Université Paris Descartes, Sorbonne Paris Cité
UFR Mathématiques et Informatique
Laboratoire MAP5, CNRS-UMR 8145

Cours de statistique des diffusions.

V. Genon-Catalot.

Cours de Statistique des diffusions.

V. Genon-Catalot.

Contents

Chapter 1

Introduction: observations et cadres asymptotiques.

Supposons que l'on dispose d'observations d'un processus $(\xi_t, t \in [0,T])$ qui évolue suivant l'équation différentielle stochastique:

$$d\xi_t = b(t, \xi_t, \theta_0)dt + \sigma(t, \xi_t, \theta_0)dW_t, \quad \xi_0 = \eta \tag{1.1}$$

où W est un mouvement brownien standard réel, η est une variable aléatoire réelle indépendante de $(W_t, t \geq 0)$. Nous appellerons diffusions cette classe de processus stochastiques. (Ceci constitue un abus de langage. Processus de diffusion signifie plus exactement processus de Markov fort à trajectoires continues, ce qui n'est pas le cas du modèle ci-dessus car il n'est pas fortement markovien). On a une idée de la forme des fonctions déterministes

$$b(t, x, \theta) : \mathbb{R}^+ \times \mathbb{R} \times \Theta \to \mathbb{R} \quad \text{(dérive)},$$

$$\sigma(t, x, \theta) : \mathbb{R}^+ \times \mathbb{R} \times \Theta \to \mathbb{R} \quad \text{(coefficient de diffusion)}.$$

Ces fonctions dépendent toutefois d'un paramètre inconnu $\theta \in \Theta \subset \mathbb{R}^p$. Le processus observé correspond à la valeur θ_0 de ce paramètre. On dira que θ_0 est la vraie valeur (inconnue) du paramètre. La loi de la variable initiale peut également dépendre de θ_0.

Le problème que l'on se pose ici est celui de l'estimation de θ_0 à partir des observations de la trajectoire dont on dispose. Notre but est donc de construire des estimateurs $\hat{\theta}$ de θ_0, fonctions des observations et d'en étudier les propriétés. La méthode de construction d'estimateurs dépendra de la nature des observations, chaque type d'observations engendrant un problème statistique propre.

1.1 Différents types d'observations.

Nous nous intéresserons principalement aux deux cas suivants.

- Observation discrétisée.

  Le point de vue naturel et le plus réaliste sur le plan pratique consiste à supposer que l'on observe la trajectoire ξ à des instants déterministes $t_1, \ldots, t_n$ avec $0 \leq t_1 < t_2 < \ldots <$

$t_n \leq T$. Le modèle statistique est alors défini par l'observation du n-uple $(\xi_{t_1}, \ldots, \xi_{t_n})$ et les estimateurs seront des fonctions mesurables $\hat{\theta}_n(\xi_{t_1}, \ldots, \xi_{t_n})$. On étudiera le cas où $t_i = i\Delta$ avec $T = n\Delta$, c'est-à-dire le cas où les observations sont régulièrement espacées, Δ étant le pas de la discrétisation.

- Observation continue.

  On peut aussi supposer que l'on observe en temps continu la trajectoire ξ_t sur tout un intervalle de temps $[0, T]$. Le modèle statistique est ainsi défini par l'observation de la fonction $(\xi_t, t \in [0, T])$ et les estimateurs seront des fonctions $\hat{\theta}_T(\xi_t, t \in [0, T])$.

  Cette hypothèse est admise dans toute la littérature portant sur la statistique des processus à temps continu. Elle est justifiée par le fait qu'elle permet l'obtention de résultats théoriques. Elle est aussi justifiée par le fait que l'on dispose désormais, notamment dans le domaine de la finance, de données très rapprochées dans le temps (cotation "en continu" d'indices boursiers par exemple [1]).

1.2 Différentes méthodes d'estimation.

Comme le montre la présentation du modèle, nous aborderons ici seulement le problème d'estimation paramétrique. Les différentes méthodes d'estimation que nous envisagerons sont les suivantes:

- Maximum de vraisemblance.

  Ceci suppose de calculer une vraisemblance de l'observation. Cette méthode est possible pour une observation continue, mais plus difficile pour une observation en temps discret ou tout autre type d'observation incomplète de la trajectoire (comme par exemple, l'observation de temps de passage par différents niveaux).

- Minimum de contraste ou fonctions d'estimation.

  Lorsque la vraisemblance exacte de l'observation est difficile à exploiter, on construira des pseudo-vraisemblances (contrastes), ou des pseudo-fonctions scores (fonctions d'estimation).

- Méthodes empiriques.

  Entrent dans ce cadre toutes les méthodes basées sur des théorèmes limites (théorèmes ergodiques, par exemple) associés à différentes fonctionnelles des observations (méthodes de moments, de moments généralisés).

[1] Cependant, il faut noter que les modèles proposés pour les données très rapprochées à l'intérieur d'une même journée ("intra-day") sont plutôt des modèles non markoviens, tels que, par exemple, les modèles à volatilité stochastique.

1.3 Différents cadres asymptotiques.

L'étude et la comparaison des estimateurs n'est possible, en règle générale, que dans un cadre asymptotique. En statistique des diffusions, il existe plusieurs approches qui diffèrent suivant les propriétés des modèles ou suivant les observations dont on dispose.

1. Pour une observation continue, on pourra faire tendre la longueur T de l'intervalle d'observation vers l'infini. En-dehors d'hypothèses d'ergodicité sur le modèle, il sera difficile de faire une théorie générale. En revanche, les conditions assurant l'ergodicité sont très simples pour les diffusions unidimensionnelles.

2. Pour une observation discrétisée, on fera tendre le nombre n d'observations vers l'infini. Dans ce cas, une distinction devenue classique est la suivante.

   (a) Soit on suppose que le pas de discrétisation Δ est fixe. C'est le point de vue qui paraît le plus naturel. Le pas de discrétisation représente l'échelle de temps: observations mensuelles, hebdomadaires, journalières, etc. Toutefois, ce point de vue est celui qui soulève le plus de difficultés théoriques (vraisemblance exacte non explicite) et restreint aussi l'étude aux modèles ergodiques. (Les modèles non ergodiques ne font pas l'objet d'une théorie générale et sont à traiter au cas par cas).

   (b) Soit on suppose que, simultanément, $n \to +\infty$ et $\Delta = \Delta_n \to 0$. Au concret, cela correspond à affiner de plus en plus l'échelle de temps: jour, minute, seconde, etc. A ce stade, une autre distinction, devenue classique également, est apparue, liée aux propriétés des processus d'Ito.

      i. On peut supposer que $T = n\Delta_n$ reste fixe (observations de plus en plus rapprochées à l'intérieur d'un intervalle de temps de longueur fixe),

      ii. ou que $T = T_n = n\Delta_n$ tend vers l'infini (observations de plus en plus proches et durée totale de l'observation de plus en plus grande).

   Selon le cas, on identifiera les paramètres inconnus présents dans le coefficient de diffusion, même sans connaissance de la dérive, avec peu d'hypothèses sur le modèle, ou tous les paramètres, à condition d'imposer des hypothèses sur le modèle (comme ci-dessus, ce sont des hypothèses d'ergodicité qui permettront une étude théorique générale).

3. On peut supposer que l'on observe n trajectoires i.i.d. $(\xi^{(i)}(t), t \leq T), i = 1, \dots, n$ en continu ou en discrétisé, avec T fixé et considérer l'asymptotique classique $n \to +\infty$.

4. Enfin, soit $\varepsilon > 0$ un "petit" paramètre, et supposons que le coefficient de diffusion soit égal à $\varepsilon\sigma(t, x)$ et connu. Le modèle étudié apparaît comme une "petite" perturbation de l'équation différentielle ordinaire

$$dx(t) = b(t, x(t), \theta_0)dt.$$

   Le cadre asymptotique est alors $\varepsilon \to 0$. C'est un cadre classique comme les précédents, dont l'un des avantages est qu'il requiert peu d'hypothèses sur le modèle. Le cas d'une observation continue de la trajectoire sur un intervalle de longueur fixe est traité de façon exhaustive dans les ouvrages de Kutoyants (1986 et 1994). Le cas de l'observation discrétisée est étudié dans quelques articles. Nous avons choisi de ne pas le détailler ici. On peut noter cependant que le cas de la petite variance s'apparente au cas où l'observation est une moyenne d'observations indépendantes et identiquement distribuées de processus de diffusion. Voyons l'exemple particulier suivant.

Exemple: Considérons $X^1, \dots X^n$ n processus réels indépendants, de même loi, régis par l'équation différentielle stochastique linéaire:

$$dX_t^i = (a(t)X_t^i + b(t))dt + dW_t^i \quad X_0^i = x$$

où $W^1, \dots W^n$ sont n mouvements browniens indépendants et les fonctions $a(t), b(t)$ sont déterministes. Alors, le processus $\xi_t = \frac{1}{n}(X_t^1 + \dots + X_t^n)$ vérifie:

$$d\xi_t = (a(t)\xi_t + b(t))\ dt + \frac{1}{\sqrt{n}} dB_t^n \quad \xi_0 = x$$

où $B_t^n = \frac{1}{n}(W_t^1 + \dots + W_t^n)$ est un mouvement brownien standard.

Ainsi, (ξ_t) est régi par une équation différentielle stochastique de petite variance $\varepsilon = \frac{1}{\sqrt{n}}$.

Notons que toute contrainte sur le modèle se traduit par une restriction de l'espace des paramètres comme on le verra sur les exemples.

1.4 Exemples de modèles

Les processus de diffusion sont utilisés dans de nombreux domaines et, en particulier, en finance.

1.4.1 Actifs financiers

Un modèle classique, pour le prix des actifs financiers (actions, stocks, indices boursiers) est celui du mouvement brownien géométrique. Si (S_t) désigne le prix de l'actif à l'instant t, on postule que

$$dS_t = S_t(rdt + \sigma dW_t).$$

Lorsque r et σ sont des constantes, c'est le modèle de Black et Scholes (1973). Pour mieux représenter la réalité des cours, les constantes r et σ ont été remplaçées par des fonctions $r = r(t, s)$ et $\sigma = \sigma(t, s)$. La quantité $\sigma(t, S_t)$ est la volatilité de l'actif, et $r(t, S_t)$ le taux de rendement. Ces fonctions contiennent des paramètres inconnus à estimer et l'on dispose des observations des prix cotés.

1.4.2 Taux d'intérêt

Pour modéliser les taux d'intérêt, diverses equations différentielles stochastiques ont été proposées. En particulier, pour les taux à court terme, si r_t désigne un tel taux à la date t, une famille de modèles classiques est la suivante:

$$dr_t = \alpha(\beta - r_t)dt + \sigma r_t^\delta dWt.$$

Le cas $\delta = 0$ est le modèle de Vasicek(1977, 2016) , le cas $\delta = 1/2$ est celui de Cox-Ingersoll-Ross (1985). Ces équations font donc apparaître les paramètres $\alpha, \beta, \sigma, \delta$.

1.4.3 Equations différentielles stochastiques en biologie.

Les modélisations par des processus de diffusion existent dans de nombreux domaines autres que la finance.

1.5 Modèles multidimensionnels

Les modèles multidimensionnels peuvent aussi être envisagés. L'étude statistique en est plus difficile, surtout lorsqu'on aborde les propriétés asymptotiques. La caractérisation de l'ergodicité et le calcul des distributions stationnaires sont les difficultés qui n'existent pas en dimension 1.

Dans (1.1), on lira une équation vectorielle où le processus ξ appartient à $\mathbb{R}^d$, le mouvement brownien W est m-dimensionnel, la fonction $b(t, x, \theta)$ est définie sur $\mathbb{R}^+ \times \mathbb{R}^d \times \theta$ à valeurs dans $\mathbb{R}^d$, la fonction $\sigma(t, x, \theta)$ définie sur le même ensemble que b est à valeurs dans $\mathbb{R}^d \otimes \mathbb{R}^m$.

1.6 Prérequis.

Pour aborder ce cours, il est nécessaire de connaître les outils principaux du calcul stochastique: mouvement brownien, intégrales stochastiques d'Ito, formule d'Ito, martingales à temps continu, formule de Girsanov, notions sur les équations différentielles stochastiques. Ce cours pourra ainsi apprendre à manipuler ces outils dans un but statistique. Des résultats qui ne sont pas toujours étudiés dans un cours de base de calcul stochastique, seront fournis en compléments au fur et à mesure des nécessités.

1.7 Références bibliographiques

1. Basawa, I.V. et Prakasa Rao, B.L.S., 1980. *Statistical inference for stochastic processes*, Academic press.

2. Black, F., Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, 637-654.

3. Cox, J.C., Ingersoll, J.E. and Ross, S.A., 1985. A theory of the term structure of interest rates, *Econometrica*, **53**, 385-407.

4. Gikhman, I.I. et Skorohod, A.V. *The theory of stochastic processes, I*, Springer, 1974, *II*, Springer, 1975, *III*, Springer, 1979.

5. Ibragimov, I.A. et Has'minskii, R.Z., 1981. *Statistical estimation, Asymptotic theory*, Springer.

6. Ikeda, N. et Watanabe, S., 1981. *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha.

7. Karatsas, I. et Shreeve, S.E., 1991. *Brownian motion and stochastic calculus*, second edition, Graduate texts in Mathematics, Springer.

8. Karlin, S. et Taylor, H.M., 1981. *A second course in stochastic processes*, Academic Press.

9. Kutoyants, Yu.A., 1984. *Parameter estimation for stochastic processes*, Research and Exposition in Mathematics 6, Heldermann.

10. Kutoyants, Yu. A., 2004. *Statistical inference for ergodic diffusion processes*, Springer.

11. Lamberton, D. et Lapeyre, B., 1991. *Introduction au calcul stochastique appliqué à la finance*, Mathématiques et applications, Ellipses, Paris.

12. Liptser, R.S. et Shiryaev, A.N., 1977. *Statistics of random processes I, General Theory*, Springer, Second edition, 2001.

13. Liptser, R.S et Shiryaev, A.N., 1978. *Statistics of random processes II, Applications*, Springer, Second edition, 2001.

14. Prakasa Rao, B.L.S., 1999. *Statistical inference for diffusion type processes*, London and Oxford University press.

15. D. Revuz et Marc Yor, 1991. *Continuous martingales and Brownian motion*, Springer, Berlin.

16. Vasicek, O.A., 1977. An equilibirium characterization of the term structure, *Journal of Financial Economics*, **5**, 177-188.

17. Vasicek, O.A., 2016. *Finance, Economics, and Mathematics.* Wiley, Hoboken, New Jersey.

Chapter 2

Deux classes d'exemples explicites

Les deux (classes de) modèles suivants présentent l'intérêt d'être totalement explicites. On peut faire tous les calculs correspondants à une observation discrétisée, ce qui n'est pas le cas général, et illustrer ainsi les méthodes statistiques et les cadres asymptotiques.

2.1 Mouvement brownien avec dérive

2.1.1 Description du modèle

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $(W_t, t \geq 0)$ un mouvement brownien standard sur Ω. On observe le processus

$$\xi_t = \theta_0 t + \sigma_0 W_t$$

aux instants déterministes $t_i = i\Delta, i = 1, \dots, n$ et l'on posera $T = n\Delta$. Les paramètres inconnus sont $(\theta_0, \sigma_0) \in \mathbb{R} \times (0, +\infty)$. (Il est impossible d'identifier le signe de σ_0, nous convenons ici que ce paramètre est positif. De fait, le paramètre naturel que l'on estimera est σ_0^2). Le processus (ξ_t) n'est pas défini par une équation différentielle stochastique, son expression en fonction de (W_t) est totalement explicite.

2.1.2 Modèle statistique canonique, vraisemblance.

Le modèle statistique canonique associé à l'observation $(\xi_{t_i}, i = 1, \dots, n)$ est défini par:

- l'espace des observations, c'est-à-dire $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$,
- la loi de probabilité des observations, c'est-à-dire la loi de probabilité sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ du vecteur $(\xi_{t_i}, i = 1, \dots, n)$, que nous noterons $P^n_{\theta_0,\sigma_0}$.

Cette loi dépend du couple inconnu (θ_0, σ_0) et ne connaissant pas la "vraie valeur" (θ_0, σ_0), nous sommes amenés à considérer la famille de lois $P^n_{\theta,\sigma}$, indexée par (θ, σ), qui est celle du vecteur $(\xi^{\theta,\sigma}_{t_i}, i = 1, \dots, n)$ défini par:

$$\xi^{\theta,\sigma}_t = \theta t + \sigma W_t.$$

D'où le modèle canonique $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)), P^n_{\theta,\sigma}), (\theta, \sigma) \in \mathbb{R} \times (0, +\infty)$. Nous pouvons à présent calculer une vraisemblance du modèle. Rappelons :

Définition 2.1.1. *On appelle fonction de vraisemblance toute fonction*

$$(\theta,\sigma) \to L(\theta,\sigma) = \frac{dP^n_{\theta,\sigma}}{d\mu}(\xi_{t_i}, i = 1,\dots,n)$$

où μ est une mesure positive et σ-finie dominant toutes les lois $P^n_{\theta,\sigma}$ et $(\xi_{t_i}, i = 1,\dots,n)$ représente l'échantillon observé.

La loi $P^n_{\theta,\sigma}$ se calcule aisément en utilisant la propriété que les v.a. $(\xi^{\theta,\sigma}_{t_i} - \xi^{\theta,\sigma}_{t_{i-1}}, i = 1,\dots,n)$ sont indépendantes et identiquement distribuées de loi gaussienne $\mathcal{N}(\theta\Delta, \sigma^2\Delta)$ (on a posé $t_0 = 0$). La densité de $P^n_{\theta,\sigma}$ par rapport à la mesure de Lebesgue λ de R^n vaut:

$$\frac{dP^n_{\theta,\sigma}}{d\lambda}(x_i, i = 1,\dots,n) = \frac{1}{\sigma^n(2\pi\Delta)^{n/2}}\exp(-\frac{1}{2\sigma^2\Delta}\sum_{i=0}^{n-1}(x_{i+1} - x_i - \theta\Delta)^2)$$

(On a posé $x_0 = 0$). En prenant comme mesure dominante la mesure $\mu = \frac{1}{(2\pi\Delta)^{n/2}}\lambda$, on obtient une fonction de vraisemblance égale à:

$$L(\theta,\sigma) = \exp \ell(\theta,\sigma)$$

avec

$$\ell(\theta,\sigma) = -\frac{1}{2}(n\log(\sigma^2) - \frac{1}{2\sigma^2\Delta}\sum_{i=0}^{n-1}(\xi_{t_{i+1}} - \xi_{t_i} - \theta\Delta)^2).$$

Définition 2.1.2. *On appelle estimateur du maximum de vraisemblance toute solution $(\hat{\theta},\hat{\sigma})$ de l'équation:*

$$L(\hat{\theta},\hat{\sigma}) = \sup_{\theta\in\mathbb{R},\sigma>0)} L(\theta,\sigma).$$

Il est plus facile de maximiser la fonction $\ell(\theta,\sigma)$. La résolution du système :

$$\partial\ell/\partial\theta = 0, \quad \partial\ell/\partial\sigma^2 = 0$$

aboutit à l'unique solution : (avec $T = n\Delta$)

$$\hat{\theta} = \frac{\xi_T}{T} \tag{2.1}$$

$$\hat{\sigma^2} = \frac{1}{T}\sum_{i=0}^{n-1}(\xi_{t_{i+1}} - \xi_{t_i} - \hat{\theta}\Delta)^2 = \frac{1}{T}\sum_{i=0}^{n-1}(\xi_{t_{i+1}} - \xi_{t_i})^2 - \hat{\theta}^2\Delta. \tag{2.2}$$

On vérifie aisément qu'en ce point l est maximum. Pour ce faire, on peut introduire la fonction $\Phi(x) = x - 1 - \log x, x > 0$, qui est positive et nulle si et seulement si $x = 1$, et l'on obtient:

$$\ell(\theta,\sigma) = -\frac{n}{2}(\Phi(\frac{\hat{\sigma^2}}{\sigma^2}) + 1 + \log\hat{\sigma^2} + \frac{\Delta}{\sigma^2}(\hat{\theta} - \theta)^2) \le \ell(\hat{\theta},\hat{\sigma}).$$

Cette étude élémentaire fournit, dans ce cas, des estimateurs explicites et dont on peut calculer la loi. Ce ne sont bien sûr pas les seuls estimateurs possibles. Par exemple, on peut considérer pour estimer σ_0^2 :

$$\tilde{\sigma^2} = \frac{1}{T}\sum_{i=0}^{n-1}(\xi_{t_{i+1}} - \xi_{t_i})^2 \tag{2.3}$$

2.1.3 Comportement asymptotique des estimateurs du maximum de vraisemblance.

Dans ce paragraphe, nous étudions la loi des estimateurs calculés ci-dessus et leur comportement asymptotique suivant les différents cadres asymptotiques présentés en introduction.

Δ fixe, n (et donc $T = n\Delta$) tend vers l'infini

Proposition 2.1.1. *Les estimateurs (2.1) et (2.2) sont fortement consistants, c'est-à-dire:*

$$\hat{\theta} \to \theta_0, \quad \hat{\sigma^2} \to \sigma_0^2, \quad p.s..$$

De plus, $T^{1/2}(\hat{\theta} - \theta_0)$ a la loi gaussienne $\mathcal{N}(0, \sigma_0^2)$, et $n^{1/2}(\hat{\sigma^2} - \sigma_0^2)$ converge en loi vers la loi $\mathcal{N}(0, 2\sigma_0^4)$. Enfin, on a la convergence en loi jointe de $(T^{1/2}(\hat{\theta} - \theta_0), n^{1/2}(\hat{\sigma^2} - \sigma_0^2))$ vers la loi $\mathcal{N}(0, \sigma_0^2) \otimes \mathcal{N}(0, 2\sigma_0^4)$

Preuve. Pour $\hat{\theta}$, on écrit :

$$\hat{\theta} = \theta_0 + \sigma_0 \frac{W_T}{T}.$$

On sait que $\frac{W_T}{T}$ converge p.s. vers 0 lorsque T tend vers $+\infty$ (voir appendice) et que $\frac{W_T}{T^{1/2}}$ suit, pour tout T, la loi $\mathcal{N}(0, 1)$.

Pour $\hat{\sigma^2}$, introduisons les variables

$$\varepsilon_i = \frac{W_{t_i} - W_{t_{i-1}}}{\Delta^{1/2}}, i = 1, \ldots, n$$

qui sont indépendantes et de même loi $\mathcal{N}(0, 1)$. On obtient, en remplaçant ξ_{t_i} par $\theta_0 t_i + \sigma_0 W_{t_i}$, après simplifications:

$$\hat{\sigma^2} = \sigma_0^2 \left(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 - \Delta(\frac{W_T}{T})^2 \right).$$

Remarquons que $E(\varepsilon_i^2) = 1$ et $E(\varepsilon_i^4) = 3$, d'où $\mathrm{Var}(\varepsilon_i^2) = 2$. La loi des grands nombres implique:

$$\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 \to 1 \quad \text{p.s..}$$

On en déduit la consistance forte de $\hat{\sigma^2}$. De plus,

$$n^{1/2}(\hat{\sigma^2} - \sigma_0^2) = \sigma_0^2 n^{1/2}(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 - 1) - \sigma_0^2(\Delta)^{1/2} \frac{W_T}{T^{1/2}} \frac{W_T}{T}.$$

Dans la somme ci-dessus, le second terme tend en probabilité vers 0 et le premier terme converge en loi vers la loi $\mathcal{N}(0, 2\sigma_0^4)$. D'où le résultat (voir appendice).

Pour la convergence en loi jointe, il suffit d'appliquer le théoreème de limite centrale bi-dimensionnel à $(\frac{1}{n^{1/2}} \sum_{i=1}^{n} \varepsilon_i, (n^{1/2}(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 - 1))$. $\square$

Remarquons que $\tilde{\sigma^2} = \hat{\sigma^2} + \Delta\hat{\theta}^2$ converge vers $\sigma_0^2 + \Delta\theta_0^2$ et n'est donc pas dans ce cadre asymptotique un estimateur consistant de σ_0^2.

$\Delta = \Delta_n$ tend vers 0 quand n tend vers l'infini et T fixe (et donc $\Delta_n = T/n$)

La loi de $\hat{\theta}$ est la loi $\mathcal{N}(\theta_0, \frac{\sigma_0^2}{T})$ qui est une loi fixe (ne dépendant pas de n). L'estimateur du maximum de vraisemblance de θ_0 n'a donc ici aucune propriété asymptotique. En revanche, le paramètre σ_0^2 est estimé de façon consistante.

Proposition 2.1.2. *Les estimateurs $\hat{\sigma^2}$ et $\tilde{\sigma^2}$ sont (faiblement) consistants: ils convergent en probabilité vers σ_0^2 quand n tend vers l'infini. De plus, $n^{1/2}(\hat{\sigma^2} - \sigma_0^2)$ converge en loi vers la loi $\mathcal{N}(0, 2\sigma_0^4)$ et $n^{1/2}(\tilde{\sigma^2} - \hat{\sigma^2})$ converge en probabilité vers 0. Par conséquent, les deux estimateurs sont asymptotiquement équivalents et ont la même loi asymptotique.*

Preuve. Etudions directement la convergence en loi (le résultat impliquant la consistance). Notons désormais ε_i^n les v.a. ε_i de la preuve précédente:

$$\varepsilon_i^n = \frac{W_{t_i^n} - W_{t_{i-1}^n}}{\Delta_n^{1/2}} \quad t_i^n = i\Delta_n = iT/n.$$

Ces v.a. forment un tableau triangulaire de v.a. indépendantes et de même loi $\mathcal{N}(0,1)$. Le théorème de limite centrale s'applique encore pour donner la même convergence en loi:

$$Z_n = n^{1/2}(\frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i^n)^2 - 1) \to \mathcal{N}(0,2).$$

Le résultat de la proposition s'obtient alors en remarquant que:

$$n^{1/2}(\hat{\sigma^2} - \sigma_0^2) = \sigma_0^2(Z_n - \frac{W_T^2}{n^{1/2}}),$$

et

$$n^{1/2}(\hat{\sigma^2} - \tilde{\sigma^2}) = -\hat{\theta}^2\frac{T}{n^{1/2}}$$

□

Il faut noter que le résultat de consistance découle directement de la propriété de la variation quadratique du mouvement brownien. La consistance forte peut aussi être prouvée au prix d'une étude un peu plus fine.

$\Delta = \Delta_n$ tend vers 0 et $T = T_n = n\Delta_n$ tend vers l'infini quand n tend vers l'infini

Enonçons seulement le résultat dont nous laisserons la preuve en exercice.

Proposition 2.1.3. *Les estimateurs $\hat{\theta}$ et $\hat{\sigma^2}$ sont consistants et le couple $\Big((n\Delta_n)^{1/2}(\hat{\theta} - \theta_0), n^{1/2}(\hat{\sigma^2} - \sigma_0^2)\Big)$ converge en loi vers la loi $\mathcal{N}(0, \sigma_0^2) \otimes \mathcal{N}(0, 2\sigma_0^4)$*

Pour prouver le résultat, on fera appel au théorème de limite centrale bidimensionnel appliqué au couple $(\varepsilon_i^n, (\varepsilon_i^n)^2 - 1))$.

Il faut noter que le couple $(\hat{\theta}, \tilde{\sigma^2})$ vérifie le même résultat de convergence en loi que $(\hat{\theta}, \hat{\sigma^2})$ à condition d'imposer une contrainte supplémentaire sur la convergence du pas de discrétisation qui est $n\Delta_n^2 \to 0$. Ainsi, si $\Delta_n = n^{-\alpha}$, on doit avoir $0 < \alpha < 1$ pour que $n\Delta_n$ tende vers l'infini et la vitesse de $\hat{\theta}$ est $n^{(1-\alpha)/2}$, plus lente que $\sqrt{n}$. Avec la contrainte supplémentaire, on doit avoir $\alpha > 1/2$, et la vitesse de $\hat{\theta}$ est plus lente que $n^{1/4}$. La vitesse des deux estimateurs de σ est toujours $\sqrt{n}$.

Les trois cadres asymptotiques que l'on vient de voir peuvent être repris pour des modèles de diffusions plus généraux. Les résultats dépendront cependant le plus souvent des propriétés des modèles.

2.1.4 Prix des actifs

Dans le célèbre modèle de Black et Scholes, le prix d'un actif financier (ou d'un stock) est modélisé par un processus (S_t) de la forme :

$$dS_t = S_t(\mu_0 dt + \sigma_0 dW_t),$$

où μ_0, σ_0 sont des constantes déterministes. Ceci signifie que (S_t) est un mouvement brownien géométrique et

$$\xi_t = \log S_t = \log S_0 + (\mu_0 - \frac{\sigma_0^2}{2})t + \sigma_0 W_t$$

est un mouvement brownien avec dérive. Les estimateurs du maximum de vraisemblance de $\theta_0 = \mu_0 - \frac{\sigma_0^2}{2}$ et σ_0^2 basés sur l'observation discrétisée (ξ_{t_i}) ont été calculés plus haut. On en déduit:

$$\hat{\sigma^2} = \frac{1}{T}\sum_{i=0}^{n-1}(\xi_{t_{i+1}} - \xi_{t_i})^2 - \Delta(\frac{\xi_T}{T})^2 \qquad \hat{\theta} = \frac{\xi_T}{T} + \frac{\hat{\sigma^2}}{2}$$

2.1.5 Généralisations

La méthode de calcul de la vraisemblance et des estimateurs du maximum de vraisemblance dans le modèle du mouvement brownien avec dérive peut être généralisée à l'ensemble des modèles gaussiens de la forme :

$$\xi_t = x + \int_0^t \mu(s, \theta_0)ds + \int_0^t \sigma(s, \theta_0)dW_s \tag{2.4}$$

où la condition initiale $x \in \mathbb{R}$ est connue, les fonctions $(t, \theta) \to \mu(t, \theta)$, $(t, \theta) \to \sigma(t, \theta)$ sont déterministes connues et telles que, pour tout $t \geq 0$ et tout $\theta \in \Theta$,

$$\int_0^t |\mu(s, \theta)|ds \ < \ +\infty, \quad \int_0^t \sigma^2(s, \theta)ds \ < \ +\infty.$$

Toutefois, on ne pourra pas, en général, obtenir d'estimateurs ayant une expression explicite en fonction des observations. Le calcul par une méthode numérique sera possible dans la mesure où la vraisemblance est explicite. Mais, l'étude asymptotique devra être faite de façon implicite uniquement à l'aide de la fonction de vraisemblance. De plus, la plupart des résultats asymptotiques (convergence ou existence de loi asymptotique lorsque n tend vers l'infini) dépendront de chaque forme de fonction de dérive et de coefficient de diffusion.

2.1.6 Simulations de trajectoires

Lorsqu'on ne dispose pas de données réelles, ou pour valider la méthode d'estimation que l'on a choisie, il est intéressant de simuler les données. Ceci est particulièrement vrai lorsqu'on sait faire une simulation exacte de la trajectoire observée. C'est le cas pour les modèles gaussiens du paragraphe précédent qui sont des processus à accroissements indépendants gaussiens. En effet, pour simuler $(\xi_{t_i}, i = 1, \ldots n)$ lorsque (ξ_t) est défini par (2.4), si $(\varepsilon_1, \ldots, \varepsilon_n)$ est un n-échantillon de la loi gaussienne centrée réduite, il suffit de construire:

$$\xi_0 = x, \quad \xi_{t_i} = \xi_{t_{i-1}} + \int_{t_{i-1}}^{t_i} \mu(s, \theta_0)ds + (\int_{t_{i-1}}^{t_i} \sigma^2(s, \theta_0)ds)^{1/2}\varepsilon_i, \quad i = 1, \ldots, n.$$

2.1.7 Exercices

Calculer et, dans la mesure du possible, étudier, les estimateurs du maximum de vraisemblance basés sur l'observation $(\xi_{t_i}, i = 1, \ldots, n)$ (où $t_i = i\Delta$) dans les modèles suivants.

- $d\xi_t = \theta_0 t dt + \sigma_0 dW_t, \quad \xi_0 = 0.$
- $d\xi_t = \theta_0 t dt + \sigma_0 t^{1/2} dW_t, \quad \xi_0 = 0.$
- $d\xi_t = \theta_0 dt + \sigma_0(1 + t^2)^{1/2} dW_t, \quad \xi_0 = 0.$

Reprendre les mêmes hypothèses asymptotiques que plus haut.

2.2 Processus d'Ornstein-Uhlenbeck

2.2.1 Description du modèle

Considérons le processus défini par l'équation différentielle stochastique :

$$d\xi_t = \theta\xi_t dt + \sigma dW_t \quad \xi_0 = \eta, \tag{2.5}$$

où $(W_t, t \geq 0)$ est un mouvement brownien standard, défini sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$, η est une v.a. indépendante de W, de loi $\nu_{(\theta,\sigma)}$ et $\theta \in \mathbb{R}$, $\sigma > 0$. Ce modèle est particulièrement riche et intéressant sur le plan statistique et l'on peut l'étudier sous tous les angles car tous les calculs sont explicites (ce qui est exceptionnel pour un modèle de diffusion).

Résolution de l'équation différentielle stochastique. Loi du processus solution.

Parmi les rares modèles pour lesquels l'équation différentielle stochastique admet une solution explicite en fonction de η et W, on trouve les équations dont la dérive $(b(t, x))$ et le coefficient de diffusion $(\sigma(t, x))$ sont linéaires en la variable d'état x. Le processus d'Ornstein-Uhlenbeck en fait partie avec la propriété supplémentaire que sa loi est complètement calculable.

Proposition 2.2.1. *Le processus défini par l'équation (2.5) s'écrit:*

$$\xi_t = \eta \exp{(\theta t)} + X_t \quad \textit{avec} \quad X_t = \sigma \exp{(\theta t)} \int_0^t \exp{(-\theta s)} dW_s. \tag{2.6}$$

En particulier, (X_t) est la solution de l'équation (2.5) correspondant à la condition initiale $X_0 = 0$. De plus, pour tout $t \geq 0, \Delta \geq 0$, la v.a.

$$Z_{t,\Delta} = \xi_{t+\Delta} - \exp{(\theta\Delta)}\xi_t$$

est indépendante de $\mathcal{F}_t = \sigma(\eta, W_s, s \leq t)$, et de loi gaussienne $\mathcal{N}(0, \sigma^2 v(\theta))$ où

$$v(\theta) = \frac{\exp 2\theta\Delta - 1}{2\theta} \tag{2.7}$$

Preuve. En s'appuyant sur l'équation différentielle ordinaire $dx_t = \theta x_t dt$, on cherche ξ_t de la forme $\xi_t = \exp{(\theta t)}\ C_t$, $C_0 = \xi_0 = \eta$. Par la formule d'Ito, on a:

$$dC_t = d(\exp{(-\theta t)}\ \xi_t) = -\theta \exp{(-\theta t)}\ \xi_t dt + \exp{(-\theta t)}\ d\xi_t = \sigma \exp{(-\theta t)}\ dW_t.$$

D'où, $C_t = \eta + \sigma \displaystyle\int_0^t \exp{(-\theta s)}\ dW_s$, ce qui donne le résultat voulu.

Pour la seconde partie de l'énoncé, il suffit de découper:

$$\xi_{t+\Delta} = \eta \exp{(\theta(t+\Delta))} + \sigma \exp{(\theta(t+\Delta))}\ (\int_0^t \ldots + \int_t^{t+\Delta} \ldots).$$

On en déduit:

$$Z_{t,\Delta} = \sigma \exp{(\theta(t+\Delta))}\ \int_t^{t+\Delta} \exp{(-\theta s)}\ dW_s.$$

Le résultat d'indépendance découle de l'indépendance de η et W et des propriétes du mouvement brownien (l'intégrale stochastique de t à $t+\Delta$ est indépendante de $(W_s, s \leq t)$). Pour le calcul de loi, on sait que $Z_{t,\Delta}$ est une v.a. gaussienne centrée et de variance:

$$\sigma^2 \exp{(2(\theta+\Delta))}\ \int_t^{t+\Delta} \exp{(-2\theta s)} ds.$$

Un calcul élémentaire montre que cette quantité vaut $\sigma^2 v(\theta)$ avec $v(\theta)$ donné dans l'énoncé. □

Corollaire 2.2.1. *Si η est de loi gaussienne $\mathcal{N}(m, \tau^2)$ (et, en particulier, déterministe), le processus (ξ_t) est gaussien avec, pour $t \geq 0, \Delta \geq 0$:*

$$E(\xi_t) = m \exp{(\theta t)}, \quad Var(\xi_t) = \tau^2\ \exp{(2\theta t)} + \sigma^2\ \frac{\exp{(2\theta t)} - 1}{2\theta}, \quad Cov(\xi_{t+\Delta}, \xi_t) = \exp{(\theta\Delta)}\ Var(\xi_t).$$

Si $\theta < 0$, et si η a la loi $\mathcal{N}(0, \frac{\sigma^2}{-2\theta})$, le processus est stationnaire strict.

Preuve. Soient, pour $k \geq 1$, $(a_1, \ldots, a_k)$ des nombres réels et $(t_1, \ldots, t_k)$ des réels positifs ou nuls. On a:

$$S_k = \sum_{i=1}^{k} a_i \, \xi_{t_i} = \alpha_k \eta + T_k,$$

où $\alpha_k = \displaystyle\sum_{i=1}^{k} \exp(\theta t_i) \, a_i$ et

$$T_k = \sum_{i=1}^{k} a_i \, X_{t_i} = \int_0^{+\infty} f_k(s) \, dW_s$$

avec

$$f_k(s) = \sigma \sum_{i=1}^{k} a_i \exp(\theta(t_i - s)) \, 1_{[0,t_i]}(s).$$

Ainsi, il apparaît que T_k est une v.a. gaussienne (et donc le processus (X_t) est gaussien), et S_k, étant la somme de deux v.a. gaussiennes indépendantes, est gaussienne. Le processus (ξ_t) est gaussien. La v.a. η et le processus (X_t) sont indépendants.

Pour calculer l'espérance et la variance de ξ_t, on utilise la formule (2.6) et l'indépendance de η et X_t. Pour calculer la covariance de ξ_t et $\xi_{t+\Delta}$, on utilise l'indépendance de ξ_t et $Z_{t,\Delta}$.

Si $\theta < 0$, et si, pour les paramètres de la loi de η, on prend $m = 0$ et $\tau^2 = \frac{\sigma^2}{-2\theta}$, on constate que ξ_t a pour tout t la loi de η et que la covariance de ξ_t et $\xi_{t+\Delta}$ ne dépend que de Δ. Le processus (ξ_t) étant gaussien, ceci suffit à prouver qu'il est stationnaire strict. $\square$

Corollaire 2.2.2. *Soient, pour $i \in N$, $t_i = i\Delta$ et $Y_i = \xi_{t_i}$. La suite (Y_i) est un processus auto-régressif d'ordre 1 (AR(1)) tel que (cf (2.7)):*

$$Y_{i+1} = \exp(\theta\Delta) \, Y_i + \sigma v^{1/2}(\theta) \, \varepsilon_{i+1}, \tag{2.8}$$

où $(\varepsilon_i, i \geq 1)$ est une suite de v.a. i.i.d. de loi $\mathcal{N}(0,1)$, indépendante de $Y_O = \eta$.

Si $\theta < 0$, et si η a la loi $\mathcal{N}(0, \frac{\sigma^2}{-2\theta})$, le processus (Y_i) est stationnaire strict et ergodique.

Preuve. Il suffit de poser

$$\varepsilon_i = \frac{Z_{t_i,\Delta}}{\sigma \, v^{1/2}(\theta)} \tag{2.9}$$

La stationnarité se déduit immédiatement de la stationnarité de (ξ_t). L'ergodicité résulte de ce que (Y_i) est gaussien et $\mathrm{Cov}(Y_i, Y_{i+n}) = \exp(n\theta\Delta) \, \frac{\sigma^2}{-2\theta}$ tend vers 0 lorsque n tend vers l'infini. $\square$

Comportement asymptotique

La proposition suivante décrit le comportement asymptotique de la v.a. ξ_t lorsque t tend vers l'infini.

Proposition 2.2.2.
- *Si $\theta < 0$, ξ_t converge en loi, lorsque t tend vers l'infini vers la loi gaussienne $\mathcal{N}(0, \frac{\sigma^2}{-2\,\theta})$.*

- *Si $\theta > 0$, $\xi_t \ \exp(-\theta t)$ converge dans $L^2(P)$ vers la v.a. $Z = \eta + \sigma \displaystyle\int_0^{+\infty} \exp(-\theta t)\ dW_t$.*

Preuve. On a $\xi_t = \exp(\theta t)\ \eta + X_t$ (cf Proposition (2.2.1)). Si $\theta < 0$, le premier terme tend vers 0 p.s.. Il suffit alors de prouver que X_t converge en loi vers la loi voulue.

Or, pour tout t, X_t est gaussienne. Par suite, X_t converge en loi si et seulement si les limites, lorsque t tend vers l'infini, de EX_t et $\mathrm{Var}X_t$ existent et sont finies. Et, dans ce cas, la loi limite de X_t est la loi gaussienne de moyenne la limite des EX_t, de variance la limite des $\mathrm{Var}X_t$. On a:

$$EX_t = 0 \quad \text{et} \quad \mathrm{Var}X_t = \sigma^2 \frac{1 - \exp(2\theta t)}{-2\theta},$$

d'où le résultat.

Si $\theta > 0$, la v.a. Z de l'énoncé est bien définie, dans $L^2(P)$, puisque $\int_0^{+\infty} \exp(-2\theta t) dt = (2\theta)^{-1} < \infty$. Et

$$E(\xi_t \exp(-\theta t) - Z)^2 = \sigma^2\ E(\int_t^{+\infty} \exp(-\theta s)\ dW_s)^2 = \int_t^{+\infty} \exp(-2\theta s) ds = \exp(-2\theta t)$$

tend vers 0. $\square$

Remarquons que, si $\theta = 0$, on a tout simplement $\xi_t = \eta + W_t$.

2.2.2 Observation discrétisée: modèle statistique canonique, vraisemblance.

Notons $\xi_t^{\theta,\sigma}$ le processus solution de l'équation (2.5), avec la condition initiale $\eta = x$ déterministe et connue (ce que nous supposerons ici pour simplifier). Supposons que l'on observe le processus $\xi_t = \xi_t^{\theta_0,\sigma_0}$ aux instants $(t_i = i\Delta, i = 1, \ldots, n)$ et l'on posera $T = n\Delta$. Les paramètres inconnus sont $(\theta_0, \sigma_0) \in \mathbb{R} \times (0, +\infty)$. La construction du modèle statistique canonique associé à l'observation $(\xi_{t_i}, i = 1, \ldots, n)$ suit le même plan que plus haut. Il est défini par $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_{\theta,\sigma}^n), (\theta, \sigma) \in \mathbb{R} \times (0, +\infty)$ où

- $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ est l'espace des observations,
- $P_{\theta,\sigma}^n$ est la loi de probabilité du vecteur $(\xi_{t_i}^{\theta,\sigma}, i = 1, \ldots, n)$.

La vraie valeur inconnue du paramètre est (θ_0, σ_0) représentée par le processus observé (ξ_t). Nous pouvons à présent calculer une vraisemblance du modèle.

Proposition 2.2.3. *Une fonction de vraisemblance associée à l'observation $(\xi_{t_i}, i = 1, \ldots, n)$ est définie par:*

$$L(\theta, \sigma) = L(\theta, \sigma; \xi_{t_i}, i = 1, \ldots, n) = \exp \ell(\theta, \sigma)$$

avec

$$\ell(\theta, \sigma) = -\frac{n}{2}(\log(\sigma^2) + \log(v(\theta))) - \frac{1}{2\sigma^2 v(\theta)} \sum_{i=0}^{n-1} (\xi_{t_{i+1}} - \exp(\theta\Delta)\ \xi_{t_i})^2,$$

et $v(\theta)$ est donné en (2.7).

Preuve. En prenant pour mesure dominante $\mu = (2\pi)^{-(n/2)} dx_1 \dots dx_n$, on calcule la densité $\frac{dP^n_{\theta,\sigma}}{d\mu}$ en utilisant le fait que les v.a. :

$$\xi^{\theta,\sigma}_{t_{i+1}} - \exp(\theta\Delta)\xi^{\theta,\sigma}_{t_i}$$

sont i.i.d. de loi $\mathcal{N}(0, \sigma^2 \, v(\theta))$. $\Box$

Les estimateurs du maximum de vraisemblance de (θ_0, σ_0) peuvent être calculés explicitement. Pour simplifier les calculs, il est commode de poser :

$$a = \exp(\theta\Delta) \quad \rho^2 = \sigma^2 \, v(\theta),$$

qui sont les paramètres naturels du modèle AR(1) et sont en correspondance biunivoque (et en C^1-difféomorphisme) avec les paramètres initiaux ($(a, \rho^2) \in (0, \infty) \times (0, \infty) \to (\theta, \sigma^2) \in (0, \infty) \times (0, \infty)$ est une bijection C^1 ainsi que sa bijection réciproque). On résout sans difficulté le système

$$\frac{\partial \ell}{\partial a} = 0 \quad \frac{\partial \ell}{\partial \rho^2} = 0.$$

Et l'on vérifie aisément que l'unique couple solution correspond bien à un maximum pour l. Les formules suivantes sont classiques pour les estimateurs de (a_0, ρ_0^2):

$$\hat{a} = \frac{\sum_{i=1}^n \xi_{t_{i-1}}\xi_{t_i}}{\sum_{i=1}^n \xi^2_{t_{i-1}}}, \quad \hat{\rho}^2 = \frac{1}{n}\sum_{i=1}^n \left(\xi_{t_i} - \hat{a}\xi_{t_{i-1}}\right)^2.$$

En termes de paramètres initiaux, on obtient:

$$\hat{\theta} = \frac{1}{\Delta} \log\left(\frac{\sum_{i=1}^n \xi_{t_{i-1}}\xi_{t_i}}{\sum_{i=1}^n \xi^2_{t_{i-1}}}\right),$$

et

$$\hat{\sigma}^2 = \frac{1}{nv(\hat{\theta})}\sum_{i=1}^n \left(\xi_{t_i} - \exp(\hat{\theta}\Delta)\xi_{t_{i-1}}\right)^2.$$

On voit que $\hat{\theta}$ n'est défini que lorsque l'argument du logarithme est positif (alors que $\hat{a}$ est toujours bien défini, p.s.).

La loi exacte des estimateurs ici obtenus est plus difficile à étudier que dans le modèle de mouvement brownien avec dérive. Il est nécessaire de faire une étude asymptotique. Reprenons les trois cadres asymptotiques décrits plus haut.

- Pas Δ fixe et n tend vers l'infini : cette étude est classique s'agissant d'un modèle à temps discret de type AR(1). On sait étudier le comportement asymptotique des estimateurs de (a_0, ρ_0^2) suivant les trois cas : $a_0 < 1$ (cas ergodique, noter que dans notre modèle, a_0 est forcément positif), $a_0 = 1$ (cas de la marche aléatoire), $a_0 > 1$ (cas explosif). Rappelons brièvement les étapes de cette étude.

On obtient aisément les relations, en posant $Y_i = \xi_{t_i}$(cf (2.9)):

$$\hat{a} - a_0 = \rho_0 \frac{\sum_{i=1}^n \varepsilon_i \, Y_{i-1}}{\sum_{i=1}^n Y_{i-1}^2},$$

et

$$n^{1/2}(\hat{\rho}^2 - \rho_0^2) = \frac{1}{n^{1/2}} \sum_{i=1}^n (\varepsilon_i^2 - 1) + R_n(Z_n),$$

où $Z_n = (\sum_{i=1}^n \varepsilon_i \;\; Y_{i-1}, \sum_{i=1}^n Y_{i-1}^2)$ et les v.a. ε_i sont celles du corollaire (2.2.2) pour les valeurs (a_0, ρ_0^2). L'étude (non triviale) du couple Z_n convenablement normalisé permet de montrer que:

- si $a_0 < 1$, $n^{1/2}(\hat{a} - a_0)$ converge en loi vers une loi gaussienne centrée;
- si $a_0 = 1$, $n(\hat{a} - 1)$ converge en loi vers une loi non gaussienne (celle de $\frac{\int_0^1 B_s dB_s}{\int_0^1 B_s^2 ds}$, avec (B_s) mouvement brownien standard);
- si $a_0 > 1$, $a_0^n(\hat{a} - a_0)$ converge en loi vers une loi de Cauchy (la loi de ε/Z où ε est gaussienne centrée réduite, Z est gaussienne centrée et ε et Z sont indépendantes).

Dans les trois cas, ces résultats permettent de vérifier que le terme $R_n(Z_n)$ de la formule $n^{1/2}(\hat{\rho}^2 - \rho_0^2)$ converge en probabilité vers 0. De ce fait, on a toujours, pour $n^{1/2}(\hat{\rho}^2 - \rho_0^2)$, une convergence en loi vers une loi gaussienne $\mathcal{N}(0, 2\rho_0^4)$. Pour en déduire les résultats correspondants pour $\hat{\theta}$ et $\hat{\sigma}^2$, il suffit d'appliquer la formule de Taylor à la fonction f telle que $(\theta, \sigma^2) = f(a, \rho^2)$ (cf appendice). De cette application, on déduira que $\hat{\theta}$ hérite de la vitesse de $\hat{a}$ et que $\hat{\sigma}^2$ a presque exactement le même comportement asymptotique que $\hat{\rho}^2$ savoir $n^{1/2}(\hat{\sigma}^2 - \sigma_0^2)$ converge en loi vers $\mathcal{N}(0, 2\sigma_0^4)$.

- Pas $\Delta = \Delta_n$ tendant vers 0 et longueur $T = n\Delta_n$ fixe de l'intervalle d'observation :

Donnons quelques détails qui seront complétés dans les chapitres ultérieurs. Regardons tout d'abord $\hat{\theta}$. Puisque T est fixe, on a: (ici, $t_i = t_i^n = i\frac{T}{n}$)

$$\sum_{i=1}^n \xi_{t_{i-1}}(\xi_{t_i} - \xi_{t_{i-1}}) \to \int_0^T \xi_s \, d\xi_s \quad \text{et} \quad \sum_{i=1}^n \xi_{t_{i-1}}^2 \, \Delta_n \to \int_0^T \xi_s^2 ds \tag{2.10}$$

La première convergence (vers l'intégrale stochastique) a lieu en probabilité (propriété d'approximation discrète d'une intégrale stochastique; de fait, on peut vérifier directement la convergence dans $L^2(\mathbb{P})$). La seconde (vers l'intégrale ordinaire) a lieu p.s.. Introduisons l'estimateur

$$\tilde{\theta} = \frac{\sum_{i=1}^n \xi_{t_{i-1}}(\xi_{t_i} - \xi_{t_{i-1}})}{\sum_{i=1}^n \xi_{t_{i-1}}^2 \, \Delta_n}$$

qui converge en probabilité vers la v.a. (bien définie car son dénominateur est strictement positif p.s.)

$$\theta_T = \frac{\int_0^T \xi_s \, d\xi_s}{\int_0^T \xi_s^2 ds}.$$

On peut écrire:

$$\hat{\theta} = \frac{1}{\Delta_n} \log(1 + \Delta_n \tilde{\theta}) \sim \tilde{\theta} \sim \theta_T,$$

où l'équivalent est en probabilité. On voit que la probabilité que $\hat{\theta}$ soit bien défini tend vers 1, mais que les estimateurs $\hat{\theta}$, $\tilde{\theta}$ ne sont pas consistants. Nous retrouverons θ_T comme estimateur de θ basé sur l'observation en temps continu de la trajectoire (ξ_t) sur tout l'intervalle $[0, T]$.

