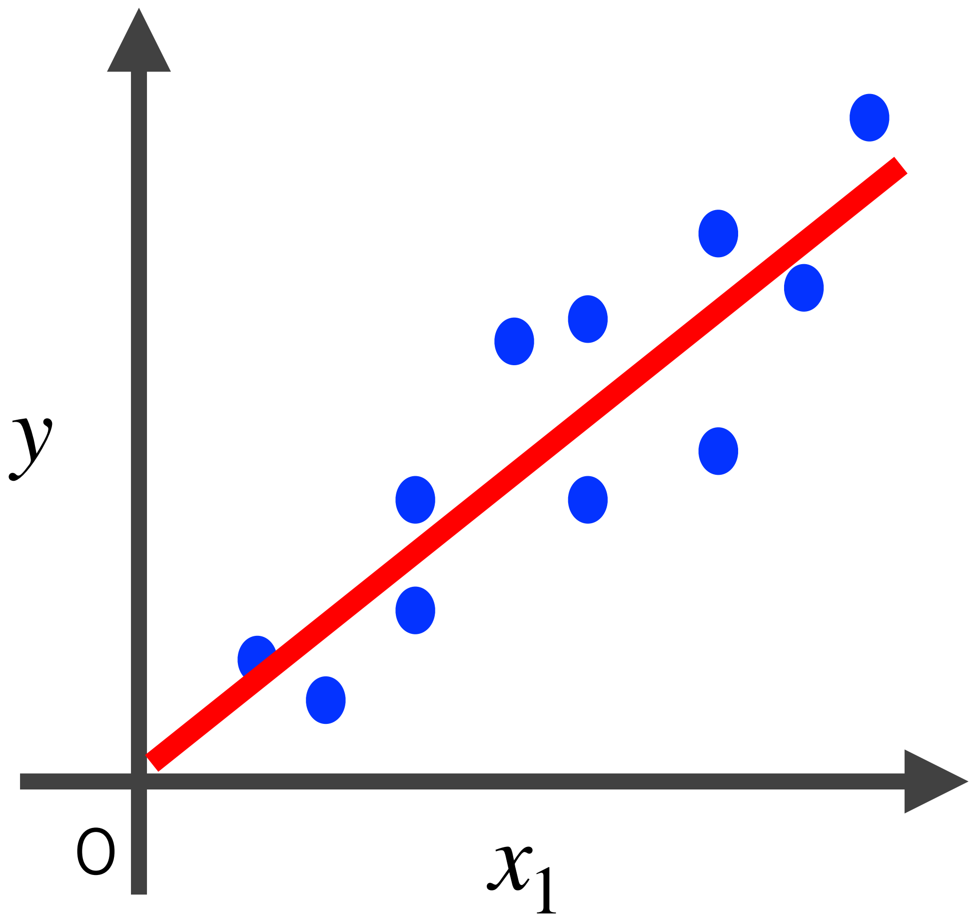
Brain Code Camp

Polynomial Regression, Overfitting and Regularization

Itthi Chatnuntawech

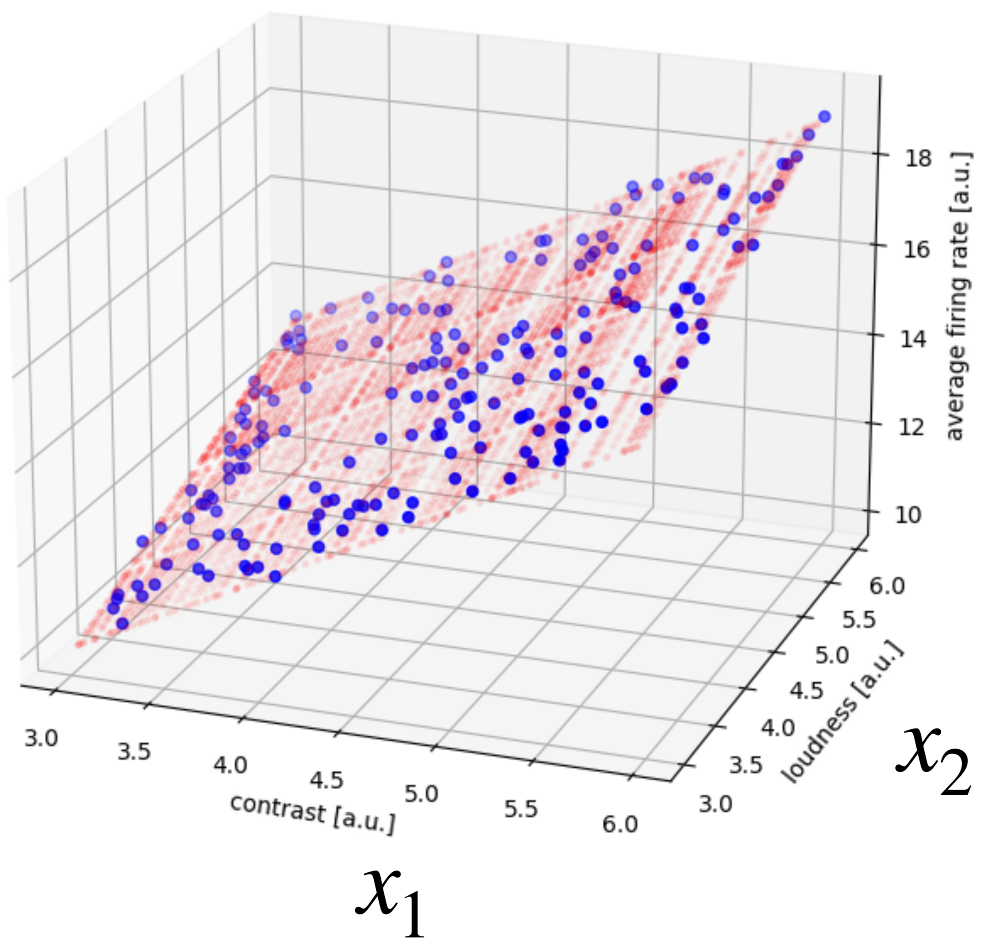
Linear Models

$$y = w_0 + w_1 x_1$$



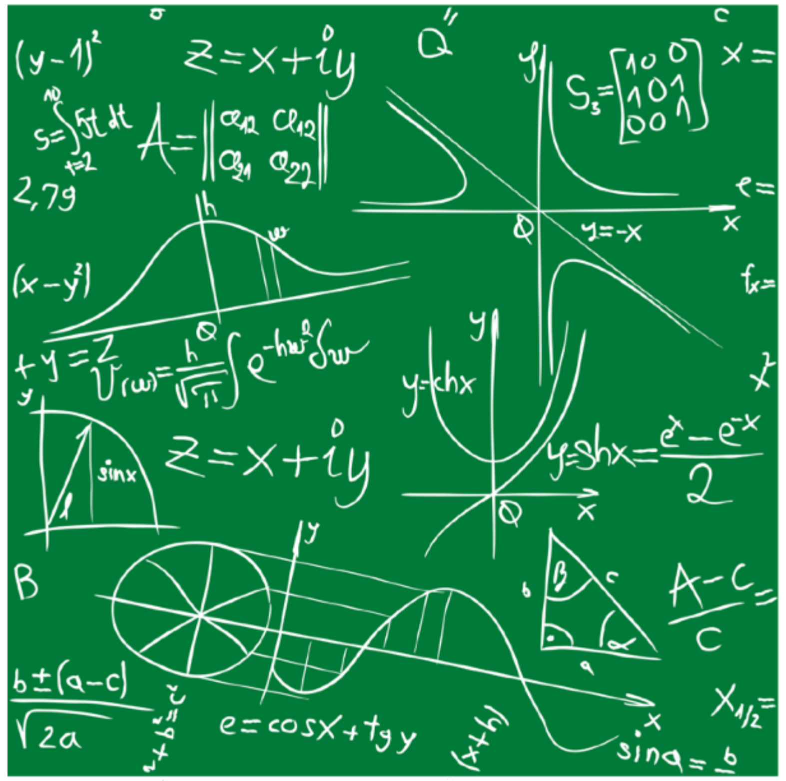
