Make Models Great Again

by optimally restricting parameters to make non-identifiable models provably identifiable

Mohamed Tarek

Pumas-AI Inc.

Jun 27, 2024

Introduction

Before vs Now

Over the years in PMx

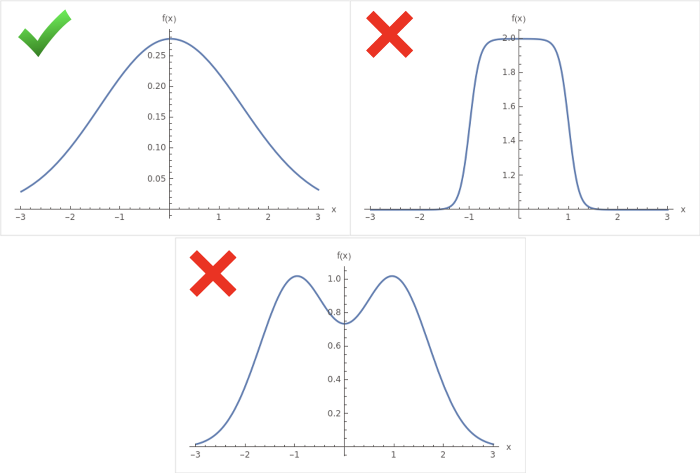
- Model complexity has significantly gone up
  - 10s to 100s of compartments, and
  - 10s to 1000s of parameters.
- Data available for each analysis remains small and often sparse.
- When the data available is small and sparse, increasing the model complexity risks over-fitting.

What is Model Identifiability?

- The ability to identify unique parameter values that fit the data.
  - Global structural identifiability
  - Global practical identifiability
  - Local structural identifiability
  - Local practical identifiability
- Structural vs practical
  - Structural identifiability: assumes infinite data, depends on the model structure only
  - Practical identifiability: assumes a specific experiment design, depends on both the model structure and experiment design

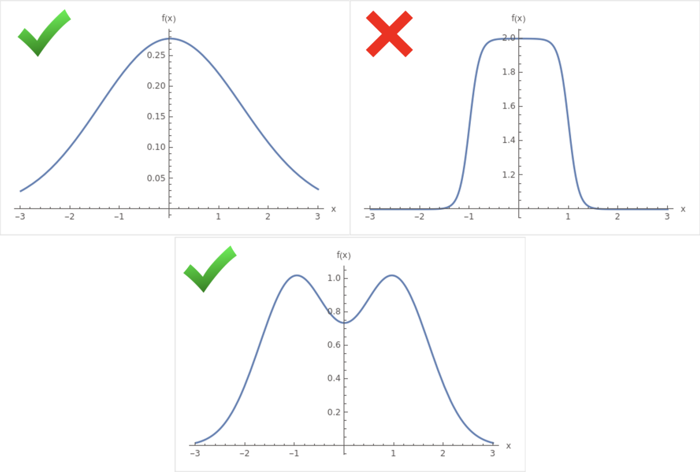
What is Model Identifiability?

Globally identifiable model likelihood curve



What is Model Identifiability?

Locally identifiable model likelihood curve



What is Model Identifiability?

- *Ideally*, we want **global practical identifiability** given the experiment design. But that can be **difficult to prove** for an arbitrary model.
- The second best thing is **local practical identifiability** (LPI) given the experiment design which is luckily **easier to prove** for any given θ using the **Fisher information**.

Fisher information (matrix)

The Fisher information matrix (FIM) is:

$$M(\theta) = E_y\left[\frac{\partial l(\theta; y, d)}{\partial \theta} \cdot \frac{\partial l(\theta; y, d)}{\partial \theta}^T\right]$$

- θ: population parameters
- y: response variable of all subjects
- d: experiment design which includes the number of subjects, subject covariates, sample times, dose levels, etc.
- $l(\theta; y, d)$: log of the marginal likelihood of θ (marginalising any subject-specific random effects)

Fisher information (matrix)

- If l is twice differentiable wrt θ, the FIM is also equivalently given by:

$$M(\theta) = E_y\left[-\frac{\partial^2 l(\theta; y, d)}{\partial\theta \cdot \partial\theta^T} \right]$$

- The positive definiteness (non-singularity) of $M(\theta)$ is a sufficient condition for LPI in the neighbourhood of θ.
- It is **only a sufficient but not a necessary condition** in general.
- Some identifiable models can have a singular FIM.

