Free denoising

Kamil Szpojankowski

joint work with Maxime Fevrier and Alexandru Nica

arXiv:2412.20792

Probabilistic Operator Algebra Seminar, UC Berkeley

24th March 2025

Problem description

Free denoising

Setup

- a, b selfadjoint and free,
- a represents a "signal",
- b represents a "noise".
- We view $a + b$ a noisy version of a. More generally, we consider "noisy observations" of the form $P(a, b)$, where P is polynomial or expression like $a^{1/2}ba^{1/2}$.

Free denoising

Setup

- a, b selfadjoint and free,
- a represents a "signal",
- b represents a "noise".
- We view $a + b$ a noisy version of a. More generally, we consider "noisy observations" of the form $P(a, b)$, where P is polynomial or expression like $a^{1/2}ba^{1/2}$.

Goal

Approximate the signal a by a function of the noisy element $P(a, b)$.

Free denoising

Setup

- a, b selfadjoint and free,
- a represents a "signal",
- b represents a "noise".
- We view $a+b$ a noisy version of a. More generally, we consider "noisy observations" of the form $P(a,b)$, where P is polynomial or expression like $a^{1/2}ba^{1/2}$.

Goal

Approximate the signal a by a function of the noisy element $P(a,b)$.

Framework

We solve this in the framework of tracial W^* probability space $(\mathcal{A}, \varphi)$, where for $c \in \mathcal{A}$ we have L^2 norm given by $\|c\|_2^2 = \varphi(cc^*)$. If we want to minimize $\|a - h(a+b)\|_2$, then the solution is given by $h(a+b) = E(a|a+b)$.

Subordination reversed

Subordination (Biane, Voiculescu)

There exists $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that

$$E\left(\frac{1}{z-a-b} \mid a\right) = \frac{1}{\omega(z)-a}$$

Subordination reversed

Subordination (Biane, Voiculescu)

There exists $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that

$$E\left(\frac{1}{z-a-b} \mid a\right) = \frac{1}{\omega(z)-a}$$

- We reverse the roles of a and $a+b$ this is, in spirit, a non-commutative version of Bayes' formula.
- We know that $E((a+b)^n|a)$ is again a polynomial of degree n in a.
- The above cannot be true for the reversed conditioning. Bożejko and Bryc showed that if $E(a|a+b)$ is linear function of $a+b$ and $E(a^2|a+b)$ is quadratic, then a, b come from free Meixner family.

Recap conditional expectation

Conditional expectation

If $x, y \in \mathcal{A}$, are selfadjoint then $E(x \mid y) = h(y)$, where h is a real-valued bounded Borel function on $Spec(y)$ such that

$$\varphi(xg(y)) = \varphi(h(y)g(y)), \quad \text{for every } g \text{ bounded Borel function on } Spec(y).$$

Recap conditional expectation

Conditional expectation

If $x, y \in \mathcal{A}$, are selfadjoint then $E(x \mid y) = h(y)$, where h is a real-valued bounded Borel function on $Spec(y)$ such that

$$\varphi(xg(y)) = \varphi(h(y)g(y)), \quad \text{for every } g \text{ bounded Borel function on } Spec(y).$$

How we compute it?

We want to take $x = a$ and $y = a + b$ and develop a systematic way of calculating $E(f(a)|a + b)$ for f bounded Borel functions on $Spec(a)$.

Two methods, both inspired by classical probability approach:

- Suitable point of view on the joint distribution of $(a, a + b)$ – overlap measure.
- Interpretation of conditional expectation as Radon-Nikodym derivative – conditional freeness of Bożejko, Leinert and Speicher.

Overlap measure

Classical probability approach

Classical probability

For (X,Y) with joint distribution $\mu_{(X,Y)}$ in order to find the conditional expectation $\mathbb{E}(f(X)|Y)$ we need to find the conditional distribution $\mu_{X|Y=t}$ (aka disintegration of $\mu_{(X,Y)}$) and then $\mathbb{E}(f(X)|Y) = h(Y)$ with

$$h(t) = \int_{\mathbb{R}} f(x)\mu_{X|Y=t}(dx).$$

Classical probability approach

Classical probability

For (X, Y) with joint distribution $\mu_{(X,Y)}$ in order to find the conditional expectation $\mathbb{E}(f(X)|Y)$ we need to find the conditional distribution $\mu_{X|Y=t}$ (aka disintegration of $\mu_{(X,Y)}$) and then $\mathbb{E}(f(X)|Y) = h(Y)$ with

$$h(t) = \int_{\mathbb{R}} f(x)\mu_{X|Y=t}(dx).$$

Example

If distribution of (X, Y) has a density wrt Lebesgue measure on $\mathbb{R}^2$ then

