

1-2-2019

Optimization via Simulation Based on Neural Network

Shihui Wu

1. Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710051, China; ;

Xiaodong Liu

1. Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710051, China; ;

Shao Yue

2. The Equipment Department, PLAAF, Beijing 100036, China;

Zhang Fa

1. Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710051, China; ;

See next page for additional authors

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Optimization via Simulation Based on Neural Network

Abstract

Abstract: To improve the efficiency of optimization via simulation (OvS), an OvS method based on neural network is proposed. Taking advantage of the approximation ability of neural network to nonlinear input-output relationship, *neural network's outputs are used as substitutes for simulation results to reduce the required simulation runs. Samples are generated by simulation according to the three proposed samples selection methods. Owing to its advantages on learning speed, network stability and parameters selection, generalized regression neural network (GRNN) is adopted to train the samples. The trained GRNN forms a regression surface that represents the relationship between simulation inputs and outputs, which makes it feasible to use GRNN output as substitutes for simulation runs. Optimization algorithms are applied to search for the best solution on the regression surface.* Experiments are carried out with some typical test functions, and the feasibility and effectiveness of our method are demonstrated.

Keywords

neural network, optimization via simulation, regression surface, samples selection

Authors

Shihui Wu, Xiaodong Liu, Shao Yue, Zhang Fa, and Minxiang Yang

Recommended Citation

Wu Shihui, Liu Xiaodong, Shao Yue, Zhang Fa, Yang Minxiang. Optimization via Simulation Based on Neural Network[J]. Journal of System Simulation, 2018, 30(1): 36-44.

一种基于神经网络的仿真优化方法

吴诗辉¹, 刘晓东¹, 邵悦², 张发¹, 杨闽湘¹

(1. 空军工程大学装备管理与无人机工程学院, 陕西 西安 710051; 2. 空军装备部外场部, 北京 100036)

摘要: 为提高仿真优化问题求解效率, 提出了一种基于神经网络的仿真优化方法。利用神经网络对非线性输入输出关系的逼近能力, 由神经网络输出值代替仿真结果以减少所需仿真次数。按照提出的3种样本选择策略, 由仿真模型产生一定数量的样本, 借助广义回归神经网络在学习速度、网络稳定性、参数选取方面的独特优势, 对样本进行训练, 生成能够反映仿真模型输入输出关系的回归曲面, 以实现用神经网络输出值代替仿真结果, 利用优化算法对回归曲面进行寻优。通过对典型测试函数进行实验, 证明了方法的可行性和有效性。

关键词: 神经网络; 仿真优化; 回归曲面; 样本选择

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2018) 01-0036-09

DOI: 10.16182/j.issn1004731x.joss.201801005

Optimization via Simulation Based on Neural Network

Wu Shihui¹, Liu Xiaodong¹, Shao Yue², Zhang Fa¹, Yang Minxiang¹

(1. Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710051, China;

2. The Equipment Department, PLAAF, Beijing 100036, China)

Abstract: To improve the efficiency of optimization via simulation (OvS), an OvS method based on neural network is proposed. Taking advantage of the approximation ability of neural network to nonlinear input-output relationship, *neural network's outputs are used as substitutes for simulation results to reduce the required simulation runs. Samples are generated by simulation according to the three proposed samples selection methods. Owing to its advantages on learning speed, network stability and parameters selection, generalized regression neural network (GRNN) is adopted to train the samples. The trained GRNN forms a regression surface that represents the relationship between simulation inputs and outputs, which makes it feasible to use GRNN output as substitutes for simulation runs. Optimization algorithms are applied to search for the best solution on the regression surface.* Experiments are carried out with some typical test functions, and the feasibility and effectiveness of our method are demonstrated.

Keywords: neural network; optimization via simulation; regression surface; samples selection

引言

对于复杂系统而言, 广泛存在的大量随机因

素、系统内各实体之间的相互影响等特征, 都决定了对复杂系统的描述难以用解析方法实现。仿真的最大优势就是能够很好地阐释带有不确定性和随机性的问题, 因此, 通常采用仿真建模和大量的仿真实验来描述复杂系统。当我们需要对复杂系统的输入参数进行优化设计, 以得到最优或满意的输出指标时, 就需要用到仿真优化(OvS, Optimization



收稿日期: 2015-11-30 修回日期: 2016-01-31;
基金项目: 国家自然科学基金(71571109, 61601501);
作者简介: 吴诗辉(1982-), 男, 湖北武汉, 博士, 讲师, 研究方向为装备论证; 刘晓东(1966-), 男, 陕西户县, 博士, 教授, 博导, 研究方向为装备经济管理。

<http://www.china-simulation.com>

via Simulation)。OvS 问题不同于确定性优化，因为它的目标函数是由仿真模拟出来的统计值，由于对目标函数的评估带有随机性，需要仿真多次，计算量非常大。随着计算机性能的不提高，OvS 在许多领域得到了广泛地应用，比如交通控制^[1]、铁路集装箱的装卸策略^[2]、多级库存策略^[3]、航迹规划^[4]等。由于这些具体工程问题本身存在大量的随机性和复杂性，都是以仿真结果作为策略优化选择的依据，通常需要用到 OvS。

