

# 全国第六届研究生数学建模竞赛



题 目

## 我国城镇登记失业率的数学建模

摘

要：

本文在深入分析了在中国国情下可能对城镇登记失业率产生影响的因素，建立了基于主成分分析法的 BP 神经网络模型，并得到很好的仿真预测效果。

对于问题一，选取了 1989 年到 2008 年间涉及消费、投资、政府购买、进出口等方面的 11 个指标，首先对指标分别进行定性分析，然后应用主成分分析法定量分析了指标之间得相关关系，并得到 3 个主要成分。

对于问题二，通过 BP 神经网络建立了城镇登记失业率与主成分之间关系的模型，并与多元线性回归模型、逐步回归模型和最小二乘法比较，发现 BP 神经网络的仿真效果比较好，绝对平均误差为 1.23%，相关系数为 0.9977。

对于问题三，在分地区研究失业率与其影响因素时，选取了有代表性的 6 个地区，分别是黑龙江、北京、上海、广东、四川和新疆。为了使模型更精确，根据城镇登记失业率所具有的地区特性，在分地区研究城镇登记失业率时，增加考虑了一些指标，最后选取 19 个指标，建立了基于主成分分析法的 BP 神经网络模型，显示出各地区城镇登记失业率模型差异性较大。

对于问题四，在考虑了国家的有关决策和规划后，重点选取了与之有关的 GDP、出口总额、固定资产的投资在 07 年到 08 年的月份和季度数据，最后经过仿真，预测出在 2009 年和 2010 年上半年失业率分别为 4.25 和 4.14，表明我国失业率还会呈上升趋势，但是在 2010 年会有所改善。

最后，根据 BP 神经网络的预测结果，以及我国的有关决策和规划，对提高失业率提出了关于统计数据、单目标发展、多目标之间协调的建议。

本文所采用的模型具有较强的适用性和普遍性，对其它经济指标的分析和预测都有一定的指导意义。

**关键词：**主成分分析法，BP 神经网络，失业率仿真预测

**参赛队号** 1048611

**队员姓名** 叶华 朱婵元 褚召辉

**参赛密码** \_\_\_\_\_

**(由组委会填写)**

## 目录

1. 问题重述.....	3
2. 符号说明.....	4
3. 问题分析与假设.....	4
4. 城镇失业率主要影响因素.....	5
4.1 城镇失业率影响因素分析.....	5
4.1.1 经济增长.....	5
4.1.2.通货膨胀.....	6
4.1.3 进出口额.....	7
4.1.4.产业结构.....	8
4.1.5.固定资产投资和最终消费支出.....	9
4.1.6.劳动力人口数量和劳动需求供给.....	9
4.1.7.技术进步.....	10
4.2 城镇失业率主要影响因素模型.....	11
4.2.1 主成分分析法赋权的原理.....	14
4.2.2 主成分分析的计算.....	14
5. 城镇登记失业率的神经网络预测模型的建立.....	18
5.1 人工神经网络模型.....	18
5.2 城镇登记失业率的神经网络预测模型的建立.....	19
5.3 模型的计算.....	21
6. 分地区建立关于城镇登记失业率的数学模型.....	24
6.1 各省份数据.....	24
6.2 数据的预处理.....	25
6.3 模型的计算.....	26
7. 预测模型的建立.....	30
8. 针对我国现状提出相关意见.....	32
9. 模型评价.....	33
10. 参考文献.....	34

## 1. 问题重述

失业、经济增长和通货膨胀为宏观经济中特别重要的三个指标，就业（或者失业）是社会、国民经济中极其重要的问题。按照已有研究，就业可以定义为三个月内有稳定的收入或与用人单位有劳动聘用关系。失业的统计方法各国差异较大，我国采用城镇登记失业率，是指城镇登记失业人数同城镇从业人数与城镇登记失业人数之和的比。其中，城镇登记失业人员是指有非农业户口，在一定的劳动年龄内（16 岁以上及男 50 岁以下、女 45 岁以下），有劳动能力，无业而要求就业，并在当地就业服务机构进行求职登记的人员。但由于统计口径不同，存在一定的差异，有些历史数据也较难获得。

从经济学的角度，影响就业（或者失业）的因素很多。从宏观层面上，消费、投资、政府购买和进出口都是重要的因素；而从中观层面，不同地区、不同产业也会表现出不同的特征。当然，中央政府调整宏观经济政策（包括财政政策和货币政策），以及对不同地区和不同产业实行不同的扶持政策都会对就业产生巨大的影响。

就我国的现实情况，2008 年我国经济社会经受了历史罕见的考验，GDP 依然保持 9% 以上平稳较快增长，城镇新增就业 1113 万人，城镇登记失业率为 4.2%。2009 年我国就业面临更大的挑战，一是国际金融危机导致国际市场需求难以在短期内复苏；二是今年我国经济增长速下滑；三是国内消费需求乏力；四是一些行业产能过剩与市场预期不确定导致企业投资不足，所以就业形势十分严峻。

为此，中央政府从 08 年 10 月开始实施了 40000 亿元的投资计划，确定了十大产业振兴计划，采取扩大国内消费需求的措施，提高对外开放水平以增加出口。同时，中央财政拟投入 420 亿元资金实施积极的就业政策。09 年我国在就业方面的目标：城镇新增就业 900 万人以上，城镇登记失业率控制在 4.6% 以内（以上数据取自温家宝总理的政府工作报告）。

利用近年来我国有关的统计数据并结合一年多来我国国民经济的运行数据（参见下面网站，也可以对比其他国家的统计数据）就我国就业人数或城镇登记失业率研究如下问题。

1. 对有关统计数据进行分析，寻找影响就业的主要因素或指标。
2. 建立城镇就业人数或城镇登记失业率与上述主要因素或指标之间联系的数学模型。
3. 对上述数学模型从包含主要的经济社会指标、分行业、分地区、分就业人群角度，尝试建立比较精确的数学模型。（由于时间限制，建议适度即可）
4. 利用所建立的关于城镇就业人数或城镇登记失业率的数学模型，根据国家的有关决策和规划对 2009 年及 2010 年上半年的我国就业前景进行仿真（可以根据模型的需要对未来的情况作适当的假设）。
5. 根据所建立的数学模型和仿真结果，对提高我国城镇就业人口数或减少城镇登记失业率提出建议。

## 2. 符号说明

符号	在文中的意义	符号	在文中的意义
GDP	国内生产总值	$X_i$	与失业率有关的指标
$A$	主成分的荷载矩阵	$F_i$	主成分
$M$	各主成分对应的权矩阵	$Y$	主成分的综合得分
MAPE	绝对平均误差	R	相关系数
$Z$	输出数据可信度		

## 3. 问题分析与假设

### 3.1 问题分析

从经济学的角度，影响就业（或者失业）的因素很多。从宏观层面上，消费、投资、政府购买和进出口都是重要的因素；而从中观层面，不同地区、不同产业也会表现出不同的特征。当然，中央政府调整宏观经济政策（包括财政政策和货币政策），以及对不同地区和不同产业实行不同的扶持政策都会对就业产生巨大的影响。

就我国的现实情况，2008 年我国经济社会经受了历史罕见的考验，GDP 依然保持 9% 以上平稳较快增长，城镇新增就业 1113 万人，城镇登记失业率为 4.2%。2009 年我国就业面临更大的挑战，一是国际金融危机导致国际市场需求难以在短期内复苏；二是今年我国经济增速下滑；三是国内消费需求乏力；四是一些行业产能过剩与市场预期不确定导致企业投资不足，所以就业形势十分严峻。

所以讨论就业和失业的问题就成为了必须提上日程的问题。

首先讨论与城镇登记就业率有关的各种因素，因为因素很多，所以本文选择了一些从宏观层面上影响就业率的指标，比如经济的增长、最终消费的支出、固定资产投资、进出口额、通货膨胀、产业结构、劳动力人口数量、劳动力需求供给以及技术进步等等。

如果本文讨论这些指标在一个长期，大概 20 年左右，对城镇登记失业率的影响，这个数据量是非常庞大的，计算有一定得困难，再者本文选择的指标中肯定存在着信息冗余的现象，用人工很难具体的操作，鉴于此，本文采用主成分分析法对搜集到的数据进行处理，并力保原始数据信息丢失最少的情况下，对高维变量空间进行降维。

在多因素与失业率之间的关系建模中，可以考虑神经网络、多元回归模型、

逐步回归模型和最小二乘拟合等方法，但神经网络有其他方法不具备的优点，既可以表现出非线性的映射关系，又可以以尽可能高的精度逼近相应的映射，所以在应用中选择使用最广泛的 BP 神经网络模型。

在使用 BP 神经网络建模过程中，本文利用主成分分析法将原指标替换成维数比较小但仍能保持数据信息完整性的主成分，这样可以使运算更有效率，得到的结果更精确。

在第三问中本文对地区进行分类，选取在中国区域上比较有代表性的几个省份进行建模。建模的过程与上一步类似，只是因为本文分的比较细，所以从整个中国来看，可以更精确的得到想要得到的数据。

下面的两问是在上面建立的模型基础上得到的进行仿真预测，并针对相应的预测提出合理化建议。

### 3.2 假设

- 1、假设未来不会有大的对经济产生巨大影响的自然灾害。
- 2、忽略本文选取因素之外的因素对问题产生的影响。
- 3、不考虑严重影响国内正常生产生活的不可抗力因素。
- 4、不考虑自然失业率的影响。

## 4. 城镇失业率主要影响因素

### 4.1 城镇失业率影响因素分析

从总体上讲，我国的失业状况与劳动力人口数量、经济发展状况等有着密切的关系，具体而言，宏观经济层面上的经济的增长、消费、投资、进出口额、通货膨胀，以及劳动力市场方面的劳动力人口数量和劳动力需求供给，包括产业结构和技术进步等，都会对失业率产生重要影响。

本文考虑的影响城镇登记失业率的因素为：经济的增长、最终消费的支出、固定资产投资、进出口额、通货膨胀、产业结构、劳动力人口数量、劳动力需求供给以及技术进步。

#### 4.1.1 经济增长

从世界各国的经济发展的经验来看，经济增长和失业是有某种程度的相关关系的。

在我国的宏观经济统计数据中，可以用国内生产总值的增长速度来代表经济增长速度，这也是国际上通用的指标。

图 1 是我国在 1989–2008 年，这 20 年的 GDP 增长率变化趋势和失业率变化趋势的比较：

经济增长和失业之间的关系是很复杂的，在研究经济增长和失业之间的关系

中，著名的奥肯定律指出 GDP 增长和失业的负相关关系，当然，这一结论是根据美国的统计资料估计出来的，它对于美国的 GDP 增长和失业关系是准确的，但是，从图中可以看出，对于我国的特殊国情而言，我国的经济增长和失业之间的关系呈现非一致性，。

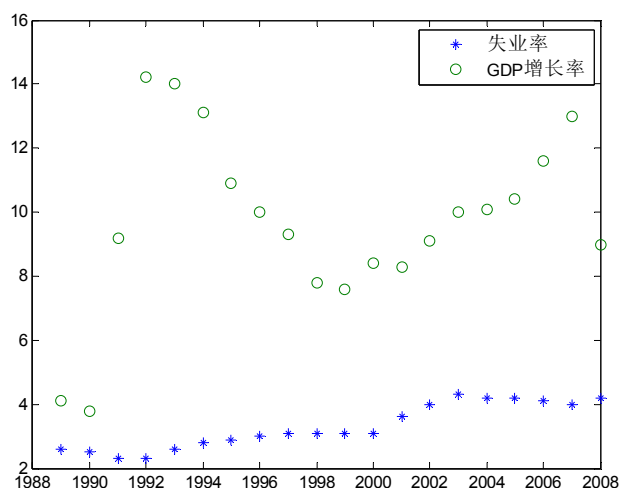


图 1 GDP 增长率变化趋势和失业率变化趋势的比较

中国经济处于转型期间，多种矛盾交织，市场规则处于薄弱阶段，单纯根据 GDP 增长和失业率两者不能得到较为准确的分析框架，必须考虑多种因素的相互交织影响，因此，需要在综合考虑其他因素的情况下，运用相关年限的数据定量分析经济增长与失业的关系。

#### 4.1.2. 通货膨胀

通货膨胀会导致企业利润的上升，从而扩大生产规模，使就业人数增加。本文使用商品零售价格指数对通货膨胀率进行量化，初步分析通货膨胀率和失业率之间的关系，图 2 为我国在 1989—2008 年，这 20 年的商品零售价格指数和失业率之间的关系。

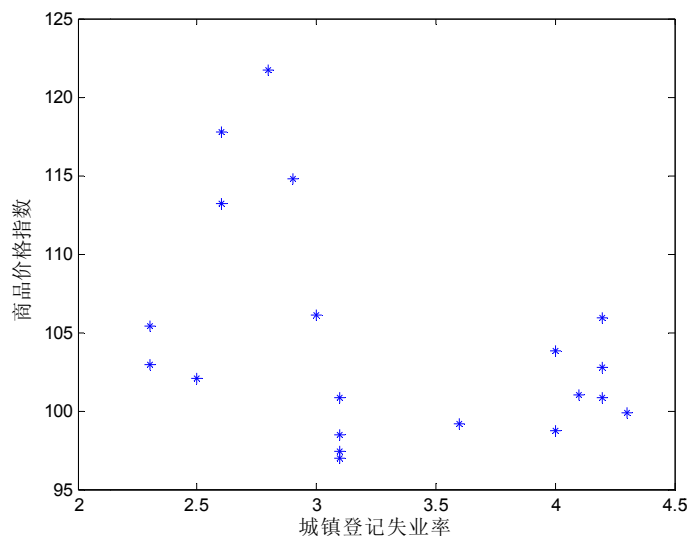


图 2 商品零售价格指数和失业率之间的关系

一般而言，根据经济理论可知，衡量通货膨胀与失业率之间的关系的是菲利普斯曲线<sup>[1]</sup>。

菲利普斯曲线是表示通货膨胀率与失业率之间相互关系的曲线，其表明失业率与通货膨胀率存在负相关关系：失业率越低，通货膨胀率越高；失业率越高，通货膨胀率越低。

菲利普斯曲线表示为图 3。

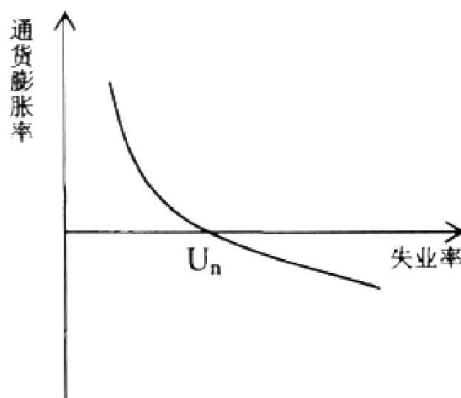


图 3 菲利普斯曲线

从图 2 可以看出，我国的城镇登记失业率和消费价格指数之间不符合菲利普斯曲线。

#### 4.1.3 进出口额

进出口额在我国经济中占有重要地位，特别是由于劳动密集产品在我国出口中占有很大比重，和就业高度相关。

对 1989—2008 年的进口额、出口额进行初步分析，发现两者的变化呈现相一致的情况，图 4 为 1989—2008 年的进口额、出口额在 GDP 中的比重变化趋势，

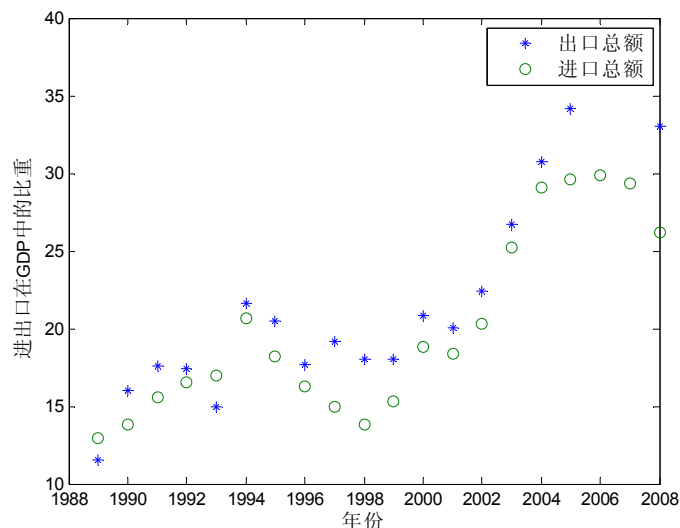


图 4 进出口额变化趋势

从整体上看，进出口额呈增加的趋势，由于近几年的金融危机，进出口额短期内减少。

创造就业机会与扩大产品出口密切相关，出口的增加等于有效地扩大了我国的就业空间，迅速发展吸收了大量劳动力，特别了经济的发展，进口对部分行业的影响也很大，增加进口和扩大出口二者是辩证的统一。

#### 4.1.4.产业结构

不同的产业结构比例体现了一个国家的不同的资本有机构成，产业结构中对失业率影响较为大的两个因素就是第二产业和第三产业。

第二产业、第三产业能够吸收从第一产业中转移出来的劳动力，从而对失业率产生重要影响。第二产业和第三产业就业容量大，如果能够增加就业岗位，就能拓宽就业渠道。

图5 是1989年~2008年这20年内第二、三产业对GDP的贡献率（百分数）

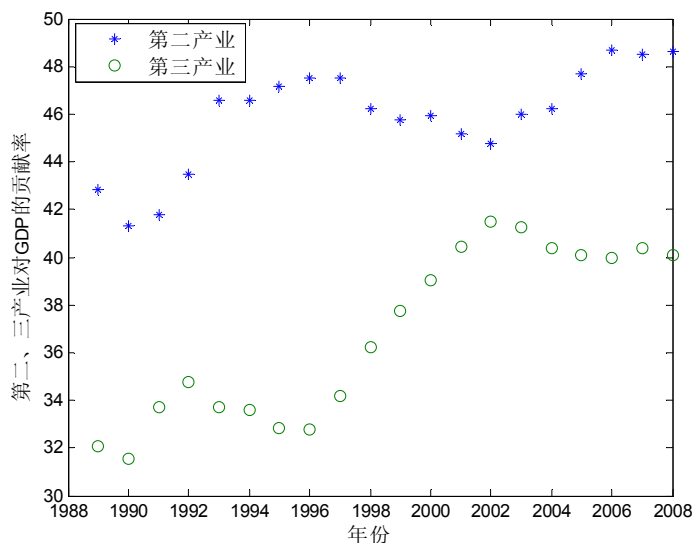


图 5 第二、三产业对 GDP 的贡献率



图 5 反映了近 20 年来我国产业结构基本趋势的。1988 年，第二产业和第三产业相差较大，第二产业发展较为平缓，第三产业发展速度较快，特别在 1996 年发展速度加快，近几年发展速度减慢。

