A orthonormal basis for Radial Basis Function approximation

9th ISAAC Congress
Krakow, August 5 - 9, 2013

Gabriele Santin,
joint work with S. De Marchi

Department of Mathematics.
Doctoral School in Mathematical Sciences,
Applied Mathematics Area
1 Data: $\Omega \subset \mathbb{R}^s$, $X \subset \Omega$, test function $f$

- $\Omega \subset \mathbb{R}^s$, 
- $X = \{x_1, \ldots, x_n\} \subset \Omega$
- $f_1, \ldots, f_n$, where $f_i = f(x_i)$ for a suitable $f$
1 Data: \( \Omega \subset \mathbb{R}^s, X \subset \Omega, \) test function \( f \)

- \( \Omega \subset \mathbb{R}^s, \)
- \( X = \{x_1, \ldots, x_n\} \subset \Omega \)
- \( f_1, \ldots, f_n, \) where \( f_i = f(x_i) \) for a suitable \( f \)

2 Interp. setting:

- \( V(\Omega) = \text{span}\{v_1, \ldots, v_n\} \subset C(\Omega) \)
- \( s_f(x) = \sum_{j=1}^{n} \alpha_j v_j(x) \in V(\Omega) \)
1 Data: $\Omega \subset \mathbb{R}^s$, $X \subset \Omega$, test function $f$

- $\Omega \subset \mathbb{R}^s$,
- $X = \{x_1, \ldots, x_n\} \subset \Omega$
- $f_1, \ldots, f_n$, where $f_i = f(x_i)$ for a suitable $f$

2 Interp. setting:

- $V(\Omega) = \text{span}\{v_1, \ldots, v_n\} \subset C(\Omega)$
- $s_f(x) = \sum_{j=1}^{n} \alpha_j \, v_j(x) \in V(\Omega)$

3 Aim: $s_f = f$ on $X$, i.e. $(v_j(x_i))_{ij} \cdot \alpha_j = f(x_i)$
\( V(\Omega) \) has to be an **Haar space**, i.e. \( \det((v_j(x_i))_{ij}) \neq 0 \ \forall X \subset \Omega. \)
$V(\Omega)$ has to be an Haar space, i.e. $\det((v_j(x_i))_{ij}) \neq 0 \forall X \subset \Omega$.

**Haar-Mairhuber-Curtis**

There is no $n$-Haar spaces of functions in $\mathbb{R}^s$ if $s > 1$ and $n > 1$.
\( V(\Omega) \) has to be an Haar space, i.e. \( \det((v_j(x_i))_{ij}) \neq 0 \ \forall X \subset \Omega \).

Haar-Mairhuber-Curtis

There is no \( n \)-Haar spaces of functions in \( \mathbb{R}^s \) if \( s > 1 \) and \( n > 1 \).

Data-dependent basis
Kernel $K \in C(\Omega \times \Omega)$, basis the translates $\{K(\cdot, x_1), \ldots, K(\cdot, x_n)\}$. 

Kernel $K \in C(\Omega \times \Omega)$, basis the translates $\{K(\cdot, x_1), \ldots, K(\cdot, x_n)\}$.

- $K$ positive definite (p.d. kernel matrix $A_{ij} = K(x_i, x_j) \in \mathbb{R}^{n \times n}$)
- $K$ radial, i.e. $K(\cdot, \cdot) = \phi(\varepsilon \| \cdot - \cdot \|_2)$, $\phi \in C([0, \infty], \mathbb{R})$
Kernel $K \in C(\Omega \times \Omega)$, basis the translates $\{K(\cdot, x_1), \ldots, K(\cdot, x_n)\}$.

- $K$ positive definite (p.d. kernel matrix $A_{ij} = K(x_i, x_j) \in \mathbb{R}^{n \times n}$)
- $K$ radial, i.e. $K(\cdot, \cdot) = \phi(\varepsilon \| \cdot - \cdot \|_2)$, $\phi \in C([0, \infty], \mathbb{R})$

Are there spaces where this is in fact an approximation scheme?
Kernel $K \in C(\Omega \times \Omega)$, basis the translates $\{K(\cdot, x_1), \ldots, K(\cdot, x_n)\}$.

- $K$ positive definite (p.d. kernel matrix $A_{ij} = K(x_i, x_j) \in \mathbb{R}^{n \times n}$)
- $K$ radial, i.e. $K(\cdot, \cdot) = \phi(\varepsilon \| \cdot - \cdot \|_2)$, $\phi \in C([0, \infty], \mathbb{R})$

Are there spaces where this is in fact an approximation scheme?

**Native Hilbert space**

$N_K(\Omega) := \text{span}\{K(\cdot, x), x \in \Omega\}$

$(f, g)$ induced by $A$.

