Kernel relaxation for space-filling design[1]

Luc Pronzato

Laboratoire I3S, Université Côte d'Azur, CNRS, France

— joint work with Anatoly Zhigljavsky, Cardiff University —

July 2023

[1]Partially funded by project ANR INDEX (ANR-18-CE91-0007)

1 Three space-filling criteria

$\mathscr{X}$ = a compact subset of $\mathbb{R}^d$, $\mathscr{X} = \mathrm{cl}(\mathrm{int}(\mathscr{X}))$; f: $\mathscr{X} \to \mathbb{R}$
use pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, \ldots, n$, to approximate or integrate f over $\mathscr{X}$

With little prior information about f
➞ observe "everywhere"
➞ choose a design $\mathbf{X}_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ space-filling in $\mathscr{X}$

1 Three space-filling criteria

$\mathscr{X}$ = a compact subset of $\mathbb{R}^d$, $\mathscr{X} = \mathrm{cl}(\mathrm{int}(\mathscr{X}))$; f: $\mathscr{X} \to \mathbb{R}$
use pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, \ldots, n$, to approximate or integrate f over $\mathscr{X}$

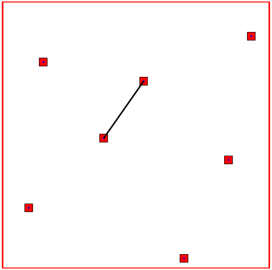
With little prior information about f
➞ observe "everywhere"
➞ choose a design $\mathbf{X}_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ space-filling in $\mathscr{X}$

We shall consider three "classical" criteria:

1. Packing radius
2. Covering radius
3. L_s-quantisation error

1/ Packing radius: maximise $\mathrm{PR}(\mathbf{X}_n) \triangleq \frac{1}{2} \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$

$$\mathrm{PR}(\mathbf{X}_n) = \text{separation radius} = \frac{1}{2}\ \text{Maximin distance criterion}$$

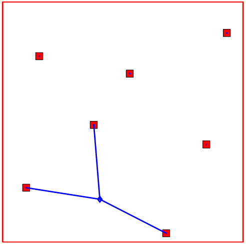


→ can often be related to numerical stability issues

→ easy to compute, but pushes points to the boundary of $\mathscr{X}$

2/ Covering radius: minimise $\mathrm{CR}(\mathbf{X}_n) = \mathrm{CR}_{\mathscr{X}}(\mathbf{X}_n) \triangleq \max_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|$

$\mathrm{CR}(\mathbf{X}_n)$ = fill distance = dispersion = miniMax distance criterion



→ we are never far from a design point

→ more difficult to compute

→ appears in bounds on approximation error (Narcowich et al., 2005; Schaback and Wendland, 2006)

3/ L_s-mean quantisation error:

μ a prob. measure on $\mathscr{X}$ (equiv. to Lebesgue measure Λ)

$$\text{minimise } E_{s,\mu}(\mathbf{X}_n) \triangleq \left[\int_{\mathscr{X}} \left[\underbrace{\min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|}_{=d(\mathbf{x},\mathbf{X}_n)} \right]^s \mathrm{d}\mu(\mathbf{x}) \right]^{1/s}, \; s > 0$$

3/ L_s-mean quantisation error:

μ a prob. measure on $\mathscr{X}$ (equiv. to Lebesgue measure Λ)

$$\text{minimise } E_{s,\mu}(\mathbf{X}_n) \triangleq \left[\int_{\mathscr{X}} \Big[\underbrace{\min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|}_{=d(\mathbf{x},\mathbf{X}_n)} \Big]^s \mathrm{d}\mu(\mathbf{x}) \right]^{1/s}, \; s > 0$$

Optimal $\mathbf{X}_n^* =$ n-optimal set of centers,
a Voronoi partition of $\mathbb{R}^d$ w.r.t. $\mathbf{X}_n^* =$ n-optimal quantiser;
see (Graf and Luschgy, 2000)

3/ L_s-mean quantisation error:

μ a prob. measure on $\mathscr{X}$ (equiv. to Lebesgue measure Λ)

$$\text{minimise } E_{s,\mu}(\mathbf{X}_n) \triangleq \left[\int_{\mathscr{X}} \Big[\underbrace{\min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|}_{=d(\mathbf{x},\mathbf{X}_n)} \Big]^s \, \mathrm{d}\mu(\mathbf{x}) \right]^{1/s}, \; s > 0$$

Optimal $\mathbf{X}_n^* =$ n-optimal set of centers,

a Voronoi partition of $\mathbb{R}^d$ w.r.t. $\mathbf{X}_n^* =$ n-optimal quantiser;

see (Graf and Luschgy, 2000)

$E_{s,\mu}(\mathbf{X}_n) \rightarrow$ bounds on worst-case approximation and integration errors

(Krieg and Sonnleitner, 2020)

$\rightarrow$ general bounds on integ. error for Lipschitz functions (Pagès, 1997)

3/ L_s-mean quantisation error:

μ a prob. measure on $\mathscr{X}$ (equiv. to Lebesgue measure Λ)

$$\text{minimise } E_{s,\mu}(\mathbf{X}_n) \triangleq \left[\int_{\mathscr{X}} \Big[\underbrace{\min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|}_{=d(\mathbf{x},\mathbf{X}_n)} \Big]^s \, \mathrm{d}\mu(\mathbf{x}) \right]^{1/s}, \quad s > 0$$

Optimal $\mathbf{X}_n^* =$ <u>n-optimal set of centers</u>,

a Voronoi partition of $\mathbb{R}^d$ w.r.t. $\mathbf{X}_n^* =$ <u>n-optimal quantiser</u>;

see (Graf and Luschgy, 2000)

$E_{s,\mu}(\mathbf{X}_n)$ → bounds on worst-case approximation and integration errors

(Krieg and Sonnleitner, 2020)

→ general bounds on integ. error for Lipschitz functions (Pagès, 1997)

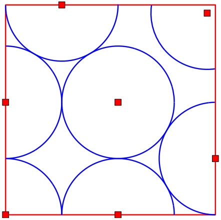
For $s \geq 1$, $E_{s,\mu}(\mathbf{X}_n^*) = \min_{\mathbf{X}_n} E_{s,\mu}(\mathbf{X}_n) = \inf_{\mu_n \in \mathscr{P}_n} W_s(\mu, \mu_n)$, with

- $\mathscr{P}_n =$ set of discrete probabilities on $\mathbb{R}^d$ supported on n points
- $W_s(\mu, \mu') = L_s$-Wasserstein (or Kantorovich) metric

  $= \inf_M \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \|\mathbf{x} - \mathbf{x}'\|^s \, \mathrm{d}M(\mathbf{x}, \mathbf{x}') \right)^{1/s}$ with M having marginals μ and μ'

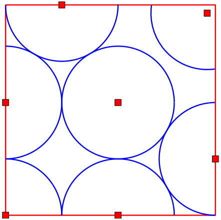
Example 1:

① Packing $d = 2, n = 7$
(radius=PR($\mathbf{X}_n$))

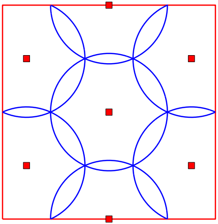


Example 1:

① Packing $d = 2, n = 7$
(radius=PR($\mathbf{X}_n$))

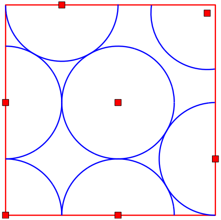


② Covering $d = 2, n = 7$
(radius=CR($\mathbf{X}_n$))

