Nonlinear Markov chains and nonlinear Lévy processes as dynamic LLN for interacting particles. [1]

Vassili N. Kolokoltsov[2]

[1] Bristol, November 2010.

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

Abstract and Plan

Highlights: (i) Nonlinear Markov process: future depends on the past not only via its present position, but also its present distribution.
(ii) A nonlinear Markov semigroup can be considered as a nonlinear deterministic dynamic system, though on a weird state space of measures. However: probabilistic interpretation makes the difference.
(iii) Including control, we extend the analysis to nonlinear controlled Markov processes and games.
Plan:
(i) Nonlinear Markov process as LLN
(ii) Nonlinear Lévy processes.

First-order PDE link with deterministic Markov processes

$$\frac{\partial S}{\partial t} = \left(b(x), \frac{\partial S}{\partial x} \right) = \sum_{j=1}^{d} b^j(x) \frac{\partial S}{\partial x_j}, \quad x \in \mathbf{R}^d,\ t \geq 0. \quad (1)$$

The solutions to the ODE $\dot{x} = b(x)$ are called the *characteristics* of the linear first-order PDE (1).
For $S_0 \in C^1 \mathbf{R}^d)$, $S(t,x) = S_0(X_t^x)$ solves the equation (1).

Proposition

PDE and deterministic Markov processes. *Let Ω be a (closed) polyhedron in $\mathbf{R}^d$ and $b(x)$ a vector-valued function on Ω of the class $C^1(\Omega)$. Assume that for any $x \in \Omega$ there exists a unique solution to the ODE $\dot{X}_t^x = b(X_t^x)$ with the initial condition x that stays in Ω for all times. Then the operators $T_t f(x) = f(X_t^x)$ form a Feller semigroup in $C(\Omega)$.*

Interacting particles: state space

Suppose our initial state space is a finite collection $\{1, ..., d\}$, which can be interpreted as the types of a particle (say, possible opinions of individuals on a certain subject). Let $Q(\mu) = (Q_{ij})(\mu)$ be a family of Q-matrices depending on a vector μ from the simplex

$$\Sigma_d = \{\mu = (\mu_1, ..., \mu_d) \in \mathbf{R}_+^d : \sum_{j=1}^{d} = 1\},$$

as on a parameter. Each such matrix specifies a Markov chain on $\{1, ..., d\}$ with the intensity of jumps

$$|Q_{ii}| = -Q_{ii}(\mu) = \sum_{j \neq i} Q_{ij}(\mu).$$

Full state space S: sequences of d non-negative integers $N = (n_1, ..., n_d)$ (numbers of particles in each state). $|N| = n_1 + ... + n_d$ the total number of particles in state N.

Interacting particles: Markov chain

For $i \neq j$ and a state N with $n_i > 0$ denote by N^{ij} the state obtained from N by removing one particle of type i and adding a particle of type j, that is n_i and n_j are changed to $n_i - 1$ and $n_j + 1$ respectively.
Markov chain generator:

$$Lf(N) = \sum_{i=1}^{d} n_i Q_{ij}(N/|N|)[f(N^{ij}) - f(N)]. \qquad (2)$$

Probabilistic description via $|Q_{ii}|(N/|N|)$-exponential random waiting times.
Such processes are usually called *mean -field interacting* Markov chains (as their transitions depend on the empirical measure $N/|N|$ or the mean field).

Interacting particles: scaling

Denote $h = 1/|N|$. Scaling and normalizing:

$$L_h f(N/|N|) = \sum_{i=1}^{d} \sum_{j=1}^{d} \frac{n_i}{|N|} |N| Q_{ij}(N/|N|)[f(N^{ij}/|N|) - f(N/|N|)], \tag{3}$$

or equivalently (e_i basis in $\mathbf{R}^d$)

$$L_h f(x) = \sum_{i=1}^{d} \sum_{j=1}^{d} x_i Q_{ij}(x) \frac{1}{h}[f(x - he_i + he_j) - f(x)], \quad x \in h\mathbf{Z}_+^d. \tag{4}$$

