

Inverse Problems and a Lévy Process Solution

Martin Bøgsted Hansen
Department of Mathematical Sciences
Aalborg University: DK-9220 Aalborg Øst, Denmark

Katja Ickstadt
Department of Mathematics
Darmstadt University of Technology: 64289 Darmstadt, Germany

Robert L. Wolpert
Institute of Statistics and Decision Sciences
Duke University: Durham, NC 27708-0251, USA

00-30*†

Revised October 11, 2000, 09:37 EDT
Durham, NC

Abstract

We propose a new method of making inference about an unknown measure $F(ds)$ upon observing some values of the integral $g(x) = \int k(x, s)F(ds)$ of a known kernel $k(x, s)$, using Lévy processes as Bayesian prior distributions for modeling uncertainty about $F(ds)$, using simulation-based MCMC methods. The methods are illustrated with two problems in polymer chemistry.

Key words: Lévy process; Fredholm integral equations; Inference; Inverse problem; Polymer.

* Available at url: <http://www.isds.duke.edu/pub/papers>

† The authors acknowledge gratefully the support of US NSF grant DMS-9626829, and of US EPA grant XXX-0000000. We would like to thank Robert S. Anderssen for suggesting the problem, Otto Glatter for making data available to us, and Jesper Møller and the Department of Mathematical Sciences of Aalborg University for facilitating our collaboration.

1 Introduction

In many problems arising in science and engineering we wish to learn about an unknown measure $F(ds)$ on a measurable space \mathcal{S} , on the basis of observations $\{Y_i\}_{i \in I}$ of a finite set of conditionally independent quantities (indexed by some set I) whose sampling distributions are determined by the parameters given by

$$Y_i \sim \psi_i(y | g_i), \quad g_i \equiv \int_{\mathcal{S}} k_i(s) F(ds), \quad i \in I, \quad (1)$$

each g_i a linear functional of the uncertain F . In this paper we usually take the observations Y_i to have independent lognormal distributions with median g_i and precision parameter τ_i , so that

$$Z_i \equiv \sqrt{\tau_i} \log(Y_i/g_i) \sim \text{No}(0, 1) \quad (2)$$

would be *i.i.d.* standard normal random variables.

Examples include:

1. **Prior Elicitation**, where $F(ds)$ would represent an expert's prior distribution, $k_i(s)$ the value of an observable quantity O_i , and Y_i the expert's prediction of $\mathbb{E}[O_i] = \int_{\mathcal{S}} k_i(s) F(ds)$ under his or her prior distribution $F(ds)$;
2. **Moment Problems**, where we infer an unknown probability distribution $F(ds)$ from its moments, the expectations of $k_i(s) \equiv s^i$, or more generally from the expectations of arbitrary functions $k_i(s)$;
3. **Diffusion**, where $F(ds)$ might represent the distribution at time $t = 0$ of some environmental pollutant, $k_i(s) = K_{t_i}(x_i, s)$ the Fickian diffusion kernel giving the concentration at later times $t_i > 0$ and locations x_i arising from an initial unit concentration at $s \in \mathcal{S}$, and Y_i the reported measurements (subject to lognormal error) at sites x_i of the concentration $g_i = g(x_i) = \int_{\mathcal{S}} K_{t_i}(x_i, s) F(ds)$;
4. **Nonparametric Bayes** The density functions of symmetric unimodal distributions $g_s(x)$ on $\mathcal{X} = \mathbb{R}$ and concave-CDF distributions $g_c(x)$ on $\mathcal{X} = \mathbb{R}_+$ have well-known representations $g_s(x) = \int_{|x|}^{\infty} (2s)^{-1} F(ds)$ and $g_c(x) = \int_x^{\infty} s^{-1} F(ds)$ respectively, each of form Eqn. (1) for $g_i \equiv g(x_i)$ with kernels $k_s(x, s) = (2s)^{-1} \mathbf{1}_{|x| < s}$ and $k_c(x, s) = s^{-1} \mathbf{1}_{x < s}$, respectively, on $\mathcal{X} \times \mathcal{S}$ with $\mathcal{S} = \mathbb{R}_+$;

5. **Poisson/gamma point process models** Models introduced and applied in (Wolpert and Ickstadt, 1998a; Best *et al.*, 2000) feature Cox point processes with uncertain Poisson intensities $g_i \equiv g(x_i)$ of form Eqn. (1) with Gaussian kernels $k(x, s)$ and measures $F(ds)$ modeled as gamma random fields. The methods described in (Wolpert and Ickstadt, 1998a) are more efficient than the present ones for the conjugate Poisson/gamma processes employed there but the present methods, which neither require nor exploit conjugacy, allow the gamma random field to be replaced by heavier-tailed Lévy random fields to achieve a more robust Bayesian analysis.

