

COMP2261 ARTIFICIAL INTELLIGENCE / MACHINE LEARNING

Polynomial Regression

Dr SHI Lei

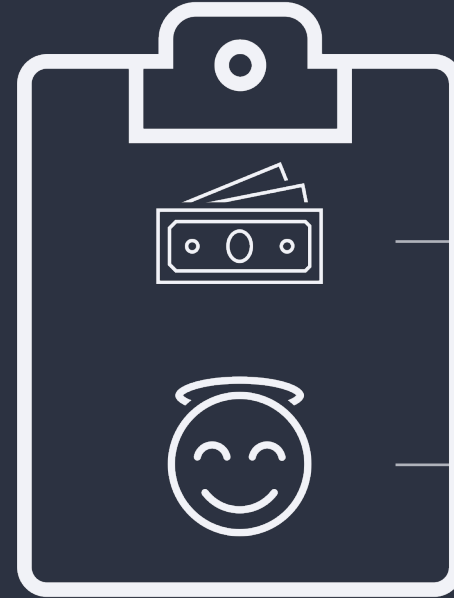
Learning Objectives

- Understand what is a polynomial regression model
- Understand how to transform features
- Understand how to solve parameters

EXAMPLE. annual income to predict happiness



500 people in Durham



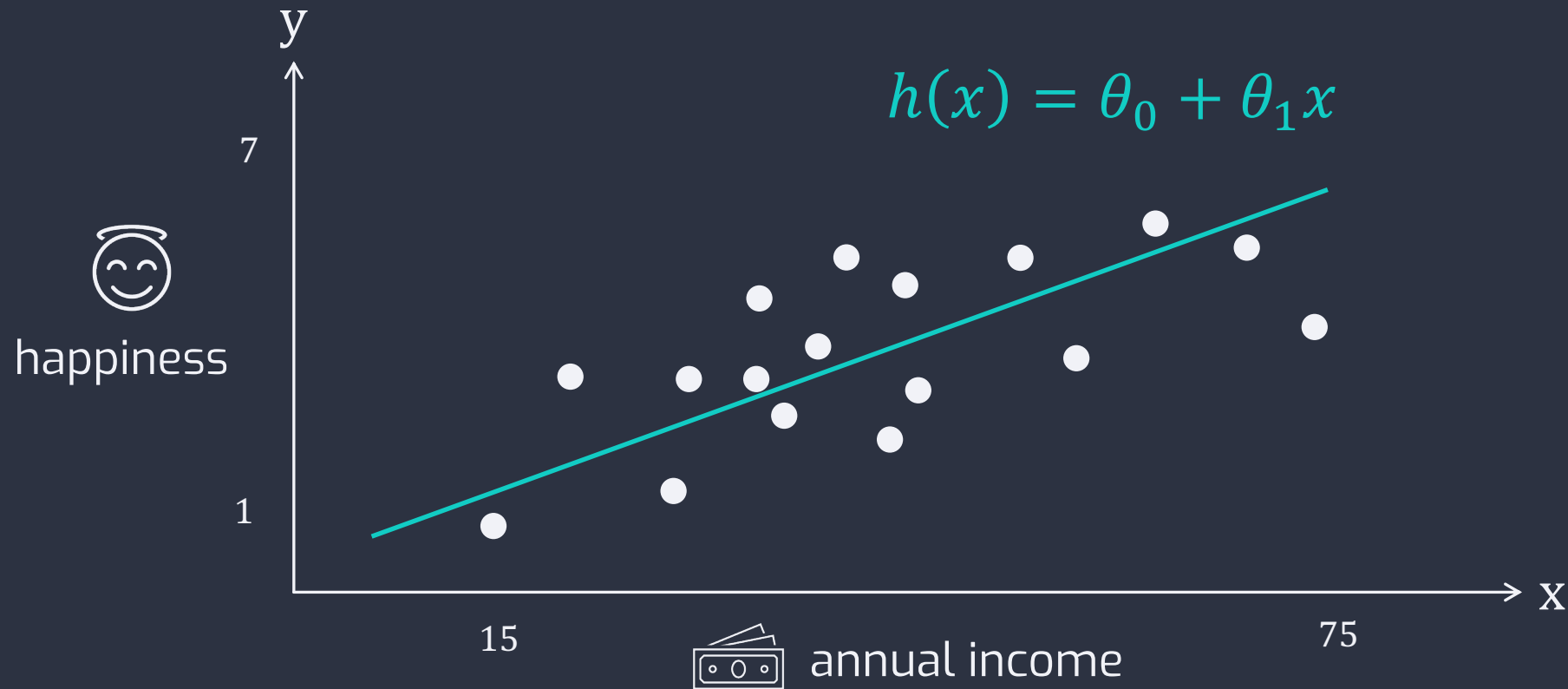
