Bases for Conditionally Positive Definite Kernels

Maryam Pazouki[a,1], Robert Schaback[a]

[a] *Institut für Numerische und Angewandte Mathematik*
Universität Göttingen
Lotzestraße 16-18
D–37073 Göttingen
Germany

Abstract

This paper extends a previous one [?] to the important case of *conditionally* positive kernels such as thin–plate splines or polyharmonic kernels. The goal is to construct well–behaving bases for interpolation on a finite set $X \subset \mathbb{R}^d$ by translates $K(\cdot, x)$ for $x \in X$ of a fixed kernel $K : \Omega \times \Omega \to \mathbb{R}$ which is conditionally positive definite of order $m > 0$. Particularly interesting cases are bases of Lagrange or Newton type, and bases which are orthogonal or orthonormal either discretely (i.e. via their function values on X) or as elements of the underlying "native" space for the given kernel, which is a direct product of a Hilbert space with the space $\P_m^d$ of d–variate polynomials of order up to m. All of these cases are considered, and relations between them are established. It turns out that there are many more possibilities fo basis construction than in the unconditionally positive definite situation $m = 0$, and these possibilities are sorted out systematically. Some numerical examples are provided.

Keywords: kernels, radial basis functions, conditionally positive definite kernels, unisolvent set.

AMS classification: 41A05,41063, 41065, 65D05, 65D15

1. Introduction

If interpolation on a finite set $X \subset \mathbb{R}^d$ of *centers* is done by *translates* $K(\cdot, x)$ for $x \in X$ of a fixed *positive definite kernel* $K : \Omega \times \Omega \to \mathbb{R}$, it is well–known that the basis spanned by the translates often is badly conditioned, while the interpolation is a well–behaving map [?] in function space. This calls for other choices of bases for the space

$$S(X) := \text{ span } \{K(\cdot, x) \ : \ x \in X\} \tag{1}$$

Email addresses: `m.pazouki@math.uni-goettingen.de` (Maryam Pazouki), `schaback@math.uni-goettingen.de` (Robert Schaback)

[1]Sponsored by Deutsche Forschungsgemeinschaft via Graduiertenkolleg 1023

and the paper [**?**] provides various useful solutions. This covers kernels like

- the Gaussian $K(x,y) = \exp(-\|x-y\|_2^2)$,
- inverse multiquadrics $K(x,y) = (1+\|x-y\|_2^2)^{-n}$, $n > 0$,
- and Wendland's compactly supported kernels like $K(x,y) = (1-\|x-y\|_2)_+^4(1+4\|x-y\|_2)$.

But there are other kernels, e.g.

- multiquadrics $K(x,y) = (-1)^{\lceil\beta/2\rceil}(1+\|x-y\|_2^2)^{\beta/2}$, $\beta \in (0,\infty)\setminus 2\mathbb{Z}$,,
- powers $K(x,y) = (-1)^{\lceil\beta/2\rceil}\|x-y\|_2^\beta$, $\beta \in (0,\infty)\setminus 2\mathbb{Z}$, or
- thin–plate splines $K(x,y) = (-1)^{1+\beta/2}\|x-y\|_2^\beta \log\|x-y\|_2$, $\beta \in 2\mathbb{Z}$

that are not covered by [**?**] because the kernels are not positive definite, but only *conditionally* positive definite of some positive order $m = \lceil\beta/2\rceil$, $\lceil\beta/2\rceil$, and $m = 1+\beta/2$ in the above three cases, respectively.

To introduce this notion, we first need the space $¶_m^d$ of d–variate real–valued polynomials of order at most m which has the dimension $Q := \binom{m-1+d}{d}$, and we have to restrict the admissible point sets by

Definition 1. *A subset $X = \{x_1,\dots,x_N\} \subset \Omega \subset \mathbb{R}^d$ is called* $¶_m^d$–unisolvent*, if the only function in $¶_m^d$ that vanishes on X is zero.*

Definition 2. *A kernel $K : \Omega\times\Omega \to \mathbb{R}^d$ on a set $\Omega \subseteq \mathbb{R}^d$ is* conditionally positive (semi–) definite of order $m \geq 0$*, if for all point sets $X = \{x_1,\dots,x_N\} \subset \Omega$ that are $¶_m^d$–unisolvent, the quadratic forms*

$$Q(c) := \sum_{j,k=1}^N c_j c_k K(x_j,x_k),\ c \in \mathbb{R}^N \tag{2}$$

are positive (semi–) definite on the subspaces

$$M := \left\{ c \in \mathbb{R}^N \ : \ \sum_{j=1}^N c_j p(x_j) = 0 \text{ for all } p \in ¶_m^d \right\} \tag{3}$$

of coefficients satisfying discrete moment conditions *of order m on X.*

The space (1) must now be replaced by

$$S := ¶_m^d + \left\{ \sum_{x_j\in X} c_j K(x_j,\cdot) \ : \ X \text{ is } ¶_m^d - \text{ unisolvent},\ c \in M \right\}, \tag{4}$$

and this coincides with (1) for $m = 0$. For the rest of the paper, we always assume X to be $¶_m^d$–unisolvent and the kernel K to be conditionally positive definite of order m, suppressing X and m in the notation.

