CSCI 347 Data Mining

More about distance measures and similarity measures

One-Hot encoding

- Appropriate for nominal data.
- Create a binary column for each possible value in the category.
  - Ex: X_3 is replaced by columns $X_{3A}, X_{3B}, X_{3C}, X_{3D}, X_{3F}, X_{3G}$
  - If category A is present in the row, then the observed value for the new column X_{3A} is 1, and all the other new columns are 0.

$$D = \begin{array}{c|cccc} & X_1 & X_2 & X_3 & X_4 \\ x_1 & 0.2 & 23 & A & 5.7 \\ x_2 & 0.4 & 1 & B & 5.4 \\ x_3 & 1.8 & 0.5 & D & 5.2 \\ x_4 & 5.6 & 50 & C & 5.1 \\ x_5 & -0.5 & 34 & F & 5.3 \\ x_6 & 0.4 & 19 & G & 5.4 \\ x_7 & 1.1 & 11 & A & 5.5 \end{array} \longrightarrow D = \begin{array}{c|ccccccccc} & X_1 & X_2 & X_{3A} & X_{3B} & X_{3C} & X_{3D} & X_{3F} & X_{3G} & X_4 \\ x_1 & 0.2 & 23 & 1 & 0 & 0 & 0 & 0 & 0 & 5.7 \\ x_2 & 0.4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 5.4 \\ x_3 & 1.8 & 0.5 & 0 & 0 & 0 & 1 & 0 & 0 & 5.2 \\ x_4 & 5.6 & 50 & 0 & 0 & 1 & 0 & 0 & 0 & 5.1 \\ x_5 & -0.5 & 34 & 0 & 0 & 0 & 0 & 1 & 0 & 5.3 \\ x_6 & 0.4 & 19 & 0 & 0 & 0 & 0 & 0 & 1 & 5.4 \\ x_7 & 1.1 & 11 & 1 & 0 & 0 & 0 & 0 & 0 & 5.5 \end{array}$$

One-Hot encoding

$$D = \begin{array}{c|ccccccccc} & X_1 & X_2 & X_{3A} & X_{3B} & X_{3C} & X_{3D} & X_{3E} & X_{3G} & X_4 \\ x_1 & 0.2 & 23 & 1 & 0 & 0 & 0 & 0 & 0 & 5.7 \\ x_2 & 0.4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 5.4 \\ x_3 & 1.8 & 0.5 & 0 & 0 & 0 & 1 & 0 & 0 & 5.2 \\ x_4 & 5.6 & 50 & 0 & 0 & 1 & 0 & 0 & 0 & 5.1 \\ x_5 & -0.5 & 34 & 0 & 0 & 0 & 0 & 1 & 0 & 5.3 \\ x_6 & 0.4 & 19 & 0 & 0 & 0 & 0 & 0 & 1 & 5.4 \\ x_7 & 1.1 & 11 & 1 & 0 & 0 & 0 & 0 & 0 & 5.5 \end{array}$$

$$\|x_1 - x_2\|_2 = \sqrt{\sum_{k=1}^{9} (x_{1k} - x_{2k})^2} =$$

$$= \sqrt{(0.2-0.4)^2 + (23-1)^2 + (5.7-5.4)^2 + (1-0)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$
$$= 22.05$$

Dot product and binary data

- If we take the one-hot encoding data, then the dot product of two vectors (projection of one-hot encoded data) is the number of matching categorical values.

$$D = \begin{array}{c|c} & X_4 \\ \hline x_1 & A \\ x_2 & B \\ x_3 & C \\ x_4 & A \\ x_5 & B \\ x_6 & C \\ x_7 & C \end{array} \quad \longrightarrow \quad \begin{array}{c|ccc} & X_{4A} & X_{4B} & X_{4C} \\ \hline x_1 & 1 & 0 & 0 \\ x_2 & 0 & 0 & 1 \\ x_3 & 0 & 0 & 1 \\ x_4 & 1 & 0 & 0 \\ x_5 & 0 & 1 & 0 \\ x_6 & 0 & 0 & 1 \\ x_7 & 0 & 0 & 1 \end{array}$$

$$x_1^T x_2 = 1(0) + 0(0) + 0(1) = 0$$

$$s(x_1, x_2) = 0 \; because \; x_{14} = A \; and \; x_{24} = B$$

Dot product and binary data

- If we take the one-hot encoding data, then the dot product of two vectors (projection of one-hot encoded data) is the number of matching categorical values.

