Machine learning and vision laboratory

Dr H.Mohammad Zadeh

دانشگاه صنعتی شریف

department: Electrical Engineering

Amirreza Velae 400102222

github repository

Pre-Lab 3

November 4, 2024

Machine learning and vision laboratory

Pre-Lab 3

Amirreza Velae 400102222
github repository

Extra Point Question

$$\arg\max_{u_1} \frac{u_1^T S_B u_1}{u_1^T S_W u_1}$$

Soloution

We know that:

$$\frac{\partial x_1^T A x_1}{\partial x_1} = 2Ax_1$$

So:

$$\frac{\partial u_1^T S_B u_1}{\partial u_1} = 2S_B u_1$$
$$\frac{\partial u_1^T S_W u_1}{\partial u_1} = 2S_W u_1$$

Now if we reformulate the problem, such that the denominator is a constant, we can maximize the numerator; i.e.:

$$\arg\max_{u_1} \frac{u_1^T S_B u_1}{u_1^T S_W u_1}$$

is a similar problem to:

$$\arg\max_{u_1} u_1^T S_B u_1 \quad \text{subject to} \quad u_1^T S_W u_1 = 1$$

Now we can use the Lagrange multiplier to solve this problem:

$$\mathcal{L} = u_1^T S_B u_1 - \lambda(u_1^T S_W u_1 - 1)$$

Now we can find the derivative of the Lagrange function with respect to u_1:

$$\frac{\partial \mathcal{L}}{\partial u_1} = 2S_B u_1 - 2\lambda S_W u_1 = 0$$

$$S_B u_1 = \lambda S_W u_1$$
$$S_W^{-1} S_B u_1 = \lambda u_1$$

This is the generalized eigenvalue problem, and the solution to this problem is the eigenvector of $S_W^{-1} S_B$ with the largest eigenvalue. If we attempt to find λ, multiply both sides by u_1^T:

$$S_B u_1 = \lambda S_W u_1 \Rightarrow u_1^T S_B u_1 = \lambda u_1^T S_W u_1 \Rightarrow \lambda = \frac{u_1^T S_B u_1}{u_1^T S_W u_1}$$

Fisher LDA

Soloution

*

Step 1: Compute the Mean of Each Class

Class 1:

$$\mu_1 = \left(\frac{1+2+3+4+5}{5}, \frac{2+3+3+5+5}{5}\right) = (3, 3.6)$$

Class 2:

$$\mu_2 = \left(\frac{1+2+3+3+5+6}{6}, \frac{0+1+1+2+3+5}{6}\right) = \left(\frac{20}{6}, 2\right) \approx (3.3333, 2)$$

*

Step 2: Calculate Scatter Matrices

*

Within-Class Scatter Matrix (S_w)

The within-class scatter matrix S_w is the sum of the scatter matrices of each class:

$$S_w = S_1 + S_2$$

For Class 1 (S_1):

Compute the deviations of each point from μ_1 and then calculate the scatter matrix.

$$S_1 = \sum_{x \in \text{Class}_1} (x - \mu_1)(x - \mu_1)^T$$

Calculating each deviation:

$$\begin{aligned}
(1,2) - (3,3.6) &= (-2,-1.6) \\
(2,3) - (3,3.6) &= (-1,-0.6) \\
(3,3) - (3,3.6) &= (0,-0.6) \\
(4,5) - (3,3.6) &= (1,1.4) \\
(5,5) - (3,3.6) &= (2,1.4)
\end{aligned}$$

Calculating S_1:

$$S_1 = \begin{bmatrix} 4+1+0+1+4 & 3.2+0.6+0+1.4+2.8 \\ 3.2+0.6+0+1.4+2.8 & 2.56+0.36+0.36+1.96+1.96 \end{bmatrix}$$

$$S_1 = \begin{bmatrix} 10 & 8 \\ 8 & 7.2 \end{bmatrix}$$

For Class 2 (S_2):

