Machine Learning

Tuan Nguyen

Ngày 3 tháng 3 năm 2021

Overview

When models meet data

Dataset split

Overfitting

Ridge regression

Lasso regression

ElasticNet regression

Data

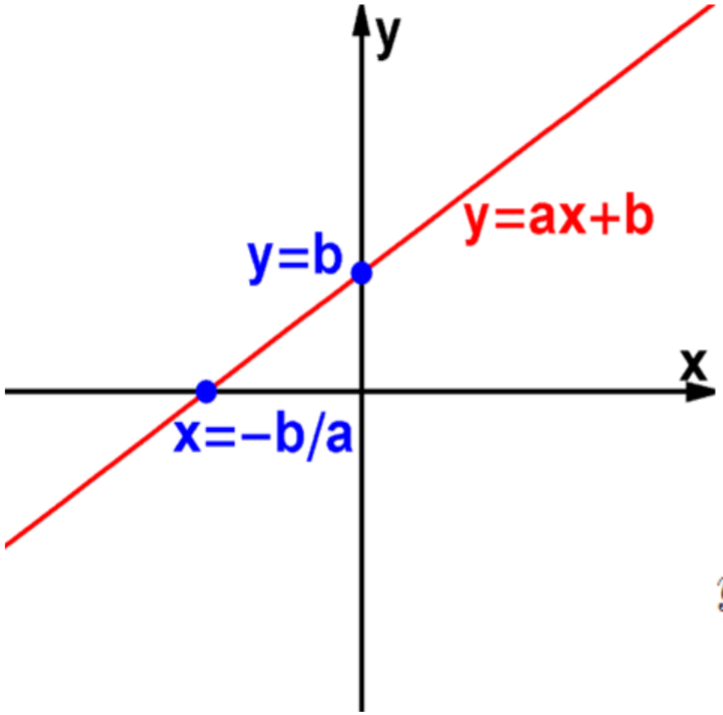
Diện tích	Giá
30	448.524
32.4138	509.248
34.8276	535.104
37.2414	551.432
39.6552	623.418
42.069	625.992

Data

$$\begin{bmatrix} 30 & 448.524 \\ 32.4138 & 509.248 \\ 34.8276 & 535.104 \\ 37.2414 & 551.432 \\ 39.6552 & 623.418 \\ 42.069 & 625.992 \end{bmatrix}$$

Hình 1: Matrix

Model



$$y = w_1 * x + w_0$$

Hình 2: Model and its parameters

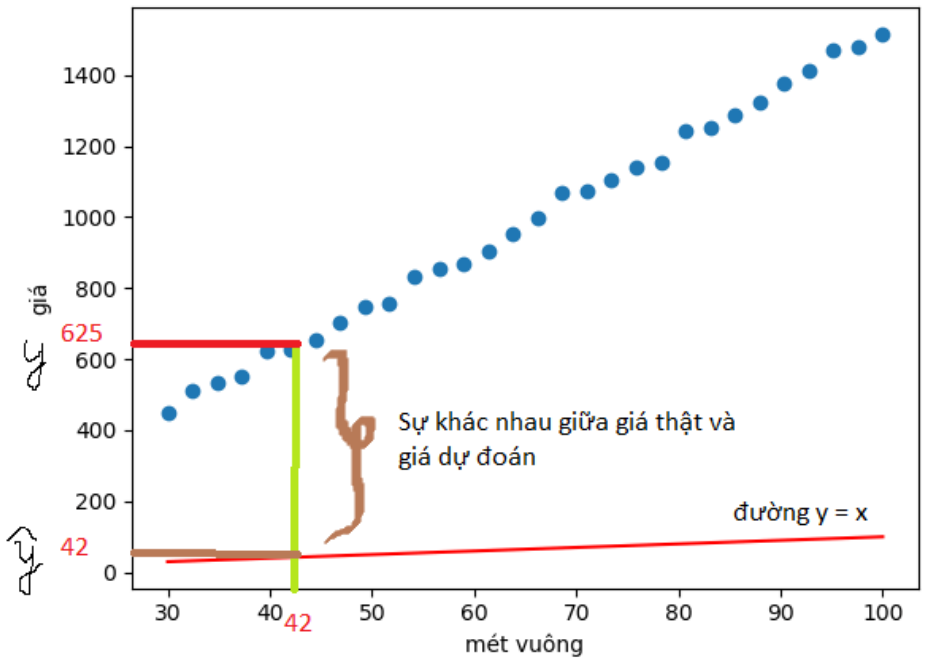
Learning

There are two main steps in ML:

- Training phase: find a model or parameters of model that performs best on unseen data.
- Prediction phase: predict unseen data.

