

文章编号: 1001-2486(2011)01-0120-05

基于参数稳态判定的两阶段序贯仿真分析策略*

朱一凡, 赵新, 李群

(国防科技大学 信息系统与管理学院, 湖南 长沙 410073)

摘要:两阶段序贯分析方法是仿真(模拟)实验研究中的一种常用方法,它虽然可以通过控制实验给出符合精度要求的结果数据,但以其现有形式有时却无法避免冗余计算的存在和实验过早终止的可能。在回顾了现有终态仿真解算终止策略特点的基础上,分析了所需样本总数的估计值与当前可用样本间的变化关系,提出了一种基于参数稳态判定的改进两阶段序贯方法。该方法增加了策略实施的主动性,可以在所需采样样本总数量 n 值较大的情况下,有效改善两阶段方法中对 n 值的初始估计,从而改进计算效率并能保证解算结果的准确性。

关键词: 决策分析; 建模仿真; 序贯分析; 参数稳态判定; 敏捷性; 有效性

中图分类号: TP391.9 **文献标识码:** A

Two-stage Sequential Stopping Procedure with Parameter Steady-state Estimate for Terminating Simulation

ZHU Yi-fan, ZHAO Xin, LI Qun

(College of Information Systems and Management, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: Two-stage sequential stopping procedure is used for terminating simulation analysis. It can control precision of result data that is required, but cannot avoid redundancy computation and early stopping. Firstly, the existing stopping procedures of terminating simulation were reviewed; then, the relationship between the quantity estimating of the samples total needed and the current amount of the samples usable was analyzed; finally, a modified method based on parameter steady-state estimate was proposed. This method can improve estimating of n (samples total needed) if the n is biggish, and ensure the validity of the simulation result. It can increase go-aheadism for the procedure actualizing.

Key words: decision-making; modeling and simulation; sequential analysis; parameter steady-state estimate; agility; validity

建模仿真是复杂问题决策分析的重要支撑技术,它为高层问题分析提供了定量化的描述和分析依据。决策者可以通过调整策略以及模型形式等方式,灵活地对问题进行全面探索和深入分析。但随着问题复杂程度的增加,不断增大的模型规模加深了模型仿真(模拟)解算代价与高层决策分析敏捷性需求之间的矛盾。作为一种比较成熟的方法,两阶段序贯仿真分析在工程实践中得到了广泛的应用,但它并不能彻底解决结果分析精确性和敏捷性之间的矛盾,而相关改进方法的研究^[1-5]多以改善对结果精度的控制为目的,且大多属于一种“被动”的方法,难以兼顾对冗余计算的改善。

仿真实验解算终止策略的选取决定了解算执行效率和最终结果精度^[6-8],而以置信区间分析为基础的所需采样样本总次数 n 值的评估计算,是其研究中的核心问题之一。对初始 n 值估计方法的改进,可以改善序贯方法的精度或效率,继而成为模型整体解算效率提高的有效途径。本文

基于以往成果,在分析了现有终态仿真解算终止策略特点的基础上,针对 n 值计算随可用样本演进变化的特性,综合考虑策略的敏捷性(即尽可能地避免不必要的采样解算)和有效性(即在两阶段设计中对所需采样样本总数量 n 值评估控制的可靠性)要求,提出了一种基于参数(即 n 的估计值)稳态判定的改进方法。该方法增加了策略的主动性,可以在所需样本值较大的情况下,有效改善两阶段方法中对 n 值的初始估计效果,使其能够支持复杂问题仿真分析模型的解算,以满足高层决策分析的特殊要求。

1 现有仿真终止策略概述

仿真解算终止方法的核心,一般是计算出解算所需的重复采样次数 n 值,这也是仿真终止分析(输出分析)中最关键的问题。基于传统统计学的置信区间,可以采用解析或序贯方法确定 n 值,而其具体的实现方式和策略一般有解析方法

* 收稿日期: 2010-04-12

基金项目: 国家自然科学基金资助项目(60974073, 60974074); 国家部委基金资助项目(9140C640505)

