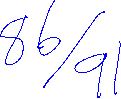
**Stephen Kim**

**MAT220 – Furno**



**4/13/20**

**Homework 2**

**Part A:**

1. **3.1 – Distance between Boolean vectors. Suppose that x and y are Boolean n-vectors, which means that each of their entries is either 0 or 1. What is their distance ||x – y||?**

x = (x1, x2, …, xn) for xn ϵ {0, 1}, y = (y1, y2, …, yn) yn ϵ {0, 1}

||x – y||2 = (x1 – y1)2 + (x2 – y2)2 + (x3 – y3)2 + (x4 – y4)2 + (x5 – y5)2

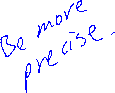
||x – y|| = √[(x1 – y1)2 + (x2 – y2)2 + … + (xn – yn)2]

Example: n = 5, x = (1, 0, 1, 0, 1), y = (1, 1, 1, 1, 0)

||x – y|| = √[(1 – 1)2 + (0 – 1)2 + (1 – 1)2 + (0 + 1)2 + (1 – 1)2]

= √(0 + 1 + 0 + 1 + 0) = √2

The maximum distance is √n if one vector is all 0’s and the other vector is all 1’s.



The minimum distance is 0 if both vectors are all 0’s or both vectors are all 1’s.

The distance (dist) between Boolean vectors is 0 < dist(x, y) < √n.

1. **3.4 – Norm identities. Verify that the following identities hold for any two vectors a and b of the same size.**
2. **(a + b)T(a – b) = ||a||2 - ||b||2**

= aTa – bTb



= aTa + bTa – aTb – bTb

= (a + b)T(a – b)

a = (a1, a2, …, an), b = (b1, b2, …, bn) where ai, bi ϵ R

a + b = [(a1 + b1), (a2 + b2), …, (an + bn)]

a – b = [(a1 – b1), (a2 – b2), …, (an – bn)]

(a + b)T(a – b) = [(a1 + b1), (a2 + b2), …, (an + bn)] \* [(a1 – b1), (a2 – b2), …, (an – bn)]

= (a1 + b1) (a1 – b1) + (a2 + b2) (a2 – b2) + … + (an + bn) (an – bn)

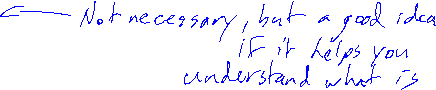
= (a12 – b12) + (a22 – b22) + … + (an2 – bn2)



= (a12 + a22 + … + an2) – (b12 + b22 + … + bn2) = ||a||2 - ||b||2

Example: a = (4, 1, 9), b = (8, 4, 7)

(a + b)T(a – b) = [(4 + 8), (1 + 4), (9 + 7)] \* [(4 – 8), (1 – 4), (9 – 7)]



= (12, 5, 16) \* (-4, -3, 2)

= -48 + (-15) + 32 = -31



||a||2 - ||b||2 = [√[(4)2 + (1)2 + (9)2]]2 – [√[(8)2 + (4)2 + (7)2]]2

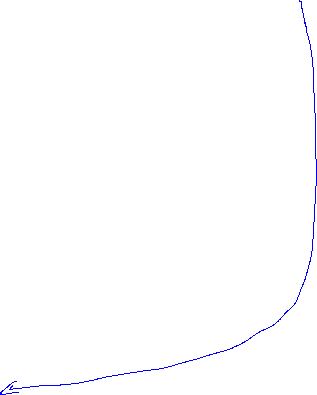
= (16 + 1 + 81) – (64 + 16 + 49)

= 98 – 129 = -31

1. **||a + b||2 + ||a – b||2 = 2(||a||2 + ||b||2). This is called the parallelogram law.**

||a + b||2 + ||a – b||2 = 2(||a||2 + ||b||2)

= (a1 + b1)2 + (a2 + b2)2 + … + (an + bn)2 + (a1 – b1)2 + (a2 – b2)2 + … + (an – bn)2



= [(a1 + b1)2 + (a1 – b1)2] + [(a2 + b2)2 + (a2 – b2)2] + … + [(an + bn)2 + (an – bn)2]

= [a12 + 2a1b1 + b12 + a12 – 2a1b1 + b12] + [a22 + 2a2b2 + b22 + a22 – 2a2b2 + b22] + … +

