Structural Results and Explicit Solution for Two-Player LQG Systems on a Finite Time Horizon

Laurent Lessard, UC Berkeley
Ashutosh Nayyar, UC Berkeley

IEEE Conference on Decision and Control
Firenze, Italy
December 13, 2013
Optimal decentralized control is difficult in general

- Linear policies are not optimal
- Finding best linear controller is difficult
- Unknown structure

Some decentralized control problems are easy!
- LQG with nested information
Problem formulation (nested)

\[ x_{t+1}^1 = f_1(x_t^1, u_t^1, w_t^1) \]
\[ y_t^1 = g_1(x_t^1, v_t^1) \]
\[ \mathcal{I}_t^1 = \{y_0^1, \ldots, y_{t-1}^1\} \]

\[ x_{t+1}^2 = f_2(x_t^1, x_t^2, u_t^1, u_t^2, w_t^2) \]
\[ y_t^2 = g_2(x_t^1, x_t^2, v_t^2) \]
\[ \mathcal{I}_t^2 = \{y_0^1, \ldots, y_{t-1}^1, y_0^2, \ldots, y_{t-1}^2\} \]
Problem formulation (LQG)

Linear dynamics

\[
\begin{align*}
x_{t+1}^1 &= f_1(x_t^1, u_t^1, w_t^1) \\
y_t^1 &= g_1(x_t^1, v_t^1)
\end{align*}
\]

\[
\begin{align*}
x_{t+1}^2 &= f_2(x_t^1, x_t^2, u_t^1, u_t^2, w_t^2) \\
y_t^2 &= g_2(x_t^1, x_t^2, v_t^2)
\end{align*}
\]

Quadratic cost

Minimize the expected quadratic cost

\[
J = \mathbb{E} \left( \sum_{t=0}^{T-1} q_t(x_t^1, x_t^2, u_t^1, u_t^2) + q_T(x_T^1, x_T^2) \right)
\]

where \( u_t^1 = k_1(I_t^1) \) and \( u_t^2 = k_2(I_t^2) \).

Gaussian disturbances

Initial state \( \{x_0^1, x_0^2\} \) and disturbances \( \{w_t^1, w_t^2, v_t^1, v_t^2\} \) for \( t = 0, \ldots, T - 1 \) are mutually independent and jointly Gaussian.
Optimal two-player controller has the form

\[ u_t^1 = K^{11} \zeta_t \]
\[ u_t^2 = K^{21} \zeta_t + K^{22} \xi_t \]

where \( \zeta_t = \mathbf{E}(x_t | \mathcal{I}_t^1) \) and \( \xi_t = \mathbf{E}(x_t | \mathcal{I}_t^2) \).

Estimates \( \zeta_t \) and \( \xi_t \) can be computed \textit{recursively}. Similar to the Kalman Filter.

All estimation and control gains are computable in \( \mathcal{O}(T) \).
Partial nestedness

Problem is partially nested (Ho and Chu, 1972)

Therefore,

- There exists an optimal policy that is linear
- Person-by-person optimality implies global optimality
Centralized sufficient statistics

Controller may use entire information history

\[ \mathcal{I}_t = \{y_0, \ldots, y_{t-1}\} \]

Optimal controller is \textit{linear} of the form

\[ u_t = K_t \xi_t, \]

where

\[ \xi_t = \mathbb{E}(x_t | \mathcal{I}_t) \]
Two-player sufficient statistics

\[
\begin{align*}
\mathbb{E} \left( \begin{bmatrix} x_t \\ \mathcal{I}_t^1 \\ \mathcal{I}_t^2 \end{bmatrix} \mid \mathcal{I}_t^2 \right) &= \begin{bmatrix} \xi_t \\ \mathcal{I}_t^1 \\ \mathcal{I}_t^2 \end{bmatrix}, \quad \text{where } \xi_t = \mathbb{E}(x_t \mid \mathcal{I}_t^2).
\end{align*}
\]
Two-player sufficient statistics
Two-player sufficient statistics

