Communications in Applied Analysis, **22**, No. 1 (2018), 35-42 ISSN: 1083-2564

AN ERROR ESTIMATION IN BACKUS-GILBER MOVING LEAST-SQUARES APPROXIMATION

SVETOSLAV NENOV

Department of Mathematics
University of Chemical Technology and Metallurgy
Sofia 1756, BULGARIA

ABSTRACT: Our goal is to establish a bound of the error in moving least-squares approximation in terms of coefficients of the approximation and error of nearest data points.

AMS Subject Classification: 93E24

Key Words: moving least-squares approximation, error estimation, Backus-Gilber approach, singular values, largest singular value

Received: May 11, 2017; **Accepted:** September 16, 2017;
Published: September 17, 2018. **doi:** 10.12732/caa.v22i1.3
Dynamic Publishers, Inc., Acad. Publishers, Ltd. http://www.acadsol.eu/caa

1. STATEMENT

Let us remind the definition of moving least-squares approximation and a basic result.

Let:

1. $\mathcal{D}$ be a compact in $\mathbb{R}^d$.

2. $\boldsymbol{x}_i \in \mathcal{D}$, $i = 1, \ldots, m$; $\boldsymbol{x}_i \neq \boldsymbol{x}_j$, if $i \neq j$.

3. $f : \mathcal{D} \to \mathbb{R}$ be a continuous function.

4. $p_i : \mathcal{D} \to \mathbb{R}$ be continuous functions, $i = 1, \ldots, l$. The functions $\{p_1, \ldots, p_l\}$ are linearly independent in $\mathcal{D}$ and let $\mathcal{P}_l$ be their linear span.

5. $W : [0, \infty) \to (0, \infty)$ be a continuous and positive function.

Usually the basis in $\mathcal{P}_l$ is constructed by monomials. For example: $p_l(\boldsymbol{x}) = x_1^{k_1} \dots x_d^{k_d}$, where $\boldsymbol{x} = (x_1, \dots, x_d)$, $k_1, \dots k_d \in \mathbb{N}$, $k_1 + \dots + k_d \leq l - 1$. In this case, we will use the notation $\boldsymbol{x}^{\boldsymbol{k}}$, $|\boldsymbol{k}| \leq l - 1$. In the case $d = 1$, the standard basis is $\{1, x, \dots, x^{l-1}\}$.

Following [3], we will use the following definition. The *moving least-squares approximation* of order l at a fixed point $\boldsymbol{x}$ is the value of $p^*(\boldsymbol{x})$, where $p^* \in \mathcal{P}_l$ is minimizing the least-squares error

$$\sum_{i=1}^{m} W(\|\boldsymbol{x} - \boldsymbol{x}_i\|)\,(p(\boldsymbol{x}) - f(\boldsymbol{x}_i))^2$$

among all $p \in \mathcal{P}_l$.

The approximation is "local" if weight function W is fast decreasing as its argument tends to infinity and interpolation is achieved if $W(0) = \infty$. We define additional function $w : [0, \infty) \to [0, \infty)$, such taht:

$$w(r) = \begin{cases} \dfrac{1}{W(r)}, & \text{if } r > 0, \\ \dfrac{1}{W(r)}, & \text{if } r = 0 \text{ and } W(0) < \infty, \\ 0, & \text{if } r = 0 \text{ and } W(0) = \infty. \end{cases}$$

Some examples of $W(r)$ and $w(r)$, $r \geq 0$:

$$\begin{array}{ll} W(r) = e^{-\alpha^2 r^2} & \text{exp-weight,} \\ W(r) = r^{-\alpha^2} & \text{Shepard weights,} \\ W(r) = r^2 e^{-\alpha^2 r^2} & \text{McLain weight,} \\ w(r) = e^{\alpha^2 r^2} - 1 & \text{see Levin's works.} \end{array}$$

Below: $\|\cdot\| = \|\cdot\|_2$ is the 2-norm, $\|\cdot\|_1$ is 1-norm in $\mathbb{R}^d$; the superscript ^{t} denotes transpose of real matrix; $\boldsymbol{I}$ is the identity matrix. Let

$$\boldsymbol{E} = \begin{pmatrix} p_1(\boldsymbol{x}_1) & p_2(\boldsymbol{x}_1) & \cdots & p_l(\boldsymbol{x}_1) \\ p_1(\boldsymbol{x}_2) & p_2(\boldsymbol{x}_2) & \cdots & p_l(\boldsymbol{x}_2) \\ \vdots & \vdots & & \vdots \\ p_1(\boldsymbol{x}_m) & p_2(\boldsymbol{x}_m) & \cdots & p_l(\boldsymbol{x}_m) \end{pmatrix},$$

$$\boldsymbol{D} = 2\begin{pmatrix} w(\|\boldsymbol{x}-\boldsymbol{x}_1\|) & 0 & \cdots & 0 \\ 0 & w(\|\boldsymbol{x}-\boldsymbol{x}_2\|) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w(\|\boldsymbol{x}-\boldsymbol{x}_m\|) \end{pmatrix},$$

