Weak dependence in stochastic programming

Michal Houda

Faculty of Mathematics and Physics, Charles University in Prague

May 17, 2007

I.

Coefficients of weak dependence and convergence of integrated empirical process

Weak dependence
Notation

- $(\Omega, \mathcal{A}, \mathbb{P})$ ... probability space
- $(X, \mathbb{B})$ ... measurable space (value space)
- $\{\xi_t\}_{-\infty}^{+\infty}$ ... X-valued stochastic process with discrete or continuous time
- $\mathcal{B}_a^b$ ... σ-algebra generated by events

$$\{\xi_{t_1} \in A_{t_1}, \dots, \xi_{t_n} \in A_{t_n}\}$$

  where $(a \leq)\ t_1 \leq \cdots \leq t_n\ (\leq b)$, n are arbitrary, $A_{t_1}, \dots, A_{t_n}$ are $\mathcal{B}$-measurable sets
- $\mathcal{B}_1, B_2$ ... two arbitrary σ-algebras of subsets of Ω

Definition

The process $\{\xi_t\}$ is *m*-dependent if $\mathcal{B}^{a}_{-\infty}$ and $\mathcal{B}^{+\infty}_{b}$ are independent when $b - a > m$

Example: moving average process MA(m) is $(m+1)$-dependent (but not *m*-dependent)

Strong mixing (α-mixing) coefficient

$$\alpha(\mathcal{B}_1, \mathcal{B}_2) = \sup_{A \in \mathcal{B}_1, B \in \mathcal{B}_2} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$$

$$\alpha(t) = \sup_s \alpha(\mathcal{B}^s_{-\infty}, \mathcal{B}^{+\infty}_{s+t})$$

Definition

The process $\{\xi_t\}$ is α-mixing if $\alpha(t) \to 0$ as $t \to +\infty$

- α-coefficient measures direct covariance dependence
- range: $\alpha(\mathcal{B}_1, \mathcal{B}_2) \leq 1/4$

Example: autoregressive process AR(m) with normal increments is strong mixing (but not with binomial increments)

Absolute regularity (β-mixing) coefficient

$$\beta(\mathcal{B}_1, \mathcal{B}_2) = \mathbb{E} \operatorname*{esssup}_{B \in \mathcal{B}_2} |\mathbb{P}(B|\mathcal{B}_1) - \mathbb{P}(B)|$$

$$\beta(t) = \sup_s \beta(\mathcal{B}^s_{-\infty}, \mathcal{B}^{+\infty}_{s+t})$$

Definition

The process $\{\xi_t\}$ is β-mixing if $\beta(t) \to 0$ as $t \to +\infty$

- $\beta(\mathcal{B}_1, \mathcal{B}_2) = \sup_{\{A_i\},\{B_i\}} \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|$

  $\{A_i\} \subset \mathcal{B}_1, \{B_i\} \subset \mathcal{B}_2$ are partitions of Ω

*-mixing (φ-mixing) coefficient

$$\varphi(\mathcal{B}_1, \mathcal{B}_2) = \sup_{A \in \mathcal{B}_1, B \in \mathcal{B}_2} \left| \frac{\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)\mathbb{P}(B)} \right|$$

$$\varphi(t) = \sup_s \varphi(\mathcal{B}^s_{-\infty}, \mathcal{B}^{+\infty}_{s+t})$$

