Inequalities for Dependent Variables

Sergey Utev

School of Mathematical Sciences, University of Nottingham

June 2010

1. going back (1990)
2. moving closer (2000)
3. jumping to now (2010)
4. Back to the future (2000)

Banach space type 2

Banach space type 2

Banach space B is said to have type 2

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B \sum_{j=1}^{n} E|X_i|^2$$

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B \sum_{j=1}^{n} E|X_i|^2$$

A cotype 2 is for the reversed inequality to $\geq c'_B$.

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B \sum_{j=1}^{n} E|X_i|^2$$

A cotype 2 is for the reversed inequality to $\geq c'_B$.
E.g. L_p are type 2 for $p \geq 2$ and cotype 2 for $p \in [1, 2]$ and if B has type 2 and cotype 2 then it is topologically equivalent to a Hilbert space.

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B \sum_{j=1}^{n} E|X_i|^2$$

A cotype 2 is for the reversed inequality to $\geq c'_B$.
E.g. L_p are type 2 for $p \geq 2$ and cotype 2 for $p \in [1, 2]$ and if B has type 2 and cotype 2 then it is topologically equivalent to a Hilbert space.
Roughly, variables X_i are "artificially dependent" through the geometry of the Banach space B.

Banach space type 2

Banach space B is said to have type 2
if there exists a finite positive constant c_B such that for all n and independent variables X_i, $i = 1 \ldots, n$ with $E|X_i|^2 < \infty$, $EX_i = 0$,

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B \sum_{j=1}^{n} E|X_i|^2$$

A cotype 2 is for the reversed inequality to $\geq c'_B$.
E.g. L_p are type 2 for $p \geq 2$ and cotype 2 for $p \in [1, 2]$ and if B has type 2 and cotype 2 then it is topologically equivalent to a Hilbert space.
Roughly, variables X_i are "artificially dependent" through the geometry of the Banach space B.
Example. Let X_i be independent commutative self adjoint operators/random matrices and the norm is via some statistics of eigenvalues. Start with the spectral decomposition $X_i = U^* D_i U$ then D_i are in general dependent via U (or the spectral measure)

Banach space type 2 and mixing

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Do we need to introduce Banach spaces of "mixing" type 2?

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Do we need to introduce Banach spaces of "mixing" type 2?

ϕ-mixing type 2? or ρ-mixing type 2?

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Do we need to introduce Banach spaces of "mixing" type 2?

ϕ-mixing type 2? or ρ-mixing type 2? mixingale type 2?

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Do we need to introduce Banach spaces of "mixing" type 2?

ϕ-mixing type 2? or ρ-mixing type 2? mixingale type 2?

Actually, Banach spaces of dependent type (cotype) 2 are going back to Pisier (1975) (who introduced martingale type (cotype) 2 via so-called p-smoothness.

Banach space type 2 and mixing

What about if X_i satisfy some mixing conditions?

That is on top of the "artificial dependence" we add a true one?

In the above example $X_i = U^* D_i U$, D_i are now in general "doubly" dependent = via U and due to the dependence of X_i

Do we need to introduce Banach spaces of "mixing" type 2?

ϕ-mixing type 2? or ρ-mixing type 2? mixingale type 2?

Actually, Banach spaces of dependent type (cotype) 2 are going back to Pisier (1975) (who introduced martingale type (cotype) 2 via so-called p-smoothness.

Recently, Markov type 2 spaces (in a slightly different way) were introduced and found to be useful in the so-called extension problem (Naor, Peres etc)

Result (1990)

Result (1990)

It appears that we do NOT need an extra definition under the following ϕ-mixing condition

Result (1990)

It appears that we do NOT need an extra definition under the following ϕ-mixing condition

$$\phi := \sum_{j=1}^{\infty} \phi^{1/2}(2^j) < \infty\,, \quad \phi(n) = \text{usual } \phi\text{-mixing coefficient}$$

Result (1990)

