ISOTROPIC PROBABILITY MEASURES IN INFINITE DIMENSIONAL SPACES

(Inverse Problems/Prior Information/Stochastic Inversion)

George Backus

Institute of Geophysics & Planetary Physics, A-025
University of California, San Diego
La Jolla, CA 92093

September 4, 1987
Abstract. Every isotropic probability measure on the space $\mathbb{R}^\infty$ of real sequences $x=(x_1, x_2, ...)$ is a convex combination of the measure concentrated at 0 and a member of $I_0(\mathbb{R}^\infty)$, the set of all isotropic probability measures $p_\infty$ on $\mathbb{R}^\infty$ with $p_\infty(\{0\})=0$. Each $p_\infty \in I_0(\mathbb{R}^\infty)$ is completely determined by any one of its finite-dimensional marginal distributions $p_n$. Each $p_n$ has a density function $f_n$ with $dp_n(x_1, ..., x_n)=dx_1 \cdots dx_nf_n(x_1^2 + \cdots + x_n^2)$. Each $f_n$ is completely monotone in $0<\xi<\infty$ (hence analytic in the right complex $\xi$ half-plane), and

$$
\pi^{n/2}\Gamma(n/2)^{-1} \int_0^\infty d\xi \xi^{n/2-1} f_n(\xi) = 1.
$$

Every $f$ which satisfies these two conditions is $f_n$ for a unique $p_\infty \in I_0(\mathbb{R}^\infty)$. Hence the equation

$$
\pi \int_\xi d\zeta f_2(\zeta) = \int_0^\infty d\mu(t)e^{-t\xi}
$$

defines a bijection between $I_0(\mathbb{R}^\infty)$ and the set of all probability measures $\mu$ on $0 \leq t < \infty$. If $p_\infty \in I_0(\mathbb{R}^\infty)$ then $p_\infty(\{x: \sum_{i=1}^\infty x_i^2 < \infty\})=0$, so $p_\infty$ is not a "softened" or "fuzzy" version of the inequality $\sum_{i=1}^\infty x_i^2 \leq 1$. If the prior information in a linear inverse problem consists of this inequality and nothing else, stochastic inversion and Bayesian inference are both unsuitable inversion techniques.
Introduction. Let $\mathbb{R}$ be the real numbers, $\mathbb{R}^n$ the linear space of all real $n$-tuples, and $\mathbb{R}^\infty$ the linear space of all infinite real sequences $x = (x_1, x_2, \ldots)$. Let $P_n : \mathbb{R}^\infty \rightarrow \mathbb{R}^n$ be the projection operator with $P_n(x) = (x_1, \ldots, x_n)$. Let $p_\infty$ be a probability measure on the smallest $\sigma$-ring of subsets of $\mathbb{R}^\infty$ which includes all of the cylinder sets $P_n^{-1}(B_n)$, where $B_n$ is an arbitrary Borel subset of $\mathbb{R}^n$. Let $p_n$ be the marginal distribution of $p_\infty$ on $\mathbb{R}^n$, so $p_n(B_n) = p_\infty(P_n^{-1}(B_n))$ for each $B_n$. A measure on $\mathbb{R}^n$ is "isotropic" if it is invariant under all orthogonal transformations of $\mathbb{R}^n$. The measure $p_\infty$ will be called isotropic if all its marginal distributions $p_n$ are isotropic. The set of all isotropic probability distributions on $\mathbb{R}^\infty$ will be written $I(\mathbb{R}^\infty)$. The present note describes all members of $I(\mathbb{R}^\infty)$. The result calls into question both stochastic inversion and Bayesian inference, as currently used in many geophysical inverse problems.

Necessary Conditions for Isotropy. Let $0 = (0, 0, \cdots)$ and let $p_0$ be the member of $I(\mathbb{R}^\infty)$ such that $p_0((0)) = 1$. If $p_\infty \in I(\mathbb{R}^\infty)$ and $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$, then $\alpha p_\infty + \beta p_\infty \in I(\mathbb{R}^\infty)$. Conversely, if $p_\infty \in I(\mathbb{R}^\infty)$ and $p_\infty((0)) = \beta$, then $p_\infty = (1-\beta)\bar{p}_\infty + \beta p_0$ where $\bar{p}_\infty \in I(\mathbb{R}^\infty)$ and $\bar{p}_\infty((0)) = 0$. Therefore it is necessary to study only those $p_\infty \in I(\mathbb{R}^\infty)$ for which $p_\infty((0)) = 0$. They constitute the subset $I_0(\mathbb{R}^\infty)$ of $I(\mathbb{R}^\infty)$.

