Multivariate localization

Benjamin Ménétrier

Last update: January 17, 2023

Contents

1 Framework

1.1 State vector definition

The state vector $\mathbf{x} \in \mathbb{R}^n$ is split into Q concatenated sub-vectors $\mathbf{x}_q$ (called "groups"), each one gathering P_q sub-vectors $\mathbf{x}_{p,q}$ of size n_q (called "variables"). Each sub-vector $\mathbf{x}_{p,q}$ corresponds to the data for a given physical variable (e.g. zonal wind, temperature, humidity, etc.). In a given group, all variables have the same size (n_q). Thus:

$$\mathbf{x} = \left(\begin{array}{c} \mathbf{x}_1 \\ \hline \vdots \\ \hline \mathbf{x}_Q \end{array}\right) \quad \text{and} \quad \mathbf{x}_q = \left(\begin{array}{c} \mathbf{x}_{1,q} \\ \hline \vdots \\ \hline \mathbf{x}_{P_q,q} \end{array}\right) \tag{1}$$

As a consequence:

- The total number of variables is $P = \sum_{q=1}^{Q} P_q$.
- Each group $\mathbf{x}_q$ has a size $P_q n_q$, so the full state vector size is $n = \sum_{q=1}^{Q} P_q n_q$.

The main idea behind this distribution of variables into groups is to apply the same localization to all variables within a given group.

1.2 Illustration setup

A simple example will be used as an illustration throughout this note. We define a vector $\mathbf{x}$ with $Q = 3$ groups, where $P_1 = 3$, $P_2 = 1$ and $P_3 = 2$:

$$\mathbf{x} = \left(\begin{array}{c} \mathbf{x}_{1,1} \\ \mathbf{x}_{2,1} \\ \mathbf{x}_{3,1} \\ \hline \mathbf{x}_{1,2} \\ \hline \mathbf{x}_{1,3} \\ \mathbf{x}_{2,3} \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} \mathbf{x}_1 \ (1^{\text{st}} \text{ group of variables}) \\ \left.\vphantom{a}\right\} \mathbf{x}_2 \ (2^{\text{nd}} \text{ group of variables}) \\ \left.\vphantom{\begin{array}{c}a\\a\end{array}}\right\} \mathbf{x}_3 \ (3^{\text{rd}} \text{ group of variables}) \end{array} \tag{2}$$

1.3 Localization square-root

The localization matrix $\mathbf{L} \in \mathbb{R}^{n \times n}$ is applied to $\mathbf{x}$. The main issue is to design a positive semi-definite $\mathbf{L}$ matrix, and the simplest solution is to build it as:

$$\mathbf{L} = \mathbf{U}\mathbf{U}^{\mathrm{T}} \tag{3}$$

where $\mathbf{U} \in \mathbb{R}^{n \times m}$ is called a "square-root" of the localization matrix $\mathbf{L}$. The control vector $\mathbf{v} \in \mathbb{R}^m$ is such that:

$$\mathbf{x} = \mathbf{U}\mathbf{v} \tag{4}$$

We can define a square-root component $\mathbf{U}_q \in \mathbb{R}^{n_q \times m_q}$ for each group.

2 Possible implementations

2.1 Univariate localization

For the case where there is no cross-localization between variables, the square-root $\mathbf{U}$ can be defined as:

$$\mathbf{U} = \left(\begin{array}{ccc|c|cc} \mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{U}_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{U}_1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \mathbf{U}_2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \mathbf{U}_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{U}_3 \end{array}\right) \tag{5}$$

Thus:

$$\mathbf{L} = \left(\begin{array}{ccc|c|cc} \mathbf{U}_1\mathbf{U}_1^\mathrm{T} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{U}_1\mathbf{U}_1^\mathrm{T} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{U}_1\mathbf{U}_1^\mathrm{T} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \mathbf{U}_2\mathbf{U}_2^\mathrm{T} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^\mathrm{T} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^\mathrm{T} \end{array}\right) \tag{6}$$

In this case, the total control vector size is $m = \sum_{q=1}^{Q} P_q m_q$

2.2 Duplicated localization

It is possible to define the cross-localization between the variables of a same group as duplicates of the diagonal localization, by defining $\mathbf{U}$ as:

$$
\begin{aligned}
\mathbf{U} &= \left(\begin{array}{c|c|c}
\mathbf{U}_1 & 0 & 0 \\
\mathbf{U}_1 & 0 & 0 \\
\mathbf{U}_1 & 0 & 0 \\
\hline
0 & \mathbf{U}_2 & 0 \\
\hline
0 & 0 & \mathbf{U}_3 \\
0 & 0 & \mathbf{U}_3
\end{array}\right) \\
&= \left(\begin{array}{c|c|c}
\mathbf{I}_1 & 0 & 0 \\
\mathbf{I}_1 & 0 & 0 \\
\mathbf{I}_1 & 0 & 0 \\
\hline
0 & \mathbf{I}_2 & 0 \\
\hline
0 & 0 & \mathbf{I}_3 \\
0 & 0 & \mathbf{I}_3
\end{array}\right)
\left(\begin{array}{c|c|c}
\mathbf{U}_1 & 0 & 0 \\
\hline
0 & \mathbf{U}_2 & 0 \\
\hline
0 & 0 & \mathbf{U}_3
\end{array}\right)
\end{aligned}
\tag{7}
$$

