Construction of linear and nonlinear Markov processes via SDEs driven by nonlinear Lévy noise. [1]

Vassili N. Kolokoltsov[2]

[1]Humboldt Kollege 2010, Kiev, August 2010.

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

Abstract and Plan

1. We introduce the notion of nonlinear Markov process, first for a finite state space.
2. We present a general construction of Markov processes and semigroups (both linear and nonlinear) via SDEs driven by nonlinear Levy noise.
3. Including control, we extend the analysis to nonlinear controlled Markov processes and games.

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.

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

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

Stochastic representation for Φ:

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

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*.

Convergence to a stationary regime

Proposition

(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*

$$P_{ij_0}^{k_0}(\mu) \geq \delta \tag{2}$$

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 semigroup (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.
Stochastic representation for Φ^t:

$$\Phi_j^t(\mu) = \sum_i \mu_i P_{ij}(t, \mu), \quad t \geq 0, \mu \in \Sigma_n, \tag{3}$$

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 (3) 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.

Generator of a nonlinear Markov semigroup

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 (nonlinear Chapman-Kolmogorov equation) implies that $\Phi^t(\mu)$ solves the Cauchy problem

$$\frac{d}{dt}\Phi^t(\mu) = A(\Phi^t(\mu)), \quad \Phi^0(\mu) = \mu. \tag{4}$$

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

Generator: stochastic representation

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

$$A_j(\mu) = \sum_{i=1}^{n} \mu_i Q_{ij}(\mu) = (\mu Q(\mu))_j, \tag{5}$$

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.

Example: replicator dynamics (RD)

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

$$\begin{cases} \dfrac{dx}{dt} = (y - z)x \\ \dfrac{dy}{dt} = (z - x)y \\ \dfrac{dz}{dt} = (x - y)z \end{cases} \tag{6}$$

Its generator has a clear stochastic representation with

$$Q(\mu) = \begin{pmatrix} -z & 0 & z \\ x & -x & 0 \\ 0 & y & -y \end{pmatrix} \tag{7}$$

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 $\sigma = X + L + Y + Z$:

$$\begin{cases} \dot{x}(t) = -\sigma\lambda x(t)y(t) \\ \dot{l}(t) = \sigma\lambda x(t)y(t) - \alpha l(t) \\ \dot{y}(t) = \alpha l(t) - \mu y(t) \\ \dot{z}(t) = \mu y(t) \end{cases} \tag{8}$$

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 (5) with the infinitesimal stochastic matrix

$$Q(\mu) = \begin{pmatrix} -\lambda y & \lambda y & 0 & 0 \\ 0 & -\alpha & \alpha & 0 \\ 0 & 0 & -\mu & \mu \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (9)$$

where $\mu = (x, I, y, z)$, yielding a natural probabilistic interpretation to the dynamics (8).

Continuous state spaces (general nonlinear Markov semigroups)

General kinetic equation in the weak form:

$$\frac{d}{dt}(f, \mu_t) = (L_{\mu_t} f, \mu_t), \quad \mu_t \in \mathcal{P}(\mathbf{R}^d), \quad \mu_0 = \mu, \tag{10}$$

(that should hold, say, for all $f \in C_c^2(\mathbf{R}^d)$), where

$$L_\mu f(x) = \frac{1}{2}(G(x,\mu)\nabla, \nabla) f(x) + (b(x,\mu), \nabla f(x))$$
$$+ \int (f(x+y) - f(x) - (\nabla f(x), y)\mathbf{1}_{B_1}(y))\nu(x,\mu,dy). \tag{11}$$

They play indispensable role in the theory of interacting particles (mean field approximation) and exhaust all positivity preserving evolutions on measures subject to certain mild regularity assumptions. They include Vlasov, Boltzmann, Smoluchovski, Landau-Fokker-Planck equations, McKean diffusions and many other models.

Nonlinear Markov process: definition, approach via SDE

A resolving semigroup $U_t : \mu \mapsto \mu_t$ of the Cauchy problem for this equation specifies a so called *generalized or nonlinear Markov process* $X(t)$, whose distribution μ_t at time t can be determined by the formula $U_{t-s}\mu_s$ from its distribution μ_s at any previous moment s.

