Our reference : UMT/FTKKI/100-54/30(6)
Date : 26 July 2022

Dr Julan Hernadi,
Department of Mathematics
Universitas Ahmad Dahlan
Kampus 4 UAD, Jalan Ringroad Selatan, Kragilan, Tamanan,
Banguntapan Bantul, Yogyakarta 56191,
Indonesia

Dr.,

INVITATION AS A SPEAKER FOR INTERNATIONAL CONFERENCE ON EMERGING TECHNOLOGIES AND SUSTAINABILITY (ICETS) 2022

The above matter is referred to.

2. Faculty of Ocean Engineering Technology and Informatics (FTKKI) will be hosting an International Conference on Emerging Technologies and Sustainability (ICETS) 2022, scheduled on 27 – 29 December 2022 at Universiti Malaysia Terengganu. The conference will gather researchers in various fields to share their knowledge and experiences. The theme of this conference is "Engaging minds, empowering technology and sustainability".

3. We would like to invite you as our speaker for one of the sessions during the conference. We would appreciate it if you could disseminate your knowledge with the audience in your respective field. Please let us know the title of your talk along with its respective abstract which should be emailed to fakhratulz@umt.edu.my.

4. We hope you will be able to fit this event in your busy schedule. The participants and organizers will benefit a lot from your expertise on the topic. It will be our greatest pleasure too if you can disseminate the information regarding this conference amongst your colleagues, many thanks in advance for your kind help.

Thank you.

Best regards,

Associate Professor Ts. Dr. Ahmad Nazri Bin Dagang
Chairman
International Conference on Emerging Technologies and Sustainability (ICETS) 2022
Faculty of Ocean Engineering Technology and Informatics
Universiti Malaysia Terengganu

s.k. - Sekretariat ICETS 2022
- Fail

International Conference on Emerging Technologies and Sustainability, UMT
December $27^{th}-29^{th}$ 2022, Terengganu, Malaysia.

The Optimal Measurement Strategy for Estimating Mathematical Model Parameters

Julan HERNADI
Department of Mathematics, Ahmad Dahlan University, Yogyakarta - Indonesia
julan.hernadi@mat.uad.ac.id

Abstract. In most cases, the mathematical model that describes a real-world phenomenon contains parameters that must be estimated using observational data. A small sample containing a lot of information about parameters is much more significant than another large sample but lacks information content. The presence of noise disturbance in the data observation due to an inaccurate instrument reading or a rounding mistake of the numerical data would disrupt the estimation quality. An intuitive consideration is that an observation data set can contain a lot of information about certain parameters but less for others. Determining the observation data that collectively provides the most information for all parameters is a major issue in the optimal design of measurement. This article implements the strategy of selecting observational data for estimating parameters of mathematical model using the sensitivity function and the optimal-criteria based on the Fisher Information Matrix (FIM).

Keywords: Optimal criteria, FIM, sensitivity function, inverse problem.

2020 Mathematics Subject Classification: 00A71; 35R30; 49K20; 62K05; 65L08.

DAY 3 PARALLEL SESSION (DECEMBER 29, 2022) – Morning session

PARALLEL ROOM 1
Date : 29 December 2022
Venue : Dewan Persidangan 2 (DP2)
Time : 9:55 am – 12:35 pm

*Link: https://umt.webex.com/umt/j.php?MTID=m84a64b3402f69c44c2bb60d2b4de5a70
Meeting number: 2510 238 3636
Password: Q3gaC9tyJK5

TIME	INVITED SPEAKER
9:55 am	Application of carbon-based materials in water treatment **Speaker: Dr. Thanh-Binh Nguyen***
10:40 am	The Optimal Measurement Strategy for Estimating Mathematical Model Parameters **Speaker: Dr. Julan Hernadi***

