Tikhonov Regularisation for (Large) Inverse Problems

Melina Freitag

Department of Mathematical Sciences
University of Bath

17th ILAS Conference
Braunschweig, Germany
23rd August 2011

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

Outline

Ill-posed Problems

Given an operator A we wish to solve

$$Af = g.$$

It is well-posed if

Ill-posed Problems

Given an operator A we wish to solve

$$Af = g.$$

It is well-posed if

- solution exits

Ill-posed Problems

Given an operator A we wish to solve

$$Af = g.$$

It is well-posed if

- solution exits
- solution is unique

Ill-posed Problems

Given an operator A we wish to solve

$$Af = g.$$

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

$$Af = g.$$

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

An Illustrative Example

Fredholm first kind integral equation in 1D

$$g(x) = \int_0^1 k(x - x')f(x')dx' =: (Af)(x), \quad 0 < x < 1$$

- f light source intensity as a function of x
- g image intensity
- k kernel representing blurring effects, e.g. $k(x) = C \exp\left(-\frac{x^2}{2\gamma^2}\right)$, C, γ are positive parameters.

An Illustrative Example

Fredholm first kind integral equation in 1D

$$g(x) = \int_0^1 k(x-x')f(x')dx' =: (Af)(x), \quad 0 < x < 1$$

- f light source intensity as a function of x
- g image intensity
- k kernel representing blurring effects, e.g. $k(x) = C\exp\left(-\frac{x^2}{2\gamma^2}\right)$, C, γ are positive parameters.

Discretisation

- use a piecewise smooth source $\mathbf{f}$
- determine $\mathbf{A}$ using standard numerical quadrature;

$$(\mathbf{A})_{ij} = hC\exp\left(-\frac{((i-j)h)^2}{2\gamma^2}\right), \quad 1 \le i,j \le n, \quad h = \frac{1}{n}$$

$\gamma = 0.05$, $C = \frac{1}{\gamma\sqrt{2\pi}}$.

Forward Problem

Given f and the kernel k, determine the blurred image $g = Af$, or the discrete version

$$\mathbf{g} = \mathbf{A}\mathbf{f}.$$

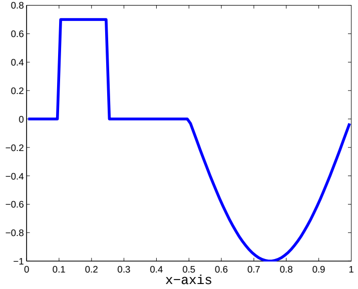


Figure: True solution of source function $\mathbf{f}$

Forward Problem

Given f and the kernel k, determine the blurred image $g = Af$, or the discrete version

$$\mathbf{g} = \mathbf{A}\mathbf{f}.$$

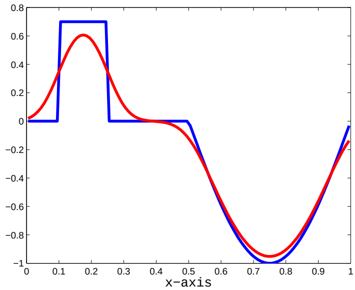


Figure: True solution **f** and blurred image **g**

Inverse Problem

Given the kernel k, and the blurred image g, determine the source f from $g = Af$, solve the discrete linear system

$$\mathbf{g} = \mathbf{A}\mathbf{f}.$$

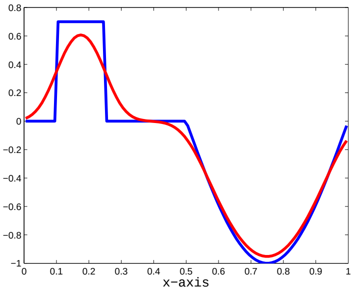


Figure: True solution and blurred image $\mathbf{g}$

Inverse Problem

Given the kernel k, and the blurred image g, determine the source f from $g = Af$, solve the discrete linear system

$$\mathbf{g} = \mathbf{A}\mathbf{f}.$$

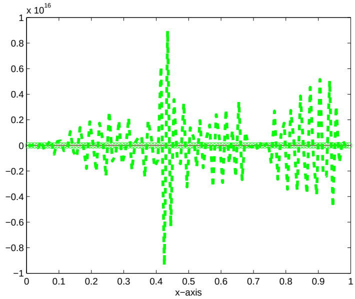


Figure: Naive Solution

Inverse Problem

Problem: data $\mathbf{g}$ are observed and contain noise and $\mathbf{A}$ is ill-conditioned:

$$\mathbf{g}_{\text{exact}} + \mathbf{e} = \mathbf{A}\mathbf{f},$$

$\mathbf{e}$ is unknown white noise.

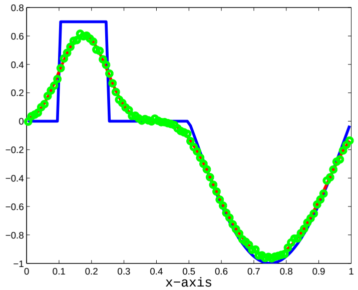


Figure: True solution and discrete noisy data

Inverse Problem

Problem: data $\mathbf{g}$ are observed and contain noise and $\mathbf{A}$ is ill-conditioned:

$$\mathbf{g}_{\text{exact}} + \mathbf{e} = \mathbf{A}\mathbf{f},$$

$\mathbf{e}$ is unknown white noise.

Singular Value Decomposition

Let

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

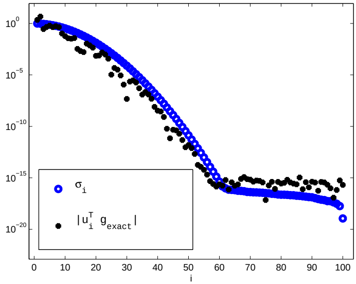
where

- $\mathbf{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_r)$ and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$
- $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$

Inverse Problem - Regularisation needed

Least squares solution (with and without noise)

$$\mathbf{f}_{\text{exact}} = \mathbf{A}^{\dagger}\mathbf{g}_{\text{exact}} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{g}_{\text{exact}}}{\sigma_i} \mathbf{v}_i$$



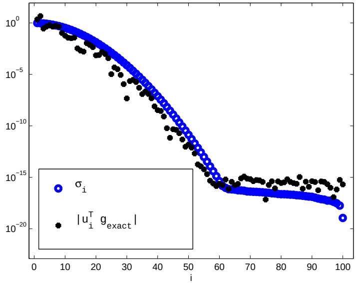
Inverse Problem - Regularisation needed

Least squares solution (with and without noise)

$$\mathbf{f}_{\text{exact}} = \mathbf{A}^{\dagger}\mathbf{g}_{\text{exact}} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{g}_{\text{exact}}}{\sigma_i}\mathbf{v}_i$$

$$\mathbf{f} = \mathbf{A}^{\dagger}\mathbf{g} = \mathbf{A}^{\dagger}(\mathbf{g}_{\text{exact}} + \mathbf{e}) = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{g}_{\text{exact}}}{\sigma_i}\mathbf{v}_i + \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i}\mathbf{v}_i$$

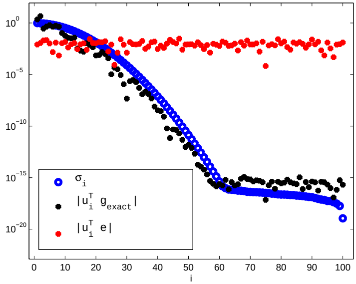
$$= \mathbf{f}_{\text{exact}} + \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i}\mathbf{v}_i$$



Inverse Problem - Regularisation needed

Least squares solution (with and without noise)

$$
\begin{aligned}
\mathbf{f}_{\text{exact}} &= \mathbf{A}^{\dagger}\mathbf{g}_{\text{exact}} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{g}_{\text{exact}}}{\sigma_i}\mathbf{v}_i \\
\mathbf{f} &= \mathbf{A}^{\dagger}\mathbf{g} = \mathbf{A}^{\dagger}(\mathbf{g}_{\text{exact}} + \mathbf{e}) = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{g}_{\text{exact}}}{\sigma_i}\mathbf{v}_i + \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i}\mathbf{v}_i \\
&= \mathbf{f}_{\text{exact}} + \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i}\mathbf{v}_i
\end{aligned}
$$



Tikhonov Regularisation

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

α regularisation parameter.

