# REPRODUCING KERNEL BANACH SPACES

PANDO G. GEORGIEV†, LUIS SÁNCHEZ-GONZÁLEZ‡, PANOS M. PARDALOS†

† Center for Applied Optimization, University of Florida, Gainesville, USA
‡ Department of Mathematical Analysis, Universidad Complutense de Madrid, Spain

Abstract. We extend the idea of Reproducing Kernel Hilbert Spaces to Banach spaces, developing a theory of pairs of Reproducing Kernel Banach Spaces (RKBS) without the requirement of existence of semi-inner product (which requirement is already explored in another construction of RKBS). We present several natural examples, which involve RKBS of functions with supremum norm and with $\ell_p$-norm ($1 \leq p \leq \infty$). Special attention is devoted to the case of a pair of RKBS $(B, B^\sharp)$ in which $B$ has sup-norm and $B^\sharp$ has $\ell_1$-norm. Namely, we show that if $(B, B^\sharp)$ is generated by an universal kernel and $B$ is furnished with the sup-norm, then $B^\sharp$, furnished with the $\ell_1$-norm is linearly isomorphically embedded in the dual of $B$. We reformulate the classical classification problem (Support Vector Machine classifiers) to RKBS and suggest that it will have sparse solutions when the RKBS is furnished with the $\ell_1$-norm.

## 1. Introduction

We will consider Banach spaces over either the field of real numbers, $\mathbb{R}$, or of complex numbers, $\mathbb{C}$. We will use $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$.

**Definition 1.1.** *Let $X$ and $Y$ be two sets, let $K : X \times Y \to \mathbb{K}$ be a function, let $V$ and $V^\sharp$ two Banach spaces composed by functions defined on $Y$ and $X$, respectively. We shall say that the pair $(V, V^\sharp)$ is a pair of* ***reproducing kernel Banach spaces (RKBS) with the reproducing kernel*** *$K$ provided that*

1. *for all $f \in V$, $||f||_V = 0$ if and only if $f(y) = 0$ for all $y \in Y$,*
2. *for all $g \in V^\sharp$, $||g||_{V^\sharp} = 0$ if and only if $g(x) = 0$ for all $x \in X$,*
3. *the point evaluation functionals are continuous on $V$ and $V^\sharp$, i.e. for every $x \in X$ and $y \in Y$ the functionals $\delta_y : V \to \mathbb{K}$ and $\delta_x : V^\sharp \to \mathbb{K}$ defined as $\delta_y(f) = f(y)$ and $\delta_x(g) = g(x)$ for all $f \in V$ and $g \in V^\sharp$ are continuous,*
4. *$K(x, \cdot) \in V$ and $K(\cdot, y) \in V^\sharp$ for every $x \in X$ and $y \in Y$, and*
5. *there exists a bilinear form $<\cdot,\cdot>_K$ in $V \times V^\sharp$ such that*

$$<f, K(\cdot, y)>_K = f(y) \qquad \text{for all } y \in Y \text{ and } f \in V, \tag{1.1}$$

*and*

$$<K(x, \cdot), g>_K = g(x) \qquad \text{for all } x \in X \text{ and } g \in V^\sharp. \tag{1.2}$$

Let us note that the first difference with RKHS is that we are not making assumptions about the norm, so the RKBS could be non-unique.

Let $(V, V^\sharp)$ be a pair of RKBS of real-valued functions on the sets $X$ and $Y$ with reproducing kernel $K$. Let $W = \{f_1 + if_2 : f_1, f_2 \in V\}$ and $W^\sharp = \{f_1 + if_2 : f_1, f_2 \in V^\sharp\}$, which are vector spaces of complex-valued functions on $X$ and $Y$. If we set the bilinear form,

$$<f_1 + if_2, g_1 + ig_2>_W = (<f_1, g_1>_K - <f_2, g_2>_K) + i(<f_1, g_2>_K + <f_2, g_1>_K),$$

and the norms

$$||f_1 + if_2||_W = ||f_1||_V + ||f_2||_V \qquad \text{and} \qquad ||g_1 + ig_2||_{W^\sharp} = ||g_1||_{V^\sharp} + ||g_2||_{V^\sharp},$$

we have that the pair $(W, W^\sharp)$ is a pair of RKBS of complex-valued functions on $X$ and $Y$ with reproducing kernel $K$ (which is real-valued). We call $W$ and $W^\sharp$ the complexifications of $V$ and $V^\sharp$. Since every real-valued RKBS can be complexified preserving the reproducing kernel, we shall consider only complex-valued reproducing kernel Banach spaces.

In Section 2, we construct a wide class of RKBS under two general assumptions. In this process, we follow and improve the technics developed in [2], removing some assumptions. In particular, we obtain the following theorem:

**Theorem 1.2.** *Let $X$ and $Y$ be two sets and $K : X \times Y \to \mathbb{C}$ a function. Let us denote $B_0 = \operatorname{span}\{K(x,\cdot) : x \in X\}$ and $B_0^\sharp = \operatorname{span}\{K(\cdot, y) : y \in Y\}$. Let us suppose that there is a norm $||\cdot||_{B_0}$ in $B_0$ satisfying:*

**(HN1)** *The evaluation functionals are continuous of $B_0$.*

**(HN2)** *If $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $(B_0, ||\cdot||_{B_0})$ such that $f_n(y) \to 0$ for all $y \in Y$, then $||f_n||_{B_0} \to 0$.*

*Then, there are $(B, ||\cdot||_B)$ and $(B^\sharp, ||\cdot||_{B^\sharp})$ – Banach completions of $B_0$ and $B_0^\sharp$, respectively, such that $(B, B^\sharp)$ is a pair of RKBS with the reproducing kernel $K$. Furthermore, if we denote by $<\cdot,\cdot>_K$ the bilinear form in $B \times B^\sharp$, we state*

$$|<f, g>_K| \le ||f||_B ||g||_{B^\sharp} \qquad \textit{for all } f \in B \textit{ and } g \in B^\sharp, \textit{and}$$

$$||g||_{B^\sharp} = \sup_{\substack{f \in B \\ ||f||_B \le 1}} |<f, g>_K| \quad \textit{for any } f \in B \textit{ and } g \in B^\sharp.$$

In Section 3, we give some examples of RKBS. In particular, we obtain RKBS when the kernel is either a bounded function, continuous and bounded function or it is in $C_0(X)$. Moreover, we obtain a simple condition in the kernel in order to obtain RKBS with $\ell_1$-norm. Lastly, in this section we obtain a class of RKBS with $\ell_p$-norm following the construction given in [2] for RKBS with the $\ell_1$-norm.

