Nonparametric Minimax Estimation of the Volatility in High-Frequency Models Corrupted by Noise

J. Schmidt-Hieber
Joint work with Axel Munk and Tony Cai

Institut für Mathematische Stochastik, Göttingen

`www.stochastik.math.uni-goettingen.de`
`schmidth@math.uni-goettingen.de`

Statistical Regularisation

Introduction

Nonparametric Minimax Estimation of the Volatility in High-Frequency Models Corrupted by Noise

Schmidt-Hieber

Models

Estimation

Numerical Results

Summary/ Outlook

Statistical Inverse Problems: Given observations of the model

$$Y = Kf + \epsilon$$

where

- K is an operator
- ϵ measurement noise

estimate (reconstruct) the function f.

- If K is linear, we say that this is a linear inverse problem.
- If the errors are not identically distributed, than the noise process is called heteroscedastic.

The models

- We consider two models.
- For $i = 1, \ldots, n$

$$Y_{i,n} = \int_0^{i/n} \sigma(s)\, dW_s + \tau\left(\frac{i}{n}\right) \epsilon_{i,n}. \tag{1}$$

$$\tilde{Y}_{i,n} = \sigma\left(\frac{i}{n}\right) W_{i/n} + \tau\left(\frac{i}{n}\right) \epsilon_{i,n}, \tag{2}$$

- σ, τ are deterministic, unknown, positive functions.
- $\epsilon_{i,n}$, i.i.d., $\mathrm{E}\left(\epsilon_{i,n}\right) = 0$, $\mathrm{E}\left(\epsilon_{i,n}^2\right) = 1$, $\mathrm{E}\left(\epsilon_{i,n}^4\right) < \infty$.
- $(W_t)_{t \geq 0}$ is a Brownian motion.
- $(\epsilon_{1,n}, \ldots, \epsilon_{n,n})$ and $(W_t)_{t \geq 0}$ are considered to be independent for $i = 1, \ldots, n$.

Connections to inverse problems

The models can be seen as linear statistical inverse problems with random operator. For $i = 1, \ldots, n$

$$Y_{i,n} = (K\sigma)\left(\frac{i}{n}\right) + \text{heteroscedastic noise},$$
$$\tilde{Y}_{i,n} = \left(\tilde{K}\sigma\right)\left(\frac{i}{n}\right) + \text{heteroscedastic noise},$$

where

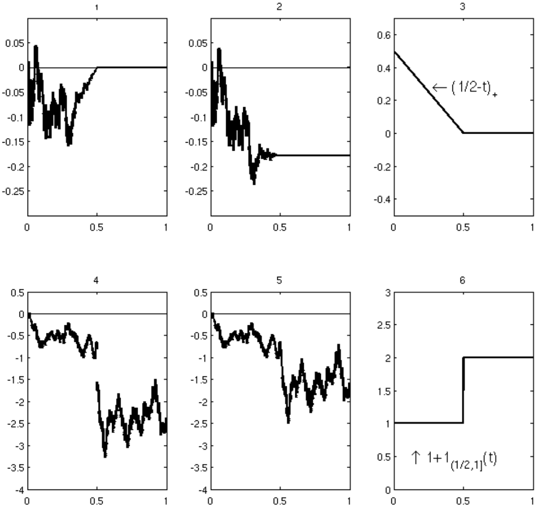
$$(K\sigma)(t) = \int_0^t \sigma(s)\, dW_s,$$
$$\left(\tilde{K}\sigma\right)(t) = \sigma(t)\, W_t.$$

Statistical Problem

Estimation of the functions σ^2 and τ^2, pointwise.

Differences

$\sigma(t)\, W_t$ $\quad \int_0^t \sigma(s)\, dW_s$ $\quad \sigma$



From now on we will only consider Model (1). Estimation and theoretical results are similar for the second model.

Origin

- The model stems from high-frequency modeling of stock returns (Fan et al. '03, Barndorff-Nielsen et al. '06).
- These models are however much more general. Economic theory indicates that σ is stochastic itself.
- Theory focusses so far only on estimation of the integrated moments of volatility (Ait-Sahalia et *al.* '05, Podolskij and Vetter '09), i.e.

$$\int_0^t \sigma^{2p}(s)\, ds \quad p = 1, 2, \ldots$$

- Remarkable exceptions are Malliavin and Mancino '05 and Hoffmann '99.

