Resolution of sharp fronts in the presence of model error in variational data assimilation

Melina Freitag

Department of Mathematical Sciences
University of Bath

Computational Mathematics and Applications Seminar
Rutherford Appleton Laboratory
11th February 2010

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

Tikhonov regularisation

4DVar and Tikhonov regularisation

Motivation: Results from image processing

Application of $L_1$-norm regularisation in 4DVar
Find an estimate $x_i$ at time $i$ for the true state of the atmosphere $x_i^{\text{Truth}}$.

Observations $y_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes
Data Assimilation in NWP

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

A priori information $x_i^B$
- background state (usual previous forecast)

Observations $y_i$
- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes
Data Assimilation in NWP

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

A priori information $x_i^B$

- background state (usual previous forecast)

Models

- a model for the atmosphere (imperfect)

$$x_{i+1} = M(x_i)$$

Observations $y_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes
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 (usual previous forecast)

### Models

- a model for the atmosphere (imperfect)
  \[ \mathbf{x}_{i+1} = M(\mathbf{x}_i) \]

- a function linking model space and observation space (imperfect)
  \[ \mathbf{y}_i = H(\mathbf{x}_i) \]

### Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes
Data Assimilation in NWP

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

A priori information $x_i^B$
- background state (usual previous forecast)

Models
- a model for the atmosphere (imperfect)
  $$x_{i+1} = M(x_i)$$
- a function linking model space and observation space (imperfect)
  $$y_i = H(x_i)$$

Observations $y_i$
- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Assimilation algorithms
- find an (approximate) state of the atmosphere $x_i$ at times $i$ (usually $i = 0$)
- forecast future states of the atmosphere
- $x_i^A$: Analysis (estimation of the true state after the DA)
Schematics of Data Assimilation

Figure: Background state $\mathbf{x}^B$
Schematics of Data Assimilation

Figure: Observations $y$
Schematics of Data Assimilation

Figure: Analysis $x^A$ (consistent with observations and model dynamics)
Data Assimilation in NWP

Underdeterminacy

- Size of the state vector $\mathbf{x}$: $432 \times 320 \times 50 \times 7 = \mathcal{O}(10^7)$
Data Assimilation in NWP

Underdeterminacy

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

- Size of the state vector $\mathbf{x}$: $432 \times 320 \times 50 \times 7 = \mathcal{O}(10^7)$
- Number of observations (size of $\mathbf{y}$): $\mathcal{O}(10^5 - 10^6)$
- Operator $H$ (nonlinear!) maps from state space into observations space: $\mathbf{y} = H(\mathbf{x})$
Error variables

Error statistics

- background error $\varepsilon^B = \mathbf{x}^B - \mathbf{x}^{Truth}$ and covariance
  $\mathbf{B} = (\varepsilon^B - \overline{\varepsilon}^B)(\varepsilon^B - \overline{\varepsilon}^B)^T$

- observation error $\varepsilon^O = \mathbf{y} - H(\mathbf{x}^{Truth})$ and covariance
  $\mathbf{R} = (\varepsilon^O - \overline{\varepsilon}^O)(\varepsilon^O - \overline{\varepsilon}^O)^T$

- analysis error $\varepsilon^A = \mathbf{x}^A - \mathbf{x}^{Truth}$ and covariance
  $\mathbf{A} = (\varepsilon^A - \overline{\varepsilon}^A)(\varepsilon^A - \overline{\varepsilon}^A)^T$

- minimise analysis error $\text{tr}(\mathbf{A}) = \|\varepsilon^A - \overline{\varepsilon}^A\|^2$
Error variables