The goal of this paper is to provide useful bases for S and to exhibit relations between these bases. Due to their dependence on the given unisolvent set X, we call these bases *data–dependent*. The case $m = 0$ of [**?**] will serve as a guideline, but it will turn out that it does not generalize in a straightforward way, since there are many possibilities to proceed. For instance, one can preselect a polynomial basis and complete it by $N-Q$ other functions, or let all basis functions contain some polynomial part. Another option is to preselect a minimal unisolvent subset of Q points first and then use a "reduced" kernel. Finally, one can try to reduce the conditional positive definite case to the unconditionally positive definite case by a suitable change of the kernel, but even this change can be done in different ways. These different approaches sometimes lead to the same result.

The paper starts with a recollection of basic notation and results, and then looks at bases from a general point of view, i.e. not assigning a special rôle to polynomials, and not reordering points. We then specialize to bases induced by "projecting polynomials away" and go over to bases that are partially orthonormal in the semi–Hilbert spaces behind conditionally positive definite kernels. In particular, we shall construct bases related to Cholesky and SVD decompositions of certain matrices generated via the polynomial projectors of the previous section. A special class of "*points and polynomials first*" bases arises if basis functions and points are reordered to let the first $Q = \dim \P^d_m$ points form a $\P^d_m$–unisolvent set, and let the basis start with a basis of $\P^d_m$. We analyze these bases and finally consider an important special case, i.e. the *Newton* basis that can be calculated adaptively and effectively. It is a natural generalization of the Newton basis in [**? ?**].

The paper closes with numerical examples illustrating the results.

2. Notation and Basic Facts

We first fix a basis $p_1, \ldots, p_Q$ of $\P^d_m$ and $Q := \binom{m-1+d}{d} = \dim \P^d_m$. If $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ is a $\P^d_m$–unisolvent set, the $N \times Q$ matrix P with entries $p_i(x_j),\quad j = 1, \ldots, N,\ i = 1, \ldots, Q$ has full rank Q, and this implies $N \geq Q$. Enlarging unisolvent sets will not destroy unisolvency, and each unisolvent set contains a subset of Q points which is unisolvent itself. We define the symmetric *kernel matrix*

$$A := (K(x_i, x_j))_{1 \leq i,j \leq N} \tag{5}$$

and we know that it defines a positive definite quadratic form (2) on the subspace (3).

Lemma 1 ([**?**]). *The* $(N+Q) \times (N+Q)$ *matrix*

$$\mathcal{A} := \begin{pmatrix} A & P \\ P^T & 0_{Q \times Q} \end{pmatrix} \tag{6}$$

is nonsingular, and interpolation of data on X *is uniquely possible.* □

If data on X are supplied as a vector $f \in \mathbb{R}^N$, solving the system

$$\mathcal{A}\begin{pmatrix} c \\ b \end{pmatrix} = \begin{pmatrix} A & P \\ P^T & 0_{Q\times Q} \end{pmatrix}\begin{pmatrix} c \\ b \end{pmatrix} = \begin{pmatrix} f \\ 0_{Q\times 1} \end{pmatrix}$$

yields vectors $c \in M \subseteq \mathbb{R}^N$ and $b \in \mathbb{R}^Q$ with moment conditions $P^T c = 0$. With such vectors, we define

$$s_c := \sum_{x_j \in X} c_j K(x_j, \cdot), \quad p_b := \sum_{i=1}^{Q} b_i p_i \in ¶_m^d \tag{7}$$

to get the interpolant $s := s_c + p_b \in S$ to the given data. Now Lemma 1 implies

Lemma 2. *In (4), the sum is direct, and the space S has dimension $N = |X|$.* □

For later use, we need that the bilinear forms

$$\begin{array}{rcl} (c, c') & := & c^T A c' \\ & =: & (s_c . s_{c'}) \end{array} \tag{8}$$

on M and the functions of the form s_c are inner products, due to $(c, c) = Q(c)$. We shall use the same notation for these inner products, since there is no possible confusion. Constructions of orthogonal or orthonormal bases will use these inner products later.

3. General Bases

We now collect necessary properties of general data–dependent bases $w = (w_1, \dots, w_N)$ for a given $¶_m^d$–unisolvent set X of $N \geq Q$ points. There will be no special rôle of polynomials, and no reordering of the points of X.

Definition 3. *The* value matrix *of a basis $w = (w_1, \dots, w_N)$ is the $N \times N$ matrix*

$$V_w = (w_i(x_j))_{1 \leq i,j \leq N} \in \mathbb{R}^{N \times N}$$

where j is the row and i is the column index.

Clearly, this matrix is necessarily nonsingular because the basis must allow unique interpolation on X. The value matrix V_w defines a basis w uniquely. To construct the basis functions from the values, consider the system

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix}\begin{pmatrix} C_w \\ B_w \end{pmatrix} = \begin{pmatrix} V_w \\ 0 \end{pmatrix} \tag{9}$$

that expresses the $N \times N$ matrix C_w and the $Q \times N$ matrix B_w uniquely in terms of V_w. The *moment conditions* are employed via $P^T C_w = 0$, while the *coefficient* matrices C_w and B_w satisfy

$$AC_w + PB_w = V_w. \tag{10}$$

The associated basis w is then determined by the columns of C_w and B_w via

$$w_k = \sum_{j=1}^{N} c_{jk} K(x_k, \cdot) + \sum_{i=1}^{Q} b_{ik} p_i, \ 1 \le k \le N \tag{11}$$

where we omitted w in the notation of the matrix elements.

We derive a few simple facts from (9):

Theorem 1. *The $N \times N$ matrix C_w has rank $N - Q$, and the $(N+Q) \times N$ matrix*

$$\begin{pmatrix} C_w \\ B_w \end{pmatrix} = \mathcal{A}^{-1} \begin{pmatrix} V_w \\ 0 \end{pmatrix}$$

has rank N.

