Ann. Inst. Statist. Math.
Vol. 54, No. 4, 758–767 (2002)

BLIND DECONVOLUTION WHEN NOISE IS SYMMETRIC: EXISTENCE AND EXAMPLES OF SOLUTIONS*

STEVEN P. ELLIS

Unit. 42, New York State Psychiatric Institute and Columbia University, 1051 Riverside Dr., New York, NY 10032, U.S.A., e-mail: ellis@neuron.cpmc.columbia.edu

(Received May 12, 2000; revised September 25, 2001)

Abstract. The problem considered is that of identifying two finite dimensional probability distributions G and H from their convolution, $F = G * H$, when all that is known about them is that H is symmetric. This problem arises in looking for hidden structure in multivariate data, for example. It is shown that one can always find a solution in which G has no nondegenerate symmetric convolution factor. However the solution is not unique in general. Examples of such "completely asymmetric" distributions are given. Existence and examples rather than estimation are the focus of the paper.

Key words and phrases: Image analysis.

1. Introduction

In this paper we consider the problem of identifying two finite dimensional probability distributions G and H from their convolution, $F = G * H$, when all that is known about them is that, like any uni- or multi-variate normal distribution with mean 0, H is symmetric about the origin. We define what a "solution" to this problem is, show it exists, and consider uniqueness of the solution. We also give some examples. We do not consider estimation of the solution from data. Thus our interest is in existence and identifiability in deconvolution (see also Teicher (1961)). We will assume for mathematical convenience that F has finite variance.

The problem of deconvolution comes up in practical settings. In image restoration, G is the "original image" and convolution with H can be interpreted as "blurring" (Jain (1989)). The data, "the blurred image", is F itself, a density estimate of F, or a discretization of F or its density estimate. Thus in image processing nomenclature we are interested in blind deconvolution when the blurring filter is phaseless (i.e., has even transfer function) and is spatially invariant. An example of a blurred image is a positron emission tomography (PET) image (Daghighian *et al.* (1990)).

Deconvolution has been studied by applied mathematicians (see, for example, Baumeister (1987), Chapter 10). In fact, the deconvolution problem is a form of "integral equation of the first kind" (Wing (1991)). Conversely, the "backward heat problem" is really a deconvolution problem (Friedman and Littman (1994), Chapter 3). These authors treat deconvolution assuming one of the convolution factors is known.

Two statistical applications of deconvolution are errors in variables and nonpara-

*This research is supported in part by United States PHS grants MH46745, MH60995, and MH62185.

metric empirical Bayes (Berger (1985) and Maritz and Lwin (1989)). A few of the many theoretical statistical papers on deconvolution are Carroll and Hall (1988), Edelman (1988), and Fan (1993). Here again, one of the convolution factors is assumed known.

Another statistical application is looking for hidden, preferably low dimensional, structure in multivariate data. Statisticians have examined a number of approaches to this problem. The technique of principal curves (Hastie and Stuetzle (1989)) is used to find a hidden curve or surface in a multivariate point cloud. For Dasgupta and Raftery (1998) and Byers and Raftery (1998) the mechanism hiding the structure is the superposition of homogeneous Poisson "clutter".

In the papers cited in the preceding paragraph, the mechanism "hiding" the structural component, G, in the population, F, from which the data are drawn is not necessarily convolution with some noise distribution. On the other hand, Chiaromonte (1998), Section 3, considers the problem of looking for hidden structure in multivariate data when the mechanism for hiding the structure is convolution with some member, H, of a class of "noise" distributions. The problem treated in this paper can be considered as Chiaromonte's (1998) problem with the class of noise distributions taken to be the class of all symmetric distributions with finite variance.

In this paper, we only assume that H is symmetric, otherwise we are "blind" to it. The problem of blind deconvolution has been considered in the signal processing literature, but there it has a temporal dimension not considered here (see, e.g., Liu and Chen (1995)).

Given $F = G * H$ the distributions G and H are not identifiable. Suppose H_1 and H_2 are each symmetric and $H = H_1 * H_2$, so H is also symmetric. Note that unit mass at 0, δ_0, is also a symmetric distribution. Thus $F = (G*H)*\delta_0 = (G*H_1)*H_2 = G*H$ all represent F as a convolution of a distribution with a symmetric distribution. Since the symmetric factor is completely unknown, we do not know which of these expressions to regard as the solution to the deconvolution problem. There is a *prima facie* lack of uniqueness to the solution of our problem.

