

参赛密码 _____

(由组委会填写)

全国第八届研究生数学建模竞赛



题 目

房地产行业的数学建模

摘

要：

本文针对房地产问题，运用多元线性与非线性回归、因子分析、灰度预测以及层次分析等方法，首先讨论了全国房地产需求、供给和房价这三指标与其各自主要影响因素之间的函数关系，然后对全国房地产行业与国民经济其他行业的关系模型进行分析并对我国房地产行业的态势的走向进行仿真，最后以天津市为例，对其房地产行业的可持续发展指标进行评估并给出结论和建议。

针对问题一 (①住房的需求、②供给与③价格)，一方面，通过相关性分析以及主成分分析，得出影响房地产需求的主要因素(①人均 GDP；②年总销售房价值；③城镇年人均可支配收入；④年储蓄存储值；⑤城镇人口；⑥城镇居民人均建筑面积；⑦城镇就业人数)与供给的主要因素(①房地产企业本年经营总收入；②商品房本年施工面积；③商品房本年新开工面积；④人口自然增长率；⑤房地产业增加值；⑥房地产企业本年土地转让收入；⑦房地产企业本年资产负债率；⑧商品房本年销售价格；⑨人均国内生产总值)。使用多元非线性回归法建立影响房地产的供需模型；另一方面，利用灰度预测和灰色关联度分析，建立房地产价格灰度模型。得出未来 5 年内商品房供需面积以及价格均呈现稳步上升的趋势。

针对问题二 (④房地产行业与国民经济其他行业关系、⑤房地产行业态势分析)，一方面，利用多元线性回归方法，采用 Matlab 软件编程，通过建立影响全国房地产行业增加值指数的三个单元线性回归模型，得出房地产业发展与国民经济总值，交通运输、仓储、邮政业，批发、零售业发展的相关系数依次为 0.9929, 0.9794, 0.9847，接近于 1，均呈现了高度的相关性。并据此建立房

地产行业增加值指数和其影响因素的多元线性回归模型，得出房地产投资每增加 1%，GDP 增加 0.5374%，交通运输、仓储、邮政业增加值指数减小 2.252%，批发和零售业增加值指数增加 4.37%。另一方面，利用投机价格衡量法，建立房地产行业发展稳定度模型，采用 SPSS 软件进行数据分析，计算出房地产行业发稳定度均值为 0.76，说明全国房地产行业的发展稳定度临介于警戒和安全区域之间，需要对房地产行业可持续发展进行规划。

针对问题三 (⑥房地产行业可持续发展)，由问题 2 中全国房地产行业态势，以天津市为例，从经济、人口、环境、资源这四个因素，通过因子分析法和层次分析法，建立天津市房地产行业可持续发展模型，然后利用 SPSS 和 Matlab 软件进行仿真，计算出房地产行业可持续发展指标。总体来看，天津市的可持续发展能力良好，2008 年达到历史最高值，说明天津市房地产行业正处于一个历史的最佳时期，全国其他城市可以借鉴天津市房地产业发展模型，努力维持该行业的平稳发展。

关键词：回归分析；灰度预测；层次分析；因子分析；投机价格

1. 问题的提出与重述

1.1 问题提出

近几年来，中央出台了一系列房地产调控政策，经济手段和行政手段并用，从抑制需求、增加供给、加强监管等方面对中国房地产市场进行了全方位的调控。为了积极响应国家的号召，并针对房地产行业的实际情况，提出了利用数学建模的方法量化分析并解决房地产行业的一系列实际应用问题。

1.2 问题重述

房地产行业既是国民经济的支柱产业之一，又是与人民生活密切相关的行业之一，同时自身也是一个庞大的系统，该系统的状态和发展对国民经济的整个态势和全国人民的生活水平影响很大。近年来，我国房地产业发展迅速，不仅为整个国民经济的发展做出了贡献，而且为改善我国百姓居住条件发挥了决定性作用。但同时房地产业也面临较为严峻的问题和挑战，引起诸多争议，各方都坚持自己的观点，然而多是从政策层面、心理层面和资金层面等因素来考虑，定性分析多于定量分析。显然从系统的高度认清当前房地产行业的态势、从定量角度把握各指标之间的数量关系、依据较为准确的预见对房地产行业进行有效地调控、深刻认识房地产行业的经济规律进而实现可持续发展是解决问题的有效途径。因此通过建立数学模型研究我国房地产问题是一个值得探索的方向。

请你们利用附录中提供的及可以查找到的资料建立房地产行业的数学模型，建议包括

1. 住房需求模型；
2. 住房供给模型；
3. 房地产行业与国民经济其他行业关系模型；
4. 对我国房地产行业态势分析模型；
5. 房地产行业可持续发展模型；
6. 房价模型等。

并利用模型进行分析，量化研究该行业当前的态势、未来的趋势，模拟房地产行业经济调控策略的成效。希望在深化认识上取得进步，产生若干结论和观点。如果仅就其中几个问题建立模型也是适宜的，对利用附件给的天津市的数据建模并进行分析同样鼓励。由于对房地产问题已经有许多研究成果和讨论材料，引用其他人的成果和数据，尤其对于定量分析的成果，务必注明参考文献，提请研究生特别注意。

研究房地产问题并不需要很多、很深的专业知识，问题也不难理解。你们完全可以独立自主地提出自己希望解决的房地产中的新问题，建立相应的数学模型予以解决，所建的每个模型要系统、深入，至少应该自成兼容系统，数据可靠，结论和观点有较多的数据支撑、有较强的说服力、有实际应用价值。

2. 附录表格中部分数据添加的说明

通过对国家房地产政策的研究以及房地产行业的实际情况，本文对附录中的全国年度库的相关表格进行了科学性处理，补全了表格中的有关缺项。关于添加的数据现做如下说明：

- 经济适用房本年销售面积表中 1991-1996 年的销售面积都记为 0，依据是国家有关经济适用房的相关政策的出台，1997 年之后才有经济适用房的出现。
- 经济适用房本年销售面积表中 2010 年的销售面积记为 2660.89，依据是国家统计局给出的相关数据加以科学计算。
- 经济适用房商品房本年销售价格表中 1991-1996 年的销售价格都记为 0，依据是国家有关经济适用房的相关政策的出台，1997 年之后才有经济适用房的出现。
- 经济适用房商品房本年销售价格表中 2010 年的销售价格记为 2364，依据是国家统计局给出的相关数据加以科学计算。
- 商品房本年销售价格表中 2010 年的销售价格记为 5032，依据是国家统计局给出的相关数据加以科学计算。
- 住宅商品房本年销售价格表中 2010 年的销售价格记为 4800，，依据是国家统计局给出的相关数据加以科学计算。
- 储蓄存款表中 1991、1992、2010 年的存款分别记为 9244.9、11757.3、303302，依据的是国家统计局的权威统计数据。
- 城镇居民人居建筑面积表中 2007-2010 的人居建筑面积分别记为 28、28.8、30、30.9，依据是国家统计局给出的相关数据加以科学计算。
- 城镇年末从业人员数表中 2010 从业人数记为 32200，依据是国家统计局的权威统计数据。

3. 问题分析

3.1 背景分析

住房是居民的基本生活需求，在全面建设小康社会阶段，随着经济社会的发展和人民生活水平的提高，城镇住房的增量需求和改善需求日益旺盛，是房地产业持续发展的动力。供不应求将是未来几十年中国房地产市场的主导趋势。

房地产行业是指以土地和建筑物为经营对象，从事房地产开发、建设、经营、管理以及维修、装饰和服务的集多种经济活动为一体的多家企业的集合。近几年来，我国房地产业发展迅速，不仅为整个国民经济的发展做出了贡献，而且为改善我国百姓居住条件发挥了决定性作用。但同时房地产业也面临较为严峻的问题和挑战，引起诸多争议，各方都坚持自己的观点，然而多是从政策层面、心理层面和资金层面等因素来考虑，定性分析多于定量分析。为保持经济健康稳定的发展，中央政府综合运用经济、法律和必要的行政手段，以区别对待和循序渐进的方式，对房地产业连续出台了一系列宏观调控政策。但房地产市场仍然存在住房供给结构不合理、部分城市房价上涨太快、中低收入居民住房难以满足等问题。因此，为房地产行业建立数学模型，并利用这些模型进行分析，量化研究该行业当前的态势、未来的趋势，模拟房地产行业经济调控策略势在必行。

3.2 具体问题分析

本题要求我们建立一个房地产行业的数学模型，其中包括住房需求模型、住房供给模型、房地产行业与国民经济其他行业关系模型、对我国房地行业态势分析模型、房地产行业可持续发展模型和房价模型等一系列模型。利用以上模型进行分析，量化研究该房地产行业当前的态势、未来的趋势，模拟房地产行业经济调控策略的成效。在深化认识该行业态势的基础上，得出影响房地产行业的相关因素，解决房地产行业的问题有很大的实际应用价值。

对于问题一 (①住房的需求、②供给与③价格)，需要建立住房需求模型。本文依托国家统计局数据库提供的全国 1991 年至 2010 年房地产市场的一系列原始数据，应建立模型的需要，将数据进行了一定的技术处理，统一口径，数据库覆盖了 7 个宏观指标。这 7 个宏观指标分别为：①人均 GDP；②年总销售房价值；③城镇年人均可支配收入；④年储蓄存储值；⑤城镇人口；⑥城镇居民人均建筑面积；⑦城镇就业人数。利用 Person、Kendall 和 Spearman 方法进行相关性分析，判断出以上 7 个指标均与住房需求有较大的相关性，继而用主成分分析法消除指标间的共线性，取累计贡献率为 96.43% 的第一主成分，再用回归拟合房地产的年总销售面积并建立走势预测模型。

然后建立住房供给模型。引用需求的数学模型，并依托国家统计局数据库提供的全国 1991 年至 2010 年房地产市场的一系列原始数据，选取 9 个宏观指标为：①房地产企业本年经营总收入；②商品房本年施工面积；③商品房本年新开工面积；④人口自然增长率；⑤房地产业增加值；⑥房地产企业本年土地转让收入；⑦房地产企业本年资产负债率；⑧商品房本年销售价格；⑨人均国内生产总值。利用 Person、Kendall 和 Spearman 方法进行相关性分析，判断出以上 9 个指标均与住房供给有较大的相关性，继而用主成分分析法消除指标间的共线性，取累计贡献率为 91.76% 的第一主成分，再用回归拟合商品房年竣工面积并建立走势预测模型。

最后，建立房价模型。由于往年全国房价的数据列拟合是一条较为单调的曲线，故全国房市就是一个灰色系统的典型，因此，本文建立灰色 GM(1,1)预测模型来预测房价并分析影响因素的灰色关联度。根据国家统计局提供的相关资料，本文首先选取 1991 年至 2010 年的商品房销售价格的原始数据来预测房价。然后选取 1991 年至 2009 年的影响房价的各项指标数据来进行灰色关联度分析。根据房价模型以及指标的灰色关联度分析结果，本文预测了全国房价的走势。

对于问题二 (④房地产行业与国民经济其他行业关系、⑤房地产行业态势分析)，需要建立房地产行业与国民经济其他行业的关系模型。依托国家数据库的有效数据，从房地产行业增加指数值、国内生产总值指数、交通运输仓储和邮政业增加值指数以及批发和零售业增加值指数 4 个代表性方面进行分析。利用线性回归模型建立房地产行业增加指数值与国内生产总值指数、交通运输仓储和邮政业增加值指数以及批发和零售业增加值指数的线性关系。然后利用多元线性回归模型，建立房地产行业增加值与影响因素的数学模型，并分析了其理论意义。

同时，需要建立对我国房地产行业态势分析模型。根据国家统计局提供的资料，选取全国 2010.7-2011.6 的相关数据，利用投机价格模型对全国房地产业的稳定度进行评估。利用投机价格衡量法，建立房地产业发展稳定度模型，采用 SPSS 软件进行数据分析，计算出房地产行业发稳定度均值为 0.76，说明全国房地产业的发展稳定度临介于警戒和安全区域之间，需要对房地产行业可持续发展进行规划。

对于问题三 (⑥房地产行业可持续发展)，需要建立房地产行业可持续发展模型。运用主观和客观相结合的方法，根据房地产可持续发展评价的原则，选取国家统计局提供的天津市的相关数据，构建适合天津市发展状况的评价指标体系。在此基础上运用层次分析法(AHP)来确立各级指标的权重，构建可持续发展模型，对天津市可持续发展程度进行合理评估及预测。

4. 模型假设

- 假设世界政局近几年不会出现非常大的波动；
- 假设国家统计局所提供的数据都是准确无误的；
- 假设城市经济发展水平用人均 GDP 来表示；
- 假设忽略消费者偏好(如有无学校、绿化率、停车位、热水供应状态、通信、房屋建筑形式等)对住房价格的影响；
- 假设忽略消费成本(如交通费用、物业费用、停车费用等)对房价的影响；
- 假设一个地区有多个房地产商家，不存在垄断现象，即忽略一些炒作对房价的影响；
- 假设地价在一定时间内变化幅度不大。

5. 符号说明

Y_1	商品房房产的年总销售面积	X_8	房地产企业本年经营总收入
Y_2	商品房房产的年总竣工面积	X_9	商品房本年施工面积
Y_3	房地产行业增加值指数	X_{10}	商品房本年新开工面积
$x^{(1)}(t+1)$	商品房的销售价格	X_{11}	人口自然增长率
X_1	人均 GDP	X_{12}	房地产业增加值
X_2	年总销售商品房价值	X_{13}	房地产企业本年土地转让收入
X_3	城镇年人均可支配收入	X_{14}	房地产企业本年资产负债率
X_4	年储蓄存储值	X_{15}	商品房本年销售价格
X_5	城镇人口	X_{16}	交通运输、仓储和邮政业增加值指数
X_6	城镇居民人均建筑面积	X_{17}	批发和零售业增加值指数
X_7	城镇就业人数		

