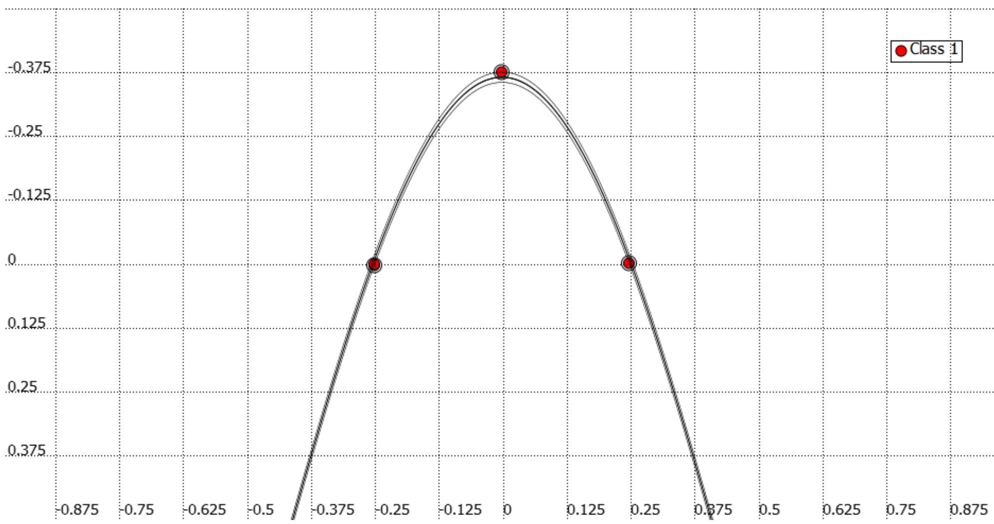
EXERCISE SESSION Nonlinear Regression: ADVANCED MACHINE LEARNING COURSE – EPFL – Lecturer A. Billard

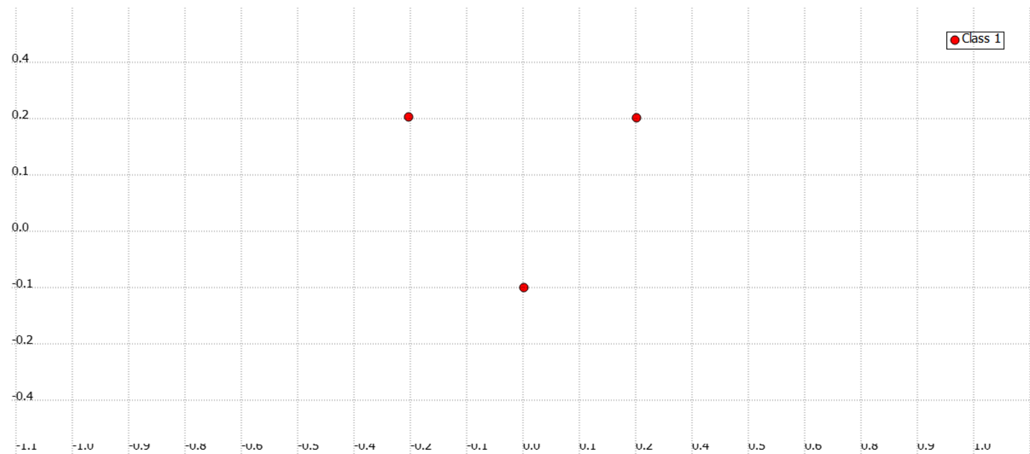
Exercise 1: Support Vector Regression

1) Show with a 2-dimensional schematic (where the first coordinate is used to predict the second coordinate, i.e. $x_2 = f\left(x_1\right) = \sum_{i=1}^{M}\left(\alpha_i^* - \alpha_i\right)k\left(x_1^i, x_1\right) + b$, that you can fit any combination of points when using SVR with Gaussian kernel. How many data-points at minimum do you need as support vectors? Can the ε-tube have an effect on this minimum number of points?

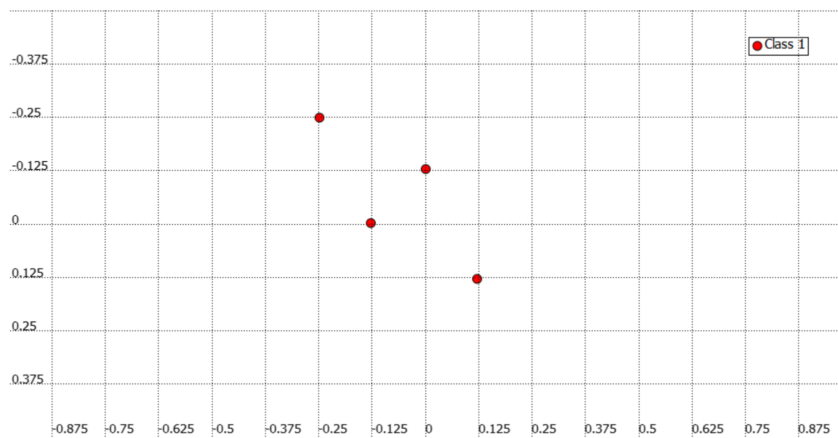
2) The inferred function f below was fit with Gaussian kernel and a kernel width of 0.8. What is the effect of increasing the kernel width on the function?



3) What minimum order of a homogeneous polynomial kernel do you need to achieve good regression on the set of 3 points below? And how many support vectors do you need at minimum?



4) The dataset below cannot be fitted with polynomial kernel of order 3. Why not? Does increasing the order of the polynomial help fit these points?



5) Give examples of datasets you can fit with polynomial kernel of order 3.

<u>Exercise 2: Gaussian Process Regression and Gaussian Mixture Regression</u>

Show analytically that GPR and GMR can become equivalent under certain restrictions.

<u>Exercise 3: Equivalence between GPR and SVR</u>

a) Using an rbf kernel, for what value of the parameter b in SVR, the regression value far away from the data would be the same as the mean GPR value.

b) How would the formulation in SVR change if it is required to learn the SVR function with a fixed value of $b = 0$ (Hint: What is the effect of b in the dual).

c) Show that if noise is not considered in both GPR and SVR, the modified SVR above (fixed b) and the GP mean regression values are equivalent.