J. Aust. Math. Soc. 74 (2003), 35–42

ON THE GEOMETRY OF $L^p(\mu)$ WITH APPLICATIONS TO INFINITE VARIANCE PROCESSES

R. CHENG, A. G. MIAMEE and M. POURAHMADI

(Received 20 March 2001; revised 25 October 2001)

Communicated by V. Stefanov

Abstract

Some geometric properties of L^p spaces are studied which shed light on the prediction of infinite variance processes. In particular, a Pythagorean theorem for L^p is derived. Improved growth rates for the moving average parameters are obtained.

2000 *Mathematics subject classification*: primary 60G25; secondary 46B25.

1. Introduction

A discrete-time process $\{X_t\}$ with $X_t \epsilon L^p(\Omega, \mathscr{F}, P)$ is said to be *p-stationary* if for all integers $n \geq 1, t_1, \dots, t_n, h$ and scalars $c_1, \dots, c_n$,

$$E\left|\sum_{k=1}^{n} c_k X_{t_k+h}\right|^p = E\left|\sum_{k=1}^{n} c_k X_{t_k}\right|^p .$$

Thus, 2-stationary processes are, indeed, the familiar and well-developed second-order stationary processes. However, when $1 < p < 2$, p-stationary processes do not even have a well-defined notion of covariance or spectrum, so that neither the spectral-domain nor the time-domain techniques are as effective as they have been for 2-stationary processes [1, 2, 5, 6]. The *innovation process* $\{\epsilon_t\}$ of $\{X_t\}$ is defined by $\epsilon_t = X_t - P_{H_{t-1}} X_t$, where $P_{H_{t-1}} X_t$ stands for the metric projection of X_t onto $H_{t-1} = \overline{\text{sp}}\{X_{t-1}, X_{t-2}, \dots\}$ in the norm of $L^p(\Omega, \mathscr{F}, P)$.

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

It is known, [5], that any nondeterministic p-stationary process can be written as

$$X_t = \epsilon_t + \sum_{k=1}^{n} a_k X_{t-k} + E_{t,n} = \epsilon_t + \sum_{k=1}^{n} b_k \epsilon_{t-k} + V_{t,n}, \tag{1.1}$$

for any $n \geq 1$, where $\{a_k\}$ and $\{b_k\}$ are unique sequences of scalars called the autoregressive (AR) and moving average (MA) parameters of $\{X_t\}$, and $V_{t,n}$, $E_{t,n} \in H_{t-n-1}$. The second representation in (1.1) is called a finite Wold decomposition of $\{X_t\}$. If the success of characterization of regularity of 2-stationary processes is any clue, then the norm-convergence of $\sum_{k=1}^{n} b_k \epsilon_{t-k}$ as $n \to \infty$, should play a central role in the study of regularity of p-stationary processes. This question of convergence is, in turn, related to the growth of the MA coefficients $\{b_k\}$; it is known, [5], that $b_k = O(2^k)$. An improved bound is obtained in the present work for the p-stationary case, using geometric properties specific to $L^p(\mu)$ spaces. Among these is a Pythagorean theorem for L^p, derived using elementary means.

2. The geometry of $L^p(\mu)$

The notion of Birkhoff orthogonality in a normed linear space is central to this work. Let x and y be elements of a Banach space $\mathscr{X}$. We write $x \perp_{\mathscr{X}} y$ if $\|x + \alpha y\| \geq \|x\|$ for all scalars α. Note that the relation $\perp_{\mathscr{X}}$ is generally not symmetric or linear. If $\mathscr{X} = L^p(\mu)$, we will write $x \perp_p y$ for $x \perp_{\mathscr{X}} y$.

A Banach space $\mathscr{X}$ is said to be *uniformly convex* if for any $\epsilon \in (0, 2]$ there exists a $\delta_\epsilon > 0$ such that the conditions $\|x\| \leq 1$, $\|y\| \leq 1$, and $\|x - y\| \geq \epsilon$ together imply that $\|x + y\|/2 \leq 1 - \delta_\epsilon$. Here is a useful criterion for uniform convexity.

