

# 非规则三维数据的曲面拟合方法

徐安凤<sup>1</sup>, 李金莱<sup>1</sup>, 姚春光<sup>2</sup>

XU An-feng<sup>1</sup>, LI Jin-lai<sup>1</sup>, YAO Chun-guang<sup>2</sup>

1. 南阳师范学院 计算机系, 河南 南阳 473061

2. 国防科技大学 电子科学与工程学院, 长沙 410073

1. Computer Department of Nanyang Normal University, Nanyang, Henan 473061, China

2. Electronic Science and Engineering College, National University of Defense Technology, Changsha 410073, China

E-mail: xuaf@sina.com.cn

XU An-feng, LI Jin-lai, YAO Chun-guang. A surface fitting method for irregular 3-D data. Computer Engineering and Applications, 2009, 45(20): 234-235.

**Abstract:** This paper gives a surface fitting method for irregular 3-D data, the BP net presented needn't delete the irregular data, then it holds the information integrality, furthermore, the surface given is smooth and takes on good continuity, the local details are comprehensive, and it has local derivations at any point.

**Key words:** irregular data; surface fitting; Back-Propagation

**摘要:** 给出了一种非规则三维数据的曲面拟合方法, 该方法给出的网络模型不需要删除奇异数据, 从而可以保持数据信息的完整性, 此外, 该方法给出的拟合曲面平滑, 连续性好, 局部细节丰富, 且处处可偏导。

**关键词:** 非规则数据; 曲面拟合; Back-Propagation

DOI: 10.3778/j.issn.1002-8331.2009.20.068 文章编号: 1002-8331(2009)20-0234-02 文献标识码: A 中图分类号: TP391

## 1 引言

曲面拟合是用连续曲面近似地刻画或比拟空间离散点组所表示的坐标之间函数关系的一种数据处理方法, 通过实验或观测得到量  $x$ 、 $y$  和  $z$  的一组数据对  $(x_i, y_i, z_i = f(x_i, y_i))$  ( $i=1, 2, \dots, m$ ), 其中各  $(x_i, y_i)$  是彼此不同的, 期望用一类与数据的背景材料规律相适应的解析表达式,  $y=f(x, y, c)$  反映量  $(x, y)$  与  $z$  之间的依赖关系, 即在一定意义下“最佳”地逼近或拟合已知数据。  $y=f(x, y, c)$  常称作拟合模型, 式中  $\bar{c}=(c_1, c_2, \dots, c_m)$  是一些待定参数。当  $\bar{c}$  在  $f$  中线性出现时, 称为线性模型, 否则称为非线性模型。

研究了一种特殊条件下的曲面拟合方法, 在给出的样本数据对  $(x_i, y_i, z_i = f(x_i, y_i))$  ( $i=1, 2, \dots, m$ ) 中, 有部分  $(x_i, y_i)$  数据是相同的, 即  $x_i = x_j, y_i = y_j, z_i \neq z_j, i \neq j$ , 也就是说一个  $(x_i, y_i)$  对应多个不同的  $z_i$ , 这是一种非理想的情况, 因为根据函数映射关系, 同一  $(x_i, y_i)$  只能有一个  $z_i$  与之对应, 所以这种数据无法使用传统的曲面拟合方法解决。但是, 这种非规则的数据相比规则的数据具有更重要的工程意义, 因为随着试验或观测条件的不同, 即使是同一个状态对  $(x_i, y_i)$ , 也可能出现不同的观测结果  $z_i$ , 观测者无法辨别这些不同观测结果  $z_i$  的正确与否, 因此它们具有相同的参考价值。

通过建立人工神经网络 Back-Propagation 模型的方法对这一问题进行了研究, 把  $(x_i, y_i, z_i = f(x_i, y_i))$  ( $i=1, 2, \dots, m$ ) 作为样本数据传递给前馈神经网络, 通过误差后向传播调整各层之间的权值和阈值, 逐步形成一个全局逼近的网络模型, 然后通过提取网络模型的权值、阈值和传输函数, 可以构造一个以  $x_i$ 、 $y_i$  为输入变量, 以模型的权值和阈值为系数的非线性函数  $z = f(x, y, \bar{c})$ , 计算表明, 该方法给出的模型相比其他方法曲面拟合结果更为准确。