Definition

The process $\{\xi_t\}$ is φ-mixing if $\varphi(t) \to 0$ as $t \to +\infty$

Uniform mixing (ϕ-mixing) coefficient

$$\phi(\mathcal{B}_1, \mathcal{B}_2) = \sup_{A \in \mathcal{B}_1, B \in \mathcal{B}_2} \left| \frac{\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)} \right|$$

$$\phi(t) = \sup_s \phi(\mathcal{B}_{-\infty}^s, \mathcal{B}_{s+t}^{+\infty})$$

Definition

The process $\{\xi_t\}$ is ϕ-mixing if $\phi(t) \to 0$ as $t \to +\infty$

Complete regularity (ρ-mixing) coefficient

$$\rho(\mathcal{B}_1, \mathcal{B}_2) = \sup_{\eta_1, \eta_2} \left| \frac{\mathbb{E}\eta_1\eta_2 - \mathbb{E}\eta_1\mathbb{E}\eta_2}{\sqrt{\mathrm{var}\,\eta_1\,\mathrm{var}\,\eta_2}} \right|$$

$$\phi(t) = \sup_a \rho(\{\xi_s, s \leq a\}, \{\xi_s, s \geq a + t\})$$

η_1, η_2 are $\mathcal{B}_1$-, $\mathcal{B}_2$-measurable random variables

Definition

The process $\{\xi_t\}$ is ϕ-mixing if $\phi(t) \to 0$ as $t \to +\infty$

Relationships among mixing conditions

General relationships:

$$\varphi \;\Rightarrow\; \phi \;\Rightarrow\; \begin{array}{c}\beta \\ \\ \rho\end{array} \;\Rightarrow\; \alpha$$

Strictly stationnary Gaussian sequences:

$$\begin{array}{c}m\text{-dep.} \\ \Updownarrow \\ \varphi \\ \Updownarrow \\ \phi\end{array} \;\Rightarrow\; \beta \;\Rightarrow\; \begin{array}{c}\rho \\ \Updownarrow \\ \alpha\end{array}$$

Relationships among mixing conditions

Various limiting theorems remain valid with weakly dependent sequences. Example of CLT:

Theorem (Mori, Yoshihara (1986))

Let

- $\{\xi_i\}$ *. . . strong mixing sequence with* $\alpha(n)$
- $\mathbb{E}\xi_1 = 0$, $\mathbb{E}\xi_1^2 < +\infty$
- $S_0 = 0$, $S_n = \sum_{j=1}^{n} \xi_j$
- $s_n^2 = ES_n^2$

Then

$$\frac{S_n}{s_n} \longrightarrow^d N(0;1)$$

iif $\left\{(\frac{S_n}{s_n})^2\right\}_{n=1}^{+\infty}$ *is uniformly integrable, i.e.,*

$$\lim_{a\to+\infty} \sup_{n\geq 1} \int_{|\frac{S_n}{s_n}|>a} \frac{S_n^2}{s_n^2} \mathrm{d}\mathbb{P} = 0$$

$$\sqrt{N}\, W(\mu_N, \mu) = \int_{-\infty}^{+\infty} \sqrt{N}\, |F_N(t) - F(t)|\, \mathrm{d}t \tag{1}$$

- $F_N(t) = \frac{1}{N} \sum_{i=1}^{N} I_{(-\infty;t]}(\xi_i), \quad t \in \mathbb{R}$ ... empirical distribution function
- μ_N ... corresponding probability measure
- μ ... probability measure with finite first moment and distribution function F
- $\xi_1, \ldots, \xi_N$ ... iid sample from μ

Classical result for μ uniform distribution on $[0; 1]$:

$$\int_0^1 \sqrt{N} \left| \frac{1}{N} \sum_{i=1}^{N} I_{(0;t]}(\xi_i) - F(t) \right| \mathrm{d}t \to_d \int_0^1 |\mathbb{U}(t)| \mathrm{d}t \tag{2}$$

Distribution of RHS is known explicitly in this case.
Shorack, Wellner (1986)

