

数学建模在实际生产计划中的应用

——基于 NARX 神经网络的地区用电预测

桂改花 苑占江 李祖猛 林泽楷

广东科学技术职业学院 广东 珠海 519090

摘要: 中国经济的快速增长带动了电力需求的迅猛增长,同时,电力消费使得环境压力加大、能源矛盾更加紧迫。本文针对预测某一个月未来用电量需求以及求解最小发电成本的问题,建立 NARX 神经网络时间序列模型和 INPL 整数非线性规划模型进行分析求解。本文首先将数据导入 Matlab 得到原数据的折线图,由图中可知其存在缺失值和异常值,再采用拉格朗日插值法对箱型图改进来剔除异常值和补全缺失值,然后可以得出结论:在每天不同时间段内,23:00-7:00 为用电低峰期,18:00-23:00 为用电高峰期。而在工作日和休息日之间,休息日相较于工作日来说用电量较多。对未来一个月的用电量预测运用了 NARX 神经网络时间序列模型进行求解,从而得到 5 月份的预测用电量。通过找出发电机最优组合来求每天电力生产总成本的最值,以每天电力生产总成本作为目标函数,并建整数非线性规划模型。根据题目的条件及相关数据,通过分析各时段的成本得出:各时段的电力生产总成本,然后对各时段成本求和得到目标成本函数。根据题目所给的已知条件进行合理的假设下,分析确定模型的约束条件。通过 lingo 软件编程求解,确定不同型号发电机在不同时段的使用数量,找出最优解,得到电产过程中每天的最小成本。

关键词: 拉格朗日插值法; NARX; 神经网络; 整数非线性规划

Application of mathematical modeling in practical production planning

-- Regional electricity prediction based on NARX neural network

Gaihua Gui, Zhanjiang Yuan, Zumeng Li, Zekai Lin

Guangdong Vocational College of Science and Technology, Zhuhai, Guangdong, 519090

Abstract: The rapid growth of China's economy has driven the rapid growth of the demand for electricity. At the same time, the consumption of electricity has increased the pressure on the environment and made the energy conflict more urgent. In this paper, the NARX neural network time series model and INPL integer nonlinear programming model are established to solve the problem of predicting the future electricity consumption demand of a certain month and solving the minimum generation cost. In this paper, the data is first imported into Matlab to obtain the broken line chart of the original data, from which it can be seen that there are missing values and outliers. Then, Lagrange interpolation method is used to improve the box diagram to eliminate outliers and complete missing values. Then, it can be concluded that in different time periods every day, 23:00 to 7:00 is the peak period of electricity consumption, and 18:00-23:00 is the peak of electricity consumption. And between working days and rest days, rest days use more electricity than working days. NARX neural network time series model is used to predict the electricity consumption in the next month, so as to get the predicted electricity consumption in May. By finding out the optimal combination of generators, the minimum value of the total cost of daily power production is obtained. The total cost of daily power production is taken as the objective function, and the integer nonlinear programming model is established. According to the conditions of the topic and relevant data, through the analysis of the cost of each period: the total cost of power production in each period, and then the sum of the cost of each period to get the target cost function. According to the given conditions given by the title under reasonable assumptions, analyze and determine the constraints of the model. Through the lingo software programming solution, to determine the use of different types of generators in different periods of time, find out the optimal solution, to get the minimum daily cost in the process of power production.

Keywords: Lagrange interpolation; NARX neural network; Integer nonlinear programming

一、引言

全国大学生数学建模竞赛是国家教委高教司和中国工业

与应用数学学会共同主办的面向全国大学生的群众性科技活动,目的在于激励学生学习数学的积极性,提高学生建立数学模

型和运用计算机技术解决实际问题的综合能力,鼓励学生踊跃参加课外科技活动,开拓知识面,培养创新精神。高职院校的培养目标是培养在生产、建设、管理、服务第一线工作的高级技术应用型人才,开设应用数学和数学建模课程符合高职院校的培养目标,是高职院校课程改革的有益尝试。

