ARTICLE IN PRESS

Automatica ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

On the stability of reproducing kernel Hilbert spaces of discrete-time impulse responses[☆]

Tianshi Chen[a], Gianluigi Pillonetto[b,*]

[a] School of Science and Engineering and Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen, China
[b] Department of Information Engineering, University of Padova, Padova, Italy

ARTICLE INFO

Article history:
Received 26 October 2017
Received in revised form 16 April 2018
Accepted 27 April 2018
Available online xxxx

Keywords:
Linear system identification
Reproducing kernel Hilbert spaces
BIBO stability
Kernel-based regularization
Inverse problems

ABSTRACT

Reproducing kernel Hilbert spaces (RKHSs) have proved themselves to be key tools for the development of powerful machine learning algorithms, the so-called regularized kernel-based approaches. Recently, they have also inspired the design of new linear system identification techniques able to challenge classical parametric prediction error methods. These facts motivate the study of the RKHS theory within the control community. In this note, we focus on the characterization of *stable* RKHSs, i.e. RKHSs of functions representing stable impulse responses. Related to this, working in an abstract functional analysis framework, Carmeli et al. (2006) has provided conditions for an RKHS to be contained in the classical Lebesgue spaces $\mathscr{L}^p$. In particular, we specialize this analysis to the discrete-time case with $p = 1$. The necessary and sufficient conditions for the stability of an RKHS are worked out by a quite simple proof, more easily accessible to the control community.

1. Introduction

Reproducing Kernel Hilbert Spaces (RKHS) were developed in the seminal works (Aronszajn, 1950; Bergman, 1950) and possess important properties. They are in one-to-one correspondence with the class of positive definite kernels and have also an important interpretation in the context of Gaussian processes (Aravkin, Bell, Burke, & Pillonetto, 2015; Kimeldorf & Wahba, 1971; Lukic & Beder, 2001). RKHSs have been introduced within the machine learning community (Girosi, 1997) leading, in conjunction with Tikhonov regularization theory (Bertero, 1989; Tikhonov & Arsenin, 1977), to the development of new powerful algorithms (Cucker & Smale, 2001; Drucker, Burges, Kaufman, Smola, & Vapnik, 1997; Evgeniou, Pontil, & Poggio, 2000; Schölkopf & Smola, 2001). Many of these regularized approaches recover an infinite-dimensional function f from a finite set of noisy measurements y_i by optimizing an objective over a suitable RKHS $\mathscr{H}$. As an example, the so called regularization networks (Poggio & Girosi, 1990; Wahba, 1990) take the form

$$\arg\min_{f \in \mathscr{H}} \sum_{i=1}^{N} (y_i - L_i[f])^2 + \gamma \|f\|^2_{\mathscr{H}}, \tag{1}$$

where $i, N \in \mathbb{N}$, γ is a positive scalar, $L_i : \mathscr{H} \to \mathbb{R}$ is linear and continuous and $\|\cdot\|^2_{\mathscr{H}}$ is the regularizer given by the RKHS (squared) norm which restores the well-posedness.

This kind of regularized paradigm has been recently introduced also in the linear system identification context. In particular, in Pillonetto and De Nicolao (2010), f is thought of as the unknown impulse response of a linear time-invariant system, y_i is the noisy output and the L_i is defined by the convolution between f and the system input. A key point is to include in the KRHS $\mathscr{H}$ the available system information. For this purpose, Pillonetto & De Nicolao (2010) has introduced a new RKHS, defined by the so called *stable spline* kernel, which embeds both smoothness and exponential stability. Equipped with this RKHS, the estimator (1), and its extensions developed in Chen, Andersen, Ljung, Chiuso, and Pillonetto (2014), Chen, Ohlsson, and Ljung (2012), Pillonetto, Chiuso, and De Nicolao (2011), have proved to challenge consolidated approaches such as classical Prediction Error Methods (PEM) (Ljung, 1999; Söderström & Stoica, 1989), equipped with the classical model structure selection methods, e.g., Akaike's information criterion and cross validation; see Pillonetto, Dinuzzo, Chen, Nicolao, and Ljung (2014) for a survey on the interplay between regularization, machine learning and system identification.