TIME	PAPER ID	PAPER TITLE/AUTHOR
11:35 am	8671*	Green Mobile Power Generation: Potential Design and Challenges **Nurliana Farhana Salehuddin, Wan Muhammad Faris Wan Ramli, Ridzuwan Mohd Jais, Noraziah Muda and Ahmad Faizal Ahmad Zamli**
11:50 am	1718*	Effect of the position of the drain in a Homogeneous Embankment Dams with a Vertical Drain **Abdelkader Djehiche, Mostefa Gafsi and Chaima Lalia Boumaàza**
12:05 pm	9158*	Simulation of a Ground-mounted PV System on the Unused Land of Mud Flow Area of Lumpur Lapindo, Indonesia **Elieser Tarigan**
12:20 pm	4286*	Case Study of the Effect of Tilt Angle and Temperature on Output Power and Efficiency of Photovoltaic Using Two Solar Irradiation Measurement Methods **Ojak Abdul Rozak, Mohd Zamri Ibrahim, Muhamad Zalani Daud, Tri Eka Octavian and Syaiful Bakhri**

PARALLEL ROOM 2
Date : 29 December 2022
Venue : Dewan Persidangan 1 (DP1)
Time : 9:55 am – 12:35 pm

TIME	INVITED SPEAKER

The Optimal Measurement Strategy for Estimating Mathematical Model Parameters

Julan HERNADI[1]

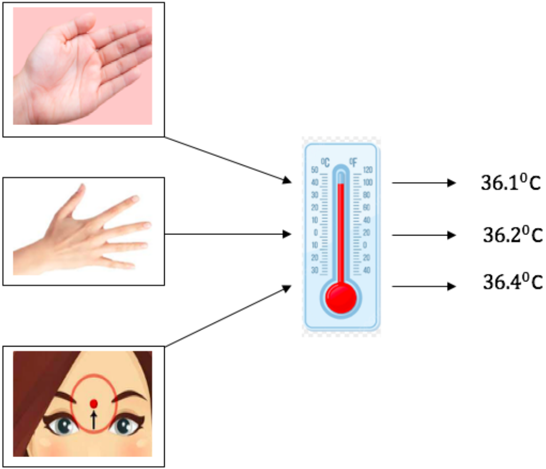
[1]Department of Mathematics and Scientific Computing
Universitas Ahmad Dahlan (UAD), Yogyakarta - Indonesia

International Conference on Emerging Technologies and Sustainability (ICETS 2022), Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 27-29 Dec 2022

Motivation

- A Mathematical models which describes some real phenomena mostly involve parameters that must be estimated by experiment data as a result of measurement.
- Parameter estimation refers to the determination from observed data of unknown parameters such that the predicted respons (system model) and the process observation are close.
- A postulate: Each experimental data set contains different information about parameters from one another.
- A small sample containing a lot of information about parameters is much more significant than another large sample but lacks information content.
- How to carry out measurement for obtaining sample containing the most information about the parameters?

Ilustration 1



Which measurement generates the most accurately of the body temperature?

Ilustration 2: Theoretical Model & Discrete Observation

Consider the explicit parameterized-model given by:

$$f(x,t;\theta) = \exp(-\theta\pi^2 t)\sin(2\pi x)\cos(20x)$$

and the observation given by:

$$z(t) = f(t,\theta_0) + \epsilon(t), t \in [0,T],$$

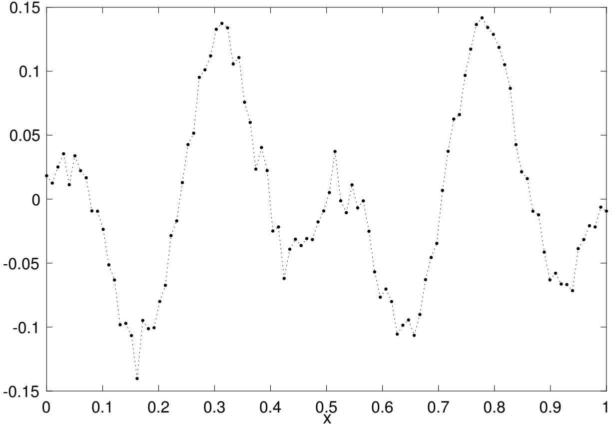
where θ_0 is the nominal parameter, the noise $\epsilon(t)$ is assumed to be zero mean with constant variance.

