EE365: Code for Dynamic Programming

Example: Inventory model

- inventory level $x_t \in \{0, 1, \dots, C\}$
- new stock added $u_t \in \{0, 1, \dots, C\}$
- $x_{t+1} = x_t - w_t + u_t$
- demand $\mathbf{Prob}(w_t = 0, 1, 2) = (0.7, 0.2, 0.1)$

Example: Inventory model with ordering policy

- stage costs
  - fixed cost is o for ordering
  - sx for holding stock x
- add constraints $2 - x_t \leq u_t \leq C - x_t$ (so $x_{t+1} \in \{0, 1, \ldots, C\}$ for any d_t)
- otherwise stage cost is $g_t(x, u) = \begin{cases} sx + o & \text{if } u > 0 \\ sx & \text{otherwise} \end{cases}$
- final cost $g_T = 0$
- constants $C = 6$, $T = 50$, $x_0 = 6$, $s = 0.1$, $o = 1$

Data structures

problem data in the form of arrays

- dynamics $x_{t+1} = f(x_t, u_t, w_t)$ specified by `f` $\in \mathbb{R}^{n \times m \times p}$
- stage cost $g(x, u)$ specified by `g` $\in \mathbb{R}^{n \times m}$
- final cost $g_T(x)$ specified by `gfinal` $\in \mathbb{R}^{n}$
- distribution of w specified by `wdist` $\in \mathbb{R}^{p}$

Functions

- `value(f, g, gfinal, wdist, T)`

  returns `pol` $\in \mathbb{R}^{n\times T}$

  `v` $\in \mathbb{R}^{n\times(T+1)}$

- `cloop(f, g, pol)`

  returns `fcl` $\in \mathbb{R}^{n\times p\times T}$

  `gcl` $\in \mathbb{R}^{n\times T}$

- `ftop(fcl, wdist)` returns `P` $\in \mathbb{R}^{n\times n\times T}$

Computing the value function

```
value(f, g, gfinal, wdist, T)
```

given f $\in \mathbb{R}^{n\times m\times p}$, g $\in \mathbb{R}^{n\times m}$, gfinal $\in \mathbb{R}^{n}$, wdist $\in \mathbb{R}^{p}$, T $\in \mathbb{N}$

let $v_T^\star(x) = g_T(x)$

for $t = T-1, \ldots, 0$

find optimal policy for time t in terms of $v_{t+1}^\star$:

$$\mu_t^\star(x) \in \operatorname*{argmin}_u \left(g(x,u) + \mathbf{E}\, v_{t+1}^\star(f(x,u,w_t))\right)$$

find $v_t^\star$ using $\mu_t^\star$:

$$v_t^\star(x) = \min_u \left(g(x,u) + \mathbf{E}\, v_{t+1}^\star(f(x,u,w_t))\right)$$

return $\mu^\star =$ pol $\in \mathbb{R}^{n\times T}$ and v $\in \mathbb{R}^{n\times(T+1)}$

Computing the closed-loop dynamics

`cloop(f, g, pol)`

given `f` $\in \mathbb{R}^{n \times m \times p}$, `g` $\in \mathbb{R}^{n \times m}$, `pol` $\in \mathbb{R}^{n \times T}$

$f_t^{\mathsf{cl}}(x, w) = f(x, \mu_t(x), w)$

$g_t^{\mathsf{cl}}(x) = g(x, \mu_t(x))$

return `fcl` $\in \mathbb{R}^{n \times p \times T}$ and `gcl` $\in \mathbb{R}^{n \times T}$

- computes the closed-loop dynamics and cost
- time-varying policy results in time-varying closed-loop system

Computing the transition matrix

```
ftop(fcl, wdist)
```

given `fcl` $\in \mathbb{R}^{n \times p \times T}$ and `wdist` $\in \mathbb{R}^{p}$

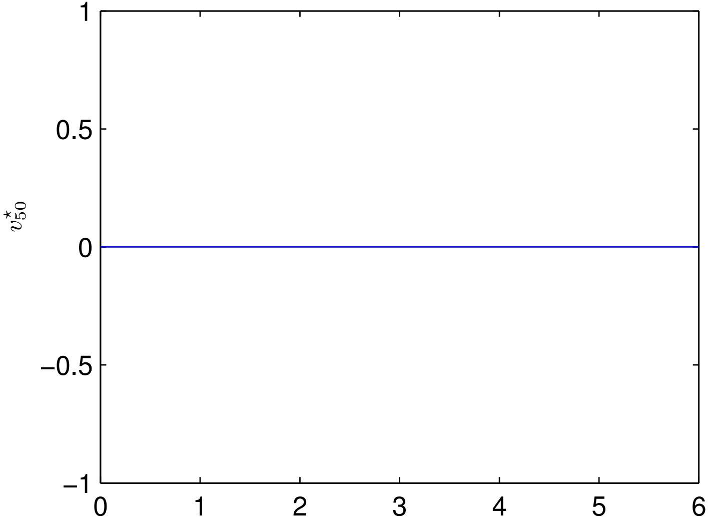
$(P_t)_{ij} = \sum \{\mathbf{Prob}(w) \mid w \in \mathcal{W} \text{ and } f_t^{\mathsf{cl}}(i, w) = j\}$