6. 模型的建立及求解

6.1 问题一的模型建立及求解

对于问题一 **(①住房的需求、②供给与③价格)**，需要建立住房需求模型。本文依托国家统计局数据库提供的全国 1991 年至 2010 年房地产市场的一系列原始数据，应建立模型的需要，将数据进行了一定的技术处理，统一口径，数据库覆盖了 7 个宏观指标。这 7 个宏观指标分别为：①人均 GDP；②年总销售房价值；③城镇年人均可支配收入；④年储蓄存储值；⑤城镇人口；⑥城镇居民人均建筑面积；⑦城镇就业人数。利用 Person、Kendall 和 Spearman 方法进行相关性分析，判断出以上 7 个指标均与住房需求有较大的相关性，继而用主成分分析法消除指标间的共线性，取累计贡献率为 96.43% 的第一主成分，再用回归拟合房地产的年总销售面积并建立走势预测模型。

然后建立住房供给模型。引用需求的数学模型，并依托国家统计局数据库提供的全国 1991 年至 2010 年房地产市场的一系列原始数据，选取 9 个宏观指标为：①房地产企业本年经营总收入；②商品房本年施工面积；③商品房本年新开工面积；④人口自然增长率；⑤房地产业增加值；⑥房地产企业本年土地转让收入；⑦房地产企业本年资产负债率；⑧商品房本年销售价格；⑨人均国内生产总值。利用 Person、Kendall 和 Spearman 方法进行相关性分析，判断出以上 9 个指标均与住房供给有较大的相关性，继而用主成分分析法消除指标间的共线性，取累计贡献率为 91.76% 的第一主成分，再用回归拟合商品房年竣工面积并建立走势预测模型。

最后，建立房价模型。由于往年全国房价的数据列拟合是一条较为单调的曲线，故全国房市就是一个灰色系统的典型，因此，本文建立灰色 GM(1,1) 预测模型来预测房价并分析影响因素的灰色关联度。根据国家统计局提供的相关资料，本文首先选取 1991 年至 2010 年的商品房销售价格的原始数据来预测房价。然后选取 1991 年至 2009 年的影响房价的各项指标数据来进行灰色关联度分析。根据房价模型以及指标的灰色关联度分析结果，本文预测了全国房价的走势。

6.1.1 住房需求模型的建立及求解

该部分模型的建立依托于国家统计局数据库提供的一系列原始数据，应模型建立的需要，将原始数据进行了一定的技术处理，统一口径，数据库覆盖了 7 个宏观指标。这 7 个宏观指标为：①人均 GDP；②年总销售房价值；③城镇年人均可支配收入；④年储蓄存储值；⑤城镇人口；⑥城镇居民人均建筑面积；⑦城镇就业人数。

根据国家统计局数据库及权威网站资料查询、计算得出上述 7 个宏观指标的历年数据如表 6-1 所示。

表 6-1 历年宏观指标

年份	人均 GDP	总销售的商品房价值	城镇年人均可支配收入	储蓄存款	城镇人口	城镇居民人均建筑面积	城镇就业人数
单位	元	万元	元	亿元	万人	平方米	万人
1991	1892.8	2378011.56	1700.6	9244.9	31202.72	14.2	17465
1992	2311.1	4267415.7	2026.6	11757.3	32175.16	14.8	17861
1993	2998.4	8634091.81	2577.4	15203.5	33172.91	15.2	18262
1994	4044	10187563.15	3496.2	21518.8	34169.24	15.7	18653
1995	5045.7	12578350.54	4282.95	29662.2	35173.54	16.3	19040
1996	5845.9	14268140.46	4838.9	38520.8	37304.17	17	19922
1997	6420.2	17993309.49	5160.3	46279.8	39449.06	17.8	20781
1998	6796	25138273.9	5425.1	53407.5	41607.79	18.7	21616
1999	7158.5	29884556.09	5854	59621.8	43748.37	19.4	22412
2000	7857.7	39361618.56	6279.98	64332.4	45906.31	20.3	23151
2001	8621.7	48633823	6859.6	73762.4	48064.33	20.8	23940
2002	9398.1	60318652.5	7702.8	86910.7	50212.28	22.8	24780
2003	10542	79539889.17	8472.2	103617.7	52375.7	23.7	25639
2004	12335.6	106207495.9	9421.6	119555.4	54282.99	25	26476
2005	14185.4	175780345	10493	141051	56212	26.1	27331
2006	16499.7	208272754.7	11759.5	161587.3	57705.67	27.1	28310
2007	20169.5	298898638.1	13785.8	172534.2	59378.77	28	29350
2008	23707.7	250685354	15780.76	217885.4	60663.95	28.8	30210
2009	25575.5	443548155	17174.65	260767.3	62185.54	30	31120
2010	29762	525084671.2	19109.4	303302	66584.08	30.9	32200

6.1.1.1 住房需求模型的建立

1. 相关性分析公式的建立

相关系数是在线性相关条件下，用来测定两个变量之间相关关系密切程度的统计指标。它是由两个变量(X, Y)的协方差、X 变量的标准差和 Y 变量的标准差三个指标结合生成的，其基本计算公式如下：

$$r(X, Y) = \frac{\sigma^2(X, Y)}{\sigma(X) \sigma(Y)} \quad (1)$$

其中， $r(X, Y)$ 表示相关系数， $\sigma^2(X, Y)$ 表示 X、Y 两个变量的协方差， $\sigma(X)$ 表示 X 变量的标准差， $\sigma(Y)$ 表示 Y 变量的标准差。协方差的计算如下：

$$\sigma^2(X, Y) = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n} \quad (2)$$

其中 $\bar{X} = \frac{\sum X}{n}$ 是变量 X 的算术平均数， $\bar{Y} = \frac{\sum Y}{n}$ 是变量 Y 的算术平均数。

根据上式建立 Pearson 相关系数公式:

$$r(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

其中, 线性相关系数 r 说明了两个变量之间线性相关的密切程度和方向, $|r|$ 越大, 说明 X 与 Y 相关性越强, 反之则越弱。

然后建立 Kendall 相关系数公式:

$$K_{\tau-b} = \frac{\sum_{i < j} \xi(X_i, X_j, Y_i, Y_j)}{\sqrt{\frac{1}{2}n(n-1) - T_x} \sqrt{\frac{1}{2}n(n-1) - T_y}} \quad (4)$$

其中, $-1 \leq K_{\tau-b} \leq 1$ 。

最后建立 Spearman 相关系数公式:

$$\rho(X, Y) = 1 - \frac{6 \sum_{i=1}^n [R(X_i) - R(Y_i)]^2}{n(n^2 - 1)} = 1 - \frac{6T}{n(n^2 - 1)} \quad (5)$$

2. 主成分分析模型的建立

通过主成分分析, 将商品房房产的年总销售面积与七个宏观指标重新组合成一组新的相互无关的综合指标来代替原来的指标。

由此, 定义主成分为

$$\begin{aligned} F_1 &= a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + a_{14}X_4 + a_{15}X_5 + a_{16}X_6 + a_{17}X_7 \\ F_2 &= a_{21}X_1 + a_{22}X_2 + a_{23}X_3 + a_{24}X_4 + a_{25}X_5 + a_{26}X_6 + a_{27}X_7 \\ &\dots\dots\dots \\ F_7 &= a_{71}X_1 + a_{72}X_2 + a_{73}X_3 + a_{74}X_4 + a_{75}X_5 + a_{76}X_6 + a_{77}X_7 \end{aligned} \quad (6)$$

其中, F_i 是主成分, X_i 是指标列向量。上述方程要求:

$a_{1i}^2 + a_{2i}^2 + \dots + a_{7i}^2 = 1, i = 1, \dots, 7$ 且系数 a_{ij} 由下列原则决定:

(1) F_i 与 F_j 不相关; (2) F_1 是 X_1, \dots, X_7 的一切线性组合中方差最大的, F_2 是与 F_1 不相关的 X_1, \dots, X_7 一切线性组合中方差最大的, F_7 是与 F_1, \dots, F_6 都不相关的 X_1, \dots, X_7 一切线性组合中方差最小的。

最后求出第 i 个主成分的贡献率 $d_i = \frac{\lambda_i}{\sum_{i=1}^7 \lambda_i}$, 这个值越大, 表明第 i 主成分综

合 x_1, \dots, x_7 信息能力越强。前 k 个主成分的累积贡献率定义为 $\frac{\sum_{j=1}^k \lambda_j}{\sum_{i=1}^7 \lambda_i}$ 。在这里，主成分累积贡献率 $\geq 90\%$ 作为主成分个数的 m 的选择依据。

6.1.1.2 住房需求模型的求解

1. 相关性分析

分别利用 **Matlab** 和 **SPSS** 软件，得出房产的年总销售面积和 7 个指标之间的相关系数，具体表 6-2 如下：

表 6-2 房产的年总销售面积和 7 个宏观指标的相关性分析

分析方法	人均 GDP	总销售的房价值	城镇年人均可支配收入	储蓄存储	城镇人口	城镇居民人均建筑面积	城镇就业人数
kendall	0.979	0.989	0.979	0.979	0.979	0.979	0.979
spearman	0.997	0.998	0.997	0.997	0.997	0.997	0.997
pearson	0.981	0.979	0.979	0.983	0.933	0.954	0.951

表 6-3 7 个宏观指标内部的相关性分析（Pearson）

1						
0.967	1					
0.997	0.952	1				
0.995	0.970	0.994	1			
0.938	0.860	0.96	0.943	1		
0.955	0.886	0.974	0.959	0.995	1	
0.958	0.887	0.976	0.960	0.997	0.998	1

由表 6-2 和表 6-3 可以看出，无论是房产的年总销售面积与宏观指标，还是宏观指标内部之间相关系数的绝对值都比较接近于 1，也就是说他们之间的相关性都比较强，因此 7 个宏观指标都保留下来做主成分分析。

2. 主成分分析

根据相关性结果，利用 **Matlab** 软件进行主成分分析，结果如表 6-4 所示。

表 6-4：数据 1991-2010

主成分	特征值	方差贡献率	累积贡献率
1	6.74991	0.96427	0.96427
2	0.22030	0.03147	0.99574
3	0.02062	0.00295	0.99869
4	0.00493	0.00070	0.99939
5	0.00309	0.00044	0.99983
6	0.00065	0.00009	0.99992
7	0.00049	0.00007	0.99999

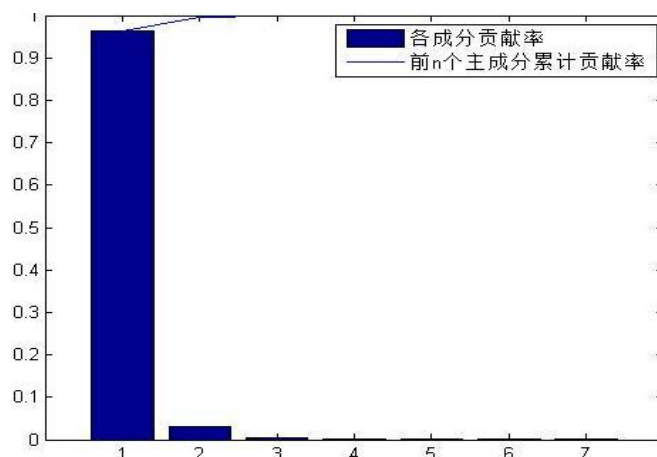


图 6-1 各成分贡献率

从表 6-4 和图 6-1 可以看出第一个特征值的累积贡献率已达 96.427%，说明第一个主成分基本包含了全部指标具有的信息，所以仅取第一个主成分，并计算出相应的特征向量：

表 6-5 第一特征向量

因素	X_1	X_2	X_3	X_4	X_5	X_6	X_7
特征向量	0.3814	0.3651	0.3837	0.3819	0.3748	0.3790	0.3795

得出相应得回归模型为

$$X = 0.3814X_1 + 0.3651X_2 + 0.3837X_3 + 0.3819X_4 + 0.3748X_5 + 0.3790X_6 + 0.3795X_7 \quad (7)$$

3. 回归拟合

通过主成分分析，得到了既与房产的年总销售面积关系密切，又能充分反映指标信息的新的综合指标。下面就这个新的综合指标与年总销售面积做回归，结果如表 6-6 所示。

表 6-6 回归拟合

	RSq	F	df1	Df2	Sig	Constant	b1	b2	b3
Linear	0.96	423.9	1	18	5.8E-14	9,542.4	0.001		
Logarithmic	0.85	103.5	1	18	6.8E-09	-280,265.7	18,876.9		
Inverse	0.28	6.8	1	18	1.7E-02	44,084.2	-6E+10		
Quadratic	0.99	1,879.4	2	17	1.1E-20	4,579.2	0.001	-2E-12	
Cubic	0.99	3,938.8	3	16	3.7E-23	2,718.1	0.001	-6E-12	1.5E-20
Compound	0.71	44.0	1	18	3.1E-06	10,089.8	1.000		
Power	0.99	3,122.3	1	18	1.2E-21	0.2	0.693		
S	0.57	24.2	1	18	1.0E-04	10.4	-3E+06		
Growth	0.71	44.0	1	18	3.1E-06	9.2	0.000		
Exponential	0.71	44.0	1	18	3.1E-06	10,089.8	0.000		
Logistic	0.71	44.0	1	18	3.1E-06	0.0	1.000		

