University of Toronto- Time series club
Lecture 2
Introduction to function fitting

Lim, Kyuson

05/18/2022

Today's outline

- Understand function fitting: parametric and nonparametric methods.

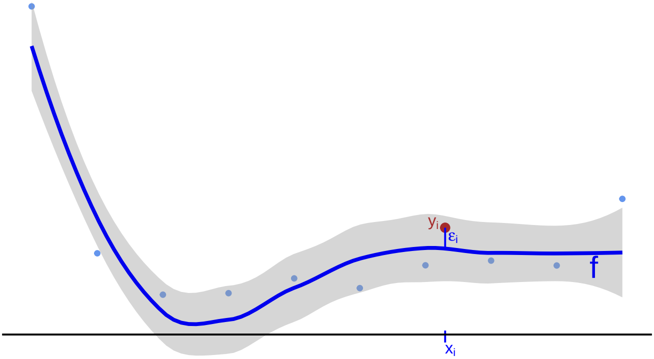
Function fitting

- Model an output as a function of many inputs,
  - Transcriptomic profiles -> Chance of developing disease.
  - Loss triangle -> Loss reserving.
  - Songs profiles -> Song(s) recommendation.
  - Image pixels -> What is in the image.
- Let $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^T$ be all the inputs.
- Let $f(\boldsymbol{x}_i)$ be input to output relationship.

A diagram

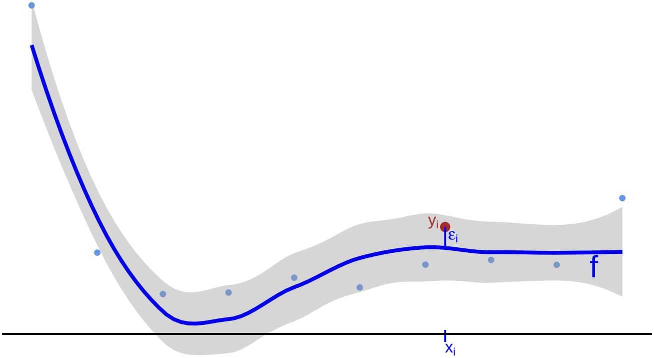
- We allow y_i to be different from $f(\boldsymbol{x}_i)$, perturbed by an observation-specific noise ϵ_i.

$$y_i = f(\boldsymbol{x}_i) + \epsilon_i.$$



Why this decomposition

- f represents systematic information that inputs provide about y_i.
- ϵ_i source of variation whose source is unknown (random error term which is independent of inputs and has mean zero).

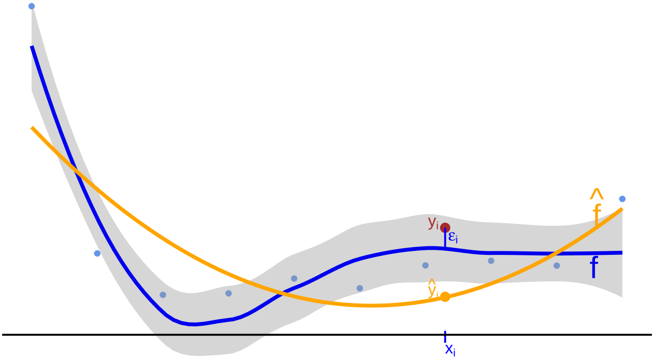


Why estimate f?

- Why not just visualize the data?
  - We want quantitative estimates, not visual summaries.
- Reason 1: Prediction
  - Inputs may be much easier to collect than output.
  - Not typically concerned with the exact form of $\hat{f}$.
- Reason 2: Inference
  - We may care about the form of f. i.e. is there particular inputs relevant at all?

Estimation and prediction

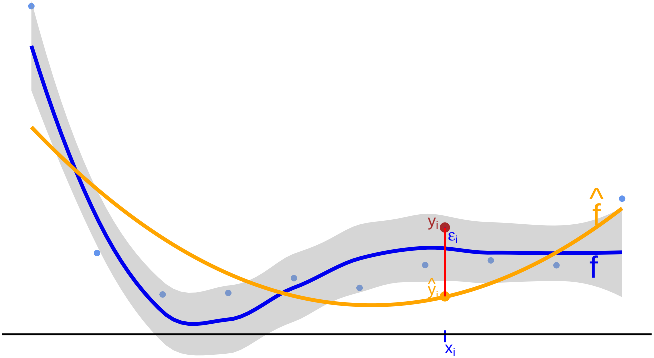
- In practice, we don't know f.
- We estimate it from the data and denote it $\hat{f}$.
- To predict y_i for some input $\boldsymbol{x}_i$, we'd use $\hat{y}_i = \hat{f}(x_i)$.



