Nonparametric estimation of log-concave densities

Jon A. Wellner

University of Washington

- Talk at U. C. Berkeley
  November 4, 2009
- *Email: jaw@stat.washington.edu*
  http: //www.stat.washington.edu/jaw/jaw.research.html
- Based on joint work with Fadoua Balabdaoui, Kaspar Rufibach, and Arseni Seregin

Outline

1. Log-concave densities on $\mathbb{R}^1$
2. Nonparametric estimation, log-concave on $\mathbb{R}^1$
3. Limit theory at a fixed point in $\mathbb{R}^1$
4. Estimation of the mode, log-concave density on $\mathbb{R}^1$
5. Generalizations: $s-$concave densities on $\mathbb{R}^1$ and $\mathbb{R}^d$
6. Summary; problems and open questions

1. Log-concave densities on $\mathbb{R}^1$

Suppose that

$$p(x) = \exp(\varphi(x)) = \exp\left(-(-\varphi(x))\right)$$

where φ is concave (and $-\varphi$ is convex). The class of all densities p on $\mathbb{R}$ of the form is called the class of *log-concave* densities, $\mathcal{P}_{log-concave}$.

Properties of log-concave densities:

- A density p on $\mathbb{R}$ is log-concave if and only if its convolution with any unimodal density is again unimodal (Ibragimov, 1956).

1. Log-concave densities on $\mathbb{R}^1$

Suppose that

$$p(x) = \exp(\varphi(x)) = \exp\left(-(-\varphi(x))\right)$$

where φ is concave (and $-\varphi$ is convex). The class of all densities p on $\mathbb{R}$ of the form is called the class of *log-concave* densities, $\mathcal{P}_{log-concave}$.

Properties of log-concave densities:

- A density p on $\mathbb{R}$ is log-concave if and only if its convolution with any unimodal density is again unimodal (Ibragimov, 1956).
- Every log-concave density p is unimodal (but need not be symmetric).

- Many parametric families are log-concave, for example:
  - Normal (μ, σ^2)
  - Uniform(a, b)
  - Gamma(r, λ) for $r \geq 1$
  - Beta(a, b) for $a, b \geq 1$

- Many parametric families are log-concave, for example:
  - Normal (μ, σ^2)
  - Uniform(a, b)
  - Gamma(r, λ) for $r \geq 1$
  - Beta(a, b) for $a, b \geq 1$
- t_r densities with $r > 0$ are not log-concave

- Many parametric families are log-concave, for example:
  - Normal (μ, σ^2)
  - Uniform(a, b)
  - Gamma(r, λ) for $r \geq 1$
  - Beta(a, b) for $a, b \geq 1$
- t_r densities with $r > 0$ are not log-concave
- Tails of log-concave densities are necessarily sub-exponential

- Many parametric families are log-concave, for example:
    - Normal (μ, σ^2)
    - Uniform(a, b)
    - Gamma(r, λ) for $r \geq 1$
    - Beta(a, b) for $a, b \geq 1$
- t_r densities with $r > 0$ are not log-concave
- Tails of log-concave densities are necessarily sub-exponential
- $\mathcal{P}_{log-concave} =$ the class of "Polyá frequency functions of order 2", PFF_2, in the terminology of Schoenberg (1951) and Karlin (1968). See Marshall and Olkin (1979), chapter 18, and Dharmadhikari and Joag-Dev (1988), page 150. for nice introductions.

2. Nonparametric estimation, log-concave on $\mathbb{R}^1$

- The (nonparametric) MLE $\widehat{p}_n$ exists (Rufibach, Dümbgen and Rufibach).