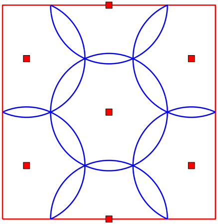


Example 1:

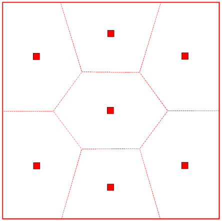
① Packing $d = 2, n = 7$ (radius=PR($\mathbf{X}_n$))



② Covering $d = 2, n = 7$ (radius=CR($\mathbf{X}_n$))

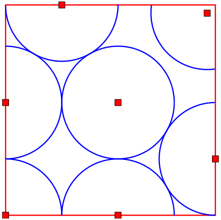


③ L_2-quantisation, $d = 2, n = 7$

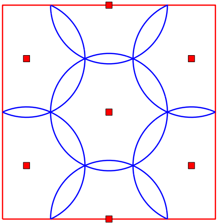


Example 1:

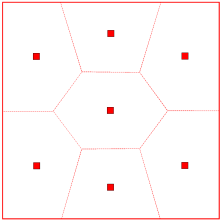
① Packing $d = 2, n = 7$ (radius=PR($\mathbf{X}_n$))



② Covering $d = 2, n = 7$ (radius=CR($\mathbf{X}_n$))



③ L_2-quantisation, $d = 2, n = 7$



→ Maximise PR($\mathbf{X}_n$), minimise CR($\mathbf{X}_n$) and $E_s(\mathbf{X}_n)$

2 Kernel relaxation

Some notation:

$\|\cdot\|$ is always the Euclidean norm in $\mathbb{R}^d$

For K a kernel on $\mathscr{X} \times \mathscr{X}$ ($\rightarrow$ RKHS $\mathcal{H}_K$ when K Pos. Def.),
ν a signed measure on $\mathscr{X}$, define:

$$\mathscr{E}_K(\nu) \triangleq \int_{\mathscr{X}^2} K(\mathbf{x},\mathbf{x}')\,\mathrm{d}\nu(\mathbf{x})\mathrm{d}\nu(\mathbf{x}') = \text{energy of } \nu$$

$$P_{K,\nu}(\mathbf{x}) \triangleq \int_{\mathscr{X}} K(\mathbf{x},\mathbf{x}')\,\mathrm{d}\nu(\mathbf{x}') = \text{potential of } \nu \text{ at } \mathbf{x}$$

$$[P_{K,\nu}(\cdot) = \text{kernel imbedding of } \nu \text{ into } \mathcal{H}_K]$$

Note that $\int_{\mathscr{X}} P_{K,\nu}(\mathbf{x})\,\mathrm{d}\nu(\mathbf{x}) = \mathscr{E}_K(\nu)$

Relaxation: example of the distance function

$d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$

Relaxation: example of the distance function $d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$

For $q > 0$, $d(\mathbf{x}, \mathbf{X}_n) = \left(\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right)^{-1/q}$

$$\frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \qquad \max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \qquad \leq \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}$$

Relaxation: example of the distance function $d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$

For $q > 0$, $d(\mathbf{x}, \mathbf{X}_n) = \left(\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right)^{-1/q}$

$$\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \quad \max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \quad \leq \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}$$

$$\left[\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q} \leq \left[\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q} \leq \left[\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q}$$

Relaxation: example of the distance function

$$d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$$

For $q > 0$, $d(\mathbf{x}, \mathbf{X}_n) = (\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q})^{-1/q}$

$$\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}$$

$$\left[\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q} \leq [\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}]^{-1/q} \leq \left[\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q}$$

$$n^{-1/q}\left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q} \leq d(\mathbf{x}, \mathbf{X}_n) \leq \left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q}$$

with $K_q(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^{-q}$ the Riesz kernel with parameter q

(K_q is singular, does not define a RKHS)

and ξ_n the empirical measure on $\mathbf{X}_n$

Relaxation: example of the distance function $d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$

For $q > 0$, $d(\mathbf{x}, \mathbf{X}_n) = \left(\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right)^{-1/q}$

$$\frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q} \leq \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}$$

$$\left[\sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q} \leq [\max_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^{-q}]^{-1/q} \leq \left[\frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x} - \mathbf{x}_i\|^{-q}\right]^{-1/q}$$

$$\boxed{n^{-1/q}} \left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q} \leq d(\mathbf{x}, \mathbf{X}_n) \leq \left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q}$$

with $K_q(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^{-q}$ the Riesz kernel with parameter q
(K_q is singular, does not define a RKHS)
and ξ_n the empirical measure on $\mathbf{X}_n$

For n fixed, $\left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q} \to d(\mathbf{x}, \mathbf{X}_n)$, uniformly in $\mathbf{x}$ and $\mathbf{X}_n$, as $q \to \infty$

Relaxation with other kernels than K_q:

$K(\mathbf{x},\mathbf{x}') = \psi(\|\mathbf{x}-\mathbf{x}'\|)$ (isotropic) with
ψ: $[0,\infty) \mapsto (0,+\infty]$ continuous, strictly decreasing ($\psi(0)=\infty$ is allowed)
→ inverse function $\psi^{\langle -1 \rangle}(u)$ defined for $u \in (0, \psi(0)]$ and strictly decreasing too

Relaxation with other kernels than K_q:

$K(\mathbf{x},\mathbf{x}') = \psi(\|\mathbf{x}-\mathbf{x}'\|)$ (isotropic) with
ψ: $[0,\infty) \mapsto (0,+\infty]$ continuous, strictly decreasing ($\psi(0)=\infty$ is allowed)
→ inverse function $\psi^{\langle -1 \rangle}(u)$ defined for $u \in (0,\psi(0)]$ and strictly decreasing too

In particular: γ-exponential family $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$
$\gamma = 1$ ➠ exponential (Matérn 1/2) kernel, $\gamma = 2$ ➠ Gaussian kernel
(K is not PD for $\gamma > 2$ but not necessarily important...)

Relaxation with other kernels than K_q:

$K(\mathbf{x},\mathbf{x}') = \psi(\|\mathbf{x}-\mathbf{x}'\|)$ (isotropic) with
ψ: $[0,\infty) \mapsto (0,+\infty]$ continuous, strictly decreasing ($\psi(0)=\infty$ is allowed)
➞ inverse function $\psi^{\langle -1\rangle}(u)$ defined for $u \in (0,\psi(0)]$ and strictly decreasing too

In particular: γ-exponential family $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$
$\gamma = 1$ ➠ exponential (Matérn 1/2) kernel, $\gamma = 2$ ➠ Gaussian kernel
(K is not PD for $\gamma > 2$ but not necessarily important...)

$$\boxed{-\gamma\,\ell^\gamma \log n} - \gamma\,\ell^\gamma \log P_{K_{e,\gamma,\ell},\xi_n}(\mathbf{x}) \leq d^\gamma(\mathbf{x},\mathbf{X}_n) \leq -\gamma\,\ell^\gamma \log P_{K_{e,\gamma,\ell},\xi_n}(\mathbf{x})\,.$$

For n fixed, $-\gamma\,\ell^\gamma \log P_{K_{e,\gamma,\ell},\xi_n}(\mathbf{x}) \to d^\gamma(\mathbf{x},\mathbf{X}_n)$, uniformly in $\mathbf{x}$ and $\mathbf{X}_n$, as $\ell \to 0$

