Electronic Journal of Probability

Vol. 1 (1996) Paper no. 7, pages 1–14.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol1/paper7.abs.html

CONDITIONAL MOMENT REPRESENTATIONS FOR DEPENDENT RANDOM VARIABLES

Włodzimierz Bryc

Department of Mathematics
PO Box 210025
University of Cincinnati
Cincinnati, OH 45 221-0025

bryc@ucbeh.san.uc.edu

Abstact: The question considered in this paper is which sequences of p-integrable random variables can be represented as conditional expectations of a fixed random variable with respect to a given sequence of σ-fields. For finite families of σ-fields, explicit inequality equivalent to solvability is stated; sufficient conditions are given for finite and infinite families of σ-fields, and explicit expansions are presented.

Keywords: alternating conditional expectation, inverse problems, ACE

AMS subject classification: 62J12

Partially supported by C. P. Taft Memorial Fund and Office of Naval Research

Submitted to EJP on November 20, 1995. Final version accepted on April 13, 1996.

Conditional moment representations for dependent random variables

Włodzimierz Bryc[1]
Department of Mathematics
PO Box 210025
University of Cincinnati
Cincinnati, OH 45 221-0025

April 13, 1996

Abstract

The question considered in this paper is which sequences of p-integrable random variables can be represented as conditional expectations of a fixed random variable with respect to a given sequence of σ-fields. For finite families of σ-fields, explicit inequality equivalent to solvability is stated; sufficient conditions are given for finite and infinite families of σ-fields, and explicit expansions are presented.

1 Introduction

We analyze which sequences of random variables $\{X_j\}$ can be represented as conditional expectations

$$E(Z|\mathcal{F}_j) = X_j. \tag{1}$$

of a p-integrable random variable Z with respect to a given sequence $(\mathcal{F}_j)$ of σ-fields. The martingale theory answers this question for families of increasing σ-fields $(\mathcal{F}_j)$. We are interested in other cases which include σ-fields generated by independent, or Markov dependent (see [3]), random variables. In particular, given a random sequence ξ_j and p-integrable random variables $X_j = f_j(\xi_j)$, we analyze when there exists $Z \in L_p$ such that

$$X_j = E(Z|\xi_j). \tag{2}$$

This is motivated by our previous results for independent random variables and by the alternating conditional expectations (ACE) algorithm of Breiman & Friedman [4]. In [4] the authors are interested in the L_2-best additive prediction Z of a random variable Y based on the finite number of the predictor variables $\xi_1, \ldots, \xi_d$. The solution (ACE) is based on the fact that the best additive predictor $Z = \phi_1(\xi_1) + \ldots + \phi_d(\xi_d)$ satisfies the conditional moment constraints (2).

Relation (1) defines an inverse problem, and shares many characteristics of other inverse problems, c. f. Groetsch [9]. Accordingly, our methods partially rely on (non-constructive) functional analysis. We give sufficient conditions for the solvability of (1) in terms of *maximal correlations*. We also show that (2) has solution for finite $d < \infty$, if the joint density of $\xi_1, \ldots, \xi_d$ with respect to the product of marginals is bounded away from zero and $EX_i = EX_j$.

[1]Partially supported by C. P. Taft Memorial Fund and Office of Naval Research Grant N00014-93-0043.

We are interested in both finite, and infinite sequences, extending our previous results in [5, 6]. In this paper we concentrate on the p-integrable case with $1 < p < \infty$. The extremes $p = 1$ or $p = \infty$ seem to require different assumptions. For infinite sequences of independent r. v. all three cases $1 < p < \infty$, $p = 1$, and $p = \infty$ are completely solved in [6]. For finite sequences of dependent σ-fields, Kellerer [10] and Strassen [16] can be quoted in connection with conditional expectations problem for bounded random variables ($p = \infty$) case. For pairs of σ-fields the case $1 < p < \infty$ is solved in [5].

2 Notation and results

For $2 \leq d \leq \infty$, let $\{\mathcal{F}_j\}_{1\leq j\leq d}$ be a given family of σ-fields. By $L_p^0(\mathcal{F})$ we denote the Banach space of all p-integrable $\mathcal{F}$-measurable centered random variables, $1 \leq p \leq \infty$. By $\mathcal{E}_j$ we denote the conditional expectation with respect to $\mathcal{F}_j$. For $d < \infty$ by $\bigoplus_{j=1}^d L_p(\mathcal{F}_j)$ we denote the set of sums $Z = Z_1 + \ldots + Z_d$, where $Z_j \in L_p(\mathcal{F}_j)$.

We shall analyze the following problems.

- For all *consistent* $X_j \in L_p$ find $Z \in L_p$ satisfying (1) and such that

$$E|Z|^p \text{ is minimum.} \tag{3}$$

- For all *consistent* $X_j \in L_p$ find additive $Z \in L_p$ satisfying (1); additive means that

$$Z = \sum_{j=1}^{d} Z_j, \text{ where } Z_j \in L_p(\mathcal{F}_j). \tag{4}$$

(for $d = \infty$ the series in (4) is assumed to converge absolutely in L_p)

The above statements do not spell out the consistency conditions which will be explicit in the theorems.

