Cramer-Rao bound and absolute sensitivity in chemical reaction networks

[DL, Yuki Sughiyama, Tetsuya J. Kobayashi, *arXiv:2401.06987*, **2024**.]

Dimitri Loutchko

April 2nd, 2024

FRK2024 Seminar

Overview

Introduction & motivation

Self-introduction

- currently: Postdoc in in the Quantitative biology lab, UTokyo
- 2019: PhD in the Fritz-Haber Institute of the Max-Planck Society under Gerhard Ertl
- studied chemistry (Humboldt-University, Berlin) and mathematics (Free University, Berlin)

Gerhard Ertl

- During PhD, I worked on protein dynamics and stochastic thermodynamics
- Now I work mainly on CRN theory, trying to combine thermodynamics with geometry
- Also on CRS theory (Kauffman's autocatalytic sets)
- Interested in the geometry of thermodynamics more generally

Introduction & motivation

CRN theory in the Quantitative biology lab, UTokyo, http://research.crmind.net/

Tetsuya J. Kobayashi

Yuki Sughiyama (now in Tohoku University)

Atsushi Kamimura

- Research on CRN as part of a CREST project on information physics since 2020
- Motivated by [Craciun, G., Dickenstein, A., Shiu, A., & Sturmfels, B. **(2009)**. Toric dynamical systems. Journal of Symbolic Computation]
- Complex-balanced and equilibrium CRN are described by toric varieties = exponential families in statistics

Introduction & motivation

Today's talk

Based on the preprint [DL, Yuki Sughiyama, Tetsuya J. Kobayashi, *arXiv:2401.06987*, **2024**. (currently under review in SIAM: Journal on Applied Mathematics)]

- Another example of the idea that there is a correspondence

$$\text{vectors of chemical concentrations} \longleftrightarrow \text{probability distributions on finite spaces}$$

- More concretely: Vectors of linear conserved quantities η in CRN depend on the choice of a basis for $\mathrm{Ker}[S^T]$,
- The sensitivity of concentration vectors x in the steady-state manifold $\mathcal{V}^{ss}$ is (infinitesimally) captured by the sensitivity matrix $\chi = \frac{\partial x}{\partial \eta}$
- Goal: Define and study quantities that fulfill the same role as χ but are basis independent, hence *absolute*.

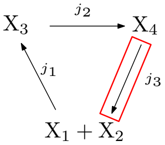
Preliminaries: Deterministic CRN (chemical reaction networks).

Basic notation

- n chemicals $X_1, X_2, \ldots, X_n$ with concentration vector $(x_1, \ldots, x_n) \in X := \mathbb{R}^n_{>0}$.
- m reactions $R_1, R_2, \ldots, R_m$

$$R_j : \sum_{i=1}^{n} S^-_{ij} X_i \to \sum_{i=1}^{n} S^+_{ij} X_i,$$

  with fluxes $j_r \in \mathbb{R}_{>0}$, and the flux vector
- Flux vector $j = (j_1, \ldots, j_m) \in \mathbb{R}^m_{>0}$
- Stoichiometric $n \times m$-matrix $S = S^+ - S^-$.
- Deterministic dynamics $\frac{\mathrm{d}x}{\mathrm{d}t} = Sj$



$$S = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Preliminaries: Vectors of conserved quantities

- Any vector $u \in \mathrm{Ker}[S^T]$ yields the conserved quantity $\eta := \langle u, x \rangle$ as

$$\frac{\mathrm{d}\langle u, x\rangle}{\mathrm{d}t} = \langle u, Sj \rangle = \langle S^T u, j \rangle = 0.$$

