Feedback LQ Nash Derivation

David Fridovich-Keil

November 14, 2021

1 Problem Formulation

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

$$x_{t+1} = A_t x_t + \sum_{j=1}^{N} B_t^j u_t^j \,. \tag{1}$$

Further, assume we have a feedback information structure, i.e. u_t^{i*} depends upon t and also x_t. Note that this is obviously an oversimplification in many cases (where full state observability is not practical), yet it is also distinct from—and in many cases far closer to reality than—an open-loop information structure, and the two have different solutions.

Each of N players is trying to minimize the following quadratic cost:

$$J^i = \frac{1}{2} \sum_{t=1}^{T} \left[\left(x_t^T Q_t^i + 2 q_t^{iT} \right) x_t + \sum_{j=1}^{N} \left(u_t^{jT} R_t^{ij} + 2 r_t^{ijT} \right) u_t^j \right] \,. \tag{2}$$

2 Solution

We are looking for a *feedback* solution, i.e., a set of functions γ_t^i such that $u_t^i \equiv \gamma_t^i(x_t)$. We start with by writing down the coupled Hamilton-Jacobi equations for each player's value function:

$$\begin{aligned} V_t^i(x_t) &= \min_{u_t^i} \left\{ \frac{1}{2} \Big((x_t^T Q_t^i + 2 q_t^{iT}) x_t + \sum_{j=1}^{N} (u_t^{jT} R_t^{ij} + 2 r_t^{ijT}) u_t^j \Big) + V_{t+1}^i(x_{t+1}) \right\} & (3) \\ &= \min_{u_t^i} \left\{ \frac{1}{2} \Big((x_t^T Q_t^i + 2 q_t^{iT}) x_t + \sum_{j=1}^{N} (u_t^{jT} R_t^{ij} + 2 r_t^{ijT}) u_t^j \Big) + V_{t+1}^i \Big(A_t x_t + \sum_{j=1}^{N} B_t^j u_t^j \Big) \right\} \,, & (4) \end{aligned}$$

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

Now, if every player is operating at Nash their value functions and associated optimal controls should be mutually consistent. Furthermore, we can infer that each value function 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. Unlike the derivation of the open-loop case, in this case our derivation will be constructive in nature.

We begin by supposing that the value functions are quadratic forms (where possible, trying to match naming with Başar and Olsder), i.e.

$$V_t^i(x_t) = \frac{1}{2} \left(x_t^T Z_t^i + 2 \zeta_t^{iT} \right) x_t + n_t^i \,, \tag{5}$$

with $Z_{T+1}^i = 0, \zeta_{T+1}^i = 0$, and $n_{T+1}^i = 0$ to be consistent with the final value condition above.

At any time t, we can plug into (4):

$$V_t^i(x_t) = \min_{u_t^i} \Bigg\{ \frac{1}{2}\Big(\big(x_t^T Q_t^i + 2q_t^{iT}\big)x_t + \sum_{j=1}^N \big(u_t^{jT} R_t^{ij} + 2r_t^{ijT}\big)u_t^j \Big) + \frac{1}{2}\Big(\big(A_t x_t + \sum_{j=1}^N B_t^j u_t^j\big)^T Z_{t+1}^i + 2\zeta_{t+1}^{iT} \Big)\big(A_t x_t + \sum_{j=1}^N B_t^j u_t^j\big) + n_{t+1}^i \Bigg\} . \tag{6}$$

We can find the feedback control law by finding the minimizer of (6). To do so, we assume strong convexity and set the gradient to zero, as follows:

$$0 = R_t^{ii} u_t^i + r_t^{ii} + B_t^{iT} Z_{t+1}^i \big(A_t x_t + \sum_{j=1}^N B_t^j u_t^j\big) + B_t^{iT}\zeta_{t+1}^i . \tag{7}$$

