Parameter estimation using simulated annealing for S-system models of biochemical networks

Orland Gonzalez
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
S-systems review

• Coupled ODEs
  – Power-laws

• Special form of GMAs
  – Production and degradation fluxes are aggregated

\[
\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}}
\]

for \( i = 1, 2, ..., n \)

\( n \): number of dependent variables

\( m \): number of independent variable
S-system example

Very simple symbolic setup!

$X_1$ production:

$\dot{X}_1^+ = \alpha_1 X_0^{g_{10}} X_2^{g_{12}}$

where: $g_{10} > 0$ since $X_1$ is produced from $X_0$

$g_{12} < 0$ since $X_1$ is inhibited by $X_2$

$X_1$ degradation:

$\dot{X}_1^- = \beta_1 X_1^{h_{11}}$

where: $h_{11} > 0$ since we expect the current state of $X_1$ affects its own degradation rate.

Therefore:

$\dot{X}_1 = \alpha_1 X_0^{g_{10}} X_2^{g_{12}} - \beta_1 X_1^{h_{11}}$
S-system example

Very simple symbolic setup!

Whole Network:

\[ \dot{X}_1 = \alpha_1 X_0^{g_{10}} X_2^{g_{12}} - \beta_1 X_1^{h_{11}} \]
\[ \dot{X}_2 = \alpha_2 X_1^{g_{21}} - \beta_2 X_2^{h_{22}} \]

where \( \alpha_2 = \beta_1 \) and \( h_{11} = g_{21} \) since the amount of \( X_1 \) degraded is the amount of \( X_2 \) that is produced.
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
Parameter Estimation

• From literature
  – From $K_m$s and other enzyme data

• From perturbation experiments
  – From a system in steady state: a single perturbation is made

• From time series data (biochemical profiles)
  – Inverse problem
  – Curve fitting
Inverse problem ex.

Given the network:

And the profile:

Problem: Find the parameters that fit the data.

\[
\begin{align*}
\dot{X}_1 &= X_3 X_5 - X_1 \\
\dot{X}_2 &= X_1 - X_2 \\
\dot{X}_3 &= X_2 - X_3 X_4 \\
\dot{X}_4 &= X_1 - X_4
\end{align*}
\]
Challenges

• Mathematical redundancy
  – Multiple possible solutions
• Computational complexity
  – Abundant local minima
  – Extreme computational requirements
    • due to numerical integration

An early EXTREME example (Kikuchi et. al. 2003, *Bioinformatics*):

• System with 5 variables
• Artificial data
• Cluster of 1040 Pentium-III 933MHz processors

Each algorithmic loop took about 10 hours!!!
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
Simulated annealing

- Optimization technique
  - roughly analogous to glass and metal creation
  - melting and then slow cooling
- Properties
  - solutions less fit than the current one can be accepted (probabilistically)
  - can „escape“ from local minima
Pseudo code

```
SIMULATED_ANNEALING
1      initialize h
2      ctr ← 0
3      T ← T₀
4      while T > Tₘᵟₙᵢₙ
5        for ctr ← 1 to ctrₘₐₓ
6          h' ← h
7            for each kinetic parameter p of h'
8              p ← p + k · ln(√E + 1) · N(\bar{x}, \sigma)
9              ΔE ← Error(h') - Error(h)
10             if ΔE ≤ 0
11                h ← h'
12             else
13                h ← h' with probability \( e^{-\frac{\Delta E}{T}} \)
14        end for
15      lower T
16      end while
```

create initial solution \( h \) (random)
initialize „temperature“ (melting)
slow cooling
create new candidate solution \( h' \)
probabilistically accept \( h' \)
lower the temperature (gradually)

Error:

\[
E = \chi^2 = \frac{\sum_{i=1}^{n} (\hat{X}_{it} - X_{it})^2}{n \cdot N}
\]
evaluated at discrete time points \( t \)
2 Aspects of SA

• Annealing schedule
  – temperature control
  – \( t \leftarrow \alpha t \) where \( \alpha < 1 \)

• Perturbation function
  – creation of candidate solution
  – initial form used

\[
p = p + c \times N(\bar{x}, \sigma) \quad \text{for each parameter } p
\]

where \( c \) is a constant
Perturbation function

\[ p = p + c \cdot N(\bar{x}, \sigma) \quad \text{for each parameter } p \]

