ML: Algo and Theory

SS 18

Tutor:

1	2	3	4	$\sum$

Florence Lopez (3878792)

Jennifer Them (3837649)

Assignment 6

(Due 05. Juni 2018)

Exercise 1

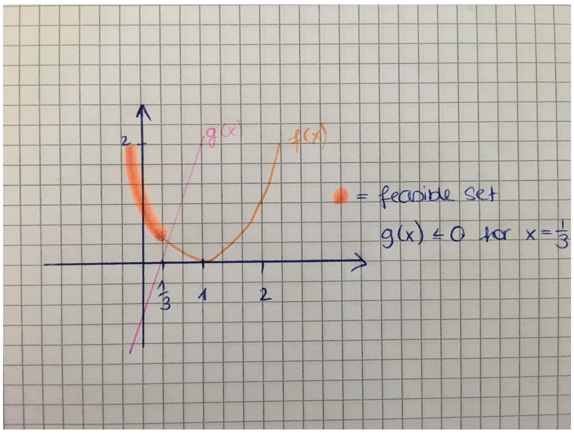
(a)

An optimization problem is convex if when the objective and the constraint function are convex. This is the case when its second derivative $f''(x) \geq 0$.

$f(x) = (x-1)^2 \Rightarrow f''(x) = 2 \geq 0$ therefore the function is convex.

The constraint function is ≤ 0 when $x \leq \frac{1}{3}$. $g(x)'' = 0$ and therefore convex.

You see on the graph that the objective function is a parabola and the constraint function is a line, which imposes that they are convex.



(b)

Lagrangian: $L(x, \alpha) = f(x) + \alpha g(x)$

$L = (x-1)^2 + \alpha(3x-1)$

Compute derivatives: $\nabla f(x) + \alpha \nabla g(x)$

$\nabla_x L = 2(x-1) + 3\alpha \stackrel{!}{=} 0$

$\nabla_\alpha L = 3x - 1 \stackrel{!}{=} 0$

I $2x + 3\alpha - 2$

II $3x - 1 = 0$

$\Rightarrow x = \frac{1}{3}$

$\Rightarrow \alpha = \frac{4}{9}$

The saddle point is at $x = \frac{1}{3}$.

(c)

see Code

(d)

λ is lagrangian multiplier $L(x, \lambda) = f(x) + \lambda g(x)$
Dual function: $d(x) = inf_x L(x, \lambda)$
$= inf_x f(x) + \lambda g(x)$
$= inf_x (x-1)^2 + \lambda(3x-1)$
$= inf_x (x^2 - 2x + 1 + 3x\lambda - \lambda)$
For fixed x, $L(x, \lambda)$ is linear in λ and therefore concave. The dual function as a pointwise infimum over concave functions is concave as well (see lecture slide 1245). .

Exercise 2

a.) The primal problem of the hard margin SVM is in general given by:
minimize$_{\omega \in \mathbb{R}^2, b \in \mathbb{R}} \frac{1}{2}||\omega||^2$ subject to: $Y_i(\langle \omega, X_i \rangle + b) \geq 1 \forall i$.
In our case this leads to the following primal problem:
minimize$_{\omega_1, \omega_2 \in \mathbb{R}, b \in \mathbb{R}} f(\omega, b) = \frac{1}{2}\omega_1^2 + \frac{1}{2}\omega_2^2$ subject to:

1. $g_1(\omega, b) = 0\omega_1 + 0\omega_2 + b + 1 \leq 0$
2. $g_2(\omega, b) = -1\omega_1 - 2\omega_2 - b + 1 \leq 0$
3. $g_3(\omega, b) = 1\omega_1 - 2\omega_2 - b + 1 \leq 0$

This leads to the following Lagrangian $L(\omega, b, \alpha)$:
$L(\omega, b, \alpha) = f(\omega, b) + \alpha_1 g_1(\omega, b) + \alpha_2 g_2(\omega, b) + \alpha_3 g_3(\omega, b)$
$L(\omega, b, \alpha) = \frac{1}{2}\omega_1^2 + \frac{1}{2}\omega_2^2 + \alpha_1 b + \alpha_1 - \alpha_2\omega_2 - 2\alpha_2\omega_2 - \alpha_2 b + \alpha_2 + \alpha_3\omega_1 - 2\alpha_3\omega_2 - \alpha_3 b + \alpha_3$
$L(\omega, b, \alpha) = \frac{1}{2}\omega_1^2 + \frac{1}{2}\omega_2^2 + (-\alpha_2 + \alpha_3) \cdot \omega_1 + (-2\alpha_2 - 2\alpha - 3) \cdot \omega_2 + (\alpha_1 - \alpha_2 - \alpha_3) \cdot b + \alpha_1 + \alpha_2 + \alpha_3$.

b.) The dimension of ω is 2, the dimension of α is 3.

