Singular Value Decomposition and Linear Discriminant Analysis

Rhithika G.
Department of Computer Science, Cochin University of Science and Technology

Course: M.Tech Computer Science and Engineering with specialization in Data Science and Artificial Intelligence

Assignment : Machine Learning Algorithms ; (`rhithikag@pg.cusat.ac.in`)

Introduction

In most machine learning algorithms, the complexity depends on the number of input dimensions as well as on the size of the data sample. And for reduced memory and computation, we reduce the dimensionality of the problem. Thus we decrease the complexity of the inference algorithm during testing and we save the cost of extracting it etc. Some dimensionality reduction techniques are SVD - Singular Value Decomposition and LDA - Linear Discriminant Analysis. These techniques are discussed below.

Singular Value Decomposition (SVD)

The singular value decomposition allows us to decompose any [N {X} d] rectangular matrix. SVD of a matrix factorization of the given matrix is a factorization of that matrix into three matrices. In matrix factorization, a large matrix is written as a product of two or three matrices. The idea is that although the data may be too large, either it is sparse, or there is a high correlation and it can be represented in a space of fewer dimensions.

SVD of m X n matrix A is given by the formula:

$$\mathbf{A=UWV^T}$$

U: m X m matrix of the orthonormal eigenvectors of AA^T.

V^T: Transpose of an n X n matrix containing the orthonormal eigenvectors of A^TA.

W: an n X n diagonal matrix of the singular values which are the square roots of the eigenvalues of A^TA.

Example of SVD:

Consider a 3 X 2 matrix A = $\begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix}$

- **STEP 1: Find $\mathbf{AA^T}$:**

$$AA^T = \begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} . \begin{pmatrix} 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 4 \end{pmatrix}$$

- **STEP 2: Calculate the eigenvalues and eigenvector for the matrix $\mathbf{AA^T}$:**

The characteristic equation for the above matrix is:

$| AA^T - \lambda I | = 0$

$$\begin{pmatrix} 8-\lambda & 4 & 4 \\ 4 & 4-\lambda & 0 \\ 4 & 0 & 4-\lambda \end{pmatrix} = -\lambda^3 + 16\lambda^2 - 48\lambda$$

$= -\lambda.(\lambda^2 - 16\lambda + 48)$

$= -\lambda.(\lambda - 4).(\lambda - 12) = 0$

Therefore, the eigenvalues are :

$\lambda_1 = 12$, $\lambda_2 = 4$, $\lambda_3 = 0$

Now we will get the singular values:

$\sigma 1 = 2\sqrt{3}$

$\sigma_2 = 2$

$\sigma_3 = 0$

Now, we compute the eigenvectors of AA^T.

eigenvector v corresponding to the eigenvalue λ *we have,*

$AA^T * v = \lambda * v$

Then:

$AA^T * v - \lambda * v = (AA^T - \lambda * I) * v = 0$

For $\lambda_1 = 12$

$(A - \lambda_1 \; *I) * v = (A - 12*I) * v$

$$\begin{pmatrix} 8-12 & 4 & 4 \\ 4 & 4-12 & 0 \\ 4 & 0 & 4-12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -4 & 4 & 4 \\ 4 & -8 & 0 \\ 4 & 0 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which can be written as:

$\frac{x_1}{[(-8*-8)-(0*0)]} = \frac{-x_2}{[(4*-8)-(4*0)]} = \frac{x_3}{[(4*0)-(4*-8)]} = a$

$\frac{x_1}{64} = \frac{-x_2}{-32} = \frac{x_3}{32} = a$

i.e, which can also be written as:

$\frac{x_1}{2} = \frac{x_2}{1} = \frac{x_3}{1} = a$

Therefore,

$x_1 = 2a$

$x_2 = 1a$

$x_3 = 1a$

Hence the eigenvector v_1 for $\lambda_1 = 12$

$$v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

The orthonormal set of eigenvector v_1 is:

$$\begin{pmatrix} \frac{2}{\left[\sqrt{(2^2+1^2+1^2)}\right]} \\ \frac{1}{\left[\sqrt{(2^2+1^2+1^2)}\right]} \\ \frac{1}{\left[\sqrt{(2^2+1^2+1^2)}\right]} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$$

