Superefficient drift estimation on the Wiener space

Nicolas Privault and Anthony Réveillac

Abstract - In the framework of nonparametric functional estimation for the drift of a Brownian motion X_t we construct Stein type estimators of the form $X_t + D_t \log F$ which are superefficient when $\sqrt{F}$ is a superharmonic functional on the Wiener space for the Malliavin derivative D.

Estimation suroptimale de dérive sur l'espace de Wiener

Résumé - Dans le cadre de l'estimation fonctionnelle nonparamétrique de la dérive d'un mouvement brownien X_t nous construisons des estimateurs de type Stein de la forme $X_t + D_t \log F$, qui sont suroptimaux lorsque $\sqrt{F}$ est une fonctionnelle surharmonique sur l'espace de Wiener pour la dérivée de Malliavin D.

Version française abrégée

Il est bien connu que l'estimateur du maximum de vraisemblance $\hat{\mu}$ de la moyenne $\mu \in \mathbb{R}^d$ d'un vecteur gaussien X dans $\mathbb{R}^d$ de covariance $\sigma^2 \mathrm{I}$ sous une probabilité $\mathbb{P}_\mu$ est égal à X lui-même, et atteint la borne de Cramer-Rao $\sigma^2 d$. Dans [2], Stein a construit des estimateurs de $\mu \in \mathbb{R}^d$ de la forme $X + \sigma^2 \operatorname{grad} \log f(X)$ en utilisant la formule d'intégration par parties sur $\mathbb{R}^d$ par rapport à la densité gaussienne. Ces estimateurs sont suroptimaux lorsque $\sqrt{f}$ est une fonction surharmonique sur $\mathbb{R}^d$. Dans cette Note nous construisons des estimateurs nonparamétriques de type Stein pour la dérive déterministe d'un mouvement brownien X_t de variance σ^2 en utilisant la formule d'intégration par parties du calcul de Malliavin et l'analyse harmonique sur l'espace de Wiener. Ici, le processus $\hat{u} = (X_t)_{t\in[0,T]}$ est considéré comme un estimateur du maximum de vraisemblance (EMV) de sa propre dérive $(u_t)_{t\in[0,T]}$ sous $\mathbb{P}_u^\sigma$, qui atteint la borne de Cramer-Rao $\sigma^2 T^2/2$ sur tous les estimateurs adaptés et sans biais de $(u_t)_{t\in[0,T]}$. Dans ce contexte nous construisons aussi des estimateurs de Bayes admissibles, et montrons que $\hat{u}$ est minimax. A l'aide du gradient de Malliavin D nous obtenons l'identité

$$\mathbb{E}_u^\sigma \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] = \sigma^2 \frac{T^2}{2} + 4\sigma^4 \, \mathbb{E}_u^\sigma \left[\frac{\Delta\sqrt{F}}{\sqrt{F}} \right],$$

où Δ est un laplacien sur l'espace de Wiener et F est une variable aléatoire positive suffisamment régulière. Il suit en particulier de cette identité que $(X_t + \sigma^2 D_t \log F)_{t\in[0,T]}$ est un estimateur suroptimal de $(u_t)_{t\in[0,T]}$ lorsque $\sqrt{F}$ est surharmonique sur l'espace de Wiener. Les estimateurs de u ainsi construits sont minimax, et ils sont non seulement biaisés mais aussi anticipants par rapport à la filtration brownienne $(\mathcal{F}_t)_{t\in[0,T]}$. Comme exemple nous considérons une situation simple où les coefficients du développement en série de Paley-Wiener de la trajectoire browienne observée sont connus. Nous construisons des fonctionnelles de Wiener surharmoniques comme fonctionnelles cylindriques, avec

$$D_t \log F = -(n-2)\sqrt{\frac{2}{T}} \sum_{k=1}^{n} \frac{\displaystyle\int_0^T \dot{h}_k(s) dX_s \sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right)}{\left|\int_0^T \dot{h}_1(s) dX_s\right|^2 + \cdots + \left|\int_0^T \dot{h}_n(s) dX_s\right|^2}, \qquad t \in [0,T],$$

pour $n \geq 3$. Le gain de l'estimateur $X + \sigma^2 D \log F$, comparé avec celui de l'EMV $\hat{u}$, est proche de $11,38\%$ pour $n = 4$. Les résultats présentés dans cette Note sont extraits de [1] où ils sont prouvés dans le cadre plus général où σ dépend du temps et u est un processus adapté.

