Original citation:

Permanent WRAP url:
http://wrap.warwick.ac.uk/61242

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A note on versions:
The version presented in WRAP is the published version or, version of record, and may be cited as it appears here. For more information, please contact the WRAP Team at: publications@warwick.ac.uk

http://wrap.warwick.ac.uk/
Parameter estimation of two dimensional two component Gaussian Mixtures

Nilantha Katugampala and Roland Wilson
Abstract

Multiresolution Gaussian Mixture Models (MGMM) can be used to represent image and video data in video annotation and retrieval. Preliminary experiments were carried out to estimate the model parameters for two-dimensional data. An iterative algorithm similar to Expectation-Maximisation (EM) is used to estimate the model parameters. The suitability of Akaike's Information Criterion (AIC) as a measure of model fit is also evaluated. AIC was successful for most of the synthetic data sets used in the experiments, however further work is required to develop a more consistent criterion for model fit.
1. Introduction

Multiresolution Gaussian Mixture Models (MGMM) has been proposed for modelling image and video data in video annotation and retrieval [1]. Preliminary experiments were carried out to estimate the model parameters for two-dimensional data, and to obtain a measure of model fit.

Two-dimensional Gaussian data of 100 points were generated with mean \( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \) and covariance \( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) [2]. Two different Gaussian populations of 100 points each were obtained by applying different transformations to the initial data. Transformed populations were merged to obtain a combined population of 200 points [3]. Then extracting the original populations from the combined population was attempted by using an iterative algorithm as follows:

1. Initialise \( \mu_1, \mu_2, \Sigma_1, \text{ and } \Sigma_2 \) of two-dimensional Gaussian populations.
2. Evaluate
   
   \[
   g_1(x_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_1|^{1/2}} \exp\left[\left(-\frac{1}{2}(x_i - \mu_1)\Sigma_1^{-1}(x_i - \mu_1)\right)\right]
   \]
   
   \[
   g_2(x_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_2|^{1/2}} \exp\left[\left(-\frac{1}{2}(x_i - \mu_2)\Sigma_2^{-1}(x_i - \mu_2)\right)\right]
   \]
   
   for all points, \( x_i, 1 \leq i \leq 200 \) of the combined population.

3. Classify \( x_i \); if \( g_2(x_i) > g_1(x_i) \) then \( x_i \) belongs to \( g_2 \) else \( x_i \) belongs to \( g_1 \).
4. Recompute \( \mu_1, \mu_2, \Sigma_1, \text{ and } \Sigma_2 \) from the classified \( x_i \), and go to step 2.

The above procedure was carried out until no further changes of \( x_i \) between \( g_1 \) and \( g_2 \) were observed, i.e. convergence was reached. Where \( \mu_1, \mu_2, \Sigma_1, \text{ and } \Sigma_2 \) represent the mean points and the covariances of the Gaussian populations.

The algorithm was tested with five different initial combined populations, each obtained by merging two transformed Gaussian populations. For each initial combined population two sets of experiments were carried out by differently initialising \( \mu_1, \mu_2, \Sigma_1, \text{ and } \Sigma_2 \).

**Experiment A**

100 random pairs of initial mean points, \( \mu_1 \) and \( \mu_2 \) are used. The initial mean points are uniformly distributed in a square area around the mean of the combined population. The length, \( a \) of the square is given by equation 1. The Initial covariance matrices, \( \Sigma_1, \text{ and } \Sigma_2 \) are set equal to the covariance matrix of the combined population, \( S_\gamma \). The set of experiments carried out with these initial conditions are referred to as Experiment A.
\[ a = 6|S_g|^{1/2} \]  

(1)

**Experiment B**

Separation was also carried out 100 times with random pairs of initial covariances, \( \Sigma_1 \), and \( \Sigma_2 \). The initial covariances are computed by scaling \( S_g \) by a randomly selected factor from the range 0.01, 0.02, 0.03, ..., 1.00. The initial mean points, \( \mu_1 \) and \( \mu_2 \) for this second set of experiments are set equal to the mean of the combined population. The set of experiments carried out with these initial conditions are referred to as Experiment B.

**Akaike’s Information Criterion (AIC)**

AIC [4] is computed as a measure of model fit to compare how well each separated solution performs against their combined counterparts. The difference of the AIC values by modelling the combined population using one Gaussian and a mixture of two Gaussians is estimated. Two versions are derived with different assumptions, when there are two Gaussian components. In Equation 2 (AIC_1) a data point is included in the likelihood estimation of a particular Gaussian component only if that data point contributed to the estimation of its parameters, i.e. belonged to that component in the classification step of the algorithm described above. In contrast equation 3 (AIC_2) evaluates the likelihood over all the data points for the complete Gaussian mixture model. In either case positive values favour two components.

\[
AIC_1 = \left\{ \frac{N}{2} \ln|S_g| + 5 \right\} - \left\{ \frac{N_1}{2} \ln|S_{g_1}| - N_1 \ln\left( \frac{N_1}{N} \right) + \frac{N_2}{2} \ln|S_{g_2}| - N_2 \ln\left( \frac{N_2}{N} \right) + 11 \right\} 
\]

(2)

\[
AIC_2 = \left\{ - \sum_{i=1}^{N} \ln g(x_i) + 5 \right\} - \left\{ - \sum_{i=1}^{N_1} \ln[\pi_1 g_1(x_i) + \pi_2 g_2(x_i)] + 11 \right\}
\]

(3)

**Further reading**

The application of MGMM in computer vision as a representation of image data and motion is described in [5]. The technique is illustrated by applying for image segmentation. The use of Expectation Maximisation (EM) Algorithm to estimate the parameters of finite mixtures is demonstrated in [6], and applied to image segmentation. AIC and Minimum Description Length (MDL) information criteria are employed to determine the number of segments in an image. An algorithm, which integrates the EM algorithm and the information criterion, Minimum Message Length (MML) has been, reported [7]. The inclusion of the information criterion within the parameter estimation process increases the ability of the algorithm to escape from local maxima. An introduction to model selection intended for nonspecialists is given in [8]. The introduction is extended to explanations of AIC and Bayesian Information
Criterion (BIC). Application of AIC and MDL for selecting the optimal number of components for a Gaussian mixture model is described in [9]. An intuitive derivation of AIC is given in [10]. A geometrical interpretation of AIC is given in [11].

The following sections describe the details of the experiments and their results. The final section draws some conclusions from the results of the experiments.
2. Similar mean points and covariances of different radii

The two initial populations are shown in blue and green in Figure 2.1. The mean of the blue population is depicted by ‘+’ and the mean of the green population by ‘*’. The dashed ellipses represent the covariances of the individual populations. The major axes are equal to the square roots of the corresponding eigenvalues. The solid ellipse and the black dot depict the covariance and the mean of the combined population respectively. Table 2.1 shows the statistics of the initial populations.

Table 2.1 Statistics of the initial populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.13641</td>
<td>-17.89735</td>
<td>-3.77466</td>
</tr>
<tr>
<td>Green</td>
<td>0.00183</td>
<td>0.691264</td>
<td>-0.06027</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.06729</td>
<td>9.299083</td>
<td>-1.8921</td>
</tr>
</tbody>
</table>

The two initial populations were merged and extracting the original populations using the algorithm described in section 1. Introduction was carried out with initial conditions of Experiments A and B.

There were 43 different solutions or convergences from both Experiments A and B. Experiment A gave rise to 36 solutions and Experiment B gave rise to 8 solutions, i.e. one solution came from both Experiments A and B. 4 pairs of initial mean points of Experiment A did not separate the combined population, i.e. insufficient number of points (less than three) in one component and the rest are in the other. The following section presents some of the selected converged solutions. The solutions are selected such that they resulted from more initial conditions for, either Experiment A or B, or the aggregate of both of them, if the solution resulted from Experiments A and B.
2.1 Examples of separated populations

*Example 1: resulted from Experiment A*

Figure 2.2 depicts a selected converged solution for the initial populations shown in Figure 2.1. Table 2.2 shows the statistics of the converged populations. 15 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 22.266667 and standard deviation 4.767482. Some of these initial pairs of mean points are depicted in Figure 2.3. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

**Table 2.2 Statistics of the separated populations**

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.108565</td>
<td>5.230056</td>
<td>-0.372275</td>
</tr>
<tr>
<td></td>
<td>1.997297</td>
<td>-0.372275</td>
<td>6.420510</td>
</tr>
<tr>
<td>Green</td>
<td>-0.336643</td>
<td>15.411473</td>
<td>-5.830666</td>
</tr>
<tr>
<td></td>
<td>-3.982810</td>
<td>-5.830666</td>
<td>9.783339</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.06729</td>
<td>9.299083</td>
<td>-1.8921</td>
</tr>
<tr>
<td></td>
<td>-0.36485</td>
<td>-1.8921</td>
<td>16.29498</td>
</tr>
</tbody>
</table>
The estimated AIC values are:

\[
\begin{align*}
AIC_1 &= -40.873256 \\
AIC_2 &= -6.097164
\end{align*}
\]

