On the Approximation Power of Local Least Squares Polynomials

Oleg Davydov

Universität Giessen, Mathematisches Institut, D-35392 Giessen, Germany.
oleg.davydov@math.uni-giessen.de
http://www.uni-giessen.de/~gcn5/davydov/

Abstract

We discuss the relationship between the norm of the local discrete least squares polynomial approximation operator, the minimal singular value $\sigma_{\min}(P_\Xi)$ of the matrix P_Ξ of the evaluations of the basis polynomials, and the norming constant of the set of data points Ξ with respect to the space of polynomials. Since these three quantities are equivalent up to bounded constants, and since $\sigma_{\min}(P_\Xi)$ can be efficiently computed, it is feasible to use $\sigma_{\min}(P_\Xi)$ as a tool for distinguishing good local point constellations, which is useful for scattered data fitting. In addition, we give a simple new proof of a bound by Reimer for the norm of the interpolation operators on the sphere and extend it to discrete least squares operators.

1 Introduction

Let Ω be a bounded domain in $\mathbb{R}^d$, $d \geq 1$, and let $\Xi = \{\xi_1, \ldots, \xi_m\}$ be a set of scattered points in Ω. Given the values $f|_\Xi = (f(\xi_1), \ldots, f(\xi_m))^T$ of an otherwise unknown function $f : \Omega \to \mathbb{R}$, we want to reconstruct f from these data. The *least squares method* consists in choosing some linear independent functions $p_1, \ldots, p_n$ on Ω, $n \leq m$, and computing the coefficients $a_1, \ldots, a_n \in \mathbb{R}$ that minimize the ℓ_2 norm of the residual on Ξ,

$$\|f|_\Xi - p|_\Xi\|_2 = \Big(\sum_{i=1}^{m} |f(\xi_i) - p(\xi_i)|^2 \Big)^{1/2},$$

with $p = a_1 p_1 + \cdots + a_n p_n \in \mathcal{P} := \operatorname{span}\{p_1, \ldots, p_n\}$. Let $\mathcal{P}|_\Xi := \operatorname{span}\{p_1|_\Xi, \ldots, p_n|_\Xi\}$. If $\dim \mathcal{P}|_\Xi = n$, then the least squares solution is unique, and we denote it by $L_{\mathcal{P},\Xi} f$. Note that the minimum norm solution available in the case of a rank deficient problem ($\dim \mathcal{P}|_\Xi < n$) seems less useful since in general it does not reproduce the elements of $\mathcal{P}$ exactly.

The computation of least squares approximation $L_{\mathcal{P},\Xi} f$ of f is expensive if m and n are large. To obtain a scattered data fitting algorithm with *linear complexity* with respect to the size of data, a *two-stage method* [8] can be employed which consists in 1) covering the original domain Ω with a number of subdomains Ω_k each containing only a small subset $\Xi_k = \Xi \cap \Omega_k$ of Ξ, computing *local* approximations to the data in Ξ_k, and 2) using the information obtained from these local approximations to build the final approximation of the (possibly huge) original data set. The least squares method can be employed in the local approximation stage, especially to deal with "real world" data usually contaminated with errors or just containing undesirable "high frequency" components.

If $\mathcal{P}$ is chosen to be the space Π_q^d of algebraic polynomials in d variables of a suitable degree q, then $n = \binom{d+q}{d}$. To achieve high approximation order, it is desirable to choose q such that n is only a little smaller then m. However, this is not always possible due to the rank deficiency or ill-conditioning of the least squares problem, which is especially difficult to control if $\xi_1, \ldots, \xi_m \in \Xi_k$ are unevenly distributed in Ω_k. This difficulty can in principle be overcome by constructing, for each Ξ_k, a suitable subspace of higher degree polynomials (least interpolation space [2]). If, however, the polynomial degree is not allowed to exceed a fixed small value, then a common practical approach is to choose larger sets $\Xi_k \subset \Xi$, with m substantially greater than n, see *e.g.* [4] where it is suggested to use for local least squares approximation $m = 11$ points if $\mathcal{P} = \Pi_2^2$ with $n = 6$ and $m = 15$ points if $\mathcal{P} = \Pi_3^2$ with $n = 10$. However, even these higher m provide no guaranty that the matrix

$$P_{\Xi_k} := [p_j(\xi_i) \, : \, i = 1, \ldots, m, \quad j = 1, \ldots, n]$$

of the local least squres problem will be always well-conditioned. Moreover, for some data, this method may lead to the use of inappropriately distant points for the local approximation.