目前，OvS 已成为国际上系统仿真和运筹学等领域共同关注的热点和前沿课题，每年的 Winter Simulation 会议都将其列为一个国际前沿课题来研讨。Hong 和 Nelson^[5]对 OvS 的已有成果进行了很好的综述。对于简单系统，策略的可取值空间小，常用的方法是枚举法，通过列出所有可能的策略组合并进行仿真，选出最佳的组合^[6]。然而，对于类似作战模拟、现实交通、大型库存网络等复杂系统，仿真一次时间往往需要分钟、小时甚至是天的数量级，要完成所有策略组合的仿真，所需时间几乎是天文数字，因此，大量学者针对复杂系统的仿真优化开展了研究^[7]。纵观各类 OvS 方法，大致可分为两种：一是以最小的优化次数收敛于局部或全局最优解，比如 COMPASS 法^[8]、黄金区域搜索法^[9]、基于梯度的方法^[10]、启发式算法^[4]等；二是通过减少每次仿真所需代价的方法^[2]，结合启发式算法来提高寻优速度。从这些方法看，都是以仿真输出结果作为目标函数或适应度函数，利用优化算法进行寻优。笔者通过研究神经网络对类似仿真这样的黑箱问题的输入输出关系的回归时，发现训练好的神经网络能够很好的实现对仿真结果的预测，因此，我们提出一种新的仿真优化方法，其核心是利用神经网络取代仿真模型，以解决仿真优化的瓶颈问题，我们称之为-基于神经网络的仿真优化方法。

1 基本原理

典型的传统仿真优化基本原理如图 1 所示^[11]，建立实际系统的仿真模型，优化算法产生系统性能

初始参数，输入仿真模型，将仿真模型的输出作为解的评价值返回优化算法；由优化算法进行评估，按照一定的搜索策略给出更优的性能参数，将此性能参数作为新的输入再进行仿真，评估结果，不断重复以上过程，直到得出优化参数。其实质是将仿真结果看作目标函数或适应度函数进行寻优。

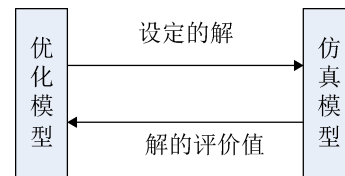


图 1 传统仿真优化的基本原理
Fig. 1 Basic principle of traditional OvS

基于神经网络的仿真优化基本原理如图 2 所示，利用仿真模型在允许的生成一定数量的样本，然后借助神经网络强大的预测能力，通过对样本的学习，实现由神经网络模型取代仿真模型对优化算法中的解进行评价。其实质是将神经网络模型看作目标函数或适应度函数进行寻优。这种思路能够有效解决传统仿真优化方法存在的仿真代价过高问题，更适用于求解复杂系统的仿真优化。

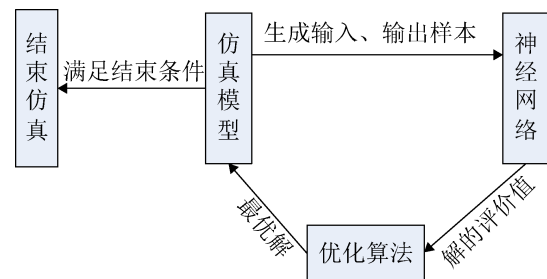


图 2 基于神经网络的仿真优化的基本原理
Fig.2 Basic principle of OvS method based on neural network

2 基于神经网络的仿真优化设计

2.1 神经网络选择

广义回归神经网络(GRNN)是美国学者 Donald F. Specht^[12]于 1991 年提出的一种神经网络，是建立在非参数估计基础上的非线性回归的径向基神经网络。实验表明^[13-14]，GRNN 在逼近能力和学习

速度上较 RBF 网络和 BP 网络有更强的优势, 网络最后收敛于样本量积聚较多的优化回归面。GRNN 一个重要特点是在样本数据较少时, 预测效果也较好, 这是选用 GRNN 对仿真输出进行预测的重要原因, 因为只需较少的仿真样本, 就能够得到较好的预测精度^[14]。另外, 相比其他神经网络, GRNN 人为调节的参数只有 1 个, 即平滑因子 σ (smoothing parameter), 网络学习全部依赖所选样本, 该特点决定了网络能最大程度避免人为影响。

2.2 GRNN 原理

GRNN 由 4 层构成, 分别为输入层、模式层、求和层和输出层。假设网络输入为 $\mathbf{X}=(x_1, x_2, \dots, x_n)^T$, 其输出为 $\mathbf{Y}=(y_1, y_2, \dots, y_k)^T$ 。GRNN 的基本原理是, 输入为 \mathbf{X} 的条件下, \mathbf{Y} 的预测输出为:

$$\hat{\mathbf{Y}}(\mathbf{X}) = \frac{\sum_{i=1}^n \mathbf{Y}_i \exp\left(\frac{-D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(\frac{-D_i^2}{2\sigma^2}\right)} \quad (1)$$

式中: $D_i^2 = (\mathbf{X} - \mathbf{X}_i)^T (\mathbf{X} - \mathbf{X}_i)$, 表示输入变量 \mathbf{X} 与其对应的样本 \mathbf{X}_i 之间的 Euclid 距离的平方; \mathbf{X} 为网络输入变量; \mathbf{X}_i 为第 i 个输入层神经元对应的学习样本; σ 为平滑因子, 是 GRNN 网络需要确定的唯一参数。