- it is the unique space where $K$ is the reproducing kernel
- $N_K(X) := \text{span}\{K(\cdot, x) : x \in X\}$ is a finite subspace of $N_K(\Omega)$
- the interpolant operator $f \mapsto s_f$ is a projection
- $\|f - s_f\| \to 0$ as $n \to \infty$
**Problem:** the standard basis of $\mathcal{N}_K(X)$ is unstable and not flexible

**Question 1**
Is it possible to find a “better” basis $\mathcal{U}$ of $\mathcal{N}_K(X)$?

**Question 2**
How to embed informations about $K$ and $\Omega$ in $\mathcal{U}$?

**Question 3**
Can we extract $\mathcal{U}' \subset \mathcal{U}$ s.t. $s'_f$ is as good as $s_f$?
Q1: It is possible to find a “better” basis?

Change of basis (Pazouki-Schaback 2011):

Let $A_{ij} = K(x_i, x_j) \in \mathbb{R}^{n \times n}$. Any basis $\mathcal{U}$ arises from a factorization $A = V_{\mathcal{U}} \cdot C_{\mathcal{U}}^{-1}$, where $C_{\mathcal{U}}$ is the matrix of change of basis, $V_{\mathcal{U}} = (u_j(x_i))_{1 \leq i, j \leq n}$.

Each $\mathcal{N}_K(\Omega)$-orthonormal basis $\mathcal{U}$ arises from a decomposition $A = B^T \cdot B$ with $V_{\mathcal{U}} = B^T$ and $C_{\mathcal{U}} = B^{-1}$.

Each $\ell_2(X)$-orthonormal basis $\mathcal{U}$ arises from a decomposition $A = Q \cdot B$ with $Q^T \cdot Q = I$, $V_{\mathcal{U}} = Q$ and $C_{\mathcal{U}} = B$.

The best bases in terms of stability are the $\mathcal{N}_K(\Omega)$-o.n. ones.
A “natural” basis:

Let $K$ be a continuous, positive definite kernel on a bounded set $\Omega \subset \mathbb{R}^s$. The operator

$$
\mathcal{T} : \mathcal{N}_K(\Omega) \to \mathcal{N}_K(\Omega), \quad \mathcal{T}[f] = \int_{\Omega} K(\cdot, y)f(y)dy
$$

is bounded, compact and self-adjoint. It has an enumerable set of eigenvalues $\{\lambda_j\}_{j>0}$ and e.vectors $\{\varphi_j\}_{j>0}$ which are a basis for $\mathcal{N}_K(\Omega)$. Moreover

$$
\{\varphi_j\}_{j>0} \quad \text{are orthonormal in } \mathcal{N}_K(\Omega) \\
\{\varphi_j\}_{j>0} \quad \text{are orthogonal in } L_2(\Omega), \quad \|\varphi_j\|_{L_2(\Omega)}^2 = \lambda_j
$$

$$
\lambda_j \xrightarrow{j \to \infty} 0, \quad \sum_{j>0} \lambda_j = K(0,0)|\Omega|
$$
Q2: How to embed informations about $K$ and $\Omega$ in $U$?

Idea: Approximate the integral equation $\lambda_j \varphi_j(x) = \mathcal{T}[\varphi_i](x)$ with the symmetric Nyström method, with a cubature formula $(X, W)$:

\{\lambda_j, \varphi_j\}_{j>0}$ are approximated by e.values/e.vectors of $\sqrt{W} \cdot A \cdot \sqrt{W}$.
**Q2:** How to embed informations about $K$ and $\Omega$ in $U$?

**Idea:** Approximate the integral equation $\lambda_j \varphi_j(x) = T[\varphi_i](x)$ with the symmetric Nyström method, with a cubature formula $(X, W)$: \(\{\lambda_j, \varphi_j\}_{j>0}\) are approximated by e.values/e.vectors of $\sqrt{W} \cdot A \cdot \sqrt{W}$.

**Definition:**

A weighted SVD basis $U$ is a basis for $\mathcal{N}_K(X)$ s.t.

\[
V_u = \sqrt{W}^{-1} \cdot Q \cdot \Sigma, \quad C_u = \sqrt{W} \cdot Q \cdot \Sigma^{-1}
\]

where $\sqrt{W} \cdot A \cdot \sqrt{W} = Q \cdot \Sigma^2 \cdot Q^T$ is a SVD (and unitary diagon.).
This basis is in fact an approximation of the “natural” one (provided $w_i > 0$, $\sum_{i=1}^{n} w_i = |\Omega|$):

**Properties:**

- $\sigma_j^2 \ u_j(x) = T_n[u_j](x) = (u_j, K(\cdot, x))_{\ell_2,w(x)} \ \forall x \in \Omega, \ 1 \leq j \leq n$