Il est possible d'étudier directement $\hat{\sigma}^2$. La preuve de la consistance est facile, la loi asymptotique est plus difficile à obtenir. Plutôt que de faire cela, introduisons l'estimateur $\tilde{\sigma}^2$ basé sur la variation quadratique :

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_{i=1}^{n} (\xi_{t_i} - \xi_{t_{i-1}})^2.$$

Cet estimateur sera étudié avec précision plus loin et l'on montrera que $n^{1/2}(\tilde{\sigma}^2 - \sigma_0^2)$ converge en loi vers la loi $\mathcal{N}(0, 2\sigma_0^4)$. Or, en utilisant les convergences (2.10), il est facile de montrer que:

$$n^{1/2}(\tilde{\sigma}^2 - \hat{\sigma}^2) = \tilde{\sigma}^2 n^{-1/2} 0_P(1) + n^{-1} 0_P(1),$$

où $0_P(1)$ signifie une v.a. bornée en probabilité (ici, qui converge en probabilité). Par conséquent, les deux estimateurs sont, dans ce cadre asymptotique, asymptotiquement équivalents et $\hat{\sigma}^2$ a même loi asymptotique (recentré et renormalisé) que $\tilde{\sigma}^2$.

- Pas $\Delta = \Delta_n$ tendant vers 0 et longueur $T = n\Delta_n$ tendant vers l'infini: ce cas nécessite plus de travail et de résultats préliminaires. L'étude en est prématurée et reportée plus loin.

2.2.3 Extensions du modèle et simulations de trajectoires

Supposons que l'on observe le processus défini par:

$$d\xi_t = (a(t, \theta_0)\xi_t + b(t, \theta_0))dt + \sigma(t, \theta_0)dW_t, \quad \xi_0 = \eta$$

avec η v.a. réelle indépendante de (W_t), et où les fonctions $a(t, \theta), b(t, \theta), \sigma(t, \theta)$ sont des fonctions déterministes connues de (t, θ) telles que, pour tout (t, θ):

$$\int_0^t (|a(s, \theta)| + |b(s, \theta)| + \sigma^2(s, \theta))ds < \infty.$$

On peut recommencer toute l'étude faite aux paragraphes précédents. Toutefois, les calculs et les expressions obtenues seront plus complexes.

Posons $a(t, \theta_0) = a(t)$, $b(t, \theta_0) = a(t)$, $\sigma(t, \theta_0) = a(t)$, $A(t) = \int_0^t a(s)ds$, $\alpha_i = \dfrac{A(t_i)}{A(t_{i-1})}$,

$$m_i = \exp A(t_i) \left(\int_{t_{i-1}}^{t_i} b(s) \exp(-A(s))\, ds \right),$$

$$\sigma_i = \exp A(t_i) \left(\int_{t_{i-1}}^{t_i} \sigma^2(s) \exp\left(-2A(s)\right) ds \right)^{1/2}.$$

On a:

$$\xi_t = \exp A(t) \left(\eta + \int_0^t \exp\left(-A(s)\right)(b(s)\, ds + \sigma(s)\, dW_s) \right),$$

et:

$$\xi_{t_i} = m_i + \alpha_i\, \xi_{t_{i-1}} + \sigma_i \varepsilon_i$$

où $(\varepsilon_i, i = 1, \ldots, n)$ est un n-échantillon de la loi $\mathcal{N}(0,1)$. Ces formules permettent de calculer la densité jointe de (ξ_{t_i}) et de simuler ce n-uple.

Pour conclure, insistons sur le fait que les calculs exacts effectués sur les deux classes d'exemples du chapitre ne pourront être reproduits pour d'autres modèles. Ces calculs reposent en effet sur la connaissance explicite de la loi jointe du n-uple (ξ_{t_i}).

2.2.4 Exercices

Exercice 1. Résoudre l'équation différentielle stochastique $d\xi_t = \theta_0 f(t)\xi_t dt + \sigma_0 dW_t, \quad \xi_0 = 0$. Etudier la loi du processus solution. Calculer les estimateurs du maximum de vraisemblance de (θ_0, σ_0^2) basés sur l'observation $(\xi_{t_i}, i = 1, \ldots, n)$ avec $t_i = i\Delta$ lorsque $f(t) = t$.

Exercice 2. Résoudre $d\xi_t = \alpha_0(\gamma_0 - \xi_t)dt + \sigma_0 dW_t, \quad \xi_0 = \eta$, avec η v.a. gaussienne indépendante de (W_t). Etudier la loi du processus solution. Donner la loi de la v.a. η qui conduise à un processus stationnaire. Calculer les estimateurs du maximum de vraisemblance de $(\alpha_0, \gamma_0, \sigma_0^2)$ basés sur une observation discrétisée de pas Δ comme ci-dessus. (On pourra poser $a_0 = \alpha_0\gamma_0, b_0 = -\gamma_0$ et calculer d'abord les e.m.v. de ces paramètres).

Exercice 3. (Exercice de simulation) Simuler une discrétisation de la trajectoire du processus des exercices précédents et calculer sur les données simulées, les différents estimateurs vus dans le cours ou obtenus dans la résolution pour tous les paramètres. Faire des répétitions des simulations pour obtenir des moyennes et écarts-types estimés. Commenter les résultats. Il conviendra de tenter différentes valeurs des paramètres, du pas de la discrétisation et du nombre d'observations.

2.3 Appendice

Dans ce chapitre, nous avons fait appel aux notions de:

- Processus gaussien, processus à accroissements indépendants.
- Mouvement brownien standard, intégrales de Wiener.
- Processus stationnaire strict, au sens L^2.

Rappelons de plus quelques résultats qui ont été utiles dans le chapitre.

Résultat 2.3.1. *Si* $(W_t, t \geq 0)$ *est un mouvement brownien standard, alors* $\frac{W_t}{t}$ *tend p.s. vers* 0 *quand* t *tend vers l'infini.*

Résultat 2.3.2. *Si* $t_i = t_i^n = iT/n$, *avec* $T > 0$ *fixé, alors* $\displaystyle\sum_{i=1}^{n}(W_{t_i} - W_{t_{i-1}})^2$ *converge dans* $L^2(P)$ *vers* T.

Résultat 2.3.3. *(Théorème de Slutsky) Si* X_n, Y_n *sont deux suites de variables aléatoires réelles telles que* X_n *converge en loi vers (la loi de)* X *et* Y_n *converge en probabilité vers* 0*, alors* $X_n + Y_n$ *converge en loi vers* X *et* $X_n Y_n$ *converge en probabilité vers* 0.

Résultat 2.3.4. *(Convergence en loi d'une suite de v.a. gaussiennes). Soit* (X_n)*, une suite de v.a.réelles telle que, pour tout* n, X_n *est de loi gaussienne* $\mathcal{N}(m_n, \sigma_n^2)$*. Alors, la suite* (X_n) *converge en loi si et seulement si les deux suites* (m_n) *et* (σ_n^2) *convergent dans* $\mathbb{R}$*. Dans ce cas, la loi limite de la suite* (X_n) *est la loi gaussienne* $\mathcal{N}(m, \sigma^2)$ *où* m *est la limite de* (m_n) *et* σ^2 *est la limite de* (σ_n^2)*. Si* $\sigma^2 = 0$*, la loi limite est la loi de la v.a. constante* m.

Voici un résultat d'utilisation très courante en statistique.

Résultat 2.3.5. *Soit* X_n *une suite de v.a. à valeurs dans* $\mathbb{R}^d$*,* m *un vecteur fixe (déterministe) de* R^d *et* $\varphi(n)$ *une suite de nombres positifs tendant vers l'infini avec* n*. On suppose que* $\varphi(n)(X_n - m)$ *converge en loi vers la loi d'un vecteur aléatoire* Z*. Soit* $f : \mathbb{R}^d \to \mathbb{R}^l$ *une fonction de classe* C^1*. Alors,* $\varphi(n)(f(X_n) - f(m))$ *converge en loi vers la loi de* $Df(m)Z$ *(produit de la matrice jacobienne* $Df(m)$ *par le vecteur* Z*).*

L'énoncé ci-dessus suppose que toutes les composantes de X_n convergent en loi à la même vitesse et se démontre par une simple application de la formule de Taylor et du théorème de Slutsky. Il est possible d'étendre ce résultat au cas où les différentes composantes de $X_n - m$ convergent en loi à des vitesses distinctes. Dans cette hypothèse, on normalise le vecteur $X_n - m$ en le multipliant à gauche par une matrice diagonale dont les termes diagonaux $\varphi_i(n)$ sont les vitesses respectives des composantes. Pour trouver la limite en loi correspondante de $f(X_n) - f(m)$, par application de la formule de Taylor, on devra étudier chaque composante pour faire apparaître une vitesse résultante (la plus lente de celles de la combinaison linéaire des composantes de $X_n - m$ fournie par l'ordre 1 du développement de Taylor).

Rappelons quelques définitions concernant l'ergodicité des processus stationnaires à temps discret (On se limitera ici à l'indexation par les entiers positifs). Notons τ l'opérateur de translation de $R^{\mathbb{N}}$ défini par : $\tau((y_i, i \in \mathbb{N}) = (y_{i+1}, i \in \mathbb{N})$. Un événement B de la tribu $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ est dit invariant (par τ) si $\tau^{-1}(B) = B$. Un processus $(Y_i, i \in \mathbb{N})$ à valeurs réelles stationnaire strict est dit ergodique si, pour tout événement B invariant, on a $\mathbb{P}(Y \in B) = 0$ ou 1.

Résultat 2.3.6. *(Ergodicité des processus gaussiens stationnaires) Un processus réel gaussien* (Y_i) *stationnaire (strict ou* L^2*) est ergodique si et seulement si Cov(*Y_i, Y_{i+n}*) (qui ne dépend que de* n*) tend vers* 0 *lorsque* n *tend vers l'infini.*

Il s'ensuit la possibilité d'appliquer le théorème ergodique.

Résultat 2.3.7. *(Théorème ergodique) Soit* (Y_i) *un processus réel stationnaire strict et ergodique et* $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ *une fonction borélienne telle que* $\mathbb{E}|f((Y_i))| < \infty$. *Alors, lorsque* n *tend vers l'infini,*

$$\frac{1}{n}\sum_{k=0}^{n} f((Y_{i+k})) \to Ef((Y_i)) \quad p.s..$$

Les théorèmes de limite centrale associés font appel à la notion de coéfficient de mélange et ne sont pas rappelés ici. (voir, par exemple, P. Hall et C.C. Heyde, 1980, P. Doukhan, 1994).

2.4 Références bibliographiques

1. Hall, P. et Heyde, C.C., 1980. *Martingale limit theory and its application*, Academic press.

2. Doukhan, P., 1994. *Mixing, Properties and Examples*, Springer, Lecture Notes in Statistics.

3. Cottrell, M., Genon-Catalot, V., Duhamel, C. et Meyre, T., 1999.*Exercices de probabilités, licence, maîtrise, école d'ingénieurs*, Cassini. Nouvelle édition: 2005.

Chapter 3

Estimation d'un paramètre du coefficient de diffusion. Observation discrète de pas tendant vers 0 sur un intervalle de temps fixe.

Dans ce chapitre, nous abordons l'étude de l'estimation d'un paramètre inconnu présent dans le coefficient de diffusion à partir d'une observation discrétisée de la trajectoire du processus de pas tendant vers 0 sur un intervalle de temps de longueur fixe. Avec une telle observation, il n'est pas possible d'estimer des paramètres inconnus de dérive et la dérive peut d'ailleurs être supposée connue ou inconnue. La dérive apparaît plutôt comme un paramètre nuisible au sens statistique, c'est-à-dire un terme inconnu mais qui n'intéresse pas le statisticien. C'est ainsi souvent le cas pour les modèles de la finance: le paramètre d'intérêt, sur lesquels portent les efforts de modélisation concrète, est le coefficient de diffusion. Ce terme est interprétable par les utilisateurs comme une mesure (chiffrée) de risque. Il sert pour le calcul de prix d'actifs dérivés (options, par exemple), pour l'obtention de courbes des taux d'intérêt, alors que la dérive ne sert pas pour ces calculs.

Le problème statistique est difficile. On ne peut pas, en général, calculer de vraisemblance explicite. D'autre part, l'étude théorique de la vraisemblance (Donahl, 1987, Gobet, 2001) montre que les lois asymptotiques des estimateurs du maximum de vraisemblance ne sont pas gaussiennes sauf cas particuliers (le modèle statistique est LAMN). On peut cependant obtenir facilement des estimateurs consistants soit de type empirique soit obtenus en minimisant un contraste approprié.

Dans un premier temps, nous allons étudier le cas d'un paramètre linéaire dans le coefficient de diffusion pour lequel on obtient des estimateurs explicites avec lois asymptotiques gaussiennes. Le modèle visé est donc le suivant:

$$d\xi_t = b(t, \xi_t)dt + \sigma h(\xi_t)dW_t, \quad \xi_0 = \eta, \tag{3.1}$$

où $b(t, x)$ est une fonction inconnue, σ est le paramètre inconnu à estimer, la fonction h est connue. Les observations sont prises aux instants $t_i = t_i^n = iT/n, i = 0, \ldots, n,$ avec T positif connu.

Puis, nous donnerons des exemples d'estimateurs consistants dans un cadre plus général. L'ensemble requiert des compléments de calcul stochastique par lesquels nous commençons.

3.1 Compléments

Rappelons tout d'abord quelques définitions.

On suppose donné un espace de probabilité filtré satisfaisant *aux conditions habituelles* (de Dellacherie et Meyer) noté $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Définition 3.1.1. *Le processus $(W_t, t \geq 0)$ défini sur Ω est un $\mathbb{P}$-$(\mathcal{F}_t)$ mouvement brownien si:*

1) $(W_t, t \geq 0)$ est un mouvement brownien standard,

2) $(W_t, t \geq 0)$ est adapté à $(\mathcal{F}_t)$,

3) Pour tout (s, t) avec $0 \leq s \leq t$, $W_t - W_s$ est indépendant de $\mathcal{F}_s$.

Exemple. Un exemple usuel s'obtient de la façon suivante. Soit η une v.a. réelle définie sur Ω indépendante de $(W_t, t \geq 0)$ et, pour $t \geq 0$, soit $(\mathcal{F}_t) = \sigma(\eta, W_s, s \leq t))$. Il est facile de vérifier que (W_t) est un $(\mathcal{F}_t)$ mouvement brownien (exercice).

Si (W_t) est un $\mathbb{P}$-$(\mathcal{F}_t)$ mouvement brownien, alors, (W_t) et $(W_t^2 - t)$ sont des $\mathbb{P}$-$(\mathcal{F}_t)$ martingales. La réciproque est vraie. En fait, la caractérisation de Paul Lévy donne un résultat plus fort. Nous devons rappeler la définition suivante.

Définition 3.1.2. *Le processus (M_t) défini sur Ω trajectoires continues, adapté à $(\mathcal{F}_t)$, est une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ s'il existe une suite croissante (T_n) de temps d'arrêt de $(\mathcal{F}_t)$ telle que le processus arrêté $(M_{t\wedge T_n}, t \geq 0)$ soit une vraie martingale relative à $\mathbb{P}$-$(\mathcal{F}_t)$. (La suite de temps d'arrêt est dite "localisante").*

Exemple. Voici un exemple de suite localisante pour une martingale locale (M_t) trajectoires continues. Considrer $T_n = \inf\{t \geq 0, |M_t| \geq n\}$. De plus, pour tout n, la martingale $(M_{t\wedge T_n})$ est uniformément bornée.

Théorème 3.1.1. *(Caractérisation de Paul Lévy) Le processus (W_t) est un $\mathbb{P}$-$(\mathcal{F}_t)$ mouvement brownien si et seulement si (W_t) et $(W_t^2 - t)$ sont des $\mathbb{P}$-$(\mathcal{F}_t)$ martingales locales trajectoires continues nulles en 0.*

Le théorème suivant est connu sous le nom d'inégalités de domination de Lenglart.

Théorème 3.1.2. *Soit $(M_t)_{t\geq 0}$ une martingale locale relative à $(\mathcal{F}_t)$, continue et nulle en 0. Alors pour tout $T > 0$, (T déterministe ou temps d'arrêt relatif à $(\mathcal{F}_t)$), pour tous $h, \eta > 0$, on a :*

$$P(\sup_{t\leq T} |M_t| > h) \leq \frac{\eta}{h^2} + P(< M >_T > \eta) \tag{3.2}$$

Corollaire 3.1.1. *Soit $(M_t^n)_{t\geq 0}$ une suite de martingales locales relatives à $(\mathcal{F}_t)$, continues et nulles en 0. Si $< M^n >_T$ converge en probabilité vers 0 lorsque n tend vers l'infini, alors $\sup_{t\leq T} |M_t^n|$ converge en probabilité vers 0.*

Proposition 3.1.1. *(Approximation discrète de l'intégrale stochastique) Soient (H_t) un processus continu à gauche, adapté à $(\mathcal{F}_t)$ et $(M_t)_{t\geq 0}$ une martingale locale continue relative à $(\mathcal{F}_t)$. Pour $T > 0$ fixé, posons $t_i = t_i^n = iT/n$, $i = 0, \ldots, n$. On a, en probabilité, lorsque n tend vers l'infini:*

$$\sum_{i=1}^{n} H_{t_{i-1}}(M_{t_i} - M_{t_{i-1}}) \to \int_0^T H_t dM_t$$

Théorème 3.1.3. *(Variation quadratique des processus d'Ito) Soit (X_t) un processus à valeurs dans $\mathbb{R}^d$ vérifiant, pour $i = 1, \ldots, d$:*

$$dX_t^i = b_t^i \, dt + \sum_{l=1}^{m} a_t^{il} \, dW_t^l$$

où $W = (W^1, \ldots, W^m)$ est un mouvement brownien de $\mathbb{R}^m$, et les processus b^i, a^{il} sont continus et adaptés. Soit (H_t) un autre processus continu adapté. On a, en probabilité, lorsque n tend vers l'infini:

$$\sum_{i=1}^{n} H_{t_{i-1}}(X_{t_i}^j - X_{t_{i-1}}^j)(X_{t_i}^k - X_{t_{i-1}}^k) \to \int_0^T H_t(a_t a_t^*)_{jk} \, dt$$

(a^* désigne la matrice transposée de la matrice a).

Preuve. Nous supposons que $d = 1 = m$ pour simplifier. On a:

$$(X_{t_i} - X_{t_{i-1}})^2 = X_{t_i}^2 - X_{t_{i-1}}^2 - 2X_{t_{i-1}}(X_{t_i} - X_{t_{i-1}}).$$

Par la formule d'Ito, on en déduit:

$$(X_{t_i} - X_{t_{i-1}})^2 = 2\int_{t_{i-1}}^{t_i} (X_t - X_{t_{i-1}})dX_t + \int_{t_{i-1}}^{t_i} a_t^2 dt.$$

Par suite, $\sum_{i=1}^{n} H_{t_{i-1}}(X_{t_i} - X_{t_{i-1}})^2 = A_1 + A_2 + A_3$ où:

$$A_1 = 2\sum_{i=1}^{n} H_{t_{i-1}} \int_{t_{i-1}}^{t_i} (X_t - X_{t_{i-1}})b_t dt, \quad A_2 = 2\int_0^T Z_t^n dW_t, \quad A_3 = \sum_{i=1}^{n} H_{t_{i-1}} \int_{t_{i-1}}^{t_i} a_t^2 \, dt$$

avec

$$Z_t^n = \sum_{i=1}^{n} 1_{]t_{i-1}, t_i]}(t) \; H_{t_{i-1}}(X_t - X_{t_{i-1}})a_t.$$

Posons

$$\rho_n(X) = \sup_{s,t\in[0,T], |t-s|\leq T/n} |X_t - X_s|.$$

Le processus (X_t) étant à trajectoires continues, son module de continuité $\rho_n(X)$ sur $[0, T]$ ci-dessus tend vers 0 en probabilité lorsque n tend vers l'infini. On obtient donc la majoration suivante pour A_1:

$$|A_1| \;\leq\; U_T \rho_n(X)$$

où $U_T = 2T\sup_{s\leq T}|H_s b_s|$ est une variable aléatoire fixe, bien définie grâce à la continuité des processus (H_s) et (b_s). On a donc $A_1 \to 0$.

Pour prouver que A_2 tend vers 0, il suffit de prouver que $\int_0^T (Z_s^n)^2 ds$ tend vers 0. Or, on a, pour ce terme, une majoration analogue à celle de A_1. En effet,

$$\int_0^T (Z_s^n)^2 ds \ \leq \ V_T \rho_n^2(X)$$

avec $V_T = T\sup_{s\leq T} H_s^2 a_s^2$. Enfin, A_3, par les propriétés classiques des intégrales de Stieltjes-Riemann, converge vers la limite cherchée $\int_0^T H_s\, a_s^2 ds$. □

3.2 Estimation d'un paramètre linéaire dans le coefficient de diffusion

On suppose donné un espace de probabilité filtré satisfaisant aux conditions habituelles noté $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ et $(W_t, t \geq 0)$ est un $\mathbb{P}$-$(\mathcal{F}_t)$ mouvement brownien défini sur Ω.

Dans ce paragraphe, on suppose que les fonctions $(t,x) \to b(t,x)$ et $x \to h(x)$ sont définies et continues respectivement sur $\mathbb{R}^+ \times \mathbb{R}$ et $\mathbb{R}$ et telles que l'équation différentielle stochastique (3.1) admette un unique processus solution défini sur l'espace $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ du mouvement brownien (W_t) et adapté à la filtration $(\mathcal{F}_t)$. Pour $T > 0$ fixé, l'observation est celle du n+1-uple $(\xi_{t_i}, 0 \leq i \leq n)$. La fonction b peut être connue ou non, la fonction h est connue.

3.2.1 Coefficient de diffusion constant

Nous étudions d'abord le cas $h = 1$. Considérons l'estimateur de σ^2 défini par

$$\tilde{\sigma}^2 = \frac{1}{T}\sum_{i=1}^n (\xi_{t_i} - \xi_{t_{i-1}})^2 \tag{3.3}$$

Théorème 3.2.1. *L'estimateur (3.3) est (faiblement) consistant et on a la convergence en loi:*

$$n^{1/2}(\tilde{\sigma}^2 - \sigma^2) \to \mathcal{N}(0, 2\sigma^4)$$

Preuve. Montrons directement la convergence en loi (qui implique la consistance). Posons $H_s = b(s, \xi_s)$. On a:

$$(\xi_{t_i} - \xi_{t_{i-1}}) = \int_{t_{i-1}}^{t_i} H_s ds + \sigma(W_{t_i} - W_{t_{i-1}}).$$

Donc,

$$n^{1/2}(\tilde{\sigma}^2 - \sigma^2) = \sigma^2 n^{1/2}\left(\frac{1}{T}\sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 \ - \ 1\right) + A + B + C,$$

avec

$$\begin{aligned} A &= n^{1/2}\frac{1}{T}\sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} H_s ds\right)^2, \quad B = n^{1/2}\frac{2\sigma}{T}\frac{T}{n}\sum_{i=1}^{n} H_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}), \\ C &= n^{1/2}\frac{2\sigma}{T}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(H_s - H_{t_{i-1}})ds(W_{t_i} - W_{t_{i-1}}). \end{aligned}$$

On a déjà vu que

$$n^{1/2}\left(\frac{1}{T}\sum_{i=1}^{n}(W_{t_i} - W_{t_{i-1}})^2 \ - \ 1\right) \to \mathcal{N}(0,2).$$

Il reste donc à montrer que les trois termes A, B, C tendent en probabilité vers 0. Pour A, on remarque:

$$0 \le A \le n^{-1/2} Z_T$$

où $Z_T = T\sup_{s\le T} H_s^2$ est une v.a. fixe. Pour B, on a:

$$B = n^{-1/2}2\sigma(\int_0^T H_s dW_s + o_P(1)),$$

où $o_P(1)$ est un terme tendant vers 0 en probabilité. Pour C, on applique l'inégalité de Schwarz,

$$|C| \le 2\sigma\rho_n(H)\left(\sum_{i=1}^{n}(W_{t_i} - W_{t_{i-1}})^2\right)^{1/2}$$

où $\rho_n(H) = \sup_{s,t\le T, |s-t|\le T/n}|H_t - H_s|$ tend vers 0 et la variation quadratique du mouvement brownien sur $[0,T]$ tend vers T. □

On a ainsi obtenu un estimateur de σ sans connaissance de la dérive $b(t,x)$. Si cette fonction est connue, on peut envisager d'autres estimateurs comme par exemple:

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{i=1}^{n}(\xi_{t_i} - \xi_{t_{i-1}} - \frac{T}{n}b(t_{i-1}, \xi_{t_{i-1}}))^2.$$

(Cet estimateur est asymptotiquement équivalent à $\tilde{\sigma}^2$ au sens où $n^{1/2}(\tilde{\sigma}^2 - \hat{\sigma}^2)$ tend vers 0 en probabilité).

3.2.2 Paramètre linéaire

Dans ce paragraphe, nous ne supposons plus $h = 1$, mais nous rajoutons les hypothèses suivantes.

- La fonction h est de classe C^1 et $\mathbb{P}(h(\xi_t) > 0, \forall t \ge 0) = 1$.

Dans ce cas, pour $F(x) = \int_{x_0}^{x} \frac{1}{h(u)} du$, le processus $Y_t = F(\xi_t) - F(\eta) = \int_{\eta}^{\xi_t} \frac{1}{h(u)} du$ est bien défini et vérifie:

$$dY_t = \sigma dW_t + K_t dt$$

avec

$$K_t = \frac{b(t, \xi_t)}{h(\xi_t)} - \frac{\sigma^2}{2} h'(\xi_t).$$

L'étude faite au paragraphe précédent montre que l'estimateur

$$\bar{\sigma}^2 = \frac{1}{T} \sum_{i=1}^{n} (F(\xi_{t_i}) - F(\xi_{t_{i-1}}))^2. \tag{3.4}$$

a les mêmes propriétés asymptotiques que $\tilde{\sigma}^2$.

Exemples. Déterminer un estimateur de σ^2 dans les modèles suivants.

- $d\xi_t = \alpha(\beta - \xi_t)dt + \sigma \xi_t dW_t, \quad \xi_0 = x_0$ (si $x_0 > 0$, il est possible de déterminer α, β, σ tels que $\xi_t > 0, \forall t \geq 0$ p.s.).
- $d\xi_t = b(t, \xi_t)dt + \sigma(1 + \xi_t^2)dW_t, \quad \xi_0 = x_0$.

3.3 Estimateurs empiriques

Le théorème (3.1.3) permet de construire des estimateurs consistants des paramètres présents dans le coefficient de diffusion sans connaissance de la dérive. Voici quelques exemples. Dans le modèle (3.1), les estimateurs

$$\frac{1}{T} \sum_{i=1}^{n} \frac{(\xi_{t_i} - \xi_{t_{i-1}})^2}{h^2(\xi_{t_{i-1}})} \tag{3.5}$$

et

$$\frac{\sum_{i=1}^{n} (\xi_{t_i} - \xi_{t_{i-1}})^2}{\sum_{i=1}^{n} (T/n) h^2(\xi_{t_{i-1}})} \tag{3.6}$$

sont consistants pour σ^2.

Dans le modèle

$$d\xi_t = b(t, \xi_t)dt + (\theta^2 + \xi_t^2)dW_t, \quad \xi_0 = x_0,$$

l'estimateur

$$(1/T) \sum_{i=1}^{n} \left((\xi_{t_i} - \xi_{t_{i-1}})^2 - (T/n) h^2(\xi_{t_{i-1}}) \right)$$

est un estimateur consistant de θ^2.

3.4 Estimation d'un paramètre du coefficient de diffusion par contraste.

Nous poursuivons l'étude de l'estimation d'un paramètre inconnu présent dans le coefficient de diffusion à partir d'une observation discrétisée de la trajectoire du processus de pas tendant vers 0 sur un intervalle de temps de longueur fixe. Nous considérons à présent un modèle paramétrique général pour le coefficient de diffusion :

$$d\xi_t = b(t,\xi_t)dt + \sigma(t,\xi_t,\theta_0)dW_t, \quad \xi_0 = \eta, \tag{3.7}$$

où $b(t,x)$ est une fonction qui peut être connue ou inconnue, θ_0 est le paramètre inconnu à estimer. Les observations sont prises aux instants $t_i = t_i^n = iT/n, i = 0,\ldots,n$, avec T positif connu et le cadre asymptotique est n tend vers l'infini. Ce cadre asymptotique permet d'estimer θ_0, que la dérive soit connue ou inconnue. En revanche, la dérive ne peut pas être estimée de manière consistante dans ce cadre asymptotique.

3.5 Estimateurs de minimum de contraste.

Pour la plupart des modèles (même avec dérive connue), la vraisemblance exacte de l'observation discrétisée $(\xi_{t_i}, i = 1,\ldots,n)$ n'a pas une forme explicite. Dans un tel cas, une approche classique consiste à construire une fonction des observations et des paramètres inconnus du modèle que l'on utilise en remplacement de la vraisemblance exacte. Une telle fonction s'appelle un contraste (on donnera une définition plus précise plus loin). Pour les processus de diffusion observés de façon discrète, le schéma d'Euler fournit un contraste naturel.

Précisons le modèle. On suppose que l'on observe le processus (ξ_t) défini par:

$$d\xi_t = b(t,\xi_t)dt + \sigma(t,\xi_t,\theta_0)dW_t, \quad \xi_0 = x_0, \tag{3.8}$$

où $\theta_0 \in \Theta \subset \mathbb{R}^p$ est un paramètre inconnu. La fonction de dérive $b(t,x)$ peut être connue ou inconnue et dépendre ou non de θ_0. La fonction $(t,x,\theta) \to \sigma(t,x,\theta)$ a une forme connue. La valeur initiale x_0 est ici supposée connue (ou observée). Le processus (W_t) est un mouvement brownien défini sur l'espace de probabilité $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), P)$, relatif à la filtration $(\mathcal{F}_t)$. Pour estimer θ_0, on observe la trajectoire aux instants $t_i = t_i^n = \frac{iT}{n}$ pour $i = 1,\ldots,n$ avec T fixe.

Le processus observé est l'un des processus du modèle:

$$d\xi_t^\theta = b(t,\xi_t^\theta)dt + \sigma(t,\xi_t^\theta,\theta)dW_t, \quad \xi_0^\theta = x_0, \tag{3.9}$$

celui qui correspond à la vraie valeur θ_0 (inconnue) du paramètre ($\xi = \xi^{\theta_0}$). Nous devons faire des hypothèses sur le modèle.

Nous supposons que:

Hypothèse H1: les fonctions de $\mathbb{R}^+ \times \mathbb{R}$ dans $\mathbb{R}$, $(t,x) \to b(t,x)$ et, pour tout θ, $(t,x) \to \sigma(t,x,\theta)$ sont continues et telles que l'équation différentielle stochastique (3.9) admette un unique processus solution défini sur Ω, adapté à la filtration du mouvement brownien (W_t). De plus, nous supposons que:

Hypothèse H2:

$$\mathbb{P}(\forall t \in [0,T], \sigma(t,\xi_t,\theta) > 0) = 1. \tag{3.10}$$

Des conditions suffisantes pour que ces hypothèses soient réalisées seront précisées plus loin.

Pour simplifier les notations (mais sans restreindre la validité des résultats), nous supposerons que le modèle est autonome pour tout θ, c'est-à-dire que:

$$b(t,x) = b(x), \quad \sigma(t,x,\theta) = \sigma(x,\theta). \tag{3.11}$$

3.5.1 Contraste issu du schéma d'Euler

Considérons le modèle annexe $(Y_0, Y_1, \ldots, Y_n)$ défini par $Y_0 = x_0$ et pour $i = 1, \ldots, n$,

$$Y_i = Y_{i-1} + \Delta_n^{1/2}\sigma(Y_{i-1},\theta)\varepsilon_i, \tag{3.12}$$

où $\Delta_n = T/n$ et

$$\varepsilon_i = \varepsilon_i^n = \frac{W_{t_i} - W_{t_{i-1}}}{\Delta_n^{1/2}}. \tag{3.13}$$

La densité de $(Y_1, \ldots, Y_n)$ est explicite et vaut, à une constante multiplicative près connue, (avec $y_0 = x_0$)

$$L_n(\theta, y_1, \ldots, y_n) = \prod_{i=1}^{n} \sigma^{-1}(y_{i-1},\theta) \exp\left(-\frac{(y_i - y_{i-1})^2}{2\Delta_n \sigma^2(y_{i-1},\theta)}\right). \tag{3.14}$$

Nous pouvons à présent définir le processus suivant, fonction de θ et des observations:

$$U_n(\theta) = -2\log L_n(\theta, (\xi_{t_i}, i = 1, \ldots, n)) = \sum_{i=1}^{n} \log \sigma^2(\xi_{t_{i-1}},\theta) + \sum_{i=1}^{n} \frac{(\xi_{t_i} - \xi_{t_{i-1}})^2}{\Delta_n\, \sigma^2(\xi_{t_{i-1}},\theta)} \tag{3.15}$$

L'estimateur $\hat{\theta}_n$ associé à $U_n(\theta)$ est défini comme toute solution de l'équation

$$U(\hat{\theta}_n) = \inf_{\theta\in\Theta} U_n(\theta). \tag{3.16}$$

3.5.2 Consistance des estimateurs de minimum de contraste

Voyons tout d'abord la propriété de contraste. Posons $\phi(x) = x - 1 - \log x, x > 0$. Cette fonction est positive et nulle si et seulement si $x = 1$.

Proposition 3.5.1. *Lorsque n tend vers l'infini, en probabilité,*

$$\frac{1}{n}(U_n(\theta) - U_n(\theta_0)) \to K_T(\theta_0,\theta)$$

où

$$K_T(\theta_0,\theta) = \frac{1}{T}\int_0^T \phi(\frac{\sigma^2(\xi_s,\theta_0)}{\sigma^2(\xi_s,\theta)})ds. \tag{3.17}$$

Preuve. On étudie, pour θ fixé, la limite de $T\,U_n(\theta)/n$. Le premier terme fait aparaître la somme de Riemann suivante:

$$\sum_{i=1}^{n} \frac{T}{n} \log \sigma^2(\xi_{t_{i-1}}, \theta)$$

qui converge p.s., puisque (ξ_t) est à trajectoires continues, vers

$$\int_0^T \log \sigma^2(\xi_s, \theta) ds.$$

Le second terme s'écrit, puisque $n\Delta_n = T$:

$$\sum_{i=1}^{n} \frac{(\xi_{t_i} - \xi_{t_{i-1}})^2}{\sigma^2(\xi_{t_{i-1}}, \theta)}.$$

Comme (ξ_t) est le processus observé, qui est régi par la valeur θ_0 du paramètre, le terme précédent converge en probabilité vers

$$\int_0^T \frac{\sigma^2(\xi_s, \theta_0)}{\sigma^2(\xi_s, \theta)} ds.$$

En faisant la différence $(U_n(\theta) - U_n(\theta_0))/n$, on fait apparaître la fonction ϕ et la limite cherchée. $\square$

Le résultat de cette proposition nous permet d'introduire une hypothèse dite d'identifiabilité, qui justifie le nom de contraste et permettra de prouver la consistance des estimateurs basés sur $U_n(\theta)$. Nous supposons donc que:
Hypothèse H3 Pour tout $\theta \neq \theta_0$,

$$\mathbb{P}(K_T(\theta_0, \theta) > 0) = 1. \tag{3.18}$$

Vu les propriétés de la fonction ϕ, et la continuité des trajectoires, cette hypothèse est équivalente à

$$\mathbb{P}(\forall s \in [0, T], \sigma^2(\xi_s, \theta_0) = \sigma^2(\xi_s, \theta)) = 0.$$

Il s'agit d'une hypothèse qui semble difficile à vérifier. Cependant, cette hypothèse de nature aléatoire ne doit pas étonner s'agissant d'un modèle statistique où aucune propriété d'ergodicité n'intervient dans le cadre asymptotique.

Pour les estimateurs de minimum de contraste, de même que pour les estimateurs de maximum de vraisemblance, deux situations sont possibles. Ou bien, le calcul par la formule (3.16) fournit des formules explicites pour les estimateurs. On peut alors étudier sur les formules leur comportement asymptotique. Sinon, les estimateurs sont définis implicitement par l'équation (3.16). On peut les calculer numériquement sur chaque jeu de données. Mais, on doit étudier leur comportement par une méthode théorique. Les hypothèses H1-H2-H3 sont insuffisantes. Il faut rajouter des conditions de régularité par rapport au paramètre θ.
Hypothèse H0 Θ est compact.
Hypothèse H4 la fonction $\sigma(x, \theta)$ est continue en $(x, \theta) \in \mathbb{R} \times U$, où U est un ouvert contenant Θ, ses dérivées partielles $\frac{\partial \sigma}{\partial \theta_j}(x, \theta)$, $j = 1, \dots, p$ existent et sont continues en $(x, \theta) \in \mathbb{R} \times U$.

On a alors le résultat.

Théorème 3.5.1. *Sous les hypothèses H0-H4, tout estimateur de minimum de contraste défini par l'équation (3.16) est consistant, c'est-à-dire, que, lorsque n tend vers l'infini,*

$$\forall h > 0, \quad \mathbb{P}(|\hat{\theta}_n - \theta_0| > h) \to 0.$$

La preuve est analogue à celle de Dacunha-Castelle et Duflo (1983, Chapitre 3, paragraphe 3.2.3). Il suffit de l'adapter au cas où la fonction de contraste $K_T(\theta_0, \theta)$ est aléatoire.

Exemples. On admettra que les hypothèses sont satisfaites pour que les modèles suivants soient bien définis.

1. Ecrire le contraste issu du schéma d'Euler lorsque le modèle est défini comme suit, avec σ inconnu:
$$d\xi_t = b(t, \xi_t)dt + \sigma h(\xi_t)dW_t, \quad \xi_0 = x_0. \tag{3.19}$$
En déduire l'estimateur de minimum de contraste associé pour σ^2.

2. Même question avec le modèle
$$d\xi_t = b(t, \xi_t)dt + \sigma(\xi_t)^\alpha dW_t, \quad \xi_0 = x_0, \tag{3.20}$$
avec comme paramètres inconnus (σ, α).

3.5.3 Lois asymptotiques des estimateurs de minimum de contraste.

Avec des hypothèses supplémentaires de régularité par rapport à θ sur la fonction $\sigma(x, \theta)$, il est possible d'obtenir des lois asymptotiques pour les estimateurs de minimum de contraste précédents à la vitesse $\sqrt{n}$. Cependant, ces lois asymptotiques ne sont pas gaussiennes sauf cas particulier (comme les exemples vus au chapitre précédent). On obtient des lois limites pour $\sqrt{n}(\hat{\theta}_n - \theta_0)$ qui sont mélanges en variance de lois gaussiennes. Pour plus de détails, nous renvoyons à Genon-Catalot et Jacod (1993) et à Jacod (1996).

3.6 Références bibliographiques

1. Dacunha-Castelle, D. et Duflo, M., 1983.*Probabilités et statistiques 2, Problèmes à temps mobile*, Masson.

2. Donahl, G. , 1987. On estimating the diffusion coefficient, *J. Appl. Prob.*, **24**, 105-114.

3. Genon-Catalot, V. et Jacod, J. , 1993. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. Henri Poincaré*, **29**, 119-151.

4. Gobet, E., 2001. Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach, *Bernoulli*, **7**, 899-912.

5. Jacod, J. 1996. On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de Probabilités XXXI* Editeurs: Azema, J., Emery, M., Yor, M., Lecture Notes in Mathematics 1655, 232-246.

Chapter 4

Observation en temps continu d'une diffusion unidimensionnelle.

Nous entamons l'étude de l'estimation d'un paramètre inconnu présent dans la dérive à partir d'une observation de la trajectoire en temps continu sur un intervalle de temps $[0,T]$. Nous devons alors supposer connu le coefficient de diffusion: en effet, le coefficient de diffusion est *identifié* lorsqu'on dispose d'une telle observation. Ceci résulte notamment de l'étude faite de l'estimation de paramètres présents dans le coefficient de diffusion à partir d'une observation discrétisée sur un intervalle de temps fixe. Si l'on dispose de toute la trajectoire sur $[0,T]$, l'on est "à l'infini" de l'asymptotique précédente.

Nous considérons à présent un modèle paramétrique pour la dérive :

$$d\xi_t = b(t,\xi_t,\theta_0)dt + \sigma(t,\xi_t)dB_t, \quad \xi_0 = \eta, \tag{4.1}$$

où la fonction $(t,x,\theta) \to b(t,x,\theta)$ a une forme connue, θ_0 est le paramètre inconnu à estimer et la fonction $\sigma(t,x)$ est supposée connue. Le processus (B_t) est un mouvement brownien défini sur l'espace de probabilité $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$, relatif à la filtration $(\mathcal{F}_t)$. La v.a. initiale η a une loi pouvant dépendre de θ_0 et elle est $\mathcal{F}_0$-mesurable. Pour estimer θ_0, on observe la trajectoire (ξ_t) pour $t \in [0,T]$. La vraie valeur du paramètre est θ_0 qui appartient à un ensemble $\Theta \subset \mathbb{R}^p$. Notre but est de définir et étudier des estimateurs de θ_0 fonction de $(\xi_t, t \in [0,T])$.

4.1 Loi de probabilité d'un processus à temps continu.

Pour ce paragraphe, on pourra consulter Billingsley (1968) ou Rogers et Williams (1987, 2000). Nous construisons dans ce paragraphe le modèle statistique canonique associé à l'observation $(\xi_t, t \in [0,T])$. Soit $T > 0$, et notons C_T l'espace des fonctions $x = (x(t), t \in [0,T])$ définies sur $[0,T]$ à valeurs réelles et continues, muni de la topologie de la convergence uniforme. Cette topologie est définie par la distance

$$d(x,y) = \sup_{0 \leq t \leq T} |x(t) - y(t)|. \tag{4.2}$$

On note $\mathcal{C}_T$ la tribu borélienne associée à cette topologie.

4.1.1 Processus canonique.

Pour $t \in [0,T]$, on note X_t l'application coordonnée d'indice t de C_T, définie par:

$$X_t : C_T \to \mathbb{R}, \quad X_t(x) = x(t). \tag{4.3}$$

Pour tout t, X_t est continue comme application de C_T dans $\mathbb{R}$ et donc borélienne: donc, X_t est une v.a.r. définie sur $(C_T, \mathcal{C}_T)$ et $(X_t, t \in [0,T])$ est un processus défini sur $(C_T, \mathcal{C}_T)$. De plus, pour tout $x \in C_T$, $t \to X_t(x) = x(t)$ est continue par construction. Donc, (X_t) est un processus à trajectoires continues, appelé processus canonique. On peut définir la filtration canonique: $\mathcal{C}_t = \sigma(X_s, s \leq t)$, pour tout $t \in [0,T]$.

Théorème 4.1.1. *On a $\mathcal{C}_T = \sigma(X_s, s \in [0,T]) = \sigma(\mathcal{U})$ où $\mathcal{U}$ est la classe des cylindres de $\mathcal{C}_T$, définie par*

$$\mathcal{U} = \{\{x \in C_T, (x(t_1), \dots, x(t_k)) \in A\}, k \geq 1, A \in \mathcal{B}(\mathbb{R}^k), 0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq T\} \tag{4.4}$$

Preuve. Par définition, la tribu $\sigma(X_s, s \in [0,T])$ est exactement la classe des ensembles de la forme $\{x \in C_T, (x(t_1), x(t_2), \dots) \in A\}$ où $A \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ et $(t_1, t_2, \dots, t_n, \dots)$ est une suite de de réels $t_n \in [0,T]$ pour tout n. Ceci permet d'avoir l'égalité $\sigma(X_s, s \in [0,T]) = \sigma(\mathcal{U})$.

Comme chaque application X_s est $\mathcal{C}_T$-mesurable, $\sigma(X_s, s \in [0,T]) \subset \mathcal{C}_T$. Pour l'autre inclusion, on utilise le fait que la topologie de C_T est séparable. Par suite, $\mathcal{C}_T$ est engendrée par les boules $B(x_0, r)$ de la forme

$$B(x_0, r) = \{x \in C_T, d(x_0, x) \leq r\}$$

avec $x_0 \in C_T$ et $r > 0$. La continuité des éléments de C_T permet d'écrire

$$B(x_0, r) = \{x \in C_T, \sup_{q \in \mathbb{Q} \cap [0,T]} |x(q) - x_0(q)| \leq r\}. \tag{4.5}$$

Donc,

$$B(x_0, r) = \bigcap_{q \in \mathbb{Q} \cap [0,T]} \{x \in C_T, |x(q) - x_0(q)| \leq r\} = \bigcap_{q \in \mathbb{Q} \cap [0,T]} \{|X_q - x_0(q)| \leq r\}. \tag{4.6}$$

Ainsi, $B(x_0, r) \in \sigma(X_s, s \in [0,T])$. Ce qui montre que $\mathcal{C}_T \subset \sigma(X_s, s \in [0,T])$. Finalement, on a l'égalité cherchée. $\square$

4.1.2 Loi de probabilité sur l'espace des fonctions continues.

Soit $(\xi_t, t \in [0,T])$ un processus défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs réelles et à trajectoires continues, c'est-à-dire:

i) pour tout ω, $t \to \xi_t(\omega)$ est une fonction continue sur $[0,T]$, soit encore $\xi^T(\omega) = (\xi_t(\omega), t \in [0,T]) \in C_T$,

ii) pour tout $t \in [0,T]$, $\omega \to \xi_t(\omega)$ est une v.a. réelle.

On a:

Proposition 4.1.1. *L'application $\omega \to \xi^T(\omega)$ de Ω dans C_T est $\mathcal{F} - \mathcal{C}_T$ mesurable. Autrement dit, ξ^T est une v.a. à valeurs dans $(C_T, \mathcal{C}_T)$.*

Proof. Comme $\mathcal{C}_T = \sigma(X_s, s \le T)$, ξ^T est $\mathcal{C}_T$- mesurable si et seulement si, pour tout $t \in [0,T], X_t \circ \xi^T = \xi_t$ est mesurable, ce qui est le cas. □

Définition 4.1.1. *La loi de probabilité P^T sur $(C_T, \mathcal{C}_T)$ image de $\mathbb{P}$ par la v.a. ξ^T s'appelle la loi de probabilité du processus $(\xi_t, t \in [0,T])$. Elle est définie par $P^T(B) = \mathbb{P}(\xi^T \in B), B \in \mathcal{C}_T$.*

Exemples:

- Prenons $a_1, a_2, \ldots, a_k$ des réels, et $0 \le t_1 \le t_2 \ldots \le t_k \le T$. Soit $B = \{x \in C_T, x(t_1) \le a_1, \ldots, x(t_k) \le a_k\}$. En utilisant le processus canonique, on peut écrire $B = \{X_{t_1} \le a_1, \ldots, X_{t_k} \le a_k\}$. On a
$$\mathbb{P}(\xi^T \in B) = \mathbb{P}(\xi_{t_1} \le a_1, \ldots, \xi_{t_k} \le a_k) = P^T(X_{t_1} \le a_1, \ldots, X_{t_k} \le a_k).$$
- Soit $f : C_T \to \mathbb{R}$ une application mesurable,
$$\mathbb{E}_{\mathbb{P}} f(\xi^T) = \int_\Omega f(\xi^T) d\mathbb{P} = \int_{C_T} f(x) dP^T(x) = \int_{C_T} f(X) dP^T = E_{P^T} f(X).$$
où $X = (X_t, t \in [0,T])$.

Proposition 4.1.2. *La loi de probabilité P^T est caractérisée par ses répartitions finies: P^T est parfaitement connue si l'on connaît*
$$P^T(X_{t_1} \le a_1, \ldots, X_{t_k} \le a_k),$$
pour tout $k \ge 1$, tout $(a_1, a_2, \ldots, a_k)$ et tout $(t_1, \ldots, t_k) \in [0,T]$.

Proof. Si l'on connaît les fonctions de répartitions des $(X_{t_1}, \ldots, X_{t_k})$ sous P^T, alors la loi de tous les vecteurs $(X_{t_1}, \ldots, X_{t_k})$ est connue. Donc, la probabilité P^T est connue sur la classe des cylindres $\mathcal{U}$. Or cette classe engendre la tribu $\mathcal{C}_T$, est stable par intersection finie et contient une suite croissante d'ensembles B_n telle que $\cup_n B_n = C_T$. De ce fait, deux probabilités P_1 et P_2 qui coïncident sur $\mathcal{U}$ coïncident sur $\mathcal{C}_T$. □

Exemples:

1. Soit $B^T = (B_t, t \in [0,T])$ un mouvement brownien standard défini sur $(\Omega, \mathcal{F}, \mathbb{P})$. La loi de B^T sur $(C_T, \mathcal{C}_T)$ s'appelle la mesure de Wiener sur $[0,T]$, en général notée W^T. Sur $(C_T, \mathcal{C}_T, W^T)$, le processus canonique $(X_t, t \in [0,T])$ est un mouvement brownien standard. En effet, c'est un processus à trajectoires continues et
   - $W^T(X_0 = 0) = \mathbb{P}(B_0 = 0) = 1$,
   - $$W^T(X_{t_1} \le a_1, \ldots, X_{t_k} \le a_k) = \mathbb{P}(B_{t_1} \le a_1, \ldots, B_{t_k} \le a_k).$$
2. Soit $f : \mathbb{R}^+ \to \mathbb{R}$ une fonction localement intégrable et soit $\xi_t^f = \int_0^t f(s)ds + B_t$. Soit P_f^T la loi de probabilité de $(\xi_t^f, t \in (0,T])$. Sur $(C_T, \mathcal{C}_T, P_f^T)$, le processus canonique est un mouvement brownien de dérive f. Le processus $(B_t^f, t \in [0,T])$ défini par $B_t^f = X_t - \int_0^t f(s)ds$ est un mouvement brownien standard.

4.2 Mouvement brownien avec dérive.

4.2.1 Prérequis.

On suppose connues les propriétés du processus défini par l'intégrale de Wiener $\int_0^t f(s)dB_s, t \geq 0)$ où $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ est localement de carré intégrable (pour tout $t \geq 0$, $\int_0^t f^2(s)ds < +\infty$) ainsi que le théorème de Cameron-Martin. On pourra consulter par exemple Revuz et Yor (1991).

