Some functionals of Brownian motion connected with estimation of montone and convex functions

Joint work with:

Piet Groeneboom (Delft)
Moulinath Banerjee (Seattle)

Talk given at University of Washington, Seattle,
February 26, 2001.

Email: jaw@stat.washington.edu
http: //www.stat.washington.edu/jaw/jaw.research.html

OUTLINE:

1. **Introduction: the white-noise (Gaussian) problems**
   - Problem A: Estimator; monotone case.
   - Problem B: Likelihood ratio test; monotone case.
   - Problem C: Estimator; convex case.
2. **Estimator; monotone case**
3. **Likelihood ratio test; monotone case**
4. **Estimator; convex case**
5. **Further Problems**

1. Introduction: the white-noise (Gaussian) problems

Problem A. Monotone function in white noise: estimation

- $\{W(t) : t \in R\}$ be two-sided Brownian motion starting from 0;
- $dW(t) =$ "white noise" .
- Observe $X(t)$, $t \in [-c, c]$ or $t \in (-\infty, \infty)$ where

$$\begin{aligned} dX(t) &= f(t)dt + \sigma dW(t) \\ &= \text{"monotone function"} + \text{"white noise"}\,; \\ X(t) &= F(t) + \sigma W(t)\,. \end{aligned}$$

- Problem: estimate f.
- Likelihood: via Cameron-Martin

$$\frac{dP_{f,\sigma}}{dP_{0,\sigma}} = \exp\left(\int_{-c}^{c} f(t)dX(t) - \frac{1}{2}\int_{-c}^{c} f^2(t)dt\right)\,.$$

The MLE $\widehat{f}_c$ of f maximizes

$$(1)\ \int_{-c}^{c} f(t)dX(t) - \frac{1}{2}\int_{-c}^{c} f^2(t)dt$$

over monotone functions f. Equivalently, $\widehat{f}_c$ minimizes

$$\begin{aligned}(2) \quad & \int_{-c}^{c} f^2(t)dt - 2\int_{-c}^{c} f(t)dX(t) \\ & \text{“} = \text{”} \int_{-c}^{c} (f(t) - x(t))^2 dt - \int_{-c}^{c} x^2(t)dt\end{aligned}$$

where $x(t) = dX(t)/dt$ doesn't exist!

Problem B. Monotone function in white noise, testing:

- Observe

$$\begin{aligned} dX(t) &= f(t)dt + \sigma dW(t) \\ &= \text{“monotone function”} + \text{“white noise”} \end{aligned}$$

  as in Problem A.
- Problem: test $H_0: \ f(t_0) = \theta_0$ versus $H_1: \ f(t_0) \neq \theta_0$. Without loss, take $t_0 = 0$, $\theta_0 = 0$.

Step 1: The *constrained* MLE $\widehat{f}_c^0$ of f maximizes (1) subject to the constraint $f(0) = 0$.

Step 2: The likelihood ratio statistic is

$$\lambda_{c,\sigma} \equiv \frac{\sup_{f \in \mathcal{F}_c} dP_{f,\sigma}/dP_{0,\sigma}}{\sup_{f \in \mathcal{F}_c, f(0)=0} dP_{f,\sigma}/dP_{0,\sigma}} = \frac{dP_{\widehat{f}_c,\sigma}/dP_{0,\sigma}}{dP_{\widehat{f}_c^0,\sigma}/dP_{0,\sigma}}$$

Problem: Distribution of $2 \log \lambda_{c,\sigma}$?
Distribution of $\lim_{\sigma \to 0} 2 \log \lambda_{c,\sigma}$?