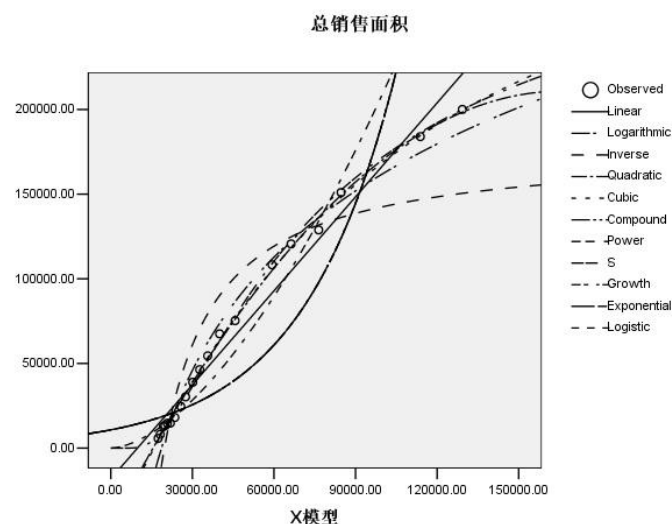


图 6-2 各种拟合方法对应的曲线

根据统计学中 RSq 与 F 越大, 拟合程度越好, 同时也可以从表 6-6 和图 6-2 可以看出, 三次曲线 **Cubic** 拟合的最好, 曲线的数学形式为:

$$Y_1 = 2718.1 \times 10^3 - 0.001151 X^2 + 9914 \times 10^{-20} \quad (8)$$

商品房产年总销售面积未来 5 年的预测数据如表 6-7 所示, 其中单位为万/平方米。

表 6-7 未来 5 年商品房产年总销售面积(万/平方米)

年份	预计商品房产年总销售面积
2011	106620
2012	121430
2013	143220
2014	175430
2015	222410

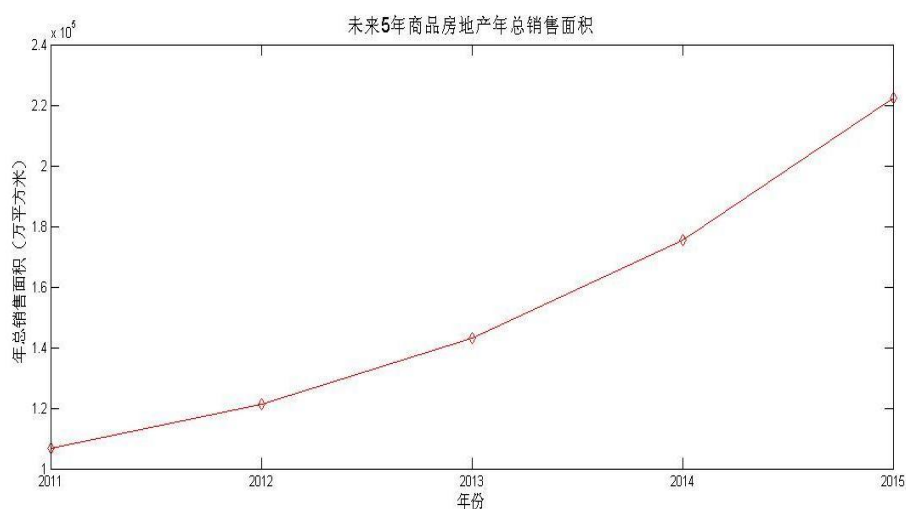


图 6-3 未来 5 年商品房产年总销售面积

由图 6-3 的曲线可以看出，未来 5 年之内商品房总销售面积呈稳步上升的发展趋势。

6.1.2 住房供给模型建立及求解

该部分模型的建立依托于国家统计局数据库提供的一系列原始数据，应模建立的需要，将原始数据进行了一定的技术处理，统一口径，数据库覆盖了 9 个宏观指标。这 9 个宏观指标为：①房地产企业本年经营总收入；②商品房本年施工面积；③商品房本年新开工面积；④人口自然增长率；⑤房地产业增加值；⑥房地产企业本年土地转让收入；⑦房地产企业本年资产负债率；⑧商品房本年销售价格；⑨人均国内生产总值。

根据国家统计局数据库及一些网站、资料查询，计算得出如表 6-8 所示。

表 6-8 宏观指标数据

年份	年经营 总收入	本年施 工面积	本年新开 工面积	人口增 长率	房地产 增加值	土地转 让收入	资产 负债 率	销售 价格	人均 GDP
单位	万元	万平方米	万平方米	%	亿元	万元	%	元	元
1997	22184557	44985.5	14026.98	10.06	2921.1	1032847	76.2	1997	6420.2
1998	29512078	50770.1	20387.9	9.14	3434.5	1322454	76.1	2063	6796
1999	30260108	56857.6	22579.41	8.18	3681.8	1032492	76.1	2053	7158.5
2000	45157119	65896.9	29582.64	7.58	4149.1	1296054	75.6	2112	7857.7
2001	54716555	79411.7	37394.18	6.95	4715.1	1889894	75	2170	8621.7
2002	70778478	94104	42800.52	6.45	5346.4	2251311	74.9	2250	9398.1
2003	91372734	117526	54707.53	6.01	6172.7	2797200	75.8	2359	10542
2004	133144608	140451.4	60413.86	5.87	7174.1	4100917	74.1	2778	12335.6
2005	147693468	166053.3	68064.44	5.89	8516.4	3414314	72.7	3168	14185.4
2006	180468000	194786.4	79252.83	5.28	10370.5	3006000	74.1	3367	16499.7
2007	233971000	236318.2	95401.53	5.17	13809.7	4279000	74.4	3864	20169.5
2008	266968000	283266.2	102553.37	5.08	14738.7	4668000	72.3	3800	23707.7
2009	346062000	320368.2	116422.05	5.05	18654.7	4980000	73.5	4681	25575.5

6.1.2.1 住房供给模型的建立

由于住房需求与住房供给是一个相互对应的问题，所以我们采用问题一住房需求的模型对问题二进行建模。

6.1.2.2 住房供给模型的求解

1. 相关性分析

分别利用 Matlab 和 SPSS 软件，得出商品房的年总竣工面积和 9 个指标之间的相关系数，具体如表 6-9 所示。

表 6-9 商品房的年总竣工面积和 9 个宏观指标的相关性分析

分析 方法	年经营 总收入	本年施 工面积	本年新开 工面积	人口 增长率	房地产 增加值	土地转 让收入	资产负 债率	销售 价格	人均 GDP
kendall	1.000	0.846	-0.753	0.949	1.000	0.846	-0.753	0.949	1.000
spearman	1.000	1.000	1.000	-0.995	1.000	0.951	-0.895	0.989	1.000
pearson	0.964	0.976	0.992	-0.915	0.954	0.946	-0.857	0.953	0.963

表 6-10 9 个宏观指标内部的相关性分析 (Pearson)

1								
0.996	1							
0.984	0.990	1						
-0.808	-0.828	-0.891	1					
0.996	0.993	0.978	-0.791	1				
0.932	0.935	0.953	-0.874	0.909	1			
-0.794	-0.819	-0.822	0.781	-0.764	-0.837	1		
0.992	0.984	0.971	-0.779	0.990	0.907	-0.782	1	
0.994	0.998	0.983	-0.804	0.994	0.920	-0.809	0.982	1

由表 6-9 和表 6-10 可以看出, 无论是商品房的年总竣工面积与宏观指标, 还是宏观指标内部之间相关系数的绝对值都比较接近于 1, 也就是说他们之间的相关性都比较强, 因此 9 个宏观指标都保留下来做主成分分析。

2. 主成分分析

根据相关性结果, 利用 Matlab 软件进行主成分分析。

表 6-11: 数据 1997-2009

主成分	特征值	方差贡献率	累积贡献率
1	8.25808	0.91757	0.91757
2	0.40864	0.04541	0.96298
3	0.23006	0.02556	0.98854
4	0.07739	0.00860	0.99714
5	0.01931	0.00215	0.99929
6	0.00373	0.00041	0.9997
7	0.00181	0.00020	0.9999
8	0.00079	0.00009	0.99999
9	0.00017	0.00001	1

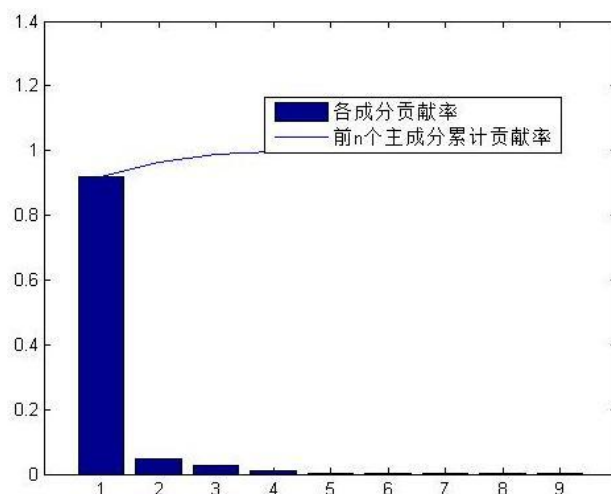


图 6-4 各成分贡献率

从表 6-11 和图 6-4 可以看出第一个特征值的累积贡献率已达 91.757%，说明第一个主成分基本包含了全部指标具有的信息，所以仅取第一个主成分，并计算出相应的特征向量：

表 6-12 第一特征向量

因素	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_1
特征向量	0.3439	0.3457	0.3466	-0.3041	0.3409	0.3338	-0.2979	0.3397	0.3434

得出相应得回归模型为

$$X = 0.3439X_8 + 0.3457X_9 + 0.3466X_{10} - 0.3041X_{11} + 0.3409X_{12} + 0.3338X_{13} - 0.2979X_{14} + 0.3397X_{15} + 0.3434X_1 \quad (9)$$

3. 回归拟合

通过主成分分析，得到了既与房产的年总销售面积关系密切，又能充分反映指标信息的新的综合指标。下面就这个新的综合指标与年总销售面积做回归，结果如表 6-13 所示。

表 6-13 回归拟合

	RSq	F	df1	df2	Sig	Constant	b1	b2	b3
Linear	0.93	148.0	1	11	1E-07	18,486.7	5E-04		
Logarithmic	0.97	474.5	1	11	2E-10	-321,057.9	20,980.9		
Inverse	0.81	49.1	1	11	2E-05	61,788.9	-4E+11		
Quadratic	0.98	250.1	2	10	2E-09	10,902.4	9E-04	-3E-12	
Cubic	0.98	208.6	3	9	1E-08	6,692.3	1E-03	-1E-11	4E-20
Compound	0.81	48.4	1	11	2E-05	20,825.9	1E+00		
Power	0.97	489.8	1	11	1E-10	2.2	5E-01		
S	0.93	153.0	1	11	8E-08	11.1	-1E+07		
Growth	0.81	48.4	1	11	2E-05	9.9	1E-08		
Exponential	0.81	48.4	1	11	2E-05	20,825.9	1E-08		
Logistic	0.81	48.4	1	11	2E-05	5E-05	1E+00		

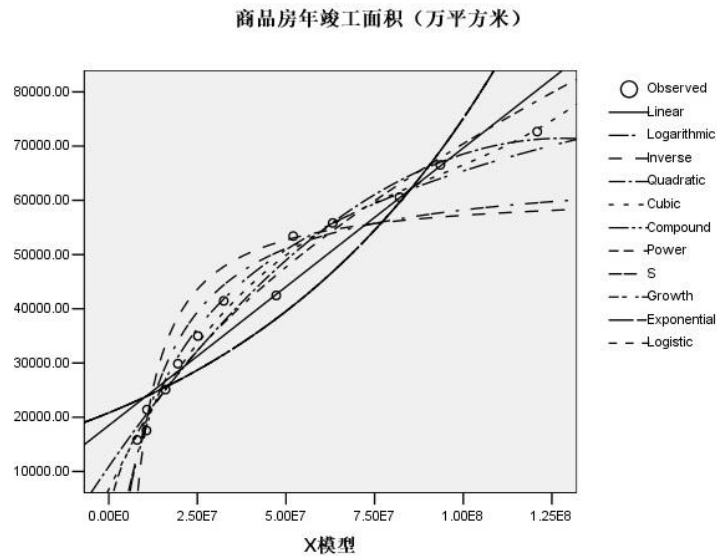


图 6-5 各种拟合方法对应的曲线

根据统计学中 RSq 与 F 越大, 拟合程度越好, 同时也可以从表 6-14 和图 6-5 可以看出, 三次曲线 **Cubic** 拟合的最好, 曲线的数学形式为:

$$Y_2 = 6692.358 - 0.001342X^1 + X^2 + 1.16 \times 10^{-20} \quad (10)$$

预测未来 5 年商品房产年竣工面积为: (单位: 万平方米)

表 6-14 未来 5 年商品房产年竣工面积(万/平方米)