Error statistics

- background error $\varepsilon^B = x^B - x^{Truth}$ and covariance $B = (\varepsilon^B - \bar{\varepsilon}^B)(\varepsilon^B - \bar{\varepsilon}^B)^T$
- observation error $\varepsilon^O = y - H(x^{Truth})$ and covariance $R = (\varepsilon^O - \bar{\varepsilon}^O)(\varepsilon^O - \bar{\varepsilon}^O)^T$
- analysis error $\varepsilon^A = x^A - x^{Truth}$ and covariance $A = (\varepsilon^A - \bar{\varepsilon}^A)(\varepsilon^A - \bar{\varepsilon}^A)^T$
- minimise analysis error $\text{tr}(A) = \|\varepsilon^A - \bar{\varepsilon}^A\|^2$

Assumptions

- Nontrivial errors: $B$, $R$ are positive definite
- Unbiased errors: $x^B - x^{Truth} = y - H(x^{Truth}) = 0$
- Uncorrelated errors: $(x^B - x^{Truth})(y - H(x^{Truth}))^T = 0$
Optimal least-squares estimator

**Cost function**
Solution of the variational optimisation problem $x^A = \arg \min J(x)$ where

$$J(x) = (x - x^B)^T B^{-1} (x - x^B) + (y - H(x))^T R^{-1} (y - H(x))$$

$$= J_B(x) + J_O(x)$$

$\Rightarrow$ Three-dimensional variational data assimilation (3DVar)
Optimal least-squares estimator

Cost function
Solution of the variational optimisation problem $\mathbf{x}^A = \arg \min J(\mathbf{x})$ where

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

$$= J_B(\mathbf{x}) + J_O(\mathbf{x})$$

⇒ Three-dimensional variational data assimilation (3DVar)

Interpolation equations

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

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

⇒ Optimal interpolation
Bayesian interpretation

Non-Gaussian PDF’s (probability density function)

- $P(x)$ is a priori PDF (background)
- $P(y|x)$ is the observation PDF (likelihood of the observations given background $x$)
Bayesian interpretation

Non-Gaussian PDF’s (probability density function)

- $P(x)$ is a priori PDF (background)
- $P(y|x)$ is the observation PDF (likelihood of the observations given background $x$)
- $P(x|y)$ conditional probability of the model state given the observations,
  Bayes theorem:

$$\arg_{x} \max P(x|y) = \arg_{x} \max \frac{P(x)P(y|x)}{P(y)}$$
Bayesian interpretation

Non-Gaussian PDF’s (probability density function)

- $P(x)$ is a priori PDF (background)
- $P(y|x)$ is the observation PDF (likelihood of the observations given background $x$)
- $P(x|y)$ conditional probability of the model state given the observations, Bayes theorem:

$$\arg_{x} \max P(x|y) = \arg_{x} \max \frac{P(x)P(y|x)}{P(y)}$$

Gaussian PDF’s

$$P(x|y) = c_1 \exp \left( -(x - x^B)^T B^{-1} (x - x^B) \right) \cdot c_2 \exp \left( -(y - H(x))^T R^{-1} (y - H(x)) \right)$$

$x^A$ is the maximum a posteriori estimator of $x^{\text{Truth}}$. Maximising $P(x|y)$ equivalent to minimising $J(x)$
Four-dimensional variational assimilation (4DVar)

Minimise the cost function

\[
J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i))
\]

subject to model dynamics \( x_i = M_{0\rightarrow i} x_0 \)

Figure: Copyright: ECMWF
4DVar analysis

Model dynamics

Strong constraint: model states $x_i$ are subject to

$$x_i = M_{0\rightarrow i}x_0$$

nonlinear constraint optimisation problem (hard!)
4DVar analysis

Model dynamics
Strong constraint: model states $x_i$ are subject to

$$x_i = M_{0\rightarrow i}x_0$$

nonlinear constraint optimisation problem (hard!)

