

关于我国就业人数的数学模型及前景预测

一、问题描述

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。

请研究生参考就业问题的研究成果，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，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。
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。
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你们的咨询建议。

二、基本假设

- 1、假设国家统计局网、国研网以及和讯网上的统计数据真实有效。
- 2、假设未来不会有大的自然灾害等严重影响国内正常生产生活的不可抗拒因素。
- 3、假设影响城镇就业人口的因素不包括国外经济以及形势等相关数据。
- 4、假设东（或西）部几个省市的平均指标能够代表东（或西）部的指标。

三、变量说明

符号	变量类型	意义	备注
n_y	常量	相关指标年份数据的总个数（样本数据容量）	$n_y=12$
n_s	常量	相关指标季度数据的总个数（样本数据容量）	$n_s=22$
y'	向量	全国就业总人数	
y	向量	全国城镇就业总人数	$y=(y_1, y_2, \dots)$
\hat{y}	向量	全国城镇就业总人数预测值	$\hat{y}=(\hat{y}_1, \hat{y}_2, \dots)$
x_i	向量	影响就业的第 i 个指标	$i=1, \dots, 18$
y_j	标量	全国城镇就业总人数的第 j 个样本数据	$j=n_y \text{ or } n_s$
x_{ij}	标量	第 i 个就业指标的第 j 个样本数据	$i=1, \dots, 18; j=n_y \text{ or } n_s$
\hat{y}_i	标量	全国城镇就业总人数第 i 个样本预测值	$j=n_y \text{ or } n_s$
b_i	标量	回归分析中第 i 个回归系数	$i=0, 1, \dots, m$
ε	向量	随机波动误差向量	

四、就业指标选取

就业是国家的民生大计，影响就业的因素错综复杂。本节首先基于经济学理论初步筛选出影响就业的相关宏观指标；其次，对近年来的数据进行统计分析进一步选取对就业影响相对较大的一些指标；然后，建立城镇就业人数与选取的指标之间的数学模型；最后，考虑分地区、分就业人群等方面，对数学模型进行扩充和完善。

4.1 基于经济学的就业因素和数据采集、处理

4.1.1 基于经济学理论的就业因素初步选取

经过长期的研究和探索，西方经济学家对就业问题（或失业问题）形成了很多行之有效的理论。本文借鉴西方的就业相关理论，从中寻找可能影响中国就业的指标。

凯恩斯经济理论[1]认为：就业水平取决与国家对产品和劳务的总需求水平，国家经济危机和就业不足的根本原因在于私人投资和消费不足而造成的有效需求不足，要实现充分就业，政府必须积极干预。社会消费品总额、商品零售价价格、工业品出厂价格、职工平均工资、居民消费价格指数（CPI）以及政府的干预，包括 M2 货币供应量等因素，与就业有着或多或少的联系。

英国著名学者菲利普斯提出失业与通胀交替理论[2]，发现失业率与通货膨胀率存在此消彼长的关系，通货膨胀率等于工资变动率。但是在中国，文献[3]指出菲利普斯曲线不能很好的解释我国改革开放以来的实际经济情况。我国经济体制转轨以来，由于通货膨胀率的大起大落导致菲利普斯曲线出现“陡峭型变形”。通货膨胀率与国内生产总值（GDP）、对外贸易的进出口总量、国家固定资产投资、商品零售价格、工业品出厂价

格指数等因素存在一定的关系。

配第·克拉克定理[2]揭示经济发展过程中产业结构变化过程：农业劳动力比例下降，第二、三产业等非农业劳动力比例上升。世界经济的发展从劳动力结构、产业结构等方面都表现为由“一、二、三”向“三、二、一”的转变趋势。第二、三产业影响到劳动力结构的分配，对就业有一定的影响。第一产业就业人数、第二产业就业人数、第三产业就业人数、第一产业 GDP、第二产业 GDP、第三产业 GDP 等因素与产业结构息息相关。

奥肯定律[2]指出经济增长与失业率存在密切的关系。文献[4]分析中国特色的经济增长与失业率上升并存现象，指出该现象与人口总数、产业化结构的不合理等因素相关。

基于以上经济学分析，初步选定影响就业的 18 指标如下：

人口总数 x_1	CPI x_2	第三产业就业人数 x_3
工业品出厂价格指数 x_4	商品零售价格指数 x_5	进出口总额 x_6
M2 货币供应量 x_7	社会消费品总额 x_8	社会固定资产投资 x_9
第一产业就业人数 x_{10}	第二产业就业人数 x_{11}	GDP（国内生产总值） x_{12}
第一产业 GDP x_{13}	第二产业 GDP x_{14}	第三产业 GDP x_{15}
人均 GDP x_{16}	就业人员平均工资 x_{17}	财政支出 x_{18}

4.1.2 相关指标数据采集

参考国家统计局网站的国家年度统计年鉴[5]和国研网上的年度宏观数据[6],采集了关于从 1996 年至 2007 年间 18 个初步就业因素数据以及就业人数、城镇就业人数数据。1996-2007 年就业及其相关指标年份数据表，详见附录一。

参考国家统计局网站的季度数据[7]、国研网上的季度宏观数据[8]、月度宏观数据[9]以及和讯网上的月度数据[10],采集了从 1996 年第一季度至今关于城镇就业及其 9 个相关就业因素的季度数据。2004 年第 1 季度至今城镇就业及其相关指标季度数据表，详见附录二。

4.1.3 相关指标数据预处理

在附录二中的季度数据表中，缺少关于人口总数以及第三产业就业人口数的季度数据，仅有相关年度数据。为此，我们采取先数据拟合然后插值的方法，对人口总数以及第三产业就业人口总数的季度数据进行插值。数据拟合的对象为附录一中的年度数据。

利用商业软件 SPSS13.0 进行非线性方程拟合。通过对线性、二次多项式、三次多项式、指数方程、对数方程等 11 种方程进行比较，根据 R^2 最大原则，选取三次多项式时， $R^2=1$ 。因此，选取三次多项式来拟合插值。

人口总数的三次多项式拟合函数为：

$$y = 121060 + 1396.76 \times t - 61.21 \times t^2 + 1.82 \times t^3 \quad (1)$$

第三产业就业人口总数采用三次多项式拟合：

$$y = 18077.18 + 72.77 \times t + 101.56 \times t^2 - 3.88 \times t^3 \quad (2)$$

图 1 给出了人口总数和第三产业就业人口总数的拟合函数、年数据以及 2004 年后季度插值数值图。

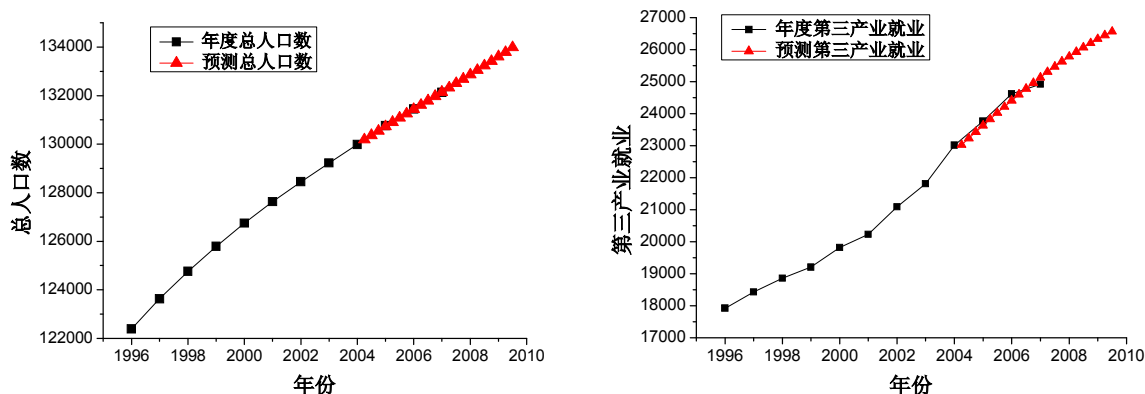


图 1 人口总数和第三产业就业人口总数的年数据以及 2004 年后内插季度数值
由附录一和附录二的数据可知，指标数据有两个明显的特点：

①数据量级相差很大：就业人数为 10^5 量级；总人口数为 10^9 量级；居民价格消费指数以及商品零售价格指数均为 10^2 量级。

②指标单位不统一：就业人数以及总人口数的单位为人；就业人员平均工资报酬单位为元；GDP 的单位为亿元；居民价格消费指数以及商品零售价格指数为无量纲。

为了消除指标数据量级的差别和单位的不统一，对各指标进行归一化处理。归一化处理将矢量的范数置为 1，消除数据量级的差别，归一化后的统计数据介于 0-1 之间；将有量纲的表达式，经过变换化为无量纲，消除单位不统一的影响。

以指标 x_i 年数据为例，对其进行归一化的表达式为：

$$x'_{ij} = \frac{x_{ij}}{\|x_i\|}, \quad \|x_i\| = \left(\sum_{j=1}^{n_y} x_{ij}^2 \right)^{1/2} \quad (3)$$

为了简便，归一化后的指标符号沿用归一化前的指标符号。无特别说明，本文后面均采用归一化的数据。对附录一和附录二中的表格数据进行归一化后的表格分别见附录三和附录四。

4.2 基于统计学分析的就业相关指标

基于经济学原理挑选出来的就业相关指标具有一定的主观性，还需要对就业相关数据进行分析。图 2 给出了全国就业人数和各个变量的年度数据，利用统计学软件 SPSS 分析影响就业的相关指标。

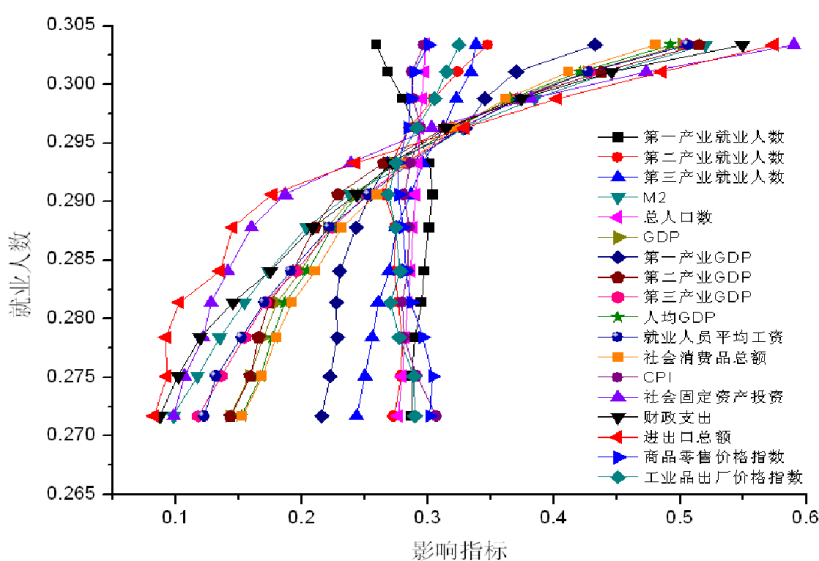


图 2 全国就业人数和各个变量的年度数据

4.2.1 双变量相关分析

双变量相关分析[11]主要是分析两个变量之间是否存在线性相关关系。利用 SPSS 软件中的 Bivariate 功能分别对因变量就业人数和各个指标变量进行双变量相关分析。

图 2 给出各指标分别和就业人数进行相关性分析得到的相关系数图。图中横坐标依次为全国就业人数、第一产业就业人数、第二产业就业人数、第三产业就业人数、M2、人口总数、GDP、第一产业 GDP、第二产业 GDP、第三产业 GDP、人均 GDP、就业人员平均工资、社会消费品总额、CPI、社会固定资产投资、财政支出、对外贸易进出口总额、商品零售价格指数、工业品出厂价格指数。纵坐标为指标与全国就业人数相关性系数。

从图 3 可以看出，全国总人口数、第三产业就业人数、M2、第三产业 GDP、就业人员平均工资、社会消费品总额、财政支出共六个指标与全国就业人数相关性系数超过 0.95；GDP、第二产业 GDP、人均 GDP、社会固定资产投资、对外贸易进出口总额这五个指标与全国就业人数相关性系数超过 0.9；第一产业就业人数、CPI、商品零售价格指数与全国就业人数相关性系数为负，并且系数较小。

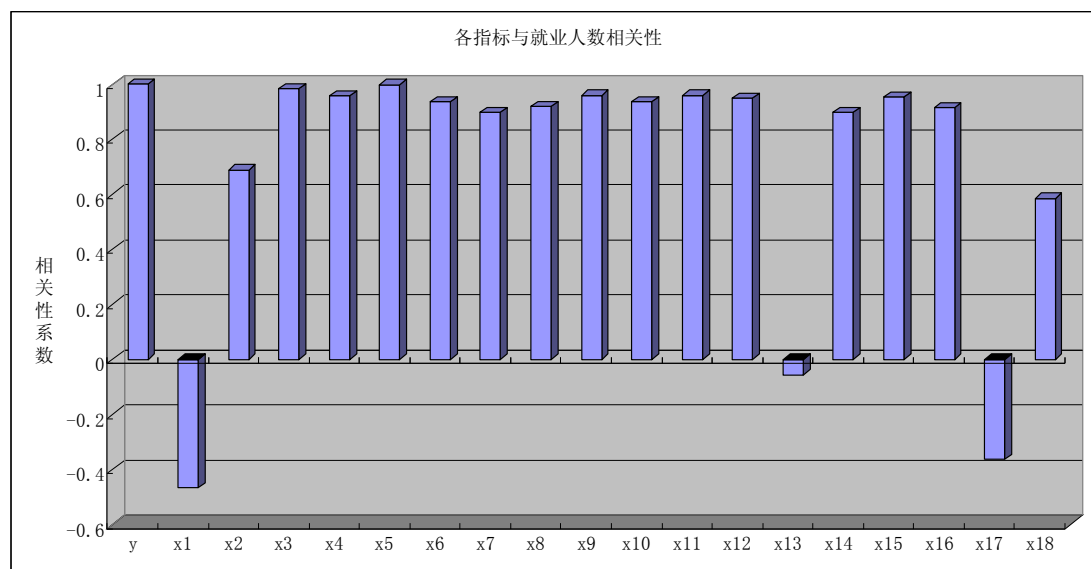


图 3 各指标与就业人数相关性系数图

相关性系数较低的指标并不意味着和就业人数不相关，因为相关指标之间可能存在着高度的依赖性关系或者非线性关系。

4.2.2 逐步回归分析

对多个自变量对因变量的影响程度，可以用多元回归方法进行分析。在就业问题的研究中，影响就业的因素很多，而这些因素之间可能存在着高度的相互依赖性关系时，就会给回归系数的估计带来不合理的解释。

逐步回归分析算法[12]是从一个自变量开始，视自变量对因变量的作用显著程度从大到小地依次逐个引入回归方程。当引入的自变量由于后面变量的引入变得不显著时，就要将其剔除掉。引入一个自变量或从回归方程中剔除一个自变量，为逐步回归的一步。对于每一步都要进行 F 检验，以确保每次引入新的显著性变量前回归方程中只包含对因变量作用显著的变量。这个过程反复进行，直至既无不显著的变量从回归方程中剔除，又无显著变量可引入回归方程为止。所以，逐步回归分析可以有效地从众多影响因变量的因素中挑选出贡献最大的变量，在它们和因变量的观测数据基础上建立最优的回归方程。

