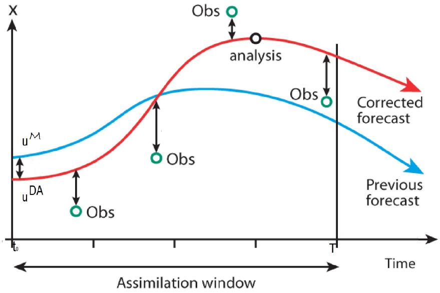
Ab-initio Functional Decomposition of Kalman Filter: a feasibility analysis on Constrained Least Squares Problems

L. D'Amore, R. Cacciapuoti, V. Mele

Department of Mathematics "R. Caccioppoli", University of Naples Federico II

PPAM 2019 - Białystok, Poland
September 9, 2019

Data Assimilation (DA) problem and Variational (Var) formulation.



Given

- $y = \{y_k\}_{k=0,\dots,r+1} \in \mathbb{R}^{(r+2)\cdot nobs}$: the vector of the observations;
- $u_0 \in \mathbb{R}^{NP}$: the vector of the state at time t_0;
- $G \in \mathbb{R}^{(r+2)\cdot nobs \times NP = [H_0 \;\; H_1 \;\; \dots \;\; H_{r+1}]^T}$: the block matrix composed by linear approximation of a observation mapping;
- $\mathbf{Q} = diag(Q_0, Q_1, \dots, Q_{r+1})$, $\mathbf{R} = diag(R_0, R_1 \dots, R_{r+1})$: covariance matrices of the errors on observations and model, respectively.

The **4D-Var DA problem** concerns the computation of:

$$u^{DA} = argmin_{u \in \mathbb{R}^{NP \cdot N}} J(u),$$

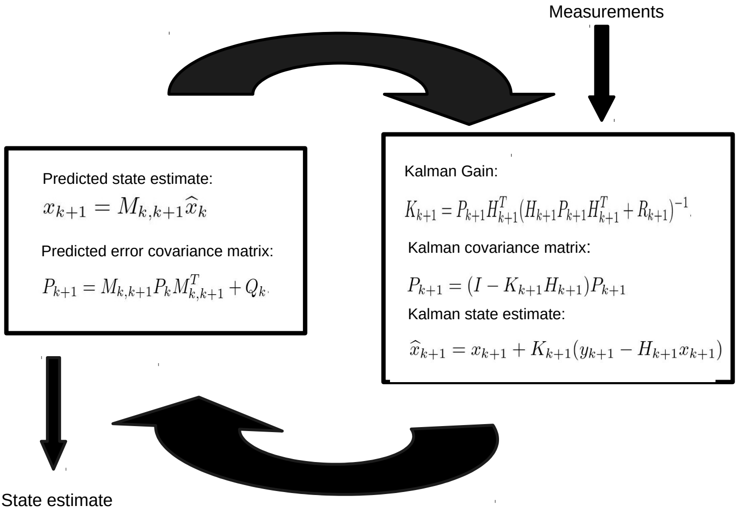
with

$$J(u) = \alpha ||u - u^{\mathcal{M}}||^2_{\mathbf{B}^{-1}} + ||Gu - y||^2_{\mathbf{R}^{-1}}.$$

Introduction and motivations

- DA problem usually used to handle a huge amount of data, so, it is a large and computationally expensive problem;
- many assimilation techniques have been developed for solving DA problem, we focus on Kalman Filter (KF);
- KF dates back to 1960, when R.E. Kalman provided recursive algorithm to compute the solution of a data filtering and prediction problem;
- we consider Constrained Least Square (CLS) model as this is the prototype model of variational DA;
- we introduce a new formulation of DD for KF where the innovation mainly lies in the decomposition ab initio of the whole KF computational method.

KF procedure



CLS problem

We consider the overdetermined linear system

$$S: \quad Ax = y$$

where

$$A = \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m \times n}; \; y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m} \quad m > n, \; r \in \mathbb{N}.$$

and $R = diag(R_0, R_1, \ldots, R_{r+1}) \in \mathbb{R}^{(r+2)\cdot m \times (r+2)\cdot m}$ the weight matrix.
This formulation allows us to apply KF (recursive algorithm).