Estimation of σ^2 and τ^2

- Non-parametric approach necessary.
- We do not have independent observations.
- Transformation that diagonalizes the process depends on the unkown quantities $\sigma(t)$ and $\tau(t)$ and can not be computed explicitly in general.

Estimation

Step 1(4):

Consider the weighted increments $\Delta_i Y(g) := g(i/n)(Y_{i+1,n} - Y_{i,n})$ and observe that

$$\Delta_i Y(g) \approx n^{-1/2} (g\sigma)\left(\frac{i}{n}\right)\eta_{i,n} + (g\tau)\left(\frac{i}{n}\right)\underbrace{(\epsilon_{i+1,n} - \epsilon_{i,n})}_{\text{MA}(1)},$$

where $\eta_{i,n} \sim \mathcal{N}(0,1)$, iid.

Step 1(4):

Consider the weighted increments $\Delta_i Y(g) := g\left(i/n\right)\left(Y_{i+1,n} - Y_{i,n}\right)$ and observe that

$$\Delta_i Y\left(g\right) \approx n^{-1/2}\left(g\sigma\right)\left(\frac{i}{n}\right)\eta_{i,n} + \left(g\tau\right)\left(\frac{i}{n}\right)\underbrace{\left(\epsilon_{i+1,n} - \epsilon_{i,n}\right)}_{\text{MA(1)}},$$

where $\eta_{i,n} \sim \mathcal{N}\left(0,1\right)$, iid.

- The noise term "dominates".
- Important: The variance of the first term is of order $1/n$ whereas the eigenvalues of the covariance of the MA$\left(1\right)$-process behave like i^2/n^2.
- In spectral domain, we can use the first $\sqrt{n}$ observations.
- This can be viewed as the degree of ill-posedness of the problem.

Estimation

Step 2(4):

- The increment process $\Delta Y(g) := (\Delta_1 Y(g), \ldots, \Delta_{n-1} Y(g))$ is stationary, if σ, τ, g are constants.
- Stationary processes are "almost" diagonalized by Discrete Fourier Transforms.
- Therefore, we transform by DST, i.e. we consider

$$Z := D_n \quad (\Delta Y(g)),$$

where

$$D_n := \left(\sqrt{\frac{2}{n}} \sin\left(\frac{ij\pi}{n}\right) \right)_{i,j=1,\ldots,n-1}.$$

Estimation

Step 3(4):

- Let $Z := D_n\left(\Delta Y\left(g\right)\right)$ and define the estimator

$$\widehat{\langle \tau^2, g^2 \rangle} := \frac{1}{n-m} \sum_{i=m+1}^{n-1} \lambda_i^{-1} Z_i^2,$$

  where $\lambda_i = 4\sin^2\left(i\pi/\left(2n\right)\right)$ and $m = m_n$, s.t. $m/\sqrt{n} \to \infty$ and $m/n \to 0$.

- Under smoothness assumptions on σ, τ, $\widehat{\langle \tau^2, g^2 \rangle}$ estimates $\int_0^1 \tau^2(s) g^2\left(s\right) ds$ at a convergence rate $n^{-1/2}$.

Estimation

- Similar for $\widehat{\langle \sigma^2, g^2 \rangle}$

$$\widehat{\langle \sigma^2, g^2 \rangle} := \sqrt{n} \sum_{i=[n^{1/2}]+1}^{2[n^{1/2}]} Z_i^2 - \underbrace{\lambda_i \widehat{\langle \tau^2, g^2 \rangle}}_{\text{bias correction}} .$$

- Under smoothness assumptions on σ, τ, $\widehat{\langle \sigma^2, g^2 \rangle}$ estimates $\int_0^1 \sigma^2(s) g^2(s)\, ds$ at a convergence rate $n^{-1/4}$.

Step 4(4):

- We are able to estimate $\langle \sigma^2, g^2 \rangle$ for all sufficiently smooth g.
- We use this to construct a series estimator.