*Example 2: resulted from Experiment A*
Table 2.3  Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.845951</td>
<td>17.220419</td>
<td>-9.174362</td>
</tr>
<tr>
<td></td>
<td>-4.734508</td>
<td>-9.174362</td>
<td>11.459727</td>
</tr>
<tr>
<td>Green</td>
<td>0.243083</td>
<td>5.803619</td>
<td>-0.886202</td>
</tr>
<tr>
<td></td>
<td>1.376908</td>
<td>-0.886202</td>
<td>7.577729</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.06729</td>
<td>9.299083</td>
<td>-1.8921</td>
</tr>
<tr>
<td></td>
<td>-0.36485</td>
<td>-1.8921</td>
<td>16.29498</td>
</tr>
</tbody>
</table>

Figure 2.4 depicts another selected converged solution for the initial populations shown in Figure 2.1. Table 2.3 shows the statistics of the converged populations. 9 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 9.111111 and standard deviation 2.131481. These initial pairs of mean points are depicted in Figure 2.5. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

Figure 2.5  Examples of successful initial pairs of mean points

The estimated AIC values are:

\[
AIC_1 = -29.856318 \\
AIC_2 = -2.940483
\]
Example 3: resulted from Experiment B

Figure 2.6 depicts another selected converged solution for the initial populations shown in Figure 2.1. Table 2.4 shows the statistics of the converged populations. 25 pairs of initial covariances of Experiment B converged to this solution, with average number of iterations for convergence 6.080000 and standard deviation 0.890842. Some of these initial pairs of covariances are depicted in Figure 2.7. Each colour represents a pair.

![Figure 2.6 Separated populations](image)

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.022878</td>
<td>0.530273</td>
<td>-0.234742</td>
</tr>
<tr>
<td>Green</td>
<td>-0.113518</td>
<td>18.421613</td>
<td>-3.639661</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.06729</td>
<td>9.299083</td>
<td>-1.8921</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[ AIC_1 = 24.941031 \]
\[ AIC_2 = 48.409106 \]
Figure 2.7 Examples of successful initial pairs of covariances

Example 4: resulted from Experiment B

Figure 2.8 Separated populations
Table 2.5 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.006140</td>
<td>-0.088675</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.563759</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.177817</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>-0.146843</td>
<td>18.750181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.664030</td>
<td>-3.7949999</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.7949999</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30.503256</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>-0.067290</td>
<td>9.299083</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.36485</td>
<td>-1.8921</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.29498</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.8 depicts another selected converged solution for the initial populations shown in Figure 2.1. Table 2.5 shows the statistics of the converged populations. 19 pairs of initial covariances of Experiment B converged to this solution, with average number of iterations for convergence 3.526316 and standard deviation 0.499307. Some of these initial pairs of covariances are depicted in Figure 2.9. Each colour represents a pair.

The estimated AIC values are:

\[
\text{AIC}_1 = 24.978408 \\
\text{AIC}_2 = 48.466776
\]

Figure 2.9 Examples of successful initial pairs of covariances
Example 5: resulted from Experiments A and B

![Image of separated populations](image)

Figure 2.10 Separated populations

Table 2.6 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.011093</td>
<td>0.584931</td>
<td>-0.062541</td>
</tr>
<tr>
<td>Green</td>
<td>-0.135979</td>
<td>19.941135</td>
<td>-4.148803</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.06729</td>
<td>9.299083</td>
<td>-1.8921</td>
</tr>
</tbody>
</table>

Figure 2.10 depicts another selected converged solution for the initial populations shown in Figure 2.1. Table 2.6 shows the statistics of the converged populations. 10 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 17.500000 and standard deviation 2.655184. These initial pairs of mean points are depicted in Figure 2.11. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points. 36 pairs of initial covariances of Experiment B also converged to this solution, with average number of iterations for convergence 5.361111 and standard deviation 0.535038. Some of these initial pairs of covariances are depicted in Figure 2.12. Each colour represents a pair.
The estimated AIC values are:

$$AIC_1 = 26.022754$$
$$AIC_2 = 48.602174$$
The 4 pairs of initial mean points of Experiment A that did not separate the combined population in Figure 2.1 are depicted in Figure 2.13. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

Figure 2.13  Examples of unsuccessful initial pairs of mean points
3. Non overlapping populations

The two initial populations are shown in blue and green in Figure 3.1. The mean of the blue population is depicted by ‘+’ and the mean of the green population by ‘*’. The dashed ellipses represent the covariances of the individual populations. The major axes are equal to the square roots of the corresponding eigenvalues. The solid ellipse and the black dot depict the covariance and the mean of the combined population respectively. Table 3.1 shows the statistics of the initial populations.

Table 3.1 Statistics of the initial populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.289687 0.522580</td>
<td>-5.255479 -7.694632</td>
<td>100</td>
</tr>
<tr>
<td>Green</td>
<td>-12.319016 -12.247319</td>
<td>17.290340 6.193848</td>
<td>100</td>
</tr>
<tr>
<td>Combined</td>
<td>-6.304351 -5.862370</td>
<td>47.449096 37.652935</td>
<td>200</td>
</tr>
</tbody>
</table>

The two initial populations were merged and extracting the original populations using the algorithm described in section 1. Introduction was carried out with initial conditions of Experiments A and B.

There was only 1 solution or convergence from both Experiments A and B. In fact the solution is identical to the initial populations. 10 pairs of initial mean points of Experiment A and 8 pairs of initial covariances of Experiment B did not separate the combined population, i.e. insufficient number of points (less than three) in one component and the rest are in the other. The following section presents the converged solution.
3.1 The Separated population resulted from Experiments A and B

Figure 3.2 depicts the converged solution for the initial populations shown in Figure 3.1. Table 3.2 shows the statistics of the converged populations. 90 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 8.111111 and standard deviation 3.790176. Some of these initial pairs of mean points are depicted in Figure 3.3. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points. 92 pairs of initial covariances of Experiment B also converged to this solution, with average number of iterations for convergence 8.576087 and standard deviation 1.854687. Some of these initial pairs of covariances are depicted in Figure 3.4. Each colour represents a pair.

Table 3.2 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.289687</td>
<td>5.255479</td>
<td>-7.694632</td>
</tr>
<tr>
<td></td>
<td>0.522580</td>
<td>-7.694632</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>-12.319016</td>
<td>17.290340</td>
<td>6.193848</td>
</tr>
<tr>
<td></td>
<td>-12.247319</td>
<td>6.193848</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>-6.304351</td>
<td>47.449096</td>
<td>37.652935</td>
</tr>
<tr>
<td></td>
<td>-5.862370</td>
<td>37.652935</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.3 Examples of successful initial pairs of mean points

Figure 3.4 Examples of successful initial pairs of covariances

The estimated AIC values are:

\[ AIC_1 = 243.378149 \]
\[ AIC_2 = 243.388687 \]
The 10 pairs of initial mean points of Experiment A that did not separate the combined population in Figure 2.1 are depicted in Figure 3.5. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

Figure 3.5 Examples of unsuccessful initial pairs of mean points

The 8 pairs of initial covariances of Experiment B that did not separate the combined population in Figure 2.1 are depicted in Figure 3.6. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

Figure 3.6 Examples of unsuccessful initial pairs of covariances
4. Overlapping populations

The two initial populations are shown in blue and green in Figure 4.1. The mean of the blue population is depicted by ‘+’ and the mean of the green population by ‘**’. The dashed ellipses represent the covariances of the individual populations. The major axes are equal to the square roots of the corresponding eigenvalues. The solid ellipse and the black dot depict the covariance and the mean of the combined population respectively. Table 4.1 shows the statistics of the initial populations.

Table 4.1 Statistics of the initial populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.130864</td>
<td>5.349469</td>
<td>-7.930679</td>
</tr>
<tr>
<td>Green</td>
<td>-7.992005</td>
<td>19.316026</td>
<td>8.396025</td>
</tr>
<tr>
<td>Combined</td>
<td>-3.930571</td>
<td>28.827996</td>
<td>19.583653</td>
</tr>
</tbody>
</table>

The two initial populations were merged and extracting the original populations using the algorithm described in section 1. Introduction was carried out with initial conditions of Experiments A and B.