- Let q denote the dimension of $\mathrm{Ker}[S^T]$, choose a basis $\{u_i\}_{i=1}^q$ of $\mathrm{Ker}[S^T]$, and write $U = (u_1, \dots, u_q)$ for the respective $n \times q$ matrix of basis vectors.
- This gives the map

$$U^T : X \to \mathbb{R}^q.$$

- For any initial condition $x_0 \in X$ with $\eta := U^T x_0$, the reaction dynamics is confined to the *stoichiometric polytope* (or *stoichiometric compatibility class*)

$$P(\eta) := \{x \in \mathbb{R}^n_{\geq 0} | U^T x = \eta\}.$$

- The range of meaningful parameters η is given by

$$H := U^T X \subset \mathbb{R}^q.$$

Preliminaries: Setup for this talk

- Interested in the steady state manifold

$$\mathcal{V}^{ss} := \{x \in X \text{ such that } Sj = 0\}$$

- Assume that, locally at $x \in \mathcal{V}^{ss}$, the map $U^T : \mathcal{V}^{ss} \to \tilde{H} \subset H$ has a differentiable inverse

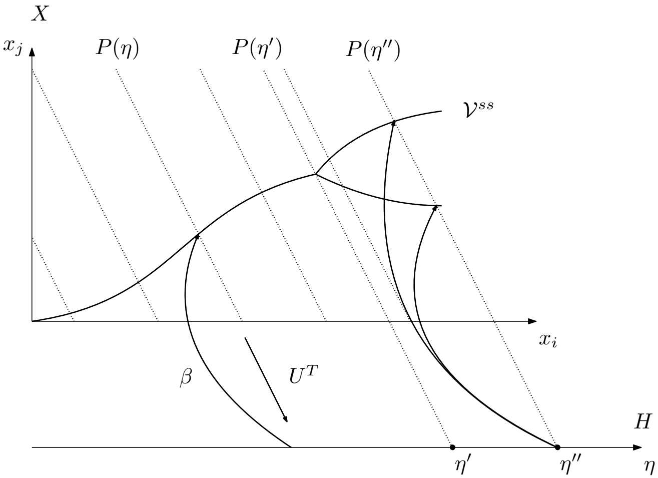
$$\beta : \tilde{H} \to X.$$

- Definition: the sensitivity matrix χ is the Jacobian matrix

$$\chi := D_\eta \beta = \frac{\partial x}{\partial \eta}.$$

- Remark: For complex-balanced CRN, the existence of this section is ensured by Birch's theorem [Craciun, Gheorghe, et al., Journal of Symbolic Computation **(2009)**]

Preliminaries: Setup



Absolute sensitivity: Motivation

From now on, fix $x \in \mathcal{V}^{ss}$ and $\eta \in H$ such that $x = \beta(\eta)$.

Main idea

- Sensitivity matrix: Perturb η, then the linear response is given by χ.
- Matrix of absolute sensitivities: Perturb x, let it relax to the new steady state, then the linear response is given by A.
- More explicitly, perturb X_i by δx_i. Then the concentration change of X_j is $\alpha_{i \to j}\delta x_i$, to first order in δx_i.

Formalize

- Perturb the concentration of X_i by δx_i, then $x \mapsto x + \Delta x$ with $\Delta x = (0, \dots, 0, \delta x_i, 0, \dots, 0)$.
- The vector η changes by $\Delta\eta = U^T \Delta x$
- The adjusted steady state is $\beta(\eta + \Delta\eta)$
- Linearize:

$$\beta(\eta + \Delta\eta) = \beta(\eta) + D_\eta\beta(\Delta\eta) + \mathcal{O}(\|\Delta\eta\|^2) = x + D_\eta\beta(U^T\Delta x) + \mathcal{O}(\|\Delta x\|^2).$$

Absolute sensitivity: Motivation

Formalize

- Perturb the concentration of X_i by δx_i, then $x \mapsto x + \Delta x$ with $\Delta x = (0, \ldots, 0, \delta x_i, 0, \ldots, 0)$.
- The vector η changes by $\Delta\eta = U^T \Delta x$
- The adjusted steady state is $\beta(\eta + \Delta\eta)$
- Linearize:

$$\beta(\eta + \Delta\eta) = \beta(\eta) + D_\eta\beta(\Delta\eta) + \mathcal{O}(\|\Delta\eta\|^2) = x + D_\eta\beta(U^T\Delta x) + \mathcal{O}(\|\Delta x\|^2).$$

- The linear change of the concentration of X_j is

$$[D_\eta\beta(U^T\Delta x)]_j = \left[\frac{\partial x}{\partial\eta} U^T \Delta x\right]_j = \sum_{k=1}^{q} \frac{\partial x_j}{\partial\eta_k} u_{ik}\delta x_i,$$