Basis Functions

- Consider the $L^2[0,1]$ ONS, $\{\psi_k\} = \{1, \sqrt{2}\cos(k\pi t), k = 1, \ldots\}$.
- Further we introduce $f_k : [0,1] \to \mathbb{R}$,

$$f_k(x) := \psi_k(x/2), \quad k = 0, 1, \ldots$$

Note

$$f_k^2(x) = 1 + \cos(k\pi x) =: \psi_0(x) + 2^{-1/2}\psi_k(x), \quad k \geq 1.$$

- $\psi_i \psi_j = 2^{-1/2}(\psi_{i-j} + \psi_{i+j})$
- $\sin(\frac{2i-1}{2}\pi)\sin(\frac{2j-1}{2}\pi) = 2^{-3/2}(\psi_{i-j} + \psi_{i+j+1})$

The pointwise estimator

The estimator of $\sigma^2(t)$ is then given by

$$\hat{\sigma}_N^2(t) = \langle \widehat{\sigma^2, f_0^2} \rangle + 2 \sum_{i=1}^{N} \left(\langle \widehat{\sigma^2, f_i^2} \rangle - \langle \widehat{\sigma^2, f_0^2} \rangle \right) \cos(i\pi t),$$

where N is some threshold parameter.

Assumptions

Function space: (Truncated) Sobolev s-ellipsoid

$$\begin{aligned}\Theta_s^b &= \Theta_s^b(\alpha, C, [l,u]) \\ &= \Big\{ f \in L^2[0,1] : l \leq f \leq u, \quad \exists \, (\theta_n)_n \, , \\ &\qquad \text{s. t. } f(x) = \theta_0 + 2\sum_{i=1}^{\infty} \theta_i \cos(i\pi x) \, , \sum_{i=1}^{\infty} i^{2\alpha}\theta_i^2 \leq C \Big\}\end{aligned}$$

We always assum that $l > 0, u < \infty$.
Characterisation:
For any q odd, $q < \alpha \in \mathbb{N}$, $f^{(q)}(0) = f^{(q)}(1) = 0$ and

$$\int_0^1 (f^{(\alpha)})^2(x)dx \leq \tilde{C}$$

Theoretical results of the estimator

Upper bound on the risk, Munk, S-H '08

Suppose $Q, \bar{Q} > 0$ are fixed constants. Assume model (1) and $\alpha > 3/4, \ \beta > 5/4$. Then it holds for $N^* = n^{1/(4\alpha+2)}$

$$\sup_{\tau^2 \in \Theta_s^b(\beta, \bar{Q}), \ \sigma^2 \in \Theta_s^b(\alpha, Q)} \operatorname{MISE}\left(\hat{\sigma}^2_{N^*}\right) = O\left(n^{-\alpha/(2\alpha+1)}\right).$$

Note that this is "half" of the minimax rate in nonparametric regression. Recall: Eigenvalues $\lambda_i \sim i^2/n^2$ are of order $O(1/n)$ as long as $i = O\left(\sqrt{n}\right)$.

Lower bound

- Many results known about lower bounds, mainly for independent observations and regression.
- Here:
  - Estimation of the scale of a Brownian motion.
  - Dependent observations.

Idea of proof

- Similar as in nonparametric regression: We use a multiple testing argument.
- Main problem: Bound Kullback-Leibler divergence between two multivariate centered normal random variables.
- Results by Golubev '08 and Reiß '08.: Bounds for Hellinger distance of multivariate centered normal r. v.s under the restriction that eigenvalues of covariance matrix are uniformly bounded or as in Reiß '08 that one covariance matrix is the identity.
- However for our purpose one has to allow that eigenvalues tend to 0 and infinity.