$$\lim_{|N|\to\infty,\, N/|N|\to x} L_h f(N/|N|) = \Lambda f(x) = \sum x_i Q_{ij}(x)[\frac{\partial f}{\partial x_j} - \frac{\partial f}{\partial x_i}](x)$$

$$= \sum [x_i Q_{ik}(x) - x_k Q_{ki}(x)] \frac{\partial f}{\partial x_k}(x).$$

Interacting particles: LLN limit

The limiting operator Λf is a first-order PDO with characteristics equation

$$\dot{x}_k = \sum_{i \neq k} [x_i Q_{ik}(x) - x_k Q_{ki}(x)] = \sum_{i=1}^{k} x_i Q_{ik}(x), \quad k = 1, ..., d, \tag{5}$$

called the *kinetic equations* for the process of interaction described above. The characteristics specify the dynamics of the deterministic Markov Feller process in Σ_d defined via the generator Λ (and Proposition 1 above).

Proposition

Let $Q_{ij}(\mu) \in C^2(\Sigma)$. The processes hN_t converge in distribution to the deterministic process X_t^x given by the kinetic equation (characteristics) above.

Nonlinear Markov semigroups

A semigroup of measurable transformations of probability measures is called a *nonlinear Markov semigroup* . One can show that, under mild regularity assumption, each nonlinear Markov semigroup on Σ_d (the set of probability laws on $\{1, ..., d\}$) arises from equations of the form (5) (*stochastic representation for a nonlinear Markov chain*).
We have shown that the solution to these equations describe the *dynamic law of large numbers*(LLN) (as the limit $|N| \to \infty$) of the mean-field interacting Markov chains.
Applications in both direction: continuous – discrete – continuous.

Interacting particles: fluctuations

$$Z_t = \frac{hN_t^N - X_t^x}{\sqrt{h}} \tag{6}$$

when $hN \to x$ and $(hN - x)/\sqrt{h}$ tend to a finite limit, as $h = 1/|N| \to 0$.

Z_t is a time non-homogeneous Markov with the propagator generated by the family of operators

$$A_t f = O_t f + O(\sqrt{h}),$$

where

$$O_t f(y) = \frac{1}{2} \sum_{i,j=1}^{d} X_{t,i} Q_{ij}(X_t) \left(\frac{\partial^2 f}{\partial y_j^2} - 2 \frac{\partial^2 f}{\partial y_j \partial y_i} + \frac{\partial^2 f}{\partial y_i^2} \right)$$

$$+ \sum_{i,j=1}^{d} [y_i Q_{ij}(X_t) + X_{t,i}(\nabla Q_{ij}(X_t), y)] \left(\frac{\partial f}{\partial y_j} - \frac{\partial f}{\partial y_i} \right). \tag{7}$$

Interacting particles: CLT

Operator O_t is a second-order PDO with a linear drift and position-independent diffusion coefficients. Hence it generates a Gaussian diffusion process (a kind of time non-homogeneous OU process). By the same argument as in Proposition 2, we arrive at the following.

Proposition

Dynamic CLT for simplest mean-field interactions. *Let all the elements $Q_{ij}(\mu)$ (of a given family of Q-matrices) belong to $C^2(\Sigma)$. Then the process of fluctuations* (6) *converge (both in the sense of convergence of propagators and the distributions on paths) to the Gaussian diffusion generated by the second order PDO* (7).

The examples of Markov chains of type (2) are numerous. For instance, in modeling a pool of voters, the transition $N \to N^{ij}$ is interpreted as the change of opinion.

Interacting particles: binary and ternary interactions

Similarly one can model binary, ternary or generally *kth order interaction*.

