Math 314. Week 12 worksheet (§6.4, §6.5).

RECALL: The projection of $\mathbf{v}$ onto a subspace H in $\mathbb{R}^n$ is

- the unique vector $\hat{\mathbf{v}} \in H$ such that $\mathbf{v} - \hat{\mathbf{v}} \in H^{\perp}$; and also
- the closest vector to $\mathbf{v}$ in H, which means $||\mathbf{v} - \mathbf{h}|| \geq ||\mathbf{v} - \mathrm{proj}_H(\mathbf{v})||$ for all points $\mathbf{h} \in H$.

If $\{\mathbf{u_1}, \dots, \mathbf{u_t}\}$ is an *orthogonal* basis for H, then we have a formula

$$\mathrm{proj}_H(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u_1}}{\mathbf{u_1} \cdot \mathbf{u_1}}\mathbf{u_1} + \dots + \frac{\mathbf{v} \cdot \mathbf{u_t}}{\mathbf{u_t} \cdot \mathbf{u_t}}\mathbf{u_t}.$$

A. GRAM-SCHMIDT PROCESS. Let $\mathbf{v_1} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\mathbf{v_2} = \begin{bmatrix}2\\0\\2\end{bmatrix}$. We will use the Gram-Schmidt process to replace $\{\mathbf{v_1}, \mathbf{v_2}\}$ by an orthonormal set $\{\mathbf{u_1}, \mathbf{u_2}\}$ with the same span.

(1) Our first goal is to make an orthogonal set $\{\mathbf{w_1}, \mathbf{w_2}\}$. Keep the first vector the same: $\mathbf{w_1} = \mathbf{v_1}$. For the second vector take $\mathbf{w_2} = \mathbf{v_2} - \frac{\mathbf{v_2} \cdot \mathbf{w_1}}{\mathbf{w_1} \cdot \mathbf{w_1}}\mathbf{w_1}$.

(2) Convince yourself[1] that $\{\mathbf{w_1}, \mathbf{w_2}\}$ is an orthogonal set.

(3) Our next goal is to turn the orthogonal set into an orthonormal set. Divide each vector $\mathbf{w_i}$ by a scalar to turn it into a unit vector $\mathbf{u_i}$.

(4) Convince yourself that $\{\mathbf{u_1}, \mathbf{u_2}\}$ is an orthonormal set.

(5) Use your answer to compute the projection of the vector $\begin{bmatrix}3 & 2 & 1\end{bmatrix}^T$ onto $H = \mathrm{Span}\{\mathbf{v_1}, \mathbf{v_2}\}$.

B. GRAM-SCHMIDT AGAIN. Let $\mathbf{v_1}, \mathbf{v_2}$ be as in the previous problem, and $\mathbf{v_3} = \begin{bmatrix}3 & 2 & 1\end{bmatrix}^T$. In this problem, we apply the Gram-Schmidt process to the set $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$.

(1) Our first goal is to make an orthogonal set $\{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\}$. Here are the formulas:
   - $\mathbf{w_1} = \mathbf{v_1}$
   - $\mathbf{w_2} = \mathbf{v_2} - \frac{\mathbf{v_2} \cdot \mathbf{w_1}}{\mathbf{w_1} \cdot \mathbf{w_1}}\mathbf{w_1}$
   - $\mathbf{w_3} = \mathbf{v_3} - \frac{\mathbf{v_3} \cdot \mathbf{w_1}}{\mathbf{w_1} \cdot \mathbf{w_1}}\mathbf{w_1} - \frac{\mathbf{v_3} \cdot \mathbf{w_2}}{\mathbf{w_2} \cdot \mathbf{w_2}}\mathbf{w_2}$.

   In each step we are subtracting off the projection of the next $\mathbf{v_i}$ onto the span of the previous $\mathbf{w_i}$'s we already computed.

(2) Our second goal is to turn the orthogonal set into an orthonormal set. Divide each vector $\mathbf{w_i}$ by a scalar to turn it into a unit vector $\mathbf{u_i}$.

(3) Use your answer to find a basis for $H^{\perp}$, where $H = \mathrm{Span}\{\mathbf{v_1}, \mathbf{v_2}\}$.

If $A\mathbf{x} = \mathbf{b}$ is a linear system, then a *least-squares solution* to $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}}$ such that $||\mathbf{b} - A\mathbf{v}|| \geq ||\mathbf{b} - A\hat{\mathbf{x}}||$ for all vectors $\mathbf{v}$. That is, $A\hat{\mathbf{x}}$ is as close as possible to $\mathbf{b}$.

If $\hat{\mathbf{x}}$ is a least-squares solution to $A\mathbf{x} = \mathbf{b}$, then $A\hat{\mathbf{x}} = \mathrm{proj}_{\mathrm{Col}(A)}(\mathbf{b})$.

