# Introduction to Christoffel-Darboux kernels for data analysis

Edouard Pauwels
collaboration with François Bachoc and Jean-Bernard Lasserre and Mihai Putinar and Trang-May Vu.

Journée PMNL

Montpellier, October 2022

# Content of the presentation

Introduction of Christoffel-Darboux kernels and Christoffel function.

Overview of first properties.

Statistical aspects and application to support inference.

# Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

# Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

- A CD kernel $K^{\mu}$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$

# Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

- A CD kernel $K^{\mu}$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^{\mu} : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^{\mu}(\mathbf{x}, \mathbf{x})$.

# Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

- A CD kernel $K^\mu$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^\mu : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^\mu(\mathbf{x}, \mathbf{x})$.
- $K^\mu$, is a polynomial, we actually have $(K_d^\mu)_{d\in\mathbb{N}}$, where $K_d^\mu$ is of degree $2d$, $d \in \mathbb{N}$.

# Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

- A CD kernel $K^{\mu}$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^{\mu}\colon \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^{\mu}(\mathbf{x}, \mathbf{x})$.
- $K^{\mu}$, is a polynomial, we actually have $(K_d^{\mu})_{d\in\mathbb{N}}$, where $K_d^{\mu}$ is of degree $2d$, $d \in \mathbb{N}$.
- It captures information on $\mu$ (support, density).

## Motivation

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

- A CD kernel $K^{\mu}$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^{\mu} : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^{\mu}(\mathbf{x}, \mathbf{x})$.
- $K^{\mu}$, is a polynomial, we actually have $(K_d^{\mu})_{d \in \mathbb{N}}$, where $K_d^{\mu}$ is of degree $2d$, $d \in \mathbb{N}$.
- It captures information on $\mu$ (support, density).
- It is easily computed from moments of the measure.

**Mathematical context:**

- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

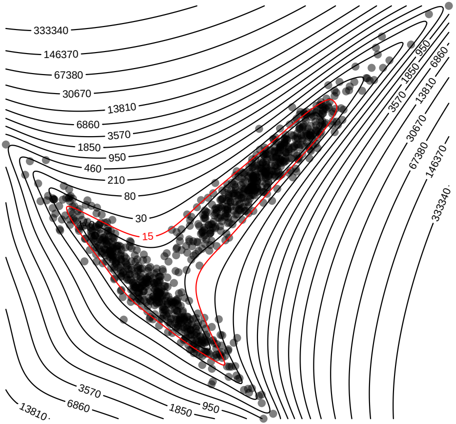
- A CD kernel $K^{\mu}$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^{\mu}: \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^{\mu}(\mathbf{x}, \mathbf{x})$.
- $K^{\mu}$, is a polynomial, we actually have $(K_d^{\mu})_{d \in \mathbb{N}}$, where $K_d^{\mu}$ is of degree $2d$, $d \in \mathbb{N}$.
- It captures information on $\mu$ (support, density).
- It is easily computed from moments of the measure.
- Moments (pseudomoments) of measures are outputs of Lassere's Hierarchy.

**Mathematical context:**

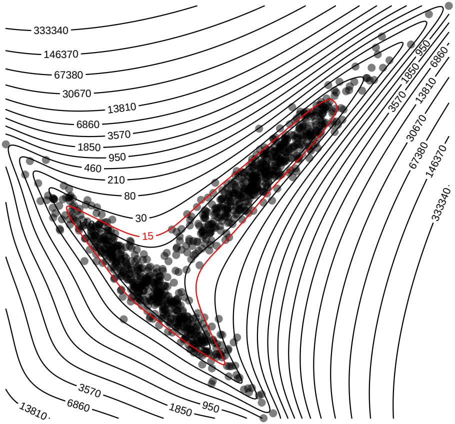
- Christoffel-Darboux (CD) kernels and orthogonal polynomials $\sim$ 19-th century.
- Important in approximation theory (convergence of generalized Fourier series).
- There is still a lot of activity going on.
- Recently used in, data science, polynomial optimization contexts.

**In a nutshell**

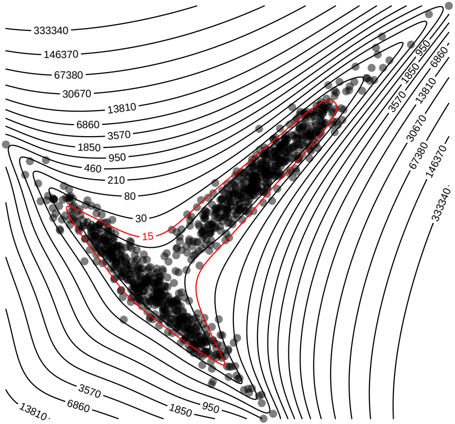
- A CD kernel $K^\mu$ depends on a (probability) measure $\mu$ on a Euclidean space $\mathbb{R}^p$
- We have $K^\mu : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, and will often consider $K^\mu(\mathbf{x}, \mathbf{x})$.
- $K^\mu$, is a polynomial, we actually have $(K_d^\mu)_{d \in \mathbb{N}}$, where $K_d^\mu$ is of degree $2d$, $d \in \mathbb{N}$.
- It captures information on $\mu$ (support, density).
- It is easily computed from moments of the measure.
- Moments (pseudomoments) of measures are outputs of Lassere's Hierarchy.
- Moments correspond to empirical averages in a statistical context.


333340
146370
67380
30670
13810
6860
3570
1850
950
460
210
80
30
15
950
6860
1850
3570
13810
30670
67380
146370
333340
3570
6860
13810
1850
950

# How does it look like?




- $(\mathbf{x}_i)_{i=1}^N$ is a set of points in $\mathbb{R}^2$ (black dots).
- $\mu$ is the empirical average $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_i}$.
- The CD kernel is a function on $\mathbb{R}^2 \times \mathbb{R}^2$ (level sets of $K_d^\mu(\mathbf{x}, \mathbf{x})$, $d = 4$).




- $(\mathbf{x}_i)_{i=1}^N$ is a set of points in $\mathbb{R}^2$ (black dots).
- $\mu$ is the empirical average $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_i}$.
- The CD kernel is a function on $\mathbb{R}^2 \times \mathbb{R}^2$ (level sets of $K_d^\mu(\mathbf{x}, \mathbf{x})$, $d = 4$).

**Plan for today:** Introduction of these objects and first properties.

# Outline

1. CD kernel, Christoffel function, orthogonal polynomials, moments
2. Quantitative asymptotics
3. Empirical measures statistical aspects
4. Application to support inference from sample

$\mu$: Borel probability measure in $\mathbb{R}^p$ .
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$\mu$: Borel probability measure in $\mathbb{R}^p$ .
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \qquad \mapsto \qquad \langle\!\langle P, Q \rangle\!\rangle_\mu := \int PQ d\mu,$$

defines a valid scalar product on $\mathbb{R}_d[X]$.

# Christoffel-Darboux kernel

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \quad \mapsto \quad \langle\langle P, Q \rangle\rangle_\mu := \int PQ d\mu,$$

defines a valid scalar product on $\mathbb{R}_d[X]$.

