

文章编号:1000-582X(2005)02-0137-04

基于BP神经网络的库存动态预测及其应用*

罗兵,黄万杰,杨帅

(重庆大学经济与工商管理学院,重庆 400030)

摘要:利用库存的历史数据,建立了动态环境下一个基于BP神经网络的初始库存模型,采用提前结束训练法克服因样本量不足而产生的网络过适应问题,并通过网络的训练获得一个关系简单的库存模型。每次获得新样本数据后对模型进行重构,达到了动态建模和分析的目的,并用实例进行了说明,为库存系统的管理决策提供了理论依据。

关键词:BP神经网络;动态建模;库存管理;数据挖掘;过适应

中图分类号:F253.4

文献标识码:A

市场经济条件下外部环境是瞬息万变的,这对库存动态建模提出了很高的要求。库存系统中的各个因素有着复杂的内在联系,为获得准确的库存预测值需要很多的变量(即多维变量决策),用传统的方法不能处理此类问题。如用经济订购批量(EOQ)模型建立一个函数关系式,也只能是一个近似的结果。实际上不必搞清楚系统内部的机理,仅通过输入和输出的信息,就能掌握系统的整体运行特性,这就是所谓的黑箱方法。神经网络即是一种黑箱方法,它具有非线性动力系统的特性和处理大量输入、输出数据的能力,能模仿人脑的识别能力,揭示库存数据中的复杂关系,并建立预测模型^[1],获得巨大的经济效益。

由于BP(back propagation)算法具有广泛的适用性,使得它在1986年被Rumelhart D E等^[2]提出后,很快就成为应用最为广泛的多级网络训练算法。目前,已有学者将BP神经网络技术用于供应链与库存管理。Reyes-Aldasoro C C等^[3]于1999年采用神经网络技术与动态规划结合建立了按地理分散分布的零售公司的混合结构模型;Partovi F Y等^[4]于2002年利用BP人工神经网络对库存的ABC分类法进行了研究;Bansal K等^[5]于1998年使用基于神经网络的DM和KDD技术解决一个大医药分销公司的库存问题,他们通过对比传统的统计技术以决定应选用的最好的神经

网络类型,并用一个实例说明使用这个神经网络能在维持同一水平的顾客需求满足率时减少总库存的50%;轩超亭,黄培清等^[6]于2000年系统地总结了神经网络技术在供应链管理中的应用,并用BP网对上海市场自行车的需求量进行了预测;王东旭、沈益民等^[7]于2001年用BP神经网络模型预测了ERP的安全供货库存;何炎祥、李峰等^[8]于2002年介绍了使用基于数据挖掘和知识发现的神经网络技术来解决库存问题的一般方法。然而BP神经网络的收敛速度很慢,许多学者提出了改进算法,Nguyen D和Widrow B^[9]通过加入权值 W 和阈值 θ 得到一个初始双层网络,这种方法通常能成数量级地缩短神经网络的训练时间,这对于规模较大的网络训练具有重要意义。Hagan M T和Menhaj M^[10]在Rumelhart^[2]的基础上,提出了BP神经网络的一个改进算法,它不用计算Hessian矩阵,因此能获得很快的训练速度。

笔者在Hagan M T和Menhaj M^[10]研究的基础上,借助BP人工神经网络的特性^[11]:学习性、普化性(去噪音、容残缺的能力)、分布性(系统受损时仍可正常工作)和实时性,通过对企业历史数据的挖掘,进行库存系统模式识别,建立了一个时变动态模型,在进行灵敏度分析后,将其应用于某电子企业,仿真结果表明此方法实用有效。

* 收稿日期:2004-10-23

基金项目:国家青年社科基金(02cjj027)

作者简介:罗兵(1964-),男,重庆市人,重庆大学副教授,博士,主要研究方向:物流与供应链管理和决策优化。

1 模型的建立

1.1 BP 网络的确定

BP 神经网络包括输入层、隐层、输出层。一般地, 设 BP 网络的样本集为:

