CrossMark

Generalized approach to the problem of regression

Dariusz Partyka · **Józef Zając**

Abstract As a result of studying and solving certain extremal problems defined on the basis of finite and infinite dimensional pseudo-Hilbert space, the authors present a generalization of the classical regression idea. Given discrete or continuous empiric data, the class of solutions is uniquely determined and expressed in a new form of the *regression function sequences*, both within the asynchronous and synchronous type in a given pseudo-Hilbert space. With the help of this new technique, a large variety of observed phenomena in different areas of practical and theoretical sciences can be precisely described and investigated.

Keywords   Approximation · Asynchronous regression structure · Hilbert space · Measurable space · Orthogonal decomposition · Orthogonal sequence · Optimization · Probabilistic space · Regression structure · Regression function · Synchronous regression structure

D. Partyka · J. Zając
Institute of Mathematics and Information Technology,
The State School of Higher Education in Chełm,
Pocztowa 54, 22-100 Chełm, Poland

J. Zając
e-mail: jzajac@kul.lublin.pl

D. Partyka (✉)
Institute of Mathematics and Computer Science,
The John Paul II Catholic University of Lublin,
Al. Racławickie 14, P.O. Box 129, 20-950 Lublin, Poland
e-mail: partyka@kul.lublin.pl

🖄 Birkhäuser

1 Introduction

The contemporary world is characterized by an increasing influx of information. It is possible to achieve some of them in applied selected moments of time, whereas others can be observed in continuous time system. In this avalanche of information, one has to know how to find both relational and functional dependence. The latter is often too complicated to capture and describe by means of simple mathematical expression.

Approximate functional dependences which would describe a number of interesting phenomena with assigned accuracy should be sought. The study of appropriately constructed approaching functions can lead to detection of not yet discovered dependences, as well as assessments separate and combined effects caused by several observed variables. This has a huge significance, especially in situations where dependence, expressed in terms of physical, chemical and biological laws, between observed parameters is unknown.

A particular, although simplified, example of a solution to such a set of scientific problems is a method of linear regression with its various modifications, formulated on the basis of probability calculus; cf. [1,11,14,15,17]. This method led to a series of implementations and experienced numerous theoretical modifications, crucial due to the seriousness of the implementation problem; cf. [4–6,13,15,16].

The regression problem considered in this paper has the form of solution of a properly formulated extremal problem well defined and stated in the both finite and infinite dimensional pseudo-Hilbert space environment.

A short presentation of the classical approach to the regression problem can be formulated as follows.

For any $p, q \in \mathbb{R}$ set

$$\mathbb{Z}_{p,q} := \{k \in \mathbb{Z} : p \leq k \leq q\} \quad \text{and} \quad \mathbb{Z}_{p,\infty} := \{k \in \mathbb{Z} : p \leq k\}.$$

Let $\mathcal{F}$ be the family of all functions $\mathbb{R} \ni t \mapsto at + b$, where $a, b \in \mathbb{R}$, and let $x, y : \mathbb{Z}_{0,n} \to \mathbb{R}$ be arbitrarily given sequences. It is well known that if x is not a constant sequence, then there exists the unique $f_0 \in \mathcal{F}$ satisfying the following condition

$$\sum_{k=0}^{n} (f(x_k) - y_k)^2 \geq \sum_{k=0}^{n} (f_0(x_k) - y_k)^2, \quad f \in \mathcal{F}. \tag{1.1}$$

In fact, the function f_0 is of the form $f_0(t) = a_0 t + b_0$ as $t \in \mathbb{R}$, where

$$a_0 := \frac{(n+1)\sum_{k=0}^{n} x_k y_k - \sum_{k=0}^{n} x_k \sum_{k=0}^{n} y_k}{(n+1)\sum_{k=0}^{n} x_k^2 - \left(\sum_{k=0}^{n} x_k\right)^2} \quad \text{and}$$

$$b_0 := \frac{\sum_{k=0}^{n} y_k - a_0 \sum_{k=0}^{n} x_k}{n+1} ; \tag{1.2}$$

cf. e.g. [17] and [4]. The function f_0 is usually said to be the *regression line* for the empiric sequences $x, y : \mathbb{Z}_{0,n} \to \mathbb{R}$. In view of (1.1), the function f_0 has a natural interpretation as an optimal function with the smallest quadratic deviation from the empiric observations $\{(x_k, y_k) : k \in \mathbb{Z}_{0,n}\}$. The function f_0 plays the very essential role in different areas of the applied mathematics; cf. e.g. [6] and [4]. It shows that the mentioned above extremal problem, can be considerably qeneralized and solved, which is a subject of this paper. To this end we introduce the *regression structures* $\mathfrak{R} := (A, B, \delta; x, y)$, where:

I.1 *A, B are given nonempty sets;*

I.2 $x : \Omega_1 \to A$ and $y : \Omega_2 \to B$ *for some nonempty sets Ω_1 and Ω_2;*

I.3 $\delta : (\Omega_1 \to B) \times (\Omega_2 \to B) \to \overline{\mathbb{R}}$.

A set $\mathcal{F}$ is called the *functional model* of $\mathfrak{R}$ if $\mathcal{F} \subset (A \to B)$, where $A \to B$ denotes the class of all functions acting from A to B.

According to the extremal problem (1.1), the components δ, x and y as well as a functional model $\mathcal{F}$ of $\mathfrak{R}$ have the following interpretations:

- $\mathcal{F}$ is a theoretic functional model of the observed phenomena, i.e. $\mathcal{F}$ consists of all functions describing theoretically the considered phenomena;
- $x : \Omega_1 \to A$ and $y : \Omega_2 \to B$ are empirical functions derived from an experiment or observation, called the *empirical data functions* in the sequel;
- δ is a deviation criterion of the theoretic functions from the empirical ones.

Given a regression structure $\mathfrak{R}$ and a functional model $\mathcal{F}$ of $\mathfrak{R}$ we seek the optimal theoretic functions $f_0 \in \mathcal{F}$ which are the best fitted to the empirical data—represented by the empirical data functions x and y—with respect to the criterion δ. To be more precise, we consider the extremal problem of determining all functions $f_0 \in \mathcal{F}$ minimizing the functional

$$\mathcal{F} \ni f \to F(f) := \delta(f \circ x, y) \in \overline{\mathbb{R}}, \tag{1.3}$$

i.e. all functions $f_0 \in \mathcal{F}$ satisfying the following inequality

$$F(f) \geq F(f_0), \quad f \in \mathcal{F}. \tag{1.4}$$

The set of all $f_0 \in \mathcal{F}$ satisfying the inequality (1.4) will be denoted by $\mathrm{Reg}(\mathcal{F}, \mathfrak{R})$. Each function $f_0 \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ is said to be the *regression function in $\mathcal{F}$ with respect to $\mathfrak{R}$*. The problem of describing all regression functions in $\mathcal{F}$ with respect to $\mathfrak{R}$ we call the *regression problem for $\mathcal{F}$ with respect to $\mathfrak{R}$*.

Example 1.1 Consider an electric circuit with direct current. According to Ohm's law the voltage V depends on the intension I by the equality $V = RI$, where the multiplier R is the resistance of the circuit. We want to determine the parameter R by means of measurements samples of intensity and voltage represented by a sequence

$\mathbb{Z}_{0,n} \ni k \mapsto (i_k, v_k)$. To this end we consider the regression structure $\mathfrak{R}$, where $A := \mathbb{R}$, $B := \mathbb{R}$, the empiric data functions are defined by

$$\mathbb{Z}_{0,n} \ni k \mapsto x(k) := i_k \quad \text{and} \quad \mathbb{Z}_{0,n} \ni k \mapsto y(k) := v_k,$$

and, as a criterion of deviation δ, we take the least squares method, i.e.

$$\delta(f, g) := \sum_{k=0}^{n} (f(k) - g(k))^2, \quad f, g : \mathbb{Z}_{0,n} \to \mathbb{R}. \tag{1.5}$$

The theoretic functional model $\mathcal{F}$ is represented by linear functions $\mathbb{R} \ni t \mapsto rt$ for $r \in \mathbb{R}$. Calculating the critical point of the function

$$\mathbb{R} \ni r \mapsto \sum_{k=0}^{n} (r i_k - v_k)^2$$

we obtain

$$R = \sum_{k=0}^{n} i_k v_k \Big/ \sum_{k=0}^{n} i_k^2. \tag{1.6}$$

In what follows we shall study the regression problem for a wide range of theoretic functional models $\mathcal{F}$ and regression structures $\mathfrak{R}$ involving a generalized variant of quadratic deviation applied in (1.1). The main idea of our approach was presented in the paper [12] where we confined ourselves to the case where the theoretic functional model $\mathcal{F}$ is a finite-dimensional linear set with respect to the standard operations of adding and multiplying complex-valued functions. In the present paper we consider the general case where $\mathcal{F}$ is an arbitrary linear set. We also provide various examples motivating our general approach to the regression problem and discuss the, so called, diagonal case much related to the classical approach to the regression problem.

2 Asynchronous regression structures

Given a nonempty set Ω and a σ-field $\mathcal{B}$ of its subsets we denote by $L(\Omega, \mathcal{B})$ the set of all complex valued functions on Ω, measurable with respect to $\mathcal{B}$. We write $\mathbf{L}(\Omega, \mathcal{B})$ for the linear space supported by $L(\Omega, \mathcal{B})$ and standard operations of adding and multiplying functions. For a given measure $\mu : \mathcal{B} \to [0; +\infty]$ and $p \geq 1$, let $L^p(\Omega, \mathcal{B}, \mu)$ stand for the class of all functions $f \in L(\Omega, \mathcal{B})$ such that

$$\|f\|_{\mu,p} := \left(\int_{\Omega} |f|^p \mathrm{d}\mu \right)^{1/p} < +\infty. \tag{2.1}$$

We recall that for each $p \geq 1$, the class $L^p(\Omega, \mathcal{B}, \mu)$ is a linear set in $\mathbf{L}(\Omega, \mathcal{B})$ and $\| \cdot \|_{\mu,p}$ is a pseudo-norm defined on the linear space $(L^p(\Omega, \mathcal{B}, \mu), +, \cdot)$ satisfying the following condition

$$\|f\|_{\mu,p} = 0 \Longrightarrow \mu(\{t \in \Omega : f(t) \neq 0\}) = 0$$

and the structure

$$\mathbf{L}^p(\Omega, \mathcal{B}, \mu) := (\mathrm{L}^p(\Omega, \mathcal{B}, \mu), +, \cdot, \|\cdot\|_{\mu,p})$$

is a complete pseudo-normed space, i.e. $\mathbf{L}^p(\Omega, \mathcal{B}, \mu)$ is a pseudo-Banach space; cf. e.g. [2, Sect. 2.4], [8, Example 1.2.4]. If $p = 2$, then the structure

$$(\mathrm{L}^2(\Omega, \mathcal{B}, \mu), +, \cdot, \langle\cdot|\cdot\rangle_\mu)$$

is a pseudo-Hilbert space, where

$$\langle f|g\rangle_\mu := \int_\Omega f\overline{g}\,\mathrm{d}\mu, \quad f, g \in \mathrm{L}^2(\Omega, \mathcal{B}, \mu).$$

Hence and by (2.1),

$$\|f\|_{\mu,2} = \|f\|_\mu := \sqrt{\langle f|f\rangle_\mu}, \quad f \in \mathrm{L}^2(\Omega, \mathcal{B}, \mu).$$

From now on we shall study the family of regression structures

$$\mathfrak{R} := (A, B, \delta; x, y)$$

satisfying the following additional assumptions:

II.1 $B = \mathbb{R}$ *or* $B = \mathbb{C}$;

II.2 *There exist a σ-field $\mathcal{B}$ of subsets of the cartesian product $\Omega_1 \times \Omega_2$ and a measure $\mu : \mathcal{B} \to [0; +\infty]$ such that the function δ satisfies for all $u : \Omega_1 \to B$ and $v : \Omega_2 \to B$ the following equality*

$$\delta(u, v) = \int_{\Omega_1 \times \Omega_2} |u(t_1) - v(t_2)|^2 \mathrm{d}\mu(t_1, t_2) \tag{2.2}$$

provided the function

$$\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto |u(t_1) - v(t_2)|$$

is $\mathcal{B}$-measurable, and $\delta(u, v) = +\infty$, otherwise.

Then the regression problem for $\mathcal{F}$ with respect to $\mathfrak{R}$ means the extremal problem of determining all functions $f_0 \in \mathcal{F}$ minimizing the functional F and satisfying, according to (1.3) and (2.2), the following equality

$$F(f) = \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) - y(t_2)|^2 \mathrm{d}\mu(t_1, t_2), \quad f \in \mathcal{F}.$$

Definition 2.1 Any regression structure $\mathfrak{R}$ satisfying the conditions II.1 and II.2 is said to be the *asynchronous regression structure*; *real asynchronous regression structure* as $B = \mathbb{R}$ and *complex asynchronous regression structure* as $B = \mathbb{C}$.

Example 2.2 Consider a regression structure $\mathfrak{R}$ defined as follows. Given $n, m \in \mathbb{N}$ let $\Omega_1 := \mathbb{Z}_{0,n}$ and $\Omega_2 := \mathbb{Z}_{0,m}$. Let $\mathcal{B}$ be the family of all subsets of the cartesian product $\Omega_1 \times \Omega_2$. Obviously, the family $\mathcal{B}$ is a σ-field of subsets of $\Omega_1 \times \Omega_2$, and hence we can define a unique measure $\mu : \mathcal{B} \to [0; +\infty)$ satisfying the condition

$$\mu(\{(k, l)\}) = \rho_{k,l}, \quad k \in \Omega_1, l \in \Omega_2, \tag{2.3}$$

where $\Omega_1 \times \Omega_2 \ni (k, l) \mapsto \rho_{k,l} \in \mathbb{R}$ is a given non-negative function. Then, for any functions $\Omega_1 \ni t \mapsto u(t) \in B$ and $\Omega_2 \ni t \mapsto v(t) \in B$, we derive from (2.2),

$$
\begin{aligned}
\delta(u, v) &= \int_{\Omega_1 \times \Omega_2} |u(t_1) - v(t_2)|^2 d\mu(t_1, t_2) \\
&= \sum_{(k,l) \in \Omega_1 \times \Omega_2} \int_{\{(k,l)\}} |u(t_1) - v(t_2)|^2 d\mu(t_1, t_2) \\
&= \sum_{k=0}^{n} \sum_{l=0}^{m} \rho_{k,l} |u(k) - v(l)|^2.
\end{aligned}
\tag{2.4}
$$

In particular, assuming $m = n$ and setting

$$\rho_{k,l} := \begin{cases} 1 & \text{as} \quad k = l, \\ 0 & \text{as} \quad k \neq l, \end{cases} \tag{2.5}$$

we conclude from (2.4) that

$$\delta(u, v) = \sum_{k=0}^{n} |u(k) - v(k)|^2. \tag{2.6}$$

Combining then (1.3) with (2.4) and (2.5) we can see that

$$F(f) = \sum_{k=0}^{n} |f \circ x(k) - y(k)|^2 = \sum_{k=0}^{n} |f(x_k) - y_k|^2, \quad f \in \mathcal{F}, \tag{2.7}$$

for given empirical data functions $\mathbb{Z}_{0,n} \ni k \mapsto x_k \in A$ and $\mathbb{Z}_{0,n} \ni k \mapsto y_k \in B$, where $A := \mathbb{R}$, $B := \mathbb{R}$ and $\mathcal{F}$ is the family of all functions $\mathbb{R} \ni t \mapsto at + b$ as $a, b \in \mathbb{R}$. Therefore δ is exactly the classical square deviation used in (1.1).

Fix a regression structure $\mathfrak{R}$ satisfying the properties II.1–II.2. We consider the family $\mathcal{L}_1(\mathfrak{R})$ of all functions $f : A \to B$ such that

$$\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto f \circ x(t_1)$$

is a $\mathcal{B}$-measurable function and

$$\int_{\Omega_1 \times \Omega_2} |f \circ x(t_1)|^2 d\mu(t_1, t_2) < +\infty. \tag{2.8}$$

We shall also consider the family[1] $\mathcal{L}_2(\mathfrak{R})$ of all functions $g : B \to B$ such that

$$\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto g \circ y(t_2)$$

is a $\mathcal{B}$-measurable function and

$$\int_{\Omega_1 \times \Omega_2} |g \circ y(t_2)|^2 d\mu(t_1, t_2) < +\infty. \tag{2.9}$$

From (2.8) and the inequality

$$|z\overline{w}| \leq \frac{1}{2}(|z|^2 + |w|^2), \quad z, w \in \mathbb{C}, \tag{2.10}$$

it follows that the functional

$$\mathcal{L}_1(\mathfrak{R}) \times \mathcal{L}_1(\mathfrak{R}) \ni (u, v) \to \langle u|v \rangle := \int_{\Omega_1 \times \Omega_2} u \circ x(t_1) \overline{v \circ x(t_1)} d\mu(t_1, t_2) \tag{2.11}$$

is well defined. Hence, $\langle u|u \rangle \geq 0$ as $u \in \mathcal{L}_1(\mathfrak{R})$, and so the functional

$$\mathcal{L}_1(\mathfrak{R}) \ni u \mapsto \|u\| := \sqrt{\langle u|u \rangle} = \left(\int_{\Omega_1 \times \Omega_2} |u \circ x(t_1)|^2 d\mu(t_1, t_2) \right)^{1/2}, \tag{2.12}$$

is also well defined.

Lemma 2.3 *The structure[2] $\mathrm{H}(\mathfrak{R}) := (\mathcal{L}_1(\mathfrak{R}), +, \cdot, \langle \cdot | \cdot \rangle)$ is a complex (resp. real in case $B = \mathbb{R}$) pseudo-Hilbert space, i.e. the structure $(\mathcal{L}_1(\mathfrak{R}), +, \cdot)$ is a linear space and the following properties*

$$\begin{aligned}
\langle \alpha u + \beta v | w \rangle &= \alpha \langle u|w \rangle + \beta \langle v|w \rangle ; \\
\langle u|v \rangle &= \overline{\langle v|u \rangle}; \\
\langle u|u \rangle &\geq 0, \tag{2.13}
\end{aligned}$$

hold for all $\alpha, \beta \in B$ and $u, v, w \in \mathcal{L}_1(\mathfrak{R})$ as well as for every sequence $\mathbb{N} \ni n \mapsto f_n \in \mathcal{L}_1(\mathfrak{R})$ satisfying the Cauchy condition

$$\|f_n - f_m\| \to 0 \quad as \ n, m \to \infty \tag{2.14}$$

---

[1] Note that the space was redefined as compared to the one in the paper [12]. This simplifies considerations.

[2] Here and in the sequel the symbols "+" and "·" denote the standard operations of adding and multiplying functions in the context of real or complex valued functional spaces.

there exists $f \in \mathcal{L}_1(\mathfrak{R})$ *such that*

$$\|f - f_n\| \to 0 \quad as \; n \to \infty. \tag{2.15}$$

Proof By (2.10) we get

$$|z + w|^2 \leq 2(|z|^2 + |w|^2), \quad z, w \in \mathbb{C}. \tag{2.16}$$

Hence and by (2.8) we see that for all $\lambda_1, \lambda_2 \in B$ and $u, v \in \mathcal{L}_1(\mathfrak{R})$,

$$\int_{\Omega_1 \times \Omega_2} |(\lambda_1 u + \lambda_2 v) \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2) \; \leq 2|\lambda_1|^2 \int_{\Omega_1 \times \Omega_2} |u \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2)$$
$$+ 2|\lambda_2|^2 \int_{\Omega_1 \times \Omega_2} |v \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2)$$
$$< +\infty.$$

Thus $\lambda_1 u + \lambda_2 v \in \mathcal{L}_1(\mathfrak{R})$ for all $\lambda_1, \lambda_2 \in B$ and $u, v \in \mathcal{L}_1(\mathfrak{R})$, and so the structure $(\mathcal{L}_1(\mathfrak{R}), +, \cdot)$ is a linear space as a linear subspace of the linear space $(A \to B, +, \cdot)$. Moreover, by (2.11) we obtain all the properties (2.13).

It remains to prove the completeness of $H(\mathfrak{R})$. The mapping $x : \Omega_1 \to A$ induces the σ-field

$$\mathcal{B}_x := \{V \in 2^A : x^{-1}(V) \times \Omega_2 \in \mathcal{B}\},$$

where 2^A denotes the family of all subsets of the set A, and the measure

$$\mathcal{B}_x \ni V \to \mu_x(V) := \mu(x^{-1}(V) \times \Omega_2).$$

Fix $u \in \mathcal{L}_1(\mathfrak{R})$. Since the function

$$\Omega_1 \times \Omega_2 \ni (t_1, t_2) \mapsto u \circ x(t_1)$$

is $\mathcal{B}$-measurable we see that for every Borel set $U \subset B$,

$$x^{-1}(u^{-1}(U)) \times \Omega_2 = (u \circ x)^{-1}(U) \times \Omega_2 \in \mathcal{B},$$

and hence $u^{-1}(U) \in \mathcal{B}_x$. Thus u is $\mathcal{B}_x$-measurable. Moreover, from (2.12) and Lemma 7.3 it follows that

$$\|u\|^2 = \int_{\Omega_1 \times \Omega_2} |u \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2) = \int_A |u(t)|^2 \mathrm{d}\mu_x(t) = \|u\|_{\mu_x, 2}^2. \tag{2.17}$$

Therefore

$$\mathcal{L}_1(\mathfrak{R}) = \mathrm{L}^2(A, \mathcal{B}_x, \mu_x). \tag{2.18}$$

Consider a sequence $\mathbb{N} \ni n \mapsto f_n \in \mathcal{L}_1(\mathfrak{R})$ satisfying the Cauchy condition (2.14). From (2.17) it follows that this sequence is also a Cauchy sequence in the space $\mathbf{L}^2(A, \mathcal{B}_x, \mu_x)$, which is complete. Thus

$$\|f_n - f\|_{\mu_x, 2} \to 0 \quad \text{as } n \to \infty \tag{2.19}$$

for some $f \in \mathrm{L}^2(A, \mathcal{B}_x, \mu_x)$. Combining (2.19) with (2.18) and (2.17) we conclude that $f \in \mathcal{L}_1(\mathfrak{R})$ and the condition (2.15) holds, which completes the proof. $\qquad \square$

Remark 2.4 The properties (2.13) yield the well known Schwarz inequality

$$|\langle u|v \rangle| \le \|u\| \|v\|, \quad u, v \in \mathcal{L}_1(\mathfrak{R}); \tag{2.20}$$

cf. e.g. [8, Thm 13.1.1]. From (2.11)–(2.13) and (2.20) it follows that the functional $\|\cdot\|$ is a pseudo-norm on the linear space $(\mathcal{L}_1(\mathfrak{R}), +, \cdot)$, i.e. for any $u, v \in \mathcal{L}_1(\mathfrak{R})$ and $\lambda \in \mathbb{C}$:

$$\begin{aligned} &\|u + v\| \le \|u\| + \|v\|, \\ &\quad \|\lambda u\| = \lambda \|u\|, \\ &\qquad \|u\| \ge 0 \quad \text{and} \quad \|u\| = 0 \implies \mu(\{(t_1, t_2) \in \Omega_1 \times \Omega_2 : u \circ x(t_1) \ne 0\}) = 0; \end{aligned} \tag{2.21}$$

cf. e.g. [8, Corollary 13.1.1], [2, Sect. 2.4.8]. Moreover, by Lemma 2.3 we know that the pseudo-norm $\|\cdot\|$ satisfies the completeness condition (2.14) $\implies$ (2.15). Thus the structure $(\mathcal{L}_1(\mathfrak{R}), +, \cdot, \|\cdot\|)$ is a pseudo-Banach space.

By (2.8)–(2.10) we see that for each $g \in \mathcal{L}_2(\mathfrak{R})$ the functional

$$\mathcal{L}_1(\mathfrak{R}) \ni u \to g^*(u) := \int_{\Omega_1 \times \Omega_2} u \circ x(t_1) \overline{g \circ y(t_2)} \mathrm{d}\mu(t_1, t_2) \tag{2.22}$$

is well defined.

