https://doi.org/10.26485/0459-6854/2018/68.3/2 PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2018 Vol. LXVIII

Recherches sur les déformations no. 3

pp. 21–38

Dedicated to the memory of
Professor Yurii B. Zelinskii

Agnieszka Szumera

PROBABILISTIC REGRESSION STRUCTURES

Summary

A new approach generalizing the classical regression idea has been widely presented in [5] and [6] in the environment of an arbitrary Hilbert space. The problem of transforming this idea to a probability space is considered in the present paper.

Keywords and phrases: nonlinear regression, polynomial regression, probability space, regression functions, regression structure

Introduction

In [5] and [6] the general problem of regression was discussed and solved. The authors introduced the concept of the *regression structures* $\mathfrak{R} := (A, B, \delta; x, y)$, where:

I.1. A and B are nonempty sets;

I.2. $x\colon \Omega_1 \to A$ and $y\colon \Omega_2 \to B$ are functions defined on given nonempty sets Ω_1 and Ω_2; they can be interpreted as experimental data of the regression model. Therefore we call them *empirical data functions*;

I.3. $\delta\colon (\Omega_1 \to B) \times (\Omega_2 \to B) \to \overline{\mathbb{R}}$ is a function which can be interpreted as a deviation criterion of the theoretic functions from the empirical data.

For a given regression structure $\mathfrak{R}$ we consider the family of functions $\mathcal{F}$ included in the family $A \to B$ of all functions acting from A to B, i.e. $\mathcal{F} \subset (A \to B)$. The family $\mathcal{F}$ is said to be a *theoretic functional model* of the observed phenomena, i.e., $\mathcal{F}$ consists of all functions describing theoretically the considered phenomena. In the sequel we will restrict our considerations to the case where $B = \mathbb{R}$ or $B = \mathbb{C}$ and $\mathcal{F}$ is a linear set with respect to the standard operations of adding and multiplying functions, i.e.,

$$f + g, \quad \lambda \cdot f \in \mathcal{F} \quad \text{for} \quad f, g \in \mathcal{F} \quad \text{and} \quad \lambda \in B \; .$$

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

$$\mathcal{F} \ni f \to F(f) := \delta(f \circ x, y) \in \overline{\mathbb{R}} \; , \tag{0.1}$$

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

$$F(f) \geq F(f_0) \quad \text{for} \quad f \in \mathcal{F} \; . \tag{0.2}$$

The set of all $f_0 \in \mathcal{F}$ satisfying the inequality (0.2) will be denoted by $\mathrm{Reg}(\mathcal{F}, \mathfrak{R})$; c.f. [6]. 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* $\mathfrak{R}$.

Given a nonempty set Ω and σ-field $\mathcal{B}$ of its subsets, we denote by $\mathrm{L}(\Omega, \mathcal{B})$ the family of all complex valued functions on Ω, measurable with respect to $\mathcal{B}$. Further on we denote by $\mathbf{L}(\Omega, \mathcal{B})$ the linear space supported by the set $\mathrm{L}(\Omega, \mathcal{B})$ and equipped with the standard operations of adding and multiplying of functions, i.e., $\mathbf{L}(\Omega, \mathcal{B}) := (\mathrm{L}(\Omega, \mathcal{B}), +, \cdot)$.

For a given measure $\mu \colon \mathcal{B} \to [0, +\infty)$ and $p \geq 1$, let $\mathrm{L}^p(\Omega, \mathcal{B}, \mu)$ stand for the class of all functions $f \in \mathrm{L}(\Omega, \mathcal{B})$ such that

(0.3)

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

$$\|f\|_p = 0 \iff \mu(\{\omega \in \Omega \colon f(\omega) \neq 0\}) = 0 \; . \tag{0.4}$$

Hence the structure

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

is a pseudo-Banach space, i.e., a complete pseudo-normed space.

1. Probabilistic regression structure

Following the general concept of regression structures, cf. [6, Definition 2.1 and Definition 7.1], we introduce a special type of regression structures on the basis of probability theory.

Definition 1.1. By a *probabilistic regression structure* we mean any regression structure $\mathfrak{P} := (A, B, \delta; x, y)$ determined by a probability space $\mathcal{P} = (\Omega, \mathcal{A}, P)$, which satisfies the following conditions:

II.1. A is nonempty set and $B = \mathbb{R}$ or $B = \mathbb{C}$;
II.2. $x \colon \Omega \to A$ and $y \colon \Omega \to B$;
II.3. the function $\delta \colon (\Omega \to B) \times (\Omega \to B) \to \overline{\mathbb{R}}$ satisfies the equality

$$\delta(u, v) = \int_\Omega |u(\omega) - v(\omega)|^2 \,\mathrm{d}P(\omega) \;, \tag{1.1}$$

provided both the functions u and v are $\mathcal{A}$-measurable, and $\delta(u, v) = +\infty$ otherwise.

Under the above conditions the regression problem for a probabilistic regression structure $\mathfrak{P}$ is the extremal problem of determining all functions $f_0 \in \mathcal{F}$ minimizing the functional F given – in the wake of (0.1) i (1.1) – by the following formula

$$F(f) = \delta(f \circ x, y) = \int_\Omega |f \circ x(\omega) - y(\omega)|^2 \,\mathrm{d}P(\omega), \qquad f \in \mathcal{F} \;. \tag{1.2}$$

For a given probabilistic regression structure $\mathfrak{P}$ we define

$$\mathcal{A}_x := \{V \in 2^A \colon x^{-1}(V) \in \mathcal{A}\} \tag{1.3}$$

and

$$\mathcal{A}_x \ni V \mapsto P_x(V) := P(x^{-1}(V)) \;. \tag{1.4}$$

It is clear that $\mathcal{A}_x$ is a σ-field on A and P_x is a probability measure on $\mathcal{A}_x$.

For the further discussion we quote the following fact, cf. [1], [2].

Theorem 1.2. *For every measurable space $(\Omega, \mathcal{A}, P)$ and every function $x \colon \Omega \to A$, the structure $(A, \mathcal{A}_x, P_x)$ is also a measurable space. Moreover, for every $\mathcal{A}_x$-measurable function $u \colon A \to B$,*

$$u \in \mathrm{L}^1(A, \mathcal{A}_x, P_x) \Longleftrightarrow u \circ x \in \mathrm{L}^1(\Omega, \mathcal{A}, P)$$

as well as

$$\int_\Omega u \circ x(\omega) \mathrm{d}P(\omega) = \int_A u(t) \mathrm{d}P_x(t), \qquad u \in \mathrm{L}^1(A, \mathcal{A}_x, P_x) \;. \tag{1.5}$$

