Analysis of the stability and accuracy of multivariate polynomial approximation by discrete least squares with evaluations in random or low-discrepancy point sets

Giovanni Migliorati

MATHICSE-CSQI, École Polytechnique Fédérale de Lausanne

Analysis with random points: joint work with Fabio Nobile (EPFL), Raul Tempone (KAUST), Albert Cohen (UPMC), Abdellah Chkifa (UPMC) and Erik von Schwerin (KTH).

Analysis with low-discrepancy points: joint work with Fabio Nobile.

Outline

1. Discrete least squares on multivariate polynomial spaces
2. Stability and accuracy with evaluations in random points
3. Stability and accuracy with evaluations in low-discrepancy point sets
4. Conclusions

Notation and definitions

For any $d \geq 1$, $\Gamma := [-1,1]^d$ and any real numbers $\alpha, \beta > -1$, define

$$\rho(y) := \mathcal{B}(\alpha,\beta)^{-d} \prod_{i=1}^{d} (1-y_i)^{\alpha}(1+y_i)^{\beta}, \quad y \in \Gamma,$$

$$\langle f_1, f_2 \rangle_{L^2_\rho(\Gamma)} := \int_\Gamma f_1(y) f_2(y) \rho(y) dy, \quad \langle f_1, f_2 \rangle_M := \frac{1}{M} \sum_{m=1}^{M} f_1(y_m)\, f_2(y_m),$$

$$\| \cdot \|_{L^2_\rho} := \langle \cdot, \cdot \rangle_{L^2_\rho}^{1/2}, \quad \| \cdot \|_M := \langle \cdot, \cdot \rangle_M^{1/2},$$

with $y_1, \ldots, y_M$ being any points in Γ, either realizations of i.i.d. random variables $Y_1, \ldots, Y_M \overset{\text{i.i.d.}}{\sim} \rho$ or deterministically given (e.g. low-discrepancy point sets). Given univariate L^2_ρ-orthonormal polynomials $(\varphi_k)_{k \geq 0}$ and a multi-index set $\Lambda \subset \mathbb{N}_0^d$, for any $\nu \in \Lambda$ we define

$$\psi_\nu(y) := \prod_{i=1}^{d} \varphi_{\nu_i}(y_i), \quad y \in \Gamma,$$

$$\mathbb{P}_\Lambda := \mathrm{span}\left\{ \psi_\nu : \nu \in \Lambda \right\}.$$

Markov and Nikolskii inequalities for multivariate polynomials with downward closed multi-index sets

Definition (Downward closed multi-index set)

Λ is downward closed if $(\boldsymbol{\nu} \in \Lambda \quad \text{and} \quad \boldsymbol{\nu}' \leq \boldsymbol{\nu}) \Rightarrow \boldsymbol{\nu}' \in \Lambda$.

Lemma (M. 2014)

In any dimension, for any Λ downward closed and any $\alpha, \beta \in \mathbb{N}_0$ it holds

$$\|u\|^2_{L^\infty(\Gamma)} \leq (\#\Lambda)^{2\max\{\alpha,\beta\}+2} \|u\|^2_{L^2_\rho(\Gamma)}, \quad \forall u \in \mathbb{P}_\Lambda(\Gamma).$$

Lemma (M. 2014)

In any dimension and for any Λ downward closed, when $\alpha = \beta = 0$ (Legendre polynomials), it holds

$$\left\| \frac{\partial^d}{\partial y_1 \cdots \partial y_d} u \right\|^2_{L^2_\rho(\Gamma)} \leq 4^{-d} (\#\Lambda)^4 \|u\|^2_{L^2_\rho(\Gamma)}, \quad \forall u \in \mathbb{P}_\Lambda(\Gamma).$$

Discrete least squares on polynomial spaces

For any smooth (analytic) real-valued (or Hilbert-valued) function $\phi : \Gamma \to \mathbb{R}$, we define its continuous and discrete L^2 projections over $\mathbb{P}_\Lambda$ as

$$\Pi_\Lambda \phi := \underset{v \in \mathbb{P}_\Lambda}{\operatorname{argmin}} \|\phi - v\|_{L^2_\rho}, \qquad \Pi_\Lambda^M \phi := \underset{v \in \mathbb{P}_\Lambda}{\operatorname{argmin}} \|\phi - v\|_M.$$

Algebraic formulation: design matrix $[D]_{ij} = \psi_j(y_i)$, right-hand side $[b]_i = \phi(y_i)$, for any $i = 1, \ldots, M$ and $j = 1, \ldots, \#\Lambda$.

