Lab Three: Fisher And ROC

name: Huizhi Yu
number: ********
email: ********

1 Class Boundaries and Posterior Probabilities

In **Figure 1**, the left column shows the probability density contours under three different cases, with 200 random sample points overlaid for each of Class 1 and Class 2. The right column shows the corresponding posterior probability contours. In the probability density contour plots, contours of different colors represent the distribution of Class 1 and Class 2 at various probability density levels. The sample points are marked as circles, which helps to observe the positioning of each class's data points within their respective distribution regions.

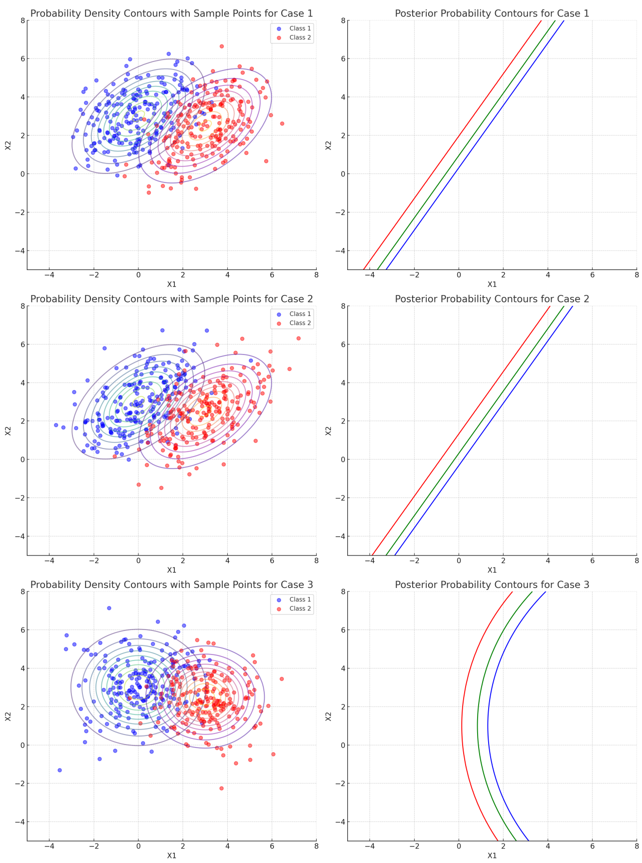

Figure 1: Probability Density and Posterior Probability Contours

In **Case 1**, the class means are $\mathbf{m}_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$ and $\mathbf{m}_2 = \begin{pmatrix} 3 \\ 2.5 \end{pmatrix}$, with a shared covariance matrix of $C_1 = C_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, and equal priors of $P_1 = P_2 = 0.5$. Since the covariance matrices are identical for both classes, the posterior probability contours appear as parallel lines, indicating a linearly separable decision boundary.

In **Case 2**, the class means and covariance matrices remain the same as in Case 1, but the priors are different, set as $P_1 = 0.7$ and $P_2 = 0.3$. In this scenario, while the probability density distribution of each class does not change, the posterior probability contours shift slightly due to the difference in prior probabilities, though the boundary remains linear.

In **Case 3**, the class means are the same as before, but the covariance matrices differ, with $C_1 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $C_2 = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}$, while the priors are equal again, $P_1 = P_2 = 0.5$. Due to the different covariance matrices, the probability density contours for Class 1 and Class 2 exhibit distinct ellipse sizes and orientations, and the posterior probability contours change from parallel lines to curves, indicating a non-linear decision boundary.

2 Fisher LDA and ROC Curve

2.1 Projection onto Fisher Discriminant Direction

To analyze the separation of the two classes, we project the data onto the Fisher discriminant direction. The data generated for this experiment consists of two classes that follow a two-dimensional Gaussian distribution. The mean vectors for the classes are $\mathbf{m}_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$ and $\mathbf{m}_2 = \begin{bmatrix} 3 \\ 2.5 \end{bmatrix}$, and they share a common covariance matrix defined as

$$C_1 = C_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

The Fisher discriminant direction is calculated as:

$$\mathbf{w}_F = (C_1 + C_2)^{-1}(\mathbf{m}_1 - \mathbf{m}_2),$$

This direction provides the optimal direction for maximizing the inter-class distance while minimizing the intra-class variance.