{(X, Y) | X 为输入向量, Y 为 X 对应的理想输出向量}

输入、输出向量的维数是由问题直接决定的, 隐层的层数和各个隐层神经元的个数则是与问题相关的。BP 网一般都选用三级网络(即 1 个隐层)。

学习过程由正向计算和反向传播组成。通过反复训练学习直到误差信号最小或学习结果令人满意为止, 故模型又称 BP 反向传播模型。

加入权值 W 和阈值 θ 的 BP 神经网络的数学表达式为:

$$Y = f(\sum W_j X_j + \theta_j) \tag{1}$$

误差函数 E 的减小采用最速下降法。当 E 是平方和(在训练前向网络中这是很典型的)的形式时, E 中的 Hessian 矩阵近似为

$$H = J^T J \tag{2}$$

同时梯度为

$$\Delta E = J^T e \tag{3}$$

其中 e 是一个网络误差矢量, J 是含有 e 对权值 W 和阈值 θ 的一阶导数的雅可比矩阵, 它可以通过计算一个比 Hessian 矩阵简单的标准的反向传递算法来计算出。

LM 算法使用下式来近似计算 Hessian 矩阵:

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \tag{4}$$

当标量 μ 等于零时, 它就是使用 Hessian 近似矩阵的牛顿法; 当 μ 很大的时候, 它变成一个小步长的梯度下降法。牛顿法能更快速、更准确地靠近误差极小值, 因此 LM 算法的目标是要尽快地向牛顿法转换。故 μ 在每次迭代成功之后均减少。这样, 在算法的每次迭代中误差函数总是减少的。这个算法是训练中等规模(可多到数百个权值)的前向神经网络最快的方法。

1.2 网络学习与测试

一个神经网络需要的训练样本数与问题的规模、复杂程度、网络学习算法等有一定的关系。有一个粗略的估计方法, 训练样本的数目约为连接权个数的两倍。但实际中往往无法获得足够的样本, 这会导致在神经网络训练期间发生过适应(overfitting)^[12], 即网络

对训练样本的误差已变得很小, 但对新的样本的误差却很大。如图 1 所示。

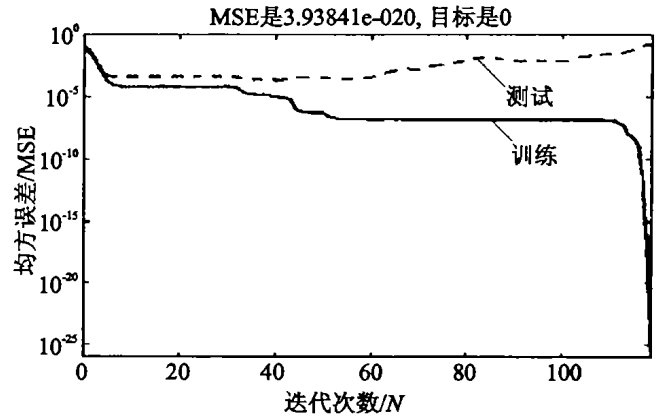


图 1 BP 人工神经网络的过适应

为了克服这个问题, 采用了提前结束(early stopping)训练法: 将企业的历史数据组成的样本集分为 3 个样本数目相近的子集: 训练集(Training Set), 有效集(Validation Set), 测试集(Test Set), 则样本数目可以减少一些。监测有效集误差的收敛性。当有效集曲线上上升时(有效集误差发散), 即使训练误差仍未达到既定精度, 也要提前结束训练。

用 MATLAB 语言建立初始库存模型, 输入样本数据进行训练。当训练曲线收敛到既定精度或有效曲线上上升时则停止训练, 并用测试集监测网络的普化性。若训练曲线经过很多次迭代后仍不能收敛, 则修改隐层神经元的个数或调整权值矩阵 W 的初值, 反复试验。达到既定精度后 MATLAB 生成一个库存模型, 即可进行分析与预测工作。

1.3 分析与预测

灵敏度分析(Sensitivity Analysis)用于评估一个给定的输入变量对网络输出的影响。改变该变量的输入, 而其他输入变量为某固定值。期间, 监测网络输出的改变。由这种形式的分析得到的知识是形如“IF X 减少 5% THEN Y 增加 8%”的规则。由此可以对库存进行很好的仿真, 找出库存的瓶颈环节与重点控制处, 从而改善库存系统。

预测第 $i+1$ 期的 EOQ, 是将第 i 期影响库存的因素输入训练好的模型中得到期望的预测值。虽然此方法得到的库存量不一定是最优值, 但却很接近。它最大的优点是能通过网络的训练, 构造出原本很复杂的库存模型, 并在后继工作中采用新的样本, 达到动态建模适应新环境的目的, 故在整个计划期中是最优的。样本集变得越丰富, 对以后的建模越有利。

2 实例

选择某电子企业生产的汽车分电器库存系统,用ABC分类法得出A类物资之一为点火模块,故对点火模块进行实例建模。对系统做出了如下假设:1)订货量与价格无关;2)瞬时补货。

