The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

The Geometry of Sloppiness

EMILIE DUFRESNE
School of Mathematics and Statistics
University of Nottingham

(with Heather Harrington, and Dhruva Raman)

Discrete Methods in Molecular Biology,
SIAM AN17, Pittsburgh, PA, USA
12 July 2017

Sloppiness

Understanding the uncertainty in parameter estimation

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

OPEN ACCESS Freely available online PLoS COMPUTATIONAL BIOLOGY

Universally Sloppy Parameter Sensitivities in Systems Biology Models

Ryan N. Gutenkunst[1*], Joshua J. Waterfall[2], Fergal P. Casey[3], Kevin S. Brown[4], Christopher R. Myers[5], James P. Sethna[1]

1 Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York, United States of America, 2 Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York, United States of America, 3 UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Ireland, 4 Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, United States of America, 5 Cornell Theory Center, Cornell University, Ithaca, New York, United States of America

Quantitative computational models play an increasingly important role in modern biology. Such models typically involve many free parameters, and assigning their values is often a substantial obstacle to model development. Directly measuring in vivo biochemical parameters is difficult, and collectively fitting them to other experimental data often yields large parameter uncertainties. Nevertheless, in earlier work we showed in a growth-factor-signaling model that collective fitting could yield well-constrained predictions, even when it left individual parameters very poorly constrained. We also showed that the model had a "sloppy" spectrum of parameter sensitivities, with eigenvalues roughly evenly distributed over many decades. Here we use a collection of models from the literature to test whether such sloppy spectra are common in systems biology. Strikingly, we find that every model we examine has a sloppy spectrum of sensitivities. We also test several consequences of this sloppiness for building predictive models. In particular, sloppiness suggests that collective fits to even large amounts of ideal time-series data will often leave many parameters poorly constrained. Tests over our model collection are consistent with this suggestion. This difficulty with collective fits may seem to argue for direct parameter measurements, but sloppiness also implies that such measurements must be formidably precise and complete to usefully constrain many model predictions. We confirm this implication in our growth-factor-signaling model. Our results suggest that sloppy sensitivity spectra are universal in systems biology models. The prevalence of sloppiness highlights the power of collective fits and suggests that modelers should focus on predictions rather than on parameters.

Citation: Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, et al. (2007) Universally sloppy parameter sensitivities in systems biology models. PLoS Comput Biol 3(10): e189. doi:10.1371/journal.pcbi.0030189

An equivalence relation on parameter space

Mathematical models with perfect data

- M is a mathematical model describing the behaviour of a variable $x \in X \subseteq \mathbb{R}^m$ depending on a parameter $p \in P \subseteq \mathbb{R}^r$ and with measurable output $y \in Y \subseteq \mathbb{R}^n$.
- Specify a choice of "perfect data" z.

An equivalence relation on parameter space

Mathematical models with perfect data

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

- (M, z) is a mathematical model with a choice of "perfect data" z.
- This choice of perfect data induces an equivalence relation $\sim_{M,z}$ on the parameter space P.

- (M, z) is a mathematical model with a choice of "perfect data" z.
- This choice of perfect data induces an equivalence relation $\sim_{M,z}$ on the parameter space P.

Example (Real-analytic ODE systems)

$\dot{x} = f(p, x)$, $y = g(x)$, where f and g are real-analytic functions and X and P are real-analytic manifolds.

and for $p \in P$ perfect data is chosen to be

- Time series data: $(y(t_1), \ldots, y(t_N))$ for $0 \leqslant t_1 < \cdots < t_N$.
- Steady state data:
  $\{y \in Y \mid f(p, x) = 0,\ y = g(x)(,\ x(0) \in X_0)\}$.

M describes the behaviour of the variable $x \in \mathbb{R}_{\geqslant 0}$ depending on $(a, b) \in \mathbb{R}^2_{\geqslant 0}$ via

$$x(a, b, t) = e^{-at} + e^{-bt}.$$

Case Study: the sum of exponentials

The model equivalence relation and its equivalence classes

M describes the behaviour of the variable $x \in \mathbb{R}_{\geqslant 0}$ depending on $(a, b) \in \mathbb{R}^2_{\geqslant 0}$ via

$$x(a, b, t) = e^{-at} + e^{-bt}.$$

Perfect data produced by the parameter (a, b):

- The perfect measurements $(e^{-at_1} + e^{-bt_1}, \ldots, e^{-at_N} + e^{-bt_N})$ for some N and timepoints $t_1 < \cdots < t_N$.

