

Porównanie różnych podejść typu ODE do modelowania sieci regulacji genów

Michał Woźniak

8 stycznia 2010

Plan prezentacji

- 1 Wprowadzenie
 - Praca źródłowa
 - Sieci regulacji genów
- 2 Modele ODE
 - Założenia
 - Funkcja Hill'a
 - Modele dyskretne
 - Modele ODE
- 3 Reprezentatywny przykład
 - Przykład
 - Modele dla przykładu
 - Model dyskretny
- 4 Analiza modeli
 - Wprowadzenie
 - Analiza teoretyczna
 - Przykłady zachowania modeli
 - Linearyzacja - modele CPWLM i SPWLM
 - Efekty dyskretyzacji
- 5 Podsumowanie

Praca źródłowa

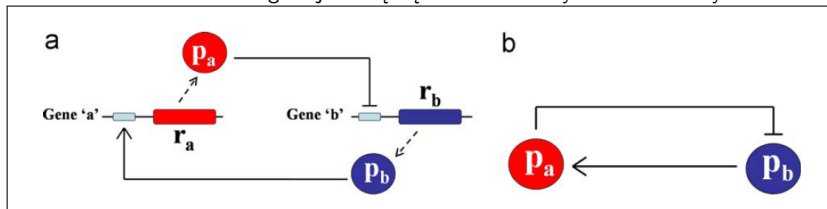
Comparing different ODE modelling approaches for gene regulatory networks

A. Polynikis, S.J. Hogan, M. di Bernardo

Journal of Theoretical Biology

Sieci regulacji genów

Sieć regulacji genów jest to sposób organizacji zależności w procesie tworzenia białek. Podstawowe modele sieci regulacji dzielą się na deterministyczne i stochastyczne.



Przykłady podejść do modelowania

Przykłady podejść do modelowania sieci regulacji genów:

- **równania różniczkowe**
- modele stochastyczne
- sieci boolowskie
- modele Bayesowskie
- ...

Ordinary Differential Equations (ODE) Model

Zazwyczaj do modelowania dynamiki procesów transkrypcji i translacji w podejściu ODE używa się takich równań:

- Transkrypcja: $\frac{dr_i}{dt} = F(f_i^R(p_1), f_i^R(p_2), \dots, f_i^R(p_n)) - \gamma_i r_i$,
- Translacja: $\frac{dp_i}{dt} = f_i^P(r_i) - \delta_i p_i$,

gdzie:

- funkcje $f_i^R(p_j) : \mathcal{R} \rightarrow \mathcal{R}$ wyrażają zależność koncentracji mRNA od koncentracji białka p_j ,
- funkcja $F : \mathcal{R}^n \rightarrow \mathcal{R}$ zazwyczaj jest zdefiniowana jako suma lub iloczyn funkcji f_j^R ,
- funkcja $f_i^P(r_i)$ opisuje dynamikę translacji mRNA r_i w białko p_i ,
- stałe γ_i, δ_i reprezentują parametry degradacji mRNA i białek i-tego genu.

Funkcja Hill'a

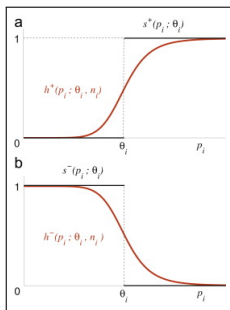
Eksperymenty potwierdzają monotoniczny kształt funkcji f_i^R , która rośnie gdy białko p_i jest aktywatorem, a maleje gdy p_i jest inhibitorem. Funkcja Hill'a to użyteczna funkcja spełniająca tę własność. Funkcja Hill'a:

- dla aktywatora: $h^+(p_i; \theta_i, n_i) = \frac{p_i^{n_i}}{p_i^{n_i} + \theta_i^{n_i}}$
- dla inhibitora: $h^-(p_i; \theta_i, n_i) = 1 - h^+(p_i; \theta_i, n_i) = \frac{\theta_i^{n_i}}{p_i^{n_i} + \theta_i^{n_i}}$
- parametr θ_i , to próg ekspresji
- parametr n_i , nazywa się współczynnikiem Hill'a, kontroluje skokowość funkcji

