Stein's Method: The Continuous Case

Elizabeth Meckes

American Institute of Mathematics
and
Case Western Reserve University

February 19, 2008

A Motivating Example

A Motivating Example

Let X be a uniformly distributed point of the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$.

A Motivating Example

Let X be a uniformly distributed point of the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$. It has long been known that, say, X_1 is approximately distributed as a standard Gaussian, as $n \to \infty$.

A Motivating Example

Let X be a uniformly distributed point of the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$. It has long been known that, say, X_1 is approximately distributed as a standard Gaussian, as $n \to \infty$. In 1987, Diaconis and Freedman found a rate of convergence in this theorem.

A Motivating Example

Let X be a uniformly distributed point of the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$. It has long been known that, say, X_1 is approximately distributed as a standard Gaussian, as $n \to \infty$. In 1987, Diaconis and Freedman found a rate of convergence in this theorem. Specifically, they showed

Theorem (Diaconis-Freedman)

Let X be a randomly chosen point of $\sqrt{n}\mathbb{S}^{n-1}$, and let X_1 be its first coordinate. If Z is a standard Gaussian random variable, then

$$d_{TV}(X_1, Z) \leq \frac{8}{n-4}.$$

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

A couple of straightforward options:

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

A couple of straightforward options:

- Pick a coordinate at random and change its sign.

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

A couple of straightforward options:

- Pick a coordinate at random and change its sign.
- Pick two coordinates at random and switch them.

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

A couple of straightforward options:

- Pick a coordinate at random and change its sign.
- Pick two coordinates at random and switch them.

In fact, any symmetry of the sphere will similarly result in a possible exchangeable pair.

Example (cont.)

One could try to prove this theorem with Stein's method of exchangeable pairs.

Question: How to construct the exchangeable pair?

A couple of straightforward options:

- Pick a coordinate at random and change its sign.
- Pick two coordinates at random and switch them.

In fact, any symmetry of the sphere will similarly result in a possible exchangeable pair.

Problem: The possibilities above throw away a tremendous amount of information. They both would also work for the cube, for example, which has far fewer symmetries than the sphere. In fact, since they work for the cube as well, it is unrealistic to expect a convergence rate better than $O\left(\frac{1}{\sqrt{n}}\right)$, whereas the correct rate (Diaconis-Freedman) is $O\left(\frac{1}{n}\right)$.

Infinitesimal rotations

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$. Then the matrix $UA_\epsilon U^T$ is a rotation by $\arcsin(\epsilon)$ in a random two-dimensional subspace of $\mathbb{R}^n$.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$. Then the matrix $UA_\epsilon U^T$ is a rotation by $\arcsin(\epsilon)$ in a random two-dimensional subspace of $\mathbb{R}^n$.

Make an exchangeable pair of random points on the sphere of radius $\sqrt{n}$ as follows.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$. Then the matrix $UA_\epsilon U^T$ is a rotation by $\arcsin(\epsilon)$ in a random two-dimensional subspace of $\mathbb{R}^n$.

Make an exchangeable pair of random points on the sphere of radius $\sqrt{n}$ as follows. Choose X according to surface measure on $\sqrt{n}\mathbb{S}^{n-1}$,

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.

Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$. Then the matrix $UA_\epsilon U^T$ is a rotation by $\arcsin(\epsilon)$ in a random two-dimensional subspace of $\mathbb{R}^n$.

Make an exchangeable pair of random points on the sphere of radius $\sqrt{n}$ as follows. Choose X according to surface measure on $\sqrt{n}\mathbb{S}^{n-1}$, and let $X_\epsilon = UA_\epsilon U^T X$.

Infinitesimal rotations

Idea: Try to make use of the spherical symmetry in constructing the exchangeable pair.

Fix $\epsilon > 0$, and let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$.
Let U be a random orthogonal matrix, distributed according to Haar measure on $\mathcal{O}_n$. Then the matrix $UA_\epsilon U^T$ is a rotation by $\arcsin(\epsilon)$ in a random two-dimensional subspace of $\mathbb{R}^n$.

Make an exchangeable pair of random points on the sphere of radius $\sqrt{n}$ as follows. Choose X according to surface measure on $\sqrt{n}\mathbb{S}^{n-1}$, and let $X_\epsilon = UA_\epsilon U^T X$. In fact, this gives an entire family of exchangeable pairs (X, X_ϵ), parametrized by ϵ.

Infinitesimal rotations (cont.)

It turns out that when using this exchangeable pair, one can take the limit as $\epsilon \to 0$.

Infinitesimal rotations (cont.)

It turns out that when using this exchangeable pair, one can take the limit as $\epsilon \to 0$. Taking this approach yields:

Theorem

Let X be distributed uniformly on $\sqrt{n}\mathbb{S}^{n-1}$, and let X_1 denote its first coordinate. If Z is a standard normal random variable, then

$$d_{TV}(X_1, Z) \leq \frac{2\sqrt{3}}{n-1}.$$

Infinitesimal rotations (cont.)

It turns out that when using this exchangeable pair, one can take the limit as $\epsilon \to 0$. Taking this approach yields:

Theorem

Let X be distributed uniformly on $\sqrt{n}\mathbb{S}^{n-1}$, and let X_1 denote its first coordinate. If Z is a standard normal random variable, then

$$d_{TV}(X_1, Z) \leq \frac{2\sqrt{3}}{n-1}.$$

Note that the bound is in total variation distance.

Infinitesimal rotations (cont.)

