2.3 Extreme points

In (Fig. 2.3-1) the results for a prediction using a value for σ resulting out of the wiggle method is shown. The circles are the available training samples used for GRNN in the prediction. The dashed line is the result of the prediction. The prediction fits the training samples very well in the range from x=0.1 to x=0.3 and x=0.6 to x=0.9. The training sample at x=0.0 and the prediction deviate from each other significantly. The prediction for the training sample at x=1.0 deviates respectively. The training samples at x=0.4, x=0.5 and x=0.6 and the prediction deviate as well but not as significantly as for the values at the edges.

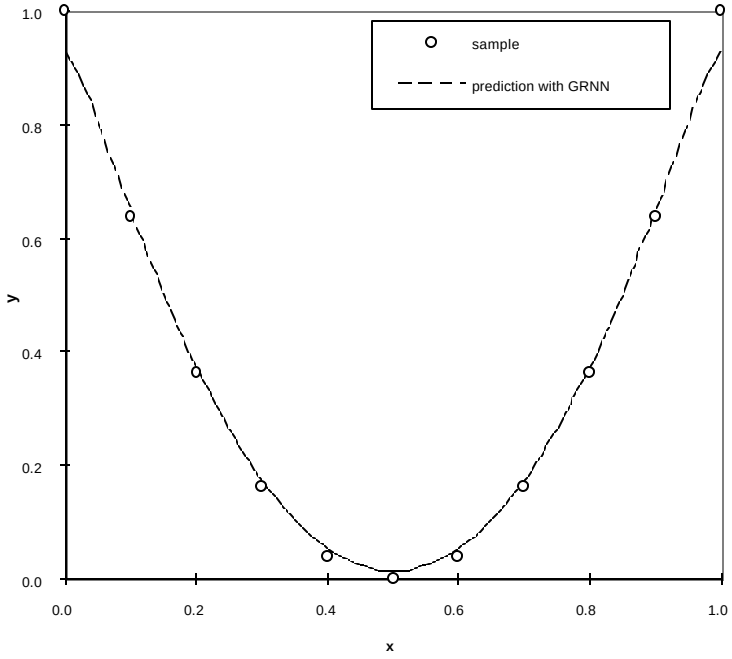

Fig. 2.3-1 GRNN with Problem at extreme points

Every prediction of any point is influenced by all the training samples but mostly by the surrounding ones. If the surrounding training samples all have smaller values than the prediction at this point will be influenced such that it will as well follow the trend of the other training points. For a maximum in $y=f(X)$, GRNN will predict the maximum at a smaller value. For a minimum in $y=f(X)$, GRNN will predict the minimum with a bigger value. For extreme value at the edges of the area of the training samples, the same underestimation or overestimation happens (Fig. 2.3-1). For edges in addition to the influence of the neighboring points with smaller values more points missing on the other side of the border to influence the prediction and therefore the prediction levels off even more.

The underestimation or overestimation of extreme points cannot be influenced by the selection of sigma without other impact on the prediction. The leveling off towards the edges introduces another inflection point. This must therefore be taken into account as the number of inflection points is selected.

2.4 Unequally spaced data

The normal distribution has a symmetrical distribution around the mean value. This is the bell shape of the normal distribution. The normal distributions for GRNN are chosen such that every training sample is the mean of one normal distribution. Training samples that are not equally spaced influence the prediction in areas of higher density of the training samples differently than in areas of low density of the training samples (Fig. 2.4-1, Fig. 2.4-2).

The training samples in (Fig. 2.4-1) and (Fig. 2.4-2) originate from the same function that was used to get the training samples in (Fig. 2.3-1). The samples at x=0.1 was moved closer to the sample at x=0.2. The sample at x=0.3 was moved closer to the sample at x=0.4 and the sample at 0.6 was moved to x=0.55. This was done to test what influence unequally spaced data has on the prediction. The smoothness parameter was again selected using the wiggle method. For the example in (Fig. 2.4-1) two inflections were allowed, for the example in (Fig. 2.4-2) four inflections were allowed. The result in (Fig. 2.4-1) shows a smooth curve. The precision of the prediction at the training points is not very high for the first example. The curve is very smooth and does not show any wiggles. The two only inflection points are located towards the edge, because of the previous mentioned effects at borders (Section 3.4).

