

Esercizio su SVM per classificazione binaria

Ricordiamo che, per un insieme linearmente separabile, la formulazione duale per una SVM binaria è la seguente:

$$\begin{aligned} \max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j \\ \text{soggetto a: } \forall i \in \{1, \dots, n\} : \alpha_i \geq 0 \text{ e } \sum_{i=1}^n y_i \alpha_i = 0. \end{aligned}$$

I valori ottimi delle α_i^* determinano l'iperpiano ottimo secondo la formula:

$$\vec{w}^* = \sum_{i=1}^n y_i \alpha_i^* \vec{x}_i$$

a meno del valore di b che viene determinato ricordando che per un qualsiasi vettore di supporto \vec{x}_s deve valere $y_s(\vec{w}^* \cdot \vec{x}_s + b^*) = 1$ e quindi considerando un esempio positivo ($y_s = 1$)

$$b^* = 1 - \vec{w}^* \cdot \vec{x}_s$$

Esercizio 1

Si consideri il seguente insieme di apprendimento con 3 esempi:

- 1: ($[1, 1]$, +1)
- 2: ($[3, 2]$, +1)
- 3: ($[2, 3]$, -1)

si calcoli il vettore \vec{w}^* e la soglia b^* corrispondenti all'iperpiano ottimo separatore, cioè la soluzione restituita da una Support Vector Machine (senza utilizzo di kernel).

Soluzione:

Si noti che da una ispezione visuale degli esempi si riconosce subito che l'insieme di apprendimento è linearmente separabile. Pertanto, per trovare la soluzione usiamo la corrispondente formulazione duale per la SVM. In particolare, dobbiamo massimizzare, rispetto alle variabili duali, la funzione:

$$L(\vec{\alpha}) = \sum_{i=1}^3 \alpha_i - \frac{1}{2} \sum_{i,j=1}^3 y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j$$

dove $y_1 = y_2 = +1$, $y_3 = -1$, $\vec{x}_1 \equiv [1, 1]$, $\vec{x}_2 \equiv [3, 2]$, $\vec{x}_3 \equiv [2, 3]$, rispettando i vincoli

$$\alpha_1 \geq 0, \quad \alpha_2 \geq 0, \quad \alpha_3 \geq 0$$

e

$$\sum_{i=1}^3 y_i \alpha_i = \alpha_1 + \alpha_2 - \alpha_3 = 0.$$

Sapendo che,

$$\begin{aligned} \vec{x}_1 \cdot \vec{x}_1 &= 2 \\ \vec{x}_2 \cdot \vec{x}_2 &= 13 \\ \vec{x}_3 \cdot \vec{x}_3 &= 13 \\ \vec{x}_1 \cdot \vec{x}_2 &= 5 \\ \vec{x}_1 \cdot \vec{x}_3 &= 5 \\ \vec{x}_2 \cdot \vec{x}_3 &= 12 \end{aligned}$$

la funzione lagrangiana può essere scritta come:

$$L(\vec{\alpha}) = \alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2} (2\alpha_1^2 + 13\alpha_2^2 + 13\alpha_3^2 + 10\alpha_1\alpha_2 - 10\alpha_1\alpha_3 - 24\alpha_2\alpha_3)$$

Poiché deve vale il vincolo $\alpha_3 = \alpha_1 + \alpha_2$, possiamo eliminare da $L()$ la variabile α_3 :

$$\begin{aligned} L(\alpha_1, \alpha_2) &= 2\alpha_1 + 2\alpha_2 - \frac{1}{2} (2(\alpha_1 + \alpha_2)^2 + 13\alpha_2^2 + 13\alpha_3^2 + \\ &\quad 10\alpha_2(\alpha_1 + \alpha_2) - 10\alpha_3(\alpha_1 + \alpha_2) - 24\alpha_2\alpha_3) \\ &= 2\alpha_1 + 2\alpha_2 - \frac{5}{2}\alpha_1^2 - \alpha_2^2 - \alpha_1\alpha_2 \end{aligned}$$

Il massimo per la funzione $L(\alpha_1, \alpha_2)$ si ottiene ponendo le derivate prime rispetto a α_1 e α_2 uguali a 0:

$$\begin{aligned}\frac{\partial L(\alpha_1, \alpha_2)}{\partial \alpha_1} &= 2 - 5\alpha_1 - \alpha_2 = 0 \\ \frac{\partial L(\alpha_1, \alpha_2)}{\partial \alpha_2} &= 2 - 2\alpha_2 - \alpha_1 = 0\end{aligned}$$

Tale sistema lineare ha soluzione $\alpha_1^* = \frac{2}{9} > 0$, $\alpha_2^* = \frac{8}{9} > 0$, e usando l'equazione $\alpha_3 = \alpha_1 + \alpha_2$ otteniamo $\alpha_3^* = \frac{10}{9} > 0$. Si noti che, in questo caso, tutti gli esempi sono di supporto perché nessuna variabile duale è nulla. Il vettore dei pesi ottimo è quindi

$$\vec{w}^* = \sum_{i=1}^3 y_i \alpha_i^* \vec{x}_i = \left[\frac{2}{3}, -\frac{4}{3} \right]$$

Infine, calcoliamo il valore ottimo di b utilizzando il vettore di supporto 1:

$$b^* = 1 - \vec{w}^* \cdot \vec{x}_1 = \frac{5}{3}$$