Entropy on Wiener Space

Hans Föllmer
Humboldt University Berlin

23. May 2019, Paris

Conférence en l'honneur de
Nicole El Karoui

60/70's :

Séminaire de Probabilités
Strasbourg / Paris

Nicole : VII, IX, XI, ...

- Probabilistic Potential theory

1971 : thèse (with A. Reinhard, B. Roynette) :
processus de Hunt / standard

- "théorie générale"

- contrôle stochastique
(Ecole d'Eté de St. Flour 1979)

Ecole d'Eté de St. Flour 1986
(F., Papanicolau, Barndorff-Nielsen)

Since late 80's :

Probability $\rightleftarrows$ Finance

a great intellectual adventure ("high risk": uncharted territory, controversial, ...!)

"$\longleftarrow$" :

- new decomposition theorems ("universal" Doob-Meyer, non-linear Riesz)
- BSDE
- risk measures

"$\longrightarrow$" :

- tremendous impact via DEA / Master "Probabilités et Finance"

1988/89: Sabbatical at BNP Paribas / Caisse des Dépots

- focus on the risks of longevity (ongoing)

In this talk: some connections between St. Flour 1986 and the

"entropic risk measure"

$$\rho_\beta(F) := \frac{1}{\beta} \log E_P[e^{-\beta F}]$$

(name coined by Nicole)

$$= \sup_{Q:\, H(Q|P)<\infty} \left(E_Q[-F] - \frac{1}{\beta} H(Q|P) \right)$$

for $F \in L^\infty(\Omega, \mathcal{F}, P)$

where

$$H(Q|P) := \begin{cases} E_Q\left[\log \frac{dQ}{dP}\right] & \text{if } Q \ll P \\ +\infty & \text{if not} \end{cases}$$

= relative entropy of Q w.r.t. P

In this talk: P = Wiener measure

" $\longleftarrow$ "

1. Relative Entropy and Optimal Transport on Wiener Space

$\Omega = C_0[0,1]$, $\quad \omega_t := \omega(t)$

$\mathcal{H} :=$ Cameron-Martin space

$$\|\omega\|_{\mathcal{H}} := \begin{cases} \left(\int_0^1 \dot{\omega}(t)^2 dt\right)^{\frac{1}{2}} & \text{if } \omega \text{ is absolutely continuous with } \dot{\omega} \in L^2[0,1] \\ +\infty & \text{else} \end{cases}$$

$P =$ Wiener measure on $(\Omega, \mathcal{F})$

$H(Q|P) < \infty$

$\Longrightarrow$ St. Flour $\quad \exists$ "intrinsic drift" $(b_t^Q)_{0 \le t \le 1}$

i) $\dfrac{dQ}{dP} = \exp\left[\int_0^1 b_t^Q(\omega)\, d\omega - \frac{1}{2}\int_0^1 (b_t^Q)^2(\omega)\, dt\right]$

ii) $d\omega = d\omega^Q + b^Q(\omega)\, dt \quad (*)$

$\Longrightarrow$

$$H(Q|P) = \frac{1}{2} E_Q\left[\int_0^1 (b_t^Q)^2(\omega)\, dt\right]$$

Remark:

$Q_s := \{ Q \mid H(Q|P) < \infty$ and

(*) has unique strong solution

given by $\phi : C_0[0,1] \to C_0[0,1] \}$

"adapted"

for $Q \in Q_s$: $\omega = \phi(\omega^Q)$,

$\hat{b}_t^Q := b_t^Q \circ \phi$

$\Rightarrow$

$$H(Q|P) = \frac{1}{2} E_P\Big[\underbrace{\int_0^1 (\hat{b}_t^Q)^2 dt}_{= \|\hat{B}^Q\|_{\mathcal{H}}^2} \Big]$$

In fact (Boué-Dupuis 1998, Lehec 2013)

$$S_P(F) = \sup_{Q \in Q_s} \left(E_Q[-F] - \frac{1}{2} H(Q|P) \right)$$

hence

$$= \sup_{\substack{\hat{B} \in \mathcal{H} \\ \text{adapt.}}} E_P\left[-F(\omega + \hat{B}) - \frac{1}{2} \|\hat{B}\|_{\mathcal{H}}^2 \right]$$

