NONLINEAR MARKOV GAMES*

Vassili N. Kolokoltsov†

Abstract

I am going to put forward a program of the analysis of a new class of stochastic games that I call nonlinear Markov games, as they arise as a (competitive) controlled version of nonlinear Markov processes (an emerging field of intensive research, see e.g. [3], [4],[5]). This class of games can model a variety of situation for economics and epidemics, statistical physics, and pursuit - evasion processes. The discussion below is mostly taken from the author's monograph in preparation [1]. I shall start by introducing the (yet not very well known) concept of nonlinear Markov chains.

1 Nonlinear Markov chains

A discrete in time and space *nonlinear Markov semigroup* Φ^k, $k \in \mathbf{N}$, is specified by an arbitrary continuous mapping $\Phi : \Sigma_n \to \Sigma_n$, where the simplex

$$\Sigma_n = \{\mu = (\mu_1, ..., \mu_n) \in \mathbf{R}_+^n : \sum_{i=1}^n \mu_i = 1\}$$

represents of course the set of probability laws on the finite state space $\{1, ..., .n\}$. For a measure $\mu \in \Sigma_n$ the family $\mu^k = \Phi^k \mu$ can be considered as an evolution of measures on $\{1, ..., .n\}$. But it does not yet define a random process (finite-dimensional distributions are not specified). In order to get a process one has to choose a *stochastic representation* for Φ, i.e. to write it down in the form

$$\Phi(\mu) = \{\Phi_j(\mu)\}_{j=1}^n = \{\sum_{i=1}^n P_{ij}(\mu)\mu_i\}_{i=1}^n, \tag{1.1}$$

where $P_{ij}(\mu)$ is a family of stochastic matrices ($\sum_{j=1}^d P_{ij}(\mu) = 1$ for all i), depending on μ (nonlinearity!), whose elements specify the *nonlinear transition probabilities*. For any given $\Phi : \Sigma_n \mapsto \Sigma_n$ a representation (1.1) exists, but is not unique. For instance, there exists a unique representation (1.1) with an additional condition that all matrices $P_{ij}(\mu) = \tilde{P}_{ij}(\mu)$ are one dimensional:

$$\tilde{P}_{ij}(\mu) = \Phi_j(\mu), \quad i, j = 1, ..., n. \tag{1.2}$$

*Talk given at Adversarial and Stochastic Elements in Autonomous Systems Workshop, 22-24 March 2009, at FDIC, 3501 Fairfax Drive, Arlington, VA 22226.

†Department of Statistics, University of Warwick, Coventry CV4 7AL, UK. Email: v.kolokoltsov@warwick.ac.uk

Once a stochastic representation (1.1) for a mapping Φ is chosen one can naturally define, for any initial probability law $\mu = \mu^0$, a stochastic process i_l, $l \in \mathbf{Z}_+$, called a *nonlinear Markov chain*, on $\{1, ..., n\}$ in the following way. Starting with an initial position i_0 distributed according to μ one then chooses the next point i_1 according to the law $\{P_{i_0 j}(\mu)\}_{j=1}^n$, the distribution of i_1 becoming $\mu^1 = \Phi(\mu)$:

$$\mu_j^1 = \mathbf{P}(i_1 = j) = \sum_{i=1}^n P_{ij}(\mu)\mu_i = \Phi_j(\mu).$$

Then one chooses i_2 according to the law $\{P_{i_1 j}(\mu_1)\}_{j=1}^n$, etc. The law of this process at any given time k is $\mu^k = \Phi^k(\mu)$, i.e. is given by the semigroup. However, now the finite-dimensional distributions are defined as well. Namely, say for a function f of two discrete variables, one has

$$\mathbf{E}f(i_k, i_{k+1}) = \sum_{i,j=1}^n f(i,j)\mu_i^k P_{ij}(\mu^k).$$

In other words, this process can be defined as a time non-homogeneous Markov chain with the transition probabilities at time $t = k$ being $P_{ij}(\mu^k)$.

Clearly the finite-dimensional distributions depend on the choice of the representation (1.1). For instance, in case of the simplest representation (1.2) one has

$$\mathbf{E}f(i_0, i_1) = \sum_{i,j=1}^n f(i,j)\mu_i \Phi_j(\mu),$$

so that the discrete random variables i_0 and i_1 turn out to be independent.