$$f_{X|Y=t}(x) = \frac{f_{(X,Y)}(x, t)}{f_Y(t)}.$$

Classical probability approach

Classical probability

For (X, Y) with joint distribution $\mu_{(X,Y)}$ in order to find the conditional expectation $\mathbb{E}(f(X)|Y)$ we need to find the conditional distribution $\mu_{X|Y=t}$ (aka disintegration of $\mu_{(X,Y)}$) and then $\mathbb{E}(f(X)|Y) = h(Y)$ with

$$h(t) = \int_{\mathbb{R}} f(x)\mu_{X|Y=t}(dx).$$

Example

If distribution of (X, Y) has a density wrt Lebesgue measure on $\mathbb{R}^2$ then

$$f_{X|Y=t}(x) = \frac{f_{(X,Y)}(x,t)}{f_Y(t)}.$$

Moral: joint distribution gives access to conditional expectations!

Overlap measure

We return to general selfadjoint elements $x, y \in \mathcal{A}$.

Definition

Non-commutative joint distribution is the linear functional $\mu_{(x,y)} : \mathbb{C}\langle X_1, X_2\rangle \to \mathbb{C}$ given by

$$\mu_{(x,y)}(P) = \varphi(P(x, y)).$$

Overlap measure

We return to general selfadjoint elements $x, y \in \mathcal{A}$.

Definition

Non-commutative joint distribution is the linear functional $\mu_{(x,y)} : \mathbb{C}\langle X_1, X_2\rangle \to \mathbb{C}$ given by

$$\mu_{(x,y)}(P) = \varphi(P(x, y)).$$

Observation: $\mu_{(x,y)}$ contains more information than what is needed for conditional expectations. We need to find h such that $\varphi(f(x)g(y)) = \varphi(h(y)g(y))$.

Overlap measure

We return to general selfadjoint elements $x, y \in \mathcal{A}$.

Definition

Non-commutative joint distribution is the linear functional $\mu_{(x,y)} : \mathbb{C}\langle X_1, X_2\rangle \to \mathbb{C}$ given by

$$\mu_{(x,y)}(P) = \varphi(P(x,y)).$$

Observation: $\mu_{(x,y)}$ contains more information than what is needed for conditional expectations. We need to find h such that $\varphi(f(x)g(y)) = \varphi(h(y)g(y))$.

Idea: Is there a probability measure on $\mathbb{R}^2$ such that

$$\varphi(f(x)g(y)) = \iint_{\mathbb{R}^2} f(s)g(t)d\mu(s,t)\,?$$

Overlap measure existence

Proposition

There exists a probability measure μ on the Borel sigma-algebra of $\mathbb{R}^2$, uniquely determined, such that the following holds:

$$\begin{cases} \text{For every bounded Borel functions } f, g \\ \text{one has that } \int_{\mathbb{R}^2} f(s)g(t)d\mu(s,t) = \varphi(f(x)g(y)). \end{cases}$$

We will refer to this probability measure μ as the *overlap measure* of x and y, and we will denote it as $\mu_{x,y}^{(\mathrm{ov})}$.

Overlap measure existence

Proposition

There exists a probability measure μ on the Borel sigma-algebra of $\mathbb{R}^2$, uniquely determined, such that the following holds:

$$\begin{cases} \text{For every bounded Borel functions } f, g \\ \text{one has that } \int_{\mathbb{R}^2} f(s)g(t)d\mu(s,t) = \varphi(f(x)g(y)). \end{cases}$$

We will refer to this probability measure μ as the *overlap measure* of x and y, and we will denote it as $\mu_{x,y}^{(\mathrm{ov})}$.

This was known since long ago in the operator algebra literature (Connes 1976).

Overlap measure existence

Proposition

There exists a probability measure μ on the Borel sigma-algebra of $\mathbb{R}^2$, uniquely determined, such that the following holds:

$$\begin{cases} \text{For every bounded Borel functions } f, g \\ \text{one has that } \int_{\mathbb{R}^2} f(s)g(t)d\mu(s,t) = \varphi(f(x)g(y)). \end{cases}$$

We will refer to this probability measure μ as the *overlap measure* of x and y, and we will denote it as $\mu_{x,y}^{(\mathrm{ov})}$.

This was known since long ago in the operator algebra literature (Connes 1976).

Idea of the proof: For any $s, t \in \mathbb{R}$ define $P(s) = 1_{(-\infty,s]}(x)$ and $Q(t) = 1_{(-\infty,t]}(y)$ and verify that $F(s,t) = \varphi(P(s)Q(t))$ is a two dimensional cdf.

Overlap measure existence

Proposition

There exists a probability measure μ on the Borel sigma-algebra of $\mathbb{R}^2$, uniquely determined, such that the following holds:

$$\begin{cases} \text{For every bounded Borel functions } f, g \\ \text{one has that } \int_{\mathbb{R}^2} f(s)g(t)d\mu(s,t) = \varphi(f(x)g(y)). \end{cases}$$

We will refer to this probability measure μ as the *overlap measure* of x and y, and we will denote it as $\mu_{x,y}^{(\mathrm{ov})}$.

