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 $x=(x_1,x_2,\ldots)$ 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 $p_\omega$ on $\mathbb{R}^\infty$ with $p_\omega(\{0\})=0$. Each $p_\omega \in I_0(\mathbb{R}^\infty)$ is completely determined by any one of its finite-dimensional marginal distributions $p_n$. Each $p_n$ has a density function $f_n$ with $dp_n(x_1,\ldots,x_n)=dx_1\cdots dx_n f_n(x_1^2+\cdots+x_n^2)$. 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 $p_\omega \in I_0(\mathbb{R}^\infty)$. Hence the equation

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

defines a bijection between $I_0(\mathbb{R}^\infty)$ and the set of all probability measures $\mu$ on $0\leq t<\infty$. If $p_\omega \in I_0(\mathbb{R}^\infty)$ then $p_\omega(\{x : \sum_{i=1}^\infty x_i^2<\infty\})=0$, so $p_\omega$ 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 $R$ be the real numbers, $R^n$ the linear space of all real $n$-tuples, and $R^\infty$ the linear space of all infinite real sequences $x = (x_1, x_2, \ldots)$. Let $P_n : R^\infty \to R^n$ 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 $R^\infty$ which includes all of the cylinder sets $P_n^{-1}(B_n)$, where $B_n$ is an arbitrary Borel subset of $R^n$. Let $p_n$ be the marginal distribution of $p_\infty$ on $R^n$, so $p_n(B_n) = p_\infty(P_n^{-1}(B_n))$ for each $B_n$.

A measure on $R^n$ is "isotropic" if it is invariant under all orthogonal transformations of $R^n$. 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 $R^\infty$ will be written $I(R^\infty)$. The present note describes all members of $I(R^\infty)$. 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 $p_\infty^0$ be the member of $I(R^\infty)$ such that $p_\infty^0(\{0\}) = 1$. If $p_\infty \in I(R^\infty)$ and $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$, then $\alpha p_\infty + \beta p_\infty^0 \in I(R^\infty)$.

Conversely, if $p_\infty \in I(R^\infty)$ and $p_\infty(\{0\}) = \beta$, then $p_\infty = (1-\beta)p_\infty^0 + \beta p_\infty^0$ where $p_\infty^0 \in I(R^\infty)$ and $p_\infty^0(\{0\}) = 0$. Therefore it is necessary to study only those $p_\infty \in I(R^\infty)$ for which $p_\infty(\{0\}) = 0$.

They constitute the subset $I_0(R^\infty)$ of $I(R^\infty)$.

If $p_\infty \in I_0(R^\infty)$, 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

$$F_n(0) = 1 \quad [2a]$$

$$F_n(\infty) = \lim_{\xi \to \infty} F_n(\xi) = 0 \quad [2b]$$

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

$$0 \leq F_n(A) \leq F_n(\alpha) \leq F_N(\alpha) \leq 1 \quad [2c]$$

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

September 4, 1987
Theorem 1: Suppose \( p_n \in I_0(R^n) \) 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 \( R^n \)

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

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

\[
F_n(\alpha) = \pi^{n/2} \Gamma(n/2)^{-1} \int_{0}^{\infty} d\xi \xi^{n/2-1} f_n(\xi).
\]

The \( f_n \) are related by

\[
f_n(\xi) = \frac{d}{d\xi} f_{n+1}(\xi) \quad \text{for } \xi \neq 0.
\]

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

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

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

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

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

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

\[
\phi_{n+1}(w) = 1 - 1S(n-1)1 1S(n)1^{-1/2} \int_{0}^{w} d\xi \xi^{n/2}(1-\xi)^{-1/2}.
\]

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

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

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.
\[ \int_\beta^B dF_{n+1}(\eta)\phi_{n+1}(\xi, \eta) + \int_\beta^B d\eta F_{n+1}(\eta) \partial_\eta \phi_{n+1}(\xi, \eta) \]

\[ = F_{n+1}(B)\phi_{n+1}(\xi/B) - F_{n+1}(\beta)\phi_{n+1}(\xi/\beta). \]

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

\[ \xi^{-n/2} \xi^+ = \int d\eta \eta^{-1/2} \eta^{-\xi^+} \]

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

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

leads to

\[ \xi^{-n/2} \xi^+ = (n/2) \int_\xi^\infty d\xi \xi^{-n/2} \xi^{-1/2} \xi^{-1/2} F_{n+2}(\xi). \]