2.3 GRNN 中平滑因子的确定

平滑因子是决定 GRNN 预测效果的唯一参数。 σ 必须大于 0, 当其趋向于 0 时, $\hat{\mathbf{Y}}(\mathbf{X})$ 和学习样本非常接近, 而非样本点的预测效果会非常差, 这说明网络的泛化能力差。 σ 越大, 则回归曲面越光滑, 当 σ 非常大的时候, $\hat{\mathbf{Y}}(\mathbf{X})$ 近似于所有样本的因变量均值。根据文献[15], 通常 σ 取值为 0.01~1 之间时能够得到理想的结果。

目前, 确定平滑因子的方法主要有两种: holdout 方法^[12]和遗传算法(GA)^[16]。

为求取平滑因子 σ , Specht^[12]提出了一种 holdout 方法, 即每次从学习样本中取出一个样本作为预测样本, 以其余样本作为学习样本组成一个 GRNN 网络, 然后用该网络来对取出的样本进行预测。通过重复以上步骤, 得到每次组成网络预测结

果与样本值之间的均方差(MSE, mean-squared error), 根据预测得到的最小均方差来选择最佳的平滑因子 σ 。Specht 认为, 由于 MSE 随 σ 变化的最小值很容易收敛, 因此, 不需要经过多次尝试就可以得到使得 MSE 最小的平滑因子, 我们可将这个最佳平衡因子用到 GRNN 网络。对神经网络拟合效果的评判通常可有两种方法, 即均方根误差法和相关系数法^[17], 相关系数又分为线性相关系数 (Pearson Linear Correlation Coefficient, CC) 和 Spearman 等级相关系数 (Spearman Rank Order Correlation Coefficient, SROCC)^[18]。本文选用线性相关系数法(CC)进行评判, 利用 holdout 法选择最佳的 σ 。

2.4 基于 GRNN 的仿真优化模型

2.4.1 理论基础

定理 1 如果一个模型 $\mathbf{y}=f(\mathbf{x})$ 等效于仿真模型, 那么, 对于仿真模型的参数优化问题, 就可以转化为对 $\mathbf{y}=f(\mathbf{x})$ 问题的优化。

证明: 如果把仿真模型的输入输出关系用 $\mathbf{Y}=\mathbf{g}(\mathbf{x})+\varepsilon$ 表示, 这里的 $\mathbf{g}(\mathbf{x})$ 为输入参数为 \mathbf{x} 时的仿真期望值, 显然它是一个黑箱模型, ε 表示由于随机因素引起的仿真输出的噪声, 其分布可能与 \mathbf{x} 有关, 令 $E(\varepsilon)=0$ 。

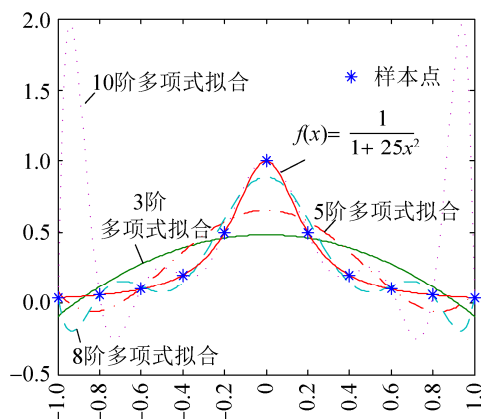
由于模型 $\mathbf{y}=f(\mathbf{x})$ 等效于仿真模型, 那么对于任意一个 \mathbf{x} , 必然有: $E(\mathbf{g}(\mathbf{x})+\varepsilon)=\mathbf{g}(\mathbf{x})=f(\mathbf{x})$ 。因此, 对于 $\mathbf{y}=f(\mathbf{x})$ 问题, 如果找到了最优解 $\mathbf{y}^*=f(\mathbf{x}^*)=\min f(\mathbf{x})$, 显然 $\mathbf{Y}^*=\mathbf{g}(\mathbf{x}^*)=\min \mathbf{g}(\mathbf{x})$ 。这意味着仿真模型也在 \mathbf{x}^* 处取得最优或满意解, 即该问题可转化为对 $\mathbf{y}=f(\mathbf{x})$ 问题的优化。

定理 2 如果一个模型 $\mathbf{y}=f(\mathbf{x})$, 其样本足够多, 且利用样本能够准确绘制出函数的回归曲面, 则必然能够利用回归曲面找到模型的最优解。

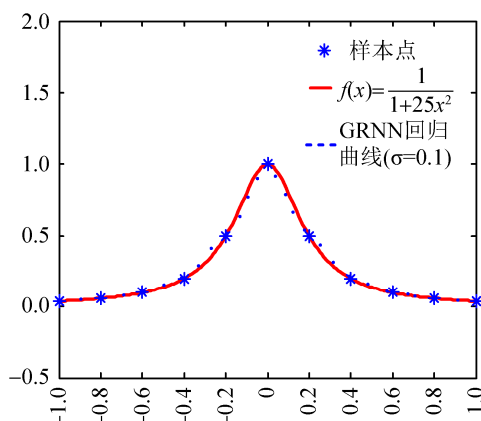
这个结论是显然的。目标函数的准确回归曲面上, 必然可以找到一个最优点, 对于求最小值问题, 最优点为曲面的最低谷, 对于求最大值问题, 最优点为曲面的最高峰。样本在足够多的情况下, 必然

