1. The Burkholder inequality

Let $\{(X_i, \mathcal{F}_i) : i = 1, \dots, n\}$ be a martingale. Define martingale indcrements $\xi_1 = X_1$ and $\xi_i = X_i - X_{i-1}$ for $i = 2, \dots, n$. The quadratic variation process is defined as

$$Q_i = \xi_1^2 + \dots \xi_i^2 \qquad \text{for } i = 1, \dots, n.$$

The Burkholder inequality shows that, as far as $\mathcal{L}^p$ norms are concerned, $\sqrt{Q_n}$ and X_n increase at the same rate: for each $p > 1$ there exist positive constants c_p and C_p such that

<1> $$c_p \|\sqrt{Q_n}\|_p \le \|X_n\|_p \le C_p \|\sqrt{Q_n}\|_p$$

In this note I will mostly follow the method presented by Burkholder (1973) to establish these inequalities.

The argument for the lower bound to $\|X_n\|_p$ is the more delicate. In fact, once we have the lower bound the duality argument for the upper bound is quite simple.

2. From tail bounds to norms

The method of proof for the first inequality in <1> relies on a handy result that relates norm bounds to bounds for tail probabilities.

<2> **Lemma.** *Suppose W and Z are nonnegative random variables for which there exists a constants $\beta > 0$ and C for which*

$$t\mathbb{P}\{W > \beta t\} \le C\mathbb{P}Z\{W > t\} \qquad \text{for all } t > 0.$$

Then for each $p > 1$ we have $\|W\|_p \le Cq\beta^p \|Z\|_p$ where $q = p/(p-1)$.

Proof. Note that $1/q + 1/p = 1$. Multiply both sides of the assumed inequality by pt^{p-2}, integrate with respect to t, then invoke Fubini.

$$\begin{aligned}
\|W/\beta\|_p^p &= \int_0^\infty pt^{p-1}\mathbb{P}\{W/\beta > t\} \\
&\le C\int_0^\infty pt^{p-2}\mathbb{P}X\{W > t\} \\
&= \frac{Cp}{p-1}\mathbb{P}XW^{p-1} \\
&\le \frac{Cp}{p-1}\|X\|_p\|W^{p-1}\|_q \qquad \text{by Hölder} \\
&= Cq\|X\|_p\|W\|_p^{p/q} \qquad \text{because } q(p-1) = p.
\end{aligned}$$

Thus

$$\|W\|_p = \|W\|_p^{p-p/q} \le Cq\beta^p\|X\|_p.$$

□

Remark. If you are worried about the possiblity that $\|W\|_p$ might be infinite, which would make the division by $\|W\|_p^{p/q}$ in the last step of the proof illegal, see Problem [1].

<3> **Corollary.** *(Doob's inequality) Let $M_n = \max_{i \le n} X_i$ for a nonnegative submartingale $\{X_i : i = 1, \dots, n\}$. Then $\|M_n\|_p \le q\|X_n\|_p$ for each $p > 1$.*

Proof. For a fixed $t > 0$, define $\tau = \inf\{i : X_i > t\}$. The set $F := \{\tau \le n\}$ is $\mathcal{F}_{\tau\wedge n}$-measurable and $\{M_n > t\} = F = \{X_{\tau\wedge n} > t\}$. By the Stopping Time Lemma (as in Section 6.2 of Pollard 2001), we get

<4> $$t\mathbb{P}\{M_n > t\} = t\mathbb{P}\{X_{\tau\wedge n} > t\}F \le \mathbb{P}X_{\tau\wedge n}F \le \mathbb{P}X_nF = \mathbb{P}X_n\{M_n > t\}.$$

□ Invoke Lemma <2> with $W = M_n$ and $X = X_n$ and $C = \beta = 1$.

3. The lower bound in Burkholder's inequality

The first inequality in <1> also holds for nonnegative submartingales. In fact, the result for nonnegative submartingales implies the analogous result for martingales via the inequalities for the submartingales $\{X_n^+ : i = 1, \dots, n\}$ and $\{X_n^- : i = 1, \dots, n\}$. See Problem [2]

There is no loss of generality, therefore, if I assume for the rest of this Section that $\{(X_i, \mathcal{F}_i) : i = 1, \dots, n\}$ is a nonnegative submartingale with increments ξ_i and $M_n = \max_{i\le n} X_i$ For convenience of notation define $X_0 = Q_0 = 0$ and $\xi_i = 0$ for $i > n$, so that $X_i = X_n$ for $i \ge n$.