A basis w is uniquely defined by either a nonsingular $N \times N$ value matrix V_w or by a $N \times N$ matrix C_w and a $Q \times N$ matrix B_w such that $P^T C_w = 0$ holds and $\binom{C_w}{B_w}$ has rank N. $\square$

Theorem 2. *Given a nonsingular value matrix V_w, the corresponding basis w is uniquely defined and independent of basis changes in $\P_m^d$. In particular, the coefficient matrix C_w is unique, while the B_w matrix and the polynomials change.*

Proof: Each basis change in $\P_m^d$ is given by the transition $Q := P \cdot C$ with a nonsingular $Q \times Q$ matrix C. Then, by easy calculations, the matrix B_w goes over to $C^{-1} B_w$ while C_w is unchanged. In (11), the basis functions w_j are not changed, because the matrix C changes both the b_{ik} and the p_i in a way that cancels out. $\square$

At this point we can already define the unique Lagrange basis L by requiring that the value matrix V_L is the identity matrix. Its construction matrices are then uniquely defined by

$$\begin{pmatrix} C_L \\ B_L \end{pmatrix} = \begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} I_{N \times N} \\ 0 \end{pmatrix},$$

and (11) can be used to evaluate the basis everywhere.

Using (10) and $P^T C_w = 0$ we see that the matrix

$$C_w^T A C_w = C_w^T V_w$$

is symmetric. If we use the inner product (8), we can form the Gramian matrix

$$G_w := C_w^T A C_w = C_w^T V_w \tag{12}$$

of the inner products of the s_c parts of the basis functions. Now Theorem 1 yields

Corollary 1. *The Gramian G_w is symmetric and positive semidefinite with rank $N - Q$. In particular, it is impossible to have a full orthonormal basis of N functions of S if $Q > 0$.* $\square$

We shall later consider bases with $N-Q$ orthonormal functions.

The matrix of discrete ℓ_2 inner products on X is

$$H = V_w^T V_w,$$

since the rows of V_w correspond to points, while the columns correspond to functions. Clearly, this matrix is positive definite, and there are plenty of ℓ_2–orthonormal bases of S, in particular the Lagrange basis. For completeness, we note the simple fact

Theorem 3. *All ℓ_2–orthonormal bases w on X arise from orthogonal value matrices V_w.* □

4. Polynomial Projectors

We now want to get rid of the polynomial part in (10). To this end, we need a $N \times N$ matrix Π with $\Pi P = 0$ and full rank $N-Q$. There are many ways to get such a matrix.

Since P has rank Q, the $Q \times Q$ matrix $P^T P$ is invertible, and we can proceed from (9) via

$$\begin{array}{rcl} P^T A C_w + P^T P B_w & = & P^T V_w, \\ B_w & = & (P^T P)^{-1} P^T (V_w - A C_w) \end{array}$$

to get

$$\Pi A C_w = \Pi V_w \tag{13}$$

with the symmetric projector

$$\Pi := I - P(P^T P)^{-1} P^T \tag{14}$$

that arises naturally here. Note that equation (13) generalizes the identity $AC_w = V_w$ obtained in the unconditionally positive definite case.

Another way to get an $N \times N$ matrix Π with $\Pi P = 0$ and full rank $N-Q$ is to reorder the points of X to let the set $X_Q := \{x_1, \dots, x_Q\}$ be unisolvent and to let the polynomial basis be a Lagrange basis on X_Q. Then we have $P = \binom{I_{Q\times Q}}{P_2}$ and can define

$$\Pi = \left(\begin{array}{cc} 0_{Q\times Q} & 0_{Q \times N-Q} \\ -P_2 & I_{N-Q\times N-Q} \end{array} \right). \tag{15}$$

If seen as actions on data vectors, Π of (14) replaces the data by the error of the ℓ_2–optimal polynomial recovery on X, while (15) replaces the data by the error of interpolation in X_Q. In both cases, the data of polynomials are mapped to zero, i.e. $\Pi P = 0$ holds. However, the matrix Π of (14) has the advantage to be symmetric, idempotent, and independent of the choice of basis in $¶_m^d$.

In all cases, the maximal rank of Π and the property $\Pi P = 0$ imply

Theorem 4. 1. *For any vector $z \in \mathbb{R}^N$, the vector $\Pi^T z$ satisfies the moment conditions.*

2. *For all vectors c satisfying the moment conditions there is a representation as $c = \Pi^T z$ for some $z \in \mathbb{R}^N$.* □

Theorem 5. *The matrix $\Pi A \Pi^T$ is symmetric and positive semidefinite with rank $N - Q$.*

Proof: The symmetric quadratic form $z \mapsto z^T \Pi A \Pi^T z$ is positive semidefinite, because the vectors $\Pi^T z$ satisfy the moment conditions. Its kernel is the kernel of Π and thus equal to the range of P, i.e. of dimension Q. □

If we multiply (10) from the left by Π, we get (13) also for the general situation. Furthermore, Theorem 4 implies that we can factorize

$$C_w = \Pi^T F_w \tag{16}$$

with a nonunique but nonsingular $N \times N$ matrix F_w. In particular, a very natural candidate for a basis w^* is definable by

$$C_{w^*} := \Pi^T$$

due to the properties of Π. In the two special cases for Π provided above, the matrix $\binom{\Pi^T}{P^T}$ has rank N, and thus Theorem 1 is applicable, yielding a basis in both cases.