年份	预计商品房产年竣工面积
2011	87088
2012	99301
2013	113380
2014	138774
2015	161380

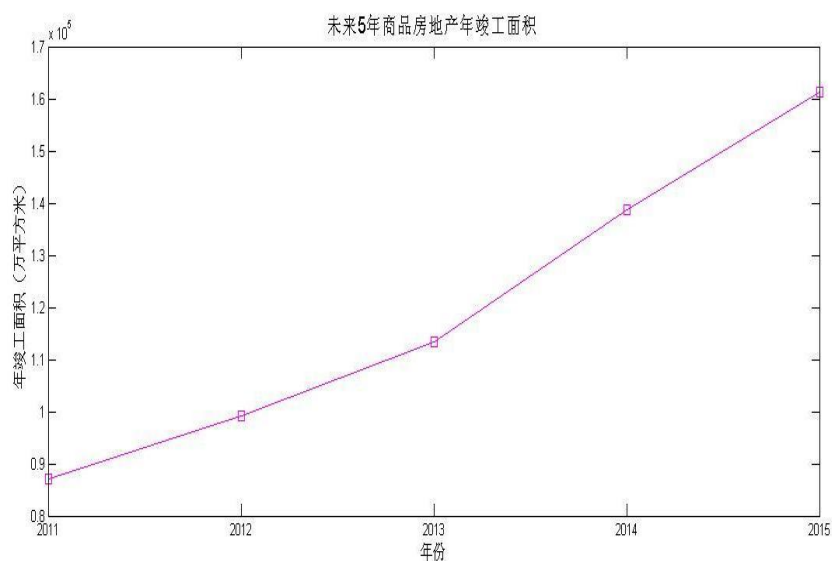


图 6-6 未来 5 年商品房产年竣工面积

由图 6-6 的曲线可以看出，未来 5 年之内商品房年竣工面积呈稳步上升的发展趋势。

6.1.3 房价模型的建立及求解

问题需要我们运用全国 1991 年-2010 年房地产市场的有效数据来建立房价预测模型。由于数据量单调，我们运用以上数据建立灰色预测模型。同时找到可能和商品房销售价格有关的影响因素，运用灰色关联分析各个因素和商品房售价的关系。观察各自的关联度，分析得到影响商品房价格的重要因素。

6.1.3.1 灰色房价模型的建立与求解

1. 建立原始数据矩阵：

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(20)\}$$

2. 有原始数据计算一次累加序列：

$$x^{(1)}(1) = x^{(0)}(1)$$

$$x^{(1)}(2) = x^{(0)}(1) + x^{(0)}(2)$$

.....

$$x^{(1)}(20) = x^{(0)}(1) + x^{(0)}(2) + \dots + x^{(0)}(20)$$

3. 构造数据矩阵 B 和数据向量 Y：

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2}[x^{(1)}(19) + x^{(1)}(20)] & 1 \end{bmatrix} \quad Y = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(20)]^T$$

4. 发展灰数 α 和内生控制灰数 μ 的确定：

$$U = \begin{bmatrix} \alpha \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (11)$$

5. 预测模型的建立：根据灰色理论，预测模型的一般形式为微分方程，即：

$$x^{(1)}(k+1) - \left[x^{(0)}(k+1) + \frac{\mu}{\alpha} \right] e^{-\alpha k} + \frac{\mu}{\alpha} \quad (12)$$

由于该模型求得的跟年份预测值为一次累加数据，需将

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \quad (13)$$

6. 模型的求解：

根据国家统计局提供的相关资料，我们选取 1991 年至 2010 年的商品房销售价格的数据来预测房价。每年的商品房销售价格如下表 6-15 所示：

表 6-15 商品房销售价格表

年份	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
商品房销售价格(元/平方米)	786	995	1291	1409	1591	1806	1997	2063	2053	2112
年份	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
商品房销售价格(元/平方米)	2170	2250	2359	2778	3168	3367	3864	3800	4681	5032

利用 Matlab 编程求解得到商品房销售价格的计算公式为：

$$x^{(1)}(t+1) = 13615.42e^{0.079\frac{t}{11}} - (124 \quad (14)$$

其中 t 表示从 2010 年开始的第几年。

7. 残差检验及修正：

残差模型：若用原始经济时间序列 $x^{(0)}$ 建立的 GM(1, 1) 模型检验不合格或精度不理想时，要对建立的 GM(1, 1) 模型进行残差修正或提高模型的预测精度。修正的方法是建立 GM(1, 1) 的残差模型。

通过 Matlab 编程计算出 1991-2010 年的各年房价销售残差如下表 6-16 所示：

表 6-16 商品房销售价格残差表

年份	预计商品房销售价格（元/平方米）	年份	预计商品房销售价格（元/平方米）
1991	0	2001	-139.8297
1992	-520.7992	2002	-251.236
1993	69.2503	2003	-349.5034
1994	86.0088	2004	-154.9462
1995	158.3778	2005	-7.9877
1996	254.6622	2006	-72.1691
1997	317.109	2007	139.8408
1998	243.9032	2008	-232.7654
1999	83.1619	2009	314.0555
2000	-21.0706	2010	303.1843
残差平均值		10.9627	

通过 Matlab 计算得出各年的残差值从而得到残差平均值为 10.9627。对上面计算出的商品房销售价格的计算公式进行更正后的计算式为：

$$x^{(1)}(t+1) = 13615.42e^{0.079\frac{t}{11}} - (12444.887 \quad (15)$$

8. 模型的计算

根据上述公式以及 Matalab 编程得出 2011-2015 年得商品房预计销售价格如下表 6-17 所示:

表 6-17 预计商品房销售价格表

年份	预计商品房销售价格 (元/平方米)
2011	5120.7
2012	5545
2013	6004.5
2014	6502.1
2015	7040.9

9. 模型精度检验:

后验差比值 C 和误差概率 P 是检验模型精度的常用方法, 公认的精度评价标准如下表 6-18 所示:

表 6-18 精度评价标准表

	C 值	P 值
1 级 (好)	<0.35	≥0.95
2 级 (合格)	<0.5	≥0.8
3 级 (勉强)	<0.65	≥0.7
4 级 (不合格)	≥0.65	<0.7

用 Matlab 求解得到模型 C= 0.1992, P=1 。对照上表可以认为该模型的精度很好。

6.1.3.2 灰色关联度模型与求解:

1. 关联度定义: 衡量指标序列与母指标相似程度的测度。
2. 选定母指标: 选取商品房销售价格为母指标。
3. 原始数据无量纲化处理:

$$x_i(j)' = \frac{x_i(j)}{\sum_{j=1}^{20} x_i(j)} 20 \quad (16)$$

4. 构造关联度公式:

$$\varepsilon_i(j) = \frac{\min \min |x_0(k)' - x_i(k)'| + \rho \max \max |x_0(k)' - x_i(k)'|}{|x_0(k)' - x_i(k)'| + \rho \max \max |x_0(k)' - x_i(k)'|} \quad (17)$$

其中: $x_i(j)' = \frac{x_i(j)}{\sum_{j=1}^{20} x_i(j)} 20$, $\rho = 0.5$

5. 关联度计算:

$$R_1 = \frac{1}{8} \sum_{i=1}^{20} \varepsilon_1(k)$$

$$R_2 = \frac{1}{8} \sum_{i=1}^{20} \varepsilon_2(k) \quad (18)$$

.....

$$R_{20} = \frac{1}{8} \sum_{i=1}^{20} \varepsilon_{20}(k)$$

6. 数据整合和模型求解：

根据国家统计局提供的相关资料，我们选取 1991 年至 2009 年的各项指标的数据来进行灰色关联度分析。每年的各项指标如下表 6-19 所示：

表 6-19 各年指标值表

指标 年份	商品房本年 销售价格（元 /平方米）	人均 GDP （元）	房地产企业 本年经营总 收入（万元）	房地产企业本 年土地转让收 入（万元）	在岗职 工平均 工（元）	房地产 业增加 值指数	流通中 现金 （亿元）
1991	786	1892.8	2840325	153810	2340	112	3177.8
1992	995	2311.1	5285565	427420	2711	134.7	4336
1993	1291	2998.4	11359074	839281	3371	110.8	5864.7
1994	1409	4044	12881866	959357	4538	112	7288.6
1995	1591	5045.7	17316624	1943981	5500	112.4	7885.3
1996	1806	5845.9	19687850	1203378	6210	104	8802
1997	1997	6420.2	22184557	1032847	6470	104.1	10177.6
1998	2063	6796	29512078	1322454	7479	107.7	11204.2
1999	2053	7158.5	30260108	1032492	8346	105.9	13455.5
2000	2112	7857.7	45157119	1296054	9371	107.1	14652.7
2001	2170	8621.7	54716555	1889894	10870	111	15688.8
2002	2250	9398.1	70778478	2251311	12422	109.9	17278
2003	2359	10542	91372734	2797200	14040	109.8	19746
2004	2778	12335.6	133144608	4100917	16024	105.9	21468.3
2005	3168	14185.4	147693468	3414314	18364	112.2	24031.7
2006	3367	16499.7	180468000	3006000	21001	115.5	27072.6
2007	3864	20169.5	233971000	4279000	24932	124.4	30375.2
2008	3800	23707.7	266968000	4668000	29229	101	34219
2009	4681	25575.5	346062000	4980000	32736	111.3	38246

利用上述关联度计算公式，在 Matlab 中编程计算出各项指标与母指标商品房本年销售价格的关联度 R_i 的值，具体见下表 6-20。

表 6-20 各项指标与母指标商品房本年销售价格的关联度值

指标 i	关联度 R_i
商品房本年销售价格（元/平方米） a_1	1
人均 GDP（元） a_2	0.7625
房地产企业本年经营总收入（万元） a_3	0.7798
房地产企业本年土地转让收入（万元） a_4	0.7076
在岗职工平均工资（元） a_5	0.7209
房地产业增加值指数（可比价，上年=100） a_6	0.6759
流通中现金（亿元） a_7	0.7189

根据灰色预测相关理论，关联度大于 0.6 即可接受。由上表中数据分析可知，各个因素与母指标商品房本年销售价格的关联度均大于 0.6，相对较大，可以认为对商品房本年销售价格有较大影响。将各个因素按照关联度大小排序 $a_3 > a_2 > a_5 > a_7 > a_4 > a_6$ ，因此影响商品房本年销售价格的重要因素由大到小为：

①房地产企业本年经营总收入；②人均 GDP；③在岗职工平均工资；④流通中现金；⑤房地产企业本年土地转让收入；⑥房地产增加值指数。特别地，人均 GDP 和房地产企业本年经营总收入两项指标均达到 0.75，所以分析得出此两项指标对商品房本年销售价格有很大的影响。

6.2 问题二的模型建立及求解

对于问题二 ④房地产行业与国民经济其他行业关系、⑤房地产行业态势分析，需要建立房地产行业与国民经济其他行业的关系模型。依托国家数据库的有效数据，从房地行业增加指数值、国内生产总值指数、交通运输仓储和邮政业增加值指数以及批发和零售业增加值指数 4 个代表性方面进行分析。利用线性回归模型建立房地行业增加指数值与国内生产总值指数、交通运输仓储和邮政业增加值指数以及批发和零售业增加值指数的线性关系。然后利用多元线性回归模型，建立房地产行业增加值与影响因素的数学模型，并分析了其理论意义。

同时，需要建立对我国房地产行业态势分析模型。根据国家统计局提供的资料，选取全国 2010.7-2011.6 的相关数据，利用投机价格模型对全国房地产业的稳定度进行评估。利用投机价格衡量法，建立房地产业发展稳定度模型，采用 SPSS 软件进行数据分析，计算出房地产行业发稳定度均值为 0.76，说明全国房地产业的发展稳定度临介于警戒和安全区域之间，需要对房地产行业可持续发展进行规划。

6.2.1 房地产行业与国民经济其他行业关系模型的建立及求解

6.2.1.1 房地产行业与国民经济其他行业关系模型的建立

1. 相关分析

相关关系式反映现象之间确实存在的，但关系数值不固定的相互依存关系。相关系数是说明变量之间在直线相关条件下相关关系密切程度和方向的统计分析指标。其定义公式为：

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (19)$$

2. 回归分析

回归分析就是对具有相关关系的变量之间数量变化的一般关系进行测定，确定一个相关的数学表达式，以便于进行估计或预测的统计方法。设 $x_1, x_2, x_3, \dots, x_p$ 是 p 个可以精确测量或可控制的变量。如果变量 y 与 $x_1, x_2, x_3, \dots, x_p$ 之间的内在联系是线性的，那么进行 n 次试验，则可得 n 组数据： $(y_i, x_{i1}, x_{i2}, \dots, x_{ip})$ ， $i=1, 2, \dots, n$ 。它们之间的关系可表示为：

$$\begin{aligned} y_1 &= b_0 + b_1 x_{11} + b_2 x_{12} + \dots + b_p x_{1p} + \varepsilon_1 \\ y_2 &= b_0 + b_1 x_{21} + b_2 x_{22} + \dots + b_p x_{2p} + \varepsilon_2 \\ &\dots\dots\dots \\ y_n &= b_0 + b_1 x_{n1} + b_2 x_{n2} + \dots + b_p x_{np} + \varepsilon_n \end{aligned} \quad (20)$$

其中， $b_0, b_1, b_2, \dots, b_p$ 是 $p+1$ 个待估参数， ε_i 表示第 i 次试验中的随机因素对 y_i 的影响。为简单起见，将此 n 个方程表达式表示成矩阵形式：

$$Y = XB + \varepsilon \quad (21)$$

上式便是 p 元线性回归的数学模型。

为了求出多元线性回归模型中的参数 $b_0, b_1, b_2, \dots, b_p$ ，可采用最小二乘法，即在其数学模型所属的函数类中找一个近似函数，使得这个近似函数在已知的对应数据上尽可能和真实函数接近。

