Optimal Rate for Random Kitchen Sinks – Journey to Empirical Process Land

Zoltán Szabó

Joint work Bharath K. Sriperumbudur (PSU)

Gatsby Unit, Research Talk
May 18, 2015

Task

- Given:

$$k(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}^T(\mathbf{x}-\mathbf{y})} \mathrm{d}\mathbb{S}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \cos\left(\boldsymbol{\omega}^T(\mathbf{x}-\mathbf{y})\right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}).$$

- $s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})$: Monte-Carlo estimator of $\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y})$ using $(\boldsymbol{\omega}_j)_{j=1}^m \overset{i.i.d.}{\sim} \mathbb{S}$.
- **Goal**: uniform large deviation inequality

$$\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| \le \epsilon\right) \ge 1 - f(\epsilon, d, m, \mathcal{K}).$$

Why?

- [Rahimi and Recht, 2007] = random kitchen sinks:
  - Existing proof: contains several errors.
  - $\mathcal{O}\left(\sqrt{\frac{\log(m)}{m}}\right)$ rate: not optimal.
  - $\mathbf{p} = \mathbf{q} = \mathbf{0}$.
- Wanted rate: $\mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$.
- Connections: nonparametric EP, ∞-D exp. family fitting.
- Interest in statistical learning theory, empirical processes.

High-level proof

1. Empirical process form:

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| = \sup_{g\in\mathcal{G}} \left|\mathbb{S}g - \mathbb{S}_m g\right| = \left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}}.$$

2. $\left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}}$ concentrates by its bounded difference property:

$$\left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}} \precsim \mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m} \left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}} + \frac{1}{\sqrt{m}}.$$

3. $\mathcal{G}$ is a uniformly bounded, separable Carathéodory family $\Rightarrow$

$$\mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m} \left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}} \precsim \mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m} \mathcal{R}\left(\mathcal{G}, (\boldsymbol{\omega}_j)_{j=1}^m\right).$$

High-level proof

4. Using Dudley's entropy integral:

$$\mathcal{R}\left(\mathcal{G},(\boldsymbol{\omega}_j)_{j=1}^m\right) \lesssim \frac{1}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r.$$

5. $\mathcal{G}$ is smoothly parameterized by a compact set $\Rightarrow$

$$\sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)} \leq \frac{f\left((\boldsymbol{\omega}_j)_{j=1}^m\right)}{\sqrt{r}} \Rightarrow \mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m}\mathcal{R}\left(\mathcal{G},(\boldsymbol{\omega}_j)_{j=1}^m\right) \lesssim \frac{1}{\sqrt{m}}.$$

6. Putting together:

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}}k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| \lesssim \frac{1}{\sqrt{m}} + \frac{1}{\sqrt{m}} = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

Notations

- For $\mathbf{p}, \mathbf{q} \in \mathbb{N}^d$ multi-indices and $\mathbf{w} \in \mathbb{R}^d$, let

$$|\mathbf{p}| = \sum_{j=1}^{d} |p_j|, \quad \partial^{\mathbf{p},\mathbf{q}} g(\mathbf{x}, \mathbf{y}) = \frac{\partial^{|\mathbf{p}|+|\mathbf{q}|} g(\mathbf{x}, \mathbf{y})}{\partial x_1^{p_1} \cdots \partial x_d^{p_d} \partial y_1^{q_1} \cdots \partial y_d^{q_d}}, \quad \mathbf{w}^{\mathbf{p}} = \prod_{j=1}^{d} w_j^{p_j}.$$

Notations

- For $\mathbf{p}, \mathbf{q} \in \mathbb{N}^d$ multi-indices and $\mathbf{w} \in \mathbb{R}^d$, let

$$|\mathbf{p}| = \sum_{j=1}^{d} |p_j|, \quad \partial^{\mathbf{p},\mathbf{q}} g(\mathbf{x}, \mathbf{y}) = \frac{\partial^{|\mathbf{p}|+|\mathbf{q}|} g(\mathbf{x}, \mathbf{y})}{\partial x_1^{p_1} \cdots \partial x_d^{p_d} \partial y_1^{q_1} \cdots \partial y_d^{q_d}}, \quad \mathbf{w}^{\mathbf{p}} = \prod_{j=1}^{d} w_j^{p_j}.$$

- $k \to \partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x}, \mathbf{y}) \to s^{\mathbf{p},\mathbf{q}}(\mathbf{x}, \mathbf{y})$: $h_0 := \cos$,

$$k(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}^d} h_0 \left(\boldsymbol{\omega}^T (\mathbf{x} - \mathbf{y})\right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}),$$

Notations

- For $\mathbf{p}, \mathbf{q} \in \mathbb{N}^d$ multi-indices and $\mathbf{w} \in \mathbb{R}^d$, let

$$|\mathbf{p}| = \sum_{j=1}^{d} |p_j|, \quad \partial^{\mathbf{p},\mathbf{q}} g(\mathbf{x}, \mathbf{y}) = \frac{\partial^{|\mathbf{p}|+|\mathbf{q}|} g(\mathbf{x}, \mathbf{y})}{\partial x_1^{p_1} \cdots \partial x_d^{p_d} \partial y_1^{q_1} \cdots \partial y_d^{q_d}}, \quad \mathbf{w}^{\mathbf{p}} = \prod_{j=1}^{d} w_j^{p_j}.$$