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \ \boldsymbol{c} = \begin{pmatrix} p_1(\boldsymbol{x}) \\ p_2(\boldsymbol{x}) \\ \vdots \\ p_l(\boldsymbol{x}) \end{pmatrix}, \ \boldsymbol{F} = \begin{pmatrix} f(\boldsymbol{x}_1) \\ f(\boldsymbol{x}_2) \\ \vdots \\ f(\boldsymbol{x}_m) \end{pmatrix}.$$

Through the article, we assume the following conditions (H1):

(H1.1) $1 \in \mathcal{P}_l$.

(H1.2) $1 \leq l \leq m$.

(H1.3) $\mathrm{rank}(\boldsymbol{E}) = l$.

(H1.4) w is a smooth function.

(H1.5) The basis in $\mathcal{P}_l$ is

$$\left\{\boldsymbol{x}^{\boldsymbol{k}} : |\boldsymbol{k}| \leq l-1\right\} = \left\{x_1^{k_1} \dots x_d^{k_d} : k_1 + \cdots + k_d \leq l-1\right\}.$$

The equivalent statement of the moving least-squares minimization problem is the following constrained problem:

$$\text{Minimum of quadratic form } Q = \sum_{i=1}^{m} w(\boldsymbol{x}, \boldsymbol{x}_i) a_i^2, \tag{1.1}$$

$$\text{subject to } \sum_{i=1}^{m} a_i p_j(\boldsymbol{x}_i) = p_j(\boldsymbol{x}), \ j = 1, \dots l. \tag{1.2}$$

Theorem 1.1 (see [3]). Let the conditions (H1.1)-(H1.4) hold true. Then:

1. The matrix $\boldsymbol{E}^t \boldsymbol{D}^{-1} \boldsymbol{E}$ is non-singular.

2. The approximation defined by the moving least-squares method is

$$\hat{L}(f) = \langle \boldsymbol{a}, \boldsymbol{F} \rangle = \sum_{i=1}^{m} a_i f(\boldsymbol{x}_i), \tag{1.3}$$

where

$$\boldsymbol{a} = \boldsymbol{A}_0 \boldsymbol{c} \quad \text{and} \quad \boldsymbol{A}_0 = \boldsymbol{D}^{-1} \boldsymbol{E} \left(\boldsymbol{E}^t \boldsymbol{D}^{-1} \boldsymbol{E} \right)^{-1}. \tag{1.4}$$

3. If $w(0) = 0$, then the approximation is interpolatory.

Remark 1.1. *Let us mark:* $\boldsymbol{D}$, $\boldsymbol{c}$, *and* $\boldsymbol{a}$ *depend on* $\boldsymbol{x}$; *the matrix* $\boldsymbol{E}$ *and the vector* $\boldsymbol{F}$ *do not depend* $\boldsymbol{x}$.

2. APPROXIMATION ORDER

Let the hypothesis (H1) be valid, $\boldsymbol{x}$ be a fixed point, $\boldsymbol{x} \neq \boldsymbol{x}_i$, $i = 1, \ldots, m$, and let an integer k_0 (not unique in general case) be choosen such that

$$h = \|\boldsymbol{x} - \boldsymbol{x}_{k_0}\| = \min \{ \|\boldsymbol{x} - \boldsymbol{x}_i\| : i = 1, \ldots, m \}.$$

Remark 2.1. *Let us mark the following reproduction property: for all* $p \in \mathcal{P}_l$, *we have*

$$\sum_{i=1}^{m} p(\boldsymbol{x}_i) a_i = p(\boldsymbol{x}).$$

or if we set $\boldsymbol{P} = \begin{pmatrix} p(\boldsymbol{x}_1) & p(\boldsymbol{x}_2) & \cdots & p(\boldsymbol{x}_m) \end{pmatrix}^t$ *then*

$$p(\boldsymbol{x}) = \langle \boldsymbol{a}, \boldsymbol{P} \rangle.$$

Indeed, let $p \in \mathcal{P}_l$. *Then there exist constants* $b_1, \ldots, b_l$ *such that*

$$p(\boldsymbol{x}) = \sum_{k=1}^{l} b_k p_k(\boldsymbol{x})$$

and using (1.2), *we receive*

$$\begin{aligned} p(\boldsymbol{x}) &= \sum_{k=1}^{l} b_k \sum_{i=1}^{m} a_i p_k(\boldsymbol{x}_i) = \sum_{i=1}^{m} a_i \sum_{k=1}^{l} b_k p_k(\boldsymbol{x}_i) \\ &= \sum_{i=1}^{m} a_i p(\boldsymbol{x}_i) = \langle \boldsymbol{a}, \boldsymbol{P} \rangle. \end{aligned}$$