For $\lambda_2 = 4$

*(A- λ_2 *I)* v = (A- 4*I) * v:*

$$\begin{pmatrix} 8-4 & 4 & 4 \\ 4 & 4-4 & 0 \\ 4 & 0 & 4-4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 4 \\ 4 & 0 & 0 \\ 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which can be written as:

$\frac{x_1}{[(4*0)-(4*0)]} = \frac{-x_2}{[(4*0)-(4*4)]} = \frac{x_3}{[(4*0)-(4*4)]} = a$

$\frac{x_1}{0} = \frac{-x_2}{-16} = \frac{x_3}{-16} = a$

i.e, which can also be written as:

$\frac{x_1}{0} = a : x_1 = 0a$

$\frac{-x_2}{-16} = \frac{x_3}{-16} = a$

i.e, $-x_2 = x_3 = a$

x_2=-1a, x_3= 1a

Therefore,

x_1=0a

x_2=-1a

x_3=1a

Hence the eigenvector v_2 *for* λ_2= 4

$$v_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

The orthonormal set of eigenvector v_2 *is :*

$$\begin{pmatrix} \frac{0}{\left[\sqrt{(0^2+-1^2+1^2)}\right]} \\ \frac{-1}{\left[\sqrt{(0^2+-1^2+1^2)}\right]} \\ \frac{1}{\left[\sqrt{(0^2+-1^2+1^2)}\right]} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

For λ_3 = 0

(A- λ_3 *I)* v = (A- 0*I) * v:

$$\begin{pmatrix} 8-0 & 4 & 4 \\ 4 & 4-0 & 0 \\ 4 & 0 & 4-0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 & 4 & 4 \\ 4 & 4 & 0 \\ 4 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which can be written as:

$$\frac{x_1}{[(4*0)-(4*4)]} = \frac{-x_2}{[(8*0)-(4*4)]} = \frac{x_3}{[(8*4)-(4*4)]} = a$$

$$\frac{x_1}{-16} = \frac{-x_2}{-16} = \frac{x_3}{16} = a$$

i.e, which can also be written as:

$$\frac{x_1}{-1} = \frac{x_2}{1} = \frac{x_3}{1} = a$$

Therefore,

x_1=-1a

x_2=1a

x_3=1a

Hence the eigenvector v_3 *for* λ_3 = 0

$$v_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

The orthonormal set of eigenvector v_3 *is :*

$$\begin{pmatrix} \frac{-1}{\left[\sqrt{(-1^2+1^2+1^2)}\right]} \\ \frac{1}{\left[\sqrt{(-1^2+1^2+1^2)}\right]} \\ \frac{1}{\left[\sqrt{(-1^2+1^2+1^2)}\right]} \end{pmatrix} = \begin{pmatrix} \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

- **STEP 3: Form matrix U of 3 x 3:**

matrix U= $\begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}$

- **STEP 4:Form matrix $\mathbf{V^T}$:**

Matrix V is the orthonormal set of eigenvectors of A^TA

A^TA= $\begin{pmatrix} 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} . \begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$

The eigenvalues of A^TA*:*

The characteristic equation for the above matrix is:

$|A^TA- \lambda I |= 0$

$|A^TA - \lambda I |= \begin{pmatrix} 8-\lambda & 4 \\ 4 & 8-\lambda \end{pmatrix} = (8 - \lambda)^2- (4 * 4)$

$=(64 - 16\lambda+ \lambda^2-16)$

$=(\lambda^2+16\lambda+48)$

$=(\lambda- 4).(\lambda- 12)=0$

Therefor, the eigenvalues of A^TA *are :*

$\lambda_1= 12$, $\lambda_2= 4$

Now we will get the singular values of A^TA*:*

$\sigma 1= 2\sqrt{3}$

$\sigma_2= 2$

The eigenvector v of A^TA*:*

$A^TA*v=\lambda*v$

Then:

$A^TA*v-\lambda*v=(A^TA-\lambda*I)*v=0$

For $\lambda_1= 12$*,*

the eigenvector v1 for λ_1*= 12,*

(A^T*-* λ_1 **I)* v = (*A^T*- 12*I) * v*

$$\begin{pmatrix} 8-12 & 4 \\ 4 & 8-12 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence we get the equations,

