A stochastic formula for the entropy and applications

Joseph Lehec

Université Paris-Dauphine

Asymptotic Geometric Analysis
Fields Institute
Toronto, nov. 4th, 2010

Setting

B: a standard Brownian motion on $\mathbb{R}^n$ starting from 0

P: be the corresponding heat semi-group

- $P_t f(x) = \mathrm{E}\, f\big(x + B(t)\big)$
- $\partial_t P_t f = \Delta P_t f / 2$

Setting

B: a standard Brownian motion on $\mathbb{R}^n$ starting from 0
P: be the corresponding heat semi-group

- $P_t f(x) = \mathrm{E}\, f\big(x + B(t)\big)$
- $\partial_t P_t f = \Delta P_t f/2$

Throughtout a drift is any process $\big(u(t)\big)_{t\geq 0}$ adapted to the underlying filtration.
This filtration may be $\mathcal{F}_t = \sigma\big(B(s), s \in [0,t]\big)$ or larger.

Borell's formula

Let γ_n be the Gaussian measure on $\mathbb{R}^n$ (law of $B(1)$)

Borell's formula

Let γ_n be the Gaussian measure on $\mathbb{R}^n$ (law of $B(1)$)

Laplace transform

Let $f\colon \mathbb{R}^n \to \mathbb{R}$, we define $\mathrm{L}(f) := \log\left(\int_{\mathbb{R}^n} \mathrm{e}^f \, \mathrm{d}\gamma_n\right)$.

Borell's formula

Let γ_n be the Gaussian measure on $\mathbb{R}^n$ (law of $B(1)$)

Laplace transform

Let $f\colon \mathbb{R}^n \to \mathbb{R}$, we define $\mathrm{L}(f) := \log\Big(\int_{\mathbb{R}^n} \mathrm{e}^f \, \mathrm{d}\gamma_n\Big)$.

Borell's formula

For all function f on $\mathbb{R}^n$ (mild conditions on f)

$$\mathrm{L}(f) = \sup_u \left(\mathrm{E}\, f\Big(B(1) + \int_0^1 u(s)\, \mathrm{d}s\Big) - \frac{1}{2} \int_0^1 |u(s)|^2 \, \mathrm{d}s \right),$$

the supremum is over all drifts u.

Comments on Borell's formula

The formula is not due to Borell, though he should be credited for the idea of using it to prove functional inequalities such as

- Prékopa-Leindler inequality.
- Brascamp-Lieb inequality.

Entropy

Relative entropy

Let μ be a probability measure on $\mathbb{R}^n$.
Assume that μ has a density, and let $f = \mathrm{d}\mu/\mathrm{d}\gamma_n$.

$$\mathrm{H}(\mu) = \int f \log(f)\ \mathrm{d}\gamma_n = \int \log(f)\ \mathrm{d}\mu.$$

Entropy

Relative entropy

Let μ be a probability measure on $\mathbb{R}^n$.
Assume that μ has a density, and let $f = \mathrm{d}\mu/\mathrm{d}\gamma_n$.

$$\mathrm{H}(\mu) = \int f \log(f)\ \mathrm{d}\gamma_n = \int \log(f)\ \mathrm{d}\mu.$$

Remarks

- $\mathrm{H}(\mu) \geq 0$
- $\mathrm{H}(\mu) = 0 \Leftrightarrow \mu = \gamma_n$.

Entropy

Relative entropy

Let μ be a probability measure on $\mathbb{R}^n$.
Assume that μ has a density, and let $f = \mathrm{d}\mu / \mathrm{d}\gamma_n$.

$$\mathrm{H}(\mu) = \int f \log(f) \ \mathrm{d}\gamma_n = \int \log(f) \ \mathrm{d}\mu.$$

Remarks

- $\mathrm{H}(\mu) \geq 0$
- $\mathrm{H}(\mu) = 0 \Leftrightarrow \mu = \gamma_n$.

Legendre duality

For all probability measure μ

$$\mathrm{H}(\mu) = \sup_f \left(\int f \ \mathrm{d}\mu - \mathrm{L}(f) \right).$$

The formula

Theorem

Let μ be a probability measure on $\mathbb{R}^n$ with smooth positive density.

$$\mathrm{H}(\mu) = \frac{1}{2} \inf \Big(\mathrm{E} \int_0^1 |u(s)|^2 \, \mathrm{d}s \Big).$$

Infimum on all drifts u such that $B(1) + \int_0^1 u(s) \, \mathrm{d}s$ has law μ.

The formula

Theorem

Let μ be a probability measure on $\mathbb{R}^n$ with smooth positive density.

$$\mathrm{H}(\mu) = \frac{1}{2} \inf \Big(\mathrm{E} \int_0^1 |u(s)|^2 \, \mathrm{d}s \Big).$$

Infimum on all drifts u such that $B(1) + \int_0^1 u(s) \, \mathrm{d}s$ has law μ.

Besides

Let $f = \mathrm{d}\mu / \mathrm{d}\gamma_n$. The infimum is attained for some drift v which

- solves the SDE: $v(t) = \nabla \ln P_{1-t} f\big(B(t) + \int_0^t v(s) \, \mathrm{d}s\big)$.
- is a martingale, in particular $\mathrm{E}\, v(t) = \mathrm{bar}(\mu)$ for all t.