This was known since long ago in the operator algebra literature (Connes 1976).

Idea of the proof: For any $s, t \in \mathbb{R}$ define $P(s) = 1_{(-\infty,s]}(x)$ and $Q(t) = 1_{(-\infty,t]}(y)$ and verify that $F(s,t) = \varphi(P(s)Q(t))$ is a two dimensional cdf.

Observation: Marginals of $\mu_{x,y}^{(\mathrm{ov})}$ are equal to μ_x and μ_y (to see that take $g \equiv 1$ and $f \equiv 1$ respectively). So $\mu_{x,y}^{(\mathrm{ov})}$ is a coupling of μ_x and μ_y.

Overlap measure for matrices

Consider the case $\mathcal{A} = \mathcal{M}_N(\mathbb{C})$ and $\varphi = \mathrm{tr}_N$ for some $N \in \mathbb{N}$, and $x, y \in \mathcal{A}$ both selfadjoint.

Case of matrices

Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of x, and let $\rho_1, \ldots, \rho_N$ be the eigenvalues of y, counted with multiplicities.

Overlap measure for matrices

Consider the case $\mathcal{A} = \mathcal{M}_N(\mathbb{C})$ and $\varphi = \mathrm{tr}_N$ for some $N \in \mathbb{N}$, and $x, y \in \mathcal{A}$ both selfadjoint.

Case of matrices

Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of x, and let $\rho_1, \ldots, \rho_N$ be the eigenvalues of y, counted with multiplicities.

Let $u_1, \ldots, u_N$ and $v_1, \ldots, v_N$ be orthonormal bases for $\mathbb{C}^N$, such that $x(u_k) = \lambda_k u_k$ for all $1 \leq k \leq N$ and $y(v_\ell) = \rho_\ell v_\ell$ for all $1 \leq \ell \leq N$.

Overlap measure for matrices

Consider the case $\mathcal{A} = \mathcal{M}_N(\mathbb{C})$ and $\varphi = \mathrm{tr}_N$ for some $N \in \mathbb{N}$, and $x, y \in \mathcal{A}$ both selfadjoint.

Case of matrices

Let $\lambda_1, \dots, \lambda_N$ be the eigenvalues of x, and let $\rho_1, \dots, \rho_N$ be the eigenvalues of y, counted with multiplicities.

Let $u_1, \dots, u_N$ and $v_1, \dots, v_N$ be orthonormal bases for $\mathbb{C}^N$, such that $x(u_k) = \lambda_k u_k$ for all $1 \leq k \leq N$ and $y(v_\ell) = \rho_\ell v_\ell$ for all $1 \leq \ell \leq N$.

Then: $\mu_{x,y}^{(\mathrm{ov})} = \sum_{k,\ell=1}^{N} \frac{|\langle u_k\,,\,v_\ell \rangle|^2}{N}\, \delta_{(\lambda_k, \rho_\ell)}$, where "$\delta_{(\lambda,\rho)}$" stands for the Dirac mass concentrated at the point $(\lambda, \rho) \in \mathbb{R}^2$.

Overlap measure for matrices

Consider the case $\mathcal{A} = \mathcal{M}_N(\mathbb{C})$ and $\varphi = \mathrm{tr}_N$ for some $N \in \mathbb{N}$, and $x, y \in \mathcal{A}$ both selfadjoint.

Case of matrices

Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of x, and let $\rho_1, \ldots, \rho_N$ be the eigenvalues of y, counted with multiplicities.

Let $u_1, \ldots, u_N$ and $v_1, \ldots, v_N$ be orthonormal bases for $\mathbb{C}^N$, such that $x(u_k) = \lambda_k u_k$ for all $1 \leq k \leq N$ and $y(v_\ell) = \rho_\ell v_\ell$ for all $1 \leq \ell \leq N$.

Then: $\mu_{x,y}^{(\mathrm{ov})} = \sum_{k,\ell=1}^{N} \frac{|\langle u_k \, , \, v_\ell \rangle|^2}{N} \, \delta_{(\lambda_k, \rho_\ell)}$, where "$\delta_{(\lambda,\rho)}$" stands for the Dirac mass concentrated at the point $(\lambda, \rho) \in \mathbb{R}^2$.

Observation: In this case $\mu_{x,y}^{(\mathrm{ov})}$ is absolutely continuous with respect to the product measure $\mu_x \times \mu_y$.

Absolute continuity

Definition

Suppose that the overlap measure $\mu_{x,y}^{(\mathrm{ov})}$ is absolutely continuous with respect to the direct product of its marginals μ_x and μ_y. The Radon-Nikodym derivative

$$\mathrm{o}_{x,y} := \frac{d\mu_{x,y}^{(\mathrm{ov})}}{d(\mu_x \times \mu_y)}$$

will be referred to as the *overlap function* of x and y.