It turns out that when using this exchangeable pair, one can take the limit as $\epsilon \to 0$. Taking this approach yields:

Theorem

Let X be distributed uniformly on $\sqrt{n}\mathbb{S}^{n-1}$, and let X_1 denote its first coordinate. If Z is a standard normal random variable, then

$$d_{TV}(X_1, Z) \leq \frac{2\sqrt{3}}{n-1}.$$

Note that the bound is in total variation distance.
This recovers the result of Diaconis-Freedman, which is sharp (up to the constant).

The real motivating example

The real motivating example

In the early 1990's, Persi Diaconis conjectured that for k fixed and $M \in \mathcal{O}_n$ a random (Haar-distributed) orthogonal matrix, the random variable $\text{Tr}(M^k)$ coverged exponentially, or even super-exponentially, fast to Gaussian as $n \to \infty$.

The real motivating example

In the early 1990's, Persi Diaconis conjectured that for k fixed and $M \in \mathcal{O}_n$ a random (Haar-distributed) orthogonal matrix, the random variable $\mathrm{Tr}(M^k)$ coverged exponentially, or even super-exponentially, fast to Gaussian as $n \to \infty$.

Charles Stein proved that the rate of convergence was better than $O(n^{-r})$ for any r, using the approach described in the example on the previous slides.

The real motivating example

In the early 1990's, Persi Diaconis conjectured that for k fixed and $M \in \mathcal{O}_n$ a random (Haar-distributed) orthogonal matrix, the random variable $\mathrm{Tr}(M^k)$ coverged exponentially, or even super-exponentially, fast to Gaussian as $n \to \infty$.

Charles Stein proved that the rate of convergence was better than $O(n^{-r})$ for any r, using the approach described in the example on the previous slides.

Shortly thereafter, Kurt Johansson proved that the rate of convergence is $O(e^{-cn})$ for some $c > 0$. He also proved an analogous result for the symplectic group, and a slightly better result ($O(n^{-\delta n})$) for the unitary group.

An abstract normal approximation theorem

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) *is a family of exchangeable pairs defined on a common probability space, such that* $\mathbb{E}W = 0$ *and* $\mathbb{E}W^2 = \sigma^2$.

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) *is a family of exchangeable pairs defined on a common probability space, such that* $\mathbb{E}W = 0$ *and* $\mathbb{E}W^2 = \sigma^2$. *Suppose there is a function* $\lambda(\epsilon)$ *and random variables* E, E' *such that*

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) *is a family of exchangeable pairs defined on a common probability space, such that* $\mathbb{E}W = 0$ *and* $\mathbb{E}W^2 = \sigma^2$. *Suppose there is a function* $\lambda(\epsilon)$ *and random variables* E, E' *such that*

1. $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} -W + E'$.

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) *is a family of exchangeable pairs defined on a common probability space, such that* $\mathbb{E}W = 0$ *and* $\mathbb{E}W^2 = \sigma^2$*. Suppose there is a function* $\lambda(\epsilon)$ *and random variables* E, E' *such that*

1. $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} -W + E'$.
2. $\frac{1}{2\lambda(\epsilon)\sigma^2}\mathbb{E}\left[(W_\epsilon - W)^2 \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} 1 + E$.

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) *is a family of exchangeable pairs defined on a common probability space, such that* $\mathbb{E}W = 0$ *and* $\mathbb{E}W^2 = \sigma^2$. *Suppose there is a function* $\lambda(\epsilon)$ *and random variables* E, E' *such that*

1. $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W \middle| W\right] \xrightarrow[\epsilon \to 0]{L_1} -W + E'$.
2. $\frac{1}{2\lambda(\epsilon)\sigma^2}\mathbb{E}\left[(W_\epsilon - W)^2 \middle| W\right] \xrightarrow[\epsilon \to 0]{L_1} 1 + E$.
3. $\frac{1}{\lambda(\epsilon)}\mathbb{E}|W_\epsilon - W|^3 \xrightarrow{\epsilon \to 0} 0$.

An abstract normal approximation theorem

The following theorem is an abstraction of the approach taken by Stein in analyzing $\mathrm{Tr}(M^k)$.

Theorem (M)

Suppose that (W, W_ϵ) is a family of exchangeable pairs defined on a common probability space, such that $\mathbb{E}W = 0$ and $\mathbb{E}W^2 = \sigma^2$. Suppose there is a function $\lambda(\epsilon)$ and random variables E, E' such that

1. $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} -W + E'$.
2. $\frac{1}{2\lambda(\epsilon)\sigma^2}\mathbb{E}\left[(W_\epsilon - W)^2 \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} 1 + E$.
3. $\frac{1}{\lambda(\epsilon)}\mathbb{E}|W_\epsilon - W|^3 \xrightarrow{\epsilon\to 0} 0$.

Then if Z is a standard normal random variable,

$$d_{TV}(W, Z) \leq \mathbb{E}|E| + \sqrt{\frac{\pi}{2}}\mathbb{E}|E'|.$$

Example 1

Linear functions on the classical compact matrix groups

Example 1

Linear functions on the classical compact matrix groups

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let A be a fixed $n \times n$ matrix over $\mathbb{R}$, with $\mathrm{Tr}(AA^T) = n$.

Example 1

Linear functions on the classical compact matrix groups

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let A be a fixed $n \times n$ matrix over $\mathbb{R}$, with $\mathrm{Tr}(AA^T) = n$. Define $W := \mathrm{Tr}(AM)$. W is a general linear combination of the entries of M.

Example 1

Linear functions on the classical compact matrix groups

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let A be a fixed $n \times n$ matrix over $\mathbb{R}$, with $\mathrm{Tr}(AA^T) = n$. Define $W := \mathrm{Tr}(AM)$. W is a general linear combination of the entries of M.

It was shown by d'Aristotile, Diaconis, and Newman (2003) that the distribution of W is asymptotically Gaussian, as $n \to \infty$.