设 $c_0, c_1, c_2, \dots, c_p$ 分别为 $b_0, b_1, b_2, \dots, b_p$ 的最小二乘估计，则多元回归方程（即近似函数）为：

$$y = c_0 + c_1 x_1 + c_2 x_2 + \dots + c_p x_p \quad (22)$$

其中， $c_0, c_1, c_2, \dots, c_p$ 叫做回归方程的回归系数。

6.2.1.2 房地产行业与国民经济其他行业关系模型的求解

1. 指标的选取

影响房地产行业与国民经济其他行业关系的因素很多，由于本题附件只提供了房地产行业增加值、国内生产总值指数、交通运输仓储和邮政业增加值指数以及批发和零售业增加值指数，所以本文对这 4 个代表性指数进行分析。房地产行业与国民经济其他行业相关影响因素如下表所示，表中的指数值均是相对于 1978 年的对应数据。

表 6-21 指数值指标数据

	房地产行业 增加值指数	国内生产 总值指数	交通运输、仓储和 邮政业增加值指数	批发和零售业 增加值指数
1991	501.7	307.6	321.4	345.8
1992	675.9	351.4	353.7	382.2
1993	748.6	400.4	398.1	414.9
1994	838.2	452.8	432	448.9
1995	942.5	502.3	479.4	485.9
1996	980.5	552.6	532.4	523
1997	1021	603.9	581.3	568.8
1998	1099.4	651.2	642.9	605.9
1999	1164.7	700.9	721.2	658.6
2000	1247.5	759.9	783	720.7
2001	1384.6	823	852	786.2
2002	1521.8	897.8	912.7	855.5
2003	1671	987.8	968.6	940.5
2004	1769.6	1087.4	1108.9	1002.2
2005	1986.1	1210.4	1233.1	1132.8
2006	2293.5	1363.8	1356	1353.3
2007	2852.1	1557	1516	1626.9
2008	2879.5	1707	1627.1	1884.7
2009	3204.3	1862.5	1687.9	2112.8

2. 房地产行业增加值与国民生产总值 GDP 的关系

以国民生产总值 GDP X_1 为横坐标，以房价 Y_3 为纵坐标，将上述数据做一元线性回归。设

$$Y_3 = a + b * X_1$$

通过 Matlab 编程计算得出相关参数如下表 6-22 所示。

表 6-22 房地产行业增加值指数与国民生产总值指数的关系参数表

参数	参数估计值	参数置信区间
a	17.9783	[-54.8206, 90.7774]
b	1.6949	[1.6217, 1.7682]

$$R^2=0.9929, F=2.3841e+03, p=1.0165e-19 < 0.0001$$

由此得到房地产行业增加值指数与国民生产总值 GDP 指数的线性关系为：

$$Y_3 = 17.9783 + 1.6949 * X_1 \quad (23)$$

Matlab 中画出房地产行业增加值指数与国民生产总值指数的关系图如图 6-7 所示。

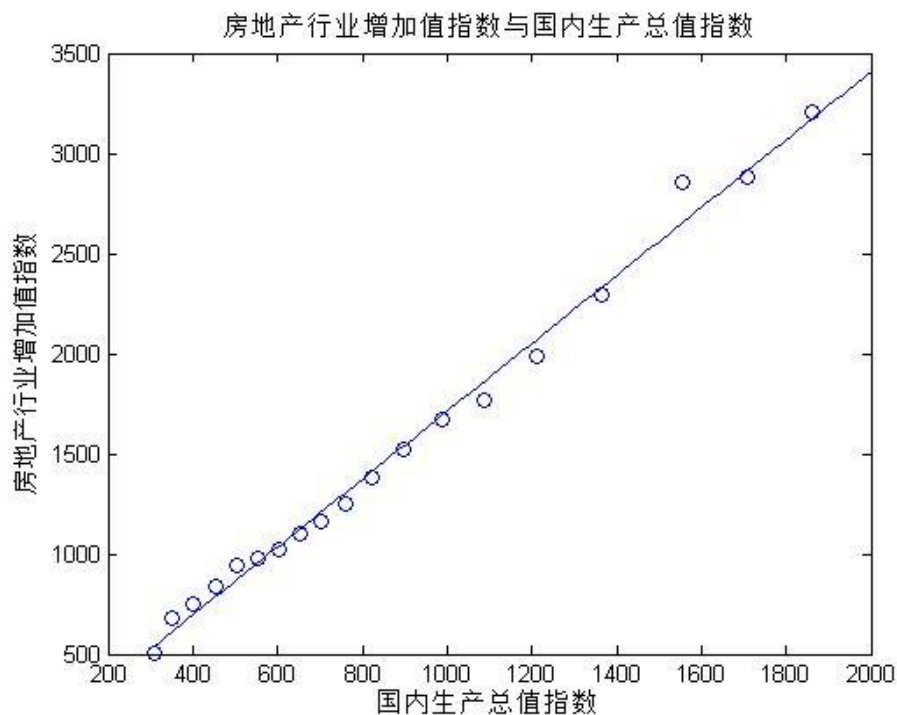


图 6-7 房地产行业增加值指数与国民生产总值指数的关系图

从上图可以看出，房地产行业增加值与国民生产总值 GDP 的相关性系数趋近于 1，说明两者的关系紧密。

3. 房地产行业增加值与交通运输、仓储和邮政业增加值的关系

运用上述同样的方法，Matlab 编程得到相关参数如下表 6-23 所示。

表 6-23 房地产行业增加值指数与交通运输、仓储和邮政业增加值指数的关系参数表

参数	参数估计值	参数置信区间
a	-37.3952	$[-1.6588e+02, 91.0860]$
b	1.7866	$[1.6541, 1.9192]$

$$R^2 = 0.9794, F = 8.0874e+02, p = 8.9131e-16 < 0.0001$$

由此得到房地产行业增加值指数与交通运输、仓储和邮政业增加值指数的线性关系为：

$$Y_3 = -37.3952 + 1.7866X_3 \quad (24)$$

Matlab 中画出房地产行业增加值指数与交通运输、仓储和邮政业增加值指数的关系图如图 6-8 所示。

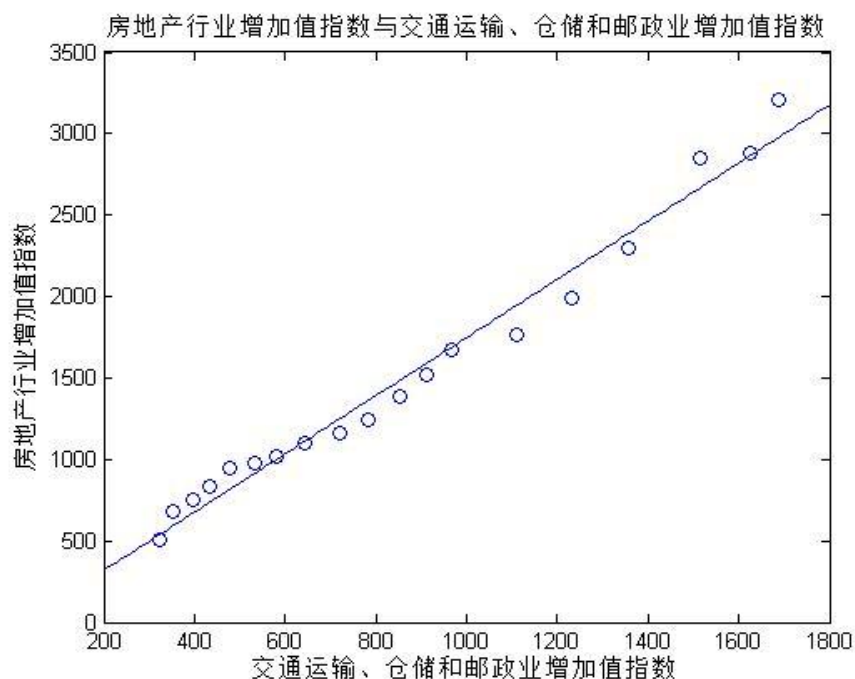


图 6-8 房地产行业增加值指数与交通运输、仓储和邮政业增加值指数的关系图

从上图可以看出，房地产行业增加值与交通运输、仓储和邮政业增加值的相关性系数趋近于 1，说明两者的关系紧密。

4. 房地产行业增加值与批发和零售业增加值的关系

运用上述同样的方法，Matlab 编程得到相关参数如下表 6-24。

表 6-24 房地产行业增加值指数与批发和零售业增加值指数的关系参数表

参数	参数估计值	参数置信区间
a	1.6569e+02	[66.7622, 2.6462e+02]
b	1.5214	[1.4245, 1.6183]

$$R^2 = 0.9847, F = 1.0975e+03, p = 6.9470e-17 < 0.0001$$

由此得到房地产行业增加值指数与批发和零售业增加值指数的线性关系为：

$$Y_3 = 1.6569 \times 10^2 + 1.5214 X_1 \quad (25)$$

Matlab 中画出房地产行业增加值指数与批发和零售业增加值指数的关系图如图 6-9 所示。

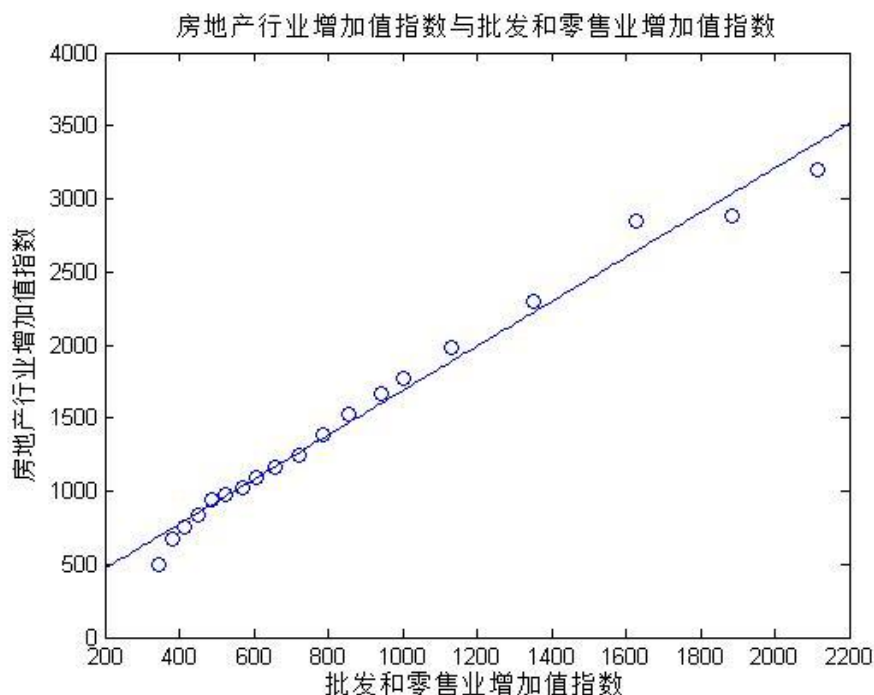


图 6-9 房地产行业增加值指数与批发和零售业增加值指数的关系图

由图 6-9 可以看出，房地产行业增加值与批发和零售业增加值的相关性系数趋近于 1，说明两者的关系紧密。

5. 多元回归模型的建立及求解

通过上述的三个一元回归，可以看出国内生产总值、交通运输仓储和邮政业增加值以及批发和零售业增加值对于房地产行业增加值的影响都比较大。先对这三个因素变量与房地产行业增加值指数做多元回归。

设 $Y = a + bX_1 + cX_{16} + dX_{17}$ ，编写 Matlab 程序运行可得房地产行业增加值与三个因素变量之间关系的参数表如下表 6-25 所示。

表 6-25 房地产行业增加值与三个因素变量之间关系的参数表

参数	参数估计值	参数置信区间
a	54.8083	$[-24.8395, 1.3446e+02]$
b	1.8607	$[-0.3593, 4.0806]$
c	-0.4439	$[-1.9640, 1.0761]$
d	0.2284	$[-0.5837, 1.0405]$

