_k)$.

* purchase has to be made in N periods.

* denote $x_{k+1} = w_k$.

DP Algorithm:

$$J_N(x_N) = x_N = w_{N-1}$$

$$J_k(x_k) = \min\{x_k, \; E J_{k+1}(w_k)\}. \quad \text{(Eq 3)}$$

$J_k(x_k)$: optimal cost-to-go when the current price is x_k and the material has not been purchased.

Terminate state T: material has been bought, cost-to-go is zero.

Optimal Policy: purchase if $x_k < \alpha_k$.

not purchase if $x_k > \alpha_k$.

$$\alpha_k = E[J_{k+1}(w_k)]$$

$\alpha_1, \cdots, \alpha_{N-1}$ can be solved recursively.

(Eq3) $\Rightarrow$ $\alpha_k = E[J_{k+1}(w_k)]$

$$= \int_0^{\alpha_{k+1}} w_k \, dP(w_k) + \int_{\alpha_{k+1}}^{\infty} \alpha_{k+1} \, dP(w_k)$$

$$= \alpha_{k+1}(1 - P(\alpha_{k+1})) + \int_0^{\alpha_{k+1}} w \, dP(w)$$

terminal condition :

$$\alpha_{N-1} = \int_0^{\infty} w \, dP(w) = E[w]$$

Correlated Prices : $w_0, w_1, \cdots, w_{N-1}$

$w_k = x_{k+1}$, $k = 0, 1, \cdots, N-1$. $x_{k+1} = \lambda x_k + \xi_k$

λ scalar , $0 \le \lambda \le 1$, $\xi_0, \xi_1, \cdots, \xi_{N-1}$ i.i.d r.v. $\xi_i \ge 0$

DP Algorithm:

$J_N(x_N)$.

$J_k(x_k) = \min\{x_k . E[J_{k+1}(\lambda x_k + \xi_k)]\}$.

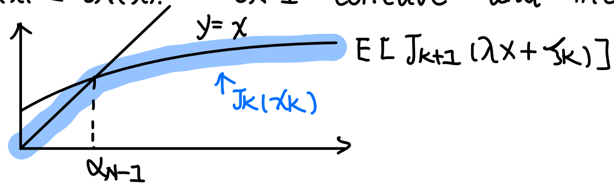
$J_{N-1}(x_{N-1}) = \min\{x_{N-1}, \lambda x_{N-1} + \bar{\xi}\}$. $\bar{\xi} = E[\xi_{N-1}]$

Optimal policy at time $N-1$:

* purchase if $x_{N-1} < \alpha_{N-1}$
* not purchase if $x_{N-1} > \alpha_{N-1}$

where $\alpha_{N-1} = \dfrac{\bar{\xi}}{1-\lambda}$ ($\alpha_{N-1} = \lambda \alpha_{N-1} + \bar{\xi}$, critical price)

$J_{N-1}(x) \le J_N(x)$. J_{N-1} concave and increasing in x



Using the fact in the DP algorithm, we can show

$J_k(x) \le J_{k+1}(x)$, $\forall x, k$, J_k is concave, increasing in x.

Furthermore, since $\bar{\xi} = E\{\xi_k\} > 0 \Rightarrow E\, J_{k+1}(\xi_k) > 0$

thus, optimal policy for period k:

* purchase if $x_k < \alpha_k$
* not purchase if $x_k > \alpha_k$

α_k is unquie solution of $x = E\, J_{k+1}(\lambda x + \xi_k)$

Because $J_k(x) \le J_{k+1}(x) \Rightarrow \alpha_{k-1} \le \alpha_k \le \alpha_{k+1}$

General stopping problems and one-step-lookahead rule

* Background

Termination state : T, go to T incur cost $t(x_k)$

last stage must go to T and incur cost $t(x_N)$

* DP Algorithm.

$J_N(x_N) = t(x_N)$. (Eq 4).

