Inverse Problems: Basic Ideas and Applications to Biology

References:

1. C. Groetsch, Inverse Problems in the Mathematical Sciences, Vieweg-Verlag, Braunschweig, Wiesbaden, 1993.

2. M. Hanke and O. Scherzer, Inverse Problems Light: Numerical Differentiation, Amer. Math. Monthly, Vol 108 (2001), 512-521.

3. A. Kirsch, An Introduction to Mathematical Theory of Inverse Problems, Springer-Verlag, New York, 1996.

4. J. Mazumdar, An Introduction to Mathematical Physiology and Biology, 2nd Ed., Cambridge University Press 1999.

OUTLINE OF TWO LECTURES:

1. Basic Ideas and Theory

   (a) Some Examples

   (b) A General Setting

   (c) The Regularization Idea

   (d) *Sample Calculations

2. Applications to Biology Models

   (a) A Pharamacokinetic Model

   (b) An Exponential Growth Model

   (c) A Size Dependent Population Model

*Using the MATLAB GUI Numderiv, as time permits.

LECTURE 1: BASIC IDEAS AND THEORY

Some Examples

What are we talking about? A *direct problem* is the sort of thing we traditionally think about in mathematics:

$$\text{Question} \longrightarrow \text{Answer}$$

An *inverse problem* looks like this:

$$\text{Question} \longleftarrow \text{Answer}$$

Example 1. The game played on TV show "Jeopardy": given the answer, say the question.

Actually, this schematic doesn't quite capture the real flavor of inverse problems. It should look more like

$$\text{Question} \longleftarrow \text{(Approximate) Answer}$$

Example 2. The game played on TV show "Wheel of Fortune".

Direct problem: if someone says the message, you spell it.

Inverse problem: if someone spells the message (approximately), you say it.

Here are some *mathematical* examples:

Example 3. Polynomial roots.

Direct problem: given a polynomial, find all its roots.

Inverse problem: given a list of roots, find a polynomial with exactly those roots.

Note. This is one of those (rare) situations in which the inverse problem is much easier than the direct problem!

Example 4. The $m \times n$ matrix A, $n \times 1$ vector $\mathbf{x}$ and $m \times 1$ vector $\mathbf{b}$ satisfy $A\mathbf{x} = \mathbf{b}$.

Direct problem: given $A, \mathbf{x}$ compute $\mathbf{b}$.

Inverse problem: given $A, \mathbf{b}$, compute $\mathbf{x}$.

Example 5. (Differentiation)

Direct problem: given $f(x) \in C[0,1]$, find the indefinite integral

$$F(x) = \int_0^x f(y)\, dy.$$

Inverse problem: given $F(0) = 0$ and $F(x) \in C^1[0,1]$, find $F'(x)$.

(Recall that $C[0,1]$ ($C^1[0,1]$) is the set of all continuous functions (functions with continuous derivative) on the interval $[0,1]$.)

Example 6. (Parameter identification) Heat flows in a steady state through an insulated inhomogeneous rod with a known heat source and the temperature held at zero at the end-points. Under modest restrictions, the temperature function $u(x)$ obeys the law

$$-\Big(k(x)u'\Big)' = f(x),\ \ 0 < x < 1$$

with boundary conditions $u(0) = 0 = u(1)$, thermal conductivity $k(x)$, $0 \le x \le 1$ and $f(x)$ determined by the heat source.

Direct Problem: given parameters $k(x), f(x)$, find $u(x) = u(x; k)$.

Inverse Problem: given $f(x)$ and measurement of $u(x)$, find $k(x)$.

The mapping from parameter set (inputs) to solutions (outputs) is not a linear function, i.e., $u(x; k_1 + k_2) \neq u(x; k_1) + u(x; k_2)$, and its inverse may not even be defined – so $k(x)$ is not necessarily identifiable from any outputs!

A General Setting

What makes inverse problems any different from direct problems? In some cases, there really isn't much difference. What mathematicians know and love is the kind of problem defined by Hadamard around the beginning of the twentieth century. A **well-posed problem** is characterized by three properties:

1. The problem has a solution.

2. The solution is unique.

3. The solution is *stable*, that is, it varies continuously with the given parameters of the problem.

A problem that is not well-posed is called **ill-posed**.

