Kernel Fisher Discriminant Analysis

Advanced Institute for Artificial Intelligence – AI2

https://advancedinstitute.ai

Introdução

A técnica de *Kernel Fisher Discriminant Analysis* - (KFDA) é uma **generalização não linear** para LDA. Novamente, faremos uso dos *kernels* para tornar LDA uma técnica de projeção não linear. Inicialmente, iremos abordar a versão para classificação com duas classes.

Seja, então, $\mathcal{X} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_m, y_m)\}$ um conjunto de dados tal que $\boldsymbol{x}_i \in \mathbb{R}^n$ e $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^{n'}$ uma função de mapeamento não linear tal que $n' > n$. Ademais, temos que $\mathcal{Y} = \{\omega_1, \omega_2\}$ de tal forma que $y_i \in \mathcal{Y}$, $i = 1, 2, \ldots, m$. No KFDA busca-se maximizar o seguinte critério:

$$J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{S}_B^\phi \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{S}_A^\phi \boldsymbol{w}}. \tag{1}$$

Note que o critério é o mesmo que LDA. O que muda, porém, é a maneira com a qual temos que calcular as matrizes de espalhamento interclasses $\boldsymbol{S}_B^\phi$ e intraclasses $\boldsymbol{S}_A^\phi$.

Neste caso, temos que:

$$\boldsymbol{S}_B^\phi = (\boldsymbol{\mu}_1^\phi - \boldsymbol{\mu}_2^\phi)(\boldsymbol{\mu}_1^\phi - \boldsymbol{\mu}_2^\phi)^T. \tag{2}$$

Além disso, temos que:

$$\boldsymbol{S}_A = \boldsymbol{S}_1 + \boldsymbol{S}_2, \tag{3}$$

em que

$$\boldsymbol{S}_i = \sum_{\boldsymbol{x}_j \in \omega_1} (\phi(\boldsymbol{x}_j) - \boldsymbol{\mu}_i^\phi)(\phi(\boldsymbol{x}_j) - \boldsymbol{\mu}_i^\phi)^T, \ i \in \{1, 2\}. \tag{4}$$

Temos que a média de cada classe é dada por:

$$\boldsymbol{\mu}_i^\phi = \frac{1}{m_i} \sum_{j=1}^{m_i} \phi(\boldsymbol{x}_j), \ i \in \{1, 2\}, \tag{5}$$

em que m_i denota a quantidade de elementos da classe ω_i.