Bachelor Thesis

Reproducing Kernel Spaces and Kernels

carried out at the

Institute for
Analysis and Scientific Computing
University of Technology Vienna

under the supervision of

Dr. Harald Woracek

by

Riel Bllakcori
matriculation number: 11711304

Vienna, May 8, 2022

Statutory Declaration in Lieu of Oath

I hereby declare in lieu of an oath that I have completed the present Bachelor's thesis independently and without illegitimate assistance from third parties. I have used no other than the specified sources and aids.

Vienna, Date

signature of the author

Contents

1. Introduction

In functional analysis a reproducing kernel Hilbert space (RKHS) is a Hilbert space such that all point evaluation functionals are continuous. Reproducing kernel Hilbert spaces are always associated with kernel functions that reproduce the evaluation of some 'x' on a function in a sense that this can be done by taking the scalar product of the function with the kernel. These type of spaces were first introduced in the first half of the 20th century and appeared to be quite useful in a wide range of applications, including complex analysis, quantum mechanics, statistics and machine learning. Within the theory of indefinite inner product spaces, reproducing kernel Pontryagin spaces play a similarly important role. In this Bachelor's thesis we will consider the kind of spaces described above. Our goal will be to give the reader a good overview about the topic of different types of reproducing kernel spaces and kernel functions. Ultimately, we want to construct such spaces from functions that appear to have the same properties as these kernel functions.

The thesis starts with some relevant basics from linear algebra and functional analysis. Subsequently, we define reproducing kernel Hilbert spaces and make a first construction of a RKHS out of a hermitian positive definite kernel in Theorem 3.7. The proof is based upon [Kal19, Chapter 3.2]. Then we introduce conditionally positive definite kernels, a generalization of positive definite kernels and construct a RKHS with regard to them. Our approach is based on [Wen04, Chapter 10.3]. We continue by defining and characterizing indefinite inner product spaces, with its most prominent representative the Pontryagin space. In this section we extend our theory to those kind of spaces. The results and proofs are taken from many authors, mostly [BzCS12] and [Wor10]. Lastly, in the Appendix we discuss further indefinite inner product spaces and an application in machine learning for the interested reader.

In this Bachelor's thesis previous knowledge in functional analysis, first and foremost inner product spaces and Hilbert spaces, is recommended.

2. Basics

Let us recall some standard vocabulary from linear algebra and functional analysis. The following is based on [Wor10, Chapter 1] and [BKW20, Chapter 3].

2.1. Definition. Let $\mathcal{L}$ be a vector space. An inner product on $\mathcal{L}$ is a map

$$[.,.] : \mathcal{L} \times \mathcal{L} \to \mathbb{C}$$

such that

- $[x + y, z] = [x, z] + [y, z], \quad x, y, z \in \mathcal{L}.$
- $[\alpha x, y] = \alpha[x, y], \quad x, y \in \mathcal{L}, \alpha \in \mathbb{C}.$
- $[x, y] = \overline{[y, x]}, \quad x, y \in \mathcal{L}.$

If $[.,.]$, is an inner product on $\mathcal{L}$, we will speak of $(\mathcal{L}, [.,.])$ as an inner product space.

2.2. Definition. Let $(\mathcal{L}, [.,.])$ be an inner product space. A linear subspace $\mathcal{M}$ of $\mathcal{L}$ is called

$$\begin{array}{rlrl} \text{positive definite} & :\Leftrightarrow [x,x] > 0 & \text{positive semi-definite} & :\Leftrightarrow [x,x] \geq 0 \\ \text{negative definite} & :\Leftrightarrow [x,x] < 0 & \text{negative semi-definite} & :\Leftrightarrow [x,x] \leq 0 \end{array}$$

for all $x \in \mathcal{M}\backslash\{0\}$.

2.3. Lemma. *(Cauchy-Schwarz) Let $(\mathcal{L}, [.,.])$ be a semi-definite inner product space. Then*

$$|[x, y]| \leq |[x, x]|^{\frac{1}{2}} \cdot |[y, y]|^{\frac{1}{2}}, \quad x, y \in \mathcal{L}.$$

Proof. Consider the case that $\mathcal{L}$ is positive semi-definite, the case that $\mathcal{L}$ is negative semi-definite is settled analogously.

Set $A := [x, x], B := |[x, y]|$, and $C := [y, y]$, and let $\alpha \in \mathbb{C}, |\alpha| = 1$ be such that $\alpha[y, x] = B$. We have

$$0 \leq [x - t\alpha y, x - t\alpha y] = [x, x] - t\alpha[y, x] - t\bar{\alpha}[x, y] + t^2[y, y], \quad t \in \mathbb{R}$$

i.e. $A - 2tB + t^2C \geq 0$ for all $t \in \mathbb{R}$. If $C = 0$, thus also $B = 0$. If $C \neq 0$, we choose $t = \frac{B}{C}$ to obtain $AC - B^2 \geq 0$.

$\square$

2.4. Definition. Let $(H, (.,.))$ be a positive definite inner product space. H is called a Hilbert space if it is complete, i. e. $(H, \|\cdot\|)$ is a Banach space with respect to the induced norm $\|x\| := (x,x)^{\frac{1}{2}}$, $x \in H$.

2.5. Definition. Let H be a vector space with some inner product $(.,.)$. A pair consisting of the inner product space $(\hat{H}, (.,.))$ and a mapping $\iota : H \to \hat{H}$, is called a Hilbert space completion of $(H, (.,.))$, if the following properties are satisfied.

- $(\hat{H}, (.,.))$ is a Hilbert space.
- ι is linear and isometric, i. e. $(\iota x, \iota y) = (x, y)$ for all $x, y \in H$.
- $\overline{\iota(H)} = \hat{H}$,

where the closure is understood with respect to the Hilbert space norm.

Two completions $\langle (\hat{H}_1, (.,.)_1), \iota_1 \rangle$ and $\langle (\hat{H}_2, (.,.)_2), \iota_2 \rangle$ of $(H, (.,.))$ are called isomorphic, if there exists an isometric isomorphism $\varphi : \hat{H}_1 \to \hat{H}_2$ such that $\iota_2 = \varphi \circ \iota_1$.

$$\begin{tikzcd}
 & H \arrow[dl, hook', "\iota_1"'] \arrow[dr, hook, "\iota_2"] & \\
\hat{H}_1 \arrow[rr, dotted, "\varphi"'] & & \hat{H}_2
\end{tikzcd}$$

3. Reproducing Kernel Hilbert Spaces

3.1. Definition of Reproducing Kernel Hilbert Spaces and some general properties

The following is extracted from [Kal19, Chapter 3.1].

3.1. Definition. Let Ω be a non-empty set and let H be a vector space of complex valued functions, $H \subseteq \mathbb{C}^\Omega$. Let $(.,.)$ be an inner product on H such that $(H, (.,.))$ is a Hilbert space. We call H a Reproducing Kernel Hilbert space (RKHS) of functions on Ω if for every $x \in \Omega$ the point evaluation functional

$$\iota_x : \begin{cases} H & \to & \mathbb{C} \\ f & \mapsto & f(x) \end{cases} \tag{3.1}$$

is continuous.

3.2. Remark. Since point evaluation functionals are linear and H is a Hilbert space the following statements are equivalent.

1. ι_x is continuous for every $x \in H$.
2. ι_x is a bounded operator, i. e. for arbitrary $x \in \Omega$ there exists some positive constant $M_x > 0$ such that for all $f \in H$

$$|\iota_x(f)| = |f(x)| \leq M_x \|f\|.$$

3. From Riesz representation theorem we obtain a unique function $k_x \in H$ such that ι_x can be represented by the inner product via $\iota_x(f) = f(x) = (f, k_x)$ for all $f \in H$.

3.3. Definition. Let $(H, (.,.))$ be a RKHS and let $k_x \in H$ the unique function such that $f(x) = (f, k_x)$. Then we call the function

$$K : \begin{cases} \Omega \times \Omega & \to & \mathbb{C} \\ (x, y) & \mapsto & (k_y, k_x) \end{cases} \tag{3.2}$$

the reproducing kernel function of $(H, (.,.))$.

It is characteristic to reproducing kernel Hilbert spaces and satisfies the so-called reproducing kernel properties.

- $K(x, .) \in H$ for all $x \in \Omega$.

- $(f, K(x, .)) = f(x)$ for every $f \in H, x \in \Omega$.

3.4. Lemma. *Let $(H, (.,.))$ be a reproducing kernel Hilbert space on Ω, and let K be the corresponding reproducing kernel. Then the following statements hold.*

i) $K(x, y) = \overline{K(y, x)}$ *for all* $x, y \in H$.

ii) $\sum_{i,j=1}^{N} \overline{\lambda_i} \lambda_j K(x_i, x_j) \geq 0$ *for all* $N \in \mathbb{N}, x_1, \ldots, x_N \in \Omega, \lambda_1, \ldots, \lambda_N \in \mathbb{C}$.

iii) *The functions* $k_x \in H$ *are dense, i. e.* $cls\{k_x : x \in \Omega\} = H$.

Proof.

i) For all $x, y \in \Omega$ we have

$$K(x, y) = (k_y, k_x) = \overline{(k_x, k_y)} = \overline{K(y, x)}.$$

ii) for $N \in \mathbb{N}, x_1, \ldots, x_N \in \Omega, \lambda_1, \ldots, \lambda_N \in \mathbb{C}$ we get

$$\sum_{i,j=1}^{N} \bar{\lambda}_i \lambda_j \underbrace{K(x_i, x_j)}_{=(k_i, k_j)} = \left(\sum_{i=1}^{N} \lambda_i k_{x_i}, \sum_{j=1}^{N} \lambda_j k_{x_j} \right) \geq 0.$$

iii) Take $f \in \mathrm{cls}\{k_x : x \in \Omega\}^{\perp} = \mathrm{span}\{k_x : x \in \Omega\}^{\perp}$ so we can conclude

$$f(x) = (f, k_x) = 0$$

for every $x \in \Omega$, implying $f \equiv 0$.

