Can we reconstruct Localized Features from Non-Local Observations?

the role of observation and background errors

Olaf Stiller

Deutscher Wetterdienst, Data Assimilation Unit, Frankfurter Str. 135, Offenbach, Germany

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

(1) Can the DA system represent localised features?

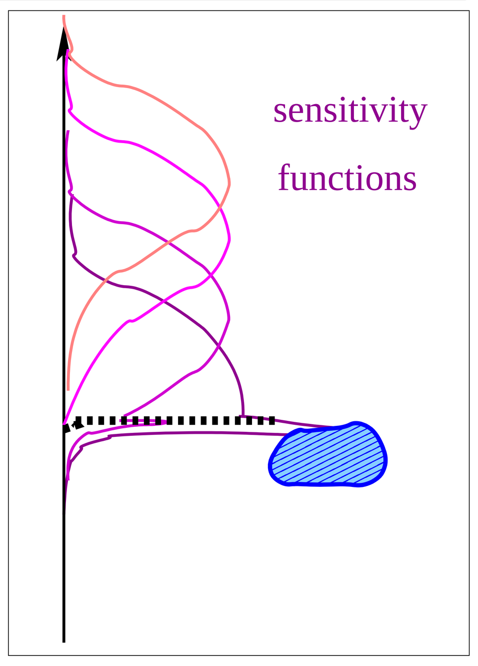
- Most modern DA systems determine the analysis increments

$$\mathbf{x}^a = \mathbf{X}^a - \mathbf{X}^b$$

via a costfunction

$$J(\mathbf{x}) = \frac{1}{2}\left\{\mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} + [\mathbf{y} - \mathbf{H}\mathbf{x}]^T\,\mathbf{R}^{-1}\,[\mathbf{y} - \mathbf{H}\mathbf{x}]\right\} \qquad (1)$$

- Many observations (particularly satellite radiances) are strongly nonlocal
- DA systems have to deal with strongly localised features like cloud tops, inversions, etc.



Question: Can the costfunction minimum describe such localised features?

Assume: Observations are locally very dense
(as true for IR radiances from hyperspectral sounders)

(2) If observation errors were negligible: The Pseudo-Inverse Solution (PI)

General solution for the cost function minimum $\mathbf{x}^a$ (of Eq.(1)):

$$\mathbf{x}^a = \left[\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \qquad (2)$$

For vanishing obs errors ($\mathbf{R} \to 0$) this yields the Pseudo Inverse (**PI**) (assume for the moment that $\left[\mathbf{H}\mathbf{B}\mathbf{H}^T\right]^{-1}$ exists)

$$\check{\mathbf{x}}^a = \mathbf{B}\mathbf{H}^T\left[\mathbf{H}\mathbf{B}\mathbf{H}^T\right]^{-1}\mathbf{y} \qquad (3)$$

with

$$\mathbf{H}_i\;\check{\mathbf{x}}^a = \mathbf{y}_i\;.$$

The **PI**

- is consistent with all observations $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, .., \mathbf{y}_p\}$
- describes localised features **as detailed as** the **density of observations** allows
- potentially **amplifies noise**

(3) Finite observation errors degrade representation of localised feature

The general solution (2) for the costfunction minimum can be written as

$$\mathbf{x}^a = \frac{\sum_\tau G_\tau \check{\mathbf{x}}^a_\tau}{\sum_\tau G_\tau} \qquad (4)$$

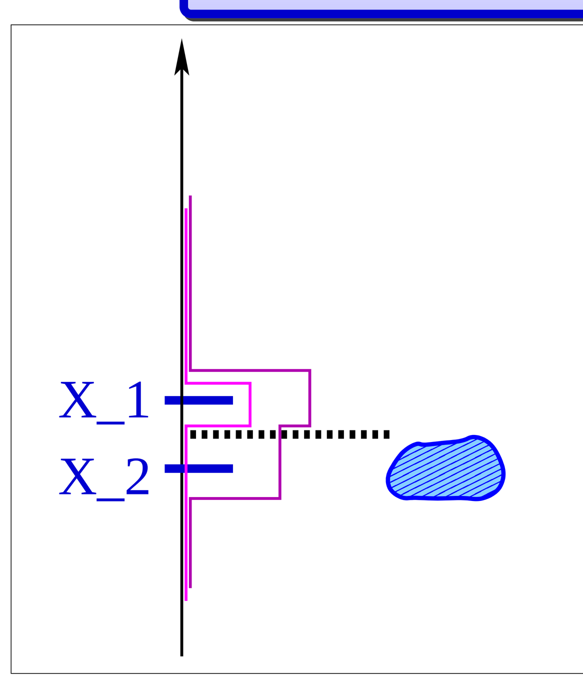
with $\check{\mathbf{x}}^a_\tau$: **PI** corresponding to **observation subset** $\tau = \{\tau_1, \tau_2, .., \tau_k\} \subset \{1, 2, .., p\}$

$\sum_\tau$: sum over all observation subsets τ

The weights $G_\tau = \frac{\det\left(\mathbf{H}_\tau\mathbf{B}\mathbf{H}_\tau^T\right)}{\prod_{j\in\tau} r_j}$

- are larger the smaller the observation errors r_j
- are smaller the more the observation operators $\mathbf{H}_\tau$ overlap
- $(G_\tau\check{\mathbf{x}}^a_\tau)$ reduces to zero if $\left(\mathbf{H}_\tau\mathbf{B}\mathbf{H}_\tau^T\right)^{-1}$ does not exist.

