MSDM5004
Numerical Methods and Modeling in Science
Spring 2024

Lecture 3

Prof Yang Xiang
Hong Kong University of Science and Technology

Error bound

Theorem

Suppose $x_0, x_1, \ldots, x_n$ are distinct numbers in the interval $[a, b]$ and $f \in C^{n+1}[a, b]$. Then, for each x in $[a, b]$, a number $\xi(x)$ (generally unknown) between $x_0, x_1, \ldots, x_n$, and hence in (a, b), exists with

$$f(x) = P(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n)$$

where $P(x)$ is the interpolating polynomial given by

$$P(x) = f(x_0)L_{n,0}(x) + \cdots + f(x_n)L_{n,n}(x) = \sum_{k=0}^{n} f(x_k)L_{n,k}(x)$$

Chapter 4

Least Squares Fitting

An example

We would like to find a straight line given by

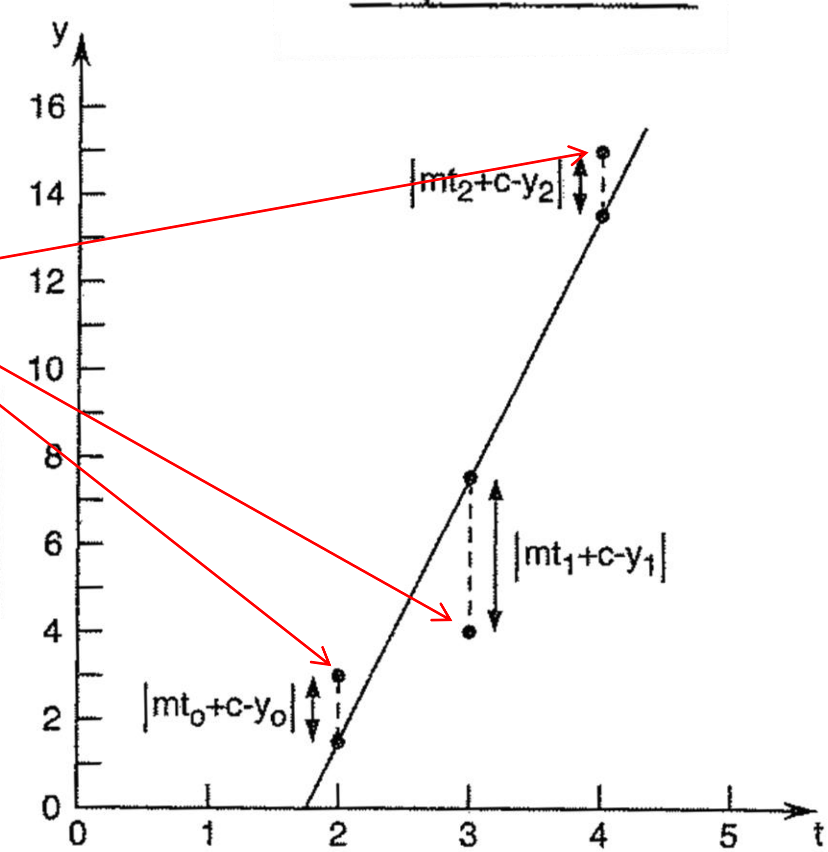
$$y = mt + c$$

that fits the experimental data.

Experimental Data :

i	0	1	2
t_i	2	3	4
y_i	3	4	15

The three data points are not exactly on a straight line.



This is a <u>linear least-squares problem</u>:

linear in unknown parameters *m* and *c*

We can represent our problem as a system of three linear equations of the form

$$
\begin{aligned}
2m + c &= 3 \\
3m + c &= 4 \\
4m + c &= 15.
\end{aligned}
$$

We can write this system of equations as

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

where

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 3 \\ 4 \\ 15 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} m \\ c \end{bmatrix}.$$

This linear system has no exact solution.

