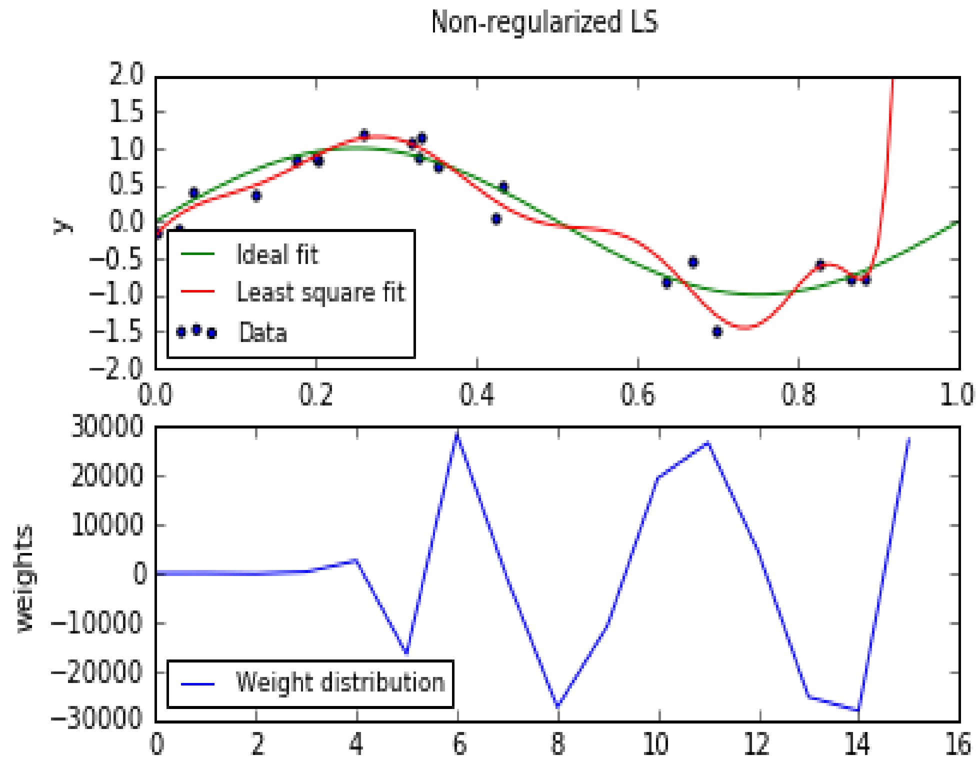
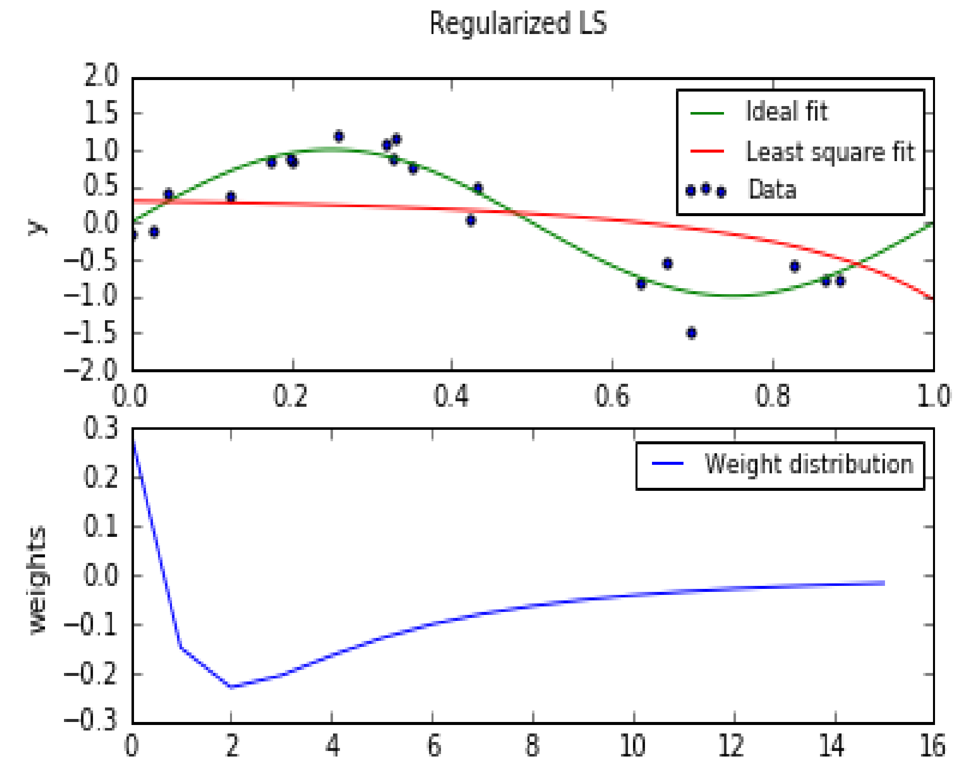


Left figure shows that the Non-regularized least square fit is basically an example over-fit because it adapts to the data rather than generalizing. We can observe that the value of the weights is high.



Right figure shows that the regularized least square fit generalizes the problem. We also can observe that the value of the weights is not high because of the penalty regularization constant.

X=(mxn) matrix input,**w**=(mx1) weight, **y**=(nx1) output, **λ**>0 is regularization constant, **I**=Identity matrix.

Regularised least square problem is of form:

min[1/2.(wt.X-y)t (wt.x-y) +λ/2.wt.w]

H(w)= min[1/2 (wt.X-y)t (wt.x-y) +λ/2.wt.I.w]

d(H(w))/d(w)=X(Xt.w-y)+λ.I.w=0

X.Xt.w-Xy+λ.I.w=0

(X.Xt+λ.I)w=XY

w=(X.Xt+λI)-1.(XY)

Regularized least square