The purpose of this paper is to draw attention to the fact that the conditioning of the matrix P_{Ξ_k} is not only the issue of numerical stability of the computation of least squares. Indeed, the reciprocal of the minimal singular value $\sigma_{\min}(P_\Xi)$ of P_Ξ provides a bound for the norm of the least squares operator $L_{\mathcal{P},\Xi}$ if both m and n are small. Therefore, the approximation power of local least squares depends on $\sigma_{\min}(P_\Xi)$ and the best approximation from $\mathcal{P}$. Since $\sigma_{\min}(P_\Xi)$ can be efficiently computed for a small matrix P_Ξ by well known numerical algorithms, it is feasible to use it as a tool to decide whether a particular portion of data is suitable for building local least squares approximation from $\mathcal{P}$ with reasonable approximation power. If $\sigma_{\min}(P_\Xi)$ is too small, then either Ξ or $\mathcal{P}$ should be modified, *e.g.* by adding more points to Ξ or using an appropriate subspace of $\mathcal{P}$. A two-stage algorithm for fitting large irregularly distributed scattered data sets employing the conditioning of the local observation matrices P_{Ξ_k} is studied in [3, 5].

The paper is organized as follows. In Section 2 we discuss the relationship between the norm of the discrete least squares approximation operator, the minimal singular value $\sigma_{\min}(P_\Xi)$, and the *norming constant* $\nu(\mathcal{P}, \Xi)$. As a by-product, we obtain a new proof of a known bound for the norm of the interpolation operators on the sphere [7], and extend it to the discrete least squares operators. Section 3 illustrates the above concepts in the univariate case, when they are also related to the *separation distance* of Ξ, while Section 4 is devoted to a discussion of the least squares multivariate polynomial approximation.

2 Bounds for $\|L_{\mathcal{P},\Xi}\|$ and approximation error

Let $p_1, \ldots, p_n$ be linearly independent continuous functions on $\Omega \subset \mathbb{R}^d$ spanning a linear space $\mathcal{P}$. Since all norms on a finite dimensional linear space are equivalent, there are positive constants K_1, K_2 such that

$$K_1\|a\|_2 \leq \|\sum_{j=1}^{n} a_j p_j\|_{C(\Omega)} \leq K_2\|a\|_2 \tag{2.1}$$

for any coefficient vector $a = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$.

Given $\Xi = \{\xi_1, \ldots, \xi_m\} \subset \Omega$, we consider the matrix $P_\Xi \in \mathbb{R}^{m\times n}$ as defined in the introduction. Obviously, $\operatorname{rank} P_\Xi = \dim \mathcal{P}|_\Xi$. If P_Ξ has full rank, then $\dim \mathcal{P}|_\Xi = n$, and the least squares approximation $L_{\mathcal{P},\Xi}f$ is uniquely determined, giving rise to the operator $L_{\mathcal{P},\Xi} : C(\Omega) \to \mathcal{P} \subset C(\Omega)$.

It is easy to see that $L_{\mathcal{P},\Xi}$ exactly reproduces the elements of $\mathcal{P}$, *i.e.*,

$$L_{\mathcal{P},\Xi}p = p, \qquad \text{all } p \in \mathcal{P}. \tag{2.2}$$

Therefore, a standard argument shows that

$$\|f - L_{\mathcal{P},\Xi}f\|_{C(\Omega)} \leq (1 + \|L_{\mathcal{P},\Xi}\|)E(f,\mathcal{P})_{C(\Omega)}, \tag{2.3}$$

where $E(f,\mathcal{P})_{C(\Omega)}$ denotes the error of the best approximation of f from $\mathcal{P}$ in Chebyshev norm,

$$E(f,\mathcal{P})_{C(\Omega)} := \inf_{p\in\mathcal{P}} \|f - p\|_{C(\Omega)}.$$

Thus, an estimate for $\|L_{\mathcal{P},\Xi}\|$ immediately gives an upper bound for $\|f - L_{\mathcal{P},\Xi}f\|_{C(\Omega)}$.