- $k \to \partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x}, \mathbf{y}) \to s^{\mathbf{p},\mathbf{q}}(\mathbf{x}, \mathbf{y})$: $h_0 := \cos$,

$$k(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}^d} h_0 \left(\boldsymbol{\omega}^T (\mathbf{x} - \mathbf{y}) \right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}),$$

$$\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}^d} \boldsymbol{\omega}^{\mathbf{p}} (-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|} \left(\boldsymbol{\omega}^T (\mathbf{x} - \mathbf{y}) \right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}),$$

Notations

- For $\mathbf{p},\mathbf{q} \in \mathbb{N}^d$ multi-indices and $\mathbf{w} \in \mathbb{R}^d$, let

$$|\mathbf{p}| = \sum_{j=1}^d |p_j|, \quad \partial^{\mathbf{p},\mathbf{q}} g(\mathbf{x},\mathbf{y}) = \frac{\partial^{|\mathbf{p}|+|\mathbf{q}|} g(\mathbf{x},\mathbf{y})}{\partial x_1^{p_1} \cdots \partial x_d^{p_d} \partial y_1^{q_1} \cdots \partial y_d^{q_d}}, \quad \mathbf{w}^{\mathbf{p}} = \prod_{j=1}^d w_j^{p_j}.$$

- $k \to \partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) \to s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})$: $h_0 := \cos$,

$$k(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^d} h_0\left(\boldsymbol{\omega}^T(\mathbf{x}-\mathbf{y})\right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}),$$

$$\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^d} \boldsymbol{\omega}^{\mathbf{p}}(-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\boldsymbol{\omega}^T(\mathbf{x}-\mathbf{y})\right) \mathrm{d}\mathbb{S}(\boldsymbol{\omega}),$$

$$s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^d} \boldsymbol{\omega}^{\mathbf{p}}(-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\boldsymbol{\omega}^T(\mathbf{x}-\mathbf{y})\right) \mathrm{d}\mathbb{S}_m(\boldsymbol{\omega}),$$

where $h_\ell = \cos^{(\ell)}$, $\mathbb{E}_{\boldsymbol{\omega}\sim\mathbb{S}}[|\boldsymbol{\omega}^{\mathbf{p}+\mathbf{q}}|] < \infty$.

Step-1: empirical process form

- Notation: $\mathbb{S}g = \int g(\omega)\mathrm{d}\mathbb{S}(\omega)$, $\mathbb{S}_m g = \int g(\omega)\mathrm{d}\mathbb{S}_m(\omega) = \frac{1}{m}\sum_{j=1}^m g(\omega_j)$.

Step-1: empirical process form

- Notation: $\mathbb{S}g = \int g(\omega)\mathrm{d}\mathbb{S}(\omega)$, $\mathbb{S}_m g = \int g(\omega)\mathrm{d}\mathbb{S}_m(\omega) = \frac{1}{m}\sum_{j=1}^m g(\omega_j)$.
- Reformulation of the objective:

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| = \sup_{g\in\mathcal{G}} \left|\mathbb{S}g - \mathbb{S}_m g\right| =: \left\|\mathbb{S} - \mathbb{S}_m\right\|_{\mathcal{G}},$$

where

$$\mathcal{G} = \{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{K}_\Delta\},$$
$$\mathcal{K}_\Delta = \mathcal{K} - \mathcal{K} = \{\mathbf{x} - \mathbf{y} : \mathbf{x},\mathbf{y} \in \mathcal{K}\},$$
$$g_{\mathbf{z}} : \omega \in supp(\mathbb{S}) \mapsto \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}\right) \in \mathbb{R}.$$

Step-2: bounded difference property of $\|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

McDiarmid inequality: Let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_m \in D$ be independent r.v.-s, and $f : D^m \to \mathbb{R}$ satisfy the bounded diff. property ($\forall r$):

$$\sup_{u_1,\ldots,u_m,u_r' \in D} \left| f(u_1, \ldots, u_m) - f(u_1, \ldots, u_{r-1}, u_r', u_{r+1}, \ldots, u_m) \right| \leq c_r.$$

Then for $\forall \beta > 0$

$$\mathbb{P}\left(f(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_m) - \mathbb{E}\left[f(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_m)\right] \geq \beta\right) \leq e^{-\frac{2\beta^2}{\sum_{r=1}^m c_r^2}}.$$

Step-2: bounded difference property of $\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$

Our choice: $f(\omega_1,\ldots,\omega_m) := \|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1,\ldots,\omega_{r-1},\omega_r,\omega_{r+1},\ldots,\omega_m) - f(\omega_1,\ldots,\omega_{r-1},\omega_r',\omega_{r+1},\ldots,\omega_m)| =$$

$$= \left| \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) \right| - \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) + \frac{1}{m}\left[g(\omega_r) - g(\omega_r')\right] \right| \right|$$

Step-2: bounded difference property of $\|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

