Lecture Summary: Information Patterns

Vijay Gupta[*] and Richard M. Murray [†]

Motivation: The complexity of being able to obtain the solution of an optimal control, or even stabilization, problem depends crucially on the information available to the controller at every time step. Even the classical LQG problem may turn extremely difficult if the implicit assumption about the controller having access to all previous control inputs and measurements is not valid. An example is the Witsenhausen counterexample framework shown in Figure 1. If at step 2, the controller has access to either the measurement or the control

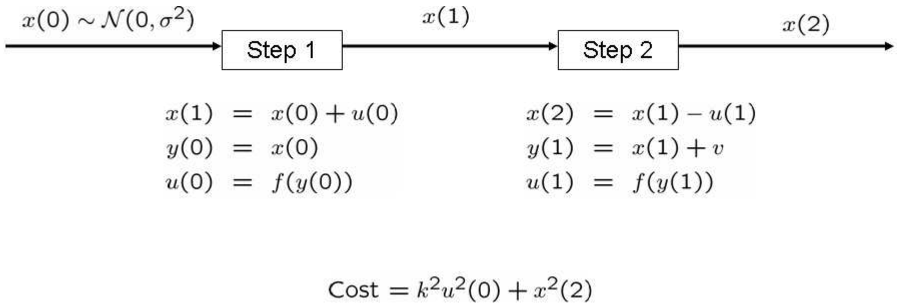

Figure 1:

at time step 1, this is an LQG problem. However, with the given information availability, the optimal controller and cost are unknown. In particular, linear controllers are no longer optimal unlike LQG. This issue is even more important in distributed control problems, where different controllers may have access to different information sets at any time.

General Framework: Consider a distributed control system with state vector $x(.)$ that possibly consists of states of several sub-systems stacked up, M sensors, and K control systems with specified information availability at every time step. The operation is in the following sequence:

- A random initial state $x(0)$ is given.
- The M sensors make measurements, $\{y^i(1)\} = \{g^i(1, x(0), w^i(1))\}$, where $w^i(1)$ is some noise component.
- The controls are calculated and implemented, $\{u^i(1)\} = \{\gamma^i(1, y^{i_1}(1), y^{i_2}(1), \cdots, y^{i_n}(1))\}$.
- The state transitions $x(1) = f(1, x(0), u^1(1), u^2(1), \cdots, u^K(1), v(1))$, where $v(1)$ is a noise component.
- The above steps are repeated $N-1$ times.

The cost function is generally of the form $\sum_{k=1}^{N} h(k, x(k), u^1(k), u^2(k), \cdots, u^K(k))$. The designer aims to pick design γ that specifies the (possibly time-varying) control laws γ^i for each i to minimize the expected cost. Constraints can be specified through feasible functional forms and information patterns.

Information Pattern: Consider the sets of pairs of indices

$$Y(k) = \{(\tau, m) | \tau = 1, \cdots, k, m = 1, \cdots, M\} \qquad k = 1, 2, \cdots, N$$
$$U(k) = \{(\tau, m) | \tau = 1, \cdots, k-1, m = 1, \cdots, K\} \qquad k = 2, \cdots, N+1.$$

[*]Department of Electrical Engineering, University of Notre Dame, vgupta2@nd.edu

[†]Control and Dynamical Systems, California Institute of Technology murray@caltech.edu

Define a data basis at time k as a pair (A,B), where $A \subset Y(k)$, and $B \subset U(k)$. An information pattern is the assignment of a data basis at time k, denoted by $(Y_m(k), U_m(k))$, to each index pair (k,m) in $U(N+1)$. Some examples are

Perfect Recall: If either $N = 1$, or for all controllers m, $Y_m(k) \subset Y_m(k+1)$ and $U_m(k) \subset U_m(k+1)$, $\forall k = 1, \cdots, N-1$.

Classical: If information pattern is independent of subscript, and information pattern is perfect recall. If there is only one controller, then the pattern is strictly classical.

Delayed Sharing Pattern: All K controllers have an associated sensor. They receive their own data immediately, and share it with all other controllers with a delay of n steps. Moreover, they have perfect recall.

LQG problem: The particular structure, and simplicity, of the classical LQG solution can be explained through the following four assertions.

- Any problem with a classical information pattern is equivalent to a problem with strictly classical information pattern.
- For a problem with strictly classical information pattern, there exists a conditional distribution of $x(k-1)$ given the variables $(y(1), \cdots, y(k), u(1), \cdots, u(k-1))$, independent of the design γ of the controller. Further, if $F(k)$ be this distribution, then there is no loss of performance if the controller is restricted to be of the form $\gamma(k) = \phi(k, F(k))$.
- For a strictly classical linear Gaussian system, $F(k)$ is Gaussian with covariance independent of data and mean affine in data.
- For a strictly classical linear system with quadratic cost, obtain the optimal control law $\phi(k, x(k-1))$ by assuming that the noise random variables are set at their mean values, and the initial distribution of state $x(0)$ is a point mass as mean of its distribution. Then, the optimal control law is given by $\phi(k, \bar{F}(k))$ where $\bar{F}(k)$ is the mean of the conditional distribution of $x(k-1)$. Further both $\phi(k, .)$ and $\bar{F}(k)$ are affine and can be determined independently.

Partially Nested Information Structure: Since the optimal control problem is difficult for general information patterns, it is of interest to look for specific information pattern for which optimal control is known. Partially nested information structure is one such pattern. A partially nested information pattern is one for which for any two controllers i and j, if $u_i(k)$ affects information at controller j at time t, then the controller j at time t has access to all the information that controller i had at time k (from which $u_i(k)$ was calculated). For a quadratic cost, if the information structure is partially nested, the optimal control is affine.