

Estimation et Contrôle Adaptatif pour les Modèles de Régression

Bernard Bercu

Séminaire de Statistique et Santé Publique

21 Novembre 2006, Université Bordeaux 2



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



On considère le modèle de régression

$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



On considère le modèle de régression

$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



On considère le modèle de régression

$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



On considère le modèle de régression

$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



On considère le modèle de régression

$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



On considère le modèle de régression

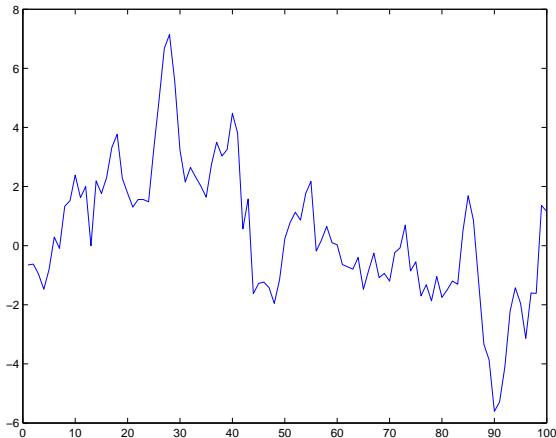
$$X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}$$

- X_n → L'observation,
- Φ_n → Le vecteur de régression,
- U_n → Le contrôle dont on a le choix,
- ε_n → Le bruit.

On se fixe deux objectifs à atteindre simultanément

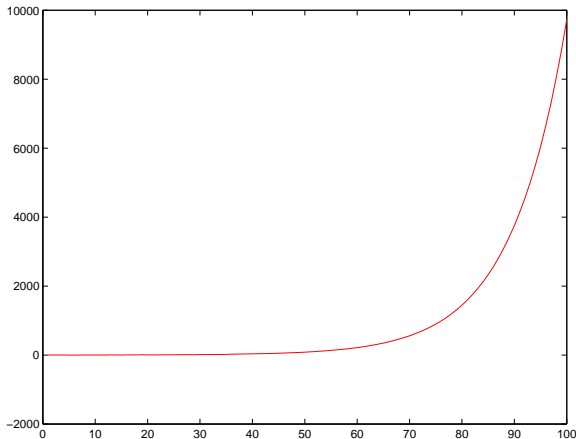
- Estimer le paramètre inconnu θ ,
- Contrôler la dynamique du processus (X_n).



Simulation autorégressif stable $|\theta| < 1$ 

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}$$

Simulation autorégressif explosif $|\theta| > 1$



$$X_{n+1} = \theta X_n + \varepsilon_{n+1}$$



Applications Industrielles

- **Aéronautique** → Contrôle de trajectoires d'aéronefs,
- **Biochimie** → Contrôle de fermenteurs,
- **Biostatistique** → Contrôle de dynamique de populations.



Applications Industrielles

- **Aéronautique** → Contrôle de trajectoires d'aéronefs,
- **Biochimie** → **Contrôle de fermenteurs,**
- **Biostatistique** → Contrôle de dynamique de populations.



Applications Industrielles

- **Aéronautique** → Contrôle de trajectoires d'aéronefs,
- **Biochimie** → Contrôle de fermenteurs,
- **Biostatistique** → Contrôle de dynamique de populations.



Plan

- 1 Introduction
 - Objectifs
 - **Algorithmes d'estimation**
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



On estime θ par l'estimateur qui minimise

$$\Delta_n(\theta) = \frac{1}{2} \sum_{k=0}^{n-1} a_k (X_{k+1} - U_k - \theta^t \Phi_k)^2.$$

