Nationaal Lucht- en Ruimtevaartlaboratorium
National Aerospace Laboratory NLR

NLR-TP-2001-626

Nash equilibria in risk-sensitive dynamic games

M.B. Klompstra

This report is based on an article published in IEEE Transactions on Automatic Control, Vol. 45, No. 7, July 2000. Publisher Item Identifier S 0018-9286(00)06086-4.

Customer:	National Aerospace Laboratory NLR
Working Plan number:	L.1.A
Owner:	National Aerospace Laboratory NLR
Division:	Air Transport
Distribution:	Unlimited
Classification title:	Unclassified
	October 2003

NASH EQUILIBRIA IN RISK-SENSITIVE DYNAMIC GAMES

M.B. Klompstra

National Aerospace Laboratory NLR, P.O. Box 90502, 1006 BM Amsterdam, The Netherlands

e-mail: klompstr@nlr.nl

ABSTRACT

Dynamic games in which each player has an exponential cost criterion are referred to as risk-sensitive dynamic games. In this paper, Nash equilibra are considered for such games. Feedback risk-sensitive Nash equilibrium solutions are derived for two-person discrete-time linear-quadratic nonzero-sum games, both under complete state observation and shared partial observation.

KEYWORDS

Kalman filter, linear quadratic control, Nash, risk-sensitivity.

Contents

(17 pages in total)

1 Introduction

This paper studies discrete-time dynamic games where each player has an exponential-of-integral cost criterion. For short, the latter game is referred to as a *risk-sensitive dynamic game*. Risk-sensitive control is well studied in literature as an extension of classical control. The study of this extension to dynamic games seems to be new. As result of this, we develop Nash equilibrium solutions both under complete observation and shared partial observation.

Linear-Exponential-Gaussian (LEG) control has been introduced in the early 1970's [1], [2]. In Jacobson [1], LEG control for discrete time with perfect state observation is treated. Jacobson also showed an equivalence of the optimal LEG control with the solutions of deterministic (cooperative and noncooperative) zero-sum quadratic games. In Speyer *et al.* [2], LEG control for discrete time with partial state observation is treated. For the case with costs only on the terminal state, it turned out that the feedback control law is a linear function of the current state. For general linear-exponential-quadratic-Gaussian (LEQG) control with costs on the intermediate states, [2] also obtained an optimal controller which is a linear function of the smoothed estimate of the *entire* state history.

Subsequently, Whittle [3] and [4], completed these results and characterized the solution in terms of a certainty equivalence principle. It was Whittle too who introduced the name *risk-sensitive LQG control*. Whittle assumes that the control at the current time is a function of the observation history up to the *previous* time, and obtained the solution for general LEQG control for discrete time with partial state observation. Jaensch and Speyer [5] and Fan *et al.* [6] extend the results of Whittle for the slightly more natural assumption that the control at the current time is a function of the observation history up to the *current* time.

LEQG control in continuous time was treated by Bensoussan and Van Schuppen, [7] for the partially observable case. Bensoussan [8] gives a good characterization of both the complete and partial observation cases.

In James *et al.* [9] finite horizon partially observed risk-sensitive stochastic control for discrete-time for nonlinear systems was considered. As in Whittle [3] and [4], they consider control with one-step-delayed observation. Their approach was motivated by the method used by Bensoussan and Van Schuppen [7] and the well-known separation method for risk-neutral control.

Collings *et al.* [10] present the output feedback discrete-time risk-sensitive LQG control solution derived via the methods in [7], [9], with a one-step delayed observation. With these methods, the solution is obtained without appealing to a certainty equivalence principle. In James and Baras [11], new results are presented concerning the certainty equivalence principle under certain standard assumptions.

In this paper, we go beyond LEQG control by considering discrete-time Nash equilibrium solutions in dynamic games with exponential cost criteria. For dynamic games without exponential cost criteria, i.e. the risk-neutral case, we refer to Başar and Olsder [12]. The risk-sensitive dynamic game problem is formulated in Section 2 for the special case of a two-person linear-quadratic nonzero-sum game. In this section, also the definition of a risk-sensitive Nash equilibrium is given. In Section 3, two theorems for the feedback risk-sensitive Nash equilibrium solution are derived, one for the complete state observation case and one for the shared partial observation case. The derivation leading to these these theorems is mainly based on the results of the book of Başar and Olsder [12] on the theory of noncooperative dynamic games, and the detailed proofs of Whittle [3] Jaensch and Speyer [5] and Fan *et al.* [6] on risk-sensitive control. In Section 4, both theorems are applied to a two-person linear-quadratic nonzero-sum game. First, the feedback risk-sensitive Nash equilibrium solution in the complete observable case is derived. Next, it is shown that the feedback Nash equilibrium solution in the shared partial observable case can be constructed from the feedback risk-sensitive Nash equilibrium design and two risk-modified Kalman filters.