- $N_K(\Omega)$-orthonormal

- $\ell_{2,w}(X)$-orthogonal, $\|u_j\|_{\ell_{2,w}(X)}^2 = \sigma_j^2, \ 1 \leq j \leq n$

- $\sum_{j=1}^{n} \sigma_j^2 = K(0, 0) |\Omega|$
Interpolant: \( s_f(x) = \sum_{j=1}^{n} (f, u_j)u_j(x) \quad \forall x \in \Omega \)

WDLS: \( s_f^m := \arg\min \left\{ \|f - g\|_{\ell^2, w(x)} : g \in \text{span}\{u_1, \ldots, u_m\} \right\} \)

WLDS as truncation:

Let \( f \in \mathcal{N}_K(\Omega), \ 1 \leq m \leq n. \) Then \( \forall x \in \Omega \)

\[
 s_f^m(x) = \sum_{j=1}^{m} \frac{(f, u_j)_{\ell^2, w(x)}}{(u_j, u_j)_{\ell^2, w(x)}} u_j(x) = \sum_{j=1}^{m} (f, u_j)u_j(x)
\]
WSVD Basis
Approximation

Interpolant: \( s_f(x) = \sum_{j=1}^{n} (f, u_j) u_j(x) \quad \forall x \in \Omega \)

WDLS: \( s^m_f := \argmin \left\{ \|f - g\|_{\ell^2, w(x)} : g \in \text{span}\{u_1, \ldots, u_m\} \right\} \)

**WLDS as truncation:**

Let \( f \in \mathcal{N}_k(\Omega), 1 \leq m \leq n. \) Then \( \forall x \in \Omega \)

\[
{s^m_f (x) = \sum_{j=1}^{m} \frac{(f, u_j)_{\ell^2, w(x)}}{(u_j, u_j)_{\ell^2, w(x)}} u_j(x) = \sum_{j=1}^{m} (f, u_j) u_j(x)}
\]

**Q3:** Can we extract \( \mathcal{U}' \subset \mathcal{U} \) s.t. \( s'_f \) is as good as \( s_f \)?

We can take \( \mathcal{U}' = \{u_1, \ldots, u_m\} \)
Q3: Can we extract $\mathcal{U}' \subset \mathcal{U}$ s.t. $s'_f$ is as good as $s_f$?

We can take $\mathcal{U}' = \{u_1, \ldots, u_m\}$

- recall that $\|u_j\|_{\ell_2^w(X)} = \sigma_j^2 \to 0$
- we can choose $m$ s.t. $\sigma_{m+1}^2 < \text{tol}$
- we don’t need $u_j, j > m$
If we define the pseudo-cardinal functions as $\tilde{\ell}_i = s_{\ell_i}^m$, we get

$$\tilde{\ell}_i(x) = \sum_{j=1}^{m} \frac{u_j(x_i)}{\sigma_j^2} u_j(x), \quad s_f^m(x) = \sum_{i=1}^{n} f(x_i) \tilde{\ell}_i(x).$$

**Generalized Power Function and Lebesgue constant:**

If $f \in N_K(\Omega)$, $|f(x) - s_f^m(x)| \leq P_{K,X}^{(m)}(x)\|f\|_{N_K(\Omega)} \quad \forall x \in \Omega$, where

$$P_{K,X}^{(m)}(x) = K(x, x) - \sum_{j=1}^{m} u_j(x)^2.$$

Moreover, $\|s_f^m\|_\infty \leq \tilde{\Lambda}_X \|f\|_X.$
Figure: Franke’s test function, \textit{lens}, IMQ Kernel, \(\varepsilon = 1\) (RMSE)
WSVD Basis
An Example

Figure: Franke’s test function, \textit{lens}, IMQ Kernel, \( \varepsilon = 1 \) (RMSE)

Problem: We have to compute the whole basis before the truncation → Krylov methods
Consider $A_{ij} = K(x_i, x_j)$, $b_i = f(x_i)$, $1 \leq i, j \leq n$

- define the Krylov subspace
  $\mathcal{K}_m(A, b) = \text{span}\{b, Ab, \ldots, A^{m-1}b\}$

- compute a o.n. basis $\{p_1, \ldots, p_m\}$ of $\mathcal{K}_m(A, b)$

- define the (tridiagonal) matrix $H_m$ which represents the projection of $A$ into $\mathcal{K}_m(A, b)$

We get

$$AP_m = P_{m+1}\overline{H}_m$$

where $\overline{H}_m = \begin{bmatrix} H_m \\ h_{m+1,m}e_m^T \end{bmatrix}$ and $P_m = [p_1, \ldots, p_m]$. 
Consider a SVD \( \overline{H}_m = U_m \Sigma^2_m V_m \), where \( U_m \in \mathbb{R}^{(m+1) \times (m+1)} \), \( V_m \in \mathbb{R}^{m \times m} \), \( \Sigma^2_m = [\tilde{\Sigma}_m^2, 0]^T \) and \( \tilde{\Sigma}_m^2 = \text{diag}(\sigma_{m,1}^2, \ldots, \sigma_{m,m}^2) \).

