

Machine Learning SS2013

Ulrike von Luxburg
Assignment 05

Arne Schröder

Falk Oswald

Angel Bakardzhiev

May 13, 2013

Exercise 1

Task

Write the given linear program in the standard form by determining **A**, **b**, **c**.

Answer

Substitute x_3 with $x'_3 = -x_3$:

$$\begin{aligned} \text{Minimize} \quad & x_1 - 2x_2 - 4x'_3 \\ \text{subject to} \quad & x_1 - x_2 \leq -1 \\ & 3x_1 - 2x'_3 \leq -1 \\ & -2x_1 + 5x'_3 \leq -4 \\ & x_1, x_2, x'_3 \leq 0 \end{aligned}$$

Standard form:

Minimize $c^T x$
subject to $Ax \leq b$
and $x \leq 0$

with

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 3 & 0 & -2 \\ -2 & 0 & 5 \end{bmatrix}, b = \begin{bmatrix} -1 \\ -1 \\ -4 \end{bmatrix}, c = \begin{bmatrix} 1 \\ -2 \\ -4 \end{bmatrix}$$

Exercise 2

Task 1

Find the derivative of the Lagrangian with respect to the vector x and set it to zero ($\nabla_x L(x, \lambda_1, \lambda_2) = 0$). Solve it with respect to λ_2 .

Solution

$$\begin{aligned}\nabla_x L(x, \lambda_1, \lambda_2) &= c^T + \lambda_1^T A + \lambda_2^T = 0 \\ &\Leftrightarrow c + A\lambda_1 + \lambda_2 = 0 \\ &\Leftrightarrow \lambda_2 = -c - A\lambda_1\end{aligned}$$

Task 2

Replace λ_2 from part 1 in $L(x, \lambda_1, \lambda_2)$ and simplify it.

Solution

$$\begin{aligned}L(x, \lambda_1, \lambda_2) &= c^T x + \lambda_1^T (Ax - b) + \lambda_2^T x \\ &= c^T x + \lambda_1^T (Ax - b) + (-c - A\lambda_1)^T x \\ &= c^T x + \lambda_1^T (Ax - b) - c^T x - \lambda_1^T A^T x \\ &= \lambda_1^T ((A - A^T)x - b)\end{aligned}$$

Exercise 3

Task

Derive the dual problem of the hard margin SVM.

Solution

$$\begin{aligned}f_0(w, b) &= \frac{1}{2} \|w\|^2 \\ f_i(w, b) &= 1 - y_i(w^T x_i - b)\end{aligned}$$

$$\Rightarrow L(w, b, \lambda) = \frac{1}{2} \|w\|^2 + \sum_{i=1}^m \lambda_i \cdot (1 - y_i(w^T x_i - b))$$

We now define $g(\lambda) := \inf_{(w, b)} L(w, b, \lambda)$, which gives us the dual problem

Exercise 4

Task

For any $\gamma > 0$ the hyperplane defined by the solution (w, b) to **(I)** $\min_{w, b} \frac{1}{2} \|w\|^2$, subject to $y_i \cdot (\langle w, x_i \rangle - b) \geq 1$, is the same as the hyperplane defined by the solution (w', b') defined by **(II)** $\min_{w', b'} \frac{1}{2} \|w'\|^2$, subject to $y_i \cdot (\langle w', x_i \rangle - b') \geq \gamma$.

Answer

For $\gamma' > 0$, minimizing $\frac{1}{2} \|w\|^2$ is equivalent to minimizing $\gamma' \cdot \frac{1}{2} \|w\|^2$. This especially applies for $\gamma' = \frac{1}{2} \gamma^{-2}$.

Given the problem **(II)** we know that

$$\begin{aligned} y_i \cdot (\langle w', x_i \rangle - b) &\geq \gamma \\ \Leftrightarrow \gamma^{-1} \cdot y_i \cdot (\langle w', x_i \rangle - b') &\geq 1 \\ \Leftrightarrow y_i \cdot (\langle \gamma^{-1} \cdot w', x_i \rangle - \gamma^{-1} \cdot b') &\geq 1. \end{aligned}$$

We also can see that $\frac{1}{2} \|\gamma^{-1} \cdot w\|^2 = \frac{1}{2} \gamma^{-1} \cdot \frac{1}{2} \|w\|^2$. And using the assumption made at the beginning, we can conclude that **(II)** can be solved by the solution $(\gamma \cdot w, \gamma \cdot b)$.

Now we need to show that the two hyperplanes H, H' defined by (w, b) and $(\gamma \cdot w, \gamma \cdot b)$ coincide. Given a point $x \in H$ we know that

$$\begin{aligned} \langle w, x \rangle + b &= 0 \\ \Leftrightarrow \langle \gamma \cdot w, x \rangle + \gamma \cdot b &= \gamma \cdot (\langle w, x \rangle + b) = \gamma \cdot 0 = 0. \end{aligned}$$

This means that also $x \in H'$ and because x was arbitrary $H \subset H'$. And because the last implication is an equivalence, we know that $H = H'$.