$\mathrm{bar}(\mu) := \int x \, \mathrm{d}\mu(x)$.

Proof: Upper bound (1)

- Let u be a drift such that $B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ. Then

$$\mathrm{E}\log(f)\Big(B(1) + \int_0^1 u(s)\ \mathrm{d}s\Big) = \mathrm{H}(\mu).$$

Proof: Upper bound (1)

- Let u be a drift such that $B(1) + \int_0^1 u(s)\,\mathrm{d}s$ has law μ. Then

$$\mathrm{E}\log(f)\Big(B(1) + \int_0^1 u(s)\,\mathrm{d}s\Big) = \mathrm{H}(\mu).$$

- Let $F(t,x) = \log(P_{1-t}f)(x)$ and

$$M^u(t) = F\Big(t, B(t) + \int_0^t u(s)\,\mathrm{d}s\Big) - \int_0^t |u(s)|^2/2\,\mathrm{d}s$$

Proof: Upper bound (1)

- Let u be a drift such that $B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ. Then

$$\mathrm{E}\log(f)\Big(B(1) + \int_0^1 u(s)\ \mathrm{d}s\Big) = \mathrm{H}(\mu).$$

- Let $F(t,x) = \log(P_{1-t}f)(x)$ and

$$M^u(t) = F\Big(t, B(t) + \int_0^t u(s)\ \mathrm{d}s\Big) - \int_0^t |u(s)|^2/2\ \mathrm{d}s$$

- Then $M^u(0) = \log(P_1 f)(0) = \log(\int f\ \mathrm{d}\gamma_n) = 0$.

Proof: Upper bound (1)

- Let u be a drift such that $B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ. Then

$$\mathrm{E} \log(f)\Big(B(1) + \int_0^1 u(s)\ \mathrm{d}s\Big) = \mathrm{H}(\mu).$$

- Let $F(t,x) = \log(P_{1-t}f)(x)$ and

$$M^u(t) = F\Big(t, B(t) + \int_0^t u(s)\ \mathrm{d}s\Big) - \int_0^t |u(s)|^2/2\ \mathrm{d}s$$

- Then $M^u(0) = \log(P_1 f)(0) = \log(\int f\ \mathrm{d}\gamma_n) = 0$.
- And $\mathrm{E}\, M^u(1) = \mathrm{H}(X) - \mathrm{E} \int_0^1 |u(s)|^2/2\ \mathrm{d}s$.

Proof: Upper bound (1)

- Let u be a drift such that $B(1) + \int_0^1 u(s)\,\mathrm{d}s$ has law μ. Then

$$\mathrm{E}\log(f)\Big(B(1) + \int_0^1 u(s)\,\mathrm{d}s\Big) = \mathrm{H}(\mu).$$

- Let $F(t,x) = \log(P_{1-t}f)(x)$ and

$$M^u(t) = F\Big(t, B(t) + \int_0^t u(s)\,\mathrm{d}s\Big) - \int_0^t |u(s)|^2/2\,\mathrm{d}s$$

- Then $M^u(0) = \log(P_1 f)(0) = \log(\int f\,\mathrm{d}\gamma_n) = 0$.
- And $\mathrm{E}\,M^u(1) = \mathrm{H}(X) - \mathrm{E}\int_0^1 |u(s)|^2/2\,\mathrm{d}s$.
- If we prove that M^u is a super-martingale, then in particular $\mathrm{E}\,M^u(0) \geq \mathrm{E}\,M^u(1)$ and we are done.

Proof: Upper bound (2)

- $F(t,x) = \log(P_{1-t}f)(x)$ yields $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.

Proof: Upper bound (2)

- $F(t,x) = \log(P_{1-t}f)(x)$ yields $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.
- Recall that

$$M^u(t) = F\big(t, B(t) + \int_0^t u(s)\ \mathrm{d}s\big) - \int_0^t |u(s)|^2/2\ \mathrm{d}s.$$

Proof: Upper bound (2)

- $F(t,x) = \log(P_{1-t}f)(x)$ yields $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.
- Recall that

$$M^u(t) = F\big(t, B(t) + \int_0^t u(s)\ \mathrm{d}s\big) - \int_0^t |u(s)|^2/2\ \mathrm{d}s.$$

- By Itô's formula (omitting variables)

$$\begin{aligned} \mathrm{d}M^u &= \partial_t F\ \mathrm{d}t + \nabla F \cdot (\mathrm{d}B + u\ \mathrm{d}t) + \Delta F/2\ \mathrm{d}t - |u|^2/2\ \mathrm{d}t \\ &= \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2\ \mathrm{d}t. \end{aligned}$$

Proof: Upper bound (2)

- $F(t,x) = \log(P_{1-t} f)(x)$ yields $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.
- Recall that

$$M^u(t) = F\Big(t, B(t) + \int_0^t u(s)\ \mathrm{d}s\Big) - \int_0^t |u(s)|^2/2\ \mathrm{d}s.$$

- By Itô's formula (omitting variables)

$$\begin{aligned}\mathrm{d}M^u &= \partial_t F\ \mathrm{d}t + \nabla F \cdot (\mathrm{d}B + u\ \mathrm{d}t) + \Delta F/2\ \mathrm{d}t - |u|^2/2\ \mathrm{d}t \\ &= \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2\ \mathrm{d}t.\end{aligned}$$

- So M^u is a super-martingale.

