

El Kernel de Christoffel-Darboux para Análisis de Datos

Leandro Bentancur

Seminario de Optimización y Aprendizaje Automático

14 de noviembre de 2024

Índice

Introducción

Función de Christoffel

Núcleos reproductores

Función de Christoffel Regularizada

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

1. Fijamos $d \in \mathbb{N}$ y notamos $s(d) = \binom{d+2}{2}$.

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

1. Fijamos $d \in \mathbb{N}$ y notamos $s(d) = \binom{d+2}{2}$.
2. Sea $\mathbf{v}_d(\mathbf{x}) = (1, x, y, x^2, xy, y^2, \dots, xy^{d-1}, y^d)$ el vector con todos los monomios $x^i y^j$ tal que $i + j \leq d$.

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

1. Fijamos $d \in \mathbb{N}$ y notamos $s(d) = \binom{d+2}{2}$.
2. Sea $\mathbf{v}_d(\mathbf{x}) = (1, x, y, x^2, xy, y^2, \dots, xy^{d-1}, y^d)$ el vector con todos los monomios $x^i y^j$ tal que $i + j \leq d$.
3. Construimos $X_d \in \mathbb{R}^{N \times s(d)}$ tal que su i -ésima fila es el vector $\mathbf{v}(\mathbf{x}_i)$, y la matriz real simétrica $M_d \in \mathbb{R}^{s(d) \times s(d)}$ con filas y columnas indexadas en los monomios tal que

$$M_d = \frac{1}{N} X_d^T X_d.$$

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

1. Fijamos $d \in \mathbb{N}$ y notamos $s(d) = \binom{d+2}{2}$.
2. Sea $\mathbf{v}_d(\mathbf{x}) = (1, x, y, x^2, xy, y^2, \dots, xy^{d-1}, y^d)$ el vector con todos los monomios $x^i y^j$ tal que $i + j \leq d$.
3. Construimos $X_d \in \mathbb{R}^{d \times s(d)}$ tal que su i -ésima fila es el vector $\mathbf{v}(\mathbf{x}_i)$, y la matriz real simétrica $M_d \in \mathbb{R}^{s(d) \times s(d)}$ con filas y columnas indexadas en los monomios tal que

$$M_d = \frac{1}{N} X_d^T X_d.$$

4. Definimos el polinomio

$$p_d(\mathbf{x}) = \mathbf{v}_d(\mathbf{x})^T M_d^{-1} \mathbf{v}_d(\mathbf{x}).$$

Introducción

Consideremos una nube de puntos en el plano $\mathbf{x}_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, N$. Vamos a seguir los siguientes pasos:

1. Fijamos $d \in \mathbb{N}$ y notamos $s(d) = \binom{d+2}{2}$.
2. Sea $\mathbf{v}_d(\mathbf{x}) = (1, x, y, x^2, xy, y^2, \dots, xy^{d-1}, y^d)$ el vector con todos los monomios $x^i y^j$ tal que $i + j \leq d$.
3. Construimos $X_d \in \mathbb{R}^{d \times s(d)}$ tal que su i -ésima fila es el vector $\mathbf{v}_d(\mathbf{x}_i)$, y la matriz real simétrica $M_d \in \mathbb{R}^{s(d) \times s(d)}$ con filas y columnas indexadas en los monomios tal que

$$M_d = \frac{1}{N} X_d^T X_d.$$

4. Definimos el polinomio

$$p_d(\mathbf{x}) = \mathbf{v}_d(\mathbf{x})^T M_d^{-1} \mathbf{v}_d(\mathbf{x}).$$

5. Graficamos los conjuntos de nivel $S_\gamma = \{\mathbf{x} \in \mathbb{R}^2 \mid p_d(\mathbf{x}) = \gamma\}$ para algunos valores de γ , entre ellos en particular para $\gamma = s(d)$.