□

3.5. Definition. Let Ω be a non-empty set and $K : \Omega \times \Omega \to \mathbb{C}$ be a function. Then K is called

a) hermitian kernel on Ω if for all $x, y \in \Omega$

$$K(x, y) = \overline{K(y, x)}$$

is satisfied.

b) positive semi-definite kernel on Ω if for all $N \in \mathbb{N}, x_1, \ldots, x_N \in \Omega, \lambda_1, \ldots, \lambda_N \in \mathbb{C}$

$$\sum_{i,j=1}^{N} \overline{\lambda_i} \lambda_j K(x_i, x_j) \geq 0$$

holds.

3.6. Remark.

- Let $K : \mathbb{C} \to \mathbb{C}$. Then K is a hermitian kernel on Ω, if and only if for all choices of $N \in \mathbb{N}$ and $x_1, ..., x_n \in \Omega$ the Gram matrix of K

  $$G(x_1, ..., x_N) := (K(x_j, x_k))_{j,k=1}^{N}$$

  is hermitian. It is a hermitian positive semi-definite kernel, if and only if all Gram matrices are non-negative.

- We have seen in Lemma 3.4 that the reproducing kernel function of a RKHS is a hermitian positive semi-definite kernel.

- Summing up, we can conclude that the assumptions made about the continuity of the evaluation mappings as well as the existence of a reproducing kernel function are both sufficient and necessary for H being a RKHS. They are indeed equivalent since the continuity of the evaluation mappings

  $$\iota_y(f) = (f, K(y, .)), f \in H$$

  is asserted from the continuity of the inner product.

These observations suggest the idea of building a RKHS for a given hermitian positive semi-definite kernel.

3.2. Standard Construction of Reproducing Kernel Hilbert Spaces

Starting with a function $K : \Omega \times \Omega \to \mathbb{C}$ one can ask whether a space H exists such that K satisfies some kind of "reproducing property", i. e. $f(x) = (f, K(x, .))$ for $f \in H$ and $x \in \Omega$. The first construction we are going to consider is the standard construction of a reproducing kernel Hilbert space given a hermitian positive semi-definite kernel. Our goal in this section will be to prove the following theorem which is extracted from [Kal19, Chapter 3.2]. Later, we will explore other constructions made on different assumptions on our function K.

3.7. Theorem. *Let Ω be a non-empty set and $K : \Omega \times \Omega \to \mathbb{C}$ be a hermitian positive semi-definite kernel. Then there exists a unique reproducing kernel Hilbert space such that K is the corresponding reproducing kernel.*

Proof. Consider the set

$$Y := \{\Phi : \Omega \to \mathbb{C} \mid \Phi(x) \neq 0 \text{ for only finitely many } x \in \Omega\} \tag{3.3}$$

and let us endow this space with a hermitian positive semi-definite sesquilinerform

$$\big\langle \Phi, \Psi \big\rangle_Y = \sum_{x,y \in \Omega} \Phi(x) \Psi(y) K(x,y)$$

which is well-defined since only finitely many terms appear in the sum. Now take a look at the following subspace $Y^\perp = \{\Phi \in Y | \langle \Phi, \Psi \rangle_Y = 0 \text{ für alle } \Psi \in Y\}$.

From Cauchy-Schwarz inequality we get for $\Phi \in Y^\circ = \{\Phi \in Y | \langle \Phi, \Phi \rangle_Y = 0\}$

$$0 \leq |\big\langle \Phi, \Psi \big\rangle_Y| \leq \|\Phi\|_Y \|\Psi\|_Y = 0,$$

thus $\Phi \in Y^\perp$ respectively $Y^\circ = Y^\perp$. Thus we obtain by the quotient space Y/Y° equipped with the scalar product

$$\big\langle \Phi + Y^\circ, \Psi + Y^\circ \big\rangle := \big\langle \Phi, \Psi \big\rangle_Y$$

a suitable candidate for our pre-Hilbert space.

Now let us consider the associated Hilbert space completion $(\hat{Y}, \langle ., . \rangle)$. Notice alongside that every $\Phi \in Y$ can be written $\Phi(x) = \sum_{t \in \Omega, \Phi(t) \neq 0} \Phi(t) \delta_t(x)$ and accordingly $Y/Y^\circ = \text{span}\{\delta_t + Y^\circ : t \in \Omega\}$ holds.

Lastly, regard the linear mapping

$$\Lambda : \begin{cases} \hat{Y} \to \mathbb{C}^\Omega \\ \hat{f} \mapsto (t \mapsto \big\langle \hat{f}, \delta_t + Y^\circ \big\rangle). \end{cases}$$

It holds $0 = \Lambda \hat{f}(t) = \langle \hat{f}, \delta_t + Y^\circ \rangle$ for all $t \in \Omega$ implying $\hat{f} \perp Y/Y^\circ$ and thus $f \equiv 0$. This conludes the injectivity of Λ.

$H := \Lambda(\hat{Y})$ will denote our RKHS via

$$\big(f, g\big)_H = \big\langle \Lambda^{-1} f, \Lambda^{-1} g \big\rangle.$$

The reproducing kernel property is satisfied since

$$f(x) = \Lambda(\hat{f})(x) = \big\langle \hat{f}, \delta_x + Y^\circ \big\rangle = \big(\Lambda(\hat{f}), \Lambda(\delta_x + Y^\circ)\big)_H = \big(f, \Lambda(\delta_x + Y^\circ)\big)_H.$$

Now consider the associated kernel function to our candidate H, which matches with our original kernel as follows from

$$\begin{aligned} K_H(x,y) &:= \big(\Lambda(\delta_y + Y^\circ), \Lambda(\delta_x + Y^\circ)\big)_H \\ &= \big\langle \delta_y + Y^\circ, \delta_x + Y^\circ \big\rangle = \big\langle \delta_y, \delta_x \big\rangle_Y \\ &= \sum_{s,t \in \Omega} \delta_y(s) \overline{\delta_x(t)} K(s,t) = K(x,y). \end{aligned}$$

Thus we have shown the existence of a reproducing kernel Hilbert space H but we still need to prove uniqueness. Consider a second reproducing kernel Hilbert space $\tilde{H} \subseteq C^{\Omega}$ such that K is the belonging kernel function. For given $t \in \Omega$ let us denote by $k_t \in H$ respectively $\tilde{k_t} \in \tilde{H}$ the elements fulfilling the reproducing kernel property $f(t) = (f, k_t)$ and $\tilde{f}(t) = (\tilde{f}, \tilde{k_t})$ for all $f \in H$, $\tilde{f} \in \tilde{H}$. By assumption we already know that

$$(k_y, k_x)_H = K(x, y) = (\tilde{k_y}, \tilde{k_x})_{\tilde{H}}$$

which concludes by linearity that

$$\left(\sum_{n=1}^{N} \lambda_n k_{y_n}, \sum_{n=1}^{M} \mu_n k_{x_n}\right)_H = \left(\sum_{n=1}^{N} \lambda_n \tilde{k_{y_n}}, \sum_{n=1}^{M} \mu_n \tilde{k_{x_n}}\right)_{\tilde{H}}$$

In particular

$$V : \begin{cases} \operatorname{span}\{k_t : t \in \Omega\} \to \operatorname{span}\{\tilde{k_t} : t \in \Omega\} \\ \qquad \sum_{n=1}^{N} \lambda_n k_{y_n} \mapsto \sum_{n=1}^{N} \lambda_n \tilde{k_{y_n}} \end{cases}$$

describes a well-defined isometric linear mapping. Both the domain $\operatorname{span}\{k_t : t \in \Omega\}$ as well as the image $\operatorname{span}\{\tilde{k_t} : t \in \Omega\}$ are dense in H and $\tilde{H}$ respectively. Accordingly, we get a unitary, bounded and linear continuation $U : H \to \tilde{H}$, satisfying

$$f(t) = (f, k_t)_H = (U(f), U(k_t))_{\tilde{H}} = (U(f), \tilde{k_t})_{\tilde{H}} = U(f)(t)$$

for every $f \in H$, $t \in \Omega$. Thus U is the identity mapping and $H = \tilde{H}$.

□

3.8. Remark.

Hermitian positive semi-definite kernels and reproducing kernel Hilbert spaces can be treated equally in the following way. Every reproducing kernel is hermitian positive semi-definite, and every positive semi-definite kernel defines a unique RKHS, of which it is the unique reproducing kernel.

3.3. Examples of Reproducing Kernel Hilbert Spaces

3.9. Example. The space

$$\mathfrak{l}^2(\mathbb{N}_0) = \left\{ (a_n)_{n \in \mathbb{N}_0} \in \mathbb{C}^{\mathbb{N}_0} \middle| \sum_{n=0}^{\infty} |a_n^2| < +\infty \right\}$$

equipped with the scalar product

$$((a_n)_{n \in \mathbb{N}_0}, (b_m)_{m \in \mathbb{N}_0})_{\mathfrak{l}^2(\mathbb{N}_0)} = \sum_{n=0}^{\infty} a_n \overline{b_n}$$

is a RKHS since the elements of this space are functions from $\Omega = \mathbb{N}_0$ to $\mathbb{C}$ and the point evaluation functionals satisfy

$$|\iota_m((a_n)_{n\in\mathbb{N}_0})| = |a_m| \leq \|(a_n)_{n\in\mathbb{N}_0}\|_{l^2(\mathbb{N}_0)}.$$

The reproducing kernel function of this RKHS is simply given by the following function

$$\delta_m(n) \begin{cases} 0, & \text{if} \quad m \neq n \\ 1, & \text{if} \quad m = n. \end{cases}$$

via $K(m,l) = (\delta_l, \delta_m)$ where $m, l \in \mathbb{N}_0$.