2. Nonparametric estimation, log-concave on $\mathbb{R}^1$

- The (nonparametric) MLE $\widehat{p}_n$ exists (Rufibach, Dümbgen and Rufibach).
- $\widehat{p}_n$ can be computed: R-package “logcondens” (Dümbgen and Rufibach)

2. Nonparametric estimation, log-concave on $\mathbb{R}^1$

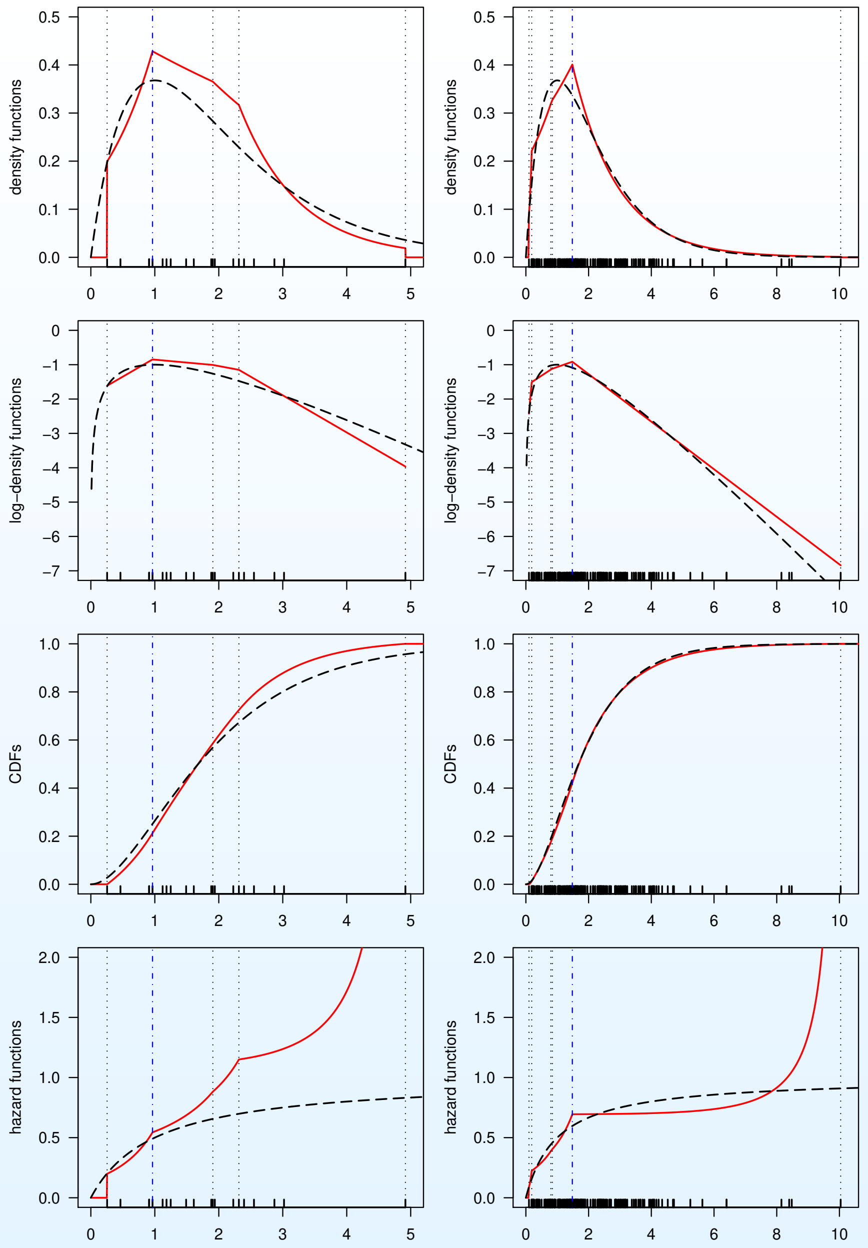
- The (nonparametric) MLE $\widehat{p}_n$ exists (Rufibach, Dümbgen and Rufibach).
- $\widehat{p}_n$ can be computed: R-package “logcondens” (Dümbgen and Rufibach)
- In contrast, the (nonparametric) MLE for the class of unimodal densities on $\mathbb{R}^1$ does not exist. Birgé (1997) and Bickel and Fan (1996) consider alternatives to maximum likelihood for the class of unimodal densities.