$J_k(x_k) = \min\left[t(x_k), \min\limits_{u_k \in U(x_k)} E\{g(x_k, u_k, w_k) + J_{k+1}(f(x_k, u_k, w_k))\}\right]$ (Eq 5)

it is optimal to stop at k for state x. is

$T_k = \{x \mid t(x) \le \min\limits_{u \in U(x)} E\{g(x,u,w) + J_{k+1}(f(x,u,w))\}\}$

(Eq 4) + (Eq 5) $\Rightarrow$ $J_{N-1}(x) \le J_N(x)$.

+ DP Algorithm $\Rightarrow$ $J_k(x) \le J_{k+1}(x)$.

By Exercise 1.23, we could obtain that

$T_0 \subset \cdots \subset T_k \subset T_{k+1} \subset \cdots \subset T_{N-1}$.

Consider a condition guaranteeing that all T_k are equal

suppose the set T_{N-1} is absorbing in the sense that if a state belongs to T_{N-1} and termination is not selected, the next state will also be in T_{N-1}.

$f(x, u, w) \in T_{N-1}$, for all $x \in T_{N-1}$, $u \in U(x)$, w.

If the equation holds $\Rightarrow T_N = T_{N-1}$.

$J_{N-1}(x) = t(x)$, $\forall x \in J_{N-1}$ by defintion of T_{N-1}

Denote $(*)$: $\min_{u \in U(x)} E\{g(x, u, w) + J_{N-1}(f(x, u, w))\}$.

$= \min_{u \in U(x)} E\{g(x, u, w) + t(f(x, u, w))\} \geq t(x)$

$T_{N-1} = \{x \mid t(x) \leq \min_{u \in U(x)} E\{g(x, u, w) + t(f(x, u, w))\}$.

therefore, stopping is optimal for all $x_{N-2} \in T_{N-1}$

$\Rightarrow T_{N-1} \subset T_{N-2} \Rightarrow T_{N-1} = T_{N-2}$.

Similarly, we have $T_k = T_{N-1}$. $\forall k$.

Example 4.4.1 (Asset selling with Past offers Retained)

asset selling problem + rejected offers can be accepted at a later time.

If the asset is not sold at time k,

$x_{k+1} = \max(x_k, w_k)$

DP Algorithm: $V_N(x_N) = x_N$

$V_k(x_k) = \max[x_k, (1+r)^{-1} E\{V_{k+1}(\max(x_k, w_k))\}]$

One-step stopping set

$T_{N-1} = \{x \mid x \geq (1+r)^{-1} E[\max(x, w)]\}$, an alternative characterization $T_{N-1} = \{x \mid x \geq \bar{\alpha}\}$. $\bar{\alpha}$ is obtained

$$(1+r)\bar{\alpha} = P(\bar{\alpha})\bar{\alpha} + \int_{\bar{\alpha}}^{\infty} w \, dP(w)$$

$\forall k$, if $x \in T_{N-1}$, then sell. otherwise wait optimal

verify T_{N-1} is absorbing:

$\forall \quad x \in T_{N-1}$, $x \geq \bar{\alpha}$, $\max(x, w) \geq x \geq \bar{\alpha}$

Example 4.4.2 (The Rational Burglar).

Background:

* the burglar might choose to retire at time k, then he gets his accumulated earning x_k
* Otherwise. he breaks into a house and gains w_k, but can be caught with probability p
* If he is caught, then he loses all his gains up to time k and has to terminate.
* The amounts w_k ind. $\bar{w} = E[w_k]$.
* Aim: maximize the burglar's expected earning over N nights.

Formulate as a DP.

* two actions: retire or continue.
* State space: a real line + retirement + caught.

* DP Algorithm:

$$J_N(x_N) = x_N.$$

$$J_k(x_k) = \max\left(x_k,\ (1-p)\,E[J_{k+1}(x_k+w_k)]\right).$$

* One-step stopping set:

$$T_{N-1} = \{x \mid x \geq (1-p)(x+\bar{w})\} = \{x \mid x \geq \frac{(1-p)\bar{w}}{p}\}.$$

$$x_k + w_k \geq x_k \geq \frac{(1-p)\bar{w}}{p}$$

If $x_k \in T_{N-1} \Rightarrow x_{k+1} \in T_{N-1} \Rightarrow T_{N-1}$ absorbing

* optimal policy : $x \in T_{N-1}$ stop (retire)

$x \notin T_{N-1}$, continuous