In this paper we try to resolve this issue by taking as our problem that of removing as much symmetry from F as possible. Call a distribution with no nondegenerate symmetric convolution factor "completely asymmetric". ("Degenerate" means having all mass at just one point.) From the standpoint of practical image analysis, say, it is reasonable to assume that original images typically do not have nondegenerate symmetric convolution factors, i.e., they are completely asymmetric. Blurring, on the other hand, is known to occur sometimes in image acquisition. So if the observed image does have a nondegenerate symmetric convolution factor, that factor is likely to be an artifact. The problem considered, then, is that of extracting from an arbitrary distribution with finite variance a completely asymmetric convolution factor. In the case described in the last paragraph, unless $H_1 = \delta_0$, G might be a solution but $G * H$ and $G * H_1$ cannot be solutions.

We show that the problem of removing all symmetric blurring from a probability distribution has a solution. Unfortunately, we shall see that the solution may still not be unique. However might it be that in many situations met in practice the solution is unique? In order to answer this question, sufficient conditions on F for the uniqueness of its completely asymmetric factor should be found. We do not attempt to do this in this paper.

Actually, in practice it might not be necessary to formulate the problem in such generality. One might be willing to make some assumptions concerning the symmetric factor (Chiaromonte (1998), Section 3). This leads to an expanded class of asymmetric distributions (see Corollary 1) and, *prima facie*, a larger class of distributions for which the deblurring problem is identifiable.

We give precise definitions and statements of results in Section 2. We briefly summarize our findings in Section 3. An appendix contains the proofs of the results.

2. Results

Let F, G, and H be (Borel probability) distributions on $\mathbb{R}^p$ ($\mathbb{R}$ = real line; p = positive integer) having finite variances. Say that F has "finite variance" if $s(F) \equiv \int |x|^2 F(dx) < \infty$. A distribution, H, is "symmetric" if $H(-B) = H(B)$ for every Borel set $B \subset \mathbb{R}^p$. Here, $-B \equiv \{-x \in \mathbb{R}^p : x \in B\}$. Thus in two or more dimensions a distribution may be invariant under reflection across a linear subspace but not be symmetric by this definition. Call a distribution F "completely asymmetric" if whenever H is symmetric and $F = G * H$ then $H = \delta_0$, where δ_x = unit mass at x. Recall that the convolution, $G * H$ is defined by $(G * H)(B) = \int H(B - x)G(dx)$, where $B \subset \mathbb{R}^p$ is Borel and, for $x \in \mathbb{R}^p$, $B - x = \{y - x \in \mathbb{R}^p : y \in B\}$.

These ideas can be re-expressed in terms of random vectors. Let X, Y, and Z be random vectors with distributions F, G, and H, resp. (respectively). Then F has finite variance if $E(|X|^2) < \infty$. H is symmetric if and only if Z and $-Z$ have the same distribution. It follows that the convolution of two symmetric distributions is again symmetric. F is completely asymmetric if whenever Y and Z are independent, H is symmetric and $X = Y + Z$, then $Z = 0$ almost surely (a.s.).

The following is the basic existence theorem for completely asymmetric factors. We give proofs of the results in this paper in the Appendix. A referee suggests that a similar result with a similar proof might apply to "clutter" removal as described in Byers and Raftery (1998) and Dasgupta and Raftery (1998).

PROPOSITION 1. *If F has finite variance then there exist distributions G and H such that G is completely asymmetric, H is symmetric, and*

$$F = G * H. \tag{2.1}$$

Since the support of the convolution of two distributions is the closure of the sum of their supports, any distribution on the line of the form $\lambda\delta_a + (1-\lambda)\delta_b$, where $0 < \lambda < 1/2$ and $a \neq b$, is completely asymmetric. Using this fact one sees that in general the decomposition (2.1) is not unique. For consider the following simple example on the line. Let $G = (3/5)\delta_{-1} + (2/5)\delta_1$, $H = (1/2)\delta_{-1} + (1/2)\delta_1$, and $\mu = (1/5)\delta_{-4} - (1/10)\delta_{-2} + (4/5)\delta_0 - (1/10)\delta_2 + (1/5)\delta_4$. Thus G is completely asymmetric, H and μ are symmetric, but μ is a signed measure. Nonetheless, it turns out that $G * \mu$ and $\mu * H$ are both probability distributions. Let $F = (G * \mu) * H = G * (\mu * H)$. Now, $F = G * (\mu * H)$ is a representation of the type (2.1). By Proposition 1, $G * \mu$ also has a decomposition of the form (2.1). That is, there is a completely asymmetric distribution G' and a symmetric distribution H' s.t. (such that) $G * \mu = G' * H'$. However μ is not a probability measure so $H' \neq \mu$. Therefore $G \neq G'$ and $F = G' * (H' * H)$ is a distinct representation of F in the form (2.1).

