SI231b: Matrix Computations

Lecture 9: Least Squares and Orthogonal Projection

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Oct. 10, 2022

Brief Overview

Overdetermined System: $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ $(m > n)$, the equation

- often has no solution, since

$$\mathbf{b} \in \mathbb{R}^m, \quad \text{while } \mathcal{R}(\mathbf{A}) \text{ is a subspace (at most of dimensional } n\text{) of } \mathbb{R}^m$$

- has unique solution when

$$\mathbf{b} \in \mathcal{R}(\mathbf{A}) \text{ and } \operatorname{rank}(\mathbf{A}) = n$$

- has infinite solutions when

$$\mathbf{b} \in \mathcal{R}(\mathbf{A}) \text{ and } \operatorname{rank}(\mathbf{A}) < n$$

In practice, we need to find the full rank least square (LS) solution $\mathbf{x}_{LS}$,

$$\mathbf{x}_{LS} = \arg\min \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2,$$

where $\|\cdot\|_2$ represents the vector 2-norm and $\mathbf{A}$ is full rank.

Outline

- Motivation Applications
- Geometric Interpretation of Least Square
- Projection onto Subspaces
- Orthogonal Projection

In many applications, we can use the following representation

$$\mathbf{y} = \mathbf{A}\mathbf{x},$$

or

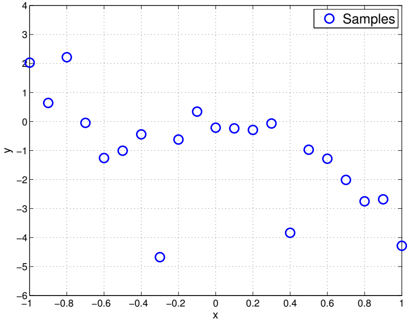
$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v},$$

where

- $\mathbf{y}$ is known (given data);
- $\mathbf{A}$ is given or stipulated;
- $\mathbf{x}$ is to be determined;
- $\mathbf{v}$ models the noise or error.

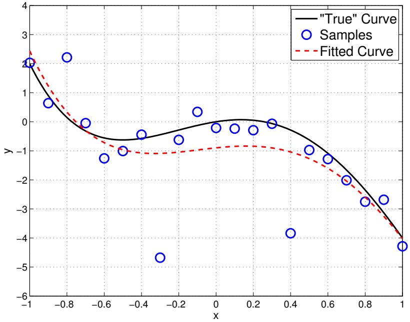
Data Fitting

Given a set of input-output data pairs $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, \ldots, m$, find a function $f(x)$ that fits the data well.



Data Fitting Using Polynomials

Applying a polynomial model $f(x) = \sum_{i=0}^{p} a_i x^i$ and use LS

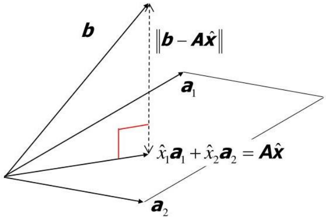


"True" curve: the true $f(x)$, $p = 5$.

Fitted curve: estimated $f(x)$, $\mathbf{a}$ obtained by LS with $p = 5$.

$$\mathbf{x}_{LS} = \arg\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2$$

1. find $\tilde{\mathbf{b}} \in \mathcal{R}(\mathbf{A})$ such that $\|\mathbf{b} - \tilde{\mathbf{b}}\|_2$ is minimized
   - recall the distance between two vectors using vector norms
2. solve $\mathbf{A}\mathbf{x}_{LS} = \tilde{\mathbf{b}}$ to obtain $\mathbf{x}_{LS}$



Question: how to obtain $\tilde{\mathbf{b}} \in \mathcal{R}(\mathbf{A})$?

Projectors

A projector is a square matrix that satisfies

$$\mathbf{P}^2 = \mathbf{P}.$$

- such a matrix is called idempotent
- geometric interpretation?