选择影响库存总成本的几个主要的、相关性较强的因素作为网络的输入(下标*i*表示第*i*阶段):

N_i 为阶段序数(基期为第1阶段); Q_i 为实际订货批量(件/次); T_i 为 Q_i 使用的天数; S_i 为期初库存量; P_i 为单价(元/件); $C1_i$ 为单位存货保管费用(元/件); $C2_i$ 为一次订购费用(元/次); $C3_i$ 为单位存货缺货损失(元/件); EOQ_i 为第*i*+1阶段最低订货批量(件/次)。

即输入层神经元为8个($N_i, Q_i, T_i, S_i, P_i, C1_i, C2_i, C3_i$),每阶段均可变;隐层按Hagan M T和Menhaj M^[10]选一层,神经元个数是输入层神经元个数的1/3~1/2倍,文中选4个;输出层神经元为 EOQ_i ,它是第*i*阶段后根据最低库存费用来确定的,而不是实际发生的订货批量。BP神经网络的有老师学习即是将输出信号(网络实际输出)和教师信号 EOQ_i 的误差进行比较,通过反复训练学习,直到误差信号最小或学习结果令人满意为止。

样本数据来源于1997.1~2001.8共56个订购阶段(即56个样本)。对样本集采用随机分组:训练集20个,有效集和测试集均为18个。学习的效果如图2所示。

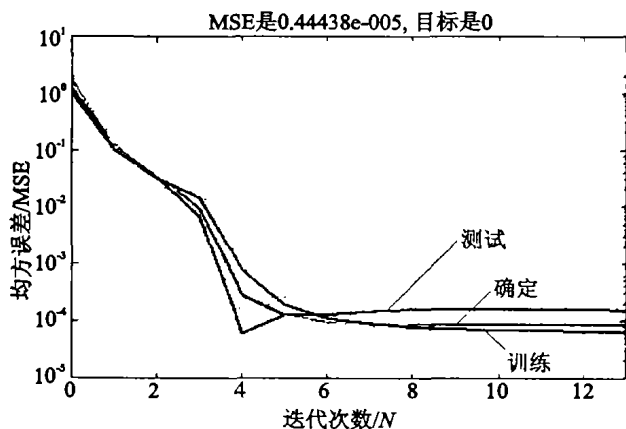


图2 BP神经网络的学习效果曲线

经过13次迭代后网络收敛,得到各权值 W 及阈值 θ 后可得出一个简单的库存模型关系式:

$$[X \cdot W_1 + \theta_1] \cdot W_2 + \theta_2 = EOQ \quad (5)$$

其中

$$X = (N_i, Q_i, T_i, S_i, P_i, C1_i, C2_i, C3_i)$$

$$W_1 = \begin{bmatrix} 333.3942 & 92.95 & -2.7706 & 291.4773 \\ 104.232 & 49.9143 & 87.6767 & 40.4041 \\ -1638.6594 & -2295.9265 & -1390.9006 & 273.3248 \\ 55.349 & -49.2376 & 69.4467 & 13.0048 \\ -1435.6086 & 2365.0292 & 435.7755 & -472.2908 \\ 2528237.715 & 1659832.173 & 3235803.725 & 4591380.265 \\ 167.3123 & 2964.7583 & -2045.9947 & 2748.2895 \\ -1077.2527 & -1229.163 & 1495.8694 & 865.0956 \end{bmatrix}$$

