Sensor motion planning in distributed parameter systems using Turing’s measure of conditioning

Dariusz Uciński* and YangQuan Chen**

*Institute of Control and Computation Engineering
University of Zielona Góra, Poland
**Department of Electrical and Computer Engineering
Utah State University, Logan, USA
Motivations for the talk

- Calibration of smog prediction models
- Data assimilation in meteorology and oceanography
- Smart material systems
- Fault detection and isolation in DPSs
- Groundwater resources management
- Recovery of valuable minerals and hydrocarbon
- Inspection in hazardous environments
Sensor network – an array of sensors of diverse type interconnected by a communication network.

A large number of inexpensive, miniature and low-power SN nodes can be deployed throughout a physical space, providing dense sensing close to physical phenomena.

WSNs incorporate technologies from sensing, communication and computing.
Mobile Actuator–Sensor Network

Project developed in the Center for Self-Organizing and Intelligent Systems, Utah State University, Logan, USA.

- 10 small differentially-driven mobile robots
- COTS
- Mark III chassis
- MICA2 controller board
- Infrared sensors for obstacle avoidance
- Optical sensors for fog concentration detection
- Odometry
MAS-net: Global view
MAS-net: Tracking fog boundary
Let $\Omega \subset \mathbb{R}^2$ be a region with boundary $\partial \Omega$.

$$\frac{\partial y}{\partial t} = \mathcal{F}\left(x, t, y, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial^2 y}{\partial x_1^2}, \frac{\partial^2 y}{\partial x_2^2}, \theta \right), \quad \begin{cases} x \in \Omega \\ t \in T \end{cases}$$

subject to appropriate I & BCs, where

- $x$ – spatial variable, $t$ – time, $T = (0, t_f)$;
- $y = y(x, t)$ – state variable;
- $\mathcal{F}$ – some known function;
- $\theta \in \mathbb{R}^m$ – vector of unknown parameters.
System description

Let $\Omega \subset \mathbb{R}^2$ be a region with boundary $\partial \Omega$.

$$\frac{\partial y}{\partial t} = \mathcal{F} \left( x, t, y, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial^2 y}{\partial x_1^2}, \frac{\partial^2 y}{\partial x_2^2}, \theta \right), \quad \left\{ \begin{array}{l} x \in \Omega \\ t \in T \end{array} \right.$$  

subject to appropriate I & BCs, where

- $x$ – spatial variable, $t$ – time, $T = (0, t_f)$;
- $y = y(x, t)$ – state variable;
- $\mathcal{F}$ – some known function;
- $\theta \in \mathbb{R}^m$ – vector of unknown parameters.
Output equation

For pointwise moving sensors we have

\[ z_j(t) = y(x^j(t), t; \theta) + \varepsilon_j(t), \quad t \in T \]

for \( j = 1, \ldots, N \), where

\( \varepsilon_j(\cdot) \) – white Gaussian measurement noise.
Least-squares criterion

The LS estimate of $\theta$ is the one which minimizes

$$J(\theta) = \frac{1}{2} \sum_{j=1}^{N} \int_{0}^{t_f} \left[ z_j(t) - \hat{y}(x^j(t), t; \theta) \right]^2 dt$$

where $\hat{y}(\cdot, \cdot; \theta)$ stands for the solution to the state equation corresponding to a given value of $\theta$. 

Sensor motion planning in distributed parameter systems using Turing's measure of conditioning – p. 9/2
Determine sensor trajectories $x^1(\cdot), \ldots, x^N(\cdot)$ so as to maximize the identification accuracy.
Optimal measurement problem

Determine sensor trajectories $x^1(\cdot), \ldots, x^N(\cdot)$ so as to maximize the **identification accuracy**.
Optimal measurement problem

Determine sensor trajectories $x^1(\cdot), \ldots, x^N(\cdot)$ so as to maximize the identification accuracy.

Make use of the Cramér-Rao inequality:

$$\text{cov} \, \hat{\theta} = E\left\{ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right\} \geq M^{-1}$$

We have $\text{cov} \, \hat{\theta} = M^{-1}$ provided that an estimator is efficient!
Optimal measurement problem

Determine sensor trajectories $x^1(\cdot), \ldots, x^N(\cdot)$ so as to maximize the identification accuracy.