Proof: Equality case

- From the previous slide (omitting variables)

$$\mathrm{d}M^u = \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2 \ \mathrm{d}t$$

Proof: Equality case

- From the previous slide (omitting variables)

$$\mathrm{d}M^u = \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2 \ \mathrm{d}t$$

- Recalling variables, if v solves the SDE

$$\begin{aligned} v(t) &= \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big) \\ &= \nabla \log(P_{1-t} f)\Big(B(t) + \int_0^t v(s)\ \mathrm{d}s\Big) \end{aligned}$$

Proof: Equality case

- From the previous slide (omitting variables)

$$\mathrm{d}M^u = \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2 \ \mathrm{d}t$$

- Recalling variables, if v solves the SDE

$$\begin{aligned} v(t) &= \nabla F\Big(t, B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \\ &= \nabla \log(P_{1-t} f)\Big(B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \end{aligned}$$

then M^v is a martingale and

$$\mathrm{E} \log(f)\Big(B(1) + \int_0^1 v(s) \ \mathrm{d}s\Big) = \mathrm{E} \int_0^1 |v(s)|^2/2 \ \mathrm{d}s.$$

Proof: Equality case

- From the previous slide (omitting variables)

$$\mathrm{d}M^u = \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2 \ \mathrm{d}t$$

- Recalling variables, if v solves the SDE

$$\begin{aligned} v(t) &= \nabla F\Big(t, B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \\ &= \nabla \log(P_{1-t} f)\Big(B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \end{aligned}$$

  then M^v is a martingale and

$$\mathrm{E} \log(f)\Big(B(1) + \int_0^1 v(s) \ \mathrm{d}s\Big) = \mathrm{E} \int_0^1 |v(s)|^2/2 \ \mathrm{d}s.$$

- It only remains to prove that $B(1) + \int_0^1 v(s) \ \mathrm{d}s$ has law μ.

Proof: Equality case

- From the previous slide (omitting variables)

$$\mathrm{d}M^u = \nabla F \cdot \mathrm{d}B - |\nabla F - u|^2/2 \ \mathrm{d}t$$

- Recalling variables, if v solves the SDE

$$\begin{aligned} v(t) &= \nabla F\Big(t, B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \\ &= \nabla \log(P_{1-t}f)\Big(B(t) + \int_0^t v(s) \ \mathrm{d}s\Big) \end{aligned}$$

  then M^v is a martingale and

$$\mathrm{E} \log(f)\Big(B(1) + \int_0^1 v(s) \ \mathrm{d}s\Big) = \mathrm{E} \int_0^1 |v(s)|^2/2 \ \mathrm{d}s.$$

- It only remains to prove that $B(1) + \int_0^1 v(s) \ \mathrm{d}s$ has law μ.
- This follows from Girsanov's formula.

Proof: Optimal drift is a martingale

- Optimal drift:

$$v(t) = \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big).$$

Proof: Optimal drift is a martingale

- Optimal drift:

$$v(t) = \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big).$$

- By Itô's formula again

$$\mathrm{d}v = \partial_t \nabla F\ \mathrm{d}t + \nabla^2 F(\mathrm{d}B + v\ \mathrm{d}t) + \Delta(\nabla F)/2\ \mathrm{d}t.$$

Proof: Optimal drift is a martingale

- Optimal drift:

$$v(t) = \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big).$$

- By Itô's formula again

$$\mathrm{d}v = \partial_t \nabla F\ \mathrm{d}t + \nabla^2 F(\mathrm{d}B + v\ \mathrm{d}t) + \Delta(\nabla F)/2\ \mathrm{d}t.$$

- Recall that $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.

Proof: Optimal drift is a martingale

- Optimal drift:

$$v(t) = \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big).$$

- By Itô's formula again

$$\mathrm{d}v = \partial_t \nabla F\ \mathrm{d}t + \nabla^2 F(\mathrm{d}B + v\ \mathrm{d}t) + \Delta(\nabla F)/2\ \mathrm{d}t.$$

- Recall that $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.
- So

$$\begin{aligned}\partial_t \nabla F &= -\frac{1}{2}\big(\nabla(\Delta F) + \nabla(|\nabla F|^2)\big)\\ &= -\frac{1}{2}\Delta(\nabla F) - \nabla^2 F(\nabla F).\end{aligned}$$

Proof: Optimal drift is a martingale

- Optimal drift:

$$v(t) = \nabla F\Big(t, B(t) + \int_0^t v(s)\ \mathrm{d}s\Big).$$

- By Itô's formula again

$$\mathrm{d}v = \partial_t \nabla F\ \mathrm{d}t + \nabla^2 F(\mathrm{d}B + v\ \mathrm{d}t) + \Delta(\nabla F)/2\ \mathrm{d}t.$$

- Recall that $\partial_t F = -(\Delta F + |\nabla F|^2)/2$.
- So

$$\begin{aligned}\partial_t \nabla F &= -\frac{1}{2}\big(\nabla(\Delta F) + \nabla(|\nabla F|^2)\big)\\ &= -\frac{1}{2}\Delta(\nabla F) - \nabla^2 F(\nabla F).\end{aligned}$$

- Thus $\mathrm{d}v = \nabla^2 F(\mathrm{d}B)$ and v is a martingale.

