Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros Industriales

Máster en Automática y Robótica

Applied Artificial Intelligence

Assignment 2.1: Scatter Matrixes

Josep María Barberá Civera (17048)

February 13, 2024

ASSIGNMENT 2.1

Scatter Matrixes

1.1 Methodology

In this exercise you are asked to use arrays of labeled two-dimensional points. From these arrays calculate the scattering matrices for each data set as well as the joint scattering matrices for all of them.

To solve the problem we have implemented a function called `Scatter_Matrices` that only requires as input the vector of points and the vector of labels, this function is responsible for calculating how many groups there are and compute the associated matrices: those of each group (`Si`), those relating the groups (`Sw` and `Sb`) and the total (`St`).

These calculations as well as the traces of the previous matrices have been performed for the unnormalized data as well as for the normalized data.

1.2 Results

Not Normalized Data

P Values:

$$S1 = \begin{bmatrix} 22.9550 & 18.3158 \\ 18.3158 & 22.5492 \end{bmatrix} \quad \text{tr}(S1) = 45.5042$$

$$S2 = \begin{bmatrix} 67.8157 & 58.1329 \\ 58.1329 & 69.4335 \end{bmatrix} \quad \text{tr}(S2) = 137.2492$$

$$Sw = \begin{bmatrix} 90.7707 & 76.4488 \\ 76.4488 & 91.9827 \end{bmatrix} \quad \text{tr}(Sw) = 182.7534$$

$$Sb = \begin{bmatrix} 14.4762 & 45.8251 \\ 45.8251 & 145.0617 \end{bmatrix} \quad \text{tr}(Sb) = 159.5378$$

$$St = \begin{bmatrix} 105.2469 & 122.2739 \\ 122.2739 & 237.0444 \end{bmatrix} \quad \text{tr}(St) = 342.2913$$

T Values:

$$S1 = \begin{bmatrix} 18.4888 & 15.3679 \\ 15.3679 & 24.1245 \end{bmatrix} \quad \text{tr}(S1) = 42.6134$$

$$S2 = \begin{bmatrix} 66.9194 & 58.2243 \\ 58.2243 & 71.6451 \end{bmatrix} \quad \text{tr}(S2) = 138.5645$$

$$Sw = \begin{bmatrix} 85.4082 & 73.5922 \\ 73.5922 & 95.7696 \end{bmatrix} \quad \text{tr}(Sw) = 181.1778$$

$$Sb = \begin{bmatrix} 18.6147 & 55.3335 \\ 55.3335 & 164.4826 \end{bmatrix} \quad \text{tr}(Sb) = 183.0973$$

$$St = \begin{bmatrix} 104.0229 & 128.9257 \\ 128.9257 & 260.2522 \end{bmatrix} \quad \text{tr}(St) = 364.2751$$

Normalized Data

P Values:

$$S1 = \begin{bmatrix} 65.2138 & 34.6719 \\ 34.6719 & 28.4429 \end{bmatrix} \quad S2 = \begin{bmatrix} 192.6602 & 110.0459 \\ 110.0459 & 87.5811 \end{bmatrix} \quad Sw = \begin{bmatrix} 257.8740 & 144.7179 \\ 144.7179 & 116.0240 \end{bmatrix}$$

$$\mathrm{tr}(S1) = 93.6567 \quad \mathrm{tr}(S2) = 280.2414 \quad \mathrm{tr}(Sw) = 373.8980$$

$$Sb = \begin{bmatrix} 41.1260 & 86.7471 \\ 86.7471 & 182.9760 \end{bmatrix} \quad St = \begin{bmatrix} 299.0000 & 231.4650 \\ 231.4650 & 299.0000 \end{bmatrix}$$

$$\mathrm{tr}(Sb) = 224.1020 \quad \mathrm{tr}(St) = 598.0$$

T Values:

$$S1 = \begin{bmatrix} 53.1437 & 27.9270 \\ 27.9270 & 27.7163 \end{bmatrix} \quad S2 = \begin{bmatrix} 192.3508 & 105.8069 \\ 105.8069 & 82.3120 \end{bmatrix} \quad Sw = \begin{bmatrix} 245.4945 & 133.7339 \\ 133.7339 & 110.0283 \end{bmatrix}$$

$$\mathrm{tr}(S1) = 80.8600 \quad \mathrm{tr}(S2) = 274.6628 \quad \mathrm{tr}(Sw) = 355.5229$$

$$Sb = \begin{bmatrix} 53.5055 & 100.5536 \\ 100.5536 & 188.9717 \end{bmatrix} \quad St = \begin{bmatrix} 299.0000 & 234.2875 \\ 234.2875 & 299.0000 \end{bmatrix}$$

$$\mathrm{tr}(Sb) = 242.4771 \quad \mathrm{tr}(St) = 598.0$$

1.3 Discussions and Results

Result will be discussed during the class.

