Homework – 6

Recitation Exercises

Rutul Mehta (A20476293)

Exercise 9.2.1

Exercise 9.2.1 : Three computers, A, B, and C, have the numerical features listed below:

Feature	A	B	C
Processor Speed	3.06	2.68	2.92
Disk Size	500	320	640
Main-Memory Size	6	4	6

We may imagine these values as defining a vector for each computer; for instance, A's vector is $[3.06, 500, 6]$. We can compute the cosine distance between any two of the vectors, but if we do not scale the components, then the disk size will dominate and make differences in the other components essentially invisible. Let us use 1 as the scale factor for processor speed, α for the disk size, and β for the main memory size.

(a) In terms of α and β, compute the cosines of the angles between the vectors for each pair of the three computers.

(b) What are the angles between the vectors if $\alpha = \beta = 1$?

(c) What are the angles between the vectors if $\alpha = 0.01$ and $\beta = 0.5$?

! (d) One fair way of selecting scale factors is to make each inversely proportional to the average value in its component. What would be the values of α and β, and what would be the angles between the vectors?

a)

- Vector A = $[3.06, 500\alpha, 6\beta]$
- Vector B = $[2.68, 320\alpha, 4\beta]$
- Vector C = $[6, 4\alpha, 6\beta]$

- Cosine angle between A and B:

$$cos\theta 1 = \frac{\big((3.06 * 1)(2.68 * 1)\big) + \big((500 * \alpha)(320 * \alpha)\big) + \big((6 * \beta)(4 * \beta)\big)}{\left(\sqrt{(3.06)^2 + (500\alpha)^2 + (6\beta)^2}\right) * \left(\sqrt{(2.68)^2 + (320\alpha)^2 + (4\beta)^2}\right)}$$

$$= \frac{8.2008 + (160000\alpha^2) + 24\beta^2}{\left(\sqrt{9.3636 + 250000\alpha^2 + 36\beta^2}\right) * \left(\sqrt{7.1824 + 102400\alpha^2 + 16\beta^2}\right)}$$

- Cosine angle between B and C

$$cos\theta 2 = \frac{7.8256 + (204800\alpha^2) + 24\beta^2}{\left(\sqrt{7.1824 + 102400\alpha^2 + 16\beta^2}\right) * \left(\sqrt{8.5264 + 409600\alpha^2 + 36\beta^2}\right)}$$

- Cosine angle between A and C

$$cos\theta 3 = \frac{8.9352 + (320000\alpha^2) + (36\beta^2)}{(\sqrt{9.3636 + 250000\alpha^2 + 36\beta^2})(\sqrt{8.5264 + 409600\alpha^2 + 36\beta^2})}$$

b)

For $\alpha = 1\ and\ \beta = 1$

- Cosine angle between A and B

$$cos\theta 1 = \frac{8.2008 + 160000 + 24}{(\sqrt{9.3636 + 250000 + 36}) * (\sqrt{7.1824 + 102400 + 16})} = 1$$

$$So, the\ angle\ between\ A\ and\ B, \theta 1 = 0$$

- Cosine angle between B and C

$$cos\theta 2 = \frac{7.8256 + 204800 + 24}{(\sqrt{7.1824 + 102400 + 16}) * (\sqrt{8.5264 + 409600 + 36})} = 1$$

$$So, the\ angle\ between\ B\ and\ C, \theta 2 = 0$$

- Cosine angle between A and C

$$cos\theta 3 = \frac{8.9352 + 320000 + 36}{(\sqrt{9.3636 + 250000 + 36}) * (\sqrt{8.5264 + 409600 + 36})} = 1$$

$$So, the\ angle\ between\ A\ and\ C, \theta 3 = 0$$

C)

For $\alpha = 0.01\ and\ \beta = 0.5$

When we substitute the value of $\alpha\ and\ \beta$ to the derived equation of a), we will get following values.

