A Burkholder-type Inequality for Discrete Martingales

James Davidson
University of Exeter

August 2020

Abstract

The upside Burkholder moment inequality is extended to $0 < p \leq 2$ for L_2-bounded martingales by adapting a result for continuous time martingales due to A. A. Novikov.

The Burkholder square-function inequalities ([1], Theorem 3.2) are of the form

$$c_p \|Q(S_n)\|_p \leq \|S_n\|_p \leq C_p \|Q(S_n)\|_p \text{ for } 1 < p < \infty$$

where $S_n = \sum_{t=1}^n X_t$ is a L_p-bounded martingale, $Q(S_n) = \left(\sum_{t=1}^n X_t^2\right)^{1/2}$ and c_p, C_p are positive constants depending only on p. These are martingale counterparts of the Marcinkiewicz-Zygmund ([5]) inequalities for independent sequences and have important applications in the proof of limit theorems for dependent processes. Alternative proofs can be found in Hall and Heyde ([3]) and Chow and Teicher ([2]).

The question of interest is whether either of these inequalities holds for $p \leq 1$. A. A. Novikov ([6]) gave the corresponding inequality for $p > 0$ in the case of stochastic integrals with respect to Brownian motion. He also gives a counter-example to show that the downside inequality does not hold with $p = 1$. His result is reported and proved in Proposition 3.26 of [4].

The purpose of this note is to adapt Novikov's result to the discrete martingale case to match the Burkholder formulation, as follows. The major difference is that here the method of proof requires L_2-boundedness.

Theorem Let S_n be an L_2-bounded martingale. For $0 < p \leq 2$ there exists $C_p > 0$ depending only on p such that

$$\|S_n\|_p \leq C_p \|Q(S_n)\|_p. \tag{1}$$

Proof For the case $p = 2$ the inequality of (1) can be set to equality with $C_2 = 1$, to reproduce the orthogonality property of the martingale. Hence, consider $p < 2$. The following argument applies except in one case, that where $X_1 = \cdots = X_{n-1} = 0$ a.s. In this case $\|S_n\|_p = \|Q(S_n)\|_p = \|X_n\|_p$ and (1) holds as an equality with $C_p = 1$.

For convenience of notation write $m = p/2 < 1$ and also let Q_n stand for $Q(S_n)$. Let $B_n = 2\sum_{t=2}^n \sum_{s=1}^{t-1} X_s X_t$ so that $S_n^2 = Q_n^2 + B_n$. Then for $\delta > 0$ and $\mu > 0$ to be chosen and $n \geq 1$ define

$$Y_n = \delta(\mu + Q_n^2) + S_n^2 = \delta\mu + (1 + \delta)Q_n^2 + B_n. \tag{2}$$

With $m < 1$ the function x^m is concave so that $2^{m-1}(x^m + y^m) \leq (x + y)^m$ for $x, y \geq 0$. Apply this formula to the first equality of (2) and take expectations to give the inequality

$$2^{m-1}\left(\delta^m \mathrm{E}(\mu + Q_n^2)^m + \mathrm{E}|S_n|^{2m}\right) \leq \mathrm{E}(Y_n^m). \tag{3}$$

Next write the telescoping sum

$$Y_n^m = Y_1^m + \sum_{t=2}^{n}(Y_t^m - Y_{t-1}^m). \tag{4}$$

Letting the second equality of (2) define Y_t, note that $Q_t^2 - Q_{t-1}^2 = X_t^2$ and so

$$Y_t - Y_{t-1} = (1+\delta)X_t^2 + \Delta B_t$$

where $\Delta B_t = 2\sum_{s=1}^{t-1} X_s X_t$. Taylor's expansions of the terms in (4) to second order for $t = 2,\dots,n$ yield

$$\begin{aligned} Y_t^m - Y_{t-1}^m = {} & mY_{t-1}^{m-1}\big((1+\delta)X_t^2 + \Delta B_t\big) \\ & + \tfrac{1}{2}m(m-1)\big(Y_{t-1} + \theta_t((1+\delta)X_t^2 + \Delta B_t)\big)^{m-2} \\ & \times ((1+\delta)X_t^2 + \Delta B_t)^2 \end{aligned} \tag{5}$$

with $\theta_t \in [0,1]$. Since $m < 1$ the second-order term in (5) is non-positive. For the case $t = 1$, noting that $Q_1^2 = S_1^2 = X_1^2$ and $Q_0 = B_0 = B_1 = 0$, write

$$Y_1^m = \big(\delta\mu + (1+\delta)X_1^2\big)^m = mY_0^{m-1}(1+\delta)X_1^2 \tag{6}$$

where $Y_0 = \delta\mu + (1+\delta)\theta_1$ and θ_1 is defined by the second equality of (6). With μ small, $\theta_1 \approx m^{-1/(m-1)}X_1^2$. Combine (4) with (6) and also substitute (5) omitting the nonpositive final terms. Taking expectations, noting $\mathrm{E}(\Delta B_t|\mathcal{F}_{t-1}) = 0$ and then applying the LIE, gives

$$\mathrm{E}(Y_n^m) \le m(1+\delta)\sum_{t=1}^{n}\mathrm{E}(Y_{t-1}^{m-1}X_t^2). \tag{7}$$

By choice of μ the functions Y_{t-1} are bounded away from zero and hence the expectations exist, for each $n \ge 1$.

