



CIMAT

Centro de Investigación en Matemáticas, A.C.

Free Berry-Esseen theorem via Stein's method

Arturo Jaramillo Gil

Centro de Investigación en Matemáticas (CIMAT)

Joint work with a great mathematician, great fan of Checo Pérez, and most of all: a great friend



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

Goal

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

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

Goal

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

Moment of honesty...

Classical Stein's method

Greek letters are probabilities and $\langle \mu, h \rangle := \int_{\mathbb{R}} h(x) \mu(dx)$.

Classical Stein's method

Greek letters are probabilities and $\langle \mu, h \rangle := \int_{\mathbb{R}} h(x) \mu(dx)$.

Lemma (Stein's lemma)

Consider the Stein identity

$$\langle \mu, \iota \cdot Df - D^2 f \rangle = 0, \tag{1}$$

where D denotes derivative and ι the identity function. If (1) holds for many choices of f , then μ is standard Gaussian.

Classical Stein's method

Greek letters are probabilities and $\langle \mu, h \rangle := \int_{\mathbb{R}} h(x) \mu(dx)$.

Lemma (Stein's lemma)

Consider the Stein identity

$$\langle \mu, \iota \cdot Df - D^2 f \rangle = 0, \tag{1}$$

where D denotes derivative and ι the identity function. If (1) holds for many choices of f , then μ is standard Gaussian.

Stein Heuristics

If (1) "more or less holds", then μ is "more or less semicircular".

Classical Stein's method

Greek letters are probabilities and $\langle \mu, h \rangle := \int_{\mathbb{R}} h(x) \mu(dx)$.

Lemma (Stein's lemma)

Consider the Stein identity

$$\langle \mu, \iota \cdot Df - D^2 f \rangle = 0, \quad (1)$$

where D denotes derivative and ι the identity function. If (1) holds for many choices of f , then μ is standard Gaussian.

Stein Heuristics

If (1) "more or less holds", then μ is "more or less semicircular".

Why Stein's identity would be natural?

How to cook up Stein's identity? (personal favorite recipe)

Let γ be the standard Gaussian law and N a r.v. with law γ .

How to cook up Stein's identity? (personal favorite recipe)

Let γ be the standard Gaussian law and N a r.v. with law γ . Consider the semigroup $\{P_\theta\}_{\theta \geq 0}$,

$$P_\theta[h](x) = \mathbb{E}[f(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

How to cook up Stein's identity? (personal favorite recipe)

Let γ be the standard Gaussian law and N a r.v. with law γ . Consider the semigroup $\{P_\theta\}_{\theta \geq 0}$,

$$P_\theta[h](x) = \mathbb{E}[f(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

Then, for a given probability μ , and a test function h ,

$$\langle \gamma, h \rangle - \langle \mu, h \rangle = \langle \mu, P_\infty[h] \rangle - \langle \mu, P_0[h] \rangle.$$

How to cook up Stein's identity? (personal favorite recipe)

Let γ be the standard Gaussian law and N a r.v. with law γ . Consider the semigroup $\{P_\theta\}_{\theta \geq 0}$,

$$P_\theta[h](x) = \mathbb{E}[f(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

Then, for a given probability μ , and a test function h ,

$$\langle \gamma, h \rangle - \langle \mu, h \rangle = \langle \mu, P_\infty[h] \rangle - \langle \mu, P_0[h] \rangle.$$

This operator 'interpolates' towards Gaussian.

How to cook up Stein's identity?

Deriving and integrating, we get

$$\langle \gamma, h \rangle - \langle \mu, h \rangle = \int_0^\infty \frac{d}{d\theta} \langle \mu, P_\theta[h] \rangle d\theta = \langle \mu, \int_0^\infty \mathcal{L}[P_\theta[h]] d\theta \rangle.$$

where

$$\mathcal{L}[g](x) := xDg(x) - D^2g(x) = (\iota \cdot Dg - D^2g)(x).$$

What have we achieved so far?

If h is regular enough,

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

where $\mathcal{S}[h]$ is the operator

$$\mathcal{S}[h] := \int_0^\infty (P_\theta[h] - P_\infty[h]) d\theta.$$

What have we achieved so far?

