

一种交叉熵算法的宏观交通模型标定方法

傅恺延^{1,2}, 钟任新^{1,2}, 黄云萍^{1,2}, 熊建辉^{1,2}

(1. 中山大学工学院, 广州 510006; 2. 广东省智能交通系统重点实验室, 广州 510006)

摘要: 针对交通仿真模型标定问题往往存在大量的局部最优解, 要求求解算法具备优秀的搜索能力与收敛性质的情况, 提出了结合蒙特卡罗与重要样本策略逐步逼近最优概率密度函数——基于交叉熵算法的交通模型标定方法。实例分别采用已知全局最优值的合成数据、实际观测数据对元胞传输模型进行参数标定, 并对比遗传算法进行验证。实例验证结果表明, 交叉熵算法具有高效的标定效率及优异的收敛性质。

关键词: 智能交通; 交通模型标定; 交叉熵算法; 高效收敛

中图分类号: U491.112

文献标志码: A

文章编号: 2095-2783(2017)14-1627-07

Cross entropy method for calibrating the first-order macroscopic traffic flow models

FU Kaiyan^{1,2}, ZHONG Renxin^{1,2}, HUANG Yunping^{1,2}, XIONG Jianhui^{1,2}

(1. School of Engineering, Sun Yat-Sen University, Guangzhou 510006, China; 2. Guangdong Provincial Key Laboratory of Intelligent Transportation Systems, Guangzhou 510006, China)

Abstract: A number of local optimal solutions for the calibration formulation requires that the solving algorithms have an excellent search ability and convergence property. A cross entropy method approaching the optimal probability density function with Monte Carlo and important sampling strategy is thus proposed to solve the calibration problem. Empirical studies calibrate the cell transmission model with synthetic data and it was observed that the data contrasted with genetic algorithm. It was concluded that the cross entropy method has high calibration efficiency and excellent convergence property.

Keywords: intelligent transportation; traffic models calibration; cross entropy method; efficient convergence

交通仿真模型在精确模拟交通状态之前, 必须对其参数进行标定。以元胞传输模型(cell transmission model, CTM)为例, 其仿真准确程度主要取决于宏观基本图的参数, 可通过最小二乘法逐步从历史宏观数据中直接计算求得^[1]。然而, 基于大量历史数据堆叠的最小二乘法虽然求解结构简单, 计算效率高, 但是采用最大观测流量确定道路通行能力存在较大的误差, 并且容易忽视交通流中的异常情形与小概率事件, 不便于分析异常交通流的形成原因与疏散机理。

相对于最小二乘法, 通过随机搜索算法求解模型输出值与观测数据的差异最小值的优化问题, 是实现模型参数标定的常用方法。随机搜索算法的主要思想是把参数可行域划分为多个子集, 按照一定的原理将参数逐步收敛至全局最优值附近, 包括模拟退火算法、多启发式算法、遗传算法等^[2]。作为求解优化问题的标杆算法, 遗传算法对于待求解优化问题在可导方面不做任何要求, 在交通模型标定方面得到了广泛的研究与应用^[2-3]。然而, 上述构建的优化问题往往存在着大量的局部最优值, 要求求解算法具备优秀的搜索能力与收敛性质; 此外, 随着交通模型的复杂度与标定数据的丰富度增加, 随机搜

索算法需要沉重的计算负荷实现确定性或随机性的模型标定, 降低了模型的标定效率。

作为新颖的随机搜索算法, 交叉熵算法(cross entropy method, CEM)基于蒙特卡罗与重要样本的策略, 常用于小概率事件的概率估计与优化问题的求解, 并拓展应用于交通模型标定与信号配时控制优化, 呈现出众的搜索能力与计算效率^[4-6]。因此, 针对交通仿真模型标定存在上述的问题, 本文提出基于交叉熵算法的交通模型标定框架, 通过预设参数构建人工合成数据, 标定CTM的基本图验证交叉熵算法的全局搜索能力, 并与遗传算法的标定性能进行对比。

1 交叉熵算法

总体来说, 交叉熵算法将优化问题转换为某一小概率事件估计问题, 通过蒙特卡罗与重要样本策略求解, 从而获得优化问题的解^[7-8]。

考虑以下的最小值优化问题:

$$\gamma^* = \min_{x \in \chi} S(x). \quad (1)$$

式中: γ^* 为函数 $S(x)$ 的最小值; x 为定义在函数空间 χ 中的参数样本。首先初始化 1 个函数空间 χ 中的含参数 ν 的概率密度函数, 用 $f(x; \nu)$ 表示。在求

收稿日期: 2017-04-26

基金项目: 高等学校博士学科点专项科研基金资助项目(20130171120032); 国家自然科学基金资助项目(51308559)

第一作者: 傅恺延(1991-), 男, 硕士研究生, 主要研究方向为交通信息工程与控制

通信作者: 钟任新, 副教授, 主要研究方向为动态交通监控、最优控制等, zhrenxin@mail.sysu.edu.cn

解 $S(x)$ 最小值的过程中,定义 1 个小概率事件 $S(x) \leq \gamma$ (γ 为 1 个充分接近最小值 γ^* 的阈值),其概率的求解为

$$\ell(\gamma) = P_u(S(x) \leq \gamma) = E_u(I_{\{S(x) \leq \gamma\}}). \quad (2)$$