KL distance bound

Munk, S-H '08

Let $X \sim \mathcal{N}(\mu, \Sigma_0)$ and $Y \sim \mathcal{N}(\mu, \Sigma_1)$ and denote by P_X and P_Y the corresponding probability measures. Assume $0 < C\Sigma_0 \leq \Sigma_1$ for some constant $0 < C \leq 1$. Then

$$d_{KL}(P_Y, P_X) \leq \frac{1}{4C^2} \left\| \Sigma_0^{-1}\Sigma_1 - I_n \right\|_F^2 .$$

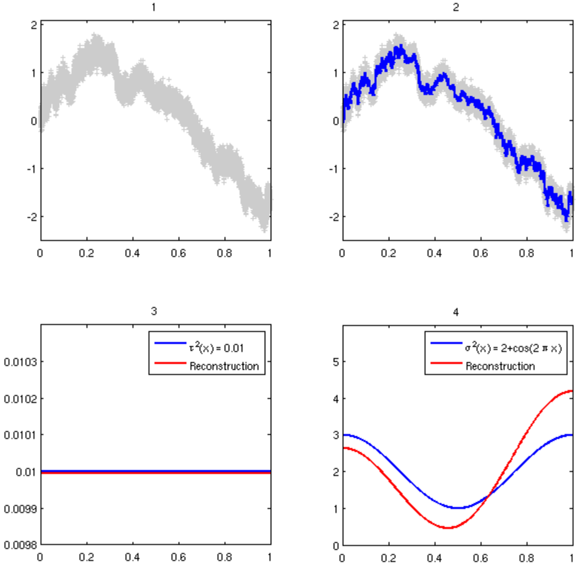
Lower bound

Munk, S-H '08

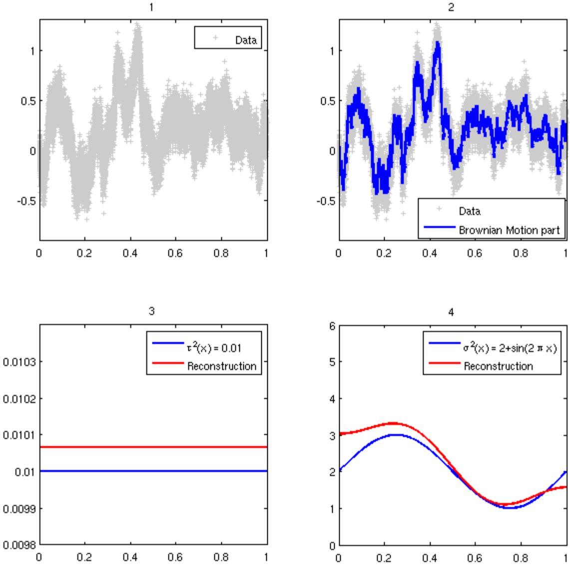
Assume model (2) or model (1), $\alpha \in \mathbb{N}^*$ and $\tau > 0$. Then there exists a $C > 0$, such that

$$\varliminf_{n\to\infty} \inf_{\hat{\sigma}_n^2} \sup_{\sigma^2 \in \Theta_s^b(\alpha,Q)} \mathrm{E}\left(n^{\frac{\alpha}{2\alpha+1}} \left\| \hat{\sigma}^2 - \sigma^2 \right\|_2^2 \right) \geq C.$$

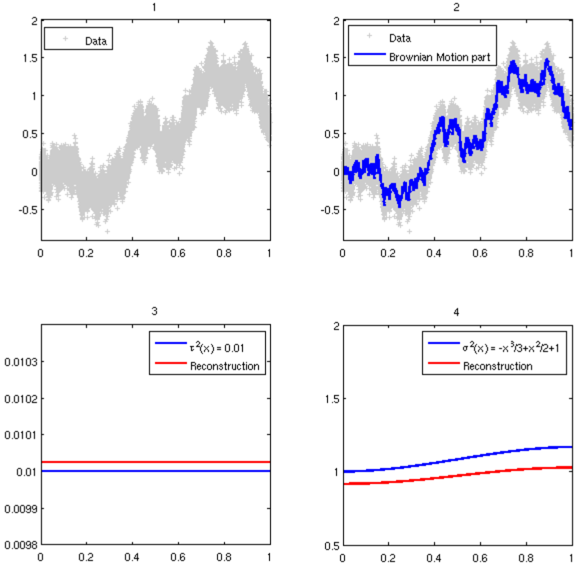
n=25.000, normal error, periodic bound., ∞-smooth



