Chapter IV
Orthogonal Discriminant Analysis Methods

ABSTRACT

In this chapter, we first give a brief introduction to Fisher linear discriminant, Foley-Sammon discriminant, orthogonal component discriminant, and application strategies for solving the SSS problems. We then present two novel orthogonal discriminant analysis methods, orthogonalized Fisher discriminant and Fisher discriminant with Schur decomposition. At last, we compare the performance of several main orthogonal discriminant analysis methods under various SSS strategies.

4.1 INTRODUCTION

4.1.1 Fisher Linear Discriminant

Fisher linear discriminant (FLD) (Duda, Hart, & Stork, 2001) operates by learning a discriminant matrix which maps a \( d \)-dimensional input space into an \( r \)-dimensional feature space by maximizing the multiple Fisher discriminant criterion.

Specifically, a Fisher discriminant matrix is an optimal solution of the following optimization model:

$$
\max_{W \in \mathbb{R}^{d \times r}} J_F(W) = \frac{W^T S_b W}{W^T S_w W} \quad (4.1)
$$
Here $W \in \mathbb{R}^{d \times r}$ is an arbitrary matrix, and $S_B$ and $S_W$ are the between- and within-class scatter matrices, and $|A|$ is the determinant of a square matrix $A$.

The between-class scatter matrix $S_B$ and the within-class scatter matrix $S_W$ are defined as follows,

$$S_B = \sum_{i=1}^{l} N_i (m_i - m)(m_i - m)^T,$$  \hspace{1cm} (4.2)

and

$$S_W = \sum_{i=1}^{l} S_i = \sum_{i=1}^{l} \sum_{x \in \omega_i} (x - m_i)(x - m_i)^T.$$  \hspace{1cm} (4.3)

Here $N_i$ and $m_i$ are respectively the number and the mean of samples from the $i$th class $\omega_i$, $m$ the mean of samples from all classes, and $l$ the number of classes.

It has been proved that if $S_W$ is nonsingular, the matrix composed of unit eigenvectors of the matrix $S_W^{-1}S_B$ corresponding to the first $r$ largest eigenvalues is an optimal solution of the optimization model defined in Eq. (4.1) (Wilks, 1962). The matrix $S_W^{-1}S_B$ is the Fisher discriminant matrix commonly used in Fisher linear discriminant.

Since the matrix $S_W^{-1}S_B$ is usually asymmetric, Fisher discriminant vectors, i.e. column vectors of the Fisher discriminant matrix are unnecessary orthogonal to each other.

### 4.1.2 Foley-Sammon Discriminant

Many researchers think that it is helpful to eliminate linear dependencies among discriminant vectors by making them orthogonal to each other. Foley-Sammon discriminant (FSD) is a feature extraction method that does this using the optimal discriminant vectors. Optimal discriminant vectors are derived from the multiple Fisher discriminant criterion which is subject to orthogonality constraints initially for binary classification tasks (Sammon, 1970; Foley & Sammon, 1975). In detail, the first discriminant vector of FSD is the first Fisher discriminant vector, i.e. a unit eigenvector of the matrix $S_W^{-1}S_B$ corresponding to the largest eigenvalue. After the first $k$ ($1 \leq k < r$) discriminant vectors $w_1, ..., w_k$ have been calculated, the $(k+1)$th discriminant vector $w_{k+1}$ of FSD is then one of the optimal solutions to the following optimization model

$$\max_{w'w=0, i=1, ..., k} J_0(w) = \frac{w^T S_B w}{w^T S_W w}.$$  \hspace{1cm} (4.4)
Related Content

Markov Chain for Multimodal Biometric Rank Fusion
www.igi-global.com/chapter/markov-chain-multimodal-biometric-rank/76163?camid=4v1a

Overview of Movement Analysis and Gait Features
Russell Best and Rezaul Begg (2006). Computational Intelligence for Movement Sciences: Neural Networks and Other Emerging Techniques (pp. 1-69).
www.igi-global.com/chapter/overview-movement-analysis-gait-features/6806?camid=4v1a
Establishing A-Priori Performance Guarantees for Robot Missions that Include Localization Software
www.igi-global.com/article/establishing-a-priori-performance-guarantees-for-robot-missions-that-include-localization-software/182506?camid=4v1a

Abortion and the Right to Life in the International Bibliography
www.igi-global.com/chapter/abortion-and-the-right-to-life-in-the-international-bibliography/125245?camid=4v1a