1 Notation and preliminaries

The maximum likelihood estimator $\hat{\mu}$ of the mean $\mu \in \mathbb{R}^d$ of a Gaussian random vector X in $\mathbb{R}^d$ with covariance $\sigma^2 \mathrm{I}$ under a probability $\mathbb{P}_\mu$ is well-known to be equal to X itself. It is efficient in the sense that it attains the Cramer-Rao bound

$$\sigma^2 d = \mathbb{E}_\mu[\|X - \mu\|_d^2] = \inf_Z \mathbb{E}_\mu[\|Z - \mu\|_d^2],$$

over all unbiased estimators Z of $\mu \in \mathbb{R}^d$. In [2], Stein used integration by parts with respect to the Gaussian density to prove the identity

$$\mathbb{E}_\mu[\|X + \sigma^2 \operatorname{grad} \log f(X) - \mu\|_d^2] = \sigma^2 d + 4\sigma^4 \sum_{i=1}^{d} \mathbb{E}_\mu\left[\frac{\partial_i^2 \sqrt{f}(X)}{\sqrt{f}(X)}\right],$$

which shows that if $\sqrt{f}$ is superharmonic on $\mathbb{R}^d$, then $X + \sigma^2 \operatorname{grad} \log f(X)$ is a superefficient estimator of μ. In this Note we construct nonparametric superefficient estimators of the deterministic drift $(u_t)_{t\in[0,T]}$ of a Brownian motion X_t, using the integration by parts formula of the Malliavin calculus and harmonic analysis on the Wiener space. Let $(\Omega, H, \mathbb{P}^\sigma)$ denote the Wiener space, where $\Omega = \mathcal{C}_0([0,T])$ is the space of continuous functions on $[0,T]$ starting at 0,

$$H = \left\{ v : [0,T] \to \mathbb{R} \ : \ v(t) = \int_0^t \dot{v}(s) ds, \ t \in [0,T], \ \dot{v} \in L^2([0,T]) \right\}$$

is the Cameron-Martin space with inner product $\langle v_1, v_2 \rangle_H = \int_0^T \dot{v}_1(s)\dot{v}_2(s) ds$, $v_1, v_2 \in H$, and $\mathbb{P}^\sigma$ is the Wiener measure with variance $\sigma^2 > 0$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by the canonical process $(X_t)_{t\in[0,T]}$, and for $u \in \Omega$, let $\mathbb{P}_u^\sigma$ denote the translation of the Wiener measure on Ω by u, i.e. $(X_t - u_t)_{t\in[0,T]}$ is a standard Brownian motion with variance $\sigma^2 > 0$ under $\mathbb{P}_u^\sigma$. We fix $(h_n)_{n\geq 1}$ a total subset of H and let $\mathcal{S}$ denote the space of cylindrical functionals of the form

$$F = f_n\left(\int_0^T \dot{h}_1(s) dX_s, \ldots, \int_0^T \dot{h}_n(s) dX_s\right), \tag{1.1}$$

where f_n is in the space of infinitely differentiable rapidly decreasing functions on $\mathbb{R}^n$, $n \geq 1$. The H-valued Malliavin derivative D, see [3] and references therein, is defined as

$$D_t F = \sum_{i=1}^{n} h_i(t) \partial_i f_n \left(\int_0^T \dot{h}_1(s) dX_s, \ldots, \int_0^T \dot{h}_n(s) dX_s \right),$$

for $F \in \mathcal{S}$ of the form (1.1). Let $\delta : L^2_u(\Omega; H) \to L^2_u(\Omega)$ denote the divergence operator under $\mathbb{P}^\sigma_u$, which satisfies the integration by parts formula

$$\mathbb{E}^\sigma_u[F\delta(v)] = \sigma^2\, \mathbb{E}^\sigma_u[\langle v, DF \rangle_H], \qquad v \in \mathrm{Dom}(\delta). \tag{1.2}$$

It is well-known that D and δ are closable, and their domains will be denoted by $\mathrm{Dom}(D)$ and $\mathrm{Dom}(\delta)$. We define the Laplacian Δ by

$$\Delta F = \mathrm{trace}_{L^2([0,T])^{\otimes 2}} DDF = \int_0^T D_s D_s F ds, \qquad F \in \mathcal{S},$$

i.e.

$$\Delta F = \sum_{i,j=1}^{n} \int_0^T h_i(s) h_j(s) ds \; \partial_i \partial_j f_n \left(\int_0^T \dot{h}_1(s) dX_s, \ldots, \int_0^T \dot{h}_n(s) dX_s \right),$$

for F of the form (1.1). Note that unlike the Gross Laplacian Δ_G defined by $\Delta_G F = \mathrm{trace}_{H^{\otimes 2}} DDF$, the operator Δ is closable, and its domain will be denoted by $\mathrm{Dom}(\Delta)$.