- Cosine angle between A and B

$$cos\theta 1 = 0.99$$

$$So, the\ angle\ between\ A\ and\ B, \theta 1 = 8.1096$$

❖ Cosine angle between B and C

$cos\theta2 = 0.96$

$So, the\ angle\ between\ B\ and\ C, \theta2 = 16.2602$

❖ Cosine angle between A and C

$cos\theta3 = 0.99$

$So, the\ angle\ between\ A\ and\ C, \theta3 = 8.1096$

d)

$\alpha\ and\ \beta\ inversely\ proportional\ to\ the\ average\ of\ their\ respective\ components.$

$$\alpha = \frac{3}{500 + 320 + 640} = \frac{3}{1460}$$

$$\beta = \frac{3}{6 + 4 + 6} = \frac{3}{16}$$

❖ Cosine angle between A and B:

$$cos\theta1 = \frac{8.2008 + \left(160000(\frac{3}{1460})^2\right) + 24(\frac{3}{16})^2}{\left(\sqrt{9.3636 + 250000(\frac{3}{1460})^2 + 36(\frac{3}{16})^2}\right) * \sqrt{7.1824 + 102400(\frac{3}{1460})^2 + 16(\frac{3}{16})^2}} = 0.99$$

$So, the\ angle\ between\ A\ and\ B, \theta1 = 8.1096$

Same way when we substitute the value of $\alpha\ and\ \beta$ we will get following values.

❖ Cosine angle between B and C

$cos\theta2 = 0.99$

$So, the\ angle\ between\ B\ and\ C, \theta2 = 8.1$

❖ Cosine angle between A and C

$cos\theta3 = 0.99$

$So, the\ angle\ between\ A\ and\ C, \theta3 = 8.1$

Exercise 9.2.3

Exercise 9.2.3: A certain user has rated the three computers of Exercise 9.2.1 as follows: A: 4 stars, B: 2 stars, C: 5 stars.

(a) Normalize the ratings for this user.

(b) Compute a user profile for the user, with components for processor speed, disk size, and main memory size, based on the data of Exercise 9.2.1.

a)

Average rating = $\frac{(4+2+5)}{3} = \frac{11}{3}$

Normalised rating for A = $4 - \frac{11}{3} = \frac{1}{3}$

Normalised rating for B = $2 - \frac{11}{3} = \frac{-5}{3}$

Normalised rating for C = $5 - \frac{11}{3} = \frac{4}{3}$

b)

Value of the Processor speed = $\left(3.06 * \frac{1}{3}\right) + \left(2.68 * \frac{-5}{3}\right) + \left(2.92 * \frac{4}{3}\right) = 0.4467$

Value of disk size = $\left(500 * \frac{1}{3}\right) + \left(320 * \frac{-5}{3}\right) + \left(640 * \frac{4}{3}\right) = 486.667$

Value for main memory size = $\left(6 * \frac{1}{3}\right) + \left(4 * \frac{-5}{3}\right) + \left(6 * \frac{4}{3}\right) = 3.3333$

Exercise 9.3.1

Exercise 9.3.1: Figure 9.8 is a utility matrix, representing the ratings, on a 1–5 star scale, of eight items, a through h, by three users A, B, and C. Compute the following from the data of this matrix.

(a) Treating the utility matrix as boolean, compute the Jaccard distance between each pair of users.

(b) Repeat Part (a), but use the cosine distance.

(c) Treat ratings of 3, 4, and 5 as 1 and 1, 2, and blank as 0. Compute the Jaccard distance between each pair of users.

(d) Repeat Part (c), but use the cosine distance.

(e) Normalize the matrix by subtracting from each nonblank entry the average value for its user.

