Aggregate Data and the Prohorov Metric Framework: Efficient Gradient Computation

H.T. Banks and Jared Catenacci
Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695-8212 USA

November 30, 2015

Abstract

We discuss efficient methods for computing gradients in inverse problems for estimation of distributions for individual parameters in models where only aggregate or population level data is available. The ideas are illustrated with two examples arising in applications.

Key words: inverse problems, least squares estimation, estimation of probability distributions, splines, Prohorov metric

Mathematics Subject Classification: 34A55,45Q05,65Y20

1 Introduction

For years even simple population models based on individual models (see, e.g., the Hare-Lynx models [19, p. 30] and the bacterial growth and diffusion models [19, p. 33], [20], [21, p. 139]) have been based on aggregate population level data for parameter estimation and validation. However, with increased interest in uncertainty quantification and recognition that statistical models for the data collection procedures drive uncertainty statements about the parameters in the underlying mathematical models, the interest in determining correct statistical models as part of parameter estimation or inverse problems has grown. Moreover, it is now recognized that aggregate data is widely (and frequently incorrectly) employed to quantify uncertainty in individual models. This occurs in a ubiquitous range of applied problems including food chemistry efforts [14, 18, 22], tracking of labeled substances in proliferating cell populations (e.g., Propagons or prion seeds in amyloid growth in yeast [12, 13, 16, 17, 23]), as well as structured population models in marine population studies such as those for mosquitofish [2] and shrimp [6]. In such individual models, one has a mathematical model which describes the behavior of one "individual" which is characterized by a single parameter set which must be estimated using population level or aggregate data.

In a second class of problems (the aggregate model case), the dynamic mathematical models explicitly depend upon a distribution that must be estimated using aggregate data. This is the case in electromagnetic interrogation problems with a distribution of polarization

permittivity and relaxation time parameters for molecules [5,7,8,11], in HIV cellular models [3, 4], and in wave propagation in viscoelastic materials [9–11]. Again in these examples, only aggregate data is available to estimate the imbedded probability distributions.

One method for such non-parametric estimation problems of a probability measure is through the Prohorov Metric Framework (PMF) [1,11] developed specifically to treat aggregate data problems (for a summary see [11, Chapter 5]). The PMF provides a theoretical and computational framework in which to estimate an unknown probability measure for which the space $\mathcal{P}(\Omega)$ of probability measures over a compact set Ω is approximated by a finite dimensional space $\mathcal{P}^N(\Omega)$ of dimension N. There are many choices for the approximating space $\mathcal{P}^N(\Omega)$; two popular choices involve using a basis of Dirac measures (zero order splines) or piecewise linear splines to approximate the distributions. In this presentation, our goal is to show how the gradient of a least squares objective function can be found in a efficient manner for inverse problems involving the estimation of a probability measure using the PMF.

2 Problem framework

We assume to have a mathematical model for a dynamical system which is dependent upon a probability measure G as well as Euclidean parameters $\mathbf{q} \in \mathcal{Q}$. We assume that the solution to this system can be obtained either analytically or numerically and denote the solution as $u(x, t; G, \mathbf{q})$. Furthermore we assume that we have a set of observations

$$y_j = u(x_j, t_j; G_0, \mathbf{q}_0) + \epsilon_j, \quad j = 1, ..., n,$$

where G_0 and $\mathbf{q}_0$ are the true or nominal probability measure and parameters, respectively, and ϵ_j is a realization of the measurement error in the observation process.