Remark 1.3. It is well known that the function

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \times \mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni (u, v) \mapsto \langle u|v\rangle := \int_A u(t) \cdot \overline{v(t)} \,\mathrm{d}P_x(t) \tag{1.6}$$

is well defined and the following properties

$$\begin{array}{rcl} \langle \lambda_1 u + \lambda_2 v|w\rangle &=& \lambda_1 \langle u|w\rangle + \lambda_2 \langle v|w\rangle \;; \\ \langle u|v\rangle &=& \overline{\langle v|u\rangle} \;; \\ \langle u|u\rangle &\geq& 0 \;, \end{array} \tag{1.7}$$

hold for all $\lambda_1, \lambda_2 \in B$ and $u, v, w \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. Moreover the functional

$$\mathrm{L}^2(A, \mathcal{A}_x, P_x) \ni u \mapsto \|u\| := \sqrt{\langle u|u\rangle} = \left(\int_A |u(t)|^2 \mathrm{d}P_x(t)\right)^{1/2} \tag{1.8}$$

has the following properties

$$\|\lambda u\| = |\lambda| \cdot \|u\| \qquad \text{and} \qquad \|u + v\| \leq \|u\| + \|v\|$$

as well as

$$\|u\| = 0 \iff P_x(\{t \in A : u(t) \neq 0\}) = 0$$

for all $\lambda \in B$ and $u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, cf. [9]. Therefore, $\|\cdot\|$ is a pseudo-norm on the linear space $(\mathrm{L}^2(A, \mathcal{A}_x, P_x), +, \cdot)$.
From the properties (1.7) the following Schwarz inequality

$$|\langle u|v\rangle| \leq \|u\| \cdot \|v\| , \qquad u, v \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \tag{1.9}$$

can be derived in the standard way, cf. [11].
Since the space $\mathbf{L}^2(A, \mathcal{A}_x, P_x)$ is complete, cf. [1], we see that the structure $\mathbf{H}(\mathfrak{P}) := (\mathrm{L}^2(A, \mathcal{A}_x, P_x), +, \cdot, \langle\cdot|\cdot\rangle)$ is a pseudo-Hilbert space (complex if $B = \mathbb{C}$ or real if $B = \mathbb{R}$), i.e., the structure $(\mathrm{L}^2(A, \mathcal{A}_x, P_x), +, \cdot, \|\cdot\|)$ is a pseudo-Banach space.

Similarly to (1.3) and (1.4) we see that

$$\mathcal{A}_y := \{V \in 2^B : y^{-1}(V) \in \mathcal{A}\} \tag{1.10}$$

is a σ-field on B and

$$\mathcal{A}_y \ni V \mapsto P_y(V) := P(y^{-1}(V)) \tag{1.11}$$

is a probabilistic measure on $\mathcal{A}_y$.

Remark 1.4. Given $u \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ and $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ we see that $|u|^2 \in \mathrm{L}^1(A, \mathcal{A}_x, P_x)$ and $|g|^2 \in \mathrm{L}^1(B, \mathcal{A}_y, P_y)$. Since $|u|^2 \circ x = |u \circ x|^2$ and $|g|^2 \circ y = |g \circ y|^2$, we conclude from (1.3), (1.4), (1.10), (1.11) and Theorem 1.2 that $u \circ x,\ g \circ y \in \mathrm{L}^2(\Omega, \mathcal{A}, P)$ and

$$\int_\Omega |u \circ x(\omega)|^2 \mathrm{d}P(\omega) = \int_A |u(t)|^2 \mathrm{d}P_x(t), \qquad u \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) .$$

Hence

$$M_g := \left(\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} < +\infty$$

and applying the Schwarz inequality for Lebesgue integral we have

$$\begin{aligned}\int_\Omega &|u\circ x(\omega)\cdot\overline{g\circ y(\omega)}|\mathrm{d}P(\omega) \qquad (1.12)\\ &\le\left(\int_\Omega |u\circ x(\omega)|^2\mathrm{d}P(\omega)\right)^{1/2}\cdot\left(\int_\Omega |g\circ y(\omega)|^2\mathrm{d}P(\omega)\right)^{1/2}\\ &=M_g\left(\int_A |u(t)|^2\mathrm{d}P_x(t)\right)^{1/2}=M_g\cdot\|u\|\ .\end{aligned}$$

Therefore for every $g\in \mathrm{L}^2(B,\mathcal{A}_y,P_y)$ the functional $g^*\colon \mathrm{L}^2(A,\mathcal{A}_x,P_x)\to B$ is well-defined by the formula

$$\mathrm{L}^2(A,\mathcal{A}_x,P_x)\ni u\to g^*(u):=\int_\Omega u\circ x(\omega)\overline{g\circ y(\omega)}\mathrm{d}P(\omega)\ . \tag{1.13}$$

and by (1.12) we obtain

$$|g^*(u)|\le M_g\cdot\|u\|\ ,\qquad u\in \mathrm{L}^2(A,\mathcal{A}_x,P_x)\,. \tag{1.14}$$

Thus g^* is a linear and bounded functional on $(\mathrm{L}^2(A,\mathcal{A}_x,P_x),+,\cdot,\|\cdot\|)$ for every $g\in \mathrm{L}^2(B,\mathcal{A}_y,P_y)$.

2. The regression problem for the probabilistic regression structures

Let $\mathfrak{P}:=(A,B,\delta;x,y)$ be a probabilistic regression structure determined by a probability space $\mathcal{P}=(\Omega,\mathcal{A},P)$. Then for a given $g\colon B\to B$,

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

is a probabilistic regression structure determined by $\mathcal{P}$. We interpret the function g as a *scaling function* of the data function y.

From now on we shall study the regression problem for $\mathcal{F}$ with respect to $\mathfrak{P}_g$, where $\mathcal{F}$ is a linear functional model with standard operations of adding and multiplying functions.

The following result is a counterpart of [6, Lemma 3.1].

Theorem 2.1. *If $\mathcal{F}\ne\emptyset$ is a linear set in $\mathbf{H}(\mathfrak{P})$ and $g\in \mathrm{L}^2(B,\mathcal{A}_y,P_y)$, then for every $f\in\mathcal{F}$ the following condition holds:*

$$f\in \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)\Longleftrightarrow \langle h|f\rangle=g^*(h)\,,\qquad h\in\mathcal{F}. \tag{2.1}$$

Proof. Given $g\in \mathrm{L}^2(B,\mathcal{A}_y,P_y)$ we define the functional

$$(A\to B)\ni f\mapsto F_g(f):=\delta(f\circ x,g\circ y).$$

From the property II.3 it follows that

$$F_g(f)=\int_\Omega |f\circ x(\omega)-g\circ y(\omega)|^2\,\mathrm{d}P(\omega),\qquad f\in\mathcal{F}\ . \tag{2.2}$$

Fix $f, h \in \mathcal{F}$ and $\lambda \in B$. Then by (2.2), we have

$$\begin{aligned}
F_g(f+\lambda h) &= \int_\Omega |(f+\lambda h)\circ x(\omega) - g\circ y(\omega)|^2 \,\mathrm{d}P(\omega) \\
&= \int_\Omega |f\circ x(\omega) + \lambda h\circ x(\omega) - g\circ y(\omega)|^2 \,\mathrm{d}P(\omega) \\
&= \int_\Omega \Bigg(|f\circ x(\omega) - g\circ y(\omega)|^2 + 2\mathrm{Re}\left[(f\circ x(\omega) - g\circ y(\omega))\overline{\lambda h\circ x(\omega)}\right] \\
&\qquad + |\lambda|^2 |h\circ x(\omega)|^2 \Bigg) \mathrm{d}P(\omega) \\
&= \int_\Omega |f\circ x(\omega) - g\circ y(\omega)|^2 \,\mathrm{d}P(\omega) \\
&\qquad + 2\int_\Omega \mathrm{Re}\left[(f\circ x(\omega) - g\circ y(\omega))\overline{\lambda h\circ x(\omega)}\right] \,\mathrm{d}P(\omega) \\
&\qquad + |\lambda|^2 \int_\Omega |h\circ x(\omega)|^2 \,\mathrm{d}P(\omega) \ .
\end{aligned}$$