Line
1 feature

$$y = w_0 + w_1 x_1 + w_2 x_2$$



Plane
2 features

$$y = w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_p x_p$$

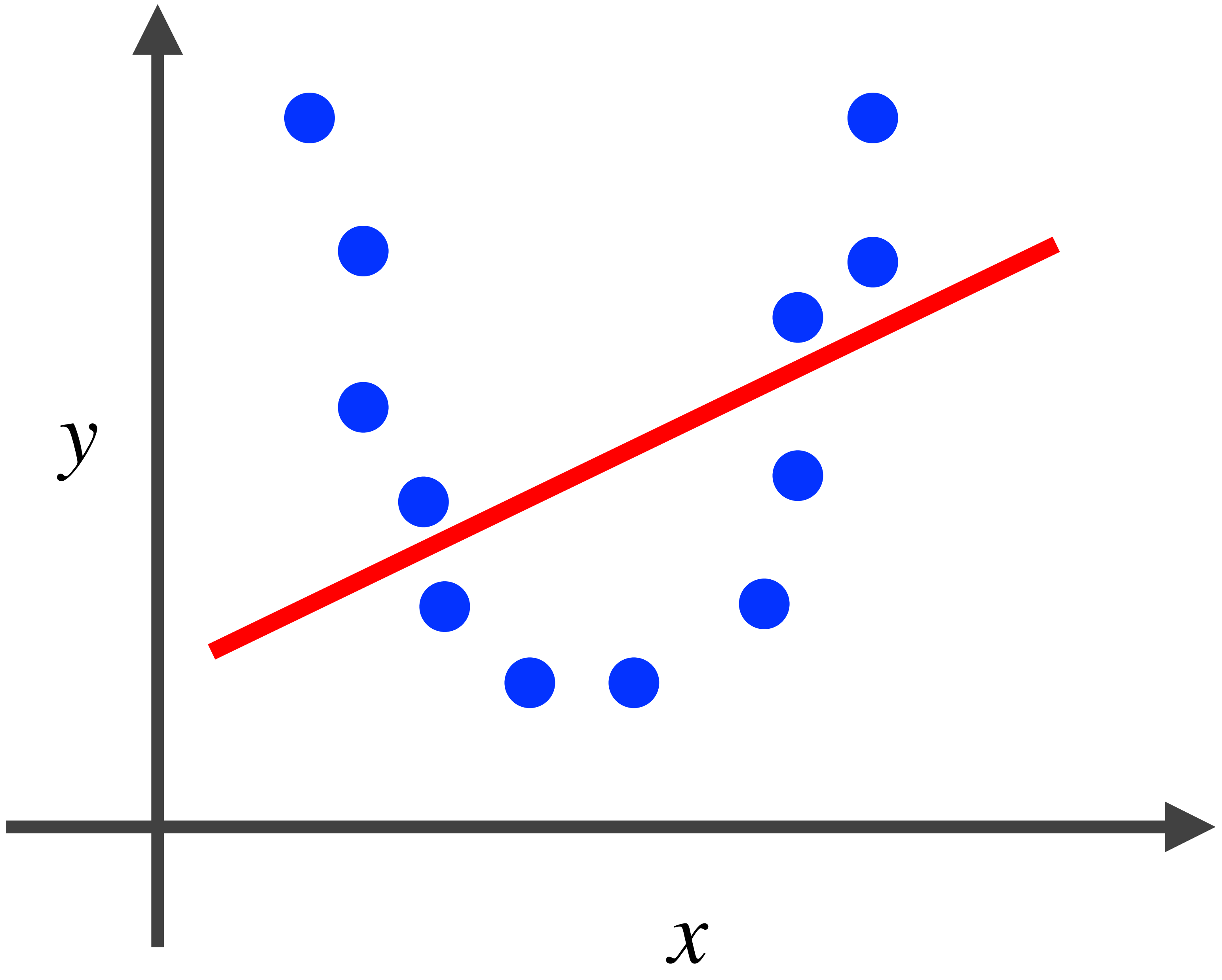


Hyperplane
p features

Nonlinear Relationship