2. Nonparametric estimation, log-concave on $\mathbb{R}^1$

- The (nonparametric) MLE $\widehat{p}_n$ exists (Rufibach, Dümbgen and Rufibach).
- $\widehat{p}_n$ can be computed: R-package “logcondens” (Dümbgen and Rufibach)
- In contrast, the (nonparametric) MLE for the class of unimodal densities on $\mathbb{R}^1$ does not exist. Birgé (1997) and Bickel and Fan (1996) consider alternatives to maximum likelihood for the class of unimodal densities.
- Consistency and rates of convergence for $\widehat{p}_n$:
  Dümbgen and Rufibach, (2007); Pal, Woodroofe and Meyer (2007).

2. Nonparametric estimation, log-concave on $\mathbb{R}^1$

- The (nonparametric) MLE $\widehat{p}_n$ exists (Rufibach, Dümbgen and Rufibach).
- $\widehat{p}_n$ can be computed: R-package "logcondens" (Dümbgen and Rufibach)
- In contrast, the (nonparametric) MLE for the class of unimodal densities on $\mathbb{R}^1$ does not exist. Birgé (1997) and Bickel and Fan (1996) consider alternatives to maximum likelihood for the class of unimodal densities.
- Consistency and rates of convergence for $\widehat{p}_n$: Dümbgen and Rufibach, (2007); Pal, Woodroofe and Meyer (2007).
- Pointwise limit theory? Yes! Balabdaoui, Rufibach, and Wellner (2007).


density functions
log−density functions
CDFs
hazard functions

- Assumptions

- Assumptions
    - p is log-concave, $\quad p(x_0) > 0$.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C \exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C \exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.
- Process H_k determined by limit Fenchel relations:

- Assumptions
  - p is log-concave, $p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.
- Process H_k determined by limit Fenchel relations:
  - $H_k(t) \leq Y_k(t)$ for all $t \in \mathbb{R}$

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.
- Process H_k determined by limit Fenchel relations:
  - $H_k(t) \leq Y_k(t)$ for all $t \in \mathbb{R}$
  - $\int_{\mathbb{R}} (H_k(t) - Y_k(t)) dH_k^{(3)}(t) = 0$.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.
- Process H_k determined by limit Fenchel relations:
  - $H_k(t) \leq Y_k(t)$ for all $t \in \mathbb{R}$
  - $\int_{\mathbb{R}} (H_k(t) - Y_k(t))dH_k^{(3)}(t) = 0$.
  - $H_k^{(2)}$ is concave.

- Assumptions
  - p is log-concave, $\quad p(x_0) > 0$.
  - If $\varphi''(x_0) \neq 0$, then $k = 2$;
    otherwise, k is the smallest integer such that
    $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi^{(k)}(x_0) \neq 0$.
  - $\varphi^{(k)}$ is continuous in a neighborhood of x_0.
- Example $p(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.
- Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.
- Process H_k determined by limit Fenchel relations:
  - $H_k(t) \leq Y_k(t)$ for all $t \in \mathbb{R}$
  - $\int_{\mathbb{R}} (H_k(t) - Y_k(t))dH_k^{(3)}(t) = 0$.
  - $H_k^{(2)}$ is concave.

- Pointwise limit theorem for $\widehat{p}_n(x_0)$:

$$\begin{pmatrix} n^{k/(2k+1)}(\widehat{p}_n(x_0) - p(x_0)) \\ n^{(k-1)/(2k+1)}(\widehat{p}_n'(x_0) - p'(x_0)) \end{pmatrix} \to_d \begin{pmatrix} c_k H_k^{(2)}(0) \\ d_k H_k^{(3)}(0) \end{pmatrix}$$

where

$$c_k \equiv \left(\frac{p(x_0)^{k+1} |\varphi^{(k)}(x_0)|}{(k+2)!} \right)^{1/(2k+1)},$$

$$d_k \equiv \left(\frac{p(x_0)^{k+2} |\varphi^{(k)}(x_0)|^3}{[(k+2)!]^3} \right)^{1/(2k+1)}.$$

