Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Matematica

Burkholder's Sharp L^p estimates for *martingale transforms*

Tesi di Laurea in Analisi Armonica

Relatore:
Chiar.mo Prof.
NICOLA ARCOZZI

Presentata da:
ANGELICA LONGETTI

V Sessione
Anno Accademico 2015/2016

Abstract

L'argomento della tesi è una disuguaglianza L^p per trasformate di martingala. La trasformata di una martingala X_n si ottiene moltiplicando le sue differenze per una sequenza prevedibile H_n, ottenendo così la sequenza delle differenze della martingala trasformata $(H \cdot X)_n$.
E' interessante che siano stati identificate le stime esatei in L^p per queste trasformate.

Nella presente tesi si discuterà sulla tecnica di dimostrazione, dovuta al probabilista Burkholder
Questa si sviluppa due parti: (i) la prova della disuguaglianza tramite una "funzione di Bellman" a due variabili con determinate proprietà e (ii) la prova dell'esistenza di tale funzione, che viene costruita esplicitamente.
Ciò che è sorprendente è che Burkholder sia stato in grado di individuarla.

La ricerca si è successivamente ampliata ad altre disuguaglianze, con applicazioni a vari problemi di analisi stocastica, generalizzando i risultati e le idee di Burkholder in differenti contesti. Si tratta di un campo di ricerca corrente in continuo sviluppo.

This thesis is about some sharp inequalities for martingale transforms. The martingale transform acts on a martingale X_n by multiplying its difference sequence times a fixed bounded, predictable process H_n, obtaing the difference sequence of the transformed martingale $(H \cdot X)_n$.
It came as a surprise that sharp L^p bounds of such transforms could be found and proved.

In this thesis shall explain Burkholder's technique of proof.It rests on two steps: (i) the inequality holds if a suitable, two-variables "Bellman function" with certain properties is known and (ii) the existence of such Bellman function is proved.
What is surprising is the fact that Burkholder was able to find an explicit expression for the Bellman function for the problem L^p estimates we consider.

Much research has been done later, extending Burkholder's methods and ideas to various inequalities, applying it to various problems in stochastic analysis, extending it to different contexts.

I wuold like to express my thanks and gratitude
to my supervisor and to my family.

Contents

Introduction

In this thesis I'm going to introduce some selected concepts about Martingales, aiming to prove very important theorems about optimality, in particular the Burkholder's Theorem.

After preliminary notions then I'll introduce the Doob's optional stopping theorem, first for submartingales uniformely integrable and then for martingales, exhibiting the conditions under which the theorem holds. To follow the famous application: the Secretary Problem.

The main topic are the Burkholder's L^p sharp inqualities for martingale transforms [4][5] for $1 < p < \infty$.
The martingale transform acts on a martingale X_n by multiplying its difference sequence times a fixed bounded, predictable process H_n, obtaing the difference sequence of the transformed martingale $(H \cdot X)_n$.
The Burkholder results and ideas about that are nowdays important: the "method of Bellman's functions" [6], which is a direct generalization of Burkholder's method, is a current area of intensive and important research.
Burkholder's proof rests on two steps:
I. The inequality is proved if a two-variables "Bellman's function" $z = U(x, y)$ having certain properties exist.
II. The existence of such function is proved [4], or, even better, the function is exhibited [5][2]. Once the function U is exhibited, verifying its relevant properties is an long, but standard exercise in multivariable calculus.
What is surprising is the fact that Burkholder was able to find it.
A few remarks about the proof are in order.
First, I have considered the case of dyadic martingales only, because I thought some passages were more transparent in this case and because the modern theory of Bellman functions is set in the dyadic setting (the proof extends without changes to the general case of discrete martingales).
Second, I give the proof for $p \geq 2$ only beacuse I thought it was enough to consider one range of the exponent to get acquainted with the techniques (the proof for $1 < p \leq 2$ is similar in spirit, but is is does not simply follows from "passing to the conjugate exponent").

Third, I did not put in the thesis the sequence of examples showing that the constant is the best one. Because of the short time writing I preferred to concentrate on the presentation of the main result.

Chapter 1

Martingales and Prerequisites

Before beginning our discussion on specific properties of Martingales, we need to remember some important prerequisites such as the conditional expectiation, i.e. given information, the way in which the probability of events changes.

Notation: We'll use X_n instead of $\{X_n\}_{n\geq 0}$ such as in [1].

1.1 Conditional Expectation and Probability

Definition 1.1 (**Conditional Expectation**)**.** Given are a probability space $(\Omega, \mathcal{F}_0, P)$, a $\sigma - field$ $\mathcal{F} \subset \mathcal{F}_0$, and a random variable $X \in \mathcal{F}_0$ with $E \mid X \mid < \infty$, we define the *conditional expectation* of X given $\mathcal{F}$, $E\,(X \mid \mathcal{F})$,to be any random variable Y such that:

(i) $Y \in \mathcal{F}$, i.e., Y is $\mathcal{F}$-measurable,
(ii) for all $A \in \mathcal{F}$, $\int_A X\, dP$=$\int_A Y\, dP$.

Any Y satisfying (i) and (ii) is said to be a **version of** $E\,(X \mid \mathcal{F})$. The first thing to be settled is that the conditional expectation exists and is unique.

Proof. , if Y' also satisfies (i) and (ii) then

$$\int_A Y\, dP = \int_A Y'\, dP \text{ for all } A \in \mathcal{F}.$$

Taking $A = \{Y - Y' \geq \varepsilon > 0\}$, we see

$$0 = \int_A X - X\, dP = \int_A Y - Y'\, dP \geq \varepsilon.$$

so $P(A) = 0$. Since this holds for all ε we have $Y \leq Y'$ a.s., and interchanging the roles of Y and Y', we have Y=Y' a.s.

(Technically, all equalities such as $Y{=}E\,(X\mid\mathcal{F})$ should be written Y=$E\,(X\mid\mathcal{F})$ a.s.).

□

Lemma 1.1.1.
If Y satisfies (i) and (ii), then it is integrable.

Proof. Letting $A{=}\{Y>0\}\in\mathcal{F}$, using (ii) twice, and then adding

$$\int_A Y\,dP{=}\int_A X\,dP\leq\int_A\mid X\mid\,dP,$$
$$\int_{A^c}-Y\,dP{=}\int_{A^c}-X\,dP\leq\int_{A^c}\mid X\mid\,dP.$$

So we have $E\mid Y\mid\leq E\mid X\mid$.

□

1.1.1 Properties

Theorem 1.1.2.
(a) Linearity of Conditional Expectation:

$$E(aX+Y\mid\mathcal{F}){=}aE(X\mid\mathcal{F})+E(Y\mid\mathcal{F}),$$

(b) If $X\leq Y$ then

$$E(X\mid\mathcal{F})\leq E(Y\mid\mathcal{F}),$$

(c) If $X_n\geq 0$ and $X_n\uparrow X$ with $E(X)<\infty$ then

$$E(X_n\mid\mathcal{F})\uparrow E(X\mid\mathcal{F}).$$

Observation 1.
By appying the last result to Y_1-Y_n we see that if $Y_n\downarrow Y$ and we have $E\mid Y_1\mid,E\mid Y\mid<\infty$ then $E(Y_n\mid\mathcal{F})\downarrow E(Y\mid\mathcal{F})$.

