6.231 DYNAMIC PROGRAMMING

LECTURE 13

LECTURE OUTLINE

- Suboptimal control
- Certainty equivalent control
- Implementations and approximations
- Issues in adaptive control

PRACTICAL DIFFICULTIES OF DP

- The curse of modeling
- The curse of dimensionality
  - Exponential growth of the computational and storage requirements as the number of state variables and control variables increases
  - Quick explosion of the number of states in combinatorial problems
  - Intractability of imperfect state information problems
- There may be real-time solution constraints
  - A family of problems may be addressed. The data of the problem to be solved is given with little advance notice
  - The problem data may change as the system is controlled – need for on-line replanning

CERTAINTY EQUIVALENT CONTROL (CEC)

- Replace the stochastic problem with a deterministic problem
- At each time k, the uncertain quantities are fixed at some "typical" values
- Implementation for an imperfect info problem. At each time k:

(1) Compute a state estimate $\overline{x}_k(I_k)$ given the current information vector I_k.

(2) Fix the w_i, $i \geq k$, at some $\overline{w}_i(x_i, u_i)$. Solve the deterministic problem:

$$\text{minimize} \;\; g_N(x_N) + \sum_{i=k}^{N-1} g_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big)$$

subject to $x_k = \overline{x}_k(I_k)$ and for $i \geq k$,

$$u_i \in U_i, \quad x_{i+1} = f_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big).$$

(3) Use as control the first element in the optimal control sequence found.

ALTERNATIVE IMPLEMENTATION

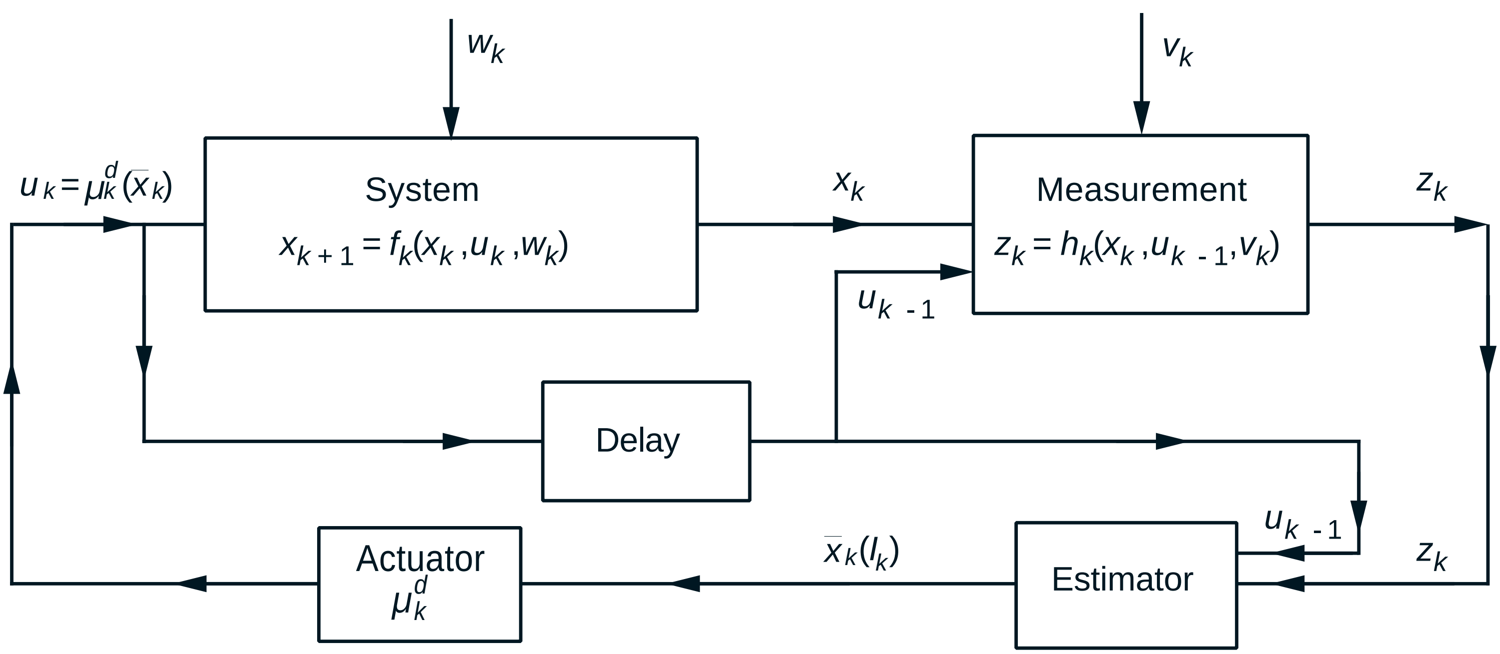
- Let $\{\mu_0^d(x_0), \ldots, \mu_{N-1}^d(x_{N-1})\}$ be an optimal controller obtained from the DP algorithm for the deterministic problem