Absolute sensitivity: Definition

Definition

The *absolute sensitivity* $\alpha_{i\to j}$ of X_j with respect to X_i at a point $x \in \mathcal{V}^{ss}$ is defined as

$$\alpha_{i\to j} := \sum_{k=1}^{q} \frac{\partial x_j}{\partial \eta_k} u_{ik}$$

and the *absolute sensitivity* of the chemical X_i is $\alpha_i := \alpha_{i\to i}$. The $n \times n$ matrix A of absolute sensitivities is given by

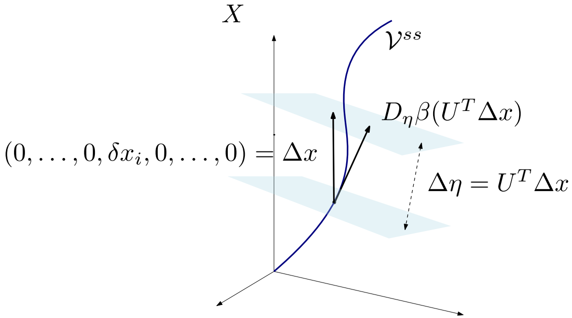
$$A_{ij} = \alpha_{j\to i}$$

and the vector α of absolute sensitivities is given by the diagonal elements of A, i.e., $\alpha = (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{R}^n$.

Remark

The matrix of absolute sensitivities is given by $A = \chi U^T$.

Absolute sensitivity: Geometry



$$\beta(\eta + \Delta\eta) = \beta(\eta) + D_\eta \beta(\Delta\eta) + \mathcal{O}(\|\Delta\eta\|^2) = x + D_\eta \beta(U^T \Delta x) + \mathcal{O}(\|\Delta x\|^2).$$

Absolute sensitivity: Basic properties

Theorem

The matrix of absolute sensitivities A is independent of the choice of a basis of $\mathrm{Ker}[S^T]$. Moreover, the equality

$$\mathrm{Tr}[A] = \sum_{i=1}^{n} \alpha_i = q$$

holds, whereby $q = \dim \mathrm{Ker}[S^T]$.

Proof

Follows directly from the definition: U' denote another matrix of basis vectors, i.e., $U' = UB$ for some $B \in \mathrm{GL}(q)$. Then $\eta' = (U')^T x$ satisfies $\eta' = (U')^T x = B^T \eta$, where $\eta = U^T x$ and

$$A = \frac{\partial x}{\partial \eta} U^T = \frac{\partial x}{\partial \eta'} \frac{\partial \eta'}{\partial \eta} U^T = \frac{\partial x}{\partial \eta'} B^T U^T = \frac{\partial x}{\partial \eta'} (U')^T .$$

Absolute sensitivity: Basic properties

Theorem

The matrix of absolute sensitivities A is independent of the choice of a basis of $\mathrm{Ker}[S^T]$. Moreover, the equality

$$\mathrm{Tr}[A] = \sum_{i=1}^{n} \alpha_i = q$$

holds, whereby $q = \dim \mathrm{Ker}[S^T]$.

Proof

The second claim is verified by differentiating $\eta_j = \sum_{i=1}^{n} u_{ij} x_i$, with respect to η_j and summung over all j:

$$q = \sum_{i=1}^{n} \sum_{j=1}^{q} \frac{\partial x_i}{\partial \eta_j} u_{ij} = \sum_{i=1}^{n} \alpha_i .$$

Definition of quasi-thermostatic CRN

Definition

A CRN is quasi-thermostatic if

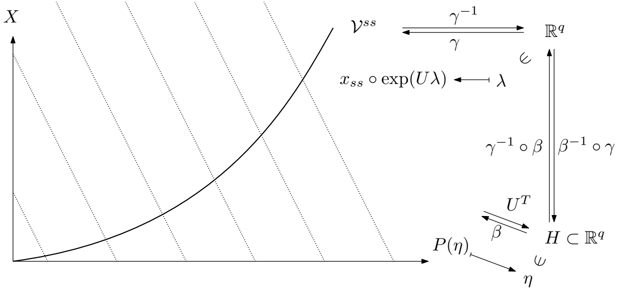
$$\mathcal{V}^{ss} = \{x \in X | \log x - \log x^{ss} \in \mathrm{Ker}[S^T]\}$$

holds for some base point x^{ss}, following [Horn, F. **(1972)**. Necessary and sufficient conditions for complex balancing in chemical kinetics. Archive for Rational Mechanics and Analysis].