An earlier version of this paper has been presented at the 1995 ACC [13].

2 Problem formulation

We will consider a two-person discrete-time linear-quadratic dynamic game with exponential cost criteria. The system model is described by

$$x_{k+1} = A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2 + w_k \tag{1}$$

where $x_k \in \mathbb{R}^n$, $u_k^i \in \mathbb{R}^{m_i}$, and $w_k \in \mathbb{R}^l$. The measurement model is

$$z_k = H_k x_k + v_k \tag{2}$$

where $z_k \in \mathbb{R}^p$ and $v_k \in \mathbb{R}^p$. The random variable x_0 is normally distributed with mean $\hat{x}_0$ and with covariance matrix P_0, the processes $\{w_k\}$ and $\{v_k\}$ are assumed to be zero-mean, jointly Gaussian, independent random variables for all $k = 0, 1, \ldots, N$ with known covariance matrices $Q_k^w > 0$ and $R_k^v > 0$, respectively.

Cost functional for player i is given by

$$J^i(\theta_i) = \mathbb{E}\left\{-\theta_i \exp(-\theta_i\, \Psi_0^i)\right\} \tag{3}$$

where the random cost Ψ_0^i are

$$\Psi_0^i = \frac{1}{2} x_N' Q_N^i x_N + \frac{1}{2} \sum_{k=0}^{N-1} (x_k' Q_k^i x_k + u_k^{i'} R_k^{ii} u_k^i + u_k^{j'} R_k^{ij} u_k^j). \tag{4}$$

It is assumed that all matrices are of appropriate dimensions, Q_k^i are symmetric and $Q_k^i \geq 0$ for $k = 0, 1, \ldots, N$ and $R_k^{ii} > 0$ for $k = 0, 1, \ldots, N-1$. The parameters θ_i are some scalar constants (positive or negative) which can be characterized as the *risk-sensitivity* parameter; they measure the optimizers' sensitivity to risk. If θ_i is positive, then player i behaves as if unobservables would take values in his advantage, which is an optimistic attitude (i.e., risk-seeking). If θ_i is negative, then player i behaves as if unobservables would take values in his disadvantage, which is a pessimistic attitude (i.e., risk-averse). It is assumed that the players have different risk-sensitivity parameters.

The objective of both players is to minimize their cost function over the class of control laws. It is assumed that both players know the state as well as the cost functions. During the evolution of the game it is assumed that both players have the same information on the (either fully or partially observed) state and know each other's control function up to the previous time. The *information space* $\mathcal{W}_k$ of both players is the same at stage k and is defined as

$$\mathcal{W}_k = \left\{x_0, x_1, \ldots, x_k;\; u_0^i, u_1^i, \ldots, u_{k-1}^i;\; i = 1, 2;\; k\right\}$$

for complete state observation, and

$$\mathcal{W}_k = \left\{z_0, z_1, \ldots, z_k;\; u_0^i, u_1^i, \ldots, u_{k-1}^i;\; i = 1, 2;\; k\right\}$$

for shared partial state observation.

The available information η_k^i is a subset of the information space $\mathcal{W}_k$. The strategy γ_k^i maps the available information η_k^i into the control, that is $\gamma_k^i(\cdot) : \mathcal{W}_k \to \mathbb{R}^{m_i}$, so that $u_k^i = \gamma_k^i(\eta_k^i)$. For given values of θ_1 and θ_2, an admissible strategy pair $(\gamma^{1^*}, \gamma^{2^*})$ constitutes a *risk-sensitive Nash equilibrium solution* if the following inequalities are satisfied for $\gamma^i \in \Gamma^i$, $i = 1, 2$

$$\begin{aligned} J^{1^*} &\triangleq J^1(\gamma^{1^*}, \gamma^{2^*}\,;\,\theta_1) \leq J^1(\gamma^1, \gamma^{2^*}\,;\,\theta_1) \\ J^{2^*} &\triangleq J^2(\gamma^{1^*}, \gamma^{2^*}\,;\,\theta_2) \leq J^2(\gamma^{1^*}, \gamma^2\,;\,\theta_2) \end{aligned}$$

where $\gamma^i \triangleq \left(\gamma_0^i, \gamma_1^i, \gamma_2^i, \ldots, \gamma_{N-1}^i\right)$. The problem is to determine the *feedback* risk-sensitive Nash equilibrium. Therefore it is assumed that the available information of player i for the complete state observation case is $\eta_k^i = \{x_k\}$ and for the shared partial state observation case this is $\eta_k^i = \{z_k\}$. In this paper, the Nash equilibrium solution is characterized for both the complete and the shared partial observation case.

