Regularization in Variational Data Assimilation

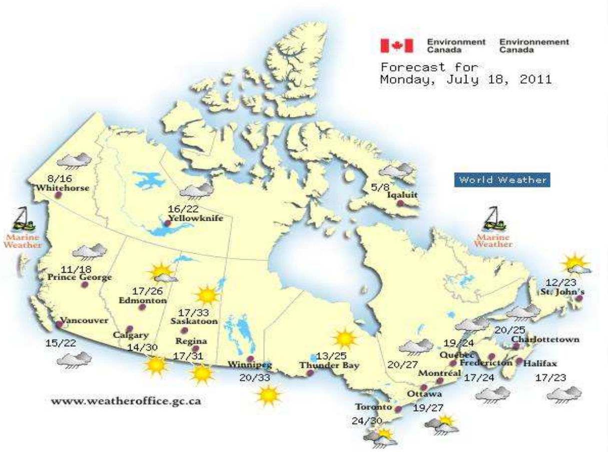
Melina Freitag

Department of Mathematical Sciences
University of Bath

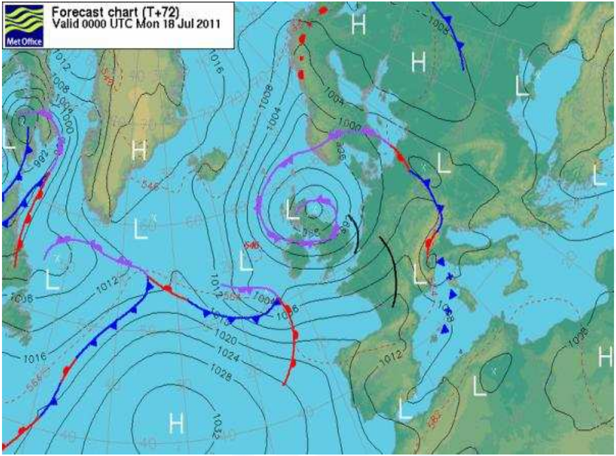
ICIAM 2011, Vancouver
Minisymposium MS49: Variational Data Assimilation
18th July 2011

joint work with C.J. Budd (Bath) and N.K. Nichols (Reading)

Weather forecast for today



The MetOffice weather forecast for today



Outline

Data Assimilation in NWP

Find an estimate $\mathbf{x}_i$ at time i for the true state of the atmosphere $\mathbf{x}_i^{\text{Truth}}$.

Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Data Assimilation in NWP

Find an estimate $\mathbf{x}_i$ at time i for the true state of the atmosphere $\mathbf{x}_i^{\text{Truth}}$.

A priori information $\mathbf{x}_i^B$

- background state (previous forecast)

Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Data Assimilation in NWP

Find an estimate $\mathbf{x}_i$ at time i for the true state of the atmosphere $\mathbf{x}_i^{\text{Truth}}$.

A priori information $\mathbf{x}_i^B$

- background state (previous forecast)

Models

- an operator linking state space and observation space (imperfect)

$$\mathbf{y}_i = H_i(\mathbf{x}_i)$$

Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Data Assimilation in NWP

Find an estimate $\mathbf{x}_i$ at time i for the true state of the atmosphere $\mathbf{x}_i^{\text{Truth}}$.

A priori information $\mathbf{x}_i^B$

- background state (previous forecast)

Models

- an operator linking state space and observation space (imperfect)

$$\mathbf{y}_i = H_i(\mathbf{x}_i)$$

- a model for the atmosphere (imperfect)

$$\mathbf{x}_{i+1} = M_{i+1,i}(\mathbf{x}_i)$$

Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Data Assimilation in NWP

Find an estimate $\mathbf{x}_i$ at time i for the true state of the atmosphere $\mathbf{x}_i^{\text{Truth}}$.

A priori information $\mathbf{x}_i^B$

- background state (previous forecast)

Models

- an operator linking state space and observation space (imperfect)

$$\mathbf{y}_i = H_i(\mathbf{x}_i)$$

- a model for the atmosphere (imperfect)

$$\mathbf{x}_{i+1} = M_{i+1,i}(\mathbf{x}_i)$$

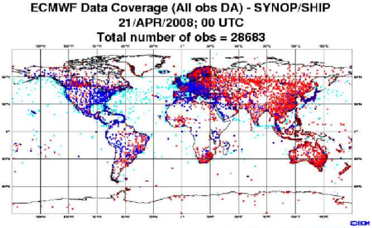
Observations $\mathbf{y}_i$

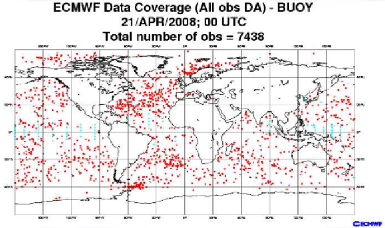
- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

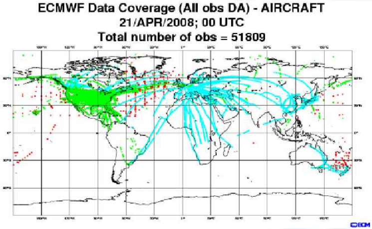
Assimilation algorithms

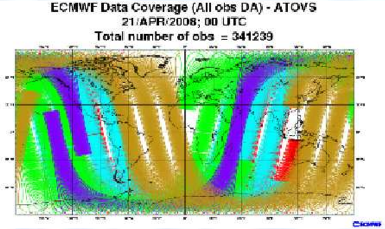
- find an (approximate) state of the atmosphere $\mathbf{x}_i$ at times i (usually $i = 0$)
- $\mathbf{x}_i^A$: Analysis (estimation of the true state after the DA)
- forecast future states of the atmosphere

