Nonlinear Markov games

Vassili N. Kolokoltsov*

Abstract

A program of the analysis of a new class of stochastic games is put forward, which 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. [2], [5], [6]). This class of games can model a variety of situation for economics and epidemics, statistical physics, and pursuit - evasion processes. The discussion below will be presented i more detail in the author's monograph [1].

1. Nonlinear Markov chains

A discrete-time, discrete-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 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 an evolution of measures on $\{1, ..., n\}$. But it does not yet define a random process, because finite-dimensional distributions are not specified. In order to obtain a process we have to choose a *stochastic representation* for Φ, i.e. to write it down in the form

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

where $P_{ij}(\mu)$ is a family of stochastic matrices 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. There exists a unique representation (1.1) with the additional condition that all matrices $P_{ij}(\mu)$ are one dimensional:

$$P_{ij}(\mu) = \Phi_j(\mu), \quad i, j = 1, ..., n. \qquad (1.2)$$

*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 we 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 we then choose the next point i_1 according to the law $\{P_{i_0 j}(\mu)\}_{j=1}^n$, the distribution of i_1 becoming $\mu^1 = \Phi(\mu)$:

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

Then we choose i_2 according to the law $\{P_{i_1 j}(\mu^1)\}_{j=1}^n$, and so on. The law of this process at any given time k is $\mu^k = \Phi^k(\mu)$; that is, it 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, we have

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

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

Clearly the finite-dimensional distributions depend on the choice of representation (1.1). For instance, for the simplest representation (1.2) we have

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

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

Once representation (1.1) is chosen, we can also define the transition probabilities P^k_{ij} at time $t = k$ recursively as

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

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

$$\Phi^k_j(\mu) = \sum_{i=1}^n P^k_{ij}(\mu)\mu_i$$

ISBN 978-963-311-370-7

and

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

Remark 1. *In practical examples of the general model* (1.1) *the transition probabilities often depend on the law* μ *via its basic characteristics like standard deviation or expectation. See e.g. Frank [5], where we can also find some elementary examples of deterministic nonlinear Markov chains, for which the transitions are certain once the distribution is known, i.e. where* $P_{ij}(\mu) = \delta^j_{j(i,\mu)}$ *for a given deterministic mapping* $(i,\mu) \mapsto j(i,\mu)$.

We can establish nonlinear analogs of many results known for the usual Markov chains. For example, let us present the following simple fact about 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 exist* $j_0 \in [1,n]$, *time* $k_0 \in \mathbf{N}$ *and positive* δ *such that*

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

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 nonhomogeneous Markov process. □

Remark 2. *The convergence of* $P^m_{ij}(\mu)$ *as* $m \to \infty$ *can be shown by a standard argument. We introduce the bounds*

$$m_j(t,\mu) = \inf_i P^t_{ij}(\mu), \quad M_j(t,\mu) = \sup_i P^t_{ij}(\mu),$$

then we deduce from the semigroup property that $m_j(t,\mu)$ *(resp.* $M_j(t,\mu)$*) is an increasing (resp. decreasing) function of t, and finally we deduce from* (1.3) *that*

$$M_j(t+k_0,\mu) - m_j(t+k_0,\mu) \leq (1-\delta)\left(M_j(t,\mu) - m_j(t,\mu)\right),$$

*implying the required convergence. Alternatively, it can be established by the fixed point argument in the space of vectors factorized by constants, see e.g. [**?**], [4].*

We turn now to nonlinear chains in continuous time. A *nonlinear Markov semigroup* in continuous time and with 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 $P(t,\mu) = \{P_{ij}(t,\mu)\}^n_{i,j=1}$ 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

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

If (1.5) holds, 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 we can define the corresponding stochastic process started at $\mu \in \Sigma_n$ as a time nonhomogeneous Markov chain with transition probabilities from time s to time t

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

To show existence of a stochastic representation (1.5) we can use the same idea as in the discrete-time case and define $P_{ij}(t,\mu) = \Phi^t_j(\mu)$. However, this is not a natural choice from the point of view of stochastic analysis. The natural choice should arise from a generator that is reasonable from the point of view of the theory of Markov processes.