(4) Example: A simple model problem



- 2 degrees of freedom
- 2 observations

$$\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix} = \begin{pmatrix}h_1 & 0\\ h_0 & h_2\end{pmatrix}\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix}$$
$$= \quad \mathbf{H} \quad \mathbf{x}$$

$$\mathbf{B} = \begin{pmatrix}b_1 & b_0\\ b_0 & b_2\end{pmatrix}\;,\;\mathbf{R} = \begin{pmatrix}r_1 & 0\\ 0 & r_2\end{pmatrix}$$

$$\mathbf{x}^a = \frac{G_{\{1\}}\check{\mathbf{x}}^a_{\{1\}} + G_{\{2\}}\check{\mathbf{x}}^a_{\{2\}} + G_{\{1,2\}}\check{\mathbf{x}}^a_{\{1,2\}}}{1 + G_{\{1\}} + G_{\{2\}} + G_{\{1,2\}}}$$

$$G_{\{i\}} = \frac{\hat{B}_{ii}}{r_i} \qquad G_{\{1,2\}} = \frac{\hat{B}_{11}}{r_1}\frac{\hat{B}_{22}}{r_2}\;\frac{\det(B)h_2^2}{(b_1h_0 + b_0h_2)^2 + \det(B)}$$

$$\hat{B} = \mathbf{H}\mathbf{B}\mathbf{H}^T$$

The **PI** related to both observations ($\tau = \{1, 2\}$) yields an exact reconstruction (assuming obs errors are sufficiently small)

$$\check{\mathbf{x}}^a_{\{1,2\}} = \frac{\mathbf{y}_1}{h_1}\begin{pmatrix}1\\ -\frac{h_0}{h_2}\end{pmatrix} + \frac{\mathbf{y}_2}{h_2}\begin{pmatrix}0\\ 1\end{pmatrix}$$

Single observation **PI**s, on the other hand, smear out the signal from one observation to both levels by distributing it statistically according to **H** and **B**.

$$\check{\mathbf{x}}^a_{\{1\}} = \frac{\mathbf{y}_1}{h_1}\begin{pmatrix}1\\ \frac{b_0}{b_1}\end{pmatrix}\;;\;\check{\mathbf{x}}^a_{\{2\}} = \frac{\mathbf{y}_2}{\hat{B}_{22}}\left\{h_0\begin{pmatrix}b_1\\ b_0\end{pmatrix} + h_2\begin{pmatrix}b_0\\ b_2\end{pmatrix}\right\}$$

The weighting factors $G_{\{i\}}$ and $G_{\{1,2\}}$ act as a filter. The 2 obs **PI** $\check{\mathbf{x}}^a_{\{1,2\}}$ amplifies noise particularly when $\det(\mathbf{B})$ or h_2 are very small. One has, e.g.,

$$\left(G_{\{1,2\}}\check{\mathbf{x}}^a_{\{1,2\}}\right) \to h_2 \quad \text{as} \quad h_2 \to 0$$

(5) Summary and conclusions

The main mathematical result

- A novel way of writing the costfunction minimum $\mathbf{x}^a$ has been presented - see Eq.(4).
- This expands $\mathbf{x}^a$ in a sum over Pseudo Inverses (**PI**)s (see Eq.(3)), each corresponding to a different subset τ of the available observations (τ: index set with $\tau = \{\tau_1, \tau_2, .., \tau_k\} \subset \{1, 2, .., p\}$, where p is the total number of observations).

The role of the Pseudo Inverse (PI)

- The **PI** for a given subset τ leads to an analysis state which is completely consistent with all the observations from τ. It can therefore be regarded as a direct **transformation of the observations τ into model space**.
- However: the **PI** is generally not optimal:
  - The **PI** neglects observation error
  - The **PI** tends to amplify noise

The expansion of $\mathbf{x}^a$ in terms of PIs (see Eq. (4))

- The coefficients G_τ in the expansion (4) show to which extent different observation sets τ contribute to the analysis increments $\mathbf{x}^a$. There are two limiting cases
  1. Obs errors are **very small**:
     - dominant are **PI**s $\check{\mathbf{x}}^a_\tau$ of the largest observation sets τ for which $\left[\mathbf{H}_\tau\mathbf{B}\left(\mathbf{H}_\tau\right)^T\right]^{-1}$ exists.
     - the spatial accuracy is the maximally achievable accuracy given the observation density.
  2. Obs errors are **very large**:
     - dominant are single obs **PI**s $\check{\mathbf{x}}^a_{\{k\}}$
     - they smear out signals from individual observations
- The coefficients G_τ filter the noise.
  - G_τ is very small if observation errors exceed the required precision.

$$G_\tau = \frac{\det\left(\widetilde{\mathbf{B}}_\tau\right)}{\prod_{i\in\tau} R_i}$$

$\widetilde{\mathbf{B}}_\tau$: background correlation matrix in obs space.

$\det\left(\widetilde{\mathbf{B}}_\tau\right)$ gives a measure for the overlap of obs-operators

$R_i = r_i / \left\{\mathbf{H}_\tau\mathbf{B}\left(\mathbf{H}_\tau\right)^T\right\}_{ii}$
(obs-/background error)
Normalised obs error

Conclusions

- The expansion of $\mathbf{x}^a$ in terms of PIs shows to which extent measured degrees of freedom (which are non-local!) are exploited for reconstructing spatial features.
- Large obs errors
  ⇒ degrade spatial resolution
  (not only decrease weight of obs in assimilation process)

Reconstruction of localized features

- requires small obs errors.
- Obs errors have to be smaller the more
  - observation operators overlap.
  - observations contradict statistical expectations from **B** matrix.

For proofs and further discussion:

O. Stiller, *The role of observation and background errors for reconstructing localized features from non-local observations.*, Physica **D**, 275C, pp.43-53 (2014)

For free online access: http://www.sciencedirect.com/science/article/pii/S0167278914000177