Administrative Details

Topics for Today
First Half
• (4.2) Projections

Second Half
• (4.3) Least Squares

First Half
Introduction
A geometric example again
Consider the matrix

\[ P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \]

What is the geometric interpretation of \( P \)? Looking at the column space helps.

What is the orthogonal complement of \( \text{col}(P) \)? What is the geometric explanation of this answer?

Projection onto a line

Derivation
Let \( \vec{a} \) be a line that goes through the origin. Our goal is to find the point \( \vec{p} \) on the line \( \vec{a} \) closest to some arbitrary point \( \vec{p} \). What else do we need for the projection?
**Example**
Compute the orthogonal projection of \((0, 1, 1)\) onto \((0, 2, 0)\).

**Example**
Compute the orthogonal projection of \((0, 1, 1)\) onto \((0, 2, 2)\).

**Example**
Compute the orthogonal projection of \((0, 1, 1)\) onto \((2, 2, 2)\).

**The Projection Matrix**
Rewrite the projection formula to use a matrix for the last example. What happens if we project twice?
A Generalization: Projection onto a subspace

Introduction

The work in the last section can be generalized to vector spaces with a dimension higher than one. This problem may be stated as follows:

Given \( n \) linearly independent vectors \( \{\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\} \subset \mathbb{R}^n \), find the combination

\[
\hat{x}_1\vec{a}_1 + \hat{x}_2\vec{a}_2 + \cdots + \hat{x}_n\vec{a}_n
\]

closest (in the sense of the orthogonal projection) to the given vector \( \vec{b} \).

The key to solving this is to keep the error, \( \vec{b} - A\vec{x} \), orthogonal to the subspace described by the \( \vec{a}s \).

Secondary Results

Why do we usually care about the case where \( A \) is rectangular? What happens if \( A \) is square?
Example

Find a projection of the vector

\[ \vec{b} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \]

onto the plane through the origin \( x + y + z = 0 \).
Second Half

Least Squares: A motivation

Consider the heat transfer equation from last time:

\[ Q = hA(\theta_0 - \theta_\infty) \]

Pretend that you are a chemical engineer attempting to calculate the heat transfer coefficient \( h \) in this equation. You have measured \( Q \) as a function of \( \theta_0 \) for five hundred values of \( \theta_0 \). \( A \) and \( \theta_\infty \) are known constants. How do you use the data to calculate \( h \)? If we try to put the data in a matrix, then what will we get?

A More Immediate Example

This is nearly the same example Strang uses to begin the chapter. Fit a line to the following three points:

\[ (0, 0), (1, 2), \text{ and } (2, 4). \]

This equation has no solution, but we can find a best solution by projecting.

The summary to this method is that if the equation \( Ax = \vec{b} \) is not solvable because it is overconstrained, then you should solve \( A^T \vec{Ax} = A^T \vec{b} \) instead.
An Algebraic Derivation

Fitting a more complicated set of data

The great strength of least squares is that the problem only needs to be linear in the coefficients for which we solve, not the underlying equation.

A common technique in chemistry is to build a calibration curve that relates known input and known output. For example, one can calculate lead concentration in water from spectroscopy where both the absorption $A$ and the concentration $C$ are known (from prepared samples) by a relationship like

$$C = xA^2 + yA + z.$$  

Assume you have a table of (several hundred) known absorption and concentration values. How may you find $x, y$, and $z$?