To estimates the nominal paramater θ_0, a sample $x_1, x_2, \cdots, x_n$ should be taken and the following SSE to be minimized:

$$\mathcal{J}(\theta) = \sum_{k=1}^{n} |f(x_k,t;\theta) - z_k(t)|^2 .$$

How to choose the such sample $x_1, x_2, \cdots, x_n$: how many and where the positions?

Example of measurements contaminated by noises



The graph was generated for $t = 0.2$ and $\theta_0 = 1$ and $\sigma = 0.01$. Let Ω be the candidate sensor location consists of 100 points evenly distributed on the interval $[0, 1]$.

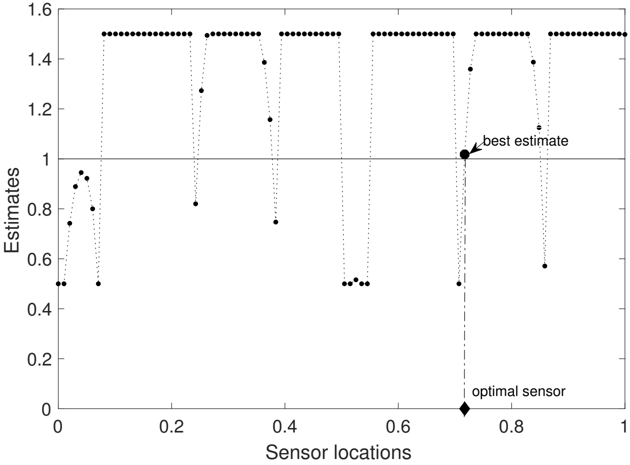
Estimates obtained from several sensor locations

Sensor	Points of measurements	Estimate $\hat{\theta}$
1	30 points uniformly	1.3580
2	10 first points	1.0009
3	10 last points	1.5000
4	$x_1 = 0.4545, x_2 = 0.5455, x_3 = 0.6162$	1.2610

Some facts:

- The number of points within sensors does not guarantee for a better estimate. For example, sensor 1 contains 30 points worse than sensor 2 with only 10 points. Sensor 3 and sensor 4 are also worse.
- Even, the best single point sensor is given by $x = 0.7172$ which produces $\hat{\theta} = 1.0818$.
- In case the nominal parameter θ_0 was not assumed beforehand, the quality of the estimate is measured by standard error of the estimates throughout the simulation.

Reliability of single point sensors



The minimum variance given by $x = 1.0818$ also the most accurate, i.e. the optimal single point sensor.

System description of mathematical model

- Lumped-parameter system (the nonlinear dynamical system):

$$\frac{dX}{dt} = F(t, X; \theta), X(t_0) = X_0, \tag{1}$$

$X : [t_0, t_f] \to \mathbb{R}^N$ state variable, $F : [t_0, t_f] \times \mathbb{R}^N \to \mathbb{R}^N$ some known function, $\theta = (\theta_1, \theta_2, \cdots, \theta_p)$ the parameter vector to be estimated.

- Distributed-parameter system:

$$\frac{\partial y}{\partial t} = \mathcal{F}\left(x, t, y, \frac{\partial y}{\partial x}, \frac{\partial^2 y}{\partial x^2}; \theta\right), x = (x_1, x_2) \in \Omega, t \in (0, t_f), \tag{2}$$

$y = y(x, t)$ the state variable, $\mathcal{F}$ some known function,
$\theta = (\theta_1, \theta_2, \cdots, \theta_p)$. The system accompanied by the BC and IC:

$$\mathcal{B}(x, t, y, \frac{\partial y}{\partial x}; \theta) = 0, x \in \partial\Omega, t \in (0, t_f) \text{ and } y(x, 0) = y_0(x), x \in \Omega.$$

The parameter model θ is estimated through measurement data.

