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

Melina Freitag

Department of Mathematical Sciences
University of Bath

Inverse Problems and Optimal Control for PDEs
Warwick Mathematics Institute
26th May 2011

joint work with C.J. Budd (Bath) and N.K. Nichols (Reading)
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
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
Data Assimilation in NWP

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 (previous forecast)

Observations $y_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

Melina Freitag

Resolution of sharp fronts
Data Assimilation in NWP

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

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

- background state (previous forecast)

Models

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

Observations $\mathbf{y}_i$

- Satellites
- Ships and buoys
- Surface stations
- Aeroplanes

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

Models

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

$$y_i = H_i(x_i)$$

- a model for the atmosphere (imperfect)

$$x_{i+1} = M_{i+1,i}(x_i)$$

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 (previous forecast)

Models

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

\[ y_i = H_i(x_i) \]

- a model for the atmosphere (imperfect)

\[ x_{i+1} = M_{i+1,i}(x_i) \]

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 $i = 0$)
- $x_i^A$: Analysis (estimation of the true state after the DA)
- forecast future states of the atmosphere
Observations

ECMWF Data Coverage (All obs DA) - SYNOP/SHIP
21/APR/2008; 00 UTC
Total number of obs = 28683

ECMWF Data Coverage (All obs DA) - BUOY
21/APR/2008; 00 UTC
Total number of obs = 7438

ECMWF Data Coverage (All obs DA) - AIRCRAFT
21/APR/2008; 00 UTC
Total number of obs = 51809

ECMWF Data Coverage (All obs DA) - ATOVS
21/APR/2008; 00 UTC
Total number of obs = 341239
Schematics of Data Assimilation

Figure: Background state $x^B$
Schematics of Data Assimilation

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

Under-determinacy

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

Error statistics

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

Error statistics

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

Assumptions

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

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

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

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

⇒ Three-dimensional variational data assimilation (3DVar)
Optimal least-squares estimator

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

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

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

⇒ Three-dimensional variational data assimilation (3DVar)

Interpolation equations

$$x^A = x^B + K(y - H(x^B))$$, where

$$K = BH^T (HBH^T + R)^{-1} \quad K \ldots \text{gain matrix}$$

⇒ Optimal interpolation
Four-dimensional variational assimilation (4DVar)

Minimise the cost function

\[
J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i))
\]

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

\[\text{Figure: Copyright: ECMWF}\]
Four-dimensional variational assimilation (4DVar)

Minimise the cost function

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i)) \]

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

Figure: Copyright: ECMWF
Four-dimensional variational assimilation (4DVar)

Minimise the cost function

$$J(x_0) = \frac{1}{2}(x_0 - x_B^0)^T B^{-1}(x_0 - x_B^0) + \frac{1}{2} \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1}(y_i - H_i(x_i))$$

subject to model dynamics $x_i = M_{0 \rightarrow i} x_0$.

Figure: Copyright ECMWF
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:

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

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

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 $J(x)$
Minimisation of the 4DVar cost function

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

\[
\nabla \nabla J(x_0^k) \Delta x_0^k = -\nabla J(x_0^k)
\]

\[
x_0^{k+1} = x_0^k + \Delta x_0^k
\]

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

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

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

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

$k \geq 0$

- Use approximate Hessian - Gauß-Newton method

$$
\nabla J(x_0) = B^{-1}(x_0 - x_0^B) - \sum_{i=1}^n M_{i,0}(x_0)^T H_i^T R_i^{-1}(y_i - H_i(x_i)),
$$

and

$$
\nabla \nabla J(x_0) = B^{-1} + \sum_{i=1}^n M_{i,0}(x_0)^T H_i^T R_i^{-1} H_i M_{i,0}(x_0).
$$
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

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i)) \]

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

4DVar minimises

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} \sum_{i=0}^{n} (y_i - H_i(x_i))^T R_i^{-1} (y_i - H_i(x_i)) \]

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

or

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]

where

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

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

and \( \hat{R} \) is block diagonal with \( R_i, i = 0, \ldots, n \) on the diagonal.
Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]
Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]

Gauß-Newton method

\[
\nabla \nabla J(x_0^k) \Delta x_0^k = -\nabla J(x_0^k) \\
x_0^{k+1} = x_0^k + \Delta x_0^k
\]
Relation between 4DVar and Tikhonov regularisation