Loss function

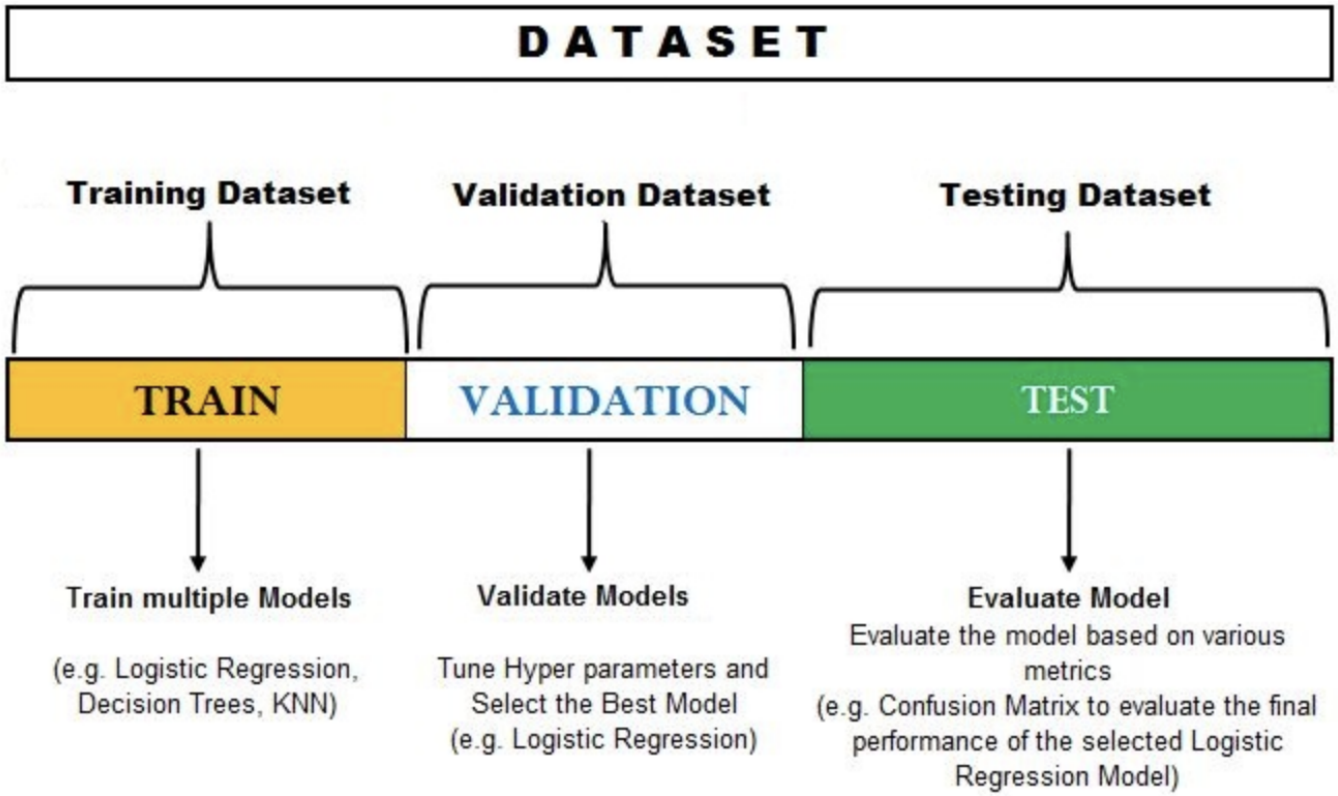
Mean Squared Error (MSE): $L = \frac{1}{2N}\sum_{i=1}^{N}(w_0 + w_1 x_i - y_i)^2$