Normal equations:

$$D^\top D \, \boldsymbol{\beta} = D^\top b,$$

with $\boldsymbol{\beta}$ containing the coefficients of the expansion $\Pi_\Lambda^M \phi = \sum_{\nu \in \Lambda} \beta_\nu \psi_\nu$.

We define also the matrix $G := D^\top D / M$.

Optimality of discrete least squares in the L^2_ρ norm

In any dimension, with any index set Λ and any ρ with bounded support:

Proposition (M., Nobile, von Schwerin and Tempone, FoCM 2014)

For any (random or deterministic) choice of M points in Γ it holds

$$\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho} \leq \left(1 + \sqrt{|||G^{-1}|||}\right) \inf_{v \in \mathbb{P}_\Lambda} \|\phi - v\|_{L^\infty}.$$

Proof

Theorem (M., Nobile, von Schwerin and Tempone, FoCM 2014)

Given M points in Γ, being realizations of random variables independent and identically distributed w.r.t. ρ, it holds

$$\lim_{M \to +\infty} |||G^{-1}||| = \lim_{M \to +\infty} |||G||| = 1, \quad almost\ surely.$$

Proposition (M., Nobile, von Schwerin and Tempone, FoCM 2014)

$$cond(G) = |||G|||\ |||G^{-1}|||.$$

Norm equivalence on $\mathbb{P}_\Lambda$ (case of random points)

Find $\delta \in (0,1)$ such that

$$(1-\delta)\|v\|^2_{L^2_\rho} \leq \|v\|^2_M \leq (1+\delta)\|v\|^2_{L^2_\rho}, \quad \forall v \in \mathbb{P}_\Lambda,$$

with high probability.

Since $\|v\|^2_M = M^{-1}\langle D\mathbf{v}, D\mathbf{v}\rangle^2_{\mathbb{R}^{\#\Lambda}} = \langle G\mathbf{v}, \mathbf{v}\rangle^2_{\mathbb{R}^{\#\Lambda}}$ and $\|v\|^2_{L^2_\rho} = \langle \mathbf{v}, \mathbf{v}\rangle^2_{\mathbb{R}^{\#\Lambda}}$, the matrix G satisfies

$$|||G||| = \sup_{v\in\mathbb{P}_\Lambda\setminus\{v\equiv 0\}} \frac{\|v\|^2_M}{\|v\|^2_{L^2_\rho}}, \qquad |||G^{-1}||| = \sup_{v\in\mathbb{P}_\Lambda\setminus\{v\equiv 0\}} \frac{\|v\|^2_{L^2_\rho}}{\|v\|^2_M}.$$

Hence, norm equivalence on $\mathbb{P}_\Lambda$ w.h.p. iff concentration bounds

$$1-\delta \leq |||G||| \leq 1+\delta,$$
$$\frac{1}{1+\delta} \leq |||G^{-1}||| \leq \frac{1}{1-\delta},$$
$$|||G-\mathbb{I}||| \leq \delta,$$

again with high probability.

Given any L^2_ρ-orthonormal polynomial basis $(\psi_\nu)_{\nu\in\Lambda}$ of $\mathbb{P}_\Lambda$, define

$$K(\Lambda) := \sup_{y\in\Gamma}\left(\sum_{\nu\in\Lambda}|\psi_\nu(y)|^2\right) = \sup_{v\in\mathbb{P}_\Lambda}\frac{\|v\|^2_{L^\infty}}{\|v\|^2_{L^2_\rho}}.$$

Lemma (Chkifa, Cohen, M., Nobile and Tempone, 2013)

In any dimension and for any downward closed Λ it holds

$$K(\Lambda) \le (\#\Lambda)^{\ln 3/\ln 2}, \textit{with tensorized Chebyshev 1st kind polynomials.}$$

Lemma (M. 2014)

In any dimension, for any downward closed Λ and any $\alpha,\beta\in\mathbb{N}_0$ it holds

$$K(\Lambda) \le (\#\Lambda)^{2\max\{\alpha,\beta\}+2}, \textit{with tensorized Jacobi polynomials.}$$

These bounds are quite general, and set the ground for adaptive polynomial approximation based on discrete least squares.