If $p_\infty \in I_0(\mathbb{R}^\infty)$, for every $\xi$ in $0 \leq \xi < \infty$ define

$$F_n(\xi) = p_\infty((x : x_1^2 + \cdots + x_n^2 > \xi)).$$

Then $F_n$ is right semi-continuous, and

$$F_n(0) = 1$$

and

$$F_n(\infty) = \lim_{\xi \to \infty} F_n(\xi) = 0.$$

Also, if $n \leq N$ and $\alpha \leq A$, then

$$0 \leq F_n(A) \leq F_n(\alpha) \leq F_N(\alpha) \leq 1.$$

Properties sufficient to characterize the members of $I_0(\mathbb{R}^\infty)$ are given in

September 4, 1987
Theorem 1: Suppose \( p_m \in I_0(R^m) \) and \( F_n \) given by [1]. Then for each integer \( n \geq 1 \), \( F_n(\xi) \) is analytic in the open right half plane of complex \( \xi \). There is a function \( f_n(\xi) \), also analytic there, such that for every Borel subset \( B_n \) of \( R^n \)

\[
p_n(B_n) = \int_{B_n} dx_1 \cdots dx_n f_n(x_1^2 + \cdots + x_n^2).
\]

[3a]

In particular, if \( 0 \leq \alpha \leq \infty \) then

\[
F_n(\alpha) = \pi^{n/2} \Gamma(n/2)^{-1} \int_0^{\infty} d\xi \xi^{n/2-1} f_n(\xi).
\]

[3b]

The \( f_n \) are related by

\[
f_n(\xi) = \int_\xi d\eta (\eta - \xi)^{-1} f_{n+1}(\eta)
\]

[3c]

\[
f_{n+1}(\xi) = -\pi^{-1} \partial_\xi \int_\xi d\eta (\eta - \xi)^{-1} f_n(\eta)
\]

[3d]

\[
f_n(\xi) = \pi \int_\xi d\eta f_{n+2}(\eta)
\]

[3e]

\[
f_{n+2}(\xi) = -\pi^{-1} \partial_\xi f_n(\xi)
\]

[3f]

For every \( \beta \) in \( 0 \leq \beta < \infty \)

\[
\lim_{n \to \infty} F_n(\beta) = 1.
\]

[3g]

PROOF: Let \( S(n-1) \) denote the unit sphere in \( R^n \), and let \( |S(n-1)| \) be its \((n-1)\)-dimensional Euclidean content, \( 2\pi^{n/2} \Gamma(n/2)^{-1} \). Let \( |S(n-1)| \phi_n(\omega) \) be the content of the part of \( S(n-1) \) where \( x_n^2 \leq 1 - \omega \). Then

\[
\phi_{n+1}(w) = 1 - |S(n-1)||S(n)|^{-1} \int_0^w d\xi \xi^{n/2}(1-\xi)^{-1}.
\]

Since \( p_n \) is the marginal distribution on \( R^n \) of \( p_{n+1} \) on \( R^{n+1} \),

\[
F_n(\xi) = -\int_\xi dF_{n+1}(\eta) \phi_{n+1}(\xi|\eta),
\]

[4a]

the right side being a Stieltjes integral. For any \( \beta \) and \( B \) satisfying \( \xi < \beta < B \), \( \partial_\eta \phi_{n+1}(\xi|\eta) \) is continuous in \( \beta \leq \eta \leq B \), so integration by parts (1) permits the conclusion.
\[ B \int dF_{n+1}(\eta) \phi_{n+1}(\xi | \eta) + B \int d\eta F_{n+1}(\eta) \partial_\eta \phi_{n+1}(\xi | \eta) = F_{n+1}(B) \phi_{n+1}(\xi | B) - F_{n+1}(\beta) \phi_{n+1}(\xi | \beta). \]

Here let \( \beta \to \xi^+ \) and \( B \to \infty \). The integrated parts tend to zero, so the Lebesque bounded convergence theorem permits (4a) to be rewritten

\[ \xi^{-n/2} F_n(\xi) = \int d\eta \eta^{-(n+1)/2} F_{n+1}(\eta) (\eta - \xi)^{-1/2}. \]