- Pointwise limit theorem for $\widehat{\varphi}_n(x_0)$:

$$\begin{pmatrix} n^{k/(2k+1)}(\widehat{\varphi}_n(x_0)-\varphi(x_0)) \\ n^{(k-1)/(2k+1)}(\widehat{\varphi}'_n(x_0)-\varphi'(x_0)) \end{pmatrix} \to_d \begin{pmatrix} C_k H_k^{(2)}(0) \\ D_k H_k^{(3)}(0) \end{pmatrix}$$

where

$$C_k \equiv \left(\frac{|\varphi^{(k)}(x_0)|}{p(x_0)^k (k+2)!} \right)^{1/(2k+1)},$$

$$D_k \equiv \left(\frac{|\varphi^{(k)}(x_0)|^3}{p(x_0)^{k+1}[(k+2)!]^3} \right)^{1/(2k+1)}.$$

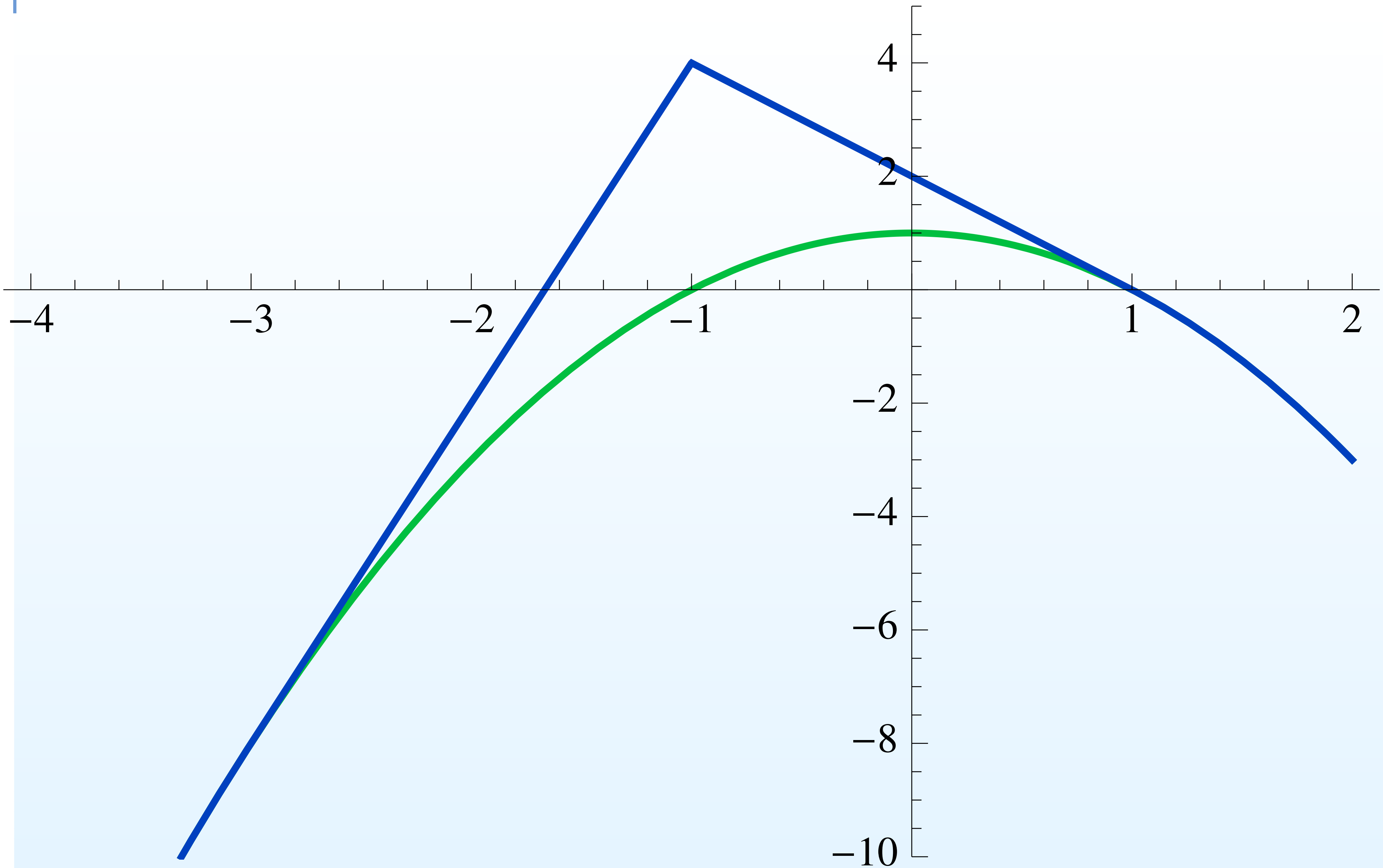
3. Estimation of the mode

Let $x_0 = M(p_0)$ be the *mode* of the log-concave density p_0, recalling that $\mathcal{P}_{log-concave} \subset \mathcal{P}_{unimodal}$. Lower bound calculations using G. Jongbloed's perturbation of a convex decreasing density, but now perturbing φ_0 yields:

Proposition. If $p_0 \in \mathcal{P}_{log-concave}$ satisfies $p_0(x_0) > 0$, $p_0''(x_0) < 0$, and p_0'' is continuous in a neighborhood of x_0, and T_n is any estimator of the mode $x_0 \equiv M(p_0)$, then with P_n corresponding to $p_{\epsilon_n} \equiv \exp(\varphi_{\epsilon_n})$ with $\epsilon_n \equiv \nu n^{-1/5}$ and $\nu \equiv 2p_0''(x_0)^2/(5p_0(x_0))$,

$$\liminf_{n\to\infty} n^{1/5} \inf_{T_n} \max\left\{E_{n,P_n}|T_n - M(p_n)|, E_{n,P}|T_n - M(p_0)|\right\}$$

$$\geq \frac{1}{4}\left(\frac{5/2}{10e}\right)^{1/5}\left(\frac{p_0(x_0)}{p_0''(x_0)^2}\right)^{1/5}.$$


4
2
−2
−4
−6
−8
−10
−4
−3
−2
−1
1
2