[an2 + 2anbn + bn2 + an2 – 2anbn + bn2]

= [2a12 + 2b12] + [2a22 + 2b22] + … + [2an2 + 2bn2]

= 2(a12 + a22 + … + an2) + 2(b12 + b22 + … + bn2)

= 2(||a||2 + ||b||2)



Example: a = (3, 1, 2), b = (6, 3, 0), n = 3

||a + b||2 + ||a – b||2 = ||(3 + 6), (1 + 3), (2 + 0)||2 + ||(3 – 6), (1 – 3), (2 – 0)||2

= ||(9, 4, 2)||2 + ||(-3, -2, 2)||2

= (81 + 16 + 4) + (9 + 4 + 4)



= 118

2(||a||2 + ||b||2) = 2(||(3, 1, 2)||2 + ||(6, 3, 0)||2)

= 2[(9 +1 + 4) + (36 + 9 + 0)]

= 118

1. **3.10 – Nearest neighbor document. Consider the 5 Wikipedia pages in table 3.1 on page 51. What is the nearest neighbor of (the word count histogram vector of) ‘Veterans Day’ among the others? Does the answer make sense?**

The nearest neighbor is Memorial Day at with a distance of 0.095. This makes sense because Veterans Day and Memorial Day are both holidays and have more in common than the other three subjects of Academy awards, Golden Globe Awards, and the Superbowl. Therefore it is not surprising if the documents have the most words in common, resulting in the lowest calculation of distance.

1. **3.11 – Neighboring electronic health records. Let x1, …, xN be n-vectors that contain n features extracted from a set of N electronic health records (EHRs), for a population of N patients. (The features might involve patient attributes and current and past symptoms, diagnoses, test results, hospitalizations, procedures, and medications.) Briefly describe in words a practical use for identifying the 10 nearest neighbors of a given EHR (as measured by their associated feature vectors), among the other EHRs.**

This would find the 10 people who have the closest medical histories out of a collection of N patients. One possible practical use for this could be to assist insurance companies in finding a price range for a customer by comparing someone to other people with very similar medical histories and setting a price based on other people who probably have similar health risks. Another practical use could be for doctors to see if a treatment may work for a patient by seeing if one of those 10 nearest neighbors may have undergone a similar procedure in the past. The results may have a high probability of being similar. Another not as practical use would be to find other people that are very similar to start a support group for some health concern.



1. **3.16 – Effect of scaling and offset on average and standard deviation. Suppose x is an n-vector and α and β are scalars.**
2. **Show that avg(αx + β1) = αavg(x) + β**

avg(αx + β1) = avg(αx) + avg(β1)

= avg(α) \* avg(x) + avg(β) \* avg(1)

= α \* avg(x) + β \* 1

= αavg(x) + β

Example:

x = (1, 2, 3, 4), α = 2, β = 3

avg(αx + β1) = avg(2(1, 2, 3, 4) + 3(1))

= avg(2(1, 2, 3, 4)) + avg(1(3))

= avg(2, 4, 6, 8) + avg(3)

= 5 + 3 = 8

αavg(x) + β = 2(avg(1, 2, 3, 4)) + 3

= 2(5/2) + 3

= 8

1. **Show that std(αx + β1) = |α|std(x)**

std(αx + β1) = std(αx) // property of adding a constant

= |α|std(x) // property of multiplying by a scalar

std(αx + β1) = rms(αx + β1 – (α(avg(x) + β)1) // RMS prediction error

= rms(αx – α(avg(x)1)

= |α|rms(x – avg(x)1)

= |α|std(x)

Example:

x = (1, 2, 3), α = 2 β = 6, n = 3

std(αx + β1) = std(2(1, 2, 3)

= std(2, 4, 6)

= √[((2 – 4)2 + (4 – 4)2 + (6 – 2)2)/3]

= √[4 + 0 + 4)/3]

= √(8/3)

|α|std(x) = |2|std(1, 2, 3)

= 2 \* √[((1 – 2)2 + (2 – 2)2 + (3 – 2)2)/3]

= 2 \* √[2 + 0 + 2)/3]

= √(8/3)

1. **3.19 – Norm of sum. Use the formulas (3.1) and (3.6) to show the following:**