In section 4, we sketch the main tasks in Statistical Learning Theory and reformulate Support Vector Machine classification problem for the case of RKBS with $\ell_1$-norm.

## 2. Constructing Reproducing Kernel Banach Spaces

Let $X$ and $Y$ be sets and $K : X \times Y \to \mathbb{C}$ a function. Following the ideas given in [2] and the proof of the uniqueness of RKHS in [1], we construct a pair of Banach spaces which is a RKBS with the reproducing kernel $K$.

Let us introduce the vector spaces

$$B_0 := \operatorname{span}\{K(x,\cdot) : x \in X\} \subset \mathbb{C}^Y \qquad \text{and} \qquad B_0^\sharp := \operatorname{span}\{K(\cdot, y) : y \in Y\} \subset \mathbb{C}^X.$$

Let us define the bilinear form $<\cdot,\cdot>_K$ on $B_0 \times B_0^\sharp$ as

$$<f, g>_K = <\sum_{j=1}^n \alpha_j K(x_j,\cdot), \sum_{k=1}^m \beta_k K(\cdot, y_k)>_K := \sum_{j,k} \alpha_j \beta_k K(x_j, y_k), \tag{2.1}$$

for every $f = \sum_{j=1}^n \alpha_j K(x_j,\cdot) \in B_0$ and $g = \sum_{k=1}^m \beta_k K(\cdot, y_k) \in B_0^\sharp$.

**Properties 2.1.** *It is straightforward to verify that*

(i) $<\cdot,\cdot>_K$ *is a bilinear form on* $B_0 \times B_0^\sharp$.
(ii) $K(x,y) = <K(x,\cdot), K(\cdot,y)>_K$ *for all* $x \in X$ *and* $y \in Y$.
(iii) $<f, K(\cdot,y)>_K = f(y)$ *and* $<K(x,\cdot), g>_K = g(x)$ *for all* $f \in B_0$, $g \in B_0^\sharp$, $x \in X$ *and* $y \in Y$.

Let us suppose (HN1), i.e. there is a norm $||\cdot||_{B_0}$ on $B_0$ such that the evaluation functionals are continuous on $B_0$. We follow [1] and [2] to obtain a Banach completion of $B_0$. Let $\{f_n\}_{n=1}^\infty$ be a Cauchy sequence in $B_0$, since the point evaluation functionals are continuous, the sequence $\{f_n(y)\}_{n=1}^\infty$ is Cauchy for every $y \in Y$. So, we can define $f(y) := \lim_{n\to\infty} f_n(y)$ for all $y \in Y$, the space

$$B := \{f : Y \to \mathbb{C} : \textit{ there exists } \{f_n\}_{n=1}^\infty \textit{ a Cauchy sequence in } B_0 \text{ such that } f_n(y) \to f(y) \text{ for all } y \in Y\},$$

and the function $||\cdot||_B$ on $B$ as

$$||f||_B := \lim_{n\to\infty} ||f_n||_{B_0},$$

for all $f \in B$, where $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0$ such that $f_n(y) \to f(y)$ on $Y$.

**Proposition 2.2.** *If* $(B_0, ||\cdot||)$ *satisfies the hypothesis (HN2), then* $||\cdot||_B$ *is well-defined, and the pair* $(B, ||\cdot||_B)$ *is a Banach space.*

*Proof.* The function $||\cdot||_B$ is well-defined by (HN2). It is easy to see that $||\cdot||_B$ is a norm in $B$, only note that $f_n(y) \to 0$ for all $y \in Y$ whenever $||f_n||_{B_0} \to 0$ since the point evaluation functionals are continuous (hypothesis (HN1)). Thus, $||f||_B = 0$ if and only if $f = 0$.

Finally, we shall show that $(B, ||\cdot||_B)$ is a Banach space. Given $\{f_n\}_{n=1}^\infty$ a Cauchy sequence in $B$, if there exists $n_0 \in \mathbb{N}$ such that $f_n \in B_0$ for all $n \geq n_0$ then, by definition of $B$, there exists $f \in B$ such that $f_n \to f$ in $B$ (i.e., $||f_n - f||_B \to 0$). Let us suppose that there is infinitely many functions $f_n \notin B_0$, we can choose $g_n \in B_0$ such that $||g_n - f_n||_B < \frac{1}{n}$ for all $n \in \mathbb{N}$. Then, $\{g_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0$. Therefore, there exists $g \in B$ such that $g_n \to g$ in $B$, and consequently, $f_n \to g$.

Let us notice that (HN2) is also a necessary hypothesis. Indeed, if $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0$ such that $f_n(y) \to 0$ for all $y \in Y$, and $(B, ||\cdot||_B)$ is well-defined. Then

$$0 = ||0||_B = \lim_{n\to\infty} ||f_n||_{B_0}.$$

□

**Properties 2.3.**

(i) *Every function* $f \in B$ *such that* $||f||_B = 0$ *satisfies* $f(y) = 0$ *for all* $y \in Y$.
(ii) *The point evaluation functionals are continuous on* $B$.

*Proof.* Since (i) is straightforward, we only show the property (ii). Given $f \in B$, there is a Cauchy sequence $\{f_n\}_{n=1}^\infty$ in $B_0$ such that $f_n(y) \to f(y)$ for all $y \in Y$ and $||f||_B = \lim_{n\to\infty} ||f_n||_{B_0}$. Then

$$|\delta_y(f)| = |f(y)| = |\lim_{n\to\infty} f_n(y)| = |\lim_{n\to\infty} \delta_y(f_n)| \leq \lim_{n\to\infty} |\delta_y(f_n)| \leq ||\delta_y|| \lim_{n\to\infty} ||f_n||_{B_0} = ||\delta_y|| ||f||_B.$$

□

Now, let us define the following function on $B_0^\sharp$.

$$||g||_{B_0^\sharp} := \sup_{\substack{f \in B_0 \\ ||f||_{B_0} \leq 1}} |<f, g>_K| \qquad \text{for all } g \in B_0^\sharp.$$

**Proposition 2.4.** *$|| \cdot ||_{B_0^\sharp}$ is a norm in $B_0^\sharp$.*

*Proof.* Let us suppose that there is $g \in B_0^\sharp$ with $||g||_{B_0^\sharp} = 0$. Then, for every $x \in X$ we have

(i) if $||K(x, \cdot)||_{B_0} \leq 1$, then $|g(x)| = |<K(x, \cdot), g>_K| \leq ||g||_{B_0^\sharp} = 0$,

(ii) if $||K(x, \cdot)||_{B_0} > 1$, then $f := \frac{1}{||K(x,\cdot)||_{B_0}} K(x, \cdot) \in B_0$ and $||f||_{B_0} = 1$, so $\frac{1}{||K(x,\cdot)||_{B_0}} |g(x)| = |<f, g>_K| \leq ||g||_{B_0^\sharp} = 0$, and $g(x) = 0$.

