Biological Information Modeling

September 7, 2018

Abstract

Keywords:
Mathematics Subject Classification 2000:

1 Introduction

The describing of the biological problems :
1. the background information (by WX)
2. the expression of the biological experiments (by WX)

2 Mathematical Formulations

1. We use the following notations :

(i) P_R stands by ...

(ii) T_R stands by ...

(iii) P_L stands by ...

(iv) T_L stands by ...

(v) a stands by ...

(vi) b stands by ...

(vii) c stands by ...

All these functions are positive.

The differential equations for the evolutions of genes :

$$\begin{aligned} \dot{P}_R &= a - cP_R, \\ \dot{T}_R &= a - b(T_R - P_R), \end{aligned} \tag{2.1}$$

where $\dot{P}_R$ and $\dot{T}_R$ are the derivatives of the two functions $P_R(t)$ and $T_R(t)$ with respect to t and the coefficients a, b and c are also functions of t.

2. The formulations for labeled genes from the labeling time t^*.

Since the labeling time is very small with respect to the evolution of genes, the coefficient $b = 0$ for labeling genes and $M_L(t^*) = 0$ at the labeling time. Hence solving (2.1), we obtain the formulation of the labeling genes:

$$\begin{aligned} P_L(t) &= \frac{a_t}{c_t}(1 - e^{-c_t t_L}), \\ T_L(t) &= a_t t_L. \end{aligned} \tag{2.2}$$

3 Fundamental Assumptions

In [1], it is assumed that

(i) $a(t)$ is a ... in each time-periodic.

(ii) $b(t)$ is a ... in each time-periodic.

(iii) $c(t)$ is a ... in each time-periodic.

Hence solving (2.1) and (2.2), we obtain the formulations in [1] as follows :

$$\begin{aligned} a &= ... \\ b &= ... \\ c &= ... \end{aligned}$$

While the problem is :
for two almost same genes, the correlations of a, b and c are very small.

4 Improvement of Models

To overcome this problem, we assume that the coefficients a, b and c are continuous functions and the variations of a, b and c are very small, i.e.,

$$\int_{t_0}^{t} |\dot{a}(s)|^2 + |\dot{b}(s)|^2 + |\dot{c}(s)|^2 \mathrm{d}s$$

is small as possible. The reason is that the ... of each gene does not change very fast.

5 Improvement of Mathematical Formulations

Let $(P_R(t_0), T_R(t_0), P_L(t_0), T_L(t_0))$ and $(P_R(t_1), T_R(t_1), P_L(t_1), T_L(t_1))$ be obtained at two time points and t_L be the given labeling-time.

(i) We insert several time points in $[t_0, t_1]$ labeled by $s_i = t_0 + \frac{i(t_1 - t_0)}{N}$, where N is an integer and $i = 0, \ldots, N$. We label the approximation values of the functions a, b and c at each time periodic $[s_i, s_{i+1}]$ by a^i, b^i and c^i, $i = 0, \ldots, N-1$ and the approximations values of the functions a, b and c at the time point $s_N = t_1$ by a^N, b^N and c^N.

(ii) With $(P_L(t_0), T_L(t_0))$, $(P_L(t_1), T_L(t_1))$ and t_L, we obtain from (2.2) that

$$\begin{aligned} a^0 &= T_L(t_0)/t_L \\ a^N &= T_L(t_1)/t_L \\ \frac{c^0}{1 - e^{-c^0 t_L}} &= \frac{T_L(t_0)}{P_L(t_0) t_L} \\ \frac{c^N}{1 - e^{-c^N t_L}} &= \frac{T_L(t_1)}{P_L(t_1) t_L} \end{aligned}$$

(iii) With the approximation coefficients a^i, b^i and c^i, we solve (2.1) in $[s_i, s_{i+1}]$ by the given initial value (P_R^i, T_R^i) to obtain the approximation value (P_R^{i+1}, T_R^{i+1}) for $i = 0, \dots, N-1$,

$$\begin{pmatrix} P_R^{i+1} \\ T_R^{i+1} \end{pmatrix} = \begin{pmatrix} P_R^{i} \\ T_R^{i} \end{pmatrix} + h \begin{pmatrix} a^i - c^i P_R^i \\ a^i - b^i (T_R^i - P_R^i). \end{pmatrix}$$

Here $h = (t_1 - t_0)/N$ and $(P_R^0, T_R^0) = (P_R(t_0), T_R(t_0))$.

