NONLINEAR MARKOV GAMES*

Vassili N. Kolokoltsov†

I will discuss 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). This class of games can model a variety of situation for economics and epidemics, statistical physics, and pursuit - evasion processes. Further discussion of this topic is given in my monograph 'Nonlinear Markov processes and kinetic equations', CUP 2010.

Nonlinear Markov process: future depends on the past not only via its present position, but also its present distribution.

A nonlinear Markov semigroup can be considered as a nonlinear 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 extend it further by turning back to deterministic control, but of measures.

Important: introducing stochasticity in control destroys the max-plus linearity of the Bellman operator, the introduction of a nonlinear (distribution dependent) control restores this linearity.

Nonlinear Markov chains (discrete time)

A *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 $\{1, ..., .n\}$.

The family $\mu^k = \Phi^k \mu$ is the 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

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

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) exists, but is not unique. For instance, one can choose matrices $P_{ij}(\mu) = \tilde{P}_{ij}(\mu)$ to be one-dimensional:

$$\tilde{P}_{ij}(\mu) = \Phi_j(\mu), \quad i, j = 1, ..., n. \tag{2}$$

Once a stochastic representation (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). For instance, in case of the simplest representation (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) 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. *(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) *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*

$$P_{ij_0}^{k_0}(\mu) \geq \delta \tag{3}$$

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

Nonlinear Markov chains (continuous time)

A *nonlinear Markov semigroup* with the finite state space $\{1, ..., n\}$ is 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:

$$\sum_{i=1}^n \mu_i P_{ij}(t+s,\mu) = \sum_{k,i} \mu_k P_{ki}(t,\mu) P_{ij}(s, \sum_{l=1}^n P_{l.}(t,\mu)\mu_l). \tag{4}$$

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

$$\Phi_j^t(\mu) = \sum_i \mu_i P_{ij}(t,\mu), \quad t \geq 0, \mu \in \Sigma_n. \tag{5}$$

If (5) holds, the equation (4) represents just the semigroup identity $\Phi^{t+s} = \Phi^t \Phi^s$.

Once a stochastic representation (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 (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

$$\frac{d}{dt}\Phi^t(\mu) = A(\Phi^t(\mu)), \quad \Phi^0(\mu) = \mu. \qquad (6)$$

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

$$A_j(\mu) = \sum_{i=1}^{n} \mu_i Q_{ij}(\mu) = (\mu Q(\mu))_j, \qquad (7)$$

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 obvious, but is not difficult to get.

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 (6)) from a given operator A, which in applications is usually given directly in its stochastic representation.

Examples: Lotka-Volterra, replicator dynamics, epidemics

The nonlinear Markov semigroups are present in abundance among the popular models of natural and social sciences.

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

$$\begin{cases} \frac{dx}{dt} = (y-z)x \\ \frac{dy}{dt} = (z-x)y \\ \frac{dz}{dt} = (x-y)z \end{cases} \tag{8}$$

Its generator has a clear stochastic representation with

$$Q(\mu) = \begin{pmatrix} -z & 0 & z \\ x & -x & 0 \\ 0 & y & -y \end{pmatrix} \tag{9}$$

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

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

$$\begin{cases} \dfrac{dx}{dt} = x(\alpha - \beta y) \\ \dfrac{dy}{dt} = -y(\gamma - \delta x) \end{cases} \qquad (10)$$

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 (10) is path-wise equivalent to the dynamics (8), i.e. there is a continuous mapping taking the phase portrait of system (10) to the one of system (8).