Proof. To prove (a), we need to check that the right-hand side is a version og the left. It clearly is $\mathcal{F}$-measurable. To check (ii), we observe that if $A\in\mathcal{F}$ then by linearity of the integral and the defining properties of $E(X\mid\mathcal{F})$ and $E(Y\mid\mathcal{F})$,

$$\int_A\{aE(X\mid\mathcal{F})+E(Y\mid\mathcal{F})\}\,,dP{=}a\int_A E(X\mid\mathcal{F})\,dP{+}\int_A E(Y\mid\mathcal{F})\,dP{=}$$
$${=}a\int_A X\,dP{+}\int_A Y\,dP\;{=}\int_A aX+Y\,dP$$

which proves (a).

Using the definition

$$\int_A E(X \mid \mathcal{F})\, dP = \int_A X\, dP \leq \int_A Y\, dP = \int_A E(Y \mid \mathcal{F})\, dP.$$

Letting $A = \{E(X \mid \mathcal{F}) - E(Y \mid \mathcal{F}) \geq \varepsilon > 0\}$, we see that the indicated set has probability 0 for all $\varepsilon > 0$, and we have proved (b).

Let $Y_n = X - X_n$. It suffices to show that $E(Y_n \mid \mathcal{F}) \downarrow 0$.
Since $Y_n \downarrow$, (b) implies $Z_n \equiv E(Y_n \mid \mathcal{F}) \downarrow$ a limit Z_∞.If $A \in \mathcal{F}$ then

$$\int_A Z_n\, dP = \int_A Y_n\, dP.$$

Letting $n \longrightarrow \infty$, noting $Y_n \downarrow 0$, and using the dominated convergence theorem gives that $\int_A Z_\infty\, dP = 0$ for all $A \in \mathcal{F}$, so $Z_\infty \equiv 0$.

□

Theorem 1.1.3.
If $\mathcal{F} \subset \mathcal{G}$ and $E(X \mid \mathcal{G}) \in \mathcal{F}$ then $E(X \mid \mathcal{F}) = E(X \mid \mathcal{G})$.

Proof. By assumption $E(X \mid \mathcal{G}) \in \mathcal{F}$. To check the other part of definition we note that if $A \in \mathcal{F} \subset \mathcal{G}$ then

$$\int_A X\, dP = \int_A E(X \mid \mathcal{G})\, dP.$$

□

Theorem 1.1.4 (Tower property).
If $\mathcal{F}_1 \subset \mathcal{F}_2$ then $E(E(X \mid \mathcal{F}_2) \mid \mathcal{F}_1) = E(X \mid \mathcal{F}_1)$.

Proof. Noticing that $E(X \mid \mathcal{F}_1) \in \mathcal{F}_2$, and if $A \in \mathcal{F}_1 \subset \mathcal{F}_2$ then

$$\int_A E(X \mid \mathcal{F}_1)\, dP = \int_A X\, dP = \int_A E(X \mid \mathcal{F}_2)\, dP.$$

□

Theorem 1.1.5.
If $X \in \mathcal{F}$ and $E \mid X \mid, E \mid XY \mid < \infty$ then

$$E(XY \mid \mathcal{F}) = XE(Y \mid \mathcal{F}).$$

Proof. The right-hand side $\in \mathcal{F}$, so we have to check the tower property. To do this, we use the usual four-step procedure. FIrst, suppose $X = 1_B$ with $B \in \mathcal{F}$. In this case, if $A \in \mathcal{F}$

$$\int_A 1_B E(X \mid \mathcal{F})\, dP = \int_{A\cap B} E(Y \mid \mathcal{F})\, dP = \int_{A\cap B} Y\, dP = \int_A 1_B Y\, dP,$$

so the tower property holds. The last result extends to simple X by linearity. If $X, Y \geq 0$, let X_n be simple random variables that $\uparrow X$, and use the monotone convergence rheorme to conclude that

$$\int_A XE(Y \mid \mathcal{F})\, dP = \int_A XY\, dP.$$

To prove the result in general, split X and Y into their positive and negative parts.

□

1.1.2 Regular Conditional Probabilities

Definition 1.2 (**Regular Conditional Probabilities**)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space, $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ a measurable map, and $\mathcal{G}$ a σ-field $\subset \mathcal{F}$. $\mu : \Omega x \mathcal{S} \to [0, 1]$ is said to be a **regular conditional distribution** for X given $\mathcal{G}$ if
(i) For each A,$\omega \to \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$.
(ii) For a.e. $\omega, A \to \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$.
When $S = \Omega$ and X is the identity map, μ is called a *regular conditional probability.*

1.2 Martingales

When we are talking about gambling, the stochastic process *martingale* is representing the notion of a fair game, in which we have no profit or loss for every gamble on average, regardless of the past gambles.
In other words we can think about martingale X_n as the fortune at time n of a gambler who is betting on a fair game; submartingale as the outcome on a favorable game and supermartingale on a unfavorable game.
Martingales are not used just for gambling but they have applications on stochastic modelling. Let see some more formal definitions and important theorems about it.

1.2.1 Martingales, supermartingales, submartingale

Definition 1.3 (**Filtration**)**.** An increasing sequence of σ-fieds $\mathcal{F}_n$.

Definition 1.4 (**Adapted process**)**.** A sequence X_n is said to be *adapted process* to $\mathcal{F}_n$ if $X_n \in \mathcal{F}_n$ for all n.

Definition 1.5 (**Martingale**). A sequence X_n is said to be *martingale* (with the respect to $\mathcal{F}_n$) if
(i) $E \mid X_n \mid < \infty$,
(ii) X_n is adapted to $\mathcal{F}_n$,
(iii) $E(X_n + 1 \mid \mathcal{F}_n)$=$X_n$ for all n.

`Example. Simple random walk.` Consider the successive tosses of a fair coin and let $\xi_n = 1$ if nth tossis heads and $\xi_n = -1$ if nth toss is tails. Let $X_n = \xi_1 + \xi_2 + ... + \xi_n$ and $\mathcal{F}_n = \sigma(\xi_1, ...\xi_n)$ for $n \geq 1$, $X_0 = 0$ and $\mathcal{F}_0$=$\{\emptyset, \Omega\}$.