式中: $x = \{x_1, x_2, \dots, x_n\}$ 为服从概率密度函数 $f(x; u)$ (令 $v = u$) 分布的参数样本; P_u 、 E_u 分别为 x 的概率与期望; $I(\cdot)$ 为指标函数,当 $S(x) \leq \gamma$ 成立时 $I_{\{S(x) \leq \gamma\}} = 1$, 反则 $I_{\{S(x) \leq \gamma\}} = 0$ 。通过上述方法,1 个常规的优化问题转换为小概率事件的估计问题:求解优化问题的目标转换为最大化 γ 使得 $\ell(\gamma)$ 充分接近 0。

求解上述小概率事件包括 2 个主要步骤:一是给定 γ 计算 $\ell(\gamma)$;二是根据给定的 γ 搜索更大的 γ 。第 1 个步骤中,通过无止尽的样本采样计算 $\ell(\gamma)$ 是不切合实际的做法,取而代之是生成一定数量的样本,通过蒙特卡罗的方法估计 $\ell(\gamma)$ 。具体来说,生成 N 个服从 $f(x; u)$ 分布的样本, $\ell(\gamma)$ 估算为

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^N I_{\{S(x_i) \leq \gamma\}}. \quad (3)$$

然而,通过蒙特卡罗估算的概率 $\ell(\gamma)$ 存在较大的误差。另一方面,当 γ 充分接近 γ^* 时, $\ell(\gamma)$ 不断接近 0,也就是 $S(x) \leq \gamma$ 将成为 1 个小概率事件,需要更多的样本精确估计概率 $\ell(\gamma)$ 。因此,交叉熵算法采样重要样本策略 (important sampling technique) 提高概率估计精确。

简单来说,重要样本策略采用另一概率密度函数 $z(x; \theta)$ 估计概率 $\ell(\gamma)$:

$$\hat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^N I_{\{S(x_i) \leq \gamma\}} \frac{f(x_i; u)}{z(x_i; \theta)}. \quad (4)$$

定义

$$z^*(x; \theta) = \frac{I_{\{S(x) \leq \gamma\}} f(x; u)}{\ell(\gamma)}, \quad (5)$$

将 z 替换为 z^* , 可得:

$$\hat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^N I_{\{S(x_i) \leq \gamma\}} \frac{f(x_i; u)}{z^*(x_i; \theta)} = \ell(\gamma). \quad (6)$$

因为 $\ell(\gamma)$ 未知,则当前优化问题的求解转换为最优概率密度函数 z^* 的估计:即通过交叉熵距离尝试寻找 1 个充分接近 z^* 的概率密度函数。对于 2 个连续的概率密度函数,其交叉熵距离计算为

$$D_{CE}(z^* \parallel f) = \int z^*(x; \theta) \ln z^*(x; \theta) dx - \int z^*(x; \theta) \ln f(x; \nu) dx. \quad (7)$$

最小化上式的交叉熵距离,等价于寻找 ν 使得:

$$\max_{\nu} \int z^*(x; \theta) \ln f(x; \nu) dx. \quad (8)$$

根据 z^* 的定义代入上式,可得

$$\max_{\nu} \int \frac{I_{\{S(x) \leq \gamma\}} f(x; u)}{\ell(\gamma)} \ln f(x; \nu) dx. \quad (9)$$

联立 $\ell(\gamma)$ 的定义,可获得下述等价的优化问题:

$$\nu^* = \operatorname{argmax}_{\nu} E_u I_{\{S(x) \leq \gamma\}} \ln f(x; \nu). \quad (10)$$

再次采用重要样本策略,通过另一个参数 w , 上述优化问题可写成:

$$\nu^* = \operatorname{argmax}_{\nu} E_w I_{\{S(x) \leq \gamma\}} W(x; u, w) \ln f(x; \nu). \quad (11)$$

其中定义

$$W(x; u, w) = \frac{f(x; u)}{f(x; w)}. \quad (12)$$

$W(x; u, w)$ 作为概率密度函数 $f(\cdot; u)$ 与 $f(\cdot; w)$ 对于 x 的似然比率。因此, ν^* 的估计求解为:

$$\hat{\nu}^* = \operatorname{argmax}_{\nu} \frac{1}{N} \sum_{i=1}^N I_{\{S(x_i) \leq \gamma\}}; \quad (13)$$

$$W(x_i; u, w) \ln f(x_i; \nu) = \operatorname{argmax}_{\nu} \hat{D}(\nu).$$

式中 x_1, x_2, \dots, x_N 为服从概率密度函数 $f(\cdot; w)$ 的随机样本。对于 \hat{D} 为凸函数,且对于 ν 可导的优化问题,其解可通过下式求得:

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(x_i) \leq \gamma\}} W(x_i; u, w) \nabla \ln f(x_i; \nu) = 0. \quad (14)$$