Comments

- This proof is very similar to that of Borell

Comments

- This proof is very similar to that of Borell
- It is also reminiscent of works by Föllmer in the 80s

1 Introduction: Borell's formula

2 Stochastic formula for the entropy

3 Applications

Transportation cost inequality

Definition

Let μ and ν be probability measures on $\mathbb{R}^n$.

$$\mathrm{W}_2(\mu,\nu) = \inf\Big(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \, \mathrm{d}\pi(x,y)\Big)^{1/2}.$$

Infimum on all probability measure π on $\mathbb{R}^n \times \mathbb{R}^n$
having marginals μ and ν.

Transportation cost inequality

Definition

Let μ and ν be probability measures on $\mathbb{R}^n$.

$$\mathrm{W}_2(\mu,\nu) = \inf\Big(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \,\mathrm{d}\pi(x,y)\Big)^{1/2}.$$

Infimum on all probability measure π on $\mathbb{R}^n \times \mathbb{R}^n$
having marginals μ and ν.

Transportation Inequality (Talagrand)

$$\mathrm{W}_2(\mu,\gamma_n)^2 \le 2\,\mathrm{H}(\mu).$$

Proof

- Let B be a Brownian motion and u be a drift such that $X := B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ.

Proof

- Let B be a Brownian motion and u be a drift such that $X := B(1) + \int_0^1 u(s)\,\mathrm{d}s$ has law μ.
- Then $\big(X, B(1)\big)$ is a coupling of (μ, γ_n) so

$$\mathrm{W}_2(\mu, \gamma_n)^2 \le \mathrm{E}|X - B(1)|^2.$$

Proof

- Let B be a Brownian motion and u be a drift such that $X := B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ.
- Then $\big(X, B(1)\big)$ is a coupling of (μ, γ_n) so

$$\mathrm{W}_2(\mu, \gamma_n)^2 \leq \mathrm{E}|X - B(1)|^2.$$

- By Jensen

$$\mathrm{E}|X - B(1)|^2 = \mathrm{E}\Big|\int_0^1 u(s)\ \mathrm{d}s\Big|^2 \leq \mathrm{E}\int_0^1 |u(s)|^2\ \mathrm{d}s$$

Proof

- Let B be a Brownian motion and u be a drift such that $X := B(1) + \int_0^1 u(s)\ \mathrm{d}s$ has law μ.
- Then $\big(X, B(1)\big)$ is a coupling of (μ, γ_n) so

$$\mathrm{W}_2(\mu, \gamma_n)^2 \leq \mathrm{E}|X - B(1)|^2.$$

- By Jensen

$$\mathrm{E}|X - B(1)|^2 = \mathrm{E}\Big|\int_0^1 u(s)\ \mathrm{d}s\Big|^2 \leq \mathrm{E}\int_0^1 |u(s)|^2\ \mathrm{d}s$$

- Taking infimum on u we get $\mathrm{W}_2(\mu, \gamma_n)^2 \leq 2\,\mathrm{H}(\mu)$.

Log-Sobolev inequality

Fisher Information

Let μ be an absolutely continuous probability measure on $\mathbb{R}^n$.
Let $f = \mathrm{d}\mu / \mathrm{d}\gamma_n$.

$$\mathrm{I}(\mu) = \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, \mathrm{d}\gamma_n = \int_{\mathbb{R}^n} |\nabla \log(f)|^2 \, \mathrm{d}\mu$$

Log-Sobolev inequality

Fisher Information

Let μ be an absolutely continuous probability measure on $\mathbb{R}^n$.
Let $f = \mathrm{d}\mu/\mathrm{d}\gamma_n$.

$$\mathrm{I}(\mu) = \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, \mathrm{d}\gamma_n = \int_{\mathbb{R}^n} |\nabla \log(f)|^2 \, \mathrm{d}\mu$$

Logarithmic Sobolev inequality (Gross)

$$\mathrm{H}(\mu) \leq \frac{1}{2} \mathrm{I}(\mu).$$

Proof

- Let $f = \mathrm{d}\mu/\mathrm{d}\gamma$. Optimal drift for μ:

$$v(t) = \nabla \log(P_{1-t} f)(B(t) + \int_0^t v(s)\,\mathrm{d}s).$$

Proof

- Let $f = \mathrm{d}\mu/\mathrm{d}\gamma$. Optimal drift for μ:

$$v(t) = \nabla \log(P_{1-t} f)(B(t) + \int_0^t v(s)\ \mathrm{d}s).$$

- Since $B(1) + \int_0^1 v(s)\ \mathrm{d}s$ has law μ

$$\mathrm{E}|v(1)|^2 = \mathrm{E}\Big|\nabla \log(f)\big(B(1) + \int_0^1 v(s)\ \mathrm{d}s\big)\Big|^2 = \mathrm{I}(\mu).$$

Proof

- Let $f = \mathrm{d}\mu/\mathrm{d}\gamma$. Optimal drift for μ:

$$v(t) = \nabla \log(P_{1-t} f)(B(t) + \int_0^t v(s)\ \mathrm{d}s).$$

- Since $B(1) + \int_0^1 v(s)\ \mathrm{d}s$ has law μ

$$\mathrm{E}|v(1)|^2 = \mathrm{E}\Big|\nabla \log(f)\big(B(1) + \int_0^1 v(s)\ \mathrm{d}s\big)\Big|^2 = \mathrm{I}(\mu).$$

- v martingale $\Rightarrow$ $|v|^2$ sub-martingale.