能够反映原始曲面的变化规律, 则回归曲面可以代替原始曲面求解最优解。

用图 3 来解释曲面回归与最优值预测的关系。对于一个显式目标函数的最大值优化问题, 假设通过仿真计算, 在解空间上获得了一些样本点, 通过拟合, 对非样本点进行估计, 得到基于样本的目标函数曲面的回归。如果回归曲面的精度足够高, 就可以利用常用的优化算法, 如遗传算法、模拟退火等方法, 找到全局最优点。这就是利用基于样本的目标函数曲面回归来进行最优值预测的基本思路。



(a) 利用不同阶数多项式拟合的效果图



(b) 利用 GRNN 回归曲线拟合的效果图

图 3 利用多项式拟合和 GRNN 回归曲线的效果图

Fig. 3 Polynomial fitting results and the GRNN fitting results

例如, 对于 $f(x) = \frac{1}{1+25x^2}$, 其中 $-1 \leq x \leq 1$ 。

我们在 -1 和 1 之间均匀取 11 个点作为样本(含边界点), 利用 Matlab 的多项式拟合函数 `polyfit`, 对样本采取不同阶数(3, 5, 8, 10 阶)的多项式进行

拟合, 观察拟合的效果如图 3(a)所示。同时, 利用 GRNN 进行拟合, 观察拟合的效果如图 3(b)所示。显然 GRNN 具有非常高的拟合精度, 可替代原始函数进行优化。

2.4.2 仿真优化模型的求解步骤

模型的求解步骤如下:

Step 1: 在输入参数解空间, 随机均匀生成一定数量的输入参数样本;

Step 2: 将生成的输入参数样本代入仿真模型, 得到输出结果的期望值;

Step 3: 将仿真得到的输入输出样本数据代入到 GRNN 网络, 并利用 holdout 方法或 GA 得到最佳的平滑因子, 并将最佳平滑因子作为 GRNN 网络的平滑因子;

Step 4: 设置一组测试样本(在输入参数解空间内随机生成, 不能与输入参数样本相同), 判断 GRNN 网络的回归效果是否满足要求, 如果不满足, 则返回 Step 1, 利用仿真模型生成更多的样本, 并重新生成 GRNN 网络; 如果满足要求, 则进入下一步;

Step 5: 利用优化算法在输入参数解空间上进行寻优, 用训练好的 GRNN 网络的输出值评价解, 优化算法可以采用常用的确定问题优化算法, 如遗传算法、模拟退火、粒子群算法等;

Step 6: 将找到的最优解代入仿真模型, 判断最优解的 GRNN 预测值与仿真结果的误差, 如果误差过大(如超过 5%), 则返回 Step 1, 利用仿真模型在找到的最优解附近生成更多的样本, 以对该部分回归曲面进行精确化, 重新生成 GRNN 网络; 如果误差满足要求, 则认为找到最优解或满意解。

2.4.3 仿真优化模型的样本选择策略

从求解步骤可以看出, 这里的关键问题在于基于样本的目标函数曲面回归的准确性, 如果回归的误差过大, 则利用回归曲面优化很可能得出错误的结论。因此, 确保神经网络在目标函数曲面回归的准确性至关重要。那么, 如何以较少的样本保证回

归曲面的准确性呢? 样本选取多少合适, 选择哪些样本最具代表性? 这就是样本选择策略问题。笔者提出3种样本选择策略。

假设需要通过仿真优化对参数集 $\mathbf{X} = \{x_1, x_2, \dots, x_k\}$ 进行优化, 其中 $x_i \in [a_i, b_i]$ 。

(1) 等间隔样本采集法

对于 x_i , 以等间隔 Δ_i 获取样本, 则得到样本集为: $\{a_i, a_i + \Delta_i, a_i + 2\Delta_i, \dots, b_i\}$ 。这里的 Δ_i 可由式(2)确定:

$$\Delta_i = (b_i - a_i) / (n_{i\max} - 1) \quad (2)$$

式中: $n_{i\max}$ 表示 x_i 最多允许的样本数量。

由于总的最大样本数 N_{\max} 是各变量样本数的乘积, 故:

$$N_{\max} = \prod_{i=1}^k n_{i\max} \quad (3)$$

这里的 N_{\max} 可根据下式确定:

$$N_{\max} = T_{\max} / \bar{T}_{\text{sim}} \quad (4)$$

式中: T_{\max} 表示允许的最大仿真时间; \bar{T}_{sim} 表示平均完成一次仿真所需时间。

在没有任何可用信息的时候, 可将样本数均匀分配给各个变量, 则由式(3)得:

$$n_{i\max} = \sqrt[k]{N_{\max}}, \quad i = 1, 2, \dots, k \quad (5)$$

(2) 基于参数精度的样本采集法

通常, 对于实际的工程问题, 参数的选取都有精度要求, 比如有的参数精度要求高, 可达 10^{-4} , 有的参数只需要整数即可。假设 x_i 的精度要求为 ε , 则可考虑取 $\Delta_i = 5\varepsilon \sim 10\varepsilon$ 。比如, x_i 的精度要求为 0.01, 则可以取 $\Delta_i = 0.1$ 。具体取何值, 可在允许的最大仿真时间 T_{\max} 内, 选取尽可能小的 Δ_i 。