There were 4 different solutions or convergences from both Experiments A and B. Experiment A gave rise to 4 solutions and Experiment B gave rise to 1 solution, i.e. one solution came from both Experiments A and B. 11 pairs of initial mean points of Experiment A and 8 pairs of initial covariances of Experiment B did not separate the combined population, i.e. insufficient number of points (less than three) in one component and the rest are in the other. The following section presents the converged solutions.
4.1 The separated populations

*Example 1: resulted from Experiment A*

![Figure 4.2 Separated populations](image)

Figure 4.2 depicts a converged solution for the initial populations shown in Figure 4.1. Table 4.2 shows the statistics of the converged populations. 4 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 2.000000 and standard deviation 0.000000. These initial pairs of mean points are depicted in Figure 4.3. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>5.236881</td>
<td>0.188004</td>
<td>-0.110598</td>
</tr>
<tr>
<td></td>
<td>-10.222348</td>
<td>-0.110598</td>
<td>0.382498</td>
</tr>
<tr>
<td>Green</td>
<td>-4.117661</td>
<td>27.662336</td>
<td>20.938038</td>
</tr>
<tr>
<td></td>
<td>-5.131446</td>
<td>20.938038</td>
<td>32.829553</td>
</tr>
<tr>
<td>Combined</td>
<td>-3.930571</td>
<td>28.827996</td>
<td>19.583653</td>
</tr>
<tr>
<td></td>
<td>-5.233264</td>
<td>19.583653</td>
<td>32.688591</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[
\text{AIC}_1 = 9.700011 \\
\text{AIC}_2 = 9.723727
\]
**Example 2: resulted from Experiment A**

Figure 4.4 depicts another converged solution for the initial populations shown in Figure 4.1. Table 4.3 shows the statistics of the converged populations. 3 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 8.333333 and standard deviation 0.942809. These initial pairs of mean points are depicted in Figure 4.5. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

The estimated AIC values are:

\[
AIC_1 = -12.832056 \\
AIC_2 = 10.115188
\]
Table 4.3 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-6.278740</td>
<td>25.652553</td>
<td>30.724216</td>
</tr>
<tr>
<td></td>
<td>-4.993368</td>
<td>30.724216</td>
<td>47.086077</td>
</tr>
<tr>
<td>Green</td>
<td>-0.016954</td>
<td>9.614170</td>
<td>3.519682</td>
</tr>
<tr>
<td></td>
<td>-5.633091</td>
<td>3.519682</td>
<td>8.437002</td>
</tr>
<tr>
<td>Combined</td>
<td>-3.930571</td>
<td>28.827996</td>
<td>19.583653</td>
</tr>
<tr>
<td></td>
<td>-5.233264</td>
<td>19.583653</td>
<td>32.688591</td>
</tr>
</tbody>
</table>

Figure 4.5 Examples of successful initial pairs of mean points

Example 3: resulted from Experiment A

Figure 4.6 Separated populations

Figure 4.6 depicts another converged solution for the initial populations shown in Figure 4.1. Table 4.4 shows the statistics of the converged populations. 1 pair of
initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 2.000000. This initial pair of mean points is depicted in Figure 4.7.

**Table 4.4 Statistics of the separated populations**

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>5.354657</td>
<td>0.195188</td>
<td>0.016894</td>
</tr>
<tr>
<td></td>
<td>-10.571228</td>
<td>0.016894</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>-4.071970</td>
<td>27.931109</td>
<td>20.647903</td>
</tr>
<tr>
<td></td>
<td>-5.151976</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>-3.930571</td>
<td>28.827996</td>
<td>19.583653</td>
</tr>
<tr>
<td></td>
<td>-5.233264</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[
AIC_1 = 9.403133 \\
AIC_2 = 9.407357
\]

![Figure 4.7 An Example of a successful initial pair of mean points](image-url)
Example 4: resulted from Experiments A and B

Figure 4.8 Separated populations

Table 4.5 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.182389</td>
<td>5.380257</td>
<td>-7.985082</td>
</tr>
<tr>
<td>Green</td>
<td>-8.211406</td>
<td>17.300327</td>
<td>7.366931</td>
</tr>
<tr>
<td>Combined</td>
<td>-3.930571</td>
<td>28.827996</td>
<td>19.583653</td>
</tr>
</tbody>
</table>

Figure 4.8 depicts another converged solution for the initial populations shown in Figure 4.1. Table 4.5 shows the statistics of the converged populations. 81 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 11.259259 and standard deviation 6.653279. Some of these initial pairs of mean points are depicted in Figure 4.9. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points. 92 pairs of initial covariances of Experiment B also converged to this solution, with average number of iterations for convergence 9.195652 and standard deviation 0.679500. Some of these initial pairs of covariances are depicted in Figure 4.10. Each colour represents a pair.

The estimated AIC values are:

\[
\text{AIC}_1 = 217.402522 \\
\text{AIC}_2 = 218.516788
\]
Figure 4.9 Examples of successful initial pairs of mean points

Figure 4.10 Examples of successful initial pairs of covariances
The 11 pairs of initial mean points of Experiment A that did not separate the combined population in Figure 4.1 are depicted in Figure 4.11. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

![Figure 4.11 Examples of unsuccessful initial pairs of mean points](image)

The 8 pairs of initial covariances of Experiment B that did no separate the combined population in Figure 4.1 are depicted in Figure 4.12.

![Figure 4.12 Example of unsuccessful initial pairs of covariances](image)
5. Similar mean points and covariances of different directions

The two initial populations are shown in blue and green in Figure 5.1. The mean of the blue population is depicted by ‘+’ and the mean of the green population by ‘*’. The dashed ellipses represent the covariances of the individual populations. The major axes are equal to the square roots of the corresponding eigenvalues. The solid ellipse and the black dot depict the covariance and the mean of the combined population respectively. Table 5.1 shows the statistics of the initial populations.

Table 5.1 Statistics of the initial populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.029255</td>
<td>6.619194</td>
<td>-8.915960</td>
</tr>
<tr>
<td></td>
<td>-0.130021</td>
<td>-8.915960</td>
<td>15.605562</td>
</tr>
<tr>
<td>Green</td>
<td>-0.182076</td>
<td>25.260471</td>
<td>9.415431</td>
</tr>
<tr>
<td></td>
<td>-0.175361</td>
<td>9.415431</td>
<td>4.310522</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665</td>
<td>15.945671</td>
<td>0.251468</td>
</tr>
<tr>
<td></td>
<td>-0.152691</td>
<td>0.251468</td>
<td>9.958556</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.251468</td>
<td>200</td>
</tr>
</tbody>
</table>

The two initial populations were merged and extracting the original populations using the algorithm described in section 1. Introduction was carried out with initial conditions of Experiments A and B.

There were 20 different solutions or convergences from both Experiments A and B. Experiment A gave rise to 19 solutions and Experiment B gave rise to 2 solutions, i.e. one solution came from both Experiments A and B. 2 pairs of initial mean points of Experiment A did not separate, i.e. insufficient number of points (less than three) in one component and the rest are in the other. The following section presents some of the selected converged solutions. The solutions are selected such that they resulted from more initial conditions for, either Experiment A or B, or the aggregate of both of them, if the solution resulted from Experiments A and B.
5.1 Examples of separated populations

*Example 1: resulted from Experiment A*

Figure 5.2 depicts a selected converged solution for the initial populations shown in Figure 5.1. Table 5.2 shows the statistics of the converged populations. 19 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 16.157895 and standard deviation 1.724839. Some of these initial pairs of mean points are depicted in Figure 5.3. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

Table 5.2 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.262281</td>
<td>25.529872</td>
<td>9.592733</td>
</tr>
<tr>
<td>Blue</td>
<td>-0.102120</td>
<td>9.592733</td>
<td>4.250640</td>
</tr>
<tr>
<td>Green</td>
<td>0.047849</td>
<td>6.503647</td>
<td>-8.889448</td>
</tr>
<tr>
<td>Green</td>
<td>-0.202260</td>
<td>-8.889448</td>
<td>15.548479</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665</td>
<td>15.945671</td>
<td>0.251468</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.152691</td>
<td>0.251468</td>
<td>0.958556</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[
AIC_1 = 67.017429 \\
AIC_2 = 92.308354
\]
**Figure 5.3** Examples of successful initial pairs of mean points

**Example 2: resulted from Experiment A**

**Figure 5.4** Separated populations

Figure 5.4 depicts another selected converged solution for the initial populations shown in Figure 5.1. Table 5.3 shows the statistics of the converged populations. 13 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 9.769231 and standard deviation 2.606319. Some of these initial pairs of mean points are depicted in Figure 5.5. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.
Table 5.3 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>0.767138, 1.876953</td>
<td>13.033658, -1.151955</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>-1.151955</td>
<td>4.235560</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>-1.361650, -3.073399</td>
<td>17.462402, -3.946541</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>-3.946541</td>
<td>3.735555</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665, -0.152691</td>
<td>15.945671, 0.251468</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0.251468</td>
<td>9.958556</td>
<td></td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[
\text{AIC}_1 = -29.976402 \\
\text{AIC}_2 = -8.092659
\]

Figure 5.5 Examples of successful initial pairs of mean points
**Example 3: resulted from Experiment A**