Approximate SVD (Novati-Russo 2013:)

Let \( \overline{U}_m = P_{m+1} U_m, \overline{V}_m = P_m V_m \), then

- \( A \overline{V}_m = \overline{U}_m \Sigma^2_m, A \overline{U}_m = \overline{V}_m \Sigma^2_m \)
- the first \( m \) singular values of \( A \) are well approx. by \( \sigma_{m,j}^2 \)
- If \( m = n \), in exact arithmetic the triplet \( (P_{m+1} \tilde{U}_m, \tilde{\Sigma}_m, P_m V_m) \) is a SVD of \( A \), where \( \tilde{U}_m \) is \( U_m \) without the last column.
The new basis

**Definition**

Recall:

- $AP_m = P_{m+1} \overline{H}_m$
- $\overline{H}_m = U_m \Sigma_m^2 V_m^T$
- $\Sigma_m^2 = [\tilde{\Sigma}_m^2, 0]^T$
- $\tilde{U}_m$ is $U_m$ without the last column.

**Definition:**

The sub-basis $\mathcal{U}_m$ is a set $\{u_1, \ldots, u_m\} \subset N_k(X)$ defined by

$$V_u = P_{m+1} \tilde{U}_m \tilde{\Sigma}_m, \quad C_u = P_m V_m \tilde{\Sigma}_m^{-1}.$$
The new basis

Properties

Properties:

The sub-basis $\mathcal{U}_m$ has the following properties for each $1 \leq m \leq n$:

- it is $\mathcal{N}_k(\Omega)$-orthonormal
- it is $\ell_{2,w}(X)$-orthogonal with $\|u_j\|_{\ell_{2,w}(X)} = \sigma_{m,j}^2, 1 \leq j \leq m$
- if $m = n$ it is the SVD basis $\mathcal{U}$ ($W = I$)

Using this basis we get $\forall f \in \mathcal{N}_k(\Omega)$

$$s'_f(x) = \sum_{j=1}^{m} \frac{(f, u_j)_{\ell_{2,w}(X)}}{\sigma_{m,j}^2} u_j(x) = \sum_{j=1}^{m} (f, u_j) u_j(x) \quad \forall x \in \Omega$$

(and $P_{k,x}^{(m)}(x)$, $\tilde{\Lambda}_x$ as before)
Numerical Results

Stopping rule

$$\left( \overline{H}_m \right)_{m+1,m} \approx \sigma^2_{m,j}$$

Figure: Gaussian kernel, $\varepsilon = 1$, square, $n = 200$ e.s. points, $f \in \mathcal{N}_K(\Omega)$
Numerical Results
Comparison with the SVD basis

Figure: Gaussian kernel, $\varepsilon = 1$, disk, trig.-gauss. points, $f \in \mathcal{N}_K(\Omega)$
Numerical Results

An example

Figure: IMQ kernel, $\varepsilon = 1$, cutted-disk, $n = 1600$ random points, $m = 260$, $f(x, y) = \exp(|x - y|) - 1$
Numerical Results
Lebesgue Constants and Power Function

Figure: IMQ kernel, $\varepsilon = 1$, cutted-disk, $n = 1600$ random points, $m = 260$, $f(x, y) = \exp(|x - y|) - 1$
Further work

Further investigation is needed:

- a better stopping rule
- understand the decay rate of $P^{(m)}_{k,X}$
- understand the growing rate of $\tilde{\Lambda}_X$
- understand how $X, \varepsilon$ influence $s'_f$
Further work

Further investigation is needed:

- a better stopping rule
- understand the decay rate of $P^{(m)}_{k,x}$
- understand the growing rate of $\tilde{\Lambda}_x$
- understand how $X, \varepsilon$ influence $s'_f$

Thank you for your attention
Fast fitting of radial basis functions: methods based on preconditioned GMRES iteration. 

S. De Marchi, G. Santin.
A new stable basis for radial basis function interpolation. 

G. Fasshauer, M. J. McCourt.
Stable evaluation of Gaussian radial basis functions interpolants. 

A. C. Faul, M. J. D. Powell.
Krylov subspace methods for radial basis function interpolation. 

P. Novati, M. R. Russo.
A GCV based Arnoldi-Tikhonov regularization method. 

M. Pazouki, R. Schaback.
Bases for kernel-based spaces. 

M. Vianello and G. Da Fies.
Algebraic cubature on planar lenses and bubbles. 