Nonparametric estimation
under Shape Restrictions

Jon A. Wellner

University of Washington, Seattle

Statistical Seminar, Frejus, France
August 30 - September 3, 2010
Outline: Five Lectures on Shape Restrictions

• L1: Monotone functions: maximum likelihood and least squares
• L2: Optimality of the MLE of a monotone density (and comparisons?)
• L3: Estimation of convex and $k$–monotone density functions
• L4: Estimation of log-concave densities: $d = 1$ and beyond
• L5: More on higher dimensions ... ... and some open problems
Outline: Lecture 5

• A: Some multivariate shape-constrained classes
  ▶ Log-concave & $h$–convex on $\mathbb{R}^d$
  ▶ Block-decreasing on $\mathbb{R}_d$
  ▶ Scale mixtures of uniform on $\mathbb{R}_d$
  ▶ $h$–convex on $\mathbb{R}_d$ with $h$ increasing.

• B: Review of available theory; MLEs for multivariate classes.

• C: Some alternative classes in $\mathbb{R}$:
  Bondesson’s hyperbolically monotone classes.

• D: Open problems and questions: $\mathbb{R}^1$

• E: Some problems and questions: $\mathbb{R}^d$
A. Some multivariate shape-constrained classes

- “Block decreasing” densities on $R^{+d} = [0, \infty)^d$
- Monotone with non-negative increments on rectangles (as for a multivariate d.f.)
- Convex and decreasing
- $k$-monotone; completely monotone
- log–concave; $s$–concave; $h$– transform of convex (or concave)
A. Some multivariate shape-constrained classes

Block-decreasing densities on \( \mathbb{R}^{+d} = [0, \infty)^d \)

- \( \mathcal{BD}(\mathbb{R}^d) = \left\{ f : [0, \infty)^d \rightarrow \mathbb{R}^{+} \middle| \int f(x)dx = 1, f(x + he_j) \leq f(x) \text{ for all basis vectors } e_j, j = 1, \ldots, d, h > 0 \right\} \).


- Biau and Devroye (2003): global minimax lower bounds ... and showed that a generalization of Birgé’s histogram estimator achieves the bounds.

- Pavlides (2008, 2009): asymptotic minimax lower bounds for estimation of \( f(x_0) \)
A. Some multivariate shape-constrained classes
A. Some multivariate shape-constrained classes
A. Some multivariate shape-constrained classes

Monotone with non-negative increments on rectangles
(as for a multivariate d.f.)

="Scale mixtures of uniform densities" on $\mathbb{R}^{+d}$

$$f(x) = \int_{\mathbb{R}^{+d}} \frac{1}{\prod_{j=1}^{d} y_j} 1_{[0,y]}(x) dG(y)$$

for some probability distribution $G$ on $\mathbb{R}^{+d}$.

Example: $dG(y_1, y_2) = (y_1 y_2)^{-2} g(1/y_1, 1/y_2, \theta) dy_1 dy_2$ with

$$g(u, v, \theta) = \{(1 + \theta u)(1 + \theta v) - \theta\} \exp(-u - v - \theta uv), \quad \theta = .4,$$
A. Some multivariate shape-constrained classes
A. Some multivariate shape-constrained classes

Convex and decreasing on \( \mathbb{R}^d \)

- Seregin (2010)'s increasing convex transformed classes with \( h(x) = x \), so that \( f(x) = \varphi(x) \) with \( \varphi \) convex (or convex and decreasing).

Example: \( f(x) = \exp(-|x|)1_{(0,\infty)^d}(x) \).
A. Some multivariate shape-constrained classes

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

- \[ f(x) = \exp(\varphi(x)) = \exp(-(-\varphi(x))) \]
  where $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ is concave (so $-\varphi$ is convex).