针对 1996-2007 年各指标的年度数据以及全国就业人数年度数据，利用 SPSS 软件中的 step regression 功能进行分析，得到结果如下：

	Step1	Step2	Step3	Step4	...	Step9
逐步选取指标	总人口数	总人口数 CPI	总人口数 CPI 第三产业就业人数	总人口数 CPI 第三产业就业人数 工业品出厂价格指数		总人口数 CPI 第三产业就业人数 工业品出厂价格指数 全国商品零售价格指数 出口总额 M2 社会消费品总额 社会固定资产投资

表 1 全国就业人数与各指标逐步回归所选取指标

逐步回归分析结果的评判参数如下：拟合判断系数 $R^2=0.99$ ，调整的判断系数 $\bar{R}^2=0.99$ ，回归估计的标准差 $S=0.0001035$ 。由评判参数可知，逐步回归分析效果很好。

从表 1 可以看出，全国总人口数、CPI、第三产业就业人数、工业品出厂价格指数、全国商品零售价格指数、对外贸易进出口总额、M2、社会消费品总额、社会固定资产投资这九个指标综合对全国就业人数的影响很大。

4.2.3 就业指标数据选取

利用 1996-2007 年度数据对影响就业的相关指标分析，双变量相关分析和逐步回归分析两种方法的结果不尽相同。大致情况如下：

①全国总人口数、第三产业就业人数、M2、社会消费品总额、社会固定资产投资、对外贸易进出口总额这六个指标在两次的分析中均对就业人数较大的影响。

②GDP、人均 GDP、第二产业 GDP、第三产业 GDP、就业人员平均工资、财政支出这个六个指标，在双变量相关分析中影响较大，但是在逐步回归分析中影响较小。由于逐步回归分析中选取的指标或者指标之间的组合能够取代这些指标对就业的影响。因此在因素选取中剔出这六个指标。

③CPI、工业品出厂价格指数、全国商品零售价格指数这三个指标，在双变量相关分析中影响较小，但是在逐步回归分析中影响较大。由于相关性系数较低的指标并不意味着和就业人数不相关，因为相关性系数只是变量间线性相关强弱的度量，变量间还可能存在非线性相关关系或者多个变量的合作效应。所以，这三个指标对就业的影响不能忽视。

综合上述情况可知，选定影响就业较大的 9 个指标如下：

人口总数 x_1	CPI x_2	第三产业就业人数 x_3
工业品出厂价格指数 x_4	商品零售价格指数 x_5	进出口总额 x_6
M2 x_7	社会消费品总额 x_8	社会固定资产投资 x_9

五、城镇就业人数模型建立与求解

5.1 关于城镇就业人数的初略数学模型

通过对年度数据的统计分析，选定了影响就业的 9 个主要指标。本节旨在建立城镇就业人数与上述 9 个主要指标之间数学模型。

利用已有的年度数据或者季度数据，以城镇就业人数 y 为因变量，影响就业的 9 个指标 (x_1, \dots, x_9) 为自变量，寻找线性或非线性回归函数 $f(b_0, b_1, \dots, b_m, x_{1i}, x_{2i}, \dots, x_{9i})$ ，使得城镇就业人数的样本数据与估计的城镇就业人数的误差平方和最小。

城镇就业人数数学模型 I 为：

$$\begin{aligned} \text{目标函数: } & \begin{cases} y = f(b_0, b_1, \dots, b_m; x_1, x_2, \dots, x_9) + \varepsilon \\ E(\varepsilon) = 0 \end{cases} \\ s.t. \quad & \begin{cases} \min R(b_0, b_1, \dots, b_m) = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \\ \hat{y}_i = f(b_0, b_1, \dots, b_m, x_{1i}, x_{2i}, \dots, x_{9i}) \quad i = 1, 2, \dots, n \\ m < n - 1 \\ n = \begin{cases} n_y & \text{选取年度数据} \\ n_s & \text{选取季度数据} \end{cases} \\ f(b_0, b_1, \dots, b_m, x_{1i}, x_{2i}, \dots, x_{9i}) \in C(x) \\ b_0, b_1, \dots, b_m \in R \end{cases} \end{aligned}$$

5.2 数学模型 I 的求解

由于回归函数的类型尚未确定，首先分析各个变量与城镇就业人数年度数据之间的关系。图 3 给出了归一化后的城镇就业人数和各个影响变量年度数据关系。选取的 9 个影响就业指标并非都与因变量之间呈线性关系，如图 4 所示。

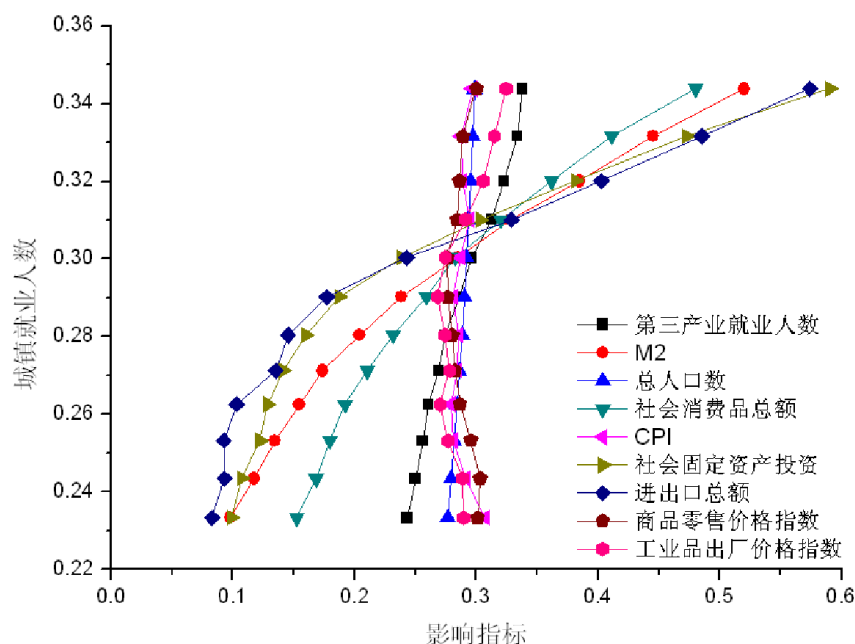


图 4 归一化后的城镇就业人数和各个影响变量年度数据

下面将分别选取线性函数和非线性函数作为回归函数，寻求城镇就业人数与各个影响指标之间的联系。

5.2.1 基于年度数据的非线性回归

利用 SPSS 软件对各个变量分别和城镇就业人数进行曲线拟合，采用 7 种不同形式的函数（线性函数、对数函数、倒数函数、二次函数、三次函数、幂函数、指数函数）进行拟合。表 2 给出不同函数拟合各指标分别与城镇就业人数评价的决定系数 R^2 值，此值越接近 1 说明拟合效果越好。

	人口	CPI	三产就业	工业出厂	商品零售	进出口总	M2	消费总额	社会投资
Linear	.985	.673	.983	.417	.081	.894	.959	.947	.869
Log	.983	.675	.988	.404	.079	.972	.999	.994	.970
Inverse	.980	.677	.991	.390	.078	.964	.944	.993	.989
Quadratic	.999	.673	.989	.569	.143	.947	.993	.994	.954
Cubic	.999	.693	.989	.566	.145	.979	.999	.999	.987
Power	.994	.669	.981	.355	.109	.956	.998	.982	.950
Exponential	.996	.667	.971	.368	.111	.859	.935	.921	.831
选择函数	Linear	Linear	Inverse	Linear	Linear	Log	Log	Log	Inverse

表 2 采取不同函数拟合各指标分别与城镇就业人数的效果

由于样本容量不是很大，在保证精度的前提下，应选择尽可能简单的函数，根据各个函数的表达式形式，也就是要首先要尽可能选择线性函数、对数函数、倒数函数，其次是二次函数、幂函数、指数函数，最后是三次函数。

基于上述基本原则，选择的非线性回归函数形式为：

$$y = b_0 + b_1x_1 + b_2x_2 + \frac{b_3}{x_3} + b_4x_4 + b_5x_5 + b_6 \ln x_6 + b_7 \ln x_7 + b_8 \ln x_8 + \frac{b_9}{x_9}$$

针对城镇就业人数的年度数据以及各相关指标的年度数据，利用 SPSS 软件 Regression 功能下的 Curve Estimation 模块，对参数 b_0, b_1, \dots, b_9 进行估计，得到城镇就业人数与各个影响指标的回归方程为：

$$y = -0.104056 + 1.52088x_1 - 0.012969x_2 + \frac{0.018367}{x_3} + 0.025674x_4 - 0.045976x_5 + 0.002430 \ln x_6 + 0.015411 \ln x_7 + 0.052460 \ln x_8 - \frac{0.000798}{x_9}$$

表 3 给出了非线性回归参数估计值、标准差以及 95%置信区间。

Parameter	Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
b0	-.104056	.457190	-2.071183	1.863072
b1	1.520888	1.339750	-4.243589	7.285365
b2	-.012969	.150841	-.661987	.636049

b3	.018367	.035573	-.134689	.171423
b4	.025674	.136878	-.563265	.614612
b5	-.045976	.176877	-.807017	.715065
b6	.002430	.018366	-.076594	.081455
b7	.015411	.047019	-.186895	.217717
b8	.052460	.041717	-.127034	.231953
b9	-.000798	.004350	-.019514	.017918

表3 非线性回归参数估计值、标准差以及置信区

表4给出了非线性回归的主要统计量。由于是非线性回归，不能用方差分析方法来进行F检验，但由于回归平方和比残差平方和数值大得多，说明回归效果比较好。 R^2 值为0.999856说明残差平方和在总平方和中占得比例非常小，同样说明回归效果比较好。

Source	Sum of Squares	df	Mean Squares
Regression	1.000039	10.000000	.100004
Residual	.000002	2.000000	.000001
Uncorrected Total	1.000042	12.000000	
Corrected Total	.013851	11.000000	
R squared =	1 - (Residual Sum of Squares) / (Corrected Sum of Squares) = 0.999856		

表4 非线性回归统计量

表5给出了不同年度下实际城镇就业人数与非线性回归预计城镇就业人数的比较情况。第二、三、四列数据是经过归一化处理的城镇就业人数相关数据，第五、六、七列是重新还原经过归一化处理的数据。图5给出了实际城镇就业人数与预测城镇就业人数的比较图。从表5和图5可以看出，实际与预测的结果误差较小。

年份	实际城镇就业人数归一化处理值	预计城镇就业人数处理值	数据归一化后绝对误差	实际城镇就业人数（万人）	预测城镇就业人数（万人）	城镇就业人数绝对误差（万人）
1996	.2333	.2331	.0002	19922.00	19901.91515	20.08484851
1997	.2434	.2441	-.0007	20781.00	20841.08747	-60.08746666
1998	.2532	.2531	.0001	21616.00	21609.50118	6.498820925
1999	.2625	.2617	.0008	22412.00	22343.76317	68.23682906
2000	.2712	.2709	.0003	23151.00	23129.25274	21.74725638
2001	.2804	.2810	-.0006	23940.00	23991.58369	-51.58368755
2002	.2902	.2906	-.0004	24780.00	24811.22498	-31.2249808
2003	.3003	.3002	.0001	25639.00	25630.86627	8.133725965
2004	.3101	.3098	.0003	26476.00	26450.50757	25.49243272
2005	.3201	.3207	-.0006	27331.00	27381.14195	-50.14195231
2006	.3316	.3312	.0004	28310.00	28277.62462	32.37538321
2007	.3438	.3437	.0001	29350.00	29344.86588	5.134115969

表5 基于年数据非线性回归预计值与实际数据比较

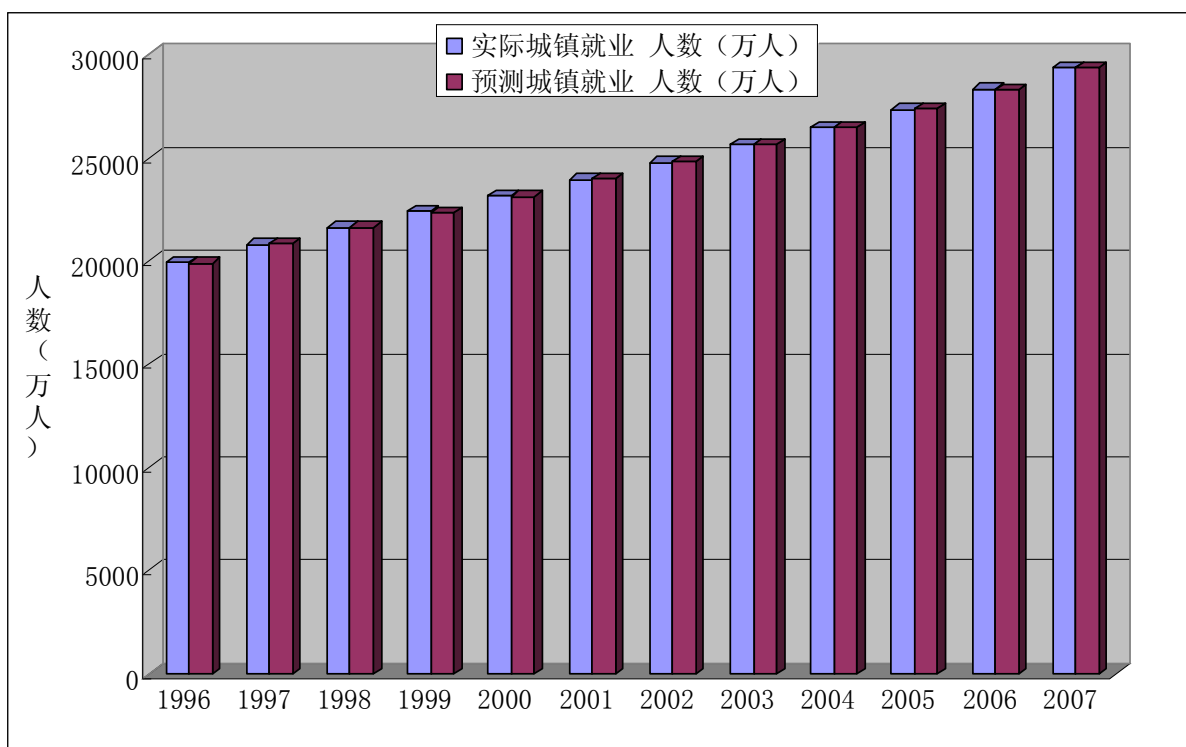


图 5 实际城镇就业人数与预测城镇就业人数的比较

5.2.2 基于年度数据的线性回归

从表 2 可以看出，线性函数对各指标与城镇就业人数的拟合均较好，CPI、工业出厂价格指数稍微差一些，但是若干变量间还可能存在合作效应。本小节尝试利用线性回归方法分析对城镇就业人数的模型求解。