We claim that X_n, $n \geq 0$, is a martingale with the respect to $\mathcal{F}_n$. To prove this, we observe that $X_n \in \mathcal{F}_n$,$E \mid X_n \mid < \infty$, and ξ_{n+1} is independent of $\mathcal{F}_n$, so using the linearity of conditional expectation and Bayes' Formula:

$$E(X_{n+1} \mid \mathcal{F}_n)=E(X_n \mid \mathcal{F}_n)+E(\xi_{n+1} \mid \mathcal{F}_n)=X_n + E\xi_{n+1}=X_n$$

Note that, in this, example, $\mathcal{F}_n$=$\sigma(X_1, ..., X_n)$ and $\mathcal{F}_n$ is the smallest filtration that X_n is adapted to.

Definition 1.6 (**Supermartingale**). A sequence X_n is said to be *supermartingale* (with the respect to $\mathcal{F}_n$) if
(i) $E \mid X_n \mid < \infty$,
(ii) X_n is adapted to $\mathcal{F}_n$,
(iii) $E(X_{n+1} \mid \mathcal{F}_n) \leq X_n$ for all n.

`Example.` If the coin tosses considered above have $P(\xi_n = 1) \leq 1/2$ then the computation just completed shows $E(X_{n+1} \mid \mathcal{F}_n) \leq X_n$, i.e., X_n is a supermartingale. In this case, X_n corresponds to betting on an unfavorable game.

Definition 1.7 (**Submartingale**). A sequence X_n is said to be *submartingale* (with the respect to $\mathcal{F}_n$) if
(i) $E \mid X_n \mid < \infty$,
(ii) X_n is adapted to $\mathcal{F}_n$,
(iii) $E(X_{n+1} \mid \mathcal{F}_n) \geq X_n$ for all n.

1.2.2 Doob's decomposition Theorem

The Doob's decomposition theorem says that in a probability space we can make an almost surely decomposition from *every* $\mathcal{F}_n$-adapted *stochatic process* X_n with $E \mid X_n \mid < \infty$ to *a martingale* and a *predictable process.*
It's worth considering this result about submartingales and supermartingales.

Theorem 1.2.1 (Doob's decomposition of submartingales).
Any submartingale X_n, $n \geq 0$ can be written in a unique way as X_n=M_n+A_n, where M_n is a martingale and A_n is a predictable increasing sequence with $A_0 = 0$.

Proof. We want X_n=$M_n + A_n$, $E(M_n \mid \mathcal{F}n-1)$=$Mn-1$, and $A-n \in \mathcal{F}n-1$. So we must have:

$$E(X_n \mid \mathcal{F}_{n-1}){=}E(M_n \mid \mathcal{F}_{n-1}){+}E(A_n \mid \mathcal{F}_{n-1}){=}$$

$$={Mn-1+A_n}{=}X_{n-1}-A_{n-1}+A_n.$$

and it follows that:
(a) $A_n - A_{n-1}$=$E(X_n \mid \mathcal{F}_{n-1}) - X_{n-1}$,
(b) M_n=$X_n - A_n$.
Now $A_0 = 0$ and M_0=X_0 by assumption, so we have A_n and M_n defined for all time and we have proved uniqueness.
To check that our recipe works, we observe that $A_n - A_{n-1} \geq 0$ since X_n is a submartingale and induction shows $A_n \in \mathcal{F}n-1$. To see that M_n is a martingale, we use (b), $A_n \in \mathcal{F}n-1$ and (a):

$$E(M_n \mid \mathcal{F}n-1){=}E(X_n-A_n \mid \mathcal{F}_{n-1}){=}$$

$$={E(X_n \mid \mathcal{F}_{n-1})-A_n}{=}X_{n-1}-A_{n-1}{=}M_{n-1}.$$

Which completes the proof.

□

Observation 2.
If we want to decompose a supermartingale X_n we'll obtain a martingale M_n and a predictable decreasing sequence A_n with $A_0 = 0$, the proof is similar.

1.2.3 Predictable sequences and the impossibility of beating the system

Suppose that $\{H_n : n \geq 1\}$ is the stake of a gambler on game (at time) n. The gambler has to base his decision on H_n on the history of the game up to time $n-1$ (we are saying that H_n is $\mathcal{F}_{n-1}$-measurable).

Definition 1.8 (Predictable Sequence). Let $\mathcal{F}_n$, $n \geq 0$ be a filtration. H_n is said to be a *predictable process* if $H_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$.

In practice, supposing that the game consists of flipping a coin and that for each dollar the gambler bets he wins one dollar when the coin comes up heads and loses his dollar when the coin comes up tails.

The winnings at time n are $H_n(X_n - X_{n-1})$ and the total winnings up to the time n are given by $(H \cdot X)_n = \sum_{m=1}^{n} H_m(X_m - X_{m-1})$.

`Example. Martingale.` This is a famous gambling system defined by $H_1 = 1$ and for $n \geq 2$,

$$H_n = 2H_{n-1} \text{ if } X_{n-1} - X_{n-2} = -1 \text{ and } H_n = 1 \text{ if } X_{n-1} - X_{n-2} = 1.$$

In other words the gambler doubles his bet when he loses, so if he loses k times and then wins, the net winning will be $-1 - 2... - 2^k = 1$. This system seems to be a "sure thing" as long as $P((X_m - X_{m-1}) = 1)$.

We want to know if the gambler can chose H_n such that the expected total winnings are positive.

Definition 1.9 (Martingale Transform of X_n by H_n). The process $(H \cdot X)_n$ is called *martingale transform of X_n by H_n.*

Theorem 1.2.2 (No way to beat on unfavorable game).
Let X_n, $n \geq 0$, be a supermartingale. If $H_n \geq 0$ is predictable and each H_n is bounded then $(H \cdot X)_n$ is a supermartingale.

Proof. Using the fact thtat contitional expectation is linear, $(H \cdot X)_n \in \mathcal{F}_n$, $H_n \in \mathcal{F}_{n-1}$, and Theorem 1.1.5, we have

$$E((H \cdot X)_{n+1} \mid \mathcal{F}_n) = (H \cdot X)_n + E(H_{n+1}(X_{n+1} - X_n) \mid \mathcal{F}_n) =$$

$$= (H \cdot X)_n + H_{n+1} E((X_{n+1} - X_n) \mid \mathcal{F}_n) \leq (H \cdot X)_n.$$

Since $E((X_{n+1} - X_n) \mid \mathcal{F}_n) \leq 0$ and $H_{n+1} \geq 0$.

□

Observation 3.
The same result is obviously true for submartingales and for martingales (in the last case, withouth the restriction $H_n \geq 0$).

1.2.4 Stopping Time

Now we are interessed to introduce a new concept of time, closely related to the concept of a gambling system, noticing the property of martingales $E(X_n)$=$E(X_0)$, $n \geq 0$, which can be extended to $E(X_N)$=$E(X_0)$, $N \leq n$. We can think of stopping time as the time a gambler stop gambling, considering that the decision to stop the game at time n must be measurable with the respect to the information the gambler has at that time.

Definition 1.10 (**Stopping Time**). A random variable N is said to be a stopping time if $\{N = n\} \in \mathcal{F}_n$ for all $n < \infty$.

