

Estimation fonctionnelle et contrôle adaptatif

B. Bercu

Université de Bordeaux

INRIA Bordeaux Sud-Ouest, Equipe CQFD

23 Octobre 2008



Outline

- 1 Estimation d'une densité
 - Estimateur de Parzen-Rosenblatt
 - Estimateur de Robbins-Monro
 - Résultats pour Wolverton-Wagner

- 2 Contrôle adaptatif
 - Régression linéaire contrôlée
 - Wolverton-Wagner en contrôle adaptatif

Outline

- 1 Estimation d'une densité
 - Estimateur de Parzen-Rosenblatt
 - Estimateur de Robbins-Monro
 - Résultats pour Wolverton-Wagner

- 2 Contrôle adaptatif
 - Régression linéaire contrôlée
 - Wolverton-Wagner en contrôle adaptatif

Objectif

Soit (X_n) une suite de variables aléatoires réelles **iid**, de densité inconnue f . Soit K une fonction **positive, bornée**, appelée **noyau**, telle que

$$\int_{\mathbb{R}} K(x) dx = 1, \quad \int_{\mathbb{R}} xK(x) dx = 0,$$

$$\int_{\mathbb{R}} K^2(x) dx = \sigma^2.$$

Objectif. Estimer la densité f via un estimateur à noyau.



Choix du noyau

- **Noyau Uniforme**

$$K_a(x) = \begin{cases} \frac{1}{2a} & \text{si } |x| \leq a, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau d'Epanechnikov**

$$K_b(x) = \begin{cases} \frac{3}{4b} \left(1 - \frac{x^2}{b^2}\right) & \text{si } |x| \leq b, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau Gaussien**

$$K_c(x) = \frac{1}{c\sqrt{2\pi}} \exp\left(-\frac{x^2}{2c^2}\right).$$



Choix du noyau

- **Noyau Uniforme**

$$K_a(x) = \begin{cases} \frac{1}{2a} & \text{si } |x| \leq a, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau d'Epanechnikov**

$$K_b(x) = \begin{cases} \frac{3}{4b} \left(1 - \frac{x^2}{b^2}\right) & \text{si } |x| \leq b, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau Gaussien**

$$K_c(x) = \frac{1}{c\sqrt{2\pi}} \exp\left(-\frac{x^2}{2c^2}\right).$$



Choix du noyau

- **Noyau Uniforme**

$$K_a(x) = \begin{cases} \frac{1}{2a} & \text{si } |x| \leq a, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau d'Epanechnikov**

$$K_b(x) = \begin{cases} \frac{3}{4b} \left(1 - \frac{x^2}{b^2}\right) & \text{si } |x| \leq b, \\ 0 & \text{sinon.} \end{cases}$$

- **Noyau Gaussien**

$$K_c(x) = \frac{1}{c\sqrt{2\pi}} \exp\left(-\frac{x^2}{2c^2}\right).$$



Parzen-Rosenblatt

Soit (h_n) une suite > 0 et décroissante vers 0, appelée **fenêtre**.
On peut estimer f par l'estimateur de **Parzen-Rosenblatt**

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right).$$

La fonction de répartition empirique associée à (X_n) est

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}_{\{X_i \leq x\}}.$$



Parzen-Rosenblatt, Justification

La **dérivée par différentiation** de F_n est donnée par

$$\begin{aligned} f_n(x) &= \frac{1}{h_n} \left(F_n\left(x + \frac{h_n}{2}\right) - F_n\left(x - \frac{h_n}{2}\right) \right) \\ &= \frac{1}{nh_n} \sum_{i=1}^n \mathbf{I} \left\{ x - \frac{h_n}{2} \leq X_i \leq x + \frac{h_n}{2} \right\} = \frac{1}{nh_n} \sum_{i=1}^n \mathbf{I} \left\{ \frac{1}{2} \leq \frac{X_i - x}{h_n} \leq \frac{1}{2} \right\} \\ &= \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \end{aligned}$$

où K est le **noyau uniforme** avec $a = 1/2$.