return `P` $\in \mathbb{R}^{n \times n \times T}$

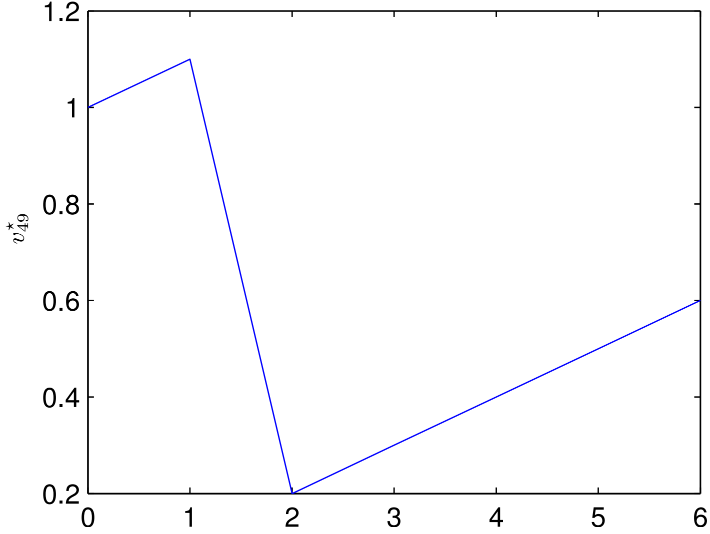
- given the system $x_{t+1} = f_t^{\mathsf{cl}}(x_t)$
- computes the time-varying transition matrix P_t such that

$$(P_t)_{ij} = \mathbf{Prob}(x_{t+1} = j \mid x_t = i)$$

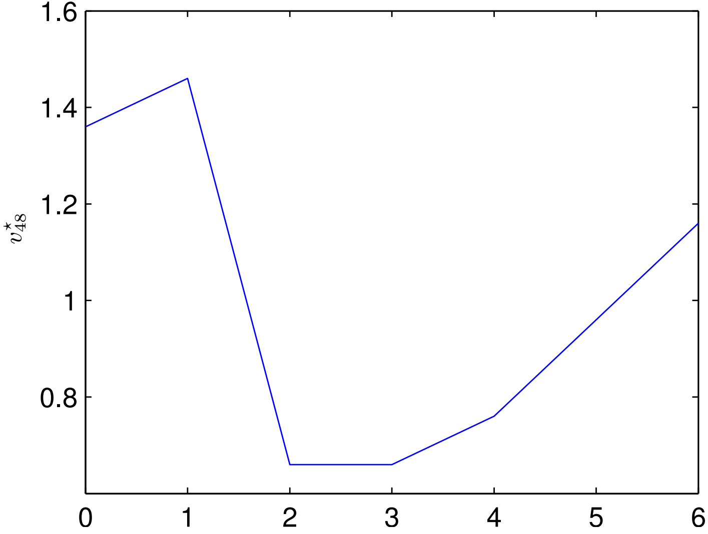
Example: Inventory model



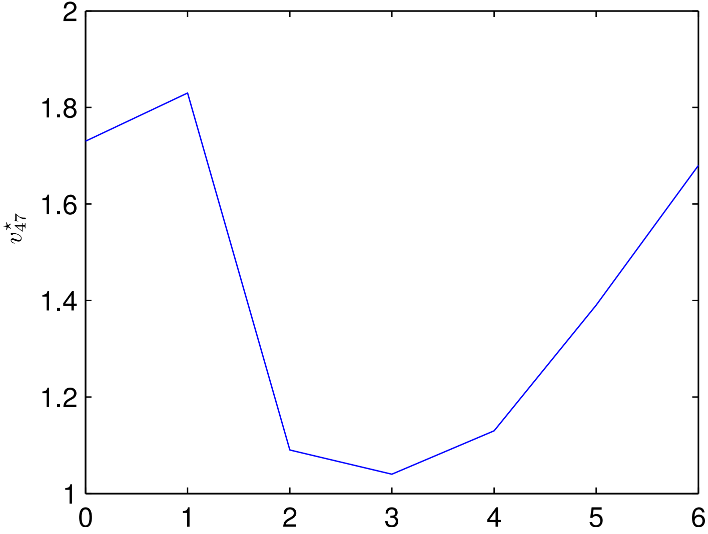
Example: Inventory model



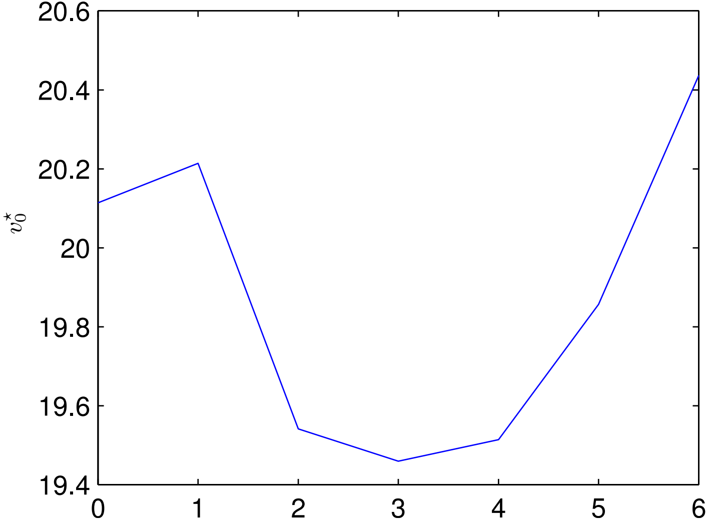
Example: Inventory model



Example: Inventory model



Example: Inventory model



Example: Inventory model

- optimal policy vs. heuristic policy

$$\mu^*(x) = \begin{cases} 4 - x & \text{if } x = 0 \text{ or } 1 \\ 0 & \text{otherwise} \end{cases} \qquad \mu^{\text{heur}}(x) = \begin{cases} 6 - x & \text{if } x = 0 \text{ or } 1 \\ 0 & \text{otherwise} \end{cases}$$

- expected total costs: $J^\star = 20.44$, $J^{\text{heur}} = 23.13$
- heuristic policy over-orders!