Remark 2.1 *If (1) can be solved, then there exists a minimal solution Z. This can be easily recovered from the Komlós law of large numbers [11].*

2.1 Maximal correlations

Maximal correlation coefficients play a prominent role below; for another use see also [4, Section 5]. The following maximal correlation coefficient is defined in [5]. Let

$$\tilde{\rho}(\mathcal{F},\mathcal{G}) = \sup\{corr(X,Y) : X \in L_2(\mathcal{F}), Y \in L_2(\mathcal{G}), E(X|\mathcal{F}\cap\mathcal{G}) = 0\}.$$

Notice that $\tilde{\rho}(\mathcal{F},\mathcal{G}) = 0$ for independent $\mathcal{F},\mathcal{G}$ but also for increasing σ-fields $\mathcal{F} \subset \mathcal{G}$. If the intersection $\mathcal{F}\cap\mathcal{G}$ is trivial, $\tilde{\rho}$ coincides with the usual *maximal correlation coefficient*, defined in general by

$$\rho(\mathcal{F},\mathcal{G}) = \sup_{X\in L_2(\mathcal{F}),Y\in L_2(\mathcal{G})} corr(X,Y). \tag{5}$$

Given $d \leq \infty$, σ-fields $\{\mathcal{F}_j\}_{j\leq d}$, and a finite subset $T \subset I := \{1, 2, \ldots, d\}$ put

$$\mathcal{F}_T = \sigma(\mathcal{F}_j : j \in T).$$

Define *pairwise maximal correlation* r by

$$r = \sup_{i \neq j} \rho(\mathcal{F}_i, \mathcal{F}_j)$$

and *global maximal correlation*

$$R = \sup_{T \cap S = \emptyset} \rho(\mathcal{F}_T, \mathcal{F}_S).$$

For $p = 2$ a version of R based on additive r. v. will also play a role. Let

$$R_\ell = \sup \left\{ corr(U, V) : U = \sum_{j \in T} X_j, V = \sum_{j \in S} X_j, X_j \in L_2(\mathcal{F}_j), T \cap S = \emptyset \right\}.$$

Clearly, $r \leq R_\ell \leq R$. All three coefficients coincide for two σ-fields $d = 2$ case. One can easily see that $R_\ell = 0$ and $R = 1$ can happen already for $d = 3$.

2.2 Main results

In Section 2.4 we present complete solution of (1) for the two σ-fields case. For general families of σ-fields, there seems to be little hope to get existence and uniqueness results as precise as for $d = 2$. As Logan & Shepp [12] point out, complications arise even in relatively simple situations. One source of difficulties is the possibility of linear dependence between vectors in $L_p^0(\mathcal{F}_j)$. Suitable assumptions on maximal correlation coefficients exclude this possibility.

The following result extends [6, Corollary 1] to infinite sequences of dependent families of σ-fields.

Theorem 2.1 *(i) Fix $1 < p < \infty$ and suppose $R < 1$. Then equation (1) is solvable for Z for all $X_j \in L_p^0(\mathcal{F}_j)$ such that*

$$E(\sum_j |X_j|^2)^{p/2} < \infty,$$

and the solution is unique.
(ii) If $R_\ell < 1$ then for all $X_j \in L_2^0(\mathcal{F}_j)$ such that $\sum_j EX_j^2 < \infty$ there exists a unique additive solution Z to (1), and it satisfies

$$E|Z|^2 \leq \frac{1 + R_\ell}{1 - R_\ell} \sum_j E|X_j|^2.$$

If one is not interested in sharp moment estimates for Z and only finite families $d < \infty$ are of interest, then one can iterate Theorem 2.11 for a pair of σ-fields, relaxing the assumption that $R < 1$. By Lemma 3.2, iterated Theorem 2.11 yields the following.

Theorem 2.2 *If $d < \infty$,*

$$\rho_\star = \max_{1 \leq j \leq d} \rho(\mathcal{F}_{\{1,\ldots,j\}}, \mathcal{F}_{j+1}) < 1, \tag{6}$$

and $1 < p < \infty$, then equation (1) has an additive solution Z to (1) for all $X_j \in L_p^0(\mathcal{F}_j)$.

The following criterion for solvability of the additive version of (1) uses the pairwise maximal correlation r and gives explicit alternative to ACE. For $d = 2$ the assumptions are close to [5], except that we assume $p = 2$ and (implicitly) linear independence.

Theorem 2.3 *If $d < \infty$, $r < \frac{1}{d-1}$, and $p = 2$, then for all $X_j \in L_2(\mathcal{F}_j)$ with $EX_i = EX_j$ there exists a unique $Z \in L_2$ such that (1) and (4) hold.*
Moreover, the solution is given by the explicit series expansion

$$\begin{aligned} Z &= EX_1 \\ &+ \sum_{k=0}^{\infty}(-1)^k \sum_{i_1\in I} \sum_{i_2\in I\setminus i_1} \dots \sum_{i_k\in I\setminus i_{k-1}} \mathcal{E}_{i_1}\dots\mathcal{E}_{i_k} \sum_{j\in I\setminus i_k}(X_j - EX_j) \end{aligned} \tag{7}$$

(with the convention $\sum_{j\in\emptyset} X_j = 0$).
Furthermore,

$$Var(Z) \leq \frac{1}{1-r(d-1)} \sum_{j=1}^{d} Var(X_j).$$

Results in [4, Proposition 5.2] can be recovered from maximal correlation methods as follows. For finite families of σ-fields, Lemma 3.2 states inequality (12) which is equivalent to solvability of (1). This inequality is verified in Lemma 3.5 under the assumptions motivated by Breiman & Friedman [4].