Tikhonov Regularisation

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

α regularisation parameter.
Solution $\mathbf{f}_\alpha$ to the minimisation problem

$$\min_{\mathbf{f}} \left\{ \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \alpha^2 \|\mathbf{f}\|_2^2 \right\}.$$

Tikhonov Regularisation

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

α regularisation parameter.
Solution $\mathbf{f}_\alpha$ to the minimisation problem

$$\min_{\mathbf{f}} \left\{ \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \alpha^2 \|\mathbf{f}\|_2^2 \right\}.$$

Least squares solution $\mathbf{f}_\alpha$ to the linear system

$$\begin{bmatrix} \mathbf{A} \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix}.$$

Tikhonov Regularisation

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

α regularisation parameter.
Solution $\mathbf{f}_\alpha$ to the minimisation problem

$$\min_{\mathbf{f}} \left\{ \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \alpha^2 \|\mathbf{f}\|_2^2 \right\}.$$

Least squares solution $\mathbf{f}_\alpha$ to the linear system

$$\begin{bmatrix} \mathbf{A} \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix}.$$

Normal equations

$$(\mathbf{A}^T\mathbf{A} + \alpha^2 \mathbf{I})\mathbf{f}_\alpha = \mathbf{A}^T\mathbf{g}.$$

Tikhonov Regularisation

Regularisation parameter α

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

Filter factor $\dfrac{\sigma_i^2}{\sigma_i^2 + \alpha^2}$ as diagonal entries of the filter matrix $\boldsymbol{\Psi}$

Tikhonov Regularisation

Regularisation parameter α

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

Filter factor $\dfrac{\sigma_i^2}{\sigma_i^2 + \alpha^2}$ as diagonal entries of the filter matrix $\mathbf{\Psi}$

Regularisation and perturbation error

$$\begin{aligned}
\mathbf{f}_\alpha &= \mathbf{V\Psi U}^T\mathbf{g}, \quad \mathbf{g} = \mathbf{g}_{\text{exact}} + \mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{g}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{U\Sigma V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e}
\end{aligned}$$

Tikhonov Regularisation

Regularisation parameter α

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

Filter factor $\dfrac{\sigma_i^2}{\sigma_i^2 + \alpha^2}$ as diagonal entries of the filter matrix $\mathbf{\Psi}$

Regularisation and perturbation error

$$\begin{aligned}
\mathbf{f}_\alpha &= \mathbf{V\Psi U}^T\mathbf{g}, \quad \mathbf{g} = \mathbf{g}_{\text{exact}} + \mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{g}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{U\Sigma V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
\mathbf{f}_{\text{exact}} - \mathbf{f}_\alpha &=
\end{aligned}$$

Tikhonov Regularisation

Regularisation parameter α

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

Filter factor $\dfrac{\sigma_i^2}{\sigma_i^2 + \alpha^2}$ as diagonal entries of the filter matrix $\mathbf{\Psi}$

Regularisation and perturbation error

$$\begin{aligned}
\mathbf{f}_\alpha &= \mathbf{V\Psi U}^T\mathbf{g}, \quad \mathbf{g} = \mathbf{g}_{\text{exact}} + \mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{g}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{U\Sigma V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
\mathbf{f}_{\text{exact}} - \mathbf{f}_\alpha &= (\mathbf{I} - \mathbf{V\Psi V}^T)\mathbf{f}_{\text{exact}} -
\end{aligned}$$

Regularisation error

Tikhonov Regularisation

Regularisation parameter α

Regularised solution of the form

$$\mathbf{f}_\alpha = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \frac{\mathbf{u}_i^T \mathbf{g}}{\sigma_i} \mathbf{v}_i$$

Filter factor $\dfrac{\sigma_i^2}{\sigma_i^2 + \alpha^2}$ as diagonal entries of the filter matrix $\mathbf{\Psi}$

Regularisation and perturbation error

$$\begin{aligned}
\mathbf{f}_\alpha &= \mathbf{V\Psi U}^T\mathbf{g}, \quad \mathbf{g} = \mathbf{g}_{\text{exact}} + \mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{g}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{U\Sigma V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
&= \mathbf{V\Psi V}^T\mathbf{f}_{\text{exact}} + \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e} \\
\mathbf{f}_{\text{exact}} - \mathbf{f}_\alpha &= (\mathbf{I} - \mathbf{V\Psi V}^T)\mathbf{f}_{\text{exact}} - \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e}
\end{aligned}$$

Regularisation error Perturbation error

Tikhonov Regularisation

Regularisation and perturbation error

$$\mathbf{f}_{\text{exact}} - \mathbf{f}_\alpha \quad = \quad (\mathbf{I} - \mathbf{V\Psi V}^T)\mathbf{f}_{\text{exact}} - \mathbf{V\Psi\Sigma}^{-1}\mathbf{U}^T\mathbf{e}$$

