Singular values of covariance matrices under localization

Travis Harty

1 Optimal linear transform

We update our state $x \sim N(\mu, \mathbf{P})$ using observation $y = \mathbf{H}\mu + \epsilon$, where $\epsilon \sim N(0, \mathbf{R})$, $\dim x = N_x$, and $\dim y = N_y$. We will perform this update by first transforming our variables using linear transformations $\mathbf{T}_x$ and $\mathbf{T}_y$ for x and y respectively such that $\mathbf{T}_x x = \tilde{x} \sim N(\tilde{\mu}, \mathbf{I})$ and $\mathbf{T}_y y = \tilde{y} = \mathbf{\Sigma}\tilde{\mu} + \tilde{\epsilon}$, where $\mathbf{\Sigma}$ is diagonal, and $\tilde{\epsilon} \sim N(0, \mathbf{I})$.

Take,

$$x' = \mathbf{P}^{-1/2}x \text{ and } y' = \mathbf{R}^{-1/2}y.$$

We then have,

$$\begin{aligned} y &= \mathbf{H}\mu + \epsilon \\ \mathbf{R}^{1/2}y' &= \mathbf{H}\mathbf{P}^{1/2}\mu' + \epsilon \\ y' &= \mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^{1/2}\mu' + \epsilon', \end{aligned}$$

where $\epsilon \sim N(0, \mathbf{I})$. We can then take the singular value decomposition of $\mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^{1/2}$ yielding,

$$\begin{aligned} y' &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mu' + \epsilon' \\ \mathbf{U}^T y' &= \mathbf{\Sigma}(\mathbf{V}^T\mu') + \tilde{\epsilon} \\ \tilde{y} &= \mathbf{\Sigma}\tilde{\mu} + \tilde{\epsilon} \end{aligned}$$

where $\mathbf{\Sigma}$ is diagonal and $\tilde{\epsilon} \sim N(0, \mathbf{I})$. We have $\mathbf{T}_x = \mathbf{V}^T\mathbf{P}^{-1/2}$ and $\mathbf{T}_y = \mathbf{U}^T\mathbf{R}^{-1/2}$.

2 Dimension reduction optimal linear transform

We again update our state $x \sim N(\mu, \mathbf{P})$ with observations $y = \mathbf{H}x + \epsilon$, where $\epsilon \sim N(0, \mathbf{R})$. We will now calculate

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

and choose N_{λ_x} and N_{λ_y} that are the number of eigenvalues to keep for P and R respectively such that,

$$\dim(\mathbf{Q}_x) = (N_x, N_{\lambda_x});\ \dim(\mathbf{\Lambda}_x) = (N_{\lambda_x}, N_{\lambda_x})$$

and

$$\dim(\mathbf{Q}_y) = (N_y, N_{\lambda_y});\ \dim(\mathbf{\Lambda}_y) = (N_{\lambda_y}, N_{\lambda_y})$$

We can then repeat the above calculations while reducing the transformed variables' dimensions. Take,

$$x' = \mathbf{\Lambda}_x^{-1/2}\mathbf{Q}_x^T x \text{ and } y' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T y$$

Note that we have,

$$x \approx \mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}x'$$
$$y \approx \mathbf{Q}_y\mathbf{\Lambda}_y^{1/2}y'$$

where there is equality when N_{λ_x} (N_{λ_y}) is the true rank of $\mathbf{P}$ ($\mathbf{R}$).

Assuming that N_{λ_x} and N_{λ_y} are the ranks of P and R respectively, we have,

