Machine Learning
Prof. S. Harmeling

Monday, 29. October 2018
DUE 10:30 Monday, 5. November 2018

**Exercise set #4**

1. **MLE for the mean of a Gaussian distribution**
   Calculate the MLE estimator of $\mu$ for a univariate Gaussian likelihood $p(y|\mu) = \mathcal{N}(y|\mu, \sigma^2)$.

   *15 points*

2. **Bayesian linear regression**
   Assuming the model from the lecture

$$\begin{aligned} p(w) &= \mathcal{N}(w|w_0, V_0) \\ p(y|X, w) &= \mathcal{N}(y|Xw, \Sigma) \end{aligned}$$

   derive the mean and variance of the posterior $p(w|X, y) = \mathcal{N}(w|w_n, V_n)$. For this write out the exponent of the product of prior and likelihood and bring it to the form of a multivariate second-order polynomial. Then apply Lemma 4.6 (3) from lecture 4 to get expressions for $\eta$ and $\Lambda$. Reordering gives you expressions for $w_n$ and $V_n$.

   *40 points*

3. **Polynomial linear regression (programming task)**
   Goal of this task is to fit a polynomial through data points $(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$. Assume that the outcome $y = (y_1, \ldots, y_n)^T$ follows a normal distribution $\mathcal{N}(y|Xw, \sigma^2 I)$, where

$$X = \begin{bmatrix} | & | & | & & | \\ 1 & x & x^2 & \ldots & x^d \\ | & | & | & & | \end{bmatrix}$$

   (a) Write a function that generates the matrix $X$ for $x = (x_1, \ldots, x_n)^T$.
   (b) Implement the estimator $w_{\text{MLE}}$.
   (c) Implement a function that calculates the error $\text{MSE}(w) = \frac{1}{n}(Xw - y)^T(Xw - y)$.
   (d) Try to find a good polynomial degree $d < 20$ that leads to a small test error. Plot your best solution together with the training data.
   (e) Plot the training and test errors against the degree of the polynomial. A paper-pencil plot on squared paper is fine. What do you observe?
   (f) Implement the estimator $w_{\text{RIDGE}}$.
   (g) Find a good combination of $d$ and $\lambda$ that gives you a small test error. Is it smaller than the test error of the optimal solution from (d)?

   If you have problems with this exercise, let us know! Training data and test data are available in sciebo.

   *45 points*