Connection to optimal transport on Wiener space via

Talagrand's inequality (1996):

I countable, Q_I and $P_I = \prod_{i \in I} N(0,1)$ on $\mathbb{R}^I$ $\Longrightarrow$

Wasserstein distance

$$W(Q_I, P_I) := \inf_{\hat{P}} E_{\hat{P}}\Big[\underbrace{\|X_I - Y_I\|^2}_{:= \sum_{i\in I}(X_i - Y_i)^2}\Big]^{\frac{1}{2}}$$

"coupling" of Q_I, P_I $\left\{ \begin{array}{l} X_I = (X_i)_{i\in I} \\ Y_I = (Y_i)_{i \in I} \\ \text{on some } (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}) \\ X_I \sim P_I,\ Y_I \sim Q_I \end{array} \right.$

$$\leq \sqrt{2H(Q_I | P_I)}$$

$|I| = 1$: integration by parts

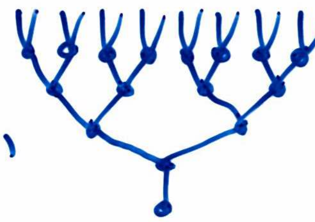
$|I| < \infty$: by induction, conditionally

$|I| = \infty$: by martingale convergence

– translates to Wiener space via Lévy-Ciesielski:

I = binary tree

$(e_i)_{i\in I}$ = Schauder functions



$$C_0[0,1] \ni \omega = \sum_{i\in I} X_i(\omega)\, e_i \;\hat{=}\; (X_i(\omega))_{i\in I} \in \mathbb{R}^I$$

$$\|\omega\|^2_{\mathcal{H}} = \sum_{i\in I} X_i^2(\omega)$$

Wiener measure P on $C_0[0,1] \;\hat{=}\; P_I := \prod_{i\in I} N(0,1)$

$\qquad\qquad Q \qquad$ " $\qquad \hat{=}\; Q_I$

$\Longrightarrow$

$$W^2_{\mathcal{H}}(Q,P) := \inf_{\substack{\text{couplings}\\ (X,Y)\text{ of } P,Q}} E_{\hat{P}}\Big[\underbrace{\|Y-X\|^2_{\mathcal{H}}}_{=\sum_{i\in I}(Y_i-X_i)^2}\Big]$$

$$\underset{\substack{\text{Talagrand}\\ \text{on } \mathbb{R}^I}}{\le} \; 2\,H(Q_I|P_I) = 2\,H(Q|P)$$

Thus:

$$(*)\quad \boxed{\mathcal{W}_{\mathcal{H}}(Q,P) \le \sqrt{2H(Q|P)}}$$

Talagrand's inequality on Wiener space

(cf. Feyel, Üstünel 2002)

direct proof on Wiener space:
(Lehec 2013)

$H(Q|P) < \infty \Rightarrow X := \omega^Q,\ Y := \omega$
is an "adaptive" coupling of P, Q on $(\Omega, \mathfrak{F}, Q)$

with

$$E_Q[\|X-Y\|_{\mathcal{H}}^2] = E_Q\left[\int_0^1 (b_t^Q)^2 dt\right]$$
$$= 2H(Q|P)$$

$\Rightarrow (*)$

(Corollary: Talagrand's inequality on $\mathbb{R}^1$)

In fact (Lassalle 2015):

$$\boxed{2H(Q|P) = W^2_{\mathcal{H},\,\text{adapted}}(Q,P)}$$

<u>Anticipating coupling</u> may do better!

Example from probabilistic potential theory:

$$Q := \int \nu(dx)\, \underbrace{P_0^x}_{\substack{\text{Brownian bridge} \\ \text{from } 0 \text{ to } x}}$$

$\Rightarrow$

$$H(Q|P) = H(\nu|\mu), \quad \mu := N(0,1)$$

optimal <u>adaptive</u> coupling: $X := \omega$,

$Y :=$ h-path process with drift $\nabla \log \underbrace{h(\cdot,t)}_{\text{space-time harmonic}}$, $\quad h(\cdot,1) = \frac{d\nu}{d\mu}(\cdot)$

optimal anticipating coupling:

$$X := \omega, \quad Y_t := \omega_t - t\omega_1 + t\,T(\omega_1) \quad (0 \le t \le 1)$$

with $T := q_\nu \circ \phi$

(transforms μ into ν)