PROPOSITION 2.1. *A Banach space $\mathscr{X}$ is uniformly convex if and only if the conditions $\|x_n\| \leq 1$, $\|y_n\| \leq 1$ and $\lim_{n\to\infty} \|(x_n + y_n)/2\| = 1$ together imply that $\lim_{n\to\infty} \|x_n - y_n\| = 0$.*

It is known that for $1 < p < \infty$, the spaces $L^p(\mu)$ are uniformly convex. For these results and additional information on Banach spaces see [3, page 353].

Suppose that M is a closed subspace of a Banach space $\mathscr{X}$. For $x \in \mathscr{X}$ consider the problem of minimizing $\|x - y\|$ over $y \in M$. When $\mathscr{X}$ is uniformly convex, then the extremal vector y is uniquely determined by x and M. In that situation the metric projection mapping $y = P_M x$ is characterized by

$$P_M x \in M \quad \text{and} \quad x - P_M x \perp_{\mathscr{X}} M. \tag{2.1}$$

If P_M is the metric projection mapping, then

$$\|P_M x\| \leq 2\|x\| \tag{2.2}$$

for all $x \in \mathscr{X}$. This is because

$$\|P_M x\| = \|P_M x - x + x\| \leq \|x - P_M x\| + \|x\| \leq 2\|x\|.$$

We shall see that this bound, derived from general norm properties, can be sharpened when $\mathscr{X} = L^p(\mu)$. Furthermore, from (1.1) and repeated application of (2.2) it follows that

$$|b_m| \leq 2^m \frac{\|X_0\|}{\|\epsilon_0\|} \tag{2.3}$$

for all m. This bound will also be sharpened when using properties special to $L^p(\mu)$ spaces.

Uniform convexity interacts with metric projection in the following way.

LEMMA 2.2. *Suppose that the Banach space $\mathscr{X}$ is uniformly convex, M is a closed subspace of $\mathscr{X}$, and $x \perp_{\mathscr{X}} M$. If $y_m \in M$, and* $\lim \|x + y_m\| = \|x\|$, *then* $\lim \|y_m\| = 0$.

PROOF. The assertion is trivial if $x = 0$. Otherwise, put $X_m = x/\|x + y_m\|$ and $Y_m = (x + y_m)/\|x + y_m\|$. Note that $\|X_m\| \leq 1$, since $x \perp_{\mathscr{X}} y_m$, and $\|Y_m\| = 1$. Furthermore,

$$\frac{\|x\|}{\|x + y_m\|} \leq \frac{\|x + y_m/2\|}{\|x + y_m\|} = \|(X_m + Y_m)/2\| \leq 1.$$

By assumption, $\lim \|x\|/\|x + y_m\| = 1$, which then forces $\lim \|(X_m + Y_m)/2\| = 1$. Now Proposition 2.1 gives

$$\begin{aligned}\lim \|y_m\| &= \|x\| \lim(\|y_m\|/\|x\|) \\ &= \|x\| \lim(\|y_m\|/\|x + y_m\|) = \|x\| \lim(\|X_m - Y_m\|) = 0. \qquad \square\end{aligned}$$

It is known that the metric projection onto a subspace is norm continuous in a strictly convex, locally compact Banach space [3, page 344]. Here is the result for a uniformly convex space.

PROPOSITION 2.3. *Let M be a closed subspace of a uniformly convex Banach space $\mathscr{X}$. If $x \in \mathscr{X}$, $x_m \in \mathscr{X}$, and* $\lim \|x_m - x\| = 0$, *then* $\lim \|P_M x_m - P_M x\| = 0$.