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

$$\mathbf{Q}_y\mathbf{\Lambda}_y^{1/2}y' = \mathbf{H}\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mu' + \epsilon$$
$$y' = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mu' + \epsilon',$$

where $\epsilon' \sim N(0, \mathbf{I}_{N_{\lambda_y}})$. We can then take the singular value decomposition of $\mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}$ yielding,

$$y' = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mu' + \epsilon'$$
$$\mathbf{U}^T y' = \mathbf{\Sigma}(\mathbf{V}^T\mu') + \tilde{\epsilon}$$
$$\tilde{y} = \mathbf{\Sigma}\tilde{\mu} + \tilde{\epsilon}$$

where $\mathbf{\Sigma}$ is diagonal and $\tilde{\epsilon} \sim N(0, \mathbf{I}_{N_{\lambda_y}})$. We have $\mathbf{T}_x = \mathbf{V}^T\mathbf{\Lambda}_x^{-1/2}\mathbf{Q}_x^T$ and $\mathbf{T}_y = \mathbf{U}^T\mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T$. Note that we have,

$$\mathbf{T}_x\mathbf{T}_x^{-R} = \mathbf{V}^T\mathbf{\Lambda}_x^{-1/2}\mathbf{Q}_x^T\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mathbf{V} = \mathbf{I}_{N_{\lambda_x}},$$

but there is no left inverse. However, the right inverse is something like a left inverse in that

$$\mathbf{T}_x^{-R}\mathbf{T}_x = \mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mathbf{V}\mathbf{V}^T\mathbf{\Lambda}_x^{-1/2}\mathbf{Q}_x^T$$
$$\mathbf{T}_x^{-R}\mathbf{T}_x = \mathbf{Q}_x\mathbf{Q}_x^T$$

which is equal to the identity matrix only if P is full rank, and all eigenvectors are kept. However, if μ is in the span of the columns of $\mathbf{Q}_x$, then we have,

$$\mathbf{T}_x^{-R}\mathbf{T}_x\mu = \mathbf{Q}_x\mathbf{Q}_x^T\mu = \mu$$

even if $\mathbf{T}_x$ does not have a left inverse. Furthermore, if μ is not in the span of the columns of $\mathbf{Q}_x$, this is possible even if the range of $\mathbf{P}$ and $\mathbf{Q}_x$ are equal, then we have,

$$(\mathbf{T}_x^{-R}\mathbf{T}_x + \mathbf{I}_{N_x} - \mathbf{Q}_x\mathbf{Q}_x^T)\mu = \mathbf{Q}_x\mathbf{Q}_x^T\mu + \mu - \mathbf{Q}_x\mathbf{Q}_x^T\mu = \mu.$$

We can therefore always recover the parts of μ that are removed because of the representation of μ in the column space of $\mathbf{Q}_x$.

3 Multi-scale optimal transform

Suppose that we have two ensembles representing the state x. One ensemble, $\mathbf{X}_c$ will have many members, but will have a lower resolution and only represent large scale structures of the state. The other ensemble, $\mathbf{X}$ will have fewer members, but will have a finer resolution and therefore be able to represent fine scale structures of the state. We then have,

$$\dim \mathbf{X}_c = N_{x_c}, N_{e_c}$$

and

$$\dim \mathbf{X} = N_x, N_e$$

where $N_{x_c} < N_x$ and $N_{e_c} > N_e$.

Taking the singular value decomposition of the ensemble of coarse perturbations, $\mathbf{X}_c^*$, will give us the eigenvalues and eigenvectors of the sample covariance matrix derived from $\mathbf{X}_c$.

$$\mathbf{X}_c^* = \left(\mathbf{X}_c - \frac{1}{N}\bar{\mathbf{X}}_c\mathbf{1}_c\right) / \sqrt{N_{e_c} - 1}$$

$$\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}\mathbf{V}_{temp}^T = \mathbf{X}_c^*$$

$$\mathbf{Q}_x\mathbf{\Lambda}_x\mathbf{Q}_x^T = (\mathbf{X}_c^*)(\mathbf{X}_c^*)^T$$

where $\mathbf{1}_c$ is a matrix of all ones of with the same dimension as $\mathbf{X}_c$ and $\bar{\mathbf{X}}_c$ is the sample mean of $\mathbf{X}_c$. Assuming that we know $\mathbf{Q}_y$ and $\mathbf{\Lambda}_y$, we can then calculate $\mathbf{U}_c$ and $\mathbf{V}_c$,

$$\mathbf{U}_c\mathbf{\Sigma}_c\mathbf{V}_c^T = \mathbf{\Lambda}_y^{-1/2}\mathbf{Q}_y^T\mathbf{H}\mathbf{Q}_x\mathbf{\Lambda}_x^{1/2}.$$