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor A. Pedro Aguiar under the direction of Editor André L. Tits. https://topnonprofits.com/examples/nonprofit-mission-statements/.

* Corresponding author.
E-mail addresses: tschen@cuhk.edu.cn (T. Chen), giapi@dei.unipd.it (G. Pillonetto).

https://doi.org/10.1016/j.automatica.2018.05.017
0005-1098/

All of these facts motivate a more careful study of RKHS theory within the control community. This note is in particular devoted to characterize all *stable RKHSs*, i.e. RKHS whose elements are stable impulse responses (absolutely integrable for continuous-time case and absolutely summable for discrete-time case). As first pointed out by Dinuzzo (2015), this problem is connected with the work (Carmeli, Vito, & Toigo, 2006), which has provided the conditions for an RKHS to be contained in the classical Lebesgue spaces L^p. The proof provided in Carmeli et al. (2006) involves advanced abstract functional analysis concepts and results, e.g., the measure theory, and is nontrivial to understand. In this paper, we specialize this analysis to the case $p = 1$, working in discrete-time (the function domain is the set of natural numbers). We provide the necessary and sufficient conditions for the stability of an RKHS through a quite simple proof, which in contrast with (Carmeli et al., 2006) only relies on the closed graph theorem and some basic concepts of weak convergence, e.g. see Megginson (1998), Zeidler (1995), and thus is more easily accessible to the control community.

This note is organized as follows. In Section 2, we provide a brief introduction to RKHS theory, and in Section 3, we review the concept of stable RKHSs and their characterizations. We then provide a simple proof for the characterization of stable RKHSs in Section 4 and conclude the note in Section 5.

2. Reproducing kernel Hilbert spaces

2.1. RKHSs and their relationship with kernels

We start by recalling some basic facts about RKHSs and in particular their characterization via the concept of a *kernel*; see e.g., Kennedy and Sadeghi (2013, Chapter 10) for systematic treatment.

In the following, f denotes a real function over the domain $\mathscr{X}$. Recall that a Hilbert space $\mathscr{H}$ of functions f is a complete vector space endowed with an inner product $\langle\cdot,\cdot\rangle_{\mathscr{H}}$. AN RKHS is a special Hilbert space of functions f where it is also assumed that pointwise evaluations are continuous linear functionals on $\mathscr{H}$. This means that, for any $x \in \mathscr{X}$ there exists a scalar $A < \infty$ (possibly dependent on x) such that

$$|f(x)| \leq A\|f\|_{\mathscr{H}}, \quad \forall f \in \mathscr{H}. \tag{2}$$

Then, RKHSs are defined as follows.

Definition 2.1 (*RKHS*)**.** A Hilbert space of functions $f : \mathscr{X} \to \mathbb{R}$ is called an RKHS if (2) holds.

A fundamental characterization of an RKHS can be obtained by the concept of positive semidefinite kernel.

Definition 2.2 (*Positive Semidefinite Kernel*)**.** A symmetric function $K : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ is called *positive semidefinite kernel* if, for any $m \in \mathbb{N}$, $c_1, \ldots, c_m \in \mathbb{R}$, and $x_1, \ldots, x_m \in \mathscr{X}$, it holds that

$$\sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j K(x_i, x_j) \geq 0.$$

In addition, given a kernel K, we define the *kernel section* $K_x(\cdot)$ centered at $x \in \mathscr{X}$ as the function $\mathscr{X} \to \mathbb{R}$ defined by

$$K_x(y) = K(x, y), \ \forall y \in \mathscr{X}.$$

The connection between positive semidefinite kernels and RKHS is illustrated in the next result.