## 2 非规则三维数据的分布

研究了 659 组观测数据的曲面拟合, 即  $(x_i, y_i, z_i = f(x_i, y_i))$  ( $i=1, 2, \dots, 659$ ), 这些数据是在不同的试验环境和观测条件下得到的, 其中有约 16% 的  $(x_i, y_i)$  是相同的, 但其对应的  $z_i$  值不同, 它们具有同样的研究价值。从空间分布来看, 这些点局部呈散乱分布, 整体分布与  $x$  呈近似反比、与  $y$  呈近似正比关系。

部分  $(x_i, y_i)$  对应两个甚至多个  $z_i$  点, 可以看出待拟合曲面的数据是非理想的, 由于传统的曲面拟合方法要求  $(x_i, y_i)$  和  $z_i$  一一对应, 从而无法使用传统的拟合方法解决, 一个通常的处理方法是根据先验知识对相同的数据进行选择删除, 即约束  $(x_i, y_i)$  只对应一个  $z_i$  值, 但是在本文的工程应用中, 没有先验

**作者简介:** 徐安凤(1973-), 女, 讲师, 主要研究方向: 优化算法; 李金莱(1977-), 女, 讲师, 主要研究方向: 计算机应用; 姚春光(1975-), 男, 博士, 主要研究方向: 无线通信系统。

收稿日期: 2008-04-18 修回日期: 2008-07-30

知识可以作为依据, 无法判断这些重复数据正确与否, 它们有同等重要的参考价值, 不能进行选择删除。文中给出的 Back-Propagation 模型可以容忍重复数据, 仅仅对曲面拟合的精度有影响, 但其拟合结果完全可以满足工程需要。

### 3 Back-Propagation 曲面拟合

基于误差反向传递算法的多层前馈神经网络(BP 网络)是目前应用最广、通用性最好的网络, 能用于函数逼近、模式识别和分类等。网络是由输入层、中间层、输出层组成的阶层型神经网络, 中间层可扩展为多层。相邻层之间各神经元进行全连接, 而每层各神经元之间无连接, 如图 1 所示。网络通过训练数据进行学习, 当一组训练数据提供给网络后, 各神经元获得网络的输入响应产生连接权值和阈值, 然后按减小期望输出与实际输出误差的方向, 从输出层经各中间层逐层修正各连接权和阈值, 回到输入层。此过程反复交替进行, 直至网络的全局误差趋向给定的极小值, 即完成学习的过程。

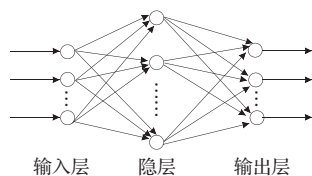


图 1 BP 网络结构图

一个神经元节点的结构如图 2 所示。其中  $x_1, x_2, \dots, x_i, \dots, x_n$  表示来自神经元  $1, 2, \dots, i, \dots, n$  的输入,  $w_1, w_2, \dots, w_i, \dots, w_n$  表示神经元  $1, 2, \dots, i, \dots, n$  与第  $j$  个神经元的权值, 在本文的应用中选择  $n=2$ ,  $b_j$  为阈值,  $y_j$  为第  $j$  个神经元的输出, 那么有

$$s_j = \sum_{i=1}^n w_{ji} \cdot x_i + b_j \quad (1)$$

$$y_j = f(s_j) = f\left(s_j = \sum_{i=1}^n w_{ji} \cdot x_i + b_j\right) \quad (2)$$