3.10. Example. Consider the Hardy space

$$H_2 := \left\{ f : \mathbb{D} \to \mathbb{C} : f(z) = \sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} |a_n|^2 < \infty \right\}$$

We will show that it is a RKHS on the unit disk $\Omega = \mathbb{D} = \{z \in \mathbb{C} : 0 < |z| < 1\}$. Due to $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ we get $\lim_{n\to\infty} a_n = 0$, which means that $\sum_{n=0}^{\infty} a_n z^n$ has a convergence radius greater or equal to 1. Thus $f \in H_2$ seem to be analytic functions on $\mathbb{D}$. We can easily check that

$$\phi : \begin{cases} \ell^2(\mathbb{N}_0) \to H_2 \\ (a_n) \mapsto \sum_{n=0}^{\infty} a_n z^n \end{cases}$$

is an isomorphism, where H_2 is considered as a linear subspace of $\mathbb{C}^{\mathbb{D}}$. Now let us set $(f,g) := \left(\phi^{-1}f, \phi^{-1}g\right)_{\ell^2(N_0)} = \sum_{n=0}^{\infty} a_n \overline{b_n}$ for $f(z) = \sum a_n z^n, g(z) = \sum b_n z^n$. From this H_2 defines a Hilbert space. For $w \in \mathbb{D}$ we get

$$f(w) = \sum_{n=0}^{\infty} a_n w^n = \left(\sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} \overline{w^n} z^n \right) = \left(\sum_{n=0}^{\infty} a_n z^n, \frac{1}{1 - \bar{w}z} \right)$$

where the function $z \mapsto \frac{1}{1-\bar{w}z}$ lies in H_2 because $\sum_{n=0}^{\infty} |w|^{2n} = \frac{1}{1-|w|^2} < \infty$. In particular $f \mapsto f(w)$ is continuous on H_2.Thus H_2 is a Reproducing Kernel Hilbert space. Obviously it holds $k_w(z) = \frac{1}{1-\bar{w}z}$, and the kernel function is given by $K(z,w) = (k_w, k_z) = k_w(z) = \frac{1}{1-\bar{w}z}$.

3.11. Example. The Sobolev space $W^{1,2}([0,1]) = \{f : f : [0,1] \to \mathrm{C}, f \in AC[0,1], f' \in L^2[0,1]\}$ with the inner product

$$(f_1, f_2)_{W^{1.2}} = \int_0^1 \left(f_1 \cdot \bar{f}_2 + f_1' \cdot \bar{f}_2' \right) d\lambda$$

defines a RKHS on the set $\Omega = [0,1]$. Consider the differential operator

$$D := \left\{(f;g) \in L^2[0,1] \times L^2[0,1] : f \in AC, f' = g\right\}.$$

We will see that D is a closed subspace of $L^2[0,1] \times L^2[0,1]$. If we restrict the summation scalar product of $L^2[0,1] \times L^2[0,1]$ on D, it follows that $((f_1;g_1),(f_2;g_2)) = (f_1,f_2)_{W^{1,2}}$. Thus $W^{1,2}$ is as a closed subspace of a Hilbert space a Hilbert space by itself. It follows for $f \in W^{1,2}$

$$f(1) - f(x) = \int_0^1 f'(t)\chi_{[x,1]}(t)d\lambda(t)$$

and due to absolute continuity

$$\int_0^1 t \cdot f'(t)d\lambda(t) = t \cdot f(t)|_0^1 - \int_0^1 f(t)d\lambda(t) = f(1) - \int_0^1 f(t)d\lambda(t)$$

which concludes if we substitute $f(1)$

$$f(x) = \int_0^1 t \cdot f'(t)d\lambda(t) + \int_0^1 f(t)d\lambda(t) - \int_0^1 f'(t)\chi_{[x,1]}(t)d\lambda(t)$$

Ultimately the point evaluation can be derived by 3 integrals. They all are continuous linear functions depending on f, which can be seen as for the first summand via Cauchy-Schwarz inequality

$$\left|\int_0^1 t \cdot f'(t)dt\right|^2 \leq \int_0^1 |t|^2 dt \cdot \int_0^1 \left|f'(t)\right|^2 dt \leq \frac{1}{3} \cdot \|f\|_{W^{1,2}}^2$$

4. Conditionally Positive Definite Kernels

4.1. Useful Properties of CP Kernels

This chapter consists of definitions and results from [BzCS12] and [AB09].

4.1. Definition. Suppose that $\kappa \in \mathbb{N}_0$ is some natural number. A hermitian kernel $K : \Omega \times \Omega \to \mathbb{C}$ is called conditionally positive (semi-)definite of order κ, if there exists a κ dimensional space $U \subseteq \mathbb{C}^{\Omega}$ of complex-valued functions on Ω, such that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i \overline{c_j} K(x_i, x_j) > (\geq) \quad 0 \tag{4.1}$$

for all choices of $n \in \mathbb{N}, x_1, \ldots, x_n \in \Omega$, and $c_1, \ldots, c_n \in \mathbb{C}$ satisfying

$$\forall u \in U : \quad \sum_{i=1}^{n} c_i u(x_i) = 0. \tag{4.2}$$

4.2. Remark.

- Note that any conditionally positive (semi-)definite kernel of order κ is also conditionally positive (semi-)definite of any higher order, since (4.2) gets more restrictive when U becomes larger.

- Let us consider the Gram matrix

  $$G(x_1, \ldots, x_n) := (K(x_i, x_j))_{i,j=1}^{n} \tag{4.3}$$

  for arbitrary $x_1, \ldots, x_n \in \Omega$. We remember that the Gram matrix of a positive semi-definite kernel has no negative eigenvalues, whereas now it can be the case. In matrix form the definition of conditionally positive semi-definite kernels can be written as follows

  $$\bar{\mathbf{c}}^T G \mathbf{c} \geq 0$$

  for all $c \in \mathbb{C}^{\mathbb{N}}$ satisfying

  $$\bar{\mathbf{u}}^T \mathbf{c} = 0 \quad \forall u \in U,$$

where $\mathbf{u} = (u(x_1), \ldots, u(x_n))^T$ and the complex conjugation is understood componentwise.

4.3. Definition. Suppose $\kappa \in \mathbb{N}_0$. A hermitian kernel $K : \Omega \times \Omega \to \mathbb{C}$ is said to have κ negative squares, if for every choice of $n \in \mathbb{N}, x_1, \ldots, x_n \in \Omega$, the Gram matrix of K has at most κ negative eigenvalues, and there are $y_1, \ldots, y_m \in \Omega$ such that $G(y_1, \ldots, y_m)$ has exactly κ negative eigenvalues (counted with multiplicities).

4.4. Proposition. *Every conditionally positive semi-definite kernel of order $\kappa \in \mathbb{N}$ has at most κ many negative squares.*

Proof. Choose $n \geq \kappa$ points $x_1, \ldots, x_n \in \Omega$ and consider the Gram matrix $G(x_1, \ldots, x_n)$. We already know that for every $\mathbf{c} \in \mathbb{C}^n$

$$\bar{\mathbf{c}}^T G \mathbf{c} \geq 0$$

holds, whenever

$$\bar{\mathbf{u}}^T \mathbf{c} = 0 \quad \forall u \in U.$$

For such $\mathbf{c} \in \mathbb{C}^n$ this means in particular that

$$\mathbf{c} \in \operatorname{span}\left\{ \left(\overline{u(x_i)}\right)_{i=1}^n : u \in U \right\}^{\perp} \subseteq \mathbb{C}^n.$$

Consider the following semi-linear mapping

$$\Psi_{x_1,\ldots,x_n} : \begin{cases} U \to & \mathbb{C}^n \\ u \mapsto & \begin{pmatrix} \overline{u(x_1)} \\ \vdots \\ \overline{u(x_n)} \end{pmatrix} \end{cases}$$

For a given basis $\{u_1, ..., u_\kappa\}$ in U we can imply by the principle of continuation from Linear Algebra that

$$\begin{aligned} \kappa = \dim U &= \dim \operatorname{span}\{u_1, ..., u_\kappa\} \\ &\geq \dim \operatorname{span}\{\Psi_{x_1,\ldots,x_n}(u_1), ..., \Psi_{x_1,\ldots,x_n}(u_\kappa)\} = \dim \Psi_{x_1,\ldots,x_n}(U). \end{aligned}$$

For a given negative eigenvalue and its corresponding eigenvector, it holds

$$\tilde{c} \in \operatorname{span}\left\{ \left(\overline{u(x_i)}\right)_{i=1}^n : u \in U \right\}^{\perp\perp} = \overline{\Psi_{x_1,\ldots,x_n}(U)}.$$

Thus, there can only be a maximal of κ many negative eigenvalues since there can only be a maximal of κ many linearly independent corresponding eigenvector.

$\square$

4.5. Definition. $\Xi = \{\xi_1, \dots, \xi_N\} \subseteq \Omega$ is said to be U-unisolvent if the linear application

$$L^{\Xi} : U \ni u \mapsto (u(\xi_1), \dots, u(\xi_N)) \in \mathbb{C}^N$$

is injective, or equivalently, if the only $u \in U$ which vanishes on every $\xi_i \in \Xi$ is the zero function $0 \in U$. A U-unisolvent set Ξ is said to be minimal if $|\Xi| = \dim U = \kappa$.

4.6. Remark. Consider a minimal U-unisolvent set $\Xi = \{\xi_1, \dots, \xi_\kappa\} \subseteq \Omega$ where $|\Xi| = \dim U = \kappa$. Then we know that for every $i \in \{1, ..., \kappa\}$ there exists a function $u_i \in U$ such that for all $k \in \{1, ..., \kappa\}$ it holds

$$u_i(\xi_k) = \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k. \end{cases}$$

The functions $\{u_i : i = 1, \dots, \kappa\}$ are often denoted as the associated Lagrange basis. A U-unisolvent set is such that functions in U can be interpolated exactly by this set of centers. Thus every $u \in U$ can be written as

$$u(x) = \sum_{i=1}^{\kappa} u(\xi_i) u_i(x).$$

4.2. Construction of Hilbert Spaces via CP Kernels

The following constructions and ideas are put together from [Wen04, Chapter 10.3].
For a given conditionally positive definite function one can ask how a Hilbert space can be constructed such that our kernel satisfies some reproducing property?