5. Partially Orthonormal Bases

Following Corollary 1, we now can ask for all bases which have Gramians G_o of (12) which are $N \times N$ diagonal matrices with Q zeros and $N - Q$ ones on the diagonal. With an $N \times N$ matrix Π of (14) with rank $N - Q$ and $\Pi P = 0$, we can use (16) and write down the necessary equation

$$C_w^T A C_w = F_w^T \Pi A \Pi^T F_w = G_o = G_o^2.$$

Since F_w is nonsingular (but also nonunique), we get the decomposition

$$\Pi A \Pi^T = (F_w^T)^{-1} G_o^2 F_w^{-1} = E_w^T E_w$$

with the matrix

$$E_w := G_o F_w^{-1}.$$

Theorem 6. *All partially orthonormal bases of S arise from factorizations*

$$\Pi\, A \Pi^T = E^T\, E \tag{17}$$

with $N \times N$ matrices E of rank $N - Q$ that have the property that $E = G_o F^{-1}$ with a nonsingular $N \times N$ matrix F, and where G_o is a $N \times N$ diagonal matrix with Q zeros and $N - Q$ ones on the diagonal.

Proof: We have just seen that the conditions are necessary. To prove sufficiency, we start with E and F and define a basis w via $C_w := \Pi^T F$. Then

$$\begin{aligned} C_w^T A C_w &= F^T \Pi A \Pi^T F \\ &= F^T E^T E F \\ &= F^T {F^T}^{-1} G_o^T G_o F^{-1} F \\ &= G_o. \end{aligned}$$

□

Corollary 2. *These partially orthonormal bases are not unique. Even if C_w is fixed along the above lines, the matrices B_w for application of Theorem 1 are still free.* □

We now consider partially orthonormal bases that arise from pivoted Cholesky decompositions of the symmetric positive semidefinite $N \times N$ matrix $\Pi\, A\, \Pi^T$ of rank $N - Q$. We assume that after reordering of points and N steps of the Cholesky algorithm, we get

$$\Pi\, A\, \Pi^T = L\, G_o\, L^T$$

with a nonsingular lower triangular matrix L and a diagonal matrix G_o with the first $N - Q$ diagonal elements being one and the final Q diagonal elements being zero.

Then, in the above context, $F = (L^T)^{-1}$ and $C = \Pi^T (L^T)^{-1}$ and $G_o = C^T V = L^{-1} \Pi V$ leading to

$$\Pi V = L G_0,$$

generalizing what we had [**?** , Section 6] in the Newton case for unconditionally positive definite kernels.

If we decompose in SVD style

$$\Pi\, A\, \Pi^T = U\, G_o\, U^T$$

with an $N \times N$ orthogonal matrix U and a nonngegative matrix G_o of singular values with exactly $N - Q$ positive ones, then

$$C = \Pi^T U,\ \Pi V = U G_0,$$

again generalizing what we have in the standard case [**?** , Section 6].

6. Points and Polynomials First

Definition 4. *Let the points of a $\P_m^d$–unisolvent set $X = \{x_1, \ldots, x_N\}$ be ordered in such a way that the first Q points form a $\P_m^d$–unisolvent subset $X_Q = \{x_1, \ldots, x_Q\} \subseteq X$, and assume that $p_1, \ldots, p_Q$ are a basis of $\P_m^d$. Then any data–dependent basis w consisting of $w_1 = p_1, \ldots, w_Q = p_Q$ and $N - Q$ other basis functions $w_{Q+1}, \ldots, w_N$ is called a* points and polynomials first *(PPF) basis.*

We now proceed to characterize all PPF bases. We define $P^T = (P_1^T, P_2^T)$ with matrices $P_1 \in \mathbb{R}^{Q\times Q}$, $P_2 \in \mathbb{R}^{(N-Q)\times Q}$ using our ordering, and let the value matrix of the complete basis take the form

$$V_w := \begin{pmatrix} P_1 & V_P \\ P_2 & V_R \end{pmatrix}$$

with matrices $V_P \in \mathbb{R}^{Q\times(N-Q)}$, $V_R \in \mathbb{R}^{(N-Q)\times(N-Q)}$. By construction, the matrix P_1 is nonsingular. Splitting the matrices A and $\mathcal{A}$ similarly, we get

$$\mathcal{A} = \begin{pmatrix} A_{11} & A_{12} & P_1 \\ A_{12}^T & A_{22} & P_2 \\ P_1^T & P_2^T & 0 \end{pmatrix}.$$

Note that since we chose a fixed polynomial basis first, the matrices P_1 and P_2 are determined by this basis choice. When taking a polynomial Lagrange basis on X_Q to start with, we get $P_1 = I$.

Theorem 7. *The identity (9) necessarily is of the form*

$$\begin{pmatrix} A_{11} & A_{12} & P_1 \\ A_{12}^T & A_{22} & P_2 \\ P_1^T & P_2^T & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & C_P \\ 0 & C_R \\ I & B \end{pmatrix} = \begin{pmatrix} P_1 & V_P \\ P_2 & V_R \\ 0 & 0 \end{pmatrix} \tag{18}$$

with a $Q\times(N-Q)$ matrix B and a nonsingular $(N-Q)\times(N-Q)$ matrix C_R. If C_R is fixed, all choices of either B or V_P are possible, and each such choice defines a PPF basis.