We know from (4) that the optimal choice of u_t^i will be a function of the state x_t, and from (7) we see that it must be an affine function, i.e.,

$$u_t^{i*} = -P_t^i x_t - \alpha_t^i . \tag{8}$$

Plugging this into (7) we obtain

$$0 = -R_t^{ii}(P_t^i x_t + \alpha_t^i) + r_t^{ii} + B_t^{iT} Z_{t+1}^i \big(A_t x_t - \sum_{j=1}^N B_t^j (P_t^j x_t + \alpha_t^j)\big) + B_t^{iT}\zeta_{t+1}^i , \tag{9}$$

from which we can form two similar systems of equations to find the P's and α's (obtained by realizing the equation must be satisfied for all values of x_t), i.e.,

$$(R_t^{ii} + B_t^{iT} Z_{t+1}^i B_t^i) P_t^i + B_t^{iT} Z_{t+1}^i \sum_{j\neq i} B_t^j P_t^j = B_t^{iT} Z_{t+1}^i A_t , \tag{10}$$

$$(R_t^{ii} + B_t^{iT} Z_{t+1}^i B_t^i) \alpha_t^i + B_t^{iT} Z_{t+1}^i \sum_{j\neq i} B_t^j \alpha_t^j = B_t^{iT}\zeta_{t+1}^i . \tag{11}$$

From this we see that it is necessary to derive expressions for the Z's and ζ's as well. To do so, we plug in and rearrange the expression in (6). Note that we have removed the $\min_{u_t^i}$ because we are already satisfying it for these choices of P and α:

$$V_t^i(x_t) = \frac{1}{2}\Big(\big(x_t^T Q_t^i + 2q_t^{iT}\big)x_t + \sum_{j=1}^N \big((P_t^j x_t + \alpha_t^j)^T R_t^{ij} - 2r_t^{ijT}\big)(P_t^j x_t + \alpha_t^j) \Big) + \frac{1}{2}\Big(\big(A_t x_t - \sum_{j=1}^N B_t^j (P_t^j x_t + \alpha_t^j)\big)^T Z_{t+1}^i + 2\zeta_{t+1}^{iT} \Big)\big(A_t x_t - \sum_{j=1}^N B_t^j (P_t^j x_t + \alpha_t^j)\big) + n_{t+1}^i \tag{12}$$

$$= \frac{1}{2} x_t^T \left[Q_t^i + \sum_{j=1}^N P_t^{jT} R_t^{ij} P_t^j + \Big(A_t - \sum_{j=1}^N B_t^j P_t^j\Big)^T Z_{t+1}^i \Big(A_t - \sum_{j=1}^N B_t^j P_t^j\Big) \right] x_t + \left[q_t^i + \sum_{j=1}^N \big(P_t^{jT} R_t^{ij} \alpha_t^j - P_t^{jT} r_t^{ij}\big) + \Big(A_t - \sum_{j=1}^N B_t^j P_t^j\Big)^T \Big(\zeta_{t+1}^i - Z_{t+1}^i \sum_{j=1}^N B_t^j \alpha_t^j\Big)^T \right]^T x_t + \frac{1}{2}\left[\sum_{j=1}^N (\alpha_t^{jT} R_t^{ij} - 2r_t^{ijT})\alpha_t^j - \Big(2\zeta_{t+1}^i - Z_{t+1}^i \sum_{j=1}^N B_t^j \alpha_t^j\Big)^T \sum_{j=1}^N B_t^j \alpha_t^j \right] + n_{t+1}^i . \tag{13}$$

We recognize the terms in brackets as the recursions for parameters of the quadratic value function. That is, setting

$$F_t = A_t - \sum_{j=1}^N B_t^j P_t^j , \tag{14}$$

$$\beta_t = -\sum_{j=1}^{N} B_t^j \alpha_t^j \,, \tag{15}$$

we obtain the recursions

$$Z_t^i = Q_t^i + \sum_{j=1}^{N} P_t^{jT} R_t^{ij} P_t^j + F_t^T Z_{t+1}^i F_t \,,\ Z_{T+1} = 0 \,, \tag{16}$$

$$\zeta_t^i = q_t^i + \sum_{j=1}^{N} \left(P_t^{jT} R_t^{ij} \alpha_t^j - P_t^{jT} r_t^{ij} \right) + F_t^T \left(\zeta_{t+1}^i + Z_{t+1}^i \beta_t \right)^T \,,\ \zeta_{T+1} = 0 \,, \tag{17}$$

$$n_t^i = \frac{1}{2} \left[\sum_{j=1}^{N} (\alpha_t^{jT} R_t^{ij} - 2 r_t^{ijT}) \alpha_t^j - \left(2\zeta_{t+1}^i - Z_{t+1}^i \sum_{j=1}^{N} B_t^j \alpha_t^j \right)^T \sum_{j=1}^{N} B_t^j \alpha_t^j \right] + n_{t+1}^i \,,\ n_{T+1} = 0 \,. \tag{18}$$

This completes our derivation by contruction.