Given a set of observations y_j at the points (x_j, t_j), $j = 1, ..., n$, we would like to estimate the unknown parameters $\mathbf{q} \in \mathcal{Q} \subset \mathbb{R}^\kappa$ and the unknown distribution $G(\theta) \in \mathcal{P}(\Omega)$, where $\mathcal{P}(\Omega)$ is the set of admissible probability measures on $\Omega \subset \mathbb{R}$. Thus, we would like to solve

$$(G, \mathbf{q}) = \underset{(G,\mathbf{q})\in(\mathcal{P}(\Omega)\times\mathcal{Q})}{\arg\min} J(G, \mathbf{q}), \tag{2.1}$$

where

$$J(G, \mathbf{q}) = \sum_{j=1}^{n} \left(y_j - u(t_j, x_j; G, \mathbf{q})\right)^2. \tag{2.2}$$

We note that (2.1) is an infinite-dimensional optimization problem. Thus, we need to approximate the infinite dimensional space $\mathcal{P}(\Omega)$ with a finite dimensional space $\mathcal{P}^N(\Omega)$ in order to have a computationally tractable finite-dimensional optimization problem

$$(\widehat{G}, \widehat{\mathbf{q}}) = \underset{(G,\mathbf{q})\in(\mathcal{P}^N(\Omega)\times\mathcal{Q})}{\arg\min} J(G, \mathbf{q}). \tag{2.3}$$

We will consider two finite-dimensional spaces, $\mathcal{P}_D^N(\Omega)$ and $\mathcal{P}_S^N(\Omega)$, to approximate $\mathcal{P}(\Omega)$. The space $\mathcal{P}_D^N$ involves the use of Dirac measures, and the space $\mathcal{P}_S^N$ involves the use of piecewise linear splines. We define these two spaces as

$$\mathcal{P}_D^N(\Omega) = \left\{ G \in \mathcal{P}(\Omega) \;\middle|\; G = \sum_{m=1}^{N} \alpha_m \Delta_{z_m}, \text{where } \alpha_m \geq 0 \text{ and } \sum_{m=1}^{N} \alpha_m = 1 \right\}, \tag{2.4}$$

and

$$\mathcal{P}_S^N(\Omega) = \left\{ G \in \mathcal{P}(\Omega) \;\middle|\; G' = \sum_{m=1}^{N} \alpha_m l_m(\theta), \text{where } \alpha_m \geq 0 \text{ and } \sum_{m=1}^{N} \alpha_m \int_{\Omega_m} l_m(\xi) d\xi = 1 \right\}, \tag{2.5}$$

where Δ_{z_m} is a Dirac measure with atom at z_m, and l_m is the mth linear spline element with support Ω_m. With both of these spaces we have reduced the infinite-dimensional problem to a finite-dimensional problem in which we only need to estimate the parameters $\mathbf{q}$ and the weights $\boldsymbol{\alpha} = \{\alpha_m\}_{m=1}^N$. Hence, when using the Delta approximation method we have the minimization problem

$$(\widehat{\boldsymbol{\alpha}}, \widehat{\mathbf{q}}) = \underset{(\boldsymbol{\alpha},\mathbf{q}) \in (\mathbb{R}_D^N \times \mathcal{Q})}{\arg\min} \; J(\boldsymbol{\alpha}, \mathbf{q}), \tag{2.6}$$

where

$$\mathbb{R}_D^N = \left\{ \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T \;\middle|\; \alpha_m \geq 0, \text{ and } \sum_{m=1}^{N} \alpha_m = 1 \right\}.$$

Using the spline method we have the minimization problem

$$(\widehat{\boldsymbol{\alpha}}, \widehat{\mathbf{q}}) = \underset{(\boldsymbol{\alpha},\mathbf{q}) \in (\mathbb{R}_S^N \times \mathcal{Q})}{\arg\min} \; J(\boldsymbol{\alpha}, \mathbf{q}), \tag{2.7}$$

where

$$\mathbb{R}_S^N = \left\{ \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T \;\middle|\; \alpha_m \geq 0, \text{ and } \sum_{m=1}^{N} \alpha_m \int_{\Omega_m} l_m(\xi) d\xi = 1 \right\}.$$

In solving (2.6) or (2.7) one may wish to use a gradient based nonlinear optimization method, particularly if there are a large number of parameters to be estimated, which is the case in problems where N needs to be taken large in order to obtain a reasonable approximation. This requires the the computation of

$$\nabla J(\boldsymbol{\alpha}, \mathbf{q}) = \left[\frac{\partial J}{\partial \alpha_1}, ..., \frac{\partial J}{\partial \alpha_N}, \frac{\partial J}{\partial q_1}, ..., \frac{\partial J}{\partial q_\kappa} \right]^T. \tag{2.8}$$