The first step in the argument will show that

<5> $$t\mathbb{P}\{Q_n > t^2, M_n \le t\} \le 2\mathbb{P}X_n,$$

which might seem to be a long way from the sort of inequality we need for Lemma <2>. However, if we replace $\{X_i\}$ by the submartingale $X_i\{\sigma \le i\}$, for a suitable stopping time σ, then the analog of <5> for the new submartingale will lead quickly to a stronger result: for each $c_0 \ge 1$,

<6> $$t\mathbb{P}\{Q_n > (2 + c_0)t^2, M_n \le t\} \le 2\mathbb{P}X_n\{Q_n > c_0t^2\} \qquad \text{for each } t > 0.$$

If we define $W = \max(M_n, \sqrt{Q_n/c_0})$ and $\beta = \sqrt{1 + 2/c_0}$ then it will follow that

$$\begin{aligned}
&t\mathbb{P}\{W > \beta t\}\\
&\quad\le t\mathbb{P}\{M_n > t\} + t\mathbb{P}\{W > \beta t, M_n \le t\}\\
&\quad\le \mathbb{P}X_n\{M_n > t\} + t\mathbb{P}\{Q_n^2 > c_0\beta^2t^2, M_n \le t\} \qquad \text{by <4> and } W \text{ definition}\\
&\quad\le 3\mathbb{P}X_n\{W > t\} \qquad \text{by <6>.}
\end{aligned}$$

From Lemma <2> we will then get

$$\|\sqrt{Q_n}\|_p \le \sqrt{c_0}\|W\|_p \le \sqrt{c_0}3q(1 + 2/c_0)^{p/2}\|X_n\|_p$$

The choice $c_0 = 1$ gives a value $1/c_p = 3q3^{p/2}$ in inequality <1>. Burkholder (1973, page 22) chose $c_0 = p$ to get $1/c_p = 9q\sqrt{p}$, which is certainly sharper when p is large. I do not know whether the rate at which $1/c_p$ increases with p is important.

Thus the proof of the first inequality in <1> reduces to showing <5> and then finding the stopping time σ that leads to <6>.

Proof of inequality <5>. Define $Z_k := \sum_i\{1 \le i \le k\}X_{i-1}\xi_i$, so that

$$X_k^2 = \left(\sum\nolimits_{i=1}^k \xi_i\right)^2 = Q_k + 2Z_k \qquad \text{for } k = 0, 1, \dots, n.$$

Also define $\tau := \inf\{i \le n : X_i > t\}$ with the usual convention that $\inf\emptyset = +\infty$. Note that $\{M_n \le t\} = \{\tau = \infty\}$. Thus

$$(*) := t\mathbb{P}\{Q_n > t^2, M_n \le t\} \le \mathbb{P}Q_n\{M_n \le t\}/t = \mathbb{P}\left(X_n^2 - 2Z_n\right)\{M_n \le t\}/t.$$

When $\tau = \infty$ we have $Z_n = Z_{\tau\wedge n}$ and $X_n^2 = X_{\tau\wedge n}^2 \le tX_{\tau\wedge n}$, which implies

$$(*) \le \mathbb{P}\left(tX_{\tau\wedge n} - 2Z_{\tau\wedge n}\right)\{\tau = \infty\}/t \le 2\mathbb{P}\left(X_{\tau\wedge n} - Z_{\tau\wedge n}/t\right)\{\tau = \infty\}$$

On the set where $\tau \wedge n = k$ we have

$$\begin{aligned} tX_k &\ge X_{k-1}X_k = X_{k-1}^2 + \xi_k X_{k-1} \\ &= \tfrac{1}{2}\left(Q_{k-1} + 2Z_k\right) + \tfrac{1}{2}X_{k-1}^2 + \xi_k X_{k-1} \\ &\ge Z_k. \end{aligned}$$

That is, $tX_{\tau\wedge n} - Z_{\tau\wedge n} \ge 0$ everywhere.

Moreover, an appeal to the Stopping Time Lemma (as in Problem 6.3 of Pollard 2001) shows that $\{(X_{\tau\wedge k}, \mathcal{F}_k) : k = 1, \dots, n\}$ is a submartingale with $\mathbb{P}X_{\tau\wedge n} \le \mathbb{P}X_n$. Only slightly more subtle (see Problem [3]) is the fact that the stopped process $\{Z_{\tau\wedge k} : k = 0, 1, \dots, n\}$ is also a submartingale, with $\mathbb{P}Z_{\tau\wedge n} \ge \mathbb{P}Z_0 = 0$. We can therefore conclude that

$$(*) \le 2\mathbb{P}\left(X_{\tau\wedge n} - Z_{\tau\wedge n}/t\right) \le 2\mathbb{P}X_n,$$

as asserted by <5>.

Proof of inequality <6>. Define $\sigma := \inf\{i \le n : Q_i > c_0 t^2\}$. First check that $Y_i := X_i\{\sigma \le i\}$ is a nonnegative submartingale: If $F \in \mathcal{F}_{i-1}$ then

$$\mathbb{P}Y_i F \ge \mathbb{P}FX_i\{\sigma \le i-1\} \ge \mathbb{P}FY_{i-1}$$

because $F\{\sigma \le i-1\} \in \mathcal{F}_{i-1}$ and X_i is a submartingale. The new submartingale has increments

$$\begin{aligned} \eta_i := Y_i - Y_{i-1} &= X_i\{\sigma = i\} + (X_i - X_{i-1})\{\sigma \le i-1\} \\ &= X_\sigma\{\sigma = i\} + \xi_i\{\sigma \le i-1\} \end{aligned}$$

and quadratic variation

$$\widetilde{Q}_k := \textstyle\sum_i \{1 \le i \le k\}\eta_i^2 = \left(X_\sigma^2 + Q_n - Q_\sigma\right)\{\sigma \le k\}$$