Example 1

Linear functions on the classical compact matrix groups

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let A be a fixed $n \times n$ matrix over $\mathbb{R}$, with $\mathrm{Tr}(AA^T) = n$. Define $W := \mathrm{Tr}(AM)$. W is a general linear combination of the entries of M.

It was shown by d'Aristotile, Diaconis, and Newman (2003) that the distribution of W is asymptotically Gaussian, as $n \to \infty$.

One way to interpret this result is as a comparison between random orthogonal matrices and Gaussian matrices (since if M is a Gaussian random matrix, $\mathrm{Tr}(AM)$ is exactly Gaussian).

Example 1

Linear functions on the classical compact matrix groups

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let A be a fixed $n \times n$ matrix over $\mathbb{R}$, with $\mathrm{Tr}(AA^T) = n$. Define $W := \mathrm{Tr}(AM)$. W is a general linear combination of the entries of M.

It was shown by d'Aristotile, Diaconis, and Newman (2003) that the distribution of W is asymptotically Gaussian, as $n \to \infty$.

One way to interpret this result is as a comparison between random orthogonal matrices and Gaussian matrices (since if M is a Gaussian random matrix, $\mathrm{Tr}(AM)$ is exactly Gaussian). Other notable such comparisons are due to Tiefeng Jiang, who showed that truncations of random orthogonal matrices are close, in various senses, to Gaussian matrices of the same size.

The following refinement of the result of d'Aristotile, Diaconis, and Newman can be proved using the abstract normal approximation theorem.

The following refinement of the result of d'Aristotile, Diaconis, and Newman can be proved using the abstract normal approximation theorem.

Theorem (M)

Let W be as above. Then if Z is a standard normal random variable,

$$d_{TV}(W, Z) \leq \frac{2\sqrt{3}}{n-1}.$$

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere.

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M)

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M), and define M_ϵ and W_ϵ by

$$M_\epsilon = UA_\epsilon U^T M \qquad\qquad W_\epsilon = \mathrm{Tr}(AM_\epsilon).$$

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M), and define M_ϵ and W_ϵ by

$$M_\epsilon = UA_\epsilon U^T M \qquad\qquad W_\epsilon = \mathrm{Tr}(AM_\epsilon).$$

- In this example, $\lambda(\epsilon) = \frac{\epsilon^2}{n}$ and $E = \frac{1}{n-1}\left[1 - \mathrm{Tr}((AM)^2)\right]$, and $E' = 0$.

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M), and define M_ϵ and W_ϵ by

$$M_\epsilon = UA_\epsilon U^T M \qquad\qquad W_\epsilon = \mathrm{Tr}(AM_\epsilon).$$

- In this example, $\lambda(\epsilon) = \frac{\epsilon^2}{n}$ and $E = \frac{1}{n-1}\left[1 - \mathrm{Tr}((AM)^2)\right]$, and $E' = 0$.
- The next natural question is: What about the unitary group?

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M), and define M_ϵ and W_ϵ by

$$M_\epsilon = UA_\epsilon U^T M \qquad\qquad W_\epsilon = \mathrm{Tr}(AM_\epsilon).$$

- In this example, $\lambda(\epsilon) = \frac{\epsilon^2}{n}$ and $E = \frac{1}{n-1}\left[1 - \mathrm{Tr}((AM)^2)\right]$, and $E' = 0$.
- The next natural question is: What about the unitary group? The analogous theorem would be the convergence of $W := \mathrm{Tr}(AM)$ for $M \in \mathcal{U}_n$ and A a fixed matrix over $\mathbb{C}$ to standard complex Gaussian.

Some comments

- The exchangeable pair is constructed exactly as in the case of the sphere. Let $A_\epsilon = \begin{bmatrix} \sqrt{1-\epsilon^2} & \epsilon \\ -\epsilon & \sqrt{1-\epsilon^2} \end{bmatrix} \oplus \mathrm{I}_{n-2}$, let U be a random orthogonal matrix (independent of M), and define M_ϵ and W_ϵ by

$$M_\epsilon = UA_\epsilon U^T M \qquad\qquad W_\epsilon = \mathrm{Tr}(AM_\epsilon).$$

- In this example, $\lambda(\epsilon) = \frac{\epsilon^2}{n}$ and $E = \frac{1}{n-1}\left[1 - \mathrm{Tr}((AM)^2)\right]$, and $E' = 0$.
- The next natural question is: What about the unitary group? The analogous theorem would be the convergence of $W := \mathrm{Tr}(AM)$ for $M \in \mathcal{U}_n$ and A a fixed matrix over $\mathbb{C}$ to standard complex Gaussian. This is a multivariate question and will be postponed for now.

Example 2

Eigenfunctions of the Laplacian.

Example 2

Eigenfunctions of the Laplacian.

If one were to try to generalize the previous result, there are many possible directions.

Example 2
Eigenfunctions of the Laplacian.

If one were to try to generalize the previous result, there are many possible directions.

One such direction is the following:

Example 2

Eigenfunctions of the Laplacian.

If one were to try to generalize the previous result, there are many possible directions.

One such direction is the following:
The orthogonal group $\mathcal{O}_n$ is, among other things, a manifold.
Can one do anything similar to the analysis above for more general manifolds?

Example 2

Eigenfunctions of the Laplacian.

If one were to try to generalize the previous result, there are many possible directions.

One such direction is the following:
The orthogonal group $\mathcal{O}_n$ is, among other things, a manifold. Can one do anything similar to the analysis above for more general manifolds?

We will see that the appropriate analog of linear functions for manifolds without a linear structure is the collection of eigenfunctions of the Laplacian.