Our choice: $f(\omega_1, \ldots, \omega_m) := \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1, \ldots, \omega_{r-1}, \omega_r, \omega_{r+1}, \ldots, \omega_m) - f(\omega_1, \ldots, \omega_{r-1}, \omega'_r, \omega_{r+1}, \ldots, \omega_m)| =$$

$$= \left| \sup_{g \in \mathcal{G}} \left| \mathbb{S}g - \frac{1}{m} \sum_{j=1} g(\omega_j) \right| - \sup_{g \in \mathcal{G}} \left| \mathbb{S}g - \frac{1}{m} \sum_{j=1} g(\omega_j) + \frac{1}{m} \left[g(\omega_r) - g(\omega'_r) \right] \right| \right|$$

$$\stackrel{(*)}{\leq} \frac{1}{m} \sup_{g \in \mathcal{G}} \left| g(\omega_r) - g(\omega'_r) \right|$$

Step-2: bounded difference property of $\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$

Our choice: $f(\omega_1,\ldots,\omega_m) := \|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1,\ldots,\omega_{r-1},\omega_r,\omega_{r+1},\ldots,\omega_m) - f(\omega_1,\ldots,\omega_{r-1},\omega_r',\omega_{r+1},\ldots,\omega_m)| =$$

$$= \left| \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) \right| - \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) + \frac{1}{m}\left[g(\omega_r) - g(\omega_r')\right] \right| \right|$$

$$\overset{(*)}{\leq} \frac{1}{m}\sup_{g\in\mathcal{G}} \left| g(\omega_r) - g(\omega_r') \right| \leq \frac{1}{m}\sup_{g\in\mathcal{G}} \left(|g(\omega_r)| + |g(\omega_r')| \right)$$

Our choice: $f(\omega_1,\ldots,\omega_m) := \|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1,\ldots,\omega_{r-1},\omega_r,\omega_{r+1},\ldots,\omega_m) - f(\omega_1,\ldots,\omega_{r-1},\omega_r',\omega_{r+1},\ldots,\omega_m)| =$$

$$= \left| \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) \right| - \sup_{g\in\mathcal{G}} \left| \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) + \frac{1}{m}\left[g(\omega_r)-g(\omega_r')\right] \right| \right|$$

$$\overset{(*)}{\leq} \frac{1}{m}\sup_{g\in\mathcal{G}} |g(\omega_r)-g(\omega_r')| \leq \frac{1}{m}\sup_{g\in\mathcal{G}} \left(|g(\omega_r)| + |g(\omega_r')|\right)$$

$$\leq \frac{1}{m}\left[\sup_{g\in\mathcal{G}} |g(\omega_r)| + \sup_{g\in\mathcal{G}} |g(\omega_r')|\right]$$

Step-2: bounded difference property of $\|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

Our choice: $f(\omega_1, \ldots, \omega_m) := \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1, \ldots, \omega_{r-1}, \omega_r, \omega_{r+1}, \ldots, \omega_m) - f(\omega_1, \ldots, \omega_{r-1}, \omega_r', \omega_{r+1}, \ldots, \omega_m)| =$$

$$= \left| \sup_{g \in \mathcal{G}} \left| \mathbb{S}g - \frac{1}{m} \sum_{j=1} g(\omega_j) \right| - \sup_{g \in \mathcal{G}} \left| \mathbb{S}g - \frac{1}{m} \sum_{j=1} g(\omega_j) + \frac{1}{m} \left[g(\omega_r) - g(\omega_r') \right] \right| \right|$$

$$\stackrel{(*)}{\leq} \frac{1}{m} \sup_{g \in \mathcal{G}} |g(\omega_r) - g(\omega_r')| \leq \frac{1}{m} \sup_{g \in \mathcal{G}} \left(|g(\omega_r)| + |g(\omega_r')| \right)$$

$$\leq \frac{1}{m} \left[\sup_{g \in \mathcal{G}} |g(\omega_r)| + \sup_{g \in \mathcal{G}} |g(\omega_r')| \right] \leq \frac{1}{m} \left[|\omega_r^{\mathbf{p}+\mathbf{q}}| + |(\omega_r')^{\mathbf{p}+\mathbf{q}}| \right]$$

Step-2: bounded difference property of $\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$

Our choice: $f(\omega_1,\ldots,\omega_m) := \|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$.

$$|f(\omega_1,\ldots,\omega_{r-1},\omega_r,\omega_{r+1},\ldots,\omega_m) - f(\omega_1,\ldots,\omega_{r-1},\omega_r',\omega_{r+1},\ldots,\omega_m)| =$$

$$= \left|\sup_{g\in\mathcal{G}}\left|\mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j)\right| - \sup_{g\in\mathcal{G}}\left|\mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j) + \frac{1}{m}\left[g(\omega_r) - g(\omega_r')\right]\right|\right|$$

$$\overset{(*)}{\leq} \frac{1}{m}\sup_{g\in\mathcal{G}}\left|g(\omega_r) - g(\omega_r')\right| \leq \frac{1}{m}\sup_{g\in\mathcal{G}}\left(\left|g(\omega_r)\right| + \left|g(\omega_r')\right|\right)$$

$$\leq \frac{1}{m}\left[\sup_{g\in\mathcal{G}}\left|g(\omega_r)\right| + \sup_{g\in\mathcal{G}}\left|g(\omega_r')\right|\right] \leq \frac{1}{m}\left[\left|\omega_r^{\mathbf{p}+\mathbf{q}}\right| + \left|(\omega_r')^{\mathbf{p}+\mathbf{q}}\right|\right]$$

$$\leq \frac{2S_{k,\mathbf{p},\mathbf{q}}}{m},$$

where $S_{k,\mathbf{p},\mathbf{q}} := \sup_{\omega\in supp(\mathbb{S})}\left|\omega^{\mathbf{p}+\mathbf{q}}\right|$.