On the other hand, the limit theory of Balabdaoui, Rufibach, and Wellner (2007) noted in the previous section implies that the mode estimator derived from the MLE $\widehat{p}_n$ of p, namely $\widehat{M}_n \equiv M(\widehat{p}_n) \equiv \min\{u : \widehat{p}_n(u) = \sup_t \widehat{p}_n(t)\}$, satisfies, assuming that

- $\varphi^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$,
- $\varphi^{(k)}(x_0) \neq 0$, and
- $\varphi^{(k)}$ is continuous in a neighborhood of x_0;

$$n^{1/(2k+1)}(\widehat{M}_n - M(p_0)) \to_d \left(\frac{(4!)^2 p_0(x_0)}{p_0''(x_0)^2} \right)^{1/(2k+1)} M(H_k^{(2)})$$

where $M(H_k^{(2)}) = \text{argmax}(H_k^{(2)})$.
Note that when $k = 2$ this agrees with the lower bound calculation, at least up to absolute constants.

4. Generalizations: $s-$concave densities on $\mathbb{R}^d$

- Three generalizations:
  - $\log -$concave densities on $\mathbb{R}^d$ (Cule, Samworth, and Stewart, 2008)
  - $s-$concave densities on $\mathbb{R}^d$ (Seregin, 2009)
  - $h-$transformed convex densities on $\mathbb{R}^d$ (Seregin, 2009)

4. Generalizations: $s-$concave densities on $\mathbb{R}^d$

- Three generalizations:
  - $\log -$concave densities on $\mathbb{R}^d$ (Cule, Samworth, and Stewart, 2008)
  - $s-$concave densities on $\mathbb{R}^d$ (Seregin, 2009)
  - $h-$transformed convex densities on $\mathbb{R}^d$ (Seregin, 2009)
- A density p on $\mathbb{R}^d$ is log-concave if $p(x) = \exp(\varphi(x))$ with φ concave.

