3.3 The Full Procrustes Mean in a fixed basis

To avoid having to sample the estimated covariance surface $\hat{C}(s,t)$ on a large grid when calculating its leading eigenfunction, it might be preferable to calculate this eigenfunction from the vector of basis coefficients directly. After choosing a basis representation $b = (b_1, \ldots, b_k)$ with $b_j : \mathbb{R} \to \mathbb{R}$ real-valued basis functions, we want to estimate complex coefficients $\theta_j \in \mathbb{C}$ so that the Full Procrustes mean of SRV curves is given by $\hat{\mu}(t) = \sum_{j=1}^{k} \hat{\theta}_j b_j(t) = b^T \hat{\theta}$:

$$
\begin{aligned}
\hat{\mu} &= \underset{\theta:||b^T\theta||=1}{\text{argmax}} \sum_{i=1}^{n} \langle b^T\theta, q_i \rangle \langle q_i, b^T\theta \rangle \\
&= \underset{\theta:||b^T\theta||=1}{\text{argmax}} \sum_{k,l} \sum_{i=1}^{n} \langle b_k\theta_k, q_i \rangle \langle q_i, b_l\theta_l \rangle \\
&= \underset{\theta:||b^T\theta||=1}{\text{argmax}} \sum_{k,l} \bar{\theta}_k \theta_l \sum_{i=1}^{n} \langle b_k, q_i \rangle \langle q_i, b_l \rangle \\
&= \underset{\theta:||b^T\theta||=1}{\text{argmax}} \, \theta^H S \theta
\end{aligned}
$$

where the matrix $S = \{\sum_{i=1}^{n} \langle b_k, q_i \rangle \langle q_i, b_l \rangle\}_{k,l}$ has to be estimated from the observed SRV curves. We can further simplify S to

$$
\begin{aligned}
S_{kl} &= \sum_{i=1}^{n} \int_0^1 \bar{b}_k(t) q_i(t) dt \int_0^1 \bar{q}_i(s) b_l(s) ds \\
&= \int_0^1 \int_0^1 \bar{b}_k(t) \underbrace{\left(\sum_{i=1}^{n} q_i(t) \bar{q}_i(s) \right)}_{=n\,\hat{C}(s,t)} b_l(s) ds dt \\
&= n \int_0^1 \int_0^1 \bar{b}_k(t) \hat{C}(s,t) b_l(s) ds dt
\end{aligned}
$$

with $\hat{C}(s,t) = \frac{1}{n} \sum_{i=1}^{n} q_i(s) \overline{q_i(t)}$ the sample analogue to the complex population covariance function $C(s,t) = \mathbb{E}[q(s)\overline{q(t)}]$. We may estimate $C(s,t)$ via tensor product splines, so that $\hat{C}(s,t) = \sum_{k,l} \hat{\xi}_{kl} b_k(t) b_l(s)$, where $b_j(t)$, $j = 1, \ldots, k$ are the same real valued basis functions as used for the mean and $\hat{\xi}_{kl}$ are the estimated complex coefficients.

We can then further simplify S_{kl}

$$
\begin{aligned}
S_{kl} =& n \int_0^1 \int_0^1 b_k(t) \left(\sum_{p,q} \hat{\xi}_{pq} b_q(t) b_p(s) \right) b_l(s) ds dt \\
=& n \sum_{p,q} \hat{\xi}_{pq} \int_0^1 \int_0^1 b_k(t) b_q(t) b_p(s) b_l(s) ds dt \\
=& n \sum_{p,q} \hat{\xi}_{pq} \langle b_k, b_q \rangle \langle b_p, b_l \rangle \\
=& n \sum_{p,q} \hat{\xi}_{pq} g_{kq} g_{pl}
\end{aligned}
$$

where g_{ij}, $i, j = 1, \ldots, k$ are the elements of the Gram matrix $G = bb^T$ with $G = \mathbb{I}_k$ in the special case of an orthogonal basis. We can then write the write the matrix S as a function of the estimated coefficient matrix $\hat{\Xi} = (\hat{\xi}_{ij})_{i,j=1,\ldots,k}$:

$$
S =n \, G \hat{\Xi} G
$$

The full Procrustes mean of SRV curves is then given by the solution to the optimization problem

$$
\begin{aligned}
\hat{\mu} &= \underset{\theta}{\operatorname{argmax}} \, n \, \theta^H G \hat{\Xi} G \theta \quad \text{subj. to} \quad ||b^T \theta|| = 1 \\
&= \underset{\theta : ||b^T \theta|| = 1}{\operatorname{argmax}} \; \theta^H G \hat{\Xi} G \theta \quad \text{subj. to} \quad \theta^H G \theta = 1
\end{aligned}
$$

One may solve this by using Lagrange optimization with the Langrangian

$$
\mathcal{L}(\theta, \lambda) = \theta^H G \hat{\Xi} G \theta - \lambda (\theta^H G \theta - 1)
$$

$$\mathcal{L}(\Theta,\lambda) = \Theta^H G \hat{\Xi} G \Theta - \lambda(\Theta^H G \Theta - 1)$$

- Instead of calculating the derivative wrt. Θ, we can treat $\mathbb{C}^k \equiv \mathbb{R}^{2k}$ and calculate the derivatives wrt. $Re(\Theta)$ and $Im(\Theta)$.

$$\begin{aligned}\mathcal{L}(Re(\Theta), Im(\Theta), \lambda) = &\ Re(\Theta)^T GSG\, Re(\Theta) + i\, Re(\Theta)^T GSG\, Im(\Theta) \\ &- i\, Im(\Theta)^T GSG\, Re(\Theta) + Im(\Theta)^T GSG\, Im(\Theta) \\ &- \lambda\left(Re(\Theta)^T G\, Re(\Theta) + Im(\Theta)^T G\, Im(\Theta) - 1 \right)\end{aligned}$$

Then:

(1) $\frac{\partial \mathcal{L}}{\partial Re(\Theta)} = G(S+S^T)G\, Re(\Theta) + i\, G(S-S^T)G\, Im(\Theta) - 2\lambda G Re(\Theta) \overset{!}{=} 0$

(2) $\frac{\partial \mathcal{L}}{\partial Im(\Theta)} = G(S+S^T)G\, Im(\Theta) + i\, G(S-S^T)G\, Re(\Theta) - 2\lambda G\, Im(\Theta) \overset{!}{=} 0$

using $S+S^T = 2R$, $S-S^T = 2i\underline{I}$:

(1)' $GRG\, Re(\Theta) - GIG\, Im(\Theta) = \lambda G\, Re(\Theta)$

(2)' $GRG\, Im(\Theta) - GIG\, Re(\Theta) = \lambda G\, Im(\Theta)$

or:

$$\begin{pmatrix} GRG & -GIG \\ -GIG & GRG \end{pmatrix} \begin{pmatrix} Re(\Theta) \\ Im(\Theta) \end{pmatrix} = \lambda \cdot \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} Re(\Theta) \\ Im(\Theta) \end{pmatrix}$$

with the solution being the leading eigenvector of

$$\begin{pmatrix} G^{-1} & 0 \\ 0 & G^{-1} \end{pmatrix} \begin{pmatrix} GRG & -GIG \\ -GIG & GRG \end{pmatrix} = \begin{pmatrix} RG & -\underline{I}G \\ -\underline{I}G & RG \end{pmatrix}$$