General distribution

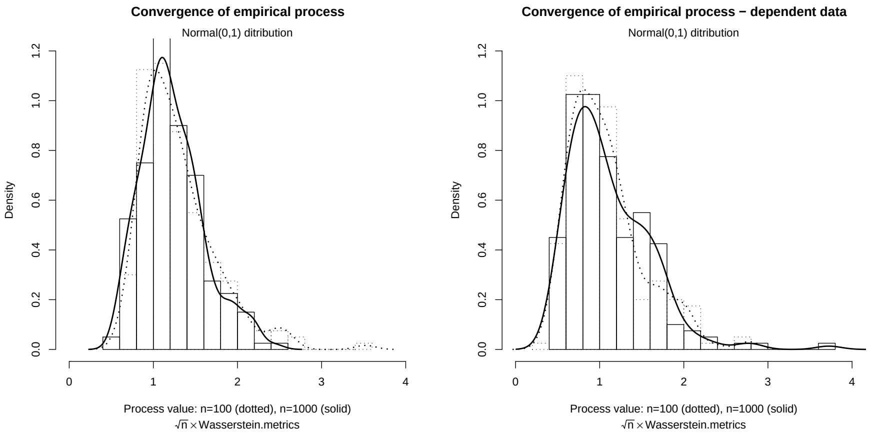
$$\int_{-\infty}^{+\infty} \sqrt{N} \left| \frac{1}{N} \sum_{i=1}^{N} I_{(-\infty;t]}(\xi_i) - F(t) \right| \mathrm{d}t \to_d \int_{-\infty}^{+\infty} |\mathbb{U}(F(t))| \mathrm{d}t. \qquad (3)$$

Del Barrio, Giné, Matrán (1999): (3) is valid if (and only if)

$$\int_{-\infty}^{+\infty} \sqrt{F(t)(1 - F(t))} \mathrm{d}t < +\infty$$

(In fact: if some processes Y_N converge weakly in $L_1(\mathbb{R})$ to Y, then, among others, $||Y_N||_{L_1} \to_d ||Y||_{L_1}$ where $||g||_{L_1} = \int_{-\infty}^{\infty} g(t)\mathrm{d}t$ for each non-negative $g \in L_1(\mathbb{R})$.)

- idea: convergence is proved for iid data, but some weak dependence property would not make difficulties (as CLT is valid with weak dependence)
- illustration: on simple MA(1) process $\xi_k := 0.5\zeta_k + 0.5\zeta_{k-1}$ with normal distribution – comparison of independent and weakly dependent and samples
- still to do: AR process (α-mixing for some class of continuous distributions)
- still to do: modify the proof of Del Bario et al. (1999) – involving the appropriate condition from theory of weakly dependent sequences



Convergence of empirical process
Normal(0,1) ditribution
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
1
2
3
4
Process value: n=100 (dotted), n=1000 (solid)
√n × Wasserstein.metrics
Convergence of empirical process − dependent data
Normal(0,1) ditribution
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
1
2
3
4
Process value: n=100 (dotted), n=1000 (solid)
√n × Wasserstein.metrics

II.

Convexity of chance-constrained programs – independent and dependent rows

$$\min F_0(x) \quad \text{subject to} \quad \mathbb{P}\big(h(x;\xi) \geq 0)\big) \geq p \tag{4}$$

- $x \in \mathbb{R}^m$ ... decision vector
- $\xi : \Omega \to \mathbb{R}^s$ ... s-dim. random vector defined on $(\Omega, \mathcal{A}, \mathbb{P})$
- $h : \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}^d$ ... vector-valued mapping
- $p \in [0;1]$ ... (prescribed) probability level

Denote

- $\mu = \mathbb{P} \circ \xi^{-1}$ ... distribution of ξ
- $F = F_\mu$ ... distribution function of ξ
- $H(x) = \{\xi \in \mathbb{R}^m : h(x;\xi) \geq 0)\}$
- $M(p) = \{x \in \mathbb{R}^m : \mathbb{P}(H(x)) = \mu(H(x)) \geq p\}$ ... set of feasible decisions

$$\min F_0(x) \quad \text{subject to} \quad \int_{\mathbb{R}^s} p - \chi_{H(x)}(\xi)\ \mu(\mathrm{d}\xi) \le 0 \tag{5}$$