Observations









Schematics of Data Assimilation

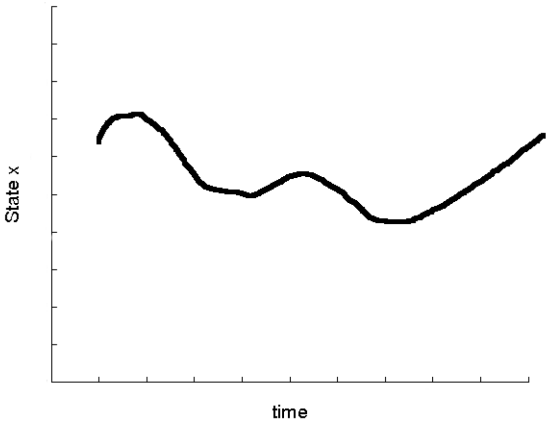


Figure: Background state $\mathbf{x}^B$

Schematics of Data Assimilation

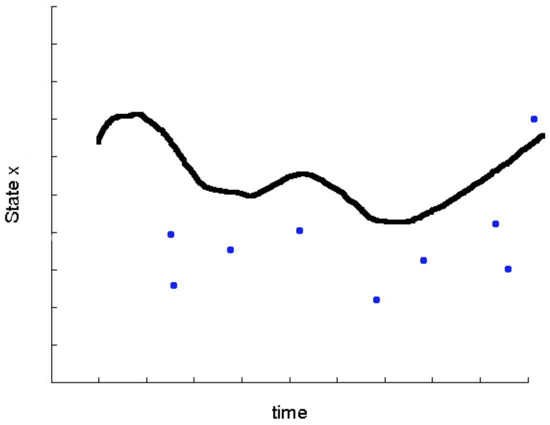


Figure: Observations **y**

Schematics of Data Assimilation

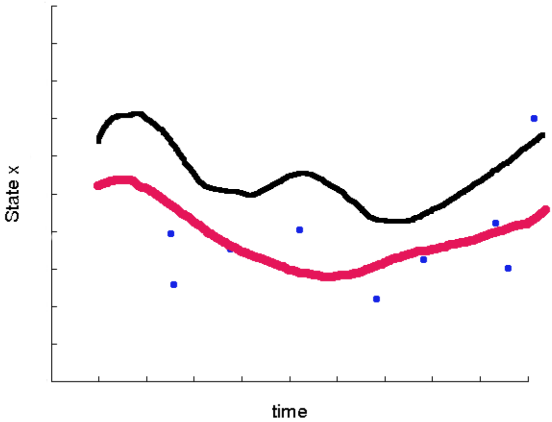


Figure: Analysis $\mathbf{x}^A$ (consistent with observations and model dynamics)

Data Assimilation in NWP

Under-determinacy

- Size of the state vector $\mathbf{x}$: $432 \times 320 \times 50 \times 7 = \mathcal{O}(10^7)$
- Number of observations (size of $\mathbf{y}$): $\mathcal{O}(10^5 - 10^6)$

Data Assimilation in NWP

Under-determinacy

- Size of the state vector $\mathbf{x}$: $432 \times 320 \times 50 \times 7 = \mathcal{O}(10^7)$
- Number of observations (size of $\mathbf{y}$): $\mathcal{O}(10^5 - 10^6)$

Assumptions

- background error $\varepsilon^B = \mathbf{x}^B - \mathbf{x}^{\text{Truth}}$ and covariance matrix $\mathbf{B} = \overline{(\varepsilon^B - \overline{\varepsilon}^B)(\varepsilon^B - \overline{\varepsilon}^B)^T}$
- observation error $\varepsilon^O = \mathbf{y} - H(\mathbf{x}^{\text{Truth}})$ and covariance matrix $\mathbf{R} = \overline{(\varepsilon^O - \overline{\varepsilon}^O)(\varepsilon^O - \overline{\varepsilon}^O)^T}$
- Non-trivial errors: $\mathbf{B}$, $\mathbf{R}$ are positive definite
- Unbiased errors: $\overline{\mathbf{x}^B - \mathbf{x}^{\text{Truth}}} = \overline{\mathbf{y} - H(\mathbf{x}^{\text{Truth}})} = 0$
- Uncorrelated errors: $\overline{(\mathbf{x}^B - \mathbf{x}^{\text{Truth}})(\mathbf{y} - H(\mathbf{x}^{\text{Truth}}))^T} = 0$

Optimal least-squares estimator

Cost function

Solution to the optimisation problem $\mathbf{x}^A = \arg\min J(\mathbf{x})$ where

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}^B)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^B) + \frac{1}{2}(\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x})) \\ &= J_B(\mathbf{x}) + J_O(\mathbf{x}) \end{aligned}$$

⇒Three-dimensional variational data assimilation (3DVar)

Optimal least-squares estimator

Cost function

Solution to the optimisation problem $\mathbf{x}^A = \arg\min J(\mathbf{x})$ where

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2}(\mathbf{x}-\mathbf{x}^B)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^B) + \frac{1}{2}(\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x})) \\ &= J_B(\mathbf{x}) + J_O(\mathbf{x}) \end{aligned}$$

⇒Three-dimensional variational data assimilation (3DVar)

Interpolation equations

$$\mathbf{x}^A = \mathbf{x}^B + \mathbf{K}(\mathbf{y} - H(\mathbf{x}^B)), \quad \text{where}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad \mathbf{K}\ldots \text{gain matrix}$$

⇒ Optimal interpolation

Four-dimensional variational assimilation (4DVar)

Minimise the cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$.

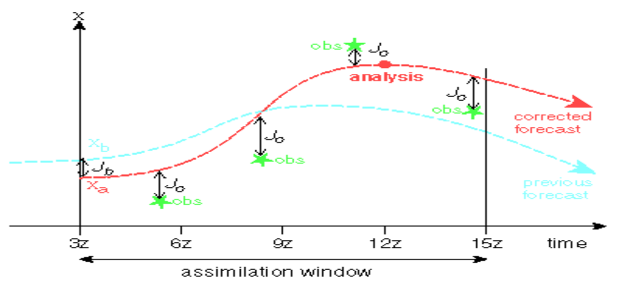


Figure: Copyright:ECMWF

Four-dimensional variational assimilation (4DVar)

Minimise the cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\rightarrow i}\mathbf{x}_0$.

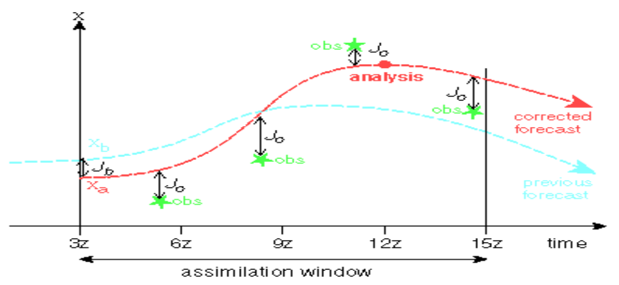


Figure: Copyright:ECMWF

Four-dimensional variational assimilation (4DVar)

Minimise the cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$.