Parzen-Rosenblatt, Justification

La **dérivée par différentiation** de F_n est donnée par

$$\begin{aligned} f_n(x) &= \frac{1}{h_n} \left(F_n\left(x + \frac{h_n}{2}\right) - F_n\left(x - \frac{h_n}{2}\right) \right) \\ &= \frac{1}{nh_n} \sum_{i=1}^n \mathbf{I} \left\{ x - \frac{h_n}{2} \leq X_i \leq x + \frac{h_n}{2} \right\} = \frac{1}{nh_n} \sum_{i=1}^n \mathbf{I} \left\{ \frac{1}{2} \leq \frac{X_i - x}{h_n} \leq \frac{1}{2} \right\} \\ &= \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \end{aligned}$$

où K est le **noyau uniforme** avec $a = 1/2$.



Robbins-Monro

Soit (γ_n) une suite > 0 et décroissante vers 0, appelée **pas de l'algorithme**. On peut estimer f par l'estimateur récursif de type **Robbins-Monro**

$$\hat{f}_n(x) = \hat{f}_{n-1}(x) + \gamma_n \left(Z_n(x) - \hat{f}_{n-1}(x) \right)$$

avec

$$Z_n(x) = \frac{1}{h_n} K \left(\frac{X_n - x}{h_n} \right).$$

Remarque. Il s'agit de trouver le zéro de $L(y) = y - f(x)$.



Wolverton-Wagner et Deheuvels

Si $\gamma_n = 1/n$, on retrouve l'estimateur de **Wolverton-Wagner**

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i} K\left(\frac{X_i - x}{h_i}\right).$$

De plus, si

$$\gamma_n = \frac{h_n}{s_n} \quad \text{avec} \quad s_n = \sum_{i=1}^n h_i,$$

on retrouve l'estimateur de **Deheuvels**

$$\tilde{f}_n(x) = \frac{1}{s_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_i}\right).$$



Wolverton-Wagner

Théorème

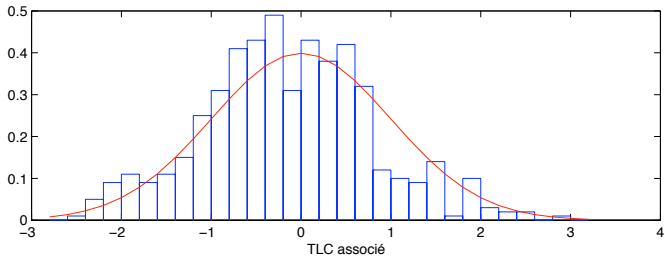
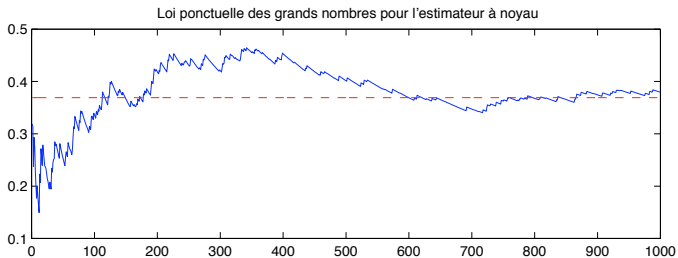
Si f est dérivable à dérivée bornée et si la fenêtre $h_n = 1/n^\alpha$ avec $0 < \alpha < 1$, on a

$$(LGN) \quad \lim_{n \rightarrow \infty} \hat{f}_n(x) = f(x) \quad p.s.$$

De plus, si $1/5 < \alpha < 1$, on a

$$(TLC) \quad \sqrt{nh_n} (\hat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \frac{\sigma^2 f(x)}{1 + \alpha} \right).$$





Outline

- 1 Estimation d'une densité
 - Estimateur de Parzen-Rosenblatt
 - Estimateur de Robbins-Monro
 - Résultats pour Wolverton-Wagner

- 2 Contrôle adaptatif
 - Régression linéaire contrôlée
 - Wolverton-Wagner en contrôle adaptatif

On considère le processus contrôlé

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

- X_n \longrightarrow L'observation,
- U_n \longrightarrow Le contrôle dont on a le choix,
- ε_n \longrightarrow Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Estimer la densité f du bruit.