Hence, by (2.2), (1.8) and (1.5), we get

$$\begin{aligned}
F_g(f+\lambda h) = F_g(f) + |\lambda|^2\|h\|^2 &+ 2\mathrm{Re}\int_\Omega f\circ x(\omega)\overline{\lambda h\circ x(\omega)}\,\mathrm{d}P(\omega) \\
&- 2\mathrm{Re}\int_\Omega g\circ y(\omega)\overline{\lambda h\circ x(\omega)}\,\mathrm{d}P(\omega) \ .
\end{aligned}$$

From (1.6), (1.13) and (1.5) we conclude that

$$F_g(f+\lambda h) = F_g(f) + |\lambda|^2\|h\|^2 + 2\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] \ .$$

Therefore, for $\lambda \in B$ and $f, h \in \mathcal{F}$, we have

$$F_g(f+\lambda h) - F_g(f) = 2\mathrm{Re}\left[\lambda\Big(\langle h|f\rangle - g^*(h)\Big)\right] + |\lambda|^2\|h\|^2 \ . \tag{2.3}$$

Fix $f \in \mathcal{F}$ satisfying $\langle h|f\rangle = g^*(h),\ h \in \mathcal{F}$. Applying (2.3) with $\lambda := 1$ we obtain

$$F_g(f+h) - F_g(f) = \|h\|^2 \geq 0$$

and so

$$F_g(f+h) \geq F_g(f), \qquad h \in \mathcal{F} \ ,$$

which means that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$.

Conversely, assume now that $f \in \mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Then from (2.3) we conclude that

$$2\mathrm{Re}\,[\lambda(\langle h|f\rangle - g^*(h))] + |\lambda|^2\|h\|^2 \geq 0 \ , \qquad h \in \mathcal{F},\ \lambda \in B \ . \tag{2.4}$$

Replacing h by $-h$ in (2.4) we get

$$-2\mathrm{Re}\,[\lambda(\langle h|f\rangle - g^*(h))] + |\lambda|^2\|h\|^2 \geq 0 \ . \tag{2.5}$$

Combining (2.4) and (2.5) we can see that

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

Fixing $h \in \mathcal{F},\ \alpha \in \mathbb{R}$ and assuming that $\lambda = |\lambda|e^{i\alpha}$ we get

$$-\frac{1}{2}|\lambda|\|h\|^2 \le \operatorname{Re}\left[e^{i\alpha}(\langle h|f\rangle - g^*(h))\right] \le \frac{1}{2}|\lambda|\|h\|^2 \ .$$

In the limiting case as $|\lambda| \to 0$, the following equality holds

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

Choosing $\alpha \in \{0, \frac{\pi}{2}\}$ we conclude that $\langle h|f\rangle - g^*(h) = 0$ for $h \in \mathcal{F}$, which completes the proof. □

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

$$\Theta := \{h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \colon \|h\| = 0\}$$

is linear. We call it the *null set* of $\mathbf{H}(\mathfrak{P})$. As a matter of fact 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:

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

Corollary 2.2. *If $\mathcal{F} \neq \emptyset$ is a linear set in $\mathbf{H}(\mathfrak{P})$ and $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$, then*

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F} \cap \operatorname{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g) \ . \tag{2.6}$$

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

Proof. Fix $f, h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. If $\|h\| = 0$, then by the Schwarz inequality (1.9) and (1.14) it follows that

$$|\langle h|f\rangle| \le \|h\|\|f\| = 0 \quad \text{and} \quad |g^*(h)| \le \left(\int_\Omega |g \circ y(\omega)|^2 \mathrm{d}P(\omega)\right)^{1/2} \|h\| = 0 \ .$$

Hence

$$\langle h|f\rangle = 0 = g^*(h) \ , \quad f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x),\ h \in \Theta \ . \tag{2.7}$$

Assume that $f \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$ and $h \in \Theta + \mathcal{F}$ are given. Then $h = h_0 + h_1$ for some $h_0 \in \Theta$ and $h_1 \in \mathcal{F}$. Applying now (2.7) and Theorem 2.1 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, $f \in \mathcal{F} \subset \Theta + \mathcal{F}$. From Theorem 2.1 it follows that $f \in \mathcal{F} \cap \operatorname{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g)$, and so

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \mathcal{F} \cap \operatorname{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g) \ . \tag{2.8}$$

Conversely, assume now that $f \in \mathcal{F} \cap \operatorname{Reg}(\Theta + \mathcal{F}, \mathfrak{P}_g)$ and $h \in \mathcal{F}$ are given. Since $h \in \Theta + \mathcal{F}$, we conclude from Theorem 2.1, that

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

Thus applying Theorem 2.1 once more, we get $f \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$, and so

$$\mathcal{F} \cap \operatorname{Reg}(\Theta \cap \mathcal{F}, \mathfrak{P}_g) \subset \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) .$$

Combining this inclusion with the inclusion (2.8) we derive the equality (2.6). Since $\Theta \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, the equalities in (2.7) hold for all $f, h \in \Theta$. Then Theorem 2.1 yields $\operatorname{Reg}(\Theta, \mathfrak{P}_g) \supset \Theta$, whereas the opposite inclusion is obvious.

Thus $\operatorname{Reg}(\Theta, \mathfrak{P}_g) = \Theta$. If now $\mathcal{F} \subset \Theta$, then the equality (2.6) takes the form $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F}$, which proves the theorem. □

By $S^\perp$ we denote the orthogonal complement of $S \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ in the space $\mathbf{H}(\mathfrak{P})$, i.e.,

$$S^\perp := \{f \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \colon \langle h|f\rangle = 0 \text{ for } h \in S\} .$$

Theorem 2.3. *If $\mathcal{F} \neq \emptyset$ is a closed and linear set in $\mathbf{H}(\mathfrak{P})$ and $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$, then $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ and $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + f$ for each $f \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Moreover, if $\mathcal{F} \subset S := (g^*)^{-1}(0)$, then $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta$. Otherwise $(\mathcal{F} \cap S)^\perp \cap \mathcal{F} \setminus \Theta \neq \emptyset$ and*

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + \frac{\overline{g^*(h)}}{\|h\|^2} h , \qquad h \in (\mathcal{F} \cap S)^\perp \cap \mathcal{F} \setminus \Theta . \tag{2.9}$$

Proof. Assume that $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ and choose arbitrarily $f \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$ and $f' \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$. If $f' \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$ then, by Theorem 2.1,

$$\langle h|f\rangle = g^*(h), \quad h \in \mathcal{F} , \tag{2.10}$$

and $\langle h|f'\rangle = g^*(h)$ for $h \in \mathcal{F}$. Hence, setting $h := f - f'$ we conclude from (2.10) that

