The Elements of statistical learning:

Ex 2.1:

Suppose each of K-classes has an associated target t_K, which is a vector of all zeros, except a one in the Kth position. Show that classifying to the largest elements of $\hat{y}$ amounts to choosing the closest target, $\min_K \|t_K - \hat{y}\|$, if the elements of $\hat{y}$ sum to one.

Simplification:

t_1 serait $[1,0,0]$
t_2 serait $[0,1,0]$
t_3 serait $[0,0,1]$

$\hat{y} \Rightarrow [0.7, 0.2, 0.1]$

Answer:

We need to prove:

$$\underset{k}{\operatorname{argmax}}\ \hat{y}_k = \underset{K}{\operatorname{argmin}}\ \|t_K - \hat{y}\|^2 \qquad (1)$$

$$\operatorname{Argmin}_K \|\hat{y} - t_k\| = \operatorname{argmin} \|\hat{y} - t_k\|^2$$
$$= \operatorname{argmin}_K \sum_{i=1}^{K} (y_i - (t_k)_i)^2$$

constante par rapport à K $(a-b)^2$

$$= \operatorname{argmin}_K \sum_{i=1}^{K} \left((y_i)^2 - 2y_i (t_k)_i + (t_k)_i^2\right)$$

$$= \operatorname{argmin}_K \sum_{i=1}^{K} \left(-2y_i (t_k)_i + (t_k)_i^2\right)$$

Rappel:

1) la propriété $\sum (t_k)_i^2 = 1$ (propriété vérifié)

Exemple: $t_1 = [1,0,0]$ | Pour $t_1 = 1^2 + 0^2 + 0^2 = 1$

$t_2 = [0,1,0]$ | Pour t_2 ...

$t_3 = [0,0,1]$

2) la propriété $\sum y_i (t_k)_i = y_k$ (propriété vérifié)

Exemple: $\sum y_i (t_2)_i = y_1 \times 0 + y_2 \times 1 + y_3 \times 0 = y_2$

Suite:

$$= -2 \sum_{i=1}^{K} y_i (t_k)_i + \sum_{i=1}^{K} (t_k)_i^2 = -2y_k + 1$$

$$\operatorname{argmin}_K (y - t_k) = \operatorname{argmin}_k (-2y_k + 1)$$

Constante

$$= \operatorname{argmin}_k (-2y_k)$$
$$= \operatorname{argmax}_k (y_k)$$

minimiser un nombre neg revient à maximiser le nombre pos.