Let $p \in \mathcal{P}_l$. Combining the previous remark and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |f(\boldsymbol{x}) - \hat{L}(f)(\boldsymbol{x})| &= |f(\boldsymbol{x}) - p(\boldsymbol{x}) + p(\boldsymbol{x}) - \langle \boldsymbol{a}, \boldsymbol{F} \rangle| \\ &\leq |f(\boldsymbol{x}) - p(\boldsymbol{x})| + |\langle \boldsymbol{a}, \boldsymbol{P} \rangle - \langle \boldsymbol{a}, \boldsymbol{F} \rangle| \\ &= |f(\boldsymbol{x}) - p(\boldsymbol{x})| + |\langle \boldsymbol{a}, \boldsymbol{P} - \boldsymbol{F} \rangle| \\ &\leq |f(\boldsymbol{x}) - p(\boldsymbol{x})| + \|\boldsymbol{a}\| \, \|\boldsymbol{P} - \boldsymbol{F}\| \,. \end{aligned}$$

Introducing the classical infinity norm ($\mathcal{D}$ is a compact subset in $\mathbb{R}^d$):

$$\|f - p\|_\infty = \max \{ |f(\boldsymbol{x}) - p(\boldsymbol{x})| : \boldsymbol{x} \in \mathcal{D} \} \,,$$

we have:

$$\begin{aligned} |f(\boldsymbol{x}) - p(\boldsymbol{x})| &\leq \|f - p\|_\infty \,, \\ \|\boldsymbol{P} - \boldsymbol{F}\|_2^2 &= \sum_{i=1}^{l} (p(x_i) - f(x_i))^2 \\ &\leq m \max \left\{ (p(x_i) - f(x_i))^2 : i = 1, \dots, l \right\} \\ &= m \left(\max \{ |p(x_i) - f(x_i)| : i = 1, \dots, l \} \right)^2 \\ &= m \, \|f - p\|_\infty^2 \,. \end{aligned}$$

Therefore, for any $p \in \mathcal{P}_l$ we have

$$|f(\boldsymbol{x}) - \hat{L}(f)(\boldsymbol{x})| \leq (1 + \sqrt{m} \, \|\boldsymbol{a}\|) \, \|f - p\|_\infty \,, \tag{2.1}$$

or we have to analyze the following two terms:

1. The infinity norm of $f - p$. In Subsection 2.1, we will construct an upper bound of $\|f - p\|_\infty$, following [1].

2. The norm of Lebesgue function $\|\boldsymbol{a}\|_1$. In Subsection 2.2 we will construct an upper estimate for the Lebesgue function $\|\boldsymbol{a}\|_1$ or for the equivalent norm $\|\boldsymbol{a}\|_2$ in the case of Backus-Gilber aproach.

2.1. THE NORM $\|f - p\|_\infty$

Let f be $(N + 1)$-times continuously differentiable function in $\mathcal{D}$ and let all derivatives of f be bounded functions in $\mathcal{D}$.

The inequality (2.1) is valid for any polinomial $p \in \mathcal{P}_l$. Then it is valid for the Taylor's polinomial for f at $\boldsymbol{x}$ of degree N on the one dimensional segment with endpoints $\boldsymbol{x}$ and $\boldsymbol{x}_1$. Hence there exists a constant $C_{k_0} > 0$ such that

$$\|f - p\|_\infty \leq C_{k_0} \|\boldsymbol{x} - \boldsymbol{x}_{k_0}\|^{N+1}.$$

Then

$$|f(\boldsymbol{x}) - \hat{L}(f)(\boldsymbol{x})| \leq \left(1 + \sqrt{m}\,\|\boldsymbol{a}\|\right) C_{k_0} h^{N+1}. \tag{2.2}$$

Obviously, we may use

$$C_{k_0} = \frac{1}{(N+1)!} \max \left\{ \left| f_\alpha^{(N+1)} \left(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{x}_{k_0}\right) \right| : \alpha \in [0,1] \right\},$$

but in this case C_{k_0} depends on $\boldsymbol{x}$. To avoid this, let $C_{k_0} = \max\{C_{k_0}(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{D}\}$.

