Convergence des moments dans le TLCPS pour les martingales vectorielles

Peggy Cénac - IUT et MAP5, Université Paris 5

Groupe de travail en statistique au MAP5 - 26 janvier 2007

Plan

1. Théorème de la limite centrale presque sûr (TLCPS)
   - Cas scalaire
   - Cas vectoriel
   - Théorème principal
2. Applications statistiques
   - Modèles de regression linéaire
   - Modèles autorégressifs linéaires
   - Processus de branchement avec immigration
3. Algorithmes stochastiques
4. Perspectives

Théorème de la limite centrale presque sûr (TLCPS)

Théorèmes de la limite centrale

- Soit (ξ_n) une suite de variables i.i.d. avec $\mathbb{E}[\xi_n] = 0$ et $\mathbb{E}[\xi_n^2] = \sigma^2$.
- $Z_n \stackrel{\text{def}}{=} \xi_1 + \ldots + \xi_n$.

Pour toute fonction h continue bornée,

▷ (TLC) $$\lim_{n\to\infty} \mathbb{E}\left[h\left(\frac{Z_n}{\sqrt{n}}\right)\right] = \int_{\mathbb{R}} h(x)dG(x),$$

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

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

TLCPS - Cadre martingales

- Ce théorème a été démontré par Brosamler, Schatte et dans sa forme présente par Lacey.
- Le TLCPS a aussi été établi dans un cadre martingales par Chaâbane, Chaâbane et Maâouia, Lifshits.
- Soit (ε_n) une suite de différences de martingale adaptée à $\mathbb{F}$, (Φ_n) une suite de v.a. adaptée à $\mathbb{F}$, la transformée de martingale réelle

$$M_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k + M_0,$$

- le coefficient d'explosion f_n associé à Φ_n

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

TLCPS - Cadre martingales (2)

Théorème (Version simplifiée, Chaâbane, 1996)

Si $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ *p.s. et*

$$\forall \delta > 0 \quad \sum_{n=1}^{\infty} s_{n-1}^{-1} \mathbb{E}[(M_n - M_{n-1})^2 \mathbb{1}_{\{|M_n - M_{n-1}| > \delta \sqrt{s_{n-1}}\}} | \mathcal{F}_{n-1}] < \infty \quad p.s.$$

$$\exists a > 0 \quad \sum_{n=1}^{\infty} s_{n-1}^{-a} \mathbb{E}[|M_n - M_{n-1}|^{2a} \mathbb{1}_{\{|M_n - M_{n-1}| \leq \sqrt{s_{n-1}}\}} | \mathcal{F}_{n-1}] < \infty \quad p.s.$$

alors pour toute fonction h continue bornée,

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

Convergence des moments

Que se passe-t-il si h n'est pas bornée ?

Théorème (Bercu, 2004)

On suppose que presque sûrement $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ *et* f_n *tend vers zéro. S'il existe un entier* $p \geq 1$ *et un réel* $a > 2p$ *tels que*

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

alors on a

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

TLCPS pour martingales vectorielles

Théorème (Chaâbane et Maâouia, 1998)

(M_n) est une martingale de carré intégrable. S'il existe une suite déterministe (U_n) de matrices réelles inversibles satisfaisant des conditions de croissance régulière, sous les hypothèses

$$i) \qquad U_n^{-1}\langle M\rangle_n U_n^{-1} \xrightarrow[n\to\infty]{\text{p.s.}} C, \quad \text{où } C \text{ est une matrice, aléatoire ou non,}$$

$$ii) \qquad \forall\delta>0,\quad \sum_{k=1}^{n}\mathbb{E}\Big[\|\, U_n^{-1}(M_k-M_{k-1})\,\|^2\, \mathbb{1}_{\{\|U_n^{-1}(M_k-M_{k-1})\|\geq\delta\}}|\mathcal{F}_{k-1}\Big] \xrightarrow[n\to\infty]{\text{p.s.}} 0,$$

(M_n) satisfait le TLCPS

$$\frac{1}{\log(\det U_n)^2}\sum_{k=1}^{n}\Big(1-\Big(\frac{\det U_k}{\det U_{k+1}}\Big)^2\Big)\delta_{U_k^{-1}M_k} \Longrightarrow Y \quad p.s.,$$

où Y est la loi mélange $C^{1/2}G$, G étant un vecteur gaussien standard indépendant de C. Si C n'est pas aléatoire, la mesure limite est la loi gaussienne $\mathcal{N}(0,C)$.