- Exponentially decaying tails; does not include multivariate $t$-densities.
A. Some multivariate shape-constrained classes

- $s$–convex densities and $h$– convex densities
  (Koenker and Mizera; Seregin, Seregin and Wellner)

\[ f(x) = h(\varphi(x)) \]

where $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ is convex, $h : \mathbb{R} \mapsto \mathbb{R}^+$ is decreasing and continuous; e.g. $h_s(u) \equiv (1 + u/s)^{-s}$ with $s > d$. Larger classes than log-concave: includes multivariate $t_n$ for $d < s \leq n + d$. 
**B: Review of available theory:**

**MLEs for multivariate classes**

Block decreasing densities on $\mathbb{R}^+^d$

<table>
<thead>
<tr>
<th>Problem</th>
<th>Lower Bound</th>
<th>Upper Bound / MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B (local)</td>
<td>$\prod_{j=1}^{d} \left( \frac{\partial f}{\partial x_j}(x)f(x) \right)^{1/(d+2)}$</td>
<td>$? \quad \text{rate: } n^{1/(d+2)} \quad ??$ $\text{const: } ??$</td>
</tr>
<tr>
<td>C (global)</td>
<td>Biau and Devroye (2003) $\text{rate: } n^{1/(d+2)}$</td>
<td>Biau and Devroye (2003) analogues of Birgé’s histogram estimators</td>
</tr>
</tbody>
</table>
B: Review of available theory:

MLEs for multivariate classes

Scale mixtures of uniform on $\mathbb{R}^+d$

<table>
<thead>
<tr>
<th>Problem</th>
<th>Lower Bound</th>
<th>Upper Bound / MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C (global)</td>
<td>? (hints from entropy bounds of Blei, Gao, and Li)</td>
<td>?? ?? ??</td>
</tr>
</tbody>
</table>
B: Review of available theory: MLEs for multivariate classes

Log concave densities on $\mathbb{R}^d$

<table>
<thead>
<tr>
<th>Problem</th>
<th>Lower Bound</th>
<th>Upper Bound / MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B (local)</td>
<td>Seregin (2010) rate: $n^2/(d+4)$ [ \left{ f^{d+2}(x)\text{curv}_x(\varphi) \right}^{1/(d+4)} ] [ \text{curv}_x(\varphi) = \det \nabla^2 \varphi(x) ]</td>
<td>?? ?? ??</td>
</tr>
<tr>
<td>C (global)</td>
<td>?? conjectures: Seregin and W (2010)</td>
<td>?? ?? ??</td>
</tr>
</tbody>
</table>
## B: Review of available theory:

**MLEs for multivariate classes**

$s$–convex and $h$–convex densities on $\mathbb{R}^d$

<table>
<thead>
<tr>
<th>Problem</th>
<th>Lower Bound</th>
<th>Upper Bound / MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B (local)</td>
<td>$Seregin$ (2010) rate: $\frac{n^2}{(d+4)} \left{ \frac{f(x) \text{curv}_x(\varphi)}{h'(\varphi(x))^4} \right}^{1/(d+4)}$</td>
<td>MLE &amp; LSE rate inefficient (d &gt; 4) ?? ??</td>
</tr>
<tr>
<td>C (global)</td>
<td>?? ?? ??</td>
<td>?? LSE rate inefficient, (d &gt; 4)?? or (d \geq 4)??</td>
</tr>
</tbody>
</table>
Let $k \geq 1$ be an integer. $f : (0, \infty) \rightarrow \mathbb{R}^+$ is hyperbolically monotone of order $k$ ($HM_k$) if, for each fixed $u > 0$, the function

$$H(w) \equiv f(uv)f(u/v), \quad w \equiv \frac{1}{2} \left( v + \frac{1}{v} \right) \geq 1,$$

is such that

$$(-1)^j H^{(j)}(w) \geq 0, \quad \text{for } j = 0, \ldots, k - 1,$$

$$(-1)^k H^{(k-1)}(w) \text{ is right continuous and decreasing.}$$

If $f$ is hyperbolically monotone for all $k$, $f$ is said to be hyperbolically completely monotone ($HCM$ or $HM_\infty$).

Note that $f(uv)f(u/v)$ is always a function of $w$ by symmetry, and for $v \geq 1 \quad v = w + \sqrt{w^2 - 1}$. 
C: Some alternative classes on $\mathbb{R}$:

Bondesson’s hyperbolically monotone classes

Example 1. Half-normal distributions

$$f(x) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) 1_{(0,\infty)}(x)$$

has

$$H(w) = f(uv)f(u/v) = \frac{2}{\pi\sigma^2} \exp\left(-\frac{u^2}{2\sigma^2} \left(v + \frac{1}{v}\right)^2 + \frac{u^2}{\sigma^2}\right)$$

$$= \frac{2}{\pi\sigma^2} \exp\left(-\frac{u^2}{2\sigma^2} w^2 + \frac{u^2}{\sigma^2}\right).$$

For fixed $u > 0$ $w \mapsto H(w)$ is decreasing, but $-H'$ is not. Thus $f \in HM_1$ while $f \not\in HM_2$.