4. Generalizations: $s-$concave densities on $\mathbb{R}^d$

- Three generalizations:
  - $\log -$concave densities on $\mathbb{R}^d$ (Cule, Samworth, and Stewart, 2008)
  - $s-$concave densities on $\mathbb{R}^d$ (Seregin, 2009)
  - $h-$transformed convex densities on $\mathbb{R}^d$ (Seregin, 2009)
- A density p on $\mathbb{R}^d$ is log-concave if $p(x) = \exp(\varphi(x))$ with φ concave.
- Some properties:
  - Any $\log -$concave p is unimodal
  - The level sets are closed convex sets
  - Convolutions of log-concave distributions are log-concave
  - Marginals of log-concave distributions are log-concave

What is known so far concerning nonparametric estimation of a log-concave density p on $\mathbb{R}^d$?

- The MLE $\widehat{p}_n$ is well-defined, with support the convex hull of the observed data, for $n > d$.

What is known so far concerning nonparametric estimation of a log-concave density p on $\mathbb{R}^d$?

- The MLE $\widehat{p}_n$ is well-defined, with support the convex hull of the observed data, for $n > d$.
- Tails of log-concave decay at least exponentially fast.

What is known so far concerning nonparametric estimation of a log-concave density p on $\mathbb{R}^d$?

- The MLE $\widehat{p}_n$ is well-defined, with support the convex hull of the observed data, for $n > d$.
- Tails of log-concave decay at least exponentially fast.
- Computable for small d via Shor's $r-$algorithm: Cule, Samworth, and Stewart (2008); R-package LogConcDEAD.

What is known so far concerning nonparametric estimation of a log-concave density p on $\mathbb{R}^d$?

- The MLE $\widehat{p}_n$ is well-defined, with support the convex hull of the observed data, for $n > d$.
- Tails of log-concave decay at least exponentially fast.
- Computable for small d via Shor's $r-$algorithm: Cule, Samworth, and Stewart (2008); R-package LogConcDEAD.
- Consistency and limit under model miss-specification: Samworth (2009), August Oberwolfach talk.
  When the true p is not log-concave, then $\widehat{p}_n$ converges to the closest p in $\mathcal{P}_{\log -concave}$ in the Kullback-Leibler sense.

What is known so far concerning nonparametric estimation of a log-concave density p on $\mathbb{R}^d$?

- The MLE $\widehat{p}_n$ is well-defined, with support the convex hull of the observed data, for $n > d$.
- Tails of log-concave decay at least exponentially fast.
- Computable for small d via Shor's $r-$algorithm: Cule, Samworth, and Stewart (2008); R-package LogConcDEAD.
- Consistency and limit under model miss-specification: Samworth (2009), August Oberwolfach talk.
  When the true p is not log-concave, then $\widehat{p}_n$ converges to the closest p in $\mathcal{P}_{\log - concave}$ in the Kullback-Leibler sense.
- Some promising applications: Cule, Samworth, Stewart (2008); Walther (2009).

Generalization to $r-$concave densities: A density p on $\mathbb{R}^d$ is $r-$concave on $C \subset \mathbb{R}^d$ if

$$p(\lambda x + (1-\lambda)y) \geq M_r(p(x), p(y); \lambda)$$

for all $x, y \in C$ and $0 < \lambda < 1$ where

$$M_r(a, b; \lambda) = \begin{cases} ((1-\lambda)a^r + \lambda b^r)^{1/r}, & r \neq 0, a, b > 0, \\ 0, & r < 0, ab = 0 \\ a^{1-\lambda} b^{\lambda}, & r = 0. \end{cases}$$

