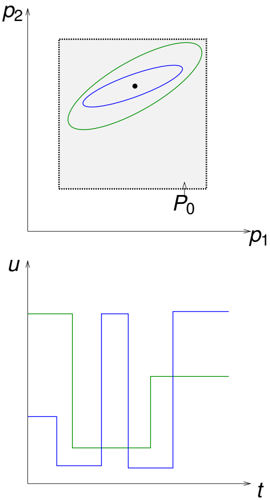
Model-Based Optimal Experiment Design for Nonlinear Parameter Estimation Using Exact Confidence Regions

A.R. Gottu Mukkula and R. Paulen†
Technische Universität Dortmund
†Present address: Slovak University of Technology in Bratislava

July 13, 2017

Motivation

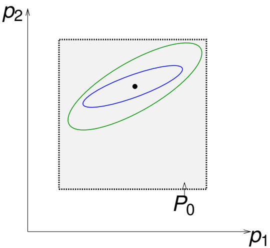
Linearized Confidence Regions

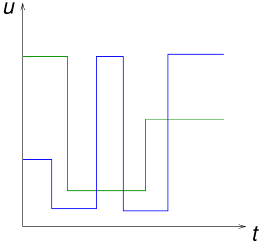


Optimal Experiment Design

Motivation

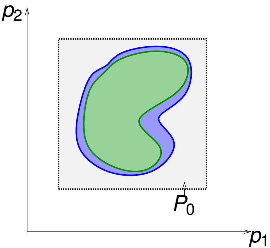
Linearized Confidence Regions

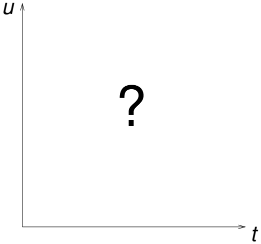




Optimal Experiment Design

Exact Confidence Regions





Optimal Experiment Design

From Data to Model

1. Select the model structure.

Dynamic model or spatially discretized PDEs:

$$\hat{y} = F(u,p) \begin{cases} \text{state equation:} & \dot{x} = f(x,u,p),\ x(0) = g(u,p) \\ \text{output equation:} & \hat{y} = h(x,p) \end{cases}$$

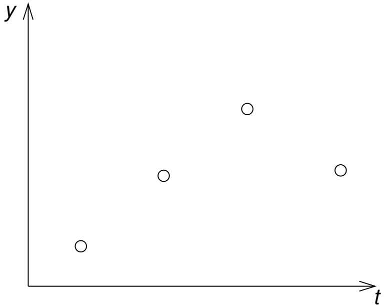
Steady-state or discretized model:

$$\hat{y} = F(u,p) \begin{cases} \text{state equation:} & 0 = f(x,u,p) \\ \text{output equation:} & \hat{y} = h(x,p) \end{cases}$$

$\hat{y}$ - outputs; u - experimental conditions

From Data to Model

2 Perform experiments. Gather data.

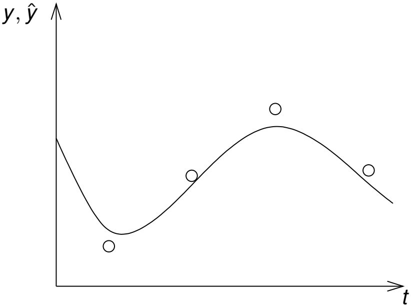


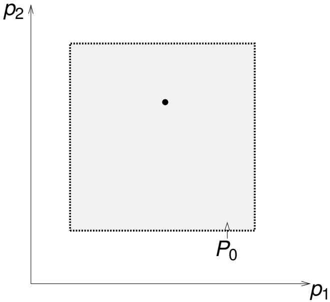
From Data to Model

3 Estimate unknown parameters.

$$\hat{p} = \arg\min_{p \in P_0} J = \arg\min_{p \in P_0} \|y - \hat{y}\|_2^2$$

$$\text{s.t. } \hat{y} = F(u, p)$$



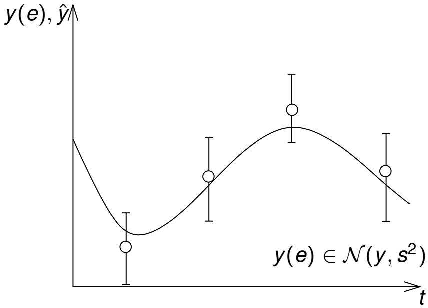


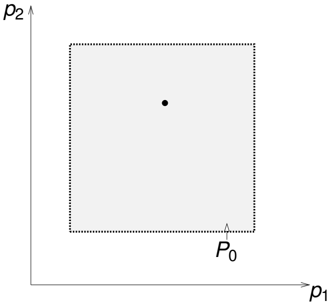
From Data to Model

4 Analyze the quality of estimates.

$$\hat{p} = \arg\min_{p \in P_0} J = \arg\min_{p \in P_0} \|y(e) - \hat{y}\|_2^2$$

