

细胞信号转导网络输入信号的最优实验设计

贾建芳^{1,2}, 刘太元¹, 岳红¹, 王宏¹

(1. 中国科学院自动化研究所, 北京, 100080; 2. 中北大学自动化系, 山西, 太原, 030051)

摘要: 细胞信号转导网络的结构复杂, 规模庞大, 具有高度非线性, 数学模型的参数和变量数目很多, 有效估计未知参数是系统生物学主要的研究问题之一。提出以 Fisher 信息矩阵和估计参数的协方差矩阵为基础, 利用不同的优化实验设计方法, 研究 NF- κ B 信号转导网络输入信号 IKK 初浓度的最优设计问题, 分析模型参数的置信区间, 为参数估计设计优化实验。仿真结果表明, 采用 E-最优实验设计准则, 获得输入信号 IKK 初浓度为 0.06 μ M, 能够提供较多的实验信息, 提高参数估计的精度。

关键词: 最优实验设计; Fisher 信息矩阵; 细胞信号转导网络; 参数估计

中图分类号: Q 332

文献标识码: A

文章编号: 1001-4160(2007)08-1033-10

Optimal experimental design for input signal of cell signal transduction networks

Jia Jianfang^{1,2}, Liu Taiyuan¹, Yue Hong¹ and Wang Hong¹

(1. Institute of Automation, Chinese Academy of Sciences, Beijing, 100080, China; 2. Department of Automation, North University of China, Taiyuan, 030051, Shanxi, China)

Abstract: Cell signal transduction networks have complex structure and large size. A mathematical model of cell signal transduction networks is highly nonlinear, and it includes a lot of parameters and variables. How to effectively estimate the unknown parameters is a major problem of systems biology. Based on Fisher information matrix and covariance matrix of estimated parameters, optimal experimental design for the initial concentration of input signal IKK of NF- κ B signaling pathway was studied by means of various criterions. The confidence interval of parameters was analyzed and optimal experiment for parameter estimation was proposed. According to the E-optimal experimental design criterion, the initial concentration of IKK is set to 0.06 μ M. The simulation results demonstrate that more information for parameter estimation is obtained and the accuracy of parameter estimation is improved evidently.

Key words: optimal experimental design, fisher information matrix, cell signal transduction networks, parameter estimation

Jia JF, Liu TY, Yue H and Wang H. Optimal experimental design for input signal of cell signal transduction networks. Computers and Applied Chemistry, 2007, 24(8):1033-1038.

1 引言

细胞信号网络具有很强的非线性, 随着模型复杂程度的增加, 参数的数量更是成倍增长^[1]。模型参数一般通过生化实验获得, 但测量数据的不完整和测量噪声的干扰, 直接妨碍参数辨识。在细胞活体内精确测量内源性蛋白质的浓度往往十分困难。此外, 一些生化反应的动力学参数也难以确定, 即使在试管中利用纯组分通过实验测得了速率常数, 也可能与细胞活体内的动力学参数截然不同^[2]。因此, 通过反应机理和实验分析, 根据不同的研究目的, 如何建立系统的“核心”数学模型, 求得系统的关键组分和关键反应, 有效辨识模型参数, 已经成为系统生物学的研究热点之一^[3]。

实验设计是实验过程的依据, 是实验数据处理的前提, 如何优化实验设计, 提高实验效率, 提取更多的实验信息是人们广泛关注的问题。在做一项实验时, 实验条件诸多限制, 例如: 输入信号的类型、输入和输出的振幅或功率、允许

持续进行实验的时间、最大和最小的采样速度、采样频率等的限制^[4]。输出信号的测量受实验技术和检测仪器的制约, 因此, 当选定一种实验信息的量度准则之后, 对输入信号和采样频率做有约束的最优化计算, 从中得到最优的实验设计。实际应用中, 通过计算机采样一般是等间隔的, 因此最优输入信号的设计问题, 成为系统辨识实验和设计的主要内容。

文献[4]就水处理过程的参数估计, 分析实验设计的自由度和约束条件, 研究不同条件下的最优实验设计, 说明实验的信息量极大依赖于实验设计, 通过改变输入的初始浓度, 延长间歇反应操作的实验, 能够提高参数估计的精度。文献[5]研究应用在系统生物学中的最优实验设计, 设计 MAP-Kinase 信号途径的输入信号, 讨论各种最优实验设计准则的优劣。为了估计信号转导途径模型的参数, 文献[6]讨论参数估计的数值方法, 选择最优采样时间以便从测量数

