Ester Mariucci

Distances for Lévy processes

Joint work with Markus Reiß
Heidelberg, 19 July 2018

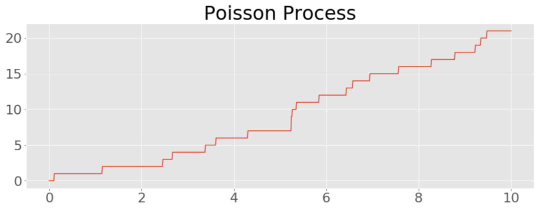
Institut für Mathematische Optimierung
Otto-von-Guericke-Universität Magdeburg

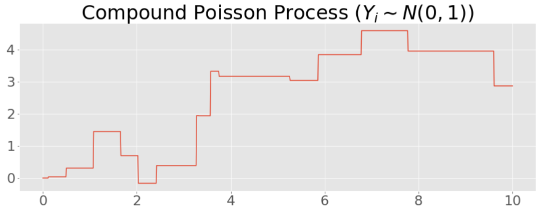
DFG-Graduiertenkolleg
MATHEMATISCHE
KOMPLEXITÄTSREDUKTION

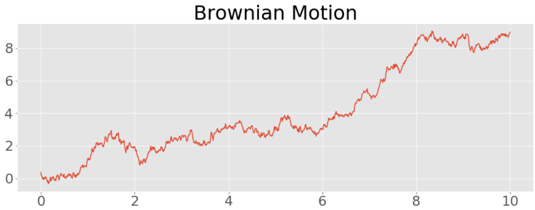
Table of contents

Lévy processes

Lévy processes are stochastic processes with independent, stationary increments and trajectories a.s. càdlàg (right continuous with left limits).







Why Lévy processes?

- Lévy processes are the fundamental building blocks in stochastic models whose evolution in time exhibits sudden changes in value.
- They play a central role in many fields of science such as *mathematical science*, *actuarial science*, *biology*, *physics*, *engineering*, etc...
- There is a one to one correspondence between Lévy processes and infinitely divisible distributions.

Lévy triplet

$$X_t = \sigma W_t + bt + \text{Jumps}_t, \quad W \text{ standard Brownian motion.}$$

The law of X is determined by its Lévy triplet (b, σ^2, ν):

- $b \in \mathbb{R}$ is the **drift**,
- $\sigma^2 \in \mathbb{R}_{\geq 0}$ is the **diffusion coefficient**,
- ν **Lévy measure**: $\nu : \nu(\{0\}) = 0$ and $\int_{\mathbb{R}} (x^2 \wedge 1)\nu(dx) < \infty$.

  Example: $\nu(dx) = \lambda \delta_1(dx)$, $\frac{\nu(dx)}{dx} = \frac{e^{-x}}{x} \mathbf{1}_{x>0}$, $\frac{\nu(dx)}{dx} = \frac{e^{-x}}{x^2} \mathbf{1}_{x>0}$.

  Interpretation of ν: $\forall B \in \mathscr{B}(\mathbb{R})$

$$\nu(B) = \frac{1}{t}\mathbb{E}\left[\sum_{0<s\leq t} \mathbf{1}_B(\Delta X_s)\right]$$

  i.e. is the average number of jumps whose magnitude falls in B.

Lévy-Itô decomposition

X is a Lévy process of Lévy triplet (b, σ^2, ν) iff:

$$X_t = \sigma W_t + bt + \lim_{\eta \to 0} \Big(\sum_{s \le t} \Delta X_s \mathbb{I}_{\eta < |\Delta X_s| \le 1} - t \int_{\eta < |x| \le 1} x \nu(dx) \Big) + \sum_{s \le t} \Delta X_s \mathbb{I}_{|\Delta X_s| > 1},$$

where

- W is a standard Brownian motion;
- X^S is a centred martingale describing the small jumps ($\Delta X_s = X_s - \lim_{r \uparrow s} X_r$):

$$\Big(\sum_{s \le t} \Delta X_s \mathbb{I}_{\eta < |\Delta X_s| \le 1} - t \int_{\eta < |x| \le 1} x \nu(dx) \Big) \xrightarrow[\eta \to 0]{L^2} X^S;$$

- $X^B := \left(\sum_{s \le t} \Delta X_s \mathbb{I}_{|\Delta X_s| > 1} \right)_{t \ge 0}$ is a compound Poisson process;
- W, X^S and X^B are independent of each other.

Lévy-Khintchine formula

Equivalently, X is a Lévy process of Lévy triplet (b, σ^2, ν) iff:

$$\mathbb{E}\left[e^{iuX_t}\right] = \exp(-t\psi(u)), \quad \forall u \in \mathbb{R},$$

where

$$\psi(u) = -iub + \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} \left(1 - e^{iux} + (iux\mathbb{I}_{|x|\le 1})\right)\nu(dx).$$

Lévy-Itô decomposition

If X is a Lévy process of Lévy triplet (b, σ^2, ν) then, for all $0 < \varepsilon \leq 1$

$$
\begin{aligned}
X_t &= \sigma W_t + b(\varepsilon)t + \lim_{\eta \to 0} \left(\sum_{s \leq t} \Delta X_s \mathbb{I}_{\eta < |\Delta X_s| \leq \varepsilon} - t \int_{\eta < |x| \leq \varepsilon} x \nu(dx) \right) + \sum_{s \leq t} \Delta X_s \mathbb{I}_{|\Delta X_s| > \varepsilon} \\
&= \sigma W_t + b(\varepsilon)t + X_t^S(\varepsilon) + X_t^B(\varepsilon),
\end{aligned}
$$

where

$$
\sigma^2 \in \mathbb{R}_{\geq 0}, \quad b(\varepsilon) = b - \int_{\varepsilon < |x| \leq 1} x \nu(dx).
$$

The problem we consider

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$, $j = 1, 2$, be two Lévy processes observed at times $0 = t_0 < \cdots < t_n = T$ and let $\Delta_i X := X_{t_i} - X_{t_{i-1}}$.