## Christoffel-Darboux kernel

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \quad \mapsto \quad \langle\langle P, Q \rangle\rangle_\mu := \int PQ d\mu,$$

defines a valid scalar product on $\mathbb{R}_d[X]$.

$(\mathbb{R}_d[X], \langle\langle \cdot, \cdot \rangle\rangle_\mu)$ is a *finite dimensional, Hilbert space* of functions from $\mathbb{R}^p$ to $\mathbb{R}$.

# Reproducing Kernel Hilbert Space (RKHS)

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn . . .

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn ...

Aronszajn (1950): $\mathcal{X}$ is a set and $\mathcal{H}$ a Hilbert space of real valued functions on $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\| \cdot \|_{\mathcal{H}}$.

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn ...

Aronszajn (1950): $\mathcal{X}$ is a set and $\mathcal{H}$ a Hilbert space of real valued functions on $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\| \cdot \|_{\mathcal{H}}$.

The following are equivalent:

- For any $x \in \mathcal{X}$, $h \mapsto h(x)$ is continuous on $\mathcal{H}$.

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn ...

Aronszajn (1950): $\mathcal{X}$ is a set and $\mathcal{H}$ a Hilbert space of real valued functions on $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\| \cdot \|_{\mathcal{H}}$.

The following are equivalent:

- For any $x \in \mathcal{X}$, $h \mapsto h(x)$ is continuous on $\mathcal{H}$.
- There is a unique symmetric positive definite $K : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{R}$ such that,

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn ...

Aronszajn (1950): $\mathcal{X}$ is a set and $\mathcal{H}$ a Hilbert space of real valued functions on $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\| \cdot \|_{\mathcal{H}}$.

The following are equivalent:

- For any $x \in \mathcal{X}$, $h \mapsto h(x)$ is continuous on $\mathcal{H}$.
- There is a unique symmetric positive definite $K : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{R}$ such that,
  - for any $x \in \mathcal{X}$, $K(x, \cdot) \in \mathcal{H}$
  - for any $x \in \mathcal{X}$, $h \in \mathcal{H}$

$$\langle h, K(x, \cdot) \rangle_{\mathcal{H}} = h(x).$$

**Hilbert space method (first half of 20-th century):** Zarembda, Mercer, Moore, Szegö, Bergman, Bochner, Kolmogorov, Aronszajn ...

Aronszajn (1950): $\mathcal{X}$ is a set and $\mathcal{H}$ a Hilbert space of real valued functions on $\mathcal{X}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\| \cdot \|_{\mathcal{H}}$.

The following are equivalent:

- For any $x \in \mathcal{X}$, $h \mapsto h(x)$ is continuous on $\mathcal{H}$.
- There is a unique symmetric positive definite $K : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{R}$ such that,
  - for any $x \in \mathcal{X}$, $K(x, \cdot) \in \mathcal{H}$
  - for any $x \in \mathcal{X}$, $h \in \mathcal{H}$

$$\langle h, K(x, \cdot) \rangle_{\mathcal{H}} = h(x).$$

$\mathcal{H}$ is called RKHS and $K$ is *the* reproducing kernel of $\mathcal{H}$.

# Reproducing Kernel Hilbert Space (RKHS)

$\mathcal{H} = (\mathbb{R}_d[X], \langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu)$ is a Reproducing Kernel Hilbert Space (RKHS):

- Evaluation is continuous with respect to coefficients.
- Finite dimension, all norms are equivalent: $\| \cdot \|_\mu$ and any norm on coefficients.

# Reproducing Kernel Hilbert Space (RKHS)

$\mathcal{H} = (\mathbb{R}_d[X], \langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu)$ is a Reproducing Kernel Hilbert Space (RKHS):

- Evaluation is continuous with respect to coefficients.
- Finite dimension, all norms are equivalent: $\|\cdot\|_\mu$ and any norm on coefficients.

**Reproducing property:** For all $d \in \mathbb{N}$, there exists $K_d^\mu : \mathbb{R}^p \times \mathbb{R}^p \mapsto \mathbb{R}$, symmetric such that for all $\mathbf{z} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{z}, \cdot) \in \mathbb{R}_d[X].$$

$K_d^\mu$ satifies the reproducing property, for all $P \in \mathbb{R}_d[X]$ and $\mathbf{z} \in \mathbb{R}^p$,

$$P(\mathbf{z}) = \langle\!\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \rangle\!\rangle_\mu = \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x})$$

$\mathcal{H} = (\mathbb{R}_d[X], \langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu)$ is a Reproducing Kernel Hilbert Space (RKHS):

- Evaluation is continuous with respect to coefficients.
- Finite dimension, all norms are equivalent: $\| \cdot \|_\mu$ and any norm on coefficients.

**Reproducing property:** For all $d \in \mathbb{N}$, there exists $K_d^\mu : \mathbb{R}^p \times \mathbb{R}^p \mapsto \mathbb{R}$, symmetric such that for all $\mathbf{z} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{z}, \cdot) \in \mathbb{R}_d[X].$$

$K_d^\mu$ satifies the reproducing property, for all $P \in \mathbb{R}_d[X]$ and $\mathbf{z} \in \mathbb{R}^p$,

$$P(\mathbf{z}) \quad = \quad \langle\!\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \rangle\!\rangle_\mu \quad = \quad \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x})$$

**Christoffel-Darboux kernel**: $K_\mu^d$ is *the* reproducing kernel of $\mathcal{H}$.

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d)\times s(d)}$ (integral coordinate-wise).

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \dots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d) \times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \dots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d)\times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d \colon \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d)\times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

**Proof:** $c_P \in \mathbb{R}^{s(d)}$, coefficients. Verify the reproducing property: $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$,

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \dots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d)\times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

**Proof:** $c_P \in \mathbb{R}^{s(d)}$, coefficients. Verify the reproducing property: $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$,

$$\left\langle\!\left\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \right\rangle\!\right\rangle_\mu = \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x}) = \int c_P^T \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) d\mu(\mathbf{x})$$

## Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \dots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d) \times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

**Proof:** $c_P \in \mathbb{R}^{s(d)}$, coefficients. Verify the reproducing property: $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$,

$$\left\langle\!\left\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \right\rangle\!\right\rangle_\mu = \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x}) = \int c_P^T \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) d\mu(\mathbf{x})$$

$$= c_P^T \left( \int \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T d\mu(\mathbf{x}) \right) M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z})$$

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d) \times s(d)}$ (integral coordinate-wise).
- Let $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q \colon \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

**Proof:** $c_P \in \mathbb{R}^{s(d)}$, coefficients. Verify the reproducing property: $P \colon \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$,

$$\left\langle\!\left\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \right\rangle\!\right\rangle_\mu = \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x}) = \int c_P^T \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) d\mu(\mathbf{x})$$

$$= c_P^T \left( \int \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T d\mu(\mathbf{x}) \right) M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) = c_P^T \mathbf{v}_d(\mathbf{z}) = P(\mathbf{z})$$

# Computation from moments

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

- Let $\{P_i\}_{i=1}^{s(d)}$ be any basis of $\mathbb{R}_d[X]$,
- $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \dots, P_{s(d)}(\mathbf{x}))^T$.
- $M_{\mu,d} = \int \mathbf{v}_d \mathbf{v}_d^T d\mu \in \mathbb{R}^{s(d)\times s(d)}$ (integral coordinate-wise).
- Let $P : \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$, and $Q : \mathbf{x} \mapsto c_Q^T \mathbf{v}_d(\mathbf{x})$, then

$$c_Q^T M_{\mu,d} c_P = \int (c_Q^T \mathbf{v}_d(\mathbf{x}))(\mathbf{v}_d(\mathbf{x})^T c_P) d\mu(\mathbf{x}) = \int P(\mathbf{x}) Q(\mathbf{x}) d\mu(\mathbf{x}).$$

$M_{\mu,d}$ is invertible and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, $K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y})$.