After this process, we obtain an approximation value (P_R^N, T_R^N) to the final value $(P_R(t_1), T_R(t_1))$.

(iv) The changing rate of the coefficients in the time periodic $[s_i, s_{i+1}]$ are approximated by

$$|\dot{a}(s)|^2 \approx \left(\frac{a^{i+1} - a^i}{h}\right)^2,$$
$$|\dot{b}(s)|^2 \approx \left(\frac{b^{i+1} - b^i}{h}\right)^2,$$
$$|\dot{c}(s)|^2 \approx \left(\frac{c^{i+1} - c^i}{h}\right)^2,$$

for $i = 0, \dots, N-1$. Hence the total change of the coefficients are approximated by

$$J(a, b, c) = \frac{1}{2} \sum_{i=0}^{N-1} \frac{(a^{i+1} - a^i)^2}{h} + \frac{(b^{i+1} - b^i)^2}{h} + \frac{(c^{i+1} - c^i)^2}{h},$$

which has an equivalent quadrature form

$$J(X) = \frac{1}{2} X^T A X + B^T X + C,$$

where $X = (a^1, \cdots, a^{N-1}, b^0, \cdots, b^N, c^1, \cdots, c^{N-1})^T \in \mathbb{R}^{3N-1}$ and the matrix $A \in \mathbb{R}^{(3N-1)\times(3N-1)}$, the vector $B \in \mathbb{R}^{3N-1}$ and $C \in \mathbb{R}$ are given by

$$A = \frac{N}{2(t_1 - t_0)} \begin{pmatrix} \bar{I}_{N-1} & & \\ & I_{N+1} & \\ & & \bar{I}_{N-1} \end{pmatrix}_{(3N-1)\times(3N-1)},$$
$$B = \frac{N}{t_1 - t_0} \begin{pmatrix} -a^0, & 0_{(N-3)}, & -a^N, & 0_{(N+1)}, & -c^0, & 0_{(N-3)}, & -c^N \end{pmatrix}^T,$$
$$C = \frac{N}{2(t_1 - t_0)} ((a^0)^2 + (a^N)^2 + (c^0)^2 + (c^N)^2),$$

where

$$I_N = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ 0 & \cdots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}_{N\times N},$$

and

$$\bar{I}_N = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ 0 & \cdots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}_{N\times N}.$$

(v) The coefficients a, b and c are solved by the following optimal problem with positive and nonlinear constraints, i.e.,

$$\begin{aligned} X = &\arg\min J(X) \\ \text{s.t. } & X > 0, \\ & P_R^N(X) = P_R(t_1), \\ & T_R^N(X) = T_R(t_1). \end{aligned}$$

6 Improvement of Solving Programs

(i) Solving the differential equations with a higher-order method such as Runge's method.

$$\begin{aligned}
\begin{pmatrix} K_1^i \\ L_1^i \end{pmatrix} &= \begin{pmatrix} a^i - c^i P_R^i \\ a^i - b^i(T_R^i - P_R^i) \end{pmatrix}, \\
\begin{pmatrix} K_2^i \\ L_2^i \end{pmatrix} &= \begin{pmatrix} a^i - c^i(P_R^i + h/2K_1^i) \\ a^i - b^i((T_R^i + h/2L_1^i) - (P_R^i + h/2K_1^i)) \end{pmatrix}, \\
\begin{pmatrix} K_3^i \\ L_3^i \end{pmatrix} &= \begin{pmatrix} a^i - c^i(P_R^i + h/2K_2^i) \\ a^i - b^i((T_R^i + h/2L_2^i) - (P_R^i + h/2K_2^i)) \end{pmatrix}, \\
\begin{pmatrix} K_4^i \\ L_4^i \end{pmatrix} &= \begin{pmatrix} a^i - c^i(P_R^i + hK_3^i) \\ a^i - b^i((T_R^i + hL_3^i) - (P_R^i + hK_3^i)) \end{pmatrix}, \\
\begin{pmatrix} P_R^{i+1} \\ T_R^{i+1} \end{pmatrix} &= \begin{pmatrix} P_R^i \\ T_R^i \end{pmatrix} + \frac{h}{6}\left(\begin{pmatrix} K_1^i \\ L_1^i \end{pmatrix} + 2\begin{pmatrix} K_2^i \\ L_2^i \end{pmatrix} + 2\begin{pmatrix} K_3^i \\ L_3^i \end{pmatrix} + \begin{pmatrix} K_4^i \\ L_4^i \end{pmatrix}\right).
\end{aligned}$$