Note: this definition of projectors include both

- orthogonal projectors (key in our course)
- oblique projectors (will not be addressed)

Question: onto which subspace does $\mathbf{P}$ project?

Answer: $\mathcal{R}(\mathbf{P})$

How to distinguish orthogonal and oblique projection?

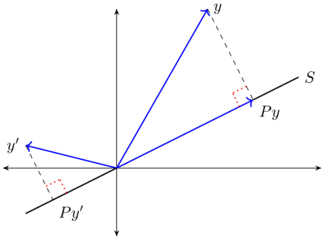


Figure 1: orthogonal projection

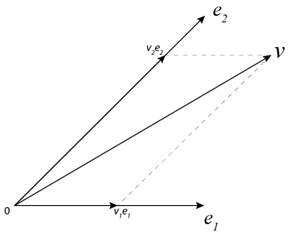


Figure 2: oblique projection

Answer: the projection direction $\mathbf{Pv} - \mathbf{v}$

Note: $\mathbf{P}(\mathbf{Pv} - \mathbf{v}) = \mathbf{0}$, which means $(\mathbf{Pv} - \mathbf{v}) \in \mathcal{N}(\mathbf{P})$.

If $\mathbf{P}$ is a projector, then $\mathbf{I}-\mathbf{P}$ is also a projector (why?)

$$(\mathbf{I}-\mathbf{P})^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P}^2 = (\mathbf{I}-\mathbf{P})$$

The projector $\mathbf{I}-\mathbf{P}$ is called the complementary projector of $\mathbf{P}$.

Question: onto which subspace does $\mathbf{I}-\mathbf{P}$ project?

Answer: $\mathcal{N}(\mathbf{P}) = \mathcal{R}(\mathbf{I}-\mathbf{P})$

First, $\mathcal{N}(\mathbf{P}) \subset \mathcal{R}(\mathbf{I}-\mathbf{P})$ (give your explanation here)

Second, $\mathcal{R}(\mathbf{I}-\mathbf{P}) \subset \mathcal{N}(\mathbf{P})$ (you are supposed to work it out indepently)

Then,

$$\mathcal{R}(\mathbf{I}-\mathbf{P}) = \mathcal{N}(\mathbf{P}) \text{ and } \mathcal{R}(\mathbf{P}) = \mathcal{N}(\mathbf{I}-\mathbf{P})$$

$$\mathcal{R}(\mathbf{P}) \cap \mathcal{N}(\mathbf{P}) = \{\mathbf{0}\}$$

Projection onto Subspaces

Suppose $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, then there is a projector $\mathbf{P}$ such that $\mathcal{R}(\mathbf{P}) = \mathcal{U}$ and $\mathcal{N}(\mathbf{P}) = \mathcal{W}$, we say that $\mathbf{P}$ is a projector onto $\mathcal{U}$ along $\mathcal{W}$.

Previous analysis show that the projector $\mathbf{P} \in \mathbb{R}^{m \times m}$ separates $\mathbb{R}^m$ into two subspaces

- $\mathcal{R}(\mathbf{P})$
- $\mathcal{N}(\mathbf{P})$

and

$$\mathbb{R}^m = \mathcal{R}(\mathbf{P}) \oplus \mathcal{N}(\mathbf{P}) \quad \text{can you prove this?}$$

$\mathbf{P}$ projects $\mathbb{R}^m$ onto $\mathcal{R}(\mathbf{P})$ along $\mathcal{N}(\mathbf{P})$.

Orthogonal projector

An orthogonal projector $\mathbf{P}$ is the one that projects onto a subspace $\mathcal{U}$ along a subspace $\mathcal{W}$ when $\mathcal{U}$ and $\mathcal{W}$ are orthogonal.

Warning: orthogonal projectors are not orthogonal matrices.

Theorem

A projector $\mathbf{P}$ is orthogonal if and only if $\mathbf{P} = \mathbf{P}^T$.

Proof ?