

Multivariate localization

Benjamin Ménétrier

Last update: October 8, 2018

Documentation for the code "BUMP", distributed under the CeCILL-C license.
Copyright ©2015-... UCAR, CERFACS, METEO-FRANCE and IRIT

Contents

1	Framework	2
2	Univariate specific blocks	2
3	Multivariate specific blocks	3
3.1	State grid interpolation	3
3.2	Control grid interpolation	3
3.3	Interpolation error	4
3.4	Localization amplitude	5
4	Common block	6
4.1	Basic method	6
4.2	Weighted variant	6
5	Applications	7

1 Framework

A localization matrix $\mathbf{L} \in \mathbb{R}^{n \times n}$ is applied to a multivariate state vector $\mathbf{x} \in \mathbb{R}^n$, built as the concatenation of p sub-vectors:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_p \end{pmatrix} \quad (1)$$

where \mathbf{x}_i is the sub-vector for variable i , of size n_i . Obviously, $n = \sum_{i=1}^p n_i$.

The localization matrix is split into blocks:

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_{11} & \cdots & \mathbf{L}_{1p} \\ \vdots & \ddots & \vdots \\ \mathbf{L}_{p1} & \cdots & \mathbf{L}_{pp} \end{pmatrix}$$

The problem is to design a positive semi-definite localization matrix:

- \mathbf{L}_{ii} should be positive semi-definite,
- the constraint on \mathbf{L}_{ij} for $i \neq j$ is less obvious...

The solution is to build a square-root $\mathbf{U} \in \mathbb{R}^{n \times m}$ of the localization matrix \mathbf{L} , such that $\mathbf{L} = \mathbf{U}\mathbf{U}^T$. The control vector \mathbf{v} such that $\mathbf{x} = \mathbf{U}\mathbf{v}$ is defined in the control space of size m .

2 Univariate specific blocks

The simplest method is to design \mathbf{U} as:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{U}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{U}_p \end{pmatrix} \quad (2)$$

where $\mathbf{U}_i \in \mathbb{R}^{n_i \times m_i}$, m_i being the size of the control sub-vector related to variable i with $m = \sum_{i=1}^p m_i$. The resulting localization matrix is:

$$\mathbf{L} = \mathbf{U}\mathbf{U}^T = \begin{pmatrix} \mathbf{U}_1\mathbf{U}_1^T & 0 & \cdots & 0 \\ 0 & \mathbf{U}_2\mathbf{U}_2^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{U}_p\mathbf{U}_p^T \end{pmatrix} \quad (3)$$

In this case, auto-localization blocks are well defined:

$$\mathbf{L}_{ii} = \mathbf{U}_i \mathbf{U}_i^T \quad (4)$$

but cross-localization blocks \mathbf{L}_{ij} for $i \neq j$ are zero.

Advantages: well-adapted for large differences between variables length-scales and/or for small cross-correlation (sampling noise is canceled).

Drawbacks: ill-adapted for significant cross-correlations (signal is lost).

Cost: p applications of $\bar{\mathbf{U}}_i$.

3 Multivariate specific blocks

3.1 State grid interpolation

A second method is to perform interpolations to and from a common state grid of size \bar{n} . $\mathbf{S}_i^s \in \mathbb{R}^{n_i \times \bar{n}}$ is defined as the interpolation from the common state grid to the state grid of variable i . We gather these matrices into a global state interpolation matrix \mathbf{S}^s :

$$\mathbf{S}^s = \begin{pmatrix} \mathbf{S}_1^s & 0 & \dots & 0 \\ 0 & \mathbf{S}_2^s & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{S}_p^s \end{pmatrix} \quad (5)$$

A specific block $\bar{\mathbf{U}}_i \in \mathbb{R}^{\bar{n} \times m}$ is applied to each variable i :

$$\mathbf{U} = \mathbf{S}^s \begin{pmatrix} \bar{\mathbf{U}}_1 \\ \vdots \\ \bar{\mathbf{U}}_p \end{pmatrix} \quad (6)$$

The resulting localization matrix is:

$$\mathbf{L} = \mathbf{U} \mathbf{U}^T = \mathbf{S}^s \begin{pmatrix} \bar{\mathbf{U}}_1 \bar{\mathbf{U}}_1^T & \dots & \bar{\mathbf{U}}_1 \bar{\mathbf{U}}_p^T \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{U}}_p \bar{\mathbf{U}}_1^T & \dots & \bar{\mathbf{U}}_p \bar{\mathbf{U}}_p^T \end{pmatrix} \mathbf{S}^{sT} \quad (7)$$

In this case, each auto- or cross-localization block is given by:

$$\mathbf{L}_{ij} = \mathbf{S}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{S}_j^{sT} \quad (8)$$