Lemma 2.5 *The structure $(\mathcal{L}_2(\mathfrak{R}), +, \cdot)$ is a complex (resp. real in the case where $B = \mathbb{R}$) linear space and for each $g \in \mathcal{L}_2(\mathfrak{R})$ the functional g^* is bounded on $\mathrm{H}(\mathfrak{R})$ and the supremum norm of g^* satisfies the following inequality*

$$\begin{aligned} \sup\{&|g^*(f)| : f \in \mathcal{L}_1(\mathfrak{R}) \text{ and } \|f\| \le 1\} \\ &\le \left(\int_{\Omega_1 \times \Omega_2} |g \circ y(t_2)|^2 \mathrm{d}\mu(t_1, t_2) \right)^{1/2}. \end{aligned} \tag{2.23}$$

Proof From the inequality (2.16) and by (2.9) we see that for all $\lambda_1, \lambda_2 \in B$ and $g_1, g_2 \in \mathcal{L}_2(\mathfrak{R})$,

$$\int_{\Omega_1 \times \Omega_2} |(\lambda_1 g_1 + \lambda_2 g_2) \circ y(t_2)|^2 d\mu(t_1, t_2) \leq 2|\lambda_1|^2 \int_{\Omega_1 \times \Omega_2} |g_1 \circ y(t_2)|^2 d\mu(t_1, t_2)$$

$$+ 2|\lambda_2|^2 \int_{\Omega_1 \times \Omega_2} |g_2 \circ y(t_2)|^2 d\mu(t_1, t_2)$$

$$< +\infty.$$

Thus $\lambda_1 g_1 + \lambda_2 g_2 \in \mathcal{L}_2(\mathfrak{R})$ for all $\lambda_1, \lambda_2 \in B$ and $g_1, g_2 \in \mathcal{L}_2(\mathfrak{R})$. Hence $\mathcal{L}_2(\mathfrak{R})$ is a linear set with respect to standard operations of adding and multiplying functions. Then the structure $(\mathcal{L}_2(\mathfrak{R}), +, \cdot)$ is a linear space as a linear subspace of the linear space $(B \to B, +, \cdot)$.

The linearity of the functional g^* comes out from the algebraic properties of the Lebesgue integral. From Schwarz's integral inequality and from (2.22), (2.8), (2.9) and (2.12) it follows that for all $f \in \mathcal{L}_1(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$,

$$|g^*(f)| \leq \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) \overline{g \circ y(t_2)}| d\mu(t_1, t_2)$$

$$\leq \left(\int_{\Omega_1 \times \Omega_2} |f \circ x(t_1)|^2 d\mu(t_1, t_2) \right)^{1/2} \left(\int_{\Omega_1 \times \Omega_2} |g \circ y(t_2)|^2 d\mu(t_1, t_2) \right)^{1/2}$$

$$\leq \left(\int_{\Omega_1 \times \Omega_2} |g \circ y(t_2)|^2 d\mu(t_1, t_2) \right)^{1/2} \|f\|,$$

which yields (2.23).   □

3 Solution of the regression problem

Given a regression structure $\mathfrak{R} := (A, B, \delta; x, y)$ satisfying the properties I.1–I.3 we see that

$$\mathfrak{R}_g := (A, B, \delta; x, g \circ y)$$

is a regression structure for each function $g : B \to B$. From now on we shall study the regression problem for $\mathcal{F}$ with respect to $\mathfrak{R}_g$, where $\mathfrak{R}$ is an arbitrarily given regression structure satisfying the assumptions II.1–II.2, $g : B \to B$ is a fixed function and $\mathcal{F}$ is a linear functional model of $\mathfrak{R}$ with respect to the standard operations of adding and multiplying functions, i.e. $f + g \in \mathcal{F}$ and $\lambda f \in \mathcal{F}$ for any $f, g \in \mathcal{F}$ and $\lambda \in B$. If additionally $\mathcal{F} \subset \mathcal{L}_1(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$, then the regression problem means the extremal problem of determining all functions $f_0 \in \mathcal{F}$ which are minimizing the functional F_g defined by

$$F_g(f) = \delta(f \circ x, g \circ y) = \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) - g \circ y(t_2)|^2 d\mu(t_1, t_2), \ f \in \mathcal{F}.$$

$$(3.1)$$

We shall start our research with the following basic characterization of the regression functions.

Lemma 3.1 *If $\mathcal{F} \neq \emptyset$ is a linear set in $H(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$, then for every $f \in \mathcal{F}$ the following property holds*

$$f \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g) \Longleftrightarrow \langle h|f \rangle = g^*(h), \quad h \in \mathcal{F}. \tag{3.2}$$

Proof Fix $f, h \in \mathcal{F}$, $g \in \mathcal{L}_2(\mathfrak{R})$ and $\lambda \in B$. By (3.1) we have

$$
\begin{aligned}
F_g(f + \lambda h) &= \int_{\Omega_1 \times \Omega_2} |(f + \lambda h) \circ x(t_1) - g \circ y(t_2)|^2 \mathrm{d}\mu(t_1, t_2) \\
&= \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) - g \circ y(t_2) + \lambda h \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2) \\
&= \int_{\Omega_1 \times \Omega_2} |f \circ x(t_1) - g \circ y(t_2)|^2 \mathrm{d}\mu(t_1, t_2) \\
&\quad + 2 \int_{\Omega_1 \times \Omega_2} \text{Re}[(f \circ x(t_1) - g \circ y(t_2))\overline{\lambda h \circ x(t_1)}] \mathrm{d}\mu(t_1, t_2) \\
&\quad + |\lambda|^2 \int_{\Omega_1 \times \Omega_2} |h \circ x(t_1)|^2 \mathrm{d}\mu(t_1, t_2).
\end{aligned}
$$

Hence and by (3.1), (2.22) as well as by (2.11) and (2.12) we get

$$
\begin{aligned}
F_g(f + \lambda h) &= F_g(f) + |\lambda|^2 \|h\|^2 - 2\,\text{Re}\left[\overline{\lambda}\,\overline{g^*(h)}\right] \\
&\quad + 2\,\text{Re}\left[\overline{\lambda} \int_{\Omega_1 \times \Omega_2} f \circ x(t_1)\overline{h \circ x(t_1)} \mathrm{d}\mu(t_1, t_2)\right] \\
&= F_g(f) + |\lambda|^2 \|h\|^2 - 2\,\text{Re}\left[\overline{\lambda}\,\overline{g^*(h)}\right] + 2\,\text{Re}\left[\overline{\lambda}\,\langle f|h\rangle\right].
\end{aligned}
$$

By this and by (2.13),

$$
F_g(f + \lambda h) - F_g(f) = 2\,\text{Re}\left[\lambda(\langle h|f\rangle - g^*(h))\right] + |\lambda|^2 \|h\|^2,
$$
$$
f, h \in \mathcal{F}, \ \lambda \in B. \tag{3.3}
$$

Assume now that $f \in \mathcal{F}$ satisfies the right hand side condition in (3.2). Then setting $\lambda := 1$ we deduce from (3.3) that

$$F_g(f + h) - F_g(f) = \|h\|^2 \geq 0, \quad h \in \mathcal{F},$$

and so $f \in \text{Reg}(\mathcal{F}, \mathfrak{R})$. Conversely, suppose that $f \in \text{Reg}(\mathcal{F}, \mathfrak{R})$. Then (3.3) yields

$$2\,\text{Re}[\lambda(\langle h|f\rangle - g^*(h))] + |\lambda|^2 \|h\|^2 \geq 0, \quad h \in \mathcal{F}, \ \lambda \in B. \tag{3.4}$$

Replacing h by $(-h)$ in (3.4) we get

$$-2\,\text{Re}[\lambda(\langle h|f\rangle - g^*(h))] + |\lambda|^2 \|h\|^2 \geq 0, \quad h \in \mathcal{F}, \ \lambda \in B. \tag{3.5}$$

Combining (3.4) and (3.5) we have

$$-\frac{1}{2}|\lambda|^2\|h\|^2 \le \mathrm{Re}[\lambda(\langle h|f\rangle - g^*(h))] \le \frac{1}{2}|\lambda|^2\|h\|^2, \quad h \in \mathcal{F}, \ \lambda \in B,$$

and consequently,

$$-\frac{1}{2}|\lambda|\|h\|^2 \le \mathrm{Re}[\mathrm{e}^{\mathrm{i}\alpha(\lambda)}(\langle h|f\rangle - g^*(h))] \le \frac{1}{2}|\lambda|\|h\|^2, \quad h \in \mathcal{F}, \ \lambda \in B\backslash\{0\}, \quad (3.6)$$

where $\alpha(\lambda) \in [0; 2\pi)$ is the unique number satisfying the equality $\lambda = |\lambda|\mathrm{e}^{\mathrm{i}\alpha(\lambda)}$. Thus (3.6) yields, in the limiting case as $|\lambda| \to 0$, the following equality

$$\mathrm{Re}[\mathrm{e}^{\mathrm{i}\alpha}(\langle h|f\rangle - g^*(h))] = 0, \quad h \in \mathcal{F}, \ \alpha \in \mathbb{R}.$$

Choosing appropriately α we can see that

$$\langle h|f\rangle - g^*(h) = 0, \quad h \in \mathcal{F},$$

which completes the proof.   □

By the properties of a pseudo-norm we see that the following set

$$\Theta := \{h \in \mathcal{L}_1(\mathfrak{R}) : \|h\| = 0\}$$

is linear. We call it the *null set* of H($\mathfrak{R}$). As a matter of fact the set Θ is the closed ball with radius 0 and center at the zero function θ, defined by $\theta(t) := 0$ for $t \in A$.

We may extend the standard operations of adding and multiplying functions by a constant to any sets $F_1, F_2 \subset (A \to B)$ as follows:

$$F_1 + F_2 := \{f_1 + f_2 : f_1 \in F_1, \ f_2 \in F_2\};$$
$$\lambda \cdot F_1 := \{\lambda f_1 : f_1 \in F_1\}, \quad \lambda \in B;$$
$$f + F_1 := \{f\} + F_1 \text{ and } F_1 + f := F_1 + \{f\}, \quad f \in (A \to B).$$

Corollary 3.2 *If $\mathcal{F} \ne \emptyset$ is a linear set in H($\mathfrak{R}$) and $g \in \mathcal{L}_2(\mathfrak{R})$, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g). \tag{3.7}$$

If additionally $\mathcal{F} \subset \Theta$, then $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathcal{F}$.

Proof Fix $f, h \in \mathcal{L}_1(\mathfrak{R})$. If $\|h\| = 0$, then from the Schwarz's inequality (2.20) and Lemma 2.5 it follows that

$$|\langle h|f\rangle| \le \|h\|\|f\| = 0$$

and

$$|g^*(h)| \leq \left(\int_{\Omega_1 \times \Omega_2} |g \circ y(t_2)|^2 \mathrm{d}\mu(t_1, t_2) \right)^{1/2} \cdot \|h\| = 0.$$

Hence

$$\langle h|f \rangle = 0 = g^*(h), \quad f \in \mathcal{L}_1(\mathfrak{R}_g), \quad h \in \Theta. \tag{3.8}$$

Assume that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$ and $h \in \Theta + \mathcal{F}$. Then $h = h_0 + h_1$ for some $h_0 \in \Theta$ and $h_1 \in \mathcal{F}$. Applying now Lemma 3.1 and (3.8) we see that

$$\langle h|f \rangle = \langle h_0|f \rangle + \langle h_1|f \rangle = 0 + g^*(h_1) = g^*(h_0) + g^*(h_1) = g^*(h), \quad h \in \Theta + \mathcal{F}.$$

By definition, $f \in \mathcal{F} \subset \Theta + \mathcal{F}$. Applying Lemma 3.1 once more, we obtain

$$f \in \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g),$$

and consequently

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \subset \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g). \tag{3.9}$$

Conversely, assume now that

$$f \in \mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g)$$

and $h \in \mathcal{F}$. Since $h \in \Theta + \mathcal{F}$, we conclude from Lemma 3.1 that

$$\langle h|f \rangle = g^*(h), \quad h \in \mathcal{F}.$$

Then Lemma 3.1 says that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$, and so

$$\mathcal{F} \cap \mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g) \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g).$$

Combining this with the inclusion (3.9) we obtain the equality (3.7).

Since $\Theta \subset \mathcal{L}_1(\mathfrak{R})$, the equalities in (3.8) hold for all $f, h \in \Theta$. Then Lemma 3.1 yields

$$\mathrm{Reg}(\Theta, \mathfrak{R}_g) = \Theta.$$

If now $\mathcal{F} \subset \Theta$, then the equality (3.7) takes the form $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathcal{F}$, which proves the theorem. $\qquad \square$

Given a set $S \subset \mathcal{L}_1(\mathfrak{R})$ we denote by $\mathrm{lin}(S)$ the family of all linear combinations $\sum_{k=1}^{n} \lambda_k v_k$ where $n \in \mathbb{N}$, $\mathbb{Z}_{1,n} \ni k \mapsto \lambda_k \in B$ and $\mathbb{Z}_{1,n} \ni k \mapsto v_k \in S$ provided $S \neq \emptyset$, and we write $\mathrm{lin}(S) := \{\theta\}$ otherwise. Obviously, $\mathrm{lin}(S)$ is a linear set and

$$S \subset \mathrm{lin}(S) = \mathrm{lin}(\mathrm{lin}(S)) \subset V \tag{3.10}$$

for every linear set $V \subset \mathcal{L}_1(\mathfrak{R})$ such that $S \subset V$. Thus $\mathrm{lin}(S)$ is the smallest linear subset of $\mathcal{L}_1(\mathfrak{R})$ containing S and

$$\mathrm{lin}(S) = S \Longleftrightarrow S \text{ is a linear set.} \tag{3.11}$$

Moreover, for any subsets S_1 and S_2 of $\mathcal{L}_1(\mathfrak{R})$,

$$\mathrm{lin}(S_1 \cup S_2) = \mathrm{lin}(S_1) + \mathrm{lin}(S_2). \tag{3.12}$$

Given a nonempty set $S \subset \mathcal{L}_1(\mathfrak{R})$ we denote by $S^\perp$ the orthogonal complement of S in the space $\mathrm{H}(\mathfrak{R})$, i.e.

$$S^\perp := \{f \in \mathcal{L}_1(\mathfrak{R}) : \langle f|h \rangle = 0 \text{ for } h \in S\}. \tag{3.13}$$

It is easy to check that for any nonempty subset S of $\mathcal{L}_1(\mathfrak{R})$:

$$\Theta \subset S^\perp \text{ and } S^\perp \text{ is a linear closed set in the space } \mathrm{H}(\mathfrak{R}); \tag{3.14a}$$

$$\mathcal{L}_1(\mathfrak{R}) = \mathrm{cl}(\mathrm{lin}(S)) + S^\perp \text{ and } \mathrm{cl}(\mathrm{lin}(S)) = (S^\perp)^\perp; \tag{3.14b}$$

$$\text{if } S \text{ is a linear set, then } \mathcal{L}_1(\mathfrak{R}) = \mathrm{cl}(S) + S^\perp \text{ and } \mathrm{cl}(S) = (S^\perp)^\perp, \tag{3.14c}$$

where $\mathrm{cl}(V)$ stands for the closure of a set $V \subset \mathcal{L}_1(\mathfrak{R})$ in the space $\mathrm{H}(\mathfrak{R})$. As a matter of fact the proofs of these properties are slight modifications of the well known properties, that hold in a Hilbert space; cf. e.g. [8, Sect. 13.4]. The following theorem is motivated by Lemma 3.1 and the well known representation of a linear and continuous functional in a Hilbert space by Riesz; cf. e.g. [8, Thm. 13.4.2].

Theorem 3.3 *If $\mathcal{F} \neq \emptyset$ is a linear set in $\mathrm{H}(\mathfrak{R})$, $g \in \mathcal{L}_2(\mathfrak{R})$ and $S := (g^*)^{-1}(0)$, then:*

(i) *if $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$, then $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + f$ for each $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$;*
(ii) *if $(\mathcal{F} \cap S)^\perp \cap \mathcal{F} \backslash \Theta \neq \emptyset$, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + \frac{\overline{g^*(f)}}{\|f\|^2} f, \quad f \in (\mathcal{F} \cap S)^\perp \cap \mathcal{F} \backslash \Theta. \tag{3.15}$$

(iii) $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta \cap \mathcal{F}$ *iff* $\mathcal{F} \subset S$;
(iv) $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$ *iff* $\mathcal{F} \subset (\mathcal{F} \cap S) + (\mathcal{F} \cap S)^\perp$.

Proof Assume that $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$. Then $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$ for some $f \in \mathcal{F}$, and by Lemma 3.1,

$$\langle h|f \rangle = g^*(h), \quad h \in \mathcal{F}. \tag{3.16}$$

From Lemma 3.1 it also follows that each $f' \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ satisfies the equality (3.16) with f replaced by f', and setting $h := f - f'$ we obtain

$$\|h\|^2 = \langle h|f - f' \rangle = \langle h|f \rangle - \langle h|f' \rangle = g^*(h) - g^*(h) = 0.$$

Thus $f' \in (\Theta \cap \mathcal{F}) + f$, and then

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \subset (\Theta \cap \mathcal{F}) + f.$$

Conversely, if $f' \in (\Theta \cap \mathcal{F}) + f$, then by (3.16) and by (2.20) we see that for every $h \in \mathcal{F}$,

$$|\langle h | f' \rangle - \langle h | f \rangle| = |\langle h | f' - f \rangle| \le \|h\| \|f' - f\| = 0.$$

Hence and by (3.16),

$$\langle h | f' \rangle = \langle h | f \rangle = g^*(h), \quad h \in \mathcal{F}.$$

Applying Lemma 3.1 once again we see that $f' \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$, and so

$$(\Theta \cap \mathcal{F}) + f \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g).$$

Both the inclusions yield the equality in (i).

Assume now that

$$(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta \ne \emptyset$$

and consider any fixed $f \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta$. Then $\|f\| > 0$ and

$$\frac{\overline{g^*(f)}}{\|f\|^2} f \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F}. \tag{3.17}$$

If $g^*(f) = 0$, then $f \in \mathcal{F} \cap S$, which gives $\|f\|^2 = \langle f | f \rangle = 0$. This means that $f \in \Theta$, which contradicts our assumption. Therefore $g^*(f) \ne 0$, and consequently

$$h - \frac{g^*(h)}{g^*(f)} f \in \mathcal{F} \cap S, \quad h \in \mathcal{F}. \tag{3.18}$$

Combining (3.17) with (3.18) we have

$$\left\langle h \left| \frac{\overline{g^*(f)}}{\|f\|^2} f \right\rangle = \left\langle h - \frac{g^*(h)}{g^*(f)} f \left| \frac{\overline{g^*(f)}}{\|f\|^2} f \right\rangle + \left\langle \frac{g^*(h)}{g^*(f)} f \left| \frac{\overline{g^*(f)}}{\|f\|^2} f \right\rangle \right.$$

$$= 0 + \left\langle \frac{g^*(h)}{g^*(f)} f \left| \frac{\overline{g^*(f)}}{\|f\|^2} f \right\rangle = \frac{g^*(h)}{g^*(f)} \overline{\left(\frac{\overline{g^*(f)}}{\|f\|^2} \right)} \langle f | f \rangle = g^*(h), h \in \mathcal{F}.$$

Then Lemma 3.1 shows that

$$\frac{\overline{g^*(f)}}{\|f\|^2} f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g),$$

which together with the property (i) yields (3.15). This proves the property (ii).

If $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta \cap \mathcal{F}$, then

$$\theta \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g),$$

and applying Lemma 3.1 we see that

$$g^*(h) = \langle h | \theta \rangle = 0, \quad h \in \mathcal{F},$$

which means that $\mathcal{F} \subset S$. Conversely, if $\mathcal{F} \subset S$, then

$$\langle h | \theta \rangle = 0 = g^*(h), \quad h \in \mathcal{F},$$

which shows, by Lemma 3.1, that

$$\theta \in \text{Reg}(\mathfrak{R}_g).$$

Hence and by the property (i),

$$\text{Reg}(\mathfrak{R}_g) = \Theta \cap \mathcal{F},$$

which yields the property (iii).

Assume now that $f \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g)$ for some $f \in \mathcal{F}$. From (3.16) it follows that

$$\langle h | f \rangle = 0 \quad \text{for } h \in \mathcal{F} \cap S,$$

and so

$$f \in \mathcal{F} \cap (\mathcal{F} \cap S)^{\perp}. \tag{3.19}$$

Consider first the case where $f \notin \Theta$ and suppose that $g^*(f) = 0$. Then $f \in \mathcal{F} \cap S$, and by (3.19), $\|f\|^2 = \langle f | f \rangle = 0$. Hence $f \in \Theta$, which contradicts our assumption. Thus $g^*(f) \neq 0$, and so for each $h \in \mathcal{F}$,

$$h_S := h - \frac{g^*(h)}{g^*(f)} f \in \mathcal{F} \cap S \quad \text{and} \quad h - h_S = \frac{g^*(h)}{g^*(f)} f \in (\mathcal{F} \cap S)^{\perp}, \tag{3.20}$$

which leads to

$$h = h_S + (h - h_S) \in (\mathcal{F} \cap S) + (\mathcal{F} \cap S)^{\perp}.$$

Therefore

$$\mathcal{F} \subset (\mathcal{F} \cap S) + (\mathcal{F} \cap S)^{\perp}, \tag{3.21}$$

provided $f \notin \Theta$. If $f \in \Theta$, then from the properties (i) and (iii) it follows that

$$\mathcal{F} \subset \mathcal{F} \cap S,$$

and obviously the inclusion (3.21) holds. This way the property (iv) was proved in the direction ($\Rightarrow$). Conversely, assume that the inclusion (3.21) holds. If

$$(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta \neq \emptyset,$$

then by the property (ii), $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$. Therefore we may confine ourselves to the case where

$$(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta = \emptyset.$$

Fix now $h \in \mathcal{F}$. By (3.21), $h = h_1 + h_2$ for some $h_1 \in \mathcal{F} \cap S$ and $h_2 \in (\mathcal{F} \cap S)^{\perp}$. Then

$$h_2 \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \subset \Theta.$$

By (3.8) we see that $\Theta \subset S$, which implies

$$h = h_1 + h_2 \in (\mathcal{F} \cap S) + \Theta \subset S.$$

Hence $\mathcal{F} \subset S$, and by the property (iii) we see that

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta \cap \mathcal{F} \neq \emptyset.$$

Thus we have showed that the inclusion (3.21) implies that

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset,$$

which completes the proof of the property (iv). $\square$

Corollary 3.4 *If $\mathcal{F} \neq \emptyset$ is a closed and linear set in $\text{H}(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$, then $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$ and $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta + f$ for each $f \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g)$. Moreover, if $\mathcal{F} \subset S := (g^*)^{-1}(0)$, then $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta$; otherwise $(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta \neq \emptyset$ and*

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta + \frac{\overline{g^*(f)}}{\|f\|^2} f, \quad f \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta. \tag{3.22}$$

Proof Since $\mathcal{F}$ is a closed set in $\text{H}(\mathfrak{R})$, we have

$$\Theta = \text{cl}(\{\theta\}) \subset \text{cl}(\mathcal{F}) = \mathcal{F},$$

and then

$$\Theta \cap \mathcal{F} = \Theta. \tag{3.23}$$

Since $g \in \mathcal{L}_2(\mathfrak{R})$, it follows from Lemma 2.5 that g^* is a continuous functional on $H(\mathfrak{R})$. Hence S is also a closed set in $H(\mathfrak{R})$, and so $\mathcal{F} \cap S$ is a closed set in $H(\mathfrak{R})$. Then each $h \in \mathcal{F}$ has an orthogonal projection h_S onto $\mathcal{F} \cap S$, i.e.

$$h_S \in \mathcal{F} \cap S \quad \text{and} \quad \langle h - h_S | g \rangle = 0, \quad g \in \mathcal{F} \cap S; \tag{3.24}$$

cf. [8, Sect. 13.3] and [10]. Hence

$$h = h_S + (h - h_S) \in (\mathcal{F} \cap S) + (\mathcal{F} \cap S)^{\perp},$$

which yields the inclusion (3.21). Applying now (iv) of Theorem 3.3 we conclude that $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$. If $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$, then from (i) of Theorem 3.3 and the equality (3.23) it follows that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + f = \Theta + f.$$

If $\mathcal{F} \subset S$, then from (iii) of Theorem 3.3 and the equality (3.23) we see that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta \cap \mathcal{F} = \Theta.$$

Otherwise there exists

$$h \in \mathcal{F} \backslash S = \mathcal{F} \backslash (\mathcal{F} \cap S)$$

and the property (3.24) holds. From (3.8) it follows that

$$\Theta \subset S.$$

If $h - h_S \in \Theta$, then by (3.24),

$$h = h_S + (h - h_S) \in (\mathcal{F} \cap S) + \Theta = \mathcal{F} \cap S,$$

which is impossible. Therefore $h - h_S \notin \Theta$, which together with (3.24) means that

$$h - h_S \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \backslash \Theta.$$

Then the condition (ii) of Theorem 3.3 and the equality (3.23) imply the equality (3.22), which completes the proof. $\square$

We end this section with an important application of Corollaries 3.2 and 3.4.

