ASSIGNMENT 10, DUE FRIDAY 8 DECEMBER

1. Let V be a complex vector space having an inner product $\langle -, - \rangle \colon V \times V \to \mathbb{C}$. Prove that if S is any subset of V, then

$$S^{\perp} = \{u \in V \mid \langle u, v \rangle = 0 \text{ for all } v \in S\}.$$

is a subspace of V.

2. Use the Gram-Schmidt algorithm to obtain an orthonormal basis of

$$\operatorname{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

under the dot product on $\mathbb{R}^4$.

3. Let V be a complex vector space having an inner product $\langle -, - \rangle \colon V \times V \to \mathbb{C}$ and let U be a subspace of V. Prove that $V = U \oplus U^{\perp}$. (Hint: Extend a basis of U to a basis of V and use Gram-Schmidt.)

4. Let V be a complex vector space with inner product

$$\langle -, - \rangle \colon V \times V \longrightarrow \mathbb{C}.$$

Let U be a subspace of V and let v be any vector in V. We saw in class that since $V = U \oplus U^{\perp}$, we may write $v = u_0 + w_0$ with $u_0 \in U$ and $w_0 \in U^{\perp}$. Because

$$v - u_0 = w_0 \in U^{\perp},$$

it follows that u_0 is a best least squares approximation to v in U. Prove that u_0 is *the only* best least squares approximation to v in U. In other words, let $u_1 \in U$ be another best least squares approximation to v and show that $u_1 = u_0$. (Hint: This does not require the triangle inequality at all, but the Pythagorean theorem might help.)

5. Find the line $y = mx + b$ that best fits the set of data

x	0	1	2	3
y	-2	3	7	10

.

Plot the data points and their best fit line.

6. The theory of linear regression extends to that of quadratic regression, i.e., given a set of data, it is also possible to find a quadratic polynomial that best fits it. In the setup for linear regression, we started with a set of data

x	x_1	x_2	$\dots$	x_N
y	y_1	y_2	$\dots$	y_N

and then wanted to find values m and b such that

$$\sum_{i=1}^{N}(y_i - (mx_i + b))^2 \qquad (*)$$

is as small as possible, the reason being that the resulting line $y = mx + b$ will be as close to the data points as possible. To solve for m and b, we formed the matrix and vector

$$X = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}.$$

We then observed for any vector $\mathbf{v} = (m, b)^t \in \mathbb{R}^2$ that the quantity $(*)$ is equal to $\|\mathbf{y} - X\mathbf{v}\|^2$. Solving the linear regression problem therefore amounted to finding the vector $\mathbf{v}_0 \in \mathbb{R}^2$ such that

$$\|\mathbf{y} - X\mathbf{v}_0\| \leq \|\mathbf{y} - X\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{R}^2.$$

To adapt this idea to quadratic regression, we need to find scalars a, b and c such that

$$\sum_{i=1}^{N}(y_i - (ax_i^2 + bx_i + c))^2 \qquad (**)$$

is as small as possible. The resulting parabola $y = ax^2 + bx + c$ will then be as close to the data points as possible.

Find the $N \times 3$ matrix X and vector $\mathbf{y} \in \mathbb{R}^N$ such that the quantity $(**)$ is equal to $\|\mathbf{y} - X\mathbf{v}\|^2$ for any $\mathbf{v} = (a, b, c)^t \in \mathbb{R}^3$. Use these to find the parabola

$$y = ax^2 + bx + c$$

that best fits the data

x	-2	0	1	2
y	5	1	1	5

Plot the data points and their best fit parabola.