On the set $D := \{Q_n > (2 + c_0)t^2, M_n \le t\}$ we must have $\sigma \le n$ and $\max_{i\le n} Y_i \le t$. Also the inequality $-X_{i-1} \le \xi_i \le X_i$ implies that $\max_{i\le n}|\xi_i| \le M_n \le t$. It follows that, on D, we have

$$\widetilde{Q}_n \ge Q_n - (Q_{\sigma-1} + \xi_\sigma^2) > (2 + c_0)t^2 - (c_0 t^2 + t^2) = t^2.$$

These facts give

$$t\mathbb{P}D \le \mathbb{P}\{\widetilde{Q}_n > t^2, \max_{i\le n} Y_i \le t\}$$

which, by <5> applied to Y_i instead of X_i, is less than

$$2\mathbb{P}Y_n = 2\mathbb{P}X_n\{\sigma \le n\} = 2\mathbb{P}X_n\{Q_n > c_0 t^2\}.$$

□ Inequality <6> follows.

4. The upper bound in Burkholder's inequality

For a random variable Γ with $\|T\|_q \le 1$, Hölder's inequality gives $\mathbb{P}(\Gamma X_n) \le \|X_n\|_p$. If $\|X_n\|_p \ne 0$, equality is achieved by the choice $\Gamma = |X_n|^{p-1}\operatorname{sgn}(X_n)/\|X_n\|_p^{p/q}$. Thus

<7> $$\|X_n\|_p = \sup\{\mathbb{P}(\Gamma X_n) : \|\Gamma\|_q \le 1\}.$$

This duality property will lead to the upper bound for $\|X_n\|_p$.

The sequence $\mathbb{P}(\Gamma \mid \mathcal{F}_i)$ for $i = 1, \dots, n$, is a martingale. Write $\gamma_1, \dots, \gamma_n$ for its increments and G_n for its quadratic variation $\sum_{i \le n} \gamma_i^2$. Then

$$
\begin{aligned}
\mathbb{P}(\Gamma X_n) &= \mathbb{P}\left(\textstyle\sum_{i \le n} \gamma_i\right)\left(\textstyle\sum_{i \le n} \xi_i\right) \\
&= \mathbb{P}\textstyle\sum_{i \le n} \gamma_i \xi_i \qquad \text{because } \mathbb{P}\gamma_i\xi_j = 0 = \mathbb{P}\xi_i\gamma_j \text{ for } i < j \\
&\le \mathbb{P}\sqrt{G_n Q_n} \qquad \text{by Cauchy-Schwartz} \\
&\le \|\sqrt{G_n}\|_q \|\sqrt{Q_n}\|_p \qquad \text{by Hölder} \\
&\le c_p^{-1} \|\Gamma\|_q \|\sqrt{Q_n}\|_p \qquad \text{by the lower bound from inequality <1>.}
\end{aligned}
$$

Take the supremum over Γ with $\|\Gamma\|_q \le 1$, or just choose Γ to achieve the supremum in <7>, to get the Burkholder upper bound with $C_p = 1/c_p$.

5. Problems

[1] Suppose $\|Z\|_p$ in Lemma <2> is finite. Replace β by $\max(1, \beta)$. Explain why the inequality for $\mathbb{P}\{W > \beta t\}$ still holds if W is replaced by $W \wedge k$, for any positive integer k. Deduce that $\|W \wedge k\|_p$ is bounded by a multiple of $\|Z\|_p$ that doesn't depend on k. Let k tend to infinity to deduce that $\|W\|_p < \infty$.

[2] Let $\{X_n : i = 1, \dots, n\}$ be a martingale.

(i) Show that $\|X\|_p \ge \max(\|X_n^+\|_p, \|X_n^-\|_p)$.

(ii) Show that $Q_n \le \sum 2 \sum_{i \le n} \left((X_i^+ - X_{i-1}^+)^2 + (X_i^+ - X_{i-1}^+)^2\right)$

(iii) Deduce the lower bound from <1> for the martingale from the analogous lower bounds for the nonnegative submartingales $\{X_i^+ : i = 1, \dots, n\}$ and $\{X_i^- : i = 1, \dots, n\}$.

[3] With $Z_k := \sum_i \{1 \le i \le k\} X_{i-1}\xi_i$ as in Section 3:

(i) Show that $Z_{\tau \wedge k} = \sum_{i=1}^k X_{i-1}\xi_i\{\{i \le \tau\}$. Deduce that the kth increment of the stopped process is bounded in absolute value by $t|\xi_k|$, which is integrable.

(ii) If $F \in \mathcal{F}_{k-1}$, show that $\mathbb{P} F X_{k-1}\xi_i\{\{k \le \tau\} = 0$. Deduce that the stopped process is a submartinagle.

References

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

Pollard, D. (2001), *A User's Guide to Measure Theoretic Probability*, Cambridge University Press.

15 June 2007 Burkholder, version: June07 ©David Pollard