Fall 2022 ME424 Modern Control and Estimation

Lecture Note 3
Least Squares and Basic System Identification

Prof. Wei Zhang
Department of Mechanical and Energy Engineering
SUSTech Institute of Robotics
Southern University of Science and Technology

zhangw3@sustech.edu.cn
https://www.wzhanglab.site/

Outline

- **Least-Squares Problem Formulation**

- Solution to Linear Least-Squares Problems

- Linear Least-Squares Examples

- Applications to System ID

- Nonlinear Least Squares

- Last lecture: obtain discrete-time linear state space model from
  - physical process
  - given continuous time state space model
  - given nonlinear state space model
  - given discrete time transfer function

- The goal of this lecture note:
  - learn how to build model based on observed input-output data pairs
    - General case beyond the scope of this course
    - Focus on special case, where first obtain transfer function model from input-output data pairs, and then obtain the corresponding state space model

  - Main method: **Least Squares**

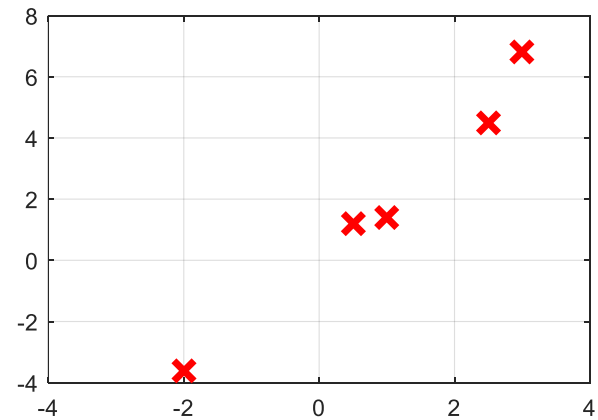# Least-Squares Problem Formulation:

- Measurement Equation:

$$y = g(\theta) + v$$

- $y \in R^m$: measurements data

- $\theta \in \Theta \subseteq R^n$: parameter to be estimated, where Θ is the constraint set for feasible parameters

- $v \in R^m$: unknown measurement noise

- $g: R^n \to R^m$: known (possibly) nonlinear function relates θ with measurement y

- Least Squares Example:

i	1	2	3	4	5
x	1	0.5	-2	3	2.5
y	1.4	1.2	-3.6	6.8	4.5

Least-Squares Problem Formulation:

- **Problem Statement**: Find the best parameter in the constraint set Θ that minimizes the difference between the model and the measured data

$$\min_{\theta \in \Theta} J(\theta) = \min_{\theta \in \Theta} ||y - g(\theta)||^2$$

- **Linear Least Squares**: $g(\theta) = H\theta$, where $H \in R^{m \times n}$ is a given deterministic matrix

Outline

- Least-Squares Problem Formulation

- **Solution to Linear Least-Squares Problems**

- Linear Least-Squares Examples

- Applications to System ID

- Nonlinear Least Squares

Optimization of multivariable function

- 1st –order necessary condition for optimality of $J(\theta)$

- Matrix calculus:

  - If $f: \mathbb{R}^n \to R^m$, then $\dfrac{\partial f}{\partial x}(\mathrm{x}) = Df(x) =$

Optimization of multivariable function

- **Gradient**: For scaler valued multivariate function $f : R^n \rightarrow R$, its gradient is defined as:

- For $f : R^n \rightarrow R$, notational convention $\nabla f(x) = \left(\frac{\partial f}{\partial x}(x) \right)^T$

- Some references use $\frac{\partial f}{\partial x}$ to denote gradient

- Directional derivative: $Df(x; d) = \lim\limits_{\alpha \to 0} \frac{f(x+\alpha d) - f(x)}{h}$

- Some calculus examples:
  - $f(x) = Ax$

  - $f(x) = x^T A \, x$

- **Exercise**: compute $\frac{\partial f}{\partial x}(x)$, where $x \in R^n$, and $f(x) = x^T x \cdot x$

- Derivation of linear least square solutions
  - Normal equation:

Solution with full rank H

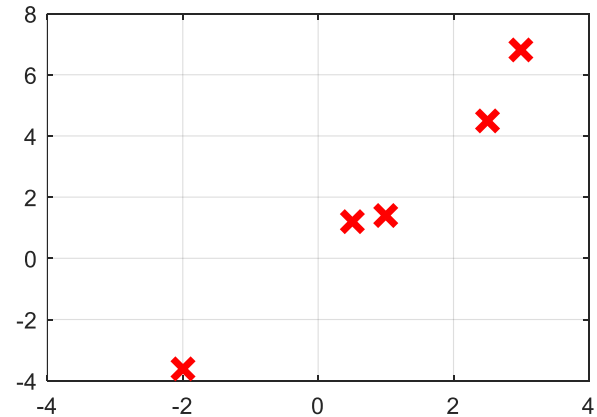# Geometric interpretation of linear least squares

Outline

- Least-Squares Problem Formulation

- Solution to Linear Least-Squares Problems

- **Linear Least-Squares Examples**

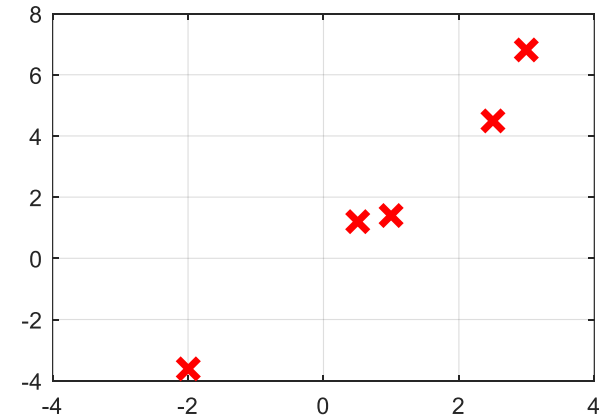- Applications to System ID

- Nonlinear Least Squares

- Linear Least Squares Example:

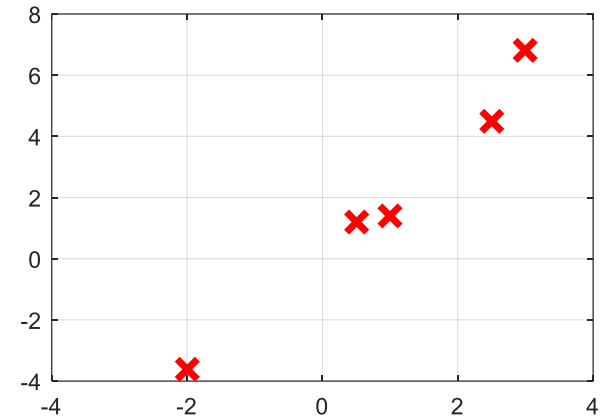| i	1	2	3	4	5
x	1	0.5	-2	3	2.5
y	1.4	1.2	-3.6	6.8	4.5

- Assume $y = \alpha x + \beta$

- Change hypothesis, assume $y = be^{ax}$

- Change hypothesis, assume that $y = \alpha_0 + \alpha_1 x + \alpha_2 x^2$

CLEAR Lab @ SUSTech

Outline

- Least-Squares Problem Formulation

- Solution to Linear Least-Squares Problems

- Linear Least-Squares Examples

- **Applications to System ID**

- Nonlinear Least Squares

Application to System Identification for Linear Systems

- ARX(p, q) model :(Autoregressive with exogenous input)

$$y(k) + \alpha_1 y(k-1) + \cdots + \alpha_p y(k-p)$$
$$= \beta_0 u(k) + \beta_1 u(k-1) + \cdots + \beta_q u(k-q) + v(k)$$

  - $v(k)$: noise signal

- Model parameter: $\theta = \left[\alpha_1, \ldots, \alpha_p, \beta_0, \beta_1, \ldots, \beta_q\right]^T$

- One-step predictor:

$$\hat{y}(k|\theta) = -\alpha_1 y(k-1) - \cdots - \alpha_p y(k-p)$$
$$+\beta_0 u(k) + \beta_1 u(k-1) + \cdots + \beta_q u(k-q)$$

- System ID problem for ARX model:

Given data pairs $\{(u(k), y(k))\}_{k \leq N}$, find the parameter vector θ that minimizes cost:

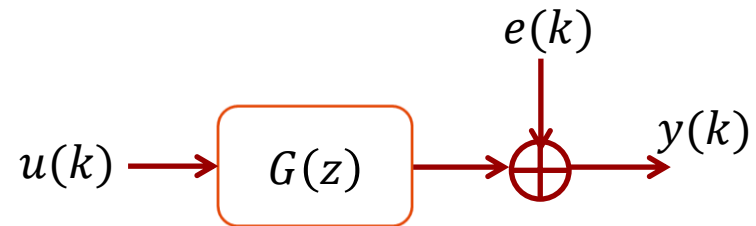- $J(\hat{\theta}) = \sum_{k=1}^{N} ||\hat{y}(k|\hat{\theta}) - y(k)||^2$