survey

annual income
£15k -- £75k

happiness
0 -- 10

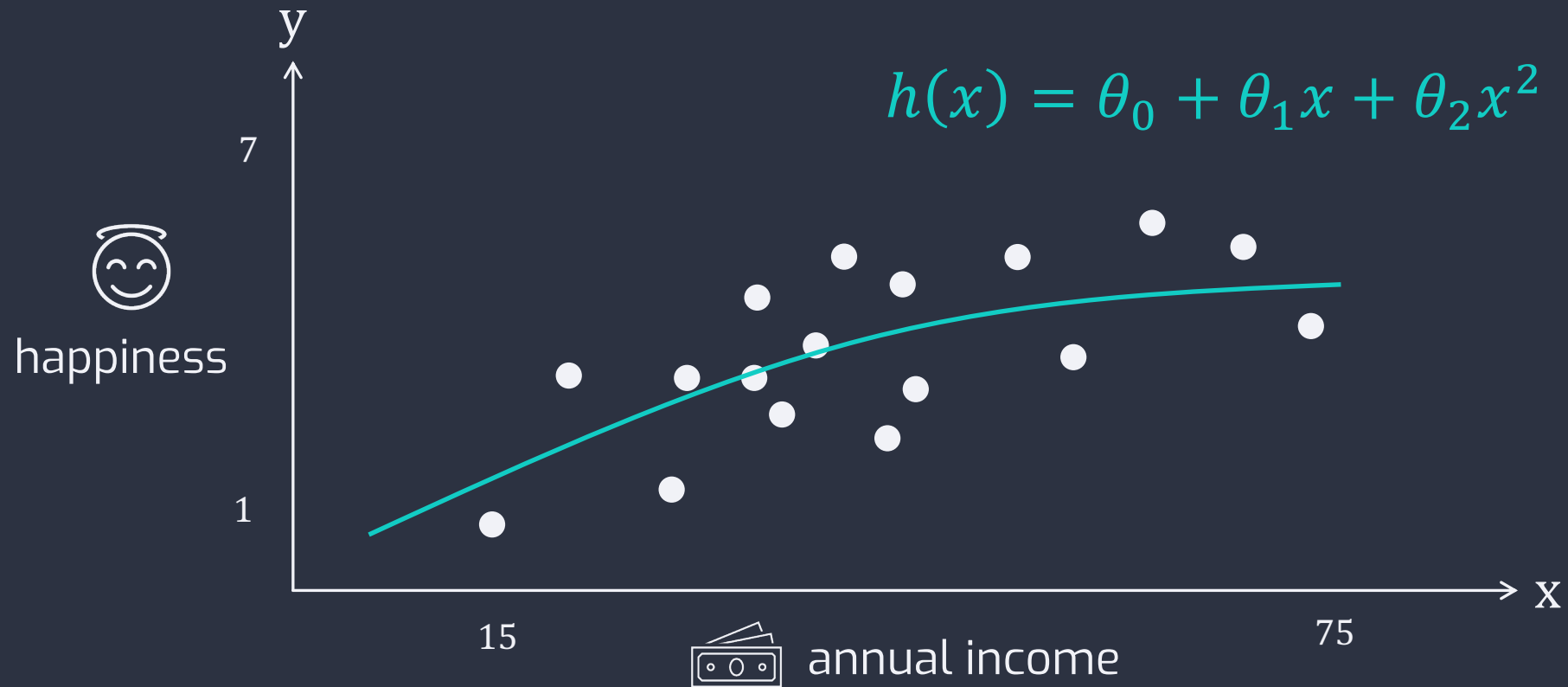
EXAMPLE. annual income to predict happiness

Simple linear regression: a linear regression model with a single explanatory variable (x).



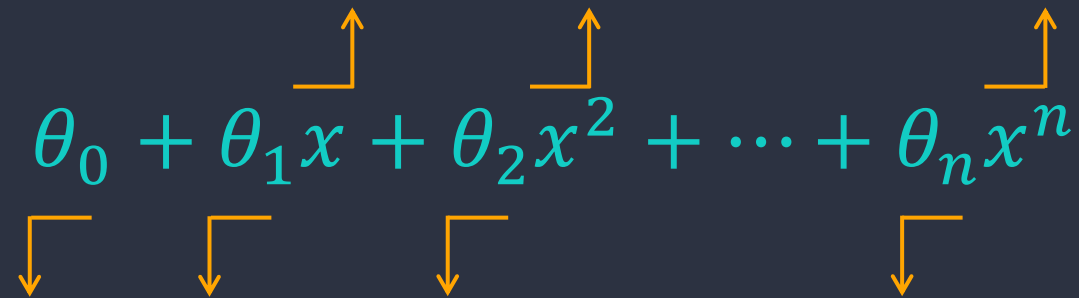
EXAMPLE. annual income to predict happiness

Polynomial linear regression: a linear regression model with higher order.