Hình 3: Loss function

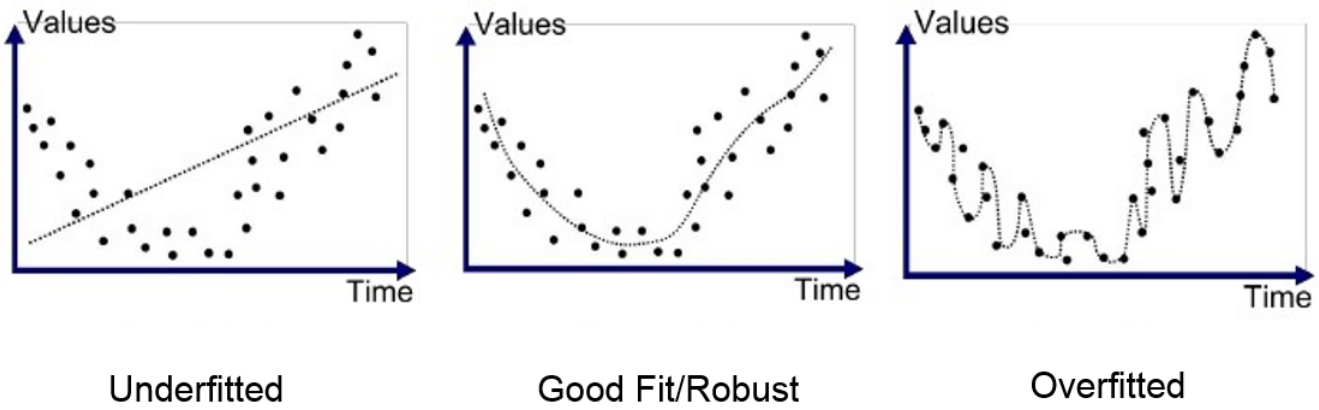
Other loss functions for regression problem?

Dataset split



Hình 4: Train test split

Overfitting



Hình 5: Overfitting and underfitting

Training set error	1%	15%	0.5%
Validation set error	11%	16%	1%

How to solve?

- Underfitting: increase complexity of model
- Overfitting:
  - Add more data
  - Regularization: L1, L2, Dropout,...
  - Early stopping
  - ...