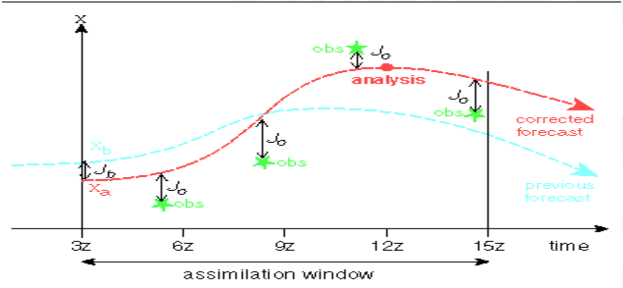


Figure: Copyright:ECMWF

Minimisation of the 4DVar cost function

- Use Newton's method in order to solve $\nabla J(\mathbf{x}_0) = 0$, that is

$$\begin{aligned} \nabla\nabla J(\mathbf{x}_0^k)\Delta\mathbf{x}_0^k &= -\nabla J(\mathbf{x}_0^k) \\ \mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta\mathbf{x}_0^k \end{aligned}$$

$k \geq 0$

Minimisation of the 4DVar cost function

- Use Newton's method in order to solve $\nabla J(\mathbf{x}_0) = 0$, that is

$$\begin{aligned} \nabla\nabla J(\mathbf{x}_0^k)\Delta\mathbf{x}_0^k &= -\nabla J(\mathbf{x}_0^k) \\ \mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta\mathbf{x}_0^k \end{aligned}$$

  $k \geq 0$

- Use approximate Hessian - Gauß-Newton method

$$\nabla J(\mathbf{x}_0) = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) - \sum_{i=1}^{n} \mathbf{M}_{i,0}(\mathbf{x}_0)^T \mathbf{H}_i^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i)),$$

  and

$$\nabla\nabla J(\mathbf{x}_0) = \mathbf{B}^{-1} + \sum_{i=1}^{n} \mathbf{M}_{i,0}(\mathbf{x}_0)^T \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{M}_{i,0}(\mathbf{x}_0).$$

Outline

Introduction

4DVar and Tikhonov regularisation

Application of L_1-norm regularisation in 4DVar

Motivation: Results from image processing

L_1-norm regularisation in 4DVar

Examples

Relation between 4DVar and Tikhonov regularisation

4DVar minimises

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{i0}\mathbf{x}_0$

Relation between 4DVar and Tikhonov regularisation

4DVar minimises

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{i0}\mathbf{x}_0$

or

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))$$

where

$$\hat{H} = [H_0^T, (H_1 M_{10}(t_1, t_0))^T, \ldots (H_n M_{n0}(t_n, t_0))^T]^T$$
$$\hat{\mathbf{y}} = [\mathbf{y}_0^T, \ldots, \mathbf{y}_n^T]^T$$

and $\hat{\mathbf{R}}$ is block diagonal with $\mathbf{R}_i$, $i = 0, \ldots, n$ on the diagonal.

Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))$$

Gauß-Newton method

$$\begin{aligned} \nabla\nabla J(\mathbf{x}_0^k)\Delta\mathbf{x}_0^k &= -\nabla J(\mathbf{x}_0^k) \\ \mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta\mathbf{x}_0^k \end{aligned}$$

Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^B) + \frac{1}{2}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0))$$

Gauß-Newton method

$$\begin{aligned}(\mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})\Delta\mathbf{x}_0^k &= -\mathbf{B}^{-1}(\mathbf{x}_0^k - \mathbf{x}_0^B) + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1}(\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0)) \\ \mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta\mathbf{x}_0^k\end{aligned}$$

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\begin{aligned}
\mathbf{B} &= \sigma_B^2 \mathbf{C}_B \\
\hat{\mathbf{R}} &= \sigma_R^2 \mathbf{C}_R \\
\mathbf{b} &= \mathbf{C}_R^{-\frac{1}{2}} (\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0)) \\
\mathbf{A} &= \mathbf{C}_R^{-\frac{1}{2}} \hat{\mathbf{H}} \mathbf{C}_B^{\frac{1}{2}} \\
\mu^2 &= \frac{\sigma_R^2}{\sigma_B^2}
\end{aligned}$$

Gauß-Newton method

$$\begin{aligned}
(\mathbf{B}^{-1} + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}) \Delta \mathbf{x}_0^k &= -\mathbf{B}^{-1}(\mathbf{x}_0^k - \mathbf{x}_0^B) + \hat{\mathbf{H}}^T \hat{\mathbf{R}}^{-1} (\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0)) \\
\mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta \mathbf{x}_0^k
\end{aligned}$$

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\begin{aligned}
\mathbf{B} &= \sigma_B^2 \mathbf{C}_B \\
\hat{\mathbf{R}} &= \sigma_R^2 \mathbf{C}_R \\
\mathbf{b} &= \mathbf{C}_R^{-\frac{1}{2}} (\hat{\mathbf{y}} - \hat{H}(\mathbf{x}_0)) \\
\mathbf{A} &= \mathbf{C}_R^{-\frac{1}{2}} \hat{\mathbf{H}} \mathbf{C}_B^{\frac{1}{2}} \\
\mu^2 &= \frac{\sigma_R^2}{\sigma_B^2}
\end{aligned}$$

Gauß-Newton method

$$\begin{aligned}
(\mu^2 \mathbf{I} + \mathbf{A}^T \mathbf{A}) \mathbf{C}_B^{-\frac{1}{2}} \Delta \mathbf{x}_0^k &= -\mu^2 \mathbf{C}_B^{-\frac{1}{2}} (\mathbf{x}_0^k - \mathbf{x}_0^B) + \mathbf{A}^T \mathbf{b} \\
\mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \Delta \mathbf{x}_0^k
\end{aligned}$$

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\mathbf{z}^k = \mathbf{C}_B^{-\frac{1}{2}}(\mathbf{x}_0^k - \mathbf{x}_0^B)$$

Gauß-Newton method

$$(\mu^2\mathbf{I} + \mathbf{A}^T\mathbf{A})(\mathbf{z}^{k+1} - \mathbf{z}^k) = -\mu^2\mathbf{z}^k + \mathbf{A}^T\mathbf{b}$$

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\mathbf{z}^k = \mathbf{C}_B^{-\frac{1}{2}}(\mathbf{x}_0^k - \mathbf{x}_0^B)$$

Gauß-Newton method

$$(\mu^2\mathbf{I} + \mathbf{A}^T\mathbf{A})(\mathbf{z}^{k+1} - \mathbf{z}^k) = -\mu^2\mathbf{z}^k + \mathbf{A}^T\mathbf{b}$$