3 Maximisation of the packing radius

ℓ_q-relaxation of $\mathrm{PR}(\mathbf{X}_n) = \frac{1}{2} \min_{i,j} \|\mathbf{x}_i - \mathbf{x}_j\|$

Maximise $\underbrace{\frac{1}{2}\left(\frac{2}{n(n-1)}\sum_{i<j}\|\mathbf{x}_i - \mathbf{x}_j\|^{-q}\right)^{-1/q}}$ for some $q > 0$ (large)

3 Maximisation of the packing radius

ℓ_q**-relaxation of** $\mathrm{PR}(\mathbf{X}_n) = \frac{1}{2} \min_{i,j} \|\mathbf{x}_i - \mathbf{x}_j\|$

Maximise $\underbrace{\frac{1}{2}\left(\frac{2}{n(n-1)} \sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}\right)^{-1/q}}_{\to\ \mathrm{PR}(\mathbf{X}_n)\ \text{as}\ q\to\infty}$ for some $q > 0$ (large)

equivalent to minimise $\mathscr{E}_{K_q}{}^{\neq}(\mathbf{X}_n) = \frac{2}{n(n-1)} \sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$ (= discrete energy)

3 Maximisation of the packing radius

ℓ_q**-relaxation of** $\mathrm{PR}(\mathbf{X}_n) = \frac{1}{2}\min_{i,j}\|\mathbf{x}_i - \mathbf{x}_j\|$

Maximise $\underbrace{\frac{1}{2}\left(\frac{2}{n(n-1)}\sum_{i<j}\|\mathbf{x}_i - \mathbf{x}_j\|^{-q}\right)^{-1/q}}_{\to\ \mathrm{PR}(\mathbf{X}_n)\ \text{as}\ q\to\infty}$ for some $q > 0$ (large)

equivalent to minimise $\mathscr{E}_{K_q}{}^{\neq}(\mathbf{X}_n) = \frac{2}{n(n-1)}\sum_{i<j}\|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$ (= discrete energy)

➞ continuous version:

PB1: minimise $\boxed{\int_{\mathscr{X}^2}\|\mathbf{x} - \mathbf{x}'\|^{-q}\,\mathrm{d}\xi(\mathbf{x})\mathrm{d}\xi(\mathbf{x}')} = \mathscr{E}_{K_q}(\xi)$ w.r.t. ξ

3 Maximisation of the packing radius

ℓ_q**-relaxation of** $\mathrm{PR}(\mathbf{X}_n) = \frac{1}{2}\min_{i,j}\|\mathbf{x}_i - \mathbf{x}_j\|$

Maximise $\underbrace{\frac{1}{2}\left(\frac{2}{n(n-1)}\sum_{i<j}\|\mathbf{x}_i - \mathbf{x}_j\|^{-q}\right)^{-1/q}}_{\to\ \mathrm{PR}(\mathbf{X}_n)\ \text{as}\ q\to\infty}$ for some $q > 0$ (large)

equivalent to minimise $\mathscr{E}_{K_q}{}^{\neq}(\mathbf{X}_n) = \frac{2}{n(n-1)}\sum_{i<j}\|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$ (= discrete energy)

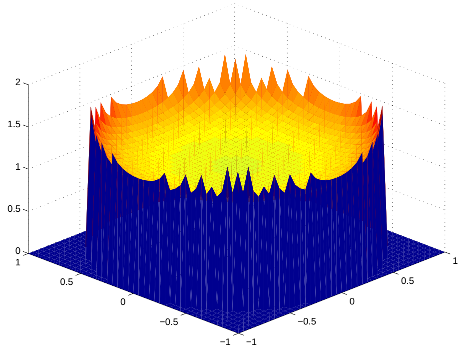
→ continuous version:

PB1: minimise $\boxed{\int_{\mathscr{X}^2}\|\mathbf{x} - \mathbf{x}'\|^{-q}\,\mathrm{d}\xi(\mathbf{x})\mathrm{d}\xi(\mathbf{x}')} = \mathscr{E}_{K_q}(\xi)$ w.r.t. ξ

For $q \geq d$, $\mathscr{E}_{K_q}(\xi)$ is infinite for any non zero ξ
For $0 < q < d$ there exists a minimum-energy probability measure $\xi^+_{K_q}$
When $d-2 < q < d$, $\xi^+_{K_q}$ has a density $\varphi^+_{K_q}$ in $\mathscr{X}$
and $P_{K_q,\xi^+_{K_q}}(\mathbf{x})$ is constant in $\mathscr{X}$; see e.g. (Landkof, 1972)

Example 2: $\mathscr{X} = \mathscr{B}_d(\mathbf{0}, \rho)$

$d-2 < q < d$: $\varphi^+_{K_q}(\mathbf{x}) = \frac{C}{(\rho^2 - \|\mathbf{x}\|^2)^{(d-q)/2}}$, $\mathbf{x} \in \mathscr{X}$ (Landkof, 1972, p. 163)



$d = 2,\ q = 7/4$

(when $d > 2$ and $2 < q \leq d-2$, $\xi^+_{K_q}$ is uniform on the sphere $\mathcal{S}_d(\mathbf{0}, \rho)$)

Other kernel relaxations? $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$

Other kernel relaxations? $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$

Maximise $\underbrace{-\frac{\gamma\ell^\gamma}{2^\gamma}\,\log\left[\frac{2}{n(n-1)}\sum_{i<j} K_{\gamma,\ell}(\mathbf{x}_i,\mathbf{x}_j)\right]}$ for some $\ell > 0$ (small)

Other kernel relaxations? $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$

Maximise $\underbrace{-\frac{\gamma\ell^\gamma}{2^\gamma} \log\left[\frac{2}{n(n-1)} \sum_{i<j} K_{\gamma,\ell}(\mathbf{x}_i,\mathbf{x}_j)\right]}_{\to \mathrm{PR}^\gamma(\mathbf{X}_n) \text{ as } \ell\to 0}$ for some $\ell > 0$ (small)

➞ continuous version:

PB1: minimise $\boxed{\int_{\mathscr{X}^2} K_{\gamma,\ell}(\mathbf{x},\mathbf{x}')\,\mathrm{d}\xi(\mathbf{x})\mathrm{d}\xi(\mathbf{x}')} = \mathscr{E}_{K_{\gamma,\ell}}(\xi)$

Other kernel relaxations? $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$, $\gamma > 0$

Maximise $\underbrace{-\frac{\gamma\ell^\gamma}{2^\gamma}\log\left[\frac{2}{n(n-1)}\sum_{i<j} K_{\gamma,\ell}(\mathbf{x}_i,\mathbf{x}_j)\right]}_{\to \mathrm{PR}^\gamma(\mathbf{X}_n) \text{ as } \ell\to 0}$ for some $\ell > 0$ (small)

➞ continuous version:

PB1: minimise $\boxed{\int_{\mathscr{X}^2} K_{\gamma,\ell}(\mathbf{x},\mathbf{x}')\,\mathrm{d}\xi(\mathbf{x})\mathrm{d}\xi(\mathbf{x}')} = \mathscr{E}_{K_{\gamma,\ell}}(\xi)$

There always exists a minimum-energy probability measure $\xi^+_{K_{\gamma,\ell}}$

For $K_{1,\ell}(\mathbf{x},\mathbf{x}') = \exp(-\|\mathbf{x}-\mathbf{x}'\|/\ell)$ (the exponential kernel):