The straight line of best fit is the one that minimizes

$$\|Ax - b\|^2 = \sum_{i=0}^{2} (mt_i + c - y_i)^2 .$$

Solution: To minimize

$$f(m, c) = \|A\mathbf{x} - \mathbf{b}\|^2 = \sum_{i=0}^{2} (mt_i + c - y_i)^2.$$

At the minimizer, we have

$$\begin{cases} \dfrac{\partial f}{\partial m} = \displaystyle\sum_{i=0}^{2} 2t_i(mt_i + c - y_i) = 0 \\ \dfrac{\partial f}{\partial c} = \displaystyle\sum_{i=0}^{2} 2(mt_i + c - y_i) = 0. \end{cases} \quad \begin{cases} 29m + 9c = 78 \\ 9m + 3c = 22. \end{cases} \quad \begin{cases} m = 6 \\ c = -\dfrac{32}{3}. \end{cases}$$

$$\boxed{\text{vector } \mathbf{x} = (x_1, x_2, \cdots, x_n)^T, \text{ its } (l_2) \text{ norm } \|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots x_n^2}.}$$

Least-squares fitting

Least-squares fitting problem: To minimize

$$\|A\mathbf{x} - \mathbf{b}\|^2.$$

Solution: Let

$$f(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2 = (A\mathbf{x} - \mathbf{b})^T(A\mathbf{x} - \mathbf{b}).$$

The minimizer satisfies

$$\mathbf{0} = \nabla f(\mathbf{x}) = 2A^T(A\mathbf{x} - \mathbf{b}) = 2A^T A\mathbf{x} - 2A^T\mathbf{b}.$$

or

$$(A^T A)\mathbf{x} = A^T\mathbf{b}.$$

Normal Equations

Thus the minimizer is

$$\mathbf{x}^* = (A^T A)^{-1}A^T\mathbf{b}.$$

In the example

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 3 \\ 4 \\ 15 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} m \\ c \end{bmatrix}.$$

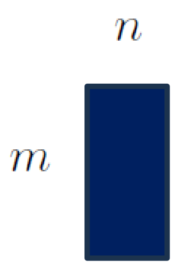
$$A^T A = \begin{pmatrix} 29 & 9 \\ 9 & 3 \end{pmatrix} \qquad A^T \mathbf{b} = \begin{pmatrix} 78 \\ 22 \end{pmatrix}$$

Using the least-squares solution formula, the fitting is

$$\mathbf{x}^* = \begin{pmatrix} m^* \\ c^* \end{pmatrix} = (A^T A)^{-1} A^T \mathbf{b} = \begin{pmatrix} 6 \\ -\frac{32}{3} \end{pmatrix}.$$

This lemma guarantees that the matrix product is invertible in the least-squares solution formula

Lemma *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$, $m \geq n$. *Then,* $\operatorname{rank} \mathbf{A} = n$ *if and only if* $\operatorname{rank} \mathbf{A}^\top \mathbf{A} = n$ *(i.e., the square matrix* $\mathbf{A}^\top \mathbf{A}$ *is nonsingular).* □



Example. Fit the data points $(2,1)$, $(3,6)$, $(5,4)$, and $(7,-15)$ using the quadratic model

$$f(t) = a_0 + a_1 t + a_2 t^2.$$

Solution. The fitting requires

This is also a linear least-squares problem: linear in unknown parameters a_0, a_1,a_2

$$\begin{cases} a_0 + 2a_1 + 2^2 a_2 = 1 \\ a_0 + 3a_1 + 3^2 a_2 = 6 \\ a_0 + 5a_1 + 5^2 a_2 = 4 \\ a_0 + 7a_1 + 7^2 a_2 = -15 \end{cases}$$

We write it as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 7 & 49 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 6 \\ 4 \\ -15 \end{pmatrix}.$$