CLS problem

CLS problem refers to the computation of $\widehat{x}$ such that:

$$CLS: \quad \widehat{x} = \arg\min_{x \in \mathbb{R}^n} J(x) \tag{1}$$

where $J(x) = ||Ax - y||_R^2 = \sum_{k=0}^{r+1} ||H_k x - y_k||_{R_k}^2$.

Variational-Kalman Filter (Var-KF) formulation

Var Kalman problem

The **Var Kalman problem** is to compute for each $k = 0, 1, \ldots, r$ the vector $\widehat{x}_{k+1}$:

$$\begin{aligned} \widehat{x}_{k+1} &= argmin_{x_{k+1} \in \mathbb{R}^n} J_{k+1}(x_{k+1}) \\ &= \{\|x_{k+1} - M_{k,k+1}\widehat{x}_k\|^2_{Q_k} + \|y_{k+1} - H_{k+1}x_{k+1}\|^2_{R_{k+1}}\} \end{aligned}$$

We note that VAR-KF is indeed a CLS problem.

KF procedure applied to CLS problem.

Measurements

Predicted state estimate:

$$x_{k+1} = M_{k,k+1}\widehat{x}_k$$

Predicted error covariance matrix:

$$P_{k+1} = M_{k,k+1}P_k M_{k,k+1}^T + Q_k$$

Kalman Gain:

$$K_{k+1} = P_{k+1}H_{k+1}^T(H_{k+1}P_{k+1}H_{k+1}^T + R_{k+1})^{-1}$$

Kalman covariance matrix:

$$P_{k+1} = (I - K_{k+1}H_{k+1})P_{k+1}$$

Kalman state estimate:

$$\widehat{x}_{k+1} = x_{k+1} + K_{k+1}(y_{k+1} - H_{k+1}x_{k+1})$$

State estimate

Where:

$$\begin{aligned} M_{k,k+1} &:= I_{n,n} \\ Q_k &:= O_{n,n} \\ P_0 &:= (H_0^T R_0 H_0)^{-1} \end{aligned}$$

2. DD-CLS

We consider the case $r = 0$ i.e.

$$A = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \in \mathbb{R}^{(m_0+m_1)\times n}; \; y = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \in \mathbb{R}^{(m_0+m_1)} \quad m_0 > n, \; m_1 \in \mathbb{N}.$$

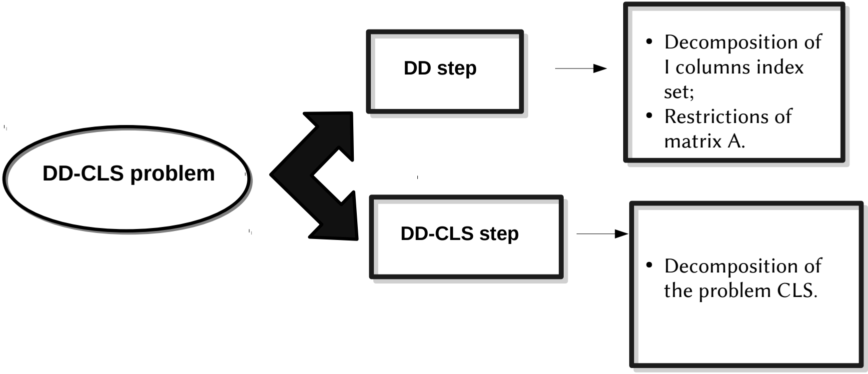
and let $I = \{1, \ldots, n\}$ be the columns index set of $A \in \mathbb{R}^{(m_0+m_1)\times n}$.

2. DD-CLS

We consider the case $r = 0$ i.e.

$$A = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \in \mathbb{R}^{(m_0+m_1)\times n}; \; y = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \in \mathbb{R}^{(m_0+m_1)} \quad m_0 > n, \; m_1 \in \mathbb{N}.$$

and let $I = \{1, \ldots, n\}$ be the columns index set of $A \in \mathbb{R}^{(m_0+m_1)\times n}$.