On a donc

$$\hat{\theta}_n = \mathbf{S}_{n-1}^{-1}(\mathbf{a}) \sum_{k=0}^{n-1} a_k \Phi_k (X_{k+1} - U_k),$$

$$\mathbf{S}_n(\mathbf{a}) = \sum_{k=0}^n a_k \Phi_k \Phi_k^t.$$

- Pour l'algorithme des **moindres carrés ordinaires**, on a

$$\mathbf{a}_n = \mathbf{1}.$$



Moindres carrés pondérés

- Pour l'algorithme des **moindres carrés pondérés**, on a

$$a_n = \left(\frac{1}{\log s_n} \right)^{1+\gamma} \quad \text{avec} \quad s_n = \sum_{k=0}^n \|\Phi_k\|^2$$

et $\gamma > 0$. On a toujours la décomposition

$$\hat{\theta}_n - \theta = \mathbf{S}_{n-1}^{-1}(\mathbf{a}) M_n(\mathbf{a})$$

$$M_n(\mathbf{a}) = \sum_{k=0}^{n-1} a_k \Phi_k \varepsilon_{k+1}.$$



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - **Contrôle adaptatif**
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



Contrôle adaptatif

On cherche à poursuivre une trajectoire de référence donnée (x_n) . On utilise le **contrôle adaptatif de poursuite**

$$U_n = x_{n+1} - \hat{\theta}_n^t \Phi_n.$$

On a alors

$$x_{n+1} - x_{n+1} = \pi_n + \varepsilon_{n+1}$$

avec

$$\pi_n = (\theta - \hat{\theta}_n)^t \Phi_n.$$

La poursuite est d'autant plus efficace que π_n est proche de 0.



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - **Optimisation**
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



Optimisation

On suppose que (ε_n) satisfait la **loi forte des grands nombres**

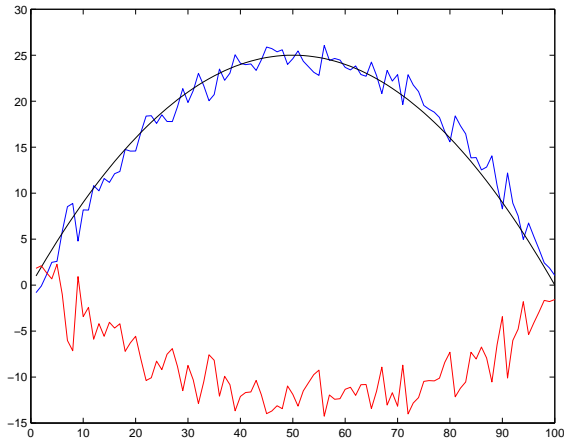
$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \varepsilon_k^2 = \sigma^2 \quad \text{p.s.}$$

avec $\sigma^2 > 0$. On dira que la **poursuite est optimale** si

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - x_k)^2 = \sigma^2 \quad \text{p.s.}$$