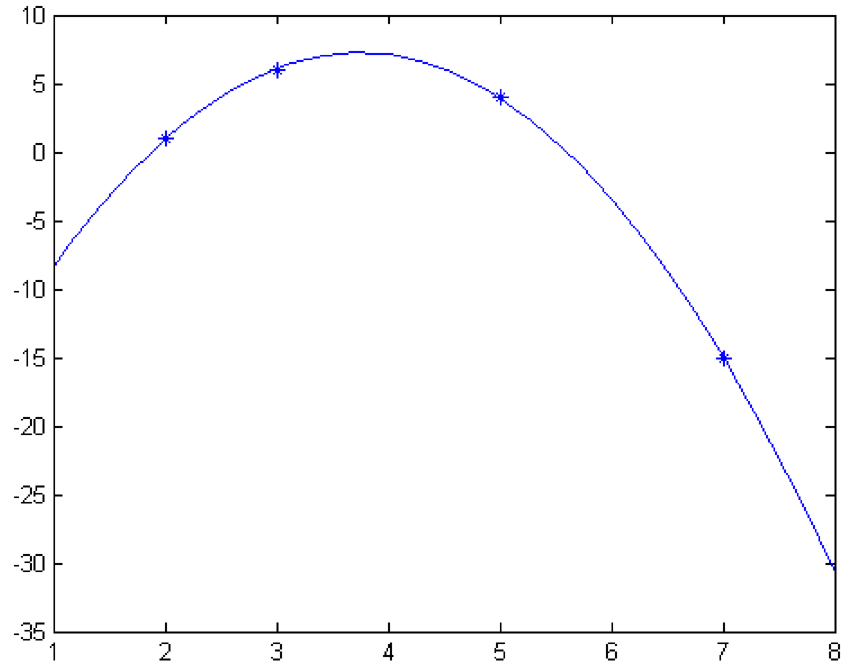
The least-squares solution is

$$\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b} = \begin{pmatrix} -21.9422 \\ 15.6193 \\ -2.0892 \end{pmatrix}.$$

```
%Least-squartes fitting
a=[1 2 4;1 3 9;1 5 25;1 7 49];
b=[1;6;4;-15];
xstar=inv(a'*a)*(a'*b)
%
%Generate the figure
t=linspace(1,8,50);
a0=xstar(1);
a1=xstar(2);
a2=xstar(3);
f=a0+a1*t+a2*t.^2;
plot(t,f)
hold on
plot(2,1,'*')
plot(3,6,'*')
plot(5,4,'*')
plot(7,-15,'*')
hold off
```

Plot the curve

Plot the data points



Chapter 5

Singular Value Decomposition (SVD)

Review

A is an $n \times n$ matrix

If A has n linearly independent eigenvalues, then there exists an invertible matrix P and a diagonal matrix D, such that

$$A = PDP^{-1}.$$

If A is symmetric, then there exists an orthogonal matrix Q and a diagonal matrix D, such that

$$A = QDQ^T.$$

$n \times n$ matrix Q is an orthogonal matrix: $Q^{-1} = Q^T$.

Definition of SVD

A is an $m \times n$ matrix

Its singular value decomposition (SVD) is

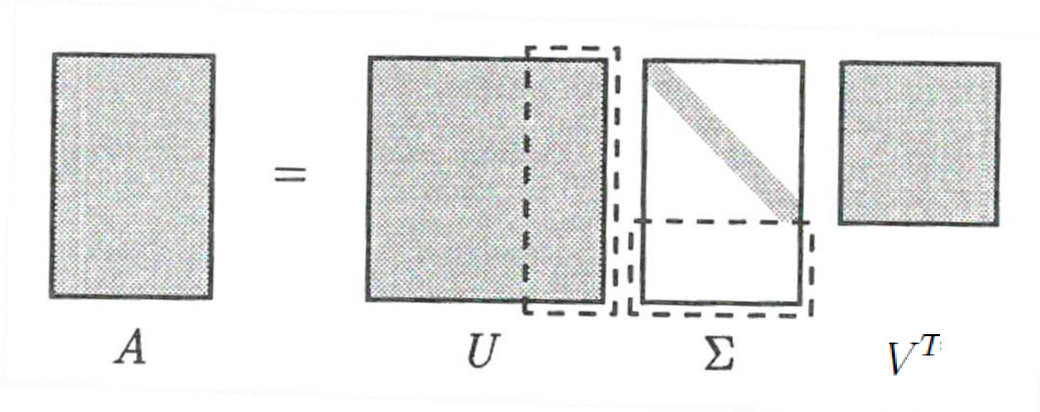
$$A = U\Sigma V^T,$$

where U is an $m \times m$ orthogonal matrix,

V is an $n \times n$ orthogonal matrix,

Σ is an $m \times n$ rectangular diagonal matrix whose diagonal entries are nonnegative and in nonincreasing order.