Source of error

The process introduce two source of errors.

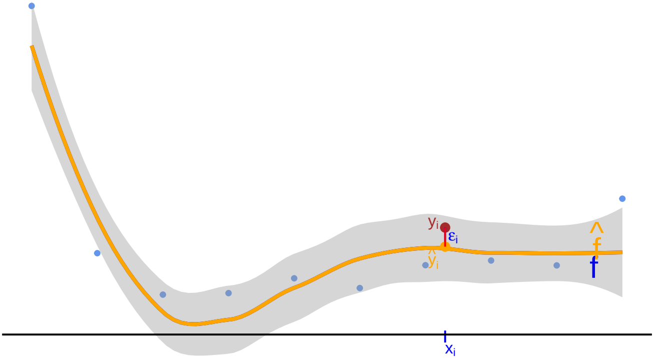
- **Reducible error/Approximation error: $\hat{f}$ isn't close to f.**
  - This error is reducible (using a better algorithm).
- Irreducible error: y_i isn't close to $f(\boldsymbol{x}_i)$.
  - Incur this error even f is known.



Source of error

The process introduce two source of errors.

- Reducible error/Approximation error: $\hat{f}$ isn't close to f.
  - This error is reducible (using a better algorithm).
- **Irreducible error: y_i isn't close to $f(\boldsymbol{x}_i)$.**
  - Incur this error even f is known.



Source of errors

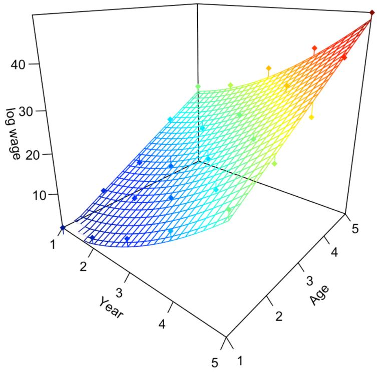
- If we consider the training set, then the decomposition of the average, or expected value, of the squared difference between the actual value of Y and the predicted $\hat{Y} = \hat{f}(X)$ is as follows.

$$\begin{aligned}\mathbb{E}\left(Y - \hat{Y}\right)^2 &= \mathbb{E}\left[f(X) + \epsilon - \hat{f}(X)\right]^2 \\ &= \underbrace{\mathbb{E}\left(\left[f(X) - \hat{f}(X)\right]^2\right)}_{\text{Reducible}} + \underbrace{\mathbb{V}(\epsilon)}_{\text{Irreducible}} \quad . \end{aligned} \tag{1}$$

- If $\hat{f}$ and X are fixed, $E\left[f(X) - \hat{f}(X)\right]^2 = \left[f(X) - \hat{f}(X)\right]^2$.
- We will learn techniques for estimating f by minimizing the reducible error.

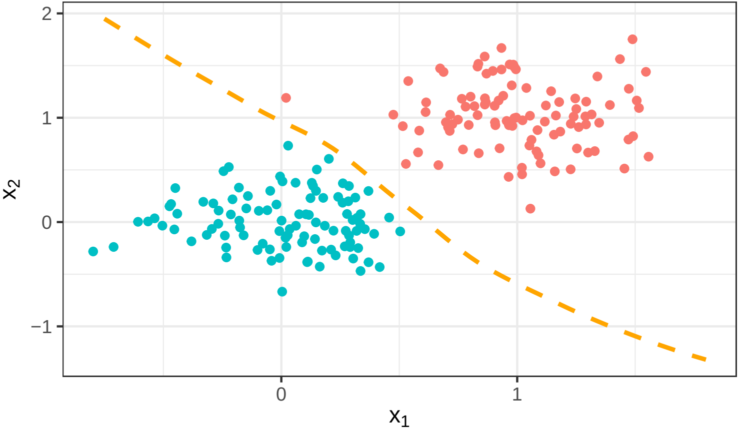
Extending the diagram

- The function fitting applies to high-dimensional $\boldsymbol{x}_i$ and general y_i.
- Here $\boldsymbol{x}_i$ are two dimensional (year, age) and f is a two dimensional surface.



Extending the diagram

- Here the response y_i is binary (green and red).



References

- Function Fitting Intro by Kris Sankaran.
- **ISLR** Sections 2.1, 2.2, 3.2, 3.5.