Funkcja schodkowa

Dla sieci wielu genów często nie mogą być wyznaczone rozwiązania równań różniczkowych, w których występuje funkcja Hill'a. W związku z tym często analizuje się schodkową wersję funkcji Hill'a, postaci:

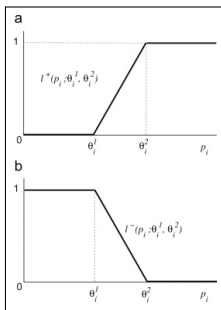
$$s^+(p_i, \theta_i) = \mathbf{1}_{\{p_i > \theta_i\}}, \quad s^-(p_i, \theta_i) = 1 - s^+(p_i, \theta_i)$$



Funkcja trójkątna

Niestety, funkcja schodkowa nie jest ciągła, co intuicyjnie źle nadaje się do modelowania procesów regulacji. Aby rozwiązać ten problem stosuje się funkcję "trójkątną":

$$I^+(p_i, \theta_i^1, \theta_i^2) = \begin{cases} 0 & \text{jeżeli: } p_i < \theta_i^1 \\ 1 & \text{jeżeli: } p_i > \theta_i^2 \\ \mu p_i + \lambda & \text{wpp} \end{cases} \quad I^-(p_i, \theta_i^1, \theta_i^2) = 1 - I^+(p_i, \theta_i^1, \theta_i^2)$$



Założenie quasi-steady-state dla koncentracji mRNA

Dla wielu modeli przyjmowane jest upraszczające założenie, że koncentracja mRNA znajduje się w stanie stabilnym, bazuje ono na fakcie, że dynamika mRNA jest znacznie szybsza niż białek, co prowadzi do szybszego uzyskania stanu stabilnego przez mRNA. Powyższe założenie możemy modelować matematycznie przyjmując: $\dot{r}_i = 0$, wtedy:

$$r_i = \frac{1}{\gamma_i} F(f_i^R(p_1), f_i^R(p_2), \dots, f_i^R(p_n)),$$

po podstawieniu do równania opisującego translację:

$$\dot{p}_i = f_i^P \left(\frac{1}{\gamma_i} F(f_i^R(p_1), f_i^R(p_2), \dots, f_i^R(p_n)) \right) - \delta_i p_i$$

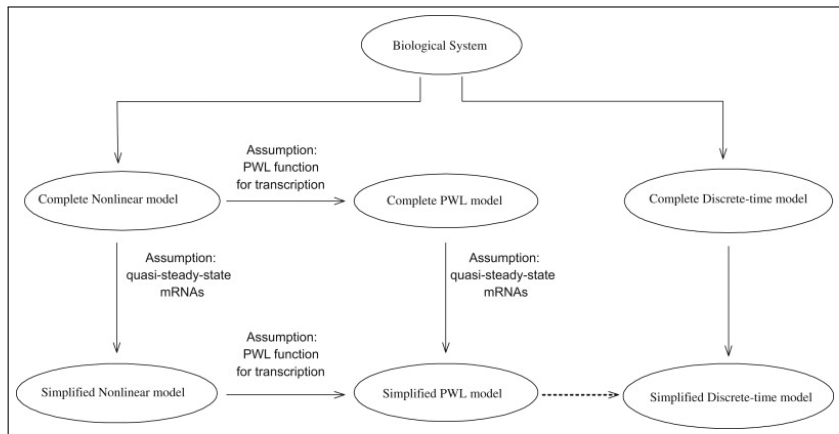
Modele z czasem dyskretnym

Idea: Napisanie modelu ODE (równań różniczkowych) i jego dyskretyzacja. Modele dyskretne wykorzystują założenie stanu quasi-stabilnego dla koncentracji mRNA.

