Centralized minimax control

Mukul Gagrani and Ashutosh Nayyar

Computer Engineering Technical Report Number CENG-2016-02

Ming Hsieh Department of Electrical Engineering – Systems
University of Southern California
Los Angeles, California 90089-2562

September 2016

Centralized minimax control

Mukul Gagrani and Ashutosh Nayyar

I. System Model

Consider a centralized minimax control problem for a finite time horizon of T. The state dynamics are given as

$$X_{t+1} = f_t(X_t, U_t, W_t), \tag{1}$$

where $U_t \in \mathcal{U}_t$ is the control action and $W_t \in \mathcal{W}_t$ is the input disturbance. The initial state X_1 is assumed to lie in the set $\mathcal{X}_1$. The controller has observations of the following form

$$Y_t = h_t(X_t, V_t), \tag{2}$$

where $V_t \in \mathcal{V}_t$ is the measurement noise. We assume that all the variables take values in finite sets.

Define $Q = (X_1, W_{1:T-1}, V_{1:T-1})$ to be the vector of all the uncertainity in the system. We assume that all the components of Q are independent in the sense that Q takes value in the product set $\mathcal{Q} = \mathcal{X}_1 \times \mathcal{W}_1 \times \cdots \mathcal{W}_{T-1} \times \mathcal{V}_1 \times \cdots \times \mathcal{V}_{T-1}$. The information available to the controller at time t is $Y_{1:t}, U_{1:t-1}$. The controller maps this information to the control action at time t as follows:

$$U_t = g_t(Y_{1:t}, U_{1:t-1}). \tag{3}$$

The collection $\mathbf{g} = (g_1, \ldots, g_T)$ is called the control strategy. The cost incurred by a control strategy $\mathbf{g}$ is given as

$$J(\mathbf{g}) = \max_{q \in \mathcal{Q}} c(X_T). \tag{4}$$

The objective is to find the control strategy $\mathbf{g}^*$ which minimizes (4).

Consider a realization $y_{1:t}, u_{1:t-1}$ of the controller's information at time t. Define $\hat{Q}(y_{1:t}, u_{1:t-1})$ to be the set of all vectors $q \in \mathcal{Q}$ which are consistent with $y_{1:t}, u_{1:t-1}$ i.e. q belongs to $\hat{Q}(y_{1:t}, u_{1:t-1})$ if and only if $q \in \mathcal{Q}$ and the vectors $q, y_{1:t}, u_{1:t-1}$ together satisfy the system and measurement equations (1), (2) for all times until t. Similarly, define $\hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)$ to be the set of all feasible values of Y_{t+1} given that the realization of controller's information at time t is $y_{1:t}, u_{1:t-1}$ and u_t is the action taken at time t. Then, [1] shows that the optimal strategy $\mathbf{g}^*$ is given by the following dynamic program.

Lemma 1. *Define*

$$V_T(y_{1:T}, u_{1:T-1}) = \max_{q \in \hat{Q}(y_{1:T}, u_{1:T-1})} c(X_T), \tag{5}$$

$$V_t(y_{1:t}, u_{1:t-1}) = \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)} V_{t+1}(y_{1:t}, u_{1:t-1}, y_{t+1}, u_t), \quad t = 1, \ldots, T-1. \tag{6}$$

The minimizing u_t in (6) *is the optimal action when controller's information is $y_{1:t}, u_{1:t-1}$.*

In the case of terminal cost problem, the value functions can be written as functions of the set of feasible values of the current state X_t consistent with the current information [1]. Define the state uncertainty set $\Pi_t(y_{1:t}, u_{1:t-1})$ as the set of all possible values of the state X_t consistent with the information $y_{1:t}, u_{1:t-1}$. For brevity, we write $\Pi_t(y_{1:t}, u_{1:t-1})$ as Π_t when the realization of information is clear from the context. The time evolution of the state uncertainty sets is characterized by the following lemma.