基于数学建模竞赛的创新型人才培养是指以数学建模竞赛为依托,将数学建模竞赛作为一个项目进行系统完整的运作,从而促进数学建模竞赛和课程教学改革的良性循环,形成

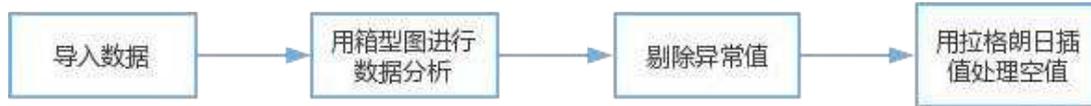


图1 数据预处理流程图

有助于提高学生的创新意识和创新能力的创新型人才培养。数学建模竞赛在整个学风建设中的作用、地位以及对素质教育和创新人才培养的独特功能,使之成为高职院校的一个重点课题,数学建模竞赛的开展,有利于激发学生的学习积极性和主动性,有利于提高学生的实践能力、创新能力、综合运用知识的能力,从而推动相关课程的教学改革,提高教学质量,培养创新型人才。把数学建模知识用到实际生产中,比如在电力生产计划中的应用,可以让学生把所学用到实际生产中。

随着社会可持续发展大环境的形成,电力发展也成为我国全面发展经济促进社会稳定的重要资源之一。由于疫情过后,我国不同地区所呈现出来的供电需求飞速上涨,并且自2021年后疫情时代的来临,我国整体经济已经在逐渐恢复中,在此背景下,我国外贸出口飞速提升,因此也拉动了整体的电力消费需求^[1]。能源供需矛盾凸现,电力建设速度加快,电力需求供不应求,进而导致了拉闸限电的再次出现。如何深化电力行业供给侧结构性改革,提高发展质量和效益是当今社会密切关注的问题,同时也吸引了各方对电力需求的预测和分析。

二、案例引入

数据为2021年1月1日至2022年4月30日该地区每日每小时的详细用电情况。为对该地区的电力需求进行预测和分析,本文联系相关知识以及查找相关文献,通过建立数学模型,解决以下问题:

问题一:结合所给的数据,利用统计方法分析该地区的用电特征,讨论不同时间段是否存在显著的用电差异以及预测未来一个月的电力需求。

问题二:结合问题一的结果和不同发电机的相关数据,在不考虑发电机连续工作时长和成本的情况下,分析为使该月发电的总成本最小,在每个时间段应分别使用哪些发电机。

三、案例分析

1. 问题一的分析 and 解决

(1) 针对问题一,题目要求用统计方法分析用电特征并判断其差异,再建立模型预测未来一个月的电力需求,本文首

先对数据通过拉格朗日插值法进行数据预处理,再进行数据分析可以得到484天内的用电量统计,从而知道其用电量差异。再根据问题背景和数据可知,影响用电量的因素多种多样,故对其的预测不能只建立在单一的理论,而应综合考虑各方面的因素上,而神经网络处理对处理这类复杂问题有独特之处。对于用电量的预测,本文根据各种影响因素建立一个神经网络时间序列模型对问题进行处理分析,利用已知数据训练网络,从而实现对用电量的预测。

(2) 统计分析用电特征

将数据导入Matlab绘制折线图得知该数据中存在3个异常值点,此外,数据中还存在少量缺失值。采用拉格朗日插值法^[2]对箱型图改进后能直接替换异常值,简化了处理过程又能够得到客观准确的结果。可以发现484天内平均每天的用电规律,其中用电低峰期为23:00-7:00,用电最高峰期为18:00-23:00。将7天作为一个周期判断可得工作日的用电量较少,休息日的用电量较多。