Corollary 3.5 *If $\mathcal{F} \neq \emptyset$ is a finite dimensional linear set in $H(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$, then*

$$\mathrm{cl}(\mathcal{F}) = \Theta + \mathcal{F} \quad \text{and} \quad \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset.$$

Proof Given a linear set $\mathcal{F}$ in $H(\mathfrak{R})$ and $f \in cl(\mathcal{F})$ there exists a sequence $\mathbb{N} \ni n \mapsto f_n \in \mathcal{F}$ such that

$$\|f - f_n\| \to 0 \quad \text{as } n \to \infty. \tag{3.25}$$

Assume first that $\dim(\mathcal{F}) = 1$, where $\dim(X)$ stands for the dimension of a linear set X. Then $\mathcal{F} = \lin(\{h\})$ for certain $h \in \mathcal{L}_1(\mathfrak{R})$. If $\|h\| = 0$, then obviously $\mathcal{F} \subset \Theta$, which gives $cl(\mathcal{F}) \subset \Theta + \mathcal{F}$. Otherwise $\|h\| > 0$, and there exists a sequence $\mathbb{N} \ni n \mapsto \lambda_n \in B$ such that $f_n = \lambda_n h$ for $n \in \mathbb{N}$. By (3.25) we have

$$|\lambda_n - \lambda_m| = \frac{1}{\|h\|}\|f_n - f_m\| \to 0 \quad \text{as } n, m \to \infty.$$

Hence $|\lambda - \lambda_n| \to 0$ as $n \to \infty$ for certain $\lambda \in B$, which together with (3.25) gives

$$\|f - \lambda h\| \leq \|f - f_n\| + \|f_n - \lambda h\| \to 0 \quad \text{as } n \to \infty.$$

We conclude from this that

$$(f - \lambda h) \in \Theta,$$

hence that

$$f \in \Theta + \lambda h \subset \Theta + \mathcal{F},$$

and finally that

$$cl(\mathcal{F}) \subset \Theta + \mathcal{F}. \tag{3.26}$$

Fix $n \in \mathbb{N}$ and suppose that (3.26) holds provided $\dim(\mathcal{F}) \leq n$. Assume now that $\dim(\mathcal{F}) \leq n+1$. If $\mathcal{F} \subset \Theta$, then the inclusion (3.26) evidently holds. Otherwise there exists $h \in \mathcal{F} \backslash \Theta$ and we may consider the following sequences

$$\mathbb{N} \ni n \mapsto f_n' := \frac{\langle f_n | h \rangle}{\|h\|^2} h \quad \text{and} \quad \mathbb{N} \ni n \mapsto f_n'' := f_n - f_n'.$$

Setting

$$\mathcal{F}' := \lin(\{h\}) \quad \text{and} \quad \mathcal{F}'' := \mathcal{F} \cap (\mathcal{F}')^\perp$$

we can see that

$$f_n' \in \mathcal{F}' \quad \text{and} \quad f_n'' \in \mathcal{F}''$$

for $n \in \mathbb{N}$. Let

$$f' := \langle f | h \rangle \|h\|^{-2} h \quad \text{and} \quad f'' := f - f'.$$

From (3.25) and Schwarz's inequality (2.20) it follows that

$$\|f' - f_n'\| = \frac{|\langle f - f_n | h \rangle|}{\|h\|} \leq \|f - f_n\| \to 0 \quad \text{as } n \to \infty,$$

and consequently

$$\|f'' - f_n''\| \leq \|f - f_n\| + \|f' - f_n'\| \to 0 \quad \text{as } n \to \infty.$$

Therefore

$$f' \in \mathrm{cl}(\mathcal{F}') \quad \text{and} \quad f'' \in \mathrm{cl}(\mathcal{F}''),$$

which gives

$$f = f' + f'' \in \mathrm{cl}(\mathcal{F}') + \mathrm{cl}(\mathcal{F}'').$$

Then applying the first part of the proof we deduce that

$$\mathrm{cl}(\mathcal{F}) \subset \mathrm{cl}(\mathcal{F}') + \mathrm{cl}(\mathcal{F}'') \subset \Theta + \mathcal{F}' + \mathrm{cl}(\mathcal{F}''), \tag{3.27}$$

because $\dim(\mathcal{F}') = 1$. By definition, $\mathcal{F}' \cap \Theta = \{\theta\}$, and so

$$\mathcal{F}' \cap \mathcal{F}'' = \{\theta\}.$$

Hence

$$\dim(\mathcal{F}'') \leq \dim(\mathcal{F}) - \dim(\mathcal{F}') \leq n,$$

and by the assumption,

$$\mathrm{cl}(\mathcal{F}'') \subset \Theta + \mathcal{F}''.$$

Combining this with (3.27) we obtain

$$\mathrm{cl}(\mathcal{F}) \subset \Theta + \mathcal{F}' + \Theta + \mathcal{F}'' = \Theta + \mathcal{F}' + \mathcal{F}'' \subset \Theta + \mathcal{F}.$$

Applying now the mathematical induction we conclude that the inclusion (3.26) holds for every finite dimensional linear set $\mathcal{F}$ in $\mathrm{H}(\mathfrak{R})$. The inverse inclusion is obvious, because $\Theta = \mathrm{cl}(\{\theta\}) \subset \mathrm{cl}(\mathcal{F})$ and $\mathcal{F} \subset \mathrm{cl}(\mathcal{F})$. Therefore

$$\mathrm{cl}(\mathcal{F}) = \Theta + \mathcal{F}.$$

By Corollary 3.4,

$$\mathrm{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g) = \mathrm{Reg}(\mathrm{cl}(\mathcal{F}), \mathfrak{R}_g) \neq \emptyset,$$

and so

$$f \in \text{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g) \quad \text{for certain } f \in \Theta + \mathcal{F}.$$

Applying now Corollary 3.4 once more we can see that

$$\text{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g) = \Theta + f.$$

Since $f = h + f_0$ for some $h \in \Theta$ and $f_0 \in \mathcal{F}$, we conclude from Corollary 3.2 that

$$f_0 = -h + f \in \mathcal{F} \cap (\Theta + f) = \mathcal{F} \cap \text{Reg}(\Theta + \mathcal{F}, \mathfrak{R}_g) = \text{Reg}(\mathcal{F}, \mathfrak{R}_g),$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

4 Orthogonal decompositions of regression functions

In this section we establish various results dealing with the orthogonality properties of regression functions. Given $f, g \in \mathcal{L}_1(\mathfrak{R})$ we will write

$$f \perp g \quad \text{iff } \langle f | g \rangle = 0.$$

We extend this relation to any nonempty sets $F, G \subset \mathcal{L}_1(\mathfrak{R})$ in the following manner:

$$F \perp G \quad \text{iff} \quad \langle f | g \rangle = 0, \quad f \in F, \ g \in G. \tag{4.1}$$

In particular, we write $f \perp G$ iff $\{f\} \perp G$ and $F \perp g$ iff $F \perp \{g\}$. Given $p, q \in \mathbb{Z}$, $p \leq q$, and a sequence $\mathbb{Z}_{p,q} \ni k \mapsto \mathcal{F}_k$ of nonempty sets in the space $\text{H}(\mathfrak{R})$, we write $\sum_{k=p}^{q} \mathcal{F}_k$ for the set of all $\sum_{k=p}^{q} f_k$, where $\mathbb{Z}_{p,q} \ni k \mapsto f_k \in \mathcal{F}_k$. Obviously,

$$\sum_{k=1}^{2} \mathcal{F}_k = \mathcal{F}_1 + \mathcal{F}_2.$$

In the sequel we will use the following auxiliary properties that hold for any nonempty subsets S_1 and S_2 of $\mathcal{L}_1(\mathfrak{R})$ such that $S_1 \perp S_2$:

$$\text{cl}(S_1) \perp \text{cl}(S_2); \tag{4.2a}$$
$$\text{cl}(S_1 + S_2) = \text{cl}(S_1) + \text{cl}(S_2); \tag{4.2b}$$
$$S_1 \cap S_2 \subset \Theta; \tag{4.2c}$$
$$\text{lin}(S_1) \perp \text{lin}(S_2); \tag{4.2d}$$
$$h_1 + h_2 = h_1' + h_2' \implies h_1 - h_1', h_2 - h_2' \in \Theta, \quad h_1, h_1' \in S_1, \ h_2, h_2' \in S_2; \tag{4.2e}$$
$$\|h_1 + h_2\|^2 = \|h_1\|^2 + \|h_2\|^2, \quad h_1 \in S_1, \ h_2 \in S_2. \tag{4.2f}$$

The proofs of these properties are slight modifications of the well known properties, that hold in a Hilbert space; cf. e.g. [8, Corollary 13.1.2, Sect. 13.4].

For any $p \in \mathbb{Z}$ and a sequence $\mathbb{Z}_{p,\infty} \ni k \mapsto \mathcal{F}_k$ of nonempty sets in the space $H(\mathfrak{R})$, we denote by $\sum_{k=p}^{\infty} \mathcal{F}_k$ the set of all $f \in \mathcal{L}_1(\mathfrak{R})$ for which there exists a sequence $\mathbb{Z}_{p,\infty} \ni k \mapsto f_k \in \mathcal{F}_k$ such that

$$\left\| f - \sum_{k=p}^{n} f_k \right\| \to 0 \quad \text{as } n \to \infty.$$

In particular, we set

$$\sum_{k=p}^{\infty} f_k := \sum_{k=p}^{\infty} \{f_k\}$$

for any sequence $\mathbb{Z}_{p,\infty} \ni k \mapsto f_k \in \mathcal{L}_1(\mathfrak{R})$.

We first prove the following two auxiliary lemmas.

Lemma 4.1 *Given $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_{p,\infty} \cup \{\infty\}$ let $\mathbb{Z}_{p,q} \ni k \to \mathcal{F}_k$ be a sequence such that $\mathcal{F}_k$ is a nonempty linear set in the space $H(\mathfrak{R})$ for every $k \in \mathbb{Z}_{p,q}$. If*

$$\mathcal{F}_k \perp \mathcal{F}_l, \quad k, l \in \mathbb{Z}_{p,q}, \, k \neq l, \tag{4.3}$$

then $\sum_{k=p}^{q} \mathrm{cl}(\mathcal{F}_k)$ is a linear and closed set in $H(\mathfrak{R})$ and

$$\mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{k=p}^{q} \mathcal{F}_k \right) \right) = \mathrm{cl}\left(\sum_{k=p}^{q} \mathcal{F}_k \right) = \sum_{k=p}^{q} \mathrm{cl}(\mathcal{F}_k). \tag{4.4}$$

Proof If $q \in \mathbb{N}$, then by definition,

$$\mathrm{lin}\left(\bigcup_{k=p}^{q} \mathcal{F}_k \right) = \sum_{k=p}^{q} \mathcal{F}_k,$$

and so the first equality in (4.4) holds. Therefore we may assume that $q = \infty$. Let now

$$f \in \mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k \right)$$

be given. By definition,

$$f = \sum_{k=p}^{n} \lambda_k f_k$$

for some $n \in \mathbb{Z}_{p,\infty}$ and sequences $\mathbb{Z}_{p,n} \ni k \mapsto f_k \in \mathcal{F}_k$ and $\mathbb{Z}_{p,n} \ni k \mapsto \lambda_k \in B$. Hence $f \in \sum_{k=p}^{n} \mathcal{F}_k$, and consequently,

$$\mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right) \subset \sum_{k=p}^{\infty} \mathcal{F}_k,$$

which gives

$$\mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right)\right) \subset \mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right). \tag{4.5}$$

Conversely, assume that $f \in \sum_{k=p}^{\infty} \mathcal{F}_k$. Then

$$\left\| f - \sum_{k=p}^{n} f_k \right\| \to 0 \quad \text{as } n \to \infty$$

for certain sequence $\mathbb{Z}_{p,\infty} \ni k \mapsto f_k \in \mathcal{F}_k$. By definition,

$$\sum_{k=p}^{n} f_k \in \mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right)$$

for every $n \in \mathbb{Z}_{p,\infty}$. Thus

$$f \in \mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right)\right),$$

and consequently

$$\mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right) \subset \mathrm{cl}\left(\mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right)\right)\right) = \mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{k=p}^{\infty} \mathcal{F}_k\right)\right),$$

which together with the inclusion (4.5) leads to the first equality in (4.4).

It remains to prove the second equality in (4.4). From the property (4.2b) it follows that

$$\mathrm{cl}\left(\sum_{k=p}^{q} \mathcal{F}_k\right) = \sum_{k=p}^{q} \mathrm{cl}(\mathcal{F}_k), \quad q \in \mathbb{Z}_{p,\infty}, \tag{4.6}$$

and hence the second equality in (4.4) holds for every $q \in \mathbb{Z}_{p,\infty}$.

Consider now the infinite case $q = \infty$. Assume that

$$f \in \mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right).$$

Then there exists a sequence

$$\mathbb{Z}_{p,\infty} \ni n \to f_n' \in \sum_{k=p}^{\infty} \mathcal{F}_k$$

such that

$$\|f - f_n'\| \to 0 \quad \text{as } n \to \infty. \tag{4.7}$$

From (3.14c) and (3.14a) we deduce that each set $\mathrm{cl}(\mathcal{F}_k)$, as $k \in \mathbb{Z}_{p,\infty}$, is linear and closed in $\mathrm{H}(\mathfrak{R})$, and hence there exists an orthogonal projection f_k of f onto the set $\mathrm{cl}(\mathcal{F}_k)$. Therefore

$$f - f_k \perp \mathrm{cl}(\mathcal{F}_k) \quad \text{and} \quad f_k \in \mathrm{cl}(\mathcal{F}_k), \quad k \in \mathbb{Z}_{p,\infty}. \tag{4.8}$$

From (4.3), (4.2a) and (4.8) it follows that for all $m, l \in \mathbb{Z}_{p,\infty}$ and $h \in \mathrm{cl}(\mathcal{F}_l)$,

$$\left\langle f - \sum_{k=p}^{m} f_k \,\middle|\, h \right\rangle = \left\langle f - f_l - \sum_{k \in \mathbb{Z}_{1,m}\setminus\{l\}}^{m} f_k \,\middle|\, h \right\rangle = \langle f - f_l | h \rangle - \sum_{k \in \mathbb{Z}_{1,m}\setminus\{l\}}^{m} \langle f_k | h \rangle = 0.$$

Hence

$$f - \sum_{k=p}^{m} f_k \perp \mathrm{cl}(\mathcal{F}_l), \quad m, l \in \mathbb{Z}_{p,\infty}. \tag{4.9}$$

Given $m \in \mathbb{Z}_{p,\infty}$ and $h \in \sum_{k=p}^{m} \mathrm{cl}(\mathcal{F}_k)$ there exists a sequence

$$\mathbb{Z}_{1,m} \ni k \to h_k \in \mathrm{cl}(\mathcal{F}_k)$$

such that $h = \sum_{k=p}^{m} h_k$. From this and (4.9) we see that

$$\left\langle f - \sum_{k=p}^{m} f_k \,\middle|\, h \right\rangle = \left\langle f - \sum_{k=p}^{m} f_k \,\middle|\, \sum_{l=p}^{m} h_l \right\rangle = \sum_{l=p}^{m} \left\langle f - \sum_{k=p}^{m} f_k \,\middle|\, h_l \right\rangle = 0.$$

This together with (4.8) leads to

$$f - \sum_{k=p}^{m} f_k \perp \sum_{k=p}^{m} \mathrm{cl}(\mathcal{F}_k) \quad \text{and} \quad \sum_{k=p}^{m} f_k \in \sum_{k=p}^{m} \mathrm{cl}(\mathcal{F}_k), \quad m \in \mathbb{Z}_{p,\infty}. \tag{4.10}$$

Given $\varepsilon > 0$ we conclude from (4.7) that $\| f - f_n' \| < \varepsilon/2$ for some $n \in \mathbb{Z}_{p,\infty}$. Since $f_n' \in \sum_{k=p}^{\infty} \mathcal{F}_k$, there exists a sequence $\mathbb{Z}_{p,\infty} \ni k \to f_k'' \in \mathcal{F}_k$ such that

$$\left\| f_n' - \sum_{k=p}^{m} f_k'' \right\| \to 0 \quad \text{as } m \to \infty,$$

and so for certain $m_\varepsilon \in \mathbb{Z}_{p,\infty}$,

$$\left\| f_n' - \sum_{k=p}^{m} f_k'' \right\| < \frac{\varepsilon}{2}, \quad m \in \mathbb{Z}_{m_\varepsilon,\infty}. \tag{4.11}$$

From (4.10) it follows that $f - \sum_{k=p}^{m} f_k$ is an orthogonal projection of f onto the set $\sum_{k=p}^{m} \mathrm{cl}(\mathcal{F}_k)$ for every $m \in \mathbb{Z}_{p,\infty}$. Since

$$\sum_{k=p}^{m} f_k'' \in \sum_{k=p}^{m} \mathcal{F}_k,$$

we conclude from (4.2f) that

$$\left\| f - \sum_{k=p}^{m} f_k \right\| \leq \left\| f - \sum_{k=p}^{m} f_k'' \right\|, \quad m \in \mathbb{Z}_{p,\infty}.$$

Combining this with (4.11) we see that

$$\left\| f - \sum_{k=p}^{m} f_k \right\| \leq \| f - f_n' \| + \left\| f_n' - \sum_{k=p}^{m} f_k'' \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad m \in \mathbb{Z}_{m_\varepsilon,\infty},$$

which means that $f \in \sum_{k=p}^{\infty} \mathrm{cl}(\mathcal{F}_k)$. Thus

$$\mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k \right) \subset \sum_{k=p}^{\infty} \mathrm{cl}(\mathcal{F}_k). \tag{4.12}$$

Conversely, assume that

$$f \in \sum_{k=p}^{\infty} \mathrm{cl}(\mathcal{F}_k).$$

Then

$$\left\| f - \sum_{k=p}^{n} f_k \right\| \to 0 \quad \text{as } n \to \infty$$

for a sequence $\mathbb{Z}_{p,\infty} \ni k \mapsto f_k \in \mathrm{cl}(\mathcal{F}_k)$. From (4.6) it follows that

$$\sum_{k=p}^{n} f_k \in \sum_{k=p}^{n} \mathrm{cl}(\mathcal{F}_k) = \mathrm{cl}\left(\sum_{k=p}^{n} \mathcal{F}_k\right) \subset \mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right), \quad n \in \mathbb{Z}_{p,\infty}.$$

Thus

$$f \in \mathrm{cl}\left(\mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right)\right) = \mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right),$$

and so

$$\sum_{k=p}^{\infty} \mathrm{cl}(\mathcal{F}_k) \subset \mathrm{cl}\left(\sum_{k=p}^{\infty} \mathcal{F}_k\right),$$

which together with the inclusion (4.12) leads to the second equality in (4.4).

Since $\mathrm{lin}\left(\bigcup_{k=p}^{q} \mathcal{F}_k\right)$ is a linear set in $\mathrm{H}(\mathfrak{R})$, we deduce from (3.14c), (3.14a) and (4.4) that $\sum_{k=p}^{q} \mathrm{cl}(\mathcal{F}_k)$ is a linear and closed set in $\mathrm{H}(\mathfrak{R})$ for every $q \in \mathbb{Z}_{p,\infty} \cup \{\infty\}$, which is the desired conclusion. □

Lemma 4.2 *Given $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_{p,\infty} \cup \{\infty\}$ let $\mathbb{Z}_{p,q} \ni k \to \mathcal{F}_k$ be a function such that $\mathcal{F}_k$ is a nonempty linear set in the space $\mathrm{H}(\mathfrak{R})$ for every $k \in \mathbb{Z}_{p,q}$. If $g \in \mathcal{L}_2(\mathfrak{R})$, $\mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g) \neq \emptyset$ for every $k \in \mathbb{Z}_{p,q}$ and the condition (4.3) holds, then*

$$\mathcal{F} \cap \left(\sum_{k=p}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g)\right) \subset \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g), \tag{4.13}$$

where $\mathcal{F} := \sum_{k=p}^{q} \mathcal{F}_k$.

Proof Fix

$$f \in \mathcal{F} \cap \left(\sum_{k=p}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g)\right).$$

By definition, there exists a sequence $\mathbb{Z}_{p,q} \ni k \mapsto f_k \in \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g)$ such that

$$f \in \sum_{k=p}^{q} \{f_k\}. \tag{4.14}$$

Applying Lemma 3.1, with $\mathcal{F}$ replaced by any $\mathcal{F}_k$, we see that

$$\langle h | f_k \rangle = g^*(h), \quad h \in \mathcal{F}_k, \ k \in \mathbb{Z}_{p,q}. \tag{4.15}$$

We will show that

$$\langle h | f \rangle = g^*(h), \quad h \in \mathcal{F}. \tag{4.16}$$

Given $h \in \mathcal{F}$ there exists a sequence $\mathbb{Z}_{p,q} \ni k \mapsto h_k \in \mathcal{F}_k$ such that

$$h \in \sum_{k=p}^{q} \{h_k\}.$$

Consider first the case where $q \in \mathbb{Z}$. By (4.3) and (4.15) we get

$$\langle h | f \rangle = \left\langle \sum_{k=p}^{q} h_k \Big| \sum_{l=p}^{q} f_l \right\rangle = \sum_{k=p}^{q} \sum_{l=p}^{q} \langle h_k | f_l \rangle \tag{4.17}$$

$$= \sum_{k=p}^{q} \langle h_k | f_k \rangle = \sum_{k=p}^{q} g^*(h_k) = g^* \left(\sum_{k=p}^{q} h_k \right) = g^*(h),$$

which yields (4.16).

Consider now the case of $q = \infty$. For each $n \in \mathbb{Z}_{p,\infty}$ set

$$f'_n := \sum_{k=p}^{n} f_k \quad \text{and} \quad h'_n := \sum_{k=p}^{n} h_k.$$

Applying the Schwarz inequality (2.20) we conclude from (4.3) that for all $n, m \in \mathbb{Z}_{p,\infty}$ with $n < m$,

$$\langle h'_m - h'_n | f'_n \rangle = \left\langle \sum_{k=n+1}^{m} h_k \Big| \sum_{l=p}^{n} f_l \right\rangle = \sum_{k=n+1}^{m} \sum_{l=p}^{n} \langle h_k | f_l \rangle = 0.$$

Applying again the Schwarz inequality (2.20) we see that for every $n \in \mathbb{Z}_{p,\infty}$,

$$|\langle h - h'_n | f'_n \rangle| = |\langle h - h'_m | f'_n \rangle + \langle h'_m - h'_n | f'_n \rangle|$$
$$\leq \|h - h'_m\| \|f'_n\| + |\langle h'_m - h'_n | f'_n \rangle| \to 0 \quad \text{as } m \to \infty,$$

and hence

$$\langle h - h'_n | f'_n \rangle = 0, \quad n \in \mathbb{Z}_{p,\infty}. \tag{4.18}$$

Likewise, we obtain

$$\langle h'_n | f - f'_n \rangle = 0, \quad n \in \mathbb{Z}_{p,\infty}. \tag{4.19}$$

Analysis similar to that in (4.17) shows that

$$\langle h'_n | f'_n \rangle = g^*(h'_n), \quad n \in \mathbb{Z}_{p,\infty}. \tag{4.20}$$

By Lemma 2.5, the functional g^* is continuous on $\mathrm{H}(\mathfrak{R})$, and so

$$g^*(h'_n - h) \to 0 \quad \text{as } n \to \infty.$$

Combining this with (4.18), (4.19) and (4.20) as well as applying the Schwarz inequality (2.20) we obtain

$$
\begin{aligned}
| \langle h | f \rangle - g^*(h) | &= | \langle h - h'_n + h'_n | f - f'_n + f'_n \rangle - g^*(h) | \\
&= | \langle h - h'_n | f - f'_n \rangle + \langle h - h'_n | f'_n \rangle + \langle h'_n | f - f'_n \rangle + \langle h'_n | f'_n \rangle - g^*(h) | \\
&\le | \langle h - h'_n | f - f'_n \rangle | + | g^*(h'_n) - g^*(h) | \\
&\le \| h - h'_n \| \| f - f'_n \| + | g^*(h'_n - h) | \to 0 \quad \text{as } n \to \infty,
\end{aligned}
$$

which yields (4.16). Hence Lemma 3.1 shows that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$. Therefore the inclusion (4.13) holds, which is our claim. $\square$

The basic result in this section is the following orthogonal decomposition theorem in the countable case.