It is straightforward to prove the other properties of the norm. □

**Proposition 2.5.** *For every $f \in B_0$ and $g \in B_0^\sharp$*

$$|<f, g>_K| \leq ||f||_{B_0} ||g||_{B_0^\sharp}. \tag{2.2}$$

*Proof.* If $||f||_{B_0} = 0$ then $f = 0$ and the inequality is obvious. If $||f||_{B_0} \neq 0$, we define $h = \frac{1}{||f||_{B_0}} f \in B_0$ and $||h||_{B_0} = 1$, then

$$|<f, g>_K| = ||f||_{B_0} \, |<h, g>_K| \leq ||f||_{B_0} ||g||_{B_0^\sharp}.$$

□

**Corollary 2.6.** *The point evaluation functionals are continuous on $B_0^\sharp$ with the norm $|| \cdot ||_{B_0^\sharp}$.*

*Proof.*

$$|\delta_x(g)| = |g(x)| = |<K(x, \cdot), g>_K| \leq ||K(x, \cdot)||_{B_0} ||g||_{B_0^\sharp}.$$

□

In the same way as before, we can define a Banach completion of $B_0^\sharp$ which yields a space of functions, given by

$$B^\sharp := \{g : X \to \mathbb{C} : \text{there exists a Cauchy sequence } \{g_n\}_{n=1}^\infty \text{ in } B_0^\sharp \text{ such that } g_n(x) \to g(x) \text{ for all } x \in X\}.$$

**Proposition 2.7.** *Given $g \in B^\sharp$ and $\{g_n\}_{n=1}^\infty$ – any Cauchy sequence in $B_0^\sharp$ such that $g_n(x) \to g(x)$ on $X$, the function*

$$||g||_{B^\sharp} := \lim_{n \to \infty} ||g_n||_{B_0^\sharp}$$

*is well-defined and $(B^\sharp, || \cdot ||_{B^\sharp})$ is a Banach space.*

*Proof.* Let us show that $|| \cdot ||_{B^\sharp}$ is well-defined. First of all, notice that it is sufficient to prove that $||g_n||_{B_0^\sharp} \to 0$ whenever $\{g_n\}$ is a Cauchy sequence in $B_0^\sharp$ such that $g_n(x) \to 0$ for all $x \in X$. Fixed $\varepsilon > 0$ there is $N_0 \geq 1$ such that

$$||g_m - g_n||_{B_0^\sharp} = \sup_{\substack{f \in B_0 \\ ||f||_{B_0} \leq 1}} |<f, g_m - g_n>_K| < \varepsilon/2 \qquad \text{for all } n, m \geq N_0.$$

Since $g_n(x) \to 0$ for all $x \in X$, given $f = \sum_{j=1}^{p} \alpha_j K(x_j, \cdot) \in B_0$ with $||f||_{B_0} \le 1$, there is $m_0$ $(= m_0(\varepsilon, f))$ such that

$$|<f, g_m>_K| = \left| \sum_{j=1}^{p} \alpha_j g_m(x_j) \right| < \varepsilon/2 \qquad \text{for all } m \ge m_0.$$

Then, for $n \ge N_0$, let us take $m \ge \max\{N_0, m_0\}$, thus

$$|<f, g_n>_K| \le |<f, g_n - g_m>_K| + |<f, g_m>_K| < \varepsilon.$$

Hence,

$$||g_n||_{B_0^\sharp} = \sup_{\substack{f \in B_0 \\ ||f||_{B_0} \le 1}} |<f, g_n>_K| < \varepsilon \qquad \text{for all } n \ge N_0, \text{ and} \qquad ||g_n||_{B_0^\sharp} \to 0.$$

□

**Properties 2.8.**

(i) *Every function $g \in B^\sharp$ such that $||g||_{B^\sharp} = 0$ satisfies $g(x) = 0$ for all $x \in X$.*

(ii) *The point evaluation functionals are continuous on $B^\sharp$.*

*Proof.* The proofs are the same as Properties 2.3, notice that $|\delta_x(g)| \le ||K(x, \cdot)||_{B_0} ||g||_{B^\sharp}$ for every $x \in X$ and $g \in B^\sharp$. □

Let us extend the bilinear form $<\cdot, \cdot>_K$ from $B_0 \times B_0^\sharp$ to $B \times B^\sharp$. Let us note that there is no Hahn-Banach theorem for bilinear mappings. So, we have to use the "density" of $B_0$ and $B_0^\sharp$ in $B$ and $B^\sharp$, respectively.

- Let us define $<\cdot, \cdot>_K : B \times B_0^\sharp \to \mathbb{C}$ by

$$<f, g>_K := \lim_{n \to \infty} <f_n, g>_K,$$

where $f \in B$, $g \in B_0^\sharp$ and $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_0$ such that $f_n(y) \to f(y)$ for all $y \in Y$. Using inequality (2.2), $\{<f_n, g>_K\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{C}$, the limit exits and is well-defined. Clearly, the form $<\cdot, \cdot>_K$ is bilinear in $B \times B_0^\sharp$. Finally,

$$|<f, g>_K| = \left| \lim_{n \to \infty} <f_n, g>_K \right| \le \lim_{n \to \infty} ||f_n||_{B_0} ||g||_{B_0^\sharp} = ||f||_B ||g||_{B_0^\sharp} \tag{2.3}$$

for all $f \in B$ and $g \in B_0^\sharp$, where $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_0$ such that $f_n(y) \to f(y)$ for all $y \in Y$.

- Let us define $<\cdot, \cdot>_K : B \times B^\sharp \to \mathbb{C}$ by

$$<f, g>_K := \lim_{n \to \infty} <f, g_n>_K,$$

where $f \in B$, $g \in B^\sharp$ and $\{g_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_0^\sharp$ such that $g_n(x) \to g(x)$ for all $x \in X$. Using inequality (2.3) instead of inequality (2.2), we prove in a similar way that $<\cdot, \cdot>_K$ is a bilinear form on $B \times B^\sharp \to \mathbb{C}$ and

$$|<f, g>_K| \le ||f||_B ||g||_{B^\sharp} \tag{2.4}$$

for all $f \in B$ and $g \in B^\sharp$.

