Linear Algebra

Suppose that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

that is,

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

Thus, a linear system, $A\mathbf{x} = \mathbf{b}$, of m equations in n unknowns can be written as

$$\boxed{x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n = \mathbf{b}}$$

from which we conclude that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is expressible as a linear combination of the column vectors of A.

THEOREM 4.7.1 *A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.*

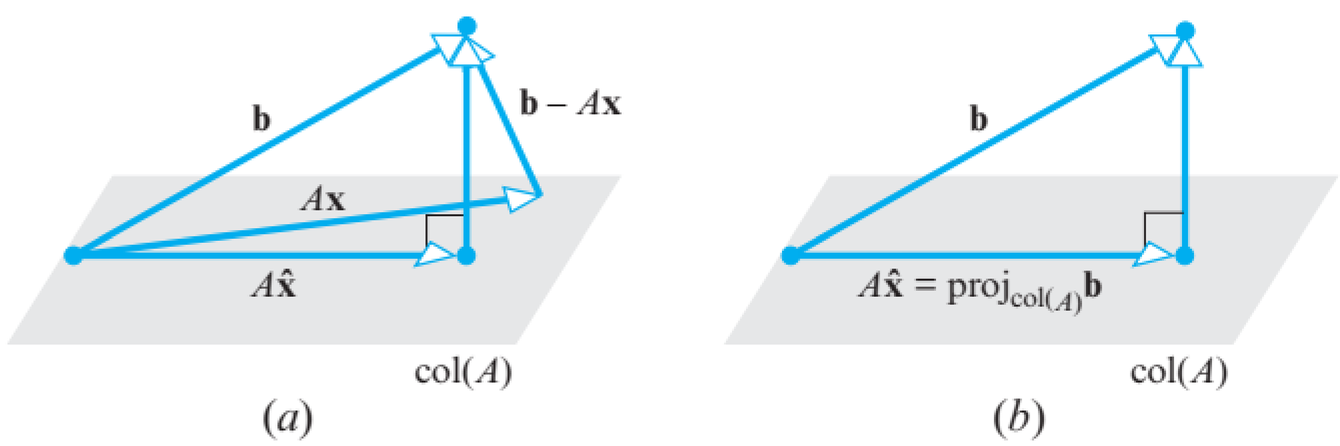
Best Approximation; Least Squares

- There are many applications in which some linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns should be consistent on physical grounds but fails to be so because of measurement errors in the entries of A or $\mathbf{b}$.
- In such cases one looks for vectors that come as close as possible to being solutions in the sense that they minimize $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m.
- In this section we will discuss methods for finding such minimizing vectors.

- Suppose that $A\mathbf{x} = \mathbf{b}$ is an *inconsistent* linear system of m equations in n unknowns in which we suspect the inconsistency to be caused by errors in the entries of A or $\mathbf{b}$.

- Since no exact solution is possible, we will look for a vector $\mathbf{x}$ that comes as "close as possible" to being a solution in the sense that it minimizes $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m.

- You can think of $A\mathbf{x}$ as an approximation to $\mathbf{b}$ and $\|\mathbf{b} - A\mathbf{x}\|$ as the *error* in that approximation—the smaller the error, the better the approximation.



Least Squares Problem Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, find a vector $\mathbf{x}$ in R^n that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. We call such a vector, if it exists, a ***least squares solution*** of $A\mathbf{x} = \mathbf{b}$, we call $\mathbf{b} - A\mathbf{x}$ the ***least squares error vector***, and we call $\|\mathbf{b} - A\mathbf{x}\|$ the ***least squares error***.

- That being the case, to find a least squares solution of $A\mathbf{x} = \mathbf{b}$ is equivalent to finding a vector $A\hat{\mathbf{x}}$ in the column space of A that is closest to $\mathbf{b}$ in the sense that it minimizes the length of the vector $\mathbf{b} - A\mathbf{x}$.

- This is illustrated in Figure 6.4.1a, which also suggests that $A\hat{\mathbf{x}}$ is the orthogonal projection of $\mathbf{b}$ on the column space of A, that is, $A\hat{\mathbf{x}} = \text{proj}_{\text{col}(A)}\mathbf{b}$.