Proof: For a PPF basis, the system (9) has the general form

$$\begin{pmatrix} A_{11} & A_{12} & P_1 \\ A_{12}^T & A_{22} & P_2 \\ P_1^T & P_2^T & 0 \end{pmatrix} \cdot \begin{pmatrix} R & C_P \\ S & C_R \\ T & B \end{pmatrix} = \begin{pmatrix} P_1 & V_P \\ P_2 & V_R \\ 0 & 0 \end{pmatrix}$$

and consequently

$$\begin{array}{ccccccc} A_{11}R & + & A_{12}S & + & P_1T & = & P_1, \\ A_{12}^TR & + & A_{22}S & + & P_2T & = & P_2, \\ P_1^TR & + & P_2^TS & & & = & 0. \end{array}$$

A simple elimination then yields $\tilde{A}S = 0$ with the matrix

$$\tilde{A} := A_{22} - A_{12}^T(P_1^T)^{-1}P_2^T - P_2P_1^{-1}A_{12} + P_2P_1^{-1}A_{11}(P_1^T)^{-1}P_2^T. \tag{19}$$

which is nonsingular due to Theorem 8 below. Thus $S = 0$, and nonsingularity of P_1 then implies $R = 0$ and $T = I$, proving (18) to be necessary.

By elimination in (18), we get necessary equations

$$\begin{array}{rcl} C_P & = & -(P_1^T)^{-1}P_2^TC_R, \\ B & = & P_1^{-1}\left(V_P+(A_{11}(P_1^T)^{-1}P_2^T-A_{12})C_R\right), \\ V_R & = & \tilde{A}C_R+P_2P_1^{-1}V_P \end{array} \tag{20}$$

with the symmetric positive definite $(N-Q)\times(N-Q)$ matrix $\tilde{A}$ from (19). The nonsingularity of the value matrix implies that $V_R - P_2P_1^{-1}V_P$ must be nonsingular, and since this is $\tilde{A}C_R$, we get that C_R necessarily is nonsingular. The above arguments can be pursued backwards to show that each nonsingular C_R and arbitrary choices of either B or V_P yield a PPF basis. □

Theorem 8. *The matrix $\tilde{A}$ is positive definite, and it is independent of the polynomial basis chosen.*

Proof: For an arbitrary vector $z\in\mathbb{R}^{N-Q}$ we get moment conditions

$$(P_1^T, P_2^T)\left(\begin{array}{c} -(P_1^T)^{-1}P_2^Tz \\ z \end{array}\right) = 0$$

and thus

$$\left(\begin{array}{c} -(P_1^T)^{-1}P_2^Tz \\ z \end{array}\right)^T A \left(\begin{array}{c} -(P_1^T)^{-1}P_2^Tz \\ z \end{array}\right) = z^T\tilde{A}z \geq 0$$

by conditional positive definiteness of the kernel. If the quadratic form is zero, then $z=0$. A basis change in $¶_m^d$ will not change $P_2P_1^{-1}$, because P_1 and P_2 go over to P_1C and P_2C with a nonsingular $Q\times Q$ matrix C. □

For a split

$$V_w = \left(\begin{array}{cc} P_1 & V_P \\ P_2 & V_R \end{array}\right)$$

of the corresponding value matrix and

$$\tilde{V}_w := V_R - P_2P_1^{-1}V_P$$

we have

$$\tilde{A}C_R = \tilde{V}_w. \tag{21}$$

There is a particularly simple special situation:

Definition 5. *A PPF basis is called* canonical*, if it satisfies $V_P = 0$.*

For canonical PPF bases, we have the simple identity

$$\tilde{A}C_R = V_R \tag{22}$$

between nonsingular $(N-Q)\times(N-Q)$ matrices. Either V_R or C_R can be prescribed for a canonical PPF basis to define it uniquely.

The identities (21) and (22) generalize what we had in the unconditionally positive definite case. It will turn out later that canonical PPF bases arise from using a "reduced" kernel.

In principle, one may put $B := 0$, and $C_R := I$. This special case of a PPF basis will be treated in the next section after a short detour.

Theorem 9. *All PPF bases w with orthonormal functions $w_{Q+1},\ldots,w_N$ arise from a factorization*

$$\tilde{A} = (C_R^T)^{-1}(C_R)^{-1} = E^T E,$$

where $E := C_R^{-1}$. Then the value matrix is

$$\tilde{V}_w = (C_R^T)^{-1}$$

and for a canonical PPF basis we have $V_R = (C_R^T)^{-1}$. □

Proof: By a simple recalculation of (12) using

$$C_w = \begin{pmatrix} 0 & C_P \\ 0 & C_R \end{pmatrix}.$$

we get the Gramian of a PPF basis w to be

$$G_w = C_R^T \tilde{A} C_R.$$

If this is the identity, the first assertion follows. The rest is a consequence of (21) and (22). □

This is similar to the results in [**?** , Section 6] for the unconditionally positive definite case. There, different decompositions like Cholesky and SVD were used to construct orthonormal bases. Here, partially orthonormal PPF bases arise from SVD or Cholesky decompositions of the matrix $\tilde{A}$.

We now want to determine the discretely orthonormal PPF bases. If we split the value matrix V_w by reordering the Q unisolvent points, then

$$V_w = \begin{pmatrix} P_1 & V_P \\ P_2 & V_R \end{pmatrix}.$$

To construct all discretely orthonormal PPF bases, we start from $H = V_w^T V_w = I$ and get

$$\begin{aligned} P_1^T P_1 + P_2^T P_2 &= I \\ V_P^T V_P + V_R^T V_R &= I \\ P_1^T V_P + P_2^T V_R &= 0, \end{aligned} \tag{23}$$

where P_1 is nonsingular. Then from (21) we get

$$\tilde{A}C_R = (I + P_2(P_1)^{-1}(P_1^T)^{-1}P_2^T)V_R,$$

which is another generalization of what we know in the unconditionally positive definite case.

