# 2-Player LQ Stackelberg Game Derivation

Hamzah Khan

November 5, 2022

## 1 Problem Formulation

Assume we are operating in discrete time with the following linear system:

$$x_{t+1} = A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2. \tag{1}$$

Furthermore, let us assume a feedback information structure where $u_t^{i*}$ depends on $t$ and $x_t$. This assumption is an oversimplification for many interesting problems (consider any system without full state state observability), but it is frequently closer to reality than an open-loop information structure.

We have two players attempting to minimize quadratic costs over some horizon $T$. Let us assume that player 1 is able to predict what player 2 will do at time $t$ in response to the input $u_t^1$. This implies that when we optimize for player 1's objective, we assume that player 2's input $u_t^2(u_t^1)$ is a function of $u_t^1$. For player 2's objective, we treat $u_t^2$ as a variable.

$$J^1 = \frac{1}{2} \sum_{t=1}^{T} x_t^\intercal Q_t^1 x_t + u_t^{1\intercal} R_t^{11} u_t^1 + u_t^{2\intercal}(u_t^1) R_t^{12} u_t^2(u_t^1) \tag{2}$$

$$J^2 = \frac{1}{2} \sum_{t=1}^{T} x_t^\intercal Q_t^2 x_t + u_t^{2\intercal} R_t^{22} u_t^2 + u_t^{1\intercal} R_t^{21} u_t^1 \tag{3}$$

The coupled optimization for each player then takes the following form.

$$\begin{aligned} &\min_{u_{1:T}^i} J^i\left(x_{1:T}, u_{1:T}^1, u_{1:T}^2\right) \\ &\text{s.t. } x_{t+1} = A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 \end{aligned} \tag{4}$$

We present the solution to this Stackelberg game in the next section.

## 2 Solution

We are looking for a *feedback* equilibrium, i.e. a set of functions $\gamma_t^i$ such that $u_t^i \equiv \gamma_t^i(x_t)$.

If every player is operating at Stackelberg, their value functions and associated optimal controls should be mutually consistent. Furthermore, we can infer that each value function $V_t^i$ is quadratic because each is a sum of quadratic functions. Therefore, we see that the optimal controls for each player, at each time, will be an affine function of the state. Let us presuppose the form of the value function and derive the form of the control law and thereby the specific value functions themselves. Our derivation will be constructive in nature and we begin by supposing the value functions have quadratic forms, i.e.

$$V_t^i(x_t) = \frac{1}{2} x_t^\intercal L_t^i x_t, \tag{5}$$

where $L_t^i > 0, L_t^i = L_t^{i\intercal}, L_{T+1}^i = 0$.

Furthermore, we will soon see that the optimal choice for $u_t^i$ will be a linear function of the state $x_t$, i.e.

$$u_t^{i*}(x_t) = -S_t^i x_t, \tag{6}$$

where $S_t^i > 0, S_t^i = S_t^{i\intercal}$

From here, we solve for the four unknowns $S_t^1, S_t^2, L_t^1, L_t^2$ to obtain recursions that let us solve this dynamic programming problem. We note that the leader and follower cases are different and thus we solve them separately.

## 2.1 Solving for Player 2's Unknowns

We start by writing down the Hamilton-Jacobi equations for player 2's value function:

$$V_t^2(x_t) = \min_{u_t^2} \left\{ \frac{1}{2} \left[ x_t^\intercal Q_t^2 x_t + u_t^{2\intercal} R_t^{22} u_t^2 + u_t^{1\intercal} R_t^{21} u_t^1 \right] + V_{t+1}^2(x_{t+1}) \right\}; \qquad u_t^1 \text{ given} \tag{7}$$

$$= \min_{u_t^2} \left\{ \frac{1}{2} \left[ x_t^\intercal Q_t^2 x_t + u_t^{2\intercal} R_t^{22} u_t^2 + u_t^{1\intercal} R_t^{21} u_t^1 \right] + V_{t+1}^2(A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2) \right\}; \quad u_t^1 \text{ given,} \tag{8}$$

with final value $V_{T+1}^i(x_{t+1}) = 0$.