$R^2=0.9945$, $F=8.9676e+02$, $p=3.8797e-17 < 0.0001$

所以可得线性关系为：

$$Y_3 = 54.8083 + 1.8607X_1 - 0.4439X_{16} + 0.2284X_{17} \quad (26)$$

上式即为确定的最终模型。

残差分析：可以通过残差图分析残差的有效性，期望的残差值应该在 0 上下随机摆动，根据此可以寻找到异常值。图 6-10 给出了 1991-2009 各年数据的

残差值图。

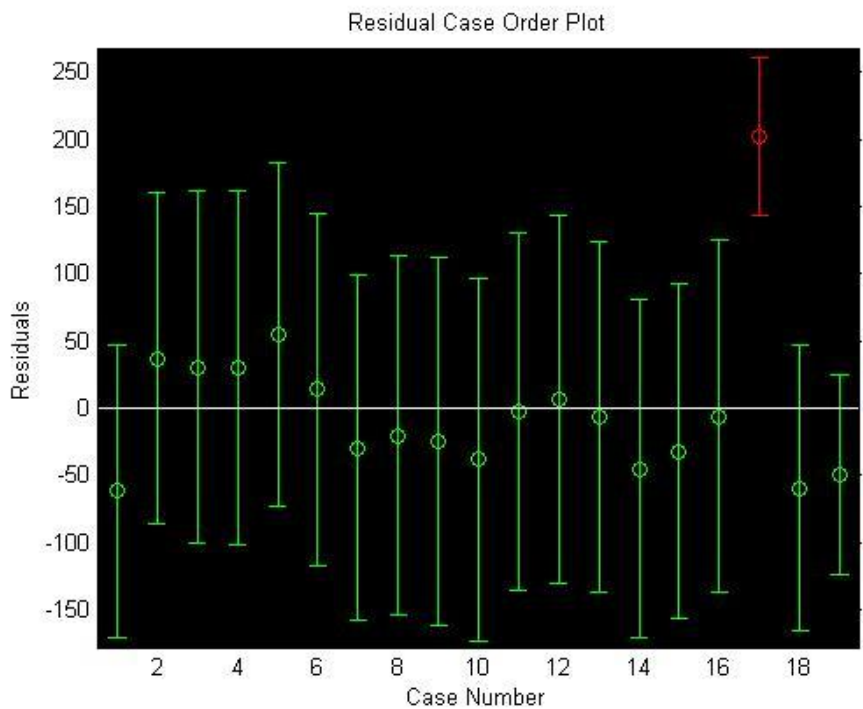


图 6-10 房地产行业增加值残差图

图 6-10 是按照样本顺序绘制的残差图，红色表示该样本为异常值。由该图可知 2007 年的相关数据是有异常的。

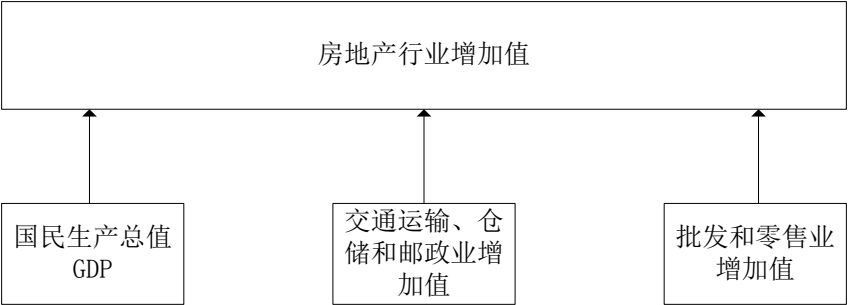


图 6-11 影响房地产行业增加值的相关因素图

- 上述的线性回归模型以及图 6-11 的理论意义如下：
- (1) 房地产行业投资与国民生产总值增长之间存在长期稳定的均衡关系。由多元线性回归得出的方程可知，房地产投资对国民生产总值增长影响的长期弹性为 0.5374，即房地产投资每增加 1%，GDP 增加 0.5374%。可见，从长期来看，我国房地产投资对国民生产总值增长的贡献是十分显著的。当房地产投资增长速度与国民生产总值增长速度相协调，国家宏观经济政策得当的时候，房地产投资对国民生产总值增长有很大的促进作用，反之，有阻碍作用。
 - (2) 由多元线性回归得出的方程可知，交通运输、仓储和邮政业值每增加 1%，房地产行业值就减少 0.4439%。可见，从长期来看，我国交通运输、仓储和邮政业对房地产行业投资有一定的阻碍。
 - (3) 房地产行业投资与批发和零售业值增长之间存在长期稳定的均衡关

系。由多元线性回归得出的方程可知，批发和零售业值增长对房地产投资影响的长期弹性为 0.2284，即批发和零售业值每增加 1%，房地产行业增加值上涨 0.2284%。可见，从长期来看，我国批发和零售业对房地产行业的贡献是十分显著的。当批发和零售行业值增长速度与房地产投资增长速度相协调，批发和零售行业对房地产投资有很大的促进作用，反之，有阻碍作用。

6.2.2 房地产行业态势稳定度模型的建立及求解

6.2.2.1 房地产行业态势稳定度模型的建立

本题选取房地产价格作为房地产行业态势分析的依据。房地产价格出现脱离市场基础价值的高估，从数量上看是指资产价格高出市场基础价值的那部分。本文将资产的基本价值或市场决定的合理价格部分称之为“标准价格”，高出市场基础价值的那部分称之为投机价格。即：

$$\text{房地产价格} = \text{标准价格} + \text{投机价格}$$

1. 投机价格的计算

按照这一思路，房地产价格由两个部分组成：一是房地产使用价值，这是房地产所有者从房地产使用中获得的收益，即房地产的“标准价格”。对住宅而言，房地产使用价值是使用、消费住宅获得的满足和享受。二是房地产资产预期资本收益的现值，这是房地产所有者预期价格变化导致其资本收益的增减量，即房地产投机部分价格。用公式表示为：

$$P_t = P_t^0 + P_t^m = P_t^0 + \frac{P_{t+1}^m}{1+i_t} \quad (27)$$

其中， P_t 为第 t 期的房地产价格， P_t^0 为房地产使用价值， P_t^m 为房地产资产预期资本收益的现值， P_{t+1}^m 为房地产资产下一期预期资本收益， i_t 为第 t 期的利率。

方程(27)中的房地产“标准价格” P_t^0 为假定资本收益预期为零时的房地产价格，Abraham and Hendershott 证明了房地产基本价格即“标准价格”可以由建造成本、雇用率、成年工作人员的真实收入和税后利率等变量来解释。假设第 t 期的房地产“标准价格”是当期房屋建造成本 c_t 、居民人均可支配收入 y_t 和银行利率 i_t 的线性函数，因此，房地产“标准价格”可表示为：

$$P_t^0 = f_1(c_t, y_t, i_t) \quad (28)$$

方程(27)中的房地产资产下一期预期资本收益 P_{t+1}^m 为房地产投机部分价格，根据房地产投资者的正反馈机制的投机心理，人们往往根据过去价格的趋势来形成对未来价格的预期，从而决定当期的交易行为，可以认为它与前一期的房地产价格增长率 G 有关，因此房地产投机部分价格 P_{t+1}^m 可表示为：

$$P_{t+1}^m = f_2(g_t) \quad (29)$$

将方程(28)和方程(29)带入价格方程(27)中。可以得到房地产投机价格构成模型

$$P_t = f_1(c_t, y_t, i_t) + \frac{f_2(g_t)}{(1+i_t)} \quad (30)$$

2. 回归分析确定房地产“标准价格”

首先确立房地产“标准价格”计算模型,确定房地产“标准价格”根据方程(30)可建立近似线性方程为:

$$P_t = \alpha_0 + \alpha_1 c_t + \alpha_2 y_t + \alpha_3 i_t + \alpha_4 \frac{g_t}{(1+i_t)} + \mu_t \quad (31)$$

这里, 房地产“标准价格”为:

$$P_t^0 = f_1(c_t, y_t, i_t) = \alpha_0 + \alpha_1 c_t + \alpha_2 y_t + \alpha_3 i_t \quad (32)$$

房地产投机部分价格为:

$$P_{t+1}^m = f_2(g_t) = \alpha_4 \frac{g_t}{(1+i_t)} \quad (33)$$

式(31)-(33)中, P_t 为第 t 期的房地产价格, c_t 为第 t 期的房屋建造成本, y_t 为第 t 期的居民人均可支配收入, i_t 为第 t 期的短期贷款利率, g_t 为第 t 期的房地产价格实际增长率, μ_t 是随机误差项。

3. 房地产行业稳定度测定

房地产行业的稳定度可以较直观地反映房地产业发展的态势。通过上一步的回归分析得出各个参数的数值, 将“标准价格” P_t^0 和 P_t 代入房地产稳定度公式:

$$K = (1 - \frac{P_t - P_t^0}{P_t^0}) * 100 \quad (34)$$

根据 K 的判断标准:

- (1) $K=(80-100\%)$, 房地产市场处于安全区;
- (2) $K=(60\%-79\%)$, 房地产市场处于警戒区;
- (3) $K=(50\%-59\%)$, 房地产市场处于危险区;
- (4) $K<50\%$, 房地产市场处于严重危险区。

6.2.2.2 房地产行业态势稳定度模型的求解

1. 指标的选取

根据国家统计局提供的资料, 选取全国 2010.7-2011.6 的相关数据, 利用上述模型对全国房地产业的稳定度进行评估。所选取的指标及数据如表 6-26 所示。

表 6-26 全国房地产业稳定度测定指标

时 间	当期房屋 建造成本	居民人均可 支配收入	短期银 行利率	房地 产价格	住宅业价格 实际增长率
Jul-10	3992.28	1525.57	5.31	4741.42	4.25
Aug-10	4191.33	1558.33	5.31	5130.15	8.2
Sep-10	4105.69	1571.33	5.31	5263.71	2.6
Oct-10	4431.1	1592	5.56	5470.49	3.93
Nov-10	3920.02	1725.78	5.56	5226.69	-4.46
Dec-11	3844.99	1855.28	5.81	4677.6	-10.51
Feb-11	4247.14	1987.33	6.06	4909.99	4.97
Mar-11	4233.26	1858.33	6.06	5168.82	5.27
Apr-11	4377.58	1749.33	6.31	5411.1	4.69
May-11	4461.11	1693	6.31	5654.13	4.49
Jun-11	4037.51	1871.58	6.31	5196.28	-8.1

2. 模型的求解

利用 SPSS 软件对以上数据进行统计分析, 将所有的变量引入回归模型, 对模型进行统计分析、方差分析、相关性分析和回归分析。分析结果如下表 6-27 至表 6-30 所示。

表6-27 统计量描述

Model Summary(b)										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.861	.742	.570	195.74031	.742	4.312	4	6	.055	2.653

表 6-27 是回归模型的一般性统计量, 包括均值、复相关系数以及预测值的标准差。回归方程的复相关系数 $R^2=0.861$, 说明得出的回归方程的紧密度较高, 各自变量选取得较恰当, 变量系数配合得较好。

表 6-28 方差分析

ANOVA(b)						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	660903.843	4	165225.961	4.312	.055(a)
	Residual	229885.619	6	38314.270		
	Total	890789.462	10			

表 6-28 是方差分析表, 从表中可以看出, 在自由度为 4 的条件下, 回归方

程的统计量 F 的值为 4.312，最后一列的“sig”说明本文选取的 4 个自变量与因变量之间完全无限性关系的概率为 0.55。虽然表中 0.55 说明文中所选取的 c_t 、 y_t 、 i_t 和 g_t 与因变量 p_t 无线性关系的概率较大，但是如采用 Stepwise 选项进行变量剔除， c_t 、 y_t 、 i_t 和 g_t 的 sig 值分别为 0.016、0.005、0.021 和 0.001。这表明在这四个自变量中，只有短期利率相对于其他变量而言，与住宅价格 p_t 的线性关系不明显。

表 6-29 相关性分析

Correlations						
		房价	造价	可支配收入	利率	实际增长率
Pearson Correlation	房价	1.000	0.747	-0.219	0.327	0.329
	造价	0.747	1.000	-0.096	0.377	0.703
	可支配收入	-0.219	-0.096	1.000	0.719	-0.397
	利率	0.327	0.377	0.719	1.000	-0.167
	实际增长率	0.329	0.703	-0.397	-0.167	1.000
Sig. (1-tailed)	房价	.	0.004	0.258	0.163	0.162
	造价	0.004	.	0.389	0.127	0.008
	可支配收入	0.258	0.389	.	0.006	0.113
	利率	0.163	0.127	0.006	.	0.312
	实际增长率	0.162	0.008	0.113	0.312	.
N	房价	11	11	11	11	11
	造价	11	11	11	11	11
	可支配收入	11	11	11	11	11
	利率	11	11	11	11	11
	实际增长率	11	11	11	11	11

表 6-29 是关于相关性的分析结果，从表中可以看出， c_t 、 y_t 、 i_t 和 g_t 与因变量 p_t 的相关性分别是 0.747、-0.219、0.327 和 0.329，各变量的显著性分别是 0.008、0.006、0.034 和 0.001。由于自变量的显著性越大，表明关于该自变量与因变量的不相关的假设越不能被否定，因而可以说明短期利率与因变量 p_t 的相关性不强。按照这个顺序排列，影响有小到大的依次是 i_t 、 c_t 、 y_t 和 g_t 。

表 6-30 回归系数表

Coefficients(a)								
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
(Constant)	66.176	2,088.065		0.032	0.976			
造价	-1.401	0.668	0.971	2.097	0.081	0.747	0.650	0.435
可支配收入	0.966	0.715	-0.497	-1.351	0.225	-0.219	-0.483	-0.280
利率	166.222	330.118	0.233	0.504	0.633	0.327	0.201	0.104
实际增长率	24.873	19.009	0.512	-1.308	0.239	0.329	-0.471	-0.271

回归系数表反映了偏回归系数，它们是回归方程的系数。在 **Standardized Coefficients** 这一列中，标准化了的偏回归系数可以真实的反映哪个变量更为重要。 i_t 的 **Beta** 值 0.005 是最小的，说明其对住宅业投机价格的影响不大，从表中的数值可以看出，住宅业价格实际增长率 g_t 为 0.512，说明人们来的价格预期在很大程度上影响着投机价格的形成。表中还列出了统计量 t 。通过以上分析，得出住宅业投机价格回归方程：

$$P_t = 66.176 - 1c_t + 0.1y_t - 0.005i_t + 0.512\frac{g_t}{1+i_t} \quad (35)$$

将前文的当期房屋建造成本 c_t 、居民人均可支配收入 y_t 和银行利率 i_t 带入 (32) 式中，求出全国房地产的标准价格 P_t^0 ，再按照公式 (34) 求出房地产业发展稳定度，以此来判断全国房地产业的发展态势，得出房地产业的发展稳定情况，下表就是经计算得出的 2010 年 7 月至 2011 年 6 月的房地产业发展稳定度。如表 6-31 所示。

表 6-31 全国房地产业发展稳定度

	2010-07	2010-08	2010-09	2010-10	2010-11	2010-12	2011-02	2011-03	2011-04	2011-05	2011-06
K	0.746	0.751	0.752	0.737	0.747	0.733	0.816	0.817	0.800	0.730	0.734