1.4 Relevant Code

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                Master in Robotics
%          Applied Artificial Intelligence
%
% Assinment 2.1: Scatter Matrices
% Student: Josep Barbera Civera
% ID: 17048
% Date: 12/02/2024
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% load data_D2_C2
% - Compute the scatter matrices S1, S2, SW, SB, ST and
   their traces
% - Compute the same matrices for the normalized data and
    their traces

load data_D2_C2.mat
```

```
16
17 %% Plotting Data
18 % plot(p.value(1,1:100),p.value(2,1:100),'+'); hold all;
19 % plot(p.value(1,101:300),p.value(2,101:300),'+');
20
21 %% Accesing Data
22 pvalues = p.value;
23 plabels = p.class;
24 tvalues = t.value;
25 tlabels = t.class;
26 %% Not normalized data
27 disp("----------- Not Normalized Data --------------")
28 disp("----- P Values -----")
29 pmatrices = Scatter_matrices(pvalues, plabels);
30 S1 = pmatrices{1}
31 trace1 = trace(S1)
32 S2 = pmatrices{2}
33 trace2 = trace(S2)
34 Sw = pmatrices{3}
35 trace_Sw = trace(Sw)
36 Sb = pmatrices{4}
37 trace_Sb = trace(Sb)
38 St = pmatrices{5}
39 trace_St = trace(St)
40
41 disp("----- T Values -----")
42 tmatrices = Scatter_matrices(tvalues, tlabels);
43 S1 = tmatrices{1}
44 trace1 = trace(S1)
45 S2 = tmatrices{2}
46 trace2 = trace(S2)
47 Sw = tmatrices{3}
48 trace_Sw = trace(Sw)
49 Sb = tmatrices{4}
50 trace_Sb = trace(Sb)
51 St = tmatrices{5}
52 trace_St = trace(St)
53
54 %% Normalized data
55 disp("-------------- Normalized Data --------------")
56 [~, N] = size(pvalues);
57 [~, Nt] = size(tvalues);
58 meanp = mean(pvalues')';
59 stdp = std(pvalues')';
60 for i = 1:N
61     pn(:,i) = (pvalues(:,i) - meanp)./stdp;
62 end
63 meant = mean(tvalues')';
```

```
stdt = std(tvalues')';
for i = 1:Nt
    tn(:,i) = (tvalues(:,i) - meant)./stdt;
end

disp("----- P Values -----")
pmatrices = Scatter_matrices(pn, plabels);
S1 = pmatrices{1}
trace1 = trace(S1)
S2 = pmatrices{2}
trace2 = trace(S2)
Sw = pmatrices{3}
trace_Sw = trace(Sw)
Sb = pmatrices{4}
trace_Sb = trace(Sb)
St = pmatrices{5}
trace_St = trace(St)

disp("----- T Values -----")
tmatrices = Scatter_matrices(tn, tlabels);
S1 = tmatrices{1}
trace1 = trace(S1)
S2 = tmatrices{2}
trace2 = trace(S2)
Sw = tmatrices{3}
trace_Sw = trace(Sw)
Sb = tmatrices{4}
trace_Sb = trace(Sb)
St = tmatrices{5}
trace_St = trace(St)

% Scatter_matrices waits a struct with a vector of data
   points (i.e. 2x100) (coordinates) and a vector with a
   label associated to this vector (i.e. 1x100).

function matrix_cell = Scatter_matrices(values, labels)
        % first we compute the number of labels
        unique_labels = unique(labels);
        N = length(unique_labels);
        % now we create as many vectors as labels: with
           cell arrays
        vectors_cell = cell(1, N);
        matrix_cell = cell (1, N+3);

        for i = 1:N
            label = unique_labels(i);
            indices = labels == label; % Find indices
               corresponding to the current label
```

```
            vectors_cell{i} = values(:, indices); % Store
                coordinates associated with the label
        end
        % Scatter matrix for each group is computed
        for i = 1:N
            data = vectors_cell{i};
            Sc = cov(data')*(length(data)-1);
            matrix_cell{i} = Sc;
        end
        % Scatter matrix within the groups is computed
        Sw = zeros(N);
        for i = 1:N
            Sw = Sw + matrix_cell{i};
        end
        s = N + 1;
        matrix_cell{s} = Sw;
        % Scatter matrix between the groups is computed
        Sb = zeros(N);
        m_x = mean(values(1,:));
        m_y = mean(values(2,:));
        m = [m_x; m_y];
        for i = 1:N
            data = vectors_cell{i};
            m_c_x = mean(data(1,:));
            m_c_y = mean(data(2,:));
            m_c = [m_c_x; m_c_y];
            Sb = Sb + length(data)*(m_c-m)*(m_c-m).';
        end
        s = s + 1;
        matrix_cell{s} = Sb;
        % Total Scatter matrix is computed
        St = (values - m)*(values - m).';
        s = s + 1;
        matrix_cell{s} = St;
end
```