Figure 5.6 depicts another selected converged solution for the initial populations shown in Figure 5.1. Table 5.4 shows the statistics of the converged populations. 11 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 11.636364 and standard deviation 2.993105. These initial pairs of mean points are depicted in Figure 5.7. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>2.505050</td>
<td>7.502316</td>
<td>2.186188</td>
</tr>
<tr>
<td></td>
<td>-0.884802</td>
<td>2.186188</td>
<td>8.855087</td>
</tr>
<tr>
<td>Green</td>
<td>-3.566380</td>
<td>6.126537</td>
<td>3.578999</td>
</tr>
<tr>
<td></td>
<td>0.817782</td>
<td>3.578999</td>
<td>9.768981</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665</td>
<td>15.945671</td>
<td>0.251468</td>
</tr>
<tr>
<td></td>
<td>-0.152691</td>
<td>0.251468</td>
<td>9.958556</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[ \text{AIC}_1 = -36.462414 \]
\[ \text{AIC}_2 = -12.850395 \]
Figure 5.7 Examples of successful initial pairs of mean points

**Example 4: resulted from Experiment B**

Figure 5.8 depicts another selected converged solution for the initial populations shown in Figure 5.1. Table 5.5 shows the statistics of the converged populations. 13 pairs of initial covariances of Experiment B converged to this solution, with average number of iterations for convergence 3.307692 and standard deviation 0.461538. Some of these initial pairs of covariances are depicted in Figure 5.9. Each colour represents a pair.
Table 5.5 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.101504</td>
<td>2.085454</td>
<td>-0.271056</td>
</tr>
<tr>
<td></td>
<td>-0.285712</td>
<td>1.333963</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>-0.108440</td>
<td>25.185796</td>
<td>0.600432</td>
</tr>
<tr>
<td></td>
<td>-0.064010</td>
<td>0.600432</td>
<td>15.688624</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665</td>
<td>15.945671</td>
<td>0.251468</td>
</tr>
<tr>
<td></td>
<td>-0.152691</td>
<td>9.958556</td>
<td></td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[ \text{AIC}_1 = -32.433444 \]
\[ \text{AIC}_2 = -0.615810 \]

Figure 5.9 Examples of successful initial pairs of covariances
Example 5: resulted from Experiments A and B

Figure 5.10 depicts another selected converged solution for the initial populations shown in Figure 5.1. Table 5.6 shows the statistics of the converged populations. 17 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 21.529412 and standard deviation 4.827832. Some of these initial pairs of mean points are depicted in Figure 5.11. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points. 87 pairs of initial covariances of Experiment B also converged to this solution, with average number of iterations for convergence 25.241379 and standard deviation 1.694925. Some of these initial pairs of covariances are depicted in Figure 5.12. Each colour represents a pair.

Table 5.6 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-0.312857</td>
<td>26.943785 10.072134</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>-0.115919</td>
<td>10.072134 4.463628</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>0.078070</td>
<td>6.120799  -8.444677</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>-0.185300</td>
<td>14.829154 8.444677</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>-0.105665</td>
<td>15.945671 0.251468</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>-0.152691</td>
<td>0.251468 9.958556</td>
<td></td>
</tr>
</tbody>
</table>
The estimated AIC values are:

\[ \text{AIC}_1 = 67.207818 \]
\[ \text{AIC}_2 = 92.509745 \]
The 2 pairs of initial mean points of Experiment A that did not separated the combined population in Figure 5.1 are depicted in Figure 5.13. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

Figure 5.13 Examples of unsuccessful initial pairs of mean points
6. Parallel populations

The two initial populations are shown in blue and green in Figure 6.1. The mean of the blue population is depicted by ‘+’ and the mean of the green population by ‘*’. The dashed ellipses represent the covariances of the individual populations. The major axes are equal to the square roots of the corresponding eigenvalues. The solid ellipse and the black dot depict the covariance and the mean of the combined population respectively. Table 6.1 shows the statistics of the initial populations.

**Table 6.1 Statistics of the initial populations**

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-7.997678</td>
<td>0.767930</td>
<td>21.853479</td>
</tr>
<tr>
<td></td>
<td>-0.406998</td>
<td>0.019526</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.019526</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>21.853479</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>8.036238</td>
<td>0.918393</td>
<td>-0.007768</td>
</tr>
<tr>
<td></td>
<td>-0.621375</td>
<td>-0.007768</td>
<td>24.077201</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.007768</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24.077201</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>0.019280</td>
<td>65.114783</td>
<td>-0.853445</td>
</tr>
<tr>
<td></td>
<td>-0.514187</td>
<td>-0.853445</td>
<td>22.976829</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.853445</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.976829</td>
<td></td>
</tr>
</tbody>
</table>

The two initial populations were merged and extracting the original populations using the algorithm described in section 1. Introduction was carried out with initial conditions of Experiments A and B.

There were 28 different solutions or convergences from both Experiments A and B. Experiment A gave rise to 12 solutions and Experiment B gave rise to 16 solutions, i.e. none of the solutions came from both Experiments A and B. 6 pairs of initial mean points of Experiment A and 41 initial covariances of Experiment B did not separate the combined population, i.e. insufficient number of points (less than three) in one component and the rest are in the other. The following section presents some of the selected converged solutions. The solutions are selected such that they resulted from more initial conditions for either Experiment A or B.
6.1 Examples of separated populations

Example 1: resulted from Experiment A

Figure 6.2 depicts a selected converged solution for the initial populations shown in Figure 6.1. In fact the solution is identical to the original populations. Table 6.2 shows the statistics of the converged populations. 75 pairs of initial mean points of Experiment A converged to this solution, with average number of iterations for convergence 7.840000 and standard deviation 4.242766. Some of these initial pairs of mean points are depicted in Figure 6.3. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

Table 6.2 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>8.036238</td>
<td>0.918393</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>-0.621375</td>
<td>-0.007768</td>
<td>24.077201</td>
</tr>
<tr>
<td>Green</td>
<td>-7.997678</td>
<td>0.767930</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>-0.406998</td>
<td>0.019526</td>
<td>21.853479</td>
</tr>
<tr>
<td>Combined</td>
<td>0.019280</td>
<td>65.114783</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>-0.514187</td>
<td>-0.853445</td>
<td>22.976829</td>
</tr>
</tbody>
</table>

Figure 6.2 Separated populations
The estimated AIC values are:

\[ AIC_1 = 290.564960 \]
\[ AIC_2 = 290.564960 \]

*Example 2: resulted from Experiment A*
Figure 6.4 depicts another selected converged solution for the initial populations shown in Figure 6.1. Table 6.3 shows the statistics of the converged populations. 6 pairs of initial mean points of Experiment A converged to this solution, with average number of interactions for convergence 7.000000 and standard deviation 4.760952. These initial pairs of mean points are depicted in Figure 6.5. The ‘+’ and ‘*’ in each colour represent each pair of random initial mean points.

Table 6.3 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-1.173667</td>
<td>63.754093</td>
<td>-5.457491</td>
</tr>
<tr>
<td>Green</td>
<td>1.165445</td>
<td>63.741106</td>
<td>-5.061874</td>
</tr>
<tr>
<td>Combined</td>
<td>0.019280</td>
<td>65.114783</td>
<td>-0.853445</td>
</tr>
</tbody>
</table>

Figure 6.5 Examples of successful initial pairs of mean points

The estimated AIC values are:

\[ AIC_1 = -40.506219 \]
\[ AIC_2 = -14.646831 \]
Example 3: resulted from Experiment B

Figure 6.6 depicts another selected converged solution for the initial populations shown in Figure 6.1. Table 6.4 shows the statistics of the converged populations. 16 pairs of initial covariances of Experiment B converged to this solution, with average number of iterations for convergence 12.750000 and standard deviation 0.661438. Some of these initial pairs of covariances are depicted in Figure 6.7. Each colour represents a pair.

Table 6.4 Statistics of the separated populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>-1.367471, 0.047295</td>
<td>-1.388535, 38.484512</td>
<td>111</td>
</tr>
<tr>
<td>Green</td>
<td>1.748823, -1.214462</td>
<td>1.996177, 2.752211</td>
<td>89</td>
</tr>
<tr>
<td>Combined</td>
<td>0.019280, -0.514187</td>
<td>-0.853445, 22.976829</td>
<td>200</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[ AIC_1 = -72.807954 \]
\[ AIC_2 = -23.447301 \]
Example 4: resulted from Experiment B

Figure 6.7 Examples of successful initial pairs of covariances

Figure 6.8 Separated populations
Figure 6.8 depicts another selected converged solution for the initial populations shown in Figure 6.1. Table 6.5 shows the statistics of the converged populations. 7 pairs of initial covariances of Experiment B converged to this solution, with average number of iterations for convergence 12.857143 and standard deviation 2.166536. These initial pairs of covariances are depicted in Figure 6.9. Each colour represents a pair.