2 Maximum likelihood and Bayes estimators

An estimator ξ of $u \in H$ is called unbiased if

$$\mathbb{E}^\sigma_u[\xi_t] = u_t, \quad t \in [0,T], \quad u \in H,$$

and adapted if $(\xi_t)_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted. Here, the process $\hat{u} = (X_t)_{t\in[0,T]}$ will be considered as an unbiased estimator of its own drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}^\sigma_u$. In particular, given N independent samples $(X^1_t)_{t\in[0,T]}, \ldots, (X^N_t)_{t\in[0,T]}$ of $(X_t)_{t\in[0,T]}$, the process $\bar{X}_t := (X^1_t + \cdots + X^N_t)/N$, $t \in [0,T]$, is an unbiased and consistent estimator of $(u_t)_{t\in[0,T]}$ as N goes to infinity. Moreover, $\hat{u} := (X_t)_{t\in[0,T]}$ is a maximum likelihood estimator (MLE) of $(u_t)_{t\in[0,T]}$ in the sense that it formally maximizes the Girsanov density

$$\frac{d\mathbb{P}^\sigma_v}{d\mathbb{P}^\sigma} = \exp\left(\int_0^T \frac{\dot{v}_s}{\sigma^2} dX_s - \frac{1}{2} \int_0^T \frac{\dot{v}^2_s}{\sigma^2} ds \right)$$

in $v \in H$. Moreover it attains the Cramer-Rao bound $\sigma^2 T^2/2$ presented in the next proposition.

Proposition 1. *For any unbiased and adapted estimator ξ of u we have*

$$\mathbb{E}^\sigma_u\left[\int_0^T |\xi_t - u_t|^2 dt \right] \geq \mathbb{E}^\sigma_u\left[\int_0^T |X_t - u_t|^2 dt \right] = \sigma^2 \frac{T^2}{2}, \qquad u \in H. \tag{2.1}$$

For any $\tau > 0$ and $v \in H$, the Bayes risk

$$\int_\Omega \mathbb{E}_z^\sigma \left[\int_0^T |\xi_t - z_t|^2 dt \right] d\mathbb{P}_v^\tau(z) \tag{2.2}$$

of any estimator $(\xi_t)_{t\in[0,T]}$ on Ω under the prior distribution $\mathbb{P}_v^\tau$ is uniquely minimized by

$$\xi_t^{\tau,v} := \frac{\sigma^2}{\tau^2+\sigma^2} v_t + \frac{\tau^2}{\tau^2+\sigma^2} X_t, \qquad t \in [0,T],$$

which has risk $\tau^2\sigma^2T^2/(2(\tau^2+\sigma^2))$. Clearly $\xi^{\tau,v}$ is unique in the sense that it is the only estimator to minimize (2.2), hence it is also admissible. However $\xi^{\tau,v}$ is not minimax since for all $u \in H$, its mean square error under $\mathbb{P}_u^\sigma$ is equal to

$$\mathbb{E}_u^\sigma \left[\int_0^T |\xi_t^{\tau,v} - u_t|^2 dt \right] = \frac{\sigma^2\tau^4}{(\tau^2+\sigma^2)^2} \frac{T^2}{2} + \frac{\sigma^4}{(\tau^2+\sigma^2)^2} \|u-v\|^2_{L^2([0,T])}.$$

Since the Bayes risk $\tau^2\sigma^2T^2/(2(\tau^2+\sigma^2))$ of $\xi^{\tau,v}$, $\tau \in \mathbb{R}$, converges as $\tau \to \infty$ to the Cramer-Rao bound $\sigma^2T^2/2$, it follows that the maximum likelihood estimator $\hat{u} = (X_t)_{t\in[0,T]}$ is minimax, i.e. for all $u \in \Omega$ we have

$$\sigma^2 \frac{T^2}{2} = \mathbb{E}_u^\sigma \left[\int_0^T |X_t - u_t|^2 dt \right] = \inf_\xi \sup_{v\in\Omega} \mathbb{E}_v^\sigma \left[\int_0^T |\xi_t - v_t|^2 dt \right]. \tag{2.3}$$

3 Superefficient drift estimators

We aim to construct a superefficient estimator of u of the form $X + \xi$, with mean square error strictly smaller than the Cramer-Rao bound when ξ is a suitably chosen stochastic process. This estimator will be biased and possibly anticipating with respect to the Brownian filtration. For this we carry over Stein's argument to the Wiener space using the duality relation (1.2) between the gradient and divergence operators.