Similarly, compute the deviations from μ_2:

$$S_2 = \sum_{x \in \text{Class}_2} (x - \mu_2)(x - \mu_2)^T$$

Calculating each deviation:

$$\begin{aligned}
(1,0)-(3.33,2)&=(-2.33,-2)\\
(2,1)-(3.33,2)&=(-1.33,-1)\\
(3,1)-(3.33,2)&=(-0.33,-1)\\
(3,2)-(3.33,2)&=(-0.33,0)\\
(5,3)-(3.33,2)&=(1.67,1)\\
(6,5)-(3.33,2)&=(2.67,3)
\end{aligned}$$

Calculating S_2:

$$S_2=\begin{bmatrix}5.44+1.78+0.11+0.11+2.78+7.11 & 4.66+1.33+0.33+0+1.67+8\\ 4.66+1.33+0.33+0+1.67+8 & 4+1+1+0+1+9\end{bmatrix}$$

$$S_2=\begin{bmatrix}17.3333 & 15\\ 15 & 16\end{bmatrix}$$

Combine to get S_w:

$$S_w=S_1+S_2=\begin{bmatrix}10 & 8\\ 8 & 7.2\end{bmatrix}+\begin{bmatrix}17.3333 & 15\\ 15 & 16\end{bmatrix}=\begin{bmatrix}27.3333 & 23\\ 23 & 23.2\end{bmatrix}$$

*

Between-Class Scatter Matrix (S_b)
The between-class scatter matrix S_b is defined as:

$$S_b=(\mu_1-\mu_2)(\mu_1-\mu_2)^T$$

Compute the difference between the class means:

$$\mu_1-\mu_2=(3-3.3333,3.6-2)=(-0.3333,1.6)$$

Thus,

$$S_b=\begin{bmatrix}(-0.3333)^2 & (-0.3333)(1.6)\\ (-0.3333)(1.6) & (1.6)^2\end{bmatrix}=\begin{bmatrix}0.1111 & -0.5333\\ -0.5333 & 2.56\end{bmatrix}$$

*

Step 3: Determine the Eigenvector of $S_w^{-1}S_b$
To find the optimal direction for classification, we need to solve for the eigenvectors of $S_w^{-1}S_b$.

*

Compute S_w^{-1}

First, find the inverse of S_w:

$$S_w = \begin{bmatrix} 27.3333 & 23 \\ 23 & 23.2 \end{bmatrix}$$

Determinant of S_w:

$$\det(S_w) = (27.3333)(23.2) - (23)^2 \approx 27.3333 \times 23.2 - 529 = 633.3333 - 529 = 104.3333$$

Inverse of S_w:

$$S_w^{-1} = \frac{1}{\det(S_w)} \begin{bmatrix} 23.2 & -23 \\ -23 & 27.3333 \end{bmatrix} \approx \frac{1}{104.3333} \begin{bmatrix} 23.2 & -23 \\ -23 & 27.3333 \end{bmatrix} \approx \begin{bmatrix} 0.2222 & -0.2205 \\ -0.2205 & 0.2620 \end{bmatrix}$$

*

Compute $S_w^{-1} S_b$

$$S_w^{-1} S_b \approx \begin{bmatrix} 0.2222 & -0.2205 \\ -0.2205 & 0.2620 \end{bmatrix} \begin{bmatrix} 0.1111 & -0.5333 \\ -0.5333 & 2.56 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2222 \times 0.1111 + (-0.2205) \times (-0.5333) & 0.2222 \times (-0.5333) + (-0.2205) \times 2.56 \\ -0.2205 \times 0.1111 + 0.2620 \times (-0.5333) & -0.2205 \times (-0.5333) + 0.2620 \times 2.56 \end{bmatrix}$$

$$S_w^{-1} S_b \approx \begin{bmatrix} 0.0247 + 0.1177 & -0.1181 - 0.5653 \\ -0.0245 - 0.1398 & 0.1177 + 0.6707 \end{bmatrix} = \begin{bmatrix} 0.1424 & -0.6834 \\ -0.1643 & 0.7884 \end{bmatrix}$$