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

Thus $f' \in \Theta + f$ for $f' \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$, and so $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \subset \Theta + f$. Conversely, suppose that $f' \in \Theta + f$. Then, by Schwarz inequality (1.9), we see that for every $h \in \mathcal{F}$,

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

Hence, and by (2.10), we get $\langle h|f'\rangle = \langle h|f\rangle = g^*(h)$ for $h \in \mathcal{F}$. Since $\mathcal{F}$ is closed and linear in $\mathbf{H}(\mathfrak{P})$, we see that $\Theta \subset \mathcal{F}$ and so $\Theta + f \subset \mathcal{F}$. Applying Theorem 2.1 we see that $f' \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$ for $f' \in \Theta + f$, and so $\Theta + f \subset \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$. This inclusion together with the inverse one yields the equality $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + f$, provided $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, and so we obtain the following implication

$$\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset \quad \Rightarrow \quad \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + f . \tag{2.11}$$

Assume now that $\mathcal{F} \subset S$. Then

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

which shows, by Theorem 2.1, that $\theta \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g)$. Hence and by (2.11) we see that $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta + \theta = \Theta$. It remains to consider the case where the inclusion $\mathcal{F} \subset S$ does not hold. If so, then $\mathcal{F} \cap S \neq \mathcal{F}$. By the assumption $\mathcal{F}$ is a closed set in $\mathbf{H}(\mathfrak{P})$. Since $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$, g^* is a continuous functional on $\mathbf{H}(\mathfrak{P})$, and so S is also a closed set in $\mathbf{H}(\mathfrak{P})$. Therefore $\mathcal{F} \cap S$ is a closed set in $\mathbf{H}(\mathfrak{P})$, and consequently

$$\Theta \subset \mathcal{F} \cap S \neq \mathcal{F} \ . \tag{2.12}$$

Hence $\mathcal{F} \backslash (\mathcal{F} \cap S) \neq \emptyset$. Since $\mathcal{F} \cap S$ is closed in $\mathbf{H}(\mathfrak{P})$, it follows that each $h \in \mathcal{F} \backslash (\mathcal{F} \cap S)$ has an orthogonal projection h_S onto $\mathcal{F} \cap S$, i.e.,

$$h_S \in \mathcal{F} \cap S \quad \text{and} \quad \langle h - h_S | h' \rangle = 0 \ , \qquad h' \in \mathcal{F} \cap S \ . \tag{2.13}$$

Hence $h - h_S \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F}$. If $h - h_S \in \Theta$, then from (2.12) and (2.13) it follows that $h = h_S + (h - h_S) \in \mathcal{F} \cap S + \Theta = \mathcal{F} \cap S$, which is impossible. Therefore $h - h_S \notin \Theta$, and so $h - h_S \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \setminus \Theta$. Thus $(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \setminus \Theta \neq \emptyset$. Given $h \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \setminus \Theta$ we see that $\|h\| \neq 0$, and so $g^*(h) \neq 0$. Hence, for each $h' \in \mathcal{F}$,

$$h'_S := h' - \frac{g^*(h')}{g^*(h)} h \in \mathcal{F} \cap S \qquad \text{and} \qquad h' - h'_S = \frac{g^*(h')}{g^*(h)} h \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \ . \tag{2.14}$$

Since

$$\frac{\overline{g^*(h)}}{\|h\|^2} h \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \ ,$$

we conclude from (2.14) that

$$\begin{aligned} \left\langle h' \middle| \frac{\overline{g^*(h)}}{\|h\|^2} h \right\rangle &= \left\langle h' - h'_S \middle| \frac{\overline{g^*(h)}}{\|h\|^2} h \right\rangle = \left\langle \frac{g^*(h')}{g^*(h)} h \middle| \frac{\overline{g^*(h)}}{\|h\|^2} h \right\rangle \\ &= \frac{g^*(h')}{g^*(h)} \overline{\left(\frac{\overline{g^*(h)}}{\|h\|^2} \right)} \langle h | h \rangle = g^*(h') \ , \qquad h' \in \mathcal{F} \ . \end{aligned}$$

Applying now Theorem 2.1, we see that

$$f := \frac{\overline{g^*(h)}}{\|h\|^2} h \in \operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \ , \qquad h \in (\mathcal{F} \cap S)^{\perp} \cap \mathcal{F} \setminus \Theta \ . \tag{2.15}$$

Therefore $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$ and, combining (2.15) with (2.11), we derive the equality (2.9) provided the inclusion $\mathcal{F} \subset S$ does not hold.
In both the cases $\operatorname{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$, which completes the proof. □

3. Calculating procedure of the regression functions

Write $\mathbb{Z}_{p,q} := \{n \in \mathbb{Z} \colon p \leq n \leq q\}$ and $\mathbb{Z}_p := \{n \in \mathbb{Z} \colon p \leq n\}$ for $p, q \in \mathbb{Z}$. In particular $\mathbb{N} = \mathbb{Z}_1$. Given a nonempty set $S \subset \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, we denote by $\operatorname{lin}(S)$ the set 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 $\mathbb{Z}_{1,n} \ni k \mapsto v_k \in S$. It is easy to check that $\operatorname{lin}(S)$ is the smallest linear subset of $\mathrm{L}^2(A, \mathcal{A}_x, P_x)$ containing S.

Assume that $\mathcal{F}$ is arbitrarily chosen linear and closed set in the space $\mathbf{H}(\mathfrak{P})$ and $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ is given. Then by Theorem 2.3 we conclude that $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) \neq \emptyset$. Moreover, Theorem 2.3 enables us to find regression functions in $\mathcal{F}$ with respect to $\mathfrak{P}_g$ provided we can determine the linear set $(\mathcal{F} \cap S)^{\perp} \cap \mathcal{F}$. This is rather difficult task, in general. However in the case where the set $\mathcal{F}$ is finitely dimensional we can effectively calculate all the regression functions in $\mathcal{F}$ with respect to $\mathfrak{P}_g$ in terms of a given base of this space. Obviously, this case is most essential from the practical point of view and will be considered later on.

For every $f, h \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, we will write $f \perp h$ if $\langle f|h \rangle = 0$. 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 $\mathbf{H}(\mathfrak{P})$, 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, $\sum_{k=1}^{2} \mathcal{F}_k = \mathcal{F}_1 + \mathcal{F}_2$.