2. DD-CLS: DD step.

DD step:

DD step:

- decomposition of I of into 2 sets (domains)

$$I_1 = \{1, ..., n_1\}, \quad I_2 = \{n_1 - s + 1, ..., n\}, \tag{2}$$

where $s \geq 0$ is the number of indexes in common, $|I_1| = n_1 > 0$, $|I_2| = n_2 > 0$, and the intersection (overlap) of sets

$$I_{1,2} = \{n_1 - s + 1, ..., n_1\}, \tag{3}$$

2. DD-CLS: DD step.

DD step:

- decomposition of I of into 2 sets (domains)

$$I_1 = \{1, ..., n_1\}, \quad I_2 = \{n_1 - s + 1, ..., n\}, \tag{2}$$

where $s \geq 0$ is the number of indexes in common, $|I_1| = n_1 > 0$, $|I_2| = n_2 > 0$, and the intersection (overlap) of sets

$$I_{1,2} = \{n_1 - s + 1, ..., n_1\}, \tag{3}$$

- restrictions of A to I_1 and I_2 defined in (2)

$$A_1 = A|_{I_1} \in \mathbb{R}^{(m_0+m_1)\times n_1}, \quad A_2 = A|_{I_2} \in \mathbb{R}^{(m_0+m_1)\times n_2}. \tag{4}$$

2. DD-CLS: DD-CLS step.

DD-CLS step:

2. DD-CLS: DD-CLS step.

DD-CLS step:

- given $x_2^0 \in \mathbb{R}^{n_2}$, according the Alternating Schwarz method (ASM), the DD approach consists in solving for $n = 0, 1, 2, ...$ the following overdetermined linear systems::

$$S_1^{n+1}: \quad A_1 x_1^{n+1} = b - A_2 x_2^n \quad S_2^{n+1}: \quad A_2 x_2^{n+1} = b - A_1 x_1^{n+1}. \quad (5)$$

This means to solve

$$\begin{aligned} P_1^{n+1}: \ \widehat{x}_{1,1}^{n+1} &= \arg\min_{x_1^{n+1} \in \mathbb{R}^{n_1}} J_1(x_1^{n+1}, x_2^n) \\ &= \arg\min_{x_1^{n+1} \in \mathbb{R}^{n_1}} \left[J|_{(I_1, I_2)}(x_1^{n+1}, x_2^n) + \mu \cdot \mathcal{O}_{1,2}(x_1^{n+1}, x_2^n) \right] \end{aligned} \quad (6)$$

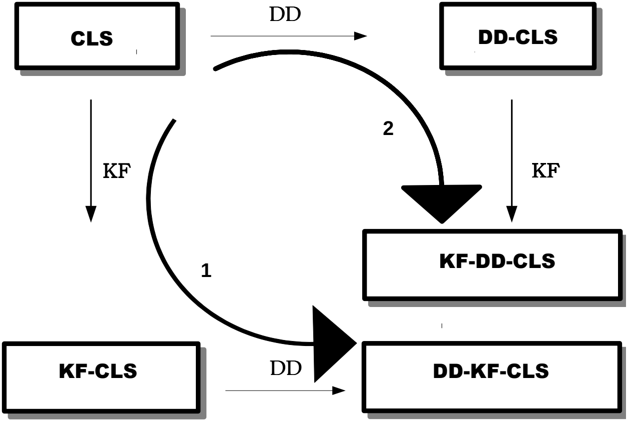
$$\begin{aligned} P_2^{n+1}: \ \widehat{x}_{2,1}^{n+1} &= \arg\min_{x_2^{n+1} \in \mathbb{R}^{n_2}} J_2(x_2^{n+1}, x_1^{n+1}) \\ &= \arg\min_{x_2^{n+1} \in \mathbb{R}^{n_2}} \left[J|_{(I_2, I_1)}(x_2^{n+1}, x_1^{n+1}) + \mu \cdot \mathcal{O}_{1,2}(x_2^{n+1}, x_1^{n+1}) \right] \end{aligned} \quad (7)$$

where

$$J|_{(I_i, I_j)}(x|_{I_i}, x|_{I_j}) = ||H_0|_{I_i} x|_{I_i} - (y_0 + H_0|_{I_j} x|_{I_j})||_{R_0}^2 + ||H_1|_{I_i} x|_{I_i} - (y_1 + H_1|_{I_j} x|_{I_j})||_{R_1}^2, \quad (8)$$