Case Study: the sum of exponentials

The model equivalence relation and its equivalence classes

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

M describes the behaviour of the variable $x \in \mathbb{R}_{\geqslant 0}$ depending on $(a, b) \in \mathbb{R}^2_{\geqslant 0}$ via

$$x(a, b, t) = e^{-at} + e^{-bt}.$$

Perfect data produced by the parameter (a, b):

- The perfect measurements $(e^{-at_1} + e^{-bt_1}, \ldots, e^{-at_N} + e^{-bt_N})$ for some N and timepoints $t_1 < \cdots < t_N$.

Taking measurements at two distinct nonzero timepoints induces the same equivalence relation on P as knowing the value of $e^{-at} + e^{-bt}$ for all t (the continuous data), that is for $0 < t_1 < \cdots < t_N$, we have

$$\sim_{M,t_1,\ldots,t_N} = \sim_{M,\infty}.$$

Case Study: the sum of exponentials

The model equivalence relation and its equivalence classes

M describes the behaviour of the variable $x \in \mathbb{R}_{\geqslant 0}$ depending on $(a, b) \in \mathbb{R}^2_{\geqslant 0}$ via

$$x(a, b, t) = e^{-at} + e^{-bt}.$$

- It is easy to see that $(a, b) \sim_{M,\infty} (b, a)$.
- One can show that the equivalence class of (a, b) is

$$\begin{array}{ll} \{(a, b), (b, a)\} & \text{if } a \neq b \\ \{(a, a)\} & \text{if } a = b. \end{array}$$

Model prediction maps

Geometric realizations of the quotient $P/{\sim_{M,z}}$.

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Definition

A model prediction map is a function $\phi \colon P \to \mathbb{R}^N$ giving the perfect data (the model predictions) as a function of the parameter which factors through the set-theoretic quotient $P/{\sim_{M,z}}$ and is injective on the equivalence classes.

- The existence of a model prediction map requires that the perfect data z produced for the parameter p can be identified with a point of $\mathbb{R}^N$ for some N.

Model prediction maps

Geometric realizations of the quotient $P/{\sim_{M,z}}$.

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Definition

A model prediction map is a function $\phi\colon P \to \mathbb{R}^N$ giving the perfect data (the model predictions) as a function of the parameter which factors through the set-theoretic quotient $P/{\sim_{M,z}}$ and is injective on the equivalence classes.

- The existence of a model prediction map requires that the perfect data z produced for the parameter p can be identified with a point of $\mathbb{R}^N$ for some N.

Remark

The sloppiness literature calls the closure of the image of a model prediction map a "model manifold", although it is not a manifold in general.

Noisy data and sloppiness

The case of additive Gaussian measurement noise

- $z \sim \mathcal{N}(\phi(p), \Sigma)$, with probability density function $\psi(p, z)$,

Noisy data and sloppiness

The case of additive Gaussian measurement noise

- $z \sim \mathcal{N}(\phi(p), \Sigma)$, with probability density function $\psi(p, z)$,

This induces a premetric on P via the Kullback-Liebler divergence:

$$\begin{aligned} d(p', p) &= \int_Z \psi(p, z) \log\left(\frac{\psi(p, z)}{\psi(p', z)}\right) dz \\ &= \left(\frac{1}{2}\big(\phi(p') - \phi(p)\big)^T \Sigma^{-1} \big(\phi(p') - \phi(p)\big)\right). \end{aligned}$$

Noisy data and sloppiness

The case of additive Gaussian measurement noise

- $z \sim \mathcal{N}(\phi(p), \Sigma)$, with probability density function $\psi(p, z)$,

This induces a premetric on P via the Kullback-Liebler divergence:

$$d(p', p) = \left(\frac{1}{2} \big(\phi(p') - \phi(p)\big)^T \Sigma^{-1} \big(\phi(p') - \phi(p)\big) \right).$$

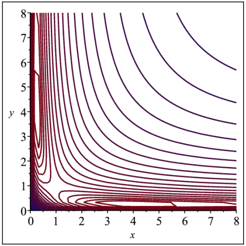
Definition (A qualitative definition of sloppiness)

We say that a mathematical model (M, ϕ, ψ, d_P) is sloppy at p_0 if in a neighborhood of p_0 the premetric d diverges significantly from the reference metric d_P on parameter space.

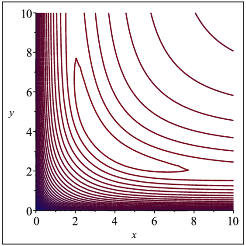
Sloppiness is a "local" property

Sloppiness is not uniform in the parameter space.

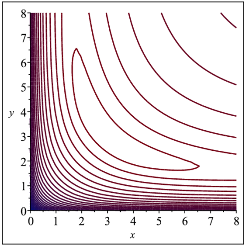
Case Study: The sum of exponentials



$(a, b) = (4, 1/8)$.