It appears that we do NOT need an extra definition under the following ϕ-mixing condition

$$\phi := \sum_{j=1}^{\infty} \phi^{1/2}(2^j) < \infty \,, \quad \phi(n) = \text{usual } \phi\text{-mixing coefficient}$$

Th 1 (1990): There exists an absolute constant c such that

Result (1990)

It appears that we do NOT need an extra definition under the following ϕ-mixing condition

$$\phi := \sum_{j=1}^{\infty} \phi^{1/2}(2^j) < \infty\,, \quad \phi(n) = \text{usual } \phi\text{-mixing coefficient}$$

Th 1 (1990): There exists an absolute constant c such that

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \leq c_B e^{c\phi} \sum_{j=1}^{n} E|X_i|^2$$

Result (1990)

It appears that we do NOT need an extra definition under the following ϕ-mixing condition

$$\phi := \sum_{j=1}^{\infty} \phi^{1/2}(2^j) < \infty \,, \quad \phi(n) = \text{usual } \phi\text{-mixing coefficient}$$

Th 1 (1990): There exists an absolute constant c such that

$$E\Big|\sum_{j=1}^{n} X_i\Big|^2 \le c_B e^{c\phi} \sum_{j=1}^{n} E|X_i|^2$$

[Initially, I could do it only with moments of order $2+\delta$ which bothered me. Magda pointed me to a paper by Wlodek Bryc. Bernoulli congress - Upsalla, jaywalking with a peperoni pizza and long math discussions]

Sketch of the proof-(i) Max inequality

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i)

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1 \leq m \leq n} |S_m|^p)$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1 \leq m \leq n} |S_m|^p) \leq$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1 \leq m \leq n} |S_m|^p) \leq c \Big[$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\leq m\leq n} |S_m|^p) \leq c\Big[E\Big(\max_{1\leq k\leq n} |X_k|^p\Big)$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\leq m\leq n} |S_m|^p) \leq c\Big[E\Big(\max_{1\leq k\leq n} |X_k|^p\Big) +$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\le m\le n} |S_m|^p) \le c\Big[E\Big(\max_{1\le k\le n} |X_k|^p\Big) + \max_{1\le k\le n} (E|S_{k,n}|)^p\Big]$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\le m\le n}|S_m|^p) \le c\left[E\left(\max_{1\le k\le n}|X_k|^p\right) + \max_{1\le k\le n}(E|S_{k,n}|)^p\right]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\le m\le n}|S_m|^p) \le c\Big[E\Big(\max_{1\le k\le n}|X_k|^p\Big) + \max_{1\le k\le n}(E|S_{k,n}|)^p\Big]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\leq m\leq n} |S_m|^p) \leq c\Big[E\Big(\max_{1\leq k\leq n} |X_k|^p\Big) + \max_{1\leq k\leq n}(E|S_{k,n}|)^p\Big]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.
Let $M_k = \max_{1\leq m\leq k} |S_m|$, $A_k(x) = \{M_{k-1} \leq x, M_k > x\}$. Then

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\leq m\leq n}|S_m|^p) \leq c\Big[E\Big(\max_{1\leq k\leq n}|X_k|^p\Big) + \max_{1\leq k\leq n}(E|S_{k,n}|)^p\Big]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.
Let $M_k = \max_{1\leq m\leq k}|S_m|$, $A_k(x) = \{M_{k-1} \leq x, M_k > x\}$. Then

$$\{|S_n| \geq x + y + z)\} \subseteq \{\max_{1\leq k\leq n}|X_k| > y\} \cup \cup_{k=1}^n A_k(x) \cap \{|S_n - S_k| > z\}$$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E\left(\max_{1\le m\le n} |S_m|^p\right) \le c\left[E\left(\max_{1\le k\le n} |X_k|^p\right) + \max_{1\le k\le n} (E|S_{k,n}|)^p\right]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.
Let $M_k = \max_{1\le m\le k} |S_m|$, $A_k(x) = \{M_{k-1} \le x, M_k > x\}$. Then