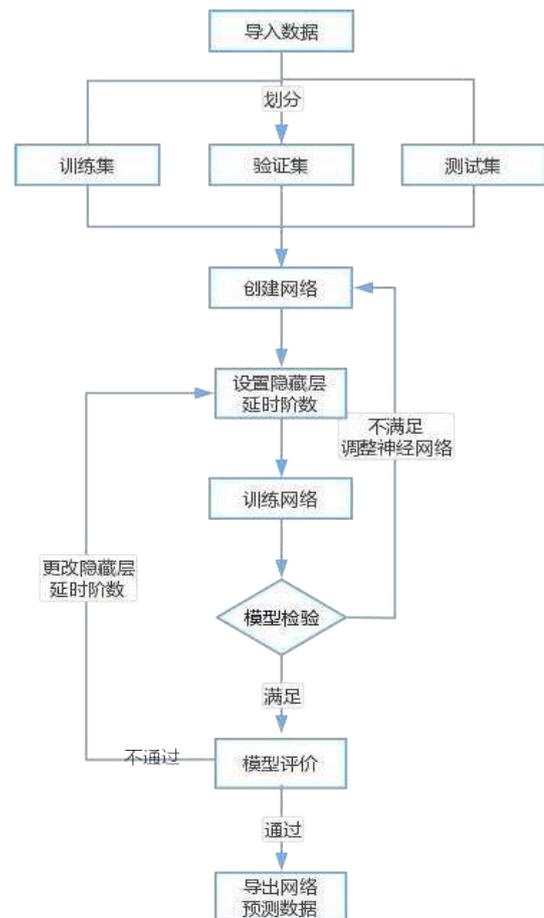


图2 神经网络预测模型建立

(3) 模型建立

NATX 神经网络是属于递归神经网络范畴的带外部输入的非线性自回归模型^[3], 具有输入相应的反馈回路, 有理论证明, NARX 神经网络可用于预测时间序列, 且通常可以保留信息的时间是常规递归神经网络的 2 ~ 3 倍, 因此得到国际上广泛认可^[4]。本文采用神经网络时间序列模型对用电量的预测进行求解, 其特点是可以根据过去的变化趋势利用统计学的方式预测未来, 且在考虑发展趋势的同时, 注重周期性变化对时间点的影响。

Step1: 创建网络

由于一个计算模型 $M(\varphi(t))$, $\varphi(t)$ 表示与时间相关的输入, 利用输入和输出的数据集建立描述 M 的数学模型, 把每天对应的时间段放入时间区间 $[T_1, T_2]$ 以时间间隔 d 离散为一系列时刻 $t_k (k = 1, 2, \dots, N_T)$ 。

NARX 模型将一个给定时刻的输出表示为过去输入和输出在当前时刻之前时刻的值的函数, 表达式即

$$y_n(t_k) = \Gamma(z(t_k)) + \varepsilon_t(t_k)$$

由题得出 NARX 模型:

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-n_y), x(t-1), x(t-2), \dots, x(t-n_x))$$

其中算式中的 f 为非线性函数; $y(t)$ 为 t 时刻的目标向量; $x(t)$ 为外部输入向量;

$y(t-1), y(t-2), \dots, y(t-n_y)$ 延时后的目标向量; $x(t-1), x(t-2), \dots, x(t-n_x)$ 为延时后的外部输入向量, 采用神经网络时间序列模型对用电量的预测进行求解, 其特点是可以根据过去的变化趋势利用统计学的方式预测未来, 且在考虑发展趋势的同时, 注重周期性变化对时间点的影响。建立网络结构如下。NARX 神经网络拓扑结构如下图 3 所示:

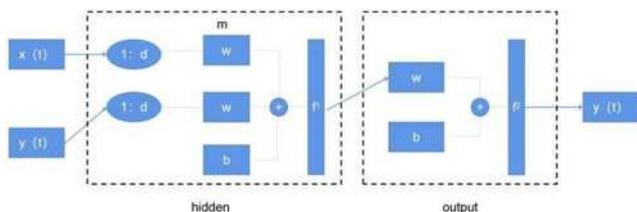


图 3 NARX 神经网络拓扑结

多层网络可以实现的预报和分类任务、高度非线性关系和非线性分类界限的一维和二维预报, 则由英汉神经元的非线性激励函数来完成。网络权值控制着非线性激励函数的特征, 训练在接触到权值特征时开始进行不断地调整, 这样可以使激励函数逐渐逼近期望的响应, 而网络预报误差逐渐下降至指定误差阈值以下地过程就称作为网络学习。NARX 神经采用 LM (Levenberg-Marquardt) 方法作为训练函数, LM 方法综合了