The most common method for the approximating the above partial derivatives is to use a finite difference. If a forward difference is used, then this results in the approximations

$$\begin{aligned} \frac{\partial J}{\partial q_k} &\approx \frac{J(\boldsymbol{\alpha}, \mathbf{q} + h_k \mathbf{e}_k^\kappa) - J(\boldsymbol{\alpha}, \mathbf{q})}{h_k} \\ \frac{\partial J}{\partial \alpha_k} &\approx \frac{J(\boldsymbol{\alpha} + h_k \mathbf{e}_k^N, \mathbf{q}) - J(\boldsymbol{\alpha}, \mathbf{q})}{h_k}, \end{aligned} \tag{2.9}$$

where $\mathbf{e}_k^m$ is the k-th standard unit basis vector of length $m = \kappa$ or $m = N$. Since the evaluation of $J(\boldsymbol{\alpha}, \mathbf{q})$ is already required, we must evaluate only the first term in each of the above numerators. This requires a total of $1 + \kappa + N$ evaluations of the cost functional J.

2.1 Individual models

In the situation of an individual model, we have a model that depends wholly on a single parameter set. We denote this individual model by $v(x,t;\theta,\mathbf{q})$, where θ is a parameter upon which the individual model is dependent, but where the values are expected to vary across the population. In contrast, the parameters $\mathbf{q}$ are assumed to be population level parameters, i.e., $\mathbf{q}$ is not expected to vary significantly across individuals. In this situation, we can formulate a population level model as

$$u(x,t;G,\mathbf{q}) = \int_\Omega v(x,t;\xi,\mathbf{q})dG(\xi). \tag{2.10}$$

If the Dirac mass approximation scheme is used, we obtain the model

$$u(x,t;\boldsymbol{\alpha},\mathbf{q}) = \sum_{m=1}^{N} \alpha_m v(x,t;z_m,\mathbf{q}), \tag{2.11}$$

and if the spline approximation scheme is used, we obtain

$$u(x,t;\boldsymbol{\alpha},\mathbf{q}) = \sum_{m=1}^{N} \alpha_m \int_{\Omega_m} v(x,t;\xi,\mathbf{q}) l_m(\xi)d\xi. \tag{2.12}$$

Using either approximation scheme, if one computes the gradient of the objective function according to (2.8)−(2.9) then $v(t,x;\cdot,\mathbf{q})$ must be evaluated at least $N^2+(\kappa+1)N$ times. If a p point quadrature is used to numerical evaluate (2.12), then $v(t,x;\cdot,\mathbf{q})$ will be evaluated $p(N^2+(\kappa+1)N)$ times.

Observe that

$$\begin{aligned}
\frac{\partial J}{\partial q_k} &= -2\sum_{j=1}^{n} \left(y_j - u(x_j,t_j;\boldsymbol{\alpha},\mathbf{q})\right) \frac{\partial u(x_j,t_j;\boldsymbol{\alpha},\mathbf{q})}{\partial q_k} \\
\frac{\partial J}{\partial \alpha_k} &= -2\sum_{j=1}^{n} \left(y_j - u(x_j,t_j;\boldsymbol{\alpha},\mathbf{q})\right) \frac{\partial u(x_j,t_j;\boldsymbol{\alpha},\mathbf{q})}{\partial \alpha_k}.
\end{aligned} \tag{2.13}$$

The derivatives of $u(x,t;\boldsymbol{\alpha},\mathbf{q})$ can be obtained from the sensitivity equations [11, 15]. However, for complex models the sensitivity equations can be difficult to derive and in most cases the term $\frac{\partial}{\partial q_k}u(x,t;\boldsymbol{\alpha},\mathbf{q})$ will still need to be approximated by a finite difference. If a forward difference is used, then we have

$$\frac{\partial u(x,t;\boldsymbol{\alpha},\mathbf{q})}{\partial q_k} \approx \frac{u(t,x;\boldsymbol{\alpha},\mathbf{q}+h_k\mathbf{e}_k^\kappa) - u(t,x;\boldsymbol{\alpha},\mathbf{q})}{h_k}. \tag{2.14}$$

