A REPORT ON
`NISOCSol`
AN ALGORITHM FOR APPROXIMATING MARKOVIAN EQUILIBRIA IN DYNAMIC GAMES WITH COUPLED-CONSTRAINTS

JACEK B. KRAWCZYK & JEFFREY D. AZZATO

ABSTRACT. In this report, we outline a method for approximating a Markovian (or feedback-Nash) equilibrium of a dynamic game, possibly subject to coupled-constraints. We treat such a game as a "multiple" optimal control problem. A method for approximating a solution to a given optimal control problem via backward induction on Markov chains was developed in [Kra01]. A Markovian equilibrium may be obtained numerically by adapting this backward induction approach to a stage Nikaidô-Isoda function (described in [KZ06]).

2006 Working Paper
School of Economics and Finance

Correspondence should be addressed to:

Jacek B. Krawczyk. Faculty of Commerce & Administration, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand. Fax: +64-4-4712200 Email: `J.Krawczykatvuw.ac.nz` Webpage: http://www.vuw.ac.nz/staff/jacek_krawczyk

Contents

1. Introduction

This report draws from [Kra06]. It also provides some technical remarks on the use of the new routines constituting `NISOCSol`.

In [KT06] a finite horizon *unconstrained* dynamic game was solved through backward induction. At each stage, a Markovian (feedback-Nash) equilibrium was computed as a solution to a system of the Bellman equations where each Bellman equation characterised a player's value function. Due to the lack of constraints, the computations were rather straightforward albeit sensitive to the horizon length.

The backward induction technique will be also applied in this paper. However, instead of simultaneous maximisation of the players' value functions (as in [KT06]), constrained min-maximisation of the *stage* Nikaidô-Isoda function (see [NI55]) will be performed with each player's value function characterised by their Bellman equation. This method appears satisfactory in the *constrained dynamic game* context.

In this paper, some background on Markovian (feedback-Nash) solutions is given in Section 2. The technical remarks of Section 3 explain how a feedback solution to a dynamic game can be obtained.

2. Markovian (feedback-Nash) equilibrium

2.1. Preliminaries.

Denote the game value (Bellman function) for player i from state $\mathbf{x}^{(\tau)}$ at time τ by the *payoff-to-go* (or *continuation payoff*) function

$$F_i^{(\tau)}\big(\mathbf{x}^{(\tau)}; \mathbf{u}\big) := \sum_{t=\tau}^{T-1} \rho^t \left(\phi_i\big(\mathbf{x}^{(t)}\big) - c_{3,i}\big(u_i^{(t)}\big)^2 \right) + \rho^T k\big(x_i^{(T)}\big) \tag{1}$$

when the players are using strategy $\mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3})$. Here, $\phi_i(\cdot)$ is the static (concave) payoff for player i contingent on full utilisation of production capacity, $c_{3,i}(\cdot)^2$ is a quadratic investment cost,[†] $k(\cdot)$ is the capacity's scrap value function and ρ the discount factor.

Definition 1. *We say that* strategies $\mathbf{u_1^*}$, $\mathbf{u_2^*}$ *and* $\mathbf{u_3^*}$ *constitute a* feedback-Nash (or Markovian subgame-perfect) equilibrium *on* $\mathbb{R}_+^3$ *if*

1. *they are admissible at every* $\mathbf{x}^{(t)} \in \mathbb{R}_+^3$, *and*
2. *for every history* $\mathbf{x}^{(\tau)}, \mathbf{x}^{(\tau+1)}, \ldots, \mathbf{x}^{(t)}$; *and for every admissible* $(\mathbf{u_1}, \mathbf{u_{-1}^*})$, $(\mathbf{u_2}, \mathbf{u_{-2}^*})$ *and* $(\mathbf{u_3}, \mathbf{u_{-3}^*})$, *the inequalities*

$$F_i^{(t)}(\mathbf{x}^{(t)}; \mathbf{u}^*) \geqslant F_i^{(t)}(\mathbf{x}^{(t)}; \mathbf{u}_i, \mathbf{u}_{-i}^*) \qquad (i = 1, 2, 3) \tag{2}$$

hold.