Theorem 10. *Canonial PPF bases are not discretely orthonormal.*

Proof Canonical PPF bases have $V_P = 0$. Then V_R is nonsingular, and the system (23) shows that P_2 must be zero. Thus all polynomials must be zero on all points, contradicting unisolvency. □

7. Standard Basis

In a straightforward attempt to construct a basis, we would like to use the $K(x, x_j)$ directly, but this is not allowed since we have to obey the moment conditions. Assume $p_1, \ldots, p_Q$ be a Lagrange basis of polynomial interpolation on $x_1, \ldots, x_Q$. Then we can reproduce

$$p(y) = \sum_{m=1}^{Q} p(x_m)p_m(y) \text{ for all } p \in \P,\ y \in \Omega,$$

and for the set $\{y\} \cup \{x_1, \ldots, x_Q\}$ we have a coefficient vector

$$1, -p_1(y), \ldots, -p_Q(y)$$

that satisfies the moment conditions on that set, i.e.

$$1 \cdot p(y) - \sum_{m=1}^{Q} p(x_m)p_m(y) = 0 \text{ for all } p \in \P.$$

This allows to define the functions

$$s_y(x) := K(x, y) - \sum_{m=1}^{Q} p_m(y)K(x, x_m),\ x \in \Omega$$

for all $y \in \Omega \setminus \{x_1, \ldots, x_Q\}$, which are in S.

Then a standard basis generalizing the translates $K(x, x_j)$ consists of $p_1, \ldots, p_Q$ and of simple linear combinations of translates of K, i.e.

$$s_j(x) := K(x, x_j) - \sum_{m=1}^{Q} p_m(x_j)K(x, x_m),\ Q + 1 \le j \le N.$$

If we add $s_j := p_j$ for $1 \le j \le Q$, we have N functions and need to prove linear independence. Define $P^T = (I_{Q\times Q}, P_2^T)$ due to the Lagrange property on the first Q points. We split A and $\mathcal{A}$ to get

$$\mathcal{A} = \left(\begin{array}{ccc} A_{11} & A_{12} & I \\ A_{12}^T & A_{22} & P_2 \\ I & P_2^T & 0 \end{array} \right).$$

The value matrix has P as its left $N \times Q$ submatrix. The right $N \times (N-Q)$ submatrix is

$$\left(\begin{array}{c} A_{12} \\ A_{22} \end{array} \right) - \left(\begin{array}{c} A_{11} \\ A_{12}^T \end{array} \right) \cdot P_2^T$$

due to the definition of the basis. The identity (9) then is

$$\left(\begin{array}{ccc} A_{11} & A_{12} & I \\ A_{12}^T & A_{22} & P_2 \\ I & P_2^T & 0 \end{array} \right) \cdot \left(\begin{array}{cc} 0 & -P_2^T \\ 0 & I \\ I & 0 \end{array} \right) = \left(\begin{array}{cc} I & A_{12} - A_{11}\, P_2^T \\ P_2 & A_{22} - A_{12}^T\, P_2^T \\ 0 & 0 \end{array} \right)$$

and we see that we have got a PPF basis with $B = 0$ and $C_R = I$.

The Gramian follows from (12) and takes the form

$$\left(\begin{array}{cc} 0 & 0 \\ 0 & \tilde{A} \end{array} \right)$$

with the symmetric matrix

$$\tilde{A} = P_2 A_{11} P_2^T - P_2 A_{12} - A_{12}^T P_2^T + A_{22}$$

that we know from (19) in a more general form. This generalizes the fact that the kernel matrix A itself is the Gramian of the standard basis in the unconditionally positive definite case.

8. Back to Lagrange Bases

Starting from a general $\P_m^d$–unisolvent set X, the standard Lagrange basis $u_1(x), \ldots, u_N(x)$ has the value matrix $V_u = I$ and is defined via

$$\left(\begin{array}{cc} A & P \\ P^T & 0 \end{array} \right) \left(\begin{array}{c} u(x) \\ v(x) \end{array} \right) = \left(\begin{array}{c} K_X(x) \\ p(x) \end{array} \right)$$

with

$$\begin{array}{rcl} K_X(x)^T & := & (K(x_1, x), \ldots, K(x_Q, x)) \in \mathbb{R}^N, \\ p(x)^T & := & (p_1(x), \ldots, p_Q(x)) \in \mathbb{R}^Q, \\ u(x)^T & := & (u_1(x), \ldots, u_N(x)) \in \mathbb{R}^N, \\ v(x)^T & := & (v_1(x), \ldots, v_Q(x)) \in \mathbb{R}^Q. \end{array}$$

The construction matrices of the Lagrange basis follow from Theorem 1 by setting $V_w = I$.

It is tempting to ask for PPF Lagrange bases. But:

Theorem 11. *For $N > Q$ there is no PPF Lagrange basis.*

Proof: If there were a PPF Lagrange basis, we must have $P_2 = 0$ in (18). But then all polynomials including 1 must vanish on $x_{Q+1}, \ldots, x_N$. □

But there clearly is a canonical PPF basis that is partially Lagrange in the sense that $V_R = I$ and $V_P = 0$. It has $C_R = \tilde{A}^{-1}$ by (22), and the equations (20) yield B and C_P for this case. Since not necessarily $P_2 = 0$, the basis is not Lagrange in the true sense.

