8.5   Least Squares Solutions to Inconsistent Systems

Performance Criterion:

8. (f) Find the least-squares approximation to the solution of an inconsistent system of equations. Solve a problem using least-squares approximation.

   (g) Give the least squares error and least squares error vector for a least squares approximation to a solution to a system of equations.

Recall that an inconsistent system is one for which there is no solution. Often we wish to solve inconsistent systems and it is just not acceptable to have no solution. In those cases we can find some vector (whose components are the values we are trying to find when attempting to solve the system) that is "closer to being a solution" than all other vectors. The theory behind this process is part of the second term of this course, but we now have enough knowledge to find such a vector in a "cookbook" manner.

Suppose that we have a system of equations $A\mathbf{x} = \mathbf{b}$. Pause for a moment to reflect on what we know and what we are trying to find when solving such a system: We have a system of linear equations, and the entries of A are the coefficients of all the equations. The vector $\mathbf{b}$ is the vector whose components are the right sides of all the equations, and the vector $\mathbf{x}$ is the vector whose components are the unknown values of the variables we are trying to find. So we know A and $\mathbf{b}$ and we are trying to find $\mathbf{x}$. If A is invertible, the solution vector $\mathbf{x}$ is given by $\mathbf{x} = A^{-1}\mathbf{b}$. If A is not invertible there will be no solution vector $\mathbf{x}$, but we can usually find a vector $\bar{\mathbf{x}}$ (usually spoken as "ex-bar") that comes "closest" to being a solution. Here is the formula telling us how to find that $\bar{\mathbf{x}}$:

> Theorem 8.5.1: **The Least Squares Theorem:** Let A be an $m \times n$ matrix and let $\mathbf{b}$ be in $\mathbb{R}^m$. If $A\mathbf{x} = \mathbf{b}$ has a least squares solution $\bar{\mathbf{x}}$, it is given by
> $$\bar{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

$$
\begin{aligned}
1.3x_1 + 0.6x_2 &= 3.3 \\
\diamond \textbf{ Example 8.5(a):} \text{ Find the least squares solution to} \quad 4.7x_1 + 1.5x_2 &= 13.5 \\
3.1x_1 + 5.2x_2 &= -0.1
\end{aligned}
$$

First we note that if we try to solve by row reduction we get no solution; this is an overdetermined system because there are more equations than unknowns. The matrix A and vector $\mathbf{b}$ are

$$
A = \begin{bmatrix} 1.3 & 0.6 \\ 4.7 & 1.5 \\ 3.1 & 5.2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 3.3 \\ 13.5 \\ -0.1 \end{bmatrix}
$$

Using a calculator or *MATLAB*, we get

$$
\bar{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} 3.5526 \\ -2.1374 \end{bmatrix}
$$

A classic example of when we want to do something like this is when we have a bunch of (x, y) data pairs from some experiment, and when we graph all the pairs they describe a trend. We then want to find a simple function $y = f(x)$ that best models that data. In some cases that function might be a line, in other cases maybe it is a parabola, and in yet other cases it might be an exponential function. Let's try to make the connection between this and linear algebra. Suppose that we have the data points $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, and when we graph these points they arrange themselves in roughly a line, as shown below and to the left. We then want to find an equation

of the form $a + bx = y$ (note that this is just the familiar $y = mx + b$ rearranged and with different letters for the slope and y-intercept) such that $a + bx_i \approx y_i$ for $i = 1, 2, ..., n$, as shown below and to the right.
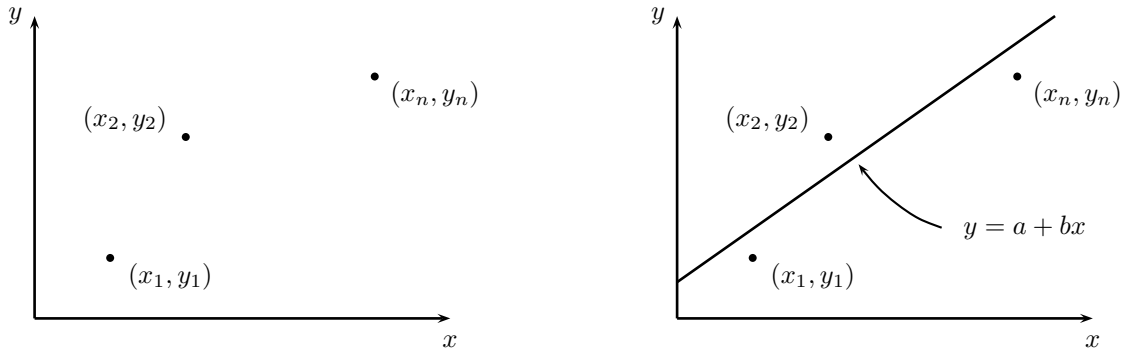

If we substitute each data pair into $a + bx = y$ we get a system of equations which can be thought of in several different ways. Remember that all the x_i and y_i are known values - the unknowns are a and b.

$$
\begin{array}{c}
a + x_1 b = y_1 \\
a + x_2 b = y_2 \\
\vdots \\
a + x_n b = y_n
\end{array}
\qquad \Longleftrightarrow \qquad
\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix}
\begin{bmatrix} a \\ b \end{bmatrix}
=
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}
\qquad \Longleftrightarrow \qquad A\mathbf{x} = \mathbf{b}
$$