4.7. Definition. Let $K : \Omega \times \Omega \to \mathbb{C}$ be a conditionally positive definite kernel of order $\kappa \in \mathbb{N}_0$, with a corresponding κ dimensional space $U \subseteq \mathbb{C}^{\Omega}$ of complex-valued functions on Ω. Then we consider the space

$$\begin{aligned} F_K(\Omega) := \Bigg\{ & \sum_{j=1}^{N} \alpha_j K(\cdot, x_j) : N \in \mathbb{N}, \alpha \in \mathbb{C}^N, x_1, \dots, x_N \in \Omega \\ & \text{with } \sum_{j=1}^{N} \alpha_j u(x_j) = 0 \text{ for all } u \in U \Bigg\} \end{aligned} \tag{4.4}$$

endowed with the inner product

$$\left(\sum_{j=1}^{N}\alpha_j K(.,x_j),\sum_{k=1}^{M}\beta_k K(.,y_k)\right)_K := \sum_{j=1}^{N}\sum_{k=1}^{M}\alpha_j\overline{\beta_k}K(y_k,x_j). \tag{4.5}$$

4.8. Remark.

- We will see that $F_K(\Omega)$ happens to be an appropriate candidate for our Pre-Hilbert space.
- To show well-definedness of our inner product let us consider $f=\sum_{k=1}^{M}\beta_k K(.,y_k)\in F_K(\Omega)$ and $\sum_{j=1}^{N}\alpha_j K(.,x_j)=0\in F_K(\Omega)$ such that we can conclude

$$(0,f)_K=\sum_{j=1}^{N}\sum_{k=1}^{M}\alpha_j\overline{\beta_k}K(y_k,x_j)=\sum_{j=1}^{N}\overline{\beta_k}\underbrace{\left(\sum_{k=1}^{M}\alpha_j K(y_k,x_j)\right)}_{=0}=0.$$

- In general our goal is similar to the positive definite kernel case. We want to consider the completion with respect to $(.,.)_K$. In order to do so, we need to interpret the elements of the completion as functions. Contrary to the general construction, $K(.,x)$ is not always in $F_K(\Omega)$ for arbitrary $x\in\Omega$. Thus it will not make sense to use the property $f(x):=(f,K(.,x))_K$ to make the functional

$$\iota_x:\begin{cases}\overline{F_K(\Omega)}^{(.,.)_K} & \to \mathbb{C}\\ f & \mapsto f(x)\end{cases}$$

continuous. We need to show this in another way.

4.9. Definition. Let $\Xi\subseteq\Omega$ be a minimal U-unisolvent set and let $\{u_i : i=1,\dots,\kappa\}$ denote the associated Lagrange basis. Let us define the following useful notations.

a) The projection operator on U

$$\Pi_U:\mathbb{C}^\Omega\ni f\mapsto \Pi_U(f)=\sum_{k=1}^{\kappa}f(\xi_k)u_k\in U. \tag{4.6}$$

b) (i) The set of complex measures with finite support $\mathcal{M}$

$$\boldsymbol{\mu}\in\mathcal{M}\Leftrightarrow \exists \mathbf{x}_1,\dots,\mathbf{x}_N\in\Omega \text{ and } \mu_1,\dots,\mu_N\in\mathbb{C},\boldsymbol{\mu}=\sum_{k=1}^{N}\mu_k\iota_{\mathbf{x}_k}$$

where $\iota_{\mathbf{x}}$ denotes the Dirac measure on Ω.

(ii) The subspace $\mathcal{M}_U$ of measures lying in $\mathcal{M}$ and vanishing on U :

$$\boldsymbol{\mu} \in \mathcal{M}_U \Leftrightarrow \boldsymbol{\mu}(u) = 0, \forall u \in U$$

where $\boldsymbol{\mu}(f) = \sum_{k=1}^{N} \mu_k f(\mathbf{x}_k)$ denotes the integral of $f \in \mathbb{C}^{\Omega}$.

We can endow this space with an inner product

$$< \boldsymbol{\mu}, \boldsymbol{\lambda} >_{\mathcal{M}_U, K} := \Big(\boldsymbol{\mu}(x \mapsto K_x), \boldsymbol{\lambda}(x \mapsto K_x) \Big)_K$$

c) Lastly, let us define the following functional

$$\iota_{(x)} := \iota_x - \sum_{j=1}^{\kappa} u_j(x) \iota_{\xi_j} \tag{4.7}$$

and the function

$$G(.,x) := K(.,x) - \sum_{j=1}^{\kappa} u_j(x) K(.,\xi_j). \tag{4.8}$$

4.10. Remark.

- Notice how the point evaluation functionals ι_x can be interpreted as an integral operator applied to the corresponding Dirac measures on Ω. Thus the notation is uniform and consistent.
- We can write the function $G(.,x) = K(.,x) - \Pi_U K(.,x)$ via the operator from (4.6). Note that $G(.,\xi_k) = 0$ for $k \in \{1,\ldots,\kappa\}$.
- For $u \in U$ the Lagrange interpolation is exact. We obtain $\iota_{(x)}(u) = 0$. Consequently we get that $\iota_{(x)} \in F_K(\Omega)'$. Now let us extend this idea to the completion $\mathcal{F}_K(\Omega)$ by continuity.

4.11. Lemma. *Consider the linear mapping*

$$R : \mathcal{F}_K(\Omega) \to C(\Omega); \quad f \mapsto (x \mapsto (f, G(.,x))_K). \tag{4.9}$$

The range lies indeed in $C(\Omega)$ and R is injective.

Proof. Our mapping is well-defined since

$$|R(f)(x) - R(f)(y)| = |(f, G(.,x) - G(.,y))_K| \leq \|f\|_K \|G(.,x) - G(.,y)\|_K.$$

Now we will show injectivity. Take $f \in \mathcal{F}_K(\Omega)$ such that $Rf = 0$.

$$Rf = 0 \Leftrightarrow \forall x \in \Omega : (f, G(.,x))_K = 0$$

Now consider some arbitrary $h = \sum_{j=1}^{N} \alpha_j K(.,x_j) \in F_K(\Omega)$. This function can also be written in the following way

$$\begin{aligned}\sum_{J=1}^{N} \alpha_j G(\cdot, x_j) &= \sum_{J=1}^{N} \alpha_j \left(K(\cdot, x) - \sum_{k=1}^{\kappa} u_k(x_j) K(\cdot, \xi_k) \right) \\ &= \sum_{J=1}^{N} \alpha_j K(\cdot, x) - \sum_{k=1}^{\kappa} K(\cdot, \xi_k) \underbrace{\sum_{J=1}^{N} \alpha_j u_k(x_j)}_{=0} = h\end{aligned}$$

This implies

$$(f,h)_K = \sum_{j=1}^{N} \alpha_j (f, G(.,x_j))_K = 0$$

for arbitrary $h \in F_K(\Omega)$, thus $f \in F_K(\Omega)^{\perp} = \overline{F_K(\Omega)}^{\perp} = \{0\}$.

□

4.12. Remark. $R(\mathcal{F}_K(\Omega))$ does not suit as a candidate for our Hilbert space such that the reproducing kernel is K. Considering how R acts on $F_K(\Omega)$ we can construct reproducing kernel Hilbert spaces. The following space is often called native space in literature (compare [Wen04]). Our next goal will be to make $R(\mathcal{F}_K(\Omega))$ as well as the native space a RKHS.

4.13. Definition. Consider the following space

$$\mathcal{N}_K(\Omega) := R(\mathcal{F}_K(\Omega)) \oplus U \tag{4.10}$$

equipped with a semi-inner product

$$\left(f,g\right)_{\mathcal{N}_K(\Omega)} := \left(R^{-1}(f - \Pi_U f), R^{-1}(g - \Pi_U g)\right)_K \tag{4.11}$$

4.14. Proposition. *The bilinear form $(.,.)_{\mathcal{N}_K(\Omega)}$ is an inner product on the space*

$$R\left(\mathcal{F}_K(\Omega)\right) = \left\{f \in \mathcal{N}_K(\Omega) : f\left(\xi_k\right) = 0, 1 \le k \le \kappa\right\},$$

which makes this space a Hilbert space. Moreover, this space is a RKHS with the reproducing kernel

$$\Phi(x,y) :=< \iota_{(x)}, \iota_{(y)} >_{\mathcal{M}_U,K} \tag{4.12}$$

Proof. We know that the linear mapping $R : \mathcal{F}_K(\Omega) \to R(\mathcal{F}_K(\Omega))$ is isometric and bijective. Since $\mathcal{F}_K(\Omega)$ is a Hilbert space, so is $R(\mathcal{F}_K(\Omega))$. In particular the bilinear form $(.,.)_{\mathcal{N}_K(\Omega)}$ becomes an inner product on $R(\mathcal{F}_K(\Omega))$. Now let us compute the reproducing kernel.

$$\begin{aligned}
< \iota_{(x)}, \iota_{(y)} >_{\mathcal{M}_U,K} \underset{(4.7)}{=} & < \iota_x - \sum_{j=1}^{\kappa} u_j(x)\iota_{\xi_j}, \iota_y - \sum_{i=1}^{\kappa} u_j(y)\iota_{\xi_i} >_{\mathcal{M}_U,K} \\
= & K(x,y) - \sum_{k=1}^{\kappa} u_k(x) K(\xi_k, y) - \sum_{\ell=1}^{\kappa} u_\ell(y) K(x, \xi_l) \\
& + \sum_{k=1}^{\kappa}\sum_{\ell=1}^{\kappa} u_k(x) u_\ell(y) K(\xi_k, \xi_l)
\end{aligned}$$