4.2.2 Vraisemblance.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, et $(B_t, t \in [0, T])$ un mouvement brownien standard sur cet espace. Supposons que l'on observe sur tout l'intervalle de temps $[0, T]$ le processus

$$\xi_t = \xi_t^{\theta_0} = \int_0^t f(\theta_0, s)ds + B_t, \tag{4.7}$$

où θ_0 est un paramètre inconnu appartenant à un ensemble de paramètres $\Theta \subset \mathbb{R}^p$ et $f : \Theta \times \mathbb{R}^+ \rightarrow \mathbb{R}$ est une fonction telle que, pour tout θ, $f(\theta, .)$ est continue. Cette hypothèse simplificatrice correspond aux applications statistiques et implique notamment que $f(\theta, .)$ est localement intégrable et de carré intégrable ce qui est utile pour la définition de l'intégrale ordinaire et de l'intégrale de Wiener de $f(\theta, .)$. La vraie valeur (inconnue) θ_0 est à estimer à partir de l'observation $(\xi_t, t \in [0, T])$. Le modèle statistique canonique associé à cette observation est défini par:

- l'espace des observations $(C_T, \mathcal{C}_T)$,
- la loi de probabilité $P_{\theta_0}^T$ de l'observation $(\xi_t, t \in [0, T])$.

La vraie valeur étant inconnue, on considère le modèle $(C_T, \mathcal{C}_T, P_\theta^T)_{\theta \in \Theta}$ où P_θ^T est la loi de $(\xi_t^\theta = \int_0^t f(\theta, s)ds + B_t, t \in [0, T])$

Théorème 4.2.1. *Pour tout θ, les lois P_θ^T et W^T sont équivalentes et l'on a:*

$$\frac{dP_\theta^T}{dW^T}(x) = \exp\,[\int_0^T f(\theta, s)dX_s(x) - (1/2)\int_0^T f^2(\theta, s)ds]. \tag{4.8}$$

Dans la formule ci-dessus, on peut omettre x. L'intégrale $\int_0^T f(\theta, s)dX_s$ est une intégrale stochastique par rapport au processus canonique. Cette intégrale a un sens sous W^T et sous P_θ^T: sous W^T, (X_t) est un mouvement brownien standard, sous P_θ^T, $(X_t - \int_0^t f(\theta, s)ds)$ est un mouvemen brownien standard. Pour prouver le théorème, nous avons besoin du lemme de théorie de la mesure suivant.

Lemme 4.2.1. *1. Soient P, Q deux mesures positives sur $(\mathcal{X}, \mathcal{B})$. Si $dP/dQ = Z$ et si $Q(Z = 0) = 0$, alors, $dQ/dP = 1/Z$.*

2. Si P_1, P_2, Q sont des mesures positives sur $(\mathcal{X}, \mathcal{B})$ avec $dP_1/dQ = Z_1$, $dP_2/dQ = Z_2$. Si $P_1(Z_2 = 0) = 0$, $dP_1/dP_2 = Z_1/Z_2$.

Montrons le théorème.

Preuve. Posons $f(\theta, s) = f(s)$, $\xi^\theta = \xi^f$, $P_\theta^T = P_f^T$. La v.a.

$$Z_T = \exp\,[-\int_0^T f(s)dB_s - (1/2)\int_0^T f^2(s)ds]$$

est strictement positive et $\mathbb{E}(Z_T) = 1$ de sorte que $\tilde{\mathbb{P}} = Z_T\mathbb{P}$ est une probabilité sur Ω. D'après le théorème de Cameron-Martin, sur $(\Omega, \tilde{\mathbb{P}})$, le processus $(B_t + \int_0^t f(s)ds = \xi_t^f, t \leq T)$ est un mouvement brownien standard. Donc, pour tout $C \in \mathcal{C}_T$,

$$\tilde{\mathbb{P}}(\xi \in C) = W^T(C) = \mathbb{E}(Z_T 1_{(\xi\in C)}).$$

Or, $dB_s = d\xi_s^f - f(s)ds$, donc:

$$Z_T = \exp\,[-\int_0^T f(s)d\xi_s^f + (1/2)\int_0^T f^2(s)ds].$$

Ainsi,

$$W^T(C) = \int_{\mathcal{C}_T} 1_{(X\in C)} \exp\,[-\int_0^T f(s)dX_s + (1/2)\int_0^T f^2(s)ds]dP_T^f.$$

Ceci démontre que

$$dW^T/dP_f^T = \exp\,[-\int_0^T f(s)dX_s + (1/2)\int_0^T f^2(s)ds].$$

On peut appliquer le lemme pour obtenir

$$dP_f^T/dW^T = \exp\,[\int_0^T f(s)dX_s - (1/2)\int_0^T f^2(s)ds].$$

□

On en déduit le corollaire:

Corollaire 4.2.1. *Pour tout couple (θ, θ'), les lois P_θ^T et $P_{\theta'}^T$ sont équivalentes et*

$$\frac{dP_\theta^T}{dP_{\theta'}^T} = \exp\,[\int_0^T (f(\theta,s) - f(\theta',s))dX_s - (1/2)\int_0^T (f^2(\theta,s) - f^2(\theta',s))ds].$$

Preuve. Il suffit de faire le quotient des densités par rapport à W^T. □

Nous pouvons définir la vraisemblance.

Définition 4.2.1. *La vraisemblance associée à l'observation $(\xi_t, t \leq T)$ est la fonction définie par:*

$$\theta \to dP_\theta^T/dW^T(\xi) = \exp\,[\int_0^T f(\theta,s)d\xi_s - (1/2)\int_0^T f^2(\theta,s)ds],$$

où $\xi = \xi^{\theta_0}$ est le processus observé.

Soit $\ell_T(\theta) = \log L_T(\theta)$.

Définition 4.2.2. *On appelle estimateur du maximum de vraisemblance de θ toute solution $\hat{\theta}_T$ de l'équation*

$$\ell_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} \ell_T(\theta). \tag{4.9}$$

Exemples:

1. Soit le processus observé $\xi_t = \theta_0 t + B_t$. Il fait partie du modèle général $\xi_t^\theta = \theta t + B_t$ de dérive $f(\theta, t) = \theta$. La log-vraisemblance s'écrit $l_T(\theta) = \theta \xi_T - (\theta^2/2)T$. L'estimateur du maximum de vraisemblance est unique et vaut $\hat{\theta}_T = \xi_T/T$. Il suit la loi $\mathcal{N}(\theta_0, 1/T)$.

2. Soit le processus $\xi_t = \theta_0(t^2/2) + B_t$. Il fait partie du modèle général $\xi_t^\theta = \theta(t^2/2) + B_t$ de dérive $f(\theta, t) = \theta t$. La log-vraisemblance s'écrit $l_T(\theta) = \theta \int_0^T t d\xi_t - (\theta^2/2)(T^3/3)$. L'estimateur du maximum de vraisemblance est unique et vaut $\hat{\theta}_T = \int_0^T t d\xi_t/(T^3/3)$. Il suit la loi $\mathcal{N}(\theta_0, 3/T^3)$.

3. Traiter le cas de $d\xi_t = (a_0 t + b_0)dt + dB_t$, $\xi_0 = 0$.

4.2.3 Extensions de la formule de vraisemblance.

Soit (B_t) un mouvement brownien défini sur $(\Omega, \mathcal{F}, \mathbb{P})$. Supposons que l'on observe le processus

$$\xi_t = \xi_t^{\theta_0} = x_0 + \int_0^t f(\theta_0, s)ds + \int_0^t \sigma(s)dB_s,$$

où $x_0 \in \mathbb{R}$ est connu, $f(\theta, .)$ est, pour tout θ, continue, comme ci-dessus, $\sigma : \mathbb{R}^+ \to \mathbb{R}$ est continue, strictement positive sur $[0, T]$ et connue. Soit

$$M_t = x_0 + \int_0^t \sigma(s)dB_s.$$

On note P_θ^T la loi du processus $(\xi_t^\theta, t \in [0, T])$ (correspondant à la valeur θ du paramètre) et P^T loi du processus $(M_t, t \in [0, T])$. Nous omettons la référence à la fonction σ dans les notations car cette fonction est connue.

Théorème 4.2.2. *Pour tout θ, les lois P_θ^T et P^T sont équivalentes et l'on a:*

$$\frac{dP_\theta^T}{dP^T}(x) = \exp\left[\int_0^T \frac{f(\theta, s)}{\sigma^2(s)} dX_s(x) - (1/2) \int_0^T \frac{f^2(\theta, s)}{\sigma^2(s)} ds\right] := L_T(\theta, x). \tag{4.10}$$

Sous P_θ^T, le processus $(B_t^\theta = \int_0^t \frac{dX_s - f(\theta,s)ds}{\sigma(s)}, t \in [0, T])$ est un mouvement brownien standard.

Preuve. Soit

$$Z_T = \exp\left[- \int_0^T \frac{f(\theta, s)}{\sigma(s)} dB_s - (1/2) \int_0^T \frac{f^2(\theta, s)}{\sigma^2(s)} ds\right].$$

Sous $\tilde{\mathbb{P}} = Z_T \mathbb{P}$, le processus $(W_t^\theta = B_t + \int_0^t \frac{f(\theta,s)}{\sigma(s)} ds, t \leq T)$ est un mouvement brownien standard et $(\xi_t^\theta = \int_0^t \sigma(s) dW_s^\theta, t \leq T)$ a donc la loi P^T. Donc, pour $B \in \mathcal{C}_T$,

$$P^T(B) = \tilde{\mathbb{P}}(\xi^\theta \in B) = \int_\Omega 1_{(\xi^\theta \in B)} Z_T d\mathbb{P}.$$

Or,

$$Z_T = \exp\,[-\int_0^T \frac{f(\theta,s)}{\sigma^2(s)} d\xi_s^\theta + (1/2)\int_0^T \frac{f^2(\theta,s)}{\sigma^2(s)} ds].$$

D'où

$$P^T(B) = \int_{C_T} 1_{(X \in B)} \left(L_T(\theta, X)\right)^{-1} dP_\theta^T.$$

Ce qui permet de conclure.

Sous P_θ^T, le processus $(X_t - \int_0^t f(\theta,s)ds = N_t^\theta)$ est une martingale gaussienne continue et $< N^\theta >_t = \int_0^t \sigma^2(s)ds$ (la filtration est la filtration canonique). En posant

$$B_t^\theta = \int_0^t \frac{dN_s^\theta}{\sigma(s)},$$

on obtient une martingale locale continue telle que $< B^\theta >_t = \int_0^t \frac{\sigma^2(s)}{\sigma^2(s)} ds = t$, donc un mouvement brownien. □

On obtient ainsi, pour ce modèle, la vraisemblance:

$$\theta \to \frac{dP_\theta^T}{dP^T}(\xi) = \exp\,[\int_0^T \frac{f(\theta,s)}{\sigma^2(s)} d\xi_s - (1/2)\int_0^T \frac{f^2(\theta,s)}{\sigma^2(s)} ds] = L_T(\theta, \xi).$$

Remarque 4.2.1. *La filtration canonique est celle définie par le processus canonique, c'est-à-dire $(\mathcal{C}_t = \sigma(X_s, s \leq t), t \leq T)$. Si nécessaire, on peut considérer la filtration rendue continue à droite $(\mathcal{C}_t^+ = \cap_{h>0} \mathcal{C}_{t+h})$.*

4.2.4 Estimateurs du maximum de vraisemblance: comportement asymptotique lorsque T tend vers l'infini.

Considérons l'exemple particulier suivant:

$$d\xi_t = \theta_0 f(t)dt + \sigma(t)dB_t, \quad \xi_0 = 0,$$

où f, σ sont connues, continues sur $\mathbb{R}^+$ et $\sigma > 0$. La fonction de log-vraisemblance s'écrit:

$$\ell_T(\theta) = \log L_T(\theta, \xi) = \theta \int_0^T \frac{f(s)}{\sigma^2(s)} d\xi_s - (\theta^2/2)\int_0^T \frac{f^2(s)}{\sigma^2(s)} ds.$$

D'où, l'estimateur du maximum de vraisemblance:

$$\hat{\theta}_T = \frac{\int_0^T \frac{f(s)}{\sigma^2(s)} d\xi_s}{\int_0^T \frac{f^2(s)}{\sigma^2(s)} ds}.$$

Utilisant la loi du modèle, on obtient:

$$\hat{\theta}_T = \theta_0 + \frac{\int_0^T \frac{f(s)}{\sigma(s)} dB_s}{\int_0^T \frac{f^2(s)}{\sigma^2(s)} ds}.$$

L'estimateur $\hat{\theta}_T$ a la loi gaussienne de moyenne θ_0 (il est sans biais), de variance I_T^{-1} avec

$$I_T = \int_0^T \frac{f^2(s)}{\sigma^2(s)} ds.$$

Le comportement asymptotique lorsque T tend vers l'infini de l'estimateur dépend de I_T. On ne peut pas faire de théorie générale. La consistance et la vitesse de l'estimateur dépendront de la forme de f et σ.

Exemple. Pour $f(t) = 1, \sigma(t) = \sqrt{1+t^2}$, $I_T = \text{Arctan}(T) \to \pi/2$ quand T tend vers l'infini. L'estimateur n'est même pas consistant.

4.2.5 Conditions initiales et coefficient de diffusion.

Considérons les processus suivants, avec f, σ, σ_0 continues:

$$\xi_t^{x,f,\sigma} = x + \int_0^t f(s) ds + \int_0^t \sigma(s) dB_s$$

et

$$\xi_t^{x_0,\sigma_0} = x_0 + \int_0^t \sigma_0(s) dB_s$$

de lois respectives $P_{x,f,\sigma}^T$ et P_{x_0,σ_0}^T sur $(C_T, \mathcal{C}_T)$. On a:

Proposition 4.2.1. *1. Si $x_0 \neq x$, les lois $P_{x,f,\sigma}^T$ et P_{x_0,σ_0}^T sont étrangères.*

2. Si les fonctions σ et σ_0 ne sont pas identiques sur $[0, T]$, les lois $P_{x,f,\sigma}^T$ et P_{x_0,σ_0}^T sont étrangères.

Le point 1) est immédiat. Le point 2) vient des propriétés des variations quadratiques des intégrales de Wiener. Il justifie le fait que, lors d'une observation en temps continu sur $[0, T]$ d'une trajectoire, on suppose le coefficient de diffusion connu: il est, en effet, *identifié* par ce type d'observation.

4.3 Estimation paramétrique de dérive pour un modèle de diffusion: généralités.

Avant de traiter la partie statistique, il convient de rappeler les propriétés probabilistes principales qui seront utiles.

4.3.1 Prérequis.

Désormais, nous aurons besoin d'intégrales stochastiques d'Itô de la forme $M_t = \int_0^t H_s dB_s$ où (B_t) est un $(\mathcal{F}_t)$-mouvement brownien sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$, (H_t) est un processus continu à gauche adapté à $(\mathcal{F}_t)$. Rappelons que la tribu prévisible sur $\Omega \times \mathbb{R}^+$ est la tribu engendrée par ces processus. Les filtrations seront toujours supposées vérifier les *conditions habituelles*.

Il faut savoir que (M_t) est une martingale locale continue nulle en 0, relative à $(\mathcal{F}_t)$, de variation quadratique $< M >_t = \int_0^t H_s^2 ds$, c'est-à-dire telle que $(M_t^2 - < M >_t)$ est aussi une martingale locale. Pour vérifier que (M_t) est une vraie martingale, nous utiliserons:

- Si, pour tout t, $\mathbb{E}(\int_0^t H_s^2 ds) = \int_0^t \mathbb{E} H_s^2 ds < +\infty$, alors (M_t) est une vraie martingale, de carré intégrable, $(M_t^2 - < M >_t)$ est aussi une vraie martingale.
- Si, pour tout t, $\mathbb{E} \sup_{s \leq t} |M_s| < +\infty$, alors, (M_t) est une vraie martingale.

Nous utiliserons souvent le théorème d'arrêt des martingales sous la forme suivante:

Théorème 4.3.1. *Soit (M_t) une martingale locale continue nulle en 0, relative à la filtration $(\mathcal{F}_t)$. Pour tout temps d'arrêt T, le processus arrêté $(M_{t \wedge T}, t \geq 0)$ est une martingale locale relative à la filtration $(\mathcal{F}_t)$.*

Dans ce qui suit, nous nous appuierons sur d'autres résultats importants: le théorème de Girsanov, le théorème de changement de temps de Dambis, Dubins et Schwarz (D-D-S). Enfin, l'outil essentiel est la formule d'Itô pour les semi- martingales d'Itô continues. (cf par exemple, Revuz et Yor (1991)).

4.3.2 Equations différentielles stochastiques: théorèmes d'existence et d'unicité de solutions.

Pour simplifier, nous ne considèrerons que des équations différentielles stochastiques (EDS) autonomes et nous donnons une version simple des théorèmes qui suffit pour les exemples que nous envisageons par la suite. Pour des théorèmes d'existence et d'unicité de solutions plus généraux, nous renvoyons, par exemple, à Karatsas et Shreeve (1997) ou tout autre ouvrage sur les équations différentielles stochastiques (voir références).

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, muni d'une filtration $(\mathcal{F}_t, t \geq 0)$ satisfaisant aux *conditions habituelles* et (B_t) un $\mathbb{P}$-$(\mathcal{F}_t, t \geq 0)$ mouvement brownien standard. Considérons l'équation différentielle stochastique:

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dB_t, \quad \xi_0 = \eta, \tag{4.11}$$

où les fonctions $b, \sigma : \mathbb{R} \to \mathbb{R}$ sont <u>continues</u> et η est un v.a. <u>réelle</u>.

Théorème 4.3.2. *Considérons les hypothèses suivantes:*

- η *est* $\mathcal{F}_0$*-mesurable.*
- *Les fonctions* b, σ *sont localement lipschitziennes: pour tout* $N > 0$*, il existe* $L_N > 0$ *tel que, pour tous* x, y *vérifiant* $|x| \leq N, |y| \leq N$,

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq L_N |x - y|.$$

- *Les fonctions* b, σ *sont à croissance sous-linéaire:*

$$\exists K > 0, \forall x \in \mathbb{R}, |b(x)| + |\sigma(x)| \leq K(1 + |x|).$$

- $\mathbb{E}(\eta^2) < +\infty$.

Alors, l'EDS (4.11) admet un processus solution $(\xi_t, t \geq 0)$ *défini sur* Ω *et tel que:*

- $\xi_0 = \eta$ *et la trajectoire* $(\xi_t, t \geq 0)$ *est continue* $\mathbb{P}$*- p.s..*
- $(\xi_t, t \geq 0)$ *est adapté à* $(\mathcal{F}_t, t \geq 0)$*.*
- *Si* ξ^1 *et* ξ^2 *sont deux processus solutions vérifiant les deux propriétés précédentes, alors* $\mathbb{P}(\forall t \geq 0, \xi^1_t = \xi^2_t) = 1$ *(*ξ^1 *et* ξ^2 *sont dits indistinguables).*
- *Pour tout* $t \geq 0$*,* $\sup_{s \leq t} \mathbb{E}(\xi_s^2) < +\infty$*.*

On dit qu'il y a existence et unicité de solution forte non explosive.

Remarque 4.3.1.

- *Si* b, σ *sont globalement lipschitziennes, elles sont localement lipschitziennes et à croissance sous-linéaire.*
- *Si* b, σ *sont de classe* C^1 *sur* $\mathbb{R}$*, elles sont localement lipschitziennes.*
- *Si* $\eta = x_0$ *est déterministe, elle est évidemment de carré intégrable et* ξ_t *l'est aussi pour tout* t*.*
- *Si* $\mathbb{E}\eta^2 < +\infty$*,* $M_t = \int_0^t \sigma(\xi_s) ds$ *vérifie*

$$\mathbb{E}(M_t^2) = \mathbb{E} \int_0^t \sigma^2(\xi_s) ds \leq 2K^2 \mathbb{E} \int_0^t (1 + \xi_s^2) ds = 2K^2 \int_0^t (1 + \mathbb{E}(\xi_s^2)) ds < +\infty,$$

donc, (M_t) *est une vraie martingale (pas seulement locale) nulle en* 0*, et centrée. (Avec la continuité des trajectoires, on montre que* $(\omega, t) \to \xi_t(\omega)$ *est conjointement mesurable, ce qui permet l'utilisation de Fubini et l'interversion de l'espérance et de l'intégrale ordinaire).*

Le théorème 4.3.2 couvre un ensemble important d'EDS et s'étend aux modèles multidimensionnels. Il ne couvre pas le cas important où $\sigma(x) = \sqrt{x^+}$ qui n'est pas localement lipschitzienne. Le théorème suivant répond à la question mais ne s'applique qu'aux modèles unidimensionnels.

Théorème 4.3.3. *Considérons les hypothèses suivantes:*

- η *est* $\mathcal{F}_0$*-mesurable.*
- *La fonction* b *est lipschitzienne, la fonction* σ *est Höldérienne d'exposant* $\alpha \in [1/2, 1]$*:*

$$\exists L > 0, \forall x, y, |b(x) - b(y)| \leq L|x - y|, \quad |\sigma(x) - \sigma(y)| \leq L|x - y|^{\alpha}.$$

- $\mathbb{E}(\eta^2) < +\infty$.

Alors, l'EDS (4.11) admet un processus solution $(\xi_t, t \geq 0)$ *défini sur* Ω *et tel que:*

- $\xi_0 = \eta$ *et la trajectoire* $(\xi_t, t \geq 0)$ *est continue* $\mathbb{P}$*-p.s. continue.*
- $(\xi_t, t \geq 0)$ *est adapté à* $(\mathcal{F}_t, t \geq 0)$*.*
- *Si* ξ^1 *et* ξ^2 *sont deux processus solutions vérifiant les deux propriétés précédentes, alors* $\mathbb{P}(\forall t \geq 0, \xi_t^1 = \xi_t^2) = 1$ *(*ξ^1 *et* ξ^2 *sont indistinguables).*
- *Pour tout* $t \geq 0$, $\sup_{s \leq t} \mathbb{E}(\xi_s^2) < +\infty$.

Encore une fois, nous ne cherchons pas à donner la version la plus générale du résultat mais celle qui convient à nos exemples. Notons que, sous les hypothèses du Théorème 4.3.3, les fonctions b, σ sont sous-linéaires.

Dans toute la suite, les EDS que nous considèrerons seront supposées vérifier l'un ou l'autre de ces théorèmes.

4.3.3 Loi de probabilité des solutions.

Théorème 4.3.4. *Sous les hypothèses du théorème 4.3.2 ou 4.3.3, la loi de probabilité* P^T *sur* $(C_T, \mathcal{C}_T)$ *d'un processus solution* $(\xi_t, t \in [0, T])$ *ne dépend que des fonctions* b, σ *et de la loi de probabilité* μ *de la v.a. initiale* η*. Elle ne dépend pas de l'espace* $(\Omega, \mathcal{F}, \mathbb{P})$*, de la v.a.* η*, du mouvement brownien utilisés pour construire le processus solution. Ceci siginifie que la loi* P^T *que l'on peut noter* $P^T_{\mu,b,\sigma}$ *est unique.*

Le théorème 4.3.4 exprime la propriété suivante, valable sous des hypothèses génerales: l'unicité de solution forte implique l'unicité en loi.

4.3.4 Continuité absolue des lois de solutions: vraisemblance.

Considérons le modèle paramétrique pour la dérive :

$$d\xi_t = b(\xi_t, \theta_0)dt + \sigma(\xi_t)dB_t, \quad \xi_0 = x_0. \tag{4.12}$$

On suppose connus la valeur initiale x_0, la fonction $\sigma(x)$, la fonction $b(x, \theta)$. La vraie valeur θ_0 est inconnue, appartient à un ensemble $\Theta \subset \mathbb{R}^p$, et doit être estimée à partir de l'observation d'une trajectoire $(\xi_t, t \in [0, T])$. Le processus (B_t) est un mouvement brownien défini sur l'espace de probabilité $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$, relatif à la filtration $(\mathcal{F}_t)$. Pour écrire la vraisemblance associée à ce problème, nous introduisons les hypothèses suivantes.

(I) Pour tout $\theta \in \Theta$, la fonction $x \to b(x, \theta)$ et la fonction $x \to \sigma(x)$ sont telles que les hypohèses du théorème 4.3.2 ou 4.3.3 sont vérifiées. On note P_θ^T la loi du processus $(\xi_t^\theta, t \in [0, T])$ correspondant à la condition initiale x_0, la dérive $b(., \theta)$, le coefficient de diffusion $\sigma(.)$. Ainsi, nous obtenons le modèle statistique $(C_T, \mathcal{C}_T, (P_\theta^T, \theta \in \Theta))$.

(II) Pour tout $\theta \in \Theta$,

$$P_\theta^T(\sigma^2(X_t) > 0, \forall t \in [0, T]) = \mathbb{P}(\sigma^2(\xi_t^\theta) > 0, \forall t \in [0, T]) = 1.$$

L'hypothèse (I) exprime l'unicité forte de solution pour l'EDS associée à $x_0, b(., \theta), \sigma(.)$. L'hypothèse (II) est satisfaite lorsque $\sigma(x) > 0$ pour tout $x \in \mathbb{R}$. Nous donnerons plus loin des conditions suffisantes pour que (II) soit satisfaite même lorsque σ peut s'annuler.

Théorème 4.3.5. *Sous les hypothèses (I)-(II), pour tout couple (θ, θ'), les lois P_θ^T et $P_{\theta'}^T$ sont équivalentes et:*

$$\frac{dP_\theta^T}{dP_{\theta'}^T}(X) = \exp\left[\int_0^T \frac{b(X_t, \theta) - b(X_t, \theta')}{\sigma^2(X_t)} dX_t - (1/2)\int_0^T \frac{b^2(X_t, \theta) - b^2(X_t, \theta')}{\sigma^2(X_t)} dt\right].$$

De plus, sur $(C_T, \mathcal{C}_T, P_\theta^T)$, le processus

$$(W_t^\theta = \int_0^t \frac{dX_s - b(X_s, \theta)ds}{\sigma^2(X_s)}, t \in [0, T])$$

est un mouvement brownien standard.

Corollaire 4.3.1. *La fonction de vraisemblance associée à l'observation $(\xi_t = \xi_t^{\theta_0})$ est définie (à une fonction de ξ ne dépendant pas de θ près) par:*

$$\theta \to L_T(\theta)$$

avec

$$\ell_T(\theta) = \log L_T(\theta) = \int_0^T \frac{b(\xi_t, \theta)}{\sigma^2(\xi_t)} d\xi_t - (1/2)\int_0^T \frac{b^2(\xi_t, \theta)}{\sigma^2(\xi_t)} dt.$$

L'estimateur du maximum de vraisemblance $\hat{\theta}_T$ est défini comme toute solution de l'équation:

$$\ell_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} \ell_T(\theta).$$

Preuve. Nous démontrons le théorème 4.3.5 dans un cadre très simplifié. On supposera que $\sigma(x) = 1$, $x_0 = 0$ et $|b(x, \theta)| \leq K$. La preuve est proche de celle faite lorsque la dérive est déterministe mais s'appuie sur le théorème de Girsanov au lieu du théorème de Cameron-Martin. Il faut en plus utiliser l'existence et l'unicité de solution pour l'EDS.

Par hypothèse, le processus ($\xi_t^\theta = B_t + \int_0^t b(\xi_s^\theta, \theta)ds, t \leq T$) existe, est unique (à une indistinguabilité près), a pour loi P_θ^T sur $(C_T, \mathcal{C}_T)$. Posons

$$Z_T = \exp\left[-\int_0^T b(\xi_s^\theta, \theta)dB_s - (1/2)\int_0^T b^2(\xi_s^\theta, \theta)ds\right].$$

L'hypothèse sur $b(x, \theta)$ implique que

$$\mathbb{E}_\mathbb{P} \exp\left[(1/2)\int_0^T b^2(\xi_s^\theta, \theta)ds\right] \leq \mathbb{E}_\mathbb{P} \exp\left(K^2 T/2\right) < +\infty.$$

D'après le critère de Novikov (cf, par exemple, Revuz et Yor, 1991, Chapitre VIII, Proposition 1.15), Z_T est une v.a. strictement positive et telle que $\mathbb{E}_\mathbb{P} Z_T = 1$. Sur $(\Omega, \mathcal{F}_T)$, on définit la probabilité $\tilde{\mathbb{P}} = Z_T \mathbb{P}$. Le théorème de Girsanov (cf Revuz et Yor, 1991, Chapitre VIII, Théorème 1.4) implique que, sur $(\Omega, \mathcal{F}_T, \tilde{\mathbb{P}})$, le processus ($\xi_t^\theta = B_t + \int_0^t b(\xi_s^\theta, \theta)ds, t \leq T$) est un mouvement brownien standard. Donc, si W^T désigne la mesure de Wiener sur $(C_T, \mathcal{C}_T)$ et $A \in \mathcal{C}_T$,

$$W^T(A) = \tilde{\mathbb{P}}((\xi_t^\theta, t \in [0, T]) \in A) = \int_\Omega 1_{\xi^\theta \in A} Z_T d\mathbb{P}.$$

Or,

$$Z_T = \exp\left[-\int_0^T b(\xi_s^\theta, \theta)d\xi_s^\theta + (1/2)\int_0^T b^2(\xi_s^\theta, \theta)ds\right].$$

Ainsi,

$$\frac{dW^T}{dP_\theta^T}(X) = \exp\left[-\int_0^T b(X_s, \theta)dX_s + (1/2)\int_0^T b^2(X_s, \theta)ds\right].$$

D'où le résultat pour $\frac{dP_\theta^T}{dW^T}$ ainsi que pour $\frac{dP_\theta^T}{dP_{\theta'}^T}$. $\square$

Nous admettrons le résultat sous sa forme générale. En ce qui concerne la vraisemblance, on fixe une valeur connue θ^* et on considère comme fonction de vraisemblance

$$\theta \to \frac{dP_\theta^T}{dP_{\theta^*}^T}(\xi).$$

On voit ainsi que la partie "utile" de cette fonction (celle qui dépend de θ et des observations ξ) est justement $L_T(\theta)$.

Pour la vraisemblance, consulter Revuz et Yor (1991) Théorème 1.10, p. 344, Liptser et Shiryaev (Chapitre VII, 2001), Jacod et Shiryaev (1987, 2003).

4.3.5 Etude des estimateurs du maximum de vraisemblance.

Dans le cadre asymptotique T tend vers $+\infty$, on cherche à répondre aux questions usuelles suivantes. Soit $\hat{\theta}_T$ un estimateur du maximum de vraisemblance (EMV), a-t-on:

1. $\hat{\theta}_T - \theta_0$ converge vers 0 en $\mathbb{P}$-probabilité (consistance faible) ou $\mathbb{P}$-p.s. (consistance forte).

2. Peut-on trouver une fonction $\varphi(T, \theta_0)$ tendant vers l'infini lorsque T tend vers l'infini, telle que $\varphi(T, \theta_0)(\hat{\theta}_T - \theta_0)$ converge en loi.

Deux cas sont possibles:

- On sait résoudre explicitement l'équation de vraisemblance et l'on peut étudier $\hat{\theta}_T$ à partir de son expression explicite en fonction de (ξ_t).
- En l'absence de formule explicite pour $\hat{\theta}_T$, il faut l'étudier directement sur l'équation de vraisemblance. C'est le cas le plus fréquent.

Dans les deux cas, la formule de vraisemblance étant explicite, on peut calculer l'EMV sur les données, au moins numériquement. Bien entendu, il faudra faire une approximation discrète des intégrales stochastiques et ordinaires.

Exemple:
Considérons le modèle
$$d\xi_t = \theta_0 b(\xi_t)dt + \sigma(\xi_t)dB_t, \xi_0 = x_0,$$
avec b, σ fonctions connues. La log-vraisemblance s'écrit:
$$\ell_T(\theta) = \theta \int_0^T \frac{b(\xi_s)}{\sigma^2(\xi_s)} d\xi_s - \frac{\theta^2}{2} \int_0^T \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds.$$
L'estimateur du maximum de vraisemblance vaut:
$$\hat{\theta}_T = \frac{\int_0^T \frac{b(\xi_s)}{\sigma^2(\xi_s)} d\xi_s}{\int_0^T \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds}.$$
Cette formule est explicite et permet de calculer l'EMV sur les données. Pour étudier la loi de l'estimateur, on reporte dans cette formule l'expression de $d\xi_t$ en fonction de dB_t. On obtient:
$$\hat{\theta}_T = \theta_0 + \frac{M_T}{< M >_T},$$
avec $M_T = \int_0^T \frac{b(\xi_s)}{\sigma(\xi_s)} dB_s$. La loi exacte est difficile à étudier. En revanche, on peut avoir des résultats asymptotiques lorsque T tend vers l'infini s'appuyant sur les deux théorèmes suivants.

Théorème 4.3.6. *Soit $M_t = \int_0^t H_s dB_s$ où (H_s) est un processus continu à gauche adapté à la filtration $(\mathcal{F}_t)$ du mouvement brownien (B_t). Si $< M >_\infty = \int_0^{+\infty} H_s^2 ds = +\infty$ p.s.,alors $M_t / < M >_t$ tends vers 0 p.s. lorsque t tend vers l'infini.*

Preuve. Cette propriété généralise celle qui est vraie pour le mouvement brownien, à savoir B_t/t tend vers 0 p.s. lorsque t tend vers l'infini. Pour la prouver, appliquons le théorème de Dambis-Dubins-Schwarz (théorème 1.6, p. 170, Revuz et yor, 1991). Soit $T_t = \inf\{u; \int_0^u H_s^2 ds \geq t\}$. L'hypothèse $< M >_\infty = +\infty$ p.s. implique que $B'_t = M_{T_t}$ est un mouvement brownien et que $M_t = B'_{<M>_t}$. D'où le résultat puisque $< M >_t$ tend vers l'infini p.s., en appliquant la propriété vraie pour (B'_t). $\square$

Théorème 4.3.7. *Si* $\frac{<M>_T}{T}$ *tend en probabilité vers une constante déterministe* $\sigma^2 \geq 0$ *lorsque* T *tend vers l'infini, alors*

$$\frac{M_T}{\sqrt{T}} \rightarrow_{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

lorsque T *tend vers l'infini. (Si* $\sigma^2 = 0$, $\frac{M_T}{\sqrt{T}}$ *converge en probabilité vers* 0*).*

Preuve. Traitons seulement le cas $\sigma^2 > 0$ (l'autre cas est élémentaire). Dans ce cas, notons que l'hypothèse implique, puisque $< M >$ est un processus croissant, que $< M >_\infty = +\infty$ p.s..

On introduit le processus $H_t^T = \frac{H_t}{\sqrt{T}}$, si $t \in [0, T]$, $H_t^T = \sigma$, si $t > T$, et le temps d'arrêt

$$\tau_T = \inf \{t \geq 0; \int_0^t (H_s^T)^2 ds \geq \sigma^2\}.$$

D'après la définition de H_t^T, $\tau_T \leq T + 1$ est un temps d'arrêt uniformément borné. De plus,

$$\int_0^{t \wedge \tau_T} (H_s^T)^2 ds \leq \sigma^2, \text{ pour tout } t, \text{ et } \int_0^{\tau_T} (H_s^T)^2 ds = \sigma^2, \tag{4.13}$$

car le processus $(\int_0^t (H_s^T)^2 ds, t \geq 0)$ est à trajectoires continues. Nous allons montrer successivement que:

$$\zeta_T = \int_0^{\tau_T} H_t^T dB_t \quad \text{a la loi } \mathcal{N}(0, \sigma^2), \tag{4.14}$$

$$Z_t - \zeta_T \rightarrow_{\mathbb{P}} 0 \quad \text{où } Z_T = \int_0^T H_t^T dB_t, \tag{4.15}$$

ce qui permettra de conclure.
Montrons (4.14) en utilisant (4.13). Pour λ réel, Soit

$$L_t = \exp(i\lambda \int_0^t H_s^T dB_s + \frac{\lambda^2}{2} \int_0^t (H_s^T)^2 ds).$$

Une application de la formule d'Itô montre que (L_t) est une martingale locale:

$$L_t = 1 + i\lambda \int_0^t L_s H_s^T dB_s.$$

Par suite, $(L_{t \wedge \tau_T})$ l'est aussi. Comme

$$|L_{t \wedge \tau_T}| \leq \exp(\frac{\lambda^2}{2} \int_0^{t \wedge \tau_T} (H_s^T)^2 ds) \leq \exp(\frac{\lambda^2}{2} \sigma^2),$$

$(L_{t \wedge \tau_T})$ est une martingale locale uniformément bornée, donc une vraie martingale. Par suite,

$$\mathbb{E} L_{t \wedge \tau_T} = \mathbb{E} L_0 = 1.$$

En faisant tendre t vers l'infini, et en appliquant le théorème de convergence dominée, on déduit $\mathbb{E} L_{\tau_T} = 1$, ce qui s'écrit:

$$\mathbb{E}(\exp(i\lambda\zeta_T)) = \exp(-\frac{\lambda^2}{2}\sigma^2).$$

Cette relation, valable pour tout réel λ, montre que la fonction caractéristique de ζ_T est celle de la loi $\mathcal{N}(0, \sigma^2)$.
Montrons (4.15). On a:

$$\zeta_T - Z_T = \int_0^{T+1} H_s^T (1_{s \leq \tau_T} - 1_{s \leq T}) dB_s.$$

D'après l'inégalité de Lenglart,

$$\mathbb{P}(|\zeta_T - Z_T| > h) \leq \frac{\eta}{h^2} + \mathbb{P}(A_T),$$

avec

$$A_T = \{\int_0^{T+1} (H_s^T)^2 \, (1_{s \leq \tau_T} - 1_{s \leq T})^2 \, ds > \eta\}.$$

Etudions l'événement A_T. D'une part, $A_T \cap (\tau_T = T) = \emptyset$. D'autre part, on a:

$$(1_{s \leq \tau_T} - 1_{s \leq T})^2 \, 1_{(\tau_T < T)} = 1_{(\tau_T < s \leq T)},$$

et, sur $(\tau_T < T)$,

$$\int_{\tau_T}^T (H_s^T)^2 ds = \frac{1}{T} \int_0^T H_s^2 ds - \int_0^{\tau_T} (H_s^T)^2 ds = \frac{1}{T} \int_0^T H_s^2 ds - \sigma^2.$$

Donc,

$$A_T \cap (\tau_T < T) = (\tau_T < T) \cap (\frac{1}{T} \int_0^T H_s^2 ds - \sigma^2 > \eta).$$

De façon analogue,

$$A_T \cap (\tau_T > T) = (\tau_T > T) \cap (\sigma^2 - \frac{1}{T} \int_0^T H_s^2 ds > \eta).$$

On en déduit:

$$\mathbb{P}(A_T) \leq \mathbb{P}(|\frac{1}{T} \int_0^T H_s^2 ds - \sigma^2| > \eta).$$

Ce dernier terme tend vers 0 quand T tend vers l'infini, par hypothèse. Ceci permet d'achever la preuve de (4.15).$\square$

On déduit le corollaire suivant.

Corollaire 4.3.2. *On suppose que (H_s^i), $i = 1, \ldots, p$ sont p processus continus à gauche, adaptés à $(\mathcal{F}_t)$, tels que, pour $1 \leq i, j \leq p$, lorsque T tend vers l'infini,*

$$\frac{1}{T} \int_0^T H_s^i H_s^j ds \to \sigma_{ij}$$

en probabilité, où les σ_{ij} sont des constantes déterministes. Si la matrice $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq p}$ est définie positive, alors le vecteur aléatoire vérifie, lorsque T tend vers l'infini,

$$\frac{1}{\sqrt{T}} \int_0^T H_s^i dB_s \to_{\mathcal{L}} \mathcal{N}_p(0, \Sigma).$$

Preuve. Pour $u_1, \ldots, u_p$ des nombres réels, on pose $H_s = \sum_{i=1}^p u_i H_s^i$ et on applique le théorème précédent:

$$\frac{1}{\sqrt{T}} \int_0^T H_s dB_s \to_{\mathcal{L}} \mathcal{N}_p(0, u'\Sigma u).$$

□

Remarque 4.3.2.
- *La preuve du théorème 4.3.7 permet de montrer que si* (H_s) *est un processus continu à gauche adapté tel que* $\int_0^{+\infty} H_s^2 ds = +\infty$ *p.s., alors* $T_a = \inf\{t, \int_0^t H_s^2 ds \geq a\} < +\infty$ *p.s. et* $\int_0^{T_a} H_s dB_s$ *est une v.a. gaussienne centrée de variance* a.

- *Dans la mesure où* τ_T *est uniformément borné, on aurait pu conclure immédiatement que* $L_{\tau_T} = 1$.

- *L'hypothèse que la limite* σ^2 *(ou* Σ *dans le cas multidimensionnel) est déterministe est capitale pour le théorème. Si la limite est une v.a. aléatoire, on peut seulement déduire la tension de* $M_T/\sqrt{T}$.

- *La preuve ci-dessus est donnée dans Kutoyants (1984).*

- *dans cette preuve, on peut remplacer* $\sqrt{T}$ *par* $\varphi(T, \theta_0)$ *tendant vers l'infini avec* T.

Revenons au modèle

$$d\xi_t = \theta_0 b(\xi_t)dt + \sigma(\xi_t)dB_t, \xi_0 = x_0$$

avec l'EMV de θ_0:

$$\hat{\theta}_T = \frac{\int_0^T \frac{b(\xi_s)}{\sigma^2(\xi_s)} d\xi_s}{\int_0^T \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds},$$

et

$$\hat{\theta}_T = \theta_0 + \frac{M_T}{< M >_T},$$

avec $M_T = \int_0^T \frac{b(\xi_s)}{\sigma(\xi_s)} dB_s$. Si l'on sait prouver que $< M >_\infty = \infty$ p.s., l'estimateur est fortement consistant. Si l'on sait prouver que $< M >_T / \sqrt{T}$ converge en probabilité vers une limite déterministe σ^2, alors, on a:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \to_{\mathcal{L}} \mathcal{N}(0, 1/\sigma^2).$$

4.3.6 Simulation de trajectoires.

Pour une reférence bibliographique sur la simulation de trajectoires de diffusions, citons Iacus (2010).

La simulation exacte d'une discrétisation de la trajectoire d'une diffusion s'appuie sur la propriété de Markov de ce processus dont nous reparlerons plus loin. La loi conditionnelle de ξ_{t+h} sachant $(\xi_s, s \leq t)$ est identique à la loi conditionnelle de ξ_{t+h} sachant ξ_t (loi de transition). Celle-ci admet, sous des hypothèses raisonnables, une densité. Simuler une discrétisation revient donc à simuler suivant la loi de transition. Cependant, ceci n'est pas simple car, en règle générale, les lois de transition ne sont pas connues de manière explicite. C'est pourquoi d'autres méthodes ont été proposées.

Dans le cas de densité de transition inconnue (non explicite), jusqu'aux résultats de l'article Beskos, A., Papaspiliopoulos, O. et Roberts, G.O., 2006 (et d'autres qui ont suivi), il n'existait pas d'algorithme permettant la simulation exacte de v.a. issues d'une diffusion. La méthode classique consiste à faire de la simulation approchée, basée sur des discrétisations de l'EDS: discrétisation au premier ordre (schéma d'Euler), éventuellement à des ordres supérieurs (schéma de Milstein, de Platen et Wagner, ...). Au début des années 2000, l'équipe de Gareth Roberts (Warwick) a trouvé un algorithme ingénieux et simple à la fois, basé sur la formule de vraisemblance et la méthode du rejet . Les algorithmes proposés ne couvrent pas tous les modèles et imposent des restrictions sur les coefficients de dérive et de diffusion. Néanmoins, pour ces modèles, le résultat est une simulation exacte.

Simulation exacte par densité de transition connue.

Nous visons ici deux exemples qui sont liés, le processus d'Ornstein-Uhlenbeck et le processus de Cox-Ingersoll-Ross (CIR) ou diffusion en racine carrée. La discrétisation du processus d'Ornstein-Uhlenbeck a ceci de particulier qu'elle est un processus autoregressif Gaussien AR(1) explicit, donc, facile à simuler. L'explication est fournie au Chapitre 2. Nous ne la redonnons pas.
Le modèle de CIR est ainsi appelé depuis l'article paru en 1985 pour modéliser la structure à terme des taux d'intérêts. Il est donné par l'équation différentielle stochastique suivante:

$$d\xi_t = (a\xi_t + b)dt + \sigma\sqrt{\xi_t^+}dB_t, \quad \xi_0 = x > 0. \tag{4.16}$$

L'existence et unicité de solution pour cette EDS est obtenue en application du Théorème 4.3.3. L'inférence statistique pour ce processus est étudiée en détail dans les articles d'Overbeck, 1998, Overbeck et Ryden, 1997, Ben Alaya et Kebaier, 2012, Ben Alaya et Kebaier, 2013. Pour certaines valeurs de a, b, σ, on peut obtenir ξ_t à l'aide de processus d'Ornstein-Uhlenbeck indépendants. Considérons $X_i(t), i = 1, \ldots, d$, d processus d'Ornstein-Uhlenbeck indépendants et de même loi définis par:

$$dX_i(t) = \theta X_i(t)dt + cdW_i(t), \quad X_i(0) = x_i$$

où $(W_1, \ldots, W_d)$ sont d mouvements browniens indépendants. Posons $\xi_t = \sum_{i=1}^d X_i^2(t)$. Par la formule d'Ito, on a:

$$d\xi_t = 2X_i(t)dX_i(t) + dc^2dt = (2\theta X_i(t) + dc^2)dt + 2c\sum_{i=1}^d X_i(t)dW_i(t), \xi_0 = \sum_{i=1}^d x_i^2.$$

Posons

$$\beta(t) = \int_0^t 1_{\xi_t>0}\frac{\sum_{i=1}^d X_i(t)dW_i(t)}{\sqrt{\xi_t}}dt + \int_0^t 1_{\xi_t=0}d\beta'(t)$$

où $(\beta'(t))$ est un mouvement brownien standard indépendant de $(W_1, \ldots, W_d)$.On vérifie aisément que $(\beta(t))$ est une martingale locale de variation quadratique $\langle \beta, \beta \rangle_t = t)$. Donc, $(\beta(t))$ est un mouvement brownien standard. On peut écrire:

$$d\xi_t = (2\theta X_i(t) + dc^2)dt + 2c\sqrt{\xi_t}d\beta(t).$$

Le processus ci-dessus a la loi d'un processus CIR avec $a = 2\theta$, $b = dc^2$ (d est entier), $\sigma = 2c$. On peut donc le simuler.

Simulation approchée: schéma d'Euler, schéma d'Alfonsi.

La méthode de simulation par le schéma d'Euler est simple et valable pour un très grand nombre de modèles mais elle n'est qu'approchée. Soit (ε_i) des v.a. i.i.d. de loi gaussienne centrée réduite et soit δ un pas de discrétisation petit. Soit le processus défini par:

$$d\xi_t^x = b(\xi_t^x)dt + \sigma(\xi_t^x)dB_t, \xi_0^x = x. \tag{4.17}$$

On construit:

$$Y_0 = x, \quad Y_i = Y_{i-1} + \delta b(Y_{i-1}) + \sqrt{\delta}\sigma(Y_{i-1})\varepsilon_i, i \geq 1.$$

Ensuite, on choisit un entier p et on peut considérer que $(Y_{ip\delta})$ est une approximation de la trajectoire discrétisée $\xi_{i\Delta}^x$ avec $\Delta = p\delta$.

La méthode du schéma d'Euler s'applique mal pour la diffusion CIR car il s'agit d'un processus positif et le schéma d'Euler simule des variables gaussiennes qui peuvent prendre des valeurs négatives. Le schéma de discrétisation proposé par Alfonsi (2005) est un schéma approché implicite composé de v.a. positives.): Voici les arguments sur lesquels repose la preuve d'Alfonsi. On écrit que, pour $\delta = T/n$:

$$\begin{aligned}
\xi_t &= x + \int_0^t (a\xi_s + b)dt + \sigma \int_0^T \sqrt{\xi_s}dB_s \\
&= x + \lim_{n\to+\infty} \left[\sum_{i:i\delta<t} (a\xi_{(i+1)\delta} + b)\delta + \sigma \sum_{i:i\delta<t} \sqrt{\xi_{(i+1)\delta}}(B_{(i+1)\delta} - B_{i\delta}) \right] \\
&- \lim_{n\to+\infty} \left[\sigma \sum_{i:i\delta<t} (\sqrt{\xi_{(i+1)\delta}} - \sqrt{\xi_{i\delta}})(B_{(i+1)\delta} - B_{i\delta}) \right].
\end{aligned}$$

Comme $d < \sqrt{\xi_s}, B >_s = (\sigma/2)ds$, on obtient:

$$\xi_t = x + \lim_{n\to+\infty} \left[\sum_{i:i\delta<t} (a\xi_{(i+1)\delta} + b - \frac{\sigma^2}{2})\delta + \sigma \sum_{i:i\delta<t} \sqrt{\xi_{(i+1)\delta}}(B_{(i+1)\delta} - B_{i\delta}) \right].$$

Par suite, pour δ petit, on obtient l'approximation suivante:

$$\xi_{(i+1)\delta} \cong \xi_{i\delta} + (a\xi_{(i+1)\delta} + b - \frac{\sigma^2}{2})\delta + \sigma\sqrt{\xi_{(i+1)\delta}}(B_{(i+1)\delta} - B_{i\delta}).$$

On introduit le polynôme du second degré $P(x) = (1-a\delta)x^2 - \sigma x(B_{(i+1)\delta} - B_{i\delta}) - (\xi_{i\delta} + (b - \frac{\sigma^2}{2})\delta)$. Si $b > \frac{\sigma^2}{2}$, $\xi_{i\delta} \geq 0$ et δ assez petit pour que $1 - a\delta > 0$, alors $P(0) < 0$ et on peut prendre la solution positive de l'équation du second degré donnant $\sqrt{\xi_{i+1)\delta}}$ en fonction de $\sqrt{\xi_{i\delta}}$ et obtenir:

$$\xi_{(i+1)\delta} \cong \left[\frac{1}{2(1-a\delta)}\left(\sigma B_{(i+1)\delta} - B_{i\delta} + \sqrt{\sigma^2(B_{(i+1)\delta} - B_{i\delta})^2 + 4(1-a\delta)(\xi_{i\delta} + (b - \frac{\sigma^2}{2})\delta}\right)\right]^2.$$

D'où, pour δ petit, $b > \frac{\sigma^2}{2}$, le processus suivant est une discrétisation approchée de (4.16):

$$Y_0 = x > 0, \quad Y_{i+1} = \left[\frac{1}{2(1-a\delta)}\left(\sigma\sqrt{\delta}\varepsilon_{i+1} + \sqrt{\sigma^2\delta\varepsilon_{i+1}^2 + 4(1-a\delta)(Y_i + (b - \frac{\sigma^2}{2})\delta}\right)\right]^2 \tag{4.18}$$

où (ε_i) sont i.i.d de loi $\mathcal{N}(0,1)$.

