1.3 Polynomial Regression

Definition

Polynomial regression is a form of linear regression in which the relationship between the x and the y is modelled as an k^{th} degree polynomial

How to use

- Polynomial models are useful in situations where the analyst know that **curvilinear effects** are present in the true response function.
- Polynomial models are also useful as possible complex **nonlinear relationship**.

Expanding Simple Linear Regression

Quadratic model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$$

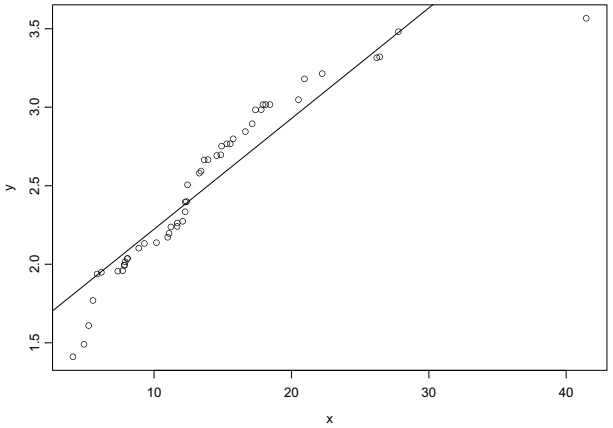
General polynomial model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1^3 + \ldots + \beta_k x_1^k$$

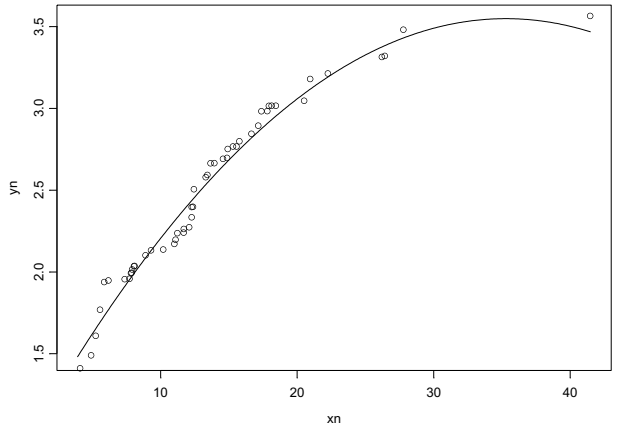
Adding one or more polynomial terms to the model.

Any independent variable, x_i, which appears in the polynomial regression model as x_i^k is called a **k^{th}-degree term**.

Polynomial model shapes



Linear



Quadratic

Adding one more terms to the model significantly improves the model fit.

EXAMPLE

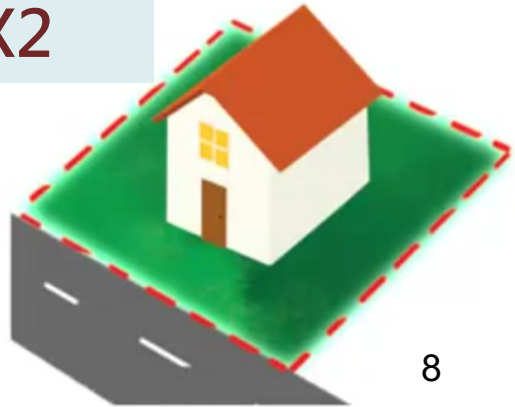
Housing prices prediction

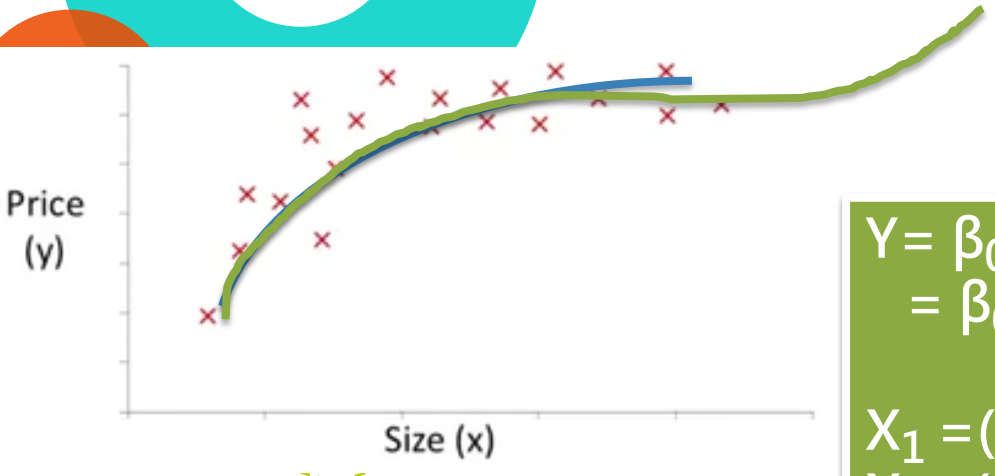
$$Y = \beta_0 + \beta_1 * \text{frontage} + \beta_2 * \text{depth}$$