PROOF. Observe that

$$\begin{aligned}\|x - P_M x\| &\leq \|x - P_M x_m\| \leq \|x - x_m\| + \|x_m - P_M x_m\| \\ &\leq \|x - x_m\| + \|x_m - P_M x\| \leq \|x - x_m\| + \|x_m - x\| + \|x - P_M x\| \\ &= 2\|x - x_m\| + \|x - P_M x\|.\end{aligned}$$

It follows that $\lim \|x - P_M x_m\| = \|x - P_M x\|$. Applying Lemma 2.2, and using the orthogonality condition $(x - P_M x) \perp_{\mathscr{X}} M$, we get $\lim \|P_M x_m - P_M x\| = 0$. $\square$

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

The following inequalities constitute a parallelogram law for $L^p(\mu)$.

PROPOSITION 2.4. *If $2 \le p < \infty$, then for any f and g in $L^p(\mu)$*

$$2(\|f\|^p + \|g\|^p) \le \|(f+g)\|^p + \|(f-g)\|^p \tag{2.4}$$

$$\le 2^{p-1}(\|f\|^p + \|g\|^p). \tag{2.5}$$

If $1 < p \le 2$, then for any f and g in $L^p(\mu)$

$$2^{p-1}(\|f\|^p + \|g\|^p) \le \|(f+g)\|^p + \|(f-g)\|^p \tag{2.6}$$

$$\le 2(\|f\|^p + \|g\|^p). \tag{2.7}$$

Equality holds in (2.4) *and* (2.7), *if and only if $fg = 0$ a.e.; equality holds in* (2.5) *and* (2.6) *if and only if $f = \pm g$ a.e.*

PROOF. For $p \ge 2$, see [3, page 55ff]. For $1 < p < 2$, consider the parameter $r = 4/p$, and apply the previous result. □

Note that as p tends to 2 in either direction the Hilbert space case results; the inequalities are sharp in this limited sense. From the parallelogram law, we get a Pythagorean theorem for $L^p(\mu)$. Again, there are two cases.

PROPOSITION 2.5. *Suppose that $X, Y \in L^p(\mu)$, $X \perp_p Y$, and $\lambda = (2^{p-1} - 1)^{-1/p}$. Then,*

$$\|X\|^p + \lambda^p \|Y\|^p \le \|X + Y\|^p, \quad \textit{if } 2 \le p < \infty, \tag{2.8}$$

$$\|X + Y\|^p \le \|X\|^p + \lambda^p \|Y\|^p, \quad \textit{if } 1 < p \le 2. \tag{2.9}$$

PROOF. We apply (2.4) in the form

$$\left\|\tfrac{1}{2}(f+g)\right\|^p + \left\|\tfrac{1}{2}(f-g)\right\|^p \le \tfrac{1}{2}(\|f\|^p + \|g\|^p). \tag{2.10}$$

Now taking $f = X$ and $g = X + Y$ in (2.10) we get

$$\|X + \tfrac{1}{2}Y\|^p + \|\tfrac{1}{2}Y\|^p \le \tfrac{1}{2}\|X\|^p + \tfrac{1}{2}\|X + Y\|^p.$$

Apply (2.10) repeatedly, taking $f = X$ and $g = X + (1/2^n)Y$, $n = 1, 2, 3, \ldots, N$, will result in

$$2^N\|X + (1/2^{N+1})Y\|^p + 2^N\|(1/2^{N+1})Y\|^p + \cdots + 2^1\|(1/2^{1+1})Y\|^p + 2^0\|(1/2^{0+1})Y\|^p$$
$$\le (2^{N-1} + \cdots + 2^1 + 2^0 + 2^{-1})\|X\|^p + \|X + Y\|^p/2.$$

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

Simplifying, taking N to infinity, and using $\|X + (1/2^N)Y\| \geq \|X\|$, we finally get

$$\|X\|^p + \frac{1}{2^{p-1} - 1}\|Y\|^p \leq \|X + Y\|^p. \tag{2.11}$$

Note that the condition $X \perp_p Y$ implies that the quantity $\|X + \alpha Y\|$ is critical when $\alpha = 0$. It follows that $\lim_{N\to\infty} 2^N(\|X + (1/2^N)Y\|^p - \|X\|^p) = 0$, and the estimate leading to (2.8) is asymptotically sharp.