**Proof:** $c_P \in \mathbb{R}^{s(d)}$, coefficients. Verify the reproducing property: $P : \mathbf{x} \mapsto c_P^T \mathbf{v}_d(\mathbf{x})$,

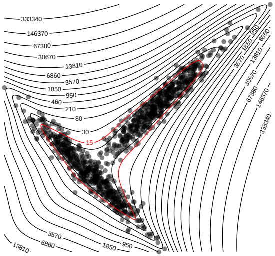
$$\langle\!\langle P(\cdot), K_d^\mu(\mathbf{z}, \cdot) \rangle\!\rangle_\mu = \int P(\mathbf{x}) K_d^\mu(\mathbf{z}, \mathbf{x}) d\mu(\mathbf{x}) = \int c_P^T \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) d\mu(\mathbf{x})$$

$$= c_P^T \left( \int \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T d\mu(\mathbf{x}) \right) M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{z}) = c_P^T \mathbf{v}_d(\mathbf{z}) = P(\mathbf{z})$$

**Remark:**

- It does not depend on the choice of the basis.
- If $\mathbf{v}_d$ is the monomial basis, then we recover the usual moment matrix.

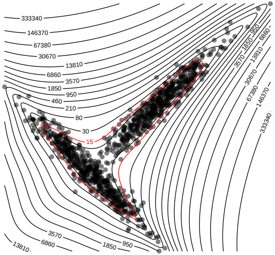
# (Practical computation: empirical measures)




**Empirical measure:** $\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$.

**Polynomial basis:** Choose $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.

# (Practical computation: empirical measures)



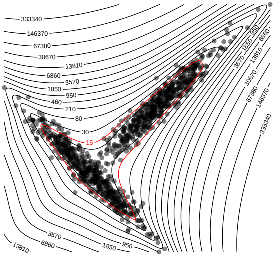

**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$.

**Polynomial basis:** Choose $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.

**Empirical moments:** $D \in \mathbb{R}^{N \times s(d)}$ rows given by $\mathbf{v}_d(X_i)$, $i = 1 \ldots N$ (design matrix)

$$M_{\mu_N,d} = \frac{1}{N} D^T D.$$

# (Practical computation: empirical measures)




**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$.

**Polynomial basis:** Choose $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.

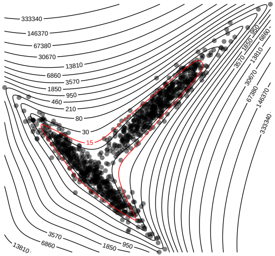
**Empirical moments:** $D \in \mathbb{R}^{N\times s(d)}$ rows given by $\mathbf{v}_d(X_i)$, $i = 1 \ldots N$ (design matrix)

$$M_{\mu_N,d} = \frac{1}{N}D^T D.$$

**Inverse moment matrix:** for all $\mathbf{x} \in \mathbb{R}^p$,

$$K_d^{\mu_N}(\mathbf{x},\mathbf{x}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu_N,d}^{-1}\mathbf{v}_d(\mathbf{x}).$$

**What's wrong?**




**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i}$.

**Polynomial basis:** Choose $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$.

**Empirical moments:** $D \in \mathbb{R}^{N \times s(d)}$ rows given by $\mathbf{v}_d(X_i)$, $i = 1 \ldots N$ (design matrix)

$$M_{\mu_N,d} = \frac{1}{N} D^T D.$$

**Inverse moment matrix:** for all $\mathbf{x} \in \mathbb{R}^p$,

$$K_d^{\mu_N}(\mathbf{x},\mathbf{x}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu_N,d}^{-1} \mathbf{v}_d(\mathbf{x}).$$

# Relation with orthogonal polynomials

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Relation with orthogonal polynomials**

Let $\{P_i\}_{i=1}^{s(d)}$ be any orthonormal basis of $\mathbb{R}_d[X]$ (*w.r.t.* $\langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu$),

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Relation with orthogonal polynomials**

Let $\{P_i\}_{i=1}^{s(d)}$ be any orthonormal basis of $\mathbb{R}_d[X]$ (*w.r.t.* $\langle\langle \cdot, \cdot \rangle\rangle_\mu$), then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{s(d)} P_i(\mathbf{x}) P_i(\mathbf{y}).$$

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Relation with orthogonal polynomials**

Let $\{P_i\}_{i=1}^{s(d)}$ be any orthonormal basis of $\mathbb{R}_d[X]$ (*w.r.t.* $\langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu$), then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{s(d)} P_i(\mathbf{x}) P_i(\mathbf{y}).$$

**Proof:** $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$, in this basis $M_{\mu,d} = I$

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Relation with orthogonal polynomials**

Let $\{P_i\}_{i=1}^{s(d)}$ be any orthonormal basis of $\mathbb{R}_d[X]$ (*w.r.t.* $\langle\langle \cdot, \cdot \rangle\rangle_\mu$), then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{s(d)} P_i(\mathbf{x}) P_i(\mathbf{y}).$$

**Proof:** $\mathbf{v}_d : \mathbf{x} \mapsto (P_1(\mathbf{x}), \ldots, P_{s(d)}(\mathbf{x}))^T$, in this basis $M_{\mu,d} = I$

**Remark:** monomial basis, Gram-Schmitt provides a canonical way to construct such a basis. This is at the heart of the (rich) theory of orthogonal polynomials.

# Christoffel function

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

# Christoffel function

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\begin{aligned} \Lambda_d^\mu : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \end{aligned}$$

# Christoffel function

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\Lambda_d^\mu : \mathbb{R}^p \mapsto [0,1]$$
$$\mathbf{z} \mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \quad = \quad \frac{1}{K_d^\mu(\mathbf{z},\mathbf{z})}.$$

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\begin{aligned} \Lambda_d^\mu : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \quad = \quad \frac{1}{K_d^\mu(\mathbf{z},\mathbf{z})}. \end{aligned}$$

For any $\mathbf{z} \in \mathbb{R}^p$ and $P \in \mathbb{R}_d[X]$ such that $P(\mathbf{z}) = 1$

$$P(\mathbf{z})^2 \quad = \quad 1 \quad = \quad \left( \int P(\mathbf{y}) K_d^\mu(\mathbf{z},\mathbf{y}) d\mu(\mathbf{y}) \right)^2$$

reproducing property,

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\begin{aligned} \Lambda_d^\mu : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \quad = \quad \frac{1}{K_d^\mu(\mathbf{z},\mathbf{z})}. \end{aligned}$$