3 Risk-sensitive Nash equilibrium strategies

In this section two theorems are derived for feedback Nash equilibrium solutions in risk-sensitive dynamic games, one for complete state observation and one for shared partial observation.

First the following proposition for the risk-sensitive dynamic game problem is derived, a similar result for risk-sensitive optimal control was given in Whittle [3, Theorem 1], Jaensch and Speyer [5, Theorem 3.1], and Fan *et al.* [6, Theorem 3.1].

Proposition 1 *Let* $S^i = \Psi_0^i + \theta_i^{-1}\Psi_D$*, where*

$$\begin{aligned}\Psi_D &= \tfrac{1}{2}(x_0-\hat{x}_0)'P_0^{-1}(x_0-\hat{x}_0)+\tfrac{1}{2}\,v_0'(R_0^v)^{-1}v_0\\ &\quad+\tfrac{1}{2}\sum_{k=0}^{N-1}\left\{w_k'(Q_k^w)^{-1}w_k+v_{k+1}'(R_{k+1}^v)^{-1}v_{k+1}\right\}.\end{aligned} \tag{5}$$

If for $i=1,2$ *the function* S^i *is minimized with respect to* $u_k^i,\dots,u_{N-1}^i$ *and extremized with respect to* $x_0,\dots,x_N$ *and* $z_{k+1},\dots,z_N$ *for a given value of* $\mathcal{W}_k$*, where the order of optimization is irrelevant and the infima and extrema are attained and denoted by the value* $(u_k^{1^*},u_k^{2^*})$*, then the strategy pair (or control law) at stage* k *defined by* $\left(\gamma_k^{1^*},\gamma_k^{2^*}\right)$ *is the risk-sensitive Nash equilibrium at stage* k*.*

Proof: In order to find the risk-sensitive Nash equilibrium solution for the problem with cost-functionals of exponential-quadratic form as given in (3), it is shown in [14, Appendix], that the problem reduces to one in which for each player a Dynamic Programming recursion must to be solved. The Dynamic Programming recursion for player i, a recursion in terms of $\Phi_k^i(\mathcal{W}_k)$, is given as

$$\Phi_k^i(\mathcal{W}_k)=\min_{u_k^i}\ \operatorname*{ext}_{z_{k+1}}\ \Phi_{k+1}^i(\mathcal{W}_{k+1}) \tag{6}$$

with boundary condition[1]

$$\Phi_N^i(\mathcal{W}_N)=\operatorname*{ext}_{x_0\dots x_N}\left(\Psi_0^i+\theta_i^{-1}\Psi_D\right). \tag{7}$$

Here, we use the term θ_i-extremizing, abbreviated as "ext", to denote an operation in which one minimizes when $\theta_i \geq 0$ and maximizes when $\theta_i < 0$.

It turns out that for every ith player recursion (6), in which we also have to take into account possible dependence of $u_k^{i^*}$ on x_k, must be satisfied. The minimizing control is denoted by $u_k^{i^*}$. The recursion means that for every ith player the following equation in which $(u_k^j,\dots,u_{N-1}^j)=(\gamma_k^{j^*},\dots,\gamma_{N-1}^{j^*})$ is substituted, must be solved

[1] $\Psi_0^i+\theta_i^{-1}\Psi_D$ is a function of $x_0\dots x_N,\ \ u_0^i\dots u_{N-1}^i,\ \ u_0^j\dots u_{N-1}^j$ and $z_0\dots z_N$.

$$\Phi_k^i(\mathcal{W}_k) = \min_{U_{i,k}^{N-1}} \operatorname*{ext}_{Z_{k+1}^N} \operatorname*{ext}_{X_N} \left(\Psi_0^i + \theta_i^{-1}\Psi_D\right). \tag{8}$$

In general, here we use the notation $X_k = (x_0, x_1, \ldots, x_k)$, $X_k^N = (x_k, x_{k+1}, \ldots, x_N)$ and $U_{i,k}^{N-1} = \left(u_k^i, u_{k+1}^i, \ldots, u_{N-1}^i\right)$, with a similar convention for other variables.