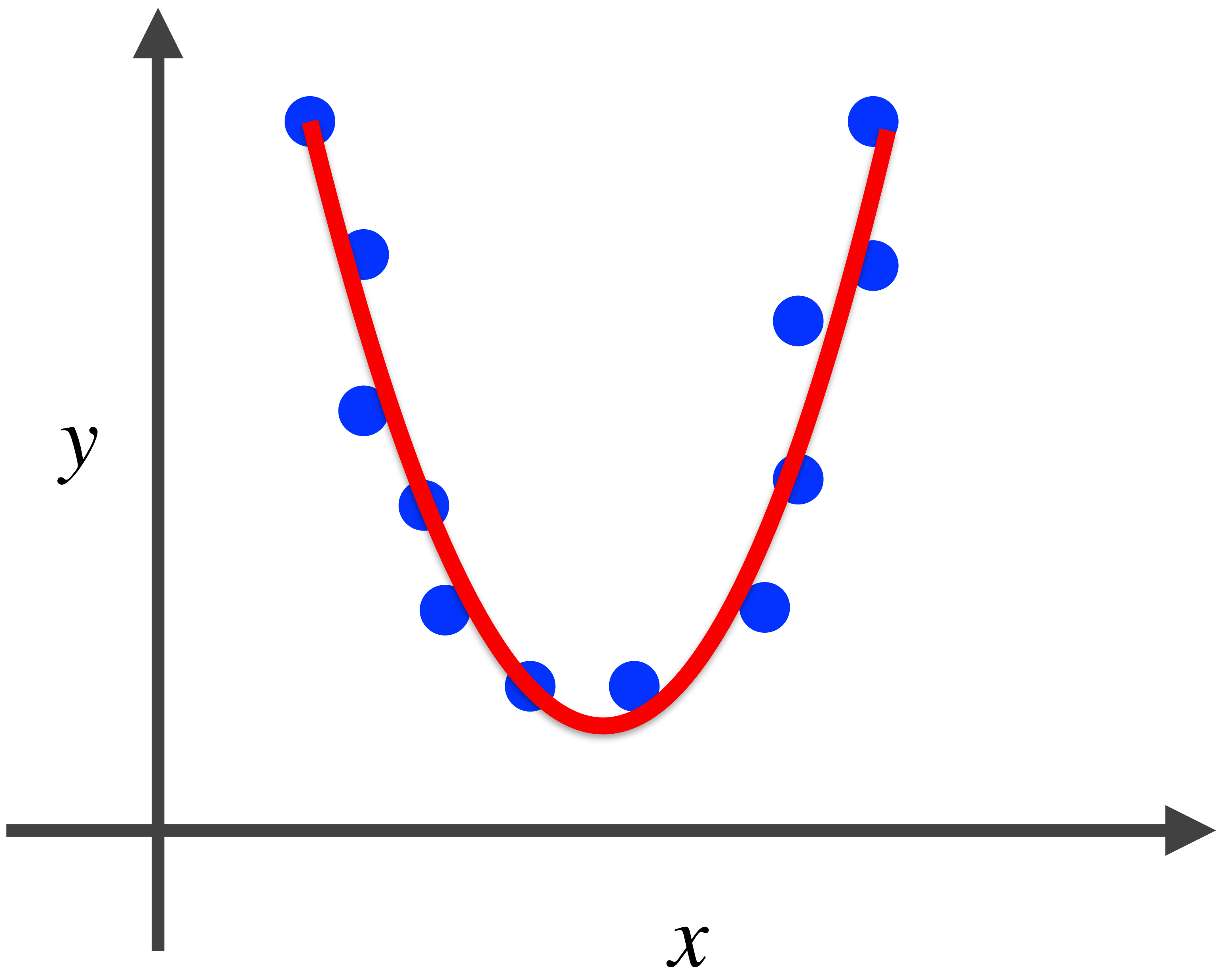
$$y = mx + c$$

$$y = w_0 + w_1 x_1$$



$$y = ax^2 + bx + c$$

$$y = w_0 + w_1 x + w_2 x^2$$



Brain Code Camp

Simple Polynomial

A linear model

$$y = w_0 + w_1x_1 + w_2x_2 + \ldots + w_px_p$$

It is linear in both parameters $(w_1, w_2, \ldots, w_p)$ and variables $(x_1, x_2, \ldots, x_p)$

There is no nonlinear terms such as $w_i^p, w_iw_j, w_i^kw_j^l, e^{w_i}$ and $log(w_i)$ and $x_i^p, x_ix_j, x_i^kx_j^l, e^{x_i}$ and $log(x_i)$.

