

Méthodes informatiques pour la biologie systémique et synthétique

Examen partiel, Cours C2-19 MPRI,

16 novembre 2009

N.B. Cette partie de l'examen est à rendre sur une copie séparée.

La dynamique des processus biochimiques se décrit traditionnellement en biologie mathématique par des équations différentielles ordinaires (EDO). Les modèles à base de règles de réactions cinétiques contiennent quant à eux plus d'information concernant notamment la structure des interactions. A un modèle de n règles de réactions sur v variables x_1, \dots, x_v ,

$$\{e_i \text{ for } l_i \Rightarrow r_i\}_{i=1, \dots, n}$$

où l_i et r_i sont des multi-ensembles de variables supposés différents, $l_i \neq r_i$, et e_i est une expression arithmétique quelconque sur les variables (non supposée positive à ce stade), on associe l'EDO, pour $1 \leq j \leq v$

$$\frac{dx_j}{dt} = \sum_{i=1}^n (r_i(x_j) - l_i(x_j)) * e_i \quad (1)$$

où $r_i(x_k)$ (resp. $l_i(x_j)$) est la multiplicité (coefficient stochiométrique) de x_j dans l_i (resp. r_i).

Dans cette partie, nous nous intéressons au problème inverse de décompilation d'une EDO en un ensemble de règles de réactions, et aux conditions d'unicité du modèle réactionnel. Dans la première sous-partie, on traite d'un problème d'analyse dimensionnelle par inférence de types. Dans la seconde sous-partie, on étudie les conditions d'unicité du processus de décompilation. Ces deux sous-parties sont totalement indépendantes.

1 Inférence de types (8 points)

L'analyse dimensionnelle permet de vérifier la cohérence d'une équation par rapport aux dimensions physiques des formules: temps T , quantité de matière N , longueur L , masse M , température θ , intensité électrique I , intensité lumineuse J . Une dimension est ainsi un vecteur de 7 nombres entiers relatifs. Par exemple $N^1 * L^{-3}$ est la dimension d'une concentration, exprimée en nombre de moles (quantité de matière N) par litre (volume L^3). L'addition, la soustraction et la comparaison de grandeurs ne sont possibles qu'entre des expressions de même dimension.

Dans une EDO paramétrée portant sur des variables de concentration, on peut chercher à inférer la dimension des paramètres, de façon à détecter d'éventuelles erreurs. Pour simplifier, on ne considèrera que la dimension temps. Sous cette hypothèse simplificatrice, les concentrations x_j sont donc sans dimension (T^0), et les membres des équations différentielles $\frac{dx_j}{dt}$ sont en dimension T^{-1} .

1.1

Dans un article publié présentant un modèle du cycle cellulaire, une typo s'est glissée dans les équations suivantes:

$$\frac{d[\text{cycE}]}{dt} = \varepsilon(k_7' + k_7[\text{E2F}_A]) - V_8[\text{cycE}] - k_{25}[\text{CycE}][\text{Kip1}] + k_{25r}[\text{CycE} : \text{Kip1}] + V_6[\text{CycE} : \text{Kip1}]$$

$$\frac{d[\text{cycE} : \text{Kip1}]}{dt} = k_{25}[\text{CycE}][\text{Kip1}] - k_{25r}[\text{CycE} : \text{Kip1}] - V_6[\text{CycE} : \text{Kip1}] - V_8[\text{CycE} : \text{Kip1}]$$

$$\frac{d[\text{cycA}]}{dt} = \varepsilon k_{29}[\text{E2FA}][\text{mass}] - k_{30}[\text{Cdc20}][\text{cycA}] - k_{25}[\text{CycA}][\text{Kip1}] + k_{25r}[\text{CycA} : \text{Kip1}] + V_6[\text{CycA} : \text{Kip1}]$$

...

$$V_6 = k_6' + k_6(\eta_E[\text{CycE}] + \eta_A[\text{CycA}] + \eta_B[\text{CycB}])$$

$$V_8 = k_8' \frac{k_8(\psi_E([\text{CycE}] + [\text{CycA}]) + \psi_B[\text{CycB}])}{J_8 + [\text{cycE}_T]}$$

...

