Pollution Source Identification

Arthur Wiedner
CE 204
December 17, 2007

Outline

1 Introduction
   • Why do we care?
   • Problem definition

2 Solving
   • A simple case: closed-form solution
   • A numerical method

3 Next Step
Outline

1 Introduction
   • Why do we care?
   • Problem definition

2 Solving
   • A simple case: closed-form solution
   • A numerical method

3 Next Step

Why is this problem interesting?

• 300,000 to 400,000 contaminated sites

Figure: Water pollution
Why is this problem interesting?

- 30,000 to 40,000 contaminated sites
- Expected cleanup cost: 1 trillion $ (4000$ per US citizen)

Figure: Water pollution

Why is this problem interesting?

- 30,000 to 40,000 contaminated sites
- Expected cleanup cost: 1 trillion $ (4000$ per US citizen)
- Chemical fingerprinting and records are seldom sufficient to determine the source.

Figure: Water pollution
Forward problem

Governing equation:

\[
\frac{\partial C}{\partial t} = \nabla \left[ D \nabla C \right] - \nabla \cdot \left[ v \cdot C \right] - \lambda C \\
C(x, 0) = C_0(t), \quad 0 \leq t \leq T_{\text{Eng}} \\
C(x, 0) = C_0(x) \quad \text{initial concentration profile} \\
C_{\text{Sensor}} (t) = f(t) \quad \text{measurements}
\]

Some analytical (closed-form) solutions in simple cases. Solved numerically most of the time.

Inverse problem

- The problem of finding the source and possibly the time history of the pollution is a “time inversion” problem.
- Due to the irreversibility of diffusion it is fundamentally ill-posed.
  This means that we do not know in general:
  1. If a solution exists
  2. If it is unique
  3. How much it is sensitive to the observed data.
- Since the problem is severely underdetermined we have to make assumptions on the unknown function.
Outline

1. Introduction
   - Why do we care?
   - Problem definition

2. Solving
   - A simple case: closed-form solution
   - A numerical method

3. Next Step

Several methods are used

- Use of closed-form solutions
- Brute-force forward solving and optimization
- Probability and backward solving
- Statistical Estimation
A groundwater model

\[ \frac{\partial C}{\partial t} = D \Delta C + q \]

We can consider the equation far from the sources, and a decay rate for the contaminant:

\[ \frac{\partial C}{\partial t} = D \Delta C - \lambda C \]

Figure: Henri Darcy

We can add carefully chosen boundary conditions to the model:

\[ \lim_{|x| \to \infty} C(x, t) = G_0 \]

\[ \int_{\Omega} (C(x, t) - G_0) \, dx = M_0 \delta(t - \tau_0) \]

\[ \lim_{t \to 0} (C(x, t) - G_0) = 0 \]

The solution to the direct problem is then:

\[ C(x, t) = \frac{M_0 \delta(t - \tau_0)}{\sqrt{4\pi (t - \tau_0)}} \exp \left( \frac{-||x - \xi||^2}{4D(t - \tau_0)} \right) + G_0 \]
Identifying the source

We can take the log:

\[ \frac{|x - x_0|^2}{2k(x - x_0)} - \lambda(t - t_0) - \frac{\pi}{2} t(t - t_0) + \ln(M_0) - \frac{\pi}{2} t \ln(4D) \]

We also introduce additional parameters to simplify the notation:

\[ \kappa = (2D)^{-\frac{1}{2}} \quad \text{and} \quad \alpha = \ln(M_0) - \frac{\pi}{2} t \frac{1}{2} \ln(4D) \]

and consider the accordingly modified measurements \( \ln(C(x, y, z) - C_0) = \gamma \). We thus have:

\[ (k, \lambda, \kappa, \alpha) = \frac{|x - x_0|^2}{(t_0 - t)^2} - \lambda(t_0 - t) - \frac{\pi}{2} t(t_0 - t_0) + \alpha \]

Theorem

For a space of any dimension \( \mathbb{R}^n \), four measurements of \( C(x_1, x_2) \) at two different times \( t_1 > t_2 \) and \( t_1 > t_0 \) common at each of distinct locations \( x_1 \) and \( x_2 \) suffice to express \( t_0 \) in closed form:

\[ t_0 = \frac{(y_{1,1} - y_{2,1})(y_{2,2} - y_{2,1})}{(y_{2,1} - y_{2,1})(y_{2,2} - y_{2,1}) - (y_{2,1} - y_{2,1})(y_{2,2} - y_{2,1})} \]

Two additional measurements at the same locations at a third time lead to closed-form solution for the decay rate \( \lambda \) and for auxiliary parameter \( \alpha \).

(They are solutions of a system function of the \( t_0 \) with a non-zero determinant.)
With the condition that sensor locations \((x_i)\) are in general position

\[ \forall m \in \{2 \ldots n+1\} \] no subset of \(m\) points

is entirely included in an affine (hyper)plane of dimension \(n+2\).

**Theorem**

For \(n+2\) measurements of \((x_k, t_k)\) at the same time \(t_k > t_0\) at each

\(n+2\) points \(x_k \neq x_0\) in general position suffice to express the location of

the source \(x_0\) in closed form as well as parameter \(v\) and \(D_0\).

By using the previous equations and by substituting the values of \(t_0, \alpha, \lambda\n\)
in \((k,f)\) we get

\[ \kappa(\|x - x_0\|^2 = -2(\alpha - t_0) \left( v + \lambda(t_0 - t_0) + \frac{\alpha}{2} \right) \right) \]

By forming the ratio of the following form \((h \neq f)\) we obtain \(n+1\)
independent equations:

\[ \frac{\|x - x_0\|^2}{\|x - x_0\|^2} = r_{h,A,h} \]

Those are the equations of the median plane if \(r_{h,A,h} = 1\) or of a sphere

which center is the centroid of \((x_0, 1)\) and \((x_0 - r_{h,A,h})\). \(x_0\) is the

intersection of all these.
**Probabilistic model**

We define the probability density function $F_q$ as $\frac{\tilde{C}(x, t)}{M_0}$. For a divergence-free flow and non-reactive transport:

$$\frac{\partial F_q}{\partial t} = -q \cdot \nabla F_q + \nabla \left[ D \cdot \nabla F_q \right] \quad (1)$$

$$F_q(x_0, 0) = \delta(x - x_0) \quad (2)$$

The backward model:

$$\frac{\partial F_{\rho}}{\partial t} = q \cdot \nabla F_{\rho} + \nabla \left[ D \cdot \nabla F_{\rho} \right] \quad (3)$$

$$F_{\rho}(x_0, 0) = \delta(x - x_0) \quad (4)$$

The boundary conditions are

$$\nabla F_{\rho} = 0 \text{ on } \partial \Omega$$

**Concept**

**Figure**: Backward model concept.

$$F_q(x_1, t_f) = F_{\rho}(x_1, t_f)$$

where $t_f = _{t_{\text{source}} - t}$

$$C_q(x_1, t_f) = \frac{1}{M_0} C(x_1, t_f)^2 \delta q$$
Solving

We can run the backward flow model and watch for the isocontour obtained by equation:

$$G_t(x_0, y_0) = \frac{1}{M} \cdot C(x_0, y_0)^2 \cdot dx_0$$

- Running this from a single point does not yield a unique solution. But it can be done from several points since the equations are true for all points.
- Since the equation is linear we can use as initial condition for the backward model the whole concentration profile at detection. We can then backtrack it until it shrinks discovering the source.
What to do next?

For the project:
- Numerical example: Have the simulation working.
- After the project: Maybe investigating on other methods, there is still a lot to do.

Questions?