• Problems
  – does not converge for too large \( c \)
  – frequently gets trapped with too small \( c \)

• What we did
  – convert \( c \) to a function of the error \( E \)
Construction of p-function

\[ E_{it} = (X_{it} - \hat{X}_{it})^2 \]

error of a specific metabolite \( X_i \) and a specific time \( t \)

\[ X_{it} \approx \hat{X}_{it} = X_{it-1} + \dot{X}_{it-1} \Delta t \]

first order approx. of \( X_{it} \) given \( X_{it-1} \)

substitution yields

\[ \sqrt{E_{it}} = X_{it} - (X_{it-1} + \dot{X}_{it-1} \Delta t) \]

\[ = X_{it} - \left( X_{it-1} + \left( \alpha_i \prod_{j=1}^{n+m} X_{jt-1}^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_{jt-1}^{h_{ij}} \right) \Delta t \right) \]
Suppose it is possible to reduce $E_{it}$ to a factor of $p < 0$ by perturbing a particular $g_{ik}$ by $\Delta g$, this relationship is described by

$$\sqrt{pE_{it}} = X_{it} - \left( X_{it-1} + \left( X_{k_{it-1}}^{\Delta g} \prod_{j=1}^{n+m} X_{jt-1}^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_{jt-1}^{h_{ij}} \right) \Delta t \right)$$

Subtracting this from the original relationship, doing algebraic manipulations, shifting to logarithmic space, and isolating $\Delta g$ yields

$$\Delta g = c_1 \ln \left( \frac{(1 - \sqrt{p}) \sqrt{E_{it}}}{\alpha_i \Delta t \prod_{j=1}^{n+m} X_{jt-1}^{g_{ij}} + 1} \right)$$
Direct use of equation

\[ \Delta g = c_1 \ln \left( \frac{(1 - \sqrt{p}) \sqrt{E_{it}}}{\alpha_i \Delta t \prod_{j=1}^{n+m} X_{ij}^{g_{ij}} + 1} \right) \]

• Problems
  – Computationally expensive
    • Aggregation for each parameter over all possible \( t \)
  – How to aggregate?

• Note: Aggregation actually makes it possible to increase the overall error \( E \).
Indirect use of equation

• Defines a roughly linear relationship between $\Delta g$ and $c_1 \ln(\sqrt{E_{it}})$

Note: Consistent with BST theory of linearization in logarithmic space.

Perturbation function:

$$p = p + l \times \ln(\sqrt{E} + 1) \times N(\bar{x}, \sigma)$$

where $l$ is externally optimized.
Graph of perturbation function

- Perturbation magnitudes go down with the error
- Encourages exploration of promising areas of the solution space
  - The lower $E$ gets, the smaller the average magnitudes are
Demonstration

Generate artificial data

Pretend we do not know the true parameters.
Try to estimate them using the generated artificial data set (4 sets in this example)

\[
\begin{align*}
\dot{X}_1 &= 20X_3^{-0.8} X_5 - 10X_1^{0.5} \\
\dot{X}_2 &= 8X_1^{0.5} - 3X_2^{0.75} \\
\dot{X}_3 &= 3X_2^{0.75} - 5X_3^{0.5} X_4^{0.2} \\
\dot{X}_4 &= 2X_1^{0.5} - 6X_4^{0.8} \\
X_5 &= 0.9
\end{align*}
\]
Parameter estimation

<table>
<thead>
<tr>
<th>True parameter set(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1) 20.0 0 0 -0.8 0 1.0</td>
</tr>
<tr>
<td>(X_3) 3.0 0 0.75 0 0 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W/o structure identification(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1) 20.3541 0 0 -0.8163 0 1.0117</td>
</tr>
<tr>
<td>(X_3) 2.7970 0 0.7809 0 0 0</td>
</tr>
</tbody>
</table>

- 18 free parameters
- Initial solution randomly created
  - [0, 90] for rate constants
  - [-2, 2] for kinetic orders (based on general observations)
- 10 hours on a 1.5GHz Intel Celeron machine
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
Decoupling / slope estimation