Next define $\tilde{Y}_t = \mu + Q_t^2 + \delta^{-1}S_t^2$ for $t = 1,\dots,n-1$ and $\tilde{Y}_0 = \mu + (1+\delta^{-1})\theta_1$. Since $Y_{t-1}^{m-1} = \delta^{m-1}\tilde{Y}_{t-1}^{m-1}$ inequality (7) has the equivalent form

$$\mathrm{E}(Y_n^m) \le m(1+1/\delta)\delta^m\sum_{t=1}^{n}\mathrm{E}(\tilde{Y}_{t-1}^{m-1}X_t^2). \tag{8}$$

Note that $\tilde{Y}_{t-1} = \mu + Q_t^2 + \delta^{-1}S_{t-1}^2 - X_t^2$. The fact that there exists $\delta > 0$ small enough that

$$\sum_{t=1}^{n}\mathrm{E}(\tilde{Y}_{t-1}^{m-1}X_t^2) \le \sum_{t=1}^{n}\mathrm{E}((\mu + Q_t^2)^{m-1}X_t^2) \tag{9}$$

is shown by contradiction. Suppose that for some t, $\mathrm{E}(\tilde{Y}_{t-1}^{m-1}X_t^2) > \mathrm{E}((\mu+Q_t^2)^{m-1}X_t^2)$ held for all $\delta > 0$. Either $S_{t-1}^2 = 0$ with probability 1 or letting $\delta \downarrow 0$ would give $\mathrm{E}((\mu+Q_t^2)^{m-1}X_t^2) \le 0$, which is impossible unless $X_1 = \cdots = X_t = 0$. Therefore, $\sum_{t=1}^{n}\mathrm{E}(\tilde{Y}_{t-1}^{m-1}X_t^2) > \sum_{t=1}^{n}\mathrm{E}((\mu+Q_t^2)^{m-1}X_t^2)$ for all $\delta > 0$ implies that $S_{t-1}^2 = 0$ a.s. for $t = 2,\dots,n$ which is the exceptional case identified above.

By the mean value theorem there exist for $t = 2,\dots,n$ r.v.s $\eta_t \in [0,1]$ such that

$$m(\mu + Q_{t-1}^2 + \eta_t X_t^2)^{m-1}X_t^2 = (\mu + Q_t^2)^m - (\mu + Q_{t-1}^2)^m.$$

For the case $t = 1$ the same equality holds with $Q_0 = 0$ where for small enough μ, $\eta_1 \approx m^{-1/(m-1)} < 1$. Since $Q_{t-1}^2 + \eta_t X_t^2 \le Q_t^2$ and $m < 1$ it follows similarly to (9) that

$$\begin{aligned} m\sum_{t=1}^{n} \mathrm{E}((\mu + Q_t^2)^{m-1} X_t^2) &\le m\sum_{t=1}^{n} \mathrm{E}\big((\mu + Q_{t-1}^2 + \eta_t X_t^2)^{m-1} X_t^2\big) \\ &= \mathrm{E}(\mu + Q_1^2)^m + \sum_{t=2}^{n}\big(\mathrm{E}(\mu + Q_t^2)^m - \mathrm{E}(\mu + Q_{t-1}^2)^m\big) \\ &= \mathrm{E}(\mu + Q_n^2)^m. \end{aligned} \tag{10}$$

Combining (3), (7), (9) and (10) gives

$$2^{m-1}(\delta^m \mathrm{E}(\mu + Q_n^2)^m + \mathrm{E}|S_n|^{2m}) \le (1 + 1/\delta)\delta^m \mathrm{E}(\mu + Q_n^2)^m$$

which rearranges, after restoring $p = 2m$, as

$$\mathrm{E}|S_n|^p \le (2^{1-p/2}(1 + 1/\delta) - 1)\delta^{p/2}\mathrm{E}(\mu + Q_n^2)^{p/2}).$$

Since μ is arbitrary it can be set as small as desired. Letting $\mu \to 0$, the proof of (1) is completed by setting

$$C_p = (2^{1-p/2}(1 + 1/\delta_p) - 1)^{1/p}\delta_p^{1/2}$$

where δ_p is largest value of δ that satisfies (9) for every $n \ge 1$. ∎

To gain a feel for the value of δ_p in this result consider the 'worst case' $t = 2$, where in (9) it can be verified that the two expectations differ by $\delta^{-1}X_1^2$ appearing in the left-hand expression where X_2^2 appears on the right. The solution defines a sufficient condition; when n is large, satisfying (9) becomes a matter of $\delta^{-1}S_{t-1}^2$ dominating X_t^2 'on average'. Thus, asymptotically δ_p may be set simply to minimize C_p as a function of p; for example for $p = 1$ the minimum of 1.53 is found at $\delta = 3.4$.

References

[1] Burkholder, D.L. (1973), 'Distribution function inequalities for martingales', *Annals of Probability* 1, 19-42.

[2] Chow, Y. S. and Teicher, H. (1978), *Probability Theory: Independence, Inter-changeability and Martingales,* Springer-Verlag, Berlin.

[3] Hall, P and Heyde, C. C. (1980), *Martingale Limit Theory and its Application,* Academic Press, New York and London.

[4] Karatzas, Ioannis and Shreve, Steven E. (1991), *Brownian Motion and Stochastic Calculus,* 2nd Edn. Springer-Verlag, New York.

[5] Marcinkiewicz, J. and A. Zygmund (1938), Quelques theoremes sur les fonctions independantes. *Studia Mathematica* 7, 104-120.

[6] Novikov, A. A. (1971) 'On moment inequalities for stochastic integrals', *Theory of Probability and its Applications* 16(3) 538-541.