Theorem 3.1. *Given $p \in \mathbb{N}$ let $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \setminus \Theta$ be an orthogonal sequence in $\mathbf{H}(\mathfrak{P})$, i.e.,*

$$h_k \perp h_j, \qquad k, j \in \mathbb{Z}_{1,p}\,,\ k \neq j\ . \tag{3.1}$$

If $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$, then

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \right\}\ , \tag{3.2}$$

where

$$\mathcal{F} := \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\})\ . \tag{3.3}$$

Proof. Fix $p \in \mathbb{N}$ and a sequence $\mathbb{Z}_{1,p} \ni k \mapsto h_k \in \mathrm{L}^2(A, \mathcal{A}_x, P_x) \setminus \Theta$ satisfying the assumptions. From (3.3) and (3.1) it follows that $\mathcal{F}_0 := \Theta + \mathcal{F}$ is a closed set in $\mathbf{H}(\mathfrak{P})$. Therefore $\mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) \neq \emptyset$ by the assumption $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ and Theorem 2.3. If $g^*(h_k) = 0$ for $k \in \mathbb{Z}_{1,p}$, then by (3.3), $\mathcal{F}_0 \subset S := (g^*)^{-1}(0)$. From Theorem 2.3 it follows that $\mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) = \Theta$. Hence, and by Corollary 2.2, we conclude that $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \Theta \cap \mathcal{F}$. Fix $h \in \mathcal{F} \cap \Theta$. By (3.3) there exists a sequence $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B$ such that $h = \sum_{k=1}^{p} \lambda_k h_k$. From (3.1) it follows that

$$\sum_{k=1}^{p} |\lambda_k|^2 \|h_k\|^2 = \|h\|^2 = 0.$$

By the assumption, $\|h_k\| > 0$ for $k \in \mathbb{Z}_{1,p}$. Therefore $\lambda_k = 0$ for $k \in \mathbb{Z}_{1,p}$ and so $h = \theta$. Consequently

$$\mathcal{F} \cap \Theta = \{\theta\}\ . \tag{3.4}$$

Thus $\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\theta\}$, and so the equality (3.2) holds.
Assume, in contrary, that $g^*(h_k) \neq 0$ for some $k \in \mathbb{Z}_{1,p}$. Then $\mathcal{F}_0 \setminus S \neq \emptyset$ and applying again Theorem 2.3 we can see that $(\mathcal{F}_0 \cap S)^{\perp} \cap \mathcal{F}_0 \setminus \Theta \neq \emptyset$ as well as that the equality (2.9) holds. Thus we have to find an element $h \in \mathcal{F}$ such that $h \in (\mathcal{F}_0 \cap S)^{\perp} \cap \mathcal{F}_0 \setminus \Theta$.

Then by (3.3) there exists a sequence $\mathbb{Z}_{1,p} \ni k \mapsto \lambda_k \in B$ such that

$$h = \sum_{k=1}^{p} \lambda_k \cdot h_k \ . \tag{3.5}$$

If $p = 1$, then $h = \lambda_1 h_1$ and $g^*(h_1) \neq 0$. Hence $\lambda_1 \neq 0$, and $h_1 \in \mathcal{F}_0 \setminus \Theta$. Moreover, for any $f \in \mathcal{F}_0 \cap S$ there exist $\lambda \in B$ and $f_0 \in \Theta$ such that $f = f_0 + \lambda h_1$. Since $f \in S$, we obtain

$$0 = g^*(f) = g^*(f_0) + \lambda g^*(h_1) = \lambda g^*(h_1) \ ,$$

and so $\lambda = 0$. Therefore, $f = f_0 \in \Theta$, which gives $\langle h_1 | f \rangle = 0$. Hence $h_1 \in (\mathcal{F}_0 \cap S)^{\perp}$, and we see that $h = \lambda_1 h_1 \in (\mathcal{F}_0 \cap S)^{\perp} \cap \mathcal{F}_0 \setminus \Theta$. Then Theorem 2.3 leads to

$$\mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) = \Theta + \frac{\overline{g^*(h)}}{\|h\|^2} h = \Theta + \frac{\overline{g^*(h_1)}}{\|h_1\|^2} h_1 \ . \tag{3.6}$$

It remains to consider the case where $p > 1$. Without lost of generality we may assume now that $g^*(h_1) \neq 0$. Since $h_k - \frac{g^*(h_k)}{g^*(h_1)} h_1 \in S$ for $k \in Z_{1,p}$ and $h \in (\mathcal{F}_0 \cap S)^{\perp} \cap \mathcal{F}_0$ we have

$$h \perp h_k - \frac{g^*(h_k)}{g^*(h_1)} h_1, \quad k \in Z_{1,p} \ .$$

Combining this with (3.1) and (3.5) we see that for each $j \in \mathbb{Z}_{1,p}$,

$$\begin{aligned}
0 =& \Big\langle h \Big| h_j - \frac{g^*(h_j)}{g^*(h_1)} h_1 \Big\rangle = \langle h | h_j \rangle - \Big\langle h \Big| \frac{g^*(h_j)}{g^*(h_1)} h_1 \Big\rangle \\
=& \Big\langle \sum_{k=1}^{p} \lambda_k \cdot h_k \Big| h_j \Big\rangle - \overline{\left(\frac{g^*(h_j)}{g^*(h_1)}\right)} \Big\langle \sum_{k=1}^{p} \lambda_k \cdot h_k \Big| h_1 \Big\rangle \\
=& \sum_{k=1}^{p} \lambda_k \langle h_k | h_j \rangle - \overline{\left(\frac{g^*(h_j)}{g^*(h_1)}\right)} \sum_{k=1}^{p} \lambda_k \langle h_k | h_1 \rangle \\
=& \lambda_j \langle h_j | h_j \rangle - \lambda_1 \overline{\left(\frac{g^*(h_j)}{g^*(h_1)}\right)} \langle h_1 | h_1 \rangle = \lambda_j \|h_j\|^2 - \lambda_1 \overline{\left(\frac{g^*(h_j)}{g^*(h_1)}\right)} \|h_1\|^2 \ .
\end{aligned}$$

Hence

$$\lambda_j = \frac{\lambda_1}{\|h_j\|^2} \overline{\left(\frac{g^*(h_j)}{g^*(h_1)}\right)} \|h_1\|^2, \qquad j \in \mathbb{Z}_{1,p} \ .$$

This together with (3.5) leads to

$$h = \sum_{k=1}^{p} \lambda_k \cdot h_k = \sum_{k=1}^{p} \frac{\lambda_1}{\|h_k\|^2} \overline{\left(\frac{g^*(h_k)}{g^*(h_1)}\right)} \|h_1\|^2 \cdot h_k = \frac{\lambda_1}{g^*(h_1)} \|h_1\|^2 \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \ ,$$

whence $\lambda_1 \neq 0$. By (3.1) we see that

$$\|h\|^2 = \left| \frac{\lambda_1}{g^*(h_1)} \|h_1\|^2 \right|^2 \cdot \left\| \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \right\|^2 = \frac{|\lambda_1|^2 \cdot \|h_1\|^4}{|g^*(h_1)|^2} \cdot \sum_{k=1}^{p} \frac{|g^*(h_k)|^2}{\|h_k\|^2} \ .$$

Moreover,

$$\begin{aligned}\overline{g^*(h)} =&\overline{g^*\left(\frac{\lambda_1}{g^*(h_1)}\|h_1\|^2 \cdot \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k\right)} \\ =&\frac{\overline{\lambda_1}\|h_1\|^2}{g^*(h_1)} \cdot \sum_{k=1}^{p} \frac{g^*(h_k)}{\|h_k\|^2} \cdot \overline{g^*(h_k)} = \frac{\overline{\lambda_1}\|h_1\|^2}{g^*(h_1)} \cdot \sum_{k=1}^{p} \frac{|g^*(h_k)|^2}{\|h_k\|^2} \ .\end{aligned}$$

