Journal of Theoretical Biology 343 (2014) 199–207

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Parameter estimation for metabolic networks with two stage Bregman regularization homotopy inversion algorithm

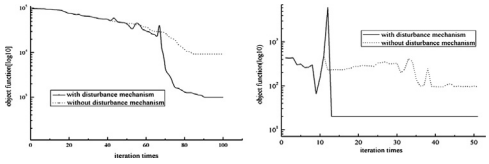
Hong Wang [a,c], Xi-cheng Wang [a,b,*]

[a] School of Computer Science and Technology, Dalian University of Technology, Dalian, China
[b] State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, China
[c] Department of Computer Science, Dalian Neusoft Institute of Technology, Dalian, China

HIGHLIGHTS

- A two-stage variable factors Bregman regularization homotopy method is proposed to identify the parameters of metabolic network.
- A disturbance mechanism for escaping the local optimum is introduced to close the global optimization solution.
- Three metabolic network inverse problems are investigated; the results show that our method performs better than several popular methods.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 May 2013
Received in revised form
11 September 2013
Accepted 13 September 2013
Available online 20 September 2013

Keywords:
Strong nonlinear
Large scale
Ordinary differential equation

ABSTRACT

Metabolism is a very important cellular process and its malfunction contributes to human disease. Therefore, building dynamic models for metabolic networks with experimental data in order to analyze biological process rationally has attracted a lot of attention. Owing to the technical limitations, some unknown parameters contained in models need to be estimated effectively by means of the computational method. Generally, problems of parameter estimation of nonlinear biological network are known to be ill condition and multimodal. In particular, with the increasing amount and enlarging the scope of parameters, many optimization algorithms often fail to find a global solution. In this paper, two-stage variable factor Bregman regularization homotopy method is proposed. Discrete homotopy is used to identify the possible extreme region and continuous homotopy is executed for the purpose of stability of path tracing in the special region. Meanwhile, Latin hypercube sampling is introduced to get the good initial guess value and a perturbation strategy is developed to jump out of the local optimum. Three metabolic network inverse problems are investigated to demonstrate the effectiveness of the proposed method.

1. Introduction

Biological processes are often represented in the form of networks such as protein–protein interaction networks (Wei et al., 2004), regulatory networks (Yang et al., 2008), signaling pathways (Nassiri et al., 2012) or metabolic pathways (Jeong et al., 2001).The aim of studying biological networks is to understand the structure and function of complex biological systems by combining experimental data with mathematical modeling and advanced computational techniques. Metabolism, as a main function in cell, is involved in many disease controls. To build the dynamic models for the metabolic networks with experimental data has attracted many researchers' attention recently.

Mathematical models are an invaluable tool for analysis, design, optimization and control of the biological systems under consideration. In the context of metabolic network, MM (Michaelis–Menten) model or power law model is frequently adopted to quantitatively describe the dynamic properties of biological reactions. Owing to the technical limitation, some kinetic parameters

* Corresponding author at: Department of Computer Science and Technology, DaLian University of Technology, DaLian, Liao Ning 116023, China.
Tel.: +86 0411 84706223.
E-mail address: forget1life@gmail.com (X.-c. Wang).

0022-5193/$ - see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.jtbi.2013.09.020

are very difficult, expensive, time consuming or impossible to measure directly (Schwacke and Voit, 2005; Liu and Wang, 2009). For these reasons, it is a key issue to estimate the unknown parameters from measurements of other quantities for biochemical modeling. Parameter estimation for mathematical models in biology networks presents four characteristics, which make this problem very difficult to solve: the models are highly nonlinear, there are a large number of parameters to be estimated, the estimated parameters vary in a large range and the information content of the available experimental data is frequently scarce (Villaverde et al., 2012).

In past years, many researchers give a lot of attention to this problem in the system biology. Various optimization methods are developed to evaluate the parameters of biological network, such as evolutionary algorithms (Tsai and Wang, 2005; Saha et al., 2008; Auliac et al., 2008; Ho et al., 2007), simulated annealing (Wei et al., 2004), hybrid or cooperation optimization algorithm (Liu and Wang, 2009; Villaverde et al., 2012; Balsa-Canto et al., 2008, 2005; Liu and Wang, 2010). In view of the large range of parameters, some researchers are inclined to narrow the parameter search space (Tucker et al., 2007; Tucker and Moulton, 2006; Kutalik et al., 2007) or use Kalman filtering (Lillacci and Khammash, 2010). In order to decrease the parameter numbers estimated simultaneously, sensitivity analysis is imported to simple models (Ashyraliyev et al., 2009; Wu et al., 2008). However, main problems associated with these optimization methods are that they tend to be computationally expensive and may trap into local optimum if the nonlinearity is strengthened.