Here,

$$u_i^{(t)} = \mathbf{u}_i\big(\mathbf{x}^{(t)}, t\big) \tag{3}$$

[†]Payoff concavity and a quadratic investment cost are not crucial for the method—weak convex-concavity of the Nikaidô-Isoda function suffices.

1

for each i, and each strategy is a map

$$
\begin{array}{ccc}
& \mathbf{u}_i & \\
\mathbf{R}_+^3 \times \mathcal{T} & \longrightarrow & \mathbf{R} \qquad (i = 1, 2, 3) \\
\cup & & \cup \\
(\mathbf{x}^{(t)}, t) & \longmapsto & u_i^{(t)} \qquad (t \in \mathcal{T} := \{0, \dots, T-1\})
\end{array}
\tag{4}
$$

The symbol $\mathbf{u}_{-i}$ will be used to represent the other players' strategies.

So, we wish to solve problem (2) for strategies (3).

2.2. A solution method.

2.2.1. *Stage-dependent Nikaidô-Isoda functions.* A coupled-constraint game is a difficult construct. A solution concept suitable for this type of game is a "generalised" Nash equilibrium (see [AD54, CKK04, HK97, HP07, Kra05b, KU00, McK59, Ros65]). A numerical approach may be necessary to obtain such a general solution—this was the case for a static game solved in [KZ06], and also for the open-loop game of [Kra05b].

For Markovian games, a solution consists of strategies understood as maps rather than time profiles. The latter may also be obtained for Markovian games as *realisations* of strategies in form of (4).

SOCSol4L (see [AK06]; also [Kra01, Kra05a]) is a suite of MATLAB® routines capable of approximating optimal strategies to intertemporal decision problems having a single planner.[1] Hence, SOCSol4L can rather straightforwardly be used to approximate the Pareto optimal strategies.

Here we explain the usage of SOCSol4L to compute feedback-Nash equilibrium strategies. In this case, the solution strategies will be maximisers of the Nikado-Isoda functions formulated *stage-by-stage* from the players' value functions $F_i(\mathbf{x}^{(\tau)}; \mathbf{u})$ (see (1)) and iterated backward in time.[2] The equilibrium solution (also, equilibrium existence, uniqueness and the algorithm convergence) will be contingent upon the *stage* Nikado-Isoda functions' weak convex-concavity and the constraint set's convexity. These can be verified.

In practical terms, to solve a dynamic game (2) in admissible feedback strategies (3) we need to combine results of the Convergence Theorem (see e.g., [Kra05b, KU00]) with the Bellman optimality principle. The use of optimality principle implies that *stage games* will be solved backward in time. Hence, the Convergence Theorem will be applied at each stage (backward in time) with the role of a "one-off" (or static) utility being played by the *utility-to-go* $F_i^{(t)}(\mathbf{x}^{(t)}; \mathbf{u})$.

2.2.2. *Stage Games.* Let $\mathcal{U}_{(-i)}^{(t)}$ denote the i-th player's strategy set at time t. In a coupled-constraint context, this set depends on the other players' strategies (indicated

---

[1] I.e., SOCSol4L solves optimal control problems.
[2] The dynamic game at hand is finite horizon.

by the subscript $(-i)$). It follows from Section 2.1 that

$$\mathcal{U}^{(t)}_{(-1)} \cup \mathcal{U}^{(t)}_{(-2)} \cup \mathcal{U}^{(t)}_{(-3)} = \mathbf{U}^{(t)} \subseteq \mathbb{R}^3.$$