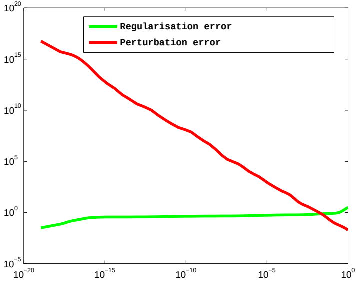


Figure: Regularisation and perturbation error

Tikhonov Regularisation

Illustrative example

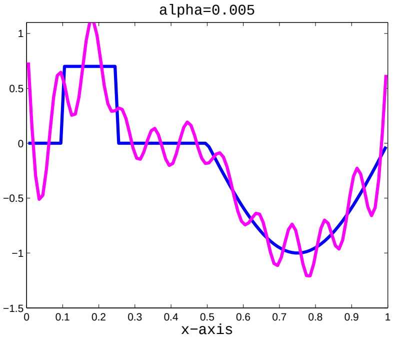


Figure: α too small

Tikhonov Regularisation

Illustrative example

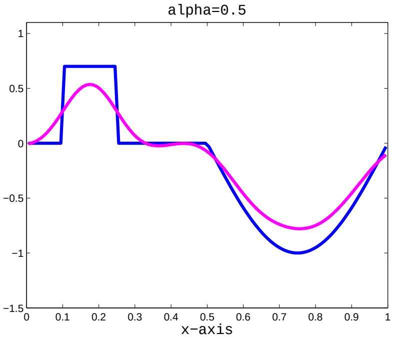


Figure: α too large

Tikhonov Regularisation

Illustrative example

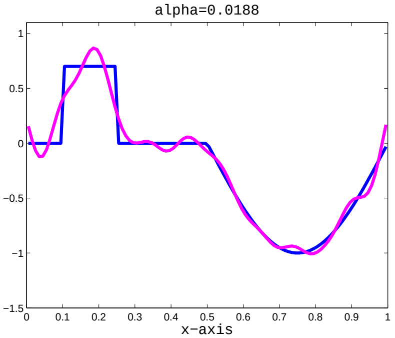


Figure: Good Value for α

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
- Planes

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
- Planes

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
- Planes

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
- Planes

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
- Planes

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

Schematics of Data Assimilation

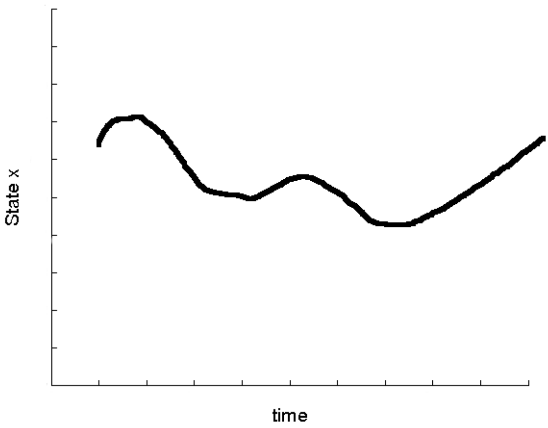


Figure: Background state $\mathbf{x}^B$

Schematics of Data Assimilation

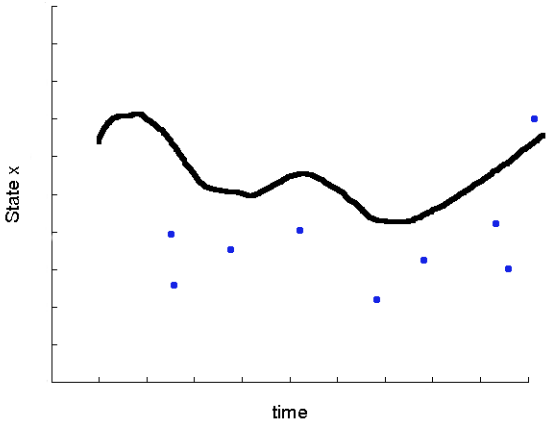


Figure: Observations $\mathbf{y}$

Schematics of Data Assimilation

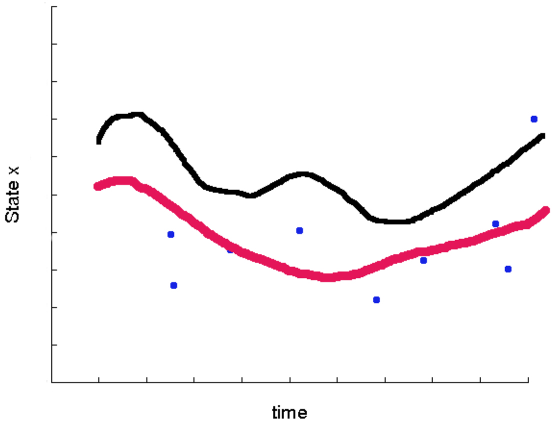
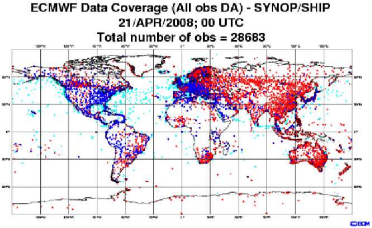
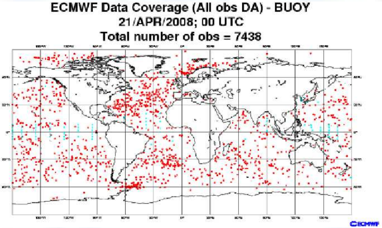
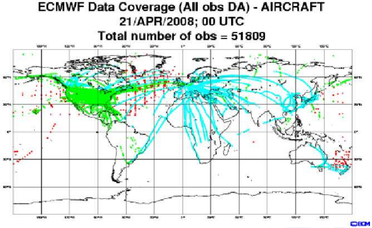


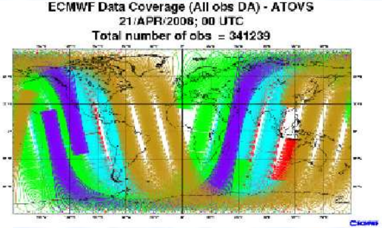
Figure: Analysis $\mathbf{x}^A$ (consistent with observations and model dynamics)

Observations









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
- 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

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

$\Rightarrow$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

Bayesian interpretation

Non-Gaußian PDF's (probability density function)

- $P(\mathbf{x})$ is a priori PDF (background)
- $P(\mathbf{y}|\mathbf{x})$ is the observation PDF (likelihood of the observations given background $\mathbf{x}$)

Bayesian interpretation

Non-Gaußian PDF's (probability density function)

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

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

Bayesian interpretation

Non-Gaußian PDF's (probability density function)

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

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

Gaußian PDF's

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

$\mathbf{x}^A$ is the maximum a posteriori estimator of $\mathbf{x}^{\text{Truth}}$.

Bayesian interpretation

Non-Gaußian PDF's (probability density function)

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

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

Gaußian PDF's

$$\begin{aligned} P(\mathbf{x}|\mathbf{y}) \quad &= \quad c_1 \exp\left(-(\mathbf{x}-\mathbf{x}^B)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^B)\right) \cdot \\ &\qquad c_2 \exp\left(-(\mathbf{y}-H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}))\right) \end{aligned}$$

$\mathbf{x}^A$ is the maximum a posteriori estimator of $\mathbf{x}^{\text{Truth}}$.
Maximising $P(\mathbf{x}|\mathbf{y})$ equivalent to minimising $J(\mathbf{x})$

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 $\mathbf{x}_i = M_{i,0}\mathbf{x}_0$.

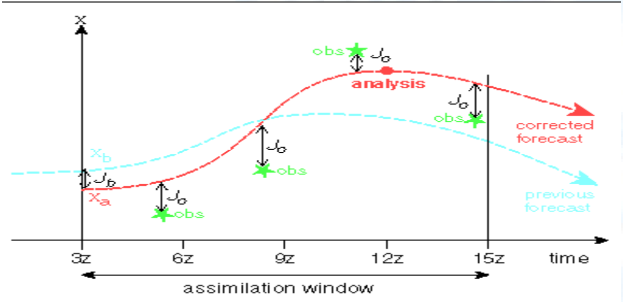


Figure: Copyright:ECMWF

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 $\mathbf{x}_i = M_{i,0}\mathbf{x}_0$.

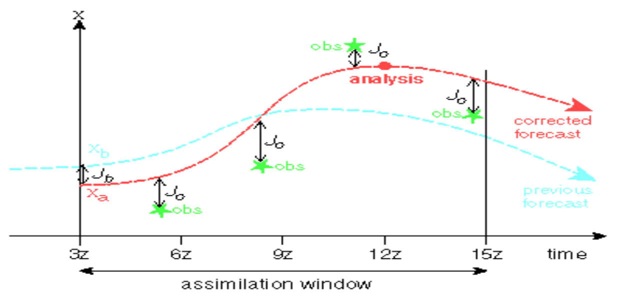


Figure: Copyright:ECMWF

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 $\mathbf{x}_i = M_{i,0}\mathbf{x}_0$.

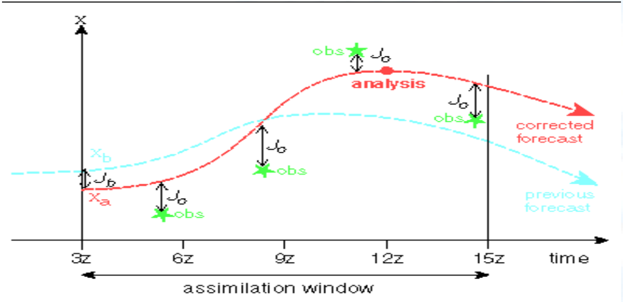


Figure: Copyright:ECMWF

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

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

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 $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$

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 $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$

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

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

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

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

$$\begin{aligned}\mathbf{B} &= \sigma_B^2 \mathbf{C}_B \\ \hat{\mathbf{R}} &= \sigma_R^2 \mathbf{C}_R\end{aligned}$$

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

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

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

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

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

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

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

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

Gauß-Newton method

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

Relation between 4DVar and Tikhonov regularisation

Variable transform

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

Gauß-Newton method

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

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

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

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

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

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

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

Normal equations

Least squares solution

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

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

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

Normal equations

Least squares solution

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

at each Gauß-Newton method step or

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

Tikhonov regularisation

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

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

Normal equations

Least squares solution

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

at each Gauß-Newton method step or

$$\|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{g}\|_2^2 + \alpha^2\|\mathbf{z}^{k+1}\|_2^2$$