In the case $1 < p \leq 2$, we turn to (2.7), with $f = X$ and $g = X + Y$. This yields

$$\tfrac{1}{2}\|X\|^p + \tfrac{1}{2}\|X + Y\|^p \leq \|X + \tfrac{1}{2}Y\|^p + \|\tfrac{1}{2}Y\|^p. \tag{2.12}$$

Repeating this argument with $f = X$ and $g = X + (1/2^n)Y$, $n = 1, 2, 3, \ldots, N$ results in

$$(2^N - 1)\|X\|^p + \|X + Y\|^p \leq 2^N\|X + (1/2^N)Y\|^p + \frac{1}{2^{p-1} - 1}\|Y\|^p.$$

Rearranging, we find that

$$\|X + Y\|^p \leq \|X\|^p + \frac{1}{2^{p-1} - 1}\|Y\|^p + 2^N(\|X + (1/2^N)Y\|^p - \|X\|^p).$$

As N tends to infinity, the last term vanishes, because $X \perp_p Y$. □

Note that (2.9) can be sharper than the triangle inequality. There is a pleasing symmetry in Proposition 2.5; also, it yields the familiar Hilbert space case as p tends to 2 in either direction.

The constant $\lambda = (2^{p-1} - 1)^{-1/p}$ appearing in (2.8) and (2.9) might not be optimal, however, since the estimates in the proof are generally not sharp. One might wonder whether the value $\lambda = 1$ is always possible. The following example shows that it is not.

Let $\mathscr{X} = l^3(\{1, 2\})$, and consider $f = (1/4, 1)$ and $g = (-1, 1/16)$ in $\mathscr{X}$. Then $f \perp_3 g$, and $\|f\|^3 = 65/64$, $\|g\|^3 = 4097/4096$, $\|f + g\|^3 = 6641/4096$. In order that $\|f\|^3 + \lambda^3\|g\|^3 \leq \|f + g\|^3$, it is necessary that $\lambda^3 \leq 2481/4097$.

The Pythagorean inequalities give rise to improved bounds on the coefficient growth in the finite Wold decomposition (1.1). As before, we write $\lambda = (2^{p-1} - 1)^{-1/p}$.

3. Application

The geometric results of Section 2 are applied to prediction of a L^p stationary process $\{X_t\}$. We obtain norm convergence of the finite prediction, improved bounds on the MA coefficients and improved bounds on the norm of the metric projection.

Let $\hat{X}$ be the projection of X_0 based on the infinite past $\{\ldots, X_{-3}, X_{-2}, X_{-1}\}$, and $\hat{X}(m)$ be the projection of X_0 based on the finite past $\{X_{-m}, \ldots, X_{-3}, X_{-2}, X_{-1}\}$.

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

THEOREM 3.1. *If* $\{X_t\}_{t=-\infty}^{\infty}$ *is a* p*-stationary process, then the finite predictors* $\hat{X}(m)$ *of* X_0 *converge in norm to its infinite predictor* $\hat{X}$.

PROOF. Let $\{Y_m\}_{m=-\infty}^{\infty}$ be a sequence such that $Y_m \in \operatorname{sp}\{X_{-m}, \dots, X_{-3}, X_{-2}, X_{-1}\}$ and $\lim \|Y_m - \hat{X}\| = 0$; such a sequence exists since $\hat{X} \in \overline{\operatorname{sp}}\{\dots, X_{-3}, X_{-2}, X_{-1}\}$. With the above definitions we have

$$\|X_0 - \hat{X}\| \le \|X_0 - \hat{X}(m)\| \le \|X_0 - Y_m\| \le \|X_0 - \hat{X}\| + \|\hat{X} - Y_m\|.$$

