The Contraction Method on $\mathcal{C}([0,1])$ and Donsker's Theorem

Henning Sulzbach
J. W. Goethe-Universität Frankfurt a. M.

9^{th} German Open Conference on Probability and Statistics
Leipzig, March 2, 2010

joint work with Ralph Neininger

Contraction method for recursive stochastic processes

Contraction method for recursive stochastic processes

Usual situation

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} \circ X^r_{I_r^{(n)}} + b^{(n)}$$

with $r.v.$ $(X_n), b^{(n)}$ taking values in some function space S, $A_r^{(n)}$ random functions from S to S, and independent copies $(X_n^1), \dots, (X_n^K)$ of (X_n).

Contraction method for recursive stochastic processes

Usual situation

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} \circ X^r_{I_r^{(n)}} + b^{(n)}$$

with $r.v.$ $(X_n), b^{(n)}$ taking values in some function space S, $A_r^{(n)}$ random functions from S to S, and independent copies $(X_n^1), \ldots, (X_n^K)$ of (X_n).

If $A_r^{(n)} \to A_r$ and $b^{(n)} \to b$ for some S valued processes A_r, b, this suggests

$$X_n \to X,$$

where X solves $X \stackrel{d}{=} \sum_{r=1}^{K} A_r \circ X^{(r)} + b$ (uniquely).

Example - The FIND algorithmn

Example - The FIND algorithmn

- Limit (FIND-) process in $D([0,1])$ by an almost sure construction [Grübel, Rösler '96],

Example - The FIND algorithmn

- Limit (FIND-) process in $D([0,1])$ by an almost sure construction [Grübel, Rösler '96],
- Adapted version in [Knof, Rösler '08+],

Example - The FIND algorithmn

- Limit (FIND-) process in $D([0,1])$ by an almost sure construction [Grübel, Rösler '96],
- Adapted version in [Knof, Rösler '08+],
- Contraction method in $D([0,1])$ equipped with (weaker) L^p-topology in [Eickmeyer, Rüschendorf '07]

Example - Profile of Search Trees

Example - Profile of Search Trees

- Martingale methods ([Drmota, Chauvin, Jabbour-Hattab '01], [Chauvin, Klein, Marckert, Rouault '05])

Example - Profile of Search Trees

- Martingale methods ([Drmota, Chauvin, Jabbour-Hattab '01], [Chauvin, Klein, Marckert, Rouault '05])
- Contraction method in function space S of square-integrable analytic functions

$$f : D \subseteq \mathbb{C} \to \mathbb{C}$$

equipped with norm

$$||f|| = \left(\int_D |f(z)|^2 d\lambda^2(z) \right)^{\frac{1}{2}}$$

S is a separable Hilbert space. [Drmota, Janson, Neininger '08]

Donsker's Theorem

Let $X_1, X_2, \ldots$ be iid r.v. with $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1$ and $\mathbb{E}|X_1|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$.

Donsker's Theorem

Let $X_1, X_2, \ldots$ be iid r.v. with $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1$ and $\mathbb{E}|X_1|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$.

$$S^n = (S_t^n)_{t\in[0,1]}$$

with

$$S_t^n = \frac{1}{\sqrt{n}} \left(\sum_{k=1}^{\lfloor nt \rfloor} X_k + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right), \quad t \in [0,1]$$

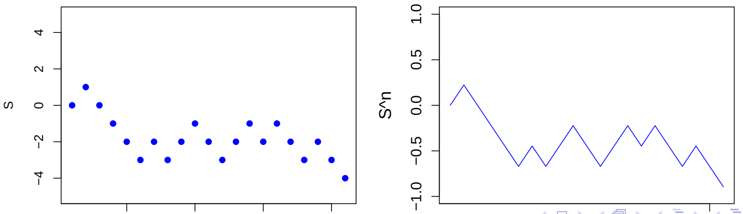
Donsker's Theorem

Let $X_1, X_2, \ldots$ be iid r.v. with $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1$ and $\mathbb{E}|X_1|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$.

$$S^n = (S^n_t)_{t\in[0,1]}$$

with

$$S^n_t = \frac{1}{\sqrt{n}} \left(\sum_{k=1}^{\lfloor nt \rfloor} X_k + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right), \quad t \in [0,1]$$