#### 4.1.5.固定资产投资和最终消费支出

固定资产投资的增长不仅通过对劳动力的即时需求而扩大当时的就业，而且也会通过项目的逐步完成和形成生产力，而对就业产生“时滞性”的利好。

图 6 是 1989 年~2008 年这 20 年内固定资产投资在 GDP 中的比重与失业率的关系：

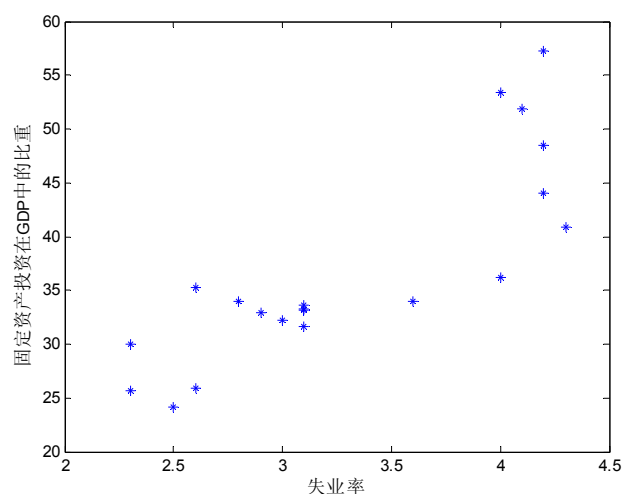


图 6 固定资产投资在 GDP 中的比重与失业率的关系

由于多方面因素的交织，有些阶段两者之间的关系不太明显，两者没有显著的关系。消费是生产的起点和目标，消费扩大了才有生产的扩大、就业的增加。最终消费支出反映消费需求，包括居民消费支出和政府消费支出，居民对物质的大量需求，会使相关行业保持增长，吸纳更多的劳动力，政府消费扩张在就业部门中也有非常显著的效应。

#### 4.1.6.劳动力人口数量和劳动需求供给

劳动力人口数量<sup>[2]</sup>决定着劳动力的供给，从市场需求<sup>[3]</sup>上来说，失业是由于商品市场的过度供给导致劳动力市场的过度供给造成的，失业的增加是因为商品市场的总需求下降而引起的，只有通过刺激有效需求才能真正地降低失业率。

本文中用城镇居民实际工资指数，间接反映劳动需求供给，工资是指劳动力的价值或价格，实际工资则为经过商品价格指数调整后的货币工资。由于城镇登记失业率主要是通过城镇登记失业人数进行统计，所以这里用城镇人口所占比例，反映劳动力人口数量。

通过 1989 年~2008 年多时期横截面数据，初步分析城镇居民实际工资、城镇人口所占比例和失业率之间的关系。

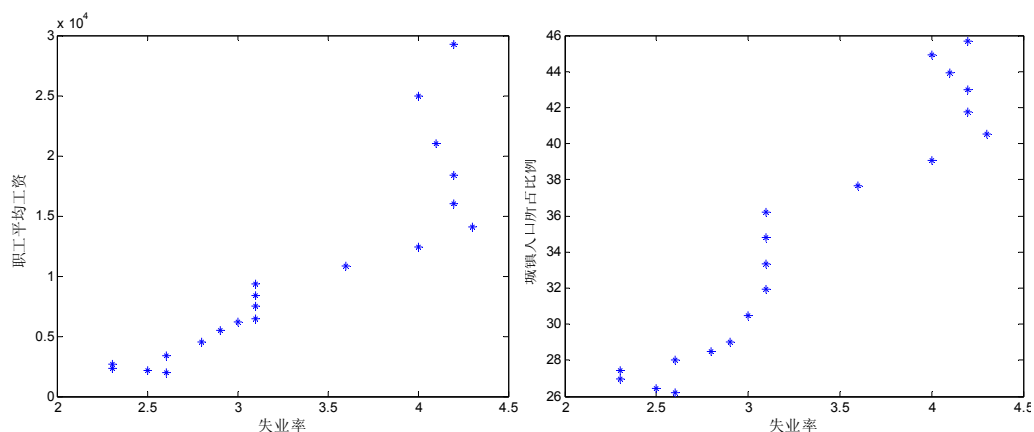


图 7 工资和失业率之间的关系      图 8 城镇人口比例与失业率之间的关系

从整体上来看，城镇居民实际工资的提高，会使失业率增加，特别是当工资高于一定阶段时，失业率就保持在一个较高的阶段。

在完全自由竞争的市场经济条，工资水平的高低完全取决于劳动力的供求关系，而在国家宏观调控的现代市场经济条件下，工资水平除了受市场影响之外，还受国家政策和工会组织的影响，如最低工资法以及工会与企业主签定的劳动合同等。因此，在现代市场经济条件下，工资水平不单被动地受劳动力供求关系的影响，而且也会主动地发挥作用，影响劳动力的供给与需求。

城镇人口所占比例增加，也会引起失业率的增加，这是因为城市化引起农村人口转化为城镇人口，从而增加了城镇劳动力供给，由此可能造成城镇失业率上升。

#### 4.1.7. 技术进步

近代西方失业理论体系中，技术进步也被认为是影响失业率<sup>[4]</sup>的一个重要因素，这在中国也是成立的。

从就业容量和就业结构两方面来说，技术变化使劳动生产率提高，技术进步大大节约生产过程中的劳动使用数量，消灭了许多传统岗位，使社会对劳动力的需求下降，同时，技术变化引起产业结构的变动，又间接地影响就业结构的变动，劳动就业的主导方向从第一产业转移到第二产业再转移到第三产业。技术变化同时促进各个产业、各部门内部劳动就业结构发生变动。工业生产部门内部就业由传统技术部门向现代技术部门以及高新技术部门转移。发展现代技术部门和高新技术部门，是当今技术变化的必然结果，也是积极有效解决就业的重要途径。技术变化使得第三产业内部就业结构发生变化，由传统服务业向社会服务业和生产服务业转化。

一方面，由于生产效率的提高，技术进步使社会对劳动力的需求下降，对失业率有提高作用；另一方面，技术进步创造了更多的新型部门，使得经济总体对劳动的需求增加，又会使失业率降低。

不管是从短期来看，还是从长远来看，技术进步对失业的影响是一个很复杂的现象。

## 4.2 城镇失业率主要影响因素模型

失业、经济增长和通货膨胀为宏观经济中特别重要的三个指标，就业（或者失业）是社会、国民经济中极其重要的问题。按照已有研究，就业可以定义为三个月内有稳定的收入或与用人单位有劳动聘用关系。失业的统计方法各国差异较大，我国采用城镇登记失业率，是指城镇登记失业人数同城镇从业人数与城镇登记失业人数之和的比。

从经济学的角度，影响就业（或者失业）的因素很多。从宏观层面上，消费、投资、政府购买和进出口都是重要的因素；而从中观层面，不同地区、不同产业也会表现出不同的特征。当然，中央政府调整宏观经济政策（包括财政政策和货币政策），以及对不同地区和不同产业实行不同的扶持政策都会对就业产生巨大的影响。

由 4.1 的介绍并考虑到我国国内对失业产生影响的因素，本文考虑的影响城镇登记失业率的因素为：经济的增长、最终消费的支出、固定资产投资、进出口额、通货膨胀、产业结构、劳动力人口数量、劳动力需求供给以及技术进步。经过查找国家权威数据，本文选取了 11 个指标作为与失业率有关的因素。这些因素分别为城镇登记失业率、职工平均工资、城镇人口所占比例、商品零售价格指数、GDP 增长率、第二产业 GDP/GDP、第三产业 GDP/GDP、出口总额/GDP、进口总额/GDP、固定资产投资/GDP、技术市场成交额/GDP、最终消费支出贡献率。

表 1 各指标在 1989 年至 2008 年之间的数据

数据来源：《中国统计年鉴》

指标 年份	城镇 登记 失业率/%	职工平 均工资 /元	城镇人 口所占 比例/%	商品零 售价格 指数	GDP 增 长率	第二产业 GDP/GDP	第三产 业 GDP/GD P	出口总 额/GDP	进口总 额/GDP	固定资 产投资 /GDP	技术市 场成交 额/GDP	最终消 费支出 贡献率
1989	2.6	1935	26.21	117.8	4.1	42.8311	32.0639	11.5111	12.9464	25.9551	47.96	39.6
1990	2.5	2140	26.41	102.1	3.8	41.3407	31.5432	15.9944	13.79	24.1967	40.2	47.8
1991	2.3	2340	26.94	102.9	9.2	41.7887	33.685	17.5704	15.6036	25.6846	43.2	65.1
1992	2.3	2711	27.46	105.4	4.2	43.4546	34.7555	17.3689	16.5034	30.0113	56.01	72.5
1993	2.6	3371	27.99	113.2	14	46.5684	33.7232	14.9567	16.9418	35.2576	58.7	59.5
1994	2.8	4538	28.51	121.7	13.1	46.5693	33.5695	21.623	20.665	33.9648	47.47	30.2
1995	2.9	5500	29.04	114.8	10.9	47.175	32.8627	20.482	18.1731	32.9298	44.08	44.7
1996	3	6210	30.48	106.1	10	47.5366	32.7723	17.6693	16.2376	32.2775	42.15	60.1
1997	3.1	6470	31.91	100.8	9.3	47.539	34.1739	19.1973	14.95	31.5818	44.45	37
1998	3.1	7479	33.35	97.4	7.8	46.2122	36.2318	18.0464	13.7746	33.6557	50.95	57.1

指标 年份	城镇 登记 失业率/%	职工平 均工资 /元	城镇人 口所占 比例/%	商品零 售价格 指数	GDP 增 长率	第二产业 GDP/GDP	第三产 业 GDP/GD P	出口总 额/GDP	进口总 额/GDP	固定资 产投资 /GDP	技术市 场成交 额/GDP	最终消 费支出 贡献率
1999	3.1	8346	34.78	97	7.6	45.7571	37.7727	18.02	15.3176	33.2914	58.54	74.7
2000	3.1	9371	36.22	98.5	8.4	45.9165	39.0204	20.7978	18.7864	33.1783	63.5	65.1
2001	3.6	10870	37.66	99.2	8.3	45.1527	40.4556	20.0851	18.3842	33.9368	72.96	50
2002	4	12422	39.09	98.7	9.1	44.7898	41.4675	22.3945	20.3023	36.1497	73.13	43.6
2003	4.3	14040	40.53	99.9	10	45.969	41.2337	26.7171	25.1766	40.9111	79.71	35.3
2004	4.2	16024	41.76	102.8	10.1	46.2253	40.3815	30.7129	29.0445	44.0819	83.44	38.7
2005	4.2	18364	42.99	100.8	10.4	47.6836	40.0796	34.1933	29.6226	48.4526	82.42	38.2
2006	4.1	21001	43.9	101	11.6	48.6789	39.9774	36.6144	29.9055	51.9047	85.79	38.7
2007	4	24932	44.94	103.8	13	48.5022	40.3722	37.4527	29.3691	53.3699	85.5	39.4
2008	4.2	29229	45.68	105.9	9	48.6191	40.0728	33.0069	26.1599	57.3024	88.64	39.7

#### 4.2.1 主成分分析法赋权的原理

主成分是利用降维思想,把多指标转化为少数几个综合指标的多元统计分析方法.这些主要指标是原指标的线性组合,且彼此之间互不相关,它可以在力保原始数据信息丢失最少的情况下,对高维变量空间<sup>[5]</sup>进行降维而得到。

##### 主成分分析法原理及数学模型

假设观测指标共有  $p$  个,分别为  $x_1, x_2, \dots, x_p$ , 将这些指标综合为一个综合指标的方法很多,最简单的方法是将这些指标用线性的方法将它们组合起来,即

$$y = a_1x_1 + a_2x_2 + \dots + a_px_p \quad (4.1)$$

由于各指标组合系数的不同,就得到不同的综合指标,其中反映原指标变动程度最大的综合指标称为第一主成分,次之为第二主成分, ..., 第  $k$  个综合指标为第  $k$  个主成分;若记原始观测变量  $x_1, x_2, \dots, x_p$ , 第一主成分为  $y_1$ , 第二主成分为  $y_2$ , ..., 第  $k$  主成分为  $y_k$ , 则有

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p, \\ &\dots \\ y_k &= a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kp}x_p. \end{aligned} \quad (4.2)$$

并且有  $\text{var}(y_1) > \text{var}(y_2) > \dots > \text{var}(y_k)$ 。由此可见,主成分分析法就是在保证原始数据信息丢失最小的前提下,使原始变量减少为具有代表意义的少数的几个新的变量,它们更集中更典型地反映了研究对象的特征<sup>[6]</sup>。

#### 4.2.2 主成分分析的计算

用表 1 中除了失业率之外的所有指标进行主成分分析法的分析,本文中的程序均在 MATLAB 7.4 环境下运行。

表 2 方差分解的主成分分析表

component	特征值	贡献率	累计贡献率
1	7.16272	0.6512	0.6512
2	1.6988	0.1544	0.8056
3	1.0416	0.0947	0.9003
4	0.487295	0.0443	0.9446
5	0.278147	0.0253	0.9699
6	0.219297	0.0199	0.9898
7	0.0675421	0.0061	0.9959
8	0.0340958	0.0031	0.9990
9	0.00563504	0.0005	0.9995
10	0.00429897	0.0004	0.9999
11	0.000575602	0.0001	1.0000

由上表可知贡献率大于 90%时,主成分数为 3,即有 3 个主成分来代替原来的 11 个指标。

表 3 初始因子载荷矩阵  $A$ 

指标	主成分 1	主成分 2	主成分 3
X <sub>1</sub> 职工平均工资	-0.9612	0.0651	-0.0811
X <sub>2</sub> 城镇人口所占比例	-0.9695	0.2213	-0.0523
X <sub>3</sub> 商品零售价格指数	0.3141	-0.87	-0.1722
X <sub>4</sub> GDP 增长率	-0.3475	-0.562	0.6917
X <sub>5</sub> 第二产业 GDP/GDP	-0.7139	-0.3653	0.2178
X <sub>6</sub> 第三产业 GDP/GDP	-0.8537	0.3978	0.0509
X <sub>7</sub> 出口总额/GDP	-0.9582	-0.0834	-0.0347
X <sub>8</sub> 进口总额/GDP	-0.939	-0.1618	-0.0334
X <sub>9</sub> 固定资产投资/GDP	-0.9646	-0.1358	0.0303
X <sub>10</sub> 技术市场成交额/GDP	-0.9376	0.1565	0.0007
X <sub>11</sub> 最终消费支出贡献率	0.4811	0.4529	0.6862

用初始因子载荷矩阵中的数据除以主成分相对应的特征根开平方根便得到两个主成分中每个指标所对应的系数。得到的3个主成分如下：

$$F = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = M^T X = \begin{pmatrix} -0.3663 & 0.1958 & -0.2790 \\ -0.3679 & 0.3609 & -0.2241 \\ 0.2094 & -0.7156 & -0.4066 \\ -0.2203 & -0.5752 & 0.8149 \\ -0.3157 & -0.4637 & 0.4573 \\ -0.3452 & 0.4839 & 0.2211 \\ -0.3658 & -0.2216 & -0.1825 \\ -0.3621 & -0.3086 & -0.1791 \\ -0.3670 & -0.2827 & 0.1706 \\ -0.3618 & 0.3035 & 0.0259 \\ 0.2592 & 0.5163 & 0.8117 \end{pmatrix}^T \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \\ X_8 \\ X_9 \\ X_{10} \\ X_{11} \end{pmatrix} \quad (4.3)$$

用第一主成分  $F_1$  中每个指标所对应的系数乘上第一主成分  $F_1$  所对应的贡献率再除以所提取两个主成分的两个贡献率之和，然后加上  $F_2$ 、 $F_3$  中每个指标所对应的系数乘上第二主成分  $F_2$ 、 $F_3$  所对应的贡献率再除以所提取两个主成分的两个贡献率之和，即可得到综合得分  $Y$  模型：

$$Y = -0.2607X_1 - 0.2278X_2 - 0.01140X_3 - 0.1722X_4 - 0.2598X_5 - 0.1435X_6 - 0.3218X_7 - 0.3337X_8 - 0.2960X_9 - 0.2069X_{10} + 0.3614X_{11} \quad (4.4)$$

第一主成分的贡献率为 65.12%，其中  $X_{11}$  和  $X_3$  与第一主成分之间是正相关的关系并且在第一主成分上有较高的载荷，这两个指标分别是最终消费支出贡献率和商品零售价格指数，两者都是与消费有关的指标，而失业率在消费、物价等都比较高的情况下是与其成正相关的关系， $X_1$  职工平均工资、 $X_2$  城镇人口所占比例、 $X_9$  固定资产投资/GDP、 $X_7$  出口总额/GDP、 $X_8$  进口总额/GDP、 $X_{10}$  技术市场成交额/GDP、 $X_6$  第三产业 GDP/GDP、 $X_5$  第二产业 GDP/GDP 这些指标在第一主成分占有很高的负相关的荷载，它们都是经济学上对经济发展起到很重要作用的指标，对失业率都有着很大的影响。

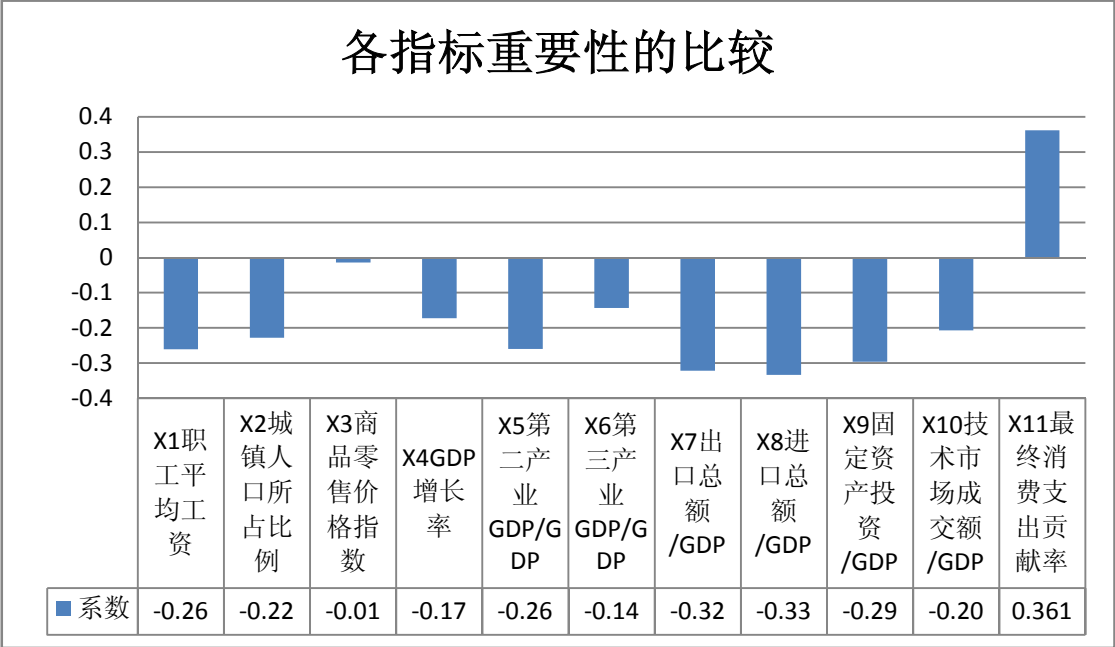
第二主成分的贡献率为 15.44%，与其正相关的占有较高载荷的是  $X_{11}$  和  $X_6$ ，说明消费和第三产业在经济增长中对于吸收劳动人口起到了积极地作用，这与本文在新闻报道中得到的信息一致。