Problem C. Convex function in white noise: estimation

- $\{W(t) : t \in R\}$ be two-sided Brownian motion starting from 0;
- $dW(t) =$ "white noise" .
- Observe $X(t)$, $t \in [-c, c]$ or $t \in (-\infty, \infty)$ where

$$\begin{aligned} dX(t) &= f(t)dt + \sigma dW(t) \\ &= \text{"convex function"} + \text{"white noise"} . \end{aligned}$$

- Problem: estimate f.
- Likelihood: via Cameron-Martin

$$\frac{dP_{f,\sigma}}{dP_{0,\sigma}} = \exp\left(\int_{-c}^{c} f(t)dX(t) - \frac{1}{2}\int_{-c}^{c} f^2(t)dt \right) .$$

2. Estimator; monotone case (Problem A)

Let $\|g\|_c \equiv \sup_{t\in[-c,c]} |g(t)|$.

$\mathcal{F}_c = \{f : [-c,c] \mapsto R |\, f \text{ is monotone } \nearrow,\ \|f\|_c \le K\}$.

For $f \in \mathcal{F}_c$, set

$$F(t) = \begin{cases} \int_0^t f(s)ds, & t \ge 0 \\ -\int_t^0 f(s)ds, & t \le 0\,. \end{cases}$$

Theorem 1. $\widehat{f}_c$ is the Maximum Likelihood Estimator of f in $\mathcal{F}_c$ (and the Least Squares Estimator) if and only if:

(i) $\lambda_2 + \int_t^c d\{\widehat{F}_c(u) - X(u)\} \ge 0$ for all $t \in (-c,c]$;
(ii) $-K(\lambda_1+\lambda_2) - \int_{-c}^c \widehat{f}_c(u) d\{\widehat{F}_c(u - X(u)\} = 0$;
(iii) $\lambda_1 - \lambda_2 - \int_{-c}^c d\{\widehat{F}_c(u) - X(u)\} = 0$.

In fact

$$\begin{aligned} \lambda_1 &= \int_{\{u:\widehat{f}_c(u)=-K\}} d\{\widehat{F}_c(u) - X(u)\} \\ \lambda_2 &= \int_{\{u:\widehat{f}_c(u)=K\}} d\{\widehat{F}_c(u) - X(u)\}\,. \end{aligned}$$

Question:

$$\lim_{c\to\infty} \widehat{f}_c(t) = \ ???$$

Special case: $f_0(t) = 2t$, $F_0(t) = t^2$.

Theorem 2. For $X \sim P_{f_0}$ on $(-\infty, \infty)$, there is a uniquely defined random continuous function $\widehat{F} = \widehat{F}_\infty$ satisfying the following conditions:

(i) $\widehat{F}(t) \le X(t)$ for each $t \in R$.

(ii) $\widehat{F}$ has a monotone (left) derivative $\widehat{f}$.

(iii) $\widehat{F}$ satisfies

$$\int_R \{X(t) - \widehat{F}(t)\} d\widehat{f}(t) = 0\,.$$

$\widehat{F} =$ the greatest convex minorant of $X(t) = W(t) + t^2$.
$\widehat{f} \equiv \mathbb{S}$ the slope process of the greatest convex minorant $\widehat{F}$.

Question(s):

1. What is the distribution of the process $\widehat{f}$?
2. What is the distribution of $\widehat{f}(0)$?

Answers to 1 and 2:

Groeneboom (1985) and (1988)!
Groeneboom and Wellner (2000).

3. Likelihood ratio test: the monotone case (Problem B)

Theorem 3. There exists an almost surely uniquely defined random continuous function $\widehat{F}^0$ on $(-\infty, 0)$ and $(0, \infty)$ (and one jump at 0) satisfying the following conditions:

(i) $\widehat{F}^0(t) \le X(t)$ for each $t \in R$.

(ii) $\widehat{F}^0$ has a monotone (left) derivative $\widehat{f}^0$ satisfying $\widehat{f}^0(0) = 0$.

(iii) $\widehat{F}^0$ satisfies

$$\int_R \{X(t) - \widehat{F}^0(t)\} d\widehat{f}^0(t) = 0\,.$$

In fact, $\widehat{F}^0$ is (are) the *constrained greatest convex minorant(s)* of $X(t) = t^2 + W(t)$, and $\widehat{f}^0 \equiv \mathbb{S}^0$ is the corresponding *slope process(es)*.

