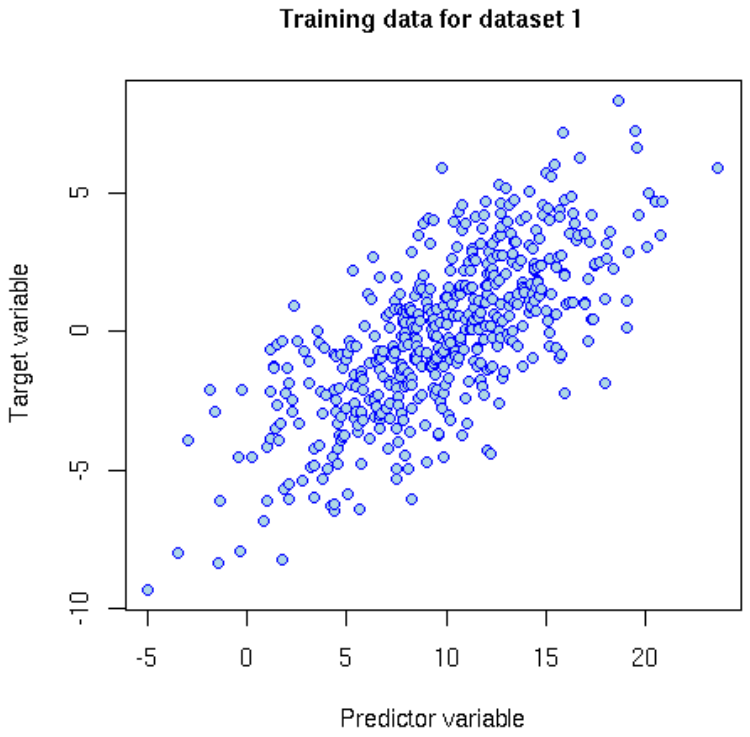
Linear regression project



Machine Learning; Mon Apr 21, 2008

Motivation

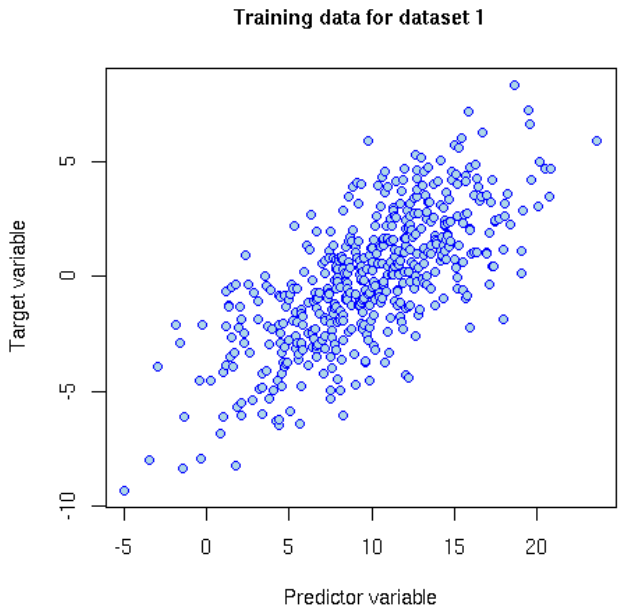
To get a feeling for the material so far, you now get a ***mandatory*** project.

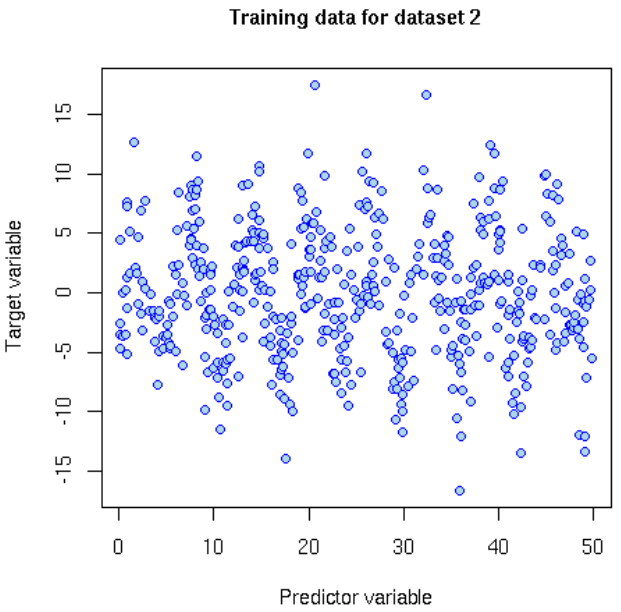
It should help you link the theory with practical machine learning problem solving.

The exercise

Problem: You are given a data set with 500 predictor variables and corresponding 500 target values.

Train a program to predict targets for new predictor variables where the targets are unknown.





The exercise

Download training predictor and target values for the five datasets.

Select your feature basis functions and construct your model matrix Φ.

Obtain weight vector by solving $\Phi^{\mathrm{T}}\Phi\mathbf{w}=\Phi^{\mathrm{T}}\mathbf{t}$.

Download new predictors and use the trained model to predict corresponding targets as

$$\hat{t} = y(x, \mathbf{w}) = \mathbf{w}^T \phi(x)$$

URL: http://www.daimi.au.dk/~cstorm/courses/ML_f08/project1.html

The exercise

Download training predictor and target values for the five datasets.

Select your feature basis functions and construct your model matrix Φ.

It might be a good idea to consider more than one model i.e. set of basis functions, and select the best performing.

Watch out for over-fitting! You have plenty of data, so use it for both training and test data!

to predict corresponding targets as

$$\hat{t} = y(x, \mathbf{w}) = \mathbf{w}^T \phi(x)$$

URL: http://www.daimi.au.dk/~cstorm/courses/ML_f08/project1.html

Handing it in

If you choose to hand in this project, you can email me your target values for the non-training data by **12:00 Fry May 9** and I will check them up against the true values.