第三主成分的贡献率为 9.47%，随相比前两者较小，但仍起到了很重要的作用，与其正相关且荷载比较高的是  $X_{11}$  和  $X_4$ ，这两者一个为最终消费支出贡献率



一个为 GDP 增长率，这两者一个反应国民消费，一个反应国民经济增长，它们与失业率的关系在前面也已讨论过。

表 4 各指标重要性的比较



由表中系数的比较可知，无论是正相关还是负相关，本文选取的指标基本上都是可以反映自身的信息量的。

表 5 主成分的得分

年份	主成分 1	主成分 2	主成分 3	年份	主成分 1	主成分 2	主成分 3
1989	-0.2106	-0.0394	0.0415	1999	-0.2901	-0.0202	0.0779
1990	-0.2111	-0.0296	0.0465	2000	-0.3219	-0.0274	0.0730
1991	-0.2284	-0.0373	0.0778	2001	-0.3447	-0.0311	0.0614
1992	-0.2553	-0.0487	0.1010	2002	-0.3718	-0.0361	0.0588
1993	-0.2700	-0.0592	0.0913	2003	-0.4192	-0.0455	0.0550
1994	-0.3001	-0.0771	0.0654	2004	-0.4518	-0.0474	0.0562
1995	-0.2828	-0.0603	0.0683	2005	-0.4798	-0.0491	0.0563
1996	-0.2676	-0.0453	0.0766	2006	-0.5087	-0.0525	0.0599
1997	-0.2844	-0.0498	0.0582	2007	-0.5326	-0.0558	0.0636
1998	-0.2828	-0.0312	0.0668	2008	-0.5363	-0.0418	0.0484

## 5. 城镇登记失业率的神经网络预测模型的建立

### 5.1 人工神经网络模型

人工神经网络(artificial neural network)是由大量的、简单元件(神经元)广泛相联结而成的非线性的、动态的复杂网络信息处理系统,它是在现代神经学研究成果基础上提出的,能模拟人脑的若干基本功能<sup>[7]</sup>。它具有并行分布的信息处理结构,可以通过“自学习”或“训练”的方式完成某一特定的工作。它可以从积累的工作案例中学习知识,尽可能多地把各种定性或定量的因素作为变量加以输入,从而建立各种影响因素与结论之间的高度非线性映射,采用自适应模式识别方法来完成预测工作。人工神经网络模型尤其是对处理内部规律不甚了解、不能用一组规则或方程进行描述的复杂的、开放的非线性系统显得较为优越。

人工神经网络模型一般由处理单元、激活状态、单元输出、连接模式、激活规则、学习规则等六部分组成。一个多层神经网络中包含有很多个信息处理单元,分布于不同的层次中。根据每项输入和相应的权重获取一个综合信号,当信号超过阈值则激活神经元而产生输出。各类影响因素和最终输出结果之间可以假定存在一种映射,即输出结果 =  $f$  (影响因素)。为了寻求最佳的映射关系  $f$ , 将训练样本集合和输入、输出转化为一种非线性关系,通过对简单非线性函数的复合,从而建立一个高度的非线性映射关系  $f$ , 最终实现输出值的最优逼近。

在人工神经网络的实际应用中,最多的是采用前馈反向传播网络(back propagation network, 简称 bp 网络)或它的变化形式。

典型的 BP 网络是含有一个隐含层的三层结构的网络。

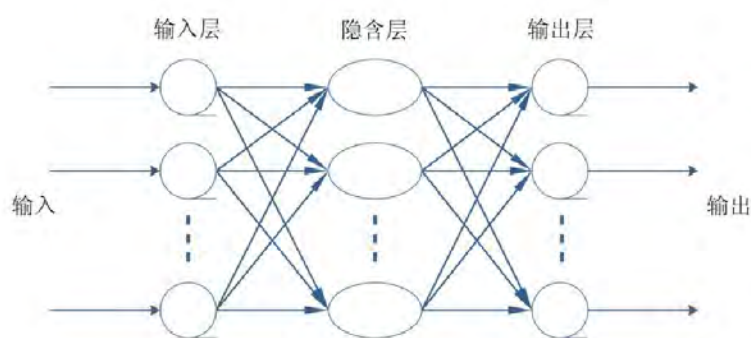


图 9 含有一个隐含层的三层结构的网络

BP 网络学习算法流程如图

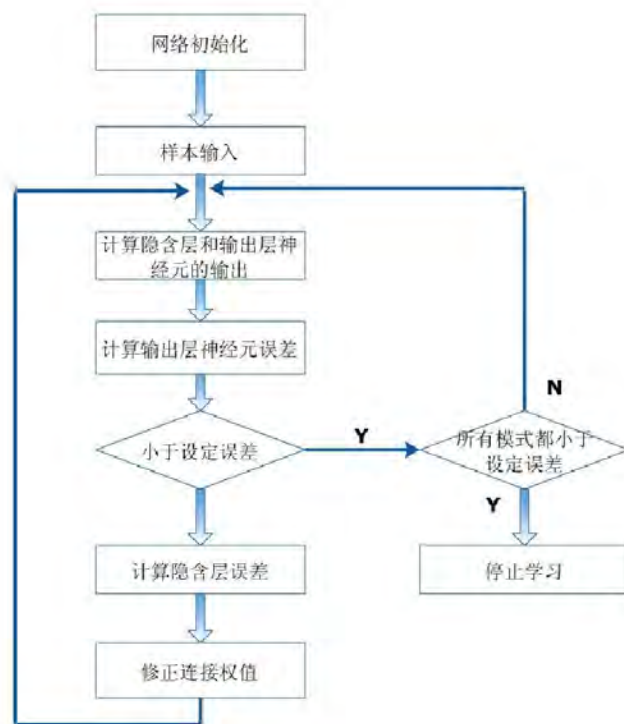


图 10 BP 网络模型算法流程图

## 5.2 城镇登记失业率的神经网络预测模型的建立

本文选取了在上一节中讨论过的 11 个指标来讨论它们与城镇登记失业率之间的关系，在应用中本文选择用主成分分析法将 11 个指标进行降维，用 3 个主成分来代替这 11 个指标。

### BP 网络设计

网络设计是一个综合性问题，它应满足多种不同要求，例如，希望所设计的网络有较好的推广能力，易于硬件实现，训练速度快等等，其中有较好的推广能力是最主要的。一般说来，推广能力决定于 3 个主要因素，即问题本身的复杂程度、网络结构以及样本量大小。在一般情况下，旅游需求预测研究中样本的数量是一定的，因此可归结为在样本量一定的情况下，如何选择网络规模的问题。在进行 BP 网络预测模型设计时，本文中主要考虑了以下因素：网络的层数、每层中的神经元个数、初始值的选择、学习速率和期望误差<sup>[8]</sup>。

为了确保训练后人工神经网络模型的泛化能力和预测能力，应在满足精度要求的前提下取尽可能紧凑的网络结构。楼文高<sup>[9]</sup>指出：一般地，随着隐层节点数的增加，训练样本误差减少，但趋势逐渐减缓（也可能有局部的波动，但总体趋势减小），而检验（测试）样本误差先减小，当发生过拟合或过学习时误差出现增大现象。因此，通过绘制训练、检验（样本）误差和隐层节点数的二维图可确定隐层的节点数。训练误差不再明显下降而又没有发生过拟合时的节点数就是隐层的合理节点数。

## BP 网络模型的建立

在 BP 网络预测模型中，对于网络的输入和输出，可以假设存在一映射  $f$ 。为了寻求  $f$  的最佳映射值，BP 网络模型将样本集合的输入、输出转化为非线性优化，通过对简单的非线性函数的复合，建立一个高度的非线性映射关系，实现  $f$  值的最优逼近。

在本模型中，网络输入为 1989 年到 2008 年的 11 个指标经主成分分析法<sup>[10]</sup>得到的 3 个主成分的得分（数据可由表 5 得到），输出为 1989 年到 2008 年城镇登记失业率数。

在实际应用中本文将各个参数设置为如下：训练次数设置为 10000，训练目标设置为  $1e-4$ ，学习率设置 0.01（该值应设置为较小值，太大会在开始加快收敛）。在 MATLAB 神经网络工具箱中集成了多种训练算法，主要有：附加动量法(traindm.m)、自适应学习速率(trainda.m, traingdx.m)、RPROP 方法(trainrp.m)、共扼梯度法(trainegfm, traingcp.m, trainegb.m, trainseg.m)、拟牛顿法(trainbfg.m, trainoss)以 Levenberg-Marquardt 方法(trainlm.m)，在 BP 网络中，很难确定哪一个问种训练算法是最好的，本模型通过对多种算法的试验，最终选择了 traingda.m 作为训练函数。

在通过训练样本进行网络的训练好后，就是对网络预测性能的考查，即将在数据中选取出来的测试集作为网络的输入，其仿真结果令人满意，达到预期的效果。

在这里介绍一下本文用到的误差分析工具，用绝对平均误差 MAPE、相关系数 R、输出数据可信度 Z 这 3 个参数来评价模型的模拟精度。因为它们既不受样本规模的影响，又不受样本单位的制约，不同模型间可比性强。它们分别可以用下列公式求得：

$$MAPE = \frac{\sum_{i=1}^n |x_i - y_i| / y_i}{n} \times 100\% \quad (5.1)$$

$$R = \frac{\sum_{i=1}^n (x_i \times y_i)}{\sqrt{\sum_{i=1}^n (x_i)^2 \times \sum_{i=1}^n (y_i)^2}} \quad (5.2)$$

$$Z = \frac{\sum_{i=1}^n j}{n} \times 100\% \quad \left( \text{如果 } \frac{|x_i - y_i|}{y_i} \leq 0.15 \text{ 则 } j=1; \text{ 否则 } j=0 \right) \quad (5.3)$$

5.3 模型的求解

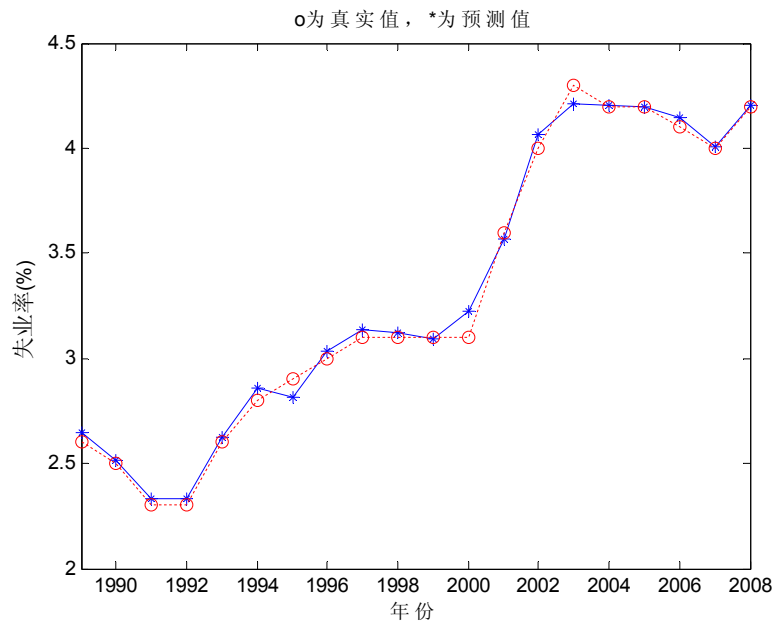


图 11 真实值与预测值曲线

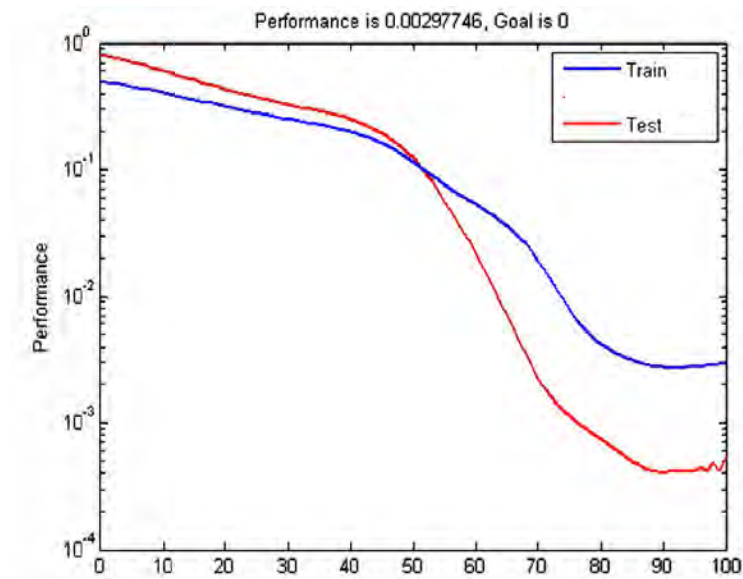


图 12 网络训练误差变化曲线

表 6 城镇登记失业率网络模型输入层到隐含层权值及隐含层阈值

输入层到隐含层权值			隐含层阈值
2.273	-0.8199	0.0327	-1.3165
0.1069	-1.7814	-0.9181	0.208
1.6531	0.4337	-1.0723	2.0005

表 7 城镇登记失业率网络模型隐含层到输出层权值及输出层阈值

隐含层到输出层权值			输出层阈值
2.1783	2.0801	0.3255	-2.9985
-1.9894	-0.9796	-2.0302	2.3078
0.3473	-1.5334	-2.5745	-1.6766
-1.1267	-2.3217	-1.5616	1.0016
-1.3498	-1.4011	2.3001	0.3259
-2.0245	1.5334	1.5342	-0.4557
0.9724	0.1477	2.8492	1.0015
1.6334	1.3813	-2.062	1.7142
-0.2785	-2.5825	1.4689	-2.3959
1.9468	0.528	-2.116	3.1322

下面分别用多元回归模型、逐步回归模型和最小二乘拟合三种方法对城镇登记失业率进行拟合。得到如下结果：

多元回归模型

$$y = 2.8613 - 1.78279 \times 10^{-5} X_1 + 0.1509 X_2 - 0.0158 X_3 + 0.0099 X_4 - 0.0448 X_5 + 0.0095 X_6 - 0.1165 X_7 + 0.1087 X_8 + 0.0394 X_9 - 2.5913 X_{10} - 0.0151 X_{11}$$

逐步线性回归

$$y = -0.3909 + 4.32693 \times 10^{-5} X_1 + 0.1367 X_2 - 0.01265 X_{11}$$

最小二乘拟合（与时间 t 的关系）

$$y = 7.00707 \times 10^{-6} - 4.51985 \times 10^{-4} t + 0.0112 t^2 - 0.1327 t^3 + 0.7831 t^4 - 2.0097 t^5 + 4.0533 t^6$$

它们拟合的结果如下表所示：

表 8 各种拟合的结果比较

年份	真实值	神经网络 仿真值	多元回归 仿真值	逐步回归 仿真值	最小二乘 拟合
1989	2.6	2.6467	2.5954	2.6082	2.7047
1990	2.5	2.5156	2.4438	2.523	2.2692
1991	2.3	2.3317	2.3567	2.3679	2.2886
1992	2.3	2.3342	2.2221	2.3293	2.4752
1993	2.6	2.6286	2.6778	2.5377	2.6699
1994	2.8	2.8575	2.9018	2.929	2.8061
1995	2.9	2.815	2.7078	2.7764	2.8776
1996	3	3.0318	2.9327	2.7478	2.9126
1997	3.1	3.1353	3.1777	3.2243	2.9521
1998	3.1	3.1186	3.1105	3.1232	3.0333
1999	3.1	3.0894	3.0438	3.0586	3.1787
2000	3.1	3.2245	3.2972	3.3326	3.3891
2001	3.6	3.5688	3.5756	3.6557	3.643

年份	真实值	神经网络 仿真值	多元回归 仿真值	逐步回归 仿真值	最小二乘 拟合
2002	4	4.0636	3.9241	3.865	3.9
2003	4.3	4.2082	4.2159	4.0969	4.1099
2004	4.2	4.2047	4.2337	4.1363	4.2263
2005	4.2	4.193	4.2075	4.2095	4.2256
2006	4.1	4.1454	4.0508	4.2135	4.1309
2007	4	4.0034	4.0531	4.1768	4.0414
2008	4.2	4.2047	4.2098	4.0882	4.1659

表 9 BP 神经网络与其他模型误差的比较

	神经网络	多元回归	逐步回归	最小二乘拟合
绝对平均误差 MAPE	1.23%	2.05%	3.00%	2.93%
相关系数 R	0.9977	0.9928	0.9839	0.9852
可信度 Z	100%	100%	100%	100%

MAPE、R、Z 是模型预测值与实际值之间精确度性能的评价指标。其中 MAPE 是绝对平均误差，值越小，表示模型的预测结果越好；R 是相关系数，值越大表示模型的预测结果越好；Z 是输出数据的可信度，值越大表示模型的预测结果越好。从表 3 可以看出，BP 神经网络模型三个参数都比其他模型的预测结果要好，其仿真效果比其他模型好，说明其能更精确地拟合城镇失业率与各指标之间的关系，因此，用 BP 神经网络模型来建立城镇登记失业率与上述 11 个指标之间联系的数学模型是可行的，而且也是比较科学的。