线性回归函数形式为：

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9$$

其中指标 x_1, \dots, x_9 分别为人口总数，CPI，第三产业就业人数，工业品出厂价格指数，商品零售价格指数，进出口总额，M2，社会消费品总额，社会固定资产投资。

针对城镇就业人数的年度数据以及各相关指标的年度数据，利用 SPSS 软件 Regression 功能下的 Linear 模块，对参数 b_0, b_1, \dots, b_9 进行估计，得到城镇就业人数与各个影响指标的回归方程为：

$$y = -0.561198 + 2.925547x_1 + 0.017422x_2 - 0.02203x_3 - 0.0773x_4 - 0.021458x_5 + 0.03409x_6 + 0.071356x_7 + 0.043968x_8 - 0.020761x_9$$

由表 6 中的 R2、修正的 R2 和标准误差值可知，线性回归的效果不错。

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999995(a)	.999990	.999943	.000268

表6 线性回归常用统计量数据

表7给出了线性回归的方差分析结果。F检验值为12553.228388，其显著性概率值为0.00008，即拒绝回归系数均为0的假设。

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.001209	9.000000	.000134	12553.228388	.000080(a)
	Residual	.000000	2.000000	.000000		
	Total	.001209	11.000000			

表7 线性回归方差分析结果

表8给出了线性回归系数分析数据。T值为t检验为偏回归系数为0的假设检验值，Sig.为偏回归系数为0的假设检验的显著性水平值。方差分析是对整个回归方程的显著性检验，它与单独进行每一个偏回归系数的显著性检验不一定等效。表8中各个参数的显著性水平Sig.差别比较大，大部分低于0.5。虽然个别偏回归系数不显著，但是从整体上而言，回归方程的系数是显著的。

	UnStandardized Coefficient		Standardized Coefficient	T	Sig.	Correlation		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	-.561198	.078837		-7.118461	.019169	-.561198	.078837	
第三产业就业人数	-.022023	.081898	-.020693	-.268905	.813202	-.022023	.081898	-.020693
M2	.071356	.029945	.276028	2.382922	.140043	.071356	.029945	.276028
人口	2.925547	.219350	.586564	13.337344	.005575	2.925547	.219350	.586564
社会消费品总额	.043968	.036419	.128809	1.207267	.350734	.043968	.036419	.128809
CPI	.017422	.033842	.003845	.514803	.657938	.017422	.033842	.003845
社会固定资产投资	-.020761	.032548	-.093971	-.637852	.588855	-.020761	.032548	-.093971
进出口总额	.034090	.036160	.163091	.942749	.445324	.034090	.036160	.163091
商品零售价格指数	-.021458	.039044	-.005749	-.549584	.637776	-.021458	.039044	-.005749
工业品出厂价格指数	-.077300	.039302	-.039708	-1.966793	.188098	-.077300	.039302	-.039708

表8 线性回归系数分析数据

表9给出了线性回归的总体统计量。回归平方和比残差平方和数值大得多， R^2 值为0.999856说明残差平方和在总平方和中占得比例非常小，说明回归效果比较好。

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.233342	.343833	.286675	.0354848	12
Residual	-.0002472	.0001821	.0000000	.0001144	12
Std. Predicted Value	-1.503	1.611	.000	1.000	12
Std. Residual	-.921	.678	.000	.426	12

表9 线性回归的总体统计量

图6描述的是城镇就业人数预测值与各变量的残差散点分布图，大部分观测量随机地落在 ± 0.5 的范围内，预测值与各变量的残差值之间没有明显的关系，所以回归方程满足线性与方差齐次性假设。

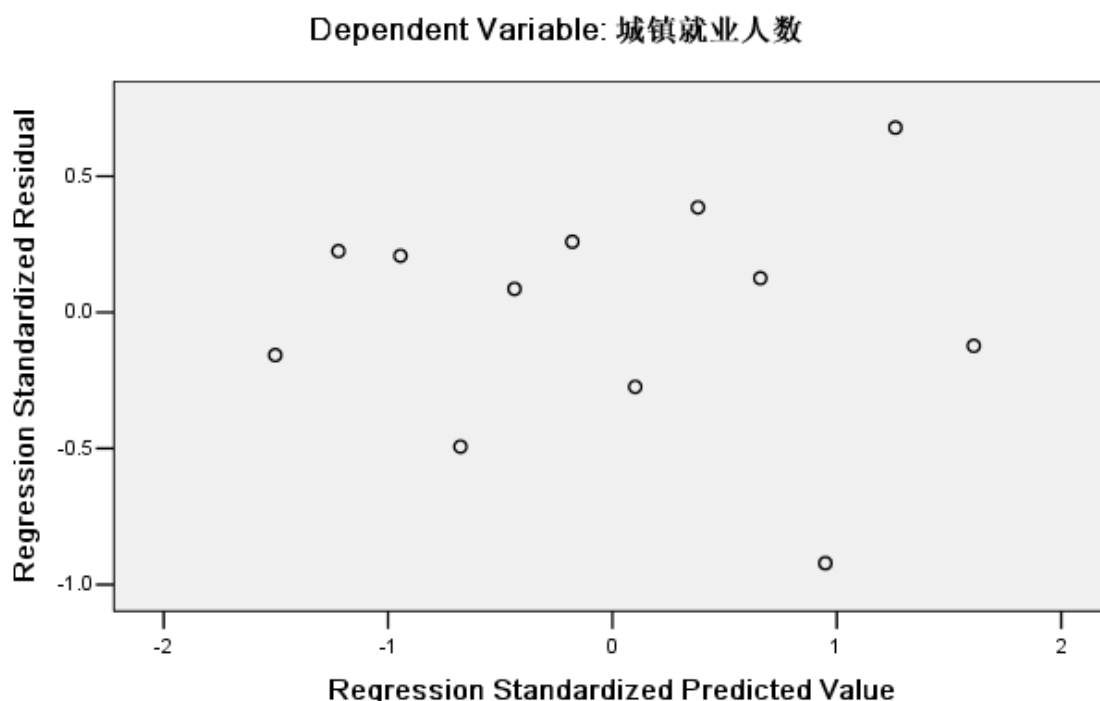


图 6 城镇就业人数预测值与各变量的残差散点分布

表10给出了不同年度下实际城镇就业人数与线性回归预计城镇就业人数的比较情况。从表10可以看出，实际与预测的结果误差相对较小。

年份	实际城镇就业人数归一化处理值	预计城镇就业人数归一化处理值	数据归一化值的绝对误差	实际城镇就业人数（万人）	预测城镇就业人数（万人）	城镇就业人数绝对误差（万人）
1996	.2333	.23334	-0.00004	19922.00	19922.40618	-0.406183822
1997	.2434	.24334	0.00006	20781.00	20776.1992	4.800802386
1998	.2532	.25314	0.00006	21616.00	21612.91635	3.08364887
1999	.2625	.26263	-0.00013	22412.00	22423.16592	-11.16592122
2000	.2712	.27118	0.00002	23151.00	23153.15895	-2.158948011
2001	.2804	.28033	0.00007	23940.00	23934.37956	5.620444369
2002	.2902	.29027	-0.00007	24780.00	24783.04981	-3.04981134
2003	.3003	.30020	0.00010	25639.00	25630.86627	8.133725965
2004	.3101	.31007	0.00003	26476.00	26473.55998	2.440021352
2005	.3201	.32035	-0.00025	27331.00	27351.2592	-20.25919683
2006	.3316	.33142	0.00018	28310.00	28296.40806	13.59193691
2007	.3438	.34383	-0.00003	29350.00	29355.96519	-5.96519321

表10 基于年度数据线性回归预计值与实际数据比较

5.2.3 基于年度数据的非线性回归和线性回归比较

根据 1996-2007 年的全国城镇就业人数和 9 个指标的年度数据，分别进行了非线性回归和线性回归。将两者的计算结果及其评价指标进行对比：

(1) 决定系数 R^2 值比较：非线性回归的决定系数值为 0.999856，线性回归的决定

系数值为 0.999990。两种方法的决定系数都很高，非常接近于 1，回归效果都比较好。

(2) 回归系数标准差的比较：表 11 给出了线性回归和非线性回归的系数标准差比较。从表 11 可以看出，线性回归的系数标准差相对非线性回归的系数标准差要小一些。

	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
nonlinear	0.4572	1.3400	0.1508	0.0356	0.1368	0.1769	0.0184	0.0470	0.04172	0.0044
linear	0.0788	0.2193	0.0338	0.0819	0.0393	0.0390	0.0362	0.0299	0.0364	0.0325

表 11 线性回归和非线性回归的系数标准差比较

(3) 预计值误差的比较：根据表 5 和表 10 中的城镇就业人数实际与预测的误差数据，线性回归预测和非线性回归预测的误差如图 7 所示。从图 7 可以看出，线性回归所预测的城镇就业人数更加符合真实的数据。

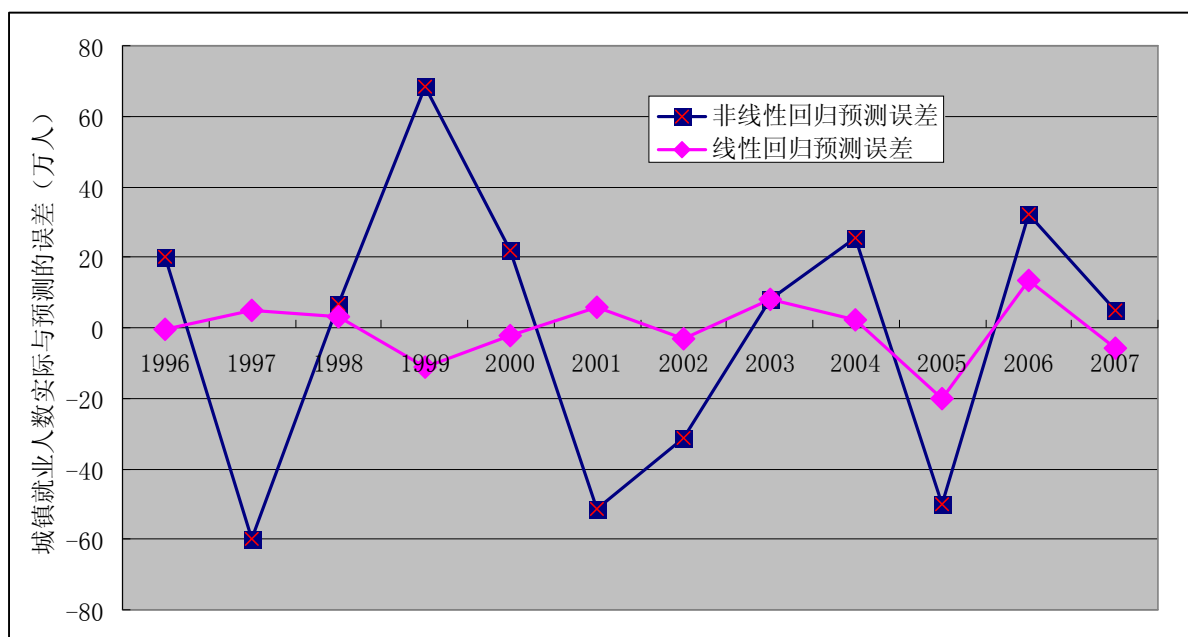


图 7 不同回归方法的就业人数实际与预测的误差数据比较

以 1994-2007 年城镇就业人数和各指标的数据为样本，根据 (1) - (3) 的分析，线性回归对于分析城镇就业问题稍逊一筹，通过九个相关指标较真实全面地反映城镇就业人数问题。

5.2.4 基于季度数据的线性回归

以 04 年第一季度到 09 年第二季度数据为样本，利用线性回归寻找城镇就业人口与其相关指标之间关系。

线性回归函数形式为：

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9$$

针对城镇就业人数的季度数据以及各相关指标的季度数据，利用 SPSS 软件 Regression 功能下的 Linear 模块，对参数 b_0, b_1, \dots, b_9 进行估计，得到城镇就业人数与各个影响指标的回归方程为：

$$y = -3.291899 + 18.410435x_1 + 0.00673x_2 - 1.370213x_3 - 0.019534x_4 + 0.052679x_5 - 0.05026x_6 - 0.025268x_7 - 0.00176x_8 + 0.000807x_9$$

由表 12 中的 R²、修正的 R² 和标准误差值可知，线性回归的效果不错。

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999988(a)	.999977	.999960	.000128

表12 基于季度数据的线性回归常用统计量

表13给出了线性回归的方差分析结果。F检验值为70363.131054，其显著性概率值为0.00000，即拒绝回归系数均为0的假设。

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.008571	9.000000	.000952	57676.713101	.000000(a)
	Residual	.000000	12.000000	.000000		
	Total	.008571	21.000000			

表13 基于季度数据的线性回归方差分析结果

表14给出了基于季度数据的线性回归系数分析数据。表8中各个参数的显著性水平Sig.只有个别大于0.5。虽然个别偏回归系数不显著，但是从整体上而言，回归方程的系数是显著的。

	UnStandardized Coefficient		Standardized Coefficient	T	Sig.
	B	Std. Error	Beta		
(Constant)	-3.291899	.250428		-13.145068	.000000
总人口数	18.410435	1.337652	1.719142	13.763250	.000000
第三产业就业人数	-1.370213	.155607	-.643287	-8.805603	.000001
M2	-.025268	.018494	-.070780	-1.366230	.196923
全国商品零售价格指数	.052679	.022334	.014166	2.358621	.036142
工业品出厂价格指数	-.019534	.014982	-.008384	-1.303891	.216729
社会消费品零售总额	-.001760	.002197	-.004718	-.801170	.438609
CPI	.006730	.020900	.001788	.322014	.752983
进出口总额	-.005026	.006143	-.015665	-.818293	.429140
固定资产投资	.000807	.001443	.003768	.559435	.586155

表14 基于季度数据的线性回归系数分析

图8描述的是基于季度数据的城镇就业人数预测值与各变量的残差散点分布图，大部分观测量随机地落在±1.0的范围内，预测值与各变量的残差值之间没有明显的关系，所以回归方程满足线性与方差齐次性假设。