Solution to the optimisation problem

Cost function

\[ J(x_0) = \frac{1}{2} (x_0 - x_0^B)^T B^{-1} (x_0 - x_0^B) + \frac{1}{2} (\hat{y} - \hat{H}(x_0))^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \]

Gauß-Newton method

\[
\begin{align*}
(B^{-1} + \hat{H}^T \hat{R}^{-1} \hat{H}) \Delta x_0^k &= -B^{-1} (x_0^k - x_0^B) + \hat{H}^T \hat{R}^{-1} (\hat{y} - \hat{H}(x_0)) \\
x_0^{k+1} &= x_0^k + \Delta x_0^k
\end{align*}
\]
Relation between 4DVar and Tikhonov regularisation

Variable transform
Set

Gauß-Newton method

\[
(B^{-1} + \hat{H}^T \hat{R}^{-1} \hat{H}) \Delta x^k_0 = -B^{-1}(x^k_0 - x^B_0) + \hat{H}^T \hat{R}^{-1}(\hat{y} - \hat{H}(x_0))
\]

\[
x^{k+1}_0 = x^k_0 + \Delta x^k_0
\]
Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

\[ B = \sigma_B^2 C_B \]
\[ \hat{R} = \sigma_R^2 C_R \]

Gauß-Newton method

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

Variable transform
Set

\[
B = \sigma_B^2 C_B \\
\hat{R} = \sigma_R^2 C_R \\
b = C_R^{-\frac{1}{2}} (\hat{y} - \hat{H}(x_0))
\]

Gauß-Newton method

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

Variable transform

Set

\[ \mathbf{B} = \sigma_B^2 \mathbf{C}_B \]
\[ \hat{\mathbf{R}} = \sigma_R^2 \mathbf{C}_R \]
\[ \mathbf{b} = \mathbf{C}_R^{-\frac{1}{2}}(\hat{\mathbf{y}} - \hat{\mathbf{H}}(\mathbf{x}_0)) \]
\[ \mathbf{A} = \mathbf{C}_R^{-\frac{1}{2}} \hat{\mathbf{H}} \mathbf{C}_B^{-\frac{1}{2}} \]

Gauß-Newton method

\[
\begin{align*}
(B^{-1} + \hat{H}^T\hat{R}^{-1}\hat{H})\Delta x_0^k &= -B^{-1}(x_0^k - x_0^B) + \hat{H}^T\hat{R}^{-1}(\hat{y} - \hat{H}(x_0)) \\
x_0^{k+1} &= x_0^k + \Delta x_0^k
\end{align*}
\]
Relation between 4DVar and Tikhonov regularisation

Variable transform

Set

\[
\begin{align*}
B &= \sigma_B^2 C_B \\
\hat{R} &= \sigma_R^2 C_R \\
b &= C_R^{-\frac{1}{2}} (\hat{y} - \hat{H}(x_0)) \\
A &= C_R^{-\frac{1}{2}} \hat{H} C_B^{\frac{1}{2}} \\
\mu^2 &= \frac{\sigma_R^2}{\sigma_B^2}
\end{align*}
\]

Gauß-Newton method

\[
\begin{align*}
(B^{-1} + \hat{H}^T \hat{R}^{-1} \hat{H}) \Delta x_0^k &= -B^{-1}(x_0^k - x_0^B) + \hat{H}^T \hat{R}^{-1}(\hat{y} - \hat{H}(x_0)) \\
x_0^{k+1} &= x_0^k + \Delta x_0^k
\end{align*}
\]
Relation between 4DVar and Tikhonov regularisation

Variable transform
Set

\[ B = \sigma_B^2 C_B \]
\[ \hat{R} = \sigma_R^2 C_R \]
\[ b = C_R^{-\frac{1}{2}} (\hat{y} - \hat{H}(x_0)) \]
\[ A = C_R^{-\frac{1}{2}} \hat{H} C_B^{\frac{1}{2}} \]
\[ \mu^2 = \frac{\sigma_R^2}{\sigma_B^2} \]

Gauß-Newton method

\[ (\mu^2 I + A^T A)C_B^{-\frac{1}{2}} \Delta x_k^0 = -\mu^2 C_B^{-\frac{1}{2}} (x_k^0 - x_B^0) + A^T b \]
\[ x_0^{k+1} = x_0^k + \Delta x_0^k \]
Relation between 4DVar and Tikhonov regularisation