作者简介: 朱一凡(1963—), 男, 教授, 博士。

(即利用 $n \geq \left(\frac{t_{n-1,1-\alpha/2} S(n)}{\epsilon} \right)^2$ 确定 n 的估计值 n_0 ^[8-9])、两阶段序贯方法^[1-3] 及一般序贯方

法^[3,6-8] 等 3 类形式,如图 1 所示。

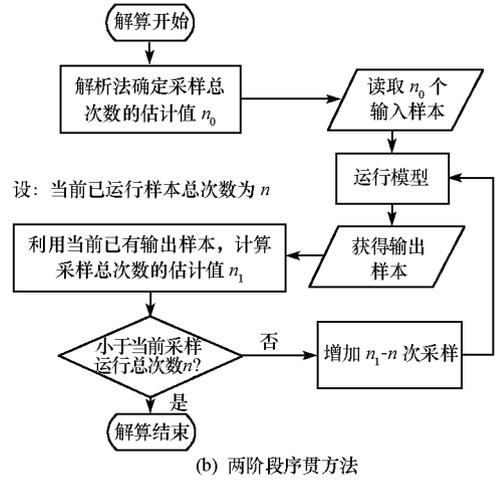
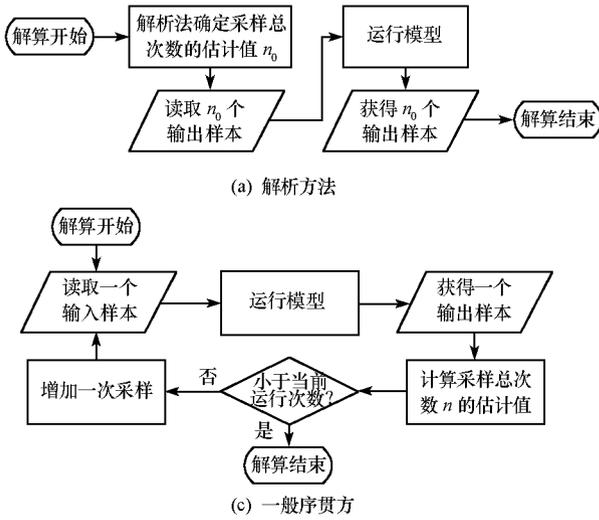


图 1 3 类方法流程示意图
Fig.1 Processes of three methods

以上 3 类方法各有优缺点,如表 1 所示。第一类方法通过解析方式进行 n 值估计并以此为依据进行采样和解算控制,它的优点是思路简单,缺点是难以准确控制采样总数和最终解算结果的精度,因此一般适用于采样总数较小、模型解算代价较小、模型中不确定性因素变化较简单的情况;第二类方法在前者的基础上对结果解算精度进行了有限地分析,可以有效控制解算结果精度,因此增强了其在模型中不确定性因素变化较复杂情况下的适用程度;第三类方法是在每次采样解算后都进行解算精度评估和 n 值估计,优点是设计思路简单,可以比较准确地控制采样数量和解算结果精度,缺点是对 n 值的估计较大程度上增加了整体仿真解算的负担,因此一般适用于 n 值估计计算成本相对很小或实际所需采样总数(n 值)不大的情况。

程一般可以被划分为“确定初始的 n 值”和“进行补充采样需求判定”两大部分功能,不同的只是在其具体实现形式和标准:第一类方法中省略了“进行补充采样需求判定”功能;第三类方法则将两部分综合到一起进行;而第二类方法的实质就是基于这种两大部分功能所给出的两阶段设计。这两大部分功能都是基于 n 值计算而实现的。然而,一般方法中假定 n 次独立重复运行结果总体方差 σ^2 的估计值 $S^2(n)$ 随着运行次数的增加无显著变化,因此才对 n 值进行估算。但这种假定在当前样本数并不是很大的情况下并不成立,它反映在基于这种关系所进行的 n 值估算过程中,就使得当前已有解算样本数与 n 值估计之间有如图 2 所示的性质和变化关系。

表 1 确定仿真解算重复采样次数的 3 类方式及其特点
Tab.1 Three kinds of methods for calculating simulation resamplings