Finding Least Squares Solutions

One way to find a least squares solution of $A\mathbf{x} = \mathbf{b}$ is to calculate the orthogonal projection $\text{proj}_W \mathbf{b}$ on the column space W of A and then solve the equation

$$A\mathbf{x} = \text{proj}_W \mathbf{b} \tag{2}$$

However, we can avoid calculating the projection by rewriting (2) as

$$\mathbf{b} - A\mathbf{x} = \mathbf{b} - \text{proj}_W \mathbf{b}$$

and then multiplying both sides of this equation by A^T to obtain

$$A^T(\mathbf{b} - A\mathbf{x}) = A^T(\mathbf{b} - \text{proj}_W \mathbf{b})$$

Since $\mathbf{b} - \text{proj}_W \mathbf{b}$ is the component of $\mathbf{b}$ that is orthogonal to the column space of A, it follows from Theorem 4.8.7(b) that this vector lies in the null space of A^T, and hence that

$$A^T(\mathbf{b} - \text{proj}_W \mathbf{b}) = \mathbf{0} \tag{3}$$

Thus, (3) simplifies to

$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$$

which we can rewrite as

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

This is called the ***normal equation*** or the ***normal system*** associated with $A\mathbf{x} = \mathbf{b}$.

THEOREM 6.4.2 *For every linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system*

$$A^TA\mathbf{x} = A^T\mathbf{b} \tag{5}$$

is consistent, and all solutions of (5) *are least squares solutions of $A\mathbf{x} = \mathbf{b}$. Moreover, if W is the column space of A, and $\mathbf{x}$ is any least squares solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on W is*

$$\operatorname{proj}_W \mathbf{b} = A\mathbf{x} \tag{6}$$

▶ EXAMPLE 1 Unique Least Squares Solution

Find the least squares solution, the least squares error vector, and the least squares error of the linear system

$$\begin{aligned} x_1 - x_2 &= 4 \\ 3x_1 + 2x_2 &= 1 \\ -2x_1 + 4x_2 &= 3 \end{aligned}$$

Solution

It will be convenient to express the system in the matrix form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

It follows that

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

so the normal system

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

is

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Solving this system yields a unique least squares solution, namely,

$$\boxed{x_1 = \tfrac{17}{95}, \quad x_2 = \tfrac{143}{285}}$$

The least squares error vector is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{95}{57} \end{bmatrix} = \begin{bmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{bmatrix}$$

and the least squares error is

$$\|\mathbf{b} - A\mathbf{x}\| \approx 4.556$$

▶ EXAMPLE 2 Infinitely Many Least Squares Solutions

Find the least squares solutions, the least squares error vector, and the least squares error of the linear system

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 2 \\ x_1 - 4x_2 + 3x_3 &= -2 \\ x_1 + 10x_2 - 7x_3 &= 1 \end{aligned}$$

Solution

The matrix form of the system is $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

It follows that

$$A^TA = \begin{bmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{bmatrix} \quad \text{and} \quad A^T\mathbf{b} = \begin{bmatrix} 5 \\ 22 \\ -15 \end{bmatrix}$$

so the normal system

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

is $\begin{bmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 22 \\ -15 \end{bmatrix}$

so the augmented matrix for the normal system $A^TA\mathbf{x} = A^T\mathbf{b}$ is

$$\hat{A} = \left[\begin{array}{ccc|c} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right]$$

The reduced row echelon form of this matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{7} & \frac{2}{7} \\ 0 & 1 & -\frac{5}{7} & \frac{13}{84} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

from which it follows that there are infinitely many least squares solutions, and that they are given by the parametric equations

$$\boxed{\begin{aligned} x_1 &= \tfrac{2}{7} - \tfrac{1}{7}t \\ x_2 &= \tfrac{13}{84} + \tfrac{5}{7}t \\ x_3 &= t \end{aligned}}$$

As a check, let us verify that all least squares solutions produce the same least squares error vector and the same least squares error. To see that this is so, we first compute

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{bmatrix} \begin{bmatrix} \frac{2}{7} - \frac{1}{7}t \\ \frac{13}{84} + \frac{5}{7}t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{7}{6} \\ -\frac{1}{3} \\ \frac{11}{6} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ -\frac{5}{3} \\ -\frac{5}{6} \end{bmatrix}$$