3.2 Control grid interpolation

An alternative to this method is to perform interpolations to and from a common control grid of size \bar{m} . $\mathbf{S}_i^c \in \mathbb{R}^{m_i \times \bar{m}}$ is defined as the interpolation from the common vector grid to the control

grid of variable i . We can write a global interpolation matrix \mathbf{S}^c as:

$$\mathbf{S}^c = \begin{pmatrix} \mathbf{S}_1^c & 0 & \cdots & 0 \\ 0 & \mathbf{S}_2^c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{S}_p^c \end{pmatrix} \quad (9)$$

A specific block $\mathbf{U}_i \in \mathbb{R}^{n_i \times m_i}$ is applied to each variable:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_p \end{pmatrix} \mathbf{S}^c \quad (10)$$

The resulting localization matrix is $\mathbf{L} = \mathbf{U}\mathbf{U}^T$ with each auto- or cross-localization block given by:

$$\mathbf{L}_{ij} = \mathbf{U}_i \mathbf{S}_i^c \mathbf{S}_j^{cT} \mathbf{U}_j^T \quad (11)$$

3.3 Interpolation error

We assume that every interpolation \mathbf{S}^s or \mathbf{S}^c can be expanded as $\mathbf{S} = \mathbf{T} + \mathbf{E}$ where \mathbf{T} is the "true" interpolation, i.e. the best we can get with a linear operator, and $\mathbf{E} = \mathbf{S} - \mathbf{T}$ the resulting error. Thus:

- For the state grid interpolation with specific blocks:

$$\begin{aligned} \mathbf{L}_{ij} &= \mathbf{S}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{S}_j^{sT} \\ &= \mathbf{T}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{T}_j^{sT} + \mathbf{E}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{T}_j^{sT} + \mathbf{T}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{E}_j^{sT} + \mathbf{E}_i^s \bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T \mathbf{E}_j^{sT} \end{aligned} \quad (12)$$

and similarly for the state grid interpolation with a common block.

- For the control grid interpolation:

$$\begin{aligned} \mathbf{L}_{ij} &= \mathbf{U}_i \mathbf{S}_i^c \mathbf{S}_j^{cT} \mathbf{U}_j^T \\ &= \mathbf{U}_i \mathbf{T}_i^c \mathbf{T}_j^{cT} \mathbf{U}_j^T + \mathbf{U}_i \mathbf{E}_i^c \mathbf{T}_j^{cT} \mathbf{U}_j^T + \mathbf{U}_i \mathbf{T}_i^c \mathbf{E}_j^{cT} \mathbf{U}_j^T + \mathbf{U}_i \mathbf{E}_i^c \mathbf{E}_j^{cT} \mathbf{U}_j^T \end{aligned} \quad (13)$$

Assuming that interpolation errors are random with a short correlation length-scale, the control grid interpolation looks like a better option since a smoothing is always applied as the last operator. However, even with the state grid interpolation:

- either the interpolation error is smoothed in subsequent operators,
- or the interpolation error is itself applied on smoothed fields at high-resolution, leading to small interpolation errors.

As a conclusion, we don't think that interpolation error is a major issue that should favor one approach or the other.

3.4 Localization amplitude

We denote $\bar{\mathbf{u}}$ the k^{th} column of $\bar{\mathbf{U}}_i$ and $\bar{\mathbf{u}}'$ the k^{th} column of $\bar{\mathbf{U}}_j$. The k^{th} diagonal coefficient of $\bar{\mathbf{U}}_i \bar{\mathbf{U}}_j^T$ is given by $\langle \bar{\mathbf{u}}, \bar{\mathbf{u}}' \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product. The Cauchy-Schwartz inequality imposes that:

$$|\langle \bar{\mathbf{u}}, \bar{\mathbf{u}}' \rangle| \leq \sqrt{\langle \bar{\mathbf{u}}, \bar{\mathbf{u}} \rangle \langle \bar{\mathbf{u}}', \bar{\mathbf{u}}' \rangle} \quad (14)$$

Thus, the cross-localization amplitude between variables i and j is necessarily smaller than the geometric mean of the auto-localization amplitudes for variables i and j . Besides, the more the vectors $\bar{\mathbf{u}}$ and $\bar{\mathbf{u}}'$ have different shapes, the smaller their inner product and the smaller the cross-localization amplitude.