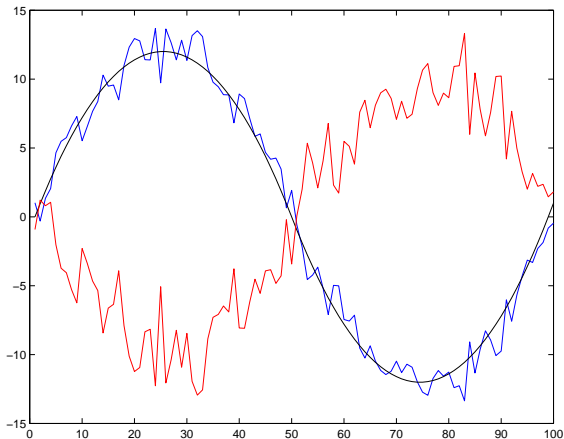
Simulation autorégressif contrôlé



$$X_{n+1} = \theta X_n + U_n + \varepsilon_{n+1}$$

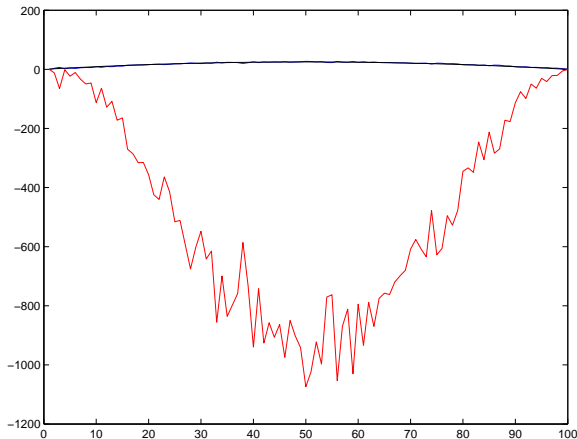


Simulation autorégressif contrôlé



$$X_{n+1} = \theta X_n + U_n + \varepsilon_{n+1}$$

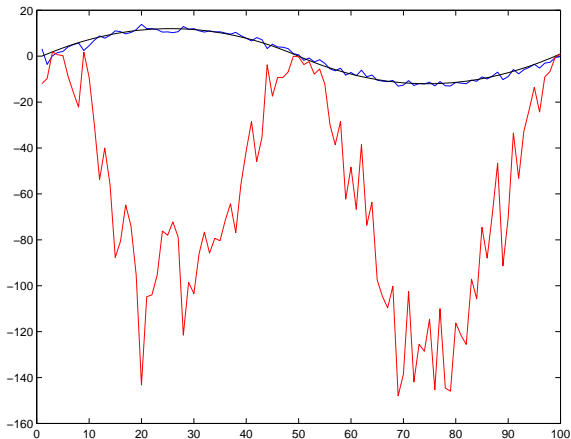
Simulation autorégressif contrôlé



$$X_{n+1} = \theta X_n^2 + U_n + \varepsilon_{n+1}$$



Simulation autorégressif contrôlé



$$X_{n+1} = \theta X_n^2 + U_n + \varepsilon_{n+1}$$



- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



Soit (ε_n) une suite adaptée à $\mathbb{F} = (\mathcal{F}_n)$ avec

$$\mathbb{E}[\varepsilon_{n+1} | \mathcal{F}_n] = 0 \quad \text{et} \quad \mathbb{E}[\varepsilon_{n+1}^2 | \mathcal{F}_n] = \sigma^2 > 0$$

Pour une suite scalaire (Φ_n) adaptée à \mathbb{F} , on étudie la **transformée de martingale**

$$M_n = \sum_{k=1}^n \Phi_{k-1} \varepsilon_k.$$

On appelle **coefficient d'explosion** associé à (Φ_n)

$$f_n = \frac{\Phi_n^2}{s_n} \quad \text{avec} \quad s_n = \sum_{k=0}^n \Phi_k^2.$$



Première loi des grands nombres

On supposera dans toute la suite que

$$\lim_{n \rightarrow \infty} s_n = +\infty \quad \text{p.s.}$$

Théorème (Première LGN)

On a

$$(LGN) \quad \lim_{n \rightarrow \infty} \frac{M_n}{s_{n-1}} = 0 \quad \text{p.s.}$$

Remarque. Si (s_n) converge, alors (M_n) converge p.s.



Seconde loi des grands nombres

Théorème (Seconde LGN)

On suppose que, pour $a > 2$,

$$(H_1) \quad \sup_{n \geq 0} \mathbb{E} [|\varepsilon_{n+1}|^a | \mathcal{F}_n] < \infty \quad p.s.$$

Alors, on a

$$\left(\frac{M_n^2}{S_{n-1}} \right) = \mathcal{O}(\log s_n) \quad p.s.$$

$$\sum_{k=1}^n f_k \left(\frac{M_k^2}{S_{k-1}} \right) = \mathcal{O}(\log s_n) \quad p.s.$$



Loi forte quadratique

Théorème (Loi forte quadratique)

Sous (H_1) , si le coefficient d'explosion $f_n \rightarrow 0$ p.s., on a

$$(LFQ) \quad \lim_{n \rightarrow \infty} \frac{1}{\log s_n} \sum_{k=1}^n f_k \left(\frac{M_k^2}{s_{k-1}} \right) = \sigma^2 \quad p.s.$$

Remarque. On va voir que la **LFQ** donne la convergence du moment d'ordre 2 dans le **TLCPS** associé à (M_n) .



