Linear Algebra
[KOMS120301] - 2023/2024

15.3 - Applications

Dewi Sintiari

Computer Science Study Program
Universitas Pendidikan Ganesha

Week 15 (December 2023)

Learning objectives

After this lecture, you should be able to:

1. explain the concept of "Least Square Problem";
2. compute the least square solution, error vector, and the least square error;
3. explain how least square is applied to fit a polynomial curve to data;
4. explain how least square is applied to approximate a function.

Part 1: Least Square

What is least square? (1)

Let us given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations and n variables, that is not consistent due to errors in the entries of A or $\mathbf{b}$.

Possible solution $\rightarrow$ look for a vector $\hat{\mathbf{x}}$ that comes as "close as possible" to being a solution.

What is least square? (1)

Let us given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations and n variables, that is not consistent due to errors in the entries of A or $\mathbf{b}$.

Possible solution $\rightarrow$ look for a vector $\hat{\mathbf{x}}$ that comes as "close as possible" to being a solution.

Problem (Least Squares Problem)

Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n variables, find a vector $\mathbf{x}$ in $\mathbb{R}^n$ that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ w.r.t. the Euclidean inner product on $\mathbb{R}^m$.

Question: Can you explain why do we minimize $\|\mathbf{b} - A\mathbf{x}\|$?

What is least square? (1)

Let us given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations and n variables, that is not consistent due to errors in the entries of A or $\mathbf{b}$.

Possible solution $\rightarrow$ look for a vector $\hat{\mathbf{x}}$ that comes as "close as possible" to being a solution.

Problem (Least Squares Problem)

Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n variables, find a vector $\mathbf{x}$ in $\mathbb{R}^n$ that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ w.r.t. the Euclidean inner product on $\mathbb{R}^m$.

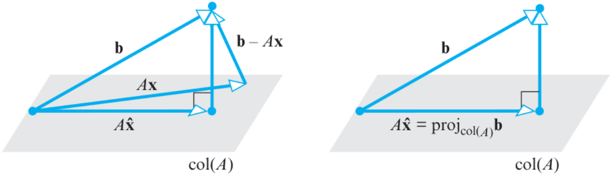
Question: Can you explain why do we minimize $\|\mathbf{b} - A\mathbf{x}\|$?

Terminology: $\mathbf{x}$ is called least squares solution, $\mathbf{b} - A\mathbf{x}$ is called least squares error vector, and $\|\mathbf{b} - A\mathbf{x}\|$ is called least squares error.

What is least square? (2)

Question: Can you explain why do we minimize $\|\mathbf{b} - A\mathbf{x}\|$?

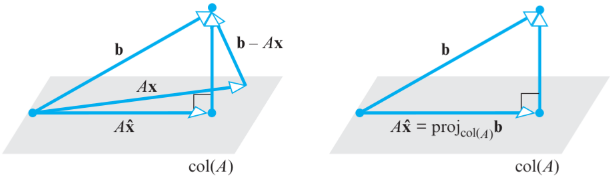
Finding solution of the least square $\mathbf{b} - A\mathbf{x}$ is equivalent to **finding the vector $A\hat{\mathbf{x}}$** in the column space A that is **the closest to $\mathbf{b}$**.



What is least square? (2)

Question: Can you explain why do we minimize $\|\mathbf{b} - A\mathbf{x}\|$?

Finding solution of the least square $\mathbf{b} - A\mathbf{x}$ is equivalent to **finding the vector $A\hat{\mathbf{x}}$** in the column space A that is **the closest to b**.

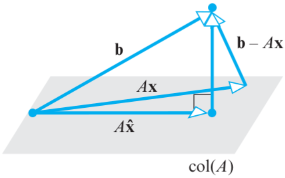


Question: Why is it named **"least square"**?

What is least square? (2)

Question: Can you explain why do we minimize $\|\mathbf{b} - A\mathbf{x}\|$?

Finding solution of the least square $\mathbf{b} - A\mathbf{x}$ is equivalent to **finding the vector $A\hat{\mathbf{x}}$** in the column space A that is **the closest to b**.



Question: Why is it named **"least square"**?

Let $\mathbf{b} - A\mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$. Hence: $\|\mathbf{b} - A\mathbf{x}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

How do we make the error $\mathbf{e} = \mathbf{b} - A\mathbf{x}$ to be as small as possible?

Theorem (Best Approximation: the vector closest to $\mathbf{b}$)

If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V. Then:

$$\|\mathbf{b} - proj_W \mathbf{b}\| \leq \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ in W, where $\mathbf{w} \neq proj_W \mathbf{b}$

How do we make the error $\mathbf{e} = \mathbf{b} - A\mathbf{x}$ to be as small as possible?

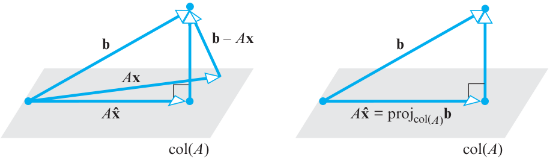
Theorem (Best Approximation: the vector closest to $\mathbf{b}$)

If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V. Then:

$$\|\mathbf{b} - proj_W\mathbf{b}\| \leq \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ in W, where $\mathbf{w} \neq proj_W\mathbf{b}$

$\rightarrow$ This means that $\text{proj}_W\mathbf{b}$ is the best approximation to $\mathbf{b}$ from W.



$\rightarrow$ A least square solution is a vector $\mathbf{x}$ satisfying: $A\mathbf{x} = \text{proj}_{\text{col}(A)}\mathbf{b}$.

How to obtain a solution of least square problem?