\[
\begin{align*}
\mathbb{E} \left( \begin{bmatrix} x_t \\ \xi_t \end{bmatrix} \mid \mathcal{I}_t^1 \right) & = \begin{bmatrix} \zeta_t \\ \zeta_t \end{bmatrix} \quad \text{and} \quad \zeta_t = \mathbb{E}(x_t \mid \mathcal{I}_t^1). \\
\end{align*}
\]
Theorem
The optimal two-player controller is linear, of the form

\[
\begin{align*}
  u_1^t &= K^{11} \zeta_t \\
  u_2^t &= K^{21} \zeta_t + K^{22} \xi_t
\end{align*}
\]

where \( \zeta_t = \mathbb{E}(x_t \mid \mathcal{I}_t^1) \) and \( \xi_t = \mathbb{E}(x_t \mid \mathcal{I}_t^2) \)
Main results

- Optimal two-player controller has the form

\[
\begin{align*}
    u_t^1 &= K_{11} \zeta_t \\
    u_t^2 &= K_{21} \zeta_t + K_{22} \xi_t
\end{align*}
\]

where \( \zeta_t = \mathbb{E}(x_t | \mathcal{I}_t^1) \) and \( \xi_t = \mathbb{E}(x_t | \mathcal{I}_t^2) \).

- Estimates \( \zeta_t \) and \( \xi_t \) can be computed recursively. Similar to the Kalman Filter.

- All estimation and control gains are computable in \( \mathcal{O}(T) \).
Centralized estimator dynamics

The sufficient statistic $\xi_t = \mathbf{E}(x_t | I_t)$ computed recursively:

$$\xi_{t+1} = A\xi_t + Bu_t - L_t(y_t - C\xi_t)$$
$$u_t = K_t \xi_t$$
Two-player estimator dynamics

Then the optimal *two-player* controller is of the form

\[
\begin{align*}
\zeta_{t+1} &= A\zeta_t + BK_t\zeta_t - \hat{L}_t(y_t - C\zeta_t) \\
\xi_{t+1} &= A\xi_t + Bu_t - L_t(y_t - C\xi_t) \\
u_t &= K_t\zeta_t + \hat{K}_t(\xi_t - \zeta_t)
\end{align*}
\]

Again, \(\zeta_t = \mathbb{E}(x_t | \mathcal{I}_t^1)\) and \(\xi_t = \mathbb{E}(x_t | \mathcal{I}_t^2)\).

- Special structure: \(\hat{L}_t \sim \begin{bmatrix} \ast & 0 \\ \ast & 0 \end{bmatrix}\) and \(\hat{K}_t \sim \begin{bmatrix} 0 & 0 \\ \ast & \ast \end{bmatrix}\).
Main results

- Optimal two-player controller has the form

\[
\begin{align*}
u_t^1 &= K_{11}^1 \zeta_t \\
u_t^2 &= K_{21}^2 \zeta_t + K_{22}^2 \xi_t
\end{align*}
\]

where \( \zeta_t = \mathbb{E}(x_t | \mathcal{I}_t^1) \) and \( \xi_t = \mathbb{E}(x_t | \mathcal{I}_t^2) \).

- Estimates \( \zeta_t \) and \( \xi_t \) can be computed recursively. Similar to the Kalman Filter.

- All estimation and control gains are computable in \( \mathcal{O}(T) \).
Centralized recursion

The optimal centralized controller is of the form

\[ \xi_{t+1} = A \xi_t + Bu_t - L_t (y_t - C \xi_t) \]
\[ u_t = K_t \xi_t \]

The \( L_t \) gains satisfy a **forward** recursion

The \( K_t \) gains satisfy a **backward** recursion

\[ \Sigma_0 = \Sigma_{\text{init}} \]
\[ \Sigma_{t+1} = \psi(\Sigma_t) \]
\[ L_t = h(\Sigma_t) \]
\[ P_T = P_{\text{final}} \]
\[ P_t = \phi(P_{t+1}) \]
\[ K_t = k(P_{t+1}) \]
Coupling diagram

- each arrow is a function evaluation
- can be computed in $O(T)$
Two-player estimator dynamics