Simplifications

- **Causality** (forecast expressed as product of intermediate forecast steps)

  $$x_i = M_{i,i-1}M_{i-1,i-2}\ldots M_{1,0}x_0$$

- **Tangent linear hypothesis** ($H$ and $M$ can be linearised)

  $$y_i - H_i(x_i) = y_i - H_i(M_{0\rightarrow i}x_0) = y_i - H_i(M_{0\rightarrow i}x_0^B) - H_iM_{0\rightarrow i}(x_0 - x_0^B)$$

  $M$ is the tangent linear model.

- **unconstrained quadratic optimisation problem** (easier).
Minimisation of the 4DVar cost function

- Use Newton’s method in order to solve $\nabla J(x_0) = 0$, that is

$$\nabla^2 J(x_0^k) \Delta x_0^k = -\nabla J(x_0^k)$$

$$x_0^{k+1} = x_0^k + \Delta x_0^k$$

$k \geq 0$
Minimisation of the 4DVar cost function

- Use Newton’s method in order to solve $\nabla J(x_0) = 0$, that is
  \[
  \nabla \nabla J(x_0^k) \Delta x_0^k = -\nabla J(x_0^k) \\
  x_0^{k+1} = x_0^k + \Delta x_0^k
  \]
  for $k \geq 0$

- Use approximate Hessian - Gaussian-Newton method
  \[
  \nabla J(x_0) = B^{-1}(x_0 - x_0^B) - \sum_{i=1}^{N} M_{i,0}(x_0)^T H^T R^{-1}(y_i - H(x_i)),
  \]
  and
  \[
  \nabla \nabla J(x_0) = B^{-1} + \sum_{i=1}^{N} M_{i,0}(x_0)^T H^T R^{-1} H M_{i,0}(x_0).
  \]
Ill-posed problems

Given an operator $A$ we wish to solve

$$Ax = b$$

it is well-posed if

- solution exists
- solution is unique
- is stable ($A^{-1}$ continuous)
Ill-posed problems

Given an operator $A$ we wish to solve

$$Ax = b$$

it is well-posed if

- solution exits
- solution is unique
- is stable ($A^{-1}$ 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 $A$ decay to zero
A way out of this - Tikhonov regularisation

Solution to the minimisation problem

\[ x_\alpha = \arg \min \left\{ \|Ax - b\|^2 + \alpha \|x\|^2 \right\} \]

\[ = (A^T A + \alpha I)^{-1} A^T b \]

\[ = (V \Sigma^T U^T U \Sigma V^T + \alpha V V^T)^{-1} V \Sigma^T U^T b \]

\[ = V \text{diag} \left( \frac{s_i^2}{s_i^2 + \alpha s_i} \right) U^T b \]

where \(\alpha\) is called the regularisation parameter.
Solution to the minimisation problem

\[
x_\alpha = \arg \min \left\{ ||Ax - b||^2 + \alpha ||x||^2 \right\}
\]

\[
= (A^T A + \alpha I)^{-1} A^T b
\]

\[
= (V \Sigma^T U^T U \Sigma V^T + \alpha V V^T)^{-1} V \Sigma^T U^T b
\]

\[
= V \text{diag} \left( \frac{s_i^2}{s_i^2 + \alpha s_i} \right) U^T b
\]

\[
x_\alpha = \sum_{i=1}^{n} \frac{s_i^2}{s_i^2 + \alpha s_i} \frac{u_i^T b}{v_i} v_i
\]

where \( \alpha \) is called the regularisation parameter.
Relation between 4DVar and Tikhonov regularisation

4DVar minimises

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i)) \]

subject to model dynamics \( x_i = M_{0 \rightarrow i} x_0 \)
Relation between 4DVar and Tikhonov regularisation

4DVar minimises

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i)) \]

subject to model dynamics \( x_i = M_{0\rightarrow i} x_0 \)

or

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]

where

\[ \hat{H} = [H_0^T, (H_1 M(t_1, t_0))^T, \ldots, (H_n M(t_n, t_0))^T]^T \]

\[ \hat{y} = [y_0^T, \ldots, y_n^T]^T \]

and \( \hat{R} \) is block diagonal with \( R_i \) on diagonal.
Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]
Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]