*- 4**x_1*+ 4**x_2*=0*

*4**x_1*- 4**x_2*=0*

*4*x_1*=4*x_2

i.e,

x_1*=* x_2*=a*

Then,

x_1*=1a,* x_2*=1a*

Hence the eigenvector v_1 *for* λ_1*= 12*

$$v_1=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

The orthonormal set of eigenvector v_1 *is :*

For λ_2*= 4,*

the eigenvector v2 for λ_2*= 4,*

(A^T*-* λ_2 ** I) * v = (*A^T*- 4 * I) * v*

$$\begin{pmatrix} 8-4 & 4 \\ 4 & 8-4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence we get the equations,

*4**x_1*+ 4**x_2*=0*

*4**x_1*+ 4**x_2*=0*

*4*x_1*=-4*x_2

i.e,

$x_1= -x_2$*=a*

Then,

x_1*= -1a,* x_2*= 1a*

Hence the eigenvector v_2*for* λ_1*= 12*

$$v_2=\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

The orthonormal set of eigenvector v_2 is :

Form matrix V of 2 X 2:

$$\text{matrix } V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Hence matrix V^T of 2 X 2:

$$\text{matrix } V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

- **STEP 5:Form matrix W of 2 X 2:**

  W is an n X n diagonal matrix of the singular values which are the square roots of the eigenvalues of A^TA.

  $$\text{matrix } W = \begin{pmatrix} 2\sqrt{3} & 0 \\ 0 & 2 \end{pmatrix}$$

- **STEP 6:SVD Composition of 3 X 2 matrix A:**

Therefore, the SVD composition is given below:

$$\boldsymbol{A = UWV^T} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} 2\sqrt{3} & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Linear Discriminant Analysis

Linear Discriminant Analysis or Normal Discriminant Function Analysis is a powerful algorithm that can be used to determine the best separation between two or more classes. With LDA, you can quickly and easily identify which class a particular data point belongs to. This makes LDA a key tool for solving classification problems.

Linear Discriminant Analysis, or LDA, is a machine learning algorithm that is used to find the Linear Discriminant function that best classifies or discriminates or separates two classes of data points. LDA is a supervised learning algorithm, which means that it requires a labeled training set of data points in order to learn the Linear Discriminant function. Once the Linear Discriminant function has been learned, it can then be used to predict the class label of new data points. LDA is similar to PCA (principal component analysis) in the sense that LDA reduces the dimensions. However, the main purpose of LDA is to find the line (or plane) that best separates data points belonging to different classes. The key idea behind LDA is that the decision boundary should be chosen such that it maximizes the distance between the means of the two classes while simultaneously minimizing the variance within each class's data or within-class scatter.

It is considered a pre-processing step for modeling differences in ML and applications of pattern classification. Whenever there is a requirement to separate two or more classes having multiple features efficiently, the LDA model is the most common technique to solve such classification problems.

The following are some of the benefits of using LDA:

- LDA is used for classification problems.
- LDA is a powerful tool for dimensionality reduction.
- LDA is not susceptible to the "curse of dimensionality" like many other machine learning algorithms.

LDA working and the steps involved in the process:

- The first step is to calculate the means and standard deviation of each feature.
- Within class scatter matrix and between class scatter matrix is calculated.
- These matrices are then used to calculate the eigenvectors and eigenvalues.
- LDA chooses the k eigenvectors with the largest eigenvalues to form a transformation matrix.
- LDA uses this transformation matrix to transform the data into a new space with k dimensions.
- Once the transformation matrix transforms the data into new space with k dimensions, LDA can then be used for classification or dimensionality reduction.

The following are some examples of how LDA can be used in practice:

- LDA can be used for classification, such as classifying emails as spam or not spam.
- LDA can be used for dimensionality reduction, such as reducing the number of features in a dataset.
- LDA can be used to find the most important features in a dataset.

Curvature	Diameter	Quality Control Result
2.95	6.63	Passed
2.53	7.79	Passed
3.57	5.65	Passed
3.16	5.47	Passed
2.58	4.46	Not Passed
2.16	6.22	Not Passed
3.27	3.52	Not Passed

Table 1:

Example for Linear Discriminant Analysis:

A Factory "ABC" produces very expensive and high-quality chip rings their qualities are measured in terms of curvature and diameter. The result of quality control by experts is given in Table 1.

