Regularization for Empirical Orthogonal Functions

August 4, 2020

Unregularized problem We want to solve

$$\min_{\mathbf{F}^T} ||\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T||_2^2 \tag{1}$$

With $\mathbf{D}$ the $nm \times l$ data matrix, $\mathbf{F}$ the $1 \times nm$ filter and $\mathbf{P}$ the $1 \times l$ a-posteriori matrix.

To find a global minimizer, we rewrite Eq. (1) and set to 0 the deerivative with respect to $\mathbf{F}^T$.

$$\begin{aligned}
\frac{\partial||\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T||_2^2}{\partial\mathbf{F}^T} &= \frac{\partial(\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T)^T(\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T)}{\partial\mathbf{F}^T} = \mathbf{0}_{nm} \\
\Leftrightarrow 2\mathbf{D}\mathbf{D}^T\mathbf{F}^T - 2\mathbf{D}\mathbf{P}^T &= \mathbf{0}_{nm} \\
\Leftrightarrow \mathbf{F}^T &= \underbrace{(\mathbf{D}\mathbf{D}^T)^{-1}\mathbf{D}}_{(\mathbf{D}^T)^+}\mathbf{P}^T
\end{aligned}$$

Which leads to the optimal filter

$$\mathbf{F} = ((\mathbf{D}^T)^+\mathbf{P}^T)^T \tag{2}$$

Regularized problem in regular form We add a Tikhonov regularization of the filter

$$\min_{\mathbf{F}^T} ||\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T||_2^2 + \lambda||\mathbf{F}||_2^2 \tag{3}$$

Similarly, the solution is

$$\mathbf{F} = ((\mathbf{D}\mathbf{D}^T + \lambda I_{nm})^{-1}\mathbf{D}\mathbf{P}^T)^T \tag{4}$$

With I_{nm} the $nm \times nm$ identity matrix. Unfortunately, we can't recognize a pseudo-inverse like we did above. We can't rewrite a problem in a way that allows us to compute the SVD of $\mathbf{D}^T$ then take its pseudo-inverse.

Regularized problem in modified form To avoid this problem, Guyon rewrites the regularization term in the original quadratic term

$$\min_{\mathbf{F}^T} ||\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T||_2^2 + \lambda||\mathbf{F}||_2^2 = \min_{\mathbf{F}^T} ||(\underbrace{\mathbf{D}|\lambda I_{nm}}_{\mathbf{D}_R})^T\mathbf{F}^T - (\underbrace{\mathbf{P}|\mathbf{0}_{nm}}_{\mathbf{P}_R})^T||_2^2 \tag{5}$$

By expanding $||(a\ b)||_2^2$ into $a^2 + b^2$ we end up as the same expression than Eq. (3). The solution is now

$$\mathbf{F} = ((\mathbf{D}_R^T)^+\mathbf{P}_R^T)^T \tag{6}$$

We can now use pseudo-inverses but the matrix dimensions is much larger (m is the number of pixels in the dark zone).