UNIVERSITÄT
SALZBURG

Kernels, Sequences and Approximations

Peter Zinterhof

Technical Report 2016-05 August 2016

Department of Computer Sciences

Jakob-Haringer-Straße 2
5020 Salzburg
Austria
www.cosy.sbg.ac.at

Technical Report Series

Kernels, Sequences and Approximations
P. Zinterhof, Salzburg

A standard problem: Given a domain E and a function $f(x)$, find $x_1, ..., x_N$ such that

$$|\int_E f(x)d\mu(x) - \frac{1}{N}\sum_{n=1}^{N} f(x_n)| \quad \text{is small!}$$

The answer depends on the domain E and the class H of functions $f \in H$.

The problem of approximation and interpolation of functions is similar. Because of the generality of the questions the literature is knowingly infinite.

I consider the following problem: Given a domain E and a class of functions H and points $x_1, x_2, ...x_N \in E$. What can one say about the quality of the "mesh" $x_1, ..., x_N$ in view of integration, interpolation and approximation?

A possible starting point: let $E = [0, 1)^s, s = 1, 2, 3...$

Let $H_s^\alpha = \{f(x_1, x_2, ...x_s) = \sum_{m_1,...m_s=-\infty}^{\infty} \hat{f}(m_1, ...m_s) exp(2\pi i(m_1 x_1 + ... + m_s x_s))\}$,
such that for $\bar{m} = max(1, |m|), m \in \mathbb{Z}$, holds

$$|\hat{f}(m_1, ...m_s)| \leq \frac{c}{(\overline{m}_1, ...\overline{m}_s)^\alpha} = O\Big(\frac{1}{(\overline{m}_1, ...\overline{m}_s)^\alpha}\Big), \alpha > 1.$$

The exponent α is connected with differentiability properties of the function f. "Korobow-classes"

A mesh $\vec{x_1}, ..., \vec{x_N} \in E$ will be good, if it is "uniform distributed" in E. A classical measure for the quality of uniform distribution is essentially due to H. Weyl, who introduced the so called discrepancy, which we will not consider in this frame.

Some years ago I introduced the so called Diaphony F_N of the mesh $\vec{x_1}, ..., \vec{x_N} \in E[0, 1)^s$: $\vec{x_n} = (x_{1n}, ..., x_{sn}), n = 1, ..., N$:

$$F_N := \Big(\frac{1}{N^2}\sum_{k,l=1}^{N}\sum_{\substack{\vec{m}=-\infty \\ \vec{m}\neq\vec{0}}}^{\infty} \frac{1}{(\overline{m}_1, \overline{m}_2...\overline{m}_s)^2} exp(2\pi i(m_1(x_{1k} - x_{1l}) + ... + m_s(x_{sk} - x_{sl}))\Big)^{\frac{1}{2}}$$

It could be made much simpler:

$$\text{Let } B_2(x) = 1 - \frac{\pi^2}{6} + \frac{\pi^2}{2}(1 - 2\{x\})^2 = \sum_m \frac{1}{\overline{m}^2} exp(2\pi i x m)$$

We get now

$$F_N = (\frac{1}{N^2}\sum_{k,l=1}^{N}\prod_{j=1}^{s} B_2(x_{kj}-x_{lj})-1)^{\frac{1}{2}}$$

$$\text{Theorem:} \quad F_N \xrightarrow{N->\infty} 0 \Leftrightarrow (\vec{x_n})_{n=1}^{\infty} \text{ is u.d. mod 1}$$

$$\text{Theorem:} \quad |\frac{1}{N}\sum_{n=1}^{N} f(\vec{x_n}) - \int_E f(\vec{x})d\vec{x}| \le CF_N^{\alpha}$$

$$\text{Typically:} \quad F_N = O\Big(\frac{1}{N^{1-\epsilon}}\Big), \quad F_N = O\Big(\frac{\ln^s N}{N}\Big)$$

Very Bad: cassical cartesian product rules:

$$F_N = O\Big(\frac{1}{N^{\frac{1}{s}}}\Big) \quad \text{,it is sharp!}$$

curse of dimension

! The frequent occurrences of 2 and $\frac{1}{2}$ call for Hilbert space approach !!