Polynomial

data x , taken to increasingly high power

$$\theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n$$
The diagram shows the polynomial equation $\theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n$ in teal. Below the equation, four orange arrows point downwards from the terms θ_0 , $\theta_1 x$, $\theta_2 x^2$, and $\theta_n x^n$ to the text 'coefficients, scaling data'. Above the equation, three orange arrows point upwards from the terms $\theta_1 x$, $\theta_2 x^2$, and $\theta_n x^n$ to the text 'data x, taken to increasingly high power'.

coefficients, scaling data

Polynomial

$$\boxed{\theta_0} + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n$$

$\theta_0 x^0$ $\boxed{2} + 4x + \frac{3}{7}x^2 + 5.6x^3$

The diagram illustrates the relationship between the general polynomial form and a specific example. In the general form, the coefficient θ_0 is enclosed in an orange box. An orange arrow points from this box down to the term $\theta_0 x^0$. Another orange arrow points from the constant term 2 in the specific polynomial $2 + 4x + \frac{3}{7}x^2 + 5.6x^3$ up to the θ_0 box, indicating that θ_0 is equal to 2.

Polynomial

$$\theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n$$

$$2 + 4x + \frac{3}{7}x^2 + 5.6x^3$$

$$8 + 9.1x^3$$

$$8 + 0x + 0x^2 + 9.1x^3$$

Polynomial

Order: the highest coefficient

$$\theta_0 + \theta_1 x + \theta_2 x^2 + \cdots + \theta_n x^n$$

$$2 + 4x + \frac{3}{7}x^2 + 5.6x^3$$

third order polynomial

$$8 + 9.1x^3$$

Polynomial Regression

$$h(x) = \theta_0 + \theta_1 x \quad \text{first order polynomial}$$

$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 \quad \text{second order polynomial}$$

$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_n x^n \quad \text{n-th order polynomial}$$

Fit a Polynomial Regression with Nonlinearities

$$h(x) = \theta_0 + \theta_1 x + \theta_2 \boxed{x^2} + \cdots + \theta_n \boxed{x^n}$$

nonlinear

Fit a Polynomial Regression with Nonlinearities

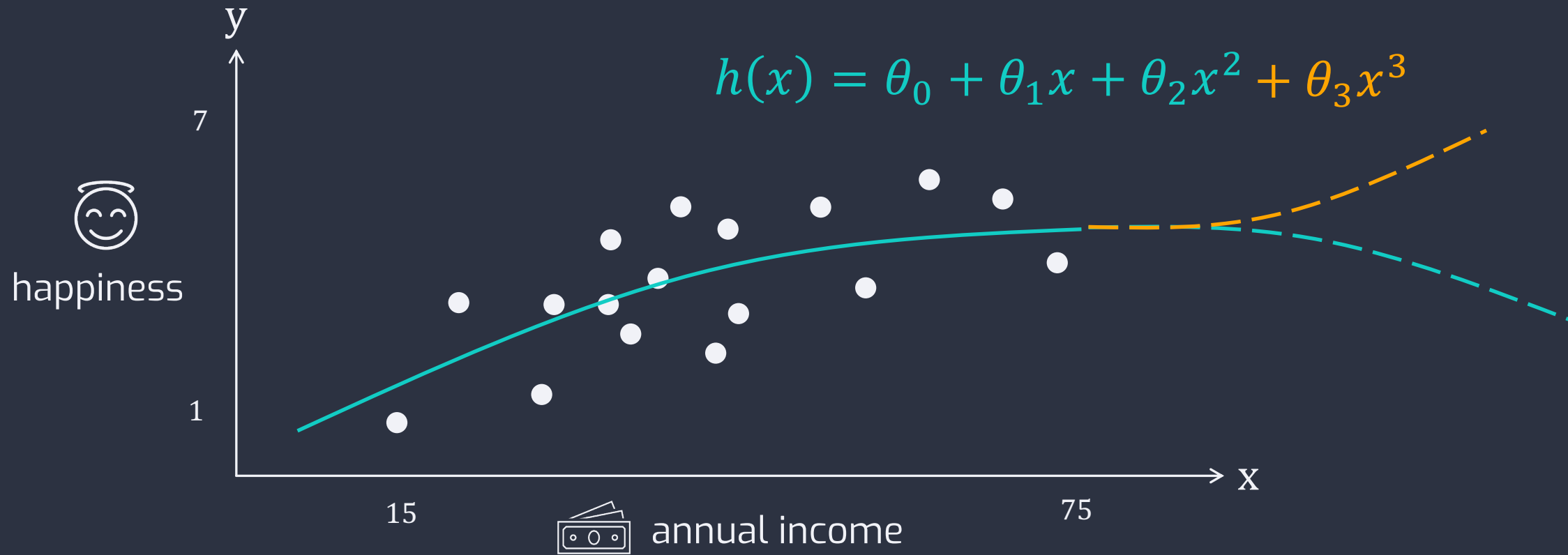
$$h(x) = \boxed{\theta_0} + \boxed{\theta_1}x + \boxed{\theta_2}x^2 + \cdots + \boxed{\theta_n}x^n$$

The coefficients are all linear.