$$\text{s.t. } \hat{y} = F(u, p)$$

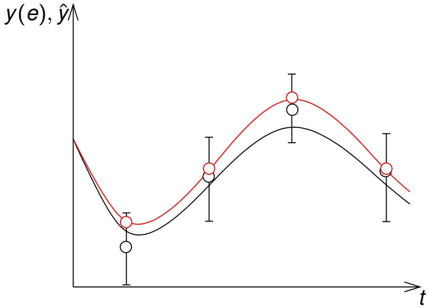


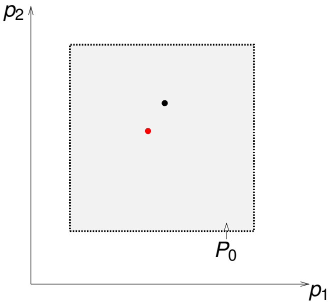


From Data to Model

4 Analyze the quality of estimates.

$$\hat{p} = \arg\min_{p \in P_0} J = \arg\min_{p \in P_0} \|y(e) - \hat{y}\|_2^2$$
$$\text{s.t. } \hat{y} = F(u, p)$$

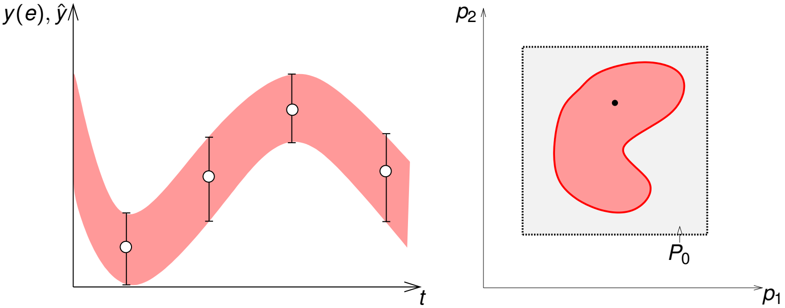




From Data to Model

4 Analyze the quality of estimates.

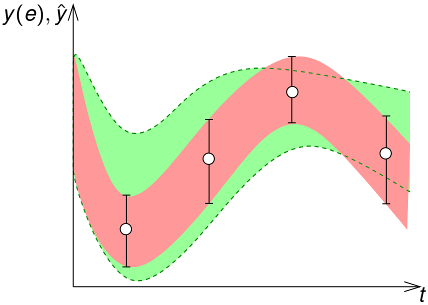
$$J(p) - J(\hat{p}) \leq n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

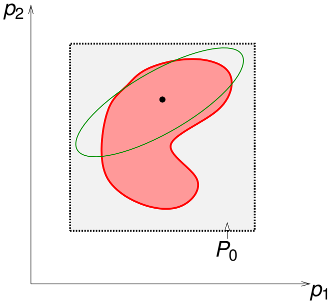


From Data to Model

4 Analyze the quality of estimates.

$$(p - \hat{p})^T \underbrace{\sum_{k=1}^{N} \frac{\partial \hat{y}(k)}{\partial p^T} \frac{\partial \hat{y}(k)}{\partial p}}_{\text{Fisher Inform. Mtx. (FIM)}} (p - \hat{p}) \leq n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

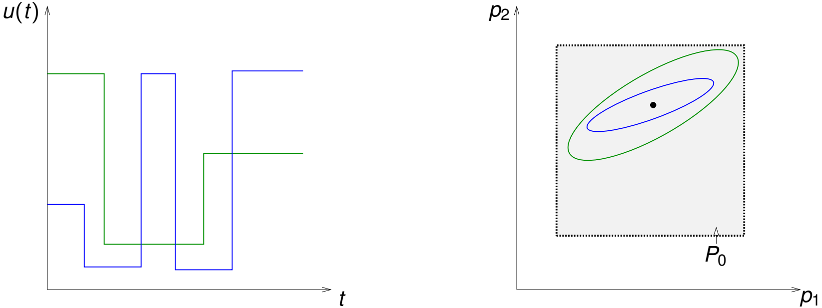




Linearized Optimal Experiment Design (OED)

- Idea: Improve the reliability of parameter estimates by optimizing some measure of Fisher Information Matrix (FIM).