Thus, the function $\Psi_0^i + \theta_i^{-1}\Psi_D$ is minimized with respect to the decisions of player i currently unmade and θ_i-extremized with respect to all quantities currently unobservable. ■

3.1 Decomposition

The recursion equation (8) can be decomposed into a forward recursion P_k^i and a backward recursion F_k^i. Together with Proposition 1, this results in two theorems as stated below. First, the decomposition itself is shown.

At stage k, the observation z_k is available and decision u_k^i has not yet been taken. The past function for the ith player at k is a function of $z_0, \ldots, z_k;\ u_0^i, \ldots, u_{k-1}^i$. The future function for the ith player at k is a function of $z_{k+1}, \ldots, z_N;\ u_k^i, \ldots, u_{N-1}^i$. According to Proposition 1 the order of optimization is irrelevant and thus the decomposition of (8) yields

$$\Phi_k^i(\mathcal{W}_k) = \operatorname*{ext}_{x_k} \left\{ P_k^i(x_k, \mathcal{W}_k) + F_k^i(x_k) \right\} \tag{9}$$

where the functions $P_k^i(x_k, \mathcal{W}_k)$ and $F_k^i(x_k)$ are given below.

The *past function* $P_k^i(x_k, \mathcal{W}_k)$ of player i at stage k is defined as

$$\begin{aligned} P_k^i(x_k, \mathcal{W}_k) \ \triangleq\ & \operatorname*{ext}_{X_{k-1}} \Bigg[\tfrac{1}{2}\,\theta_i^{-1}(x_0 - \hat{x}_0)' \left(\tilde{W}_0^i\right)^{-1}(x_0 - \hat{x}_0) \\ & + \frac{1}{2}\,\theta_i^{-1} v_0' \,(R_0^v)^{-1} v_0 + \frac{1}{2}\sum_{j=0}^{k-1} \Big\{ g_j^i\big(x_j, u_j^i, u_j^l\big) \\ & + \theta_i^{-1}\big(n_j\big(x_{j+1}, u_j^i, u_j^l, x_j\big) + m_{j+1}(z_{j+1}, x_{j+1})\big) \Big\} \Bigg]. \end{aligned} \tag{10}$$

The *future function* $F_k^i(x_k)$ of player i at stage k is defined as

$$\begin{aligned} F_k^i(x_k) \ \triangleq\ & \min_{U_{i,k}^{N-1}} \operatorname*{ext}_{X_{k+1}^N} \operatorname*{ext}_{Z_{k+1}^N} \Bigg[\frac{1}{2} x_N' Q_N x_N + \frac{1}{2}\sum_{j=k+1}^{N} \theta_i^{-1} m_j(z_j, x_j) \\ & + \frac{1}{2}\sum_{j=k}^{N-1} \Big\{ g_j^i\big(x_j, u_j^i, u_j^l\big) + \theta_i^{-1} n_j\big(x_{j+1}, u_j^i, u_j^l, x_j\big) \Big\} \Bigg] \\ =\ & \min_{U_{i,k}^{N-1}} \operatorname*{ext}_{X_{k+1}^N} \Bigg[\frac{1}{2} x_N' Q_N x_N + \frac{1}{2}\sum_{j=k}^{N-1} \Big\{ g_j^i\big(x_j, u_j^i, u_j^l\big) \\ & + \theta_i^{-1} n_j\big(x_{j+1}, u_j^i, u_j^l, x_j\big) \Big\} \Bigg] \end{aligned} \tag{11}$$

with

$$g_k^i\big(x_k,u_k^i,u_k^j\big) \triangleq x_k' Q_k^i x_k + u_k^{i'} R_k^{ii} u_k^i + u_k^{j'} R_k^{ij} u_k^j$$
$$m_k(z_k,x_k) \triangleq (z_k - H_k x_k)' \, (R_k^v)^{-1} \, (z_k - H_k x_k)$$

and

$$\begin{aligned} n_k\big(x_{k+1},u_k^i,u_k^j,x_k\big) &\triangleq \big(x_{k+1} - A_k x_k - B_k^i u_k^i - B_k^j u_k^j\big)' \\ &\quad \times (Q_k^w)^{-1}\big(x_{k+1} - A_k x_k - B_k^i u_k^i - B_k^j u_k^j\big) \end{aligned}$$

for $i = 1,2$ and $j \neq i$.