Rate constants (h^{-1})

$k_1' = 0.1, k_1 = 0.6, k_2' = 0.05, k_2 = 20, k_2'' = 1, k_3' = 7.5, k_3 = 140, k_4 = 40, k_5 = 20, k_6' = 10, k_6 = 100, k_7' = 0, k_7 = 0.6, k_8' = 0.1, k_8 = 2, k_9 = 2.5, k_{10} = 5, k_{11}' = 0, k_{11} = 1.5, k_{12} = 1.5, k_{13} = 5, k_{14} = 2.5, k_{15} = 0.25, k_{16} = 0.25, k_{17}' = 0.35, k_{17} = 10, k_{18} = 10, k_{19}' = 0, k_{19} = 20, k_{20} = 10, k_{22} = 1, k_{23}' = 0.005, k_{23} = 1, k_{24} = 1000, k_{24r} = 10, k_{25} = 1000, k_{25r} = 10, k_{26} = 10.000, k_{26r} = 200, k_{27} = 0.2, k_{28} = 0.2, k_{29} = 0.05, k_{30} = 20, k_{31} = 0.7, k_{32} = 1.8, k_{33} = 0.05, k_{34} = 0.05, \mu = 0.061$

Dimensionless constants

$J_1 = 0.1, J_3 = J_4 = 0.01, J_8 = 0.1, J_{13} = 0.005, J_{14} = 0.005, J_{15} = 0.1, J_{17} = 0.3, J_{31} = J_{32} = 0.01, K_{21} = 1, [\text{E2F}_T] = 5, [\text{PP1}_T] = 1, [\text{Rb}_T] = 10, \phi_E = 25, \phi_B = 2, \gamma_A = 0.3, \gamma_B = 1, \eta_E = 0.5, \eta_A = 0.5, \eta_B = 1, \lambda_D = 3.3, \lambda_E = 5, \lambda_A = 3, \lambda_B = 5, \psi_E = 1, \psi_B = 0.05, \varepsilon = 1$

...

Déterminer l'erreur par analyse dimensionnelle et proposer une correction de l'équation.

Réponse: V_8 a une dimension T^{-2} incompatible avec son utilisation dans les équations différentielles. Une correction possible est $V_8 = k_8' + \frac{k_8 \dots}{J_8 \dots}$.

1.2

On suppose l'ensemble des expressions arithmétiques \mathcal{E} défini par la grammaire suivante:

$$e ::= p \mid x \mid e + e \mid e - e \mid e * e \mid e / e \mid e \hat{\ } n$$

où n représente un nombre entier relatif sans dimension, p un paramètre à valeur réelle, et x une variable de concentration.

Définir formellement, par induction structurale, la fonction partielle de typage

$$[] : \mathcal{E} \rightarrow \mathbb{Z}$$

qui associe à toute expression arithmétique sa dimension en temps lorsqu'elle existe, supposant données les dimensions $[p]$ des paramètres.

Réponse: On a $[p]$ qui est donné pour chaque paramètre, $[x] = 0$ pour chaque variable, $[e_1 + - e_2] = [e_1]$ si $[e_1] = [e_2]$, indéfini sinon, $[e_1 * e_2] = [e_1] + [e_2]$, $[e_1 / e_2] = [e_1] - [e_2]$, $[e \hat{\ } n] = n * [e]$

1.3

Donner un algorithme d'inférence de la dimension des paramètres dans un système d'EDO sur des variables de concentration

$$\frac{dx_j}{dt} = f_j(p_1, \dots, p_k, x_1, \dots, x_v)$$

$1 \leq j \leq v$.

Réponse: En posant $[f_j(p_1, \dots, p_k, x_1, \dots, x_v)] = -1$ pour $1 \leq j \leq v$, et en décomposant les expressions suivant la définition précédente, on se ramène à un système d'équations linéaires sur les dimensions des paramètres $[p_1], \dots, [p_k]$.

En résolvant ce système par éliminations successives, le résultat est soit une erreur s'il n'y a pas de solution (équation $i = i'$ avec $i \neq i'$), soit la dimension des paramètres s'il existe une unique solution (système d'équations de la forme $[p_1] = i_1, \dots, [p_k] = i_k$), soit indéfini en présence d'équations non simplifiables comme $[p] + [p'] = i$.