$\hat{\mathbf{x}}$ is a least-squares solution to $A\mathbf{x} = \mathbf{b} \iff \hat{\mathbf{x}}$ is a solution to $A^TA\mathbf{x} = A^T\mathbf{b}$. The system $A^TA\mathbf{x} = A^T\mathbf{b}$ is called the system of *normal equations*.

C. LEAST SQUARES Let $A = \begin{bmatrix}-3\\2\end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix}5\\-1\end{bmatrix}$.

(1) Is the system $A\mathbf{x} = \mathbf{b}$ consistent?

(2) Write and solve the normal equations for this system. What is the least-squares solution?

(3) Draw the vector corresponding to the column of A, the vector $\mathbf{b}$, and the vector $A\hat{\mathbf{x}}$. Interpret their relationship in the context of least-squares solutions.

[1]Hint: $\frac{\mathbf{v_2} \cdot \mathbf{w_1}}{\mathbf{w_1} \cdot \mathbf{w_1}}\mathbf{w_1} = \mathrm{proj}_{\mathrm{Span}\{\mathbf{w_1}\}}(\mathbf{v_2})$

D. LEAST SQUARES AGAIN. Let $A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$.

(1) Is the system $A\mathbf{x} = \mathbf{b}$ consistent?
(2) Write and solve the normal equations for this system. What is the least-squares solution?
(3) Draw the vector corresponding to the column of A, the vector $\mathbf{b}$, and the vector $A\hat{\mathbf{x}}$. Interpret their relationship in the context of least-squares solutions.

E. LINE OF BEST FIT. In this problem, we will consider lines of the form $y = ax + b$ in the plane $\mathbb{R}^2$.

(1) Plug in each of the points $(3, -2)$, $(5, -1)$, $(7, 0)$ into the equation $y = ax + b$ to get a linear system, and solve for $\begin{bmatrix} a \\ b \end{bmatrix}$. Do these points lie on a line? If so, which one?
(2) Repeat the same process for the points $(1, 5)$, $(2, 4)$, $(5, 3)$.
(3) Solve the normal equations for the linear system you found in the previous part. This gives a recipe for a line $y = \hat{a}x + \hat{b}$. Graph it and the points from the previous part.
(4) The line you found in part (3) is the closest line to the points $(1, 5)$, $(2, 4)$, $(5, 3)$ in a precise sense. Use the definition of least-squares solution to make this precise.
(5) Here is some totally made up data on advertising spending and sales for local dog food companies:

spending (thousands of dollars)	0	0	.3	1.2	3.6	8.9	46.7
sales (thousands of dollars)	4	18	45	19	40	29	346

How would you find the line that best fits these data points?

F. LINEARLY INDEPENDENCE REVIEW. Consider the set of polynomials $S = \{t^2, (t-1)^2, (t-2)^2, (t-3)^2\}$.

(1) Is any element of S a scalar multiple of another one?
(2) Is the set S linearly independent?[2]
(3) What is $\mathrm{Span}(S)$? Can you find a basis for it?

G. LINEAR TRANSFORMATION REVIEW. Consider the function $E : P_2 \to \mathbb{R}^4$ given by $E(p(t)) = \begin{bmatrix} p(0) \\ p(1) \\ p(2) \\ p(3) \end{bmatrix}$.

(1) Check carefully that E is a linear transformation.
(2) What is the kernel of E?
(3) What is the range of E? Find a basis for it.

H. PARABOLA OF BEST FIT. A parabola is the graph of a polynomial of degree two[3] on the plane.

(1) Do the points $(-1, 1)$, $(0, 2)$, $(4, 3)$ lie on a parabola?
(2) What about $(-1, 1)$, $(0, 2)$, $(4, 3)$, and $(5, 4)$?
(3) Find the parabola of best fit for the four points in the last problem.
(4) Explain why any three points with different x-coordinates lie on a parabola.
(5) Explain why any set of points with different x-coordinates have a unique parabola of best fit.

I. $\mathcal{B}$-MATRIX REVIEW. Consider the function $D : P_2 \to P_2$ given by $D(p(t)) = \frac{dp}{dt}$.

(1) The set of polynomials $\mathcal{B} = \{t^2, (t-1)^2, (t-2)^2\}$ is a basis for P_2. So is $\mathcal{C} = \{t^2, t, 1\}$. Find the change-of-coordinates matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$.
(2) Find the $\mathcal{B}$-matrix of D.

[2] This is NOT the same question as the previous one.
[3] Let's include lines as special cases of parabolas, so we can go with degree at most two, rather than exactly two.