On considère le processus contrôlé

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

- X_n → L'observation,
- 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 ,
- Estimer la densité f du bruit.



On considère le processus contrôlé

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

- X_n → L'observation,
- 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 ,
- Estimer la densité f du bruit.



On considère le processus contrôlé

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

- X_n → L'observation,
- 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 ,
- Estimer la densité f du bruit.



On considère le processus contrôlé

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

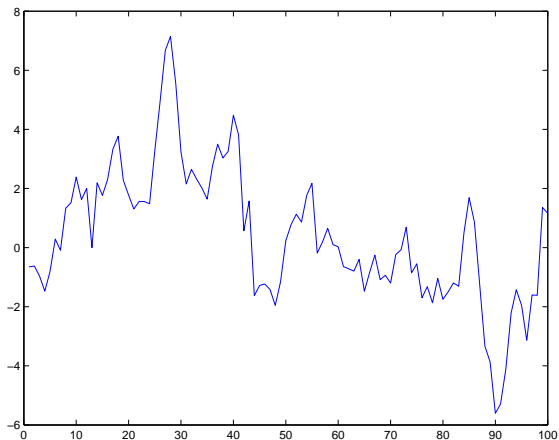
- X_n \longrightarrow L'observation,
- U_n \longrightarrow Le contrôle dont on a le choix,
- ε_n \longrightarrow Le bruit.

On se fixe deux objectifs à atteindre simultanément

- Estimer le paramètre inconnu θ ,
- Estimer la densité f du bruit.

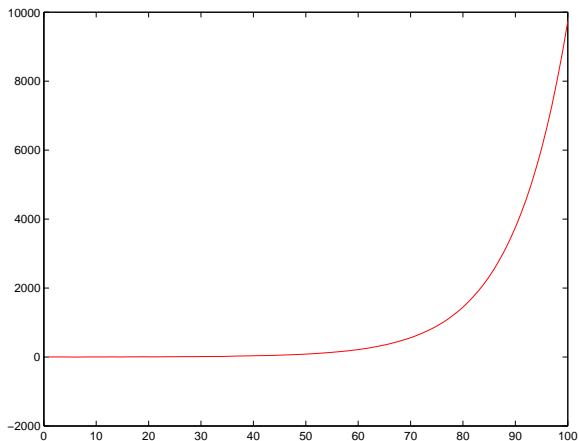


Simulation autorégressif stable $|\theta| < 1$



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

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.

Estimation et contrôle

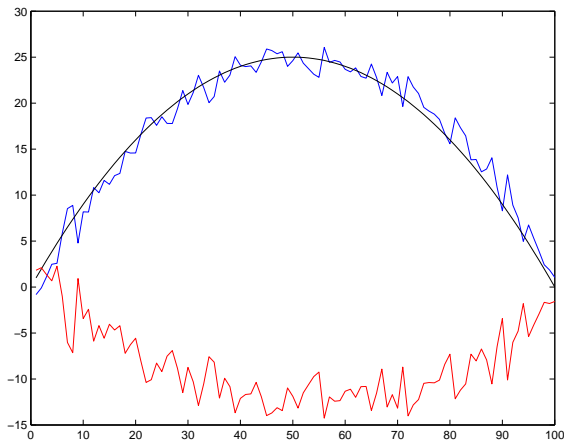
On estime θ par l'estimateur des **moindres carrés**

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

On utilise le **contrôle adaptatif de poursuite**

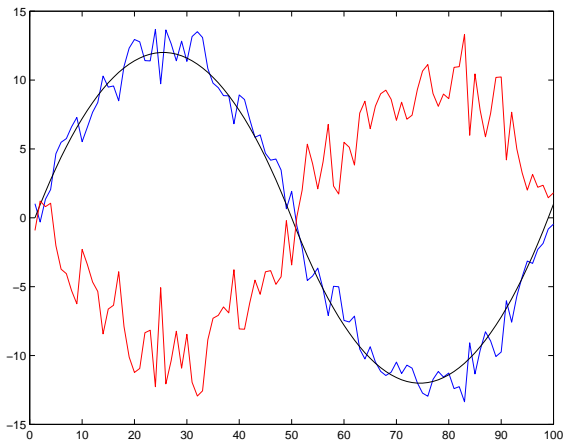
$$U_n = x_{n+1} - \hat{\theta}_n \Phi_n.$$