Donsker's Theorem

Donsker's Theorem

$$S_t^n = \frac{1}{\sqrt{n}} \left(\sum_{k=1}^{\lfloor nt \rfloor} X_k + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right), \quad t \in [0,1]$$

Theorem (Donsker, 1951)

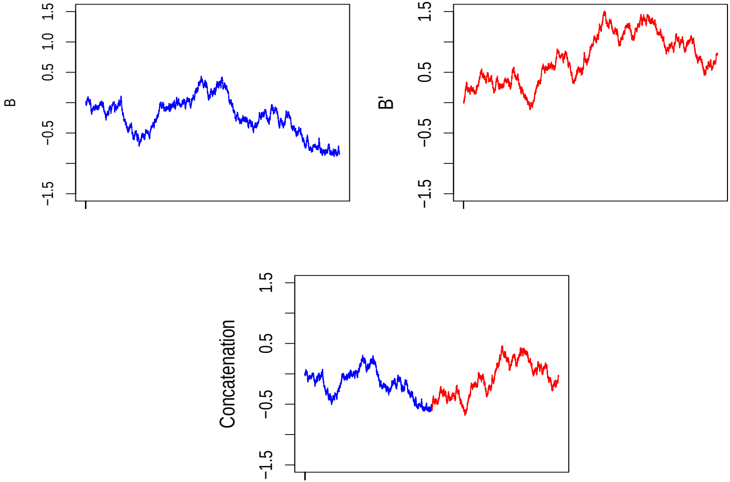
$S^n \xrightarrow{d} B$ in $(\mathcal{C}([0,1]), ||\cdot||_{\text{sup}})$, where B is a standard Brownian Motion.

Decomposition of Brownian Motion

In time:

Decomposition of Brownian Motion

In time:



Decomposition of Brownian Motion

$$
\begin{aligned}
(B_t)_t &\stackrel{d}{=} \left(\frac{1}{\sqrt{2}}\left(\mathbf{1}_{\{t\leq 1/2\}}B_{2t} + \mathbf{1}_{\{t>1/2\}}B_1\right) + \frac{1}{\sqrt{2}}\mathbf{1}_{\{t>1/2\}}B'_{2t-1}\right)_t \\
&= \frac{1}{\sqrt{2}}\varphi_2(B) + \frac{1}{\sqrt{2}}\psi_2(B'), \qquad (1)
\end{aligned}
$$

Decomposition of Brownian Motion

$$
\begin{aligned}
(B_t)_t &\stackrel{d}{=} \left(\frac{1}{\sqrt{2}} \left(\mathbf{1}_{\{t \le 1/2\}} B_{2t} + \mathbf{1}_{\{t > 1/2\}} B_1 \right) + \frac{1}{\sqrt{2}} \mathbf{1}_{\{t > 1/2\}} B'_{2t-1} \right)_t \\
&= \frac{1}{\sqrt{2}} \varphi_2(B) + \frac{1}{\sqrt{2}} \psi_2(B'), \qquad (1)
\end{aligned}
$$

with (continuous and linear) functions $\varphi_\beta, \psi_\beta : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$

$$
\begin{aligned}
\varphi_\beta(f)(t) &= \mathbf{1}_{\{t \le 1/\beta\}} f(\beta t) + \mathbf{1}_{\{t > 1/\beta\}} f(1), \\
\psi_\beta(f)(t) &= \mathbf{1}_{\{t \le 1/\beta\}} f(0) + \mathbf{1}_{\{t > 1/\beta\}} f\left(\frac{\beta t - 1}{\beta - 1} \right).
\end{aligned}
$$

Decomposition of Brownian Motion

$$
\begin{aligned}
(B_t)_t &\stackrel{d}{=} \left(\frac{1}{\sqrt{2}} \left(\mathbf{1}_{\{t \le 1/2\}} B_{2t} + \mathbf{1}_{\{t > 1/2\}} B_1 \right) + \frac{1}{\sqrt{2}} \mathbf{1}_{\{t > 1/2\}} B'_{2t-1} \right)_t \\
&= \frac{1}{\sqrt{2}} \varphi_2(B) + \frac{1}{\sqrt{2}} \psi_2(B'), \qquad (1)
\end{aligned}
$$

with (continuous and linear) functions $\varphi_\beta, \psi_\beta : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$

$$
\begin{aligned}
\varphi_\beta(f)(t) &= \mathbf{1}_{\{t \le 1/\beta\}} f(\beta t) + \mathbf{1}_{\{t > 1/\beta\}} f(1), \\
\psi_\beta(f)(t) &= \mathbf{1}_{\{t \le 1/\beta\}} f(0) + \mathbf{1}_{\{t > 1/\beta\}} f\left(\frac{\beta t - 1}{\beta - 1} \right).
\end{aligned}
$$