Assume that $f^{(i)}(\mathbf{x}^{(t)}, u^{(t)}_i)$ and $k(x^{(T)})$ are concave, with

$$f^{(i)}(\mathbf{x}^{(t)}, u^{(t)}_i) := \phi_i(\mathbf{x}^{(t)}) - c_{3,i}(u^{(t)}_i)^2 \tag{5}$$

and

$$f^{(i)}_T(x^{(T)}_i) := k_i(x^{(T)}_i). \tag{6}$$

Define $V^{(t)}_i(\mathbf{x}^{(t)})$, an *optimal value function* for player i at stage t, by

$$V^{(t)}_i(\mathbf{x}^{(t)}) := \max_{u^{(t)}_i \in \mathcal{U}^{(t)}_{(-i)}} F^{(t)}_i\left(\mathbf{x}^{(t)}; u^{(t)}_i, \mathbf{u}_{-i}(\mathbf{x}^{(t)}, t)\right) \tag{7}$$

for each $t = 0, \ldots, T - 1$; and

$$V^{(i)}_T(\mathbf{x}^{(T)}, T) := f^{(i)}_T(x^T_i),$$

where

$$F^{(t)}_i\left(\mathbf{x}^{(t)}; u^{(t)}_i, \mathbf{u}_{-i}(\mathbf{x}^{(t)}, t)\right) := f^{(i)}(\mathbf{x}^{(t)}, u^{(t)}_i) + \varrho V^{(t+1)}_i(\mathbf{x}^{(t+1)}). \tag{8}$$

The following theorem[3] establishes a basis for using dynamic programming as a computational technique for Markovian (feedback-Nash) equilibria in dynamic games.

Theorem 1. *If there exist value functions $V^{(t)}_i(\mathbf{x}^{(t)})$ and strategies $\mathbf{u}_i(\mathbf{x}^{(t)}, t)$ which satisfy equations (7)–(8) for each $t = T - 1, \ldots, 1, 0$; where $\mathbf{x}^{(t)}$ is a vector of state variables observable at t, then the strategy*

$$\mathbf{u}^* := (\mathbf{u}^*_i, \mathbf{u}^*_{-i})$$

constitutes a Markovian (feedback-Nash) equilibrium for the dynamic game under a feedback information pattern.[4] Moreover, the value functions $V^{(t)}_i(\mathbf{x}^{(t)})$ represent player i's optimal utility for the game starting at $(\mathbf{x}^{(t)}, t)$. In particular,

$$V^{(0)}_i(\mathbf{x}^{(0)}) = \Phi_i(\mathbf{x}^{(0)}; \mathbf{u}^*).$$

If each stage game $(F^{(\cdot)}_i, F^{(\cdot)}_{-i})$ is concave, then it makes sense to ask whether the stage games have unique equilibria. We observe that at least the last game "played" to maximise the utility-to-go functions $F^{(T-1)}_i$ and $F^{(T-1)}_{-i}$ is concave. This is necessary for the game's solution. Weak convex-concavity of the Nikaidô-Isoda functions for stage games (or diagonal strict concavity of these games) can be established stage-wise backward in time to demonstrate solution uniqueness.

Notice that in equilibrium,

$$V^{(t)}_i(\mathbf{x}^{(t)}) = F^{(t)}_i(\mathbf{x}^{(t)}; \mathbf{u}^*)$$

---

[3]"Standard" in dynamic games, see [BO82, Theorem 6.6].
[4]Such an equilibrium is subgame perfect.

for each $i = 1, 2, 3$ and each $t = T - 1, \ldots, 1, 0$. Hence, if player i played strategy $\mathbf{v}_i$ while the rest were playing $\mathbf{u}^*$, then

$$F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{v}_i | \mathbf{u}^*\big) \leqslant F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{u}^*\big)$$

for each $i = 1, 2, 3$ and each $t = T - 1, \ldots, 1, 0$. However, if $\mathbf{u}$ is not an equilibrium, then player i might improve his or her utility-to-go by playing the strategy $\mathbf{v}_i$ (given that the other players play $\mathbf{u}$). Consequently, the expression

$$\psi_i^{(t)}(\mathbf{u}, \mathbf{v}) = F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{v}_i | \mathbf{u}\big) - F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{u}\big) \qquad t = T - 1, \ldots, 0$$

might be positive. This suggests that we can define a *stage* Nikaidô-Isoda function whose maximisation will reveal the equilibrium $\mathbf{u}^*$ under certain concavity conditions of $F_i^{(t)}(\cdot\,; \mathbf{u})$, where $i = 1, 2, 3$ and $t = T - 1, \ldots, 1, 0$.