式中 ∇ 为函数对于 ν 的梯度。

值得注意的是,上述生成随机样本的概率密度函数为任意含参数的概率密度函数,且与待优化问题的函数形式无关。文献[9]指出选取正态分布作为概率密度函数可以简化交叉熵算法的求解结构。基于此,交叉熵算法求解优化问题的步骤为:

步骤 1 初始化 ($l = 0$), 确定含参正态分布的均值 $\hat{\mu}_0$ 与标准差 $\hat{\sigma}_0$ 。

步骤 2 令 $l = l + 1$, 并生成 N 个服从正态分布 $N(\hat{\mu}_{l-1}, \hat{\sigma}_{l-1}^2)$ 的随机样本 (x_1, x_2, \dots, x_N) 。

步骤 3 计算所有随机样本的目标函数值 $\hat{S}(x_i)$ 。

步骤 4 将目标函数值从小到大进行排序,取前 $\omega\%$ 的目标函数值对应的随机样本作为重要样本 N_E 。

步骤 5 通过重要样本更新正态分布的均值与标准差: $\hat{\mu}_l = \sum x_j / \omega\% N$, $\hat{\sigma}_l^2 = \sum (x_j - \hat{\mu}_l)^2 / \omega\% N$, 其中 $j = 1, 2, \dots, (\omega\% N)$ 。

步骤 6 参数平滑: $\hat{\mu}_l = \beta \hat{\mu}_l + (1 - \beta) \hat{\mu}_{l-1}$, $\hat{\sigma}_l = \beta \hat{\sigma}_l + (1 - \beta) \hat{\sigma}_{l-1}$ 。

步骤 7 若 $\hat{\sigma}_l < \epsilon$, 返回步骤 2; 否则,取 $\hat{\mu}_l$ 作为优化问题的解,对应的 $\hat{S}(\hat{\mu}_l)$ 为其最小值。

上述步骤中,交叉熵算法参数的意义及确定准则如下^[7,9]: 1) N 是随机样本的生成数量,本文取 1 000; 2) $\omega\%$ 是重要样本占总体样本的百分数,取值范围为 0% 至 100%, 本文 $\omega\%$ 设置为 1%; 3) β 是参数平滑系数,让概率密度函数参数平滑收敛变化,本文 β 取 0.7; 4) ϵ 为算法收敛的预设阈值。

2 改进元胞传输模型

CTM 在时间和空间上离散化路段交通状态,从

而演化与分析路段的宏观交通状态^[10],同时采用基本图描述路段的流量、速度、密度三者之间的关系,其中三角形基本图如图 1 所示。基本图的参数包括自由流速度 v_f 、拥堵波速 v_j 、最大通行流量 Q_M 、最大通行流量密度 k_c 和拥堵密度 k_j 。

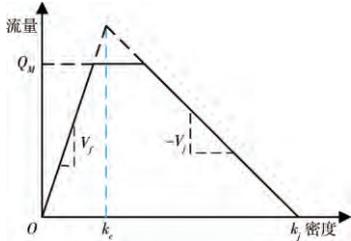


图 1 三角形基本图及其参数

路段交通密度主要衡量道路交通拥堵的程度,常作为交通监控与交通管制等的控制指标之一^[11-12]。基本图中交通流量到交通密度的映射值有 2 个,而从交通密度到交通流量的映射却是唯一的。基于此,传统基于交通流量的 CTM,拓展到以交通密度为演化量的改进元胞传输模型(modified cell transmission model, MCTM)^[13]。MCTM 允许不均等长度的元胞,提高了重构仿真高速公路交通流状态的灵活性。与 CTM 类似,MCTM 划分元胞的示意图如图 2 所示。

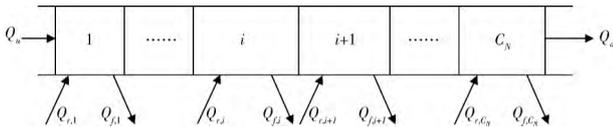


图 2 MCTM 元胞结构示意图

以元胞的上下边界流量 Q_u 与 Q_d 、上下匝道流量 $Q_{r,i}$ 与 $Q_{f,i}$ 作为 MCTM 的边界条件,根据流量守恒定律,元胞 i 在单位时间间隔 ΔT 内交通密度为

$$\rho_{i,t+1} = \rho_{i,t} + \frac{\Delta T}{L_i} (Q_{i,t} - Q_{i+1,t} + Q_{r,i,t} - Q_{f,i,t}) \quad (15)$$

式中: $\rho_{i,t+1}$ 为元胞 i 在第 $t+1$ 个时间间隔内的交通密度; $Q_{i,t}$ 、 $Q_{i+1,t}$ 分别为在时间间隔 t 内进入元胞 i 与元胞 $i+1$ 的交通流量; ΔT 为时间间隔; L_i 为元胞 i 的元胞长度; $Q_{r,i,t}$ 、 $Q_{f,i,t}$ 分别为进入上匝道与离开下匝道的交通流量。