Recall that

$$\lambda_{c,\sigma} \equiv \frac{\sup_{f \in \mathcal{F}_c} dP_{f,\sigma}/dP_{0,\sigma}}{\sup_{f \in \mathcal{F}_c, f(0)=0} dP_{f,\sigma}/dP_{0,\sigma}} = \frac{dP_{\widehat{f}_c,\sigma}/dP_{0,\sigma}}{dP_{\widehat{f}_c^0,\sigma}/dP_{0,\sigma}}$$

Theorem 4. (Banerjee and Wellner, 2000). Suppose that f satisfies $f(0) = 0$ and $f'(0) > 0$. Under $X \sim P_{f,\sigma}$,

$$2 \log \lambda_{c,\sigma} \to_d \mathbb{D} = \int \{(\mathbb{S}(z))^2 - (\mathbb{S}^0(z))^2\} dz$$

as $\sigma \to 0$.

Proof sketch: By Brownian scaling and approximation of f in a neighborhood of 0 by a linear function, the problem reduces to finding the limit of $2\log\lambda_{c,1}$ as $c \to \infty$ under the canonical monotone function $f_0(t) = 2t$. Then we have

$$
\begin{aligned}
2\log\lambda_{c,1} &= 2\left\{\int_{-c}^{c} \widehat{f}_c dX - \frac{1}{2}\int_{-c}^{c} \widehat{f}_c^2 dt \right. \\
&\qquad \left. - \int_{-c}^{c} \widehat{f}_c^0 dX - \frac{1}{2}\int_{-c}^{c} (\widehat{f}_c^0)^2 dt\right\} \\
&= 2\int_{-c}^{c} (\widehat{f}_c - \widehat{f}_c^0) dX - \int_{-c}^{c} \{\widehat{f}_c^2 - \widehat{f}_c^{02}\} dt \\
&\to 2\int_{-\infty}^{\infty} (\widehat{f} - \widehat{f}^0) dX - \int_{-\infty}^{\infty} \{\widehat{f}^2 - \widehat{f}^{02}\} dt \qquad (3)
\end{aligned}
$$

where $\widehat{f}$ and $\widehat{f}^0$ are characterized by Theorems 2 and 3 respectively. Thus

$$\int (X - \widehat{F}) d\widehat{f} = 0 \quad \text{and} \quad \int (X - \widehat{F}^0) d\widehat{f}^0 = 0\,;$$

Thus via integration by parts,

$$\int (\widehat{f} - \widehat{f}_0) dX = \int \{\widehat{f}^2 - \widehat{f}^{02}\} dt\,,$$

and using this in (3) shows that the limit equals

$$\int_{-\infty}^{\infty} \{\widehat{f}^2 - \widehat{f}^{02}\} dt = \int \{\{\mathbb{S}(t)\}^2 - \{\mathbb{S}^0(t)\}^2\} dt \equiv \mathbb{D}\,.$$

Question 3: What is the distribution of $\mathbb{D}$?
Only Monte-carlo evidence so far: Banerjee and Wellner (2000).

4. Estimator; convex case (Problem C)

Let $\|g\|_c \equiv \sup_{t\in[-c,c]} |g(t)|$.

$\mathcal{F}_c = \{f : [-c,c] \mapsto R |\, f \text{ is convex }, \, f(-c) = k_1, \, f(c) = k_2\}$.

For $f \in \mathcal{F}_c$, set

$$F(t) = \begin{cases} \int_0^t f(s)ds, & t \geq 0 \\ -\int_t^0 f(s)ds, & t \leq 0 \,, \end{cases}$$

and

$$H(t) = \begin{cases} \int_0^t F(s)ds = \int_0^t \int_0^s f(u)duds, & t \geq 0 \\ -\int_t^0 F(s)ds = -\int_t^0 \int_s^0 f(u)duds, & t \leq 0 \,, \end{cases}$$