## 6. 分地区建立关于城镇登记失业率的数学模型

### 6.1 各省份数据

由中国的区域分布情况，本文选取了比较具有代表性的 6 个省份的数据来代表中国的基本区域分布，它们是黑龙江、北京、上海、广东、四川和新疆。

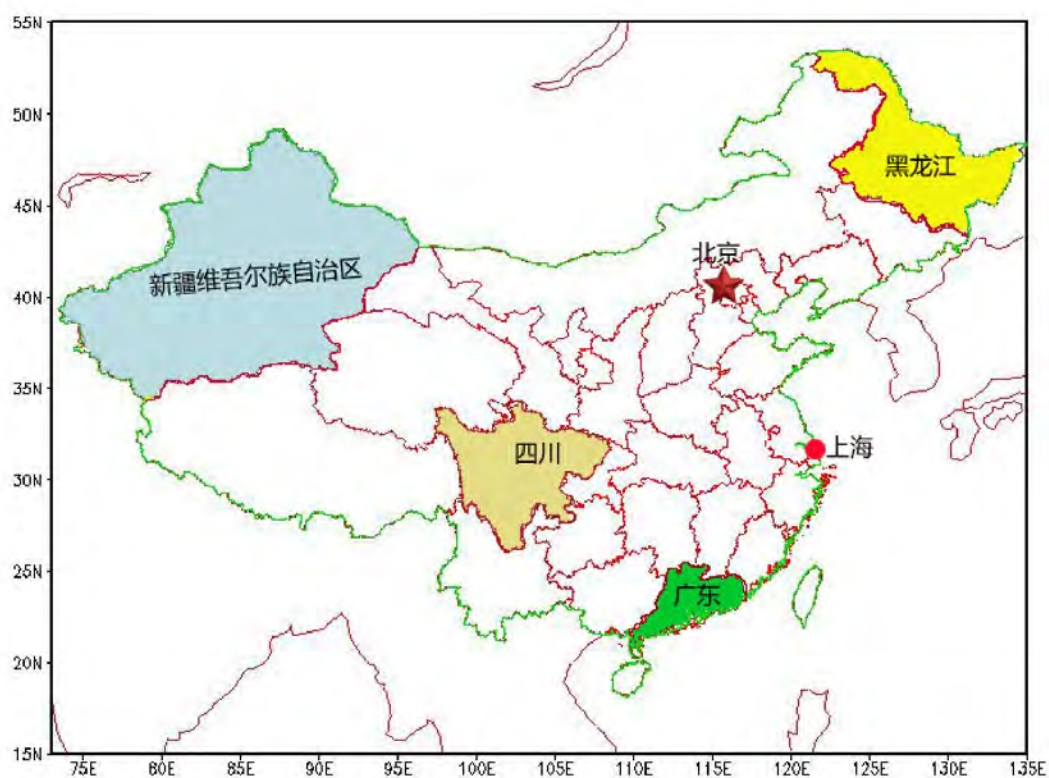


图 13 重要省份在中国地图的分布

各重要省份 1999 年到 2007 年 19 个指标的数据见附录（数据表）



## 6.2 数据的预处理

为了使模型更精确一些，同时，根据城镇登记失业率所具有的地区特性，在分地区建立城镇登记失业率的模型时，增加考虑了一些指标，最后选取的指标为 19 个。

对上述 19 个指标用主成分分析法降维，剔出冗余信息。本文在下面的计算中用主成分的得分来代替上面的 19 个指标，方法同上。荷载为各指标与各主成分的对对应系数。

表 10 各省主成分与荷载

注：表中序号对应于失业率的 19 个指标

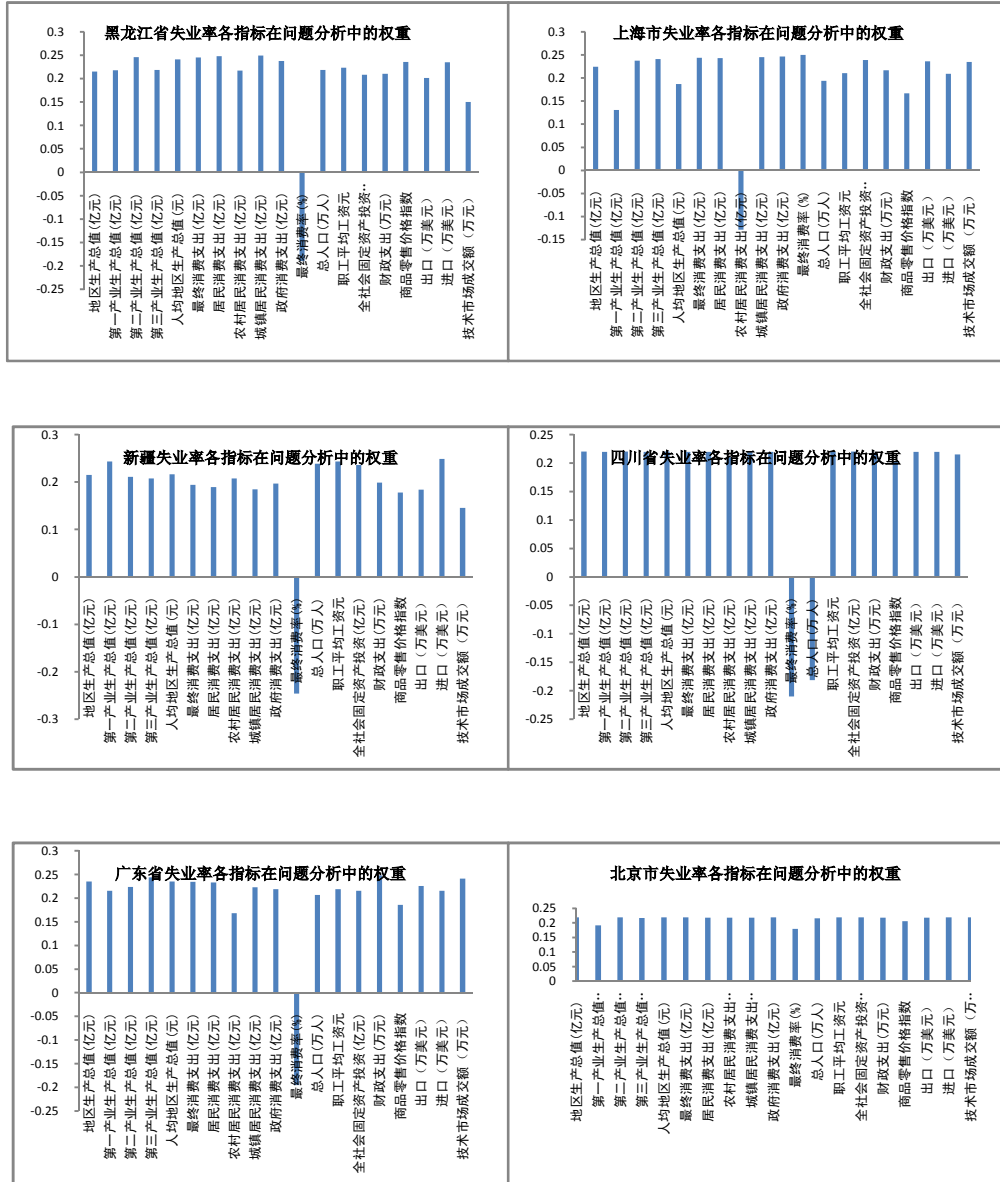
指标	黑龙江		上海		新疆		四川		广东		北京	
	主成分 1	主成分 2	主成分 1	主成分 2	主成分 1	主成分 2	主成分 1		主成分 1	主成分 2	主成分 1	
1	0.9900	-0.0563	0.9993	0.0018	0.9980	-0.0140	0.9993		0.9958	0.0074	0.9902	
2	0.9925	-0.0390	0.7209	-0.6553	0.9837	0.0801	0.9914		0.9850	-0.0586	0.7624	
3	0.9700	0.1213	0.9960	0.0539	0.9946	-0.0299	0.9985		0.9986	-0.0160	0.9932	
4	0.9942	-0.0341	0.9929	0.0865	0.9949	-0.0467	0.9932		0.9901	0.0633	0.9800	
5	0.9929	0.0442	0.9726	-0.1939	0.9971	-0.0099	0.9986		0.9985	0.0066	0.9983	
6	0.9867	0.0939	0.9928	0.1094	0.9850	-0.1498	0.9957		0.9991	0.0045	0.9919	
7	0.9685	0.1518	0.9937	0.0989	0.9779	-0.1908	0.9932		0.9915	0.0036	0.9864	
8	0.9782	-0.0336	-0.0695	-0.9850	0.9931	-0.0484	0.9098		0.2347	0.9650	0.9833	
9	0.9503	0.1963	0.9914	0.1199	0.9671	-0.2389	0.9852		0.9912	-0.0152	0.9843	
10	0.9967	0.0222	0.9898	0.1362	0.9825	-0.1190	0.9919		0.8611	0.0077	0.9953	
11	-0.8408	0.1425	0.9170	0.2859	-0.7175	-0.5719	-0.9101		-0.8434	0.0814	0.6662	
12	0.5497	0.7192	0.9110	-0.0583	0.9465	0.0780	-0.6755		0.8860	-0.0341	0.9659	
13	0.9964	-0.0117	0.9967	-0.0217	0.9925	0.0740	0.9908		0.9792	-0.0261	0.9981	
14	0.9864	-0.1244	0.9946	0.0646	0.9979	0.0253	0.9990		0.9959	-0.0640	0.9943	
15	0.9895	-0.1033	0.9961	-0.0038	0.9827	-0.1026	0.9784		0.9778	0.1224	0.9880	
16	0.9032	0.0909	0.8930	-0.3544	0.7722	-0.0673	0.8419		0.8469	-0.2116	0.8784	
17	0.9725	-0.1968	0.9947	0.0457	0.9634	-0.2384	0.9952		0.9963	-0.0073	0.9826	
18	0.9909	0.0116	0.9938	-0.0271	0.7140	0.6304	0.9949		0.9936	-0.0625	0.9924	
19	0.6984	-0.5582	0.9929	0.0410	0.1117	0.9048	0.9504		0.9431	0.0933	0.9924	

下面将主成分的得分

表 11 各省主成分与对应得分

年份	黑龙江		上海		新疆		四川		广东		北京	
	主成分 1	主成分 2	主成分 1	主成分 2	主成分 1	主成分 2	主成分 1		主成分 1	主成分 2	主成分 1	
1999	1.1145	0.0636	1.0959	-0.1548	1.0097	-0.0557	0.9120		0.9546	0.0865	1.0558	
2000	1.2188	0.0633	1.2381	-0.1583	1.1739	-0.0289	0.9574		1.1024	0.0920	1.2315	
2001	1.3352	0.0832	1.3533	-0.1644	1.2209	0.0073	1.1273		1.1959	0.0993	1.3899	
2002	1.4675	0.0799	1.4901	-0.1663	1.3677	0.0247	1.2436		1.3541	0.1030	1.5341	
2003	1.6163	0.0778	1.7175	-0.1685	1.5987	0.0747	1.4437		1.5631	0.0924	1.7400	
2004	1.8780	0.0763	2.0284	-0.1660	1.8070	0.0955	1.7683		1.7741	0.0854	2.0324	
2005	2.0660	0.0429	2.3935	-0.0870	2.0921	0.0230	2.0408		2.2696	0.0944	2.5770	
2006	2.3922	0.0232	2.7428	-0.0931	2.4157	-0.0595	2.4559		2.5649	0.0969	2.9956	
2007	2.9777	-0.0738	3.2099	-0.0974	2.9515	-0.1163	3.0728		3.0422	0.1100	3.5676	

表 12 各省份失业率指标在问题分析中的权重



### 6.3 模型的求解

应用本文建立的 BP 神经网络模型，将各个省份的指标数据经过主成分分析法降维后代入模型中，得到各个省份更具体的 BP 神经网络模型。

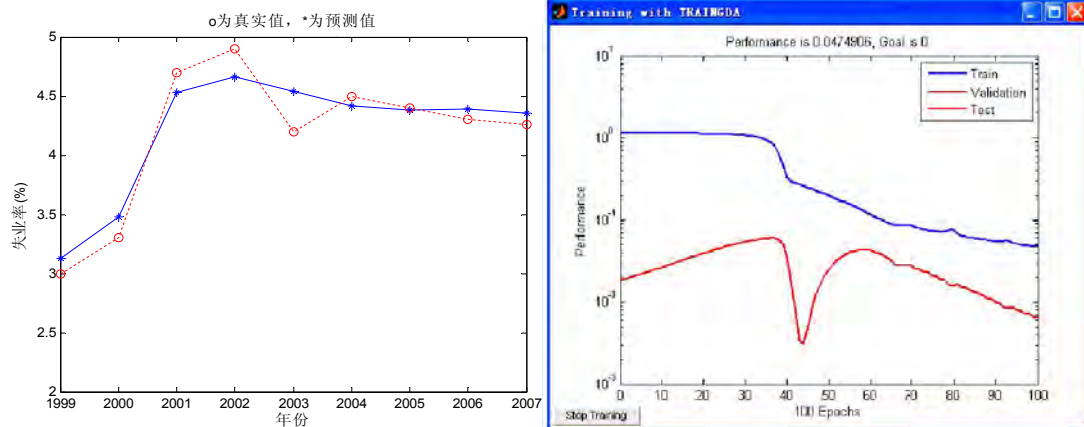


图 14 黑龙江省仿真结果 (2-5-1)

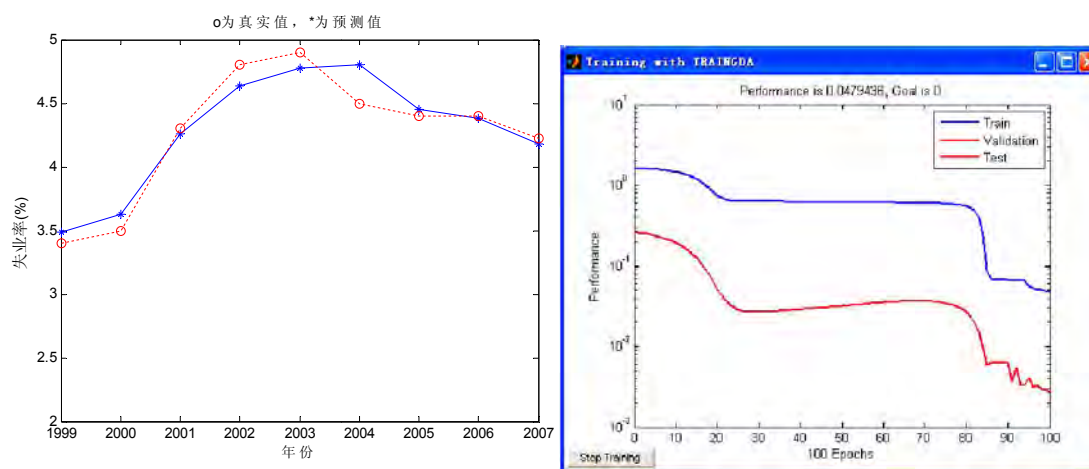


图 15 上海市仿真结果 (2-5-1)

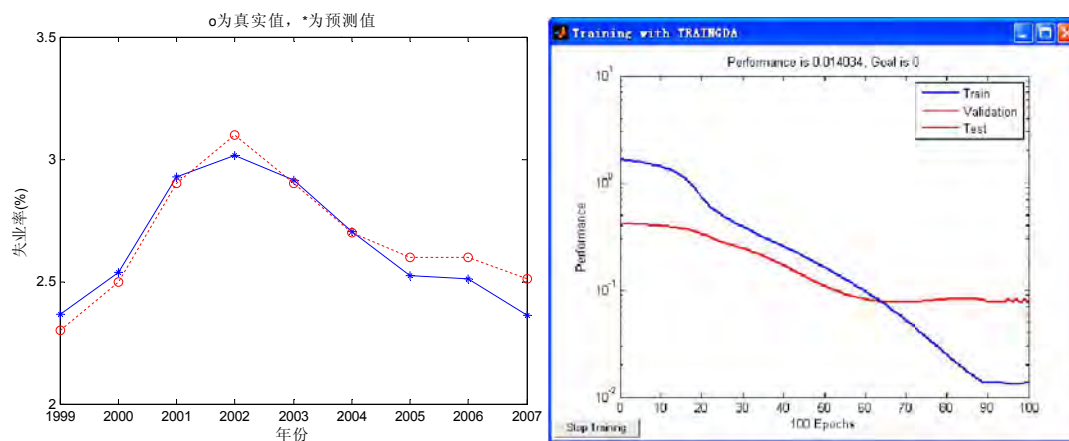


图 16 广东省仿真结果 (2-6-1)

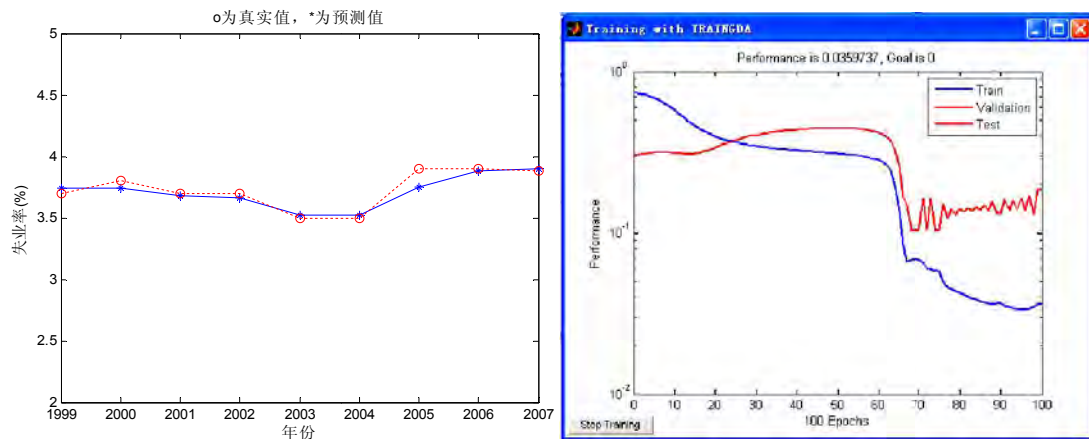


图 17 新疆维吾尔自治区仿真结果 (2-8-1)

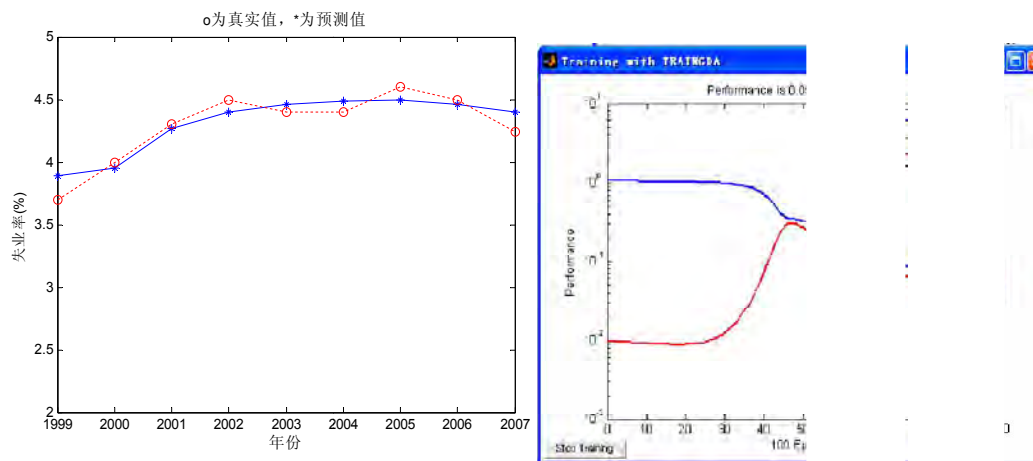


图 18 四川省仿真结果 (1-3-1)

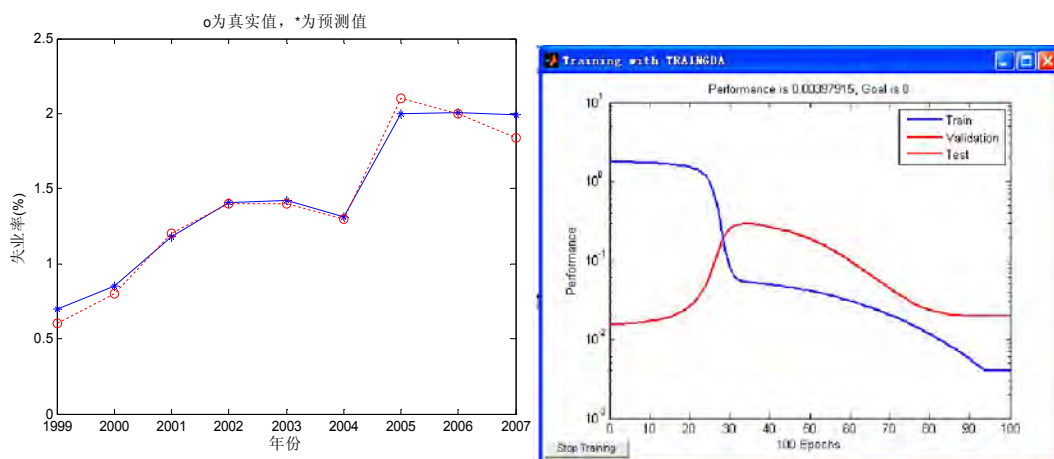


图 20 北京市仿真结果 (1-4-1)