Theorem (Uniqueness)

Let $X = (X_t)_{t \in [0,1]}$ be a continuous real-valued stochastic process satisfying (1). Then there exists a constant $\sigma \ge 0$ such that σX is a standard Brownian Motion.

Theorem (Uniqueness)

Let $X = (X_t)_{t \in [0,1]}$ be a continuous real-valued stochastic process satisfying (1). Then there exists a constant $\sigma \geq 0$ such that σX is a standard Brownian Motion.

Theorem (Uniqueness)

Let $X = (X_t)_{t\in[0,1]}$ be a continuous real-valued stochastic process satisfying (1). Then there exists a constant $\sigma \geq 0$ such that σX is a standard Brownian Motion.

The Theorem follows directly from

Lemma

A real valued random variable X is centered and normally distributed if and only if it holds

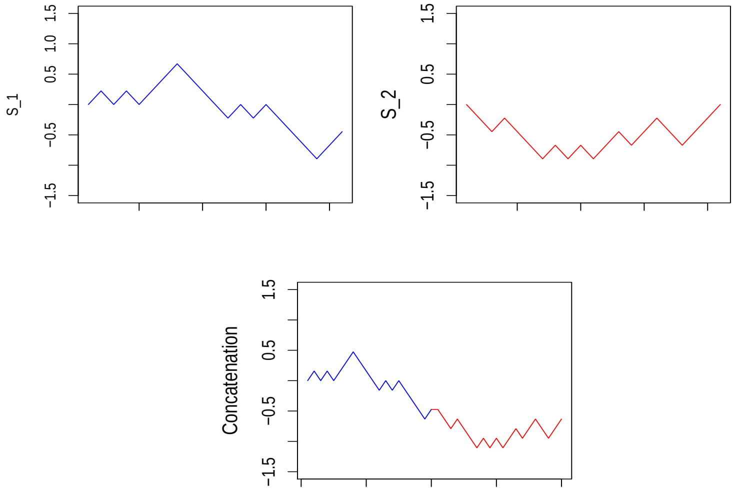
$$X \stackrel{d}{=} \frac{X + X'}{\sqrt{2}}$$

where X' is an independent copy of X.

Proof: e.g. Lukacs [75]

Time-decomposition of the random walk

Time-decomposition of the random walk



Time-decomposition of the random walk

It holds:

$$S^n \stackrel{d}{=} \sqrt{\frac{\lceil n/2 \rceil}{n}} \varphi_{\frac{n}{\lceil n/2 \rceil}} \left(S^{\lceil n/2 \rceil} \right) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \psi_{\frac{n}{\lceil n/2 \rceil}} \left(\widehat{S}^{\lfloor n/2 \rfloor} \right), \quad (2)$$

Time-decomposition of the random walk

It holds:

$$S^n \stackrel{d}{=} \sqrt{\frac{\lceil n/2 \rceil}{n}} \varphi_{\frac{n}{\lceil n/2 \rceil}} \left(S^{\lceil n/2 \rceil} \right) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \psi_{\frac{n}{\lceil n/2 \rceil}} \left(\widehat{S}^{\lfloor n/2 \rfloor} \right), \quad (2)$$

Now

$$\sqrt{\frac{\lceil n/2 \rceil}{n}} \to \frac{1}{\sqrt{2}}, \quad \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \to \frac{1}{\sqrt{2}}.$$

suggests weak convergence $S^n \to B$.