Let $\mathcal{P}_r$ denote the class of all $r-$concave densities on C. For $r \leq 0$ it suffices to consider $C = \mathbb{R}^d$, and it is almost immediate from the definitions that if $p \in \mathcal{P}_r$ for some $r \leq 0$, then

$$p(x) = \left\{ \begin{array}{ll} g(x)^{1/r}, & r < 0 \\ \exp(-g(x)), & r = 0 \end{array} \right\} \qquad \text{for } g \text{ convex.}$$

- Long history: Avriel (1972), Prékopa (1973), Borell (1975), Rinott (1976), Brascamp and Lieb (1976)
- Nice connections to $t-$concave measures.
- Known now in math-analysis as the Borell, Brascamp, Lieb inequality
- One way to get heavier tails than log-concave!

This motivates the following definitions:

Definition 1. (Seregin) Say that $h : \mathbb{R} \to \mathbb{R}^+$ is a decreasing transformation if, with $y_0 \equiv \sup\{y : \ h(y) > 0\}$, $y_\infty \equiv \inf\{y : h(y) < \infty\}$,

- $h(y) = o(y^{-\alpha})$ for some $\alpha > d$ as $y \to \infty$.
- If $y_\infty > -\infty$, then $h(y) \asymp (y - y_\infty)^{-s}$ for some $s > d$ as $y \searrow y_\infty$.
- If $y_\infty = -\infty$, then $h(y)^\gamma h(-Cy) = o(1)$ as $y \to -\infty$ for some $\gamma, C > 0$.
- h is continuously differentiable on (y_∞, y_0).

Examples. $h(x) = x^{-s}$ with $s > 0$ and $h(x) = \exp(-x)$ are both decreasing transformations.

For the definition of increasing transformations, let $y_0 \equiv \inf\{y : \ h(y) > 0\}$ and $y_\infty \equiv \sup\{y : h(y) < \infty\}$.

Definition 2. (Seregin) Say that $h : \mathbb{R} \to \mathbb{R}^+$ is a increasing transformation if

- $h(y) = o(|y|^{-\alpha})$ for some $\alpha > d$ as $y \to -\infty$.
- If $y_\infty < \infty$, then $h(y) \asymp (y_\infty - y)^{-s}$ for some $s > d$ as $y \nearrow y_\infty$.
- h is continuously differentiable on (y_0, y_∞).

Examples. $h(x) = \max\{x, 0\}$ and $h(x) = \exp(x)$ are both increasing transformations.

Definition 3. (Seregin) For h either a decreasing transformation and $C = \mathbb{R}^d$ or h an increasing transformation and $C = \mathbb{R}^d_+$, define the class of convex - transformed densities $\mathcal{P}_h$ to be the collection of all densities of the form

$$p(x) \equiv p_g(x) = h(g(x))1_C(x), \ \ x \in \mathbb{R}^d, \ \ g \ \text{ convex.}$$

Theorem 1. (Seregin).
• For h an increasing transformation the MLE $\widehat{p}_n$ for $\mathcal{P}_h$ exists almost surely.
• For h a decreasing transformation, the MLE $\widehat{p}_n$ for $\mathcal{P}_h$ exists if $n \ge d(1+\gamma)$ if $y_\infty = -\infty$ and if $n \ge d + sd^2/(\alpha(s-d))$ if $y_\infty > -\infty$.

Theorem 2. (Seregin).
• For any decreasing model $\mathcal{P}_h$ the sequence of maximum likelihood estimators $\widehat{p}_n = h \circ \widehat{g}_n$ is Hellinger consistent: with $H^2(P,Q) = 2^{-1} \int \{\sqrt{dP} - \sqrt{dQ}\}^2$,

$$H(\widehat{p}_n, p_0) \rightarrow_{a.s.} 0. \tag{1}$$

• Suppose that for a decreasing model $\mathcal{P}_h$ we have $p_0 = h(g_0)$ satisfying
(a) g_0 is bounded.
(b) If $d > 1$ then $\int_{\mathbb{R}_+^d} \log(1/(|\underline{x}| \wedge 1)) p_0(x) dx < \infty$.
Then the MLE $\widehat{p}_n$ over $\mathcal{P}_h$ is Hellinger consistent; i.e. (1) holds.