Corollary 2.4 ([4], Proposition 5.2) *If $d < \infty$, vector spaces $L_2^0(\mathcal{F}_j)$ are linearly independent, and for all $1 \leq j, k \leq d, k \neq j$, the operators*

$$\mathcal{E}_j : L_2^0(\mathcal{F}_k) \to L_2^0(\mathcal{F}_j)$$

are compact, then for all square integrable $X_1, \dots, X_d$ with equal means, there exists a unique additive solution Z to (1).

2.3 Conditioning with respect to random variables

We now state sufficient conditions for solvability of (1) in terms of joint distributions for finite families $d < \infty$ of σ-fields generated by random variables $\mathcal{F}_j = \sigma(\xi_j)$.

We begin with the density criterion that gives explicit estimate for R, and was motivated by [15]. By Lemma 3.2, it implies that (1) has a unique additive solution Z for all $1 < p < \infty$. Although it applies both to discrete and continuous distributions (typically, the density in the statement is with respect to the product of marginals), it is clear that the result is far from being optimal.

Theorem 2.5 *Suppose there is a product probability measure $\mu = \mu_1 \otimes \dots \otimes \mu_d$ such that the distribution of $\xi_1, \dots, \xi_d$ on $\mathbb{R}^d$ is absolutely continuous with respect to μ and its density f is bounded away from zero and infinity, $0 < b \leq f(x_1, \dots, x_d) \leq B < \infty$. Then*

$$R \leq 1 - \frac{b}{B^2}.$$

Next we give sufficient conditions in terms of bivariate densities only. The result follows from [14, page 106, Exercise 15] and Corollary 2.4 and is stated for completeness only.

Proposition 2.6 ([4]) *Suppose $d < \infty$ and for every pair of $i \neq j$ the density $f_{i,j}$ of the distribution of (ξ_i, ξ_j) with respect to the product measure $\mu_{i,j} = \mu_i \otimes \mu_j$ of the marginals exists and*

$$\max_{i \neq j} \int \int f_{i,j}^2(x, y) d\mu_i(x) d\mu_j(y) < \infty. \tag{8}$$

If vector spaces $L_2^0(\mathcal{F}_j)$ are linearly independent $p = 2$, then $R_\ell < 1$. In particular, (1) has a unique additive solution Z for all square integrable $X_1, \ldots, X_d$ with equal means.

In general, linear independence is difficult to verify (vide [12], where it fails). The following consequence of Proposition 2.6 gives a relevant "density criterion".

Corollary 2.7 *Suppose the density f of the distribution of $\xi_1, \ldots, \xi_d$ $(d < \infty)$ with respect to the product of marginals $\mu = \mu_1 \otimes \ldots \otimes \mu_d$ exists. If f is strictly positive, i.e.,*

$$\mu(\{(x_1, \ldots, x_d) : f(x_1, \ldots, x_d) = 0\}) = 0$$

and $\int f^2 d\mu < \infty$ then there exists an additive solution to (1) for all $X_j \in L_2(\mathcal{F}_j)$ such that $EX_i = EX_j$.

In relation to Theorem 2.5, one should note that the lower bound on the density is of more relevance. (On the other hand, in Theorem 2.5 we use the density with respect to arbitrary product measure rather than the product of marginals.)

Proposition 2.8 *Let f be the density of the absolute continuous part of the distribution of $\xi_1, \ldots, \xi_d$ $(d < \infty)$ with respect to the product of marginals $\mu = \mu_1 \otimes \ldots \otimes \mu_d$. If f is bounded away from zero, i.e., there is $b > 0$ such that $\mu(\{(x_1, \ldots, x_d) : f(x_1, \ldots, x_d) \geq b\}) = 1$, then (12) holds for all $1 < q < \infty$. In particular, for $1 < p < \infty$ and $X_j \in L_p(\mathcal{F}_j)$ such that $EX_i = EX_j$ there exists an additive solution to (1).*

2.4 Results for two σ-fields

This case is rather completely settled. Most of the results occurred in various guises in the literature. They are collected below for completeness, and to point out what to aim for in the more general case.

The following shows that for $d = 2$ there is at most one solution of (1) and (4). (Clearly, there is no Z if X_1, X_2 are not consistent, e.g., if $EX_1 \neq EX_2$.)

Proposition 2.9 *Given $X_j \in L_p(\mathcal{F}_j)$, $1 \leq p \leq \infty$, there exists at most one $Z = Z_1 + Z_2 + Z' \in L_1$ such that (1) holds, $Z_j \in L_p(\mathcal{F}_j)$, and $E(Z_j | \mathcal{F}_1 \cap \mathcal{F}_2) = 0$.*

Since best additive approximations satisfy (1), uniqueness allows to consider the inverse problem (1) instead. This is well known, c.f., [8].

Corollary 2.10 *If $p = 2$ and the best additive approximation $Z = Z_1 + Z_2 + Z'$ of $Y \in L_2$ (i.e., Z minimizing $E\left(Y - E(Y|\mathcal{F}_1 \cap \mathcal{F}_2) - (Z_1 + Z_2)\right)^2$) exists, then it is given by the solution to (1).*

The following result points out the role of maximal correlation and comes from [5].