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean</th>
<th>Covariance</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>1.340049</td>
<td>63.036136</td>
<td>3.710449</td>
</tr>
<tr>
<td>Green</td>
<td>-1.500314</td>
<td>63.190146</td>
<td>5.381791</td>
</tr>
<tr>
<td>Combined</td>
<td>0.019280</td>
<td>65.114783</td>
<td>-0.853445</td>
</tr>
</tbody>
</table>

The estimated AIC values are:

\[
AIC_1 = -40.649783 \\
AIC_2 = -13.701798
\]

Figure 6.9 Examples of successful initial pairs of covariances
The 6 pairs of initial mean points of Experiment A that did not separate the combined population in Figure 6.1 are depicted in Figure 6.10. These initial conditions resulted in insufficient number of points (less than three) in one component and the rest are in the other.

Figure 6.10 Examples of unsuccessful initial pairs of mean points

8 of the 41 pairs of initial covariances of Experiment B that did no separate the combined population in Figure 6.1 are depicted in Figure 6.11.

Figure 6.11 Example of unsuccessful initial pairs of covariances
7. Conclusions

The combined population presented in section 3. Non overlapping populations resulted in a unique solution for all the initial conditions of Experiments A and B, excluding the initial conditions, which did not separate the population. All the other combined populations resulted in more than one solution. However the combined population in section 3 is the easiest to separate, since it consists of individual populations with different mean points as well as covariances, without any overlap of the data points. As the overlap between the individual populations increases there are more solutions. However not all of them are equally acceptable.

An example of more solutions without any overlap is given in section 6. Parallel populations. In this example the y coordinates of the mean points are equal as well as the covariances. Thus an acceptable solution may only be achieved by initial mean points with offsets in the x direction.

The initial conditions, which did not separate the combined populations had one Gaussian distribution with a higher probability than the other one for all the points. This is possible due to one mean point being further away than the other one from both individual populations in the case of Experiment A, and one covariance being concentrated in an area without any points in the case of Experiment B.

Positive AIC values do not necessarily indicate a good model fit. Example 2 of section 4. Overlapping populations demonstrates a very poor fit of the model parameters for the given data. However AIC2 is positive. AIC1 is always less than or equal to AIC2 and in this example it is negative, indicating the poor fit. A negative AIC1 reliably indicates a poor fit. However neither a positive AIC1 nor a positive AIC2 reliably indicate a good fit. Examples 1 and 3 of section 4. Overlapping populations demonstrates poor model fits, and both AIC1 and AIC2 are positive. However in these examples one population has only about 2% or less of the total data points. Therefore a combination of AIC and the proportions of the data points have the potential to develop a reliable measure of model fit. The AIC values of clearly good model fits are remarkably higher than the ambiguous cases, usually greater than 50. Furthermore AIC1 and AIC2 become closer as the overlap between the two Gaussian components of the mixture model reduces.

Therefore an improved criterion may be used to obtain a measure of model fit and to identify acceptable solutions. However there is no guarantee that the best possible solution is among the solutions achieved in the process since the dependence of convergences on the initial conditions.
8. Appendix

Matlab code 1

% Generates the two initial Gaussian populations and save them in % gauss1.txt and gauss2.txt % The mean points and covariances of initial populations as well as % the combined population are saved in meancov.txt % The combined population is separated using an iterative algorithm, % with 200 different initial conditions % The initial and final means and covariances, number of iterations % of convergence, number of samples in converged populations and the % AIC values are % saved in results.txt % All the final converged solutions are also saved in the files % f<i>gauss1.txt and f<i>gauss2.txt, where 1 <= i <= 200

clear; hf = 1; % handle of the current figure
n1 = 100; % number of data points in the first component
n2 = 100; % number of data points in the second component
n = n1+n2;
N_itr = 100; % number of iterations in the classification
x1 = randn(2,n1); % 2*n1 Gaussian matrix
x2 = randn(2,n2); % 2*n2 Gaussian matrix

As1 = [1,0;0,5]; % scaling
Ah1 = [1,0;2,1]; % shearing
theta = 0; % rotation
Ar1 = [cos(theta), -sin(theta); sin(theta), cos(theta)];

As2 = [1,0;0,5]; % scaling
Ah2 = [1,2;0,1]; % shearing
theta = 0; % rotation
Ar2 = [cos(theta), -sin(theta); sin(theta), cos(theta)];

c1 = [-8,0;0,0]; % translation
c2 = [8,0;0,0];
mu1 = ones(2,n1);
mu1 = c1*mu1;
mu2 = ones(2,n2);
mu2 = c2*mu2;

% generate two 2-D Gaussian populations
g1 = Ar1*As1*x1 + mu1;
g2 = Ar2*As2*x2 + mu2;

% concatenate g1 and g2 to produce 2*(n = n1+n2) matrix G
G = [g1,g2];
% set the initial conditions using mean and covariance of G
muG = mean(G,2);
covG = cov(G',1); % maximum likelihood covariance instead of unbiased one
muG1 = mean(g1,2);
covG1 = cov(g1',1); % maximum likelihood covariance instead of unbiased one
muG2 = mean(g2,2);
covG2 = cov(g2',1); % maximum likelihood covariance instead of unbiased one

% plot the second raw of g1 versus the first raw of g1
x1 = [1,0];
x1 = x1*g1; % first raw of g1
x2 = [0,1];
x2 = x2*g1; % second raw of g1
figure(hf);
plot(x1,x2, 'b.');
hf = hf + 1;

% plot the second raw of g2 versus the first raw of g2
x1 = [1,0];
x1 = x1*g2; % first raw of g2
x2 = [0,1];
x2 = x2*g2; % second raw of g2
figure(hf);
plot(x1,x2, 'g.');
hf = hf + 1;

% plot the second raw of G versus the first raw of G
figure(hf); hold on; plot(x1,x2, 'g.');
x1 = [1,0]; x1 = x1*g1; x2 = [0,1]; x2 = x2*g1;
plot(x1,x2, 'b.');

% plot an ellipse to represent the covarariance of G
ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.');
axis equal; hold off; hf = hf + 1;

% save the initial populations, means, and covariances
f_mmeancov = fopen('data\mean cov.txt', 'wt');
fprintf(f_mmeancov, '%f\n%f\n%f\n%f\n%f\n%f\n%f\n%f\n%f', muG1, muG2, muG);
fprintf(f_mmeancov, '%f\n%f\n%f\n%f\n%f\n%f\n%f\n%f\n%f', covG1', covG2', covG');
fclose(f_mmeancov);
f_gauss1 = fopen('data\gauss1.txt', 'wt');
fprintf(f_gauss1, '%f %f\n', g1);
fclose(f_gauss1);
f_gauss2 = fopen('data\gauss2.txt', 'wt');
fprintf(f_gauss2, '%f %f\n', g2);
fclose(f_gauss2);

r = det(covG)^0.25; % an estimate of the radius of the combined population
r = 6*r;

f_results = fopen('data\results.txt', 'wt');
% run with different initial conditions
for ii = 1:200,
    if ii > 100
        muG1 = muG;
        muG2 = muG;
        covG1 = covG*ceil(100*rand(1))/100;
        covG2 = covG*ceil(100*rand(1))/100;
    else
        muG1 = muG + [r*(rand(1)-0.5); r*(rand(1)-0.5)];
        muG2 = muG + [r*(rand(1)-0.5); r*(rand(1)-0.5)];
        covG1 = covG;
        covG2 = covG;
    end
end
fprintf(f_results, '%f\n%f\n\n%f\n\n%f\n\n%f\n\n%f\n\n%f\n\n%f\n\n', muG1, muG2);
fprintf(f_results, '\n\n%f %f \n%f %f \n%f %f \n%f %f \n%f %f \n%f %f \n%f %f %f\n\n', covG1',
covG2');

% do N_itr iterations of the classification
cl_old = zeros(n,1); % memory for the previous
classification
for a = 1:N_itr,

% constant factors of the 2-D Gaussian density functions
fac1 = 1/(2*pi*sqrt(det(covG1)));
fac2 = 1/(2*pi*sqrt(det(covG2)));
exp1 = -covG1\eye(2)/2; % constant factors of the exponent
exp2 = -covG2\eye(2)/2; % inv(A) = A\eye(size(A)); about 2-3
times faster
MUG1 = ones(1,n); MUG1 = muG1*MUG1;
MUG2 = ones(1,n); MUG2 = muG2*MUG2;

% classify according to the probabilities
p1 = fac1*exp(diag((G-MUG1)'*exp1*(G-MUG1)));
p2 = fac2*exp(diag((G-MUG2)'*exp2*(G-MUG2)));
cl = p2 > p1;
k = sum(cl); j = n - k;

G1 = zeros(2,j); G2 = zeros(2,k);
j = 1; k = 1;
for i = 1:n,
if cl(i)
G2(2*k-1) = G(2*i-1);
G2(2*k) = G(2*i);
k = k + 1;
else
G1(2*j-1) = G(2*i-1);
G1(2*j) = G(2*i);
j = j + 1;
end
end

