Exercises for
Pattern Recognition
Peter Fischer, Shiyang Hu
Assignment 10, 13./16.01.2014

SOLUTION

General Information:

Exercises (1 SWS):	Tue 12:15 – 13:45 (0.154-115) and Fri 08:15 – 09:45 (0.151-115)
Certificate:	Oral exam at the end of the semester
Contact:	peter.fischer@fau.de, shiyang.hu@fau.de

Support Vector Machine & Constrained Optimization

Exercise 1 In this exercise, you will create a *Support Vector Machine* (SVM) classifier for a two-class problem. We perform classification in the original feature space (i.e. you don't need to apply kernel functions to lift the features into a space of higher dimensionality). Also, we solve the optimization in the primal form. For non-linear optimization, we suggest to apply the Matlab function `fmincon`, which provides the constrained optimization algorithm that is required to optimize the hyperplane. The caller to this function has to provide a handle to the objective function and a linear system of inequalities.

Decision boundary for SVM: $F(\boldsymbol{x}) = \boldsymbol{\alpha}^T \boldsymbol{x} + \alpha_0$.
Hard margins
Objective function: $\operatorname{argmin}_{\boldsymbol{\alpha}} \frac{1}{2} \|\boldsymbol{\alpha}\|^2 = f'(\boldsymbol{x}')$
Constraints:

$$
\begin{aligned}
y_i\left(\boldsymbol{\alpha}^T \boldsymbol{x}_i + \alpha_0\right) - 1 &\geq 0 \\
y_i \alpha_1 x_{1i} + y_i \alpha_2 x_{2i} + y_i \alpha_0 &\geq 1 \\
\begin{pmatrix} y_i x_{1i} & y_i x_{2i} & y_i \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_0 \end{pmatrix} &\geq 1 \\
\underbrace{-\begin{pmatrix} y_i x_{1i} & y_i x_{2i} & y_i \end{pmatrix}}_{\boldsymbol{A}'_i} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_0 \end{pmatrix}}_{\boldsymbol{x}'} &\leq \underbrace{-1}_{\boldsymbol{b}'}
\end{aligned}
$$

The size of $\boldsymbol{A}'$ depends on the number of training samples.
$\boldsymbol{x}'$ depends on the dimensionality of the features.
In MATLAB, use fmincon with linear equality constraints $\boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b}'$ and objective function $f'(\boldsymbol{x}')$.

The gradient is calculated as:

$$\frac{\partial f'}{\partial \alpha_1} = \alpha_1$$
$$\frac{\partial f'}{\partial \alpha_2} = \alpha_2$$
$$\frac{\partial f'}{\partial \alpha_0} = 0$$

(a) Implement the SVM with hard margin constraints.

(b) Extend your implementation with soft margin constraints.

Soft margins
Objective function: $\operatorname{argmin}_{\boldsymbol{\alpha}} \frac{1}{2} \|\boldsymbol{\alpha}\|^2 + C \sum_i \xi_i = f'(\boldsymbol{x}')$
Constraints of the decision boundary:

$$y_i \left(\boldsymbol{\alpha}^T \boldsymbol{x}_i + \alpha_0\right) - 1 + \xi_i \geq 0$$
$$-\left[y_i \alpha_1 x_{1i} + y_i \alpha_2 x_{2i} + y_i \alpha_0 + \xi_i\right] \leq -1$$
$$-\begin{pmatrix} y_i x_{1i} & y_i x_{2i} & y_i & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_0 \\ \xi_i \end{pmatrix} \leq -1$$

Positivity constraint for slackness variables:

$$\xi_i \geq 0$$
$$-\begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_0 \\ \xi_i \end{pmatrix} \leq 0$$

All constraints in matrix formulation:

$$-\begin{pmatrix} y_1 x_{11} & y_1 x_{21} & y_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \ddots & \vdots \\ y_N x_{1N} & y_N x_{2N} & y_N & 0 & \dots & 1 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_0 \\ \xi_1 \\ \vdots \\ \xi_N \end{pmatrix} \leq -\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

The number of parameters in $\boldsymbol{x}'$ now also depends on the number of training samples. For each training sample, a slackness variable ξ_i is estimated.

(c) Visualize the decision boundary of the SVM with hard/soft margin constraints for different training sets.

Exercise 2 Find the largest rectangle that can be inscribed in the ellipse $x^2 + 2 \cdot y^2 = 4$. In this example, it is sufficient to consider rectangles which are centered at the origin.
See visualization in `ellipse_fitting.m`

(a) What is the area of such rectangles? Formulate the constrained optimization problem.

$$A(x,y) = 2x \cdot 2y; \max A(x,y) \text{ s. t. } x^2 + 2 \cdot y^2 = 4$$

(b) Solve using the Lagrange multiplier method.
Turn maximization into minimization in order to be able to use normal Lagrangian multipliers (add)

$$\min -A(x,y) \text{ s. t. } x^2 + 2 \cdot y^2 = 4$$

Setup objective function, calculate derivatives w.r.t the parameters

$$\begin{aligned}
\min D(x,y,\lambda) &= -4xy + \lambda\left(x^2 + 2y^2 - 4\right) \\
\frac{\partial D(x,y,\lambda)}{\partial x} &= -4y + 2\lambda x \stackrel{!}{=} 0 \\
\frac{\partial D(x,y,\lambda)}{\partial x} &= -4x + 4\lambda y \stackrel{!}{=} 0 \\
\frac{\partial D(x,y,\lambda)}{\partial \lambda} &= x^2 + 2y^2 - 4 \stackrel{!}{=} 0
\end{aligned}$$

Solve for λ, substitute the variables and solve using the constraint

$$\begin{aligned}
4y = 2\lambda x &\rightarrow \lambda = \frac{2y}{x} \\
-4x + \frac{8y^2}{x} = 0 &\rightarrow x^2 = 2y^2 \\
2y^2 + 2y^2 = 4 &\rightarrow y^2 = 1 \\
y = \pm 1 &\rightarrow x = \pm\sqrt{2}
\end{aligned}$$

The area of the largest rectangle is $4\sqrt{2}$.