The following result describes a class of absolutely continuous completely asymmetric distributions on the real line. It suggests that very sharp edges in an original image can be discerned through any kind of symmetric blurring (see Example 1 below).

PROPOSITION 2. *Let F be a distribution with finite variance and support in $[a, \infty)$, where a is finite. Suppose F has a density, f, satisfying*

Condition 1. *f is locally bounded on (i.e., bounded on compact subsets of) (a, ∞) and*

Condition 2. $\lim_{\epsilon \downarrow 0} f(a + \epsilon) = \infty$.

Then F is completely asymmetric.

The reason a distribution F satisfying the hypotheses of Proposition 2 is completely asymmetric is, roughly speaking, as follows (see the Appendix for a rigorous proof). Suppose $F = G * H$, where H is symmetric and nondegenerate. Since the support, supp F, of F is bounded on the left, for some finite $b > 0$, $-$ inf supp $H =$ sup supp $H =$ (inf supp G) $- a = b$. The left "edge" of H and the left edge of G come together in convolution to produce the infinite spike in f at a. Since H is symmetric, the right edge of H is the same as its left edge. But this means that the right edge of H and the left edge of G must combine in convolution to produce an infinite spike in f at $a + 2b$, contradicting Condition 1.

If F is a distribution on the line, define its "reflection" to be the distribution that assigns to a Borel set $B \subset \mathbb{R}$ the probability $F(-B)$. Clearly, the reflection of a completely asymmetric distribution is also completely asymmetric. Thus Proposition 2 also provides examples of absolutely continuous completely asymmetric distributions on the line with long left hand tails.

In the proof of Proposition 2, we make no use of any moment properties of the distributions involved. Thus Proposition 2 holds for a stronger definition of complete asymmetry involving no assumptions concerning moments.

One can also weaken the definition of complete asymmetry. So far we have put no conditions on the symmetric factor H, except that it have a finite second moment. If additional conditions are put on H then the class of distributions having no symmetric convolution factor satisfying the additional conditions obviously broadens. An examination of the proof of Proposition 2 yields the following generalization.

COROLLARY 1. *Let $c \in (0, \infty]$ and let F be a distribution with finite variance and support in $[a, \infty)$, where a is finite. Suppose F has a density, f, satisfying*

Condition 1. *f is locally bounded on $(a, a + 3c]$ and*

Condition 2. $\lim_{\epsilon \downarrow 0} f(a + \epsilon) = \infty$.

Then F has no nondegenerate symmetric convolution factor with support in $(-c, c)$.

Let X be a two-dimensional random vector with distribution F. Let v be a nonrandom unit vector in $\mathbb{R}^2$. Define the "projection", F^v, of F onto the line spanned by v to be the distribution of the random variable $v \cdot X$, where "$\cdot$" indicates the usual Euclidean inner product. (Define the projection f^v of a density f similarly.) All projections of F are determined by $\{F^v : v = (\cos\theta, \sin\theta), \theta \in (0, \pi]\}$ (essentially the "Radon transform" of F (Deans (1993))). It is easy to see that if $F = G * H$ then $F^v = G^v * H^v$. If H is symmetric, so are all its projections. If H is also nondegenerate not all its projections will be degenerate. Therefore *if all projections of F are completely asymmetric then so is F*. Exploiting this fact and Proposition 2, we generate an example of an absolutely continuous completely asymmetric distribution on $\mathbb{R}^2$.

Example 1. Let $1/2 < \alpha < 1$. Consider the following density. For convenience it is written in polar coordinates but still think of it as a density on x-y space.

$$f(x,y) = \begin{cases} c\theta(1-r)^{-\alpha}, & \text{if } \; 0 < r < 1 \text{ and } 0 < \theta < \pi; \\ 0, & \text{otherwise}, \end{cases} \tag{2.2}$$

where $(x, y) = (r\cos\theta, r\sin\theta)$. Here, c is a positive constant chosen to make f integrate (in x, y) to 1. In the Appendix we show that such a c exists and that all projections of f (or their reflections) satisfy Conditions 1 and 2 of Proposition 2. Hence all projections of the distribution, F, having density f are completely asymmetric. Therefore F is also completely asymmetric.

Note that f blows up on the semicircular arc $\{(\cos\theta, \sin\theta),\ \theta \in (0, \pi)\}$. It appears that the semicircular shape is not crucial. Probably it suffices only that the arc along which the density blows up have appropriate Hölder continuity (Gilbarg and Trudinger (1998), pp. 52 and 94). Thus it appears that certain sharply defined "corners" in an image can be recovered after any amount of symmetric blurring. With some care, more complex examples based on Corollary 1 instead of Proposition 2 might be constructed.