收稿日期: 2007-02-16; 修回日期: 2007-04-02

基金资助: 中国科学院海外杰出学者基金资助项目 (2004-1-4)

作者简介: 贾建芳 (1973—), 男, 山西人, 博士研究生, 讲师。

据中获得更多信息,降低参数辨识的误差。同时考虑初值、采样时间、时变和时不变变量的值,以及实验时间等约束的基础上,文献[7]设计了基于系统 Fisher 信息矩阵的评价函数,仿真表明在相同条件下,最优实验设计找出更有价值的信息。文献[8]从概率论的角度提出一种新的研究思路,假定辨识参数与概率分布相符,如何设计输入信号使辨识参数以一定的概率落入某一区间,以便获得容易控制的概率。文献[9]提出以 Fisher 信息矩阵和 Hessian 矩阵为基础,分别估计生态系统非线性动态模型参数的置信区间,阐明它们的主要差别在于输出函数的二次导数项。

本文以 TNF- α 诱导 NF- κ B 信号转导网络为研究对象,讨论以 Fisher 信息矩阵为基础的最优实验设计准则,研究输入信号初浓度的最优值和参数的置信区间,通过信号转导网络的计算机模拟为优化实验设计提供参考依据。

2 最优实验设计方法

细胞信号转导网络为非线性动态自治系统,状态方程为:

$$\dot{x}(t, \theta) = f(x(t, \theta), t, \theta), x(t_0, \theta) = x_0 \quad (1)$$

输出测量方程为:

$$y(t, \theta) = Cx(t, \theta) \quad (2)$$

式中, $x \in R^n$ 为系统状态变量向量, $y \in R^m$ 为系统测量输出, $\theta \in R^l$ 为模型参数, $f(\cdot)$ 是描述系统结构的非线性函数, 矩阵 C 选择系统的可观测状态。在初时刻 t_0 , 状态变量 x 的初值为 x_0 。当模型参数 θ 未知, 参数估计的目的是发现一组尽可能使模型预测和实测数据一致的参数值。假设参数的真实值为 θ , 估计值为 $\hat{\theta}$, 则性能指标函数为:

$$J(\hat{\theta}) = \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^N [y_{i,k}^{exp}(t_i, \theta) - y_{i,k}^{mod}(t_i, \hat{\theta})]^2 \quad (3)$$

式中, $y_{i,k}^{exp}(t_i, \theta)$ 表示参数为真实值的实验测量数据, $y_{i,k}^{mod}(t_i, \hat{\theta})$ 表示参数为估计值时通过模型计算得到的数据, $i=1, 2, \dots, m$ 表示可测状态的数目, $k=1, 2, \dots, N$ 表示采样数据点。为了获得参数的估计值, 就是寻找使性能指标函数 $J(\hat{\theta})$ 最小的参数值, 这样参数估计就转化为最优化问题。

由于细胞信号转导网络的参数很多, 单纯采用常规的最优化算法计算量极大, 耗费时间极多。因此, 在参数辨识前预先分析系统模型和系统参数, 研究其敏感性, 引入 Fisher 信息矩阵, 利用优化实验设计准则确定最优输入信号, 为实验过程提供指导, 有针对性地测量实验数据, 节约建模时间和成本。

2.1 Fisher 信息矩阵

考虑系统模型(1), 由参数向量 θ 微小的改变量 $\delta\theta$ 而致 t 时刻观测向量 $y(t)$ 的改变量为

$$\delta y(t) = \frac{\partial y(t)}{\partial \theta} \delta \theta \quad (4)$$

因此, 如果测量噪声服从高斯分布, 则在 $[t_1, t_2, \dots, t_N]$ 时刻观测序列的总改变量定义为:

$$\begin{aligned} & \sum_{i=1}^N [\delta y(t_i)]^T Q^{-1} [\delta y(t_i)] \\ &= \sum_{i=1}^N \delta \theta^T \left[\frac{\partial y(t_i)}{\partial \theta} \right]^T Q^{-1} \left[\frac{\partial y(t_i)}{\partial \theta} \right] \delta \theta \\ &= \delta \theta^T \left\{ \sum_{i=1}^N \left[\frac{\partial y(t_i)}{\partial \theta} \right]^T Q^{-1} \left[\frac{\partial y(t_i)}{\partial \theta} \right] \right\} \delta \theta \\ &= \delta \theta^T [FIM(\theta)] \delta \theta \end{aligned} \quad (5)$$