Iterating this formula once, reversing orders of integration, and invoking the identity

\[ \int d\eta (\xi - \eta)^{-1/2} (\eta - \xi)^{-1/2} = \pi \]

leads to

\[ \xi^{-n/2} F_n(\xi) = (n/2) \int d\xi \xi^{-(n+2)/2} F_{n+2}(\xi). \]  

By induction on \( n \), it follows that \( F_n(\xi) \) is infinitely differentiable in \( 0 < \xi < \infty \). If we define

\[ f_n(\xi) = -\pi^{-n/2} \Gamma(n/2) \xi^{1-n/2} \partial_\xi F_n(\xi), \]

then \( f_n \) is also infinitely differentiable in \( 0 < \xi < \infty \) and [2b] yields [3b]. Then [3a] follows by straightforward integration theory. Then the definition of marginal distributions implies

\[ f_n(x_1^2 + \cdots + x_n^2) = \int_{-\infty}^{\infty} dx_{n+1} f_{n+1}(x_1^2 + \cdots + x_{n+1}^2), \]  

which is [3c] with \( \xi = x_1^2 + \cdots + x_2^2, \eta = x_1^2 + \cdots + x_{n+1}^2 \). Also,

\[ f_n(x_1^2 + \cdots + x_n^2) = \int_{-\infty}^{\infty} dx_{n+1} \int_{-\infty}^{\infty} dx_{n+2} f_{n+2}(x_1^2 + \cdots + x_{n+2}^2), \]  

which is [3e]. Then [3f] follows from [3e], and [3d] follows from [3f] and [3c] with \( n \) replaced by \( n-1 \). To prove analyticity, note that if \( q \) is an integer \( \geq 0 \) and if \( 0 < \alpha < \beta \), then by Taylor's theorem with remainder

\[ F_2(\alpha) - F_2(\beta) = \sum_{i=1}^{q} \frac{(-\alpha)^i}{i!} F_2(\beta) + \frac{1}{\alpha} \int d\xi (\xi - \alpha)^q (-\partial_\xi)^{q+1} F_2(\xi). \]
But \((-\partial_\xi)^{\beta} F_2=\pi^1 f_2\), so by [3b]

\[
\frac{1}{q!} \int_A d\xi \xi^q (-\partial_\xi)^{q+1} F_2(\xi) = F_{2q+2}(\alpha) - F_{2q+2}(\beta). \tag{6b}
\]

Hence, the Lebesque bounded convergence theorem implies that as \(\alpha \to 0\) the integral in [6a] converges to \(1-F_{2q+2}(\beta)\). Therefore

\[
F_{2q+2}(\beta) - F_2(\beta) = \sum_{i=1}^{q} \frac{\beta^i}{i!} (-\partial_\xi)^{i} F_2(\beta). \tag{6c}
\]

All terms in the sum [6c] are nonnegative, and \(F_{2q+2}(\beta) \leq 1\), so the series

\[
\sum_{i=1}^{\infty} \frac{(-\beta)^i}{i!} F_2^{(i)}(\beta) \tag{6d}
\]

converges absolutely (here \(F_2^{(i)} = \partial_\xi F_2\)). Therefore, the power series for \(F_2(\xi)\) at \(\xi = \beta\) converges absolutely for all complex \(\xi\) in the closed disk \(1|\xi - \beta| \leq \beta\). Since \(\beta\) is arbitrary, \(F_2(\xi)\) is analytic for all complex \(\xi\) with positive real part. By [5a], so is \(f_2(\xi)\) and then by [3c,d] so is \(f_n(\xi)\) for every \(n \geq 1\). Hence so is \(F_n(\xi)\) for every \(n \geq 1\). Furthermore, since [6d] converges, Abel’s theorem (2) implies that

\[
F_2(0) - F_2(\beta) = \sum_{i=1}^{\infty} \frac{\beta^i}{i!} (-\partial_\xi)^{i} F_2(\beta). \tag{6e}
\]

Together, [6e], [6c] and [2a] imply [3g].