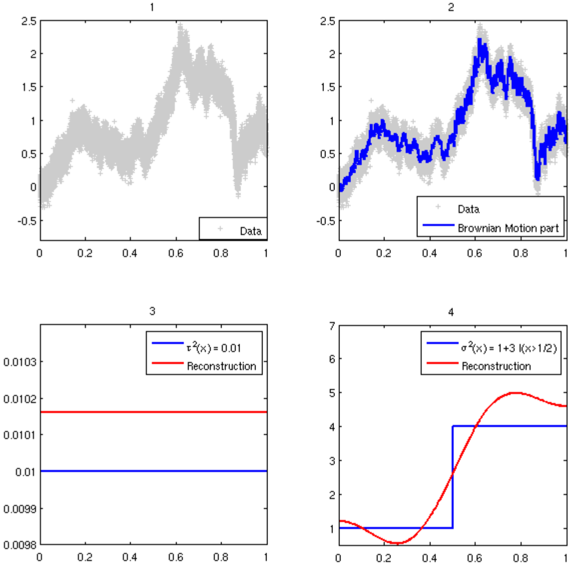
n=25.000, low smoothness: $\alpha < 3/2$



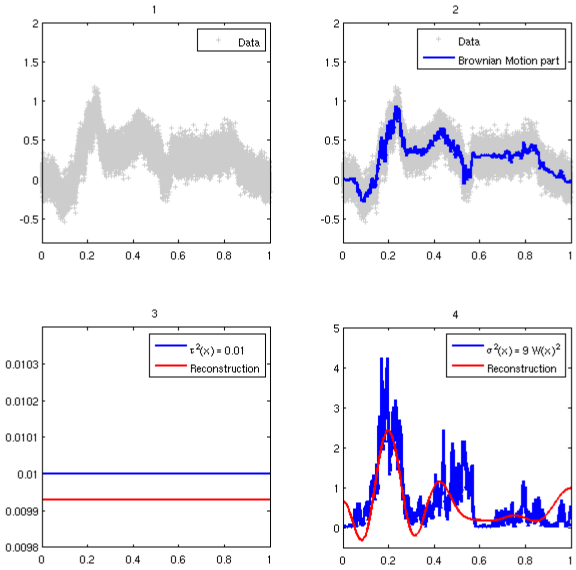
n=25.000, normal error, $\alpha < 7/2$



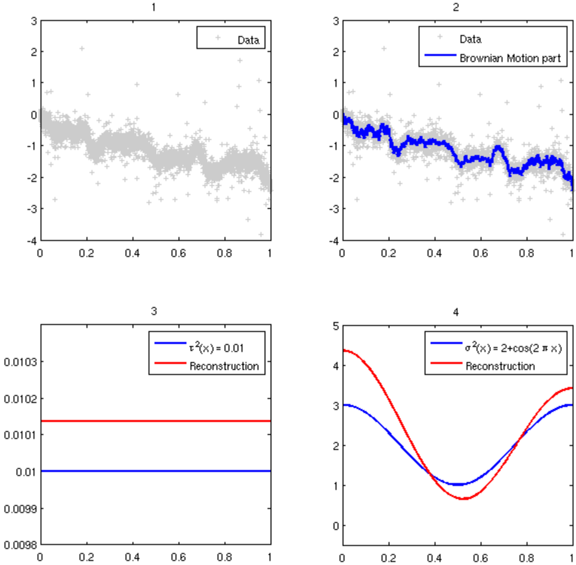
low smoothness: jump volatility



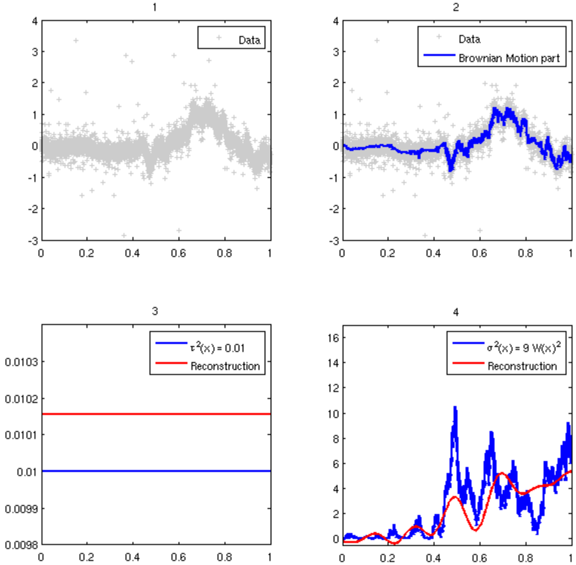
low smoothness: oscillating volatility



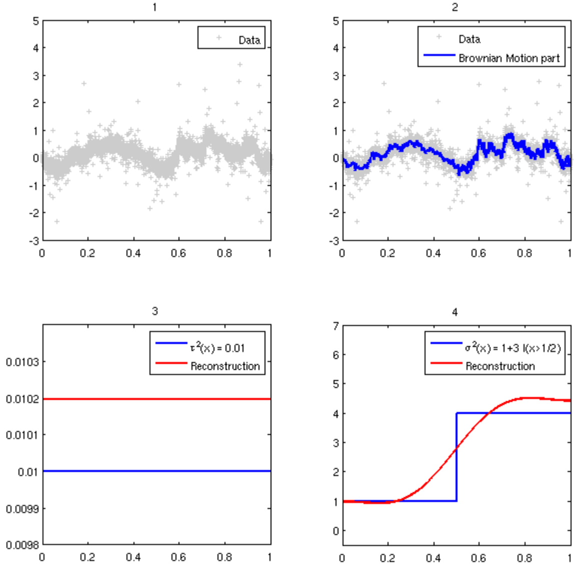
Robustness: t_2-distribution



Oscillating vol., t_2-distr.



Jump volatility, t_2-distr.



Summary

- Microstructure noise models with deterministic volatility provide some insight into volatility estimation.
- Viewing these models as a statistical inverse problem problems reveals similarities to deconvolution.
- Degree of ill posedness corresponds to $1/2$.
- Our approach relies heavily on Fourier methods and hence on a minimal smoothness of the estimated functions. Nevertheless, it seems quite robust.

Summary/Outlook

- Fourier type estimator achieves optimal (global) rates of convergence.
- Fast computable $O\left(Nn\log n\right)$.
- Open issues: Adaptation, locally adaptive basis, ...

References

- Ait-Sahalia, Y., Mykland, P. and Zhang, L. (2005). *How often to sample a continuous-time process in the presence of market microstructure noise.* Review of Financial Studies, Vol. 18, 351-416.
- Barndorff-Nielsen, O., Hansen, R., Lunde, A. and Stephard, N. (2006). *Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise.* Working Paper.