Observation 4.
If we have H_n=$1_{\{N \geq n\}}$, then $\{N \geq n\}$=$\{N \leq n-1\}^c \in \mathcal{F}_{n-1}$, so H_n is predictable, and it follows from Theorem 1.2.1 that $(H \cdot X)_n$=$X_{N\wedge n} - X_0$ is a supermartingale.

Theorem 1.2.3.
Let $X = \{X_n : n \geq 0\}$ be a martingale and N a stopping time w.r.t. X, then the stopped process $\hat{X}$=$\left\{\hat{X}_n : n \geq 0\right\}$ is a martingale, where:

$$\hat{X} := \begin{cases} X_n, & \text{if } N > n \\ X_N, & \text{if } N \leq n \end{cases} = X_{N\wedge n}.$$

Since $\hat{X}_0 = X_0$, we conclude that $E(\hat{X}_n)$=$E(X_0)$, $n \geq 0$.

Proof. (i) Since $\mid \hat{X}_n \mid \leq \max_{0\leq k\leq n} \mid X_k \mid \leq \mid X_0 \mid + ... + \mid X_n \mid$, we conclude that $E \mid \hat{X}_n \mid \leq E(\mid X_0 \mid) + ... + E(\mid X_n \mid) < \infty$.
(ii) by definition of $\hat{X}$.
(iii) It is sufficient to use $\mathcal{F}_n$=$\sigma\{X_0, ...X_n\}$ since $\sigma\left\{\hat{X}_0, ...\hat{X}_n\right\} \subset \mathcal{F}_n$ by the stopping time property that $\{N > n\}$ is determined by $\{X_0, ..., X_n\}$. Noticing that both $\hat{X}_n = X_n$ and $\hat{X_{n+1}}$=X_{n+1} if $N > n$, and $\hat{X}_{n+1}$=$\hat{X}_n$ if $N \leq n$ yelds

$$\hat{X}_{n+1} = \hat{X}_n + (X_{n+1} - X_n)1_{\{N\geq n\}};$$

Thus

$$E(\hat{X}_{n+1} \mid \mathcal{F}_n) = \hat{X}_n + E((X_{n+1} - X_n)1_{\{N\geq n\}} \mid \mathcal{F}_n) =$$

$$= \hat{X}_n + 1_{\{N\geq n\}} E((X_{n+1} - X_n) \mid \mathcal{F}_n) =$$

$$= \hat{X}_n + 1_{\{N\geq n\}} \cdot 0 = \hat{X}_n.$$

□

Theorem 1.2.4.
If N is a stopping time and X_n a supermartingale, then $X_{N\wedge n}$ is a supermartingale.

Theorem 1.2.5.
If X_n is a submartingale and N is a stopping time with $P(N \leq k) = 1$ then

$$EX_0 \leq EX_N \leq EX_k.$$

Proof. The Theorem above implies that $X_{N\wedge n}$ is a supermartingale, so it follows that

$$E(X_0)=(EX_{N\wedge 0}) \leq E(X_{N\wedge k})=E(X_n).$$

To prove the other inequality, let $K_n=1_{N<n}=1N \leq n-1$. K_n is predictable, so Theorem 1.2.2 implies $(K \cdot X)_n=X_n - X_{N\wedge n}$ is a submartingale and it follows that

$$E(X_k) - E(X_N)=E((K \cdot X)_n) \geq E((K \cdot X)_0)=0.$$

□

1.2.5 Convergence

This gives sufficient condition for the almost sure convergence of martingales X_n to a limiting random variable.

Theorem 1.2.6.
If X_n is a submartingale w.r.t. $\mathcal{F}_n$ and φ is an increasing convex function with $E \mid \varphi(X_n) \mid \leq \infty$ for all n, then $\varphi(X_n)$ is a submartingale w.r.t. $\mathcal{F}_n$.
Consequently:
(i) If X_n is a submartingale then $(X_n - a)^+$ is a submartingale.
(ii) If X_n is a supermarginale then $X_n \wedge a$ is a supermartingale.

Proof. By Jensen's inequality and the assumpions

$$E(\varphi(X_{n+1}) \mid \mathcal{F}_n) \geq \varphi(E(X_{n+1}) \mid \mathcal{F}_n)) \geq \varphi(X_n).$$

□

Theorem 1.2.7 (Upcrossing inequality)**.**
If X_m,$m \geq 0$, is a submartingale then

$$(b-a)EU_n \leq E(X_n - a)^+ - E(X_0 - a)^+ \ .$$

Proof. Let $Y_m = a + (X_m - a)^+$. By Theorem above, Y_m is a submartingale. Clearly, it upcrosses $[a, b]$ the same number of times that X_m does, and we have $(b - a)U_n \leq (H \cdot Y)_n)$, since each upcrossing results
in a profit $\geq (b - a)$ and a final incomplete upcrossing (if there is one) makes a nonnegative contribution to the right-hand side. Let $K_m = 1 - H_m$.
Clearly, $Y_n - Y_0 {=} (H \cdot Y)_n + (K \cdot Y)_n$, and it follows from Theorem 1.2.2 that $E(K \cdot Y)_n \geq E(K \cdot Y)_0 {=} 0$ so $E(H \cdot Y)_n \leq E(Y_n - Y_0)$, proving the desired inequality.

□

Theorem 1.2.8 (**Martingale Convergence Theroem**).
If X_n is a submartingale with $\sup E(X_n^+) < \infty$ then as $n \to \infty$, X_n converges a.s. to a limit X with $E \mid X \mid < \infty$.

Proof. Since $(X - a)^+ \leq X^+ + \mid a \mid$, the upcrossing inequality implies that $E(U_n) \leq (\mid a \mid + E(X_n^+))/(b - a)$.
As $n \uparrow \infty$, $U_n \uparrow U$ the number of upcrossing of $[a, b]$ by the whole sequence, so if $supE(X_n^+) < \infty$ then $E(U) < \infty$ a.s.
Since the last conclusion holds for all rational a and b,

$$\cup_{a,b \in \mathbb{Q}} \liminf X_n < a < b < limsup X_n \text{ has probability } 0$$

and hance $\limsup X_n {=} \liminf X_n$ a.s., i.e., $\lim X_n$ exists a.s. Faout's lemma guarantees $E(X^+) \leq \liminf E(X_n^+) < \infty$, so $X < infty$ a.s.
To see $X > -\infty$, we observe that

$$E(X_n^-) {=} E(X_n^+) - E(X_n) \leq E(X_n^+) - E(X_0);$$

(since X_n is a submartingale), so another application of Fatou's lemma shows

$$E(X^-) \leq \liminf_{n \to \infty} E(X_n^-) \leq \sup_n E(X_n^+) - E(X_0) < \infty.$$

□

Chapter 2

Martingales in Optimisation Problem

2.1 Optional Stopping Theorem

Now we want to find the conditions under which we can prove that if X_n is a submartingale, $M \leq N$ are stopping times, then $E(X_M) \leq E(X_N)$. The key to this is the following definition:

Definition 2.1 (Uniformly Integrable). A collection of random variables X_i, $i \in I$, is said to be *uniformly integrable* if

$$\lim_{n\to\infty}(\sup_{i\in I} E(\mid X_i \mid), \mid X_i \mid) > M).$$

A trivial example of a uniformly integrable family is a collection of random variables that are dominated by an integrable random variable, i.e., $\mid X_i \mid \leq Y$ where $E(Y) < 1$.