Simulation exacte: algorithme rétropectif de Beskos *et al.*

Nous donnons brièvement les étapes de l'algorithme *rétrospectif* de simulation exacte d'une trajectoire discrétisée de diffusion proposé par Beskos *et al.* (2004), dans le cas le plus simple. Considérons le processus défini par:

$$d\xi_t^x = \alpha(\xi_t^x)dt + dB_t, \xi_0^x = x. \tag{4.19}$$

où $\alpha(.)$ est une fonction réelle définie sur $\mathbb{R}$ et (B_t) est un mouvement brownien standard. Il est possible de produire une réalisation exacte de la variable aléatoire ξ_Δ^x pour $\Delta > 0$ sous les hypothèses suivantes:

- $\alpha(.)$ est de classe C^1 et il existe des constantes c_1, c_2 telles que $c_1 \leq \alpha^2 + \alpha' \leq c_2$.
- Soit $A(\xi) = \int_0^\xi \alpha(u)du$. La fonction $h(\xi) = \exp(A(\xi) - (\xi - x)^2/2\Delta)$ est intégrable sur $\mathbb{R}$.
- Il est possible de simuler une réalisation exacte d'une v.a. de densité proportionnelle à h.

On pose $\Phi(.) = \frac{1}{2}(\alpha^2(.) + \alpha'(.) - c_1)$ et $M = (c_2 - c_1)/2$ de sorte que $0 \leq \Phi \leq M$.

Rappelons que C_Δ est l'espace des fonctions réelles continues définies sur $[0, \Delta]$, muni de la tribu $\mathcal{C}_\Delta$ associée à la topologie de la convergence uniforme sur $[0, \Delta]$. On a noté $X = (X_t, 0 \leq t \leq \Delta)$ le processus canonique de C_Δ. Soit P_Δ^x la loi sur $(C_\Delta, \mathcal{C}_\Delta)$ de $(\xi_t^x, 0 \leq t \leq \Delta)$. Soit W_Δ^x la loi sur C_Δ de $(x + B_t, 0 \leq t \leq \Delta)$. On sait que la loi conditionnelle $W_\Delta^x(.|X_\Delta = y)$ sur C_Δ est la loi d'un pont brownien partant de x finissant en y, c'est-à-dire la loi de $(x + \frac{t}{\Delta}(y - x) + B_t - \frac{t}{\Delta}B_\Delta, t \in [0, \Delta])$.
On définit la probabilité Z_Δ^x sur C_Δ comme suit:

- Sous Z_Δ^x, X_Δ est une v.a. de densité proportionnelle à $h(.)$,
- $Z_\Delta^x(.|X_\Delta = y) = W_\Delta^x(.|X_\Delta = y)$.

Alors, on peut montrer:

Proposition 4.3.1. *La probabilité P_Δ^x est absolument continue par rapport à Z_Δ^x de densité*

$$\frac{dP_\Delta^x}{dZ_\Delta^x} \propto \exp\left(-\int_0^\Delta \Phi(X_s)ds\right).$$

Proposition 4.3.2. *Soit τ une v.a. de Poisson de paramètre ΔM et, sachant $\tau = k$, soient $((T_i, V_i), i = 1, \dots, k)$ des v.a. i.i.d. de loi uniforme sur $[0, \Delta] \times [0, M]$. Alors,*

$$N(ds, dv) = \sum_{i=1}^{\tau} \delta_{(T_i,V_i)}(ds, dv)\ 1_{\tau \geq 1}$$

est une mesure aléatoire de Poisson d'intensité $\lambda(ds, dv) = 1_{[0,\Delta]}(s)1_{[0,M]}(v)dsdv$ (δ_z désigne la masse de Dirac en z).

Soit $B(X) = \{(s, v); 0 \leq s \leq \Delta, 0 \leq v \leq \Phi(X_s)\}$. On a:

$$P(N(B(X)) = 0|X) = \exp\left(-\int_0^\Delta \Phi(X_s)ds\right).$$

Rappelons la méthode de simulation par rejet.

Proposition 4.3.3. *Soient μ et ν deux probabilités sur l'espace $(S, \mathcal{S})$. On suppose que l'on sait simuler ν et que $\frac{d\mu}{d\nu} = \frac{1}{\varepsilon}f$ où $f \leq 1$. Soit $((Y_n, I_n), n \geq 1)$ une suite de v.a. i.i.d. à valeurs dans $S \times \{0, 1\}$ et telles que Y_i a la loi ν et $P(I_i = 1|Y_i = y) = f(y)$, $y \in S$. Posons $\kappa = \min\{i \geq 1, I_i = 1\}$. On a $P(\kappa < \infty) = 1$ et Y_κ a la loi μ.*

Enfin, nous pouvons décrire l'algorithme rétrospectif pour ξ_Δ^x.

- (Etape 1) Simuler $X_\Delta = y$ selon une densité proportionnelle à h, (par exemple en utilisant la Proposition 4.3.3);

- (Etape 2) Simuler $\tau = k$, et $((T_i, V_i) = (t_i, v_i), i = 1, \dots, k)$ selon la Proposition 4.3.2,

- (Etape 3) Simuler $(X_{t_i} = x_i, i = 1, \dots, k)$ selon un pont brownien issu de x à l'instant 0 finissant en y à l'instant Δ, aux instants obtenus lors de l'étape 2;

- (Etape 4) Calculer l'indicatrice

  $$I = \prod_{i=1}^{k} 1_{(\Phi(x_i) \leq v_i)}.$$

  Si $I = 1$, c'est-à-dire si $N(B(X)) = 0$, accepter X (en application de Proposition 4.3.3). Sinon, retour à l'étape 1.

A la fin, une réalisation exacte $(x, X_{t_1} = x_1, \dots, X_{t_k} = x_k, X_\Delta = y)$ selon P_Δ^x est obtenue. Ainsi, $X_\Delta = y$ est une réalisation exacte de ξ_Δ^x. Lorsqu'une trajectoire discrète est acceptée, il est possible de continuer de la construire entre les instants t_i en simulant des ponts browniens indépendants issus de x_i à l'instant t_i, finissant en x_{i+1} à l'instant t_{i+1}.

En gardant seulement la dernière valeur y, on peut itérer l'algorithme partant de y et produire une réalisation exacte pour l'instant 2Δ, etc... (On utilise ici la propriété de Markov de (ξ_t)).

L'hypothèse $\Phi(.)$ majorée est une restriction assez forte. Un autre algorithme simple (utilisant d'autres outils probabilistes) est décrit dans Beskos et al (2004) qui suppose seulement $\limsup_{u\to+\infty}\Phi(u)<\infty$ ou $\limsup_{u\to-\infty}\Phi(u)<\infty$.

Pour les modèles de coefficient de diffusion $\sigma(x)$ non constant, si σ est strictement positive et de classe C^1, la transformation $F'(x)=1/\sigma(x)$ permet de se ramener au cas où le coefficient de diffusion est constant. On trouvera dans l'article Comte et al. (2007) des exemples de modèles pouvant être simulés par cet algorithme. Les simulations de trajectoires permettent de créer des données. Les données sont utilisées pour implémenter des estimateurs non paramétriques de la dérive et du coefficient de diffusion.

4.4 Exercices

Exercice 1: Soit $(B_t, t\geq 0)$ est un mouvement brownien standard défini sur l'espace $(\Omega,\mathcal{F},\mathbb{P})$, η une v.a. réelle définie sur Ω, indépendante de $(B_t, t\geq 0)$. On considère l'équation différentielle stochastique :

$$d\xi_t=\theta_0\xi_t dt+\sigma(1+\xi_t^2)^{1/2}dB_t,\quad \xi_0=x_0,$$

où $\theta_0\in\mathbb{R}$ est inconnu, $\sigma>0, x_0\in\mathbb{R}$ sont connus.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution adapté à la filtration $\mathcal{F}_t=\sigma(\eta, B_s, s\leq t), t\geq 0$.
2. Ecrire la vraisemblance de l'observation $(\xi_t, t\in[0,T])$ et calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0 basé sur cette observation.
3. Simuler une discrétisation de (ξ_t) à l'aide de l'algorithme de Beskos *et al* pour différentes valeurs de θ_0,σ,x_0. Discrétiser l'EMV de θ_0 et calculer les valeurs obtenues à partir des données simulées. Répéter les simulations et calculer la moyenne et l'écart-type empirique de l'estimateur.

Exercice 2: Mêmes questions avec

$$d\xi_t=\theta_0\frac{\xi_t}{(1+\xi_t^2)^{1/2}}dt+\sigma dB_t,\quad \xi_0=x_0,$$

où $\theta_0\in\mathbb{R}$ est inconnu, $\sigma>0, x_0\in\mathbb{R}$ sont connus.

Exercice 3: Soit $(B_t, t\geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega,\mathcal{F},\mathbb{P})$, $x_0>0$ fixé. On considre le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t=(a\xi_t+b)dt+\sigma\xi_t dB_t,\quad \xi_0=x_0, \tag{4.20}$$

où $\sigma>0$, $a,b\in\mathbb{R}$, $x_0>0$.

1. Résoudre l'équation différentielle stochastique $dY_t = aY_t dt + \sigma Y_t dB_t, \quad Y_0 = 1$.

2. En posant $\eta_t = 1/Y_t$ et $Z_t = \xi_t \eta_t$, trouver l'l'équation différentielle stochastique vérifée par Z_t. En déduire le processus solution (ξ_t) de (4.20) en fonction de (Y_t).
Solution:
$$\xi_t = Y_t(x_0 + b\int_0^t ds/Y_s), \quad Y_t = \exp\left[(a - (\sigma^2/2))t + \sigma B_t\right].$$

3. On suppose que $a = a_0 \in \mathbb{R}$ et $b = b_0 \geq 0$ sont inconnus. Ecrire la vraisemblance de l'observation $(\xi_t, t \in [0,T])$. Calculer les estimateurs du maximum de vraisemblance de (a_0, b_0) basés sur cette observation.
Solution: Puisque $\xi_t > 0$ pour tout t, la densité de $P^T_{a,b}$ par rapport à $P^T_{0,0}$ existe pour tout couple (a,b) tel que $a \in \mathbb{R}, b \geq 0$ et la log-vraisemblance s'écrit:
$$\ell_T(a,b) = \sigma^{-2}\left(\int_0^T \frac{a\xi_s + b}{\xi_s^2} d\xi_s - \frac{1}{2}\int_0^T \frac{(a\xi_s + b)^2}{\xi_s^2} ds\right).$$
Le vecteur EMV $(\hat{a}_T\ \hat{b}_T)'$ est solution du système linéaire
$$\begin{pmatrix} \hat{a}_T \\ \hat{b}_T \end{pmatrix} = \mathcal{I}_T \begin{pmatrix} \int_0^T \frac{d\xi_s}{\xi_s} \\ \int_0^T \frac{d\xi_s}{\xi_s^2} \end{pmatrix},$$
avec
$$\mathcal{I}_T = \begin{pmatrix} T & \int_0^T \frac{ds}{\xi_s} \\ \int_0^T \frac{ds}{\xi_s} & \int_0^T \frac{ds}{\xi_s^2} \end{pmatrix}.$$
Avec l'inégalité de Cauchy-Schwarz, on vérifie que cette matrice est définie positive (le cas d'égalité n'est pas possible).

4.5 Observations indépendantes de trajectoires.

Supposons que l'on observe sur $[0,T]$ les N processus définis par:
$$d\xi_i(t) = b(\xi_i(t), \theta_0)dt + \sigma(\xi_i(t))dB_i(t), \quad \xi_i(0) = x_0 \tag{4.21}$$
où $B_1, \ldots, B_N$ sont N mouvements browniens standards indépendants. Dans ce cas, les processus $(\xi_i(t), t \in [0,T]), i = 1, \ldots, N)$ sont *i.i.d.*. On se place sous les hypothèses de la section 4.3.4. La vraisemblance associée à cette observation est donc :
$$\theta \to L_{N,T}(\theta) = \prod_{i=1}^N L_T^{(i)}(\theta)$$
avec
$$\ell_{N,T}(\theta) = \log L_{N,T}(\theta) = \sum_{i=1}^N \left(\int_0^T \frac{b(\xi_i(t),\theta)}{\sigma^2(\xi_i(t))} d\xi_i(t) - (1/2)\int_0^T \frac{b^2(\xi_i(t),\theta)}{\sigma^2(\xi_i(t))} dt\right).$$

On peut considérer ici l'asymptotique usuelle N tend vers l'infini avec T fixé. L'étude asymptotique des EMV est beaucoup plus simple. Par exemple, si la dérive est de la forme $b(x,\theta) = \theta\, b(x)$ avec $b(x)$ fonction connue, on obtient l'EMV

$$\hat{\theta}_{N,T} = \frac{\sum_{i=1}^{N} \int_0^T \frac{b(\xi_i(s))}{\sigma^2(\xi_i(s))} d\xi_i(s)}{\sum_{i=1}^{N} \int_0^T \frac{b^2(\xi_i(s))}{\sigma^2(\xi_i(s))} ds}.$$

Pour calculer sa loi asymptotique, on utilise l'équation (4.21):

$$\hat{\theta}_{N,T} = \theta_0 + \frac{\sum_{i=1}^{N} \int_0^T \frac{b(\xi_i(s))}{\sigma^2(\xi_i(s))} dB_i(s)}{\sum_{i=1}^{N} \int_0^T \frac{b^2(\xi_i(s))}{\sigma^2(\xi_i(s))} ds} = \theta_0 + \frac{M_N(T)}{< M_N > (T)}.$$

Si $V := \mathbb{E}_{\theta_0}(\int_0^T \frac{b^2(\xi_1(s))}{\sigma^2(\xi_1(s))} ds < +\infty$, on peut apppliquer la loi des nombres pour le dénominateur normalisé par $1/N$, le théorème de limite centrale pour le numérateur normalisé par $1/\sqrt{N}$. Ce qui permet de déduire:

$$\sqrt{N}(\hat{\theta}_{N,T} - \theta_0) \to_{\mathcal{L}} \mathcal{N}(0, V^{-1}).$$

4.6 Arrêt aléatoire.

Il est parfois intéressant d'observer le processus jusqu'à un instant aléatoire τ. Le modèle statistique canonique associé à un telle observation doit s'écrire différemment. Soit $C(\mathbb{R}^+, \mathbb{R}) := C$ l'espace des fonctions continues définies sur $\mathbb{R}^+$ à valeurs réelles muni de la tribu borélienne $\mathcal{C}$ associée à la topologie de la convergence uniforme sur tout compact.
On définit le processus canonique $(X_t, t \geq 0)$ de $(C, \mathcal{C})$ par $X_t(x) = x(t)$ pour $x \in C$. On a comme plus haut $\mathcal{C} = \sigma(X_t, t \geq 0$.
On définit la filtration canonique: $\mathcal{C}_t^0 = \sigma(X_s, s \leq t)$ pour tout $t \geq 0$ et sa version continue à droite $\mathcal{C}_t = \cup_{s \geq t} \mathcal{C}_s^0$. Le processus canonique (X_t) est continu et adapté à la filtration $\mathcal{C}_t, t \geq 0$.
Sous les hypothèses du théorème 4.3.2 ou 4.3.3, la loi de probabilité P sur $(C, \mathcal{C})$ du processus solution $(\xi_t, t \geq 0)$ solution de (4.11) ne dépend que des fonctions b, σ et de la loi de probabilité μ de la v.a. initiale η.
On se place sous les hypothèses de la section 4.3.4. Pour le modèle (4.12), on définira un processus de vraisemblance $(L_t(\theta), t \geq 0$ où

$$L_t(\theta) = \exp\left[\int_0^t \frac{b(\xi_s, \theta)}{\sigma^2(\xi_s)} d\xi_s - (1/2) \int_0^t \frac{b^2(\xi_s, \theta)}{\sigma^2(\xi_s)} ds\right]$$

représente la vraisemblance *arrêtée* à l'instant t. L'estimateur du maximum de vraisemblance $\hat{\theta}_t$ à l'instant t est défini comme toute solution de l'équation:

$$L(\hat{\theta}_t) = \sup_{\theta \in \Theta} L_t(\theta).$$

Soit τ un temps d'arrêt de la filtration $(\mathcal{F}_t)$ fini p.s.. La vraisemblance de l'observation arrêtée au temps d'arrêt τ est $L_\tau(\theta)$.

Comme exemple d'application, considérons de nouveau le modèle

$$d\xi_t = \theta_0 b(\xi_t)dt + \sigma(\xi_t)dB_t, \xi_0 = x_0,$$

avec b, σ fonctions connues, observé jusqu'à l'instant τ. L'estimateur du maximum de vraisemblance arrêté à τ vaut:

$$\hat{\theta}_\tau = \frac{\int_0^\tau \frac{b(\xi_s)}{\sigma^2(\xi_s)} d\xi_s}{\int_0^\tau \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds}.$$

Examinons le cas particulier où, pour un $a > 0$,

$$\tau = \tau_a = \inf\{t \geq 0, \int_0^t \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds \geq a\}.$$

En supposant que $\tau_a < +\infty$ p.s. (avec des hypothèses sur b et σ permettant de le vérifier),

$$\int_0^{\tau_a} \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds = a.$$

L'estimateur arrêté vaut donc:

$$\hat{\theta}_{\tau_a} = \frac{1}{a} \int_0^{\tau_a} \frac{b(\xi_s)}{\sigma^2(\xi_s)} d\xi_s.$$

Proposition 4.6.1. *Si $\tau_a < +\infty$ p.s., $\sqrt{a}(\hat{\theta}_{\tau_a} - \theta_0)$ a exactement la loi gaussienne $\mathcal{N}(0, 1)$.*

Preuve. Soit $\lambda \in \mathbb{R}$. On considère

$$M_t(\lambda) = exp[i\lambda \int_0^t \frac{b(\xi_s)}{\sigma(\xi_s)} dB_s + \frac{\lambda^2}{2} \int_0^t \frac{b^2(\xi_s)}{\sigma^2(\xi_s)} ds].$$

Le processus $(M_t(\lambda), t \geq 0)$ est une martingale locale de la filtration $(\mathcal{F}_t)$. D'après le théorème d'arrêt, il en est de même de $(M_{t\wedge\tau_a}(\lambda), t \geq 0)$. Comme

$$|M_{t\wedge\tau_a}(\lambda)| \leq \exp[\frac{\lambda^2}{2} a],$$

$(M_{t\wedge\tau_a}(\lambda), t \geq 0)$ est une vraie martingale. Ce qui implique: $\mathbb{E}(M_{t\wedge\tau_a}(\lambda)) = 1$. En faisant tendre t vers l'infini, on conclut par application de théorème de convergence dominée, que $\mathbb{E}(M_{\tau_a}(\lambda)) = 1$. Autrement dit,

$$\mathbb{E} \exp[i\lambda \int_0^{\tau_a} \frac{b(\xi_s)}{\sigma(\xi_s)} dB_s] = \exp[-\frac{\lambda^2}{2} a]$$

Ainsi, $\int_0^{\tau_a} \frac{b(\xi_s)}{\sigma(\xi_s)} dB_s$ a exactement la loi gaussienne $\mathcal{N}(0, a)$ ce qui donne le résultat.□

4.7 Diffusions multidimensionnelles.

Dans ce paragraphe, nous donnons la formule de vraisemblance pour les modèles multidimensionnels. Considérons le modèle d-dimensionnel:

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dW_t, \quad X_0^\theta = x, \tag{4.22}$$

où $x \in \mathbb{R}^d$, $b : \Theta \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, $W = (W^1, \ldots, W^d)'$ est un mouvement brownien d-dimensionnel ($W^1, \ldots, W^d$ sont des mouvements browniens standards independants). On pose:

$$C(x) = \sigma(x)\sigma(x)'. \tag{4.23}$$

On suppose que:

- Les hypothèses du théorème 4.3.2 (multidimensionnel) sont vérifiées pour $x \to b(\theta, x)$ pour tout θ et pour $x \to \sigma(x)$.
- $\forall x \in \mathbb{R}^d, C(x)$ est inversible.

On note P_θ^T la loi du processus $(X_t^\theta, t \in [0,T])$ défini par (4.22) sur l'espace $C([0,T], \mathbb{R}^d)$ des fonctions continues sur $[0,T]$ à valeurs dans $\mathbb{R}^d$, muni de la tribu borélienne associée à la convergence uniforme. On note Q_T la loi du processus $(Y_t, t \in [0,T])$ défini par:

$$dY_t = \sigma(Y_t)dW_t, \quad Y_0 = x.$$

Proposition 4.7.1. *La loi P_θ^T est absolument continue par rapport à Q_T et*

$$\frac{dP_\theta^T}{dQ_T} = \exp\left[\int_0^T b(\theta, X_s)'C^{-1}(X_s)dX_s - \frac{1}{2}\int_0^T b(\theta, X_s)'C^{-1}(X_s)b(\theta, X_s)ds\right].$$

((X_t) est le processus canonique de $C([0,T], \mathbb{R}^d)$).

4.8 Commentaires bibliographiques

Pour ce chapitre, on pourra consulter les ouvrages suivants. Pour les équations différentielles stochastiques, Ikeda, N. et Watanabe, S., 1981, Karatsas, I. et Shreeve, S.E., 1991, Gikhman, I.I. et Skorohod, A.V., 1974 et 1979, Revuz, D. et Yor, M., 1991, Rogers, L.C.G. et Williams, D.,1987 et 2000. Pour la statistique asymptotique en général et l'étude asymptotique des estimateurs du maximum de vraisemblance, Ibragimov, I.A. et Has'minskii, R.Z., 1981. Pour la statistique des processus en général et la statistique des diffusions en particulier, Iacus, S.M., 2010, Jacod, J. et Shiryaev, A.N., 1983, Liptser, R.S. et Shiryaev, A.N., 1977 et 1978, Gikhman, I.I. et Skorohod, A.V., 1974 et 1979, Kessler, M., Lindner, A. et Sørensen, M. (éditeurs), 2012, Kuechler, U. et Sørensen, M., 1997, Kutoyants, Yu.A., 1984 et 2004, Prakasa Rao, B.L.S., 1999.

Références bibliographiques

1. Alfonsi, A., 2005. On the discretization schemes for the CIR (and Bessel squares) processes. *Monte-Carlo methods and Applications*, **11**, 355-384.

2. Ben Alaya, M. et Kebaier, M. , 2012. Parameter estimation for the square root diffusions: ergodic and non ergodic cases. *Stochastic models* **28**, 609-634.

3. Ben Alaya, M. et Kebaier, M. , 2013. Asymptotic behavior of the maximum likelihood estimator for ergodic and non ergodic suare-root diffusions. *Stochastic Analysis and Applications*, **31**, 552-573.

4. Beskos, A., Papaspiliopoulos, O. and Roberts, G.O., 2006. Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12**, 1077-1098.

5. Billingsley, P., 1968. *Convergence of probability measures*, Wiley.

6. Comte F., Genon-Catalot V. et Rozenholc Y., 2007. Penalized nonparametric mean square estimation of the coefficients of diffusion processes. Bermoulli, bf 13, 514-543.

7. Cox, J.C. Ingersoll, J.E. et Ross, S.A., 1985. A Theory of the Term Structure of Interest rates. *Econometrica*, **53**, 385-407.

8. Iacus, S.M., 2010. *Simulation and Inference for Stochastic Differential Equations, with R-examples.* Springer Series in Statistics. Springer New York.

9. Ibragimov, I.A. et Has'minskii, R.Z., 1981. *Statistical estimation, Asymptotic theory*, Springer.

10. Ikeda, N. et Watanabe, S., 1981. *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha.

11. Jacod, J. et Shiryaev, A.N., 1983. *Limit theorems for stochastic processes*, Springer, 2003 (2nde édition).

12. Karatsas, I. et Shreeve, S.E., 1991. *Brownian motion and stochastic calculus, second edition*, Graduate texts in Mathematics, Springer. 4ème édition, 1997.

13. Liptser, R.S. et Shiryaev, A.N., 1977. *Statistics of random processes I, General Theory*, Springer. Seconde édition, 2001.

14. Liptser, R.S. et Shiryaev, A.N., 1978. *Statistics of random processes II, Applications*, Springer. Seconde édition, 2001.

15. Gikhman, I.I. et Skorohod, A.V.. *The theory of stochastic processes*,I, Springer, 1974, II, Springer, 1975, III, Springer, 1979.

16. Kessler, M., Lindner, A. et Sørensen, M. (éditeurs), 2012. *Statistical Methods for Stochastic Differential Equations.* CRC Press, Chapman & Hall, Boca Raton.

17. Kuechler, U. et Sørensen, M., 1997. *Exponential falilies of Stochastic processes.* Springer series in Statistics, Springer.

18. Kutoyants, Yu.A., 1984. *Parameter estimation for stochastic processes*, Research and Exposition in Mathematics 6, Heldermann.

19. Kutoyants, Yu. A., 2004. *Statistical inference for ergodic diffusion processes*, Springer.

20. Overbeck, L., 1998. Estimation for continuous branching processes. *Scandinavian Journal of Statistics*, **25**, 111-126.

21. Overbeck, L. et Ryden, T., 1997. Estimation in the Cos-Ingersoll-Ross model. *Econometric Theory*, **13**, 430-461.

22. Prakasa Rao, B.L.S., 1999. *Statistical inference for diffusion type processes*, London and Oxford University press.

23. Revuz, D. et Yor, M., 1991. *Continuous martingales and Brownian motion*, Springer, Berlin.

24. Rogers, L.C.G. et Williams, D., 2000. *Diffusions, Markov processes and martingales, Volume 1, Foundations*, Cambridge university press, (Seconde édition).

25. Rogers, L.C.G. et Williams, D., 1987. *Diffusions, Markov processes and martingales, Volume 2, Itô calculus*, Wiley.

Chapter 5

Processus d'Ornstein-Uhlenbeck. Estimateur du maximum de vraisemblance basé sur une observation en temps continu.

Dans ce chapitre, nous étudions l'estimateur du maximum de vraisemblance exact du paramètre θ_0 du modèle

$$d\xi_t = \theta_0 \xi_t dt + \sigma dB_t, \quad \xi_0 = x, \tag{5.1}$$

basé sur l'observation en temps continu de la trajectoire $(\xi_t, t \in [0,T])$. Le processus (B_t) est un mouvement brownien standard et la condition initiale $x \in \mathbb{R}$ est déterministe et connue. La log-vraisemblance s'écrit:

$$\ell_T(\theta) = \theta \int_0^T \xi_t d\xi_t - \frac{\theta^2}{2} \int_0^T \xi_t^2 dt.$$

L'EMV est égal à:

$$\hat{\theta}_T = \frac{\int_0^T \xi_t d\xi_t}{\int_0^T \xi_t^2 dt}. \tag{5.2}$$

La formule ci-dessus est celle qui permet de calculer l'estimateur sur les données observées $(\xi_t, t \in [0,T])$. Une autre formule ne contenant pas d'intégrale stochastique peut être utilisée:

$$\hat{\theta}_T = \frac{\xi_T^2 - x^2 - T}{2 \int_0^T \xi_t^2 dt}. \tag{5.3}$$

Bien entendu, pour calculer effectivement cet estimateur, il faudra utiliser une discrétisation de l'intégrale du dénominateur.

Nous avons déjà vu que l'équation différentielle stochastique (5.1) a une solution explicite:

$$\xi_t = x \exp(\theta_0 t) + \exp(\theta_0 t) \int_0^t \exp(-\theta_0 s) dB_s. \tag{5.4}$$

Pour étudier la loi de l'estimateur, on utilise la formule faisant apparaître le mouvement brownien (qui n'est pas observé), obtenue en remplaçant dans la formule (5.2) $d\xi_t$ par son expression en fonction de dB_t, dans l'intégrale stochastique du numérateur. Ce qui conduit à la formule:

$$\hat{\theta}_T = \theta_0 + \frac{\int_0^T \xi_t dB_t}{\int_0^T \xi_t^2 dt}. \tag{5.5}$$

On a:

$$\hat{\theta}_T = \theta_0 + \frac{M_T}{< M >_T}, \tag{5.6}$$

où

$$M_t = \int_0^T \xi_t dB_t \quad \text{et} \quad < M >_T = \int_0^T \xi_t^2 dt. \tag{5.7}$$

Nous allons étudier la consistance et la loi asymptotique de cet estimateur lorsque T tend vers l'infini. L'intérêt du modèle est qu'il est l'un des rares pour lesquels une étude complète est possible, pour toutes les valeurs de θ_0. Nous distinguerons les trois cas $\theta_0 < 0$ (récurrent positif), $\theta_0 = 0$ (récurrent nul, le processus observé est un mouvement brownien) et $\theta_0 > 0$ (transient). L'estimateur est consistant dans tous les cas, mais les lois asymptotiques sont différentes.

Pour simplifier un peu les calculs, nous supposerons que $x = 0$. Les résultats analogues pour $x \neq 0$ peuvent être obtenus sans difficulté. Notons que

$$(\ell_T)'(\theta_0) = \int_0^T \xi_t d\xi_t - \theta_0 \int_0^T \xi_t^2 dt = M_T, \quad (\ell_T)''(\theta_0) = -\int_0^T \xi_t^2 dt = -< M >_T .$$

La dérivée première en θ_0 de la log-vraisemblance est une martingale de variation quadratique appelée *information de Fisher stochastique*, ici égale à $< M >_T$. Dans ce modèle, on a de plus, la dérivée seconde en θ_0 exactement égale à $-< M >_T$.

5.1 Cas $\theta_0 < 0$.

Proposition 5.1.1. *Lorsque $\theta_0 < 0$, $< M >_T /T = (1/T)\int_0^T \xi_s^2 ds$ converge p.s. et dans L^2 vers* $1/(2|\theta_0|)$

Preuve. La convergence p.s. sera montrée plus loin à l'aide du théorème ergodique pour les diffusions récurrentes positives.
En utilisant la forme explicite de (ξ_t), on peut calculer la limite L^2. En effet, avec $x = 0$, on a

$$\xi_t = \exp{(\theta_0 t)} \int_0^t \exp{(-\theta_0 s)} dB_s. \tag{5.8}$$

On sait ainsi que (ξ_t) est un processus gaussien centré de fonction de covariance :

$$\mathrm{Cov}(\xi_s, \xi_{s'}) = \exp{(2\theta_0(s' - s))} \frac{\exp{(2\theta_0 s)} - 1}{2\theta_0}, \quad s \leq s'. \tag{5.9}$$

De plus, la formule suivante, valable pour les processus gaussiens, donne la fonction de covariance du processus (ξ_t^2):

$$E(\xi_s^2 \xi_{s'}^2) = E(\xi_s^2)E(\xi_{s'}^2) + 2(E(\xi_s \xi_{s'}))^2 = \mathrm{Var}(\xi_s)\mathrm{Var}(\xi_{s'}) + 2(\mathrm{Cov}(\xi_s, \xi_{s'}))^2, \quad s \leq s'. \tag{5.10}$$

Pour obtenir la convergence dans L^2 cherchée, il suffit de montrer que:

$$\mathbb{E}(\frac{1}{T} < M >_T) \to 1/(2|\theta_0|), \quad \mathbb{E}(\frac{1}{T} < M >_T)^2 \to (1/(2|\theta_0|))^2. \tag{5.11}$$

On a:

$$\mathbb{E}(\frac{1}{T} < M >_T) = \frac{1}{T}\int_0^T \mathbb{E}\xi_s^2 ds = \frac{1}{T}\int_0^T (\frac{e^{2\theta_0 s} - 1}{2\theta_0} ds = -\frac{1}{2\theta_0} + \frac{e^{2\theta_0 T} - 1}{(2\theta_0)^2 T} \to -\frac{1}{2\theta_0},$$

puisque $\theta_0 < 0$. Ensuite, on utilise:

$$\mathbb{E}(\frac{1}{T} < M >_T)^2 = \frac{1}{T^2}\int_0^T \int_0^T \mathbb{E}(\xi_s^2 \xi_{s'}^2) ds ds' = 2\int_{0 \leq s \leq s' \leq T} \mathbb{E}(\xi_s^2 \xi_{s'}^2) ds ds'.$$

Le calcul de l'intégrale double est élémentaire (un peu pénible) et on obtient (5.11) en utilisant (5.9) et (5.10). □

Remarque 5.1.1. *On a utilisé la propriété suivante des vecteurs gaussiens. Si* $(\xi_1, \xi_2, \xi_3, \xi_4)$ *est un vecteur gaussien, alors:*

$$\mathbb{E}\xi_1\xi_2\xi_3\xi_4 = \mathbb{E}\xi_1\xi_2\mathbb{E}\xi_3\xi_4 + \mathbb{E}\xi_1\xi_3\mathbb{E}\xi_2\xi_4 + \mathbb{E}\xi_1\xi_4\mathbb{E}\xi_2\xi_3$$

qui se démontre en dérivant 4 fois la fonction caractéristique du vecteur.

Corollaire 5.1.1. *L'EMV* $\hat{\theta}_T$ *vérifie, lorsque* T *tend vers l'infini:*

$$\hat{\theta}_T \to \theta_0 \ p.s. \ et \ \sqrt{T}(\hat{\theta}_T - \theta_0) \to_{\mathcal{L}} \mathcal{N}(0, 2|\theta_0|).$$

Preuve. $< M >_T$ converge en probabilité vers $+\infty$ et comme $< M >_T$ est croissant avec T, la convergence a lieu p.s.. Donc, $M_T / < M >_T$ tend vers 0 p.s. et $\hat{\theta}_T$ est fortement consistant. Pour la convergence en loi, on écrit:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = \frac{M_T}{\sqrt{T}} \frac{T}{< M >_T}.$$

On conclut par le théorème de Slutsky. □

5.2 Cas où $\theta_0 = 0$.

Proposition 5.2.1. *Si* $\theta_0 = 0$, $T\hat{\theta}_T$ *a une loi fixe, ne dépendant pas de* T, *la loi de* $\frac{\int_0^1 W_u dW_u}{\int_0^1 W_u^2 du}$ *où* (W_u) *est un mouvement brownien standard. L'EMV est donc consistant et converge en loi à la vitesse* T *avec limite en loi non gaussienne.*

Preuve. Introduisons le mouvement brownien standard $(B_u^T = B_{uT}/\sqrt{T}, u \geq 0)$. L'EMV s'écrit:

$$\hat{\theta}_T = \frac{\int_0^T B_s dB_s}{\int_0^T B_s^2 ds} = \frac{(B_T^2 - T)/2}{\int_0^T B_s^2 ds} = \frac{T((B_1^T)^2 - 1)/2}{T^2 \int_0^1 (B_u^T)^2 du} = \frac{\int_0^1 B_u^T dB_u^T}{T \int_0^1 (B_u^T)^2 du}$$

□

5.3 Cas où $\theta_0 > 0$.

Proposition 5.3.1. *On suppose $\theta_0 > 0$. Soient $Z = \int_0^{+\infty} e^{-\theta_0 s} dB_s$ et*

$$m_T(\theta_0) = \int_0^T e^{2\theta_0 s} ds = \frac{e^{2\theta_0 T} - 1}{2\theta_0}.$$

Quand T tend vers l'infini,

$$\frac{< M >_T}{m_T(\theta_0)} \to_{L^1} Z^2.$$

Preuve. Comme $\theta_0 > 0$, $\int_0^{+\infty} e^{-2\theta_0 s} ds = 1/(2\theta_0) < +\infty$.. Donc, la v.a. Z est bien définie et a la loi gaussienne centrée, de variance $1/(2\theta_0)$. En utilisant (5.8), on voit que, lorsque t tend vers l'infini,

$$\frac{\xi_t}{e^{\theta_0 t}} := N_t = \int_0^t e^{-\theta_0 s} dB_s \to_{L^2} Z. \tag{5.12}$$

On en déduit aisément que:

$$\left(\frac{\xi_t}{e^{\theta_0 t}}\right)^2 \to_{L^1} Z^2. \tag{5.13}$$

Etudions la différence

$$D_T = \frac{1}{\int_0^T e^{2\theta_0 s} ds} \int_0^T (\xi_s e^{-\theta_0 s})^2 e^{2\theta_0 s} ds - Z^2 := D_T^{(1)} + D_T^{(2)},$$

avec, pour tout $T_0 \in [0, T]$,

$$D_T^{(1)} = \frac{1}{m_T(\theta_0)} C_{T_0}, \quad C_{T_0} = \int_0^{T_0} [(\xi_s e^{-\theta_0 s})^2 - Z^2] e^{2\theta_0 s} ds,$$

et

$$D_T^{(2)} = \frac{1}{m_T(\theta_0)} \int_{T_0}^T e^{2\theta_0 s} [(\xi_s e^{-\theta_0 s})^2 - Z^2] ds.$$

En utilisant la convergence L^1 (5.13), pour tout $\varepsilon > 0$, on fixe T_0 tel que, pour tout $s \geq T_0$,

$$\mathbb{E}|(\xi_s e^{-\theta_0 s})^2 - Z^2| \leq \varepsilon.$$

Ainsi, pour tout $T \geq T_0$, $\mathbb{E}|D_T^{(2)}| \leq \varepsilon$. D'autre part, lorsque T tend vers l'infini,

$$\mathbb{E}|D_T^{(1)}| \leq \frac{1}{m_T(\theta_0)} \mathbb{E}|C_{T_0}| \to 0.$$

D'où le résultat. □

Théorème 5.3.1. *L'EMV est fortement consistant et:*

1. *$(m_T(\theta_0))^{1/2}(\hat{\theta}_T - \theta_0) \to_{\mathcal{L}} \frac{U}{Z}$ où U et Z sont indépendantes, U a la loi $\mathcal{N}(0,1)$ et Z a la loi $\mathcal{N}(0, 1/(2\theta_0))$ (loi de la v.a. Z ci-dessus).*

2. $(< M >_T)^{1/2}(\hat{\theta}_T - \theta_0) \to_{\mathcal{L}} \mathcal{N}(0,1)$

Preuve. La consistance forte résulte de ce que $< M >_\infty = \infty$ p.s.. Posons $K_t = \int_0^t e^{\theta_0 s} dB_s$. Par application de la formule d'Itô, on a (voir la notation (5.12)):

$$M_T = \int_0^T N_s dK_s = K_T N_T - \int_0^T K_s dN_s - T = K_T(Z - \int_T^{+\infty} e^{-\theta_0 s} dW_s) - \int_0^T K_s dN_s - T.$$

On introduit la v.a. gaussienne centrée réduite suivante:

$$U_T = \frac{1}{(m_T(\theta_0))^{1/2}} K_T.$$

On a:

$$\frac{1}{(m_T(\theta_0))^{1/2}} M_T = Z U_T + R_T,$$

avec

$$R_T = -U_T \int_T^{+\infty} e^{-\theta_0 s} dB_s - \frac{1}{(m_T(\theta_0))^{1/2}} \int_0^T K_s dN_s - \frac{T}{(m_T(\theta_0))^{1/2}}.$$

Nous allons montrer que R_T tend vers 0 en probabilité et que le couple (Z, U_T) converge en loi vers $\mathcal{N}(0, 1/(2\theta_0)) \otimes \mathcal{N}(0,1)$.
Ce dernier point est simple car le couple (Z, U_T) est gaussien, de lois marginales fixes égales respectivement à la loi $\mathcal{N}(0, 1/(2\theta_0))$ et à la loi $\mathcal{N}(0,1)$. Il reste à vérifier que

$$\mathrm{Cov}(Z, U_T) = \mathbb{E}(Z U_T) = \frac{T}{(m_T(\theta_0))^{1/2}}$$

tend vers 0 lorsque T tend vers l'infini ce qui est bien le cas.

Etudions R_T. Le troisième terme tend vers 0 quand T tend vers l'infini. Pour le premier terme, on voit que U_T a une loi fixe, la loi $\mathcal{N}(0,1)$ et $\int_T^{+\infty} e^{-\theta_0 s} dB_s$ tend vers 0 dans L^2. Donc le produit $U_T \int_T^{+\infty} e^{-\theta_0 s} dB_s$ tend vers 0 dans L^1. Pour le deuxième terme de R_T, on calcule:

$$\mathbb{E}\left(\frac{1}{(m_T(\theta_0))^{1/2}} \int_0^T K_s dN_s\right)^2 = \frac{1}{m_T(\theta_0)} \mathbb{E} \int_0^T e^{-2\theta_0 s} (\int_0^s e^{\theta_0 u} dB_u)^2 ds$$

$$= \frac{1}{m_T(\theta_0)} \int_0^T e^{-2\theta_0 s} (\int_0^s e^{2\theta_0 u} du)^2 ds \leq \frac{T}{2\theta_0 m_T(\theta_0)} \to 0.$$

On obtient $R_T \to 0$ dans L^1. En conclusion, on a obtenu:

$$\frac{1}{(m_T(\theta_0))^{1/2}} \int_0^T \xi_s dB_s = Z U_T + o_P(1),$$

$$\frac{1}{m_T(\theta_0)} \int_0^T \xi_s^2 ds = Z^2 + o_P(1),$$

avec (Z, U_T) qui converge en loi vers (Z, U) de loi $\mathcal{N}(0, 1/(2\theta_0)) \otimes \mathcal{N}(0, 1)$. On en déduit:

$$(m_T(\theta_0))^{1/2}(\hat{\theta}_T - \theta_0) = \frac{U_T}{Z} + o_P(1) \to_{\mathcal{L}} \frac{U}{Z},$$

$$\int_0^T \xi_s^2 ds(\hat{\theta}_T - \theta_0) = \frac{ZU_T}{|Z|} + o_P(1) \to_{\mathcal{L}} U.$$

Ce qui achève la preuve. □

Remarque 5.3.1. • *L'EMV converge à la vitesse $(m_T(\theta_0))^{1/2}$ beaucoup plus rapide que dans les deux cas précédents. Sa loi limite avec la normalisation déterministe $(m_T(\theta_0))^{1/2}$ est non gaussienne. Il s'agit d'une loi dite mélange en variance de lois gaussiennes, en l'occurence une loi de Cauchy (quotient de deux gaussiennes indépendantes centrées). Avec la normalisation aléatoire $(< M >_T)^{1/2}$, la loi limite de l'EMV est gaussienne.*

- *Dans le livre Genon-Catalot et Picard (1993, chapitre 3, p.79-84), ce modèle est étudié par la méthode de Sweeting (1980) qui s'applique aux modèles à information de Fisher aléatoire.*

5.4 Exercices

Exercice 1: (Modèle à deux paramètres) On considère l'EDS:

$$d\xi_t = \alpha_0(\xi_t - \gamma_0)dt + dB_t, \quad \xi_0 = x_0,$$

où $x_0 \in \mathbb{R}$ est connu, $\theta_0 = (\alpha_0, \gamma_0) \in \mathbb{R}^2$ est inconnu.

1. Poser $Y_t = \xi_t - \gamma_0$ et résoudre l'EDS.
2. Poser $a_0 = \alpha_0, b_0 = -\alpha_0\gamma_0$. Calculer les estimateurs du maximum de vraisemblance de (a_0, b_0).
3. Etudier les estimateurs précédents en distinguant les trois cas: $\alpha_0 < 0, = 0, > 0$. (On trouvera que l'estimateur de a_0 a trois comportements différents, tandis que l'estimateur de b_0 converge toujours à la même vitesse).
4. Etudier les estimateurs de (α_0, γ_0) déduits de ceux de (a_0, b_0) (γ_0 n'est pas défini pour $\alpha_0 = 0$).

Exercice 2: (Exemple proposé par Mishra et Prakasa Rao (1985)). On considère l'EDS:

$$d\xi_t = \theta_0 t\xi_t dt + dB_t, \quad \xi_0 = x_0,$$

où $x_0 \in \mathbb{R}$ est connu, $\theta_0 \geq 0$ est inconnu.

1. Résoudre l'EDS.
2. Calculer l'estimateur du maximum de vraisemblance de θ_0.
3. Etudier l'estimateur précédent en distinguant les deux cas: $\theta = 0, > 0$.

5.5 Processus d'Ornstein-Uhlenbeck multidimensionnel.

Considérons l'équation différentielle stochastique linéaire multidimensionnelle:

$$dX_t = (A + BX_t)dt + \Sigma dW_t, \quad X_0 = \eta \tag{5.14}$$

où A est un vecteur de $\mathbb{R}^d$, B et Σ sont des matrices $d \times d$, W est un mouvement brownien de $\mathbb{R}^d$, η est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ indépendant de W. L'équation admet un unique processus solution (X_t) qui a pour dimension d. (On aurait pu supposer que Σ et de dimension $d \times m$ et que W est un mouvement brownien de $\mathbb{R}^m$)

Pour obtenir une résolution explicite simple de l'équation (5.14), nous supposons:

- (H1) La matrice B est inversible.
- (H2) La matrice B est diagonalisable dans $\mathbb{R}$.

Sous (H1), le processus $Y_t = B^{-1}A + X_t$ est solution de:

$$dY_t = BY_t dt + \Sigma dW_t, \quad Y_0 = B^{-1}A + \eta. \tag{5.15}$$

Nous résolvons donc (5.15). Sous (H2), on peut trouver une matrice réelle inversible P telle que $P^{-1}BP = D = \text{diag}(\lambda_1, \ldots, \lambda_d)$ avec $\prod_{i=1,\ldots,d} \lambda_i \neq 0$. Si l'on pose $Z_t = P^{-1}Y_t$, le processus (Z_t) est solution de l'équation:

$$dZ_t = DZ_t dt + \Gamma dW_t, \quad Z_0 = P^{-1}Y_0, \tag{5.16}$$

avec $\Gamma = P^{-1}\Sigma$. Puisque D est diagonale, on peut résoudre cette équation comme dans le cas unidimensionnel en cherchant Z_t sous la forme $Z_t^i = K_t^i \exp \lambda_i t, i = 1, \ldots, d$. Posons $e^{Dt} = \text{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_d t})$. Il vient:

$$K_t = K_0 + \int_0^t e^{-Ds}\Gamma dW_s. \tag{5.17}$$

Ainsi,

$$Z_t = e^{Dt}K_t = e^{Dt}Z_0 + e^{Dt}\int_0^t e^{-Ds}\Gamma dW_s, \tag{5.18}$$

et

$$Y_t = Pe^{Dt}P^{-1}Y_0 + Pe^{Dt}\int_0^t e^{-Ds}\Gamma dW_s. \tag{5.19}$$

Puisque $B = PDP^{-1}$, on a $e^{Bt} = Pe^{Dt}P^{-1}$ et $e^{-Bt} = Pe^{-Dt}P^{-1}$. Donc,

$$Y_t = e^{Bt}Y_0 + e^{Bt}\int_0^t e^{-Bs}\Sigma dW_s. \tag{5.20}$$

Le processus (Y_t) est gaussien, d'espérance $e^{Bt}\mathbb{E}(Y_0)$, de matrice de covariance:

$$\text{Cov}(Y_t) = e^{Bt}\text{Cov}(Y_0)e^{B't} + e^{Bt}\int_0^t e^{-Bs}\Sigma\Sigma' e^{-B's}ds\, e^{B't}. \tag{5.21}$$

Donc,

$$\mathrm{Cov}(Y_t) = \int_0^t e^{Bu}\Sigma\Sigma' e^{B'u} du + e^{Bt}\mathrm{Cov}(Y_0)e^{B't}. \tag{5.22}$$

De plus, pour $t, h \geq 0$,

$$Y_{t+h} = e^{Bh}Y_t + e^{B(t+h)} \int_t^{t+h} e^{-Bs}\Sigma dW_s. \tag{5.23}$$

Conditionnellement à $(Y_s, s \leq t)$, Y_{t+h} est gaussien, d'espérance $e^{Bh}Y_t$, de matrice de covariance:

$$e^{-B(t+h)} \int_t^{t+h} e^{-Bs}\Sigma\Sigma' e^{-B's} ds e^{B'(t+h)} = \int_0^h e^{Bu}\Sigma\Sigma' e^{B'u} du = P \int_0^h e^{Du}\Gamma\Gamma' e^{Du} du P'$$

Notons $\Gamma = (\gamma_{ij})$. Alors, $\Gamma\Gamma' = (\alpha_{ij})$ avec $\alpha_{ij} = \sum_{l=1}^d \gamma_{il}\gamma_{jl}$. On a:

$$\int_0^h e^{Du}\Gamma\Gamma' e^{Du} = \left(\frac{e^{(\lambda_i+\lambda_j)h} - 1}{\lambda_i + \lambda_j} \alpha_{ij} \right).$$

Proposition 5.5.1. *Le processus (X_t) admet une loi stationnaire si et seulement si, pour tout $i = 1, \ldots, d$, $\lambda_i < 0$. Dans ce cas, la loi stationnaire est unique et égale à la loi gaussienne d'espérance $-B^{-1}A$, de matrice de covariance*

$$V = \int_0^{+\infty} e^{Bu}\Sigma\Sigma' e^{B'u} du = P \int_0^{+\infty} e^{Ds}\Gamma\Gamma' e^{Ds} ds P' = \left(-\frac{\alpha_{ij}}{\lambda_i + \lambda_j} \right)_{1 \leq i,j \leq d}.$$

Preuve. Raisonnons sur (Z_t). Le vecteur gaussien Z_t converge en loi lorsque t tend vers l'infini si et seulement son espérance converge vers un vecteur de $\mathbb{R}^d$ et sa matrice de covariance converge vers une matrice positive ou nulle. Ceci n'est possible que si tous les λ_i sont strictement négatifs. Dans ce cas, $\mathbb{E}(Z_t) \to 0$ et $\mathrm{Cov}(Z_t) \to K := \int_0^{+\infty} e^{Ds}\Gamma\Gamma' e^{Ds} ds$. On en déduit que Y_t converge en loi vers la loi $\mathcal{N}_d(0, V = PKP')$. et le résultat pour $X_t = Y_t - B^{-1}A$.

Si nous supposons que Z_0 suit la loi $\mathcal{N}_d(0, K)$, alors Z_t suit la loi gaussienne centrée de matrice de covariance

$$e^{Dt}Ke^{Dt} + e^{Dt} \int_0^t e^{-Ds}\Gamma\Gamma' e^{-Ds} ds \, e^{Dt} = \left(-\alpha_{ij} \frac{e^{(\lambda_i+\lambda_j)t}}{\lambda_i + \lambda_j} \right) + \left(\frac{e^{(\lambda_i+\lambda_j)t} - 1}{\lambda_i + \lambda_j} \alpha_{ij} \right) = K.$$

La loi $\mathcal{N}_d(0, K)$ est donc l'unique loi stationnaire pour (Z_t).

5.6 Rérérences bibliographiques

1. Genon-Catalot V., Picard D., 1993. *Eléments de statistique asymptotique*, Springer France.

2. Mishra M.N., Prakasa Rao, B.L.S., 1985. Asymptotic study of the maximum likelihood estimator for non-homogeneous diffusion processes. Statistics and Decisions **3**, 193-203.

3. Sweeting T.J., (1980), Uniform asymptotic normality of the maximum likelihood estimator, Annals of Statist., **8**, 1375-1381.