Introducción

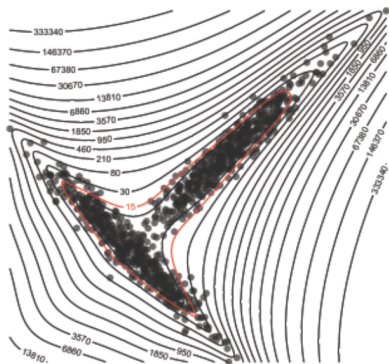


Figure 1.1 $n = 4$; $N = 1000$; level sets S_γ (in red for $\gamma = \binom{2+4}{2}$).

Introducción

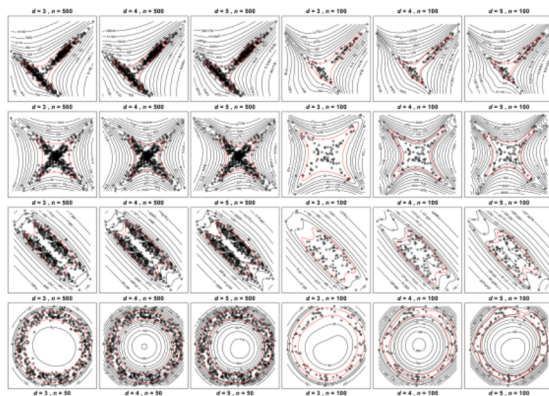


Figure:

Introducción

- ▶ Para detectar la forma de Ω muchas veces alcanza considerar valores bajos de d (al menos en dimensión $p = 2, 3$).

Introducción

- ▶ Para detectar la forma de Ω muchas veces alcanza considerar valores bajos de d (al menos en dimensión $p = 2, 3$).
- ▶ Algunos ejemplos de aplicaciones de esta herramienta son:
 - ▶ Detección de anomalías
 - ▶ Estimación de densidad
 - ▶ Aprendizaje de variedades

Introducción

- ▶ Para detectar la forma de Ω muchas veces alcanza considerar valores bajos de d (al menos en dimensión $p = 2, 3$).
- ▶ Algunos ejemplos de aplicaciones de esta herramienta son:
 - ▶ Detección de anomalías
 - ▶ Estimación de densidad
 - ▶ Aprendizaje de variedades
- ▶ Calcular el polinomio p_d requiere solamente una iteración sobre los datos y no necesita optimización.

Introducción

- ▶ Para detectar la forma de Ω muchas veces alcanza considerar valores bajos de d (al menos en dimensión $p = 2, 3$).
- ▶ Algunos ejemplos de aplicaciones de esta herramienta son:
 - ▶ Detección de anomalías
 - ▶ Estimación de densidad
 - ▶ Aprendizaje de variedades
- ▶ Calcular el polinomio p_d requiere solamente una iteración sobre los datos y no necesita optimización.
- ▶ El cuello de botella computacional está en invertir la matriz.

Notación

Notamos $\mathbb{R}[x] = \mathbb{R}[x_1, \dots, x_p]$ al espacio de polinomios reales en p -variables. Para $\alpha = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^p$, notamos x^α al monomio $x_1^{\alpha_1} \dots x_p^{\alpha_p}$, los cuales forman una base de $\mathbb{R}[x]$.

Para $d \in \mathbb{N}$, notamos $\mathbb{R}[x]_d \subset \mathbb{R}[x]$ al subespacio de polinomios reales generado por los monomios de grado (total) menor o igual a d , ie, x^α tal que $\|\alpha\|_1 \leq d$. Tenemos que $\dim \mathbb{R}[x]_d = s(d, p) = \binom{d+p}{p}$.

Función de Christoffel

Sea $\Omega \in \mathbb{R}^p$ un compacto con interior no vacío y sea μ una medida de probabilidad soportada en Ω con densidad f respecto a la medida de Lebesgue.

Función de Christoffel

Sea $\Omega \in \mathbb{R}^P$ un compacto con interior no vacío y sea μ una medida de probabilidad soportada en Ω con densidad f respecto a la medida de Lebesgue.