Proof

- Let $f = \mathrm{d}\mu/\mathrm{d}\gamma$. Optimal drift for μ:

$$v(t) = \nabla \log(P_{1-t} f)(B(t) + \int_0^t v(s)\ \mathrm{d}s).$$

- Since $B(1) + \int_0^1 v(s)\ \mathrm{d}s$ has law μ

$$\mathrm{E}|v(1)|^2 = \mathrm{E}\Big|\nabla \log(f)\big(B(1) + \int_0^1 v(s)\ \mathrm{d}s\big)\Big|^2 = \mathrm{I}(\mu).$$

- v martingale $\Rightarrow$ $|v|^2$ sub-martingale.
- Hence

$$\mathrm{H}(\mu) = \frac{1}{2}\int_0^1 \mathrm{E}|v(s)|^2\ \mathrm{d}s \leq \frac{1}{2}\,\mathrm{E}|v(1)|^2 = \frac{1}{2}\,\mathrm{I}(\mu).$$

Shannon's inequality

Definition

X a random vector on $\mathbb{R}^n$ having density f with respect to the Lebesgue measure

$$\mathrm{S}(X) = \int f \log(f) \, \mathrm{d}x = \mathrm{E} \log(f)(X).$$

Shannon's inequality

Definition

X a random vector on $\mathbb{R}^n$ having density f with respect to the Lebesgue measure

$$\mathrm{S}(X) = \int f \log(f) \,\mathrm{d}x = \mathrm{E} \log(f)(X).$$

Remark

$$\mathrm{S}(X) = \mathrm{H}(X) - \frac{n}{2} \log(2\pi) - \frac{1}{2} \mathrm{E}|X|^2.$$

Shannon's inequality

Definition

X a random vector on $\mathbb{R}^n$ having density f with respect to the Lebesgue measure

$$\mathrm{S}(X) = \int f \log(f)\,\mathrm{d}x = \mathrm{E}\log(f)(X).$$

Remark

$$\mathrm{S}(X) = \mathrm{H}(X) - \frac{n}{2}\log(2\pi) - \frac{1}{2}\,\mathrm{E}|X|^2.$$

Shannon's Inequality

X, Y independant random vectors, $\lambda \in (0,1)$

$$\mathrm{S}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq (1-\lambda)\,\mathrm{S}(X) + \lambda\,\mathrm{S}(Y).$$

Proof (1)

Let B be a Brownian motion and u be a drift satisfying

- $B(1) + \int_0^1 u(s)\ \mathrm{d}s = X$ in law.
- $\mathrm{H}(X) = \mathrm{E} \int_0^1 |u(s)|^2/2\ \mathrm{d}s$.
- $\mathrm{E}\, u(s) = \mathrm{E}\, X$ for all s.

Proof (1)

Let B be a Brownian motion and u be a drift satisfying

- $B(1) + \int_0^1 u(s)\ \mathrm{d}s = X$ in law.
- $\mathrm{H}(X) = \mathrm{E} \int_0^1 |u(s)|^2/2\ \mathrm{d}s$.
- $\mathrm{E}\, u(s) = \mathrm{E}\, X$ for all s.

Let C be a Brownian motion independant of B and v be a drift satisfying

- $C(1) + \int_0^1 v(s)\ \mathrm{d}s = Y$ in law.
- $\mathrm{H}(Y) = \mathrm{E} \int_0^1 |v(s)|^2/2\ \mathrm{d}s$.
- $\mathrm{E}\, v(s) = \mathrm{E}\, Y$ for all s.

Proof (2)

- Let $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$ and $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$.

Proof (2)

- Let $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$ and $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$.
- Then W is a standard Brownian motion and

$$W + \int_0^1 w(s)\, \mathrm{d}s \stackrel{\text{(law)}}{=} \sqrt{1-\lambda}X + \sqrt{\lambda}Y.$$

Proof (2)

- Let $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$ and $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$.
- Then W is a standard Brownian motion and

$$W + \int_0^1 w(s)\ \mathrm{d}s \stackrel{\text{(law)}}{=} \sqrt{1-\lambda}X + \sqrt{\lambda}Y.$$

- Hence $\mathrm{H}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq \mathrm{E} \int_0^1 |w(s)|^2/2\ \mathrm{d}s$.

Proof (2)

- Let $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$ and $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$.
- Then W is a standard Brownian motion and

$$W + \int_0^1 w(s)\,\mathrm{d}s \overset{\text{(law)}}{=} \sqrt{1-\lambda}X + \sqrt{\lambda}Y.$$

- Hence $\mathrm{H}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq \mathrm{E}\int_0^1 |w(s)|^2/2\,\mathrm{d}s$.
- $\mathrm{E}|w(s)|^2 = (1-\lambda)|u(s)|^2 + \lambda|v(s)|^2 + 2\sqrt{\lambda(1-\lambda)}\,\mathrm{E}(X)\cdot\mathrm{E}(Y)$.