高斯-牛顿训练方法和最速下降法的优点, 使由二阶函数导数导致的沿误差上升而产生的误差变为沿误差负梯度的方向减小。

根据上述模型, 预测出 5 月份的用电量。

表 1 5 月份用电量预测

日期时段	0:00-1:00	1:00-2:00	2:00-3:00	...	23:00-24:00
2022/5/1	14903	15108	14723	...	29713
2022/5/2	16694	16942	16699	...	35487
2022/5/3	12149	13317	13142	...	23148
2022/5/4	13173	13306	13327	...	23761
2022/5/5	12744	12496	12378	...	21953
...
2022/5/30	13081	15004	14682	...	18337
2022/5/31	12780	12179	12149	...	21824

2. 问题二的分析和解决

本题研究的是发电机最优组合问题, 表 1 中有 10 种不同型号的发电机, 每种发电机都有启动成本、固定成本和边际成本, 并且规定了 7 个不同的供电时段。题目要求 7 个时段的发电机组既能够满足用电需求, 同时也满足成本费用最小, 不造成过多浪费。其中固定成本和边际成本较容易确定, 而启动成本 (电机启动时所花费的代价) 则难于有效地确定。因此, 我们不能将各个时段分开分别求最优解。

由于每种发电机成本都是由启动成本、固定成本和边际成本三种动态约束条件组成, 本题我们采用动态模型^[2]来求解, 我们将这个时段作为一个整体, 考虑前后两相邻时段 发电机开启的数量, 尽量在前一时段的基础上减少启动成本。因为如果后一个时段某种 型号的电机数量减小, 则只需要关闭电机即可, 它不需要付出任何代价, 这样就可以大大减少时段的启动成本。

(1) 确认目标函数

电力生产的总成本是由发电机的启动成本、固定成本及边际成本组成。为了提出一种发电机的最优组合使得电力成本最小的数学模型, 本文通过分析局部优化因素, 通过求解各时段发电机最小成本。而总成本中的启动成本是动态的, 我们建立整数动态规划模型。

Step1:

i 个不同型号的发电机在不同时间段的启动成本为 W_{ij} , 在第一时段, 发电机的启动不受其他影响, 按正常成本计算, 其数学表达式为:

$$W_{ij} = q_j * x_{ij}$$

Step2:

第 i 时段发电机的总成本 W_i 由发电机的启动成本 W_{ij} 固定成本 T_j 及边际成本 M_j 组成,

其数学表达式为:

$$W_{ij} = q_j * x_{ij} \quad W_i = W_{ij} + T_j * x_{ij} + (p_j - p_{j \min}) * M_j * x_{ij}$$

Step3:

建立每天发电机的总成本的目标函数：

$$W_{ij} = q_j * x_{ij} \min W = \sum_{i=1}^7 \sum_{j=1}^{10} W_i \quad (i, j \in N)$$

(2) 约束条件的建立

由于各种型号的发电机的输出功率不应低于其最小输出功率，高于最大输出功率，所以在不同时段发电机的输出功率有上、下限约束：

$$W_{ij} = q_j * x_{ij} p_{j \min} \leq p_{ij} \leq p_{j \max}$$

因为 10 种不同发电机的数量是固定的，其数学表达式为：

$$w_{ij} = q_j * x_{ij} \text{ s.t. } \begin{cases} 0 \leq x_{ij} \leq K_j \\ p_{j \min} \leq p_{ij} \leq p_{j \max} \end{cases} \quad (1 \leq i \leq 7, 1 \leq j \leq 10)$$