$$\{|S_n| \ge x + y + z)\} \subseteq \{\max_{1\le k\le n} |X_k| > y\} \cup \cup_{k=1}^{n} A_k(x) \cap \{|S_n - S_k| > z\}$$

by ϕ-mixing applied to term $A_k(x) \cap \{|S_n - S_k| > z\}$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E\left(\max_{1\leq m\leq n} |S_m|^p\right) \leq c\left[E\left(\max_{1\leq k\leq n} |X_k|^p\right) + \max_{1\leq k\leq n} (E|S_{k,n}|)^p\right]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.
Let $M_k = \max_{1\leq m\leq k} |S_m|$, $A_k(x) = \{M_{k-1} \leq x, M_k > x\}$. Then

$$\{|S_n| \geq x + y + z)\} \subseteq \{\max_{1\leq k\leq n} |X_k| > y\} \cup \cup_{k=1}^{n} A_k(x) \cap \{|S_n - S_k| > z\}$$

by ϕ-mixing applied to term $A_k(x) \cap \{|S_n - S_k| > z\}$
$P(|S_n| \geq x + y + z) \leq P(\max_{1\leq m\leq n} |X_k| > y) + \eta P(M_n \geq x)$

Sketch of the proof-(i) Max inequality

The proof was based on two ingredients:
(i) The adaptation of Magda's Max inequality in the form (with $p = 2$)

$$E(\max_{1\leq m\leq n} |S_m|^p) \leq c\Big[E\Big(\max_{1\leq k\leq n} |X_k|^p\Big) + \max_{1\leq k\leq n}(E|S_{k,n}|)^p\Big]$$

where $S_n = X_1 + \ldots + X_n$, $S_{k,n} = X_k + \ldots + X_n$.
The argument was based on the following adaptation of Skorohod trick.
Let $M_k = \max_{1\leq m\leq k} |S_m|$, $A_k(x) = \{M_{k-1} \leq x, M_k > x\}$. Then

$$\{|S_n| \geq x + y + z)\} \subseteq \{\max_{1\leq k\leq n} |X_k| > y\} \cup \cup_{k=1}^{n} A_k(x) \cap \{|S_n - S_k| > z\}$$

by ϕ-mixing applied to term $A_k(x) \cap \{|S_n - S_k| > z\}$
$P(|S_n| \geq x + y + z) \leq P(\max_{1\leq m\leq n} |X_k| > y) + \eta P(M_n \geq x)$
where $\eta = \phi(1) + \max_{1\leq k\leq n} P(|S_n - S_k| > z)$ and then apply the Levy-Cohn argument to the max variable M_n.

Sketch of the proof-(ii) dyadic induction

Sketch of the proof-(ii) dyadic induction

(ii)

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c($$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3})$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\left[n + E\max_{1\leq k\leq n}|S_k|^2\right]$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\Big[n + E \max_{1\leq k\leq n} |S_k|^2\Big]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\Big[n + E \max_{1\leq k\leq n} |S_k|^2\Big]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\left[n + E \max_{1\leq k\leq n} |S_k|^2\right]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling $P(|X - Y| \geq t) \leq 2\phi(1)P(|Y| \geq t/2)$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\left[n + E\max_{1\leq k\leq n}|S_k|^2\right]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling
$P(|X - Y| \geq t) \leq 2\phi(1)P(|Y| \geq t/2)$
Then approximate $S_{n/2} \approx S'_{n/4} + S''_{n/4}$ and carry on until $h \leq n/2^k < 2h$ for some large but fixed h to get

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\left[n + E\max_{1\leq k\leq n}|S_k|^2\right]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling $P(|X - Y| \geq t) \leq 2\phi(1)P(|Y| \geq t/2)$
Then approximate $S_{n/2} \approx S'_{n/4} + S''_{n/4}$ and carry on until $h \leq n/2^k < 2h$ for some large but fixed h to get

$$E\|S_n - (S^{(1)}_{n/2^k} + \ldots + S^{(2^k)}_{n/2^k})\|^2$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\Big[n + E \max_{1\leq k\leq n} |S_k|^2\Big]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling
$P(|X - Y| \geq t) \leq 2\phi(1)P(|Y| \geq t/2)$
Then approximate $S_{n/2} \approx S'_{n/4} + S''_{n/4}$ and carry on until $h \leq n/2^k < 2h$ for some large but fixed h to get

$$E\|S_n - (S^{(1)}_{n/2^k} + \ldots + S^{(2^k)}_{n/2^k})\|^2 \leq \text{ error, summable by Gronwall lemma}$$