上述图标题中列出的数字为训练过程中得到最优节点数。

下面的表格为本文应用 BP 神经网络建模过程中得到的结果，输入层到隐含层的权值，隐含层到输出层的权值，以及相应层的阈值。

表 13 各省份 BP 神经网络模型建立的相关权值和阈值

黑龙江	输入层到隐含层权值		隐含层阈值
	0.7223	2.0608	-1.92
	-2.3196	0.1989	-1.923
	隐含层到输出层权值		输出层阈值
	2.7943	1.4112	-3.1301
	1.6995	-2.5222	-1.6912
	-1.5788	-2.7467	-0.032
	-2.3898	-2.148	-1.4852
	-3.0603	2.016	-2.7891

新疆	输入层到隐含层权值		隐含层阈值
	-1.4257	1.3601	1.9559
	-1.2172	1.6179	-2.0649
	隐含层到输出层权值		输出层阈值
	-3.913	-0.7014	3.947
	-3.9451	0.2894	2.8327
	-3.1754	2.3634	1.6986
	3.4806	1.8854	-0.566
	3.9442	-0.3513	0.5657
	0.326	4.0201	1.6155
	3.8731	-0.8243	2.8284
	-3.1106	-2.4504	-3.9598

上海	输入层到隐含层权值		隐含层阈值
	-1.0849	-1.6522	1.9753
	-2.1507	1.0183	-1.2708
	隐含层到输出层权值		输出层阈值
	1.6585	-2.8802	-2.9483
	2.7429	1.6475	-1.8487
	-3.0531	-0.7827	-0.036
	2.003	-2.4058	1.5652
	1.33	-2.8339	3.1305

广东	输入层到隐含层权值		隐含层阈值
	-0.5894	-1.8407	2.213
	-2.499	-0.2245	-1.7912
	隐含层到输出层权值		输出层阈值
	1.81	-2.9216	-3.5437
	2.8923	-2.2789	-2.1709
	-2.7018	-2.2221	0.5684
	-3.3804	0.5838	-0.6874
	3.2741	1.0197	2.0575
	0.3118	-3.3881	3.4677

四川	输入层到隐含层权值		隐含层阈值
	1.5403		0.5116
	隐含层到输出层权值		输出层阈值
	4.2		-4.1998
	-4.1968		0.0162
	-4.876		-3.6205

北京	输入层到隐含层权值		隐含层阈值
	1.64		0.3392
	隐含层到输出层权值		输出层阈值
	5.6		-5.6
	5.6001		-1.8669
	5.5804		1.9205
	-6.1609		-5.0552

表 14 各省份的仿真误差

	黑龙江	四川	广东	上海	新疆	北京
绝对平均误差 MAPE	2.74%	2.22%	2.32%	2.53%	1.15%	4.51%
相关系数 R	0.9691	0.9258	0.9571	0.962	0.9288	0.9908
可信度 Z	100%	100%	100%	100%	100%	88.89%

由上表的各种误差比较可知，应用 BP 神经网络得到的结果，MAPE 均小于 5%，相关系数均在 92%以上，可信度均比较高，有很好的仿真拟合结果，说明本文分地区研究失业率与各指标之间关系时建立的模型较合理。

通过这 6 个省份模型的建立，就像是在全国范围内织了一张很大的网，将全国各个地区联系在了一起，可以更加清楚的研究整个国家的失业率与各个指标之间的关系。

## 7. 预测模型的建立

中央政府从 2008 年 10 月进行了大量资金的投资计划，并且采取了扩大内需消费需求的措施，提高对外开放水平以增加出口，影响失业率的一些指标因为国家的决策和规划而引起了较大的改变，相对于其他的指标来说，改变较大的指标对失业率变化的影响大大增加，本模型主要是在 2008 年的特殊国情下，根据国家的决策对 2009 年和 2010 年上半年的失业率进行仿真，所以本模型中，选取 GDP、固定资产投资、和出口额最为影响失业率的主要主表，并且假设其他指标对失业率的影响远远小于这些指标，不予考虑。

为了能够利用所建立的城镇城镇登记失业率的 BP 神经网络模型对 2009 年及 2010 年上半年的失业率进行仿真，本文对 GDP、固定资产投资和出口额未来的数值进行了预测，图 20 为 2007 年 12 月到 2009 年 7 月的出口额趋势图。

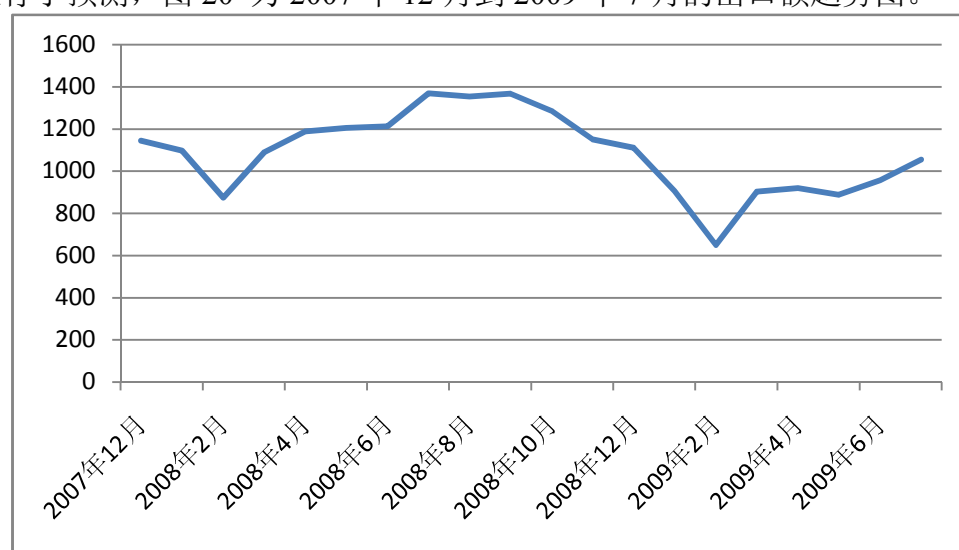


图 20 商品零售价格指数月度变化趋势

出口额变化趋势比较精确的反应了其走势，用月度的数据来进行年度的预测可以提高预测的准确性，因此可以通过数据逼近和插值来预计未来的数值。同样的，可以对 GDP、固定资产投资进行预测，所用数据和预测结果如下：

表 15 GDP 季度数据和预测值（至本季度累计值）（亿元）

季度	2007- I	2007- II	2007-III	2007-IV	2008- I	2008- II
GDP 值	139862.2	65745	300670	208025	134726	63475
季度	2008-III	2008-IV	2009- I	2009- II	2009 年度	2010 上半 年度
GDP 值	257306	174428	112458	53058	420800	549400

表 16 出口额月份数据和预测值

月份	2007.12	2008.1	2008.2	2008.3	2008.4	2008.5	2008.6	2008.7	2008.8	2008.9	2008.10
投资额	1144.2	1096.4	873.7	1089.6	1187.7	1205	1211.8	1368.4	1353.9	1367.6	1285.3
月份	2008.11	2008.12	2009.1	2009.2	2009.3	2009.4	2009.5	2009.6	2009.7	2009 年度	2010 上半 年
投资额	1149.9	1110.6	904.8	648.7	902.1	919.2	887.2	954.7	1054.2	42080	54940

表 17 固定资产投资月份数据和预测值

月份	2008 年 1-2 月	2008 年 1-3 月	2008 年 1-4 月	2008 年 1-5 月	2008 年 1-6 月	2008 年 1-7 月	2008 年 1-8 月	2008 年 1-9 月	2008 年 1-10 月	2008 年 1-11 月
出口额	8121.3	18316.9	28410.1	40264.2	58436.0	72160.1	84919.7	99870.7	113189.1	127614.1
月份	2008 年 1-12 月	2009 年 1-2 月	2009 年 1-3 月	2009 年 1-4 月	2009 年 1-5 月	2009 年 1-6 月	2009 年 1-7 月	2009 度	2010 上 半年度	
出口额	148167.2	10275.8	23562.0	37082.3	53520.3	78098.4	95932.0	13124	15269	

在得到 GDP、固定资产投资和出口额在 2009 年和 2010 年上半年的数值后，输入到训练好的 BP 神经网络中，得到失业率的预测值，预测结果为：

**2009 年的失业率：4.2520**

**2010 年上半年的失业率：4.1402**

## 8. 针对我国现状提出相关意见

通过对城镇登记失业率模型的仿真，可以看出，在 2009 年下半年和 2010 年上半年，我国失业率还会呈上升趋势，但是在 2010 年会有所改善。

对这一趋势，可以从两个方面来进行解释：

(1) 在自然因素方面，2008 年的奥运会，在一定程度上促进了经济的发展，扩大了市场需求，对就业起到了不可忽视的作用，同时，2008 年下半年爆发的金融危机，对我国的就业造成重要影响。

(2) 在政府方面，中央政府实施了投资计划，并确定十大振兴产业计划，提高对外开放增加出口，采取一系列措施来扩大市场需求，对就业起到很大的促进作用。虽然在短期内，我国的经济没有复苏，但是从长远来看，就业形势能够达到改善。

针对我国的特殊国情，通过数学模型的有效分析，结合国家的有关决策和规划，本文给出如下一些建议：

### 1. 数据的准确性

要能够正确反映现实，必须要能得到可观的现实数据，以便正确作出规划，但是在各个指标的统计方面存在一些问题

#### (1) 失业率的不完善性

我国采用的是“城镇登记失业率”，分子是登记的城镇失业人数，分母是城镇从业的人数与登记失业人数之和。由于国家统计局公布的失业率数字只是登记失业率，而这个指标与市场经济国家使用的失业率反映了不尽相同的内容，因而许多经验研究的结论并不可靠。我国登记失业率这个指标有时候不能确切地反映就业形势的好坏，因而影响国家判断经济形势。

城镇登记失业率存在一些问题：在城镇单位从业人员中，不包括使用的农村劳动力。下岗工人转换成失业工人，使失业率增加但劳动力就业状况可能并未糟糕。

国家统计局从 2005 年底就已经开始就城镇调查失业率进行试点，到目前为止共进行了 7 次失业调查，城镇调查失业率实施起来可能会有一定困难，但是城镇调查失业率更接近现实失业情况，所以失业率的统计方法可以朝这个方向改善。

#### (2) 数据的不规则性

由于中国国情的特殊性，经济市场的不稳定，多种矛盾交织，经济框架和技术平台比较薄弱，很多重要的指标在实践序列上呈现不规则性，在某些阶段不符合经济理论。另外，由于一些自然灾害等突发事件对经济的影响，产生了很多异常数据；

### 2. 单方面指标的发展

本文分析了影响就业的多个因素，每个因素对于就业的发展都起到了不可忽视的作用，所以，最基本的一点就是对于各方面的指标采取相应的措施，例如大力发展第三产业，加大教育投入，大力发展人力资源开发力度等。

### 3. 多指标之间的相互影响

由于影响失业率的各个指标之间也是有相互关系的，所以在从单个指标出发，制定有关对策时，还需要综合考虑多个指标之间的交织。



从失业率、经济增长和通货膨胀三个因素的相互交织作用出发,经济增长速度越快就能越多的吸收失业人群,降低失业率,但是,从整个经济运行来讲,经济增长速度过快可能会引起通货膨胀,从而增加失业率;反之,经济增长速度过低,就不能增加就业岗位,导致社会稳定程度的降低,影响改革政策。

这在一方面也可以解释前面定性分析经济增长和失业率之间的关系时,两者之间呈现的非一致性。

总而言之,目前,我国经济社会正经历一个重大的考验,因此,通过数学模型深入分析城镇失业率的相关因素,根据模型的结果采取相应的措施,处理好经济与就业之间的关系,对找准宏观调控的着力点十分重要。

## 9. 模型评价

通过模型建立及实例分析,本文可以得到以下结论:

本模型的优点:

1) 在建模中,可以充分考虑到各个影响城镇登记失业率的因素,通过主成分分析,提炼出较少的与线性无关的主成分作为输入变量进行神经网络模拟,运行效率和精度很高。

2) 采用 BP 神经网络主成分分析法建立的预测模型是可行的。

本模型还存在着一些不足和值得改进的地方:

1)有关 BP 网络的稳定性方法的理论尚不完善,关于隐层层次与节点数只能由试算确定,使得该方法在使用中具有一定的局限性。

2)基于主成分分析法的 BP 神经网络模拟对于受多种因素影响的城镇登记失业率,具有较强的实用性,并充分考虑了与失业率有关因素的复杂性与随机性,这方面的应用外,其应用的领域、运用的范围可在进一步拓宽,相关模型尚需进一步研究和改进。

## 参考文献

- [1] 陈学彬. 对我国经济运行中的菲利普斯曲线关系和通胀预期的实证分析[J]. 财经研究, 1996 (8), 页码: 3-8.
- [2] 程红莉, 刘强. 区域失业率差异影响因素的实证分析[J]. 统计与信息论坛, 2003 年, 18 ( 3): 90-92.
- [3] 田力. 影响失业率的主要因素及降低失业率的对策[J]. 金融理论与教学, 2003 (1): 46-49
- [4] 袁志刚. 失业经济学[M]. 上海: 上海三联, 上海人民出版社, 1997. 76-85.
- [5] 麻荣永, 郑二伟, 王魁, 李建. 基于主成分分析法的广西水资源可持续利用综合评价[J]. 广西大学学报(自然科学版), 2008 (3): 17
- [6] 雷钦礼. 经济管理多元统计分析[M]. 北京: 中国统计出版社, 2002. 153-154
- [7] 王士同等. 问题求解的人工智能: 神经网络方法[M]. 北京: 气象出版社, 1995
- [8] 吴江华, 葛兆帅, 杨达源. 基于人工神经网络的国际入境旅游需求的定量分析与预测[J]. 旅游学刊, 2002, 17 (3): 55-59
- [9] 楼文高. 海水水质评价的人工神经网络模型研究[J]. 海洋环境科学, 2000, 20 (4): 49-53
- [10] 何明, 过秀成. BP 神经网络主成分分析法在汽车保有量预测中的应用[J]. 交通与计算机. 2004. 7 (25): 96-99

附录：

### 黑龙江省失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

年份	失业率 %	地区生产总值(亿元)	第一产业生产总值(亿元)	第二产业生产总值(亿元)	第三产业生产总值(亿元)	人均地区生产总值(元)	最终消费支出(亿元)	居民消费支出(亿元)	农村居民消费支出(亿元)	城镇居民消费支出(亿元)	政府消费支出(亿元)	最终消费率 (%)	总人口(万人)	职工平均工资元	全社会固定资产投资(亿元)	财政支出(万元)	商品零售价格指数	出口(万美元)	进口(万美元)	技术市场成交额(万元)
1999	3	2,897.41	377.23	1,587.76	932.42	7,660.04	1,730.80	1,297.67	328.29	969.38	433.13	59.74	3,792.00	7,094.00	751.66	3,390,342.00	96.1	95,714.00	124,116.00	157,121.00
2000	3.3	3,253.00	357	1,868.55	1,027.45	8,561.65	1,871.54	1,394.21	343.16	1,051.05	477.33	58.31	3,689.00	7,835.00	832.64	3,818,736.00	97.8	145,118.00	153,519.00	152,382.00
2001	4.7	3,390.13	409.3	1,998.74	1,152.96	9,348.91	2,110.54	1,534.64	374.68	1,159.96	575.9	60	3,811.00	8,910.00	963.58	4,782,724.00	100	161,166.00	177,226.00	111,035.00
2002	4.9	3,637.20	447	2,169.15	1,266.01	10,184.00	2,287.75	1,653.11	391.74	1,261.37	634.64	59.7	3,813.00	9,926.00	1,046.17	5,318,682.00	98.5	198,665.00	236,251.00	120,110.00
2003	4.2	4,057.40	500.8	2,532.45	1,396.75	11,615.00	2,465.29	1,771.63	405.63	1,365.98	693.66	58.2	3,815.00	11,038.00	1,166.18	5,649,080.00	99.7	287,426.00	245,514.00	121,165.40
2004	4.5	4,750.60	587.76	3,155.33	1,559.92	13,897.00	2,793.53	1,958.36	401.74	1,556.62	835.17	57	3,817.00	12,557.00	1,430.80	6,975,516.00	103	368,069.00	310,831.00	125,715.10
2005	4.4	5,511.50	684.6	2,971.68	1,855.22	14,434.00	2,660.75	1,841.27	434.72	1,406.55	819.48	48.3	3,820.00	14,458.00	1,737.30	7,877,854.00	100	606,944.00	349,658.00	142,584.50
2006	4.3	6,188.90	737.59	3,365.31	2,086.00	16,195.00	2,961.21	1,964.76	455.48	1,509.28	996.45	47.8	3,823.00	16,505.00	2,236.00	9,685,255.00	102	843,595.00	442,060.00	156,934.00
2007	4.3	7,065.00	915.38	3,695.58	2,454.04	18,478.00	3,514.32	2,288.69	540.45	1,748.24	1,225.63	49.7	3,824.00	19,386.00	2,833.50	13,552,232.00	106	1,225,712.00	503,947.00	350,209.00

# 上海市失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

年份	失业率 %	地区生产总值(亿元)	第一产业生产总值(亿元)	第二产业生产总值(亿元)	第三产业生产总值(亿元)	人均地区生产总值(元)	最终消费支出(亿元)	居民消费支出(亿元)	农村居民消费支出(亿元)	城镇居民消费支出(亿元)	政府消费支出(亿元)	最终消费率 (%)	总人口(万人)	职工平均工资元	全社会固定资产投资(亿元)	财政支出(万元)	商品零售价格指数	出口(万美元)	进口(万美元)	技术市场成交额(万元)
1999	3.4	4,034.96	80	1,953.98	2,000.98	30,804.75	1,719.48	1,352.86	204.41	1,148.45	366.62	42.61	1,474.00	16,641.00	1,855.76	5,335,364.00	97.3	1,879,958.00	1,981,854.00	366,324.00
2000	3.5	4,551.15	83.2	2,163.68	2,304.27	34,546.98	1,947.10	1,521.05	213.74	1,307.31	426.05	42.78	1,674.00	18,531.00	1,869.38	6,085,621.00	96.4	2,535,233.00	2,935,569.00	738,952.00
2001	4.3	5,210.12	85.5	2,355.53	2,509.81	37,382.00	2,149.07	1,663.73	229.67	1,434.06	485.34	43.4	1,614.00	21,781.00	2,004.64	7,081,382.00	98.6	2,762,133.00	3,326,199.00	1,061,603.00
2002	4.8	5,741.03	88.24	2,564.69	2,755.83	40,646.00	2,455.67	1,902.19	241.82	1,660.37	553.48	45.4	1,625.00	23,959.00	2,213.72	8,623,847.00	98.7	3,203,739.00	4,058,972.00	1,202,170.00
2003	4.9	6,694.23	90.64	3,130.72	3,029.45	46,718.00	2,769.74	2,122.90	250.66	1,872.24	646.84	44.3	1,711.00	27,304.00	2,499.14	10,884,386.00	99	4,845,296.00	6,388,659.00	1,427,790.20
2004	4.5	8,072.83	96.71	3,788.22	3,565.34	55,307.00	3,261.42	2,476.19	252.18	2,224.01	785.23	43.8	1,742.00	30,085.00	3,050.30	13,825,254.00	101	7,350,526.00	8,650,467.00	1,716,963.40
2005	4.4	9,164.10	80.34	4,452.92	4,620.92	51,474.00	4,418.99	3,271.64	184.02	3,087.62	1,147.35	48.3	1,778.00	34,345.00	3,509.70	16,462,550.00	99.4	9,071,752.00	9,561,922.00	2,317,328.20
2006	4.4	10,366.37	93.8	5,028.37	5,244.20	57,695.00	5,079.76	3,763.15	202.19	3,560.96	1,316.61	49	1,815.00	41,188.00	3,900.00	17,955,660.00	100	11,358,927.00	11,393,493.00	3,095,095.00
2007	4.2	12,188.85	101.84	5,678.51	6,408.50	66,367.00	6,016.31	4,455.52	233.16	4,222.36	1,560.79	49.4	1,858.00	49,310.00	4,420.37	23,642,915.00	102	14,384,611.00	13,900,777.00	3,548,877.00