Absolute continuity

Definition

Suppose that the overlap measure $\mu_{x,y}^{(\mathrm{ov})}$ is absolutely continuous with respect to the direct product of its marginals μ_x and μ_y. The Radon-Nikodym derivative

$$\mathrm{o}_{x,y} := \frac{d\mu_{x,y}^{(\mathrm{ov})}}{d(\mu_x \times \mu_y)}$$

will be referred to as the *overlap function* of x and y.

In the matrix example we have $\mathrm{o}_{x,y}(\lambda_k, \rho_l) = N \mid \langle u_k \, , \, v_\ell \rangle \mid^2$.

Conditional expectation via overlap function

Conditional expectation from the overlap function

Suppose x, y are such that the overlap function exists. Then we have

$$\varphi(f(x)g(y)) = \iint_{\mathbb{R}^2} f(s)g(t)d\mu_{x,y}^{(\mathrm{ov})}(s,t) = \iint_{\mathbb{R}^2} f(s)g(t)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\mu_y(dt)$$

Conditional expectation via overlap function

Conditional expectation from the overlap function

Suppose x, y are such that the overlap function exists. Then we have

$$\varphi(f(x)g(y)) = \iint_{\mathbb{R}^2} f(s)g(t)d\mu^{(\mathrm{ov})}_{x,y}(s,t) = \iint_{\mathbb{R}^2} f(s)g(t)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\mu_y(dt)$$

$$= \int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} f(s)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\right)}_{:=h(t)} g(t)\mu_y(dt) = \varphi(h(y)g(y)).$$

Conditional expectation via overlap function

Conditional expectation from the overlap function

Suppose x, y are such that the overlap function exists. Then we have

$$\varphi(f(x)g(y)) = \iint_{\mathbb{R}^2} f(s)g(t)d\mu^{(\mathrm{ov})}_{x,y}(s,t) = \iint_{\mathbb{R}^2} f(s)g(t)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\mu_y(dt)$$

$$= \int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} f(s)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\right)}_{:=h(t)} g(t)\mu_y(dt) = \varphi(h(y)g(y)).$$

Hence $E(f(x)|y) = h(y)$.

Conditional expectation via overlap function

Conditional expectation from the overlap function

Suppose x, y are such that the overlap function exists. Then we have

$$\varphi(f(x)g(y)) = \iint_{\mathbb{R}^2} f(s)g(t)d\mu^{(\mathrm{ov})}_{x,y}(s,t) = \iint_{\mathbb{R}^2} f(s)g(t)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\mu_y(dt)$$

$$= \int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} f(s)\mathrm{o}_{x,y}(s,t)\mu_x(ds)\right)}_{:=h(t)} g(t)\mu_y(dt) = \varphi(h(y)g(y)).$$

Hence $E(f(x)|y) = h(y)$.

Observation: Overlap measure $\mu^{(\mathrm{ov})}_{x,y}$ is not always absolutely continuous with respect to $\mu_x \times \mu_y$, take for example $y = x$, then $\mu^{(\mathrm{ov})}_{x,y}$ is concentrated on a subset of the line $\{(s,s), s \in \mathbb{R}\} \subset \mathbb{R}^2$.

Overlap measure in the free additive case

Theorem 1

Let a, b be free, with $a \sim \mu$ and $b \sim \nu$ and assume that neither of μ, ν is a point mass. Then $d\mu^{(\mathrm{ov})}_{a,a+b}(s,t) = \mathrm{o}_{a,a+b}(s,t)\mu(ds)\, \mu \boxplus \nu(dt)$ where

$$\mathrm{o}_{a,a+b}(s,t) = \begin{cases} -\frac{1}{\pi}\frac{1}{f_{\mu\boxplus\nu}(t)} \,\mathrm{Im}\left(\frac{1}{\omega(t)-s}\right) & \text{if } t \in U, \\ \frac{1}{\mu(\{s\})} 1_{\omega(t)=s} & \text{if } t \text{ is an atom of } \mu \boxplus \nu \end{cases}$$

is defined in the $\mu \times (\mu \boxplus \nu)$ a.e. sense.

Here $f_{\mu\boxplus\nu}$ is the density of the absolutely continuous part of $\mu \boxplus \nu$ and U is the set where the density is positive, ω as before denotes the additive subordination function.