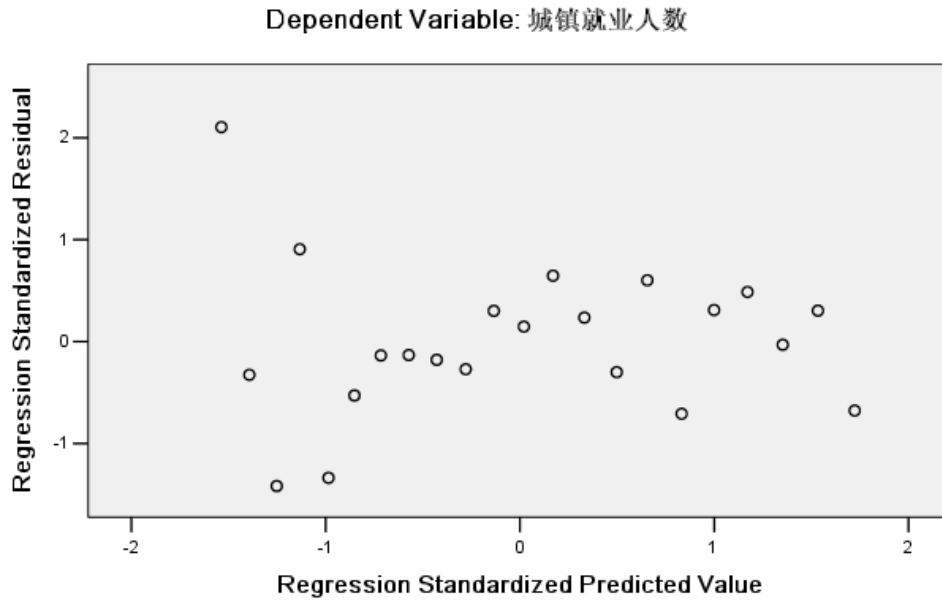


图 8 基于季度数据的城镇就业人数预测值与各变量的残差散点分布

表 15 给出了不同季度下实际城镇就业人数与线性回归预计城镇就业人数的比较情况。从表 15 可以看出，实际与预测的结果误差较小。

季度	实际城镇就业人数归一化处理值	预测城镇就业人数归一化处理值	数据归一化处理后值绝对误差	实际城镇就业人数（万人）	预测城镇就业人数（万人）	城镇就业人数绝对误差（万人）
04-1	.3125	.31223	0.00027	266577	266600.0524	23.05241
04-2	.3151	.31514	0.00004	26906	26902.58483	-3.41517
04-3	.3178	.31798	0.00018	27148	27132.6317	-15.3683
04-4	.3205	.32038	0.00012	27331	27341.24552	10.24552
05-1	.3232	.32337	0.00017	27609	27594.4855	-14.5145
05-2	.3260	.32607	0.00007	27839	27833.02345	-5.97655
05-3	.3288	.32882	0.00002	28074	28072.29241	-1.70759
05-4	.3317	.33172	0.00002	28331	28329.29241	-1.70759
06-1	.3346	.33462	0.00002	28569	28567.29241	-1.70759
06-2	.3376	.33763	0.00003	28826	28823.43862	-2.56138
06-3	.3406	.34056	0.00004	29076	29079.41517	3.415172
06-4	.3437	.34368	0.00002	29350	29351.70759	1.707586
07-1	.3468	.34672	0.00008	29602	29608.83034	6.830344
07-2	.3500	.34997	0.00003	29880	29882.56138	2.561379
07-3	.3533	.35334	0.00004	30167	30163.58483	-3.41517
07-4	.3566	.35652	0.00008	30439	30445.83034	6.830344
08-1	.3600	.36009	0.00009	30744	30736.31586	-7.68414
08-2	.3635	.36346	0.00004	31031	31034.41517	3.415172
08-3	.3670	.36694	0.00006	31329	31334.12276	5.122758
08-4	.3706	.37060	0.00000	31641	31641	0
09-1	.3743	.37426	0.00004	31954	31957.41517	3.415172
09-2	.3780	.37809	0.00009	32281	32273.31586	-7.68414

表 15 基于季度数据线性回归预计值与实际数据比较

5.2.5 基于季度数据和年度数据的线性回归比较

前面分别做基于 1996-2007 年年度数据和 2004 年第一季度至 2009 年第二季度的季度数据的线性回归，由于样本容量以及数据的差异，线性回归的结果有所差异：

(1) 决定系数 R^2 值比较：基于年度数据和季度数据的线性回归的决定系数值分别为 0.999990 和 0.999977。两种方法的决定系数都很高，非常接近于 1，回归效果都比较好。

(2) 回归系数标准差的比较：表 16 给出了线性回归和非线性回归的系数标准差比较。从表 11 可以看出，线性回归的系数标准差相对非线性回归的系数标准差要小一些。

	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
linear_year	0.0788	0.2193	0.0338	0.0819	0.0393	0.0390	0.0362	0.0299	0.0364	0.0325
linear_season	0.2504	1.3377	0.01556	0.01849	0.0223	0.0149	0.0021	0.0209	0.0061	0.0014

表 16 基于年度数据和季度数据的线性回归系数标准差比较

(3) 预计值误差的比较：年度数据和季度数据在 04 年至 07 年存在交集。根据表 10 和表 15 中的城镇就业人数实际与预测的误差数据，基于年度数据的线性回归预测和基于季度数据的线性回归预测的误差如图 9 所示。从图 7 可以看出，两种不同数据的最大误差都差不多，但是基于季度数据的线性回归更加集中于 0 附近，而且随着时间的递增，实际与预测的误差越来越趋于 0。

(4) 样本容量的比较：年度数据的样本个数为 12，季度数据的样本个数为 22。大样本容量的数据相对来说真实可靠一些。

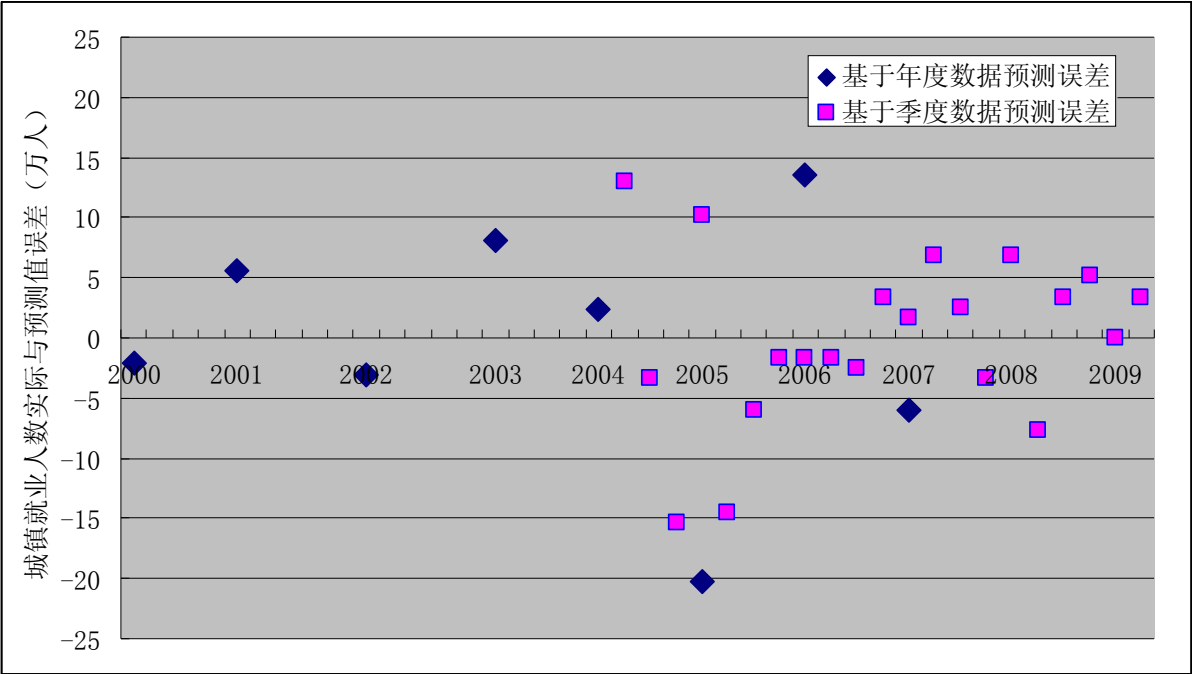


图 9 不同回归方法的就业人数实际与预测的误差数据比较

基于 (1) - (4) 的分析，综合考虑近期数据对数据预测的影响更大，认为基于季度数据的线性回归更能较真实全面地反映城镇就业人数问题，能够更好的为预测 09 年

下半年以及 2010 年的数据。

六、基于地区、行业等城镇就业人数模型建立与求解

在文中第四部分中，模型 I 的影响就业的指标包括第三产业就业人数、M2、人口总数、社会消费品总额、CPI、社会固定资产投资、对外贸易进出口总额、商品零售价格指数、工业品出厂价格指数。其中 M2、社会消费品总额、CPI、社会固定资产投资、对外贸易进出口总额、商品零售价格指数、工业品出厂价格指数为主要经济指标。本文中的就业人群划分依据按照学历。本节将从分行业、分地区、分就业人群的角度来分析城镇就业人员与影响就业的指标之间的关系。

6.1 基于地区、行业、就业人群的城镇就业人数模型建立

不同地区的城镇就业人数差异很大，不同的行业的城镇就业人数也大相径庭，不同人群的就业人数也有所差异。为此，下面将分别建立基于不同地区、不同行业、不同就业人群的城镇就业人数模型。

6.1.1 基于不同地区的城镇就业人数模型

由于地区的城镇就业人数与地区的经济指标等有密切的关系。以某地区为例，依旧考虑影响就业的 9 个主要指标 (x_1, \dots, x_9) ，选取相应的该地区九个影响指标 $(x_1^{(a)}, x_2^{(a)}, \dots, x_9^{(a)})$ 作为自变量，某地区的城镇就业人数 $y^{(a)}$ 为因变量，寻找线性回归函数 $f^{(a)}(b_0^{(a)}, b_1^{(a)}, \dots, b_9^{(a)}; x_1^{(a)}, x_2^{(a)}, \dots, x_9^{(a)})$ ，建立城镇就业人数与上述 9 个主要指标之间数学模型。

因此，基于地区的城镇就业人数数学模型 II 为：

$$\begin{cases} y^{(a)} = f^{(a)}(b_0^{(a)}, b_1^{(a)}, \dots, b_9^{(a)}; x_1^{(a)}, x_2^{(a)}, \dots, x_9^{(a)}) + \varepsilon \\ E(\varepsilon) = 0 \end{cases}$$

$$s.t. \begin{cases} \min R^{(a)}(b_0^{(a)}, b_1^{(a)}, \dots, b_9^{(a)}) = \sqrt{\sum_{i=1}^n (y_i^{(a)} - \hat{y}_i^{(a)})^2} \\ \hat{y}_i^{(a)} = b_0^{(a)} + b_1^{(a)}x_{1i}^{(a)} + \dots + b_9^{(a)}x_{9i}^{(a)} \quad i = 1, 2, \dots, n \\ b_0^{(a)}, b_1^{(a)}, \dots, b_9^{(a)} \in R \end{cases}$$

6.1.2 基于不同行业的城镇就业人数模型

由于不同行业的城镇就业人数与全国的经济指标等因素由密切关系。以某行业为例，选取影响就业的九个全国性指标 (x_1, \dots, x_9) 作为自变量，某行业的城镇就业人数 $y^{(b)}$ 为因变量，寻找线性回归函数 $f^{(b)}(b_0^{(b)}, b_1^{(b)}, \dots, b_9^{(b)}; x_1, x_2, \dots, x_9)$ ，建立城镇就业人数与上述 9 个主要指标之间数学模型。

因此，基于行业的城镇就业人数数学模型 III 为：

$$\begin{cases} y^{(b)} = f^{(b)}(b_0^{(b)}, b_1^{(b)}, \dots, b_9^{(b)}; x_1, x_2, \dots, x_9) + \varepsilon \\ E(\varepsilon) = 0 \end{cases}$$

$$s.t. \begin{cases} \min R^{(b)}(b_0^{(b)}, b_1^{(b)}, \dots, b_9^{(b)}) = \sqrt{\sum_{i=1}^n (y_i^{(b)} - \hat{y}_i^{(b)})^2} \\ \hat{y}_i^{(b)} = b_0^{(b)} + b_1^{(b)}x_{1i} + \dots + b_9^{(b)}x_{9i} \quad i = 1, 2, \dots, n \\ b_0^{(b)}, b_1^{(b)}, \dots, b_9^{(b)} \in R \end{cases}$$

6.1.3 基于不同学历的城镇就业人数模型

由于不同就业人群的城镇就业人数与全国的经济指标等因素有密切关系。考虑不同就业人群按照不同的学历划分，以大专及其以上学历的城镇就业人口 $y^{(e)}$ 作为因变量，影响就业的九个全国性指标 (x_1, \dots, x_9) 作为自变量，寻找线性回归函数 $f^{(e)}(b_0^{(e)}, b_1^{(e)}, \dots, b_9^{(e)}; x_1, x_2, \dots, x_9)$ ，建立城镇就业人数与上述 9 个主要指标之间数学模型。

因此，基于不同学历教育的城镇就业人数数学模型 IV 为：

$$\begin{cases} y^{(e)} = f^{(e)}(b_0^{(e)}, b_1^{(e)}, \dots, b_9^{(e)}; x_1, x_2, \dots, x_9) + \varepsilon \\ E(\varepsilon) = 0 \end{cases}$$

$$s.t. \begin{cases} \min R^{(e)}(b_0^{(e)}, b_1^{(e)}, \dots, b_9^{(e)}) = \sqrt{\sum_{i=1}^n (y_i^{(e)} - \hat{y}_i^{(e)})^2} \\ \hat{y}_i^{(e)} = b_0^{(e)} + b_1^{(e)}x_{1i} + \dots + b_9^{(e)}x_{9i} \quad i = 1, 2, \dots, n \\ b_0^{(e)}, b_1^{(e)}, \dots, b_9^{(e)} \in R \end{cases}$$

6.2 数学模型求解

数学模型的求解按照不同地区、不同行业、不同就业人群分别进行求解。

6.2.1 基于不同地区的城镇就业人数模型求解及比较

不同地区的指标季度数据采集主要参考中国统计年鉴[7]，为了方便计算，对江苏、浙江、福建、山东四省的指标季度数据进行平均，代表东部地区的指标季度数据；同样的，选取四川、贵州、云南三省的平均季度数据作为西部地区季度数据，具体数据见附录五。此外，由于 M2 是全国性指标，因此 M2 的地区性季度数据用全国性的季度数据代替。

6.2.1.1 东部地区的城镇就业人数模型求解

选取东部地区的九个影响指标 $(x_1^{(a1)}, x_2^{(a1)}, \dots, x_9^{(a1)})$ 作为自变量，东部地区的城镇就业人数 $y^{(a1)}$ 为因变量，利用 SPSS 软件采取先筛选指标然后进行线性回归分析方法，对

参数 $b_0^{(a1)}, b_1^{(a1)}, \dots, b_9^{(a1)}$ 进行估计, 得到东部地区的城镇就业人数与各个影响指标的回归方程为:

$$y^{(a1)} = 0.105746 + 0.308086x_2^{(a1)} + 0.042327x_4^{(a1)} + 0.079469x_5^{(a1)} + 0.161814x_8^{(a1)} + 0.026476x_9^{(a1)}$$

其中指标 $x_2^{(a1)}, x_4^{(a1)}, x_5^{(a1)}, x_8^{(a1)}, x_9^{(a1)}$ 分别为东部地区 CPI, 东部地区工业品出厂价格指数, 东部地区商品零售价格指数, 东部社会消费品总额, 东部社会固定资产投资。

表 17 给出了东部地区城镇就业人数逐步回归分析表。选取的影响因素有: 东部社会消费品总额, 东部固定资产投资, 东部 CPI、M2、东部工业品出厂价格指数。

Model		UnStandardized Coefficient		Standardized Coefficient	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	.202603	.007536		26.884700	.000000
	东部社会消费品总额	.244621	.028196	.928703	8.675624	.000002
<2-4> 5	(Constant)	.105746	.028041		3.771108	.005457
	东部社会消费品总额	.161814	.056718	.614328	2.852941	.021382
	东部固定资产投资	.026476	.004588	.284199	5.771237	.000419
	东部CPI	.308086	.095246	.184414	3.234644	.011973
	M2	.079469	.064372	.285195	1.234514	.252041
	东部工业品出厂价格指数	.042327	.076054	.037872	.556546	.593060

