1D Inverse Heat Equation

An illustrative example is the inversion for the initial condition for a one-dimensional heat equation.
We consider a rod of lenght L, and we are interested in reconstructing the initial temperature profile m given some noisy measurements d of the temperature profile at a later time T.

Forward problem

Given

- The initial temperature profile $u(x,0) = m(x)$,
- The termal diffusivity k,
- A prescribed temperature $u(0,t) = u(L,t) = 0$ at the extremities of the rod,

solve the heat equation

$$\begin{cases} \dfrac{\partial u}{\partial t} - k\dfrac{\partial^2 u}{\partial x^2} = 0 & \forall x \in (0.L),\ \forall t \in (0,T) \\ u(x,0) = m(x) & \forall x \in [0,L] \\ u(0,t) = u(L,t) = 0 & \forall t \in (0,T], \end{cases}$$

and observe the temperature at the final time T: $\mathcal{F}(m) = u(x,T)$.

Analytical solution to the forward problem

Verify that if $m(x) = \sin(n\frac{\pi}{L}x), \quad n = 1, 2, 3, \cdots$
then $u(x,t) = e^{-k(n\frac{\pi}{L})^2 t}\sin(n\frac{\pi}{L}x)$ is the unique solution to the heat equation.

Inverse problem

Given the forward model $\mathcal{F}$ and a noisy measurament d of the temperature profile at time T, find the initial temperature profile m such that

$$\mathcal{F}(m) = d.$$

Ill-posedness of the inverse problem

Consider a perturbation $\delta m(x) = \epsilon \sin(n\frac{\pi}{L}x)$, where $\epsilon > 0$ and $n = 1, 2, 3, \cdots$.
Then, by linearity of the forward model $\mathcal{F}$, the corresponding perturbation $\delta d(x) = \mathcal{F}(m + \delta m) - \mathcal{F}(m)$ is $\delta d(x) = \epsilon e^{-k(n\frac{\pi}{L})^2 T} \sin(n\frac{\pi}{L}x)$, which converges to zero as $n \longrightarrow +\infty$.
Hence the ratio between δm and δd can become arbitrary large, which shows that the stability requirement for well-posedness can not be satisfied.

Discretization

We use finite differences in space and Implicit Euler in time.

$\longrightarrow$ Semidiscretization in space

We divide the $[0, L]$ interval in n_x subintervals of the same lenght $h = \frac{L}{n_x}$, and we denote with $u_j(t) := u(j \cdot h, t)$ the value of the temperature at point $x_j = j \cdot h$ and time t.

We use a centered finite difference approximation of the second derivative in space and write $\frac{\partial u_j(t)}{\partial t} - k \frac{u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)}{h^2}$ for $j = 1, 2, \cdots, n_x - 1$, with the boundary condition $u_0(t) = u_{n_x}(t) = 0$.

By letting $\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_{n_x-1}(t) \end{pmatrix}$

be the vector collecting the values of the temperature u at the points $x_j = j \cdot h$, we then write the system of ordinary differential equations (ODEs):

$$\frac{\partial}{\partial t}\mathbf{u}(t) + K\mathbf{u}(t) = \mathbf{0},$$

where $K \in \mathbb{R}^{(n_x-1)\times(n_x-1)}$ is the tridiagonal matrix given by

$$K = \frac{k}{h^2} \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \cdots & \cdots & \cdots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

$\longrightarrow$ Time discretization

We subdivide the time interval (0,T] in n_t time step of size $\Delta t = \frac{T}{n_t}$. By letting $\mathbf{u}^{(i)} = \mathbf{u}(i \cdot \Delta t)$ denote the discretized temperature profile at time $t_i = i \cdot \Delta t$, the Implicit Euler scheme reads

$$\frac{\mathbf{u}^{(i+1)} - \mathbf{u}^{(i)}}{\Delta t} + K\mathbf{u}^{(i+1)} = \mathbf{0}, \qquad \text{for } i = 0, 1, \cdots, n_t - 1.$$

After simple algebraic manipulations and exploiting the initial condition $u(x, 0) = m(x)$, we then obtain

$$\begin{cases} \mathbf{u}^{(i+1)} = (I + \Delta t \cdot K)^{-1}\mathbf{u}^{(i)} \\ \mathbf{u}^{(0)} = \mathbf{m} \end{cases}$$

or equivalently, $\mathbf{u}^{(i)} = (I + \Delta t \cdot K)^{-i}\mathbf{m}$.

Código Parte 1

assembleMatrix generates the finite difference matrix $(I + \Delta t \cdot K)$ and solveFwd evaluates the forward model $F\mathbf{m} = (I + \Delta t \cdot K)^{-n_t}$ $\mathbf{m}$.

