Origins of sloppiness in biological networks

Joshua Waterfall

DOE CSGF Meeting
June 20, 2006
Growth factor signaling in PC12 cells

\[
d\frac{[SOSActive]}{dt} = \frac{k_{EGF}[boundEGFR][SOSInactive]}{[SOSInactive] + K_{mEGF}}
\]
Prediction for Erk activation by EGF with PI3K inhibitor

10 time-series experiments with ~15% uncertainties
All parameters known within a factor of one million
All 48 parameters known within +/- 25%
47 parameters known within +/- 25%, Km for Mek phosphatase known within +/- factor of 30
10 time-series experiments with ~15% uncertainties
All parameters known within a factor of one million
Parameter Fluctuations and Systems Fluctuations

\[ \chi^2(\theta) = \sum_{\text{species}} \frac{1}{T} \int_0^T \frac{(y(\theta,t) - y(\theta_0,t))^2}{\sigma^2} \, dt \]

\[ H_{ij} = \frac{\partial^2 \chi^2}{\partial \log \theta_i \partial \log \theta_j} \]
(a) Growth factor signaling in PC12 cells, (b) Quantum Monte Carlo (Jastrow parameters), (c) radioactivity from pharmaceutical isotopes, (d) 48 random exponentials from log uniform(±25.), (e) random symmetric matrix (GOE), (f) products of GOE random matrices, (g) Wishart matrix, (h) fitting monomials (Hilbert matrix)
Sums of exponentials

\[ y(x) = \sum_{i=0}^{n} e^{-\gamma_i x} \]

\[ H_{ij} = \frac{2\gamma_i \gamma_j}{(\gamma_i + \gamma_j)^3} \]

Linear Correlations

\[ a(\vec{p}) = \vec{n} \cdot \vec{p} \]

Hessian is covariance matrix of data.

Wishart statistics for eigenvalues.

\[ H = \sum_i \vec{p}_{(i)} \vec{p}_{(i)}^T \]

SLOPPY

NOT SLOPPY
Sums of monomials

\[ H_{ij} = \frac{2}{i + j + 1} \]

Hessian = Hilbert matrix

\[ \sqrt{|H|} = \text{volume change under basis transformation} \]

Hessians

\[ H_{ij} = \delta_{ij} \]

Sums of Shifted Legendre polynomials

\[ \text{eigenvalue} \]

SLOPPY

NOT SLOPPY
The Vandermonde Ensemble

Assumptions for proof:
1. Parameters are nearly degenerate.
2. Cost function is sum of squares of residuals.
3. Each residual is a symmetric function of all the parameters.

\[ \theta_i = \theta_0 + \epsilon_i \quad C(\bar{\theta}) = \sum_k r_k^2(\{m_i\}) \quad m_i = \sum_j \epsilon_j^i \]

\[ J_{ij} = \frac{\partial r_i}{\partial \theta_j} = \sum_k \frac{\partial r_i}{\partial m_k} k \epsilon_j^{k-1} = A_{ik} V_{kj} \]

\[ H = J^T J = V^T A^T A V \]

\[ \det(V) = \prod_{i<j} (\epsilon_i - \epsilon_j) \propto \epsilon^{N(N-1)/2} \]
The Vandermonde Ensemble

\[ H = J^T J = V^T A^T AV = W \tilde{H} W^T \]

\[ \Sigma U^T A^T A U \Sigma = \tilde{H} \]

\[ \begin{bmatrix} \epsilon^0 \\ \vdots \\ \epsilon^{n-1} \end{bmatrix} \begin{bmatrix} \epsilon^0 & \cdot & \cdot & \cdot \\ & \epsilon^0 & \cdot & \cdot \\ & & \ddots & \cdot \\ & & & \cdot \epsilon^{n-1} \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \]

\[ \tilde{H}_{ij} = O(\epsilon^{i+j-2}) \]

\[ \tilde{\lambda}^{(i)} = O(\epsilon^{2(i-1)}) \]

\[ \tilde{\nu}^{(i)} = \delta_{ij} + O(\epsilon^{|i-j|}) \]

\[ \lambda^{(i)} = O(\epsilon^{2(i-1)}) \]

\[ \nu^{(i)} = W_{ij} + O(\epsilon) \]
Conclusions

- Large biological networks are governed by a small number of combinations of parameters.
- The origin of this sloppiness is the redundant behavior of the constituent biochemical reactions.
- We have found a new class of “sloppy” models, which includes those of biological networks, and are defined by being a Vandermonde transformation of a random symmetric matrix.

Acknowledgements

Fergal Casey  
Ryan Gutenkunst  
Kevin Brown  
Christopher Myers  
James Sethna

Veit Elser  
Piet Brouwer

DOE  
CSGF  
USDA