Fisher Information Matrix (FIM) & Cramer-Rao Inequality

- Lumped-parameter system (Thomaseth & Cobelli, 1999):

$$M = \sum_{i=1}^{N} \frac{1}{\sigma^2(t_i)} \nabla_\theta^T f(t_i, \theta)\ \nabla_\theta f(t_i, \theta), \qquad (3)$$

where $f(t, \theta)$ the output model, $\nabla_\theta f := \frac{\partial f(t,\theta)}{\partial \theta}$ denotes the vector gradient of parameters.

- Distributed-parameter system (Ucinski, 2005):

$$M = \sum_{j=1}^{N} \frac{1}{\sigma^2(t)} \int_0^{t_f} \left(\frac{\partial y(x^j, t)}{\partial \theta} \right)^T \left(\frac{\partial y(x^j, t)}{\partial \theta} \right) dt \qquad (4)$$

- Cramer-Rao inequality:

$$\text{cov}(\hat{\theta}) \succeq M^{-1}, \qquad (5)$$

where $\succeq$ denotes the Löwner ordering of symmetric matrices, i.e. $\text{cov}(\hat{\theta}) - M^{-1}$ is a nonnegative definite.

Some Optimal Criteria (Quantitative Reference)

The objective is to find $t^* = \{t_1, t_2, \cdots, t_N\}$ for LPS or $x^* = \{x_1, x_2, \cdots, x_N\}$ for DPS which minimizes some real-valued function $\mathcal{J}$ defined on all possible FIMs.

- D-optimal, which maximizes the determinant of FIM, i.e.

$$\mathcal{J}(F) := -\ln \det(F)$$

  where "det" denotes the determinant of matrix.

- E-optimal, which maximizes the spectral radius of FIM, i.e.

$$\mathcal{J}(F) := \lambda_{\max}(F^{-1})$$

  where $\lambda_{\max}$ is the largest eigen value.

- A-optimal or SE-optimal, which maximizes the trace of FIM, i.e.

$$\mathcal{J}(F) := \mathrm{tr}(F^{-1})$$

  where "tr" stands for the trace of matrix, viz. the sum of entries on the main diagonal.

Example 1 - DPS

The IBVP describing the temperature distribution $y = y(x,t)$ of a rod thin (Hernadi, 2022):

$$\begin{aligned} \frac{\partial y}{\partial t}(x,t) &= \theta \frac{\partial^2 y}{\partial x^2}(x,t), \;\; x \in (0,1), \;\; t \in (0,t_f) && (6) \\ y(x,0) &= \sin(\pi x) && (7) \\ y(0,t) &= y(1,t) = 0, \;\; t \in (0,t_f) && (8) \end{aligned}$$

where the parameter θ stands for the diffusivity of the material forming the rod.

How to find x^1 the location which contains the most information about parameter θ?

The solution of model trivially given by $y(x,t) = e^{-\theta\pi^2 t}\sin(\pi x)$, and so the FIM is obtained as

$$M(x) = \underbrace{\frac{\pi^4}{\sigma^2}\int_0^{t_f} t^2 e^{-2\theta\pi^2 t}\,dt}_{>0}\,\sin^2(\pi x).$$

This FIM can be regarded as a matrix of single element.

Example 1 (cont...)

The FIM at sensor point x^j is written as

$$M(x^j) = k \sin^2(\pi x^j)$$

where $k = \frac{\pi^4}{\sigma^2} \int_0^{t_f} t^2 e^{-2\theta\pi^2 t} dt > 0$. It reaches maximum at

$$\pi x^j = \pi/2 \text{ or } x^j = 0.5.$$

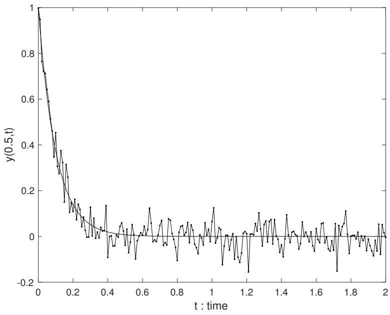
D-optimal criteria

$$\mathcal{J}(M(x)) = -\ln\det M(x)$$

attains the minimum at $x = 0.5$. Similarly argument for other optimal criteria.