The problem we consider

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$, $j = 1, 2$, be two Lévy processes observed at times $0 = t_0 < \cdots < t_n = T$ and let $\Delta_i X := X_{t_i} - X_{t_{i-1}}$.

Goal: To better understand the geometry of the class of discretely observed Lévy processes.

Read: To find a "good" distance D that allows to quantify

$$D\big((\Delta_i X^1)_{i=1}^n, (\Delta_i X^2)_{i=1}^n\big) \leq \ ???$$

and provide sharp bounds for D that do not separate the continuous part from the discontinuous one. E.g., for $n = 1$, we do **not** want a bound of the form:

$$D(X_t^1, X_t^2) \leq f\big(b_1, b_2, \sigma_1^2, \sigma_2^2, d(\nu_1, \nu_2)\big)$$

for some function f and distance d.

Difference between continuous and discrete observations

Consider, for instance, the class of α-stable Lévy processes.

- On the Skorokhod path space $D([0,T])$ α-stable processes, $\alpha \in (0,2)$, induce laws singular to that of Brownian motion ($\alpha = 2$).
- But, their respective marginals at $t_k = kT/n$, $k = 0, \dots, n$ for n fixed, have equivalent laws, which even converge in total variation distance as $\alpha \to 2$ to those of Brownian motion.

Total variation distance for marginals of Lévy processes

Theorem (Liese)

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$, $j = 1, 2$, be two Lévy processes.
Then, the total variation distance between the laws of X_t^1 and X_t^2 is bounded as:

$$\|\mathscr{L}(X_t^1) - \mathscr{L}(X_t^2)\|_{TV} \leq 2\sqrt{1 - \Big(1 - \frac{H^2(\mathcal{N}(t\tilde{b}_1, t\sigma_1^2), \mathcal{N}(t\tilde{b}_2, t\sigma_2^2))}{2}\Big)^2 \exp(-tH^2(\nu_1, \nu_2))}$$

with $\tilde{b}_1 = b_1 - \int_{-1}^1 x\nu_1(dx)$, $\tilde{b}_2 = b_2 - \int_{-1}^1 x\nu_2(dx)$.

F. Liese, Estimates of Hellinger integrals of infinitely divisible distributions. Kybernetika, 1987.

Weak convergence for marginals of Lévy processes

Theorem (Gnedenko, Kolmogorov)

Marginals of Lévy processes $X^n = (X^n_t)_{t\geq 0}$ with Lévy triplets (b_n, σ_n^2, ν_n) converge weakly to marginals of a Lévy process $X = (X_t)_{t\geq 0}$ with Lévy triplet (b, σ^2, ν) if and only if

$$b_n \to b \text{ and } \sigma_n^2 \delta_0 + (x^2 \wedge 1)\nu_n(dx) \xrightarrow{w} \sigma^2 \delta_0 + (x^2 \wedge 1)\nu(dx),$$

where δ_0 is the Dirac measure in 0 and $\xrightarrow{w}$ denotes weak convergence of finite measures.

Example: Let $X(\varepsilon) \sim \Big(0, 0, \frac{\delta_{-\varepsilon}+\delta_\varepsilon}{2\varepsilon^2}\Big)$, i.e. $X_t(\varepsilon) = \varepsilon\big(N_t^1(\varepsilon) - N_t^2(\varepsilon)\big)$ with N^1 and N^2 independent Poisson processes of intensity $\frac{1}{2\varepsilon^2}$. Then

$$\mathscr{L}(X_t(\varepsilon)) \xrightarrow{w} \mathcal{N}(0, t), \quad \text{as } \varepsilon \downarrow 0.$$

B.V. Gnedenko and A.N. Kolmogorov, Limit distributions for sums of independent random variables. Addison-Wesley, 1954.

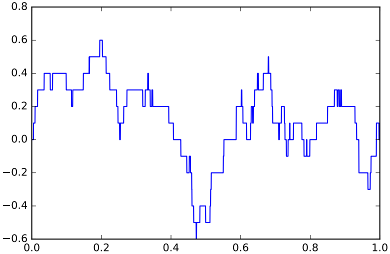
Example

Let $X(\varepsilon) \sim (0, 0, \nu_\varepsilon)$, where $\nu_\varepsilon = \frac{\delta_{-\varepsilon}+\delta_\varepsilon}{2\varepsilon^2}$. We have that $\mathscr{L}(X_t(\varepsilon)) \xrightarrow{w} \mathcal{N}(0, t)$ but

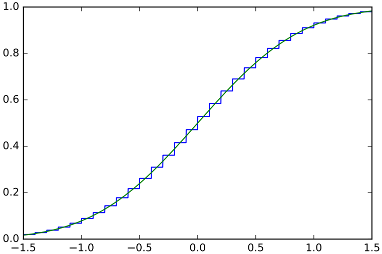
$$\|\mathscr{L}(X_t(\varepsilon)), \mathcal{N}(0, t)\|_{TV} = 1$$

whereas

$$\mathscr{W}_1(\mathscr{L}(X_t(\varepsilon)), \mathcal{N}(0, t)) = \int |F_X(u) - \phi(u)| du \to 0.$$



(a) A trajectory of $X_t(\varepsilon = \frac{1}{10})$.



(b) CDF of $X_t(\varepsilon = \frac{1}{10})$ and $\mathcal{N}(0, t)$.