**COROLLARY 1:** If one of the marginal distributions \(p_n\) is known, \(p_\infty\) is completely determined.

**COROLLARY 2:** Let \(H(\alpha)\) be the set of \(x\) in \(R^n\) with \(\sum x_i^2 < \alpha\). Then \(p_\infty(H(\infty)) = 0\). This follows immediately from [3g] and the fact that \(H(\infty)\) is the monotone limit of the sets \(H(\alpha)\) (3).

**Sufficient Conditions for Isotropy.** Let \(M(n)\) be the set of infinitely differentiable real-valued functions \(f\) on the open half-line \(0 < \xi < \infty\) such that

\[
\pi^{n/2} \Gamma(n/2)^{-1} \int_0^\infty d\xi \xi^{n/2-1} f(\xi) = 1. \tag{7a}
\]

September 4, 1987
and also for every integer $q \geq 0$ and every $\xi$ in $0 < \xi < \infty$

$$(-\partial_\xi)^q f(\xi) \geq 0.$$  \hspace{1cm} [7b]

Note that if $p_\infty \in I_0(R^n)$ and $f_\infty$ comes from $p_\infty$ via [3a] then $f_\infty \in M(n)$. The converse is also true, and to prove it we need

**LEMMA 1:** Suppose $n \geq 1$ and $f \in M(n)$. Then

1. $\lim_{\xi \to \infty} \xi^{n/2} f(\xi) = 0$  \hspace{1cm} [8a]
2. $\lim_{\xi \to 0} \xi^{n/2} f(\xi) = 0$  \hspace{1cm} [8b]
3. $f(\xi) = \int_\xi^\infty d\eta (-\partial_\eta f(\eta))$  \hspace{1cm} [8c]
4. $(n/2) \int_0^\infty d\xi \xi^{n/2-1} f(\xi) = \int_0^\infty d\xi \xi^{n/2} [-\partial_\xi f(\xi)]$  \hspace{1cm} [8d]
5. $-\pi^{-1} \partial_\xi f \in M(n+2)$  \hspace{1cm} [8e]

**PROOF:** Let $m = n/2 - 1$ and let $0 < \alpha < A < \infty$. Integration by parts gives

$$A \int_a^\infty d\xi \xi^m f(\xi) = A^{m+1} f(A) - \alpha^{m+1} f(\alpha) + \int_\alpha^A d\xi \xi^{m+1} [-\partial_\xi f(\xi)].$$  \hspace{1cm} [9a]

Fix $\alpha$. The integral on the right in [9a] increases as $A \to \infty$ and yet is bounded, so it has a limit. Therefore $\lim_{A \to \infty} A^{m+1} f(A)$ exists. By [7a] it cannot be positive, so we have [8a], and hence [8c], and also

$$A \int_a^\infty d\xi \xi^m f(\xi) = -\alpha^{m+1} f(\alpha) + \int_\alpha^\infty d\xi \xi^{m+1} [-\partial_\xi f(\xi)].$$  \hspace{1cm} [9b]

As $\alpha$ decreases to 0, the integral on the right in [9b] increases, and that on the left has a finite limit, so $\alpha^{m+1} f(\alpha)$ approaches either $+\infty$ or a nonnegative limit. Then [7a] requires [8b], and [9b] converges to [8d]. Then [8c] follows from [8d] and [7b].

Now we can prove

**THEOREM 2:** Suppose $n$ is a nonnegative integer and $f \in M(n)$. Then there is a $p_\infty \in I_0(R^n)$ whose marginal distribution $p_\infty$ on $R^n$ is given by [3a] with $f_\infty = f$.
PROOF: For every integer $q \geq 0$, define $f_{n+2q}(\xi) = \pi^{-q} (-\partial_\xi)^q f(\xi)$. If $N-n$ is a nonnegative even integer, induction on [8c] implies

$$f_N(x_1^2 + \cdots + x_N^2) = \int dx_{N+1} \int dx_{N+2} f_{N+2}(x_1^2 + \cdots + x_{N+2}^2).$$  \[10a\]

If $N-n$ is a nonnegative odd integer, define $f_N$ from $f_{N+1}$ via [3c]. Then

$$f_N(x_1^2 + \cdots + x_N^2) = \int dx_{N+1} f_{N+1}(x_1^2 + \cdots + x_{N+1}^2).$$  \[10b\]

That [10b] also holds when $N-n$ is nonnegative and even follows from [10a]. Therefore [10b] holds for all $N \geq n$. Use it inductively to define $f_N$ for $1 \leq N < n$. For $N=n$, [7a] implies

$$\int_{R^n} dx_1 \cdots dx_N f_N(x_1^2 + \cdots + x_N^2) = 1,$$  \[10c\]

and then [10b] implies [10c] for all $N \geq 1$. Thus the probability distributions $p_N$ on $R^N$ given by $f_N$ via [3a] satisfy the Kolmogorov consistency condition. Then the existence of $p_\infty$ follows from Kolmogorov’s Fundamental Theorem (4).