For the canonical convex function $f_0(t) = 12t^2$,
$F_0(t) = 4t^3$, $H_0(t) = t^4$.
We also define

$$Y(t) \equiv \int_0^t X(s)ds = \int_0^t W(s)ds + H(t)$$

where now

$$\begin{aligned} dX(t) &= f(t) + dW(t) \\ &= \text{“convex”} + \text{“white noise”} \,. \end{aligned}$$

Theorem 5. $\widehat{f}_c$ is the Maximum Likelihood Estimor (MLE) (and the Least Squares Estimator) over $\mathcal{F}_c$ if and only if:
(i) $\widehat{H}_c(t) \geq Y(t)$ for $t \in [-c, c]$;
(ii) $\int_{(-c,c)} (\widehat{H}(t) - Y(t)) d\widehat{f}_c'(t) = 0$;
(iii) $\widehat{f}(-c) = k_1$, $\widehat{f}(c) = k_2$.

Question:

$$\lim_{c \to \infty} \widehat{f}_c(t) = \ ???$$

Special case:

$f_0(t) = 12t^2$, $F_0(t) = 4t^3$, $H_0(t) = t^4$.

Theorem 6. For $X \sim P_{f_0}$ on $(-\infty, \infty)$, there is a uniquely defined random continuous function $\widehat{H} = \widehat{H}_\infty$ satisfying the following conditions:
(i) $\widehat{H}(t) \geq Y(t)$ for each $t \in R$.
(ii) $\widehat{H}$ has a convex second derivative $\widehat{f}$.
(iii) $\widehat{H}$ satisfies

$$\int_R \{\widehat{H}(t) - Y(t)\} d\widehat{H}^{(3)}(t) = 0\,.$$

$\widehat{H}$ = the "invelope process" of Y, is piecewise cubic;
$\widehat{F} = \widehat{H}'$ is piecewise quadratic;
$\widehat{f} = \widehat{H}''$ is piecewise linear (and convex);
$\widehat{f'} = \widehat{H}^{(3)}$ is piecewise constant (and "estimates" $f_0'(t) = 24t$).

Questions:

4. What is the distribution of the process $\widehat{f}$?
5. What is the distribution of $\widehat{f}(0)$?
6. What is the joint distribution of $(\widehat{f}(0), \widehat{f'}(0))$?

No answers to 4-6 yet.

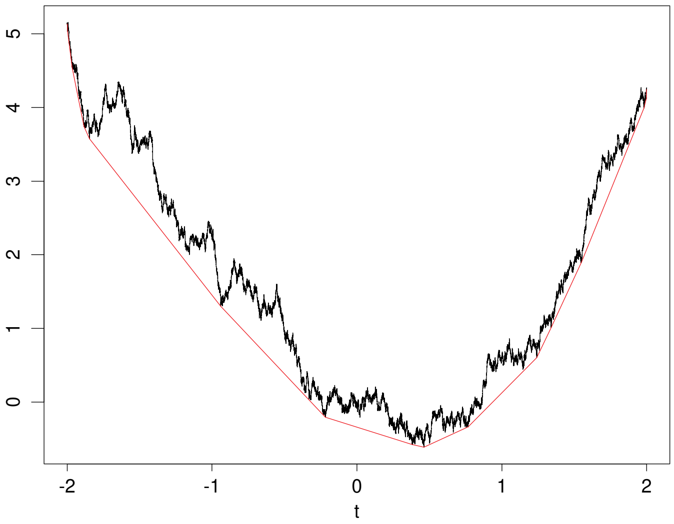

Figure 1: Greatest Convex Minorant and $W(t) + t^2$.