进入元胞 i 的交通流量为

$$Q_{i,t} = \min\{\bar{D}_{i-1,t}, \bar{R}_{i,t}\} \quad (16)$$

其中: $\bar{D}_{i-1,t} = D_{i-1,t} - Q_{f,i-1,t}$; $\bar{R}_{i,t} = R_{i,t} - Q_{r,i,t}$ 。

式中 $D_{i-1,t}$ 为元胞 $i-1$ 的发送函数(sending function),定义为元胞 $i-1$ 在单位时间间隔内,自由流状态下能流出的最大交通流量,即:

$$D_{i-1,t} = \min\{v_{f,i-1} \rho_{i-1,t}, Q_{M,i-1}\} \quad (17)$$

$R_{i,t}$ 为元胞 i 的接收函数(receiving function),定义为元胞 i 在单位时间间隔内,拥堵状态下能够流入的

最大交通流量,即

$$R_{i,t} = \min\{Q_{M,i}, v_{c,i}(k_{j,i} - \rho_{i,t})\} \quad (18)$$

发送函数与接收函数均受基本图参数的影响。此外,MCTM 必须满足车辆在自由流状态下,单位时间间隔内的行驶距离不能超过元胞的长度,即 $v_{f,i} \Delta T \leq L_i$ 。

针对边界可能处于交通拥堵状态,不能及时疏散全部交通流量的问题,添加以下限制条件:

$$Q_{1,t} = \begin{cases} \min\{\bar{Q}_{u,t}, R_{1,t}\}, & \rho_{u,t} \leq k_{c,1}; \\ R_{1,t}, & \text{其他} \end{cases}; \quad (19)$$

$$Q_{c_{N+1},t} = \begin{cases} D_{c_{N+1},t}, & \rho_{d,t} \leq k_{c,c_N} \\ \min\{D_{c_{N+1},t}, \bar{Q}_{d,t}\}, & \text{其他} \end{cases} \quad (20)$$

式中: $\bar{Q}_{u,t}$ 、 $\bar{Q}_{d,t}$ 分别上下游边界的检测交通流量; $\rho_{u,t}$ 、 $\rho_{d,t}$ 分别为上下游边界的检测交通密度。

为了统一模型的表达形式,定义

$$Q_{u,t} = \bar{Q}_{u,t} \delta_{1,t}; \quad Q_{d,t} = \bar{Q}_{d,t} \delta_{2,t} \quad (21)$$

其中 $\delta_{1,t}$ 和 $\delta_{2,t}$ 定义为:

$$\delta_{1,t} = \begin{cases} 1, & \rho_{u,t} \leq k_{c,1} \\ \infty, & \text{其他} \end{cases}; \quad (22)$$

$$\delta_{2,t} = \begin{cases} \infty, & \rho_{d,t} \leq k_{c,c_N} \\ 1, & \text{其他} \end{cases} \quad (23)$$

因此,MCTM 的边界限制条件改写为

$$Q_{1,t} = \min\{Q_{u,t}, \bar{R}_{1,t}\}; \quad (24)$$

$$Q_{c_{N+1},t} = \min\{\bar{D}_{c_{N+1},t}, Q_{d,t}\} \quad (25)$$

综上所述,MCTM 交通流的传播与交通密度的演算由发送函数与接收函数共同确定。

3 实例验证

3.1 参数标定模型

建立用于标定 MCTM 参数的模型,目标为最小化模型输出与观测数据之间差异:

$$x^* = \arg \min_{x \in X} S(M(x), O) \quad (26)$$

式中: x^* 为模型参数的最优值; $S(M(x), O)$ 为目标函数的形式,用于衡量模型输出 $M(x)$ 与观测数据 O 之间的差异。输入边界交通流量,MCTM 的输出通过式(15)、式(16)、式(24)与式(25)联合求解。

基于三角形基本图的 MCTM 的参数有 5 个,只需确定其中 3 个,另外 2 个可通过参数之间的关系求得^[14]。选择自由流速度 v_f 、拥堵波速 v_j 、最大通行流量 Q_M 作为决策变量,因为上述 3 个参数决定三角形基本图的斜率与顶点,其余 2 个参数最大通行流量密度与拥堵密度通过下式计算:

$$k_c = \frac{Q_M}{v_f}; k_j = \frac{v_f + v_j}{v_f v_j} Q_M \quad (27)$$

采用平均绝对百分误差(mean absolute percentage error, MAPE)衡量模型输出密度与观测密度之

间的差异:

$$S(M(x), O) = \frac{1}{C_N} \sum_{i=1}^{C_N} \sum_{t=1}^{T_N} \left| \frac{M_{i,t}(x) - O_{i,t}}{O_{i,t}} \right|. \quad (28)$$

式中 $M_{i,t}(x)$ 、 $O_{i,t}$ 分别为元胞 i 在时间间隔 t 内的模型输出密度与观测密度。

3.2 路段信息与数据处理

试验路段位于英国伦敦 M25 高速公路,是交叉口编号为 15 的路段,全长 2 km。图 3 给出了试验路段的结构与各个传感器的分布位置以及对应的传感器编号,每个传感器间隔 500 m。上下边界(包括上下匝道)的传感器数据作为标定的边界条件与校验数据,中间传感器数据仅作为校验数据。每个元胞的长度划分为 250 m,仿真时间间隔为 5 s。