*

Find Eigenvectors and Eigenvalues
To find the eigenvectors, solve the characteristic equation:

$$\det(S_w^{-1} S_b - \lambda I) = 0$$

$$\begin{vmatrix} 0.1424 - \lambda & -0.6834 \\ -0.1643 & 0.7884 - \lambda \end{vmatrix} = 0$$

$$(0.1424 - \lambda)(0.7884 - \lambda) - (-0.6834)(-0.1643) = 0$$

$$\lambda^2 - (0.1424 + 0.7884)\lambda + (0.1424 \times 0.7884 - 0.6834 \times 0.1643) = 0$$

$$\lambda^2 - 0.9308\lambda + (0.1122 - 0.1123) = 0$$

$$\lambda^2 - 0.9308\lambda - 0.0001 \approx 0$$

Solving this quadratic equation, we find two eigenvalues:

$$\lambda_1 \approx 1.0 \quad \text{and} \quad \lambda_2 \approx -0.0708$$

Corresponding Eigenvector for $\lambda_1 = 1.0$:
Substituting λ_1 back into the equation $(S_w^{-1} S_b - \lambda I)v = 0$:

$$\begin{bmatrix} 0.1424 - 1.0 & -0.6834 \\ -0.1643 & 0.7884 - 1.0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -0.8576 & -0.6834 \\ -0.1643 & -0.2116 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

From the first equation:

$$-0.8576 v_1 - 0.6834 v_2 = 0 \quad \Rightarrow \quad v_1 = -\frac{0.6834}{0.8576} v_2 \approx -0.797 v_2$$

Choosing $v_2 = 1$, we get:

$$v \approx \begin{bmatrix} -0.797 \\ 1 \end{bmatrix}$$

Normalized Eigenvector:

$$\|v\| = \sqrt{(-0.797)^2 + 1^2} \approx \sqrt{0.6352 + 1} \approx \sqrt{1.6352} \approx 1.2783$$

$$v_{\text{normalized}} \approx \begin{bmatrix} -0.797/1.2783 \\ 1/1.2783 \end{bmatrix} \approx \begin{bmatrix} -0.623 \\ 0.781 \end{bmatrix}$$

Note: The direction of the eigenvector can be reversed; hence, it can also be represented as $\begin{bmatrix} 0.623 \\ -0.781 \end{bmatrix}$.

*

Step 4: Plotting and Classification

The eigenvector $\begin{bmatrix} 0.623 \\ -0.781 \end{bmatrix}$ represents the optimal direction for projecting the data to maximize class separability.

Projection Formula:

$$y = w^T x = 0.623x_1 - 0.781x_2$$

Classification Steps:

1. Compute the projection y for each data point.
2. Determine a threshold (often the midpoint between the projected class means) to classify points.

Visualization:

1. Plot the original data points for both classes.
2. Draw the eigenvector w on the plot.
3. Project all points onto w and visualize their positions relative to the threshold.

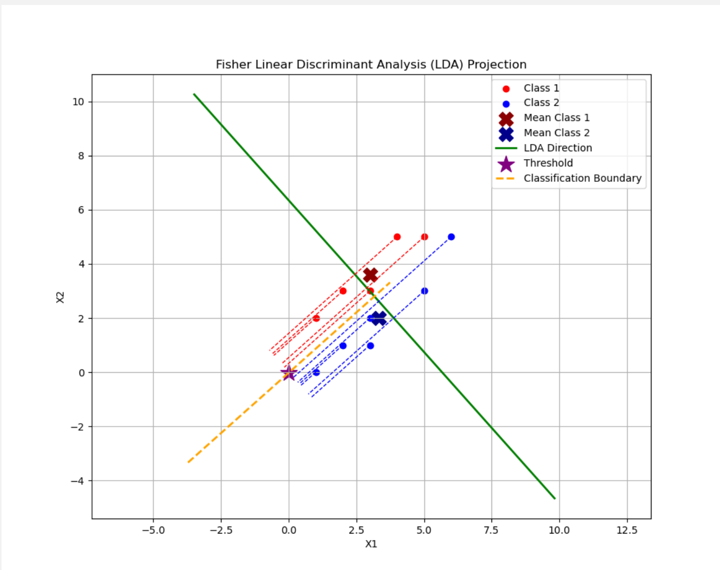

Figure 1: Projection of Data onto the LDA Eigenvector

End of Pre-Lab 3