REFERENCES

ASIMOV, D. (1985). The grand tour. *SIAM J. Sci. Statist. Comput.* **6** 128–143.

BREIMAN, L. and FRIEDMAN, J. H. (1985). Estimating optimal transformations for multiple regression and correlation (with discussion). To appear in *J. Amer. Statist. Assoc.* Also published as Technical Report Orion-010 (1982), Dept. of Statist., Stanford University.

DIACONIS, P. and FREEDMAN, D. (1984). Asymptotics of graphical projection pursuit. *Ann. Statist.* **12** 793–815.

FISHERKELLER, M. A., FRIEDMAN, J. H. and TUKEY, J. W. (1974). Prim-9: an interactive multidimensional data display and analysis system. SLAC-Pub-1408, Stanford Linear Accelerator Center, Stanford, California.

FRIEDMAN, J. H. and STUETZLE, W. (1981a). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823.

FRIEDMAN, J. H. and STUETZLE, W. (1981b). Projection pursuit methods for data analysis. In *Modern Data Analysis.* (A. F. Siegel and R. Launer, eds.) Academic, New York.

FRIEDMAN, J. H., STUETZLE, W. and SCHROEDER, A. (1984). Projection pursuit density estimation. *J. Amer. Statist. Assoc.* **79** 599–608. Also published as Stanford Technical Report Orion-002 (1981), Dept. of Statist., Stanford University.

HENRY, D. H. (1983). Multiplicative models in projection pursuit. SLAC-Report 261, Stanford Linear Accelerator Center, Stanford, California.

HUBER, P. J. (1975). Applications vs. abstraction: the selling out of mathematical statistics? *Suppl. Adv. Appl. Probab.* **7** 84–89.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

PING CHENG AND C. F. J. WU

Academia Sinica, Beijing and University of Wisconsin, Madison

Professor Huber's stimulating paper has greatly advanced our knowledge of the projection pursuit methodology. Our discussion will be confined to the convergence of the projection pursuit density approximation method (PPDA). In Proposition 14.3 he proved the uniform and L_1-convergence of the PPDA by assuming that the density f can be deconvoluted with a Gaussian component. This is a very strong smoothness condition on f. Our original attempt was to prove his conjecture that the convergence still holds under more general smoothness condition on f. Failing this, we have instead found a smoothed version of the PPDA that converges uniformly and in L_1 to f with no smoothness condition required on f. Our modification is described as follows.

Let $\{g^{(k)}\}$ be the sequence of approximating densities defined in Proposition 14.3. Define the *smoothed* approximating density $\bar{g}^{(k)}$ by convoluting $g^{(k)}$ with a normal density

$$\bar{g}^{(k)} = g^{(k)} * N(0, \sigma_k^2 I_d), \tag{1}$$

where σ_k satisfies

$$\tau_k \sigma_k^{-2d} \to 0 \quad \text{and} \quad \sigma_k \to 0 \quad \text{as} \quad k \to \infty \tag{2}$$

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

and

$$\tau_k = \sup_{\|a\|=1} E_f\left(\log \frac{f_a(x)}{g_a^{(k)}(x)}\right) = E_f\left(\log \frac{f_{a_k}(x)}{g_{a_k}^{(k)}(x)}\right).$$

Since $g^{(0)}$ is the normal density with the same mean and covariance as f, $E_f(f, g^{(0)}) < \infty$, which implies $\tau_k \to 0$ from the discussion at the beginning of Section 14.

THEOREM. *In general,*

$$\int |\bar{g}^{(k)}(x) - f(x)|\, dx \to 0 \quad \text{as} \quad k \to \infty.$$

If $f(x)$ is uniformly continuous,

$$\sup_x |\bar{g}^{(k)}(x) - f(x)| \to 0 \quad \text{as} \quad k \to \infty.$$

PROOF. We treat the convergence of $\bar{g}^{(k)} - \bar{f}^{(k)}$ and $\bar{f}^{(k)} - f$ separately, where

$$\bar{f}^{(k)} = f * N(0, \sigma_k^2 I_d).$$

The characteristic functions (ch.f.) of $\bar{g}^{(k)}$ and $\bar{f}^{(k)}$, $\psi^{(k)}(t)\exp(-\sigma_k^2 \|t\|^2/2)$ and $\psi(t)\exp(-\sigma_k^2 \|t\|^2/2)$, where $\psi^{(k)}$ and ψ are the ch.f. of $g^{(k)}$ and f, are absolutely integrable. From the inversion theorem,

$$\begin{aligned}
\sup_x |\bar{g}^{(k)}(x) - \bar{f}^{(k)}(x)| &\le (2\pi)^{-d} \int |\psi^{(k)}(t) - \psi(t)| \exp(-\sigma_k^2 \|t\|^2/2)\, dt \\
&\le (2\pi)^{-d}(2\tau_k)^{1/2} \int \exp(-\sigma_k^2 \|t\|^2/2)\, dt \qquad (3)\\
&= (2\tau_k)^{1/2}(\sigma_k)^{-d} \to 0
\end{aligned}$$

from (2). The last inequality follows from the proof of Huber's Proposition 14.2 and Lemma 12.3.