The Laplacian

Let (M, g) be a compact Riemannian manifold. Without loss of generality, assume that

$$\int_M 1 \, d\,vol = 1.$$

The Laplacian

Let (M, g) be a compact Riemannian manifold. Without loss of generality, assume that

$$\int_M 1 \, d\, vol = 1.$$

For coordinates $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1}^{n}$ on M, define

$$(G(x))_{ij} = g_{ij}(x) = \left\langle \left.\frac{\partial}{\partial x_i}\right|_x, \left.\frac{\partial}{\partial x_j}\right|_x \right\rangle$$

$$g(x) = \det(G(x)) \qquad g^{ij}(x) = (G^{-1}(x))_{ij}.$$

The Laplacian

Let (M, g) be a compact Riemannian manifold. Without loss of generality, assume that

$$\int_M 1 \, d\,vol = 1.$$

For coordinates $\left\{ \frac{\partial}{\partial x_i} \right\}_{i=1}^{n}$ on M, define

$$(G(x))_{ij} = g_{ij}(x) = \left\langle \left. \frac{\partial}{\partial x_i} \right|_x , \left. \frac{\partial}{\partial x_j} \right|_x \right\rangle$$

$$g(x) = \det(G(x)) \qquad g^{ij}(x) = (G^{-1}(x))_{ij}.$$

The Laplacian Δ is defined on real-valued functions on M by

$$\Delta f(x) = \frac{1}{\sqrt{g}} \sum_{j,k} \frac{\partial}{\partial x_j} \left(\sqrt{g} g^{jk} \frac{\partial f}{\partial x_k} \right).$$

Eigenfunctions

A function f on M is an eigenfunction of the Laplacian if

$$\Delta f = \lambda f$$

for some $\lambda \in \mathbb{R}$.

Eigenfunctions

A function f on M is an eigenfunction of the Laplacian if

$$\Delta f = \lambda f$$

for some $\lambda \in \mathbb{R}$.

Some facts:

- The set of eigenvalues λ is of the form

$$0 \geq \lambda_1 \geq \lambda_2 \downarrow -\infty,$$

  with each associated eigenspace finite-dimensional.

Eigenfunctions

A function f on M is an eigenfunction of the Laplacian if

$$\Delta f = \lambda f$$

for some $\lambda \in \mathbb{R}$.

Some facts:

- The set of eigenvalues λ is of the form

$$0 \geq \lambda_1 \geq \lambda_2 \downarrow -\infty,$$

with each associated eigenspace finite-dimensional.

- Eigenspaces for distinct eigenvalues are orthogonal in $L_2(M)$ and $L_2(M)$ is the direct sum of the eigenspaces.

Eigenfunctions

A function f on M is an eigenfunction of the Laplacian if

$$\Delta f = \lambda f$$

for some $\lambda \in \mathbb{R}$.

Some facts:

- The set of eigenvalues λ is of the form

  $$0 \geq \lambda_1 \geq \lambda_2 \downarrow -\infty,$$

  with each associated eigenspace finite-dimensional.
- Eigenspaces for distinct eigenvalues are orthogonal in $L_2(M)$ and $L_2(M)$ is the direct sum of the eigenspaces.
- Each eigenfunction f is C^∞ on M.

Main Result

Let $\{f_i\}_{i=1}^k$ be a sequence of functions on M with $\Delta f_i = -\mu_i f_i$ and $\int_M f_i f_j dvol = \delta_{ij}$.

Main Result

Let $\{f_i\}_{i=1}^{k}$ be a sequence of functions on M with $\Delta f_i = -\mu_i f_i$ and $\int_M f_i f_j dvol = \delta_{ij}$. Define the function $f : M \to \mathbb{R}$ by

$$f(x) = \sum_{i=1}^{k} a_i f_i(x),$$

where the a_i are scalars such that $\sum a_i^2 = 1$. Let X be a random point of M.

Main Result

Let $\{f_i\}_{i=1}^k$ be a sequence of functions on M with $\Delta f_i = -\mu_i f_i$ and $\int_M f_i f_j dvol = \delta_{ij}$. Define the function $f : M \to \mathbb{R}$ by

$$f(x) = \sum_{i=1}^{k} a_i f_i(x),$$

where the a_i are scalars such that $\sum a_i^2 = 1$. Let X be a random point of M. Then for $\mu > 0$,

$$d_{TV}(f(X), Z) \leq \frac{2}{\mu} \left[\mathbb{E} \left| \|\nabla f(X)\|^2 - \mathbb{E}\|\nabla f(X)\|^2 \right| + \left(1 + \frac{\sqrt{\pi}}{2\sqrt{2}}\right) \sqrt{\sum_i a_i^2 (\mu_i - \mu)^2} \right],$$

where Z is a standard Gaussian random variable on $\mathbb{R}$.

Main Result

Let $\{f_i\}_{i=1}^k$ be a sequence of functions on M with $\Delta f_i = -\mu_i f_i$ and $\int_M f_i f_j dvol = \delta_{ij}$. Define the function $f : M \to \mathbb{R}$ by

$$f(x) = \sum_{i=1}^{k} a_i f_i(x),$$

where the a_i are scalars such that $\sum a_i^2 = 1$. Let X be a random point of M. Then for $\mu > 0$,

$$d_{TV}(f(X), Z) \leq \frac{2}{\mu} \left[\mathbb{E} \left| \|\nabla f(X)\|^2 - \mathbb{E}\|\nabla f(X)\|^2 \right| \right.$$

where Z is a standard Gaussian random variable on $\mathbb{R}$.