where $\mathbf{I}_q \in \mathbf{R}^{n_q,n_q}$ is the identity matrix of size n_q. Thus:

$$
\mathbf{L} = \left(\begin{array}{ccc|c|cc}
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\
0 & 0 & 0 & 0 & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}}
\end{array}\right)
\tag{8}
$$

In this case, the total control vector size is $m = \sum_{q=1}^{Q} m_q$

2.3 Duplicated and weighted localization

A variant of the previous case uses group-specific lower triangular matrices $\mathbf{S}^q \in \mathbb{R}^{P_q \times P_q}$ to define $\mathbf{U}$ as:

$$
\begin{aligned}
\mathbf{U} &= \left(\begin{array}{ccc|c|cc}
S^1_{1,1}\mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\
S^1_{2,1}\mathbf{U}_1 & S^1_{2,2}\mathbf{U}_1 & 0 & 0 & 0 & 0 \\
S^1_{3,1}\mathbf{U}_1 & S^1_{3,2}\mathbf{U}_1 & S^1_{3,3}\mathbf{U}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & S^2_{1,1}\mathbf{U}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & S^3_{1,1}\mathbf{U}_3 & 0 \\
0 & 0 & 0 & 0 & S^3_{2,1}\mathbf{U}_3 & S^3_{2,2}\mathbf{U}_3
\end{array}\right) \\
&= \left(\begin{array}{ccc|c|cc}
S^1_{1,1}\mathbf{I}_1 & 0 & 0 & 0 & 0 & 0 \\
S^1_{2,1}\mathbf{I}_1 & S^1_{2,2}\mathbf{I}_1 & 0 & 0 & 0 & 0 \\
S^1_{3,1}\mathbf{I}_1 & S^1_{3,2}\mathbf{I}_1 & S^1_{3,3}\mathbf{I}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & S^2_{1,1}\mathbf{I}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & S^3_{1,1}\mathbf{I}_3 & 0 \\
0 & 0 & 0 & 0 & S^3_{2,1}\mathbf{I}_3 & S^3_{2,1}\mathbf{I}_3
\end{array}\right)
\left(\begin{array}{ccc|c|cc}
\mathbf{U}_1 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{U}_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{U}_1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & \mathbf{U}_2 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \mathbf{U}_3 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{U}_3
\end{array}\right)
\end{aligned}
\tag{9}
$$

Thus:

$$
\mathbf{L} = \left(\begin{array}{ccc|c|cc}
W^1_{1,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{1,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{1,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
W^1_{2,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{2,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{2,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
W^1_{3,1}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{3,2}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & W^1_{3,3}\mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & W^2_{1,1}\mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & W^3_{1,1}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & W^3_{1,2}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\
0 & 0 & 0 & 0 & W^3_{2,1}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & W^3_{2,2}\mathbf{U}_3\mathbf{U}_3^{\mathrm{T}}
\end{array}\right)
\tag{10}
$$

where $\mathbf{W}^q = \mathbf{S}^q\mathbf{S}^{q\mathrm{T}}$ provides the weights applied to each covariance between variables within the group q. In practice, it is easier to set $\mathbf{W}^q$ and to compute $\mathbf{S}^q$ using a Cholesky decomposition. It also seems reasonable to set the diagonal elements of $\mathbf{W}^q$ to one, even if it is not mandatory.

In this case, the total control vector size is $m = \sum_{q=1}^{Q} P_q m_q$

2.4 Multivariate localization

For the case where the cross-localization between variables is implicitly given as the product of square-roots, $\mathbf{U}$ can be reduced to a column:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_1 \\ \mathbf{U}_1 \\ \hline \mathbf{U}_2 \\ \hline \mathbf{U}_3 \\ \mathbf{U}_3 \end{pmatrix} = \left(\begin{array}{c|c|c} \mathbf{I}_1 & 0 & 0 \\ \mathbf{I}_1 & 0 & 0 \\ \mathbf{I}_1 & 0 & 0 \\ \hline 0 & \mathbf{I}_2 & 0 \\ \hline 0 & 0 & \mathbf{I}_3 \\ 0 & 0 & \mathbf{I}_3 \end{array}\right) \begin{pmatrix} \mathbf{U}_1 \\ \hline \mathbf{U}_2 \\ \hline \mathbf{U}_3 \end{pmatrix} \tag{11}$$

An additional assumption is required here: the size of the group sub-vectors must be the same for all groups, i.e. $m_1 = ... = m_Q = \overline{m}$. Thus:

$$\mathbf{L} = \left(\begin{array}{ccc|c|cc} \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_1\mathbf{U}_3^{\mathrm{T}} \\ \hline \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_2\mathbf{U}_3^{\mathrm{T}} \\ \hline \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \\ \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_1^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_2^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_3\mathbf{U}_3^{\mathrm{T}} \end{array}\right) \tag{12}$$

In this case, the total control vector size is $m = \overline{m}$.