Theorem 2.1.1.
If X_n is a uniformly integrable submartingale then for any stopping time N,$X_{N\wedge n}$ is uniformly integrable.

Proof. X_n^+ is a submartingale, so Theorem 1.2.5 implies $E(X_{N\wedge n}^+) \leq E(X_n^+)$. Since X_n^+ is uniformely integrable, it follows from the definition that $\sup_n E(X_{N\wedge n}^+) \leq \sup_n E(X_n^+) < \infty$.
Using the Martingale Convergence Theorem now gives $X_{N\wedge n}^+ \to X_n$ a.s. (here X_∞=$\lim_n X_n$) and $E \mid X_N \mid < \infty$. With this established, the rest is easy. We write

$$E(\mid X_{N\wedge n} \mid ; \mid X_{N\wedge n} \mid > K) = E(\mid X_N \mid ; \mid X_N \mid > K, N \leq n) + E(\mid X_N \mid ; \mid X_N \mid > K, N > n)$$

Since $E \mid X_N \mid \leq \infty$ and X_n is uniformly integrable, if K is large then each term is $\varepsilon/2$.

□

From the last computation in the proof above, we get:

Theorem 2.1.2.
If $E|X_N| < \infty$ and $X_n 1_{(N>n)}$ is uniformly integrable, then $X_{N\wedge n}$ is uniformly integrable.

From the Theorem 2.1.1 we also immediately get:

Theorem 2.1.3.
If X_n is a uniformly integrable submartingale then for any stopping time $N \leq \infty$, we have $E(X_0) \leq E(X_N) \leq E(X_\infty)$, where $X_\infty = \lim X_n$.

Proof. Theorem 1.2.5 implies $E(X_0) \leq E(X_{N\wedge n}) \leq E(X_n)$.
Letting $n \to \infty$ and observing that Theorem 2.1.1 implies $X_{N\wedge n} \to X_N$ and $X_n \to X_\infty$ in L^1 gives the desired result.

□

Theorem 2.1.4 (Optional Stopping Theorem).
If $L \leq M$ are stopping times and $Y_{M\wedge n}$ is uniformly integrable submartingale, then $E(Y_L) \leq E(Y_M)$ and

$$Y_L \leq E(Y_M \mid \mathcal{F}_L).$$

Proof. Use the inequality $E(X_n) \leq E(X_\infty)$ in Theorem 2.1.3 with $X_n = Y_{M\wedge n}$ and $N = L$. To prove the second result, let $A \in \mathcal{F}_L$ and $N = \begin{cases} L, & on\ A \\ M, & on\ A^c \end{cases}$ is stopping time.
Using the first result now shows $E(Y_N) \leq E(Y_M)$. Since $N = M$ on A^c, it follows from the last inequality and the definition of conditional expectation that

$$E(Y_L; A) \leq E(Y_M; A) = E(E(Y_M \mid \mathcal{F}_L); A).$$

Taking $A_\varepsilon = \{Y_L - E(Y_M \mid \mathcal{F}_L) > \varepsilon\}$, we conclude $P(A_\varepsilon) = 0$ for all $\varepsilon > 0$ and the desired result follows.

□

It is worth considering the stopping theorem on martingales, explicating the conditions that make sure that we have $E(X_n) = E(X_0)$:

Theorem 2.1.5 (**Doob's optional stopping Theorem**).
Let T be a stopping time and X_n a martingale. Then X_T is integrable and $E(X_T)$=$E(X_0)$
if one of the following conditions holds:
(i) T is bounded,
(ii) T is almost surely finite and X is bounded,
(iii) $E(T) < \infty$ and there is $K > 0$ such that, $\mid X_n - X_{n-1} \mid \leq K$ for all n.

Proof. We assume that X_n is a supermaringale. Then $X_{T\wedge n}$ is a supermartingale by Theorem 1.2.4 and in particular, it is integrable,
and $E(X_{T\wedge n}) - X_0) \leq 0$.
For (i) $E(X_T) \leq E(X_0)$ follows by chosing n=N.
For (ii) $E(X_T) \leq E(X_0)$ letting $n \to \infty$ and use dominated convergence.
For (iii) we observe that $\mid X_{T\wedge n} - X_0 \mid = \mid \sum_{k=1}^{T\wedge n}(X_k - X_{k-1}) \mid \leq KT$ and we can use dominated convergence to have $E(X_T) \leq E(X_0)$.
Applying the previuos considerations to supermartingales $-X_n$ we have the statement.

□

`Example. The first run of three sixes.` We have a fair die throwing independently at each time step.
A gambler wins a fixed amount of money as soon as the first rum of three consecutive sixes appear. We want to know which is the mean number of throws until the gambler wins for the first time.
Let X_1,X_2,... be the sequence of random variables representing the outcomes of the throws.
We have $P(X_i = k)$=1/6 for every $k \in \{1, ..., 6\}$. Let $\mathcal{F}(n)$=$\sigma(X_1, ..., X_n)$ and T be the first time three consecutive sixes appear.
T is a stopping time and we are looking for $E(T)$.
Before each time n a gambler bets 1€that the nth throw will show six.
If he loses, he leaves, if he wins he receives 6€, all of which he bets on the event that $(n + 1)$st throw will be six again and so on if he loses he leaves and if he wins he will bet in the third throw and so forth.
T is a stopping time satisfying condition (iii) of Doob's optional stopping theorem, so we have $E(T)$=$6 + 6^2 + 6^3$=258 which is the expected money spent by the gamblers.
At time T the last gambler won 6€,the one before 36€and the one before 216€.All other gamblers have lost their post.

More formally, let be S_n=$(1 + 6 + ... + 6^k)$ the total stakes of all gamblers at time n if there is a run of k sixes, and let M_n=$S_n - n$, in particular M_0=1.

Then $\{M_n\}$ is a martingale, indeed
$E(M_{n+1}\mathcal{F}_n)$=$(5/6)(1-(n+1))+(1/6)(6S_n+1-(n+1))$=$S_n-n$=$M_n$.
Need to argue that $E(T)<\infty$:
Observing $T=k$ we need to have at least one number which is not six in every tuple $(X_{3m+1},X_{3m+2}),X_{3m+3}$ for $3m+3<k$ hence
$P(T=k)\leq(1-1/6^3)^{(k-3)/3}$=$(215/216)^{(k-3)/3}$ so $E(T)$=$\sum_{k=1}^{\infty}kP(T=k)$ converges.
We consider the stopped martingale M^T, then considering that $E(T)<\infty$ as we have seen and $\mid M_n^T-M_{n-1}^T\mid\leq 260$, by part (iii) of Doob's Optional Stopping Theorem, we have 1=$E(M_0)$=$E(M_T)$=$1+6+6^2+6^3-E(T)$.