Example 2. Uniform $(a,b)$ If $f(x) = (b - a)^{-1} 1_{(a,b)}(x)$ with $0 \leq a < b$, then for $v \geq 1$

$$H(w) = f(uv)f(u/v) = (b - a)^{-2} 1\{a < u/v \leq uv < b\}$$

so $H(w) = 1\{1 \leq v < V_u \equiv \min\{u/a, b/u\}\}$. Thus $H$ is decreasing and $f \in HM_1$. 
Exercise 1. Show that \( f(x) = Cx^{\beta-1}(1 + cx)^{-\gamma}1_{(0,\infty)}(x), \) with \( \beta, \gamma, c \geq 0 \) and \( C \) a normalizing constant, satisfies \( f \in HM_\infty. \)

Exercise 2. (log-normal) \( f(x) = C \exp(-(\log x - \mu)^2/2\sigma^2) \) satisfies \( f \in HM_\infty. \)

Exercise 3. \( f(x) = (a - x)^{\gamma-1}1_{(0,\infty)}(x) \) satisfies \( f \in HM_{\lfloor \gamma \rfloor}. \)

Theorem 1. (Bondesson, 1997) If \( X \) and \( Y \) are independent random variables such that \( X \sim f \in HM_k \) and \( Y \sim g \in HM_k \), then \( XY \sim HM_k \) and \( X/Y \sim HM_k. \)

Theorem 2. (Bondesson, 1992) \( X \sim f \in HM_1 \) if and only if \( \log X \sim e^x f(e^x) \) is log-concave.
C: Some alternative classes on $\mathbb{R}$:

Bondesson’s hyperbolically monotone classes

Putting these two results together:

- Transform the hyperbolically monotone classes from $\mathbb{R}^+$ to $\mathbb{R}$:
  \[ \mathcal{HM}_k \circ \exp \equiv \{ g(x) = e^x f(e^x) : f \in \mathcal{HM}_k \} \]
  \[ \equiv \text{log-hyperbolically } k-\text{monotone} \]

- $\mathcal{HM}_k \circ \exp$ is closed under convolution.

- $\mathcal{HM}_k \circ \exp$ have the same degree of smoothness as the $k+1$-monotone densities

- $\mathcal{HM}_\infty \circ \exp$ contains the Gaussian distributions (by Exercise 2).

**Conclusion:** The classes $\mathcal{HM}_k \circ \exp \setminus \mathcal{HM}_\infty \circ \exp$ provide a nice analogue of the $k$-monotone classes on $\mathbb{R}^+$ for $\mathbb{R}$ with nice closure properties.
D: Open problems and questions: $\mathbb{R}^1$

- Are there “natural” switching relations for the $k$–monotone MLE’s and / or LSE’s?
- More connections to convexity theory?
- Pointwise rates of convergence for the $k$–monotone MLE’s?
- Rates of convergence under degenerate mixing, $G = \delta_1$?
- Rates of convergence for the MLE’s of $G$ (inverse problems)?
- Global rates of convergence in $L_1$ and Hellinger metrics, log-concave classes?
- Theory for natural discrete shape-constrained classes? (Monotone, convex-decreasing, completely monotone, ....)
- MLE’s for Bondesson’s $HM_k$ classes?
E: Open problems and questions: $\mathbb{R}^d$

- **local rates** and **global rates** for shape constrained estimators in $\mathbb{R}^d$?

- Local (pointwise) limiting distribution theory for MLE’s and other natural divergence-based estimators?

- When are the MLE’s rate (in-)efficient?
  
  **Conjecture 1:** Block decreasing: inefficient for $d > 2$.
  
  **Conjecture 2:** Log-concave and $s$–concave: inefficient for $d > 4$.

- How to penalize or sieve or ... to obtain rate efficient estimators in these classes for higher dimensions?

- Do there exist natural shape-constraints with smoothness $> 2$ for which MLE’s are rate-efficient and which have natural preservations properties under convolution, marginalization, and so forth?
E: Open problems and questions: $\mathbb{R}^d$

- Faster and more efficient algorithms?

- Faster and more efficient algorithms?!
Je vous remerci!