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

$$\begin{cases} \dot{X}(t) = -\lambda X(t)Y(t) \\ \dot{L}(t) = \lambda X(t)Y(t) - \alpha L(t) \\ \dot{Y}(t) = \alpha L(t) - \mu Y(t) \\ \dot{Z}(t) = \mu Y(t) \end{cases} \qquad (11)$$

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

$$\begin{cases} \dot{x}(t) = -\sigma\lambda x(t)y(t) \\ \dot{l}(t) = \sigma\lambda x(t)y(t) - \alpha l(t) \\ \dot{y}(t) = \alpha l(t) - \mu y(t) \\ \dot{z}(t) = \mu y(t) \end{cases} \qquad (12)$$

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 (7) with the infinitesimal stochastic matrix

$$Q(\mu)=\begin{pmatrix} -\lambda y & \lambda y & 0 & 0 \\ 0 & -\alpha & \alpha & 0 \\ 0 & 0 & -\mu & \mu \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (13)$$

where $\mu=(x,l,y,z)$, yielding a natural probabilistic interpretation to the dynamics (12).

Discrete nonlinear Markov games and controlled processes

A nonlinear Markov semigroup is after all just a deterministic dynamic system (with some special features). 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.

We shall work in the competitive control setting (game theory), 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

$$(BS)(\mu) = \min_u \max_v [g(u, v, \mu) + S(\nu(u, v, \mu))] \tag{14}$$

for a given final cost function S on Σ_n.

In case of no competition (only one control parameter), this turns to

$$(BS)(\mu) = \min_u [g(u, v, \mu) + S(\nu(u, v, \mu))] \tag{15}$$

MAX-PLUS linear!

According to the dynamic programming principle, the dynamic multi-step game solution is given by the iterations $B^k S$.

Long horizon problem: behavior of the optimal cost $B^k S(\mu)$ as $k \to \infty$.

The function $\nu(u, v, \mu)$ can be interpreted as the controlled version of the mapping Φ specifying a nonlinear discrete time Markov semigroup.

Assume a stochastic representation is chosen:

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

with stochastic matrices P_{ij}. If g describes the averages over the random transitions, then

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

with certain real coefficients g_{ij} and

$$(BS)(\mu) = \min_u \max_v \Big[\sum_{i,j=1}^{n} \mu_i P_{ij}(u, v, \mu) g_{ij} + S\left(\sum_{i=1}^{n} \mu_i P_{i\cdot}(u, v, \mu) \right) \Big]. \quad (16)$$

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:

$$(\bar{B}S)_i = \min_u \max_v \sum_{j=1}^n P_{ij}(u,v)(g_{ij} + S_j). \quad (17)$$

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 (algebraically: max-plus eigenvector and eigenvalue).

Proposition 2. *If the mapping ν is a contraction uniformly in u, v, i.e. if*

$$\|\nu(u,v,\mu^1) - \nu(u,v,\mu^2)\| \le \delta \|\mu^1 - \mu^2\| \quad (18)$$

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

$$\|g(u,v,\mu^1) - g(u,v,\mu^2)\| \le C \|\mu^1 - \mu^2\| \quad (19)$$

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

$$B(S) = \lambda + S, \quad (20)$$

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

$$\|B^m g - m\lambda\| \le \|S\| + \|S - g\|, \quad (21)$$

$$\lim_{m \to \infty} \frac{B^m g}{m} = \lambda. \quad (22)$$

Proof. By homogeneity $B(h+S) = h + B(S)$ one can project B to the operator $\tilde{B}$ on the quotient space $\tilde{C}(\Sigma_n)$ of $C(\Sigma_n)$ with respect to constant functions.

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 (18) and (19) 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 (20) holds.

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 (21). 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 Kolokoltsov 1992).

Nonlinear max-plus spectral theory?!

[1] V. Kolokoltsov. Nonlinear Markov Semigroups and Interacting Lévy Type Processes. Journ. Stat. Physics **126:3** (2007), 585-642.
[2] T.D. Frank. Nonlinear Markov processes. Phys. Lett. A **372:25** (2008), 4553-4555.
[3] M. Zak. Quantum Evolution as a Nonlinear Markov Process. Foundations of Physics Letters **15:3** (2002), 229-243.
[4] V. N. Kolokoltsov. Nonlinear Markov processes and kinetic equations. Monograph. To appear in Cambridge University Press 2010.