カーネル法による数値積分とその理論

統計的機械学習研究センター 特任研究員 金川元信

共同研究者：Bharath K. Sriperumbudur (Penn. State U.) 福水健次 (統数研)

Task: Numerical Integration

$$\int_{[0,1]^d} f(x)dx \approx \frac{1}{n}\sum_{i=1}^n f(X_i)$$

Generate samples

$$X_1, \ldots, X_n \in [0,1]^d$$

to minimize the **worse case error**

$$\mathrm{e}_H(X_1, \ldots, X_n) := \sup_{f \in H, \|f\|_H \leq 1} \left| \int_{[0,1]^d} f(x)dx - \frac{1}{n}\sum_{i=1}^n f(X_i) \right|$$

H : reproducing kernel Hilbert space (RKHS)

$\|f\|_H$: norm (represents smoothness)

Problem: Misspecified Case

- To achieve the rate $O(n^{-\alpha-1/2+\varepsilon})$ QMC requires α to be known

  i.e. QMC uses the assumption $f \in H^\alpha$ to generate samples $X_1, \ldots, X_n$

- Question: what if the assumption is misspecified? i.e.

$$f \notin H^\alpha$$

- Existing theory does not provide convergence guarantees

Quasi Monte Carlo (QMC)

deterministically generate

$$X_1, \ldots, X_n \in [0,1]^d$$

so that

$$\mathrm{e}_{H^\alpha}(X_1, \ldots, X_n) = O(n^{-\alpha-1/2+\varepsilon})$$

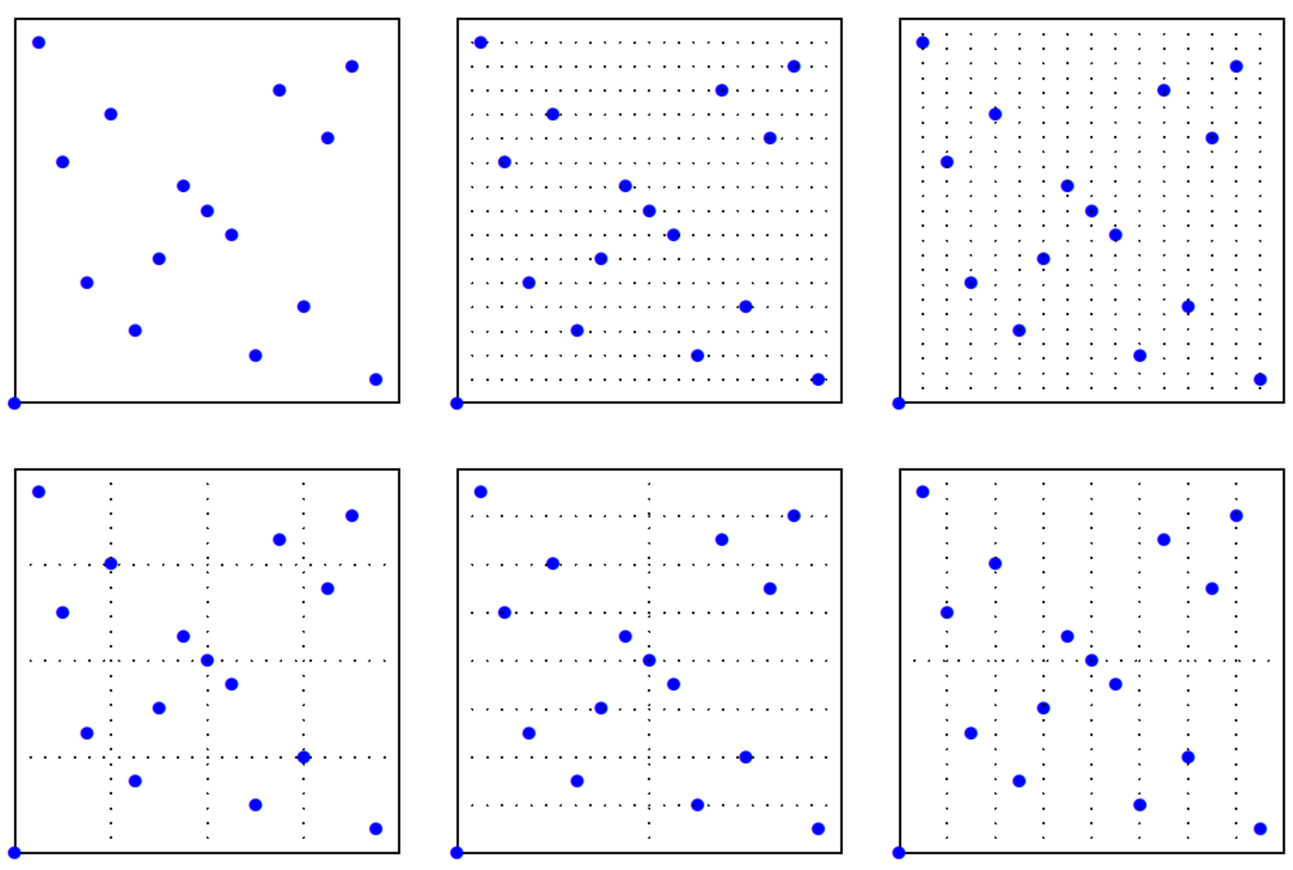
where

H^α : Sobolev RKHS of degree α (in each variable)

$\alpha \in \mathbb{N}$: smoothness of functions (num. of differentiability)

$\varepsilon > 0$: arbitrarily small constant

Illustration of QMC (taken from [1])
- every elementary interval of volume 1/16 contains exactly one point



Cf. Monte Carlo integration:

$$X_1, \ldots, X_n \sim \mathrm{unif}([0,1]^d)$$

$$\Rightarrow \mathrm{e}_{H^\alpha}(X_1, \ldots, X_n) = O_p(n^{-1/2})$$

Main Result

Define an integral operator

$$T : L_2([0,1]^d) \to L_2([0,1]^d)$$

$$Tg := \int_{[0,1]^d} k(\cdot, x) g(x) dx, \quad \forall g \in L_2([0,1]^d)$$

where $k(x', x)$ is the reproducing kernel of H^α

Theorem:

- Suppose $X_1, \ldots, X_n$ satisfy

$$\mathrm{e}_{H^\alpha}(X_1, \ldots, X_n) = O(n^{-\alpha-1/2+\varepsilon})$$

- Assume there is $\theta \in (0,1]$ such that

$$f \in \mathrm{Range}(T^{\theta/2})$$

- Then

$$\left| \int_{[0,1]^d} f(x)dx - \frac{1}{n}\sum_{i=1}^n f(X_i) \right| = O(n^{-\theta(\alpha+1/2-\varepsilon)})$$

Range assumption [2]:

$$\theta = 1 \Rightarrow \mathrm{Range}(T^{\theta/2}) = H^\alpha$$

$$\theta = 0 \Rightarrow \mathrm{Range}(T^{\theta/2}) = L_2([0,1]^d)$$

$$0 < \theta < 1 \Rightarrow \mathrm{Range}(T^{\theta/2}) = [L_2([0,1]^d), H^\alpha]_{\theta,2}$$

(Interpolation space)

References:

[1] Dick et al. (2013). High-dimensional integration: The quasi-Monte Carlo way. *Acta Numerica*, 22:133-288.

[2] Smale and Zhou (2007). Learning theory estimates via integral operators and their approximations. *Constructive Approximation*, 26:153–172.