9. Newton Basis, Iterative Construction

We start with $X_Q := \{x_1, \ldots, x_Q\}$ to be unisolvent, and there we take the standard Lagrange polynomial basis $p_1, \ldots, p_Q$ on X_Q. We need all values $p_i(x_j)$, $1 \le j \le N$, $1 \le i \le Q$ for further calculation, and we shall extend this basis to a full basis $v_1, \ldots, v_N$, setting $v_j = p_j$, $1 \le j \le Q$. We already have $v_i(x_j) = 0$ for $1 \le i < j \le Q$. We shall need the definition of the *reduced kernel*

$$\begin{aligned} K_Q(x,y) \quad := \quad & K(x,y) - \sum_{j=1}^{Q} p_j(x) K(y, x_j) - \sum_{k=1}^{Q} p_k(y) K(x, x_k) \\ & - \sum_{j=1}^{Q} \sum_{k=1}^{Q} p_j(x) p_k(y) K(x_j, x_k) \end{aligned}$$

which is symmetric and unconditionally positive definite on $\Omega \setminus X_Q$ and vanishes if one of the arguments is in X_Q. The second fact follows by direct evaluation, and the first follows from Theorem 8 and observing that the matrix $\tilde{A}$ of (19) is the kernel matrix for K_Q on $X \setminus X_Q$.

We store the actual values of the *power function*

$$P_Q^2(x) := K_Q(x,x)$$

on the points of X. They are zero on X_Q. For what follows, we start with $m := Q$.

Now, for induction, we assume that we already have a construction of a basis $v_1, \ldots, v_m$ for some $Q \le m < N$ with the orthonormality properties

$$(v_i, v_j) = \delta_{ij},\ Q+1 \le i, j \le m.$$

Formally, we also assume that we have the values of the power function $P_m(x)$ on X. Note that the associated *power kernel* satisfies the recursion

$$K_m(x,y) = K_{m-1}(x,y) - \frac{K_{m-1}(x, x_m) K_{m-1}(y, x_m)}{K_{m-1}(x_m, x_m)}$$

for $m > Q$ (see [**? ?**]) and has the properties

$$K_m(x_j, y) = K_m(x, x_j) = 0,\ 1 \le j \le m,\ P_m^2(x) = K_m(x, x) \text{ for all } x, y \in \Omega.$$

Then we define

$$x_{m+1} := \text{ arg max } \{P_m^2(x)\ :\ x \in X\}$$

and stop if $P_m(x_{m+1})$ is zero or very small, because then we are done. Now we formally define

$$v_{m+1}(x) := \frac{K_m(x, x_{m+1})}{\sqrt{K_m(x_{m+1}, x_{m+1})}} \text{ for all } x \in \Omega$$

to get that

$$v_{m+1}(x_j) = 0,\ 1 \le j \le m,$$

as required. But we have to show how v_{m+1} and P_{m+1} can be calculated efficiently on X.

We see immediately that

$$P_{m+1}^2(x) = P_m^2(x) - v_{m+1}^2(x) = P_Q^2(x) - \sum_{j=Q+1}^{m+1} v_j^2(x)$$

holds by construction and induction. Thus we only need v_{m+1} on X. We consider the recursion

$$\begin{array}{rcl} K_j(x, y) & = & K_{j-1}(x, y) - \dfrac{K_{j-1}(x, x_j) K_{j-1}(y, x_j)}{K_{j-1}(x_j, x_j)} \\ & = & K_{j-1}(x, y) - v_j(x) v_j(y) \end{array}$$

that boils down to

$$K_m(x, x_{m+1}) = K_Q(x, x_{m+1}) - \sum_{j=Q+1}^{m} v_j(x_{m+1}) v_j(x)$$

which is computable from the values $K_Q(x, x_{m+1})$ for all $x \in X$. But these are obtainable from

$$\begin{array}{rcl} K_Q(x, x_{m+1}) & = & K(x, x_{m+1}) - \displaystyle\sum_{j=1}^{Q} p_j(x) K(x_{m+1}, x_j) - \sum_{k=1}^{Q} p_k(x_{m+1}) K(x, x_k) \\ & & - \displaystyle\sum_{j=1}^{Q} \sum_{k=1}^{Q} p_j(x) p_k(x_{m+1}) K(x_j, x_k) \end{array}$$

at reasonable cost.

It remains to show

$$(v_{m+1}, v_j) = 0,\ Q + 1 \le j \le m,$$

and this follows from

$$(K_Q(x, x_{m+1}), v_j) = v_j(x_{m+1}),\ Q + 1 \le j \le m$$

because these v_j vanish on $x_1, \ldots, x_Q$. Furthermore,

$$\begin{aligned} P_m^2(x_{m+1})(v_{m+1}, v_{m+1}) &= (K_m(x, x_{m+1}), K_m(x, x_{m+1})) \\ &= K_m(x_{m+1}, x_{m+1}) \\ &= P_m^2(x_{m+1}) \end{aligned}$$

by what is known for power kernels from [**?**]. This proves orthonormality.