其中 $FIM(\theta) = \sum_{i=1}^N \left[\frac{\partial y(t_i)}{\partial \theta} \right]^T Q^{-1} \left[\frac{\partial y(t_i)}{\partial \theta} \right]$, 称 Fisher 信息矩阵, 表示在噪声情况下测量模型参数可辨识性的信息度, 反映参数向量变化和与之对应的观测变化之间的关系, 而且它还有一非常重要的性质, 是一个半正定对称矩阵, 其所有特征值非负。令 $S = \frac{\partial y(t)}{\partial \theta}$, 则 Fisher 信息矩阵 $FIM(\theta)$ 可以表示为^[10]

$$FIM(\theta) = S^T Q^{-1} S \quad (6)$$

式中, S 称为绝对敏感性矩阵, $Q = \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & \ddots & \vdots \\ 0 & 0 & Q_m \end{bmatrix}$ 称

为测量协方差矩阵, 其中 $Q_i = \begin{bmatrix} \sigma_{i,1}^2 & 0 & 0 \\ 0 & \ddots & \vdots \\ 0 & 0 & \sigma_{i,N}^2 \end{bmatrix}$, m

表示系统测量状态的个数, N 表示采样的个数, $\sigma_{i,j}$ 表示第 i 个测量状态在第 j 个采样时刻的标准差。

2.2 估计参数的置信区间

在试图估计某参数之前, 必须判断其可否辨识, 实际的可辨识性受制于测量噪声和测量数据, 如果有较大测量误差, 或者只有很少数据, 那么参数估计就不可靠, 称为实际不可辨识。

根据数理统计学的 Cramer-Rao 不等式^[11], 估计参数的协方差矩阵 Σ 能够表示为:

$$\Sigma \geq [FIM(\theta)]^{-1} \quad (7)$$

协方差矩阵 Σ 的对角元素, 给出了估计参数的方差。因此, 当等式成立时, 称参数估计达到了 Cramer-Rao 下界, 它为置信区间的估计和动态实验设计提供了理论依据。当测量噪声服从高斯分布时, 则估计参数 $\hat{\theta}_i$ 的 95% 置信区间可以表示为:

$$[\hat{\theta}_i - 1.96\sigma_{\theta_i}, \hat{\theta}_i + 1.96\sigma_{\theta_i}] \quad (8)$$

其中, σ_{θ_i} 表示第 i 个估计参数的标准差。因此, 在估计参数的 95% 置信区间内, 当参数值为零时, 该参数就是不可辨识的。估计参数的相对误差为:

$$\text{error\%} = \frac{1.96\sigma_{\theta_i}}{\hat{\theta}_i} \quad (9)$$

2.3 最优实验设计准则

建立细胞信号转导网络的数学模型,精确辨识模型中未知参数是系统生物学基本任务之一。然而,信号转导网络的组分多,反应本身所表现出的强非线性,以及无法获得完备的实验数据等原因,都使系统模型的辨识变得困难。实测信息度能够用估计参数的协方差矩阵 Σ 来表示,估计参数的置信区间越小,实验包含的信息量就越多。因此,可以利用 Fisher 信息矩阵 $FIM(\theta)$ 和估计参数的协方差矩阵 Σ ,以参数辨识精度及输出空间误差作为优化依据,研究细胞信号转导网络的优化实验设计。

基于 Fisher 信息矩阵 $FIM(\theta)$ 的实验信息度量方式分别为^[4]:

(1) A-最优设计准则,简称 A-准则:

$$\min |tr(FIM(\theta)^{-1})| \quad (10)$$

(2) D-最优设计准则,简称 D 准则:

$$\max |\det(FIM(\theta))| \quad (11)$$

(3) E-最优设计准则,简称 E-准则:

$$\max |\lambda_{\min}(FIM(\theta))| \quad (12)$$

(4) 修正 E-最优设计准则,简称修正 E-准则:

$$\min \left\{ \frac{\lambda_{\max}(FIM(\theta))}{\lambda_{\min}(FIM(\theta))} \right\} \quad (13)$$

其中, $tr(FIM(\theta)^{-1})$ 表示 Fisher 信息矩阵的逆矩阵的秩; $\det(FIM(\theta))$ 表示 Fisher 信息矩阵的行列式, $\lambda_{\min}(FIM(\theta))$ 和 $\lambda_{\max}(FIM(\theta))$ 分别表示 Fisher 信息矩阵的最小特征值和最大特征值。