**Proposition 2.9.** *For all $g \in B^\sharp$*

$$||g||_{B^\sharp} = \sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g>_K| .$$

*Proof.* By inequality (2.4) it is clear that $\sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g>_K| \leq ||g||_{B^\sharp}$. Let us prove the reverse inequality. If $g \in B_0^\sharp$, then $||g||_{B^\sharp} = ||g||_{B_0^\sharp} = \sup_{\substack{f \in B_0 \\ ||f||_{B_0} \leq 1}} |<f, g>_K| \leq \sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g>_K|$. If $g \in B^\sharp$ and $\{g_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0^\sharp$ such that $g_n(x) \to g(x)$ on $X$, then $||g_n - g||_{B^\sharp} \to 0$, and

$$|<f, g_n>_K| \leq |<f, g_n - g>_K| + |<f, g>_K| \leq ||g_n - g||_{B^\sharp} + |<f, g>_K| ,$$

for all $f \in B$ with $||f||_B \leq 1$. Thus,

$$\begin{aligned} ||g||_{B^\sharp} &= \lim_{n\to\infty} ||g_n||_{B_0^\sharp} \leq \lim_{n\to\infty} \sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g_n>_K| \\ &\leq \lim_{n\to\infty} \left( ||g_n - g||_{B^\sharp} + \sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g>_K| \right) = \sup_{\substack{f \in B \\ ||f||_B \leq 1}} |<f, g>_K| . \end{aligned}$$

□

Let us denote by $<\cdot, \cdot> : B \times B^* \to \mathbb{C}$ the evaluation map $<x, y^*> = y^*(x)$ for all $x \in B$ and $y^* \in B^*$. Let us notice that the mapping $\mathcal{L}$ from the Banach space $B^\sharp$ to the dual space $B^*$ of $B$ defined as

$$(\mathcal{L}(g))(f) = <f, \mathcal{L}(g)> := <f, g>_K \qquad \text{for all } f \in B \text{ and } g \in B^\sharp, \tag{2.5}$$

is an embedding from $B^\sharp$ to $B^*$, i.e. it is an isometric and linear mapping.

So, we can define $\phi : X \to B$ and $\phi^* : Y \to B^*$ as

$$\phi(x) = K(x, \cdot) \in B \qquad \text{and} \qquad \phi^*(y) = \mathcal{L}(K(\cdot, y)) \in B^*,$$

and they satisfy

$$K(x, y) = <\phi(x), \phi^*(y)>.$$

Now, we have all ingredients to prove Theorem 1.2.

*Proof of Theorem 1.2.* It only remains to check the equalities (1.1) and (1.2). Let us prove equality (1.1) (equality (1.2) can be proved similarity). Given $f \in B$ and $y_0 \in Y$, there is a Cauchy sequence $\{f_n\}_{n=1}^\infty$ in $B_0$ such that $f_n(y) \to f(y)$ for all $y \in Y$. Thus,

$$f(y_0) = \lim_{n\to\infty} f_n(y_0) = \lim_{n\to\infty} <f_n, K(\cdot, y_0)>_K.$$

By inequality (2.4), the map $h \in B \to <h, K(\cdot, y_0)>_K$ is a bounded linear functional on $B$, then $f(y_0) = \lim_{n\to\infty} <f_n, K(\cdot, y_0)>_K = <f, K(\cdot, y_0)>_K$. □

## 3. Some examples of RKBS

3.1. **Pairs of RKBS defined on Banach spaces.** Let $X$ be a Banach space and $X^*$ be its dual space. For a given $p \in \mathbb{N}$, define a function $K : X \times X^* \to \mathbb{C}$ such that

$$K(x, y^*) = (<x, y^*>_X + 1)^p,$$

where $<\cdot,\cdot>_X : X \times X^* \to \mathbb{C}$ denotes the evaluation map $<x, y^*>_X = y^*(x)$ for all $x \in X$ and $y^* \in X^*$. We can define

$$B_0 := \operatorname{span}\{K(x,\cdot) : x \in X\} \qquad \textit{and} \qquad B_0^\sharp := \operatorname{span}\{K(\cdot, y) : y \in X^*\},$$

and the norm in $B_0$ given by

$$||f||_{B_0} := \sup_{y\in X^*} \frac{|f(y)|}{1 + \sum_{j=1}^{p} ||y||^j}.$$

**Properties 3.1.**
(1) *If* $f = \sum_{j=1}^{n} \alpha_j K(x_j, \cdot) \in B_0$ *then* $||f||_{B_0} \le \sum_{j=1}^{n} |\alpha_j|(1 + ||x_j||)^p < \infty$.
(2) $||\cdot||_{B_0}$ *is a norm on* $B_0$ *and satisfies the hypothesis (HN1) and (HN2).*

*Proof.* It is easy to see that it is a norm, only notice that if $||f||_{B_0} = 0$ then $f(y) = 0$ for all $y \in X^*$, i.e. $f = 0$.

Let us take $y \in X^*$, then

$$|\delta_y(f)| = |f(y)| = (1 + \sum_{j=1}^{p} ||y||^j) \frac{|f(y)|}{1 + \sum_{j=1}^{p} ||y||^j} \le (1 + \sum_{j=1}^{p} ||y||^j)||f||_{B_0},$$

for every $f \in B_0$. Then, the evaluation functionals are continuous and $||\cdot||$ satisfies the hypothesis (HN1). Let us prove the hypothesis (HN2). Let us take a Cauchy sequence $\{f_n\}$ in $B_0$ such that $f_n(y) \to 0$ for every $y \in X^*$. Since $\{f_n\}$ is a Cauchy sequence in $B_0$, for every $\varepsilon > 0$ there is $N_0 \ge 1$ such that for every $n, m \ge N_0$,

$$||f_n - f_m||_{B_0} < \varepsilon/2.$$

Fixed $y \in X^*$, since $f_m(y) \to 0$, then there is $m_0 = m_0(\varepsilon, y) \in \mathbb{N}$ such that for every $m \ge m_0$

$$\frac{|f_m(y)|}{1 + \sum_{j=1}^{p} ||y||^j} < \varepsilon/2.$$

So, for every $n \ge N_0$, let us take $m \ge \max\{N_0, m_0\}$ and we have that

$$\frac{|f_n(y)|}{1 + \sum_{j=1}^{p} ||y||^j} \le \frac{|f_n(y) - f_m(y)|}{1 + \sum_{j=1}^{p} ||y||^j} + \frac{|f_m(y)|}{1 + \sum_{j=1}^{p} ||y||^j} < ||f_n - f_m||_{B_0} + \varepsilon/2 < \varepsilon.$$