Theorem 4.3 *Given $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_{p,\infty} \cup \{\infty\}$ let $\mathbb{Z}_{p,q} \ni k \to \mathcal{F}_k$ be a function such that $\mathcal{F}_k$ is a nonempty closed linear set in the space $\mathrm{H}(\mathfrak{R})$ for all $k \in \mathbb{Z}_{p,q}$. If $g \in \mathcal{L}_2(\mathfrak{R})$ and the condition (4.3) holds, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \sum_{k=p}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g), \tag{4.21}$$

where $\mathcal{F} := \sum_{k=p}^{q} \mathcal{F}_k$.

Proof Since the sets $\mathcal{F}_k$, as $k \in \mathbb{Z}_{p,q}$, are closed in $\mathrm{H}(\mathfrak{R})$, we conclude from Corollary 3.4, with $\mathcal{F}$ replaced by each $\mathcal{F}_k$, that

$$\mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g) = \Theta + f_k, \quad k \in \mathbb{Z}_{p,q}, \tag{4.22}$$

for some sequence $\mathbb{Z}_{p,q} \ni k \mapsto f_k \in \mathcal{F}_k$. Fix $f \in \sum_{k=p}^{q} \{f_k\}$. From Lemma 4.2 it follows that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$. By Lemma 4.1 we deduce that $\mathcal{F}$ is a linear and closed set in $\mathrm{H}(\mathfrak{R})$. Then by Corollary 3.4 we see that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta + f.$$

From this and (4.22) we conclude that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta + f = \Theta + \sum_{k=p}^{q}\{f_k\} = \sum_{k=p}^{q}(\Theta + f_k) = \sum_{k=p}^{q}\mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g),$$

and so the equality (4.21) holds, which is our claim. $\qquad\qquad\qquad\qquad\qquad$ □

We now extend the decomposition (4.21) to arbitrary family of closed linear sets in the space $H(\mathfrak{R})$. To this end we write

$$\|V\| := \sup_{f \in V} \|f\|$$

for any nonempty set $V \subset \mathcal{L}_1(\mathfrak{R})$. Then the following theorem holds.

Theorem 4.4 *Given a set* $\mathcal{A} \neq \emptyset$ *let* $\mathcal{A} \ni \alpha \to \mathcal{F}_\alpha$ *be a function such that* $\mathcal{F}_\alpha$ *is a nonempty closed linear set in the space* $H(\mathfrak{R})$ *for every* $\alpha \in \mathcal{A}$*. If* $g \in \mathcal{L}_2(\mathfrak{R})$ *and*

$$\mathcal{F}_\alpha \perp \mathcal{F}_\beta, \quad \alpha, \beta \in \mathcal{A}, \alpha \neq \beta, \tag{4.23}$$

then

$$\mathcal{A}' := \{\alpha \in \mathcal{A} : \|\mathrm{Reg}(\mathcal{F}_\alpha, \mathfrak{R}_g)\| > 0\}$$

is a countable set and for every $q \in \mathbb{N} \cup \{\infty\}$ *and every injective mapping* σ *of* $\mathbb{Z}_{1,q}$ *onto* $\mathcal{A}'$*, the equality*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \sum_{k=1}^{q}\mathrm{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g), \tag{4.24}$$

where

$$\mathcal{F} := \mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}}\mathcal{F}_\alpha\right)\right),$$

holds provided $\mathcal{A}' \neq \emptyset$ *and*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta$$

otherwise.

Proof Since $\mathcal{F}$ and $\mathcal{F}_\alpha$ as $\alpha \in \mathcal{A}$ are closed sets in $H(\mathfrak{R})$ we see that

$$\Theta \subset \mathcal{F} \quad \text{and} \quad \Theta \subset \mathcal{F}_\alpha, \quad \alpha \in \mathcal{A}. \tag{4.25}$$

Hence and by Corollary 3.4 we deduce that there exist $f \in \mathcal{F}$ and a function

$$\mathcal{A} \ni \alpha \mapsto f_\alpha \in \mathcal{F}_\alpha$$

such that

$$\text{Reg}(\mathcal{F}, \mathfrak{R}) = \Theta + f \quad \text{and} \quad \text{Reg}(\mathcal{F}_\alpha, \mathfrak{R}_g) = \Theta + f_\alpha, \quad \alpha \in \mathcal{A}. \tag{4.26}$$

Given $p, q \in \mathbb{N}$ with $p < q$, let us consider an injective sequence

$$\mathbb{Z}_{p,q} \ni k \to \gamma(k) \in \mathcal{A}.$$

Write

$$\mathcal{F}_\gamma := \mathcal{F} \cap \left(\sum_{k=p}^{q} \mathcal{F}_{\gamma(k)} \right)^{\perp}.$$

From (4.23) and (3.14c) it follows that

$$\mathcal{F} = \mathcal{F}_\gamma + \sum_{k=p}^{q} \mathcal{F}_{\gamma(k)}, \tag{4.27}$$

and applying now Theorem 4.3 we obtain

$$\Theta + f = \Theta + f_\gamma + \sum_{k=p}^{q} f_{\gamma(k)}, \tag{4.28}$$

for some $f_\gamma \in \text{Reg}(\mathcal{F}_\gamma, \mathfrak{R}_g)$. From this, by (4.23) and (4.2f) we have

$$\|f\|^2 = \left\| f_\gamma + \sum_{k=p}^{q} f_{\gamma(k)} \right\|^2 = \|f_\gamma\|^2 + \sum_{k=p}^{q} \|f_{\gamma(k)}\|^2 \geq \sum_{k=p}^{q} \|f_{\gamma(k)}\|^2. \tag{4.29}$$

Fix now $m \in \mathbb{N}$ and assume that the set

$$\mathcal{A}_m := \left\{ \alpha \in \mathcal{A} : \|f_\alpha\| \geq \frac{1}{m} \right\}$$

is not finite. Then there exists an injective sequence $\sigma : \mathbb{N} \to \mathcal{A}_m$. Applying the inequality (4.29) with $p := 1$ and γ replaced by σ restricted to the set $\mathbb{Z}_{1,q}$ we see that

$$q \leq m^2 \|f\|^2$$

for every $q \in \mathbb{N}$, which is impossible. This means that the set $\mathcal{A}_m$ is finite for every $m \in \mathbb{N}$. Since

$$\mathcal{A}' = \bigcup_{m \in \mathbb{N}} \mathcal{A}_m$$

it follows that $\mathcal{A}'$ is a countable set.

From Lemma 3.1 and (4.26) it follows that

$$\langle h | f_\alpha \rangle = g^*(h), \quad h \in \mathcal{F}_\alpha, \; \alpha \in \mathcal{A}. \tag{4.30}$$

Setting

$$\mathcal{A}'' := \{\alpha \in \mathcal{A} : \|\operatorname{Reg}(\mathcal{F}_\alpha, \mathfrak{R}_g)\| = 0\}$$

we see that

$$\mathcal{A}' \cup \mathcal{A}'' = \mathcal{A} \neq \emptyset.$$

The remaining part of the proof will be divided naturally into three cases: $\mathcal{A}'' = \emptyset$, $\mathcal{A}' = \emptyset$ and $\mathcal{A}' \neq \emptyset \neq \mathcal{A}''$.

The case I, where $\mathcal{A}'' = \emptyset$. Then $\mathcal{A}' = \mathcal{A} \neq \emptyset$. Since $\mathcal{A}'$ is a countable set, there exist $q \in \mathbb{N} \cup \{\infty\}$ and an injective mapping σ of $\mathbb{Z}_{1,q}$ onto $\mathcal{A}$. Setting $\mathcal{F}'_k := \mathcal{F}_{\sigma(k)}$, $k \in \mathbb{Z}_{1,q}$, we conclude from Lemma 4.1 that

$$\sum_{k=1}^{q} \mathcal{F}'_k = \operatorname{cl}\left(\operatorname{lin}\left(\bigcup_{k=1}^{q} \mathcal{F}'_k\right)\right) = \operatorname{cl}\left(\operatorname{lin}\left(\bigcup_{k=1}^{q} \mathcal{F}_{\sigma(k)}\right)\right) = \operatorname{cl}\left(\operatorname{lin}\left(\bigcup_{\alpha \in \mathcal{A}} \mathcal{F}_\alpha\right)\right) = \mathcal{F}.$$

Applying now Theorem 4.3 we can see that

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{R}_g) = \sum_{k=p}^{q} \operatorname{Reg}(\mathcal{F}'_k, \mathfrak{R}_g) = \sum_{k=1}^{n} \operatorname{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g),$$

which yields the equality (4.24).

The case II, where $\mathcal{A}' = \emptyset$. Then $\mathcal{A}'' = \mathcal{A} \neq \emptyset$. Fix $h \in \mathcal{F}$. By the definition of $\mathcal{F}$ it follows that there exists a sequence

$$\mathbb{N} \ni n \mapsto h_n \in \mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}} \mathcal{F}_\alpha \right)$$

such that

$$\|h - h_n\| \to 0 \quad \text{as } n \to \infty. \tag{4.31}$$

For each $n \in \mathbb{N}$ there exist $n' \in \mathbb{N}$ and a sequence $\mathbb{Z}_{1,n'} \ni k \mapsto \alpha_k \in \mathcal{A}$ such that

$$h_n \in \mathcal{F}'_n := \sum_{k=1}^{n'} \mathcal{F}_{\alpha_k}.$$

Then there exists a sequence

$$\mathbb{Z}_{1,n'} \ni k \mapsto h_{n,\alpha_k} \in \mathcal{F}_{\alpha_k}$$

such that

$$h_n = \sum_{k=1}^{n'} h_{n,\alpha_k}.$$

Applying now (4.30) and the Schwarz inequality (2.20) we obtain

$$|g^*(h_n)| = \left| g^*\left(\sum_{k=1}^{n'} h_{n,\alpha_k} \right) \right| \leq \sum_{k=1}^{n'} |g^*(h_{n,\alpha_k})| = \sum_{k=1}^{n'} |\langle h_{n,\alpha_k} | f_{\alpha_k} \rangle|$$

$$\leq \sum_{k=1}^{n'} \|h_{n,\alpha_k}\| \cdot \|f_{\alpha_k}\| \leq \sum_{k=1}^{n'} \|h_{n,\alpha_k}\| \cdot \|\mathrm{Reg}(\mathcal{F}_{\alpha_k}, \mathfrak{R}_g)\| = 0.$$

Thus $g^*(h_n) = 0$ for every $n \in \mathbb{N}$. Since g^* is a continuous functional on the space $\mathrm{H}(\mathfrak{R})$, we conclude from (4.31) that $g^*(h) = 0$ for every $h \in \mathcal{F}$. Then Lemma 3.1 shows that $\theta \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$. Hence and by Corollary 3.4 we obtain

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = \Theta + \theta = \Theta,$$

which proves the theorem in the case where $\mathcal{A}' = \emptyset$.

The case III, where $\mathcal{A}' \neq \emptyset \neq \mathcal{A}''$. Setting

$$\mathcal{F}' := \mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}'} \mathcal{F}_\alpha \right)$$

and

$$\mathcal{F}'' := \mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}''} \mathcal{F}_\alpha\right)$$

we conclude from (4.23) that $\mathcal{F}' \perp \mathcal{F}''$. Then (4.2a) yields

$$\mathrm{cl}(\mathcal{F}') \perp \mathrm{cl}(\mathcal{F}'').$$

Moreover, by (3.12) and (4.2b) we obtain

$$\mathrm{cl}(\mathcal{F}') + \mathrm{cl}(\mathcal{F}'') = \mathrm{cl}(\mathcal{F}' + \mathcal{F}'') = \mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}'} \mathcal{F}_\alpha\right) + \mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}''} \mathcal{F}_\alpha\right)\right)$$

$$= \mathrm{cl}\left(\mathrm{lin}\left(\bigcup_{\alpha \in \mathcal{A}' \cup \mathcal{A}''} \mathcal{F}_\alpha\right)\right) = \mathcal{F}.$$

Theorem 4.3 now implies

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathrm{Reg}(\mathrm{cl}(\mathcal{F}'), \mathfrak{R}_g) + \mathrm{Reg}(\mathrm{cl}(\mathcal{F}''), \mathfrak{R}_g). \tag{4.32}$$

Applying our theorem in the already proved case I we have

$$\mathrm{Reg}(\mathrm{cl}(\mathcal{F}'), \mathfrak{R}_g) = \sum_{k=1}^{q} \mathrm{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g), \tag{4.33}$$

where $q \in \mathbb{N} \cup \{\infty\}$ and σ is an injective mapping of $\mathbb{Z}_{1,q}$ onto $\mathcal{A}'$. Applying once again our theorem in the already proved case II we obtain

$$\mathrm{Reg}(\mathrm{cl}(\mathcal{F}''), \mathfrak{R}_g) = \Theta. \tag{4.34}$$

Combining the equalities (4.32), (4.33) and (4.34) we conclude from Corollary 3.4 that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta + \sum_{k=1}^{q} \mathrm{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g) = \sum_{k=1}^{q} (\Theta + \mathrm{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g))$$

$$= \sum_{k=1}^{q} \mathrm{Reg}(\mathcal{F}_{\sigma(k)}, \mathfrak{R}_g),$$

which yields the equality (4.24) in the third case, and the proof is complete. □

Corollary 4.5 *Given* $q \in \mathbb{N} \cup \{\infty\}$ *let* $\mathbb{Z}_{1,q} \ni k \to \mathcal{F}_k$ *be a sequence such that* $\mathcal{F}_k$ *is a nonempty closed linear set in the space* $\mathrm{H}(\mathfrak{R})$ *for every* $k \in \mathbb{Z}_{1,p}$. *If* $\mathcal{F}$ *is a*

closed linear set in the space $\mathrm{H}(\mathfrak{R})$, $g \in \mathcal{L}_2(\mathfrak{R})$, $\mathcal{F}_k \subset \mathcal{F}$ *for every* $k \in \mathbb{Z}_{1,p}$ *and the condition* (4.3) *holds, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathrm{Reg}(\mathcal{F}_0, \mathfrak{R}_g) + \sum_{k=1}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g), \qquad (4.35)$$

where $\mathcal{F}_0 := \mathcal{F} \cap \left(\sum_{k=1}^{p} \mathcal{F}_k\right)^{\perp}$.

Proof Since $\mathcal{F}$ is a nonempty closed linear set in the space $\mathrm{H}(\mathfrak{R})$, we conclude from (3.14a) that $\mathcal{F}_0$ is also a nonempty closed linear set in the space $\mathrm{H}(\mathfrak{R})$. Moreover, $\mathcal{F}_0 \perp \mathcal{F}_k$ for every $k \in \mathbb{Z}_{1,p}$ and, by (3.14c),

$$\sum_{k=0}^{p} \mathcal{F}_k = \mathcal{F}_0 + \sum_{k=1}^{p} \mathcal{F}_k = \mathcal{F}.$$

Then Theorem 4.3 shows that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \sum_{k=0}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g) = \mathrm{Reg}(\mathcal{F}_0, \mathfrak{R}_g) + \sum_{k=1}^{q} \mathrm{Reg}(\mathcal{F}_k, \mathfrak{R}_g),$$

and the equality (4.35) is proved.   □

5 The regression functions calculating procedure

Given an asynchronous regression structure $\mathfrak{R}$ let $\mathcal{F}$ be a nonempty linear set in the space $\mathrm{H}(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$. If $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset$, then Theorem 3.3 enables us to find regression functions in $\mathcal{F}$ with respect to $\mathfrak{R}_g$, provided we can determine the linear set $(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F}$, which is rather difficult task, in general. However in the case where $\mathcal{F}$ is spanned by a finite system of functions we can effectively calculate all the regression functions in $\mathcal{F}$ with respect to $\mathfrak{R}_g$ in terms of these functions; cf. [12, Corollary 3.2]. Obviously, this case is the most essential one from the practical point of view.

In what follows we will show how to use directly Lemma 3.1 in order to determine all regression functions in $\mathcal{F}$ with respect to $\mathfrak{R}_g$, provided

$$\mathcal{F} = \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\}) \qquad (5.1)$$

for some $p \in \mathbb{N}$ and $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$. By (5.1), each $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$ is of the form $f = \sum_{k=1}^{p} \lambda_k h_k$ for a sequence $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B$. From Lemma 3.1 it follows that f satisfies the condition (3.2), which leads to the following equalities

$$g^*(h_l) = \langle h_l | f \rangle = \left\langle h_l \,\Big|\, \sum_{k=1}^{p} \lambda_k h_k \right\rangle = \sum_{k=1}^{p} \overline{\lambda_k} \, \langle h_l | h_k \rangle, \quad l \in \mathbb{Z}_{1,p}.$$

This way we obtain the following linear equation system of the Gramm-Schmidt type

$$\sum_{k=1}^{p} \lambda_k \langle h_k | h_l \rangle = \overline{g^*(h_l)}, \quad l \in \mathbb{Z}_{1,p}, \tag{5.2}$$

with respect to the variables $\lambda_1, \lambda_2, \ldots, \lambda_p \in B$. All solutions of the equation system (5.2) determine the set $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g)$. The system (5.2) simplifies itself much if we assume that the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ is *orthogonal* in the space $\mathrm{H}(\mathfrak{R})$, i.e.

$$h_k \perp h_l, \quad k, l \in \mathbb{Z}_{1,p}, \ k \neq l. \tag{5.3}$$

Then $\langle h_k | h_l \rangle = 0$ for $k \neq l$, and by (5.2) we obtain the unique solution

$$\lambda_k = \frac{\overline{g^*(h_k)}}{\|h_k\|^2}, \quad k \in \mathbb{Z}_{1,p}, \tag{5.4}$$

provided $\|h_k\| > 0$ as $k \in \mathbb{Z}_{1,p}$. Thus $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \{f\}$, where

$$f = \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k; \tag{5.5}$$

cf. [12, Thm. 3.1].

Using orthogonal decomposition properties from the previous section we can develop this idea in the form of the following theorem.

Theorem 5.1 *Given $p \in \mathbb{N} \cup \{\infty\}$ let $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R}) \backslash \Theta$ be an orthogonal sequence in the space $\mathrm{H}(\mathfrak{R})$ and $g \in \mathcal{L}_2(\mathfrak{R})$. If $p \in \mathbb{N}$, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k, \tag{5.6}$$

where $\mathcal{F} := \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$. If $p = \infty$, then

$$\mathrm{Reg}(\mathrm{cl}(\mathcal{F}), \mathfrak{R}_g) = \Theta + \sum_{k=1}^{\infty} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k. \tag{5.7}$$

Proof Fix $p \in \mathbb{N} \cup \{\infty\}$ and an orthogonal sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{F}$ in $\mathrm{H}(\mathfrak{R})$.

Consider first the case where $p = 1$. If $g^*(h_1) = 0$, then $\mathcal{F} \subset S$, where $S := (g^*)^{-1}(0)$. Then by (iii) of Theorem 3.3,

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \Theta \cap \mathcal{F},$$

and so the equality (5.6) obviously holds. Assume in contrary, that $g^*(h_1) \neq 0$. Then $\mathcal{F} \cap S = \{\theta\}$, and so $(\mathcal{F} \cap S)^\perp = \mathcal{L}_1(\mathfrak{R})$. Hence

$$h_1 \in (\mathcal{F} \cap S)^\perp \cap \mathcal{F} \backslash \Theta,$$

and applying now (ii) of Theorem 3.3 we derive the equality (5.6). Thus the equality (5.6) holds for every one-dimensional linear set $\mathcal{F} \subset \mathcal{L}_1(\mathfrak{R})$.

Assume now that $p \in \mathbb{N}$ or $p = \infty$. Setting $\mathcal{F}_k := \text{lin}(\{h_k\})$ for $k \in \mathbb{Z}_{1,p}$ we see that each set $\mathcal{F}_k$ is a one-dimensional linear subset of $\mathcal{L}_1(\mathfrak{R})$. Hence

$$\text{Reg}(\mathcal{F}_k, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}_k) + \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k, \quad k \in \mathbb{Z}_{1,p}. \tag{5.8}$$

If $p \in \mathbb{N}$, then from Lemma 4.2 we conclude that

$$f := \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \in \mathcal{F} \cap \sum_{k=1}^{p} \text{Reg}(\mathcal{F}_k, \mathfrak{R}_g) \subset \text{Reg}(\mathcal{F}, \mathfrak{R}_g),$$

and consequently the equality (5.6) follows from (i) of Theorem 3.3. Thus it remains to consider the case where $p = \infty$. Since

$$\text{cl}(\mathcal{F}_k) = \Theta + \mathcal{F}_k \quad \text{as } k \in \mathbb{Z}_{1,\infty},$$

we deduce from (5.8) and Corollary 3.2 that for every $k \in \mathbb{Z}_{1,\infty}$,

$$\frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \in \text{Reg}(\mathcal{F}_k, \mathfrak{R}_g) = \mathcal{F}_k \cap \text{Reg}(\Theta + \mathcal{F}_k, \mathfrak{R}_g) \subset \text{Reg}(\text{cl}(\mathcal{F}_k), \mathfrak{R}_g).$$

Then Corollary 3.4 yields

$$\text{Reg}(\text{cl}(\mathcal{F}_k), \mathfrak{R}_g) = \Theta + \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k, \quad k \in \mathbb{Z}_{1,\infty}. \tag{5.9}$$

Since the sequence $\mathbb{Z}_{1,\infty} \ni k \mapsto h_k \in \mathcal{F}$ is orthogonal, we deduce from Lemma 4.1 and (4.2a) that

$$\text{cl}(\mathcal{F}) = \sum_{k=1}^{\infty} \text{cl}(\mathcal{F}_k) \quad \text{and} \quad \text{cl}(\mathcal{F}_k) \perp \text{cl}(\mathcal{F}_l), \quad k, l \in \mathbb{Z}_{1,\infty}, \quad k \neq l. \tag{5.10}$$

Applying now Theorem 4.3 we conclude from (5.9) and (5.10) that

$$\text{Reg}(\text{cl}(\mathcal{F}), \mathfrak{R}_g) = \sum_{k=1}^{\infty} \text{Reg}(\text{cl}(\mathcal{F}_k), \mathfrak{R}_g) = \sum_{k=1}^{\infty} \left(\Theta + \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \right),$$

which proves the equality (5.7).   □

As far as applications are concerned we will study theoretic models

$$\mathcal{F} = \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$$

spanned by sequences $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ which in general are not orthogonal in the space $H(\mathfrak{R})$, because the pseudo-inner product $\langle \cdot | \cdot \rangle$ depends on the empirical data function $x : \Omega_1 \to A$ and measure μ. Then we can not apply Theorem 5.1 directly. However, in such cases we can ortogonalize these sequences. To this end we recall that for a given $p \in \mathbb{N} \cup \{\infty\}$, a sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k' \in \mathcal{L}_1(\mathfrak{R})$$

is said to be an *orthogonalization* of a sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R}),$$

provided

$$h_k' \in H_k \cap H_{k-1}^{\perp} \backslash \Theta, \quad k \in \mathbb{Z}_{1,p}, \tag{5.11}$$

where $H_0 := \Theta$ and $H_k := \mathrm{lin}(\{h_m : m \in \mathbb{Z}_{1,k}\})$, $k \in \mathbb{Z}_{1,p} \cup \{p\}$. The following lemma gives a sufficient and necessarily condition for a sequence to be ortogonalized.

Lemma 5.2 *Given $p \in \mathbb{N} \cup \{\infty\}$, a sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ is linearly independent and $\Theta \cap H_p = \{\theta\}$ iff the following condition*

$$H_k \cap H_{k-1}^{\perp} \backslash \Theta \neq \emptyset, \quad k \in \mathbb{Z}_{1,p}, \tag{5.12}$$

holds.

Proof Fix $p \in \mathbb{N}$ and a sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$. Assume first that the condition (5.12) holds. Then there exists a sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k' \in \mathcal{L}_1(\mathfrak{R})$ satisfying (5.11). Setting

$$H_p' := \mathrm{lin}(\{h_k' : k \in \mathbb{Z}_{1,p}\})$$

we shall prove that $H_p' = H_p$. Since $H_{k-1} \subset H_k$ for $k \in \mathbb{Z}_{1,p}$, the sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k'$$

is orthogonal. Given a sequence $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B$ assume that $\sum_{k=1}^{p} \lambda_k h_k' \in \Theta$. Then for every $l \in \mathbb{Z}_{1,p}$,

$$0 = \left\langle \sum_{k=1}^{p} \lambda_k h_k' \middle| h_l' \right\rangle = \sum_{k=1}^{p} \lambda_k \langle h_k' | h_l' \rangle = \lambda_l \|h_l'\|^2,$$

and hence $\lambda_l = 0$, because $\|h'_l\| > 0$ by (5.11). Therefore the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ is linearly independent and

$$H'_p \cap \Theta = \{\theta\}. \tag{5.13}$$

Moreover, from (5.11) it follows that

$$h'_k \in H_k \subset H_p \quad \text{as } k \in \mathbb{Z}_{1,p},$$

and consequently $H'_p \subset H_p$. Thus

$$p = \dim(H'_p) \leq \dim(H_p) \leq p,$$

which implies $\dim(H_p) = p$. Hence $H'_p = H_p$. Combining this with (5.13) we have

$$\Theta \cap H_p = \Theta \cap H'_p = \{\theta\}.$$

Moreover the equality $\dim(H_p) = p$ implies that the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ is linearly independent. This proves the lemma in the direction ($\Leftarrow$), provided $p \in \mathbb{N}$.