The kernel Φ is a symmetric function and satisfies $\Phi(.,y) = RG(.,y)$ for all $y \in \Omega$. This shows that $\Phi(.,y)$ is indeed an element of $R(\mathcal{F}_K(\Omega))$ for every $y \in \Omega$. Finally, for $f \in R(\mathcal{F}_K(\Omega))$ and $x \in \Omega$ we obtain

$$\begin{aligned}
(f, \Phi(.,x))_{\mathcal{N}_K(\Omega)} &= \left(R^{-1} f, R^{-1}\Phi(.,x)\right)_K \\
&= \left(R^{-1} f, G(.,x)\right)_K \\
&= \Pi_U f(x) + \left(R^{-1}(f - \Pi_U f), K(.,x) - \Pi_U K(.,x)\right)_K \\
&= f(x)
\end{aligned}$$

□

4.15. Theorem. *The space $\mathcal{N}_K(\Omega)$ from (4.10) together with the inner product*

$$(f,g) := (f,g)_{\mathcal{N}_K(\Omega)} + \sum_{k=1}^{\kappa} f(\xi_k)\, g(\xi_k) \tag{4.13}$$

is a RKHS. Its reproducing kernel is given as

$$\tilde{K}(x,y) = \Phi(x,y) + \sum_{k=1}^{\kappa} u_k(x) u_k(y). \tag{4.14}$$

Proof. Obviously, we obtain a hermitian, non-negative sesquilinear form. To show positive definiteness consider for $f \in \mathcal{N}_K(\Omega)$

$$0 = (f, f) = (f, f)_{\mathcal{N}_\uparrow(\Omega)} + \sum_{k=1}^{Q} |f(\xi_k)|^2 .$$

Since $(f, f)_{\mathcal{N}_K(\Omega)} = 0$ we get that $f \in U$. From $\sum_{k=1}^{Q} |f(\xi_k)|^2 = 0$, we can imply $f(\xi_k) = 0$ for $1 \leq k \leq \kappa$. Since $\Xi = \{\xi_1, \ldots, \xi_N\}$ is U-unisolvent, f must be equal to 0.

Let us continue with the reproducing kernel property. First consider the following since u_ℓ is a Lagrangian basis for Ξ and $\Phi(\xi_k, \cdot) = 0$.

$$\begin{aligned}
\sum_{k=1}^{\kappa} f(\xi_k) \tilde{K}(\xi_k, x) &= \sum_{k=1}^{\kappa} f(\xi_k) \Phi(\xi_k, x) + \sum_{k=1}^{\kappa} f(\xi_k) \sum_{\ell=1}^{\kappa} u_\ell(x) u_\ell(\xi_k) \\
&= \sum_{k=1}^{\kappa} f(\xi_k) u_k(x).
\end{aligned}$$

In summary, we obtain the reproducing property

$$\begin{aligned}
f(x) &= \Pi_U f(x) + (f, \Phi(\cdot, x))_{\mathcal{N}_K(\Omega)} \\
&= \sum_{k=1}^{\kappa} f(\xi_k) \tilde{K}(\xi_k, x) + (f, \tilde{K}(\cdot, x))_{\mathcal{N}(\Omega)} \\
&= (f, \tilde{K}(\cdot, x))
\end{aligned}$$

since $\Phi(., x)$ and $\tilde{K}(., x)$ differ only by a polynomial.

$\square$

5. Pontryagin Spaces

5.1. Definition and Vocabulary

The first definitions and remarks are based on [BzCS12].
For given Hilbert space consider its corresponding antispace, i.e. the vector space endowed with its negative inner product.
5.1. Definition. Let Π be a space endowed with a scalar product. Π is now called a Pontryagin space if it can be written as a direct sum

$$\Pi = N \oplus N^{\perp} \tag{5.1}$$

where $\left(N^{\perp}, \langle\cdot,\cdot\rangle_{N\perp}\right)$ is a Hilbert space and $(N, \langle\cdot,\cdot\rangle_N)$ is the antispace of a finite-dimensional Hilbert space. Every decomposition of Π of the form (5.1) is called a fundamental decomposition.

5.2. Remark.

- For a given fundamental decomposition the Pontryagin space Π is equipped with an indefinite inner product

$$(x, y) = \langle x_+, y_+\rangle_{N^{\perp}} + \langle x_-, y_-\rangle_N\,, \quad x_+, y_+ \in N^{\perp}, x_-, y_- \in N. \tag{5.2}$$

- The subspace N denotes the negative subspace of Π. From the definition we immediately get that $(v, v) < 0$ for all $v \in N\backslash\{0\}$. Note that N is maximal with respect to this property, i.e., not included in any larger subspace.

- The subspace N is not unique in general. Nevertheless the dimension is unique and is called the index of the Pontryagin space.

- For a given Pontryagin space $(\Pi, (\cdot,\cdot))$ its induced Hilbert space is naturally given via the fundamental decomposition

$$[x, y] = (x_+, y_+) - (x_-, y_-)\,, \quad x_+, y_+ \in N^{\perp}, x_-, y_- \in N \tag{5.3}$$

  Let us denote by P_+ and P_- the projections to the according definite subspaces $N^{\perp}$ and N. Then we can define the fundamental symmetry operator $J := P_+ - P_-$ which satisfies

$$[x, y] = (Jx, y) \quad x, y \in \Pi$$

5.2. Well-Definedness and Useful Characterizations of Pontryagin Spaces

The following results are inspired by proofs from [KW14, Chapter 2] as well as [Wor10, Chapter 2.6].

5.3. Theorem. *For a given Pontryagin space* Π *the norm induced topology* $\mathcal{T} = \mathcal{T}(\|\cdot\|_{[\cdot,\cdot]})$ *does not depend on the choice of the fundamental decomposition since all the norms are equivalent.*

Proof. Let $N_1 \oplus N_1^\perp$ and $N_2 \oplus N_2^\perp$ be two different fundamental decompositions of our Pontryagin space $(\Pi, (\cdot,\cdot))$. Let us denote by $\|\cdot\|_1$ and $\|\cdot\|_2$ the corresponding norms by the induced scalar products. Ultimately, we want to show $\|\cdot\|_1 \sim \|\cdot\|_2$, i. e. there exist $\alpha, \beta > 0$ such that

$$\alpha\|x\|_1 \leq \|x\|_1 \leq \beta\|x\|_2$$

for all $x \in \Pi$. Denote by $P_+^1, P_-^1$and J_1 the fundamental projections and fundamental symmetry regarding $\|\cdot\|_1$ and equivalently P_+^2, P_-^2 and J_2 for $\|\cdot\|_2$. Since Π is not degenerated our fundamental symmetries are bijective and it holds

$$[x,y]_1 = (J_1 x, y), \quad [x,y]_2 = (J_2 x, y)\,, \quad x, y \in \Pi.$$

Because we know $(x,y) = [x,y]_1$ for $x, y \in N_1^\perp$, let us consider the Hilbert space $(N_1^\perp, [\cdot,\cdot]_1)$ and the following linear functionals

$$\varphi_y : \begin{cases} N_1^\perp & \to \mathbb{C} \\ \quad x & \mapsto (x,y) \end{cases}, \quad y \in \Pi$$

Obviously these functionals are part of the topological dual space, $\varphi_y \in \big(N_1^\perp, [\cdot,\cdot]_1)\big)'$ for every $y \in \Pi$. Thus by Cauchy-Schwarz inequality

$$|\varphi_y(x)| = |(x,y)| = |[x,y]_2| \leq \|x\|_2\|y\|_2, \quad x \in N_1^\perp, y \in \Pi.$$

we get that the family

$$\{\varphi_y : y \in \Pi, \|y\|_2 \leq 1\} \subseteq \Big(N_1^\perp, [\cdot,\cdot]_1\Big)'$$

is pointwise bounded. Applying the principle of uniform boundedness we obtain that the family is norm bounded, i. e.

$$\gamma := \sup_{\substack{y\in\Pi \\ \|y\|_2\leq 1}} \|\varphi_y\|_1 < \infty.$$

Here $\|\cdot\|_1$ denotes the norm with respect to $[\cdot,\cdot]_1$. Now we can conclude

$$\|x\|_2 = \sup_{\substack{y\in\Pi \\ \|y\|_2\leq 1}} |(x,y)| = \sup_{\substack{y\in\Pi \\ \|y\|_2\leq 1}} |\varphi_y(x)| \leq \sup_{\substack{y\in\Pi \\ \|y\|_2\leq 1}} \|\varphi_y\| \cdot \|x\|_1 = \gamma\|x\|_1, \quad x \in N_1^\perp$$

From this it follows

$$\begin{aligned}\left(\left\|P_+^1x\right\|_2\right)^2 \leq \gamma^2\left\|P_+^1x\right\|_1^2 = \gamma^2\left[P_+^1x, P_+^1x\right]_1 =& \gamma^2\left(P_+^1x, P_+^1x\right)\\ =& \gamma^2\left(P_+^1x, x\right) \leq \gamma^2\left\|P_+^1x\right\|_2\|x\|_2, \quad x \in \Pi.\end{aligned}$$

By dividing with $\left\|P_+^1x\right\|_2$ we get that $\left\|P_+^1x\right\|_2 \leq \gamma^2\|x\|_2$. In summary this concludes

$$\begin{aligned}\|x\|_1^2 = [x,x]_1 = (J_1x, x) \leq& \|J_1x\|_2\|x\|_2\\ =& \|(2P_+^1 - I)x\|_2\|x\|_2 \leq \left(2\gamma^2+1\right)\|x\|_2^2, \quad x \in \Pi.\end{aligned}$$

and we have showed that there exists a positive constant $\beta := \sqrt{2\gamma^2+1} > 0$ such that $\|x\|_1 \leq \beta\|x\|_2$ for all $x \in \Pi$. Out of symmetry reasons, we can analogously construct another positive constant $\frac{1}{\alpha}$ such that $\|x\|_2 \leq \frac{1}{\alpha}\|x\|_1$ respectively $\alpha\|x\|_2 \leq \|x\|_1$. Thus we get $\|\cdot\|_1 \sim \|\cdot\|_2$.