(3) 回归曲面的样本增补法

GRNN 网络通常收敛于样本量积聚较多的回归面, 如果测试样本和学习样本在最优解周围都数量偏少, 则最优解附近的回归面将难以得到准确反映, 从而影响对最优解的寻优。为此, 提出回归曲面的样本增补法。该方法类似于 Google 地图的定位, 首先大面积的低分辨率拍照, 对于重点地区则增加照片的分辨率和精细程度。具体步骤如下:

1) 按照等间隔样本采集法, 允许的时间范围内在参数的取值区间获取一定数量的样本 Φ_1 ;

2) 利用 GRNN 神经网络对回归曲面进行绘制, 执行 2.4.2 优化算法的 Step3~5, 快速找出可能的最优解所在曲面区间, 假设该区间为 $[a_i + p\Delta_i, a_i + q\Delta_i]$, 在该区间段再次采取等间隔样本采集法, 选取少量的样本 Φ_2 , 以实现该区间段曲面的精细化回归;

3) 将样本集 Φ_1 和 Φ_2 合并为一个样本集, 再次执行 2.4.2 优化算法的 Step3~5(注意, 执行 Step 4 时的测试样本必须选取 1 个最优解可能区间上的点), 从而得出更加精确的解集。

3 仿真实验与分析

为了验证基于神经网络的仿真优化方法的有效性, 我们选取典型测试函数(Levy No.5 Function)的最小值寻优问题进行试验。

$$f(\mathbf{X}) =$$

$$\sum_{i=1}^5 [i \cos((i-1)x_1 + i)] \sum_{j=1}^5 [j \cos((j+1)x_2 + j)] + (x_1 + 1.42513)^2 + (x_2 + 0.80032)^2$$

其中, $-5 \leq x_1, x_2 \leq 5$ 。有 760 个局部最小和 1 个全局最小点, 其中最小点为 $\mathbf{X}^* = (-1.368, -1.424, 8)$, $f(\mathbf{X}^*) = -176.1375$ 。

尽管这是一个确定性问题的寻优, 但是我们把它视为一个黑箱模型, 仍然利用 GRNN 网络来对函数产生的样本进行回归, 如果找到最优解, 则在一定程度上也验证了本文方法的正确性。其中, GRNN 网络利用 Matlab 中的 newgrnn 函数构建。

(1) 基于 GRNN 神经网络的仿真优化

①首先在解空间上设置样本点, 假设 x_1 每隔 0.05 取一个样本, x_2 每隔 0.1 取一个样本, 如 $x_1 = \{-5, -4.95, -4.9, \dots, 4.95, 5\}$, $x_2 = \{-5, -4.9, -4.8, \dots, 4.9, 5\}$, 则 x_1 和 x_2 两两组合得到 20301 个样本点(见表 1 中基于参数精度的样本采集法)。随机在解空间上取 10 个点作为测试点, 要求测试点与样本点不同。

表 1 不同样本选择策略对 Levy No.5 函数的优化
Tab. 1 Optimization of Levy No.5 function using different samples selection methods

| 采样策略 | 寻优时间/s | 样本间隔及数量 | σ | 本文方法 |
|--------------------------|--|---|----------|--|
| 等间隔 样本采 集法 | $N_{\max}=500$ | $\Delta=\{0.468\ 1, 0.468\ 1\}$ | 0.27 | $\mathbf{X}_{\min}=(-0.314\ 8, -0.780\ 8)$ GRNN's output = -80.076 4 $f(\mathbf{X}_{\min})=-142.522\ 8$ (陷入局部最小) |
| | $N_{\max}=1\ 000$ | $\Delta=\{0.326\ 6, 0.326\ 6\}$ | 0.18 | $\mathbf{X}_{\min}=(-1.357\ 0, -1.422\ 9)$ GRNN's output = -130.339 $f(\mathbf{X}_{\min})=-174.100\ 4$ |
| | $N_{\max}=5\ 000$ | $\Delta=\{0.143\ 5, 0.143\ 5\}$ | 0.09 | $\mathbf{X}_{\min}=(-1.299\ 2, -1.423\ 7)$ GRNN's output = -160.481 9 $f(\mathbf{X}_{\min})=-176.087\ 3$ |
| | $N_{\max}=10\ 000$ | $\Delta=\{0.101\ 0, 0.101\ 0\}$ | 0.05 | $\mathbf{X}_{\min}=(-1.290\ 6, -1.444\ 4)$ GRNN's output = -169.451 4 $f(\mathbf{X}_{\min})=-175.086\ 5$ |
| 基于参 数精度 的样本 采集法 | 精度要求 0.01 | $\Delta=\{0.1, 0.1\}$, 总样本数 10 201 | 0.06 | $\mathbf{X}_{\min}=(-1.304\ 8, -1.418\ 2)$ GRNN's output = -168.851 7 $f(\mathbf{X}_{\min})=-176.036\ 3$ |
| | x_1 精度要求 0.01, x_2 精度要求 0.05 | $\Delta=\{0.05, 0.1\}$, 总样本数 20 301 | 0.06 | $\mathbf{X}_{\min}=(-1.307\ 4, -1.418\ 2)$ GRNN's output = -168.619 7 $f(\mathbf{X}_{\min})=-176.038\ 6$ |
| 回归曲 面的样 本增补 法 | 第一次搜索 $N_{\max}=1000$ | $\Delta=\{0.326\ 6, 0.326\ 6\}$, 总样本数 1 000 | 0.18 | $\mathbf{X}_{\min}=(-1.357\ 0, -1.422\ 9)$ GRNN's output = -130.339 $f(\mathbf{X}_{\min})=-174.100\ 4$ |
| | 第二次搜索, 增补区间: $x_1: [-1.4, -1.3]$, $x_2: [-1.5, -1.4]$ | 增补区间上的 $\Delta=\{0.025, 0.025\}$ 新增样本数为 25, 总样本数 1 025 | 0.01 | $\mathbf{X}_{\min}=(-1.293, -1.425)$ GRNN's output = -176.063 6 $f(\mathbf{X}_{\min})=-175.981\ 3$ |