Assume that $|\phi| \leq \tau$ almost surely w.r.t. ρ and define

$$T_\tau(t) := \mathrm{sign}(t) \min\{\tau, |t|\}, \qquad \widetilde{\Pi}_\Lambda^M := T_\tau(\Pi_\Lambda^M).$$

Theorem (Chkifa, Cohen, M., Nobile and Tempone, 2013)

For any $\gamma > 0$ and any downward closed Λ, if M is such that

$$K(\Lambda) \leq \frac{0.15}{1+\gamma} \frac{M}{\ln M}$$

then, for any $\phi \in L^\infty(\Gamma)$ with $\|\phi\|_{L^\infty} \leq \tau$, it holds that

$$Pr\,(cond(G) \leq 3) \geq 1 - 2M^{-\gamma},$$

$$Pr\left(\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho} \leq (1+\sqrt{2}) \inf_{v \in \mathbb{P}_\Lambda} \|\phi - v\|_{L^\infty} \right) \geq 1 - 2M^{-\gamma},$$

$$\mathbb{E}\left(\|\phi - \widetilde{\Pi}_\Lambda^M \phi\|^2_{L^2_\rho} \right) \leq \left(1 + \frac{0.6}{(1+\gamma)\ln M} \right) \|\phi - \Pi_\Lambda \phi\|^2_{L^2_\rho} + 8\tau^2 M^{-\gamma}.$$

($\delta = 1/2$ everywhere!)

Discrete least squares with deterministic points: multivariate case with Chebyshev density in $[0,1]^d$

Deterministic points introduced by Zhou, Narayan and Xu:

$$y_j = \cos\left(\frac{2\pi}{M}(j,\dots,j^d)\right) \in [-1,1]^d, \qquad j = 1,\dots,M,$$

asymptotically distributed according to the Chebyshev density.

Theorem (Zhou, Narayan and Xu, arXiv 2014)

In any dimension d and with the Chebyshev density, if M is a prime number and

$$M \geq 4^{d+1} d^2 (\#\Lambda)^2$$

then it holds that

$$\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho} \leq \left(1 + \frac{4}{d^2 \#\Lambda}\right) \inf_{v \in \mathbb{P}_\Lambda} \|\phi - v\|_{L^\infty}.$$

The proof uses arguments from number theory.

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Given any set of M points $y_1, \ldots, y_M \in [0,1]^d$ and any set $\emptyset \neq U \subseteq \{1, \ldots, d\}$, we define its local discrepancy

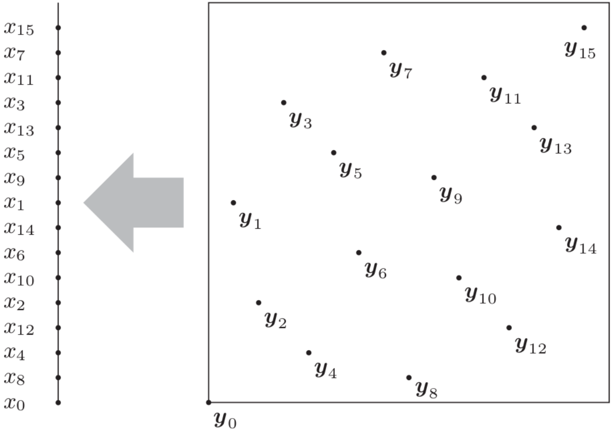
$$\Delta^U(t,1) := \frac{1}{M} \sum_{i=1}^{M} \prod_{q \in U} \mathbb{I}_{[0,t^q]}(y_i^q) - \prod_{q \in U} t^q, \qquad t \in [0,1]^{|U|},$$

and its star-discrepancy

$$D^{*,U} := \sup_{t \in [0,1]^{|U|}} |\Delta^U(t,1)|.$$

Values of components in $\{1, ..., d\} \setminus U$ are frozen to 1.

Example $d = 2$, $U = \{2\}$, $\{1,2\} \setminus U = \{1\}$



(picture from J.Dick, F.Pillichshammer: Digital Nets and Sequences, 2010)

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Let $t \geq 0$, $m \geq 1$, $d \geq 1$ and $b \geq 2$ be integers with $t \leq m$.