Since $\mathbf{b} - A\mathbf{x}$ does not depend on t, all least squares solutions produce the same error vector, namely

$$\boxed{\|\mathbf{b} - A\mathbf{x}\| = \sqrt{\left(\tfrac{5}{6}\right)^2 + \left(-\tfrac{5}{3}\right)^2 + \left(-\tfrac{5}{6}\right)^2} = \tfrac{5}{6}\sqrt{6}}$$

$$\hat{A} = \left[\begin{array}{cccc} 1 & 12/11 & -7/11 & \vdots & 5/11 \\ 12 & 120 & -84 & \vdots & 22 \\ -7 & -84 & 59 & \vdots & -15 \end{array}\right]$$

$$\hat{A} = \left[\begin{array}{ccccc} 1 & 12/11 & -7/11 & \vdots & 5/11 \\ 12 & 120 & -84 & \vdots & 22 \\ -7 & -84 & 59 & \vdots & -15 \end{array}\right] \begin{array}{c} R_1 \\ R_2 \to R_2 - 12R_1 \\ R_3 \to R_3 + 7R_1 \end{array}$$

$$\hat{A} = \left[\begin{array}{ccccc} 1 & 12/11 & -7/11 & \vdots & 5/11 \\ 0 & 1176/11 & -840/11 & \vdots & 182/11 \\ 0 & -840/11 & 600/11 & \vdots & -130/11 \end{array}\right] \begin{array}{c} R_1 \\ R_2 \to R_2 - 12R_1 \\ R_3 \to R_3 + 7R_1 \end{array}$$

$$\hat{A} = \left[\begin{array}{ccccc} 1 & 12/11 & -7/11 & \vdots & 5/11 \\ 0 & 1 & -5/7 & \vdots & 13/84 \\ 0 & -840/11 & 600/11 & \vdots & -130/11 \end{array}\right] \begin{array}{c} R_1 \\ R_2 * (11/1176) \\ R_3 \to R_3 + 7R_1 \end{array}$$

THEOREM 6.4.4 *If A is an $m \times n$ matrix with linearly independent column vectors, then for every $m \times 1$ matrix $\mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution. This solution is given by*

$$\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b} \tag{9}$$

Moreover, if W is the column space of A, then the orthogonal projection of $\mathbf{b}$ on W is

$$\text{proj}_W \mathbf{b} = A\mathbf{x} = A(A^TA)^{-1}A^T\mathbf{b} \tag{10}$$

▶ EXAMPLE 3 A Formula Solution to Example 1

Use Formula (9) to find the least squares solution of the linear system in Example 1.

Solution

$$\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$$

$$= \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

$$\boxed{\mathbf{x} = \frac{1}{285}\begin{bmatrix} 21 & 3 \\ 3 & 14 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix}}$$

In Exercises 1–2, find the associated normal equation.

1. $\begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$

2. $\begin{bmatrix} 2 & -1 & 0 \\ 3 & 1 & 2 \\ -1 & 4 & 5 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$

Solution If we call the system $A\mathbf{x} = \mathbf{b}$, then the associated normal system is

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$$

which simplifies to

$$\boxed{\begin{bmatrix} 21 & 25 \\ 25 & 35 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 20 \\ 20 \end{bmatrix}}$$

▶ In Exercises 3–6, find the least squares solution of the equation $A\mathbf{x} = \mathbf{b}$.

3. $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 4 & 5 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$

4. $A = \begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -2 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 9 \\ 3 \end{bmatrix}$

6. $A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & -2 & 2 \\ 2 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 0 \\ 6 \\ 0 \\ 6 \end{bmatrix}$

Solution **5.** The associated normal system is

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & -2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -2 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & -2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 9 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 7 & 4 & -6 \\ 4 & 3 & -3 \\ -6 & -3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 18 \\ 12 \\ -9 \end{bmatrix}$$

This system has solution

$$\boxed{\begin{aligned} x_1 &= 12, \\ x_2 &= -3, \\ x_3 &= 9 \end{aligned}}$$

The orthogonal projection of **b** on the column space of A is $A\mathbf{x}$, or

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -2 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 12 \\ -3 \\ 9 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 9 \\ 0 \end{bmatrix}$$

which can be written as $\underline{(3, 3, 9, 0)}$.

THANK YOU