Applying now (2.9) we obtain

$$\begin{aligned}\mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) =&\Theta + \frac{\overline{g^*(h)}}{\|h\|^2} \cdot h = \Theta + \frac{\frac{\overline{\lambda_1}\|h_1\|^2}{g^*(h_1)} \cdot \sum_{k=1}^{p} \frac{|g^*(h_k)|^2}{\|h_k\|^2}}{\frac{|\lambda_1|^2 \cdot \|h_1\|^4}{|g^*(h_1)|^2} \cdot \sum_{k=1}^{p} \frac{|g^*(h_k)|^2}{\|h_k\|^2}} \cdot h \\ =&\Theta + \frac{\overline{g^*(h_1)}}{\lambda_1\|h_1\|^2} \cdot \frac{\lambda_1}{g^*(h_1)} \cdot \|h_1\|^2 \cdot \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} \cdot h_k \\ =&\Theta + \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \ .\end{aligned}$$

Hence, and by (3.6), we see that for each $p \in \mathbb{N}$,

$$\mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) = \Theta + f, \tag{3.7}$$

where, in view of (3.3),

$$f := \sum_{k=1}^{p} \frac{\overline{g^*(h_k)}}{\|h_k\|^2} h_k \in \mathrm{lin}(\{h_k \colon k \in \mathbb{Z}_{1,p}\}) = \mathcal{F} \ . \tag{3.8}$$

From Corollary 2.2, (3.7), (3.8) and (3.4) it follows that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \mathcal{F} \cap \mathrm{Reg}(\mathcal{F}_0, \mathfrak{P}_g) = \mathcal{F} \cap (\Theta + f) = (\mathcal{F} \cap \Theta) + f = \{f\} \ .$$

This yields the equality (3.2), which completes the proof. □

As far as applications are concerned we will study theoretic models $\mathcal{F}$ spanned by sequences $\mathbb{Z}_{1,p} \ni k \mapsto h_k$ which are not, in general, orthogonal in the space $\mathbf{H}(\mathfrak{P})$, because the pseudo-inner product $\langle \cdot | \cdot \rangle$ depends on the empirical data function $x \colon \Omega \to A$ and probability measure P. Therefore we can not apply Theorem 3.1 directly. However, in such a case we may orthogonalize those sequences. To this end we may use the generalized Gram - Schmidt orthogonalization method, saying that,

$$h'_1 := h_1 \quad \text{and} \quad h'_n := h_n - \sum_{k=1}^{n-1} \lambda(h_n, h'_k) \cdot h'_k \ , \qquad n \in \mathbb{Z}_{2,p} \ , \tag{3.9}$$

where λ is defined by

$$\mathrm{L}^2(A,\mathcal{A}_x,P_x)\times\mathrm{L}^2(A,\mathcal{A}_x,P_x)\ni(u,v)\mapsto\lambda(u,v):=\begin{cases}\dfrac{\langle u|v\rangle}{\|v\|^2} & \text{if}\quad \|v\|>0\ ,\\ 0 & \text{if}\quad \|v\|=0\ .\end{cases}\tag{3.10}$$

Corollary 3.2. *Given $p\in\mathbb{N}$ and $\mathbb{Z}_{1,p}\ni k\mapsto h_k\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)$ let $\mathbb{Z}_{1,p}\ni k\mapsto h'_k$ be a sequence defined by* (3.9). *If $g\in\mathrm{L}^2(B,\mathcal{A}_y,P_y)$ and*

$$\|h'_k\|>0\ ,\qquad k\in\mathbb{Z}_{1,p}\ ,\tag{3.11}$$

then

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\left\{\sum_{k=1}^{p}\frac{\overline{g^*(h'_k)}}{\|h'_k\|^2}h'_k\right\}\ ,\tag{3.12}$$

where $\mathcal{F}$ is given by (3.3).

Proof. Under the assumption we see that $h'_k\perp h'_l$ for $k,l\in\mathbb{Z}_{1,p}$ such that $k\neq l$. From (3.3) and (3.9) it follows that $\mathrm{lin}(\{h'_k\colon k\in\mathbb{Z}_{1,p}\})=\mathcal{F}$. Moreover, by (3.11), $h'_k\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)\setminus\Theta$ for $k\in\mathbb{Z}_{1,p}$. Thus, applying Theorem 3.1 for the sequence $\mathbb{Z}_{1,p}\ni k\mapsto h_k$, replaced by its orthogonal associate $\mathbb{Z}_{1,p}\ni k\mapsto h'_k$ we derive the equality (3.12), which is our claim. □

Remark 3.3. From [6, Lemma 5.2] it follows the condition (3.11) holds if and only if a sequence $\mathbb{Z}_{1,p}\ni k\mapsto h_k$ is linearly independent and $\mathcal{F}\cap\Theta=\{\theta\}$. In particular, the condition (3.11) holds provided a sequence $\mathbb{Z}_{1,p}\ni k\mapsto h_k$ is linearly independent and the functional is a norm in $(\mathrm{L}^2(A,\mathcal{A}_x,P_x),+,\cdot)$.

4. Examples and complementary remarks

In this section we present examples and comments which illustrate our considerations from the previous section. From now on we always assume that $\mathfrak{P}=(A,B,\delta;x,y)$ is a given probabilistic regression structure determined by a probability space $\mathcal{P}=(\Omega,\mathcal{A},P)$ and $g\in\mathrm{L}^2(B,\mathcal{A}_y,P_y)$ is arbitrarily fixed.

Example 4.1. Let us consider the case where the functional model $\mathcal{F}$ is spanned by one arbitrarily fixed function $h_1\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)\setminus\Theta$, i.e., $\mathcal{F}=\mathrm{lin}(\{h_1\})$. Applying Theorem 3.1 we can see that

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)=\left\{\frac{\overline{g^*(h_1)}}{\|h_1\|^2}h_1\right\}\ .\tag{4.1}$$

Using the expected value operator for the probability space $\mathcal{P}$ we conclude from the formula (1.13) that

$$g^*(h_1)=\int_\Omega(h_1\circ x(\omega))\cdot\overline{g\circ y(\omega)}\mathrm{d}P(\omega)=\mathrm{E}[(h_1\circ x)\cdot\overline{g\circ y}]\ ,\tag{4.2}$$

and from the formula (1.8) and Theorem 1.2 that

$$\|h_1\|^2 = \int_A |h_1(t)|^2 \mathrm{d}P_x(t) = \int_\Omega |h_1 \circ x(\omega)|^2 \mathrm{d}P(\omega) = \mathrm{E}[|h_1 \circ x|^2] \ . \tag{4.3}$$

Hence we can rewrite (4.1) in terms of the expected value as follows

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \frac{\mathrm{E}[(\overline{h_1} \circ x) g \circ y]}{\mathrm{E}[|h_1 \circ x|^2]} \cdot h_1 \right\} \ . \tag{4.4}$$

Given $\alpha \in \mathbb{Z}_0$ suppose that $A = B$, $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ and $h_1 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ where $h_1(t) := t^\alpha$ and $g(t) := t$ for $t \in B$. Then

$$\mathrm{E}[(\overline{h_1} \circ x)(g \circ y)] = \mathrm{E}[\overline{x}^\alpha \cdot y] \quad \text{and} \quad \mathrm{E}[|h_1 \circ x|^2] = \mathrm{E}[|x|^{2\alpha}] \ ,$$

and so (4.4) implies

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}) = \left\{ A \ni t \mapsto \frac{\mathrm{E}[\overline{x}^\alpha \cdot y]}{\mathrm{E}[|x|^{2\alpha}]} \cdot t^\alpha \right\} \ , \tag{4.5}$$

provided x is not equal 0 a.s. on Ω.
If x is a real random variable, then putting $\alpha := 1$ in (4.5) we see that $\mathrm{E}[xy]$ can be expressed by means of regression functions $\mathrm{Reg}(\mathcal{F}, \mathfrak{P})$ and $\mathrm{E}[x^2]$. Putting $\alpha := 0$ in (4.5) we obtain

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}) = \{A \in t \mapsto \mathrm{E}[y]\} \ . \tag{4.6}$$

Notice that the equality (4.6) is still valid even if $A \neq B$.

