Math 342W Lecture 15

$g(x) = a_1(x_1)^n + \dots + a_p(x_p)^n$, $n \geq 1$

↳ a_j's are continuous functions and this $g(x)$ is called a general additive model

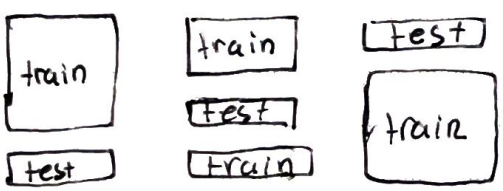
Interaction of features

By interacting features we are able to capture differential slopes of the given features

→ Pretty hard to overfit with a large enough (n)

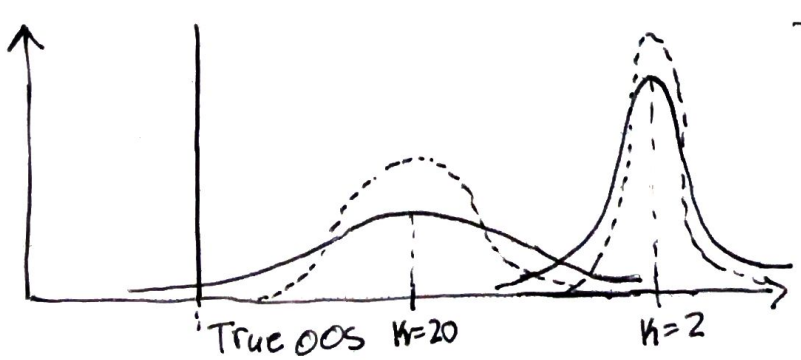
Validation / Cross Validation

The goal of these validations is to lower the variance in our metrics, particularly out of sample error metrics.



Each test from the folds returns $\vec{e}_i$ for the i^{th} fold

$$\vec{e}_{cv} = \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \\ \end{bmatrix} \Rightarrow OOS\_SE = \sqrt{\frac{1}{n}\sum(\vec{e}_i - \bar{e})^2} \Rightarrow \begin{matrix} OOS\_SE_1 \\ \vdots \\ OOS\_SE_n \end{matrix} \Rightarrow \sqrt{\frac{1}{k-1}\sum(SE_\ell - \overline{SE})^2}$$



——— = Before C.V

- - - - = After C.V

⇒ Less Variance in $\frac{OOS}{2}$