Chapter 6

Equations différentielles stochastiques autonomes.

Dans ce chapitre, nous présentons des propriétés des équations différentielles stochastiques autonomes utiles pour l'inférence statistique.

6.1 Modèle et hypothèses. Propriété de Markov.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, muni d'une filtration $(\mathcal{F}_t, t \geq 0)$ satisfaisant aux *conditions habituelles* et (B_t) un $\mathbb{P}$-$(\mathcal{F}_t, t \geq 0)$ mouvement brownien standard. Considérons l'équation différentielle stochastique:

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dB_t, \quad \xi_0 = \eta, \tag{6.1}$$

où η est une v.a. réelle, définie sur Ω, de loi μ, $\mathcal{F}_0$-mesurable, les fonctions $b, \sigma : \mathbb{R} \to \mathbb{R}$ vérifient les hypothèses suivantes:

(1) ou bien : Conditions (A')
b, σ sont définies et de classe C^1 sur $\mathbb{R}$ et

$$\exists K > 0, \;\; \forall x, \; |b(x)| + |\sigma(x)| \leq K(1 + |x|).$$

(2) ou bien: Conditions (B')
b, σ sont définies sur $\mathbb{R}$, b est lipschitzienne, σ est Höldérienne d'exposant $\alpha \in [1/2, 1]$:

$$\exists L > 0, \forall x, y, |b(x) - b(y)| \leq L|x - y|, \quad |\sigma(x) - \sigma(y)| \leq L|x - y|^{\alpha}.$$

Alors, l'EDS (6.1) vérifie les conditions du théorème 4.3.2 ou du théorème 4.3.3. Par conséquent, l'équation (6.1) admet un unique (à une indistinguabilité près) processus solution $(\xi_t, t \geq 0)$ à trajectoires continues, adapté à $(\mathcal{F}_t, t \geq 0)$. La loi de probabilité de $(\xi_t, t \in [0, T])$ sur $(C_T, \mathcal{C}_T)$ ne dépend que de (μ, b, σ) (*cf* Théorème 4.3.4). Nous pouvons énoncer:

Théorème 6.1.1. *Le processus solution (ξ_t) est un processus de Markov, de probabilité de transition homogène dans le temps qui ne dépend que des fonctions b et σ. Ceci signifie:*

- *Pour tout $B \in \mathcal{B}(\mathbb{R})$ et pour tous s, t vérifiant $0 \leq s \leq t$, ($\mathbb{P}$-p.s.):*

$$\mathbb{P}(\xi_t \in B|\mathcal{F}_s) = \mathbb{P}(\xi_t \in B|\xi_s).$$

- *La probabilité conditionnelle $\mathbb{P}(\xi_t \in .|\xi_s = x)$ ne dépend que de $t - s$. Plus précisément, il existe une fonction de transition $P_t(x, B)$ définie sur $\mathbb{R}^+ \times \mathbb{R} \times \mathcal{B}(\mathbb{R})$ telle que, pour tout $B \in \mathcal{B}(\mathbb{R})$ et pour tous s, t vérifiant $0 \leq s \leq t$, ($\mathbb{P}$-p.s.):*

$$P_{t-s}(\xi_s, B) = \mathbb{P}(\xi_t \in B|\xi_s)$$

  et qui jouit des propriétés suivantes:

  - $\forall t \geq 0, P_t(x, .)$ *est une probabilité sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,
  - $\forall x \in \mathbb{R}$, $P_0(x, .) = \delta_x$,
  - $\forall B \in \mathcal{B}(\mathbb{R})$, $(t, x) \to P_t(x, B)$ *est borélienne sur* $(\mathbb{R}^+ \times \mathbb{R})$,
  -

$$\forall s, t \geq 0, \ \ \forall B, \ \ \forall x, \ \ P_{t+s}(x, B) = \int_{\mathbb{R}} P_s(y, B) P_t(x, dy).$$

Pour $f : \mathbb{R} \to \mathbb{R}$, on pose, lorsque l'intégrale a un sens,

$$P_t f(x) = \int_{\mathbb{R}} f(y) P_t(x, dy).$$

On peut déduire du théorème que $(t, x) \to P_t f(x)$ est borélienne sur $(\mathbb{R}^+ \times \mathbb{R})$ et

$$\forall s, t \geq 0, \ \ \forall x, \ \ P_{t+s} f(x) = P_t(P_s f)(x).$$

On a ($\mathbb{P}$-p.s.):

$$\mathbb{E}[f(\xi_{t+s})|\xi_s] = P_t f(\xi_s).$$

Soit (ξ_t^x) le processus solution de (6.1) correspondant à la condition initiale $\xi_0^x = x$. On a, pour toute loi initiale μ, lorsque ces expressions ont un sens:

$$P_t f(x) = \mathbb{E} f(\xi_t^x) = \mathbb{E}[f(\xi_t)|\xi_0 = x] \quad \mu - \text{p.s.}.$$

La loi de probabilité de $(\xi_t^x, t \geq 0)$ est identique à la loi conditionnelle de $(\xi_t, t \geq 0)$ sachant $\xi_0 = x$ pour toute loi initiale μ (la loi de (ξ_t^x) est une version régulière de la loi conditionnelle de (ξ_t) sachant $\xi_0 = x$).

6.2 Densités de transition.

Sous d'assez faibles hypothèses sur b, σ (*cf* par exemple, Rogers et Williams, 1987, p.253), les probabilités de transition $P_t(x, dy)$ admettent une densité par rapport à la mesure de Lebesgue:

$$P_t(x, dy) = p_t(x, y) dy.$$

Nous supposerons ces hypothèses vérifiées et donnons quelques exemples de densités de transition explicites.

6.2.1 Exemples.

Exemple 1. Mouvement brownien avec dérive: $d\xi_t = \theta dt + \sigma dB_t$,

$$p_t(x,y) = \frac{1}{\sigma\sqrt{2\pi\, t}} \exp\,[-\frac{1}{2\sigma^2}(y-x-\theta t)^2].$$

Exemple 2. Processus d'Ornstein-Uhlenbeck: $d\xi_t = \theta\xi_t dt + \sigma dB_t$. On a vu que

$$\xi_t^x = xe^{\theta t} + \sigma e^{\theta t}\int_0^t e^{-\theta s} dB_s.$$

D'où,

$$p_t(x,y)dy = \mathcal{N}(xe^{\theta t}, \sigma^2\ \frac{e^{2\theta t}-1}{2\theta}).$$

Exemple 3. Mouvement brownien géométrique: $d\xi_t = \theta\xi_t dt + \sigma\xi_t dB_t$. Alors,

$$\xi_t^x = xe^{[(\theta-(\sigma^2/2))t+\sigma B_t]}$$

et $p_t(x,y)$ est une densité log-normale qui peut être calculée.

Exemple 4. Processus de C.I.R. (diffusion en racine carrée)(*cf* Cox *et al.*, 1984): $d\xi_t = k(\theta-\xi_t)dt + \sigma\sqrt{\xi_t^+}dB_t$. Avec les conditions $k>0, 2k\theta/\sigma^2 \geq 1$, on peut montrer que, pour tout $x>0$, $\mathbb{P}(\xi_t^x > 0, \forall t \geq 0) = 1$. Avec les notations

$$q = (2k\theta/\sigma^2) - 1, \quad c = 2k/(\sigma^2(1-e^{-kt})), \quad u = cxe^{-kt}, \quad ,v = cy,$$

$$p_t(x,y) = c\exp\left(-u-v\right)\,(v/u)^{q/2} I_q(2\sqrt{uv}))$$

où I_q est la fonction de Bessel modifiée d'indice q, de première espèce. Plus concrètement,

$$\xi_t^x e^{kt} \sim (\chi')^2(\delta, x/t'), \quad t' = \frac{\sigma^2}{4k}(e^{kt}-1), \quad \delta = 4k\theta/\sigma^2 = 2q+2$$

où $(\chi')^2(\delta, x/t')$ désigne la loi du chi-deux décentrée de nombre de degrés de liberté δ et de paramètre de non-centralité x/t'.

Ce dernier exemple montre que ces densités de transition même lorqu'elles sont explicites, ont des expressions complexes.

Exercice. Soit $W^i, i = 1,\ldots,n$ n mouvements browniens standards indépendants, $d\xi_t^i = \theta\xi_t^i dt + \sigma dW_t^i, \xi_0^i = x_0^i$, n processus d'Ornstein-Uhlenbeck indépendants. On pose $X_t = \sum_{i=1}^n (\xi_t^i)^2$. Montrer que (X_t) vérifie l'équation différentielle stochastique:

$$dX_t = (2\theta X_t + n\sigma^2)dt + 2\sigma\sqrt{X_t^+}d\beta_t, X_0 = \sum_{i=1}^n (x_0^i)^2,$$

où (β_t) est un mouvement brownien standard.

On pourra consulter le livre de Karlin et Taylor (1981, chapitre 15) pour plus de détails et d'exemples sur les densités de transition.

6.2.2 Distributions de dimension finie.

Connaissant la loi initiale μ (loi de ξ_0) et les probabilités de transition, il est possible de calculer la loi de tout vecteur $(\xi_{t_0}, \xi_{t_1}, \ldots, \xi_{t_n})$ pour $0 \leq t_0 < t_1 < \ldots < t_n$.

Proposition 6.2.1. *1. Pour $t > 0$ et B borélien de $\mathbb{R}$,*

$$\mathbb{P}(\xi_t \in B) = \int_B dy(\int_{\mathbb{R}} \mu(dx)p_t(x,y)).$$

La loi de ξ_t est la loi, notée $\mu P_t(dy)$, de densité $\int_{\mathbb{R}} \mu(dx)p_t(x,y)$.

2. Pour $0 < t_1 < \ldots < t_n$ et $B_0, B_1, \ldots, B_n$ boréliens,

$$\mathbb{P}(\xi_0 \in B_0, \xi_{t_1} \in B_1, \ldots, \xi_{t_n} \in B_n) =$$

$$\int_{x_0 \in B_0, x_1 \in B_1, \ldots, x_n \in B_n} \mu(dx_0)dx_1 \ldots, dx_n p_{t_1}(x_0, x_1)p_{t_2-t_1}(x_1, x_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, x_n).$$

Preuve. Prenons $t > 0$, B_0, B des boréliens:

$$\mathbb{P}(\xi_0 \in B_0, \xi_t \in B) = \mathbb{E}[1_{\xi_0 \in B_0}\mathbb{E}(1_{\xi_1 \in B}|\mathcal{F}_0)] = \mathbb{E}[1_{\xi_0 \in B_0}P_t(\xi_0, B)] = \int_{B_0} \mu(dx_0) \int_B p_t(x_0, y)dy.$$

A partir de $n > 1$, on applique la propriété de Markov:

$$\begin{aligned}
\mathbb{P}(\xi_0 \in B_0, \xi_{t_1} \in B_1, \ldots, \xi_{t_n} \in B_n) &= \mathbb{E}[1_{(\xi_0 \in B_0, \ldots, \xi_{t_{n-1}} \in B_{n-1})}\mathbb{E}(1_{(\xi_{t_n} \in B_n)}|\mathcal{F}_{n-1})] \\
&= \mathbb{E}[1_{(\xi_0 \in B_0, \ldots, \xi_{t_{n-1}} \in B_{n-1})}P_{t_n - t_{n-1}}(\xi_{t_{n-1}},, B_n)].
\end{aligned}$$

On conclut par récurrence sur n. $\square$

Insistons sur le fait que la propriété de Markov dit que la loi de $(\xi_0^x, \xi_{t_1}^x, \ldots, \xi_{t_n}^x)$, pour $0 < t_1 < \ldots < t_n$, est identique à la loi, sachant $\mathcal{F}_t$ sur $\xi_t = x$, de $(\xi_t, \xi_{t+t_1}, \ldots, \xi_{t+t_n})$. Si $t_i = i\Delta$, et $\mu = \delta_x$, le vecteur $(\xi_{i\Delta}, i = 1, \ldots, n)$ a pour densité sur $\mathbb{R}^n$,

$$\prod_{i=1}^n p_\Delta(x_{i-1}, x_i) \quad \text{avec} \quad x_0 = x.$$

Ceci donne une formule exacte pour la vraisemblance de l'observation $(\xi_{i\Delta}, i = 1, \ldots, n)$. Le problème est qu'en dehors de quelques cas, la formule n'est pas explicite.

6.2.3 Une représentation de la densité de transition d'une diffusion.

Coefficient de diffusion constant.

Considérons l'équation différentielle stochastique unidimensionnelle:

$$d\xi_t = b(\xi_t)dt + dW_t, \quad \xi_0 = x \tag{6.2}$$

où b est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ et $x \in \mathbb{R}$. Nous donnons une représentation et un encadrement de la densité de transition $p_t(x,y)$ qui utilisent les hypothèses suivantes:

- (H1) b est de classe C^1.
- (H2) $|b| \leq M$ et $|b'| \leq k$.

Proposition 6.2.2. *Posons $B(y) = \int_0^y b(u)du$. Sous (H1), on a*

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}} \exp\left(B(y) - B(x) - \frac{(y-x)^2}{2t}\right) E(\exp(-\frac{t}{2}\int_0^1 g((1-u)x + uy + \sqrt{t}B_u^0)du)),$$

où $g = b^2 + b'$ et $(B_u^0, u \in [0,1])$ est un pont brownien standard.

Sous (H1)-(H2),

$$\begin{aligned} \frac{1}{\sqrt{2\pi t}} \exp[-M|y-x| - (M^2+k)(t/2) - \frac{(y-x)^2}{2t}] &\leq p_t(x,y) \\ &\leq \frac{1}{\sqrt{2\pi t}} \exp[M|y-x| + k\frac{t}{2} - \frac{(y-x)^2}{2t}]. \end{aligned}$$

On en déduit :

$$\frac{1}{\sqrt{2\pi t}} \exp[-\frac{5}{2}M^2 t - k\frac{t}{2} - \frac{3(y-x)^2}{4t}] \leq p_t(x,y) \leq \frac{1}{\sqrt{2\pi t}} \exp[(k\frac{t}{2} + 2M^2 t - \frac{(y-x)^2}{4t}].$$

Preuve. Notons P_x^b la loi de $(\xi_s, 0 \leq s \leq t)$, W_x la loi de $(x + W_s, 0 \leq s \leq t)$ et $(X_s, 0 \leq s \leq t)$ le processus des coordonnées canoniques de $C([0,t], \mathbb{R})$. On a:

$$\frac{dP_x^b}{dW_x} = \exp\left(\int_0^t b(X_s)dX_s - \frac{1}{2}\int_0^t b^2(X_s)ds\right) = \exp\left(B(X_t) - B(x) - \frac{1}{2}\int_0^t g(X_s)ds\right).$$

Prenons φ une fonction positive. On obtient:

$$E(\varphi(\xi_t)) = \mathbb{E}_{P_x^b}\varphi(X_t) = E_{W_x}\left(\varphi(X_t)\exp\left(B(X_t) - B(x) - \frac{1}{2}\int_0^t g(X_s)ds\right)\right).$$

Ainsi,

$$\int \varphi(y)p_t(x,y)dy = E_{W_x}\left(\varphi(X_t)\exp\left(B(X_t) - B(x)E_{W_x}(\exp(-\frac{1}{2}\int_0^t g(X_s)ds)|X_t\right)\right).$$

Calculons:

$$E_{W_x}(\exp(-\frac{1}{2}\int_0^t g(X_s)ds)|X_t = y) = E_{W_x}(\exp(-\frac{t}{2}\int_0^1 g(X_{ut})du)|X_t = y).$$

Sous W_x, le processus (X_s) est un mouvement brownien issu de x. De plus, pour $u \in [0,1]$,

$$X_{ut} = x + u(X_t - x) + \sqrt{t}B_u^0,$$

où

$$B_u^0 = \frac{1}{\sqrt{t}}(X_{ut} - x - u(X_t - x)), 0 \leq u \leq 1.$$

Or, sous W_x, $((B^0_u), X_t)$ est un processus gaussien tel que:

$$E(B^0_u) = 0, E(B^0_u B^0_v) = u(1-v), 0 \le u \le v \le 1, E(B^0_u X_t) = 0.$$

Donc, c'est un pont brownien standard indépendant de X_t. Par suite,

$$E_{W_x}(\exp(-\frac{1}{2}\int_0^t g(X_s)ds)|X_t = y) = E_{W_x}(\exp(-\frac{t}{2}\int_0^1 g(x + u(y-x) + \sqrt{t}B^0_u)du)).$$

Ensuite, on écrit: $M|y-x| = 2M\sqrt{t}|y-x|/2\sqrt{t} \le 2M^2 t + (y-x)^2/4t$. D'où les résultats. ■

Coefficient de diffusion non constant.

Il est possible de généraliser la formule de représentation de la desnité de transition au cas où

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dW_t, \quad \xi_0 = x \tag{6.3}$$

où b, σ sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et $x \in \mathbb{R}$ en rajoutant l'hypothèse :

- (H3) σ est de classe C^2, $\sigma(u) \ge \sigma_0 > 0$ pour tout u.

Dans ce cas, en introduisant la fonction $F(.) = \int_0^{\cdot} \frac{1}{\sigma(u)}du$ (strictement croissante), le processus $Y_t = F(\xi_t)$ vérifie l'équation différentielle stochastique:

$$dY_t = \alpha(Y_t)dt + dW_t,$$

avec $\alpha(y) = \frac{b(F^{-1}(y)}{\sigma(F^{-1}(y)} - \frac{1}{2}\sigma'(F^{-1}(y))$. La densité de transition $p_t(x, x')$ de ξ est liée à la densité de transition $q_t(y, y')$ de Y par la relation:

$$p_t(x, x') = q_t(F(x), F(x'))\frac{1}{\sigma(x')}.$$

On rajoute l' hypothèse:

- (H4) $\sigma \le \sigma_1$, σ', σ'' sont bornées.

Dans ce cas, la fonction α est bornée par M_1, la fonction α' est bornée par k_1, et $(y-x)^2/\sigma_1^2 \le (F(y) - F(x))^2 \le (y-x)^2/\sigma_0^2$. Si l'on pose $C_1(t) = \sqrt{2}\exp[(k_1 + 2M_1^2)(t/2)]$, on obtient:

$$\frac{1}{\sigma_1}C_1(t)^{-1}\mathcal{N}(x, t/2)(y) \le p_t(x, y) \le \frac{1}{\sigma_0}C_1(t)\mathcal{N}(x, 2t)(y).$$

6.3 Récurrence sur un intervalle.

Pour ce paragraphe, on pourra consulter par exemple Ikeda et Watanabe (1981), Karlin et Taylor (1981) ou encore Rogers et Williams (1990). On suppose toujours vérifiées les hypothèses (A') ou (B') et, de plus, on suppose qu'il existe un intervalle ouvert (ℓ, r) avec $-\infty \le \ell < r \le +\infty$ tel que:

- $\forall x \in (\ell, r), \ \ \sigma^2(x) > 0,$
- b de classe C^1 sur (ℓ, r),
- σ^2 de classe C^2 sur (ℓ, r).

On rappelle que $(\xi_t^x, t \geq 0)$ désigne le processus solution de (6.1) de condition initiale $\xi_0^x = x$. Pour $x, y \in (\ell, r)$, on définit

$$T_{x,y} = \inf\{t \geq 0, \xi_t^x = y\} \tag{6.4}$$

le premier instant de passage de la trajectoire $(\xi_t^x, t \geq 0)$ en y avec la convention habituelle $\inf \emptyset = +\infty$. Pour a, b tels que $\ell < a \leq x \leq b < r$, on note:

$$T = T_{x,a} \wedge T_{x,b} = \inf\{t \geq 0, \xi_t^x = a \text{ ou } b\} \tag{6.5}$$

le temps de sortie de (a, b) de la trajectoire démarrant de $x \in [a, b]$. Lorsque $b \uparrow r$, $T_{x,b} \uparrow$ et converge vers une limite que nous noterons T_{x,r^-}. De même, $T_{x,\ell^+} = \lim_{a \downarrow \ell} \downarrow T_{x,a}$. Enfin, on pose

$$e_x = T_{x,r^-} \wedge T_{x,\ell^+}. \tag{6.6}$$

Nous allons, dans ce paragraphe, définir un critère assurant que

$$\forall x \in (\ell, r), \ \ \mathbb{P}(e_x = +\infty) = 1 = \mathbb{P}(\limsup_{t \to +\infty} \xi_t^x = r) = \mathbb{P}(\liminf_{t \to +\infty} \xi_t^x = \ell). \tag{6.7}$$

Autrement dit, d'une part, la trajectoire issue de $x \in (\ell, r)$ reste indéfiniment dans cet intervalle. De plus, elle oscille indéfiniment entre les deux frontières ℓ et r. Sur $(e_x = +\infty)$, $\forall t \geq 0$, $\sigma^2(\xi_t^x) > 0$. Nous déduirons en particulier de (6.7) que

$$\mathbb{P}(\forall t \geq 0, \ \sigma^2(\xi_t^x) > 0) = 1.$$

Remarque 6.3.1. *La première égalité de (6.7) est immédiate lorsque $(\ell, r) = \mathbb{R}$. En effet, si $T_{x,r^-} < +\infty$, la continuité des trajectoires implique que, pour tout $x < b < r$, $T_{x,b} < T_{x,r^-} < +\infty$ et $\xi_{T_{x,b}} = b \to \xi_{T_{x,r^-}} = r$ lorsque $b \uparrow r$. Ceci n'est possible que si $r \in \mathbb{R}$. Par conséquent, si $r = +\infty$, $T_{x,r^-} = +\infty$ p.s.. De même, si $\ell = -\infty$, $T_{x,\ell^+} = +\infty$ p.s.. Ainsi, si $(\ell, r) = \mathbb{R}$, $e_x = +\infty$ p.s. ce que nous savons déjà puisque $\mathbb{P}(\xi_t^x \in \mathbb{R}, \forall t \geq 0) = 1$.*

La deuxième partie des égalités (6.7) exprime la propriété de récurrence sur (ℓ, r).

6.3.1 Probabilités de sortie d'intervalles bornés.

Nous allons montrer que T défini en (6.5) est tel que $\mathbb{P}(T < +\infty) = 1$ et calculer les probabilités de sortie de l'intervalle (a, b):

$$\mathbb{P}(\xi_T^x = a) = \mathbb{P}(T_{x,a} < T_{x,b}), \qquad \mathbb{P}(\xi_T^x = b) = \mathbb{P}(T_{x,b} < T_{x,a}),$$

ainsi que $\mathbb{E}(T)$ et plus généralement $\mathbb{E}(\int_0^T g(\xi_s^x) ds$ pour toute fonction $g : [a, b] \to \mathbb{R}$ continue et positive. Cette étude fait intervenir l'opérateur L défini sur l'ensemble des fonctions f de classe C^2 sur $\mathbb{R}$ par:

$$Lf(x) = \frac{\sigma^2(x)}{2} f''(x) + b(x) f'(x). \tag{6.8}$$

Proposition 6.3.1. • *Pour toute fonction $f \in C^2(\mathbb{R})$,*

$$(M_t^{f,x} = f(\xi_t^x) - f(x) - \int_0^t Lf(\xi_s^x)ds, t \geq 0)$$

est une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$.

- *Pour toute fonction $f \in C^2((\ell, r))$, $(M_{t\wedge T}^{f,x}, t \geq 0)$ est une vraie martingale de carré intégrable et*

$$\mathbb{E}f(\xi_{t\wedge T}^x) = f(x) + \mathbb{E}(\int_0^{t\wedge T} Lf(\xi_s^x)ds).$$

Preuve. En appliquant la formule d'Itô, on obtient que

$$M_t^{f,x} = \int_0^t f'(\xi_s^x)\sigma(\xi_s^x)dB_s$$

est une intégrale stochastique donc une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$.
D'après le théorème d'arrêt, $(M_{t\wedge e_x}^{f,x})$ est encore une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ et ceci est valable pour toute fonction $f \in C^2((\ell, r))$. Par suite, $(M_{t\wedge e_x\wedge T}^{f,x})$ est encore une martingale locale. Comme $T < e_x$, $M_{t\wedge e_x\wedge T}^{f,x} = M_{t\wedge T}^{f,x}$ et $(M_{t\wedge T}^{f,x})$ est une martingale locale.
Par définition de T, pour tout $t \leq T$, $a \leq \xi_t^x \leq b$. Ceci implique:

$$\int_0^{t\wedge T} \left(f'(\xi_s^x)\sigma(\xi_s^x)\right)^2 ds \leq t \sup_{a\leq u\leq b} (f'(u)\sigma(u))^2.$$

On en déduit que:

$$\mathbb{E}(M_{t\wedge T}^{f,x})^2 = \mathbb{E} < M^{f,x} >_{t\wedge T} = \mathbb{E}\int_0^{t\wedge T} \left(f'(\xi_s^x)\sigma(\xi_s^x)\right)^2 ds < +\infty.$$

Ainsi, la martingale locale $(M_{t\wedge T}^{f,x})$ est une vraie martingale, de carré intégrable, nulle en 0, ce qui implique, pour tout $t \geq 0$,

$$\mathbb{E}(M_{t\wedge T}^{f,x}) = \mathbb{E}f(\xi_{t\wedge T}^x) - f(x) - \mathbb{E}(\int_0^{t\wedge T} Lf(\xi_s^x)ds) = 0.$$

■

On en déduit les probabilités de sortie de l'intervalle (a, b).

Théorème 6.3.1. *On a $\mathbb{P}(T < +\infty) = 1$ et*

$$\begin{aligned} \mathbb{P}(\xi_T^x = a) = \mathbb{P}(T_{x,a} < T_{x,b}) &= \frac{S(b) - S(x)}{S(b) - S(a)} \quad \textit{(sortie par a)} \\ \mathbb{P}(\xi_T^x = b) = \mathbb{P}(T_{x,b} < T_{x,a}) &= \frac{S(x) - S(a)}{S(b) - S(a)} \quad \textit{(sortie par b)}, \end{aligned}$$

où S est définie comme toute solution de l'équation différentielle ordinaire $LS = 0$.

Avant de prouver le théorème, résolvons l'équation $LS = 0$ sur (ℓ, r). On a:

$$\frac{S''}{S'} = -2\frac{b}{\sigma^2}, \text{ d'où } s(x) = S'(x) = \exp\left(-2\int_{x_0}^{x} \frac{b(u)}{\sigma^2(u)} du\right),$$

$x_0 \in (\ell, r)$, et $S(x) = \int_{x_1}^{x} s(t)dt$, $x_1 \in (\ell, r)$. La fonction S appelée *fonction d'échelle* dépend de deux constantes d'intégration (x_0, x_1) mais les probabilités de sortie données par le théorème n'en dépendent pas. Si $S(x) = x$ $(b = 0)$, le processus est *dans son échelle naturelle*: c'est une martingale locale et il a les mêmes probabilités de sortie d'intervalle qu'un mouvement brownien standard. On retiendra surtout la dérivée de la fonction d'échelle:

$$s(x) = \exp\left(-2\int_{x_0}^{x} \frac{b(u)}{\sigma^2(u)} du\right). \tag{6.9}$$

Preuve du théorème. D'après la proposition précédente,

$$S(\xi_{t\wedge T}^x) - S(x) = \int_0^{t\wedge T} s(\xi_s^x)\sigma(\xi_s^x)dB_s := M_t^T$$

est une martingale locale uniformément bornée puisque:

$$\sup_{t\geq 0} |S(\xi_{t\wedge T}^x) - S(x)| \leq 2 \sup_{a\leq u\leq b} |S(u)| := C.$$

Par conséquent, cette martingale converge p.s. et dans L^2 vers une v.a. M_∞^T qui vérifie:

$$c^2\mathbb{E}(T) \leq \mathbb{E}(M_\infty^T)^2 = \mathbb{E}\left(\int_0^T s^2(\xi_s^x)\sigma^2(\xi_s^x)ds\right) \leq C^2,$$

où $c^2 = \inf_{a\leq u\leq b} s^2(u)\sigma^2(u)$, ce qui implique $\mathbb{E}(T) < +\infty$. La v.a. ξ_T^x est donc bien définie et ne prend que les deux valeurs a et b. De plus, $\mathbb{E}(M_\infty^T) = \mathbb{E}(M_0^T) = 0 = \mathbb{E}S(\xi_T^x) - S(x)$. Par suite,

$$S(a)\mathbb{P}(\xi_T^x = a) + S(b)\mathbb{P}(\xi_T^x = b) = S(x).$$

En rajoutant l'équation

$$\mathbb{P}(\xi_T^x = a) + \mathbb{P}(\xi_T^x = b) = 1,$$

on obtient le résultat ∎

6.3.2 Récurrence et temps d'atteinte de niveaux.

Définition 6.3.1. *La famille de processus $(\xi_t^x, t \geq 0), x \in (\ell, r)$ est dite récurrente sur (ℓ, r) si*

$$\forall x, y \in (\ell, r), \quad \mathbb{P}(T_{x,y} < +\infty) = 1.$$

Théorème 6.3.2. *(i) Si $S(\ell) = \lim_{x\downarrow\ell} S(x) = -\infty$ et $S(r) = \lim_{x\uparrow r} S(x) = +\infty$, alors*

$$\forall x, y \in (\ell, r), \quad \mathbb{P}(T_{x,y} < +\infty) = 1.$$

(ii) De plus, $\mathbb{P}(e_x = +\infty) = 1 = \mathbb{P}(\limsup_{t\to+\infty} \xi_t^x = r) = \mathbb{P}(\liminf_{t\to+\infty} \xi_t^x = \ell)$.

Preuve. Nous démontrons uniquement (i). Choisissons d'abord $\ell < x < y < r$ et soit a tel que $\ell < a < x < y < r$. Posons $T = T_{x,a} \wedge T_{x,y}$. Si $a \downarrow \ell$, puisque $S(\ell) = -\infty$,

$$\mathbb{P}(T_{x,y} < T_{x,a}) = \frac{S(x) - S(a)}{S(y) - S(a)} \to 1.$$

Comme $\mathbb{P}(T_{x,y} < +\infty) \geq \mathbb{P}(T_{x,y} < T_{x,a})$, on en déduit l'implication:

$$S(\ell) = -\infty \Rightarrow \forall x < y, \mathbb{P}(T_{x,y} < +\infty) = 1.$$

De façon analogue,

$$S(r) = +\infty \Rightarrow \forall x > y, \mathbb{P}(T_{x,y} < +\infty) = 1.$$

■

Remarque 6.3.2. • *Il est plus commode d'exprimer les conditions* $S(\ell) = -\infty, S(r) = +\infty$ *comme la divergence de l'intégrale de la fonction* $s = S'$ *aux deux bornes* ℓ, r*:*

$$\int_{\ell} s(x)dx = +\infty, \qquad \int^{r} s(x)dx = +\infty. \tag{6.10}$$

- *Lorsque* $S(\ell) = -\infty, S(r) = +\infty$*, la fonction d'échelle* $S : (\ell, r) \to \mathbb{R}$ *est strictement croissante et* $S^{-1} : \mathbb{R} \to (\ell, r)$ *est bien définie. Le processus* $Y_t = S(\xi_t^x)$ *vérifie*

$$dY_t = S' \circ S^{-1}(Y_t)\sigma \circ S^{-1}(Y_t)dB_t.$$

Il est dans son échelle naturelle et $\limsup_{t\to+\infty} Y_t = +\infty, \liminf_{t\to+\infty} Y_t = -\infty$.

6.3.3 Temps de sortie d'intervalles bornés.

Pour $\ell < a \leq x \leq b < r$, nous étudions dans ce paragraphe $\mathbb{E}\, T$ pour $T = T_{x,a} \wedge T_{x,b}$ et plus genéralement $\mathbb{E} \int_0^T g(\xi_s^x)ds$ pour $g : [a, b] \to \mathbb{R}^+$ continue et positive.

Théorème 6.3.3. *(i) Soit* u *la solution sur* (ℓ, r) *de l'équation* $Lu = -1$ *avec conditions aux bords* $u(a) = u(b) = 0$*. On a, pour tout* $x \in [a, b]$*,*

$$\mathbb{E}\, T = u(x).$$

(ii) Soit v *la solution sur* (ℓ, r) *de l'équation* $Lv = -g$ *avec conditions aux bords* $v(a) = v(b) = 0$ *et* $g : (\ell, r) \to \mathbb{R}^+$ *continue et positive ou nulle sur* $[a, b]$*. On a, pour tout* $x \in [a, b]$*,*

$$\mathbb{E}(\int_0^T g(\xi_s^x)ds) = v(x).$$

(iii) La résolution de $Lv = -g$, $v(a) = v(b) = 0$ conduit à:

$$\begin{aligned} v(x) &= 2\frac{S(x)-S(a)}{S(b)-S(a)} \int_x^b (S(b)-S(u))g(u)m(u)du \\ &+ 2\frac{S(b)-S(x)}{S(b)-S(a)} \int_a^x (S(u)-S(a))g(u)m(u)du \end{aligned}$$

où

$$m(u) = \frac{1}{\sigma^2(u)s(u)}. \tag{6.11}$$

La fonction v ne dépend pas des constantes x_0, x_1 choisies pour calculer s, S. Si $g(x) \geq 0, x \in [a,b]$, alors $v(x) > 0$, $x \in (a,b)$.

(iv) Pour tout $f \in C^2((\ell, r))$,

$$Lf = \frac{1}{2m}\left(\frac{f'}{s}\right)'.$$

La mesure de densité m sur (ℓ, r) s'appelle *la mesure de vitesse* de la famille de processus $(\xi_t^x, t \geq 0), x \in (\ell, r)$. En utilisant *la fonction de Green*

$$G(x,u) = \frac{S(x)-S(a)}{S(b)-S(a)}(S(b)-S(u))1_{(a\leq x\leq u\leq b)} + \frac{S(b)-S(x)}{S(b)-S(a)}(S(u)-S(a))1_{(a\leq u\leq x\leq b)},$$

$v(x) = 2\int G(x,u)g(u)m(u)du$.

Preuve du théorème. D'après la proposition 6.3.1, si $Lu = -1$,

$$u(\xi_{t\wedge T}^x) - u(x) + t\wedge T = \int_0^{t\wedge T} u'(\xi_s^x)\sigma(\xi_s^x)dB_s$$

est une vraie martingale. Par suite,

$$\mathbb{E}(t\wedge T) = u(x) - \mathbb{E}(u(\xi_{t\wedge T}^x)).$$

On a vu que $T < +\infty$. D'une part, $\mathbb{E}(t\wedge T) \nearrow \mathbb{E}(T)$ lorsque t tend vers l'infini (théorème de convergence monotone). D'autre part, $\xi_{t\wedge T}^x \to \xi_T^x$. Par continuité de u, $u(\xi_{t\wedge T}^x) \to u(\xi_T^x) = 0$ puisque $u(a) = u(b) = 0$. On a

$$\sup_{t\geq 0} |u(\xi_{t\wedge T}^x)| \leq \sup_{y\in[a,b]} |u(y)|.$$

Le théorème de convergence dominée de Lebesgue permet de conclure que $\mathbb{E}(u(\xi_{t\wedge T}^x)) \to 0$ ce qui permet d'obtenir (i).

Pour (ii), on procède de manière analogue en utilisant la martingale:

$$v(\xi_{t\wedge T}^x) - v(x) + \int_0^{t\wedge T} g(\xi_s^x)ds = \int_0^{t\wedge T} v'(\xi_s^x)\sigma(\xi_s^x)dB_s.$$

L'égalité (iv) est obtenue par un calcul élémentaire. Pour (iii), résolvons $Lu = -1$ avec les conditions aux bords $u(a) = u(b) = -1$. Une première intégration conduit à:

$$u'(x) = -2s(x)\int_a^x m(t)dt + \beta s(x)$$

où $\beta \in \mathbb{R}$ est une constante d'intégration. Puis,

$$u(x) = -2\int_a^x s(t)dt \int_a^t m(u)du + \beta(S(x) - S(a)) + \alpha$$

où $\alpha \in \mathbb{R}$ est une seconde constante d'intégration. Avec $u(a) = 0$, on obtient $\alpha = 0$. Avec $u(b) = 0$, on obtient:

$$\beta = \frac{2}{S(b) - S(a)}\int_a^b s(t)dt \int_a^t m(u)du.$$

D'où,

$$u(x) = \frac{2(S(x) - S(a))}{S(b) - S(a)}\int_a^b s(t)dt \int_a^t m(u)du - 2\int_a^x s(t)dt \int_a^t m(u)du.$$

On intervertit les intégrations:

$$u(x) = \frac{2(S(x) - S(a))}{S(b) - S(a)}\int_a^b (S(b) - S(u))m(u)du - 2\int_a^x (S(x) - S(u))m(u)du.$$

On découpe $\int_a^b \ldots = \int_a^x \ldots + \int_x^b \ldots$. Il vient:

$$\begin{aligned} u(x) &= \frac{2(S(x) - S(a))}{S(b) - S(a)}\int_b^x (S(b) - S(u))m(u)du \\ &+ \frac{2(S(x) - S(a))}{S(b) - S(a)}\int_a^x (S(b) - S(u))m(u)du - 2\int_a^x (S(x) - S(a) + S(a) - S(u))m(u)du. \end{aligned}$$

On a

$$\begin{aligned} (S(x) - S(a))\left(\frac{S(b) - S(u)}{S(b) - S(a)} - 1\right) &- (S(a) - S(u)) \\ = (&S(x) - S(a))\frac{S(a) - S(u)}{S(b) - S(a)} - (S(a) - S(u)) \\ = &(S(a) - S(u))\frac{S(b) - S(x)}{S(b) - S(a)}. \end{aligned}$$

Ce qui donne le résultat. La résolution de $Lv = -g$ est analogue. ∎

6.3.4 Autres moments

Il est possible de calculer tous les moments de $T = T_{x,a} \wedge T_{x,b}$, de $\int_0^T g(\xi_s^x)ds$.

Soit $g : [a,b] \to \mathbb{R}^+$, continue. On pose $v(x) = v_1(x)$ (cf Théorème 6.3.3) et par récurrence, soit v_n la fonction solution de

$$Lv_n = -n\, v_{n-1}\, g, \quad v_n(a) = v_n(b) = 0. \tag{6.12}$$

Alors ,

$$\mathbb{E}\left(\int_0^T g(\xi_s^x)ds\right)^n = v_n(x).$$

Nous ébauchons la preuve pour $g = 1$ et renvoyons à Fitzsimmons et Pitman (1999) pour plus de détails.
Posons $u(x) = u_1(x)$.La fonction u_1 est définie sur $[a,b]$ continue et positive. La fonction u_2 se calcule par la formule du Théorème 6.3.3 en remplaçant g par $2u_1$. De même, u_2 est est définie sur $[a,b]$ continue et positive. Pour prouver que $u_2(x) = \mathbb{E}T^2$, on utilise une martingale adéquate. Par la formule d'Ito,

$$tu_1(\xi_t^x) = \int_0^t sLu_1(\xi_s^x)ds + \int_0^t u_1(\xi_s^x)ds + \int_0^t su_1'(\xi_s^x)\sigma(\xi_s^x)dB_s.$$

On applique cette formule en $t \wedge T$ et on utilise le fait que $Lu_1 = -1$:

$$(t\wedge T)u_1(\xi_{t\wedge T}^x) + (t\wedge T)^2/2 - \int_0^{t\wedge T} u_1(\xi_s^x)ds = \int_0^{t\wedge T} su_1'(\xi_s^x)\sigma(\xi_s^x)dB_s.$$

Le membre de droite est une vraie martingale. Donc,

$$\mathbb{E}(t\wedge T)^2 = -\mathbb{E}2\ (t\wedge T)u_1(\xi_{t\wedge T}^x) + \mathbb{E}\int_0^{t\wedge T} 2\ u_1(\xi_s^x)ds.$$

D'après le théorème de convergence monotone, $\mathbb{E}(t\wedge T)^2 \uparrow \mathbb{E}T^2$, $\mathbb{E}\int_0^{t\wedge T} 2\ \ u_1(\xi_s^x)ds \uparrow \mathbb{E}\int_0^T 2\ u_1(\xi_s^x)ds$. D'autre part, $(t\wedge T)u_1(\xi_{t\wedge T}^x) \to Tu_1(\xi_T^x) = 0$. Comme

$$|(t\wedge T)u_1(\xi_{t\wedge T}^x| \leq T \sup_{a\leq y\leq b}|u(y)|$$

avec $\mathbb{E}T < +\infty$, on en déduit $\mathbb{E}(t\wedge T\ u_1(\xi_{t\wedge T}^x)) \to 0$ et la relation

$$\mathbb{E}T^2 = \mathbb{E}\int_0^T 2\ u_1(\xi_s^x)ds = u_2(x).$$

Enfin, le membre de droite est solution de l'équation $Lu_2 = -2\ u_1, u_2(a) = u_2(b) = 0$.

Pour calculer, $\mathbb{E}T^3$, on part de $tu_2(\xi_t^x)$ et on utilise les relations $Lu_2 = -2u_1, L(Lu_2) = -2Lu_1 = 2 \ldots$. Et ainsi de suite, pour tous les moments.

On peut aussi calculer la transformée de Laplace de $\int_0^T g(\xi_s^x)ds$ pour $g \geq 0$. Soit f la solution de

$$Lf = -g\ f, \quad f(a) = f(b) = 1.$$

Alors,

$$\mathbb{E}\exp\left(-\int_0^T g(\xi_s^x)ds\right) = f(x).$$

(Ecrire les conditions pour que $M_t = \exp\left(-\int_0^t g(\xi_s^x)ds\right)f(\xi_t^x))$ soit une martingale locale).

L'équation $Lf = -g\ f$ est une équation différentielle du second ordre. Si ses coefficients sont constants (cas du mouvement brownien avec drift), on la résout aisément. Sinon, il n'y a pas de solution explicite simple. Dans certains cas, (dérive linéaire, coefficient de diffusion $\sigma(x) = 1, \sqrt{x}, \sqrt{x-1)(1-x)}$), on obtient des solutions qui sont des *fonctions spéciales* (cf par exemple Abramowitz et Stegun, 1964).

6.3.5 Récurrence positive.

Définition 6.3.2. *La famille de processus $(\xi_t^x, t \geq 0), x \in (\ell, r)$ est dite récurrente positive sur (ℓ, r) si*

$$\forall x, y \in (\ell, r), \quad \mathbb{E}(T_{x,y}) < +\infty.$$

Théorème 6.3.4. *Si $\int_\ell s(x)dx = +\infty = \int^r s(x)dx$ et si, de plus, $\int_\ell^r m(x)dx < +\infty$, alors, $\forall x, y \in (\ell, r), \quad \mathbb{E}(T_{x,y}) < +\infty$. Dans ce cas, on a, si $x < y$,*

$$\mathbb{E}(T_{x,y}) = 2\int_x^y (S(y) - S(u))m(u)du + 2(S(y) - S(x))\int_\ell^x m(u)du.$$

Si $x > y$,

$$\mathbb{E}(T_{x,y}) = 2(S(y) - S(x))\int_x^r m(u)du + 2\int_x^y (S(y) - S(u))m(u)du.$$

Preuve. Si $x < y$, considérer $\ell < a < x < y$. Alors, $T_{x,a} \to +\infty$ p.s. lorsque $a \downarrow \ell$ et $T_{x,a} \wedge T_{x,y} \uparrow T_{x,y}$. On passe à la limite grâce au théorème de convergence monotone (voir théorème 6.3.3). Raisonnement analogue pour $x > y$. ■

Remarque 6.3.3. *On peut calculer la transformée de Laplace de $T_{x,y}$. Soit f une solution positive et strictement croissante de l'équation différentielle ordinaire $Lf = \lambda\ f$, alors, pour $x < y$, et pour tout $\lambda > 0$,*

$$\mathbb{E}(\exp(-\lambda T_{x,y})) = \frac{f(x)}{f(y)}.$$

Pour cela, on cherche la condition sur f pour que $f(\xi_t^x)\exp(-\lambda t)$ soit une martingale locale:

$$f(\xi_t^x)\exp(-\lambda t) = f(x) + \int_0^t \exp(-\lambda s)\left(Lf(\xi_s^x) - \lambda f(\xi_s^x)\right)ds + \int_0^t f'(\xi_s^x))\sigma(\xi_s^x))dB_s.$$

Si f est telle que $Lf = \lambda\ f$, on obtient que $f(\xi_t^x)\exp(-\lambda t) = f(x) + \int_0^t f'(\xi_s^x))\sigma(\xi_s^x))dB_s$ est une martingale locale ainsi que $f(\xi_{t\wedge T_{x,y}}^x)\exp(-\lambda(t \wedge T_{x,y}))$. Comme f est positive et croissante,

$$0 < f(\xi_{t\wedge T_{x,y}}^x)\exp(-\lambda(t \wedge T_{x,y})) \leq f(y).$$

On obtient que $f(\xi^x_{t\wedge T_{x,y}})\exp(-\lambda t\, T_{x,y})$ *est une martingale locale uniformément bornée donc une vraie martingale. Il en résulte:*

$$\mathbb{E} f(\xi^x_{t\wedge T_{x,y}})\exp(-\lambda(t\wedge T_{x,y})) = f(x).$$

En appliquant le théorème de convergence dominée, pour t *tendant vers l'infini, on obtient:*

$$\mathbb{E} f(y)\exp(-\lambda T_{x,y}) = f(x).$$

C'est le résultat.
Pour $b(x)=\theta\geq 0, \sigma(x)=\sigma$ *(mouvement brownien avec dérive positive ou nulle), on trouve*

$$\mathbb{E}\exp(-\lambda T_{x,y}) = \exp[-\frac{(y-x)}{\sigma^2}\left(\sqrt{\theta^2+2\lambda\sigma^2}-\theta\right)].$$

6.3.6 Théorèmes limites.

Dans ce paragraphe, nous supposons vérifées les conditions de récurrence positive sur (ℓ, r): $s(x)=\exp(-2\int_{\cdot}^{x} b(u)/\sigma^2(u)du)$ est la dérivée de la fonction d'échelle, $m(x)=1/(\sigma^2(x)s(x))$ est la densité de la mesure de vitesse,

$$\int_{\ell} s(x)dx = +\infty = \int^{r} s(x)dx, \quad M=\int_{\ell}^{r} m(x)dx<+\infty. \tag{6.13}$$

La fonction

$$\pi(x)=\frac{m(x)}{M}1_{(\ell,r)} \tag{6.14}$$

est une densité de probabilité et l'on a le théorème:

Théorème 6.3.5. *Soit* (ξ_t) *le processus solution de l'équation différentielle stochastique (6.1).*

- *Soit* $f:(\ell,r)\to\mathbb{R}$ *borélienne telle que* $\int|f(x)|\pi(x)dx<+\infty$*, pour toute loi initiale* μ *(loi de* ξ_0*),*
$$\frac{1}{T}\int_0^T f(\xi_s)ds \to_{T\to+\infty} \int f(x)\pi(x)dx, \qquad p.s..$$
- *Pour toute loi initiale* μ*,* ξ_t *converge en loi vers* $\pi(x)dx$ *quand* t *tend vers* $+\infty$*.*

Pour la preuve, nous renvoyons à Rogers et Williams (1987, p.300 et s.). Le premier point est le *théorème ergodique.*

6.3.7 Distributions invariantes

Définition 6.3.3. *La distribution* μ *est dite stationnaire ou invariante si* $\xi_0\sim\mu$ *implique que, pour tout* $t\geq 0$*,* $\xi_t\sim\mu$*. C'est-à-dire*

$$\mu \;\; invariante \;\; \Leftrightarrow \forall t\geq 0,\;\; \mu P_t=\mu.$$

En vertu de la propriété de Markov, si μ est une distribution invariante et si $\xi_0 \sim \mu$, alors le processus (ξ_t) est stationnaire strict. On peut démontrer que, sous les conditions de récurrence positive, le processus (6.1) admet une unique probabilité invariante égale à $\pi(x)dx$. Nous commençons par la condition nécessaire.

Proposition 6.3.2. • *Soit $h : (\ell, r) \to \mathbb{R}$ de classe C^2 vérifiant $h \geq 0$ et $\int_\ell^r h(x)dx = 1$. Si $\mu = h(x)dx$ est une distribution invariante, alors h est solution de l'équation différentielle ordinaire*

$$(h\frac{\sigma^2}{2})'' - (hb)' = 0. \tag{6.15}$$

- *Si $\int_\ell s(x)dx = +\infty = \int^r s(x)dx$ (conditions de récurrence), la seule solution positive de (6.15) est $h(x) = Km(x)$ où $m(x)$ est la densité de la mesure de vitesse.*

Preuve. Soit $f : (\ell, r) \to \mathbb{R}$ une fonction de classe C^2 à support compact inclus dans (ℓ, r). Supposons que $\mu = h(x)dx$ soit une distribution invariante avec $h(x) = 0$ pour $x \notin (\ell, r)$ et que $\xi_0 \sim \mu$. Dans ce cas, pour tout $t \geq 0$,

$$\mathbb{E}f(\xi_0) = \int f d\mu = \mathbb{E}f(\xi_t).$$

D'autre part,

$$f(\xi_t) = f(\xi_0) + \int_0^t Lf(\xi_s)ds + \int_0^t f'(\xi_s)\sigma(\xi_s)dB_s.$$

Comme f est C^2 avec ună support compact inclus dans (ℓ, r),

$$\mathbb{E}\int_0^t (f'(\xi_s)\sigma(\xi_s))^2 ds = \int_0^t ds \int (f'(x)\sigma(x))^2 d\mu(x) = t\int (f'(x)\sigma(x))^2 d\mu(x) < +\infty$$

ce qui implique $\mathbb{E}\int_0^t f'(\xi_s)\sigma(\xi_s)dB_s = 0$. Ainsi,

$$\mathbb{E}f(\xi_t) = \mathbb{E}f(\xi_0) + \mathbb{E}\int_0^t Lf(\xi_s)ds.$$

Donc,

$$\mathbb{E}\int_0^t Lf(\xi_s)ds = t\int Lf(x)d\mu(x) = 0.$$

Après intégrations par parties:

$$\int Lf(x)h(x)dx = \int f\left((h\frac{\sigma^2}{2})'' - (hb)'\right)dx = 0.$$

Comme cette relation est vraie pour toute fonction f de classe C^2 à support compact dans (ℓ, r), on en déduit que $(h\frac{\sigma^2}{2})'' - (hb)' = 0$ p.p. donc partout puisque qu'on a supposé h, σ^2 de classe C^2 et b de classe C^1.

