This is the readme file for the R script "JRSS C codes cleft and jaw data", which contains R codes related to Section 6.3, Test Results for Smile Data 1 and Smile Data 2: Deviation from Control Subjects, in the paper. In this section, the pre-surgery data and post-surgery data are compared with the control subjects from the cleft lip data separately. Here, the 10 landmarks over the lip are used.

The codes have been divided into 6 chunks with indices ranged from 0 to 5.

**Code chunk 0** contains codes for reading in the data for control subjects and convert these data into arrays at the three frames. These arrays are saved under the names:

1. con\_fir\_arr;

2. con\_mid\_arr;

3. con\_la\_arr.

Further, we scale the subjects from orthognathic data so that their scale is in mm. The scaled pre-surgery data are saved under the names:

1. jawpre\_arr\_cen\_fir;

2. jawpre\_arr\_cen\_mid;

3. jawpre\_arr\_cen\_la,

whereas the scaled post-surgery data are saved under the names:

4. jawpost\_arr\_cen\_fir;

5. jawpost\_arr\_cen\_mid;

6. jawpost\_arr\_cen\_la.

**Code chunk 1** contains the codes for computing signed elementary feature vectors defined by equations (2.1) - (2.3) in Section 2.2. This chunk has been divided into 3 sub-chunks:

(a). 1.1, where the vectors for control subjects are computed at the three frames and are saved under the names:

(1). d\_con\_fir;

(2). d\_con\_mid;

(3). d\_con\_la.

(b). 1.2, where the vectors for pre-surgery data are computed at the three frames and are saved under the names:

(1). d\_pre\_fir\_lip;

(2). d\_pre\_mid\_lip;

(3). d\_pre\_la\_lip,

(c). 1.3, where the vectors for post-surgery data are computed at the three frames and are saved under the names:

(1). d\_post\_fir\_lip;

(2). d\_post\_mid\_lip;

(3). d\_post\_la\_lip.

**Code chunk 2** involves codes for computing the composite asymmetry scores. This chunk has been divided into 5 sub-chunks:

(a). 2.1, where the \phi^\*\_{L\_1} defined in equation (3.9) in Section 3.1 is computed for control subjects, pre- and post-surgery data at the three frames. The results are saved in three matrices, phistar\_L1\_con\_lip, phistar\_L1\_pre\_lip and phistar\_L1\_post\_lip, where each row is for each subject and each column is for each frame.

(b). 2.2, where the \phi\_{L\_1} defined in equation (3.6) in Section 3.1 is calculated for control subjects, pre- and post-surgery data at the three frames. The results are saved in three matrices, phi\_L1\_con, phi\_L1\_pre\_lip and phi\_L1\_post\_lip, where each row is for each subject and each column is for each frame.

(c). 2.3, where the \phi\_{L\_2} defined in equation (3.7) in Section 3.1 is calculated for control subjects, pre- and post-surgery data at the three frames. The results are saved in three matrices, phi\_L2\_con, phi\_L2\_pre\_lip and phi\_L2\_post\_lip, where each row is for each subject and each column is for each frame.

(d). 2.4, where we use weight as 1 for landmark pairs and 2 for solos. Then \phi\_{L\_1} defined in equation (3.6) in Section 3.1 is calculated for control subjects, pre- and post-surgery data at the three frames. The results are saved in three matrices, phi\_L1\_con\_wei, phi\_L1\_pre\_wei and phi\_L1\_post\_wei, where each row is for each subject and each column is for each frame.

(e). 2.5, where we use weight as 1 for landmark pairs and 2 for solos. Then \phi\_{L\_2} defined in equation (3.7) in Section 3.1 is calculated for control subjects, pre- and post-surgery data at the three frames. The results are saved in three matrices, phi\_L2\_con\_wei, phi\_L2\_pre\_wei and phi\_L2\_post\_wei, where each row is for each subject and each column is for each frame.

**Code chunk 3** contains the codes related to univariate tests. The two-sample t-tests for comparing pre- and post-surgery data with control subjects are performed in 3.1. The Mann-Whitney U tests are carried out in 3.2.

**Code chunk 4** contains the codes for carrying out feature selection, where the two-sample t-values are computed in 4.1 and bootstrap is performed in 4.2. Codes for Mann-Whitney U tests on \phi\_{L\_1} and \phi\_{L\_2} computed on various subsets of landmarks are in 4.3.

**Code chunk 5** contains the codes for Fisher's and Pearson's methods. We first calculate the separate p-values in 5.1. Then, Fisher's and Pearson's methods are executed in 5.2 and 5.3 respectively.