Note that we are already required to compute the term $u(x,t;\boldsymbol{\alpha},\mathbf{q})$ for the cost function $J(\boldsymbol{\alpha},\mathbf{q})$, so an efficiently implemented optimization scheme will take advantage of this information rather than computing the term multiple times.

Since the coefficients α_k appear linearly in the population level model we can compute the derivatives of $u(x,t;\boldsymbol{\alpha},\mathbf{q})$ exactly in (2.13). Thus,

$$\frac{\partial u(x,t;\boldsymbol{\alpha},\mathbf{q})}{\partial \alpha_k} = v(t,x;z_k,\mathbf{q}) \tag{2.15}$$

if using the Dirac approximation method, and

$$\frac{\partial u(x,t;\boldsymbol{\alpha},\mathbf{q})}{\partial \alpha_k} = \int_{\Omega_m} v(t,x;\xi,\mathbf{q}) l_m(\xi) d\xi \tag{2.16}$$

if using the spline approximation method. In either case, these values are already required to be computed to obtain the cost function $J(\boldsymbol{\alpha},\mathbf{q})$. Taking advantage of these precomputed values greatly reduces the computational expense in approximating the gradient of $J(\boldsymbol{\alpha},\mathbf{q})$. If we compute the gradient according to (2.13)−(2.16), then only $(\kappa+1)N$ evaluations of $v(x,t;\cdot,\mathbf{q})$ are required ($p(\kappa+1)N$ if using a p point quadrature rule to approximate (2.16)). Additionally, since we are computing the derivative exactly in this case, there is zero truncation error in the derivative computation of (2.15) or (2.16) and the only source of truncation error is from (2.14).

2.2 Aggregate models

In the case of an aggregate model, the model is explicitly dependent upon a probability measure G. Thus, our model $u(x,t;G,\mathbf{q})$ does not have the form of (2.10). However, $u(x,t;G,\mathbf{q})$ will include at least one term of the form

$$\int_{\Omega} f(x,t;\xi;\mathbf{q}) dG(\xi), \tag{2.17}$$

and again we will be able to exploit the linearity of the approximation terms to reduce computational times. Using the PMF approximation, the model reduces to $u(x,t;\boldsymbol{\alpha},\mathbf{q})$, which now depends on a term of the form

$$\sum_{m=1}^{N} \alpha_m f(x,t;z_m,\mathbf{q}), \tag{2.18}$$

if using the Dirac approximation scheme, and if the spline approximation scheme is used we obtain

$$\sum_{m=1}^{N} \alpha_m \int_{\Omega_m} f(x,t;\xi,\mathbf{q}) l_m(\xi) d\xi. \tag{2.19}$$

Hence, just as in the case of using an individual model, computing the gradient of $J(\boldsymbol{\alpha},\mathbf{q})$ according to (2.13) requires $(\kappa+1)N$ evaluations of $f(x,t;\cdot,\mathbf{q})$, whereas computing the gradient according to (2.9) requires $N^2+(\kappa+1)N$ evaluations of $f(x,t;\cdot,\mathbf{q})$. Again, as before if using a p point quadrature rule to approximate (2.19) then the number of evaluations is multiplied by p.

3 Example: Sinko-Streifer model

Consider the case where an individual Sinko-Streifer model can be used to model the size-structured population. The model provided here is adapted from [2] where the goal was to estimate individual growth rates for a mosquito fish population, but only aggregate data was available. A similar problem arose in the population modeling examples for shrimp [6].