Normal equations

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\mathbf{z}^k = \mathbf{C}_B^{-\frac{1}{2}}(\mathbf{x}_0^k - \mathbf{x}_0^B)$$

Gauß-Newton method

$$(\mu^2\mathbf{I} + \mathbf{A}^T\mathbf{A})(\mathbf{z}^{k+1} - \mathbf{z}^k) \quad = \quad -\mu^2\mathbf{z}^k + \mathbf{A}^T\mathbf{b}$$

Normal equations

Least squares solution

$$\left\| \begin{bmatrix} \mathbf{A} \\ \mu\mathbf{I} \end{bmatrix} (\mathbf{z}^{k+1} - \mathbf{z}^k) + \begin{bmatrix} -\mathbf{b} \\ \mu\mathbf{z}^k \end{bmatrix} \right\|_2^2 \rightarrow \min$$

at each Gauß-Newton method step

Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

$$\mathbf{z}^k = \mathbf{C}_B^{-\frac{1}{2}}(\mathbf{x}_0^k - \mathbf{x}_0^B)$$

Gauß-Newton method

$$(\mu^2\mathbf{I} + \mathbf{A}^T\mathbf{A})(\mathbf{z}^{k+1} - \mathbf{z}^k) \quad = \quad -\mu^2\mathbf{z}^k + \mathbf{A}^T\mathbf{b}$$

Normal equations

Least squares solution

$$\left\| \begin{bmatrix} \mathbf{A} \\ \mu\mathbf{I} \end{bmatrix} (\mathbf{z}^{k+1} - \mathbf{z}^k) + \begin{bmatrix} -\mathbf{b} \\ \mu\mathbf{z}^k \end{bmatrix} \right\|_2^2 \to \min$$

at each Gauß-Newton method step or

$$\|\mathbf{A}\mathbf{z}^{k+1} - (\mathbf{A}\mathbf{z}^k + \mathbf{b})\|_2^2 + \mu^2\|\mathbf{z}^{k+1}\|_2^2$$

Tikhonov regularisation

Ill-posed problems

Given an operator **A** we wish to solve

$$\mathbf{Az} = \mathbf{c}$$

it is well-posed if

- solution exits
- solution is unique
- is stable ($\mathbf{A}^{-1}$ continuous)

Ill-posed problems

Given an operator $\mathbf{A}$ we wish to solve

$$\mathbf{A}\mathbf{z} = \mathbf{c}$$

it is well-posed if

- solution exits
- solution is unique
- is stable ($\mathbf{A}^{-1}$ continuous)

but ..

In finite dimensions existence and uniqueness can be imposed, but

- discrete problem of underlying ill-posed problem becomes ill- conditioned
- singular values of $\mathbf{A}$ decay to zero

Ill-posed problems

Given an operator $\mathbf{A}$ we wish to solve

$$\mathbf{Az} = \mathbf{c}$$

it is well-posed if

- solution exits
- solution is unique
- is stable ($\mathbf{A}^{-1}$ continuous)

but ..

In finite dimensions existence and uniqueness can be imposed, but

- discrete problem of underlying ill-posed problem becomes ill- conditioned
- singular values of $\mathbf{A}$ decay to zero
- Tikhonov regularization

$$\begin{aligned}\mathbf{z} =& \arg\min \left\{ \|\mathbf{Az} - \mathbf{c}\|^2 + \mu^2 \|\mathbf{z}\|^2 \right\} \\ =& (\mathbf{A}^T\mathbf{A} + \mu^2 \mathbf{I})^{-1} \mathbf{A}^T \mathbf{c} \\ =& (\mathbf{V}\boldsymbol{\Sigma}^T \mathbf{U}^T \mathbf{U} \Sigma \mathbf{V}^T + \mu^2 \mathbf{V}\mathbf{V}^T)^{-1} \mathbf{V}\boldsymbol{\Sigma}^T \mathbf{U}^T \mathbf{c} \\ =& \mathbf{V} \mathrm{diag} \left(\frac{s_i^2}{s_i^2 + \mu^2} \frac{1}{s_i} \right) \mathbf{U}^T \mathbf{c} = \mathbf{z}_\mu = \sum_{i=1}^{n} \frac{s_i^2}{s_i^2 + \mu^2} \frac{u_i^T \mathbf{c}}{s_i} \mathbf{v}_i\end{aligned}$$

Outline

Results from image deblurring: L_1 regularisation

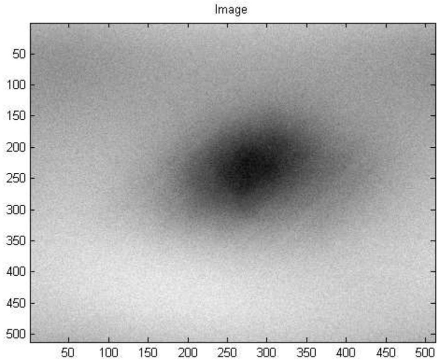


Figure: Blurred picture

Results from image deblurring: L_1 regularisation

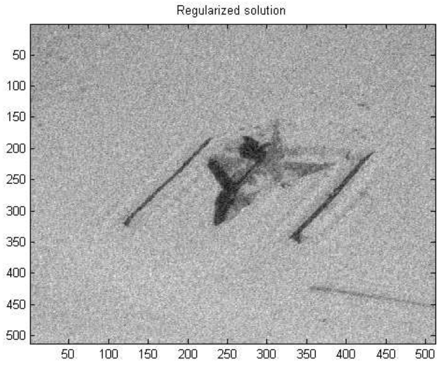


Figure: Tikhonov regularisation $\min \left\{ \|\mathfrak{A}\mathfrak{x} - \mathfrak{b}\|_2^2 + \alpha \|\mathfrak{x}\|_2^2 \right\}$

Results from image deblurring: L_1 regularisation

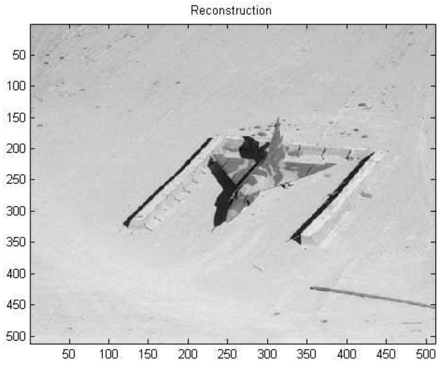


Figure: L_1-norm regularisation $\min\left\{\|\mathfrak{A}\mathfrak{x}-\mathfrak{b}\|_2^2+\alpha\|\mathfrak{x}\|_1\right\}$

3 Regularisation Methods

4DVar

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_2^2$$

3 Regularisation Methods

4DVar

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_2^2$$

L_1-norm regularisation

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_1$$

3 Regularisation Methods

4DVar

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_2^2$$

L_1-norm regularisation

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_1$$

Total Variation regularisation

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2 \|\mathbf{z}^{k+1}\|_2^2 + \beta \|\mathbf{D}\mathbf{x}_0^{k+1}\|_1$$

where $\mathbf{x}_0^{k+1} = \mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{x}_0^B$ and $\mathbf{D}$ is a matrix approximating the derivative of the solution.