Main Result

Let $\{f_i\}_{i=1}^k$ be a sequence of functions on M with $\Delta f_i = -\mu_i f_i$ and $\int_M f_i f_j dvol = \delta_{ij}$. Define the function $f : M \to \mathbb{R}$ by

$$f(x) = \sum_{i=1}^{k} a_i f_i(x),$$

where the a_i are scalars such that $\sum a_i^2 = 1$. Let X be a random point of M. Then for $\mu > 0$,

$$d_{TV}(f(X), Z) \leq \frac{2}{\mu}\left[\mathbb{E}\left|\|\nabla f(X)\|^2 - \mathbb{E}\|\nabla f(X)\|^2\right| + \left(1 + \frac{\sqrt{\pi}}{2\sqrt{2}}\right)\sqrt{\sum_i a_i^2(\mu_i - \mu)^2}\right],$$

where Z is a standard Gaussian random variable on $\mathbb{R}$.

Idea of proof

Idea of proof

Make an exchangeable pair of random points (X, X_ϵ) on M as follows.

Idea of proof

Make an exchangeable pair of random points (X, X_ϵ) on M as follows.

1. Let X be distributed according to normalized volume measure on M.

Idea of proof

Make an exchangeable pair of random points (X, X_ϵ) on M as follows.

1. Let X be distributed according to normalized volume measure on M.
2. Pick V, independent of X, according to normalized surface area measure on $\mathbb{S}^{n-1}$, where $n = dim(M)$.

Idea of proof

Make an exchangeable pair of random points (X, X_ϵ) on M as follows.

1. Let X be distributed according to normalized volume measure on M.
2. Pick V, independent of X, according to normalized surface area measure on $\mathbb{S}^{n-1}$, where $n = dim(M)$.
3. Let $X_\epsilon = \exp_X(\epsilon V)$.

Idea of proof

Make an exchangeable pair of random points (X, X_ϵ) on M as follows.

1. Let X be distributed according to normalized volume measure on M.
2. Pick V, independent of X, according to normalized surface area measure on $\mathbb{S}^{n-1}$, where $n = dim(M)$.
3. Let $X_\epsilon = \exp_X(\epsilon V)$.
4. For a given eigenfunction f, apply the infinitesimal version of Stein's theorem to $W = f(X)$ and $W_\epsilon = f(X_\epsilon)$.

Why eigenfunctions?

Why eigenfunctions?

Recall condition 1 of the theorem:

$$\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W \middle| W\right] \xrightarrow[\epsilon\to 0]{L_1} -W + E'.$$

Why eigenfunctions?

Recall condition 1 of the theorem:

$$\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[W_\epsilon - W\middle|W\right] \xrightarrow[\epsilon\to 0]{L_1} -W + E'.$$

Fact:

For $f : M \to \mathbb{R}$ smooth and $n = dim(M)$,

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\mathbb{E}\left[f(X_\epsilon) - f(X)\middle|X\right] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\int_{S_x(\epsilon)\subseteq M}\left[f(y) - f(X)\right]d\overline{y}\\ &= \frac{1}{2n}\Delta f(X).\end{aligned}$$

Application 1: The Sphere

Application 1: The Sphere

Some facts:

Application 1: The Sphere

Some facts:

- Eigenfunctions of Δ on $\mathbb{S}^{n-1}$ are the restrictions of homogeneous harmonic polynomials on $\mathbb{R}^n$ to $\mathbb{S}^{n-1}$.

Application 1: The Sphere

Some facts:

- Eigenfunctions of Δ on $\mathbb{S}^{n-1}$ are the restrictions of homogeneous harmonic polynomials on $\mathbb{R}^n$ to $\mathbb{S}^{n-1}$.
- For such a polynomial of degree ℓ, the corresponding eigenvalue is $-\ell(\ell+n-2)$.

Application 1: The Sphere

Some facts:

- Eigenfunctions of Δ on $\mathbb{S}^{n-1}$ are the restrictions of homogeneous harmonic polynomials on $\mathbb{R}^n$ to $\mathbb{S}^{n-1}$.
- For such a polynomial of degree ℓ, the corresponding eigenvalue is $-\ell(\ell + n - 2)$.
- Eigenfunctions on $\mathbb{S}^{n-1}$ can be expressed in terms of special functions; in particular, the harmonic projection (in L_2) of $g(x) = x_n^\ell$ is the Gegenbauer polynomial $C_\ell^{\frac{n-2}{2}}(x_n)$, where

$$C_\ell^k(t) = \frac{2^\ell}{\Gamma(k)} \sum_{j=0}^{\lfloor \ell/2 \rfloor} \frac{(-1)^j \Gamma(k + \ell - j)}{2^{2j} j! (\ell - 2j)!} t^{\ell - 2j}.$$

Let ℓ be odd and define $p_\ell^{(k)}(x) := AC_\ell^{\frac{n-2}{2}}(x_k)$ for $1 \leq k \leq n$, where A is chosen so that $\mathbb{E}(p_\ell^{(k)}(X))^2 = 1$.

Let ℓ be odd and define $p_\ell^{(k)}(x) := AC_\ell^{\frac{n-2}{2}}(x_k)$ for $1 \leq k \leq n$, where A is chosen so that $\mathbb{E}(p_\ell^{(k)}(X))^2 = 1$. Define

$$p(x) := \sum_{k=1}^{n} a_k p_\ell^{(k)}(x) \qquad (\mathit{with} \sum_{i=1}^{n} a_i^2 = 1).$$

Let ℓ be odd and define $p_\ell^{(k)}(x) := AC_\ell^{\frac{n-2}{2}}(x_k)$ for $1 \leq k \leq n$, where A is chosen so that $\mathbb{E}(p_\ell^{(k)}(X))^2 = 1$. Define

$$p(x) := \sum_{k=1}^{n} a_k p_\ell^{(k)}(x) \qquad (\textit{with} \sum_{i=1}^{n} a_i^2 = 1).$$