We assume that the growth rate varies according to each individual, but the death (or removal rate) is constant across the population. For simplicity we assume that there is no recruitment into the system. The model is given by

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(gv) &= -\mu v, \quad x_0 < x < x_1, \quad t > 0 \\ v(x,0) &= \Phi(x) \\ g(x_0,t)v(x_0,t) &= 0 \\ g(x_1,t) &= 0, \end{aligned} \tag{3.1}$$

where $v(t,x)$ represents the population density, and t and x denote time and size, respectively, $g(x)$ is the size dependent growth rate term, and μ is the removal rate. From [2], the admissible growth rates of an individual fish we will consider are of the form

$$g(x;\theta,\gamma) = \begin{cases} \theta(\gamma - x), & x_0 \le x \le \gamma, \\ 0, & \text{otherwise}, \end{cases}$$

where θ and γ denote the intrinsic growth rate and maximum size, respectively. For simplicity we assume that $\gamma = 1$. The collection of admissible growth rates is given by

$$\mathcal{G} = \{g(x;\theta) \mid \theta \in \Omega\},$$

where Ω is a compact set.

The solution to (3.1) can be found using the method of characteristics. Data was simulated according to

$$y_{ij} = \int_\Omega v(x_i, t_j; \theta, \mu) dG_0(\theta) + \epsilon_{ij}, \quad i = 1, ..., n_x, \ j = 1, ..., n_t \tag{3.2}$$

where $v(x,t;\theta,\mu)$ is the solution to (3.1), and ϵ_{ij} are realizations of a normally distributed random variable with 0 mean and variance 0.01. The distribution G_0 was taken to be a normal distribution with mean 4.5 and variance 0.25, and the death rate was chosen to be $\mu_0 = 1.0$. The initial condition was taken as

$$\Phi(x) = \begin{cases} \sin^2 10\pi x & 0 \le x \le 0.1 \\ 0 & x > 0.1 \end{cases} \tag{3.3}$$

and is assumed to be known. Hence, in this example we need to estimate the probability measure G and the removal rate μ. G is estimated using the Delta approximation scheme, where the nodes were placed in a uniform grid over the interval $[3,6]$.

The Matlab routine `fmincon` was used to preform the resulting optimization problems for 100 independent simulated data sets. We considered 2 methods for computing the gradients. For method 1, we computed the gradient using a forward difference of the objective function $J(\boldsymbol{\alpha}, \mu)$. This method is equivalent to the default method for computing the gradient if the user does not supply the gradient to `fmincon`. For method 2 we computed the gradient according to (2.13)−(2.15). In Figure 1 we depict the average cpu time required to complete the optimization using both methods for 100 independent data sets as N, the number of elements in the approximation scheme, increases.

Using method 1 requires $N^2 + N(\kappa + 1)$ evaluations of $v(x,t;\theta,\mu)$, whereas method 2 only requires $(\kappa + 1)N$ evaluations. This agrees with Figure 1 where method 1 exhibits an approximately quadratic increase in time to preform the optimization as N increases, and method 2 has an approximately linear increase in time.

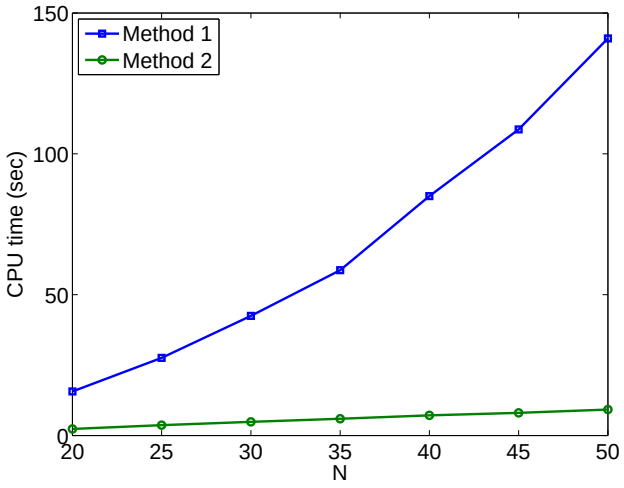

Figure 1: The average cpu time (for 100 independent data sets) required to complete the optimization as N increases using both methods.