Variable transform

\[ z^k = C_B^{-\frac{1}{2}}(x_0^k - x_B^0) \]

Gauß-Newton method

\[
\begin{align*}
(\mu^2 I + A^T A)C_B^{-\frac{1}{2}}\Delta x_0^k &= -\mu^2 C_B^{-\frac{1}{2}}(x_0^k - x_B^0) + A^T b \\
x_0^{k+1} &= x_0^k + \Delta x_0^k
\end{align*}
\]
Relation between 4DVar and Tikhonov regularisation

Variable transform
Set

\[ z^k = C_B^{-\frac{1}{2}} (x_0^k - x_0^B) \]

Gauß-Newton method

\[(\mu^2 I + A^T A)(z^{k+1} - z^k) = -\mu^2 z^k + A^T b\]
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
\[ z^k = C_B^{-\frac{1}{2}} (x_0^k - x_0^B) \]

Gauß-Newton method

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

Normal equations

Least squares solution

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

at each Gauß-Newton method step
Relation between 4DVar and Tikhonov regularisation

Variable transform
Set
\[ z^k = C_B^{-\frac{1}{2}}(x_0^k - x_0^B) \]

Gauß-Newton method

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

Normal equations

Least squares solution

\[ \| \begin{bmatrix} A & \mu I \end{bmatrix} (z^{k+1} - z^k) + \begin{bmatrix} b \\ \mu z^k \end{bmatrix} \|_2^2 \rightarrow \text{min} \]

at each Gauß-Newton method step or

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

Tikhonov regularisation
Ill-posed problems

Given an operator $A$ we wish to solve

$$Az = c$$

it is well-posed if
- solution exits
- solution is unique
- is stable ($A^{-1}$ continuous)
Ill-posed problems

Given an operator $A$ we wish to solve

$$Az = c$$

it is **well-posed** if

- solution exists
- 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
Ill-posed problems

Given an operator $A$ we wish to solve

$$Az = c$$

it is well-posed if

- solution exists
- 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
- Tikhonov regularization

$$z = \text{arg min} \left\{ \|Az - c\|^2 + \mu^2 \|z\|^2 \right\}$$

$$= (A^T A + \mu^2 I)^{-1} A^T c$$

$$= (V \Sigma^T U^T U \Sigma V^T + \mu^2 V V^T)^{-1} V \Sigma^T U^T c$$

$$= V \text{diag} \left( \frac{s_i^2}{s_i^2 + \mu^2} \frac{1}{s_i} \right) U^T c = z_\mu = \sum_{i=1}^{n} \frac{s_i^2}{s_i^2 + \mu^2} \frac{u_i^T c}{s_i} v_i$$
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
The blurring process as a linear model

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

$$\mathbf{r} = \text{vec}(\mathbf{X}) \in \mathbb{R}^N, \quad \mathbf{b} = \text{vec}(\mathbf{B}) \in \mathbb{R}^N$$

are related by the linear model

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

where $\mathbf{A}$ is a blurring matrix.
Blurred and exact images - Need regularisation techniques!

Standard technique: Tikhonov regularisation - Least squares

$$r_\alpha^2 = \operatorname{arg\ min} \left\{ \| A r - b \|_2^2 + \alpha \| r \|_2^2 \right\}$$
Blurred and exact images - Need regularisation techniques!

Standard technique: Tikhonov regularisation - Least squares

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

\(L_1\) regularisation

In image processing, \(L_1\)-norm regularisation provides edge preserving image deblurring!

\[
x_\alpha^1 = \arg \min \left\{ \|Ax - b\|_1 + \alpha \|x\|_2^2 \right\}
\]
Results from image deblurring: $L_1$ regularisation

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

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

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

In image processing, \textit{L}_1\text{-norm regularisation provides edge preserving image deblurring!}

- \textit{L}_1\text{-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!
3 Regularisation Methods

4DVar

$$\min_{z^{k+1}} \|Az^{k+1} - c\|_2^2 + \mu^2 \|z^{k+1}\|_2^2$$
3 Regularisation Methods