****

3.1 = norm of sum

3.6 = norm of sum via angles

1. **a ꓕ b if and only if ||a + b|| = √(||a||2 + ||b||2)**

||a + b|| = √(||a||2 + 2aTb + ||b||2)

= √(||a||2 + 2||x||||y||cosθ + ||b||2)

= √(||a||2 + 0 + ||b||2)

= √(||a||2 + ||b||2)

a ꓕ bwhen 2||a||||b||cosθ = 2aTb = 0

1. **Nonzero vectors a and b make an acute angle if and only if ||a + b|| > √(||a||2 + ||b||2)**

2aTb > 0

||a||2 + ||b||2 + 2aTb > ||a||2 + ||b||2

||a + b||2 > ||a||2 + ||b||2

||a + b|| > √(||a||2 + ||b||2)

1. **Nonzero vectors a and b make an obtuse angle if and only if ||a + b|| < √(||a||2 + ||b||2)**

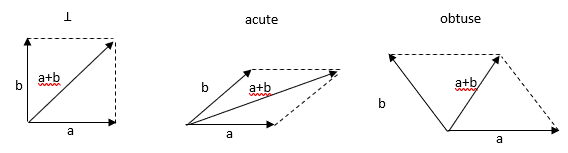
2aTb < 0

||a||2 + ||b||2 + 2aTb < ||a||2 + ||b||2

||a + b||2 < ||a||2 + ||b||2

||a + b|| < √(||a||2 + ||b||2)

1. **Draw a picture illustrating each case in 2-D**

****

1. **4.3 – Linear separation in 2-way partitioning. Clustering a collection of vectors into k = 2 groups is called 2-way partitioning, since we are partitioning the vectors into 2 groups, with index sets G1 and G2. Suppose we run k-means, with k = 2, on the n-vectors x1, …, xN. Show that there is a nonzero vector w and a scalar v that satisfy**

**wTxi + v ≥ 0 for i ϵ G1, wTxi + v ≤ 0 for i ϵ G2.**

**In other words, the affine function f(x) = wTx + v is greater than or equal to zero on the first group, and less than or equal to zero on the second group. This is called linear separation of the two groups (although affine separation would be more accurate). Hint. Consider the function ||x – z1||2 - ||x – z2||2, where z1 and z2 are the group representatives.**

||x – z1||2 - ||x – z2||2

If ||x – z2||2 ≥ 0 🡪 ||x – z1|| ≥ ||x – z2|| 🡪 x in G1

If ||x – z2||2 ≤ 0 🡪 ||x – z1|| ≤ ||x – z2|| 🡪 x in G2

||x – z1||2 = ||x||2 + ||z1||2 – 2xTz1

- ||x – z1||2 = - ||x||2 - ||z1||2 + 2xTz1

||z1||2 - ||z2||2 + 2xT(z2 – z1) ≤ 0

0 ≤ ||z2||2 - ||z1||2 + 2(z2 – z1)Tx

v = ||z2||2 - ||z1||2

w = 2(z1 – z2)



**Part B:**

1. **Compute the following for *u* = (2*,*4), *v* = (−3*,*1), and *w* = (−3*,*−5). Show your work.**
2. **(1/2)u 🡪** [(1/2)(2), (1/2)(4)] = (1, 2)



1. **– w 🡪**  – (-3, -5) = (-(-3), -(-5)) = (3, 5)



1. **– 3v 🡪**  – 3(-3, 1) = ((-3)(-3), (-3)(1)) = (9, -3)
2. **u + v 🡪** (2, 4) + (-3, 1) = ((2 + (-3)), (4 + 1)) = (-1, 5)



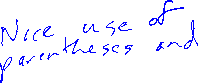
1. **v – w 🡪** (-3, 1) – (-3, -5) = ((-3 – (-3)), (1 – (-5))) = (0, 6)



1. **2u + 3w 🡪** 2(2, 4) + 3(-3, -5) = (4, 8) + (-9, -15) = (-5, -7)
2. **5u – 4v 🡪** 5(2, 4) – 4(-3, 1) = (10, 20) – (-12, 4) = (22, 16)
3. **-2u + 2v – 3w 🡪** -2(2, 4) + 2(-3, 1) – 3(-3, -5) = (-4, -8) + (-6, 2) – (-9, -15) = (-1, 9)
4. **u \* v 🡪** (2, 4)(-3, 1) = (2 \* (-3)) + (4 \* 1) = -6 + 4 = -2



