Fundamentals of Linear Algebra and Optimization

Hard Margin Support Vector Machine; Version II

Jean Gallier and Jocelyn Quaintance

CIS Department
University of Pennsylvania
jean@cis.upenn.edu

May 7, 2020

Converting from Affine to Quadratic Functional

Since $\delta > 0$ (otherwise the data would not be separable into two disjoint sets), we can divide the affine constraints by δ to obtain

$$
\begin{aligned}
w'^{\top} u_i - b' &\geq 1 & \qquad i &= 1, \ldots, p \\
-w'^{\top} v_j + b' &\geq 1 & \qquad j &= 1, \ldots, q,
\end{aligned}
$$

except that now, w' is not necessarily a unit vector.

Converting from Affine to Quadratic Functional

To obtain the distances to the hyperplane H, we need to divide by $\|w'\|$ and then we have

$$
\begin{aligned}
\frac{w'^{\top} u_i - b'}{\|w'\|} &\geq \frac{1}{\|w'\|} & \qquad i = 1, \ldots, p \\
\frac{-w'^{\top} v_j + b'}{\|w'\|} &\geq \frac{1}{\|w'\|} & \qquad j = 1, \ldots, q,
\end{aligned}
$$

Converting from Affine to Quadratic Functional

To obtain the distances to the hyperplane H, we need to divide by $\|w'\|$ and then we have

$$\begin{aligned} \frac{w'^\top u_i - b'}{\|w'\|} &\geq \frac{1}{\|w'\|} & \qquad i = 1, \ldots, p \\ \frac{-w'^\top v_j + b'}{\|w'\|} &\geq \frac{1}{\|w'\|} & \qquad j = 1, \ldots, q, \end{aligned}$$

which means that the shortest distance from the data points to the hyperplane is $\delta = 1/\|w'\|$.

The Optimization Problem (SVM_{h2})

Therefore, we wish to *maximize* $1/\|w\|$, that is, to *minimize* $\|w\|$, so we obtain the following optimization Problem (SVM_{h2}):

The Optimization Problem (SVM_{h2})

Therefore, we wish to *maximize* $1/\|w\|$, that is, to *minimize* $\|w\|$, so we obtain the following optimization Problem (SVM_{h2}):

Hard margin SVM (SVM_{h2}):

$$
\begin{aligned}
&\text{minimize} \quad \frac{1}{2}\|w\|^2 \\
&\text{subject to} \\
&\qquad w^\top u_i - b \geq 1 \qquad i = 1, \ldots, p \\
&\qquad -w^\top v_j + b \geq 1 \qquad j = 1, \ldots, q.
\end{aligned}
$$

Solving (SVM_{h2}) *Via the KKT Conditions*

The objective function $J(w) = 1/2\,\|w\|^2$ is *convex*, so the last proposition of the KKT lesson applies and gives us a necessary and sufficient condition for having a minimum in terms of the KKT conditions.

Solving (SVM_{h2}) *Via the KKT Conditions*

The objective function $J(w) = 1/2\,\|w\|^2$ is *convex*, so the last proposition of the KKT lesson applies and gives us a necessary and sufficient condition for having a minimum in terms of the KKT conditions.

Observe that the trivial solution $w = 0$ is impossible, because the blue constraints would be

$$-b \geq 1,$$

Solving (SVM_{h2}) *Via the KKT Conditions*

The objective function $J(w) = 1/2\,\|w\|^2$ is *convex*, so the last proposition of the KKT lesson applies and gives us a necessary and sufficient condition for having a minimum in terms of the KKT conditions.

Observe that the trivial solution $w = 0$ is impossible, because the blue constraints would be

$$-b \geq 1,$$

that is $b \leq -1$, and the red constraints would be

$$b \geq 1,$$

but these are contradictory.

Solving (SVM_{h2}) *via the Lagrangian*

Our goal is to find w and b, and optionally, $\delta = 1/\|w\|$. In theory this can be done using the KKT conditions but in the present case it is much more efficient to solve the dual.