Define the *stage* Nikadô-Isoda function at state $\mathbf{x}^{(t)}$ by

$$\Psi^{\mathbf{x}^{(t)}} \colon \mathbf{U}^{(t)} \times \mathbf{U}^{(t)} \to \mathbb{R} \colon (\mathbf{u}, \mathbf{v}) \mapsto \sum_{i=1}^{3} \left(F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{v}_i | \mathbf{u}\big) - F_i^{(t)}\big(\mathbf{x}^{(t)}; \mathbf{u}\big) \right) \qquad (9)$$

for each $t = T - 1, \ldots, 1, 0$. Obviously,

$$\Psi^{\mathbf{x}^{(t)}}(\mathbf{u}, \mathbf{u}) \equiv 0 \qquad (10)$$

for every $\mathbf{u} \in \mathbf{U}^{(t)}$.

Each summand in (9) can be thought of as the *improvement* in the player's value function that he (or she) will receive by changing his (or her) strategy from $\mathbf{u}_i$ to $\mathbf{v}_i$ while all other players continue to play $\mathbf{u}_{-i}$. The Nikaidô-Isoda function thus represents the sum of these value function improvements. Note that its *maximum* for a given $\mathbf{u}$ is always non-negative as a consequence of (10). Moreover, (9) is non-positive everywhere when either $\mathbf{u}$ or $\mathbf{v}$ is a Markovian equilibrium strategy, for in an equilibrium situation no player can unilaterally improve their payoff, and so in this case each summand is at most zero.

From here, we conclude that when the Nikaidô-Isoda function (9) cannot be made (significantly) positive at each stage for a given $\mathbf{u}$, we have (approximately) reached the Markovian equilibrium point. This observation is useful in constructing a termination condition for our algorithm: choose an $\varepsilon > 0$ such that the equilibrium has been reached to a sufficient degree of precision when $\max_{\mathbf{v} \in \mathbf{U}} \Psi^{\mathbf{x}^{(t)}}(\mathbf{u}^s, \mathbf{v}) < \varepsilon$, where $\mathbf{u}^s \in \mathbf{U}$ is computed at the current iteration s.

More formally, we shall compute a "normalised" Markovian equilibrium which will be a Nash equilibrium under weak convex-concavity of $\Psi^{\mathbf{x}^{(t)}}(\mathbf{u}, \mathbf{v})$.

Definition 2. *We call an admissible strategy* $\mathbf{u}^* \in \mathbf{U}$ *a* Nash normalised Markovian equilibrium *if*

$$\max_{\mathbf{v} \in \mathbf{U}} \Psi^{\mathbf{x}^{(t)}}(\mathbf{u}^*, \mathbf{v}) = 0.$$

A Nash normalised Markovian equilibrium $\mathbf{u}^*$ is a unique feedback-Nash equilibrium if

(1) $\Psi^{\mathbf{x}^{(t)}}(\mathbf{u}, \mathbf{v})$ is weakly convex-concave for every t, or

(2) $\sum_{i=1}^{n} F_i^{(t)}(\mathbf{x}^{(t)}; \mathbf{u})$ is diagonally strictly concave for every t (see [KT06], where this feature was examined in a dynamic game).

3. NUMERICAL RESULTS THROUGH NISOCSOL

This section should be read in conjunction with [AK06], where the calls to SOCSol4L are explained. Here, we explain the differences that the user should be familiar with in order to be able to call NISOCSol.

3.1. **State and time discretisation.** These are specified as for SOCSol4L.

3.2. **Options.** In this version of NISOCSol, it is assumed that each player has only one dimension of control. Consequently, the control dimension should be set to the number of players.

