CIMAT
Centro de Investigación en Matemáticas, A.C.

Free Berry-Esseen theorem via Stein's method

Arturo Jaramillo Gil (joint work with Mario Diaz)

Centro de Investigación en Matemáticas (CIMAT)

Objective

Let $\{\mu_{k,n}\ ;\ k,n \geq 1\}$ be a sequence of centered probability measures, and define

$$\nu_n \ := \mu_{1,n} \boxplus \cdots \boxplus \mu_{n,n}.$$

Objective

Let $\{\mu_{k,n} \ ; \ k, n \geq 1\}$ be a sequence of centered probability measures, and define

$$\nu_n \ := \mu_{1,n} \boxplus \cdots \boxplus \mu_{n,n}.$$

Under the condition $Var[\nu_n] = 1$ plus (other technical conditions), we have that $\nu_n \overset{Law}{\approx} \mathbf{s}$, where $\mathbf{s}$ is the standard semicircular distribution.

Objective

Let $\{\mu_{k,n}\ ;\ k,n \geq 1\}$ be a sequence of centered probability measures, and define

$$\nu_n \ := \mu_{1,n} \boxplus \cdots \boxplus \mu_{n,n}.$$

Under the condition $Var[\nu_n] = 1$ plus (other technical conditions), we have that $\nu_n \overset{Law}{\approx} \mathbf{s}$, where $\mathbf{s}$ is the standard semicircular distribution.

Goal

Bound $d_{TV}(\nu_n, \mathbf{s})$ in a probabilistic way.

Elements of free probability

Let $\mathcal{A}$ be a unital C^*-algebra and $\tau : \mathcal{A} \to \mathbb{C}$ a positive unital linear functional.

Elements of free probability

Let $\mathcal{A}$ be a unital C^*-algebra and $\tau : \mathcal{A} \to \mathbb{C}$ a positive unital linear functional.

Definition
Let $\{A_n\}_{n\geq 1}$ be subalgebras of $\mathcal{A}$. Define $\bar{a} := a - \tau[a]$. We say that $\{A_n\}_{n\geq 1}$ are freely independent if

$$\tau[\bar{a}_1 \bar{a}_2 \cdots \bar{a}_k] = 0, \tag{1}$$

for $a_1, \ldots, a_k$ alternating. Sums of free random variables yields the free convolution $\boxplus$.

Recipe for cooking up a free Stein identity

Definition
Let $\{P^*_\theta\}_{\theta\geq 0}$ be operators **over measures**, defined by

$$P^*_\theta[\mu] := Law(e^{-\theta}X + \sqrt{1-e^{-2\theta}}Y),$$

with $X \sim \mu$ and $Y \sim m_1[\mu] + \sqrt{Var[\mu]}\mathbf{s}$.

Recipe for cooking up a free Stein identity

Definition
Let $\{P_\theta^*\}_{\theta\geq 0}$ be operators **over measures**, defined by

$$P_\theta^*[\mu] := Law(e^{-\theta}X + \sqrt{1-e^{-2\theta}}Y),$$

with $X \sim \mu$ and $Y \sim m_1[\mu] + \sqrt{Var[\mu]}\mathbf{s}$.

Let $\langle \cdot, \cdot \rangle$ denote dual pairing of measures and functions. Observe that

Recipe for cooking up a free Stein identity

Definition
Let $\{P^*_\theta\}_{\theta \geq 0}$ be operators **over measures**, defined by

$$P^*_\theta[\mu] := Law(e^{-\theta} X + \sqrt{1 - e^{-2\theta}} Y),$$

with $X \sim \mu$ and $Y \sim m_1[\mu] + \sqrt{Var[\mu]}\mathbf{s}$.

Let $\langle \cdot, \cdot \rangle$ denote dual pairing of measures and functions. Observe that

$$\langle \mathbf{s}, h \rangle - \langle \mu, h \rangle = \langle P^*_\infty[\mu], h \rangle - \langle P^*_0[\mu], h \rangle$$

Recipe for cooking up a Stein identity

Deriving and integrating, we get

$$\langle \boldsymbol{s}, h\rangle - \langle \mu, h\rangle = \int_0^\infty \frac{d}{d\theta}\langle P_\theta^*[\mu], h\rangle d\theta.$$

Recipe for cooking up a Stein identity

Deriving and integrating, we get

$$\langle \boldsymbol{s}, h\rangle - \langle \mu, h\rangle = \int_0^\infty \frac{d}{d\theta}\langle P_\theta^*[\mu], h\rangle d\theta.$$

Lemma

$$\frac{d}{d\theta}\langle P_\theta^*[\mu], h\rangle = \langle P_\theta^*[\mu] \otimes P_\theta^*[\mu], \mathcal{L}_\boxplus[h]\rangle,$$

where

$$\mathcal{L}_\boxplus[h](x, y) := xDh(x) - \partial Dh,$$

for D denoting derivative and

$$\partial g(x, y) := (g(x) - g(y))/(x - y).$$

What have we achieved in the free case?