Least mixed norm solutions

Solve

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2\|\mathbf{z}^{k+1}\|_2^2$$

using Least squares and

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \mu^2\|\mathbf{z}^{k+1}\|_2^2 + \beta\|\mathbf{D}\mathbf{x}_0^{k+1}\|_1$$

using quadratic programming (see Fu/Ng/Nikolova/Barlow 2006).

Least mixed norm solutions

Consider

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta\|\mathbf{D}\mathbf{x}_0^{k+1}\|_1$$

where $\mathbf{x}_0^{k+1} = \mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{x}_0^B$

Least mixed norm solutions

Consider

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta\|\mathbf{D}\mathbf{x}_0^{k+1}\|_1$$

where $\mathbf{x}_0^{k+1} = \mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{x}_0^B$

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta\|\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{D}\mathbf{x}_0^B\|_1$$

Least mixed norm solutions

Consider

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta\|\mathbf{D}\mathbf{x}_0^{k+1}\|_1$$

where $\mathbf{x}_0^{k+1} = \mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{x}_0^B$

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta\|\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{D}\mathbf{x}_0^B\|_1$$

Set

$$\mathbf{v} = \beta\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \beta\mathbf{D}\mathbf{x}_0^B.$$

and split $\mathbf{v}$ into its positive and negative part:

$$\mathbf{v} = \mathbf{v}^+ - \mathbf{v}^-$$

where

$$\begin{aligned} \mathbf{v}^+ &= \max(\mathbf{v}, 0) \\ \mathbf{v}^- &= \max(-\mathbf{v}, 0) \end{aligned}$$

Least mixed norm solutions

With

$$\mathbf{v} = \beta \mathbf{D} \mathbf{C}_B^{\frac{1}{2}} \mathbf{z}^{k+1} + \beta \mathbf{D} \mathbf{x}_0^B$$

and

$$\mathbf{v} = \mathbf{v}^+ - \mathbf{v}^-$$

the solution to

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta \|\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \mathbf{D}\mathbf{x}_0^B\|_1$$

is equivalent to

Least mixed norm solutions

With

$$\mathbf{v} = \beta \mathbf{D} \mathbf{C}_B^{\frac{1}{2}} \mathbf{z}^{k+1} + \beta \mathbf{D} \mathbf{x}_0^B$$

and

$$\mathbf{v} = \mathbf{v}^+ - \mathbf{v}^-$$

the solution to

$$\min_{\mathbf{z}^{k+1}} \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 + \beta \|\mathbf{D}\mathbf{C}_B^{\frac{1}{2}} \mathbf{z}^{k+1} + \mathbf{D}\mathbf{x}_0^B\|_1$$

is equivalent to

$$\min_{\mathbf{z}^{k+1}, \mathbf{v}^+, \mathbf{v}^-} \left\{ \mathbf{1}^T \mathbf{v}^+ + \mathbf{1}^T \mathbf{v}^- + \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 \right\}$$

subject to

$$\begin{aligned} \beta \mathbf{D}\mathbf{C}_B^{\frac{1}{2}} \mathbf{z}^{k+1} + \beta \mathbf{D}\mathbf{x}_0^B &= \mathbf{v}^+ - \mathbf{v}^- \\ \mathbf{v}^+, \mathbf{v}^- &\geq \mathbf{0}. \end{aligned}$$

Least mixed norm solutions

$$\min_{\mathbf{z}^{k+1},\mathbf{v}^+,\mathbf{v}^-}\left\{\mathbf{1}^T\mathbf{v}^+ + \mathbf{1}^T\mathbf{v}^- + \|\mathbf{A}\mathbf{z}^{k+1}-\mathbf{c}\|_2^2\right\}$$

subject to

$$\begin{aligned}\beta\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1}+\beta\mathbf{D}\mathbf{x}_0^B &= \mathbf{v}^+-\mathbf{v}^-\\ \mathbf{v}^+,\mathbf{v}^- &\geq \mathbf{0}.\end{aligned}$$

or

Least mixed norm solutions

$$\min_{\mathbf{z}^{k+1},\mathbf{v}^+,\mathbf{v}^-} \left\{ \mathbf{1}^T\mathbf{v}^+ + \mathbf{1}^T\mathbf{v}^- + \|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{c}\|_2^2 \right\}$$

subject to

$$\begin{aligned} \beta\mathbf{D}\mathbf{C}_B^{\frac{1}{2}}\mathbf{z}^{k+1} + \beta\mathbf{D}\mathbf{x}_0^B &= \mathbf{v}^+ - \mathbf{v}^- \\ \mathbf{v}^+, \mathbf{v}^- &\geq \mathbf{0}. \end{aligned}$$

or

$$\min_{\mathbf{w}} \left\{ \frac{1}{2}\mathbf{w}^T\mathbf{G}\mathbf{w} + \mathbf{g}^T\mathbf{w} \right\}$$

subject to

$$\mathbf{E}\mathbf{w} = \mathbf{e} \quad \text{and} \quad \mathbf{F}\mathbf{w} \geq \mathbf{0}.$$

where

$$\mathbf{G} = \begin{bmatrix} 2\mathbf{A}^T\mathbf{A} & & \\ & \mathbf{0} & \\ & & \mathbf{0} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} -2\mathbf{A}^T\mathbf{b} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{0} & & \\ & -\mathbf{I} & \\ & & -\mathbf{I} \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} \beta\mathbf{D}\mathbf{C}_B^{\frac{1}{2}} & -\mathbf{I} & \mathbf{I} \end{bmatrix} \quad \mathbf{w} = \begin{bmatrix} \mathbf{z}^{k+1} & \mathbf{v}^+ & \mathbf{v}^- \end{bmatrix}^T \quad \mathbf{e} = -\beta\mathbf{D}\mathbf{x}_0^B$$

Outline

Example 1 - Linear advection equation

$$u_t + u_z = 0,$$

on the interval $z \in [0, 1]$, with periodic boundary conditions. The initial solution is a square wave defined by

$$u(z,0) = \begin{cases} 0.5 & 0.25 < z < 0.5 \\ -0.5 & z < 0.25 \quad \text{or} \quad z > 0.5. \end{cases}$$

This wave moves through the time interval, the model equations are defined by the upwind scheme

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta z}(U_j^n - U_{j-1}^n),$$
$$U_0^{n+1} = U_N^{n+1},$$

where $j = 1, \ldots, N$, $\Delta z = \frac{1}{N}$ and n is the number of time steps. We take $N = 100$, $\Delta t = 0.005$.