Lemma 2. *The state uncertainty set Π_{t+1} at time $t+1$ can be derived using Π_t, y_{t+1} and u_t i.e.*

$$\Pi_{t+1} = \Phi_t(\Pi_t, y_{t+1}, u_t). \tag{7}$$

Proof. Given Π_t, u_t we know that any feasible value of x of X_{t+1} must satisfy (1) for some $w_{t+1} \in \mathcal{W}_{t+1}$ and $x_t \in \Pi_t$. Also, after receiving the observation y_{t+1} at time $t+1$ we can refine our belief on X_{t+1} and construct Π_{t+1} as follows:

$$\Pi_{t+1} = \{x : h_{t+1}(x, v) = y_{t+1},\ x = f_t(x_t, u_t, w), \text{ for some } x_t \in \Pi_t,\ w \in \mathcal{W}_{t+1},\ v \in \mathcal{V}_{t+1}\}. \tag{8}$$

From the above equation it is straightforward to conclude that Π_{t+1} is a transformation of Π_t, y_{t+1}, u_t and hence we can write $\Pi_{t+1} = \Phi_t(\Pi_t, y_{t+1}, u_t)$. $\square$

Now, we can state the result which simplifies the dynamic program of [1] by specifying the value functions in terms of the state uncertainty set Π_t.

Theorem 1. *The dynamic program in Lemma 1 can be simplified to the following: For each possible state uncertainty set π_t, $t = T, T-1, \ldots, 1$, define*

$$V_T(\pi_T) = \max_{x_T \in \pi_T} c(x_T) \tag{9}$$

$$V_t(\pi_t) = \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \hat{Y}_{t+1}(\pi_t, u_t)} [V_{t+1}(\Phi_t(\pi_t, u_t, y_{t+1}))], \quad t = 1, \ldots, T-1, \tag{10}$$

where $\hat{Y}_{t+1}(\pi_t, U_t) = \{y : y = h_{t+1}(f_t(x_t, u_t, w), v)$ for some $x_t \in \pi_t, w \in \mathcal{W}_{t+1}, v \in \mathcal{V}_{t+1}\}$. A minimizing u_t in (10) is an optimal action when the state uncertainty set is π_t.

Proof. Let π_t be the state uncertainty set corresponding to the realization $y_{1:t}, u_{1:t-1}$. Using the results of [1] it is clear that the value functions can be written as a function of π_t. In the maximization problem in (5), each $q \in \hat{Q}(y_{1:T}, u_{1:T-1})$ produces an $x_T \in \pi_T$. Conversely, in the maximization problem of (9), each $x_T \in \pi_T$ is produced by some $q \in \hat{Q}(y_{1:T}, u_{1:T-1})$. Therefore, (5) and (9) are equivalent.

Proceeding inductively, assume that $V_{t+1}(y_{1:t+1}, u_{1:t}) = V_{t+1}(\pi_{t+1})$, where π_{t+1} is the state uncertainty set corresponding to the realization $y_{1:t+1}, u_{1:t}$. Now, in order to show the equivalence of (6) and (10) we need to show that the sets $\hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)$ and $\hat{Y}_{t+1}(\pi_t, u_t)$ are equal. Let $y \in \hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)$. This implies there exists $w_{t+1} \in \mathcal{W}_{t+1}$, $v_{t+1} \in \mathcal{V}_{t+1}$ and a x_t consistent with $y_{1:t}, u_{1:t-1}$ which together with y satisfy (1) and (2) for time $t+1$. Also, x_t consistent with $y_{1:t}, u_{1:t-1}$ implies that $x_t \in \pi_t$. Therefore $y \in \hat{Y}_{t+1}(\pi_t, u_t)$ and hence

$$\hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t) \subset \hat{Y}_{t+1}(\pi_t, u_t). \tag{11}$$