设计辨识实验需要考虑选择可测量的输出状态、测量时间和测量频率,还需要考虑以何种方式控制输入的变化。一般,测量输出的选择受实验仪器和实验条件的限制,实验设计的自由度仅仅体现在输入函数的设计,因此,优化实验设计可简化为寻找能够产生最大实验信息度的输入函数。另外,对于非线性模型,由于 Fisher 信息矩阵是与模型参数有关,所以优化实验设计受模型参数的影响,当参数值变化时,优化实验设计也需要相应改变。

3 细胞信号转导网络仿真研究

3.1 NF- κ B 信号转导网络

NF- κ B (nuclear transcription factor-kappa B) 是普遍存在于各种细胞类型中的一种核转录因子,调节参与免疫反应和炎症过程的各种基因^[12-14]。利用 Hoffmann 等建立的 TNF- α 诱导 NF- κ B 信号转导网络的数学模型作为研究对象,模型包括 26 个反应物和 64 个生化反应^[13],其中 24 个反应物是动态的,它们的浓度被定义为状态向量 $x = [x_1 \ x_2 \ \dots \ x_{24}]^T$ 。表 1 是 24 个状态变量所代表的反应物及其初浓度^[13,15]。

根据表 1 的初浓度,系统模型仿真运行 12 小时后,达到稳定,保存此时 24 个状态变量值作为新的初值。然后,给状态变量 x_1 所示反应物 IKK 施加幅值为 0.1 的阶跃扰动,仿

真运行 6.5 小时,观察 x_9 所示的 NF- κ Bn 产生周期性振荡。阶跃信号 IKK 的变化曲线和 x_9 表示的 NF- κ Bn,在 IKK 作用下的输出信号变化曲线分别如图 1 和图 2 所示。

表 1 NF- κ B 信号转导网络的状态变量初始浓度

Table I Initial concentration of states variable of NF- κ B signal transduction networks.

State	Components	Initial Value (μ M)	State	Components	Initial Value (μ M)
x_1	IKK	0	x_{13}	I κ B α n	0
x_2	IKK I κ B α	0	x_{14}	I κ B β n	0
x_3	IKK I κ B β	0	x_{15}	I κ B ϵ n	0
x_4	IKK I κ B ϵ	0	x_{16}	I κ B α -t	0
x_5	iKK I κ B α -NF- κ B	0	x_{17}	I κ B β -t	0
x_6	IKK I κ B β -NF- κ B	0	x_{18}	I κ B ϵ -t	0
x_7	IKK I κ B ϵ -NF- κ B	0	x_{19}	I κ B α -NF- κ B	0
x_8	NF- κ B	0.1	x_{20}	I κ B β -NF- κ B	0
x_9	NF- κ Bn	0	x_{21}	I κ B ϵ -NF- κ B	0
x_{10}	I κ B α	0	x_{22}	I κ B α n-NF- κ Bn	0
x_{11}	I κ B β	0	x_{23}	I κ B β n-NF- κ Bn	0
x_{12}	I κ B ϵ	0	x_{24}	I κ B ϵ n-NF- κ Bn	0

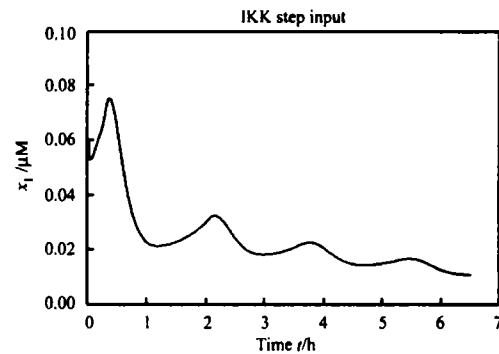


Fig. 1 Time course of IKK input signal.

图 1 输入信号 IKK 变化曲线

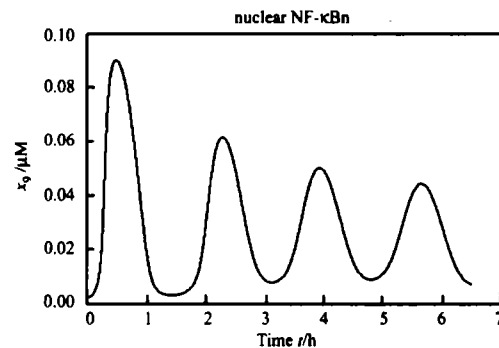


Fig. 2 Time course of nuclear NF- κ B output signal.