□

5.4. Proposition. *Let $\langle\Pi, [.,.]\rangle$ be a Pontryagin space, let $x_n \in \Pi, n \in \mathbb{N}$, and $x \in \Pi$. Then the following hold:*

(i) We have $\lim_{n\to\infty} x_n = x$ with respect to $\mathcal{T}$, if and only if there exists a dense subset D of Π, such that

$$\lim_{n\to\infty}[x_n, x_n] = [x,x], \quad \lim_{n\to\infty}[x_n, y] = [x,y], y \in D \tag{5.4}$$

(ii) The sequence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to $\mathcal{T}$ (i.e. with respect to some norm inducing $\mathcal{T}$), if and only if there exists a maximal negative subspace $\mathcal{M}$ of Π, such that

$$\lim_{n,m\to\infty}[x_n - x_m, x_n - x_m] = 0, \quad \lim_{n,m\to\infty}[x_n - x_m, y] = 0, y \in \mathcal{M} \tag{5.5}$$

Proof. Let $\mathfrak{J} = (\Pi_+, \Pi_-)$ be any fundamental decomposition of Π. Then $|[x,y]| \leq \|x\|_{\mathfrak{J}}\|y\|_{\mathfrak{J}}$, $x, y \in \Pi$. Thereby, we may take $D := \Pi$.

Conversely, assume that D is a dense subset of Π such that (5.4) holds. Clearly, we may assume without loss of generality that D is a linear subspace of Π. First let us construct a maximal negative subspace $\mathcal{M}$ which is contained in D. Since we know that $\Pi = \overline{D}$ consider a maximal negative subspace $\mathcal{N}$ of Π with dimension κ. Write $\mathcal{N} = \operatorname{span}\{x_1, \dots, x_\kappa\}$. Then the matrix $A := ([x_i, x_j])_{i,j=1}^\kappa$ is negative definite, i.e. all zeros of the polynomial

$$p(\lambda) := \det(A - \lambda I)$$

are negative. Since a polynomial depends continuously on its coefficients in the topology of locally uniform convergence, there exists some $\epsilon > 0$ such that the polynomial $\det(A' - \lambda I)$ has exclusively negative zeros whenever $\|A' - A\| < \epsilon$. Here $\|\cdot\|$ denotes some matrix norm.

Since D is dense in $\overline{D}$, there exist elements $x'_1, \ldots, x'_\kappa \in D$, such that

$$\left\| \left([x'_i, x'_j] \right)_{i,j=1}^{\kappa} - ([x_i, x_j])_{i,j=1}^{\kappa} \right\| < \epsilon$$

By what we said above, this implies that the matrix $\left(\left[x'_i, x'_j \right] \right)_{i,j=1}^{\kappa}$ is negative definite. Thus the space

$$\mathcal{M} := \operatorname{span} \left\{ x'_1, \ldots, x'_\kappa \right\} \leq D$$

is a maximal negative subspace of Π contained in D.

Now let us denote by $\mathfrak{J} := \left(\mathcal{M}^\perp, \mathcal{M} \right)$ the fundamental decomposition of Π and by $P_\mathfrak{J}^+$ and $P_\mathfrak{J}^-$ the corresponding projections.

Consider the sequences $\left(P_\mathfrak{J}^+ x_n \right)_{n \in \mathbb{N}}$ and $\left(P_\mathfrak{J}^- x_n \right)_{n \in \mathbb{N}}$. The second relation in (5.4), together with our choice of $\mathcal{M}$ as a subspace of D, gives

$$\lim_{n \to \infty} \left[P_\mathfrak{J}^- x_n, y \right] = \lim_{n \to \infty} [x_n, y] = [x, y] = \left[P_\mathfrak{J}^- x, y \right], y \in \mathcal{M}$$

Since $\mathcal{M}$ is finite-dimensional and negative definite, this implies that

$$\lim_{n \to \infty} P_\mathfrak{J}^+ x_n = P_\mathfrak{J}^+ x \quad \text{ with respect to } \; \| \cdot \|_\mathfrak{J} |_\mathcal{M} = (-[.,.])^{\frac{1}{2}}$$

In total $\lim_{n \to \infty} x_n = x$ with respect to $\| \cdot \|_\mathfrak{J}$. Assume that (5.5) holds. We argue similarly, and show that both of $\left(P_\mathfrak{J}^+ x_n \right)_{n \in \mathbb{N}}$ and $\left(P_\mathfrak{J}^- x_n \right)_{n \in \mathbb{N}}$ are Cauchy sequences in the norm $\| \cdot \|_\mathfrak{J}$. First,

$$\lim_{n,m \to \infty} \left[P_\mathfrak{J}^- x_n - P_\mathfrak{J}^- x_m, y \right] = 0, y \in \mathcal{M}$$

Again finite dimensionality implies that $(P_\mathfrak{J}^+ x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in the norm $(-[.,.])^{\frac{1}{2}}$. Next, we compute

$$\begin{aligned} &\lim_{n,m \to \infty} \left[P_\mathfrak{J}^+ x_n - P_\mathfrak{J}^+ x_m, P_\mathfrak{J}^+ x_n - P_\mathfrak{J}^+ x_m \right] = \\ &= \lim_{n,m \to \infty} \left([x_n - x_m, x_n - x_m] - \left[P_\mathfrak{J}^- x_n - P_\mathfrak{J}^- x_m, P_\mathfrak{J}^- x_n - P_\mathfrak{J}^- x_m \right] \right) = 0. \end{aligned}$$

□

5.5. Remark. With the knowledge about convergence of our topology, we could have shown the uniqueness of the norm induced topology from Theorem 5.3 in the following way.

Shortcut Proof. Let $N_1 \oplus N_1^\perp$ and $N_2 \oplus N_2^\perp$ be two different fundamental decompositions of our Pontryagin space $(\Pi, (\cdot, \cdot))$. Let us denote by $\| \cdot \|_1$ and $\| \cdot \|_2$ the corresponding norms by the induced scalar products. Consider the identical mapping $\mathrm{id} : \langle \Pi, \| \cdot \|_1 \rangle \to \langle \Pi, \| \cdot \|_2 \rangle$.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in Π, where $x_n \xrightarrow{\|\cdot\|_1} x, x_n \xrightarrow{\|\cdot\|_2} x'$. It holds

$$(x_n, y) \to (x, y), (x_n, y) \to \left(x', y\right), \quad y \in \Pi$$

and thus $x = x'$. Now we know by the closed graph theorem that id is continuous and by the open mapping theorem we obtain a homeomorphism, $\mathcal{T}(\|\cdot\|_1) = \mathcal{T}(\|\cdot\|_2)$.

□

5.6. Proposition. *Let $\langle\Pi, [\cdot,\cdot]\rangle$, be an inner product space. Then Π is a Pontryagin space, if and only if Π is nondegenerated and there exists a positive subspace $\mathcal{L} \leq \Pi$ with finite codimension, i. e. $\dim \Pi/\mathcal{L} < \infty$, such that $\mathcal{L}$ is complete with respect to $\mathcal{T}|_{\mathcal{L}}$.*

Proof. If Π is a Pontryagin space, choose a fundamental decomposition $\mathfrak{J} = (\Pi_+, \Pi_-)$, and set $\mathcal{L} := \Pi_+$. Then, clearly, $\mathcal{L}$ is a positive subspace of Π and $\dim \Pi/\mathcal{L} < \infty$. Since $\|x\|_{\mathfrak{J}} = [x, x]^{\frac{1}{2}}, x \in \Pi_+$, the subspace $\mathcal{L}$ is moreover complete with respect to $\|\cdot\|_{\mathfrak{J}}$.

On the other hand, assume that a subspace $\mathcal{L}$ with the stated properties exists. First of all, for every negative subspace $\mathcal{N} \leq \Pi$, it holds $\mathcal{N} \cap \mathcal{L} = \{0\}$, and hence $\dim \mathcal{N} \leq \dim \Pi/\mathcal{L}$. Let us denote by κ the dimension of a maximal negative subspace of Π. Thus

$$\kappa \leq \dim \Pi/\mathcal{L} < \infty$$

In particular, Π is decomposable and the topology $\mathcal{T}$ is well-defined. Choose a fundamental decomposition $\mathfrak{J} = (\Pi_+, \Pi_-)$ of Π, and consider the positive definite inner product $(.,.)_{\mathfrak{J}}$. Since the subspace $\mathcal{L}$ is complete with respect to $(.,.)_{\mathfrak{J}}$, we obtain

$$\mathcal{L} \oplus_{\mathfrak{J}} \mathcal{L}^{(\perp)_{\mathfrak{L}}} = \Pi$$

However, since

$$\dim \mathcal{L}^{(\perp)_{\mathfrak{J}}} = \dim \Pi/\mathcal{L} < \infty$$

also the space $\mathcal{L}^{(\perp)_{\mathfrak{J}}}$ is complete with respect to $(\cdot,\cdot)_{\mathfrak{J}}$. Thus Π is $(\cdot,\cdot)_{\mathfrak{J}}$ -complete, and we conclude that Π is a Pontryagin space.

□

5.3. Examples for Indefinite Spaces

The following examples are directly taken from [Wor07, Chapter A].

5.7. Example. (Pseudo-Euclidean space) A pseudo-Euclidean space $\mathcal{E} = \mathbb{R}^{(p,q)}$ is a real vector space equipped with a nondegenerate, indefinite inner product $\langle\cdot,\cdot\rangle_{\mathcal{E}}$. $\mathcal{E}$ admits a direct orthogonal decomposition $\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-$, where $\mathcal{E}_+ = \mathbb{R}^p$ and $\mathcal{E}_- = \mathbb{R}^q$ and the inner product is positive definite on $\mathcal{E}_+$ and negative definite on $\mathcal{E}_-$. The space $\mathcal{E}$ is, therefore, characterized by the signature (p, q).