4 Example: Reflectance spectroscopy model

Here we describe an example from [5] where the model is an aggregate model. In this project, the goal is to develop a noninvasive technique to characterize the degradation of a complex nonmagnetic dielectric material by assessing the small physical and chemical changes in the material using reflectance spectroscopy. This involves determining the components of the permittivity of the dielectric medium using the measured spectral responses. The distributed relative permittivity of the dielectric medium is described by

$$\widehat{\varepsilon}_r(k;G,\mathbf{q}) = \varepsilon_\infty - \int_\Omega \frac{k_p^2}{k^2 - ik/\tau - k_0^2} dG(k_0). \tag{4.1}$$

In the above equation, ε_∞ denotes the relative permittivity of the dielectric medium at infinite frequency, k is the wavenumber ($k = \omega/(2\pi c)$, where ω is the angular frequency and c is the speed of light), k_0 represents the resonance wavenumbers, and τ denotes the relaxation time. The composite parameter k_p is given by $k_p = k_0\sqrt{\varepsilon_s - \varepsilon_\infty}$ with ε_s being the relative permittivity of the medium at zero frequency, $i = \sqrt{-1}$ is the imaginary unit, and $\theta = k_0 \in \Omega \subset \mathbb{R}$. If we assume that a monochromatic uniform wave is incident at an angle of $\phi = 45°$ on a plane interface between free space and a nonmagnetic dielectric medium with the electric field composed of the parallel and perpendicular polarizations in equal weights, then the reflection coefficient is given by

$$R(k;G,\mathbf{q}) = \frac{1}{2}\left(|r_\perp(k;G,\mathbf{q})|^2 + |r_\parallel(k;G,\mathbf{q})|^2\right), \tag{4.2}$$

where

$$r_\perp(k;G,\mathbf{q}) = \frac{\cos\phi - \sqrt{\widehat{\varepsilon}_r(k;G,\mathbf{q}) - \sin\phi}}{\cos\phi + \sqrt{\widehat{\varepsilon}_r(k;G,\mathbf{q}) - \sin\phi}}, \tag{4.3}$$

and

$$r_\parallel(k;G,\mathbf{q}) = \frac{\sqrt{1 - \sin^2\phi/\widehat{\varepsilon}_r(k;G,\mathbf{q})} - \sqrt{\widehat{\varepsilon}_r(k;G,\mathbf{q})}\cos\phi}{\sqrt{1 - \sin^2\phi/\widehat{\varepsilon}_r(k;G,\mathbf{q})} + \sqrt{\widehat{\varepsilon}_r(k;G,\mathbf{q})}\cos\phi}. \tag{4.4}$$

In this application, the reflectance $R(k; G, \mathbf{q})$ is measured at various wave numbers k in order to determine the distribution $G(\theta) = G(k_0)$ of resonance wave numbers as well as the parameters $\mathbf{q} = [\varepsilon_s, \varepsilon_\infty, \tau]^T$. Data sets which were collected using a Bruker Vertex 80V FTIR spectrometer have been provided by researchers at the Air Force Research Lab at Wright-Patterson Air Force Base. For a full description of the model, data collection, and subsequent inverse problems, see [5].

In this case the spline approximation scheme was used to estimate the probability measure $G(\theta)$, thus the permittivity model can be written as

$$\widehat{\varepsilon}_r(k; G, \mathbf{q}) = \varepsilon_\infty - \sum_{m=1}^{N} \alpha_m \int_{\Omega_m} \frac{\theta^2(\varepsilon_s - \varepsilon_\infty)}{k^2 - ik/\tau - \theta^2} l_m(\theta) d\theta. \tag{4.5}$$

Again we use two methods to compute the gradient of $J(\alpha, \mathbf{q})$, where method 1 computes the gradient according to (2.9) and method 2 employs (2.13) where the integral terms in (4.5) were computed only once. In Figure 2 we show the average over 100 trials of the cpu time required to preform the first 10 iterations of the optimization problems as N is increased. As expected, we see that method 1 increases quadratically and method 2 linearly.