Definition

Let $(\mathcal{X}, d)$ be a Polish metric space. Given $p \in [1, \infty)$, let

$$\mathcal{P}_p(\mathcal{X}) := \{\mu : \mathbb{E}_\mu[d(X, x_0)^p] < \infty, \text{ for some } x_0 \in \mathcal{X}\}.$$

Definition

Given $p \geq 1$, for any two probability measures $\mu, \nu \in \mathcal{P}_p(\mathcal{X})$, the **Wasserstein distance of order p** between μ and ν is

$$\mathscr{W}_p(\mu, \nu) = \inf \left\{ \left[\mathbb{E}[d(X', Y')^p] \right]^{\frac{1}{p}}, \ \text{law}(X') = \mu, \ \text{law}(Y') = \nu \right\},$$

where the infimum is taken over all random variables X' and Y' having laws μ and ν, respectively.

Notation: If $\text{law}(X) = \mu$ and $\text{law}(Y) = \nu$, we will write $\mathscr{W}_p(X, Y)$ instead of $\mathscr{W}_p(\mu, \nu)$.

Properties

The Wasserstein distances have the following properties:

1. For all $p \geq 1$, $\mathscr{W}_p(\cdot,\cdot)$ is a metric on $\mathcal{P}_p(\mathcal{X})$.
2. If $1 \leq p \leq q$, then $\mathcal{P}_q(\mathcal{X}) \subseteq \mathcal{P}_p(\mathcal{X})$, and $\mathscr{W}_p(\mu,\nu) \leq \mathscr{W}_q(\mu,\nu)$ for every $\mu,\nu \in \mathcal{P}_q(\mathcal{X})$.
3. Given a sequence $(\mu_n)_{n\geq 1}$ and a probability measure μ in $\mathcal{P}_p(\mathcal{X})$

$$\lim_{n\to\infty} \mathscr{W}_p(\mu_n,\mu) = 0$$

   if and only if $(\mu_n)_{n\geq 1}$ converges to μ weakly and

$$\lim_{n\to\infty} \int_{\mathcal{X}} d(x,x_0)^p \mu_n(dx) = \int_{\mathcal{X}} d(x,x_0)^p \mu(dx).$$

4. The infimum is actually a minimum; i.e., there exists a pair (X^*,Y^*) of jointly distributed $\mathcal{X}$-valued random variables with $\mathrm{law}(X^*) = \mu$ and $\mathrm{law}(Y^*) = \nu$, such that

$$\mathscr{W}_p(\mu,\nu)^p = \mathbb{E}[d(X^*,Y^*)^p].$$

Some useful properties

(P1) Homogeneity and sub-additivity:

$$\mathscr{W}_p(X+Z, Y+Z) \le \mathscr{W}_p(X,Y), \quad \mathscr{W}_p(cX, cY) = |c|\mathscr{W}_p(X,Y).$$

$$\mathscr{W}_p(X_1+\cdots+X_n, Y_1+\cdots+Y_n) \le \sum_{i=1}^n \mathscr{W}_p(X_i, Y_i).$$

Some useful properties

(P1) Homogeneity and sub-additivity:

$$\mathscr{W}_p(X+Z, Y+Z) \leq \mathscr{W}_p(X,Y), \quad \mathscr{W}_p(cX, cY) = |c|\mathscr{W}_p(X,Y).$$

$$\mathscr{W}_p(X_1+\cdots+X_n, Y_1+\cdots+Y_n) \leq \sum_{i=1}^n \mathscr{W}_p(X_i, Y_i).$$

(P2) A good behavior with respect to product measures: if

$$d(x,y) = \left(\sum_{i=1}^n |x_i - y_i|^r\right)^{1/r}, \quad r \geq 1;$$

then,

$$\mathscr{W}_p(\mu^{\otimes n}, \nu^{\otimes n})^p \leq n^{\frac{p}{r}} \mathscr{W}_p(\mu,\nu)^p.$$

Some useful properties

(P1) Homogeneity and sub-additivity:

$$\mathscr{W}_p(X+Z, Y+Z) \leq \mathscr{W}_p(X, Y), \quad \mathscr{W}_p(cX, cY) = |c|\mathscr{W}_p(X, Y).$$

$$\mathscr{W}_p(X_1 + \cdots + X_n, Y_1 + \cdots + Y_n) \leq \sum_{i=1}^{n} \mathscr{W}_p(X_i, Y_i).$$

(P2) A good behavior with respect to product measures: if

$$d(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^r \right)^{1/r}, \quad r \geq 1;$$

then,

$$\mathscr{W}_p(\mu^{\otimes n}, \nu^{\otimes n})^p \leq n^{\frac{p}{r}} \mathscr{W}_p(\mu, \nu)^p.$$

(P3) An accurate central limit theorem.

CLT for $\mathscr{W}_p$, $1 \le p \le 2$

Theorem (Esseen $p=1$, (1958); Rio $1 \le p \le 2$ (2009))

Let $(Y_i)_{i\ge1}$ be a sequence of centred i.i.d. random variables with finite and positive variance σ^2. For any $n \ge 1$ and any $1 \le p \le 2$, there exists some positive constant C depending only on p such that

$$\mathscr{W}_p\left(\frac{1}{\sqrt{n\sigma^2}}\sum_{i=1}^n Y_i, \mathcal{N}(0,1)\right) \le C\frac{\left(\mathbb{E}\left[|Y_1|^{p+2}\right]\right)^{1/p}}{\sqrt{n}\left(\mathrm{Var}[Y_1]\right)^{\frac{p+2}{2p}}}.$$

E. Rio, Upper bounds for minimal distances in the central limit theorem. Ann. Inst. Henri Poincaré, 2009.

Wasserstein distances for the marginals of Lévy processes

Theorem (M., Reiß)