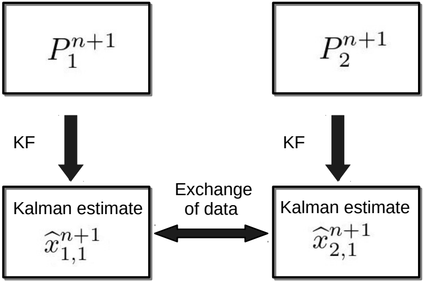
for $i, j = 1, 2$, $\mathcal{O}_{1,2}$ is appropriate overlapping operator and $\mu > 0$ is the regularization parameter.

Validation scheme

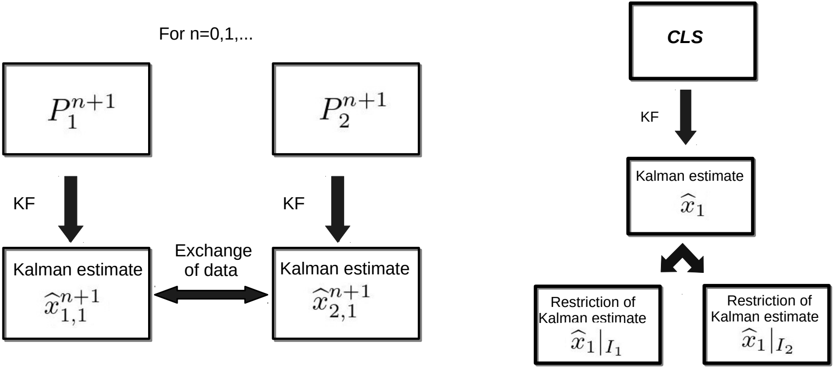


1. DD-KF-CLS VS 2. KF-DD-CLS

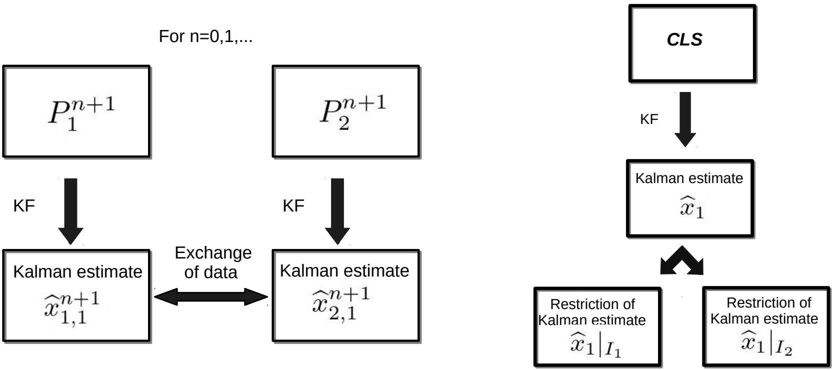
For n=0,1,...



1. DD-KF-CLS VS 2. KF-DD-CLS



1. DD-KF-CLS VS 2. KF-DD-CLS



We show that

$$\widehat{x}_{1,1}^{n+1} \xrightarrow[n\to\infty]{} \widehat{x}_1\big|_{I_1}, \qquad \widehat{x}_{2,1}^{n+1} \xrightarrow[n\to\infty]{} \widehat{x}_1\big|_{I_2}.$$

Theorem

Let

$$S: \quad Ax = y$$

be the overdetermined linear system where

$$A = \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m \times n}; \; y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m} \quad m > n, \; r \in \mathbb{N}.$$

Let $\widehat{x}_1 \in \mathbb{R}^n$ be the Kalman estimate of the solution $\widehat{x}$ of problem

$$CLS: \quad \widehat{x} = \arg\min_{x \in \mathbb{R}^n} J(x)$$

where $J(x) = ||Ax - y||_R^2$ and $R \in \mathbb{R}^{(r+2)\cdot m \times (r+2)\cdot m}$ the weight matrix and $\widehat{x}_{1,1}^{n+1}$, $\widehat{x}_{2,1}^{n+1}$ be Kalman estimates of local problems P_1^{n+1}, P_2^{n+1}, then