## 新疆失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

新疆	失业率 %	地区生产总值(亿元)	第一产业生产总值(亿元)	第二产业生产总值(亿元)	第三产业生产总值(亿元)	人均地区生产总值(元)	最终消费支出(亿元)	居民消费支出(亿元)	农村居民消费支出(亿元)	城镇居民消费支出(亿元)	政府消费支出(亿元)	最终消费率 (%)	总人口(万人)	职工平均工资元	全社会固定资产投资(亿元)	财政支出(万元)	商品零售价格指数	出口(万美元)	进口(万美元)	技术市场成交额(万元)
1999	3.7	1,168.55	268.51	460.71	439.33	6,469.73	745.2	530.37	170.24	360.13	214.83	63.77	1,774.00	7,611.00	526.65	1,662,788.00	96.2	102,734.00	73,801.00	42,822.00
2000	3.8	1,364.36	288.18	586.84	489.34	7,469.81	899.9	585.69	182.05	403.64	314.21	65.96	1,925.00	8,717.00	610.39	1,909,529.00	98.3	120,413.00	105,991.00	66,168.00
2001	3.7	1,491.60	288.12	630.37	566.99	7,913.00	854.6	541.09	176	365.09	313.51	57.5	1,876.00	10,278.00	706	2,633,232.00	103	66,837.00	110,027.00	82,387.00
2002	3.7	1,612.65	305	672.1	621.18	8,382.00	948.92	600.67	190.85	409.82	348.25	59.4	1,905.00	11,605.00	800.09	3,611,729.00	97.9	130,850.00	138,320.00	100,699.00
2003	3.5	1,886.35	412.9	796.84	667.87	9,700.00	1,037.59	617.26	202.68	414.58	420.33	55.3	1,934.00	13,255.00	973.39	3,684,676.00	99.2	254,055.00	222,931.00	120,394.90
2004	3.5	2,209.09	444.7	1,010.07	745.38	11,199.00	1,150.81	663.52	222.49	441.03	487.29	52.3	1,963.00	14,484.00	1,147.20	4,210,446.00	101	304,627.00	258,825.00	133,370.60
2005	3.9	2,604.19	509.99	1,164.79	929.41	13,108.00	1,266.52	764.26	238.98	525.28	502.26	48.6	2,010.00	15,558.00	1,339.10	5,190,179.00	99.4	503,892.00	290,157.00	80,029.30
2006	3.9	3,045.26	527.8	1,459.30	1,058.16	15,000.00	1,584.53	853.89	253.58	600.31	730.64	52	2,050.00	17,819.00	1,567.10	6,784,723.00	102	713,923.00	196,404.00	76,084.00
2007	3.9	3,523.16	628.72	1,647.55	1,246.89	16,999.00	1,930.76	1,013.48	295.52	717.96	917.28	54.8	2,095.00	21,434.00	1,850.84	9,435,653.00	105	1,150,217.00	221,366.00	71,724.00

## 四川省失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

年份	失 业 率 %	地区生产 总值(亿 元)	第一产 业生产 总值(亿 元)	第二产 业生产 总值(亿 元)	第三产 业生产 总值(亿 元)	人均地区 生产总值 (元)	最终消 费支出 (亿元)	居民消 费支出 (亿元)	农村居 民消费 支出(亿 元)	城镇居 民消费 支出(亿 元)	政府消 费支出 (亿元)	最终 消费 率 (%)	总人口 (万人)	职工平均 工资元	全社会 固定资 产投资 (亿元)	财政支出(万 元)	商品 零 售 价 格 指 数	出口(万美 元)	进口(万美 元)	技术市场 成交额(万 元)
1999	3.7	3,711.61	941.02	1,556.68	1,213.91	4,451.90	2,247.20	1,826.99	1,103.39	723.6	420.21	60.55	8,550.00	7,249.00	1,224.40	3,634,981.00	97.3	113,851.00	132,996.00	125,931.00
2000	4	4,010.25	945.58	1,700.49	1,364.18	4,783.77	2,522.87	2,059.20	1,254.76	804.44	463.67	62.91	8,329.00	8,323.00	1,418.04	1,876,433.00	97.7	139,437.00	115,083.00	104,150.00
2001	4.3	4,293.49	981.68	1,756.86	1,683.22	5,250.00	2,691.47	2,076.80	1,194.73	882.07	614.67	60.9	8,640.00	9,934.00	1,617.52	5,940,981.00	101	158,234.00	151,644.00	126,311.00
2002	4.5	4,725.01	1,027.62	1,982.44	1,865.06	5,766.00	2,894.10	2,216.29	1,221.28	995.01	677.81	59.4	8,673.00	11,183.00	1,902.72	7,016,201.00	99.4	271,163.00	175,690.00	77,524.00
2003	4.4	5,333.09	1,128.61	2,266.06	2,061.65	6,418.00	3,155.90	2,413.60	1,303.26	1,110.34	742.3	57.8	8,700.00	12,441.00	2,336.34	7,322,993.00	100	320,871.00	242,558.00	128,686.30
2004	4.4	6,379.63	1,394.26	2,690.00	2,471.76	8,113.00	3,824.88	2,943.58	1,201.78	1,741.80	881.3	58.3	8,725.00	14,063.00	2,818.40	8,952,534.00	104	397,970.00	288,728.00	165,640.10
2005	4.6	7,385.11	1,481.14	3,067.23	2,836.74	9,060.00	4,357.69	3,366.47	1,328.28	2,038.19	991.22	59	8,212.00	15,826.00	3,585.20	10,821,769.00	101	470,161.00	320,036.00	190,823.30
2006	4.5	8,637.81	1,595.48	3,775.19	3,267.14	10,546.00	4,824.88	3,686.82	1,397.96	2,288.86	1,138.06	55.9	8,169.00	17,852.00	4,412.90	13,473,951.00	102	662,412.00	439,670.00	259,323.00
2007	4.2	10,505.30	2,032.00	4,641.30	3,832.00	12,893.00	5,671.56	4,285.21	1,563.20	2,722.01	1,386.35	54	8,127.00	21,312.00	5,639.80	20,805,029.00	105	860,596.00	577,217.00	303,878.00

# 广东省失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

年份	失业率 %	地区生产总值(亿元)	第一产业生产总值(亿元)	第二产业生产总值(亿元)	第三产业生产总值(亿元)	人均地区生产总值(元)	最终消费支出(亿元)	居民消费支出(亿元)	农村居民消费支出(亿元)	城镇居民消费支出(亿元)	政府消费支出(亿元)	最终消费率 (%)	总人口 (万人)	职工平均工资元	全社会固定资产投资(亿元)	财政支出(万元)	商品零售价格指数	出口（万美元）	进口（万美元）	技术市场成交额（万元）
1999	2.3	8,464.31	1,021.30	4,264.32	3,178.69	11,728.29	4,511.64	3,435.00	1,411.18	2,023.82	1,076.64	53.3	7,270.00	12,245.00	2,937.02	9,658,990.00	96.7	7,768,679.00	6,265,562.00	344,528.00
2000	2.5	9,662.23	1,000.06	4,868.75	3,793.42	12,885.40	5,336.22	3,754.80	1,471.03	2,283.77	1,581.42	55.23	8,642.00	13,823.00	3,145.13	10,803,189.00	99.9	9,191,770.00	7,818,118.00	482,104.00
2001	2.9	12,039.25	1,004.35	5,341.61	4,301.75	13,729.93	5,841.32	3,907.02	1,531.21	2,375.81	1,934.30	54.9	7,783.00	15,682.00	3,484.43	13,213,314.00	98.7	9,542,085.00	8,106,590.00	539,722.00
2002	3.1	13,502.42	1,032.80	5,935.63	4,801.30	15,030.00	6,701.15	4,450.00	1,551.55	2,898.45	2,251.15	56.9	7,859.00	17,814.00	3,850.78	15,210,792.00	98.5	11,846,274.00	10,263,357.00	684,532.00
2003	2.9	15,844.64	1,093.52	7,307.08	5,225.27	17,213.00	7,566.13	4,900.01	1,416.37	3,483.64	2,666.12	55.5	7,954.00	19,986.00	4,813.20	16,956,324.00	100	15,284,823.00	13,067,654.00	805,729.90
2004	2.7	18,864.62	1,245.42	8,890.29	5,903.75	19,707.00	8,774.75	5,930.47	1,415.60	4,514.87	2,844.28	54.7	8,304.00	22,116.00	5,870.00	18,529,500.00	103	19,157,104.00	16,555,958.00	572,650.90
2005	2.6	22,366.54	1,428.27	11,339.93	9,598.34	24,435.00	11,533.44	8,989.70	1,419.88	7,569.82	2,543.74	51.6	9,194.00	23,959.00	6,977.90	22,890,691.00	102	23,815,883.00	18,980,614.00	1,124,739.80
2006	2.6	26,204.47	1,577.12	13,431.82	11,195.53	28,332.00	12,892.81	10,015.29	1,483.48	8,531.81	2,877.52	49.2	9,304.00	26,186.00	7,973.40	25,533,399.00	102	30,194,643.00	22,525,267.00	1,070,257.00
2007	2.5	31,084.40	1,695.57	15,939.10	13,449.73	33,151.00	15,166.71	11,873.01	1,554.61	10,318.40	3,293.70	48.8	9,449.00	29,443.00	9,294.26	36,829,589.00	103	36,931,609.00	26,486,987.00	1,328,448.00



# 北京市失业率与各指标的数据（1999-2007 年）

数据来源：原始数据来源《中国统计年鉴》 国研网整理

年份	失 业 率 %	地区生 产总值 (亿元)	第一 产业 生产 总值 (亿元)	第二产 业生产 总值(亿 元)	第三产 业生产 总值(亿 元)	人均地区 生产总值 (元)	最终消 费支出 (亿元)	居民消 费支出 (亿元)	农村 居民 消费 支出 (亿元)	城镇居 民消费 支出(亿 元)	政府消 费支出 (亿元)	最终 消费 率 (%)	总人口 (万人)	职工平均 工资元	全社会 固定资 产投资 (亿元)	财政支出(万 元)	商 品 零 售 价 格 指 数	出口（万美 元）	进口（万美 元）	技术市场成 交额（万元）
1999		2,174.46	87.48	840.23	1,246.75	19,846.30	954.14	633.76	112.52	521.24	320.38	43.88	1,257.00	14,054.00	1,171.16	3,551,932.00	98.8	990,352.00	2,445,599.00	921,889.00
2000		2,478.76	89.97	943.51	1,445.28	22,459.66	1,221.33	808.53	125.92	682.61	412.8	49.27	1,382.00	16,350.00	1,280.46	4,429,969.00	98.9	1,196,813.00	3,765,376.00	1,402,871.00
2001	1.2	3,710.52	93.08	1,030.60	1,721.97	25,523.00	1,467.71	913.95	131.99	781.96	553.76	51.6	1,383.00	19,155.00	1,513.32	5,591,063.00	98.8	1,177,236.00	3,972,572.00	1,910,065.00
2002	1.4	4,330.40	98.05	1,116.53	1,998.13	28,449.00	1,699.81	1,049.21	147.41	901.8	650.6	52.9	1,423.00	21,852.00	1,796.14	6,283,496.00	98.4	1,261,386.00	3,989,142.00	2,211,738.00
2003	1.4	5,023.77	95.64	1,311.86	2,255.60	32,061.00	1,967.87	1,209.27	163.18	1,046.09	758.6	53.7	1,456.00	25,312.00	2,169.26	7,348,043.00	98.2	1,688,682.00	5,161,335.00	2,653,574.00
2004	1.3	6,060.28	102.9	1,610.37	2,570.04	37,058.00	2,264.55	1,354.23	172.15	1,182.08	910.32	52.9	1,493.00	29,674.00	2,528.20	8,982,756.00	99.2	2,056,926.00	7,400,647.00	4,249,975.00
2005	2.1	6,886.31	97.99	2,026.51	4,761.81	45,444.00	3,539.49	2,248.02	184.92	2,063.10	1,291.47	51.4	1,538.00	34,191.00	2,827.20	10,583,114.00	99.7	3,086,590.00	9,464,052.00	4,895,922.00
2006	2	7,870.28	98.04	2,191.43	5,580.81	50,467.00	4,205.16	2,615.26	191.21	2,424.05	1,589.90	53.4	1,581.00	40,117.00	3,296.40	12,968,389.00	100	3,795,398.00	12,008,265.00	6,973,256.00
2007	1.8	9,353.32	101.26	2,509.40	6,742.66	58,204.00	5,082.80	3,039.03	226.93	2,812.10	2,043.77	54.3	1,633.00	46,507.00	3,907.20	18,290,601.00	101	4,892,639.00	14,407,337.00	8,825,603.00

## 附件：程序

### 一. 函数组：

#### 1. 主成分计算函数之一

```
%cwfac.m 计算贡献率和载荷
function [result,newi]=cwfac(vector);
fprintf('相关系数矩阵:\n')
std=CORRCOEf(vector)    %计算相关系数矩阵
fprintf('特征向量(vec)及特征值(val): \n')
[vec,val]=eig(std)    %求特征值(val)及特征向量(vec)
newval=diag(val);
[y,i]=sort(newval);    %对特征根进行列排序, y 为排序结果, i 为索引
fprintf('特征根排序: \n')
for z=1:length(y)
    newy(z)=y(length(y)+1-z);
end
fprintf('%g\n',newy)
rate=y/sum(y);
fprintf('\n 贡献率: \n')
newrate=newy/sum(newy)
sumrate=0;
newi=[];
for k=length(y):-1:1
    sumrate=sumrate+rate(k);
    newi(length(y)+1-k)=i(k);
    if sumrate>0.99 break;
end
%记下累积贡献率大 99%的特征值的序号放入 newi 中
fprintf('主成分数: %g\n',length(newi));
fprintf('主成分位置及数据 zhu、主成分载荷 result: \n')
for p=1:length(newi)
    for q=1:length(y)
        result(q,p)=sqrt(newval(newi(p)))*vec(q,newi(p));
    end
end
%计算载荷
disp(result)
disp(i)
```

#### 2. 主成分计算函数之二

```
%cwprint.m 读入数据文件; 调用以上三个函数并输出结果
function print=cwprint(filename,a,b);
%filename 为文本文件文件名, a 为矩阵行数(样本数), b 为矩阵列数(变量指标数)
fid=fopen(filename,'r')
vector=fscanf(fid,'%g',[a b]);
fprintf('标准化结果如下: \n')
v1=cwstd(vector)
```

```
result=cwfac(v1);
cwscore(v1,result);
```

### 3. 主成分计算函数之三

```
%cwscore.m, 计算得分 计算各主成分得分、综合得分并排序
function score=cwscore(vector1,vector2);
sco=vector1*vector2;
csum=sum(sco,2);
[newcsum,i]=sort(-1*csum);
[newi,j]=sort(i);
fprintf(' 计算得分: \n')
score=[sco,csum,j]
%得分矩阵: sco 为各主成分得分; csum为综合得分; j 为排序结果
```

### 4. 主成分计算函数之四

```
%cwstd.m, 用总和标准化法标准化矩阵
function std=cwstd(vector)
cwsum=sum(vector,1); %对列求和
[a,b]=size(vector); %矩阵大小,a 为行数,b 为列数
for i=1:a
    for j=1:b
        std(i,j)= vector(i,j)/cwsum(j);
    end
end
```

### 5. 算绝对平均误差

```
%算绝对平均误差 mape.m
function mape_r=mape(x,y)
abs0=abs(x-y);
mape_r=sum(abs0./y)/length(x);
```

### 6. 算可信度 Z

```
%算可信度 z.m
function z_r=z(x,y)
abs0=abs(x-y);
err=abs0./y;
for i=1:length(x)
    if abs(err(i))<=0.15
        j(i)=1;
    else j(i)=0;
    end
end
```

```
z_r=sum(j)/length(x);
```

## 二. 各题运算程序

### 1. 主成分分析程序

(1) 计算相关系数和得分 `yuanshishuju.m`

```
a=[1935, 2140, 2340, 2711, 3371, 4538, 5500, 6210, 6470, 7479, 8346, 9371, 10870, 12422, 14040, 16024, 18364, 21001, 24932, 29229
```

```
26. 21, 26. 41, 26. 94, 27. 46, 27. 99, 28. 51, 29. 04, 30. 48, 31. 91, 33. 35, 34. 78, 36. 22, 37. 66, 39. 09, 40. 53, 41. 76, 42. 99, 43. 9, 44. 94, 45. 68
```

```
117. 8, 102. 1, 102. 9, 105. 4, 113. 2, 121. 7, 114. 8, 106. 1, 100. 8, 97. 4, 97, 98. 5, 99. 2, 98. 7, 99. 9, 102. 8, 100. 8, 101, 103. 7949, 105. 9
```

```
16992. 32, 18667. 82, 21781. 5, 26923. 48, 35333. 92, 48197. 86, 60793. 73, 71176. 59, 78973. 03, 84402. 28, 89677. 05, 99214. 55, 109655. 17, 120332. 69, 135822. 76, 159878. 34, 183217. 4, 211923. 5, 257306, 300670
```

```
7278, 7717. 4, 9102. 2, 11699. 5, 16454. 43, 22445. 4, 28679. 46, 33834. 96, 37543, 39004. 19, 41033. 58, 45555. 88, 49512. 29, 53896. 77, 62436. 31, 73904. 31, 87364. 58, 103162, 124799, 146183
```

```
5448. 4, 5888. 42, 7337. 1, 9357. 38, 11915. 73, 16179. 76, 19978. 46, 23326. 24, 26988. 15, 30580. 47, 33873. 44, 38713. 95, 44361. 61, 49898. 9, 56004. 73, 64561. 29, 73432. 87, 84721. 4, 103880, 120487
```

```
1956, 2985. 8, 3827. 1, 4676. 3, 5284. 8, 10421. 8, 12451. 8, 12576. 4, 15160. 7, 15231. 6, 16159. 8, 20634. 4, 22024. 4, 26947. 9, 36287. 9, 49103. 3, 62648. 1, 77594. 59, 93455. 6272, 99241. 707828
```

```
2199. 9, 2574. 3, 3398. 7, 4443. 3, 5986. 2, 9960. 1, 11048. 1, 11557. 4, 11806. 5, 11626. 1, 13736. 4, 18638. 8, 20159. 2, 24430. 3, 34195. 6, 46435. 8, 54273. 7, 63376. 856, 73284. 5612, 78654. 917127
```

```
4410. 38, 4517, 5594. 5, 8080. 09, 12457. 88, 16370. 33, 20019. 26, 22974. 03, 24941. 11, 28406. 17, 29854. 71, 32917. 73, 37213. 49, 43499. 9093, 55566. 6059, 70477. 44, 88773. 61, 109998. 1624, 137323. 9381, 172291
```

```
81. 5 75. 04 94. 1 150. 8 207. 4 228. 8 268 300 351 430 525 630 800 880 1082. 7 1334 1510 1818 2200 2665];
```

```
y=a';
```