Step-2: (*) = reverse triangle inequality with sup

- Lemma: $\mathcal{G}$: set of functions, $a, b : \mathcal{G} \to \mathbb{R}$ maps; then

$$\left| \sup_{g \in \mathcal{G}} |a(g)| - \sup_{g \in \mathcal{G}} |a(g) + b(g)| \right|$$

Step-2: (*) = reverse triangle inequality with sup

- Lemma: $\mathcal{G}$: set of functions, $a, b : \mathcal{G} \to \mathbb{R}$ maps; then

$$\left| \sup_{g \in \mathcal{G}} |a(g)| - \sup_{g \in \mathcal{G}} |a(g) + b(g)| \right| \leq \sup_{g \in \mathcal{G}} |b(g)|.$$

Step-2: (*) = reverse triangle inequality with sup

- Lemma: $\mathcal{G}$: set of functions, $a, b : \mathcal{G} \to \mathbb{R}$ maps; then

$$\left| \sup_{g\in\mathcal{G}} |a(g)| - \sup_{g\in\mathcal{G}} |a(g) + b(g)| \right| \leq \sup_{g\in\mathcal{G}} |b(g)|.$$

- Proof: combine

$$sup_{g\in\mathcal{G}} |a(g) + b(g)| \leq \sup_{g\in\mathcal{G}} \left(|a(g)| + |b(g)|\right) \leq \sup_{g\in\mathcal{G}} |a(g)| + \sup_{g\in\mathcal{G}} |b(g)|,$$

Step-2: (*) = reverse triangle inequality with sup

- Lemma: $\mathcal{G}$: set of functions, $a, b : \mathcal{G} \to \mathbb{R}$ maps; then

$$\left| \sup_{g\in\mathcal{G}} |a(g)| - \sup_{g\in\mathcal{G}} |a(g) + b(g)| \right| \leq \sup_{g\in\mathcal{G}} |b(g)|.$$

- Proof: combine

$$sup_{g\in\mathcal{G}} |a(g) + b(g)| \leq \sup_{g\in\mathcal{G}} \left(|a(g)| + |b(g)| \right) \leq \sup_{g\in\mathcal{G}} |a(g)| + \sup_{g\in\mathcal{G}} |b(g)|,$$

$$\begin{aligned} \sup_{g\in\mathcal{G}} |a(g)| &= \sup_{g\in\mathcal{G}} |a(g) + b(g) - b(g)| \\ &\leq \sup_{g\in\mathcal{G}} |a(g) + b(g)| + \sup_{g\in\mathcal{G}} |b(g)|. \end{aligned}$$

Step-2: (*) = reverse triangle inequality with sup

- Lemma: $\mathcal{G}$: set of functions, $a, b : \mathcal{G} \to \mathbb{R}$ maps; then

$$\left|\sup_{g\in\mathcal{G}} |a(g)| - \sup_{g\in\mathcal{G}} |a(g) + b(g)|\right| \le \sup_{g\in\mathcal{G}} |b(g)|.$$

- Proof: combine

$$\mathit{sup}_{g\in\mathcal{G}} |a(g) + b(g)| \le \sup_{g\in\mathcal{G}} (|a(g)| + |b(g)|) \le \sup_{g\in\mathcal{G}} |a(g)| + \sup_{g\in\mathcal{G}} |b(g)|,$$

$$\sup_{g\in\mathcal{G}} |a(g)| = \sup_{g\in\mathcal{G}} |a(g) + b(g) - b(g)| \le \sup_{g\in\mathcal{G}} |a(g) + b(g)| + \sup_{g\in\mathcal{G}} |b(g)|.$$

$$\Rightarrow \pm \left[\sup_{g\in\mathcal{G}} |a(g)| - \sup_{g\in\mathcal{G}} |a(g) + b(g)|\right] \le \sup_{g\in\mathcal{G}} |b(g)|.$$

- Our choice: $a(g) = \mathbb{S}g - \frac{1}{m}\sum_{j=1} g(\omega_j)$, $b(g) = \frac{1}{m}[g(\omega_r) - g(\omega_r')]$.

Step-2

Applying McDiarmid to f $[D = supp(\mathbb{S})$, $c_r = \frac{2S_{k,\mathbf{p},\mathbf{q}}}{m}$; $\tau = e^{-\frac{m\beta^2}{2S^2_{k,\mathbf{p},\mathbf{q}}}}]$: with probability $1-\tau$

$$\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}} \leq \underbrace{\mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m}\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}}_{\text{Step-3: bounding this term}} + \frac{S_{k,\mathbf{p},\mathbf{q}}\sqrt{\log\left(\frac{1}{\tau^2}\right)}}{\sqrt{m}}.$$

Step-3: bounding $\mathbb{E}_{\omega_1,\ldots,\omega_m} \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

$\mathcal{G} = \{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{K}_\Delta\}$ is a separable Carathéodory family, i.e.