Let $\mathcal{E} = \mathbb{R}^{(p,q)}$ be a pseudo-Euclidean space. Then $\langle\cdot,\cdot\rangle_{\mathcal{E}}$ can be expressed by the traditional $\langle\cdot,\cdot\rangle$ in a Euclidean space $\mathbb{R}^{p+q}$ as

$$\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{E}} = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{p+q} x_i y_i = \mathbf{x}^\top \mathcal{J}_{pq}\mathbf{y} = \langle \mathbf{x}, \mathcal{J}_{pq}\mathbf{y}\rangle$$

where

$$\mathcal{J}_{pq} = \begin{bmatrix} I_{p\times p} & 0 \\ 0 & -I_{q\times q} \end{bmatrix}$$

and $I_{p\times p}$ and $I_{q\times q}$ are the identity matrices. If x_+ and x_- stand for the orthogonal projections of x onto $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively, then $\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{E}} = \langle \mathbf{x}_+, \mathbf{y}_+\rangle - \langle \mathbf{x}_-, \mathbf{y}_-\rangle$ .The indefinite 'norm' of a non-zero vector x becomes $\|\mathbf{x}\|^2_{\mathcal{E}} = \langle \mathbf{x}, \mathbf{x}\rangle_{\mathcal{E}} = \mathbf{x}^\top \mathcal{J}_{pq}\mathbf{x}$, which can have any sign. Based on the inner product, the pseudo-Euclidean distance is defined analogously to the Euclidean case.

$$d^2_{\mathcal{E}}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2_{\mathcal{E}} = \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle_{\mathcal{E}} = \langle \mathbf{x}-\mathbf{y}\rangle^\top \mathcal{J}_{pq}\langle \mathbf{x}-\mathbf{y}\rangle$$

5.8. Example. Consider the space $\mathcal{L} := \ell^2(\mathbb{N}_0)$ endowed with the inner product

$$\left[(\xi_j)_{j\in\mathbb{N}_0}, (\eta_j)_{j\in\mathbb{N}_0}\right] := -\xi_0\overline{\eta_0} + \sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j\overline{\eta_j}$$

This inner product is obviously not degenerated and indefinite. Denote by $\mathcal{M}$ the following subspace

$$\mathcal{M} := \left\{ (\xi_j)_{j\in\mathbb{N}_0} : \xi_0 = \sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j \right\}$$

For $(\xi_j)_{j\in\mathbb{N}_0} \in \mathcal{M}$ it holds

$$\left[(\xi_j)_{j\in\mathbb{N}_0}, (\xi_j)_{j\in\mathbb{N}_0}\right] = -\sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j \cdot \sum_{j=1}^{\infty} \frac{1}{2^j}\overline{\xi_j} + \sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j\overline{\xi_j}$$

With respect to the Cauchy-Schwarz inequality for the weighted ℓ^2-space $\ell^2\left(\left(\frac{1}{2^n}\right)_{j\in\mathbb{N}}\right)$ we get

$$\left|\sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j\right|^2 = \left|\left((\xi_j)_{j\in\mathbb{N}}, (1)_{j\in\mathbb{N}}\right)\right|^2 \leq$$

$$\leq \left((\xi_j)_{j\in\mathbb{N}}, (\xi_j)_{j\in\mathbb{N}}\right) \cdot \left((1)_{j\in\mathbb{N}}, (1)_{j\in\mathbb{N}}\right) = \sum_{j=1}^{\infty} \frac{1}{2^j}\xi_j\overline{\xi_j} \cdot 1$$

where equality holds if and only if $(\xi_j)_{j\in\mathbb{N}}$ is a constant sequence. The one and only constant sequence in $\ell^2(\mathbb{N})$ is the zero sequence, thus we obtain

$$\left[(\xi_j)_{j\in\mathbb{N}_0}, (\xi_j)_{j\in\mathbb{N}_0}\right] > 0, (\xi_j)_{j\in\mathbb{N}_0} \in \mathcal{M}\backslash\{0\}$$

That means $\mathcal{M}$ is positive definite. Furthermore, it obviously holds $\dim \mathcal{L}/\mathcal{M} = 1$, and because $\mathcal{L}$ is indefinite itself, $\mathcal{M}$ has to be maximal non negative.

Now consider $(\eta_j)_{j\in\mathbb{N}_0} \in \mathcal{M}^\perp$. The sequence $x^{(n)} := (1, 0, \ldots, 0, 2^n, 0, \ldots)$ where the entry "2^n" lies at the n-th place, belongs to $\mathcal{M}$. Thus we get

$$0 = \left[(\eta_j)_{j\in\mathbb{N}_0}, x^{(n)}\right] = -\eta_0 + \eta_n$$

We see that $(\eta_j)_{j\in\mathbb{N}_0}$ is constant, and thus equal to 0. This means $\mathcal{M}^\perp = \{0\}$. Especially $\mathcal{M}^\perp$ is not maximal non-positive.

5.4. Kernel Reproducing Pontryagin Spaces

This proportion consists definitions and results from [BzCS12] as well as proofs from [Wor14, Chapter 4].

5.9. Definition. Let Π be a Pontryagin space of functions $f : \Omega \to \mathbb{C}$, where Ω is some non-empty set. A kernel K mapping $\Omega \times \Omega$ into the space $\mathbb{C}$ is called reproducing kernel of Π if the following two properties are satisfied.

(i) $K(\cdot, x)$ belongs to Π for every $x \in \Omega$,

(ii) $(f, K(\cdot, x)) = f(x)$ for every $x \in \Omega$ and $f \in \Pi$.

If such a kernel exists, Π is called reproducing kernel Pontryagin space.

5.10. Remark. Similar to the Hilbert space case the reproducing property is tied up with the fact that all evaluation mappings

$$\iota_x : \begin{cases} \Pi & \to & \mathbb{C} \\ f & \mapsto & f(x) \end{cases}$$

are continuous for every $x \in \Omega$. In fact, there exists an analogous result to the Riesz representation theorem.

Now we want to continue with the main construction of this chapter.

5.11. Theorem. *Let $K : \Omega \times \Omega \to \mathbb{C}$ be a conditionally positive definite kernel of order κ. Then there exists a unique Pontryagin space of index γ, $\gamma \leq \kappa$, having K as its reproducing kernel.*

Proof. Let us consider our candidate for our Pontryagin space $\mathcal{T} = \operatorname{span}\{K_x : x \in \Omega\}$, where $K_x = K(\cdot, x)$. For any two elements $f, g \in \mathcal{T}$ we can find points $x_1, \dots, x_n \in X$ and complex numbers $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{C}$ such that

$$f = \sum_{i=1}^{n} \alpha_i K_{x_i} \text{ and } g = \sum_{i=1}^{n} \beta_i K_{x_i}$$

Since it holds

$$\sum_{i=1}^{n} \overline{\beta_i} f(x_i) = \sum_{i,j=1}^{n} \alpha_j \overline{\beta_i} K(x_i, x_j) = \sum_{j=1}^{n} \alpha_j g(x_j)$$

we can define a hermitian indefinite sesquilinearform

$$(f, g) = \sum_{i,j=1}^{n} \alpha_j \overline{\beta_i} K(x_i, x_j).$$

It can easily be checked that our sesquilinearform satisfies the reproducing kernel property for every $f \in \mathcal{T}$ and $x \in \Omega$.

$$(f, K_x) = (\sum_{i=1}^{n} \alpha_i K_{x_i}, K_x) = \sum_{i=1}^{n} \alpha_i K_{x_i}(x) = f(x)$$

Note that this property also guarantees non-degeneracy. Assume there exists $f_0 \in \mathcal{T}$ with $(f_0, f) = 0$ for all $f \in \mathcal{T}$. In particular this holds for $f = K_x$ implying

$$0 = (f_0, K_x) = f_0(x)$$

for all $x \in \Omega$. Thus f_0 must be equal to 0.

It remains to show that any maximal negative subspace of $\mathcal{T}$ is of dimension $\gamma, \gamma \leq \kappa$. Therefore, choose $n \geq \kappa$ points $x_1, \dots, x_n \in \Omega$ and consider the Gram matrix

$$\mathrm{G} = \left[\left(K_{x_j}, K_{x_i}\right)\right]_{i,j=1}^{n}.$$

From the defining equations for conditional positive definiteness we obtain that

$$\mathbf{a}^* \mathrm{G} \mathrm{a} \geq 0$$

for all $\mathbf{a} \in \mathbb{C}^n$ satisfying

$$\sum_{i=1}^{n} \overline{a_i} u(x_i) = 0 \quad \forall u \in U$$

or, equivalently,

$$\mathbf{a} \in \{[u(x_i)]_{i=1}^n : u \in U\}^{\perp} \subseteq \mathbb{C}^n.$$

Since $\dim U = \kappa$, G can have at most κ negative eigenvalues. Any negative subspace in $\operatorname{span}\{K_{x_1}, \dots, K_{x_n}\}$ is of dimension less than the number of negative eigenvalues of the Gram matrix. Let γ be the maximal number of negative eigenvalues occurring for arbitrary choices of finitely many points. A negative subspace cannot be infinite dimensional since the existence of a negative subspace of dimension greater than κ immediately results in a contradiction.