- Approximate slopes
  - Finite difference, splines, etc...
- Equate them directly to the ODE's
- Advantages
  - removes the need for numerical integration
  - runtime dropped to under 2 mins per iteration!
- Disadvantages
  - additional search bias
  - susceptibility to noise
  - not always applicable
Structure identification

- Try to identify structure
  - simple due to canonical form of S-systems
  - leave all/most parameters free
  - let optimization determine structure

- Note: still in infancy but can be useful for limited applications
  - ex. leaving only alternative parameters free (alternative regulations)
...cont’d

For Constituent X1:

<table>
<thead>
<tr>
<th></th>
<th>$a_1$</th>
<th>$g_{11}$</th>
<th>$g_{12}$</th>
<th>$g_{13}$</th>
<th>$g_{14}$</th>
<th>$g_{15}$</th>
<th>$\beta_1$</th>
<th>$h_{11}$</th>
<th>$h_{12}$</th>
<th>$h_{13}$</th>
<th>$h_{14}$</th>
<th>$h_{15}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>18.2654</td>
<td>0</td>
<td>-0.0595</td>
<td>-0.8355</td>
<td>-0.0594</td>
<td>1.0156</td>
<td>8.6238</td>
<td>0.5567</td>
<td>-0.0700</td>
<td>0.0896</td>
<td>-0.0294</td>
<td>-0.1333</td>
</tr>
<tr>
<td>Step 2</td>
<td>20.4407</td>
<td>0</td>
<td>0</td>
<td>-0.8199</td>
<td>0</td>
<td>1.0316</td>
<td>10.0049</td>
<td>0.5049</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0142</td>
</tr>
<tr>
<td>Step 3</td>
<td>20.3541</td>
<td>0</td>
<td>0</td>
<td>-0.8163</td>
<td>0</td>
<td>1.0117</td>
<td>10.0037</td>
<td>0.5039</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Outline

• S-systems quick review
• Definition of the problem
• Simulated annealing
  – Perturbation function
  – Demonstration
• General techniques
  – Decoupling
  – Structure identification
• The cadBA network in *E. coli*
The cadBA network in *E. coli*

- Believed to be part of coping mechanism for acidic environments
- Has regulatory (transcriptional) and metabolic elements
Existing cadBA model (CMA)

**Similarity:** We used the same data set.

**Difference:** We tried to avoid speculating on dynamics of components with no data.

\[
\frac{d}{dt} [\text{CmaC}] = -k_1 [\text{LysP}] [\text{CmaC}] + k_2 [\text{Cma}] [\text{LysP} \cdot \text{CmaC}]
+ k_3 [\text{H}_{\text{wA}}] [\text{CmaC}] - k_{29} [\text{CmaC}] \tag{1}
\]

\[
\frac{d}{dt} [\text{CmaC}^*] = -k_4 [\text{Cma}] [\text{CmaC}^*] + k_5 [\text{LysP} \cdot \text{CmaC}^*]
+ k_6 [\text{H}_{\text{wA}}] [\text{CmaC}^*] - k_{29} [\text{CmaC}^*] \tag{2}
\]

\[
\frac{d}{dt} [\text{LysP} \cdot \text{CmaC}] = k_1 [\text{LysP}] [\text{CmaC}] - k_7 [\text{LysP} \cdot \text{CmaC}]
+ k_8 [\text{H}_{\text{wA}}] [\text{LysP} \cdot \text{CmaC}] + k_{38} [\text{LysP} \cdot \text{CmaC}] \tag{3}
\]

\[
\frac{d}{dt} [\text{LysP} \cdot \text{CmaC}^*] = k_9 [\text{Cma}] [\text{LysP} \cdot \text{CmaC}^*] - k_10 [\text{LysP} \cdot \text{CmaC}^*]
+ k_{38} [\text{LysP} \cdot \text{CmaC}^*] \tag{4}
\]

\[
\frac{d}{dt} [\text{cadBA}] = \frac{k_5 [\text{CmaC}^*]^2}{[\text{CmaC}^*]^2 + k_8} - k_11 [\text{cadBA}] \tag{5}
\]

\[
\frac{d}{dt} [\text{CmaA}] = k_{12} [\text{cadBA}] - k_{12} [\text{CmaA}] \tag{6}
\]

\[
\frac{d}{dt} [\text{CmaB}] = k_{12} [\text{cadBA}] - k_{13} [\text{CmaB}] \tag{7}
\]