For any $\mathbf{z} \in \mathbb{R}^p$ and $P \in \mathbb{R}_d[X]$ such that $P(\mathbf{z}) = 1$

$$\begin{aligned} P(\mathbf{z})^2 \quad = \quad 1 \quad &= \quad \left( \int P(\mathbf{y}) K_d^\mu(\mathbf{z},\mathbf{y}) d\mu(\mathbf{y}) \right)^2 \\ &\leq \quad \int P^2 d\mu \times \int K_d^\mu(\mathbf{z},\mathbf{y})^2 d\mu(\mathbf{y}) \end{aligned}$$

reproducing property, Cauchy-Schwartz

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\begin{aligned} \Lambda_d^\mu : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \quad = \quad \frac{1}{K_d^\mu(\mathbf{z},\mathbf{z})}. \end{aligned}$$

For any $\mathbf{z} \in \mathbb{R}^p$ and $P \in \mathbb{R}_d[X]$ such that $P(\mathbf{z}) = 1$

$$\begin{aligned} P(\mathbf{z})^2 \quad = \quad 1 \quad &= \quad \left( \int P(\mathbf{y}) K_d^\mu(\mathbf{z},\mathbf{y}) d\mu(\mathbf{y}) \right)^2 \\ &\leq \quad \int P^2 d\mu \times \int K_d^\mu(\mathbf{z},\mathbf{y})^2 d\mu(\mathbf{y}) \quad = \quad K_d^\mu(\mathbf{z},\mathbf{z}) \int P^2 d\mu \end{aligned}$$

reproducing property, Cauchy-Schwartz , reproducing property.

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).

**Christoffel function**

$$\begin{aligned} \Lambda_d^\mu : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\} \quad = \quad \frac{1}{K_d^\mu(\mathbf{z},\mathbf{z})}. \end{aligned}$$

For any $\mathbf{z} \in \mathbb{R}^p$ and $P \in \mathbb{R}_d[X]$ such that $P(\mathbf{z}) = 1$

$$\begin{aligned} P(\mathbf{z})^2 \quad = \quad 1 \quad &= \quad \left( \int P(\mathbf{y}) K_d^\mu(\mathbf{z},\mathbf{y}) d\mu(\mathbf{y}) \right)^2 \\ &\leq \quad \int P^2 d\mu \times \int K_d^\mu(\mathbf{z},\mathbf{y})^2 d\mu(\mathbf{y}) \quad = \quad K_d^\mu(\mathbf{z},\mathbf{z}) \int P^2 d\mu \end{aligned}$$

reproducing property, Cauchy-Schwartz , reproducing property.

Equality for $P(\cdot) = K_d^\mu(\mathbf{z},\cdot)/K_d^\mu(\mathbf{z},\mathbf{z})$.

# Historical remarks

Univariate case (complex and real) since beginning of 20-th century:

- quadrature, interpolation, approximation
- orthogonal polynomials
- potential theory
- random matrices/polynomials
- ...

A few contributors

- Szegö, Erdös, Turan, Freud, Totik, Máté, Nevai, ...

Still an object of very active research (asymptotics, multivariate case).

# Outline

1. CD kernel, Christoffel function, orthogonal polynomials, moments

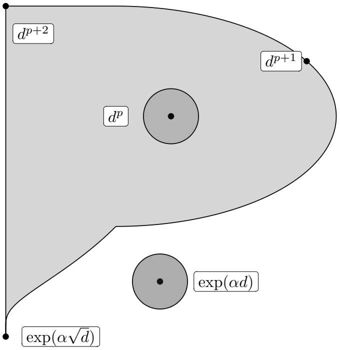
2. Quantitative asymptotics

3. Empirical measures statistical aspects

4. Application to support inference from sample

# Main idea

$\mu$: Lebesgue restricted to $S \subset \mathbb{R}^p$, compact, non-empty interior.
Order of growth of the CD kernel.



# The unit euclidean ball (Bos, Xu)

$\omega_p$ is the area of the $p$ dimensional unit sphere in $\mathbb{R}^{p+1}$.

$\omega_p$ is the area of the $p$ dimensional unit sphere in $\mathbb{R}^{p+1}$.

**Lebesgue measure on the ball:** Let $\lambda_B$ be the restriction of Lebesgue measure to the unit Euclidean ball $B \subset \mathbb{R}^p$. We have

$$K_d^{\lambda_B}(0,0) \leq \frac{s(d)}{\omega_p} \frac{(d+p+1)(d+p+2)(2d+p+6)}{(d+1)(d+2)(d+3)} = O(d^p)$$

$$K_d^{\lambda_B}(\mathbf{x},\mathbf{x}) = 2\binom{p+d+1}{d} - \binom{p+d}{d} = O(d^{p+1}), \qquad \|\mathbf{x}\| = 1.$$

**Exercise:** Show that if $\mu(A) \geq \nu(A)$ for all measurable set $A$, then for all $d$, $K_d^\mu \leq K_d^\nu$.

# Smooth boundary

**Exercise:** Show that if $\mu(A) \geq \nu(A)$ for all measurable set $A$, then for all $d$, $K_d^\mu \leq K_d^\nu$.

**Lebesgue measure on a set with non empty interior:** Let $S \subset \mathbb{R}^p$ have non empty interior. Then for all $\mathbf{x} \in \mathrm{int}(S)$,

$$K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^p)$$

**Exercise:** Show that if $\mu(A) \geq \nu(A)$ for all measurable set $A$, then for all $d$, $K_d^\mu \leq K_d^\nu$.

**Lebesgue measure on a set with non empty interior:** Let $S \subset \mathbb{R}^p$ have non empty interior. Then for all $\mathbf{x} \in \mathrm{int}(S)$,

$$K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^p)$$

**Proof:** $\mathbf{x} \in \mathrm{int}(S)$, there is a ball $B_r \subset S$ of radius $r$ and center $\mathbf{x}$. Consider $\lambda_{B_r} \leq \lambda_S$.

# Smooth boundary

**Exercise:** Show that if $\mu(A) \geq \nu(A)$ for all measurable set $A$, then for all $d$, $K_d^\mu \leq K_d^\nu$.

**Lebesgue measure on a set with non empty interior:** Let $S \subset \mathbb{R}^p$ have non empty interior. Then for all $\mathbf{x} \in \mathrm{int}(S)$,

$$K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^p)$$

If in addition the boundary of $S \subset \mathbb{R}^p$ is a smooth embedded hypersurface in $\mathbb{R}^p$. Then

$$\sup_{\mathbf{x} \in S} K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^{p+1}).$$

**Proof:** $\mathbf{x} \in \mathrm{int}(S)$, there is a ball $B_r \subset S$ of radius $r$ and center $\mathbf{x}$. Consider $\lambda_{B_r} \leq \lambda_S$.