Then, $||f_n||_{B_0} \to 0$. □

Then, according to Theorem 1.2, we conclude that there are $(B, ||\cdot||_B)$ and $(B^\sharp, ||\cdot||_{B^\sharp})$ Banach completions of $B_0$ and $B_0^\sharp$, respectively, such the pair $(B, B^\sharp)$ is a pair of RKBS on $X^*$ and $X$ with the reproducing kernel $K$. Furthermore, if we take the map $\mathcal{L} : B^\sharp \to B^*$ defined as equation (2.5), and denote by $<\cdot,\cdot> : B \times B^* \to \mathbb{C}$ the evaluation map $<x, y^*> = y^*(x)$ for all $x \in B$ and $y^* \in B^*$, we can define $\phi : X \to B$ and $\phi^* : Y \to B^*$ as

$$\phi(x) = K(x,\cdot) = (<x,\cdot>_X + 1)^p \in B \qquad \text{and} \qquad \phi^*(y) = \mathcal{L}(K(\cdot, y)) = \mathcal{L}((<\cdot, y^*>_X + 1)^p) \in B^*,$$

and satisfies that

$$K(x, y) = <\phi(x), \phi^*(y)>.$$

3.2. **Bounded, continuous and $C_0(X)$ kernels.** Let $X$ be a set, and $F(X)$ be one of the following space of functions:

- $\mathcal{B}(X)$, the space of bounded functions on $X$.
- $C(X) \cap \mathcal{B}(X)$, the space of continuous and bounded functions on $X$.
- $C_0(X)$, the space of continuous functions $f : X \to \mathbb{C}$ such that for all $\varepsilon > 0$, the set $\{x \in X : |f(x)| \geq \varepsilon\}$ is compact.

Let $K : X \times X \to \mathbb{C}$ be a function such that for every $x \in X$ the functions $K(x, \cdot) \in F(X)$. Consider the vector spaces

$$B_0 := \operatorname{span}\{K(x, \cdot) : x \in X\} \subset F(X) \qquad \text{and} \qquad B_0^\sharp := \operatorname{span}\{K(\cdot, x) : x \in X\},$$

and define the function $|| \cdot ||_{B_0} : B_0 \to \mathbb{C}$ by

$$||f||_{B_0} := \sup\{|f(x)| : x \in X\} \qquad \text{for every } f \in B_0.$$

**Proposition 3.2.** *$|| \cdot ||_{B_0}$ is a norm on $B_0$ and satisfies the hypothesis (HN1) and (HN2).*

*Proof.* It is straightforward to verify that $|| \cdot ||_{B_0}$ is a norm on $B_0$.

Given $x \in X$, let us consider $\delta_x : B_0 \to \mathbb{C}$ the point evaluation functional, i.e., $\delta_x(f) = f(x)$ for $f \in B_0$. Then

$$|\delta_x(f)| = |f(x)| \leq \sup\{|f(y)| : y \in X\} = ||f||_{B_0}.$$

Let us show that it satisfies (HN2). We have to prove that $||f_n||_{B_0} \to 0$ whenever $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0$ such that $f_n(x) \to 0$ on $X$. Let $\varepsilon > 0$ fixed. On the one hand, there is $N_0 \geq 1$ such that

$$\sup\{|f_n(x) - f_m(x)| : x \in X\} < \frac{\varepsilon}{2} \qquad \text{for every } n, m \geq N_0.$$

On the other hand, for a fixed $x \in X$ there exists $m_0 \in \mathbb{N}$ such that $|f_m(x)| < \varepsilon/2$ for all $m \geq m_0$, since $f_m(x) \to 0$. Thus, let us take $m \geq \max\{N_0, m_0\}$ and

$$|f_n(x)| \leq |f_n(x) - f_m(x)| + |f_m(x)| < \varepsilon$$

for all $n \geq N_0$. Then $||f_n||_{B_0} \to 0$. □

Then, by Theorem 1.2, we conclude that there are $(B, || \cdot ||_B)$ and $(B^\sharp, || \cdot ||_{B^\sharp})$ Banach completions of $B_0$ and $B_0^\sharp$, respectively, such that the pair $(B, B^\sharp)$ is a pair of RKBS on $X$ with the reproducing kernel $K$. In fact, the completion of $B_0$ is given by

$$B := \{f : X \to \mathbb{C} : \textit{ there exists } \{f_n\}_{n=1}^\infty \textit{ a Cauchy sequence in } B_0 \text{ such that } f_n(x) \to f(x) \text{ for all } x \in X\},$$

and the norm $|| \cdot ||_B$ on $B$ by

$$||f||_B := \lim_{n\to\infty} ||f_n||_{B_0},$$

for all $f \in B$, where $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_0$ such that $f_n(x) \to f(x)$ for all $x \in X$.

**Proposition 3.3.** *$B \subset F(X)$.*

*Proof.* If $f \in B$, then there exists a Cauchy sequence $\{f_n\}_{n=1}^\infty$ on $B_0$ such that $f_n(x) \to f(x)$ for all $x \in X$. Recall that $|| \cdot ||_{B_0} = || \cdot ||_\infty$.

- $B \subset \mathcal{B}(X)$. There is $n_0 \in \mathbb{N}$ such that $||f_n - f_m||_{B_0} < 1$ for all $n, m \geq n_0$. Then,

$$|f(x)| \leq \lim_n |f_n(x)| \leq \lim_n ||f_n||_{B_0} \leq ||f_{n_0}||_{B_0} + 1, \qquad \text{for all } x \in X.$$

• If $\{f_n\} \subset C(X)$ then $f \in C(X)$, since if a sequence of continuous functions uniform converges to a function $f$, then $f$ is continuous.