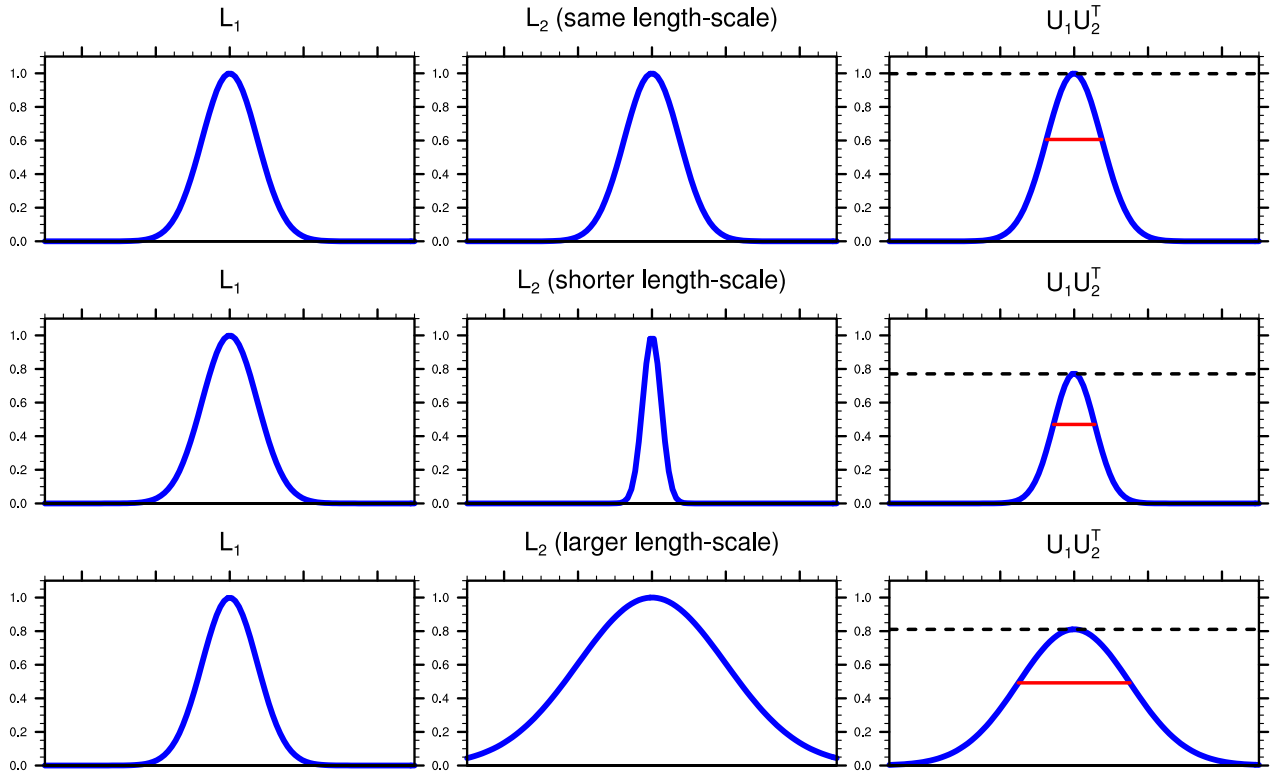


Figure 1: 1D illustration of the product of two localization square-root, with Gaussian localization functions.

Advantages: well-adapted for large differences between variables length-scales.

Drawbacks: no control on the cross-localization amplitude.

Cost: p applications of $\bar{\mathbf{U}}_i$.

4 Common block

4.1 Basic method

A third method is to apply a common block $\bar{\mathbf{U}} \in \mathbb{R}^{\bar{n} \times m}$ to all variables, with a state grid interpolation:

$$\mathbf{U} = \mathbf{S}^s \begin{pmatrix} \bar{\mathbf{U}} \\ \vdots \\ \bar{\mathbf{U}} \end{pmatrix} = \mathbf{S}^s \begin{pmatrix} \mathbf{I}_{\bar{n}} \\ \vdots \\ \mathbf{I}_{\bar{n}} \end{pmatrix} \bar{\mathbf{U}} \quad (15)$$

where $\mathbf{I}_{\bar{n}} \in \mathbb{R}^{\bar{n} \times \bar{n}}$ is the identity matrix. It should be noted that to apply \mathbf{U} , the matrix $\bar{\mathbf{U}}$ needs to be applied only once. The resulting localization matrix is:

$$\mathbf{L} = \mathbf{U}\mathbf{U}^T = \mathbf{S}^s \begin{pmatrix} \bar{\mathbf{U}}\bar{\mathbf{U}}^T & \dots & \bar{\mathbf{U}}\bar{\mathbf{U}}^T \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{U}}\bar{\mathbf{U}}^T & \dots & \bar{\mathbf{U}}\bar{\mathbf{U}}^T \end{pmatrix} \mathbf{S}^{sT} \quad (16)$$

In this case, each auto- or cross-localization block is given by:

$$\mathbf{L}_{ij} = \mathbf{S}_i^s \bar{\mathbf{U}}\bar{\mathbf{U}}^T \mathbf{S}_j^{sT} \quad (17)$$

It should be noted that this approach is inconsistent with the optimal linear filtering theory. Indeed, the cross-correlation between different variables, even at zero separation, is lower than 1 (i.e. the value for auto-correlations). Thus, the cross-localization should be lower than the auto-localization at zero separation, whereas $[\bar{\mathbf{U}}\bar{\mathbf{U}}^T]_{ii}$ is applied in both case.