In numerical analysis we are frequently cautioned to make sure that a problem is well posed before we design solution algorithms.

Let's return to some of our examples. We saw earlier that inverse problems need not be linear processes. Nonetheless, linear problems are a focal point of inverse theory. If linear inverse problems present problems, we can only imagine how much more difficult nonlinear problems might be.

As a case in point, consider Example 4 ($A\mathbf{x} = \mathbf{b}$); multiplication by the $m \times n$ matrix A can be thought of as a linear operator (read "linear function") $K : \mathbb{R}^n \to \mathbb{R}^m$. In this language, the direct problem is to find $y = Kx$, given A, x, and the inverse problem is to find x, given K, y and $Kx = y$.

Likewise, Example 5 (differentiation) is a linear problem. There is a customary way of expressing this example in the language of "integral equations." Define $k(x, y)$ by

$$k(x,y) = \begin{cases} 0, & \text{if } 0 \le x < y \le 1 \\ 1, & \text{if } 0 \le y \le x \le 1 \end{cases}$$

so that $F(x) = \int_0^x f(y)\,dy$ can be expressed as

$$F(x) = \int_0^x k(x,y) f(y)\,dy = \int_0^1 k(x,y) f(y)\,dy$$

This is a *Volterra integral equation of the first kind* if the last term is deleted and a *Fredholm integral equation of the first kind* if the middle term is deleted. Denote by K the operator that maps $f(x)$ to $F(x)$. The punch line is the same as in Example 4 ($A\mathbf{x} = b$), that is, solving the linear operator equation $Kf = F$ for F is the direct problem, and solving it for f is the inverse problem.

So what's the fuss? The direct problem of computing F from $F = Kf$ is easy and the solution to the inverse problem is $f = K^{-1}F$, right?

Wrong! All of Hadamard's well-posedness requirements fall by the wayside, even for the "simple" inverse problem of solving for $\mathbf{x}$ with $A\mathbf{x} = \mathbf{b}$ a linear system.

1. The system

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

has no solution!

2. The system

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has infinitely many solutions.

3. The system

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$$

has no solution for $\varepsilon \neq 0$ and infinitely many for $\varepsilon = 0$, so solutions do not vary continuously with the parameter ε.

The Regularization Idea

How can we extricate ourselves from these problems, if only for the "simple" system $A\mathbf{x} = \mathbf{b}$?

Existence: We use an old trick: least squares, which finds the $\mathbf{x}$ that minimizes the size of the residual (squared) $\|\mathbf{b} - A\mathbf{x}\|^2$. This turns out to be equivalent to solving the *normal equations*

$$A^T A\mathbf{x} = A^T\mathbf{b},$$

a system which is guaranteed to have a solution. Further, we can see that if $A\mathbf{x} = \mathbf{b}$ has *any* solution, then every solution to the normal equations is a solution to $A\mathbf{x} = \mathbf{b}$. This trick extends to more abstract linear operators K of equations $Kx = y$ using the concept of "adjoint" operators K^* which play the part of a transpose matrix A^T.

Uniqueness: We "regularize" the problem. Rather than give the general theory, we'll illustrate it by one particular kind of regularization, called *Tikhonov* regularization. One introduces a regularization parameter $\alpha > 0$ in such a way that small α give us a problem that is "close" to the original. In the case of the normal equations, one can show that minimizing the modified residual

$$\|\mathbf{b} - A\mathbf{x}\|^2 + \alpha \|\mathbf{x}\|^2$$

leads to the linear system $\left(A^T A + \alpha I\right)\mathbf{x} = A^T\mathbf{b}$, where I is the identity matrix. One can show the coefficient matrix $A^T A + \alpha I$ is always nonsingular. Therefore, the problem has a unique solution. What should we do about α? This is one of the more fundamental (and intriguing) problems of inverse theory. Let's analyze one of our simple systems for insight, say

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Invariably, our input data for the inverse problem, $(1,1)$, has error in it, say we have $(1+\delta_1, 1+\delta_2)$ for data instead. Let $\delta = \delta_1 + \delta_2$. The regularized system becomes

$$\begin{bmatrix} 2+\alpha & 2 \\ 2 & 2+\alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2+\delta \\ 2+\delta \end{bmatrix} = (2+\delta) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which has unique solution

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2+\alpha & 2 \\ 2 & 2+\alpha \end{bmatrix}^{-1} (2+\delta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{2+\delta}{4+\alpha} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Observe that if the input error δ were 0, all we would have to do is let $\alpha \to 0$ and we would get the valid solution $\frac{1}{2}(1,1)$. But given that the input error is not zero, taking the limit as $\alpha \to 0$ gives us a worse approximation to a solution than we would otherwise get by choosing $\alpha \approx 2\delta$.