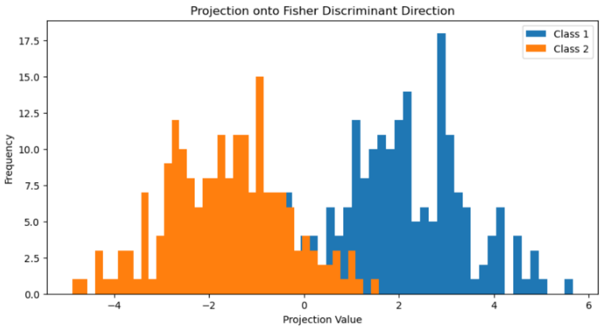

Figure 2: Projection onto Fisher Discriminant Direction

Figure 2 shows the distribution of Class 1 (blue) and Class 2 (orange) after projecting onto the Fisher discriminant direction. It shows that the two classes are relatively well-separated along this direction, with minimal overlap. This separation demonstrates the effectiveness of the Fisher discriminant direction in distinguishing between the two classes.

2.2 Receiver Operating Characteristic (ROC) Curve

Figure 3 illustrates the ROC curve based on the projection values along the Fisher discriminant direction. The blue curve captures the model's classification performance, while the red dashed line serves as a baseline for random guessing. With an AUC value of **0.98**, close to the ideal score of 1, this result reflects strong performance in differentiating between Class 1 and Class 2.

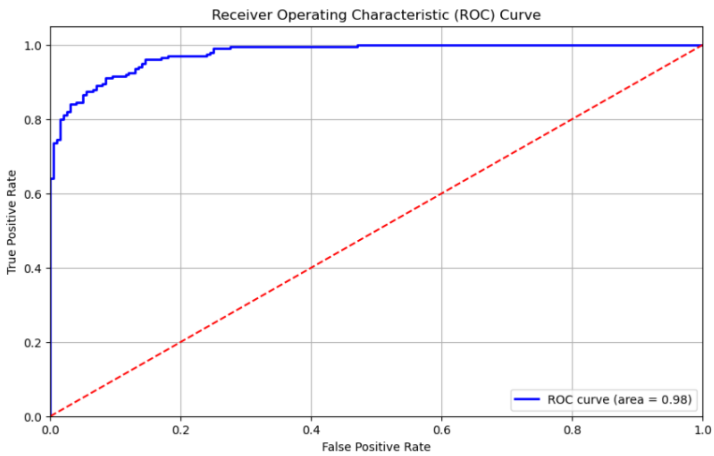

Figure 3: ROC Curve

2.3 ROC Curves for Different Projection Directions

Figure 4 presents a comparison of ROC curves and AUC values across three different projection directions. The blue curve corresponds to the Fisher discriminant direction, achieving an AUC of **0.98**; the purple dotted line represents the mean difference direction with an AUC of **0.95**; and the green dashed line reflects a random direction with an AUC of **0.90**. The Fisher discriminant direction, having the highest AUC, clearly delivers the strongest classification performance, followed by the mean difference direction, with the random direction showing the lowest performance. This indicates that the Fisher discriminant direction is particularly effective at enhancing class separation, leading to improved classification accuracy.

AUC essentially represents the likelihood that the classifier will rank a randomly chosen positive sample higher than a randomly chosen negative sample. An AUC close to 1 suggests the classifier is highly effective at distinguishing between classes. Therefore, the high AUC achieved by the Fisher discriminant direction reinforces its advantage in optimizing separation and enhancing overall classification performance.

