Given a training data set, the input data samples are randomly generated subject to Circularly Symmetric Gaussian Distribution, the output data is generated by a given target linear function, Which is given by

$$d_n = W^T x_n \quad (1)$$

Where W denotes the coefficient vector which is deterministic, x_n denotes the nth input data samples, and d_n denotes the true output of the linear function.

The observations $d_n^{obervation}$ are disturbed by additive white Gaussian noise, denoted by ω. The signal to noise ratio (SNR) is defined as the ratio between the expectation of true output data and the variance of the Gaussian noise. Thus the observations can be given by

$$d_n^{observation} = d_n + \omega \quad (2)$$

The experiments are performed by first training support vector regression (SVR) or minimum mean square error (MMSE) estimators by a randomly generated training data set, then the prediction performances are evaluated by prediction risk of the new testing data set(mean square error of the deviation between the true function output and the output of the regression estimation, that is $\mathrm{MSE} = ||d_n - \mathrm{f}(x_n)||^2$), where d_n is the true output of the linear function and $\mathrm{f}(x_n)$ is the estimation output by SVR or MMSE estimator. The testing data set is uniform distributed but independent to training data set and are generated randomly, the size of the testing data set is 1000.

Because the prediction accuracy is a random variable determined by different realizations of the training-testing process, thus the training-testing procedure is repeated independent in a large number of realizations (100) by different and independent training and testing data sets. The MSE is the average of all the MSE in different realizations.

The first experiment is designed to test the performances of SVR and MMSE given different numbers of training date set. The SNR is fixed by 20dB.

The dimension of the linear function as shown in (1) is 30, the number of training data sets are [30, 40, 60, 100, 150].

The MSE are evaluated by a large number of testing data samples (1000) over 100 training-testing realizations.

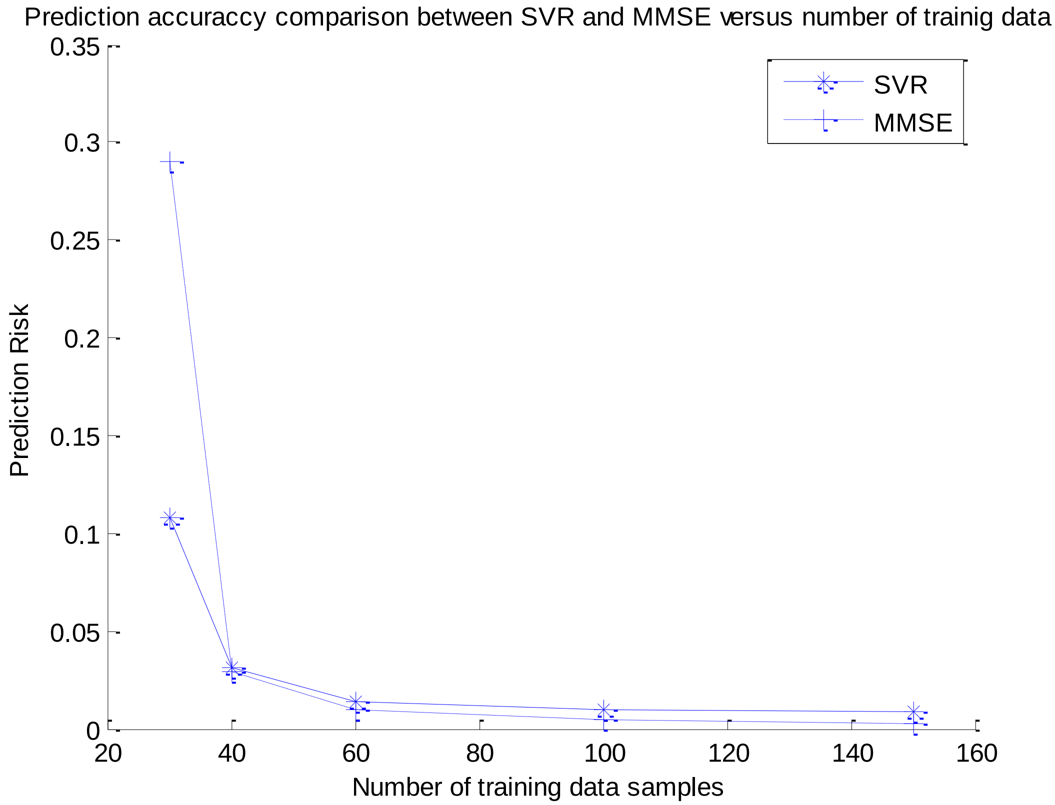

Figure 1

The hyper-parameters settings of SVR are

C=1.000000

ϵ =0.0000001000

tolerence of the duality gap=0.0010000000

Figure 1 shows the prediction risk comparison between the SVR and MMSE, it indicate the SVR can outperform MMSE in a small training data set (30), but performs worse when it comes to large number of training data.