Once the representation (1.1) is chosen, one can also define the transition probabilities P_{ij}^k in time $t = k$ recursively as

$$P_{ij}^k(\mu) = \sum_{m=1}^n P_{im}^{k-1}(\mu) P_{mj}(\mu^{k-1}).$$

The semigroup identity $\Phi^{k+l} = \Phi^k \Phi^l$ implies that

$$\Phi_j^k(\mu) = \sum_{i=1}^n P_{ij}^k(\mu)\mu_i$$

and

$$P_{ij}^k(\mu) = \sum_{m=1}^n P_{im}^l(\mu) P_{mj}^{k-l}(\mu^l), \quad l < k.$$

One can get nonlinear analogs of many results from the usual Markov chains. For example, let us present the following simple fact on the long time behavior.

Proposition 1.1 *(i) For any continuous $\Phi : \Sigma_n \to \Sigma_n$ there exists a stationary distribution, i.e. a measure $\mu \in \Sigma_n$ such that $\Phi(\mu) = \mu$. (ii) If a representation* (1.1) *for Φ is chosen in such a way that there exists a $j_0 \in [1, n]$, a time $k_0 \in \mathbf{N}$ and a positive δ such that*

$$P_{ij_0}^{k_0}(\mu) \geq \delta \tag{1.3}$$

for all i, μ, then $\Phi^m(\mu)$ converges to a stationary measure for any initial μ.

Proof. Statement (i) is a consequence of the Browder fixed point principle. Statement (ii) follows from the representation (given above) of the corresponding nonlinear Markov chain as a time non-homogeneous Markov process.

We shall turn now to nonlinear chains in continuous time. A *nonlinear Markov semigroup* with continuous time and the finite state space $\{1, ..., n\}$ is defined as a semigroup Φ^t, $t \geq 0$, of continuous transformations of Σ_n. As in the case of discrete time the semigroup itself does not specify a process. A *continuous family of nonlinear transition probabilities* on $\{1, ..., n\}$ is a family $P(t, \mu) = \{P_{ij}(t, \mu)\}_{i,j=1}^n$ of stochastic matrices depending continuously on $t \geq 0$ and $\mu \in \Sigma_n$ such that the following *nonlinear Chapman-Kolmogorov equation* holds:

$$\sum_{i=1}^n \mu_i P_{ij}(t+s, \mu) = \sum_{k,i} \mu_k P_{ki}(t, \mu) P_{ij}(s, \sum_{l=1}^n P_{l.}(t, \mu)\mu_l). \tag{1.4}$$

This family is said to yield a *stochastic representation* for the Markov semigroup Φ^t whenever

$$\Phi_j^t(\mu) = \sum_i \mu_i P_{ij}(t, \mu), \quad t \geq 0, \mu \in \Sigma_n. \tag{1.5}$$

If (1.5) holds, the equation (1.4) represents just the semigroup identity $\Phi^{t+s} = \Phi^t \Phi^s$.

Once a stochastic representation (1.5) for the semigroup Φ^k is chosen one can define the corresponding stochastic process started at $\mu \in \Sigma_n$ as a time nonhomogeneous Markov chain with the transition probabilities from time s to time t being

$$p_{ij}(s, t, \mu) = P_{ij}(t - s, \Phi^s(\mu)).$$

To get the existence of a stochastic representation (1.5) one can use the same idea as for the discrete time case and define

$$\tilde{P}_{ij}(t, \mu) = \Phi_j^t(\mu).$$

However, this is not a natural choice from the point of view of stochastic analysis. The natural choice should correspond to a reasonable generator.

Namely, assuming the semigroup Φ^t is differentiable in t one can define the *(nonlinear) infinitesimal generator* of the semigroup Φ^t as the nonlinear operator on measures given by

$$A(\mu) = \frac{d}{dt}\Phi^t|_{t=0}(\mu).$$

The semigroup identity for Φ^t implies that $\Phi^t(\mu)$ solves the Cauchy problem

$$\frac{d}{dt}\Phi^t(\mu) = A(\Phi^t(\mu)), \quad \Phi^0(\mu) = \mu. \tag{1.6}$$

As follows from the invariance of Σ_n under this dynamics, the mapping A is *conditionally positive* in the sense that $\mu_i = 0$ for a $\mu \in \Sigma_n$ implies $A_i(\mu) \geq 0$ and is also *conservative* in the sense that A maps the measures from Σ_n to the space of the signed measures