$\Longrightarrow$

$$E_P\left[\|Y-X\|^2_{\mathcal{H}}\right] = E_P\left[(T(\omega_1)-\omega_1)^2\right]$$

$= t\,(T(\omega_1)-\omega_1)$

$$= \int (T(x)-x)^2\,\mu(dx)$$

$$= \int_0^1 (q_\nu(\alpha)-q_\mu(\alpha))^2\,d\alpha$$

$$= W^2(\nu,\mu) < 2H(\nu|\mu)$$

if $\nu \neq N(m,1)$ for some $m \in \mathbb{R}^1$

Corollary: Talagrand's inequality $\Longrightarrow$

$$\rho_\beta(F) \le \rho_{W,\beta}(F) := \sup_Q \left(E_Q[F] - \frac{1}{2\beta} W^2(Q,P) \right)$$

So far: restricted to measures $Q \ll P$ (fixed volatility pattern)

2. Beyond absolute continuity: specific relative entropy

$D_N := \{k2^{-N} \mid k = 0, \dots, 2^N - 1\}$

$\mathcal{F}_N := \sigma(\{\omega_t \mid t \in D_N\})$

Wasserstein distance along D_N:

$$W_N^2(Q,P) := \inf_{\substack{\text{couplings} \\ (Y_t)_{t \in D_N}, (X_t)_{t \in D_N} \\ \text{of } Q|_{\mathcal{F}_N},\ P|_{\mathcal{F}_N}}} E_{\hat{P}}\Big[\underbrace{\sum_{t_i \in D_N} \Delta_i (Y-X)^2}\Big]$$

= quadratic variation of $Y - X$ along D_N

$= \frac{1}{2^N} \sum_{i \in I_N} (Y_i - X_i)^2$

$$\underset{\text{Talagrand for } I_N}{\leq} \frac{1}{2^N}\, \underbrace{2 H(Q|P)\big|_{\mathcal{F}_N}}_{=: H_N(Q|P)}$$

Thus:

① $$\omega^2_{\mathcal{H}}(Q,P) \overset{\text{martingale convergence}}{=} \lim_{N\uparrow\infty} 2^N \omega^2_N(Q,P)$$

$$\leq 2H(Q|P)$$

- another proof of Talagrand's inequality on Wiener space

② Beyond absolute continuity:

$$\omega^2_{\langle\cdot\rangle}(Q,P) := \inf_{\substack{\text{couplings}\\ X,Y \text{ of } P,Q}} E_{\hat{P}}\left[\langle Y-X\rangle_1\right]$$

$$\underset{\text{martingale convergence}}{=} \lim_{N\uparrow\infty} \omega^2_N(Q,P)$$

$$\leq 2 \underbrace{\lim_{N\uparrow\infty} 2^{-N} H_N(Q|P)}$$

"specific relative entropy" $=: h(Q|P)$

cf. Nina Gantert (thesis Bonn, 1991):

$h(Q|P)$ is the key to large deviations in the ergodic behavior

$$\sum_{t_i \in D_N} \delta_{t_i} (\omega_{t_{i+1}} - \omega_{t_i})^2 \xrightarrow[N \uparrow \infty]{} \text{Lebesgue measure}$$

of quadratic variation weakly under P

Thus:

$$\boxed{W_{\langle \cdot \rangle}(Q,P) \leq \sqrt{2h(Q|P)}}$$

— a new (rescaled) version of Talagrand's inequality on Wiener space, beyond the absolutely continuous case

Illustration:

① $Q :=$ law of $X := \omega + B$

$B \in C_0[0,1]$ deterministic

a) Cameron-Martin:

$H(Q|P) < \infty \iff B \in \mathcal{H}$

$$H(Q|P) = \frac{1}{2} \|B\|_{\mathcal{H}}^2 = \frac{1}{2} \mathcal{W}^2(Q,P)$$

b) Nina Gantert:

$\exists\, h(Q|P) < \infty \iff$

B has quadratic variation $\langle B \rangle$ along dyadic partitions,

$$h(Q|P) = \frac{1}{2} \langle B \rangle_1 = \frac{1}{2} \mathcal{W}^2_{\langle \cdot \rangle}(Q,P)$$

(2) Q = martingale measure on $C_0[0,1]$

such that

$$d\langle\omega\rangle_t = \sigma_t^2(\omega)\,dt, \quad \sigma_t \neq 0$$

$\Rightarrow$ $\hat{\omega}_t^Q := \int_0^t \frac{1}{\sigma_s(\omega_s)}\,d\omega_s$ Wiener process under Q

$\Rightarrow$ $(\omega, \hat{\omega}^Q)$ on $(\Omega, \mathcal{F}, Q)$ is an (adapted) coupling of Q, P

such that

$$E_Q[\langle\omega - \hat{\omega}^Q\rangle_1] = E_Q\Big[\int_0^1 \underbrace{(\sigma_t(\omega) - 1)^2}_{\leq \frac{1}{2}\left(\log\frac{1}{\sigma_t^2} + (\sigma_t^2 - 1)\right)}\,dt\Big]$$

$$\leq \frac{1}{2} E_Q\Big[\int_0^1 \log\frac{1}{\sigma_t^2} + (\sigma_t^2 - 1)\,dt\Big]$$

$$\leq \quad 2\,h(Q|P)$$

N. Gantert

($=$ if σ_t deterministic)

Thus:

$$\boxed{\omega_{\langle\cdot\rangle}(Q,P) \leq \sqrt{2\,h(Q|P)}}$$

3. Rescaling the entropic risk measure:

$h(Q|P)$ instead of $H(Q|P)$

"fine structure":

$$\hat{\mathcal{F}} := \bigcap_{N=1}^{\infty} \sigma(\underbrace{\{X_i \mid i \in I - I_N\}}_{\text{cf. Lévy-Ciesielski}})$$

$\Rightarrow$

$$P = 0\text{-}1 \text{ on } \hat{\mathcal{F}}$$

i.e., not sensitive to "risk" in the fine structure

$\Rightarrow$

$$\rho_\beta(\hat{F}) = \frac{1}{\beta} \log E_P[e^{-\beta \hat{F}}]$$

$$= E_P[-\hat{F}]$$

for $\hat{\mathcal{F}}$-measurable $\hat{F}$

Example: quadratic variation

$$\hat{F} := \langle \omega \rangle_1 = \lim_{N \uparrow \infty} \underbrace{\frac{1}{2^N} \sum_{i \in I_N} X_i^2(\omega)}$$

$= 1$ P-a.s.

$= \text{quadratic variation along } D_N$

$=: \hat{F}_N$

However:

$\beta_N := \beta 2^N \Rightarrow$

$$\hat{\rho}_\beta(\hat{F}) := \lim_{N \uparrow \infty} \frac{1}{\beta_N} \log E_P[e^{-\beta_N \hat{F}_N}]$$

$$= -\frac{1}{\beta} \log \sqrt{1+2\beta}$$

$$> -1 = \rho_\beta(\hat{F})$$

Moreover:

$$\boxed{\hat{\rho}_\beta(\hat{F}) = \sup_{Q \in \mathcal{Q}} \left(E_Q[-\hat{F}] - \frac{1}{\beta} h(Q|P) \right)}$$

$$\leq \sup_{Q \in \mathcal{Q}} \left(E_Q[-\hat{F}] - \frac{1}{2\beta} \omega^2_{\langle \cdot \rangle}(Q,P) \right)$$

Talagrand

a typical case of

" $\longleftarrow$ "

(original motivation:
dynamics in the space of
martingale measures as a source
of "bubbles", cf.
Jarrow, Protter, Shimbo 2010
Biagini, F., Nedelcu 2014)

no claim to relevance!

in contrast to current work of

Nicole: " $\longrightarrow$ "

in particular:

the (collective/actuarial/...)
risks of longevity

May longevity come to you,

Nicole,

in its best possible form, including many enjoyable encounters with all of us in the future

(in terms of "scientific longevity", you are already a model for all of us!)

— thanks for your tremendous contribution, and my

best wishes!