2.1.1 The Secretary Problem

We consider a known number of items presented one by one in random order, i.e. such that all $n!$ possible orders being equally likely.
We can rank at any time the items that have so far been presented in order of usefulness. As each item is presented we must either accept it, in which case the process stops, or reject it, when the next item in the sequence is presented and we have to do the same choice as before.
Our aim is to maximize the probability that the item we choose is the best of the n items available.
Since we cannot never to go back and choose a previously-presented item which, in retrospect, turns out to be best, we clearly need to balance the danger of stopping too soon and accepting an item when an even better one might be still to come, and the danger of going on too long and finding that the best item was alreasy rejected.

`Example.` There are N candidates for a job interview. Let X_i be the ith candidate. The boss interviews each in turn and he must decide wheter to accept or reject the candidate, with not recall of an eventual rejected candidate.

Theorem 2.1.6.
Let X_i, $i\in 1,...,N$ be random variable uniformly distribuited on $[0,1]$, the stopping time T^*=inf$\{n>0:X_n>\alpha_n\}$ maximises $E(X_T)$, for α_N=0 and α_{n-1}=$1/2+\alpha_n^2/2$ for $1\leq n\leq N$.

Proof. We want to show that for any $0\leq\alpha\leq 1$, we have $E(X_n\vee\alpha)$=$1/2+\alpha^2/2$. It sufficies to notice that

$$E(X_n\vee\alpha)=\int_0^1 x\vee\alpha\,dx=\int_0^\alpha\alpha\,dx+\int_\alpha^1 x,dx=\alpha^2+1/2-\alpha^2/2=1/2+\alpha^2/2$$

Now, for any stopping time T, the process Y defined by

$$Y_0{=}\alpha_0, \text{ and } Y_n{=}(X_{T\wedge n}) \vee \alpha_n \text{ for } n \leq 1$$

Is a submartingale, indeed, on the event $\{T \leq n-1\}$, we have

$$E(Y_n \mid \mathcal{F}_{n-1}){=}E(X_{T\wedge n} \vee \alpha_n \mid \mathcal{F}_{n-1}){=}X_T \vee \alpha_n \leq X_T \vee \alpha_{n-1}{=}Y_{n-1}.$$

Using that α_n is decreasing. On the event $\{T \leq n-1\}$, we have

$$E(Y_n \mid \mathcal{F}_{n-1}){=}E(X_{T\wedge n} \vee \alpha_n){=}\ \alpha_n^2/2 + 1/2$$

which shows the supermartingale property.
Let' see that for $T = T^*$ the process Y is a martingale. Indeed, on the $\{T^* \leq n-1\}$, we have, from above,

$$E(Y_n \mid \mathcal{F}_{n-1}){=}X_{T^*} \vee \alpha_n{=}X_{T^*} \text{ as } X_{T^*} > \alpha_{T^*} \geq \alpha_{n-1} \geq \alpha_n.$$

Noticing that $Yn-1{=}X_{T^*} \vee \alpha_{n-1}{=}X_{T^*}$, on the event $T^* \geq n$ we have, as before,

$$E(Y_n \mid \mathcal{F}_{n-1}){=}E(X_{T\wedge n} \vee \alpha_n){=}\ \alpha_n^2/2 + 1/2{=}Y_{n-1}.$$

In the end we show that for any stopping time T, we have $E(X_T) \leq E(X_{T^*})$. For this we use Doob's optimal stopping theorem (noticing that all stopping times are bounded), to see that, for arbitrary stopping times

$$E(X_T) \leq E(X_T \vee \alpha_T){=}E(Y_T) \leq E(Y_0){=}\alpha_0$$

and, for the special choice T^*,

$$E(X_{T^*}){=}E(X_{T^*} \vee \alpha_{T^*}){=}E(Y_{T^*}) \leq E(Y_0){=}\alpha_0$$

and this completes the proof.

□

2.2 Burkholder's Sharp L^p Estimate for Martingale Tranfsorm

Before beginning our considerations about the Sharp inequality, we need to introduce the space, the σ-algebra and the probability that we are going to use.
Our space is $\Omega = (0,1]$, on which we can define subintervals such $(\frac{(j-1)}{2^n}, \frac{j}{2^n}]$ with $1 \leq j \leq 2^n$, the probability is $P(E) = \mid E \mid$ i.e. the Lebesgue-measure, so $P(((\frac{(j-1)}{2^n}, \frac{j}{2^n}]]){=}1/2^n$.
Let $\mathcal{F}_n{=}\sigma((\frac{(j-1)}{2^n}, \frac{j}{2^n}])$, $1 \leq j \leq 2^n$, the $\bigcup_n \mathcal{F}_n{=}\mathcal{F}$ which is Borel-σ-algebra.

Theorem 2.2.1.
Let X_n be a martingale in $(\Omega, \mathcal{F}_n, P)$ with the property that $\sup_n E \mid X_n \mid^p < +\infty$, $1 < p < +\infty$, *then X_n is uniformly integrable (i.e. there exist* $\lim_{n\to\infty} X_n = X$ *a.s. with* $X \in \mathcal{F}=\bigcup_n \mathcal{F}_n$ *and* $X_n=E(X \mid \mathcal{F}_n)$*).*

Theorem 2.2.2 (**Burkholder's Sharp Inequality**).
Let X_n be a martingale with the property that $\sup_n E \mid X_n \mid^p < +\infty$, $1 < p < +\infty$ *and let H_n be a predictable sequence such that* $\mid H_n \mid \leq 1$ *a.s.*
Let $d_n = X_n - X_{n-1}$, $n \geq 1$, *and define the martingale transform* $Y_n=(H \cdot X)_n=\sum_{m=1}^n H_m(X_m - X_{m-1})$.
Then:
Y_n is martingale in $(\Omega, \mathcal{F}_n, P)$, and

$$E \mid Y_n \mid^p \leq c(p)^p E \mid X_n \mid^p,$$

$$\text{with } c(p)=p^* - 1=\begin{cases} p-1, & \text{if } p \geq 2 \\ p'-1, & \text{if } 1 < p < 2, p' \text{ such that } 1/p + 1/p' = 1 \end{cases}$$

$$p^* = \max\{p, p/p-1\}.$$

Moreover by the Theorem above we have that there exist $Y=\lim_{n\to\infty} Y_n = Y$ *a.s. and then* $E \mid Y \mid^p | \leq c(p)^p E \mid X \mid^p$ *and the constant $c(p)$ is the best possible.*

Proof. $X : \Omega = (0, 1] \to \mathbb{R}$ is Borel-measurable, so if $X_n=E(X \mid \mathcal{F}_n)$ then $E(X \mid \mathcal{F}_n)^p=\int_0^1 \mid X(t) \mid^p \, dt < \infty$.
$E(X \mid \mathcal{F}_n)$ is a random variable which is constant in $I=((j-1)/2^n, j/2^n]$:

$$E(X \mid \mathcal{F}_n)(t) = \begin{cases} (1/\mid I \mid) \int_I \mid X(t) \mid \, dt = (1/P(I))E(X1_I), & \text{if } t \in I \\ 0, & \text{if } t \notin I \end{cases}$$

We prove that $E \mid Y \mid^p \leq c(p)^p E \mid X \mid^p$ by considering the function $V : \mathbb{R}^2 \to \mathbb{R}$ such that $V(x, y) =\mid y \mid^p -c(p)^p \mid x \mid^p$, the goal is to show that $E(V(X_n, Y_n)) \leq 0$.