$$\theta_1 = (-57.5639 \quad -55.7388 \quad -60.0852 \quad -59.5093)$$

$$W_2 = (-1.1809 \quad 0.01879 \quad -0.030479 \quad 0.0026351)^T$$

$$\theta_2 = -0.42646$$

上述这些值与具体的库存系统有关,且对同一库存系统也不唯一,随每一次网络训练而不同,但都能反映某特定库存系统各因素之间的关系。

对 P_i 进行灵敏度分析。让其他输入变量固定,改变 P_i 的值,监测到的网络输出 EOQ_i ,如图3所示。

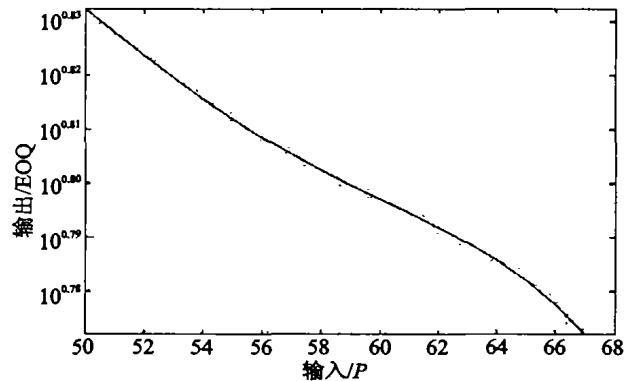


图3 P_i 变化对 EOQ_i 的影响

从图3可以看出, P_i 越大, EOQ_i 越小。同理可对其他变量进行分析。

预测量为2001年9月的 EOQ_{57} 。把 $N_{57} = 57, Q_{57} = 6400, T_{57} = 23, S_{57} = 15, P_{57} = 58.88, C1_{57} = 0.02312, C2_{57} = 73, C3_{57} = 78$ 输入训练好的神经网络,得 $EOQ_{57} = 6344$ 。当本阶段完成后,重新确定 EOQ_{57} 的值(不一定是6344),则可得到一个新的样本。由此不断地积累新的样本,并与原样本混合,用于下一阶段的重新建模,则企业的库存模型可以越来越精确。

3 结语

文中的方法能处理很复杂的问题,约束条件比传统的库存模型也少得多,并充分利用了BP人工神经网络学习性和时变性的优点,达到动态建模和分析的目的。人工神经网络技术是一种数据挖掘技术,历史数据的获取与整理是关键,在实际中应尽早建立和完善企业的数据库。由于神经网络的可解释性及模型的

稳健性较差,许多学者仍对其心存疑虑,解决这些问题是今后研究的重点。

参考文献:

- [1] THALL N. Neural Forecasts: A Retail Sales Booster[J]. *Discount Merchandiser*, 1992, 32: 41 - 46.
- [2] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning Representations by Back-propagating Errors[J]. *Nature*, 1986, 323: 533 - 536.
- [3] REYES-ALDASORO C C, GANGULY A R, LEMUS G, et al. A Hybrid Model Based on Dynamic Programming, Neural Networks, and Surrogate Value for Inventory Optimization Applications[J]. *The Journal of the Operational Research Society*, 1999, 50: 85 - 94.
- [4] PARTOVI F Y, ANANDARAJAN M. Classifying Inventory Using an Artificial Neural Network Approach[J]. *Computers & Industrial Engineering*, 2002, 41: 389 - 404.
- [5] BANSAL K, VADHAVKAR S, GUPTA A. Neural Networks Based Forecasting Techniques for Inventory Control Applications[J]. *IEEE*, 1998, 2(1): 97 - 102.
- [6] 轩超亭, 黄培清, 吕冬. 神经网络技术在供应链管理中的应用[J]. *工业工程与管理*, 2000, (3): 41 - 44.
- [7] 王东旭, 沈益民, 王志强. 用 B - P 神经网络模型预测 ERP 的安全供货库存[J]. *计算机应用*, 2001, 21(3): 53 - 55.
- [8] 何炎祥, 李峰, 宋志凯, 等. 神经网络技术在库存管理中的应用[J]. *计算机工程与应用*, 2002, 38: 182 - 184.
- [9] NGUYEN D, WIDROW B. Improving the Learning Speed of 2-layer Neural Networks by Choosing Initial Values of the Adaptive Weights[A]. *Proceedings of the International Joint Conference on Neural Networks*[C]. Washington, 1990, 3: 21 - 26.
- [10] HAGAN M T, MENHAJ M. Training Feedforward Networks with the Marquardt Algorithm[J]. *IEEE Transactions on Neural Networks*, 1994, 5(6): 989 - 993.
- [11] 韩家炜, Michelin K. 数据挖掘: 概念与技术[M]. 北京: 机械工业出版社, 2001.
- [12] 蒋宗礼. 人工神经网络导论[M]. 北京: 高等教育出版社, 2001.

Application of BP Neural Networks in Inventory Dynamic Modeling

LUO Bing, HUANG Wan-jie, YANG Shuai

(College of Economics and Business Administration, Chongqing University, Chongqing 400030, China)

Abstract: The authors establish an original inventory model based on BP neural networks in dynamic environment using the inventory history data, and overcome the net overfitting problem occurred with insufficient samples by using early stopping method. After a simple-relation inventory model training is attained, they get inventory model and analyze it dynamically by reconstructing the model after getting new sample data. An example is provided to illustrate the model. The theoretical evidence is provided for the inventory system to make management decision.

Key words: BP neural networks; dynamic modeling; inventory management; data mining; overfitting

(编辑 刘道芬)