Exercise set #6

Please submit your solutions in teams of two using the sciebo file-drop folder. The link is available in ILIAS. For the formatting please stick to the `submission_guideline.pdf` that you can find on sciebo. In the case of multiple uploads we will consider the latest. Uploads after the deadline will be deleted without further notice.

1. **Equality constrained least squares problem**
   Let $m \geq n \geq k$, $y \in \mathbb{R}^m, b \in \mathbb{R}^k$ and $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{k\times n}$ having full rank. Consider the following optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \|Ax - y\|_2^2 \\ \text{s.t.} \quad & Bx = b. \end{aligned}$$

   Find a matrix $P \in \mathbb{R}^{(n+k)\times(n+k)}$ and a vector $p \in \mathbb{R}^{n+k}$ such that solving

$$P \begin{bmatrix} x \\ \lambda \end{bmatrix} = p$$

   gives a critical point for the optimization problem. You can assume that P is invertible.

   Hint: Start with defining the Lagrangian function for this problem. After that calculate its gradient with respect to x and λ and set it to zero.

   20 points

2. **Dual problem of ν-SVM for the non-separable case**
   The ν-SVM is another realization of the SVM for the non-separable case. Instead of the parameter C it uses $\nu \in [0, 1]$ which is the proportion of support vectors that lie on the wrong side of the hyperplane. The *primal* problem of a ν-SVM is given as:

$$\begin{aligned} \min_{w,b,\xi,\rho} \quad & \frac{1}{2}||w||^2 - \nu\rho + \frac{1}{n}\sum_{i=1}^{n} \xi_i \\ \text{s.t.} \quad & y_i(w^T x_i + b) \geq \rho - \xi_i \text{ for all } i \\ & \xi_i \geq 0 \text{ for all } i \\ & \rho \geq 0 \end{aligned}$$

   Derive the dual problem by completing the following subtasks:

   (a) Define the Lagrangian by introducing dual variables $\alpha_i \geq 0$, $\beta_i \geq 0$ and $\delta \geq 0$.
   (b) Calculate the derivatives wrt w, b, ξ and ρ.
   (c) Set the derivatives to zero and plug the resulting equations back into the Lagrangian to eliminate w, b, ξ and ρ.
   (d) Do you get the constraints $\frac{1}{n} \geq \alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i \geq \nu$? Try to derive it, if it is missing!
   (e) Write down the resulting maximization problem with constraints on α_i.

   30 points

3. **Dropping the one**
   Consider the linearly separable case of a SVM, but replace in the constraints the 1 with 0, i.e.

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}||w||^2 \\ \text{s.t.} \quad & y_i(w^T x_i + b) \geq 0 \text{ for all } i \end{aligned}$$

   Consider a data set that is separable, what will happen to w and b? Will it still work? What might go wrong? What happens if you replace 0 by 0.5?

   10 points

4. **Implement your own SVM (programming task)**
   Implement your own linear SVM for the non-separable case. Please do not use any pre-built SVM implementations.

   (a) Start with writing down the quadratic programming formulation (see section 11 slide 31) for the primal problem. Describe how you choose $Q, c, A, b, A_{eq}, b_{eq}, x_l, x_u$.

   (b) Generate training data using the following code snippet:

```
import numpy as np
np.random.seed(123)
x1 = np.random.randn(2, 20)
x2 = np.random.randn(2, 20) + 2
y1, y2 = np.ones(20), -np.ones(20)
X_train = np.hstack([x1, x2]).T
y_train = np.hstack([y1, y2]).T
```

   (c) Solve the quadratic programming problem for the given training data. You can use the `quadprog` function from the lecture.

   (d) Visualize the training data and plot the decision boundary.

   40 points