PROXIMITY/DISTANCE MEASURES-PROBLEM SOLVING

Dr. Umarani Jayaraman
Assistant Professor

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY,
DESIGN AND MANUFACTURING,
KANCHEEPURAM

Topic:

- Solved Problems

Mahalanobis distance

- We know the quadratic form distance

$$\boldsymbol{d_Q(x,y)} = \sqrt{(\boldsymbol{x}-\boldsymbol{y})^{\boldsymbol{T}}\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{y})}$$

- Replace A in quadratic form distance by inverse of covariance matrix Σ to get **Mahalanobis distance**

$$\boldsymbol{d_M(x,y)} = \sqrt{(\boldsymbol{x}-\boldsymbol{y})^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{y})}$$

Example 1:Mahalanobis Distance

- Suppose you have data for five people, and each person vector has a height, score on some test, and an age.

X	Y	Z
Height	Score	Age
64	580	29
66	570	33
68	590	37
69	660	46
73	600	55

mean

68	600	40

Steps to compute mahanalobis distance

- Find the mean vector
- Find the covariance metrics
- Find the inverse of covariance metrics
- Then multiply it with (X-Y), let say C_1
- Then $(X-Y)^T. C_1$

Example 1: Mahalanobis Distance

- The mean of the data is $(68.0, 600.0, 40.0)$.
- Now suppose you want to know how far another person, $v = (66, 640, 44)$, is from this data.
- It turns out the Mahalanobis Distance is 5.33 (no units).

Example 1:Mahalanobis Distance

- $Z = \begin{bmatrix} -4 & -20 & -11 \\ -2 & -30 & -7 \\ 0 & -10 & -3 \\ 1 & 60 & 6 \\ 5 & 0 & 15 \end{bmatrix}$
- Find the covariance matrix
- $\Sigma = \frac{1}{N-1} Z^T Z$

Example 1:Mahalanobis Distance

$$\Sigma = \frac{1}{4}\begin{bmatrix} -4 & -2 & 0 & 1 & 5 \\ -20 & -30 & -10 & 60 & 0 \\ -11 & -7 & -3 & 6 & 15 \end{bmatrix}\begin{bmatrix} -4 & -20 & -11 \\ -2 & -30 & -7 \\ 0 & -10 & -3 \\ 1 & 60 & 6 \\ 5 & 0 & 15 \end{bmatrix}$$

$$= \frac{1}{4}\begin{bmatrix} 46 & 200 & 139 \\ 200 & 5000 & 820 \\ 139 & 820 & 440 \end{bmatrix}$$

Example 1:Mahalanobis Distance

- $\Sigma = \begin{bmatrix} 11.50 & 50 & 34.75 \\ 50 & 1250 & 205 \\ 34.75 & 205 & 110 \end{bmatrix}$
- $\Sigma^{-1} = \begin{bmatrix} 3.6885 & 0.0627 & -1.2821 \\ 0.0627 & 0.0022 & -0.0240 \\ -1.2821 & -0.0240 & 0.4588 \end{bmatrix}$
- $D\left(h_q, h_t\right) = \left(h_q - h_t\right)^T A\left(h_q - h_t\right) - - - Eqn(1)$
- Here $A = \Sigma^{-1}$

Example 1:Mahalanobis Distance

- $Eqn(1)$ becomes

$$[-2 \quad 40 \quad 4]\begin{bmatrix} 3.6885 & 0.0627 & -1.2821 \\ 0.0627 & 0.0022 & -0.0240 \\ -1.2821 & -0.0240 & 0.4588 \end{bmatrix}\begin{bmatrix} -2 \\ 40 \\ 4 \end{bmatrix}$$

$$=[-9.9964 \quad -0.1325 \quad 3.4413]\begin{bmatrix} -2 \\ 40 \\ 4 \end{bmatrix}$$

$$= \sqrt{28.4573}$$

$$= 5.33$$

Quadratic form distance

$$d_Q(x,y) = \sqrt{(x-y)^T A (x-y)}$$

- For example $A_{ij}=1-c_{ij}/c_{max}$ for color histograms
- c_{ij} is bin-to-bin distance and c_{max} the maximum distance
- Note
  - If A is an identity matrix, then Euclidean
  - If A is a diagonal matric, then weighted Euclidean
  - If A is a inverse of covariance matrix, then Mahalanobis distance
  - **Is it a Metric? Yes , if A is positive definite**

Example 2: Quadratic form distance

- Given the histogram of a pure red image

  $h_q = [1, 0, 0]^T$

  and a pure orange image:

  $h_t = [0, 1, 0]^T$

- The transformation matrix $A = \begin{bmatrix} 1 & 0.9 & 0 \\ 0.9 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- Find the quadratic form distance and the Euclidean distance between them.

Example 2: Quadratic form distance

- $D(h_q, h_t) = (h_q - h_t)^T A (h_q - h_t) --- Eqn(1)$
- $h_q = [1\ 0\ \ 0]^T$
  $= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
- $h_t = [0\ \ 1\ \ \ 0]^T$
  $= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Example 2: Quadratic form distance

- $h_q - h_t = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$
- $(h_q - h_t)^T = [1 \quad -1 \quad 0] --- Eqn(2)$
- Substituting Eqn (2) and value of A in Eqn (1),
- $D\left(h_q, h_t\right) = [1 \ -1 \ 0] \begin{bmatrix} 1 & 0.9 & 0 \\ 0.9 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

Example 2: Quadratic form distance

- $$D(h_q, h_t) = [1 \ -1 \ \ 0] \begin{bmatrix} 1 & 0.9 & 0 \\ 0.9 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$
$$= [0.1 \ -0.1 \ \ 0] \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$
$$= (0.1) + (0.1) + 0$$
$$= 0.2$$

Quadratic form distance

$$d_Q(x,y) = \sqrt{(x-y)^T A (x-y)}$$

- For example $A_{ij}=1-c_{ij}/c_{max}$ for color histograms
- c_{ij} is bin-to-bin distance and c_{max} the maximum distance
- Note
  - If A is an identity matrix, then Euclidean
  - If A is a diagonal matric, then weighted Euclidean
  - If A is a inverse of covariance matrix, then Mahalanobis distance
  - **Is it a Metric? Yes , if A is positive definite**

Example 3: Euclidean Distance

- $D^2(h_q, h_t) = (h_q - h_t)^2$

$$= \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$= [1 \;+ 1 + \; 0]$$

$$= 2$$

- $D(h_q, h_t) \;= \sqrt{2}$

Summary-Metric Similarity Measures

- Solved Problems
  - Mahalanobis distance
  - Quadratic form distance
  - Euclidean distance

THANK YOU