Tenemos un producto interno $\langle \cdot, \cdot \rangle_\mu$ en el espacio de polinomios $\mathbb{R}[x]$ definido como

$$\langle p, q \rangle_\mu := \int p(x)q(x)d\mu(x).$$

Cualquier subespacio $V \subset \mathbb{R}[x]$ de dimensión finita de dimensión d tiene una base ortonormal $\{p_1, \dots, p_d\}$.

Función de Christoffel

Consideremos $(P_\alpha)_{\alpha \in \mathbb{N}^p}$ una familia de polinomios ortonormales con respecto μ . Se define el núcleo de Christoffel-Darboux $K_d^p : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$, asociado a μ , como:

$$K_d^p(x, y) := \sum_{\|\alpha\|_1 \leq d} P_\alpha(x) P_\alpha(y), \quad \forall x, y \in \mathbb{R}^p, d \in \mathbb{N}.$$

Función de Christoffel

Consideremos $(P_\alpha)_{\alpha \in \mathbb{N}^p}$ una familia de polinomios ortonormales con respecto μ . Se define el núcleo de Christoffel-Darboux $K_d^p : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$, asociado a μ , como:

$$K_d^p(x, y) := \sum_{\|\alpha\|_1 \leq d} P_\alpha(x) P_\alpha(y), \quad \forall x, y \in \mathbb{R}^p, d \in \mathbb{N}.$$

El polinomio de Christoffel $q_d^\mu : \mathbb{R}^p \rightarrow \mathbb{R}_+$ está definido por:

$$q_d^\mu(\mathbf{x}) := \sum_{\|\alpha\|_1 \leq d} P_\alpha(\mathbf{x})^2$$

La función de Christoffel $\Lambda_d^\mu : \mathbb{R}^p \rightarrow \mathbb{R}_+$ está definida por:

$$\Lambda_d^\mu(\mathbf{x}) := q_d^\mu(\mathbf{x})^{-1}$$

Matriz de Momentos

- ▶ La matriz de momentos $M_d(\mu)$ tiene entradas:

$$M_d(\mu)(\alpha, \beta) := \int_{\Omega} x^{\alpha+\beta} d\mu = \mu_{\alpha+\beta}.$$

- ▶ Ejemplo en dimensión 2:

$$M_2(\mu) := \begin{pmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ 1 & \mu_{00} & \mu_{10} & \mu_{01} & \mu_{20} & \mu_{11} & \mu_{02} \\ x_1 & \mu_{10} & \mu_{20} & \mu_{11} & \mu_{30} & \mu_{21} & \mu_{12} \\ x_2 & \mu_{01} & \mu_{11} & \mu_{02} & \mu_{21} & \mu_{11} & \mu_{02} \\ x_1^2 & \mu_{20} & \mu_{30} & \mu_{21} & \mu_{40} & \mu_{31} & \mu_{22} \\ x_1x_2 & \mu_{11} & \mu_{21} & \mu_{12} & \mu_{31} & \mu_{22} & \mu_{13} \\ x_2^2 & \mu_{02} & \mu_{12} & \mu_{03} & \mu_{22} & \mu_{13} & \mu_{14} \end{pmatrix}$$

Función de Christoffel

Theorem

La función de Christoffel también puede definirse como

$$\Lambda_d^\mu(\mathbf{x})^{-1} = v_d(\mathbf{x}) M_d(\mu)^{-1} v_d(\mathbf{x})$$

y verifica la propiedad variacional:

$$\Lambda_d^\mu(\mathbf{x}) = \min_{P \in \mathbb{R}[\mathbf{x}]_d} \left\{ \int_{\Omega} P^2(z) d\mu(z) : P(\mathbf{x}) = 1 \right\}.$$

Función de Christoffel

Theorem

La función de Christoffel también puede definirse como

$$\Lambda_d^\mu(\mathbf{x})^{-1} = v_d(\mathbf{x}) M_d(\mu)^{-1} v_d(\mathbf{x})$$

y verifica la propiedad variacional:

$$\Lambda_d^\mu(\mathbf{x}) = \min_{P \in \mathbb{R}[\mathbf{x}]_d} \left\{ \int_{\Omega} P^2(z) d\mu(z) : P(\mathbf{x}) = 1 \right\}.$$

La demostración sigue de la propiedad reproductora:

$$\int_{\Omega} K_d^p(x, y) p(y) d\mu(y) = p(x), \quad \forall p \in \mathbb{R}[\mathbf{x}]_d.$$

Propiedades de la función de Christoffel

Theorem

Sea μ una medida con soporte compacto Ω de interior no vacío.

