

Pauta Control 2 - Máquinas de Aprendizaje Computacional (INF-578)

Semestre II 2017 - Viernes 24.11.17

Respuestas:

1. (15 puntos) Si el problema es linealmente separable, el problema tiene solución usando $C = \infty$. Sin embargo, aún en este caso puede tener sentido explorar otros valores de C ya que un margen (blando) puede traducirse en un mejor error de predicción (test). Una ventaja de considerar la solución dura ($C = \infty$) es que el modelo tendrá, muy probablemente, un número menor de support vectors, lo que lo hace más simple de almacenar y operar.
2. (15 puntos) La estrategia utilizada por un árbol de clasificación para enfrentar problemas linealmente no separables consiste en la división recursiva del espacio de entrada en regiones de manera que el patrón remanente en cada una de ellas sea suficientemente simple de modelar.
3. (25 puntos) De la ayuda, si $y^{(\ell)} \text{sign}(\mathbf{w}_t^T \mathbf{x}) < 0$, entonces por el paso 7 del algoritmo

$$\mathbf{w}_{new} = \mathbf{w}_{old} + \eta y^{(\ell)} \mathbf{x}^{(\ell)}.$$

Restando en ambos lados por \mathbf{w}_* tenemos que

$$\mathbf{w}_{new} - \mathbf{w}_* = \mathbf{w}_{old} - \mathbf{w}_* + \eta y^{(\ell)} \mathbf{x}^{(\ell)}.$$

Elevando al cuadrado en ambos lados

$$\|\mathbf{w}_{new} - \mathbf{w}_*\|^2 = \|\mathbf{w}_{old} - \mathbf{w}_*\|^2 + (\eta y^{(\ell)})^2 \|\mathbf{x}^{(\ell)}\|^2 + 2\eta(y^{(\ell)} \mathbf{w}_{old} \mathbf{x}^{(\ell)} - y^{(\ell)} \mathbf{w}_* \mathbf{x}^{(\ell)}).$$

Considerando que $y^{(\ell)2} = 1$ y que $y^{(\ell)} \mathbf{w}_{old} \mathbf{x}^{(\ell)} < 0$, ya que el dato estaba mal clasificado y $y^{(\ell)} \mathbf{w}_* \mathbf{x}^{(\ell)} > 0$, ya que \mathbf{w}_* clasifica bien todos los ejemplos

$$\|\mathbf{w}_{new} - \mathbf{w}_*\|^2 < \|\mathbf{w}_{old} - \mathbf{w}_*\|^2$$

Si $\eta^2 \|\mathbf{x}^{(\ell)}\|^2 > 2\eta$. Es decir, para $\eta > \frac{2}{\|\mathbf{x}^{(\ell)}\|^2}$.

Como cada iteración produce una disminución estricta de la f.o. y esta está claramente acotada inferiormente por 0, el algoritmo debe terminar.