3.3. **User defined .m functions.** The utility functions for NISOCSol return values for the i-th player, where i is passed as an argument following those standard for SOCSol4L. If the game is defined as discrete-time with times $0, \Delta, 2\Delta, \ldots, T$; where T is a multiple of Δ, then the following should also be observed:

- TimeStep = ones$(1, \frac{T}{\Delta}) / \frac{T}{\Delta}$;
- The return value of the instantaneous cost function needs to be multplied by $\frac{T}{\Delta}$.
- The return value of the motion (or "delta") function needs to be multplied by $\frac{T}{\Delta}$.

With these modifications a typical NISOCSol call has the following form.

```
NISOCSol('DeltaFunctionFile', 'InstantaneousCostFunctionFile',
    'TerminalStateFunctionFile', StateLB, StateUB, StateStep,
    TimeStep, 'ProblemFile', Options, InitialControlValue, A,
    b, Aeq, beq, ControlLB, ControlUB,
    'UserConstraintFunctionFile');
```

3.4. **A test: a static equilibrium obtained as feedback-Nash.** The static solution of the River Basin Pollution game has been cited in many publications (e.g., [Kra05b, KU00, KZ06]). The unique coupled-constraint equilibrium is

$$\mathbf{x}^* = (21.14, 16.03, 2.73).$$

We will use this problem to test NISOCSol.

Presumably, with no investment cost and no depreciation, agents should move to this equilibrium from any "initial condition" in one step. So we set the following parameter values: $T = 1$, $c_{3,i} \equiv 0$, $\rho = 1$, $k_i(\cdot) \equiv \phi_i(\mathbf{x})$. Furthermore, we assume the system dynamics

$$x_i^{(t+1)} = (1 - \mu_i)x_i^{(t)} + u_i^{(t)}.$$

As there is no depreciation, $\mu_i \equiv 0$. We assume that
$$\phi_i(\mathbf{x}) = \underbrace{[d_1 - d_2(x_1 + x_2 + x_3)]x_i}_{\text{Revenue}} - \underbrace{(c_{1,i} + c_{2,i}x_i)x_i}_{\text{Cost}}.$$

The economic constants d_1 and d_2 determine an inverse demand law. We assign them the values 3 and 0.01 respectively. The values of the cost function coefficients $c_{1,i}$ and $c_{2,i}$ and the other model parameters are given in Table 1 below.

Player (i)	$c_{1,i}$	$c_{2,i}$	e_i	$\delta_{i,1}$	$\delta_{i,2}$
1	0.10	0.01	0.50	6.5	4.583
2	0.12	0.05	0.25	5.0	6.250
3	0.15	0.01	0.75	5.5	3.750

TABLE 1. Constants for the River Basin Pollution game.

The fourth column of Table 1 gives the emission parameters, while the fifth gives the pollution transportation and decay parameters that govern contamination levels at location $\ell = 1, 2$. The local authority imposes the pollution constraints

$$q_\ell(\mathbf{x}) = \sum_{i=1}^{3} \delta_{i,\ell} e_i x_i \leqslant K_\ell \qquad \ell = 1, 2. \tag{11}$$

We take $K_1 = K_2 = 100$.

Given this information and recalling the comments in Sections 3.1, 3.2 and 3.3, the functions Delta(u, x, t), Costi(u, x, t, i), Termi(x, i) and Constraints(u, x, ts) were defined after [AK06].

The programme's output is the null strategy for $\mathbf{x}^{(0)} = \mathbf{x}^*$ —see Figure 1, which shows strategies at $t = 0$. Players' actions $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are each plotted (from left to right) as functions of the third, second and first state variables. We see that the feedback strategies intersect with the horizontal axis (i.e., no action) for the initial condition $\mathbf{x}^*$.