This is a standard linear model!

EXAMPLE. annual income to predict happiness



EXAMPLE. annual income to predict happiness

Transforming polynomial linear regression to multivariate linear regression.

$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$



$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

Gradient descent / normal equation to get $\theta_0, \theta_1, \theta_2, \theta_3$

EXAMPLE. annual income to predict happiness

Transforming polynomial linear regression to multivariate linear regression.

$$h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$



$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

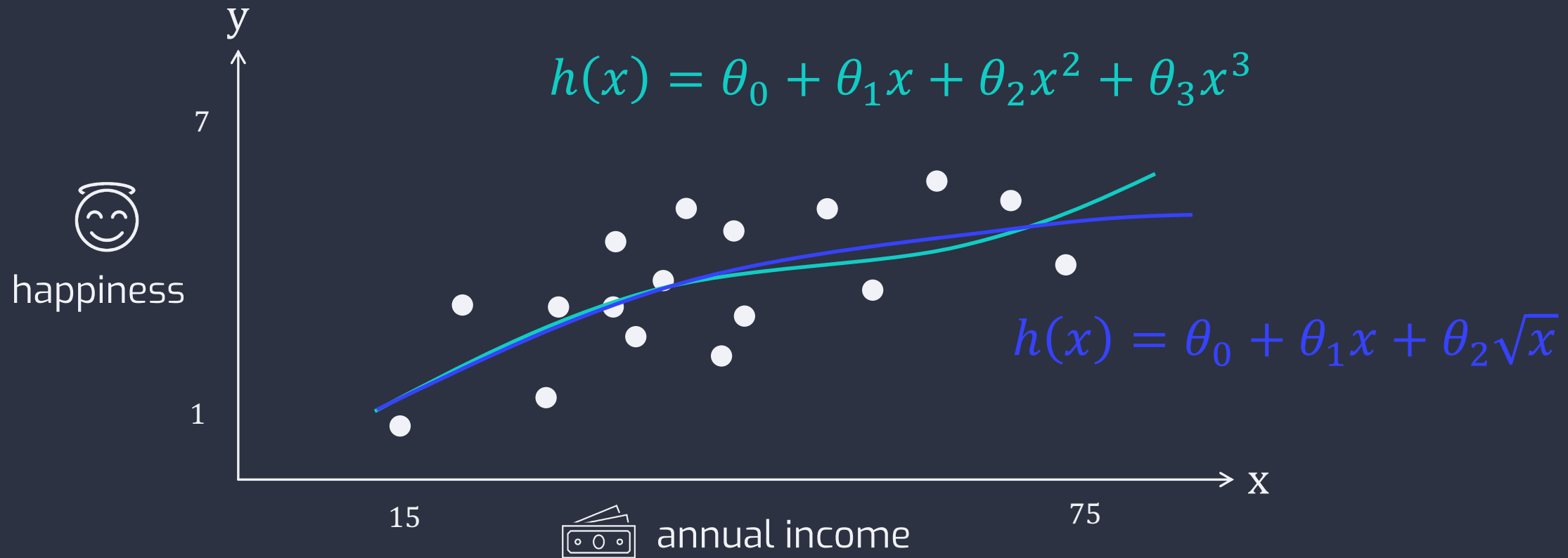
$$x_1 = x \quad 10 \dots 1000$$

$$x_2 = x^2 \quad 100 \dots 1,000,000$$

$$x_3 = x^3 \quad 1000 \dots 1,000,000,000$$

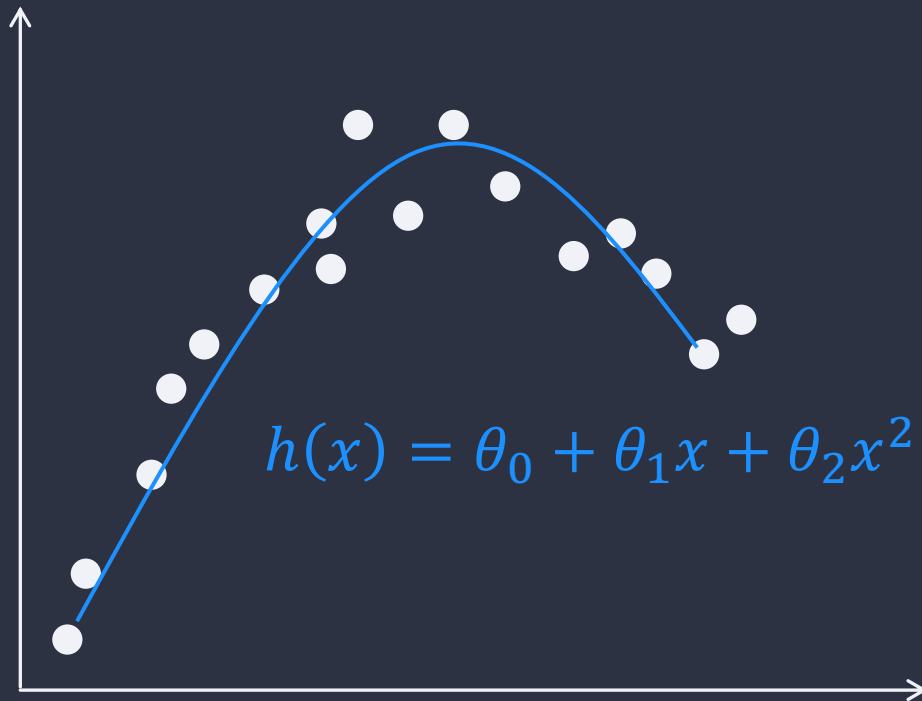
Feature scaling is important!