图 2 输出信号 NF- κ Bn 变化曲线

3.2 最优实验设计

选择状态变量 x_9 表示的反应物 NF- κ Bn 作为系统的输出信号,状态变量 x_1 表示的反应物 IKK 作为系统的阶跃输入信号,则 NF- κ B 信号转导网络数学模型表示为:

$$\begin{cases} \dot{x} = f(x, \theta), x(t_0) = x_0 \\ y = x_9 \end{cases} \quad (14)$$

NF- κ B 信号转导网络包括 64 个模型参数,经过敏感性分析可知,参数 θ_9 、 θ_{20} 和 θ_{36} 是测量输出信号 NF- κ Bn 的 3 个

敏感参数^[17], 其真实值分别为 0.0204、0.0165、0.00408, 它们代表的速率常数分别为:

θ_9 : IKK I κ B α -NF- κ B catalysis rate, s^{-1} ;

θ_{28} : I κ B α inducible mRNA synthesis rate, $\mu M^{-1} \cdot s^{-1}$;

θ_{36} : constitutive I κ B α translation rate, s^{-1} .

为了分析方便, 只考虑估计敏感参数 θ_9 、 θ_{28} 和 θ_{36} 的最优实验设计。利用参数敏感性分析法^[16,17], 输出 y 关于参数 θ_9 、 θ_{28} 和 θ_{36} 的绝对敏感性函数 S 可以表示为:

$$S = \begin{bmatrix} s_9^0(t_1) & s_{28}^0(t_1) & s_{36}^0(t_1) \\ s_9^0(t_2) & s_{28}^0(t_2) & s_{36}^0(t_2) \\ \vdots & \vdots & \vdots \\ s_9^0(t_N) & s_{28}^0(t_N) & s_{36}^0(t_N) \end{bmatrix} \quad (15)$$

其中 $s_j^0(t_k) = \left. \frac{\partial x_i(t)}{\partial \theta_j} \right|_{t=t_k}$, t_k 表示采样时刻 ($k=1, 2, \dots, N$)。

仿真实验采样时间为 400 分钟, 采样个数 $N=500$, 故 $S \in R^{500 \times 3}$ 。假设测量噪声服从高斯分布, 则 Fisher 信息矩阵可以表示为:

$$FIM(\theta) = \left(\frac{\partial x_9}{\partial \theta} \right)^T Q^{-1} \left(\frac{\partial x_9}{\partial \theta} \right) \quad (16)$$

协方差矩阵 Q 中的 $\sigma_{i,j} = \alpha \times \gamma_{i,j}^{exp}(t, \theta) + \beta$, 其中 α 表示测量相对误差系数, β 表示测量绝对误差值, $\gamma_{i,j}^{exp}(t, \theta)$ 表示第 i 个状态在 j 采样时刻的测量值。

按照 A-准则、D-准则、E-准则和修正 E-准则等 4 个准则, 输入信号 IKK 初浓度为 $0.1 \mu M$, 则参数 θ_9 、 θ_{28} 和 θ_{36} 为真实值时, 得:

$$FIM(\theta) = \begin{bmatrix} 1.318 \times 10^8 & -1.989 \times 10^8 & -8.393 \times 10^8 \\ -1.989 \times 10^8 & 3.637 \times 10^8 & 1.520 \times 10^9 \\ -8.393 \times 10^8 & 1.520 \times 10^9 & 6.363 \times 10^9 \end{bmatrix} \quad (17)$$

$$FIM(\theta)^{-1} = \begin{bmatrix} 6.919 \times 10^{-8} & -2.849 \times 10^{-7} & 7.719 \times 10^{-8} \\ -2.849 \times 10^{-7} & 3.710 \times 10^{-6} & -9.242 \times 10^{-7} \\ 7.719 \times 10^{-8} & -9.242 \times 10^{-7} & 2.312 \times 10^{-7} \end{bmatrix} \quad (18)$$

$$tr(FIM(\theta)^{-1}) = 4.011 \times 10^{-6} \quad (19)$$

$$\det(FIM(\theta)) = 3.625 \times 10^{22} \quad (20)$$

$$\lambda_{\min}(FIM(\theta)) = 2.523 \times 10^5 \quad (21)$$

$$\frac{\lambda_{\max}(FIM(\theta))}{\lambda_{\min}(FIM(\theta))} = 2.709 \times 10^4 \quad (22)$$