Entonces:

- ▶ Para todo $\mathbf{x} \in \text{int}(\Omega) : \Lambda_d(\mathbf{x}) = O(d^p)$
- ▶ Para todo $\mathbf{x} \in \text{int}(\mathbb{R}^p \setminus \Omega) : \Lambda_d(\mathbf{x}) = O(e^{\alpha d})$ para algún $\alpha > 0$.

En particular, cuando $d \rightarrow \infty$, $d^p \Lambda_d^\mu(\mathbf{x}) \rightarrow 0$ de forma exponencial cuando $\mathbf{x} \notin \Omega$.

Propiedades de la función de Christoffel

Sea μ una medida con soporte compacto Ω y una densidad continua f .

Bajo algunas hipótesis sobre Ω y μ se puede probar que

$$\lim_{d \rightarrow \infty} s(d) \Lambda_d^\mu(\mathbf{x}) = \frac{f(\mathbf{x})}{\omega(\mathbf{x})}$$

donde ω es la densidad de la medida de equilibrio asociada a Ω . La medida de equilibrio μ_Ω minimiza un funcional de energía, de manera similar a cómo la medida de equilibrio clásica en una variable compleja minimiza la energía potencial logarítmica. En general la medida de equilibrio es desconocida salvo algunos casos especiales (segmentos, circunferencias, esferas).

Núcleos reproductores

Sea (X, μ) un espacio medible. Sea $\mathcal{F} \subset L^2(X)$ un subespacio de dimensión finita con producto interno $\langle f, g \rangle = \int fg \, d\mu$.

Núcleos reproductores

Sea (X, μ) un espacio medible. Sea $\mathcal{F} \subset L^2(X)$ un subespacio de dimensión finita con producto interno $\langle f, g \rangle = \int fg \, d\mu$.

Dado $x \in X$, consideremos un funcional lineal $l_x : \mathcal{F} \rightarrow \mathbb{R}$. Por ejemplo:

- ▶ $l_x(f) = f(x)$.
- ▶ $l_x(f) = \int_{B(x, \epsilon)} f(z) d\nu$, para algún $\epsilon \in \mathbb{R}^+$ y una medida ν no negativa.

Núcleos reproductores

Sea (X, μ) un espacio medible. Sea $\mathcal{F} \subset L^2(X)$ un subespacio de dimensión finita con producto interno $\langle f, g \rangle = \int fg d\mu$.

Dado $x \in X$, consideremos un funcional lineal $l_x : \mathcal{F} \rightarrow \mathbb{R}$. Por ejemplo:

- ▶ $l_x(f) = f(x)$.
- ▶ $l_x(f) = \int_{B(x, \epsilon)} f(z) d\nu$, para algún $\epsilon \in \mathbb{R}^+$ y una medida ν no negativa.

Sea $e_1, \dots, e_N \in \mathcal{F}$ una base ortonormal. Definimos el siguiente núcleo reproductor $K : X \times X \rightarrow \mathbb{R}$ como

$$K(x, y) := \sum_{i=1}^N l_x(e_i) e_i(y)$$

que es un elemento de \mathcal{F} si lo vemos como función de x .

Núcleos reproductores

Sea (X, μ) un espacio medible. Sea $\mathcal{F} \subset L^2(X)$ un subespacio de dimensión finita con producto interno $\langle f, g \rangle = \int fg d\mu$.

Dado $x \in X$, consideremos un funcional lineal $l_x : \mathcal{F} \rightarrow \mathbb{R}$. Por ejemplo:

- ▶ $l_x(f) = f(x)$.
- ▶ $l_x(f) = \int_{B(x, \epsilon)} f(z) d\nu$, para algún $\epsilon \in \mathbb{R}^+$ y una medida ν no negativa.