In Sec. 3 we study in some detail two further examples concerning polymer structures.

1.1 Previous Attempts

A number of ways have been proposed for solving Fredholm integral equations of the first kind like Eqn. (1). A common starting point is to replace the integral in Eqn. (1) by a discrete sum $\sum_{j \in J} k_i(s_j) F_j$ for some partition $\cup S_j = \mathcal{S}$ and points $s_j \in S_j$, with $F_j \equiv F(S_j)$. The solution $F = (F_j)_{j \in J}$ to the linear system $g = K F$, with $g = (g_i)_{i \in I}$ and $K = \{k_{ij} \equiv k_i(s_j)\}_{i, j \in I, J}$, gives an approximation to the solution $F(ds)$ of Eqn. (1) on the partition.

More generally we may write Eqn. (1) in the form $g(x) = \int_{\mathcal{S}} k(x, s) F(ds)$ for x in some measure space \mathcal{X} . In the *Galerkin method* we now choose arbitrary orthonormal bases $\{\phi_i\}$ for \mathcal{X} and $\{\psi_j\}$ for \mathcal{S} and then solve the linear system of equations $g = K f$ for the vector $f = \{f_j \equiv \int_{\mathcal{S}} \psi_j(s) F(ds)\}$, where $g_i \equiv \int_{\mathcal{X}} g(x) \phi_i(x) dx$ and $K = \{k_{ij} \equiv \iint_{\mathcal{X} \times \mathcal{S}} \phi_i(x) k(x, s) \psi_j(s) dx ds\}$.

An intrinsic problem for any method of solving Fredholm integral equations of the first kind is their ill-posedness: even small fluctuations in data may lead to large fluctuations in the solution, because the integral operator tends to smooth any moderate variations. Discretization leads to an ill-conditioned linear system of equations, making all standard methods for solution unreliable. Additionally, achieving adequate resolution of F in the continuous problem usually leads to the requirement $|J| \gg |I|$, leading to a strongly underdetermined discrete problem.

Various methods have been proposed for handling large discrete ill-posed problems (see Hansen, 1998, for a recent survey). All are based on variations of the *method of regularization* (Tikhonov, 1963), choosing a solution to minimize some combination of a discrepancy penalty (often a sum of

squared residual errors) and a roughness penalty. Methods of this type have appeared in the statistics literature under names as ridge regression (Hoerl and Kennard, 1970a,b), penalized likelihood estimation (Good and Gaskins, 1971), method of sieves (Grenander, 1981) and the maximum entropy method.

There is an enormous literature on this subject dating from the early sixties covering a range of computational, mathematical and statistical and scientific aspects of regularization. For more details consult the extensive survey on numerical aspects by Hansen (1998), the research monograph on mathematical aspects by Kirsch (1996), the review papers on statistical perspectives by O’Sullivan (1986) and Stark (2000), and the description of promising new Galerkin-based wavelet methods by Abramovich *et al.* (2000, §3.3).

All regularization methods have a Bayesian interpretation, with the discrepancy penalty playing the role of a likelihood and the roughness penalty playing the role of a prior probability distribution, leading to a regularized estimate that may be interpreted as a maximum *a posteriori* (MAP) estimate. Most regularizing discretization methods lead to estimates \hat{F} of F expressed as linear combinations of suitably weighted basis functions, $\hat{F}(ds) = \sum_{j \in J} a_j \psi_j(s) ds$. The Bayesian approach now amounts to specifying a prior distribution on the roughness penalty through a joint distribution of coefficients $\{a_j\}$ in this series expansion (Kimeldorf and Wahba, 1971).

A distinction of the present approach is that it is fully Bayesian, allowing the direct incorporation of any prior evidence or belief about $F(ds)$ (not just contrived roughness penalties) and sensible measurement error models (not just discrepancy penalties), leading to full posterior distributions and posterior means for $F(ds)$ (not just posterior modes, or MAP estimates) (see also Hansen and Lauritzen, 1999). *BE MORE SPECIFIC: in this paragraph we say how our approach differs from earlier ones. The citation could then be interpreted such that people think it’s all done in Hansen and Lauritzen, 1999. Not true, but at least very misleading at the moment.*

2 Computational Method

We use the Markov chain Monte Carlo method (Gelfand and Smith, 1990; Tierney, 1994) to estimate the uncertain measure $F(ds)$. In this widely applicable approach to Bayesian inference about an uncertain parameter $\theta \in \Theta$, on the basis of a prior distribution $\pi(d\theta)$ and the likelihood function