方法适用性	解算准确性		模型单次调用成本		所需的最终样本数	
	控制需求	需求	高	低	大	小
第一类方法	○	◎	◎	◎	○	◎
第二类方法	◎	●	◎	◎	◎	●
第三类方法	●	◎	●	◎	●	◎

从图 2 中可以看出,无论初始样本如何,随着采样解算的实际运行,对 n 值估计将呈震荡变化趋势,并逐渐趋于一个比较稳定的值;但在已有解算样本数较小时, n 值的估计值 n_0 一般偏差较大且不稳定^[5]。这种不精确性对“确定初始的 n 值”和“进行补充采样需求判定”两大部分功能都可能产生影响。具体而言,前文所述的第一、二类方法在大多数情况下会得到较实际 n 值偏差较大的初始预测值;第三类方法在所需样本总数较小时对 n 值的评估也并不准确;第二类方法对补充采样需求的判定可以在一定程度上避免早期预测 n 值偏小的错误,但却无法避免预测 n 值偏大时的额外偏差。

图示说明:好●,一般◎,差○。

2 技术问题与研究设想

2.1 策略实现的问题

从功能结构的角度讲,以上几类方法属于基于 n 值计算所进行的仿真解算终止判定,而其过

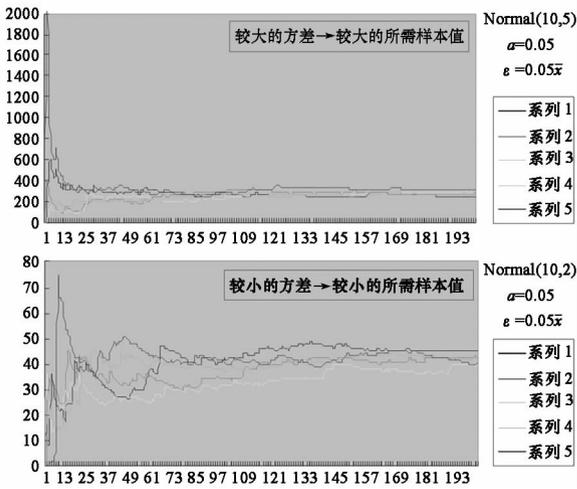


图2 采样实际运行次数与所需采样总次数估计值间的关系变化曲线^①

Fig.2 The curve of total sampling time estimated vs. amount of samples already produced

2.2 现有解决方法

针对上述问题,现有可行的解决方法大致有图形方法^[3]和工程化方法^[5,10]两种。前者可以以直观的方式将 n 值估算结果实时地显示出来,使用户可直观地了解该估算值的变化趋势,获得对结果的合理认识。后者一般则是通过给出可以满足工程实现要求的设计,以策略控制的形式尽可能保持结果分析的精确性。如文献[5]给出的“look-ahead”工程化方法,在工程上有着比较强的易用性和一定的有效性。然而,这两种方法都是对结果精度进行控制,并不能去除因 n 的初始估计值过大带来的额外解算代价。而且,图形方法虽然是通过定量计算给出,但其中有关 n 值确认等过程却需要人工参与,不便于解算评估的自动进行;而现有的工程化方法一般可以改进结果精度,但却处于“进行补充采样需求判定”阶段,且属于一种被动形式,仍然有过早终止的可能。

2.3 设计原理

图2不仅显示了 n 值估计的变化过程,揭示了传统的仿真终止策略分析方法本身存在着固有的缺陷(即在“确定初始的 n 值”阶段往往过早地确定一个不太准确的估计值 n_0),还给我们带来了进一步的启示:终态仿真采样次数估计值的计算,也可以是寻找趋向于某一近似稳定值的过程,可以通过相应的分析方法,在合理的计算代价内,得到较为准确的 n 值初始估计。这与一般的稳态分析相似,面临有类似“模型准备(warm-up)”的问题,但后者是要去掉“模型准备”时期计算出来的样本,以消除其对稳态均值拟合的影响,而前者

则是要避免因评估不稳定而在“模型准备”时期过早地终止对 n 值的预测。由于“确定初始的 n 值”功能过程正处于整个解算策略框架的前期,如果可以使用稳态判定方法来解决“确定初始的 n 值”问题,那么相比较而言会更加具有说服力,也将使整体策略的设计具有更多的主动性和可信性。