- For $d=1$, $\mathscr{X}=[0,1]$, $\xi^+_{K_\theta} = 1/(1+2\,\ell)\Lambda(\mathscr{X}) + (\delta_0+\delta_1)\ell/(1+2\,\ell)$
  ➞ $\xi^+_{K_{1,\ell}}$ has full support and $P_{K_{1,\ell},\xi^+_{K_{1,\ell}}}(\mathbf{x})$ is constant on $\mathscr{X}$
- For $d>1$, not full support: for $\mathscr{X}=\mathscr{B}_2(\mathbf{0},1)$, $\xi^+_{K_{1,\ell}}$ has positive mass on the surface of the circle and a spherically symmetric density in the interior of $\mathscr{X}$, equal to zero close to the boundary

Minimisation of $\mathscr{E}_K(\xi)$: main properties

If $\boxed{\mathscr{E}_K(\cdot) \text{ is convex}}$ ($\gamma \in (0, 2]$ for $K_{\gamma,\ell}(\mathbf{x}, \mathbf{x}') = \exp[-\|\mathbf{x} - \mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$)

- $\rightarrow$ directional derivative (for bounded kernels)

$$
\begin{aligned}
F_{\mathscr{E}_K}(\xi, \delta_{\mathbf{x}}) &= \lim_{\alpha \to 0^+} \frac{\mathscr{E}_K[(1-\alpha)\xi + \alpha\delta_{\mathbf{x}}] - \mathscr{E}_K(\xi)}{\alpha} \\
&= 2\left[P_{K,\xi}(\mathbf{x}) - \mathscr{E}_K(\xi)\right]
\end{aligned}
$$

Minimisation of $\mathscr{E}_K(\xi)$: main properties

If $\boxed{\mathscr{E}_K(\cdot) \text{ is convex}}$ ($\gamma \in (0,2]$ for $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$)

- → directional derivative (for bounded kernels)

$$\begin{aligned} F_{\mathscr{E}_K}(\xi,\delta_{\mathbf{x}}) &= \lim_{\alpha\to 0^+} \frac{\mathscr{E}_K[(1-\alpha)\xi+\alpha\delta_{\mathbf{x}}]-\mathscr{E}_K(\xi)}{\alpha} \\ &= 2\,[P_{K,\xi}(\mathbf{x})-\mathscr{E}_K(\xi)] \end{aligned}$$

- N&S condition for optimality ("Equivalence Theorem" – **ET**):
  ξ_K^+ is optimal $\Leftrightarrow F_{\mathscr{E}_K}(\xi_K^+,\delta_{\mathbf{x}}) \geq 0$ for all $\mathbf{x} \in \mathscr{X}$
  $\Leftrightarrow \boxed{P_{K,\xi_K^+}(\mathbf{x}) - \mathscr{E}_K(\xi_K^+) \geq 0 \text{ for all } \mathbf{x} \in \mathscr{X}}$

Minimisation of $\mathscr{E}_K(\xi)$: main properties

If $\boxed{\mathscr{E}_K(\cdot) \text{ is convex}}$ ($\gamma \in (0,2]$ for $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$)

- ➞ directional derivative (for bounded kernels)

$$\begin{aligned} F_{\mathscr{E}_K}(\xi,\delta_\mathbf{x}) &= \lim_{\alpha\to 0^+} \frac{\mathscr{E}_K[(1-\alpha)\xi+\alpha\delta_\mathbf{x}]-\mathscr{E}_K(\xi)}{\alpha} \\ &= 2\,[P_{K,\xi}(\mathbf{x})-\mathscr{E}_K(\xi)] \end{aligned}$$

- N&S condition for optimality ("Equivalence Theorem" – **ET**):
  ξ_K^+ is optimal $\Leftrightarrow F_{\mathscr{E}_K}(\xi_K^+,\delta_\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathscr{X}$
  $\Leftrightarrow \boxed{P_{K,\xi_K^+}(\mathbf{x}) - \mathscr{E}_K(\xi_K^+) \geq 0 \text{ for all } \mathbf{x} \in \mathscr{X}}$

- As $\int_\mathscr{X} P_{K,\xi}(\mathbf{x})\,\mathrm{d}\xi(\mathbf{x}) = \mathscr{E}_K(\xi)$ for any ξ
  $$\boxed{P_{K,\xi_K^+}(\mathbf{x}) = \mathscr{E}_K(\xi_K^+) \text{ on } \mathrm{Supp}(\xi_K^+)}$$

Minimisation of $\mathscr{E}_K(\xi)$: main properties

If $\boxed{\mathscr{E}_K(\cdot) \text{ is convex}}$ ($\gamma \in (0,2]$ for $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$)

- → directional derivative (for bounded kernels)

$$F_{\mathscr{E}_K}(\xi,\delta_\mathbf{x}) = \lim_{\alpha\to 0^+} \frac{\mathscr{E}_K[(1-\alpha)\xi+\alpha\delta_\mathbf{x}]-\mathscr{E}_K(\xi)}{\alpha} = 2\,[P_{K,\xi}(\mathbf{x})-\mathscr{E}_K(\xi)]$$

- N&S condition for optimality ("Equivalence Theorem" – **ET**):
ξ_K^+ is optimal $\Leftrightarrow F_{\mathscr{E}_K}(\xi_K^+,\delta_\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathscr{X}$
$\Leftrightarrow \boxed{P_{K,\xi_K^+}(\mathbf{x}) - \mathscr{E}_K(\xi_K^+) \geq 0 \text{ for all } \mathbf{x} \in \mathscr{X}}$

- As $\int_{\mathscr{X}} P_{K,\xi}(\mathbf{x})\,\mathrm{d}\xi(\mathbf{x}) = \mathscr{E}_K(\xi)$ for any ξ (★)
$\boxed{P_{K,\xi_K^+}(\mathbf{x}) = \mathscr{E}_K(\xi_K^+) \text{ on } \mathrm{Supp}(\xi_K^+)}$

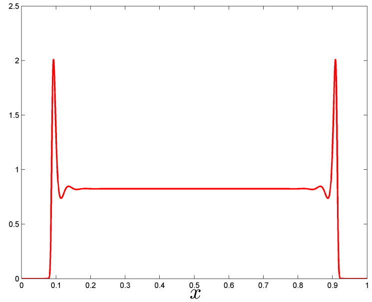
- (★) $\Rightarrow \min_{\mathbf{x}\in\mathscr{X}} P_{K,\xi}(\mathbf{x}) \leq \mathscr{E}_K(\xi)$ for any ξ, and thus
$\boxed{\xi_K^+ \text{ maximises } \min_{\mathbf{x}\in\mathscr{X}} P_{K,\xi}(\mathbf{x}) - \mathscr{E}_K(\xi)}$

Example 3: $\mathscr{X}=[0,1]$, $K(x,x')=(1+|x-x'|/\ell)\exp(-|x-x'|/\ell)$ (Matérn 3/2)

Example 3: $\mathscr{X} = [0,1]$, $K(x,x') = (1+|x-x'|/\ell)\exp(-|x-x'|/\ell)$ (Matérn 3/2)

→ Numerical determination of ξ_K^+ for $\ell = 1/(10\sqrt{3})$
(2000 points equally spaced in $[0,1]$, multiplicative algorithm)

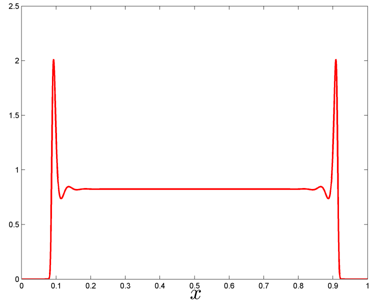
ξ_K^+ in $(0,1)$ (with $\xi_K^+((0,1)) \simeq 0.705$)