Lemma 1. *Unbiased risk estimate. For any $\xi \in L_u^2(\Omega \times [0,T])$ such that $\xi_t \in \mathrm{Dom}(D)$, $t \in [0,T]$, we have*

$$\mathbb{E}_u^\sigma \left[\|X + \xi - u\|^2_{L^2([0,T])} \right] = \sigma^2 \frac{T^2}{2} + \|\xi\|^2_{L_u^2(\Omega\times[0,T])} + 2\sigma^2\, \mathbb{E}_u^\sigma \left[\int_0^T D_t\xi_t dt \right]. \tag{3.1}$$

The next proposition specializes the above lemma to processes ξ of the form

$$\xi_t = \sigma^2 D_t \log F, \qquad t \in [0,T],$$

where F is an a.s. strictly positive and sufficiently smooth random variable.

Proposition 2. *Stein type estimator. Let $F \in \mathcal{S}$ be such that $F > 0$, $\mathbb{P}^\sigma$-a.s. and $\sqrt{F} \in \mathrm{Dom}(\Delta)$. We have*

$$\mathbb{E}_u^\sigma \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] = \sigma^2 \frac{T^2}{2} + 4\sigma^4\, \mathbb{E}_u^\sigma \left[\frac{\Delta\sqrt{F}}{\sqrt{F}} \right]. \tag{3.2}$$

Relation (3.2) extends to any $F \in \mathrm{Dom}(D)$ such that $\sqrt{F} \in \mathrm{Dom}(\Delta)$, and $F > 0$, $\Delta\sqrt{F} \leq 0$, $\mathbb{P}^\sigma$-a.s., and in particular, $X + \sigma^2 D \log F$ is a superefficient estimator of u if $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}^\sigma$-measure. As in [2], the superefficient estimators constructed in this way are minimax in the sense that for all $u \in H$ we have

$$\mathbb{E}_u^\sigma \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] < \sigma^2 \frac{T^2}{2} = \inf_\xi \sup_{v \in \Omega} \mathbb{E}_v^\sigma \left[\int_0^T |\xi_t - v_t|^2 dt \right],$$

thus showing that the MLE $\hat{u} = (X_t)_{t \in [0,T]}$ is inadmissible. In contrast to the MLE, such estimators are not only biased but also anticipating with respect to the Brownian filtration $(\mathcal{F}_t)_{t \in [0,T]}$.

4 Numerical application

We consider a simple example where the coefficients of the Paley-Wiener expansion of the observed path are known. Here we assume that $(h_k)_{k \geq 1}$ is orthonormal in $L^2([0,T])$, and additionally that $(h_k)_{k \geq 1}$ is orthogonal in H, for instance one can take $h_n(t) = \sqrt{2T^{-1}} \sin((n - 1/2)\pi t/T)$, $t \in [0,T]$, $n \geq 1$. Superharmonic functionals on the Wiener space can be constructed as cylindrical functionals, by composition with finite-dimensional functions. Given $a \in \mathbb{R}$, let

$$F_{n,a} = \left(\left| \int_0^T \dot{h}_1(t) dX_t \right|^2 + \cdots + \left| \int_0^T \dot{h}_n(t) dX_t \right|^2 \right)^{a/2}.$$

We have

$$\frac{\Delta\sqrt{F_{n,a}}}{\sqrt{F_{n,a}}} = \frac{a(n-2+a/2)/2}{\left| \int_0^T \dot{h}_1(t) dX_t \right|^2 + \cdots + \left| \int_0^T \dot{h}_n(t) dX_t \right|^2},$$

which is negative if $4 - 2n \leq a \leq 0$, and minimal for $a = 2 - n$. The estimator of u will be given by

$$D_t \log F_{n,2-n} = -(n-2)\sqrt{\frac{2}{T}} \sum_{k=1}^n \frac{\displaystyle\int_0^T \dot{h}_k(s) dX_s \sin\left(\left(k - \frac{1}{2} \right) \frac{\pi t}{T} \right)}{\left| \int_0^T \dot{h}_1(s) dX_s \right|^2 + \cdots + \left| \int_0^T \dot{h}_n(s) dX_s \right|^2},$$

with

$$\|D \log F_{n,2-n}\|^2_{L^2([0,T])} = \frac{(n-2)^2}{\left| \int_0^T \dot{h}_1(s) dX_s \right|^2 + \cdots + \left| \int_0^T \dot{h}_n(s) dX_s \right|^2}. \tag{4.1}$$