There are many questions here, e.g., how do we know in general what the best choice of regularization parameter is, if any? This and other issues would be the subject matter of a course in inverse theory.

Stability: We get this for free – for each *regularized* problem. We cannot hope to have stability for the unregularized problem $Kx = y$, even if K^{-1} is well defined (although stability happens in some cases). However, we have to look to infinite dimensional examples such as our Example 5 (operator K is integration), where it can be shown that K^{-1} (differentiation) exists but is not continuous, even though K is. Specifically, we have $K : C[0,1] \to C[0,1]$ via the rule $Kf(x) = \int_0^x f(y)\,dy$ is a one-to-one function. Measure size by the sup norm:

$$\|f\| = \sup_{0 \le x \le 1} |f(x)|$$

so that the "closeness" of $f(x)$ and $g(x)$ is determined by the number $\|f - g\|$. Then one can show that the operator K is continuous in the sense that if $f(x)$ and $g(x)$ are close, then so are $Kf(x)$ and $Kg(x)$.

Let $R = K(C[0,1])$, the range of K. Then $K^{-1} : R \to C[0,1]$ is also one-to-one. But it is not continuous. For consider the function

$$g_\varepsilon(x) = \varepsilon \sin\left(\frac{x}{\varepsilon^2}\right),$$

where $\varepsilon > 0$. We have $\|g_\varepsilon\| = \|g_\varepsilon - 0\| \leq \varepsilon$. So for small ε, $g_\varepsilon(x)$ is close to the zero function. Yet,

$$K^{-1}g_\varepsilon(x) = g_\varepsilon(x)' = \frac{\varepsilon}{\varepsilon^2}\cos\left(\frac{x}{\varepsilon^2}\right) = \frac{1}{\varepsilon}\cos\left(\frac{x}{\varepsilon^2}\right)$$

so that $\left\|K^{-1}g_\varepsilon\right\| = \frac{1}{\varepsilon}$, so that $K^{-1}g_\varepsilon$ becomes far from zero as $\varepsilon \to 0$. Hence K^{-1} is not a continuous operator.

These calculations raise a question: How do we "regularize" the differentiation operator? There is a Tikhonov regularization, but we want to show another somewhat simpler approach, which nicely illustrates the effects of noisy data.

Suppose we can sample F(x) at equally spaced nodes $x_k = kh$, $k = 0, 1, \ldots, n$ and $h = 1/n$.

Of course there is noise in the data, so instead of $F_k = F(x_k)$, we get $F_k + \delta_k$ with each $|\delta_k| \leq \delta$, say. We'll use centered differences to approximate the derivative values $F'_k = F'(x_k)$, $k = 1, \ldots, n-1$. We know from calculus that

$$\frac{F_{k+1} - F_{k-1}}{2h} = F'_k + \mathcal{O}(h^2)$$

Here "$\mathcal{O}(t)$" means roughly "a quantity no bigger than a constant multiple of t for sufficiently small t." What we *actually* calculate is

$$\begin{aligned}
\widehat{F'_k} &= \frac{\left(F_{k+1} + \delta_{k+1}\right) - \left(F_{k-1} + \delta_{k-1}\right)}{2h} \\
&= \frac{F_{k+1} - F_{k-1}}{2h} + \frac{\delta_{k+1} - \delta_{k-1}}{2h} \\
&= F'_k + \mathcal{O}(h^2) + \mathcal{O}(\frac{\delta}{h})
\end{aligned}$$

Think of h or n as the regularization parameter (recall that $h = 1/n$). In the ideal world of exact data and calculations, we would simply let $h \to 0$ to obtain ever better answers.