Make use of the Cramér-Rao inequality:

$$\text{cov} \hat{\theta} = E \left\{ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right\} \succeq M^{-1}$$

We have $\text{cov} \hat{\theta} = M^{-1}$ provided that an estimator is efficient! But what is $M$?
Average Fisher Inf. Matrix (FIM)

Up to a constant multiplier, we have

\[ M = \frac{1}{NT} \sum_{i=1}^{N} \int_{0}^{t_f} g(x^j(t), t) \ g^T(x^j(t), t) \ dt \]

where \( g(x, t) = \left( \frac{\partial y(x, t; \theta)}{\partial \theta} \right)^T \) is the sensitivity vector, \( \theta^0 \) is a preliminary estimate of \( \theta \).
Average Fisher Inf. Matrix (FIM)

Up to a constant multiplier, we have

\[ M = \frac{1}{NT} \sum_{i=1}^{N} \int_{0}^{t_f} g(x^j(t), t) g^T(x^j(t), t) \, dt \]

where \( g(x, t) = \left( \frac{\partial y(x, t; \theta)}{\partial \theta} \right)^T_{\theta=\theta^0} \) is the sensitivity vector, \( \theta^0 \) is a preliminary estimate of \( \theta \).
Ultimate formulation

Determine $x^j(\cdot)$, $j = 1, \ldots, N$ which minimize a scalar measure $\Psi$ defined on the FIM, e.g.,

$$\Psi_r(M) = \begin{cases} 
\left[ \frac{1}{m} \text{trace}(QM^{-1}Q^T) \right]^{1/r} & \text{if } \det(M) \neq 0 \\
\infty & \text{if } \det(M) = 0
\end{cases}$$
Determine $x^j(\cdot), j = 1, \ldots, N$ which minimize a scalar measure $\Psi$ defined on the FIM, e.g., minimize $\Psi_r(M) =$ \[ \begin{cases} \frac{1}{m} \text{trace}(QM^{-1}Q^T)^r \end{cases} \] if $\det(M) \neq 0$ 
\[ \infty \] if $\det(M) = 0$

$r \downarrow 0$ and $Q = I \Rightarrow$ D-optimum design

$J_D(x^1, \ldots, x^N) = \log \det(M(x^1, \ldots, x^N)) \longrightarrow \max$
Application of the D-optimality criterion may lead to very elongated confidence ellipsoids. A desirable characteristic of the solutions is that the confidence ellipsoids be as spherical as possible.
Application of the D-optimality criterion may lead to very elongated confidence ellipsoids. A desirable characteristic of the solutions is that the confidence ellipsoids be as spherical as possible.

Consider

$$\Upsilon[M] = \frac{1}{m} \sqrt{\text{trace}(M) \text{trace}(M^{-1})}$$

The minimum value above is unity, which is achievable only for spherical confidence regions.
Minimization of $\Upsilon(M)$ makes the Hessian $H(s)$ of the estimation cost well conditioned in the sense that it will yield minimization with respect to $s$ of the Frobenius condition number

$$\tilde{J}(s) = \sqrt{\text{trace}[H(s)] \text{trace}[H^{-1}(s)]}.$$
Minimization of $\Upsilon(M)$ makes the Hessian $H(s)$ of the estimation cost well conditioned in the sense that it will yield minimization with respect to $s$ of the Frobenius condition number

$$\tilde{\gamma}(s) = \sqrt{\text{trace}[H(s)] \text{trace}[H^{-1}(s)]}.$$ 

Yet direct use of Turing’s measure of conditioning only guarantees that the condition number is close to unity, while little information about the parameters may be gained.
Dynamics of sensor movements

Define the state vector

\[ s(t) = (x^1(t), x^2(t), \ldots, x^N(t)), \quad \forall \ t \in [0, t_f] \]

A general model

\[ \dot{s}(t) = f(s(t), u(t)) \quad \text{a.e. on } [0, t_f], \quad s(t_0) = s_0 \]

where \( s_0 \in \mathbb{R}^n \) is the initial sensor configuration, \( u : [0, t_f] \rightarrow \mathbb{R}^r \) is a control satisfying

\[ u_l \leq u(t) \leq u_u \quad \text{a.e. on } [0, t_f] \]
Constraints on sensor movements

The conditions of taking measurements only in a given admissible subdomain and keeping safe distances between the sensors are written as

\[ \gamma_\ell(s(t)) \leq 0, \quad \forall t \in [0, t_f], \quad \ell \in \bar{\nu} = \{1, \ldots, \nu\} \]

where \( \gamma_\ell : \mathbb{R}^n \to \mathbb{R} \) are \( C^1 \) functions. Then, the \textit{Fisher Information Matrix} is

\[ M(s) = \frac{1}{Nt_f} \sum_{j=1}^{N} \int_{0}^{t_f} g(x^j(t), t) g^T(x^j(t), t) \, dt \]
Sphericality as a critical factor