X1 X2

New variable area X = X1 * X2

$$Y = \beta_0 + \beta_1 * X$$



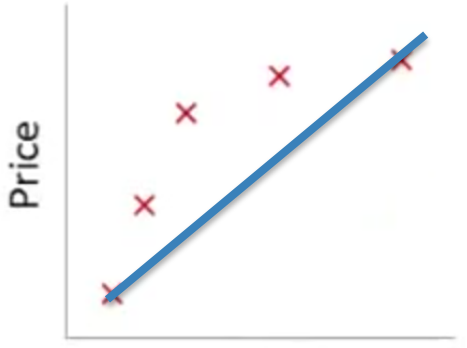


$$Y = \beta_0 + \beta_1 * X + \beta_2 * X^2$$

$$Y = \beta_0 + \beta_1 * X + \beta_2 * X^2 + \beta_3 * X^3$$
$$= \beta_0 + \beta_1(\text{size}) + \beta_2(\text{size})^2 + \beta_3(\text{size})^3$$

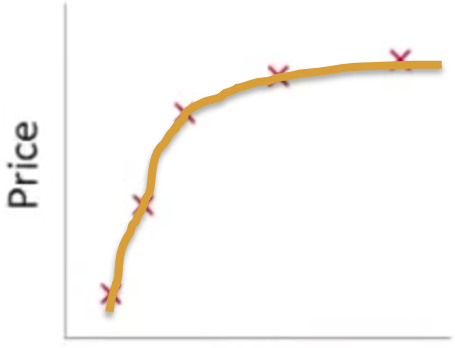
$X_1 = (\text{size})$
$X_2 = (\text{size})^2$
$X_3 = (\text{size})^3$

Overfitting



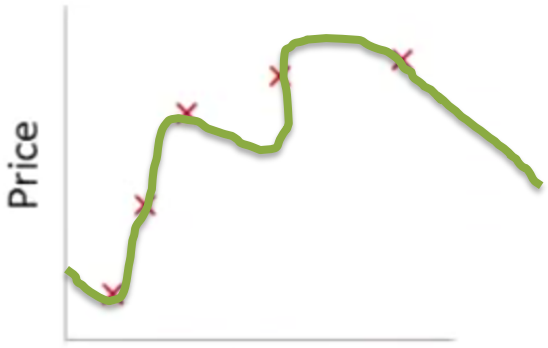
$\theta_0 + \theta_1 x$

Under fitting



$\theta_0 + \theta_1 x + \theta_2 x^2$

Just Right



$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

Over fitting

generalization

A more complex model does not always lead to better performance

Understanding the Bias-Variance Tradeoff

Error due to “bias”

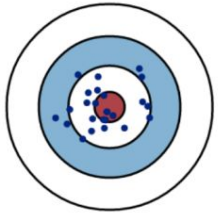
Error due to “variance”

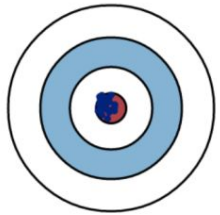
Error = Bias + Variance

Bias & Variance

概念	重點	精準定義	對象	高時的結果
Bias（偏差）	模型「準不準」	單個模型的平均誤差（預測 vs 真實）	一個模型	**容易欠擬合**（模型太簡單）
Variance（變異）	模型「穩不穩」	同一模型架構用不同資料訓練後，預測的變動程度	多個模型 同架構、不同資料	**容易過擬合**（模型太敏感）

射箭比喻





情況	模型狀態	解釋
箭都射到上方	高 Bias	方向錯了， 且錯得很固定， 永遠偏同一邊。 **模型本身方向不正確**
箭散在四處	高 Variance	每批資料結果都不一樣， 每次訓練結果都不同， **模型不穩**
箭集中在靶心附近	低 Bias + 低 Variance	好模型該有的狀態

