Data Processing

October 28, 2022

1. Pre-process (only for our data) the renogram curves by dividing: a) each baseline renogram curve by its maximum; and b) each post-furosemide (diuretic) renogram curve by the maximum of the baseline renogram curve:

$$X_i^{(1)P} = \frac{X_i^{(1)}}{\max_t\{X_i^{(1)}(t)\}} \quad \text{and} \quad X_i^{(2)P} = \frac{X_i^{(2)}}{\max_t\{X_i^{(1)}(t)\}}, \qquad i = 1, \ldots, n$$

2. Normalize the renogram renogram curves:

$$X_i^{(1)*} = \frac{X_i^{(1)P} - \overline{X}^{(1)P}}{\sqrt{\int \mathrm{Var}\{X_i^{(1)P}(t)\}dt}} \quad \text{and} \quad X_i^{(2)*} = \frac{X_i^{(2)P} - \overline{X}^{(2)P}}{\sqrt{\int \mathrm{Var}\{X_i^{(2)P}(t)\}dt}}$$

3. Normalize the scalar predictors $\mathbf{Z}_i = (Z_{i1}, \ldots Z_{ip})^\top$

$$Z_{ir}^S = \frac{Z_{ir} - \overline{Z}_p}{s(Z_r)}, \qquad i = 1, \ldots, n, \; r = 1, \ldots, p.$$

4. Scale the scalar predictors so that the variability *between* the functional and scalar predictors are comparable (between-normalization):

$$\mathbf{Z}_i^* = \omega^{1/2}\mathbf{Z}_i^S$$

where ω is a scaling factor:

$$\omega = \frac{\sum_{i=1}^n \|X_i^*\|_{\mathcal{F}}^2}{\sum_{i=1}^n \|\mathbf{Z}_i^S\|^2} = \frac{\sum_{i=1}^n [\int\{X_i^{(1)*}(t_1)\}^2 dt_1 + \int\{X_i^{(2)*}(t_2)\}^2 dt_2]}{\sum_{i=1}^n \mathbf{Z}_i^{S\top}\mathbf{Z}_i^S}.$$

5. Run the partial least squares with $\mathbf{W}_i = (X_i^*, \mathbf{Z}_i^*)$. Make sure that you input testing data are normalized based on scaling factors computed from the training data.