Nous allons écrire différemment l'équation différentielle ordinaire de h en utilisant la formule $Lf = (2m)^{-1}(f'/s)'$. Les intégrations par parties s'écrivent, pour f à support compact:

$$\int Lf\ h\ dx = \frac{1}{2}\int \left(\frac{f'}{s}\right)' \frac{h}{m} dx = -\frac{1}{2}\int \left(\frac{h}{m}\right)' \frac{f'}{s} dx = \frac{1}{2}\int f\left(\frac{1}{s}\left(\frac{h}{m}\right)'\right)' dx.$$

Ainsi, l'opérateur adjoint de L dans $\mathbb{L}^2$ s'écrit:

$$L^* h = (h\frac{\sigma^2}{2})'' - (hb)' = \frac{1}{2}\left(\frac{1}{s}\left(\frac{h}{m}\right)'\right)'$$

Résolvons

$$\left(\frac{1}{s}\left(\frac{h}{m}\right)'\right)' = 0.$$

On obtient $h/m = C_1\ S + C_2$ où S est la fonction déchelle. Pour assurer $h/m \geq 0$ sachant que $S(r) = +\infty, S(\ell) = -\infty$, il est nécessaire que $C_1 = 0$. Donc, $h(x) = C_2 m(x)$ est proportionnelle àă la densité de la mesure de vitesse. En normalisant, $h(x) = m(x)/M = \pi$ sur (ℓ, r). ∎

Réciproquement, il est possible de démontrer que π est l'unique distribution invariante. Nous admettrons ce résultat.

Remarque 6.3.4. *Notons que l'opérateur L est auto-adjoint relativement à $L^2(m)$: prenons f, g à support compact dans (ℓ, r). On a:*

$$\int g\ Lf\ m\ dx = \frac{1}{2}\int g\ \left(\frac{f'}{s}\right)' dx = -\frac{1}{2}\int g'\ \frac{f'}{s} dx = \frac{1}{2}\int f \frac{1}{m}\left(\frac{g'}{s}\right)'\ m\ dx = \int f\ Lg\ m\ dx.$$

6.4 Exercices

Exercices. Déterminer l'intervalle (ℓ, r) et les conditions de récurrence et de récurrence positive pour les modèles suivants:

$$b(x) = \alpha(\beta - x), \sigma(x) = \sigma, \sigma\sqrt{x^+}, \sigma x, \sigma\sqrt{1+x^2}.$$

En déduire s'il y a lieu, la distribution invariante.

6.5 Compléments.

6.5.1 Densité de la moyenne. Coefficient de diffusion constant.

Condition initiale déterministe.

Nous reprenons le processus (ξ_t) défini par (6.2) avec les hypothèses (H1)-(H2) et nous intéressons à la loi de la moyenne $\bar{\xi}_x = \frac{1}{p}\sum_{j=1}^{p} \xi_{j\delta}$ pour δ un pas de discrétisation. Notons d'abord que, puisque les transitions admettent une densité, la loi $\bar{\xi}_x$ admet une densité que nous noterons $\bar{\pi}_x(.)$. Pour étudier cette densité, on introduit d'abord le processus transformé suivant :

$$x_u = \frac{1}{\sqrt{\Delta}}(\xi_{u\Delta} - x), \quad \text{où} \quad \Delta = p\delta. \tag{6.16}$$

Le processus $(x_u, u \in [0,1])$ est solution de:

$$dx_u = \tilde{b}(x_u)du + d\tilde{W}_u, \quad x_0 = 0, \tag{6.17}$$

avec $\tilde{b}(t) = \sqrt{\Delta} b(t\sqrt{\Delta} + x)$, $\tilde{W}_u = W_{u\Delta}/\sqrt{\Delta}$. De plus,

$$(\bar{\xi}_x, \xi_\Delta) = (x + \sqrt{\Delta}U, x + \sqrt{\Delta}V), \quad U = \frac{1}{p}\sum_{j=1}^{p} x_{j/p}, \quad V = x_1. \tag{6.18}$$

Dans un premier temps, on étudie la loi jointe de (U, V) dont on va chercher un encadrement. Cette loi admet une densité puisque (x_u) admet des densités de transition.

Proposition 6.5.1. *Soit $p_x(u, v)$ la densité de la loi du couple (U, V) défini en (6.18). Il existe des constantes $c_1 > 0, c_2 > 0$ qui ne dépendent que des bornes de b, b' (et pas de x)telles que, pour p assez grand,*

$$c_1^{-1} \exp -c_1(u^2 + v^2) \le p_x(u, v) \le c_2^{-1} \exp -c_2(u^2 + v^2)$$

Preuve. Notons P la loi de $(x_u, u \in [0,1])$ et W la mesure de Wiener sur $[0,1]$. Soient h_0, h_1 deux fonctions positives. On a, avec la formule de Girsanov, comme plus haut (en conservant la notation x_u pour le processus canonique):

$$\mathbb{E}_P h_0(U)h_1(V) = \mathbb{E}_W h_0(U)h_1(V)L_1, \tag{6.19}$$

où

$$L_1 = \exp\left(\int_0^1 \tilde{b}(x_u)dx_u - \frac{1}{2}\int_0^1 \tilde{b}^2(x_u)du\right) = \exp(\tilde{B}(V) - \frac{1}{2}\int_0^1 \tilde{g}(x_u)du),$$

avec $\tilde{B}(v) = \int_0^v \tilde{b}(u)du$ et $\tilde{g} = \tilde{b}^2 + \tilde{b}'$. Les hypothèses sur b, b' impliquent

$$|\tilde{g}| \le C, \quad |\tilde{B}(v)| \le C\sqrt{\Delta}|v|.$$

Par suite,

$$C_1^{-1} \exp(-C\sqrt{\Delta}V) \le L_1 \le C_1 \exp(C\sqrt{\Delta}V).$$

D'où

$$C_1^{-1}\mathbb{E}_W h_0(U)h_1(V)\exp(-C\sqrt{\Delta}V) \le \mathbb{E}_P h_0(U)h_1(V) \le C_1\mathbb{E}_W h_0(U)h_1(V)\exp(C\sqrt{\Delta}V). \tag{6.20}$$

Ainsi, pour encadrer la densité jointe de (U, V) sous P, il suffit d'encadrer la densité jointe de (U, V) sous W. Or cette loi est explicite.

Lemme 6.5.1. *La loi de (U,V) sous W est la loi gaussienne de densité*

$$f(u,v)=\frac{1}{2\pi d_p^{1/2}}\exp[-\frac{1}{2d_p}(u^2+a_pv^2-2b_puv)],$$

avec $d_p=\frac{1}{12}+O(\frac{1}{p}), a_p=\frac{1}{3}+O(\frac{1}{p}), b_p=\frac{1}{2}+O(\frac{1}{p})$.

Preuve. Sous W, le couple (U,V) est gaussien centré. Il suffit de calculer sa matrice de covariance K. On a $\mathrm{Var}(V)=1$, $\mathrm{Cov}(U,V)=\frac{1}{p}\sum_{j=1}^{p}\frac{j}{p}=\frac{1}{2}+\frac{1}{p}:=b_p$ et

$$\mathrm{Var}(U)=\frac{1}{p^2}(\sum_{j=1}^{p}\frac{j}{p}+2\sum_{j=1}^{p-1}\sum_{j'=j+1}^{p}\frac{j}{p})=\frac{1}{3}+\frac{1}{2p}+\frac{1}{6p^2}:=a_p.$$

Donc,

$$K=\begin{bmatrix} a_p & b_p \\ b_p & 1 \end{bmatrix}$$

On a

$$\text{dét}K=\frac{1}{12}-\frac{1}{2p}-\frac{5}{6p^2}:=d_p.$$

D'où le résultat puisque:

$$K^{-1}=\frac{1}{d_p}\begin{bmatrix} 1 & -b_p \\ -b_p & a_p \end{bmatrix}.$$

□

Pour tout couple (u,v), $u^2+a_pv^2-2b_puv\leq(1+b_p)u^2+(a_p+b_p)v^2\leq C(u^2+v^2)$ pour p assez grand. D'autre part, on peut trouver $\varepsilon>0$ tel que, pour tout (u,v), $u^2+a_pv^2-2b_puv\geq\varepsilon(u^2+v^2)$. En effet, cette inégalité équivaut à $(1-\varepsilon)u^2+(a_p-\varepsilon)v^2-2b_puv\geq 0$ pour tout (u,v). Ceci est réalisé si $D=b_p^2-(1-\varepsilon)(a_p-\varepsilon)\leq 0$. Comme

$$D=(\frac{1}{2}+\frac{1}{p})^2-(1-\varepsilon)(\frac{1}{3}+O(\frac{1}{p})-\varepsilon)=-\frac{1}{12}+O(\frac{1}{p})+\varepsilon(\frac{4}{3}+O(\frac{1}{p}))+\varepsilon^2.$$

On aura $D<0$ si p est assez grand et ε assez petit (par exemple: $\varepsilon=1/8$ et p assez grand). La preuve de la proposition 6.5.1 est ainsi achevée.□

Par suite, la densité marginale $p_x(u)$ de U (sous P) est aussi encadrée comme suit:

$$\sqrt{\pi/c_1}\;c_1^{-1}\exp -c_1u^2\leq p_x(u)\leq\sqrt{\pi/c_1}\;c_2^{-1}\exp -c_2u^2.$$

Et la densité de $\bar{\xi}_x$ s'écrit:

$$\bar{\pi}_x(\bar{x})=\frac{1}{\sqrt{\Delta}}p_x(\frac{\bar{x}-x}{\sqrt{\Delta}})$$

Régime stationnaire

Supposons que l'équation (6.2) admette une loi stationnaire de densité $\pi(.)$ (avec les conditions habituelles sur la fonction d'échelle et la mesure de vitesse) et considérons le processus (ξ_t) en régime stationnaire. On définit $\bar{\xi} = \frac{1}{p}\sum_{j=1}^{p}\xi_{j\delta}$ et l'on note $\bar{\pi}(.)$ la densité de cette variable. On a :

$$\bar{\pi}(\bar{x}) = \int \bar{\pi}_x(\bar{x})\pi(x)dx.$$

Proposition 6.5.2. *On suppose $\pi(x) \leq \pi_1$ pour tout x. Soit $A = [a, b]$. Il existe des constantes $\bar{\pi}_0, \bar{\pi}_1$, ne dépendant que des bornes de b, b' telles que, pour p assez grand,*

$$\forall \bar{x} \in A, \bar{\pi}_0 \leq \bar{\pi}(\bar{x}) \leq \bar{\pi}_1$$

Preuve. On utilise l'encadrement de $p_x(u)$ et un changement de variables pour obtenir:

$$K_1 \int \exp\left(-c_1 t^2\right) \pi(\bar{x} + t\sqrt{\Delta})dt \leq \bar{\pi}(\bar{x}) \leq K_2 \int \exp\left(-c_2 t^2\right) \pi(\bar{x} + t\sqrt{\Delta})dt \leq C\pi_1.$$

Pour la minoration, on utilise le fait que $\pi(.)$ est minorée sur tout compact par une constante strictement positive. Pour tout $t_0 > 0$ et $\Delta \leq 1$,

$$\int \exp\left(-c_1 t^2\right) \pi(\bar{x} + t\sqrt{\Delta})dt \geq \int_0^{t_0} \exp\left(-c_1 t^2\right) \pi(\bar{x} + t\sqrt{\Delta})dt \geq t_0 \exp\left(-c_1 t_0^2\right) \inf_{u \in [a, b+t_0]} \pi(u).$$

D'où le résultat. □

6.5.2 Densité de la moyenne. Coefficient de diffusion non constant.

Condition initiale déterministe.

On considère à présent l'équation différentielle stochastique (6.3) avec les hypothèses suivantes: Les fonctions b, σ sont de $\mathbb{R}$ dans $\mathbb{R}$ et $x \in \mathbb{R}$ avec:

- (H1) b est de classe C^1,
- (H2) b, b' sont bornées,
- (H3) σ est de classe C^2 et il existe $\sigma_0 > 0$ tel que, pour tout u dans $\mathbb{R}$, $\sigma(u) \geq \sigma_0$.
- (H4) σ, σ' sont bornées.

Posons $\bar{\xi}_x = \frac{1}{p}\sum_{j=1}^{p}\xi_{j\delta}$ pour δ un pas de discrétisation. La loi de $\bar{\xi}_x$ admet une densité $\bar{\pi}_x(.)$. Pour étudier cette densité, on introduit le processus transformé suivant :

$$x_u = \frac{1}{\sqrt{\Delta}}(\xi_{u\Delta} - x), \quad \text{où} \quad \Delta = p\delta. \tag{6.21}$$

Le processus $(x_u, u \in [0, 1])$ est solution de:

$$dx_u = \tilde{b}(x_u)du + \tilde{\sigma}(x_u)d\tilde{W}_u, \quad x_0 = 0, \tag{6.22}$$

avec $\tilde{b}(t) = \sqrt{\Delta} b(t\sqrt{\Delta} + x)$, $\tilde{\sigma}(t) = \sigma(t\sqrt{\Delta} + x)$, $\tilde{W}_u = W_{u\Delta}/\sqrt{\Delta}$. On a:

$$(\bar{\xi}_x, \xi_\Delta) = (x + \sqrt{\Delta}U, x + \sqrt{\Delta}V), \quad U = \frac{1}{p}\sum_{j=1}^{p} x_{j/p}, \quad V = x_1. \tag{6.23}$$

On étudie la loi jointe de (U, V). Cette loi admet une densité puisque (x_u) admet des densités de transition. On a la proposition analogue à la proposition (6.5.1).

Proposition 6.5.3. *Soit $p_x(u, v)$ la densité de la loi du couple (U, V) défini en (6.23). Il existe des constantes $c_1 > 0, c_2 > 0$ qui ne dépendent que des bornes de b, b', σ, σ' (et pas de x) telles que, pour p assez grand,*

$$c_1^{-1} \exp -c_1(u^2 + v^2) \leq p_x(u, v) \leq c_2^{-1} \exp -c_2(u^2 + v^2)$$

La preuve de cette proposition suit pas à pas celle de Gobet et Gloter (2008). Notons que:

$$U = \int_0^1 x_u \mu_p(du),$$

où $\mu_p(du) = \frac{1}{p}\sum_{j=1}^{p} \delta_{\frac{j}{p}}(du)$. Gobet et Gloter étudient le cas où $U = \int_0^1 x_u \mu(du)$ avec μ une probabilité sur $[0, 1]$. Par rapport à l'étude de Gobet et Gloter, il y a deux différences. La première est une simplification puisqu'il n'est pas besoin de prouver que (U, V) a une densité. D'un autre côté, la mesure μ_p dépend de p et il faut tenir compte de cette dépendance dans notre cas.

Tout d'abord, nous effectuons la transformation classique sur (x_u) pour se ramener à une équation de coefficient de diffusion égal à 1. Posons

$$F(u) = \int_0^u \frac{du}{\tilde{\sigma}(u)}. \tag{6.24}$$

Notons que $F' = \frac{1}{\tilde{\sigma}}, F'' = -\frac{\tilde{\sigma}'}{\tilde{\sigma}^2}$. La fonction F est de $\mathbb{R}$ dans $\mathbb{R}$ strictement croissante, inversible d'inverse notée F^{-1}. Donc, $y_t = F(x_t)$ satisfait:

$$dy_t = \tilde{\alpha}(y_t)dt + d\tilde{W}_t, \quad y_0 = 0,$$

avec $\tilde{\alpha} = \tilde{b} \circ F^{-1} - \tilde{\sigma}' \circ F^{-1}$. On a:

$$U = \frac{1}{p}\sum_{j=1}^{p} F^{-1}(y_{j/p}), \quad V = F^{-1}(y_1). \tag{6.25}$$

Soit h_0, h_1 deux fonctions positives. Comme plus haut, on peut écrire:

$$\mathbb{E}_P h_0(U)h_1(V) = \mathbb{E}_{\tilde{W}} h_0(U)h_1(V)L_1, \tag{6.26}$$

où

$$L_1 = \exp\left(\int_0^1 \tilde{\alpha}(y_u)dy_u - \frac{1}{2}\int_0^1 \tilde{\alpha}^2(y_u)du\right) = \exp(\tilde{A}(y_1) - \frac{1}{2}\int_0^1 \tilde{h}(y_u)du),$$

avec $\tilde{A}(y) = \int_0^y \tilde{\alpha}(u)du$ et $\tilde{h} = \tilde{\alpha}^2 + \tilde{\alpha}'$. Les hypothèses sur b, b', σ, σ' impliquent

$$|\tilde{h}| \leq C, \quad |\tilde{A}(y)| \leq C\sqrt{\Delta}|y|.$$

Par suite, on peut trouver $C > 0$ tel que:

$$C^{-1} \exp(-C\sqrt{\Delta}y_1) \leq L_1 \leq C \exp(C\sqrt{\Delta}y_1).$$

Or, $y_1 = F(x_1)$ et d'après la définition de F, il existe $c > 0$ tel que

$$c^{-1}u \leq F(u) \leq cu.$$

Comme $y_1 = F(x_1) = F(V)$, on obtient pour une constante $C_1 > 0$:

$$C_1^{-1}\mathbb{E}_{\tilde{W}} h_0(U)h_1(V) \exp(-C_1\sqrt{\Delta}V) \leq \mathbb{E}_P h_0(U)h_1(V) \leq C_1 \mathbb{E}_{\tilde{W}} h_0(U)h_1(V) \exp(C_1\sqrt{\Delta}V). \tag{6.27}$$

A nouveau, il suffit d'encadrer la densité jointe de (U, V) sous $\tilde{W}$. Cette loi n'est plus explicite (voir (6.25)).

Sous $\tilde{W}$, le processus (y_t) est un mouvement brownien standard. On écrit puisque $V = F^{-1}(y_1)$:

$$\mathbb{E}_{\tilde{W}} h_0(U)h_1(V) = \mathbb{E}_{\tilde{W}}(h_1(V)\mathbb{E}_{\tilde{W}}(h_0(U)|y_1)) = \int h_1(F^{-1}(y))n(y)E_{\tilde{W}}(h_0(U)|y_1 = y)dy, \tag{6.28}$$

où $n(y)$ représente la densité gaussienne centrée réduite (densité de y_1). On a alors le lemme suivant.

Lemme 6.5.2. *Il existe $c > 0$ et $\varepsilon_1 > 0$ tels que :*

$$\mathbb{E}_{\tilde{W}}(h_0(U)|y_1)) \geq c^{-1} \exp(-cy_1^2) \int h_0(u) \exp(-cu^2)du.$$

Pour tout $\varepsilon \in]0, \varepsilon_1[$, il existe $c(\varepsilon) > 0$ tel que

$$\mathbb{E}_{\tilde{W}}(h_0(U)|y_1)) \leq c(\varepsilon)^{-1} \exp(\varepsilon y_1^2) \int h_0(u) \exp(-c(\varepsilon)u^2)du.$$

Admettant le lemme, on déduit en utilisant (6.28):

$$\mathbb{E}_{\tilde{W}} h_0(U)h_1(V) \geq c^{-1} \int h_0(u) \exp(-cu^2)du \int h_1(F^{-1}(y)) \exp(-cy^2)n(y)dy.$$

Dans la dernière intégrale, on fait le changement de variables: $v = F^{-1}(y)$, donc $dy = \frac{1}{\tilde{\sigma}(v))}dv$. On obtient:

$$\int h_1(F^{-1}(y)) \exp(-cy^2)n(y)dy \propto \int h_1(v) \exp(-(c + \frac{1}{2})F^2(v))\frac{1}{\tilde{\sigma}(v))}dv \geq c' \int h_1(v) \exp(-c'v^2)dv,$$

puisque $\tilde{\sigma}_0 \leq \tilde{\sigma}(.) \leq \tilde{\sigma}_1$ et $\frac{v}{\sigma_0} \leq F(v) \leq \frac{v}{\sigma_1}$. Ceci nous donne donc la minoration voulue de la densité jointe de (U, V) sous $\tilde{W}$ et par conséquent la minoration de la Proposition 6.5.3.

Pour obtenir la majoration, on utilise la majoration du lemme 6.5.2:

$$\mathbb{E}_{\tilde{W}} h_0(U)h_1(V) \le c(\varepsilon)^{-1} \int h_0(u)\exp(-c(\varepsilon)u^2)du \int h_1(F^{-1}(y))\exp(\varepsilon y^2)n(y)dy.$$

Comme $\exp(\varepsilon y^2)n(y) \propto \exp(-\frac{1}{2}(1-2\varepsilon)y^2) \le \exp(-\frac{1}{4}y^2)$ dès que $\varepsilon \le \frac{1}{4}$, on conclut comme pour la minoration avec le changement de variables $v = F^{-1}(y)$. La preuve de la proposition 6.5.3 est donc achevée.

Il reste à prouver le lemme 6.5.2.

Preuve du lemme 6.5.2. Sous $\tilde{W}$, (y_t) est un mouvement brownien standard et $(B_t^0 = y_t - ty_1, t \in [0,1]$ est un pont brownien standard indépendant de y_1. Par suite,

$$A := \mathbb{E}_{\tilde{W}}(h_0(U)|y_1 = y) = \mathbb{E}(h_0(\frac{1}{p}\sum_{j=1}^{p} F^{-1}(\frac{j}{p}y + B^0_{\frac{j}{p}}))) = \mathbb{E}(h_0(\int_0^1 F^{-1}(ty + B_t^0)d\mu_p(t)). \quad (6.29)$$

On a besoin de la représentation suivante du pont brownien.

Lemme 6.5.3. *Soit $(B_t^0, t \in [0,1])$ un pont brownien standard. Il admet la représentation suivante:*

$$B_t^0 = \xi\eta(t) + \bar{B}_t,$$

où $\eta(t) = t$ si $t \in [0,1/2]$, $\eta(t) = 1-t$ si $t \in [1/2,1]$, $\bar{B}_t = B_t^{(1)}, t \in [0,1/2]$ est un pont brownien sur $[0,1/2]$, $\bar{B}_t = B_t^{(2)}, t \in [1/2,1]$ est un pont brownien sur $[1/2,1]$, ξ est une variable gaussienne centrée réduite indépendante de $(\bar{B}_t, t \in [0,1]$.

Preuve. On part de $B_t^0 = W_t - tW_1$ où (W_t) est un mouvement brownien standard. Si $t \in [0,1/2]$,

$$B_t^0 = B_t^{(1)} + t\xi, \quad B_t^{(1)} = W_t - 2tW_{1/2}, \quad \xi = 2W_{1/2} - W_1.$$

Si $t \in [1/2,1]$,

$$B_t^0 = B_t^{(2)} + (1-t)\xi, \quad B_t^{(2)} = W_t - W_{1/2} - \frac{t-1/2}{1-1/2}(W_1 - W_{1/2}).$$

On vérifie le lemme.□

Nous pouvons passer à l'étude de A (voir (6.29)). Définissons la fonction

$$g_{\bar{B}}(x) = \int_0^1 F^{-1}(ty + \eta(t)x + \bar{B}_t)d\mu_p(t).$$

Comme ξ et $\bar{B}$ sont indépendants, on a:

$$A = \mathbb{E}h_0(g_{\bar{B}}(\xi)) = \mathbb{E}\int n(x)h_0(g_{\bar{B}}(x))dx.$$

Or, la fonction $x \to g_{\bar{B}}(x)$ est C^1 et sa dérivée vérifie pour une constante $c > 0$ ne dépendant que des bornes de l'encadrement de $\tilde{\sigma}$,

$$c^{-1}\int_0^1 \eta(t)d\mu_p(t) \le g'_{\bar{B}}(x) \le c\int_0^1 \eta(t)d\mu_p(t).$$

Un calcul élémentaire montre que $\int_0^1 \eta(t)d\mu_p(t) = 1/4 + O(1/p)$. Ainsi, la fonction $g_{\bar{B}}(x)$ est inversible sur $\mathbb{R}$ de dérivée majorée et minorée par une constante ne dépendant ni de $\bar{B}$ ni de y. On peut donc faire le changement de variables $u = g_{\bar{B}}(x)$. Ceci donne:

$$c^{-1} \int h_0(u) \exp(-\frac{1}{2}(g_{\bar{B}}^{-1}(u))^2) du \leq \int n(x) h_0(g_{\bar{B}}(x)) dx \leq c \int h_0(u) \exp(-(-\frac{1}{2}(g_{\bar{B}}^{-1}(u))^2) du$$

Les fonctions $g_{\bar{B}}^{-1}, g_{\bar{B}}$ sont globalement lipschitziennes de constantes de Lipschitz ne dépendant pas de $\bar{B}$. On en dÃľduit, pour tout u:

$$|g_{\bar{B}}^{-1}(u)| \leq c|u| + |g_{\bar{B}}^{-1}(0)|, \quad |g_{\bar{B}}^{-1}(0)| = |g_{\bar{B}}^{-1}(g_{\bar{B}}(0)) - g_{\bar{B}}^{-1}(0)| \leq c|g_{\bar{B}}(0)|.$$

De plus,

$$|g_{\bar{B}}(0)| \leq c(|y| + \sup_{t\in[0,1]} |\bar{B}_t|.$$

On en déduit une nouvelle minoration:

$$\int n(x) h_0(g_{\bar{B}}(x)) dx \geq c^{-1} \exp(-cy^2) \exp(-c \sup_{t\in[0,1]} \bar{B}_t^2) \int h_0(u) \exp(-cu^2) du.$$

Il suffit de prendre l'espérance de cette expression pour obtenir la minoration de (6.29) recherchée.

Pour la majoration de A, on utilise d'une part:

$$||u| - |g_{\bar{B}}(0)|| \leq |g_{\bar{B}}(g_{\bar{B}}^{-1}(u)) - g_{\bar{B}}(0)| \leq c|g_{\bar{B}}^{-1}(u)) - 0|,$$

et d'autre part, l'inégalité valable pour $x, y \in \mathbb{R}$ et $\varepsilon \in (0,1)$

$$(x-y)^2 \geq x^2 \frac{\varepsilon}{1+\varepsilon} - \varepsilon y^2.$$

On obtient:

$$(g_{\bar{B}}^{-1}(u)))^2 \geq \frac{1}{c^2}(u^2 \frac{\varepsilon}{1+\varepsilon} - \varepsilon g_{\bar{B}}(0))^2).$$

Puis,

$$-\frac{1}{2}(g_{\bar{B}}^{-1}(u)))^2 \leq -\frac{1}{2c^2} u^2 \frac{\varepsilon}{1+\varepsilon} + \frac{\varepsilon}{2c}(y^2 + \sup_{t\in[0,1]} \bar{B}_t^2).$$

Ainsi, on aboutit à:

$$\int n(x) h_0(g_{\bar{B}}(x)) dx \leq c \exp(\frac{\varepsilon}{2c} y^2) \exp(\frac{\varepsilon}{2c} \sup_{t\in[0,1]} \bar{B}_t^2) \int h_0(u) \exp(-\frac{u^2}{2c^2} \frac{\varepsilon}{1+\varepsilon}).$$

Il reste à prendre l'espérance (par rapport à $\bar{B}$) de l'expression ci-dessus pour obtenir la majoration cherchée. Or l'espérance

$$\mathbb{E} \exp(\frac{\varepsilon}{2c} \sup_{t\in[0,1]} \bar{B}_t^2)$$

est finie pour ε assez petit. La preuve du lemme 6.5.2 est achevée. Par conséquent, la preuve de la proposition 6.5.3 est également achevée.

Régime stationnaire.

On peut reproduire exactement le paragraphe 6.5.1 pour obtenir la proposition 6.5.2.

6.6 Prérequis.

Rappel: La loi du $(\chi')^2(n, \lambda)$ est la loi de $Z = \sum_{i=1}^n Y_i^2$ où $(Y_1, \ldots, Y_n$ sont indépendantes, $Y_i \sim \mathcal{N}(m_i, 1)$ et $\lambda = \sum_{i=1}^n m_i^2$ (on montre que la loi de Z ne dépend de $m_1, \ldots, m_n$ qu'à travers λ.

Nous donnons quelques résultats pour les martingales qui sont celles utilisées dans le chapitre, notamment des versions simplifiées mais pratiques du théorème d'arrêt.

Pour plus de précisions, on pourra consulter Revuz et Yor (1991) ou Karatsas et Shreve (1991).

Théorème 6.6.1. *Soit $(M_t, t \geq 0)$ une martingale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ et $S \leq T$ deux temps d'arrêt bornés ($\forall \omega, S(\omega) \leq T(\omega) \leq C$). Alors,*

$$\mathbb{E}(M_T) = \mathbb{E}(M_S) = \mathbb{E}(M_0).$$

Théorème 6.6.2. *Soit $(M_t, t \geq 0)$ une martingale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ uniformément bornée (($\sup_{t\geq 0} |M_t| \leq c$, c déterministe). Alors, il existe une v.a. M_∞ intégrable telle que, pour tout t, $M_t = \mathbb{E}(M_\infty | \mathcal{F}_t)$. De plus, M_t converge p.s. et dans tous les L^p vers M_∞.*

Théorème 6.6.3. *Soit $(M_t, t \geq 0)$ une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$(cf Chapitre 3). Pour tout temps d'arrêt T, $(M_{t\wedge T}, t \geq 0)$ une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$.*
Si $(M_t, t \geq 0)$ une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ uniformément bornée ($\sup_{t\geq 0} |M_t| \leq c$, c déterministe), alors $(M_t, t \geq 0)$ une vraie martingale.
Si $(M_t, t \geq 0)$ une martingale locale relative à $\mathbb{P}$-$(\mathcal{F}_t)$ telle que $\mathbb{E}M_t^2 < +\infty$ pour tout $t \geq 0$, alors $(M_t, t \geq 0)$ une vraie martingale de carré intégrable.

6.7 Références bibliographiques

1. M. Abramowitz, I.A. Stegun, *Handbook of mathematical functions with formulas, graphs and mathematical tables.* National Bureau of Standards Applied Mathematics Series, 55 For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

2. J. Cox, J. Ingersoll and S. Ross, A theory of the term structure of interest rates, *Econometrica* **53**, 385-467, 1984.

3. P.J. Fitzsimmons, J. Pitman, Kac's moment formula and the Feynman-Kac formula for additive functionals of a Markov process. *Stochastic processes and their applications* **79** 117-134, 1999.

4. E. Gobet, A. Gloter, LAMN property for hidden processes: the case of integrated diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, **44**, 104-128, 2008.

5. N. Ikeda et S. Watanabe. *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha, 1981.

6. I. Karatsas et S.E Shreve, *Brownian motion and stochastic calculus*, second edition, Graduate texts in Mathematics, Springer, 1991.

7. Karlin S. et Taylor H.M., *A second course in stochastic processes*, Academic Press, 1981.

8. D. Revuz et Marc Yor, *Continuous martingales and Brownian motion*, Springer, Berlin,1991,

9. L.C.G. Rogers, D. Williams, *Diffusions, Markov processes and Martingales, Volume 2: Ito calculus*, Wiley, 1990.

Chapter 7

Diffusions récurrentes positives sur un intervalle. Estimation paramétrique de dérive.

Dans ce chapitre nous étudions l'estimation de θ_0 (vraie valeur) à partir de l'observation du processus (ξ_t) défini par:

$$d\xi_t = b(\xi_t, \theta_0)dt + \sigma(\xi_t)dB_t, \qquad \xi_0 = x_0. \tag{7.1}$$

Le processus (ξ_t) est défini sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$ ne dépendant pas de T, ainsi que le mouvement brownien (B_t). On suppose x_0 connu, ainsi que la fonction $\sigma(.)$. La fonction $b(x, \theta_0)$ dépend d'un paramètre inconnu $\theta_0 \in \Theta \subset \mathbb{R}^p$.
Pour ce chapitre, nous utiliserons les résultats vus au chapitre 6.

7.1 Estimation par maximum de vraisemblance

Nous considérons des hypothèses assurant que le processus (7.1) est bien défini et possède les propriétés de récurrence positive sur un intervalle. Dans le cadre asymptotique $T \to +\infty$, on peut donner des résultats généraux de consistance et normalité asymptotique des estimateurs du maximum de vraisemblance (EMV) à la vitesse $\sqrt{T}$.

7.1.1 Hypothèses sur le modèle

Considérons les hypothèses:

(H1) Pour tout $\theta \in \Theta$, les fonctions $x \to b(x, \theta)$ et la fonction $x \to \sigma(x)$ vérifient les conditions (A') ou (B') (cf Chapitre 6, section 6.1).
On noteP_θ^T la loi de probabilité sur $(C_T, \mathcal{C}_T)$ du processus $(\xi_t^\theta, t \in [0, T])$ défini par

$$d\xi_t^\theta = b(\xi_t^\theta, \theta)dt + \sigma(\xi_t^\theta)dB_t, \qquad \xi_0^\theta = x_0. \tag{7.2}$$

(H2) Il existe un intervalle $I = (\ell, r)$ avec ℓ, r connus, $-\infty \leq \ell < r \leq +\infty$ tel que: $\forall x \in I$, $\sigma^2(x) > 0$, σ de classe C^2 sur I, $\forall \theta$, $b(,.\theta)$ de classe C^1 sur I.
Soit

$$s_\theta(u) = \exp[-2 \int_{u_0}^{u} \frac{b(v,\theta)}{\sigma^2(v)} dv], \quad u_0, u \in I. \tag{7.3}$$

On suppose que, pour tout $\theta \in \Theta$,

$$\int_\ell s_\theta(u) du = +\infty, \quad \int^r s_\theta(u) du = +\infty.$$

Soit

$$m_\theta(u) = \frac{1}{\sigma^2(u) s_\theta(u)}, \quad u \in I. \tag{7.4}$$

On suppose que, pour tout θ, $M_\theta = \int_I m_\theta(u) du < +\infty$ et l'on pose:

$$\pi_\theta(u) = \frac{1}{M_\theta} m_\theta(u) 1_{(u \in I)}. \tag{7.5}$$

On suppose que $x_0 \in I$.

Sous (H1)-(H2), le processus (ξ_t^θ) est récurrent positif sur I et admet pour unique distribution stationnaire la loi $\pi_\theta(u)du$. On a, en particulier, pour tout θ, $\mathbb{P}(\xi_t^\theta \in I, \forall t \geq 0) = 1$. Ce qui implique:

$$\mathbb{P}(\sigma^2(\xi_t^\theta) > 0, \forall t \in [0,T]) = P_\theta^T(\sigma^2(X_t) > 0, \forall t \in [0,T]) = 1$$

où l'on rappelle que (X_t) est le processus canonique de C_T.
On verra sur les exemples que (H2) implique souvent une restriction de l'espace des paramètres.
La fonction de vraisemblance du modèle $(C_T, \mathcal{C}_T, P_\theta^T)_{\theta \in \Theta}$ peut être définie par: $\theta \to L_T(\theta) = \exp \ell_T(\theta)$ avec

$$\ell_T(\theta) = \int_0^T \frac{b(\xi_s, \theta)}{\sigma^2(\xi_s)} d\xi_s - \frac{1}{2} \int_0^T \frac{b^2(\xi_s, \theta)}{\sigma^2(\xi_s)} ds$$

où $(\xi_s = \xi_s^{\theta_0})$ est le processus observé correspondant à la vraie valeur θ_0 du paramètre.

7.1.2 Consistance des estimateurs du maximum de vraisemblance

Rappelons qu'on appelle estimateur du maximum de vraisemblance (EMV) de θ_0, noté $\hat{\theta}_T$, toute solution de l'équation de vraisemblance:

$$\ell_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} \ell_T(\theta). \tag{7.6}$$

Deux cas se présentent en pratique:
- ou bien, $\hat{\theta}_T$ peut être calculé explicitement en fonction de $(\xi_t, t \in [0,T])$. On peut alors l'étudier directement;
- ou bien, $\hat{\theta}_T$ reste défini de façon implicite par l'équation ci-dessus. Il faut l'étudier par l'équation de vraisemblance. C'est le cas général.
Dans tous les cas, l'équation de vraisemblance étant explicite, l'EMV peut être calculé

numériquement à partir de l'équation. Voyons un exemple explicite.

Exemple: Considérons le modèle d'Ornstein-Uhlenbeck:

$$d\xi_t^\theta = \theta\xi_t^\theta dt + \sigma dB_t, \;\; \xi_0^\theta = 0,$$

où $\theta \in \mathbb{R}$. Soit $(\xi_t = \xi_t^{\theta_0})$ le processus observé. L'hypothèse (H1) est vérifiée. On suppose la constante $\sigma \neq 0$ connue et on prend $(\ell, r) = (-\infty, +\infty)$. Pour vérifier (H2), on calcule les fonctions s_θ, m_θ:

$$s_\theta(u) = \exp\big(-\frac{2}{\sigma^2}\int_0^u \theta v dv\big) = \exp\big(-\frac{\theta}{\sigma^2}u^2\big), \;\; m_\theta(u) = \frac{1}{\sigma^2 s_\theta(u)}.$$

On a $\int_{-\infty} s_\theta = \int^{+\infty} s_\theta = +\infty$ si et seulement si $\theta \leq 0$ et $\int_{\mathbb{R}} m_\theta(u)du < +\infty$ si et seulement si $\theta < 0$. Le modèle est récurrent positif ((H2)) pour $\theta < 0$ et la loi stationnaire est la loi gaussienne $\pi_\theta = \mathcal{N}(0, \sigma^2/2|\theta|)$. La log-vraisemblance:

$$\ell_T(\theta) = \frac{\theta}{\sigma^2}\int_0^T \xi_s d\xi_s - \frac{\theta^2}{2\sigma^2}\int_0^T \xi_s^2 ds$$

admet comme unique maximum

$$\hat{\theta}_T = \frac{\int_0^T \xi_s d\xi_s}{\int_0^T \xi_s^2 ds}.$$

Nous avons déjà étudié ce modèle et cet estimateur dans le chapitre 5. Reprenons ici cette étude dans le cas $\theta < 0$ avec l'aide des résultats de récurrence positive. En particulier, nous pouvons finir la preuve de la proposition 5.1.1. Pour étudier les propriétés asymptotiques de $\hat{\theta}_T$, on écrit:

$$\hat{\theta}_T = \frac{\int_0^T \xi_s d\xi_s}{\int_0^T \xi_s^2 ds} = \frac{\int_0^T \xi_s(\theta_0\xi_s ds + \sigma dB_s)}{\int_0^T \xi_s^2 ds} = \theta_0 + \sigma\frac{M_T}{< M >_T}$$

avec $M_T = \int_0^T \xi_s d\xi_s$ et $< M >_T = \int_0^T \xi_s^2 ds$. Comme $\int_{\mathbb{R}} x^2 d\pi_{\theta_0}(x) = \sigma^2/2|\theta_0| < +\infty$, d'après le théorème ergodique (cf Chapitre 6, 6.3.6), p.s., lorsque T tend vers l'infini,

$$\frac{1}{T}\int_0^T \xi_s^2 ds \to \sigma^2/2|\theta_0|.$$

Donc, $M_T = \int_0^T \xi_s dB_s$ vérifie $< M >_T = \int_0^T \xi_s^2 ds \to +\infty$. Par suite, $M_T/ < M >_T$ converge p.s. vers 0: $\hat{\theta}_T \to \theta_0$ p.s.. On peut aussi conclure pour la loi limite et retrouver le résultat du chapitre 5: quand T tend vers l'infini,

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = \sigma\frac{M_T/\sqrt{T}}{< M >_T /T} \to_{\mathcal{L}} \mathcal{N}(0, 2|\theta_0|).$$

En général, on ne dispose pas de formule explicite pour $\hat{\theta}_T$. Il faut l'étudier avec l'équation de vraisemblance (7.6).Cette équation peut avoir plusieurs solutions. La méthode des contrastes permet de montrer que *toute* solution de l'équation de vraisemblance converge vers la vraie valeur θ_0 en probabilité sous des hypothèses supplémentaires (cf Dacunha-Castelle et Duflo, 1983, van der Vaart, 1998). Introduisons les hypothèses:

(K1) Θ est un compact de $\mathbb{R}^p$.

(K2) (1) La fonction $(x,\theta) \to b(x,\theta)$ est continue sur $I \times \Theta$ et
(2) pour tout ω, la fonction $\theta \to \ell_T(\theta,\omega)$ est continue sur Θ.

(K3) On suppose que

$$K(\theta_0,\theta) = \frac{1}{2}\int_\ell^r \frac{(b(u,\theta)-b(u,\theta_0))^2}{\sigma^2(u)}\pi_{\theta_0}(u)du < +\infty.$$

et que $\theta \neq \theta_0 \Rightarrow K(\theta_0,\theta) > 0$.

(K4) Il existe une v.a. $Z_T \geq 0$ et une fonction $\beta(h)$ définie sur $\mathbb{R}^+$ telles que

$$\sup_{\theta,\theta'\in\Theta,|\theta-\theta'|\leq h} \frac{1}{T}|\ell_T(\theta)-\ell_T(\theta')| \leq \beta(h)Z_T$$

où, lorsque T tend vers l'infini, Z_T est bornée en probabilité et $\lim_{h\to 0}\beta(h)=0$.
$(\forall\varepsilon > 0, \exists M > 0, \exists T_0 > 0, \forall T \geq T_0,\ \mathbb{P}(|Z_T| \geq M) \leq \varepsilon)$.

L'hypothèse (K1) est simplificatrice: elle permet de prouver la consistance sans avoir recours à l'étude du processus de vraisemblance (pour cette étude, consulter Kutoyants, 2004); elle est souvent vérifiée en pratique et désormais admise dans la littérature.
L'hypothèse (K2) (2) signifie qu'il existe une version continue de la log-vraisemblance et que l'on choisit l'une de celles-ci. Une difficulté provient de ce que $\ell_T(\theta)$ contient une intégrale stochastique dépendant du paramètre θ et la seule condition (K2)(1) ne suffit pas. En pratique, il est souvent possible de vérifier (K2)(2) sur la formule explicite de $\ell_T(\theta)$ (voir l'exemple précédent) ou d'appliquer le lemme de Kolmogorov (Revuz et Yor, 1991, Theorem 2.1, p.25, par exemple).
L'hypothèse (K3) est essentielle: c'est l'hypothèse d'identifiabilité de θ. Elle signifie que le modèle est bien paramétré relativement à ce cadre asymptotique. La fonction $K(\theta_0,\theta)$ joue le rôle d'information de Kullback limite de la loi $P^T_{\theta_0}$ par rapport à P^T_θ.
L'hypothèse (K4) est classique: le module de continuité de la fonction de contraste $T^{-1}\ell_T(\theta)$ doit être contrôlé.

Proposition 7.1.1. *On suppose (H1)-(H2) (cf Section 7.1.1), (K1)-(K3).*
Alors, $U_T(\theta) = -T^{-1}\ell_T(\theta)$ est un contraste au sens suivant:

- *Pour tout ω, la fonction $\theta \to U_T(\theta)$ est continue.*
- *Pour tout θ, $U_T(\theta) - U_T(\theta_0)$ converge en probabilité vers $K(\theta_0,\theta)$.*
- *La fonction $\theta \to K(\theta_0,\theta)$ est continue sur Θ et admet un unique minimum en $\theta = \theta_0$.*

La fonction $\theta \to K(\theta_0,\theta)$ s'appelle la fonction de contraste associée à $U_T(\theta)$.

Dans ce qui suit, nous avons besoin d'utiliser des versions de la log-vraisemblance et de ses dérivées par rapport au paramètre qui ne contiennent pas d'intégrale stochastique, ceci pour pouvoir écrire $\ell_T(\hat{\theta}_T)$ sans faire de "plug-in" d'une variable aléatoire dépendant de toute la trajectoire dans une intégrale stochastique. Pour cela, posons:

$$B(x,\theta) = \int_{x_0}^{x} \frac{b^2(u,\theta)}{\sigma^2(u)}du, \qquad g(.,\theta) = \frac{\partial}{\partial x}\left[\frac{b^2(.,\theta)}{\sigma^2(.)}\right],$$

de sorte que:

$$\frac{\partial}{\partial\theta_j}B(x,\theta) = \int_{x_0}^{x} \frac{\frac{\partial}{\partial\theta_j}b^2(u,\theta)}{\sigma^2(u)}du, \quad \frac{\partial}{\partial\theta_j}g(x,\theta) = \frac{\partial}{\partial x}\left[\frac{\frac{\partial}{\partial\theta_j}b^2(x,\theta)}{\sigma^2(x)}\right].$$

On calculera de façon analogue les dérivées secondes par rapport aux composantes de θ. Alors,

$$\ell_T(\theta) = B(\xi_T,\theta) - \frac{1}{2}\int_0^T \left(\frac{b^2(\xi_s,\theta)}{\sigma^2(\xi_s)} + \sigma^2(\xi_s)g(\xi_s,\theta)\right)ds.$$

De façon analogue, on pourra écrire une version de $\frac{\partial}{\partial\theta_j}\ell_T(\theta)$ et de $\frac{\partial^2}{\partial\theta_j\partial\theta_k}\ell_T(\theta)$ ne contenant pas d'intégrale stochastique. Ce sont ces versions qui sont utilisées dans les preuves suivantes.

On a le théorème:

Théorème 7.1.1. *On suppose (H1)-(H2), (K1)-(K4). Soit $\hat{\theta}_T$ un EMV de θ_0. On a:*

$$\forall h > 0, \;\; \mathbb{P}(|\hat{\theta}_T - \theta_0| > h) \to_{T\to+\infty} 0.$$

($\hat{\theta}_T$ est consistant).

Preuve de la proposition 7.1.1 On a:

$$\ell_T(\theta) = \int_0^T \frac{b(\xi_s,\theta)}{\sigma^2(\xi_s)}\left(b(\xi_s,\theta_0)ds + \sigma(\xi_s)\right)dB_s - \frac{1}{2}\int_0^T \frac{b^2(\xi_s,\theta)}{\sigma^2(\xi_s)}ds.$$

Par suite, $U_T(\theta) - U_T(\theta_0) = \frac{M_T}{T} + \frac{<M>_T}{2T}$ avec

$$M_T = \int_0^T \frac{b(\xi_s,\theta) - b(\xi_s,\theta_0)}{\sigma(\xi_s)}dB_s, < M >_T = \int_0^T \frac{\left(b(\xi_s,\theta) - b(\xi_s,\theta_0\right)^2}{\sigma^2(\xi_s)}ds.$$

Le théorème ergodique valable sous (H1)-(H2) implique que $\frac{<M>_T}{2T}$ converge p.s. vers $K(\theta_0,\theta)$ et $\frac{M_T}{T} = \frac{M_T}{<M>_T}\frac{<M>_T}{T}$ converge p.s. vers 0. Les hypothèses impliquent le reste de la proposition. □

Preuve du théorème 7.1.1 Nous esquissons la preuve qui est celle, classique, de consistance des estimateurs de minimum de contraste. Le principe est que la propriété de contraste ainsi que l'uniformité de la convergence vers $K(\theta_0,\theta)$ permet de montrer que, en probabilité,

$$\hat{\theta}_T = \text{arginf}_\theta\; U_T(\theta) - U_T(\theta_0) \to_{T\to+\infty} \theta_0 = \text{arginf}_\theta\; K(\theta_0,\theta).$$

En effet, si, pour $h > 0$, $\hat{\theta}_T \in \{|\theta - \theta_0| > h\}$, alors,

$$\inf_{\{|\theta-\theta_0|>h\}} (U_T(\theta) - U_T(\theta_0)) = U_T(\hat{\theta}_T) - U_T(\theta_0) = \inf_{\Theta} (U_T(\theta) - U_T(\theta_0)) \leq U_T(\theta_0) - U_T(\theta_0) = 0.$$

On a donc l'inclusion

$$\{|\hat{\theta}_T - \theta_0| > h\} \subset \{\inf_{\{|\theta-\theta_0|>h\}} (U_T(\theta) - U_T(\theta_0)) \leq 0\}.$$

Il reste à prouver que $\mathbb{P}(\{\inf_{\{|\theta-\theta_0|>h\}} U_T(\theta) - U_T(\theta_0) \leq 0\})$ tend vers 0. On sait que $U_T(\theta) - U_T(\theta_0)$ tend vers $K(\theta_0, \theta) > 0$ pour $\theta \neq \theta_0$. Pour conclure, on utilise la compacité de Θ et les autres hypothèses notamment (K3)-(K4) (cf par exemple, Dacunha-Castelle et Duflo, 1983, p.93-94).□

7.1.3 Normalité asymptotique

Rajoutons les hypothèses:

(K5) La vraie valeur du paramètre $\theta_0 \in \mathring{\Theta}$.

(K6) (1) Pour $i, j = 1, \ldots, p$, les dérivées partielles $\frac{\partial}{\partial\theta_i} b(x, \theta)$, $\frac{\partial^2}{\partial\theta_i\partial\theta_j} b(x, \theta)$ sont définies et continues sur $I \times \mathring{\Theta}$.
(2) Pour tout ω, $\ell_T(\theta)$ est de classe C^2 sur $\mathring{\Theta}$ et

$$\begin{aligned}
\frac{\partial}{\partial\theta_i} \ell_T(\theta) &= \int_0^T \frac{\frac{\partial}{\partial\theta_i} b(\xi_s, \theta)}{\sigma^2(\xi_s)} (d\xi_s - b(\xi_s, \theta)ds) \\
\frac{\partial^2}{\partial\theta_i\partial\theta_j} \ell_T(\theta) &= \int_0^T \frac{\frac{\partial^2}{\partial\theta_i\partial\theta_j} b(\xi_s, \theta)}{\sigma^2(\xi_s)} (d\xi_s - b(\xi_s, \theta)ds) - \int_0^T \frac{\frac{\partial}{\partial\theta_i} b(\xi_s, \theta) \frac{\partial}{\partial\theta_j} b(\xi_s, \theta)}{\sigma^2(\xi_s)} ds
\end{aligned}$$

(K7) Pour $i = 1, \ldots, p$,

$$\int_\ell^r \frac{\left(\frac{\partial}{\partial\theta_i} b(u, \theta_0)\right)^2}{\sigma^2(u)} \pi_{\theta_0}(u) du < +\infty$$

et la matrice $\mathcal{I}(\theta_0) = [\mathcal{I}_{ij}(\theta_0)]_{1\leq i,j\leq p}$ est inversible avec

$$\mathcal{I}_{ij}(\theta_0) = \int_\ell^r \frac{\frac{\partial}{\partial\theta_i} b(u, \theta_0) \frac{\partial}{\partial\theta_j} b(u, \theta_0)}{\sigma^2(u)} \pi_{\theta_0}(u) du.$$

(K8) Pour $i, j = 1, \ldots, p$,

$$\mathcal{J}_{ij}(\theta_0) = \int_\ell^r \frac{\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j} b(u, \theta_0)\right)^2}{\sigma^2(u)} \pi_{\theta_0}(u) du < +\infty.$$

(K9) Pour toute v.a. η_T tendant vers 0 en probabilité quand T tend vers l'infini et pour $i, j = 1, \ldots, p$,

$$\sup_{|\theta-\theta_0|\leq\eta_T} \frac{1}{T} \left| \frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta) - \frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0) \right| \to_{T\to+\infty} 0$$

en probabilité.

L'hypothèse (K5) est classique: l'estimateur $\hat{\theta}_T$ étant consistant, elle assure que $\mathbb{P}(\hat{\theta}_T \in \mathring{\Theta}) \to 1$ ce qui permet d'écrire sur l'ensemble $\hat{\theta}_T \in \mathring{\Theta}$, $\frac{\partial}{\partial\theta_i}\ell_T(\hat{\theta}_T) = 0,\ j = 1, \ldots, p$.

