KB-74
Applied Data Science

Math behind Machine Learning 9

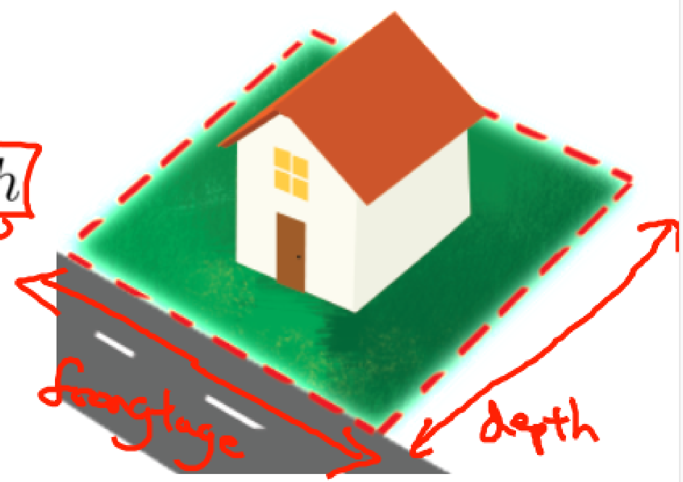
Polynomial regression, normal equation

THE HAGUE UNIVERSITY OF APPLIED SCIENCES
FACULTY IT & DESIGN

Polynomial Regression

Housing prices prediction

$$h_\theta(x) = \theta_0 + \theta_1 \times \underbrace{\boxed{frontage}}_{x_1} + \theta_2 \times \underbrace{\boxed{depth}}_{x_2}$$



Area

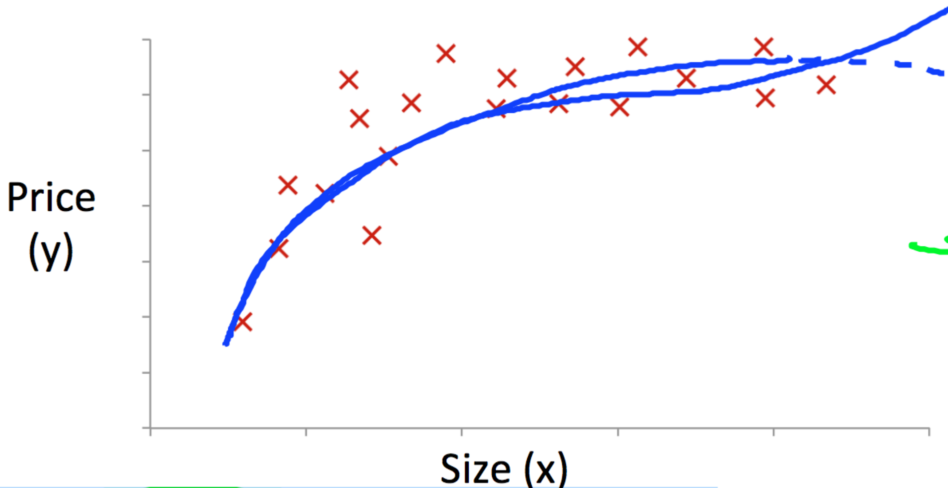
$x = frontage * depth$

$h_\theta(x) = \theta_0 + \theta_1 x$

← land area

Polynomial Regression

Polynomial regression



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

$$\boxed{\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3}$$

$$\boxed{h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3}$$
$$= \theta_0 + \theta_1(size) + \theta_2(size)^2 + \theta_3(size)^3$$