A simple polynomial

$$y = w_0 + w_1x + w_2x^2 + \ldots + w_px^p$$

It is linear in parameters $(w_1, w_2, \ldots, w_p)$, but not linear in variables because it contains $x^2, x^3, \ldots, x^p$

Brain Code Camp

Simple Polynomial

A linear model

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x_1 + \hat{w}_2 x_2 + \ldots + \hat{w}_p x_p$$

A simple polynomial

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x + \hat{w}_2 x^2 + \ldots + \hat{w}_p x^p$$

Brain Code Camp

Simple Polynomial

A linear model

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x_1 + \hat{w}_2 x_2 + \ldots + \hat{w}_p x_p$$

feature 1 feature 2 feature p

A simple polynomial

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x + \hat{w}_2 x^2 + \ldots + \hat{w}_p x^p$$

Let $z_i = x^i$

$$\hat{y} = \hat{w}_0 + \hat{w}_1 z_1 + \hat{w}_2 z_2 + \ldots + \hat{w}_p z_p$$

Simple Polynomial

A linear model

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x_1 + \hat{w}_2 x_2 + \ldots + \hat{w}_p x_p$$

feature 1 feature 2 feature p

A simple polynomial

$$\hat{y} = \hat{w}_0 + \hat{w}_1 z_1 + \hat{w}_2 z_2 + \ldots + \hat{w}_p z_p$$

sklearn.preprocessing.PolynomialFeatures

class sklearn.preprocessing.**PolynomialFeatures**(*degree=2, *, interaction_only=False, include_bias=True, order='C'*)

[source]

sklearn.linear_model.LinearRegression

class sklearn.linear_model.**LinearRegression**(**, fit_intercept=True, copy_X=True, n_jobs=None, positive=False*)

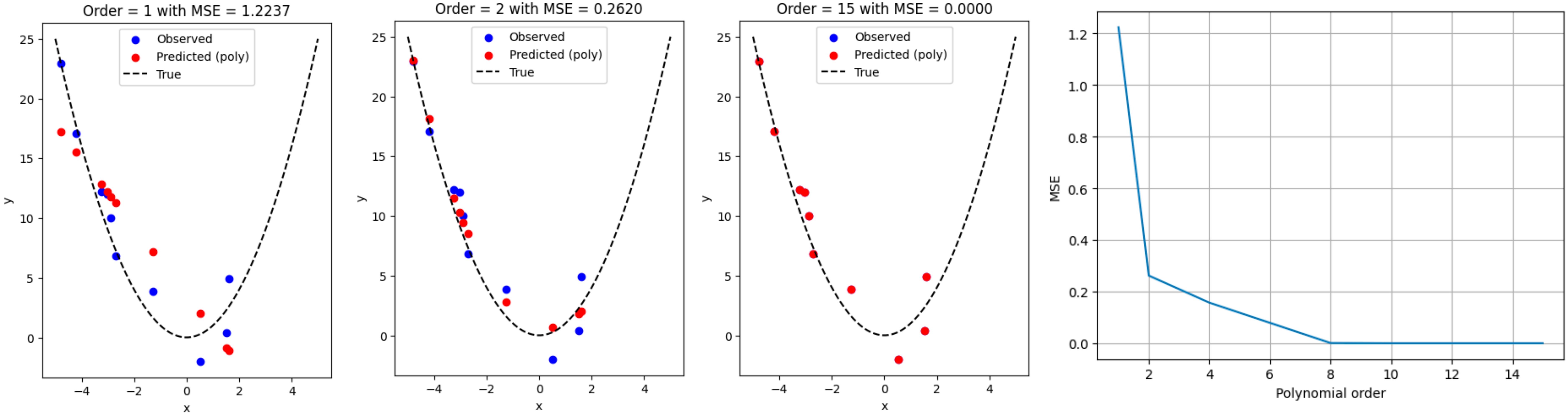
[source]

Brain Code Camp

Underfitting and Overfitting

Assume that the true relationship is $y = x^2$, and we have collected only 10 points

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x + \hat{w}_2 x^2 + \ldots + \hat{w}_p x^p$$



Underfitting and Overfitting

Assume that the true relationship is $y = x^2$, and we have collected only 10 points

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x + \hat{w}_2 x^2 + \ldots + \hat{w}_p x^p$$