Leitmotiv: Put $\varphi_n(x) = \frac{1}{n}exp(2\pi i n x), n \in \mathbb{Z}$,
and put $<\varphi_n, \varphi_m> = \delta_{nm}, H := \{f(x) = \sum_{n\in R} \tilde{f}_n\varphi_n(x)\}, \sum_n |\tilde{f}_n|^2 < \infty$,
and you get a Hilbert space of continuous functions. More general:

Let $E \neq \emptyset, K(x,y) : ExE \to \mathbb{C}$ strictly pos. def. matrix. Let $< K(x,y_1), K(x,y_2) >:= K(y_2,y_1)$, then $span\{a_1K(x,y_1)+a_2K(x,y_2)+...+a_nK(x,y_n)\} = H$ defines a Hilbertspace with rep. Kernel K(x,y):

$$< f(x), K(x,y) >= f(y), f \in H$$
$$f \Rightarrow f(x) \text{ being a cont. functional on H}$$

Bergman, Aronszajn, Moore, Zaremba, ...

We define a metric on $E := d(x,y) : ||K(t,x) - K(t,y)||$ and assume compactness of the metric space (E, d).
Let now $(x_n)_{n=1}^{\infty}$ be a sequence of points in $E, x_n \neq x_m, n \neq m$. We apply Gram-Schmidt to $K(x,x_1), K(x,x_2), ..., K(x,x_n), ...$ and get an orthonormal sequence $\tau_n(x), n \in \mathbb{N}$.
$(\tau_n(x))_{n=1}^{\infty}$ has the important property $\tau_m(x_n) = 0$ for $m > n$.

Let $f(x) = \sum_{n=1}^{\infty} \tilde{f}_n \tau_n(x) \in H$. We compute $\tilde{f}_n, n = 1, 2, ...$ recursively: $\tilde{f}_1 = f(x_1)/\tau_1(x_1)$, $\tilde{f}_2 = (f(x_2) - \tilde{f}_1\tau_1(x_2))/\tau_2(x_2), ... \tilde{f}_n = (f(x_n) - \tilde{f}_1\tau_1(x_n) - ... - \tilde{f}_{n-1}\tau_{n-1}(x_n))/\tau_n(x_n)$

The $\tau_n(x_n) \neq 0, n \in \mathbb{N}$, by Cholesky's theorem. e.g.

Main question: in which cases is the ONS $(\tau_n(x))_{n=1}^{\infty}$ an ONB:

$$f(x) = \sum_{n=1}^{\infty} \tilde{f}_n \tau_n(x) \text{ for all } f(x) \in H??$$

Theorem: $(\tau_n(x))_n$ is an ONB $\Leftrightarrow$ $(K(t, x_n))_n$ is total in H $\Leftrightarrow$ $(f(x_n) = 0 => f = 0 \quad \forall f \in H)$

Let now $K_N(x, y) := \sum_{n=1}^{N} \tau_n(x)\overline{\tau_n(y)}$, $K_N^{\perp}(x, y) = K(x, y) - K_N(x, y)$.

$K_N(x, y)$ reproduces $H_N = span\{K(x, x_1), ..., K(x, x_N)\}$, $K_N^{\perp}(x, y)$ reproduces $H \ominus H_N = H_N^{\perp}$.

Def: The N-th totality of $x_1, ...x_N, ...$ is $T_N = \max_{y \in E} ||K_N^{\perp}(x, y)|| = \max_{y} ||K_N(x, y)|| - ||K_N(x, y)|| = \max_{y}(K(y, y)^{\frac{1}{2}} - K_N(y, y)^{\frac{1}{2}}) = \max_{y} T_N(y)$

Theorem: $(x_n)_{n=1}^{\infty}$ is total in E iff $\lim_{N \to \infty} T_N = 0$.

The proof uses Arzela-Ascoli's theorem.

Hausdorff-Distances and Dispersion.

The Hausdorff-Distance between $\{x_1, ...x_N\} \subseteq E$ and E itself is defiend by

$$\delta(x_1, ...x_N) = \max_{x} \min_{n=1,...N} d(x, x_n)$$

Mostly $\delta(x_1, ...x_N)$ is called the dispersion of $x_1, ..., x_N$ in E. Well known: The sequence $(x_n)_{n=1}^{\infty}$ is dense in E iff $\lim_{N \to \infty} \delta_N = 0$

The following theorem holds:

$$0 \leq T_N \leq \delta_N$$

Corollary: If $(x_n)_{n=1}^{\infty}$ is dense in E, it is total as well. I have a Hilbert space and total sequences with arbitrary bad density properties. Furthermore I have a Hilbert space (E,K,H), where $T_N \to 0 \Leftrightarrow \delta_N \to 0$.

It means, a sequence $(x_n)_{n=1}^{\infty}$ is total iff it is dense in that special space E.