Fisher information example

Example: normal distribution with mean μ and unit variance:

$$l(\mu; y) = -\frac{1}{2}\log 2\pi - (y-\mu)^2$$

$$\frac{\partial l}{\partial \mu} = 2(y-\mu)$$

$$E_y\left[\frac{\partial l}{\partial \mu}^2\right] = E_y[4 \cdot (y-\mu)^2] = 4$$

Therefore, the model is identifiable!

Fisher information example

Let $\mu = \theta^3$

$$\begin{aligned} l(\theta; y) &= -\frac{1}{2}\log 2\pi - (y - \theta^3)^2 \\ \frac{\partial l}{\partial \theta} &= 6\theta^2 \cdot (y - \theta^3) \\ \frac{\partial^2 l}{\partial \theta^2} &= 12\theta \cdot (y - \theta^3) - 18\theta^4 \\ E_y\left[-\frac{\partial^2 l}{\partial \theta^2}\right] &= E_y[-12\theta \cdot (y - \theta^3) + 18\theta^4] \\ &= 18\theta^4 \end{aligned}$$

If the true θ is 0, the Fisher information is 0. **But the model IS identifiable!**

Methodology

Scope of this work

- Fixing local practical non-identifiability by fixing parameter values.
- Default parameters values are assumed to be given by the user.
- LPI is only guaranteed in the neighbourhood of the given values of θ.

What is the smallest number of parameters that we should fix to make the model locally identifiable at θ?

- When parameters are fixed, the new model's FIM is a sub-matrix of the original model's FIM.
- **Easier question**: what is the largest sub-matrix of the FIM $M(\theta)$ that is not singular?
- The remaining parameters not included in this sub-matrix should be fixed.
- This guarantees identifiability but it may be too conservative because the non-singular FIM is only a sufficient, not a necessary condition.

Mixed integer semi-definite programming (MISDP) formulation

$$\begin{aligned}
&\text{minimize}_x \sum_{i=1}^{n} x_i \\
&\text{s.t.} \\
&\quad C + 100 \cdot \text{diag}(x) - \epsilon \cdot I \succeq 0 \\
&\quad x_i \in \{0, 1\} \qquad \forall i = 1, \ldots, n
\end{aligned}$$

- x_i: a binary decision variable, 0 if parameter θ_i is to be estimated and 1 if the parameter is to be fixed.
- C: scaled version of the FIM $M(\theta)$, the diagonal elements larger than 1e-3 are normalized to 1 by row and column scaling.
- diag(x): diagonal matrix with diagonal elements x.

Mixed integer semi-definite programming (MISDP) formulation

$$
\begin{aligned}
&\text{minimize}_x \sum_{i=1}^{n} x_i \\
&\quad \text{s.t.} \\
&\qquad C + 100 \cdot \text{diag}(x) - \epsilon \cdot I \succeq 0 \\
&\qquad x_i \in \{0, 1\} \qquad\qquad \forall i = 1, \ldots, n
\end{aligned}
$$

- $\succeq 0$: positive semi-definite constraint.
- ϵ: small tolerance to ensure strict positive definiteness.
- Adding 100 to a diagonal element of C is analogous to specifying a **strong prior** for the corresponding parameter practically fixing it in the model.

Mixed integer semi-definite programming (MISDP) formulation

- Once an optimal solution x^* is found, an additional constraint can be added to sequentially discover distinct ways to make the model identifiable.

$$\sum_{i=1}^{n} |x_i - x_i^*| \geq 2$$

- This can be re-written as a linear constraint in x because x and x^* are binary.

- Solutions that are super-sets of other solutions are filtered as a post-processing step.