4DVar

\[
\min_{z^{k+1}} \| Az^{k+1} - c \|_2^2 + \mu^2 \| z^{k+1} \|_2^2
\]

\[\text{\textbf{L}_1\textbf{-norm regularisation}}\]

\[
\min_{z^{k+1}} \| Az^{k+1} - c \|_2^2 + \mu^2 \| z^{k+1} \|_1
\]
3 Regularisation Methods

4DVar

\[
\min_{z^{k+1}} \| Az^{k+1} - c \|^2_2 + \mu^2 \| z^{k+1} \|^2_2
\]

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

\[
\min_{z^{k+1}} \| Az^{k+1} - c \|^2_2 + \mu^2 \| z^{k+1} \|_1
\]

\[\text{Total Variation regularisation}\]

\[
\min_{z^{k+1}} \| Az^{k+1} - c \|^2_2 + \mu^2 \| z^{k+1} \|^2_2 + \beta \| Dx^{k+1}_0 \|_1
\]

where \(x^{k+1}_0 = C_B^{\frac{1}{2}} z^{k+1} + x^B_0\) and \(D\) is a matrix approximating the derivative of the solution.
Least mixed norm solutions

Solve

\[
\min_{z^{k+1}} \|Az^{k+1} - c\|^2_2 + \mu^2\|z^{k+1}\|^2_2
\]

using Least squares and

\[
\min_{z^{k+1}} \|Az^{k+1} - c\|^2_2 + \mu^2\|z^{k+1}\|_1
\]

or

\[
\min_{z^{k+1}} \|Az^{k+1} - c\|^2_2 + \mu^2\|z^{k+1}\|^2_2 + \beta\|Dx_0^{k+1}\|_1
\]

using quadratic programming (see Fu/Ng/Nikolova/Barlow 2006).
Least mixed norm solutions

Consider

$$\min_{z^{k+1}} ||Az^{k+1} - c||_2^2 + \beta ||Dx_0^{k+1}||_1$$

where $x_0^{k+1} = C_B^\frac{1}{2} z^{k+1} + x_0^B$
Least mixed norm solutions

Consider

$$\min_{z^{k+1}} \| A z^{k+1} - c \|_2^2 + \beta \| D x_0^{k+1} \|_1$$

where $x_0^{k+1} = C_B^{\frac{1}{2}} z^{k+1} + x_0^B$

$$\min_{z^{k+1}} \| A z^{k+1} - c \|_2^2 + \beta \| D C_B^{\frac{1}{2}} z^{k+1} + D x_0^B \|_1$$
Least mixed norm solutions

Consider

$$\min_{z^{k+1}} \|Az^{k+1} - c\|_2^2 + \beta \|Dx_0^{k+1}\|_1$$

where $$x_0^{k+1} = C_B^{\frac{1}{2}} z^{k+1} + x_0^B$$

$$\min_{z^{k+1}} \|Az^{k+1} - c\|_2^2 + \beta \|DC_B^{\frac{1}{2}} z^{k+1} + Dx_0^B\|_1$$

Set

$$v = \beta DC_B^{\frac{1}{2}} z^{k+1} + \beta Dx_0^B.$$ 

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

\[ v = \beta DC_B^{\frac{1}{2}} z^{k+1} + \beta Dx_0^B \]

and

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

the solution to

\[ \min_{z^{k+1}} \|Az^{k+1} - c\|_2^2 + \beta \|DC_B^{\frac{1}{2}} z^{k+1} + Dx_0^B\|_1 \]

is equivalent to
Least mixed norm solutions

With

\[ v = \beta DC_B^{1/2} z^{k+1} + \beta Dx_0^B \]

and

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

the solution to

\[
\min_{z^{k+1}} \|Az^{k+1} - c\|_2^2 + \beta \|DC_B^{1/2} z^{k+1} + Dx_0^B \|_1
\]

is equivalent to

\[
\min_{z^{k+1}, v^+, v^-} \left\{ 1^Tv^+ + 1^Tv^- + \|Az^{k+1} - c\|_2^2 \right\}
\]

subject to

\[
\beta DC_B^{1/2} z^{k+1} + \beta Dx_0^B = v^+ - v^-
\]