Tikhonov regularisation

Summary

Minimising the cost function

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

amounts to solving a Tikhonov regularised least squares problem at every step

$$\|\mathbf{A}\mathbf{z}^{k+1} - \mathbf{g}\|_2^2 + \alpha^2\|\mathbf{z}^{k+1}\|_2^2$$

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

- Low Rank Kalman Filters [Houtekamer, Mitchell 1998]

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

- Low Rank Kalman Filters [Houtekamer, Mitchell 1998]
- Model Order Reduction [Boess, Lawless, Nichols, Bunse-Gerstner 2008, 2010]

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

- Low Rank Kalman Filters [Houtekamer, Mitchell 1998]
- Model Order Reduction [Boess, Lawless, Nichols, Bunse-Gerstner 2008, 2010]
- Preconditioners [Haben, Lawless, Nichols 2010]

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

- Low Rank Kalman Filters [Houtekamer, Mitchell 1998]
- Model Order Reduction [Boess, Lawless, Nichols, Bunse-Gerstner 2008, 2010]
- Preconditioners [Haben, Lawless, Nichols 2010]
- Regularisation Parameter Estimation

Data Assimilation and Tikhonov regularisation

Issues

- Dynamic vs static
- Nonlinear Dynamics (and chaotic)
- Multiscale and Large Scale
- Many unknown parameters (B ...)
- Model error

Previous/current work

- Low Rank Kalman Filters [Houtekamer, Mitchell 1998]
- Model Order Reduction [Boess, Lawless, Nichols, Bunse-Gerstner 2008, 2010]
- Preconditioners [Haben, Lawless, Nichols 2010]
- Regularisation Parameter Estimation
- L_1-Norm Regularisation

Outline

Inverse Problems

Data Assimilation as a Large Inverse Problem

Regularisation Parameter estimation in 4DVar
- Regularisation Parameter estimation
- Example

Application of L_1-norm regularisation in 4DVar
- Motivation: Results from image processing
- L_1-norm regularisation in 4DVar
- Examples

Choosing the regularisation parameter α

$$\|\mathbf{A}\mathbf{f} - \mathbf{g}\|_2^2 + \alpha^2 \|\mathbf{f}\|_2^2$$

Generalised Cross-Validation (Golub, Heath, Wahba 1979) GCV

- If a data value is omitted, then a good choice of the reconstruction should be able to predict the missing data value as well

Generalised Cross-Validation (Golub, Heath, Wahba 1979) GCV

- If a data value is omitted, then a good choice of the reconstruction should be able to predict the missing data value as well
- Minimise the GCV functional

$$G(\alpha) = \frac{\|(\mathbf{I} - \mathbf{A}\mathbf{V}\mathbf{\Psi}\mathbf{\Sigma}^{-1}\mathbf{U}^T)\mathbf{g}\|^2}{(\mathrm{trace}(\mathbf{I} - \mathbf{A}\mathbf{V}\mathbf{\Psi}\mathbf{\Sigma}^{-1}\mathbf{U}^T))^2}$$

Generalised Cross-Validation (Golub, Heath, Wahba 1979) GCV

- If a data value is omitted, then a good choice of the reconstruction should be able to predict the missing data value as well
- Minimise the GCV functional

$$G(\alpha) = \frac{\|(\mathbf{I} - \mathbf{AV\Psi\Sigma}^{-1}\mathbf{U}^T)\mathbf{g}\|^2}{(\mathrm{trace}(\mathbf{I} - \mathbf{AV\Psi\Sigma}^{-1}\mathbf{U}^T))^2}$$

$$G(\alpha) = \frac{\sum_{i=1}^{N}\left(\frac{\mathbf{u}_i^T\mathbf{g}}{\sigma_i^2+\alpha^2}\right)^2}{\left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\alpha^2}\right)^2}$$

Generalised Cross-Validation (Golub, Heath, Wahba 1979) GCV

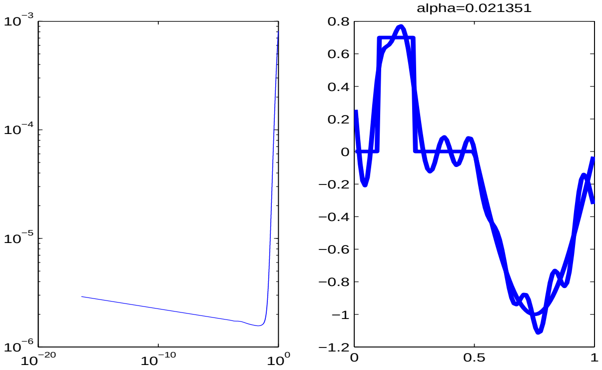


Figure: Parameter estimation using GCV

L-Curve Criterion (Hansen 1992)

- Log-log plot of the norm of the regularised solution $\|\mathbf{f}\|$ versus the corresponding residual norm $\|\mathbf{A}\mathbf{f} - \mathbf{g}\|$

L-Curve Criterion (Hansen 1992)

- Log-log plot of the norm of the regularised solution $\|\mathbf{f}\|$ versus the corresponding residual norm $\|\mathbf{A}\mathbf{f} - \mathbf{g}\|$
- Best value of α determined by maximum curvature

$$R(\alpha) = \log \|\mathbf{A}\mathbf{f}_\alpha - \mathbf{g}\|^2 \quad S(\alpha) = \log \|\mathbf{f}_\alpha\|^2$$

$$k(\alpha) = \frac{R''(\alpha)S'(\alpha) - R'(\alpha)S''(\alpha)}{(R'(\alpha)^2 + S'(\alpha))^{3/2}}$$

L-Curve Criterion (Hansen 1992)

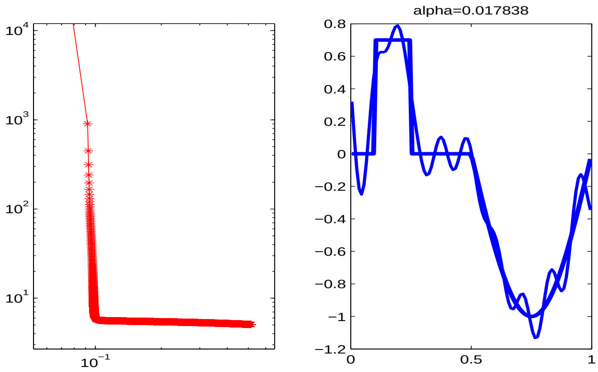


Figure: Parameter estimation using L-Curve

Discrepancy Principle (Morozov 1966)

- Choose a regularised solution such that

$$\|\mathbf{g} - \mathbf{A}\mathbf{f}_\alpha\|_2 = \tau\delta$$

where $2 \leq \tau \leq 5$

Discrepancy Principle (Morozov 1966)

- Choose a regularised solution such that

$$\|\mathbf{g} - \mathbf{A}\mathbf{f}_\alpha\|_2 = \tau\delta$$

  where $2 \leq \tau \leq 5$
- δ is the expected value of the error $\|\mathbf{e}\|$

Discrepancy Principle (Morozov 1966)

- Choose a regularised solution such that

$$\|\mathbf{g} - \mathbf{A}\mathbf{f}_\alpha\|_2 = \tau\delta$$

  where $2 \leq \tau \leq 5$
- δ is the expected value of the error $\|\mathbf{e}\|$
- Apply iterative method to $\mathbf{g} = \mathbf{A}\mathbf{f}$

Discrepancy Principle (Morozov 1966)

- Choose a regularised solution such that

$$\|\mathbf{g} - \mathbf{A}\mathbf{f}_\alpha\|_2 = \tau\delta$$

  where $2 \leq \tau \leq 5$
- δ is the expected value of the error $\|\mathbf{e}\|$
- Apply iterative method to $\mathbf{g} = \mathbf{A}\mathbf{f}$
- First steps: reduce the residual error in the singular direction associated with larger singular values

Discrepancy Principle (Morozov 1966)