The strategy realisations are the state and action time profiles shown in Figure 2. The solid lines (which are constant in time) demonstrate that no changes are made when given $\mathbf{x}^{(0)} = \mathbf{x}^*$. The dash-dotted lines show that the equilibrium point is achieved from an arbitrary initial state (here, $(22, 15, 4)$).

Lagrange multiplier *rules* have also been computed. To ensure that the standards (11) are obeyed, the government needs to inform the players what taxes would be collected should the pollution levels be exceeded. Under the feedback information pattern, the value of the Lagrange multipliers is a function of states $(x_1^{(t)}, x_2^{(t)}, x_3^{(t)})$ (here, $t = 0$). See Figure 3.
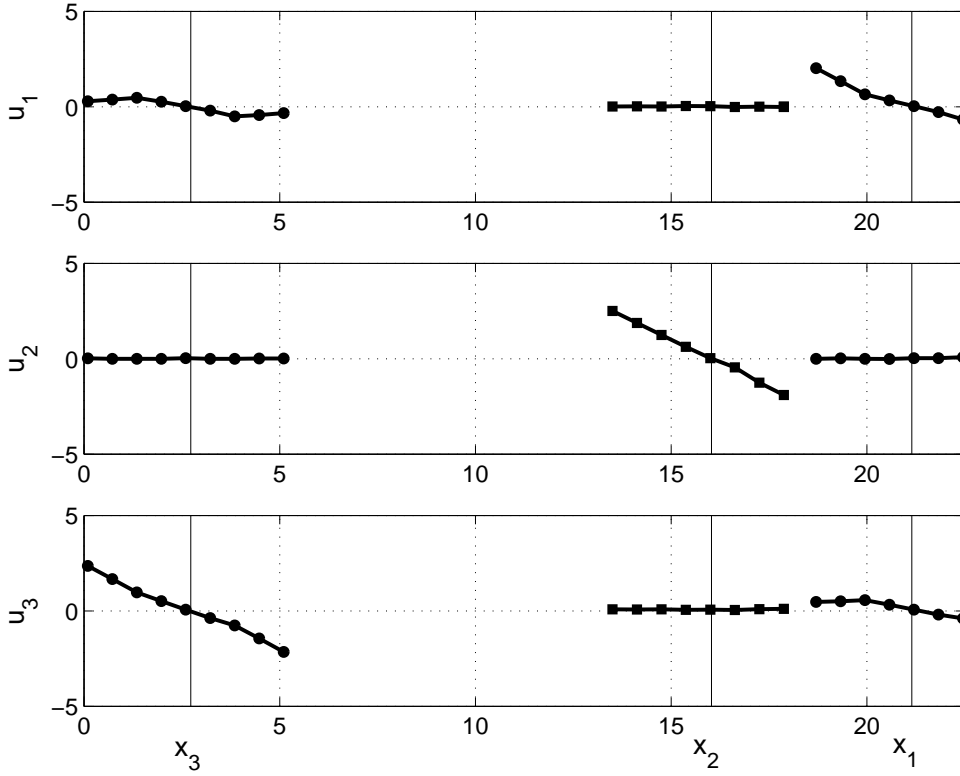
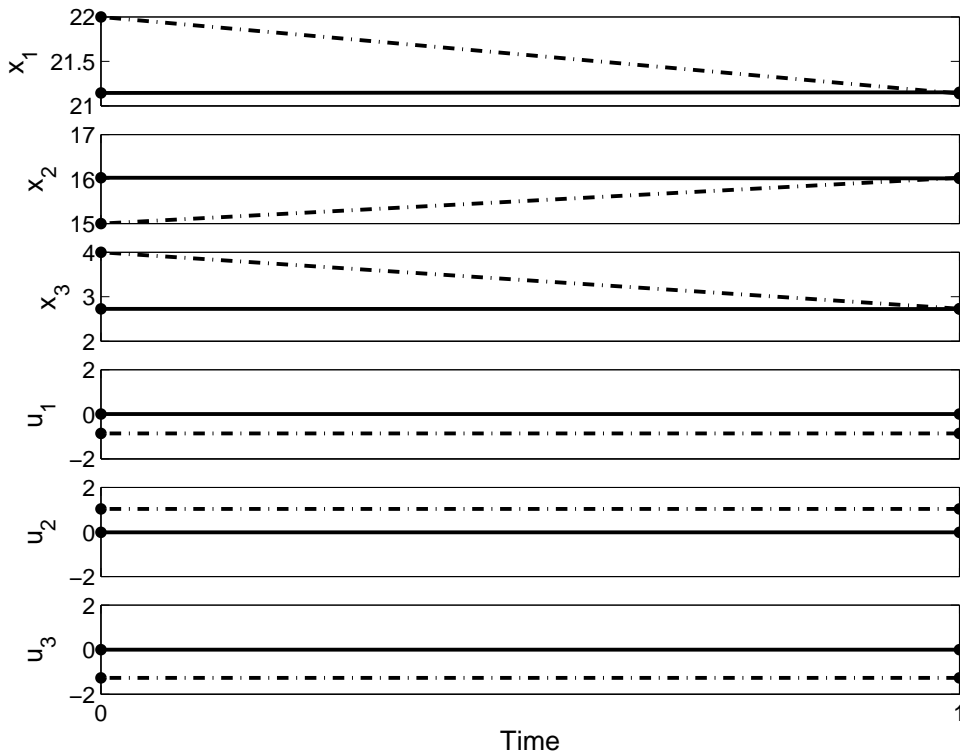FIGURE 1. One-step Markovian equilibrium strategies.