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

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

subject to

$$\beta D C^{\frac{1}{2}} B z^{k+1} + \beta D x_0^B = v^+ - v^-$$

$$v^+, v^- \geq 0.$$
Least mixed norm solutions

\[
\min_{z^{k+1}, v^+, v^-} \left\{ 1^T v^+ + 1^T v^- + \|Az^{k+1} - c\|_2^2 \right\}
\]

subject to

\[
\beta DC_B^{1/2} z^{k+1} + \beta D x_0^B = v^+ - v^- \\
v^+, v^- \geq 0.
\]

or

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

subject to

\[
E w = e \quad \text{and} \quad F w \geq 0.
\]

where

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

\[
E = \begin{bmatrix}
\beta DC_B^{1/2} & -I & 1
\end{bmatrix}, \quad w = \begin{bmatrix}
z^{k+1} & v^+ & v^-
\end{bmatrix}^T, \quad e = -\beta D x_0^B
\]
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
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

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

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

\[ U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta z} (U_j^n - U_{j-1}^n), \]

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
- $R = 0.01$.
- $B = I$ and $B = 0.1e^{-\frac{|i-j|}{2L^2}}$, where $L = 5$
Setup

- length of the assimilation window: 40 time steps
- perfect observations, noisy and sparse observations
- $R = 0.01$.
- $B = I$ and $B = 0.1e^{-\frac{|i-j|}{2L^2}}$, where $L = 5$
- use MATLAB quadprog.m
4DVar - perfect and full observations, \( B = I \)

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

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

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

**Figure: \( t = 80 \)**
L1 on the background term - perfect and full observations, \( B = I \)

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

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

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

**Figure:** \( t = 80 \)
L1 - perfect and full observations, \( B = I \)

Figure: \( t = 0 \)

Figure: \( t = 20 \)

Figure: \( t = 40 \)

Figure: \( t = 80 \)
4DVar - noisy and sparse observations, $B = I$

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
L1 - noisy and sparse observations, $B = I$

**Figure: $t = 0$**

**Figure: $t = 20$**

**Figure: $t = 40$**

**Figure: $t = 80$**
4DVar - perfect and full observations, \( \mathbf{B} = 0.1e^{-\frac{|i-j|}{2L^2}} \)

Figure: \( t = 0 \)

Figure: \( t = 20 \)

Figure: \( t = 40 \)

Figure: \( t = 80 \)
L1 - perfect and full observations, $B = 0.1e^{-\frac{|i-j|}{2L^2}}$

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
4DVar - noisy and sparse observations, $B = 0.1e^{-\frac{|i-j|}{2L^2}}$

Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$
L₁-norm regularisation in 4DVar

Examples

$B = 0.1e^{-\frac{|i-j|}{2L^2}}$

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)

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

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

\[ U_{j+1}^n = U_j^n - \frac{\Delta t}{2\Delta x} (f(U_{j+1}^n) - f(U_{j-1}^n)) + \frac{\Delta t^2}{2\Delta x^2} \left( A_{j+\frac{1}{2}} (f(U_{j+1}^n) - f(U_{j}^n)) - A_{j-\frac{1}{2}} (f(U_{j}^n) - f(U_{j-1}^n)) \right) \]
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
- \( \mathbf{R} = 0.01 \).
- \( \mathbf{B} = 0.1 e^{-\frac{|i-j|}{2L^2}} \), where \( L = 5 \)
Setup

- length of the assimilation window: 100 time steps
- noisy and sparse observations
- $R = 0.01$
- $B = 0.1e^{-\frac{|i-j|}{2L^2}}$, where $L = 5$
- use MATLAB `quadprog.m`
Lax-Friedrichs method
Optimal solution (4DVar)

\[ x_0 = x_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2} \mathbf{u}_j^T \hat{\mathbf{c}} \mathbf{v}_j, \text{ where } \mu^2 = \frac{\sigma_O^2}{\sigma_B^2}. \]
Singular value analysis - observations every 2 time steps and every 20 points in space

Optimal solution (4DVar)

\[ x_0 = x_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2} \frac{u_j^T \hat{c}}{s_j} v_j, \quad \text{where} \quad \mu^2 = \frac{\sigma_Q^2}{\sigma_B^2}. \]
4DVar - noisy and sparse observations, $B = 0.1e^{-\frac{|i-j|}{2L^2}}$

Figure: $t = 0$

Figure: $t = 50$

Figure: $t = 100$

Figure: $t = 200$
L1 - noisy and sparse observations, $B = 0.1e^{\frac{|i-j|}{2L^2}}$

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