假定阶跃输入信号 IKK 初浓度的变化范围为 $[0.01 \mu M, 1 \mu M]$, IKK 初浓度变化时, 则根据不同的设计准则, 研究 NF- κ B 信号转导网络输入信号 IKK 初浓度的变化对估计参数的 95% 置信区间的影响, 寻找最优的输入初浓度。按照 A-准则、D-准则、E-准则和修正 E-准则等得曲线如图 3 所示。图中的垂直虚线分别表示 $\min |tr(FIM(\theta)^{-1})|$ 、 $\max |\det(FIM(\theta))|$ 、 $\max |\lambda_{\min}(FIM(\theta))|$ 和 $\min \left\{ \frac{\lambda_{\max}(FIM(\theta))}{\lambda_{\min}(FIM(\theta))} \right\}$ 所对应的输入信号 IKK 的最优初浓度值, 其中 A-准则、D-准则和 E-准则条件下得输入信号 IKK 的最优初浓度为 $0.06 \mu M$, 修正 E-准则得输入信号 IKK 的最优初浓度为 $0.01 \mu M$ 。

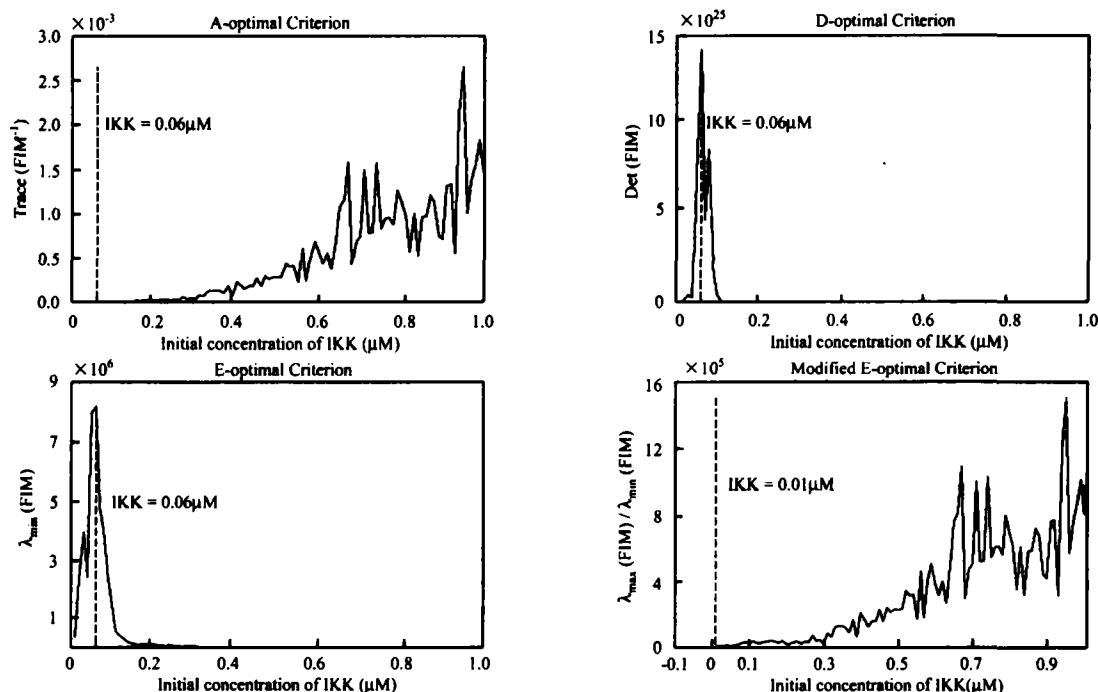


Fig. 3 Optimal experiment design criterion for initial concentration of input signal IKK.

图 3 输入信号 IKK 初始浓度的优化实验设计准则

3.3 参数估计结果

为了在二维空间直观讨论参数变化,这里只估计2个参数,分别估计 $(\hat{\theta}_{28}, \hat{\theta}_9)$ 、 $(\hat{\theta}_{36}, \hat{\theta}_{28})$ 和 $(\hat{\theta}_9, \hat{\theta}_{36})$ 的3种情况,其它参数为定值。按照修正E-准则,得输入信号IKK的最优初浓度为0.01 μM ,而其他3个准则获得输入信号IKK的最优初浓度为0.06 μM ,分别在这两种情况下估计参数对 $(\hat{\theta}_{28}, \hat{\theta}_9)$ 、 $(\hat{\theta}_{28}, \hat{\theta}_{36})$ 和 $(\hat{\theta}_9, \hat{\theta}_{36})$ 的95%置信区间,如图4所示。与A-准则、D-准则及E-准则条件下得输入信号IKK的最优初浓度相比,根据修正E-准则得的输入信号IKK最优初浓度估计参数时,估计参数 $(\hat{\theta}_{28}, \hat{\theta}_9)$ 、 $(\hat{\theta}_{28}, \hat{\theta}_{36})$ 和 $(\hat{\theta}_9, \hat{\theta}_{36})$ 的95%置信区间放大了,说明实验包含的信息量减少了。