Advantages: lower cost.

Drawbacks: ill-adapted for large differences between variables length-scales and/or for small cross-correlations (sampling noise is kept).

Cost: one application of $\bar{\mathbf{U}}$.

4.2 Weighted variant

A fourth method is to apply a lower triangular square-root of weighted common blocks $\bar{\mathbf{U}} \in \mathbb{R}^{\bar{n} \times m}$, with a state grid interpolation:

$$\mathbf{U} = \mathbf{S}^s \begin{pmatrix} Q_{1,1} \bar{\mathbf{U}} & 0 & \dots & 0 \\ Q_{2,1} \bar{\mathbf{U}} & Q_{2,2} \bar{\mathbf{U}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ Q_{p,1} \bar{\mathbf{U}} & \dots & Q_{p,p-1} \bar{\mathbf{U}} & Q_{p,p} \bar{\mathbf{U}} \end{pmatrix} \quad (18)$$

$$= \mathbf{S}^s \begin{pmatrix} Q_{1,1} \mathbf{I}_{\bar{n}} & 0 & \dots & 0 \\ Q_{2,1} \mathbf{I}_{\bar{n}} & Q_{2,2} \mathbf{I}_{\bar{n}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ Q_{p,1} \mathbf{I}_{\bar{n}} & \dots & Q_{p,p-1} \mathbf{I}_{\bar{n}} & Q_{p,p} \mathbf{I}_{\bar{n}} \end{pmatrix} \begin{pmatrix} \bar{\mathbf{U}} & 0 & \dots & 0 \\ 0 & \bar{\mathbf{U}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \bar{\mathbf{U}} \end{pmatrix} \quad (19)$$

where $\mathbf{I}_{\bar{n}} \in \mathbb{R}^{\bar{n} \times \bar{n}}$ is the identity matrix. It should be noted that to apply \mathbf{U} , the matrix $\bar{\mathbf{U}}$ needs to be applied p times. The resulting localization matrix is:

$$\mathbf{L} = \mathbf{U}\mathbf{U}^T = \mathbf{S}^s \begin{pmatrix} W_{1,1} \bar{\mathbf{U}}\bar{\mathbf{U}}^T & \dots & W_{1,p} \bar{\mathbf{U}}\bar{\mathbf{U}}^T \\ \vdots & \ddots & \vdots \\ W_{p,1} \bar{\mathbf{U}}\bar{\mathbf{U}}^T & \dots & W_{p,p} \bar{\mathbf{U}}\bar{\mathbf{U}}^T \end{pmatrix} \mathbf{S}^{sT} \quad (20)$$

where $\mathbf{W} = \mathbf{Q}\mathbf{Q}^T$. In this case, each auto- or cross-localization block is given by:

$$\mathbf{L}_{ij} = W_{i,j} \mathbf{S}_i^s \bar{\mathbf{U}}\bar{\mathbf{U}}^T \mathbf{S}_j^{sT} \quad (21)$$

The actual process would be reversed: first define \mathbf{W} , and its square-root decomposition \mathbf{Q} . Since \mathbf{W} and \mathbf{Q} are square matrices of size p , a Cholesky decomposition can work.

Advantages: precise control on the cross-localization amplitude.

Drawbacks: ill-adapted for large differences between variables length-scales.

Cost: p applications of $\bar{\mathbf{U}}$.

5 Applications

The approaches above can be used to localize 3D variables located on different points of the grid (e.g. Arakawa C grid).

It can also work to maintain a non-zero cross-localization between 3D and 2D variables, which is needed if they are correlated. The usual practice is to build an extended 3D grid with one more level at the bottom, and to copy the 2D variable on this extra level. The formalism described above can be used to understand this approach: the state grid interpolation is a simple copy of the 2D variable on the extra level, and a copy of the 3D variable on the remaining levels above.

Using a common block $\bar{\mathbf{U}}$ can lead to some issues when the 3D and 2D variables have very different length-scales (e.g. temperature and surface pressure). However, we can also use the state grid interpolation with two specific blocks $\bar{\mathbf{U}}_{3D}$ and $\bar{\mathbf{U}}_{2D}$, which are both defined on the extended grid.

The control grid interpolation approach is not suitable to cross-localize 3D and 2D variables, since it is not possible to define cross-localizations on their specific grids.