In the case of $m \geq n$



$$\Sigma = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & s_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

Each s_i is called a singular value of A.

From $A = U\Sigma V^T$, we have

$$AV = U\Sigma$$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \cdots & u_m \end{array}\right] \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & s_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

$$Av_j = s_j u_j, \qquad j = 1, 2, \cdots, n$$

It can be written as

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \cdots & u_n \end{array}\right] \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_n \end{bmatrix}$$

Reduced SVD

In the case of $m \geq n$

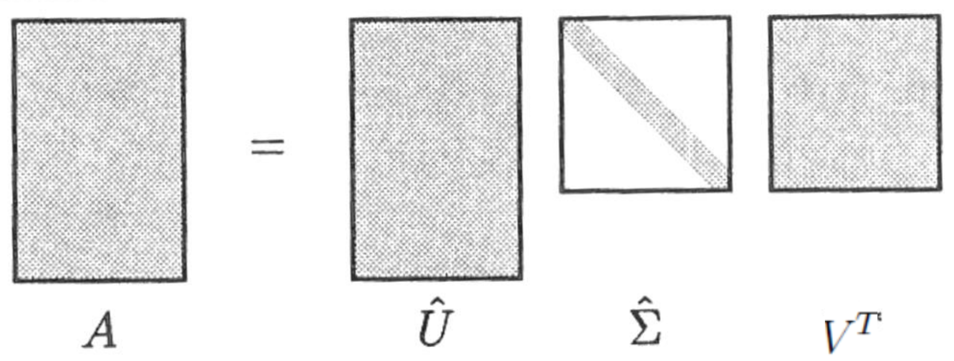
Reduced SVD

$$A = \hat{U}\hat{\Sigma}V^T,$$

where $\hat{U}$ is an $m \times n$ matrix with orthogonal columns,

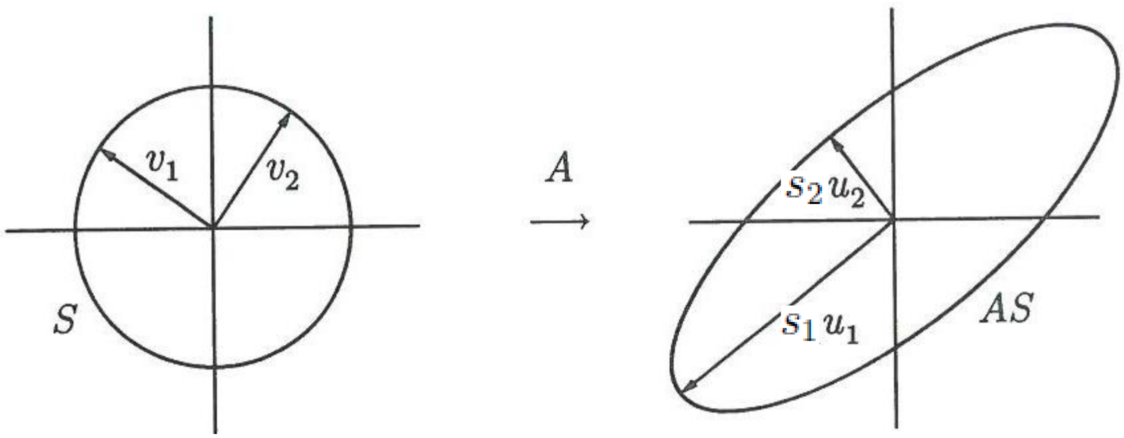
V is an $n \times n$ orthogonal matrix,

$\hat{\Sigma}$ is an $n \times n$ diagonal matrix whose diagonal entries are nonnegative and in nonincreasing order.



Geometric idea of SVD

$$Av_j = s_j u_j, \quad j = 1, 2, \cdots, n$$



Two principal semiaxes of AS