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 $\mathbf{Q}_x$

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

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

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

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

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

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 $y' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T y$, $\mu_l' = \mathbf{\Lambda}_{xl}^{-1/2}\mathbf{Q}_{lx}^T\mu$, and $\mu_s' = \mathbf{\Lambda}_{xs}^{1/2}\mathbf{Q}_{sx}\mu$ we have

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

Furthermore, because $\mathbf{P}_y^{\text{roj}} y$ is orthogonal to the columns of $\mathbf{Q}_y$, setting $\varepsilon' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T$ we have

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

In this form, we cannot perform data assimilation on μ'_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 $\mathbf{Q}_{xs}$?

We can do this by taking the singular value decompositions of the matrices multiplying our μ's in Eq. 2

$$y' = \mathbf{U}_l \mathbf{\Sigma}_l \mathbf{V}_l^T \mu'_l + \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^T \mu'_s + \mathbf{U}_\perp \mathbf{\Sigma}_\perp \mathbf{V}_\perp^T \mu + \varepsilon'$$

If we then assume that $\mathbf{U}_l$, $\mathbf{U}_s$, and $\mathbf{U}_\perp$ are all orthogonal to each other, we can then decompose y'

$$\mathbf{U}_l^T y' = \mathbf{\Sigma}_l \mathbf{V}_l^T \mu'_l + \mathbf{U}_l^T \varepsilon'$$
$$\mathbf{U}_s^T y' = \mathbf{\Sigma}_s \mathbf{V}_s^T \mu'_s + \mathbf{U}_s^T \varepsilon'$$

This however will only be true it the long and short scales as defined by the eigenvectors of $\mathbf{P}$ can truly be assimilated separately when taking $\mathbf{R}$ and $\mathbf{H}$ into account.