Setup

- length of the assimilation window: 40 time steps
- perfect observations, noisy and sparse observations
- $\mathbf{R} = 0.01$.
- $\mathbf{B} = \mathbf{I}$ and $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$, where $L = 5$
- use MATLAB `quadprog.m`

4DVar - perfect and full observations, $\mathbf{B} = \mathbf{I}$

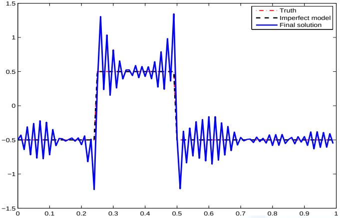


Figure: $t = 0$

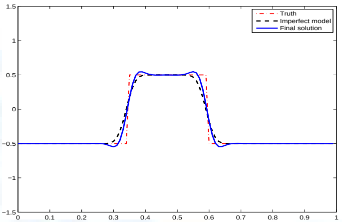


Figure: $t = 20$

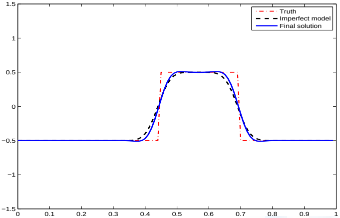


Figure: $t = 40$

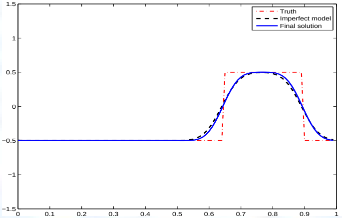


Figure: $t = 80$

L1 - perfect and full observations, $\mathbf{B} = \mathbf{I}$

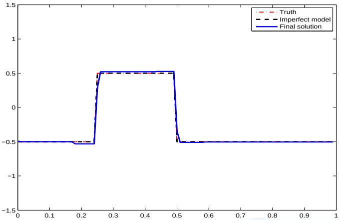


Figure: $t = 0$

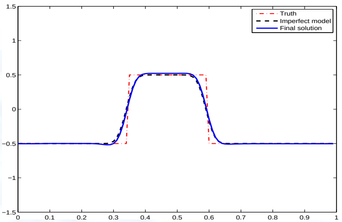


Figure: $t = 20$

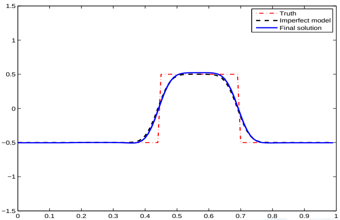


Figure: $t = 40$

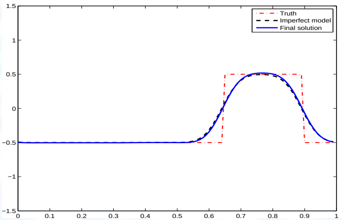


Figure: $t = 80$

4DVar - noisy and sparse observations, $\mathbf{B} = \mathbf{I}$

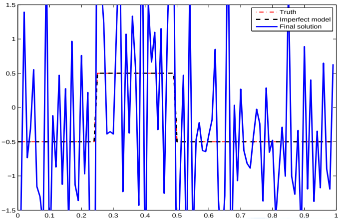


Figure: $t = 0$

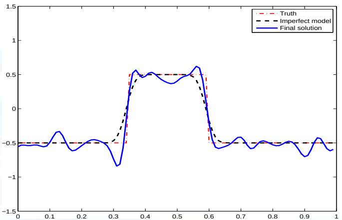


Figure: $t = 20$

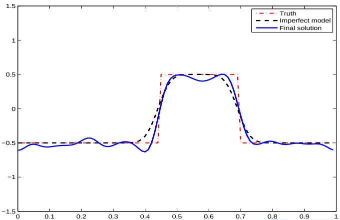


Figure: $t = 40$

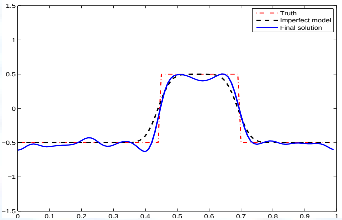


Figure: $t = 80$

L1 - noisy and sparse observations, $\mathbf{B} = \mathbf{I}$

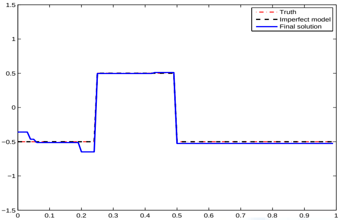


Figure: $t = 0$

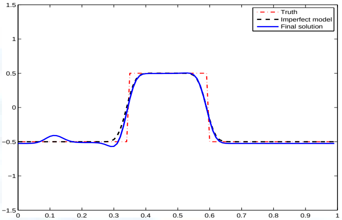


Figure: $t = 20$

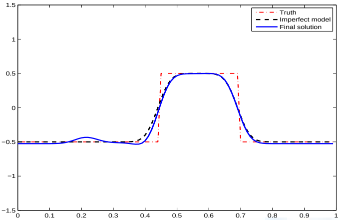


Figure: $t = 40$

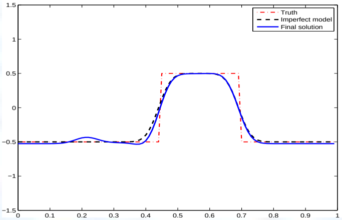


Figure: $t = 80$

4DVar - noisy and sparse observations, $\mathbf{B} = 0.1 e^{-\frac{|i-j|}{2L^2}}$