(ii) By introducing a penalty factor, the optimal problem is transformed into

$$\begin{aligned} &\min J(X) + 0.5\alpha_1(P_R^N(X) - P_R(t_1))^2 + 0.5\alpha_2(T_R^N(X) - T_R(t_1))^2 \\ &\text{s.t. } X > 0, \end{aligned}$$

where α_1, α_2 are positive large numbers. Then the gradient descent method is given by

$$\begin{aligned} X_{k+1} = X_k - \rho_k(\frac{\partial J}{\partial X}(X_k) &+ \alpha_1(P_R^N(X_k) - P_R(t_1))\frac{\partial P_R^N}{\partial X}(X_k) \\ &+ \alpha_2(T_R^N(X_k) - T_R(t_1))\frac{\partial T_R^N}{\partial X}(X_k)) \end{aligned}$$

where ρ_k is a small number such that $X_{k+1} > 0$ piecewise component. Of course

$$\frac{\partial J}{\partial X}(X_k) = AX_k + B.$$

Hence in the following, we present an approach for the approximations to $\frac{\partial P_R^N}{\partial X}(X_k)$ and $\frac{\partial T_R^N}{\partial X}(X_k)$.

(iii) Following from (2.1), in each $[s_i, s_{i+1}]$, one obtains that

$$
\begin{aligned}
\frac{d}{dt}\left(\frac{\partial P_R}{\partial X}\right)_{3N-1} &= (\eta_i)_{3N-1} - P_R(\xi_i)_{3N-1} - c_i\left(\frac{\partial P_R}{\partial X}\right)_{3N-1} \\
&=: f_i\left(P_R, \left(\frac{\partial P_R}{\partial X}\right)_{3N-1}\right), \\
\frac{d}{dt}\left(\frac{\partial T_R}{\partial X}\right)_{3N-1} &= (\eta_i)_{3N-1} - (T_R - P_R)(\theta_i)_{3N-1} \\
&\quad - b_i\left(\left(\frac{\partial T_R}{\partial X}\right)_{3N-1} - \left(\frac{\partial P_R}{\partial X}\right)_{3N-1}\right) \\
&=: g_i\left(T_R - P_R, \left(\frac{\partial P_R}{\partial X}\right)_{3N-1}, \left(\frac{\partial T_R}{\partial X}\right)_{3N-1}\right),
\end{aligned}
\tag{6.1}
$$

where $\frac{\partial P_R}{\partial X}, \frac{\partial T_R}{\partial X}, \eta_i, \xi_i, \theta_i \in \mathbb{R}^{3N-1}$ are given by

$$
\begin{aligned}
&\eta_0 = 0, \\
&\eta_i = (e_i, 0_{N+1}, 0_{N-1})^T, i = 1, \cdots, N-1, \\
&\xi_0 = 0, \\
&\xi_i = (0_{N-1}, 0_{N+1}, e_i)^T, i = 1, \cdots, N-1, \\
&\theta_i = (0_{N-1}, \bar{e}_{i+1}, 0_{N-1})^T, i = 0, \cdots, N-1.
\end{aligned}
$$

Here $\frac{\partial P_R}{\partial X}(0) = 0$, $\frac{\partial T_R}{\partial X}(0) = 0$, $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T \in \mathbb{R}^{N-1}$, $i = 1, \cdots, N-1$, and $\bar{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T \in \mathbb{R}^{N+1}$, $i = 1, \cdots, N+1$.