The amplitude of cross-localization is not one in this case. For $1 \leq q, q' \leq Q$, we denote $\mathbf{u}$ the k$^{\mathrm{th}}$ column of $\mathbf{U}_q$ and $\mathbf{u}'$ the k$^{\mathrm{th}}$ column of $\mathbf{U}_{q'}$. The k$^{\mathrm{th}}$ diagonal coefficient of $\mathbf{U}_q\mathbf{U}_{q'}^{\mathrm{T}}$ is given by $\langle \mathbf{u}, \mathbf{u}' \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product. The Cauchy-Schwartz inequality requires that:

$$|\langle \mathbf{u}, \mathbf{u}' \rangle| \leq \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{u}', \mathbf{u}' \rangle} \tag{13}$$

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

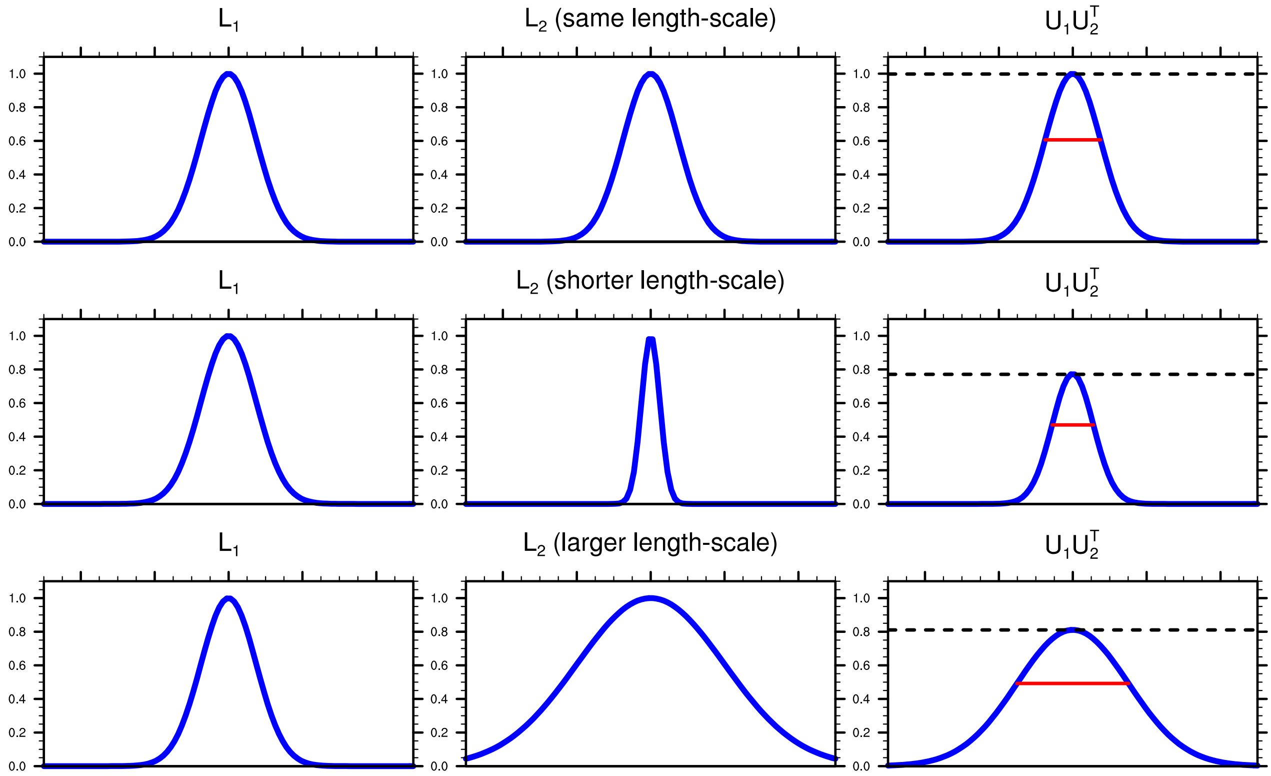

Figure 1: 1D illustration of the product of two localization square-root, with Gaussian localization functions. The dashed black line shows the function amplitude, while the solid red line shows the function length-scale.