Problems: 2.8, 3.6, 3.9, 3.24, 4.3, 5.5 from textbook, 7 and 8 from homework file. To solve this homework, I used a lot of online resources like math.stackexchange.com to double check theory.

# 1 Problem 1

## 1.1 a

The integral of the polynomial p(x) will be equal to $c_1x+1/2c_2x^2+1/3c_3x^3...1/(n)c_nx^n$. Taking this integral from $\alpha$ to $\beta$, we will get $c_1(\beta-\alpha)+1/2c_2(\beta^2-\alpha^2)+1/3c_3(\beta^3-\alpha^3)...1/(n)c_n(\beta^n-\alpha^n)$. Therefore, since c is the vector $[c_1, c_2, c_3...c_n]^T$, we can say that for some vector a, if $a^Tc=$

$$\int_{\alpha}^{\beta} p(x)\,dx$$

, then a $=[(\beta-\alpha), 1/2(\beta^2-\alpha^2), 1/3(\beta^3-\alpha^3)...1/n(\beta^n-\alpha^n)]^T$.

## 1.2 b

The derivative of the polynomial $p(\alpha)$ will be equal to $c_2+2c_3\alpha+3c_4\alpha^2...(n-1)c_n\alpha^{n-2}$. Therefore, since c is the vector $[c_1, c_2, c_3...c_n]^T$, we can say that for some vector b, if $b^Tc=p'(x)$ then b $=[0, 1, 2\alpha, 3\alpha^2...(n-1)\alpha^{n-2}]$.

# 2 Problem 2

The derivative of f(z) for some vector z is equal to $z/||z||$, since f(x) = $||x||$. Therefore, $\hat{f}(x)=||z||+zz^T(x-z)/||z||$, which we can simplify to the form of $||z||-||z||+z^Tx/||z||$.

# 3 Problem 3

Since the definitions of $||x-c||^2=(x-c)^T(x-c)$ and $||x-d||^2=(x-d)^T(x-d)$, we can say that $f(x)=(x-c)^T(x-c)-(x-d)^T(x-d)$. Expanding this out gives us $f(x)=(x^Tx-x^Tc-c^Tx+c^Tc)-(x^Tx-x^Td-d^Tx+d^Td)=-2x^Tc+||c||^2+2x^Td-||d||^2=2x^T(d-c)+||c||^2-||d||^2$. We can invert the first term to instead make the equation of the form $2(d-c)^Tx+||c||^2-||d||^2$, which is now of the form $a^Tx+b$. This is therefore affine.

# 4 Problem 4

## 4.1 a

If we have vector x = 1, then let $z_1=-1$, meaning it has distance 2 from x and angle $\pi/2$. Let there also be $z_2=10$, meaning it has distance 9 from x and angle 0. Therefore, the smaller distance does not have the smaller angle.

### 4.2 b

When a vector is normalized, then for some vector $z_i, ||z_i|| = 1, or z_i/||z_i||$. For some vectors $z_j$ and $z_i$, if $z_j$ is the closest vector to x, then we can say $||x - z_j|| \leq ||z - z_i||$, or $||x - z_j||^2 \leq ||z - z_i||^2$. This is equal to $||x||^2 + ||z_j||^2 - 2x^T z_j \leq ||x||^2 + ||z_i||^2 - 2x^T z_i$. This means that $||z_j||^2 - 2x^T z_j \leq ||z_i||^2 - 2x^T z_i$, and since $z_j \geq z_i$, this means $x^T z_j \geq x^T z_i$. If we turn these into the normalize versions of the vectors, we get $\frac{x^T z_j}{||x||||z_j||} / \geq \frac{x^T z_i}{||x||||z_i||}$. We can then apply the arccos function to both sides to find the angle between $x and z_j$ and the angle between $x and z_i$. This gives us $\angle(x, z_j) \leq (x, z_i)$ since arccos is a decreasing function, thereby proving the nearest neighbor and angle nearest neighbor are the same $z_j$.

## 5 Problem 5

When we split the vectors into two groups, we are effectively saying that those vectors in $G_1$ are closer to some vector $z_1$ that exists in $G_1$, and those vectors in $G_2$ are closer to some vector $z_2$ that exists in $G_2$. Therefore, $||x - z_1||^2 - ||x - z_2||^2 = ||x||^2 - 2z_1^T x + ||z_1||^2 - ||x||^2 + 2z_2^T x - ||z_2||^2 = -2(z_1 - z_2)^T x + ||z_1||^2 - ||z_2||^2$. Therefore, there exists a vector $w = -2(z_1 - z_2)^T x$ and a scalar $x = ||z_1||^2 - ||z_2||^2$ that satisfy this.

## 6 Problem 6

We know two vectors are orthogonal to each other when their resulting dot product is equal to 0. Therefore, for 2 n-vectors, $a$ and $b$, we need $a \cdot b = 0$ and $(a - \gamma b) \cdot b = 0$. If $b = 0$, then all vectors are orthogonal to it, so $\gamma$ can equal any constant. If $b \neq 0$, then since $(a - \gamma b)^T b = 0$, $a^T b = a^T b - b^T b\gamma$. This means $\gamma = a^T b / b^T b$.

## 7 Problem 7

The function $f(x) = x_T -$avg(x) is linear because $f(\alpha x + \beta y)$ is equal to $\alpha x_T - \alpha$ avg(x) + $\beta y_T - \beta$ avg(y) and $\alpha f(x) + \beta f(y) = \alpha x_T - \alpha$ avg(x) + $\beta y_T - \beta$ avg(y). Since this is linear, that means that for $f(x) = a^T x = x_T -$avg(x), there exists some vector a such that this is true. This vector a is equal to $[a_1, a_2 ... a_T]$, meaning $a^T x = a_1 x_1 + x_2 x_2 ... a T_{x T}$. For this to equal $x_T -$avg(x), a must be equal to the array $[0, 0 ... 1]$ with length T.

# 8 Problem 8

## 8.1 a

This statement is true, since for any 2 vectors, if the angle between the two is acute, then the resulting vector from adding the two together will be longer than both. This is held true by the triangle inequality, which will always be true when the angle between the two is acute.

## 8.2 b

For a vector a, avg(a) is calculated from summing the elements in the vector and dividing by the total length of a. The rms(a) is calculated from the norm of a divided by the square root of the number of elements in a, or $||a||/\sqrt{n}$. If we say a is made up of $[a_1, a_2...a_n]$, then avg(a) = $(a_1 + a_2... + a_n)/n$ and rms(a) = $\sqrt{a_1^2 + a_2^2... + a_n}/\sqrt{n}$. Therefore, rms(a) is always bigger than avg(a) since $||a|| \geq sum(a)$ and $n > \sqrt{n}$, so this is true.