From this we see that $\lim \|X_0 - \hat{X}(m)\| = \|X_0 - \hat{X}\|$. Applying Lemma 2.2, we get $\lim \|\hat{X}(m) - \hat{X}\| = 0$. □

THEOREM 3.2. *Suppose that* $\{X_t\}_{t=-\infty}^{\infty}$ *is a* p*-stationary process with nontrivial innovation process* $\{\epsilon_t\}_{t=-\infty}^{\infty}$, *and finite Wold decomposition* (1.1). *If* $2 \le p < \infty$, *then* $\|(1, \lambda b_1, \lambda^2 b_2, \dots)\|_{l^p} \le \|X_0\|/\|\epsilon_0\|$.

PROOF. By applying (2.8) repeatedly to the finite Wold decomposition (1.1), we get the bound

$$\|\epsilon_0\|^p + |\lambda b_1|^p \|\epsilon_1\|^p + \dots + |\lambda^N b_N|^p \|\epsilon_N\|^p + \lambda^N \|V_{0,N}\|^p \le \|X_0\|^p$$

for all N. Now drop the nonnegative term $\lambda^N \|V_{0,N}\|^p$, and let N increase without bound. □

Observe that this improves on the bound (2.2). The case $1 < p \le 2$ is more delicate, since the estimate (2.9) is not similarly useful. However, the following can be said.

PROPOSITION 3.3. *Let* $1 < p \le 2$, *and suppose that* $X \perp_p Y$. *If* κ *is a constant satisfying* $0 \le \kappa \le (2^{p-1} - 1)$, *then for any positive integer* N *satisfying*

$$N \le \frac{1}{p-1} \log_2 \left[\frac{\kappa(2^{p-1} - 1) - 1}{2^{p-1} - 2} \right],$$

we have $\kappa\|X\|^p + (1 - 2^{-N})\|Y\|^p \le \|X + Y\|^p$.

PROOF. We start with (2.7), using $f = X$ and $g = X + Y$ to get

$$2^{p-1}\|X + \tfrac{1}{2}Y\|^p + 2\|\tfrac{1}{2}Y\|^p \le \|X + Y\|^p + \|X\|^p.$$

Repeat this estimate using $f = X$ and $g = X + (1/2^n)Y$, $1 \le n \le N$, with the result

$$2^{(p-1)N}\|X + (1/2^N)Y\|^p + \left(\frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^N}\right)\|Y\|^p$$
$$\le \|X + Y\|^p + (1 + 2^{p-1} + \dots + 2^{(p-1)(N-1)}\|X\|^p.$$

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

Rearranging, and using $X \perp_p Y$, we deduce that

$$\left[2^{(p-1)N} - \frac{2^{(p-1)N} - 1}{2^{p-1} - 1}\right] \|X\|^p + (1 - 2^{-N})\|Y\|^p \le \|X + Y\|^p.$$

The constant enclosed in the square brackets is at most the value $(2^{p-1} - 1)$. For κ satisfying $0 \le \kappa \le (2^{p-1} - 1)$, we have

$$\kappa \le \left[2^{(p-1)N} - \frac{2^{(p-1)N}-1}{2^{p-1}-1}\right], \quad \text{whenever} \quad N \le \frac{1}{p-1}\log_2\left[\frac{\kappa(2^{p-1}-1)-1}{2^{p-1}-2}\right]. \qquad \square$$

The values $\kappa = (2^{p-1} - 1)$ and $N = 1$ can always be used, corresponding to the crude bound $(2^{p-1} - 1)\|X\|^p + \frac{1}{2}\|Y\|^p \le \|X + Y\|^p$. The coefficient growth estimate that results from Proposition 3.3 is the following.