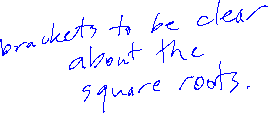
1. **||w|| 🡪** √[(-3)2 + (-5)2] = √(9 + 25) = √(36) = 6



1. **avg(u) 🡪** (2 + 4)/2 = 6/2 = 3



1. **rms(u) 🡪** √[(22 + 42)/2] = √[(4 + 16)/2] = √(20/2) = √10



1. **std(u) 🡪** √[((2 – 3)2 + (4 – 3)2)/2] = √[((-1)2 + 12)/2] = √(2/2) = √1 = 1



1. **avg(v) 🡪** (-3 + 1)/2 = (-2)/2 = -1
2. **rms(v) 🡪** √[((-3)2 + 12)/2] = √[(9 + 1)/2] = √(10/2) = √5
3. **std(v) 🡪** √[(((-3) – (-1))2 + (1 – (-1))2)/2] = √[(22 + 22)/2] = √(8/2) = √4 = 2
4. **avg(w) 🡪 [**(-3) + (-5)]/2 = (-8)/2 = -4
5. **rms(w) 🡪** √[((-3)2 + (-5)2)/2] = √[(9 + 25)/2] = √(26/2) = √13
6. **std(w) 🡪** √[(((-3) – (-4))2 + ((-5) – (-4))2)/2] = √[(12 + (-1)2)/2] = √(2/2) = √1 = 1
7. **avg(u, v, w) 🡪** [(2 + (-3) + (-3)), (4 + 1 + (-5))]/3 = (-4, 0)/3 = (-4/3, 0)
8. **rms(u, v, w) 🡪** √[(22 + (-3)2 + (-3)2), (42 + 12 + (-5)2)/3] = √[[(4 + 9 + 9),(16 + 1+ 25)]/3]

= √((22, 42)/3) = (√(22/3), √14)

1. **std(u, v, w) 🡪** √[(((2 – (-4/3))2 + ((-3) – (-4/3))2 + ((-3) – (-4/3))2), (((4 – (0))2 + ((1) – (0))2 + ((-5) – (0))2)/3]

= √[(((10/3)2 + (-5/3)2 + (-5/3)2), (42 + 12 + (-5)2))/3] = √[((100/3 + 25/3 + 25/3), (16 + 1 + 25))/3]

**=** √[(150/3, 42)/3] = (√(50/3), √14)

1. **Express v as a linear combination of the standard basis vectors *e*1 and *e*2.**

v = β1e1 + β2e2



1. **Which pair of *u*, *v*, *w* are closest to each other?**

dist(u, v) 🡪 ||u – v|| … ((2 – (-3)), (4 – 1)) = (5, 3)

= √(52 + 32) = √(25 + 9) = √36 = 6

dist(u, w) 🡪 ||u – w|| … ((2 – (-3)), (4 – (-5))) = (5, 9)

= √(52 + 92) = √(25 + 81) = √106 = 10.3

dist(v, w) 🡪 ||v – w|| … (((-3) – (-3)), (1 – (-5))) = (0, 6)

= √(02 + 62) = √(0 + 36) = √36 = 6

The distance between u and v, and the distance between v and w are both the shortest.

1. **Find the angle between *u* and *w* in radians and in degrees.**

θ = arccos(uTw/(||u||\*||w||))

uTw = (2, 4)(-3, -5) = (-6) + (-20) = -26

||u|| = √(22 + 42) = √(4 + 16) = √20

||w|| = √((-3)2 + (-5)2) = √(9 + 25) = √34

arccos(-26/(√20√34) = arccos(-0.9971) = 3.0648 radians = 175.6°

180 – 175.6 = 4.4°

1. **Find the centroid of *u*, *v*, and *w.***

[(2 + (-3) + (-3)), (4 + 1 + (-5))]/3 = (-4, 0)/3 = (-4/3, 0)

***u* = (2*,*4), *v* = (−3*,*1), and *w* = (−3*,*−5).**

1. **Compute the following for *u* = (6*,*2*,*4), *v* = (2*,*0 − 3), and *w* = (−3*,*1*,*2). Show your work.**
2. **(1/2)u 🡪** [(1/2)(6), (1/2)(2), (1/2)(4)] = (3, 1, 2)
3. **– w 🡪**  – (-3, 1, 2) = (-(-3), -(1), -(2)) = (3, -1, -2)