Theorem 2.11 ([5]) *Suppose $1 < p < \infty$ is fixed. The following conditions are equivalent:*

1. *There exists a minimal solution to (1) for all consistent X_1, X_2 in $L_p(\mathcal{F}_1), L_p(\mathcal{F}_2)$ respectively;*
2. *There exists an additive solution to (1) for all consistent X_1, X_2 in $L_p(\mathcal{F}_1), L_p(\mathcal{F}_2)$ respectively;*
3. *$\tilde{\rho} < 1$.*

The consistency condition is

$$E\{X_1|\mathcal{F}_1 \cap \mathcal{F}_2\} = E\{X_2|\mathcal{F}_1 \cap \mathcal{F}_2\}.$$

Furthermore, if $E(Z|\mathcal{F}_1 \cap \mathcal{F}_2) = 0$, the minimum norm in (3) is bounded by

$$E|Z|^2 \leq \frac{1}{1-\tilde{\rho}}(EX_1^2 + EX_2^2)$$

and the bound is sharp.
The solution Z is given by the follwoing series expansion which converges in L_p

$$Z = E(X_1|\mathcal{F}_1 \cap \mathcal{F}_2) + \sum_{k=0}^{\infty}(\mathcal{E}_2\mathcal{E}_1)^k(X_2 - \mathcal{E}_2 X_1) + \sum_{k=0}^{\infty}(\mathcal{E}_1\mathcal{E}_2)^k(X_1 - \mathcal{E}_1 X_2). \tag{9}$$

Remark 2.2 *Formula (9) resembles the expansion for the orthogonal projection of L_2 onto the closure of $(L_2(\mathcal{F}_1) \oplus L_2(\mathcal{F}_2))$ (see [1]).*

$$Z = E\{Y|\mathcal{F}_1 \cap \mathcal{F}_2\} + \sum_{k=1}^{\infty}\left(\mathcal{E}_1(\mathcal{E}_2\mathcal{E}_1)^k + \mathcal{E}_2(\mathcal{E}_1\mathcal{E}_2)^k - (\mathcal{E}_1\mathcal{E}_2)^k - (\mathcal{E}_2\mathcal{E}_1)^k\right)Y. \tag{10}$$

3 Proofs

The following uniqueness result is proved in [4] for the square-integrable case $p = 2$ (the new part is $1 < p < 2$).

Lemma 3.1 *(i) If $L_p^0(\mathcal{F}_j)$ are linearly independent, $p \geq 2$, and $d < \infty$, then for every $\{X_j\}$ in $L_p(\mathcal{F}_j)$, there is at most one solution of (1) in the additive class (4).*
(ii) If inequality (12) holds with $q = 2$, then for every $\{X_j\}$ in $L_p(\mathcal{F}_j)$, $p \geq 2$, there is at most one solution of (1) in the additive class (4).
(iii) Fix $1 < p < \infty$. If there are constants c, C such that for all centered $\{X_j\} \in L_q(\mathcal{F}_j)$ inequality

$$cE(\sum_{j=1}^{d} X_j^2)^{q/2} \leq E|\sum_{j=1}^{d} X_j|^q \leq CE(\sum_{j=1}^{d} X_j^2)^{q/2} \tag{11}$$

holds for $q = p$ and for the conjugate exponent $q = \frac{p}{p-1}$, then for every $\{X_j\}$ in $L_p(\mathcal{F}_j)$ there is at most one solution of (1) in the additive class (4).

Proof of Lemma 3.1. The case $p=2$ goes as follows. Suppose $Z=Z_1+Z_2+\dots$ has $\mathcal{E}_j(Z)=0$ for all j. Then $EZ^2=\sum_j EZZ_j=\sum_j E(Z_j\mathcal{E}_j(Z))=0$. This implies that $Z_j=0$ for all j either by linear independence, or by (12).

The second part uses the existence part of the proof of Theorem 2.1. Take $Z=\sum_j Z_j$ (L_p-convergent series) such that $\mathcal{E}_j(Z)=0$. Then by (11)

$$\|Z\|_p\leq C(E(\sum_j Z_j^2)^{p/2})^{1/p}=C\sum_j E(Z_jX_j),$$

where $E(\sum_j X_j^2)^{q/2}=1$, $1/p+1/q=1$ and $X_j\in L_q^0(\mathcal{F}_j)$. The latter holds because the conjugate space to $L_q^0(\ell_2(\mathcal{F}_j))$ is $L_p^0(\ell_2(\mathcal{F}_j))$. The existence part of the proof of Theorem 2.1 implies that there is $\tilde{Z}\in L_q$ such that $\mathcal{E}_j(\tilde{Z})=X_j$ and $\tilde{Z}=\sum_j\tilde{Z}_j$ with $\tilde{Z}_j\in L_q^0(\mathcal{F}_j)$. Therefore

$$\sum_j E(Z_jX_j)=\sum_j E(Z_j\tilde{Z})=E(Z\tilde{Z})=E(Z\sum_j\tilde{Z}_j)=\sum_j E(\tilde{Z}_j\mathcal{E}_j(Z))=0.$$

This shows $E|Z|^p=0$ and by the left hand side of (11) we have $Z_j=0$ a.s. for all j. □

Proof of Proposition 2.9. Clearly $Z'=E\{Z|\mathcal{F}_1\cap\mathcal{F}_2\}$ is uniquely determined by Z and without loosing generality we may assume $Z'=0$. Suppose that $Z=Z_1+Z_2$ satisfies $\mathcal{E}_jZ=0$. Then $Z_1=TZ_1$, where $T=\mathcal{E}_1\mathcal{E}_2$. Using this iteratively, by "alternierende Verfahren" (see [13]) we get $Z_1=(\mathcal{E}_1\mathcal{E}_2)^kZ_1\to E(Z_1|\mathcal{F}_1\cap\mathcal{F}_2)=0$. By symmetry, $Z_2=0$ and the proof of uniqueness follows. □