- Choose a regularised solution such that

$$\|\mathbf{g} - \mathbf{A}\mathbf{f}_\alpha\|_2 = \tau\delta$$

  where $2 \leq \tau \leq 5$
- δ is the expected value of the error $\|\mathbf{e}\|$
- Apply iterative method to $\mathbf{g} = \mathbf{A}\mathbf{f}$
- First steps: reduce the residual error in the singular direction associated with larger singular values
- Latter steps: singular direction associated to smaller singular values are fitted - truncate the iteration before the amplified noise takes over

Discrepancy Principle (Morozov 1966)

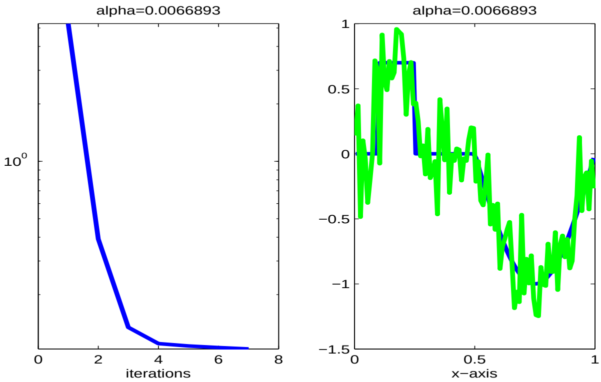


Figure: Parameter estimation using Discrepancy Principle

N-dimensional (chaotic) Lorenz system (Lorenz-95)

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \ldots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.

- $F = 8$, $N = 40$ (13 positive Lyapunov exponents).
- solver: Runge-Kutta method with time step $h = 0.01$

N-dimensional (chaotic) Lorenz system (Lorenz-95)

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \ldots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.

- $F = 8$, $N = 40$ (13 positive Lyapunov exponents).
- solver: Runge-Kutta method with time step $h = 0.01$
- a unit time $T = 1$ is associated with 5 days
- assimilation window: 5 time steps (associated with 6 hours)
- subsequent forecast: 95 time steps (associated with 5 day forecast)

N-dimensional (chaotic) Lorenz system (Lorenz-95)

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \ldots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.

- $F = 8$, $N = 40$ (13 positive Lyapunov exponents).
- solver: Runge-Kutta method with time step $h = 0.01$
- a unit time $T = 1$ is associated with 5 days
- assimilation window: 5 time steps (associated with 6 hours)
- subsequent forecast: 95 time steps (associated with 5 day forecast)
- observations are taken as noise from the truth trajectory

N-dimensional (chaotic) Lorenz system (Lorenz-95)

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \ldots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.

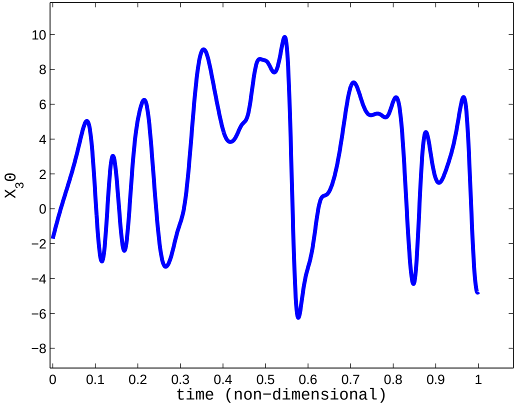
- $F = 8$, $N = 40$ (13 positive Lyapunov exponents).
- solver: Runge-Kutta method with time step $h = 0.01$
- a unit time $T = 1$ is associated with 5 days
- assimilation window: 5 time steps (associated with 6 hours)
- subsequent forecast: 95 time steps (associated with 5 day forecast)
- observations are taken as noise from the truth trajectory
- Model error introduced by parameter change $F_{mod} = 12$.

Lorenz-95 dynamics

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \ldots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.

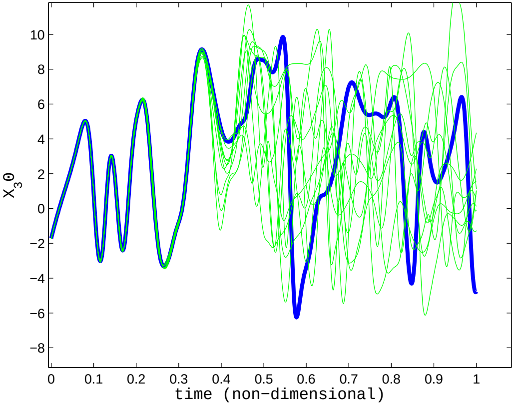


Lorenz-95 dynamics

The system is given by

$$\frac{dX_i}{dt} = -X_{i-2}X_{i-1} + X_{i-1}X_{i+1} - X_i + F, \quad i = 1, \dots, N,$$

cyclic boundary conditions $X_0 = X_N$, $X_{-1} = X_{N+1}$, $X_{N+1} = X_1$.



Initial condition error

Observation frequency	4DVAR	Discrepancy Principle	GCV	L-Curve
every 10 points	0.7729	0.7608	0.7394	0.8101
every 5 points	0.8043	0.6725	0.6510	0.7727
every 2 points	0.5492	0.3309	0.2812	0.4469

Table: Comparison RMS error - no model error in the Lorenz system

Comparison - no model error in the Lorenz system

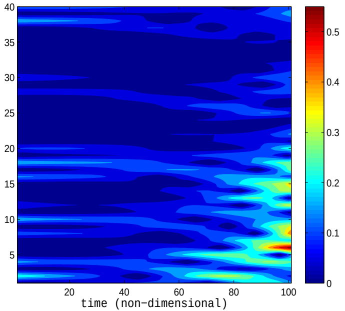


Figure: 4DVAR

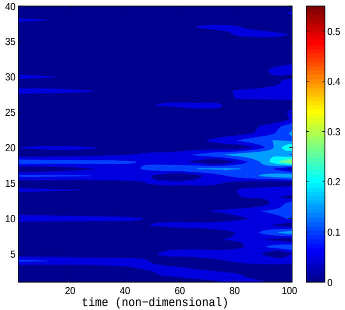


Figure: Generalised Cross-Validation

Initial condition error

Observation frequency	4DVAR	Discrepancy Principle	GCV	L-Curve
every 10 points	3.4641	3.4156	6.1941	0.8579
every 5 points	5.3430	4.4666	6.0010	0.8651
every 2 points	26.5536	5.8955	11.0836	0.7630

Table: Comparison RMS error - with model error in the Lorenz system

Comparison - with model error in the Lorenz system

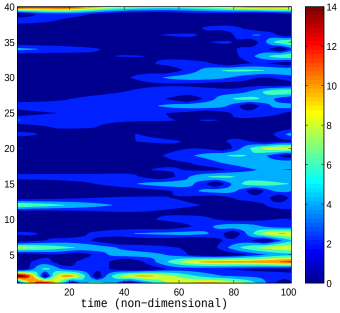


Figure: 4DVAR

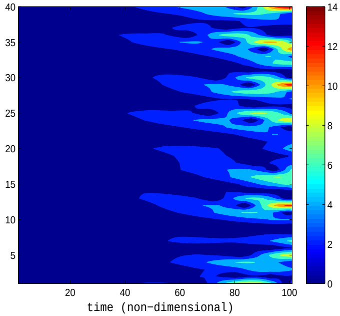


Figure: L-Curve Criterion

Outline

Results from image deblurring: L_1 regularisation

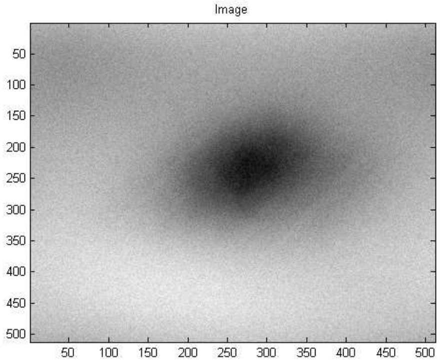


Figure: Blurred picture

Results from image deblurring: L_1 regularisation

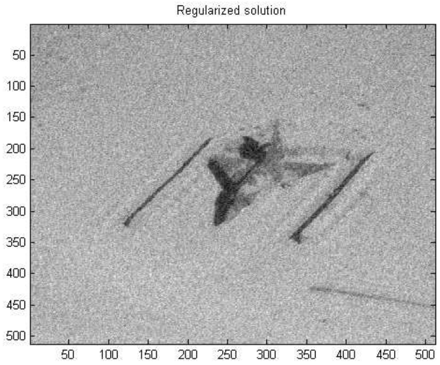


Figure: Tikhonov regularisation $\min\left\{\|\mathfrak{A}\mathfrak{x}-\mathfrak{b}\|_2^2+\alpha\|\mathfrak{x}\|_2^2\right\}$