Proof (2)

- Let $W = \sqrt{1-\lambda}B + \sqrt{\lambda}C$ and $w = \sqrt{1-\lambda}u + \sqrt{\lambda}v$.
- Then W is a standard Brownian motion and

$$W + \int_0^1 w(s)\ \mathrm{d}s \overset{(\text{law})}{=} \sqrt{1-\lambda}X + \sqrt{\lambda}Y.$$

- Hence $\mathrm{H}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) \leq \mathrm{E}\int_0^1 |w(s)|^2/2\ \mathrm{d}s$.
- $\mathrm{E}|w(s)|^2 = (1-\lambda)|u(s)|^2 + \lambda|v(s)|^2 + 2\sqrt{\lambda(1-\lambda)}\,\mathrm{E}(X)\cdot\mathrm{E}(Y)$.
- Therefore

$$\begin{aligned}\mathrm{H}(\sqrt{1-\lambda}X + \sqrt{\lambda}Y) &\leq (1-\lambda)\,\mathrm{H}(X) + \lambda\,\mathrm{H}(Y)\\ &\quad + \sqrt{\lambda(1-\lambda)}\,\mathrm{E}(X)\cdot\mathrm{E}(Y)\end{aligned}$$

which is the result.

Brascamp-Lieb inequality

Brascamp-Lieb inequality

Frame condition

Let E be a Euclidean space, $E_1, \ldots, E_m$ subspaces, $c_1, \ldots, c_m$ positive numbers, satisfying

$$\sum_{i=1}^{m} c_i P_i = \mathrm{id}_E$$

where P_i is the orthogonal projection with range E_i.

Brascamp-Lieb inequality

Frame condition

Let E be a Euclidean space, $E_1, \dots, E_m$ subspaces, $c_1, \dots, c_m$ positive numbers, satisfying

$$\sum_{i=1}^{m} c_i P_i = \mathrm{id}_E$$

where P_i is the orthogonal projection with range E_i.

Brascamp-Lieb Inequality

Under the frame condition, for all $f_i \colon E_i \to \mathbb{R}$,

$$\int_E \mathrm{e}^{\sum c_i f_i(P_i x)} \, \mathrm{d}\gamma_E(x) \leq \prod_{i=1}^{m} \Big(\int_{E_i} \mathrm{e}^{f_i} \, \mathrm{d}\gamma_{E_i} \Big)^{c_i}$$

γ_E : Gaussian measure on E.

Comments on BL inequality

$$\int_E e^{\sum c_i f_i(P_i x)} \, d\gamma_E(x) \le \prod_{i=1}^{m} \Big(\int_{E_i} e^{f_i} \, d\gamma_{E_i} \Big)^{c_i}$$

Remarks

- If $c_i = 1$ and the spaces E_i form an orthogonal decomposition of E there is equality (this is just Fubini).
- When $E_i = E$ for all i and $\sum c_i = 1$ this is just Hölder.
- Remains true with Lebesgue measures instead of Gaussian measures.

Comments on BL inequality

$$\int_E e^{\sum c_i f_i(P_i x)}\, d\gamma_E(x) \le \prod_{i=1}^m \Big(\int_{E_i} e^{f_i}\, d\gamma_{E_i}\Big)^{c_i}$$

Remarks

- If $c_i = 1$ and the spaces E_i form an orthogonal decomposition of E there is equality (this is just Fubini).
- When $E_i = E$ for all i and $\sum c_i = 1$ this is just Hölder.
- Remains true with Lebesgue measures instead of Gaussian measures.

Applications

- Connection with Young's convolution inequality.
- Hypercontractivity of the Ornstein-Uhlenbeck semi-group.
- Other geometric applications

Dual version

Dual formulation of the BL inequality (Carlen-Cordero)

Under the frame condition, for all random vector X on E

$$\mathrm{H}(X) \geq \sum_{i=1}^{m} c_i \, \mathrm{H}(P_i X)$$

Dual version

Dual formulation of the BL inequality (Carlen-Cordero)

Under the frame condition, for all random vector X on E

$$\mathrm{H}(X) \geq \sum_{i=1}^{m} c_i \, \mathrm{H}(P_i X)$$

The equivalence follows from the Legendre duality

$$\mathrm{H}(\mu) = \sup_f \Big(\int f \, \mathrm{d}\mu - \mathrm{L}(f) \Big)$$

Proof of Dual BL

- B a Brownian motion.
- u a drift satisfying $B(1) + \int_0^1 u(s)\, \mathrm{d}s = X$ in law.

Proof of Dual BL

- B a Brownian motion.
- u a drift satisfying $B(1) + \int_0^1 u(s)\ \mathrm{d}s = X$ in law.
- then $P_i B$ is a Brownian motion on E_i
- and $P_i B + \int_0^1 P_i u(s)\ \mathrm{d}s = P_i X$ in law

Proof of Dual BL

- B a Brownian motion.
- u a drift satisfying $B(1) + \int_0^1 u(s)\ \mathrm{d}s = X$ in law.
- then $P_i B$ is a Brownian motion on E_i
- and $P_i B + \int_0^1 P_i u(s)\ \mathrm{d}s = P_i X$ in law
- Thus $\mathrm{H}(P_i X) \leq \mathrm{E} \int_0^1 |P_i u(s)|^2/2\ \mathrm{d}s$

Proof of Dual BL

- B a Brownian motion.
- u a drift satisfying $B(1) + \int_0^1 u(s)\ \mathrm{d}s = X$ in law.
- then $P_i B$ is a Brownian motion on E_i
- and $P_i B + \int_0^1 P_i u(s)\ \mathrm{d}s = P_i X$ in law
- Thus $\mathrm{H}(P_i X) \leq \mathrm{E} \int_0^1 |P_i u(s)|^2/2\ \mathrm{d}s$
- Using the frame condition

$$\sum c_i \, \mathrm{H}(P_i X) \leq \mathrm{E} \int_0^1 \sum c_i |P_i u(s)|^2/2\ \mathrm{d}s$$
$$= \mathrm{E} \int_0^1 |u(s)|^2/2\ \mathrm{d}s.$$

- Taking infimum on u yields the result.