%从1989年到2008年的职工平均工资          单位为：元

```
x1=[1935, 2140, 2340, 2711, 3371, 4538, 5500, 6210, 6470, 7479, 8346, 9371, 10870, 12422, 14040, 16024, 18364, 21001, 24932, 29229];
```

%从1989年到2008年的城镇人口所占比例

```
x2=[26. 21, 26. 41, 26. 94, 27. 46, 27. 99, 28. 51, 29. 04, 30. 48, 31. 91, 33. 35, 34. 78, 36. 22, 37. 66, 3
```

9. 09, 40. 53, 41. 76, 42. 99, 43. 9, 44. 94, 45. 68];

% 商品零售价格指数

x3=[117. 8, 102. 1, 102. 9, 105. 4, 113. 2, 121. 7, 114. 8, 106. 1, 100. 8, 97. 4, 97, 98. 5, 99. 2, 98. 7, 99  
. 9, 102. 8, 100. 8, 101, 103. 7949, 105. 9];

%gdp增长率

x4=[4. 1  
3. 8  
9. 2  
4. 2  
14  
13. 1  
10. 9  
10  
9. 3  
7. 8  
7. 6  
8. 4  
8. 3  
9. 1  
10  
10. 1  
10. 4  
11. 6  
13  
9];

%第二产业/gdp

x5=[42. 8311  
41. 3407  
41. 7887  
43. 4546  
46. 5684  
46. 5693  
47. 1750  
47. 5366  
47. 5390  
46. 2122  
45. 7571  
45. 9165  
45. 1527  
44. 7898  
45. 9690  
46. 2253  
47. 6836  
48. 6789  
48. 5022  
48. 6191];

%第三产业/gdp

x6=[32. 0639  
31. 5432  
33. 6850

34.7555  
 33.7232  
 33.5695  
 32.8627  
 32.7723  
 34.1739  
 36.2318  
 37.7727  
 39.0204  
 40.4556  
 41.4675  
 41.2337  
 40.3815  
 40.0796  
 39.9774  
 40.3722  
 40.0728];

%出口总额 单位为：亿元

%x7=[1956, 2985.8, 3827.1, 4676.3, 5284.8, 10421.8, 12451.8, 12576.4, 15160.7, 15231.6, 16159.8, 20634.4, 22024.4, 26947.9, 36287.9, 49103.3, 62648.1, 77594.59, 93455.6272, 99241.707828];

x7=[11.5111  
 15.9944  
 17.5704  
 17.3689  
 14.9567  
 21.6230  
 20.4820  
 17.6693  
 19.1973  
 18.0464  
 18.0200  
 20.7978  
 20.0851  
 22.3945  
 26.7171  
 30.7129  
 34.1933  
 36.6144  
 37.4527  
 33.0069]';

%进口总额

%x8=[2199.9, 2574.3, 3398.7, 4443.3, 5986.2, 9960.1, 11048.1, 11557.4, 11806.5, 11626.1, 13736.4, 18638.8, 20159.2, 24430.3, 34195.6, 46435.8, 54273.7, 63376.856, 73284.5612, 78654.917127];

x8=[12.9464  
 13.7900  
 15.6036  
 16.5034  
 16.9418  
 20.6650  
 18.1731

16.2376  
14.9500  
13.7746  
15.3176  
18.7864  
18.3842  
20.3023  
25.1766  
29.0445  
29.6226  
29.9055  
29.3691  
26.1599]’;

%固定资产投资/gdp 单位为：亿元

x9=[25.9551  
24.1967  
25.6846  
30.0113  
35.2576  
33.9648  
32.9298  
32.2775  
31.5818  
33.6557  
33.2914  
33.1783  
33.9368  
36.1497  
40.9111  
44.0819  
48.4526  
51.9047  
53.3699  
57.3024];

%技术市场成交额/国内生产总值 单位为：亿元

%x10=[81.5 75.04 94.1 150.8 207.4 228.8 268 300 351 430 525 630 800 880 1082.7 1334 1510  
1818 2200 2665];

x10=[47.96  
40.20  
43.20  
56.01  
58.70  
47.47  
44.08  
42.15  
44.45  
50.95  
58.54  
63.50  
72.96

73.13  
 79.71  
 83.44  
 82.42  
 85.79  
 85.50  
 88.64];  
 %最终消费支出的贡献率 ——贡献率指与支出法国内生产总值增量之比  
 x11=[39.6  
 47.8  
 65.1  
 72.5  
 59.5  
 30.2  
 44.7  
 60.1  
 37.0  
 57.1  
 74.7  
 65.1  
 50.0  
 43.6  
 35.3  
 38.7  
 38.2  
 38.7  
 39.4  
 39.7]';  
 x=[1935, 2140, 2340, 2711, 3371, 4538, 5500, 6210, 6470, 7479, 8346, 9371, 10870, 12422, 14040, 16  
 024, 18364, 21001, 24932, 29229  
  
 26.21, 26.41, 26.94, 27.46, 27.99, 28.51, 29.04, 30.48, 31.91, 33.35, 34.78, 36.22, 37.66, 39.09  
 , 40.53, 41.76, 42.99, 43.9, 44.94, 45.68  
  
 117.8, 102.1, 102.9, 105.4, 113.2, 121.7, 114.8, 106.1, 100.8, 97.4, 97, 98.5, 99.2, 98.7, 99.9, 1  
 02.8, 100.8, 101, 103.7949, 105.9  
 4.1, 3.8, 9.2, 14.2, 14, 13.1, 10.9, 10, 9.3, 7.8, 7.6, 8.4, 8.3, 9.1, 10, 10.1, 10.4, 11.6, 13, 9  
 42.8311 41.3407 41.7887 43.4546 46.5684 46.5693 47.1750 47.5366  
 47.5390 46.2122 45.7571 45.9165 45.1527 44.7898 45.9690 46.2253  
 47.6836 48.6789 48.5022 48.6191  
 32.0639 31.5432 33.6850 34.7555 33.7232 33.5695 32.8627 32.7723  
 34.1739 36.2318 37.7727 39.0204 40.4556 41.4675 41.2337 40.3815  
 40.0796 39.9774 40.3722 40.0728  
 11.5111 15.9944 17.5704 17.3689 14.9567 21.6230 20.4820 17.6693  
 19.1973 18.0464 18.0200 20.7978 20.0851 22.3945 26.7171 30.7129  
 34.1933 36.6144 37.4527 33.0069  
 12.9464 13.7900 15.6036 16.5034 16.9418 20.6650 18.1731 16.2376  
 14.9500 13.7746 15.3176 18.7864 18.3842 20.3023 25.1766 29.0445  
 29.6226 29.9055 29.3691 26.1599  
 25.9551 24.1967 25.6846 30.0113 35.2576 33.9648 32.9298 32.2775  
 31.5818 33.6557 33.2914 33.1783 33.9368 36.1497 40.9111 44.0819  
 48.4526 51.9047 53.3699 57.3024  
 47.96 40.20 43.20 56.01 58.70 47.47 44.08 42.15 44.45



```

50.95    58.54    63.50    72.96    73.13    79.71    83.44    82.42    85.79    85.50
88.64
    39.6 47.8 65.1 72.5 59.5 30.2 44.7 60.1 37.0 57.1 74.7 65.1 50.0 43.6 35.3 38.7
38.2 38.7 39.4 39.7];
fprintf('x的转置')
x'
v1=cwstd(x')
result=cwfac(v1);
cwscore(v1,result);

```

```

b=[ -0.9612    0.0651   -0.0811
   -0.9695    0.2213   -0.0523
    0.3141   -0.8700   -0.1722
   -0.3475   -0.5620    0.6917
   -0.7139   -0.3653    0.2178
   -0.8537    0.3978    0.0509
   -0.9582   -0.0834   -0.0347
   -0.9390   -0.1618   -0.0334
   -0.9646   -0.1358    0.0303
   -0.9376    0.1565    0.0007
    0.4811    0.4529    0.6862];
c=[7.16272
1.6988
1.0416];
fprintf('m值\n')
for i=1:11
    for j=1:3
        m(i,j)=sign(b(i,j))*sqrt(abs(b(i,j))/c(j,1));
    end
end
disp(m)
fprintf('s值\n')
n=[ 0.6512    0.1544    0.0947 ];
for i=1:11
    s(i,1)=m(i,1)*0.6512/(0.6512 + 0.1544 + 0.0947)+m(i,2)* 0.1544/(0.6512 +
0.1544 + 0.0947)+m(i,3)* 0.0947/(0.6512 + 0.1544 + 0.0947);
end
disp(s)
mean(s)
disp(s-mean(s))

```

## (2) 计算指标的权重 xishu.m

```

%黑龙江
a=[2897.41 377.23 1587.76 932.42 7660.04 1730.80 1297.67 328.29 969.38
433.13 59.74 3792.00 7094.00 751.66 3390342.00 96.195714.00 124116.00
157121.00
3253.00 357 1868.55 1027.45 8561.65 1871.54 1394.21 343.16 1051.05 477.33
58.31 3689.00 7835.00 832.64 3818736.00 97.8145118.00 153519.00
152382.00
3390.13 409.3 1998.74 1152.96 9348.91 2110.54 1534.64 374.68 1159.96 575.9
60 3811.00 8910.00 963.58 4782724.00 100.4 161166.00 177226.00
111035.00
3637.20 447 2169.15 1266.01 10184.00 2287.75 1653.11 391.74 1261.37 634.64

```

```

59.73813.00 9926.00 1046.17 5318682.00 98.5198665.00 236251.00
120110.00
4057.40 500.8 2532.45 1396.75 11615.00 2465.29 1771.63 405.63 1365.98 693.66
58.23815.00 11038.00 1166.18 5649080.00 99.7287426.00 245514.00
121165.40
4750.60 587.76 3155.33 1559.92 13897.00 2793.53 1958.36 401.74 1556.62 835.17
57 3817.00 12557.00 1430.80 6975516.00 102.8 368069.00 310831.00
125715.10
5511.50 684.6 2971.68 1855.22 14434.00 2660.75 1841.27 434.72 1406.55 819.48
48.33820.00 14458.00 1737.30 7877854.00 100.4 606944.00 349658.00
142584.50
6188.90 737.59 3365.31 2086.00 16195.00 2961.21 1964.76 455.48 1509.28 996.45
47.83823.00 16505.00 2236.00 9685255.00 101.5 843595.00 442060.00
156934.00
7065.00 915.38 3695.58 2454.04 18478.00 3514.32 2288.69 540.45 1748.24 1225.63
49.73824.00 19386.00 2833.50 13552232.00 105.6 1225712.00 503947.00
350209.00

```

```
];
```

```
%shanghai
```

```

sh=[4034.96 80.00 1953.98 2000.98 30804.75 1719.48 1352.86 204.41 1148.45
366.62 42.61 1474.00 16641.00 1855.76 5335364.00 97.30 1879958.00
1981854.00 366324.00
4551.15 83.20 2163.68 2304.27 34546.98 1947.10 1521.05 213.74 1307.31 426.05
42.78 1674.00 18531.00 1869.38 6085621.00 96.40 2535233.00 2935569.00
738952.00
5210.12 85.50 2355.53 2509.81 37382.00 2149.07 1663.73 229.67 1434.06 485.34
43.40 1614.00 21781.00 2004.64 7081382.00 98.60 2762133.00 3326199.00
1061603.00
5741.03 88.24 2564.69 2755.83 40646.00 2455.67 1902.19 241.82 1660.37 553.48
45.40 1625.00 23959.00 2213.72 8623847.00 98.70 3203739.00 4058972.00
1202170.00
6694.23 90.64 3130.72 3029.45 46718.00 2769.74 2122.90 250.66 1872.24 646.84
44.30 1711.00 27304.00 2499.14 10884386.00 99.00 4845296.00 6388659.00
1427790.20
8072.83 96.71 3788.22 3565.34 55307.00 3261.42 2476.19 252.18 2224.01 785.23
43.80 1742.00 30085.00 3050.30 13825254.00 100.90 7350526.00 8650467.00
1716963.40
9164.10 80.34 4452.92 4620.92 51474.00 4418.99 3271.64 184.02 3087.62 1147.35
48.30 1778.00 34345.00 3509.70 16462550.00 99.40 9071752.00 9561922.00
2317328.20
10366.37 93.80 5028.37 5244.20 57695.00 5079.76 3763.15 202.19 3560.96 1316.61
49.00 1815.00 41188.00 3900.00 17955660.00 100.20 11358927.00 11393493.00
3095095.00
12188.85 101.84 5678.51 6408.50 66367.00 6016.31 4455.52 233.16 4222.36 1560.79
49.40 1858.00 49310.00 4420.37 23642915.00 102.40 14384611.00 13900777.00
3548877.00

```

```
];
```

```
%xinjiang
```

```

xin=[1168.55 268.51 460.71 439.33 6469.73 745.20 530.37 170.24 360.13
214.83 63.77 1774.00 7611.00 526.65 1662788.00 96.20 102734.00
73801.00 42822.00
1364.36 288.18 586.84 489.34 7469.81 899.90 585.69 182.05 403.64 314.21

```

```

        65.96    1925.00  8717.00  610.39   1909529.00   98.30    120413.00    105991.00
        66168.00
1491.60 288.12   630.37   566.99   7913.00  854.60   541.09   176.00   365.09   313.51
        57.50    1876.00  10278.00  706.00   2633232.00   102.50   66837.00  110027.00
        82387.00
1612.65 305.00   672.10   621.18   8382.00  948.92   600.67   190.85   409.82   348.25
        59.40    1905.00  11605.00  800.09   3611729.00   97.90    130850.00   138320.00
        100699.00
1886.35 412.90   796.84   667.87   9700.00  1037.59  617.26   202.68   414.58   420.33
        55.30    1934.00  13255.00  973.39   3684676.00   99.20    254055.00   222931.00
        120394.90
2209.09 444.70   1010.07  745.38   11199.00 1150.81  663.52   222.49   441.03   487.29
        52.30    1963.00  14484.00 1147.20   4210446.00   100.70   304627.00   258825.00
        133370.60
2604.19 509.99   1164.79  929.41   13108.00 1266.52  764.26   238.98   525.28   502.26
        48.60    2010.00  15558.00 1339.10   5190179.00   99.40    503892.00   290157.00
        80029.30
3045.26 527.80   1459.30 1058.16   15000.00 1584.53  853.89   253.58   600.31   730.64
        52.00    2050.00  17819.00 1567.10   6784723.00   101.80   713923.00   196404.00
        76084.00
3523.16 628.72   1647.55 1246.89   16999.00 1930.76 1013.48   295.52   717.96   917.28
        54.80    2095.00  21434.00 1850.84   9435653.00   105.10   1150217.00   221366.00
        71724.00
];
%sichuang
si=[3711.61 941.02   1556.68 1213.91 4451.90 2247.20 1826.99 1103.39 723.60
    420.21  60.55   8550.00 7249.00 1224.40 3634981.00 97.30  113851.00
    132996.00  125931.00
4010.25 945.58   1700.49 1364.18 4783.77 2522.87 2059.20 1254.76 804.44 463.67
    62.91   8329.00  8323.00 1418.04 1876433.00 97.70  139437.00   115083.00
    104150.00
4293.49 981.68   1756.86 1683.22 5250.00 2691.47 2076.80 1194.73 882.07 614.67
    60.90   8640.00  9934.00 1617.52 5940981.00 100.80  158234.00   151644.00
    126311.00
4725.01 1027.62 1982.44 1865.06 5766.00 2894.10 2216.29 1221.28 995.01 677.81
    59.40   8673.00  11183.00 1902.72 7016201.00 99.40  271163.00   175690.00
    77524.00
5333.09 1128.61 2266.06 2061.65 6418.00 3155.90 2413.60 1303.26 1110.34 742.30
    57.80   8700.00  12441.00 2336.34 7322993.00 100.10  320871.00   242558.00
    128686.30
6379.63 1394.26 2690.00 2471.76 8113.00 3824.88 2943.58 1201.78 1741.80 881.30
    58.30   8725.00  14063.00 2818.40 8952534.00 103.70  397970.00   288728.00
    165640.10
7385.11 1481.14 3067.23 2836.74 9060.00 4357.69 3366.47 1328.28 2038.19 991.22
    59.00   8212.00  15826.00 3585.20 10821769.00 100.60  470161.00   320036.00
    190823.30
8637.81 1595.48 3775.19 3267.14 10546.00 4824.88 3686.82 1397.96 2288.86 1138.06
    55.90   8169.00  17852.00 4412.90 13473951.00 101.70  662412.00   439670.00
    259323.00
10505.30 2032.00 4641.30 3832.00 12893.00 5671.56 4285.21 1563.20 2722.01 1386.35
    54.00   8127.00  21312.00 5639.80 20805029.00 105.30  860596.00   577217.00
    303878.00
];