Variable transformations

\[ B = \sigma_B^2 F_B \] and \( \hat{R} = \sigma_O^2 F_R \) and define new variable \( z := F_B^{-1/2} (x_0 - x_0^B) \)

\[ \hat{J}(z) = \mu^2 \|z\|_2^2 + \|F_B^{-1/2} \hat{d} - F_R^{-1/2} \hat{H} F_B^{1/2} z\|_2^2, \quad \mu^2 = \frac{\sigma_O^2}{\sigma_B^2} \]

This is Tikhonov regularisation!

\[ \hat{J}(z) = \|A z - b\|_2^2 + \alpha \|z\|_2^2 \]
Blurred and exact images

The blurring process as a linear model

- Let $\mathbf{X}$ be the exact image
- Let $\mathbf{B}$ be the blurred image

$$
\mathbf{x} = \text{vec}(\mathbf{X}) = \begin{bmatrix}
\mathbf{x}_1 \\
\vdots \\
\mathbf{x}_N
\end{bmatrix} \in \mathbb{R}^N, \quad \mathbf{b} = \text{vec}(\mathbf{B}) = \begin{bmatrix}
\mathbf{b}_1 \\
\vdots \\
\mathbf{b}_N
\end{bmatrix} \in \mathbb{R}^N
$$

are related by the linear model

$$
\mathbf{A}\mathbf{x} = \mathbf{b}
$$

where $\mathbf{A}$ is a blurring matrix.

Noise $\mathbf{b} = \mathbf{b}_{\text{exact}} + \mathbf{e}$

$$
\mathbf{x}_\text{Naive} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}_{\text{exact}} + \mathbf{A}^{-1}\mathbf{e} = \mathbf{x} + \mathbf{A}^{-1}\mathbf{e}
$$
Blurred and exact images - Need regularisation techniques!

Standard technique: Tikhonov regularisation

\[ \mathbf{x}_\alpha = \arg \min \{ \| \mathbf{A} \mathbf{x} - \mathbf{b} \|_2^2 + \alpha \| \mathbf{x} \|_2^2 \} \]
Blurred and exact images - Need regularisation techniques!

Standard technique: Tikhonov regularisation

$$x_{\alpha} = \arg \min \left\{ \|Ax - b\|_2^2 + \alpha \|x\|_2^2 \right\}$$

$L_1$ regularisation

In image processing, $L_1$-norm regularisation provides edge preserving image deblurring!

$$\min \left\{ \|Ax - b\|_2^2 + \alpha \|x\|_1 \right\}$$
Results from image deblurring: $L_1$ regularisation

Figure: Blurred picture
Results from image deblurring: $L_1$ regularisation

Figure: Tikhonov regularisation $\min \left\{ \|Ax - b\|_2^2 + \alpha \|x\|_2^2 \right\}$
Results from image deblurring: $L_1$ regularisation

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

In image processing, $L_1$-norm regularisation provides edge preserving image deblurring!

- $L_1$-norm regularisation beneficial in Data Assimilation?
- 4DVar smears out sharp fronts
In image processing, $L_1$-norm regularisation provides edge preserving image deblurring!

- $L_1$-norm regularisation beneficial in Data Assimilation?
- 4DVar smears out sharp fronts
- $L_1$-norm regularisation has the potential to overcome this problem!
Example 1

Burger’s 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

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

with \( \Delta x = \frac{1}{100} \) and \( j = 1, \ldots, N. \)
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-Friedrich 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).

\[
U_{j}^{n+1} = U_{j}^{n} - \frac{\Delta t}{2\Delta x} (f(U_{j+1}^{n}) - f(U_{j-1}^{n}))+ \\
\frac{\Delta t^2}{2\Delta x^2} \left( A_{j+1/2} (f(U_{j+1}^{n}) - f(U_{j}^{n})) - A_{j-1/2} (f(U_{j}^{n}) - f(U_{j-1}^{n})) \right)
\]
Visualisation - Truth trajectory and numerical solution