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées**
 - **Modèles linéaires**
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



On se place dans le cadre autorégressif

$$\mathbf{X}_{n+1} = \theta^t \Phi_n + \mathbf{U}_n + \varepsilon_{n+1},$$

$$\Phi_n = (\mathbf{X}_n, \dots, \mathbf{X}_{n-p+1})^t.$$

On suppose que la trajectoire de référence (x_n) vérifie

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n x_k^2 = \tau^2 \quad \text{p.s.}$$

avec $\tau^2 \geq 0$. On pose

$$S_n = \sum_{k=0}^n \Phi_k \Phi_k^t.$$



Théorème (Bercu)

Sous (H_1) , si $\ell = \sigma^2 + \tau^2$, on a

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \ell I_p \quad p.s.$$

- L'estimateur $\hat{\theta}_n$ **converge presque sûrement** vers θ

$$\| \hat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad p.s.$$

- La **poursuite est optimale**

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - x_k)^2 = \sigma^2 \quad p.s.$$



Théorème (Bercu)

Sous (H_1) , si $\ell = \sigma^2 + \tau^2$, on a

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \ell I_p \quad \text{p.s.}$$

- L'estimateur $\hat{\theta}_n$ **converge presque sûrement** vers θ

$$\| \hat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad \text{p.s.}$$

- La **poursuite est optimale**

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - x_k)^2 = \sigma^2 \quad \text{p.s.}$$



Théorème (Bercu)

Sous (H_1) , on a

$$(TLC) \quad \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{\ell} I_p\right),$$

$$(LLI) \quad \limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right) \|\hat{\theta}_n - \theta\|^2 = \frac{\sigma^2}{\ell} \quad p.s.$$

$$(LFQ) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \|\hat{\theta}_k - \theta\|^2 = \frac{\sigma^2}{\ell} \quad p.s.$$



Plan

- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées**
 - Modèles linéaires
 - Processus de branchement**
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



On considère le processus de branchement contrôlé

$$X_{n+1} = \sum_{k=1}^{X_n + U_n} Y_{n,k} + I_{n+1}$$

- X_n → Le nombre d'enfants de la n^{e} génération,
- $Y_{n,k}$ → Les descendants,
- I_n → L'immigration,
- U_n → Le contrôle dont on a le choix.

$(Y_{n,k})$ et (I_n) sont deux suites indépendantes de variables aléatoires **iid** à valeurs dans \mathbb{N} et $V_n = X_n + U_n \geq 1$.



On considère le processus de branchement contrôlé

$$X_{n+1} = \sum_{k=1}^{X_n + U_n} Y_{n,k} + I_{n+1}$$

- X_n → Le nombre d'enfants de la n^{e} génération,
- $Y_{n,k}$ → Les descendants,
- I_n → L'immigration,
- U_n → Le contrôle dont on a le choix.

$(Y_{n,k})$ et (I_n) sont deux suites indépendantes de variables aléatoires **iid** à valeurs dans \mathbb{N} et $V_n = X_n + U_n \geq 1$.



On considère le processus de branchement contrôlé

$$X_{n+1} = \sum_{k=1}^{X_n + U_n} Y_{n,k} + I_{n+1}$$

- X_n → Le nombre d'enfants de la n^{e} génération,
- $Y_{n,k}$ → Les descendants,
- I_n → L'immigration,
- U_n → Le contrôle dont on a le choix.

$(Y_{n,k})$ et (I_n) sont deux suites indépendantes de variables aléatoires **iid** à valeurs dans \mathbb{N} et $V_n = X_n + U_n \geq 1$.



On considère le processus de branchement contrôlé

$$X_{n+1} = \sum_{k=1}^{X_n + U_n} Y_{n,k} + I_{n+1}$$

- X_n → Le nombre d'enfants de la n^{e} génération,
- $Y_{n,k}$ → Les descendants,
- I_n → L'immigration,
- U_n → Le contrôle dont on a le choix.

$(Y_{n,k})$ et (I_n) sont deux suites indépendantes de variables aléatoires **iid** à valeurs dans \mathbb{N} et $V_n = X_n + U_n \geq 1$.