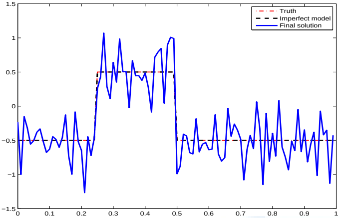


Figure: $t = 0$

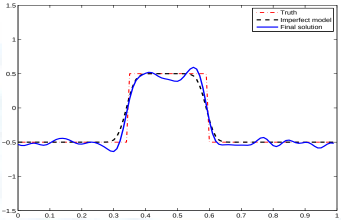


Figure: $t = 20$

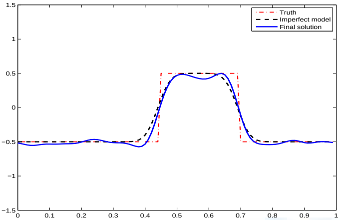


Figure: $t = 40$

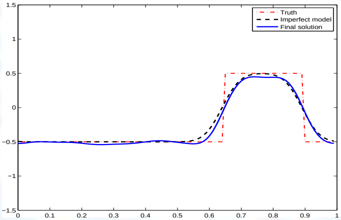


Figure: $t = 80$

L1 - noisy and sparse observations, $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$

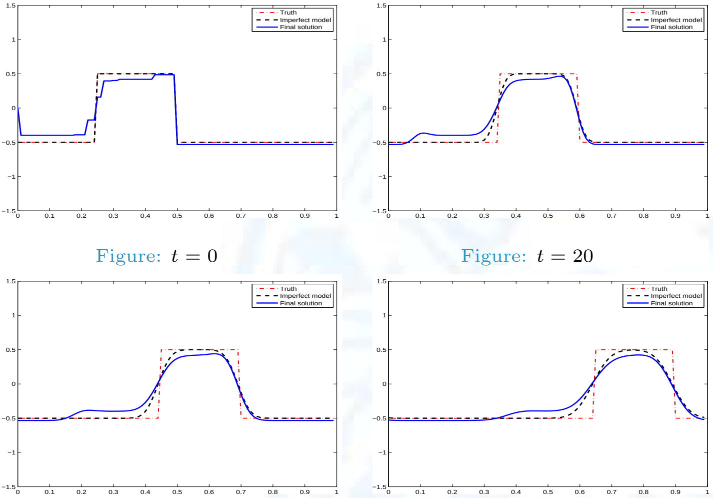


Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$

Example 2 - Burgers' equation

$$u_t + u\frac{\partial u}{\partial x} = u + f(u)_x = 0, \quad f(u) = \frac{1}{2}u^2$$

with initial conditions

$$u(x,0) = \begin{cases} 2 & 0 \leq x < 2.5 \\ 0.5 & 2.5 \leq x \leq 10. \end{cases}$$

Discretising

$$x(j) = 10(j - 1/2)\Delta x; \quad U^0(x(j)) = \begin{cases} 2 & 0 \leq x(j) < 2.5 \\ 0.5 & 2.5 \leq x(j) \leq 10. \end{cases}$$

where $j = 1, \ldots, N$, $\Delta x = \frac{1}{N}$ and n is the number of time steps. We take $N = 100$, $\Delta t = 0.001$.

Exact solution and model error

Exact solution - method of characteristics

Riemann problem

$$u(x,t) = \begin{cases} 2 & 0 \leq x < 2.5 + st \\ 0.5 & 2.5 + st \leq x \leq 10, \end{cases}$$

where $s = 1.25$

Numerical solution - model error

- the Lax-Friedrichs method (smearing out the shock)

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)).$$

- the Lax-Wendroff method (oscillations near the shock).

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)) + \frac{\Delta t^2}{2\Delta x^2}\left(A_{j+\frac{1}{2}}(f(U_{j+1}^n) - f(U_j^n)) - A_{j-\frac{1}{2}}(f(U_j^n) - f(U_{j-1}^n))\right.$$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