Modele przeanalizowane w pracy

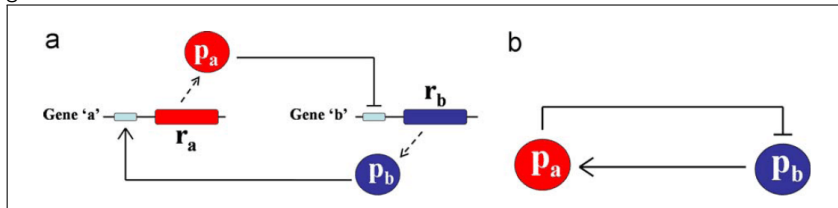
- CNM - complete nonlinear model (funkcja Hill'a dla transkrypcji, funkcje liniowe dla translacji)
- CPWLM - complete piecewise linear model (zamiast funkcji Hill'a funkcja schodkowa)
- SNM - simplified nonlinear model (CNM + założenie o quasi stabilności)
- SPWLM - simplified piecewise linear model (SNM + zamiast funkcji Hill'a funkcja schodkowa)
- discrete time model - (Coutinho et al. 2006), dyskretyzacja czasu w SPWLM

Modele przeanalizowane w pracy



Przykład

Aby zaprezentować zalety i wady wybranych modeli użyto sieci złożonej z dwóch genów:



W dalszych rozważaniach będziemy przyjmować $i = a, b$.

Model CNM dla przykładowej sieci

$$\dot{r}_a = m_a h^+(p_b; \theta_b, n_b) - \gamma_a r_a,$$

$$\dot{r}_b = m_b h^+(p_a; \theta_a, n_a) - \gamma_b r_b,$$

$$\dot{p}_a = k_a r_a - \delta_a p_a$$

$$\dot{p}_b = k_b r_b - \delta_b p_b$$

Model SNM dla przykładowej sieci

W poprzednim modelu (CNM) przyjmujemy założenie: $\dot{r}_a \simeq 0$ i $\dot{r}_b \simeq 0$, wtedy:

$$r_a = \frac{m_a}{\gamma_a} h^+(p_b; \theta_b, n_b)$$

$$r_b = \frac{m_b}{\gamma_b} h^-(p_a; \theta_a, n_a)$$

$$\dot{p}_a = k'_a h^+(p_b; \theta_b, n_b) - \delta_a p_a$$

$$\dot{p}_b = k'_b h^-(p_a; \theta_a, n_a) - \delta_b p_b$$

gdzie:

$$k'_a = \frac{m_a}{\gamma_a} k_a, \quad k'_b = \frac{m_b}{\gamma_b} k_b$$

Modele CPWLM i SPWLM dla przykładowej sieci

$$\begin{aligned}\dot{r}_a &= m_a s^+(p_b; \theta_b, n_b) - \gamma_a r_a, \\ \dot{r}_b &= m_a s^-(p_a; \theta_a, n_a) - \gamma_b r_b, \\ \dot{p}_a &= k_a r_a - \delta_a p_a \\ \dot{p}_b &= k_b r_b - \delta_b p_b\end{aligned}$$

Dla modelu SPWLM otrzymujemy:

$$\begin{aligned}\dot{p}_a &= k'_a s^+(p_b; \theta_b, n_b) - \delta_a p_a \\ \dot{p}_b &= k'_b s^-(p_a; \theta_a, n_a) - \delta_b p_b\end{aligned}$$

Model dyskretny

Zapiszmy równania z modelu SPWLM na koncentrację białek w następującej postaci:

$$\dot{\mathbf{p}} = \mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{u},$$

gdzie:

$$\mathbf{p} = \begin{pmatrix} p_a \\ p_b \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} -\delta_a & 0 \\ 0 & -\delta_b \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} k'_a & 0 \\ 0 & k'_b \end{pmatrix},$$
$$\mathbf{u} = \begin{pmatrix} s^+(p_b; \theta_b) \\ s^-(p_a; \theta_a) \end{pmatrix}.$$