Theorem

There is a constant c, depending on ℓ, such that for $p : \mathbb{S}^{n-1} \to \mathbb{R}$ as above and X a random point of $\mathbb{S}^{n-1}$,

$$d_{TV}(p(X), Z) \leq c\|\mathbf{a}\|_4^2.$$

Let ℓ be odd and define $p_\ell^{(k)}(x) := AC_\ell^{\frac{n-2}{2}}(x_k)$ for $1 \leq k \leq n$, where A is chosen so that $\mathbb{E}(p_\ell^{(k)}(X))^2 = 1$. Define

$$p(x) := \sum_{k=1}^{n} a_k p_\ell^{(k)}(x) \qquad \left(\textit{with} \sum_{i=1}^{n} a_i^2 = 1\right).$$

Theorem

There is a constant c, depending on ℓ, such that for $p : \mathbb{S}^{n-1} \to \mathbb{R}$ as above and X a random point of $\mathbb{S}^{n-1}$,

$$d_{TV}(p(X), Z) \leq c\|\mathbf{a}\|_4^2.$$

If the vector of coefficients $\{a_i\}$ is chosen to be random and uniformly distribtuted on the sphere, then there is another constant c' such that

$$\mathbb{E} d_{TV}(p(X), Z) \leq \frac{c'}{\sqrt{n}}.$$

Application 2: The Torus

Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ with metric given by B (symmetric, positive-definite):

$$(x, y)_B = \langle Bx, y \rangle .$$

Application 2: The Torus

Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ with metric given by B (symmetric, positive-definite):

$$(x, y)_B = \langle Bx, y\rangle \, .$$

The Laplacian Δ_B on $\mathbb{T}^n$ is

$$\Delta_B f(x) = \sum_{j,k} (B^{-1})_{jk} \frac{\partial^2 f}{\partial x_j \partial x_k}(x).$$

Application 2: The Torus

Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ with metric given by B (symmetric, positive-definite):

$$(x, y)_B = \langle Bx, y \rangle .$$

The Laplacian Δ_B on $\mathbb{T}^n$ is

$$\Delta_B f(x) = \sum_{j,k} (B^{-1})_{jk} \frac{\partial^2 f}{\partial x_j \partial x_k}(x).$$

Eigenfunctions of Δ_B are the real and imaginary parts of functions of the form

$$f_v(x) = e^{2\pi i (v,x)_B} = e^{2\pi i \langle Bv, x \rangle}$$

for vectors $v \in \mathbb{R}^n$ such that Bv has integer components, with corresponding eigenvalue $-\mu_v = -(2\pi \|v\|_B)^2$.

Let

$$f(x) := \Re\left(\sum_{v\in\mathcal{V}} a_v e^{2\pi i\langle Bv,x\rangle}\right)$$

for $\mathcal{V}$ a finite collection of vectors v such that Bv has integer components for each $v \in \mathcal{V}$ and $\{a_v\}_{v\in\mathcal{V}}$ is a random vector on the sphere of radius $\sqrt{2}$ in $\mathbb{R}^{\mathcal{V}}$. Assume that $v + w \neq 0$ for $v, w \in \mathcal{V}$.

Let

$$f(x) := \Re\left(\sum_{v\in\mathcal{V}} a_v e^{2\pi i\langle Bv,x\rangle}\right)$$

for $\mathcal{V}$ a finite collection of vectors v such that Bv has integer components for each $v \in \mathcal{V}$ and $\{a_v\}_{v\in\mathcal{V}}$ is a random vector on the sphere of radius $\sqrt{2}$ in $\mathbb{R}^{\mathcal{V}}$. Assume that $v + w \neq 0$ for $v, w \in \mathcal{V}$.

Theorem

For a random function f on $(\mathbb{T}^n, B)$ defined as above and for $\mu > 0$,

$$\mathbb{E}d_{TV}(f(X), Z) \le \frac{1}{|\mu|}\left[\sqrt{\frac{8(2\pi)^4}{|\mathcal{V}|(|\mathcal{V}|+2)}\sum_{v,w\in\mathcal{V}}(v,w)_B^2} + \left(2\sqrt{2}+\sqrt{\pi}\right)\sqrt{\frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}}(\mu_v-\mu)^2}\right].$$

Example

Let $B = I$. Choose $\mathcal{V}$ to be the set of vectors in $\mathbb{R}^n$ with two entries equal to one and all others zero.

Example

Let $B = I$. Choose $\mathcal{V}$ to be the set of vectors in $\mathbb{R}^n$ with two entries equal to one and all others zero. Then $|\mathcal{V}| = \binom{n}{2}$ and $\|v\|^2 = 2$ for each $v \in \mathcal{V}$.

Example

Let $B = I$. Choose $\mathcal{V}$ to be the set of vectors in $\mathbb{R}^n$ with two entries equal to one and all others zero. Then $|\mathcal{V}| = \binom{n}{2}$ and $\|v\|^2 = 2$ for each $v \in \mathcal{V}$. Given $v \in \mathcal{V}$, there are exactly $2n - 3$ elements $w \in \mathcal{V}$ (including v itself) such that $\langle v, w \rangle \neq 0$.