Simulation

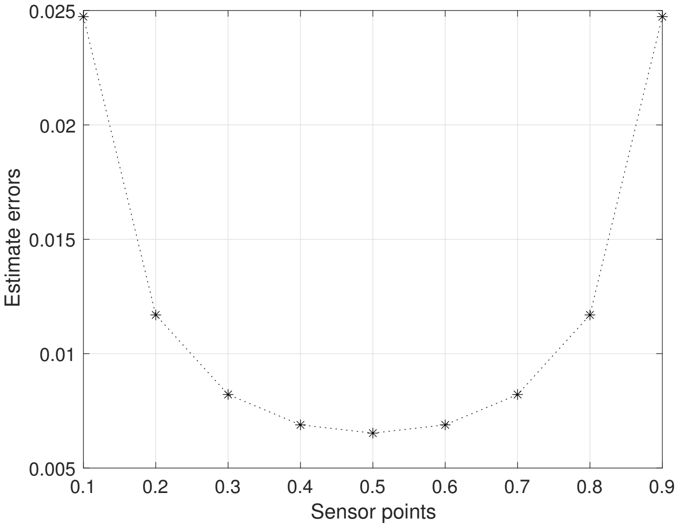
- It is assumed that $\theta_0 = 1$, $t_f = 2$, $\epsilon \sim N(0, 0.02)$.



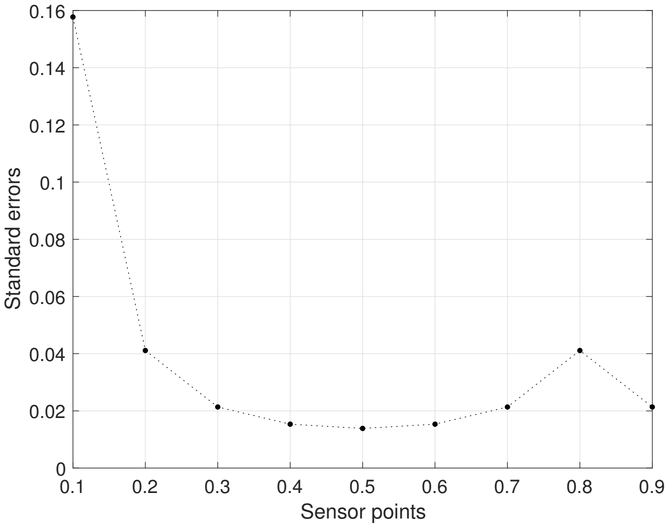
- Supposedly, $\Omega_0 = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ the sensor locations being observed.

The following experiment was performed as $L = 20$ repetitions where the noises were generated independently.

Estimate Errors



Standard Errors



Sensitivity Functions (Qualitative Reference)

The sensitivity functions, both traditional (TSF) and generalized (GSF), were frequently used as a reference for optimal measurement. Consider the dynamical model:

$$\frac{dX}{dt} = F(t, X; \theta), X(t_0) = X_0, \tag{9}$$

where $X_0 := [x_1(t_0), x_2(t_0), \cdots, x_N(t_0)]^T = [x_{01}, x_{02}, \cdots, x_{0N}]^T$ is the initial condition and $\theta = [\theta_1, \theta_2, \cdots, \theta_p]^T$ is the parameter system.

- The TSF are the derivatives of states with respect to parameter, i.e. $\frac{\partial X}{\partial \theta}$ and it is obtained from the sensitiviy equation:

$$\frac{d}{dt}\left(\frac{\partial X}{\partial \theta}\right) = \frac{\partial F}{\partial X}\frac{\partial X}{\partial \theta} + \frac{\partial F}{\partial \theta} \tag{10}$$

- A system of size N with p parameters will govern the sensitivity equation of size $N \times p$. In case the initial condition is also considered as parameter, there will be an additional equation of size $N \times N$.
- Challenges in solving the sensitivity equations: large sizes, the matrix function as unknown, and the dependence on the state equation.