EXAMPLE. annual income to predict happiness

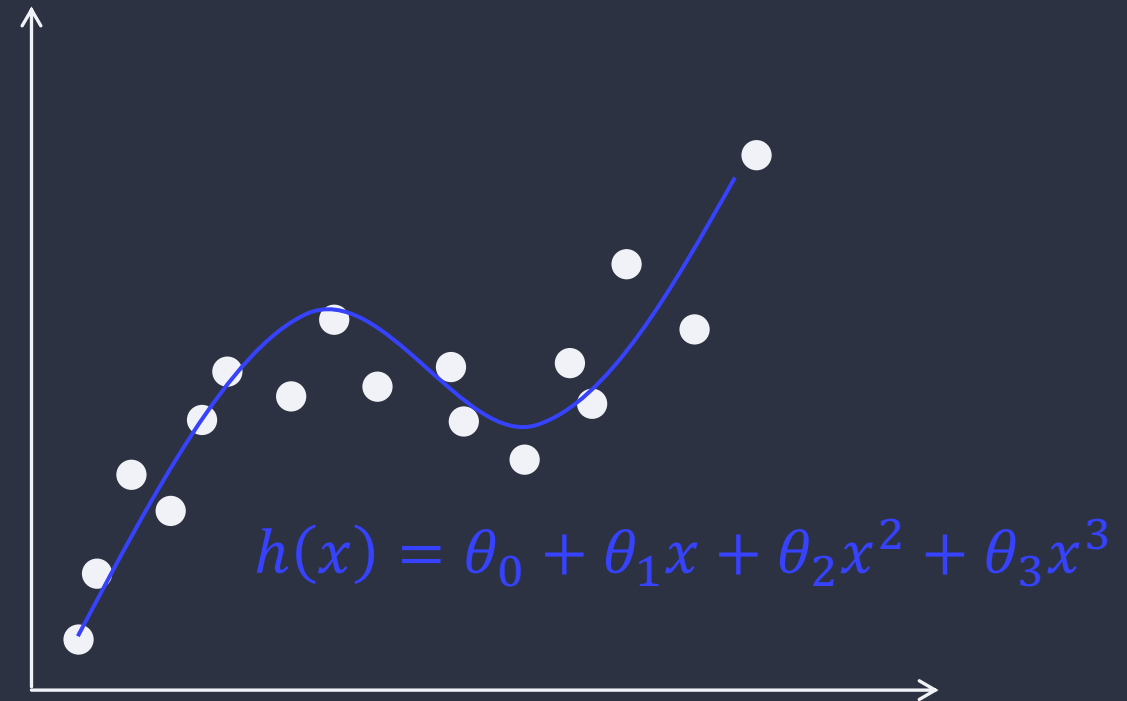


When to use Polynomial

Second-order polynomial



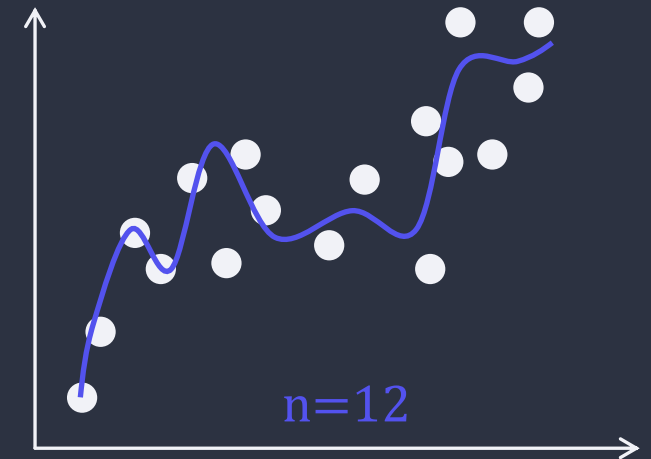
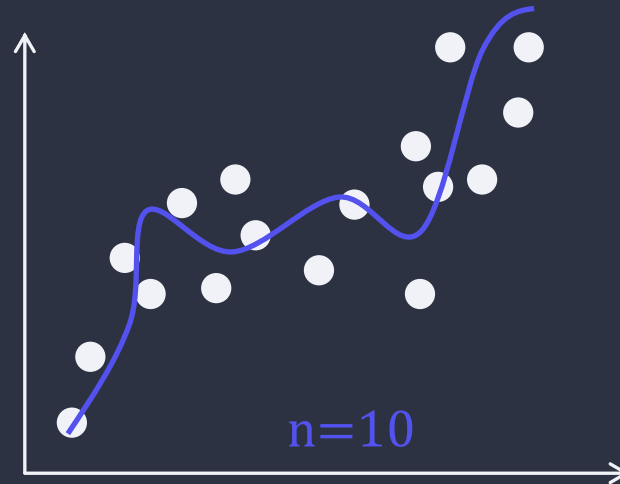