$L(\mathbf{Y} \mid \theta)$ for an observed vector \mathbf{Y} , we construct an ergodic Markov Chain $\theta^t \in \Theta$ whose stationary distribution is exactly the posterior distribution $\pi(d\theta \mid \mathbf{Y}) \propto L(\mathbf{Y} \mid \theta) \pi(d\theta)$, so that posterior expectations of arbitrary functions $\phi(\theta)$ may be approximated by ergodic averages

$$\mathbb{E} \left[\int_{\Theta} \phi(\theta) \mid \mathbf{Y} \right] \approx \frac{1}{T} \sum_{t \leq T} \phi(\theta^t).$$

The Metropolis-Hastings variation of the method begins with the selection of a Markov kernel $Q(\theta, d\theta^*)$ on Θ as a “proposal distribution” and an arbitrary initial value θ^0 . At each time $t \geq 0$ a proposed step $\theta^* \sim Q(\theta^t, d\theta^*)$ is drawn; the proposal is accepted, whereupon $\theta^{t+1} = \theta^*$, or rejected, so that $\theta^{t+1} = \theta^t$, with probabilities $\alpha(\theta^t, \theta^*)$ and $1 - \alpha(\theta^t, \theta^*)$, respectively, given by

$$\alpha(\theta, \theta^*) \equiv \min \left[1, \frac{\pi(d\theta^*) L(\mathbf{Y} \mid \theta^*) Q(\theta^*, d\theta)}{\pi(d\theta) L(\mathbf{Y} \mid \theta) Q(\theta, d\theta^*)} \right]. \quad (3)$$

Under mild conditions (Tweedie and Mengersen, 1996) the resulting sequence θ^t can be shown to be an ergodic Markov chain with the posterior distribution $\pi(d\theta \mid \mathbf{Y})$ as its stationary distribution. To implement this approach we must select a prior distribution, a measurement-error model leading to a likelihood function, and a proposal distribution.

2.1 Lévy Process Prior Distributions

Our nonparametric Bayesian approach begins by modeling the measure $F(ds)$ *a priori* as an increasing Lévy process $F \sim \text{Lv}(\nu(du ds))$ (Rogers and Williams, 1987, p. 313), expressible as a finite or countable sum

$$F(ds) = \sum_m v_m \delta_{\sigma_m}(ds) \quad (4)$$

of point masses of random size v_m at random locations σ_m drawn jointly from a Poisson point process on $\mathbb{R}_+ \times \mathcal{S}$ with intensity measure $\nu(du ds)$ satisfying an integrability condition $\iint_{\mathbb{R}_+ \times \mathcal{S}} \min(1, u) \nu(du ds) < \infty$. This will assign independent random variables $F(A_i)$ to disjoint sets $A_i \subset \mathcal{S}$, with characteristic functions

$$\mathbb{E}[e^{i\omega F(A)}] = e^{-\iint_{\mathbb{R}_+ \times A} (1 - e^{i\omega u}) \nu(du ds)}$$

and hence with (finite or infinite) prior mean $\mathbf{E}[F(A)] = \iint_{\mathbb{R}_+ \times A} u \nu(du ds)$ and variance $\mathbf{V}[F(A)] = \iint_{\mathbb{R}_+ \times A} u^2 \nu(du ds)$. Familiar examples of Lévy processes include the **Poisson process**, with $\nu(du ds) = \delta_1(du) \alpha(ds)$, assigning Poisson-distributed random variables $F(A) \sim \text{Po}(\alpha(A))$ with mean $\alpha(A)$; the **gamma process**, with $\nu(du ds) = e^{-u\beta} u^{-1} du \alpha(ds)$, assigning gamma-distributed random variables $F(A) \sim \text{Ga}(\alpha(A), \beta)$ with shape $\alpha(A)$ and rate β ; and the (one-sided) **stable process**, with $\nu(du ds) = \xi u^{-1-\xi} du \alpha(ds)$, assigning fully-skewed stable random variables $F(A) \sim \text{St}(\xi, \alpha(A))$ with stable index $\xi \in (0, 1)$ and scale $\alpha(A)$.

Note that we do not impose stationary increments (*i.e.*, $\nu(du ds)$ need not have product form $\eta(du) ds$); see Jacod and Shiryaev (1987, Chapter II, §4c) for more details about this generalization of the usual Lévy process definition, and Wolpert and Ickstadt (1998a,b) for an explicit construction, the ILM (inverse Lévy measure) algorithm.