Rozwiązując powyższe równanie otrzymujemy:

$$\mathbf{p}(t) = e^{\mathbf{A}t} \mathbf{p}(0) + (e^{\mathbf{A}t} - \mathbb{I}d) \mathbf{A}^{-1} \mathbf{B} \mathbf{u}$$

Model dyskretny II

Dla dostatecznie małego czasu T , otrzymujemy:

$$\mathbf{p}(t+T) = \begin{pmatrix} e^{-\delta_a(t+T)} & 0 \\ 0 & e^{-\delta_b(t+T)} \end{pmatrix} \begin{pmatrix} p_{a_0} \\ p_{b_0} \end{pmatrix} + \begin{pmatrix} -\frac{k'_a}{\delta_a}(e^{-\delta_a(t+T)} - 1) & 0 \\ 0 & -\frac{k'_b}{\delta_b}(e^{-\delta_b(t+T)} - 1) \end{pmatrix} \begin{pmatrix} s^+(p_{b_0}; \theta_b) \\ s^-(p_{a_0}; \theta_a) \end{pmatrix}$$

Przyjmujemy $t = 0$, wtedy otrzymujemy:

$$p_a(T) = e^{-\delta_a T} p_a(0) + \frac{k'_a}{\delta_a} (1 - e^{-\delta_a T}) s^+(p_b(0); \theta_b)$$
$$p_b(T) = e^{-\delta_b T} p_b(0) + \frac{k'_b}{\delta_b} (1 - e^{-\delta_b T}) s^-(p_a(0); \theta_a)$$

Model dyskretny III

Przyjmujemy $T = 1$, niech $p_a(0) = p_a(n)$ oraz $p_a(T) = p_a(n + 1)$, wtedy otrzymujemy rekurencyjne wzory na $p_a(n)$:

$$p_a(n + 1) = e^{-\delta_a} p_a(n) + \frac{k'_a}{\delta_a} (1 - e^{-\delta_a}) s^+(p_b(n); \theta_b)$$
$$p_b(n + 1) = e^{-\delta_b} p_b(n) + \frac{k'_b}{\delta_b} (1 - e^{-\delta_b}) s^-(p_a(n); \theta_a)$$

Robiąc kolejne uproszczenie: $\delta_a = k'_a = \delta_b = k'_b$ oraz $\alpha = e^{-\delta_a}$, otrzymujemy:

$$p_a(n + 1) = \alpha p_a(n) + (1 - \alpha) s^+(p_b(n); \theta_b)$$
$$p_b(n + 1) = \alpha p_b(n) + (1 - \alpha) s^-(p_a(n); \theta_a)$$

Ten model został zaproponowany przez Coutinho w 2006.

Analiza

Analiza została przeprowadzona dla modeli opisanych wcześniej na przykładzie sieci składającej się z dwóch genów (również jak wcześniej).

Analiza CNM i SNM

Startujemy od modelu CNM. Przyjmijmy, że znajdujemy się w stanie stabilnym, czyli: $\dot{r}_a = \dot{r}_b = \dot{p}_a = \dot{p}_b = 0$. Oznaczmy przez $\hat{r}_a, \hat{r}_b, \hat{p}_a, \hat{p}_b$ wartości koncentracji mRNA i białek w stanie stabilnym. Zgodnie ze wzorami:

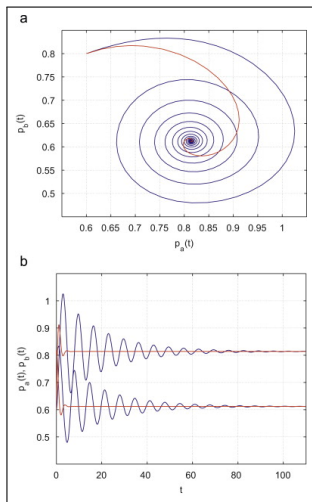
$$\dot{p}_a = k_a r_a - \delta_a p_a, \quad \dot{p}_b = k_b r_b - \delta_b p_b,$$

otrzymujemy:

$$\hat{r}_a = \frac{\delta_a}{k_a} \hat{p}_a, \quad \hat{r}_b = \frac{\delta_b}{k_b} \hat{p}_b,$$

stąd można uzyskać zależności dla: $\hat{r}_a, \hat{r}_b, \hat{r}_a, \hat{p}_b$.