\[
\frac{d}{dt} [\text{Cma}] = -f_{\text{trans}} - k_14 [\text{LysP} \cdot \text{Cma}] \tag{8}
\]

\[
\frac{d}{dt} [\text{Cma}] = f_{\text{trans}} \tag{9}
\]

\[
\frac{d}{dt} [\text{H}_{\text{wA}}] = -f_{\text{trans}} - k_4 [\text{H}_{\text{wA}}] [\text{CmaC}] \tag{10}
\]

\[
\frac{d}{dt} [\text{H}_{\text{wA}}] = f_{\text{trans}} - k_{15} [\text{CmaA}] [\text{H}_{\text{wA}}] + k_{14} [\text{LysP} \cdot \text{Cma}] [\text{H}_{\text{wA}}] \tag{11}
\]

\[
\frac{d}{dt} [\text{H}_{\text{wA}}] = -f_{\text{trans}} - k_{15} [\text{CmaA}] [\text{H}_{\text{wA}}] \tag{12}
\]

\[
\frac{d}{dt} [\text{H}_{\text{wA}}] = f_{\text{trans}} - k_{15} [\text{CmaA}] [\text{H}_{\text{wA}}] \tag{13}
\]

\[
\frac{d}{dt} [\text{LysP}] = \frac{k_{20}}{1 + \frac{[\text{LysP}]}{k_{20}}} - k_3 [\text{LysP}] \tag{14}
\]

\[
\frac{d}{dt} [\text{LysP}] = k_{38} [\text{LysP}] - k_{38} [\text{LysP}] + k_5 [\text{LysP} \cdot \text{CmaC}] - k_7 [\text{LysP} \cdot \text{CmaC}] + k_8 [\text{LysP} \cdot \text{CmaC}] \tag{15}
\]

\[
f_{\text{trans}}([\text{Cma}], [\text{Cma}], [\text{H}_{\text{wA}}], [\text{LysP}]) = \frac{k_{39} [\text{Cma}] [\text{Cma}] [\text{H}_{\text{wA}}] [\text{LysP}]}{[\text{H}_{\text{wA}}] f_{\text{trans}} + k_{22}} \tag{16}
\]
Existing cadBA model (CMA)

\[
\frac{d}{dt}[CmC] = -k_3 [LysP] [CmC] + k_2 [x_{cm}] [LysP \cdot CmC] \\
- k_8 [H_{cm}^+] [CmC] + k_4 [x_{cm}] [CmC^+] + k_{23} [CmC^+] (1)
\]

\[
\frac{d}{dt}[CmC^+] = -k_4 [x_{cm}] [CmC^+] + k_5 [x_{cm}] [LysP \cdot CmC^+] \\
+ k_8 [H_{cm}^+] [CmC] - k_{23} [CmC^+] (2)
\]

\[
\frac{d}{dt}[LysP \cdot CmC] = k_1 [LysP] [CmC] - k_3 [x_{cm}] [LysP \cdot CmC] \\
- k_6 [H_{cm}^+] [LysP \cdot CmC] + k_{24} [LysP \cdot CmC] (3)
\]

\[
\frac{d}{dt}[LysP \cdot CmC^+] = k_6 [H_{cm}^+] [LysP \cdot CmC] - k_8 [x_{cm}] [LysP \cdot CmC^+] \\
- k_{24} [LysP \cdot CmC^+] (4)
\]

\[
\frac{d}{dt}[CmBA] = \frac{k_5 [CmC]^2}{[CmC]^2} - k_9 [x_{cm}] [CmBA] (5)
\]

\[
\frac{d}{dt}[CmA] = k_{12} [x_{cm}] [CmBA] - k_{11} [CmA] (6)
\]

\[
\frac{d}{dt}[CmB] = k_{12} [x_{cm}] [CmBA] - k_{13} [CmB] (7)
\]

\[
\frac{d}{dt}[x_{cm}] = -f_{x_{cm}} (8)
\]

\[
\frac{d}{dt}[H_{cm}^+] = -f_{x_{cm}} - k_7 [H_{cm}^+] [CmC] (9)
\]

\[
\frac{d}{dt}[x_{cm}] = f_{x_{cm}} - k_{15} [CmA] [x_{cm}] [H_{cm}^+] + k_{14} [LysP] [x_{cm}] (10)
\]