Theorem 2.1 (*Moore-Aronszajn*)**.** *To every RKHS $\mathscr{H}$ there corresponds a unique positive semidefinite kernel K, called the* reproducing kernel, *such that the* reproducing property *holds:*

$$f(x) = \langle f, K_x\rangle_{\mathscr{H}}, \qquad \forall (x, f) \in (\mathscr{X}, \mathscr{H}). \tag{3}$$

Conversely, given a positive semidefinite kernel K, there exists a unique RKHS of real valued functions defined over $\mathscr{X}$ whose reproducing kernel is K.

Further remarks on the nature of an RKHS are now in order (details can be found e.g. on p. 35 of Cucker and Smale (2001)). The Moore-Aronszajn Theorem shows that the Hilbert space $\mathscr{H}$ is completely characterized by its reproducing kernel. In particular, every RKHS is generated by the kernel sections as follows. First, consider all the functions of the type

$$f(\cdot) = \sum_{i=1}^{m} c_i K_{x_i}(\cdot)$$

for any $m \in \mathbb{N}$, $c_i \in \mathbb{R}$, $x_i \in \mathscr{X}$ and any $p \in \mathbb{N}$, $d_i \in \mathbb{R}$, $y_i \in \mathscr{X}$. All of the resulting functions form a subspace equipped with the inner product $\langle f, g\rangle_{\mathscr{H}} = \sum_{i=1}^{m}\sum_{j=1}^{p} c_i d_j K(x_i, y_j)$, inducing the norm

$$\|f\|_{\mathscr{H}}^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j K(x_i, x_j). \tag{4}$$

The RKHS associated with K then corresponds to the union of this subspace and all limits of Cauchy sequences. Summarizing, we have

- all the kernel sections $K_x(\cdot)$ belong to the RKHS $\mathscr{H}$ induced by K;
- $\mathscr{H}$ contains also all the finite sums of kernel sections along with some particular infinite sums with finite norm;
- every $f \in \mathscr{H}$ is thus a linear combination of a possibly infinite number of kernel sections.

2.2. RKHS of functions defined by sequences

Hereafter, the focus is on RKHS containing impulse responses of linear discrete-time systems. Since the impulse response is defined by a sequence of real numbers, we set $\mathscr{X} = \mathbb{N}$. Moreover, we use the following sequence spaces:

$$\ell_1 = \left\{\{f_i\}_{i=1}^{\infty} : \|f\|_1 = \sum_{i=1}^{\infty}|f_i| < \infty, f_i \in \mathbb{R}\right\},$$

$$\ell_2 = \left\{\{f_i\}_{i=1}^{\infty} : \|f\|_2 = (\sum_{i=1}^{\infty}|f_i|^2)^{\frac{1}{2}} < \infty, f_i \in \mathbb{R}\right\},$$

$$\ell_\infty = \left\{\{f_i\}_{i=1}^{\infty} : \|f\|_\infty = \sup_{i\in\mathbb{N}}|f_i| < \infty, f_i \in \mathbb{R}\right\},$$

where f_i is the ith element of f with $i \in \mathbb{N}$. Also, from now *sequence*, *impulse response* and *function* are exchangeable terms, with f to denote both the sequence $\{f_i\}_{i=1}^{\infty}$ and the function $\mathbb{N} \to \mathbb{R}$ $(i \to f_i)$. Sometimes, to stress its functional nature, we write $f(\cdot)$.

The particular domain choice makes also all the introduced kernels maps from $\mathbb{N} \times \mathbb{N}$ into $\mathbb{R}$. Given an index $i \in \mathbb{N}$, the notation K_i denotes both the sequence $\{K(i,j)\}_{j=1}^{\infty}$ and the kernel section $K_i(\cdot) := K(i,\cdot)$ centered at $i \in \mathbb{N}$.

3. Stable reproducing kernel Hilbert spaces

3.1. Stable kernels and RKHS

The necessary and sufficient condition for a discrete-time linear time invariant system to be bounded-input bounded-output stable