Overlap measure in the free additive case

Overlap measure in the free additive case

Line of proof of Theorem 1, first part: use subordination

- Given additive subordination function ω one defines (Biane 1998) a family of probability measures k_s via

$$\int_{\mathbb{R}} \frac{1}{z-t} k_s(dt) = \frac{1}{\omega(z)-s}, \quad \forall\, s \in \mathbb{R} \text{ and } z \in \mathbb{C}^+.$$

Overlap measure in the free additive case

Line of proof of Theorem 1, first part: use subordination

- Given additive subordination function ω one defines (Biane 1998) a family of probability measures k_s via

$$\int_{\mathbb{R}} \frac{1}{z-t} k_s(dt) = \frac{1}{\omega(z)-s}, \quad \forall\, s \in \mathbb{R} \text{ and } z \in \mathbb{C}^+.$$

- The measures k_s together give a Feller-Markov kernel $\mathcal{K}g(s) = \int_{\mathbb{R}} g(t) k_s(dt), \quad s \in \mathbb{R}$, with

$$E(g(a+b)|a) = \mathcal{K}g(a).$$

Overlap measure in the free additive case

Line of proof of Theorem 1, first part: use subordination

- Given additive subordination function ω one defines (Biane 1998) a family of probability measures k_s via

$$\int_{\mathbb{R}} \frac{1}{z-t} k_s(dt) = \frac{1}{\omega(z)-s}, \quad \forall\, s \in \mathbb{R} \text{ and } z \in \mathbb{C}^+.$$

- The measures k_s together give a Feller-Markov kernel $\mathcal{K}g(s) = \int_{\mathbb{R}} g(t) k_s(dt), \quad s \in \mathbb{R}$, with

$$E(g(a+b)|a) = \mathcal{K}g(a).$$

Thus

$$\varphi(f(a)g(a+b)) = \varphi(f(a)\mathcal{K}g(a)) = \int_{\mathbb{R}} f(s)\mathcal{K}g(s)\mu(ds) = \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t)k_s(dt)\mu(ds).$$

Overlap measure in the additive case

Line of proof of Theorem 1, second part.

Repeat equation obtained on preceding slide:

$$\varphi(f(a)g(a+b)) = \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t) k_s(dt) \mu(ds).$$

Overlap measure in the additive case

Line of proof of Theorem 1, second part.

Repeat equation obtained on preceding slide:

$$
\begin{aligned}
\varphi(f(a)g(a+b)) &= \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t) k_s(dt) \mu(ds). \\
&\overset{?}{=} \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t) o(s,t) \mu \boxplus \nu(dt) \mu(ds).
\end{aligned}
$$

Overlap measure in the additive case

Line of proof of Theorem 1, second part.

Repeat equation obtained on preceding slide:

$$\begin{aligned}\varphi(f(a)g(a+b)) &= \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t)k_s(dt)\mu(ds). \\ &\stackrel{?}{=} \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t)o(s,t)\mu \boxplus \nu(dt)\mu(ds).\end{aligned}$$

Consequence: in order to show absolute continuity of $\mu^{(\mathrm{ov})}_{a,a+b}$ with respect to product of marginals it is enough to show that for μ a.e. s measure k_s is absolutely continuous with respect to $\mu \boxplus \nu$.

Overlap measure in the additive case

Line of proof of Theorem 1, second part.

Repeat equation obtained on preceding slide:

$$
\begin{aligned}
\varphi(f(a)g(a+b)) &= \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t) k_s(dt) \mu(ds). \\
&\stackrel{?}{=} \int_{\mathbb{R}} f(s) \int_{\mathbb{R}} g(t) o(s,t) \mu \boxplus \nu(dt) \mu(ds).
\end{aligned}
$$

Consequence: in order to show absolute continuity of $\mu^{(\mathrm{ov})}_{a,a+b}$ with respect to product of marginals it is enough to show that for μ a.e. s measure k_s is absolutely continuous with respect to $\mu \boxplus \nu$. Recall that

$$
\int_{\mathbb{R}} \frac{1}{z-t} k_s(dt) = \frac{1}{\omega(z)-s}, \quad \forall\, s \in \mathbb{R} \text{ and } z \in \mathbb{C}^+.
$$

Results of Belinschi (2006) about existence and behaviour of continuous continuation of ω to $\mathbb{R}$ together with regularity properties of $\boxplus$ allow to conclude.

Free denoiser – additive case

Theorem 2

For every bounded Borel function f we have

$$E[f(a) \mid a+b] = h(a+b),$$

where $h : \mathrm{Spec}(a+b) \to \mathbb{C}$ is defined (in $\mu \boxplus \nu$-almost everywhere sense) by

$$h(t) = \begin{cases} -\frac{1}{\pi}\frac{1}{f_{\mu\boxplus\nu}(t)} \operatorname{Im}\left(\int_{\mathbb{R}} f(s)\frac{1}{\omega(t)-s}d\mu(s)\right) & \text{if } t \in U, \\ f(\omega(t)) & \text{if } t \text{ is an atom of } \mu \boxplus \nu. \end{cases}$$