HAM (Homotopy analysis method), initially proposed by Liao (2009), is a powerful method to solve non-linear problems. In recent years, this method has been successfully used in several areas of mathematics, including optimization (Thomas et al., 2012) and nonlinear root finding (Abbasbandy, 2006). Furthermore, there are approaches derived from homotopy methods that can find global optimum in situations where other deterministic methods cannot (Thomas et al., 2012; Kuno and Seader, 1988). Essentially, parameter identification is often considered as an inverse problem (Baker et al., 2010) and biological network parameter estimation is thus a nonlinear inverse problem. However, a nonlinear inverse problem is often ill posed or ill conditioned (Eng and Ugler, 2005), in particular unstable with respect to the data noise (Liao, 2009). Iterative regularization methods are an effective method to overcome the instability of solution in nonlinear inverse problems, at the same time, the specific regularization term has the result of noise suppression (Cui et al., 2005).

In this study, a two stage Bregman regularization homotopy method is introduced as a novel strategy into biological network parameter estimation to find the global solution. For saving homotopy tracing time, the algorithm gives full play to complementary advantages of continuous and discrete homotopy, and imports variable regularly factor to reflect the forecast value and the experimental value approaching degree. At last, the algorithm is tested on three metabolic networks to show the effectiveness.

2. Method

2.1. Parameter estimation

Generally, metabolic networks are modeled using ODEs (ordinary differential equations) to describe the kinetics of an enzymatic reaction (Kutalik et al., 2005):

$$dX/dt = f(x,p) \tag{1}$$

where x, p and f denote the metabolite concentration, the parameter set to be estimated, and the reaction rate function, respectively. The reaction rate function f is nonlinear and complex with p. Integrating both sides of Eq. (1) gives

$$X(p) = \int f(x,p)dt \tag{2}$$

where $X(p)$ may only have numerical solutions. Parameter estimation often views as an optimization problem that aims to find the parameters which give the best fit to a set of experimental data by minimizing an error function (based on the discrepancy between the observed data and the simulated model). So the optimal object function for parameter estimation can be written as:

$$\begin{aligned} &\min J(p) = \sum_{j=1}^{N}\sum_{i=1}^{n} ||X_{ei}(t_j) - X_i(t_j,p)||^2 / X_{ei\ \max}^2 \\ &\text{Subject to : } dX_i/dt = f_i(t,p) \\ &\qquad p^L < p < p^U \end{aligned} \tag{3}$$

where $X_{ei}(t_j)$ is the experimental data for the ith component at $t=t_j$, $X_i(t_j,p)$is the computed concentration for the ith component at $t=t_j$, n is the number of ODEs and N is number of sampling data. p^L and p^U are the lower and upper bound constraints on the parameter p. $t_i \in T \subset [0,T]$ describes the time associated with the experimental data.

2.2. Bregman regularization method

For highly nonlinear inverse problem, a small change in measurement can lead to an enormous change in the estimated model (Engl et al., 1996) and the data noises make it worse. The regularization method is usually required for overcoming the ill-conditioned, and Tikhonov's regularization term is one of the most widely used regularization methods (Silva Neto and Cella, 2006). In order to achieve the purpose of noise suppression simultaneously, Bregman distance as an option to Tikhonov has been applied to regularization algorithm (Tan et al., 2006; Osher et al., 2005). There are three popular types of Bregman distance functions:

$$B_1(x,y) = ||x-y||^2 \tag{4}$$

$$B_2(x,y) = \sum_{i=1}^{n} [x^i \log(x^i/y^i) - x^i + y^i] \tag{5}$$

$$B_3(x,y) = \sum_{i=1}^{n} [-\log(x^i/y^i) + x^i/y^i - 1] \tag{6}$$

Inspired by reference (Tan et al., 2006), using the Bregman distance B_3 to modify the object function Eq. (3) gives:

$$\min J(p) = \sum_{j=1}^{N}\sum_{i=1}^{n} ||X_{ei}(t_j) - X_i(t_j,p)||^2 / X_{ei\ \max}^2 + \alpha B_3(p,p^*) \tag{7}$$

where α is the regularization coefficient and p^* is a priori estimate for p. After adding the regularization term, the objective function critically depends on α. When α is close to zero, Eq. (7) reverts to Eq. (3). As α approaches infinity, p^* becomes the solution of Eq. (3).

3. Homotopy method

It is difficult to obtain analytical solutions for nonlinear ODEs arising from physical systems (Odibat, 2010). Although perturbation quantity techniques have been proposed to help us understand nonlinear phenomena, they are invalid for strong nonlinear problems. Since Liao (Liao, 2009) proposed the homotopy analysis method, many types of nonlinear problems have been solved by homotopy (Molabahrami and Khani, 2009), especially nonlinear inverse problems (Cao and Han, 2011). Homotopy methods can reach a solution by tracing a path from a fairly arbitrary initial point (Kuno and Seader, 1988). More importantly, homotopy