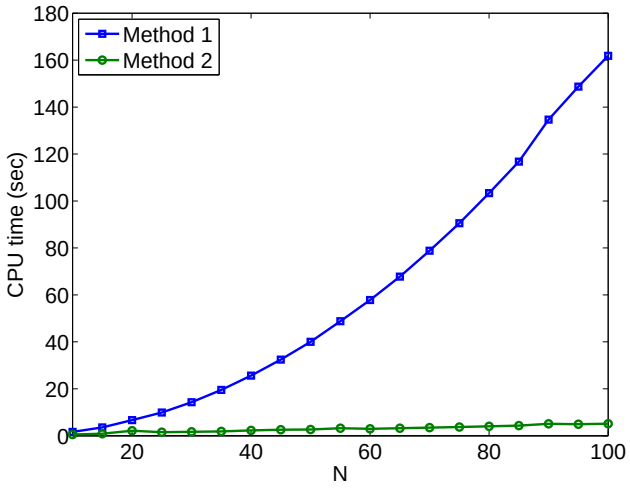

Figure 2: The average cpu time (for 100 independent data sets) required to compute the first 10 iterations of the optimization problems as N increases using both methods.

5 Conclusions

In this paper we consider the case of non-parametric estimation of a probability measure under the Prohorov Metric Framework in a least squares problem. It is demonstrated that the gradient computation can be reduced by exploiting the linearity of the coefficients to be estimated which appear in the approximation schemes under the PMF.

For individual models the number of forward solves of the underlying model v is reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, where N is the number of elements in the approximation. Due to the use of the exact partial derivatives in computing the gradient, there is no truncation error present from a finite differencing of the objective function. An example using a Sinko-Streifer

model with aggregate data is discussed and the expected linear increase in cpu time as N increases was observed.

For aggregate models the reduction of computational expense in computing the gradient of the objective function is not as straight forward. This is due directly to the fact that the model depends *explicitly* on the probability measure for aggregate models. However, we still can reduce the number of evaluations of the kernel function f in (2.17) from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$. The practical degree to which any speed up can be obtained in the inverse problem depends directly on the complexity of the kernel function f. If f is relatively cheap to evaluate, then the speed up may be negligible, even though we have reduced the number of evaluations. However, if f is costly to evaluate, then the speed up may be significant. We demonstrated this in an example arising in an application using reflectance spectroscopy, and the cpu time was observed to have the expected linear behavior.

Acknowledgments

This research was supported in part by the Air Force Office of Scientific Research under grant number AFOSR FA9550-12-1-0188, and in part by the US Department of Education Graduate Assistance in Areas of National Need (GAANN) under grant number P200A120047.

References

[1] H.T. Banks, *A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*, Chapman and Hall/CRC Press, Boca Raton, FL, 2012.

[2] H.T. Banks, L.W. Botsford, F. Kappel, and C. Wang, Modeling and estimation in size structured population models, LCDS-CCS Report 87-13, Brown University; *Proceedings 2nd Course on Mathematical Ecology*, (Triests, December 8-12, 1986) World Press (1988), Singapore, 521–541.

[3] H.T. Banks and D.M. Bortz, Inverse problems for a class of measure dependent dynamical systems, *J. Inverse and Ill-posed Problems*, **13** (2005), 103–121.

[4] H.T. Banks, D.M. Bortz and S.E. Holte, Incorporation of variability into the mathematical modeling of viral delays in HIV infection dynamics, *Mathematical Biosciences*, **83** (2003), 63–91.

[5] H.T. Banks, J. Catenacci and A. Criner, Quantifying the degradation in thermally treated ceramic matrix composite, CRSC-TR15-10, September, 2015; *International J. Applied Electromagnetics* (submitted).

[6] H.T. Banks, J.L. Davis, S.L. Ernsterberger, S. Hu, E. Artimovich, and A.K. Dhar, Experimental design and estimation of growth rate distributions in size-structured shrimp populations, CRSC-TR08-20, North Carolina State University, November 2008; *Inverse Problems* **25** (2009), 095003 (28 pages).

