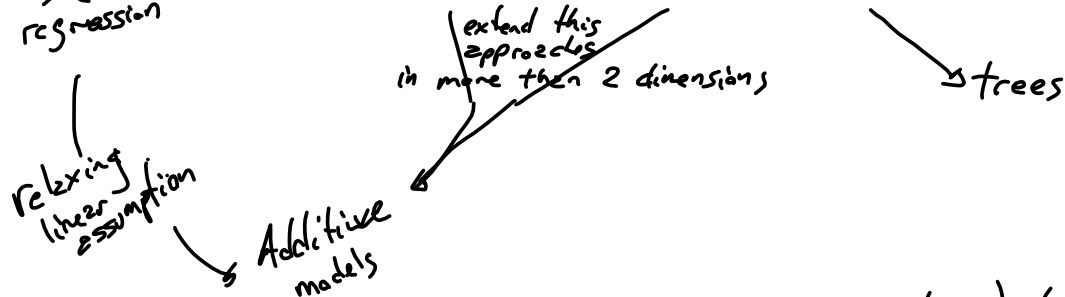
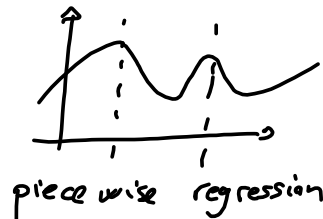
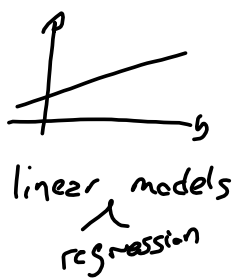


Additive models (GAM)



- nonparametric regression techniques (local regression, splines) do not really work in higher dimensions
- global approaches do not suffer from this issue (curse of dimensionality) because we impose structure

We want to combine the flexibility of the former and the possibility to use the technique in higher dimensions of the latter

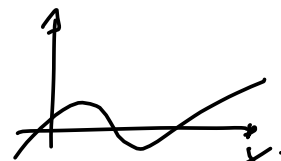
$$f(x) = f(x_1, \dots, x_p) = \alpha + \sum_{j=1}^p f_j(x_j)$$

lin regr: $\beta_j x_j$
spline

where $f_j(x_j)$ is the function that models the effect of x_j on the response, α is the intercept.

Due to identifiability issues, we need to add the constraint

$$\sum_{i=1}^n \hat{f}_j(x_{ij}) = 0 \quad \forall j = 1, \dots, p$$



E.g., using splines

$$\hat{f}_A(x) = \alpha_A + \underbrace{2 + 3x_1 + 1.5(x_1 - \bar{x}_1)_+ + 0.5(x_1 - \bar{x}_2)_+}_{\hat{f}_1^A(x_1)} + \underbrace{1 + 2x_2 + 3(x_2 - \bar{x}_3)_+ - 4(x_2 - \bar{x}_4)_+}_{\hat{f}_2^A(x_2)}$$

$$\hat{f}_B(x) = \alpha_B + \underbrace{3 + 3x_1 + 1.5(x_1 - \bar{x}_1)_+ + 0.5(x_1 - \bar{x}_2)_+}_{\hat{f}_1^B(x_1)} - \underbrace{1 + 2x_2 + 3(x_2 - \bar{x}_3)_+ - 4(x_2 - \bar{x}_4)_+}_{\hat{f}_2^B(x_2)}$$

$x_1 = 9$
 $x_2 = 9$

$\alpha = \bar{y}$

So the parameter α is uniquely equal to $\bar{y}$

To fit the additive model we use the backfitting algorithm:

1 - initialize $\alpha = \bar{y}$, $\hat{f}_j(x_j) = 0 \quad \forall j$

2 - in cycle, $j=1, 2, \dots, P, 1, 2, \dots, P, 1, \dots$

• estimate $\hat{f}_j$ using the preferred technique, keeping all other terms fixed

e.g, using splines $S(y_i - \hat{\alpha} - \sum_{k \neq j} \hat{f}_k(x_k))_{i=1, \dots, n}$

• update $\hat{f}_j \leftarrow \hat{f}_j + \hat{f}_j^*(x_{ij})$ ← theoretically not necessary
useful in practice

until all $\hat{f}_j$ stabilize

steps of the backfitting algorithm

terms	init	1	2	3	4	...	j^*
α	$\bar{y}$	$\bar{y}$	$\bar{y}$	$\bar{y}$	$\bar{y}$	...	$\bar{y}$
$f_1(x_1)$	0	$\hat{f}_1(x_1)$	$\hat{f}_1(x_1)$	$\hat{f}_1(x_1)$	$\hat{f}_1(x_1)$	...	.
$f_2(x_2)$	0	0	$\hat{f}_2(x_2)$	$\hat{f}_2(x_2)$	$\hat{f}_2(x_2)$	...	.
$f_3(x_3)$	0	0	0	$\hat{f}_3(x_3)$	$\hat{f}_3(x_3)$	...	.

Note: Red circles highlight the diagonal elements $\hat{f}_j(x_j)$ and a red 'X' is drawn under the first column of the f_j rows.

S is a general method, often splines

• we can add interactions terms,

$$f(x_1, \dots, x_p) = \alpha + \sum_{j=1}^p f_j(x_j) + \sum_{j=1}^p \sum_{k < j} f_{jk}(x_j, x_k) + \dots$$

• we can have mixture of linear and non-linear effects

$$f(x_1, \dots, x_p) = \alpha + (x_1 \dots x_{p-2}) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-2} \end{pmatrix} + f_{p-1}(x_{p-1}) + f_p(x_p)$$

$$= \alpha + X\beta + f_{z_1}(z_1) + f_{z_2}(z_2)$$

$X = (x_1, \dots, x_{p-2})$
 $z_1 = x_{p-1}$
 $z_2 = x_p$