It follows immediately that $P_k^i(x_k,\mathcal{W}_k)$ satisfies the following *forward recursion*

$$\begin{aligned} P_{k+1}^i(x_{k+1},\mathcal{W}_{k+1}) &= \operatorname*{ext}_{x_k}\Big[P_k^i(x_k,\mathcal{W}_k) + \frac{1}{2}\, g_k^i\big(x_k,u_k^i,u_k^j\big) + \\ &\quad + \frac{1}{2\theta_i}\,\big(n_k\big(x_{k+1},u_k^i,u_k^j,x_k\big) + m_{k+1}(z_k,x_k)\big)\Big] \end{aligned} \tag{12}$$

with initial condition

$$P_0^i(x_0,\mathcal{W}_0) = \frac{1}{2\theta_i}\,(x_0 - \hat{x}_0)' P_0^{-1}(x_0 - \hat{x}_0) + \frac{1}{2\theta_i}\, v_0' \,(R_0^v)^{-1} v_0.$$

Similarly, the function $F_k^i(x_k)$ satisfies the following *backward recursion*

$$\begin{aligned} F_k^i(x_k) &= \min_{u_k^i} \operatorname*{ext}_{x_{k+1}} \Big[F_{k+1}^i(x_{k+1}) \; + \frac{1}{2}\big\{g_k^i\big(x_k,u_k^i,u_k^j\big) \\ &\quad + \theta_i^{-1} n_k\big(x_{k+1},u_k^i,u_k^j,x_k\big)\big\}\Big] \end{aligned} \tag{13}$$

with terminal condition $F_N^i(x_N) = \frac{1}{2} x_N' Q_N^i x_N$.

The following two theorems follow directly from the discussion on the dynamic programming recursion (8) and on the decomposition of function S^i. The theorems concern the recursions needed to determine feedback risk-sensitive Nash equilibria. For the complete observation case, we have the following theorem.

Theorem 1
Let $u_k^{i^*}$ be the minimizing value of u_k^i in the recursion equation $F_k^i(x_k)$ for $i = 1,2$, then the feedback risk-sensitive Nash equilibrium at stage k for the complete observable case is given by $\big(\gamma_k^{1^*}(x_k), \gamma_k^{2^*}(x_k)\big)$. ■

For the shared partial observation case, we have the following theorem which is in fact an alternative form of Proposition 1.

Theorem 2

Let $u_k^{i^*}$ *be the minimizing value of* u_k^i *in the recursion equation* $F_k^i(x_k)$ *for* $i = 1, 2$*, and let vector* $\breve{x}_k^i$ *be the value of* x_k *extremizing* $P_k^i(x_k, \mathcal{W}_k) + F_k^i(x_k)$*. Then, the feedback risk-sensitive Nash equilibrium at stage* k *for the partially observable case is given by* $\left(\gamma_k^{1^*}(\breve{x}_k^1), \gamma_k^{2^*}(\breve{x}_k^2)\right)$. ■

This theorem is in fact an extension of [5, Theorem 3.2], in which only one player is considered.

Proof: Both theorems follow directly from the discussion on the dynamic programming recursion (8) and on the decomposition of function S^i, see (9)-(11). If the state is completely observable, then the feedback risk-sensitive Nash equilibrium solution is $\left(\gamma_k^{1^*}(x_k), \gamma_k^{2^*}(x_k)\right)$ and is determined from the backward recursions $F_k^i(x_k)$ for $i = 1, 2$ in (13). See also [12, Theorem 6.6] in which the risk-neutral case is considered. The feedback risk-sensitive Nash equilibrium solution for the case of shared partial observation is obtained by replacing x_k by $\breve{x}_k^i$, where $\breve{x}_k^i$ is the value of x_k extremizing $P_k^i(x_k, \mathcal{W}_k) + F_k^i(x_k)$. Theorem 2 is, in fact, an extension of, Theorem 3.2 in [5] and [6] where only one player (optimal control problem) is considered. ■

Both theorems are applied to the two-person linear-exponential-quadratic dynamic game as formulated in Section 2, the results of which are presented in the following section.

4 Evaluation of forward and backward recursions

In this section, first the feedback risk-sensitive Nash equilibrium solutions for the complete observation case is determined from the backward recursions $F_k^i(x_k)$ for $i = 1, 2$, see Corollary 1 in Section 4.1. Next, the forward recursions $P_k^i(x_k, \mathcal{W}_k)$ for $i = 1, 2$ are evaluated and finally the feedback risk-sensitive Nash equilibrium solutions for the shared partial observation case is determined, see Corollary 2 in Section 4.2.