Free denoiser – additive case

Theorem 2

For every bounded Borel function f we have

$$E[f(a) \mid a+b] = h(a+b),$$

where $h : \mathrm{Spec}(a+b) \to \mathbb{C}$ is defined (in $\mu \boxplus \nu$-almost everywhere sense) by

$$h(t) = \begin{cases} -\frac{1}{\pi} \frac{1}{f_{\mu \boxplus \nu}(t)} \operatorname{Im}\left(\int_{\mathbb{R}} f(s) \frac{1}{\omega(t)-s} d\mu(s) \right) & \text{if } t \in U, \\ f(\omega(t)) & \text{if } t \text{ is an atom of } \mu \boxplus \nu. \end{cases}$$

Note the use of "$\omega(t)$" in the formula, hence we are invoking $\omega(a+b)$ (somewhat different, from how ω is typically used).

Tweedie's formula

Classical case

Take X, Y independent (in the commutative sense) and $Y \sim \mathcal{N}(0, \sigma^2)$. The well-known Tweedie's formula in statistics says that

$$\mathbb{E}(X \mid X + Y) = X + Y + \sigma^2 g(X + Y), \quad \text{where } g(t) = \tfrac{d}{dt} \log(f_{X+Y}(t)).$$

Remarkably $h(t) = t + \sigma^2 \frac{d}{dt} \log(f_{X+Y}(t))$ depends only on the distribution of $X + Y$.

Tweedie's formula

Classical case

Take X, Y independent (in the commutative sense) and $Y \sim \mathcal{N}(0, \sigma^2)$. The well-known Tweedie's formula in statistics says that

$$\mathbb{E}(X \mid X+Y) = X + Y + \sigma^2 g(X+Y), \quad \text{where } g(t) = \tfrac{d}{dt}\log(f_{X+Y}(t)).$$

Remarkably $h(t) = t + \sigma^2 \frac{d}{dt}\log(f_{X+Y}(t))$ depends only on the distribution of $X+Y$.

Free case

If a and b are free and b has Wigner semicircle law with variance σ^2, then

$$E(a \,|\, a+b) = a + b - 2\pi\,\sigma^2\, H_{\mu\boxplus\nu}(a+b),$$

where $H_{\mu\boxplus\nu}$ is the Hilbert transform of the measure $\mu \boxplus \nu$.

This formula was found before in work of Allez, Bouchaud, Bun and Potters in the context of matrix denoising.

Multiplicative case

The multiplicative case works in an analogous way. The formulas are a bit more complicated. So we only illustrate here the example of free Poisson noise.

Free Poisson noise

If a, b are free, $a \geq 0$ and b has free Poisson distribution of parameter $\lambda > 0$ then $E(a \mid a^{1/2}ba^{1/2}) = h(a^{1/2}ba^{1/2})$ with

$$h(t) = \begin{cases} \frac{\lambda t}{|\lambda - 1 + tG_{\mu \boxtimes \nu}(t)|^2} & \text{if } t \in \operatorname{supp}(\mu \boxtimes \nu) \setminus \{0\}, \\ -\frac{\lambda}{(1-\lambda)G_{(\mu \boxtimes \nu)^{ac}}(0)} & \text{if } t = 0 \text{ when } \lambda < 1, \end{cases}$$

The above is well known shrinkage estimator of covariance matrix of Ledoit and Peche.

Free denoiser and conditional freeness

Free denoiser as a Radon–Nikodym derivative

How to go beyond additive and multiplicative case?

Our ncps is $(\mathcal{A}, \varphi)$. The goal is to find $E(f(a) \mid P(a,b)) = h(P(a,b))$ for a non-commutative polynomial P.

Theorem 3

If $f \geq 0$ and $\varphi(f(a)) = 1$, then we define a new state $\chi : \mathcal{A} \to \mathbb{C}$ by

$$\chi(c) := \varphi(\ f(a)c\ ), \quad c \in \mathcal{A}.$$

Then one has

$$\frac{d\mu^{\chi}_{P(a,b)}}{d\mu^{\varphi}_{P(a,b)}} = h \quad \text{(free denoiser).}$$

For general f we consider $f = f^+ - f^-$.

Proof

Proof of Theorem 3

$$\int_{\mathbb{R}} g(t)\mu^{\chi}_{P(a,b)}(dt) = \chi\big(\, g(\, P(a,b)\,)\,\big) = \varphi\big(\, f(a)\, g(\, P(a,b)\,)\,\big).$$

Where we only used the definition on χ.

Proof

Proof of Theorem 3

$$\int_{\mathbb{R}} g(t)\mu^{\chi}_{P(a,b)}(dt) = \chi\big(g(P(a,b) \big) = \varphi\big(f(a)\, g(P(a,b) \big).$$

Where we only used the definition on χ.