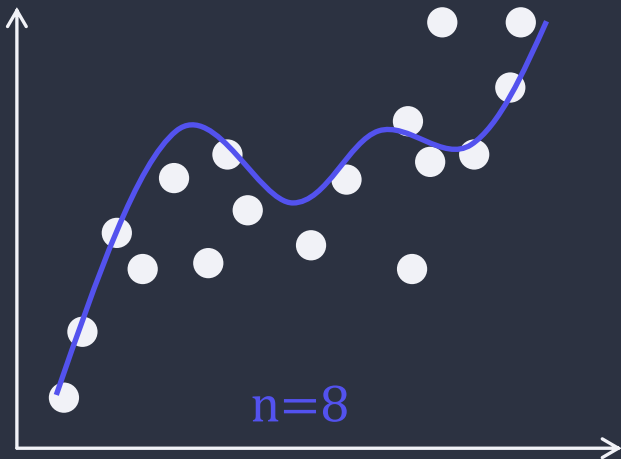
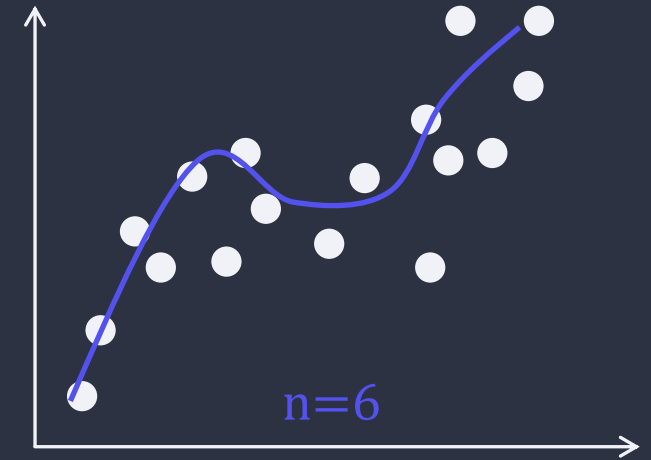
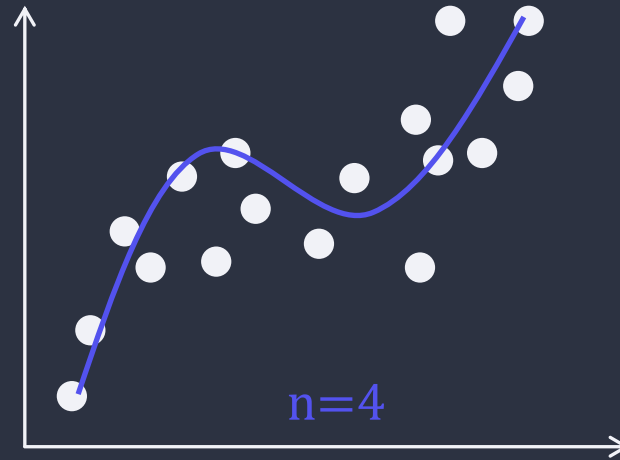
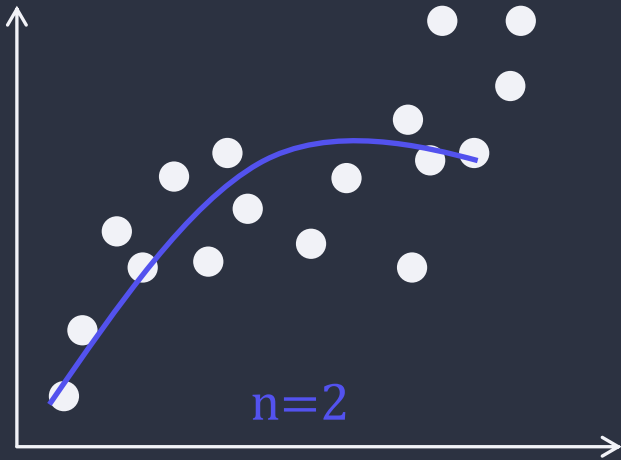
Third-order polynomial



How do we decide the order of a polynomial?