The *norming constant* $\nu(\mathcal{P},\Xi)$ of Ξ with respect to $\mathcal{P}$ [6] can be defined by

$$\nu(\mathcal{P},\Xi) = \min_{p\in\mathcal{P}} \|p|_\Xi\|_\infty / \|p\|_{C(\Omega)}. \tag{2.4}$$

Given any matrix A, we denote by $\sigma_{\min}(A)$ the *minimal singular value* of it,

$$\sigma_{\min}(A) = \min_{\|x\|_2=1} \|Ax\|_2.$$

Recall that if A has full rank, then $\sigma_{\min}(A) = \|A^+\|_2^{-1}$, where A^+ is the pseudoinverse of A, see *e.g.* [1].

Theorem 2.1 *If* $\operatorname{rank} P_\Xi = n$, *then*

$$K_1/\sigma_{\min}(P_\Xi) \leq \quad \|L_{\mathcal{P},\Xi}\| \quad \leq K_2\sqrt{m}/\sigma_{\min}(P_\Xi), \tag{2.5}$$

$$1/\nu(\mathcal{P},\Xi) \leq \quad \|L_{\mathcal{P},\Xi}\| \quad \leq \sqrt{m}/\nu(\mathcal{P},\Xi), \tag{2.6}$$

$$K_1\nu(\mathcal{P},\Xi) \leq \quad \sigma_{\min}(P_\Xi) \quad \leq K_2\sqrt{m}\,\nu(\mathcal{P},\Xi). \tag{2.7}$$

Proof: We first prove (2.5). Let $L_{\mathcal{P},\Xi}f = \sum_{j=1}^n a_j p_j$. It follows by a well-known result in numerical linear algebra that the vector $a = (a_1, \ldots, a_n)^T$ can be computed as the product of the pseudoinverse P_Ξ^+ of P_Ξ with the vector $f|_\Xi$. Therefore,

$$\|a\|_2 = \|P_\Xi^+ f|_\Xi\|_2 \leq \|P_\Xi^+\|_2 \|f|_\Xi\|_2 = \sigma_{\min}^{-1}(P_\Xi)\|f|_\Xi\|_2.$$

Since $\|L_{\mathcal{P},\Xi}f\|_{C(\Omega)} \le K_2\|a\|_2$ and $\|f|_\Xi\|_2 \le \sqrt{m}\,\|f|_\Xi\|_\infty \le \sqrt{m}\,\|f\|_{C(\Omega)}$, the upper bound in (2.5) follows. To prove the lower bound in (2.5), we choose a function $\tilde{f} \in C(\Omega)$ such that

$$\|P_\Xi^+ \tilde{f}|_\Xi\|_2 = \|P_\Xi^+\|_2\|\tilde{f}|_\Xi\|_2, \quad \|\tilde{f}|_\Xi\|_\infty = \|\tilde{f}\|_{C(\Omega)},$$

which is obviously possible. Then by (2.1) we have

$$\|L_{\mathcal{P},\Xi}\tilde{f}\|_{C(\Omega)} \ge K_1\|P_\Xi^+\tilde{f}|_\Xi\|_2 = K_1\sigma_{\min}^{-1}(P_\Xi)\|\tilde{f}|_\Xi\|_2,$$

which implies the desired lower bound since $\|\tilde{f}|_\Xi\|_2 \ge \|\tilde{f}|_\Xi\|_\infty = \|\tilde{f}\|_{C(\Omega)}$.

