Robust linear regression

It is very common to model the noise in regression model using a Gaussian distribution. However, if we have outliers in our data, this can result in a poor fit. This is because squared error penalizes deviations quadratically, so point far from the line have more affect on the fit than points near to the line. One way to achieve robustness to outliers is to replace the Gaussian distribution for the response variable with a distribution that has heavy tails, such as Laplace distribution.

Ridge regression

One problem with ML estimation is that it can result in overfitting – if we changed the data a little, the obtained coefficients would change a lot. We can encourage the parameters of MLE to be small, thus resulting in a smoother curve, by using a zero-mean Gaussian prior.

In general, adding a Gaussian prior to the parameters of a model to encourage them to be small is called ℓ𝟐 regularization or weight decay.

Перемножаем векторы значения Х