$$\Sigma_n^0 = \{\nu \in \mathbf{R}^n : \sum_{i=1}^n \nu_i = 0\}.$$

We shall say that such an A has a *stochastic representation* if it is written in the form

$$A_j(\mu) = \sum_{i=1}^n \mu_i Q_{ij}(\mu) = (\mu Q(\mu))_j, \tag{1.7}$$

where $Q(\mu) = \{Q_{ij}(\mu)\}$ is a family of infinitesimally stochastic matrices (also referred to as Q-matrices or Kolmogorov's matrices) depending on $\mu \in \Sigma_n$. Thus in stochastic representation the generator has the form of a usual Markov chain generator, though additionally depending on the present distribution. The existence of a stochastic representation for the generator is not as obvious as for the semigroup, but is not difficult to get, as shows the following statement, whose is based on the observation that as we are interested only in the action of Q on μ we can choose its action on the transversal to μ space Σ_n^0 in an arbitrary way.

Proposition 1.2 *For any differentiable in t nonlinear Markov semigroup Φ^t on Σ_n its infinitesimal generator has a stochastic representation.*

In practice, the converse problem is of more importance: not to construct the generator from a given semigroup, but to construct a semigroup (i.e. a solution to (1.6)) from a given operator A, which in applications is usually given directly in its stochastic representation. The problem of such a construction will be one of the central in this book, but in a much more general setting.

2 Examples: Lotka-Volterra, replicator dynamics, epidemics

The nonlinear Markov semigroups are present in abundance among the popular models of natural and social sciences, so that it would be difficult to distinguish the most important examples. We shall discuss here shortly three biological examples (anticipating our future analysis of the evolutionary games) and a statistical mechanics examples (anticipating the subsequent analysis of the kinetic equations).

The replicator dynamics of the evolutionary game arising from the classical paper-rock-scissors game has the form

$$\begin{cases} \dfrac{dx}{dt} = (y-z)x \\ \dfrac{dy}{dt} = (z-x)y \\ \dfrac{dz}{dt} = (x-y)z \end{cases} \tag{2.1}$$

Its generator has a clear stochastic representation (1.7) with the infinitesimal stochastic matrix

$$Q(\mu) = \begin{pmatrix} -z & 0 & z \\ x & -x & 0 \\ 0 & y & -y \end{pmatrix} \tag{2.2}$$

where $\mu = (x, y, z)$.

The famous LotkaVolterra equations describing a biological systems with two species, a predator and its prey, have the form

$$\begin{cases} \dfrac{dx}{dt} = x(\alpha - \beta y) \\ \dfrac{dy}{dt} = -y(\gamma - \delta x) \end{cases} \tag{2.3}$$

where $\alpha, \beta, \gamma, \delta$ are some positive parameters. The generator of this model is conditionally positive, but not conservative, as the total mass $x + y$ is not preserved. However, due to the existence of the integral of motion $\alpha \log y - \beta y + \gamma \log x - \delta x$, the dynamics (2.3) is path-wise equivalent to the dynamics (2.1), i.e. there is a continuous mapping taking the phase portrait of system (2.3) to the one of system (2.1).

One of the simplest deterministic models of epidemics can be written in the form of the system of 4 differential equations:

$$\begin{cases} \dot{X}(t) = -\lambda X(t)Y(t) \\ \dot{L}(t) = \lambda X(t)Y(t) - \alpha L(t) \\ \dot{Y}(t) = \alpha L(t) - \mu Y(t) \\ \dot{Z}(t) = \mu Y(t) \end{cases} \tag{2.4}$$

where $X(t), L(t), Y(t)$ and $Z(t)$ denote respectively the numbers of susceptible, latent, infectious and removed individual at time t and the positive coefficients λ, α, μ (which may actually depend on X, L, Y, Z) reflect the rates at which susceptible individuals become infected, latent individuals become infectious and infectious individuals become removed. Written in terms of the proportions $x = X/\sigma$, $y = Y/\sigma$, $l = L/\sigma$, $z = Z/\sigma$, i.e. normalized on the total mass $\sigma = X + L + Y + Z$, this system becomes

$$\begin{cases} \dot{x}(t) = -\sigma\lambda x(t)y(t) \\ \dot{l}(t) = \sigma\lambda x(t)y(t) - \alpha l(t) \\ \dot{y}(t) = \alpha l(t) - \mu y(t) \\ \dot{z}(t) = \mu y(t) \end{cases} \tag{2.5}$$