Now we still need to complete our space. Consider $\mathfrak{J} = (\mathcal{L}_+, \mathcal{L}_-)$, a fundamental decomposition of $\mathcal{T}$, and let us call (ι_+, Π_+) and $(\iota_- \Pi_-)$ the Hilbert space, respectively anti Hilbert space completion from $\mathcal{L}_+$ and $\mathcal{L}_-$. Ultimately, we obtain a space $(\iota_+ + \iota_-, \Pi_+ \oplus \Pi_-)$ which is a completion of $\mathcal{T}$ in a way that it is a Pontryagin space and it is complete with respect to $\|\cdot\|_{\mathfrak{J}}$. We will show that K fulfills the reproducing kernel property. Consider an arbitrary $f \in \Pi$ and the corresponding fundamental decomposition $f = f^+ + f^-$, where $f^+ \in \Pi_+$ and $f^- \in \Pi_-$. Consider the sequences $(f_n^+)_{n\in\mathbb{N}} \in \iota_+(\mathcal{L}_+)^{\mathbb{N}}$ and $(f_n^-)_{n\in\mathbb{N}} \in \iota_-(\mathcal{L}_-)^{\mathbb{N}}$ converging to f^+ and f^-. Since $\iota_+(\mathcal{L}_+) \oplus \iota_-(\mathcal{L}_-)$ lies dense in $\Pi_+ \oplus \Pi_-$, we can apply Proposition 5.4.We obtain for arbitrary $x \in \Omega$

$$\begin{aligned} f(x) = f^+(x) + f^-(x) &= \lim_{n\to\infty} f_n^+(x) + f_n^-(x) \\ &= \lim_{n\to\infty} (f_n^+, K(\cdot, x)) + (f_n^-, K(\cdot, x)) = (f, K(\cdot, x)). \end{aligned}$$

Now we want to show uniqueness. Let Π_1 and Π_2 be two reproducing kernel Pontryagin spaces which contain $\mathcal{T}$ isometrically and densely. The completions $\langle \subseteq, \Pi_1 \rangle$ and $\langle \subseteq, \Pi_2 \rangle$ are isomorphic. Let ω be a linear and isometric homeomorphism $\omega : \Pi_1 \to \Pi_2$ with

$$\begin{tikzcd} & \mathcal{T} \arrow[dl, hook', "\subseteq"'] \arrow[dr, hook, "\subseteq"] & \\ \Pi_1 \arrow[rr, dotted, "\omega"'] & & \Pi_2 \end{tikzcd}$$

Then $\omega(f) = f, f \in \mathcal{T}$, and hence

$$\left[\iota_x|_{\Pi_2} \circ \omega \right]\big|_{\mathcal{T}} = \left[\iota_x|_{\Pi_1} \right]\big|_{\mathcal{T}}, \quad x \in \Omega.$$

Since point evaluations are continuous in both spaces Π_1 and Π_2, it follows that $\iota_x|_{\Pi_2} \circ \omega = \iota_x|_{\Pi_1}$. Hence, $\omega(f) = f$ for all $f \in \Pi_1$. This just says that Π_1 and Π_2 are equal (since ω is isometric and homeomorphic, they are equal as Pontryagin spaces).

$\square$

A. Appendix

A.1. Further Indefinite Inner Product Spaces

The following is put together from [Wor14, Chapter 2].

1.1. Definition. A space $\langle \mathcal{K}, [.,.] \rangle$ with an indefinite inner product is called Krein space if $[.,.]$ is non-degenerate and it has a fundamental decomposition $\mathfrak{J}$ as in (5.2) such that $\mathcal{K}$ is complete with respect to $\| \cdot \|_{\mathfrak{J}}$.

1.2. Remark. Note that the dimension of the negative subspaces of a Krein space does not have to be necessarily finite. In particular, the class of the Pontryagin spaces is a subclass of the Krein spaces.

1.3. Definition. We call a triple $\langle \mathcal{A}, [\cdot,\cdot]_{\mathcal{A}}, \mathcal{O} \rangle$ an almost Pontryagin space, if $\mathcal{A}$ is a linear space, $[\cdot,\cdot]_{\mathcal{A}}$ is an indefinite inner product on $\mathcal{A}$, and $\mathcal{O}$ is a topology on $\mathcal{A}$, such that the space $\mathcal{A}$ can be decomposed as the direct and orthogonal sum

$$\mathcal{A} = \mathcal{A}_+ \oplus \mathcal{A}_- \oplus \mathcal{A}^\circ,$$

with a finite dimensional negative subspace $\mathcal{A}_-$ and an $\mathcal{O}$-closed subspace $\mathcal{A}_+$ such that $\left\langle \mathcal{A}_+, [\cdot,\cdot]_{\mathcal{A}}|_{\mathcal{A}_+ \times \mathcal{A}_+} \right\rangle$ is a Hilbert space and $\mathcal{A}^\circ = \{x \in \mathcal{A} : [x,y]_{\mathcal{L}} = 0, y \in \mathcal{A}\}$ denoting the isotropic part with $\dim \mathcal{A}^\circ < \infty$.

1.4. Definition. Let Ω be a nonempty set and let $\mathcal{K}$ be a Krein space. We call an almost Pontryagin space $\langle \mathcal{A}, [\cdot,\cdot], \mathcal{A}, \mathcal{O} \rangle$ a reproducing kernel almost Pontryagin space of $\mathcal{K}$-valued functions on Ω, if

- The elements of $\mathcal{A}$ are $\mathcal{K}$-valued functions on Ω, and the linear operations of $\mathcal{A}$ are given by pointwise addition and scalar multiplication.
- For each $x \in \Omega$ the point evaluation functional $\iota_x|_{\mathcal{A}} : \mathcal{A} \to \mathcal{K}$ is continuous with respect to the topology $\mathcal{O}$ on $\mathcal{A}$ and the Krein space topology on $\mathcal{K}$, where

$$\iota_{x,a} : \begin{cases} \mathcal{K}^\Omega \to \mathbb{C} \\ \quad f \mapsto (f(x), a)_{\mathcal{K}} \end{cases}, a \in \mathcal{A}.$$

A.2. An Application In Machine Learning

There is a vast range of applications linked to the theory of reproducing kernel spaces and kernel functions. In computer science these so called kernel methods enjoy a great reputation in the field of machine learning. A small fraction of what this theory is capable of is going to be highlighted in this section, which is extracted from [HSS08, Chapter 2].

1.5. Definition. For a given RKHS H consider the function $\Phi : \Omega \to H$, $\Phi(x) = K_x$, where K denotes the reproducing kernel function of H. We call this function the feature map and H the feature space.

1.6. Example. Suppose we are given some empirical data

$$(x_1, y_1), \dots, (x_n, y_n) \in \Omega \times \{\pm 1\}$$

and a feature map Φ. Now let us interpret $K(x, x') = (\Phi(x), \Phi(x'))$ as a similarity measure between our inputs x and x'. The advantage of using such a kernel as a similarity measure is that it allows us to construct algorithms in inner product spaces. The idea is to compute the means of the two classes in the feature space, $c_+ = \frac{1}{n_+}\sum_{\{i:y_i=+1\}} \Phi(x_i)$, and $c_- = \frac{1}{n_-}\sum_{\{i:y_i=-1\}} \Phi(x_i)$, where $n_\pm = |\{i : y_i = \pm 1\}|$. We then assign a new point $\Phi(x)$ to the class whose mean is closer to it. This leads to the prediction rule

$$\begin{aligned} y &= \operatorname{sgn}\left(\langle \Phi(x), c_+\rangle - \langle \Phi(x), c_-\rangle + \frac{1}{2}\left(\|c_-\|^2 - \|c_+\|^2\right)\right) \\ &= \operatorname{sgn}\left(\frac{1}{n_+}\sum_{\{i:y_i=+1\}} \underbrace{\langle \Phi(x), \Phi(x_i)\rangle}_{K(x,x_i)} - \frac{1}{n_-}\sum_{\{i:y_i=-1\}} \underbrace{\langle \Phi(x), \Phi(x_i)\rangle}_{K(x,x_i)} + \frac{1}{2}\left(\|c_-\|^2 - \|c_+\|^2\right)\right) \end{aligned}$$

The benefit of the kernel trick is that the objective function we are optimizing to fit the higher dimensional decision boundary only includes the dot product of the transformed feature vectors. Therefore, we can just substitute these dot product terms with the kernel function, and we do not even use $\Phi(x)$. For example a feature map could look like this $\Phi((a_1, a_2)^T) = (a_1^2, \sqrt{2}a_1a_2, a_2^2)^T$, which leads to

$$\begin{aligned} \Phi(\mathbf{a})^T \cdot \Phi(\mathbf{b}) = \begin{pmatrix} {a_1}^2 \\ \sqrt{2}a_1a_2 \\ {a_2}^2 \end{pmatrix}^T \cdot \begin{pmatrix} b_1^2 \\ \sqrt{2}b_1b_2 \\ {b_2}^2 \end{pmatrix} &= {a_1}^2{b_1}^2 + 2a_1b_1a_2b_2 + {a_2}^2{b_2}^2 \\ = (a_1b_1 + a_2b_2)^2 = \left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^T \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right)^2 &= \left(\mathbf{a}^T \cdot \mathbf{b}\right)^2 \end{aligned}$$

On the left-hand side, we have the dot product of the transformed feature vectors, which is equal to our second degree polynomial kernel function $K(a, b) = (a^Tb)^2$. Geometrically, we are finding the optimal separating hyperplane in this higher dimensional space without knowing anything about $\Phi(x)$, leading to much faster computations for algorithms.

Bibliography

[AB09] Yves Auffray and Pierre Barbillon. *Conditionally positive definite kernels: theoretical contribution, application to interpolation and approximation.* PhD thesis, INRIA, 2009.

[BKW20] Martin Blümlinger, Michael Kaltenbäck, and Harald Woracek. Funktionalanalysis. 2020.

[BzCS12] Georg Berschneider, Wolfgang zu Castell, and Stefan Schrödl. Function spaces for conditionally positive definite operator-valued kernels. *Mathematics of Computation*, 81(279):1551–1569, 2012.

[HSS08] Thomas Hofmann, Bernhard Schölkopf, and Alexander J Smola. Kernel methods in machine learning. *The annals of statistics*, 36(3):1171–1220, 2008.

[Kal19] Michael Kaltenbäck. *Funktionalanalysis 2.* 2019.

[KW14] Michael Kaltenbäck and Harald Woracek. *Geometrie im Pontryagin Raum.* 2014.

[Wen04] Holger Wendland. *Scattered data approximation*, volume 17. Cambridge university press, 2004.

[Wor07] Harald Woracek. *Operatoren im Krein Raum.* 2007.

[Wor10] Harald Woracek. *Operatortheorie im Krein Raum 1+2.* 2010.

[Wor14] Harald Woracek. Reproducing kernel almost pontryagin spaces. *Linear Algebra and its Applications*, 461:271–317, 2014.