Regularization to combat the overfitting issue

Hao Ni
University College London
The Alan Turing Institute

Overfitting and Regularization

Overfitting

When the sample size is strictly smaller than the input dimension d, there are infinite many θ such that

$$L(\theta|X, Y) = (Y - X\theta)^T(Y - X\theta) = 0,$$

and it leads to little predictive power of the estimated model in the testing dataset.

Regularization

In order to resolve the overfitting issue here, we consider the constraint optimization problem

$$\min_\beta (Y - X\beta)^T(Y - X\beta), \text{ s.t. } ||\beta|| \leq t.$$

From constraint optimization to unconstraint optimization

Lagrange multiplier

We reformulate the above constraint optimization problem as the following unconstraint one by Lagrange multiplier:

$$\min_{\lambda,\theta} L_{\text{new}}(\lambda, \theta | X, Y)$$

where

$$L_{\text{new}}(\lambda, \theta | X, Y) := (Y - X\theta)^T (Y - X\theta) + \lambda(||\theta|| - t)).$$

Therefore, it suggests us to consider the following modified loss function

$$\tilde{L}_\lambda(\theta | X, Y) = (Y - X\theta)^T (Y - X\theta) + \lambda ||\theta||,$$

where $\lambda > 0$ is a model hyper-parameter. [1, 2]

Thanks for your attention!

References I


Hao Ni, Xin Dong, Jinsong Zheng, and Guangxi Yu.
An Introduction to Machine Learning in Quantitative Finance (Chinese version).
Tsinghua University Press, 2021.


Hao Ni, Xin Dong, Jinsong Zheng, and Guangxi Yu.
An Introduction to Machine Learning in Quantitative Finance (English version).
World Scientific, 2021.