Conversely, assume that the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ is linearly independent and the equality $\Theta \cap H_p = \{\theta\}$ holds. If $\|h_k\| = 0$ for certain $k \in \mathbb{Z}_{1,p}$, then

$$h_k \in \Theta \cap H_p,$$

and so $h_k = \theta$. This is impossible, because the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ is linearly independent. Therefore

$$\|h_k\| > 0, \quad k \in \mathbb{Z}_{1,p}. \tag{5.14}$$

In particular,

$$h_1 \in H_1 \backslash \Theta = H_1 \cap H_0^\perp \backslash \Theta,$$

and so the property (5.12) holds in the case where $p = 1$. Therefore we may assume that $p \geq 2$. Then

$$q := \min(\{k \in \mathbb{Z}_{1,p} : H_k \cap H_{k-1}^\perp \backslash \Theta = \emptyset\}) \in \mathbb{Z}_{2,p}, \tag{5.15}$$

provided the condition (5.12) does not hold. Hence there exists a sequence

$$\mathbb{Z}_{1,q-1} \ni k \mapsto h'_k \in H_k \cap H_{k-1}^\perp \backslash \Theta,$$

and we can define

$$h'_q := h_q - \sum_{k=1}^{q-1} \frac{\langle h_q | h'_k \rangle}{\|h'_k\|^2} h'_k. \tag{5.16}$$

Since $H_{k-1} \subset H_k$ for $k \in \mathbb{Z}_{1,q}$, the sequence $\mathbb{Z}_{1,q-1} \ni k \mapsto h'_k$ is orthogonal, and consequently for each $l \in \mathbb{Z}_{1,q-1}$,

$$\langle h'_q | h'_l \rangle = \langle h_q | h'_l \rangle - \sum_{k=1}^{q-1} \frac{\langle h_q | h'_k \rangle}{\|h'_k\|^2} \langle h'_k | h'_l \rangle = \langle h_q | h'_l \rangle - \frac{\langle h_q | h'_l \rangle}{\|h'_l\|^2} \langle h'_l | h'_l \rangle = 0. \quad (5.17)$$

Therefore the sequence $\mathbb{Z}_{1,q} \ni k \mapsto h'_k$ is orthogonal, and following the first part of the proof we see that this sequence is linearly independent. Hence

$$\dim(\operatorname{lin}(\{h'_k : k \in \mathbb{Z}_{1,q-1}\})) = q - 1.$$

Since $h'_k \in H_k \subset H_{q-1}$ for $k \in \mathbb{Z}_{1,q-1}$, we see that

$$\operatorname{lin}(\{h'_k : k \in \mathbb{Z}_{1,q-1}\}) \subset H_{q-1}.$$

On the other hand side the sequence $\mathbb{Z}_{1,q-1} \ni k \mapsto h_k$ is also linearly independent, and consequently

$$\dim(H_{q-1}) = q - 1.$$

Thus

$$\operatorname{lin}(\{h'_k : k \in \mathbb{Z}_{1,q-1}\}) = H_{q-1},$$

which together with (5.16) and (5.17) leads to

$$h'_q \in H_q \cap H_{q-1}^\perp. \quad (5.18)$$

Suppose that $\|h'_q\| = 0$. Then

$$h'_q \in \Theta \cap H_q \subset \Theta \cap H_p = \{\theta\},$$

and consequently

$$h_q = \sum_{k=1}^{q-1} \frac{\langle h_q | h'_k \rangle}{\|h'_k\|^2} h'_k.$$

This is impossible, because the sequence $\mathbb{Z}_{1,q} \ni k \mapsto h_k$ is linearly independent. Therefore $\|h'_q\| > 0$, which together with (5.18) implies that

$$h'_q \in H_q \cap H_{q-1}^\perp \setminus \Theta,$$

contrary to (5.15). This means that the condition (5.12) holds, which proves the lemma in the direction ($\Rightarrow$), provided $p \in \mathbb{N}$.

It remains to prove the lemma in the case where $p = \infty$. Then $\mathbb{Z}_{1,p} = \mathbb{N}$. Assume first that a sequence $\mathbb{N} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ is linearly independent and the equality $\Theta \cap H_\infty = \{\theta\}$ holds. Then for every $n \in \mathbb{N}$ the sequence $\mathbb{Z}_{1,n} \ni k \mapsto h_k$ is also linearly independent and

$$H_n \cap \Theta \subset H_\infty \cap \Theta = \{\theta\},$$

which gives

$$H_n \cap \Theta = \{\theta\}.$$

Applying now the already proved finite part ($\Rightarrow$) of the lemma we see that

$$H_n \cap H_{n-1}^\perp \backslash \Theta \neq \emptyset$$

for every $n \in \mathbb{N}$, which means that the condition (5.12) holds. Conversely, assume now that a sequence

$$\mathbb{N} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$$

satisfies the condition (5.12). From the already proved finite part ($\Leftarrow$) of the lemma it follows that for every $n \in \mathbb{N}$ the sequence $\mathbb{Z}_{1,n} \ni k \mapsto h_k$ is linearly independent and $H_n \cap \Theta = \{\theta\}$. Hence each finite subsequence of the sequence $\mathbb{N} \ni k \mapsto h_k$ is linearly independent, and hence the sequence $\mathbb{N} \ni k \mapsto h_k$ is linearly independent as well. Moreover,

$$\Theta \cap H_\infty = \Theta \cap \bigcup_{n=1}^\infty H_n = \bigcup_{n=1}^\infty \Theta \cap H_n = \{\theta\}.$$

Thus the lemma holds also in the case, where $p = \infty$, which completes the proof. $\square$

From Lemma 5.2 it follows that each linearly independent sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ satisfying the condition

$$\Theta \cap H_p = \{\theta\}$$

has an associated sequence being its ortogonalization result. Such a sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k'$ may be determined by using the Gramm-Schmidt recursive method by setting

$$h_1' := h_1 \quad \text{and} \quad h_n' := h_n - \sum_{k=1}^{n-1} \frac{\langle h_n | h_k' \rangle}{\|h_k'\|^2} h_k', \quad n \in \mathbb{Z}_{2,p}. \tag{5.19}$$

Corollary 5.3 *Given* $p \in \mathbb{N} \cup \{\infty\}$ *let* $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ *be a linearly independent sequence such that* $\Theta \cap \mathcal{F} = \{\theta\}$, *where* $\mathcal{F} := \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$.

Suppose that $\mathbb{Z}_{1,p} \ni k \mapsto h'_k \in \mathcal{L}_1(\mathfrak{R})$ is a sequence satisfying (5.11) *and that $g \in \mathcal{L}_2(\mathfrak{R})$. If $p \in \mathbb{N}$, then*

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k. \tag{5.20}$$

If $p = \infty$, then

$$\mathrm{Reg}(\mathrm{cl}(\mathcal{F}), \mathfrak{R}_g) = \Theta + \sum_{k=1}^{\infty} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k. \tag{5.21}$$

In particular, the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ may be defined by (5.19).

Proof Fix $p \in \mathbb{N} \cup \{\infty\}$ and consider any sequences

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{F} \backslash \Theta$$

and

$$\mathbb{Z}_{1,p} \ni k \mapsto h'_k \in \mathcal{L}_1(\mathfrak{R})$$

satisfying the assumptions. By Lemma 5.2, the condition (5.12) holds. Therefore the last sequence exists. It may be defined for instance by (5.19). From the property (5.11) it follows that $\|h'_k\| \neq 0$ for $k \in \mathbb{Z}_{1,p}$ and $h'_k \perp h'_l$ for $k, l \in \mathbb{Z}_{1,p}$ such that $k \neq l$. Moreover, by definition,

$$\mathrm{lin}(\{h'_k : k \in \mathbb{Z}_{1,p}\}) = \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\}) = \mathcal{F}.$$

Thus, applying Theorem 5.1 to the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ replaced by its ortogonalized associate $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$, we derive the assertion. □

The following example shows how to apply Lemma 5.2 and Corollary 5.3 to numerical computation of the regression functions.

Example 5.4 Given $p \in \mathbb{N}$ let $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ be a linearly independent sequence such that $\Theta \cap \mathcal{F} = \{\theta\}$, where $\mathcal{F} := \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$. Suppose that

$$\mathbb{Z}_{1,p} \ni k \mapsto h'_k \in \mathcal{L}_1(\mathfrak{R})$$

is a sequence satisfying (5.11) and that $g \in \mathcal{L}_2(\mathfrak{R})$. Let us consider the sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto v_k := h_k - \frac{\langle h_k | h'_k \rangle}{\|h'_k\|^2} h'_k. \tag{5.22}$$

Fix $n \in \mathbb{Z}_{1,p}$ and suppose that $h'_n \perp h_n$. Since $h'_n \perp H_{n-1}$ and

$$H_n = H_{n-1} + \mathrm{lin}(\{h_n\}), \tag{5.23}$$

we have $h'_n \perp H_n$. On the other hand side $h'_n \in H_n$. Hence

$$\|h'_n\|^2 = \langle h'_n | h'_n \rangle = 0,$$

which is impossible, because $h'_n \notin \Theta$. Therefore

$$\langle h_k | h'_k \rangle \neq 0, \quad k \in \mathbb{Z}_{1,p}. \tag{5.24}$$

Since $h_n, h'_n \in H_n$, we deduce from (5.22) that $v_n \in H_n$. From this and (5.23) we conclude that $v_n - \mu h_n \in H_{n-1}$ for certain $\mu \in B$. Moreover, $h'_n \perp H_{n-1}$. Hence and by (5.22) we obtain

$$0 = \langle v_n - \mu h_n | h'_n \rangle = \langle v_n | h'_n \rangle - \mu \langle h_n | h'_n \rangle = -\mu \langle h_n | h'_n \rangle.$$

This together with (5.24) yields $\mu = 0$, and consequently $v_n \in H_{n-1}$. Moreover, from (5.22) it follows that $h_n - v_n \perp H_{n-1}$. Therefore

$$v_k \in H_{k-1} \quad \text{and} \quad h_k - v_k \perp H_{k-1}, \quad k \in \mathbb{Z}_{1,p}, \tag{5.25}$$

which means that each v_k is an orthogonal projection of h_k onto H_{k-1}. Then

$$0 = \langle h_k - v_k | h'_l \rangle = \langle h_k | h'_l \rangle - \langle v_k | h'_l \rangle, \quad k \in \mathbb{Z}_{1,p}, \, l \in \mathbb{Z}_{1,k-1},$$

and, consequently,

$$\langle h_k | h'_l \rangle = \langle v_k | h'_l \rangle, \quad k \in \mathbb{Z}_{1,p}, \, l \in \mathbb{Z}_{1,k-1}. \tag{5.26}$$

We now define the following two matrixes

$$\mathbb{Z}_{1,p} \times \mathbb{Z}_{1,p} \ni (k, l) \mapsto R_{k,l} := \langle h_k - v_k | h_l \rangle \tag{5.27}$$

as well as

$$\mathbb{Z}_{1,p} \ni k \mapsto Q_k := g^*(h_k - v_k). \tag{5.28}$$

By (5.11), $h'_1 \in H_1 \setminus \Theta$. Then both the vectors h_1 and h'_1 are collinear, and by (5.22) we get $v_1 = \theta$. Applying now the formulas (5.28) and (5.27) we obtain

$$Q_1 = g^*(h_1) \quad \text{and} \quad R_{1,l} = \langle h_1 | h_l \rangle, \quad l \in \mathbb{Z}_{1,p}. \tag{5.29}$$

Now the task is to find recursive formulas in order to determine the remaining terms.
Applying formulas (5.22) and (5.27) we obtain

$$R_{k,l} = \frac{\langle h_k | h'_k \rangle \langle h'_k | h_l \rangle}{\|h'_k\|^2}, \quad k, l \in \mathbb{Z}_{1,p}. \tag{5.30}$$

Since the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h'_k$ is orthogonal, we have

$$v_k = \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\|h'_s\|^2} h'_s, \quad k \in \mathbb{Z}_{2,p}. \tag{5.31}$$

From this, (5.24) and (5.30) we conclude that for all $k \in \mathbb{Z}_{2,p}$ and $l \in \mathbb{Z}_{1,p}$,

$$\begin{aligned} R_{k,l} &= \left\langle h_k - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\|h'_s\|^2} h'_s \middle| h_l \right\rangle = \langle h_k | h_l \rangle - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\|h'_s\|^2} \langle h'_s | h_l \rangle \\ &= \langle h_k | h_l \rangle - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\langle h_s | h'_s \rangle} \frac{\langle h_s | h'_s \rangle \langle h'_s | h_l \rangle}{\|h'_s\|^2} = \langle h_k | h_l \rangle - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\langle h_s | h'_s \rangle} R_{s,l}. \end{aligned}$$

Moreover, by (5.26) we have

$$\frac{\langle v_k | h'_s \rangle}{\langle h_s | h'_s \rangle} = \frac{\langle h_k | h'_s \rangle \langle h'_s | h_s \rangle}{\|h'_s\|^2} \cdot \frac{\|h'_s\|^2}{\langle h_s | h'_s \rangle \langle h'_s | h_s \rangle} = \frac{\overline{R_{s,k}}}{R_{s,s}}, \quad k \in \mathbb{Z}_{2,p}, \ s \in \mathbb{Z}_{1,k-1}. \tag{5.32}$$

Therefore

$$R_{k,l} = \langle h_k | h_l \rangle - \sum_{s=1}^{k-1} \frac{\overline{R_{s,k}} R_{s,l}}{R_{s,s}}, \quad k \in \mathbb{Z}_{2,p}, l \in \mathbb{Z}_{1,p}. \tag{5.33}$$

Combining (5.11) with (5.22) and (5.24) we conclude that

$$h_k - v_k = \frac{\langle h_k | h'_k \rangle}{\|h'_k\|^2} h'_k \in H_k \cap H_{k-1}^\perp \backslash \Theta, \quad k \in \mathbb{Z}_{1,p}. \tag{5.34}$$

From this, (5.28) and (5.31) it follows that for every $k \in \mathbb{Z}_{2,p}$,

$$\begin{aligned} Q_k &= g^*(h_k) - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\|h'_s\|^2} g^*(h'_s) = g^*(h_k) - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\|h'_s\|^2} g^* \left(\frac{\|h'_s\|^2}{\langle h_s | h'_s \rangle} (h_s - v_s) \right) \\ &= g^*(h_k) - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\langle h_s | h'_s \rangle} g^*(h_s - v_s) = g^*(h_k) - \sum_{s=1}^{k-1} \frac{\langle v_k | h'_s \rangle}{\langle h_s | h'_s \rangle} Q_s, \end{aligned}$$

which together with (5.32) leads to

$$Q_k = g^*(h_k) - \sum_{s=1}^{k-1} \frac{\overline{R_{s,k}}}{R_{s,s}} Q_s, \quad k \in \mathbb{Z}_{2,p}. \tag{5.35}$$

Applying now (5.34), Lemma 5.2 and Corollary 5.3 we see that

$$f := \sum_{k=1}^{p} \frac{\overline{g^*(h_k - v_k)}}{\|h_k - v_k\|^2} (h_k - v_k) \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g). \qquad (5.36)$$

Now, we wish to find a sequence $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B$ such that

$$f = \sum_{k=1}^{p} \lambda_k h_k.$$

Then (5.25) and (5.27) imply

$$\langle h_l - v_l | f \rangle = \left\langle h_l - v_l \Big| \sum_{k=1}^{p} \lambda_k h_k \right\rangle = \sum_{k=1}^{p} \langle h_l - v_l | \lambda_k h_k \rangle$$

$$= \sum_{k=1}^{p} \overline{\lambda_k} \langle h_l - v_l | h_k \rangle = \sum_{k=l}^{p} \overline{\lambda_k} R_{l,k}, \quad l \in \mathbb{Z}_{1,p}. \qquad (5.37)$$

Since $h_k - v_k \perp h_l - v_l$ as $k, l \in \mathbb{Z}_{1,p}, k \neq l$, we conclude from (5.36) and (5.28) that

$$\langle h_l - v_l | f \rangle = \sum_{k=1}^{p} \frac{g^*(h_k - v_k)}{\|h_k - v_k\|^2} \langle h_l - v_l | h_k - v_k \rangle = g^*(h_l - v_l) = Q_l, \quad l \in \mathbb{Z}_{1,p},$$

which together with (5.37) yields

$$\overline{\lambda_p} = \frac{Q_p}{R_{p,p}} \quad \text{and} \quad \overline{\lambda_l} = \frac{1}{R_{l,l}} \Big(Q_l - \sum_{k=l+1}^{p} \overline{\lambda_k} R_{l,k} \Big), \quad l \in \mathbb{Z}_{1,p-1}. \qquad (5.38)$$

Applying the initial formulas (5.29) and the recursive ones (5.33), (5.35) and (5.38) we can directly compute the coefficients λ_k, $k \in \mathbb{Z}_{1,p}$, of the regression function f in the basis $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ of $\mathcal{F}$. Therefore these formulas are suitable for numerical computation of the regression functions.

In certain cases the following corollary can be an useful tool in order to determine the regression functions.

Corollary 5.5 *Given nonempty linear sets $\mathcal{F}$ and $\mathcal{F}'$ in $\text{H}(\mathfrak{R})$ suppose that $\mathcal{F} \subset \mathcal{F}'$. If $g \in \mathcal{L}_2(\mathfrak{R})$ and $\text{Reg}(\mathcal{F}, \mathfrak{R}_g) \neq \emptyset \neq \text{Reg}(\mathcal{F}' \cap \mathcal{F}^{\perp}, \mathfrak{R}_g)$, then*

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g) + \text{Reg}(\mathcal{F}' \cap \mathcal{F}^{\perp}, \mathfrak{R}_g) \subset \text{Reg}(\mathcal{F}', \mathfrak{R}_g). \qquad (5.39)$$

Proof Setting $\mathcal{F}_1 := \mathcal{F}$ and $\mathcal{F}_2 := \mathcal{F}' \cap \mathcal{F}^\perp$ we deduce from the assumption $\mathcal{F} \subset \mathcal{F}'$ that

$$\mathcal{F}_1 + \mathcal{F}_2 = \mathcal{F}' \quad \text{and} \quad \mathcal{F}_1 \perp \mathcal{F}_2.$$

Then Lemma 4.2 shows that

$$\begin{aligned} \text{Reg}(\mathcal{F}_1, \mathfrak{R}_g) + \text{Reg}(\mathcal{F}_2, \mathfrak{R}_g) &= \mathcal{F}' \cap \text{Reg}(\mathcal{F}_1, \mathfrak{R}_g) + \mathcal{F}' \cap \text{Reg}(\mathcal{F}_2, \mathfrak{R}_g) \\ &\subset \mathcal{F}' \cap \big(\text{Reg}(\mathcal{F}_1, \mathfrak{R}_g) + \text{Reg}(\mathcal{F}_2, \mathfrak{R}_g)\big) \subset \text{Reg}(\mathcal{F}', \mathfrak{R}_g), \end{aligned}$$

which proves the inclusion (5.39). $\qquad\square$

We end this section with a few simple observations arrising from Corollary 5.3 on the structure of the set consisting of all regression functions.

Remark 5.6 Under the assumptions of Corollary 5.3, we have

$$\theta \in \Theta \cap \mathcal{F},$$

and so the equality (5.20) yields

$$f := \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g) \quad \text{and} \quad \text{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta \cap \mathcal{F}) + f, \quad (5.40)$$

provided $p \in \mathbb{N}$. Since $\Theta \cap \mathcal{F}$ is a linear set, the second equality in (5.40) shows that the class $\text{Reg}(\mathcal{F}, \mathfrak{R}_g)$ forms an affine variety in the space $H(\mathfrak{R})$.

Moreover from (5.40) we can easily deduce that the following properties are pairwise equivalent:

(i) f is a unique regression function in $\mathcal{F}$ with respect to $\mathfrak{R}_g$;
(ii) $\Theta \cap \mathcal{F} = \{\theta\}$;
(iii) $\|h\| > 0$ for every $h \in \mathcal{F} \backslash \{\theta\}$;
(iv) $\|h\| = 0 \implies h = \theta$ for every $h \in \mathcal{F}$.

If additionally the sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{F} \backslash \Theta$ satisfies the orthogonality condition (5.3), then the formulas (5.19) yield

$$h'_k = h_k \quad \text{as } k \in \mathbb{Z}_{1,p},$$

and consequently the property (5.40) remains valid after replacing h'_k by h_k as $k \in \mathbb{Z}_{1,p}$.

According to (5.40) the class $\text{Reg}(\mathcal{F}, \mathfrak{R}_g)$ is determined by the sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto \overline{g^*(h'_k)} \|h'_k\|^{-2} h'_k.$$

We call it the *regression functions sequence* (RFS) generated by a linearly independent sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathcal{L}_1(\mathfrak{R})$ satisfying the equality

$$\Theta \cap \mathcal{F} = \{\theta\},$$

where

$$\mathcal{F} := \lim(\{h_k : k \in \mathbb{Z}_{1,p}\})$$

is a functional model of $\mathfrak{R}_g$.

6 Examples

It is worth noting that our approach to the regression theory is very flexible. We provide an universal and simple theory covering classical cases of regressions where the theoretic functional model $\mathcal{F}$ is spanned by polynomials, trigonometric polynomials and other specific functions; cf. e.g. [17] and [4]. Moreover, we study the regression functions with respect to the wide range of the regression structures $\mathfrak{R}$, involving the generalized quadratic deviation (2.2) by means of certain measures μ. This simplifies much theoretical considerations on the ground of pseudo-Hilbert spaces. On the other hand side we gain the possibility of using the modified least squares method which can be more adequate in more specific situations.

In Example 1.1 the classical least squares method was used. According to the equality (2.4) in Example 2.2, this is a special case of the criterion δ with the measure μ satisfying (2.3) and (2.5). In what follows we present an example which motivate using a more sophisticated measure μ.

Example 6.1 Following Example 1.1 we want to determine now the electric circuit resistance R by means of measurements samples of intensity and voltage represented by two sequences $\mathbb{Z}_{0,n} \ni k \mapsto i_k$ and $\mathbb{Z}_{0,m} \ni k \mapsto v_k$ for some $n, m \in \mathbb{N}$. Assume that all measurements were made independently. Given a precision rate let ρ'_k be the probability that the intensity sample i_k satisfies the precision rate for $k \in \Omega_1 := \mathbb{Z}_{0,n}$ and let ρ''_l be the probability that the voltage sample v_l satisfies the precision rate for $l \in \Omega_2 := \mathbb{Z}_{0,m}$. As in Example 1.1 we consider the regression structure $\mathfrak{R}$, where $A := \mathbb{R}$, $B := \mathbb{R}$, the empiric data functions are defined by

$$\mathbb{Z}_{0,n} \ni k \mapsto x(k) := i_k \quad \text{and} \quad \mathbb{Z}_{0,m} \ni k \mapsto y(k) := v_k,$$

and as the deviation criterion δ we take the generalized quadratic deviation given by (2.2). Following Example 2.2 we define the measure μ as a unique measure satisfying the equalities (2.3). Now, we shall need to define the numbers $\rho_{k,l}$ for $k \in \Omega_1$ and $l \in \Omega_2$. Obviously, we can do it in many ways. In our particular case all measurements were made independently, so it seems to be natural to set

$$\rho_{k,l} := \rho'_k \cdot \rho''_l, \quad k \in \Omega_1, \, l \in \Omega_2. \tag{6.1}$$

Then each coefficient $\rho_{k,l}$ is equal to the probability of the event that both the measurement samples i_k and v_l satisfy simultaneously the prescribed precision rate. As a matter of fact the coefficient $\rho_{k,l}$ reflects accuracy of the measurement samples i_k and v_l, and thereby reflects accuracy of the measurements devices used for getting these samples. If a coefficient $\rho_{k,l}$ is closer to 1, then intuitively the corresponding

pair (i_k, v_l) of samples is more valuable for us. Therefore the generalized quadratic deviation criterion δ, defined by (6.1), seems to be more natural in this case as compared to the classical least squares method, where all samples (i_k, v_k) are treated equivalently and that the samples of the form (i_k, v_l) as $k \neq l$, are not considered at all.