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$ *be Lévy processes and set* $\bar\sigma_j^2(\varepsilon) = \int_{|x|\le\varepsilon} x^2 \nu_j(dx)$, $j = 1, 2$. *For all* $p \in [1, 2]$, *for all* $0 < \varepsilon \le 1$ *and all* $t \ge 0$, *the following estimate holds*

$$\begin{aligned}\mathscr{W}_p\big(\mathscr{L}(X_t^1), \mathscr{L}(X_t^2)\big) \le{} & t\big|b_1(\varepsilon) - b_2(\varepsilon)\big| \\ & + \sqrt{t\big(\sigma_1 + \bar\sigma_1(\varepsilon) - \sigma_2 - \bar\sigma_2(\varepsilon)\big)^2} \\ & + C\sum_{j=1}^{2} \min\Big(\sqrt{t}\bar\sigma_j(\varepsilon), \varepsilon\Big) \\ & + \mathscr{W}_p\big(\mathscr{L}(X_t^{1,B}(\varepsilon)), \mathscr{L}(X_t^{2,B}(\varepsilon))\big),\end{aligned}$$

for a constant C *only depending on* p.

Wasserstein distances between the marginals of Lévy processes

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$, $j = 1, 2$, be two Lévy processes. Then,

$$X_t^j \overset{\text{law}}{=} \mathcal{N}(tb_j(\varepsilon), t\sigma_j^2) + X_t^{j,S}(\varepsilon) + X_t^{j,B}(\varepsilon).$$

By sub-additivity of the Wasserstein distances

$$\mathscr{W}_p\Big(X_t^1, X_t^2\Big) \leq \mathscr{W}_p\Big(X_t^{1,B}(\varepsilon), X_t^{2,B}(\varepsilon)\Big)$$
$$+ \mathscr{W}_p\Big(\mathcal{N}\Big(tb_1(\varepsilon), t\sigma_1^2\Big) + X_t^{1,S}(\varepsilon), \mathcal{N}\Big(tb_2(\varepsilon), t\sigma_2^2\Big) + X_t^{2,S}(\varepsilon)\Big).$$

Wasserstein distances between the marginals of Lévy processes

Also,

$$
\begin{aligned}
&\mathscr{W}_p\Big(\mathcal{N}\big(tb_1(\varepsilon), t\sigma_1^2\big) + X_t^{1,S}(\varepsilon), \mathcal{N}\big(tb_2(\varepsilon), t\sigma_2^2\big) + X_t^{2,S}(\varepsilon)\Big)\\
&\leq \mathscr{W}_p\Big(X_t^{1,S}(\varepsilon), \mathcal{N}\big(0, t\bar{\sigma}_1^2(\varepsilon)\big)\Big) + \mathscr{W}_p\Big(X_t^{2,S}(\varepsilon), \mathcal{N}\big(0, t\bar{\sigma}_2^2(\varepsilon)\big)\Big)\\
&\quad + \mathscr{W}_p\Big(\mathcal{N}\big(tb_1(\varepsilon), t(\sigma_1^2 + \bar{\sigma}_1^2(\varepsilon))\big), \mathcal{N}\big(tb_2(\varepsilon), t(\sigma_2^2 + \bar{\sigma}_2^2(\varepsilon))\big)\Big).
\end{aligned}
$$

Wasserstein distances between the marginals of Lévy processes

Also,

$$
\begin{aligned}
&\mathscr{W}_p\Big(\mathcal{N}\big(tb_1(\varepsilon), t\sigma_1^2\big) + X_t^{1,S}(\varepsilon), \mathcal{N}\big(tb_2(\varepsilon), t\sigma_2^2\big) + X_t^{2,S}(\varepsilon)\Big)\\
&\le \mathscr{W}_p\Big(X_t^{1,S}(\varepsilon), \mathcal{N}\big(0, t\bar{\sigma}_1^2(\varepsilon)\big)\Big) + \mathscr{W}_p\Big(X_t^{2,S}(\varepsilon), \mathcal{N}\big(0, t\bar{\sigma}_2^2(\varepsilon)\big)\Big)\\
&\quad + \mathscr{W}_p\Big(\mathcal{N}\big(tb_1(\varepsilon), t(\sigma_1^2 + \bar{\sigma}_1^2(\varepsilon))\big), \mathcal{N}\big(tb_2(\varepsilon), t(\sigma_2^2 + \bar{\sigma}_2^2(\varepsilon))\big)\Big).
\end{aligned}
$$

Hence, we need to control

- $\mathscr{W}_p\big(\mathcal{N}(m_1, \sigma_1^2), \mathcal{N}(m_2, \sigma_2^2)\big)$,
- $\mathscr{W}_p\Big(X_t^S(\varepsilon), \mathcal{N}\big(0, t\bar{\sigma}^2(\varepsilon)\big)\Big)$,
- $\mathscr{W}_p\Big(X_t^{1,B}(\varepsilon), X_t^{2,B}(\varepsilon)\Big)$.

Wasserstein distances between Gaussian random variables

Using the fact that

$$\mathscr{W}_2\big(\mathcal{N}(m_1,\sigma_1^2),\mathcal{N}(m_2,\sigma_2^2)\big)=\sqrt{(m_1-m_2)^2+(\sigma_1-\sigma_2)^2},$$

we get, for all $p\in[1,2]$

$$\mathscr{W}_p\big(\mathcal{N}(m_1,\sigma_1^2),\mathcal{N}(m_2,\sigma_2^2)\big)\leq\sqrt{(m_1-m_2)^2+(\sigma_1-\sigma_2)^2}.$$

Comparing small jumps with their Gaussian counterpart

Theorem (M., Reiß)

Let $X \sim (b, \sigma^2, \nu)$ *and let* $\bar\sigma^2(\varepsilon) = \int_{|x|\le\varepsilon} x^2 \nu(dx)$. *For any* $p \in [1,2]$, *there exists a positive constant* C, *such that*

$$\mathscr{W}_p\Big(\mathscr{L}(X_t^S(\varepsilon)), \mathcal{N}\big(0, t\bar\sigma^2(\varepsilon)\big)\Big) \le C \min\left(\sqrt{t}\bar\sigma(\varepsilon), \left(\frac{\int_{-\varepsilon}^{\varepsilon} |x|^{p+2}\nu(dx)}{\bar\sigma^2(\varepsilon)}\right)^{1/p}\right)$$
$$\le C \min\big(\sqrt{t}\bar\sigma(\varepsilon), \varepsilon\big).$$

In particular, for $p = 1$ *the bound is* $\min(2\sqrt{t}\bar\sigma(\varepsilon), \frac{\varepsilon}{2})$.