Reversed Brascamp-Lieb inequality

Reversed Brascamp-Lieb Inequality (Barthe)

Under the frame condition, for all $h\colon E \to \mathbb{R}$ and $f_i\colon E_i \to \mathbb{R}$ satisying

$$\forall (x_1, \ldots, x_m) \in E_1 \times \cdots \times E_m, \quad h\Big(\sum_{i=1}^{m} c_i x_i\Big) \geq \sum_{i=1}^{m} c_i f_i(x_i)$$

Reversed Brascamp-Lieb inequality

Reversed Brascamp-Lieb Inequality (Barthe)

Under the frame condition, for all $h\colon E \to \mathbb{R}$ and $f_i\colon E_i \to \mathbb{R}$ satisying

$$\forall (x_1, \ldots, x_m) \in E_1 \times \cdots \times E_m, \quad h\Big(\sum_{i=1}^{m} c_i x_i\Big) \geq \sum_{i=1}^{m} c_i f_i(x_i)$$

we have

$$\int_E \mathrm{e}^h \, \mathrm{d}\gamma_E \geq \prod_{i=1}^{m} \Big(\int_{E_i} \mathrm{e}^{f_i} \, \mathrm{d}\gamma_{E_i} \Big)^{c_i}$$

Reversed Brascamp-Lieb inequality

Reversed Brascamp-Lieb Inequality (Barthe)

Under the frame condition, for all $h\colon E \to \mathbb{R}$ and $f_i\colon E_i \to \mathbb{R}$ satisying

$$\forall (x_1, \dots, x_m) \in E_1 \times \cdots \times E_m, \quad h\Big(\sum_{i=1}^m c_i x_i\Big) \geq \sum_{i=1}^m c_i f_i(x_i)$$

we have

$$\int_E \mathrm{e}^h \, \mathrm{d}\gamma_E \geq \prod_{i=1}^m \Big(\int_{E_i} \mathrm{e}^{f_i} \, \mathrm{d}\gamma_{E_i}\Big)^{c_i}$$

Remark

When $m = 2$, $E_1 = E_2 = E$, $c_1 + c_2 = 1$ this yields the log concavity of the Gaussian measure.
Related to the Brunn-Minkowski inequality.

Entropic RBL inequality

Dual version of RBL

Under the frame condition,
for all random vectors $X_1, \ldots, X_m$ on $E_1, \ldots, E_m$ respectively,

Entropic RBL inequality

Dual version of RBL

Under the frame condition,
for all random vectors $X_1, \ldots, X_m$ on $E_1, \ldots, E_m$ respectively,
there exists $Y_1, \ldots, Y_m$ satisfying $Y_i = X_i$ in law for all i and

Entropic RBL inequality

Dual version of RBL

Under the frame condition,
for all random vectors $X_1, \ldots, X_m$ on $E_1, \ldots, E_m$ respectively,
there exists $Y_1, \ldots, Y_m$ satisfying $Y_i = X_i$ in law for all i and

$$\mathrm{H}\Big(\sum_{i=1}^{m} c_i Y_i\Big) \le \sum_{i=1}^{m} c_i \,\mathrm{H}(X_i).$$

Entropic RBL inequality

Dual version of RBL

Under the frame condition,
for all random vectors $X_1, \dots, X_m$ on $E_1, \dots, E_m$ respectively, there exists $Y_1, \dots, Y_m$ satisfying $Y_i = X_i$ in law for all i and

$$\mathrm{H}\Big(\sum_{i=1}^m c_i Y_i\Big) \le \sum_{i=1}^m c_i \, \mathrm{H}(X_i).$$

Proof that this implies RBL:

Entropic RBL inequality

Dual version of RBL

Under the frame condition,
for all random vectors $X_1, \dots, X_m$ on $E_1, \dots, E_m$ respectively,
there exists $Y_1, \dots, Y_m$ satisfying $Y_i = X_i$ in law for all i and

$$\mathrm{H}\Big(\sum_{i=1}^{m} c_i Y_i\Big) \leq \sum_{i=1}^{m} c_i \, \mathrm{H}(X_i).$$

Proof that this implies RBL:

- Let $f_1, \dots, f_m, h$ satisfy the hypothesis of RBL.
- Let $X_1, \dots, X_m$ be random vectors on $E_1, \dots, E_m$ and $Y_1, \dots, Y_m$ be as above.

Entropic RBL implies RBL

- $\sum c_i f_i(x_i) \leq h(\sum c_i x_i)$ for all $(x_1, \ldots, x_m) \in E_1 \times \cdots \times E_m$.
- $Y_i = X_i$ in law for $i = 1, \ldots, m$.
- $\mathrm{H}(\sum c_i Y_i) \leq \sum c_i \, \mathrm{H}(X_i)$.