• If $\{f_n\} \subset C_0(X)$, then $f \in C_0(X)$. For any $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $||f_n - f_m||_{B_0} < \varepsilon/4$ for all $n, m \geq n_0$. Since $K = \{x \in X : |f_{n_0}(x)| \geq \varepsilon/2\}$ is compact, the closed set $\{x \in X : |f(x)| \geq \varepsilon\} \subset K$ is also compact. Indeed, let us take $x \in X$ such that $|f(x)| \geq \varepsilon$. Then there is $m_0 \in \mathbb{N}$ such that $|f(x) - f_m(x)| < \varepsilon/4$ for every $m \geq m_0$. Let us fix $m \geq n_0, m_0$. Then we have

$$\begin{aligned}|f_{n_0}(x)| &\geq |f(x)| - |f(x) - f_{n_0}(x)| \geq \varepsilon - (|f(x) - f_m(x)| + |f_m(x) - f_{n_0}(x)|) \\ &\geq \varepsilon - (\varepsilon/4 + \varepsilon/4) = \varepsilon/2,\end{aligned}$$

showing that $\{x \in X : |f(x)| \geq \varepsilon\} \subset K$, and the proof is completed. □

**Proposition 3.4.** *The norm* $||\cdot||_B$ *satisfies:*

$$||f||_B = \sup\{|f(x)| : x \in X\} \qquad \text{for all } f \in B.$$

*Proof.* Firstly, let us show that $\sup\{|f(x)| : x \in X\} \leq ||f||_B$. Let $f \in B$ and $\{f_n\}_{n=1}^\infty$ be a Cauchy sequence of $B_0$ with $f_n(x) \to f(x)$ on $X$. Then, for every $x \in X$

$$|f(x)| \leq \lim_{n\to\infty} |f_n(x)| \leq \lim_{n\to\infty} ||f_n||_{B_0} = ||f||_B.$$

Let us see the reverse inequality. If $f \in B$, then there exists a Cauchy sequence $\{f_n\}_{n=1}^\infty$ on $B_0$ such that $f_n(x) \to f(x)$ for all $x \in X$. For any $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $||f_n - f_m||_{B_0} < \varepsilon/3$ for all $n, m \geq n_0$. Furthermore, there are $x_n \in X$ for all $n$ such that $|f_n(x_n)| > ||f_n||_{B_0} - \varepsilon/3$. Since $|f_m(x_{n_0}) - f_{n_0}(x_{n_0})| \leq ||f_m - f_{n_0}||_{B_0} < \varepsilon/3$, then

$$||f_m||_{B_0} \leq ||f_{n_0}||_{B_0} + \varepsilon/3 \leq |f_{n_0}(x_{n_0})| + 2\varepsilon/3 \leq |f_m(x_{n_0})| + \varepsilon \qquad \text{for every } m \geq n_0.$$

Thus,

$$||f||_B = \lim_m ||f_m||_{B_0} \leq \lim_m |f_m(x_{n_0})| + \varepsilon = |f(x_{n_0})| + \varepsilon \leq \sup\{|f(x)| : x \in X\} + \varepsilon.$$

Since the above inequality holds for every $\varepsilon > 0$, then $||f||_B \leq \sup\{|f(x)| : x \in X\}$ for every $f \in B$. □

As in the general case, we have that for all $g \in B^\sharp$

$$||g||_{B^\sharp} = \sup_{\substack{f\in B \\ ||f||_B \leq 1}} |\langle f, g\rangle_K|\,.$$

3.3. **RKBS with $\ell_1$-norm.** Now, we can obtain a RKBS with $\ell_1$-norm in a easier way than [3]. We suppose the following condition, which is weaker that the given condition in [3].

**(H$_1$)** For every sequence $\{x_i\}_{i=1}^n \subset X$ and $\varepsilon_i = +1$ or $\varepsilon_i = -1$ for $1 \leq i \leq n$, there is $f \in B$ with $||f||_B = 1$ such that $f(x_i) = \varepsilon_i$.

This assumption is satisfied if the kernel $K$ is *universal*, i.e. $\text{span}\{K(x, .) : x \in X\}$ is dense in $C_0(X)$ (see [4]).

In [3] it has been proved that the following kernels satisfy the hypothesis (H$_1$):

- the exponential kernel, $K(s,t) = e^{-|s-t|}$ for $s, t \in \mathbb{R}$,
- the Brownian bridge kernel, $K(s,t) = \min\{s,t\} - st$ for $s, t \in (0, 1)$.

**Proposition 3.5.** *If $K$ satisfies hypothesis $(H_1)$, then, for every $g = \sum_{j=1}^{m} \beta_j K(\cdot, y_j) \in B_0^\sharp$ the norm $||\cdot||_{B^\sharp}$ satisfies:*

$$||g||_{B^\sharp} = \sum_{j=1}^{m} |\beta_j| = ||\{\beta_j\}||_1.$$

*Proof.* For every $f \in B$ with $||f||_B \leq 1$, we have

$$|<f, g>_K| \leq \sum_{j=1}^{m} |\beta_j||f(y_j)| \leq \sum_{j=1}^{m} |\beta_j|||f||_B \leq \sum_{j=1}^{m} |\beta_j|.$$

Now, if we chose $\varepsilon_i = \mathrm{sign}(\beta_i)$ then there is $f \in B$ with $||f||_B = 1$ such that $f(y_i) = \varepsilon_i$. Thus

$$||g||_{B^\sharp} \geq |<f, g>_K| = |\sum_{j=1}^{m} \beta_j f(y_j)| = \sum_{j=1}^{m} |\beta_j|.$$

□

**Proposition 3.6.** *If $K$ satisfies hypothesis $(H_1)$, then*

$$B^\sharp \subset \{\sum_{n=1}^{\infty} \beta_n K(\cdot, y_n) : y_n \in X \text{ and } \{\beta_n\}_{n=1}^{\infty} \in \ell_1\}.$$

*Furthermore,*

$$||g||_{B^\sharp} = \sum_{n=1}^{\infty} |\beta_n| \qquad \text{for every } g = \sum_{n}^{\infty} \beta_n K(\cdot, y_n) \in B^\sharp.$$

*Proof.* First of all, let us note that the function $K(x, \cdot)$ is bounded for every $x \in X$, and let us denote its bound by $M_x$.

If $g \in B^\sharp$, then there exists a Cauchy sequence $\{g_n\}_{n=1}^{\infty}$ on $B_0^\sharp$ such that $g_n(x) \to g(x)$ for all $x \in X$ and $||g_n||_{B^\sharp} \to ||g||_{B^\sharp}$. Since $g_n = \sum_{j=1}^{m_n} \beta_j^n K(\cdot, y_j^n)$, by Proposition 3.5 $||g_n||_{B^\sharp} = ||\{\beta_j^n\}_j||_1$ and $\{g_n\}$ is a Cauchy sequence in $B_0^\sharp$, the set of points $\{y_j^n\}_{j,n}$ can be ordered and denoted by $\{y_n\}_n$, and there is $\beta = \{\beta_n\}_n \in \ell_1$ such that $\{\beta_j^n\}_j \to \beta$ when $n \to \infty$.

Let us define $G = \sum_{n=1}^{\infty} \beta_n K(\cdot, y_n)$ on $X$. $G$ is well defined since $|G(x)| \leq \sum_n |\beta_n||K(x, y_n)| \leq M_x||\beta||_1$ for every $x \in X$. Moreover, since $g_n(x) \to g(x)$ for every $x \in X$ and $K(x, \cdot)$ is bounded, we conclude that $g_n(x) \to G(x)$ and $G(x) = g(x)$ for all $x \in X$.