表 17 东部地区城镇就业人数逐步回归分析表

表 18 给出了东部地区城镇就业人数逐步回归分析的总体统计参数。回归平方和为 0.002063, 残差平方和为 0.000011, 回归效果良好。F 检验值为 90.086718, 显著性水平为 0.000001, 方差分析结果表明拒绝回归系数为 0 的假设。

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	.002063	5.000000	.000413	90.086718	.000001(e)
Residual	.000037	8.000000	.000005		
Total	.002100	13.000000			

表 18 东部地区城镇就业人数逐步回归分析的总体统计参数

6.2.1.2 西部地区的城镇就业人数模型求解

选取西部地区的九个影响指标 ($x_1^{(a2)}, x_2^{(a2)}, \dots, x_9^{(a2)}$) 作为自变量, 西部地区的城镇就业人数 $y^{(a2)}$ 为因变量, 利用 SPSS 软件采取先筛选指标然后进行线性回归分析方法, 对参数 $b_0^{(a2)}, b_1^{(a2)}, \dots, b_9^{(a2)}$ 进行估计, 得到西部地区的城镇就业人数与各个影响指标的回归方程为:

$$y^{(a2)} = -0.28724 + 1.278262x_1^{(a2)} + 1.017879x_2^{(a2)} + 0.67503x_4^{(a2)} - 1.164742x_5^{(a2)} \\ - 0.065145x_6^{(a2)} + 0.32796x_7^{(a2)} + 0.021103x_8^{(a2)} + 0.021925x_9^{(a2)}$$

其中指标 $x_1^{(a2)}, x_2^{(a2)}, x_4^{(a2)}, x_5^{(a2)}, x_6^{(a2)}, x_7^{(a2)}, x_8^{(a2)}, x_9^{(a2)}$ 分别为西部地区人口总数, 西部地区CPI, 西部地区工业品出厂价格指数, 西部地区商品零售价格指数, 西部地区进出口总额, M2, 西部地区社会消费品总额, 西部地区社会固定资产投资。

表 19 给出了西部地区城镇就业人数逐步回归分析表。选取的影响因素有: 西部社会消费品总额, 西部商品零售价格指数, M2, 西部固定资产投资, 西部人口数, 西部进出口总额, 西部工业品出厂价格指数, 西部 CPI。

Model		UnStandardized Coefficient		Standardized Coefficient	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	.218555	.003921		55.740637	.000000
	西部社会消费品总额	.184893	.014671	.964236	12.602410	.000000
<2-7>略						
8	(Constant)	-.287240	.573955		-.500458	.637999
	西部社会消费品总额	.021103	.046313	.110055	.455660	.667740
	西部商品零售价格指数	-1.164742	2.168534	-.992595	-.537110	.614222
	M2	.327960	.108402	1.494963	3.025399	.029235
	西部固定资产投资	.021925	.006536	.308420	3.354398	.020231
	西部人口数	1.278262	1.973290	.090981	.647782	.545691
	西部进出口总额	-.065145	.019225	-.482901	-3.388633	.019490
	西部工业品出厂价格指数	.675030	.459704	.897615	1.468401	.201929
	西部CPI	1.017879	1.782711	.894371	.570973	.592722

表 19 西部地区城镇就业人数逐步回归分析表

表 20 给出了西部地区城镇就业人数逐步回归分析的总体统计参数。回归平方和为 0.00129, 残差平方和为 0.000011, 回归效果良好。

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	.001290	8.000000	.000161	71.692251	.000098(h)
Residual	.000011	5.000000	.000002		
Total	.001302	13.000000			

表 20 西部地区城镇就业人数逐步回归分析的总体统计参数

6.2.1.3 不同地区之间的城镇就业人数问题比较

通过 6.2.1.1 节东部地区城镇就业人数的回归分析和 6.2.1.2 节西部地区城镇就业人数的回归分析, 两个地区的就业问题比较归纳如下:

①社会消费品总额、固定资产投资、CPI、商品零售价格指数、工业品出厂价格指数这五个指标对东西部地区的城镇就业人数均有影响。

②工业品出厂价格指数、社会消费品总额、固定资产投资的增加能够增加东西部地

区的城镇就业人数。

③人口数、进出口总额、M2 对东部地区的城镇就业人数影响不大，而对西部地区的城镇就业人数影响较大，这与部分西部人群在东部就业相关、与西部经济不够发达存在有一定的影响。

④对于东部地区而言，商品零售价格指数能够增加东部地区的城镇就业人数；西部地区正好相反。这是由东西部经济水平的差异导致的。

6.2.2 基于不同行业的城镇就业人数模型求解及比较

由于行业的多种多样，为了方便计算，选取具有典型的建筑业和交通运输行业。通过中国统计年鉴[7]，采集了建筑业和交通运输业的指标季度数据，相关数据见附录六。

6.2.2.1 建筑业的城镇就业人数模型求解

针对建筑业的城镇就业人数的季度数据以及各相关指标的季度数据，利用 SPSS 软件 Regression 功能下的 Linear 模块，采取先筛选指标然后进行线性回归分析方法，对参数 $b_0^{(b1)}, b_1^{(b1)}, \dots, b_9^{(b1)}$ 进行估计，得到城镇就业人数与各个影响指标的回归方程为：

$$y^{(b1)} = -0.216064 + 1.093655x_2 - 1.096745x_3 \\ - 1.09346x_4 + 2.125905x_5 + 0.912042x_7 - 0.137392x_9$$

其中指标 $x_2, x_3, x_4, x_5, x_7, x_9$ 分别为 CPI，第三产业就业人数，工业品出厂价格指数，商品零售价格指数，M2，社会固定资产投资。

表 21 给出了建筑行业城镇就业人数逐步回归分析表。选取的影响因素有：社会固定资产投资，CPI，工业品出厂价格指数，商品零售价格指数，M2，第三产业就业人数。

Model		UnStandardized Coefficient		Standardized Coefficient	T	Sig.
		B	Std.Error	Beta		
1	(Constant)	.203950	.006164		33.086842	.000000
	社会固定资产投资	.327867	.021353	.979443	15.354288	.000000
2-7	<略>					
8	(Constant)	-.216064	.108091		-1.998904	.102082
	社会固定资产投资	-.137392	.153294	-.410433	-.896264	.411182
	CPI	1.093655	.165072	.159313	6.625314	.001180
	工业品出厂价格指数	-1.093460	.284806	-.370711	-3.839319	.012132
	商品零售价格指数	2.125905	.346918	.375946	6.127974	.001679
	M2	.912042	.234853	2.328479	3.883450	.011601
	第三产业就业人数	-1.096745	.383643	-.680135	-2.858765	.035458

表21 建筑行业城镇就业人数逐步回归分析表

表 22 给出了建筑行业城镇就业人数逐步回归分析主要统计数据。从表中可以看出，余量非常小，说明回归效果良好。

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.235332	.414329	.284048	.0537333	12
Residual	-.0037010	.0032597	.0000000	.0018765	12

Std. Predicted Value	-.907	2.425	.000	1.000	12
Std. Residual	-1.330	1.171	.000	.674	12

表22 建筑行业城镇就业人数逐步回归分析余量统计表

6.2.2.2 交通运输行业的城镇就业人数模型求解

针对交通运输行业的城镇就业人数的季度数据以及各相关指标的季度数据，利用 SPSS 软件 Regression 功能下的 Linear 模块，采取先筛选指标然后进行线性回归分析方法，对参数 $b_0^{(b2)}, b_1^{(b2)}, \dots, b_9^{(b2)}$ 进行估计，得到城镇就业人数与各个影响指标的回归方程为：

$$y^{(b2)} = 5.315713 - 12.763435x_1 - 1.186622x_2 - 5.999234x_3 - 1.309033x_4 \\ + 2.071316x_5 + 2.619048x_6 + 2.413897x_7 - 3.060083x_9$$

其中指标 $x_1, x_2, x_3, x_4, x_5, x_7, x_9$ 分别为全国总人数，CPI，第三产业就业人数，工业品出厂价格指数，商品零售价格指数，M2，社会固定资产投资。

表 23 给出了交通运输行业城镇就业人数逐步回归分析表。选取的影响因素有：商品零售价格指数，全国总人口，第三产业就业人数，CPI，社会固定资产投资，M2，进出口总额，工业品出厂价格指数。从最后一列的显著性水平可以看出，偏回归系数水平显著，即拒绝回归系数等于 0 的情况。

Model		UnStandardized Coefficient		Standardized Coefficient	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.512273	.219855		-2.330052	.042047
	商品零售价格指数	2.769206	.761594	.754556	3.636064	.004566
2-7	<略>					
8	(Constant)	5.315713	1.041073		5.105995	.014531
	商品零售价格指数	2.071316	.672803	.564394	3.078639	.054194
	全国总人口	-12.763435	3.143201	-2.602357	-4.060649	.026925
	第三产业就业人数	-5.999234	1.395853	-5.732442	-4.297899	.023170
	CPI	-1.186622	.465308	-.266342	-2.550188	.083929
	社会固定资产投资	-3.060083	.438806	-14.085404	-6.973651	.006051
	M2	2.413897	.466683	9.495781	5.172453	.014022
	进出口总额	2.619048	.521435	12.741900	5.022772	.015202
	工业品价格出厂指数	-1.309033	.646952	-.683813	-2.023384	.136211

表23 交通运输行业城镇就业人数逐步回归分析表

表 24 为交通运输行业城镇就业人数逐步回归分析主要统计数据。从表中可以看出，余量较小，回归效果良好。

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.249712	.360297	.286735	.034804	12.000000
Residual	-.005648	.004253	.000000	.002514	12.000000
Std. Predicted Value	-1.063769	2.113631	.000000	1.000000	12.000000
Std. Residual	-1.173278	.883525	.000000	.522233	12.000000

表24 交通运输行业城镇就业人数逐步回归分析余量统计表

6.2.2.3 不同行业之间的城镇就业人数问题比较

通过6.2.2.1节建筑行业的城镇就业人数的回归分析和6.2.2.2节交通运输行业的城镇就业人数的回归分析，比较两个行业的就业问题，有如下共同点：

- ①社会消费品总额、固定资产投资、CPI、商品零售价格指数、工业品出厂价格、M2 这六个指标对建筑行业和交通运输行业的城镇就业人数均有影响。
 - ②M2 的增加能够增加建筑行业和交通运输行业的城镇就业人数。因为广义货币量增加，拉动建筑行业、交通运输行业等的发展，故就业人数会有一定的增加。
- 两个行业的就业问题表现差异如下：全国人口数、进出口总额对建筑行业的城镇就业人数影响不大，而对交通运输行业的城镇就业人数影响较大。交通运输行业解决的是全国的交通问题，与人口数量息息相关。

6.2.3 基于不同学历的城镇就业人数模型求解及比较

6.2.3.1 基于不同学历的城镇就业人数模型求解

中国统计年鉴[7]上介绍了年度全国受教育程度的比例数据，假设受教育程度在全国为均匀分布，则：某年具有大专及其以上学历的城镇就业人数=某年的城镇就业人数*某年的大专及其以上学历比例。相关数据见附录七。

针对具有大专及其以上学历的城镇就业人数的季度数据以及各相关指标的季度数据，利用 SPSS 软件 Regression 功能下的 Linear 模块，采取先筛选指标然后进行线性回归分析方法，对参数 $b_0^{(e)}, b_1^{(e)}, \dots, b_9^{(e)}$ 进行估计，得到具有大专及其以上学历的城镇就业人数与各个影响指标的回归方程为

$$y^{(e)} = 0.198917 - 4.006973x_1 + 1.153569x_2 + 1.713901x_3 + 1.177451x_4 - 0.605864x_5 - 0.684547x_6 + 1.87889x_7 + 0.606276x_8 - 0.995208x_9$$

其中指标 x_1, \dots, x_9 分别为人口总数，CPI，第三产业就业人数，工业品出厂价格指数，商品零售价格指数，进出口总额，M2，社会消费品总额，社会固定资产投资。

如表 25 所示为具有大专及其以上学历的就业人数逐步回归分析情况。选取的影响因素有：第三产业就业人数，商品零售价格指数，CPI，全国总人口，工业品出厂价格指数，M2，进出口总额，社会固定资产投资，社会消费品总额。

Model		UnStandardized Coefficient		Standardized Coefficient	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.574208	.049451		-11.611702	.000000
	第三产业就业人数	2.950626	.171305	.983561	17.224381	.000000
2-8	略					
9	(Constant)	.198917	1.044418		.190457	.866532
	第三产业就业人数	1.713901	1.084964	.571312	1.579686	.254949
	商品零售价格指数	-.605864	.517243	-.057591	-1.171334	.362125
	CPI	1.153569	.448335	.090326	2.573007	.123649
	总人口数	-4.006973	2.905910	-.285009	-1.378905	.301888
	工业品出厂价格指数	1.177451	.520672	.214572	2.261408	.152142
	M2	1.878890	.396703	2.578439	4.736259	.041803
	进出口总额	-.684547	.479046	-1.161816	-1.428979	.289231

社会固定资产投资	-.995208	.431194	-1.598061	-2.308029	.147335
社会消费品总额	.606276	.482475	.630110	1.256594	.335779

表25 具有大专及其以上学历的城镇就业人数逐步回归分析表

表 26 给出了具有大专及其以上学历的城镇就业人数逐步回归分析的总体统计参数。回归平方和为 0.110031，残差平方和为 0.000025，回归效果良好。

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	.110031	9.000000	.012226	967.126787	.001033(i)
Residual	.000025	2.000000	.000013		
Total	.110056	11.000000			

表26 具有大专及其以上学历的城镇就业人数逐步回归分析的总体统计参数

6.2.3.2 基于不同学历的城镇就业人数问题比较

通过第 5 节的城镇就业人数的回归分析和 6.2.3.1 节具有大专及其以上学历的城镇就业人数的回归分析，分析大专及其以上学历对就业的影响如下：

①第三产业就业人数，商品零售价格指数，CPI，全国总人口，工业品出厂价格指数，M2，进出口总额，社会固定资产投资，社会消费品总额这九个指标对城镇就业均有影响，无论是对具有高学历的人还是没有高学历的人。

②商品零售价格指数的上升给城镇就业带来负面影响，无论是高学历人员就业或者非高学历人员的就业。由于商品零售价格指数的上升，选择就业的工资要求可能会所有提高，故给城镇就业带来一定的影响。