Next taking the conditional expectation inside we get

$$= \varphi\Big(E\big(f(a)\, g(P(a,b)) \mid P(a,b)\big) \Big) = \varphi\Big(E\big(f(a) \mid P(a,b)\big) \cdot g(P(a,b)) \Big).$$

Proof

Proof of Theorem 3

$$\int_{\mathbb{R}} g(t)\mu^{\chi}_{P(a,b)}(dt) = \chi\big(g(P(a,b))\big) = \varphi\big(f(a)\,g(P(a,b))\big).$$

Where we only used the definition on χ.

Next taking the conditional expectation inside we get

$$= \varphi\Big(E\big(f(a)\,g(P(a,b)) \mid P(a,b)\big)\Big) = \varphi\Big(E\big(f(a) \mid P(a,b)\big)\cdot g(P(a,b))\Big).$$

Since $E\big(f(a) \mid P(a,b)\big) = h(P(a,b))$ we get

$$= \varphi\big(h(P(a,b))\cdot g(P(a,b))\big) = \int_{\mathbb{R}} h(t)g(t)\mu^{\varphi}_{P(a,b)}(dt).$$

Thus

$$\int_{\mathbb{R}} g(t)\mu^{\chi}_{P(a,b)}(dt) = \int_{\mathbb{R}} g(t)h(t)\mu^{\varphi}_{P(a,b)}(dt).$$

From freeness to c-freeness

Proposition

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $(\mathcal{A}_i)_{i \in I}$ be freely independent unital subalgebras in $(\mathcal{A}, \varphi)$. Fix $i_0 \in I$ and a positive element $x \in \mathcal{A}_{i_0}$ such that $\varphi(x) = 1$. Define another state $\chi : \mathcal{A} \to \mathbb{C}$ by

$$\chi(y) := \varphi(xy), \qquad \forall y \in \mathcal{A}.$$

Then the subalgebras $(\mathcal{A}_i)_{i \in I}$ are conditionally freely independent in $(\mathcal{A}, \varphi, \chi)$.

Proof

Proof Let $n \geq 2$, $i_1 \neq i_2 \neq \ldots \neq i_n$ in I and $a_1 \in \mathcal{A}_{i_1}, \ldots, a_n \in \mathcal{A}_{i_n}$ be such that $\varphi(a_1) = \ldots = \varphi(a_n) = 0$.

Proof

Proof Let $n \geq 2$, $i_1 \neq i_2 \neq \ldots \neq i_n$ in I and $a_1 \in \mathcal{A}_{i_1}, \ldots, a_n \in \mathcal{A}_{i_n}$ be such that $\varphi(a_1) = \ldots = \varphi(a_n) = 0$. We want to show that

$$\chi(a_1 \cdots a_n) = \chi(a_1) \cdots \chi(a_n).$$

Proof

Proof Let $n \geq 2$, $i_1 \neq i_2 \neq \ldots \neq i_n$ in I and $a_1 \in \mathcal{A}_{i_1}, \ldots, a_n \in \mathcal{A}_{i_n}$ be such that $\varphi(a_1) = \ldots = \varphi(a_n) = 0$. We want to show that

$$\chi(a_1 \cdots a_n) = \chi(a_1) \cdots \chi(a_n).$$

There is $k \in \{1, \ldots, n\}$ such that $i_k \neq i_0$. Then, by free independence of $\chi(a_k) = \varphi(xa_k) = \varphi(x)\varphi(a_k) = 0$ so that the RHS vanishes.

Proof

Proof Let $n \geq 2$, $i_1 \neq i_2 \neq \ldots \neq i_n$ in I and $a_1 \in \mathcal{A}_{i_1}, \ldots, a_n \in \mathcal{A}_{i_n}$ be such that $\varphi(a_1) = \ldots = \varphi(a_n) = 0$. We want to show that

$$\chi(a_1 \cdots a_n) = \chi(a_1) \cdots \chi(a_n).$$

There is $k \in \{1, \ldots, n\}$ such that $i_k \neq i_0$. Then, by free independence of $\chi(a_k) = \varphi(xa_k) = \varphi(x)\varphi(a_k) = 0$ so that the RHS vanishes.

To prove that the LHS of vanishes, there are two cases: $i_0 \neq i_1$ or $i_0 = i_1$.

Proof

Proof Let $n \geq 2$, $i_1 \neq i_2 \neq \ldots \neq i_n$ in I and $a_1 \in \mathcal{A}_{i_1}, \ldots, a_n \in \mathcal{A}_{i_n}$ be such that $\varphi(a_1) = \ldots = \varphi(a_n) = 0$. We want to show that

$$\chi(a_1 \cdots a_n) = \chi(a_1) \cdots \chi(a_n).$$

There is $k \in \{1, \ldots, n\}$ such that $i_k \neq i_0$. Then, by free independence of $\chi(a_k) = \varphi(xa_k) = \varphi(x)\varphi(a_k) = 0$ so that the RHS vanishes.

