### Imprecise Gaussian Discriminant Classification

**Problem statement**

- **Setting:** Training data \(D = \{(x_i, y_i)\}_i \subseteq \mathbb{R} \times \mathbb{Y} \) where \(\mathbb{R} = \mathbb{R}^d\) and \(\mathbb{Y} = \{a, b, c\}\).
- **Motivation:** avoids mistakes performed by the precise models in hard-to-predict unlabeled instances by making cautious decisions (Figure 1a and 1b).

**Models**

- Defining the marginal distribution as \(D\).
- Assuming normality on \(\mathbb{R}\) stance \(x\).

**Utility-discount accuracy and \(u\)**

**Motivation**

- Training data
- Cautious decision: a extension of Gaussian Discriminant analysis aiming to quantify the lack of evidence of the component \(P_{X|Y=a}\).

**Our proposal**

- Avoids mistakes performed by the precise models in hard-to-predict unlabeled instances by making cautious decisions (Figure 1a and 1b)

**Motivation**

- Training data
- Cautious decision: an extension of Gaussian Discriminant analysis aiming to quantify the lack of evidence of the component \(P_{X|Y=a}\).

**Defining the marginal distribution as \(D\)**

**Utility**

### Decision Making

**Definition 2 (Precise ordering)**

Given a general loss function \(L_i\), a conditional probability distribution \(P_{X|Y=a}\) and a new unlabeled instance \(x\), \(m_0\) is preferred to \(m_1\) denoted by \(m_0 \succ m_1\) if \(m_0 > m_1\) under a) loss function, then \(m_0 \succ m_1 \iff \text{loss}(m_0(X=x)) < \text{loss}(m_1(X=x))\).

**Example 1**

Given a set of labels \(K = \{m_0, m_1\}\), a new unlabeled instance \(x\), and the probability estimates of the conditional distribution \(P_{X|Y=a}(x)\) at \(m_1\) and \(m_0\), making imprecise the components eigenvalues and eigenvectors of covariance matrix \(\Sigma\), i.e. \(\Sigma \trianglerighteq 0\).

**Estimating parameters by MLE on a subset \(S\) of \(D\)**

- \(\hat{\Sigma} = \hat{\Sigma}(S)\) or \(\Sigma\) has a diagonal structure of the covariance matrix, i.e. \(\Sigma = \Lambda\).

**Near-Ignorance on Gaussian Discriminant Analysis**

- Our proposal: avoids mistakes performed by the precise models in hard-to-predict unlabeled instances by making cautious decisions (Figure 1a and 1b)

**Motivation**

- Training data
- Cautious decision: an extension of Gaussian Discriminant analysis aiming to quantify the lack of evidence of the component \(P_{X|Y=a}\).

### Results

**Table 1: Data sets used in the experiments**

<table>
<thead>
<tr>
<th>Data set</th>
<th>LDA</th>
<th>ILDA</th>
<th>QDA</th>
<th>RQDA</th>
<th>Avg. Time (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>194</td>
<td>22.3</td>
<td>28.9</td>
<td>19.9</td>
<td>2.36</td>
</tr>
<tr>
<td>b</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>1.92</td>
</tr>
<tr>
<td>c</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>d</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>e</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>f</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>g</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>h</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
<tr>
<td>i</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>94.1</td>
<td>1.92</td>
</tr>
</tbody>
</table>

**Table 2: Average utility-discounted accuracies (k)**

- Increasing imprecision on the estimators has allowed us to be more cautious and to improve the prediction of classification.
- We performed since submission of ISPTA paper: Considering a diagonal structure of the covariance matrix, i.e. \(\Sigma = \Lambda\).
- Considering a set of marginal distribution \(P_{Y=a}\) instead of \(P_{Y}\) (i.e., release Assumption (A2))
- Considering the use of a generic loss function instead of zero-one loss function \(L_0\).
- What remains to be done:
  - Make imprecise the covariance matrix \(\Sigma\) by using the following set of prior distributions:
  \(\tilde{\mathbf{a}} = \sum_{i=1}^{m} \mathbf{a}_i \cdot p_i (\mathbf{ML}) \quad \mathbf{a}_i \in \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}, m = \rho\)
  - Making imprecise the components eigenvalues and eigenvectors of covariance matrix \(\Sigma\).

**Acknowledgments**

- This work was carried out within the framework of the Label MSIT, funded by the French Government, through the National Agency for Research (Reference ‘ANR-11-IDEX-0002’).

**References**