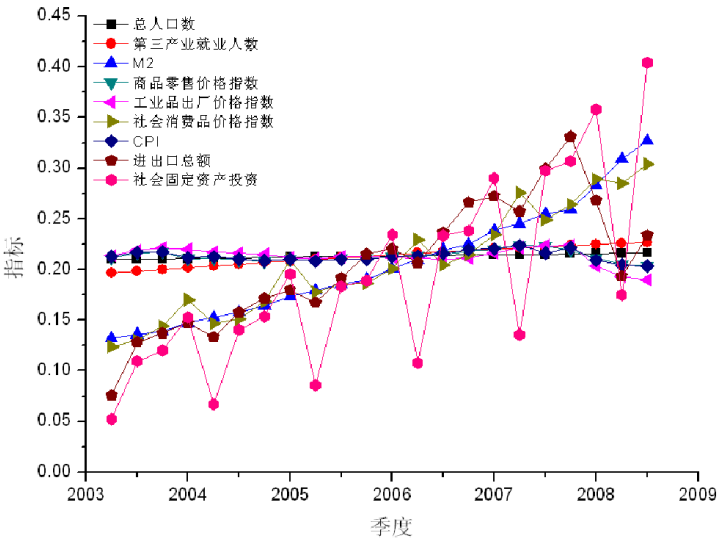
③全国总人数的增加能够给城镇就业人数带来积极的影响，却不能给具有高学历的人就业带来消极的影响。全国总人口的增加，而就业岗位数量不变，为了就业，不具备高学历的人就业要求门槛不高，而高学历的人对就业要求苛刻，因此给高学历的人带来消极的影响。

七、2010 年就业前景预测

在前面建立的数学模型的基础上对 2009 年及 2010 年上半年的就业前景进行预测。首先需要对就业的相关指标进行预测。

7.1 就业相关指标的数据预测

本小节以 2004 年第一季度至 2009 年第二季度相关指标的归一化数据为基础介绍预测的方法，1996 年至 2007 年的年度指标数据的预测方法同理可以计算。图 10 给出了 9 个相关指标随季度的变化图。由这个图形可以看出，各个指标随着时间是非线性变化的，这给指标值的准确预测带来了一定的困难，本文采用非线性回归预测法和灰色



模型预测法两种方法对指标值进行预测。

图 10 相关指标随季度的变化趋势

7.1.1 非线性回归预测

将 2004 年第一季度记为时间 1，2004 年第二季度记为时间 2，依次类推，2009 年第二季度记为时间 22，则每个指标值 x_i 都可以看成是时间 t 的连续函数 $f_i(c_1, \dots, c_m, t)$ 。从而可以建立下面的数学模型：

$$\begin{cases} x_i = f_i(c_1, \dots, c_m, t) + \varepsilon \\ E(\varepsilon) = 0 \end{cases}$$
$$s.t. \begin{cases} \min R_i(c_1, \dots, c_m) = \sqrt{\sum_{j=1}^n (x_{ij} - \hat{x}_{ij})^2} \\ \hat{x}_{ij} = f_i(c_1, \dots, c_m, t_j) \end{cases}$$

其中， c_1, \dots, c_m 为非线性方程中系数。

利用 SPSS 软件进行非线性回归分析，采用 7 种不同形式的函数（线性函数、对数函数、倒数函数、二次函数、三次函数、幂函数、指数函数）进行拟合，选择拟合精度最好的函数作为预测函数。表给出了用不同函数拟合各个指标时的 R^2 值，此值越接近 1 说明拟合效果越好。

	人口	CPI	三产就业	工业品出厂	商品零售	进出口总	M2	消费总额	社会投资
Linear	.99993	.0001	.9942	.212	.011	.733	.959	.927	.605
Log	.85963	.0008	.9005	.131	.007	.751	.736	.764	.502
Inverse	.47853	.0030	.5279	.040	.001	.515	.361	.419	.295
Quadratic	.99998	.0295	.9999	.290	.045	.799	.990	.936	.614
Cubic	1.00000	.6223	1.0000	.502	.647	.846	.995	.939	.620
Power	.86202	.0012	.9102	.135	.006	.856	.828	.840	.580
Exponential	.99997	.0004	.9911	.221	.009	.732	.992	.932	.579
选择函数	Cubic	Cubic	Cubic	Cubic	Cubic	Power	Cubic	Cubic	Cubic

表 26 不同形式函数的预测值

在获得 R 值的同时，可以得到相应的回归系数及其置信区间，如表 27 所示。

指标	预测方程式	季度	预测值	95%置信下限	95%置信上限
总人口数	C0=.210 C1=.0003 C2=-1.23E-006 C3=4.58E-008	2009-3	.21660	.21660	.21660
		2009-4	.21691	.21691	.21691
		2010-1	.21723	.21723	.21723
		2010-2	.21756	.21756	.21756
CPI	C0=.225 C1=-.006 C2=.001 C3=-2.04E-005	2009-3	.19670	.18600	.20740
		2009-4	.18876	.17586	.20166
		2010-1	.17924	.16335	.19514
		2010-2	.16803	.14833	.18773

第三产业就业	C0=.195	2009-3	.22760	.22760	.22760
	C1= .002	2009-4	.22840	.22840	.22840
	C2=1.64E-006	2010-1	.22913	.22913	.22913
	C3=-5.17E-007	2010-2	.22977	.22977	.22977
工业品出厂价格指数	C0=.228	2009-3	.18631	.16646	.20616
	C1= -.005	2009-4	.17703	.15310	.20096
	C2= .001	2010-1	.16614	.13665	.19563
	C3=-1.97E-005	2010-2	.15353	.11697	.19008
商品零售价格指数	C0=.224	2009-3	.19729	.18682	.20776
	C1= -.006	2009-4	.18924	.17662	.20186
	C2= .001	2010-1	.17958	.16403	.19514
	C3=-2.08E-005	2010-2	.16819	.14891	.18747
进出口总额	C0=.083	2009-3	.27757	.20741	.37147
	C1= .385	2009-4	.28215	.21067	.37788
		2010-1	.28662	.21385	.38415
		2010-2	.29097	.21693	.39028
M2	C0=.120	2009-3	.34205	.32876	.35534
	C1= .008	2009-4	.36382	.34780	.37984
	C2=-.0004	2010-1	.38752	.36778	.40727
	C3= 1.84E-005	2010-2	.41327	.38880	.43774
社会消费品总额	C0=.117	2009-3	.32128	.27775	.36481
	C1= .009	2009-4	.33768	.28520	.39017
	C2= -.0003	2010-1	.35541	.29074	.42009
	C3= 1.40E-005	2010-2	.37456	.29440	.45472
社会固定资产投资	C0=.061	2009-3	.36856	.17991	.55721
	C1= .017	2009-4	.39822	.17077	.62568
	C2= -.001	2010-1	.43098	.15069	.71127
	C3= 3.48E-005	2010-2	.46704	.11964	.81443

表 27 拟合的回归系数和置信区间

需要注意的是，本文中选取的非线性函数对 CPI、工业品出厂价格指数、商品零售价格指数和社会固定资产投资的回归效果比较差。为获得更好地预测结果，可以根据指标的变化趋势构造合适的非线性函数，也可以采用其他的方法进行预测，如时间序列分析预测及下面介绍的灰色模型预测等。

7.1.2 灰色模型预测

灰色预测是建立在 GM 模型基础上的预测，它通过对原始数据的处理和灰色模型的建立，发现并掌握系统的发展规律，对系统的未来状态做出科学的定量预测。灰色预测模型属于非线性拟外推类预测方法，实用性强，所需数据较少，建模灵活方便，预测精度高，从而在社会科学和自然科学各领域得到了广泛的应用。

灰色预测应用最为广泛的是 GM(1,1)模型。GM(1,1)模型是一阶单变量的微分方程动态模型，主要用于时间序列的预测。GM(1,1)模型通过对数列长度的不同取舍得到系列预测的结果，从而形成一个预测的灰色区间，即灰平面。但有时由于所得到的预测灰平

面成喇叭展开型，致使 GM(1,1)模型预测所得灰区间过大而失去意义。所以，用已知数据序列的 GM(1,1)模型进行预测时，并不是用一个模型一直预测下去，而是每个预测模型只得到一个值。同时，将这个预测值作为灰数重新补充到数据序列中。为了保持数据序列的维度不变，就要去掉第一个已知数据。这样一次反复，逐个进行预测数据的替补，不断补充最新的数据信息，使数据的灰度不断降低，直至达到一定的精度为止。所以 GM(1,1)模型有两个特点：首先是可以及时补充和利用新信息，从而提高灰区间的白化度。用补充后的新模型预测下一个值比继续使用原模型进行预测要更加科学，也更接近实际值；其次，每一次预测后都对参数做一次修正，预测值都产生在动态之中，因此使模型得到改进。

灰色系统模型 GM(1,1)预测的具体算法步骤如下：

Step1: 从所给信息得到一个等时间间隔的原始数据序列: $X^0 = \{x_1^0, x_2^0, \dots, x_m^0\}$;

Step2: 对原始的原始数据序列: X^0 进行一次累加，根据灰色系统理论，累加数据满

$$\text{足 } x_k^1 = \sum_{i=1}^k x_i^0 \quad (k=1,2,\dots,m), \text{ 由此计算累次和向量}$$

$$X^1 = \{x_1^1, x_2^1, \dots, x_m^1\};$$

Step3: 对序列 X^1 ，建立白化形式的微分方程: $\frac{dX^1}{dt} + aX^1 = b$.

其中参数 a 、 b 为微分方程的待定参数，并称参数 a 为发展系数，其大小反映的序列 X^0 的增长速度，参数 b 为灰色作用量（内生变量）。参数 a 、 b 可由最小二乘法求解。

$$\text{Step4: 记参数列为 } d = \begin{bmatrix} a \\ b \end{bmatrix}, \text{ 记 } B = \begin{bmatrix} -(x_1^1 + x_2^1)/2 & 1 \\ \vdots & \vdots \\ -(x_{m-1}^1 + x_m^1)/2 & 1 \end{bmatrix}, \quad Y_n = (x_2^0, x_3^0, \dots, x_n^0)^T$$

则 $d = (B^T B)^{-1} B^T Y_n$ ，从而 $a = d(1)$ 、 $b = d(2)$ 。

$$\text{由此可得白化形式的微分方程模型: } x_{t+1}^1 = \left(x_1^0 - \frac{b}{a}\right)e^{-at} + \frac{b}{a},$$

Step5: 经过累次减得到 $t+1$ 年的预测值: $x_{t+1}^0 = x_{t+1}^1 - x_t^1 = (1 - e^a) \left(x_1^0 - \frac{b}{a}\right)e^{-at}$

其中当 $1 \leq t \leq m$ 时， x_t^0 反映了原始数据列的变化趋势，当 $t > m$ 时， x_t^0 为预测值。

利用 Matlab 编写灰色系统 GM(1,1)预测程序（代码见附录八），预测得到相应的九个就业相关指标的近期预测数据为：

季度	总人口数	第三产业就业人数	M2	商品零售价格指数	工业品出厂价格指数	社会消费品总额	CPI	进出口总额	固定资产投资
09-3	0.2166	0.23	0.3244	0.2141	0.2055	0.3135	0.213	0.3013	.3507
09-4	0.2169	0.2315	0.3387	0.2148	0.2048	0.3259	0.219	0.3117	.3714
10-1	0.2172	0.2331	0.3536	0.2152	0.2041	0.3388	0.228	0.3226	.3932
10-2	0.2175	0.2347	0.3691	0.2153	0.2035	0.3522	0.230	0.3338	.4164

表 28 九个就业因素近期季度预测数据

灰色系统模型对各个指标的原始序列的预测的相对误差如表 29 所示。由表 29 可以看出，灰色模型对社会固定资产和进出口总额的预测较差，但相对于曲线拟合的而言，此预测获得了很好的效果。对进出口总额采用三次曲线拟合，其平均相对误差为 163.68，远大于灰色模型预测的值。

季度	总人口数	第三产业就业人数	M2	商品零售价格指数	工业品出厂价格指数	社会消费品总额	CPI	进出口总额	固定资产投资
相对误差	8.66e-005	0.003	0.018	0.0198	0.0231	0.049	0.02	0.1077	0.28

图 29 灰色系统模型对各个指标的原始序列的预测的相对误差

7.2 就业前景预测

基于季度数据的线性回归方程为：

$$y = -3.291899 + 18.410435x_1 + 0.00673x_2 - 1.370213x_3 - 0.019534x_4 + 0.052679x_5 - 0.05026x_6 - 0.025268x_7 - 0.00176x_8 + 0.000807x_9$$

基于年度数据的线性回归方程为：

$$y = -0.561198 + 2.925547x_1 + 0.017422x_2 - 0.02203x_3 - 0.0773x_4 - 0.021458x_5 + 0.03409x_6 + 0.071356x_7 + 0.043968x_8 - 0.020761x_9$$

将 5.1 得到的各个指标在 2009 年第三季度、第四季度和 2010 年第一季度、第二季度的预测值带入上面的回归方程，可以得到相应精度的就业人数的预测值，如表 30 所示。

	2009-3	2009-4	2010-1	2010-2
年度数据(利用非线性回归预测就业指标)	30557.25196	30770.70022	30984.14847	31197.59672
季度数据(利用非线性回归预测就业指标)	31319.49	31600.65	31882.89	32166.01
季度数据(利用灰色模型预测就业指标)	31126.97	31359.35	31589.26	31816.86

表 30 对 2009 年下半年和 2010 年上半年就业人数的预测值

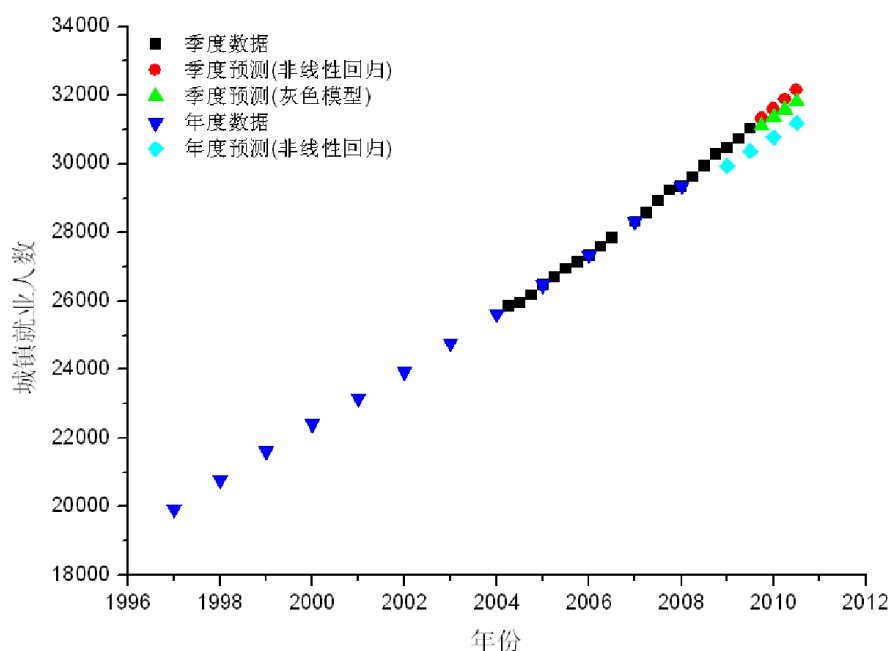


图 11 对 2009 年下半年和 2010 年上半年就业人数的预测图

	2009 年就业人数 预测值（万人）	2008 年就业人数（万人）	同比增长值（万人）
年度数据（非线性回归）	31070.70022	30463	607.70022
季度数据（非线性回归）	31600.65	30463	1137.65
季度数据（灰色模型）	31359.35	30463	896.35