We exploit the idea of nonlinear integrators combined with a certain coupling of Lévy processes in order to push forward the probabilistic construction in a way that allows the natural Lipschitz continuous dependence of the coefficients G, b, ν on x, μ. Thus obtained extension of the standard SDEs with Lévy noise represents a probabilistic counterpart of the celebrated extension of the Monge mass transformation problem to the generalized Kantorovich one.

Wasserstein-Kantorovich metrics

$$W_p(\nu_1, \nu_2) = \left(\inf_{\nu} \int |y_1 - y_2|^p \nu(dy_1 dy_2) \right)^{1/p}, \qquad (12)$$

where inf is taken over the class of probability measures ν on $\mathbf{R}^{2d}$ that couple ν_1 and ν_2.
The Wasserstein distances between the distributions in the Skorohod space $D([0, T], \mathbf{R}^d)$:

$$W_{p,T}(X_1, X_2) = \inf \left(\mathbf{E} \sup_{t \le T} |X_1(t) - X_2(t)|^p \right)^{1/p}, \qquad (13)$$

where inf is over the couplings of the random paths X_1, X_2.
To compare Lévy measures, we extend these distances to unbounded measures with a finite pth moment.

Basic well-posedness: setting

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

$$+\int (f(x+z) - f(x) - (\nabla f(x), z))\nu(x,\mu; dz) \qquad (14)$$

with $\nu(x,\mu;.) \in \mathcal{M}_2(\mathbf{R}^d)$ (has a finite second moment). Let $Y_\tau(z,\mu)$ be a family of Lévy processes depending measurably on the points z and probability measures μ in $\mathbf{R}^d$ and specified by their generators

$$L[z,\mu]f(x) = \frac{1}{2}(G(z,\mu)\nabla,\nabla)f(x) + (b(z,\mu),\nabla f(x))$$

$$+\int (f(x+y) - f(x) - (\nabla f(x), y))\nu(z,\mu; dy) \qquad (15)$$

where $\nu(z,\mu) \in \mathcal{M}_2(\mathbf{R}^d)$.

Position dependent SDE with a nonlinear noise

Our approach to solving (10) is via the solution to the following *nonlinear distribution dependent stochastic equation* with *nonlinear Lévy type integrators*:

$$X(t) = X + \int_0^t dY_s(X(s), \mathcal{L}(X(s))), \ \mathcal{L}(X) = \mu, \tag{16}$$

with a given initial distribution μ and a random variable X independent of $Y_\tau(z, \mu)$.
Euler-Ito approximation:

$$X_\mu^\tau(t) = X_\mu^\tau(l\tau) + Y_{t-l\tau}^l(X_\mu^\tau(l\tau), \mathcal{L}(X_\mu^\tau(l\tau))), \tag{17}$$

$\mathcal{L}(X_\mu^\tau(0)) = \mu$, where $l\tau < t \leq (l+1)\tau$, $l = 0, 1, 2, ...$, and $Y_\tau^l(x, \mu)$ is a collection (depending on l) of independent families of the Lévy processes $Y_\tau(x, \mu)$ introduced above.

Basic well-posedness: formulation

Theorem

Assume

$$\|\sqrt{G(x,\mu)} - \sqrt{G(z,\eta)}\| + |b(x,\mu) - b(z,\eta)|$$

$$+ W_2(\nu(x,\mu;.), \nu(z,\eta;.)) \le \kappa(|x-z| + W_2(\mu,\eta)), \quad (18)$$

$$\sup_{x,\mu}\left(\sqrt{G(x,\mu)} + |b(x,\mu)| + \int |y|^2 \nu(x,\mu,dy)\right) < \infty. \quad (19)$$

Then for any $\mu \in \mathcal{P}(\mathbf{R}^d) \cap \mathcal{M}_2(\mathbf{R}^d)$ the approximations $X_\mu^{\tau_k}$, $\tau_k = 2^{-k}$, converge to a process $X_\mu(t)$ in W^2_{2,t_0} and the resolving operators $U_t : \mu \mapsto \mu_t$ of the Cauchy problem (10) *form a nonlinear Markov semigroup. If $L[z,\mu]$ do not depend explicitly on μ the operators $T_t f(x) = \mathbf{E} f(X_x(t))$ form a conservative Feller semigroup preserving the space of Lipschitz continuous functions.*