1. $\boldsymbol{\omega} \mapsto \boldsymbol{\omega}^{\mathbf{p}}(-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\boldsymbol{\omega}^T \mathbf{z}\right)$: measurable for $\forall \mathbf{z} \in \mathcal{K}_\Delta$.

Step-3: bounding $\mathbb{E}_{\omega_1,\ldots,\omega_m} \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

$\mathcal{G} = \{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{K}_\Delta\}$ is a separable Carathéodory family, i.e.

1. $\omega \mapsto \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}\right)$: measurable for $\forall \mathbf{z} \in \mathcal{K}_\Delta$.
2. $\mathbf{z} \mapsto \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}\right)$: continuous for $\forall \omega \in supp(\mathbb{S})$.

Step-3: bounding $\mathbb{E}_{\omega_1,\ldots,\omega_m} \|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}}$

$\mathcal{G} = \{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{K}_\Delta\}$ is a separable Carathéodory family, i.e.

1. $\omega \mapsto \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}\right)$: measurable for $\forall \mathbf{z} \in \mathcal{K}_\Delta$.
2. $\mathbf{z} \mapsto \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}\right)$: continuous for $\forall \omega \in supp(\mathbb{S})$.
3. $\mathbb{R}^d$ is separable, $\mathcal{K}_\Delta \subseteq \mathbb{R}^d \Rightarrow \mathcal{K}_\Delta$: separable.

Step-3: bounding $\mathbb{E}_{\omega_1,\ldots,\omega_m} \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}}$

$\mathcal{G} = \{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{K}_\Delta\}$ is a separable Carathéodory family, i.e.

1. $\boldsymbol{\omega} \mapsto \boldsymbol{\omega}^{\mathbf{p}}(-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\boldsymbol{\omega}^T \mathbf{z}\right)$: measurable for $\forall \mathbf{z} \in \mathcal{K}_\Delta$.
2. $\mathbf{z} \mapsto \boldsymbol{\omega}^{\mathbf{p}}(-\boldsymbol{\omega})^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\boldsymbol{\omega}^T \mathbf{z}\right)$: continuous for $\forall \boldsymbol{\omega} \in supp(\mathbb{S})$.
3. $\mathbb{R}^d$ is separable, $\mathcal{K}_\Delta \subseteq \mathbb{R}^d \Rightarrow \mathcal{K}_\Delta$: separable.

Thus, by [Steinwart and Christmann, 2008, Prop. 7.10]

$$\mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m} \|\mathbb{S} - \mathbb{S}_m\|_{\mathcal{G}} \le 2\mathbb{E}_{\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_m}\Big[\underbrace{\mathcal{R}\left(\mathcal{G}, (\boldsymbol{\omega}_j)_{j=1}^m\right)}_{:=\mathbb{E}_\epsilon \sup_{g\in\mathcal{G}}\left|\frac{1}{m}\sum_{j=1}^m \epsilon_j g(\boldsymbol{\omega}_j)\right|} \Big]$$

using the uniformly boundedness of $\mathcal{G}$ $\left(\sup_{g\in\mathcal{G}} \|g\|_\infty \le S_{k,\mathbf{p},\mathbf{q}} < \infty\right)$.

Step-4: bounding $\mathcal{R}$

$$\mathcal{R}\left(\mathcal{G}, (\omega_j)_{j=1}^m\right) \le \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)} \mathrm{d}r,$$

Step-4: bounding $\mathcal{R}$

$$\mathcal{R}\left(\mathcal{G}, (\boldsymbol{\omega}_j)_{j=1}^m\right) \leq \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r,$$

where

- $L^2(\mathbb{S}_m) = L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{S}_m)$, $\|g\|_{L^2(\mathbb{S}_m)} = \sqrt{\frac{1}{m}\sum_{j=1}^m g^2(\omega_j)}$,

Step-4: bounding $\mathcal{R}$

$$\mathcal{R}\left(\mathcal{G},(\omega_j)_{j=1}^m\right) \le \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r,$$

where

- $L^2(\mathbb{S}_m) = L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{S}_m)$, $\|g\|_{L^2(\mathbb{S}_m)} = \sqrt{\frac{1}{m}\sum_{j=1}^m g^2(\omega_j)}$,
- $|\mathcal{G}|_{L^2(\mathbb{S}_m)} = \sup_{g_1,g_2\in\mathcal{G}} \|g_1 - g_2\|_{L^2(\mathbb{S}_m)}$.

Step-4: bounding $\mathcal{R}$

$$\mathcal{R}\left(\mathcal{G},(\omega_j)_{j=1}^m\right) \le \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r,$$

where

- $L^2(\mathbb{S}_m) = L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{S}_m)$, $\|g\|_{L^2(\mathbb{S}_m)} = \sqrt{\frac{1}{m}\sum_{j=1}^m g^2(\omega_j)}$,
- $|\mathcal{G}|_{L^2(\mathbb{S}_m)} = \sup_{g_1,g_2\in\mathcal{G}} \|g_1 - g_2\|_{L^2(\mathbb{S}_m)}$,
- $\mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)$: r-covering number.
  - r-net: $S \subseteq \mathcal{G}$, for $\forall g \in \mathcal{G}$ $\exists s \in S$ such that $\|g - s\|_{L^2(\mathbb{S}_m)} \le r$.
  - $\mathcal{N}$: size of the smallest r-net of $\mathcal{G}$.