根据式(8)式(9)计算估计参数的95%置信区间和误差,结果如表2所示。参数误差的分析,按修正E-准则所得的误差,比E-准则大。例如,E-准则下的参数 θ_{28} 和 θ_9 的误差

分别是8.65%、1.87%,而修正E-准则的误差则为29.11%、26.93%,误差分别增加了20.46%和25.06%。因此,针对NF- κ B信号转导网络,为了提高参数辨识的精度,可以根据E-准则,选择阶跃输入信号IKK的初浓度为0.06 μM 。

表2 估计参数 $\hat{\theta}_9, \hat{\theta}_{28}, (\hat{\theta}_{28}, \hat{\theta}_{36})$ 和 $(\hat{\theta}_9, \hat{\theta}_{36})$ 的95%置信区间
Table 2 95% confidence interval of estimated parameters
 $(\hat{\theta}_{28}, \hat{\theta}_9), (\hat{\theta}_{28}, \hat{\theta}_{36})$ and $(\hat{\theta}_9, \hat{\theta}_{36})$.

Parameters	95% Confidence interval			
	E-optimal design	Error %	Modified E-optimal design	Error %
$(\hat{\theta}_{28}, \hat{\theta}_9)$	$(\hat{\theta}_{28}): [0.01507, 0.01793]$	8.65	$(\hat{\theta}_{28}): [0.01169, 0.02130]$	29.11
	$(\hat{\theta}_9): [0.02002, 0.02078]$	1.87	$(\hat{\theta}_9): [0.01491, 0.02589]$	26.93
$(\hat{\theta}_{28}, \hat{\theta}_{36})$	$\hat{\theta}_{28}: [0.01567, 0.01732]$	5.01	$\hat{\theta}_{28}: [0.01416, 0.01887]$	14.27
	$\hat{\theta}_{36}: [0.00377, 0.00439]$	7.72	$\hat{\theta}_{36}: [0.00315, 0.00501]$	22.84
$(\hat{\theta}_9, \hat{\theta}_{36})$	$\hat{\theta}_9: [0.01907, 0.02173]$	6.52	$\hat{\theta}_9: [0.01704, 0.02379]$	16.54
	$\hat{\theta}_{36}: [0.00389, 0.00428]$	4.78	$\hat{\theta}_{36}: [0.00313, 0.00503]$	23.28

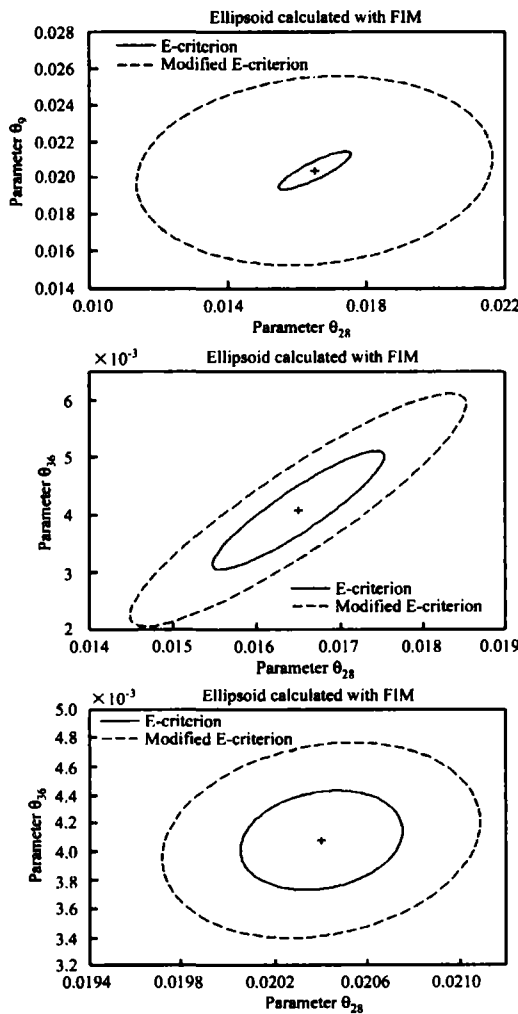


Fig. 4 95% confidence ellipses for parameters $(\hat{\theta}_{28}, \hat{\theta}_9), (\hat{\theta}_{28}, \hat{\theta}_{36})$ and $(\hat{\theta}_9, \hat{\theta}_{36})$

图4 不同实验设计准则下参数 $(\hat{\theta}_{28}, \hat{\theta}_9), (\hat{\theta}_{28}, \hat{\theta}_{36})$ 和 $(\hat{\theta}_9, \hat{\theta}_{36})$ 的95%置信区间