3 改进序贯仿真策略

3.1 框架改进

进行策略设计的目的,一是改进解算效率,避免不必要的采样计算,二是提高解算精度,减少犯两类错误的概率。为了能够在解算效率和解算结果精度间找到良好的折中,有必要使实际所进行的采样数量与最终分析出的所需采样数近似相等,且要避免过多的 n 值估计计算。这种控制上的改进,在“确定初始的 n 值”阶段进行相对来说更加有效。因此,可以给出如图3所示的策略设计。

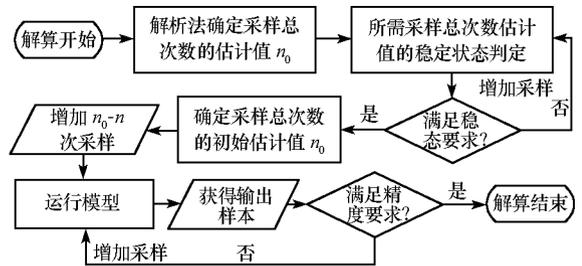


图3 基于参数稳态判定的两阶段序贯仿真策略整体框架

Fig.3 The frame of improved sequential stopping procedure

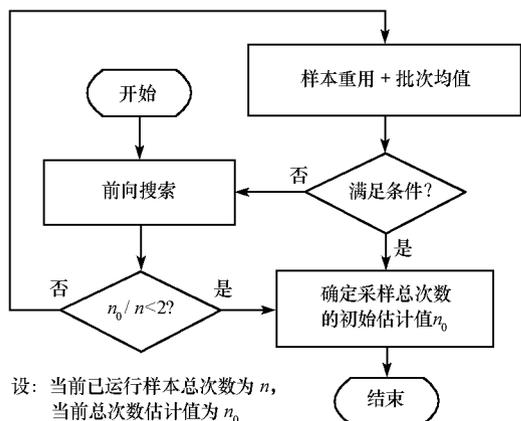
该策略框架在以往方法的基础上,添加了“采样总次数估计值的稳定状态判定”的内容,使其具有了一定的适应性,既可保持现有策略过程的有效进行,又可为解算效率和精度控制提供可靠性。

3.2 采样次数估计的稳态判定

给出重复采样次数估计值的稳态分析目的,是找到该值趋于稳态时的时刻。其思路是给出一个可以接受的误差,在此基础上计算估计值趋于稳态的时刻,并取该时刻的估计值,作为仿真解算所需重复采样运行次数的估计值。该策略根据实际采样数及所需总次数的大小和比例关系,制定了相对鲁棒的控制过程,其实施过程如图4所示。其中, n 表示当前已经进行的采样次数, n_0 表示采用当前样本估算的所需采样总次数, n_0 与 n 的

① 仿真输出结果正态总体假设下的人工数据

比值可以反映出所需采样总次数的估算风险大小,而 n 值则反映了估算所需计算成本的大小,它们是控制各流程执行的条件。



设: 当前已运行样本总次数为 n , 当前总次数估计值为 n_0

图 4 所需采样总次数估计值的稳态判定流程

Fig.4 The process to judge steady-state of the parameter estimation

①“前向搜索”的执行条件是:当前已经获得的采样样本较小,且采样总次数的估算风险适中或较小(如 $n_0/n < 2$,见图 4);

②“样本重用 + 批次均值”的执行条件是:当前已获得的样本较小(但不能太小,至少大于 15(保证其执行,参见下文)),或采样总次数的估算风险适中(如 $n_0/n > 2$,见图 4)。

(1)“前向搜索”序贯方法

执行算法流程如图 5 所示,通过改变附加采样判断标准和执行过程。该方法改进了对 n 值的初始估计,使其具有一定的稳态特质,可以满足大部分应用下 n 值估计的稳定性要求,具有在较小样本下的可用性 & 便于工程化实现等特点。