Comments

The inequality

$$\mathscr{W}_p\big(X_t^S(\varepsilon), \mathcal{N}(0, t\bar{\sigma}^2(\varepsilon))\big) \leq 2\sqrt{t}\bar{\sigma}(\varepsilon)$$

is clear from the definition of $\mathscr{W}_p$. The interest lies in the bound

$$\mathscr{W}_p\big(X_t^S(\varepsilon), \mathcal{N}(0, t\bar{\sigma}^2(\varepsilon))\big) \leq C\varepsilon,$$

which after renormalisation yields

$$\mathscr{W}_p\left(\frac{X_t^S(\varepsilon)}{\sqrt{t}\bar{\sigma}(\varepsilon)}, \mathcal{N}(0,1)\right) \leq \frac{C\varepsilon}{\sqrt{t}\bar{\sigma}(\varepsilon)}.$$

Optimality

Let $X(\varepsilon) \sim (0,0,\nu_\varepsilon)$, where $\nu_\varepsilon = \frac{\delta_{-\varepsilon}+\delta_\varepsilon}{2\varepsilon^2}$, i.e.

$$X_t(\varepsilon) = \sum_{i=1}^{N_t^\varepsilon} Y_i, \quad N_t^\varepsilon \sim \mathcal{P}(t/\varepsilon^2), \quad Y_i \sim F_\varepsilon := \frac{\delta_{-\varepsilon}+\delta_\varepsilon}{2}.$$

In particular, $\bar{\sigma}^2(\varepsilon) = 1$ and $\int_{-\varepsilon}^{\varepsilon} |x|^3 \nu_\varepsilon(dx) = \varepsilon$. Then,

$$\mathscr{W}_1\big(X_t(\varepsilon), \mathcal{N}(0,t)\big) \geq K \min(\sqrt{t}, \varepsilon), \text{ for some } K > 0.$$

Example

Let X be an α-stable Lévy process with $\nu(dx) \propto \frac{1}{|x|^{1+\alpha}}$, $\alpha \in (0,2)$.

For any $t > 0$ and $0 < \varepsilon \leq 1$

$$\mathscr{W}_p\left(\frac{X_t^S(\varepsilon)}{\bar{\sigma}(\varepsilon)}, \mathcal{N}(0,t)\right) \lesssim \min\left(\frac{\varepsilon}{\bar{\sigma}(\varepsilon)}, \sqrt{t}\right).$$

Here, $\bar{\sigma}^2(\varepsilon) \propto \frac{\varepsilon^{2-\alpha}}{2-\alpha}$. Therefore

$$\mathscr{W}_p\left(\frac{X_t^S(\varepsilon)}{\sqrt{t}\bar{\sigma}(\varepsilon)}, \mathcal{N}(0,1)\right) \lesssim \frac{\varepsilon^{\alpha/2}}{\sqrt{t}}.$$

Wasserstein distances for random sums of r.v.

Theorem (M., Reiß)

Let $(X_i)_{i\geq 0}$, $(Y_i)_{i\geq 0}$ be sequences of i.i.d. r. v. in $L_2(\mathbb{R})$ and N, N' be two positive integer-valued r.v. with N (resp. N') independent of $(X_i)_{i\geq 0}$ (resp. $(Y_i)_{i\geq 0}$). Denote by

$$Q_p := \mathscr{W}_p\bigg(\sum_{i=1}^{N} X_i, \sum_{i'=1}^{N'} Y_{i'}\bigg)^p.$$

Then:

$$Q_p \leq \begin{cases} \mathbb{E}[N]\mathscr{W}_1(X_1, Y_1) + \mathbb{E}[|Y_1|]\mathscr{W}_1(N, N') & \text{if } p = 1, \\ \mathbb{E}[N^p]\mathscr{W}_p(X_1, Y_1)^p + \mathscr{W}_p(N, N')^p\mathbb{E}[|Y_1|^p] & \text{if } 1 < p \leq 2. \end{cases}$$

Wasserstein distances for Poisson random variables

Proposition (M., Reiß)

Let N and N' be two Poisson variables of mean given by λ and λ', respectively. Denote by $m_{(p,\ell)}$ the moment of order p of a Poisson r.v. of mean ℓ. Then, for $p \geq 1$:

$$\mathscr{W}_p(N, N')^p \leq m_{(p,|\lambda-\lambda'|)}.$$

In particular,

$$\begin{aligned} \mathscr{W}_1(N, N') &\leq |\lambda - \lambda'| \\ \mathscr{W}_p(N, N')^p &\leq |\lambda - \lambda'| + |\lambda - \lambda'|^p, \quad 1 < p \leq 2. \end{aligned}$$

Wasserstein distances for the marginals of Lévy processes

Theorem (M., Reiß)

Let $X^j \sim (b_j, \sigma_j^2, \nu_j)$, $j = 1, 2$, be Lévy processes. For all $p \in [1, 2]$, for all $0 < \varepsilon \le 1$ and all $t \ge 0$, the following estimate holds

$$
\begin{aligned}
\mathscr{W}_p\big(X_t^1, X_t^2\big) \le t\big|b_1(\varepsilon) - b_2(\varepsilon)\big| &\\
&+ \sqrt{t\big(\sigma_1 + \bar{\sigma}_1(\varepsilon) - \sigma_2 - \bar{\sigma}_2(\varepsilon)\big)^2} \\
&+ C \sum_{j=1}^{2} \min\Big(\sqrt{t}\bar{\sigma}_j(\varepsilon), \varepsilon\Big) \\
&+ \mathscr{W}_p\big(X_t^{1,B}(\varepsilon), X_t^{2,B}(\varepsilon)\big),
\end{aligned}
$$

for a constant C only depending on p.

