Polynomial Regression

Polynomial Features with Linear Regression

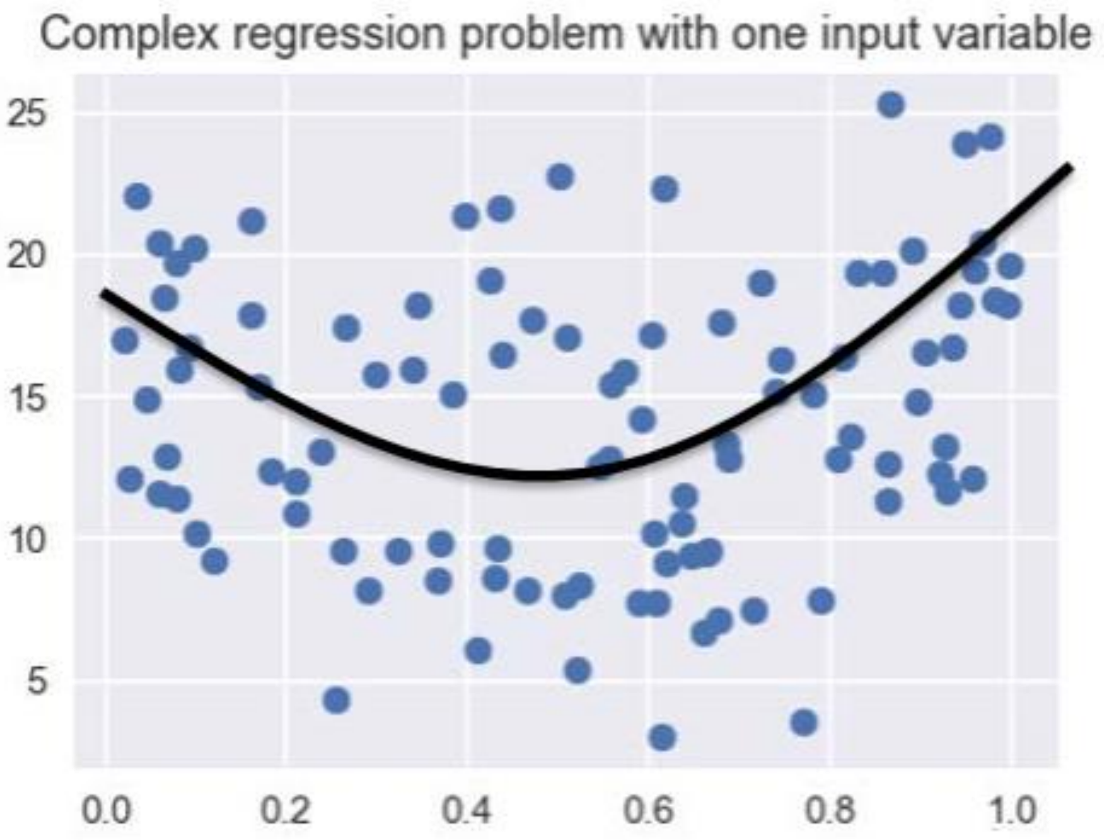
$$\boldsymbol{x} = (x_0, x_1) \quad \Longrightarrow \quad \boldsymbol{x}' = (x_0, x_1, x_0^2, x_0 x_1, x_1^2)$$

$$\hat{y} = \widehat{w}_0 x_0 + \widehat{w}_1 x_1 + \widehat{w}_{00}\, x_0^2 + \widehat{w}_{01} x_0 x_1 + \widehat{w}_{11} x_1^2 + b$$

- **Generate new features consisting of all polynomial combinations of the original two features (x_0, x_1).**
- **The *degree* of the polynomial specifies how many variables participate at a time in each new feature (above example: degree 2)**
- **This is still a weighted linear combination of features, so it's <u>still a linear model</u>, and can use same least-squares estimation method for *w* and *b*.**

Least-Squares Polynomial Regression

Original feature



With polynomial degree 2 (quadratic) feature

Polynomial Features with Linear Regression

- **Why would we want to transform our data this way?**
  - *To capture interactions between the original features by adding them as features to the linear model.*
  - *To make a classification problem easier (we'll see this later).*
- **More generally, we can apply other non-linear transformations to create new features**
  - *(Technically, these are called non-linear basis functions)*
- **Beware of polynomial feature expansion with high as this can lead to complex models that overfit**
  - *Thus, polynomial feature expansion is often combined with a regularized learning method like ridge regression.*