EE M146 Discussion 1 Solution
Introduction to Machine Learning
Instructor: Lara Dolecek
TA: Zehui (Alex) Chen, Ruiyi (John) Wu

1. Linear Algebra Review

   (a) What are the four properties of a inner product i.e. $< x, y >$?
   Solution:
      i. *Positivity* $< x, x > \geq 0$ and $< x, x >= 0 \iff x = 0$.
      ii. *Symmetry* $< x, y >=< y, x >$
      iii. *Additivity* $< x + z, y >=< x, z > + < y, z >$
      iv. *Homogeneity* $< rx, y >= r < x, y > \ \forall r \in \mathbb{R}$

   (b) What are the three properties of a norm $f(x)$?
   Solution:
      i. *Positivity* $f(x) \ \geq 0$ and $f(x) = 0 \iff x = 0$.
      ii. *Homogeneity* $f(rx) = |r|f(x) \ \forall r \in \mathbb{R}$
      iii. *Triangle Inequality* $f(x + y) \leq f(x) + f(y)$

   (c) What is the L2-norm? L1-norm? L∞-norm?
   Solution:

$$\text{L2 norm} : f(x) = (\sum_i x_i^2)^{\frac{1}{2}}$$
$$\text{L1 norm} : f(x) = \sum_i |x_i|$$
$$\text{L}\infty \text{ norm} : f(x) = \max_i |x_i|$$

   (d) What does it mean to normalize a vector?
   Solution: It means to divide each element of the vector by its norm to ensure that the norm of the resultant vector is 1. Normalizing a vector makes it into a unit vector.

   (e) What does it mean for two vectors to be orthogonal?
   Solution: There are many right answers to this. One is that two vector are orthogonal when their inner product is zero. Another answer is that two vectors are orthogonal when they are separated by exactly 90 degrees.

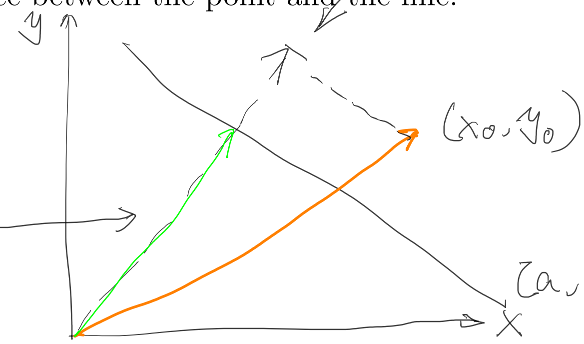
   (f) Given two vectors **w** and **v** that are not orthogonal, find two orthogonal vectors that span the same space as **w** and **v**.
   Solution:

$$\mathbf{x}_1 = \mathbf{w}$$
$$\mathbf{x}_2 = \mathbf{v} - \mathbf{v}^T\mathbf{w} \cdot \frac{\vec{w}}{\|w\|^2}$$

$$\frac{[x_0, y_0]\begin{bmatrix}a\\b\end{bmatrix}}{\sqrt{a^2+b^2}} \cdot \frac{\begin{bmatrix}a\\b\end{bmatrix}}{\sqrt{a^2+b^2}}$$

2. Given a line $ax + by + c = 0$ and a point (x_0, y_0), find the formula for the minimum distance between the point and the line.

$$-\begin{bmatrix}\frac{ac}{a^2+b^2}\\ \frac{bc}{a^2+b^2}\end{bmatrix}$$



$$[a, b]\begin{bmatrix}x\\y\end{bmatrix} + c = 0$$

$$d = \frac{\left|[x_0, y_0]\begin{bmatrix}a\\b\end{bmatrix} + c\right|}{\sqrt{a^2+b^2}}$$

3. In class, you were taught to only consider the perceptron that goes through the origin. We will now show that this formulation is sufficient to encompass the case where the perceptron does not go through the origin.

   Consider the classification function of a perceptron classifier that does not go through the origin,

$$h(x) = w^T x + b$$

   where w and b are the hyper plane parameters.

   Now, consider the classification function of a perceptron classifier that does go through the origin

$$\tilde{h}(\tilde{x}) = \tilde{w}^T \tilde{x}.$$

   Find a way to formulate $\tilde{w}$ and $\tilde{x}$ in terms of x, b, w.

$$\tilde{x} = [1, x]^T$$

$$\tilde{w} = [b, w]^T$$