L'hypothèse (K6)(2) signifie qu'il existe une version de classe C^2 de $\theta \to \ell_T(\theta)$ et que cette version peut être obtenue en dérivant sous l'intégrale stochastique par rapport à θ. Comme pour la continuité de $\ell_T(\theta)$, on vérifiera ces hypothèses soit directement sur l'expression explicite de $\ell_T(\theta)$ (voir l'exemple du modèle d'Ornstein-Uhlenbeck) soit en appliquant une théorème de dérivation sous l'intégrale stochastique. Nous renvoyons par exemple à Karandikhar, 1983 ou Hutton et Nelson, 1984.
L'hypothèse (K7) est proche de l'hypothèse d'identifiabilité (K3). Si le paramètre est de dimension 1, le nombre $\mathcal{I}(\theta_0)$ est automatiquement strictement positif.
L'hypothèse (K9) est technique et classique.
On peut énoncer:

Proposition 7.1.2. *On suppose (H1)-(H2), (K5), (K6) et (K8). Quand T tend vers l'infini,*

$$\begin{array}{rcl} \dfrac{1}{\sqrt{T}}\left(\dfrac{\partial}{\partial\theta_i}\ell_T(\theta_0), i = 1, \ldots, p\right) & \to_{\mathcal{L}} & \mathcal{N}_p(0, \mathcal{I}(\theta_0)) \\ -\dfrac{1}{T}\left[\dfrac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0)\right]_{1\leq i,j\leq p} & \to_P & \mathcal{I}(\theta_0). \end{array}$$

($\to_{\mathcal{L}}$: converge en loi, $\to_P$: converge en probabilité).

Théorème 7.1.2. *On suppose (H1)-(H2), (K1)-(K9). Alors,*

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \to_{\mathcal{L}} \mathcal{N}_p(0, \mathcal{I}^{-1}(\theta_0)).$$

Preuve de la proposition 7.1.2 Montrons le premier point. En utilisant les formules donnant $(\partial/\partial\theta_i)\ell_T(\theta)$ (voir (K6)(2)), on voit que:

$$\frac{\partial}{\partial\theta_i}\ell_T(\theta_0) = \int_0^T \frac{\frac{\partial}{\partial\theta_i}b(\xi_s, \theta_0)}{\sigma^2(\xi_s)} dB_s := M_T^i$$

avec $T^{-1} < M^i, M^j >_T \to_{T\to+\infty} \mathcal{I}_{ij}(\theta_0)$ p.s. d'après le théorème ergodique. Le théorème de limite centrale pour les intégrales stochastiques permet de conclure (voir Chapitre 6, paragraphe 6.3.6, Chapitre 4, paragraphe 4.3.5, théorèmes 4.3.6, 4.3.7, corollaire 4.3.2).
Pour le second point,

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0) = N_T^{ij} - < M^i, M^j >_T$$

où $N_T^{ij} = \displaystyle\int_0^T \sigma^{-1}(\xi_s)\frac{\partial^2}{\partial\theta_i\partial\theta_j}b(\xi_s,\theta_0)dB_s$ et M_T^i est celui du premier point. En utilisant (K8), on obtient que $T^{-1} < N^{ij} >_T$ converge p.s. vers $\mathcal{J}_{ij}(\theta_0)$. Donc,

$$\frac{1}{T}N_T^{ij} = \frac{N_T^{ij}}{< N^{ij} >_T}\frac{< N^{ij} >_T}{T} \rightarrow_{p.s.} 0.$$

D'où le second point. □

Preuve du théorème 7.1.2 Les hypothèses impliquent que $\hat{\theta}_T$ est consistant donc $\mathbb{P}(\hat{\theta}_T \in \mathring{\Theta}) \to 1$. Sur l'ensemble $\hat{\theta}_T \in \mathring{\Theta}$, on a, pour $i = 1,\ldots,p$:

$$\frac{\partial}{\partial\theta_i}\ell_T(\hat{\theta}_T) = 0 = \frac{\partial}{\partial\theta_i}\ell_T(\theta_0) + \sum_{j=1}^p\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0) + R_T^{ij}\right)(\hat{\theta}_T^j - \theta_0^j)$$

avec

$$R_T^{ij} = \int_0^1\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0 + u(\hat{\theta}_T - \theta_0)) - \frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0)\right)du.$$

En utilisant (K9), puisque $\hat{\theta}_T - \theta_0 \to 0$ en probabilité, pour tout $i,j = 1,\ldots,p$, $T^{-1}R_T^{ij} \to 0$. Par suite,

$$\mathcal{I}_T = \frac{1}{T}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell_T(\theta_0) + R_T^{ij}\right)_{1\le i,j\le p} \rightarrow_P -\mathcal{I}(\theta_0)$$

et $\mathbb{P}(\mathcal{I}_T$ inversible $) \to 1$. On peut écrire:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = -\frac{1}{\sqrt{T}}\mathcal{I}_T^{-1}\left(\frac{\partial}{\partial\theta_i}\ell_T(\theta_0)\right)_{i=1,\ldots,p} \rightarrow_{\mathcal{L}} \mathcal{N}(0,\mathcal{I}^{-1}(\theta_0)).$$

□

Exemples.

- Modèle bilinéaire: Soit le modèle $d\xi_t = (a_0\xi_t + b_0)dt + \sigma\xi_t dW_t, \xi_0 = x > 0$. On peut résoudre cette équation différentielle stochastique. Introduisons le processus

  $$Y_t = \exp\left[(a_0 - (\sigma^2/2))t + \sigma W_t\right]$$

  solution de $dY_t = a_0Y_tdt + \sigma Y_tdW_t, Y_0 = 1$ et posons $Z_t = \xi_t\eta_t$ avec $\eta_t = 1/Y_t$. Par la formule d'Itô appliquée à $F(x,y) = xy$, on obtient:

  $$dZ_t = \eta_t d\xi_t + \xi_t d\eta_t + d < \xi,\eta >_t .$$

  Or

  $$d\eta_t = \eta_t[-(a_0 - (\sigma^2/2))]dt - \sigma\eta_t dW_t + (\sigma^2/2)\eta_t dt = (\sigma^2 - a_0)\eta_t dt - \sigma\eta_t dW_t$$

  et $d < \xi,\eta >_t = -\sigma^2\eta_t\xi_t$. Donc, $dZ_t = b_0\eta_t dt$ ce qui conduit à $Z_t = Z_0 + b_0\int_0^t \eta_s ds = x + b_0\int_0^t \eta_s ds$. Finalement,

  $$\xi_t = x\exp\left[(a_0 - (\sigma^2/2))t + \sigma W_t\right] + b_0\int_0^t \exp\left[(a_0 - (\sigma^2/2))(t-s) + \sigma(W_t - W_s)\right]ds.$$

En particulier, $x>0$ et $b_0 \geq 0$ implique $\xi_t>0$ pour tout t.
Supposons que l'on observe le processus $(\xi_t, t\in[0,T])$ correspondant à la vraie valeur du paramètre (a_0,b_0), avec $b_0\geq 0$, σ étant supposé connu. La log-vraisemblance s'écrit:

$$\ell_T(a,b)=\frac{1}{\sigma^2}\left(\int_0^T \frac{a\xi_s+b}{\xi_s^2}d\xi_s-\frac{1}{2}\int_0^T\frac{(a\xi_s+b)^2}{\xi_s^2}ds\right).$$

L'estimateur du maximum de vraisemblance $\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}$ est solution de

$$I_T\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}=\begin{pmatrix}\int_0^T\frac{d\xi_s}{\xi_s}\\ \int_0^T\frac{d\xi_s}{\xi_s^2}\end{pmatrix}$$

où

$$I_T=\begin{pmatrix}T & \int_0^T\frac{ds}{\xi_s}\\ \int_0^T\frac{ds}{\xi_s} & \int_0^T\frac{ds}{\xi_s^2}.\end{pmatrix}.$$

En appliquant l'inégalité de Cauchy-Schwarz, on voit que dét$I_T\geq 0$ et $=0$ si et seulement si $1/\xi_s=1$ p.p. sur $[0,T]$ ce qui est impossible. D'où

$$\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}=I_T^{-1}\begin{pmatrix}\int_0^T\frac{d\xi_s}{\xi_s}\\ \int_0^T\frac{d\xi_s}{\xi_s^2}\end{pmatrix}.$$

Pour étudier le comportement asymptotique de $\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}$, on écrit $d\xi_t=(a_0\xi_t+b_0)dt+\sigma\xi_t dW_t$ dans la relation ci-dessus et l'on obtient:

$$\begin{pmatrix}\hat{a}-a_0\\ \hat{b}-b_0\end{pmatrix}=\sigma I_T^{-1}\begin{pmatrix}W_T\\ \int_0^T\frac{dW_s}{\xi_s}\end{pmatrix}. \tag{7.7}$$

Etudions la loi limite de cet estimateur dans les conditions de récurrence positive pour le modèle. Pour la dérive $b(x)=ax+b$ et le coefficient de diffusion $\sigma(x)=\sigma\, x$, l'intervalle $(\ell,r)=(0,+\infty)$, la fontion s vaut:

$$s(x)=x^{-2a/\sigma^2}\exp\left(2b/(\sigma^2 x),\ x>0.\right.$$

Elle vérifie $\int^{+\infty}s(x)dx=\infty$ si et seulement si $2a/\sigma^2\leq 1$ et $\int_0 s(x)dx=\infty$ si et seulement si $b>0$ ou $b=0$ et $2a/\sigma^2\geq 1$. La fonction m vaut:

$$m(x)=\exp\left[-2b/(\sigma^2x)\right]x^{2(a/\sigma^2-1)}.$$

Elle vérifie $\int_0^{+\infty}m(x)dx<\infty$ si et seulement si $b>0$ et $a<\sigma^2/2$. Les conditions de récurrence positive sont $b>0$ et $a<\sigma^2/2$ ce qui réduit l'espace des paramètres. La loi invariante π est la mesure $m(x)dx$ normalisée pour être une probabilité. On reconnaît une loi inverse Gamma, c'est-à-dire que π est la loi de $1/V$ où V suit la loi Gamma $G(1-(2a/\sigma^2),2b/\sigma^2)$.

On suppose désormais que les paramètres du processus observé satisfont $b_0 > 0$ et $a_0 < \sigma^2/2$. La matrice I_T/T converge p.s. vers

$$I = \begin{pmatrix} 1 & \int_0^{+\infty} x^{-1} d\pi_{(a_0,b_0)}(x) \\ \int_0^{+\infty} x^{-1} d\pi_{(a_0,b_0)}(x) & \int_0^{+\infty} x^{-2} d\pi_{(a_0,b_0)}(x). \end{pmatrix} = \begin{pmatrix} 1 & \mathbb{E}V_0 \\ \mathbb{E}V_0 & \mathbb{E}V_0^2 \end{pmatrix}$$

où V_0 est une v.a. de loi $G(1-(2a_0/\sigma^2), 2b_0/\sigma^2)$. Donc,

$$I = \begin{pmatrix} 1 & \frac{\sigma^2-2a_0}{2b_0} \\ \frac{\sigma^2-2a_0}{2b_0} & \frac{(\sigma^2-2a_0)(\sigma^2-a_0)}{b\sigma^2} \end{pmatrix}.$$

On a dét $I = \sigma^2(\sigma^2-2a_0)/4b_0^2 > 0$ d'après la condition $\sigma^2 > 2a_0$. D'autre part, d'après le théorème de limite centrale pour les intégrales stochastiques (Théorème 4.3.2), le vecteur $\frac{1}{\sqrt{T}}\begin{pmatrix} W_T \\ \int_0^T \frac{dW_s}{\xi_s} \end{pmatrix}$ converge en loi lorsque T tend vers l'infini vers la loi gaussienne $\mathcal{N}_2(0, I)$. On en déduit que $\sqrt{T}\begin{pmatrix} \hat{a} - a_0 \\ \hat{b} - b_0 \end{pmatrix}$ converge en loi vers $\mathcal{N}_2(0, \sigma^2 I^{-1})$.

- Modèle de Cox-Ingersoll-Ross: soit le modèle $d\xi_t = (a_0\xi_t + b_0)dt + \sigma\sqrt{\xi_t^+}dW_t, \xi_0 = x > 0$. Ce modèle est plus souvent écrit sous la forme $d\xi_t = k_0(\theta_0 - \xi_t)dt + \sigma\sqrt{\xi_t^+}dW_t$ (forme de l'équation dans l'article original de Cox-Ingersoll-Ross). Cependant, la première formulation où la dérive est linéaire par rapport aux paramètres est plus pratique pour calculer les estimateurs de (a_0, b_0). On peut ensuite revenir aux estimateurs des paramètres κ_0, θ_0 en posant $\hat{b} = \hat{\kappa}\hat{\theta}$, $-\hat{\kappa} = \hat{a}$.
  On cherche d'abord les conditions de récurrence positive sur $(\ell, r) = (0, +\infty)$. La dérivée de la fonction d'échelle s'écrit:

$$s(x) = \exp\left(-2\frac{a}{\sigma^2}x\right) x^{(-2\frac{b}{\sigma^2})}.$$

On a $\int^{+\infty} s(x)dx = +\infty$ si et seulement si $a < 0$ ou $a = 0, 2b/\sigma^2 \leq 1$, $\int_0 s(x)dx = +\infty$ si et seulement si $2b/\sigma^2 \geq 1$. La densité de la mesure de vitesse vaut:

$$m(x) = \exp\left(2\frac{a}{\sigma^2}x\right) x^{(2\frac{b}{\sigma^2}-1)}.$$

et l'on a: $\int_0^{+\infty} m(x)dx < +\infty$ si et seulement si $a < 0$ et $2b/\sigma^2 > 0$. La diffusion sera récurrente positive sur $(0, +\infty)$ si $a < 0$ et $2b/\sigma^2 \geq 1$. Avec ces restrictions de l'espace des paramètres, on sait que, p.s., $\xi_t > 0$ pour tout $t \geq 0$. De plus, l'unique loi invariante est la loi de densité proportionnelle à m(x) c'est-à-dire la loi Gamma $G(2\frac{b}{\sigma^2}, -2\frac{a}{\sigma^2})$. On peut écrire la log-vraisemblance:

$$\ell_T(a,b) = \frac{1}{\sigma^2}\left(\int_0^T (a + \frac{b}{\xi_s})d\xi_s - \frac{1}{2}\int_0^T \frac{(a\xi_s + b)^2}{\xi_s}ds\right).$$

L'estimateur du maximum de vraisemblance $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$ est solution de

$$I_T \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \int_0^T d\xi_s \\ \int_0^T \frac{d\xi_s}{\xi_s} \end{pmatrix}$$

où

$$I_T = \begin{pmatrix} \int_0^T \xi_s ds & T \\ T & \int_0^T \frac{ds}{\xi_s} \end{pmatrix}.$$

La matrice I_T/T converge p.s. quand $T \to +\infty$ vers

$$I = \begin{pmatrix} \mathbb{E}(X) & 1 \\ 1 & \mathbb{E}(X^{-1}) \end{pmatrix}$$

où X est une v.a. de loi $G(2\frac{b_0}{\sigma^2}, -2\frac{a_0}{\sigma^2})$. Donc, $\mathbb{E}(X) = -b_0/a_0$, $\mathbb{E}(X^{-1}) = -2a_0/(2b_0 - \sigma^2)$. On peut poursuivre l'étude comme dans l'exemple précédent: $\sqrt{T}\begin{pmatrix} \hat{a} - a_0 \\ \hat{b} - b_0 \end{pmatrix}$ converge en loi vers $\mathcal{N}_2(0, \sigma^2 I^{-1})$. Pour revenir aux paramètres κ, θ, on applique la méthode *delta*. Si lon écrit $(\kappa, \theta) = F(a, b)$et que l'on note J la matrice Jacobienne de F en (a_0, b_0), $\sqrt{T}\begin{pmatrix} \hat{a} - a_0 \\ \hat{b} - b_0 \end{pmatrix}$ converge en loi vers $\mathcal{N}_2(0, \sigma^2 J I^{-1} J')$

7.2 Commentaires bibliographiques

L'étude asymptotique des EMV dans divers cadres d'observations est traitée dans plusieurs ouvrages anciens et plus récents. L'approche par processus de vraisemblance, que nous n'abordons pas ici, est celle d'Ibragimov et Hasmin'minskii (1981). On trouvera dans Kutoyants (2004) et Hoepfner (2014) cette théorie avec l'accent mis sur les observations de processus et en particulier de diffusions. La méthode de contrastes est détaillée dans Dacunha-Castelle et Duflo (1983). Une version plus récente, plus détaillée et plus moderne se trouve dans van der Vaart (1998).

Dans le modèle de Cox-Ingersoll-Ross, l'étude des estimateurs du maximum de vraisemblance (qui n'est pas simple) a été traitée complètement (pas uniquement dans le cas récurrent positif) dans différents articles (Overbeck, 1998, Ben Alaya et Kebaier, 2012, 2013).

7.3 Estimateurs empiriques

Nous étudions toujours le modèle

$$d\xi_t = b(\xi_t, \theta_0)dt + \sigma(\xi_t)dB_t, \qquad \xi_0 = x_0$$

avec les hypothèses (H1)-(H2) du paragraphe 7.1.1. On dispose donc du théorème ergodique qui permet de construire des estimateurs empiriques, de la forme $\frac{1}{T}\int_0^T f(\xi_s)ds$, des moments $\int_\ell^r f(x)\pi_{\theta_0}(x)dx$ de la distribution stationnaire $\pi_{\theta_0}(x)dx$ dès que $\int_\ell^r |f(x)|\pi_{\theta_0}(x)dx < +\infty$.

Exemple: $d\xi_t = \theta_0\xi_t dt + dB_t,\ \xi_0 = 0, \theta_0 < 0, \quad \pi_{\theta_0}(x)dx = \mathcal{N}(0, 1/(2|\theta_0|)$. La fonction $f(x) = x^2$ vérifie $\int_{\mathbb{R}} x^2\pi_{\theta_0}(x)dx = 1/(2|\theta_0|$. Par suite,

$$\frac{1}{T}\int_0^T \xi_s^2 ds \to_{p.s.} -\frac{1}{2\theta_0}.$$

Donc,

$$\tilde{\theta}_T = -\frac{T}{2\int_0^T \xi_s^2 ds}$$

est un estimateur fortement consistant de θ_0. La question que nous étudions ici est celle d'obtenir un théorème de limite centrale pour $\tilde{\theta}_T - \theta_0$ convenablement normalisé.

7.3.1 Un théorème de limite centrale pour les estimateurs empiriques

Notre objectif est de montrer que, si $f : (\ell, r) \to \mathbb{R}$ vérifie $\int_\ell^r |f(x)|\pi_{\theta_0}(x)dx < +\infty$ et des hypothèses supplémentaires,

$$\sqrt{T}\left(\frac{1}{T}\int_0^T f(\xi_s)ds - \int_\ell^r f(x)\pi_{\theta_0}(x)dx\right) = \frac{1}{\sqrt{T}}\int_0^T g(\xi_s)ds,$$

où $g(x) = f(x) - \int_\ell^r f(x)\pi_{\theta_0}(x)dx$, converge en loi vers une loi gaussienne.

Dans ce qui suit, nous omettrons θ_0 dans les notations et poserons $b(x, \theta_0) = b(x), \pi_{\theta_0} = \pi, \ldots$. Rappelons que le générateur infinitésimal (formel) de (7.1) est défini sur l'ensemble des fonctions de classe C^2 sur (ℓ, r) par (voir (6.8)):

$$LF = \frac{\sigma^2}{2}F'' + bF', \quad F : (\ell, r) \to \mathbb{R}, \text{ de classe } C^2.$$

En utilisant la dérivée de la fonction d'échelle s (voir (6.9)) et la densité de la mesure de vitesse m (voir (6.11)), on a:

$$LF = \frac{1}{2m}\left(\frac{F'}{s}\right)'.$$

Enfin, la densité de la mesure invariante est égale à $\pi(x) = \frac{1}{M\sigma^2 s}$ avec $M = \int_\ell^r \pi(x)dx$.

Théorème 7.3.1. *Soit $g : (\ell, r) \to \mathbb{R}$ telle que $\int_\ell^r g^2(x)\pi(x)dx < +\infty$ et $\int_\ell^r g(x)\pi(x)dx = 0$. Supposons que l'on puisse trouver $F \in C^2((\ell, r))$ telle que*

$$LF = -g \quad \text{et} \quad V = \int_\ell^r (F'(x)\sigma(x))^2\pi(x)dx < +\infty,$$

alors:

$$\frac{1}{\sqrt{T}}\int_0^T g(\xi_s)ds \to_{\mathcal{L}} \mathcal{N}(0, V).$$

Avant de prouver le théorème, calculons F en résolvant $LF = -g$:

$$\frac{1}{2m}\left(\frac{F'}{s}\right)' = -g \Rightarrow \frac{F'}{s} = -2\int_{\ell}^{x} g(u)m(u)du$$

qui vérifie $F'(x)/s(x) \to 0$ quand $x \to \ell$ car g est m-intégrable et $F'(x)/s(x) \to 0$ quand $x \to r$ car $\int_{\ell}^{r} g(x)\pi(x)dx = 0$. Il n'est pas utile d'intégrer de nouveau pour obtenir F car F' suffit et on obtient:

$$V = 4\int_{\ell}^{r} \sigma^2(x)\left(\int_{\ell}^{x} g(u)m(u)du\right)^2 dx = 4M\int_{\ell}^{r} s(x)\left(\int_{\ell}^{x} g(u)\pi(u)du\right)^2 dx. \tag{7.8}$$

Remarque 7.3.1. *La convergence de l'intégrale définissant V n'est pas immédiate car nous sommes dans le cas où*

$$\int_{\ell} s(x)dx = \int^{r} s(x)dx = \infty.$$

Mais, elle n'est pas impossible car

$$A_g(x) := \int_{\ell}^{x} g(u)\pi(u)du = -\int_{x}^{r} g(u)\pi(u)du \tag{7.9}$$

tend vers 0 quand x tend vers ℓ ou r. Pour certains modèles, elle est toujours vérifiée. Si F et LF appartiennent à $\mathbb{L}^2(\pi)$, avec $F'(x)/s(x) \to 0$ quand $x \to \ell, r$, on a $\int_{\ell}^{r} LF(x)\pi(x)dx = 0$ et on peut montrer que:

$$\int_{\ell}^{r} \left(F'(x)\sigma(x)\right)^2 \pi(x)dx = -2\int_{\ell}^{r} LF(x)\, F(x)\pi(x)dx.$$

Démontrons cette égalité dans le cas où F est de classe C^2 à support compact dans (ℓ, r). Rappelons que $LF = (F'/s)'/2m$ où s, m désignent respectivement la dérivée de la fonction d'échelle et la densité de la mesure de vitesse, $\pi = m/M$ la densité invariante. On a, après intégration par parties:

$$\langle LF, F\rangle_{\pi} = \int \frac{1}{2m}(\frac{F'}{s})' \times F \times \frac{m}{M}dx = -\frac{1}{2M}\int \frac{F'}{s} \times F'dx = -\frac{1}{2}\int (F'\sigma)^2\pi(x)dx.$$

Pour plus de détails, nous renvoyons à Genon-Catalot, Jeantheau et Larédo (2000) ainsi qu'aux références données dans cet article. Dans cet article, les propriétés du générateur infinitésimal opérant sur l'espace $\mathbb{L}^2(\pi)$ sont étudiées en détail.

Preuve du Théorème 7.3.1 Par la formule d'Ito,

$$F(\xi_T) = F(\xi_0) + \int_0^T LF(\xi_s)ds + \int_0^T \sigma(\xi_s)F'(\xi_s)dB_s.$$

D'après les théorèmes limites du paragraphe 6.3.6, on a $\xi_T \to_{\mathcal{L}} \pi$, ce qui implique $T^{-1}(F(\xi_T) - F(\xi_0)) \to 0$ en probabilité. Par suite,

$$-\frac{1}{\sqrt{T}}\int_0^T LF(\xi_s)ds = \frac{1}{\sqrt{T}}\int_0^T \sigma(\xi_s)F'(\xi_s)dB_s + o_P(1).$$

Si $V < +\infty$, $T^{-1}\int_0^T \sigma^2(\xi_s)(F'(\xi_s))^2 ds \to_{p.s.} V$ ce qui implique

$$\frac{1}{\sqrt{T}}\int_0^T \sigma(\xi_s)F'(\xi_s)dB_s \to_{\mathcal{L}} \mathcal{N}(0,V),$$

d'où le résultat. $\square$

On peut énoncer le corollaire suivant.

Corollaire 7.3.1. *Si $g_1,\ldots,g_k$ vérifient les hypothèses du théoréme précédent, alors*

$$\frac{1}{\sqrt{T}}(\int_0^T g_1(\xi_s)ds,\ldots,\int_0^T g_k(\xi_s))'_{\mathcal{L}} \to \mathcal{N}_k(0,V)$$

où $V=(V_{ij})_{1\le i,j\le k}$ et $V_{ij}=V(g_i,g_j)=4M\int_\ell^r s(x)A_{g_i}(x)A_{g_j}(x)dx$ (cf (7.9)), à condition que la matrice V soit définie positive.

Preuve du corollaire En appliquant l'inégalité de Schwarz, on voit que $V_{ii}+V_{jj}<\infty$ implique que V_{ij} est bien défini. Pour obtenir le corollaire, il suffit d'appliquer le théorème à $g=\sum_{i=1}^k \alpha_i g_i$, $\alpha_1,\ldots,\alpha_k\in\mathbb{R}$. $\square$

Ainsi, pour appliquer le théorème 7.3.1 à $\int_0^T g(\xi_s)ds$, il faut résoudre $LF=-g$. Dans certains exemples, cette résolution est immédiate.

Exemple Considérons le modèle $d\xi_t=\alpha(\beta-\xi_t)dt+\sigma(\xi_t)dW_t$. Supposons que $\sigma^2(x)\le K(1+x^2)$, que le modèle est récurrent positif sur un intervalle (ℓ,r) de distribution stationnaire π avec $\int_\ell^r x^2 d\pi(x)<+\infty$. Si ξ_0 suit la loi π (modèle en régime stationnaire), on peut écrire:

$$\xi_t=\xi_0+\alpha\int_0^t(\beta-\xi_s)ds+\int_0^t\sigma(\xi_s)dW_s.$$

Avec notre hypothèse,

$$\mathbb{E}(\int_0^t \sigma^2(\xi_s)ds)=t\mathbb{E}(\sigma^2(\xi_0))\le Kt(1+\int_\ell^r x^2d\pi(x))<+\infty.$$

Posons $m=\mathbb{E}(\xi_0)=\int_\ell^r xd\pi(x)$. On obtient, en prenant l'espérance de ξ_t, pour tout t, $m=m+\alpha(\beta-m)t$ ce qui implique $m=\beta$. Considérons alors la fonction $g(x)=x-\beta$. Elle vérifie puisque $g'=1, g''=0$, $Lg(x)=\alpha(\beta-x)=-\alpha g(x)$. Par suite, on peut prendre $F=\alpha^{-1}g$. La variance asymptotique de $T^{-1/2}\int_0^T g(\xi_s)ds$ est égale à $\alpha^{-2}\int_\ell^r\sigma^2(x)d\pi(x)$.

Cet exemple contient par exemple les cas $\sigma(x)=1$ (processus d'Ornstein-Uhlenbeck), $\sigma(x)=\sqrt{x}$ (processus de C.I.R.), $\sigma(x)=x$ (processus bilinéaire). La fonction $g(x)=(x-\beta)$ est d'intégrale nulle par rapport à la loi invariante. Elle est une fonction propre de l'opérateur L.

7.3.2 Moments empiriques et moments de la distribution stationnaire

Les résultats précédents permettent notamment d'obtenir des estimateurs des moments de la distribution stationnaire $\pi_{\theta_0}(x)dx$. Supposons que $\int_\ell^r |x|^\gamma \pi_{\theta_0}(x)dx < +\infty$ et posons $f_p(x, \theta_0) = x^p - \mu(p, \theta_0)$ où $\mu(p, \theta_0) = \int_\ell^r x^p \pi_{\theta_0}(x)dx, p \leq \gamma$. Alors le vecteur

$$(\hat{\mu}(p) = \frac{1}{T} \int_0^T \xi_s^p ds, 1 \leq p \leq \gamma)$$

vérifie le théorème:

Théorème 7.3.2. *Supposons que $V_{\theta_0}(f_p(., \theta_0)) < +\infty$ pour $1 \leq p \leq \gamma$. Alors,*

$$\sqrt{T}(\hat{\mu}(p) - \mu(p, \theta_0), 1 \leq p \leq \gamma) \to_{\mathcal{L}} \mathcal{N}_\gamma(0, \Sigma(\theta_0))$$

où $\Sigma(\theta_0) = (V_{\theta_0}(f_i(., \theta_0), f_j(., \theta_0)))_{1 \leq i,j \leq \gamma}$.

7.4 Références bibliographiques

1. Ben Alaya, M. et Kebaier, A. (2012). Parameter Estimation for the Square-Root Diffusion: Ergodic and Nonergodic cases. *Stochastic models.* **28**, 609-634.

2. Ben Alaya, M. et Kebaier, A. (2013). Asymptotic Behavior of the Maximum Likelihood Estimator for Ergodic and non ergodic square-root diffusions. *Stochastic Analysis and Applications*, **31**, 552-573.

3. Dacunha-Castelle, D., Duflo, M., (1983). *Probabilités et statistiques, 2. Problèmes à temps mobiles*, Masson, Paris.

4. Genon-Catalot, V. Jeantheau, T. et Larédo, C. (2000). Stochastic volatility models as Hidden Markov models and statistical applications. *Bernoulli* **6**, 1051-1079.

5. Hoepfner, R. (2014). *Asymptotic statistics with a view to stochastic processes.* Walter de Gruyter GmbH, Berlin/Boston.

6. Hutton J.E., Nelson, P.I. (1984).Interchanging the order of differentiation and stochastic integration, *Stoch. Proc Appl.* **18**, 371-377.

7. Ibragimov, I.A., Has'minskii, R.Z. (1981). *Statistical Estimation, Asymptotic theory.* Springer-Verlag New York, Heidelberg, Berlin.

8. Karandikar, R.L., (1983). Interchanging the order of stochastic integration and ordinary differentiation, *Sankhia, The Indian Journal of Statistics* **45** Series **A**, 120-124.

9. Kutoyants, Yu. A. (2004). *Statistical inference for ergodic diffusion processes.* Springer-verlag, London.

10. Overbeck, L. (1998). Estimation for continuous branching processes. *Scand. J. Statist.*, **25**, 111-126.

11. D. Revuz et Marc Yor (1991). *Continuous martingales and Brownian motion*, Springer, Berlin.

12. van der Vaart, A.W. (1998) *Asymptotic statistics.* Cambridge Unversity Press, Cambridge, New York.

Chapter 8

Sujets d'examens

Ce chapitre contient l'essentiel des sujets d'examens posés pour ce cours de Statistique des diffusions au master 2 de Modélisation Aléatoire en Economie et Finance (anciennement D.E.A. de Statistique et modèles mathématiques en économie et finance).

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 30 mai 2000

Durée 3H - Documents autorisés

Exercice 1 : Soit $\xi_t = \theta_0 t + \sigma_0 W_t$, $t \geq 0$, où $\theta_0 \in \mathbb{R}, \sigma_0 > 0$ sont des paramètres inconnus et (W_t) est un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On observe la trajectoire (ξ_t) aux instants $t_i = i\Delta, i = 1, \ldots, n$, avec $\Delta > 0$ connu.

1. Calculer la vraisemblance de l'observation $(\xi_{t_i}, i = 1, \ldots, n)$ et les estimateurs du maximum de vraisemblance $\hat{\theta}, \hat{\sigma}^2$ de θ_0, σ_0^2 basés sur cette observation (on vérifiera bien qu'en $\hat{\theta}, \hat{\sigma}^2$, la vraisemblance passe par un maximum).

2. On suppose que $\Delta = \Delta_n$ tend vers 0 et $n\Delta_n$ tend vers l'infini lorsque n tend vers l'infini. Quelle est la loi de $(n\Delta_n)^{1/2}(\hat{\theta} - \theta_0)$? Démontrer que le couple
$$((n\Delta_n)^{1/2}(\hat{\theta} - \theta_0), n^{1/2}(\hat{\sigma}^2 - \sigma_0^2))$$
converge en loi vers la loi $\mathcal{N}(0, \sigma_0^2) \otimes \mathcal{N}(0, 2\sigma_0^4)$.
(On pourra introduire les v.a. $\varepsilon_i^n = \dfrac{W_{t_i^n} - W_{t_{i-1}^n}}{\Delta_n^{1/2}}$ où $t_i^n = i\Delta_n$).

3. On pose
$$\tilde{\sigma}^2 = \frac{1}{n\Delta_n} \sum_{i=1}^{n} (\xi_{t_i^n} - \xi_{t_{i-1}^n})^2.$$
Déterminer une condition sur n et Δ_n pour que $n^{1/2}(\hat{\sigma}^2 - \tilde{\sigma}^2)$ converge en probabilité vers 0. En déduire, en supposant vérifée cette condition, la loi limite de
$$((n\Delta_n)^{1/2}(\hat{\theta} - \theta_0), n^{1/2}(\tilde{\sigma}^2 - \sigma_0^2)).$$

4. Si $\Delta_n = n^{-\alpha}$, $\alpha > 0$, commenter les vitesses obtenues pour $\hat{\theta}, \hat{\sigma}^2$ et $\hat{\theta}, \tilde{\sigma}^2$ (tenant compte de la condition obtenue en 3)).

Exercice 2 : Soit $(W_t, t \geq 0)$ est un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$, η une v.a. rélle définie sur Ω, indépendante de $(W_t, t \geq 0)$. On considère l'équation différentielle stochastique :
$$d\xi_t = -\alpha \xi_t dt + \sigma(\theta + \xi_t^2)^{1/2} dW_t, \quad \xi_0 = \eta,$$
où $\alpha \in \mathbb{R}, \sigma > 0, \theta > 0$.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution adapté à la filtration $\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0$.

2. On suppose que $\theta = 1$ et α, σ inconnus.On observe la trajectoire (ξ_t) aux instants $t_i^n = i\dfrac{T}{n}$, avec $i = 0, 1, \ldots, n$, et $T > 0$ fixé connu.

   (a) On pose
   $$\hat{\sigma}^2 = \frac{1}{T} \sum_{i=1}^{n} \frac{(\xi_{t_i^n} - \xi_{t_{i-1}^n})^2}{1 + \xi^2_{t_{i-1}^n}}.$$
   Montrer que $\hat{\sigma}^2$ est un estimateur consistant (faiblement) de σ^2.

   (b) Soit $F(x) = \log(x + \sqrt{1 + x^2})$ et $Y_t = F(\xi_t)$. Montrer que (Y_t) est solution d'une équation différentielle stochastique que l'on déterminera.

   (c) On pose
   $$\tilde{\sigma}^2 = \frac{1}{T} \sum_{i=1}^{n} (Y_{t_i^n} - Y_{t_{i-1}^n})^2.$$
   Montrer que $\tilde{\sigma}^2 \to \sigma^2$ en probabilité et que $n^{1/2}(\tilde{\sigma}^2 - \sigma^2)$ converge en loi vers la loi $\mathcal{N}(0, 2\sigma^4)$.

3. On ne suppose plus $\theta = 1$, mais l'on suppose les trois paramètres α, θ, σ inconnus. On dispose de la même observation de la trajectoire qu'en 2).

   (a) Donner, pour tout $k \in \mathbb{N}$, la limite en probabilité lorsque n tend vers l'infini, de
   $$\sum_{i=1}^{n} \xi^k_{t_{i-1}^n} (\xi_{t_i^n} - \xi_{t_{i-1}^n})^2.$$

   (b) Proposer des estimateurs consistants de θ et σ^2.

Exercice 3 : (Le modèle est le même que dans l'exercice 2). On observe, pour $t \in [0, T]$,le processus (ξ_t) défini par :

$$d\xi_t = -\alpha_0\, \xi_t dt + \sigma(1 + \xi_t^2)^{1/2} dW_t, \quad \xi_0 = 0,$$

où α_0 est la vraie valeur inconnue du paramètre et σ est connu.

1. Ecrire la vraisemblance de l'observation $(\xi_t, t \in [0, T])$ et calculer l'estimateur du maximum de vraisemblance $\hat{\alpha}_T$ de α_0 basé sur cette observation.

2. Déterminer les valeurs de α_0 pour lesquelles le processus (ξ_t) est récurrent positif. Calculer, dans ce cas, la distribution stationnaire du processus (ξ_t).

3. On suppose vérifiées les conditions de récurrence positive. Montrer que $\hat{\alpha}_T$ converge p.s. vers α_0 et que $\sqrt{T}(\hat{\alpha}_T - \alpha_0)$ converge en loi, lorsque T tend vers l'infini, vers une loi limite que l'on déterminera.

   (On pourra introduire les intégrales $M(a) = \displaystyle\int_R \frac{dx}{(1 + x^2)^a}$ et utiliser la relation $M(a+1) = \frac{2a-1}{2a} M(a)$ lorsque ces quantités sont bien définies).

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 7 septembre 2000

Durée 3H - Documents autorisés

I : Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité, $(W_t, t \geq 0)$ un mouvement brownien standard défini sur Ω et η une variable aléatoire réelle strictement positive définie sur Ω, indépendante de (W_t). On pose $\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0$ et l'on considère l'équation différentielle stochastique (e.d.s.):

$$d\xi_t = \alpha(\beta - \xi_t)dt + \sigma\xi_t dW_t, \quad \xi_0 = \eta,$$

où $\beta \geq 0, \alpha \in \mathbb{R}, \sigma > 0$.

1. Vérifier que W est un $P - (\mathcal{F}_t)$ mouvement brownien et que l'e.d.s. admet un unique processus solution adapté à $(\mathcal{F}_t)$.

2. Déterminer les conditions sur α, β, σ^2 pour que le processus (ξ_t) soit récurrent positif sur $(0, +\infty)$. Calculer alors sa distribution stationnaire que l'on notera π. (On pourra poser $\lambda = \frac{2\alpha\beta}{\sigma^2}$, $a = \frac{2\alpha}{\sigma^2} + 1$. On rappelle que la fonction $\Gamma(a) = \displaystyle\int_0^{+\infty} x^{a-1} \exp(-x)\, dx$ vérifie la relation $\Gamma(a+1) = a\,\Gamma(a)$ pour tout $a > 0$.)

3. Déterminer les conditions sur α, β, σ^2 pour que $\displaystyle\int_0^{+\infty} x^2\, \pi(dx) < \infty$. Calculer

$$m = \int_0^{+\infty} x\, \pi(dx), \quad c = \int_0^{+\infty} x^2\, \pi(dx), \quad v = \int_0^{+\infty} (x - \beta)^2\, \pi(dx).$$

   Dans toute la suite du problème (I, II, III,IV), on suppose vérifiées les conditions sur α, β, σ^2 obtenues dans les questions précédentes. De plus, on suppose que η suit la loi π.

4. Calculer $E(\xi_t)$, $E(\xi_t^2)$ et $E(\xi_t - \beta)^2)$.

5. En posant $\xi_t = \beta + \exp(-\alpha t)\, C_t$, calculer dC_t et en déduire que:

$$\xi_t - \beta = \exp(-\alpha t)\left(\xi_0 - \beta + \sigma M_t\right),$$

   où (M_t) est une martingale relative à $P - (\mathcal{F}_t)$, de carré intégrable.

6. Calculer $E(\xi_t - \xi_0)^2$.

7. En déduire que, pour tous $s, t \in [0, 1]$, on a $E(\xi_t - \xi_s)^2 \leq K|t - s|$, où K est une constante ne dépendant pas de s, t.

II : Dans cette partie, on suppose que $\alpha = 1 = \sigma$ et l'on se propose d'estimer le paramètre inconnu β à partir de l'observation discrétisée $(\xi_{i\Delta_n}, i = 0, \dots, n)$, où le pas de discrétisation Δ_n est tel que

$$\Delta_n \to 0, \quad n\Delta_n \to +\infty, \quad n\Delta_n^2 \to 0.$$

On étudie l'estimateur $\tilde{\beta}_n = \frac{1}{n} \sum_{i=0}^{n-1} \xi_{i\Delta_n}$

1. Soit $Z_n = \frac{1}{n\Delta_n} \int_0^{n\Delta_n} \xi_s ds$. Montrer que, lorsque n tend vers l'infini,

$$E(\sqrt{n\Delta_n} |Z_n - \tilde{\beta}_n|) \to 0$$

2. Montrer que Z_n converge p.s. vers β lorsque n tend vers l'infini.

3. Montrer que

$$\sqrt{n\Delta_n}(Z_n - \beta)$$

   converge en loi, lorsque n tend vers l'infini, vers une loi limite que l'on précisera.

4. En déduire les propriétés asymptotiques de l'estimateur $\tilde{\beta}_n$ de β.

III : On suppose les trois paramètres α, β, σ inconnus.

1. Montrer que, pour tout $\Delta > 0$ et tout $i \geq 0$, on a la relation:

$$\xi_{(i+1)\Delta} - \beta = \exp(-\alpha\Delta)(\xi_{i\Delta} - \beta) + \rho\, \varepsilon_{i+1},$$

   où $(\varepsilon_i, i \geq 1)$ est une suite de v.a. telle que $E\varepsilon_i = 0$, $E\varepsilon_i^2 = 1$, et pour $i \neq j$, $E\varepsilon_i\varepsilon_j = 0$, et où ρ est une fonction de σ, α, β à déterminer.

2. Proposer une méthode possible d'estimation des trois paramètres β, $a = \exp(-\alpha\Delta)$, ρ^2, basée sur l'observation $(\xi_{i\Delta}, i = 0, \dots, n)$. (On utilisera la relation de la question précédente). En déduire les expressions d'estimateurs possibles pour ces paramètres (sans en étudier le comportement).

IV On pose $Y_t = \log \xi_t$.

1. Déterminer l'e.d.s. vérifiée par (Y_t).

2. On observe ξ_t aux instants $t_i^n = i/n$, $i = 0, \dots, n$, avec $n \to +\infty$. Donner un estimateur consistant et asymptotiquement normal de σ^2, basé sur cette observation.

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 2 mai 2002

Durée 3H - Documents autorisés

Problème :

- I Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité, $(W_t, t \geq 0)$ un mouvement brownien standard sur Ω, η une v.a. réelle définie sur Ω, indépendante de $(W_t, t \geq 0)$. On considère l'équation différentielle stochastique :

$$dX_t = -\theta_0 b(X_t)dt + \sigma_0 dW_t, \quad X_0 = \eta, \tag{8.1}$$

où $\theta_0 \in \mathbb{R}, \sigma_0 > 0$, $b : \mathbb{R} \to \mathbb{R}$ est une fonction non constante, de classe C^1 sur $\mathbb{R}$, bornée et de dérivée bornée. On pose $B(x) = \int_0^x b(u)\, du$.

1. Soit, pour $t \geq 0$, $\mathcal{F}_t = \sigma(\eta, W_s, s \leq t)$. Vérifier que l'équation différentielle stochastique (8.1) admet un unique processus solution adapté à la filtration $(\mathcal{F}_t, t \geq 0)$.
2. A quelles conditions le processus (X_t) est-il récurrent sur $\mathbb{R}$, récurrent positif sur $\mathbb{R}$. Exprimer, dans ce dernier cas, sa distribution stationnaire que l'on notera π_0.
3. On suppose la fonction b connue et $\eta = x_0$ déterministe connu. On suppose vérifiées les conditions de récurrence positive.
   (a) Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0 basé sur l'observation $(X_t, t \in [0, T])$, $T > 0$.
   (b) Montrer que, lorsque T tend vers l'infini, $\hat{\theta}_T$ converge p.s. vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi limite que l'on précisera.
   (c) On suppose que $b(x) = \frac{x}{1+x^2}$.
      i. Vérifier que b satisfait aux conditions de l'énoncé.
      ii. Exprimer les conditions de récurrence, de récurrence positive ainsi que la distribution stationnaire.
      iii. Sous les conditions de récurrence positive, calculer la loi limite de $\sqrt{T}(\hat{\theta}_T - \theta_0)$ lorsque T tend vers l'infini.

- II Dans cette partie, on suppose b connue, les conditions de récurrence positive vérifiées et que η suit la loi stationnaire π_0. On observe une discrétisation $(X_{i\Delta_n}, i = 0, \ldots n)$ de pas Δ_n tel que, lorsque n tend vers l'infini, Δ_n tend vers 0, $n\Delta_n$ tend vers l'infini et $n\Delta_n^2$ tend vers 0. On se propose d'étudier l'estimateur de θ_0 suivant:

$$\tilde{\theta}_n = -\frac{\sum_{i=0}^{n-1} b(X_{i\Delta_n})(X_{(i+1)\Delta_n} - X_{i\Delta_n})}{\sum_{i=0}^{n-1} \Delta_n\, b^2(X_{i\Delta_n})}$$

1. Montrer que, pour tout $t \geq 0$ et tout $h \in [0,1]$, $E(X_{t+h} - X_t)^2 \leq Ch$ où C est une constante ne dépendant ni de t ni de h.
2. Montrer que, pour toute fonction $f : R \to R$ lipschitzienne,
$$D_n(f) = \frac{1}{\sqrt{n\Delta_n}} \left(\sum_{i=0}^{n-1} \Delta_n \ f(X_{i\Delta_n}) - \int_0^{n\Delta_n} f(X_s)ds \right)$$
converge en probabilité vers 0 lorsque n tend vers l'infini.
3. Soit
$$Z_n = \frac{1}{n\Delta_n} \sum_{i=0}^{n-1} b(X_{i\Delta_n})(W_{(i+1)\Delta_n} - W_{i\Delta_n}).$$
Montrer que Z_n converge en probabilité vers 0 et que $\sqrt{n\Delta_n} Z_n$ converge en loi vers une loi limite que l'on précisera.
4. Montrer que $\tilde{\theta}_n$ converge en probabilité vers θ_0 et que $\sqrt{n\Delta_n}(\tilde{\theta}_n - \theta_0)$ converge en loi vers une loi limite que l'on précisera.

- III Dans cette partie, θ_0 est supposé connu.On suppose toujours les conditions de récurrence positive vérifiées et que η suit la loi π_0. On considère l'estimateur de σ_0^2 défini par:
$$\tilde{\sigma}_n^2 = \frac{1}{n\Delta_n} \sum_{i=1}^{n} (X_{(i+1)\Delta_n} - X_{i\Delta_n} + \theta_0 \Delta_n b(X_{i\Delta_n}))^2.$$
  1. Montrer que $\tilde{\sigma}_n^2$ converge en probabilité vers σ_0^2 lorsque n tend vers l'infini.
  2. Montrer que $\sqrt{n}(\tilde{\sigma}_n^2 - \sigma_0^2)$ converge en loi vers une loi limite que l'on précisera.

- IV On suppose θ_0 et σ_0 inconnus avec les mêmes hypothèses que dans les parties II et III. On considère pour estimer σ_0^2 l'estimateur
$$\bar{\sigma}_n^2 = \frac{1}{n\Delta_n} \sum_{i=1}^{n} (X_{(i+1)\Delta_n} - X_{i\Delta_n} + \tilde{\theta}_n \Delta_n b(X_{i\Delta_n}))^2.$$
Montrer que $\sqrt{n}(\tilde{\sigma}_n^2 - \bar{\sigma}_n^2)$ converge en probabilité vers 0 et en déduire la limite en loi de $\sqrt{n}(\bar{\sigma}_n^2 - \sigma_0^2)$.

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 18 mai 2004

Durée 3H - Documents autorisés

Exercice 1: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$, η une v.a. réelle définie sur Ω, indépendante de $(W_t, t \geq 0)$. On considère l'équation différentielle stochastique :

$$dX_t = -\theta X_t dt + (1 + X_t^2)^{1/2} dW_t, \quad X_0 = \eta,$$

où $\theta \in \mathbb{R}$.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution adapté à la filtration $(\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0)$.

2. (a) Déterminer les conditions sur θ pour que le modèle soit récurrent sur $\mathbb{R}$.

   (b) Si $\theta = -1/2$, calculer dY_t pour $Y_t = \text{Argsh}(X_t)$ et en déduire X_t (On rappelle: $\text{Argsh}(x) = \int_0^x \frac{dt}{(1+t^2)^{1/2}}$).

   (c) Déterminer les conditions sur θ pour que le modèle soit récurrent positif sur $\mathbb{R}$. Calculer, dans ce cas, la distribution stationnaire que l'on notera π_θ.

   On supposera désormais vérifiées les conditions de récurrence positive et que la v.a. initiale η suit la loi π_θ.

3. Quelles sont les propriétés du processus (X_t).

4. On suppose que l'on observe $(X_t, t \in [0, T])$ pour estimer θ. On considère l'estimateur

$$\hat{\theta}_T = -\int_0^T \frac{X_t}{1 + X_t^2} dX_t \left(\int_0^T \frac{X_t^2}{1 + X_t^2} dt \right)^{-1}.$$

   (a) Montrer que $\hat{\theta}_T$ converge p.s. vers θ lorsque T tend vers l'infini.

   (b) Déterminer la loi limite, lorsque T tend vers l'infini de $\sqrt{T}(\hat{\theta}_T - \theta)$.

5. (a) Déterminer les conditions sur θ pour que la distribution stationnaire π_θ admette un moment d'ordre 8. Ces conditions sont supposées désormais vérifiées.

   (b) Calculer, en utilisant les différentielles stochastiques de X_t, X_t^2, X_t^4, $E(X_t)$, $E(X_t^2)$, $E(X_t^4)$.

   (c) Montrer que, pour tout $t \geq 0$ et tout $h \in [0, 1]$, $E(X_{t+h} - X_t)^2) \leq C\, h$, où C est une constante.

   (d) Montrer que $\frac{1}{T} \int_0^T X_s^2 ds$ converge p.s. lorsque T tend vers l'infini vers une limite $l(\theta)$ que l'on calculera.

(e) Calculer la limite en loi de $\sqrt{T}(\frac{1}{T}\int_0^T X_s^2 ds - l(\theta)$ lors que T tend vers l'infini.

(f) En déduire un estimateur $\tilde{\theta}_T$ de θ, basé sur l'observation $(X_t, t \in [0,T])$ et démontrer sa consistance et sa normalité asymptotique.

(g) Soit Δ_n un pas de discrétisation. On observe, pour estimer θ, les variables $(X_{i\Delta_n}, i = 0, 1, \ldots, n)$. Utiliser les questions précédentes pour construire un estimateur consistant de θ basé sur l'observation $(X_{i\Delta_n}, i = 0, 1, \ldots, n)$, dans le cas où Δ_n tend vers 0 et $n\Delta_n$ tend vers l'infini.

(h) Rajouter une condition sur Δ_n pour que cet estimateur soit asymptotiquement normal et de même loi asymptotique que $\tilde{\theta}_T$.