At any time $t$, we can plug (5) into (8):

$$V_t^2(x_t) = \frac{1}{2} \min_{u_t^2} \left\{ x_t^\intercal Q_t^2 x_t + u_t^{2\intercal} R_t^{22} u_t^2 + u_t^{1\intercal} R_t^{21} u_t^1 + \left( A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 \right)^\intercal L_{t+1}^2 \left( A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2 \right) \right\} \tag{9}$$

We can find the feedback control law by finding the minimizer of (9). To do so, we assume strong convexity and set the gradient to 0 as follows:

$$0 = R_t^{22} u_t^{2*} + B_t^{2\intercal} L_{t+1}^2 \left( A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^{2*} \right). \tag{10}$$

We use the symmetry of $L_{t+1}^2$ to rearrange terms in this expression. Note that we can remove $\min_{u_t^2}$ since we already satisfy it through the first-order conditions. We take note of one more relationship before proceeding - given the control law we assume in (6), we see that

$$x_{t+1} = A_t x_t + B_t^1 u_t^{1*} + B_t^2 u_t^{2*} = (A_t - B_t^1 S_t^1 - B_t^2 S_t^2) x_t \tag{11}$$

Plugging in equations (6) and (11) to (10), we then rewrite it in the form:

$$u_t^{2*} = - \left[ R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2 \right]^{-1} B_t^{2\intercal} L_{t+1}^2 \left( A_t - B_t^1 S_t^1 \right) x_t, \tag{12}$$

which fits the form of the control law proposed in (6) and from which we can extract $S_t^2$. Thus,

$$S_t^2 = \left[ R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2 \right]^{-1} B_t^{2\intercal} L_{t+1}^2 \left( A_t - B_t^1 S_t^1 \right). \tag{13}$$

To identify $L_t^2$, we then plug $u_t^{i*} = -S_t^i x_t$ into $V_t^2$ to put the expression in the form we propose in (5) and extract $L_t^2$:

$$
\begin{aligned}
V_t^2(x_t) &= \frac{1}{2}\left[x_t^\intercal Q_t^2 x_t + x_t^\intercal S_t^{2\intercal} R_t^{22} S_t^2 x_t + x_t^\intercal S_t^{1\intercal} R_t^{21} S_t^1 x_t + x_t^\intercal \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)^\intercal L_{t+1}^2 \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right) x_t\right] \\
&= \frac{1}{2} x_t^\intercal \left[Q_t^2 + S_t^{2\intercal} R_t^{22} S_t^2 + S_t^{1\intercal} R_t^{21} S_t^1 + \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)^\intercal L_{t+1}^2 \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)\right] x_t.
\end{aligned}
$$

Thus, we see that

$$
L_t^2 = Q_t^2 + S_t^{2\intercal} R_t^{22} S_t^2 + S_t^{1\intercal} R_t^{21} S_t^1 + \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)^\intercal L_{t+1}^2 \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right), \tag{14}
$$

where $L_{T+1}^i = Q_{T+1}^i$.

## 2.2 Solving for Player 1's Unknowns

Recall for player 1's value function, player 2's input $u_t^2(u_t^1)$ is a function of $u_t^1$. We can repeat the same process as in Section 2.1, with a change in the derivation of the recursion based on our assumptions about the function $u_t^2(u_t^1)$. We start by writing down the Hamilton-Jacobi equations for player 1's value function:

$$
V_t^1(x_t) = \min_{u_t^1} \left\{ \frac{1}{2}\left[x_t^\intercal Q_t^1 x_t + u_t^{1\intercal} R_t^{11} u_t^1 + u_t^{2\intercal}(u_t^1) R_t^{12} u_t^2(u_t^1)\right] + V_{t+1}^1(x_{t+1}) \right\} \tag{15}
$$

$$
= \min_{u_t^1} \left\{ \frac{1}{2}\left[x_t^\intercal Q_t^1 x_t + u_t^{1\intercal} R_t^{11} u_t^1 + u_t^{2\intercal}(u_t^1) R_t^{12} u_t^2(u_t^1)\right] + V_{t+1}^1(A_t x_t + B_t^1 u_t^1 + B_t^2 u_t^2) \right\} \tag{16}
$$

with final value $V_{T+1}^1(x_{t+1}) = 0$.

When we, under the same assumptions of strong convexity and set the gradient to 0, we find that we must apply the chain rule to compute the derivative because of the dependence of $u_t^2$ on $u_t^1$. We can now differentiate and solve for the minimum cost control $u_t^{1*}$ as we did for $u_t^{2*}$ in (10):

$$
0 = \frac{\partial V_t^1}{\partial u_t^1} = R_t^{11} u_t^{1*} + u_t^{2\intercal}(u_t^{1*}) R_t^{12} \frac{\partial u_t^2}{\partial u_t^1} + \left(B_t^1 + B_t^2 \frac{\partial u_t^2}{\partial u_t^1}\right)^\intercal L_{t+1}^1 \left(A_t x_t + B_t^1 u_t^{1*} + B_t^2 u_t^2(u_t^1)\right) \tag{17}
$$

We will need the derivative of $u_t^{2*}(u_t^1)$ with respect to $u_t^1$ later, so we compute that here using the chain rule:

$$
\frac{\partial u_t^2}{\partial u_t^1} = -\left[R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2\right]^{-1} B_t^{2\intercal} L_{t+1}^2 B_t^1. \tag{18}
$$