% check for empty populations and not enough points to compute a
covariance
if prod(size(G1)) < 6 | prod(size(G2)) < 6
G1 = G;
muG1 = muG;
mug2 = mug;
covG1 = covG;
covG2 = covG;
j = 1;
k = 1;
break;
end

% recompute the means and covariances
muG1 = mean(G1,2);
mug2 = mean(G2,2);
covG1 = cov(G1',1); % maximum likelihood covariance instead
of unbiased one
covG2 = cov(G2',1);

% check for singularities in covariance matrices
if cond(covG1) > 1000 | cond(covG2) > 1000
G1 = G;
muG1 = muG;
mug2 = mug;
covG1 = covG;
covG2 = covG;

break;
end

48
\[ j = 1; \quad k = 1; \]
break;
end

% check for no changes in the points
cl_old = abs(cl_old - cl);
if sum(cl_old) == 0
\break;
end
cl_old = cl;
end

% end of the main classification loop

N1 = j - 1; N2 = k - 1; N = N1+N2;
L = \log(\text{det}(\text{cov}G)); L1 = \log(\text{det}(\text{cov}G1)); L2 = \log(\text{det}(\text{cov}G2));

% The lower AIC gives the better solution, common factor 2 is not used
AIC_1 = N/2*L + 5; \quad \% AIC for one component
AIC_2 = N1/2*L1 - N1*log(N1/N) + N2/2*L2 - N2*log(N2/N) + 11; \quad \% AIC for two components

% AIC +ve: two components; -ve: one component
AICs = AIC_1 - AIC_2;
if ~(AICs <= 100 \mid AICs > 100)
\quad \% \text{AICs} = 10000;
end

% log likelihood of the combined population
fac = 1/(2*pi*sqrt(\text{det}(\text{cov}G)));
fac_exp = -\text{covG}\text{\textbackslash eye}(2)/2; \quad \% constant factors of the exponent
MUG = \text{ones}(1,n); MUG = muG*MUG;
p = fac*exp(diag((G-MUG)'*fac_exp*(G-MUG)));
L = sum(log(p));
AIC_1 = -L + 5;

% log likelihoods of the classified populations
fac1 = 1/(2*pi*sqrt(\text{det}(\text{cov}G1))) * N1/N;
fac2 = 1/(2*pi*sqrt(\text{det}(\text{cov}G2))) * N2/N;
exp1 = -\text{covG1}\text{\textbackslash eye}(2)/2; \quad \% constant factors of the exponent
exp2 = -\text{covG2}\text{\textbackslash eye}(2)/2; \quad \% \text{inv}(A) = A\text{\textbackslash eye(size(A))}; \text{about 2-3 times faster}
MUG1 = \text{ones}(1,n); MUG1 = muG1*MUG1;
MUG2 = \text{ones}(1,n); MUG2 = muG2*MUG2;
p1 = fac1*exp(diag((G-MUG1)'*exp1*(G-MUG1)));
p2 = fac2*exp(diag((G-MUG2)'*exp2*(G-MUG2)));
Ls = sum(log(p1+p2));
AIC_2 = -Ls + 11;
AICC = AIC_1 - AIC_2;
if ~(AICC <= 100 \mid AICC > 100)
\quad \% \text{AICC} = 10000;
end

% save final results, populations, means, covariances, and AIC
fprintf(f_results, '\n\n\%f\n\%f\n\%f\n\%f', muG1, muG2);
fprintf(f_results, '\n\n\%f \%f \%f \%f', covG1',
covG2');
fprintf(f_results, '\n\n\%d \%d', a, N1, N2);
fprintf(f_results, '\n\n\%f \%f', AICs, AICC);
Matlab code 2

% Analyzes the file results.txt and identifies the different solutions
% A new results file nresults.txt is created with only one entry for each different solution
% The first occurrence of the solution in the 200 initial conditions, number of iterations for that occurrence, number of samples in the converged solutions, the final means and covariances, and the AIC values are copied from results.txt
% The number of initial conditions which did not separate the initial populations, and the total number of solutions are saved in nresults.txt
% For each solution the total number of initial conditions, initial conditions from Experiments A and B, the corresponding average number of iterations for convergence and the standard deviation values, and the first occurrence of the solution in the 200 initial conditions are also saved
% The initial mean points of Experiment A which did not separate the populations are saved in nmeans.txt
% The initial covariances of Experiment B which did not separate the populations are saved in ncovs.txt
% The successful initial mean points and covariances of each solution are saved in the files immeans<i>.txt and icovs<i>.txt, where 1 <= i <= number of different solutions

j = 0; k = 0;
v_itr = zeros(0); itr_index = zeros(0);
mu_itr = zeros(0); mu_index = zeros(0);
cov_itr = zeros(0); cov_index = zeros(0);
fp_arr = zeros(0); fp_cov = zeros(0); file_list = zeros(0);
mu1Array = zeros(2,0); mu2Array = zeros(2,0);

f_results = fopen('data\results.txt', 'rt');
f_nresults = fopen('data\nresults.txt', 'wt');
f_nmeans = fopen('data\nmeans.txt', 'wt');
f_ncovs = fopen('data\ncovs.txt', 'wt');

for ii = 1:200,

muG1i = fscanf(f_results, '%f', 2);
muG2i = fscanf(f_results, '%f', 2);
covG1i = fscanf(f_results, '%f', [2,2]); covG1i = covG1i';

cl
51

covG2i = fscanf(f_results, '%f', [2,2]);  covG2i = covG2i';
uG1 = fscanf(f_results, '%f', 2);
uG2 = fscanf(f_results, '%f', 2);

covG1 = fscanf(f_results, '%f', [2,2]);  covG1 = covG1';
covG2 = fscanf(f_results, '%f', [2,2]);  covG2 = covG2';

itr = fscanf(f_results, '%d', 1);
N1 = fscanf(f_results, '%d', 1);
N2 = fscanf(f_results, '%d', 1);
AICs = fscanf(f_results, '%f', 1);
AICc = fscanf(f_results, '%f', 1);
s = fscanf(f_results, '%c', 44);

if itr == 100 fprintf(1, '100 Iterations for file: %d\n', ii); end

tArray1 = uG1*ones(1,size(mu1Array,2));
tArray2 = uG2*ones(1,size(mu2Array,2));

diff1 = sum(mu1Array - tArray1, 1);
diff2 = sum(mu2Array - tArray2, 1);
diff3 = sum(mu1Array - tArray2, 1);
diff4 = sum(mu2Array - tArray1, 1);

n = 0;
for m = 1:k,
    if (diff1(m) == 0 & diff2(m) == 0) | (diff3(m) == 0 & diff4(m) == 0)
        n = m;
    end
end

if N1 == 0
    j = j + 1;
    if ii > 100
        fprintf(f_ncovs, '%f %f %f %f %f %f %f\n', covG1i',
covG2i');
    end
    else
        fprintf(f_nmeans, '%f %f %f %f\n', mu1i, mu2i);
    end
elseif n > 0
    if ii > 100
        fprintf(fp_cov(n), '%f %f %f %f %f %f %f\n', covG1i',
covG2i');
    end
    else
        cov_itr = [cov_itr, itr];  cov_index = [cov_index, n];
        fprintf(fp_arr(n), '%f %f %f %f\n', mu1i, mu2i);
        mu_itr = [mu_itr, itr];  mu_index = [mu_index, n];
        end
        v_itr = [v_itr, itr];  itr_index = [itr_index, n];
    else
        k = k + 1;  fp_arr = [fp_arr, 0];  fp_cov = [fp_cov, 0];
        file_list = [file_list, ii];
        fp_arr(k) = fopen(strcat('data\imeans', int2str(k), '.txt'),
            'wt');
        fp_cov(k) = fopen(strcat('data\icovs', int2str(k), '.txt'),
            'wt');
        if ii > 100
            fprintf(fp_cov(k), '%f %f %f %f %f %f %f\n', covG1i',
covG2i');
        end
        else
            cov_itr = [cov_itr, itr];  cov_index = [cov_index, k];
        else
            ...
fprintf(fp_arr(k), '%f %f %f\n', muG1i, muG2i);
mu_itr = [mu_itr, itr]; mu_index = [mu_index, k];
end

v_itr = [v_itr, itr]; itr_index = [itr_index, k];

mu1Array = [mu1Array, muG1];
mu2Array = [mu2Array, muG2];

fprintf(f_nresults, '%3d %3d', ii, itr);
fprintf(f_nresults, '\n%3d %3d', N1, N2);
fprintf(f_nresults, '\n\n%10f %10f\n%10f %10f\n%10f %10f\n%10f %10f', muG1, muG2);
fprintf(f_nresults, '\n\n%10f %10f %10f %10f\n%10f %10f\n%10f %10f\n%10f %10f', covG1', covG2');
fprintf(f_nresults, '\n\n%10f', AICs, AICc);
fprintf(f_nresults, '******************************************************\n\n');
end