Sea $e_1, \dots, e_N \in \mathcal{F}$ una base ortonormal. Definimos el siguiente núcleo reproductor $K : X \times X \rightarrow \mathbb{R}$ como

$$K(x, y) := \sum_{i=1}^N l_x(e_i) e_i(y)$$

que es un elemento de \mathcal{F} si lo vemos como función de x . Este núcleo verifica la propiedad reproductora:

$$\int_X K(x, y) f(y) d\mu(y) = l_x(f)$$

Núcleos reproductores

Veamos qué podemos decir de la norma de l_x .

Núcleos reproductores

Veamos qué podemos decir de la norma de l_x .

$$\begin{aligned}\|l_x\|_{op} &= \sup_{\|h\|_2=1} \left\{ \int_X K(x, y) h(y) d\mu(y) \right\} \\ &= \sup_{\|h\|_2=1} \left\{ \left(\int_X K(x, y)^2 d\mu(y) \right)^{\frac{1}{2}} \left(\int_X h(y)^2 d\mu(y) \right)^{\frac{1}{2}} \right\} \\ &\leq (l_x(K(x, y)))^{\frac{1}{2}} (1)^{\frac{1}{2}} = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}}\end{aligned}$$

Núcleos reproductores

Veamos qué podemos decir de la norma de l_x .

$$\begin{aligned}\|l_x\|_{op} &= \sup_{\|h\|_2=1} \left\{ \int_X K(x, y) h(y) d\mu(y) \right\} \\ &= \sup_{\|h\|_2=1} \left\{ \left(\int_X K(x, y)^2 d\mu(y) \right)^{\frac{1}{2}} \left(\int_X h(y)^2 d\mu(y) \right)^{\frac{1}{2}} \right\} \\ &\leq (l_x(K(x, y)))^{\frac{1}{2}} (1)^{\frac{1}{2}} = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}}\end{aligned}$$

Y la igualdad se alcanza cuando $h(\cdot) = \lambda K(x, \cdot)$.

Núcleos reproductores

Veamos qué podemos decir de la norma de l_x .

$$\begin{aligned}\|l_x\|_{op} &= \sup_{\|h\|_2=1} \left\{ \int_X K(x, y) h(y) d\mu(y) \right\} \\ &= \sup_{\|h\|_2=1} \left\{ \left(\int_X K(x, y)^2 d\mu(y) \right)^{\frac{1}{2}} \left(\int_X h(y)^2 d\mu(y) \right)^{\frac{1}{2}} \right\} \\ &\leq (l_x(K(x, y)))^{\frac{1}{2}} (1)^{\frac{1}{2}} = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}}\end{aligned}$$

Y la igualdad se alcanza cuando $h(\cdot) = \lambda K(x, \cdot)$.

Tomando norma esto es $\|h\|_2^2 = \lambda^2 l_x(K(x, y))$.

Núcleos reproductores

Veamos qué podemos decir de la norma de l_x .

$$\begin{aligned}\|l_x\|_{op} &= \sup_{\|h\|_2=1} \left\{ \int_X K(x, y) h(y) d\mu(y) \right\} \\ &= \sup_{\|h\|_2=1} \left\{ \left(\int_X K(x, y)^2 d\mu(y) \right)^{\frac{1}{2}} \left(\int_X h(y)^2 d\mu(y) \right)^{\frac{1}{2}} \right\} \\ &\leq (l_x(K(x, y)))^{\frac{1}{2}} (1)^{\frac{1}{2}} = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}}\end{aligned}$$

Y la igualdad se alcanza cuando $h(\cdot) = \lambda K(x, \cdot)$.

Tomando norma esto es $\|h\|_2^2 = \lambda^2 l_x(K(x, y))$.