Proof of Corollary 2.10. Without loss of generality we may assume $E(Y|\mathcal{F}_1\cap\mathcal{F}_2)=0$. For optimal $Z=Z_1+Z_2$ we have

$$\begin{aligned}\min &= E\left(Y-(Z_1+Z_2)\right)^2\\ &= E\left(Y-(\mathcal{E}_1(Y)-\mathcal{E}_1(Z_2)+Z_2)\right)^2+E\left(\mathcal{E}_1(Y)-\mathcal{E}_1(Z)\right)^2\\ &\geq \min +E\left(\mathcal{E}_1(Y)-\mathcal{E}_1(Z)\right)^2.\end{aligned}$$

Since the same analysis applies to $\mathcal{E}_2$, the optimal Z has to satisfy (1). By Theorem 2.9, there is only one such Z, so this one has to be the optimal one.□

Proof of Theorem 2.11. Let $\mathcal{L}_p^0$ denote the null space of the linear operator $E(\cdot|\mathcal{F}_1\cap\mathcal{F}_2)$ on L_p.

If $E(Y_j|\mathcal{F}_1\cap\mathcal{F}_2)=0$ we have

$$E(Y_1+Y_2)^2\geq EY_1^2+EY_2^2-2\tilde{\rho}(EY_1^2EY_2^2)^{1/2}\geq(1-\tilde{\rho})(EY_1^2+EY_2^2).$$

Therefore the linear operator

$$T:\mathcal{L}_2^0\to\mathcal{L}_2^0(\mathcal{F}_1)\times\mathcal{L}_2^0(\mathcal{F}_2)$$

given by $T(Y)=(\mathcal{E}_1(Y),\mathcal{E}_2(Y))$ is onto and the norm of its left inverse is bounded by $(1-\tilde{\rho})^{-1/2}$. This proves the bound $EZ^2\leq(EX_1^2+EX_2^2)/(1-\tilde{\rho})$.

Because of the explicit formula for Z, it is clear that 3⇒2; implication 2⇒1 holds by general principles (see Remark 2.1). The equivalence 1⇔3 is in [5]. □

Remark 2.1 says that if for a given sequence $\{X_j\}$ equation (1) is solvable in the additive class (4), then there exists also a minimal solution (3). The following shows that for finite families of σ-fields the solvability of both problems is actually equivalent, at least when constraints $EX_i=EX_j$ are the only ones to be used.

Lemma 3.2 *Fix $1 < p < \infty$ and suppose $d < \infty$. The following conditions are equivalent*

(i) Equation (1) has an additive (4) solution Z for all $X_j \in L_p^0(\mathcal{F}_j)$;

(ii) Equation (1) has a minimal (3) solution Z for all $X_j \in L_p^0(\mathcal{F}_j)$;

(iii) There exists $\delta = \delta(q) > 0$ such that for all $X_j \in L_q^0(\mathcal{F}_j)$

$$E|\sum_j X_j|^q \geq \delta^q \sum_j E|X_j|^q, \tag{12}$$

where $1/p + 1/q = 1$.

Moreover, if inequality (12) holds, then there exists an additive solution Z to (1) with

$$E|Z|^p \leq \frac{1}{\delta^p} \sum_j E|X_j|^p.$$

Remark 3.1 *If in addition $L_q^0(\mathcal{F}_j)$ are linearly independent, then the following equivalent condition can be added:*

(iv) $L_q^0(\mathcal{F}_1) \oplus \ldots \oplus L_q^0(\mathcal{F}_d)$ is a closed subspace of $L_q(\mathcal{F}_I)$.

Proof of Lemma 3.2. (iii)⇒(i) Consider the linear bounded operator $T : L_p \to \ell_p(L_p^0(\mathcal{F}_j))$ defined by $Z \mapsto (E(Z|\mathcal{F}_j) : j = 1, \ldots, d)$. The conjugate operator $T^\star : \ell_q \to L_q$ is given by $(X_j) \mapsto \sum_{j=1}^d X_j$.

Coercivity criterion for T being onto, see [14, Theorem 4.15], is

$$\|T^\star(X_j)\|_{L_q} \geq \delta \|(X_j)\|_{\ell_q},$$

which is (12). Therefore (i) follows.

The left-inverse operator has $\ell_p \to L_p$ operator norm $\|T^{-1}\| \leq 1/\delta$, which gives the estimate of the norm of Z as claimed.

(i)⇒(ii) If there is additive solution, then X_j are consistent and Remark 2.1 implies that there is a solution with the minimal L_p-norm.

(ii)⇒(iii) If for every sequence X_j there exist Z such that (1) holds, then the linear operator $T : L_p^0 \to L_p^0(\mathcal{F}_1) \times \ldots \times L_p^0(\mathcal{F}_d)$ given by $Z \mapsto (\mathcal{E}_j(Z))$ is onto. Therefore the conjugate operator satisfies

$$\|T^\star(X_1, \ldots, X_d)\|_q \geq \delta \|(X_1, \ldots, X_d)\|_{\ell_q(L_q(\mathcal{F}_j))}$$

and inequality (12) follows, see [14, Theorem 4.15].