- Cai, T., Munk, A. and Schmidt-Hieber, J. (2008) *Sharp minimax estimation of the variance of Brownian motion corrupted by Gaussian noise.* Stat. Sinica., to appear.
- Fan, J., Jiang, J., Zhang, C. and Zhou, Z. (2003). *Time-dependent diffusion models for term structure dynamics.* Stat. Sinica 13, 965-992.
- Gloter, A. and Jacod, J. (2001). *Diffusions with measurement errors. I. Local Asymptotic Normality.* ESAIM: Probability and Statistics, 5, 225-242.
- Gloter, A. and Jacod, J. (2001). *Diffusions with measurement errors. II. Optimal estimators.* ESAIM: Probability ans Statistics, 5, 243-260.
- Golubev, G., Nussbaum, M. and Zhou, H. (2008). *Asymptotic equivalence of spectral density estimation and Gaussian white noise.* Submitted.
- Hoffmann, M. (1999). *Adaptive estimation in diffusion processes.* Stoch. Proc. and their Appl., 79, 135-163.
- Hoffmann, M. (2002). *Rate of convergence for parametric estimation in a stochastic volatility model.* Stoch. Proc. and their Appl., 97, 147-170.
- Malliavin, P. and Mancino, M. E. (2005) *Harmonic analysis methods for nonparametric estimation of volatility.* Preprint.
- Podolskij, M. and Vetter, M. (2009) *Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps.* Bernoulli, to appear.
- Reiß, M. (2008). *Asymptotic equivalence for nonparametric regression with multivariate and random design.* Annals of Statistics, to appear.

www.stochastik.math.uni-goettingen.de/munk
munk@math.uni-goettingen.de