Distances for the increments of Lévy processes

In the Euclidean case $r = 2$, i.e. $d(x, y) := \left(\sum_{j=1}^n |x_j - y_j|^2\right)^{1/2}$:

$$\begin{aligned}
&\mathscr{W}_p\Big((X^1_{kT/n} - X^1_{(k-1)T/n})_{k=1}^n, (X^2_{kT/n} - X^2_{(k-1)T/n})_{k=1}^n\Big) \\
&\leq Tn^{-\frac{1}{2}}|b_1(\varepsilon) - b_2(\varepsilon)| + \sqrt{T}|\sigma_1 + \bar{\sigma}_1(\varepsilon) - \sigma_2 - \bar{\sigma}_2(\varepsilon)| \\
&+ C\sum_{j=1}^{2} \min(\sqrt{T}\bar{\sigma}_j(\varepsilon), \sqrt{n}\varepsilon) + \sqrt{n}\mathscr{W}_p(X^{1,B}_{T/n}(\varepsilon), X^{2,B}_{T/n}(\varepsilon)).
\end{aligned}$$

- The drift disappears as $n \to \infty$ (T fixed).
- The Gaussian part remains invariant.
- The Gaussian approximation of small jumps gives an error of order $\min(\bar{\sigma}_j(\varepsilon), \sqrt{n}\varepsilon)$.
- The bound on the larger jumps scales as $\sqrt{T}(\mathscr{W}_p(Y_1^{(1)}, Y_1^{(2)}) + \sqrt{|\lambda_1(\varepsilon) - \lambda_2(\varepsilon)|})$ (for $p = 1$ even as $T/\sqrt{n}$).

Toscani-Fourier distance

Let φ_1 (resp. φ_2) the characteristic function of P_1 (resp. P_2), i.e.

$$\varphi_1(u) = \int_{\mathbb{R}} e^{iux} P_1(dx).$$

Definition

For $s > 0$ the *Toscani-Fourier distance* of order s is defined as:

$$T_s(P_1, P_2) = \sup_{u \in \mathbb{R} \setminus \{0\}} \frac{|\varphi_1(u) - \varphi_2(u)|}{|u|^s}.$$

Proposition (M., Reiß)

For all $p \geq 1$,

$$\mathscr{W}_p(P_1, P_2) \geq \frac{1}{\sqrt{2}} T_1(P_1, P_2).$$

Wasserstein distances and Toscani-Fourier distance

Denote by φ_i the characteristic function of $X_t^i \sim (b_i, \sigma_i^2, \nu_i)$, i.e.

$$\varphi_i(u) = \exp\left(iut\left(b_i - \frac{\sigma_i^2 u^2}{2} - \int_{\mathbb{R}} \left(1 - e^{iux} + (iux\mathbb{I}_{|x|\le 1})\right)\nu_i(dx)\right)\right).$$

Thanks to the previous proposition, we have:

$$\mathscr{W}_p(X_t^1, X_t^2) \geq \frac{1}{\sqrt{2}} \sup_{u\in\mathbb{R}} \frac{|\varphi_1(u) - \varphi_2(u)|}{|u|}, \quad \forall p \geq 1.$$

Total variation distance

Let $(\mathcal{X}, \mathscr{F})$ be a measurable space and let μ and ν be two probability measures on $(\mathcal{X}, \mathscr{F})$.

Definition

The total variation distance between μ and ν is defined as

$$\|\mu - \nu\|_{TV} = \sup_{A \in \mathscr{F}} \big|\mu(A) - \nu(A)\big|.$$

The total variation distance has the following properties.

1. $\|\mu - \nu\|_{TV} = \frac{1}{2} L_1(\mu, \nu)$.
2. $\|\mu - \nu\|_{TV} = \inf \big(\mathbb{P}(X' \neq Y') : \text{law}(X') = \mu,\ \text{law}(Y') = \nu\big)$.
3. $\|\mu - \nu\|_{TV} = \frac{1}{2} \sup_{\|\Psi\|_\infty \leq 1} \big| \int_{\mathcal{X}} \Psi(x)(\mu - \nu)(dx)\big|$, the supremum being taken over all ψ: $\sup_{x \in \mathcal{X}} |\Psi(x)| \leq 1$.

Wasserstein distances and total variation distance

Theorem (See, e.g., Villani, 2009)

Let μ and ν be two probability measures on a Polish space $(\mathcal{X}, d)$. Let $p \in [1, \infty)$ and $x_0 \in \mathcal{X}$. Then

$$\mathscr{W}_p(\mu, \nu) \leq 2^{\frac{1}{p'}} \bigg(\int d(x_0, x)^p |\mu - \nu|(dx) \bigg)^{\frac{1}{p}}, \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

In particular, if $p = 1$ and the diameter of $\mathcal{X}$ is bounded by D, then

$$\mathscr{W}_1(\mu, \nu) \leq 2D \|\mu - \nu\|_{TV}.$$

C. Villani, Optimal transport: Old and new. Springer-Verlag, Berlin, 2009.

Wasserstein distance of order 1 and total variation

A real function g is of *bounded variation* if its total variation norm is finite, i.e. $\|g\|_{BV} = \sup_{P\in\mathscr{P}} \sum_{i=0}^{n_P-1} |g(x_{i+1}) - g(x_i)| < \infty$, where the supremum is taken over the set $\mathscr{P}$ of partitions of $\mathbb{R}$.