Lax-Friedrich method

\[ U(x) \]

\( \text{Truth} \)

\( \text{Lax-Friedrich} \)

\[ t = 0 \]

Lax-Wendroff method

\[ U(x) \]

\( \text{Truth} \)

\( \text{Lax-Wendroff} \)

\[ t = 0 \]
Visualisation - Truth trajectory and numerical solution

Lax-Friedrich method

\[ U(x) \]

\[ \text{Truth} \quad \text{Lax-Friedrich} \]

Figure: \( t = 25 \)

Lax-Wendroff method

\[ U(x) \]

\[ \text{Truth} \quad \text{Lax-Wendroff} \]

Figure: \( t = 25 \)
Visualisation - Truth trajectory and numerical solution

Lax-Friedrich method

![Lax-Friedrich method graph](image1)

Lax-Wendroff method

![Lax-Wendroff method graph](image2)

Figure: $t = 50$
Visualisation - Truth trajectory and numerical solution

Lax-Friedrich method

\[ U(x) \]

Truth
Lax–Friedrich

[Graph showing comparison between Truth and Lax–Friedrich methods at t = 100]

Lax-Wendroff method

\[ U(x) \]

Truth
Lax–Wendroff

[Graph showing comparison between Truth and Lax–Wendroff methods at t = 100]

Figure: \( t = 100 \)
Visualisation - Truth trajectory and numerical solution

Lax-Friedrich method

Figure: $t = 200$

Lax-Wendroff method

Figure: $t = 200$
3 Regularisation Methods

4DVar

\[
J(U^0) = \frac{1}{2} \|U_B^0 - U^0\|_{\alpha_B}^2 + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2
\]
3 Regularisation Methods

4DVar

\[
J(U^0) = \frac{1}{2} \|U_B^0 - U^0\|_{\alpha B}^2 + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2
\]

\[L_1\text{-norm regularisation}\]

\[
J(U^0) = \frac{1}{2} \|Z_B^0 - Z^0\|_p^p + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2
\]

where \(p = 1\) (or \(p = 1.0001\)) and \(Z = (\alpha B)^{-1/2}U\).
3 Regularisation Methods

4DVar

\[ J(U^0) = \frac{1}{2} \|U^0_B - U^0\|_{\alpha B}^2 + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2 \]

$L_1$-norm regularisation

\[ J(U^0) = \frac{1}{2} \|Z^0_B - Z^0\|_p^p + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2 \]

where \( p = 1 \) (or \( p = 1.0001 \)) and \( Z = (\alpha B)^{-1/2}U \).

Total Variation regularisation

\[ J(U^0) = \frac{1}{2} \|D(Z^0_B - Z^0)\|_p^p + \frac{1}{2} \sum_{i=1}^{N} \|Y_i - H_i(U_i)\|_{R_i}^2 \]

where \( D \) is a matrix approximating the derivative of the solution.
Least mixed norm solutions

Solve

\[ \min \{ \|Ax - b\|_2^2 + \alpha \|x\|_2^2 \} \]

using a Gauss-Newton method and

\[ \min \{ \|Ax - b\|_2^2 + \alpha \|Rx\|_1 \} \]

using quadratic programming tools: Let
Least mixed norm solutions

Solve

$$\min \left\{ \| Ax - b \|_2^2 + \alpha \| x \|_2^2 \right\}$$

using a Gauss-Newton method and

$$\min \left\{ \| Ax - b \|_2^2 + \alpha \| Rx \|_1 \right\}$$

using quadratic programming tools: Let

$$v = \alpha Rx.$$ 

and split $v$ into its positive and negative part:

$$v = v^+ - v^-$$

where

$$v^+ = \max(v, 0)$$
$$v^- = \max(-v, 0)$$
Least mixed norm solutions

With

\[ v = \alpha Rx. \]

and

\[ v = v^+ - v^- \]

the solution to

\[ \min \{ \|Ax - b\|^2_2 + \alpha \|Rx\|_1 \} \]

is equivalent to
Least mixed norm solutions

With

\[ \mathbf{v} = \alpha \mathbf{R} \mathbf{x}. \]

and

\[ \mathbf{v} = \mathbf{v}^+ - \mathbf{v}^- \]

the solution to

\[ \min \left\{ \| \mathbf{A} \mathbf{x} - \mathbf{b} \|_2^2 + \alpha \| \mathbf{R} \mathbf{x} \|_1 \right\} \]

is equivalent to

\[ \min_{\mathbf{x}^+, \mathbf{x}^-, \mathbf{v}^+, \mathbf{v}^-} \left\{ \mathbf{1}^T \mathbf{v}^+ + \mathbf{1}^T \mathbf{v}^- + \| \mathbf{A} \mathbf{x} - \mathbf{b} \|_2^2 \right\} \]

subject to

\[ \alpha \mathbf{R} \mathbf{x} = \mathbf{v}^+ - \mathbf{v}^- \]

\[ \mathbf{v}^+, \mathbf{v}^- \geq 0. \]
Least mixed norm solutions

\[
\min_{x^+, x^-, v^+, v^-} \left\{ 1^T v^+ + 1^T v^- + \|Ax - b\|_2^2 \right\}
\]

subject to

\[
\alpha R x = v^+ - v^-
\]

\[
v^+, v^- \geq 0.
\]

or
Least mixed norm solutions

\[
\min_{x^+, x^-, v^+, v^-} \left\{ 1^T v^+ + 1^T v^- + \|Ax - b\|^2_2 \right\}
\]

subject to

\[
\alpha Rx = v^+ - v^- \\
v^+, v^- \geq 0.
\]

or

\[
\min_w \left\{ \frac{1}{2} w^T G w + c^T w \right\}
\]

subject to

\[
Bw = 0 \quad \text{and} \quad Cw \geq 0.
\]

where

\[
G = \begin{bmatrix} 2A^T A & 0 \\ 0 & 0 \end{bmatrix}, \quad c = \begin{bmatrix} -2A^T b \\ 1 \\ 1 \end{bmatrix}
\]

\[
B = \begin{bmatrix} \alpha R & -I & I \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -I & -I \end{bmatrix},
\]

\[
w = \begin{bmatrix} x & v^+ & v^- \end{bmatrix}^T
\]
Setup

- $\Delta t = 0.001$
- length of the assimilation window: 100 time steps
- perfect observations, noisy observations, partial observations
Lax-Friedrich method
Singular value analysis - observations everywhere

Optimal solution (4DVar)

\[ x_0 = x_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2} u_j^T d \hat{v}_j, \quad \text{where} \quad \mu^2 = \frac{\sigma_O^2}{\sigma_B^2}. \]

Regularisation needed!
Singular value analysis - observations every 2 time steps and every 20 points in space

Optimal solution (4DVar)

\[ x_0 = x_0^B + \sum_{j} \frac{s_j^2}{\mu^2 + s_j^2} u_j^T \hat{d} v_j, \quad \text{where} \quad \mu^2 = \frac{\sigma_O^2}{\sigma_B^2}. \]
4DVar - perfect observations everywhere - Truth, Background and final solution

Figure: \( t = 0 \)

Figure: \( t = 50 \)

Figure: \( t = 100 \)

Figure: \( t = 200 \)
$L_1$ - perfect observations everywhere - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
4DVar vs $L_1$ regularisation - perfect observations everywhere

**Figure:** Root mean square error using 4DVar.

**Figure:** Root mean square error using $L_1$ regularisation.
4DVar - perfect observations every 2 time steps and every 20 points in space - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
$L_1$ - perfect observations every 2 time steps and every 20 points in space- Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
4DVar vs $L_1$ regularisation - observations every 2 time steps and every 20 points in space

Figure: Root mean square error using 4DVar.