**Exercise:** Show that if $\mu(A) \geq \nu(A)$ for all measurable set $A$, then for all $d$, $K_d^\mu \leq K_d^\nu$.

**Lebesgue measure on a set with non empty interior:** Let $S \subset \mathbb{R}^p$ have non empty interior. Then for all $\mathbf{x} \in \mathrm{int}(S)$,

$$K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^p)$$

If in addition the boundary of $S \subset \mathbb{R}^p$ is a smooth embedded hypersurface in $\mathbb{R}^p$. Then

$$\sup_{\mathbf{x} \in S} K_d^{\lambda_S}(\mathbf{x}, \mathbf{x}) = O(d^{p+1}).$$

**Proof:** $\mathbf{x} \in \mathrm{int}(S)$, there is a ball $B_r \subset S$ of radius $r$ and center $\mathbf{x}$. Consider $\lambda_{B_r} \leq \lambda_S$.

**Tubular neighborhood theorem:** There exists $r > 0$ such that for all $\mathbf{x} \in S$, there is a ball of radius $r$, $B_r \subset S$ such that $x \in B_r$. Consider $\lambda_{B_r} \leq \lambda_S$.

# Explicit construction: the cube $[-1,1]^p$

**Legendre Polynomials:** $P_0(t) = 0$, $P_1(t) = t$

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

$\max_{t\in[-1,1]} P_n(t) = 1$.

# Explicit construction: the cube $[-1,1]^p$

**Legendre Polynomials:** $P_0(t) = 0$, $P_1(t) = t$

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

$\max_{t\in[-1,1]} P_n(t) = 1$.

**Orthogonality:**

$$\int_{-1}^{1} P_m(t)P_n(t)dt = \frac{2}{2n+1}\delta_{mn}.$$

# Explicit construction: the cube $[-1, 1]^p$

**Legendre Polynomials:** $P_0(t) = 0$, $P_1(t) = t$

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

$\max_{t\in[-1,1]} P_n(t) = 1$.

**Orthogonality:**

$$\int_{-1}^{1} P_m(t)P_n(t)dt = \frac{2}{2n+1}\delta_{mn}.$$

**Lebesgue measure on the cube:** orthogonal polynomials given by

$$Q_\alpha(\mathbf{x}) = \prod_{i=1}^{p} \sqrt{\alpha_i + \frac{1}{2}} P_{\alpha_i}(x_i), \qquad \alpha \in \mathbb{N}_+^p, \quad |\alpha| < d$$

# Explicit construction: the cube $[-1,1]^p$

**Legendre Polynomials:** $P_0(t) = 0$, $P_1(t) = t$

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

$\max_{t\in[-1,1]} P_n(t) = 1$.

**Orthogonality:**

$$\int_{-1}^{1} P_m(t)P_n(t)dt = \frac{2}{2n+1}\delta_{mn}.$$

**Lebesgue measure on the cube:** orthogonal polynomials given by

$$Q_\alpha(\mathbf{x}) = \prod_{i=1}^{p} \sqrt{\alpha_i + \frac{1}{2}} P_{\alpha_i}(x_i), \qquad \alpha \in \mathbb{N}_+^p, \quad |\alpha| < d$$

Let $\lambda_C$ be the restriction of Lebesgue measure to the unit cube $C = [-1,1]^p$, then

$$\sup_{\mathbf{x}\in C} K_d^{\lambda_C}(\mathbf{x},\mathbf{x}) \leq \sum_{|\alpha|\leq d} \prod_{i=1}^{p} \left(\alpha_i + \frac{1}{2}\right) = O(d^{2p})$$

# Exponential lower bounds

Let $S \subset \mathbb{R}^p$ be compact and $\mu$ be a probability measure supported on $S$. Then for all $\mathbf{x}$ with $\mathrm{dist}(\mathbf{x}, S) \geq \delta > 0$, and $d \in \mathbb{N}$

$$K_d^{\mu}(\mathbf{x}, \mathbf{x}) \geq 2^{\frac{\delta d}{\delta + \mathrm{diam}(S)} - 3}.$$

Let $S \subset \mathbb{R}^p$ be compact and $\mu$ be a probability measure supported on $S$. Then for all $\mathbf{x}$ with $\mathrm{dist}(\mathbf{x}, S) \geq \delta > 0$, and $d \in \mathbb{N}$

$$K_d^\mu(\mathbf{x}, \mathbf{x}) \geq 2^{\frac{\delta d}{\delta + \mathrm{diam}(S)} - 3}.$$

**Proof:** For any $\mathbf{x} \in \mathbb{R}^p$ and $P \in \mathbb{R}_d[X]$ with $P(\mathbf{x}) = 1$,

$$K_\mu^d(\mathbf{x}, \mathbf{x}) \geq \left( \int P^2 d\mu \right)^{-1}.$$

Choose $P$ such that $P(\mathbf{x}) = 1$ and the integral is small.

*Kroó's needle polynomial*, for any $\delta > 0$, $d \in \mathbb{N}^*$, $\exists Q \in \mathbb{R}_{2d}[X]$

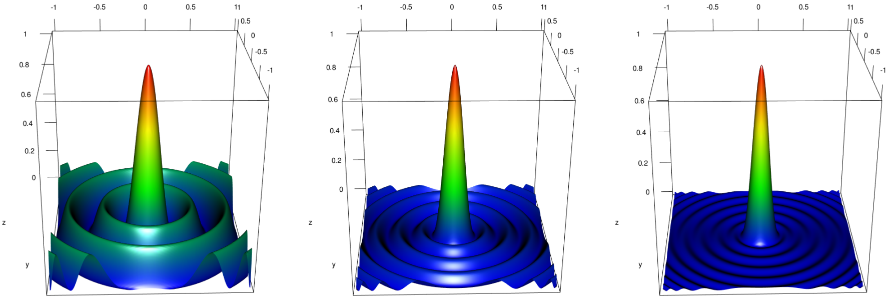
$$Q(0) = 1, \qquad |Q(\mathbf{x})| \leq 1 \text{ if } \|\mathbf{x}\| \leq 1, \qquad |Q(\mathbf{x})| \leq 2^{1-\delta d} \text{ if } \delta \leq \|\mathbf{x}\| \leq 1.$$

# Exponential lower bounds: Needle polynomial

*Kroó's needle polynomial*, for any $\delta > 0$, $d \in \mathbb{N}^*$, $\exists Q \in \mathbb{R}_{2d}[X]$

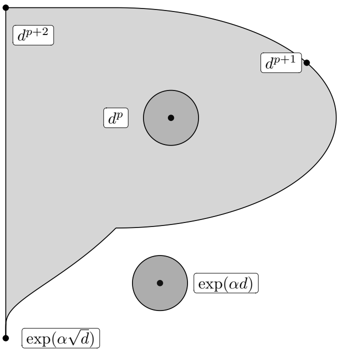
$$Q(0) = 1, \qquad |Q(\mathbf{x})| \leq 1 \text{ if } \|\mathbf{x}\| \leq 1, \qquad |Q(\mathbf{x})| \leq 2^{1-\delta d} \text{ if } \delta \leq \|\mathbf{x}\| \leq 1.$$

Example for $\delta = 0.2$ and $d = 20, 30, 40$.



# Exponential separation of the support

$\mu$: Lebesgue restricted to $S \subset \mathbb{R}^p$, compact, non-empty interior.




**Exponential growth dichotomy:** Growth of the CD kernel is

- At most polynomial in the degree $d$ in the interior of the support.
- Exponential in the degree $d$ outside the support.
- In between on the boundary of the support of $\mu$: depending on local geometry.