Entonces $\lambda = \frac{\|h\|_2}{(l_x(K(x, y)))^{\frac{1}{2}}}$, y obtenemos que la igualdad se da

para $h^* = \frac{K(x, y)}{(l_x(K(x, y)))^{\frac{1}{2}}}$.

Núcleos reproductores

Veamos ahora que sucede con la siguiente función $l_x \in L^2(X)$

$$\begin{aligned} l_x(f) &= \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) d\nu(z) = \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) \frac{d\nu}{d\mu}(z) d\mu(z) \\ &= \left\langle f(z), \frac{1}{\tau_\epsilon} \chi_{B(x,\epsilon)} \frac{d\nu}{d\mu} \right\rangle \end{aligned}$$

Núcleos reproductores

Veamos ahora que sucede con la siguiente función $l_x \in L^2(X)$

$$\begin{aligned}l_x(f) &= \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) d\nu(z) = \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) \frac{d\nu}{d\mu}(z) d\mu(z) \\ &= \left\langle f(z), \frac{1}{\tau_\epsilon} \chi_{B(x,\epsilon)} \frac{d\nu}{d\mu} \right\rangle\end{aligned}$$

Y su norma verifica $\|l_x\|^2 = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}} = \left\| \frac{1}{\tau_\epsilon} \chi_{B(x,\epsilon)} \frac{d\nu}{d\mu} \right\|^2$.

Núcleos reproductores

Veamos ahora que sucede con la siguiente función $l_x \in L^2(X)$

$$\begin{aligned}l_x(f) &= \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) d\nu(z) = \frac{1}{\tau_\epsilon} \int_X \chi_{B(x,\epsilon)}(z) f(z) \frac{d\nu}{d\mu}(z) d\mu(z) \\ &= \left\langle f(z), \frac{1}{\tau_\epsilon} \chi_{B(x,\epsilon)} \frac{d\nu}{d\mu} \right\rangle\end{aligned}$$

Y su norma verifica $\|l_x\|^2 = \left(\sum_{i=1}^N l_x(e_i)^2 \right)^{\frac{1}{2}} = \left\| \frac{\frac{1}{\tau_\epsilon} \chi_{B(x,\epsilon)}}{\frac{d\mu}{d\nu}} \right\|^2$.

Si $\kappa > \frac{d\mu}{d\nu} > \gamma > 0$ y es continua

$$\|l_x\|^2 = \int_X \frac{\frac{1}{\tau_\epsilon^2} \chi_{B(x,\epsilon)}(z)}{\frac{d\mu}{d\nu}(z)^2} d\mu(z) = \int_X \frac{\frac{1}{\tau_\epsilon^2} \chi_{B(x,\epsilon)}(z)}{\frac{d\mu}{d\nu}(z)} d\nu(z).$$

Función de Christoffel Regularizada

En el artículo “A modified Christoffel function and its asymptotic properties”, Laserre introduce una versión regularizada de la función de Christoffel, que notamos $\tilde{\Lambda}_n^\mu : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}_+$ y se define como:

$$\tilde{\Lambda}_n^\mu(\boldsymbol{\xi}, \varepsilon) := \left\{ \inf_{p \in \mathbb{R}[\mathbf{x}]_n} \left\{ \int p^2 d\mu : \int_{\mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)} p(\mathbf{x}) \frac{d\mathbf{x}}{\varepsilon^d} = 1 \right\} \right\},$$

para todo $(\boldsymbol{\xi}, \varepsilon) \in \mathbb{R}^d \times \mathbb{R}_+$ y $n \in \mathbb{N}$, donde

$\mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon) := \{\mathbf{x} : \|\mathbf{x} - \boldsymbol{\xi}\|_\infty < \varepsilon/2\}$ (con volumen de Lebesgue ε^d).