Exercice 2:

Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère l'équation différentielle stochastique (analogue à celle de l'exercice 1):

$$dX_t = -\theta X_t dt + \sigma(1 + X_t^2)^{1/2} dW_t, \quad X_0 = 0.$$

On souhaite ici estimer (θ, σ^2) à partir de l'"observation $(X_{i\Delta}, i = 1, \ldots, n)$ où Δ est un pas de discrétisation.

1. Ecrire le schéma d'Euler associé à cette équation différentielle stochastique, ainsi que sa vraisemblance.

2. Déduire de la question précédente des estimateurs $(\hat{\theta}_n, \hat{\sigma}_n^2)$ de (θ, σ^2). Que pouvez-vous dire de l'estimateur $\hat{\theta}_n$ ainsi obtenu.

3. On suppose $\Delta = \Delta_n = T/n$ avec T fixé et $n \to +\infty$.
   - Que pouvez-vous dire de l'estimation de (θ, σ^2) dans ce cadre asymptotique?
   - Proposer un estimateur consistant de σ^2.
   - en utilisant la transformation $Y_t = \mathrm{Argsh}(X_t)$ (cf exercice 1), proposer un estimateur consistant et asymptotiquement normal de σ^2.

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 7 septembre 2004

Durée 3H - Documents autorisés

Problème: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$, $x_0 > 0$ fixé. On considère le processus (X_t) défini par l'équation différentielle stochastique :

$$dX_t = (aX_t + b)dt + \sigma X_t dW_t, \quad X_0 = x_0, \tag{8.2}$$

où $\sigma > 0$, $a < \frac{\sigma^2}{2}$ et $b > 0$.

1. Montrer que le processus (X_t) est récurrent positif sur $]0, +\infty[$ et que sa distribution stationnaire π est la loi de $\frac{1}{V}$ où V est une v.a. de loi Gamma de paramètres $\alpha = 1 - \frac{2a}{\sigma^2}$ et $\lambda = \frac{2b}{\sigma^2}$. En déduire
$$\int_0^{+\infty} x^{-1}\pi(dx), \quad \int_0^{+\infty} x^{-2}\pi(dx).$$
(Rappel: la loi Gamma de paramètres α et λ est la loi de densité
$$\gamma_{\alpha,\lambda}(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x}\, x^{\alpha-1}\, 1_{(x>0)}.)$$

2. On observe le processus (X_t) sur l'intervalle $[0, T]$ aux instants $t_i = i\frac{T}{n}$, $i = 1, \ldots, n$, x_0 est connu. On se propose d'étudier l'estimateur de σ^2 défini par:
$$\tilde{\sigma}^2 = \frac{1}{T}\sum_{i=1}^{n}(Y_{t_i} - Y_{t_{i-1}})^2$$
où $Y_t = \log X_t$. On supposera pour cette étude que T est fixé et que n tend vers l'infini.
   (a) Montrer que (Y_t) est solution d'une équation différentielle stochastique de la forme
   $$dY_t = f(Y_t)dt + \sigma dW_t.$$
   (b) On pose
   $$A_n = \sqrt{n}\sigma^2\left(\frac{1}{T}\sum_{i=1}^{n}(W_{t_i} - W_{t_{i-1}})^2 \;-\; 1\right).$$
   Montrer que A_n converge en loi lorsque n tend vers l'infini et préciser sa loi limite.

(c) On pose $S_n = \sqrt{n}(\tilde{\sigma}^2 - \sigma^2)$. Montrer que $R_n = S_n - A_n$ converge vers 0 en probabilité. (Il conviendra de décomposer R_n en la somme de plusieurs termes).

(d) En déduire la limite en loi de S_n.

3. On suppose dans cette question $\sigma = \sqrt{2}$, $a = 0$. Pour estimer b, on observe le processus (X_t) sur tout l'intervalle $[0, T]$.

   (a) Calculer l'estimateur du maximum de vraisemblance $\hat{b}$ de b.

   (b) Montrer que cet estimateur est consistant lorsque T tend vers l'infini.

   (c) Montrer que $\sqrt{T}(\hat{b} - b)$ converge en loi lorsque T tend vers l'infini, vers une loi gaussienne $\mathcal{N}(0, \gamma^2)$. Calculer γ^2.

   (d) On considère l'estimateur
   $$\tilde{b} = \left(\frac{1}{T} \int_0^T \frac{ds}{X_s} \right)^{-1}$$
   Montrer que l'estimateur $\tilde{b}$ est consistant et que $\sqrt{T}(\tilde{b} - b)$ converge en loi, lorsque T tend vers l'infini, vers une loi limite que l'on précisera.

   (e) Lequel des deux estimateurs $\hat{b}$ et $\tilde{b}$ est, à votre avis, préférable et pourquoi.

   (f) Résoudre l'équation différentielle stochastique $dU_t = \sqrt{2} U_t dW_t, U_0 = 1$. Poser $Z_t = \frac{X_t}{U_t}$ et déterminer l'équation différentielle stochastique vérifiée par (Z_t). En déduire (X_t) en fonction de (U_t).

4. Dans cette question, on suppose les trois paramètres a, b, σ^2 inconnus. Pour les estimer, on observe le processus (X_t) aux instants $t_i = i\Delta$, $i = 1, \ldots, n$ où Δ est un pas de discrétisation donné.

   (a) Ecrire le schéma d'Euler associé à l' équation différentielle (8.2), ainsi que sa vraisemblance.

   (b) Déduire de la question précédente des estimateurs de (a, b, σ^2).

   (c) Proposer différents cadres asymptotiques pour étudier les estimateurs de la question précédente. Suivant le cas, que pouvez-vous dire de leur comportement asymptotique (preuve non demandée).

Université de Paris 7
D.E.A. de Statistique et modèles mathématiques en économie et finance.

Examen de Statistique des diffusions - 26 mai 2005

Durée 3H - Documents autorisés.
Les parties III et IV peuvent être traitées indépendamment. Barême indicatif: I+II+III ou I+II+IV sur 20, le reste en plus.

Problème: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = -\theta_0 \frac{\xi_t}{\sqrt{1+\xi_t^2}} dt + dW_t, \quad X_0 = \eta, \tag{8.3}$$

où $\theta_0 \in \mathbb{R}R$ est un paramètre inconnu et η est une variable aléatoire réelle indépendante de (W_t).

I

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω, adapté à la filtration $(\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0)$.
2. Pour quelles valeurs de θ_0 le processus (ξ_t) est-il récurrent sur R? Récurrent positif? Donner dans ce cas l'expression de sa distribution stationnaire π_{θ_0}.
3. On suppose que $\theta_0 > 0$ et que η suit la loi π_{θ_0}.
   - Quelle(s) propriété(s) le processus solution (ξ_t) vérifie-t-il.
   - Utiliser la formule d'Ito appliquée à ξ_t^2 pour montrer que

$$\int_R \frac{x^2}{\sqrt{1+x^2}} \pi_{\theta_0}(dx) = \frac{1}{2\theta_0}. \tag{8.4}$$

Dans toute la suite, on suppose que $\eta = 0$.

II Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0, T]$.

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0.
2. On suppose $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, v(\theta_0))$. Donner, sans chercher à la calculer explicitement, l'expression de $v(\theta_0)$.

3. On suppose toujours $\theta_0 > 0$.

   - Utiliser (8.15) pour construire un estimateur fortement consistant de θ_0 que l'on notera $\tilde{\theta}_T$.
   - Utiliser la formule d'Ito appliquée à ξ_T^2 pour montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, \gamma(\theta_0))$. Donner, sans chercher à la calculer explicitement, l'expression de $\gamma(\theta_0)$.

4. Montrer que $1 \leq v(\theta_0) \leq \gamma(\theta_0)$. Lequel des deux estimateurs est préférable?

III On suppose toujours $\theta_0 > 0$ et l'on définit, pour $a > 0$,

$$T_a = \inf\{t \geq 0, \int_0^t \frac{\xi_s^2}{1+\xi_s^2} ds \geq a\}.$$

Au lieu d'observer (ξ_t) sur $[0, T]$, on l'observe sur l'intervalle $[0, T_a]$ et l'on étudie l'estimateur $\hat{\theta}_{T_a}$.

1. Vérifier que, pour tout $a > 0$, $T_a < +\infty$ p.s. et que $T_a \to +\infty$ p.s. lorsque $a \to +\infty$.

2. Vérifier que T_a/a converge p.s. vers $v(\theta_0)$ lorsque a tend vers l'infini.

3. En utilisant la martingale locale

$$M_t(\lambda) = \exp i\lambda \int_0^t \frac{\xi_s}{\sqrt{1+\xi_s^2}} dW_s + \lambda^2/2 \int_0^t \frac{\xi_s^2}{1+\xi_s^2} ds,$$

   déterminer la loi de la v.a. $\int_0^{T_a} \frac{\xi_s}{\sqrt{1+\xi_s^2}} dW_s$

4. En déduire la loi de $\sqrt{a}(\hat{\theta}_{T_a} - \theta_0)$.

IV On suppose que $\theta_0 = 0$. On se propose d'étudier les estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$.

1. Le processus (ξ_t) est-il récurrent positif?

2. On pose $Z_T = (1/T) \int_0^T \frac{\xi_s^2}{1+\xi_s^2} ds$. En considérant le changement de variables $s = uT$ dans Z_T et le mouvement brownien $(B_u^{(T)} = W_{uT}/\sqrt{T}, u \in [0,1])$, montrer que $E(Z_T)$ tend vers 1 lorsque T tend vers l'infini. En déduire que Z_T tend en probabilité vers 1.

3. Montrer que $\hat{\theta}_T$ converge p.s. vers 0 et que $\sqrt{T}\hat{\theta}_T$ converge en loi vers la loi gaussienne $\mathcal{N}(0, 1)$.

4. On pose $X_T = (1/T)\int_0^T \frac{\xi_s^2}{\sqrt{1+\xi_s^2}} ds$. En utilisant le même changement de variables $s = uT$ que ci-dessus, montrer que

$$D_T = X_T/\sqrt{T} - \int_0^1 |B_u^{(T)}| du$$

converge vers 0 en probabilité lorsque T tend vers l'infini (on pourra étudier $E(|D_T|)$).

5. En déduire que $\tilde{\theta}_T$ converge en probabilité vers 0 et que $\sqrt{T}\tilde{\theta}_T$ converge en loi vers une loi limite que l'on précisera.

Université de Paris 7
Master M2 Modélisation aléatoire.

Examen de Statistique des diffusions. 13 septembre 2005

Durée 3H - Documents autorisés.

I: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = \sqrt{\theta_0 + \xi_t^2}\, dW_t, \quad \xi_0 = x_0, \tag{8.5}$$

où $\theta_0 > 0$ est un paramètre inconnu et x_0 est un réel connu.

I

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω.
2. Le processus (ξ_t) est-il récurrent sur R? Récurrent positif? Donner dans ce cas l'expression de sa distribution stationnaire π_{θ_0}.
3. On souhaite estimer θ_0 à partir de l'observation de ξ_t aux instants $t_i = t_i^n = iT/n$, $i = 1, \ldots, n$, où $T > 0$ est fixé. On posera $\Delta_n = T/n$.
   - Quelle est la limite en probabilité quand n tend vers l'infini de $\sum_{i=1}^n (\xi_{t_i} - \xi_{t_{i-1}})^2$. En déduire un estimateur consistant de θ_0 qui s'exprime explicitement en fonction de l'observation $(\xi_{t_i}, i \leq n)$. De façon analogue, en considérant les variables aléatoires de la forme
   $$\sum_{i=1}^n \xi_{t_{i-1}}^{2p} (\xi_{t_i} - \xi_{t_{i-1}})^2,$$
   pour p entier positif, construire d'autres estimateurs consistants de θ_0.
   - On considère $Y_1, \ldots, Y_n$ n variables aléatoires telles que:
   $$Y_0 = x_0,\ Y_i - Y_{i-1} = \sqrt{\theta + Y_{i-1}^2}\, \sqrt{\Delta_n}\, \varepsilon_i, i = 1, \ldots, n,$$
   avec $\theta > 0$ et $\varepsilon_1, \ldots, \varepsilon_n$ indépendantes et de même loi gaussienne centrée réduite. Calculer la densité jointe $f(\theta, y_1, \ldots, y_n)$ de $(Y_1, \ldots, Y_n)$.
   - On pose $U_n(\theta) = -2\Delta_n \log f(\theta, \xi_{t_1}, \ldots, \xi_{t_n})$. Calculer la limite en probabilité lorsque n tend vers l'infini, de $U_n(\theta) - U_n(\theta_0)$ (que l'on exprimera en faisant apparaître la fonction $\varphi(u) = u - 1 - \log u, u > 0$).
   - En notant $K_T(\theta_0, \theta)$ la limite obtenue précédemment, montrer que $K_T(\theta_0, \theta) \geq 0$ p.s. et $K_T(\theta_0, \theta) = 0$ p.s. si et seulement si $\theta = \theta_0$.

- Quels estimateurs peut-on construire à l'aide de $U_n(\theta)$? Quelles propriétés peut-on attendre de tels estimateurs ?

II Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = \theta_0 dt + \sqrt{2}\, \xi_t \, dW_t, \quad \xi_0 = 1, \tag{8.6}$$

où θ_0 est un paramètre inconnu.

(A)

1. Résoudre l'équation $dY_t = \sqrt{2}\, Y_t \, dW_t, Y_0 = 1$.
2. En posant $\eta_t = 1/Y_t$ et $Z_t = \xi_t \, \eta_t$, trouver l'équation différentielle stochastique vérifiée par Z_t. En déduire l'expression de $(\xi_t, t \geq 0)$ en fonction de (Y_t). Vérifier que pour tout $\theta_0 \geq 0$, et pour tout $t \geq 0$, $\xi_t > 0$.
3. Montrer que (ξ_t) est récurrent positif sur $]0, +\infty[$ si et seulement si $\theta_0 > 0$. Calculer dans ce cas sa distribution stationnaire que l'on notera π_{θ_0}.

(B) Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0, T]$.

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0.
2. On suppose $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, v(\theta_0))$. Donner l'expression de $v(\theta_0)$.
3. On suppose toujours $\theta_0 > 0$.
   - On pose $\tilde{\theta}_T = T / \int_0^T (ds/\xi_s)$.Montrer que $\tilde{\theta}_T$ converge p.s. vers θ_0 lorsque T tend vers l'infini.
   - Utiliser la formule d'Ito appliquée à $\log \xi_T$ pour montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, \gamma(\theta_0))$. Donner l'expression de $\gamma(\theta_0)$.
   - Lequel des deux estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$ est préférable? Justifier votre réponse.

(C) On suppose toujours $\theta_0 > 0$ et l'on définit, pour $a > 0$,

$$T_a = \inf\{t \geq 0, \int_0^t \frac{1}{\xi_s^2} ds \geq a\}.$$

Au lieu d'observer (ξ_t) sur $[0, T]$, on l'observe sur l'intervalle $[0, T_a]$ et l'on étudie l'estimateur $\hat{\theta}_{T_a}$.

1. Vérifier que, pour tout $a > 0$, $T_a < +\infty$ p.s. et que $T_a \to +\infty$ p.s. lorsque $a \to +\infty$.
2. Vérifier que T_a/a converge p.s. lorsque a tend vers l'infini une limite que l'on explicitera.
3. En utilisant la martingale locale

$$M_t(\lambda) = \exp i\lambda \int_0^t \frac{1}{\xi_s} dW_s + \lambda^2/2 \int_0^t \frac{1}{\xi_s^2} ds,$$

   déterminer la loi de la v.a. $\int_0^{T_a} \frac{1}{\xi_s} dW_s$
4. En déduire la loi de $\sqrt{a}(\hat{\theta}_{T_a} - \theta_0)$.

Université de Paris 7
Master M2 Modélisation aléatoire.

Examen de Statistique des diffusions. 2 février 2006.

Durée 3H - Documents autorisés.

Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)\, dW_t, \quad \xi_0 = \eta, \tag{8.7}$$

où les fonctions b et σ sont définies, de classe C^1 sur $\mathbb{R}$, lipschitziennes, et η est une variable aléatoire réelle indépendante de (W_t). On suppose de plus que b et σ vérifient les hypothèses suivantes:

- $\exists M > 0, r > 0, d \geq 0, \forall x, |x| \geq M, b(x)\text{sgn}(x) \leq -r|x|^d$ (où $\text{sign}(x) = 1$ si $x > 0$, $= -1$ si $x \leq 0$).
- $\exists \sigma_0, \sigma_1 > 0, \forall x, \sigma_0 \leq \sigma(x) \leq \sigma_1$.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω, adapté à la filtration $(\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0)$.
2. Montrer que le processus (ξ_t) est récurrent positif sur R. Donner l'expression de sa distribution stationnaire π et vérifier qu'elle admet des moments (positifs) de tous les ordres (c'est-à-dire que $\int_{\mathbb{R}} |x|^k \pi(dx) < \infty$ pour tout $k \geq 0$).
3. On suppose que $b(x) = \beta_0 - x$ où $\beta_0 \in \mathbb{R}$ est un paramètre inconnu. On suppose que la fonction σ est connue.
   (a) Calculer l'espérance de la loi stationnaire π.
   (b) On dispose de l'observation $(\xi_t, t \in [0, T])$, où $T > 0$. On pose $\tilde{\beta}_T = (1/T) \int_0^T \xi_s ds$. Montrer que $\tilde{\beta}_T$ converge p.s. lorsque T tend vers l'infini vers β_0. Etudier la convergence en loi de $\sqrt{T}(\tilde{\beta}_T - \beta_0)$.
   (c) Ecrire la vraisemblance de l'observation $(\xi_t, t \in [0, T])$. Calculer l'estimateur du maximum de vraisemblance $\hat{\beta}_T$ de β_0. Etudier cet estimateur lorsque T tend vers l'infini (consistance, loi asymptotique). Comparer avec l'estimateur de la question précédente. Lequel des deux estimateurs est préférable et pourquoi. Discuter les qualités possibles de ces deux estimateurs.
   (d) On reprend l'estimateur $\tilde{\beta}_T$ pour étudier son espérance et sa variance. On suppose que $\mathbb{E}(\eta^2) < \infty$ et l'on pose $\mathbb{E}(\eta) = m$, $\text{Var}(\eta) = c^2$.

i. On pose $X_t = e^t(\xi_t - \beta_0)$. Calculer dX_t et montrer que $X_t = X_0 + M_t$ où (M_t) est une $(\mathcal{F}_t$-martingale.

ii. Calculer $\mathbb{E}(\tilde{\beta}_T)$.

iii. Etudier $\mathrm{Var}(\tilde{\beta}_T)$ et montrer que $\mathrm{Var}(T\tilde{\beta}_T)$ converge, lorsque T tend vers l'infini, vers une limite que l'on précisera.

4. On suppose $\sigma(x) = \sigma$ inconnu et $b(x)$ inconnue. On suppose que η suit la loi stationnaire π. On souhaite estimer σ^2 à partir de l'observation de ξ_t aux instants $t_i = t_i^n = i\Delta_n$, $i = 1, \dots, n$, où Δ_n est un pas de discrétisation tel que $\Delta_n \to 0$ et $n\Delta_n \to +\infty$ lorsque n tend vers l'infini.

   (a) Montrer que $\tilde{\sigma}_n^2 = (1/(n\Delta_n) \sum_{i=1}^n (\xi_{t_i} - \xi_{t_{i-1}})^2$ est un estimateur consistant de σ^2.

   (b) On se propose de démontrer que $\sqrt{n}(\tilde{\sigma}_n^2 - \sigma^2)$ converge en loi (dans le cadre asymptotique ci-dessus) lorsque n tend vers l'infini. On utilisera la décomposition suivante:

$$(\xi_{t_i} - \xi_{t_{i-1}})^2 = \left(\int_{t_{i-1}}^{t_i} b(\xi_s) ds \right)^2 + 2\sigma(W_{t_i} - W_{t_{i-1}})\Delta_n b(\xi_{t_{i-1}}) +$$

$$2\sigma(W_{t_i} - W_{t_{i-1}}) \int_{t_{i-1}}^{t_i} (b(\xi_s) - b(\xi_{t_{i-1}}))ds + \sigma^2(W_{t_i} - W_{t_{i-1}})^2.$$

   Donner une condition sur le pas de discrétisation Δ_n pour que $\sqrt{n}(\tilde{\sigma}_n^2 - \sigma^2)$ converge en loi vers une loi limite gaussienne à préciser.

Université de Paris 7
Master de Modélisation aléatoire (2ème année).

Examen de Statistique des diffusions - 14 septembre 2006

Durée 3H - Documents autorisés

Problème: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$, $x_0 > 0$ fixé. On considère le processus (X_t) défini par l'équation différentielle stochastique :

$$dX_t = (aX_t + b)dt + \sigma X_t dW_t, \quad X_0 = x_0, \tag{8.8}$$

où $\sigma > 0$, $a < \frac{\sigma^2}{2}$ et $b > 0$.

1. Montrer que le processus (X_t) est récurrent positif sur $]0, +\infty[$ et que sa distribution stationnaire π est la loi de $\frac{1}{V}$ où V est une v.a. de loi Gamma de paramètres $\alpha = 1 - \frac{2a}{\sigma^2}$ et $\lambda = \frac{2b}{\sigma^2}$. En déduire
$$\int_0^{+\infty} x^{-1}\pi(dx), \quad \int_0^{+\infty} x^{-2}\pi(dx).$$
(Rappel: la loi Gamma de paramètres α et λ est la loi de densité
$$\gamma_{\alpha,\lambda}(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x}\, x^{\alpha-1}\, 1_{(x>0)}.)$$

2. On observe le processus (X_t) sur l'intervalle $[0, T]$ aux instants $t_i = i\frac{T}{n}$, $i = 1, \ldots, n$, x_0 est connu. On se propose d'étudier l'estimateur de σ^2 défini par:
$$\tilde{\sigma}^2 = \frac{1}{T}\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2$$
où $Y_t = \log X_t$. On supposera pour cette étude que T est fixé et que n tend vers l'infini.
   (a) Montrer que (Y_t) est solution d'une équation différentielle stochastique de la forme
   $$dY_t = f(Y_t)dt + \sigma dW_t.$$
   (b) On pose
   $$A_n = \sqrt{n}\sigma^2 \left(\frac{1}{T}\sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 \; - \; 1\right).$$
   Montrer que A_n converge en loi lorsque n tend vers l'infini et préciser sa loi limite.

(c) On pose $S_n = \sqrt{n}(\tilde{\sigma}^2 - \sigma^2)$. Montrer que $R_n = S_n - A_n$ converge vers 0 en probabilité. (Il conviendra de décomposer R_n en la somme de plusieurs termes).

(d) En déduire la limite en loi de S_n.

3. On suppose dans cette question $\sigma = \sqrt{2}$, $a = 0$. Pour estimer b, on observe le processus (X_t) sur tout l'intervalle $[0, T]$.

   (a) Calculer l'estimateur du maximum de vraisemblance $\hat{b}$ de b.

   (b) Montrer que cet estimateur est consistant lorsque T tend vers l'infini.

   (c) Montrer que $\sqrt{T}(\hat{b} - b)$ converge en loi lorsque T tend vers l'infini, vers une loi gaussienne $\mathcal{N}(0, \gamma^2)$. Calculer γ^2.

   (d) On considère l'estimateur

   $$\tilde{b} = \left(\frac{1}{T} \int_0^T \frac{ds}{X_s} \right)^{-1}$$

   Montrer que l'estimateur $\tilde{b}$ est consistant et que $\sqrt{T}(\tilde{b} - b)$ converge en loi, lorsque T tend vers l'infini, vers une loi limite que l'on précisera.

   (e) Lequel des deux estimateurs $\hat{b}$ et $\tilde{b}$ est, à votre avis, préférable et pourquoi.

   (f) Résoudre l'équation différentielle stochastique $dU_t = \sqrt{2}U_t dW_t, U_0 = 1$. Poser $Z_t = \frac{X_t}{U_t}$ et déterminer l'équation différentielle stochastique vérifiée par (Z_t). En déduire (X_t) en fonction de (U_t).

4. Dans cette question, on suppose les trois paramètres a, b, σ^2 inconnus. Pour les estimer, on observe le processus (X_t) aux instants $t_i = i\Delta$, $i = 1, \ldots, n$ où Δ est un pas de discrétisation donné.

   (a) Ecrire le schéma d'Euler associé à l' équation différentielle (8.8), ainsi que sa vraisemblance.

   (b) Déduire de la question précédente des estimateurs de (a, b, σ^2).

   (c) Proposer différents cadres asymptotiques pour étudier les estimateurs de la question précédente. Suivant le cas, que pouvez-vous dire de leur comportement asymptotique (preuve non demandée).

Université de Paris 7
Master de Modélisation aléatoire (2ème année).

Examen de Statistique des diffusions - 13 février 2007

Durée 3H - Documents autorisés.

Problème: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = -\theta_0 \frac{\xi_t}{\sqrt{1+\xi_t^2}} dt + \sigma_0 dW_t, \quad \xi_0 = 0, \tag{8.9}$$

où $\sigma_0 > 0$, $\theta_0 \in \mathbb{R}$.

1. (1 point) Pour quelles valeurs de (θ_0, σ_0) le processus (ξ_t) est-il récurrent sur $\mathbb{R}$? récurrent positif? Donner, dans ce cas, l'expression de sa distribution stationnaire π_0.
2. Dans cette question, on suppose $\sigma_0 = 1$ et θ_0 inconnu. Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0, T]$.
   (a) (1 point) Calculer l'estimateur $\hat{\theta}_T$ du maximum de vraisemblance de θ_0 basé sur cette observation.
   (b) (2 points) On suppose que $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0. Montrer que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne centrée de variance $v(\theta_0)$. Donner, sans chercher à la calculer explicitement, l'expression de $v(\theta_0)$.
   (c) On suppose toujours $\theta_0 > 0$. On pose

$$X_T = \frac{1}{T}\int_0^T \frac{\xi_t^2}{\sqrt{1+\xi_t^2}} dt, \quad \tilde{\theta}_T = 1/2X_T. \tag{8.10}$$

      i. (1 point) Montrer que $\int_{\mathbb{R}} \frac{x^2}{\sqrt{1+x^2}} \pi_0(dx) = 1/2\theta_0$.
      ii. (1 point) Montrer que $\tilde{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0.
      iii. (1,5 points) On pose $g(x) = (1/2\theta_0) - \frac{x^2}{\sqrt{1+x^2}}$ et

$$A(x, \theta_0) = \int_{-\infty}^{x} g(x)\pi_0(dx). \tag{8.11}$$

Montrer que

$$\exists C > 0, \exists \varepsilon > 0, \forall x \in \mathbb{R}, |A(x, \theta_0)| \leq C \exp\left(-(\theta_0 + \varepsilon)\sqrt{1+x^2}\right).$$

iv. (2 points) Montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne centrée de variance $\gamma(\theta_0)$. Donner, sans chercher à la calculer explicitement, l'expression de $\gamma(\theta_0)$. (voir rappel ci-dessous).

3. On suppose dans cette question que $\theta_0 = 0$ (donc $\xi_t = W_t$) et l'on se propose d'étudier les deux estimateurs précédents.

   (a) (1, 5 points) Montrer que $Y_T = \frac{1}{T}\int_0^T \frac{\xi_t^2}{1+\xi_t^2}dt$ converge en probabilité vers 1 lorsque T tend vers l'infini (on pourra utiliser le changement de variables $t = uT, u \in [0,1]$ et étudier $E(Y_T)$).

   (b) (1 point) Montrer que l'estimateur $\hat{\theta}_T$ converge p.s. vers 0 lorsque T tend vers l'infini et que $\sqrt{T}\hat{\theta}_T$ converge en loi vers une loi gaussienne centrée dont on précisera la variance.

   (c) (1 point) Montrer que $\tilde{\theta}_T$ converge en probabilité vers 0 lorsque T tend vers l'infini (étudier X_T).

   (d) (2 points) Etudier la limite en loi de $\sqrt{T}\tilde{\theta}_T$.

4. (1 point) Parmi ces deux estimateurs, $\hat{\theta}_T$ et $\tilde{\theta}_T$, y-en-a-t-il qui soit préférable à l'autre: discuter des qualités et défauts respectifs.

5. Dans cette question, on suppose les deux paramètres θ_0, σ_0 inconnus. Pour les estimer, on observe le processus (ξ_t) aux instants $t_i = i\Delta$, $i = 1, \ldots, n$ où Δ est un pas de discrétisation donné.

   (a) (1 point) Ecrire le schéma d'Euler associé à l'équation différentielle (8.9), ainsi que sa vraisemblance.

   (b) (1, 5 points) Déduire de la question précédente des estimateurs de (θ_0, σ_0^2).

   (c) (1, 5 points) Proposer différents cadres asymptotiques pour étudier les estimateurs de la question précédente. Suivant le cas, que pouvez-vous dire de leur comportement asymptotique (preuve non demandée).

Rappel: Pour un processus solution de $dX_t = b(X_t)dt + \sigma(X_t)dW_t$, on rappelle que l'opérateur, défini pour les fonctions de classe C^2 sur $\mathbb{R}$, par $Lf = \frac{\sigma^2}{2}f" + bf'$ peut s'écrire comme

$$Lf = \frac{1}{2m}\left(\frac{f'}{s}\right)',$$

où s est la dérivée de la fonction d'échelle, et m la densité de la mesure de vitesse.

Université de Paris 7
Master M2 Modélisation aléatoire.

Examen de Statistique des diffusions. 18 septembre 2007

Durée 3H - Documents autorisés.

I Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considére le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = \theta_0 dt + \sqrt{2}\, \xi_t\, dW_t, \quad \xi_0 = 1,$$

où θ_0 est un paramètre inconnu.

(A)

1. Résoudre l'équation $dY_t = \sqrt{2}\, Y_t\, dW_t, Y_0 = 1$.
2. En posant $\eta_t = 1/Y_t$ et $Z_t = \xi_t\, \eta_t$, trouver l'équation différentielle stochastique vérifiée par Z_t. En déduire l'expression de $(\xi_t, t \geq 0)$ en fonction de (Y_t). Vérifier que pour tout $\theta_0 \geq 0$, et pour tout $t \geq 0$, $\xi_t > 0$.
3. Montrer que, si $\theta_0 > 0$, (ξ_t) est récurrent positif sur $]0, +\infty[$. Calculer dans ce cas sa distribution stationnaire que l'on notera π_{θ_0}.

(B) Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0, T]$.

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0.
2. On suppose $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, v(\theta_0))$. Donner l'expression de $v(\theta_0)$.
3. On suppose toujours $\theta_0 > 0$.
   - On pose $\tilde{\theta}_T = T / \int_0^T (ds/\xi_s)$.Montrer que $\tilde{\theta}_T$ converge p.s. vers θ_0 lorsque T tend vers l'infini.
   - Utiliser la formule d'Ito appliquée à $\log \xi_T$ pour montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, \gamma(\theta_0))$. Donner l'expression de $\gamma(\theta_0)$.
   - Lequel des deux estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$ est préférable? Justifier votre réponse.

(C) On suppose toujours $\theta_0 > 0$ et l'on définit, pour $a > 0$,

$$T_a = \inf\{t \geq 0, \int_0^t \frac{1}{\xi_s^2} ds \geq a\}.$$

Au lieu d'observer (ξ_t) sur $[0, T]$, on l'observe sur l'intervalle $[0, T_a]$ et l'on étudie l'estimateur $\hat{\theta}_{T_a}$.

1. Vérifier que, pour tout $a > 0$, $T_a < +\infty$ p.s. et que $T_a \to +\infty$ p.s. lorsque $a \to +\infty$.
2. Vérifier que T_a/a converge p.s. lorsque a tend vers l'infini une limite que explicitera.
3. En utilisant la martingale locale
$$M_t(\lambda) = \exp\left(i\lambda \int_0^t \frac{1}{\xi_s} dW_s + \lambda^2/2 \int_0^t \frac{1}{\xi_s^2} ds\right),$$
déterminer la loi de la v.a. $\int_0^{T_a} \frac{1}{\xi_s} dW_s$
4. En déduire la loi de $\sqrt{a}(\hat{\theta}_{T_a} - \theta_0)$.

II: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = \sqrt{\theta_0 + \xi_t^2}\, dW_t, \quad \xi_0 = x_0,$$

où $\theta_0 > 0$ est un paramètre inconnu et x_0 est un réel connu.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω.
2. On souhaite estimer θ_0 partir de l'observation de ξ_t aux instants $t_i = t_i^n = iT/n$, $i = 1, \ldots, n$, où $T > 0$ est fixé. On posera $\Delta_n = T/n$.
   - Quelle est la limite en probabilité quand n tend vers l'infini de $\sum_{i=1}^n (\xi_{t_i} - \xi_{t_{i-1}})^2$. En déduire un estimateur consistant de θ_0 qui s'exprime explicitement en fonction de l'observation $(\xi_{t_i}, i \leq n)$. De façon analogue, en considérant les variables aléatoires de la forme
   $$\sum_{i=1}^n \xi_{t_{i-1}}^{2p} (\xi_{t_i} - \xi_{t_{i-1}})^2,$$
   pour p entier positif, construire d'autres estimateurs consistants de θ_0.
   - On considére $Y_1, \ldots, Y_n$ n variables aléatoires telles que:
   $$Y_0 = x_0,\ Y_i - Y_{i-1} = \sqrt{\theta + Y_{i-1}^2}\, \sqrt{\Delta_n}\, \varepsilon_i, i = 1, \ldots, n,$$
   avec $\theta > 0$ et $\varepsilon_1, \ldots, \varepsilon_n$ indépendantes et de même loi gaussienne centrée réduite. Calculer la densité jointe $f(\theta, y_1, \ldots, y_n)$ de $(Y_1, \ldots, Y_n)$.

- On pose $U_n(\theta) = -2\Delta_n\ \log f(\theta, \xi_{t_1}, \ldots, \xi_{t_n})$. Calculer la limite en probabilité lorsque n tend vers l'infini, de $U_n(\theta) - U_n(\theta_0)$ (que l'on exprimera en faisant apparaître la fonction $\varphi(u) = u - 1 - \log u, u > 0$).
- En notant $K_T(\theta_0, \theta)$ la limite obtenue précédemment, montrer que $K_T(\theta_0, \theta) \geq 0$ p.s. et $K_T(\theta_0, \theta) = 0$ p.s. si et seulement si $\theta = \theta_0$.
- Quels estimateurs peut-on construire l'aide de $U_n(\theta)$? Quelles propriétés peut-on attendre de tels estimateurs ?

Université Paris 7 Denis Diderot
Master M2 Modélisation aléatoire.

Examen de Statistique des diffusions. 18 février 2008

Durée 3H - Documents autorisés.

Problème : Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = \theta \xi_t dt + \sigma \, dW_t, \quad \xi_0 = \eta, \tag{8.12}$$

où θ est un paramètre réel inconnu, η est une variable aléatoire réelle indépendante de $(W_t, t \geq 0)$ et σ est connu.

1. (a) Résoudre l'équation différentielle stochastique (8.12).
   (b) Pour quelles valeurs de θ le processus (ξ_t) est-il récurrent sur R? Récurrent positif? Donner dans ce cas l'expression de sa distribution stationnaire π_θ.
   Désormais, pour toute la suite du problème (sauf en question 4), on suppose que θ est tel que le processus est récurrent positif et que η suit la loi π_θ.
   (c) Montrer qu'il existe une constante C telle que, pour tout $t \geq 0$, pour tout $h \in [0, 1]$, $E((\xi_{t+h} - \xi_t)^2) \leq Ch$,
2. On observe le processus (ξ_t) sur tout l'intervalle de temps $[0, T]$ et l'on se propose d'étudier les deux estimateurs suivants:

$$\hat{\theta}_T = \frac{\int_0^T \xi_s d\xi_s}{\int_0^T \xi_s^2 ds} \quad \text{et} \quad \tilde{\theta}_T = -\frac{\sigma^2 T}{2 \int_0^T \xi_s^2 ds}.$$

   (a) Montrer que $\hat{\theta}_T$ est un estimateur fortement consistant de θ lorsque T tend vers l'infini.
   Montrer que $\sqrt{T}(\hat{\theta}_T - \theta)$ converge en loi, lorsque T tend vers l'infini, vers une loi limite que l'on précisera.
   (b) Montrer que $\frac{1}{T} \int_0^T \xi_s^2 ds$ converge p.s. vers une limite $l(\theta)$ que l'on précisera.
   (c) En utilisant la formule d'Ito appliquée à ξ_T^2, montrer que $\sqrt{T}(\frac{1}{T} \int_0^T \xi_s^2 ds - l(\theta))$ converge en loi, lorsque T tend vers l'infini, vers une limite que l'on précisera.
   (d) Montrer que $\tilde{\theta}_T$ est un estimateur fortement consistant de θ.
   (e) Montrer que $\sqrt{T}(\tilde{\theta}_T - \theta)$ converge en loi, lorsque T tend vers l'infini, vers une loi limite que l'on précisera.
   (f) Lequel des deux estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$ est préférable. Commenter.

3. On souhaite estimer θ à partir de l'observation de ξ_t aux instants $t_i = t_i^n = i\Delta_n$, $i = 1, \ldots, n$, où Δ_n est un pas de discrétisation tel que Δ_n tend vers 0, $n\Delta_n$ tend vers l'infini lorsque n tend vers l'infini. De plus, on suppose que $n\Delta_n^2$ tend vers 0 lorsque n tend vers l'infini. Pour estimer θ, on considère les discrétisés des estimateurs précédents:
$$\bar{\theta}_n = \frac{\sum_{i=1}^n \xi_{t_{i-1}}(\xi_{t_i} - \xi_{t_{i-1}})}{\sum_{i=1}^n \xi_{t_{i-1}}^2 \Delta_n} \quad \text{et} \quad \theta_n^* = -\frac{\sigma^2 n\Delta_n}{2\sum_{i=1}^n \xi_{t_{i-1}}^2 \Delta_n}.$$
On définit
$$C_n = \frac{1}{n\Delta_n}\int_0^{n\Delta_n} \xi_s^2 ds - \frac{1}{n}\sum_{i=1}^n \xi_{t_{i-1}}^2 \quad \text{et} \quad Z_n = \sum_{i=1}^n \xi_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}),$$
ainsi que
$$D_n = \frac{1}{n\Delta_n}\left(\sum_{i=1}^n \xi_{t_{i-1}} \int_{t_{i-1}}^{t_i} \xi_s ds - \sum_{i=1}^n \xi_{t_{i-1}}^2 \Delta_n\right)$$
   (a) Montrer que $E|C_n| \leq C\Delta_n^{1/2}$, où C est une constante.
   (b) En déduire que θ_n^* est un estimateur faiblement consistant de θ et que $\sqrt{n\Delta_n}(\theta_n^* - \theta)$ converge en loi vers une loi limite que l'on précisera.
   (c) Montrer que $E|D_n| \leq C\Delta_n^{1/2}$ où C est une constante.
   (d) On pose $\xi_s^n = \sum_{i=1}^n \xi_{t_{i-1}} 1_{]t_{i-1},t_i]}(s)$. En remarquant que $Z_n = \int_0^{n\Delta_n} \xi_s^n dW_s$, montrer que $\frac{1}{\sqrt{n\Delta_n}}(Z_n - \int_0^{n\Delta_n} \xi_s dW_s)$ converge en probabilité vers 0.
   (e) Déduire de ce qui précède la loi limite de $\sqrt{n\Delta_n}(\bar{\theta}_n - \theta)$.

4. On suppose dans cette question que $\theta = 0$. Et l'on étudie à nouveau $\hat{\theta}_T$ et $\tilde{\theta}_T$.
   (a) En introduisant le processus $(B_u^{(T)} = W_{uT}/\sqrt{T}, u \in [0,1]$, montrer que la variable aléatoire $T\hat{\theta}_T$ converge en loi vers une loi limite que l'on précisera.
   (b) Montrer que $T\tilde{\theta}_T$ converge en loi vers une loi limite que l'on précisera.

Exercice: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :
$$d\xi_t = \alpha(\beta - \xi_t)dt + c\,(\xi_t^+)^{1/2}\,dW_t, \quad \xi_0 = 1. \tag{8.13}$$
où $\alpha > 0, \beta > 0$.

1. Pour quelles valeurs de α, β, c le processus est-il récurrent positif sur $(0, +\infty)$.

2. Calculer dans ce cas sa distribution stationnaire.

Université Paris Diderot
Master de modélisation aléatoire en économie et finance.

Examen de Statistique des diffusions - 4 mai 2011.

Durée 3H - Documents autorisés.
Les parties III et IV peuvent être traitées indépendamment.

Problme: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation diffrentielle stochastique :

$$d\xi_t = -\theta_0 \frac{\xi_t}{\sqrt{1+\xi_t^2}} dt + dW_t, \quad X_0 = \eta, \tag{8.14}$$

où $\theta_0 \in \mathbb{R}$ est un paramètre inconnu et η est une variable aléatoire réelle indépendante de (W_t).

I

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω, adapté la filtration $(\mathcal{F}_t = \sigma(\eta, W_s, s \leq t), t \geq 0)$.
2. Pour quelles valeurs de θ_0 le processus (ξ_t) est-il récurrent sur $\mathbb{R}$? Récurrent positif? Donner dans ce cas l'expression de sa distribution stationnaire π_{θ_0}.
3. On suppose que $\theta_0 > 0$ et que η suit la loi π_{θ_0}.
   - Quelle(s) propriété(s) le processus solution (ξ_t) vérifie-t-il.
   - Utiliser la formule d'Ito appliquée ξ_t^2 pour montrer que

$$\int_R \frac{x^2}{\sqrt{1+x^2}} \pi_{\theta_0}(dx) = \frac{1}{2\theta_0}. \tag{8.15}$$

Dans toute la suite, on suppose que $\eta = 0$.

II Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0, T]$.

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0.
2. On suppose $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, v(\theta_0))$. Donner, sans chercher à la calculer explicitement, l'expression de $v(\theta_0)$.
3. On suppose toujours $\theta_0 > 0$.

- Utiliser (8.15) pour construire un estimateur fortement consistant de θ_0 que l'on notera $\tilde{\theta}_T$.
- Utiliser la formule d'Ito applique ξ_T^2 pour montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, \gamma(\theta_0))$. Donner, sans chercher la calculer explicitement, l'expression de $\gamma(\theta_0)$.

4. Montrer que $1 \leq v(\theta_0) \leq \gamma(\theta_0)$. Lequel des deux estimateurs est préférable?

III On suppose toujours $\theta_0 > 0$ et l'on définit, pour $a > 0$,

$$T_a = \inf\{t \geq 0, \int_0^t \frac{\xi_s^2}{1+\xi_s^2} ds \geq a\}.$$

Au lieu d'observer (ξ_t) sur $[0, T]$, on l'observe sur l'intervalle $[0, T_a]$ et l'on tudie l'estimateur $\hat{\theta}_{T_a}$.

1. Vérifier que, pour tout $a > 0$, $T_a < +\infty$ p.s. et que $T_a \to +\infty$ p.s. lorsque $a \to +\infty$.
2. Vérifier que T_a/a converge p.s. vers $v(\theta_0)$ lorsque a tend vers l'infini.
3. En utilisant la martingale locale

$$M_t(\lambda) = \exp i\lambda \int_0^t \frac{\xi_s}{\sqrt{1+\xi_s^2}} dW_s + \lambda^2/2 \int_0^t \frac{\xi_s^2}{1+\xi_s^2} ds,$$

   déterminer la loi de la v.a. $\int_0^{T_a} \frac{\xi_s}{\sqrt{1+\xi_s^2}} dW_s$

4. En déduire la loi de $\sqrt{a}(\hat{\theta}_{T_a} - \theta_0)$.

IV On suppose que $\theta_0 = 0$. On se propose d'étudier les estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$.

1. Le processus (ξ_t) est-il récurrent positif?
2. On pose $Z_T = (1/T) \int_0^T \frac{\xi_s^2}{1+\xi_s^2} ds$. En considérant le changement de variables $s = uT$ dans Z_T et le mouvement brownien $(B_u^{(T)} = W_{uT}/\sqrt{T}, u \in [0, 1])$, montrer que $E(Z_T)$ tend vers 1 lorsque T tend vers l'infini. En déduire que Z_T tend en probabilité vers 1.
3. Montrer que $\hat{\theta}_T$ converge p.s. vers 0 et que $\sqrt{T}\hat{\theta}_T$ converge en loi vers la loi gaussienne $\mathcal{N}(0, 1)$.
4. On pose $X_T = (1/T) \int_0^T \frac{\xi_s^2}{\sqrt{1+\xi_s^2}} ds$. En utilisant le même changement de variables $s = uT$ que ci-dessus, montrer que

$$D_T = X_T/\sqrt{T} - \int_0^1 |B_u^{(T)}| du$$

   converge vers 0 en probabilité lorsque T tend vers l'infini (on pourra étudier $E(|D_T|)$).

5. En déduire que $\tilde{\theta}_T$ converge en probabilité vers 0 et que $\sqrt{T}\tilde{\theta}_T$ converge en loi vers une loi limite que l'on précisera.

Université de Paris Diderot
Master M2 Modélisation aléatoire.

Examen de Statistique des diffusions. 14 février 2012.

Durée 3H - Documents autorisés.

I Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considére le processus (ξ_t) défini par l'équation différentielle stochastique :

$$d\xi_t = (\theta_0(1-\xi_t) - \theta_1\xi_t)\, dt + \sqrt{\xi_t(1-\xi_t)}dW_t, \quad \xi_0 = x, \tag{8.16}$$

où $\theta_0, \theta_1 \in \mathbb{R}$ sont des paramètres inconnus, $x \in]0,1[$ est connu..

1. Pour quelles valeurs de θ_0, θ_1, le processus est-il récurrent positif sur $]0,1[$. Calculer, dans ce cas, sa distribution stationnaire.
2. On suppose désormais que $\theta_1 = \frac{1}{2}$ et $\theta_0 \geq \frac{1}{2}$. On veut estimer θ_0 à partir de l'observation $(\xi_t, t \in [0,T]$.
3. Montrer que, si $\theta_0 > 0$, (ξ_t) est récurrent positif sur $]0,+\infty[$. Calculer dans ce cas sa distribution stationnaire que l'on notera π_{θ_0}.

(B) Pour estimer θ_0, on observe le processus (ξ_t) sur l'intervalle $[0,T]$.

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_T$ de θ_0.
2. On suppose $\theta_0 > 0$. Montrer que $\hat{\theta}_T$ converge p.s. lorsque T tend vers l'infini vers θ_0 et que $\sqrt{T}(\hat{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, v(\theta_0))$. Donner l'expression de $v(\theta_0)$.
3. On suppose toujours $\theta_0 > 0$.
   - On pose $\tilde{\theta}_T = T/\int_0^T (ds/\xi_s)$.Montrer que $\tilde{\theta}_T$ converge p.s. vers θ_0 lorsque T tend vers l'infini.
   - Utiliser la formule d'Ito appliquée à $\log \xi_T$ pour montrer que $\sqrt{T}(\tilde{\theta}_T - \theta_0)$ converge en loi vers une loi gaussienne $\mathcal{N}(0, \gamma(\theta_0))$. Donner l'expression de $\gamma(\theta_0)$.
   - Lequel des deux estimateurs $\hat{\theta}_T$ et $\tilde{\theta}_T$ est préférable? Justifier votre réponse.

(C) On suppose toujours $\theta_0 > 0$ et l'on définit, pour $a > 0$,

$$T_a = \inf\{t \geq 0, \int_0^t \frac{1}{\xi_s^2} ds \geq a\}.$$

Au lieu d'observer (ξ_t) sur $[0,T]$, on l'observe sur l'intervalle $[0,T_a]$ et l'on étudie l'estimateur $\hat{\theta}_{T_a}$.

1. Vérifier que, pour tout $a > 0$, $T_a < +\infty$ p.s. et que $T_a \to +\infty$ p.s. lorsque $a \to +\infty$.
2. Vérifier que T_a/a converge p.s. lorsque a tend vers l'infini une limite que explicitera.
3. En utilisant la martingale locale
$$M_t(\lambda) = \exp\left(i\lambda \int_0^t \frac{1}{\xi_s} dW_s + \lambda^2/2 \int_0^t \frac{1}{\xi_s^2} ds\right),$$
déterminer la loi de la v.a. $\int_0^{T_a} \frac{1}{\xi_s} dW_s$
4. En déduire la loi de $\sqrt{a}(\hat{\theta}_{T_a} - \theta_0)$.

II: Soit $(W_t, t \geq 0)$ un mouvement brownien standard défini sur l'espace $(\Omega, \mathcal{F}, P)$. On considère le processus (ξ_t) défini par l'équation différentielle stochastique :
$$d\xi_t = \sqrt{\theta_0 + \xi_t^2}\, dW_t, \quad \xi_0 = x_0, \tag{8.17}$$
où $\theta_0 > 0$ est un paramètre inconnu et x_0 est un réel connu.

1. Vérifier que cette équation différentielle stochastique admet un unique processus solution sur Ω.
2. On souhaite estimer θ_0 partir de l'observation de ξ_t aux instants $t_i = t_i^n = iT/n$, $i = 1, \ldots, n$, où $T > 0$ est fixé. On posera $\Delta_n = T/n$.
   - Quelle est la limite en probabilité quand n tend vers l'infini de $\sum_{i=1}^n (\xi_{t_i} - \xi_{t_{i-1}})^2$. En déduire un estimateur consistant de θ_0 qui s'exprime explicitement en fonction de l'observation $(\xi_{t_i}, i \leq n)$. De façon analogue, en considérant les variables aléatoires de la forme
   $$\sum_{i=1}^n \xi_{t_{i-1}}^{2p} (\xi_{t_i} - \xi_{t_{i-1}})^2,$$
   pour p entier positif, construire d'autres estimateurs consistants de θ_0.
   - On considére $Y_1, \ldots, Y_n$ n variables aléatoires telles que:
   $$Y_0 = x_0,\ Y_i - Y_{i-1} = \sqrt{\theta + Y_{i-1}^2}\, \sqrt{\Delta_n}\, \varepsilon_i, i = 1, \ldots, n,$$
   avec $\theta > 0$ et $\varepsilon_1, \ldots, \varepsilon_n$ indépendantes et de même loi gaussienne centrée réduite. Calculer la densité jointe $f(\theta, y_1, \ldots, y_n)$ de $(Y_1, \ldots, Y_n)$.
   - On pose $U_n(\theta) = -2\Delta_n\ \log f(\theta, \xi_{t_1}, \ldots, \xi_{t_n})$. Calculer la limite en probabilité lorsque n tend vers l'infini, de $U_n(\theta) - U_n(\theta_0)$ (que l'on exprimera en faisant apparaître la fonction $\varphi(u) = u - 1 - \log u, u > 0$).
   - En notant $K_T(\theta_0, \theta)$ la limite obtenue précédemment, montrer que $K_T(\theta_0, \theta) \geq 0$ p.s. et $K_T(\theta_0, \theta) = 0$ p.s. si et seulement si $\theta = \theta_0$.
   - Quels estimateurs peut-on construire l'aide de $U_n(\theta)$? Quelles propriétés peut-on attendre de tels estimateurs ?