□

Proof of Remark 3.1. (iv)⇒(iii) If $L_q^0(\mathcal{F}_1) \oplus \ldots \oplus L_q^0(\mathcal{F}_d)$ is a closed subspace of $L_q(\mathcal{F}_I)$ then (12) holds. Indeed, by linear independence, the linear operator $(X_1 + \ldots + X_d) \mapsto (X_1, \ldots, X_d)$ is an injection of the Banach space $L_q^0(\mathcal{F}_1) \oplus \ldots \oplus L_q^0(\mathcal{F}_d)$ with ℓ_q norm into $L_q^0(\mathcal{F}_1) \times \ldots \times L_q^0(\mathcal{F}_d)$. Since the range is closed, the open mapping theorem ([14, Theorem 2.11]) implies (12).

(iii)⇒(iv) is trivial. □

Proof of Theorem 2.1. From the proof of Bryc & Smoleński [7, (7)] we have the left hand side of the inequality

$$cE|\sum_{j=1}^d \epsilon_j X_j|^q \leq E|\sum_{j=1}^d X_j|^q \leq CE|\sum_{j=1}^d \epsilon_j X_j|^q.$$

The right hand side is stated as [7, (7)].

By the Khinchin inequality this implies (11). (Note that a more careful analysis gives explicit estimates for the constants involved.)

For $q = 2$ the above is replaced by

$$\frac{1-R_\ell}{1+R_\ell}\sum_{j=1}^{d} EX_j^2 \leq E|\sum_{j=1}^{d} X_j|^2 \leq \frac{1+R_\ell}{1-R_\ell}\sum_{j=1}^{d} EX_j^2,$$

which is stated in [2, Lemma 1].

Existence of the solution follows now from functional analysis. Consider the bounded linear (c.f. (11)) operator $T : L_p^0 \to L_p^0(\ell_2(\mathcal{F}_j))$ defined by $Z \mapsto (E(Z|\mathcal{F}_j) : j = 1, 2, \ldots)$. The conjugate operator $T^\star : L_q^0(\ell_2) \to L_q^0$ is given by $(X_j) \mapsto \sum_{j=1}^{\infty} X_j$.

Coercivity criterion for T being onto, see [14, Theorem 4.15], is

$$\|T^\star(X_j)\|_{L_q} \geq \delta\|(X_j)\|_{L_q(\ell_2)},$$

which follows from (11). Therefore the existence of a solution to (1) follows and the minimal solution exists by Remark 2.1.

For $p = 2$ inequalities (11) show that $L_2^0(\ell_2) = \ell_2\left(L_2^0(\mathcal{F}_j)\right) \ni (X_j)$ generates the L_2 convergent series $\sum_j X_j$. Denote by H the set of random variables represented by such series. By (11) H is closed and since the orthogonal projection onto H shrinks the norm, the minimal solution to (1) has to be in H, thus it is additive (4). The left-inverse operator has $\ell_2 \to L_2$ operator norm $\|T^{-1}\| \leq (\frac{1-R_\ell}{1+R_\ell})^{1/2}$, which gives the estimate for the norm of Z as claimed.

The uniqueness follows from (11) by Lemma 3.1.□

Proof of Theorem 2.2. Use Theorem 2.11 to produce recurrently
$\mathcal{F}_{1,2}$-measurable Z^1 such that $\mathcal{E}_1(Z^1) = X_1, \mathcal{E}_2(Z^1) = X_2$;
$\mathcal{F}_{1,2,3}$-measurable Z^2 such that $\mathcal{E}_{1,2}(Z^2) = Z^1, \mathcal{E}_3(Z^2) = X_3$;
$\vdots$
$\mathcal{F}_{1,\ldots,d}$-measurable Z^d such that $\mathcal{E}_{1,\ldots,d-1}(Z^d) = Z^{d-1}, \mathcal{E}_d(Z^d) = X_d$.
This shows that for all $d < \infty$ there is a solution to (1), and hence a minimal solution exists. Therefore, by Lemma 3.2 there exists an additive solution (4).

The fact that inequality (12) holds with $q = 2$ and $\delta = (1-\rho_\star)^{d/2}$ follows recurrently from

$$E(\sum_{j=1}^{k} X_j + X_{k+1})^2 \geq (1-\rho_\star)(E(\sum_{j=1}^{k} X_j)^2 + EX_{k+1}^2).$$

For $d < \infty$ inequality (12) implies (11), which by Lemma 3.1 implies uniqueness. □

Proof of Theorem 2.3.

To verify that the series (7) converges, notice that for $j \neq i_k$

$$\|\mathcal{E}_{i_1}\ldots\mathcal{E}_{i_k}\|_{L_2^0(\mathcal{F}_j)\to L_2^0} \leq r^k.$$

Therefore

$$\|\sum_k(-1)^k \sum_{i_1\in I}\sum_{i_2\in I\setminus i_1}\ldots\sum_{i_k\in I\setminus i_{k-1}} \mathcal{E}_{i_1}\ldots\mathcal{E}_{i_k}\sum_{j\in I\setminus i_k}(X_j - EX_j)\|_2$$

$$\leq 2d \sum_k (d-1)^k r^k \max_j \|X_j\|_2.$$

Clearly, (4) holds true. We check now that Z defined by (7) satisfies (1). To this end, without loss of generality we assume $EX_j = 0$ and we verify (1) for $j = 1$ only. Splitting the sum (7) we get

$$\mathcal{E}_1(Z) = \sum_{k=0}^{\infty} (-1)^k \sum_{i_2 \in I \setminus 1} \cdots \sum_{i_k \in I \setminus i_{k-1}} \mathcal{E}_1 \mathcal{E}_{i_2} \ldots \mathcal{E}_{i_k} \sum_{j \in I \setminus i_k} (X_j - EX_j)$$

$$+ \sum_{k=0}^{\infty} (-1)^k \sum_{i_1 \in I \setminus 1} \sum_{i_2 \in I \setminus i_1} \cdots \sum_{i_k \in I \setminus i_{k-1}} \mathcal{E}_1 \mathcal{E}_{i_1} \ldots \mathcal{E}_{i_k} \sum_{j \in I \setminus i_k} (X_j - EX_j).$$

The 0-th term of the first series is X_1 and the k-th term of the first series cancels the $(k-1)$ term of the second series. Therefore $\mathcal{E}_1(Z) = X_1$.