满足电力生产的需求，根据题目要求可得：

$$W_{ij} = q_j * x_{ij} Q_i = \sum_{j=1}^{10} p_{ij} * x_{ij} \geq \delta_i \quad (1 \leq i \leq 7, 1 \leq j \leq 10)$$

(3) 确立多目标规划模型

$$W_{ij} = q_j * x_{ij} \min W$$

$$= \sum_{i=1}^7 \sum_{j=1}^{10} [W_{ij} + T_j * x_{ij} + (p_j - p_{j \min}) * M_j * x_{ij}]$$

$$W_{ij} = q_j * x_{ij} \text{ s.t. } \begin{cases} 0 \leq x_{ij} \leq K_j \\ p_{j \min} \leq p_{ij} \leq p_{j \max} \\ Q_i = \sum_{j=1}^{10} p_{ij} * x_{ij} \geq \delta_i \\ W_{ij} = q_j * x_{ij} \leq 10, x_{ij}, i, j \in N \end{cases} \quad (1 \leq i \leq 7, 1 \leq j \leq 10)$$

(4) 模型的求解

由于将各个时段各种不同型号发电机的成本进行累加，然后写出相应的约束条件，使得 Lingo 软件对题目进行求解得出发电机的最优组合结果如表 2。

表 2 各时间段启用的发电机数量

时段 型号	0-5	6-8	9-12	12-14	14-18	18-22	22-24
A1	0	4	10	1	0	0	0
A2	0	2	0	4	5	4	4
A3	7	8	8	8	3	3	1
A4	0	5	1	0	0	7	2
B1	0	3	3	6	5	2	1
B2	2	8	8	8	8	8	8
B3	2	4	2	4	2	2	1
C1	0	10	5	12	8	11	12
C2	3	9	9	8	8	4	4
C3	2	8	7	7	8	8	4

四、模型优点和推广

神经网络模型时间序列分析可以从时间序列中找出变量变化的特征、趋势以及发展规律，从而对变量的未来变化进行有效地预测。自相关性图说明了预测误差在时间上是如何关联的。对于完美的预测模型，自相关函数应该只有一个非零值，并且应在零滞后时出现（此值为均方误差）。这表示预测误差彼此完全不相干（白噪声）。动态规划任何一个 i + 1 阶段都仅仅依赖 i 阶段做出的选择。而与 i 之前的选择无关。但是动态规划不仅求出了当前状态最优值，而且同时求出了到中间状态的最优值。本文中运用的模型实现了数据多样、约束条件动态化情况下对用电量的预测以及最优化求解。还可适用于对用水量、用电量等一系列具有预测问题的求解，可以有效提高资源利用率。

参考文献：

[1] 陆一川. 新能源与传统电力的融合发展分析[J]. 集成电路应用, 2022, 9(03): 234-235.

[2] 赵媛媛, 季洁. 基于拉格朗日插值法的国省干线缺失数据恢复[J]. 无线互联科技, 2021, 18(10): 97-100.

[3] 鞠宪龙. 二阶对角递归神经网络的算法研究及应用[D]. 哈尔滨: 哈尔滨工程大学, 2011: 9.

[4] ZOUNEMAT-KERMANI M, STEPHAN D, HINKELMANN R. Multivariate NARX Neural Network in Prediction Gaseous Emissions Within the Influent Chamber of Wastewater Treatment Plants[J]. Atmospheric Pollution Research, 2019, 10(6): 1812-1822.

基金项目：广东省高等职业教育教学改革研究与实践项目人工智能技术应用专业“产教融合、精准育人”人才培养体系的构建与实践（项目编号：GDJG2021155）；广东省智慧职教工程技术研究中心（项目编号：2021A118）；2022 年《应用数学（1）》校级课程思政示范课；2021 年《应用数学（1）》校级金课。

作者简介：桂改花（1981.9—），女，汉族，山东聊城，硕士研究生，讲师；研究方向：应用数学，数学建模，数学教育。