Simulation autorégressif contrôlé



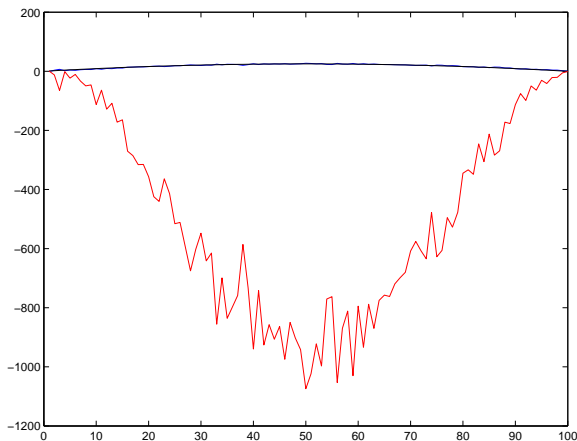
$$X_{n+1} = \theta X_n + U_n + \varepsilon_{n+1}$$

Simulation autorégressif contrôlé



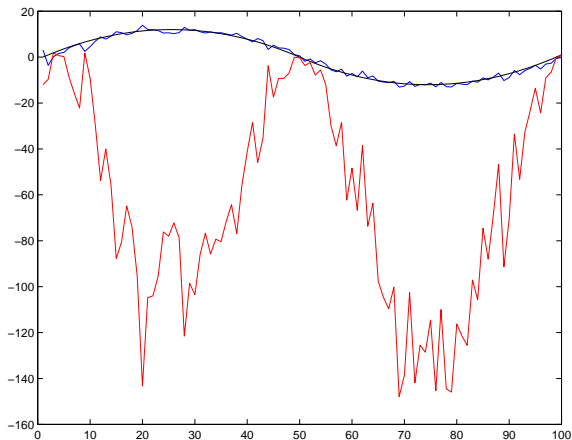
$$X_{n+1} = \theta X_n + U_n + \varepsilon_{n+1}$$

Simulation autorégressif contrôlé



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

Simulation autorégressif contrôlé



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

Si (ε_n) était observable, on pourrait estimer f par

$$f_n(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i} K \left(\frac{\varepsilon_i - \mathbf{x}}{h_i} \right).$$

Comme (ε_n) n'est pas observable, on peut **prédire** ε_{n+1} par

$$\widehat{\varepsilon}_{n+1} = X_{n+1} - \widehat{\theta}_n \Phi_n - U_n = X_{n+1} - x_{n+1}.$$

On estime alors f par l'estimateur de **Wolverton-Wagner**

$$\widehat{f}_n(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i} K \left(\frac{X_i - x_i - \mathbf{x}}{h_i} \right).$$



Wolverton-Wagner

Théorème

Si f est dérivable à dérivée bornée, si K est Lipschitzien, et si la fenêtre $h_n = 1/n^\alpha$ avec $0 < \alpha < 1$, on a

$$(LGNU) \quad \lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |\hat{f}_n(x) - f(x)| = 0 \quad p.s.$$

Wolverton-Wagner

Théorème

Si $1/5 < \alpha < 1$, on a

$$\mathbf{G}_n(\mathbf{x}) = \sqrt{nh_n}(\hat{f}_n(\mathbf{x}) - f(\mathbf{x})) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2 f(\mathbf{x})}{1 + \alpha}\right) = \mathbf{G}(\mathbf{x}).$$

Pour p points distincts x_1, \dots, x_p de \mathbb{R} , on a

$$(TLCM) \quad (\mathbf{G}_n(x_1), \dots, \mathbf{G}_n(x_p)) \xrightarrow{\mathcal{L}} (\mathbf{G}(x_1), \dots, \mathbf{G}(x_p))$$

où $G(x_1), \dots, G(x_p)$ sont des gaussiennes indépendantes.

Application. Test d'adéquation sur la densité f .

