

Máquinas de Vectores de Soporte

Margen Suave

Luis Norberto Zúñiga Morales

21 de agosto de 2022

Contenido

- 1 Margen Suave
- 2 Formulación del Problema
- 3 Resolución del Problema

MVS Margen Suave

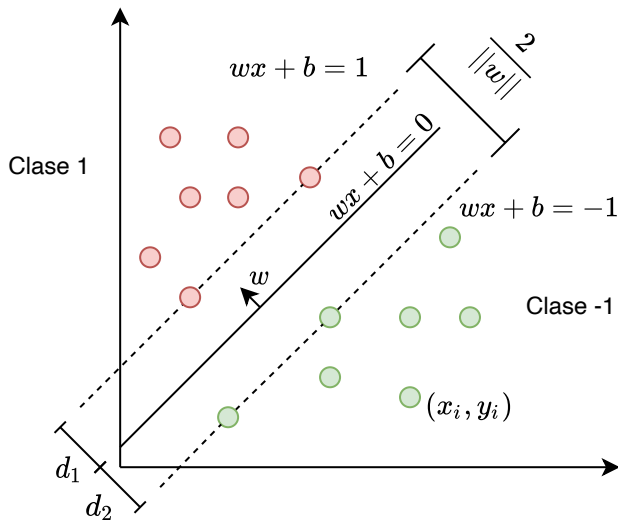


Figura: El caso linealmente separable es un sueño que rara vez se ve en la práctica.

MVS Margen Suave

Idea del Margen Suave

Margen Suave

La idea es relajar las limitantes de H_1 y H_2 para permitir puntos mal clasificados.

Lo anterior se logra introduciendo un variable de holgura positiva ξ_i , $i = 1, \dots, L$:

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 - \xi_i \qquad y_i = +1 \qquad (1)$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 + \xi_i \qquad y_i = -1 \qquad (2)$$

$$\xi_i \geq 0 \qquad \forall i \qquad (3)$$

MVS Margen Suave

Idea del Margen Suave

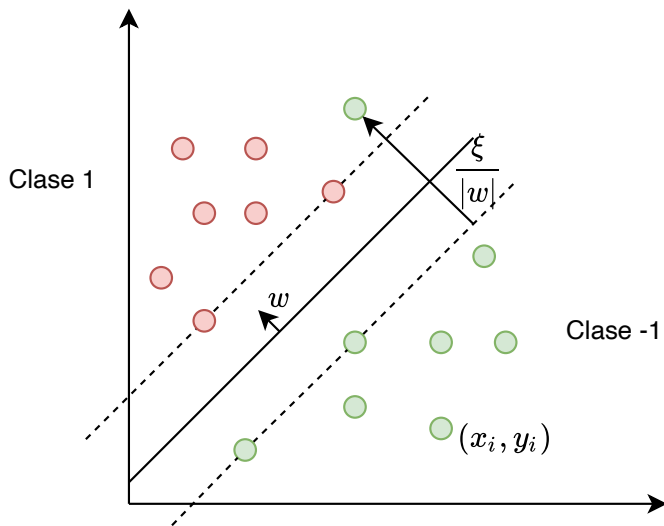


Figura: Idea del margen suave de una MVS.

MVS Margen Suave

Formulación del Problema

Ecuaciones con Margen Suave

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 - \xi_i$$

$$y_i = +1$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 + \xi_i$$

$$y_i = -1$$

$$\xi_i \geq 0$$

$$\forall i$$

Similar a la ecuación al caso lineal, las ecuaciones (1), (2) y (3) se pueden combinar en una sola.

Ejercicio

Combinar las ecuaciones (1), (2) y (3) en una sola.

MVS Margen Suave

Formulación del Problema

Ecuaciones con Margen Suave

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 - \xi_i$$

$$y_i = +1$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 + \xi_i$$

$$y_i = -1$$

$$\xi_i \geq 0$$

$$\forall i$$

Solución

$$y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \text{donde} \quad \xi_i \geq 0 \quad \forall i \quad (4)$$

MVS Margen Suave

Formulación del Problema

Para incluir esta idea en el modelo de la MVS, se adapta el problema de optimización del caso lineal de la siguiente manera:

$$\begin{aligned} \text{mín} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i \\ \text{sujeto a} \quad & y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i \end{aligned} \tag{5}$$

donde el parámetro C controla la razón de intercambio entre la penalización y el tamaño del margen.

