Optimal linear transform for multi-scale localization

Travis Harty

We update our state $x \sim N(\mu, \mathbf{P})$ using observation $y = \mathbf{H}\mu + \varepsilon$, where $\varepsilon \sim N(0, \mathbf{R})$, $\dim x = N_x$, and $\dim y = N_y$. We will look for transformations $\mathbf{T}_{xl}$, $\mathbf{T}_{yl}$, $\mathbf{T}_{xs}$, and $\mathbf{T}_{ys}$ such that $x_l = \mathbf{T}_{xl}x$ and $y_l = \mathbf{T}_{yl}y$ contain long-scale information and $x_s = \mathbf{T}_{xs}x$ and $y_s = \mathbf{T}_{ys}y$ contains short-scale information.

We begin by calculating the eigenvalue decomposition of $\mathbf{P}$ and $\mathbf{R}$,

$$\mathbf{P} = \mathbf{Q}_x\mathbf{\Lambda}_x\mathbf{Q}_x^T \text{ and } \mathbf{R} = \mathbf{Q}_y\mathbf{\Lambda}_y\mathbf{Q}_y^T.$$

We do not assume that $\mathbf{P}$ or $\mathbf{R}$ are full rank and that the decompositions are in reduced form.

We can now reform the relationship between x and y. First we must notice that even if the image of $\mathbf{P}$ is equal to the image of $\mathbf{Q}_x$, this does not mean that μ is in the image of $\mathbf{Q}_x$. This means that if we project μ onto the preimage of $\mathbf{Q}_x$ we will lose the part of μ that is in the kernel of $\mathbf{Q}_x$ and therefore must add that part back using the projection onto the kernel of $\mathbf{Q}_x$

$$\mathbf{P}_x^{\mathrm{roj}} = \mathbf{I} - \mathbf{Q}_x\mathbf{Q}_x^T,$$

with the same statement being true for y. We therefore have

$$y = \mathbf{H}\mu + \varepsilon$$

$$\mathbf{Q}_y\mathbf{\Lambda}_y^{1/2}\mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T y + \mathbf{P}_y^{\mathrm{roj}} y = \mathbf{H}\left[\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mathbf{\Lambda}_x^{-1/2}\mathbf{Q}_x^T\mu + \mathbf{P}^{\mathrm{roj}}\mu\right] + \varepsilon$$

$$\mathbf{Q}_y\mathbf{\Lambda}_y^{1/2}\mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T y + \mathbf{P}_y^{\mathrm{roj}} y = \mathbf{H}\left[\mathbf{Q}_{lx}\mathbf{\Lambda}_{lx}^{1/2}\mathbf{\Lambda}_{lx}^{-1/2}\mathbf{Q}_{lx}^T\mu + \mathbf{Q}_{sx}\mathbf{\Lambda}_{sx}^{1/2}\mathbf{\Lambda}_{sx}^{-1/2}\mathbf{Q}_{sx}^T\mu + \mathbf{P}^{\mathrm{roj}}\mu\right] + \varepsilon$$

where

$$\mathbf{Q}_x\mathbf{\Lambda}_x\mathbf{Q}_x^T = \mathbf{Q}_{lx}\mathbf{\Lambda}_{lx}\mathbf{Q}_{lx}^T + \mathbf{Q}_{sx}\mathbf{\Lambda}_{sx}\mathbf{Q}_{sx}^T, \tag{1}$$

and $\mathbf{Q}_{lx}$ and $\mathbf{Q}_{sx}$ contain the eigenvectors that contain information about the long and short scales respectively. Setting $y' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T y$, $\mu_l' = \mathbf{\Lambda}_l^{-1/2}\mathbf{Q}_{lx}^T\mu$, and $\mu_s' = \mathbf{\Lambda}_s^{1/2}\mathbf{Q}_{sx}\mu$, we have

$$\mathbf{Q}_y\mathbf{\Lambda}_y^{1/2}y' + \mathbf{P}_y^{\mathrm{roj}} y = \mathbf{H}\mathbf{Q}_{lx}\mathbf{\Lambda}_{lx}^{1/2}\mu_l' + \mathbf{H}\mathbf{Q}_{sx}\mathbf{\Lambda}_{sx}^{1/2}\mu_s' + \mathbf{H}\mathbf{P}^{\mathrm{roj}}\mu + \varepsilon.$$