On souhaite estimer les paramètres

$$\mathbb{E}[Y_{n,k}] = m,$$

$$\mathbb{E}[I_n] = \lambda,$$

$$\text{Var}(Y_{n,k}) = \sigma^2,$$

$$\text{Var}(I_n) = b^2.$$

On a l'écriture régressive

$$\mathbf{X}_{n+1} = \theta^t \Phi_n + \varepsilon_{n+1}$$

avec

$$\theta = \begin{pmatrix} m \\ \lambda \end{pmatrix} \quad \text{et} \quad \Phi_n = \begin{pmatrix} V_n \\ 1 \end{pmatrix}.$$



Moindres carrés pondérés

La suite (ε_n) est adaptée à $\mathbb{F} = (\mathcal{F}_n)$ avec $\mathbb{E}[\varepsilon_{n+1}|\mathcal{F}_n] = 0$ et

$$\mathbb{E}[\varepsilon_{n+1}^2|\mathcal{F}_n] = \sigma^2 V_n + b^2.$$

On estime θ par l'algorithme des **moindres carrés pondérés**

$$\hat{\theta}_n = \mathbf{S}_{n-1}^{-1}(\mathbf{a}) \sum_{k=0}^{n-1} \mathbf{a}_k \Phi_k \mathbf{X}_{k+1}$$

où la pondération est donnée, pour $\gamma > 0$, par

$$a_n = \frac{1}{V_n} \left(\frac{1}{\log s_n} \right)^{1+\gamma} \quad \text{avec} \quad s_n = \sum_{k=0}^n V_k.$$



On cherche à **stabiliser la dynamique** du processus (X_n) en le forçant à poursuivre une trajectoire (x_n) . On utilise

$$U_n = \max(1, P(\hat{m}_n^{-1}(x_{n+1} - \hat{\lambda}_n))) - X_n$$

où P est la projection sur \mathbb{N} et $\hat{\theta}_n = (\hat{m}_n, \hat{\lambda}_n)^t$.

Lemme

Si $\mathbf{x}_n \rightarrow \mathbf{x}$, il existe une variable aléatoire $\ell \geq 1$ telle que

$$(\log n)^{1+\gamma} \frac{S_n(\mathbf{a})}{n} \longrightarrow L = \begin{bmatrix} \ell & \mathbf{1} \\ \mathbf{1} & \ell^{-1} \end{bmatrix} \quad p.s.$$

Remarque. $\hat{\theta}_n$ ne converge pas vers θ car $\det(L) = 0$.



On utilise maintenant le **contrôle adaptatif excité**

$$U_n = \max(1, P(\hat{m}_n^{-1}(x_{n+1} - \hat{\lambda}_n))) - X_n + \xi_n$$

où (ξ_n) est une suite **exogène** de variables aléatoires **iid** bornées à valeurs dans \mathbb{N}^* , de loi ξ .

Lemme

Si $x_n \rightarrow x$ et $h = P(m^{-1}(x - \lambda))$, on a

$$(\log n)^{1+\gamma} \frac{S_n(a)}{n} \longrightarrow H = \begin{bmatrix} E[h + \xi] & 1 \\ 1 & E[(h + \xi)^{-1}] \end{bmatrix} \text{ p.s.}$$



Théorème (Bercu)

Si $\mathbf{x}_n \rightarrow \mathbf{x}$, $\hat{\theta}_n$ **converge presque sûrement** vers θ

$$\|\hat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad \text{p.s.}$$

On a également

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, W)$$

avec W calculée explicitement.



- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 **Théorème limite centrale presque sûr**
- 5 Régressions fonctionnelles contrôlées



Théorème limite centrale

Soit (ξ_n) une suite de variables aléatoires **iid** avec $\mathbb{E}[\xi_n] = 0$ et $\mathbb{E}[\xi_n^2] = \sigma^2$. Si $S_n = \xi_1 + \xi_2 + \dots + \xi_n$, on a

$$(TLC) \quad \frac{S_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

Pour toute fonction **h continue bornée**,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[h \left(\frac{S_n}{\sqrt{n}} \right) \right] = \int_{\mathbb{R}} h(x) dG(x)$$

où G est la mesure gaussienne $\mathcal{N}(0, \sigma^2)$.