Generalized Sensitivity Function (GSF)

Let $t_1, t_2, \cdots, t_n$ be the observation times, the FIM of (9):

$$F = \sum_{j=1}^{n} \frac{1}{\sigma^2(t_j)} \left[\nabla_\theta X(t_j;\theta_0)\right] \left[\nabla_\theta X(t_j;\theta_0)\right]^T, \tag{11}$$

where

$$\nabla_\theta X(t_j;\theta_0) = \begin{bmatrix} \frac{\partial X(t_j;\theta_0)}{\partial \theta_1} & \frac{\partial X(t_j;\theta_0)}{\partial \theta_2} & \cdots & \frac{\partial X(t_j;\theta_0)}{\partial \theta_p} \end{bmatrix}^T, \tag{12}$$

$$\frac{\partial X(t_j;\theta_0)}{\partial \theta_k} = \begin{bmatrix} \frac{\partial x_1(t_j;\theta_0)}{\partial \theta_k} & \frac{\partial x_2(t_j;\theta_0)}{\partial \theta_k} & \cdots & \frac{\partial x_N(t_j;\theta_0)}{\partial \theta_k} \end{bmatrix}^T. \tag{13}$$

The generalized sensitivity function (GSF) is defined as

$$G(t_\ell) = \sum_{j=1}^{\ell} \frac{1}{\sigma^2(t_j)} \left[F^{-1} \times \nabla_\theta(t_j;\theta_0)\right] \bullet \left[\nabla_\theta(t_j;\theta_0)\right], \tag{14}$$

where the notation "$\bullet$" stands for element-wise vector multiplication

Information provided by TSF, FIM, and GSF

- The TSF describe how model output trajectories change in response to modest changes in model parameters.
- The FIM measures the information content of the data corresponding to the model parameters (Banks, Holm, Kappel, 2011).
- The sharp increases of GSF indicate a high information about parameters (Thomaseth and Cobelli, 1999; Banks et all, 2011).

More detail, see Hernadi et al (2022) and references there in.

Implementation to competitive model

The simplest competitive model can be described as

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \frac{\sum_{j=1}^{N} \alpha_{ij} x_j}{K_i}\right), x_i(t_0) = x_{i0}, i = 1, 2, \cdots, N. \qquad (15)$$

For each $i, j = 1, 2, \cdots, N$:

- x_i : the size of species,
- r_i : the intrinsic growth,
- K_i : the carrying capacity, and
- $\alpha_{i,j}$: the impact of i^{th} species towards j^{th} species where $\alpha_{ii} = 1$ and $\alpha_{ij} > 0$ for $i \neq j$.
- Parameter sets: $\{r_i : i = 1, 2, \cdots, N\}$, $\{K_i : i = 1, 2, \cdots, N\}$, $\{\alpha_{i,j} : i, j = 1, 2, \cdots, N, i \neq j\}$, $\{x_{i0} : i = 1, 2, \cdots, N\}$. The number of sensitivity functions is $N^2(N+2)$.

The special case is when $N = 2$ which is well-known as the Lotka-Volterrra equation. The logistic growth model of Verhulst-Perl is this type for $N = 1$.

Numerical simulation

The simulation data are $N = 2$, $t_0 = 0$, $t_f = 150$, $x_{10} = 10$; $x_{20} = 30$, $r_1 = 0.1$, $r_2 = 0.3$, $a_{12} = 0.3$, $a_{21} = 0.25$, $K_1 = 150$, and $K_2 = 100$. We want to estimate parameters r_1 and r_2 through observation data contaminated by noises.

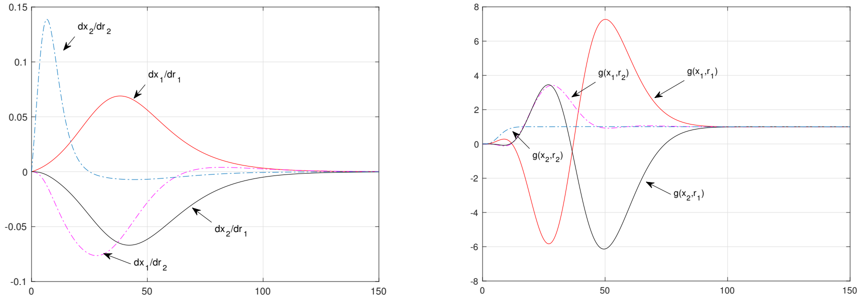


Figure: TSF $\frac{\partial X}{\partial r}$ (left) and GSF $g(X, r)$ (right)

