Machine Intelligence 1 **Exercise Sheet 10**
WS 2017/18, Obermayer/Augustin due: 24.01.2017 at 23:55

Support Vector Regression

Exercise T10.1: Regression with SVM (tutorial)

In support vector regression problems, we assume that we are given a training data set

$$\{(\underline{\mathbf{x}}^{(\alpha)}, y_T^{(\alpha)})\}, \quad \alpha \in \{1, \ldots, p\}, \qquad \underline{\mathbf{x}} \in \mathbb{R}^N, \quad y_T \in \mathbb{R},$$

and want to fit the linear regression function

$$y(\underline{\mathbf{x}}) \quad = \quad \underline{\mathbf{w}}^\top \underline{\mathbf{x}} + b\,.$$

(a) What is the ε-insensitive cost function for regression?

(b) Derive the primal problem of the ε-support vector regression (ε-SVR).

(c) The optimal ε-parameter depends linearly on the noise level in the data, which is unknown. Derive the primal problem for the ν-SVR, which adjusts ε as a primal parameter.

(d) Derive the Lagrangian of the ν-SVR.

Exercise H10.1: The dual problem of the ν-SVR (homework, 5 points)

In this exercise you will derive the dual problem of the ν-SVR.

(a) (2 points) Calculate the derivatives of the Lagrangian with respect to the primal variables.

(b) (3 points) By setting the derivatives from a) to zero and using the results to eliminate the primal variables from the Lagrangian show that the *dual problem* takes the following form:

$$\max_{\lambda_\alpha, \lambda_\alpha^*} -\tfrac{1}{2} \sum_{\alpha,\beta=1}^{p} (\lambda_\alpha^* - \lambda_\alpha)(\lambda_\beta^* - \lambda_\beta)(\underline{\mathbf{x}}^{(\alpha)})^T \underline{\mathbf{x}}^{(\beta)} + \sum_{\alpha=1}^{p} (\lambda_\alpha^* - \lambda_\alpha) y_T^{(\alpha)}$$

s.t. $\forall \alpha \in \{1, \ldots, p\}$:

$$0 \leq \lambda_\alpha \leq \tfrac{C}{p}\,, \qquad 0 \leq \lambda_\alpha^* \leq \tfrac{C}{p}\,, \qquad \sum_{\alpha=1}^{p} (\lambda_\alpha - \lambda_\alpha^*) = 0\,, \qquad \sum_{\alpha=1}^{p} (\lambda_\alpha + \lambda_\alpha^*) \leq \nu C\,.$$

Exercise H10.2: Regression with the ν-SVR (homework, 5 points)

In this exercise you will apply ν-SVR from a software package of your choice (e.g. `libsvm`) to the data set used in exercise sheet 5. The training set `TrainingRidge.csv` and the validation set `ValidationRidge.csv` can be found on ISIS. Do **not** center, whiten or expand the data before training (otherwise the following hyperparameter ranges are inadequate).

(a) (2 points) Train the ν-SVR on the training set with the standard parameters of your library ("out of the box"). Plot the model prediction for the validation set as an image plot (where colors represent the output values, the axes represent the two coordinates: x_1 and x_2). Plot the data points from the training set as colored rectangles in the same axis.
Make a second plot over x_1 and x_2 taking the true labels of the validation set. Compute the total mean squared error (MSE) between model prediction and true labels of the validation set.

(b) (2 points) Perform a 10-fold nested cross-validation with a ν-SVR and parameters $\nu = 0.5$ and $C \in 2^i$, $i \in \{-2, \ldots, 12\}$. Use a Gaussian RBF kernel with $\gamma \in 2^j$, $j \in \{-12, \ldots, 0\}$. Plot the resulting mean (test set) MSE over the folds as an image plot. Note that the RBF kernel is parametrized as in the previous sheet (with parameter γ instead of σ).

(c) (1 point) Extract the best parameter combination C and γ. Train the entire training set with these parameters and plot the model prediction for the validation set as an image plot. Compare the plot with the true labels and the results from (a). Compute also the total mean squared error for the validation set for comparison.

Total 10 points.