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \le c(\phi(n^{1/3}) + n^{-1/3})\left[n + E \max_{1\le k\le n} |S_k|^2\right]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling $P(|X - Y| \ge t) \le 2\phi(1)P(|Y| \ge t/2)$
Then approximate $S_{n/2} \approx S'_{n/4} + S''_{n/4}$ and carry on until $h \le n/2^k < 2h$ for some large but fixed h to get

$$E\|S_n - (S^{(1)}_{n/2^k} + \ldots + S^{(2^k)}_{n/2^k})\|^2 \le \text{ error, summable by Gronwall lemma}$$

where $S^{(i)}_{n/2^k}$ are independent, roughly partial sums of size h

Sketch of the proof-(ii) dyadic induction

(ii) The dyadic induction "small blocks" technique, i.e. to use the dyadic induction step in the form

$$E\|S_n - (S'_{n/2} + S''_{n/2})\|^2 \leq c(\phi(n^{1/3}) + n^{-1/3})\Big[n + E\max_{1\leq k\leq n}|S_k|^2\Big]$$

where variables $S'_{n/2}$ and $S''_{n/2}$ are independent and $S'_{n/2} =^d S_{n/2}$, $S''_{n/2} =^d S_n - S_{n/2}$ and where we applied the Bryc coupling
$P(|X-Y| \geq t) \leq 2\phi(1)P(|Y| \geq t/2)$
Then approximate $S_{n/2} \approx S'_{n/4} + S''_{n/4}$ and carry on until $h \leq n/2^k < 2h$ for some large but fixed h to get

$$E\|S_n - (S^{(1)}_{n/2^k} + \ldots + S^{(2^k)}_{n/2^k})\|^2 \leq \text{ error, summable by Gronwall lemma}$$

where $S^{(i)}_{n/2^k}$ are independent, roughly partial sums of size h
and so by B-type 2 condition $E|\text{sum}|^2 \leq c_h c_B \sum_{i=1}^n E|X_i|^2$

Open questions

Open questions

(Hinted during numerous discussions with Magda)

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?
The trouble is that lower bounds are not valid in general even for the real valued variables.

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?
The trouble is that lower bounds are not valid in general even for the real valued variables.
(ii) Can we replace the ϕ-mixing condition by the following ρ-mixing $\Sigma_1^\infty \rho(2^j) < \infty$?

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?
The trouble is that lower bounds are not valid in general even for the real valued variables.
(ii) Can we replace the ϕ-mixing condition by the following ρ-mixing $\Sigma_1^\infty \rho(2^j) < \infty$? or perhaps under stronger ρ-mixing rates?

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?
The trouble is that lower bounds are not valid in general even for the real valued variables.
(ii) Can we replace the ϕ-mixing condition by the following ρ-mixing $\Sigma_1^\infty \rho(2^j) < \infty$? or perhaps under stronger ρ-mixing rates?
Some partial results can be derived for L_p using the moment inequality under stronger mixing rates. Inequalities can be used to extend/improve asymptotic results for Banach valued statistics of dependent variables (Bingham, Bosq, Dehling, Gotze, Merlevede etc)

Open questions

(Hinted during numerous discussions with Magda)