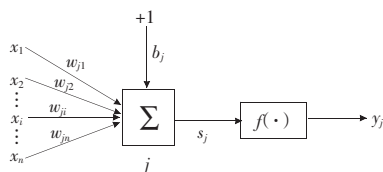


图 2 神经元节点图

输入 659 组训练数据, 用  $x^1, x^2, x^i, \dots, x^{659}$  表示, 第  $p$  组数据输入网络后的输出为  $y^p$ , 采用平方型误差函数, 那么第  $p$  个样本的误差为:

$$E_p = \frac{1}{2} \cdot (y^p - t^p)^2 \quad (3)$$

那么全局误差为:

$$E = \sum_{p=1}^{659} \frac{1}{2} \cdot (y^p - t^p)^2 \quad (4)$$

文献[1]给出了权值的调整公式, 这里推导过程从略。

$$\Delta v_{ki} = \sum_{p=1}^{659} \eta (t^p - y^p) f_2'(S) w_{jk} f_1'(S_k) x_i \quad (5)$$

其中神经元节点的阈值调整推导过程和权值完全相同。通过输入训练数据 → 计算输出误差 → 调整权值、阈值 → 输入训练数据 → 计算输出误差的循环过程调整权值、阈值, 直至输出误差

满足要求, 完成 BP 网络模型的训练, 此时可以提取网络各神经元的参数, 包括权值、阈值、传输函数和连接关系, 以二维输入  $(x_i, y_i)$  为变量, 构成以权值、阈值为系数, 以传输函数为拟合函数的函数关系式, 得到曲面拟合方程。

### 4 Back-Propagation 曲面拟合算法与性能

根据上述的 BP 神经网络原理建立双隐含层的 BP 网络, 两个隐含层的节点数分别为 10 个和 6 个, 输出节点设置为 1 个, 输入层、两个隐含层的传输函数分别设置为 tansig、logsig 和 purelin, 训练函数选择 traingdm, 函数语句如图 3 所示。

```
net=newff([-1 1;-1 1],[10,6,1],{'tansig','logsig','purelin'},'traingdm');
net.trainParam.show=5000;
net.trainParam.lr=0.05;
net.trainParam.epochs=5000;
net.trainParam.goal=0.001;
```

图 3 BP 网络模型主要源码

图 4 给出了 659 组数据的训练误差图, 可以看出, 随着训练次数的增加, 输出误差在逐渐减小。经过 5 000 次训练后, 输出误差为 0.003 65, 这和常见的理想数据的输出误差有较大的差距, 后者一般在 0.000 1 左右, 前者是后者的 300 倍左右。从理论意义上说, 该训练模型是不准确的, 但是该模型有一定的应用价值, 首先在没有先验知识的情况下, 使用其他曲面拟合方法无法解决该拟合问题; 其次, 虽然网络模型输出误差较大, 但工程应用表明其精度满足工程要求, 下面给出若干数据图形来说明这个结论。

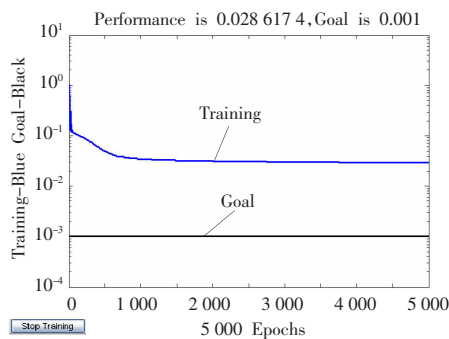


图 4 BP 网络训练误差图

图 5 给出了 BP 网络模型对样本数据的拟合效果, 图 6 给出了样本数据经过选择性删除后使用差分拟合方法对样本数据的拟合效果。从图 5 可以看出, 训练后的模型对输入样本数据的拟合效果  $(x_i, y_i)$  的输出和期望输出  $z_i$  仍然存在差异, 从这一点上来说

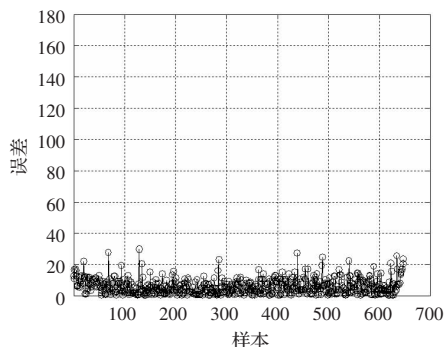


图 5 BP 模型拟合样本数据误差 (下转 239 页)