Equivalently, the steady state manifold of a quasi-thermostatic CRN can be parametrized by $\mathbb{R}^q$ as

$$\begin{aligned} \gamma : \mathbb{R}^q &\to X \\ \lambda &\mapsto x^{ss} \circ \exp(U\lambda). \end{aligned}$$

Two parametrizations



Birch's theorem: There is a parametrization of $\mathcal{V}^{ss}$ by the space H of conserved quantities given by

$$\beta : H \to X$$
$$\eta \mapsto \mathcal{V}^{ss} \cap P(\eta),$$

cf. [Horn 1972, and Craciun et al. 2009].

Absolute sensitivity

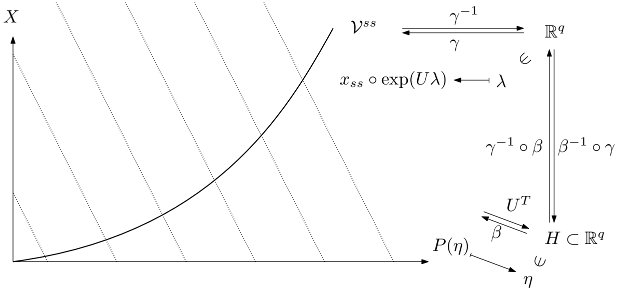


Diagram chasing yields

$$D_\eta \beta = D_\lambda \gamma \cdot D_\eta(\gamma^{-1} \circ \beta) = D_\lambda \gamma \cdot [D_\lambda(\beta^{-1} \circ \gamma)]^{-1}$$
$$= D_\lambda \gamma \cdot [D_x \beta^{-1} \cdot D_\lambda \gamma]^{-1},$$

where $\beta^{-1} = U^T$ and $D_\lambda \gamma = \mathrm{diag}(x) U$ explicitly ($\mathrm{diag}(x)$ is the $n \times n$ diagonal matrix with $\mathrm{diag}(x)_{ii} = x_i$.).

Absolute sensitivity

Diagram chasing yields

$$D_\eta \beta = D_\lambda \gamma \cdot D_\eta(\gamma^{-1} \circ \beta) = D_\lambda \gamma \cdot [D_\lambda(\beta^{-1} \circ \gamma)]^{-1}$$
$$= D_\lambda \gamma \cdot [D_x \beta^{-1} \cdot D_\lambda \gamma]^{-1},$$

where $\beta^{-1} = U^T$ and $D_\lambda \gamma = \mathrm{diag}(x)U$ explicitly ($\mathrm{diag}(x)$ is the $n \times n$ diagonal matrix with $\mathrm{diag}(x)_{ii} = x_i$.).

This yields the explicit form for the sensitivity matrix $\chi = D_\eta \beta = \mathrm{diag}(x)U \cdot [U^T \mathrm{diag}(x)U]^{-1}$ and for the matrix of absolute sensitivities

$$A = \chi \beta U^T = \mathrm{diag}(x)U \cdot [U^T \mathrm{diag}(x)U]^{-1} U^T.$$

Cramer-Rao bound for CRN

As before, fix a point $x \in \mathcal{V}^{ss}$ with coordinates $\lambda \in \mathbb{R}^q$ and $\eta \in H$, respectively.