$D =$

	X_4
x_1	A
x_2	B
x_3	C
x_4	A
x_5	B
x_6	C
x_7	C

	X_{4A}	X_{4B}	X_{4C}
x_1	1	0	0
x_2	0	0	1
x_3	0	0	1
x_4	1	0	0
x_5	0	1	0
x_6	0	0	1
x_7	0	0	1

$$x_1^T x_4 = 1(1) + 0(0) + 0(0) = 1$$

$$s(x_1, x_2) = 0 \; because \; x_{14} = A \; and \; x_{24} = B \; for \; X_4$$

$$s(x_1, x_4) = 1 \; because \; x_{14} = A \; and \; x_{44} = A \; for \; attribute \; X_4$$

Dot product and binary data

- For one-hot encoded data, the number categorical attributes d is the squared 2-norm each point:

$$D = \begin{array}{c|c} & X_4 \\ \hline x_1 & A \\ x_2 & B \\ x_3 & C \\ x_4 & A \\ x_5 & B \\ x_6 & C \\ x_7 & C \end{array} \quad \longrightarrow \quad \begin{array}{c|ccc} & X_{4A} & X_{4B} & X_{4C} \\ \hline x_1 & 1 & 0 & 0 \\ x_2 & 0 & 0 & 1 \\ x_3 & 0 & 0 & 1 \\ x_4 & 1 & 0 & 0 \\ x_5 & 0 & 1 & 0 \\ x_6 & 0 & 0 & 1 \\ x_7 & 0 & 0 & 1 \end{array}$$

$$d = \|x_i\|^2 = x_i^T x_i$$

$$\|x_1\|^2 = 1^2 + 0^2 + 0^2 = 1 = d$$

Dot product and binary data

- For one-hot encoded data, the number categorical attributes d is the squared 2-norm each point:

$$
D = \begin{array}{c|c}
 & X_4 \\
\hline
x_1 & A \\
x_2 & B \\
x_3 & C \\
x_4 & A \\
x_5 & B \\
x_6 & C \\
x_7 & C
\end{array}
\longrightarrow
\begin{array}{c|ccc}
 & X_{4A} & X_{4B} & X_{4C} \\
\hline
x_1 & 1 & 0 & 0 \\
x_2 & 0 & 0 & 1 \\
x_3 & 0 & 0 & 1 \\
x_4 & 1 & 0 & 0 \\
x_5 & 0 & 1 & 0 \\
x_6 & 0 & 0 & 1 \\
x_7 & 0 & 0 & 1
\end{array}
$$

$$d = \|x_i\|^2 = x_i^T x_i$$

$$\|x_2\|^2 = 0^2 + 0^2 + 1^2 = 1 \; = d$$

Dot product and binary data

- For one-hot encoded data, the number categorical attributes d is the squared 2-norm each point:

$D =$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	0	1	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0

$$d = \|x_i\|^2 = x_i^T x_i$$

$$\|x_2\|^2 = 0^2 + 0^2 + 1^2 + 0^2 + 1^2 = 2 = d$$

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

Recall that $XOR \oplus$

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

Recall that $XOR \oplus$

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

$$\delta_H(x_1, x_2) = (1 \oplus 0) + (0 \oplus 1) + (0 \oplus 0) + (1 \oplus 0) + (0 \oplus 1) \ = 4$$

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

$\delta_H(x_3, x_8) = ?$

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk})$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

$$\delta_H(x_3, x_8) = (0 \oplus 0) + (0 \oplus 0) + (1 \oplus 1) + (0 \oplus 0) + (1 \oplus 1)$$

Hamming distance

- Hamming Distance: number of mismatches between two vectors

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk}) = d - s$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

$$\delta_H(x_3, x_8) = (0 \oplus 0) + (0 \oplus 0) + (1 \oplus 1) + (0 \oplus 0) + (1 \oplus 1) = 0 = 2 - 2$$

Side note

$$\delta_H(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} \oplus x_{jk}) = d \; - s$$

$$\|x_i - x_j\|_2 = \sqrt{x_i^T x_i \; - 2x_i x_j + x_j^T x_j} = \sqrt{2(d - s)}$$