(i) The Cotype version under good phi mixing conditions?
The trouble is that lower bounds are not valid in general even for the real valued variables.
(ii) Can we replace the ϕ-mixing condition by the following ρ-mixing $\Sigma_1^\infty \rho(2^j) < \infty$? or perhaps under stronger ρ-mixing rates?
Some partial results can be derived for L_p using the moment inequality under stronger mixing rates. Inequalities can be used to extend/improve asymptotic results for Banach valued statistics of dependent variables (Bingham, Bosq, Dehling, Gotze, Merlevede etc)
Alternatively, it is probably easier to work with 2-smooth Banach spaces, i,e. such that $|x+y|^2 + |x-y|^2 \leq |x|^2 + D|y|^2$ for all x, y , and projective criteria, which gives for the stationary case ($P_j(X_1)$=projection operator)
$E|a_1X_1 + \ldots + a_nX_n|^2 \leq cD(\Sigma_1^n a_j^2)(\Sigma_j \|P_j(X_1)\|)^2$

ρ^*-mixing and maximal inequalities

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing?

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing?
More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing?
More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$a_n :=$

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$$a_n := \sup_k$$

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$$a_n := \sup_k \max_{0 \leq a \leq n}$$

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$$a_n := \sup_k \max_{0\leq a\leq n} E[\max_{0\leq j\leq a} |X_k + \ldots + X_{k+j}|^2]$$

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$$a_n := \sup_k \max_{0 \leq a \leq n} E[\max_{0 \leq j \leq a} |X_k + \ldots + X_{k+j}|^2]/$$

ρ^*-mixing and maximal inequalities

How to prove the Kolmogorov maximal inequality under ρ^*-mixing? More exactly, now let $\{X_k\}$ be a ρ^*-mixing sequence that is

$$\rho_n^* = \sup_{Q,T}\{\rho(\mathcal{F}_Q, \mathcal{F}_T)\} \to 0$$

where the sup is taken over all subsets $Q, T \in Z$ such that $|x - y| \geq n$ for all $x \in Q$ and $y \in T$.

Magda come up with the fascinating argument using the invariance principle! but unfortunately it needed stationarity.

The other approach was to use a so-called extremality trick that is to bound

$$a_n := \sup_k \max_{0 \leq a \leq n} E[\max_{0 \leq j \leq a} |X_k + \ldots + X_{k+j}|^2]/(\|X_k\|^2 + \ldots + \|X_{k+j}\|^2)$$

max inequality under ρ-mixing

max inequality under ρ-mixing

How it worked under ρ-mixing?

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left|\sum_{i=t+1}^{u} D_i\right| \leq \max_{t\leq j\leq u} |S_j - S_t|$$

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \le 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left| \sum_{i=t+1}^{u} D_i \right| \le \max_{t \le j \le u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \le a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where } \Big|\sum_{i=t+1}^{u} D_i\Big| \leq \max_{t\leq j\leq u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \leq a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

where $T(n) \leq (n/2)(1 + n^{-1/3})$ by a combinatoric argument.

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left| \sum_{i=t+1}^{u} D_i \right| \leq \max_{t \leq j \leq u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \leq a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

where $T(n) \leq (n/2)(1 + n^{-1/3})$ by a combinatoric argument.
Th(New result)

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left| \sum_{i=t+1}^{u} D_i \right| \leq \max_{t \leq j \leq u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \leq a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

where $T(n) \leq (n/2)(1 + n^{-1/3})$ by a combinatoric argument.
Th(New result) Assume that $\rho = \Sigma \rho(2^n) < \infty$. Then

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left| \sum_{i=t+1}^{u} D_i \right| \leq \max_{t \leq j \leq u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \leq a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

where $T(n) \leq (n/2)(1 + n^{-1/3})$ by a combinatoric argument.
Th(New result) Assume that $\rho = \Sigma\rho(2^n) < \infty$. Then

$$E \max_{1 \leq j \leq n} |X_1 + \ldots + X_j|^2 \leq e^{c\rho}(E[X_1^2] + \ldots + E[X_n^2])$$

max inequality under ρ-mixing

How it worked under ρ-mixing?