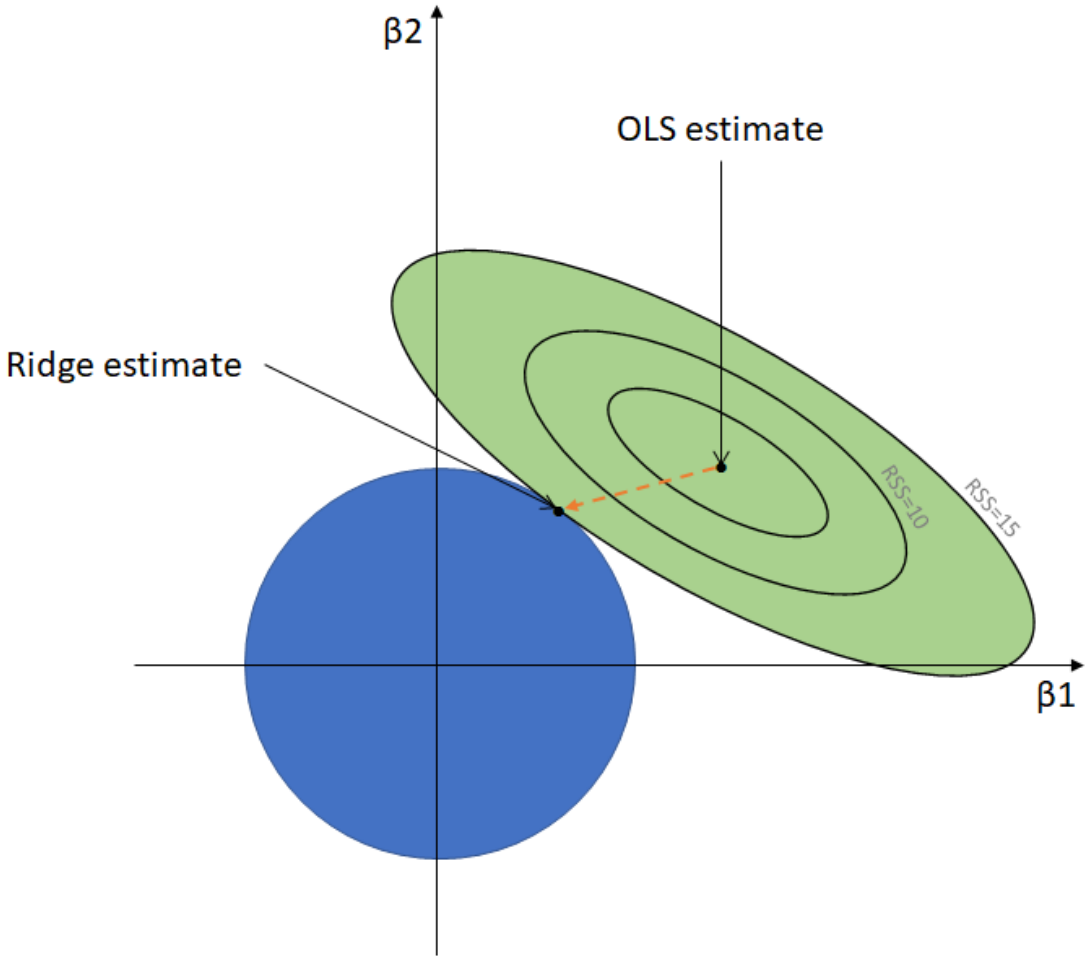
Ridge regression

$$L = \frac{1}{2N}\sum_{i=1}^{N}(w_0 + w_1 x_i - y_i)^2 + \lambda w_1^2$$

Remark:

- Loss function is added with the penalty equivalent to square of the magnitude of the all parameters.
- Ridge regression shrinks the parameters and it helps to reduce the model complexity => avoid overfitting.

Ridge regression (cont.)