MVS Margen Suave

Resumen hasta el momento

- Se definió el concepto de margen duro .

MVS Margen Suave

Resumen hasta el momento

- Se definió el concepto de **margen duro**.
- Se introdujo la idea del **margen suave** para considerar errores de clasificación.

MVS Margen Suave

Resumen hasta el momento

- Se definió el concepto de **margen duro**.
- Se introdujo la idea del **margen suave** para considerar errores de clasificación.
- Se formuló el **problema de optimización ajustado** a la idea del margen suave.



Plantear los multiplicadores de Lagrange.



Resolver las ecuaciones resultantes.



Considerar todas las restricciones del problema.

MVS Margen Suave

Resolución del Problema

Problema de Optimización

$$\begin{aligned} \text{mín} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i \\ \text{sujeto a} \quad & y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i \end{aligned}$$

MVS Margen Suave

Resolución del Problema

Problema de Optimización

$$\begin{aligned} \text{mín} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i \\ \text{sujeeto a} \quad & y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i \end{aligned}$$

Similar al caso lineal, es momento de formular el Lagrangiano:

$$\mathbb{L}_P = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i - \sum_{i=1}^L \alpha_i [y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1 + \xi_i] - \sum_{i=1}^L \mu_i \xi_i \quad (6)$$

MVS Margen Suave

Resolución del Problema

Consideraciones al momento de resolver el problema de optimización

De manera similar a la MVS Lineal, debemos encontrar \mathbf{w} , b y ahora ξ_i que minimicen

$$\begin{aligned} \text{mín} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i \\ \text{sujeto a} \quad & y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i \end{aligned}$$

y α y μ que maximicen

$$\mathbb{L}_P = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i - \sum_{i=1}^L \alpha_i [y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1 + \xi_i] - \sum_{i=1}^L \mu_i \xi_i$$

Ejercicio

Dado el Lagrangiano

$$\mathbb{L}_P = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i - \sum_{i=1}^L \alpha_i [y_i (\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1 + \xi_i] - \sum_{i=1}^L \mu_i \xi_i$$

encontrar

$$\frac{\partial \mathbb{L}_P}{\partial \mathbf{w}} = 0 \quad (?)$$

$$\frac{\partial \mathbb{L}_P}{\partial b} = 0 \quad (?)$$

$$\frac{\partial \mathbb{L}_P}{\partial \xi_i} = 0 \quad (?)$$

MVS Margen Suave

Resolución del Problema

Solución

Diferenciando con respecto a \mathbf{w} , b y ξ_i e igualando a cero:

$$\frac{\partial \mathbb{L}_P}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i \quad (7)$$

$$\frac{\partial \mathbb{L}_P}{\partial b} = \sum_{i=1}^L \alpha_i y_i = 0 \quad (8)$$

$$\frac{\partial \mathbb{L}_P}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \Rightarrow C = \alpha_i + \mu_i \quad (9)$$

Ejercicio

Sustituir

$$\frac{\partial \mathbb{L}_P}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$$

$$\frac{\partial \mathbb{L}_P}{\partial b} = \sum_{i=1}^L \alpha_i y_i = 0$$

$$\frac{\partial \mathbb{L}_P}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \Rightarrow C = \alpha_i + \mu_i$$

en la ecuación del Lagrangiano

$$\mathbb{L}_P = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L \xi_i - \sum_{i=1}^L \alpha_i [y_i (\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1 + \xi_i] - \sum_{i=1}^L \mu_i \xi_i$$