Quality of Itnerpolation:

Problem: We want to find a function $f_N(x)$, such that $f_N(x_n) = f(x_n), n = 1, ..., N$, and $f_N(x) \in span\{K(x, x_n), n = 1, ..., N\} = H_N$. Let $l_1(x), l_2(x), ..., l_H(x) \in H_N$, such that $l_m(x_n) = \delta_{nm}$.
The $l_1(x), ..., l_H(x)$ are the dual base to $K(x, x_1), ..., K(x, x_N)$.

So, $f_N(x) = \sum_{n=1}^{N} f(x_n) l_n(x) \in H_n$ fulfills the requirements. On the other hand $f(x_n) = \sum_{k=1}^{N} \tilde{f}_k \tau_k(x_n)$, because $\tau_k(x_n) = 0$ for $k > N$.

So we get $f_N(x) = \sum_{n=1}^{N} \tilde{f}_n \tau_n(x)$, because $f_N(x)$ and $f(x)$ coincide at $x_1, ..., x_N$. It follows:

$$|f(x) - f_N(x)| = |\sum_{n=N+1}^{\infty} \tilde{f}_n \tau_n(x)| \leq \Big(\sum_{n=N+1}^{\infty} |\tilde{f}_n|^2\Big)^{\frac{1}{2}} \Big(\sum_{n=N+1}^{\infty} |\tau_n(x)|^2\Big)^{\frac{1}{2}}$$

$$\max_{x \in E} = |f(x) - f_N(x)| \leq (\sum_{n=N+1}^{\infty} |\tilde{f}_n|^2)^{\frac{1}{2}} \quad T_N = ||f_N^{\perp}(x)|| \cdot T_N \leq ||f|| T_N$$

And

$$||f - f_N|| = (\sum_{n=N+1}^{\infty} |\tilde{f}_n|^2)^{\frac{1}{2}}$$

means, that the "Lagrange"-Interpolation function $f_N(x)$ is the best approximation in Hilbert space sense!

Computational Complexity:

The computation of $\tilde{f}_n, \tau_n(x), n = 1, ..., N, K_N(x, y)$ needs $O(N^2)$ operations.

Important special case: E = G, compact abelian group with Haar measure λ and $K(x, y) = k(x - y), k(x) = f(x) \star \overline{f(-x)}, f(x) \in L_2(G)$, and $x_n = x_1 \cdot n, n = 0, ..., N - 1$ is a cyclic subgroup of G. Then one can apply FFT with N log(N) operations.

I have similar results for numercical Integration of the type $I_g(f) =< f, g >$ using $f(x_1), ..., f(x_H)$.

"But that's another story"

would Rudyard Kipling say....

The Kernel $K_N(x, y) = \sum_{n=1}^{N} \tau_n(x) \overline{\tau_n(y)}$ plays the main role in interpolation and approximation of functions in H:

$$f_N(x) =< f(x), K_N(x, y) >$$

$K_N(x, y)$ reproduces the space $H_n = span\{K(x, x_1), ..., K(x, x_N)\}$. Let $Gram(K(x, x_1), ..., K(x, x_N)) = Gram_N$, then we can show, that

$$K_N(x,y) = (K(y,x_1),...,K(y,x_N)) \cdot Gram_N^{-1} \begin{pmatrix} K(x,x_1) \\ K(x,x_2) \\ ... \\ K(x,x_N) \end{pmatrix}$$

$Gram_N$ is not singular and pos. def.. So we have with a matrix $U, U^* = U^{-1}$, and $\lambda_1^2 \geq \lambda_2^2 \geq ... \geq \lambda_N^2 > 0$

$$Gram_N = U^* \begin{pmatrix} \lambda_1^2 & & \\ & ... & \\ & & \lambda_N^2 \end{pmatrix} U$$

and

$$K_N(x,y) = (K(y,x_1),...,K(y,x_N))U^* \begin{pmatrix} \frac{1}{\lambda_1^2} & & 0 \\ & ... & \\ 0 & & \frac{1}{\lambda_N^2} \end{pmatrix} U \begin{pmatrix} K(x,x_1) \\ ... \\ K(x,x_N) \end{pmatrix}$$

This means explicitly, that we have

$$K_N(x,y) = \sum_{n=1}^{N} \frac{L_n(x)\overline{L_n(y)}}{\lambda_n^2} \text{ with } L_n \perp L_m, ||L_n||^2 = \lambda_n^2$$

We know already: the "better" the $x_1, ..., x_N$ are distributed in E the better $K_N(x,y)$ will be suitable for interpolation and for integration.
So the behavior of the Eigenfunctions and Eigenvalues of $Gram_N$ play an important role. In many important cases $(E = [0,1)^s, E = c.a.g.)$ we have $\lambda_n^2 \sim N$. It means, that in this cases the arithmetic mean is asymptotically the best choice for numerical integration.

The End :)