**COROLLARY 1:** If $f \in M(n)$, $f(\xi)$ is analytic in the open right half-plane of complex $\xi$.

**COROLLARY 2:** The equation $F_2(\xi) = \int_0^\infty d\mu(t) e^{-\xi t}$ furnishes a bijection between the members of $I_0(R^m)$ and the probability measures $\mu$ on $0 \leq t < \infty$.

**PROOF:** Demanding that $f_2 \in M(2)$ is equivalent to demanding that $F_2(\xi)$ be completely monotonic on $0 \leq \xi < \infty$ (5).

**Examples and Applications.** Setting $f_2(\xi) = \pi^{-1} e^{-\xi}$ gives $f_n(\xi) = \pi^{-n/2} e^{-\xi}$. This $p_\infty$ is the gaussian with independent $x_1, x_2, \ldots$, each having mean 0 and variance 1. Setting $f_2(\xi) = \pi^{-1} \nu(\xi)^{v-1} (1+\xi)^{-v}$ with $0 < v < 1$ gives a $p_\infty$ for which $\lim_{\xi \to 0} f_n(\xi) = \infty$ if $n \leq 2$ and also if $n=1$ and $\nu/2 < v < 1$. Thus the densities $f_n(\xi)$ need not remain finite as $\xi \to 0$.

The geophysical application is to inverse theory. An infinite dimensional linear space $X$ of earth models $x$ is given, along with a finite number of linear functionals, $g_j : X \to R$, September 4, 1987
An observer measures \( D \) data \( y_i = g_i(x_E) + \varepsilon_i \) for \( i = 1, \ldots, D \). Here \( x_E \) is the correct earth model and \( \varepsilon_i \) is the error in observing \( y_i \). The observer wants to predict the value of 
\[ z = g_{D+1}(x_E). \]
Since \( \text{dim} \ X = \infty \), the problem is hopeless unless \( g_{D+1} \) is a linear combination of \( g_1, \ldots, g_D \), or unless the observer has some prior information about \( x_E \) not included among the data (6,7). One common sort of prior information is a quadratic bound on \( x_E \), a quadratic form \( Q \) on \( X \) such that \( x_E \) is known to satisfy
\[
Q(x_E, x_E) \leq 1. \tag{11}
\]
Often (11) is a bound on energy content or dissipation rate (8). In stochastic inversion and Bayesian inference, such a bound is often "softened" to a prior personal probability distribution \( p_\infty \) on \( X \) (8–10). In practice, \( X \) is truncated to an \( \mathbb{R}^n \), and \( p_\infty \) is used in the inversion.

To see why this process is questionable, complete \( X \) to a Hilbert space with the inner product \( x \cdot x' = Q(x, x') \). Let \( \xi_1, \xi_2, \ldots \) be an orthonormal basis for \( X \), and write \( x = \sum_{i=1}^{\infty} x_i \xi_i \).

Then \( X \) becomes the subset \( H(\infty) \) of \( \mathbb{R}^\infty \) defined in corollary 2 to theorem 1. The prior information (11) can now be written
\[
\sum_{i=1}^{\infty} x_i^2 \leq 1. \tag{12}
\]
If the observer wants to soften (12) to a probability distribution \( p_\infty \) without introducing new information not implied by (12), then clearly he should take \( p_\infty \in I(\mathbb{R}^\infty) \). He is unlikely to assign nonzero probability to 0, so \( p_\infty \in I_0(\mathbb{R}^\infty) \). But then \( p_\infty(X) = 0 \) by corollary 2 to theorem 1. Any prior personal probability distribution obtained by softening (12) without adding new information must deny (12) with probability 1.

Acknowledgments. The author is grateful for partial support from NSF grant EAR 85-21543 and NASA grant NAG 5-818.
References