Jaccard similarity

- The Jaccard Coefficient is a commonly used similarity measure between two categorical points.
- It is defined as the ratio of the number of matching values to the number of distinct values that appear in x_i or x_j ,across the d attributes.

$$J(x_i, x_j) = \frac{s}{2(d-s)+s} = \frac{s}{2d-s}$$

Jaccard similarity

$$J(x_i, x_j) = \frac{s}{2(d-s)+s} = \frac{s}{2d-s}$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

Jaccard similarity

$$J(x_i, x_j) = \frac{s}{2(d-s)+s} = \frac{s}{2d-s}$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

$$d = x_2^T x_2 = x_6^T x_6 = 2$$
$$s = x_2 \cdot x_6 = 1$$
$$J(x_2, x_6) = \frac{1}{2+1} = \frac{1}{3}$$

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

Jaccard similarity

$$J(x_i, x_j) = \frac{s}{2(d-s)+s} = \frac{s}{2d-s}$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

Let's calculate the $J(x_6, x_8)$

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

Jaccard similarity

$$J(x_i, x_j) = \frac{s}{2(d-s)+s} = \frac{s}{2d-s}$$

	X_1	X_2
x_1	A	H
x_2	B	L
x_3	C	L
x_4	A	L
x_5	B	H
x_6	C	L
x_7	C	H
x_8	C	L

$$d = x_6^T x_6 = x_8^T x_8 = 2$$
$$s = x_6 \cdot x_8 = 2$$
$$J(x_2, x_6) = \frac{2}{2(2-2)+2} = \frac{2}{2} = 1$$

	X_{1A}	X_{1B}	X_{1C}	X_{2H}	X_{2L}
x_1	1	0	0	1	0
x_2	0	1	0	0	1
x_3	0	0	1	0	1
x_4	1	0	0	0	1
x_5	0	1	0	1	0
x_6	0	0	1	0	1
x_7	0	0	1	1	0
x_8	0	0	1	0	1

Gower distance

- Gower distance is a similarity or dissimilarity measure designed to handle mixed data types, including numerical, categorical, ordinal, and binary data.

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$Categorical$$

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$$Numerical$$

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

$Range(k)$: $Difference\ between\ maximum\ value\ of\ the\ k^{th}$ $feature\ and\ the\ minimum\ value$

Example: Gower distance

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$D = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{array} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \\ 0.2 & 23 & A & 5.7 \\ 0.4 & 1 & B & 5.4 \\ 1.8 & 0.5 & D & 5.2 \\ 5.6 & 50 & C & 5.1 \\ -0.5 & 34 & F & 5.3 \\ 0.4 & 19 & G & 5.4 \\ 1.1 & 11 & A & 5.5 \end{array}$$

$Numerical$

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

$Categorical$

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$Range(k){:}\ Difference\ between\ maximum\ value\ of\ the\ k^{th}\ feature\ and\ the\ minimum\ value$

Example: Gower distance

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$D = \begin{array}{c|cccc} & X_1 & X_2 & X_3 & X_4 \\ x_1 & 0.2 & 23 & A & 5.7 \\ x_2 & 0.4 & 1 & B & 5.4 \\ x_3 & 1.8 & 0.5 & D & 5.2 \\ x_4 & 5.6 & 50 & C & 5.1 \\ x_5 & -0.5 & 34 & F & 5.3 \\ x_6 & 0.4 & 19 & G & 5.4 \\ x_7 & 1.1 & 11 & A & 5.5 \end{array}$$

$$G(x_1, x_4) = \frac{1}{4} \sum_{k=1}^{4} dist_k(x_{1k}, x_{4k})$$

$Numerical$

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

$Categorical$

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$Range(k): Difference\ between\ maximum\ value\ of\ the\ k^{th}$
$feature\ and\ the\ minimum\ value$