with $x(t) + y(t) + l(t) + z(t) = 1$. Subject to the often made assumption that $\sigma\lambda$, α and μ are constants, the r.h.s. is an infinitesimal generator of a nonlinear Markov chain in Σ_4. This generator depends again quadratically on its variable and has an obvious stochastic representation (1.7) with the infinitesimal stochastic matrix

$$Q(\mu) = \begin{pmatrix} -\lambda y & \lambda y & 0 & 0 \\ 0 & -\alpha & \alpha & 0 \\ 0 & 0 & -\mu & \mu \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.6}$$

where $\mu = (x, l, y, z)$, yielding a natural probabilistic interpretation to the dynamics 2.5, as explained in the previous section.

3 Discrete nonlinear Markov games and controlled processes

The theory of controlled stochastic Markov processes has a sound place in the literature, due to its wide applicability in practice. Here we shall touch upon the corresponding nonlinear extension just to indicate the possible directions of analysis.

The main point is that a nonlinear Markov semigroup is after all just a deterministic dynamic system (though on a weird state space of measures). Thus, as the stochastic control theory is a natural extension of the deterministic control, we are going to further extend it by turning back to deterministic control, but of measures, thus exemplifying the usual spiral development of science. The next 'turn of the screw' would lead to stochastic measure-valued games forming a stochastic control counterpart for the class of processes discussed in the previous section.

We shall work directly in the competitive control setting (game theory), which of course include the usual optimization as a particular case, but for simplicity only in discrete time and finite original state space $\{1, ..., n\}$. The full state space is then chosen as a set of probability measures Σ_n on $\{1, ..., n\}$.

Suppose we are given two metric spaces U, V of the control parameters of two players, a continuous transition cost function $g(u, v, \mu)$, $u \in U$, $v \in V$, $\mu \in \Sigma_n$ and a transition law $\nu(u, v, \mu)$ prescribing the new state $\nu \in \Sigma_n$ obtained from μ once the players had chosen their strategies $u \in U, v \in V$. The problem of the corresponding one-step game (with sequential moves) consists in calculating the Bellman operator

$$(BS)(\mu) = \min_u \max_v [g(u, v, \mu) + S(\nu(u, v, \mu))] \tag{3.1}$$

for a given final cost function S on Σ_n. According to the dynamic programming principle, the dynamic multi-step game solution is given by the iterations $B^k S$. Often of interest is the behavior of this optimal cost $B^k S(\mu)$ as the number of steps k go to infinity.

Remark 1 *In game theory one often assumes (but not always) that* min, max *in* (3.1) *are exchangeable (the existence of the value of the game, leading to the possibility of making simultaneous moves), but we shall not make or use this assumption.*

The function $\nu(u, v, \mu)$ can be interpreted as the controlled version of the mapping Φ specifying a nonlinear discrete time Markov semigroup discussed in Section 1. Assume a stochastic representation for this mapping is chosen, i.e.

$$\nu_j(u, v, \mu) = \sum_{i=1}^n \mu_i P_{ij}(u, v, \mu)$$

with a given family of (controlled) stochastic matrices P_{ij}. Then it is natural to assume g to describe the average over the random transitions, i.e. be given by

$$g(u, v, \mu) = \sum_{i,j=1}^n \mu_i P_{ij}(u, v, \mu) g_{ij}$$

with certain real coefficients g_{ij}. Under this assumption the Bellman operator (3.1) takes the form

$$(BS)(\mu) = \min_u \max_v [\sum_{i,j=1}^n \mu_i P_{ij}(u,v,\mu) g_{ij} + S\left(\sum_{i=1}^n \mu_i P_{i.}(u,v,\mu)\right)]. \tag{3.2}$$

We can now identify the (not so obvious) place of the usual stochastic control theory in this nonlinear setting. Namely, assume P_{ij} above do not depend on μ. But even then the set of the linear functions $S(\mu) = \sum_{i=1}^n s_i \mu^i$ on measures (identified with the set of vectors $S = (s_1, ..., s_n)$) is not invariant under B. Hence we are not automatically reduced to the usual stochastic control setting, but to a game with incomplete information, where the states are probability laws on $\{1, ..., n\}$, i.e. when choosing a move the players do not know the position precisely, but only its distribution. Only if we allow only Dirac measures μ as a state space (i.e. no uncertainty on the state), the Bellman operator would be reduced to the usual one of the stochastic game theory:

$$(\bar{B}S)_i = \min_u \max_v \sum_{j=1}^n P_{ij}(u,v)(g_{ij} + S_j). \tag{3.3}$$

As an example of a nonlinear result we shall get here an analog of the result on the existence of the average income for long lasting games.