Since $\|L_{\mathcal{P},\Xi}f\|_{C(\Omega)} \le \nu^{-1}(\mathcal{P},\Xi)\|(L_{\mathcal{P},\Xi}f)|_\Xi\|_\infty$, the upper bound in (2.6) follows by

$$\|(L_{\mathcal{P},\Xi}f)|_\Xi\|_\infty \le \|(L_{\mathcal{P},\Xi}f)|_\Xi\|_2 \le \|f|_\Xi\|_2 \le \sqrt{m}\|f\|_{C(\Omega)}.$$

To prove the lower bound, we denote by $\tilde{p}$ an element of $\mathcal{P}$ for which the minimum in (2.4) is attained and choose a function $\tilde{f} \in C(\Omega)$ such that $\tilde{f}|_\Xi = \tilde{p}|_\Xi$ and $\|\tilde{f}\|_{C(\Omega)} = \|\tilde{f}|_\Xi\|_\infty$. Then by (2.2),

$$\|L_{\mathcal{P},\Xi}\tilde{f}\|_{C(\Omega)} = \|\tilde{p}\|_{C(\Omega)} = \nu^{-1}(\mathcal{P},\Xi)\|\tilde{p}|_\Xi\|_\infty = \nu^{-1}(\mathcal{P},\Xi)\|\tilde{f}\|_{C(\Omega)},$$

which implies $\|L_{\mathcal{P},\Xi}\| \ge \nu^{-1}(\mathcal{P},\Xi)$.

We finally establish (2.7). For any $p \in \mathcal{P}$, let $p = \sum_{j=1}^n a_j p_j$ and $a = (a_1,\ldots,a_n)^T$. Then $p|_\Xi = P_\Xi a$ and hence

$$\|p|_\Xi\|_\infty \le \|P_\Xi a\|_2 \le \sqrt{m}\,\|p|_\Xi\|_\infty.$$

Since

$$\sigma_{\min}(P_\Xi) = \min_{a\in\mathbb{R}^n}\|P_\Xi a\|_2/\|a\|_2,$$

(2.7) follows by (2.1). □

In view of (2.3), the upper bound in (2.5) implies

$$\|f - L_{\mathcal{P},\Xi}f\|_{C(\Omega)} \le (1 + K_2\sqrt{m}/\sigma_{\min}(P_\Xi))E(f,\mathcal{P})_{C(\Omega)}, \tag{2.8}$$

which shows that the approximation power of discrete least squares proportionally reduces if $\sigma_{\min}(P_\Xi)$ (or $\nu(\mathcal{P},\Xi)$) is small. We will discuss some practical consequences of this fact in the next two sections.

Although $\nu(\mathcal{P},\Xi)$ gives tighter bounds for $\|L_{\mathcal{P},\Xi}\|$, $\sigma_{\min}(P_\Xi)$ has a clear practical advantage that it is easily computable by using *e.g.* the singular value decomposition of the small "local" matrix P_Ξ. On the other hand, the norming constants were used in [6, 9] to derive estimates for the approximation error of radial basis function interpolation and moving least squares, respectively.

Remark 2.2 If $p_1,\ldots,p_n$ is an *orthonormal basis* for $\mathcal{P}$, then $\|a\|_2 = \|p\|_{L_2(\Omega)}$, $p = \sum_{j=1}^n a_j p_j$, and the constants K_1, K_2 in (2.1) are closely related to *Nikolskii constants* of the space $\mathcal{P}$, namely,

$$K_1 = N_{2,\infty}^{-1}(\mathcal{P}), \qquad K_2 = N_{\infty,2}(\mathcal{P}),$$

where

$$N_{q_1,q_2}(\mathcal{P}) := \max_{p\in\mathcal{P}}\|p\|_{L_{q_1}(\Omega)}/\|p\|_{L_{q_2}(\Omega)}, \qquad 1 \le q_1, q_2 \le \infty.$$

In particular, if $\Omega = S^{d-1}$, the unit sphere in $\mathbb{R}^d$, and $\{p_1, \ldots, p_n\}$ is the set of *spherical harmonics* forming an orthonormal basis for the space $\mathcal{P} = \mathcal{H}_q^d$ of spherical polynomials of degree q in d variables, then it is not difficult to prove that $K_2 = N_{\infty,2}(\mathcal{H}_q^d) = \sqrt{n/|S^{d-1}|}$, where $|S^{d-1}|$ denotes the surface area of S^{d-1}. Therefore, for any set $\Xi \subset S^{d-1}$ with $\#\Xi = m \geq n$, we have by (2.5),

$$\|L_{\mathcal{H}_q^d,\Xi}\| \leq \sqrt{nm/|S^{d-1}|}/\sigma_{\min}(P_\Xi), \tag{2.9}$$

which recovers in the case of interpolation ($m = n$) an error bound by Reimer [7] originally proved by using Lagrangian square sums (see also [10]).