5. Summary; problems and open questions

- Summary for $\mathbb{R}^1$:

5. Summary; problems and open questions

- Summary for $\mathbb{R}^1$:
  - Reasonably good understanding of basics of nonparametric estimation of log-concave densities on $\mathbb{R}$.

5. Summary; problems and open questions

- Summary for $\mathbb{R}^1$:
  - Reasonably good understanding of basics of nonparametric estimation of log-concave densities on $\mathbb{R}$.
  - Recent progress due to Rufibach (2007); Duembgen and Rufibach (2008); Pal, Woodroofe, and Meyer (2007); Balabdaoui, Rufibach, and Wellner (2009).

5. Summary; problems and open questions

- Summary for $\mathbb{R}^1$:
  - Reasonably good understanding of basics of nonparametric estimation of log-concave densities on $\mathbb{R}$.
  - Recent progress due to Rufibach (2007); Duembgen and Rufibach (2008); Pal, Woodroofe, and Meyer (2007); Balabdaoui, Rufibach, and Wellner (2009).
  - Interesting applications of theory even for $\mathbb{R}$; e.g. Walther (2001, 2002)

5. Summary; problems and open questions

- Summary for $\mathbb{R}^1$:
  - Reasonably good understanding of basics of nonparametric estimation of log-concave densities on $\mathbb{R}$.
  - Recent progress due to Rufibach (2007); Duembgen and Rufibach (2008); Pal, Woodroofe, and Meyer (2007); Balabdaoui, Rufibach, and Wellner (2009).
  - Interesting applications of theory even for $\mathbb{R}$; e.g. Walther (2001, 2002)
  - Quite a few problems remaining.

- Summary for $\mathbb{R}^d$:

- Summary for $\mathbb{R}^d$:
  - First steps toward estimation of log-concave densities on $\mathbb{R}^d$ by Cule, Samworth, and Stewart (2008).

- Summary for $\mathbb{R}^d$:
  - First steps toward estimation of log-concave densities on $\mathbb{R}^d$ by Cule, Samworth, and Stewart (2008).
  - First consistency proof, log-concave case (perhaps with miss-specified model): Samworth (2009).

- Summary for $\mathbb{R}^d$:
    - First steps toward estimation of log-concave densities on $\mathbb{R}^d$ by Cule, Samworth, and Stewart (2008).
    - First consistency proof, log-concave case (perhaps with miss-specified model): Samworth (2009).
    - Generalizations to $r-$ concave classes with general $r \leq 0$ by Seregin (2009); he proves:
        - Existence of MLEs for both decreasing and increasing convex - transformed models.
        - Consistency of MLEs for both decreasing and increasing convex-transformed models.
        - Asymptotic minimax lower bounds for estimation of $p(x_0)$ and $M(p_0)$ in monotone-transformed convex function models.

- Problems for $\mathbb{R}$
  - Asymptotic behavior of smooth functionals: asymptotic equivalence to usual empirical estimators under minimal assumptions?
  - Global rates of convergence?
  - Limit distributional results under model miss-specification?

- Problems for $\mathbb{R}$
  - Asymptotic behavior of smooth functionals: asymptotic equivalence to usual empirical estimators under minimal assumptions?
  - Global rates of convergence?
  - Limit distributional results under model miss-specification?
- Problems for $\mathbb{R}^d$: many!
  - Faster algorithms for log-concave case?
  - Any algorithm for the $r-$concave case and for increasing models?
  - Rates of convergence of estimators at fixed points?
  - Limiting distributions at fixed points?
  - MLE's will probably be rate inefficient for $d \geq 4$. Two questions:
    (a) How to penalize or sieve to get estimators within the classes which achieve the optimal rates?
    (b) How to define interesting or natural smaller subclasses for which MLE's remain optimal?