Say, if any two particles i, j of different type (binary interaction) can be transformed to a pair of type k, l (say, two agents i, j communicated and i changed her opinion to j) with rates Q_{ij}^{kl}:

$$Lf(N) = \sum_{i \neq j} \sum_{k,l=1}^{d} n_i n_j Q_{ij}^{kl}(N/|N|)[f(N^{ij,kl}) - f(N)],$$

where $N^{ij,kl}$ is obtained from N by changing two particles of type i, j to two particles of type k, l. Appropriate scaling leads to

$$\Lambda f = \sum_{i \neq j} \sum_{k,l=1}^{d} x_i x_j Q_{ij}^{kl} \left[\frac{\partial f}{\partial x_k} + \frac{\partial f}{\partial x_l} - \frac{\partial f}{\partial x_j} - \frac{\partial f}{\partial x_i} \right](x),$$

Nonlinear Markov semigroup: stochastic representation

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.
Stochastic representation for Φ^t:

$$\Phi_j^t(\mu) = \sum_i \mu_i P_{ij}(t, \mu), \quad t \geq 0, \mu \in \Sigma_n, \tag{8}$$

where $P(t, \mu) = \{P_{ij}(t, \mu)\}_{i,j=1}^n$ is a family of stochastic matrices depending continuously on $t \geq 0$ and $\mu \in \Sigma_n$ (nonlinear transition probabilities).

Nonlinear Markov chain (continuous time)

Once a stochastic representation (8) for the semigroup Φ^t 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

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

Thus, to each trajectory of a nonlinear semigroup there corresponds a *tangent Markov process*.
Stochastic representation for the semigroup depends on the stochastic representation for the generator.

Example: replicator dynamics (RD)

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

$$\begin{cases} \frac{dx}{dt} = (y - z)x \\ \frac{dy}{dt} = (z - x)y \\ \frac{dz}{dt} = (x - y)z \end{cases} \tag{9}$$

Its generator has a clear stochastic representation with

$$Q(\mu) = \begin{pmatrix} -z & 0 & z \\ x & -x & 0 \\ 0 & y & -y \end{pmatrix} \tag{10}$$

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

Example: simplest epidemics (1)

Let $X(t), L(t), Y(t)$ 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.
Basic model, written in terms of the proportions $x = X/\sigma$, $y = Y/\sigma$, $l = L/\sigma$, $z = Z/\sigma$, where $\sigma = X + L + Y + Z$:

$$\begin{cases} \dot{x}(t) = -\sigma\lambda x(t)y(t) \\ \dot{l}(t) = \sigma\lambda x(t)y(t) - \alpha l(t) \\ \dot{y}(t) = \alpha l(t) - \mu y(t) \\ \dot{z}(t) = \mu y(t) \end{cases} \tag{11}$$

with $x(t) + y(t) + l(t) + z(t) = 1$.

Example: simplest epidemics (2)

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 with the infinitesimal stochastic matrix

$$Q(\mu) = \begin{pmatrix} -\lambda y & \lambda y & 0 & 0 \\ 0 & -\alpha & \alpha & 0 \\ 0 & 0 & -\mu & \mu \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (12)$$

where $\mu = (x, I, y, z)$, yielding a natural probabilistic interpretation to the dynamics (11).

Time nonhomogeneous Lévy processes: propagator

$$L_t f(x) = \frac{1}{2}(G_t \nabla, \nabla) f(x) + (b_t, \nabla f)(x)$$
$$+ \int [f(x+y) - f(x) - (y, \nabla f(x)) \mathbf{1}_{B_1}(y)] \nu_t(dy). \qquad (13)$$

Proposition

For L_t with coefficients continuous in t, there exists a family $\Phi^{s,t}$ of positive linear contractions in $C_\infty(\mathbf{R}^d)$ depending strongly continuously on $s \le t$ such that for any $f \in C^2_\infty(\mathbf{R}^d)$ the function $f_s = \Phi^{s,t} f$ is the unique solution in $C^2_\infty(\mathbf{R}^d)$ of the Cauchy problem

$$\dot{f}_s = -L_s f_s, \quad s \le t, \quad f_t = f. \qquad (14)$$

Time nonhomogeneous Lévy processes: duality

Time nonhomogeneous Lévy processes (or *additive process*): Markov process generated by the time-dependent family of the operators L_t:

$$\mathbf{E}(f(X_t)|X_s = x) = (\Phi^{s,t} f)(x), \quad f \in C(\mathbf{R}^d).$$

Corollary

Dual operators on measures $V^{t,s} = (\Phi^{s,t})'$ depend weakly continuously on s, t and Lipschitz continuously in the norm topology of the Banach dual $(C^2_\infty(\mathbf{R}^d))'$ to $C^2_\infty(\mathbf{R}^d)$. Moreover, for any $\mu \in \mathcal{P}(\mathbf{R}^d)$, $V^{t,s}(\mu)$ yields the unique solution of the weak Cauchy problem

$$\frac{d}{dt}(f, \mu_t) = (L_t f, \mu_t), \quad s \le t, \quad \mu_s = \mu, \tag{15}$$

which is meant to hold for any $f \in C^2_\infty(\mathbf{R}^d)$.