4.1 Complete observable risk-sensitive dynamic game

The LEQ dynamic game solution follows as a special case of Theorem 1. First some preliminary notation for Corollary 1 is given. Define

$$\left(M_k^i\right)^{-1} = \left(\tilde{M}_k^i\right)^{-1} + \theta_i\, Q_k^w. \tag{14}$$

Let N_k^i $(i = 1, 2, \quad k = 0, 1, \ldots, N)$ be appropriate dimensional matrices satisfying the set of linear matrix equations

$$\left(R_k^{11} + B_k^{1'} M_{k+1}^1 B_k^1\right) N_k^1 + B_k^{1'} M_{k+1}^1 B_k^2 N_k^2 = B_k^{1'} M_{k+1}^1 A_k \tag{15}$$

$$\left(R_k^{22} + B_k^{2'} M_{k+1}^2 B_k^2\right) N_k^2 + B_k^{2'} M_{k+1}^2 B_k^1 N_k^1 = B_k^{2'} M_{k+1}^2 A_k \tag{16}$$

where the matrices $\tilde{M}_k^i$ are obtained recursively from

$$\tilde{M}_k^1 = F_k' M_{k+1}^1 F_k + N_k^{1'} R_k^{11} N_k^1 + N_k^{2'} R_k^{12} N_k^2 + Q_k^1 \tag{17}$$

$$\tilde{M}_k^2 = F_k' M_{k+1}^2 F_k + N_k^{2'} R_k^{22} N_k^2 + N_k^{1'} R_k^{21} N_k^1 + Q_k^2 \tag{18}$$

with boundary conditions $\tilde{M}_N^1 = Q_N^1$ and $\tilde{M}_N^2 = Q_N^2$ and with

$$F_k = A_k - B_k^1 N_k^1 - B_k^2 N_k^2. \tag{19}$$

The feedback Nash equilibrium solution for the complete observation case is fully determined by the recursions for future functions of both players. The *future function* for player i is quadratic in the state variable

$$F_k^i(x_k) = \frac{1}{2} x_k' \tilde{M}_k^i x_k. \tag{20}$$

If $\theta_i < 0$, then it is necessary that $\tilde{M}_l^i + (\theta_i\, Q_k^w)^{-1} < 0$ for $l > k$.

Corollary 1 *The two-person linear-exponential-quadratic dynamic game with* $Q_k^i \geq 0$ *and* $R_k^{ij} \geq 0$ *(*$j \neq i$*), admits a* unique feedback risk-sensitive Nash equilibrium *solution, if and only if, (15)-(16) admits a unique solution set* $\left\{N_k^i; i=1,2,\ k=0,1,\ldots,N\right\}$*, in which case the equilibrium strategies are linear in the state variable*

$$\gamma_k^{i^*}(x_k) = -N_k^i x_k. \tag{21}$$

For a proof the backwards recursive equations (13) are solved here for $k = N, N-1, \ldots, 0$. This goes similar as in the proof of [12, Corollary 6.1].

Remark 1: The nonnegative definiteness requirements imposed on Q_k^i and R_k^{ij} are sufficient conditions so that the future functions (13) are strictly convex and can be minimized, but they are by no means necessary. Less stringent conditions for which the statement in Corollary 1 is still true is that

$$R_k^{ii} + B_k^{i'} M_{k+1}^i B_k^i > 0, \quad i=1,2,\ k=0,1,\ldots,N.$$

Furthermore, if the set of equations (15) and (16) admits more than one set of solutions, every such set constitutes a feedback risk-sensitive Nash equilibrium solution. See also [12, Remark 6.4].

Remark 2: A precise condition for which the set of equations (15) and (16) admits a unique solution is the invertibility for each $k = 0,1,\ldots,N$ of the matrix in which the iith block is given by $R_k^{ii} + B_k^{i'} M_{k+1}^i B_k^i$ and the ijth block by $B_k^{i'} M_{k+1}^i B_k^j$, where $i,j = 1,2$, $j \neq i$. See also [12, Remark 6.5].

4.2 Partially observable risk-sensitive dynamic game

The partially observable LEQ dynamic game solution follows as a special case of Theorem 2. First some preliminary notation for Corollary 2 is given. Let the vectors $\hat{x}_k^i$ satisfy *risk-modified Kalman filter* recursion equations

$$\begin{aligned} \hat{x}_{k+1}^i &= A_k \hat{x}_k^i + B_k^1 u_k^1 + B_k^2 u_k^2 \\ &\quad + A_k \left[\left(\tilde{W}_k^i\right)^{-1} + H_k'(R_k^v)^{-1} H_k + \theta_i Q_k^i \right]^{-1} \\ &\quad \cdot \left\{ H_k'(R_k^v)^{-1}\left(z_k - H_k \hat{x}_k^i\right) - \theta_i Q_k^i \hat{x}_k^i \right\} \end{aligned} \tag{22}$$

with initial condition $\hat{x}_0^i = \hat{x}_0$.

