



---

# TEMA 3. CASO NO LINEALMENTE SEPARABLE. SVM CONMARGEN FLEXIBLE

---

Hector Israel Heredia Cruz



Hector Israel Heredia Cruz

Tema 3

Caso no lineal separable. SVM con margen Flexible

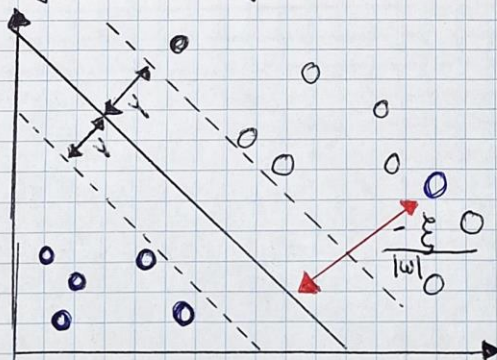
8 ISCI

Equipo 4

Scribe

El primer problema aparece cuando hay puntos sueltos de las dos clases en la zona del margen, o puntos de una clase en la zona de la otra. El caso mostrado en la Figura 4.10 no es lineal separable. El punto rojo que está entre verdes normalmente es un error o un outlier. La solución pasa por generar un margen flexible. El margen flexible admite errores en la separación a cambio de aumentar el margen. Estos errores están ponderados por la expresión  $\xi_i / |w|$ . El uso de márgenes flexibles normalmente aumenta la generalización y disminuye la varianza del modelo. El parámetro  $c$  del symtocy codifica la rigidez del margen. Cuanto es menor sea el valor de este parámetro más flexible es el margen. O dicho de otro modo, la flexibilidad del margen es inversamente proporcional al valor del parámetro  $c$ . Este parámetro se escoge normalmente como una potencia de diez ( $10^{\pm k}$ ).

Figura 4.10 Margen Flexible



El problema planteado, permitiendo que algunos datos puedan estar mal clasificados por el hiperplano obtenido, Aceptar algunos errores de clasificación equivalente a permitir violaciones en las restricciones del problema planteado. Para ello, se introducen unas nuevas variables  $\xi_i$  (variables de holgura).



Hector Israel Heredia Cruz  
Tema 3

815C11

Scribe

Caso no lineal separable, SVM con margen Flexible

Equipo 4

Por otro lado, la constante  $C$  en el problema planteado, que multiplica el término relativo al coste, actúa como término de compromiso entre el tamaño del margen (término de regularización o penalización de la complejidad) y error de la clasificación.

- Para valores pequeños de  $C$ , el primer término de la función a minimizar prevalece, por tanto se logrará un mayor margen a costa de obtener un mayor error de clasificación en el conjunto de datos de entrenamiento.

- Para valores grandes de  $C$ , el segundo término prevalece por lo que se tiene un menor margen pero también un menor error de clasificación en el conjunto de entrenamiento.