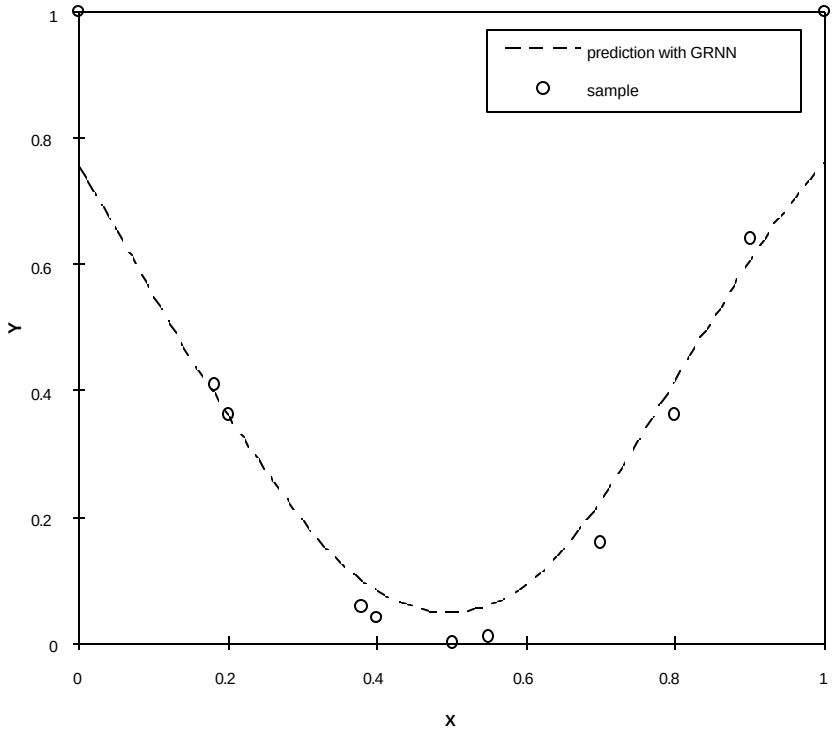

Fig. 2.4-1 GRNN with problems having unequally spaced training samples, two inflections were allowed

The training samples in (Fig. 2.4-2) are exactly the same as in the example in (Fig. 2.4-1). The number of inflection points that were allowed in the wiggle method were four instead of two. The precision at the training samples is increased but in the same instance the smoothness decreased. The curve now has four wiggles. Here again the tradeoff between smoothness and precision has to be made.

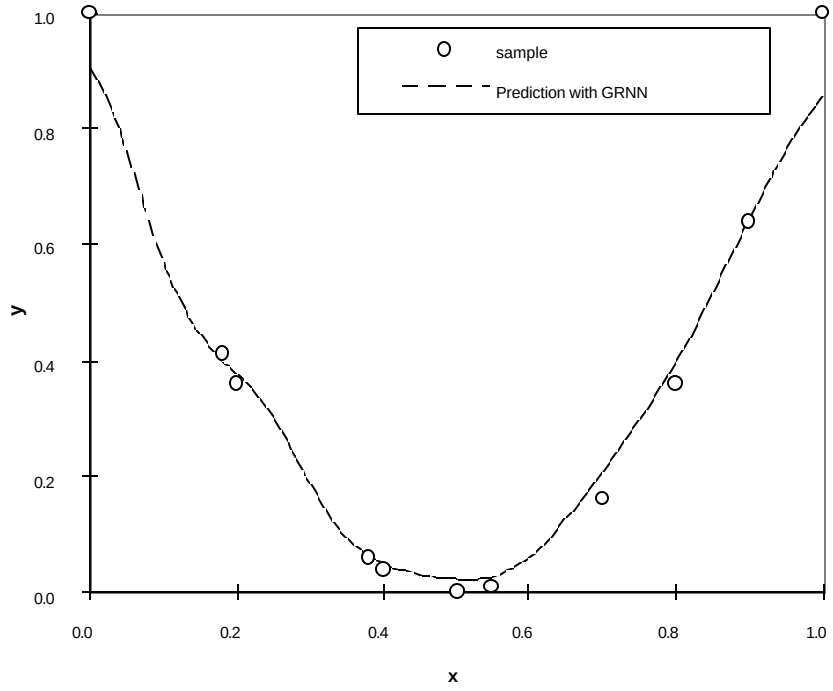

Fig. 2.4-2 GRNN with problems having unequally spaced training samples, four inflections were allowed

Compared to (Fig. 2.3-1) the curves in (Fig. 2.4-1) and (Fig. 2.4-2) are very imprecise. The matter of equally spaced data is very important. It would be very desirable to always have equally spaced data. Data resulting from measurements on technical systems is not very likely to be equally spaced. It is may be possible to obtain enough training samples to get equally spaced data by using a Nonparametric Regression [Hastie]. The nonparametric regression calculates the average value of training samples in a certain area and then sets the average equal to a new training sample that is positioned such that the new training samples are equally spaced. In (Fig. 2.4-3) the general way a nonparametric regression works is shown for an example with two input variables. The small gray circles are positions for which

measurements are available. The big circles are the borders of the area for which the average of all the samples in this are is calculated. The areas for which the average is calculated can be changed in size and the areas can overlap too. The average that was calculated is the new value of the training sample which is centered at the indicated point.