For h regular enough,

$$\langle \boldsymbol{s}, h \rangle - \langle \mu, h \rangle = \int_0^\infty \langle P_\theta^*[\mu] \otimes P_\theta^*[\mu], \mathcal{L}_\boxplus[h] \rangle d\theta.$$

Lemma (Non-commutative Stein's lemma)
A law ν is semicircular if and only if

$$\langle \nu \otimes \nu, \mathcal{L}_\boxplus[h] \rangle = 0,$$

for h smooth with second bounded derivative.

What have we achieved in the free case?

For h regular enough,

$$\langle \boldsymbol{s}, h\rangle - \langle \mu, h\rangle = \int_0^\infty \langle P_\theta^*[\mu] \otimes P_\theta^*[\mu], \mathcal{L}_\boxplus[h]\rangle d\theta.$$

Lemma (Non-commutative Stein's lemma)
A law ν is semicircular if and only if

$$\langle \nu \otimes \nu, \mathcal{L}_\boxplus[h]\rangle = 0,$$

for h smooth with second bounded derivative.

As a consequence,

$$\langle \boldsymbol{s}, h\rangle - \langle \mu, h\rangle = \int_0^\infty \langle P_\theta^*[\mu] \otimes P_\theta^*[\mu] - P_\infty^*[\mu] \otimes P_\infty^*[\mu], \mathcal{L}_\boxplus[h]\rangle d\theta.$$

What have we achieved in the free case? Part II

Writing what we have differently,

$$|\langle \mathbf{s}, h\rangle - \langle \mu, h\rangle| = |\langle \mathcal{S}^*_{\boxplus}[\mu], \mathcal{L}_{\boxplus}[h]\rangle|,$$

where

$$\mathcal{S}^*_{\boxplus}[\mu] := \int_0^\infty (P^*_\theta[\mu] \otimes P^*_\theta[\mu] - P^*_\infty[\mu] \otimes P^*_\infty[\mu]) d\theta.$$

What have we achieved in the free case? Part II

Writing what we have differently,

$$|\langle \mathbf{s}, h\rangle - \langle \mu, h\rangle| = |\langle \mathcal{S}^*_{\boxplus}[\mu], \mathcal{L}_{\boxplus}[h]\rangle|,$$

where

$$\mathcal{S}^*_{\boxplus}[\mu] := \int_0^\infty (P^*_\theta[\mu] \otimes P^*_\theta[\mu] - P^*_\infty[\mu] \otimes P^*_\infty[\mu]) d\theta.$$

Compare this with the classical case:

$$|\langle \gamma, h\rangle - \langle \mu, h\rangle| = |\langle \mu, \mathcal{L}[\mathcal{S}[h]]\rangle|,$$

where γ is the standard Gaussian and $\mathcal{S}$ is defined similarly.

New bottleneck

Dealing with the bottleneck when $\mu = \nu_n$

Let $\xi_{1,n}, \ldots, \xi_{n,n}$ be free random variables with law

$$\int_{\mathbb{R}_+} \left(P_\theta^* - P_\infty^*\right)[\mu_{k,n}]d\theta.$$

Dealing with the bottleneck when $\mu = \nu_n$

Let $\xi_{1,n}, \ldots, \xi_{n,n}$ be free random variables with law

$$\int_{\mathbb{R}_+} \left(P_\theta^* - P_\infty^*\right)[\mu_{k,n}] d\theta.$$

Define $\nu_n = \mu_{1,n} \boxplus \cdots \boxplus \mu_{n,n}$, and assume $Var(\nu_n) = 1$ and that each $\mu_{k,n}$ has small support.

Dealing with the bottleneck when $\mu = \nu_n$

Let $\xi_{1,n}, \ldots, \xi_{n,n}$ be free random variables with law

$$\int_{\mathbb{R}_+} \big(P_\theta^* - P_\infty^*\big)[\mu_{k,n}]d\theta.$$

Define $\nu_n = \mu_{1,n} \boxplus \cdots \boxplus \mu_{n,n}$, and assume $Var(\nu_n) = 1$ and that each $\mu_{k,n}$ has small support. Then

$$\langle \mathcal{S}_\boxplus^*[\nu_n], \mathcal{L}_\boxplus[h] \rangle = \tau[S_n Dh(S_n)] - \langle \mathcal{S}_\boxplus^*[\nu_n], \partial Dh \rangle,$$

where $S_n := \xi_{1,n} + \cdots + \xi_{n,n}$.

An important technicality

Corollary (Bercovici and Voiculescu)
For n large, (ν_n) supported in $[-3, 3]$

An important technicality

Corollary (Bercovici and Voiculescu)
For n large, (ν_n) *supported in* $[-3,3]$

If h is holomorphic

$$\langle \mathcal{S}^*_{\boxplus}[\nu_n], \mathcal{L}_{\boxplus}[h] \rangle = \frac{1}{2\pi i} \int_{\mathcal{R}} h(z)(\tau[S_n g_z(S_n)] - \langle \mathcal{S}^*_{\boxplus}[\nu_n], \partial g_z \rangle) dz,$$

where $\mathcal{R}$ is a contour strictly containing $[-5,5]$ and $g_z(x) := (z-x)^{-2}$.