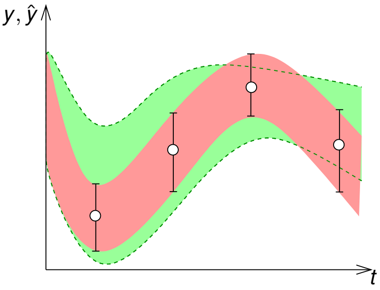
$$\min_{u(t)} \ \phi(\mathrm{FIM}) \quad \text{s.t. } \mathrm{FIM} = \sum_{k=1}^{N} \frac{\partial \hat{y}(k)}{\partial p^T} Q \frac{\partial \hat{y}(k)}{\partial p}, \ \hat{y} = F(p, u)$$

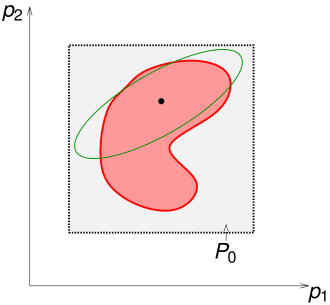


A design: $\phi(\mathrm{FIM}) = \mathrm{tr}(\mathrm{FIM})$, D design: $\phi(\mathrm{FIM}) = \det(\mathrm{FIM})$, ...

Optimal Experiment Design (OED)

- The linearized and exact confidence regions are different in general.





How to find OED using the exact confidence region?

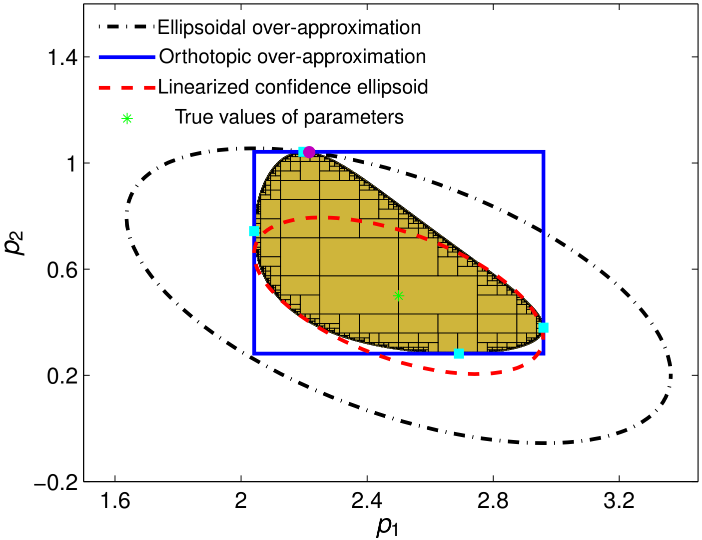
Example

- Benchmark nonlinear problem (biological oxygen demand (BOD))

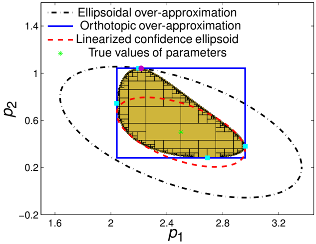
$$\hat{y} = p_1(1 - \exp(-p_2 u)), \quad u \in [0, 20],$$

  with y measured at the incubation times u.
- Random normally distributed measurement noise with standard deviation $\sigma = 0.1$.
- True values of parameters $p^* = (p_1, p_2)^T = (2.5, 0.5)^T$.
- $N = 4$, e.g. $u := (2, 2, 20, 20)^T$ (linearized D design)

Example: 2σ(95%)-confidence regions



Over-approximation of exact confidence regions



Ellipsoidal over-approximation

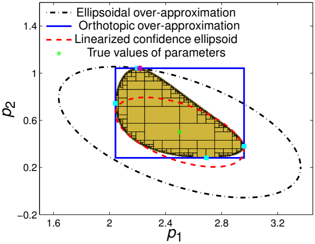
$$\max_{p,k} \quad k$$

$$\text{s.t. } J(p) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

$$(p-\hat{p})^T \text{FIM}(p-\hat{p}) =$$

$$k n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Over-approximation of exact confidence regions



Ellipsoidal over-approximation

$$\max_{p,k} \quad k$$
$$\text{s.t. } J(p) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$
$$(p - \hat{p})^T \mathrm{FIM}(p - \hat{p}) =$$
$$k n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Orthotopic over-approximation

$$\max_{\pi} \sum_{j=1}^{n_p} p_j^U - p_j^L$$
$$\text{s.t. } J(\pi_j) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Optimal Experiment Design