- Formulate as least square problem:

    given data set, $(u_1, y_1), (u_2, y_2), \ldots, (u_m, y_m)$

    - Regressor:

- Derivation continued

- **System ID Example I:**

$$G(z) = \frac{(z^2+b)}{z^3+az} \quad, \text{find best estimate for } a, b,$$

$$\text{given data set } (u_1, y_1), (u_2, y_2), \ldots, (u_{20}, y_{20})$$

- **System ID Example 2:**

$$e(k)$$

$$u(k) \longrightarrow \boxed{G(z)} \longrightarrow \bigoplus \longrightarrow y(k)$$

- $G(z) = \frac{z-1}{z-a}$, where a is an unknown scalar

- Data: $u(1) = 1, u(2) = \frac{1}{2}, u(3) = 1, y(1) = 2, y(2) = 1, y(3) = 2$

Outline

- Least-Squares Problem Formulation

- Solution to Linear Least-Squares Problems

- Linear Least-Squares Examples

- Applications to System ID
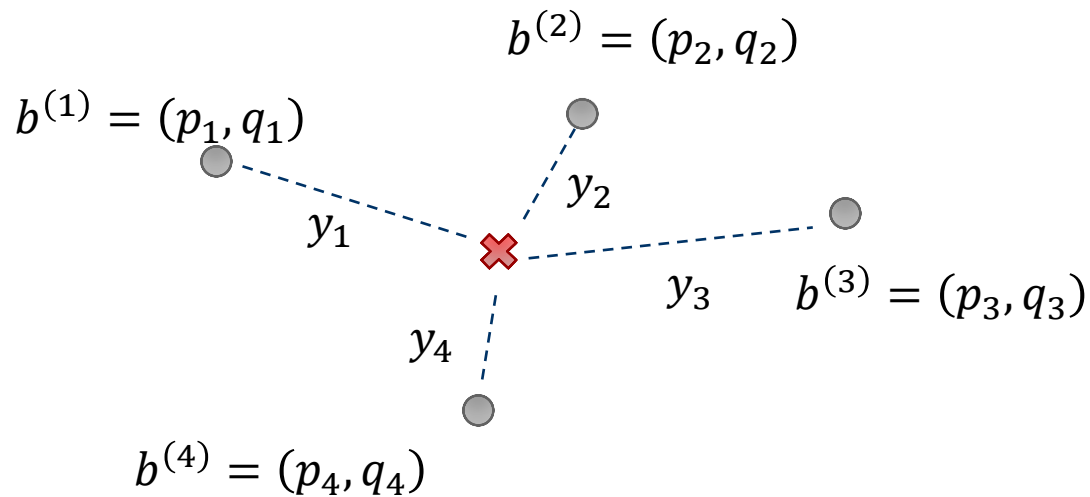
- **Nonlinear Least Squares**

- Nonlinear Least Squares:

$$\min_{\theta \in \Theta} J(\theta) = \left\|y - g(\theta)\right\|^2$$

  - For general nonlinear function $g(\theta)$, analytical solution to the above optimization is not available

  - Numerical optimization algorithms can be used to find the optimizer

$$\theta^* = argmin_{\theta \in \Theta} J(\theta)$$

- Nonlinear Least Square Example: Navigation by range measurement:

$$b^{(2)} = (p_2, q_2)$$

$$b^{(1)} = (p_1, q_1)$$

$$y_2$$

$$y_1$$

$$y_3 \qquad b^{(3)} = (p_3, q_3)$$

$$y_4$$

$$b^{(4)} = (p_4, q_4)$$

- ○ : beacons with known positions $b^{(i)} = (p_i, q_i)$
- ✖ : target with unknown position $\theta = (\theta_1, \theta_2)$
- y_i : known measured distance or range from beacon i:

  typical assumption: $y_i = \left\| b^{(i)} - \theta \right\| + v_i$

- Given measurements $y_1, y_2, \dots, y_m$, find the best target location θ

- We can choose cost function: $J(\theta) = \sum_{i=1}^{m} \left(y_i - \left| \left| b^{(i)} - \theta \right| \right| \right)^2$

- Coding Example