The matrices $\tilde{W}_k^i$ satisfy the set of *forward matrix Riccati equations*

$$\tilde{W}_{k+1}^i = A_k \left[\left(\tilde{W}_k^i\right)^{-1} + H_k'(R_k^v)^{-1} H_k + \theta_i Q_k^i \right]^{-1} A_k' + Q_k^w \tag{23}$$

with initial condition $\tilde{W}_0^i = P_0$.

For the partial observation case we must also solve the recursions for the past functions of both players. Solving these recursions yields that the *past function* for player i has the form

$$\begin{aligned} P_k^i(x_k, \mathcal{W}_k) &= \frac{1}{2\theta_i}\left(x_k - \hat{x}_k^i\right)' \left(\tilde{W}_k^i\right)^{-1}\left(x_k - \hat{x}_k^i\right) \\ &+ \frac{1}{2\theta_i}(z_k - Hx_k)' \left(R_k^v\right)^{-1}(z_k - Hx_k) + \ldots \end{aligned} \tag{24}$$

where ... indicates terms independent of x_k.

If $\theta_i < 0$, then it is necessary that $\tilde{W}_l^i + (\theta_i\, Q_k^i)^{-1} < 0$ for $l \leq k$.

Corollary 2 *The two-person linear-exponential-quadratic dynamic game and with* partial observations without delay, *admits a* unique feedback risk-sensitive Nash equilibrium *solution, if and only if, (15) and (16) admits a unique solution set* $\left\{N_k^i; i = 1, 2,\ k = 0, 1, \ldots, N\right\}$*, in which case the equilibrium strategies are*

$$\gamma_k^{i^*}\left(\breve{x}_k^i\right) = -N_k^i \breve{x}_k^i \tag{25}$$

where

$$\begin{aligned} \breve{x}_k^i &= \left[I + \theta_i \tilde{W}_k^i \tilde{M}_k^i + \tilde{W}_k^i H_k' (R_k^v)^{-1} H_k\right]^{-1} \\ &\cdot \left[\hat{x}_k^i + \tilde{W}_k^i H_k' (R_k^v)^{-1} z_k\right] \end{aligned} \tag{26}$$

Proof: As shown in Theorem 2, for the partial observable case, the feedback risk-sensitive Nash equilibrium is $\left(\gamma_k^{1^*}(\breve{x}_k^1), \gamma_k^{2^*}(\breve{x}_k^2)\right)$, where $\breve{x}_k^1$ and $\breve{x}_k^2$ are determined by extremizing $P_k^i(x_k, \mathcal{W}_k) + F_k^i(x_k)$ with respect to x_k for $i = 1$ and $i = 2$ respectively. Substitution of the results of (20) and (24) yields

$$\begin{aligned} P_k^i(x_k, \mathcal{W}_k) + F_k^i(x_k) &= \tfrac{1}{2}\,\theta_i^{-1}\Big[\left(x_k - \hat{x}_k^i\right)' \left(\tilde{W}_k^i\right)^{-1}\left(x_k - \hat{x}_k^i\right) \\ &\quad + (z_k - H_k x_k)' \left(R_k^v\right)^{-1}(z_k - H_k x_k) \\ &\quad + \theta_i x_k' \tilde{M}_k^i x_k + \ldots\Big] \\ &= \tfrac{1}{2}\,\theta_i^{-1}\Big[x_k' \left\{\left(\tilde{W}_k^i\right)^{-1} + \theta_i \tilde{M}_k^i + H_k'(R_k^v)^{-1} H_k\right\} x_k \\ &\quad - 2x_k' \left\{\left(\tilde{W}_k^i\right)^{-1}\hat{x}_k + H_k'(R_k^v)^{-1} z_k\right\} + \ldots\Big]. \end{aligned}$$

Extremizing with respect to x_k yields $\breve{x}_k^i$ as in (26). From Corollary 1 we know that $\gamma_k^{i^*}(x_k) = -N_k^i x_k$ and thus the result follows. ■

Remark 3: Equations (22) and (23) can be considered as a risk-modified Kalman filter for player i. The vector $\hat{x}_k^i$ denotes the estimate of x_k which θ_i-extremizes the past function of player i at k. If in these equations the matrices Q_k^i are zero for $k = 0, 1, \ldots, N-1$, then they reduce to one set of equations for both players which is exactly the Kalman filter.