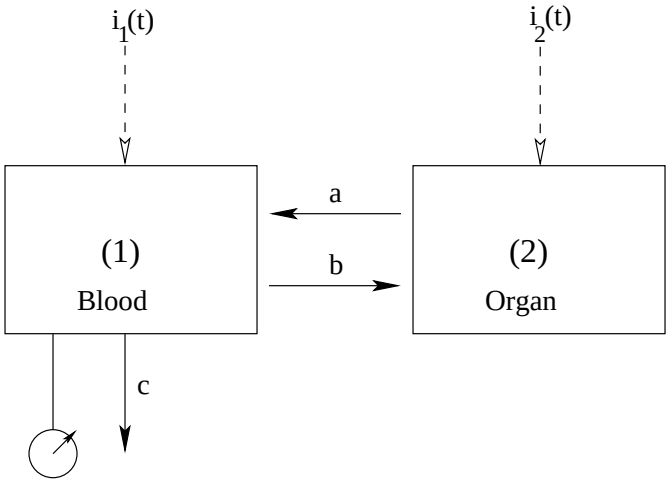
In the imperfect real world, the calculated term shows clearly that it is not wise to let $h \to 0$ in the situation of a nonzero sampling error δ. Rather we should try to balance the error terms in such a way that the theoretical error $\mathcal{O}(h^2)$ predominates. We might try $\mathcal{O}(h^2) \geq \mathcal{O}(\frac{\delta}{h})$, which suggests for some constant c, that $h^2 \geq \frac{c\delta}{h}$, that is, $h \geq (c\delta)^{1/3}$. This is a familiar lesson in elementary numerical analysis when one trys to account for the impact of finite precision calculations on numerical differentiation.

Of course, there are other approaches. One very useful idea is to approximately interpolate the data with some sort of smooth curve (like a spline) and then differentiate the smooth curve. A very charming presentation comparing centered differences and natural cubic splines was given in a recent American Math Monthly, vol 108, year 2001. Interestingly enough, a careful analysis gives the same sort of lower bounds on step size h as we saw above.

LECTURE 2: APPLICATIONS TO BIOLOGY MODELS

A Pharmacokinetic Model

Suppose a drug is introduced into the blood stream or a metabolite is already in the blood stream, and is thence exchanged with an organ within the body through bounding membranes. We'll assume that the exchange rates between these two locales are proportional to the source of metabolite being transported across the membranes and that the rate coefficients are non-negative constants. Here is a schematic for the action (t is time):



We are interested in the concentrations $x_i(t)$ of the metabolite in each compartment i. (Note: volumes have been incorporated into the rate numbers, which are *concentration* rates of exchange.) These functions are governed by a differential equation of the form (I'm appending the optional input in parentheses.)

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}' = \begin{bmatrix} -(a+c) & b \\ a & -b \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \left(+ \begin{bmatrix} i_1(t) \\ i_2(t) \end{bmatrix} \right)$$

This is an example of a "compartment model." One could imagine more compartments, more than one drug or metabolite under consideration and inputs into some of the compartments that would not be representated by rate coefficients, but by input functions.

Notations for Compartment Model: $\mathbf{x}(t)' = \mathbf{A}\mathbf{x}(t) + B\mathbf{w}(t)$

- $\mathbf{x}(t) = (x_1(t), x_2(t))$ is the **response function**.
- $\mathbf{y}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = C\mathbf{x}(t)$ is the **output function**.
- C is the **output sampling matrix**.
- $\mathbf{w}(t) = (u_1(t), u_2(t))$ is a **set of controlled inputs**.
- $\mathbf{u}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} i_1(t) \\ i_2(t) \end{bmatrix} = B\mathbf{w}(t)$ is the **input function**.
- B is the **input distribution matrix**.

With all of this notation, the governing dynamics of this compartment model can be given as

$$\mathbf{x}'(t) = A\mathbf{x}(t) + B\mathbf{w}(t)$$

Direct problem: Given the model, compute the input function $\mathbf{u}(t) = B\mathbf{w}(t)$ and output function $\mathbf{y}(t) = C\mathbf{x}(t)$.

Inverse Problem: Given the inputs and outputs, compute the matrix A.