This then allows us to use

4 Equivalence to standard KF

Sticking with the above notation, we can calculate the standard Kalman filter update and the corresponding equations:

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1},$$

$$\mu^a = \mu + \mathbf{K}(y - \mathbf{H}\mu),$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}.$$

Alternatively, in terms of our transformed equations, we have:

$$\tilde{\mathbf{K}} = \mathbf{I}_{N_{\lambda_x}}\mathbf{\Sigma}^T(\mathbf{\Sigma}\mathbf{I}_{N_{\lambda_x}}\mathbf{\Sigma}^T + \mathbf{I}_{N_{\lambda_y}})^{-1} = \mathbf{\Sigma}^T(\mathbf{\Sigma}\mathbf{\Sigma}^T + \mathbf{I}_{N_{\lambda_y}})^{-1}$$

$$\tilde{\mu}^a = \tilde{\mu} + \tilde{\mathbf{K}}(\tilde{y} + \mathbf{\Sigma}\tilde{\mu})$$

$$\tilde{\mathbf{P}}^a = (\mathbf{I}_{N_{\lambda_x}} - \tilde{\mathbf{K}}\mathbf{\Sigma})\tilde{\mathbf{P}}.$$

We can then convert $\tilde{x}^a$ and $\tilde{\mathbf{P}}^a$ back to the original x-space through,

$$\mu_1^a = \mathbf{T}_x^{-R} \tilde{\mu}^a$$

and

$$\mathbf{P}_1^a = \mathbf{T}_x^{-R} \tilde{\mathbf{P}}^a \left(\mathbf{T}_x^{-R}\right)^T .$$

Assuming that the ranks of $\mathbf{P}$ and $\mathbf{R}$ are N_{λ_x} and N_{λ_y} respectively,

$$\tilde{\mathbf{K}} = \mathbf{\Sigma}^T (\mathbf{\Sigma}\mathbf{\Sigma}^T + \mathbf{I}_{N_{\lambda_y}})^{-1}$$

$$\tilde{\mathbf{K}} = \mathbf{T}_x \mathbf{P} \mathbf{T}_x^T (\mathbf{T}_y \mathbf{H} \mathbf{T}_x^{-R})^T \left((\mathbf{T}_y \mathbf{H} \mathbf{T}_x^{-R})(\mathbf{T}_y \mathbf{H} \mathbf{T}_x^{-R})^T + \mathbf{T}_y \mathbf{R} \mathbf{T}_y^T \right)^{-1}$$

$$\tilde{\mathbf{K}} = \mathbf{T}_x \mathbf{P} \mathbf{T}_x^T \left(\mathbf{T}_x^{-R}\right)^T \mathbf{H}^T \mathbf{T}_y^T \left(\mathbf{T}_y \mathbf{H} \mathbf{T}_x^{-R} \left(\mathbf{T}_x^{-R}\right)^T \mathbf{H}^T \mathbf{T}_y^T + \mathbf{T}_y \mathbf{R} \mathbf{T}_y^T \right)^{-1}$$

$$\tilde{\mathbf{K}} = \mathbf{T}_x \mathbf{P} \mathbf{Q}_x \mathbf{Q}_x^T \mathbf{H}^T \mathbf{T}_y^T \left(\mathbf{T}_y \left(\mathbf{H} \mathbf{Q}_x \Lambda_x \mathbf{Q}_x^T \mathbf{H}^T + \mathbf{R} \right) \mathbf{T}_y^T \right)^{-1}$$

If we assume that μ is in the span of $\mathbf{Q}_x$, we have,

$$x_1^a = \mathbf{T}_x^{-R} \tilde{x}^a$$

$$x_1^a = \mathbf{T}_x^{-R} \left(\tilde{x} + \tilde{K} \right)$$