In the proposed formulation, we wish to guarantee an acceptable level of the information content of the collected observations, quantified by the D-efficiency:

\[ E_D(s) = \left\{ \frac{\text{det}(M(s))}{\text{det}(M(\hat{s}))} \right\}^{1/m} \]

where \( \hat{s} \) stands for the D-optimal trajectories obtained for the observations over the same feasible time interval.
Sphericality as a critical factor

In the proposed formulation, we wish to guarantee an acceptable level of the information content of the collected observations, quantified by the D-efficiency:

\[ E_D(s) = \left\{ \frac{\det(M(s))}{\det(M(\hat{s}))} \right\}^{1/m} \]

where \( \hat{s} \) stands for the D-optimal trajectories obtained for the observations over the same feasible time interval.
The value of $\det(M(\hat{s}))$ can be determined beforehand, and setting a reasonable positive threshold $\eta < 1$, we can introduce the constraint relation

$$E_D(s) \geq \eta,$$

which guarantees a suboptimal yet reasonable solution.
The value of $\det(M(\hat{s}))$ can be determined beforehand, and setting a reasonable positive threshold $\eta < 1$, we can introduce the constraint relation

$$E_D(s) \geq \eta,$$

which guarantees a suboptimal yet reasonable solution. This is equivalent to

$$\Psi[M(s)] \leq C$$

where $C = \Psi[M(\hat{s})] - m \log(\eta)$. 
Ultimate optimal control problem

Find \[
\min_{(s_0,u) \in \mathcal{P}} \text{trace}(M(s)) \text{trace}(M^{-1}(s))
\]
subject to
\[
h(s_0, u) \leq 0 \quad \text{and} \quad J(s_0, u) \leq 0
\]
Ultimate optimal control problem

Find

$$\min_{(s_0,u) \in \mathcal{P}} \text{trace}(M(s)) \text{trace}(M^{-1}(s))$$

subject to

$$h(s_0, u) \leq 0 \quad \text{and} \quad J(s_0, u) \leq 0$$

where

$$\mathcal{P} = \{(s_0, u) : s_0 \in X^N, u : [0, t_f] \rightarrow \mathbb{R}^r \text{ is measurable}, \ u_l \leq u(t) \leq u_u \text{ a.e. on } [0, t_f]\}$$

$$J(s_0, u) = \Psi[M(s)] - C,$$

$$h(s_0, u) = \max_{(\ell,t) \in \tilde{\nu} \times [0, t_f]} \{\gamma(\ell(s(t)))\}.$$
Reduction to canonical form

It is convenient to cast the problem in question as a problem in canonical form which is acceptable by common numerical software for solving optimal-control problems (RIOTS_95, DIRCOL, MISER).

The procedure is rather standard and reduces to augmenting the state vector by some functions of the sensitivity coefficients along the sensor trajectories, see the Proceedings.
Consider the heat equation

\[
\frac{\partial y(x, t)}{\partial t} = \frac{\partial}{\partial x_1} \left( \kappa(x) \frac{\partial y(x, t)}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left( \kappa(x) \frac{\partial y(x, t)}{\partial x_2} \right) + 20 \exp\left(-50(x_1 - t)^2\right), \quad (x, t) \in (0, 1)^3,
\]

\[y(x, 0) = 0, \quad x \in \Omega\]

\[y(x, t) = 0, \quad (x, t) \in \partial \Omega \times T\]

where \(\kappa(x) = \theta_1 + \theta_2 x_1 + \theta_3 x_2\), \(\theta_1 = 0.1\), \(\theta_2 = -0.05\), \(\theta_3 = 0.2\), \(\dot{s}(t) = u(t)\), \(s(t_0) = s_0\),
\[|u_i(t)| \leq 0.7, \quad i = 1, \ldots, 6\]
Results

D-optimum solution

Turing’s measure of cond.

D-efficiency = 0.8
Results

Pure Turing’s measure of conditioning
Conclusions

We have indicated possible ways of making the Hessian of the parameter estimation cost well conditioned while guaranteeing a reasonably small volume of the uncertainty ellipsoid for the estimates.
Conclusions

We have indicated possible ways of making the Hessian of the parameter estimation cost well conditioned while guaranteeing a reasonably small volume of the uncertainty ellipsoid for the estimates.

Common software (MATLAB Partial Differential Equation Toolbox and RIOTS_95) can be used in a relatively easy implementation and fast computations.
For the interested audience

Details beyond the talk are described in the book

For the interested audience

Details beyond the talk are described in the book


Thank you!