Example 4.2. Let us consider the case where the functional model $\mathcal{F}$ is spanned by two arbitrarily fixed functions $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, i.e., $\mathcal{F} = \mathrm{lin}(\{h_1, h_2\})$. Suppose that $\|h_1'\| > 0$ and $\|h_2'\| > 0$, where $\mathbb{Z}_{1,2} \ni k \mapsto h_k'$ is a sequence defined by (3.9). Applying Corollary 3.2 we can see that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \frac{\overline{g^*(h_1')}}{\|h_1'\|^2} h_1' + \frac{\overline{g^*(h_2')}}{\|h_2'\|^2} h_2' \right\} \ , \tag{4.7}$$

where, according to (3.9),

$$h_1' := h_1 \quad \text{and} \quad h_2' := h_2 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} h_1 \ . \tag{4.8}$$

Hence $h_2' \perp h_1$, and consequently

$$\begin{aligned} \|h_2'\|^2 =& \langle h_2' | h_2' \rangle = \langle h_2' | h_2 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} h_1 \rangle = \langle h_2' | h_2 \rangle \\ =& \langle h_2 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} h_1 | h_2 \rangle = \|h_2\|^2 - \frac{|\langle h_2 | h_1 \rangle|^2}{\|h_1\|^2} \ . \end{aligned} \tag{4.9}$$

Setting

$$\alpha_2 := \frac{\overline{g^*(h_2)} \|h_1\|^2 - \overline{g^*(h_1) \langle h_2 | h_1 \rangle}}{\|h_2\|^2 \|h_1\|^2 - |\langle h_2 | h_1 \rangle|^2} \quad \text{and} \quad \alpha_1 := \frac{\overline{g^*(h_1)} - \langle h_2 | h_1 \rangle \alpha_2}{\|h_1\|^2} \tag{4.10}$$

we conclude from (4.7), (4.8) and (4.9) that

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \{\alpha_2 h_2 + \alpha_1 h_1\} \ . \tag{4.11}$$

We can calculate the coefficients α_2 and α_1 by means of the expected value operator E for the probability space $\mathcal{P}$ using the following equalities

$$g^*(h_k) = \mathrm{E}[(h_k \circ x) \cdot \overline{g \circ y}] \quad \text{and} \quad \|h_k\|^2 = \mathrm{E}[|h_k \circ x|^2] \ , \quad k \in \mathbb{Z}_{1,2} \tag{4.12}$$

as well as

$$\langle h_2 | h_1 \rangle = \int_A h_2(t) \cdot \overline{h_1(t)} \mathrm{d}P_x(t) = \mathrm{E}[(h_2 \circ x)\overline{(h_1 \circ x)}] \ . \tag{4.13}$$

To prove them we appeal to the equalities (4.2), (4.3), (1.6) and Theorem 1.2.
In particular, suppose that $A = B$, $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$ and $h_1, h_2 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ where $g(t) := t$, $h_1(t) := 1$ and $h_2(t) := t$ for $t \in B$. Applying now the equalities (4.12), (4.13) we can rewrite the formulas (4.10) as

$$\alpha_2 = \frac{\mathrm{E}[\overline{x} \cdot y] - \mathrm{E}[\overline{x}] \cdot \mathrm{E}[y]}{\mathrm{E}[|x|^2] - (\mathrm{E}[x])^2} \quad \text{and} \quad \alpha_1 = \mathrm{E}[y] - \mathrm{E}[x] \cdot \alpha_2 \tag{4.14}$$

provided x is a not a constant a.s. on Ω. Therefore the coefficients α_2 and α_1 given by (4.14) coincide with the classical linear regression coefficients in the case of real random variables, cf. [3], [4].

Example 4.3. Let us consider the case where the functional model $\mathcal{F}$ is spanned by three arbitrarily fixed functions $h_1, h_2, h_3 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$, i.e., $\mathcal{F} = \mathrm{lin}(\{h_1, h_2, h_3\})$. Suppose that $\|h'_k\| > 0$ for $k \in \mathbb{Z}_{1,3}$, where $\mathbb{Z}_{1,3} \ni k \mapsto h'_k$ is a sequence defined by (3.9). Applying Corollary 3.2 we obtain

$$\mathrm{Reg}(\mathcal{F}, \mathfrak{P}_g) = \left\{ \sum_{k=1}^{3} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \right\} \ , \tag{4.15}$$

where, by (3.9), we have

$$\begin{aligned} h'_1 &= h_1 \ , \\ h'_2 &= h_2 - \frac{\langle h_2 | h_1 \rangle}{\|h_1\|^2} \cdot h_1 \ , \\ h'_3 &= h_3 + \eta \cdot h_2 - \frac{\langle h_3 | h_1 \rangle + \eta \cdot \langle h_2 | h_1 \rangle}{\|h_1\|^2} \cdot h_1 \\ &\text{and} \qquad \eta := \frac{\langle h_3 | h_1 \rangle \langle h_2 | h_1 \rangle - \langle h_3 | h_2 \rangle \langle h_1 | h_1 \rangle}{\|h_2\|^2 \|h_1\|^2 - |\langle h_2 | h_1 \rangle|^2} \ . \end{aligned}$$

In particular, suppose that x and y are independent real random variables with normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ respectively. Then, cf. [4],

$$\mathrm{E}[(x - \mu_1)^{2s+1}] = 0 \quad \text{and} \quad \mathrm{E}[(x - \mu_1)^{2s}] = (2s-1)!! \cdot \sigma_1^{2s} \ , \quad s \in \mathbb{N} \ . \tag{4.16}$$

Setting $A := \mathbb{R}$ and $B := \mathbb{R}$ we see that $h_1, h_2, h_3 \in \mathrm{L}^2(A, \mathcal{A}_x, P_x)$ and $g \in \mathrm{L}^2(B, \mathcal{A}_y, P_y)$, where $g(t) := t$, $h_1(t) := 1$, $h_2(t) := t$ and $h_3(t) := t^2$ for $t \in B$.

