Linear Data Chapter 7

Written by: Emily J. King

1. Assume $\vec{x}, \vec{y} \in \mathbb{R}^{1001}$ and you know the following:

$$\|\vec{x}\| = 4 \quad \|\vec{y}\| = 2.$$

For each of the following, either explicitly compute the value or explain why there is insufficient information.

   (a) Explicitly compute $\|-3\vec{y}\|$. Explain your answer.
   The norm is absolutely homogenous:

$$\|-3\vec{y}\| = |-3|\,\|\vec{y}\| = 3(2) = 6$$

   (b) Explicitly compute $\left\|\frac{\vec{x}}{\|\vec{x}\|}\right\|$. Explain your answer.
   METHOD ONE: The normalization of a vector is always a unit vector. Thus,

$$\left\|\frac{\vec{x}}{\|\vec{x}\|}\right\| = 1.$$

   METHOD TWO: The norm is absolutely homogenous:

$$\left\|\frac{\vec{x}}{\|\vec{x}\|}\right\| = \left|\frac{1}{\|\vec{x}\|}\right| \|\vec{x}\| = \frac{1}{\|\vec{x}\|}\|\vec{x}\| = 1$$

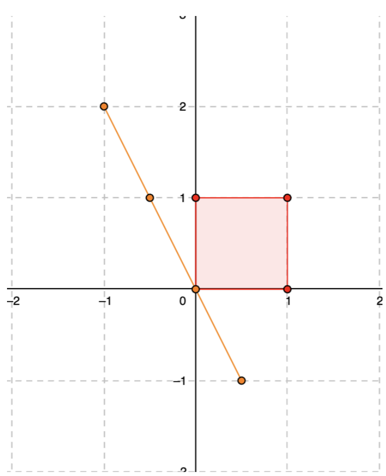
   (c) Is it possible for $\|\vec{x} + \vec{y}\| = 6$? Explain your answer.
   It is possible for $\|\vec{x} + \vec{y}\| = 6$ because the triangle inequality says that $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| = 4+2 = 6$. However, we can't know for sure if it is with the information given.

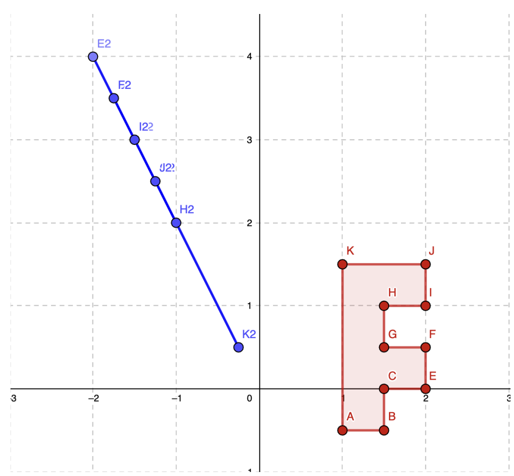
2. If $\vec{u}, \vec{v}, \vec{w}, \vec{z} \in \mathbb{R}^{300}$ are all word embeddings from GloVe and $\vec{u} - \vec{v} \approx \vec{w} - \vec{z}$, what might that mean semantically (i.e., about the meanings of the words)?
That the semantic relationship between $\vec{u}$ and $\vec{v}$ might be the same as between $\vec{w}$ and $\vec{z}$ (e.g., girl – boy and queen – king).

3. Using Python/Jupyter or Matlab/Matlab Live Script, perform the following:

- Define
$$\vec{a} = \begin{pmatrix} 12.3 \\ 0.56 \\ -1.7 \\ 0.34 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0.55 \\ 2.22 \\ -1.2 \\ 3.14 \end{pmatrix}$$
  You don't need to force them to be column vectors in Python. 1D arrays suffice.
  In Matlab:
  `a=[12.3; 0.56; -1.7; 0.34]` *(with or without semicolon)*
  `b = [0.55; 2.22; -1.2; 3.14]` *(with or without semicolon)*
  In Python:
  ```
  import numpy as np
  a=np.array([12.3, 0.56, -1.7, 0.34])
  b=np.array([0.55, 2.22, -1.2, 3.14])
  ```
- Compute the (Euclidean) norm of $\vec{a}$.
  Answer is approx. 12.4342
  In Matlab:
  `norm(a)`
  In Python:
  ```
  from numpy.linalg import norm
  norm(a)
  ```
- Compute the (Euclidean) distance between $\vec{a}$ and $\vec{b}$.
  Answer is approx. 12.2028
  In Matlab:
  `norm(a-b)` *or* `norm(b-a)`
  In Python:
  `norm(a-b)` *or* `norm(b-a)`

4. Consider the Geogebra screenshots below showing the effects of multiplication by a certain 2×2 matrix $\mathbf{B}$.





What must the determinant of $\mathbf{B}$ be? You must explain your answer to receive any credit.

Both the red square and the red 'F' are mapped to line segments (which have 0 *area). Thus, a dimension is "being squeezed out," meaning that* $\det(\mathbf{B}) = 0$*.*

If you are curious to know what $\mathbf{B}$ *is (not necessary to solve the problem), it was:*

$$\mathbf{B} = \begin{pmatrix} -1 & 0.5 \\ 2 & -1 \end{pmatrix}$$