1. **– 3v 🡪**  – 3(2, 0, (-3)) = ((-3)(2), (-3)(0), (-3)(-3)) = (-6, 0, 9)
2. **u + v 🡪** (6, 2, 4) + (2, 0, -3) = ((6 + 2), (2 + 0), (4 + (-3)) = (8, 2, 1)
3. **v – w 🡪** (2, 0, -3) – (-3, 1, 2) = ((2 – (-3)), (0 – 1), ((-3) – 2)) = (5, -1, -5)
4. **2u + 3w 🡪** 2(6, 2, 4) + 3(-3, 1, 2) = (12, 4, 8) + (-9, 3, 6)) = (3, 7, 14)
5. **5u – 4v 🡪** 5(6, 2, 4) – 4(2, 0, -3) = (30, 10, 20) – (8, 0, (-12)) = (22, 10, 32)
6. **-2u + 2v – 3w 🡪** -2(6, 2, 4) + 2(2, 0, -3) – 3(-3, 1, 2) = (-12, -4, -8) + (4, 0, 6) – (-9, 3, 6)



= (((-12) + 4 – (-9)), ((-4) + 0 – 3), ((-8) + 6 – 6)) = (1, -7, -8)



1. **u \* v 🡪** (6, 2, 4)(2, 0, -3) = (6 \* 2) + (2 \* 0) + (4 \* (-3)) = 12 + 0 + (-12) = 0



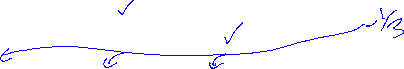
1. **||w|| 🡪** √[(-3)2 + (1)2 + (2)2] = √(9 + 25 + 4) = √(40) = 2√(10)



1. **avg(u) 🡪** (6 + 2 + 4)/3 = 12/3 = 4
2. **rms(u) 🡪** √[(62 + 22 + 42)/3] = √[(36 + 4 + 16)/3] = √(56/3)
3. **std(u) 🡪** √[((6 – 4)2 + (2 – 4)2 + (4 – 4)2)/3] = √[(22 + (-2)2 + 02)/3] = √((4 + 4 + 0)/3) = √(8/3)
4. **avg(v) 🡪** (2 + 0 + (-3))/3 = (-1)/3 = -1/3



1. **rms(v) 🡪** √[(22 + 02 + (-3)2)/3] = √[(4 + 0 + 9)/3] = √(13/3)



1. **std(v) 🡪** √[((2 - √(13/3))2 + ((0) – (√(13/3)))2 + ((-3) – √(13/3))2)/3]

= √[[(2 - √(13/3))(2 - √(13/3)) + ((√(13/3)) (√(13/3))) + ((-3) – √(13/3))((-3) – √(13/3))]/3]

= √[((4 – 4√(13/3) + (13/3)) + (13/3) + (9 + 6√(13/3) + (13/3)))/3]

= √[(((25/3) - 4√(13/3)) +(13/3) + ((40/3) + 6√(13/3)))/3]



√[((26 – 10√(13/3))/3] = √(78 + 10√39)/3

1. **avg(w) 🡪** ((-3) + (1) + (2))/3 = (0)/3 = 0
2. **rms(w) 🡪** √[((-3)2 + 12 + 22)/3] = √[(9 + 1+ 4)/3] = √(14/3)
3. **std(w) 🡪** √[(((-3) – (√(14/3)))2 + ((1) – (√(14/3)))2 + ((2) – (√(14/3)))2)/3]

= √[[((-3) – √(14/3))((-3) – √(14/3)) + (1 – √(14/3))(1 – √(14/3))) + (2 – √(14/3))(2 – √(14/3))]/3]

= √[(((41/3) + 6√(14/3)) + ((17/3) – 2√(14/3)) + ((26/3) - 4√(14/3)))/3]

= √[((41/3) + (17/3) + (26/3) + 6√(14/3) – 2√(14/3) – 4√(14/3))/3]

= √((84/3)/3) = √(28/3)