(iv) Runge's method for (6.1) from s_i with $G_i^P \approx \frac{\partial P_R}{\partial X}(s_i)$ and $G_i^T \approx \frac{\partial T_R}{\partial X}(s_i)$ to s_{i+1} with $G_{i+1}^P \approx \frac{\partial P_R}{\partial X}(s_{i+1})$ and $G_{i+1}^T \approx \frac{\partial T_R}{\partial X}(s_{i+1})$ is given by :

$$
\begin{aligned}
&\begin{pmatrix} D_1^i \\ E_1^i \end{pmatrix} = \begin{pmatrix} f_i(P_R^i, G_i^P) \\ g_i(T_R^i - P_R^i, G_i^P, G_i^T) \end{pmatrix}, \\
&\begin{pmatrix} D_2^i \\ E_2^i \end{pmatrix} = \begin{pmatrix} f_i(P_R^i + h/2K_1^i, G_i^P + h/2D_1^i) \\ g_i((T_R^i + h/2L_1^i) - (P_R^i + h/2K_1^i),, G_i^P + h/2D_1^i, G_i^T + h/2E_1^i) \end{pmatrix}, \\
&\begin{pmatrix} D_3^i \\ E_3^i \end{pmatrix} = \begin{pmatrix} f_i(P_R^i + h/2K_2^i, G_i^P + h/2D_2^i) \\ g_i((T_R^i + h/2L_2^i) - (P_R^i + h/2K_2^i),, G_i^P + h/2D_2^i, G_i^T + h/2E_2^i) \end{pmatrix}, \\
&\begin{pmatrix} D_4^i \\ E_4^i \end{pmatrix} = \begin{pmatrix} f_i(P_R^i + hK_3^i, G_i^P + hD_3^i) \\ g_i((T_R^i + hL_3^i) - (P_R^i + hK_3^i),, G_i^P + hD_3^i, G_i^T + hE_3^i) \end{pmatrix}, \\
&\begin{pmatrix} G_{i+1}^P \\ G_{i+1}^T \end{pmatrix} = \begin{pmatrix} G_i^P \\ G_i^T \end{pmatrix} + \frac{h}{6}\left(\begin{pmatrix} D_1^i \\ E_1^i \end{pmatrix} + 2\begin{pmatrix} D_2^i \\ E_2^i \end{pmatrix} + 2\begin{pmatrix} D_3^i \\ E_3^i \end{pmatrix} + \begin{pmatrix} D_4^i \\ E_4^i \end{pmatrix}\right).
\end{aligned}
$$

(v) Implementation.

(a) Given initial values $(P_R(t_0), T_R(t_0), P_L(t_0), T_L(t_0))$, $(P_R(t_1), T_R(t_1), P_L(t_1), T_L(t_1))$ and t_L.

(b) Given an integer $N \geqslant 1$, $\alpha_1, \alpha_2 \gg 1$, $\rho \in (0,1)$ and $\epsilon \ll 1$.

(c) Calculate a^0, a^N, c^0, c^N.

(d) Set $k = 0$ and give an initial guess X_0 for the other coefficients.

(e) Calculate P_R^N and T_R^N by Runge's method for (2.1).

(f) Calculate G_N^P and G_N^T by Runge's method for (6.1)

(g) Calculate the gradient

$$g_k = \frac{\partial J}{\partial X}(X_k) + \alpha_1 (P_R^N(X_k) - P_R(t_1)) \frac{\partial P_R^N}{\partial X}(X_k) + \alpha_2 (T_R^N(X_k) - T_R(t_1)) \frac{\partial T_R^N}{\partial X}(X_k)$$

(h) If $\max |g_k| < \epsilon$, the program is stopped. Otherwise update the guessing value by

$$X_{k+1} = X_k - \rho \frac{\min\{X_k\}}{\max |g_k|} g_k,$$

set $k = k + 1$ and go to Step (e).

References

[1] INSPECT