Examples

Example 1: accidental non-identifiable model

```
model = @model begin
  @param begin
    θ ∈ VectorDomain(3, lower = zeros(3))
    Ω ∈ PSDDomain(3)
    σ ∈ RealDomain(lower = 0.0)
  end
  @random η ~ MvNormal(Ω)
  @pre begin
    Ka = θ[1] * exp(η[1])
    CL = θ[2] * exp(η[2])
    Vc = θ[3] * exp(η[3])
  end
  @dynamics begin
    Depot' = -Ka * Depot
    Central' = Ka * Depot - CL / Vc * Central
  end
  @derived begin
    cp := @. Central
    dv ~ @. Normal(cp, σ)
  end
end
```

Example 1: accidental non-identifiable model

```
subject = Subject(
  events = DosageRegimen(100),
  observations = (dv = nothing,),
  time = 0.0:5.0:100.0,
)

OD.analyze_identifiability(
  model,
  subject,
  (
    θ = [1.5, 1.0, 30.0],
    Ω = I(3),
    σ = 0.1,
  ),
  nsol = 5,
)
```

Example 1: accidental non-identifiable model

```
[ Info: Computing FIM.
[ Info: Running identifiability analysis.
[ Info: Reached max `nsol` = 5.
[ Info: Removing supersets.
[ Info: The following parameters are likely locally non-identifiable:
["θ₃", "Ω₃,₁", "Ω₃,₂", "Ω₃,₃", "θ₂", "Ω₂,₂", "Ω₂,₁"].
┌ Info: Fixing any of the following parameter sets may make the model
locally identifiable. The list may not be exhaustive.
│ Set 1: ["θ₃", "Ω₃,₁", "Ω₃,₂", "Ω₃,₃"]
│ Set 2: ["θ₂", "Ω₃,₁", "Ω₂,₂", "Ω₃,₃"]
│ Set 3: ["θ₃", "Ω₂,₁", "Ω₂,₂", "Ω₃,₃"]
│ Set 4: ["θ₃", "Ω₃,₁", "Ω₂,₂", "Ω₃,₃"]
└ Set 5: ["θ₃", "Ω₃,₁", "Ω₂,₂", "Ω₃,₂"]
```

Example 2: five compartment lung model

```
lung_model = @model begin
  @param begin
    V1 ~ RealDomain(lower = 0.0, init = 10.0)
    V2 ~ RealDomain(lower = 0.0, init = 10.0)
    V4 ~ RealDomain(lower = 0.0, init = 10.0)
    fu4 ~ RealDomain(lower = 0.0, init = 1.1)
    CL ~ RealDomain(lower = 0.0, init = 2.0)
    fu1 ~ RealDomain(lower = 0.0, init = 1.1)
    fu2 ~ RealDomain(lower = 0.0, init = 1.1)
    CL_D12 ~ RealDomain(lower = 0.0, init = 1.1)
    CL_D14 ~ RealDomain(lower = 0.0, init = 1.1)
    CL_D45 ~ RealDomain(lower = 0.0, init = 1.1)
    k32 ~ RealDomain(lower = 0.0, init = 1.1)
    CL_D23 ~ RealDomain(lower = 0.0, init = 1.1)
  end
  @dynamics begin
    A1' = -CL * A1 / V1 - CL_D12 * fu1 * A1 / V1 - CL_D14 * fu1 * A1 / V1 +
            1.0     + CL_D12 * fu2 * A2 / V2 + CL_D14 * fu4 * A4 / V4
    A2' = k32 * A3 + CL_D12 * fu1 * A1 / V1 - (CL_D12 + CL_D23) * fu2 * A2 / V2
    A3' = CL_D23 * fu2 * A2 / V2 - k32 * A3
    A4' = CL_D14 * fu1 * A1 / V1 - CL_D14 * fu4 * A4 / V4 - CL_D45 * fu4 * A4 / V4 + k32 * A5
    A5' = CL_D45 * fu4 * A4 / V4 - k32 * A5
  end
  @derived begin
    y1 ~ @. Normal((A2 + A3) / V2, 0.01)
    y2 ~ @. Normal(A1 / V1, 0.01)
  end
end
```

Example 2: five compartment lung model

```
lung_subject = Subject(
  id = 1,
  time = 0.0:0.1:100,
  observations = (y1 = nothing, y2 = nothing),
)
iparams = init_params(lung_model)
ps = OD.analyze_identifiability(
  lung_model,
  lung_subject,
  iparams,
  nsol = 5,
  psd_tol = 1e-5
)
```