Proposition (M., Reiß)

Let μ and ν be two probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and G be a measure admitting a density g with respect to the Lebesgue measure. Then,

$$\|\mu * G - \nu * G\|_{TV} \leq \frac{\|g\|_{BV}}{2} \mathscr{W}_1(\mu,\nu).$$

Gaussian approximation in Total Variation distance

Corollary (M., Reiß)

Let X be a Lévy process with Lévy measure ν. Then, for all $t > 0$, $0 < \varepsilon \leq 1$ and $\Sigma > 0$ we have:

$$\Big\| X_t^S(\varepsilon) * \mathcal{N}(0, t\Sigma^2), \mathcal{N}(0, t\bar{\sigma}^2(\varepsilon)) * \mathcal{N}(0, t\Sigma^2) \Big\|_{TV} \leq \sqrt{\frac{2}{\pi t \Sigma^2}} \min\Big(2\sqrt{t\bar{\sigma}^2(\varepsilon)}, \frac{\varepsilon}{2} \Big).$$

Total variation for the increments of Lévy processes

Theorem (M., Reiß)

Let $X^i \sim (b_i, \sigma_i^2, \nu_i)$, $i = 1, 2$. For all $t > 0$, $0 < \varepsilon \le 1$ and $\sigma_i > 0$, $i = 1, 2$, we have:

$$\left\|X_t^1 - X_t^2\right\|_{TV} \le \sqrt{\frac{t}{2\pi}}\frac{|b_1(\varepsilon) - b_2(\varepsilon)|}{\sqrt{\sigma_2^2 + \bar\sigma_2^2(\varepsilon)}} + \sqrt{2}\left|1 - \frac{\sqrt{\sigma_1^2 + \bar\sigma_1^2(\varepsilon)}}{\sqrt{\sigma_2^2 + \bar\sigma_2^2(\varepsilon)}}\right|$$
$$+ \sum_{i=1}^{2}\sqrt{\frac{2}{\pi t\sigma_i^2}}\min\left(\sqrt{2}\sqrt{t\bar\sigma_i^2(\varepsilon)}, \frac{\varepsilon}{2}\right)$$
$$+ t|\lambda_1(\varepsilon) - \lambda_2(\varepsilon)| + t\min\left(\lambda_1(\varepsilon), \lambda_2(\varepsilon)\right)\left\|\frac{\nu_1^\varepsilon}{\lambda_1(\varepsilon)} - \frac{\nu_2^\varepsilon}{\lambda_2(\varepsilon)}\right\|_{TV},$$

with $\nu_j^\varepsilon = \nu_j(\cdot \cap (\mathbb{R} \setminus (-\varepsilon, \varepsilon)))$ and $\lambda_j(\varepsilon) = \nu_j^\varepsilon(\mathbb{R})$.

Total variation distance and T_j

Let μ, ν and G be real probability measures. Suppose that G is absolutely continuous wrt Lebesgue with density g.

Proposition (M., Reiß)

Suppose that g is j-times weakly differentiable with jth derivative $g^{(j)} \in L^2(\mathbb{R})$. Then,

$$\|\mu * G - \nu * G\|_{TV} \leq C_j^{1/(2j+1)} \|g^{(j)}\|_2^{2j/(2j+1)} T_j(\mu, \nu)^{2j/(2j+1)},$$

where $C_j = \max\left\{\mathbb{E}\left[|X + Y|^j\right], \mathbb{E}\left[|Z + Y|^j\right]\right\}$ with $X \sim \mu$, $Y \sim G$, $Z \sim \nu$ and Y independent of X and Z.

Total variation for the increments of Lévy processes

Let $X^i \sim (b_i, \sigma_i^2, \nu_i)$, with $\sigma_i \neq 0$, $i = 1, 2$. Then, for any $\Sigma \in (0, \sigma_1 \wedge \sigma_2)$

$$X_t^i = \mathcal{N}(0, t\Sigma^2) + \widetilde{X}_t^i, \text{ where } \widetilde{X}^i \sim (b_i, \sigma_i^2 - \Sigma^2, \nu_i).$$

Theorem (M., Reiß)

$$\left\| X_t^1 - X_t^2 \right\|_{TV} \leq \frac{\left(\sqrt{C}\, T_1(\widetilde{X}_t^1, \widetilde{X}_t^2)\right)^{\frac{2}{3}}}{(16\pi)^{\frac{1}{3}} \sqrt{t}\Sigma},$$

with $C = \max\left(\mathbb{E}\left|X_t^1\right|, \mathbb{E}\left|X_t^2\right|\right)$.

Problem: how to simplify Jacod-Reiß lower bound

$X \sim (B, C, \nu)$ Itô-semimartingale

$$B_t = \int_0^t b_s ds, \quad C_t = \int_0^t c_s ds, \quad \nu(dt, dx) = dtF_t(dx).$$

Goal: Estimate C observing $X_{\frac{i}{n}}$, $i = 0, \dots, n$ when X in

$$\mathcal{S}_A^r = \text{the set of all Itô-semimartingales with}$$
$$|b_t| + c_t + \int (|x|^r \wedge 1) F_t(dx) \leq A \text{ for all } t.$$

J. Jacod, M. Reiß, A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. Annals of Statistics, 2014.

Problem: how to simplify Jacod-Reiß lower bound

Goal: Prove that any uniform rate w_n for estimating $C(X)_1$ satisfies

$$\Psi_n \leq (n \log n)^{-(2-r)/2} \quad \text{if} \quad r > 1.$$

Strategy: Find $X^i \sim (b_i, \sigma_i^2, F_i)$ such that

- $\sigma_1^2 - \sigma_2^2 = a_n := (n \log n)^{-(2-r)/2}$, $r \in (0,2)$,
- $\int (|x|^r \wedge 1) F_i(dx) \leq K$,
- $\|\mathscr{L}((X^1_{i/n} - X^1_{(i-1)/n})_{1 \leq i \leq n}) - \mathscr{L}((X^2_{i/n} - X^2_{(i-1)/n})_{1 \leq i \leq n})\|_{TV} \to 0$ as $n \to \infty$.