Step-5: bound on $|\mathcal{G}|_{L^2(\mathbb{S}_m)}$

$$|\mathcal{G}|_{L^2(\mathbb{S}_m)} = \sup_{g_1, g_2 \in \mathcal{G}} \|g_1 - g_2\|_{L^2(\mathbb{S}_m)} \leq \sup_{g_1, g_2 \in \mathcal{G}} \left(\|g_1\|_{L^2(\mathbb{S}_m)} + \|g_2\|_{L^2(\mathbb{S}_m)} \right)$$

$$\leq \sup_{g_1 \in \mathcal{G}} \|g_1\|_{L^2(\mathbb{S}_m)} + \sup_{g_1 \in \mathcal{G}} \|g_2\|_{L^2(\mathbb{S}_m)} \overset{*}{\leq} 2\sqrt{S_{k,2\mathbf{p},2\mathbf{q}}},$$

Step-5: bound on $|\mathcal{G}|_{L^2(\mathbb{S}_m)}$

$$|\mathcal{G}|_{L^2(\mathbb{S}_m)} = \sup_{g_1,g_2\in\mathcal{G}} \|g_1 - g_2\|_{L^2(\mathbb{S}_m)} \le \sup_{g_1,g_2\in\mathcal{G}} \left(\|g_1\|_{L^2(\mathbb{S}_m)} + \|g_2\|_{L^2(\mathbb{S}_m)}\right)$$

$$\le \sup_{g_1\in\mathcal{G}} \|g_1\|_{L^2(\mathbb{S}_m)} + \sup_{g_1\in\mathcal{G}} \|g_2\|_{L^2(\mathbb{S}_m)} \overset{*}{\le} 2\sqrt{S_{k,2\mathbf{p},2\mathbf{q}}},$$

$$\sup_{g\in\mathcal{G}} \|g\|_{L^2(\mathbb{S}_m)} = \sup_{\mathbf{z}\in\mathcal{K}_\Delta} \sqrt{\frac{1}{m}\sum_{j=1}^m g_{\mathbf{z}}^2(\omega_j)}$$

$$= \sup_{\mathbf{z}\in\mathcal{K}_\Delta} \sqrt{\frac{1}{m}\sum_{j=1}^m \left[\omega_j^{\mathbf{p}}(-\omega_j)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega_j^T \mathbf{z}\right)\right]^2}$$

$$\le \sup_{\mathbf{z}\in\mathcal{K}_\Delta} \sqrt{\frac{1}{m}\sum_{j=1}^m \omega_j^{2(\mathbf{p}+\mathbf{q})}} \le \sqrt{S_{k,2\mathbf{p},2\mathbf{q}}}.$$

Step-5: bound on $\mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)$

Let $g_{\mathbf{z}_1}, g_{\mathbf{z}_2} \in \mathcal{G}$. We want to bound $\|g_{\mathbf{z}_1} - g_{\mathbf{z}_2}\|_{L^2(\mathbb{S}_m)}$. One term:

$$\begin{aligned}
&\left|\omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}_1\right) - \omega^{\mathbf{p}}(-\omega)^{\mathbf{q}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}_2\right)\right| \\
&= \left|\omega^{\mathbf{p}+\mathbf{q}}\right| \left|h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}_1\right) - h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}_2\right)\right| \\
&= \left|\omega^{\mathbf{p}+\mathbf{q}}\right| \left\|\nabla_{\mathbf{z}} h_{|\mathbf{p}+\mathbf{q}|}\left(\omega^T \mathbf{z}_c\right)\right\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\
&= \left|\omega^{\mathbf{p}+\mathbf{q}}\right| \left\|h_{|\mathbf{p}+\mathbf{q}|+1}\left(\omega^T \mathbf{z}_c\right)\omega\right\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \\
&\leq \left|\omega^{\mathbf{p}+\mathbf{q}}\right| \|\omega\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2,
\end{aligned}$$

where $\mathbf{z}_c \in (\mathbf{z}_1, \mathbf{z}_2)$, we used the convexity of $\mathcal{K}_\Delta$.

Step-5: bound on $\mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)$

- Smooth parameterization:

$$\|g_{\mathbf{z}_1} - g_{\mathbf{z}_2}\|_{L^2(\mathbb{S}_m)} \leq \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\left|\omega_j^{\mathbf{p}+\mathbf{q}}\right| \|\omega_j\|_2 \|\mathbf{z}_1 - \mathbf{z}_2\|_2\right)^2}$$

$$= \|\mathbf{z}_1 - \mathbf{z}_2\|_2 \underbrace{\sqrt{\frac{1}{m}\sum_{j=1}^{m}\left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\omega_j\|_2^2}}_{=:c}.$$

- r-net on $(\mathcal{K}_\Delta, \|\cdot\|_2) \Rightarrow r' = rc$-net on $(\mathcal{G}, L^2(\mathbb{S}_m))$.
- $M \subseteq \mathbb{R}^d$ compact set: coverable by $\left[\frac{2|M|}{s}\right]^d$ s-balls [Cucker and Smale, 2002].