```

```

%guangdong]
gu=[8464.31 1021.30 4264.32 3178.69 11728.29 4511.64 3435.00 1411.18 2023.82
    1076.64 53.30 7270.00 12245.00 2937.02 9658990.00 96.70 7768679.00
    6265562.00 344528.00
9662.23 1000.06 4868.75 3793.42 12885.40 5336.22 3754.80 1471.03 2283.77 1581.42
    55.23 8642.00 13823.00 3145.13 10803189.00 99.90 9191770.00 7818118.00
    482104.00
12039.25 1004.35 5341.61 4301.75 13729.93 5841.32 3907.02 1531.21 2375.81 1934.30
    54.90 7783.00 15682.00 3484.43 13213314.00 98.70 9542085.00 8106590.00
    539722.00
13502.42 1032.80 5935.63 4801.30 15030.00 6701.15 4450.00 1551.55 2898.45 2251.15
    56.90 7859.00 17814.00 3850.78 15210792.00 98.50 11846274.00 10263357.00
    684532.00
15844.64 1093.52 7307.08 5225.27 17213.00 7566.13 4900.01 1416.37 3483.64 2666.12
    55.50 7954.00 19986.00 4813.20 16956324.00 100.00 15284823.00 13067654.00
    805729.90
18864.62 1245.42 8890.29 5903.75 19707.00 8774.75 5930.47 1415.60 4514.87 2844.28
    54.70 8304.00 22116.00 5870.00 18529500.00 102.90 19157104.00 16555958.00
    572650.90
22366.54 1428.27 11339.93 9598.34 24435.00 11533.44 8989.70 1419.88 7569.82 2543.74
    51.60 9194.00 23959.00 6977.90 22890691.00 101.80 23815883.00 18980614.00
    1124739.80
26204.47 1577.12 13431.82 11195.53 28332.00 12892.81 10015.29 1483.48 8531.81 2877.52
    49.20 9304.00 26186.00 7973.40 25533399.00 101.50 30194643.00 22525267.00
    1070257.00
31084.40 1695.57 15939.10 13449.73 33151.00 15166.71 11873.01 1554.61 10318.40 3293.70
    48.80 9449.00 29443.00 9294.26 36829589.00 103.40 36931609.00 26486987.00
    1328448.00
];
%beijing
bei=[2174.46 87.48 840.23 1246.75 19846.30 954.14 633.76 112.52 521.24
    320.38 43.88 1257.00 14054.00 1171.16 3551932.00 98.80 990352.00
    2445599.00 921889.00
2478.76 89.97 943.51 1445.28 22459.66 1221.33 808.53 125.92 682.61 412.80
    49.27 1382.00 16350.00 1280.46 4429969.00 98.90 1196813.00 3765376.00
    1402871.00
3710.52 93.08 1030.60 1721.97 25523.00 1467.71 913.95 131.99 781.96 553.76
    51.60 1383.00 19155.00 1513.32 5591063.00 98.80 1177236.00 3972572.00
    1910065.00
4330.40 98.05 1116.53 1998.13 28449.00 1699.81 1049.21 147.41 901.80 650.60
    52.90 1423.00 21852.00 1796.14 6283496.00 98.40 1261386.00 3989142.00
    2211738.00
5023.77 95.64 1311.86 2255.60 32061.00 1967.87 1209.27 163.18 1046.09 758.60
    53.70 1456.00 25312.00 2169.26 7348043.00 98.20 1688682.00 5161335.00
    2653574.00
6060.28 102.90 1610.37 2570.04 37058.00 2264.55 1354.23 172.15 1182.08 910.32
    52.90 1493.00 29674.00 2528.20 8982756.00 99.20 2056926.00 7400647.00
    4249975.00
6886.31 97.99 2026.51 4761.81 45444.00 3539.49 2248.02 184.92 2063.10 1291.47
    51.40 1538.00 34191.00 2827.20 10583114.00 99.70 3086590.00 9464052.00
    4895922.00
7870.28 98.04 2191.43 5580.81 50467.00 4205.16 2615.26 191.21 2424.05 1589.90
    53.40 1581.00 40117.00 3296.40 12968389.00 100.20 3795398.00 12008265.00

```

```

        6973256.00
9353.32 101.26 2509.40 6742.66 58204.00 5082.80 3039.03 226.93 2812.10 2043.77
        54.30 1633.00 46507.00 3907.20 18290601.00 100.80 4892639.00 14407337.00
        8825603.00
];
%b为荷载
b=[];
v1=cwstd(bei)
result=cwfac(v1);
cwscore(v1,result);
score= cwscore(v1,result);
%b为荷载
b=[ 0.9902
    0.7624
    0.9932
    0.9800
    0.9983
    0.9919
    0.9864
    0.9833
    0.9843
    0.9953
    0.6662
    0.9659
    0.9981
    0.9943
    0.9880
    0.8784
    0.9826
    0.9924
    0.9924
];

c=[17.4334];
fprintf('m值\n')
for i=1:19
    for j=1
        m(i,j)=sign(b(i,1))*sqrt(abs(b(i,1))/c(j,1));
    end
end
disp(m)
fprintf('s值\n')
n=[ 0.9175];
for i=1:19
    s(i,1)=m(i,1)* n(1);
end
disp(s)
mean(s)
disp(s-mean(s))

```

## 2.第二问程序 main22.m

```
clear
%从1989年到2008年的城镇登记失业率
a=[2.6, 2.5, 2.3, 2.3, 2.6, 2.8, 2.9, 3, 3.1, 3.1, 3.1, 3.1, 3.6, 4, 4.3, 4.2, 4.2, 4.1, 4, 4.2;];

p5=[-0.2106    -0.0394    0.0415;
     -0.2111    -0.0296    0.0465;
     -0.2284    -0.0373    0.0778;
     -0.2553    -0.0487    0.1010;
     -0.2700    -0.0592    0.0913;
     -0.3001    -0.0771    0.0654;
     -0.2828    -0.0603    0.0683;
     -0.2676    -0.0453    0.0766;
     -0.2844    -0.0498    0.0582;
     -0.2828    -0.0312    0.0668;
     -0.2901    -0.0202    0.0779;
     -0.3219    -0.0274    0.0730;
     -0.3447    -0.0311    0.0614;
     -0.3718    -0.0361    0.0588;
     -0.4192    -0.0455    0.0550;
     -0.4518    -0.0474    0.0562;
     -0.4798    -0.0491    0.0563;
     -0.5087    -0.0525    0.0599;
     -0.5326    -0.0558    0.0636;
     -0.5363    -0.0418    0.0484];

t=a;
% p=vector';
p=p5';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);

%%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:16); trainSamples.T=guiyiTarget(:,1:16);
testSamples.P=guiyiInput(:,17:20); testSamples.T=guiyiTarget(:,17:20);

% % 设置网络参数
for j=1:100
NodeNum1=3; % 第一层节点数
NodeNum2=10; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数，TF3为输出层传输函数
%如果训练结果不理想，可以尝试更改传输函数，以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
```

```

net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2,TypeNum], {TF1 TF2 TF3}, 'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置,应设置为较小值,太大虽然会在开始加快收敛速度,但
临近最佳点时,会产生动荡,而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的p数据,BP得到的结果t

[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的数据p,BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的p数据,BP得到的归
一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p,BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);
All_error=[absTrainError absTestError error_sum];
eps=0.1;%其为3组测试数据的标准差,或者每个数据偏差在一定范围内而判别
% out=[trainOutput validateOutput testOutput];
out=[trainOutput testOutput];
mape0=mape(out,t);
if
((mape0<=0.02)&((abs(absTestError(1))<=eps)&(abs(absTestError(2))<=eps)&(abs(absTes
tError(3))<=eps)|(error_sum<=eps)))
save job_14 net
    break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

```

```

%-----
% 数据分析和绘图
%-----
figure
plot(1989:2008,[trainOutput testOutput],'-b*',1989:2008,t,'--ro');

title('o为真实值,*为预测值')
xlabel('年份');
ylabel('失业率(%)' );
axis([1989,2008,2,4.5])

figure
xx=1:length(All_error);
plot(xx,All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

```

### 3. 第三问程序

(1) 黑龙江省 main3\_hei.m

```

%黑龙江
clear
p_hei=[1.1145    0.0636
        1.2188    0.0633
        1.3352    0.0832
        1.4675    0.0799
        1.6163    0.0778
        1.8780    0.0763
        2.0660    0.0429
        2.3922    0.0232
        2.9777   -0.0738];
a6=[3;3.3;4.7;4.9;4.2;4.5;4.4;4.3;4.26;];

t=a6';
% p=vector';
p=p_hei';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);

%%%以下是自己选的几组

```



```

trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=2; % 第一层节点数
NodeNum2=5; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数, TF3为输出层传输函数
%如果训练结果不理想, 可以尝试更改传输函数, 以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2, TypeNum], {TF1 TF2 TF3},'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置, 应设置为较少值, 太大虽然会在开始加快收敛速度, 但
临近最佳点时, 会产生动荡, 而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的p数据, BP得到的结果t
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的3组数据p, BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的p数据, BP得到的归
一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p, BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);
All_error=[absTrainError absTestError error_sum];
eps=0.1;
out=[trainOutput testOutput];
mape0=mape(out,t);
if ((mape0<=0.05)&(error_sum<=eps))
save job_3_1 net
    break
end

```

```

j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out, t)
Z0=z(out, t)

%-----
% 数据分析和绘图
%-----
figure
plot(1999:2007, [trainOutput testOutput], '-b*', 1999:2007, t, '--ro');

title('o为真实值, *为预测值')
xlabel('年份');
ylabel('失业率(%)');
axis([1999, 2007, 2, 4.5])

figure
xx=1:length(All_error);
plot(xx, All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

(2) 北京 main3_jing.m
%北京
clear
p_jing=[ 1.0558
1.2315
1.3899
1.5341
1.7400
2.0324
2.5770
2.9956
3.5676];
a5=[0.6;0.8;1.2;1.4;1.4;1.3;2.1;2;1.84;];

```

```

t=a5';
% p=vector';
p=p_jing';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);
%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=1; % 第一层节点数
NodeNum2=3; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数, TF3为输出层传输函数
%如果训练结果不理想, 可以尝试更改传输函数, 以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2,TypeNum], {TF1 TF2 TF3},'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置, 应设置为较少值, 太大虽然会在开始加快收敛速度, 但
临近最佳点时, 会产生动荡, 而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的p数据, BP得到的结果t
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的数据p, BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的9组p数据, BP得到的
归一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的9组数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p, BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);

```

```

All_error=[absTrainError absTestError error_sum];
eps=0.15;out=[trainOutput testOutput];
mape0=mape(out,t);
if (mape0<=0.03)&(error_sum<=eps)
save job_3_2 net
break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

%-----
% 数据分析和绘图
%-----
figure
plot(1999:2007,[trainOutput testOutput],'-b*',1999:2007,t,'--ro');

title('o为真实值,*为预测值')
xlabel('年份');
ylabel('失业率(%)' );
axis([1999,2007,0,2.5])

figure
xx=1:length(All_error);
plot(xx,All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

(3) 上海 main3_shang.m
%上海
clear
p_shang=[1.0959 -0.1548
1.2381 -0.1583
1.3533 -0.1644
1.4901 -0.1663
1.7175 -0.1685

```

```

2.0284    -0.1660
2.3935    -0.0870
2.7428    -0.0931
3.2099    -0.0974];
a1=[3.4;3.5;4.3;4.8;4.9;4.5;4.4;4.4;4.22;];

t=a1';
% p=vector';
p=p_shang';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);
%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=2; % 第一层节点数
NodeNum2=5; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数，TF3为输出层传输函数
%如果训练结果不理想，可以尝试更改传输函数，以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput),[NodeNum1,NodeNum2,TypeNum],{TF1 TF2 TF3},'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置，应设置为较少值，太大虽然会在开始加快收敛速度，但
临近最佳点时，会产生动荡，而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的p数据，BP得到的结果t
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的数据p，BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的9组p数据，BP得到的
归一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的9组数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p，BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;

```

```

absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);

All_error=[absTrainError absTestError error_sum];
eps=0.1;out=[trainOutput testOutput];
mape0=mape(out,t);
if (mape0<=0.05)&(error_sum<=eps)
save job_3_1 net
    break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

%-----
% 数据分析和绘图
%-----
figure
plot(1999:2007,[trainOutput testOutput],'-b*',1999:2007,t,'--ro');

title('o为真实值,*为预测值')
xlabel('年份');
ylabel('失业率(%)');
axis([1999,2007,2,5])

figure
xx=1:length(All_error);
plot(xx,All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

```

```

(4) 新疆 main3_xin.m
%新疆
clear
p_xin=[1.0097    -0.0557
        1.1739    -0.0289
        1.2209     0.0073
        1.3677     0.0247
        1.5987     0.0747
        1.8070     0.0955
        2.0921     0.0230
        2.4157    -0.0595
        2.9515    -0.1163];
a2=[3.7;3.8;3.7;3.7;3.5;3.5;3.9;3.9;3.88;];

t=a2';
% p=vector';
p=p_xin';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);
%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=2; % 第一层节点数
NodeNum2=8; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数, TF3为输出层传输函数
%如果训练结果不理想, 可以尝试更改传输函数, 以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2, TypeNum], {TF1 TF2 TF3},'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置, 应设置为较少值, 太大虽然会在开始加快收敛速度, 但
临近最佳点时, 会产生动荡, 而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的数据p, BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的9组p数据, BP得到的
归一化后的结果t

```

```

trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的9组数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p, BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);

All_error=[absTrainError absTestError error_sum];
eps=0.15;out=[trainOutput testOutput];
mape0=mape(out,t);
if (mape0<=0.03)&(error_sum<=eps)
    save job_3_2 net
    break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

%-----
% 数据分析和绘图
%-----
figure
plot(1999:2007,[trainOutput testOutput],'-b*',1999:2007,t,'--ro');

title('o为真实值,*为预测值')
xlabel('年份');
ylabel('失业率(%)');
axis([1999,2007,2,5])

figure
xx=1:length(All_error);
plot(xx,All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:

```



```

theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

(5) 四川 main3_chuan.m
%四川
clear
p_chuan=[ 0.9120
0.9574
1.1273
1.2436
1.4437
1.7683
2.0408
2.4559
3.0728];
a3=[3.7;4;4.3;4.5;4.4;4.4;4.6;4.5;4.24;];

t=a3';
% p=vector';
p=p_chuan';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);
%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=1; % 第一层节点数
NodeNum2=3; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数, TF3为输出层传输函数
%如果训练结果不理想, 可以尝试更改传输函数, 以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2,TypeNum], {TF1 TF2 TF3},'traingda');%
网络创建
% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置, 应设置为较少值, 太大虽然会在开始加快收敛速度, 但
临近最佳点时, 会产生动荡, 而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);

```

```

[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的9组p数据，BP得到的结果t
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的3组数据p，BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的p数据，BP得到的归
一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);%用作变量数据p，BP得到的归一化
的结果t
testInsect = mapminmax('reverse',testSamples.T,ts);%用作变量数据t

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);

All_error=[absTrainError absTestError error_sum];
eps=0.2;out=[trainOutput testOutput];
mape0=mape(out,t);
if (mape0<=0.03)&(error_sum<=eps)
save job_3_3 net
    break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

%-----
% 数据分析和绘图
%-----
figure
plot(1999:2007,[trainOutput testOutput],'-b*',1999:2007,t,'--ro');

title('o为真实值，*为预测值')
xlabel('年份');
ylabel('失业率(%)');
axis([1999,2007,2,5])

figure

```

```

xx=1:length(All_error);
plot(xx, All_error)
title('误差变化图')

```

```

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

```

(6) 广东 main3\_dong.m

```

%广东
clear
p_dong=[ 0.9546 0.0865
1.1024 0.0920
1.1959 0.0993
1.3541 0.1030
1.5631 0.0924
1.7741 0.0854
2.2696 0.0944
2.5649 0.0969
3.0422 0.1100];
a4=[2.3;2.5;2.9;3.1;2.9;2.7;2.6;2.6;2.51];

t=a4';
% p=vector';
p=p_dong';
%mapminmax函数默认将数据归一化到[-1,1]
[guiyiInput,ps]=mapminmax(p);
[guiyiTarget,ts]=mapminmax(t);
%%以下是自己选的几组
trainSamples.P=guiyiInput(:,1:6); trainSamples.T=guiyiTarget(:,1:6);
testSamples.P=guiyiInput(:,7:9); testSamples.T=guiyiTarget(:,7:9);

% % 设置网络参数
for j=1:100
NodeNum1=2; % 第一层节点数
NodeNum2=6; % 隐层第二层节点数
TypeNum=1; % 输出维数
TF1='tansig';TF2='tansig';
TF3='tansig';%各层传输函数, TF3为输出层传输函数
%如果训练结果不理想, 可以尝试更改传输函数, 以下这些是各类传输函数
% TF1 = 'tansig';TF2 = 'logsig';
%TF1 = 'logsig';TF2 = 'purelin';
%TF1 = 'tansig';TF2 = 'tansig';
%TF1 = 'logsig';TF2 = 'logsig';
%TF1 = 'purelin';TF2 = 'purelin';
net=newff(minmax(guiyiInput), [NodeNum1,NodeNum2, TypeNum], {TF1 TF2 TF3}, 'traingda');%
网络创建

```

```

% 设置训练参数
net.trainParam.epochs=10000;%训练次数设置
net.trainParam.goal=1e-4;%训练目标设置
net.trainParam.lr=0.01;%学习率设置,应设置为较少值,太大虽然会在开始加快收敛速度,但
临近最佳点时,会产生动荡,而致使无法收敛
net.trainFcn='traingda';
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],[],[],testSamples);
[guiyiTrainOutput,Pf,Af,E,trainPerf] = sim(net,trainSamples.P,[],[],trainSamples.T);%
正常输入的9组p数据,BP得到的结果t
% [guiyiValidateOutput,Pf,Af,E,validatePerf] =
sim(net,validateSamples.P,[],[],validateSamples.T);%用作变量3的数据p,BP得到的结果t
[guiyiTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],[],testSamples.T);%用
作测试的3组数据p,BP得到的结果t
%反归一化
trainOutput = mapminmax('reverse',guiyiTrainOutput,ts);%正常输入的p数据,BP得到的归
一化后的结果t
trainInsect = mapminmax('reverse',trainSamples.T,ts);%正常输入的数据t
testOutput = mapminmax('reverse',guiyiTestOutput,ts);
testInsect = mapminmax('reverse',testSamples.T,ts);

%绝对误差计算
absTrainError = trainOutput-trainInsect;
absTestError = testOutput-testInsect;
error_sum=0;
for e=1:length(absTestError)
    error_sum=error_sum+absTestError(e).^2;
end
error_sum=sqrt(error_sum);

All_error=[absTrainError absTestError error_sum];
eps=0.2;out=[trainOutput testOutput];
mape0=mape(out,t);
if (mape0<=0.03)&(error_sum<=eps)
save job_3_4 net
    break
end
j
end
j
Min_error_sqrt=min(All_error)
Min_error_sqrt_abs=min(abs(All_error))

testOutput
testInsect
mape0
R=corrcoef(out,t)
Z0=z(out,t)

%-----
% 数据分析和绘图
%-----
figure

```

```

plot(1999:2007, [trainOutput testOutput], '-b*', 1999:2007, t, '--ro');

title('o为真实值, *为预测值')
xlabel('年份');
ylabel('失业率(%)');
axis([1999, 2007, 2, 5])

figure
xx=1:length(All_error);
plot(xx, All_error)
title('误差变化图')

%输入到隐层权值
w1=net.iw{1,1}
%隐层阈值:
theta1=net.b{1}
%隐层到输出层权值:
w2=net.lw{2,1}
%输出层阈值:
theta2=net.b{2}

```

### 三. 回归等模型程序 huigui.m

```

%%%%%%%%%多元线性回归
t=a';
p=vector;
P=[ones(length(t),1) p];
[b0,bint,r,rint,stats]=regress(t,P);
b0
stats
%求绝对平均误差
result=P*b0;
fprintf('绝对平均误差')
mape2=mape(result,t)
fprintf('误差')
error=result-t
fprintf('相对误差')
error_x=error./t
R1=corrcoef(result,t)
Z1=z(result,t)

%%%%%%%%%逐步回归
% stepwise(vector,t)
%%%结果为1, 2, 11
P2=[ones(1,length(t));s;b;cof]';
fprintf('主要指标的系数')
b2=regress(t,P2)
result2=P2*b2;
fprintf('绝对平均误差')
mape3=mape(result2,t)
fprintf('误差')

```

```

error2=result2-t
fprintf(' 相对误差')
error_x2=error2./t
R2=corrcoef(result2,t)
Z2=z(result2,t)

%%%%%%%%%%%%拟合
clc
t=1:20;
p=polyfit(t,a,6)
f=polyval(p,t);
plot(t,a,'o',t,f,'-b*')
xlabel(' 年份')
ylabel(' 城镇登记失业率')
R4=corrcoef(f,a)
mape4=mape(f,a)
z4=z(f,a)

```