Results from image deblurring: L_1 regularisation

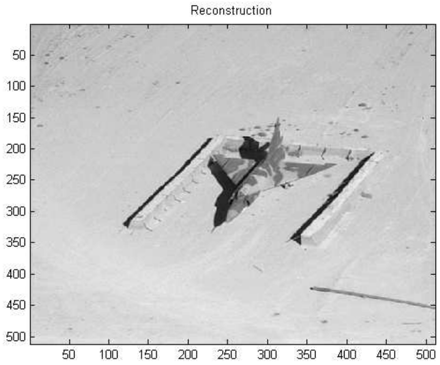


Figure: L_1-norm regularisation $\min \left\{ \|\mathfrak{A}\mathfrak{x} - \mathfrak{b}\|_2^2 + \alpha \|\mathfrak{x}\|_1 \right\}$

L_1 regularisation

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

- 4DVar smears out sharp fronts

L_1 regularisation

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

- 4DVar smears out sharp fronts
- L_1-norm regularisation has the potential to overcome this problem!

2 Regularisation Methods

4DVar

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

2 Regularisation Methods

4DVar

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

Total Variation regularisation

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

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

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

using Least squares and

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

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

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

subject to

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

where

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

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

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

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

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 $L = 5$

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

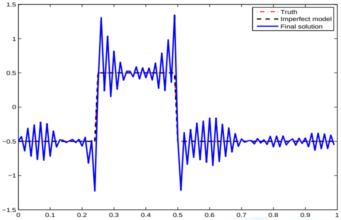


Figure: $t = 0$

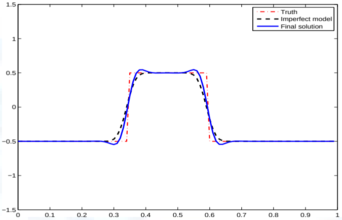


Figure: $t = 20$

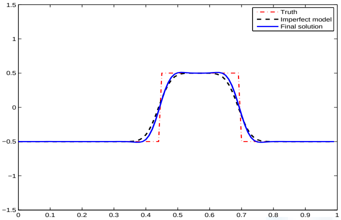


Figure: $t = 40$

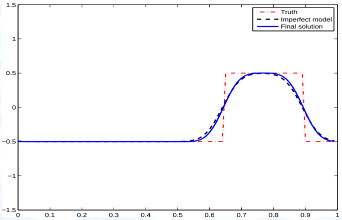


Figure: $t = 80$

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

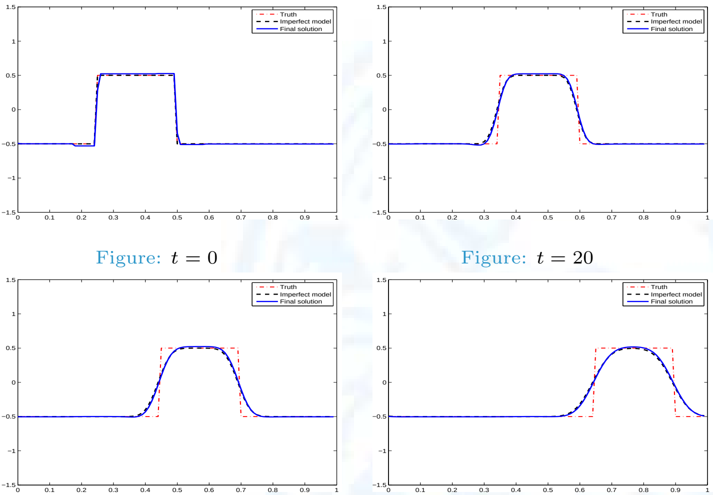


Figure: $t = 0$

Figure: $t = 20$

Figure: $t = 40$

Figure: $t = 80$

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

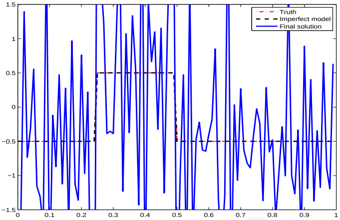


Figure: $t = 0$

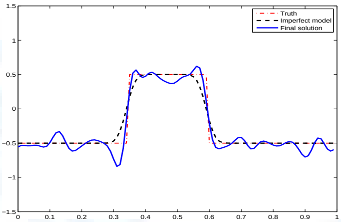


Figure: $t = 20$

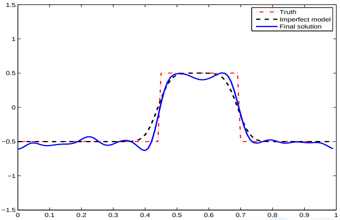


Figure: $t = 40$

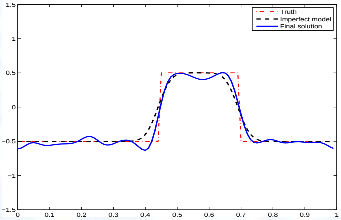


Figure: $t = 80$

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

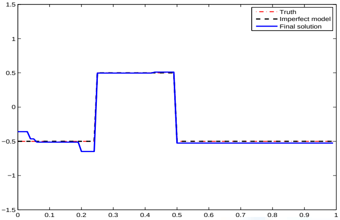


Figure: $t = 0$

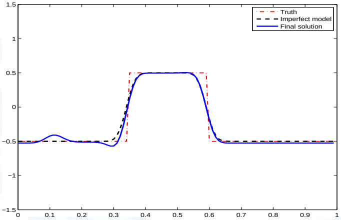


Figure: $t = 20$

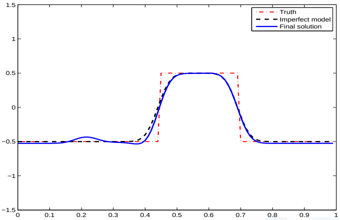


Figure: $t = 40$

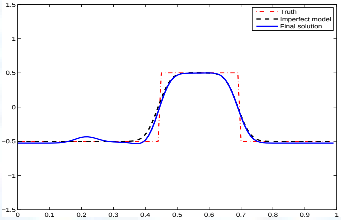


Figure: $t = 80$

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

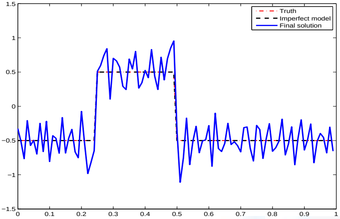


Figure: $t = 0$

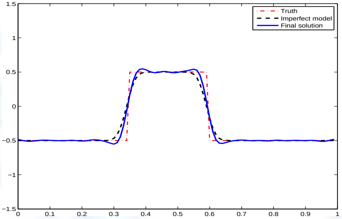


Figure: $t = 20$

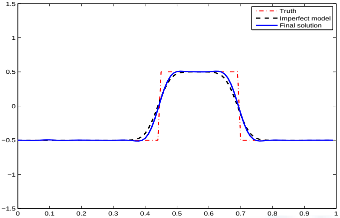


Figure: $t = 40$

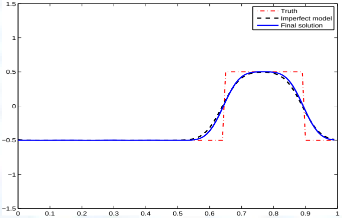


Figure: $t = 80$

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

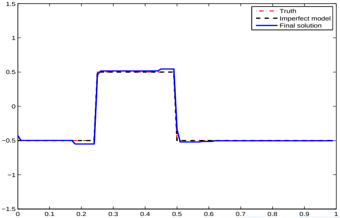


Figure: $t = 0$

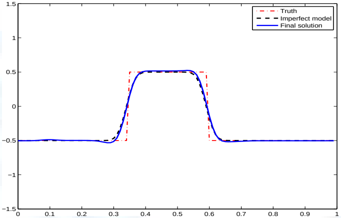


Figure: $t = 20$

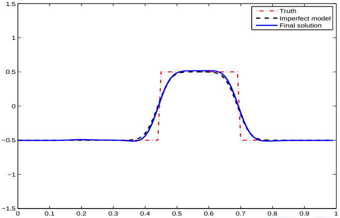


Figure: $t = 40$

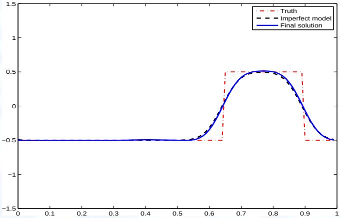


Figure: $t = 80$

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

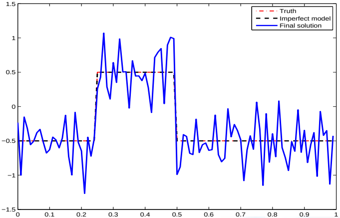


Figure: $t = 0$

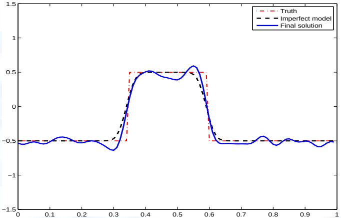


Figure: $t = 20$

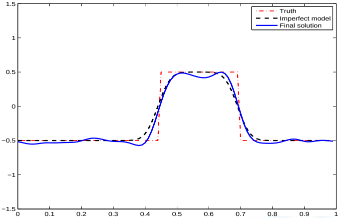


Figure: $t = 40$

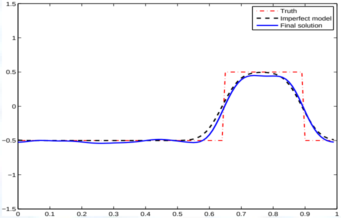


Figure: $t = 80$

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

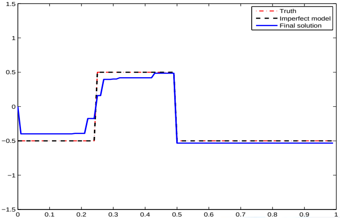


Figure: $t = 0$

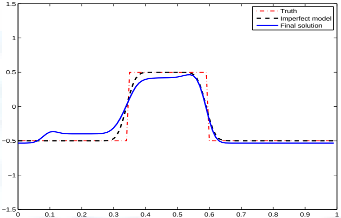


Figure: $t = 20$

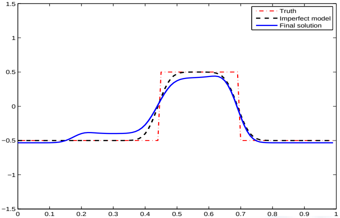


Figure: $t = 40$

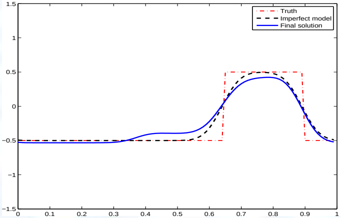


Figure: $t = 80$

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).