表 31 不同方法对就业人数预测值的比较

根据政府工作报告：2009 年度要确定新增就业人数要保持在 900 万以上，结合表 31 可以看出：采用年度数据预测，其结果明显偏小，而采用季度数据进行预测，能够取得比较好的效果。这主要是由于 2008 年以来的主要经济运行指标能够从侧面刻画国家的宏观经济调控政策。

根据表 30 和 31，由于 2008 年我国城镇新增就业人数为 1137 万，按照选取绝对误差最小的原理，预测因素指标采用曲线拟合和基于季度数据的线性回归函数，预测 2009 年和 2010 上半年就业人数为 31600.65 万人和 32166.01 万人。

7.3 基于相关政策的就业预测

在 2010 年之前，中国将启动高达 4 万亿元的经济刺激方案。据新华网报道，5 日召开的国务院常务会议确定当前进一步扩大内需、促进经济增长的十项措施。初步匡算，到 2010 年底约需投资 4 万亿元。而今年四季度先增加安排中央投资 1000 亿元，明年灾后重建基金提前安排 200 亿元，带动投资总规模达 4000 亿元。自 2009 年的第三季度到 2010 年的第二季度，社会固定资产投资将会稳步大幅增长。

假设每个季度增加 10%，则自 2009 年的第三季度到 2010 年的第二季度的社会固定资产投资额度为：0.44451、0.488961、0.537857、0.591643，采用 7.2 中的关于季度的多元线性回归方程数学模型，相应的就业人数的预测值见表 32。由 2008 年新增城镇就业人数为 1113 万人，利用非线性回归的季度数据误差最小选取预测方法，在没有宏观调控的情况下，2009 年和 2010 年的预计值为 31600.65 万人和 32166.01 万人；如果考虑宏观调控，2009 年和 2010 年的预计值为 31606.9 万人和 32174.6 万人。

	2009-4 就业人数 预测值 (万人)	与 2008 相比 增长值 (万人)	2010-2 就业人数 预测值 (万人)	与 2008 相比 增长值 (万人)
季度数据 (非线性回归)	31600.65	1137.65	32166.01	1703.01
季度数据 (非线性回归+宏观 调控)	31606.9	1143.9	32174.6	1711.6
季度数据 (灰色模型)	31359.35	896.35	31816.86	1353.86
季度数据 (灰色拟合+宏观调 控)	31367.45	904.45	31828.93	1365.93

表 32 不同方法对相关政策下的就业人数预测值比较

灰色拟合+调控能够获得比较好的预测效果。灰色模型的预测值要小于曲线回归的预测值。曲线回归无法发掘出数据之间内在的信息,只是根据数据的整体趋势进行预测。

若假设每个季度增加 25%,则自 2009 年的第三季度到 2010 年的第二季度的社会固定资产投资额度为:0.505125、0.63140625、0.789257813、0.986572266,利用基于月度的多元线性回归方程,相应的就业人数的预测值见表 33。

由 2008 年新增城镇就业人数为 1113 万人,利用非线性回归的季度数据误差最小选取预测方法,在没有宏观调控的情况下,2009 年和 2010 年的预计值为 31600.65 万人和 32166.01 万人。如果考虑宏观调控,2009 年和 2010 年的预计值为 31616.72 万人和 32201.81 万人。

	2009-4 就业 人数预测值	与 2008 相比 增长值	2010-2 就业 人数预测值	与 2008 比 增长值
季度数据 (非线性回归)	31600.65	1137.65	32166.01	1703.01
季度数据 (曲线拟合+宏观调控)	31616.72	1153.72	32201.81	1738.81
季度数据 (灰色模型)	31359.35	896.35	31816.86	1353.86
季度数据 (灰色拟合+宏观调控)	31377.26	914.26	31856.15	1393.15

表 33 不同方法对相关政策下的就业人数预测值比较

六、关于提高城镇就业人口数的建议

中国是世界上人口最多的国家,也是劳动力资源最丰富的国家,同时也是人口就业压力最大的国家。丰富的劳动力资源与稀缺的资本资源,构成了中国社会内部一对主要矛盾。就业问题已经成为我国的民生大计。在短短的十多年间,我国的城镇就业人口从 1996 年 19922 万人的飙升到 2009 年第二季度的 31032 万人。在这可喜的成就下,我们应该看到我们目前尚未就业的城镇人口以及农业人口数量依旧庞大。如何提高城镇就业人数是国家的首要任务之一。

根据 1996 年至今的城镇就业人数与社会主要经济指标数据,通过分析不同地区、不同行业、不同学历就业人群三个角度的城镇就业人数有如下结论:

(1) 商品零售价指数的提高对全国城镇就业人数有负面影响,进出口总额、社会消费品总额、社会固定资产投资的增加能够提高城镇就业人数。

(2) 工业品出厂价格指数、社会消费品总额、固定资产投资的增加能够提高东西部地区的城镇就业人数。

(3) 人口数、进出口总额、M2 对东部地区的城镇就业人数影响不大,而对西部地

区的城镇就业人数影响较大，这与东西部就业人群重心的倾斜、西部经济不够发达存在有一定的影响。

(4) M2 的增加能够提高不同行业的城镇就业人数。

(5) 商品零售价格指数的上升给城镇就业带来负面影响，无论是高学历人员就业或者非高学历人员的就业。由于商品零售价格指数的上升，选择就业的工资要求可能会所有提高，故给城镇就业带来一定的影响。

(6) 全国总人数的增加能够给城镇就业人数带来积极的影响，却不能给具有高学历的人就业消极的影响。全国总人口的增加，而就业岗位数量不变，为了就业，不具备高学历的人就业要求门槛不高，而高学历的人对就业要求苛刻，因此给高学历的人就业带来消极的影响。

据此，提出对提高城镇就业人口的建议：

(1) 控制人口快速增长：中国人口基数大，每年净增人口多，给劳动总需求很大的压力，劳动力总供给与总需求出现严重的不均衡。随着人口的增加，交通运输行业的城镇就业人数增长率会有所降低。中国人口快速的的增长是失业的主要矛头之一。

(2) 适度提高经济发展速度

我国正面临前所未有的困难和挑战。一是国际金融危机在蔓延、仍未见底。国际市场需求继续萎缩，全球通货紧缩趋势明显，不确定因素显著增多。二是受国际金融危机影响，经济增速持续下滑，已成为影响全局的主要矛盾。适度的提高经济发展速度，势必增加对外贸易进出口总额以及社会消费品总额，这两个指标直接与提高城镇就业人口正相关。

(3) 加速西部大开发

由于自然、历史、社会等原因，西部地区经济发展相对落后，人均国内生产总值仅相当于全国平均水平的三分之二，不到东部地区平均水平的 40%。通过对东西部地区的城镇就业人数分析，工业品出厂价格指数、社会消费品总额、固定资产投资的增加都能够提高东西部的就业人数。商品零售价格指数能够增加东部地区的城镇就业人数；西部地区正好相反。这是由东西部经济水平的差异导致的。东部经济水平日趋饱和，而西部的经济容易受到外界影响的波动性很大。加速西部大开发，提高西部经济的发展，大力发展工业和第三产业，直接关系到扩大内需、促进经济增长，关系到民族团结、社会稳定和边防巩固，关系到东西部协调发展和最终实现共同富裕，具有重要的现实意义和深远的历史意义。

(4) 扩大内需

扩大内需，可以通过发行国债等积极财政货币政策，启动投资市场，通过信贷等经济杠杆，启动消费市场，使得居民消费指数 CPI 和社会消费品总额上升。社会消费品总额的上升的增加能够提高东西部地区的城镇就业人数。

(5) 增加社会固定资产投资

社会固定资产投资包括基本建设投资、更新改造投资、国有单位其他固定资产投资、房地产开发投资、城镇集体固定资产投资等方面。国家的固定资产投资，比如国家计划 2008 年四季度到 2010 年底全国投资 4 万亿元，根据季度数据的回归模型，社会固定资产投资的增加能够直接提高城镇就业人数。

(6) 实施适度宽松的货币政策

改善金融调控，保证货币信贷总量满足经济发展需求，09 年预计广义货币增长 17% 左右，新增贷款 5 万亿元以上。根据季度数据的回归模型，广义货币 M2 的增加能够直接提高城镇就业人数。

(7) 解决高校毕业生就业问题

近年来高校的不断扩招，毕业生人数骤增，2002年毕业生人数为145万人，2008年毕业生达到700万人。由于每年全国总人口的不断增加，而就业岗位增加数量有限，不具备高学历的人就业要求门槛不高，而高学历的人毕业生的就业期望值较高，理想与现实存在较大差距。表现为：毕业生对自己估计过高，眼高手低，这山望着那山高；对薪水、福利待遇要求过高，不顾自身的条件，导致用人单位不敢接收；片面追求大城市，不愿去一些小城市发展；对单位选择过高，“学而优则仕”非要去当官，看重国家机关、大单位等，对一些小企业不感兴趣，致使高不成低不就。对就业要求苛刻，因此给高校毕业生就业问题越来越严重。通过加强对大学生的就业指导，更新就业观念，增加城镇就业人口。

控制人口数量，促进经济增长，创造就业机会，减少失业率，是今后我国经济发展最优先的目标和任务。路漫漫其修远兮，任重道远，群策群力，明天的就业前景会更加美好！

参考文献

- [1] 高鸿业，吴易风，现代西方经济学讲座第三讲：凯恩斯模型，教学与研究，2：60-66，1986。
- [2] 张得志，中国经济高速增长过程中的劳动就业及其失业预警研究，复旦大学博士论文，17-22，2007年5月。
- [3] 王雪标，王晓婷，我国失业-利率敏感性分析，财经问题研究，262：21-25，2005。
- [4] 彭晓燕，钟学旗，我国目前经济高速增长与失业率上升并存原因探析，宏观经济管理，1:57-59，2009。
- [5] 中华人民共和国国家统计局，年度数据，<http://www.stats.gov.cn/tjsj/ndsj/>，2009-9-18。
- [6] 国研网，年度数据，<http://www.drcnet.com.cn/DRCNet.OLAP.Web/NewSelect/SelectMacro.aspx?version=Integrated>，2009-9-18。
- [7] 中华人民共和国国家统计局，季度数据，<http://www.stats.gov.cn/tjsj/jidusj/>，2009-9-18。
- [8] 国研网，季度数据，<http://www.drcnet.com.cn/DRCNet.OLAP.Web/NewSelect/SelectMacroQuarter.aspx?version=Integrated>，2009-9-18。
- [9] 国研网，月度数据，<http://www.drcnet.com.cn/DRCNet.OLAP.Web/NewSelect/SelectMacroQuarter.aspx?version=Integrated>，2009-9-18。
- [10] 和讯网，数据查询，<http://mac.hexun.com/Default.shtml?id=I>，2009-9-18。
- [11] 陈平雁，黄浙明，SPSS 10.0统计软件应用教程，人民军医出版社，2002年。

附录一：1996–2007 年就业及相关指标年份数据表

年份	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}
1996	122389	108.3	17927	102.9	106.1	24133.8	76094.9	28360.2	22913.5	34820	16203	71176.6	14015.4	33835	23326.2	5845.9	5980	7937.6
1997	123626	102.8	18432	99.7	100.8	26967.2	90995.3	31252.9	24941.1	34840	16547	78973.0	14441.9	37543	26988.1	6420.2	6444	9233.6
1998	124761	99.2	18860	95.9	97.4	26849.7	104498.5	33378.1	28406.2	35177	16600	84402.3	14817.6	39004.2	30580.5	6796.0	7446	10798.2
1999	125786	98.6	19205	97.6	97.0	29896.2	119897.9	35647.9	29854.7	35768	16421	89677.1	14770.0	41033.6	33873.4	7158.5	8319	13187.7
2000	126743	100.4	19823	102.8	98.5	39273.2	134610.4	39105.7	32917.7	36043	16219	99214.6	14944.7	45555.9	38714.0	7857.7	9333	15886.5
2001	127627	100.7	20228	98.7	99.2	42183.6	158301.9	43055.4	37213.5	36513	16284	109655.2	15781.3	49512.3	44361.6	8621.7	10834	18902.6
2002	128453	99.2	21090	97.8	98.7	51378.2	185007.0	48135.9	43499.9	36870	15780	120332.7	16537.0	53896.8	49898.9	9398.1	12373	22053.2
2003	129227	101.2	21809	102.3	99.9	70483.5	221222.8	52516.3	55566.6	36546	16077	135822.8	17381.7	62436.3	56004.7	10542.0	13969	24650.0
2004	129988	103.9	23011	106.1	102.8	95539.1	254107.0	59501.0	70477.4	35269	16920	159878.3	21412.7	73904.3	64561.3	12335.6	15920	28486.9
2005	130756	101.8	23771	104.9	100.8	116921.8	298755.7	67176.6	88773.6	33970	18084	183217.4	22420.0	87364.6	73432.9	14053.0	18200	33930.3
2006	131448	101.5	24614	103.0	101.0	140971.4	345603.6	76410.0	109998.2	32561	19225	211923.5	24040.0	103162	84721.4	16165.0	20856	40422.7
2007	132129	104.8	24917	103.1	103.8	166740.2	403442.2	89210.0	137323.9	31444	20629	249529.9	28095.0	121381.3	100053.5	18934.0	24721	49781.4

注：

 x_1 : 人口总数 (万人) x_2 : 居民消费价格指数(上年=100) x_3 : 第三产业就业人数 (万人) x_4 : 工业品出厂价格指数(上年=100) x_5 : 商品零售价格指数(上年=100) x_6 : 进出口总额 (亿元人民币) x_7 : M2 货币供应量 (亿元) x_8 : 社会消费品零售总额 (亿元) x_9 : 社会固定资产投资 (亿元) x_{10} : 第一产业就业人数 (万人) x_{11} : 第二产业就业人数 (万人) x_{12} : GDP 国内生产总值 (亿元) x_{13} : 第一产业 GDP (亿元) x_{14} : 第二产业 GDP (亿元) x_{15} : 第三产业 GDP (亿元) x_{16} : 人均 GDP (元) x_{17} : 就业人员平均工资 (元) x_{18} : 财政支出 (亿元)