Nonlinear Lévy processes: definition

$$A_\mu f(x) = \frac{1}{2}(G(\mu)\nabla, \nabla) f(x) + (b(\mu), \nabla f)(x)$$

$$+ \int [f(x+y) - f(x) - (y, \nabla f(x))\mathbf{1}_{B_1}(y)]\nu(\mu, dy), \quad (16)$$

depending on $\mu \in \mathcal{P}(\mathbf{R}^d)$.
Nonlinear Lévy semigroup generated by A_μ: semigroup V^t of weakly continuous transformations of $\mathcal{P}(\mathbf{R}^d)$:
$\forall \mu \in \mathcal{P}(\mathbf{R}^d)$, $f \in C^2_\infty(\mathbf{R}^d)$, $\mu_t = V^t(\mu)$ solves the problem

$$\frac{d}{dt}(f, \mu_t) = (A_{\mu_t} f, \mu_t), \quad t \geq 0, \quad \mu_0 = \mu.$$

Nonlinear Lévy process with initial law μ: time nonhomogeneous Lévy process generated by the family $A_{V^t\mu} f(x)$ and started with law μ at $t = 0$.

Nonlinear Lévy processes: basic well-posedness

Theorem

Suppose the coefficients of a family (16) *depend on* μ *Lipschitz continuously in the norm of the Banach space* $(C^2_\infty(\mathbf{R}^d))'$ *dual to* $C^2_\infty(\mathbf{R}^d)$, *i.e.*

$$\|G(\mu)-G(\eta)\|+\|b(\mu)-b(\eta)\|+\int \min(1,|y|^2)|\nu(\mu,dy)-\nu(\eta,dy)|$$

$$\leq \kappa\|\mu-\eta\|_{(C^2_\infty(\mathbf{R}^d))'} = \kappa \sup_{\|f\|_{C^2_\infty(\mathbf{R}^d)}\leq 1} |(f,\mu-\eta)| \qquad (17)$$

with constant κ. *Then there exists a unique nonlinear Lévy semigroup generated by* A_μ, *and hence a unique nonlinear Lévy process.*

Condition (17) is not at all weird. It holds when the coefficients G,b, ν depend on μ via certain integrals (possibly multiple) with smooth densities.

Further models and results

Models: Boltzmann, Smoluchovski, McKean-Vlasov, Landau-Fokker-Planck evolutionary games, arbitrary interacting (mean field or kth order interaction) Feller processes, say interacting stable-like processes, including manifolds, nonlinear quantum Markov proceses, control nonlinear Markov processes and games

Results: Well-posedness, rate of convergence of LLN, and CLT, long-time behavior. In particular: outstanding problem on coagulation.

Selected bibliography

[1] V. N. Kolokoltsov. Nonlinear Markov processes and kinetic equations. Cambridge University Press, August 2010.
[2] V. N. Kolokoltsov. Nonlinear Markov Semigroups and Interacting Lévy Type Processes. Journ. Stat. Physics **126:3** (2007), 585-642.
[3] V.N. Kolokoltsov. The Lévy-Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups. arXiv:0911.5688v1 [Math.PR] (2009). Prob. Theory Related Fields. Online First DOI 10.1007/s00440-010-0293-8.
[4] V. N. Kolokoltsov. Stochastic integrals and SDE driven by nonlinear Lévy noise. To appear in D. Crisan (Ed.) "Stochastic Analysis in 2010", Springer.
[5] V.N. Kolokoltsov. Nonlinear Markov games. Proceedings of the 19th MTNS Symposium, July 2010.