6.4-5 The Gram-Schmidt Process; Least Squares

Example 1 Find an orthogonal basis for the subspace spanned by $\mathbf{x}_1 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

The Gram-Schmidt Process

The Gram-Schmidt Process gives us a simple way of producing an orthogonal basis for any non-zero subspace of $\mathbb{R}^n$.

Suppose we have a basis for a subspace W of $\mathbb{R}^n$ made of the vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$. An orthogonal basis for W is $\{v_1, v_2, \ldots, v_p\}$, where

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Example 2 Find an orthogonal basis for the subspace spanned by $\mathbf{x}_1 = \begin{pmatrix} 3 \\ -4 \\ 5 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} -3 \\ 14 \\ -7 \end{pmatrix}$, and $\mathbf{x}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$.

Consider the over-determined $A\mathbf{x} = \mathbf{b}$ equation $\begin{pmatrix} 2 & 4 \\ 3 & 1 \\ 5 & -4 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 11 \\ 8 \\ 6 \end{pmatrix}$. The equation does not have an exact solution.

Can you find a suitable $\hat{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ that gives a $\hat{\mathbf{b}}$ "close" to $\begin{pmatrix} 10 \\ 11 \\ 8 \\ 6 \end{pmatrix}$? What is the closest you can find?

Least-Squares Problems

Definition

If A is an $m \times n$ matrix and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\| \mathbf{b} - A\hat{\mathbf{x}} \| \le \| \mathbf{b} - A\mathbf{x} \|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$

The solution to a Least-Squares Problem

Suppose $\hat{\mathbf{x}}$ satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. By the Orthogonal Decomposition Theorem in Section 6.3, the projection $\hat{\mathbf{b}}$ has the property that $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to col A, so $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to each column in A. If $\mathbf{a}_j$ is any column of A, then $\mathbf{a}_j \cdot (\mathbf{b} - A\hat{\mathbf{x}}) = 0$ or $\mathbf{a}_j^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$. Since each $\mathbf{a}_j^T$ is a row of A^T, $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$. Thus,

$$A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = \mathbf{0}$$
$$A^T\mathbf{b} = A^TA\hat{\mathbf{x}}$$

This shows that each least-squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies the equation $A^TA\mathbf{x} = A^T\mathbf{b}$. This is called the **normal equations** for $A\mathbf{x} = \mathbf{b}$.

Example 3 Find a least-squares solution for $A = \begin{pmatrix} 2 & 4 \\ 3 & 1 \\ 5 & -4 \\ 3 & 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 10 \\ 11 \\ 8 \\ 6 \end{pmatrix}$.

Example 4 Find the least-squares line, $ax + b = y$ for the points (1, 3), (3, 6), (4, 5), and (6, 8). Compare with your calculator's linear regression function. Also, note the correlation coefficient R^2.

Example 5 Find the least-squares solution for the quadratic that best fits the points (0, 2), (1, −1), (2, 3), and (3, 8). Compare your answer with your calculator.

Example 6 A function f is a linear combination of x^2, $\sqrt{x}$, and $\ln(x)$. Points on the graph of f are (1, −2.5), (2, 5), (3, 17), (4, 34), and (5, 56). Find the least-squares solution for f, that is, the function that best fits the data.

Example 7 Find a point in the plane that lies closest to the four lines: $3x + 4y = 12$, $y = x - 5$, $y = 3x + 2$, and $x - y = 1$.

Example 8 Find the LR equation of the plane that is closest to the four points: (2, 3, −1), (2, 0, 5), (0, 4, −2), and (2, −3, 5). Rearrange the equation $ax + by + cz = d$ by dividing by c and solving for z.

Example 9 Find the equation of a circle, $(x - h)^2 + (y - k) = r^2$, that passes through the points (−2, 8), (1, 7) and (3, 3). *Note*: expanding the equation of a circle we get, $x^2 - 2xh + h^2 + y^2 - 2yk + k^2 = r^2$, which can be rearranged to $x^2 + y^2 = 2hx + 2yk + r^2 - h^2 - k^2$, which can also be written as $C_1x + C_2y + C_3 = x^2 + y^2$ using substitution.

Example 10 If a fourth point (−4, 4) is included in Example 6 that is not on the circle, find the *best* circle that fits the four points.

Extra: The Correlation Coefficient R^2: Most statistics texts discuss the correlation coefficient, R^2, a measure of how well a model fits the data. Suppose $\hat{\mathbf{x}}$ is a LS solution to $A\mathbf{x} = \mathbf{b}$, i.e., $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. If we let, $\mathrm{SS}(E) = \hat{\mathbf{x}}^TA^TNA\hat{\mathbf{x}}$ and $\mathrm{SS}(T) = \mathbf{b}^TN\mathbf{b}$, where $N = I - \frac{1}{n}[1]_n$. Then, $R^2 = \frac{\mathrm{SS}(E)}{\mathrm{SS}(T)}$. (Also, $\mathrm{SS}(E) = \hat{\mathbf{b}}^TN\hat{\mathbf{b}}$)

Example 11 Find R^2 for Example 4 and Example 5.