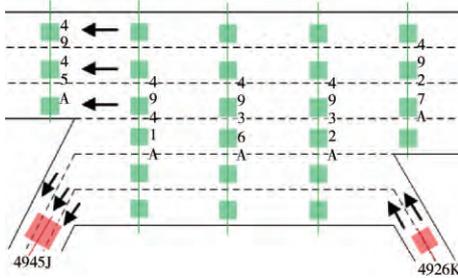


图 3 路段结构与传感器位置示意图

数据来源于英国 Motorway Incident Detection and Automatic Signaling (MIDAS) 系统,收集的数据包括交通流量、车辆速度、占有率(在统计间隔内传感器被占有的时间百分比),并估计空间交通密度,数据整合时间间隔为 1 min^[15]。数据采集时间为 2012 年 9 月 12 号 14 时至 20 时,并采用零阶差值方法生成充足的时间间隔为 5 s 的数据。

根据交叉熵算法的思想,将 MCTM 的参数标定问题转化为小概率事件的估计问题。通过离散化参数可行的子域,初始化参数正态分布的均值与标准

差,使模型输出尽可能与数据集所反映的交通状态相吻合。因此,采用最小二乘法标定的每个子路段的参数对参数进行初始化,结果如表 1 所示。元胞参数的初始值设置为元胞最接近的传感器对应的基本图参数,参数正态分布的标准差分别为 10 km/h、10 km/h、2 500 辆/h,依次对应自由流速度 v_f 、拥堵波速 v_j 和最大通行流量 Q_M 。

表 1 最小二乘法标定结果

| 元胞 | $v_f / (\text{km} \cdot \text{h}^{-1})$ | $v_j / (\text{km} \cdot \text{h}^{-1})$ | $Q_M / (\text{辆} \cdot \text{h}^{-1})$ |
|--------|---|---|--|
| 元胞 1 | 99 | 16 | 8 787 |
| 元胞 2,3 | 97 | 16 | 8 710 |
| 元胞 4,5 | 109 | 18 | 8 627 |
| 元胞 6,7 | 103 | 17 | 8 528 |
| 元胞 8 | 99 | 17 | 7 949 |

3.3 合成数据标定

由于最小化目标函数的解不一定是模型全局最优值,因此在评估算法搜索全局最优值的能力时,模型的全局最优值必须是已知的。所以,标定已知全局最优值的人工合成数据,能够清晰对比参数标定结果与全局最优值的差异,展现标定算法寻找全局最优值的能力。

实验路段的边界数据作为输入,预设各个元胞的基本图参数,记录 MCTM 仿真的各个元胞的交通密度作为人工合成数据。将上述获得的合成数据与相同的边界数据作为输入,采用交叉熵算法标定 MCTM,将标定结果与预设的参数值进行对比,采用相对误差 ($|\hat{x}^* - x^*| / x^*$) 衡量算法全局最优值的搜索能力,结果如表 2 所示。预设的基本图参数为 $x^* = \begin{cases} 110 & 110 & 115 & 115 & 105 & 105 & 100 & 100 \\ 20 & 20 & 25 & 25 & 28 & 28 & 30 & 30 \\ 9600 & 9600 & 9800 & 9800 & 10000 & 10000 & 10200 & 10200 \end{cases}$, 分别对应 8 个元胞的 v_f 、 v_j 与 Q_M 。

表 2 人工合成数据标定结果

| 元胞 | v_f | | v_j | | Q_M | | k_c | | k_j | |
|------|---|------------|---|------------|--|------------|---|------------|---|------------|
| | 标定值/ ($\text{km} \cdot \text{h}^{-1}$) | 相对 误差/% | 标定值/ ($\text{km} \cdot \text{h}^{-1}$) | 相对 误差/% | 标定值/ ($\text{辆} \cdot \text{h}^{-1}$) | 相对 误差/% | 标定值/ ($\text{辆} \cdot \text{km}^{-1}$) | 相对 误差/% | 标定值/ ($\text{辆} \cdot \text{km}^{-1}$) | 相对 误差/% |
| 元胞 1 | 110 | 0.19 | 20 | 0.91 | 9 617 | 0.17 | 87 | 0.02 | 564 | 0.62 |
| 元胞 2 | 110 | 0 | 22 | 10.19 | 9 799 | 2.07 | 89 | 2.07 | 534 | 5.91 |
| 元胞 3 | 115 | 0 | 27 | 8.98 | 9 979 | 1.82 | 87 | 1.83 | 453 | 5.07 |
| 元胞 4 | 115 | 0.01 | 24 | 2.83 | 9 745 | 0.56 | 85 | 0.55 | 486 | 1.82 |
| 元胞 5 | 105 | 0 | 27 | 3.46 | 9 930 | 0.70 | 95 | 0.71 | 462 | 2.10 |
| 元胞 6 | 105 | 0.01 | 27 | 4.57 | 9 906 | 0.94 | 94 | 0.95 | 465 | 2.80 |
| 元胞 7 | 100 | 0 | 29 | 4.01 | 10 116 | 0.82 | 101 | 0.82 | 452 | 2.36 |
| 元胞 8 | 103 | 2.86 | 33 | 9.76 | 10 492 | 2.86 | 102 | 0 | 421 | 4.84 |