Función de Christoffel Regularizada

Función de Christoffel Regularizada

- ▶ $\tilde{\Lambda}_d^\mu$ es un polinomio suma de cuadrados con forma explícita:

$$\tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon)^{-1} = \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon)^T \mathbf{M}_d(\mu)^{-1} \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon),$$

donde,

$$\tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon) := \int_{\mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)} \mathbf{v}_d(\mathbf{x}) \frac{d\mathbf{x}}{\varepsilon^d}.$$

Función de Christoffel Regularizada

- ▶ $\tilde{\Lambda}_d^\mu$ es un polinomio suma de cuadrados con forma explícita:

$$\tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon)^{-1} = \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon)^T \mathbf{M}_d(\mu)^{-1} \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon),$$

donde,

$$\tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon) := \int_{\mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)} \mathbf{v}_d(\mathbf{x}) \frac{d\mathbf{x}}{\varepsilon^d}.$$

- ▶ Con $\varepsilon > 0$ fijo, y suponiendo que f es continua y $f > 0$ en Ω , se tiene

$$\lim_{d \rightarrow \infty} \varepsilon^{-p} \tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon) = f(\zeta_\varepsilon)$$

para algún $\zeta_\varepsilon \in \mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)$, y por lo tanto $f(\zeta_\varepsilon) \approx f(\boldsymbol{\xi})$ si ε es lo suficientemente chico.

Función de Christoffel Regularizada

- ▶ $\tilde{\Lambda}_d^\mu$ es un polinomio suma de cuadrados con forma explícita:

$$\tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon)^{-1} = \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon)^T \mathbf{M}_d(\mu)^{-1} \tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon),$$

donde,

$$\tilde{\mathbf{v}}_d(\boldsymbol{\xi}, \varepsilon) := \int_{\mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)} \mathbf{v}_d(\mathbf{x}) \frac{d\mathbf{x}}{\varepsilon^d}.$$

- ▶ Con $\varepsilon > 0$ fijo, y suponiendo que f es continua y $f > 0$ en Ω , se tiene

$$\lim_{d \rightarrow \infty} \varepsilon^{-p} \tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon) = f(\zeta_\varepsilon)$$

para algún $\zeta_\varepsilon \in \mathbf{B}_\infty(\boldsymbol{\xi}, \varepsilon)$, y por lo tanto $f(\zeta_\varepsilon) \approx f(\boldsymbol{\xi})$ si ε es lo suficientemente chico.

- ▶ El crecimiento con d de $\tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, \varepsilon)^{-1}$ ($\varepsilon > 0$ fijo) y $\tilde{\Lambda}_d^\mu(\boldsymbol{\xi}, 1/d)^{-1}$, es a lo sumo polinomial si $\boldsymbol{\xi} \in \text{int}(\Omega)$ y al menos exponencial si $\boldsymbol{\xi} \in \text{int}(\mathbb{R}^p \setminus \Omega)$.

Función de Christoffel esparcificada

Recientemente, en el “artículo A sparsified Christoffel function for high-dimensional inference.”, Lasserre y Slot presentan una nueva versión de la función de Christoffel con busca aproximar el caso de medidas producto a la clase de medidas que exhiben una estructura de producto condicional representada por un modelo gráfico G .

Función de Christoffel esparcificada

Recientemente, en el “artículo A sparsified Christoffel function for high-dimensional inference.”, Lasserre y Slot presentan una nueva versión de la función de Christoffel con busca aproximar el caso de medidas producto a la clase de medidas que exhiben una estructura de producto condicional representada por un modelo gráfico G .

$$\Gamma_d^\mu(\mathbf{x}) = \Gamma_d^{\mu, G, J}(\mathbf{x}) := \frac{\prod_{v \in V} \Lambda_{\otimes d}^{\mu_{A_v}}(\mathbf{x}_{A_v})}{\prod_{e \in E} \Lambda_{\otimes d}^{\mu_{B_e}}(\mathbf{x}_{B_e})}, \quad \mathbf{x} \in \mathbb{R}^p.$$

Las medidas $\mu_{A_v}, \mu_{B_e}, v \in V, e \in E$ son las marginales de μ en las variables indexadas por subconjuntos $A_v, B_e \subseteq [p]$, que corresponden a los vértices y aristas del árbol de “unión” $J = (V, E)$ asociado con G . Los vértices de este árbol representan cliques maximales en una completación cordal de G , y sus aristas representan intersecciones entre estos cliques.