A (t, m, d)-net in base b is a point set consisting of b^m points in $[0,1)^d$ such that every elementary interval of the form

$$\prod_{q=1}^{d} \left[\frac{a_j}{b^{h_j}}, \frac{a_j + 1}{b^{h_j}} \right)$$

with each $h_j \geq 0$, $0 \leq a_j < b^{h_j}$ and $h_1 + \ldots + h_d = m - t$, contains exactly b^t points.

Example: $(0,4,2)$-net in base $b = 2$ (Hammersley points).

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Our analysis uses a type of Koksma-Hlawka inequality and low-discrepancy point sets. Starting point:

Lemma (Hlawka-Zaremba's identity)

Given M points $y_1, \dots, y_M \in [0,1]^d$, for any f with continuous mixed derivatives it holds

$$\left| \int_{[0,1]^d} f(y)dy - \frac{1}{M} \sum_{i=1}^{M} f(y_i) \right| = \sum_{\emptyset \neq U \subseteq \{1,\dots,d\}} (-1)^{|U|} \int_{[0,1]^{|U|}} \Delta^U(y^U, 1) \frac{\partial^{|U|}}{\partial y^U} f(y^U, 1)\, dy^U.$$

Lemma (Standard Koksma-Hlawka inequality)

$$\left| \int_{[0,1]^d} f(y)dy - \frac{1}{M} \sum_{i=1}^{M} f(y_i) \right| \leq D^{*,\{1,\dots,d\}} \|f\|_{HK}.$$

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Three main ingredients in our approach:

1) we prove a variant of the standard Koksma-Hlawka inequality starting from the Hlawka-Zaremba's identity:

Lemma (M., Nobile 2014)

$$\left| \|f\|^2_{L^2_\rho} - \|f\|^2_M \right| \leq \sum_{\emptyset \neq U \subseteq \{1,\dots,d\}} D^{*,U} \sum_{T \subseteq U} \left\| \frac{\partial^{|T|}}{\partial y^T} f(y^U, 1) \right\|_{L^2([0,1]^{|U|})} \left\| \frac{\partial^{|U|-|T|}}{\partial y^{U \setminus T}} f(y^T, 1) \right\|_{L^2([0,1]^{|U|})}.$$

2) Markov-type and Nikolskii-type multivariate inequalities for polynomials associated with downward closed multi-index sets (M. 2014).

3) upper bounds for the star-discrepancy of (t,m,d)-nets and (t,d)-sequences (e.g. Faure-Kritzer, Monatsh. Math. 2013).

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Consider any (t, m, d)-net in base $b \geq 2$ with quality parameter $t \geq 0$.

Theorem (M., Nobile 2014)

In any dimension d, with the uniform density and with anisotropic tensor product spaces $\mathbb{P}_\Lambda$, if

$$1 > \delta > \left(0.7\ b^t \exp\left\{\frac{b-1}{2\ln b}\right\} + \mathcal{O}(1)\right)(\#\Lambda)^2 \frac{(1+2\ln M)^d}{M}$$

then it holds that

$$cond(G) \leq \frac{1+\delta}{1-\delta},$$

$$\|\phi - \Pi_\Lambda^M \phi\|_{L^2_\rho} \leq \left(1 + \sqrt{\frac{1}{1-\delta}}\right) \inf_{v\in\mathbb{P}_\Lambda} \|\phi - v\|_{L^\infty}.$$

Similar theorem also for (t, d)-sequences (M.,Nobile 2014).

Discrete least squares with deterministic points: the multivariate case with uniform density in $[0,1]^d$

Given Λ downward closed and $U \subseteq \{1,\dots,d\}$ we define its sections by

$$\Lambda_U := \{\boldsymbol{\nu} \in \mathbb{N}_0^{|U|} : \exists\ \boldsymbol{\mu} = (\boldsymbol{\mu}_U, \boldsymbol{\mu}_{\{1,\dots,d\}\setminus U}) \in \Lambda \text{ and } \boldsymbol{\nu} = \boldsymbol{\mu}_U\}.$$

In general, for any downward closed multi-index set the condition becomes

$$1 > \delta > \min\left\{ (\#\Lambda)^4 \sum_{\emptyset \neq U \subseteq \{1,\dots,d\}} D^{*,U}, \right.$$
$$\left. \sum_{\emptyset \neq U \subseteq \{1,\dots,d\}} D^{*,U} \left(\#\Lambda_{\{1,\dots,d\}\setminus U}\right)^2 \sum_{T \subseteq U} \left(\#\Lambda_T\right)^2 \left(\#\Lambda_{U\setminus T}\right)^2 \right\}.$$

Nonoptimal when Λ is more sparse than anisotropic tensor product, compared to $M \propto (\#\Lambda)^2$ with random points and any Λ downward closed.