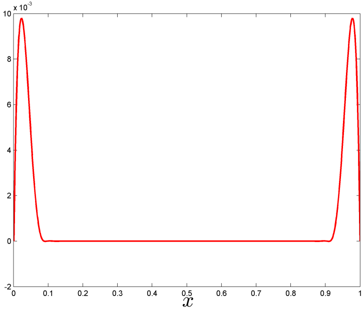
Example 3: $\mathscr{X} = [0,1]$, $K(x,x') = (1+|x-x'|/\ell)\exp(-|x-x'|/\ell)$ (Matérn 3/2)

→ Numerical determination of ξ_K^+ for $\ell = 1/(10\sqrt{3})$
(2000 points equally spaced in $[0,1]$, multiplicative algorithm)

ξ_K^+ in $(0,1)$ (with $\xi_K^+((0,1)) \simeq 0.705$)



$F_{\mathscr{E}_K}(\xi_K^+, \delta_x)$



$(\min_{x\in\mathscr{X}} F_{\mathscr{E}_K}(\xi_K^+, \delta_x) > -10^{-5})$

→ $F_{\mathscr{E}_K}(\xi_K^+, \delta_x) = 2\,[P_{K,\xi_K^+}(x) - \mathscr{E}_K(\xi_K^+)] \simeq 0$ on $\mathrm{Supp}(\xi^+) \subset \mathscr{X}$

Example 3: $\mathscr{X} = [0,1]$, $K(x,x') = (1 + |x - x'|/\ell) \exp(-|x - x'|/\ell)$ (Matérn 3/2)

→ Exact calculation of ξ_K^+ — see (P&Zh, 2023a)

$\ell \geq \ell_0 \simeq 0.3980$	$\xi_K^+ = (\delta_0 + \delta_1)/2$
$\ell_0 > \ell \geq \ell_1 \simeq 0.3180$	$\xi_K^+ = m_0(\delta_0 + \delta_1) + (1 - 2\,m_0)\delta_{1/2}$ (with $m_0 = m_0(\ell) \ldots$)
$\ell_1 > \ell$	$\xi_K^+ = m_0'(\delta_0 + \delta_1) + m_a(\delta_a + \delta_{1-a}) + \alpha\Lambda([a, 1-a])$ (with $m_0' = m_0'(\ell)$, $m_a = m_a(\ell)$, $\alpha = \alpha(\ell)$, $a = a(\ell)$)

Example 3: $\mathscr{X} = [0,1]$, $K(x,x') = (1+|x-x'|/\ell)\exp(-|x-x'|/\ell)$ (Matérn 3/2)

→ Exact calculation of ξ_K^+ — see (P&Zh, 2023a)

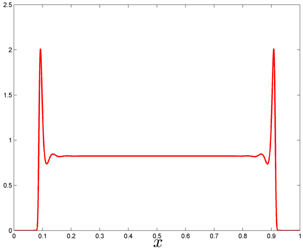
$\ell \geq \ell_0 \simeq 0.3980 \quad \xi_K^+ = (\delta_0 + \delta_1)/2$

$\ell_0 > \ell \geq \ell_1 \simeq 0.3180 \quad \xi_K^+ = m_0(\delta_0 + \delta_1) + (1 - 2\,m_0)\delta_{1/2}$ (with $m_0 = m_0(\ell)\ldots$)

$\ell_1 > \ell \quad \xi_K^+ = m_0'(\delta_0 + \delta_1) + m_a(\delta_a + \delta_{1-a}) + \alpha\Lambda([a, 1-a])$ (with $m_0' = m_0'(\ell)$, $m_a = m_a(\ell)$, $\alpha = \alpha(\ell)$, $a = a(\ell)$)

→ For $\ell = 1/(10\sqrt{3})$, $a \simeq 0.0908$ and $\xi_K^+((0,1)) \simeq 0.6913$

$\xi_K^+((0,1)) \simeq 0.705$



Discrete case: $\mathbf{X}_n \to$ empirical measure ξ_n

$\mathscr{E}_K(\xi_n) = \mathrm{var}(\widehat{\theta}^n_{OLSE})$ **in** $Y_\mathbf{x} = \theta + Z_\mathbf{x}$ **with** $\mathsf{E}\{Z_\mathbf{x}\} = 0$, $\mathsf{E}\{Z_\mathbf{x} Z_{\mathbf{x}'}\} = K(\mathbf{x}, \mathbf{x}')$

Discrete case: $\mathbf{X}_n$ ➞ empirical measure ξ_n

$\mathscr{E}_K(\xi_n) = \mathrm{var}(\widehat{\theta}^n_{OLSE})$ **in** $Y_\mathbf{x} = \theta + Z_\mathbf{x}$ **with** $\mathrm{E}\{Z_\mathbf{x}\} = 0$, $\mathrm{E}\{Z_\mathbf{x} Z_{\mathbf{x}'}\} = K(\mathbf{x}, \mathbf{x}')$

Greedy minimisation of $\mathscr{E}_K^{\neq}(\mathbf{X}_n)$, or **Conditional gradient descent for $\mathscr{E}_K(\xi)$**
(= Frank-Wolfe = Wynn's vertex-direction alg.)

(P&Zh, 2023b): if

$q = q_n$ is such that $q_n / \log n \to \infty$ in $K_q(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^{-q}$, or
$\ell = \ell_n$ is such that $\ell_n n^{1/d} (\log n)^{1/\gamma} \to 0$ in $K_{\gamma,\ell}(\mathbf{x}, \mathbf{x}')$,

then $\liminf_{n\to\infty} \mathrm{PR}(\mathbf{X}_n)/\,\mathrm{PR}^*_n \geq 1/2$ and $\limsup_{n\to\infty} \mathrm{CR}(\mathbf{X}_n)/\,\mathrm{CR}^*_n \leq 2$

Discrete case: $\mathbf{X}_n$ ➞ empirical measure ξ_n

$\mathscr{E}_K(\xi_n) = \mathrm{var}(\widehat{\theta}^n_{OLSE})$ **in** $Y_\mathbf{x} = \theta + Z_\mathbf{x}$ **with** $\mathrm{E}\{Z_\mathbf{x}\} = 0$, $\mathrm{E}\{Z_\mathbf{x} Z_{\mathbf{x}'}\} = K(\mathbf{x}, \mathbf{x}')$

Greedy minimisation of $\mathscr{E}_K^{\neq}(\mathbf{X}_n)$, or **Conditional gradient descent for** $\mathscr{E}_K(\xi)$ (= Frank-Wolfe = Wynn's vertex-direction alg.)