- Measurements were taken from intervals $[0, 20]$ and $[90, 120]$ with six evenly distributed data points from each interval.
- Three criteria LS functions:

$$LS_1(\theta) = \sum_{j=1}^{n} |x_{1j} - y_{1j}|^2 \, , LS_2(\theta) = \sum_{j=1}^{n} |x_{2j} - y_{2j}|^2$$

$$LS_3(\theta) = \sum_{j=1}^{n} \left(|x_{1j} - y_{1j}|^2 + |x_{2j} - y_{2j}|^2\right)$$

For each $k = 1, 2$; $j = 1, 2, \cdots, n$, x_{kj} : the model outputs, and y_{kj}: observations including disturbance terms.

Domain of observation	Estimates $(\hat{r}_1, \hat{r}_2)$ given by LS_1	LS_2	LS_3
$[90, 120]$	$(0.100, 0.100)$	$(3.540, 18.343)$	$(0.0012, 0.095)$
$[0, 20]$	$(0.0986, 0.2344)$	$(0.0999, 0.3002)$	$(0.1000, 0.3007)$

Commentary

- The nominal parameters were set as $r_1 = 0.1$ and $r_2 = 0.3$.
- The observation data obtained from $[90, 120]$ contains insufficient information about the parameters $[90, 120]$ and $[90, 120]$. The curves of TSF and GSF are no longer changes after $t > 90$.
- Consider the estimates obtained from the interval $[0, 20]$. LS_1 provides a good estimate for r_1 but a poor estimate for r_2. LS_2 yields very good estimates for both r_1 and r_2. This means that the data obtained through observation of x_1 contains less information about parameter r_2 than x_2 about parameter r_1. The best result is given by LS_3. Thus, observation from interval $[0, 20]$ contains a lot of information about the parameters r_1 and r_2.
- The TSF and GSF curves are look like similar behavior, i.e. $\left|\frac{dx_1}{dr_2}\right| \ll \left|\frac{dx_2}{dr_1}\right|$ (left) and $|g(x_1, r_2)| \ll |g(x_2, r_1)|$ (right).
- It is confirmed that the information content of parameters delivered by observation data of states could be identified through their sensitivity functions.

Understanding the underlying mathematical model is essential in parameter estimation.

Example #1:
The following data represents the average monthly per capita income of ten randomly selected people from village X (e.g. in USD): 534, 440, 536, 582, 524, 552, 536, 485, 515, 461.

$$\mu_X = E(X) \approx \overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i = 516.5.$$

Example #2:
Similarly to Example #1, but the residents are from village Y: 150, 90, 200, 250, 175, 2500, 210, 3200, 400, 325.

$$\mu_Y = E(Y) \approx \overline{Y} = \frac{1}{10}\sum_{i=1}^{10} Y_i = 750.$$

Information given by Example #1 makes sense, but Example #2 is misleading.

Concluding Remarks

- Two kinds of parameter systems have been considered were the lumped parameter system (LPS) and the distributed parameter system (DPS). The optimal sensor (measurement strategy) is intended to collect data of high information content about the parameters being estimated.
- Two approaches for determining optimal sensor:
  - Quantitative methods: Taking minimum the real-valued function $\mathcal{J}$ defined on all possible FIMs, i.e. D-optimality (determinant), E-optimality (smallest eigenvalue), A-optimality (trace).
  - Qualitative methods: Traditional sensitivity function (TSF), Generalized sensitivity function (GSF).
- Computation isssues:
  - Solve the state equation and sensitivity equation simultaneously $\rightarrow$ large system with state the matrix function.
  - Apply the optimization method to solve the optimality criteria $\rightarrow$ ill-conditioned, complexity computation.

End of slides

Thank you very much

Wassalamu'alaikum Wr. Wb.