Step-5: bound on $\mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)$

- Thus, by $|\mathcal{K}_\Delta| \le 2|\mathcal{K}|$ and the compactness of $\mathcal{K}_\Delta$

$$\mathcal{N}\left(\mathcal{G}, L^2(\mathbb{S}_m), r\right) \le \left(\frac{4|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\boldsymbol{\omega}_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\boldsymbol{\omega}_j\|_2^2}}{r}[+1] \right)^d .$$

Step-5: bound on $\mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)$

- Thus, by $|\mathcal{K}_\Delta| \le 2|\mathcal{K}|$ and the compactness of $\mathcal{K}_\Delta$

$$\mathcal{N}\left(\mathcal{G}, L^2(\mathbb{S}_m), r\right) \le \left(\frac{4|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\omega_j\|_2^2}}{r}[+1]\right)^d.$$

- Taking $\log(\cdot)$, using $\log(u+1) \le u$

$$\log\left[\mathcal{N}\left(\mathcal{G}, L^2(\mathbb{S}_m), r\right)\right] \le d\log\left(\frac{4|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\omega_j\|_2^2}}{r}+1\right)$$

$$\le \frac{4d|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\omega_j\|_2^2}}{r}.$$

Step-5: bound on $\mathcal{R}$

Combining the obtained

$$\mathcal{R}\left(\mathcal{G},(\boldsymbol{\omega}_j)_{j=1}^m\right) \le \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r,$$

$$|\mathcal{G}|_{L^2(\mathbb{S}_m)} \le 2\sqrt{S_{k,2\mathbf{p},2\mathbf{q}}},$$

$$\log\left[\mathcal{N}\left(\mathcal{G}, L^2(\mathbb{S}_m), r\right)\right] \le \frac{4d|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\boldsymbol{\omega}_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\boldsymbol{\omega}_j\|_2^2}}{r}$$

results

Step-5: bound on $\mathcal{R}$

Combining the obtained

$$\mathcal{R}\left(\mathcal{G},(\boldsymbol{\omega}_j)_{j=1}^m\right) \le \frac{8\sqrt{2}}{\sqrt{m}} \int_0^{|\mathcal{G}|_{L^2(\mathbb{S}_m)}} \sqrt{\log \mathcal{N}(\mathcal{G}, L^2(\mathbb{S}_m), r)}\mathrm{d}r,$$

$$|\mathcal{G}|_{L^2(\mathbb{S}_m)} \le 2\sqrt{S_{k,2\mathbf{p},2\mathbf{q}}},$$

$$\log\left[\mathcal{N}\left(\mathcal{G}, L^2(\mathbb{S}_m), r\right)\right] \le \frac{4d|\mathcal{K}|\sqrt{\frac{1}{m}\sum_{j=1}^m \left|\boldsymbol{\omega}_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\boldsymbol{\omega}_j\|_2^2}}{r}$$

results, we get $[\int_0^b r^{-\frac{1}{2}}\mathrm{d}r = 2\sqrt{b}]$

$$\mathcal{R}\left(\mathcal{G},(\boldsymbol{\omega}_j)_{j=1}^m\right) \le \frac{64\sqrt{d|\mathcal{K}|}(S_{k,2\mathbf{p},2\mathbf{q}})^{\frac{1}{4}}}{\sqrt{m}} \left(\frac{1}{m}\sum_{j=1}^m \left|\boldsymbol{\omega}_j^{2(\mathbf{p}+\mathbf{q})}\right| \|\boldsymbol{\omega}_j\|_2^2\right)^{\frac{1}{4}}.$$

Recall

$$\mathbb{E}_{\omega_1,\ldots,\omega_m}\|\mathbb{S}-\mathbb{S}_m\|_{\mathcal{G}} \le 2\mathbb{E}_{\omega_1,\ldots,\omega_m}\mathcal{R}\left(\mathcal{G},(\omega_j)_{j=1}^m\right) \overset{?}{=} \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

Taking expectation $[\mathbb{E}=\mathbb{E}_{\omega_1,\ldots,\omega_m},\ \mathcal{R}=\mathcal{R}(\mathcal{G},(\omega_j)_{j=1}^m)]$ of the derived result

$$\mathbb{E}_{\omega_1,\ldots,\omega_m}\mathcal{R} \le \frac{64\sqrt{d|\mathcal{K}|}(S_{k,2\mathbf{p},2\mathbf{q}})^{\frac{1}{4}}}{\sqrt{m}} \underbrace{\mathbb{E}\left(\frac{1}{m}\sum_{j=1}^m\left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right|\|\omega_j\|_2^2\right)^{\frac{1}{4}}}_{Q},$$

$$Q \le \left(\frac{1}{m}\sum_{j=1}^m\mathbb{E}\left|\omega_j^{2(\mathbf{p}+\mathbf{q})}\right|\|\omega_j\|_2^2\right)^{\frac{1}{4}} \le \left(\frac{1}{m}\sum_{j=1}^m C_{k,2\mathbf{p},2\mathbf{q}}\right)^{\frac{1}{4}} = (C_{k,2\mathbf{p},2\mathbf{q}})^{\frac{1}{4}},$$

using Jensen $[f(u)=u^{\frac{1}{4}}]$, $C_{k,\mathbf{p},\mathbf{q}} := \mathbb{E}_{\omega\sim\mathbb{S}}\left[\left|\omega^{\mathbf{p}+\mathbf{q}}\right|\|\omega\|_2^2\right]$.