As in Example 1.1, we consider the theoretic functional model $\mathcal{F}$ represented by linear functions $\mathbb{R} \ni t \mapsto rt$ for $r \in \mathbb{R}$. Then $\mathcal{F} = \mathrm{lin}(\{h_1\})$ where h_1 is the identity mapping on $\mathbb{R}$, i.e. $h_1(t) = t$ for $t \in \mathbb{R}$. Thus we can apply our theory from the previous sections in order to determine all regression functions in $\mathcal{F}$ with respect to $\mathfrak{R}$. The condition (2.8) obviously holds for every function $f : A \to B$, which means that $\mathcal{L}_1(\mathfrak{R}) = (\mathbb{R} \to \mathbb{R})$. From (2.12) we have

$$\|h_1\|^2 = \int_{\Omega_1 \times \Omega_2} |(h_1 \circ x)(t_1)|^2 \mathrm{d}\mu(t_1, t_2) = \sum_{k=0}^{n} \sum_{l=0}^{m} i_k^2 \rho_{k,l}. \tag{6.2}$$

It is also easily seen that each function $g : B \to B$ satisfies the condition (2.9), and so $\mathcal{L}_2(\mathfrak{R}) = (\mathbb{R} \to \mathbb{R})$. From (2.22) it follows that

$$g^*(h_1) = \int_{\Omega_1 \times \Omega_2} (h_1 \circ x)(t_1) \overline{g \circ y(t_2)} \mathrm{d}\mu(t_1, t_2) = \sum_{k=0}^{n} \sum_{l=0}^{m} i_k g(v_l) \rho_{k,l}. \tag{6.3}$$

Assume that $\|h_1\| = 0$. Then $\mathcal{F} \subset \Theta$, which implies, by Corollary 3.2, that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = \mathcal{F}.$$

Suppose for simplicity that $\rho_{k,l} > 0$ for $k \in \Omega_1$ and $l \in \Omega_2$. From (6.2) we see that $\|h_1\| = 0$ iff $i_k = 0$ for $k \in \Omega_1$. Thus the equality $\|h_1\| = 0$ means that the current intensity vanishes (current does not flow) or the current intensity is below the sensitivity of intensity measurements devices. In both the cases we are not able to determine the resistance R. This provides a natural interpretation of the equality $\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = \mathcal{F}$.

Assume now that $\|h_1\| \neq 0$. Then $h_1 \in \mathcal{F} \backslash \Theta$, and so $\Theta \cap \mathcal{F} = \{\theta\}$. Theorem 3.3 (or directly Theorem 5.1) now leads to

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = (\Theta \cap \mathcal{F}) + \frac{g^*(h_1)}{\|h_1\|^2} h_1 = \left\{ \frac{g^*(h_1)}{\|h_1\|^2} h_1 \right\}, \tag{6.4}$$

which means that the set $\mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ consists of the unique regression function

$$\mathbb{R} \ni t \to \frac{g^*(h_1)}{\|h_1\|^2} h_1(t) = \frac{g^*(h_1)}{\|h_1\|^2} t.$$

Combining this with (6.2) and (6.3) we can uniquely determine the resistance

$$R = \frac{g^*(h_1)}{\|h_1\|^2} = \sum_{k=0}^{n} \sum_{l=0}^{m} i_k g(v_l) \rho_{k,l} \bigg/ \sum_{k=0}^{n} \sum_{l=0}^{m} i_k^2 \rho_{k,l}. \tag{6.5}$$

Note that if $n = m$, g is the identity function and the coefficients $\rho_{k,l}$ are defined by (2.5), then (6.5) yields (1.6). Such a situation naturally corresponds to the sequence $\mathbb{Z}_{0,n} \ni k \to (i_k, v_k)$ of $n + 1$ simultaneous measurements of the current intension and voltage with the same precision.

The following example illustrates the usage of Corollary 5.3 in the case where the theoretic functional model $\mathcal{F}$ is spanned by two functions.

Example 6.2 Given a regression structure $\mathfrak{R}$ with $g \in \mathcal{L}_2(\mathfrak{R})$ let us consider the case where the functional model $\mathcal{F} = \mathrm{lin}(\{h_1, h_2\})$ is spanned by two linearly independent functions $h_1, h_2 \in \mathcal{L}_1(\mathfrak{R})$, such that $\mathcal{F} \cap \Theta = \{\theta\}$. Applying Corollary 5.3 we see that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = (\Theta \cap \mathcal{F}) + \sum_{k=1}^{2} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k, \tag{6.6}$$

where according to (5.19),

$$h'_1 := h_1 \quad \text{and} \quad h'_2 := h_2 - \frac{\langle h_2 | h'_1 \rangle}{\|h'_1\|^2} h'_1 = h_2 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} h_1. \tag{6.7}$$

Hence, $h'_2 \perp h_1$, and consequently

$$\|h'_2\|^2 = \|h_2\|^2 - \left\| \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} h_1 \right\|^2 = \|h_2\|^2 - \frac{|\langle h_2 | h_1 \rangle|^2}{\|h_1\|^2}. \tag{6.8}$$

Setting

$$a_2 := \frac{\overline{g^*(h'_2)}}{\|h'_2\|^2} \quad \text{and} \quad a_1 := \frac{\overline{g^*(h'_1)}}{\|h'_1\|^2} - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} a_2 \tag{6.9}$$

we conclude from (6.6) and (6.7) that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}) = (\Theta \cap \mathcal{F}) + a_2 h_2 + a_1 h_1. \tag{6.10}$$

Combining (6.9) with (6.7) and (6.8) we obtain

$$a_2 = \frac{\overline{g^*(h_2)} \|h_1\|^2 - \overline{g^*(h_1)} \langle h_2 | h_1 \rangle}{\|h_2\|^2 \|h_1\|^2 - |\langle h_2 | h_1 \rangle|^2} \quad \text{and} \quad a_1 = \frac{\overline{g^*(h_1)} - \langle h_2 | h_1 \rangle a_2}{\|h_1\|^2}. \tag{6.11}$$

In particular, if $h_2(t) = t$, $h_1(t) = 1$ and $g(t) = t$ as $t \in \mathbb{R}$, and the regression structure $\mathfrak{R}$ is defined as in Example 2.2 under the assumption that $m = n$ and the coefficients $\rho_{k,l}$ satisfy (2.5), then the equalities in (6.11) yield $a_2 = a_0$ and $a_1 = b_0$ where a_0 and b_0 are defined in (1.2).

The next example deals with applications to econometric sciences.

Example 6.3 Consider a linear econometric model of the following form

$$Y = a_0 + \sum_{k=1}^{p} a_k X_k, \tag{6.12}$$

for some $p \in \mathbb{N}$ and a sequence of coefficients $\mathbb{Z}_{0,p} \ni k \mapsto a_k \in \mathbb{R}$. This model describes the theoretic dependence the variable Y on the variables $X_1, X_2, \ldots, X_p$, which represent certain economical parameters; cf. e.g. [7, pp. 89–130], [9, pp.127–208], [3, pp. 73–93]. We seek the best such model for a given $n \in \mathbb{N}$ and experimental data series $\mathbb{Z}_{1,n} \ni k \mapsto \tilde{y}_k \in \mathbb{R}$ and $\mathbb{Z}_{1,n} \ni k \mapsto \tilde{x}_{k,l} \in \mathbb{R}$ for $l \in \mathbb{Z}_{1,p}$, with respect to the classical quadratic deviation. To be more precise we seek a sequence of coefficients $\mathbb{Z}_{0,p} \ni k \mapsto \hat{a}_k \in \mathbb{R}$ satisfying the condition

$$\sum_{k=1}^{n} \left(\hat{a}_0 + \sum_{l=1}^{p} \hat{a}_k \tilde{x}_{k,l} - \tilde{y}_k \right)^2 \leq \sum_{k=1}^{n} \left(a_0 + \sum_{l=1}^{p} a_k \tilde{x}_{k,l} - \tilde{y}_k \right)^2 \tag{6.13}$$

for every $\mathbb{Z}_{0,p} \ni k \mapsto a_k \in \mathbb{R}$. It shows that this problem can be reduced to the following regression problem. Let

$$\mathfrak{R} := (A, B, \delta; x, y)$$

be an asynchronous regression structure, where:

$$A := \mathbb{R}^p, \quad B := \mathbb{R}, \quad \Omega_1 := \mathbb{Z}_{1,n}, \quad \Omega_2 := \mathbb{Z}_{1,n}.$$

The functions $x : \Omega_1 \to A$ and $y : \Omega_1 \to B$ are defined by the formulas

$$x(k) := (\mathbb{Z}_{1,p} \ni l \mapsto \tilde{x}_{k,l}) \quad \text{and} \quad y(k) := \tilde{y}_k, \quad k \in \mathbb{Z}_{1,n},$$

and the deviation criterion δ is defined by the formula (2.6). As a functional model $\mathcal{F}$ of $\mathfrak{R}$ we set

$$\mathcal{F} := \text{lin}(\{h_l : l \in \mathbb{Z}_{0,p}\}),$$

where $h_0(v) := 1$ and $h_l(v) := v_l$ as $l \in \mathbb{Z}_{1,p}$, for every

$$v = (\mathbb{Z}_{1,p} \ni k \mapsto v_k) \in \mathbb{R}^p.$$

Then for each sequence of coefficients $\mathbb{Z}_{0,p} \ni k \mapsto a_k \in \mathbb{R}$, the function

$$f := \sum_{l=0}^{p} a_k h_k \in \mathcal{F},$$

and by (6.13) we see that

$$\delta(f \circ x, y) = \sum_{k=1}^{n} (f \circ x(k) - y(k))^2 = \sum_{k=1}^{n} \left(\sum_{l=0}^{p} a_l h_l \circ x(k) - y(k) \right)^2$$

$$= \sum_{k=1}^{n} \left(a_0 + \sum_{l=1}^{p} a_k \tilde{x}_{k,l} - \tilde{y}_k \right)^2 \geq \sum_{k=1}^{n} \left(\hat{a}_0 + \sum_{l=1}^{p} \hat{a}_k \tilde{x}_{k,l} - \tilde{y}_k \right)^2$$

$$= \sum_{k=1}^{n} \left(\sum_{l=0}^{p} \hat{a}_l h_l \circ x(k) - y(k) \right)^2 = \sum_{k=1}^{n} \left(\hat{f} \circ x(k) - y(k) \right)^2$$

$$= \delta(\hat{f} \circ x, y),$$

where

$$\hat{f} := \sum_{l=0}^{p} \hat{a}_k h_k \in \mathcal{F}.$$

Thus

$$\hat{f} \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$$

iff $\mathbb{Z}_{0,p} \ni k \mapsto \hat{a}_k \in \mathbb{R}$ is the best choice of coefficients. This way the optimal linear model (6.12) can be expressed by $(p+1)$-dimensional regression functions $\hat{f} \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$.

The next example deals in a very natural way with theoretical functional model $\mathcal{F}$ supported by complex-valued functions.

Example 6.4 Assume that a point P runs along an elliptic trajectory with a constant radial speed ω. We want to describe the trajectory by means of the location measurements samples of the point P. It is quite convenient to use here the complex plane $\mathbb{C}$, because the elliptic trajectory has a simple representation

$$f(t) := a + be^{i\omega t} + ce^{-i\omega t}, \quad t \in \mathbb{R}, \tag{6.14}$$

for certain $a, b, c \in \mathbb{C}$. Consider an asynchronous regression structure of the form $\mathfrak{R} := (A, B, \delta; x, y)$, where: $A := \mathbb{R}$ and $B := \mathbb{C}$. As a functional model $\mathcal{F}$ of $\mathfrak{R}$ we set

$$\mathcal{F} := \mathrm{lin}(\{h_0, h_1, h_2\}),$$

where $h_0(t) := 1$, $h_1(t) := e^{i\omega t}$ and $h_2(t) := e^{-i\omega t}$ for every $t \in \mathbb{R}$. From (6.14) it follows that the optimal trajectory is a 3-dimensional regression $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$. In particular, if the deviation criterion δ is defined by the formula (2.6), then the optimal trajectory $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{R})$ is best fitted to the empirical data functions $x : \mathbb{Z}_{0,n} \to \mathbb{R}$ and $y : \mathbb{Z}_{0,n} \to \mathbb{C}$ with respect to the classical square deviation. Note that a 2-dimensional regression $f \in \mathrm{Reg}(\mathcal{F}', \mathfrak{R})$, where

$$\mathcal{F}' := \mathrm{lin}(\{h_0, h_1\}),$$

is the best chosen circular trajectory.

7 The synchronous regression structures

We have considered so far the asynchronous regression structures of the form $\mathfrak{R} :=$ $(A, B, \delta; x, y)$ where the empirical data functions x and y were depended on two different parameters t_1 and t_2, respectively. In the other words the functions x and y were defined asynchronously. The dependence between these functions was given by a measure μ; see Example 6.1. In this section we define another type of regression structures where, roughly speaking, the empirical data functions x and y depend on one parameter t, i.e. they are defined synchronously. To be more precise, we consider regression structures satisfying the following three conditions:

III.1 $\Omega_1 = \Omega_2$;

III.2 $B = \mathbb{R}$ or $B = \mathbb{C}$;

III.3 *There exist a σ-field $\mathcal{A}$ of subsets of the set $\Omega := \Omega_1$ and a measure $\nu : \mathcal{A} \to [0; +\infty]$ such that the function δ satisfies the following equality*

$$\delta(u, v) = \int_{\Omega} |u(t) - v(t)|^2 \mathrm{d}\nu(t), \tag{7.1}$$

provided the function $|u - v|$ is $\mathcal{A}$-measurable, and $\delta(u, v) = +\infty$, otherwise.

Then the regression problem for $\mathfrak{R}$ means the extremal problem of determining all functions $f_0 \in \mathcal{F}$ minimising the functional F satisfying, according to (1.3) and (7.1), the following equality

$$F(f) = \int_{\Omega} |f \circ x(t) - y(t)|^2 \mathrm{d}\nu(t), \quad f \in \mathcal{F}. \tag{7.2}$$

Definition 7.1 Any regression structure $\mathfrak{R}$ satisfying the conditions III.1, III.2 and III.3 is said to be the *synchronous regression structure*; *real synchronous regression structure* as $B = \mathbb{R}$ and *complex synchronous regression structure* as $B = \mathbb{C}$.

Remark 7.2 The above type of regression structures corresponds to the classical regression theory, where $\Omega := \mathbb{Z}_{0,n}$ for certain $n \in \mathbb{N}$ and the measure ν defined on the family $\mathcal{A}$ of all subsets of Ω is such that $\nu(\{k\}) = 1$ for $k \in \Omega$. Then by (7.2),

$$\int_{\Omega} |f \circ x(t) - y(t)|^2 \mathrm{d}\nu(t) = \sum_{k=0}^{n} |f(x_k) - y_k|^2, \tag{7.3}$$

where $x_k := x(k)$ and $y_k := y(k)$ for $k \in \Omega$. Thus the deviation criterion δ coincides with the classical quadratic deviation.

Analyzing our considerations from the previous sections it is easy to see that we could develop, in much the same way, an adequate theory of the synchronous regression structures. However, we will handle this case in a different, more interesting manner. We will show that, in fact, all the properties of the regression functions with respect to the synchronous regression structure $\mathfrak{R}$ can be derived from those with respect to certain asynchronous regression structure $\mathfrak{R}^*$ associated with $\mathfrak{R}$. Therefore the regression theory for asynchronous regression structures, embraces in this sense the one for synchronous regression structures.

In the discrete case described in Remark 7.2 we can easily associate an asynchronous regression structure considered in Example 2.2 by putting (2.5). Then (2.6) yields (7.1). Therefore in such a case synchronous regression structures can be treated as a special case of asynchronous regression structures where the measure μ is focused to the diagonal of $\Omega_1 \times \Omega_2$. Therefore this method of reduction of asynchronous regression structure to synchronous regression structure will be called the *diagonal method*. It is possible to find a variant of the diagonal method in the non-discrete case as well, but it is more difficult task. We will handle this problem in what follows now.

First of all we recall the following, useful in the sequel, fact; cf. e.g. [2, Thm. 1.6.12].

Lemma 7.3 *For every measurable space $(\Omega, \mathcal{A}, \mu)$ and every function $\varphi : \Omega \to \Omega'$, the structure $(\Omega', \mathcal{A}_\varphi, \mu_\varphi)$ is also a measurable space, where*

$$\mathcal{A}_\varphi := \{S \in 2^{\Omega'} : \varphi^{-1}(S) \in \mathcal{A}\} \quad and \quad \mu_\varphi(S) := \mu(\varphi^{-1}(S)), \quad S \in \mathcal{A}_\varphi.$$

Moreover, for every $\mathcal{A}_\varphi$-measurable function $h : \Omega' \to \mathbb{C}$,

$$h \in \mathrm{L}^1(\Omega', \mathcal{A}_\varphi, \mu_\varphi) \Longleftrightarrow h \circ \varphi \in \mathrm{L}^1(\Omega, \mathcal{A}, \mu)$$

as well as

$$\int_{\varphi(\Omega)} h \mathrm{d}\mu_\varphi = \int_\Omega h \circ \varphi \mathrm{d}\mu, \quad h \in \mathrm{L}^1(\Omega', \mathcal{A}_\varphi, \mu_\varphi). \tag{7.4}$$

Consider a synchronous regression structure $\mathfrak{R} = (A, B, \eta; x, y)$, where the deviation criterion η satisfies the equality (7.1) instead of δ for a given measurable space $(\Omega, \mathcal{A}, \nu)$. Let $\mathcal{A} \otimes \mathcal{A}$ be the smallest σ-field containing all sets $U \times V$ for $U, V \in \mathcal{A}$. It is evident that

$$\Omega \ni t \mapsto \phi(t) := (t, t)$$

is an injective mapping of Ω onto the diagonal $D := \{(t, t) : t \in \Omega\}$.

Lemma 7.4 *The set*

$$\mathcal{B} := \{S_0 \cup (S_1 \cap D) \cup (S_2 \cap D') : S_0, S_1, S_2 \in \mathcal{A} \otimes \mathcal{A}\}, \tag{7.5}$$

where $D' := (\Omega \times \Omega) \setminus D$, is the smallest σ-field in $\Omega \times \Omega$ such that

$$\mathcal{A} \otimes \mathcal{A} \cup \{D\} \subset \mathcal{B}. \tag{7.6}$$

Moreover,

$$\{\phi^{-1}(S \cap D) : S \in \mathcal{B}\} = \mathcal{A} \tag{7.7}$$

and for every set $S \subset \Omega$ the implication

$$(S \times \Omega \in \mathcal{B} \text{ or } \Omega \times S \in \mathcal{B}) \Longrightarrow S \in \mathcal{A} \tag{7.8}$$

holds.

Proof We show first that $\mathcal{B}$ is a σ-field in $\Omega \times \Omega$. Given a sequence $\mathbb{N} \ni k \mapsto S_k \in \mathcal{B}$ we deduce from (7.5) that there exist sequences

$$\mathbb{N} \ni k \mapsto S'_k \in \mathcal{A} \otimes \mathcal{A}, \quad \mathbb{N} \ni k \mapsto S''_k \in \mathcal{A} \otimes \mathcal{A}$$

and

$$\mathbb{N} \ni k \mapsto S'''_k \in \mathcal{A} \otimes \mathcal{A}$$

such that

$$S_k = S'_k \cup (S''_k \cap D) \cup (S'''_k \cap D'), \quad k \in \mathbb{N}. \tag{7.9}$$

Since

$$S' := \bigcup_{k=1}^{\infty} S'_k \in \mathcal{A} \otimes \mathcal{A}, \quad S'' := \bigcup_{k=1}^{\infty} S''_k \in \mathcal{A} \otimes \mathcal{A} \quad \text{and} \quad S''' := \bigcup_{k=1}^{\infty} S'''_k \in \mathcal{A} \otimes \mathcal{A},$$

we can see by (7.5) that

$$\bigcup_{k=1}^{\infty} S_k = \bigcup_{k=1}^{\infty} [S'_k \cup (S''_k \cap D) \cup (S'''_k \cap D')]$$

$$= \bigcup_{k=1}^{\infty} S'_k \cup \left(D \cap \bigcup_{k=1}^{\infty} S''_k \right) \cup \left(D' \cap \bigcup_{k=1}^{\infty} S'''_k \right)$$

$$= S' \cup (S'' \cap D) \cup (S''' \cap D') \in \mathcal{B}.$$

Fix now $S \in \mathcal{B}$. Then

$$S = S' \cup (S'' \cap D) \cup (S''' \cap D')$$

for some S', S'', $S''' \in \mathcal{A} \otimes \mathcal{A}$. Since

$$
\begin{aligned}
S_0 &:= (\Omega \times \Omega \backslash S') \cap (\Omega \times \Omega \backslash S'') \cap (\Omega \times \Omega \backslash S''') \in \mathcal{A} \otimes \mathcal{A}, \\
S_1 &:= (\Omega \times \Omega \backslash S') \cap (\Omega \times \Omega \backslash S'') \in \mathcal{A} \otimes \mathcal{A}, \\
S_2 &:= (\Omega \times \Omega \backslash S') \cap (\Omega \times \Omega \backslash S''') \in \mathcal{A} \otimes \mathcal{A},
\end{aligned}
$$

we deduce from (7.5) that

$$
\begin{aligned}
\Omega \times \Omega \backslash S &= \Omega \times \Omega \backslash [S' \cup (S'' \cap D) \cup (S''' \cap D')] \\
&= [\Omega \times \Omega \backslash S'] \cap [\Omega \times \Omega \backslash (S'' \cap D)] \cap [\Omega \times \Omega \backslash (S''' \cap D')] \\
&= [\Omega \times \Omega \backslash S'] \cap [(\Omega \times \Omega \backslash S'') \cup D')] \cap [(\Omega \times \Omega \backslash S''') \cup D)] \\
&= S_0 \cup (S_1 \cap D) \cup (S_2 \cap D') \in \mathcal{B}.
\end{aligned}
$$

Thus we have showed that $\mathcal{B}$ is a σ-field in $\Omega \times \Omega$. Moreover,

$$
\emptyset \in \mathcal{A} \otimes \mathcal{A} \quad \text{and} \quad \Omega \times \Omega \in \mathcal{A} \otimes \mathcal{A},
$$

which shows that

$$
D = \emptyset \cup (\Omega \times \Omega \cap D) \cup (\emptyset \cap D') \in \mathcal{B}
$$

and

$$
S = S \cup (\emptyset \cap D) \cup (\emptyset \cap D') \in \mathcal{B},
$$

as $S \in \mathcal{A} \otimes \mathcal{A}$, and so the inclusion (7.6) holds.

Suppose now that $\mathcal{B}'$ is a σ-field in $\Omega \times \Omega$ satisfying (7.6) with $\mathcal{B}$ replaced by $\mathcal{B}'$. Then $D' \in \mathcal{B}'$. From (7.5) we conclude that each $S \in \mathcal{B}$ satisfies the equality

$$
S = S' \cup (S'' \cap D) \cup (S''' \cap D')
$$

for some S', S'', $S''' \in \mathcal{A} \otimes \mathcal{A}$. Since $\mathcal{A} \otimes \mathcal{A} \subset \mathcal{B}'$, it follows that $S \in \mathcal{B}'$, and so $\mathcal{B} \subset \mathcal{B}'$. Thus $\mathcal{B}$ is the smallest σ-field in $\Omega \times \Omega$ containing the diagonal D and the family $\mathcal{A} \otimes \mathcal{A}$.