For simulation purposes we construct the (nondrifted) Brownian motion $(Y_t)_{t \in [0,T]}$ via the Paley-Wiener series

$$Y_t = \sigma \frac{\sqrt{2T}}{\pi} \sum_{n=1}^\infty \eta_n \frac{\sin\left(\left(n - \frac{1}{2} \right) \frac{\pi t}{T} \right)}{\left(n - \frac{1}{2} \right)}, \tag{4.2}$$

where $(\eta_n)_{n\geq 1}$ are independent standard Gaussian random variables with unit variance under $\mathbb{P}_u^\sigma$. Unlike in classical Stein estimation, n becomes a free parameter and there is some interest in determining optimal values of n. Next we present numerical simulations which allow us to measure the efficiency of our estimators. The gain of the superefficient estimator $X + D\log F_{n,2-n}$ compared to that of the MLE is given by

$$G(u,\sigma,T,n) = -\frac{8\sigma^2}{T^2}\,\mathbb{E}_u^\sigma\left[\frac{\Delta\sqrt{F_{n,2-n}}}{\sqrt{F_{n,2-n}}}\right] = \frac{2\sigma^2}{T^2}\,\mathbb{E}_u^\sigma\left[\|D\log F_{n,2-n}\|^2_{L^2([0,T])}\right].$$

Finally we choose $u_t = \alpha t$, $t \in [0,T]$, for some $\alpha \in \mathbb{R}$. In this case, $G(\alpha,\sigma,T,n)$ converges to $(n-2)^2\frac{8}{\pi^2}\,\mathbb{E}\left[\left(\sum_{l=1}^n (2l-1)^2\,\eta_l^2\right)^{-1}\right]$ when $\alpha^{-2}\sigma^2/T$ tends to infinity, which yields 11.38% for $n=4$, and is equivalent to $\alpha^{-2}(1-2/n)^2\sigma^2/T$ as $\alpha^{-2}\sigma^2/T$ tends to 0, and to $6/(n\pi^2)$ as n goes to infinity.

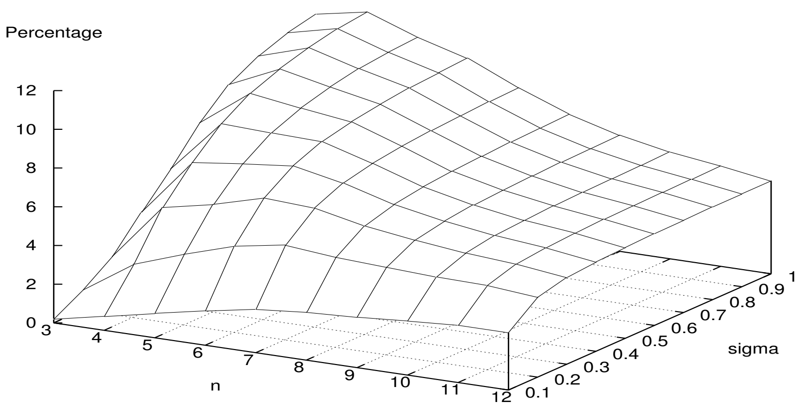

Figure 4.1: Percentage gain as a function of n and σ.
Figure 4.1: Gain en pourcentage en fonction de n et σ.

References

[1] N. Privault and A. Réveillac. Stein estimation on the Wiener space. Preprint, 2006.

[2] C. Stein. Estimation of the mean of a multivariate normal distribution. *Ann. Stat.*, 9(6):1135–1151, 1981.

[3] A.S. Üstünel. *An introduction to analysis on Wiener space*, volume 1610 of *Lecture Notes in Mathematics*. Springer Verlag, 1995.

N.P.: Département de Mathématiques, Université de Poitiers, F-86962 Futuroscope, France. nicolas.privault@math.univ-poitiers.fr
A.R.: Laboratoire de Mathématiques, Université de La Rochelle, F-17042 La Rochelle, France. anthony.reveillac@univ-lr.fr