3. Conclusions

The problem of removing all symmetric blurring has a solution. Unfortunately, that solution may not be unique. However I conjecture that for completely asymmetric distributions G of the sort met with in applications, G often can be uniquely determined from the distribution of the data. As evidence for this we found that very distinct "corners" are features that can be discerned through symmetric blurring.

Appendix: Proofs

PROOF OF PROPOSITION 1. The method used to prove Khintchine's theorem (Linnik (1964), p. 88, or in Linnik and Ostrovskii (1977), p. 79) might be modified to prove Proposition 1. We present an alternative proof that is an application of Zorn's Lemma (Halmos (1974), Section 16).

It is easy to see that

(A.1) If G is completely asymmetric then for every $x \in \mathbb{R}^p$, the distribution $G * \delta_x$ is also completely asymmetric.

Therefore without loss of generality $\int xF(dx) = 0$. Let $\mathcal{S}$ denote the set of all distributions G on $\mathbb{R}^p$ s.t.

$$s(G) \left(\equiv \int |x|^2 G(dx) \right) < \infty.$$

Let $\mathcal{M} = \mathcal{M}_F$ denote the set of all distributions, $G \in \mathcal{S}$, s.t. there exists a symmetric distribution, $H \in \mathcal{S}$ s.t. $F = G * H$. $\mathcal{M}$ is not empty since $F = F * \delta_0$. If $G \in \mathcal{M}$ then $\int xG(dx) = 0$.

If $G_1, G_2 \in \mathcal{M}$, write $G_1 \leq_{\mathcal{M}} G_2$ if there exists a symmetric distribution, $H \in \mathcal{S}$, s.t. $G_1 * H = G_2$. One easily sees that $\leq_{\mathcal{M}}$ is a partial ordering. (Chiaromonte (1998), makes use of a similar ordering on probability laws.)

The proposition amounts to saying $\mathcal{M}$ contains a minimal element with respect to $\leq_{\mathcal{M}}$. We will prove this using Zorn's Lemma (Halmos (1974), Section 16). Let

$\mathcal{G} = \{G_\alpha \in \mathcal{M} : \alpha \in A\}$ be totally ordered with respect to $\leq_{\mathcal{M}}$. Then it suffices to show that $\mathcal{G}$ has a lower bound in $\mathcal{M}$. Let $v = \inf\{s(G_\alpha) : \alpha \in A\}$. Choose a countable sequence $\{G_{\alpha(n)} \in \mathcal{G}\}$ s.t. $s(G_{\alpha(n)}) \downarrow v$ and write $G^{-n} = G_{\alpha(n)}$, $n = 1, 2, \ldots$. Since $\mathcal{G}$ is totally ordered, if $-m < -n$ then $G^{-m} \leq_{\mathcal{M}} G^{-n}$.

Claim. If $\{G^{-n}\}$ has a lower bound in $\mathcal{M}$ then so does $\mathcal{G}$. For suppose not. Let $G \in \mathcal{M}$ be a lower bound for $\{G^{-n}\}$ which is not a lower bound for $\mathcal{G}$. Then there exists $\alpha \in A$ s.t. $G \not\leq_{\mathcal{M}} G_\alpha$. Suppose $s(G_\alpha) > v$. Since $\mathcal{G}$ is totally ordered and $s(G_{\alpha(n)}) \downarrow v$, for some n we have, $G \leq_{\mathcal{M}} G^{-n} \leq_{\mathcal{M}} G_\alpha$, contradicting the way G_α was chosen. Thus $s(G_\alpha) = v$. But this implies G_α itself is a lower bound for $\mathcal{G}$. The claim is established.

Thus we need only find a lower bound for $\{G^{-n}\}$ in $\mathcal{M}$. We will do this by constructing a reversed martingale whose marginal distributions are G^{-n} , $n \geq 1$, and then applying a martingale limit theorem. If $k = 1, 2, 3, \ldots$, then there exists a symmetric distribution $K_{-k} \in \mathcal{S}$ s.t. $G^{-k+1} = G^{-k} * K_{-k}$ ($G^0 \equiv F$). For each n, let $X_{-n,n}$ be a random vector with distribution G^n and let $Y_{-1,n}, \ldots, Y_{-n,n}$ be independent random vectors independent of $X_{-n,n}$ s.t. $Y_{-k,n}$ has distribution K_{-k}. Let

$$X_{-n+k,n} = X_{-n,n} + \sum_{j=1}^{k} Y_{-n+j-1,n}, \qquad k = 1, \ldots, n.$$

Then $X_{-n+k,n}$ has distribution G^{-n+k}. Note that for each k, $Y_{-n+k,n}$ is independent of $X_{-n,n}, \ldots, X_{-n+k,n}$. Let P_n denote the joint distribution of $(X_{-n,n}, \ldots, X_{0,n})$.