For implementation purposes we truncate if necessary to achieve a Lévy measure with finite total mass $\nu(\mathbb{R}_+ \times \mathcal{S}) < \infty$, so that $F(ds)$ will be supported (almost surely) on only finitely many points, and so can be represented $F(ds) = F^\theta(ds)$ for $\theta = ((u_1, s_1), \dots, (u_M, s_M))$ in the disjoint union $\Theta = \cup_{m < \infty} (\mathbb{R}_+ \times \mathcal{S})^m$. In each of the examples we consider here the Lévy measure will have a density function $\nu(u, s)$ with respect to a Lebesgue product reference measure $m(du, ds)$ on $u \in \mathbb{R}_+$ and $s \in \mathcal{S} \subset \mathbb{R}^d$, so the prior distribution will be absolutely continuous with respect to the reference Poisson $\text{Po}(m(du ds))$ random field, with prior log density function

$$\log \pi(\theta) = \kappa + \sum_{m \leq M} \log \nu(v_m, \sigma_m), \quad (5)$$

where $\kappa = m(\mathbb{R}_+ \times \mathcal{S}) - \nu(\mathbb{R}_+ \times \mathcal{S})$ is constant.

2.2 Lognormal Likelihood Functions

Inverse Eqn. (1) with lognormal prior distributions $Y_i \sim \text{LN}(g_i, \tau_i)$ leads to a log likelihood function from Eqn. (2) of the form

$$\log L(\mathbf{Y} | \theta) = \sum_{i \in I} \left(\frac{1}{2} \log \tau_i - \frac{\tau_i}{2} \log^2 \frac{Y_i}{g_i} - \log Y_i \right),$$

with $g_i \equiv \sum_m v_m k_i(\sigma_m)$. When the τ_i are treated as certain, and the Y_i are observed, this (with a different arbitrary constant) simplifies to

$$\log L(\mathbf{Y} \mid \theta) = -\frac{1}{2} \sum_{i \in I} \tau_i \log^2(g_i/Y_i). \quad (6)$$

2.3 Proposal Distributions

In all of our examples, $\mathcal{S} \subset \mathbb{R}^d$ for some integer d and $\nu(du, ds)$ is supported on $[\epsilon, \infty) \times \mathcal{S}$ for some $\epsilon > 0$. We propose two sorts of moves: Gaussian random walk steps in σ_m^t and (on a log scale) v_m^t , removing points that step outside $[\epsilon, \infty) \times \mathcal{S}$; and immigration steps, adding new points (σ_m^*, v_m^*) from an arbitrary entrance distribution:

1. Select $m \in \{1..M^t\}$ uniformly and draw $z_v^t, z_\sigma^t \sim \text{No}(0, 1)$ independently from standard normal distributions. Set $v_m^* \equiv v_m^t e^{s_v z_v^t}$ and $\sigma_m^* \equiv \sigma_m^t + s_\sigma z_\sigma^t$, if $v_m^* > \epsilon$ and $\sigma_m^* \in \mathcal{S}$, keeping $M^* = M^t$ and all remaining $v_n^* = v_n^t$, $\sigma_n^* = \sigma_n^t$ for $n \neq m$; otherwise remove (v_m^t, σ_m^t) and set $M^* \equiv M^t - 1$.
2. Set $m \equiv M^* \equiv M^t + 1$ and introduce a new point $(v_m^*, \sigma_m^*) \sim \mu(v, \sigma) dv d\sigma$ for an arbitrary entrance distribution with density $\mu(v, \sigma)$ on $[\epsilon, \infty) \times \mathcal{S}$.

If we denote the probabilities of these two moves by p_1 and $p_2 \equiv 1 - p_1$, then $Q(\theta, d\theta^*)$ is given by:

$$Q(\theta, d\theta^*) = \begin{cases} \frac{p_1}{2\pi M s_v^2 s_\sigma^2 v_m^*} e^{-\log^2(v_m^*/v_m) / 2s_v^2 - (\sigma_m^* - \sigma_m)^2 / 2s_\sigma^2} & M^* = M \\ \frac{p_1}{M} \Phi((\epsilon - v_m)/s_v) & M^* = M - 1 \\ p_2 \mu(v_m, \sigma_m) & m = M^* = M + 1 \end{cases} \quad (7)$$

2.4 Acceptance Probability

Evidently in passing from θ to θ^* the change $\Delta M \equiv M^* - M$ in the number of point masses in $\theta \in \Theta$ can only be zero or ± 1 ; in each of these three cases the log acceptance probability is given by the sum of the log changes in prior, likelihood, and proposal probabilities as:

$$\log \alpha(\theta, \theta^*) = \begin{cases} \left[\Delta \text{llh} + \log \frac{\nu(v_m^*, \sigma_m^*) v_m}{\nu(v_m, \sigma_m) v_m^*} \right]^- & \Delta M = 0 \\ \left[\Delta \text{llh} + \log \frac{p_1 \Phi[(\epsilon - v_m)/s_v]}{p_2 M \mu(v_m, \sigma_m) v_m} \right]^- & \Delta M = -1 \\ \left[\Delta \text{llh} + \log \frac{p_2 M^* \mu(v_m^*, \sigma_m^*) v_m^*}{p_1 \Phi[(\epsilon - v_m^*)/s_v]} \right]^- & \Delta M = +1 \end{cases} \quad (8)$$