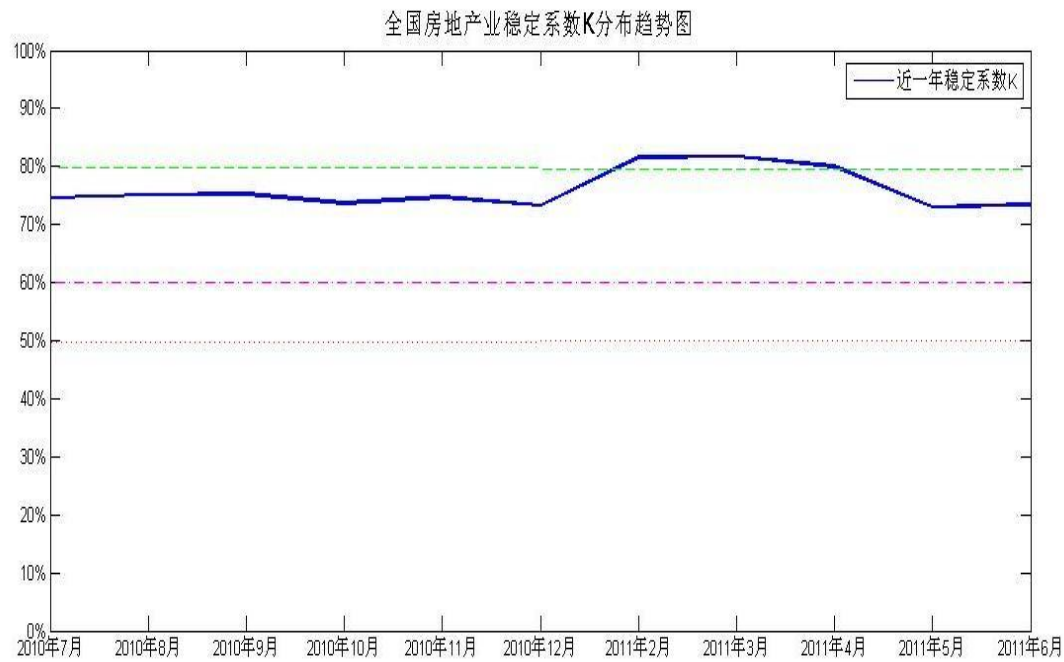


图 6-12 全国房地产行业稳定系数 K 分布趋势图

根据 K 的判断标准：

- (1) $K=(80-100\%)$ ，房地产市场处于安全区；
- (2) $K=(60-79\%)$ ，房地产市场处于警戒区；
- (3) $K=(50-59\%)$ ，房地产市场处于危险区；
- (4) $K<50\%$ ，房地产市场处于严重危险区。

由表 6-56 和图 6-12 中可以反映出，2010 年 7 月至 2010 年 12 月以及 2011 年 5 月至 2011 年 6 月这两个时间段内全国房地产行业处于警戒区；2011 年 2 月至 2011 年 4 月全国房地产行业处于安全区。总体来说，2010 年 7 月至 2011 年 6 月这段时间全国房地产行业均处于警戒区与安全区之间，由此可以看出全国房地产业发展平稳的同时也存在一定的不稳定性。

6.3 问题三的模型建立及求解

对于问题三 (⑥房地产行业可持续发展)，需要建立房地产行业可持续发展模型。运用主观和客观相结合的方法，根据房地产可持续发展评价的原则，选取国家统计局提供的天津市的相关数据，构建适合天津市发展状况的评价指标体系。在此基础上运用层次分析法(AHP)来确立各级指标的权重，构建可持续发展模型，对天津市可持续发展程度进行合理评估及预测。

6.3.1 房地产行业可持续发展模型的建立

1. 一般指标体系的建立

本文运用主观和客观相结合的方法，根据房地产可持续发展评价的原则，选取国家统计局提供的天津市的相关数据，构建适合天津市发展状况的评价指标体系。

一级指标包括四个部分：经济评价、人口评价、环境评价、资源评价。依据目前有关房地产可持续发展评价研究的报告、论文以及专家咨询等方法，结合天津市统计数据，确定了以下指标构成天津市房地产可持续发展评价体系，包括 4 个一级指标，16 个二级指标。

表 6-32 房地产行业可持续发展指标体系

可持续发展指标体系	c1. 经济评价	c11. 人均国民生产总值
		c12. 社会消费品零售总额
		c13. 固定资产投资总额
		c14. 在岗职工工资总额
		c15. 国内生产总值增速
	c2. 人口评价	c21. 年末总人口数
		c22. 人口密度
		c23. 人口自然增长率
	c3. 环境评价	c31. 城镇污水处理率
		c32. 生活垃圾无害处理率
		c33. 绿化覆盖率
	c4. 资源评价	c41. 商品房本年新开工面积
		c42. 商品房本年施工面积
		c43. 商品房本年竣工面积
		c44. 商品房本年销售面积
		c45. 房地产企业本年资产负债率

2. 数据标准化处理

数据标准化也就是统计数据的指数化。数据标准化处理主要包括数据同趋化处理和无量纲化处理两个方面。数据同趋化处理主要解决不同性质数据问题，对不同性质指标直接加总不能正确反映不同作用力的综合结果，须先考虑改变

逆指标数据性质，使所有指标对测评方案的作用力同趋化，再加总才能得出正确结果。数据无量纲化处理主要解决数据的可比性，在此我们采用指数化处理方法。指数化处理以指标的最大值和最小值的差距进行数学计算，其结果介于0—1之间。具体计算公式如下：

$$C_i = \left| \frac{(T_i - T_{\min})}{(T_{\max} - T_{\min})} \right| * 100\% \quad (36)$$

其中， C_i 是评价值， T_i 为实际观察值， T_{\max} 为实际观察值的最大值， T_{\min} 为实际观察值的最小值。

下面是天津市 2000 年到 2008 年的指标原始数据：

表 6-33 经济评价指标原始数据

代码	经济评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c11	人均 GDP	1.7	1.9	2.1	2.5	3.0	3.7	4.2	4.7	5.8
c12	消费品总额	669.8	782.3	941.3	839.4	1001.9	1112.9	1275.3	1459.2	1875.0
c13	固产总额	503.5	595.7	685.5	1017.2	1215.4	1447.6	1637.2	2236.4	3109.0
c14	工资总额	244.5	266.7	290.7	325.4	377.7	419.8	487.5	602.7	737.2
c15	生产增速	13.3	13.3	13.2	14	14.1	14.5	14.4	15.1	16.5

表 6-34 人口评价指标原始数据

代码	人口评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c21	总人口数	910.1	912	913.9	919.0	926	932.5	939.3	948.8	959.1
c22	市区人数	601.0	682.0	747.9	752.2	758.7	764.3	769.6	777.9	786.3
c23	人口密度	1386	1154	1008	1014	1022.89	1030.43	1037.48	1048.6	1062.7

表 6-35 环境评价指标原始数据

代码	环境评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c31	城镇污水处理率	37	38	40	43	44	54	62.61	73.9	78.21
c32	生活垃圾无害处理率	42	46	50	55	61	61	80.29	85.06	93.31
c33	绿化覆盖率	24	25	26.1	27.3	31.04	35.02	36	37.5	38

表 6-36 资源评价指标原始数据

代码	资源评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c41	新开工面积	617.8	569.3	825.2	838.1	1216.5	1579.4	1906.3	2114.3	2440.2
c42	施工面积	1783	1863.5	2135.6	2314.4	2865.5	3470.6	4142.6	4836.5	5704.2
c43	年竣工面积	583.5	690.5	746.4	911.3	1108.1	1479.2	1520.2	1704.4	1799.3
c44	销售面积	391.7	537.4	564	786.5	847.0	1408.4	1458.6	1548.5	1252.0
c45	资产负债率	77.4	74.2	72	73.3	73	72.8	71	70.8	71.8

运用 SPSS 软件对指标数据无量纲处理(处理中要注意数据对可持续发展是

正效应还是负效应):

表 6-37 经济评价指标标准化后数据

代码	经济评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c11	人均国民生产总值	-1.1	-1.0	-0.8	-0.5	-0.1	0.30	0.6	1.0	1.8
c12	社会消费品零售总额	-1.1	-0.8	-0.4	-0.7	-0.2	0.01	0.4	0.9	2.0
c13	固定资产投资总额	-1.0	-0.9	-0.8	-0.4	-0.2	0.07	0.3	1.0	2.0
c14	在岗职工工资总额	-1.0	-0.9	-0.7	-0.5	-0.2	0.02	0.4	1.1	1.9
c15	国内生产总值增速	-0.9	-0.9	-1.	-0.2	-0.1	0.22	0.1	0.8	2.1

表 6-38 人口评价指标标准化后数据

代码	人口评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c21	年末总人口数	-1.0	-0.9	-0.8	-0.6	-0.2	0.2	0.6	1.1	1.7
c22	市区人口数	-2.3	-0.9	0.2	0.2	0.3	0.4	0.5	0.6	0.8
c23	市区人口密度	2.4	0.5	-0.6	-0.6	-0.5	-0.4	-0.4	-0.2	-0.2

表 6-39 环境评价指标标准化后数据

代码	环境评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c31	城镇污水处理率	-0.9	-0.9	-0.8	-0.6	-0.53	0.1	0.6	1.3	1.6
c32	生活垃圾无害处理率	-1.1	-0.9	-0.7	-0.5	-0.15	-0.1	0.9	1.1	1.6
c33	绿化覆盖率	-1.2	-1.1	-0.9	-0.7	-0.01	0.7	0.9	1.1	1.2

表 6-40 资源评价指标标准化后数据

代码	资源评价指标	年份								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c41	新开工面积	-1.0	-1.1	-0.7	-0.7	-0.2	0.3	0.8	1.1	1.6
c42	施工面积	-1.0	-0.9	-0.8	-0.6	-0.3	0.1	0.6	1.1	1.7
c43	竣工面积	-1.3	-1.0	-0.9	-0.5	-0.1	0.6	0.7	1.1	1.3
c44	销售面积	-1.3	-0.9	-0.9	-0.4	-0.3	0.9	1.1	1.3	0.6
c45	资产负债率	2.2	0.6	-0.5	0.2	0.03	-0.06	-0.9	-1.0	-0.6

3. 因子分析法

因子分析法是从研究变量内部相关的依赖关系出发,把一些具有错综复杂关系的变量归结为少数几个综合因子的一种多变量统计分析方法。它的基本思想是将观测变量进行分类,将相关性较高,即联系比较紧密的分在同一类中,而不同类变量之间的相关性则较低。那么每一类变量实际上就代表了一个基本结构。即公共因子。对于所研究的问题就是试图用最少数个数的不可测的所谓公共因子的线性函数与特殊因子之和来描述原来观测的每一分量。

建立因子分析模型的目的不仅是找出主因子,更重要的是知道每个主因子的意义,以便对实际问题进行分析。如果求出主因子解后,各个主因子的典型代表变量不很突出,还需要进行因子旋转,通过适当的旋转得到比较满意的主

因子。

旋转的方法有很多，正交旋转(orthogonal rotation)和斜交旋转(oblique rotation)是因子旋转的两类方法。最常用的方法是最大方差正交旋转法(Varimax)。进行因子旋转，就是要使因子载荷矩阵中因子载荷的平方值向 0 和 1 两个方向分化，使大的载荷更大，小的载荷更小。因子旋转过程中，如果因子对应轴相互正交，则称为正交旋转；如果因子对应轴相互间不是正交的，则称为斜交旋转。常用的斜交旋转方法有 Promax 法等。

4. 因子得分

因子分析模型建立后，还有一个重要的作用是应用因子分析模型去评价每个样品在整个模型中的地位，即进行综合评价。这时需要将公共因子用变量的线性组合来表示，也即由地区经济的各项指标值来估计它的因子得分。

5. 指标的权重确定

运用层次分析法(AHP)来确立各级指标的权重。层次分析法(AHP)是将决策总是有关的元素分解成目标、准则、方案等层次，在此基础上进行定性和定量分析的决策方法。该方法是美国运筹学家匹茨堡大学教授萨蒂于本世纪 70 年代初，在为美国国防部研究‘根据各个工业部门对国家福利的贡献大小而进行电力分配’课题时，应用网络系统理论和多目标综合评价方法，提出的一种层次权重决策分析方法。这种方法的特点是在对复杂的决策问题的本质、影响因素及其内在关系等进行深入分析的基础上，利用较少的定量信息使决策的思维过程数学化，从而为多目标、多准则或无结构特性的复杂决策问题提供简便的决策方法。尤其适合于对决策结果难于直接准确计量的场合。用层次分析法确立权重方法如下：

1) 建立系统的递阶层次结构

在这一个步骤中，要求将问题所含的要素进行分组，把每一组作为一个层次，并将它按照：最高层(目标层)——若干中间层(准则层)——最低层(属性层)的次序排列起来。

2) 建立判断矩阵

这一个步骤是 AHP 决策分析中一个关键的步骤。就判断矩阵表示针对上一层次中的元素而言，评定该层次中各有关元素相对重要性程度的判断，其形式如下：

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

关于如何确定 b_{ij} 的值，一般引用数字 1-9 及其倒数作为标度。

3) 层次单排序

即在层次分析中由单一判断矩阵计算元素之间相对重要性权重。层次单排序是通过解以下特征值问题得到的

$$BW = \lambda_{\max} W$$

(其中， λ_{\max} 为判断矩阵 B 的最大特征根， W 对应于 λ_{\max} 的正规化特征向量， W 的分量 W_i 就是对应元素单排序的权重值)。

4) 检验判断矩阵的一致性

(1)计算一致性指标 CI, $CI = \frac{\lambda_{\max} - n}{n - 1}$ 。

当 CI=0 时, 判断矩阵具有完全一致性; 反之, CI 愈大, 就表示判断矩阵的一致性就越差。为了检验判断矩阵是否具有令人满意的一致性, 需要将 CI 与平均随机一致性指标 RI 进行比较。