(P&Zh, 2023b): if

$q = q_n$ is such that $q_n / \log n \to \infty$ in $K_q(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^{-q}$, or

$\ell = \ell_n$ is such that $\ell_n n^{1/d} (\log n)^{1/\gamma} \to 0$ in $K_{\gamma,\ell}(\mathbf{x}, \mathbf{x}')$,

then $\liminf_{n\to\infty} \mathrm{PR}(\mathbf{X}_n) / \mathrm{PR}_n^* \geq 1/2$ and $\limsup_{n\to\infty} \mathrm{CR}(\mathbf{X}_n) / \mathrm{CR}_n^* \leq 2$

General statements about $\mathrm{Supp}(\xi_K^+)$:

• When $\boxed{K \text{ is PD}}$, translation-invariant, with $K(\mathbf{x}, \mathbf{x}') = \Psi(\mathbf{x} - \mathbf{x}')$

➞ 2 conjectures (related to the non-existence of the continuous BLUE of β in the location model $Y(\mathbf{x}) = \beta + \varepsilon(\mathbf{x})$ where the errors $\varepsilon(\mathbf{x})$ have zero mean and covariance K)

- ξ_K^+ is not of full support when Ψ is differentiable at $\mathbf{0}_d$
- When $d > 1$, ξ_K^+ is not of full support unless K is singular

4 Minimisation of the covering radius

Minimise $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{z}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{z} - \mathbf{x}_i\|$

4 Minimisation of the covering radius

Minimise $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{z}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

ℓ_q-relaxation of $\min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

Minimise $\max_{\mathbf{z}\in\mathscr{X}} \left(\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{z}-\mathbf{x}_i\|^{-q}\right)^{-1/q}$ for some $q > 0$

4 Minimisation of the covering radius

Minimise $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{z}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{z} - \mathbf{x}_i\|$

ℓ_q-relaxation of $\min_{\mathbf{x}_i} \|\mathbf{z} - \mathbf{x}_i\|$

Minimise $\max_{\mathbf{z}\in\mathscr{X}} \left(\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{z} - \mathbf{x}_i\|^{-q}\right)^{-1/q}$ for some $q > 0$

$\Longleftrightarrow$ maximise $\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi_n}(\mathbf{z})$

→ continuous version: **PB2** maximise $\boxed{\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi}(\mathbf{z})}$

4 Minimisation of the covering radius

Minimise $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{z}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

ℓ_q-relaxation of $\min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

Minimise $\max_{\mathbf{z}\in\mathscr{X}} \left(\frac{1}{n}\sum_{i=1}^n \|\mathbf{z}-\mathbf{x}_i\|^{-q}\right)^{-1/q}$ for some $q>0$

$\iff$ maximise $\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi_n}(\mathbf{z})$

→ continuous version: **PB2** maximise $\boxed{\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi}(\mathbf{z})}$

Other kernels can be used, e.g., $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

PB2 is called **continuous polarisation**: concave, but difficult (non diff.)

4 Minimisation of the covering radius

Minimise $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{z}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

ℓ_q**-relaxation of** $\min_{\mathbf{x}_i} \|\mathbf{z}-\mathbf{x}_i\|$

Minimise $\max_{\mathbf{z}\in\mathscr{X}} \left(\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{z}-\mathbf{x}_i\|^{-q}\right)^{-1/q}$ for some $q>0$

$\iff$ maximise $\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi_n}(\mathbf{z})$

→ continuous version: **PB2** maximise $\boxed{\min_{\mathbf{z}\in\mathscr{X}} P_{K_q,\xi}(\mathbf{z})}$

Other kernels can be used, e.g., $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

PB2 is called **continuous polarisation**: concave, but difficult (non diff.)

When $\mathscr{E}_K(\cdot)$ is convex, ξ_K^+ of **PB1** maximises $\min_{\mathbf{x}\in\mathscr{X}} P_{K,\xi}(\mathbf{x}) - \mathscr{E}_K(\xi)$
if ξ_K^+ has full support, it is also optimal for **PB2**

true for $d=1$ when $K(x,x') = \psi(|x-x'|)$ with ψ convex,
and for any d when $K=K_q$ (Riesz), $q\in(d-2,d)$
(... but we conjecture it is not the case when $d>1$ for nonsingular kernels)

Doubly K-relaxed covering

relax max and min in $\mathsf{CR}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x}-\mathbf{x}_i\| \approx \max_{\mathbf{x}\in\mathscr{X}} P_{K_q,\xi_n}^{-1/q}(\mathbf{x})$

$$\text{minimise } \left\{\int_{\mathscr{X}} \left[P_{K_q,\xi}^{-1/q}(\mathbf{x})\right]^q \mathrm{d}\mu(\mathbf{x})\right\}^{1/q} \Rightarrow \text{minimise } \int_{\mathscr{X}} P_{K_q,\xi}^{-1}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$$

Doubly K-relaxed covering

relax max and min in $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x}-\mathbf{x}_i\| \approx \max_{\mathbf{x}\in\mathscr{X}} P_{K_q,\xi_n}^{-1/q}(\mathbf{x})$

$$\text{minimise } \left\{\int_{\mathscr{X}} \left[P_{K_q,\xi}^{-1/q}(\mathbf{x})\right]^q \mathrm{d}\mu(\mathbf{x})\right\}^{1/q} \Rightarrow \text{minimise } \int_{\mathscr{X}} P_{K_q,\xi}^{-1}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$$

= convex problem (strictly when K ISPD and μ equivalent to Λ) → **ET**

ξ^* is optimal $\Leftrightarrow$ for all $\mathbf{x}'\in\mathscr{X}$, $\displaystyle\int_{\mathscr{X}} \frac{K(\mathbf{x},\mathbf{x}')}{P_{K,\xi^*}^2(\mathbf{x})}\,\mathrm{d}\mu(\mathbf{x}) \leq \int_{\mathscr{X}} P_{K,\xi^*}^{-1}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$

Doubly K-relaxed covering

relax max and min in $\mathrm{CR}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x}-\mathbf{x}_i\| \approx \max_{\mathbf{x}\in\mathscr{X}} P_{K_q,\xi_n}^{-1/q}(\mathbf{x})$

$$\text{minimise } \left\{\int_{\mathscr{X}} \left[P_{K_q,\xi}^{-1/q}(\mathbf{x})\right]^q \mathrm{d}\mu(\mathbf{x})\right\}^{1/q} \Rightarrow \boxed{\text{minimise } \int_{\mathscr{X}} P_{K,\xi}^{-1}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})}$$

= convex problem (strictly when K ISPD and μ equivalent to Λ) → **ET**

$$\xi^* \text{ is optimal } \Leftrightarrow \text{ for all } \mathbf{x}'\in\mathscr{X}\,,\ \int_{\mathscr{X}} \frac{K(\mathbf{x},\mathbf{x}')}{P_{K,\xi^*}^2(\mathbf{x})}\,\mathrm{d}\mu(\mathbf{x}) \le \int_{\mathscr{X}} P_{K,\xi^*}^{-1}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$$

Discretisation of μ:

→ Replace μ by μ_N uniform on $\mathscr{X}_N = \{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(N)}\}$

→ $\boxed{\text{minimise } \Phi_1[\mathbf{M}_K(\xi)] = \mathrm{trace}\left[\frac{1}{N}\mathbf{M}_K^{-1}(\xi)\right]}$ (A-optimal design) with

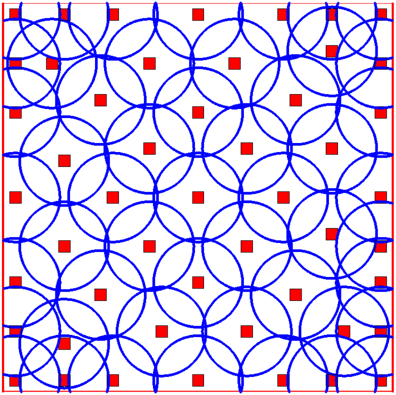
$$\mathbf{M}_K(\xi) = \mathrm{diag}\{P_{K,\xi}(\mathbf{x}^{(j)}),\ j=1,\ldots,N\} = \int_{\mathscr{X}} \mathrm{diag}\{K(\mathbf{x}^{(j)},\mathbf{x}),\ j=1,\ldots,N\}\,\mathrm{d}\xi(\mathbf{x})$$

(considered in (P&Zh, 2019) for $K = K_q$)