从表 2 看出,交叉熵算法的标定结果非常接近预设参数,特别是自由流速度、最大通行流量与最大通行流量密度,相对误差均小于 2%,甚至达到 0,与

预设参数完全相等,充分表明交叉熵算法搜索全局最优值的性能。拥堵波速存在接近 10% 的相对误差,可能的原因是实验路段涉及车道数量的变换,车

辆交织区中的交通流行为的影响因素更为复杂,同时受到路段道路物理结构与驾驶员驾驶行为(如换道行为)等其他因素影响,降低了模型对于交通流拥堵情况的识别能力,导致标定结果无法充分接近预设参数。尽管如此,标定结果的平均绝对百分误差为 0.67%,重现的交通流情况与合成数据非常吻合。

3.4 观测数据标定分析

将交叉熵算法与遗传算法(genetic algorithm, GA)进行对比来分析交叉熵算法标定观测数据的性能,从标定效率与收敛性质上评价两个算法之间的差异。

遗传算法是模拟生物进化机制的全局随机搜索算法,其内部包含选择与再生、交叉、变异等筛选机制,在每次迭代中保存优值子代的遗传信息,直到算法收敛。采用 MATLAB 的 GATBX 工具箱标定智能驾驶模型的参数,子代数量为 1 000,最大迭代次数为 100。

交叉熵算法与遗传算法按照统一的收敛准则,即目标函数的最优值 $S_{l,best}$ 与前 t_{last} 次迭代的目标函数最优值的均值之间的差值小于预设阈值 ϵ :

$$\| S_{l,best} - \frac{1}{t_{last}} \sum_{i=l-t_{last}+1}^l S_{l,best} \| \leq \epsilon. \quad (29)$$

其中 ϵ 取 5×10^{-5} 。

统计 2 种算法完成标定相同总样本数量所需的 MCTM 仿真次数及最优目标函数值各 5 次。仿真次数统计完成 MCTM 标定总共需要对 MCTM 模型进行计算的次数,仿真次数越少,标定效率越高;最

优目标函数值记录最优仿真数据与观测数据之间差异,函数最优目标数值越小,仿真结果越精确。统计结果如表 3 所示:

表 3 交叉熵算法与遗传算法的标定结果

| 编号 | 仿真次数 | | 最优目标函数值/% | |
|----|-------|-------|-----------|-------|
| | 交叉熵算法 | 遗传算法 | 交叉熵算法 | 遗传算法 |
| 1 | 43 | 100.0 | 9.69 | 10.11 |
| 2 | 45 | 100.0 | 9.81 | 10.11 |
| 3 | 40 | 92.0 | 9.68 | 10.10 |
| 4 | 45 | 100.0 | 9.74 | 10.10 |
| 5 | 42 | 100.0 | 9.64 | 10.09 |
| 平均 | 43 | 98.4 | 9.71 | 10.10 |

从表 3 可看出,交叉熵算法采用不到遗传算法一半的仿真次数,就完成了更优目标函数的搜索,表明在该数据下,交叉熵算法标定 MCTM 模型比遗传算法更高效更准确。具体来说,交叉熵算法采用重要样本策略在迭代过程中不断收敛参数的标准差,逐步逼近最优参数,收敛方向明确,速度快;而遗传算法在标定前期通过选择与再生、交叉等机制保存优值遗传信息,不断获得更优子代。但是,当优质子代逐渐增多后,遗传算法的变异机制起主要影响作用,越来越多的子代尝试向更为优质的子代变异却没有变异成功,这就导致某些子代的目标函数值出现极其不合理的值,减缓了参数的收敛速度与收敛方向。图 4 呈现了元胞 3 的拥堵波速的收敛过程。