\[
\frac{d}{dt}[H_{cm}^+] = f_{x_{cm}} - k_{15} [CmA] [x_{cm}] [H_{cm}^+] (11)
\]

\[
\frac{d}{dt}[H_{cm}^+] = -f_{x_{cm}} - k_{15} [CmA] [x_{cm}] [H_{cm}^+] (12)
\]

\[
\frac{d}{dt}[H_{cm}^+] = f_{x_{cm}} - k_{13} [CmBA] [x_{cm}] [H_{cm}^+] (13)
\]

\[
\frac{d}{dt}[LysP] = \frac{k_{16}}{1 + \frac{[H_{cm}^+]}{K_{eq}}} - k_{18} [x_{cm}] (14)
\]

\[
\frac{d}{dt}[LysP] = k_{18} [LysP] - k_{16} [LysP] + k_3 [x_{cm}] [LysP \cdot CmC^+] \\
- k_1 [LysP] [CmC] + k_2 [x_{cm}] [LysP \cdot CmC] (15)
\]

\[
f_{x_{cm}}([CmB], [x_{cm}], [H_{cm}^+], [x_{cm}]) = \frac{k_{21} [x_{cm}] [CmB] [H_{cm}^+] [x_{cm}]}{[H_{cm}^+] f_{x_{cm}} + k_{22}} (16)
\]
Problems & handling

• Data availability
  – CadC is constitutive
    • hard to measure active form
    • bypassed by redirecting regulatory signals from pH and lysine to the equation for the transcript
  – No data for internal lysine and cadaverine
    • decarboxylation and transport activities were coupled
Working model

\[
\frac{d[CadA]}{dt} = 0.2902[cadBA]^{0.8875} - 0.0329[CadA]^{0.9612} \\
\frac{d[cadBA]}{dt} = 0.1142[Lys]^{1.3308}[H^+]^{2.3677} - 4.7892[CadA]^{0.6551} \\
\frac{d[Cadav]}{dt} = 0.2978[CadA]^{0.909}[Cadav]^{-0.507}[Lys]^{0.492} \\
\frac{d[Lys]}{dt} = -0.3458[CadA]^{0.909}[Cadav]^{-0.507}[Lys]^{0.524}
\]

Network with estimated parameters

Model fit

Uncertain reason (data for lysine measured separately from the rest)
Working model

Network with estimated parameters

play around: (neutral pH)

\[
\frac{d[CadA]}{dt} = 0.2902[cadBA]^{0.8875} - 0.0329[CadA]^{0.9612} \\
\frac{d[cadBA]}{dt} = 0.1142[Lys]^{1.3308}[H^+]^{2.3677} - 4.7892[CadA]^{0.6551} \\
\frac{d[Cadav]}{dt} = 0.2978[CadA]^{0.909}[Cadav]^{-0.507}[Lys] \\
\frac{d[Lys]}{dt} = -0.3458[CadA]^{0.909}[Cadav]^{-0.507}[Lys]
\]
cadBA model

• Challenges left
  – nature of data
    • CadA in arbitrary units (activity)
    • cadBA in relative intensity

• Further work
  – include other players
    • lysP and LysP
  – determine correctness of hypothesized structure
Possible SA refinements

• Annealing schedule
  – quick convergence heuristics

• Perturbation function
  – perturbation by parameter subsets
    • probably useful for handling larger problems
  – solution histories
    • momentum terms
  – automatic control
  – direct use of derived equation
Thanks!!!!

Reference

BIOINFORMATICS

Parameter Estimation using Simulated Annealing for S-System Models of Biochemical Networks

Orland R. Gonzalez a,*, Christoph Küper b, c, Kirsten Jung b, Prospero C. Naval, Jr. a, Eduardo Mendoza d, e

aDepartment of Computer Science University of the Philippines-Diliman, bDepartment Biologie I, Bereich Mikrobiologie, Ludwig-Maximilians-Universität, cMedizinische Fakultät, Physiologisches Institut, Ludwig-Maximilians-Universität, dMathematics Department University of the Philippines-Diliman, ePhysics Department & Center for NanoScience Ludwig-Maximilians-University Munich

Associate Editor: Martin Bishop