Analiza CNM i SNM - zbieżność modeli

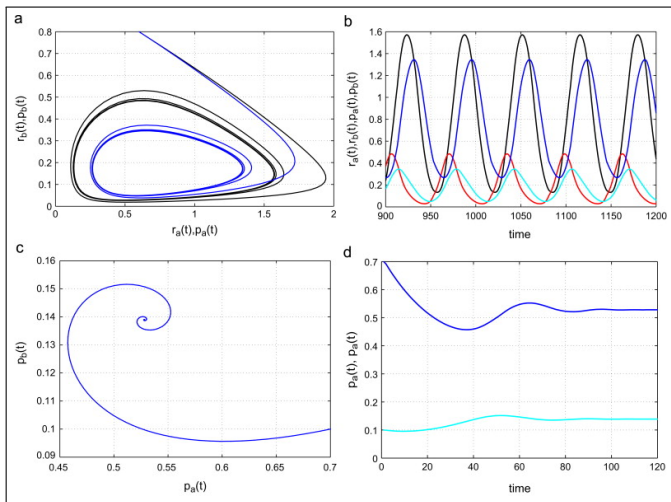


Zachowanie (p_a, p_b) :

- Czerwona linia - SNM
- Czerwona linia - CNM

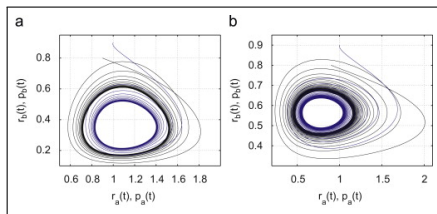
Wniosek: SNM zbiega szybciej

Analiza CNM i SNM - różnice



W pracy zostało analitycznie udowodnione, że model SNM (dla potencjalnie interesujących parametrów) jest zawsze zbieżny, natomiast CNM niekoniecznie. Dla pewnych parametrów w modelu CNM powstaje bifurkacja Hopfa- w portrecie fazowym powstaje okrąg, wokół którego koncentrują się trajektorie.

Analiza CNM - bifurkacja Hopfa



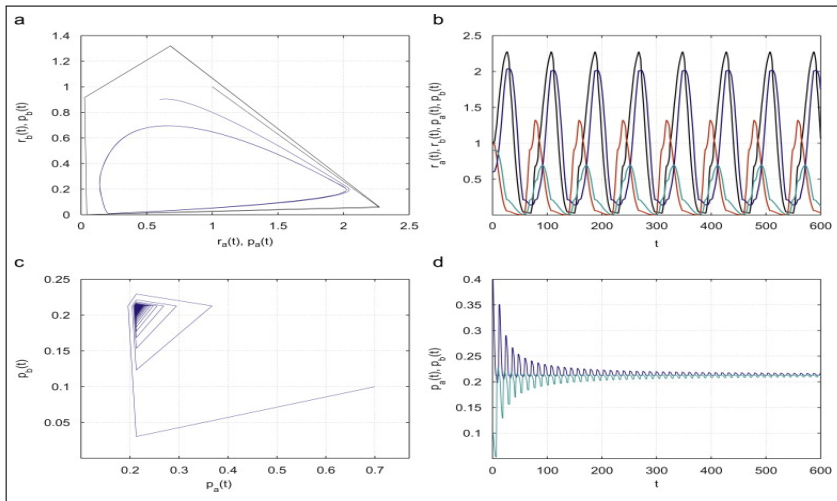
Wykresy ukazujące bifurkację Hopfa w modelu CNM w zależności od różnych parametrów n_a, n_b w modelu CNM.