f_nnresults = fopen('data\nnresults.txt', 'wt');
fprintf(f_nnresults, '%d %d\n\n', j, k);
for ii = 1:k,
    vv_itr = zeros(0); n = 0;
    for j = 1:size(itr_index,2)
        if itr_index(j) == ii
            vv_itr = [vv_itr, v_itr(j)];
            n = n + 1;
        end
    end

vmu_itr = zeros(0); mu_n = 0;
for j = 1:size(mu_index,2)
    if mu_index(j) == ii
        vmu_itr = [vmu_itr, mu_itr(j)];
        mu_n = mu_n + 1;
    end
end
if size(vmu_itr,2) == 0 vmu_itr = 0; end

vcov_itr = zeros(0); cov_n = 0;
for j = 1:size(cov_index,2)
    if cov_index(j) == ii
        vcov_itr = [vcov_itr, cov_itr(j)];
        cov_n = cov_n + 1;
    end
end
if size(vcov_itr,2) == 0 vcov_itr = 0; end

fprintf(f_nnresults, '\n%10f %10f\n%10f %10f\n%10f %10f\n%10f %10f', mean(vv_itr),
sqrt(var(vv_itr,1)));
fprintf(f_nnresults, '| %10f %10f | ', mean(vmu_itr),
sqrt(var(vmu_itr,1)));
fprintf(f_nnresults, '| %10f %10f | ', mean(vcov_itr),
sqrt(var(vcov_itr,1)));
fprintf(f_nnresults, '%3d\n\n', file_list(ii));
end
fclose(f_nnresults);
fclose(f_results);
fclose(f_nresults);
fclose(f_nmeans);
fclose(f_ncovs);

for ii = 1:k,
    fclose(fp_arr(ii));
    fclose(fp_cov(ii));
end

Matlab code 3

% Plot the initial populations and results of Experiment A

clear; hf = 1;

f_nnresults = fopen('data\nnresults.txt', 'rt');
nn = fscanf(f_nnresults, '%d', 2); ss = nn(2);
fprintf(1, 'Number of initial conditions which did not separate: %d\n', nn(1));
fprintf(1, 'Total number of solutions: %d\n', ss);
nn = fscanf(f_nnresults, '%d%d%c%c%f%f%c%c%f%f%c%c%', [18, inf]);
fclose(f_nnresults);
if ss ~= size(nn, 2)
    fprintf(2, 'Error in nnresults.txt\n');
    return;
end
tmp_nn = nn; lnn = zeros(0);
for ii = 1:4,
    rr = 0;
    for jj = 1:ss,
        if (tmp_nn(2, jj) > rr)
            kk = jj; rr = tmp_nn(2, kk);
        end
    end
    if rr > 0
        lnn = [lnn, kk]; tmp_nn(2, kk) = 0;
    else
        break;
    end
end
ss = size(lnn, 2);

f_gauss1 = fopen('data\gauss1.txt', 'rt');
g1 = fscanf(f_gauss1, '%f', [2, inf]);
fclose(f_gauss1);

f_gauss2 = fopen('data\gauss2.txt', 'rt');
g2 = fscanf(f_gauss2, '%f', [2, inf]);
fclose(f_gauss2);

% concatenate g1 and g2 to produce 2*(n = n1+n2) matrix G
G = [g1, g2];
muG = mean(G, 2); covG = cov(G, 1);
muG1 = mean(g1, 2); covG1 = cov(g1, 1);
muG2 = mean(g2, 2); covG2 = cov(g2, 1);
% plot the second row of G versus the first row of G
figure(hf); box on; hold on;
x1 = [1,0]; x1 = x1*gl; x2 = [0,1]; x2 = x2*gl; plot(x1,x2, 'b.');
x1 = [1,0]; x1 = x1*g2; x2 = [0,1]; x2 = x2*g2; plot(x1,x2, 'g.');

% plot ellipses to represent the covariance of G
ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k');
ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
axis equal; hold off; hf = hf + 1;

% plot the means
fprintf('
Means Covs 1st_file hf
');
for kk = 1:ss,
  cc = lnn(kk);
  fprintf1, '%d\t%d\t%d\t%d\n', nn(2,cc), nn(3,cc), nn(18,cc), hf);
  f_imens = fopen(strcat('data\imeans', int2str(cc), '.txt'), 'rt');
  muGi = fscanf(f_imens, '%f', [4,inf]);
  fclose(f_imens);
  a = min(24, size(muGi,2));
  if a > 0
    x1 = [1,0,0,0]; x1 = x1*muGi; x2 = [0,1,0,0]; x2 = x2*muGi;
    x3 = [0,0,1,0]; x3 = x3*muGi; x4 = [0,0,0,1]; x4 = x4*muGi;
  end
  for k = 1:a,
    switch round(6*(k/6 - floor(k/6)))
    case 1, figure(hf); box on; hold on;
      ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k');
      ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
      ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
      hf = hf + 1;
    end
  end
end

% plot the means, which didn't work
f_nmeans = fopen('data\nmeans.txt', 'rt');
muGn = fscanf(f_nmeans, '%f', [4,inf]);
fclose(f_nmeans);
a = min(12, size(muGn,2));
if a > 0
  fprintf('Figures of means which did not work starts from: %d\n', hf);
  x1 = [1,0,0,0]; x1 = x1*muGn; x2 = [0,1,0,0]; x2 = x2*muGn;
  x3 = [0,0,1,0]; x3 = x3*muGn; x4 = [0,0,0,1]; x4 = x4*muGn;
end
for k = 1:a,
switch round(6*(k/6 - floor(k/6)))
    case 1, figure(hf); box on; hold on;
        ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.');
        ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
        ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
        hf = hf + 1;
        plot(x1(k), x2(k), 'm+', x3(k), x4(k), 'm*');
    case 2, plot(x1(k), x2(k), 'c+', x3(k), x4(k), 'c*');
    case 3, plot(x1(k), x2(k), 'r+', x3(k), x4(k), 'r*');
    case 4, plot(x1(k), x2(k), 'g+', x3(k), x4(k), 'g*');
    case 5, plot(x1(k), x2(k), 'b+', x3(k), x4(k), 'b*');
    case 0, plot(x1(k), x2(k), 'y+', x3(k), x4(k), 'y*');
end
end

for kk = 1:ss,
    cc = nn(18, lnn(kk));
    f_flgauss1 = fopen(strcat('data\f', int2str(cc), 'gauss1.txt'), 'rt');
    fgauss1 = fscanf(f_flgauss1, '%f', [2, inf]);
    fclose(f_flgauss1);
    f_flgauss2 = fopen(strcat('data\f', int2str(cc), 'gauss2.txt'), 'rt');
    fgauss2 = fscanf(f_flgauss2, '%f', [2, inf]);
    fclose(f_flgauss2);
    muG1 = mean(fgauss1,2); covG1 = cov(fgauss1',1);
    muG2 = mean(fgauss2,2); covG2 = cov(fgauss2',1);

    figure(hf); box on; hold on;
    x1 = [1,0]; x1 = x1*fgauss1; x2 = [0,1]; x2 = x2*fgauss1; plot(x1,x2, 'b.');
    x1 = [1,0]; x1 = x1*fgauss2; x2 = [0,1]; x2 = x2*fgauss2; plot(x1,x2, 'g.');
    ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.');</n
Matlab code 4

% Plot the initial populations and results of Experiment B

clear; hf = 1;

f_nnresults = fopen('data\nnresults.txt', 'rt');
nn = fscanf(f_nnresults, '%d', 2); ss = nn(2);
fprintf(1, 'Number of initial conditions which did not separate: %d\n', nn(1));
fprintf(1, 'Total number of solutions: %d\n', ss);
nn = fscanf(f_nnresults, '%d%d%d%d%d%d%d%d%d%d', [18,inf]);
close(f_nnresults);
if ss ~= size(nn,2)
    fprintf(2, 'Error in nnresults.txt\n');
    return;
end

tmp_nn = nn; lnn = zeros(0);
for ii = 1:4,
    rr = 0;
    for jj = 1:ss,
        if (tmp_nn(3, jj) > rr)
            kk = jj; rr = tmp_nn(3, kk);
        end
    end
    if rr > 0
        lnn = [lnn, kk]; tmp_nn(3, kk) = 0;
    else
        break;
    end
end
ss = size(lnn, 2);

f_gauss1 = fopen('data\gauss1.txt', 'rt');
g1 = fscanf(f_gauss1, '%f', [2,inf]);
close(f_gauss1);
f_gauss2 = fopen('data\gauss2.txt', 'rt');
g2 = fscanf(f_gauss2, '%f', [2,inf]);
close(f_gauss1);