3 Univariate polynomials

Let Ω be an interval $[-h, h]$ on the real line $\mathbb{R}$, and let

$$p_j(t) = (t/h)^{j-1}, \qquad j = 1, \ldots, n.$$

Then $\mathcal{P}$ is the restriction to $[-h, h]$ of the space Π_{n-1}^1 of all univariate polynomials of degree at most $n-1$. By the well-known interpolation properties of the univariate polynomials, $\operatorname{rank} P_\Xi = n$ for any $\Xi = \{\xi_1, \ldots, \xi_m\} \subset [-h, h]$, $m \geq n$, with distinct ξ_i's.

For any $\Xi' = \{\xi_{i_1}, \ldots, \xi_{i_n}\} \subset \Xi$, let $q_{\Xi'}$ denote the *separation distance*,

$$q_{\Xi'} := \frac{1}{2} \min_{j \neq k} |\xi_{i_j} - \xi_{i_k}|.$$

The Lebesgue constant $\|L_{\mathcal{P},\Xi'}\|$ of the corresponding interpolation scheme can be easily estimated as

$$\|L_{\mathcal{P},\Xi'}\| \leq \frac{2^{n-1}}{(n-1)!}(h/q_{\Xi'})^{n-1}.$$

Since Ξ' may be any subset of Ξ of cardinality n and since $\nu(\mathcal{P}, \Xi) \geq \|L_{\mathcal{P},\Xi'}\|^{-1}$, we get

$$\nu^{-1}(\mathcal{P}, \Xi) \leq \frac{2^{n-1}}{(n-1)!}(h/q_{\Xi,n})^{n-1},$$

where

$$q_{\Xi,n} := \max_{\substack{\Xi' \subset \Xi \\ \#\Xi' = n}} q_{\Xi'}.$$

Hence, by (2.3) and (2.6),

$$\|f - L_{\Pi_{n-1}^1,\Xi} f\|_{C[-h,h]} \leq \Big(1 + \frac{\sqrt{m}\, 2^{n-1}}{(n-1)!}(h/q_{\Xi,n})^{n-1}\Big) E(f, \Pi_{n-1}^1)_{C[-h,h]}. \tag{3.1}$$

This last estimate shows that the univariate least squares polynomials have the approximation power of the best local polynomial approximation as $h \to 0$ provided $h/q_{\Xi,n}$ remains bounded. However, if the scattered points $\xi_1, \ldots, \xi_m \in [-h, h]$ are clustered together in at most $n-1$ very tight groups, then $q_{\Xi,n}$ may be arbitrarily small, thus forcing the right hand side of (3.1) to blow up. To figure out what happens to $\|f - L_{\Pi_{n-1}^1,\Xi} f\|_{C[-h,h]}$ in these circumstances, we consider the following example.

Let $h=1$, $n=2$, $f(t)=t^2-1/2$, and $\Xi=\{-\xi,0,\xi\}$ for some $0<\xi\le 1$. It is easy to see that $L_{\Pi_1^1,\Xi}f\equiv -1/2+2\xi^2/3$. Since $E(f,\Pi_1^1)_{C[-1,1]}=1/2$, we have

$$\|f-L_{\Pi_1^1,\Xi}f\|_{C[-1,1]}=1/2+|1/2-2\xi^2/3|\le 2E(f,\Pi_1^1)_{C[-1,1]}$$

even though, by a simple calculation,

$$\|L_{\Pi_1^1,\Xi}\|=1/3+1/\xi,$$

$$\sqrt{2}/\sigma_{\min}(P_\Xi)=1/\nu(\mathcal{P},\Xi)=1/q_{\Xi,2}=1/\xi\to\infty \quad \text{as} \quad \xi\to 0.$$