Several ways to go. Most common one is to use the following Garsia trick

$$M_n^2 \leq 4S_n^2 + \sum_{j=1}^{n} D_j(S_n - S_j) \quad \text{where} \left| \sum_{i=t+1}^{u} D_i \right| \leq \max_{t \leq j \leq u} |S_j - S_t|$$

and use for example, the dyadic induction together with the extremality argument to derive

$$a_n \leq a_{T(n)}(1 + c[\rho(n^{1/3}) + n^{-1/9}])$$

where $T(n) \leq (n/2)(1 + n^{-1/3})$ by a combinatoric argument.
Th(New result) Assume that $\rho = \Sigma\rho(2^n) < \infty$. Then

$$E \max_{1 \leq j \leq n} |X_1 + \ldots + X_j|^2 \leq e^{c\rho}(E[X_1^2] + \ldots + E[X_n^2])$$

However, it does require the extra rate $\Sigma\rho(2^n) < \infty$.

ρ^* - mixing case

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p$$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c$$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c \Big[\Big(E \max_{1 \le k \le n} |S_k| \Big)^p$$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c\Big[\Big(E \max_{1 \le k \le n} |S_k|\Big)^p + \sum_{j=1}^{n} E|X_j|^p + \Big(\sum_{j=1}^{n} E|X_j|^2\Big)^{p/2}\Big]$$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c \left[\left(E \max_{1 \le k \le n} |S_k| \right)^p + \sum_{j=1}^{n} E|X_j|^p + \left(\sum_{j=1}^{n} E|X_j|^2 \right)^{p/2} \right]$$

Start with X_i, with zero means and $\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = 1$ and write $X_i = X_{i,<M} + X_{i,>M}$ (truncate and centralize).

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c\left[\left(E \max_{1 \le k \le n} |S_k|\right)^p + \sum_{j=1}^{n} E|X_j|^p + \left(\sum_{j=1}^{n} E|X_j|^2\right)^{p/2}\right]$$

Start with X_i, with zero means and $\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = 1$ and write $X_i = X_{i,<M} + X_{i,>M}$ (truncate and centralize).

Tails are bounded by Chebyshev argument $\Sigma_1^n E|X_{i,>M}| \le 2M^{1/2}$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c\left[\left(E \max_{1 \le k \le n} |S_k|\right)^p + \sum_{j=1}^n E|X_j|^p + \left(\sum_{j=1}^n E|X_j|^2\right)^{p/2}\right]$$

Start with X_i, with zero means and $\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = 1$ and write $X_i = X_{i,<M} + X_{i,>M}$ (truncate and centralize).

Tails are bounded by Chebyshev argument $\Sigma_1^n E|X_{i,>M}| \le 2M^{1/2}$

And for the variables $X_{i,<M}$ (bounded by $2M$) we use the blocking procedure via the non-random stopping times

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1\leq k\leq n} |S_k|^p \leq c\Big[\Big(E \max_{1\leq k\leq n} |S_k|\Big)^p + \sum_{j=1}^n E|X_j|^p + \Big(\sum_{j=1}^n E|X_j|^2\Big)^{p/2}\Big]$$

Start with X_i, with zero means and $\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = 1$ and write $X_i = X_{i,<M} + X_{i,>M}$ (truncate and centralize).

Tails are bounded by Chebyshev argument $\Sigma_1^n E|X_{i,>M}| \leq 2M^{1/2}$

And for the variables $X_{i,<M}$ (bounded by $2M$) we use the blocking procedure via the non-random stopping times

$$m_k = \min\{m : m > m_{k-1}, \sum_{j=m_{k-1}+1}^{m} \mathrm{Var}(X_{i,<M}) > 1/M\}$$

ρ^* - mixing case

We apply the Bryc–Smolenskii–Peligrad– Gut (with $p = 2$)

$$E \max_{1 \le k \le n} |S_k|^p \le c\left[\left(E \max_{1 \le k \le n} |S_k|\right)^p + \sum_{j=1}^n E|X_j|^p + \left(\sum_{j=1}^n E|X_j|^2\right)^{p/2}\right]$$

Start with X_i, with zero means and $\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = 1$ and write $X_i = X_{i,<M} + X_{i,>M}$ (truncate and centralize).