- Denote the Jacobian of the coordinate change from η to λ by

$$g_\eta := D_\eta(\gamma^{-1} \circ \beta) = [U^T \mathrm{diag}(x) U]^{-1}.$$

- Define the $n \times n$ diagonal matrix $\mathrm{diag}\left(\frac{1}{x}\right)$ by $\mathrm{diag}\left(\frac{1}{x}\right)_{ii} = \frac{1}{x_i}$ and the $\mathrm{diag}\left(\frac{1}{x}\right)$-weighted inner product on $\mathbb{R}^n$ by

$$\langle v, w \rangle_{\frac{1}{x}} := \sum_{i=1}^{n} \frac{1}{x_i} v_i w_i.$$

- Let V be an arbitrary $n \times n$ matrix and $\overline{V}$ a $n \times n$ matrix whose column span satisfies

$$\mathrm{Span}\left[\overline{V}\right] \subset \mathrm{Im}[S].$$

Cramer-Rao bound for CRN

- The covariance matrix of I_n is defined as

$$\mathrm{Cov}(I_n) := (I_n - \overline{V})^T \mathrm{diag}\left(\frac{1}{x}\right)(I_n - \overline{V}).$$

- It is called a covariance matrix because its elements are of the form

$$\mathrm{Cov}(I_n)_{ij} := \langle e_i - \overline{V}_i, e_j - \overline{V}_j \rangle_{\frac{1}{x}}.$$

Theorem: Cramer-Rao bound for CRN

For a quasi-thermostatic CRN, let the covariance matrix $\mathrm{Cov}(V)$ be defined as above. It is bounded from below by

$$\mathrm{Cov}(I_n) \geq U g_\eta U^T,$$

where the matrix inequality is understood in the sense that the difference matrix between the left hand side and the right hand side of the inequality is positive semidefinite.

Cramer-Rao bound and absolute sensitivity

Connection between CRN and absolute sensitivity

Recall that $A = \chi\beta U^T = \mathrm{diag}(x)U \cdot [U^T \mathrm{diag}(x)U]^{-1}U^T$ and that $g_\eta = [U^T \mathrm{diag}(x)U]^{-1}$, the CRB

$$\mathrm{Cov}(I_n) \geq U g_\eta U^T,$$

yields

$$\mathrm{Cov}(I_n) \geq \mathrm{diag}\left(\frac{1}{x}\right) A$$

For $\overline{V} = 0$, the diagonal elements yield $1 \geq \alpha_i$ which is not tight as can be seen by summing over all i and comparing with the general Theorem on absolute sensitivity, giving $n \geq q$.

Tightening the bound

Goal: Minimize LHS of the CRB

$$\mathrm{Cov}(I_n) \geq \mathrm{diag}\left(\frac{1}{x}\right) A.$$

The diagonal entries of $\mathrm{Cov}(I_n)$ are given by the squared norm $\|e_i - \overline{V}_i\|^2_{\frac{1}{x}} = \langle e_i - \overline{V}_i, e_i - \overline{V}_i \rangle_{\frac{1}{x}}$, which is minimized if and only if $\overline{V}_i$ is the $\langle ., . \rangle_{\frac{1}{x}}$-orthogonal projection of e_i to $\mathrm{Im}[S]$.
Denote this projection as

$$\pi : \mathbb{R}^n \to \mathrm{diag}(x)\mathrm{Ker}[S^T].$$

This is enough to achieve equality.

Tightening the bound

The diagonal entries of $\mathrm{Cov}(I_n)$ are given by the squared norm $\|e_i - \overline{V}_i\|^2_{\frac{1}{x}} = \langle e_i - \overline{V}_i, e_i - \overline{V}_i \rangle_{\frac{1}{x}}$, which is minimized if and only if $\overline{V}_i$ is the $\langle ., . \rangle_{\frac{1}{x}}$-orthogonal projection of e_i to $\mathrm{Im}[S]$.
Denote this projection as

$$\pi : \mathbb{R}^n \to \mathrm{diag}(x)\mathrm{Ker}[S^T].$$

Lemma

For quasi-thermostatic CRN, the absolute sensitivity α_i at a point $x = (x_1, \ldots, x_n)$ is given by

$$\alpha_i = x_i \|\pi(e_i)\|^2_{\frac{1}{x}},$$

where e_i is the ith canonical unit vector.