Ellipsoidal over-approximation

$$\min_u \phi(\text{FIM}/k) \;\; \text{s.t.}$$

$$\max_{p,k} \quad k$$

$$\text{s.t.} \; J(p) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

$$(p-\hat{p})^T \text{FIM} (p-\hat{p}) =$$

$$k n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Optimal Experiment Design

Ellipsoidal over-approximation

$$\min_{u} \phi(\mathrm{FIM}/k) \;\; \text{s.t.}$$

$$\max_{p,k} \quad k$$

$$\text{s.t.} \;\; J(p) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

$$(p - \hat{p})^T \mathrm{FIM} (p - \hat{p}) =$$

$$k n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Orthotopic over-approximation

$$\min_{u} \sum_{j=1}^{n_p} p_j^U - p_j^L \;\; \text{s.t.}$$

$$\max_{\pi} \sum_{j=1}^{n_p} p_j^U - p_j^L$$

$$\text{s.t.} \;\; J(\pi_j) - J(\hat{p}) = n_p s^2 \mathcal{F}_{n_p, N-n_p, \alpha}$$

Solving bi-level program

$$
\begin{aligned}
\min_{x}\ & f(y) \\
\text{s.t.}\ & \max_{y} g(y) \\
& \quad \text{s.t. } 0 = h(x, y)
\end{aligned}
$$

- Using KKT-based reformulation of the lower-level problem

$$
\begin{aligned}
\min_{x,y,\lambda}\ & f(y) \\
\text{s.t. } 0 &= \nabla_y g(y) + \nabla_y h(x, y)\lambda \\
0 &= h(x, y)
\end{aligned}
$$

- Using nested approach

$$
\min_{x} f(y^*(x)) \quad \xleftrightarrow{x^*, y^*} \quad
\begin{aligned}
&\max_{y} g(y) \\
&\text{s.t. } 0 = h(x^*, y)
\end{aligned}
$$

Solving bi-level program

$$\min_{x} f(y)$$
$$\text{s.t.} \ \max_{y} g(y)$$
$$\text{s.t.} \ 0 = h(x, y)$$

- Using KKT-based reformulation of the lower-level problem

$$\min_{x,y,\lambda} f(y)$$
$$\text{s.t.} \ 0 = \nabla_y g(y) + \nabla_y h(x, y)\lambda$$
$$0 = h(x, y)$$

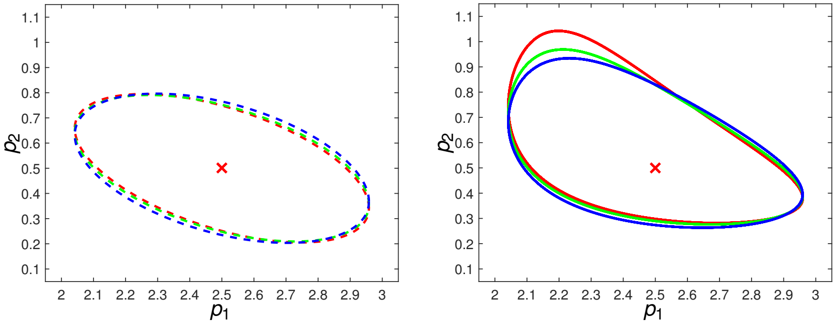
- Using nested approach

$$x_{k+1} = x_k - \left.\frac{\partial f}{\partial y}\right|_{y^*} \left.\frac{\partial y^*}{\partial x}\right|_{x_k} \quad \overset{x^*, y^*}{\longleftrightarrow} \quad \begin{aligned} &\max_{y} g(y) \\ &\text{s.t.} \ 0 = h(x^*, y) \end{aligned}$$

Example: Simulation results

$$\hat{y} = p_1(1 - \exp(-p_2 u)), \quad u \in [0, 20]$$

Linearized $\Leftarrow$ Confidence Regions $\Rightarrow$ Exact



- Linearized design: $u_A = (1.6855, 1.6855, 20, 20)^T$,
  $u_D = (2, 2, 20, 20)^T$
- Orthotope-based design: $u_{A/D} = (1.3685, 1.3685, 20, 20)^T$

Conclusions

- Problem of optimal experiment design formulated as a bilevel program.
- A tight over-approximation of a joint-confidence region realized as an orthotope.
- Computationally intensive problem but tractable for small-scale cases.
- Classical optimal experiment design does not cope well with model nonlinearity $\rightarrow$ Optimal experiment designed using the exact confidence region can differ greatly from the (classical) linearized counterpart.

Acknowledgement: Model-based Optimizing Control- from a vision to industrial reality mobocon