Cadre martingales vectorielles - Définitions

- (Φ_n) est une suite de v.a. dans $\mathbb{R}^d$ adaptée à $\mathbb{F}$,
- (M_n) est la transformée de martingales

$$M_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k + M_0.$$

- Soit S une matrice définie positive, symétrique et déterministe, on définit la matrice aléatoire

$$S_n \stackrel{\text{def}}{=} \sum_{k=0}^{n} \Phi_k \Phi_k^t + S,$$

- et le coefficient d'explosion associé à (Φ_n)

$$f_n \stackrel{\text{def}}{=} \Phi_n^t S_n^{-1} \Phi_n = \frac{d_n - d_{n-1}}{d_n} \quad \text{avec} \quad d_n \stackrel{\text{def}}{=} \det(S_n).$$

Notations

On note respectivement (H_{2p}), (H_{2p+}) et (C_{2p}) les assertions : (ε_n) est une différence de martingale telle que

$$(H_{2p}) \qquad \sup_{n\geq 0} \mathbb{E}\left[\varepsilon_{n+1}{}^{2p} \middle| \mathcal{F}_n \right] < \infty \quad \text{p.s.}$$

$$(H_{2p+}) \qquad \sup_{n\geq 0} \mathbb{E}\left[|\varepsilon_{n+1}|^{a} \middle| \mathcal{F}_n \right] < \infty \quad \text{p.s.} \quad \text{pour un réel } a > 2p,$$

$$(C_{2p}) \qquad \forall n \geq 0, \quad \mathbb{E}\left[\varepsilon_{n+1}^{2p} \middle| \mathcal{F}_n \right] \stackrel{\text{def}}{=} \sigma(2p) < \infty \quad \text{p.s.}$$

On note $\sigma(2) = \sigma^2$.

Théorème principal

Théorème (Convergence des moments dans le TLCPS)

On suppose que (ε_n) satisfaisait (C_2) et (H_{2p+}) et que f_n tend vers zéro p.s. S'il existe une suite aléatoire positive (α_n) croissante vers l'infini et une matrice inversible L telles que

$$\lim_{n\to\infty} \alpha_n^{-1} S_n = L \quad p.s.,$$

alors on a presque sûrement

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k \left(M_k^t S_{k-1}^{-1} M_k\right)^p = d\sigma^{2p} \prod_{j=1}^{p-1} (d+2j) \stackrel{\text{def}}{=} \ell(p).$$

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} \left[\left(M_k^t S_{k-1}^{-1} M_k\right)^p - \left(M_k^t S_k^{-1} M_k\right)^p \right] = \frac{p}{d} \ell(p).$$

Applications statistiques

Modèles de regression linéaire

- On considère le modèle de regression linéaire de paramètre inconnu $\theta \in \mathbb{R}^d$

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

- X_n, Φ_n, et ε_n sont respectivement l'observation scalaire, le vecteur de regression et le bruit scalaire du système.

Asymptotique des erreurs

- On choisit une suite $(\widehat{\theta}_n)$ d'estimateurs de θ. Quelle est la performance asymptotique de l'estimation et de la prédiction ?
- On se concentre sur l'erreur de prédiction $X_{n+1} - \widehat{\theta}_n^t \Phi_n$ et sur l'erreur d'estimation $\widehat{\theta}_n - \theta$.
- Il est plus approprié de considérer les erreurs cumulées de prédiction et d'estimation

$$C_n(p) = \sum_{k=0}^{n-1} (X_{k+1} - \widehat{\theta}_k^t \Phi_k)^{2p}, \quad G_n(p) = \sum_{k=1}^{n} k^{p-1} \|\widehat{\theta}_k - \theta\|^{2p}.$$

Asymptotique des erreurs (2)

- Dans le cas scalaire, sous des hypothèses de moment appropriées, des résultats asymptotiques sur les erreurs ont été établis par Bercu, en utilisant l'estimateur des moindres carrés

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k=1}^{n} \Phi_{k-1} X_k.$$

- Ces résultats sont maintenant étendus au cadre vectoriel grâce à la convergence des moments dans le TLCPS.

Estimation des moments

Un estimateur naturel du moment $\sigma(2p)$ du bruit est

$$\Gamma_n(2p) = \frac{1}{n}\sum_{k=0}^{n-1}\left(X_{k+1} - \widehat{\theta}_k^t \Phi_k\right)^{2p}.$$

Corollaire

On suppose qu'il existe

- $p \geq 2$ *vérifiant* (C_{2p}),
- *une suite aléatoire* (α_n) *et une matrice inversible* L *telles que l'hypothèse de convergence du théorème principal soit vérifiée.*

Si f_n *tend presque sûrement vers zéro alors*

$$\left(\Gamma_n(2p) - \frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^{2p}\right)^2 = \mathcal{O}\left(\frac{\log d_n}{n}\right) \quad p.s.$$