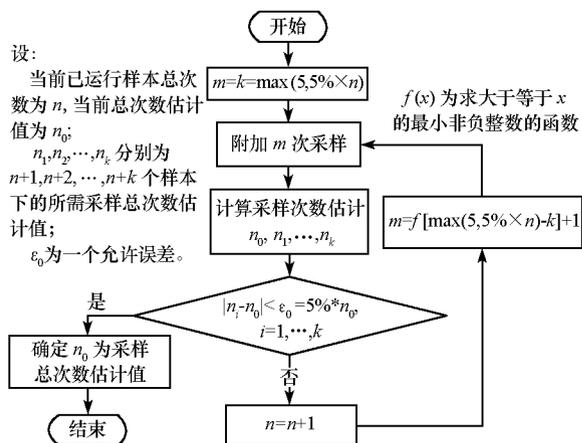


图 5 前向搜索逻辑流程

Fig.5 The process of forwards searching

但是,参考图 2 中给出的评估曲线,各种初始状态(随机种子)下的评估值变化过程存在较大差异,而 k 的取值又不宜过大,所以使用“前向搜

索”方法还是有使 n 值的估计过早结束的风险,进而使所得到的估计值过大或过小。因此,在使用该方法确定 n 的一个可以接受的待选估计值 n_0 后,还可以给出某种策略,使得 n_0 值更进一步接近 n 的理论值,进一步提高 n 值估计的可靠性。

(2)“样本重用 + 批次均值”稳态判定工程化方法

“样本重用 + 批次均值”方法设计的目的,是进一步判定当前样本数据对 n 值的估计是否进入一个相对稳定状态,而此时可以利用的样本数一般较少。因此,这里可以考虑对单次解算样本进行随机化处理后进行重用,并以此作为稳态判定的对比依据,而为了在这种样本重用处理基础上进行稳态判定中体现出差异性,有必要对这些数据进行批次均值处理^[10-11]。该方法是对“前向搜索”序贯方法的补充,是继其被动的稳态判定后,给出的相对主动的稳态判定方法。

该方法的具体过程是通过对样本的重用,对估计值 n 的批次均值处理,再进行方差上的判定。虽然批次均值处理有将评估差异淹没的风险,但是足够数量的多个序列样本重用可以将这种风险降至最低。

图 6 给出了 $m = 4$ 个重用序列样本下,使用批次处理后数据对 n 的评估值变化(基于图 2 中 Normal(10,5)的数据系列 1 所进行的处理操作)。该方法的优点是:样本重用过程减少了采样解算的计算量,批次均值处理减少了 n 值估计的偏差且为方差判定提供了差别基础,最后的方差判定处理为进一步的估计过程提供终止条件,该方法可以到达工程应用的要求,缺点是:批均值处理中的批次 n 值样本数 s 、方差处理的标准 ϵ 等不易确定,解算效果需要大量数据实验的进一步验证。

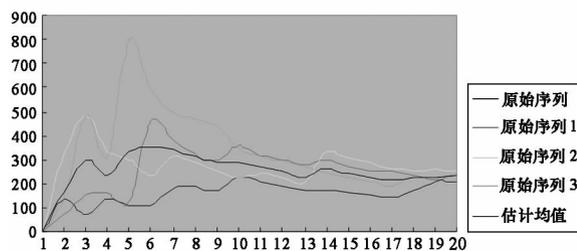


图 6 “样本重用 + 批次处理”下 n 值估计的变化曲线

Fig.6 The curve of the parameter estimation by samples reusing & batched processing

3.3 数据实验

表 2 给出了基于人工数据(即对某正态分布取随机样本)的实验比较,所使用的样本数据与图

2相同。其中,“解析法预测值”表示只使用解析方法的所需样本预测值;“look-ahead下的实际采样次数”表示在解析法预测值基础上,运用文献[5]提出的“look-ahead”方法得到的实际采样次数;“稳态判定后的预测值”表示在解析法预测值基础上,使用本文提出的采样次数估计的稳态判定流程及其中相应方法所得到的所需样本的稳定预测值,其中括号中是当前已经运行的采样次数;“采样运行后的再估计值”表示运行“稳态判定后的预测值”的个数个样本后所得到的进一步估计值;“结果精度判定后的实际采样次数”表示使用