If h is regular enough,

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

where $\mathcal{S}[h]$ is the operator

$$\mathcal{S}[h] := \int_0^\infty (P_\theta[h] - P_\infty[h]) d\theta.$$

New bottleneck

Bound uniformly $|\langle \mu, \mathcal{L}[g] \rangle|$, with $g = \mathcal{S}[h]$.

Dealing with the bottleneck for getting CLT

Consider the case

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

with $\mu_{j,n}$ centered, such that

$$\sum_{k=1}^n \text{Var}[\mu_{k,n}] = 1.$$

Dealing with the bottleneck for getting CLT

Consider the case

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

with $\mu_{j,n}$ centered, such that

$$\sum_{k=1}^n \text{Var}[\mu_{k,n}] = 1.$$

If $\xi_{1,n}, \dots, \xi_{n,n}$ are independent with $\xi_{k,n} \sim \mu_{k,n}$, then

$$\langle \nu_n, \mathcal{L}[g] \rangle = \mathbb{E}[S_n Dg(S_n)] - \mathbb{E}[D^2 g(S_n)]$$

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

A Lindeberg-type trick

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

A Lindeberg-type trick

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

$$\mathbb{E}[S_n Dg(S_n)] = \sum_{k=1}^n \mathbb{E}[\xi_{k,n} Dg(S_n)] = \sum_{k=1}^n \mathbb{E}[\xi_{k,n} (Dg(S_n) - Dg(S_n^{(k)}))]$$

A Lindeberg-type trick

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

$$\mathbb{E}[S_n Dg(S_n)] = \sum_{k=1}^n \mathbb{E}[\xi_{k,n} Dg(S_n)] = \sum_{k=1}^n \mathbb{E}[\xi_{k,n} (Dg(S_n) - Dg(S_n^k))]$$

Then,

$$\begin{aligned} \mathbb{E}[S_n Dg(S_n)] &\approx \sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^2 D^2 g(S_n^k)] = \sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^2] \mathbb{E}[D^2 g(S_n^k)] \\ &\approx \sum_{k=1}^n \mathbb{E}[|\xi_{k,n}|^2] \mathbb{E}[D^2 g(S_n)] = \mathbb{E}[D^2 g(S_n)] \end{aligned}$$

Conclusion: $\langle \nu_n, \mathcal{L}[g] \rangle \approx 0$

Elements of free probability

Let \mathcal{A} be a unital C^* -algebra and $\tau : \mathcal{A} \rightarrow \mathbb{C}$ a positive unital linear functional. We then say that the pair (\mathcal{A}, τ) is a C^* -*probability space*.

Elements of free probability

Let \mathcal{A} be a unital C^* -algebra and $\tau : \mathcal{A} \rightarrow \mathbb{C}$ a positive unital linear functional. We then say that the pair (\mathcal{A}, τ) is a C^* -probability space.

Definition (Freeness)

Let $\{A_n\}_{n \geq 1}$ be a sequence of subalgebras of \mathcal{A} . For $a \in \mathcal{A}$, denote the centering of a by $\bar{a} := a - \tau[a]$. We say that $\{A_n\}_{n \geq 1}$ are freely independent, or free, if

$$\tau[\bar{a}_1 \bar{a}_2 \cdots \bar{a}_k] = 0, \tag{2}$$

for a_1, \dots, a_k alternating algebras.

Elements of free probability

Let \mathcal{A} be a unital C^* -algebra and $\tau : \mathcal{A} \rightarrow \mathbb{C}$ a positive unital linear functional. We then say that the pair (\mathcal{A}, τ) is a C^* -probability space.

Definition (Freeness)

Let $\{A_n\}_{n \geq 1}$ be a sequence of subalgebras of \mathcal{A} . For $a \in \mathcal{A}$, denote the centering of a by $\bar{a} := a - \tau[a]$. We say that $\{A_n\}_{n \geq 1}$ are freely independent, or free, if

$$\tau[\bar{a}_1 \bar{a}_2 \cdots \bar{a}_k] = 0, \tag{2}$$

for a_1, \dots, a_k alternating algebras.

Definition (Free convolution)

Sums of free random variables yields the free convolution \boxplus .

Recipe for cooking up a free Stein identity

Our building block before was

$$P_{\theta}[h](x) = \mathbb{E}[h(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

Recipe for cooking up a free Stein identity

Our building block before was

$$P_\theta[h](x) = \mathbb{E}[h(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

now what we do?