$$\text{minimize} \quad g_N(x_N) + \sum_{k=0}^{N-1} g_k\big(x_k, \mu_k(x_k), \overline{w}_k(x_k, u_k)\big)$$

$$\text{subject to} \quad x_{k+1} = f_k\big(x_k, \mu_k(x_k), \overline{w}_k(x_k, u_k)\big), \qquad \mu_k(x_k) \in U_k$$

The CEC applies at time k the control input

$$\tilde{\mu}_k(I_k) = \mu_k^d\big(\overline{x}_k(I_k)\big)$$



CEC WITH HEURISTICS

- Solve the "deterministic equivalent" problem using a heuristic/suboptimal policy
- Improved version of this idea: At time k minimize the stage k cost and plus the heuristic cost of the remaining stages, i.e., apply at time k a control $\tilde{u}_k$ that minimizes over $u_k \in U_k(x_k)$

$$g_k\big(x_k, u_k, \overline{w}_k(x_k, u_k)\big) + H_{k+1}\Big(f_k\big(x_k, u_k, \overline{w}_k(x_k, u_k)\big)\Big)$$

where H_{k+1} is the cost-to-go function corresponding to the heuristic.

- This an example of an important suboptimal control idea:

Minimize at each stage k the sum of approximations to the current stage cost and the optimal cost-to-go.

- This is a central idea in several other suboptimal control schemes, such as limited lookahead, and rollout algorithms.

PARTIALLY STOCHASTIC CEC

- Instead of fixing all future disturbances to their typical values, fix only some, and treat the rest as stochastic.

- Important special case: Treat an imperfect state information problem as one of perfect state information, using an estimate $\overline{x}_k(I_k)$ of x_k as if it were exact.

- Multiaccess Communication Example: Consider controlling the slotted Aloha system (discussed in Ch. 5) by optimally choosing the probability of transmission of wating packets. This is a hard problem of imperfect state info, whose perfect state info version is easy.

- Natural partially stochastic CEC:

$$\tilde{\mu}_k(I_k) = \min\left[1, \frac{1}{\overline{x}_k(I_k)}\right],$$

where $\overline{x}_k(I_k)$ is an estimate of the current packet backlog based on the entire past channel history of successes, idles, and collisions (which is I_k).

SYSTEMS WITH UNKNOWN PARAMETERS

- Let the system be of the form

$$x_{k+1} = f_k(x_k, \theta, u_k, w_k),$$

where θ is a vector of unknown parameters with a given a priori probability distribution.

- To formulate this into the standard framework, introduce a state variable $y_k = \theta$ and the system

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} f_k(x_k, y_k, u_k, w_k) \\ y_k \end{pmatrix},$$

and view $\tilde{x}_k = (x_k, y_k)$ as the new state.

- Since $y_k = \theta$ is unobservable, we have a problem of imperfect state information even if the controller knows the state x_k exactly.

- Consider a partially stochastic CEC. If for a fixed parameter vector θ, we can compute the corresponding optimal policy $\{\mu_0^*(I_0, \theta), \ldots, \mu_{N-1}^*(I_{N-1}, \theta)\}$ a partially stochastic CEC applies $\mu_k^*(I_k, \hat{\theta}_k)$, where $\hat{\theta}_k$ is some estimate of θ.

THE PROBLEM OF IDENTIFIABILITY

- Suppose we consider two phases:
  - A parameter identification phase (compute an estimate $\hat{\theta}$ of θ)
  - A control phase (apply control that would be optimal if $\hat{\theta}$ were true).