Next consider $y \in \hat{Y}_{t+1}(\pi_t, u_t)$. Thus there exists a $x_t \in \pi_t$, , $w_{t+1} \in \mathcal{W}_{t+1}$ and $v_{t+1} \in \mathcal{V}_{t+1}$ which together with y satisfy (1) and (2) for time $t+1$. Also, $x_t \in \pi_t$ implies that x_t is consistent with the information $y_{1:t}, u_{1:t-1}$. Therefore, y is a feasible value of Y_{t+1} given the information $y_{1:t}, u_{1:t-1}$ and the control action u_t. Therefore, $y \in \hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)$ and hence

$$\hat{Y}_{t+1}(\pi_t, u_t) \subset \hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t). \tag{12}$$

Using (12) and (11), $\hat{Y}_{t+1}(\pi_t, u_t) = \hat{Y}_{t+1}(y_{1:t}, u_{1:t-1}, u_t)$ which concludes the proof. □

II. Alternate observation model

Consider a centralized minimax problem with the model of Section I but instead of (2), the observations are given as follows

$$Y_t = h_t(X_{t-1}, U_{t-1}) \tag{13}$$

The problem is to obtain the control strategy $\mathbf{g}^*$ which minimizes the worst case cost (4).

Define an extended state $\tilde{X}_t$ as

$$\tilde{X}_t = \begin{pmatrix} X_t \\ X_{t-1} \\ U_{t-1} \end{pmatrix} \tag{14}$$

The time evolution of $\tilde{X}_t$ is given as

$$\tilde{X}_{t+1} = \begin{pmatrix} X_{t+1} \\ X_t \\ U_t \end{pmatrix} = \begin{pmatrix} f_t(X_t, U_t, W_t) \\ X_t \\ U_t \end{pmatrix} = \tilde{f}_t(\tilde{X}_t, U_t, W_t) \tag{15}$$

We can write the observation equation in terms of the extended state as

$$Y_t = h_t(X_{t-1}, U_{t-1}) = \tilde{h}_t(\tilde{X}_t) \tag{16}$$

The cost incurred by the control strategy $\mathbf{g}$ can be redefined as

$$J(\mathbf{g}) = \max_{X_1, \{W_t\}_{t=1}^{T-1}} c(X_T) = \max_{X_1, \{W_t\}_{t=1}^{T-1}} \tilde{c}(\tilde{X}_T) \tag{17}$$

The strategy optimization problem is now an instance of the problem of Section I with $\tilde{X}_t$ as the state of the system. Hence, we can apply the result of Theorem 1 to characterize the optimal control strategy. For a realization $y_{1:t}, u_{1:t-1}$ of the controller's information at time t, the state uncertainty set of the extended state $\tilde{X}_t$ is

$$\begin{aligned}\tilde{\Pi}_t = \Big\{ & \tilde{x}_t \mid \exists\, x_1 \in \mathcal{X}_1, w_{1:t-1} \in \prod_{j=1}^{t-1} \mathcal{W}_j, \\ & \text{such that}\, y_k = \tilde{h}_k(\tilde{x}_k)\, 1 \le k \le t \\ & \tilde{x}_{k+1} = \tilde{f}_k(\tilde{x}_k, u_k, w_k),\, 1 \le k \le t-1 \Big\}\end{aligned}$$

Let Π_t be the set of values of the original state X_t consistent with $y_{1:t}, u_{1:t-1}$.

Define a mapping $\mathbb{P}$ such that

$$\mathbb{P}\begin{pmatrix} X_t \\ X_{t-1} \\ U_{t-1} \end{pmatrix} = X_t.$$

Lemma 3. *The state uncertainty set Π_t of the state X_t and $\tilde{\Pi}_t$ of the extended state $\tilde{X}_t$ are related as*

$$\Pi_t = \mathbb{P}(\tilde{\Pi}_t) \tag{18}$$

Proof. This follows straight from the definition of $\tilde{X}_t$, $\tilde{\Pi}_t, \Pi_t$. □