(f) Using the normalized matrix from Part (e), compute the cosine distance between each pair of users.

	a	b	c	d	e	f	g	h
A	4	5		5	1		3	2
B		3	4	3	1	2	1	
C	2		1	3		4	5	3

Figure 9.8: A utility matrix for exercises

a)

Jaccard Distance (X,Y) = $\frac{features\ for\ which\ both\ users\ have\ rated}{Total\ features\ rated\ by\ both\ the\ users\ (count\ each\ features\ just\ once)}$

Jaccard distance (A, B) = $\frac{4}{8} = \frac{1}{2}$

Jaccard distance (B, C) = $\frac{4}{8} = \frac{1}{2}$

Jaccard distance (A, C) = $\frac{4}{8} = \frac{1}{2}$

b)

Cosine Distance (X, Y) = $1 - \frac{\sum Product\ of\ rating\ where\ feature\ match\ in\ X\ and\ Y}{|X|*|Y|}$

Cosine Distance (A, B) = $1 - \frac{(5*3)+(5*3)+(1*1)+(3*1)}{(\sqrt{4^2+5^2+5^2+1^2+3^2+2^2})*\sqrt{3^2+4^2+3^2+1^2+2^2+1^2}} = 0.399$

Cosine Distance (B, C) = $1 - \frac{(4*1)+(3*3)+(2*4)+(1*5)}{(\sqrt{2^2+1^2+3^2+4^2+5^2+3^2})*\sqrt{3^2+4^2+3^2+1^2+2^2+1^2}} = 0.486$

Cosine Distance (A, C) = $1 - \frac{(4*2)+(5*3)+(3*5)+(2*3)}{(\sqrt{4^2+5^2+5^2+3^2+2^2+1^2})*\sqrt{2^2+1^2+3^2+4^2+5^2+3^2}} = 0.385$

c)

	a	b	c	d	e	f	g	h
A	1	1	0	1	0	0	1	0
B	0	1	1	1	0	0	0	0
C	0	0	0	1	0	1	1	1

$Jaccard\ Distance\ (X, Y) = 1 - Jaccard\ Similarity(X, Y)$

Where,

$Jaccard\ Similarity(X, Y) = \frac{a}{a+b+c}$

a = the number of attributes that equal 1 for both objects X and Y

b = the number of attributes that equal 0 for objects X but equal 1 for object Y

c = the number of attributes that equal 1 for objects X but equal 0 for object Y

d = the number of attributes that equal 0 for both objects X and Y

Jaccard distance (A, B) = $1 - \frac{2}{5} = 0.6$

Jaccard distance (B, C) = $1 - \frac{1}{6} = 0.833$

Jaccard distance (A, C) = $1 - \frac{2}{6} = 0.667$

d)

Cosine Distance (A, B) = $1-\frac{(1*1)+(1*1)}{(\sqrt{1^2+2^2+2^2+1^2+2^2})*\sqrt{1^2+1^2+1^2+1^2+2^2+1^2}} = 0.4227$

Cosine Distance (B, C) = $1-\frac{(1*1)+(1*1)+(2*2)+(1*2)}{(\sqrt{1^2+1^2+1^2+1^2+2^2+1^2})*\sqrt{2^2+1^2+1^2+2^2+2^2+1^2}} = 0.7114$

Cosine Distance (A, C) = $1-\frac{(1*2)+(2*1)+(1*2)+(2*1)}{(\sqrt{1^2+2^2+2^2+1^2+1^2+2^2})*\sqrt{2^2+1^2+1^2+2^2+2^2+1^2}} = 0.5$

e)

	a	b	c	d	e	f	g	h
A	2/3	5/3		5/3	-7/3		-1/3	-4/3
B		2/3	5/3	2/3	-4/3	-1/3	-4/3	
C	-1		-2	0		1	2	0

f)

$cosine\ Distance(A, B)$

$$= 1-\frac{\left(\frac{5}{3}*\frac{2}{3}\right)+\left(\frac{5}{3}*\frac{2}{3}\right)+\left(-\frac{7}{3}*-\frac{4}{3}\right)+(-\frac{1}{3}*-\frac{4}{3})}{\sqrt{\left(\frac{2}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(\frac{2}{3}\right)^2+\left(-\frac{4}{3}\right)^2+\left(-\frac{1}{3}\right)^2+\left(-\frac{4}{3}\right)^2}\sqrt{\left(\frac{2}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(-\frac{7}{3}\right)^2+\left(-\frac{1}{3}\right)^2+\left(-\frac{4}{3}\right)^2}}$$

