Classification

Friday, August 19, 2022 2:53 PM

Problem: We have a dataset of vectors $\vec{x}_i \in \mathbb{R}^N$ that belong to classes C_k for $k = 0, 1, \ldots, K-1$.

Solution: Linear classifiers.

$$\tilde{\vec{x}}_i = \begin{bmatrix} 1 \\ \vec{x}_i \end{bmatrix}$$

First, augment data points with a unity entry.

$\vec{a} \in \mathbb{R}^{N+1}$.

$\tilde{\vec{x}}_i \cdot \vec{a} = 0$ defines a line. So we can perform binary classification as

$$\tilde{\vec{x}}_i \cdot \vec{a} > \theta \longrightarrow \tilde{\vec{x}}_i \in \text{class } A$$
$$\tilde{\vec{x}}_i \cdot \vec{a} < \theta \longrightarrow \tilde{\vec{x}}_i \in \text{class } B$$

We can find boundary coefficients $\vec{a}$ using linear least squares.

$$T = \begin{bmatrix} 1 & t_1^{T(A)} \\ 1 & t_2^{T(A)} \\ 1 & \vdots \\ \vdots & t_1^{T(B)} \\ & \vdots \\ 1 & t_{M_B}^{T(B)} \end{bmatrix}$$

M_A : # class A data points

M_B : # class B data points

where $\vec{t}_m^{(A)}$ is a datapoint of class A.

$\vec{l} \in \mathbb{R}^{M_A + M_B}$ is the labels of the

$\vec{\ell} \in \mathbb{R}^{M_A+M_B}$ is the labels of the training data, 1 for class A and -1 for class B.

We seek $\underset{\vec{a}}{\operatorname{argmin}} \| T\vec{a} - \vec{\ell} \|^2$, which we know to be $\vec{a}^* = (T^T T)^{-1} T^T \vec{\ell}$.