Conclusions (theoretical analysis)

RANDOM POINTS: analysis w.r.t. M, d, Λ, ρ, smoothness ϕ:

- in any dimension d, proven stability and accuracy provided that
  - $M/\ln M \geq C_1(\dim(\mathbb{P}_\Lambda))^{\frac{\ln 3}{\ln 2}}$ with Chebyshev density,
  - $M/\ln M \geq C_2(\dim(\mathbb{P}_\Lambda))^2$ with uniform density,
  - $M/\ln M \geq C_3(\dim(\mathbb{P}_\Lambda))^{2\max\{\alpha,\beta\}+2}$ with $\mathrm{beta}(\alpha+1,\beta+1)$, $\alpha,\beta \geq 0$,

  with the constants C_1, C_2, C_3 being independent of d.

DETERMINISTIC POINTS: analysis w.r.t. M, d, Λ, smoothness ϕ:

- in any dimension d, proven stability and accuracy provided that
  - $M \geq \widehat{C}_1(d)(\dim(\mathbb{P}_\Lambda))^2$ with Chebyshev density and any Λ (Zhou et al.),
  - $M/(1+2\ln M)^d \geq \widehat{C}_2(\dim(\mathbb{P}_\Lambda))^2$ with uniform density and anisotropic tensor product,
  - $M/(1+2\ln M)^d \geq \widehat{C}_3(\dim(\mathbb{P}_\Lambda))^\gamma$, $2 \leq \gamma \leq 4$ with uniform density and any Λ downward closed.

  with the constant $\widehat{C}_1$ being dependent on d, and $\widehat{C}_2$, $\widehat{C}_3$ being dependent on the parameters of the (t,m,d)-net or (t,d)-sequence.

Conclusions (experience from numerics)

- In high dimensions and with smooth functions, with both random and deterministic points, it seems to be enough

$$M \propto \dim(\mathbb{P}_\Lambda)$$

  to achieve the optimal convergence rate up to a threshold. A lot of numerical evidence, but no formal proof yet.
- Deterministic points CAN outperform random points in low dimensions. What about high dimensions?
- Discrete least squares is a well-promising approximation tool for multivariate aleatory functions and PDEs with stochastic data.

References on discrete least squares with RANDOM points

A.Cohen, M.Davenport, D.Leviatan: *On the stability and accuracy of least squares approximations*. Foundations of Computational Mathematics, 2013.

G.Migliorati, F.Nobile, E.von Schwerin, R.Tempone: *Analysis of discrete L^2 projection on polynomial spaces with random evaluations*. Foundations of Computational Mathematics, 2014.

A.Chkifa, A.Cohen, G.Migliorati, F.Nobile, R.Tempone: *Discrete least squares polynomial approximation with random evaluations; application to parametric and stochastic elliptic PDEs*. submitted. Available as MATHICSE report 35-2013.

G.Migliorati, F.Nobile, E.von Schwerin, R.Tempone: *Approximation of Quantities of Interest in stochastic PDEs by the random discrete L^2 projection on polynomial spaces*, SIAM J. Sci. Comput., 2013.

G.Migliorati: *Multivariate Markov-type and Nikolskii-type inequalities for polynomials associated with downward closed multi-index sets*, submitted. Available as MATHICSE report 1-2014.

G.Migliorati: *Polynomial approximation by the random discrete L^2 projection and application to inverse problems for PDEs with stochastic data*, PhD thesis, Department of Mathematics at Politecnico di Milano and Centre de Mathématiques Appliquées at École Polytechnique, 2013.

References on discrete least squares with DETERMINISTIC points

T.Zhou, A.Narayan, Z.Xu: *Multivariate discrete least-squares approximations with a new type of collocation grid*, arkiv:1401.0894v1, 2014.

G.Migliorati, F.Nobile: *Analysis of discrete least squares on multivariate polynomial spaces with evaluations in low-discrepancy point sets*, submitted. Available as MATHICSE report 25-2014.

Thank you for your attention!