EE365: More Information Patterns

Measuring w and x

Measuring x and part of w

Measuring w and x

DP for modified information pattern

- suppose w_t is known (as well as x_t) before u_t is chosen
- typical applications: action is chosen *after* current (random) price, cost, demand, congestion is revealed
- policy has form $u_t = \mu_t(x_t, w_t)$, $\mu_t : \mathcal{X}_t \times \mathcal{W}_t \to \mathcal{U}_t$
- can map this into our standard form, but it's more natural to modify DP to handle it directly

Optimal value function when w_t is known

- define

$$v_t^\star(x) = \min_{\mu_t,\mu_{t+1},\ldots,\mu_{T-1}} \mathbf{E}\left(\sum_{\tau=t}^{T-1} g_\tau(x_\tau,u_\tau,w_\tau) + g_T(x_T)\,\middle|\, x_t = x\right)$$

  - minimization is over policies $\mu_t,\ldots,\mu_{T-1}$, functions of x *and* w
  - subject to dynamics $x_{t+1} = f_t(x_t,u_t,w_t)$
- $v_t^\star(x)$ is expected cost-to-go, using an optimal policy, if you are in state x at time t, *before w_t is revealed*

Dynamic programming for w_t known

- define $v_T^\star(x) := g_T(x)$
- for $t = T-1, \ldots, 0$,
  - find optimal policy for time t in terms of $v_{t+1}^\star$:

    $$\mu_t^\star(x, w) \in \operatorname*{argmin}_u \left(g_t(x,u,w) + v_{t+1}^\star(f_t(x,u,w))\right)$$

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

    $$v_t^\star(x) := \mathbf{E}\left(g_t(x, \mu_t^\star(x, w_t), w_t) + v_{t+1}^\star(f_t(x, \mu_t^\star(x, w_t), w_t))\right)$$

    (expectation is over w_t)
- only need to store a value function on $\mathcal{X}_t$, even though policy is a function on $\mathcal{X}_t \times \mathcal{W}_t$

Measuring x and part of w

DP for modified information pattern II

- suppose $w_t = (w_t^1, w_t^2)$ splits into independent components
- w_t^1 is known (as well as x_t) before u_t is chosen
- w_t^2 is not known before u_t is chosen
- policy has form $u_t = \mu_t(x_t, w_t^1)$, $\mu_t : \mathcal{X}_t \times \mathcal{W}_t^1 \to \mathcal{U}_t$
- can map this into our standard form, but it's more natural to modify DP to handle it directly

Optimal value function when w_t^1 is known

- define

$$v_t^\star(x) = \min_{\mu_t,\mu_{t+1},\ldots,\mu_{T-1}} \mathbf{E}\left(\sum_{\tau=t}^{T-1} g_\tau(x_\tau, u_\tau, w_\tau) + g_T(x_T) \middle| x_t = x \right)$$

  - minimization is over policies $\mu_t, \ldots, \mu_{T-1}$, functions of x *and* w^1
  - subject to dynamics $x_{t+1} = f_t(x_t, u_t, w_t)$
- $v_t^\star(x)$ is expected cost-to-go, using an optimal policy, if you are in state x at time t, *before* w_t^1 *is revealed*

Dynamic programming for w_t^1 known

- define $v_T^\star(x) := g_T(x)$
- for $t = T-1, \ldots, 0$,
  - find optimal policy for time t in terms of $v_{t+1}^\star$:

    $$\mu_t^\star(x, w^1) \in \operatorname*{argmin}_u \mathbf{E}\left(g_t(x, u, (w^1, w_t^2)) + v_{t+1}^\star(f_t(x, u, (w^1, w_t^2)))\right)$$

    (expectation is over w_t^2)
  - find $v_t^\star$ using $\mu_t^\star$:

    $$v_t^\star(x) := \mathbf{E}\left(g_t(x, \mu_t^\star(x, w_t^1), w_t) + v_{t+1}^\star(f_t(x, \mu_t^\star(x, w_t^1), w_t))\right)$$

    (expectation is over w_t)
- only need to store a value function on $\mathcal{X}_t$, even though policy is a function on $\mathcal{X}_t \times \mathcal{W}_t^1$