The construction in the paper of Jacod and Reiß is very involved. But now, with our results, it is much easier to prove the desired lower bound.

J. Jacod, M. Reiß, A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. Annals of Statistics, 2014.

A solution

Let $X \sim (0, 1 + a_n, F_n)$ and $Y \sim (0, 1, G_n)$ with F_n and G_n s.t.

$$\mathbb{E}\Big[e^{iuX_{1/n}}\Big] = \exp\Big(-\frac{u^2}{2n}(1 + a_n) - \frac{\Psi_n(u)}{n}\Big),$$
$$\mathbb{E}\Big[e^{iuY_{1/n}}\Big] = \exp\Big(-\frac{u^2}{2n} - \frac{\Phi_n(u)}{n}\Big),$$

where Φ_n and Ψ_n are real positive functions s.t.

$$\mathbb{E}\Big[e^{iuX_{1/n}}\Big] = \mathbb{E}\Big[e^{iuY_{1/n}}\Big] \quad \forall |u| < u_n := 2\sqrt{n \log n}.$$

A solution

$$X_{\frac{1}{n}} = \mathcal{N}\Big(0, \frac{1}{8n}\Big) + \widetilde{X}_{\frac{1}{n}}, \quad \widetilde{X} \sim \Big(0, \frac{7}{8} + a_n, F_n\Big),$$
$$Y_{\frac{1}{n}} = \mathcal{N}\Big(0, \frac{1}{8n}\Big) + \widetilde{Y}_{\frac{1}{n}}, \quad \widetilde{Y} \sim \Big(0, \frac{7}{8}, G_n\Big).$$

Let $\mu_n \sim \mathscr{L}(\widetilde{X}_{1/n})$, $\nu_n \sim \mathscr{L}(\widetilde{Y}_{1/n})$ and $G_n \sim \mathcal{N}\big(0, \frac{1}{8n}\big)$. Then,

$$\begin{aligned}\|\mathscr{L}(X_{1/n}) - \mathscr{L}(Y_{1/n})\|_{TV} &= \|\mu_n * G_n - \nu_n * G_n\|_{TV} \\ &\lesssim \Big(n^{3/4} T_1(\mu_n, \nu_n)\Big)^{2/3},\end{aligned}$$

Computing $T_1(\mu_n, \nu_n)$

We have

$$
\begin{aligned}
T_1(\mu_n,\nu_n) &= \sup_{u\in\mathbb{R}} \frac{\left|\exp\left(-\frac{u^2}{2n}\left(\frac{7}{8}+a_n\right)-\frac{\Psi_n(u)}{n}\right)-\exp\left(-\frac{7u^2}{16n}-\frac{\Phi_n(u)}{n}\right)\right|}{|u|}\\
&= \sup_{|u|>u_n} \frac{\exp\left(-\frac{7u^2}{16n}\right)\left|\exp\left(-\frac{u^2a_n}{2n}-\frac{\Psi_n(u)}{n}\right)-\exp\left(-\frac{\Phi_n(u)}{n}\right)\right|}{|u|}\\
&\le \frac{\exp\left(-\frac{7u_n^2}{16n}\right)}{u_n} = \frac{\exp\left(-\frac{7\times 4n\log n}{16n}\right)}{2\sqrt{n\log n}} = \frac{n^{-9/4}}{2\sqrt{\log n}}.
\end{aligned}
$$

A solution

Hence,

$$\|\mathscr{L}(X_{1/n}) - \mathscr{L}(Y_{1/n})\|_{TV} \lesssim \frac{n^{-1}}{(\log n)^{1/3}}.$$

Therefore,

$$\begin{aligned}\|\mathscr{L}((X_{i/n} - X_{(i-1)/n})_i) - \mathscr{L}((Y_{i/n} - Y_{(i-1)/n})_i)\|_{TV} &\leq 2\sqrt{n\|\mathscr{L}(X_{1/n}) - \mathscr{L}(Y_{1/n})\|_{TV}} \\ &\lesssim (\log n)^{-1/6} \to 0.\end{aligned}$$

To sum up

- Upper and lower bounds for the Wasserstein distances between the increments of two Lévy processes in terms of their Lévy triplets.
- A fine control for the Wasserstein distances between the marginal of the martingale associated with the small jumps of a Lévy process and a Gaussian distribution.
- Upper bound for the total variation distance between the marginals of any pair of Lévy processes with a non zero Gaussian component.
- Application to statistical lower bounds.

E. Mariucci, M. Reiß, Wasserstein and total variation distance between marginals of Lévy processes, to appear in Electronic Journal of Statistics.

Perspectives

Joint work in progress with A. Carpentier and C. Duval:

- Gaussian approximation of the small jumps of a Lévy process in total variation.
- A sharp upper bound for the total variation distance between the marginals of any pair of Lévy processes.

Still to be done:

- Generalization in dimension d.
- Go beyond Lévy processes: additive processes, semimartingales, ...

Perspectives

Joint work in progress with A. Carpentier and C. Duval:

- Gaussian approximation of the small jumps of a Lévy process in total variation.
- A sharp upper bound for the total variation distance between the marginals of any pair of Lévy processes.

Still to be done:

- Generalization in dimension d.
- Go beyond Lévy processes: additive processes, semimartingales, ...

Thank you for your attention!