By the Kolmogorov Extension Theorem (Ash (1972), p. 191, Theorem 4.4.3) there exists a process $\{\ldots, X_{-3}, X_{-2}, X_{-1}, X_0\}$ s.t. P_n is the joint distribution of $X_{-n}, \ldots, X_0$. Thus X_{-k} has distribution G^{-k}, $k \geq 0$.

Let $Y_{-k} = X_{-k+1} - X_{-k}$, $k \geq 1$. Then for $k \geq 1$,

- Y_{-k} has distribution K_{-k} (in particular $EY_{-k} = 0$),
- $Y_{-1}, Y_{-2}, \ldots$ are independent, and
- Y_{-k} is independent of $X_{-k-j}, j \geq 0$.

It follows that, for each i $(= 1, \ldots, p)$, $\{X_{-n}[i], n \geq 0\}$ is a (reversed) martingale. (Here, $X_{-n}[i]$ is the i-th component of the random vector X_{-n}.) Therefore by Chung (1974), Theorems 9.4.7, p. 338 and 4.5.2, p. 95, $\{X_{-n}\}$ converges a.s. to a random vector $X_{-\infty}$ with finite variance.

As a consequence, for each $n \geq 0$, $\sum_{k=n+1}^{\infty} Y_{-k}$ converges a.s. to $W_{-\infty,n} \equiv X_{-n} - X_{-\infty}$. Now, $W_{-\infty,n}$ has a symmetric distribution, $s(W_{-\infty,n}) < \infty$, and $W_{-\infty,0} + X_{-\infty}$ has distribution $G_0 = F$. Furthermore, $W_{-\infty,n}$ is independent of $X_{-\infty}$. But this means that if $G_{-\infty}$ is the distribution of $X_{-\infty}$ then $G_{-\infty} \in \mathcal{M}$ and $G_{-\infty} \leq_{\mathcal{M}} G^n$ for all n. The proposition follows.

PROOF OF PROPOSITION 2. Write $F = G * H$, where H is symmetric. By (A.1) we may assume $a = 0$. By Condition 2 of Proposition 2, $\inf \operatorname{supp} F = a = 0$ ("supp" = "support of"). Since $\operatorname{supp} F = \overline{(\operatorname{supp} G + \operatorname{supp} H)}$ (the overbar indicates closure), it follows that if $-b = \inf \operatorname{supp} H$, then $-b > -\infty$ and $b = \inf \operatorname{supp} G$. Since H is symmetric, $b = \sup \operatorname{supp} H \geq 0$. We will see that f must blow up at $a + 2b$, contradicting Condition 1, unless $b = 0$, i.e. unless F is completely asymmetric.

Assume F is not completely asymmetric. This amounts to assuming that $b > 0$. We make H and G absolutely continuous by convolving them with a symmetric, unimodal, differentiable density, whose support we will shrink to $\{0\}$. Let α_1 be a symmetric probability density with support $[-1,1]$ and unique local maximum at 0. Suppose α_1 is everywhere differentiable with bounded derivative. Then $\alpha_1(x) > 0$ if $-1 < x < 1$, $\alpha_1'(x) \geq 0$ if $x \leq 0$, and $\alpha_1'(x) \leq 0$ if $x \geq 0$. Let $\beta_1 = \alpha_1 * \alpha_1$. Then β_1 has support $[-2,2]$. For $\epsilon \in (0, b/2)$, let $\alpha_\epsilon(x) = \epsilon^{-1}\alpha_1(x/\epsilon)$, $\beta_\epsilon(x) = \epsilon^{-1}\beta_1(x/\epsilon)$. Then α_ϵ and β_ϵ have supports $[-\epsilon, \epsilon]$, $[-2\epsilon, 2\epsilon]$, resp. and $\beta_\epsilon = \alpha_\epsilon * \alpha_\epsilon$. The convolution of α_ϵ with any distribution, K, has bounded derivative, $\int \alpha_\epsilon'(x-y)K(dy)$. In particular, β_ϵ has a bounded derivative.

Define functions $g_\epsilon = G * \alpha_\epsilon$, $h_\epsilon = H * \alpha_\epsilon$, and $f_\epsilon = f * \beta_\epsilon$. Notice that

$$\operatorname{supp} g_\epsilon \subset [b-\epsilon, \infty), \qquad \operatorname{supp} h_\epsilon \subset [-b-\epsilon, b+\epsilon], \quad \text{and} \quad f_\epsilon = g_\epsilon * h_\epsilon.$$

Now,

$$g_\epsilon'(x) = \int_{[b,\infty]} \alpha_\epsilon'(x-y)G(dy).$$

Since $\alpha_\epsilon' \geq 0$ on $(-\infty, 0]$, it follows that g_ϵ is nondecreasing on $(-\infty, b]$. Similarly, h_ϵ is nondecreasing on $(-\infty, -b]$.