- A fundamental difficulty: the control process may make some of the unknown parameters invisible to the identification process.

- Example: Consider the scalar system

$$x_{k+1} = ax_k + bu_k + w_k, \qquad k = 0, 1, \ldots, N-1,$$

with the cost $E\left\{\sum_{k=1}^{N}(x_k)^2\right\}$. If a and b are known, the optimal control law is $\mu_k^*(x_k) = -(a/b)x_k$.

- If a and b are not known and we try to estimate them while applying some nominal control law $\tilde{\mu}_k(x_k) = \gamma x_k$, the closed-loop system is

$$x_{k+1} = (a + b\gamma)x_k + w_k,$$

so identification can at best find $(a + b\gamma)$ but not the values of both a and b.

CEC AND IDENTIFIABILITY I

- Suppose we have $P\{x_{k+1} \mid x_k, u_k, \theta\}$ and we use a control law μ^* that is optimal for known θ:

$$\hat{\mu}_k(I_k) = \mu^*(x_k, \hat{\theta}_k), \qquad \text{with } \hat{\theta}_k\text{: estimate of } \theta$$

There are three systems of interest:

(a) The system (perhaps falsely) believed by the controller to be true, which evolves probabilistically according to

$$P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}_k), \hat{\theta}_k\big\}.$$

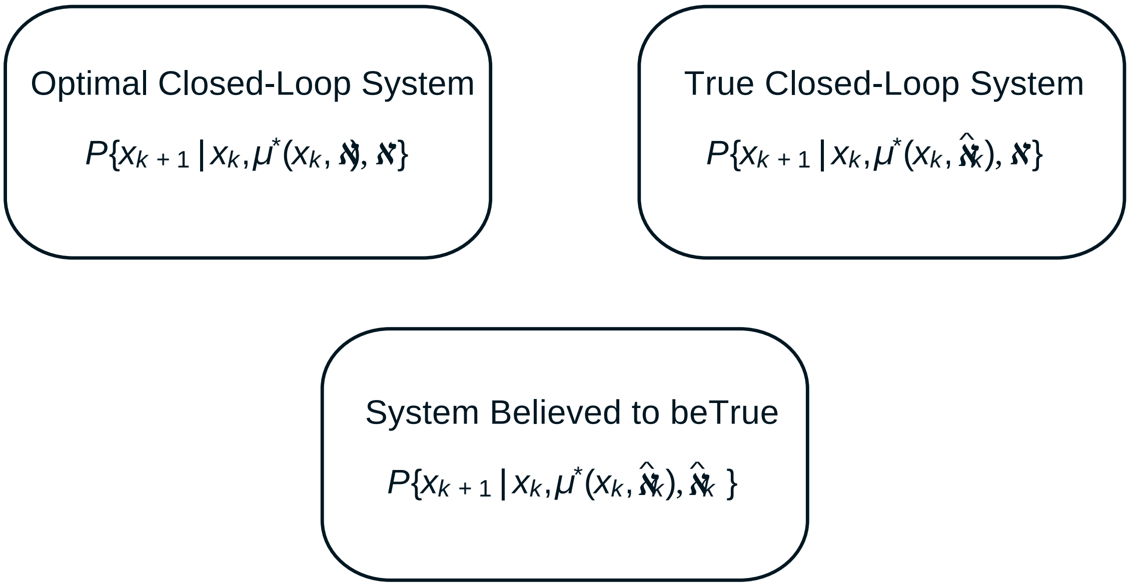
(b) The true closed-loop system, which evolves probabilistically according to

$$P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}_k), \theta\big\}.$$

(c) The optimal closed-loop system that corresponds to the true value of the parameter, which evolves probabilistically according to

$$P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \theta), \theta\big\}.$$

CEC AND IDENTIFIABILITY II



- There is a built-in mechanism for the parameter estimates to converge to a wrong value
- Assume that for some $\hat{\theta} \neq \theta$ and all x_{k+1}, x_k,

$$P\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}), \hat{\theta}\} = P\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}), \theta\}$$

i.e., there is a false value of parameter for which the system under closed-loop control looks exactly as if the false value were true.

- Then, if the controller estimates at some time the parameter to be $\hat{\theta}$, subsequent data will tend to reinforce this erroneous estimate.