$= 0.416$

$cosine\ Distance(B, C)$

$$= 1-\frac{\left(\frac{5}{3}*\frac{-2}{1}\right)+\left(\frac{2}{3}*0\right)+\left(-\frac{1}{3}*1\right)+(-\frac{4}{3}*2)}{\sqrt{\left(\frac{2}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(\frac{2}{3}\right)^2+\left(-\frac{4}{3}\right)^2+\left(-\frac{1}{3}\right)^2+\left(-\frac{4}{3}\right)^2}\sqrt{(-1)^2+(-2)^2+(0)^2+(1)^2+(2)^2+(0)^2}}$$

$= 1.739$

$cosine\ Distance(A, B)$

$$= 1-\frac{\left(\frac{2}{3}*-1\right)+\left(\frac{5}{3}*0\right)+\left(-\frac{1}{3}*2\right)+(0*-\frac{4}{3})}{\sqrt{(-1)^2+(-2)^2+(0)^2+(1)^2+(2)^2+(0)^2}\sqrt{\left(\frac{2}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(\frac{5}{3}\right)^2+\left(-\frac{7}{3}\right)^2+\left(-\frac{1}{3}\right)^2+\left(-\frac{4}{3}\right)^2}}$$

$= 1.115$

Exercise 9.4.1

Exercise 9.4.1: Starting with the decomposition of Fig. 9.10, we may choose any of the 20 entries in U or V to optimize first. Perform this first optimization step assuming we choose: (a) u_{32} (b) v_{41}.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

Figure 9.10: Matrices U and V with all entries 1

a)

u_{32}

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & x \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 1+x & 1+x & 1+x & 1+x & 1+x \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

The contribution to the sum of squares from the third row is

$(x-1)^2 + (x-2)^2 + x^2 + (x-3)^2$

To find minimum number value of this expression we will be differentiating and equating to 0,

$2 * \big((x-1) + (x-2) + x + (x-3)\big) = 0$

So, we get x=1.5

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1.5 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \\ 2.5 & 2.5 & 2.5 & 2.5 & 2.5 \\ 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

b)

$\underline{V_{14}\ (V_{41}\ is\ not\ exist)}$

$$\begin{bmatrix}1 & 1\\1 & 1\\1 & 1\\1 & 1\\1 & 1\end{bmatrix} \times \begin{bmatrix}1 & 1 & 1 & y & 1\\1 & 1 & 1 & 1 & 1\end{bmatrix} = \begin{bmatrix}2 & 2 & 2 & y+1 & 2\\2 & 2 & 2 & y+1 & 2\\2 & 2 & 2 & y+1 & 2\\2 & 2 & 2 & y+1 & 2\\2 & 2 & 2 & y+1 & 2\end{bmatrix}$$

The contribution to the sum of squares from the third row is

$(y-3)^2 + (y-3)^2 + y^2 + (y-2)^2 + (y-3)^2$

To find minimum number value of this expression, we will be differentiating and equating to 0,

$2 * \left((y-3) + (y-3) + y + (y-2) + (y-3)\right) = 0$

So, we get y=2.2

$$\begin{bmatrix}1 & 1\\1 & 1\\1 & 1\\1 & 1\\1 & 1\end{bmatrix} \times \begin{bmatrix}1 & 1 & 1 & 2.2 & 1\\1 & 1 & 1 & 1 & 1\end{bmatrix} = \begin{bmatrix}2 & 2 & 2 & 3.2 & 2\\2 & 2 & 2 & 3.2 & 2\\2 & 2 & 2 & 3.2 & 2\\2 & 2 & 2 & 3.2 & 2\\2 & 2 & 2 & 3.2 & 2\end{bmatrix}$$