This may contribute to the opinion that $\|L_{\Pi_1^1,\Xi}\|$, $\sigma_{\min}(P_\Xi)$, $\nu(\mathcal{P},\Xi)$ and $q_{\Xi,n}$ are not the right quantities to describe the behaviour of the approximation. Indeed, as the three points $-\xi,0,\xi$ coalesce, $L_{\Pi_1^1,\Xi}f$ converges to a Hermite interpolation polynomial provided the entries of P_Ξ as well as the values of $f|_\Xi$ are exact. However, if we simulate "real world" data by adding to $f(-\xi),f(0),f(\xi)$ normally distributed errors with standard deviation 10^{-4}, then the picture substantially changes. Table 1 shows that $\|f-L_{\Pi_1^1,\Xi}f\|_{C[-1,1]}$ does blow up in this case. For comparison we also include in the table the error of $\|f-L_{\Pi_0^1,\Xi}f\|_{C[-1,1]}$ for the same contaminated data.

Table 1 Average (d^1_{mean}) and maximum ($d^1_{\max}$) of $\|f-L_{\Pi_1^1,\Xi}f\|_{C[-1,1]}$ as well as maximum of $\|f-L_{\Pi_0^1,\Xi}f\|_{C[-1,1]}$ ($d^0_{\max}$) in 1000 tests with contaminated data

ξ	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
d^1_{mean}	1.06	1.56	6.63	57.3	564	5630
$d^1_{\max}$	1.24	3.39	24.9	240	2390	23900
$d^0_{\max}$	1.00018	1.00018	1.00018	1.00018	1.00018	1.00018

Thus, if $q_{\Xi,n}$ is too small, we cannot practically achieve with least squares the approximation order of $E(f,\Pi^1_{n-1})_{C[-h,h]}$ simply because the points lying too close to each other carry redundant information and we have at most $n-1$ clusters of such points. Therefore, we should adjust the polynomial degree to the given data paying attention to the trade-off between higher approximation power of higher degree polynomials and the "pollution" caused by the factor $q_{\Xi,n}^{-1}$ that increases with n. In practice one may choose maximal n such that $h/q_{\Xi,n}$ is smaller than a prescribed tolerance value $0<E<\infty$.

4 Multivariate polynomials

The situation becomes substantially more complicated when we turn to multivariate polynomials. Let Ω be a bounded domain in $\mathbb{R}^d$ and let $\{p_1,\dots,p_n\}$, $n=\binom{d+q}{d}$, be a basis of the space $\mathcal{P}=\Pi_q^d$ of polynomials in d variables of total degree q satisfying (2.1) on Ω. (For example, we may consider a properly scaled standard power basis with the center at a point in Ω or the Bernstein-Bézier basis with respect to some simplex overlapping Ω or a significant part of it.) Let, furthermore, Ξ be an arbitrary finite set of points in Ω such that $m=\#\Xi\ge n$.

The first problem we face in the case $d\ge 2$ is that the matrix P_Ξ may be rank deficient. It is clear, however, that there is no practical difference between this situation and the one when P_Ξ has full rank but is extremely ill-conditioned, *i.e.*, $\sigma_{\min}(P_\Xi)$ is very small. Moreover, (2.8) shows that even moderately small $\sigma_{\min}(P_\Xi)$ may significantly reduce the approximation power of $L_{\mathcal{P},\Xi}$. Clearly, the same can also happen in the univariate case if $q_{\Xi,n}$ is too small. The real difficulty of the multivariate case seems to be that simple characteristics of Ξ, like separation distance $q_{\Xi,n}$, do not give much information about the norm of $L_{\mathcal{P},\Xi}$. For example, six equidistant points on the unit circle in $\mathbb{R}^2$ are well separated and look reasonably distributed. However, they are not good for least squares approximation from the space Π_2^2 since the matrix P_Ξ is rank deficient. Suitably perturbed, these points will give rise to the least squares operator $L_{\Pi_2^2,\Xi}$ with a very large norm. More generally, the norm of $L_{\Pi_q^d,\Xi}$ will be large if the points in $\Xi \subset \mathbb{R}^d$ lie "too close" to an algebraic hypersurface of order q.