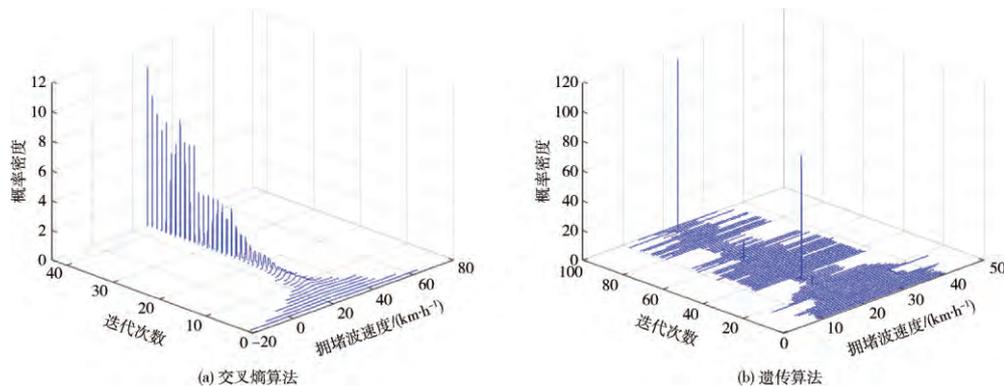


图 4 拥堵波速的收敛情况

从图 4 可看出,交叉熵算法仅需 42 次迭代即完成了算法收敛,遗传算法直到最大的迭代次数(设置为 100),也无法收敛到最优目标函数值。具体来说,交叉熵算法在迭代初期,在拥堵波速的搜索范围内随机分布,各个数值的概率密度都较低,确保算法初期拥有较大的搜索空间。随着算法的迭代收敛,拥堵波速的分布越来越集中于某个参数(最优值),甚至在第 30 次迭代后,拥堵波速的概率密度函数近乎一条直线,充分逼近最优值,满足收敛条件。另一方面,遗传算法在整个标定期间,都在尝试寻找更优的

子代,无法使所有的参数向最优值聚拢。换言之,遗传算法自始至终都在参数的可行范围内尝试寻找更优子代,与此形成鲜明对比的是交叉熵算法则不断缩小可能的参数子域,集中搜索更为精确的最优值,满足更为严格的收敛条件,如收敛平均目标函数等^[6]。

3.5 模型校验

将交叉熵算法、遗传算法与最小二乘法的标定结果作为输入,对 MCTM 的输出进行校验,与观测密度的对比时序图如图 5 所示,对应的 MAPE 分别

为 9.64%、10.09% 与 20.07%。从图 5 可清晰看出,最小二乘法在交通拥堵的状态下低估了实际的交通密度,在交通通畅的状态下却高估了实际交通密度。主要原因是最小二乘法通过观测交通流量确定的最大通行流量存在较大的误差,导致传输函数与接受函数的输出交通流量无法准确反映实际的交通情况。此外,最小二乘法独立于模型进行参数标定是

其误差较大的另一个主要原因。最小二乘法无需进行模型仿真,故其标定效率最高,但模型精度较低。而交叉熵算法与遗传算法仿真输出的交通密度几乎重合,均能够更准确地重现实际交通流的拥堵形成与疏散的大致情况,且交叉熵算法的标定效率、精度均高于遗传算法,与观测数据更为吻合。

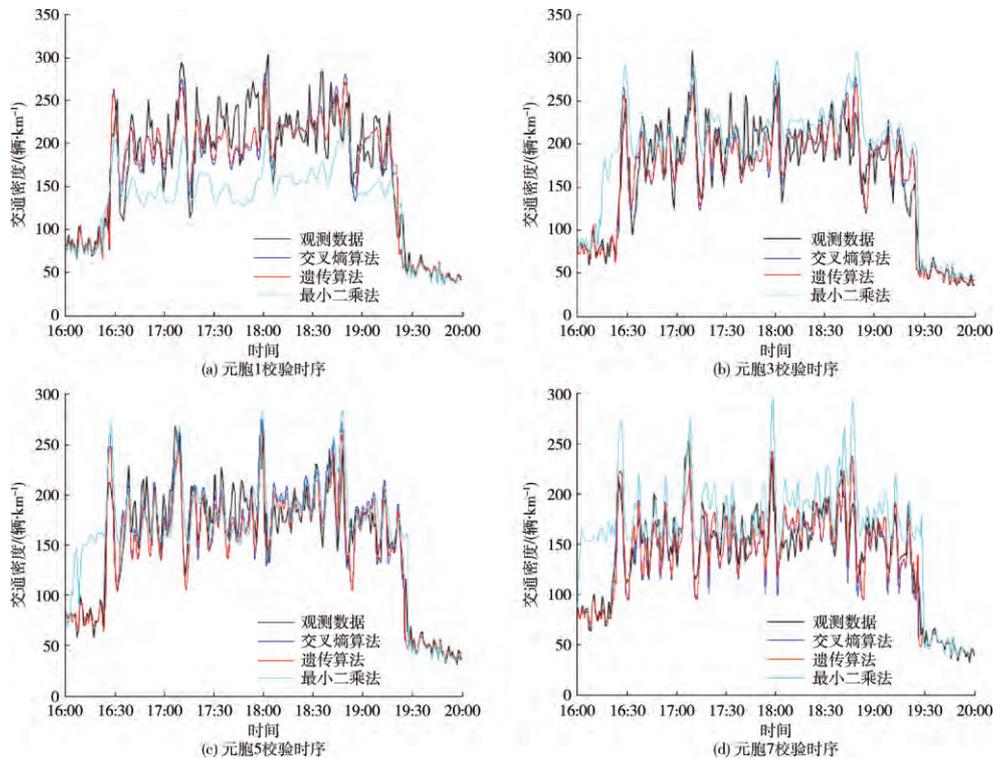


图 5 模型部分元胞校验时序

4 结 语

针对模型标定存在的低效率与低准确度的问题,笔者提出了采用交叉熵算法求解元胞传输模型参数标定问题的方法。通过将优化问题转换为小概率事件估计问题,交叉熵算法采用蒙特卡罗与重要样本策略逐步接近最优值,以概率密度函数的收敛实现模型参数的确定。实例验证结果表明,交叉熵算法有着优秀的搜索全局最优值的性能。对比遗传算法,交叉熵算法能以更少的模型仿真次数实现更精确的模型标定,体现了交叉熵算法逼近最优参数的概率密度函数优异的收敛性质,且交叉熵算法的仿真数据与观测数据更为吻合。