Proposition 3.1 *If the mapping ν is a contraction uniformly in u, v, i.e. if*

$$\|\nu(u,v,\mu^1) - \nu(u,v,\mu^2)\| \leq \delta \|\mu^1 - \mu^2\| \tag{3.4}$$

with a $\delta \in (0,1)$, where $\|\nu\| = \sum_{i=1}^n |\nu_i|$, and if g is Lipschitz continuous, i.e.

$$\|g(u,v,\mu^1) - g(u,v,\mu^2)\| \leq C\|\mu^1 - \mu^2\| \tag{3.5}$$

with a constant $C > 0$, then there exists a unique $\lambda \in \mathbf{R}$ and a Lipschitz continuous function S on Σ_n such that

$$B(S) = \lambda + S, \tag{3.6}$$

and for all $g \in C(\Sigma_n)$ we have

$$\|B^m g - m\lambda\| \leq \|S\| + \|S - g\|, \tag{3.7}$$

$$\lim_{m \to \infty} \frac{B^m g}{m} = \lambda. \tag{3.8}$$

Proof. Clearly for any constant h and a function S one has $B(h + S) = h + B(S)$. Hence one can project B to the operator $\tilde{B}$ on the factor space $\tilde{C}(\Sigma_n)$ of $C(\Sigma_n)$ with respect to constant functions. Clearly in the image $\tilde{C}_{Lip}(\Sigma_n)$ of the set of Lipschitz continuous functions $C_{Lip}(\Sigma_n)$ the Lipschitz constant

$$L(f) = \sup_{\mu^1 \neq \mu^2} \frac{|f(\mu^1) - f(\mu^2)|}{\|\mu^1 - \mu^2\|}$$

is well defined (does not depend on the choice of the representative of an equivalence class). Moreover, from (3.4) and (3.5) it follows that

$$L(BS) \leq 2C + \delta L(S),$$

implying that the set

$$\Omega_R = \{f \in \tilde{C}_{Lip}(\Sigma_n) : L(f) \leq R\}$$

is invariant under $\tilde{B}$ whenever $R > C/(1-\delta)$. As by the Arzela-Ascoli theorem, Ω_R is convex and compact, one can conclude by the Shauder fixed point principle, that $\tilde{B}$ has a fixed point in Ω_R. Consequently there exists a $\lambda \in \mathbf{R}$ and a Lipschitz continuous function $\tilde{S}$ such that (3.6) holds.

Notice now that B is non-expansive in the usual sup-norm, i.e.

$$\|B(S_1) - B(S_2)\| = \sup_{\mu \in \Sigma_n} |(BS_1)(\mu) - (BS_2)(\mu)| \leq \sup_{\mu \in \Sigma_n} |S_1(\mu) - S_2(\mu)| = \|S_1 - S_2\|.$$

Consequently, for any $g \in C(\Sigma_n)$

$$\|B^m g - B^m S\| = \|B^m(g) - m\lambda - S\| \leq \|g - S\|,$$

implying the first formula in (3.7). The second one is its straightforward corollary. This second formula also implies the uniqueness of λ (as well as its interpretation as the average income). The proof is complete.

One can extend the other results for stochastic multi-step games to this nonlinear setting, say, the turnpike theorems from [2], and then go on studying the nonlinear Markov analogs of differential games, in particular, games of pursuit.

References

[1] V. N. Kolokoltsov. Nonlinear Markov processes and kinetic equations. Monograph. To appear in CUP.

[2] V. N. Kolokoltsov. On linear, Additive, and Homogeneous Operators in Idempotent Analysis. In: Advances in Soviet Mathematics **13** (1992), Idempotent Analysis, (Eds. V.P.Maslov and S.N. Samborski), 87-101.

[3] T.D. Frank. Nonlinear Markov processes. Phys. Lett. A **372:25** (2008), 4553-4555.

[4] M. Zak. Dynamics of intelligent systems. Int. J. Theor. Phys. **39:8** (2000), 2107-2140.

[5] M. Zak. Quantum Evolution as a Nonlinear Markov Process. Foundations of Physics Letters **15:3** (2002), 229-243.