Tails are bounded by Chebyshev argument $\Sigma_1^n E|X_{i,>M}| \le 2M^{1/2}$

And for the variables $X_{i,<M}$ (bounded by $2M$) we use the blocking procedure via the non-random stopping times

$$m_k = \min\{m : m > m_{k-1}, \sum_{j=m_{k-1}+1}^{m} \mathrm{Var}(X_{i,<M}) > 1/M\}$$

and the blocks $X_{m_{k-1}+1,<M} + \ldots + X_{m_k,<M}$ admit handy moment bounds.

Probabilistic approach (in progress)

Probabilistic approach (in progress)

Lemma (New inequality)

Probabilistic approach (in progress)

Lemma (New inequality)

$$P(f(X) > x + 6y) \leq P(f(X) > x)[\rho_1^* + P(f(X) > x)] + 7P(K\|\varepsilon X\| > y)$$

Probabilistic approach (in progress)

Lemma (New inequality)

$$P(f(X) > x + 6y) \leq P(f(X) > x)[\rho_1^* + P(f(X) > x)] + 7P(K\|\varepsilon X\| > y)$$

where $X = (X_1, \ldots, X_n)$, $\varepsilon X = (\varepsilon_1 X_1, \ldots, \varepsilon_n X_n)$ and ε_i are iid Rademacher independent of X, $\|\cdot\|$ is a seminorm and f is the coordinatewise nondecreasing with the Lipschitz coefficient K, i.e. $|f(x) - f(y)| \leq K\|x - y\|$

Probabilistic approach (in progress)

Lemma (New inequality)

$$P(f(X) > x + 6y) \leq P(f(X) > x)[\rho_1^* + P(f(X) > x)] + 7P(K\|\varepsilon X\| > y)$$

where $X = (X_1, \ldots, X_n)$, $\varepsilon X = (\varepsilon_1 X_1, \ldots, \varepsilon_n X_n)$ and ε_i are iid Rademacher independent of X, $\|\cdot\|$ is a seminorm and f is the coordinatewise nondecreasing with the Lipschitz coefficient K, i.e. $|f(x) - f(y)| \leq K\|x - y\|$
Actually, it can be applied not only to derive the moment inequalities but also the maximal inequalities by treating the max seminorm.

References

(1) Dedecker, J., Merlevede, Fl., Peligrad, M. and Utev, S. (2009). Moderate deviations for stationary sequences of bounded random variables. *Ann. Inst. Henri Poincare Probab. Stat.***45** , no. 2, 453–476.

(2) Peligrad, M., Utev, S. and Wu, W. B. (2007). A maximal $\mathbb{L}_p$-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135**, no. 2, 541–550.

(3) Peligrad, M. and Utev, S. (2006). Invariance principle for stochastic processes with short memory. *In High dimensional probability* , 18–32, IMS Lecture Notes Monogr. Ser., 51, Inst. Math. Statist., Beachwood, OH, 2006.

(4) Peligrad, M. and Utev, S. (2006). Central limit theorem for stationary linear processes. *Ann. Probab.* **34**, no. 4, 1608–1622.

(5) Merlevede, Fl., Peligrad, M. and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** 1–36.

(6) Peligrad, M. and Utev, S. (2006). Another approach to Brownian motion. *Stochastic Process. Appl.* **116 (2006)**, no. 2, 279–292.

(7) Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33**, no. 2, 798–815.

(8) Utev, S. and Peligrad, M. (2003). Maximal inequalities and an invariance principle for a class of weakly dependent random variables. *J. Theoret. Probab.* **16** , no. 1, 101–115.

(9) Merlevede, Fl., Peligrad, M. and Utev, S. (1997). Sharp conditions for the CLT of linear processes in a Hilbert space. *J. Theoret. Probab.* **10**, no. 3, 681–693.

(10) Peligrad, M. and Utev, S. (1997). Central limit theorem for linear processes. *Ann. Probab.* **25**, no. 1, 443–456.