The value matrix of this canonical PPF basis is

$$\begin{pmatrix} I & 0 \\ P_2 & L \end{pmatrix}$$

with a lower triangular value matrix $L = V_R$, and we have orthonormality of $v_{Q+1}, \ldots, v_N$. Now Theorem 9 implies that we have $C_R = (L^{-1})^T$ and a Cholesky decomposition $\tilde{A} = LL^T$. This proves

Theorem 12. *The above construction generates the canonical PPF basis with $P_1 = I$, which coincides with the extension of the Lagrange basis of $¶_m^d$ by the Newton basis for $X \setminus X_Q$ and the reduced kernel.* □

10. Numerical Examples

Like in [**?**], we focus on the domain Ω defined as the unit circle with the third quadrant taken away. There, we take a fine set of points and select a subset of 15 points which is unisolvent for $¶_4^2$. This is done by a pivoted QR decomposition of the value matrix of the 15 basis functions of $¶_4^2$ on the fine grid. The left part of Figure 1 shows the selected points and the grid. For comparison, the right–hand part shows the 15 points selected by the greedy Newton strategy of section 9 when applied to the full grid.

The following plots always show the basis function # 15 for most of the bases constructed in this paper. Figure 2 starts with the standard basis of 7 and the basis using the projector Π of (15). Then Figure 3 provides the Cholesky and SVD bases of section 5. Finally, Figure 4 shows the adaptive Newton basis function v_{15} of section 9 with the associated Power Function P_{15}.

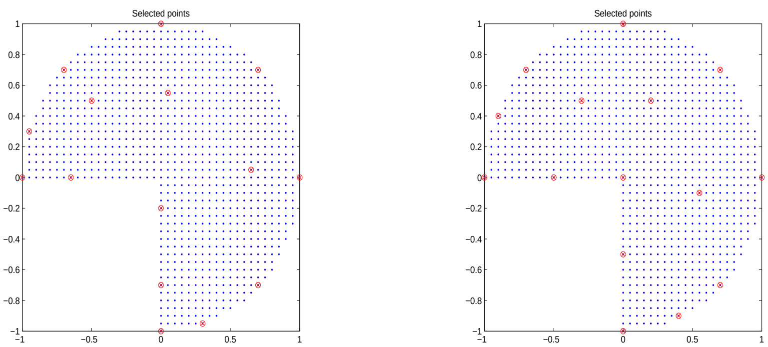

Figure 1: Selected 15 points, left: by unisolvency, right: by greedy Newton basis calculation

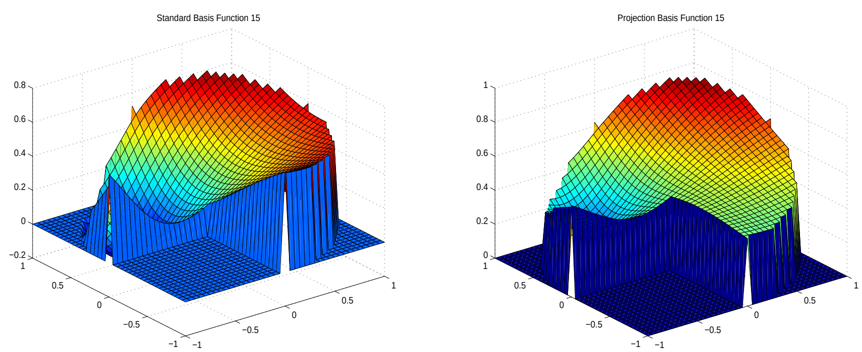

Figure 2: Basis function 15, left: standard basis, right: projection basis (15)

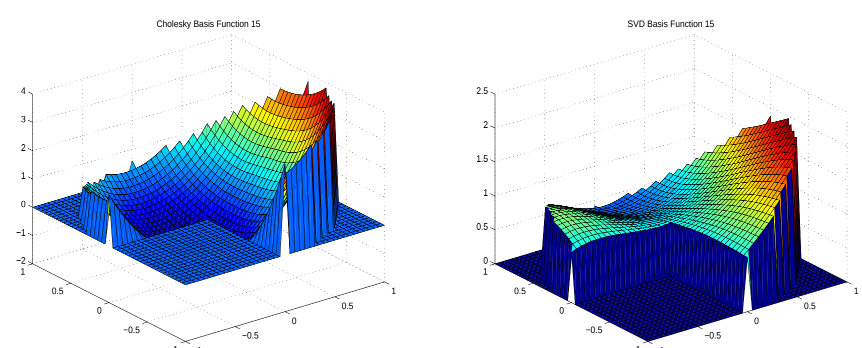

Figure 3: Basis function 15, left: Cholesky basis, right: SVD basis

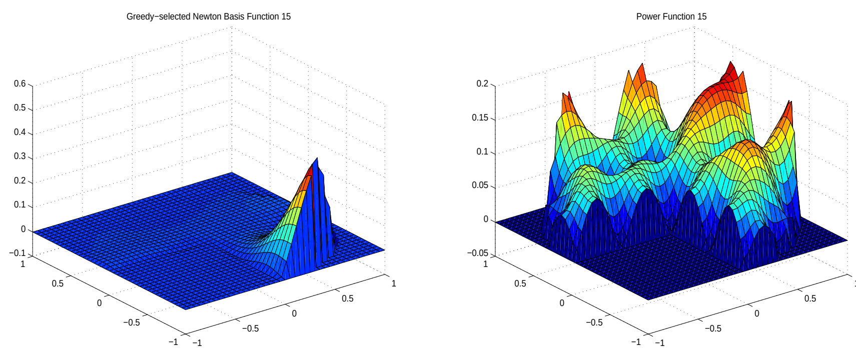

Figure 4: Newton basis function v_{15} and corresponding Power Function P_{15}