To prove the uniqueness, it suffices to notice that $r < \frac{1}{d-1}$ implies linear independence. Alternatively, suppose that both $Z = Z_1 + \ldots + Z_d$ and $Z' = Z_1' + \ldots + Z_d'$ have the same conditional moments $\mathcal{E}_1$. Then $\|Z_1 - Z_1'\|_2 = \|\mathcal{E}_1(Z_2 + \ldots + Z_d) - \mathcal{E}_1(Z_2' + \ldots + Z_d')\| \leq r \sum_{j=2}^d \|Z_j - Z_j'\|_2$, and the similar estimate holds for all other components. Therefore $\sum_{j=1}^d \|Z_j - Z_j'\|_2 \leq r(d-1) \sum_{j=1}^d \|Z_j - Z_j'\|_2$. Since $r < 1/(d-1)$, this implies that the sum vanishes, proving uniqueness.

To prove the variance estimate notice that $r < 1/(d-1)$ implies (12) with $p = 2$ and $\delta^2 = 1 - r(d-1)$. Indeed,

$$E|\sum_{j=1}^d X_j|^2 \geq \sum_{j=1}^d EX_j^2 - r \sum_{j \neq k} (EX_j^2 EX_k^2)^{1/2}.$$

The estimate now follows from the elementary inequality

$$\frac{1}{d-1} \sum_{j=1}^d \sum_{k=1, k\neq j}^d a_k a_j \leq \frac{1}{d-1} \sum_{j=1}^d \sum_{k=1, k\neq j}^d \frac{1}{2}(a_k^2 + a_j^2) = \sum_{j=1}^d a_j^2,$$

valid for arbitrary numbers $a_1, \ldots, a_d$. $\square$

Proof of Theorem 2.5. Take $U \in L_2(\mathcal{F}_S), V \in L_2(\mathcal{F}_T)$ with disjoint $S, T \subset I$ and such that $EU = EV = 0, EU^2 = EV^2 = 1, EUV = \rho$. Then

$$E(U-V)^2 = 2 - 2\rho \geq 2\frac{b}{B^2}.$$

Indeed, we have

$$E(U-V)^2 = \int_{\mathbb{R}^d} (U(\mathbf{x}) - V(\mathbf{x}))^2 f(\mathbf{x}) d\mu(\mathbf{x}) \geq b \int_{\mathbb{R}^d} (U(\mathbf{x}) - V(\mathbf{x}))^2 d\mu(\mathbf{x})$$

$$= b \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (U(\mathbf{x}) - V(\mathbf{y}))^2 d\mu(\mathbf{y}) d\mu(\mathbf{x})$$

$$\geq \frac{b}{B^2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (U(\mathbf{x}) - V(\mathbf{y}))^2 f(\mathbf{y}) d\mu(\mathbf{y}) f(\mathbf{x}) d\mu(\mathbf{x}) = \frac{2b}{B^2}.$$

$\square$

Since the above analysis can also be carried through for $E(U+V)^2$, we get the following.

Corollary 3.3 (c.f.[15] Lemma 1) *Under the assumption of Theorem 2.5, for $V_j \in L_2^0(\mathcal{F}_j)$ we have*

$$E|V_1 + \ldots + V_d|^2 \geq \frac{b}{2B^2 - b}(EV_1^2 + \ldots + EV_d^2)$$

Lemma 3.4 *Let f be the density of the absolute continuous part of the distribution of $\xi_1, \ldots, \xi_d$ ($d < \infty$) with respect to the product of marginals $\mu = \mu_1 \otimes \ldots \otimes \mu_d$. If f is strictly positive, i.e., $\mu(\{(x_1, \ldots, x_d) : f(x_1, \ldots, x_d) = 0\}) = 0$, then vector spaces $L_2^0(\mathcal{F}_j)$ are linearly independent.*

Proof of Lemma 3.4. Suppose $X_1 = X_1(\xi_1), \ldots, X_d = X_d(\xi_d) \in L_2^0$ are non-zero. Denote by μ the product of marginal measures on $\mathbb{R}^d$ and let

$$A_\epsilon = \{(x_1, \ldots, x_d) : f(x_1, \ldots, x_d) > \epsilon\}.$$

Choose $\epsilon > 0$ such that

$$\int_{A_\epsilon^c} |\sum X_j|^2 d\mu < \frac{1}{2} \sum EX_j^2.$$

Then

$$E|\sum X_j|^2 \geq \int_{A_\epsilon} |\sum X_j(x_j)|^2 f(x_1, \ldots, x_d) d\mu$$

$$\geq \epsilon \int_{A_\epsilon} |\sum X_j|^2 d\mu \geq \frac{\epsilon}{2} \sum EX_j^2 > 0.$$