For the proof of the equality (7.7) we consider the family

$$
\mathcal{A}_\phi := \{S \in 2^{\Omega \times \Omega} : \phi^{-1}(S \cap D) \in \mathcal{A}\}. \tag{7.10}
$$

Then for every sequence $\mathbb{N} \ni n \mapsto S_n \in \mathcal{A}_\phi$,

$$\phi^{-1}\left(D \cap \bigcup_{n=1}^{\infty} S_n\right) = \bigcup_{n=1}^{\infty} \phi^{-1}(D \cap S_n) \in \mathcal{A},$$

and so $\bigcup_{n=1}^{\infty} S_n \in \mathcal{A}_\phi$. Furthermore, for every $S \in \mathcal{A}_\phi$,

$$\phi^{-1}((\Omega \times \Omega \backslash S) \cap D) = \phi^{-1}(\Omega \times \Omega \cap D) \backslash \phi^{-1}(S \cap D) = \Omega \backslash \phi^{-1}(S \cap D) \in \mathcal{A},$$

and consequently,

$$\Omega \times \Omega \backslash S \in \mathcal{A}_\phi.$$

Thus $\mathcal{A}_\phi$ is a σ-field in $\Omega \times \Omega$. Since

$$(U \times V) \cap D = \phi(U \cap V), \quad U, V \in \mathcal{A}, \tag{7.11}$$

we see that for all $U, V \in \mathcal{A}$,

$$\phi^{-1}((U \times V) \cap D) = U \cap V \in \mathcal{A},$$

and therefore $U \times V \in \mathcal{A}_\phi$. Hence

$$\mathcal{A} \otimes \mathcal{A} \subset \mathcal{A}_\phi, \tag{7.12}$$

because $\mathcal{A} \otimes \mathcal{A}$ is the smallest σ-field in $\Omega \times \Omega$ containing the set $\mathcal{A} \times \mathcal{A}$. Furthermore,

$$\phi^{-1}(D \cap D) = \phi^{-1}(D) = \Omega,$$

and so $D \in \mathcal{A}_\phi$. This, together with (7.12), yields the inclusion (7.6) with $\mathcal{B}$ replaced by $\mathcal{A}_\phi$. Therefore $\mathcal{B} \subset \mathcal{A}_\phi$, because $\mathcal{B}$ is the smallest σ-field in $\Omega \times \Omega$ containing the set $\mathcal{A} \otimes \mathcal{A} \cup \{D\}$. Then by (7.10) we obtain

$$\{\phi^{-1}(S \cap D) : S \in \mathcal{B}\} \subset \mathcal{A}. \tag{7.13}$$

On the other hand side, $S \times S \in \mathcal{B}$ for any $S \in \mathcal{A}$. From (7.11) it folows that

$$\phi^{-1}(S \times S \cap D) = \phi^{-1}(\phi(S \cap S)) = S, \quad S \in \mathcal{A}.$$

Therefore the inverse inclusion to that in (7.13), and consequently, the equality (7.7) holds.

It remains to prove the implication (7.8). To this end fix a set $S \subset \Omega$. If $S \times \Omega \in \mathcal{B}$, then from (7.11) and (7.7) we obtain

$$S = \phi^{-1}(\phi(S \cap \Omega)) = \phi^{-1}(S \times \Omega \cap D) \in \mathcal{A},$$

which yields the implication

$$S \times \Omega \in \mathcal{B} \Longrightarrow S \in \mathcal{A}.$$

In the same manner we can see that

$$\Omega \times S \in \mathcal{B} \Longrightarrow S \in \mathcal{A}.$$

Both the implications lead to (7.8), which completes the proof.   □

Lemma 7.5 *The function*

$$\mathcal{B} \ni S \mapsto \mu(S) := \nu(\phi^{-1}(S \cap D)) \tag{7.14}$$

is well defined, the structure $(\Omega \times \Omega, \mathcal{B}, \mu)$ *is a measurable space and*

$$\mu(\Omega \times \Omega \backslash D) = 0. \tag{7.15}$$

Proof Since ν is a measure on the σ-field $\mathcal{A}$, we deduce from (7.7) that $\mu : \mathcal{B} \to \overline{\mathbb{R}}$ is a well defined function. Let $\mathbb{N} \ni k \mapsto S_k \in \mathcal{B}$ be a sequence of pairwise disjoint sets. Then for all $k, l \in \mathbb{N}, k \neq l$, we have

$$(S_k \cap D) \cap (S_l \cap D) = (S_k \cap S_l) \cap D = \emptyset \cap D = \emptyset.$$

Thus

$$\mathbb{N} \ni k \mapsto S_k \cap D$$

is a sequence of pairwise disjoint sets. Since ν is a measure on the σ-field $\mathcal{A}$, we deduce from (7.14) and (7.7) that

$$\mu\left(\bigcup_{k=1}^{\infty} S_k\right) = \nu\left(\phi^{-1}\left(D \cap \bigcup_{k=1}^{\infty} S_k\right)\right) = \nu\left(\bigcup_{k=1}^{\infty} \phi^{-1}(D \cap S_k)\right)$$
$$= \sum_{k=1}^{\infty} \nu(\phi^{-1}(D \cap S_k)) = \sum_{k=1}^{\infty} \mu(S_k).$$

Moreover, the following properties hold

$$\mu(S) = \nu(\phi^{-1}(S \cap D)) \geq 0, \quad S \in \mathcal{B};$$
$$\mu(\emptyset) = \nu(\phi^{-1}(\emptyset \cap D)) = \nu(\emptyset) = 0;$$
$$\mu(\Omega \times \Omega \backslash D) = \nu(\phi^{-1}((\Omega \times \Omega \backslash D) \cap D)) = \nu(\emptyset) = 0.$$

Thus μ is a measure on the σ-field $\mathcal{B}$ and the equality (7.15) holds.   □

Lemma 7.6 *Given any functions $h : \Omega \to B$ and $H : \Omega \times \Omega \to B$ suppose that one of the following conditions hold:*

(i) $H(t_1, t_2) = h(t_1)$ *for $t_1, t_2 \in \Omega$;*
(ii) $H(t_1, t_2) = h(t_2)$ *for $t_1, t_2 \in \Omega$;*
(iii) $H \circ \phi(t) = h(t)$ *for $t \in \Omega$ and H is $\mathcal{B}$-measurable.*

Then

$$h \in L^1(\Omega, \mathcal{A}, \nu) \Longleftrightarrow H \in L^1(\Omega \times \Omega, \mathcal{B}, \mu). \tag{7.16}$$

Moreover, if $h \in L^1(\Omega, \mathcal{A}, \nu)$, then

$$\int_\Omega h \, d\nu = \int_{\Omega \times \Omega} H \, d\mu. \tag{7.17}$$

Proof We prove first this lemma in the case where h is a simple function, i.e. h is a $\mathcal{A}$-measurable function with finite set of values. Then there exists $n \in \mathbb{N}$ and sequences $\mathbb{Z}_{1,n} \ni k \mapsto \lambda_k \in B$ and $\mathbb{Z}_{1,n} \ni k \mapsto V_k \in \mathcal{A}$ such that

$$\bigcup_{k=1}^n V_k = \Omega,$$

whereas

$$V_k \cap V_l = \emptyset$$

provided $k \neq l$. Moreover,

$$h = \sum_{k=1}^n \lambda_k \, I_{V_k} . \tag{7.18}$$

Here and subsequently, I_V stands for the characteristic function of a set

$$V \subset \Omega \cup \Omega \times \Omega,$$

i.e. $I_V(t) := 1$ for $t \in V$ and $I_V(t) := 0$ otherwise.
   Assume that the condition (i) holds. Then by (7.18),

$$H = \sum_{k=1}^n \lambda_k \, I_{V_k \times \Omega} .$$

Therefore by (7.18), (7.14) and the properties of Lebesgue's integral we deduce that

$$\int_\Omega |h| \mathrm{d}\nu = \int_\Omega \sum_{k=1}^n |\lambda_k| \, \mathrm{I}_{V_k} \, \mathrm{d}\nu = \sum_{k=1}^n |\lambda_k| \int_\Omega \mathrm{I}_{V_k} \, \mathrm{d}\nu$$

$$= \sum_{k=1}^n |\lambda_k| \nu(V_k) = \sum_{k=1}^n |\lambda_k| \nu(\phi^{-1}(V_k \times \Omega \cap D))$$

$$= \sum_{k=1}^n |\lambda_k| \mu(V_k \times \Omega) = \sum_{k=1}^n |\lambda_k| \int_{\Omega \times \Omega} \mathrm{I}_{V_k \times \Omega} \, \mathrm{d}\mu$$

$$= \int_{\Omega \times \Omega} \sum_{k=1}^n |\lambda_k| \, \mathrm{I}_{V_k \times \Omega} \, \mathrm{d}\mu = \int_{\Omega \times \Omega} |H| \mathrm{d}\mu. \qquad (7.19)$$

This shows, by Lemma 7.4, that the condition (i) implies (7.16). In the much similar way we show that the condition (ii) also implies (7.16). If $h \in \mathrm{L}^1(\Omega, \mathcal{A}, \nu)$, then

$$\sum_{k=1}^n |\lambda_k| \nu(V_k) < +\infty,$$

and consequently, $\nu(V_k) < +\infty$ provided $|\lambda_k| > 0$ as $k \in \mathbb{Z}_{1,n}$. Replacing $|h|$ by h in (7.19) we derive the equality (7.17) in the case where h is a simple function. Using now the standard approximation technique for integrable functions by the simple ones we extend the implications (i) $\implies$ (7.16) and (ii) $\implies$ (7.16) as well as the equality (7.17) to the general case of arbitrary functions $h : \Omega \to B$ and $H : \Omega \times \Omega \to B$; for details cf. e.g. [2, Sec. 1.6].

Assume now that the condition (iii) holds. From (7.15) it follows that

$$\int_{\Omega \times \Omega \setminus D} |H| \mathrm{d}\mu = 0, \qquad (7.20)$$

and so H is integrable on $\Omega \times \Omega$ iff it is integrable on the diagonal D. Since $H \circ \phi = h$, we conclude from Lemma 7.3 that H is integrable on D iff $h \in \mathrm{L}^1(\Omega, \mathcal{A}, \nu)$. Thus the condition (iii) implies (7.16). Furthermore, if $h \in \mathrm{L}^1(\Omega, \mathcal{A}, \nu)$, then from (7.7), (7.14), (7.20) and Lemma 7.3 it follows that

$$\int_\Omega h \mathrm{d}\nu = \int_\Omega H \circ \phi \mathrm{d}\nu = \int_D H \mathrm{d}\nu_\phi$$

$$= \int_D H \mathrm{d}\mu + \int_{\Omega \times \Omega \setminus D} H \mathrm{d}\mu = \int_{\Omega \times \Omega} H \mathrm{d}\mu.$$

This gives (7.17), and the proof is complete.   □

We are now in a position to show the fundamental result in this section.

Theorem 7.7 *Suppose that a measurable space* $(\Omega, \mathcal{A}, \nu)$ *determines a synchronous regression structure* $\mathfrak{R} = (A, B, \eta; x, y)$. *Then the measurable space* $(\Omega \times \Omega, \mathcal{B}, \mu)$, *defined in* Lemma 7.5, *determines an asynchronous regression structure* $\mathfrak{R}^* = (A, B, \delta; x, y)$, *and the following properties hold:*

(i) $\mathcal{L}_1(\mathfrak{R}^*) = L^2(A, \mathcal{A}_x, \nu_x)$ and

$$\langle h|g \rangle = \langle h|g \rangle_{\nu_x} = \langle h \circ x | g \circ x \rangle_\nu, \quad h, g \in \mathcal{L}_1(\mathfrak{R}^*); \tag{7.21}$$

$$\|h\| = \|h\|_{\nu_x} = \|h \circ x\|_\nu, \quad h \in \mathcal{L}_1(\mathfrak{R}^*); \tag{7.22}$$

(ii) $\mathcal{L}_2(\mathfrak{R}^*) = L^2(B, \mathcal{A}_y, \nu_y)$ and

$$g^*(h) = g_\nu^*(h) := \int_\Omega (h \circ x)\overline{(g \circ y)}\mathrm{d}\nu, \quad h \in \mathcal{L}_1(\mathfrak{R}^*), \, g \in \mathcal{L}_2(\mathfrak{R}^*); \tag{7.23}$$

(iii) $\delta(u, v) = \eta(u, v)$ for all $\mathcal{A}$-measurable functions $u, v : \Omega \to B$.

Moreover, for each $g \in \mathcal{L}_2(\mathfrak{R}^*)$ and any nonempty linear set $\mathcal{F} \subset \mathcal{L}_1(\mathfrak{R}^*)$,

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = \mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g^*). \tag{7.24}$$

Proof Applying the implication (i) $\implies$ (7.16) from Lemma 7.6 with h replaced by $|h|^2$ we get

$$\mathcal{L}_1(\mathfrak{R}^*) = L^2(A, \mathcal{A}_x, \nu_x).$$

Given $h, g \in L^2(A, \mathcal{A}_x, \nu_x)$ we conclude from (2.10) that

$$h\overline{g} \in L^1(A, \mathcal{A}_x, \nu_x).$$

Then Lemma 7.3 shows that

$$(h\overline{g}) \circ x \in L^1(\Omega, \mathcal{A}, \nu)$$

and

$$\langle h \circ x | g \circ x \rangle_\nu = \int_\Omega (h\overline{g}) \circ x \mathrm{d}\nu = \int_{x(\Omega)} h\overline{g}\mathrm{d}\nu_x = \int_A h\overline{g}\mathrm{d}\nu_x = \langle h|g \rangle_{\nu_x}. \tag{7.25}$$

Setting

$$H(t_1, t_2) := (h\overline{g}) \circ x(t_1) \quad \text{for all } t_1, t_2 \in \Omega$$

we see by the implication (i) $\implies$ (7.16) from Lemma 7.6 that

$$H \in L^1(\Omega \times \Omega, \mathcal{B}, \mu).$$

From (7.17) it follows that

$$\langle h|g \rangle = \int_{\Omega \times \Omega} h \circ x(t_1)\overline{g \circ x(t_1)}\mathrm{d}\mu(t_1, t_2) = \int_{\Omega \times \Omega} H\mathrm{d}\mu = \int_\Omega (h\overline{g}) \circ x\mathrm{d}\nu,$$

which together with (7.25) yields (7.21). By this we have

$$\|h\|^2 = \langle h|h\rangle = \langle h|h\rangle_{\nu_x} = \|h\|^2_{\nu_x},$$

and

$$\langle h|h\rangle_{\nu_x} = \langle h \circ x|h \circ x\rangle_{\nu} = \|h \circ x\|^2_{\nu},$$

so (7.22) is proved.

Applying the implication (ii) $\Longrightarrow$ (7.16) from Lemma 7.6 with h replaced by $|h|^2$ we get

$$\mathcal{L}_2(\mathfrak{R}^*) = L^2(B, \mathcal{A}_y, \nu_y).$$

Fix $h \in L^2(A, \mathcal{A}_x, \nu_x)$ and $g \in L^2(B, \mathcal{A}_y, \nu_y)$. Then

$$h \circ x, g \circ y \in L^2(\Omega, \mathcal{A}, \nu),$$

and by (2.10) we see that

$$(h \circ x)\overline{(g \circ y)} \in L^1(\Omega, \mathcal{A}, \nu).$$

Setting

$$H(t_1, t_2) := (h \circ x(t_1))\overline{(g \circ y(t_2))}, \quad t_1, t_2 \in \Omega,$$

we see, by the implication (ii) $\Longrightarrow$ (7.16) from Lemma 7.6, that

$$H \in L^1(\Omega \times \Omega, \mathcal{B}, \mu).$$

Then, by the implication (iii) $\Longrightarrow$ (7.16) from Lemma 7.6, we see that

$$H \circ \phi \in L^1(\Omega, \mathcal{A}, \nu),$$

and (7.17) leads to

$$g^*(h) = \int_{\Omega \times \Omega} H \mathrm{d}\mu = \int_{\Omega} H \circ \phi \mathrm{d}\nu = \int_{\Omega} (h \circ x)\overline{(g \circ y)} \mathrm{d}\nu = g_\nu^*(h),$$

which proves (7.23).

Fix $\mathcal{A}$-measurable functions $u, v : \Omega \to B$. Then the function

$$\Omega \times \Omega \ni (t_1, t_2) \mapsto H(t_1, t_2) := |u(t_1) - v(t_2)|^2$$

is $\mathcal{B}$-measurable, and by the implication (iii) $\Longrightarrow$ (7.16) from Lemma 7.6, we conclude that

$$H \circ \phi \in L^1(\Omega, \mathcal{A}, \nu)$$

iff $H \in L^1(\Omega \times \Omega, \mathcal{B}, \mu)$. If $H \in L^1(\Omega \times \Omega, \mathcal{B}, \mu)$, then (7.17) leads to

$$\delta(u, v) = \int_{\Omega \times \Omega} H d\mu = \int_{\Omega} H \circ \phi d\nu = \int_{\Omega} |u - v|^2 d\nu = \eta(u, v),$$

and the equality $\delta(u, v) = \eta(u, v)$ holds. Otherwise $\delta(u, v) = +\infty$ and $\eta(u, v) = +\infty$. Thus the property (iii) is proved.

Consider now any $g \in \mathcal{L}_2(\mathfrak{R}^*)$ and any nonempty linear set $\mathcal{F} \subset \mathcal{L}_1(\mathfrak{R}^*)$. Applying the property (iii) we see that for each $f_0 \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g^*)$,

$$\eta(f \circ x, g \circ y) = \delta(f \circ x, g \circ y) \geq \delta(f_0 \circ x, g \circ y) = \eta(f_0 \circ x, g \circ y), \quad f \in \mathcal{F},$$

and so

$$f_0 \in \text{Reg}(\mathcal{F}, \mathfrak{R}_g).$$

Thus

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g^*) \subset \text{Reg}(\mathcal{F}, \mathfrak{R}_g).$$

In the same manner we can show that

$$\text{Reg}(\mathcal{F}, \mathfrak{R}_g) \subset \text{Reg}(\mathcal{F}, \mathfrak{R}_g^*).$$

Both the inclusions yield the equality (7.24), which completes the proof. $\square$

As an application of Theorem 7.7 we will show the following counterpart of Corollary 5.3 for synchronous regression structures.

Corollary 7.8 *Let $\mathfrak{R} = (A, B, \eta; x, y)$ be a synchronous regression structure determined by a measurable space $(\Omega, \mathcal{A}, \nu)$ and let $g \in L^2(B, \mathcal{A}_y, \nu_y)$. Given $p \in \mathbb{N} \cup \{\infty\}$ suppose that $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in L^2(A, \mathcal{A}_x, \nu_x)$ is a linearly independent sequence such that $\Theta_{\nu_x} \cap \mathcal{F} = \{\theta\}$, where $\mathcal{F} := \text{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$ and*

$$\Theta_{\nu_x} := \{h \in L^2(A, \mathcal{A}_x, \nu_x) : \|h\|_{\nu_x} = 0\}. \tag{7.26}$$

Moreover, the sequence

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k' \in \mathcal{L}_1(\mathfrak{R})$$

is its orthogonalization result by virtue of the Gramm-Schmidt recursive method, i.e.

$$h_1' := h_1 \quad and \quad h_n' := h_n - \sum_{k=1}^{n-1} \frac{\langle h_n | h_k' \rangle_{v_x}}{\|h_k'\|_{v_x}^2} h_k', \quad n \in \mathbb{Z}_{2,p}. \tag{7.27}$$

If $p \in \mathbb{N}$, then

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta_{v_x} \cap \mathcal{F}) + \sum_{k=1}^{p} \frac{\overline{g_v^*(h_k')}}{\|h_k'\|_{v_x}^2} h_k'. \tag{7.28}$$

If $p = \infty$, then

$$\mathrm{Reg}(\mathrm{cl}_{v_x}(\mathcal{F}), \mathfrak{R}_g) = \Theta_{v_x} + \sum_{k=1}^{\infty} \frac{\overline{g_v^*(h_k')}}{\|h_k'\|_{v_x}^2} h_k', \tag{7.29}$$

where cl_{v_x} stands for the closure operation in the space $\mathbf{L}^2(A, \mathcal{A}_x, v_x)$.

Proof Given a measurable space $(\Omega, \mathcal{A}, v)$ assume that it determines a synchronous regression structure $\mathfrak{R} = (A, B, \eta; x, y)$. Then Theorem 7.7 shows that the measurable space $(\Omega \times \Omega, \mathcal{B}, \mu)$, defined in Lemma 7.5, determines an asynchronous regression structure $\mathfrak{R}^* = (A, B, \delta; x, y)$. From the property (i) in Theorem 7.7 and (7.26) it follows that

$$\Theta_{v_x} = \{h \in \mathcal{L}_1(\mathfrak{R}^*) : \|h\| = 0\} = \Theta \tag{7.30}$$

as well as

$$\mathrm{cl}_{v_x}(\mathcal{F}) = \mathrm{cl}(\mathcal{F}). \tag{7.31}$$

Moreover, combining (7.27) with (7.21) and (7.22) we obtain

$$h_n' := h_n - \sum_{k=1}^{n-1} \frac{\langle h_n | h_k' \rangle}{\|h_k'\|^2} h_k', \quad n \in \mathbb{Z}_{2,p},$$

and so the equality (5.19) holds. If $p \in \mathbb{N}$, then by Corollary 5.3 the equality (5.20) holds with $\mathfrak{R}$ replaced by $\mathfrak{R}^*$. This together with (7.30), (7.21), (7.22), (7.23) and (7.24) yields the equality (7.28). If $p = \infty$, then by Corollary 5.3 the equality (5.21) holds with $\mathfrak{R}$ replaced by $\mathfrak{R}^*$. This together with (7.30), (7.31), (7.21), (7.22), (7.23) and (7.24) yields the equality (7.29), which completes the proof. □

Remark 7.9 Analysing the proof of Corollary 7.8 we see that this corollary is a direct conclusion from Corollary 5.3 and Theorem 7.7. Applying Theorem 7.7 we may easily derive, in much the same way as in the proof of Corollary 7.8, the other properties of the synchronous regression structures, being the respective counterparts properties of the asynchronous regression structures, presented in Sects. 3, 4 and 5. To this end we need an additional fact that the orthogonality condition in the space $\mathrm{H}(\mathfrak{R}^*)$ coincides with the one in the space $\mathbf{L}^2(A, \mathcal{A}_x, v_x)$, which is a direct consequence from (7.21).

8 Remarks on approximation

In this section we will point out that the synchronous regression structures can be applicable in the theory of approximation. To this end let $\mathfrak{R} = (A, B, \eta; x, y)$ be a synchronous regression structure determined by a measurable space $(\Omega, \mathcal{A}, \nu)$, where $A = B = \Omega$ and x, y are the identity mapping on B. Then $\nu_x = \nu_y = \nu$, and thereby

$$\mathrm{L}^2(B, \mathcal{A}_y, \nu_y) = \mathrm{L}^2(A, \mathcal{A}_x, \nu_x) = \mathrm{L}^2(B, \mathcal{A}, \nu).$$

Let us consider a sequence

$$\mathbb{N} \ni k \mapsto h_k \in \mathrm{L}^2(B, \mathcal{A}, \nu)$$

and fix $g \in \mathrm{L}^2(B, \mathcal{A}, \nu)$. From Corollary 3.5 and Theorem 7.7 it follows that

$$\mathrm{Reg}(\mathcal{F}_n, \mathfrak{R}_g) \neq \emptyset$$

for every $n \in \mathbb{N}$, where $\mathcal{F}_n := \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,n}\})$ as $n \in \mathbb{N}$. Therefore there exits a sequence

$$\mathbb{N} \ni n \mapsto f_n \in \mathrm{Reg}(\mathcal{F}_n, \mathfrak{R}_g),$$

which can approximate the function g, as stated in the following theorem.