Example 4: $n = 50$ points in $\mathscr{X} = [0,1]^2$ for K_q with $q = 10$

➠ incremental design by vertex-direction alg. ($\mathbf{x}_1 = (1/2, 1/2)$)

($\mathscr{X}_N = 33 \times 33$ regular grid, ξ supported on the 32×32 interlaced grid)



(radius $= \mathrm{CR}(\mathbf{X}_n)$)

5 Minimisation of the L_s-mean quantisation error

ℓ_q-relaxation of $E_{s,\mu}(\mathbf{X}_n) = (\int_{\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^s \, \mathrm{d}\mu(\mathbf{x}))^{1/s}$, $s > 0$

Replace $d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|$ by $\left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q}$

→ minimise $E_{K_q,s,\mu}(\mathbf{X}_n) = \left[\int_{\mathscr{X}} \left(P_{K_q,\xi_n}(\mathbf{x})\right)^{-s/q} \mathrm{d}\mu(\mathbf{x})\right]^{1/s}$ for $s, q > 0$

→ continuous version:

PB3: minimise $\boxed{e_{K_q,s,\mu}(\xi) = \left[\int_{\mathscr{X}} P_{K_q,\xi}^{-s/q}(\mathbf{x}) \, \mathrm{d}\mu(\mathbf{x})\right]^{1/s}}$, $s, q > 0$

[= extension of (P&Zh, 2019) where only the case $q = s$ is considered]

5 Minimisation of the L_s-mean quantisation error

ℓ_q-relaxation of $E_{s,\mu}(\mathbf{X}_n) = \left(\int_{\mathscr{X}} \min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|^s \,\mathrm{d}\mu(\mathbf{x})\right)^{1/s}$, $s > 0$

Replace $d(\mathbf{x}, \mathbf{X}_n) = \min_{\mathbf{x}_i} \|\mathbf{x} - \mathbf{x}_i\|$ by $\left[P_{K_q,\xi_n}(\mathbf{x})\right]^{-1/q}$

→ minimise $E_{K_q,s,\mu}(\mathbf{X}_n) = \left[\int_{\mathscr{X}} \left(P_{K_q,\xi_n}(\mathbf{x})\right)^{-s/q} \,\mathrm{d}\mu(\mathbf{x})\right]^{1/s}$ for $s, q > 0$

→ continuous version:

PB3: minimise $\boxed{e_{K_q,s,\mu}(\xi) = \left[\int_{\mathscr{X}} P_{K_q,\xi}^{-s/q}(\mathbf{x}) \,\mathrm{d}\mu(\mathbf{x})\right]^{1/s}}$, $s, q > 0$

[= extension of (P&Zh, 2019) where only the case $q = s$ is considered]

μ equivalent to Λ, $q \in (0, d)$ → $e^s_{K_q,s,\mu}(\cdot)$ is convex and differentiable → **ET**

$$\xi_q^* \text{ is optimal} \Leftrightarrow \text{for all } \mathbf{x}' \in \mathscr{X},\ \int_{\mathscr{X}} \frac{K_q(\mathbf{x},\mathbf{x}')}{P^{1+s/q}_{K_q,\xi_q^*}(\mathbf{x})} \,\mathrm{d}\mu(\mathbf{x}) \le e^s_{K_q,s,\mu}(\xi_q^*)$$

Discretisation of μ:

→ Replace μ by μ_N uniform on $\mathscr{X}_N = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$:

→ $\boxed{\text{minimise } e^q_{K_q,s,\mu_N}(\xi) = \left[\mathrm{trace}\left(\frac{1}{N}\,\mathbf{M}^{-p}_{K_q}(\xi)\right)\right]^{1/p}}$ (Φ_p-optimal design, $p = s/q$)

$$\mathbf{M}_K(\xi) = \int_{\mathscr{X}} \mathrm{diag}\{K(\mathbf{x}^{(j)}, \mathbf{x}),\, j=1, \ldots, N\}\, \mathrm{d}\xi(\mathbf{x})$$

Relaxation with $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

$$E_{K_{\gamma,\ell},s,\mu}(\mathbf{X}_n) = \ell\,\gamma^{1/\gamma} \left(\int_{\mathscr{X}} \left[-\log P_{K_{e,\gamma,\ell},\xi_{n,e}}(\mathbf{x})\right]^{s/\gamma} \mathrm{d}\mu(\mathbf{x})\right)^{1/s}$$

Relaxation with $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

$$E_{K_{\gamma,\ell},s,\mu}(\mathbf{X}_n) = \ell\,\gamma^{1/\gamma}\left(\int_{\mathscr{X}}\left[-\log P_{K_{e,\gamma,\ell},\xi_{n,e}}(\mathbf{x})\right]^{s/\gamma}\,\mathrm{d}\mu(\mathbf{x})\right)^{1/s}$$

Simplification for $\gamma = s$

➠ $E^s_{K_{s,\ell},s,\mu}(\mathbf{X}_n) = -s\,\ell^s\int_{\mathscr{X}}\log P_{K_{e,s,\ell},\xi_{n,e}}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$ ($\rightarrow E^s_{s,\mu}(\mathbf{X}_n)$ as $\ell\rightarrow 0$)

Relaxation with $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

$$E_{K_{\gamma,\ell},s,\mu}(\mathbf{X}_n) = \ell\,\gamma^{1/\gamma}\left(\int_{\mathscr{X}}\left[-\log P_{K_{e,\gamma,\ell},\xi_{n,e}}(\mathbf{x})\right]^{s/\gamma}\,\mathrm{d}\mu(\mathbf{x})\right)^{1/s}$$

Simplification for $\gamma = s$

➠ $E^s_{K_{s,\ell},s,\mu}(\mathbf{X}_n) = -s\,\ell^s\int_{\mathscr{X}}\log P_{K_{e,s,\ell},\xi_{n,e}}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$ ($\to E^s_{s,\mu}(\mathbf{X}_n)$ as $\ell\to 0$)

➞ continuous version:

$e^s_{K_{s,\ell},s,\mu}(\cdot)$ is convex (strictly when μ equivalent to Λ and $s\in(0,2]$) ➞ **ET**

ξ^*_s is optimal $\Leftrightarrow$ for all $\mathbf{x}'\in\mathscr{X}$, $\int_{\mathscr{X}}\frac{K_{s,\ell}(\mathbf{x},\mathbf{x}')}{P_{K_{s,\ell},\xi^*_s}(\mathbf{x})}\,\mathrm{d}\mu(\mathbf{x})\le 1$

(but ξ^*_K is not of full support even when $\xi^+_{K_{s,\ell}}$ is ➞ Example 5)

Relaxation with $K_{\gamma,\ell}(\mathbf{x},\mathbf{x}') = \exp[-\|\mathbf{x}-\mathbf{x}'\|^\gamma/(\gamma\ell^\gamma)]$

$$E_{K_{\gamma,\ell},s,\mu}(\mathbf{X}_n) = \ell\,\gamma^{1/\gamma}\left(\int_{\mathscr{X}}\left[-\log P_{K_{e,\gamma,\ell},\xi_{n,e}}(\mathbf{x})\right]^{s/\gamma}\mathrm{d}\mu(\mathbf{x})\right)^{1/s}$$

Simplification for $\gamma = s$

➠ $E^s_{K_{s,\ell},s,\mu}(\mathbf{X}_n) = -s\,\ell^s\int_{\mathscr{X}}\log P_{K_{e,s,\ell},\xi_{n,e}}(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x})$ $(\to E^s_{s,\mu}(\mathbf{X}_n)$ as $\ell\to 0)$

