Dynamic Programming Lecture #11

Outline:

- Imperfect Information Set-up
- DP on Probabilities
- Example: Machine repair

Imperfect Information

- Standard problem:

$$x_{k+1} = f_k(x_k, u_k, w_k)$$
$$\min_{\pi} E\left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(x_k), w_k) \right\}$$

- What if we do not have access to x_k?
- Initial condition: x_0 random.
- Observations:

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

- Measurement disturbance (noise) v_k is random, but can depend on $(x_{k-1}, u_{k-1}, w_{k-1})$
- Implication: Cannot implement $\mu_k(x_k)$ because we never know x_k.

Examples

- Motivation #1: Partial measurements

$$y = Cx_k + v_k$$

  Examples: Position vs Velocity measurements or distributed sensors

- Motivation #2: Artificial states
- Revisit LQ:

$$x^+ = Ax + Bu + Lw$$

- Disturbance model: "white" vs "colored" (correlated with past).
- Represent w_k as

$$w_k = H\xi_k$$
$$\xi_{k+1} = F\xi_k + G\tilde{w}_k$$

  where $\tilde{w}_k$ independent.

- New dynamics:

$$\begin{pmatrix} x \\ \xi \end{pmatrix}^+ = \begin{pmatrix} A & LH \\ 0 & F \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} + \begin{pmatrix} B \\ 0 \end{pmatrix} u + \begin{pmatrix} 0 \\ G \end{pmatrix} w$$

- Cannot measure ξ...it is part of (artificial) model.

Information Formulation

- Information at time k:

$$I_0 = z_0$$

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

- Information based policy:

$$\mu_k(I_k) \quad \text{(not function of } x_k\text{)}$$

- Minimization:

$$E_{x_0, w_{\{0,\ldots,N-1\}}, v_{\{0,\ldots,N-1\}}} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(I_k), w_k) \right\}$$

- FACT: Can formulate DP on an information based system. We will pursue instead a "sufficient statistic".

Probability Propagation

- Define X_k: distribution function for x_k given information I_k.
- Interpretation: "Beliefs" on x_k. Instead of value of x_k, we have probabilities on values of x_k.
- FACT: For appropriately defined Φ_k,

$$X_{k+1} = \Phi_k(X_k, u_k, z_{k+1})$$

- Tomorrow's beliefs are a function of:
  - Today's beliefs
  - Today's control
  - Tomorrow'smeasurement
- All information encoded in X...i.e., a new "state"
- Recall previous discussion on (uncontrolled) hidden Markov models

Computation of Φ

- We have reduced partial information to full state information, but on system:

$$X_{k+1} = \Phi_k(X_k, u_k, z_{k+1})$$

  - New state: X_k
  - Control: u_k
  - New disturbance: z_{k+1}, i.e., tomorrow's measurement

- Assume "given" X_k and u_k throughout
- How to derive Φ_k?

$$\begin{aligned} P(x_{k+1} = s^*|z_{k+1} = z^*) &= \frac{P(x_{k+1} = s^* \ \& \ z_{k+1} = z^*)}{P(z_{k+1} = z^*)} \\ &= \frac{P(x_{k+1} = s^* \ \& \ z_{k+1} = z^*)}{\sum_j P(z_{k+1} = z^* \ \& \ x_{k+1} = j)} \end{aligned}$$

- By total probability:

$$P(x_{k+1} = s^*|z_{k+1} = z^*) = \frac{\sum_i P(x_{k+1} = s^* \ \& \ z_{k+1} = z^*|x_i = i)P(x_k = i)}{\sum_i P(z_{k+1} = z^*|x_{k+1} = j)P(x_k = i)}$$

- Allows for simpler computations

Example: Propagation of X

$$x_{k+1} = x_k + 1 + w_k, \quad w_k \in \{-1, 0, 1\}$$
$$z_{k+1} = x_{k+1} + v_k, \quad v_k \in \{-1, 0, 1\}$$

- Suppose we know $x_k = 0$, i.e.,

$$P(x_k = 0) = 1 \simeq X_k$$

- Main issue: What is X_{k+1} (given $I_{k+1} = I_k \cup z_{k+1}$)?

$$P(x_{k+1} = x^*|z_{k+1} = 1) = \frac{P(x_{k+1} = x^* \ \& \ z_{k+1} = 1)}{P(z_{k+1} = 1)}$$

- All possibilities:

	$w_k = -1$	0	1
$v_{k+1} = -1$	$z_{k+1} = -1$	0	1
0	0	1	2
1	1	2	3

	$w_k = -1$	0	1
$v_{k+1} = -1$	$x_{k+1} = 0$	1	2
0	0	1	2
1	0	1	2

- Result: $X_{k+1} = \{0, 1, 2\}$ with probability $\{1/3, 1/3, 1/3\}$

Example, cont (2)

	$w_k = -1$	0	1
$v_{k+1} = -1$	$z_{k+1} = -1$	0	1
0	0	1	2
1	1	2	3

	$w_k = -1$	0	1
$v_{k+1} = -1$	$x_{k+1} = 0$	1	2
0	0	1	2
1	0	1	2

- What is $P(x_{k+1} = x^*|z_{k+1} = 2)$?
- Result: $X_{k+1} = \{0, 1, 2\}$ with probability $\{0, 1/2, 1/2\}$
- What is $P(x_{k+1} = x^*|z_{k+1} = 3)$?
- Result: $X_{k+1} = \{0, 1, 2\}$ with probability $\{0, 0, 1\}$

DP on X

- Define

$$G_N(X_N) = E\left\{g_N(x_N)|I_N\right\}$$

i.e., expected terminal penalty given all information to date.

- Define

$$G_k(X_k, u_k) = E\left\{g_k(x_k, u_k, w_k)\right\}$$

- System dynamics:

$$X_{k+1} = \Phi_k(X_k, u_k, z_{k+1})$$

where probabilities of z_{k+1} depend on (X_k, u_k)

- Initialize

$$J_N(X_N) = G_N(X_N)$$

- Proceed as usual...

$$J_k(X_k) = \min_{u_k} E\left\{G_k(X_k, u_k) + J_{k+1}(\Phi_k(X_k, u_k, z_{k+1}))\right\}$$

- "Looks like" standard DP...but acting on belief dynamics.

Machine Repair

- Machine state: $= \begin{cases} 1 & \text{up} \\ 0 & \text{down} \end{cases}$.
- Control actions: $= \begin{cases} C & \text{continue (do nothing)} \\ S & \text{stop \& repair if needed} \end{cases}$
- System probabilities: If up, stay up with probability $2/3$ & fail with probability $1/3$.
- Measurement: Correct diagnosis with probability $3/4$ & false diagnosis with probability $1/4$
- Stage costs:

$$g(x,u) = \begin{cases} 2 & \text{continue with machine down} \\ 0 & \text{continue with machine up} \\ 1 & \text{stop \& repair} \end{cases}$$

- Let $p_k =$ probability machine is down at time k given information up to time k. (Then $(1 - p_k)$ is probability machine is up at time k.)
- What is p_1 as a function of p_0, u_0, & z_1? i.e., derive

$$X_1 = \Phi(X_0, u_0, z_1)$$

Machine Repair, cont (2)

- In each case, compute

$$\frac{P(\text{down } \ \& \ u_0, z_1)}{P(u_0, z_1)}$$

- $u_0 = S \ \& \ z_1 = U$:

$$\begin{cases} (1/3)(1/4) & \text{(fail)(false)} \\ (2/3)(3/4) & \text{(no fail)(true)} \end{cases} \Rightarrow \frac{1/12}{1/12 + 6/12} = 1/7$$

- $u_0 = S \ \& \ z_1 = D$:

$$\begin{cases} (1/3)(3/4) & \text{(fail)(true)} \\ (2/3)(1/4) & \text{(no fail)(false)} \end{cases} \Rightarrow \frac{1/4}{1/4 + 2/12} = 3/5$$

- $u_0 = C \ \& \ z_1 = U$:

$$\begin{cases} (1 - p_0)(1/3)(1/4) & \text{(was up)(fail)(false)} \\ p_0(1/4) & \text{(was down)(false)} \\ (1 - p_0)(2/3)(3/4) & \text{(was up)(no fail)(true)} \end{cases} \Rightarrow \frac{1 + 2p_0}{7 - 4p_0} = \frac{1 + 2}{1 + 2 + 3}$$

- $u_0 = C \ \& \ z_1 = D$:

$$\begin{cases} (1 - p_0)(1/3)(3/4) & \text{(was up)(fail)(true)} \\ (1 - p_0)(2/3)(1/4) & \text{(was up)(no fail)(false)} \\ p_0(3/4) & \text{(was down)(true)} \end{cases} \Rightarrow \frac{3 + 6p_0}{5 + 4p_0} = \frac{1 + 3}{1 + 2 + 3}$$

- In general: $p^+ = \Phi(p, u, z^+)$