$(a, b) = (4, 3)$.



$(a, b) = (3, 3)$.

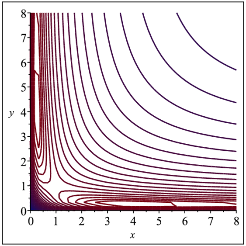
Level curves of $\sqrt{d(\cdot, (a, b))}$ for the model prediction map given by taking $t_1, t_2, t_3 = \frac{1}{3}, 1, 3$ assuming additive Gaussian noise with identity covariance.

Sloppiness is a "local" property

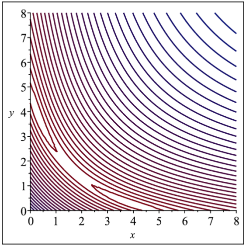
Sloppiness depends on the choice of timepoints.

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

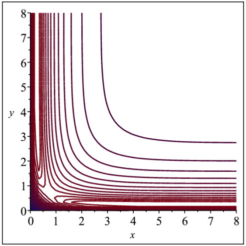
Case Study: The sum of exponentials



$t_1, t_2, t_3 = \frac{1}{3}, 1, 3$



$t_1, t_2 = \frac{1}{9}, \frac{1}{3}$



$t_1, t_2 = 1, 3$

Level curves of $\sqrt{d(\cdot, (4, 1/8))}$ for the model prediction map given by taking the stated timepoints assuming additive Gaussian noise with identity covariance.

Quantifying Sloppiness

Sloppiness in the literature: the Fisher Information Matrix

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Assume that $d_P = d_2$ is the standard Euclidean metric. Suppose that $d(\cdot, p_0)\colon P \to \mathbb{R}_{\geqslant 0}$ is twice continuously differentiable in a neighbourhood of p_0.

$$d(p, p_0) = \frac{1}{2}\langle (p - p_0), \nabla_p^2 d(p, p_0)(p - p_0)\rangle + \mathcal{O}(\|(p - p_0)\|_2),$$

Assume that $d_P = d_2$ is the standard Euclidean metric. Suppose that $d(\cdot, p_0)\colon P \to \mathbb{R}_{\geqslant 0}$ is twice continuously differentiable in a neighbourhood of p_0.

$$d(p, p_0) = \frac{1}{2}\langle (p - p_0), \nabla_p^2 d(p, p_0)(p - p_0)\rangle + \mathcal{O}(\|(p - p_0)\|_2),$$

- $\nabla_p^2 d(p, p_0)$ is known as the Fisher Information Matrix (FIM).

Assume that $d_P = d_2$ is the standard Euclidean metric. Suppose that $d(\cdot, p_0) \colon P \to \mathbb{R}_{\geqslant 0}$ is twice continuously differentiable in a neighbourhood of p_0.

$$d(p, p_0) = \frac{1}{2}\langle (p - p_0), \nabla_p^2 d(p, p_0)(p - p_0)\rangle + \mathcal{O}(\|(p - p_0)\|_2),$$

- $\nabla_p^2 d(p, p_0)$ is known as the Fisher Information Matrix (FIM).
- The FIM induces a pseudometric on parameter space

$$\begin{aligned} d_{\mathrm{FIM,p_0}} \colon P \times P &\to \mathbb{R}_{\geqslant 0} \\ (p, p') &\mapsto \frac{1}{2}\Big\langle (p - p'), \big(\nabla_p^2 d(p, p_0)\big)\,|_{p=p_0}(p - p')\Big\rangle. \end{aligned}$$

which is a linear approximation of d near p_0.

Assume that $d_P = d_2$ is the standard Euclidean metric.

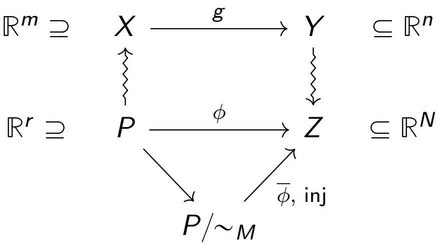
$$d(p, p_0) = \frac{1}{2}\langle (p - p_0), \nabla_p^2 d(p, p_0)(p - p_0)\rangle + \mathcal{O}(\|(p - p_0)\|_2),$$

- $\nabla_p^2 d(p, p_0)$ is known as the Fisher Information Matrix (FIM).

Definition

We say that a mathematical model (M, ϕ, ψ, d_2) is infinitesimally sloppy at a parameter p_0 if there are several orders of magnitude between the largest and smallest eigenvalues of the FIM at p_0.

We define the infinitesimal sloppiness at p_0 to be the condition number of the FIM at p_0, that is, the ratio between its largest and smallest eigenvalues.



Thank you for your attention!