Note: This inverse problem is a special sort of parameter identification problem, where the parameters are constants. The goal is to really identify them (at least in principle), so skip regularization parameters, nearby solutions, etc. If knowledge of the inputs and outputs is sufficient to compute the parameters explicitly (in principle), we say the system is **identifiable**.

Let's examine our example (without inputs) in detail and see if we can design an experiment that helps us identify the system. Here our system is

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}' = \begin{bmatrix} -(a+c) & b \\ a & -b \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

and our output function is

$$\mathbf{y}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = x_1(t).$$

We calculate the characteristic polynomial of A as

$$\begin{aligned} p(s) &= \begin{vmatrix} -(a+c)-s & b \\ a & -b-s \end{vmatrix} \\ &= s^2 + (a+b+c)s + bc \end{aligned}$$

which has discriminant

$$(a+b+c)^2 - 4cb = (b-c)^2 + a^2 + 2a(b+c).$$

This tells us that the eigenvalues of A are repeated if $b = c$ and $a = 0$, and otherwise they are real and distinct.

Before we analyze the examples, we define a **data point** to be any scalar component of the output vector or the response vector at a particular time. We will always count the first time of measurement as $t = 0$. Thus, for example, if you have initial conditions for the response function of our example system, then you have already two data points. BTW, we make no *a priori* assumptions about knowledge of initial conditions for the response function of our model. What we want to know is what data points, if any, are sufficient to identify the rate parameters of this problem.

Case 1: (Repeated roots) In this case we assume that we know a priori that the roots are repeated, i.e., $a = 0$ and $b = c$. In this case we know that the only eigenvalue to the system is given by $\lambda^2 + 2c\lambda + c^2 = 0$, so $\lambda = -c$. The general solution to the system is

$$\begin{aligned} x_1(t) &= d_1 cte^{-ct} + d_2 e^{-ct} \\ x_2(t) &= d_1 e^{-ct} \end{aligned}$$

Our first data point occurs at $t = 0$, say $x_1(0) = x_1^{(0)}$, so we get $d_2 = x_1^{(0)}$. If we tap the outputs at two later times, say t_1 and t_2, we get two nonlinear equations in d_1 and c that can be solved by a Newton's method. If one of our data points is not an output, but an initial condition, say $x_2(0) = x_2^{(0)}$, then we immediately obtain that $d_1 = y_0$ and one more output $x_1^{(1)} = x_1(t_1)$ will give us enough information to solve for c. Thus we need three (at least!) data points to identify the parameters uniquely.

N.B. Why "at least"? The problem is a little trickier than it appears: one can get the same value $x_1(t)$ at two separate times, yielding the same equations in the variable $x = ct$. A common remedy is to sample many data points, then find values of the parameters that minimize the sum of squares of residuals $\left\| x_1(t_k) - x_1^{(k)} \right\|^2$ resulting from each data point (*output least squares* method; add a regularization term to get the *penalized least squares* method.) In any case, this system is *identifiable*.

Case 2. (Distinct real roots) Here the general solution is

$$\begin{aligned} x_1(t) &= d_{11}e^{\lambda_1 t} + d_{12}e^{\lambda_2 t} \\ x_2(t) &= d_{21}e^{\lambda_1 t} + d_{22}e^{\lambda_2 t} \end{aligned}$$

where λ_1, λ_2 are the eigenvalues of A. We also have that

$$\begin{aligned} \lambda_1\lambda_2 &= bc \\ \lambda_1 + \lambda_2 &= -(a+b+c) \end{aligned}$$

so that both λ_1, λ_2 are non-positive. Thus, four outputs should suffice to determine $d_{11}, \lambda_1, d_{12}$ and λ_2. Thus $x_1(t)$ is determined in principle and all further outputs add nothing to our knowledge but confirmation of the form of $x_1(t)$. This information comes close to determining the ceofficients of A in view of the above two equations. However, these equations are insufficient to determine all three of a, b, c. Therefore, this system is *not identifiable* as it stands!

Remedies:

(1) If initial conditions are available, use them as the first two data points. Then sample three more outputs. As above, we can determine

$$x_1(t) = d_{11}e^{\lambda_1 t} + d_{12}e^{\lambda_2 t}$$

explicitly. Differentiate this solution for $x_1(t)$ and evaluate at $t = 0$ to obtain

$$d_{11}\lambda_1 + d_{12}\lambda_2 = x_1'(0) = -(a+c)x_1(0) + bx_2(0)$$

The only time this equation would not add new information is when both $x_1(0)$ and $x_2(0)$ are zero or of same magnitude and opposite sign. Neither of these conditions give physical initial conditions. This will solve the problem and in this context, the system is *identifiable*.