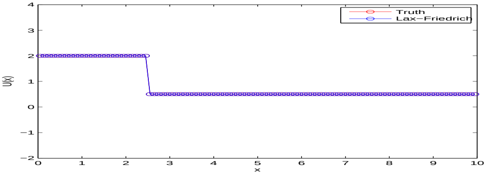


Figure: $t = 0$

Lax-Wendroff method

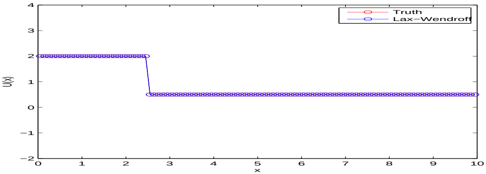


Figure: $t = 0$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

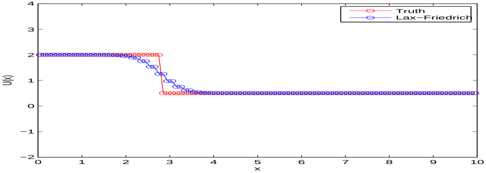


Figure: $t = 25$

Lax-Wendroff method

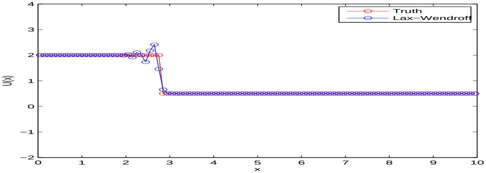


Figure: $t = 25$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

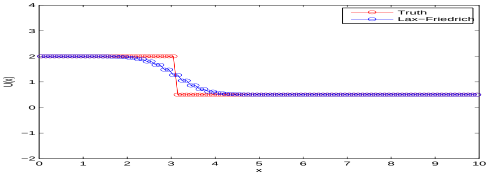


Figure: $t = 50$

Lax-Wendroff method

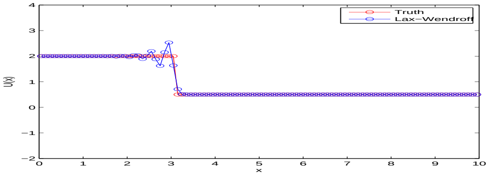


Figure: $t = 50$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

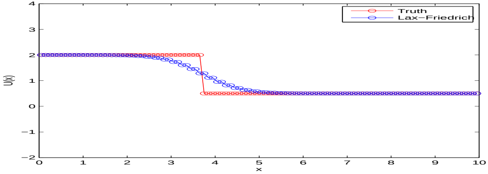


Figure: $t = 100$

Lax-Wendroff method

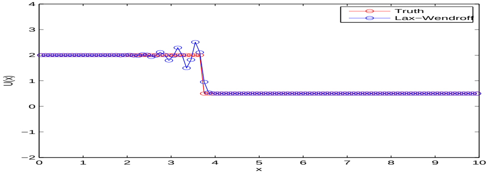


Figure: $t = 100$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

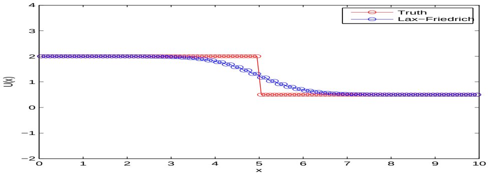


Figure: $t = 200$

Lax-Wendroff method

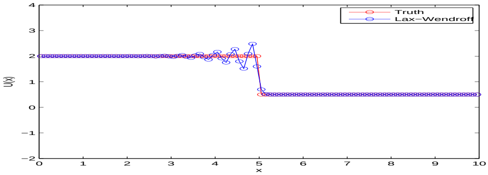


Figure: $t = 200$

Setup

- length of the assimilation window: 100 time steps
- noisy and sparse observations
- $\mathbf{R} = 0.01$.
- $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$, where $L = 5$
- use MATLAB `quadprog.m`

Lax-Friedrichs method

4DVar - noisy and sparse observations, $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$

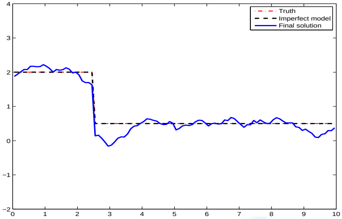


Figure: $t = 0$

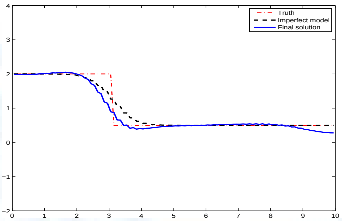


Figure: $t = 50$

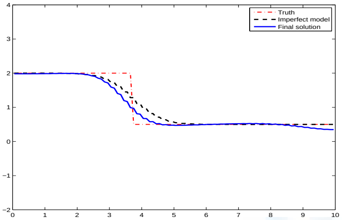


Figure: $t = 100$

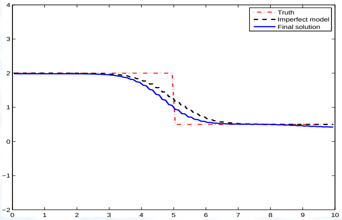


Figure: $t = 200$

L1 - noisy and sparse observations, $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$

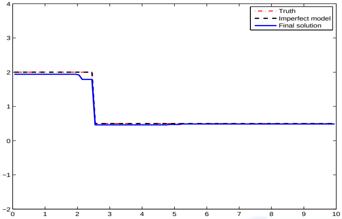


Figure: $t = 0$

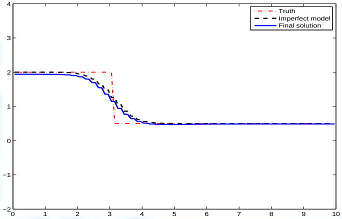


Figure: $t = 50$

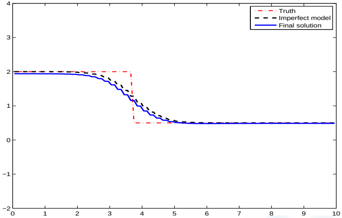


Figure: $t = 100$

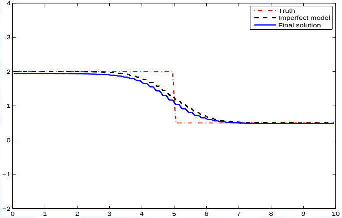


Figure: $t = 200$

Lax-Wendroff method

4DVar - noisy and sparse observations, $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$

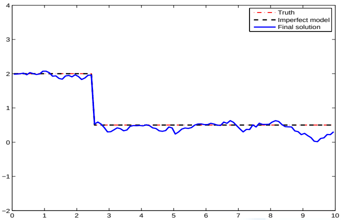


Figure: $t = 0$

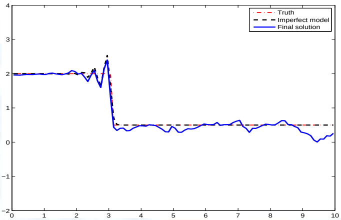


Figure: $t = 50$

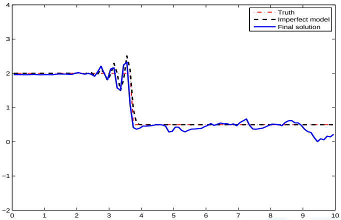


Figure: $t = 100$

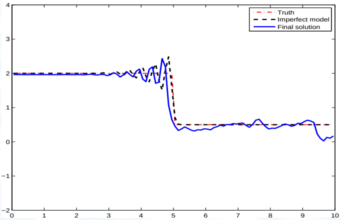


Figure: $t = 200$

L1 - noisy and sparse observations, $\mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}}$

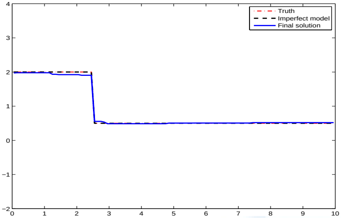


Figure: $t = 0$

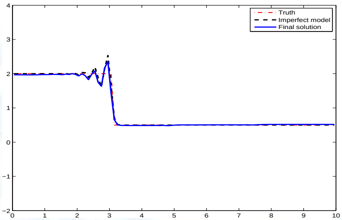


Figure: $t = 50$

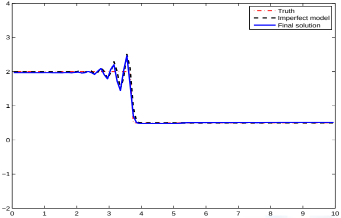


Figure: $t = 100$

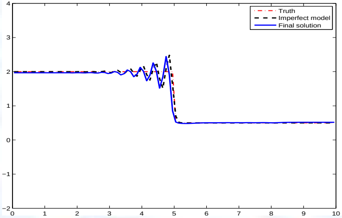


Figure: $t = 200$

Conclusions, questions and further work

- L_1-norm regularisation recovers discontinuity better than 4DVar
- Further work: analysis of methods; tests in 2D, 3D
- multiscale methods, other regularisation approaches