2.2. BACKUS-GILBER APROACH

The Backus-Gilber aproach of moving-least squares method is exactly the constrained problem (1.1), (1.2) with $l = 1$ and $p_1(\boldsymbol{x}) = 1$, see [3, Section 2].

In this case, we have the following result.

Theorem 2.1. Let the conditions (H1) hold. Let f be $(N+1)$-times continuously differentiable function in $\mathcal{D}$ and let all derivatives of f be bounded functions in $\mathcal{D}$.

Then for any $\boldsymbol{x} \in \mathcal{D}$ the following inequality holds

$$\left| f(\boldsymbol{x}) - \hat{L}(f)(\boldsymbol{x}) \right| \leq (1 + \|\boldsymbol{c}\|)\, C_{k_0} h^{N+1}. \tag{2.3}$$

Proof. Let

$$\boldsymbol{A} = \boldsymbol{D}^{-1}\boldsymbol{E}\left(\boldsymbol{E}^t\boldsymbol{D}^{-1}\boldsymbol{E}\right)^{-1}\boldsymbol{E}^t = \boldsymbol{A}_0\boldsymbol{E}^t.$$

Then (using [2, P15-SVD], let us remind that $l \leq m$)

$$\sigma_{\max}(\boldsymbol{A}_0)\,\sigma_{\min}\left(\boldsymbol{E}^t\right) \leq \sigma_{\max}\left(\boldsymbol{A}_0\boldsymbol{E}^t\right) = \sigma_{\max}(\boldsymbol{A}).$$

But (see [4])

$$\sigma_{\max}(\boldsymbol{A}) \leq \sqrt{m}.$$

On the other hand, it is not hard to calculate the singular value decomposition of the vector $\boldsymbol{E}$:

$$\boldsymbol{E} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^t,$$

where:

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{e} & \widetilde{\boldsymbol{U}} \end{pmatrix}, \qquad \boldsymbol{e} = \frac{\boldsymbol{E}}{\|\boldsymbol{E}\|_2} = \frac{1}{\sqrt{m}}\boldsymbol{E};$$

$\widetilde{\boldsymbol{U}}$ is the $(m \times (m-1))$-matrix with orthonormal columns and such that $\widetilde{\boldsymbol{U}}^t \boldsymbol{E} = \boldsymbol{0}$;

$$\boldsymbol{\Sigma} = \begin{pmatrix} \|\boldsymbol{E}\|_2 & 0 & \cdots & 0 \end{pmatrix}^t = \begin{pmatrix} \sqrt{m} & 0 & \cdots & 0 \end{pmatrix}^t, \qquad \boldsymbol{V} = \begin{pmatrix} 1 \end{pmatrix}.$$

Hence $\sigma_{\min}\left(\boldsymbol{E}^t\right) = \sigma_{\min}\left(\boldsymbol{E}\right) = \sqrt{m}$ and

$$\|\boldsymbol{A}_0\| = \sigma_{\max}\left(\boldsymbol{A}_0\right) \leq \frac{\sqrt{m}}{\sigma_{\min}\left(\boldsymbol{E}^t\right)} = 1.$$

Therefore

$$\|\boldsymbol{a}\| = \|\boldsymbol{A}_0\boldsymbol{c}\| \leq \|\boldsymbol{A}_0\| \, \|\boldsymbol{c}\| = \|\boldsymbol{c}\| .$$

The inequality (2.3) follows from (2.2). □

REFERENCES

[1] G. Fasshauer, *Multivariate Mesh-free Approximation*, http://www.math.iit.edu/∼fass/603_ch7.pdf

[2] Faryar Jabbari, *Linear System Theory II*, Chapter 3: Eigenvalue, singular values, pseudo-inverse, The Henry Samueli School of Engineering, University of California, Irvine (2015), http://gram.eng.uci.edu/∼fjabbari/me270b/me270b.html.

[3] D. Levin, The approximation power of moving least-squares, http://www.math.tau.ac.il/∼levin/

[4] Svetoslav I. Nenov, An upper estimate for the maximal singular value of a special matrix, I, *International Journal of Pure and Applied Mathematics*, **115**, No. 4 (2017), 771-775, doi10.12732/ijpam.v115i4.10.

[5] Svetoslav Nenov, An upper estimate for the maximal singular value of a special matrix, III, *International Journal of Pure and Applied Mathematics*, To appear.

[6] Kaare Brandt Petersen, Michael Syskind Pedersen, *The Matrix Cookbook*, Version: February 16, 2006, http://www.mit.edu/∼wingated/stuff_i_use/matrix_cookbook.pdf