The Zolotarev metric on a Banach space

The Zolotarev metric on a Banach space

Arbitrary Banach space $(\mathcal{B}, ||\cdot||)$ instead of $\mathcal{C}([0,1])$.
Let $\mathcal{M}(\mathcal{B})$ be the set of probability measures on $\mathcal{B}$. For $\mu, \nu \in \mathcal{M}(\mathcal{B})$ and $s > 0$ define the Zolotarev distance of μ and ν by

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}[f(X) - f(Y)]|,$$

with $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$

The Zolotarev metric on a Banach space

Arbitrary Banach space $(\mathcal{B}, ||\cdot||)$ instead of $\mathcal{C}([0,1])$.
Let $\mathcal{M}(\mathcal{B})$ be the set of probability measures on $\mathcal{B}$. For $\mu, \nu \in \mathcal{M}(\mathcal{B})$ and $s > 0$ define the Zolotarev distance of μ and ν by

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}[f(X) - f(Y)]|,$$

with $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ and for $s = m + \alpha$ with $0 < \alpha \leq 1$ and $m := \lceil s \rceil - 1 \in \mathbb{N}_0$

$$\mathcal{F}_s := \{f \in C^m(B, \mathbb{R}) : \|D^m f(x) - D^m f(y)\| \leq \|x - y\|^\alpha, \quad x, y \in B\}.$$

The Zolotarev metric on a Banach space

Arbitrary Banach space $(\mathcal{B}, ||\cdot||)$ instead of $\mathcal{C}([0,1])$.
Let $\mathcal{M}(\mathcal{B})$ be the set of probability measures on $\mathcal{B}$. For $\mu, \nu \in \mathcal{M}(\mathcal{B})$ and $s > 0$ define the Zolotarev distance of μ and ν by

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}[f(X) - f(Y)]|,$$

with $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ and for $s = m + \alpha$ with $0 < \alpha \leq 1$ and $m := \lceil s \rceil - 1 \in \mathbb{N}_0$

$$\mathcal{F}_s := \{f \in C^m(B, \mathbb{R}) : \|D^m f(x) - D^m f(y)\| \leq \|x-y\|^\alpha, \quad x, y \in B\}.$$

Set $\zeta_s(X, Y) = \zeta_s(\mathcal{L}(X), \mathcal{L}(Y))$.

Properties of the Zolotarev distance ζ_s

Properties of the Zolotarev distance ζ_s

It holds $\zeta_s(X, Y) < \infty$, if

$$\mathbb{E}||X||^s, \mathbb{E}||Y||^s < \infty, \quad \mathbb{E}[g(X, \dots, X)] = \mathbb{E}[g(Y, \dots, Y)]$$

for all $k \leq m$ and multilinear, bounded functions $g : B^k \to \mathbb{R}$. In the following assume finiteness of the considered ζ_s-distances.

Properties of the Zolotarev distance ζ_s

Lemma

ζ_s is $(s,+)$-ideal , i.e.

$$\zeta_s(\varphi(X), \varphi(Y)) \leq ||\varphi||^s \zeta_s(X, Y)$$

for any continuous and linear function $\varphi : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ with

$$||\varphi|| = \sup_{f \in \mathcal{C}([0,1]), ||f||=1} ||\varphi(f)||.$$

Furthermore

$$\zeta_s(X_1 + X_2, Y_1 + Y_2) \leq \zeta_s(X_1, Y_1) + \zeta_s(X_2, Y_2)$$

for (X_1, Y_1) and (X_2, Y_2) independent.

Proof: Zolotarev ['77]

Properties of the Zolotarev distance ζ_s

Theorem

Let B be a separable Hilbert space and $T \subseteq \mathcal{M}(B)$ be a subset of probability measures with

- $\mathbb{E}||X||^s < \infty$ *for all r.v.* X *with* $\mathcal{L}(X) \in T$*,*
- $\mathbb{E}[f(X, \ldots, X)] = \mathbb{E}[f(Y, \ldots, Y)]$ *for all r.v.* X, Y *with* $\mathcal{L}(X), \mathcal{L}(Y) \in T$ *and continuous, multilinear functions* $f : B^k \to R$ *for all* $0 \leq k < s$.