The saddle point is computed by first deriving the Lagrangian:

$$
\begin{array}{rrr}
(1) & \dfrac{\delta}{\delta\omega_1}L(\omega,b,\alpha) = \omega_1 - \alpha_2 + \alpha_3 \overset{!}{=} 0 & (**) \\
(2) & \dfrac{\delta}{\delta\omega_2}L(\omega,b,\alpha) = \omega_2 - 2\alpha_2 - 2\alpha_3 \overset{!}{=} 0 & (**) \\
(3) & \dfrac{\delta}{\delta b}L(\omega,b,\alpha) = \alpha_1 - \alpha_2 - \alpha_3 \overset{!}{=} 0 & (*) \\
(4) & \dfrac{\delta}{\delta\alpha_1}L(\omega,b,\alpha) = b + 1 \overset{!}{=} 0 & \\
(5) & \dfrac{\delta}{\delta\alpha_2}L(\omega,b,\alpha) = -\omega_1 - 2\omega_2 - b + 1 \overset{!}{=} 0 & \\
(6) & \dfrac{\delta}{\delta\alpha_3}L(\omega,b,\alpha) = \omega_1 - 2\omega_2 - b + 1 \overset{!}{=} 0 & \\
\hline
(5) - (6) & \omega_1 = 0 & \\
(4) + (5) & -2\omega_2 + 2 = 0 \Rightarrow \omega_2 = 1 & \\
(4) & b = -1 & \\
2 \cdot (1) - (2) & 4\alpha_3 - 1 = 0 \Rightarrow \alpha_3 = \dfrac{1}{4} & \\
(1) & 0 - \alpha_2 + \dfrac{1}{4} = 0 \Rightarrow \alpha_2 = \dfrac{1}{4} & \\
(3) & \alpha_1 - \dfrac{1}{4} - \dfrac{1}{4} = 0 \Rightarrow \alpha_1 = \dfrac{1}{2} &
\end{array}
$$

c.) see code. The values of $\omega, b, y_i\alpha_i$ given from the code differ very slightly from the analytic solution above. But in principal they are the same. Differences may occur, because of numerical issues.

d.) A support vector is a vector, whose Lagrangian multipliers are non-zero. In the case of a hard-margin SVM the support vectors lie on the margin. In the case of a soft-margin SVM the support vectors can lie either on the margin, inside the margin or on the wrong side of the hyperplane.
We added the points $x_4 = (0,1), y_4 = 1, x_5 = (-1,-1), y_5 = -1$. This results in only two support vectors, namely the vectors x_0, x_4. The new values of alpha are then: $\alpha_1 = -1, \alpha_4 = 1$.

Exercise 3

a.) The primal problem is: minimize$_{x\in\mathbb{R}} f(x) = x^4 - 10x^2 + x$ subject to $g(x) = x^2 - 2x - 3 \leq 0$. The feasible set of this problem can be computed by trying out. One can see that for $x \in [-1,3], g(x)$ stays ≤ 0, since for $x = -2, g(x) = 5$ and for $x = 4, g(x) = 5$, too, which is both > 0.
To solve the primal, one has to compute: $\frac{\delta f(x)}{\delta x} = 4x^3 - 20x + 1 \overset{!}{=} 0$. We solved this using WolframAlpha and got the following results: $x \approx -2.26$ and $x \approx 2.21$. Since $x \approx 2.21$ is the only solution that is in the feasible set, it is also the solution of the primal problem: $f(2.21) \approx -22.77$.

b.) The Langrangian to this problem is:
$L(x,\alpha) = f(x) + \alpha g(x) = x^4 - 10x^2 + x + \alpha x^2 - 2\alpha x - 3\alpha$. The derivative with respect to x is: $\frac{\delta L(x,\alpha)}{\delta x} = x^3 - 20x + 1 + 2\alpha - 2\alpha = x^3 - 20x + 1$.
The dual function can be computed by: $h(\alpha) = \inf_x L(x,\alpha) = \inf_x x^4 - 10x^2 + x + \alpha x^2 - 2\alpha x - 3\alpha$. We know from the exercise that the dual function $h(\alpha)$ is optimal for $\alpha = 0.5$. Therefore we can calculate the optimal solution of the dual function:
$h(\alpha) = h(0.5) = \inf_x x^4 - 10x^2 + x + \frac{1}{2}x^2 - x - \frac{3}{2} = x^4 - \frac{19}{2}x^2 - \frac{3}{2}$. To get the value of the infimum, we have to set the derivative of $h(0.5)$ to zero:
$\frac{\delta h(0.5)}{\delta x} = 4x^3 - 19x \stackrel{!}{=} 0 \Leftrightarrow 4x^3 = 19x \Leftrightarrow x^2 = \frac{19}{4} \Leftrightarrow x = -\frac{\sqrt{19}}{2}$. This solution leads us to the result of the dual problem, which is: $h(0.5) = (-\frac{\sqrt{19}}{2})^4 - \frac{19}{2} \cdot (-\frac{\sqrt{19}}{2})^2 - \frac{3}{2} = -\frac{385}{16} = -24.0625$.

c.) In this case strong duality doesn't hold, since there is a duality gap between the solution of the primal and the solution of the dual problem: $f(2.21) - h(0.5) = -22.77 - (-24.0625) = 1.2925$