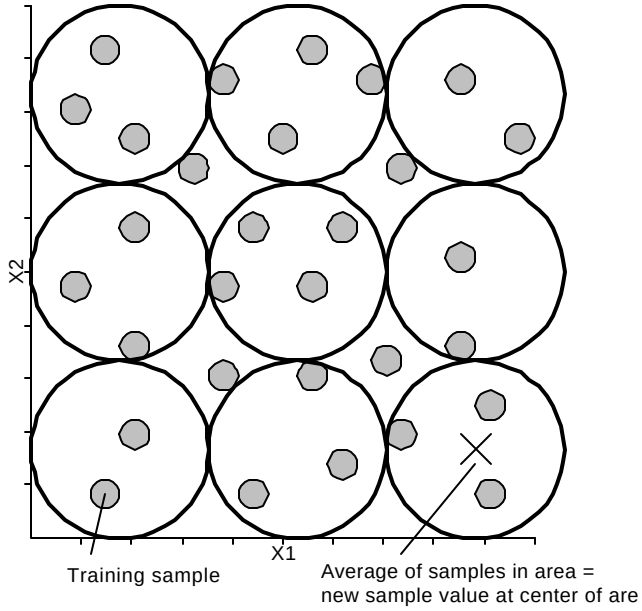

Fig. 2.4-3 Nonparametric Regression

This procedure reduces the number of training samples and causes a loss of information. Loss of information can result in the loss of significant points that include information such as information about extreme points or drastic changes in the slope. Skipping such information seems not to be a very good tactic. In addition to that, there is probably not enough data to allow a regression and a further reduction of data. Results of this appraoch are

not shown. This aproach could be applicable for problems having a large amount of data available.

2.5 Multi-Dimensions

GRNN should be able to predict the function value for a given multi-dimensional input vector. But the precision declines with increasing numbers of dimensions. This is not surprising since more things have to be taken into account than for one dimension. One problem can be that the influence of the neighboring points results in a larger deviation of the expected value. Another problem is that one smoothness parameter has to accommodate **all** the needs for **all** the dimensions. It can be that one dimension needs a larger value for the smoothness parameter than the other dimension. This especially happens if there are different numbers of samples for changes in one dimension. It was shown earlier that the smoothness parameter has a very big influence on the quality and shape of the curve.

One possibility to circumvent the impact of multi-dimensionality is to split the problem in two or more parts (Fig. 2.5-1). This is possible under the assumption of additive independence. Assume a function $Y=f(X_1,X_2)$. If a change in X_1 by DX_1 results in a constant change of DY in Y for any value of X_2, then X_1 is independent from X_2. If the same is true for X_2 then X_1 and X_2 are independent. Under this assumption of independence the problem can be split up in two parts. Y is then a function of $Y=g(X_1)+h(X_2)$. Now the different parts of Y can be predicted independently. This reduces the necessary number of training samples but also makes it necessary to get the training samples under very specific conditions. In Chapter three the approach using independence assumptions was tested on a technical example.

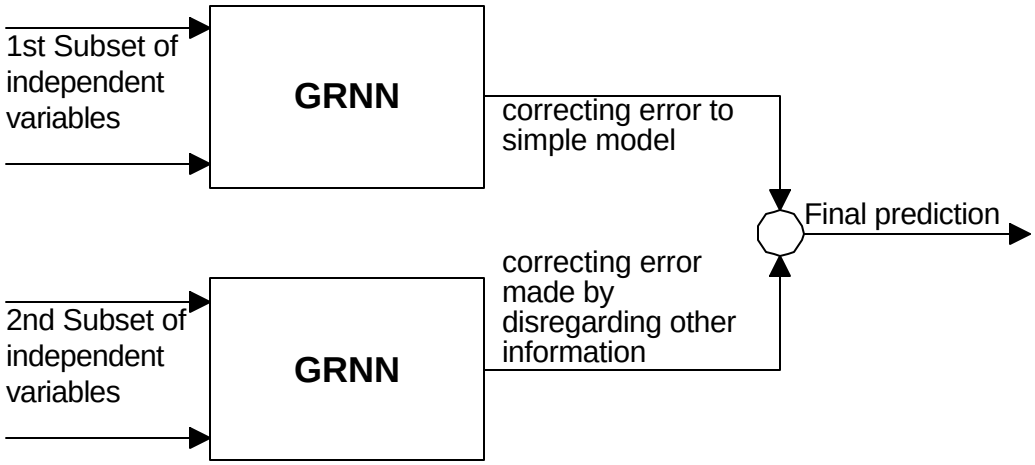

Fig. 2.5-1 A way to circumvent the influence of multi-dimensionality for independence of the inputs