②构建 GRNN 模型。这里采取 holdout 方法, 对平滑因子 σ 在 [0.01, 1] 上每隔 0.01 进行一次试算, 选取使得对测试点的预测效果最佳的 σ , 作为 GRNN 网络的参数。这里, 用随机的 10 组样本作测试, 选取“与测试样本的相关系数 CC”作为 GRNN 效果的评价标准, 编制 Matlab 程序进行循环试算, 得到不同 σ 对应的相关系数变化情况, 发现当 $\sigma=0.06$ 时, GRNN 的效果最佳, 相关系数几乎为 1。所以, 取 $\sigma=0.06$, 用样本点训练 GRNN, 将得到的 GRNN 作为 $f(\mathbf{X})$ 的替代进行寻优。

③调用标准遗传算法^[19]对 GRNN 模型进行寻优。选用 Matlab 的 ga 函数, 设置进化代数数为 100, 种群大小为 20, 交叉函数选择为启发式算法 crossoverheuristic, 经过 19.7 s 的运行, 得到 GRNN 最小值为 -168.619 0, $\mathbf{X}_{\min}=(-1.3074, -1.418\ 2)$,

如图 4 所示, 将最优解代入测试函数 $f(\mathbf{X})$, 算得对应值为 -176.038 6。可见, 采取基于神经网络的仿真优化模型得到的最优解已经非常接近实际最优解, 由于 GRNN 对原函数进行回归时, 不可能完全一致, 存在可接受的误差, 所以得到的结果可能不是最优解, 但已非常接近最优, 可认为是满意解。

该算例表明, 基于神经网络的仿真优化模型在仅知道一系列样本点的情况下, 能够找出满意解。如果把例子中的函数换成一个复杂的仿真系统, 我们可以首先利用仿真模型获取样本, 从而利用本文仿真优化方法进行寻优。

(2) 不同样本选择策略对优化效果的影响分析

为了测试不同样本选择策略对优化时间和求解精度的影响, 我们对 2.4.3 节提出的 3 种策略进行对比, 结果如表 1 所示。以回归曲面的样本增补

法为例,在第一次搜索时,采取了等间隔样本采集法,选择产生 1000 个样本点,发现可能的最优解在点(-1.3570, -1.4229)附近,此时平滑因子 σ 为 0.18,用时 6.6s。因此,对该点附近的回归曲面进行细化描述,选取增补区间为 $x_1 \in [-1.4, -1.3]$, $x_2 \in [-1.5, -1.4]$,在该段上增加 25 个样本,则根据式(2)可算得 x_1 和 x_2 的增补样本间隔均为 0.025。然后将之前的 1 000 个样本点和新增的 25 个样本点合并为一个总样本,进行第二次搜索,此时平滑因子 σ 为 0.01,调用标准遗传算法(参数设置同前),得到 GRNN 最小值为-176.0636, $\mathbf{X}_{\min} = (-1.293, -1.425)$,对应函数值为-175.981 3。可见,由于在优化解附近增补了 25 个样本,使得该部分曲面的回归效果明显提高,GRNN 输出值与实际值的误差仅为 0.047%。

可以看出,相比等间隔样本采集法和基于参数精度的样本采集法,采用回归曲面的样本增补法可以极大的减少所需样本的数量(从前两种方法所需样本数分别为 5 000 和 10 201 个,减小到只需 1 025 个),从而可大幅提升仿真优化的速度。这是因为该方法仅在可能的极小值附近的回归曲面增补少量样本,避免了在其他区域生成不必要的样本,提高了求解效率。另外,从表 1 中的寻优时间看,由

于测试函数的样本获取时间非常短,样本数改变时对应的优化时间差别不大,可以说,影响基于神经网络的仿真优化求解时间的关键在于所需样本的数量。因此,样本选择策略宜尽量选择回归曲面的样本增补法。

(3) 方法的适用性测试

为检验本文方法的适用性,我们还测试了多个其他的典型测试函数,均采用“回归曲面的样本增补法”进行寻优,结果见表 2。可见,基于神经网络的仿真优化法均能够找到与最优解非常接近的满意解,这也证明了本文方法的正确性和适用性。