where $x^- \equiv \min(x, 0)$ and in each case $\Delta \text{llh} = \frac{1}{2} \sum_{i \in I} \tau_i (\log^2 \frac{g_i^*}{Y_i} - \log^2 \frac{g_i}{Y_i})$, again with $g_i \equiv \sum_{m \leq M} v_m k_i(\sigma_m)$ and $g_i^* \equiv \sum_{m \leq M^*} v_m^* k_i(\sigma_m^*)$; in practice the latter is computed as

$$g_i^* = \begin{cases} g_i + v_m^* k_i(\sigma_m^*) - v_m k_i(\sigma_m) & \Delta M = 0 \\ g_i - v_m k_i(\sigma_m) & \Delta M = -1 \\ g_i + v_m^* k_i(\sigma_m^*) & \Delta M = +1. \end{cases} \quad (9)$$

Thus, after initializing θ^0 , the algorithm proceeds by generating proposals θ^* according to Eqn. (7), evaluating each g_i^* according to Eqn. (9) and the log acceptance probability from Eqn. (8), and updating θ^{t+1} accordingly.

3 Polymer Structure Examples

An important goal of polymer science is to identify and evaluate fundamental quantities for a given polymers from which other material functions can be calculated easily. Often, however, data are accessible only from some noisy linear transform of these fundamental quantities.

3.1 Unfolding distance distributions

Following Glatter *et al.* (1994) we consider the problem of making inference from light-scattering measurements about material characteristics of the poly(ethylene oxide)–poly(propylene oxide) triblock copolymer *P85*, whose optical properties in aqueous solution have attracted a great deal of research interest.

In small angle scattering a light beam is sent through a dilute polymer sample and the intensity of the scattered light is measured at a number of different angles. Features of the size, mass and structure of the molecules may be inferred from the intensity, as a function of angle.

In a very dilute solution (the zero-concentration limit) of *monodispersed* (uniform in size and structure) macromolecules it can be shown (Beeman *et al.*, 1957) that the mean value of the scattering intensity $I_m(h)$, upon averaging over the molecules' possible orientations and Brownian motion, is related to the incoming light intensity I_0 and to the probability distribution $\bar{P}(dr)$ of the distance between two uniformly distributed random points on

the molecule by

$$I(h) \equiv \frac{I_m(h)}{I_0} = (\rho V)^2 \int_0^\infty \frac{\sin(hr)}{hr} \bar{P}(dr),$$

where ρ is the electron density (number of electrons per unit volume of the molecule), V is the molecular volume, $h \equiv 4\pi \sin(\theta/2)/\lambda$ is the *scattering vector* length (in $1/\text{\AA}$), λ is the wavelength (in \AA) of the incoming light beam and θ is the scattering angle.

Inference is normally based on reconstructions of the *distance distribution* $P(dr) \equiv (\rho V)^2 \bar{P}(dr)/4\pi$ from which we can infer important information about the shape of the molecules. If, for instance, we assume a uniform electron density throughout the macromolecule of interest then we would expect a unimodal density of P for a spherical molecule (\bullet) and a bimodal density for a dimer with two spherical components ($\bullet\bullet$) (Glatter, 1979).

3.1.1 Distance distribution estimation

We turn now to the problem of estimating the distance distribution $P(dr)$ for three different concentrations of the triblock copolymer P85 based on a series of small-angle neutron scattering (SANS) measurements (Bergmann and Glatter, 2000). Neutron scattering is useful for exploring structure on a length scale from a few thousand Ångströms down to 5–10 Ångströms, a range that includes the molecular diameter (something less than 20nm, or 200Å).

Fig. (1) presents the measured scattering intensity as a function of the scattering vector at concentrations 1%, 0.33% and 0.05%, along with (log-scale) $\pm 2\sigma$ error bars, calculated from a delta-method approximation based on expert variance estimates (Bergmann and Glatter, 2000) and empirical mean estimates obtained from locally-smoothed scatter plots (Cleveland, 1979). The phase diagram in Glatter *et al.* (1994, Figure 15) suggests that the polymer should consist of spherical micelles at low concentrations, with a maximum diameter (from Glatter *et al.* (1994, Figure 4)) of about 15nm (150Å). We therefore choose a prior distribution for the normalized distance distribution $\bar{P}(dr)$ centered at the theoretical distance distribution for a spherical particle with uniform electron density given by $12x^2(2 - 3x + x^3) dr/D$ (Porod, 1948), where D is the diameter of the sphere and $x \equiv r/D$. We now use the relation $P(\mathbb{R}_+) = \lim_{h \rightarrow 0} \int \sin(hr)/hr P(dr) = I(0)/4\pi$ to get the appropriate scaling for the measure $P(dr)$.