Théorème limite centrale presque sûr

On a également

$$(TLCPS) \quad \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta_{\frac{S_k}{\sqrt{k}}} \implies G \quad \text{p.s.}$$

Pour toute fonction **h continue bornée**,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} h\left(\frac{S_k}{\sqrt{k}}\right) = \int_{\mathbb{R}} h(x) dG(x) \quad \text{p.s.}$$



On a déjà vu la **LGN** pour la **transformée de martingale**

$$M_n = \sum_{k=1}^n \Phi_{k-1} \varepsilon_k.$$

Le **coefficient d'explosion** associé à (Φ_n) est

$$f_n = \frac{\Phi_n^2}{S_n} \quad \text{avec} \quad S_n = \sum_{k=0}^n \Phi_k^2.$$



Théorème (Brown, Chaabane, Lifshits)

Sous (H_1) , si le coefficient d'explosion $f_n \rightarrow 0$ p.s., on a

$$(TLC) \quad \frac{M_n}{\sqrt{s_{n-1}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

De plus, si pour $\gamma > 0$,

$$\sum_{n=1}^{\infty} f_n^\gamma < \infty \quad \text{p.s.}$$

$$(TLCPS) \quad \frac{1}{\log s_n} \sum_{k=1}^n f_k \frac{\delta_{M_k}}{\sqrt{s_{k-1}}} \implies G \quad \text{p.s.}$$



Puissances de martingales

Pour toute fonction **h continue bornée**,

$$\lim_{n \rightarrow \infty} \frac{1}{\log s_n} \sum_{k=1}^n f_k h\left(\frac{M_k}{\sqrt{s_{k-1}}}\right) = \int_{\mathbb{R}} h(x) dG(x) \quad \text{p.s.}$$

Définition. On dira que (M_n) satisfait un **TLCPSP** si cette convergence a lieu pour toute fonction **polynomiale h** .

Objectif. Passer aux **puissances de martingales** dans le but d'étudier la stabilité des régressions non linéaires contrôlées.



Soit

$$v_n(p) = \frac{s_n^p - s_{n-1}^p}{s_n^p}.$$

Théorème (Bercu)

On suppose que, pour $p \geq 1$ et $a > 2p$,

$$(H_p) \quad \sup_{n \geq 0} \mathbb{E} [|\varepsilon_{n+1}|^a | \mathcal{F}_n] < \infty \quad p.s.$$

Alors, on a

$$\left(\frac{M_n^2}{s_{n-1}} \right)^p = \mathcal{O}(\log s_n) \quad p.s.$$

$$\sum_{k=1}^n v_k(p) \left(\frac{M_k^2}{s_{k-1}} \right)^p = \mathcal{O}(\log s_n) \quad p.s.$$



Théorème (Bercu)

Sous (H_p) , si le coefficient d'explosion $f_n \rightarrow 0$ p.s., on a

$$\lim_{n \rightarrow \infty} \frac{1}{\log s_n} \sum_{k=1}^n f_k \left(\frac{M_k^2}{s_{k-1}} \right)^p = \frac{\sigma^{2p} (2p)!}{2^p p!} \quad \text{p.s.}$$

Théorème (Bercu-Fort)

On suppose que (H_p) est vraie pour tout $p \geq 1$. Si $f_n \rightarrow 0$ p.s., (M_n) satisfait le **TLCPS**

$$\frac{1}{\log s_n} \sum_{k=1}^n f_k \frac{\delta_{M_k}}{\sqrt{s_{k-1}}} \implies G \quad \text{p.s.}$$



Martingales explosives

Pour tout $p \geq 1$, on pose

$$\sigma_n(p) = \mathbb{E}[\varepsilon_{n+1}^p | \mathcal{F}_n] \quad \text{p.s.}$$

Théorème (Bercu)

On se place sous (H_p) avec $\sigma_n(q) \rightarrow \sigma(q)$ pour tout $2 \leq q \leq 2p$ où $\sigma(q) = 0$ si q est impair. Si le coefficient d'explosion $f_n \rightarrow f$ p.s. avec $0 < f < 1$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \left(\frac{M_k^2}{S_{k-1}} \right)^p = I(p, f) \quad \text{p.s.}$$



Limite gaussienne

La limite $I(p, \mathbf{f})$ est donnée par

$$I(p, \mathbf{f}) = \frac{1}{1 - (1 - \mathbf{f})^p} \sum_{k=1}^p C_{2p}^{2k} \mathbf{f}^k (1 - \mathbf{f})^{p-k} \sigma(2k) I(p-k, \mathbf{f}).$$