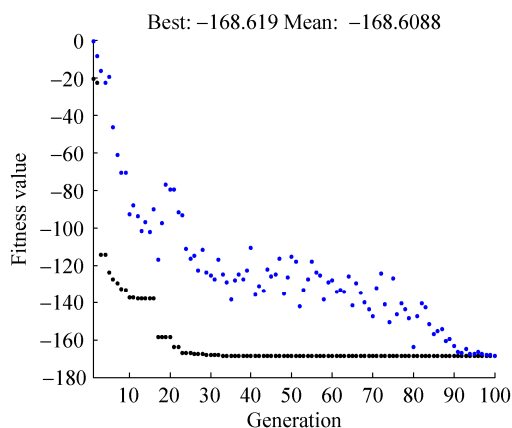


图 4 采用标准 GA 对 GRNN 网络进行寻优
Fig. 4 Optimization of GRNN regression surface using GA

表 2 几种典型测试函数的优化结果

Tab. 2 Optimization results for several typical test functions

| 测试函数 | 最优解 | 样本选择策略 | σ | 最少样本数 | 本文方法 |
|------------------------------|--|---|----------|-------|---|
| Levy No.5 Function | $\mathbf{X}^* = (-1.307, -1.4248)$ $f(\mathbf{X}^*) = -176.1375$ | 第一次搜索 1 000 个样本, $\mathbf{X}_{\min} = (-1.3570, -1.4229)$, $f(\mathbf{X}_{\min}) = -174.100 4$, 增补区间 $\{[-1.4, -1.3], [-1.5, -1.4]\}$, 增补样本 25 个 | 0.01 | 1025 | $\mathbf{X}_{\min} = (-1.293, -1.425)$ GRNN's output = -176.063 6 $f(\mathbf{X}_{\min}) = -175.981 3$ |
| Six-hump Camel-Back Function | $\mathbf{X}^* = (-0.0898, 0.7126), (0.0898, -0.7126)$ $f(\mathbf{X}^*) = -1.0316$ | 第一次搜索 1500 个样本, $\mathbf{X}_{\min} = (0.1473, -0.7388)$, $f(\mathbf{X}_{\min}) = -1.0146$, 增补区间 $\{[0.05, 0.25], [-0.85, -0.65]\}$, 增补样本 100 个 | 0.01 | 1600 | $\mathbf{X}_{\min} = (0.088 5, -0.711 5)$ GRNN's output = -1.031 1 $f(\mathbf{X}_{\min}) = -1.031 6$ |
| Goldstein-Price Function | $\mathbf{X}^* = (0, -1)$, $f(\mathbf{X}^*) = 3$ | 第一次搜索 1000 个样本, $\mathbf{X}_{\min} = (-0.006 6, -0.983 0)$, $f(\mathbf{X}_{\min}) = 3.155 3$, 增补区间 $\{[-0.05, 0], [-1.2, -0.8]\}$, 增补样本 100 个 | 0.03 | 1100 | $\mathbf{X}_{\min} = (0.003 5, -0.987 2)$ GRNN's output = 3.172 0 $f(\mathbf{X}_{\min}) = 3.063 5$ |

选取的测试函数如下:

①Six-hump Camel-Back 函数:

$$f(\mathbf{X}) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$$

式中: $-5 \leq x_1, x_2 \leq 5$ 。

②Goldstein-Price 函数:

$$f(\mathbf{X}) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$

式中: $-2 \leq x_1, x_2 \leq 2$ 。

4 结论

本文围绕提高复杂系统仿真优化的速度, 提出了一种新的仿真优化方法, 即基于神经网络的仿真优化方法。该方法针对仿真运行次数过多, 是造成仿真优化(OvS)问题速度过慢的瓶颈, 采用首先由仿真模型产生一定数量的样本, 然后利用神经网络的预测能力形成能够反映仿真输入输出关系的回归曲面, 从而实现用神经网络代替仿真模型进行适应度评价, 进而采用常用的优化方法在较短时间内实现寻优。从本文测试函数效果看, 该方法运算速度相对较快, 能够满足仿真优化的精度要求。

本文算法的优点有: (1) 所需样本数相对较少。对于仿真优化问题而言, 普通优化算法(比如遗传算法), 在寻优的时候并没有考虑计算目标函数(或适应度)所需花费的时间代价, 这是因为它的目标函数是显式的函数, 而仿真优化的目标函数是一个黑箱模型, 需要仿真来求得, 故普通优化算法在寻优中往往要计算大量的样本值, 大大超过了基于神经网络的仿真优化算法所需的样本数; (2) 样本不会重复获取。样本获取一次后, 计入总样本库, 不会重复计算; (3) 方便多台计算机并行计算。由于是分步实施: 先仿真得到样本, 后进行 GRNN 训练和优化, 因此, 在允许时间较短的情况下, 可以利用多台计算机同步获取仿真样本, 能够成倍地提高算法的效率; (4) 全局最优性。由于是在整个

可行域上采取等间隔获取样本, 并对各个局部极小区域增补样本, 可提高找到全局最优解的概率。

当然, 该方法在具体应用中还有很多值得研究的问题。比如样本选取的策略问题, 究竟选择多少训练样本和测试样本, 本文只是初步探索, 是否有更好的选择方法、是否存在更好的神经网络模型等, 这些都有待深入研究。这也是后续研究的重点。

参考文献:

- [1] 孙剑, 刘好德, 李克平. 城市干道交通信号协调控制仿真优化[J]. 同济大学学报(自然科学版), 2009, 37(11): 1467-1471.
Sun Jian, Liu Haode, Li Keping. Simulation optimization for urban arterial coordinate control [J]. Journal of Tongji University (Natural Science), 2009, 37(11): 1467-1471.
- [2] 李东, 汪定伟. 基于仿真的铁路集装箱装卸策略优化[J]. 系统仿真学报, 2009, 21(11): 3188-3191.
Li Dong, Wang Dingwei. Simulation-based optimization of loading-unloading strategies in railway container terminal[J]. Journal of System Simulation, 2009, 21(11): 3188-3191.
- [3] Tsai S C, Zheng Y X. A simulation optimization approach for a two-echelon inventory system with service level constraints [J]. European Journal of Operational Research (S0377-2217), 2013, 229(2): 364-374.
- [4] 刘娟, 张伟, 李群, 等. 基于启发式仿真优化的航迹规划方法研究[J]. 系统仿真学报, 2011, 23(12): 2592-2596.
Liu Juan, Zhang Wei, Li Qun, et al. Research of heuristic simulation optimization-based path planning[J]. Journal of System Simulation, 2011, 23(12): 2592-2596.
- [5] Hong L J, Nelson B L. A brief introduction to optimization via simulation[C]// Proceedings of the 2009 Winter Simulation Conference. Piscataway, New Jersey, USA: IEEE, 2009: 75-85.
- [6] 张会奇, 陈春良, 刘峻岩, 等. 装甲装备维修保障资源优化仿真研究[J]. 系统仿真学报, 2015, 27(1): 142-146.
Zhang Huiqi, Chen Chunliang, Liu Junyan, et al. Simulation research on armored equipment maintenance support resource optimization [J]. Journal of System Simulation, 2015, 27(1): 142-146.
- [7] Wang L F, Shi L Y. Simulation optimization: a review on

- theory and applications[J]. Acta Automatica Sinica (S1874-1029), 2013, 39(11): 1957-1968.
- [8] Hong L J, Nelson B L. Discrete optimization via simulation using COMPASS[J]. Operations Research (S0030-364X), 2006, 54(1): 115-129.
- [9] Kabirian A, Ólafsson S. Continuous optimization via simulation using Golden Region search [J]. European Journal of Operational Research (S0377-2217), 2011, 208(1): 19-27.
- [10] Khoo L P, Chen C H. Integration of response surface methodology with genetic algorithms[J]. Int J Advanced Manufacturing Technology (S0268-3768), 2001, 18(7): 483-489.
- [11] Capocchi L, Santucci J F. Discrete optimization via simulation of catchment basin management within the devsimpy framework[C]//Proceedings of the 2013 Winter Simulation Conference. IEEE, Piscataway, New Jersey, 2013: 205-216.
- [12] Specht D F. A general regression neural network[J]. IEEE Trans Neural Networks(S1045-9227), 1991, (2): 568-576.
- [13] Tomandl D, Schober A. A modified general regression neural network (MGRNN) with new, efficient training algorithms as a robust 'black box'-tool for data analysis[J]. Neural Netw(S0893-6080), 2001, 14(8): 1023-1034.
- [14] Li Q, Meng Q, Cai J, et al. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks[J]. Energy Conversion and Management (S0196-8904), 2009(50): 90-96.
- [15] Ward Systems Group, Inc. Neuroshell/Neurowindows Release 3.0 Manual, 3rd Edition[M]. Maryland, 1996.
- [16] 景涛. 基于改进广义回归神经网络的雷达故障预测[J]. 科学技术与工程, 2009, 9(15): 4492-4494.
Jing Tao. Fault prediction of radar based on modified general regression neural network[J]. Science Technology and Engineering, 2009, 9(15): 4492-4494.
- [17] Keka S, Papita D S, Siddhartha D. Response surface optimization and artificial neural network modeling of microwave assisted natural dye extraction from pomegranate rind [J]. Industrial Crops and Products (S0926-6690), 2012, (37): 408-414.
- [18] Li C F, Bovik A C, Wu X J. Preprocessing data for short-term load forecasting with a general regression neural network and a moving average filter[J]. IEEE Trans Neural Networks (S1045-9227), 2011, 22(5): 793-799.
- [19] 薛定宇, 陈阳泉. 高等应用数学问题的MATLAB求解[M]. 3版. 北京: 清华大学出版社, 2013.
Xue D Y, Chen Y Q. Advanced applied mathematical problem solutions with Matlab M]. 3rd edition. Beijing: Tsinghua University Press, 2013.
- 征的人脸表情识别[J]. 光学学报, 2014, 34(5): 1-7.
Li Yaqian, Li Yingjie, Li Haibin. Fusion of Global and Local Various Feature for Facial Expression Recognition [J]. Journal of Acta Optica Sinica, 2014, 34(5): 1-7.
- [22] Yang P, Liu Q S, Metaxas D N. Exploring facial expressions with compositional features[C]//IEEE International Conference on Computer V-ision and P attern Recognition, 2010: 2638-2644.
- (上接第35页)
- [20] 罗元, 崔叶, 王艳. 基于融合DCT和LBP特征提取的面部表情识别[J]. 半导体光电, 2014, 35(2): 330-334.
LUO Yuan, CUI Ye, WANG Yan. Face Expression Recognition Based on Fusion Feature of DCT and LBP[J]. Semiconductor Optoelectronics, 2014, 35(2): 330-334.
- [21] 李雅倩, 李颖杰, 李海滨. 融合全局与局部多样性特