附录二：2004年第1季度至2009年2季度城镇就业及相关指标季度数据表

时间	城镇就业人数	总人口数	第三产业就业人数	M2 货币供应量	商品零售价格指数	工业品出厂价格指数	社会消费品零售总额	居民消费价格指数	进出口总额	社会固定资产投资
2004 年 1 季度	26681.03	130180.5	23026.9	231654.6	101.7	103.9	4049.8	103.0	1682.1	7058.50
2004 年 2 季度	26903.02	130362.5	23229.3	238127	104.1	106.4	4250.7	105.0	2837.4	14785.50
2004 年 3 季度	27133.54	130543.4	23430.7	243740.3	104.3	107.9	4717.7	105.2	3046.9	16184.30
2004 年 4 季度	27364.07	130723.2	23630.5	257752.8	101.3	107.1	5562.5	102.4	3263.7	20592.00
2005 年 1 季度	27594.59	130902.2	23828.5	266992.7	101.4	105.6	4799.1	102.7	2952.2	9036.70
2005 年 2 季度	27833.65	131080.6	24024.3	276952.1	100.5	105.2	4935.0	101.6	3498.2	18930.30
2005 年 3 季度	28072.71	131258.4	24217.6	287591.6	100.0	104.5	5495.2	100.9	3796.4	20774.50
2005 年 4 季度	28320.31	131435.9	24407.9	303538.6	100.9	103.2	6850.4	101.6	3978.9	26355.00
2006 年 1 季度	28567.91	131613.3	24594.9	313701.8	100.2	102.5	5796.7	100.8	3711.9	11608.40
2006 年 2 季度	28824.05	131790.7	24778.3	324010.8	101.3	103.5	6057.8	101.5	4247.8	24760.00
2006 年 3 季度	29080.19	131968.2	24957.6	332747.2	101.2	103.5	6077.4	101.5	4767.6	25511.72
2006 年 4 季度	29344.87	132146.1	25132.6	351498.8	102.4	103.1	6553.6	102.8	4884.0	31592.24
2007 年 1 季度	29609.54	132324.6	25302.8	367425.6	102.5	102.7	7499.2	103.3	4577.7	14543.61
2007 年 2 季度	29882.76	132503.7	25468.0	383884.9	103.2	102.5	6685.8	104.4	5233.4	31534.21
2007 年 3 季度	30164.51	132683.7	25627.6	394204.2	104.9	102.7	7026.0	106.2	5898.1	32168.96
2007 年 4 季度	30446.26	132864.7	25781.5	417846.2	105.6	105.4	7668.4	106.5	6035.0	39167.13
2008 年 1 季度	29628.00	133047.0	25929.1	429240.9	107.8	108.0	9015.3	108.3	5705.1	18316.94
2008 年 2 季度	29960.00	133230.6	26070.2	446362.2	107.1	108.8	8123.2	104.6	6630.2	40119.04
2008 年 3 季度	30286.00	133415.7	26204.3	453133.3	105.3	109.1	8642.0	107.1	7331.9	41434.73
2008 年 4 季度	30463.00	133602.6	26331.2	496112.3	101.4	98.9	9446.5	101.2	5940.4	48296.54
2009 年 1 季度	30731.00	133791.4	26450.4	540481.2	98.5	94.0	9317.6	98.8	4287.9	23562.00
2009 年 2 季度	31032.00	133982.2	26561.6	572997.8	97.7	92.2	9941.6	98.3	5177.1	54536.00

单位： 就业人数：万人
M2：亿元

城镇就业人数：万人
社会消费品零售总额：亿元

总人口数：万人
进出口总额：亿美元

第三产业就业人数：万人
社会固定资产投资：亿元

附录三：归一化的 1996-2007 年就业及相关指标年份数据表

年份	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}
1996	0.2765	0.3068	0.2433	0.2898	0.3011	0.0831	0.0981	0.1527	0.0985	0.2870	0.2729	0.1430	0.2158	0.1435	0.1178	0.1521	0.1224	0.0876
1997	0.2793	0.2912	0.2501	0.2889	0.3035	0.0929	0.1173	0.1683	0.1072	0.2872	0.2787	0.1587	0.2223	0.1593	0.1363	0.1670	0.1319	0.1019
1998	0.2819	0.2810	0.2560	0.2771	0.2956	0.0925	0.1347	0.1798	0.1221	0.2900	0.2796	0.1696	0.2281	0.1655	0.1545	0.1768	0.1525	0.1191
1999	0.2842	0.2793	0.2606	0.2704	0.2868	0.1030	0.1546	0.1920	0.1284	0.2948	0.2766	0.1802	0.2274	0.1741	0.1711	0.1862	0.1703	0.1455
2000	0.2863	0.2844	0.2690	0.2780	0.2825	0.1352	0.1735	0.2106	0.1415	0.2971	0.2731	0.1994	0.2301	0.1933	0.1955	0.2044	0.1911	0.1753
2001	0.2883	0.2853	0.2745	0.2744	0.2802	0.1453	0.2041	0.2319	0.1600	0.3010	0.2742	0.2204	0.2430	0.2101	0.2241	0.2243	0.2218	0.2085
2002	0.2902	0.2810	0.2862	0.2684	0.2766	0.1769	0.2385	0.2593	0.1870	0.3039	0.2658	0.2418	0.2546	0.2287	0.2520	0.2445	0.2533	0.2433
2003	0.2919	0.2867	0.2960	0.2745	0.2763	0.2427	0.2852	0.2828	0.2389	0.3013	0.2708	0.2730	0.2676	0.2649	0.2829	0.2742	0.2860	0.2719
2004	0.2937	0.2943	0.3123	0.2913	0.2840	0.3290	0.3276	0.3205	0.3030	0.2907	0.2850	0.3213	0.3297	0.3135	0.3261	0.3209	0.3260	0.3143
2005	0.2954	0.2884	0.3226	0.3056	0.2864	0.4026	0.3851	0.3618	0.3817	0.2800	0.3046	0.3682	0.3452	0.3706	0.3709	0.3656	0.3726	0.3743
2006	0.2970	0.2875	0.3340	0.3148	0.2892	0.4855	0.4455	0.4115	0.4729	0.2684	0.3238	0.4259	0.3701	0.4377	0.4279	0.4205	0.4270	0.4459
2007	0.2985	0.2969	0.3382	0.3245	0.3002	0.5742	0.5201	0.4805	0.5904	0.2592	0.3474	0.5015	0.4325	0.5150	0.5054	0.4925	0.5062	0.5492

注：

 x_1 : 人口总数 x_5 : 商品零售价格指数 x_9 : 社会固定资产投资 x_{13} : 第一产业 GDP x_{17} : 就业人员平均工资 x_2 : 居民消费价格指数 x_6 : 进出口总额 x_{10} : 第一产业就业人数 x_{14} : 第二产业 GDP x_{18} : 财政支出 x_3 : 第三产业就业人数 x_7 : M2 货币供应量 x_{11} : 第二产业就业人数 x_{15} : 第三产业 GDP x_4 : 工业品出厂价格指数 x_8 : 社会消费品零售总额 x_{12} : GDP 国内生产总值 x_{16} : 人均 GDP

附录四：归一化的 2004 年第 1 季度至 2009 年 2 季度城镇就业及相关指标季度数据表

时间	城镇就业人数	总人口数	第三产业就业人数	M2 货币供应量	商品零售价格指数	工业品出厂价格指数	社会消费品零售总额	CPI	进出口总额	社会固定资产投资
2004 年 1 季度	0.3125	0.2101	0.1966	0.1323	0.2116	0.2135	0.1239	0.2130	0.0759	0.0523
2004 年 2 季度	0.3151	0.2104	0.1983	0.1360	0.2166	0.2186	0.1301	0.2171	0.1281	0.1095
2004 年 3 季度	0.3178	0.2107	0.2000	0.1392	0.2170	0.2217	0.1443	0.2175	0.1375	0.1199
2004 年 4 季度	0.3205	0.2110	0.2017	0.1472	0.2108	0.2201	0.1702	0.2117	0.1473	0.1526
2005 年 1 季度	0.3232	0.2113	0.2034	0.1525	0.2110	0.2170	0.1468	0.2124	0.1333	0.0670
2005 年 2 季度	0.3260	0.2116	0.2051	0.1582	0.2091	0.2162	0.1510	0.2101	0.1579	0.1403
2005 年 3 季度	0.3288	0.2119	0.2067	0.1643	0.2081	0.2147	0.1681	0.2086	0.1714	0.1539
2005 年 4 季度	0.3317	0.2122	0.2083	0.1734	0.2100	0.2121	0.2096	0.2101	0.1796	0.1953
2006 年 1 季度	0.3346	0.2125	0.2099	0.1792	0.2085	0.2106	0.1774	0.2084	0.1676	0.0860
2006 年 2 季度	0.3376	0.2127	0.2115	0.1851	0.2108	0.2127	0.1853	0.2099	0.1918	0.1835
2006 年 3 季度	0.3406	0.2130	0.2130	0.1901	0.2106	0.2127	0.1859	0.2099	0.2152	0.1890
2006 年 4 季度	0.3437	0.2133	0.2145	0.2008	0.2131	0.2119	0.2005	0.2126	0.2205	0.2341
2007 年 1 季度	0.3468	0.2136	0.2160	0.2099	0.2133	0.2110	0.2294	0.2136	0.2066	0.1078
2007 年 2 季度	0.3500	0.2139	0.2174	0.2193	0.2148	0.2106	0.2046	0.2159	0.2362	0.2336
2007 年 3 季度	0.3533	0.2142	0.2188	0.2252	0.2183	0.2110	0.2150	0.2196	0.2663	0.2383
2007 年 4 季度	0.3566	0.2145	0.2201	0.2387	0.2197	0.2166	0.2346	0.2202	0.2724	0.2902
2008 年 1 季度	0.3600	0.2148	0.2213	0.2452	0.2243	0.2219	0.2758	0.2239	0.2575	0.1357
2008 年 2 季度	0.3635	0.2151	0.2225	0.2550	0.2229	0.2236	0.2485	0.2163	0.2993	0.2973
2008 年 3 季度	0.3670	0.2154	0.2237	0.2589	0.2191	0.2242	0.2644	0.2215	0.3310	0.3070
2008 年 4 季度	0.3706	0.2157	0.2248	0.2834	0.2110	0.2032	0.2890	0.2093	0.2682	0.3578
2009 年 1 季度	0.3743	0.2160	0.2258	0.3087	0.2050	0.1932	0.2851	0.2043	0.1936	0.1746
2009 年 2 季度	0.3780	0.2163	0.2267	0.3273	0.2033	0.1895	0.3042	0.2033	0.2337	0.4041

附录五：东西部各主要指标数据表

时间	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
2006 年 1 季度	3189.3	6055.8	6290.3	5718.0	101.35	100.53	100.65	99.61	358.3	1013.5	12.98	99.61	1218.6	436.8	103.24	106.17
2006 年 2 季度	3192.0	6135.8	6305.2	5721.2	102.15	101.16	101.40	100.98	962.3	2586.0	14.79	100.98	1153.7	407.6	102.79	105.68
2006 年 3 季度	3231.7	6185.0	6320.9	5723.1	102.29	100.99	101.41	100.98	1528.3	4224.0	18.30	100.98	1178.3	403.7	102.03	105.13
2006 年 4 季度	3301.0	6538.0	6337.7	5723.4	102.92	102.68	101.95	102.35	2300.3	6086.8	17.04	102.35	1333.6	518.4	100.78	103.90
2007 年 1 季度	3260.3	6366.0	6355.5	5722.1	103.73	103.23	102.52	102.55	436.0	1223.8	15.33	102.55	1417.5	472.3	100.76	103.49
2007 年 2 季度	3281.7	6440.8	6374.6	5719.2	104.76	104.16	103.26	103.10	1215.5	3167.9	20.97	103.10	1352.5	472.4	101.38	103.34
2007 年 3 季度	3301.3	6495.0	6394.8	5714.4	108.68	105.90	106.86	104.53	1943.6	5114.0	22.51	104.53	1395.4	509.3	94.78	104.14
2007 年 4 季度	3507.7	6811.3	6416.5	5707.8	108.48	105.75	107.15	104.98	2925.1	7286.0	16.48	104.98	1581.3	623.3	103.16	105.94
2008 年 1 季度	3418.0	6690.5	6439.5	5699.2	109.88	107.40	109.10	107.34	545.5	1484.2	23.71	107.34	1714.8	566.5	105.41	109.49
2008 年 2 季度	3425.3	6812.3	6464.1	5688.5	107.66	106.27	107.72	106.80	1501.3	3806.5	30.94	106.80	1649.7	565.1	106.67	111.64
2008 年 3 季度	3462.0	6837.5	6490.2	5675.7	103.59	104.17	104.35	105.22	2450.5	6251.5	32.88	105.22	1711.2	624.1	107.33	111.91
2008 年 4 季度	3551.3	7022.3	6337.7	5723.4	101.83	100.70	102.20	101.23	3683.3	8751.2	28.32	101.23	1901.4	755.7	100.77	103.72
2009 年 1 季度	3480.3	6876.0	6355.5	5722.1	99.23	98.63	98.90	98.33	820.0	1758.3	22.22	98.33	1973.0	670.0	94.95	96.68
2009 年 2 季度	3499.0	6904.8	6374.6	5719.2	99.52	97.99	98.58	97.45	2171.3	4609.7	26.07	97.45	1916.5	686.4	93.18	92.20

注：

 x_1 : 西部就业人数 x_2 : 东部就业人数 x_3 : 西部人口数 x_4 : 东部人口数 x_5 : 西部 CPI x_6 : 东部 CPI x_7 : 西部商品零售价格指数 x_8 : 东部商品零售价格指数 x_9 : 西部固定资产投资 x_{10} : 东部固定资产投资 x_{11} : 西部进出口总额 x_{12} : 东部进出口总额 x_{13} : 东部社会消费品总额 x_{14} : 西部社会消费品总额 x_{15} : 东部工业品出厂价格指数 x_{16} : 西部工业品出厂价格指数

附录六：分行业就业人数表

年份	建筑业就业人数	交通运输业就业人数
1996	21218700	6740000
1997	21015100	6664000
1998	20299900	5485000
1999	20201300	5276000
2000	19943000	5094000
2001	21106600	4824192
2002	22451900	4655030
2003	24142700	5289049
2004	25578600	5202915
2005	26999249	5088998
2006	28781591	5084348
2007	71337116	5128760

单位：人

附录七：分行业就业人数表

年份	大专及以上学历
1996	677.348
1997	727.335
1998	821.408
1999	941.304
2000	1041.795
2001	1149.12
2002	1288.56
2003	1435.784
2004	1562.084
2005	1694.522
2006	1840.15
2007	1995.8

单位：万人

附录八：灰色系统模型 $GM(1,1)$ 的程序：

程序语言：MATLAB

软件类型：MATLAB 7.0.1

具体程序：

```
function [a,b,X,W]=GM(x,n,t)

% 函数功能：用灰色系统模型GM(1,1)进行预测
% 输入： x—— 原始数据序列，
%        n—— x中元素的个数
%        t—— 超前预测的数据的个数
% 输出： a,b—— GM(1,1)模型中的参数
%        X—— 前n个为原始数据预测值，剩余的为超前预测值。
%        W—— 误差向量

for i=1:n
    a=0;
    for j=1:i
        a=a+x(j);
    end
    xx(i)=a;
end
for i=1:n-1
    b(i,1)=(xx(i)+xx(i+1))/2*(-1);
    b(i,2)=1;
end
for i=2:n
    y(i-1)=x(i);
end
d=inv(b'*b)*b'*y'
a=d(1);
b=d(2);
X(1)=x(1);
for i=2:n+t
    X(i)=(1-exp(a))*(X(1)-b/a)*exp(-1*a*(i-1));
end
for i=1:n
    W(i)=x(i)-X(i);
end
```