(5.3) Aditivni modeli [$p > 1$]

Ako je $p > 1$:

- lok. regresija $\xrightarrow{npr.}$ $K_\lambda(x, x^{(i)}) = D\left(\frac{\|x - x^{(i)}\|_2}{h_\lambda(x)}\right)$
- splajnovi $\xrightarrow{npr.}$ za $p=2$, tražimo $f: \mathbb{R}^2 \to \mathbb{R}$ koja minimizira

$$\sum_{i=1}^{m} (y_i - f(x_i z))^2 + \lambda \boxed{\int\left[\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial z}\right)^2 + \left(\frac{\partial^2 f}{\partial z^2}\right)^2\right] dx\, dz}$$

↳ rješenje je tzv. thin-plate spline

(može se proširiti i na $p \geq 3$)

↳ problem za velike p ako nema dodatnih pretpostavki

[curse of dimens.]

↳ Aditivan model "pretpostavlja"

$$f^*(x) = E[Y \mid X = x] = \alpha + \sum_{j=1}^{p} f_j(x_j), \quad x = (x_1, \ldots, x_p) \in \mathbb{R}^p \qquad (5.32)$$

gdje su $f_1, \ldots, f_p$ proizvoljne (glatke) f-je.

[veća fleksibilnost od lin. modela, ali i dalje moguća interpretacija!]

Nap. 5.10 Za $c_j \in \mathbb{R}$ proizvoljne, $\tilde{f}_j := f_j + c_j$, $\forall j = 1, \ldots, p$

$$\Rightarrow \underbrace{\alpha - \sum_{j=1}^{p} c_j}_{=:\tilde{\alpha}} + \sum_{j=1}^{p} \tilde{f}_j(x_j) = \alpha + \sum_{j=1}^{p} f_j(x_j), \quad \forall x \in \mathbb{R}^p$$

↳ $\alpha, f_1, \ldots, f_p$ nisu jedinstveno određene

$\Rightarrow$ tipično pretpostavljamo $\boxed{\frac{1}{m}\sum_{i=1}^{m} f_j(x_{ij}) = 0, \quad \forall j = 1, \ldots, p}$ (5.33)

5.3.1 Backfitting algoritam

- fiksiramo metodu S ("smoother") koja svakom skupu za trening $\{(z_i, w_i) : i=1,\dots,n\}$ pridružuje $\hat{f}: \mathbb{R} \to \mathbb{R}$ $(\hat{w} := \hat{f}(z))$

npr. S je smoothing spline $\rightarrow$ fiksnim df ili df odabranim npr. GCV.

Backfitting:

(1.) $\hat{\alpha} := \frac{1}{n} \sum_{i=1}^{n} y_i =: \bar{y}$, $\hat{f}_j \overset{\text{npr.}}{:=} 0$, $\forall j = 1,\dots,p$ (ne mijenja se)

(2.) ponavljaj za $j = 1,2,\dots,p,1,2,\dots,p,\dots$

(a) $r_i := y_i - \hat{\alpha} - \sum_{\ell \neq j} \hat{f}_\ell(x_{i\ell})$, $i=1,\dots,n$

(b) $\hat{f}_j := S(\{(x_{ij}, r_i) : i=1,\dots,n\})$

(c) $\hat{f}_j := \hat{f}_j - \frac{1}{n}\sum_{i=1}^{n} \hat{f}_j(x_{ij})$ zbog (5.33),

sve dok se $\hat{f}_j$ ne mijenjaju za manje od δ.

$$\rightarrow \boxed{\hat{f}(x) := \bar{y} + \sum_{j=1}^{p} \hat{f}_j(x_j), \quad x \in \mathbb{R}^p}$$

Nap. 5.11 (i) Možemo koristiti različite metode S_j, $\forall j=1,\dots,p$.

(ii) Neke f_j možemo modelirati linearnim ili nekim drugim parametarskim modelom

$\rightarrow$ npr. $X = (\underset{\mathbb{R}^{p_1}}{Z}, \underset{\{0,1,2\}}{V}, \underset{\mathbb{R}}{W})$ te $\sum_{k=0}^{2} \alpha_k 1_{\{k=v\}}$

$$\mathbb{E}[Y \mid X=(z,v,w)] = z^T\beta + \alpha_v + f(w)$$

$\in \mathbb{R}^{p_1+2}$ — modeliramo neparametarski

(iii) Nelinearne interakcije: npr. $X = (Z, W_1, W_2)$ te

npr. koristimo lok. lin. regresiju.

$$\mathbb{E}[Y \mid X=(z,w_1,w_2)] = z^T\beta + \boxed{f(w_1,w_2)}$$