As a consultant to the factory, you get a task to set up the criteria for automatic quality control. Then, the manager of the factory also wants to test your criteria on new types of chip rings that even human experts are argued to each other. The new chip rings have a curvature of 2.81 and a diameter of 5.46.

Solutions: x = features (or independent variables) of all data. Each row (denoted by *) represents one object; each column stands for one feature.

y=group of the object (or dependent variable) of all data. Each row represents one object and it has only one column.

In our example, $$\text{x}=\begin{bmatrix} 2.95 & 6.63 \\ 2.53 & 7.79 \\ 3.57 & 5.65 \\ 3.16 & 5.47 \\ 2.58 & 4.46 \\ 2.16 & 6.22 \\ 3.27 & 3.52 \end{bmatrix} \quad and \quad y = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{bmatrix}$$

$\text{x}_k \ = data\ for\ row\ k$

For example, x_3=[3.577 5.65], x_7=[3.27 3.52]

g =number of groups in y. In our example, $g = 2$

x_i =features data for the group i . Each row represents one object;each column stands for one feature. We separate x into several groups based on the number of category in y.

$C_1 X_1$=(x_1,x_2)=[(2.95,6.63),(2.53,7.79),(3.57,5.65),(3.16,5.47)]

C_2=X_2=(x_1,x_2)=[(2.58,4.46),(2.16,6.22),(3.27,3.52)]

C_1= Class 1, C_2= Class 2. (from y)

μ_i = mean of features in group, which is average of xi

$$\mu_1 = [\tfrac{(2.95+2.53+3.57+3.16)}{4}, \tfrac{(6.63+7.79+5.65+5.47)}{4}] = \mu_1 = [3.05 \quad 6.38]$$

$$\mu_2 = [\tfrac{(2.58+2.16+3.27)}{3}, \tfrac{(4.46+6.22+3.52)}{3}] = \mu_2 = [2.67 \quad 4.73]$$

μ =global mean vector, that is mean of the whole data set.

In this example, $\mu = [2.88 \quad 5.676]$

STEP 1: Compute within class scatter matrix ($\mathbf{S}_w$) :

$$\mathrm{S}_w = S_1 + S_2$$
$$S_1 : Co-variance\ of\ C_1$$
$$S_2 : Co-variance\ of\ C_2$$
$$S_1 = \sum (X_1 - \mu_1) \ . \ (X_1 - \mu_1)^T$$
$$S_2 = \sum (X_2 - \mu_2) \ . \ (X_2 - \mu_2)^T$$

$$\mathrm{X}_1 - \mu_1 = \begin{bmatrix} -0.1 & -0.52 & 0.52 & 0.11 \\ 2.5 & 1.41 & -0.73 & -0.91 \end{bmatrix}$$

$Seperately\ ,\ calculate\ (X_1 - \mu_1)\ .\ (X_1 - \mu_1)^T$

$$First\ matrix = \begin{bmatrix} -0.1 \\ 0.25 \end{bmatrix} . \begin{bmatrix} -0.1 & 0.25 \end{bmatrix} = \begin{bmatrix} 0.01 & -0.025 \\ -0.025 & 0.0625 \end{bmatrix}$$
$$Second\ matrix = \begin{bmatrix} -0.52 \\ 1.41 \end{bmatrix} . \begin{bmatrix} -0.52 & 1.41 \end{bmatrix} = \begin{bmatrix} 0.2704 & -0.7332 \\ -0.7332 & 1.9881 \end{bmatrix}$$
$$Third\ matrix = \begin{bmatrix} 0.52 \\ -0.73 \end{bmatrix} . \begin{bmatrix} 0.52 & -0.73 \end{bmatrix} = \begin{bmatrix} 0.2704 & -0.3796 \\ -0.3796 & 0.5329 \end{bmatrix}$$
$$Fourth\ matrix = \begin{bmatrix} 0.11 \\ -0.91 \end{bmatrix} . \begin{bmatrix} 0.11 & -0.91 \end{bmatrix} = \begin{bmatrix} 0.012 & -0.100 \\ -0.100 & 0.828 \end{bmatrix}$$

