CSE 3380 – Homework #6

Assigned: Thursday, March 8, 2018
Due: Thursday, March 22, 2018 at the end of class

Note the following about the homework:

1. You must show your work to receive credit.
2. If your submission has more than one page, staple the pages. If I have to staple it, the cost is 10 points.

Assignment:
Process

1. An **affine transformation** $T : \mathbb{R}^n \to \mathbb{R}^m$ has the form $A\vec{x} + \vec{b}$. Show that when $\vec{b} \neq \vec{0}$ the transformation is not linear. Note that this can be done algebraically using the conditions of linearity. If you choose to be more explicit, then use $\vec{x} = (x_1, x_2, \ldots, x_n)^T$ and $\vec{b} = (b_1, b_2, \ldots, b_m)^T$, not $\vec{x} = (x_1, x_2)$.
2. Show that T is a linear transformation by finding a matrix that implements the mapping.

   (a)
   $$T(x_1, x_2, x_3) = (x_1, 2x_1 + 3x_2, 4x_1 + 5x_2 + 6x_3)$$

   (b)
   $$T(x_1, x_2, x_3, x_4) = (2x_2 + 4x_4, x_1 + 3x_3)$$

   (c)
   $$T(x_1, x_2) = (x_1, x_2, x_1 + x_2, x_1 - x_2)$$

Theory

3. Sometimes we wish to apply multiple transformations to a vector. For example, given transformations $T_1(\vec{x}) = A\vec{x}$ and $T_2(\vec{x}) = B\vec{x}$, we may produce $T_2(T_1(\vec{x})) = BA\vec{x}$. Generally speaking, $BA\vec{x} \neq AB\vec{x}$.

   (a) Is this the case when the two transformations are

   $$T_1(\vec{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{(a rotation)} \qquad \text{and} \qquad T_2(\vec{x}) = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \quad \text{(scaling by } k\text{)}$$

   That is, does the order in which you apply the transformations matter? Show your work.

   (b) Speaking of diagonal matrices in general, if we have the $n \times n$ matrix A and the diagonal matrix D

   $$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad \text{and} \qquad D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & d_n \end{bmatrix}$$

   is $AD = DA$?

i. Consider the case where all of D's diagonal elements have the same value (i.e., $d_1 = d_2 = \cdots = d_n$)

ii. Consider the case where some of D's diagonal elements differ.

Show your work for the general $n \times n$ matrices given above, not specific matrices.

Applications

4. **(CS application: computer graphics)** Computer graphics is one of the easiest areas within computer science to understand how linear algebra is applicable since we can see the vectors in 2D or 3D space. Operations can be applied to the vectors (scaling, rotation, etc) via matrix multiplication.

   (a) Plot the following coordinates: $A = (2, 3)$, $B = (4, 3)$, and $C = (4, 6)$. Connect the points to form triangle ABC.

   (b) Apply a transformation to the points that rotates them $90\,^\circ$ counterclockwise. Note that this will not rotate the points in place, but will rotate them about the origin. Plot the transformed triangle, being clear which point is which.

   (c) Produce a single transformation matrix that performs the equivalent of the following individual transformations (in the order given):

      i. $T_1(\vec{x})$: Translate (i.e., move) the triangle such that point A is at the origin.

      ii. $T_2(\vec{x})$: Rotate $90\,^\circ$ counterclockwise.

      iii. $T_3(\vec{x})$: Translate the rotated triangle such that point A is at it's original point.

   Then apply this transformation matrix to the ORIGINAL three points of the triangle. Plot the transformed triangle, being clear which point is which.

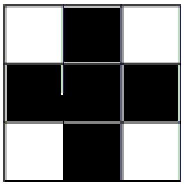
   Note that for step 4c, you will need to represent the points in homogeneous coordinates and use the appropriate 3×3 transformation matrices.

5. **(CS application: artificial intelligence)** A popular task in artificial intelligence is to classify things; popular choices are images and documents. Many methods for performing classification have been developed, but one of the simplest is called *k-means*.

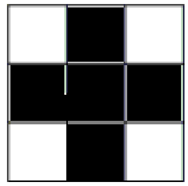
   Training samples from each of `k` classes of objects are chosen, with each object represented as a vector in $\mathbb{R}^n$ where n is the number of attributes describing each object (for example, pixels for images). Then the mean of the vectors within each class is found, which represents the center of these objects.

   To classify a new object, the distance from its vector to each mean is calculated and the object is classified as the class whose mean is closest.

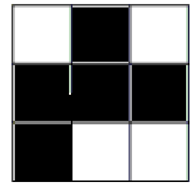
   Problem: We have two classes of images: plus symbols and X symbols. We also have two test images that we wish to classify. How do we represent an image as a vector? The value of each pixel (i.e., picture element) in the image is stored as an element of a vector. If the image has dimensions $p \times q$, then the vector will be a vector with pq elements. For this problem, each image has dimensions 3×3. When creating a vector of an image, the dark squares will have a value of 1 and the white squares will have a value of 0 in the vector.



(a) plus_1

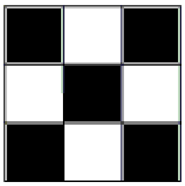


(b) plus_2

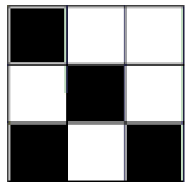


(c) plus_3

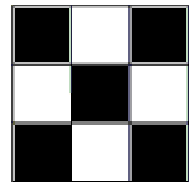
Figure 1: Training samples for class `plus`



(a) x_1



(b) x_2



(c) x_3

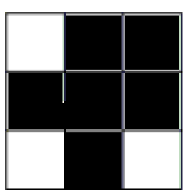
Figure 2: Training samples for class `X`

(a) Represent each training sample in Figure 1 as a 9×1 column vector and find the mean vector of this class. The mean vector will also be a 9×1 column vector and each of its elements is the average (i.e., arithmetic mean) of the elements of the training samples in that position.

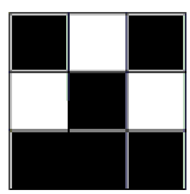
(b) Represent each training sample in Figure 2 as a 9×1 column vector and find the mean vector of this class.

(c) For each test sample in Figure 3, represent it as a 9×1 column vector. Then calculate the distance from this vector to the mean of each class and determine which class it should be assigned to. The distance measure to use is the Euclidean distance. If we have vectors $\vec{v} = (v_1, v_2, \ldots, v_n)^T$ and $\vec{w} = (w_1, w_2, \ldots, w_n)^T$, the Euclidean distance between them is

$$\mathrm{d}(\vec{v}, \vec{w}) = \sqrt{(v_1 - w_1)^2 + (v_2 - w_2)^2 + \cdots + (v_n - w_n)^2}$$



(a) t_1



(b) t_2

Figure 3: Test samples

Review Questions

These are not for credit, but instead are intended to test your understanding of the concepts. You should be able to answer these without simply regurgitating equations.

1. What is a linear transformation?

References

[Eld07] Lars Eldén. *Matrix Methods in Data Mining and Pattern Recognition*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2007.