To prove the L_1-convergence of $\bar{g}^{(k)} - \bar{f}^{(k)}$ to zero, consider

$$\begin{aligned}
&\int |\bar{g}^{(k)}(x) - \bar{f}^{(k)}(x)|\, dx \\
&\quad\le \int_{\|x\|\le R} |\bar{g}^{(k)}(x) - \bar{f}^{(k)}(x)|\, dx + \int_{\|x\|>R} \bar{g}^{(k)}(x)\, dx + \int_{\|x\|>R} \bar{f}^{(k)}(x)\, dx \\
&\quad= I_1(k) + I_2(k) + I_3(k).
\end{aligned}$$

From (3), $I_1(k) \to 0$ as $k \to \infty$.

$$
\begin{aligned}
I_2(k) &= (2\pi)^{-d/2}\sigma_k^{-d}\int_{\|x\|>R}\int_t g^{(k)}(x-t)\exp\left(-\frac{\|t\|^2}{2\sigma_k^2}\right)dt\,dx \\
&= (2\pi)^{-d/2}\int_t\int_{\|y+\sigma_k t\|>R} g^{(k)}(y)\,dy e^{-\|t\|^2/2}\,dt \\
&\le (2\pi)^{-d/2}\left\{\int_{\|t\|\le T_k} e^{-\|t\|^2/2}\,dt\int_{\|y+\sigma_k t\|>R} g^{(k)}(y)\,dy + \int_{\|t\|>T_k} e^{-\|t\|^2/2}\,dt\right\} \qquad (4)\\
&\le \int_{\|y\|>R-M} g^{(k)}(y)\,dy + (2\pi)^{-d/2}\int_{\|t\|>T_k} e^{-\|t\|^2/2}\,dt,
\end{aligned}
$$

if T_k is chosen to be $M\sigma_k^{-1}$ and $M < R$. The first term of (4) converges to $\int_{\|y\|>R-M} f(y)\,dy$ as $k \to \infty$ from the weak convergence of $g^{(k)}$ to f (Proposition 14.2). The second term converges to zero as $k \to \infty$ since $\sigma_k^{-1} \to \infty$. Therefore

$$
\limsup_{k\to\infty} I_2(k) \le \int_{\|y\|>R-M} f(y)\,dy. \tag{5}
$$

Similarly it can be shown that lim sup $I_3(k)$ is bounded above by the same expression. The upper bound in (5) can be made arbitrarily small by choosing R large. This establishes

$$
\int |\bar{g}^{(k)}(x) - \bar{f}^{(k)}(x)|\,dx \to 0. \tag{6}
$$

Next we consider

$$
\int |\bar{f}^{(k)}(x) - f(x)|\,dx \le (2\pi)^{-d/2}\int A_k(y)e^{-\|y\|^2/2}\,dy,
$$

where

$$
A_k(y) = \int |f(x - \sigma_k y) - f(x)|\,dx \to 0 \quad \text{as} \quad k \to \infty
$$

follows from a standard covergence result in measure theory (Problem 17b, Royden, 1968) since $\sigma_k \to 0$. Since $A_k(y) \le 2$,

$$
\int |\bar{f}^{(k)}(x) - f(x)|\,dx \to 0. \tag{7}
$$

The L_1-convergence of $\bar{g}^{(k)}$ to f follows from (6) and (7). To prove the uniform convergence of $\bar{f}^{(k)}$ to f, we need to assume that f is uniformly continuous. We have

$$
(8)\quad
\begin{aligned}
|f(x) - \bar{f}^{(k)}(x)| &\le (2\pi)^{-d/2} \int |f(x) - f(x - \sigma_k y)| e^{-\|y\|^2/2}\, dy \\
&\le (2\pi)^{-d/2} \biggl\{ 2 \sup_x f(x) \int_{\|y\|>R} e^{-\|y\|^2/2}\, dy \\
&\qquad + \int_{\|y\|\le R} |f(x) - f(x - \sigma_k y)| e^{-\|y\|^2/2}\, dy \biggr\},
\end{aligned}
$$

whose first term can be made arbitrarily small by choosing a large R since $\sup_x f(x) < \infty$ follows from the uniform continuity of f, and whose second term, for fixed R, can be made arbitrarily small (uniformly in x) by choosing a small σ_k (k large) again from the uniform continuity of f. This and (3) imply the uniform convergence of $\bar{g}^{(k)}$ to f. □

The uniform continuity condition on f is much weaker than the condition in Proposition 14.3 that f can be deconvoluted with a normal density.

Our last remark concerns the choice of σ_k in the smoother (1), which depends on the knowledge of τ_k. An *optimal* choice of σ_k can be obtained by equating the convergence rates of $\bar{g}^{(k)} - \bar{f}^{(k)}$ and $\bar{f}^{(k)} - f$. Let us further assume that f satisfies the Lipschitz condition of order λ

$$|f(x_1) - f(x_2)| \le C|x_1 - x_2|^\lambda,$$

where C is independent of x_1, x_2. Then $|f(x) - \bar{f}^{(k)}(x)|$ in (8) is bounded above by $C'\sigma_k^\lambda$. This and the rate $\tau_k^{1/2}\sigma_k^{-d}$ in (3) are of the same order if

$$\sigma_k = c\tau_k^{1/2(d+\lambda)}.$$

REFERENCE

ROYDEN, H. L. (1968). *Real Analysis*. 2nd ed. Macmillan, New York.

INSTITUTE OF SYSTEMS SCIENCE
ACADEMIA SINICA
BEIJING 100080, PEOPLE'S REPUBLIC OF CHINA

1210 W. DAYTON ST.
DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706

D. R. COX

Imperial College, London

Dr. Huber's scholarly paper invests the impressive techniques of projection pursuit with a halo of mathematical formalism. Key questions clearly concern the choice of properties that it is *scientifically* fruitful to pursue. My judgment, based on totally inadequate experience, is that, except in fairly extreme cases, peculiarities of univariate distributional form are often of fairly fleeting interest