Entropic RBL implies RBL

- $\sum c_i f_i(x_i) \le h(\sum c_i x_i)$ for all $(x_1, \ldots, x_m) \in E_1 \times \cdots \times E_m$.
- $Y_i = X_i$ in law for $i = 1, \ldots, m$.
- $\mathrm{H}(\sum c_i Y_i) \le \sum c_i \, \mathrm{H}(X_i)$.
- Then

$$\begin{aligned} \sum c_i \big(\mathrm{E}\, f_i(X_i) - \mathrm{H}(X_i)\big) &= \mathrm{E}\Big(\sum c_i f_i(Y_i)\Big) - \sum c_i \, \mathrm{H}(X_i) \\ &\le \mathrm{E}\, h\Big(\sum c_i Y_i\Big) - \mathrm{H}\Big(\sum c_i Y_i\Big). \end{aligned}$$

Entropic RBL implies RBL

- $\sum c_i f_i(x_i) \leq h(\sum c_i x_i)$ for all $(x_1, \ldots, x_m) \in E_1 \times \cdots \times E_m$.
- $Y_i = X_i$ in law for $i = 1, \ldots, m$.
- $\mathrm{H}(\sum c_i Y_i) \leq \sum c_i \, \mathrm{H}(X_i)$.
- Then

$$\begin{aligned}\sum c_i \big(\mathrm{E}\, f_i(X_i) - \mathrm{H}(X_i)\big) &= \mathrm{E}\Big(\sum c_i f_i(Y_i)\Big) - \sum c_i \, \mathrm{H}(X_i) \\ &\leq \mathrm{E}\, h\Big(\sum c_i Y_i\Big) - \mathrm{H}\Big(\sum c_i Y_i\Big).\end{aligned}$$

- Using $\log\big(\int \mathrm{e}^f \, \mathrm{d}\gamma\big) = \sup_X \big(\mathrm{E}\, f(X) - \mathrm{H}(X)\big)$ we obtain

$$\prod_{i=1}^{m} \Big(\int_{E_i} \mathrm{e}^{f_i} \, \mathrm{d}\gamma_{E_i}\Big)^{c_i} \leq \int_E \mathrm{e}^h \, \mathrm{d}\gamma_E.$$

Proof of Entropic RBL

Let $X_1, \ldots, X_m$ be random vectors on $E_1, \ldots, E_m$.
Let B be a Brownian motion.

Proof of Entropic RBL

Let $X_1, \ldots, X_m$ be random vectors on $E_1, \ldots, E_m$.
Let B be a Brownian motion.

- Since $P_i B$ is a Brownian motion on E_i, there exists a drift u_i satisfying $P_i B(1) + \int_0^1 u_i(s)\,\mathrm{d}s = X_i$ in law and

$$\mathrm{H}(X_i) = \mathrm{E} \int_0^1 |u_i(s)|^2/2 \,\mathrm{d}s.$$

Proof of Entropic RBL

Let $X_1, \ldots, X_m$ be random vectors on $E_1, \ldots, E_m$.
Let B be a Brownian motion.

- Since $P_i B$ is a Brownian motion on E_i, there exists a drift u_i satisfying $P_i B(1) + \int_0^1 u_i(s)\, \mathrm{d}s = X_i$ in law and
$$\mathrm{H}(X_i) = \mathrm{E} \int_0^1 |u_i(s)|^2/2 \, \mathrm{d}s.$$
- Let $Y_i = P_i B(1) + \int_0^1 u_i(s)\, \mathrm{d}s$.

Proof of Entropic RBL

Let $X_1, \ldots, X_m$ be random vectors on $E_1, \ldots, E_m$.
Let B be a Brownian motion.

- Since $P_i B$ is a Brownian motion on E_i, there exists a drift u_i satisfying $P_i B(1) + \int_0^1 u_i(s)\,\mathrm{d}s = X_i$ in law and
$$\mathrm{H}(X_i) = \mathrm{E}\int_0^1 |u_i(s)|^2/2\,\mathrm{d}s.$$
- Let $Y_i = P_i B(1) + \int_0^1 u_i(s)\,\mathrm{d}s$.
- The frame condition yields $\sum c_i Y_i = B(1) + \int_0^1 \sum c_i u_i(s)\,\mathrm{d}s$.

Proof of Entropic RBL

Let $X_1, \ldots, X_m$ be random vectors on $E_1, \ldots, E_m$.
Let B be a Brownian motion.

- Since $P_i B$ is a Brownian motion on E_i, there exists a drift u_i satisfying $P_i B(1) + \int_0^1 u_i(s)\,\mathrm{d}s = X_i$ in law and

$$\mathrm{H}(X_i) = \mathrm{E} \int_0^1 |u_i(s)|^2/2 \,\mathrm{d}s.$$

- Let $Y_i = P_i B(1) + \int_0^1 u_i(s)\,\mathrm{d}s$.
- The frame condition yields $\sum c_i Y_i = B(1) + \int_0^1 \sum c_i u_i(s)\,\mathrm{d}s$.
- Hence (using the frame condition again)

$$\begin{aligned}
\mathrm{H}\Big(\sum c_i Y_i\Big) &\le \mathrm{E} \int_0^1 \Big|\sum c_i u_i(s)\Big|^2/2 \,\mathrm{d}s \\
&\le \sum c_i \,\mathrm{E} \int_0^1 |u_i(s)|^2/2 \,\mathrm{d}s \\
&= \sum c_i \,\mathrm{H}(X_i).
\end{aligned}$$