Using Theorem 1.2 we conclude from the formula (1.6) that

$$\begin{aligned}\langle h_n|h_k\rangle &= \int_{\mathbb{R}} h_n(t)\overline{h_k(t)}\,\mathrm{d}P_x(t) = \int_{\mathbb{R}} t^{n+k-2}\,\mathrm{d}P_x(t) \\ &= \int_{\Omega} x^{n+k-2}(\omega)\,\mathrm{d}P(\omega) = \mathrm{E}[x^{n+k-2}]\ , \qquad n,k\in\mathbb{Z}_{1,3}\ .\end{aligned} \tag{4.17}$$

Combining (4.17) with (4.16) we calculate

$$\begin{aligned}&\langle h_1|h_1\rangle = 1\ , &&\langle h_2|h_1\rangle = \mu_1\ , &&\langle h_3|h_1\rangle = \mu_1^2+\sigma_1^2\ ,\\ &\langle h_2|h_2\rangle = \mu_1^2+\sigma_1^2\ , &&\langle h_3|h_2\rangle = (\mu_1^2+3\sigma_1^2)\mu_1\ .\end{aligned}$$

Hence $\eta = -2\mu_1$ and so

$$h'_1 = h_1\ ,\qquad h'_2 = h_2-\mu_1 h_1\ ,\qquad h'_3 = h_3 - 2\mu_1 h_2 + (\mu_1^2-\sigma_1^2)h_1\ . \tag{4.18}$$

Since x and y are independent, we conclude from the formula (1.13) that

$$\begin{aligned}g^*(h_k) &= \int_{\Omega} h_k\circ x(\omega)\cdot\overline{g\circ y(\omega)}\,\mathrm{d}P(\omega) = \int_{\Omega} x^{k-1}(\omega)\cdot y(\omega)\,\mathrm{d}P(\omega)\\ &= \mathrm{E}[x^{k-1}\cdot y] = \mathrm{E}[x^{k-1}]\cdot\mathrm{E}[y]\ ,\qquad k\in\mathbb{Z}_{1,3}\ .\end{aligned}$$

This together with (4.18) yields

$$g^*(h'_1) = \mathrm{E}[y] = \mu_2\ ,\qquad g^*(h'_2) = 0\ ,\quad \text{and}\quad g^*(h'_3) = 0\ .$$

Using now (4.15) we obtain

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}) = \{\mathbb{R}\ni t\mapsto \mu_2\}\ .$$

In particular for $\mu_2 := 0$ we get

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

Example 4.4. Assume that $A = B$. Let $\mathcal{F}$ be a functional model consisting of all polynomials f with coefficients in B and degree $\deg f\le p-1$, where $p\in\mathbb{N}$. Setting $B\ni t\mapsto h_k(t) := t^{k-1}$ for $k\in\mathbb{Z}_{1,p}$ we see that $\mathcal{F} = \mathrm{lin}(\{h_k\colon k\in\mathbb{Z}_{1,p}\})$. Suppose that $h_k\in\mathrm{L}^2(A,\mathcal{A}_x,P_x)$ for $k\in\mathbb{Z}_{1,p}$ and $\|h'_k\|>0$ for $k\in\mathbb{Z}_{1,p}$, where $\mathbb{Z}_{1,p}\ni k\mapsto h'_k$ is a sequence defined by (3.9). Applying Corollary 3.2 we get

$$\mathrm{Reg}(\mathcal{F},\mathfrak{P}_g) = \left\{\sum_{k=1}^{p}\frac{\overline{g^*(h'_k)}}{\|h'_k\|^2}h'_k\right\}\ . \tag{4.19}$$

According to the classical definition, cf., e.g. [3], [10], by a polynomial regression of the random variable y with respect to the random variable x, we mean each polynomial $f_0\in\mathcal{F}$ such that

$$\mathrm{E}[|f\circ x-y|^2]\ge\mathrm{E}[|f_0\circ x-y|^2]\ ,\qquad f\in\mathcal{F}\ .$$

From (1.2) it follows that

$$F(f) = \mathrm{E}[|f\circ x-y|^2]\ ,\qquad f\in\mathcal{F}\ .$$

Therefore the class of all such f_0 coincides with the class $\mathrm{Reg}(\mathcal{F},\mathfrak{P})$. Suppose that $g \in \mathrm{L}^2(B,\mathcal{A}_y,P_y)$, where $g(t) := t$ for $t \in B$. Then $\mathrm{Reg}(\mathcal{F},\mathfrak{P}) = \mathrm{Reg}(\mathcal{F},\mathfrak{P}_g)$ and by (4.19) we see that there exists the unique polynomial regression $f_0 \in \mathcal{F}$ of y with respect to x and f_0 can be determined by the following equality

$$f_0 = \sum_{k=1}^{p} \frac{\overline{g^*(h'_k)}}{\|h'_k\|^2} h'_k \ .$$

Acknowledgement

Most of the presented results were obtained by the author during Monday's Seminar on Applications of Mathematical Methods at the State School of Higher Education in Chełm.

References

[1] R. B. Ash, *Real Analysis and Probability,* Probability and Mathematical Statistics, Academic Press, Inc., New York and London, 1972.

[2] P. Billingsley, *Probability and Measure,* John Wiley & Sons, Inc., New York 1995.

[3] N. R. Draper andH. Smith, *Applied Regression Analysis,* John Wiley & Sons, Inc., New York 1998.

[4] M. Fisz, *Probability Theory and Mathematical Statistics,* John Wiley & Sons, Inc., New York 1967.

[5] D. Partyka and J. Zając, *Generalized problem of regression,* Bulletin De La Société des Sciences et des Lettres De Łódź, vol. LX (2010), pp. 77–94.

[6] D. Partyka and J. Zając, *Generalized approach to the problem of regression*, Anal. Math. Phys. (2015) 5:249-320 DOI 10.1007/s13324-014-0096-7.

[7] M. M. Rao andR. J. Swift, *Probability Theory with Applications,* Springer, New York 2006.

[8] A. Rényi, *Probability Theory,* Akadémiai Kiadó, Budapest 1970.

[9] W. Rudin, *Real and complex analysis,* McGRAF-HILL, Inc., New York 1970.

[10] G. A. F. Seber and C. J. Wild, *Nonlinear Regression,* John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.

[11] G. Strang, *Linear Algebra and its Applications,* Stamford, CT: Cengage Learning, 2005.

Institute of Mathematics and Information Technology
The State School of Higher Education in Chełm
Pocztowa 54, PL-22-100 Chełm
Poland
E-mail: aszumera@pwsz.chelm.pl

Presented by Adam Paszkiewicz at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on June 25, 2018.

PROBABILISTYCZNE STRUKTURY REGRESJI

S t r e s z c z e n i e

Nowe podejście uogólniające klasyczną koncepcję regresji jest szeroko prezentowane w [5] i [6] na gruncie przestrzeni Hilberta. W niniejszym artykule wyniki tej pracy zostały przeniesione na przestrzeń probabilistyczną, gdzie uogólnione zagadnienie regresji ma postać rozwiązania problemu ekstremalnego, zdefiniowanego na przestrzeni probabilistycznej.

Słowa kluczowe: regresja nieliniowa, regresja wielomianowa, przestrzeń probabilistyczna, funkcje regresji, struktura regresji