Elle est libre de \mathbf{f} ssi pour tout $1 \leq k \leq p$,

$$\sigma(2k) = \frac{\sigma^{2k} (2k)!}{2^k k!}.$$

Dans ce cas, on a

$$I(p, \mathbf{f}) = \frac{\sigma^{2p} (2p)!}{2^p p!} = I(p).$$



Martingales explosives

Théorème (Bercu-Fort)

On suppose que (H_p) est vraie pour tout $p \geq 1$. Si $f_n \rightarrow f$ p.s. avec $0 < f < 1$ et $I(p, f) = I(p)$, (M_n) satisfait le **TLCPS**

$$\frac{1}{n} \sum_{k=1}^n \frac{\delta_{M_k}}{\sqrt{s_{k-1}}} \implies \mathbf{G} \quad \text{p.s.}$$

Processus autorégressif stable $|\theta| < 1$

On considère le processus autorégressif stable

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}.$$

Sous (H_1) , on a $f_n \rightarrow 0$,

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \frac{\sigma^2}{(1 - \theta^2)} \quad \text{p.s.}$$

De plus, $\hat{\theta}_n \rightarrow \theta$,

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1 - \theta^2).$$



Si (H_p) est vraie pour tout $p \geq 1$, on a le **TLCPS**

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta_{\sqrt{k}(\hat{\theta}_k - \theta)} \implies \mathcal{N}(\mathbf{0}, 1 - \theta^2) \quad \text{p.s.}$$

En particulier, pour tout $p \geq 1$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n k^{p-1} (\hat{\theta}_k - \theta)^{2p} = \frac{(1 - \theta^2)^p (2p)!}{2^p p!} \quad \text{p.s.}$$



Processus autorégressif explosif $|\theta| > 1$

Sous (H_1) , $\theta^{-n}X_n$ converge p.s. vers la variable aléatoire

$$Y = X_0 + \sum_{k=1}^{\infty} \theta^{-k} \varepsilon_k.$$

On a $f_n \rightarrow (\theta^2 - 1)/\theta^2$,

$$\lim_{n \rightarrow \infty} \frac{S_n}{\theta^{2n}} = \frac{\theta^2 Y^2}{(\theta^2 - 1)} \quad \text{p.s.}$$

De plus, $\hat{\theta}_n \rightarrow \theta$ et si (ε_n) est **gaussien** et \mathcal{C} de loi de Cauchy

$$|\theta|^n (\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} c.$$



Si (ε_n) est **gaussien**, on a le **TLCPSP**

$$\frac{1}{n} \sum_{k=1}^n \delta_{|\theta|^k(\hat{\theta}_k - \theta)} \implies \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2(\theta^2 - 1)}{\gamma^2}\right) \quad \text{p.s.}$$

En particulier, pour tout $p \geq 1$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (|\theta|^k(\hat{\theta}_k - \theta))^{2p} = \frac{\sigma^{2p}(\theta^2 - 1)^p (2p)!}{\gamma^{2p} 2^p p!} \quad \text{p.s.}$$



- 1 Introduction
 - Objectifs
 - Algorithmes d'estimation
 - Contrôle adaptatif
 - Optimisation
- 2 Lois fortes des grands nombres
- 3 Régressions linéaires contrôlées
 - Modèles linéaires
 - Processus de branchement
- 4 Théorème limite centrale presque sûr
- 5 Régressions fonctionnelles contrôlées



On considère le modèle autorégressif fonctionnel d'ordre $d \geq 1$

$$X_{n+1} = \theta f(X_n, \dots, X_{n-d+1}) + U_n + \varepsilon_{n+1}.$$

On utilise l'estimateur des **moindres carrés**

$$\hat{\theta}_n - \theta = \frac{M_n}{S_{n-1}} \quad \text{avec} \quad M_n = \sum_{k=1}^n \Phi_{k-1} \varepsilon_k.$$

On choisit le **contrôle adaptatif de poursuite**

$$U_n = x_{n+1} - \hat{\theta}_n \Phi_n$$

où $\Phi_n = f(X_n, \dots, X_{n-d+1})$.



