More on least squares approximation

Suppose we have data points $(x_1, y_1),\ (x_2, y_2), \dots,\ (x_n, y_n)$ and we want to "fit" them to a line. Let

$$\mathbf{Y}_d = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}.$$

Using orthogonal projection, we obtain

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}_d.$$

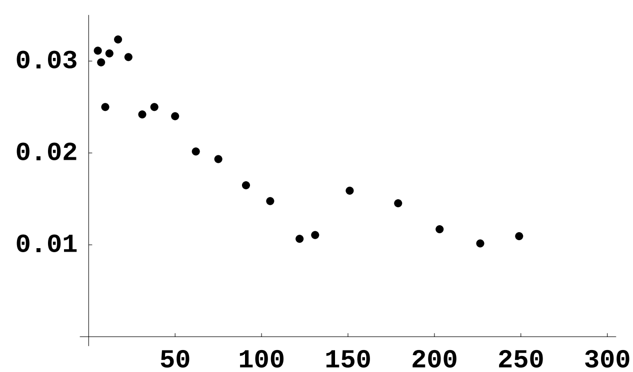
**Example.** Suppose that a company had profits of \$500,000 in year 1, \$1,000,000 in year 2, and \$2,000,000 in year 5. Model its profits with a least-squares linear model.

A more realistic example of least squares approximation

**Example.** Here are relative growth rates for the U.S. population from 1800 to 1990:

| Year | U.S. Population | Rel Growth Rate |
|---|---|---|
| 1800 | 5.3 | 0.03113 |
| 1810 | 7.2 | 0.02986 |
| 1820 | 9.6 | 0.02500 |
| 1830 | 12 | 0.03083 |
| 1840 | 17 | 0.03235 |
| 1850 | 23 | 0.03043 |
| 1860 | 31 | 0.02419 |
| 1870 | 38 | 0.02500 |
| 1880 | 50 | 0.02400 |
| 1890 | 62 | 0.02016 |
| 1900 | 75 | 0.01933 |
| 1910 | 91 | 0.01648 |
| 1920 | 105 | 0.01476 |
| 1930 | 122 | 0.01066 |
| 1940 | 131 | 0.01107 |
| 1950 | 151 | 0.01589 |
| 1960 | 179 | 0.01453 |
| 1970 | 203 | 0.01170 |
| 1980 | 226 | 0.01015 |
| 1990 | 249 | 0.01094 |

Here's a graph of these relative growth rates versus population:

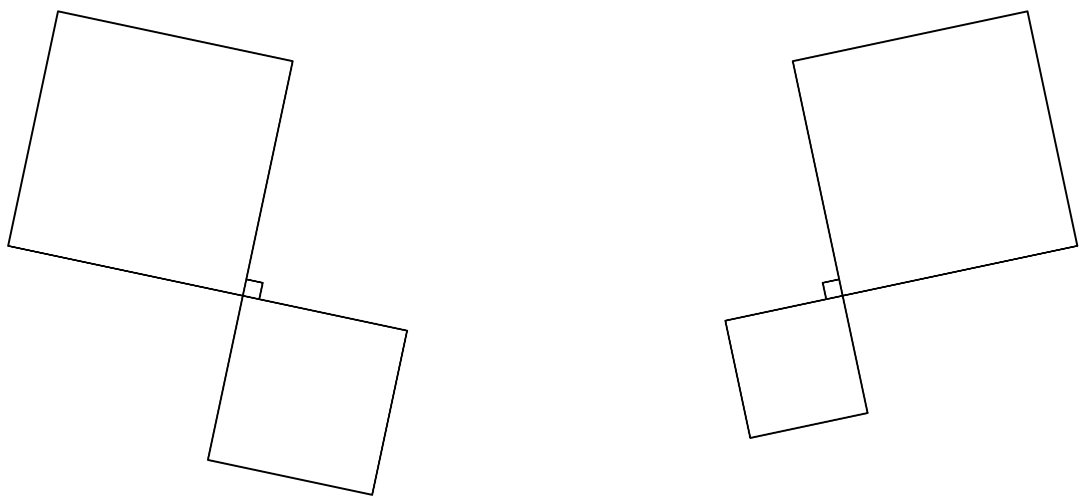


Solution sets of linear equations viewed in terms of Strang's figure

There is a version of Strang's figure for solution sets of consistent systems of linear equations. Let $\mathbf{A}$ be an $m \times n$ matrix and suppose that the linear system

$$\mathbf{Ax} = \mathbf{b}$$

is consistent for some $\mathbf{b}$ in $\mathbb{R}^m$.



This figure is one way of remembering the fact that the solution set of a consistent nonhomogeneous system consists of the sum of one particular solution of the nonhomogeneous system and the general solution of the associated homogeneous system (Theorem 6 in Chapter 1). It also suggests the second part of Exercise 23 in Section 6.1. That is, there is a unique $\mathbf{p}$ in Row $\mathbf{A}$ such that

$$\mathbf{Ap} = \mathbf{b}.$$

Symmetric matrices

Symmetric matrices arise frequently in applications. Moreover, they have a particularly nice structure that can often be used to solve the problem at hand. Today we discuss that structure.

**Definition.** A matrix $\mathbf{A}$ is symmetric if $\mathbf{A}^T = \mathbf{A}$.

**Example.** Consider

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 3 \end{bmatrix}.$$

With the aid of *Mathematica*, we see that $\mathbf{A}$ has three distinct real eigenvalues, $\lambda = 7$, $\lambda = 4$, and $\lambda = 1$. We also have three eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

We can diagonalize $\mathbf{A}$ using

**Theorem.** Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors associated to distinct eigenvalues of a symmetric matrix $\mathbf{A}$. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.