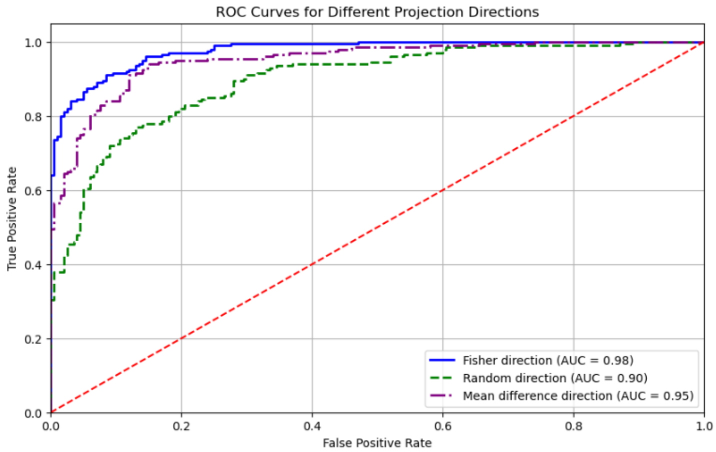

Figure 4: Comparison of ROC Curves for Different Projection Directions

2.4 ROC Curve Analysis

Using the threshold determined by the optimal balance point on the ROC curve, the model achieves a classification accuracy of 0.91.

Beyond classification accuracy, additional performance metrics offer a fuller picture of the classifier's effectiveness. Calculated at the optimal ROC threshold, these metrics include:

- **Precision**: Precision measures the accuracy of positive predictions, showing the percentage of true positives out of all samples predicted as positive.
- **Recall**: Recall reflects the classifier's ability to detect positive samples, calculated as the percentage of true positives out of all actual positive samples.
- **F1 Score**: The F1 Score, a balanced measure between Precision and Recall, is especially useful when both metrics are important to consider together.
- **Matthews Correlation Coefficient (MCC)**: MCC is a robust performance measure that takes into account true and false positives and negatives. An MCC close to 1 indicates a strong alignment between predicted and actual classifications, signifying a high level of reliability in the classifier.

The classifier based on the Fisher discriminant direction demonstrates strong classification performance at the optimal threshold. The classifier's **Precision** is 0.93, indicating that 93% of the samples predicted as positive are correctly classified. The **Recall** is also 0.93, meaning that 93% of actual positive samples are correctly identified. This high combination of precision and recall results in an **F1 Score** of 0.93, showing that the model achieves a good balance between precision and recall. Additionally, the **MCC** is 0.86, close to 1, indicating that the classifier performs well in distinguishing between positive and negative classes overall.

3 Mahalanobis Distance

3.1 The principle of Mahalanobis Distance

For sample to be classified, we calculate its Mahalanobis distance to the mean of each class. The formula for the Mahalanobis distance is:

$$D_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

where x is the sample to be classified, μ is the mean vector of the class, Σ is the covariance matrix of the class, and Σ^{-1} is the inverse of the covariance matrix. The Mahalanobis distance scales according to the covariance of each class, making the distance smaller in directions where the covariance is larger, and larger where the covariance is smaller.

For each sample, calculate its Mahalanobis distance to the mean of each class. Assign the sample to the class with the smaller Mahalanobis distance, thus completing the classification.

3.2 comparison between FDA and Mahalanobis distance

Comparing classifiers using FDA and Mahalanobis distance shows a classification accuracy improvement from 0.91 to **0.92**.

FDA optimizes a projection direction based on inter- and intra-class variance ratios, assuming similar covariance across classes. This approach is effective for classes with similar covariance structures but may overlook finer distinctions when the covariances differ.

Mahalanobis distance, by factoring in each class's covariance, adjusts distances according to the spread and orientation of the data. This refinement results in a more accurate decision boundary that takes into account variations in class shape and scale. The achieved accuracy of 0.92 suggests that Mahalanobis distance aligns better with the current data distribution, offering improved classification for samples near the boundary.