Example 2: five compartment lung model

```
[ Info: Computing FIM.
[ Info: Running identifiability analysis.
new incumbent
[ Info: Reached max `nsol` = 5.
[ Info: Removing supersets.
[ Info: The following parameters are likely locally non-identifiable:
["V4", "fu4", "CL", "fu1", "fu2", "CL_D45", "CL_D23", "V2", "CL_D12",
"CL_D14", "k32"].
┌ Info: Fixing any of the following parameter sets may make the model
locally identifiable. The list may not be exhaustive.
│ Set 1: ["V4", "fu4", "CL", "fu1", "fu2", "CL_D45", "CL_D23"]
│ Set 2: ["V2", "V4", "fu1", "CL_D12", "CL_D14", "k32"]
│ Set 3: ["V4", "fu4", "fu2", "CL_D12", "CL_D14", "CL_D45"]
│ Set 4: ["V2", "fu4", "fu1", "fu2", "CL_D14", "k32"]
└ Set 5: ["V2", "fu4", "CL", "fu1", "CL_D12", "CL_D14", "k32"]
```

Example 2: five compartment lung model

Let's verify the identifiability with simulation!

```
constantcoef = (:V4, :fu4, :CL, :fu1, :fu2, :CL_D45, :CL_D23)
sim_pop = [Subject(simobs(lung_model, lung_subject, iparams)) for _ = 1:10]
fpm = fit(lung_model, sim_pop, iparams, NaivePooled(); constantcoef)
ses = stderror(infer(fpm))
```

Example 2: five compartment lung model

12×4 DataFrame

Row	param Symbol	fitted_params Float64	true_params Float64	ses Float64
1	V1	10.028	10.0	0.0739027
2	V2	9.97279	10.0	0.0582973
3	V4	10.0	10.0	NaN
4	fu4	1.1	1.1	NaN
5	CL	2.0	2.0	NaN
6	fu1	1.1	1.1	NaN
7	fu2	1.1	1.1	NaN
8	CL_D12	1.10005	1.1	0.00782735
9	CL_D14	1.09482	1.1	0.00987285
10	CL_D45	1.1	1.1	NaN
11	k32	1.1105	1.1	0.00706723
12	CL_D23	1.1	1.1	NaN

Conclusion

Conclusion

- Choosing which parameters to fix in a locally non-identifiable model to make it locally identifiable can be posed as a tractable optimization problem.
- By enforcing the FIM sufficient (but not necessary) condition, it can be overly conservative.
- However in practice, the proposed method seems to give reasonable suggestions.
- This method can potentially be used to restrict large QSP models as part of NLME model development.

References

References

- General identifiability analysis and the FIM
  - Thomas J. Rothenberg. Identification in parametric models. Econometrica, 1971.
- FIM in NLME models
  - F. Mentre, A Mallet, and D. Baccar. Optimal design in random-effects regression models. Biometrika, 1997.
  - S. Retout and F Mentre. Further development of the fisher information matrix in nonlinear mixed-effects models with evaluation in population pharmacokinetics. Journal of biopharmaceutical statistics, 2003.

References

- Using the FIM to analyse NLME model identifiability
  - V. Shivva, K. Korell, I. Tucker, and S. Duffull. An approach for identifiability of population pharmacokinetic-pharmacodynamic models. CPT Pharmacometrics & Systems Pharmacology, 2013.
  - Stephen Dufful, A workflow for resolving model internal consistency in use-reuse settings (aka repairing unstable models). PAGANZ, 2024.
  - Dan Wright. The identifiability of a turnover model for allopurinol urate-lowering effect. PAGANZ, 2024.

References

- Five compartment lung model

  - R. Hendrickx, E. L. Bergstrom, D. L. I. Janzen, M. Friden, U. Eriksson, K. Grime, and D. Ferguson. Translational semi-physiological model to predict human plasma and lung pharmacokinetics and pulmonary efficacy after oral drug inhalation of bronchodilators. CPT Pharmacometrics & Systems Pharmacology, 2018.
  - David Lars Ivan Janzen. Structural Identifiability and Indistinguishability in Mixed-Effects Models. PhD Thesis, 2016.

Thank you!