Theorem 8.1 *For all* $f, g \in \mathrm{L}^2(B, \mathcal{A}, \nu)$ *and* $\mathbb{N} \ni n \mapsto f_n \in \mathrm{Reg}(\mathcal{F}_n, \mathfrak{R}_g)$, *if* $f \in \mathrm{Reg}(\mathrm{cl}_\nu(\mathcal{F}), \mathfrak{R}_g)$, *then*

$$\|f - f_n\|_\nu \to 0 \quad and \quad \|g - f_n\|_\nu \to \|g - f\|_\nu \quad as \ n \to \infty, \tag{8.1}$$

where $\mathcal{F} := \mathrm{lin}(\{h_k : k \in \mathbb{N}\})$. *In particular, if* $g \in \mathrm{cl}_\nu(\mathcal{F})$, *then*

$$g \in \mathrm{Reg}(\mathrm{cl}_\nu(\mathcal{F}), \mathfrak{R}_g)$$

and

$$\|g - f_n\|_\nu \to 0 \quad as \ n \to \infty. \tag{8.2}$$

Proof Assume that $f \in \mathrm{Reg}(\mathrm{cl}_\nu(\mathcal{F}), \mathfrak{R}_g)$ is arbitrarily fixed. By the definition of the regression function,

$$\|h - g\|_\nu^2 = \eta(h, g) \geq \eta(f, g) = \|f - g\|_\nu^2, \quad h \in \mathrm{cl}_\nu(\mathcal{F}),$$

and consequently,

$$\|f - g\|_\nu = \min(\{\|h - g\|_\nu : h \in \mathrm{cl}_\nu(\mathcal{F})\}). \tag{8.3}$$

Hence f is an orthogonal projection of g onto the linear and closed set $\mathrm{cl}_\nu(\mathcal{F})$ in the space $\mathbf{L}^2(B, \mathcal{A}, \nu)$, and so

$$\langle f - g | h \rangle_\nu = 0, \quad h \in \mathrm{cl}_\nu(\mathcal{F}). \tag{8.4}$$

Since

$$f \in \mathrm{Reg}(\mathrm{cl}_\nu(\mathcal{F}), \mathfrak{R}_g)$$

and $\mathcal{F} = \bigcup_{n=1}^\infty \mathcal{F}_n$, we see that there exists a sequence

$$\mathbb{N} \ni n \mapsto f_n' \in \mathcal{F}_n$$

satisfying the following property

$$\| f - f_n' \|_\nu \to 0 \quad \text{as } n \to \infty. \tag{8.5}$$

By assumption, $f_n \in \mathrm{Reg}(\mathcal{F}_n, \mathfrak{R}_g)$ for $n \in \mathbb{N}$, and therefore

$$\| f_n' - g \|_\nu^2 = \eta(f_n', g) \ge \eta(f_n, g) = \| f_n - g \|_\nu^2, \quad n \in \mathbb{N}. \tag{8.6}$$

From (8.4) it follows that

$$\langle g - f | f - f_n \rangle_\nu = 0 \quad \text{and} \quad \langle g - f | f - f_n' \rangle_\nu = 0, \quad n \in \mathbb{N}. \tag{8.7}$$

Combining this with (8.6) we see that for every $n \in \mathbb{N}$,

$$\begin{aligned}
\| g - f \|_\nu^2 + \| f - f_n \|_\nu^2 &= \| g - f \|_\nu^2 + 2 \operatorname{Re} \langle g - f | f - f_n \rangle_\nu + \| f - f_n \|_\nu^2 \\
&= \| g - f_n \|_\nu^2 \le \| g - f_n' \|_\nu^2 \\
&= \| g - f \|_\nu^2 + 2 \operatorname{Re} \langle g - f | f - f_n' \rangle_\nu + \| f - f_n' \|_\nu^2 \\
&= \| g - f \|_\nu^2 + \| f - f_n' \|_\nu^2.
\end{aligned}$$

From this and (8.5) we conclude that

$$\| f - f_n \|_\nu \le \| f - f_n' \|_\nu \to 0 \quad \text{as } n \to \infty.$$

Hence and by (8.7) we have

$$\| g - f_n \|_\nu^2 = \| g - f \|_\nu^2 + \| f - f_n \|_\nu^2 \to \| g - f \|_\nu^2 \quad \text{as } n \to \infty.$$

Both the limits yield (8.1). If now $g \in \mathrm{cl}_\nu(\mathcal{F})$, then by

$$\eta(h, g) \ge 0 = \eta(g, g), \quad h \in \mathrm{cl}_\nu(\mathcal{F}),$$

we deduce that

$$g \in \text{Reg}(\text{cl}_\nu(\mathcal{F}), \mathfrak{R}_g),$$

and (8.1) leads to (8.2), which completes the proof. $\qquad\square$

The following example illustrates how to apply synchronous regression structures in the theory of approximation.

Example 8.2 Let $\mathbb{N} \ni k \mapsto h_k \in \mathbf{L}^2(B, \mathcal{A}, \nu)$ be an orthogonal sequence in the space $\mathbf{L}^2(B, \mathcal{A}, \nu)$ such that $\|h_k\|_\nu > 0$ for $k \in \mathbb{N}$. Then for every $g \in \mathbf{L}^2(B, \mathcal{A}, \nu)$,

$$\left\| f - \sum_{k=1}^{n} \frac{\overline{\langle g|h_k\rangle_\nu}}{\|h_k\|_\nu^2} h_k \right\| \to 0 \quad \text{as } n \to \infty, \tag{8.8}$$

where f is an orthogonal projection of g onto $\text{cl}_\nu(\mathcal{F})$ in the space $\mathbf{L}^2(B, \mathcal{A}, \nu)$ and $\mathcal{F} := \text{lin}(\{h_k : k \in \mathbb{N}\})$. Moreover, the following Bessel equality holds:

$$\sum_{k=1}^{\infty} \frac{|\langle g|h_k\rangle_\nu|^2}{\|h_k\|_\nu^2} = \|g\|_\nu^2 - \|g - f\|_\nu^2. \tag{8.9}$$

Namely, let f be an orthogonal projection of g onto $\text{cl}_\nu(\mathcal{F})$ in the space $\mathbf{L}^2(B, \mathcal{A}, \nu)$. Then

$$\eta(h, g) = \|h - g\|_\nu^2 \geq \|f - g\|_\nu^2 = \eta(f, g), \quad h \in \text{cl}_\nu(\mathcal{F}),$$

and consequently,

$$f \in \text{Reg}(\text{cl}_\nu(\mathcal{F}), \mathfrak{R}_g).$$

From Corollary 7.8 it follows that

$$f_n := \sum_{k=1}^{n} \frac{\overline{\langle g|h_k\rangle_\nu}}{\|h_k\|_\nu^2} h_k \in \text{Reg}(\mathcal{F}_n, \mathfrak{R}_g), \quad n \in \mathbb{N}. \tag{8.10}$$

Then Theorem 8.1 yields (8.8). From (8.8) it follows that

$$\begin{aligned} |\|f\|_\nu^2 - \|f_n\|_\nu^2| &= (\|f\|_\nu + \|f_n\|_\nu)|\|f\|_\nu - \|f_n\|_\nu| \\ &\leq (\|f\|_\nu + \|f_n\|_\nu)\|f - f_n\|_\nu \to 0 \quad \text{as } n \to \infty. \end{aligned} \tag{8.11}$$

Since f is an orthogonal projection of g onto $\text{cl}_\nu(\mathcal{F})$ in the space $\mathbf{L}^2(B, \mathcal{A}, \nu)$, the equality (8.4) holds, and consequently

$$\|g\|_\nu^2 = \|g - f\|_\nu^2 + 2\,\text{Re}\,\langle g - f|f\rangle_\nu + \|f\|_\nu^2 = \|g - f\|_\nu^2 + \|f\|_\nu^2. \tag{8.12}$$

By the orthogonality of the sequence $\mathbb{N} \ni k \mapsto h_k$, we deduce from (8.10) that

$$\|f_n\|_\nu^2 = \sum_{k=1}^n \sum_{l=1}^n \frac{\overline{\langle g|h_k\rangle_\nu}}{\|h_k\|_\nu^2} \frac{\langle g|h_l\rangle_\nu}{\|h_l\|_\nu^2} \langle h_k|h_l\rangle_\nu = \sum_{k=1}^n \frac{|\langle g|h_k\rangle_\nu|^2}{\|h_k\|_\nu^2}, \quad n \in \mathbb{N}.$$

Combining this with (8.11) and (8.12) we get (8.9).

9 The regression in probabilistic spaces

Regression functions are very often used in the context of the probability theory. Note that this case corresponds to a synchronous regression structure

$$\mathfrak{R} = (A, B, \eta; x, y)$$

determined by a probability space $(\Omega, \mathcal{A}, P)$, where $\nu = P$, $A = B$ and $x, y : \Omega \to B$ are random variables, i.e. they are $\mathcal{A}$-measurable. Therefore all the facts discussed in Sect. 7 are still valid, in particular, in the probabilistic case. In this section we rewrite formulas (7.28) and (7.29) in terms of — specific for the probability theory — expected values of random variables and distribution functions.

First we note that usually instead of the probability measure P the distribution function $P_{x,y}$ generated by P and random variables x and y is considered. We recall that $P_{x,y}$ is the unique probability measure on the σ-field $\mathfrak{B}(B) \otimes \mathfrak{B}(B)$ satisfying the following condition

$$P_{x,y}(U \times V) = P(x^{-1}(U) \cap y^{-1}(V)), \quad U, V \in \mathfrak{B}(B). \tag{9.1}$$

Here and in the sequel, $\mathfrak{B}(B)$ denotes the σ-field of Borel sets in B. Thus our main task is to describe regression functions by means of the probability space

$$(B \times B, \mathfrak{B}(B) \otimes \mathfrak{B}(B), P_{x,y}).$$

Lemma 9.1 *For any Borel functions $g, h : B \to B$, $(g \circ x)(h \circ y) \in \mathrm{L}^1(\Omega, \mathcal{A}, P)$ iff the function*

$$B \times B \ni (t_1, t_2) \mapsto g(t_1)h(t_2)$$

is integrable with respect to the distribution function $P_{x,y}$. Moreover, if

$$(g \circ x)(h \circ y) \in \mathrm{L}^1(\Omega, \mathcal{A}, P),$$

then

$$\int_\Omega (g \circ x)(h \circ y)\mathrm{d}P = \int_{B \times B} g(t_1)h(t_2)\mathrm{d}P_{x,y}(t_1, t_2). \tag{9.2}$$

Proof We first prove this lemma in the case where g and h are simple Borel functions, i.e.

$$g = \sum_{k=1}^{n} \lambda_k \, I_{U_k} \quad \text{and} \quad h = \sum_{k=1}^{m} \mu_k \, I_{V_k}$$

for certain $n, m \in \mathbb{N}$ and sequences

$$\mathbb{Z}_{1,n} \ni k \mapsto \lambda_k \in B, \quad \mathbb{Z}_{1,m} \ni k \mapsto \mu_k \in B,$$

$$\mathbb{Z}_{1,n} \ni k \mapsto U_k \in \mathfrak{B}(B), \quad \mathbb{Z}_{1,m} \ni k \mapsto V_k \in \mathfrak{B}(B)$$

such that

$$\bigcup_{k=1}^{n} U_k = B, \quad \bigcup_{k=1}^{m} V_k = B \quad \text{and} \quad U_k \cap U_l = \emptyset = V_k \cap V_l$$

provided $k \neq l$. Then

$$g \circ x = \sum_{k=1}^{n} \lambda_k \, I_{x^{-1}(U_k)} \quad \text{and} \quad h \circ y = \sum_{k=1}^{m} \mu_k \, I_{y^{-1}(V_k)},$$

and consequently

$$\int_{\Omega} (g \circ x)(h \circ y) \mathrm{d}P = \int_{\Omega} \left(\sum_{k=1}^{n} \lambda_k \, I_{x^{-1}(U_k)} \sum_{l=1}^{m} \mu_l \, I_{y^{-1}(V_l)} \right) \mathrm{d}P \qquad (9.3)$$

$$= \int_{\Omega} \sum_{k=1}^{n} \sum_{l=1}^{m} \lambda_k \mu_l \, I_{x^{-1}(U_k)} \, I_{y^{-1}(V_l)} \, \mathrm{d}P$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{m} \lambda_k \mu_l \int_{\Omega} I_{x^{-1}(U_k) \cap y^{-1}(V_l)} \, \mathrm{d}P.$$

From (9.1) it follows that

$$\int_{B \times B} I_{U_k \times V_l}(t_1, t_2) \mathrm{d}P_{x,y}(t_1, t_2) = P_{x,y}(U_k \times V_l) = P(x^{-1}(U_k) \cap y^{-1}(V_l))$$

$$= \int_{\Omega} I_{x^{-1}(U_k) \cap y^{-1}(V_l)} \, \mathrm{d}P.$$

Combining this with (9.3) we deduce that

$$
\begin{aligned}
\int_{B \times B} g(t_1)h(t_2)\mathrm{d}P_{x,y}(t_1, t_2) &= \int_{B \times B} \sum_{k=1}^{n} \sum_{l=1}^{m} \lambda_k \mu_l \, \mathrm{I}_{U_k}(t_1) \, \mathrm{I}_{V_l}(t_2)\mathrm{d}P_{x,y}(t_1, t_2) \\
&= \sum_{k=1}^{n} \sum_{l=1}^{m} \lambda_k \mu_l \int_{B \times B} \mathrm{I}_{U_k \times V_l}(t_1, t_2)\mathrm{d}P_{x,y}(t_1, t_2) \\
&= \int_{\Omega} (g \circ x)(h \circ y)\mathrm{d}P,
\end{aligned}
$$

and so the lemma holds for any simple Borel functions g and h.

Assume now that $g, h : B \to B$ are any Borel functions. Then there exist sequences

$$
\mathbb{N} \ni n \mapsto g_n \quad \text{and} \quad \mathbb{N} \ni n \mapsto h_n
$$

of simple Borel functions such that

$$
|g_n(t)| \leq |g(t)| \quad \text{and} \quad |h_n(t)| \leq |h(t)|, \quad t \in B, \, n \in \mathbb{N}, \tag{9.4}
$$

as well as for every $t \in B$,

$$
g_n(t) \to g(t) \quad \text{and} \quad h_n(t) \to h(t) \quad \text{as } n \to \infty. \tag{9.5}
$$

By the already proved equality (9.2) for simple Borel functions we see that the equalities

$$
\int_{\Omega} (g_n \circ x)(h_n \circ y)\mathrm{d}P = \int_{B \times B} g_n(t_1)h_n(t_2)\mathrm{d}P_{x,y}(t_1, t_2) \tag{9.6}
$$

as well as

$$
\int_{\Omega} |(g_n \circ x)(h_n \circ y)|\mathrm{d}P = \int_{B \times B} |g_n(t_1)h_n(t_2)|\mathrm{d}P_{x,y}(t_1, t_2) \tag{9.7}
$$

hold for every $n \in \mathbb{N}$. Assume that

$$
(g \circ x)(h \circ y) \in \mathrm{L}^1(\Omega, \mathcal{A}, P).
$$

Since each function

$$
B \times B \ni (t_1, t_2) \mapsto |g_n(t_1)h_n(t_2)|
$$

is $\mathfrak{B}(B) \otimes \mathfrak{B}(B)$-measurable we conclude from (9.5) that the pointwise limit function

$$
B \times B \ni (t_1, t_2) \mapsto |g(t_1)h(t_2)|
$$

is also $\mathfrak{B}(B) \otimes \mathfrak{B}(B)$-measurable. Applying now Fatou's integral lemma, (9.7) and (9.4) we have

$$
\begin{aligned}
\int_{B \times B} |g(t_1)h(t_2)| \mathrm{d}P_{x,y}(t_1, t_2) &= \int_{B \times B} \liminf_{n \to \infty} |g_n(t_1)h_n(t_2)| \mathrm{d}P_{x,y} \\
&\leq \liminf_{n \to \infty} \int_{B \times B} |g_n(t_1)h_n(t_2)| \mathrm{d}P_{x,y}(t_1, t_2) \\
&= \liminf_{n \to \infty} \int_{\Omega} |(g_n \circ x)(h_n \circ y)| \mathrm{d}P \\
&\leq \int_{\Omega} |(g \circ x)(h \circ y)| \mathrm{d}P < +\infty.
\end{aligned}
$$

This means that the function

$$
B \times B \ni (t_1, t_2) \mapsto g(t_1)h(t_2)
$$

is integrable with respect to the distribution function $P_{x,y}$, provided

$$
(g \circ x)(h \circ y) \in \mathrm{L}^1(\Omega, \mathcal{A}, P).
$$

In much the same way we can justify the inverse implication.

It remains to show the equality (9.2), provided

$$
(g \circ x)(h \circ y) \in \mathrm{L}^1(\Omega, \mathcal{A}, P).
$$

The last condition implies

$$
\int_{B \times B} |g(t_1)h(t_2)| \mathrm{d}P_{x,y}(t_1, t_2) < +\infty.
$$

Applying now Lebesgue's dominated limiting integral theorem we conclude from (9.4), (9.5) and (9.6) that

$$
\begin{aligned}
\int_{\Omega} (g \circ x)(h \circ y) \mathrm{d}P &= \lim_{n \to \infty} \int_{\Omega} (g_n \circ x)(h_n \circ y) \mathrm{d}P \\
&= \lim_{n \to \infty} \int_{B \times B} g_n(t_1)h_n(t_2) \mathrm{d}P_{x,y}(t_1, t_2) \\
&= \int_{B \times B} g(t_1)h(t_2) \mathrm{d}P_{x,y}(t_1, t_2),
\end{aligned}
$$

which is the desired conclusion.   □

Since our structure $\mathfrak{R}$ is a synchronous regression structure we can use directly Corollary 7.8 in order to compute the regression functions. However, from the practical

point of view it seems to be more convenient to compute the ones in terms of the distribution function $P_{x,y}$. To this end we introduce the following two functions:

$$(f, h) \mapsto \mathrm{E}_{x,x}(f|h) := \frac{\mathrm{E}((f \circ x)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} \tag{9.8}$$

as well as

$$(g, h) \mapsto \mathrm{E}_{y,x}(g|h) := \frac{\mathrm{E}((g \circ y)\overline{(h \circ x)})}{\mathrm{E}(|h \circ x|^2)} \tag{9.9}$$

for $f \in \mathrm{L}^2(B, \mathcal{A}_x, P_x)$, $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ and $h \in \mathrm{L}^2(B, \mathcal{A}_x, P_x) \backslash \Theta_{P_x}$, where E is the expected value operator for the probability space $(\Omega, \mathcal{A}, P)$, i.e.

$$\mathrm{E}(f) := \int_\Omega f \, \mathrm{d}P, \quad f \in \mathrm{L}^1(\Omega, \mathcal{A}, P). \tag{9.10}$$

By Lemma 9.1 we obtain

$$\mathrm{E}_{x,x}(f|h) = \frac{\int_{B \times B} f(t_1)\overline{h(t_1)} \mathrm{d}P_{x,y}(t_1, t_2)}{\int_{B \times B} |h(t_1)|^2 \mathrm{d}P_{x,y}(t_1, t_2)} \tag{9.11}$$

and

$$\mathrm{E}_{y,x}(g|h) = \frac{\int_{B \times B} g(t_2)\overline{h(t_1)} \mathrm{d}P_{x,y}(t_1, t_2)}{\int_{B \times B} |h(t_1)|^2 \mathrm{d}P_{x,y}(t_1, t_2)} \tag{9.12}$$

for all $f \in \mathrm{L}^2(B, \mathfrak{B}(B), P_x)$, $g \in \mathrm{L}^2(B, \mathfrak{B}(B), P_y)$ and $h \in \mathrm{L}^2(B, \mathfrak{B}(B), P_x) \backslash \Theta_{P_x}$.

Corollary 9.2 *Let* $\mathfrak{R} = (B, B, \eta; x, y)$ *be a synchronous regression structure determined by a probabilistic space* $(\Omega, \mathcal{A}, P)$ *and let* $g \in \mathrm{L}^2(B, \mathfrak{B}(B), P_y)$. *Given* $p \in \mathbb{N} \cup \{\infty\}$ *suppose that*

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(B, \mathfrak{B}(B), P_x)$$

is a linearly independent sequence such that

$$\Theta_{P_x} \cap \mathcal{F} = \{\theta\},$$

where $\mathcal{F} := \mathrm{lin}(\{h_k : k \in \mathbb{Z}_{1,p}\})$, *and the sequence*

$$\mathbb{Z}_{1,p} \ni k \mapsto h_k' \in \mathcal{L}_1(\mathfrak{R})$$

is defined as follows

$$h_1' := h_1 \quad and \quad h_n' := h_n - \sum_{k=1}^{n-1} \mathrm{E}_{x,x}(h_n|h_k')h_k', \quad n \in \mathbb{Z}_{2,p}. \tag{9.13}$$

If $p \in \mathbb{N}$, then

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{R}_g) = (\Theta_{P_x} \cap \mathcal{F}) + \sum_{k=1}^{p} \operatorname{E}_{y,x}(g|h_k')h_k'. \qquad (9.14)$$

If $p = \infty$, then

$$\operatorname{Reg}(\operatorname{cl}_{P_x}(\mathcal{F}), \mathfrak{R}_g) = \Theta_{P_x} + \sum_{k=1}^{\infty} \operatorname{E}_{y,x}(g|h_k')h_k'. \qquad (9.15)$$

Proof From Lemma 7.3 and (9.10) it follows that for all $g, h \in \operatorname{L}^2(B, \mathfrak{B}(B), P_x)$,

$$\langle g|h \rangle_{P_x} = \int_B g\overline{h} \, \mathrm{d}P_x = \int_{\Omega} (g\overline{h}) \circ x \, \mathrm{d}P = \operatorname{E}((g \circ x)\overline{(h \circ x)}), \qquad (9.16)$$

and, in particular,

$$\|h\|_{P_x}^2 = \langle h|h \rangle_{P_x} = \operatorname{E}(|h \circ x|^2). \qquad (9.17)$$

Combining (9.8) with (9.16) and (9.17) we obtain

$$\operatorname{E}_{x,x}(g|h) = \frac{\langle g|h \rangle_{P_x}}{\|h\|_{P_x}^2}, \qquad (9.18)$$

as $g \in \operatorname{L}^2(B, \mathfrak{B}(B), P_x)$, and $h \in \operatorname{L}^2(B, \mathfrak{B}(B), P_x) \backslash \Theta_{P_x}$.

From (7.23) and (9.10) it may be concluded that for all $g \in \operatorname{L}^2(B, \mathfrak{B}(B), P_y)$ and $h \in \operatorname{L}^2(B, \mathfrak{B}(B), P_x)$,

$$g_P^*(h) = \int_{\Omega} (h \circ x)\overline{(g \circ y)} \, \mathrm{d}P = \operatorname{E}((h \circ x)\overline{(g \circ y)}). \qquad (9.19)$$

Combining (9.9) with (9.19) and (9.17) we have

$$\operatorname{E}_{y,x}(g|h) = \frac{g_P^*(h)}{\|h\|_{P_x}^2}, \qquad (9.20)$$

as $g \in \operatorname{L}^2(B, \mathfrak{B}(B), P_y)$, and $h \in \operatorname{L}^2(B, \mathfrak{B}(B), P_x) \backslash \Theta_{P_x}$.

Using (9.18) and (9.19) we deduce from Corollary 7.8 that equalities (9.13), (9.14) and (9.15) hold, which is the desired conclusion. $\qquad \square$

Remark 9.3 Note that if $\mathcal{F} \subset \operatorname{L}(B, \mathfrak{B}(B))$, then by Lemma 9.1,

$$\Theta_{P_x} \cap \mathcal{F} = \{h \in \operatorname{L}^2(B, \mathcal{A}_x, P_x) \cap \mathcal{F} : \|h\|_{P_x} = 0\}$$

$$= \{h \in \mathcal{F} : \int_{B \times B} |h(t_1)|^2 \mathrm{d}P_{x,y}(t_1, t_2) = 0\}.$$

References

1. Anderson, T.W.: Last squares and best unbiased estimates. Ann. Math. Stat. **33**, 266–272 (1982)
2. Ash, R.B.: Real Analysis and Probability, Probability and Mathematical Statistics. Academic Press Inc, New York and London (1972)
3. Baltagi, B.H.: Econometrics. Springer, Berlin (2011)
4. Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley, New York, London, Sydney (1966)
5. Durbin, I.: A note on regression when there is extraneous information about one of the coefficients. IASA **48**, 799–808 (1953)
6. Fisher, N.I., Lee, A.I.: Regression models for an angular response. Biometrics **48**(3), 665–677 (1992)
7. Goldberg, A.S.: Introductory Econometrics. Harvard Uniersity Press, London (1998)
8. Larsen, R.: Functional Analysis (an introduction). Pure and Applied Mathematics. Marcel Dekker Inc, New York (1973)
9. Maddala, G.S., Lahiri, K.: Introduction to Econometrics. Wiley, New York (2006)
10. Mlak, W.: Hilbert Spaces and Operator Theory, Mathematics and Its Applications, Kluwer Academic Publishers and PWN-Polish Scientific Publishers, Dordrecht, Boston, London and Warsaw (1991)
11. Naes, T., Helland, I.S.: Relevant components in regression. Scand J. Stat. **20**(3), 239–250 (1993)
12. Partyka D, Zając J.: Generalized problem of regression. Bull. Soc. Sci. Lett. Łódź 60 (2010), **1**, 77–94, Série: Recherches sur les déformations
13. Philips, P.C.B., Sun, Y.: Regression with an evaporating logarithmic trend. Econom. Theory **19**(4), 692–701 (2003)
14. Quandt, R.E.: The estimation of the parameter of a linear regression system obeying two separate regimes. IASA **53**, 873–890 (1968)
15. Teser, L.G.: Iterative estimation of a set linear regression equations. IASA **59**, 845–862 (1964)
16. Whiston, W.B.: Sequential Selection of Variables in Multiple Regression. University of Cincinnati, Cincinnati (1964)
17. Wiliams, E.I.: Regression Analysis. Wiley, New York (1959)