Propiedades de la función de Christoffel esparcificada

Estas funciones de Christoffel racionales verifican las siguientes propiedades:

Propiedades de la función de Christoffel esparcificada

Estas funciones de Christoffel racionales verifican las siguientes propiedades:

- ▶ Los factores de Γ_d^μ son polinomios de Christoffel asociados con marginales de μ . En particular, son sumas de cuadrados de polinomios de grado d en cada coordenada.

Propiedades de la función de Christoffel esparcificada

Estas funciones de Christoffel racionales verifican las siguientes propiedades:

- ▶ Los factores de Γ_d^μ son polinomios de Christoffel asociados con marginales de μ . En particular, son sumas de cuadrados de polinomios de grado d en cada coordenada.
- ▶ Los coeficientes de cada factor de Γ_d^μ pueden calcularse invirtiendo una matriz de tamaño a lo sumo $(d+1)^{\text{cl}(J)}$, cuyas entradas son momentos de μ , donde $\text{cl}(J) =$ denota el tamaño del clique más grande en el árbol de “unión” J . Si J se elige de forma óptima, entonces $\text{cl}(J) = \tau(G) + 1$. En total, hay a lo sumo $2p$ de estos factores en la definición de la función.

Propiedades de la función de Christoffel esparcificada

Estas funciones de Christoffel racionales verifican las siguientes propiedades:

- ▶ Los factores de Γ_d^μ son polinomios de Christoffel asociados con marginales de μ . En particular, son sumas de cuadrados de polinomios de grado d en cada coordenada.
- ▶ Los coeficientes de cada factor de Γ_d^μ pueden calcularse invirtiendo una matriz de tamaño a lo sumo $(d+1)^{\text{cl}(J)}$, cuyas entradas son momentos de μ , donde $\text{cl}(J) =$ denota el tamaño del clique más grande en el árbol de “unión” J . Si J se elige de forma óptima, entonces $\text{cl}(J) = \tau(G) + 1$. En total, hay a lo sumo $2p$ de estos factores en la definición de la función.
- ▶ Bajo algunas hipótesis sobre μ , para casi todo $\mathbf{x} \in \mathbb{R}^p$, las funciones Γ_d^μ exhiben la dicotomía deseada en el soporte:
 - ▶ Para todo $\mathbf{x} \in \text{int}(\Omega) : \Gamma_d(\mathbf{x}) = O(d^p)$
 - ▶ Para todo $\mathbf{x} \in \text{int}(\mathbb{R}^p \setminus \Omega) : \Gamma_d(\mathbf{x}) = O(e^{\alpha d})$ para algún $\alpha > 0$.

Propiedades de la función de Christoffel esparcificada

Para la estimación de densidad de μ se obtienen resultados similares a los análogos. Bajo algunas condiciones de regularidad sobre Ω y μ , definen una versión regularizada $\tilde{\Gamma}_n^{\mu, G, J}$ de las funciones de Christoffel racionales y demuestran que esto recupera nuevamente la densidad:

$$\lim_{n \rightarrow \infty} \varepsilon^{-d} \cdot \tilde{\Gamma}_d^{\mu, G, J}(\mathbf{x}; \varepsilon)^{-1} = f(\mathbf{x}) + O(\varepsilon).$$

Bibliografía

- ▶ Lasserre, J. B., Pauwels, E., & Putinar, M. (2022). The Christoffel–Darboux Kernel for Data Analysis (Vol. 38). Cambridge University Press.
- ▶ Lasserre, J. B. (2023). A modified Christoffel function and its asymptotic properties. *Journal of Approximation Theory*, 295, 105955.
- ▶ Lasserre, J. B. (2024). The Christoffel function: Applications, connections and extensions. *Numerical Algebra, Control and Optimization*, 0-0.
- ▶ Lasserre, J. B., & Slot, L. (2024). A sparsified Christoffel function for high-dimensional inference. *arXiv preprint arXiv:2409.15965*.

Gracias!