Hình 6: Ridge regression

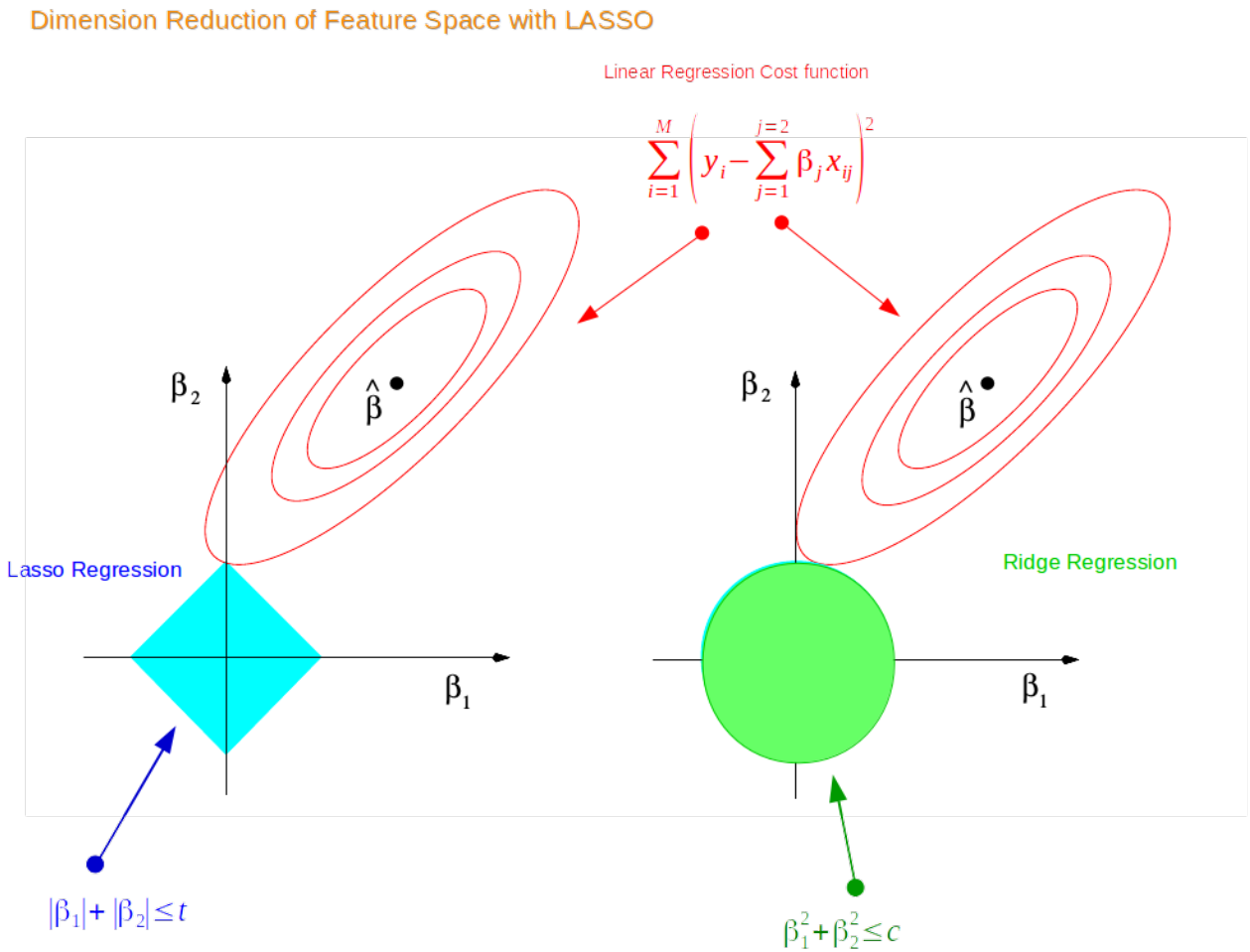
Lasso regression

$$L = \frac{1}{2N}\sum_{i=1}^{N}(w_0 + w_1 x_i - y_i)^2 + \lambda|w_1|$$

Remark:

- Loss function is added with the penalty equivalent to absolute value of the magnitude of the all parameters.
- Lasso regression not only shrinks the parameters and it helps to reduce the model complexity => avoid overfitting but also selects the important feature.

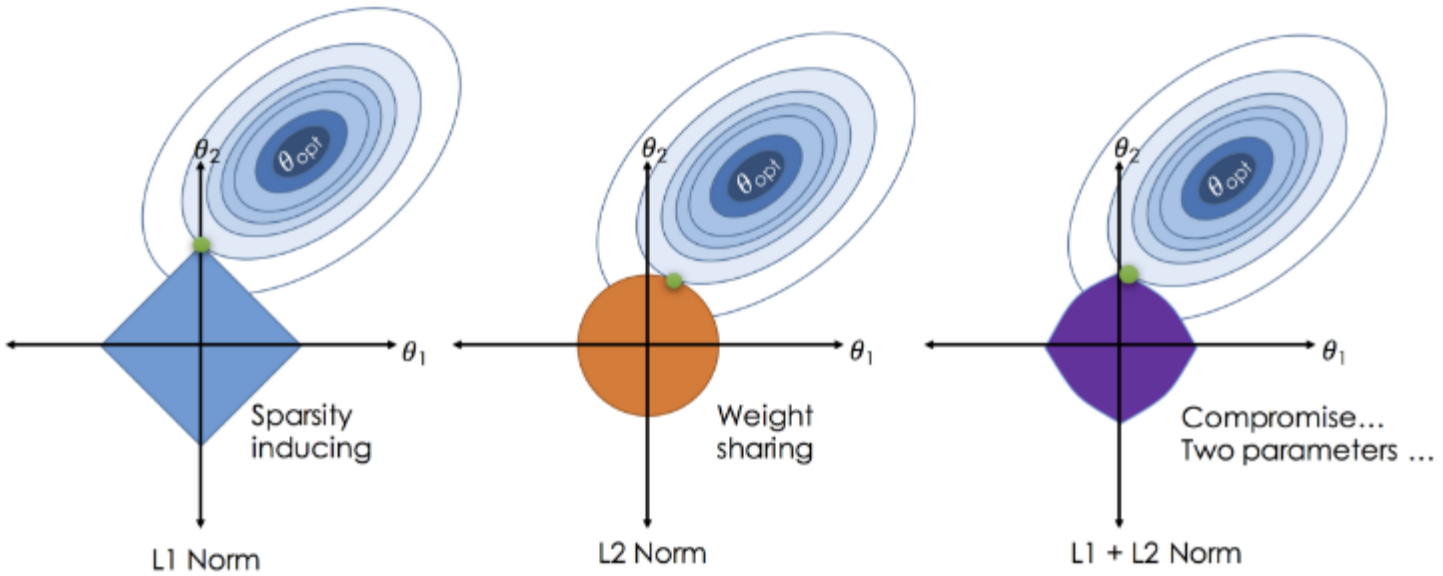
Lasso regression (cont.)



Hình 7: Lasso regression

ElasticNet regression

$$L = \frac{1}{2N}\sum_{i=1}^{N}(w_0 + w_1 x_i - y_i)^2 + \lambda(\frac{1-\alpha}{2}w_1^2 + \alpha|w_1|)$$



Hình 8: ElasticNet regression