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

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

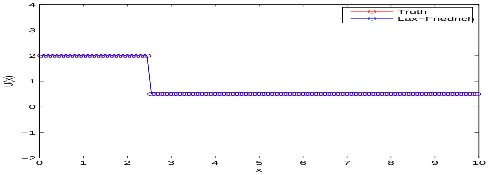


Figure: $t = 0$

Lax-Wendroff method

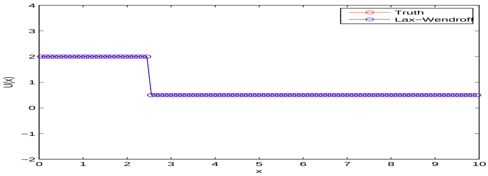


Figure: $t = 0$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

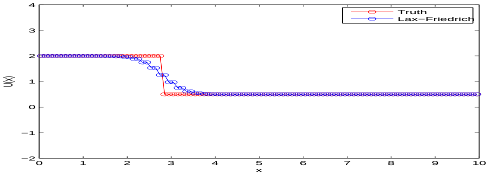


Figure: $t = 25$

Lax-Wendroff method

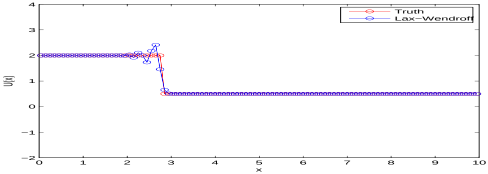


Figure: $t = 25$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

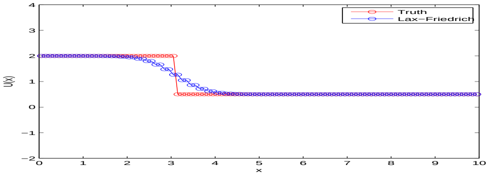


Figure: $t = 50$

Lax-Wendroff method

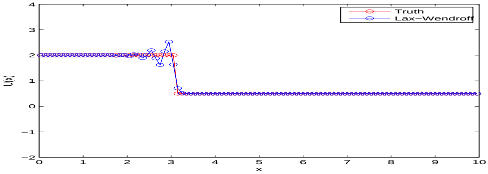


Figure: $t = 50$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

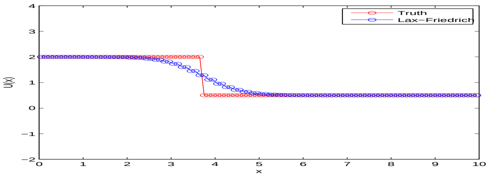


Figure: $t = 100$

Lax-Wendroff method

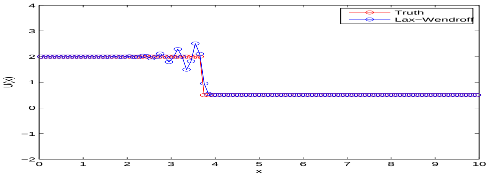


Figure: $t = 100$

Visualisation - Truth trajectory and numerical solution

Lax-Friedrichs method

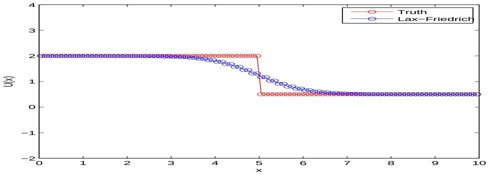


Figure: $t = 200$

Lax-Wendroff method

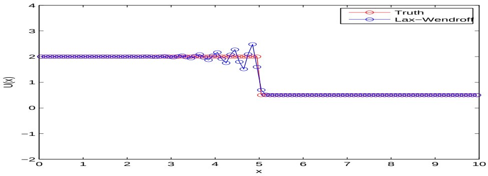


Figure: $t = 200$

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 $L = 5$

Lax-Friedrichs method

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

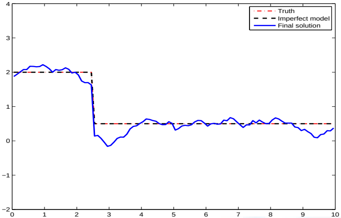


Figure: $t = 0$

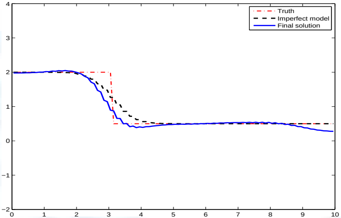


Figure: $t = 50$

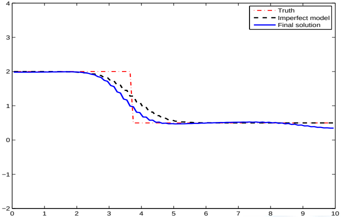


Figure: $t = 100$

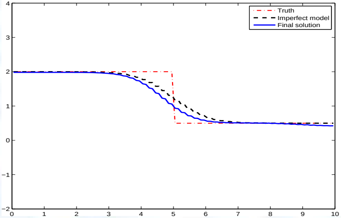


Figure: $t = 200$

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

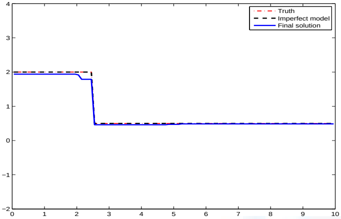


Figure: $t = 0$

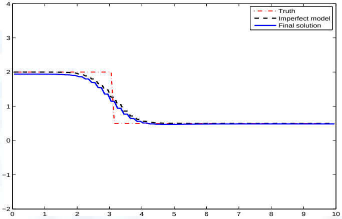


Figure: $t = 50$

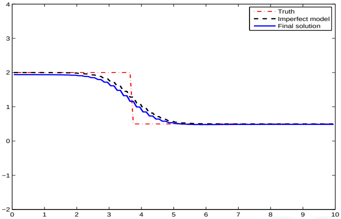


Figure: $t = 100$

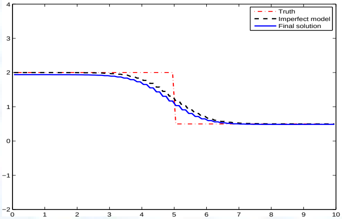


Figure: $t = 200$

Lax-Wendroff method

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

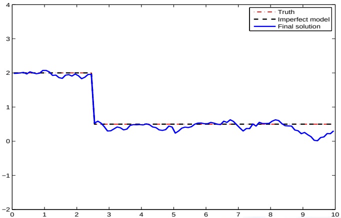


Figure: $t = 0$

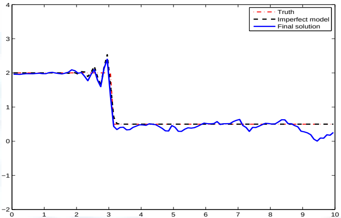


Figure: $t = 50$

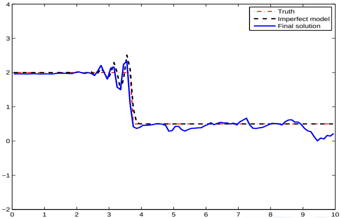


Figure: $t = 100$

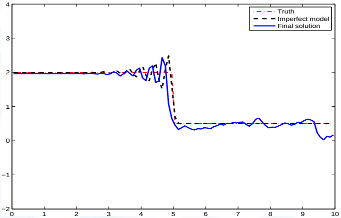


Figure: $t = 200$

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

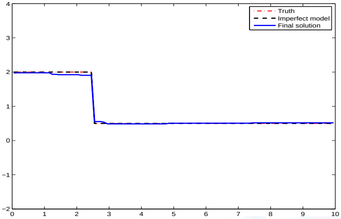


Figure: $t = 0$

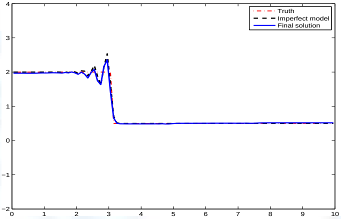


Figure: $t = 50$

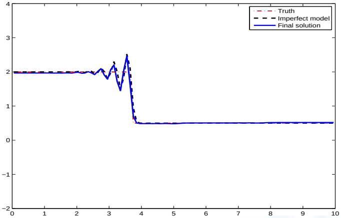


Figure: $t = 100$

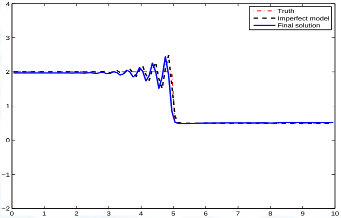


Figure: $t = 200$

Conclusions and further work

Conclusions

- regularisation parameter estimation methods improve 4DVar analysis
- L_1-norm regularisation recovers discontinuity better than 4DVar