Theorem

Let

$$S: \quad Ax = y$$

be the overdetermined linear system where

$$A = \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m \times n}; \ y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{r+1} \end{bmatrix} \in \mathbb{R}^{(r+2)\cdot m} \quad m > n, \ r \in \mathbb{N}.$$

Let $\widehat{x}_1 \in \mathbb{R}^n$ be the Kalman estimate of the solution $\widehat{x}$ of problem

$$CLS: \quad \widehat{x} = \arg\min_{x \in \mathbb{R}^n} J(x)$$

where $J(x) = ||Ax - y||_R^2$ and $R \in \mathbb{R}^{(r+2)\cdot m \times (r+2)\cdot m}$ the weight matrix and $\widehat{x}_{1,1}^{n+1}$, $\widehat{x}_{2,1}^{n+1}$ be Kalman estimates of local problems P_1^{n+1}, P_2^{n+1}, then

$$\widehat{x}_{1,1}^{n+1} \xrightarrow[n\to\infty]{} \widehat{x}_1|_{I_1}, \quad \widehat{x}_{2,1}^{n+1} \xrightarrow[n\to\infty]{} \widehat{x}_1|_{I_2}. \tag{9}$$

Experimental Results

Example

Decomposition of $I = \{1,2,3,4,5,6\}$ with overlap into I_1 and I_2 i.e. for $s = 0,1,2$, $I_1 = \{1,2,3,4\}$, $I_2 = \{4-s,...,n\}$ and the overlap set $I_{1,2} = \{4-s,...,4\}$.

For $s = 0,1,2$, we consider

- $n \equiv |I| = 6$, $n_1 \equiv |I_1| = 4$, $n_2 \equiv n + s - n_1 \equiv |I_2| = 2 + s$;
- $\widehat{x}_{i,1}^{0} \equiv zeros(n_i) \in \mathbb{R}^{n_i}$, where $zeros(n_i)$ is the null vector, $i = 1,2$; for each iteration $n = 1,2,...,nmax$, we compute
- $\widehat{x}_{1,1}^{n+1} \in \mathbb{R}^4$, $\widehat{x}_{2,1}^{n+1} \in \mathbb{R}^2$: the Kalman estimates.
- $x_s^{n+1} \in \mathbb{R}^9$ the Kalman estimate obtained as follows

$$x_s^{n+1} = \begin{cases} \widehat{x}_{1,1}^{n+1}|_{I_1 \setminus I_{1,2}} & on\ I_1 \setminus I_{1,2} \\ \frac{\mu}{2}(\widehat{x}_1^{n+1}|_{I_{1,2}} + \widehat{x}_2^{n+1}|_{I_{1,2}}) & on\ I_{1,2} \\ \widehat{x}_{2,1}^{n+1}|_{I_2 \setminus I_{1,2}} & on\ I_2 \setminus I_{1,2} \end{cases},$$

  with as regularization parameter $\mu \equiv 1$;
- $\bar{x}$ solution of normal equations by Conjugate Gradient method;
- ns_s the corresponding iterations needed to stop of the iterative procedure.

Example

In Table, for different choices of s, we report the values of the error and the relative number of iterations ns necessary to overcome the established tolerance.

Table: Values of $error = ||\bar{x} - x^{ns}||$ for different values of s.

tol	ns	$error$	s
10^{-6}	20	$6.4037e-07$	0
	17	$7.2526e-07$	1
	15	$5.1744e-07$	2

Conclusions

- We introduced a new numerical scheme based on DD approach and KF method for solving DA when the numerical data, i.e. forecast and observations, are all available at time of analysis;
- DA has long been playing a crucial role in meteorology and oceanography and more in general, in climate science;
- we are interested in real time data assimilation, so, we have only the numerical data at the time of analysis;
- we are working on decomposition of the evolution model (as model based on the shallow water equations that describing fluid flow in the atmosphere, oceans, rivers and channels) in sync with DD-KF.