Example

Let $B = I$. Choose $\mathcal{V}$ to be the set of vectors in $\mathbb{R}^n$ with two entries equal to one and all others zero. Then $|\mathcal{V}| = \binom{n}{2}$ and $\|v\|^2 = 2$ for each $v \in \mathcal{V}$. Given $v \in \mathcal{V}$, there are exactly $2n - 3$ elements $w \in \mathcal{V}$ (including v itself) such that $\langle v, w \rangle \neq 0$. Thus

$$\begin{aligned} \frac{1}{|\mathcal{V}|(|\mathcal{V}| + 2)} \sum_{v,w \in \mathcal{V}} \langle v, w \rangle^2 &\leq \frac{1}{\binom{n}{2}\left(\binom{n}{2} + 2\right)} \cdot \binom{n}{2}(4)(2n - 3) \\ &= \frac{16n - 12}{n^2 - n + 4}, \end{aligned}$$

so in this case there is a constant c such that

$$\mathbb{E} d_{TV}(f(X), Z) \leq \frac{c}{\sqrt{n}}.$$

Fancier version of the example

Let $B = diag(1 + \delta_1, \ldots, 1 + \delta_n)$, with δ_i small real numbers and let $\mathcal{V}$ be as before.

Fancier version of the example

Let $B = diag(1 + \delta_1, \ldots, 1 + \delta_n)$, with δ_i small real numbers and let $\mathcal{V}$ be as before. For $v \in \mathcal{V}$, let $v' = \left(\frac{1}{1+\delta_1} v_1, \ldots, \frac{1}{1+\delta_n} v_n\right)$ and let $\mathcal{V}' = \{v' : v \in \mathcal{V}\}$.

Fancier version of the example

Let $B = diag(1+\delta_1, \ldots, 1+\delta_n)$, with δ_i small real numbers and let $\mathcal{V}$ be as before. For $v \in \mathcal{V}$, let $v' = \left(\frac{1}{1+\delta_1} v_1, \ldots, \frac{1}{1+\delta_n} v_n\right)$ and let $\mathcal{V}' = \{v' : v \in \mathcal{V}\}$. Taking $\mu = 2(2\pi)^2$ in the theorem and $\epsilon := \max_i \left|1 - \frac{1}{1+\delta_i}\right|$, one can modify the previous argument to show that for f defined as in the theorem, using this matrix B and index set $\mathcal{V}'$, there is a constant C so that

$$\mathbb{E} d_{TV}(f(X), Z) \leq C \max\left(\frac{1}{\sqrt{n}}, \epsilon\right).$$

A multivariate version

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \stackrel{d}{=} X_\epsilon$, *such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely.*

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \stackrel{d}{=} X_\epsilon$*, such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely. Let Z be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$*.*

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \stackrel{d}{=} X_\epsilon$*, such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely. Let* Z *be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$*. Suppose there is a function* $\lambda(\epsilon)$ *and a random matrix* F *such that*

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \overset{d}{=} X_\epsilon$*, such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely. Let* Z *be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$*. Suppose there is a function* $\lambda(\epsilon)$ *and a random matrix* F *such that*

1. $\frac{1}{\lambda(\epsilon)} \mathbb{E}\left[X_\epsilon - X \middle| X\right] \xrightarrow[\epsilon \to 0]{L_1} -X.$

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \overset{d}{=} X_\epsilon$*, such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely. Let* Z *be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$*. Suppose there is a function* $\lambda(\epsilon)$ *and a random matrix* F *such that*

1. $\frac{1}{\lambda(\epsilon)} \mathbb{E}\left[X_\epsilon - X \middle| X\right] \xrightarrow[\epsilon \to 0]{L_1} -X$.
2. $\frac{1}{2\lambda(\epsilon)} \mathbb{E}\left[(X_\epsilon - X)(X_\epsilon - X)^T \middle| X\right] \xrightarrow[\epsilon \to 0]{L_1} \sigma^2 \mathrm{I}_k + \mathbb{E}\left[F \middle| X\right]$.

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \stackrel{d}{=} X_\epsilon$, *such that* $\lim_{\epsilon\to 0} X_\epsilon = X$ *almost surely. Let* Z *be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$. *Suppose there is a function* $\lambda(\epsilon)$ *and a random matrix* F *such that*

1. $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[X_\epsilon - X \middle| X\right] \xrightarrow[\epsilon\to 0]{L_1} -X$.
2. $\frac{1}{2\lambda(\epsilon)}\mathbb{E}\left[(X_\epsilon - X)(X_\epsilon - X)^T \middle| X\right] \xrightarrow[\epsilon\to 0]{L_1} \sigma^2 \mathrm{I}_k + \mathbb{E}\left[F \middle| X\right]$.
3. *For all* $\rho > 0$, $\frac{1}{\lambda(\epsilon)}\mathbb{E}\left[|X_\epsilon - X|^2 \mathbb{I}(|X_\epsilon - X| > \rho)\right] \xrightarrow{\epsilon\to 0} 0$.

A multivariate version

Theorem (Chatterjee-M)

Suppose that (X, X_ϵ) *is a family of random vectors of* $\mathbb{R}^k$ *with* $X \stackrel{d}{=} X_\epsilon$*, such that* $\lim_{\epsilon \to 0} X_\epsilon = X$ *almost surely. Let* Z *be a normal random vector with mean 0 and covariance* $\sigma^2 \mathrm{I}_k$*. Suppose there is a function* $\lambda(\epsilon)$ *and a random matrix* F *such that*

1. $\frac{1}{\lambda(\epsilon)} \mathbb{E}\left[X_\epsilon - X \middle| X\right] \xrightarrow[\epsilon \to 0]{L_1} -X.$
2. $\frac{1}{2\lambda(\epsilon)} \mathbb{E}\left[(X_\epsilon - X)(X_\epsilon - X)^T \middle| X\right] \xrightarrow[\epsilon \to 0]{L_1} \sigma^2 \mathrm{I}_k + \mathbb{E}\left[F \middle| X\right].$
3. *For all* $\rho > 0$, $\frac{1}{\lambda(\epsilon)} \mathbb{E}\left[|X_\epsilon - X|^2 \mathbb{I}(|X_\epsilon - X| > \rho)\right] \xrightarrow{\epsilon \to 0} 0.$

Then

$$d_W(X, Z) \le \frac{1}{\sigma} \mathbb{E}\|F\|_{H.S.}.$$

Application 1

Rank k projections of Haar measure on $\mathcal{O}_n$

Application 1

Rank k projections of Haar measure on $\mathcal{O}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let $A_1, \ldots, A_k$ be $n \times n$ matrices over $\mathbb{R}$.