Example: Gower distance

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$D = \begin{array}{c|cccc} & X_1 & X_2 & X_3 & X_4 \\ x_1 & 0.2 & 23 & A & 5.7 \\ x_2 & 0.4 & 1 & B & 5.4 \\ x_3 & 1.8 & 0.5 & D & 5.2 \\ x_4 & 5.6 & 50 & C & 5.1 \\ x_5 & -0.5 & 34 & F & 5.3 \\ x_6 & 0.4 & 19 & G & 5.4 \\ x_7 & 1.1 & 11 & A & 5.5 \end{array}$$

$$G(x_1, x_4) = \frac{1}{4} \sum_{k=1}^{4} dist_k(x_{1k}, x_{4k})$$

$$= \frac{1}{4} \cdot \left(\frac{|0.2 - 5.6|}{5.6 - (-0.5)} + \frac{|23 - 50|}{(50 - 0.5)} + 1 + \frac{|5.7 - 5.1|}{5.7 - 5.1} \right)$$

$$= \frac{1}{4} \left(\frac{5.2}{6.1} + \frac{27}{49.5} + 1 + \frac{0.6}{0.6} \right)$$

$$= 0.849$$

Numerical

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

Categorical

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$Range(k)$: *Difference between maximum value of the k^{th} feature and the minimum value*

Activity: Gower distance

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$G(x_1, x_7) = \frac{1}{4} \sum_{k=1}^{4} dist_k(x_{1k}, x_{7k})$$

$$D = \begin{array}{c|cccc} & X_1 & X_2 & X_3 & X_4 \\ x_1 & 0.2 & 23 & A & 5.7 \\ x_2 & 0.4 & 1 & B & 5.4 \\ x_3 & 1.8 & 0.5 & D & 5.2 \\ x_4 & 5.6 & 50 & C & 5.1 \\ x_5 & -0.5 & 34 & F & 5.3 \\ x_6 & 0.4 & 19 & G & 5.4 \\ x_7 & 1.1 & 11 & A & 5.5 \end{array}$$

Numerical

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

Categorical

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$Range(k)$: *Difference between maximum value of the* k^{th} *feature and the minimum value*

Activity: Gower distance

$$G(x_i, x_j) = \frac{1}{d} \cdot \sum_{k=1}^{d} dist_k(x_{ik}, x_{jk})$$

$$D = \begin{array}{ccccc} & X_1 & X_2 & X_3 & X_4 \\ x_1 & 0.2 & 23 & A & 5.7 \\ x_2 & 0.4 & 1 & B & 5.4 \\ x_3 & 1.8 & 0.5 & D & 5.2 \\ x_4 & 5.6 & 50 & C & 5.1 \\ x_5 & -0.5 & 34 & F & 5.3 \\ x_6 & 0.4 & 19 & G & 5.4 \\ x_7 & 1.1 & 11 & A & 5.5 \end{array}$$

$$G(x_1, x_7) = \frac{1}{4} \sum_{k=1}^{4} dist_k(x_1, x_7)$$

$$= \frac{1}{4} \cdot \left(\frac{|0.2 - 1.1|}{5.6 - (-0.5)} + \frac{|23 - 11|}{(50 - 0.5)} + 0 + \frac{|5.7 - 5.5|}{5.7 - 5.1} \right)$$

$$= \frac{1}{4} \left(\frac{0.9}{6.1} + \frac{12}{49.5} + 0 + \frac{0.2}{0.6} \right)$$

$$= 0.181$$

$Numerical$

$$dist(x_{ik}, x_{jk}) = \frac{|x_{ik} - x_{jk}|}{Range(k)}$$

$Categorical$

$$dist_k(x_{ik}, x_{jk}) = \begin{cases} 0 & if\ x_{ik} = x_{jk} \\ 1 & otherwise \end{cases}$$

$Range(k): Difference\ between\ maximum\ value\ of\ the\ k^{th}$
$feature\ and\ the\ minimum\ value$

Review Quiz

- Given any two vectors over a vector field, which represents the shortest distance between two vectors?
  A. Euclidean distance
  B. L_1 norm
  C. L_2 norm
  D. $A\ and\ C$
  E. L_∞ norm
  F. It depends on the vectors.

Review Quiz

- Given any two vectors over a vector field, which L_p norm gives the smallest value?
  A. Euclidean distance
  B. L_1 norm
  C. L_2 norm
  D. *A and C*
  E. L_∞ norm
  F. It depends on the vectors.

Review Quiz

- What is the dot product of the vectors below

$$a = \begin{pmatrix} 2 & -1 & 2 & 1 & 0 \end{pmatrix}$$
$$b = \begin{pmatrix} 5 & 0 & -1 & 3 & 9 \end{pmatrix}$$

A. 16

B. 20

C. 11

D. $\begin{pmatrix} 5 & 0 & 1 & 2 & 0 \end{pmatrix}$