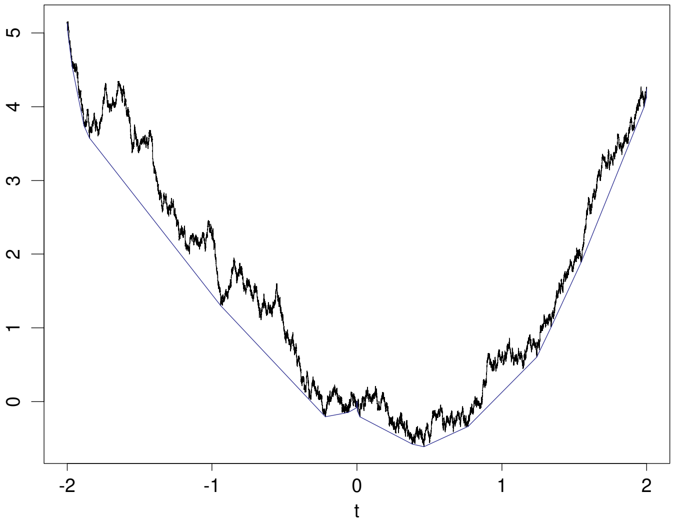

Figure 2: One-Sided Greatest Convex Minorants and $W(t) + t^2$.

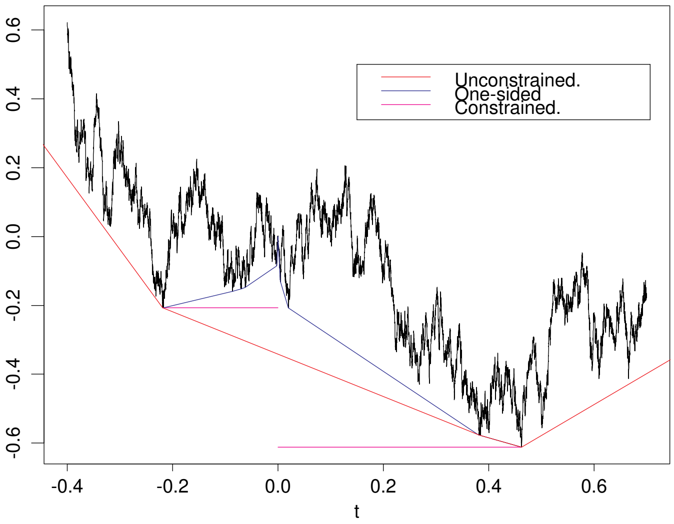

Figure 3: Close-up view of $G_{1,1}$, $\tilde{G}_{L,R}$, $G^0_{1,1}$, and $W(t) + t^2$.

5. Further Problems.

Question 3. Can we find the distribution of $\mathbb{D}$ analytically? (Likelihood ratio statistic)

Question 4. What is the distribution of the process $\widehat{f}$? (convex function estimation)

Queston 5. What is the distribution of $\widehat{f}(0)$? (convex function estimation)

Question 6. What is the joint distribution of $(\widehat{f}(0), \widehat{f}'(0))$? (convex function estimation)

Question 7. What is the likelihood ratio test for testing $f(t_0) = \theta_0$ in the convex function case?

Question 8. How can we prove that the same limit $\mathbb{D}$ arises as the limit distribution for the likelihood ratio test for a large class of such problems involving monotone functions?

Question 9. What is the appropriate contiguity theory? What is the limit distribution of the likelihood ratio statistic under local alternatives?

Question 10. What about further smoothness restrictions? (Completely monotone?)

Selected References:

- Banerjee, M. and Wellner, J. A. (2000). Likelihood ratio tests for monotone functions. *Technical Report* **377**, Department of Statistics, University of Washington.
- Groeneboom, P. (1985). Estimating a monotone density. *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer,* Vol. II, Lucien M. LeCam and Richard A. Olshen eds.
- Groeneboom, P. (1988). Brownian motion with a parabolic drift and Airy functions. *Probability Theory and Related Fields* **81**, 79 - 109.
- Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2000a). A canonical process for estimation of convex functions: The "invelope" of integrated Brownian motion $+t^4$. *Technical Report* **369**, Department of Statistics, University of Washington.
- Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2000b). Estimation of convex functions: characterizations and asymptotic theory. *Technical Report* **372**, Department of Statistics, University of Washington.
- Groeneboom, P. and Wellner, J. A. (2000). Computing Chernoff's distribution. *J. Computational and Graphical Statistics*, to appear.

- Wellner, J. A. (2000). Gaussian white noise models: some results for monotone functions. *Technical Report* **??**, Department of Statistics, University of Washington.