Theorem (The least square solutions)

For every linear system $A\mathbf{x} = \mathbf{b}$, the least square solutions of the system is given by <u>all solutions</u> of the system:

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

Moreover, if $\hat{\mathbf{x}}$ is any least square solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on the column space of A is:

$$proj_{col(A)}\mathbf{b} = A\hat{\mathbf{x}}$$

How to obtain a solution of least square problem?

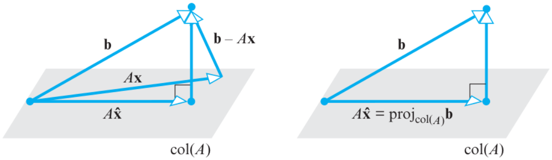
Theorem (The least square solutions)

For every linear system $A\mathbf{x} = \mathbf{b}$, the least square solutions of the system is given by all solutions of the system:

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

Moreover, if $\hat{\mathbf{x}}$ is any least square solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on the column space of A is:

$$proj_{col(A)}\mathbf{b} = A\hat{\mathbf{x}}$$



Exercise: Give a mathematical proof of correctness of first statement!

Conclusion

When $A\mathbf{x} = \mathbf{b}$ has no solution, then:

1. Multiply both sides of $A\mathbf{x} = \mathbf{b}$ by A^T, so we obtain:

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b} \tag{1}$$

2. Solve system (1), so we obtain all *least squares solutions* $\hat{\mathbf{x}}$.
3. Solve: $\mathbf{b} - A\hat{\mathbf{x}}$ to obtain the *error vector*.
4. Compute $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ to obtain the *error vector*.

Example: Unique least square solution

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

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

Example: Unique least square solution

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

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

Solution:

From the linear system, we can derive:

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

Example solution

1. **Finding solution:** Compute $A^T A$, $A^T \mathbf{b}$, and solve $A^T A\mathbf{x} = A^T \mathbf{b}$.

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

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

Now, solve linear system $A^T A\mathbf{x} = A^T \mathbf{b}$.

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

which yields a **unique least square solution**:

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

Example solution (*cont.*)

2. **Error vector:** It is given by $\mathbf{b} - A\mathbf{x}$, that is,

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

3. **Least square error:** It is given by $\|\mathbf{b} - A\mathbf{x}\|$.

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

Example solution (*cont.*)

2. **Error vector:** It is given by $\mathbf{b} - A\mathbf{x}$, that is,

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

3. **Least square error:** It is given by $\|\mathbf{b} - A\mathbf{x}\|$.

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

Question: Can you briefly explain the interpretation of this example?

Exercise: How many least square solutions?

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

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

How many least square solutions that you find?

Part 2: Applications

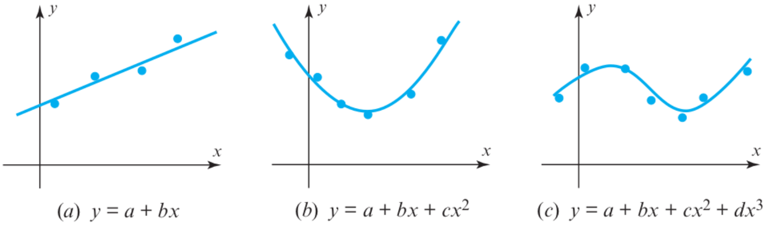
Application 1: Polynomial curve

Problem (Fitting a curve to data)

Given experimental values of pairs (x, y)*, namely:*

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

How to obtain a mathematical relationship $y = f(x)$*?*



Least squares fit for linear curve

Suppose we want to fit a straight line $y = a + bx$. Then we solve:

$$\begin{aligned} y_1 &= a + bx_1 \\ y_2 &= a + bx_2 \\ &\vdots \\ y_n &= a + bx_n \end{aligned}$$

which can be written in matrix form:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}$$

or as $A\mathbf{v} = \mathbf{y}$, where:

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}$$

Exercise: What about fitting a polynomial curve?

Suppose that we want to fit a polynomial function:

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

to n points:

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_n, y_n)$$

Hint:

1. Determine the system of n equations based on the data above.
2. Determine each component A, $\mathbf{v}$, and $\mathbf{y}$.
3. Solve the least squares.

Application 2: Function approximation

Problem (Approximation problem)

Given a function f that is continuous on an interval $[a, b]$, find the "best possible approximation" to f using only functions from a specified subspace W of $C[a, b]$.*

Example

Find the best possible approximation to:

1. e^x over $[0, 1]$ by a polynomial of the form $a_0 + a_1x + a_2x^2$
2. $\sin \pi x$ over $[-1, 1]$ by a function of the form:

$$a_0 + a_1e^x + a_2e^{2x} + a_3e^{3x}$$

3. x over $[0, 2\pi]$ by a function of the form:

$$a_0 + a_1 \sin x + a_2 \sin 2x + b_1 \cos x + b_2 \cos 2x$$

*the space of continuous functions on $[a, b]$

The mathematical meaning of "best possible approximation over $[a, b]$"

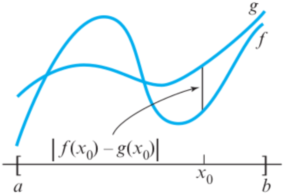


Figure 6.6.1 The deviation between f and g at x_0.

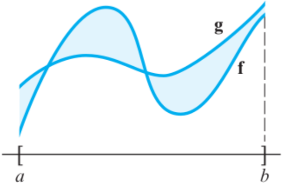


Figure 6.6.2 The area between the graphs of **f** and **g** over $[a, b]$ measures the error in approximating f by g over $[a, b]$.

Intuitive explanation