Application 1

Rank k projections of Haar measure on $\mathcal{O}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let $A_1, \ldots, A_k$ be $n \times n$ matrices over $\mathbb{R}$. Assume that $\mathrm{Tr}(A_i A_j^T) = n\delta_{ij}$.

Application 1

Rank k projections of Haar measure on $\mathcal{O}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let $A_1, \ldots, A_k$ be $n \times n$ matrices over $\mathbb{R}$. Assume that $\mathrm{Tr}(A_i A_j^T) = n\delta_{ij}$*. Define the random vector $X \in \mathbb{R}^k$ by*

$$X := (\mathrm{Tr}(A_1 M), \ldots, \mathrm{Tr}(A_k M)).$$

Application 1

Rank k projections of Haar measure on $\mathcal{O}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{O}_n$ be a Haar-distributed random orthogonal matrix, and let $A_1, \ldots, A_k$ be $n \times n$ matrices over $\mathbb{R}$. Assume that $\mathrm{Tr}(A_i A_j^T) = n\delta_{ij}$. Define the random vector $X \in \mathbb{R}^k$ by

$$X := (\mathrm{Tr}(A_1 M), \ldots, \mathrm{Tr}(A_k M)).$$

If Z is a standard Gaussian random vector, then

$$d_W(X, Z) \leq \frac{\sqrt{2}k}{n-1}.$$

A little (recent) historical context

A little (recent) historical context

In 2006, Tiefeng Jiang showed that a $p \times q$ submatrix of an $n \times n$ random orthogonal matrix is close (in total variation distance) to a Gaussian random matrix of the same size, as long as $p = o(\sqrt{n})$ and $q = o(\sqrt{n})$.

A little (recent) historical context

In 2006, Tiefeng Jiang showed that a $p \times q$ submatrix of an $n \times n$ random orthogonal matrix is close (in total variation distance) to a Gaussian random matrix of the same size, as long as $p = o(\sqrt{n})$ and $q = o(\sqrt{n})$. The result above is more general in that it is valid for any rank k projection of Haar measure when $k = o(n)$, not just those having the special form of truncation to a submatrix.

A little (recent) historical context

In 2006, Tiefeng Jiang showed that a $p \times q$ submatrix of an $n \times n$ random orthogonal matrix is close (in total variation distance) to a Gaussian random matrix of the same size, as long as $p = o(\sqrt{n})$ and $q = o(\sqrt{n})$. The result above is more general in that it is valid for any rank k projection of Haar measure when $k = o(n)$, not just those having the special form of truncation to a submatrix. The convergence rates themselves are in different metrics and are not directly comparable.

A little (recent) historical context

In 2006, Tiefeng Jiang showed that a $p \times q$ submatrix of an $n \times n$ random orthogonal matrix is close (in total variation distance) to a Gaussian random matrix of the same size, as long as $p = o(\sqrt{n})$ and $q = o(\sqrt{n})$. The result above is more general in that it is valid for any rank k projection of Haar measure when $k = o(n)$, not just those having the special form of truncation to a submatrix. The convergence rates themselves are in different metrics and are not directly comparable.

Also in 2006 (after a preliminary version of the result above had been announced), B. Collins and M. Stolz proved that random vectors of the type above defined not only for the orthogonal group, but for matrix representations of several more general homogeneous spaces, converged weakly to Gaussian.

Application 2

Rank k projections of Haar measure on $\mathcal{U}_n$

Application 2

Rank k projections of Haar measure on $\mathcal{U}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{U}_n$ be a Haar-distributed random unitary matrix, and let $A_1, \ldots, A_k$ be fixed $n \times n$ matrices over $\mathbb{C}$.

Application 2

Rank k projections of Haar measure on $\mathcal{U}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{U}_n$ be a Haar-distributed random unitary matrix, and let $A_1, \ldots, A_k$ be fixed $n \times n$ matrices over $\mathbb{C}$. Assume that $\mathrm{Tr}(A_i A_j^*) = n\delta_{ij}$.

Application 2

Rank k projections of Haar measure on $\mathcal{U}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{U}_n$ be a Haar-distributed random unitary matrix, and let $A_1, \ldots, A_k$ be fixed $n \times n$ matrices over $\mathbb{C}$. Assume that $\mathrm{Tr}(A_i A_j^*) = n\delta_{ij}$. Define the random vector $X \in \mathbb{C}^k$ by

$$X := (\mathrm{Tr}(A_1 M), \ldots, \mathrm{Tr}(A_k M)).$$

Application 2

Rank k projections of Haar measure on $\mathcal{U}_n$

Theorem (Chatterjee-M)

Let $M \in \mathcal{U}_n$ be a Haar-distributed random unitary matrix, and let $A_1, \ldots, A_k$ be fixed $n \times n$ matrices over $\mathbb{C}$. Assume that $\mathrm{Tr}(A_i A_j^*) = n\delta_{ij}$. Define the random vector $X \in \mathbb{C}^k$ by

$$X := (\mathrm{Tr}(A_1 M), \ldots, \mathrm{Tr}(A_k M)).$$

If Z is a standard complex Gaussian random vector, then for $n \geq 4$,

$$d_W(X, Z) \leq \frac{3k}{n-1}.$$