1. **avg(u, v, w) 🡪** [((6 + 2 + 4)/3), ((2 + 0 + (-3))/3), (4 + (-3) + 2)/3] = ((12/3), (-1)/3, 3/3) = (4, -1/3, 1)
2. **rms(u, v, w) 🡪** √[(62 + 22 + (-3)2), (22 + 02 + 12)/3, (42 + (-3)2 + 22)/3]

= √[[(36 + 4 + 9), (4 + 0 + 1), (16 + 9 + 4)]/3]

= √((49, 5, 29)/3) = (√(49/3), √(5/3), √(29/3))

1. **std(u, v, w) 🡪** √[(((6 – 4)2 + (2 – 4)2 + ((-3) – 4)2), (((2 – (-1/3))2 + ((0) – (-1/3))2 + (1 – (-1/3))2,

(4 – 1)2 + ((-3) – 1)2 + (2 – 1)2),)/3]

= √[(((2)2 + (-2)2 + (-7)2), ((7/3)2 + (1/3)2 + (4/3)2), (3)2 + (-4)2 + (1)2)/3]

= √[((4 + 4 + 49), (49/9 + 1/9 + 16/9), (9 + 16 + 1))/3]

**=** √[(57, 22/3, 26)/3] = (√(57), √(22/3), √(26))

1. **Express v as a linear combination of the standard basis vectors *e*1 and *e*2 and *e*3.**

v = β1e1 + β2e2+ β3e3

1. **Which pair of *u*, *v*, *w* are closest to each other?**

dist(u, v) 🡪 ||u – v|| … ((6 – 2), (2 – 0), (4 – (-3)) = (4, 2, 7)

= √(42 + 22 + 72) = √(16 + 4 + 49) = √69 = 8.31



dist(u, w) 🡪 ||u – w||… ((6 – (-3)), (2 – 1), (4 – 2) = (9, 1, 2)



= √(92 + 12 + 22) = √(81 + 1 + 4) = √86 = 9.27

dist(v, w) 🡪 ||v – w||… ((2 – (-3)), (0 – 1), ((-3) – 2) = (5, -1, -5)



= √(52 + (-1)2 + (-5)2) = √(25 + 1 + 25) = √51 = 7.14



The distance between v and w is the shortest at 7.14 and the two points are the closest.

1. **Find the angle between *u* and *w* in radians and in degrees.**

θ = arccos(uTw/(||u||\*||w||))

uTw = (6, 2, 4)(-3, 1, 2) = (-18) + 2 + 8 = -8

||u|| = √(62 + 22 + 42) = √(36 + 4 + 16) = √56

||w|| = √((-3)2 + 12 + 22) = √(9 + 1 + 4) = √14

arccos(-8/(√56√14) = arccos(-0.2857) = 1.8605 radians = 106.6°



180 – 106.6 = 73.4°



1. **Find the centroid of *u*, *v*, and *w.***

[((6 + 2 + 4)/3), ((2 + 0 + (-3))/3), (4 + (-3) + 2)/3] = ((12/3), (-1)/3, 3/3) = (4, -1/3, 1)



***u* = (6*,*2*,*4), *v* = (2*,*0 − 3), and *w* = (−3*,*1*,*2)**



1. **Find a vector that is orthogonal to both *a* = (1*,*2*,*3*,*4) and *b* = (1*,*−1*,*2*,*−2).**

A zero vector is orthogonal to any vector. 🡪 (0, 0, 0, 0)

a & b = linearly independent

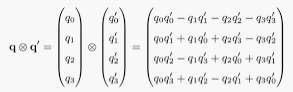
c ꓕ a & c ꓕ b …. c = (w, x, y, z)

cTa = c (1, 2, 3, 4) = 0

cTb = c (1, -1, 2, -2) = 0

1w + 2x + 3y + 4z = 0

1w – 1x + 2y – 2z = 0



(1, 2, 3, 4) (1, -1, 2, -2) = ((1\*1) – (2\*-1) – (3\*2) – (4\*-2), (2\*1) + (1\*-1) – (4\*2) + (3\*-2),

(3\*1) + (4\*-1) + (1\*2) – (2\*-2), (4\*1) – (3\*-1) + (2\*2) + (1\*-2))

= ((1 – (-2) – 6 – (-8)), (2 + (-1) – 8 + (-6)), (3 + (-4) + 2 – (-4)), (4 – (-3) + 4 + (-2)))

= (5, -13, 5, 9)