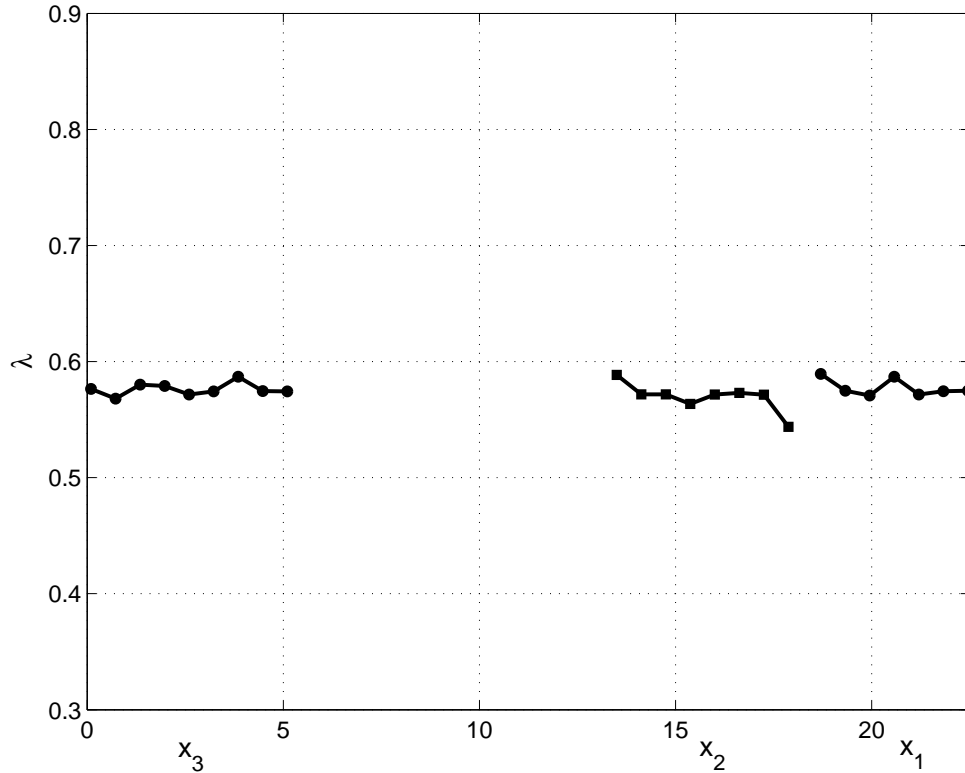FIGURE 2. One-step Markovian equilibrium states and actions.