Figure: Root mean square error using $L_1$ regularisation.
4DVar - noisy observations every 2 time steps and every 20 points in space - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
$L_1$ - noisy observations every 2 time steps and every 20 points in space-
Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
Lax-Wendroff method
Singular value analysis - observations everywhere

Optimal solution (4DVar)

\[ \mathbf{x}_0 = \mathbf{x}_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2} \frac{\mathbf{u}_j^T \hat{\mathbf{d}}}{s_j} \mathbf{v}_j, \quad \text{where} \quad \mu^2 = \frac{\sigma_O^2}{\sigma_B^2}. \]
Singular value analysis - observations every 2 time steps and every 20 points in space

Optimal solution (4DVar)

$$ \mathbf{x}_0 = \mathbf{x}_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2} \mathbf{u}_j^T \hat{d}_j \mathbf{v}_j, \quad \text{where} \quad \mu^2 = \frac{\sigma_O^2}{\sigma_B^2}. $$

Regularisation needed!
4DVar - perfect observations everywhere - Truth, Background and final solution

Figure: \(t = 0\)

Figure: \(t = 50\)

Figure: \(t = 100\)

Figure: \(t = 200\)
$L_1$ - perfect observations everywhere - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
4DVar vs $L_1$ regularisation - perfect observations everywhere

Figure: Root mean square error using 4DVar.

Figure: Root mean square error using L1 regularisation.
4DVar - perfect observations every 2 time steps and every 20 points in space - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
$L_1$ - perfect observations every 2 time steps and every 20 points in space
- Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
**4DVar vs $L_1$ regularisation** - perfect observations every 2 time steps and every 20 points in space

![Figure: Root mean square error using 4DVar.](image1)

![Figure: Root mean square error using L1 regularisation.](image2)
Example 2

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

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

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

\[
    U^n_{j+1} = U^n_j - \frac{\Delta t}{\Delta z} (U^n_j - U^n_{j-1}),
\]
\[
    U^n_0 = U^n_N,
\]

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

- $\Delta t = 0.005$
- length of the assimilation window: 40 time steps
- perfect observations, noisy observations, partial observations
4DVar - perfect observations everywhere - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
L1 - perfect observations everywhere - Truth, Background and final solution

Figure: \( t = 0 \)

Figure: \( t = 20 \)

Figure: \( t = 40 \)

Figure: \( t = 80 \)
4DVar - noisy partial observations - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
L1 - noisy partial observations - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
A different background error covariance matrix

Cost function

\[ J(x_0) = (x_0 - x_0^B)^T B^{-1}(x_0 - x_0^B) + \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1}(y_i - H_i(x_i)) \]

with

\[ B_{ij} = e^{-\frac{|i-j|}{2L^2}}, \quad L = 5 \]
4DVar - noisy partial observations - Truth, Background and final solution

**Figure: \( t = 0 \)**

**Figure: \( t = 20 \)**

**Figure: \( t = 40 \)**

**Figure: \( t = 80 \)**
L1 - noisy partial observations - Truth, Background and final solution

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
Conclusions, questions and further work

- $L_1$-norm regularisation recovers discontinuity better than 4DVar
Conclusions, questions and further work

- $L_1$-norm regularisation recovers discontinuity better than 4DVar
- $L_1$- and $L_2$-norm regularisation do equally well if no shocks/fronts are present
Conclusions, questions and further work

- $L_1$-norm regularisation recovers discontinuity better than 4DVar
- $L_1$- and $L_2$-norm regularisation do equally well if no shocks/fronts are present
- Work in progress: analysis of methods; further testing with other examples (2D, 3D, chaotic).
- multiscale methods, other regularisation approaches?