Namely, assuming the semigroup Φ^t is differentiable in t we 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 these 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 signed measures

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

We shall say that such a generator A has a *stochastic representation* if it can be written in the form

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

where $Q(\mu) = \{Q_{ij}(\mu)\}$ is a family of infinitesimally stochastic matrices (or Q=matrices) depending on $\mu \in \Sigma_n$. Thus in its stochastic representation the generator has the form of a usual Markov chain generator, though depending additionally on the present distribution. The existence of a stochastic representation for the generator is not difficult to obtain, see [1].

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

The examples of nonlinear Markov chains are numerous including Lotka-Volterra systems, general replicator dynamics of the evolutionary game theory, models of epidemics, coagulation processes, see more in [1].

2. Controlled nonlinear processes

Now we discuss the corresponding nonlinear extension of controlled processes.

Nonlinear Markov games can be considered as a systematic tool for modeling deception. In particular, in a game of pursuit - evasion, an evading object can create false objectives or hide in order to deceive the pursuit. Thus, observing this object leads not to its precise location, but to its distribution only, implying that it is necessary to build competitive control on the basis of the distribution of the present state. Moreover, by observing the action of the evading objects, one can make conclusions about its certain dynamic characteristics making the (predicted) transition probabilities depending on the observed distribution via these characteristics. This is precisely the type of situations modeled by nonlinear Markov games.

The starting point for the analysis is the observation that a nonlinear Markov semigroup is after all just a deterministic dynamic system (though on a weird state space of measures and with a specifically structured payoff function). 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 includes 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

$$(BS)(\mu) = \min_u \max_v [g(u,v,\mu) + S(\nu(u,v,\mu))] \qquad (2.1)$$

for a given final cost function S on Σ_n. According to the dynamic programming principle (see e.g. [4]), 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 3. *In game theory one often assumes (but not always) that* min, max *in* (2.1) *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 (2.1) takes the form

$$(BS)(\mu) = \min_u \max_v [\sum_{i,j=1}^n \mu_i P_{ij}(u,v,\mu) g_{ij} + S\left(\sum_{i=1}^n \mu_i P_{i\cdot}(u,v,\mu)\right)]. \qquad (2.2)$$

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:

$$(\bar{B}S)_i = \min_u \max_v \sum_{j=1}^{n} P_{ij}(u,v)(g_{ij} + S_j). \tag{2.3}$$

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

$$\|\nu(u,v,\mu^1) - \nu(u,v,\mu^2)\| \leq \delta\|\mu^1 - \mu^2\| \tag{2.4}$$

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

$$\|g(u,v,\mu^1) - g(u,v,\mu^2)\| \leq C\|\mu^1 - \mu^2\| \tag{2.5}$$

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

$$B(S) = \lambda + S, \tag{2.6}$$

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

$$\|B^m g - m\lambda\| \leq \|S\| + \|S - g\|, \tag{2.7}$$

$$\lim_{m\to\infty} \frac{B^m g}{m} = \lambda. \tag{2.8}$$

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 $\tilde{C}_{\text{Lip}}(\Sigma_n)$ of the set of Lipschitz continuous functions $C_{\text{Lip}}(\Sigma_n)$ 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 (2.4) and (2.5) it follows that

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

implying that the set

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

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 (2.6) 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 (2.7). 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). □

One can extend the other results for stochastic multi-step games to this nonlinear setting, say, the turnpike theorems from [3] (see also [4]), and then go on studying the nonlinear Markov analogs of differential games.

Acknowledgements. The work is supported by the AFOSR grant 'Nonlinear Markov Control Processes and Games'.

References

[1] V. N. Kolokoltsov. Nonlinear Markov processes and kinetic equations. Monograph. To appear in Cambridge University Press 2010.

[2] V. Kolokoltsov. Nonlinear Markov Semigroups and Interacting Lévy Type Processes. Journ. Stat. Physics **126:3** (2007), 585-642.

[3] 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.

[4] V.N. Kolokoltsov, O.A. Malafeyev. Understanding Game Theory. World Scientific, 2010.

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

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