Adding all these 5 matrices together we get ,

$$\mathrm{S}_1 = \begin{bmatrix} 0.5628/4 & -1.2378/4 \\ -1.2378/4 & 3.4115/4 \end{bmatrix}$$

$$=> S_1 = \begin{bmatrix} 0.1407 & -0.30945 \\ -0.30945 & 0.852875 \end{bmatrix}$$

similarly ,

$$S_2 = \sum (X_2 - \mu_2) \cdot (X_2 - \mu_2)^T$$

$$\mathrm{X}_2 - \mu_2 = \begin{bmatrix} -0.09 & -0.51 & 0.6 \\ -0.27 & 1.49 & -1.21 \end{bmatrix}$$

Seperately , *calculate* $(X_2 - \mu_2) \cdot (X_2 - \mu_2)^T$

$$First\ matrix = \begin{bmatrix} -0.09 \\ -0.27 \end{bmatrix} . \begin{bmatrix} -0.09 & -0.27 \end{bmatrix} = \begin{bmatrix} 0.008 & 0.024 \\ 0.024 & 0.073 \end{bmatrix}$$

$$Second\ matrix = \begin{bmatrix} -0.51 \\ 1.49 \end{bmatrix} . \begin{bmatrix} -0.51 & 1.49 \end{bmatrix} = \begin{bmatrix} 0.260 & -0.760 \\ -0.760 & 2.220 \end{bmatrix}$$

$$Third\ matrix = \begin{bmatrix} 0.6 \\ -1.21 \end{bmatrix} . \begin{bmatrix} 0.6 & -1.21 \end{bmatrix} = \begin{bmatrix} 0.36 & -0.726 \\ -0.726 & 1.464 \end{bmatrix}$$

Adding all these 5 matrices together we get ,

$$\mathrm{S}_2 = \begin{bmatrix} 0.628/3 & -1.462/3 \\ -1.462/3 & 3.757/3 \end{bmatrix}$$

$$=> S_2 = \begin{bmatrix} 0.2093 & -0.4873 \\ -0.4873 & 1.2523 \end{bmatrix}$$

$$Within\ class\ scatter\ matrix\ ,\ S_w = S_1 + S_2 = \begin{bmatrix} 0.1407 & -0.30945 \\ -0.30945 & 0.852875 \end{bmatrix} + \begin{bmatrix} 0.2093 & -0.4873 \\ -0.4873 & 1.2523 \end{bmatrix}$$

$$S_W = \begin{bmatrix} 0.35 & -0.797 \\ -0.797 & 2.10 \end{bmatrix}$$

STEP 2: Compute Between class scatter matrix ($\mathbf{S}_B$) :

$$\mathrm{S}_B = (\mu_1 - \mu_2) \cdot (\mu_1 - \mu_2)^T$$

$$(\mu_1 - \mu_2) = \begin{bmatrix} 3.05 - 2.67 \\ 6.38 - 4.73 \end{bmatrix}$$

$$Therefore,\ (\mu_1 - \mu_2) = \begin{bmatrix} 0.38 \\ 1.65 \end{bmatrix}$$

$$(\mu_1 - \mu_2) \cdot (\mu_1 - \mu_2)^T = \begin{bmatrix} 0.38 \\ 1.65 \end{bmatrix} . \begin{bmatrix} 0.38 & 1.65 \end{bmatrix}$$

$$(\mu_1 - \mu_2) \cdot (\mu_1 - \mu_2)^T = \begin{bmatrix} 0.144 & 0.627 \\ 0.627 & 2.723 \end{bmatrix}$$

STEP 3: Compute LDA Projection
$(\mathbf{W})^*$:

$$(\mathrm{S}_w)^{-1} = 1/[(0.1407) . (0.852875) - (-0.30945) . (-0.30945)] \begin{bmatrix} 0.852875 & -0.30945 \\ -0.30945 & 0.1407 \end{bmatrix}$$

$$=> \ (\mathrm{S}_w)^{-1} = \begin{bmatrix} 35.184 & 12.766 \\ 12.766 & 5.804 \end{bmatrix}$$

$$W^* = (S_w)^{-1} \, . \, (\mu_1 - \mu_2) = \begin{bmatrix} 35.184 & 12.766 \\ 12.766 & 5.804 \end{bmatrix} . \begin{bmatrix} 0.38 \\ 1.65 \end{bmatrix} = \begin{bmatrix} 34.434 \\ 14.428 \end{bmatrix}$$