Putting together: with probability at least $1-\tau$

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| \leq$$

$$\leq \frac{128\sqrt{d|\mathcal{K}|}\left(S_{k,2\mathbf{p},2\mathbf{q}} C_{k,2\mathbf{p},2\mathbf{q}}\right)^{\frac{1}{4}} + S_{k,\mathbf{p},\mathbf{q}}\sqrt{\log\left(\frac{1}{\tau^2}\right)}}{\sqrt{m}}.$$

Step-6: finish

Putting together with probability at least $1-\tau$

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| \leq$$

$$\leq \underbrace{\frac{128\sqrt{d|\mathcal{K}|}\left(S_{k,2\mathbf{p},2\mathbf{q}}C_{k,2\mathbf{p},2\mathbf{q}}\right)^{\frac{1}{4}} + S_{k,\mathbf{p},\mathbf{q}}\sqrt{\log\left(\frac{1}{\tau^2}\right)}}{\sqrt{m}}}_{=:\epsilon}.$$

Equivalently

$$\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in\mathcal{K}} \left|\partial^{\mathbf{p},\mathbf{q}} k(\mathbf{x},\mathbf{y}) - s^{\mathbf{p},\mathbf{q}}(\mathbf{x},\mathbf{y})\right| \leq \epsilon\right) \geq$$

$$\geq 1 - e^{-\frac{1}{2}\left[\frac{\epsilon\sqrt{m} - 128\sqrt{d|\mathcal{K}|}\left(S_{k,2\mathbf{p},2\mathbf{q}}C_{k,2\mathbf{p},2\mathbf{q}}\right)^{\frac{1}{4}}}{S_{k,\mathbf{p},\mathbf{q}}}\right]^2}.$$

Assumptions

Let $S_{k,\mathbf{p},\mathbf{q}} := \sup_{\omega \in supp(\mathbb{S})} \left|\omega^{\mathbf{p}+\mathbf{q}}\right|$, $C_{k,\mathbf{p},\mathbf{q}} := \mathbb{E}_{\omega\sim\mathbb{S}}\left[\left|\omega^{\mathbf{p}+\mathbf{q}}\right| \|\omega\|_2^2\right]$.
Assumptions:

- k: continuous, shift-invariant.
- $C_{k,2\mathbf{p},2\mathbf{q}} < \infty$.
- If $[\mathbf{p};\mathbf{q}] \neq \mathbf{0}$: $supp(\mathbb{S})$ is bounded.
- $\mathcal{K} \subseteq \mathbb{R}^d$: convex and compact set.

Assumptions

Let $S_{k,\mathbf{p},\mathbf{q}} := \sup_{\omega \in supp(\mathbb{S})} |\omega^{\mathbf{p}+\mathbf{q}}|$, $C_{k,\mathbf{p},\mathbf{q}} := \mathbb{E}_{\omega \sim \mathbb{S}} \left[|\omega^{\mathbf{p}+\mathbf{q}}| \, \|\omega\|_2^2 \right]$.

Assumptions:

- k: continuous, shift-invariant.
- $C_{k,2\mathbf{p},2\mathbf{q}} < \infty$.
- If $[\mathbf{p}; \mathbf{q}] \neq \mathbf{0}$: $supp(\mathbb{S})$ is bounded.
- $\mathcal{K} \subseteq \mathbb{R}^d$: convex and compact set.

Notes:

- If
  - $\mathbf{p} = \mathbf{q} = 0$: $S_{k,\mathbf{p},\mathbf{q}} = S_{k,2\mathbf{p},2\mathbf{q}} = 1$,
  - else: $supp(\mathbb{S})$: bounded $\Rightarrow S_{k,\mathbf{p},\mathbf{q}}, S_{k,2\mathbf{p},2\mathbf{q}} < \infty$.
- $\mathcal{K}$: convex, compact $\Rightarrow \mathcal{K}_\Delta$ is so.

Summary

- We proved optimal rates for random kitchen sinks.
- Slightly annoying assumptions:
  - *supp*$(\mathbb{S})$: bounded,
  - compactness of $\mathcal{K}$.
- Other open questions:
  - metrizable LCA groups,
  - error propagation in specific tasks.

Thank you for the attention!

Cucker, F. and Smale, S. (2002).
On the mathematical foundations of learning.
Bulletin of the American Mathematical Society, 39:1–49.

Rahimi, A. and Recht, B. (2007).
Random features for large-scale kernel machines.
In *Neural Information Processing Systems (NIPS)*, pages 1177–1184.

Steinwart, I. and Christmann, A. (2008).
Support Vector Machines.
Springer.

Support of a measure

- Ingredients:
  - (X, τ): topological space with a countable basis.
  - $\mathcal{B} = \sigma(\tau)$: sigma-algebra generated by τ.
  - $\mathbb{S}$: measure on $(X, \mathcal{B})$.

  Then

  $$supp(\mathbb{S}) = \overline{\cup\{A \in \tau : \mathbb{S}(A) = 0\}},$$

  i.e., the complement of the union of all open $\mathbb{S}$-null sets.
- Our choice: $X = \mathbb{R}^d$.