Erreur de prédiction

- Lien entre estimateur des moments et erreurs de prédiction :

$$n\Gamma_n(2p) = C_n(p)$$

- Sous les hypothèses du corollaire précédent, en supposant de plus que l'hypothèse (H_{2p+}) est vérifiée, pour c vérifiant $2pa^{-1} < c < 1$, on a

$$\left| \frac{1}{n}\ C_n(p) - \sigma(2p) \right|^2 = o(n^{c-1}) \quad \text{p.s.}$$

Erreur d'estimation

Corollaire

Sous les hypothèses du théorème principal, on a

$$\lim_{n\to\infty} \frac{1}{\log d_n} \sum_{k=1}^{n} f_k \left((\widehat{\theta}_k - \theta)^t S_k (\widehat{\theta}_k - \theta) \right)^p = \ell(p) \quad p.s.$$

Ainsi, puisque L est strictement définie positive, on obtient

$$G_n(p) = \mathcal{O}\left(\log d_n\right) \quad p.s.$$

Modèles autorégressifs linéaires

Dans le cas particulier de modèles autorégressifs linéaires stables

$$X_{n+1} = \sum_{k=1}^{d} \theta_k X_{n-k+1} + \varepsilon_{n+1},$$

si (ε_n) est une suite de différences de martingales vérifiant (C_{2p}) et un moment conditionnel fini d'ordre $a > 2p$, alors on a

$$\left| \frac{1}{n}\, C_n(p) - \sigma(2p) \right|^2 = o(n^{c-1}), \qquad G_n(p) = \mathcal{O}(\log n) \quad \text{p.s.}$$

Processus de branchement avec immigration

- On considère le processus de branchement avec immigration donné par la relation de récurrence

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

- On suppose que les v.a. (I_n) et $(Y_{n,k})$, indépendantes et identiquement distribuées, sont indépendantes entre elles.

$$\mathbb{E}[Y_{n,k}] \stackrel{\text{def}}{=} m, \quad \mathbb{E}[I_n] \stackrel{\text{def}}{=} \lambda,$$
$$\text{var}[Y_{n,k}] \stackrel{\text{def}}{=} \sigma^2, \quad \text{var}[I_n] \stackrel{\text{def}}{=} b^2.$$

Estimation de la moyenne

- Le modèle peut s'écrire sous la forme

$$X_{n+1} = mX_n + \lambda + \varepsilon_{n+1},$$

où (ε_n) est une suite de différences de martingale.

- Dans ce modèle, le moment conditionnel d'ordre 2 du bruit n'est pas presque sûrement borné car

$$\mathbb{E}\left[\varepsilon_{n+1}^2 \mid \mathcal{F}_n\right] = \sigma^2 X_n + b^2.$$

Estimation de la moyenne (2)

Dans le cas stable $m < 1$, l'erreur de prédiction vérifie

$$\left| \frac{1}{n} \ C_n(q) - \sigma(2q) \right|^2 = o(n^{c-1}) \quad \text{p.s.}$$

Pour l'erreur d'estimation on a

$$G_n(p) = \mathcal{O}(\log n) \quad \text{p.s.}$$

Estimation de la variance

- La définition de (ε_n) permet d'écrire la décomposition

$$\varepsilon_{n+1}^2 = \sigma^2 X_n + b^2 + v_{n+1},$$

  où (v_n) est une suite de différences de martingale.

- On utilise l'estimateur des variances $\eta^t \stackrel{\text{def}}{=} (\sigma^2, b^2)$ défini par

$$\widehat{\eta}_n \stackrel{\text{def}}{=} Q_n^{-1} \sum_{k=1}^{n} c_k^{-2} \Phi_k \left(X_{k+1} - \widehat{\theta}_k \Phi_k\right)^2, \quad Q_n \stackrel{\text{def}}{=} S + \sum_{k=1}^{n} c_k^{-2} \Phi_k \Phi_k^t$$

- Les comportements asymptotiques des erreurs d'estimation et de prédiction cumulées sont identiques au cas de l'estimation de la moyenne.

Algorithmes stochastiques

Algorithmes stochastiques

- On considère le modèle

$$Z_{n+1} = Z_n + \gamma_n \big[h(Z_n) + R_{n+1} \big] + \sigma_n \varepsilon_{n+1},$$

  où h est définie de $\mathbb{R}$ dans $\mathbb{R}$.

- (R_n) et (ε_n) sont deux perturbations réelles. Les pas (γ_n) et (σ_n) sont deux suites déterministes positives convergeant vers zéro.

-
$$\lim_{n\to\infty} \frac{1}{s_n} \sum_{k=1}^{n} \gamma_k \Big[\sqrt{\gamma_k}\,\sigma_k^{-1}(Z_k - z^*) \Big]^{2p} = \frac{\Sigma^{2p}(2p)!}{2^p p!} \quad \text{a.s.}$$

Perspectives

- Convergence composante par composante ;
- Carleman ;
- Réduire les hypothèses de convergence ;
- Cas non explosif ;