MVS Margen Suave

Resolución del Problema

Solución

Ya que $\frac{\partial \mathbb{L}_P}{\partial \mathbf{w}}$ y $\frac{\partial \mathbb{L}_P}{\partial b}$ son las mismas que en el caso lineal, obtenemos

$$\mathbb{L}_D = \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j, \quad \alpha_i \geq 0, \quad \sum_{i=1}^L \alpha_i y_i = 0 \quad (10)$$

$$= \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i H_{ij} \alpha_j, \quad H_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (11)$$

$$= \sum_{i=1}^L \alpha_i - \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{H} \boldsymbol{\alpha}, \quad \alpha_i \geq 0, \quad \sum_{i=1}^L \alpha_i y_i = 0 \quad (12)$$

Sin embargo, al considerar la ecuación

$$C = \alpha_i + \mu_i$$

y el hecho que $\mu_i \geq 0 \forall i$, se sigue que $\alpha > C$. Por lo tanto, se debe encontrar:

$$\begin{aligned} & \underset{\alpha}{\text{máx}} \quad \sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T \mathbf{H} \alpha \\ & \text{sujeto a} \quad 0 \leq \alpha_i \leq C \quad \forall_i \\ & \quad \quad \quad \sum_{i=1}^L \alpha_i y_i = 0 \end{aligned} \tag{13}$$

MVS Margen Suave

Resolución del Problema

Ejercicio

Para encontrar b , cualquier punto x_s que funcione como vector de soporte satisface la ecuación

$$y_s(\mathbf{x}_s \cdot \mathbf{w} + b) = 1 \quad (14)$$

Considerando que

$$\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$$

obtendríamos el mismo valor de b para el caso de margen duro. ¿Qué consideración debemos añadir el caso del margen suave?

MVS Margen Suave

Resolución del Problema

Solución

Para calcular b , se sigue la misma idea que en el caso lineal dado que

$$\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$$

son la misma para el caso de margen duro y suave. En este caso, los vectores de soporte son aquellos cuyos índices i cumplan la desigualdad $0 \leq \alpha_i \leq C$.

MVS Margen Suave

Implementación Práctica

1. Crear \mathbf{H} , donde $H_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$.

MVS Margen Suave

Implementación Práctica

1. Crear \mathbf{H} , donde $H_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$.
2. Elegir el valor del parámetro C , el cual permitirá penalizar clasificaciones erróneas.

MVS Margen Suave

Implementación Práctica

1. Crear \mathbf{H} , donde $H_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$.
2. Elegir el valor del parámetro C , el cual permitirá penalizar clasificaciones erróneas.
3. Encontrar las α_i que maximicen

$$\sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T \mathbf{H} \alpha$$

sujeto a las restricciones

$$0 \leq \alpha_i \leq C, \quad \sum_{i=1}^L \alpha_i y_i = 0.$$

mediante un programa para resolver problemas de optimización cuadrática.

4. Calcular $\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$.

MVS Margen Suave

Implementación Práctica

4. Calcular $\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$.
5. Determinar el conjunto de vectores de soporte S mediante la identificación de los índices i tales que $0 \leq \alpha_i \leq C$.

4. Calcular $\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$.
5. Determinar el conjunto de vectores de soporte S mediante la identificación de los índices i tales que $0 \leq \alpha_i \leq C$.
6. Calcular el valor de b mediante la ecuación

$$b = \frac{1}{N_S} \sum_{s \in S} (y_s - \sum_{k \in S} \alpha_k y_k \mathbf{x}_k \cdot \mathbf{x}_s).$$

MVS Margen Suave

Implementación Práctica

4. Calcular $\mathbf{w} = \sum_{i=1}^L \alpha_i y_i \mathbf{x}_i$.
5. Determinar el conjunto de vectores de soporte S mediante la identificación de los índices i tales que $0 \leq \alpha_i \leq C$.
6. Calcular el valor de b mediante la ecuación

$$b = \frac{1}{N_S} \sum_{s \in S} (y_s - \sum_{k \in S} \alpha_k y_k \mathbf{x}_k \cdot \mathbf{x}_s).$$

7. Cada elemento del conjunto de prueba \mathbf{x}_t se clasifica evaluando

$$y_t = \text{sgn}(\mathbf{w}^T \cdot \mathbf{x}_t + b).$$