[7] H.T. Banks and N.L. Gibson, Well-posedness in Maxwell systems with distributions of polarization relaxation parameters, CRSC-TR04-01, January, 2004; *Applied Math. Letters*, **18** (2005), 423–430.

[8] H.T. Banks and N.L. Gibson, Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters, CRSC-TR05-29, August 2005; *Quarterly of Applied Mathematics*, **64** (2006), 749–795.

[9] H.T. Banks, S. Hu, Z.R. Kenz, C. Kruse, S. Shaw, J.R. Whiteman, M.P. Brewin, S.E. Greenwald and M.J. Birch, Material parameter estimation and hypothesis testing on a 1D viscoelastic stenosis model: methodology, CRSC-TR12-09, April, 2012; *J. Inverse and Ill-posed Problems*, **21** (2013), 25–57.

[10] H.T. Banks and G.A. Pinter, A probabilistic multi scale approach to hysteresis in shear wave propagation in biotissue, CRSC-TR04-03, January, 2004; *SIAM J. Multiscale Modeling and Simulation*, **3** (2005), 395–412.

[11] H. T. Banks, S. Hu, and W. C. Thompson, *Modeling and Inverse Problems in the Presence of Uncertainty*, CRSC Press/ Taylor & Frances Publishing, Boca Raton, FL, 2014.

[12] Lee J. Byrne, Diana J. Cole, Brian S. Cox, Martin S. Ridout, Byron J. T. Morgan, and Mick F. Tuite, The number and transmission of [PSI$^+$] prion seeds (Propagons) in the yeast *Saccharomyces cerevisiae*, *PLoS ONE*, **4**(3) (2009), e4670.

[13] Brian Cox, Frederique Ness and Mick Tuite, Analysis of the generation and segretation of Propagons: Entities that propagate the [PSI$^+$] prion in yeast, *Genetics*, **165** (2003), 23–33.

[14] Christopher J. Doona, Florence E. Feeherry, and Edward W. Ross, A quasi-chemical model for the growth and death of microorganisms in foods by non-thermal and high-pressure processing, *Int. J. Food Microbiol.,* **100** (2005), 21–32.

[15] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelpha, PA, 1999.

[16] Peter Olofsson and Suzanne S. Sindi, A Crump-Mode-Jagers branching process model of prion loss in yeast, *J. Appl. Prob.,* **51** (2014), 453–465.

[17] K.J. Palmer, M.S. Ridout, and B.J.T. Morgan, Kinetic models of guanidine hydrochloride-induced curing of the yeast [PSI$^+$] prion, *J. Theoretical Biology,* **274** (2011), 1–11.

[18] E.W. Ross, I.A. Taub, C.J. Doona, F.E. Feeherry, and K. Kustin, The mathematical properties of the quasi-chemical model for microorganism growth-death kinetics in foods, *Int. J. Food Microbiol.,* **99** (2005), 157–171.

[19] S.I. Rubinow, *Introduction to Mathematical Biology*, John Wiley & Sons, New York, NY, 1975.

[20] Lee A. Segel, Ed., *Mathematical Models in Molecular and Cellular Biology*, Cambridge University Press, Cambridge, UK, 1980

[21] Lee A. Segel, *Modeling Dynamic Phenomena in Molecular and Cellular Biology*, Cambridge University Press, Cambridge, UK, 1984

[22] Vinicio Serment-Moreno, Gustavo Barbosa-Canovas, Jose Antonio Torres, and Jorge Welti-Chanes, High-pressure processing: Kinetic models for microbial and enzyme inactivation, *Food Eng. Rev.,* **6** (2014), 56–88.

[23] Motomasa Tanaka, Sean R. Collins, Brandon H. Toyama, and Jonathan S. Weissman, The physical basis of how prion conformations determine strain phenotypes, *Nature,* **442** (2006), 585–589.