Analiza teoretyczna

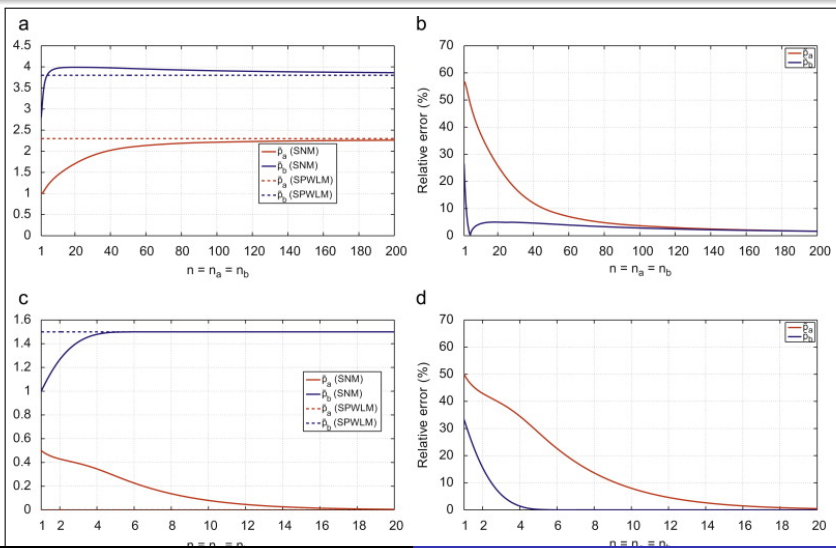
Zachowanie modeli CPWLM i SPWLM jest analogiczne do ich odpowiedników CNM i SNM opartych na funkcji Hill'a, tzn.:

- model SPWLM jest zawsze zbieżny, a CPWLM niekoniecznie
- SPWLM jest również szybciej zbieżny.

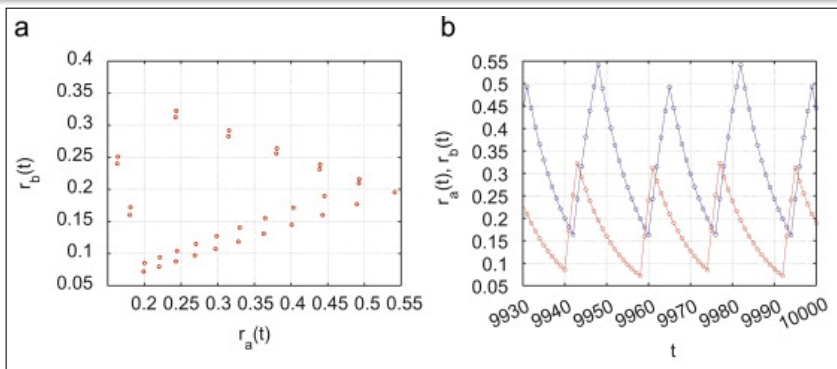
Zachowanie modeli CPWLM i SPWLM



Porównanie SNM i SPWLM



Model dyskretny



- najważniejsza różnica pomiędzy modelami dyskretnymi a ciągłymi to taka, że modele dyskretne zawsze dają rozwiązania okresowo oscylujące, niezależnie od doboru parametrów.
- dla modeli CNM i CPWLM okresowe oscylacje istnieją tylko dla specyficznych wartości parametrów
- dla modeli SNM i SPWLM układ równań nie wpada w okresowe oscylacje

Wnioski

- Często wyniki dla różnych modeli są bardzo różne, nawet w przypadku prostej sieci.
- Brakuje metod pozwalających rozstrzygnąć kompromis pomiędzy złożonością modelu a jakością predykcji.
- Nie ma pełnego zrozumienia wpływu poszczególnych założeń modeli na całościowe zachowanie modelu oraz jego jakość.