In a moment we will prove,

$$\text{(A.2)} \qquad f_\epsilon(2b) \to \infty \quad \text{as} \quad \epsilon \to 0.$$

On the other hand, for ϵ sufficiently small,

$$f_\epsilon(2b) = \int_{2b-2\epsilon}^{2b+2\epsilon} \beta_\epsilon(2b-y)f(y)dy \leq \sup_{2b-2\epsilon<y<2b+2\epsilon} f(y) \leq \sup_{b<y<3b} f(y) < \infty,$$

by Condition 1 (we are assuming $b > 0$). This contradicts (A.2). The proposition follows.

(A.2) is a consequence of two further facts:

$$\text{(A.3)} \qquad f_\epsilon(-\epsilon) \to \infty \quad \text{as} \quad \epsilon \downarrow 0, \text{ and}$$

$$\text{(A.4)} \qquad f_\epsilon(2b) \geq f_\epsilon(-\epsilon) \quad \text{for} \quad \epsilon \in (0, b/2).$$

PROOF OF (A.3). Since β_ϵ is symmetric with support $[-2\epsilon, 2\epsilon]$,

$$\begin{aligned} f_\epsilon(-\epsilon) &= \int_0^\infty \beta_\epsilon(\epsilon+y)f(y)dy = \int_0^\epsilon \beta_\epsilon(\epsilon+y)f(y)dy \\ &\geq \left[\inf_{0<y<\epsilon} f(y)\right] \int_1^2 \beta_1(z)dz. \end{aligned}$$

Since $\operatorname{supp} \beta_1 = [-2,2]$, (A.3) now follows from Condition 2.

PROOF OF (A.4). Since $\operatorname{supp} g_\epsilon \subset [b-\epsilon, \infty)$ and $\operatorname{supp} h_\epsilon \subset [-b-\epsilon, b+\epsilon]$,

$$\begin{aligned} f_\epsilon(2b) - f_\epsilon(-\epsilon) &= \int_{b-\epsilon}^b [h_\epsilon(2b-y) - h_\epsilon(-\epsilon-y)]g_\epsilon(y)dy + \int_b^\infty h_\epsilon(2b-y)g_\epsilon(y)dy \\ &\geq \left(\int_{b-\epsilon}^{b-\epsilon/2} + \int_{b-\epsilon/2}^b\right) [h_\epsilon(2b-y) - h_\epsilon(-\epsilon-y)]g_\epsilon(y)dy. \end{aligned}$$

Making the change of variables $y = 2b - z - \epsilon$ in the first integral in the last expression in the preceding and using the fact that h_ϵ is symmetric, we get,

$$f_\epsilon(2b) - f_\epsilon(-\epsilon) \geq \int_{b-\epsilon/2}^{b} [h_\epsilon(y - 2b) - h_\epsilon(-y - \epsilon)][g_\epsilon(y) - g_\epsilon(2b - y - \epsilon)]dy. \tag{A.5}$$

Now, if $y \in (b - \epsilon/2, b)$, then $-b > y - 2b > -b - \epsilon/2 > -y - \epsilon$ and $b > y > 2b - y - \epsilon$. Since h_ϵ and g_ϵ are nondecreasing on $(-\infty, -b]$ and $(-\infty, b]$, resp., it follows that the integral on the right hand side of (A.5) is nonnegative, i.e. (A.4) holds. □

PROOFS FOR EXAMPLE 1. To study the behavior of f we will use the following simpler function.

$$\tilde{f}(x,y) = \begin{cases} (1 - \sqrt{x^2 + y^2})^{-\alpha}, & \text{if } \ 0 < r < 1; \\ 0, & \text{otherwise.} \end{cases}$$

So $\tilde{f}$ is symmetric about the origin. Thus all projections of $\tilde{f}$ onto lines are the same. So for convenience project it onto the x axis. The projection is

$$\tilde{f}^1(x) \equiv \begin{cases} \displaystyle\int_{-\infty}^{\infty} \left(k^1(y;x)\right)^{-\alpha} dy, & \text{if } \ -1 < x < 1. \\ 0, & \text{otherwise,} \end{cases} \tag{A.6}$$

where, for $-1 < x < 1$,

$$k^1(y) = k^1(y;x) = \begin{cases} 1 - \sqrt{x^2 + y^2}, & \text{if } \ |y| < \sqrt{1 - x^2}; \\ 0, & \text{otherwise.} \end{cases}$$

k^1 vanishes at $y = \sqrt{1 - x^2}$ and is concave (i.e., its derivative decreases) in the interval $(0, \sqrt{1 - x^2})$. Therefore over the interval $(0, \sqrt{1 - x^2})$

$$g \leq k^1 \leq h, \tag{A.7}$$

where $g(y) = g(y;x)$ and $h(y) = h(y;x)$ are the linear functions satisfying

$$g(\sqrt{1 - x^2}) = 0, g' \equiv -\frac{k^1(0)}{\sqrt{1 - x^2}}, h(\sqrt{1 - x^2}) = 0, \text{ and } h' \equiv (k^1)'(\sqrt{1 - x^2}).$$