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

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

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

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

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

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

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

\[ F_2(\alpha) - F_2(\beta) = \sum_{i=1}^q \frac{(\beta - \alpha)^i}{i!} (-\partial_\xi)^i F_2(\beta) + \frac{1}{q} \int d\xi (\xi - \alpha)^q (-\partial_\xi)^q F_2(\xi). \]
But \((-\partial^{i}_{\xi}) F_2 = \pi^{i} f_{2i}\), so by \([3b]\) 

\[
\frac{1}{q!} \int_{\alpha}^{\beta} d\xi \xi^{q} (-\partial^{i}_{\xi}) \xi^{q+1} F_2(\xi) = F_{2q+2}(\alpha) - F_{2q+2}(\beta).
\]  

[6b] 

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

\[
F_{2q+2}(\beta) - F_2(\beta) = \sum_{i=1}^{\infty} \frac{\beta^{i}}{i!} (-\partial^{i}_{\xi}) F_2(\beta).
\]  

[6c] 

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

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

[6d] 

converges absolutely (here \(F_2^{(i)} = \partial^{i}_{\xi} F_2\)). Therefore, the power series for \(F_2(\xi)\) at \(\xi = \beta\) converges absolutely for all complex \(\xi\) in the closed disk \(1 \leq \beta \leq \xi \leq \beta \leq 1\). 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 

\[
F_2(0) - F_2(\beta) = \sum_{i=1}^{\infty} \frac{\beta^{i}}{i!} (-\partial^{i}_{\xi}) F_2(\beta).
\]  

[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 \(R^m\) with \(\sum_{i=1}^{\infty} x_i^2 < \alpha\). 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 

\[
\pi^{n/2} \Gamma(n/2)^{-1} \int_{0}^{\infty} d\xi \xi^{n/2-1} f(\xi) = 1
\]  

[7a] 

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

\[
(-\partial_\xi)^q f(\xi) \geq 0.
\]  

[7b]

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

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

\[
\lim_{\xi \to \infty} \xi^{n/2} f(\xi) = 0
\]  

[8a]

\[
\lim_{\xi \to 0} \xi^{n/2} f(\xi) = 0
\]  

[8b]

\[
f(\xi) = \int \frac{d\eta}{\xi} [-\partial_\eta f(\eta)]
\]  

[8c]

\[
(n/2) \int_0^\infty d\xi \xi^{n/2-1} f(\xi) = \int_0^\infty d\xi \xi^{n/2} [-\partial_\xi f(\xi)]
\]  

[8d]

\[
-\pi^{-1} \partial_\xi f \in M(n+2).
\]  

[8e]

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

\[
(m+1) \int_A^A d\xi \xi^m f(\xi) = A^{m+1} f(A) + \int_0^A d\xi \xi^m [-\partial_\xi f(\xi)].
\]  

[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

\[
(m+1) \int_0^\infty d\xi \xi^m f(\xi) = -\alpha^{m+1} f(\alpha) + \int_0^\infty d\xi \xi^m [-\partial_\xi f(\xi)].
\]  

[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 [8e] 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 \( p_\infty \in I_0(\mathbb{R}^\infty) \) whose marginal distribution \( p_n \) on \( \mathbb{R}^n \) is given by [3a] with \( f_n = f \).

September 4, 1987
**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

\[
 f_N(x_1^2 + \cdots + x_N^2) = \int_{R^n} \cdots \int_{R^n} dx_{N+1} f_{N+2}(x_1^2 + \cdots + x_{N+2}^2). \tag{10a}
\]

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

\[
 f_N(x_1^2 + \cdots + x_N^2) = \int_{R^n} \cdots \int_{R^n} dx_{N+1} f_{N+1}(x_1^2 + \cdots + x_{N+1}^2). \tag{10b}
\]

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

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

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 \( f_2(\xi) = \pi^{-1} [\xi e^{\nu - 1} - (1 + \xi)^{\nu - 1}] \) with \( 0 < \nu < 1 \) gives a \( p_\infty \) for which \( \lim_{\xi \to 0} f_n(\xi) = \infty \) if \( n \leq 2 \) and also if \( n = 1 \) and \( 1/2 \leq \nu < 1 \). 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, \( g_j : X \to R \),
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 $\dim X = \infty$, 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

$$Q(x_E, x_E) \leq 1.$$  \hfill [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 $R^n$, 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 $R^\infty$ defined in corollary 2 to theorem 1. The prior information [11] can now be written

$$\sum_{i=1}^{\infty} x_i^2 \leq 1.$$  \hfill [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 $p_\infty \in I(R^\infty)$. He is unlikely to assign nonzero probability to $0$, so $p_\infty \in I_0(R^\infty)$. 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