Let us next define a variable

$$
D_t = \left[R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2\right]^{-1} B_t^{2\intercal} L_{t+1}^2. \tag{19}
$$

Then, for the sake of easier substitution, we can rewrite (12) as

$$
u_t^{2*}\left(u_t^1\right) = -D_t A_t x_t + \frac{\partial u_t^2}{\partial u_t^1} u_t^1 \tag{20}
$$

and (18) as

$$
\frac{\partial u_t^2}{\partial u_t^1} = -D_t B_t^1 \tag{21}
$$

Recall Equation (17). We underline the terms we will replace in the next step.

$$0 = R_t^{11} u_t^{1*} + \underline{u_t^{2\intercal}(u_t^{1*})} R_t^{12} \underline{\underline{\frac{\partial u_t^2}{\partial u_t^1}}} + \left( B_t^1 + B_t^2 \underline{\underline{\frac{\partial u_t^2}{\partial u_t^1}}} \right)^\intercal L_{t+1}^1 \left( A_t x_t + B_t^1 u_t^{1*} + B_t^2 \underline{u_t^{2\intercal}(u_t^{1*})} \right) \tag{22}$$

We rewrite Equation 20 using Equation 21 to get

$$u_t^{2\intercal *}(u_t^{1*}) = -D_t[A_t x_t + B_t^1 u_t^{1*}] \tag{23}$$

We plug in Equation 23 for $\underline{u_t^2(u_t^1)}$ and Equation 21 for $\underline{\underline{\frac{\partial u_t^2}{\partial u_t^1}}}$ into Equation 22.

$$\begin{aligned} 0 = R_t^{11} u_t^{1*} &+ \left[-D_t[A_t x_t + B_t^1 u_t^{1*}]\right] R_t^{12} \left[-D_t B_t^1\right] \\ &+ \left(B_t^1 + B_t^2 \left[-D_t B_t^1\right]\right)^\intercal L_{t+1}^1 \left(A_t x_t + B_t^1 u_t^{1*} + B_t^2 \left[-D_t[A_t x_t + B_t^1 u_t^{1*}]\right]\right) \end{aligned} \tag{24}$$

Finally, we solve for $u_t^{1*}$, which will give us $S_t^1$.

$$\begin{aligned} u_t^{1*} = -S_t^1 x_t = -&\left[R_t^{11} + B_t^{1\intercal} D_t^T R_t^{12} D_t B_t^1 + \left(B_t^1 - B_t^2 D_t B_t^1\right)^\intercal L_{t+1}^1 \left(B_t^1 - B_t^2 D_t B_t^1\right)\right]^{-1} \\ &\cdot \left[B_t^{1\intercal} D_t^T R_t^{12} D_t + \left(B_t^1 - B_t^2 D_t B_t^1\right)^\intercal L_{t+1}^1 \left(I - B_t^2 D_t\right)\right] A_t x_t \end{aligned} \tag{25}$$

We solve for $L_t^1$ in the same manner as $L_t^2$, and we find that the recurrence takes the same form.

$$L_t^1 = Q_t^1 + S_t^{1\intercal} R_t^{11} S_t^1 + S_t^{2\intercal} R_t^{12} S_t^2 + \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)^\intercal L_{t+1}^1 \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right) \tag{26}$$

### 2.3 Recursions

$$D_t = \left[R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2\right]^{-1} B_t^{2\intercal} L_{t+1}^2 \tag{27}$$

$$\begin{aligned} S_t^1 = &\left[R_t^{11} + B_t^{1\intercal} D_t^T R_t^{12} D_t B_t^1 + \left(B_t^1 - B_t^2 D_t B_t^1\right)^\intercal L_{t+1}^1 \left(B_t^1 - B_t^2 D_t B_t^1\right)\right]^{-1} \\ &\cdot \left[B_t^{1\intercal} D_t^T R_t^{12} D_t + \left(B_t^1 - B_t^2 D_t B_t^1\right)^\intercal L_{t+1}^1 \left(I - B_t^2 D_t\right)\right] A_t \end{aligned} \tag{28}$$

$$S_t^2 = \left[R_t^{22} + B_t^{2\intercal} L_{t+1}^2 B_t^2\right]^{-1} B_t^{2\intercal} L_{t+1}^2 \left(A_t - B_t^1 S_t^1\right) \tag{29}$$

$$L_t^i = Q_t^i + S_t^{i\intercal} R_t^{ii} S_t^i + S_t^{j\intercal} R_t^{ij} S_t^j + \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right)^\intercal L_{t+1}^i \left(A_t - B_t^1 S_t^1 - B_t^2 S_t^2\right) \qquad (i, j \in [1, 2], i \neq j) \tag{30}$$

This concludes our derivation by construction.

## 3 Extension to Nonlinear Systems

TODO: Describe the iterative solver here.