FIGURE 3. Lagrange multipliers at $t = 0$ as a function of state.

4. CONCLUDING REMARKS

This report outlines a set of useful machinery for the study dynamic games (possibly) subject to coupled-constraints (typical of environmental management problems). In particular, we have seen how the Nikaidô-Isoda function can be utilised to find a coupled-constraint equilibrium through an optimisation approach.

REFERENCES

[AD54]   K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3):265–290, 1954.

[AK06]   Jeffrey D. Azzato and Jacek B. Krawczyk. SOCSol4L: An improved MATLAB® package for approximating the solution to a continuous-time stochastic optimal control problem. Working paper, School of Ecnomics and Finance, Victoria University of Wellington, Dec 2006. Available at http://mpra.ub.uni-muenchen.de/10015/ on 26/08/2008.

[BO82]   T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Academic Press, New York, 1982.

[CKK04] Javier Contreras, Matthias Klusch, and Jacek B. Krawczyk. Numerical solutions to Nash–Cournot equilibria in coupled constraint electricity markets. *IEEE T. Power Syst.*, 19(1):195–206, Feb 2004.

[HK97]   Alain Haurie and Jacek B. Krawczyk. Optimal charges on river effluent from lumped and distributed sources. *Environ. Model. Assess.*, 2(3):177–189, Oct 1997. DOI: 10.1023/A:1019049008557.

[HP07]   Benjamin F. Hobbs and J.-S. Pang. Nash–Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. 55(1):113–127, Jan/Feb 2007. DOI: 10.1287/opre.1060.0342.

[Kra01]   Jacek B. Krawczyk. A Markovian approximated solution to a portfolio management problem. *ITEM.*, 1(1), 2001. Available at http://www.item.woiz.polsl.pl/issue/journal1.htm on 26/08/2008.

[Kra05a] J. B. Krawczyk. Numerical solutions to lump-sum pension problems that can yield left-skewed fund return distributions. In Christophe Deissenburg and Richard F. Hartl, editors, *Optimal Control and Dynamic Games*, number 7 in Advances in Computational Management Science, chapter 10, pages 155–176. Springer, New York, 2005.

[Kra05b] Jacek B. Krawczyk. Coupled constraint Nash equilibria in environmental games. *Resource Energy Econ.*, 27(2):157–181, Jun 2005. DOI: 10.1016/j.reseneeco.2004.08.001.

[Kra06]   Jacek B. Krawczyk. Coupled constraint Markovian equilibria in dynamic games of compliance. Seminar at the Kyoto University Institute of Economic Research, 30 Nov 2006.

[KT06]   J. B. Krawczyk and M. Tidball. A discrete-time dynamic game of seasonal water allocation. *J. Optim. Theory Appl.*, 128(2):411–429, Feb 2006. DOI: 10.1007/s10957-006-9020-0.

[KU00]   Jacek B. Krawczyk and Stanislav Uryasev. Relaxation algorithms to find Nash equilibria with economic applications. *Environ. Model. Assess.*, 5(1):63–73, Jan 2000. DOI: 10.1023/A:1019097208499.

[KZ06]   Jacek B. Krawczyk and James Zuccollo. NIRA-3 An improved MATLAB package for finding Nash equilibria in infinite games. Working paper, School of Economics and Finance, Victoria University of Wellington, Dec 2006. Available at http://mpra.ub.uni-muenchen.de/1119/ on 26/08/2008.

[McK59]   Lionel W. McKenzie. On the existence of general equilibrium for a competitive market. *Econometrica*, 27(1):54–71, Jan 1959.

[NI55]   Hukukane Nikaidô and Kazuo Isoda. Note on noncooperative games. *Pac. J. Math.*, 5(1):807–815, 1955.

[Ros65]   J. B. Rosen. Existence and uniqueness of equilibrium points for concave *n*-person games. *Econometrica*, 33(3):520–534, Jul 1965.