Une classe de fonctions

Soit $\mathcal{C}(a, b)$ avec $a, b \in \mathbb{N}$ et $a \geq 1$, la classe des fonctions f de \mathbb{R}^d dans \mathbb{R} vérifiant pour tout $x \in \mathbb{R}^d$

$$c_1 + c_2 \|x\|^b \leq |f(x)| \leq c_3 + c_4 \|x\|^a$$

avec $b \geq 1$ si $c_1 = 0$ et $b \geq 0$ sinon.



Corollaire (Bercu-Portier)

Sous (H_a) , si $f \in \mathcal{C}(a, b)$ avec $a < 4$, on a $\hat{\theta}_n \rightarrow \theta$,

$$(\hat{\theta}_n - \theta)^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad p.s.$$

Pour tout $1 \leq p \leq a$, la **poursuite est stable d'ordre p**

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - x_k)^{2p} < \infty \quad p.s.$$

Si $\sigma_n(2p) \rightarrow \sigma(2p)$, la **poursuite est optimale d'ordre p**

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - x_k)^{2p} = \sigma(2p) \quad p.s.$$



Une hypothèse LGN

Pour $a \geq 1$, soit $\mathcal{P}(a)$ l'algèbre des polynômes à d variables et de **degré total** $\leq a$. On note (L_a) l'hypothèse $f^2 \in \mathcal{P}(2a)$ avec

$$(L_a) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=d}^n f^2(\varepsilon_k + x_k, \dots, \varepsilon_{k-d+1} + x_{k-d+1}) = \ell \quad \text{p.s.}$$

où $\ell > 0$. Sous (H_a) et (L_a) avec $a < 4$, on a

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \ell \quad \text{p.s.}$$



Corollaire (Bercu-Portier)

Sous (H_a) et (L_a) avec $a < 4$, on a $\hat{\theta}_n \rightarrow \theta$,

$$(TLC) \quad \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{\ell}\right),$$

$$(LLI) \quad \limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right) (\hat{\theta}_n - \theta)^2 = \frac{\sigma^2}{\ell} \quad p.s.$$

Finalement, pour tout $1 \leq p \leq a$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{k=1}^n k^{p-1} (\hat{\theta}_k - \theta)^{2p} = \frac{\sigma^{2p} (2p)!}{\ell^p 2^p p!} \quad p.s.$$



Processus autorégressifs polynômiaux de degré 2

On suppose que $x_n \rightarrow 0$ et $\sigma_n(p) \rightarrow \sigma(p)$ pour $1 \leq p \leq 4$. On considère les processus autorégressifs polynômiaux

$$(1) \quad X_{n+1} = \theta X_n^2 + U_n + \varepsilon_{n+1},$$

$$(2) \quad X_{n+1} = \theta X_n(1 - X_n) + U_n + \varepsilon_{n+1},$$

$$(3) \quad X_{n+1} = \theta X_n X_{n-1} + U_n + \varepsilon_{n+1}.$$

Le Corollaire est vrai avec $\ell(1) = \sigma(4)$,

$$\ell(2) = \sigma(4) - 2\sigma(3) + \sigma(2), \quad \ell(3) = \sigma(2)^2.$$



Processus autorégressifs polynômiaux de degré 3

On suppose que $x_n \rightarrow 0$ et $\sigma_n(p) \rightarrow \sigma(p)$ pour $1 \leq p \leq 6$. On considère les processus autorégressifs polynômiaux

$$(4) \quad X_{n+1} = \theta X_n^3 + U_n + \varepsilon_{n+1},$$

$$(5) \quad X_{n+1} = \theta X_n^2(1 - X_n) + U_n + \varepsilon_{n+1},$$

$$(6) \quad X_{n+1} = \theta X_n^2 X_{n-1} + U_n + \varepsilon_{n+1}.$$

Le Corollaire est vrai avec $\ell(4) = \sigma(6)$,

$$\ell(5) = \sigma(6) - 2\sigma(5) + \sigma(4), \quad \ell(6) = \sigma(4)\sigma(2).$$