Then ζ_s *is a complete metric on* T *and convergence in* ζ_s *in* T *implies weak convergence.*

Proof: Drmota, Janson, Neininger ('08)

Properties of ζ_s in $\mathcal{C}([0,1])$

Properties of ζ_s in $\mathcal{C}([0,1])$

Theorem

Let $B = \mathcal{C}([0,1])$ *and* $0 < s \leq 3$*, i.e,* $m \in \{0,1,2\}$*. Let* $(X_n)_{n\geq 1}$*, X be random variables in* $(\mathcal{C}([0,1]), \|\cdot\|)$ *where, for each* $n \geq 1$*,* X_n *is piecewise linear with intervals of length at least* r_n*. If*

$$\zeta_s(X_n, X) = o\left(\log^{-m} \frac{1}{r_n}\right), \quad as \quad n \to \infty,$$

then $X_n \to X$ *in distribution.*
Furthermore $\zeta_s(X, Y) = 0$ *implies* $\mathcal{L}(X) = \mathcal{L}(Y)$ *for any r.v.* X, Y *in* $\mathcal{C}([0,1])$*.*

Properties of ζ_s in $\mathcal{C}([0,1])$

Theorem

Let $B = \mathcal{C}([0,1])$ and $0 < s \leq 3$, i.e, $m \in \{0,1,2\}$. Let $(X_n)_{n \geq 1}$, X be random variables in $(\mathcal{C}([0,1]), \|\cdot\|)$ where, for each $n \geq 1$, X_n is piecewise linear with intervals of length at least r_n. If

$$\zeta_s(X_n, X) = o\left(\log^{-m} \frac{1}{r_n}\right), \quad \text{as} \quad n \to \infty,$$

then $X_n \to X$ in distribution.
Furthermore $\zeta_s(X,Y) = 0$ implies $\mathcal{L}(X) = \mathcal{L}(Y)$ for any r.v. X, Y in $\mathcal{C}([0,1])$.

Proof: This basically follows from a result of Barbour from the context of Stein's method.

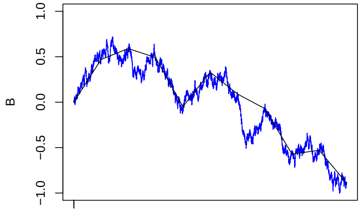
Proof of Donsker's Theorem (Sketch)

We are in the case $2 < s < 3$ and have to consider $\mathbb{E}[f(X, X)]$ for continuous, bilinear functions $f : \mathcal{C}([0,1])^2 \to \mathbb{R}$.
This is done by controlling the covariance function.

Proof of Donsker's Theorem (Sketch)

We are in the case $2 < s < 3$ and have to consider $\mathbb{E}[f(X,X)]$ for continuous, bilinear functions $f : \mathcal{C}([0,1])^2 \to \mathbb{R}$.
This is done by controlling the covariance function.
Since S^n and B do not share their covariance function (of course $\mathbb{E}[S^n_s, S^n_t] \to \min(s,t)$) we also consider the process B^n defined by

$$B^n_t = B_{\frac{\lfloor nt \rfloor}{n}} + (nt - \lfloor nt \rfloor)\left(B_{\frac{\lfloor nt \rfloor + 1}{n}} - B_{\frac{\lfloor nt \rfloor}{n}}\right).$$



Proof of Donsker's Theorem (Sketch)

An contraction argument shows

$$\zeta_{2+\varepsilon}(S^n, B^n) = O(n^{-\delta})$$

for $\delta < \varepsilon/2$.

Proof of Donsker's Theorem (Sketch)

An contraction argument shows

$$\zeta_{2+\varepsilon}(S^n, B^n) = O(n^{-\delta})$$

for $\delta < \varepsilon/2$. Together with $||B^n - B|| \to 0 \quad a.s.$

Proof of Donsker's Theorem (Sketch)

An contraction argument shows

$$\zeta_{2+\varepsilon}(S^n, B^n) = O(n^{-\delta})$$

for $\delta < \varepsilon/2$. Together with $||B^n - B|| \to 0 \quad a.s.$ Donsker's Theorem follows with

Theorem

Suppose $0 < s \leq 3$, *i.e,* $m \in \{0, 1, 2\}$. *Let* $(X_n)_{n\geq 1}$, $(Y_n)_{n\geq 1}$, Z *be random variables in* $(\mathcal{C}([0,1]), \|\cdot\|)$ *where, for each* $n \geq 1$, X_n, Y_n *are piecewise linear with intervals of length at least* r_n. *If* $Y_n \to Z$ *in distribution and*

$$\zeta_s(X_n, Y_n) = o\left(\log^{-m} \frac{1}{r_n}\right), \quad as \quad n \to \infty,$$

then $X_n \to Z$ *in distribution.*