本文所提出的补充采样判定流程后,所得到的最终采样次数;“理论采样次数”表示所需采样次数的理论值,是使用人工数据下预设的均值和方差计算出来的。

从表2中实验数据知,结果数据方差的大小决定了所需采样样本数量的多少,所提出的方法虽然不能显著提高结果数据方差较小时的解算效率,但可保证解算精度;而在所需采样样本数较大时,该方法却可以获得与理论采样次数更加接近的实际采样次数,进而给出解算效率和精度的更佳折中。

表2 人工数据实验

Tab.2 The experiment under artificial data

正态分布 及置信标准	人工数据 批次	解析法 预测值	“look-ahead” 方法下的实际 采样次数	稳态判定后 的预测值(已 有采样次数)	采样运行后 的再估计值	结果精度判定 后的实际 采样次数	理论采样 次数
<i>Normal</i> (10,5) $\alpha = 0.05$ $\epsilon = 0.05\bar{x}$	实验 I	365	365	331(47)	279	331	289
	实验 II	101	279	301(83)	309	307	
	实验 III	101	239	253(44)	309	273	
	实验 IV	159	270	243(62)	264	271	
	实验 V	405	405	328(23)	320	328	
<i>Normal</i> (10,2) $\alpha = 0.05$ $\epsilon = 0.05\bar{x}$	实验 I	18	40	42(67)	42	67	47
	实验 II	26	36	33(75)	33	75	
	实验 III	31	31	31(58)	31	58	
	实验 IV	24	44	41(47)	42	47	
	实验 V	54	54	44(52)	44	52	

4 思考与展望

描述复杂问题的模型一般具有较大的规模以及更不稳定的解算结果数据,前者带来了较大的模型采样解算的单次调用负担,后者则带来了较大的采样样本总量。这种情况下,使用本文所给出的基于参数稳态判定的两阶段序贯方法会获得较好的运行结果。该方法以相对较小的额外计算代价(参数稳态判定)改善大型模型的仿真解算效率,并为解算结果的准确度提供更可靠的保障。

当然,该方法的相关参数值(如 n_0/n 和 k 等)根据问题的不同存在变化,在本文的讨论中,它们的设定还处于初步阶段,还需要在大量数据实验基础上进一步的调整。另外,文中给出的讨论建立在仿真输出正态总体假设下的人工数据基础上,在进一步的研究中可以通过不同问题及其大量工程化数据对其进行深入验证与分析设计。

参考文献:

- [1] 张金槐,等. Bayes 试验分析方法[M]. 长沙: 国防科技大学出版社, 2007.
- [2] Marvin K N. Two-stage Stopping Procedures Based on Standardized

- Time Series [J]. Management Science, 1994, 40: 1189 - 1206.
- [3] Robinson S. Simulation: The Practice of Model Development and Use[R]. John Wiley & Sons Ltd., 2004.
- [4] Peter F, Warren B P. The Knowledge-gradient Stopping Rule for Ranking and Selection [C]//Proceedings of the 2008 Winter Simulation Conference, 2008: 305 - 312.
- [5] Hoag K, Robinson S, Davies R. Automating Des Output Analysis: How Many Replications to Run [C]//Proceedings of the 2007 Winter Simulation Conference, 2007: 505 - 512.
- [6] Marvin K N. Statistical Analysis of Simulation Output [C]//Proceedings of the 2008 Winter Simulation Conference, 2008: 62 - 72.
- [7] Averill M L. Statistical Analysis of Simulation Output Data: The Practical State of The Art [C]//Proceedings of the 2007 Winter Simulation Conference, 2007: 77 - 83.
- [8] 王维平,等. 离散事件系统建模与仿真[M]. 长沙: 国防科技大学出版社, 2006.
- [9] Morgan M, Henrion M. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis [M]. Cambridge University Press, Cambridge, United Kingdom, 1990.
- [10] Stephen E C, Jürgen B, Christian S. New Greedy Myopic and Existing Asymptotic Sequential Selection Procedures [C]//Proceedings of the 2007 Winter Simulation Conference, 2007: 289 - 296.
- [11] Emily K L, James R W. SBatch: A Spaced Batch Means Procedures for Simulation Analysis [C]//Proceedings of the 2007 Winter Simulation Conference, 2007: 463 - 471.