$$x_1 = (size)$$
$$x_2 = (size)^2$$
$$x_3 = (size)^3$$

Size: 1 - 1000

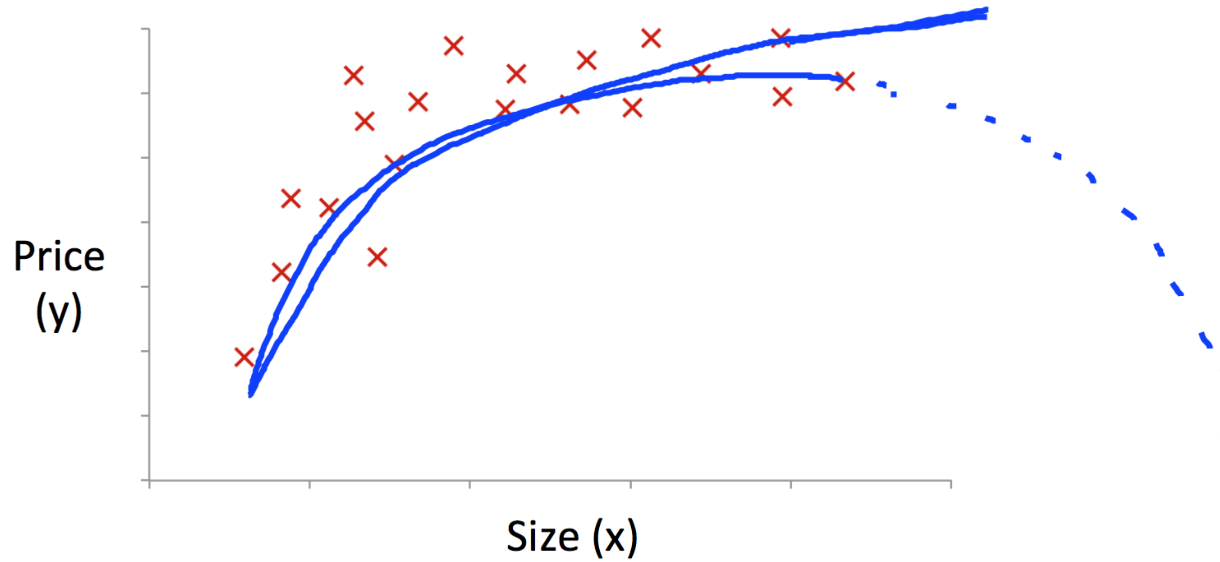
$size^2$: 1 - 1000,000

$size^3$: 1 - 10^9

Andrew N

Polynomial Regression

Choice of features



$$h_\theta(x) = \theta_0 + \theta_1(size) + \theta_2(size)^2$$

$$\boxed{h_\theta(x) = \theta_0 + \theta_1(size) + \theta_2\sqrt{(size)}}$$

Normal Equation

$$\boxed{\theta = \boxed{(X^T X)^{-1} X^T y}} \leftarrow$$

$(X^T X)^{-1}$ is inverse of matrix $X^T X$.

Set $A = X^T X$

$$\boxed{(X^T X)^{-1}} = A^{-1}$$

Octave: `pinv(X'*X)*X'*y`

$\text{pinv}(X^T * X) * X^T * y$

$\theta = (X^T X)^{-1} X^T y$

$\min_{\theta} J(\theta)$

X' X^T

~~Feature Scaling~~

$0 \leq x_1 \leq 1$

$0 \leq x_2 \leq 1000$

$0 \leq x_3 \leq 10^{-5}$ ✓

Normal Equation

m **training examples,** n **features.**

Gradient Descent	Normal Equation
• Need to choose α.	• No need to choose α.
• Needs many iterations.	• Don't need to iterate.
• Works well even when n is large.	• Need to compute $\boxed{(\boxed{X^TX})^{-1}}$
	• Slow if n is very large.

$n \times n$

$O(n^3)$

$n = 10^6$

$n = 100$

$n = 1000$

$n = 10000$

Normal Equation

Normal equation

$$\theta = (X^T X)^{-1} X^T y$$

$X^T X$

- What if $\boxed{X^T X}$ is non-invertible? (singular/ degenerate)
- Octave: `pinv(X'*X)*X'*y`

$\boxed{\text{pinv}}$

inv

Normal Equation

What if $\boxed{X^T X}$ is non-invertible?

- Redundant features (linearly dependent).
  E.g. x_1 = size in feet^2
  ~~x_2 = size in m^2~~
  $x_1 = (3.28)^2 x_2$

1m = 3.28 feet

$m = 10$
$n = 100$
$\theta \in \mathbb{R}^{101}$

- Too many features (e.g. $m \leq n$).
  - Delete some features, or use regularization.

later