$f^{-k}(x) \;=\; \sum_{l \neq k} \beta^l b^l(x)$

Algorithm BackFitting

Input M, labeled training set $\mathcal{D}$

Initialize $b^{1:M} = 0, \beta^{1:M} = 0$

repeat

for $k = 1, 2, \ldots M$

calculate $r^k(x^i) = y^i - f^{-k}(x^i)$, $i = 1 : N$

fit k-th predictor $\beta^k, b^k = \text{argmin} \hat{L}(f^{-k} + \beta b)$

until change in $\hat{L}$ small enough (or, change in $b^{1:M}$ small enough)

Output $f(x) \;=\; \sum_{k=1}^{M} \beta^k b^k(x)$