# Outline

# Empirical christoffel-Darboux kernel

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \quad \mapsto \quad \langle\langle P, Q \rangle\rangle_\mu := \int PQ d\mu,$$

defines a valid scalar product on $\mathbb{R}_d[X]$.

**Christoffel-Darboux kernel**: $K_\mu^d$ is *the* reproducing kernel of $(\mathbb{R}_d[X], \langle\langle \cdot, \cdot \rangle\rangle_\mu)$.

# Empirical christoffel-Darboux kernel

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \qquad \mapsto \qquad \langle\langle P, Q \rangle\rangle_\mu := \int PQ d\mu,$$

defines a valid scalar product on $\mathbb{R}_d[X]$.

**Christoffel-Darboux kernel**: $K_\mu^d$ is *the* reproducing kernel of $(\mathbb{R}_d[X], \langle\langle \cdot, \cdot \rangle\rangle_\mu)$.

**Empirical measure:** $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_i}$.

# Empirical christoffel-Darboux kernel

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support ).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \quad \mapsto \quad \langle\langle P, Q \rangle\rangle_\mu := \int PQ d\mu,$$

.

**Christoffel-Darboux kernel**: $K_\mu^d$ is *the* reproducing kernel of $(\mathbb{R}_d[X], \langle\langle \cdot, \cdot \rangle\rangle_\mu)$.

**Empirical measure:** $\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\mathbf{x}_i}$.

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support ).
$\mathbb{R}_d[X]$: $p$-variate polynomials of degree at most $d$ (of dimension $s(d) = \binom{p+d}{d}$).
(dimension $p$ and degree $d$ are fixed).

$$(P, Q) \qquad \mapsto \qquad \langle\!\langle P, Q \rangle\!\rangle_\mu := \int PQd\mu,$$

.

**Christoffel-Darboux kernel**: $K_\mu^d$ is *the* reproducing kernel of $(\mathbb{R}_d[X], \langle\!\langle \cdot, \cdot \rangle\!\rangle_\mu)$.

**Empirical measure:** $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_i}$.

**Example:** $P \colon \mathbf{x} \mapsto \prod_{i=1}^n \|\mathbf{x} - \mathbf{x}_i\|^2$, $\qquad \langle\!\langle P, P \rangle\!\rangle_{\mu_N} = 0$ but $P \neq 0$.

# Fix using empirical Christoffel function

**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\mathbf{x}_i}$

# Fix using empirical Christoffel function

**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_i}$

**Empirical christoffel function (no need for a valid scalar product)**

$$\begin{aligned} \Lambda_d^{\mu_N} : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \frac{1}{N}\sum_{i=1}^N P(X_i)^2 : \quad P(\mathbf{z}) = 1 \right\}. \end{aligned}$$

# Fix using empirical Christoffel function

**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_i}$

**Empirical christoffel function (no need for a valid scalar product)**

$$\begin{aligned} \Lambda_d^{\mu_N} : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \frac{1}{N} \sum_{i=1}^N P(X_i)^2 : \quad P(\mathbf{z}) = 1 \right\}. \end{aligned}$$

**Degeneracy for large** $d$**:** if $s(d) \geq N$, then

$$\Lambda_n^{\mu_N} : \mathbf{z} \mapsto \begin{cases} 0, & \mathbf{z} \neq \mathbf{x}_i,\ i = 1\ldots, N \\ \frac{1}{N}, & \text{otherwise.} \end{cases}$$

# Fix using empirical Christoffel function

**Empirical measure:** $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_i}$

**Empirical christoffel function (no need for a valid scalar product)**

$$\begin{aligned}\Lambda_d^{\mu_N} : \mathbb{R}^p &\mapsto [0,1] \\ \mathbf{z} &\mapsto \min_{P \in \mathbb{R}_d[X]} \left\{ \frac{1}{N}\sum_{i=1}^N P(X_i)^2 : \quad P(\mathbf{z}) = 1 \right\}.\end{aligned}$$

**Degeneracy for large $d$:** if $s(d) \geq N$, then

$$\Lambda_n^{\mu_N} : \mathbf{z} \mapsto \begin{cases} 0, & \mathbf{z} \neq \mathbf{x}_i,\ i = 1 \ldots, N \\ \frac{1}{N}, & \text{otherwise.} \end{cases}$$

How to ensure that $\Lambda_d^{\mu_N}$ retains the favorable properties of $\Lambda_d^{\mu}$?

# Empirical measure: statistical setting

**Statistical setting:** $(X_i)_{i\in\mathbb{N}}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$.

# Empirical measure: statistical setting

**Statistical setting:** $(X_i)_{i\in\mathbb{N}}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$.
For any $N \in \mathbb{N}$, measurable subsets $A_1, \dots, A_N$ in $\mathbb{R}^p$:

$$\mathbb{P}\left[(X_1 \in A_1) \ \& \ (X_2 \in A_2) \ \& \ \dots \ \& \ (X_N \in A_N)\right] = \prod_{i=1}^{N} \mu(A_i).$$

# Empirical measure: statistical setting

**Statistical setting:** $(X_i)_{i\in\mathbb{N}}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$.
For any $N \in \mathbb{N}$, measurable subsets $A_1, \dots, A_N$ in $\mathbb{R}^p$:

$$\mathbb{P}\left[(X_1 \in A_1) \ \& \ (X_2 \in A_2) \ \& \ \dots \ \& \ (X_N \in A_N)\right] = \prod_{i=1}^{N} \mu(A_i).$$

**Empirical measure:** For any $N \in \mathbb{N}$,

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}$$

# Empirical measure: statistical setting

**Statistical setting:** $(X_i)_{i\in\mathbb{N}}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$.
For any $N \in \mathbb{N}$, measurable subsets $A_1, \dots, A_N$ in $\mathbb{R}^p$:

$$\mathbb{P}\left[(X_1 \in A_1) \; \& \; (X_2 \in A_2) \; \& \; \dots \; \& \; (X_N \in A_N)\right] = \prod_{i=1}^{N} \mu(A_i).$$

**Empirical measure:** For any $N \in \mathbb{N}$,

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}$$

**Strong law of large numbers:** for any continuous $f$, almost surely

$$\lim_{N\to\infty} \int f(\mathbf{z}) d\mu_N(\mathbf{z}) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(X_i) = \int f(\mathbf{z}) d\mu(\mathbf{z}),$$

# Statistical consistency

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\dots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

# Statistical consistency

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\ldots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

**Consistency for fixed degree:** $d \in \mathbb{N}$ fixed, as $N \to \infty$, almost surely, uniformly in $\mathbf{x}$

$$\Lambda_d^{\mu_N}(\mathbf{x}) \to \Lambda_d^{\mu}(\mathbf{x}).$$

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\dots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

**Consistency for fixed degree:** $d \in \mathbb{N}$ fixed, as $N \to \infty$, almost surely, uniformly in $\mathbf{x}$

$$\Lambda_d^{\mu_N}(\mathbf{x}) \to \Lambda_d^{\mu}(\mathbf{x}).$$

**Finite sample concentration:** Set $m = \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x})$.

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\ldots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