To do that let introduce a the function $U : \mathbb{R}^2 \to \mathbb{R}$ such that:
(a) $U(x, y) \geq \mid y \mid^p -c(p)^p \mid X \mid^p$,
(b) $U(x, 0) \leq 0$ for all $x \in \mathbb{R}$,
(c) Taking a function $\mathbb{R} \to \mathbb{R}$ such that $t \mapsto U(x + th, y + tk)$ then this function is concave for all x, y, h, t such that $\mid k \mid \leq \mid h \mid$.

Let $U(x, y)=\alpha_p(\mid y \mid^p -(p^* - 1)) \mid x \mid)(\mid x \mid)(\mid x \mid + \mid y \mid)^{p-1}$, with $\alpha = p(1-1/p^*)^{p-1}$.

First, we show that (a),(b),(c) imply $E \mid Y_n \mid^p \leq c(p)^p E \mid X_n \mid^p$:

By (a) we have $E(\mid X_n \mid^p - (c(p) \mid Y_n \mid^p)) \leq E(U(X_n, Y_n))$ then to prove $E(X_n, Y_n) \leq 0$ it suffices to prove that $EU(X_n, Y_n) \leq 0$.
Let $I \in \mathcal{F}_{n-1}$ and $I_+, I_- \in \mathcal{F}_n$ its left,right values.

$$E(U(X_n, Y_n) \mid \mathcal{F}_{n-1})(I) = E(U(X_{n-1} + d_n, Y_{n-1} + d_n H_n) \mid \mathcal{F}_{n-1})(I) =$$

(Considering $h = d_n$ and $k = d_n H_n$ we have $\mid k \mid \leq \mid h \mid$)

by definition

$$=(1/2)U(X_{n-1}(I) + d_n(I_+), Y_{n-1}(I) + d_n(I_+)H_n(I)) + \\ +(1/2)U(X_{n-1}(I) + d_n(I_-), Y_{n-1}(I) + d_n(I_-)H_n(I)). \quad (**)$$

By (c) and definition of concavity we have $(\phi(a) + \phi(b))/2 \leq \phi((a+b)/2)$.
Let $\phi(t) = U(X_{n-1}(I) + t, Y_{n-1}(I)H_n(I)t)$ and, in particular, we have $(\phi(d(I_+)) + \phi(d(I_-))/2 \leq \phi((d(I_+) + d(I_-))/2) = \phi(0)$
i.e. we have $(\phi(d(I_+)) + \phi(d(I_-)) \leq \phi(0)$.
Then:

$$(**) = (1/2)U((X_{n-1}(I) + d_n(I_+), Y_{n-1}(I) + d_n(I_+)H_n(I)) + \\ +(X_{n-1}(I) + d_n(I_-), Y_{n-1}(I) + d_n(I_-)H_n(I))) \leq \\ \leq U(X_{n-1}(I) + \tfrac{d_n(I_+)+d_n(I_-)}{2}, Y_{n-1}(I) + \tfrac{d_n(I_+)+d_n(I_-)}{2} H_n) = \\ = U(X_{n-1}(I) + Y_{n-1}(I)).$$

Consequently $E(U(X_n, Y_n) \mid \mathcal{F}_{n-1})(I) \leq U(X_{n-1}(I), Y_{n-1}(I))$ and calculating E for both members we have $E(U(X_n, Y_n))(I) \leq E(U(X_{n-1}, Y_{n-1}))(I)$.

Repeating this recursively we'll obtain

$$E(U(X_n, Y_n))(I) \leq E(U(X_{n-1}, Y_{n-1}))(I) \leq \\ \leq E(U(X_{n-2}, Y_{n-2}))(I) \leq ... \leq E(U(X_0, Y_0))(I) = U(X_0, 0) \leq 0$$

by the property (b), concluding the first part of the proof.

Now we want to show that $U(x, y)$ satisfies (a),(b),(c) for $p \geq 2$ (the case $1 < p \leq 2$ is similar, even if not identical):

For (a):
Let $x, y > 0$ and $p > 2$ then $V(x, y) = y^p - (p-1)^p x^p$
and $U(x, y) = p(1-1/p)^{p-1}[y - (p-1)x][x+y]^{p-1}$, we need to prove $V(x, y) \leq U(x, y)$.

$U(x, y)$ and $V(x, y)$ are both homogeneous functions so we can use a change of variables:

$x + y = u, \quad y = v \cdot u, \quad x = (1 - v)u, \qquad$ with $u \geq 0$ and $0 \leq v < 1$,
then:
$V(x, y) = u^p[v^p - (p-1)^p(1-v)^p]$ and
$U(x, y) = u^p[p(1 - 1/p)^{p-1}(v - (p-1)(1-v))]$, we have:

$V(x, y) \leq U(x, y)$ iff $v^p - (p-1)^p(1-v)^p \leq p(1-1/p)^{p-1}(v - (p-1)(1-v))$.

Let $\phi(v) = v^p - (p-1)^p(1-v)^p$ and $\psi(v) = p(1-1/p)^{p-1}(v - (p-1)(1-v))$, which is a line for all $0 \leq v \leq 1$, we need to show for $0 < v_0 < v_1 < 1$:
i. $\psi(v_0) = \phi(v_0) = 0$,
ii. $\psi'(v_0) = \phi'(v_0)$,
iii. $\phi''(v) \leq 0$ in $[0, v_1]$ and $\phi''(v) \geq 0$ in $[v_1, 1]$,
iv. $\phi(1) < \psi(1)$.