The optimal *two-player* controller is of the form

\[
\begin{align*}
\zeta_{t+1} &= A\zeta_t + BK_t\zeta_t - \hat{L}_t(y_t - C\zeta_t) \\
\xi_{t+1} &= A\xi_t + Bu_t - L_t(y_t - C\xi_t) \\
u_t &= K_t\zeta_t + \hat{K}_t(\xi_t - \zeta_t)
\end{align*}
\]

- $L_t$ and $K_t$ are *the same* as in the centralized solution
- $\hat{L}_t$ and $\hat{K}_t$ have new recursions...
Two-player recursion

The $\hat{L}_t$ gains satisfy a *forward* recursion

\[
\hat{\Sigma}_0 = \Sigma_{\text{init}} \\
\hat{\Sigma}_{t+1} = \psi(\hat{\Sigma}_t, \hat{K}_t) \\
\hat{L}_t = h(\hat{\Sigma}_t, \hat{K}_t)
\]

The $\hat{K}_t$ gains satisfy a *backward* recursion

\[
\hat{P}_T = P_{\text{final}} \\
\hat{P}_t = \phi(\hat{P}_{t+1}, \hat{L}_t) \\
\hat{K}_t = k(\hat{P}_{t+1}, \hat{L}_t)
\]

These equations are *coupled*. 
Coupling diagram

- each arrow is a function evaluation
- *everything* is coupled!
Key observation

- The $\hat{L}_t$ gains satisfy a *forward* recursion

$$\hat{\Sigma}_{t+1} = \psi(\hat{\Sigma}_t, \hat{K}_t)$$
$$\hat{L}_t = h(\hat{\Sigma}_t, \hat{K}_t)$$

- Partition: $\hat{\Sigma}_t = \begin{bmatrix} \hat{\Sigma}_{11}^t & \hat{\Sigma}_{12}^t \\ \hat{\Sigma}_{21}^t & \hat{\Sigma}_{22}^t \end{bmatrix}$ and $\hat{L}_t = \begin{bmatrix} \hat{L}_{11}^t \\ \hat{L}_{21}^t \end{bmatrix}$.

- Then we have

$$\hat{\Sigma}_{11}^{t+1} = \psi_1(\hat{\Sigma}_{11}^t)$$
$$\hat{L}_{11}^t = h_1(\hat{\Sigma}_{11}^t)$$
$$\hat{\Sigma}_{21}^{t+1} = \psi_2(\hat{\Sigma}_{21}^t, \hat{K}_t)$$
$$\hat{L}_{21}^t = h_2(\hat{\Sigma}_{21}^t, \hat{K}_t)$$

where $\psi_2$ and $h_2$ are *affine functions*
Modified coupling diagram
Modified coupling diagram
Gain computation

- Compute $\Sigma_{0:T-1}$ and $\hat{\Sigma}_{0:T-1}^{11}$ (forward recursion)
- Compute $P_{0:T-1}$ and $\hat{P}_{0:T-1}^{22}$ (backward recursion)
- Solve for $(\hat{\Sigma}_{0:T-1}^{21}, \hat{L}_{0:T-1}^{21}, \hat{P}_{0:T-1}^{21}, \hat{K}_{0:T-1}^{21})$ simultaneously (sparse linear equations)

All steps are $\mathcal{O}(T)$. 
Main results

- Optimal two-player controller has the form

\[
\begin{align*}
  u_1^t &= K_{11} \zeta_t \\
  u_2^t &= K_{21} \zeta_t + K_{22} \xi_t
\end{align*}
\]

where \( \zeta_t = \mathbb{E}(x_t \mid \mathcal{I}_t^1) \) and \( \xi_t = \mathbb{E}(x_t \mid \mathcal{I}_t^2) \).

- Estimates \( \zeta_t \) and \( \xi_t \) can be computed \textit{recursively}. Similar to the Kalman Filter.

- All estimation and control gains are computable in \( \mathcal{O}(T) \).
Grazie!

Papers and additional information available at:

www.laurentlessard.com