$$h(x) = \theta_0 + \theta_1x + \theta_2x^2 + \dots + \theta_nx^n$$

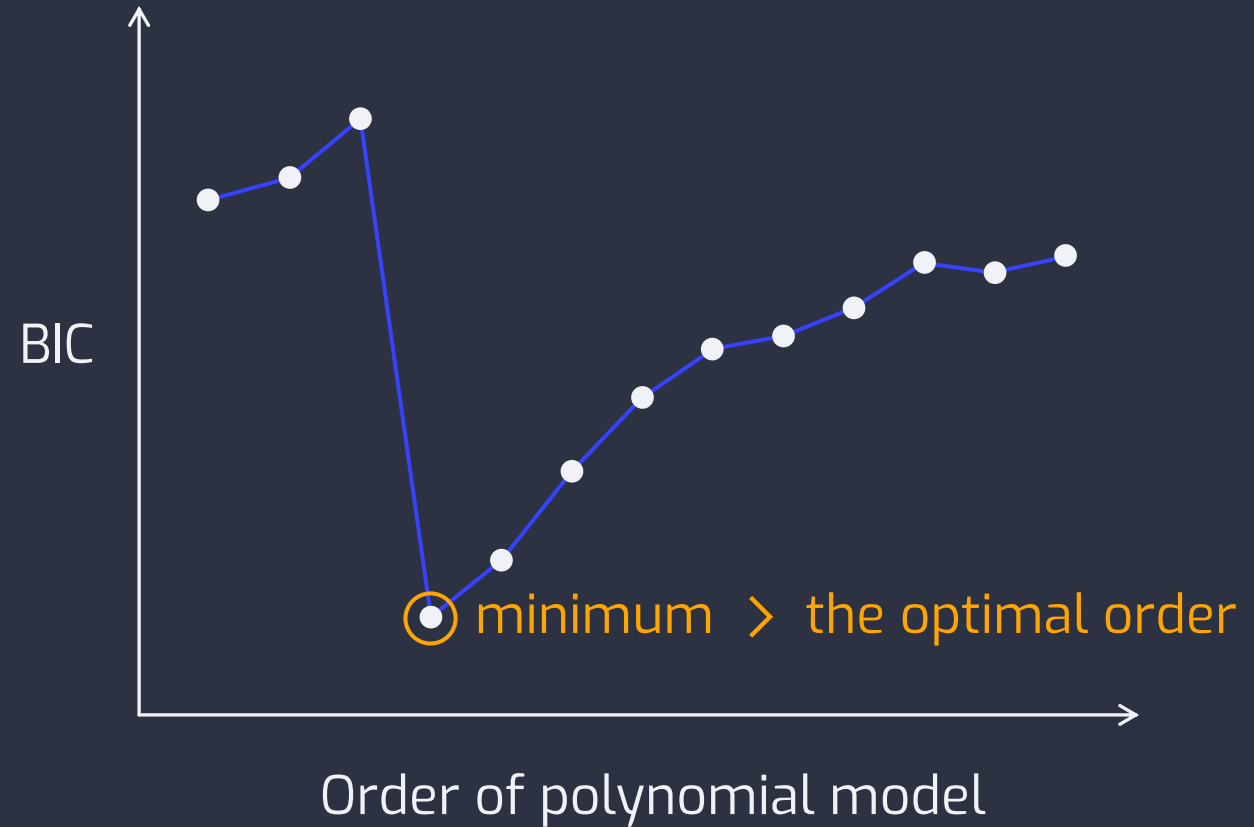
Order of Polynomial



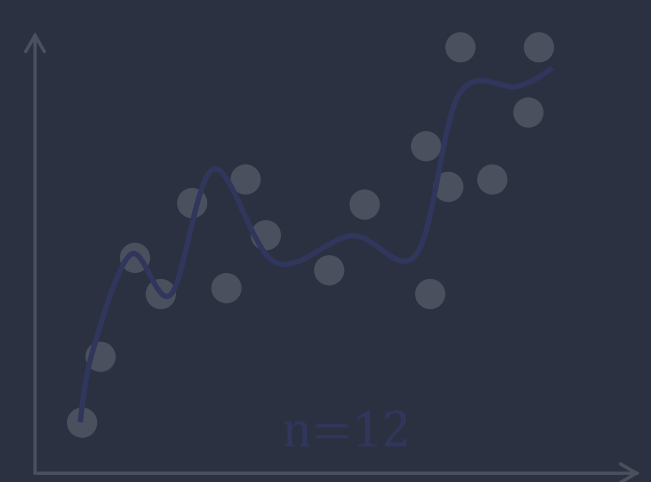
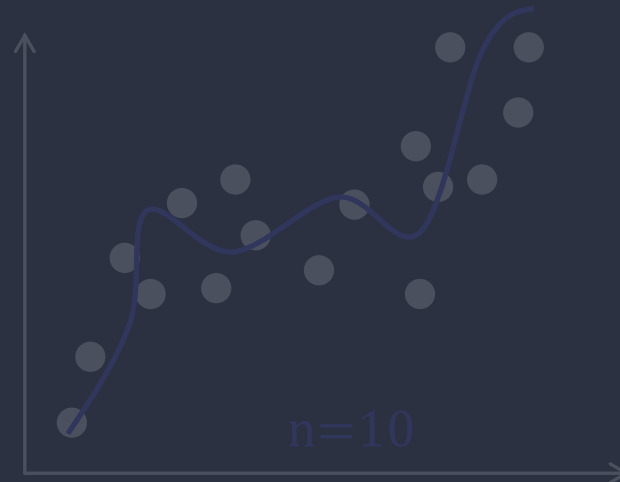
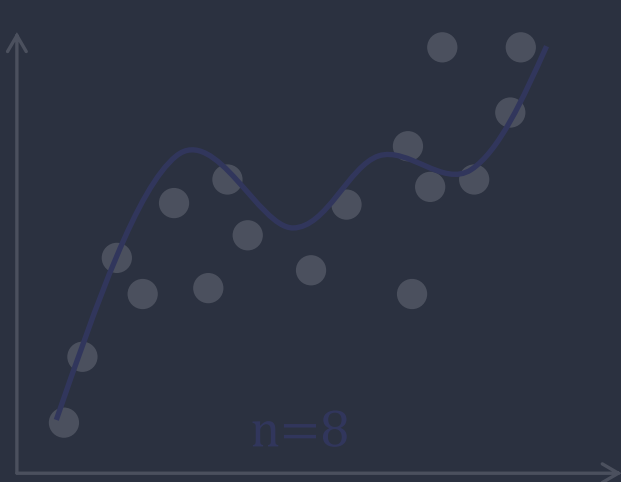
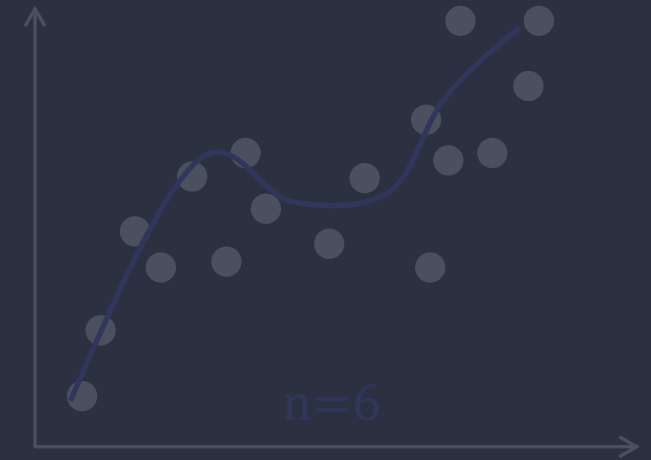
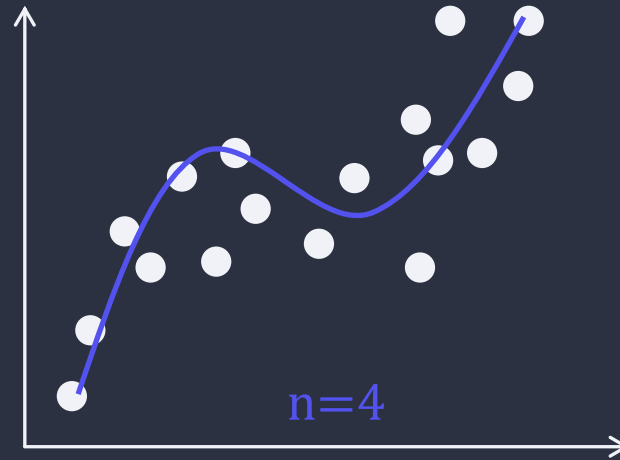
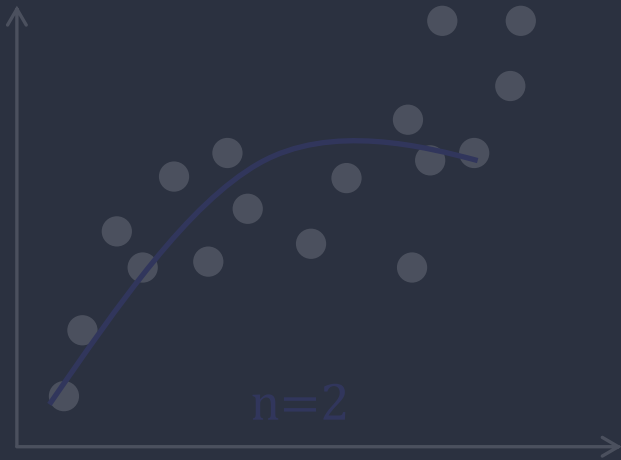
Deciding on Order of Polynomial

Bayesian Information Criterion

$$BIC_n = m \cdot \ln(SS_r) + n \cdot \ln(m)$$



Deciding on Order of Polynomial



✓ Takeaway Points

- Higher order polynomial linear model may better fit data.
- Multivariate linear regression model method to solve $\theta_0, \theta_1, \theta_2, \dots$
- Try different features $x^2, x^3, \sqrt{x}, \dots$
- BIC to decide optimal polynomial order.