From Section 1, we know that time evolution of $\tilde{\Pi}_t$ is given as $\tilde{\Pi}_{t+1} = \tilde{\Phi}(\tilde{\Pi}_t, y_{t+1}, u_t)$ for an appropriate transformation $\tilde{\Phi}$ which is characterized using Lemma 2. Hence, using Theorem 1 we can write the dynamic program to compute the optimal control strategy as follows

$$\tilde{V}_T(\tilde{\pi}_T) = \max_{\tilde{x}_T \in \tilde{\pi}_T} \tilde{c}(\tilde{x}_T) \tag{19}$$

$$\tilde{V}_t(\tilde{\pi}_t) = \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \hat{Y}_{t+1}(\tilde{\pi}_t, u_t)} [\tilde{V}_{t+1}(\tilde{\Phi}(\tilde{\pi}_t, u_t, y_{t+1}))], \quad t = 1, \dots, T-1, \tag{20}$$

where

$$\hat{Y}_{t+1}(\tilde{\pi}_t, u_t) = \{y : y = \tilde{h}_{t+1}(\tilde{f}_t(\tilde{x}_t, u_t, w_t)) \text{ for some } \tilde{x}_t \in \tilde{\pi}_t, w_t \in \mathcal{W}_t\}. \tag{21}$$

A minimizing u_t in (20) is an optimal action when the uncertainty set of the extended state is $\tilde{\pi}_t$.

We will now show that the above dynamic program in terms of the uncertainty set of the extended state can be transformed into an equivalent dynamic program in terms of the uncertainty set of the original state X_t. The next two lemmas will be useful in the subsequent derivation of the main result.

Lemma 4. *There exists a transformation $\breve{\Phi}$ such that $\mathbb{P}(\tilde{\Phi}(\tilde{\pi}_t, y_{t+1}, u_t)) = \breve{\Phi}(\mathbb{P}(\tilde{\pi}_t), y_{t+1}, u_t)$.*

Proof. Let $\mathbb{P}(\tilde{\pi}_t) = \pi_t$.

$$\begin{aligned}\mathbb{P}(\tilde{\Phi}(\tilde{\pi}_t, y_{t+1}, u_t)) &= \{\mathbb{P}(\tilde{x}) | \tilde{h}_{t+1}(\tilde{x}) = y_{t+1}, \tilde{x} = \tilde{f}_t(\tilde{x}_t, u_t, w_t),\ \tilde{x}_t \in \tilde{\pi}_t, w_t \in \mathcal{W}_t\} \\ &= \{\mathbb{P}(\tilde{x}) | h_{t+1}(\mathbb{P}(\tilde{x}_t), u_t) = y_{t+1}, \tilde{x} = \tilde{f}_t(\tilde{x}_t, u_t, w_t),\ \tilde{x}_t \in \tilde{\pi}_t, w_t \in \mathcal{W}_t\} \\ &= \{\mathbb{P}(\tilde{x}) | h_{t+1}(\mathbb{P}(\tilde{x}_t), u_t) = y_{t+1}, \tilde{x} = (f_t(\mathbb{P}(\tilde{x}_t), u_t, w_t), \mathbb{P}(\tilde{x}_t), u_t),\ \tilde{x}_t \in \tilde{\pi}_t, w_t \in \mathcal{W}_t\} \\ &= \{\mathbb{P}(\tilde{x}) | h_{t+1}(x_t, u_t) = y_{t+1}, \tilde{x} = (f_t(x_t, u_t, w_t), x_t, u_t),\ x_t \in \mathbb{P}(\tilde{\pi}_t), w_t \in \mathcal{W}_t\} \\ &= \{x | h_{t+1}(x_t, u_t) = y_{t+1}, x = f_t(x_t, u_t, w_t),\ x_t \in \pi_t, w_t \in \mathcal{W}_t\} \\ &=: \breve{\Phi}(\pi_t, y_{t+1}, u_t).\end{aligned}$$