**Consistency for fixed degree:** $d \in \mathbb{N}$ fixed, as $N \to \infty$, almost surely, uniformly in $\mathbf{x}$

$$\Lambda_d^{\mu_N}(\mathbf{x}) \to \Lambda_d^{\mu}(\mathbf{x}).$$

**Finite sample concentration:** Set $m = \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x})$.
Then, for all $\alpha > 0$, with probability at least $1-\alpha$,

$$\sup_{\mathbf{x}} |\Lambda_d^{\mu_N}(\mathbf{x}) - \Lambda_d^{\mu}(\mathbf{x})| \leq \max\left(\sqrt{\frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)}, \frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)\right)$$

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\dots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

**Consistency for fixed degree:** $d \in \mathbb{N}$ fixed, as $N \to \infty$, almost surely, uniformly in $\mathbf{x}$

$$\Lambda_d^{\mu_N}(\mathbf{x}) \to \Lambda_d^{\mu}(\mathbf{x}).$$

**Finite sample concentration:** Set $m = \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x})$.
Then, for all $\alpha > 0$, with probability at least $1-\alpha$,

$$\sup_{\mathbf{x}} |\Lambda_d^{\mu_N}(\mathbf{x}) - \Lambda_d^{\mu}(\mathbf{x})| \leq \max\left(\sqrt{\frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)}, \frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)\right)$$

**Nondegenerate regime:** $N \geq \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x}) \geq s(d)$.
Of order $d^{p+1}$ for smooth boundary.

# Statistical consistency

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\ldots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$

**Consistency for fixed degree:** $d \in \mathbb{N}$ fixed, as $N \to \infty$, almost surely, uniformly in $\mathbf{x}$

$$\Lambda_d^{\mu_N}(\mathbf{x}) \to \Lambda_d^{\mu}(\mathbf{x}).$$

**Finite sample concentration:** Set $m = \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x})$.
Then, for all $\alpha > 0$, with probability at least $1-\alpha$,

$$\sup_{\mathbf{x}} |\Lambda_d^{\mu_N}(\mathbf{x}) - \Lambda_d^{\mu}(\mathbf{x})| \leq \max\left(\sqrt{\frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)}, \frac{16m}{3N}\log\left(\frac{s(d)}{\alpha}\right)\right)$$

**Nondegenerate regime:** $N \geq \max_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^{\mu}(\mathbf{x},\mathbf{x}) \geq s(d)$.
Of order $d^{p+1}$ for smooth boundary.

**Statistical tools:** Concentration for random matrices (non commutative Bernstein).

# Outline

## Leveraging the exponential growth dichotomy

**Smooth boundary:** Let $S \subset \mathbb{R}^p$ compact, nonempty interior, boundary is a smooth embedded hypersurface. $\mu$ is the restriction of Lebesgue measures to $S$.

**Smooth boundary:** Let $S \subset \mathbb{R}^p$ compact, nonempty interior, boundary is a smooth embedded hypersurface. $\mu$ is the restriction of Lebesgue measures to $S$.

**Constants:** Then for any $\alpha > 0$, there exist constants $C_1$, $C_2$, $C_3$ which depend on $\alpha$ and can be computed from problem data,

## Leveraging the exponential growth dichotomy

**Smooth boundary:** Let $S \subset \mathbb{R}^p$ compact, nonempty interior, boundary is a smooth embedded hypersurface. $\mu$ is the restriction of Lebesgue measures to $S$.

**Constants:**Then for any $\alpha > 0$, there exist constants $C_1$, $C_2$, $C_3$ which depend on $\alpha$ and can be computed from problem data,

**Degree choice and threshold:** setting for all $N \in \mathbb{N}$,

$$d_N := \left\lfloor C_1 N^{\frac{1}{p+2}} \right\rfloor$$

$$\gamma_N := C_2 d_N^{\frac{3p}{2}}$$

$$S_N := \{\mathbf{x} \in \mathbb{R}^p,\ K_{d_N}^{\mu_N}(\mathbf{x}, \mathbf{x}) \leq \gamma_N\},$$

## Leveraging the exponential growth dichotomy

**Smooth boundary:** Let $S \subset \mathbb{R}^p$ compact, nonempty interior, boundary is a smooth embedded hypersurface. $\mu$ is the restriction of Lebesgue measures to $S$.

**Constants:** Then for any $\alpha > 0$, there exist constants $C_1$, $C_2$, $C_3$ which depend on $\alpha$ and can be computed from problem data,

**Degree choice and threshold:** setting for all $N \in \mathbb{N}$,

$$d_N := \left\lfloor C_1 N^{\frac{1}{p+2}} \right\rfloor$$

$$\gamma_N := C_2 d_N^{\frac{3p}{2}}$$

$$S_N := \{\mathbf{x} \in \mathbb{R}^p,\ K_{d_N}^{\mu_N}(\mathbf{x}, \mathbf{x}) \leq \gamma_N\},$$

**Set convergence:** it holds with probability at least $1 - \alpha$ that

$$d_H(S, S_N) \leq \frac{C_3}{N^{\frac{1}{2p+4}}} \qquad d_H(\partial S, \partial S_N) \leq \frac{C_3}{N^{\frac{1}{2p+4}}}$$

# Leveraging the exponential growth dichotomy

**Smooth boundary:** Let $S \subset \mathbb{R}^p$ compact, nonempty interior, boundary is a smooth embedded hypersurface. $\mu$ is the restriction of Lebesgue measures to $S$.

**Constants:** Then for any $\alpha > 0$, there exist constants $C_1$, $C_2$, $C_3$ which depend on $\alpha$ and can be computed from problem data,

**Degree choice and threshold:** setting for all $N \in \mathbb{N}$,

$$d_N := \left\lfloor C_1 N^{\frac{1}{p+2}} \right\rfloor$$
$$\gamma_N := C_2 d_N^{\frac{3p}{2}}$$
$$S_N := \{\mathbf{x} \in \mathbb{R}^p,\ K_{d_N}^{\mu_N}(\mathbf{x}, \mathbf{x}) \leq \gamma_N\},$$

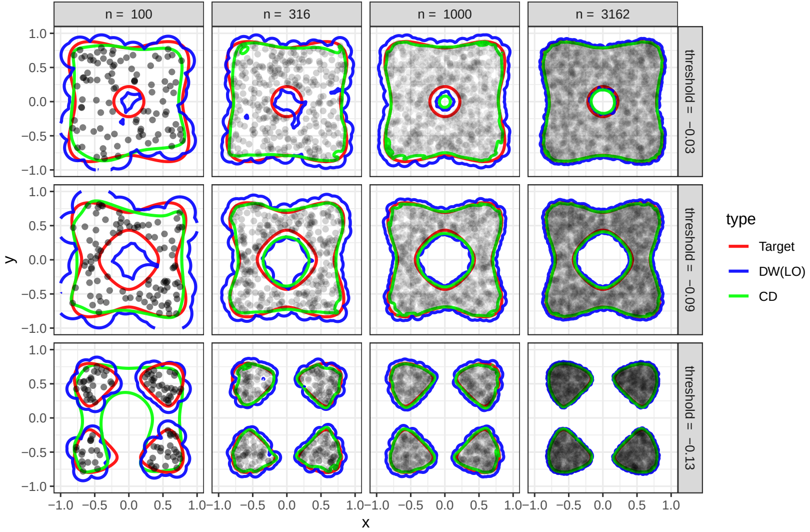
**Set convergence:** it holds with probability at least $1 - \alpha$ that

$$d_H(S, S_N) \leq \frac{C_3}{N^{\frac{1}{2p+4}}} \qquad d_H(\partial S, \partial S_N) \leq \frac{C_3}{N^{\frac{1}{2p+4}}}$$