Linear algebraic characterization of absolute sensitivity

Theorem

For quasi-thermostatic CRN, the matrix of absolute sensitivities A at a point $x \in \mathcal{V}^{ss}$ is given by

$$A = \mathrm{diag}(x)\mathrm{Cov}(I_n)$$

with $\mathrm{Cov}(I_n)_{ij} = \langle \pi(e_i), \pi(e_j) \rangle_{\frac{1}{x}}$. Thus, the absolute sensitivities are given by

$$\alpha_{i \to j} = x_j \langle \pi(e_j), \pi(e_i) \rangle_{\frac{1}{x}}.$$

Corollary

For quasi-thermostatic CRN, the absolute sensitivities α_i satisfy

$$\alpha_i \in [0, 1].$$

Example

Example: the core module of the IDHKP-IDH (IDH = isocitrate dehydrogenase, KP = kinase-phosphatase) glyoxylate bypass regulation system shown in the following reaction scheme:

$$
\begin{aligned}
&\mathrm{E} + \mathrm{I_p} \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} \mathrm{EI_p} \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} \mathrm{E} + \mathrm{I} \\
&\mathrm{EI_p} + \mathrm{I} \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} \mathrm{EI_pI} \underset{k_4^-}{\overset{k_4^+}{\rightleftharpoons}} \mathrm{EI_p} + \mathrm{I_p}.
\end{aligned} \tag{1}
$$

I is the IDH enzyme, $\mathrm{I_p}$ is its phosphorylated form, and E is the bifunctional enzyme IDH kinase-phosphatase. The system obeys approximate concentration robustness in the IDH enzyme I [LaPorte, D. C., Thorsness, P. E., & Koshland, D. E. **(1985)**. Journal of Biological Chemistry]

Absolute concentration robustness [Shinar, G., & Feinberg, M. (2010). Science.]

If $k_2^- = k_4^- = 0$, the systems exhibits absolute concentration robustness in I.

Example

Abbreviate the chemicals as $X_1 = \mathrm{E}, X_2 = \mathrm{I_p}, X_3 = \mathrm{EI_p}, X_4 = \mathrm{I}, X_5 = \mathrm{EI_pI}$ and use $x_i, i = 1, \ldots, 5$ for the respective concentrations.

What happens if $k_2^-, k_4^- > 0$?

For the complex balancing case (corresponds to the equilibrium situation), the absolute sensitivity for X_4 can be given in an analytically closed form based on the previous theorem, i.e., $\alpha_4 = x_4 \langle \pi(e_4), \pi(e_4) \rangle$.
This yields

$$\alpha_4 = \frac{1}{1 + r},$$

where r is given by the ratio

$$r = \frac{(x_2 + x_5)(x_1 + x_3) + x_1 (x_3 + 3x_5)}{x_4 (x_1 + x_3 + x_5)}.$$

Example

What happens if $k_2^-, k_4^- > 0$?

The absolute sensitivity for X_4,

$$\alpha_4 = \frac{1}{1+r},$$

is governed by r is given by the ratio

$$r = \frac{(x_2 + x_5)(x_1 + x_3) + x_1(x_3 + 3x_5)}{x_4(x_1 + x_3 + x_5)}.$$

Approximate concentration robustness

Achieved for $r \gg 0$. This is the case, for example, for $x_1 \approx x_2 \approx x_3 \approx x_5 \gg x_4$, for $x_2 \gg x_1 \approx x_4 \approx x_3 \approx x_5$ as well as for $x_1 \approx x_3 \gg x_2 \approx x_4 \approx x_5$, etc.

High sensitivity is also possible

Achieved for $r \approx 0$. For example, when $x_4 \gg x_1 \approx x_2 \approx x_3 \approx x_5$.

Summary

Caution

I just realized: On the current version of the arxiv preprint, *Remarks* became *Definitions*.

Summary

- The concept of absolute sensitivity might be more suitable to study sensitivity in CRN than then classical sensitivity matrix because the numerical values have meaning.
- Generalizes absolute concentration robustness (ACR):
  - If ACR in X_i holds, then $\alpha_{j\to i} = 0$ for all j.
  - But it might be biologically relevant that $\alpha_{j\to i} = 0$ for some but not all j.
  - Quantifies approximate concentration robustness by $\alpha_i \approx 0$.
- Similarly, hypersensitivity can be quantified by $\alpha_i > 0$.
- Can be explicitly compute for complex-balanced CRN.