- form adapted to the general stability theorem – Henrion, Römisch (1999)

$$\min F_0(x) \quad \text{subject to} \quad \mathbb{P}(h(x;\xi) < 0) \le \varepsilon \tag{6}$$

- $\varepsilon = 1 - p$ … (admissible) level of violation of the constraints
- $M(1-\varepsilon)$ … set of ε-feasible solutions (used in robust programming)

Convexity of chance constraints: known results

PRÉKOPA (1995)

Key question

When the set $M(p)$ of feasible solutions is convex?

1. Trivial result

   If $h(\cdot, \xi)$ is convex for all ξ, $M(0)$, $M(1)$ are convex
2. $M(p)$ is convex if
   - μ is a log-concave (or r-concave for $r \geq -1/s$) measure (implied by a log-concave, or $\frac{r}{1-rs}$-concave density)
   - components of h are quasi-concave (in both variables)

Parameterization of the concavity

Definition

$f : \mathbb{R}^d \to (0; +\infty)$ is *r-concave* for $r \in [-\infty; +\infty]$ if

$$f(\lambda x + (1-\lambda)y) \geq [\lambda f^r(x) + (1-\lambda) f^r(y)]^{1/r}$$

- cases $r = -\infty, 0, +\infty$ by continuity
$r = +\infty$	...	$\mathrm{RHS} = \max\{f(x), f(y)\}$
$r \in (1; +\infty)$	...	f^r is concave
$r = 1$	...	f is concave
$r = 0$	...	f is *log-concave* ($\log f$ is concave): $f(\lambda x + (1-\lambda)y) \geq f^{\lambda}(x) f^{1-\lambda}(y)$
$r < 0$	...	f^r is convex
$r = -\infty$	...	f is *quasi-concave*: $\mathrm{RHS} = \min\{f(x), f(y)\}$
- for all $r \leq r^*$, f is r-concave if it is r^*-concave
- interesting cases: $r \leq 1$

Special case: random RHS

$$\min F_0(x) \quad \text{subject to} \quad \mathbb{P}\big(g(x) \geq \xi)\big) \geq p \qquad (7)$$

- $h(x;\xi) = g(x) - \xi$
- $M(p) = \{x \in \mathbb{R}^n : F(g(x)) \geq p\}$

Required condition: components of h are quasi-concave
Problem: quasi-concavity is not preserved under addition
$\Rightarrow$ we require $g(x)$ to be convex

Idea of Henrion, Strugarek (2006):

- relax concavity condition of g;
- make more stringent concavity condition on μ.

Parameterization of the function decrease

Definition

$f : \mathbb{R} \to \mathbb{R}$ is *r-decreasing* for $r \in \mathbb{R}$ if

- it is continuous on $(0; +\infty)$, and
- there exists a threshold $t^* > 0$ such that $t^r f(t)$ is *strictly decreasing* for all $t > t^*$

- $r = 0$ ... strictly decreasing (in the classical sense)
- for all $r \leq r^*$, nonnegative f is r-decreasing if it is r^*-decreasing
- key property for marginal densities from the chance-constrained problem
- lemma: if F is distribution function with $(r+1)$-decreasing density, then $z \mapsto F(z^{-1/r})$ is concave on $(0; (t^*)^{-r})$

Convexity of the problem with RHS
Henrion, Strugarek (2006)

Theorem (Henrion, Strugarek (2006))

If

1. g_i *are* $-r_i$*-concave,*
2. ξ_i *have* $(r_i + 1)$*-decreasing densities,*
3. ξ_i *are independent,*

then $M(p)$ *is convex for* $p \geq \max F_i(t_i^*)$

Notation ($i = 1, \ldots, s$):

g_i ... components of g
ξ_i ... components of ξ
F_i ... components of F
t_i^* ... threshold of $r_i + 1$-decreasing density