To prove that the LHS of vanishes, there are two cases: $i_0 \neq i_1$ or $i_0 = i_1$.

Consider $i_0 \neq i_1$, then

$$\begin{aligned} \chi(a_1 \cdots a_n) &= \varphi(xa_1 \cdots a_n) \\ &= \varphi((x-1)a_1 \cdots a_n) + \varphi(a_1 \cdots a_n) \\ &= 0, \end{aligned}$$

where we used twice in the last line free independence of $(\mathcal{A}_i)_{i \in I}$.

Free denoiser for general polynomials

Theorem 4

Let a, b be free with distributions μ, ν. Let P be a selfadjoint polynomial then

$$E(f(a) \mid P(a,b)) = \frac{d\left((\mu, f \cdot \mu)\square_c^P(\nu,\nu)\right)}{d\left(\mu \square^P \nu\right)}(P(a,b)).$$

Free denoiser for general polynomials

Theorem 4

Let a, b be free with distributions μ, ν. Let P be a selfadjoint polynomial then

$$E(f(a) \mid P(a,b)) = \frac{d\left((\mu, f\cdot\mu)\square_c^P(\nu,\nu)\right)}{d\left(\mu\square^P\nu\right)}(P(a,b)).$$

Proof: Let h be the free denoiser, that is, the bounded Borel-measurable function $h : \mathbb{R} \to \mathbb{R}$ such that

$$E(f(a) \mid P(a,b)) = h(P(a,b)).$$

It is given by the Radon-Nikodym derivative

$$h = \frac{d\mu^{\chi}_{P(a,b)}}{d\mu^{\varphi}_{P(a,b)}}.$$

Observe that $a \sim (\mu, f\cdot\mu)$ while by freeness $b \sim (\nu,\nu)$. Then, according to Proposition $(\mu^{\varphi}_{P(a,b)}, \mu^{\chi}_{P(a,b)}) = (\mu\square^P\nu, (\mu, f\cdot\mu)\square_c^P(\nu,\nu))$.

Matrix denoising

Convergence of overlap measure

Brief comments on how our work relates to matrix denoising.

Proposition

Let $(\mathcal{A}, \varphi)$ and, for each $N \in \mathbb{N}$, (x_N, y_N) be a pair of selfadjoint non-commutative random variables in some non-commutative probability space $(\mathcal{A}_N, \varphi_N)$. If $((x_N, y_N))_{N\in\mathbb{N}}$ converges in non-commutative distribution to $(x, y) \in \mathcal{A}^2$, then the sequence $(\mu^{(\mathrm{ov})}_{x_N,y_N})_{N\in\mathbb{N}}$ of probability measures on $\mathbb{R}^2$ converges weakly to $\mu^{(\mathrm{ov})}_{x,y}$.

Convergence of overlap measure

Brief comments on how our work relates to matrix denoising.

Proposition

Let $(\mathcal{A}, \varphi)$ and, for each $N \in \mathbb{N}$, (x_N, y_N) be a pair of selfadjoint non-commutative random variables in some non-commutative probability space $(\mathcal{A}_N, \varphi_N)$. If $((x_N, y_N))_{N\in\mathbb{N}}$ converges in non-commutative distribution to $(x, y) \in \mathcal{A}^2$, then the sequence $(\mu^{(\mathrm{ov})}_{x_N,y_N})_{N\in\mathbb{N}}$ of probability measures on $\mathbb{R}^2$ converges weakly to $\mu^{(\mathrm{ov})}_{x,y}$.

Corollary

If $((x_N, y_N))_{N\in\mathbb{N}}$ converges in non-commutative distribution to (x, y) and $h : \mathbb{R} \to \mathbb{R}$ is bounded Borel-measurable μ_y-almost everywhere continuous, then

$$\varphi_N((x_N - h(y_N))^2) \underset{N\to\infty}{\longrightarrow} \varphi((x - h(y)^2).$$

Additive noise for matrices

Matrix case

In particular if we have a sequence of random matrices A_N, B_N, where:

- A_N, B_N are almost surely asymptotically free, i.e. converge in nc-distribution to (a, b).
- A_N is signal and B_N is noise,
- we observe $A_N + B_N$ and want to approximate A_N.

Additive noise for matrices

Matrix case

In particular if we have a sequence of random matrices A_N, B_N, where:

- A_N, B_N are almost surely asymptotically free, i.e. converge in nc-distribution to (a, b).
- A_N is signal and B_N is noise,
- we observe $A_N + B_N$ and want to approximate A_N.

Then (if h is almost surely continuous with respect to distribution of $a + b$) then application the free additive denoiser to $A_N + B_N$ is asymptotically optimal in the sense that

$$tr\left((A_N - h(A_N + B_N))^2\right) \to \varphi\left((a - h(a + b))^2\right).$$

Thank you!