Recipe for cooking up a free Stein identity

Our building block before was

$$P_\theta[h](x) = \mathbb{E}[h(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

now what we do?

Definition

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

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

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

Recipe for cooking up a free Stein identity

Our building block before was

$$P_\theta[h](x) = \mathbb{E}[h(e^{-\theta}x + \sqrt{1 - e^{-2\theta}}N)].$$

now what we do?

Definition

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

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

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

This operator 'interpolates' from μ standardized to s , as

$$\langle s, h \rangle - \langle \mu, h \rangle = \langle P_\infty^*[\mu], h \rangle - \langle P_0^*[\mu], h \rangle$$

Recipe for cooking up a Stein identity

Deriving and integrating, we get

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

Recipe for cooking up a Stein identity

Deriving and integrating, we get

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

Lemma (Diaz-Jaramillo)

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

where $\mathcal{L}_\boxplus[h]$ is the real function in \mathbb{R}^2 :

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

where g is the non-commutative derivative

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

What have we achieved in the free case?

For h regular enough,

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

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,

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

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,

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

What have we achieved in the free case? Pt 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? Pt 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, stating

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

New bottleneck

Bound uniformly $|\langle \rho, \mathcal{L}_{\boxplus}[g] \rangle|$, with $\rho = \mathcal{S}_{\boxplus}^*[\mu]$.

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

Consider

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

Assume $\text{Var}[\nu_n] = 1$ and that $\mu_{k,n}$ have **small supports**.

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

Consider

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

Assume $\text{Var}[\nu_n] = 1$ and that $\mu_{k,n}$ have **small supports**. If $\xi_{1,n}, \dots, \xi_{n,n}$ are free with law equal to $\int_{\mathbb{R}_+} (P_\theta^* - P_\infty^*)[\mu_{k,n}] d\theta$, then

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

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

Corollary (Superconvergence by Bercovici and Voiculescu)

For n large, $\text{Supp}(\nu_n) \subset [-3, 3]$ and $\text{Supp}(\mathcal{S}_{\boxplus}^[\nu_n]) \subset [-5, 5]$.*

The Cauchy formula yields

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

where \mathcal{R} strictly containing $[-5, 5]$ and

$$g(x) := (z - x)^{-2}.$$

Corollary (Superconvergence by Bercovici and Voiculescu)

For n large, $\text{Supp}(\nu_n) \subset [-3, 3]$ and $\text{Supp}(\mathcal{S}_{\boxplus}^[\nu_n]) \subset [-5, 5]$.*

The Cauchy formula yields

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

where \mathcal{R} strictly containing $[-5, 5]$ and

$$g(x) := (z - x)^{-2}.$$

For total variation, h is bounded by one!

For h very smooth and bounded by one, 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,

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

Non-commutative Lindeberg trick?

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

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

Non-commutative Lindeberg trick?

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

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

Do I need a commutative Taylor?

The only non-commutative calculus we need

Lemma

For non-commutative variables a, r , define

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

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

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

Consequence

We have the following expansion for all $q \geq 1$

$$\begin{aligned}\tau[S_n g_z(S_n)] &= \sum_{k=1}^n \tau[\xi_{k,n} g(S_n) (\Delta(S_n^k, \xi_{k,n}) g(S_n^k))^q] \\ &\quad + \sum_{j=1}^{q-1} \tau[g(S_n^k) (\Delta(S_n^k, \xi_{k,n}) g(S_n^k))^j].\end{aligned}$$

Consequence

We have the following expansion for all $q \geq 1$

$$\begin{aligned}\tau[S_n g_z(S_n)] &= \sum_{k=1}^n \tau[\xi_{k,n} g(S_n) (\Delta(S_n^k, \xi_{k,n}) g(S_n^k))^q] \\ &\quad + \sum_{j=1}^{q-1} \tau[g(S_n^k) (\Delta(S_n^k, \xi_{k,n}) g(S_n^k))^j].\end{aligned}$$

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

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

Finally, using freeness,

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

The unit variance condition then implies

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

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. The Mauricio Salazar trick holds.

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 Mauricio Salazar trick or Edgeworth expansions
- Implementations in large matrix problems
- Multidimensional versions
- Free stable limits

Thanks!



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.