Basic well-posedness: example

(1) $\nu(x;.) = \sum_{n=1}^{\infty} \nu_n(x;.)$, $\nu_n(x,.)$ are probability measures with

$$W_2(\nu_i(x;.), \nu_i(z;.)) \leq a_i |x - z|$$

and the series $\sum a_i^2$ converges.
It is well known that the optimal coupling of probability measures (Kantorovich problem) can not always be realized via a mass transportation (a solution to the Monge problem), thus leading to the examples when the construction of the process via standard stochastic calculus would not work.
(2) common star shape of the measures $\nu(x;.)$:

$$\nu(x;dy) = \nu(x,s,dr)\,\omega(ds), \quad r = |y|, s = y/r, \qquad (20)$$

with a certain measure ω on S^{d-1} and a family of measures $\nu(x,s,dr)$ on $\mathbf{R}_+$. This allows to reduce the general coupling problem to a much more easily handled one-dimensional one.

Another point of view: additive processes as stochastic integrals driven by nonhomogeneous noise

Additive processes: generated by a time dependent family of Lévy-Khintchine operators

$$L_t f(x) = \frac{1}{2}(G_t \nabla, \nabla) f(x) + (b_t, \nabla f)(x)$$

$$+ \int [f(x+y) - f(x) - (y, \nabla f(x)) \mathbf{1}_{B_1}(y)] \nu_t(dy), \qquad (21)$$

where for any t, G_t is a non-negative symmetric $d \times d$-matrix, $b_t \in \mathbf{R}^d$ and ν_t is a Lévy measure. The set of Lévy measures is equipped with the weak topology, where the continuous dependence of the family ν_t on t means that $\int f(y) \nu_t(dy)$ depends continuously on t for any continuous f on $\mathbf{R}^d$ with $|f(y)| \leq c \min(|y|^2, 1)$.

Additive processes (well posedness)

Proposition

Suppose

$$\sup_t (\|G_t\| + \|b_t\| + \int (1 \wedge y^2)\nu_t(dy)) < \infty,$$

and the coefficients depend continuously on t a.s. (outside a fixed zero-measure subset $S \subset \mathbf{R}$). Then there exists a unique family $\{\Phi^{s,t}\}$ of positive linear contractions in $C_\infty(\mathbf{R}^d)$ depending strongly continuously on $s \leq t$ such that for any $f \in C_\infty^2(\mathbf{R}^d)$ the functions $f_s = \Phi^{s,t} f$ belong to $C_\infty^2(\mathbf{R}^d)$ and solve a.s. (i.e. for s outside a zero-measure set) the inverse-time Cauchy problem $\dot{f}_s = -L_s f_s$, $s \leq t$, $f_t = f$.

Nonhomogeneous Lévy processes (or additive process) X_t, generated by the family $\{L_t\}$: time-nonhomogeneous cadlag Markov process with transitions $\{\Phi^{s,t}\}$.

Additive processes: interpretation as weak stochastic integrals (1)

Namely, let L_η be a family of the operators of form (21) with coefficients G_η, b_η, ν_η depending continuously on a parameter $\eta \in \mathbf{R}^n$. Let ξ_t be a curve in $\mathbf{R}^n$ with not more than countably many discontinuities and with left and right limits existing everywhere. Then the family of operators L_{ξ_t} satisfies the assumptions of Proposition 2. Clearly, the resulting propagator $\{\Phi^{s,t}\}$ does not depend on the values of ξ_t at the points of discontinuity.

Applying the Randomization Lemma to the distributions of the family of the Lévy processes $Y_t(\eta)$ (corresponding to the generators L_η, we define them on a single probability space (actually on the standard Lebesgue space) in such a way that they depend measurably on the parameter η.