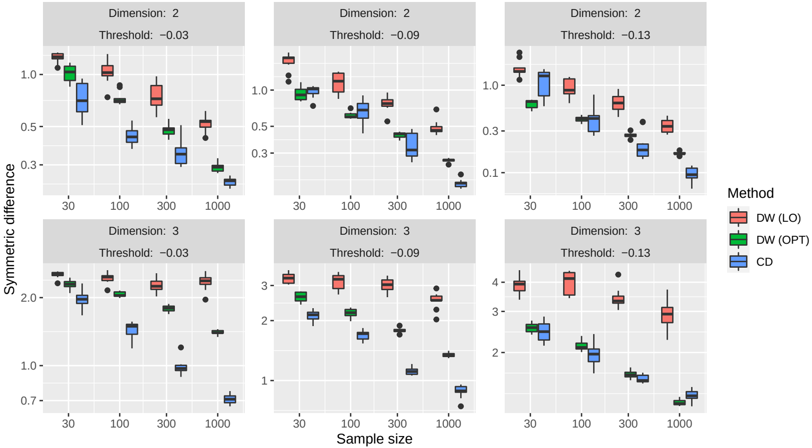
**Hausdorff Distance:** $d_H(X, Y)$ between two compact sets $X, Y$:

$$d_H(X, Y) = \max \left\{ \sup_{\mathbf{x} \in X} \inf_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|, \sup_{\mathbf{y} \in Y} \inf_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\| \right\}.$$

# Numerical illustration



# Numerical illustration



# Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).

# Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P, Q) \mapsto \int PQd\mu$.

# Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P, Q) \mapsto \int PQd\mu$.

**Inverse moment matrix:** $\mathbf{v}_d : \mathbb{R}^p \to \mathbb{R}^{s(d)}$ basis of $\mathbb{R}_d[X]$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y}).$$

# Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P, Q) \mapsto \int PQd\mu$.

**Inverse moment matrix:** $\mathbf{v}_d : \mathbb{R}^p \to \mathbb{R}^{s(d)}$ basis of $\mathbb{R}_d[X]$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y}).$$

**Variational formulation:** Christoffel function, for all $\mathbf{x} \in \mathbb{R}^p$,

$$\frac{1}{K_d^\mu(\mathbf{x}, \mathbf{x})} = \Lambda_d^\mu(\mathbf{x}) := \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\}.$$

# Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P, Q) \mapsto \int PQd\mu$.

**Inverse moment matrix:** $\mathbf{v}_d : \mathbb{R}^p \to \mathbb{R}^{s(d)}$ basis of $\mathbb{R}_d[X]$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x}, \mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y}).$$

**Variational formulation:** Christoffel function, for all $\mathbf{x} \in \mathbb{R}^p$,

$$\frac{1}{K_d^\mu(\mathbf{x}, \mathbf{x})} = \Lambda_d^\mu(\mathbf{x}) := \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\}.$$

**Exponential growth dichotomy:** as $d$ grows, $K_d^\mu(\mathbf{x}, \mathbf{x})$ goes to infinity

- At most in $O(d)$ in the interior of the support of $\mu$.
- At least exponentially outside the support of $\mu$.

## Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\ldots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$.

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P,Q) \mapsto \int PQd\mu$.

**Inverse moment matrix:** $\mathbf{v}_d : \mathbb{R}^p \to \mathbb{R}^{s(d)}$ basis of $\mathbb{R}_d[X]$, for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x},\mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y}).$$

**Variational formulation:** Christoffel function, for all $\mathbf{x} \in \mathbb{R}^p$,

$$\frac{1}{K_d^\mu(\mathbf{x},\mathbf{x})} = \Lambda_d^\mu(\mathbf{x}) := \min_{P \in \mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\}.$$

**Exponential growth dichotomy:** as $d$ grows, $K_d^\mu(\mathbf{x},\mathbf{x})$ goes to infinity

- At most in $O(d)$ in the interior of the support of $\mu$.
- At least exponentially outside the support of $\mu$.

## Takeaway

$\mu$: Borel probability measure in $\mathbb{R}^p$ (compact support, density).
$(X_i)_{i=1,\ldots,N}$, $\mathbb{R}^p$ valued random variables, *iid* with distribution $\mu$
$\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$.

**RKHS:** $K_d^\mu$ is *the* reproducing kernel of $\mathbb{R}_d[X]$ with dot product $(P,Q) \mapsto \int PQd\mu$.

**Inverse moment matrix:** $\mathbf{v}_d : \mathbb{R}^p \to \mathbb{R}^{s(d)}$ basis of $\mathbb{R}_d[X]$, for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^p$,

$$K_d^\mu(\mathbf{x},\mathbf{y}) = \mathbf{v}_d(\mathbf{x})^T M_{\mu,d}^{-1} \mathbf{v}_d(\mathbf{y}).$$

**Variational formulation:** Christoffel function, for all $\mathbf{x} \in \mathbb{R}^p$,

$$\frac{1}{K_d^\mu(\mathbf{x},\mathbf{x})} = \Lambda_d^\mu(\mathbf{x}) := \min_{P\in\mathbb{R}_d[X]} \left\{ \int P^2 d\mu : \quad P(\mathbf{z}) = 1 \right\}.$$

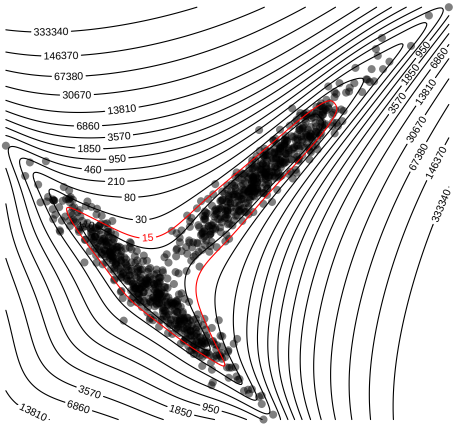
**Exponential growth dichotomy:** as $d$ grows, $K_d^\mu(\mathbf{x},\mathbf{x})$ goes to infinity

- At most in $O(d)$ in the interior of the support of $\mu$.
- At least exponentially outside the support of $\mu$.

**Statistical approximation:** $\Lambda_d^{\mu_N} \sim \Lambda_d^\mu$ provided that

$$N \geq \sup_{\mathbf{x}\in\mathrm{supp}(\mu)} K_d^\mu(\mathbf{x},\mathbf{x}) \geq s(d).$$

# Thanks



333340
146370
67380
30670
13810
6860
3570
1850
950
460
210
80
30
15
950
1850
3570
6860
13810
30670
67380
146370
333340
3570
6860
13810
1850
950