和模型複雜度的關係

簡單模型	Bias 大 Variance 小	不準 穩	欠擬合
複雜模型	Bias 小 Variance 大	準 不穩	過擬合

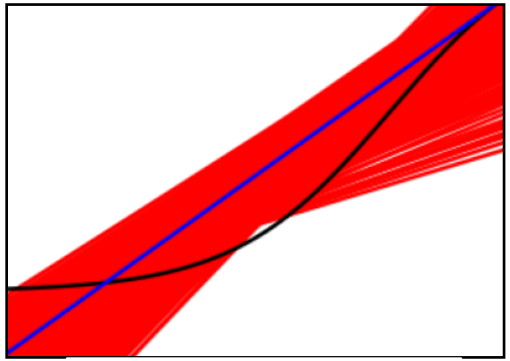
我們要的模型：正確 + 穩定

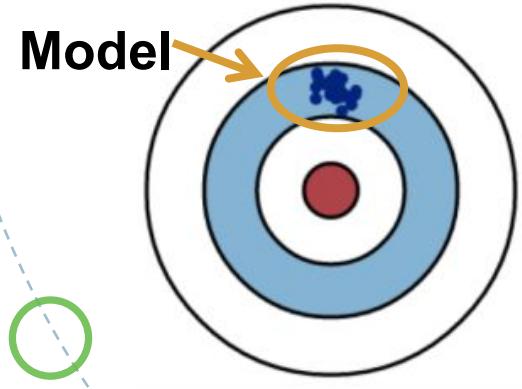
“

bias

bias

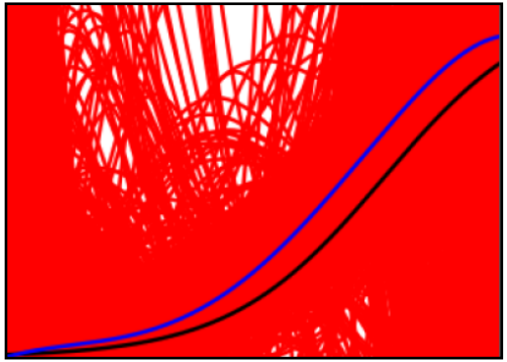
$Y = \beta_0 + \beta_1 * X$

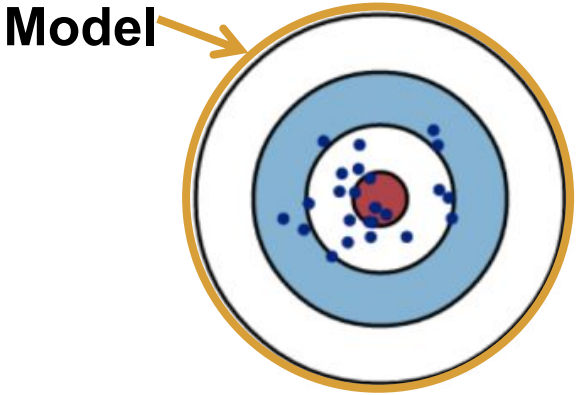




Large Bias

$Y = \beta_0 + \beta_1 * X + \beta_2 * X^2 + \beta_3 * X^3 + \beta_4 * X^4 + \beta_5 * X^5$



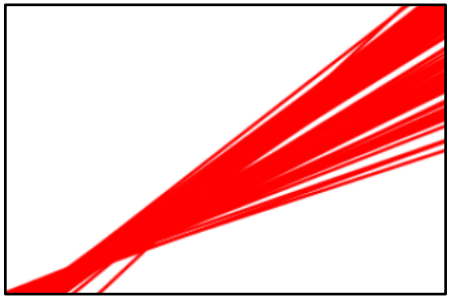


Small Bias

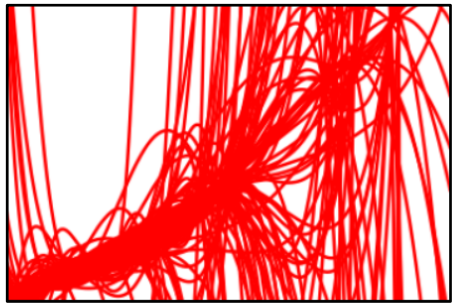
variance

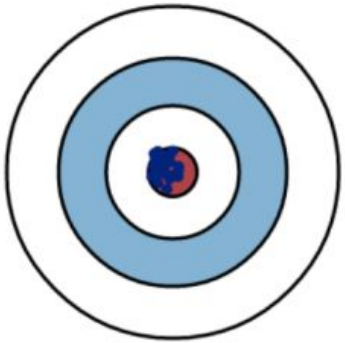
variance

$Y = \beta_0 + \beta_1 * X$

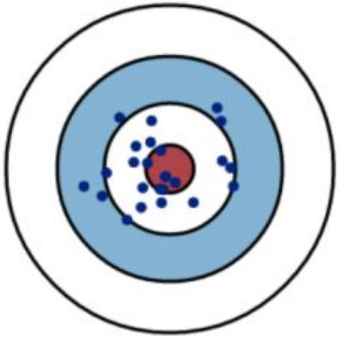


$Y = \beta_0 + \beta_1 * X + \beta_2 * X^2 + \beta_3 * X^3 + \beta_4 * X^4 + \beta_5 * X^5$