Additive processes: interpretation as weak stochastic integrals (2)

Let ξ_s, α_s be piecewise constant left continuous functions (deterministic, to begin with) with values in $\mathbf{R}^n$ and $d \times d$-matrices respectively, that is

$$\xi_s = \sum_{j=0}^{n} \xi^j \mathbf{1}_{(t_j,t_{j+1}]}(s), \quad \alpha_s = \sum_{j=0}^{n} \alpha^j \mathbf{1}_{(t_j,t_{j+1}]}(s), \tag{22}$$

where $0 = t_0 < t_1 < .... < t_{n+1}$. Then it is natural to define the stochastic integral with respect to the nonlinear Lévy noise $Y_s(\xi_s)$ by the formula

$$\int_0^t \alpha_s dY_s(\xi_s) = \sum_{j=0}^{n} \alpha^j Y^j_{t\wedge t_{j+1}-t_j}(\xi^j) \mathbf{1}_{t_j<t}, \tag{23}$$

where $Y^j_t(\eta)$ are independent copies of the families of $Y_t(\eta)$.

Additive processes: interpretation as weak stochastic integrals (3)

$\int_0^t \alpha_s dY_s(\xi_s)$ is a nonhomogeneous Lévy process constructed by Proposition 2 from the generator family

$$L_t^{\alpha,\xi} f(x) = \frac{1}{2}((\alpha_t G_{\xi_t} \alpha_t')\nabla, \nabla) f(x) + (\alpha_t b_{\xi_t}, \nabla f)(x)$$

$$+ \int [f(x + \alpha_t y) - f(x) - (\alpha_t y, \nabla f(x)) \mathbf{1}_{B_1}(y)] \nu_{\xi_t}(dy), \quad (24)$$

which coincides with L_{ξ_t} for $\alpha_t = 1$.

Additive processes: interpretation as weak stochastic integrals (4)

If ξ_t and α_t are arbitrary cadlag function, let us define its natural piecewise constant approximation as

$$\xi_t^\tau = \sum_{\tau j < t}^{n} \xi_{\tau j} \mathbf{1}_{(\tau j, \tau(j+1)]}, \quad \alpha_t^\tau = \sum_{\tau j < t}^{n} \alpha_{\tau j} \mathbf{1}_{(\tau j, \tau(j+1)]},$$

As usual the integral $\int_0^t \alpha_s dY_s(\xi_s)$ should be defined as a limit (if it exists in some sense) of the integrals over its approximations $\int_0^t \alpha_s^\tau dY_s(\xi_s^\tau)$.

Additive processes: interpretation as weak stochastic integrals (5)

Theorem

The distribution of the process of integrals $x + \int_0^t \alpha_s dY_s(\xi_s)$ *is well defined as the weak limit, as* $\tau \to 0$*, of the distributions on the Skorohod space* $D([0,T], \mathbf{R}^d)$*, of the approximating simple integrals* $x + \int_0^t \alpha_s^\tau dY_s(\xi_s^\tau)$*, and is the distribution of the Lévy process started at* x *and generated by the family* (24)*. This limit also holds in the sense of the convergence of the propagators of the corresponding nonhomogeneous Lévy processes.*

In particular, we have constructed the probability kernel on the space $D([0,T], \mathbf{R}^d)$ of cadlag paths that takes a curve ξ_t to the distribution of the integral $x + \int_0^t dY_s(\xi_s)$. An invariant measure for this kernel defines a weak solution to the stochastic equation $\xi_t = \xi_0 + \int_0^t dY_s(\xi_s)$, and its uniqueness implies the Markov property of the corresponding process.

Controlled nonlinear processes and games (1)

Consider a single control variable u and assume that μ only is observable, so that the control is based on μ. This leads to the following infinite-dimensional HJB equation

$$\frac{\partial S}{\partial t} + \max_u \left(L_{\mu,u} \frac{\delta S}{\delta \mu} + g_u, \mu \right) = 0. \tag{25}$$

If the Cauchy problem for the corresponding kinetic equation $\dot{\mu} = L^{\star}_{\mu,u}\mu$ is well posed (say, above theorem applies) uniformly for controls u from a compact set, with a solution denoted by $\mu^t(\mu, u)$ this can be resolved via discrete approximations

$$S_k(t-s) = B^k S(t), \quad k = (t-s)/\tau,$$

$$BS(\mu) = \max_u [S(\mu^\tau(\mu, u) + (g_u, \mu)].$$

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). To appear in PTRF.
[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.