4 结论

分析NF- κ B信号转导网络的数学模型,选择NF- κ Bn为系统可测输出信号,IKK为阶跃输入信号,以Fisher信息矩阵 $FIM(\theta)$ 和估计参数的协方差矩阵 Σ 为基础,研究各种准则下,阶跃输入信号IKK初浓度的变化对参数估计精度的影响。在A-准则、D-准则、E-准则条件下得输入信号IKK的最优初浓度为0.06 μM ,而修正E-准则得输入信号IKK的最优初浓度为0.01 μM 。仿真结果显示,根据修正E-准则会导致估计参数的95%置信区间增大,所以,实验可用E-准则,选择阶跃输入信号IKK的初浓度为0.06 μM 。

细胞信号转导网络的计算机模拟表明:实验的信息量大大依赖于设计的实验,通过改变输入信号IKK的初浓度,参数估计误差显著降低。当然,输入信号的初浓度,因辨识参数的不同而发生变化。因此,在实验数据采集之前,选定影响系统辨识精度的各个因素,以达到最优为准,同时综合考虑实验条件、数学论证和实施方法等的制约,根据适当的准则,合理安排各种实验因素,严格控制实验误差,有效分析实验数据,从而用较少的人力、物力和时间,最大限度获得丰富而可靠的信息。

References

- 1 Neves SR and Iyengar R. Modeling of signaling networks. *BioEssays*, 2002, 24(12):1110-1117.
- 2 Asthagiri A and Lauffenburger D. Bioengineering models of cell signaling. *Annual Review of Biomedical Engineering*, 2000, 2:31-53.
- 3 Cho KH, Shin SY and Kolch W, et al. Experimental design in systems biology, based on parameter sensitivity analysis using a Monte Carlo method; a case study for the TNF α -mediated NF- κ B signal transduction pathway. *Simulation*, 2003, 79(12):726-739.
- 4 Vanrolleghem PA, Daele MV and Dochain D. Optimal experimental design for parameter estimation in activated sludge processes. *Proceeding of the American Control Conference*, Seattle, Washington,

- June, 1995.
- 5 Faller D, Klingmuller U and Timmer J. Simulation methods for optimal experimental design in systems biology. *Simulation*, 2003, 79(12):717 - 725.
 - 6 Kutalik Z, Cho KH and Wolkenhauer O. Optimal sampling time selection for parameter estimation in dynamic pathway modeling. *Bio-systems*, 2004, 75(1-3):43 - 55.
 - 7 Asprey SP and Macchietto S. Designing robust optimal dynamic experiments. *Journal of Process Control*, 2002, 12(4):545 - 556.
 - 8 Hamby ES and Kabamba PT. Optimal design of experiments for control: a preposition viewpoint. *Proceeding of 38 American Control Conference*, 1999:3446 - 3450.
 - 9 Marsili-Libelli S, Guerrizio S and Checchi N. Confidence regions of estimated parameters for ecological systems. *Ecological modeling*, 2003, 165:127 - 146.
 - 10 Ljung L. *System Identification: Theory for the User*. NJ: Prentice Hall, Englewood Cliffs, 1999.
 - 11 Cover TM and Thomas JA. *Elements of the Information Theory*. John Wiley & Sons Inc, 1991.
 - 12 Baeuerle PA and Henkel T. Function and activation of NF- κ B in the immune system. *Annual Review of Immunology*, 2004, 12:141 - 179.
 - 13 Hoffmann A, Levchenko A and Scott ML, et al. The I κ B-NF- κ B signaling module: temporal control and selective gene activation. *Science*, 2002, 298:1241 - 1245.
 - 14 Lipniacki T, Paszek P and Brasier AR, et al. Mathematical model of NF- κ B regulatory module. *Journal of Theoretical Biology*, 2004, 228(2):195 - 215.
 - 15 Ihekwaba AEC, Broomhead DS and Grisley RL, et al. Sensitivity analysis of parameters controlling oscillatory signaling in the NF- κ B pathway: the roles of IKK and I κ B α . *Journal of Systems Biology*, 2004, 7(1):93 - 103.
 - 16 Turányi T. Sensitivity analysis of complex kinetic systems: tools and applications. *Journal of Mathematical Chemistry*, 1990, 5(3):203 - 248.
 - 17 Yue H, Brown M and Knowles J, et al. Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an NF- κ B signaling pathway. *Molecular Bio-systems*, 2006, 2(12):640 - 649.