COROLLARY 3.4. *Suppose that $\{X_t\}_{t=-\infty}^{\infty}$ is a p-stationary process with nontrivial innovation process $\{\epsilon_t\}_{t=-\infty}^{\infty}$, and finite Wold decomposition* (1.1). *If $1 < p \le 2$, then with the notation of Proposition* 3.3,

$$1 + (1 - 2^{-N})|b_1|^p + (1 - 2^{-N})^2|b_2|^p + \cdots \le \|X_0\|^p/\kappa\|\epsilon_0\|^p.$$

When p is close to 2 (greater than about 1.695), then N is greater than 1, and this is a sharper bound on the coefficient growth than (2.3).

These Pythagorean inequalities also give improved bounds on the norm of the metric projection, compared with the crude result (2.2).

COROLLARY 3.5. *Let M be a closed subspace of $L^p(\mu)$. Then*

$$\|P_M f\| \le (2^{p-1} - 1)^{1/p}\|f\|, \quad \text{if } 2 \le p < \infty.$$
$$\|P_M f\| \le (1 - 2^{-N})^{-1/p}\|f\|, \quad \text{if } 1 < p \le 2.$$

where N is any positive integer satisfying $N \le -(p-1)^{-1}\log_2(2 - 2^{p-1})$.

Again, note that when $1 < p \le 2$ we can always choose $N = 1$, which gives

$$\|P_M f\| \le 2^{1/p}\|f\|,$$

still an improvement over (2.2). Furthermore, Corollary 3.5 is sharp in the limiting sense that as p tends to 2 in either direction, we get $\|P_M f\| \le \|f\|$, which is the correct statement when $p = 2$.

Seeing Corollary 3.5, one might wonder whether $\|P_M x\|$ can actually exceed $\|x\|$. The following example shows that it can. Here, let $\mathscr{X} = l^p(\{1, 2\})$ with $p = 1.1$. Consider $f = (2, 1)$ and $g = (-2, 2^p)$. Then $f \perp_p g$. Take $x = f + g$ and $M = \text{sp}\{g\}$. Clearly, $P_M x = g$. We now compute

$$\|x\|^p = (1 + 2^p)^p \approx 3.52\ldots, \qquad \|P_M x\| = 2^p + 2^{(p^2)} \approx 4.45\ldots.$$

For more information on the norm of metric projections, see [4].

https://doi.org/10.1017/S1446788700003104 Published online by Cambridge University Press

Acknowledgement

The work of the second author was supported by US Army Research Office grant (DAAH-04-96-1-0027) and the ONR grant (N00014-89-J-1824) and third author was supported by grants from NSA (MDA 904-97-1-0013) and NSF (DMS-9707721).

References

[1] S. Cambanis, C. D. Hardin and A. Weron, 'Innovations and Wold decompositions of stable processes', *Probab. Theory Relat. Fields* **79** (1988), 1–27.

[2] R. Cheng, A. G. Miamee and M. Pourahmadi, 'Some extremal problems in $L^p(w)$', *Proc. Amer. Math. Soc.* **126** (1998), 2333–2340.

[3] G. Köthe, *Topological vector spaces I* (Springer, New York, 1969).

[4] F. Mazzone and H. Cuenya, 'A note on metric projections', *J. Approx. Theory* **81** (1995), 425–428.

[5] A. G. Miamee and M. Pourahmadi, 'Wold decomposition, prediction and parameterization of stationary processes with infinite variance', *Probab. Theory Relat. Fields* **79** (1988), 145–164.

[6] B. S. Rajput and C. Sundberg, 'On some extremal problems in H^p and the prediction of L^p-harmonizable stochastic processes', *Probab. Theory Relat. Fields* **99** (1994), 197–210.

ECI Systems and Engineering
596 Lynnhaven Parkway
Virginia Beach, VA 23452
USA
e-mail: rayc@ecihq.com

Department of Mathematics
Hampton University
Hampton, VA 23668
USA
e-mail: abolghassem.miamee@hamptonu.edu

Division of Statistics
Northern Illinois University
DeKalb, Ill. 60115
USA
e-mail: pourahm@math.niu.edu