This proves linear independence of X_j.
□

Proof of Proposition 2.8. This follows the proof of Proposition 3.4 with $\epsilon = b$. Namely,

$$E|\sum X_j|^q \geq b \int_{\mathbb{R}^d} |\sum X_j|^q d\mu \geq cE(\sum X_j^2)^{q/2}.$$

The last inequality holds by the Marcinkiewicz-Zygmund inequality, because under μ random variables X_j are independent and centered. □

Lemma 3.5 *If $d < \infty$, vector spaces $L_2^0(\mathcal{F}_j)$ are linearly independent, and for all $1 \leq j \leq d, k \neq j$, the operators $\mathcal{E}_j\mathcal{E}_k : L_2^0(\mathcal{F}_k) \to L_2^0(\mathcal{F}_j)$ are compact, then $R_\ell < 1$; hence inequality (12) holds for $q = 2$.*

Proof of Lemma 3.5. The proof is similar to the proof of Proposition 2.9 with $T = P_S P_Q$, where S, Q are disjoint and P_Q denotes the orthogonal projection onto the L_2-closure of $\bigoplus_{j \in Q} L_2^0(\mathcal{F}_j)$; operator T is compact, compare [4, Proposition 5.3]. Details are omitted. □

4 Example

The following simple example illustrates sharpness of some moment estimates.

Example 4.1 *Let $d < \infty$. Suppose $X_1, \dots, X_d$ are square-integrable, centered and have linear regressions, i.e., there are constants $a_{i,j}$ such that $E(X_i|X_j) = a_{i,j} X_j$ for all i, j (for example, this holds true for $(X_1, \dots, X_d)$ with elliptically contoured distributions, or when all X_j are two-valued). Let $C = [C_{i,j}]$ be their covariance matrix. Clearly, if either $R_\ell < 1$ or $r < 1/(d-1)$, then C is non degenerate.*

Explicit solutions illustrating Theorems 2.11, 2.1, and 2.3 are then possible:
It is easy to check that $Z = \sum_{j=1}^{d} \theta_j X_j$, where $\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_d \end{bmatrix} = C^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, satisfies $E(Z|X_j) = X_j$ for all j, and Z has additive form. In the special case when $corr(X_i, X_j) = \rho$ does not depend on $i \neq j$,

$$Z = \frac{1}{1 + (d-1)\rho}(X_1 + \dots + X_d).$$

It is easy to see directly from

$$E(X_1 + \dots + X_d)^2 = d + d(d-1)\rho \geq \frac{1-R}{1+R}$$

that

$$(d-1)\rho \geq -\frac{2R}{1+R} > -1,$$

provided $R_\ell < 1$. Therefore Z is well defined.

In particular, for $d = 2$ we have $Z = \frac{1}{1+\rho}(X_1 + X_2)$, which points out the sharpness of estimates for EZ^2 in Theorem 2.11 when $\rho < 0$.

Acknowledgement: The subject of this paper arose in the conversation with M. Pourahmadi, who also provided relevant references, helpful comments and additional materials. The author also thanks Q. Shata for helpful comments. Part of the research was done at the Institute for Mathematics and Applications of the University of Minnesota and at the Center for Stochastic Processes in Chapel Hill.

References

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

[2] R. Bradley, On the spectral density and asymptotic normality of weakly dependent random fields. Journ. Theoret. Probab. 5 (1992), 355–373.

[3] R. Bradley, Every "lower psi-mixing" Markov chain is "interlaced rho-mixing", preprint 1996.

[4] L. Breiman & J. H. Friedman, Estimating Optimal Transformations for Multiple Regression and Correlation. JASA 80 (1985), 580–598.

[5] W. Bryc, Conditional expectation with respect to dependent σ-fields. In: *Proceedings of the Seventh Conference on Probability Theory (Brasov)*, Ed. M. Iosifescu, Bucuresti, 1984, 409–411.

[6] W. Bryc & S. Kwapień, On the conditional expectation with respect to a sequence of independent σ-fields. Z. Wahrsch. Verw. Gebiete 46 (1979), 221–225.

[7] W. Bryc & W. Smoleński, Moment conditions for almost sure convergence of weakly correlated random variables. Proc. Amer. Math. Soc. 119 (1993), 629–635.

[8] A. Buja & T. Hastie & R. Tibshirani, Linear smoothers and additive models. Ann. Statist. 17 (1989), 453–510.

[9] C. Groetsch, *Inverse Problems in the Mathematical Sciences.* Vieweg, Braunschweig (Germany), 1993.

[10] H. G. Kellerer, Masstheoretische Marginalproblem. Math. Ann. 153 (1964), 168–198.

[11] J. Komlós, A generalization of a problem of Steinhaus. Acta Math. Acad. Sci. Hung. 18 (1967), 217–229.

[12] B. F. Logan & L. A. Shepp, Optimal reconstruction of a function from its projections. Duke Math. Journ. 42 (1975), 645–659.

[13] G. C. Rota, An "alternierende Verfahren" for general positive operators. Bull. Amer. Math. Soc. 68 (1962), 95–102.

[14] W. Rudin, *Functional Analysis.* McGraw, New York 1973.

[15] C. J. Stone, Additive regression and other nonparametric models. Ann. Statist. 13 (1985), 689–705.

[16] V. Strassen, The existence of probability measures with given marginals. Ann. Math. Statist. 36 (1965), 423–439.