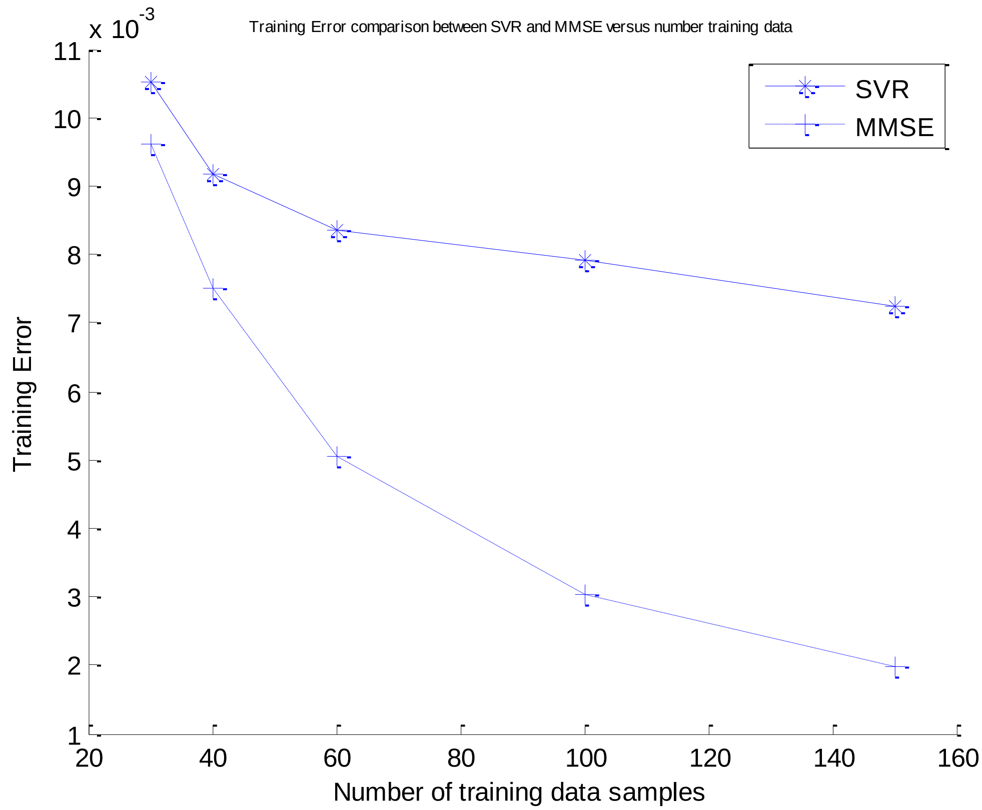

Figure 2

Figure 2 shows the MSE performances of training data set of SVR and MMSE, shows MMSE can achieve lower MSE for training data set over SVR.

Then the Second experiment is performed to evaluate the performances of SVR and MMSE under different SNRs.