2 Inférence de réactions (12 points)

Nous abordons maintenant les problèmes d'existence, unicité et inférence d'un modèle réactionnel correspondant à une EDO donnée définissant $\frac{dx_j}{dt}$ pour $1 \leq j \leq v$.

2.1

Sans aucune contrainte sur les expressions cinétiques, associer un modèle de réactions à toute EDO de la forme $\frac{dx_j}{dt} = f_j$ pour $1 \leq j \leq v$.

Réponse: $\{f_j \text{ for } - \Rightarrow x_j\}_{j=1, \dots, v}$

2.2

En se restreignant à des expressions cinétiques positives, ne portant que sur les valeurs de concentration des réactants, donner deux modèles de réactions différents correspondant à l'EDO

$$\frac{dA}{dt} = -\frac{dB}{dt} = -k * A$$

où k est un paramètre strictement positif.

Réponse: $\{k*A \text{ for } A \Rightarrow B\}$ et $\{k*A \text{ for } A+A \Rightarrow A+B\}$

2.3

En se restreignant maintenant à des cinétiques positives de loi d'action de masse, donner deux modèles de réactions différents correspondant à l'EDO

$$\frac{dA}{dt} = -\frac{dB}{dt} = -2 * k * A^2$$

Réponse: $\{k*A^2 \text{ for } A+A \Rightarrow B+B\}$ et $\{2*k*A^2 \text{ for } A = [A] \Rightarrow B\}$

2.4

Dans la suite, nous nous restreignons à des modèles de règles de réactions,

$$\{e_i \text{ for } l_i \Rightarrow r_i\}_{i=1, \dots, n}$$

dotées de cinétiques de loi d'action de masse paramétrées de la forme

$$e_i = k_i * \prod_{l=1}^v x_l^{l_i(x_i)} \quad (2)$$

où les k_i sont des paramètres formels à valeurs réelles strictement positives.
 Nous supposons que l'EDO est polynômiale, de la forme pour $1 \leq j \leq v$,

$$\frac{dx_j}{dt} = \sum_{k=1}^{n_j} s_{jk} * p_{jk} * \prod_{l=1}^v x_l^{r_{jkl}} \quad (3)$$

où $n_j \geq 1$, $s_{jk} \in \mathbb{Z}$, p_{jk} est un paramètre formel à valeur réelle positive et $r_{jkl} \in \mathbb{N}$.

Montrer que si l'on ne suppose pas de plus que les paramètres doivent être tous différents dans les règles de réactions ayant même ensemble de réactants, i.e. pour tout $1 \leq i < j \leq n$,

$$k_i = k_j \Rightarrow l_i \neq l_j \quad (4)$$

il peut encore exister plusieurs modèles de réactions vérifiant la condition (2) et produisant l'équation différentielle

$$\frac{dA}{dt} = -\frac{dB}{dt} = -k * A$$

Réponse: {k*A for A => B} et {k*A for A => -, k*A for A => A+B}

2.5

Montrer que pour chaque terme $s_{jk} * p_{jk} * \prod_{l=1}^v x_l^{r_{jkl}}$ de l'EDO (3), on peut former une règle de réaction de paramètre p_{jk} satisfaisant aux conditions (2)(4).

Réponse: $p_{jk} * \prod_{l=1}^v x_l^{r_{jkl}}$ for $\sum_{l=1}^v r_{jkl} * x_l \Rightarrow s_{jk} * x_j$

2.6

Montrer que sous les conditions (2)(4), chaque expression cinétique e_i apparaît au moins une fois dans l'EDO.

Réponse: Les termes des expressions cinétiques e_i ne peuvent pas s'annuler dans les EDO polynomiale de la forme (3) sous la condition (4).

2.7

En déduire que pour toute EDO définissant $\frac{dx_i}{dt}$ pour $1 \leq i \leq v$, il existe au plus un modèle de réactions satisfaisant aux conditions (2)(4) qui le produit.

Réponse: Les questions précédentes montrent l'unicité de e_i et l_i dans les règles. Les membres droits s'obtiennent en posant $r_i = l_i + s$, ce qui montre l'unicité du modèle hormis les cas où r_i contient une multiplicité strictement négative, pour lesquels il n'existe pas de modèle de réaction satisfaisant les conditions (2)(4).