If the data is comparatively dense in Ω, namely the *fill distance*

$$h_{\Xi,\Omega} := \sup_{x\in\Omega} \min_{\xi\in\Xi} |x-\xi|$$

does not exceed some small positive constant depending on Ω and the polynomial degree, then the estimates of the norming constant $\nu(\Pi_q^d,\Xi)$ given in [9] provide a bound for $\|L_{\Pi_q^d|_\Omega,\Xi}\|$, in view of (2.6). For example, if Ω is a ball of radius r, then $\nu(\Pi_q^d|_\Omega,\Xi) \geq 1/2$ if $h_{\Xi,\Omega} < 0.11r/q^2$.

On the other hand, without any density assumptions we can always rely on (2.8), where $\sigma_{\min}(P_\Xi)$ can be efficiently computed by well known algorithms of numerical linear algebra. In some sense, small $\sigma_{\min}(P_\Xi)$ indicates that the local data has "hidden redundancies" (*e.g.* too many points lying very close to the same straight line or the same ellipse) that prevent it from carrying enough information for a "full power" approximation of the underlying function from Π_q^d. Similar to the univariate case, but using $\sigma_{\min}(P_\Xi)$ instead of $q_{\Xi,n}$, we can adaptively choose the polynomial degree according to the following algorithm that has proven to be useful for scattered data fitting [3, 5].

Let $\Omega \subset \mathbb{R}^d$, $\Xi \subset \Omega$, $\#\Xi = m$. Denote by P_Ξ^q the matrix of the evaluations of appropriate basis functions for Π_q^d, $q \geq 0$, at the points $\xi \in \Xi$.

Algorithm 4.1 *Starting with some $q = q_0 \geq 0$ such that $\binom{d+q}{d} \leq m$, compute $\sigma_{\min}(P_\Xi^q)$. If $1/\sigma_{\min}(P_\Xi^q)$ is smaller than a prescribed tolerance $E < \infty$, then compute the least squares Π_q^d-approximation to the data in Ξ and accept it as a reliable approximation on Ω. Otherwise, repeat the same with $q = q_0 - 1$ and succesively reduce the degree q to $q_0 - 2, \ldots, 0$, while $1/\sigma_{\min}(P_\Xi^q) \geq E$. For $q = 0$ no comparison of $1/\sigma_{\min}(P_\Xi^0)$ with E is needed since $\|L_{\Pi_0^d|_\Omega,\Xi}\|$ is bounded for any Ω and Ξ.*

Note that, optionally, the *condition number* $\|P_\Xi^q\|_2/\sigma_{\min}(P_\Xi^q)$ of P_Ξ^q can be used in the above algorithm instead of $1/\sigma_{\min}(P_\Xi^q)$, as it has been formulated in [5].

Bibliography

1. Å. Björck, *Numerical Methods for Least Squares Problems,* SIAM, Philadelphia, 1996.

2. C. de Boor and A. Ron, On multivariate polynomial interpolation, *Constr. Approx.* **6** (1990), 287–302.
3. O. Davydov and F. Zeilfelder, Scattered data fitting by direct extension of local polynomials with bivariate splines, 2002, preprint.
4. T. A. Foley, Scattered data interpolation and approximation with error bounds, *Comput. Aided Geom. Design* **3** (1986), 163–177.
5. J. Haber, F. Zeilfelder, O. Davydov and H.-P. Seidel, Smooth approximation and rendering of large scattered data sets, in *Proceedings of IEEE Visualization 2001,* Th. Ertl, K. Joy and A. Varshney (eds), IEEE, 2001, 341–347, 571.
6. K. Jetter, J. Stöckler and J. Ward, Error estimates for scattered data interpolation on spheres, *Math. Comp.* **226** (1999), 733–747.
7. M. Reimer, Interpolation on the sphere and bounds for the Lagrangian square sums, *Results in Mathematics* **11** (1987), 144–164.
8. L. L. Schumaker, Two-stage methods for fitting surfaces to scattered data, in *Quantative Approximation,* R. Schaback and K. Scherer (eds), Lecture Notes 556, Springer, Berlin, 1976, 378–389.
9. H. Wendland, Local polynomial reproduction and moving least squares approximation, *IMA J. Numer. Anal.* **21** (2001), 285–300.
10. R. S. Womersley and I. H. Sloan, How good can polynomial interpolation on the sphere be?, *Advances in Comp. Math.* **14** (2001), 195–226.