5 Conclusion

Whereas risk-sensitivity in control theory is well studied in literature, risk-sensitivity in dynamic games has not been considered before in literature. In this paper results have been presented for risk-sensitive dynamic games. A two-person linear-quadratic dynamic game in discrete time has been considered with exponential cost criteria, where the players have different risk-sensitivity parameters. Nash solutions have been derived under the hypothesis that the strategies at the current time are functions of the observation history up to the *current* time. This is similar as the hypothesis in Jaensch and Speyer [5] and Fan *et al.* [6] for risk-sensitive control. This hypothesis is an extension of the results of Whittle [3] and [4], who assumes that the control at the current time is a function of the observation history up to the *previous* time.

We derived two theorems on feedback risk-sensitive Nash equilibria and it is shown that the derivation leading to these theorems is based on known results of both dynamic game theory and of risk-sensitive control theory.

In the discrete-time risk-sensitive game, as presented in this paper, the results for the complete observable risk-sensitive dynamic game are a simple extension of similar results in both dynamic game theory and risk-sensitive control theory. Even though the results for the complete observation case are not very difficult to derive, they are important results since they are the basis for the shared partial observable case. The results for this latter case are less straightforward. It is shown in this paper that the feedback Nash equilibrium for the partially observable stochastic problem can be constructed from the feedback risk-sensitive Nash equilibrium design for the complete observable stochastic problem and from two risk-modified Kalman filters. It turned out that if there are no intermediate state costs these two filters are identical and exactly the well-known Kalman filter, which is analogous to the risk-sensitive control situation. However, for the more general case with costs on the intermediate states, it turned out that for each player one gets a risk-modified Kalman filter.

Acknowledgement

The author acknowledges the helpful comments of the reviewers and of H.A.P. Blom from National Aerospace Laboratory (NLR).

References

1. D.H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, Apr. 1973.
2. J.L. Speyer, J. Deyst, and D.H. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," *IEEE Transactions on Automatic Control*, vol. 19, no. 4, pp. 358–366, Aug. 1974.
3. P. Whittle, "Risk-sensitive linear quadratic gaussian control," *Adv. Appl. Prob.*, vol. 13, pp. 764–777, 1981.
4. P. Whittle, *Risk-sensitive optimal control*, John Wiley & Sons, Inc., 1990.
5. C.R. Jaensch and J.L. Speyer, "Centralized and decentralized stochastic control problems with an exponential cost criterion," in *Proceedings of the 27th Conference on Decision and Control*, Dec. 1988, pp. 286–291, Austin, Texas.
6. C.-H. Fan, C.R. Jaensch, and J.L. Speyer, "Centralized and decentralized solutions of the linear-exponential-gaussian problem," *IEEE Transactions on Automatic Control*, vol. 39, no. 10, pp. 1986–2003, Oct. 1994.
7. A. Bensoussan and J.H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of-integral performance index," *SIAM Journal on Control and Optimization*, vol. 23, no. 4, pp. 599–613, July 1985.
8. A. Bensoussan, *Stochastic control of observable systems*, Cambridge University Press, 1992.
9. M.R. James, J.S. Baras, and R.J. Elliott, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 4, pp. 780–792, Apr. 1994.
10. I.B. Collings, M.R. James, and J.B. Moore, "An information-state approach to linear/risk-sensitive/quadratic/ Gaussian Control," in *Proceedings of the 33rd Conference on Decision and Control*, Dec. 1994, pp. 3802–3807, Florida, USA.
11. M.R. James and J.S. Baras, "Partially observed differential games, infinite-dimensional Hamilton-Jacobi-Isaacs equations, and nonlinear H_∞ control," *SIAM Journal on Control and Optimization*, vol. 34, no. 4, pp. 1342–1364, July 1996.
12. T.S. Başar and G.J. Olsder, *Dynamic noncooperative game theory*, Second Edition, Academic Plenum Press, 1995.
13. M.B. Klompstra, "Nash equilibria in risk-sensitive dynamic games," in *Proceedings of the 1995 American Control Conference*, June 1995, pp. 2458–2462, Seattle, Washington.
14. M.B. Klompstra, "Nash equilibria in risk-sensitive dynamic games," 1995, NLR TP 95101.