Finally,

$$||g||_{B^\sharp} = \lim_n ||g_n||_{B^\sharp} = \lim_n ||\{\beta_j^n\}_j||_1 = ||\beta||_1.$$

□

Summarizing, we obtain the following theorem.

**Theorem 3.7.** *Let $X$ be a set and $K : X \times X \to \mathbb{C}$ a function such that for every $x \in X$ the functions $K(x, \cdot) \in F(X)$. Let us denote $B_0 = \mathrm{span}\{K(x, \cdot) : x \in X\}$ and $B_0^\sharp = \mathrm{span}\{K(\cdot, x) : x \in X\}$. Then, there are $(B, ||\cdot||_B)$ and $(B^\sharp, ||\cdot||_{B^\sharp})$ Banach completions of $B_0$ and $B_0^\sharp$, respectively, such that they are RKBS on $X$ with the reproducing kernel $K$. Furthermore,*

$$|<f, g>_K| \leq ||f||_B||g||_{B^\sharp} \qquad \text{for all } f \in B \text{ and } g \in B^\sharp,$$

$$||f||_B = \sup\{|f(x)| : x \in X\}, \qquad ||g||_{B^\sharp} = \sup_{\substack{f\in B\\ ||f||_B\le 1}} |<f,g>_K| \quad \textit{for any } f \in B \textit{ and } g \in B^\sharp.$$

*Moreover, if $K$ satisfies hypothesis ($H_1$), then, for every $g \in B^\sharp$, there are $\beta = \{\beta_n\}_n \in \ell_1$ and $\{y_n\}_n$ a sequence of points in $X$ such that $g = \sum_{n=1}^{\infty} \beta_n K(\cdot, y_n)$, and the norm $||\cdot||_{B^\sharp}$ satisfies:*

$$||g||_{B^\sharp} = \sum_{n=1}^{\infty} |\beta_n| = ||\beta||_1.$$

3.4. **RKBS with the norm $\ell_p$.** Here we follow the construction of RKBS with $\ell_1$-norm given in [2]. Let $X$ be a set, $1 \le p, q \le +\infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Let us denote by

$$\ell_q(X) = \{\alpha = \{\alpha(x)\}_{x\in X} \in \mathbb{C}^X : \sum_{x\in X} |\alpha(x)|^q < +\infty\},$$

with the norm $\qquad ||\alpha||_q = (\sum_{x\in X} |\alpha(x)|^q)^{\frac{1}{q}}$ for all $\alpha \in \ell_q(X)$.

Suupose that $K : X \times X \to \mathbb{C}$ is a function which satisfies:

**(Hp1)** $\{K(x, y)\}_{x\in X} \in \ell_q(X)$ for every $y \in X$.

**(Hp2)** Let $\{\alpha_j\}_{j=1}^{\infty} \in \ell_p$ and $\{x_j\}_{j=1}^{\infty}$ a sequence of $X$ such that $\sum_{j=1}^{\infty} \alpha_j K(x_j, \cdot) = 0$ then $\alpha_j = 0$ for any $j$.

**Fact 3.8.** *The hypothesis (Hp1) implies that for every $y \in X$, the set $\{x \in X : K(x, y) \neq 0\}$ is countable.*

**Fact 3.9.** *When $X = \{x_j\}_{j=1}^{\infty}$ is a countable set, $\ell_q(X) = \ell_q(\mathbb{N})$.*

Let us introduce the vector spaces

$$B_0 := \operatorname{span}\{K(x, \cdot) : x \in X\} \qquad \text{and} \qquad B_0^\sharp := \operatorname{span}\{K(\cdot, x) : x \in X\},$$

and define the function $||\cdot||_{B_0} : B_0 \to \mathbb{C}$ by

$$||f||_{B_0} = ||\sum_{j=1}^{n} \alpha_j K(x_j, \cdot)||_{B_0} := \left(\sum_{j=1}^{n} |\alpha_j|^p\right)^{\frac{1}{p}}, \qquad \text{for every } f = \sum_{j=1}^{n} \alpha_j K(x_j, \cdot) \in B_0.$$

**Proposition 3.10.** *$||\cdot||_{B_0}$ is a **norm** on $B_0$ and satisfies the hypothesis (HN1) and (HN2).*

*Proof.* Using hypothesis (Hp2), it is straightforward to verify that $||\cdot||_{B_0}$ is a norm on $B_0$.

Let us prove that the evaluation functionals are continuous. Given $y \in X$, let us denote $[y]_K := \left(\sum_{x\in X} |K(x, y)|^q\right)^{\frac{1}{q}}$. Then

$$|\delta_y(f)| = |f(y)| = |\sum_{j=1}^{n} \alpha_j K(x_j, y)| \le \left(\sum_{j=1}^{n} |\alpha_j|^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^{n} |K(x_j, y)|^q\right)^{\frac{1}{q}} \le ||f||_{B_0}[y]_K.$$

Let us show that it satisfies (HN2). For every $n \in \mathbb{N}$ we can write

$$f_n(\cdot) = \sum_{j=1}^{\infty} \alpha_j^n K(x_j, \cdot),$$

where $\alpha^n = \{\alpha_j^n\}_{j=1}^{\infty}$ has finitely nonzero components, so $\alpha^n \in \ell_p(\mathbb{N})$. Since $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_0$, then $\{\alpha^n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\ell_p(\mathbb{N})$, so there exists $\alpha \in \ell_p(\mathbb{N})$ such that $\alpha^n \to \alpha$ in $\ell_p(\mathbb{N})$. Let us define

$$f(\cdot) = \sum_{j=1}^{\infty} \alpha_j K(x_j, \cdot).$$

Hence, for all $x \in X$

$$|f(x) - f_n(x)| = |\sum_{j=1}^{\infty} (\alpha_j - \alpha_j^n) K(x_j, x)| \leq ||\alpha - \alpha^n||_p [x]_K,$$

and, then, $f_n(x) \to f(x)$ for all $x \in X$. Since $f_n(x) \to 0$, then $f(x) = 0$ for all $x \in X$, by hypothesis (Hp2) we obtain that $\alpha_j = 0$ for any $j \in \mathbb{N}$ and

$$\lim_n ||f_n||_{B_0} = \lim_n ||\alpha^n||_p = ||\alpha||_p = 0.$$

Let us notice that the hypothesis (Hp2) is also necessary to proof that $||\cdot||_B$ is well-defined. □

Then, by Theorem 1.2, we can concluded that there are $(B, ||\cdot||_B)$ and $(B^\sharp, ||\cdot||_{B^\sharp})$ Banach completions of $B_0$ and $B_0^\sharp$, respectively, such that the pair $(B, B^\sharp)$ is a pair of RKBS on $X$ with the reproducing kernel $K$. In fact, the completion of $B_0$ is given by

$$B := \{f : X \to \mathbb{C} : \textit{ there exists } \{f_n\}_{n=1}^{\infty} \textit{ a Cauchy sequence in } B_0 \text{ such that } f_n(x) \to f(x) \text{ for all } x \in X\},$$

and the norm $||\cdot||_B$ on $B$ by

$$||f||_B := \lim_{n\to\infty} ||f_n||_{B_0},$$

for all $f \in B$, where $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_0$ such that $f_n(x) \to f(x)$ for all $x \in X$.