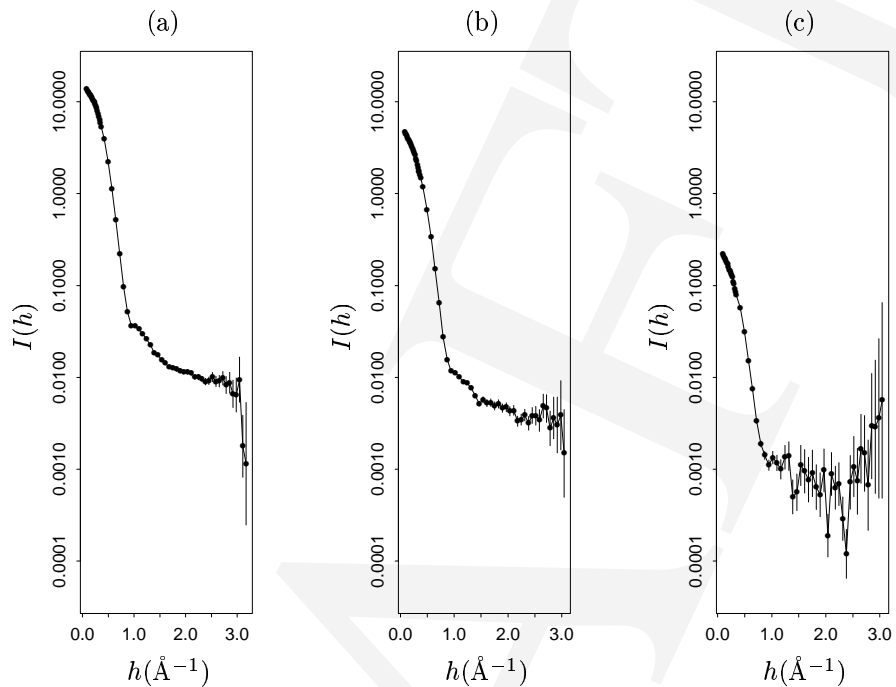


Figure 1: Measured relative scattering intensities $I(h)$ for (a) 1.0%, (b) 0.33%, and (c) 0.05% solutions of P85, with approximate 95% error-bars.

3.2 Unfolding relaxation spectra

Following Honerkamp and Weese (1989) we consider the problem of making inference about the material characteristics of a polybutadiene melt, a commercial rubber product, from the deformation measurements reported by Berger (1988).

Hooke's law relates force and deformation for a perfectly elastic compound as

$$\tau = G\gamma,$$

where τ denotes *stress* (force per unit area, or Pa), γ denotes *strain* (relative length change), and the constant G is the *elastic modulus* (also in Pa). Polymers, however, often show time dependence in their elastic response arising from *relaxation* of their macromolecule entanglements.

Boltzmann's approach (1876) was to model the elastic modulus G , now

termed a *relaxation modulus*, as time-dependent and express the stress τ at each time t as a bilinear functional of the relaxation modulus and the time derivative $\dot{\gamma}$ of the strain, or *strain rate*:

$$\tau(t) = \int_{-\infty}^t G(t-s)\dot{\gamma}(s) ds = \int_0^{\infty} G(s)\dot{\gamma}(t-s) ds. \quad (10)$$

Boltzmann took $G(t)$ to have a “fading memory”, *i.e.*, to be completely monotonically decreasing and so, by Bernstein’s Theorem (Feller, 1971, §XIII.4), representable in the form

$$G(t) = \int_0^{\infty} \exp(-tx)F(dx)$$

for some nonnegative (but not necessarily finite) measure F on \mathbb{R}_+ . It is custom in *rheology* (the study of non-Newtonian fluid mechanics) to parametrize by $\lambda = x^{-1}$ and write this in the form

$$G(t) = \int_0^{\infty} \exp(-t/\lambda) H(d\lambda)/\lambda.$$

The *relaxation spectrum* $H(d\lambda)$ can be interpreted as the contribution to the relaxation modulus of those components of the polymer with relaxation times in the volume element $d\lambda$. One can learn from $H(d\lambda)$ about the composition and complexity of the polymer. The Boltzmann model for linear viscoelasticity can also be viewed as a series of Maxwell elements (springs and dashpots) in series (Anderssen, 1999), offering a different route to inference.