→ continuous version:

$e^s_{K_{s,\ell},s,\mu}(\cdot)$ is convex (strictly when μ equivalent to Λ and $s\in(0,2]$) → **ET**

ξ^*_s is optimal $\Leftrightarrow$ for all $\mathbf{x}'\in\mathscr{X}$, $\int_{\mathscr{X}}\frac{K_{s,\ell}(\mathbf{x},\mathbf{x}')}{P_{K_{s,\ell},\xi^*_s}(\mathbf{x})}\,\mathrm{d}\mu(\mathbf{x})\le 1$

(but ξ^*_K is not of full support even when $\xi^+_{K_{s,\ell}}$ is → Example 5)

Discretisation of μ:

→ Replace μ by μ_N uniform on $\mathscr{X}_N = \{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(N)}\}$:

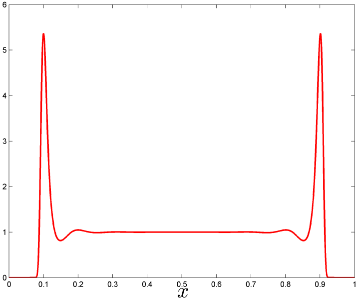
→ minimise $e^s_{K_{s,\ell},s,\mu_N}(\xi) = s\,\ell^s\,\log\left\{\det{}^{-1/N}[\mathbf{M}_{K_{s,\ell}}(\xi)]\right\}$ (D-optimal design)

$\mathbf{M}_K(\xi) = \int_{\mathscr{X}}\mathrm{diag}\{K(\mathbf{x}^{(j)},\mathbf{x}),\ j=1,\ldots,N\}\,\mathrm{d}\xi(\mathbf{x})$

Example 5: $\mathscr{X} = [0,1]$, $K(x,x') = \exp(-|x-x'|/\ell)$ (Matérn 1/2), $s = q = 1$, $\mu = \mathcal{U}_{\mathscr{X}}$ uniform on $\mathscr{X}$

Numerical determination of ξ_K^* for $\ell = 1/10$ (multiplicative algorithm):

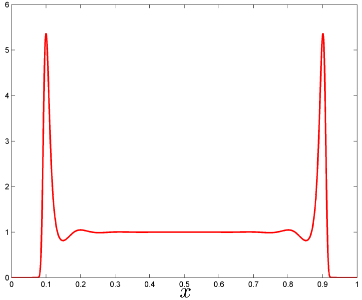
ξ_K^*



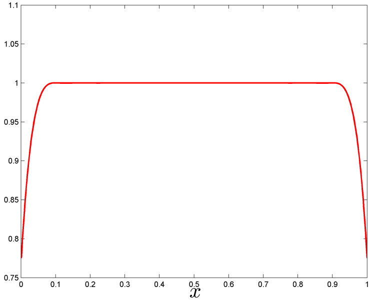
Example 5: $\mathscr{X} = [0,1]$, $K(x,x') = \exp(-|x-x'|/\ell)$ (Matérn 1/2), $s = q = 1$, $\mu = \mathcal{U}_{\mathscr{X}}$ uniform on $\mathscr{X}$

Numerical determination of ξ_K^* for $\ell = 1/10$ (multiplicative algorithm):

ξ_K^*



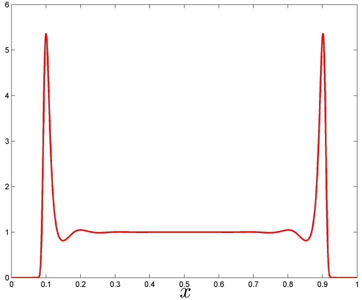
$\int \frac{K(\mathbf{x},\mathbf{z})}{P_{K,\xi_K^*}(\mathbf{z})} \, \mathrm{d}\mu(\mathbf{z})$



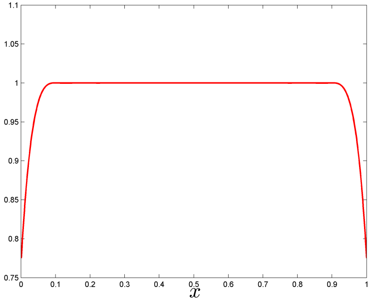
Example 5: $\mathscr{X} = [0,1]$, $K(x,x') = \exp(-|x-x'|/\ell)$ (Matérn 1/2), $s = q = 1$, $\mu = \mathcal{U}_{\mathscr{X}}$ uniform on $\mathscr{X}$

Numerical determination of ξ_K^* for $\ell = 1/10$ (multiplicative algorithm):

ξ_K^*



$\int \frac{K(\mathbf{x},\mathbf{z})}{P_{K,\xi_K^*}(\mathbf{z})}\,\mathrm{d}\mu(\mathbf{z})$



Exact solution:

$$\begin{array}{ll} \ell \geq 1/2 & \xi_K^* = \delta_{1/2} \\ 1/2 > \ell & \xi_K^* = \ell\,(\delta_\ell + \delta_{1-\ell}) + (1-2\,\ell)\Lambda([\ell, 1-\ell]) \end{array}$$

(and $\xi_K^* \to \mu = \mathcal{U}_{\mathscr{X}}$ when $\ell \to 0$)

6 Conclusions

- After kernel relaxation, the packing, covering and quantisation problems are related to standard Φ_p-optimal design, including A- and D-optimal design

6 Conclusions

- After kernel relaxation, the packing, covering and quantisation problems are related to standard Φ_p-optimal design, including A- and D-optimal design
- Gradient-type descent (vertex-direction) methods can be used for incremental constructions (→ kernel herding of machine learning)

6 Conclusions

- After kernel relaxation, the packing, covering and quantisation problems are related to standard Φ_p-optimal design, including A- and D-optimal design
- Gradient-type descent (vertex-direction) methods can be used for incremental constructions (→ kernel herding of machine learning)
- Several theoretical questions remain open and are challenging, concerning the support of a minimum-energy probability measure (related to the existence of a minimum-energy signed measure of total mass one and to the existence of the continuous BLUE in the location model with correlated errors), in connection with the properties of K

Thank you for your attention!

Graf, S., Luschgy, H., 2000. Foundations of Quantization for Probability Distributions. Springer, Berlin.

Krieg, D., Sonnleitner, M., 2020. Random points are optimal for the approximation of Sobolev functions. arXiv preprint arXiv:2009.11275.

Landkof, N., 1972. Foundations of Modern Potential Theory. Springer, Berlin.

Narcowich, F., Ward, J., Wendland, H., 2005. Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting. Mathematics of Computation 74 (250), 743–763.

Pagès, G., 1997. A space quantization method for numerical integration. Journal of Computational and Applied Mathematics 89 (1), 1–38.

Pronzato, L., Zhigljavsky, A., 2019. Measures minimizing regularized dispersion. Journal of Scientific Computing 78 (3), 1550–1570.

Pronzato, L., Zhigljavsky, A., 2023a. BLUE against OLSE in the location model: energy minimization and asymptotic considerations. Statistical Papers (to appear), doi.org/10.1007/s00362-023-01423-2, hal-03909349.

Pronzato, L., Zhigljavsky, A., 2023b. Quasi-uniform designs with asymptotically optimal and near-optimal uniformity constant. Journal of Approximation Theory 294 (105931), hal-03494864, arXiv:2112.10401.

Schaback, R., Wendland, H., 2006. Kernel techniques: from machine learning to meshless methods. Acta Numerica 15, 543–639.