Above we first see the system that results from putting each of the (x_i, y_i) pairs into the equation $a + bx = y$. After that we see the $A\mathbf{x} = \mathbf{b}$ form of the system. We must be careful of the notation here. A is the matrix whose columns are a vector in $\mathbb{R}^n$ consisting of all ones and a vector whose components are the x_i values. It would be logical to call this last vector $\mathbf{x}$, but instead $\mathbf{x}$ is the vector $\begin{bmatrix} a \\ b \end{bmatrix}$. $\mathbf{b}$ is the column vector whose components are the y_i values. Our task, as described by this interpretation, is to *find a vector* $\mathbf{x}$ *in* $\mathbb{R}^2$ *that* A *transforms into the vector* $\mathbf{b}$ *in* $\mathbb{R}^n$. Even if such a vector did exist, it couldn't be given as $\mathbf{x} = A^{-1}\mathbf{b}$ because A is not square, so can't be invertible. However, it is likely no such vector exists, but we *CAN* find the least-squares vector $\bar{\mathbf{x}} = \begin{bmatrix} a \\ b \end{bmatrix} = (A^T A)^{-1} A^T \mathbf{b}$. When we do, its components a and b are the intercept and slope of our line.

Theoretically, here is what is happening. Least squares is generally used in situations that are **over determined**. This means that there is too much information and it is bound to "disagree" with itself somehow. In terms of systems of equations, we are talking about cases where there are more equations than unknowns. Now the fact that the system $A\mathbf{x} = \mathbf{b}$ has no solution means that $\mathbf{b}$ is not in the column space of A. The least squares solution to $A\mathbf{x} = \mathbf{b}$ is simply the vector $\bar{\mathbf{x}}$ for which $A\bar{\mathbf{x}}$ is the projection of $\mathbf{b}$ onto the column space of A. This is shown simplistically below, for the situation where the column space is a plane in $\mathbb{R}^3$.
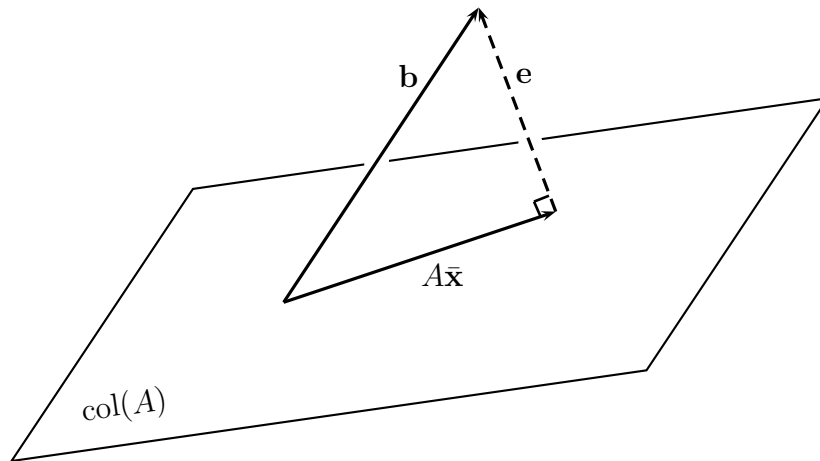
To recap a bit, suppose we have a system of equations $A\mathbf{x} = \mathbf{b}$ where there is no vector $\mathbf{x}$ for which $A\mathbf{x}$ equals $\mathbf{b}$. What the least squares approximation allows us to do is to find a vector $\bar{\mathbf{x}}$ for which $A\bar{\mathbf{x}}$ is as "close" to $\mathbf{b}$ as possible. We generally determine "closeness" of two objects by finding the difference between them. Because both $A\bar{\mathbf{x}}$ and $\mathbf{b}$ are both vectors of the same length, we can subtract them to get a vector $\mathbf{e}$ that we will call the **error vector**, shown above. The **least squares error** is then the magnitude of this vector:

<u>DEFINITION 8.5.2</u>: If $\bar{\mathbf{x}}$ is the least-squares solution to the system $A\mathbf{x} = \mathbf{b}$, the **least squares error vector** is
$$\vec{\varepsilon} = \mathbf{b} - A\bar{\mathbf{x}}$$
and the **least squares error** is the magnitude of $\vec{\varepsilon}$.

◇ **Example 8.5(b):** Find the least squares error vector and least squares error vector for the solution obtained in Example 8.5(a).

The least squares error vector is

$$\vec{\varepsilon} = \mathbf{b} - A\bar{\mathbf{x}} = \begin{bmatrix} 3.3 \\ 13.5 \\ -0.1 \end{bmatrix} - \begin{bmatrix} 1.3 & 0.6 \\ 4.7 & 1.5 \\ 3.1 & 5.2 \end{bmatrix} \begin{bmatrix} 3.5526 \\ -2.1374 \end{bmatrix} = \begin{bmatrix} -0.0359 \\ 0.0089 \\ 0.0016 \end{bmatrix}$$

The least squares error is $\|\vec{\varepsilon}\| = 0.0370$. ♠

Section 8.5 Exercises

1. Find the least squares approximating parabola for the points $(1,8)$, $(2,7)$, $(3,5)$, $(4,2)$. **Give the system of equations to be solved (in any form), and give the equation of the parabola.**