The relaxation modulus is often estimated experimentally in oscillatory shear experiments by applying periodic strains of the form

$$\gamma(t) = \gamma_0 \sin(\omega_i t),$$

with fixed amplitude γ_0 at various frequencies ω_i . The Laplace transform equation Eqn. (10) can be evaluated explicitly in the form

$$\tau(t) = \gamma_0 [G'(\omega_i) \sin(\omega_i t) + G''(\omega_i) \cos(\omega_i t)]$$

where

$$G'(\omega) \equiv \int_0^{\infty} \frac{\omega^2 \lambda^2}{1 + \omega^2 \lambda^2} H(d\lambda)/\lambda$$

and

ω_i	$G'(\omega_i)$	$G''(\omega_i)$	ω_i	$G'(\omega_i)$	$G''(\omega_i)$
2.493×10^0	2052	34526	7.680×10^1	432105	359952
3.670×10^0	4156	50445	1.144×10^2	534678	343388
5.373×10^0	8847	73294	1.654×10^2	619214	327629
7.864×10^0	18834	105329	2.433×10^2	701325	307419
1.144×10^1	37737	149699	3.539×10^2	772708	290069
1.695×10^1	74730	206936	5.238×10^2	841878	278292
2.451×10^1	136257	266220	7.529×10^2	897344	264055
3.608×10^1	223611	313420	1.114×10^3	956262	249131
5.218×10^1	324937	345321			

Table 1: Experimental measurements of storage modulus $G'(\omega_i)$ and loss modulus $G''(\omega_i)$ (both in Pa) at various frequencies ω_i (in s^{-1}) for a polybutadiene melt at 23°C.

$$G''(\omega) \equiv \int_0^\infty \frac{\omega\lambda}{1 + \omega^2\lambda^2} H(d\lambda)/\lambda$$

denote the *storage modulus* and the *loss modulus*, respectively. The in-phase storage modulus represents elastic behavior and therefore energy conservation, while the out-of-phase loss modulus represents viscous behavior and therefore energy dissipation.

3.2.1 Spectral Estimation

We now turn to the problem of estimating the relaxation spectral measure $H(d\lambda)$ for a polybutadiene melt based on the measurements of the storage and loss moduli at 23°C given in Table 1, taken from (Berger, 1988).

For a Bayesian solution we must first specify a joint prior distribution for all uncertain aspects of the problem—in this example that includes the measurement-error distribution $Y_i \sim \psi_i(y | g_i)$ and the relaxation spectral measure $H(d\lambda)$.

Rheological measurements may be expected to be less precise with larger measurements. Following Honerkamp and Weese (1989), we take the Y_i to be lognormally distributed with median g_i and precision $\tau = 25$ (diagnostic checks based on residual errors from kernel estimates support that value).

We model the uncertain spectral measure $H(d\lambda) \sim \text{Ga}(\alpha(d\lambda), \beta(\lambda))$ as a gamma random field with Lévy measure $\nu(du d\lambda) = e^{-\beta(\lambda)\nu} u^{-1} du \alpha(d\lambda)$ on \mathbb{R}_+^2 , with constant inverse scale $\beta(\lambda) \equiv 0.14$ and shape measure $\alpha(d\lambda) =$

REFERENCES

$\alpha d\lambda/\lambda$ on the interval (e^{-7}, e^{-1}) . This measure is uniform on a logarithmic scale on the approximate interval $1/\omega_{\max} < \lambda < 1/\omega_{\min}$ that Davies and Anderssen (1997) suggest is the widest range on which we can hope to learn from the data about $H(d\lambda)$. The parameters were chosen to ensure that the zero-shear viscosity $\int_0^\infty G(t) dt = H(\mathbb{R}_+) \sim \text{Ga}(6\alpha, \beta)$ would have mean $6\alpha/\beta \approx 12800$ and variance $6\alpha/\beta^2 \approx 90000$. See (Anderssen and Hansen, 2000) for more specific rheological treatment and discussion of the sampling localization theorem in the context of relaxation spectral analysis.

4 Solution

5 Discussion

References

- Abramovich, F., Bailey, T. C. and Sapatinas, T. (2000) Wavelet analysis and its statistical applications. *The Statistician*, **49**, 1–29.
- Anderssen, R. S. (1999) The pragmatics of solving industrial (real-world) inverse problems with exemplification based on the molecular weight distribution problem. *Inverse Problems*, **15**, R1–R40.
- Anderssen, R. S. and Hansen, M. B. (2000) Bayesian relaxation spectrum recovery. In preparation.
- Beeman, W. W., Kaesberg, P., Anderegg, J. W. and Webb, M. B. (1957) Size of particles and lattice defects. In *Handbuch der Physik* (ed. S. Flügge), vol. XXXII. Berlin: Springer-Verlag, pp. 321–442.
- Berger, L. (1988) *Untersuchung zum rheologischen Verhalten von Polybutadienen mit bimodaler Molmassenverteilung*. Ph.D. dissertation, Eidgenössische Technische Hochschule (ETH), Zürich.
- Bergmann, A. and Glatter, O. (2000) Personal communication.
- Best, N. G., Ickstadt, K. and Wolpert, R. L. (2000) Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association*, **95**, 1076–1088.
- Boltzmann, L. (1876) Zur Theorie der elastischen Nachwirkung. *Annals of Physical Chemistry*, **7**, 624–654.