Thus

$$g(y) = (1 - |x|)\left(1 - \frac{y}{\sqrt{1 - x^2}}\right) \quad \text{and} \quad h(y) = 1 - x^2 - y\sqrt{1 - x^2}.$$

(A.6) and (A.7) can be used to compute bounds for $\tilde{f}^1(x)$:

$$2\int_0^{\sqrt{1-x^2}} h(y)^{-\alpha} dy \leq \tilde{f}^1(x) \leq 2\int_0^{\sqrt{1-x^2}} g(y)^{-\alpha} dy. \tag{A.8}$$

Simple calculations show

$$2\int_0^{\sqrt{1-x^2}} h(y)^{-\alpha} dy = 2(1 - \alpha)^{-1} \left(1 - x^2\right)^{-\alpha} \sqrt{1 - x^2} \quad \text{and} \tag{A.9}$$

$$2\int_0^{\sqrt{1-x^2}} g(y)^{-\alpha} dy = 2(1 - \alpha)^{-1}(1 - |x|)^{-\alpha}\sqrt{1 - x^2}, \quad \text{for } \ x \in (-1, 1). \tag{A.10}$$

Write $f = c\bar{f}$. I.e., $\bar{f}(r, \theta) = \theta(1-r)^{-\alpha}$ for $0 \leq r < 1$ and $0 < \theta < \pi$ and $\bar{f} = 0$ otherwise. Then from (A.8), (A.10), and the fact that $\alpha < 1$,

$$\begin{aligned} \int_{\mathbb{R}^2} \bar{f} &\leq \pi \int_{\mathbb{R}^2} \tilde{f} = \pi \int_{-\infty}^{\infty} \tilde{f}^1(x)dx \\ &\leq 2\pi \int_{-1}^{1} \int_{0}^{\sqrt{1-x^2}} g(y;x)^{-\alpha} dy dx = 4\pi(1-\alpha)^{-1} \int_0^1 (1-x)^{-\alpha}\sqrt{1-x^2}dx \\ &\leq 4\pi(1-\alpha)^{-1}\sqrt{2} \int_0^1 (1-x)^{(1/2)-\alpha}dx = 4\pi\sqrt{2}(1-\alpha)^{-1}\left(\frac{3}{2}-\alpha\right)^{-1} < \infty \end{aligned}$$

so the promised finite positive constant "c" in (2.2) exists.

Next, we investigate the projections of f. Let $\phi \in (0, \pi]$ and let $v = (\cos\phi, \sin\phi)$. Let $f_\phi = f^v$. We show that f_ϕ is the reflection of a density satisfying the hypotheses of Proposition 2. Obviously, $f_\phi(z) = 0$ if $|z| > 1$. Because $f_\phi \leq \pi c\tilde{f}^1$, from (A.8) and (A.10) we see that f_ϕ is locally bounded on $(-1, 1)$.

Since, by assumption, $\phi \in (0, \pi]$, there exists $\theta_0 > 0$ s.t. if $z \in \mathbb{R}$ is sufficiently close to 1, $f_\phi(z)$ is an integral of f along a line segment only involving values of $\theta \in (\theta_0, \pi)$. Therefore by (A.8) and (A.9)

$$f_\phi(z) \geq c\theta_0\tilde{f}^1(z) \geq 2c\theta_0(1-\alpha)^{-1}\left(1-z^2\right)^{(1/2)-\alpha} \uparrow \infty \quad \text{as} \quad z \uparrow 1,$$

since $1/2 < \alpha < 1$.

Assume $0 < \phi < \pi$. Then, since f is supported by the upper half unit disk, $f_\phi(z) = 0$ for z in a neighborhood of -1. It follows that for $\phi \in (0, \pi)$ the reflection of f_ϕ satisfies the hypotheses of Proposition 2.