In fact, in the same way as the proof of the above proposition, it can be proved that

$$B = \{\sum_{x\in X} c_x K(x, \cdot) : \{c_x\}_{x\in X} \in \ell_p(X)\} \qquad \text{and} \qquad ||\sum_{x\in X} c_x K(x, \cdot)||_B = ||\{c_x\}_{x\in X}||_p.$$

**Theorem 3.11.** *Let $X$ be a set and $K : X \times X \to \mathbb{C}$ a function which satisfies the hypothesis (Hp1) and (Hp2). Let us denote $B_0 = \operatorname{span}\{K(x, \cdot) : x \in X\}$ and $B_0^\sharp = \operatorname{span}\{K(\cdot, x) : x \in X\}$. Then, there are $(B, ||\cdot||_B)$ and $(B^\sharp, ||\cdot||_{B^\sharp})$ Banach completions of $B_0$ and $B_0^\sharp$, respectively, such that they are RKBS on $X$ with the reproducing kernel $K$. Furthermore,*

$$|\langle f, g\rangle_K| \leq ||f||_B ||g||_{B^\sharp} \qquad \textit{for all } f \in B \textit{ and } g \in B^\sharp,$$

$$B = \{\sum_{x\in X} c_x K(x, \cdot) : \{c_x\}_{x\in X} \in \ell_p(X)\}, \qquad ||\sum_{x\in X} c_x K(x, \cdot)||_B = ||\{c_x\}_{x\in X}||_p,$$

*and*

$$||g||_{B^\sharp} = \sup_{\substack{f\in B \\ ||f||_B \leq 1}} |\langle f, g\rangle_K| \quad \textit{for any } g \in B^\sharp.$$

Several common kernels satisfy the hypothesis (Hp1) and (Hp2): the Gaussian kernel, the exponential kernel, etc.

## 4. Main Problems in Statistical Learning Theory

The Statistical Learning Theory originates in the paper of Vapnik and Chervonenkis [5] and is systematically developed in [6]. There is a huge number of publications devoted to various branches of this theory, increasing rapidly, showing the importance and applicability of this theory in practical problems. Here we sketch briefly some of its main tasks.

We are given input-output data $(\mathbf{x}, \mathbf{y})$, $\mathbf{x} = (x_1, ..., x_N), \mathbf{y} = (y_1, ..., y_N)$ and cost functional $L_{\mathbf{x},\mathbf{y}}$. The task is to find a function from an admissible class, usually Reproducing Kernel Hilbert space $H$, that minimizes the perturbed functional

$$L_{\mathbf{x},\mathbf{y}}(f) + \lambda\Phi(\|f\|_H)$$

where $\Phi$ is a regularization functional.

Some of the main problems in Statistical Learning Theory are:

– regularization networks

$$Y = \mathbb{R},\ L_{\mathbf{x},\mathbf{y}}(f) = \sum_{j\in\mathbb{N}_N} |f(x_j) - y_j|^2,\ \Phi(t) = t^2$$

– support vector machine regression

$$Y = \mathbb{R},\ L_{\mathbf{x},\mathbf{y}}(f) = \sum_{j\in\mathbb{N}_N} |f(x_j) - y_j|_\epsilon,\ \Phi(t) = t^2,$$

where $|t|_\epsilon = \max(|t| - \epsilon, 0)$ is called Vapnik $\epsilon$-insensitive norm

– support vector machine classification

$$Y = \{-1, 1\},\ L_{\mathbf{x},\mathbf{y}}(f) = \sum_{j\in\mathbb{N}_N} \max(1 - y_j f(x_j), 0),\ \Phi(t) = t^2$$

The support vector machine classification can be reformulated equivalently as:

$$\max_{\gamma,w,b}\{\gamma : y_n(<x_n, w> + b) \geq \gamma :\ 1 \leq n \leq N \text{ and } ||w|| = 1\}$$

(linear SVM), and

$$\text{(4.1)} \qquad \text{(RKHS)} \qquad \max_{\gamma,\{\alpha_j\}_j,b}\{\gamma : y_n(\sum_{j=1}^N \alpha_j K(x_n, x_j) + b) \geq \gamma : \sum_{i,j=1}^N \alpha_i\alpha_j K(x_i, x_j) = 1\}$$

(nonlinear SVM)

Our contribution to the above model is to replace RKHS with RKBS with $\ell_1$ norm, i.e. to solve the problem

$$\text{(4.2)} \qquad \text{(RKBS)} \qquad \max_{\gamma,\{\alpha_j\}_j,b}\{\gamma : y_n(\sum_{j=1}^N \alpha_j K(x_n, x_j) + b) \geq \gamma : \sum_{j=1}^N |\alpha_j| = 1\}$$

We expect sparse solutions of this problem. For the case of regularization network, sparse solutions in RKBS with $\ell_1$ norm are obtained in [2].

## Acknowledgments

Luis Sánchez-González conducted part of this research while at the Center for Applied Optimization. This author is indebted to the members of the Center and very especially to Pando Georgiev and Panos Pardalos for their kind hospitality and for many useful discussions.

## References

[1] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337-404.

[2] G. Song, H. Zhang and F.J. Hickernell, *Reproducing kernel Banach spaces with the $\ell_1$ norm*, preprint 2011.

[3] G. Song and H. Zhang, *Reproducing kernel Banach spaces with the $\ell_1$ norm II: Error Analysis for Regularized Least Square Regression*, preprint 2011.

[4] I. Steinwart, *On the Influence of the Kernel on the Consistency of Support Vector Machines*, Journal of Machine Learning Research 2 (2001) 67-93.

[5] V. N. Vapnik, and A. Y. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Theory of Probability and its Applications, Vol 16, 1971, pp 264-280.

[6] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, 1998.