REFERENCES

- Cleveland, W. (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74**, 829–836.
- Davies, A. and Anderssen, R. S. (1997) Sampling localization in determining the relaxation spectrum. *Journal of Non-Newtonian Fluid Mechanics*, **73**, 163–179.
- Feller, W. (1971) *An Introduction to Probability Theory and its Application*, vol. 2. 2nd edn. New York, NY, USA: John Wiley & Sons.
- Gelfand, A. E. and Smith, A. F. M. (1990) Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398–409.
- Glatter, O. (1979) The interpretation of real-space information from small-angle scattering experiments. *Journal of Applied Crystallography*, **12**, 166–175.
- Glatter, O., Scherf, G., Schillén, K. and Brown, W. (1994) Characterization of a poly(thylene oxide)-poly(propylene oxide) triblock copolymer (EO₂₇-PO₃₉-EO₂₇) in aqueous solution. *Macromolecules*, **27**, 6046–6054.
- Good, I. J. and Gaskins, R. A. (1971) Non-parametric roughness penalties for probability densities. *Biometrika*, **58**, 255–277.
- Grenander, U. (1981) *Abstract Inference*. New York, NY, USA: John Wiley & Sons.
- Hansen, M. B. and Lauritzen, S. L. (1999) Discussion of “Bayesian non-parametric inference for random distributions and related functions” by Walker, Laud, Smith and Damien. *Journal of the Royal Statistical Society, Series B (Methodological)*, **61**, 517.
- Hansen, P. C. (1998) *Rank-Deficient and Discrete Ill-Posed Problems*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).
- Hoerl, A. E. and Kennard, R. W. (1970a) Ridge regression : Applications to nonorthogonal problems. *Technometrics*, **12**, 69–82.
- (1970b) Ridge regression : Biased estimation for nonorthogonal problems. *Technometrics*, **12**, 55–67.
- Honerkamp, J. and Weese, J. (1989) Determination of the relaxation spectrum by a regularization method. *Macromolecules*, **22**, 4372–4377.

REFERENCES

- Jacod, J. and Shiryaev, A. N. (1987) *Limit Theorems for Stochastic Processes*, vol. 288 of *Grundlehren der mathematischen Wissenschaften*. Berlin: Springer-Verlag.
- Kimeldorf, G. S. and Wahba, G. (1971) Some results on Tchebycheffian spline functions. *Journal of Mathematical Analysis and Applications*, **33**, 82–95.
- Kirsch, A. (1996) *An Introduction to the Mathematical Theory of Inverse Problems*. New York, NY, USA: Springer-Verlag.
- O’Sullivan, F. (1986) A statistical perspective on ill-posed inverse problems (with discussion). *Statistical Science*, **1**, 502–527.
- Porod, G. (1948) Die Abhängigkeit der Röntgen-Kleinwinkelstreuung von Form und Größe der kolloiden Teilchen in verdünnten Systemen. *Acta Physica Austriaca*, **2**, 255–292.
- Rogers, L. and Williams, D. (1987) *Diffusions, Markov Processes, and Martingales*, vol. 2. New York, NY, USA: John Wiley & Sons.
- Stark, P. B. (2000) Inverse problems as statistics. In *Surveys on Solution Methods for Inverse Problems* (eds. D. Colton, H. Engl, A. Louis, J. McLaughlin and W. Rundell). New York, NY, USA: Springer-Verlag.
- Tierney, L. (1994) Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, **22**, 1701–1762.
- Tikhonov, A. N. (1963) Solution of incorrectly formulated problems and the regularization method. *Soviet Doklady*, **4**, 1035–1038.
- Tweedie, R. and Mengersen, K. L. (1996) Rates of convergence in the Hastings-Metropolis algorithm. *Annals of Statistics*, **24**, 101–121.
- Wolpert, R. L. and Ickstadt, K. (1998a) Poisson/gamma random field models for spatial statistics. *Biometrika*, **85**, 251–267.
- (1998b) Simulation of Lévy random fields. In *Practical Nonparametric and Semiparametric Bayesian Statistics* (eds. D. Dey, P. Müller and D. Sinha), vol. 133 of *Lecture Notes in Statistics*. New York, NY, USA: Springer-Verlag.