% concatenate g1 and g2 to produce 2*(n = n1+n2) matrix G
G = [g1, g2];
muG = mean(G,2); covG = cov(G,1);
muG1 = mean(g1,2); covG1 = cov(g1,1);
muG2 = mean(g2,2); covG2 = cov(g2,1);

% plot the second raw of G versus the first raw of G
figure(hf); box on; hold on;
x1 = [1,0]; x1 = x1*g1; x2 = [0,1]; x2 = x2*g1; plot(x1,x2, 'b.');
x1 = [1,0]; x1 = x1*g2; x2 = [0,1]; x2 = x2*g2; plot(x1,x2, 'g.');

% plot ellipses to represent the covariance of G
ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
axis equal; hold off; hf = hf + 1;

% plot the covariances
fprintf(\nMeans Covs 1st_file hf\n');
for kk = 1:ss,
    cc = lnn(kk);
    fprintf(1, '%d%d%d%d\n', nn(2,cc), nn(3,cc), nn(18,cc), hf);
    f_icovs = fopen(strcat('data\icovs', int2str(cc), '.txt'), 'rt');
    covGi = fscanf(f_icovs, '%f', [8,inf]);
close(f_icovs);
a = min(8, size(covGi,2));

for k = 1:a
    switch round(6*(k/6 - floor(k/6)))
        case 1, figure(hf); box on; hold on;
```matlab
% plot the covariances, which didn't work
f_nccovs = fopen('data\ncovs.txt', 'rt');
covGn = fscanf(f_nccovs, '%f', [8, inf]);
close(f_nccovs);
a = min(8, size(covGn, 2));
if a > 0
    fprintf(1, 'Figures of covariances which did not work starts from: %d\n', hf);
end

for k = 1:a
    switch round(6*(k/6 - floor(k/6)))
        case 1, figure(hf); box on; hold on;
            ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.'),
            ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+'),
            ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*'),
            h = h + 1;
        case 2, plotellipse(muG, covGn, k, 'm--');
        case 3, figure(hf); box on; hold on;
            ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.'),
            ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+'),
            ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*'),
            h = h + 1;
        case 4, plotellipse(muG, covGn, k, 'r--');
        case 5, figure(hf); box on; hold on;
            ellipse(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.'),
            ellipse(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+'),
            ellipse(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*'),
            h = h + 1;
        case 6, plotellipse(muG, covGn, k, 'g--');
    end
end
end
```
ellipsoid(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.');
ellipsoid(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
ellipsoid(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
hf = hf + 1;

plotellipsoid(muG, covGn, k, 'r--');
case 4, plotellipsoid(muG, covGn, k, 'g--');
case 5, figure(hf); box on; hold on;
    ellipsoid(muG, covG, 'k-', hf); plot(muG1(1), muG1(2), 'k.');
    ellipsoid(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');
    ellipsoid(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');
    hf = hf + 1;
    plotellipsoid(muG, covGn, k, 'b--');
case 0, plotellipsoid(muG, covGn, k, 'y--');
end
end

for kk = 1:ss,
    cc = nn(18, lnn(kk));
    f_f1gauss1 = fopen(strcat('data\f', int2str(cc), 'gauss1.txt'), 'rt');
    fgauss1 = fscanf(f_f1gauss1, '%f', [2, inf]);
    fclose(f_f1gauss1);
    f_f1gauss2 = fopen(strcat('data\f', int2str(cc), 'gauss2.txt'), 'rt');
    fgauss2 = fscanf(f_f1gauss2, '%f', [2, inf]);
    fclose(f_f1gauss2);
    muG1 = mean(fgauss1,2); covG1 = cov(fgauss1,1);
    muG2 = mean(fgauss2,2); covG2 = cov(fgauss2,1);
figure(hf); box on; hold on;
    x1 = [1,0]; x1 = x1*fgauss1; x2 = [0,1]; x2 = x2*fgauss1; plot(x1,x2, 'b.');
    x1 = [1,0]; x1 = x1*fgauss2; x2 = [0,1]; x2 = x2*fgauss2; plot(x1,x2, 'g.');
    ellipsoid(muG, covG, 'k-', hf); plot(muG(1), muG(2), 'k.');</ref>
    ellipsoid(muG1, covG1, 'k--', hf); plot(muG1(1), muG1(2), 'k+');</ref>
    ellipsoid(muG2, covG2, 'k--', hf); plot(muG2(1), muG2(2), 'k*');</ref>
    axis equal; hold off; hf = hf + 1;
end

**Auxiliary functions 1**

% Plot an ellipse to represent the population with,
% mean muG and covariance covG

function ellipse(muG, covG, style, hf)
    [v, d] = eig(covG); c = 1;
\( x = \text{linspace}(-\sqrt{d(1,1)}*c, \sqrt{d(1,1)}*c, 1000); \)
\( yp = \text{real}((c^2 - x.^2/d(1,1))*d(2,2)); \)
\( xp = [x;yp]; \)
\( yn = -yp; \)
\( xn = [x;yn]; \)
\( ev = v(1,1) + i*v(2,1); \)
\( \theta = \text{angle}(ev); \)
\( R = [\cos(\theta), -\sin(\theta); \sin(\theta), \cos(\theta)]; \)
\( \mu = \text{ones}(2,1000); \)
\( \mu_{GG} = [\mu_{G}(1),0;0,\mu_{G}(2)]; \)
\( \mu = \mu_{GG}.*\mu; \)
\( xp = R*xp + \mu; \)
\( xn = R*xn + \mu; \)
\( xg1 = [1,0]; \)
\( xg1p = xg1*xp; \)
\( xg2 = [0,1]; \)
\( xg2p = xg2*xp; \)
\( xg1 = [1,0]; \)
\( xg1n = xg1*xn; \)
\( xg2 = [0,1]; \)
\( xg2n = xg2*xn; \)

\text{figure(hf); hold on;}
\text{plot}(xg1p, xg2p, style); \text{plot}(xg1n, xg2n, style); \)

\textbf{Auxiliary functions 2}

\texttt{function\ plotellipse(muG, covGi, k, style)}

\texttt{covG1 = [covGi(1, k), covGi(2, k); covGi(3, k), covGi(4, k)];}
\texttt{covG2 = [covGi(5, k), covGi(6, k); covGi(7, k), covGi(8, k)];}
\texttt{muG1 = muG; muG2 = muG;}

\texttt{[v, d] = eig(covG1); c = 1;}
\texttt{x = linspace(-sqrt(d(1,1))*c, \sqrt{d(1,1)}*c, 1000);}
\texttt{yp = real((c^2 - x.^2/d(1,1))*d(2,2));}
\texttt{xp = [x;yp];}
\texttt{yn = -yp; xn = [x;yn];}
\texttt{ev = v(1,1) + i*v(2,1);}
\texttt{theta = angle(ev);}
\texttt{R = [\cos(\theta), -\sin(\theta); \sin(\theta), \cos(\theta)];}
\texttt{mu = \text{ones}(2,1000);}
\texttt{mu_{GG} = [\mu_{G}(1),0;0,\mu_{G}(2)];}
\texttt{mu = \mu_{GG}.*\mu;}
\texttt{xp = R*xp + \mu;}
\texttt{xn = R*xn + \mu;}
\texttt{xg11 = [1,0];}
\texttt{xg11p = xg11*xp;}
\texttt{xg12 = [0,1];}
\texttt{xg12p = xg12*xp;}
\texttt{xg11 = [1,0];}
\texttt{xg11n = xg11*xn;}
\texttt{xg12 = [0,1];}
\texttt{xg12n = xg12*xn;}

\texttt{[v, d] = eig(covG2); c = 1;}
\texttt{x = linspace(-sqrt(d(1,1))*c, \sqrt{d(1,1)}*c, 1000);}
\texttt{yp = real((c^2 - x.^2/d(1,1))*d(2,2));}
\texttt{xp = [x;yp];}
\texttt{yn = -yp; xn = [x;yn];}
\texttt{ev = v(1,1) + i*v(2,1);}
\texttt{theta = angle(ev);}
\texttt{R = [\cos(\theta), -\sin(\theta); \sin(\theta), \cos(\theta)];}
\texttt{mu = \text{ones}(2,1000);}
\texttt{mu_{GG} = [\mu_{G}(1),0;0,\mu_{G}(2)];}
\texttt{mu = \mu_{GG}.*\mu;}
\texttt{xp = R*xp + \mu;}
\texttt{xn = R*xn + \mu;}
\texttt{xg21 = [1,0];}
\texttt{xg21p = xg21*xp;}
\texttt{xg22 = [0,1];}
\texttt{xg22p = xg22*xp;}
\texttt{xg21 = [1,0];}
\texttt{xg21n = xg21*xn;}
\texttt{xg22 = [0,1];}
\texttt{xg22n = xg22*xn;}

\text{plot(xg11p, xg12p, style); plot(xg11n, xg12n, style);}
\text{plot(xg21p, xg22p, style); plot(xg21n, xg22n, style);}
9. References