Furthermore, because $\mathbf{P}_y^{\mathrm{roj}} y$ is orthogonal to the columns of $\mathbf{Q}_y$, we have

$$y' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{Q}_{lx}\mathbf{\Lambda}_{lx}^{1/2}\mu_l' + \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{Q}_{sx}\mathbf{\Lambda}_{sx}^{1/2}\mu_s' + \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{P}^{\mathrm{roj}}\mu + \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\varepsilon. \tag{2}$$

In this form, we cannot perform data assimilation individually on mu_l' or μ_s' individually because they each are using the single y' observation. The question then becomes: Can we decompose y' into observations of only long and short scales as defined though $\mathbf{Q}_{xl}$ and $\mathbf{Q}_{xs}$? If we have any hope of this, we first require that the first three terms on the left hand side of Eq. (2) be orthogonal. To check this, we first reformulate Eq. (1):

$$\mathbf{Q}_x \mathbf{\Lambda}_x \mathbf{Q}_x^T = \tilde{\mathbf{Q}}_x \tilde{\mathbf{\Lambda}}_{lx} \tilde{\mathbf{Q}}_x^T + \tilde{\mathbf{Q}}_x \tilde{\mathbf{\Lambda}}_{sx} \tilde{\mathbf{Q}}_x^T, \tag{3}$$

where $\tilde{\mathbf{\Lambda}}_{lx}$ has zeros on the diagonal that correspond to short scale eigenvectors, $\tilde{\mathbf{\Lambda}}_{sx}$ has zeros on the diagonal that correspond to long scale eigenvectors, and both have zeros on the diagonal that correspond to eigenvectors that are in the kernel of $\mathbf{P}$. If we have $\tilde{\mu}_l = \tilde{\Lambda}_{lx}^{-1/2} \tilde{\mathbf{Q}}_x^T \mu$ and $\tilde{\mu}_s = \tilde{\Lambda}_{sx}^{-1/2} \tilde{\mathbf{Q}}_x^T \mu$ then we can see that

$$\mathbf{\Lambda}_y^{-1/2} \mathbf{Q}_y^T \mathbf{H} \mathbf{Q}_{lx} \mathbf{\Lambda}_{lx}^{1/2} \mu_l' = \mathbf{\Lambda}_y^{-1/2} \mathbf{Q}_y^T \mathbf{H} \tilde{\mathbf{Q}}_x \tilde{\mathbf{\Lambda}}_{lx}^{1/2} \tilde{\mu}_l$$

and

$$\mathbf{\Lambda}_y^{-1/2} \mathbf{Q}_y^T \mathbf{H} \mathbf{Q}_{sx} \mathbf{\Lambda}_{sx}^{1/2} \mu_l' = \mathbf{\Lambda}_y^{-1/2} \mathbf{Q}_y^T \mathbf{H} \tilde{\mathbf{Q}}_x \tilde{\mathbf{\Lambda}}_{sx}^{1/2} \tilde{\mu}_s.$$

We also have

$$\tilde{y}_l^T \tilde{y}_s = \tilde{\mu}_l^T \tilde{\mathbf{\Lambda}}_{lx}^{1/2} \tilde{\mathbf{Q}}_x^T \mathbf{H}^T \mathbf{Q}_y \mathbf{\Lambda}_y^{-1/2} \mathbf{\Lambda}_y^{-1/2} \mathbf{Q}_y^T \mathbf{H} \tilde{\mathbf{Q}}_x \tilde{\mathbf{\Lambda}}_{sx}^{1/2} \tilde{\mu}_s = 0, \tag{4}$$

because of the structures of $\tilde{\mathbf{\Lambda}}_{xl}$ and $\tilde{\mathbf{\Lambda}}_{xs}$.