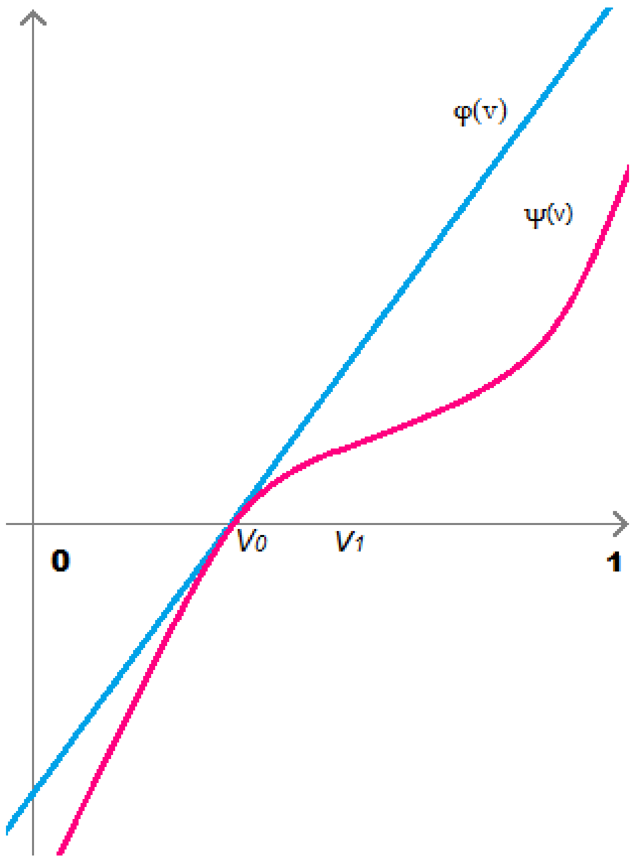

Figure 2.1:

Calculating the derivatives we obtain:
$\phi'(v)=p[v^{p-1} + (p-1)^p(1-v)^{p-1}]$,
$\phi''(v)=p(p-1)[v^{p-2} - (p-1)^p(1-v)^{p-2}]$.

To prove i. we notice that

$$\phi(v_0)=0 \text{ iff } \frac{v_0}{1-v_0}=p-1 \text{ iff } v_0=\frac{p-1}{p}=1/p' \geq 1/2.$$

Let just notice that $\psi(v_0) = 0$.

To prove ii. we have

$$\phi'(v_0)=p[(\tfrac{p-1}{p})^{p-1} + (p-1)^p(1-(\tfrac{p-1}{p})^{p-1})=$$
$$=p(p-1)^{p-1}\frac{1}{p^{p-1}}[1+(p-1)]=p^2(1-1/p)^{p-1}$$

which is the slope of ψ, so we have $\psi'(v_0) = \phi'(v_0)$.

To prove iii. we notice that

$$\phi''(v_1)=0 \text{ iff } \frac{v_1}{(1-v_1)}=(p-1)^{\frac{p}{p-2}}>p-1=\frac{v_0}{(1-v_0)}$$

because $p > p-2 > 0$ and then $v_1 > v_0$ because the function $\frac{S}{1-S}$ is an increasing function on positive quadrans.

To prove iv. we have that $\phi(1) = 1$ and $\psi(1)=p(1-1/p)^{p-1}=\frac{(p-1)^{p-1}}{p^{p-2}}$ then:

$$\phi(1) < \psi(1) \text{ iff } (p-1)^{p-1} > p^{p-2};$$

let prove that $(x-1)^\alpha - x^\alpha + \alpha x^{\alpha-1} \geq 0$ for all $x \geq 1$ and for all $\alpha \geq 1$,

$$(x-1)^\alpha - x^\alpha=\int_0^1 \frac{d}{dt}(x-t)^{\alpha-1}\,dt \geq \int_0^1(-\alpha)(x-t)^{\alpha-1}\,dt \geq$$
$$\geq \int_0^1(-\alpha)x^{\alpha-1}\,dt=(-\alpha)x^{\alpha-1},$$

by this property, for all $p \geq 2$ we have
$(p-1)^{p-1} \geq p^{p-1} - (p-1)p^{p-2}=P^{p-2}[p-(p-1)]=p^{p-2}$ i.e. we proved $(p-1)^{p-1} > p^{p-2}$ and so iv.

To prove this properties sufficies to prove the disequality.

For (b):
It is easy to see that $U(x,0)=p(1-1/p)^{p-1}(-(p-1)x)(x)^{p-1} \leq 0$ when $p \geq 2$.

For (c):
For all $x, y, h, k \in \mathbb{R}$, $x, y > 0$, $p \geq 2$ as before we have:
$U(x,y)=p(1-1/p)^{p-1}[y-(p-1)x][x+y]^{p-1}$ and moreover

$$\left\langle HessU(x,y)\binom{h}{k}, \binom{h}{k}\right\rangle = [U_{xx}(x,y)h]\cdot h + 2\,[U_{xy}(x,y)h]\cdot k + [U_{yy}(x,y)k]\cdot k$$

is the directional concavity in direction (h,k).

Calculating the second derivatives we'll obtain:
$U_{xx} = -p(p-1)[(p-1)x+y](x+y)^3$,
$U_{xy} = -p(p-1)(p-2)(x+y)^{p-3}x$,
$U_{yy} = p(p-1)(x+y)^{p-3}[y-(p-3)x]$,

then:

$$HessU(x,y) = [p(1-1/p)^{p-1}p(p-1)(x+y)^{p-3}]\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

with $a = -[(p-1)x+y]$, $b = -(p-2)x$, $c = y-(p-3)x$.

We have:

$$\begin{aligned}\left\langle HessU(x,y)\binom{h}{k}, \binom{h}{k}\right\rangle &= -[(p-1)x+y]h^2 - 2(p-2)xhk + [y-(p-1)x]k^2 = \\ &= (y+x)(K^2+h^2) - (p-2)x(h^2+2hk+k^2) = \\ &= (y+x)(K^2+h^2) - (p-2)x(h+k)^2 \le 0,\end{aligned}$$

if $|\, k \,|\le|\, h \,|$.

Let $G(t) = U(x+ht, y+kt)$, then:
$G''(t) = [U_{xx}(x(t),y(t))h]\cdot h + 2\,[U_{xy}(x(t),y(t))h]\cdot k + [U_{yy}(x(t),y(t))k]\cdot k$,

with $x(t) = x+ht$, $y(t) = y+kt$.

So $G''(t) \le 0$ whenever $|\, k \,|\le|\, h \,|$.

□

Bibliography

[1] Durrett, Rick Probability: theory and examples. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. x+428 pp. ISBN: 978-0-521-76539-8

[2] Bañuelos, Rodrigo The foundational inequalities of D. L. Burkholder and some of their ramifications. Illinois J. Math. 54 (2010), no. 3, 789-868 (2012).

[3] Peter Mörters, Lecture Notes on Martingale Theory. http://people.bath.ac.uk/maspm/martingales.pdf

[4] D. L. Burkholder, Boundary value problems and sharp inequalities for martingale transforms, Ann. Probab. 12 (1984), 647-702.

[5] D. L. Burkholder, Martingales and Fourier analysis in Banach spaces, C.I.M.E. Lectures (Varenna (Como), Italy, 1985), Lecture Notes in Mathematics, vol. 1206, Springer, Berlin, 1986, pp. 61-108.

[6] Nazarov, F. L.(1-MIS); Treĭl', S. R.(1-MIS) The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. (Russian. Russian summary) Algebra i Analiz 8 (1996), no. 5, 32–162; translation in St. Petersburg Math. J. 8 (1997), no. 5, 721-824.