Smoothing Splines

Write $g(x) = \sum_{i=1}^{n} \beta_i h_i(x)$ where h_i's are basis functions for NCS with knots at $x_1, \ldots, x_n$.

$$\sum_{i=1}^{n} [y_i - g(x_i)]^2 = (\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^t (\mathbf{y} - \mathbf{F}\boldsymbol{\beta}),$$

where $\mathbf{F}_{n\times n}$ with $\mathbf{F}_{ij} = h_j(x_i)$.

$$\begin{aligned}\int_a^b \left[g''(x)\right]^2 dx &= \int \Big[\sum_i \beta_i h_i''(x)\Big]^2 dx \\ &= \sum_{i,j} \beta_i \beta_j \int h_i''(x) h_j''(x) dx = \boldsymbol{\beta}^t \Omega \boldsymbol{\beta},\end{aligned}$$

where $\Omega_{n\times n}$ with $\Omega_{ij} = \int_a^b h_i''(x) h_j''(x) dx$.

So

$$\mathsf{RSS}_\lambda(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^t(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}) + \lambda\boldsymbol{\beta}^t\Omega\boldsymbol{\beta},$$

and the solution is

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= \arg\min_{\boldsymbol{\beta}} \mathsf{RSS}_\lambda(\boldsymbol{\beta}) \\
&= (\mathbf{F}^t\mathbf{F} + \lambda\Omega)^{-1}\mathbf{F}^t\mathbf{y} \\
\hat{\mathbf{y}} &= \mathbf{F}(\mathbf{F}^t\mathbf{F} + \lambda\Omega)^{-1}\mathbf{F}^t\mathbf{y} \\
&= S_\lambda\mathbf{y}
\end{aligned}$$

What if we use a different set of basis functions $\tilde{h}_1(x), \ldots, \tilde{h}_n(x)$?

- Demmler & Reinsch (1975): a basis with double orthogonality property, i.e.

$$\mathbf{F}^t\mathbf{F} = \mathbf{I}, \quad \Omega = \text{diag}(d_i),$$

where d_i's are arranged in an increasing order and in addition $d_2 = d_1 = 0$ (Why?).

- Using this basis, we have

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{F}^t\mathbf{F} + \lambda\Omega)^{-1}\mathbf{F}^t\mathbf{y} \\ &= (\mathbf{I} + \lambda\text{diag}(d_i))^{-1}\mathbf{F}^t\mathbf{y}, \end{aligned}$$

i.e.,

$$\hat{\beta}_i = \frac{1}{1+\lambda d_i}\hat{\beta}_i^{(\text{LS})}.$$

- Smoother matrix S_λ

$$\hat{\mathbf{y}} = \mathbf{F}\hat{\boldsymbol{\beta}} = \mathbf{F}(\mathbf{F}^t\mathbf{F} + \lambda\Omega)^{-1}\mathbf{F}^t\mathbf{y} = S_\lambda \mathbf{y}.$$

- Using D&R basis,

$$S_\lambda = \mathbf{F}\text{diag}\Big(\frac{1}{1+\lambda d_i}\Big)\mathbf{F}^t.$$

  So columns of $\mathbf{F}$ are the eigen-vectors of S_λ, which does not depend on λ.

- Effective df of a smoothing spline:

$$df(\lambda) = \text{tr}S_\lambda = \sum_{i=1}^{n} \frac{1}{1+\lambda d_i}.$$

- Check the R page to see what the DR basis functions look like.