□

Lemma 5. *For $\pi_t = \mathbb{P}(\tilde{\pi}_t)$, define $\breve{Y}_{t+1}(\pi_t, u_t) = \{y | y = h_{t+1}(x_t, u_t)$ for some $x_t \in \pi_t\}$. Then, $\hat{Y}_{t+1}(\tilde{\pi}_t, u_t) = \breve{Y}_{t+1}(\pi_t, u_t)$.*

Proof. From (21),

$$\begin{aligned}\hat{Y}_{t+1}(\tilde{\pi}_t, u_t) &= \{y : y = \tilde{h}_{t+1}(\tilde{f}_t(\tilde{x}_t, u_t, w_t)),\, \tilde{x}_t \in \tilde{\pi}_t, w_t \in \mathcal{W}_t\} \\ &= \{y : y = h_{t+1}(\mathbb{P}(\tilde{x}_t), u_t),\, \tilde{x}_t \in \tilde{\pi}_t\} \\ &= \{y : y = h_{t+1}(x_t, u_t),\, x_t \in \pi_t\} \\ &= \breve{Y}_{t+1}(\pi_t, u_t).\end{aligned}$$

$\square$

We can now write the dynamic program of (19) and (20) in terms of π_t instead of $\tilde{\pi}_t$.

Theorem 2. *Define the functions $V_t(\pi_t)$ as*

$$V_T(\pi_T) = \max_{x_T \in \pi_T} c(x_T) \tag{22}$$

$$V_t(\pi_t) = \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \check{Y}_{t+1}(\pi_t, u_t)} [V_{t+1}(\check{\Phi}(\pi_t, u_t, y_{t+1}))],\ t = 1, \dots, T-1 \tag{23}$$

Then,

$$\tilde{V}_t(\tilde{\pi}_t) = V_t(\pi_t),\ 1 \le t \le T \tag{24}$$

where $\pi_t = \mathbb{P}(\tilde{\pi}_t)$. Further, a minimizing u_t in (23) is an optimal action when the state uncertainty set is π_t.

Proof. The proof follows by induction. For $t = T$,

$$\begin{aligned}
\tilde{V}_T(\tilde{\pi}_T) &= \max_{\tilde{x}_T \in \tilde{\pi}_T} \tilde{c}(\tilde{x}_T) \\
&= \max_{\tilde{x}_T \in \tilde{\pi}_T} c(\mathbb{P}(\tilde{x}_T)) = \max_{x_T \in \mathbb{P}(\tilde{\pi}_T)} c(x_T) && (25) \\
&= \max_{x_T \in \pi_T} c(x_T) = V_T(\pi_T), && (26)
\end{aligned}$$

where (25) follows from the definition of $\tilde{c}(\cdot)$.

Now, let (24) be true for $t+1$. Then using the definition (20),

$$\begin{aligned}
\tilde{V}_t(\tilde{\pi}_t) &= \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \hat{Y}_{t+1}(\tilde{\pi}_t, u_t)} \tilde{V}_{t+1}(\tilde{\Phi}(\tilde{\pi}_t, u_t, y_{t+1})) \\
&= \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \hat{Y}_{t+1}(\tilde{\pi}_t, u_t)} V_{t+1}(\mathbb{P}(\tilde{\Phi}(\tilde{\pi}_t, u_t, y_{t+1}))) && (27) \\
&= \min_{u_t \in \mathcal{U}_t} \max_{y_{t+1} \in \check{Y}_{t+1}(\pi_t, u_t)} V_{t+1}(\check{\Phi}(\pi_t, u_t, y_{t+1})) = V_t(\pi_t), && (28)
\end{aligned}$$

where (27) follows from the induction hypothesis and (28) follows from Lemmas 4 and 5. Hence, (24) is true by induction. It also follows that the minimizing values of u_t in (20) and (23) are the same.

$\square$

REFERENCES

[1] Dimitri Bertsekas and Ian Rhodes, *Sufficiently Informative Functions and the Minimax Feedback Control of Uncertain Dynamic Systems*, IEEE transactions on Automatic control, 1973