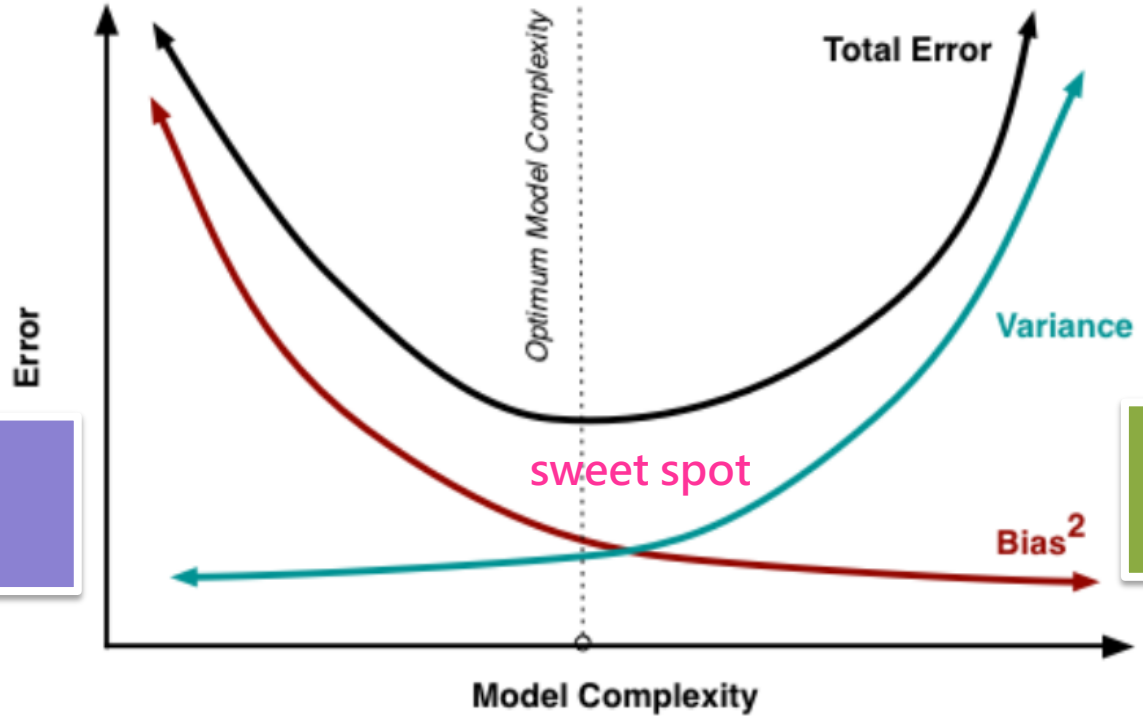
Low Variance



High Variance

Simpler model is less influenced by the sampled data


“
Tradeoff

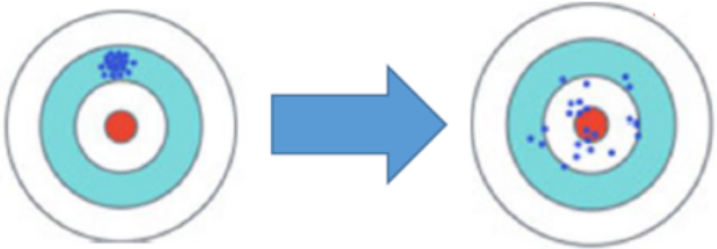


Under fitting

Over fitting

Large Bias

Small Variance



Small Bias

Large Variance

Diagnosis

Underfitting

If your model cannot even fit the data, then you have larges bias

Overfitting

The error due to variance is taken as the variability of a model prediction for a given data point

Redesign model

For large bias

① Add more variables as input
② A more complex model
③ New model architecture

For large variance

① Decrease number of variables
② Obtain more data
③ New model architecture

Model select

① There is usually trade-off between bias and variance
② Select a model that balances two kinds of error to minimize total error

Thanks!