Conclusions and further work

Conclusions

- regularisation parameter estimation methods improve 4DVar analysis
- L_1-norm regularisation recovers discontinuity better than 4DVar

Future work

- Further work: analysis of methods; convergence

Conclusions and further work

Conclusions

- regularisation parameter estimation methods improve 4DVar analysis
- L_1-norm regularisation recovers discontinuity better than 4DVar

Future work

- Further work: analysis of methods; convergence
- Extension to 2D, 3D

Conclusions and further work

Conclusions

- regularisation parameter estimation methods improve 4DVar analysis
- L_1-norm regularisation recovers discontinuity better than 4DVar

Future work

- Further work: analysis of methods; convergence
- Extension to 2D, 3D
- Multiscale methods

Weather forecast

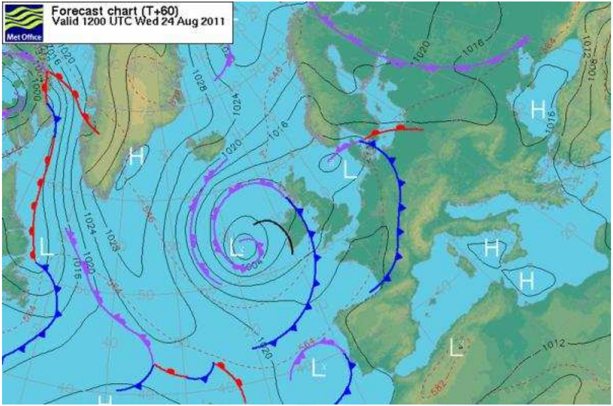


Figure: Weather forecast for Europe for Wednesday lunchtime

- C. Boess, A. Lawless, N. Nichols, and A. Bunse-Gerstner, *State estimation using model order reduction for unstable systems*, Computers & Fluids, (2010).
- C. Budd, M. Freitag, and N. Nichols, *Regularization techniques for ill-posed inverse problems in data assimilation*, Computers & Fluids, (2010).
- M. Freitag, N. Nichols, and C. Budd, *L1-regularisation for ill-posed problems in variational data assimilation*, PAMM, 10 (2010), pp. 665–668.
- H. Fu, M. K. Ng, M. Nikolova, and J. L. Barlow, *Efficient minimization methods of mixed l2-l1 and l1-l1 norms for image restoration*, SIAM J. Sci. Comput., 27 (2006), pp. 1881–1902 (electronic).
- S. Haben, A. Lawless, and N. Nichols, *Conditioning and preconditioning of the variational data assimilation problem*, Computers & Fluids, (2010).
- P. Houtekamer and H. Mitchell, *Data assimilation using an ensemble kalman filter technique*, Monthly Weather Review, 126 (1998), pp. 796–811.
- C. Johnson, B. Hoskins, and N. Nichols, *A singular vector perspective of 4d-var: Filtering and interpolation*, Quarterly Journal of the Royal Meteorological Society, 131 (2005), pp. 1–19.
- A. Lawless, N. Nichols, C. Boess, and A. Bunse-Gerstner, *Using model reduction methods within incremental four-dimensional variational data assimilation*, Monthly Weather Review, 136 (2008), pp. 1511–1522.

Thank you.

Thank you.

RICAM

Special Semester on
Multiscale Simulation & Analysis in Energy and the Environment
Linz, October 3-December 16, 2011

ÖAW

Workshop 2: October 24-28, 2011
Large-Scale Inverse Problems and Applications in the Earth Sciences