(2) Sample the outputs with input function $\mathbf{u}(t) = [u_2(t)]$. An appropriate input, e.g.,

$$u_2(t) = \begin{cases} 0, \text{ for } 0 \le t < 1 \\ 1, \text{ for } 1 \le t < \infty \end{cases}$$

(or most inputs, really) will yield sufficient information to identify the system.

An Exponential Growth Model

Suppose a population function $u(t)$ can be modeled by the exponential growth law

$$\frac{du}{dt} = ru$$

where $r = r(t)$ is a non-constant growth rate.

Direct problem: Given the model, i.e., growth rate function $r(t)$, compute the population growth function $u(t)$.

Inverse Problem: Given the population function $u(t)$, compute the growth rate function $r(t)$.

In this case, the inverse problem seems easy:

$$r(t) = \frac{du/dt}{u} = \frac{d}{dt} \ln u(t).$$

So all we have to do is take the natural log of the data and differentiate that function (or differentiate and divide by the function.)

What could be simpler, right?

Wrong! You know what the catch is from Lecture 1: numerical differentiation of data with errors in it is much more difficult than it sounds, and taking the log of the data first could make things even worse. If we have time, we'll have some fun with this example later by way of the MATLAB GUI Numderiv.

A Size Dependent Population Model

The direct problem here is a model for a population density function $v = v(x,t)$ where $t \geq 0$ is time and x, $0 \leq x \leq 1$ is a scaled size class variable. Think of x as a kind of parameter that quantifies some characteristic of the population. For example, x could be age; remember that we have scaled x so that in scaled numbers 0 is the age of a new born and 1 is the largest age achieved by any member of the populace. Another example is that x could represent scaled size. Whatever the characteristic of this parameter, we assume that members of the populace progress steadily in time from class size 0 to class size 1 and along the way give birth to newborns of class size 0 according to some class size birth rate function.

The basic parameter functions for this model:

- $v(x,t)$ is the population density of members of the species with class size x at time t.
- $g(x)$ is a size dependent growth rate.
- $\mu(x)$ is a size dependent mortality rate.
- $p(x)$ is a size dependent birth rate.
- $\psi(x)$ is the initial population distribution at time $t = 0$.

A balance law analysis yields model equations:

$$\begin{aligned} v_t + (gv)_x &= -\mu v,\ 0 \le x \le 1,\ t \ge 0 \\ g(0)v(0,t) &= \int_0^1 p(s)v(s,t)\,ds,\ t \ge 0 \\ v(x,0) &= \psi(x),\ 0 \le x \le 1 \end{aligned}$$

Direct Problem: Given $g(x)$, $\mu(x)$, $p(x)$ and $\Psi(x)$, compute the population density function $v(x,t)$.

Inverse Problem: Given all but one of the functions $g(x)$, $\mu(x)$, $p(x)$ and $\Psi(x)$, and suitable data sampling of $v(x,t)$, compute the remaining parameter function.

This is ongoing research jointly with Alexandar Densiov. Here is a sampling of what we've discovered so far. Regarding the direct problem:

1. Even the direct problem is nontrivial. One has to do a little work to show it is a well-posed problem (it is.)

2. Numerical computation of the direct problem is also nontrivial, but we have an apparently decent method.

Regarding inverse problems:

1. Suppose $g(x) = g$ is a constant, $T = 1/g$ and all parameters except μ are given. Under some reasonable restrictions, $\mu(x)$ is uniquely determined by the data $v(1,t)$, $0 \leq t \leq T$ and Volterra integral equation

$$\int_1^x \mu(s)\,ds = g \ln\left(\frac{u(1,\frac{1-x}{g})}{\phi(x)}\right), \; 0 \leq x \leq 1.$$

2. Suppose $g(x) = g$ is a constant and all parameters except $q(x)$ are given. In general, $q(x)$ cannot be uniquely determined by *any* measurements of $v(x,t)$.