An important technicality

Corollary (Bercovici and Voiculescu)
For n large, (ν_n) *supported in* $[-3,3]$

If h is holomorphic

$$\langle \mathcal{S}^*_{\boxplus}[\nu_n], \mathcal{L}_{\boxplus}[h] \rangle = \frac{1}{2\pi i} \int_{\mathcal{R}} h(z)(\tau[S_n g_z(S_n)] - \langle \mathcal{S}^*_{\boxplus}[\nu_n], \partial g_z \rangle) dz,$$

where $\mathcal{R}$ is a contour strictly containing $[-5,5]$ and $g_z(x) := (z-x)^{-2}$.

warning: we only need h must be only bounded, not holomorphic.

Moral of the story

For h holomorphic and bounded by one over $[-5, 5]$, there is a constant only depending on $\mathcal{R}$, such that

$$|\langle \mathbf{s}, h\rangle - \langle \nu_n, h\rangle| \leq C \sup_{z\in\mathcal{R}} |\tau[S_n g_z(S_n)] - \langle \mathcal{S}^*_{\boxplus}[\nu_n], \partial g_z\rangle|.$$

By an approximation argument,

$$d_{TV}(\mathbf{s}, \nu_n) \leq C \sup_{z\in\mathcal{R}} |\tau[S_n g_z(S_n)] - \langle \mathcal{S}^*_{\boxplus}[\nu_n], \partial g_z\rangle|.$$

Non-commutative Lindeberg trick?

Define $S_n^{(k)}$ as the part of S_n that does not involve $\xi_{k,n}$. Observe that

$$\tau[S_n g_z(S_n)] = \sum_{k=1}^{n} \tau[\xi_{k,n} g_z(S_n)] = \sum_{k=1}^{n} \tau[\xi_{k,n}(g_z(S_n) - g_z(S_n^{(k)}))]$$

Non-commutative Lindeberg trick?

Define $S_n^{(k)}$ as the part of S_n that does not involve $\xi_{k,n}$. Observe that

$$\tau[S_n g_z(S_n)] = \sum_{k=1}^{n} \tau[\xi_{k,n} g_z(S_n)] = \sum_{k=1}^{n} \tau[\xi_{k,n}(g_z(S_n) - g_z(S_n^{(k)}))]$$

Idea: use a bit of non-commutative Taylor

The only non-commutative calculus we need

Lemma

For non-commutative variables a, r, define

$$\Delta\left(a, r\right) := 2\mathfrak{s}[(z-a)r] - r^2,$$

where $\mathfrak{s}$ denotes the symmetrization operator. Then, for all $q \geq 1$,

$$g\left(a+r\right) = g\left(a\right)\left(\Delta\left(a, r\right) g(a)\right)^q + \sum_{j=0}^{q-1} g\left(a\right)\left(\Delta\left(a, r\right) g\left(a\right)\right)^j,$$

Consequence

For $q = 2$, the boundedness of the g_z then yields

$$\tau[S_n g_z(S_n)] = O(\sum_{k=1}^{n} \tau[|\xi_{k,n}|^3]) \\ + 2\sum_{k=1}^{n} \tau[\xi_{k,n} g_z\left(S_n^{(k)}\right) \mathfrak{s}\left[(z - S_n^{(k)})\xi_{k,n}\right] g_z\left(S_n^{(k)}\right)].$$

Consequence

Using freeness,

$$\tau[S_n g_z(S_n)] = O\left(\sum_{k=1}^{n} \tau[|\xi_{k,n}|^3]\right) + 2\sum_{k=1}^{n} \tau[|\xi_{k,n}|^2]\,\tau[g_z(S_n)(z - S_n)g_z(S_n)]\,.$$

The unit variance condition then implies

$$\sup_{z\in\mathcal{R}} |\tau[S_n g_z(S_n)] - \langle \mathcal{S}_{\boxplus}^*[\nu_n], \partial g_z\rangle| \leq C\sum_{k=1}^{n} \tau[|\xi_{k,n}|^3].$$

Wrapping things up

Theorem (Diaz-Jaramillo)
Under the above considerations,

$$d_{TV}(\mathbf{s}, \nu_n) \leq C \sum_{k=1}^{n} \int_{\mathbb{R}} |x|^3 \mu_{k,n}(dx).$$

Some improvements:

1. Neighborhoods of dependency.
2. Uniform convergence of the density.
3. Including sharper approximations under more conditions.

Some unsolved improvements:

1. Uniform convergence of the derivatives of the density.

Some questions that I thought could be interesting

- Free law of rare events
- For Boolean or monotone convolutions, can we still say something in Wasserstein distance?
- Can we change $\boxplus$ by $\boxplus_m$ and still say something?
- Extended to Edgeworth expansions
- Implementations in large matrix problems
- Multidimensional versions
- Free stable limits

Thanks!

References

Diaz M., Jaramillo A. Non-commutative Stein's method: Applications to free probability and sums of non-commutative variables.

G. P. Chistyakov and F. Gotze. Limit theorems in free probability theory.