后续研究将考虑利用交叉熵算法进行随机性标定问题的求解,并通过不同的指标来验证交叉熵算法的优异收敛性质。此外,针对样本标准差收敛过程中出现的重复样本,可以提出自适应采样策略以提高模型标定效率。

致谢:感谢英国伦敦大学学院的 Andy H. F. Chow 教授为本文提供数据。

[参考文献] (References)

- [1] MUNOZ L, SUN Xiaotan, SUN Dengfeng, et al. Methodological calibration of the cell transmission model [C]// Proceedings of the 2004 American Control Conference. Boston, MA: IEEE, 2004: 798-803.
- [2] CIUFFO B, PUNZO V. 'No Free Lunch' theorems applied to the calibration of traffic simulation models [J]. IEEE Transactions Intelligent Transportation System, 2014, 15: 553-562.
- [3] 王殿海, 陶鹏飞, 金盛, 等. 跟驰模型参数标定及验证方法[J]. 吉林大学学报(工学版), 2011, 41: 59-65. WANG Dianhai, TAO Pengfei, JIN Sheng, et al. Method of calibrating and validating car-following model [J]. Journal of Jilin University (Engineering and Technology Edition), 2011, 41: 59-65. (in Chinese)
- [4] NGODUY D, MAHER M. Calibration of second order traffic models using continuous cross entropy method [J]. Transportation Research: Part C, 2012, 24: 102-121.
- [5] MAHER M, LIU R, NGODUY D. Signal optimisation using the cross entropy method [J]. Transportation Research: Part C, 2013, 27: 76-88.
- [6] ZHONG R, FU K, SUMALEE A, et al. A cross-en-

- tropy method and probabilistic sensitivity analysis framework for calibrating microscopic traffic models [J]. *Transportation Research: Part C*, 2016, 63: 147-169.
- [7] KROESE D, POROTSKY S, RUBINSTEIN R. The cross-entropy method for continuous multi-extremal optimization [J]. *Methodology & Computing in Applied Probability*, 2006, 8: 383-407.
- [8] KROESE D, RUBINSTEIN R, GLYNN P. The cross-entropy method for estimation [M]// KROESE D, RUBINSTEIN R, GLYNN P. *Handbook of Statistics*. Chennai: Elsevier B. V. , 2013: 19-34.
- [9] de BOER P, KROESE D, MANNOR S, et al. A tutorial on the cross-entropy method [J]. *Annals of Operations Research*, 2005, 134: 19-67.
- [10] DAGANZO C F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory [J]. *Transportation Research Part B: Methodological*, 1994, 28 (4): 269-287.
- [11] CHOW A, DADOK V, DERVISOGLU G, et al. Tools for operational planning of transportation networks [C]// ASME Dynamic Systems and Control Conference. Ann Arbor Michigan: American Society of Mechanical Engineers, 2008: 1035-1042.
- [12] PAN T, SUMALEE A, ZHONG R, et al. Short-term traffic state prediction based on temporal-spatial correlation [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(3): 1242-1254.
- [13] MUNOZ L M, SUN X T, HOROWITZ R. Traffic density estimation with the cell transmission model [C]// Proceedings of the 2003 American Control Conference. Denver, Colorado, USA: American Control Conference, 2003: 3750-3755.
- [14] ZHONG R, CHEN C, CHOW A, et al. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models [J]. *Journal of Advanced Transportation*, 2016, 50(3): 363-385.
- [15] CHOW, A, LI Y, GKIOTSALITIS K. Specifications of fundamental diagrams for dynamic traffic modeling [J]. *Journal of transportation engineering*, 2015, 141 (9): 04015015.

(上接第 1589 页)

- [10] 王保云, 张荣, 袁圆, 等. 可见光遥感图像中舰船目标检测的多阶阈值分割方法 [J]. *中国科学技术大学学报*, 2011, 41(4): 293-298.
WANG Baoyun, ZHANG Rong, YUAN Yuan, et al. A new multi-level threshold segmentation method for ship targets detection in optical remote sensing images [J]. *Journal of University of Science and Technology of China*, 2011, 41(4): 293-298. (in Chinese)
- [11] CORBANE C, NAJMAN L, PECOUL E, et al. A complete processing chain for ship detection using optical satellite imagery [J]. *International Journal of Remote Sensing*, 2010, 31(22): 5837-5854.
- [12] INGLADA J. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2007, 62(3): 236-248.
- [13] CHENG Gong, HAN Junwei, GUO Lei, et al. Object detection in remote sensing imagery using a discriminatively trained mixture model [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 85: 32-43.
- [14] HAN Junwei, ZHANG Dingwen, CHENG Gong, et al. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(6): 3325-3337.
- [15] LIN Min, CHEN Qiang, YAN Shuicheng. Network in network [EB/OL]. [2017.05.16]. <https://arxiv.org/abs/1312.4400>.