It remains to show that f_π is bounded in a neighborhood of -1. Note that

$$\text{(A.11)} \qquad 2\sin\theta \geq \theta \quad \text{for} \quad \theta \in [0, \pi/3].$$

Let $\arctan(x, y)$ denote the angle in $(-\pi, \pi]$ that the ray from the origin passing through $(x, y) \in \mathbb{R}^2$, $(x, y) \neq 0$, makes with the positive x-axis. Then, if $-1 < z < -1/2 = -\cos(\pi/3)$, by (A.11), (A.8), and (A.10),

$$\begin{aligned} f_\pi(z) &= \int_{-\infty}^{\infty} f(-z, y)dy = c\int_0^{\sqrt{1-z^2}} \arctan(-z, y)k^1(y; -z)^{-\alpha}dy \\ &\leq 2c\int_0^{\sqrt{1-z^2}} \sin[\arctan(-z, y)]k^1(y; -z)^{-\alpha}dy \\ &= 2c\int_0^{\sqrt{1-z^2}} \frac{y}{\sqrt{z^2+y^2}}k^1(y; -z)^{-\alpha}dy \\ &\leq \frac{\sqrt{1-z^2}}{|z|}c\int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} k^1(y; -z)^{-\alpha}dy = \frac{\sqrt{1-z^2}}{|z|}c\tilde{f}^1(-z) \\ &\leq 4c(1-\alpha)^{-1}(1-|z|)^{-\alpha}\left(1-z^2\right) \leq 8c(1-\alpha)^{-1}(1-|z|)^{1-\alpha} \downarrow 0 \quad \text{as } z \downarrow -1. \end{aligned}$$

Since $f_\pi(z) = 0$ for $z \leq -1$, this shows that f_π is bounded in a neighborhood of -1. □

Acknowledgements

Communications with Francesca Chiaromonte helped stimulate my interest in this problem. Suggestions by the referees led to a considerable improvement in the paper.

References

Ash, R. B. (1972). *Real Analysis and Probability*, Academic Press, New York.

Baumeister, J. (1987). *Stable Solution of Inverse Problems*, Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed., Springer, New York.

Byers, S. and Raftery, A. E. (1998). Nearest-neighbor clutter removal for estimating features in spatial point processes, *J. Amer. Statist. Assoc.*, **93**, 577–584.

Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density, *J. Amer. Statist. Assoc.*, **83**, 1184–1186.

Chiaromonte, F. (1998). On multivariate structures and exhaustive reductions, *Computing Science and Statistics*, **30**, 204–213, Interface Foundation of North America, Fairfax Station, Virginia.

Chung, K. L. (1974). *A Course in Probability Theory*, 2nd ed., Academic Press, New York.

Daghighian, F., Sumida, R., Phelps, M. (1990). PET imaging: An overview and instrumentation, *Journal of Nuclear Medicine Technology*, **18**, 5–13.

Dasgupta, A. and Raftery, A. E. (1998). Detecting features in spatial point processes with clutter via model-based clustering, *J. Amer. Statist. Assoc.*, **93**, 294–302.

Deans, S. R. (1993). *The Radon Transform and Some of Its Applications*, Krieger, Malabar, Florida.

Edelman, D. (1988). Estimation of the mixing distribution for a normal mean with applications to the compound decision problem, *Ann. Statist.*, **16**, 1609–1622.

Fan, J. Q. (1993). Adaptively local one-dimensional subproblems with application to a deconvolution problem, *Ann. Statist.*, **21**, 600–610.

Friedman, A. and Littman, W. (1994). *Industrial Mathematics: A Course in Solving Real-World Problems*, SIAM, Philadelphia.

Gilbarg, F. and Trudinger, N. S. (1998). *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer, New York.

Halmos, P. R. (1974). *Naive Set Theory*, Springer, New York.

Hastie, T. and Stuetzle, W. (1989). Principal curves, *J. Amer. Statist. Assoc.*, **84**, 502–516.

Jain, A. K. (1989). *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, New Jersey.

Linnik, Yu. V. (1964). *Decomposition of Probability Distributions*, Oliver & Boyd, London.

Linnik, Ju. V. and Ostrovskii, I. V. (1977). *Decomposition of Random Variables and Vectors*, Translations of Mathematical Monographs, Vol. **48**, American Mathematical Society, Providence, Rhode Island.

Liu, J. S. and Chen, R. (1995). Blind deconvolution via sequential imputations, *J. Amer. Statist. Assoc.*, **90**, 567–576.

Maritz, J. S. and Lwin, T. (1989). *Empirical Bayes Methods*, 2nd ed., Chapman and Hall, New York.

Teicher, H. (1961). Identifiability of mixtures, *Annals Math. Statist.*, **32**, 244–248.

Wing, G. M. (1991). *A Primer on Integral Equations of the First Kind: The Problem of Deconvolution and Unfolding*, SIAM, Philadelphia.