The hyper-parameters settings of SVR are

C=1.000000

ϵ =0.0000001000

tolerence of the duality gap=0.0010000000

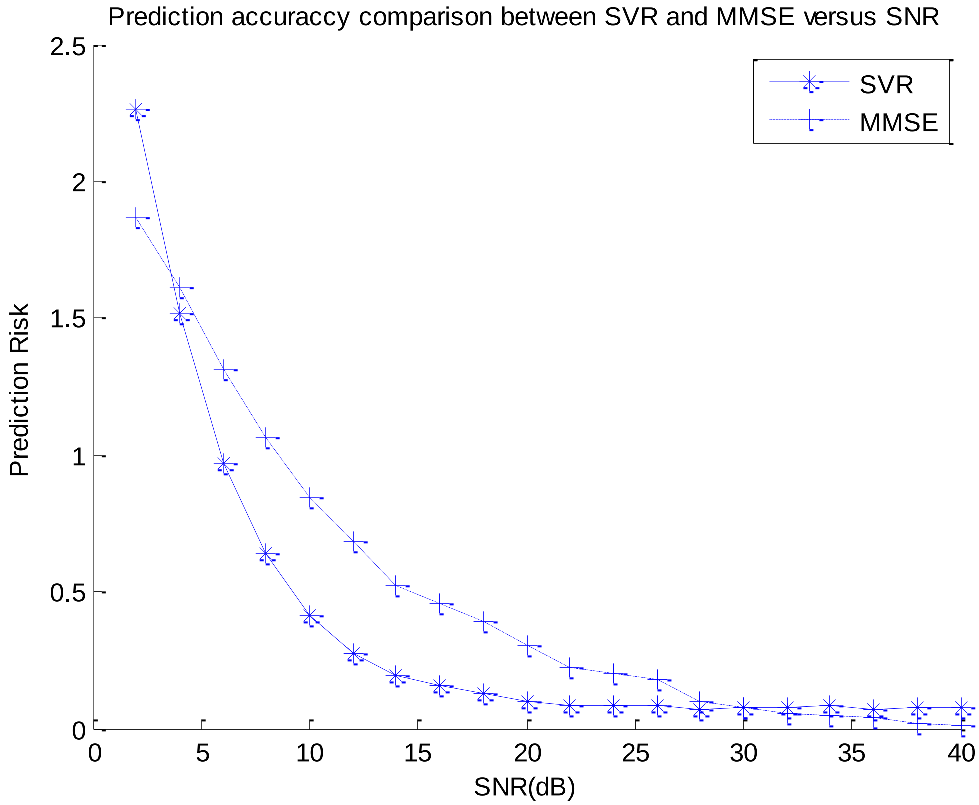

Figure 3

Figure 3 shows the prediction performances comparison between SVR and MMSE over different SNR. From figure 3, it can be seen that SVR outperforms MMSE in 5 to 30 dB region, but performs worse when the SNR gets higher.

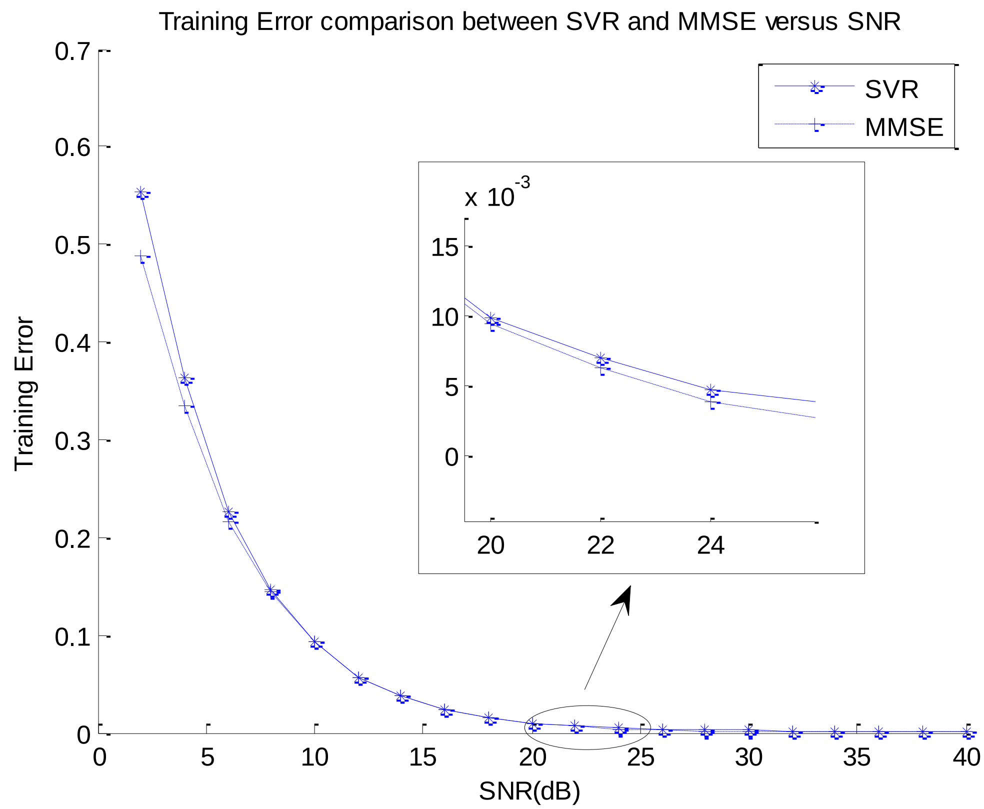

Figure 4

Figure 4 shows the MSE performance over training data set. As we can see in 5-25 dB SNR region, MMSE outperforms SVR in training error, but as shown in Figure 3, the prediction performance of SVR outperforms MMSE in this SNR region.