A naive solution to the inverse problem

If $\mathcal{F}$ is invertible a naive solution to the inverse problem $\mathcal{F}m = d$ is simply to set $m = \mathcal{F}^{-1}d$. The function naiveSolveInv computes the solution of the discretized inverse problem $\mathbf{m} = F^{-1}\mathbf{d}$ as
$\mathbf{m} = (I + \Delta t K)^{n_t}\mathbf{d}$.
It can be seen that:

- For a very coarse mesh ($n_x = 20$) and no measurement noise ($noise\_std\_dev = 0.0$) the naive solution is quite good.
- For a finer mesh ($n_x = 100$) and/or even small measurement noise ($noise\_std\_dev = 1e-4$) the naive solution is garbage.

Código Parte 2

Why does the naive solution fail?

Let $v_n = \sqrt{\frac{2}{L}} \sin(n\frac{\pi}{L}x)$ with $n = 1, 2, 3, \cdots$, then we have that $\mathcal{F} v_n = \lambda_n v_n$, where the eigenvalues $\lambda_n = e^{-kT(\frac{\pi}{L}n)^2}$.

Note 1:

- Large eigenvalues λ_n corresponds to smooth eigenfunctions v_n,
- Small eigenvalues λ_n corresponds to oscillatory eigenfuctions v_n.

The eigenvalues λ_n decays extremely fast, that is the matrix F, discretization of the forward model $\mathcal{F}$, is extremely ill conditioned.

Código Parte 3

Note 2:
The functions $v_n, n = 1, 2, 3, \cdots$, form an orthonormal basis of $L^2([0, 1])$.
That is, every function $f \in L^2([0, 1])$ can be written as

$$f = \sum_{n=1}^{\infty} \alpha_n \; v_n, \text{ where } \alpha_n = \int_0^1 \; f \; v_n \; dx.$$

Consider now the noisy problem

$$d = \mathcal{F} m_{\text{true}} + \eta,$$

where

- d is the data (noisy measurements)
- η is the noise: $\eta(x) = \sum_{n=1}^{\infty} \eta_n \; v_n(x)$
- m_{true} is the true value of the parameter that generated the data
- $\mathcal{F}$ is the forward heat equation

Then, the naive solution to the inverse problem $\mathcal{F}m = d$ is

$$m = \mathcal{F}^{-1}d = \mathcal{F}^{-1}(\mathcal{F}m_{\text{true}} + \eta) = m_{\text{true}} + \mathcal{F}^{-1}\eta =$$

$$m_{\text{true}} + \mathcal{F}^{-1}\sum_{n=1}^{\infty}\eta_n\ v_n = m_{\text{true}} + \sum_{n=1}^{\infty}\frac{\eta_n}{\lambda_n}\ v_n.$$

If the coefficients $\eta_n = \int_0^1 \eta(x)\ v_n(x)\ dx$ do not decay sufficiently fast with respect to the eigenvalues λ_n, then the naive solution is unstable.

This implies that oscillatory components can not reliably be reconstructed from noisy data since they correspond to small eigenvalues.

Regularization by filtering

Remedy is to dampen the terms corresponding to small eigenvalues and write

$$m \approx \sum_{n=1}^{\infty} \omega(\lambda_n^2)\ \lambda_n^{-1}\ \delta_n\ v_n,$$

where

- $\delta_n = \int_0^1 d(x)\ v_n(x)\ dx$ denotes the coefficients of the data d in the basis $\{v_n\}_{n=1}^{\infty}$,
- $\omega(\lambda_n^2)$ is a filter function that allows to drop/stabilize the terms corresponding to small λ_n

Consider α is a regularization parameter,

- Truncated Singular Value Decomposition:

$$\omega_\alpha(\lambda^2) = \begin{cases} 1, & \text{if } \lambda^2 \geq \alpha \\ 0, & \text{otherwise} \end{cases},$$

  Then, we have $m_{\text{svd}} = \sum_{n=1}^{N} \lambda_n^{-1}\ \delta_n\ v_n$, where N is largest index such that $\lambda_n^2 \geq \alpha$ (assuming that λ_n are sorted in a decreasing order).

- Tikhonov filter: $\omega_\alpha(\lambda^2) = \frac{\lambda^2}{\lambda^2+\alpha}$, ($\omega_\alpha(\lambda^2)$ is close to 1 when $\lambda \gg \alpha$, close to 0 when $\lambda \ll \alpha$). Then we have $m_{\text{tikh}} = \sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda_n^2+\alpha}\ \delta_n\ v_n$.

Código Parte 4