(2)查找相应的平均随机一致性指标 RI。

(3)计算一致性比例 CR, $CR = \frac{CI}{RI}$ 。

5) 合成权重的计算

6) 层次总排序的一致性检验。

6.3.2 房地产行业可持续发展模型的求解

1. 指标的 Bartlett 球性检验

Bartlett 估计因子得分可由最小二乘法或极大似然法导出。在采用因子分析法对变量进行分析以前, 需对变量进行 Bartlett 球性检验, 计算 KMO(Kaiser—Meyer—Olkin Measure of Sampling Adequacy)值的检验, 通过对这些统计量的综合考虑, 判断是否应对变量进行因子分析。使用 SPSS 软件检验如下;

表 6-41 经济因子 Bartlett 球性检验

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.701
Bartlett's Test of Sphericity	Approx. Chi-Square	102.720
	df	10
	Sig.	.000

表 6-42 人口因子 Bartlett 球性检验

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.743
Bartlett's Test of Sphericity	Approx. Chi-Square	39.305
	df	3
	Sig.	.000

表 6-43 环境因子 Bartlett 球性检验

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.767
Bartlett's Test of Sphericity	Approx. Chi-Square	32.536
	df	3
	Sig.	.000

表 6-44 资源因子 Bartlett 球性检验

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.777
Bartlett's Test of Sphericity	Approx. Chi-Square	65.669
	df	10
	Sig.	.000

2. 选取二级指标

通过 SPSS 软件对每组因子进行反旋转，得到反旋转后的载荷矩阵，如下表：

表 6-45 经济因素旋转后的载荷矩阵

Rotated Component Matrix(a)				
	Component			
	1	2	3	4
c11	.571	.594	.566	.019
c12	.545	.739	.396	.008
c13	.663	.612	.426	.069
c14	.599	.649	.458	.103
c15	.763	.515	.391	.018

表 6-46 人口因素旋转后的载荷矩阵

Rotated Component Matrix(a)		
	Component	
	1	2
c21	.250	-.168
c22	.859	.512
c23	-.986	.968

表 6-47 环境因素旋转后的载荷矩阵

Rotated Component Matrix(a)		
	Component	
	1	2
c31	.815	.569
c32	.806	.581
c33	.578	.816

表 6-48 资源因素旋转后的载荷矩阵

Rotated Component Matrix(a)				
	Component			
	1	2	3	4
c41	.851	.356	.382	-.029
c42	.889	.340	.304	.000
c43	.772	.375	.507	.078
c44	.558	.411	.721	-.009
c45	-.342	-.900	-.270	-.004

经济因素这一变量包含了 5 个测量项目，计划取贡献率的前 4 个作为特征因子。从表 6-45 矩阵可以看出，编号为 c15 的测量项目对特征因子 1 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 1 命名为“国内生产总值增速”；编号为 c12 的测量项目对特征因子 2 的作用比较明显，且明显前强于其他测项目，因此可以把特征因子 2 命名为“社会消费品零售总额”；编号 c11 的测量项目对特征因子 3 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 3 命名为“人均国民生产总值”；编号 c14 的测量项目对特征因子 4 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 4 命名为“在岗职工工资总额”。

人口因素这一变量包含了 3 个测量项目。计划取贡献率的前 2 个作为特征因子。从表 6-46 矩阵可以看出，编号为 c22 的测量项目对特征因子 1 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 1 命名为“人口数”；编号为 C23 的测量项目对特征因子 2 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 2 命名为“人口密度”。

环境因素这一变量包含了 3 个测量项目，计划取贡献率的前 2 个作为特征因子。从表 6-47 矩阵可以看出，编号为 c31 的测量项目对特征因子 1 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 1 命名为“城镇污水处理率”；编号为 c33 的测量项目对特征因子 2 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 2 命名为“绿化覆盖率”。

资源因素这一变量包含了 5 个测量项目，计划取贡献率的前 3 个作为特征因子，从表 6-48 矩阵可以看出，编号为 c42 的测量项目对特征因子 1 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 1 命名为“商品房本年施工面积”；编号为 c44 的测量项目对特征因子 2 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 2 命名为“商品房本年销售面积”；编号 c43 的测量项目对特征因子 3 的作用比较明显，且明显前强于其他测量项目，因此可以把特征因子 3 命名为“商品房本年竣工面积”。

通过筛选，最后得到 4 个一级指标，11 个二级指标：

表 6-49 筛选后指标体系：

可持续发展指标体系	c1. 经济评价	c11. 人均国民生产总值
		c12. 社会消费品零售总额
		c14. 在岗职工工资总额
		c15. 国内生产总值增速
	c2. 人口评价	c22. 人口数（市区）
		c23. 人口密度（市区）
	c3. 环境评价	c31. 城镇污水处理率
		c33. 绿化覆盖率
	c4. 资源评价	c42. 商品房本年施工面积
		c43. 商品房本年竣工面积
		c44. 商品房本年销售面积

3. 确立权重及赋权后数据处理

采用 Delphi 法，通过专家数轮打分，得到一致分数后，建立矩阵，得到其权重，赋权后得到数据如下表所示。

表 6-50 赋权后经济指标指数

代码	权重	经济评价指标	赋权后标准化指数								
			2000	2001	2002	2003	2004	2005	2006	2007	2008
c11	0.4667	人均 GDP	-0.52	-0.46	-0.39	-0.25	-0.09	0.14	0.28	0.47	0.82
c12	0.1210	零售总额	-0.13	-0.10	-0.05	-0.08	-0.03	0.00	0.05	0.11	0.24
c14	0.1805	工资总额	-0.18	-0.16	-0.13	-0.09	-0.04	0.00	0.07	0.20	0.34
c15	0.2318	总值增速	-0.21	-0.21	-0.23	-0.05	-0.03	0.05	0.02	0.18	0.49
c1	1	经济指标	-1.04	-0.93	-0.80	-0.47	-0.19	0.19	0.42	0.96	1.89

表 6-51 赋权后人口指标指数

代码	权重	人口评价指标	赋权后标准化指数								
			2000	2001	2002	2003	2004	2005	2006	2007	2008
c22	0.5987	人口数	-1.37	-0.56	0.10	0.14	0.21	0.26	0.32	0.40	0.48
c23	0.4013	人口密度	0.99	0.22	-0.25	-0.23	-0.20	-0.18	-0.15	-0.12	-0.07
c2	1	人口指标	-0.38	-0.34	-0.15	-0.09	0.01	0.08	0.17	0.28	0.41

表 6-52 赋权后环境指标指数

代码	权重	环境评价指标	赋权后标准化指数								
			2000	2001	2002	2003	2004	2005	2006	2007	2008
c31	0.5287	污水处理率	-0.96	-0.90	-0.77	-0.58	-0.52	0.10	0.65	1.36	1.64
c32	0.4713	绿化覆盖率	-1.25	-1.08	-0.88	-0.67	-0.01	0.69	0.86	1.13	1.22
c3	1	环境指标	-2.21	-1.98	-1.65	-1.25	-0.53	0.79	1.51	2.49	2.86

表 6-53 赋权后资源指标指数

代码	权重	资源评价指标	赋权后标准化指数								
			2000	2001	2002	2003	2004	2005	2006	2007	2008
c42	0.3147	施工面积	-0.32	-0.30	-0.24	-0.20	-0.08	0.05	0.20	0.35	0.55
c43	0.3995	竣工面积	-0.50	-0.41	-0.36	-0.22	-0.05	0.26	0.30	0.45	0.54
c44	0.2858	销售面积	-0.37	-0.28	-0.26	-0.12	-0.08	0.27	0.30	0.36	0.17
c4	1	资源指标	-1.19	-0.99	-0.86	-0.54	-0.21	0.58	0.80	1.16	1.26

经过统计，最后汇总，得到下表：

表 6-54 汇总指标指数

指标	权重	赋权后标准化指数								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c1 经济指标	0.141	-0.15	-0.13	-0.11	-0.07	-0.02	0.02	0.06	0.13	0.26
c2 人口指标	0.578	-0.21	-0.19	-0.08	-0.05	0.00	0.05	0.09	0.16	0.24
c3 环境指标	0.138	-0.15	-0.13	-0.11	-0.08	-0.03	0.06	0.10	0.17	0.19
c4 资源指标	0.143	-0.17	-0.14	-0.12	-0.07	-0.03	0.08	0.11	0.16	0.18
房地产业可持续发展指标	1	-0.68	-0.59	-0.42	-0.27	-0.08	0.21	0.36	0.62	0.87

4. 评价结果分析

为了对指标的判断更清晰明了，根据可持续发展度相对应的取值范围进行度量，对所有指数进行合理调整，使新指标 $C=(原指标+1)*50$ ，得到下表：

表 6-55 调整后指标指数

指标	权重	赋权后标准化指数								
		2000	2001	2002	2003	2004	2005	2006	2007	2008
c1 经济指标	0.141	42.4	43.3	44.2	46.4	48.5	51.4	53.1	56.8	63.4
c2 人口指标	0.578	39.0	40.3	45.5	47.3	50.1	52.5	54.7	58.2	62.0
c3 环境指标	0.138	42.1	43.0	44.1	45.5	48.3	53.0	55.3	58.5	59.6
c4 资源指标	0.143	41.3	42.8	43.7	46.0	48.4	54.2	55.8	58.4	59.0
房地产业可持续发展指标	1	14.8	19.4	27.5	35.2	45.3	61.1	68.9	81.9	94.0

对照天津市企业家信心指数和房地产景气指数，参考专家意见，对房地产业可持续发展进行对应取值：

表 6-56

可持续发展度	指标指数
可持续发展度低	40 以下
可持续发展度中	40-80
可持续发展度高	80 以上

根据表做出图 6-13，可以看出：

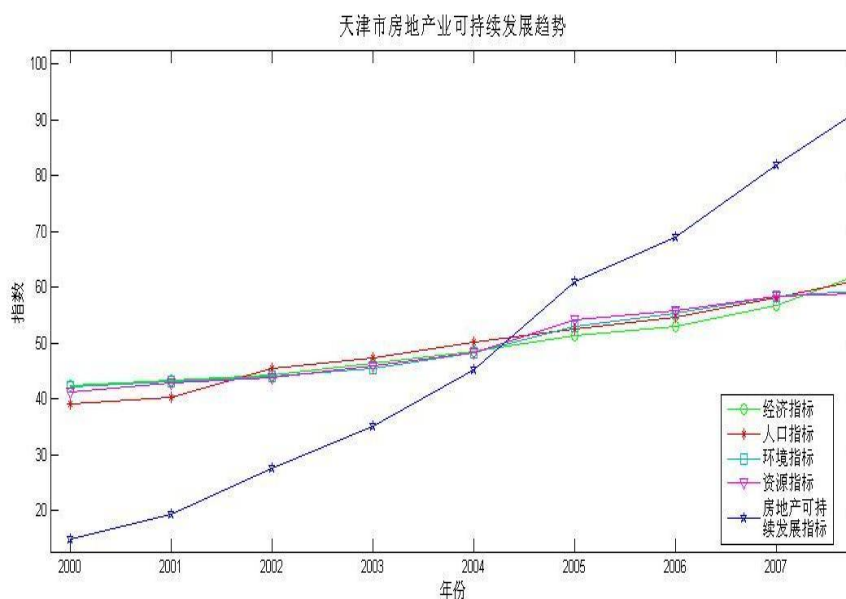


图 6-13 天津市房地产业可持续发展趋势

通过上述模型的建立及求解，得出如下结论：

- 1) 经济指标评价天津市经济呈现平稳增长，人均国民生产总值逐年升高，各指标数据呈现上涨趋势。随着人民生活水平的提高，对商品房的需求，包括第一套自住房和改善性住房的需求都保持旺盛。图中各项指标的得分较准确反应了目前的经济走势对于房地产可持续发展的支持。
- 2) 人口指标评价天津市人口保持稳定增长，而且随着经济的发展，天津市目前的城市化率还能进一步提高，从而保证对房地产业的需求。另外目前天津市的人口密度还没有达到很高的水平，城市接收人口的能力还比较强，完全能保证天津市房地产业的可持续发展。
- 3) 环境指标评价天津市作为一个旅游城市，空气、水质和绿化水平都非常好，而且总体指标表现平稳，利于可持续发展的要求。
- 4) 资源指标评价天津市 2000 年到 2008 年房地产一直处于一个高速发展的时期，房地产行业资源充沛，对于房地产业的可持续发展提供了良好的保障。

总体来看，天津市房地产业的可持续发展能力良好，房地产业现在正处于一个历史的最佳时期，各方面都基本支持着天津市的房地产业可持续发展。

7. 模型的评价、改进与推广

本文紧密围绕我国房地产问题，从当前经济社会发展现实出发，运用了全国与天津市的年度数据与月度数据，由面到点，针对问题中的六种房地产模型分别建立了适宜的数学模型，对其进行详细的阐述、分析、推导计算，最后得出未来五年内房地产的供需面积、价格趋势走向、房地产与其他行业关系以及全国房地产发展稳定度和可持续发展度等指标，较好地反映出房地产行业当前的态势、未来的趋势及经济调控策略的成效。

在房地产模型的改进：

针对问题一 (①住房的需求、②供给与③价格)，我们在数据上使用了商品房的销售面积以及竣工面积来衡量住房的需求量与供给量，而实际上商品房与住房存在某种包含与被包含关系，可能会造成一定程度上结果集的偏差，下一步改进是寻找一种更好的指标来衡量房地产住房的供需关系。

针对问题二 (④房地产行业与国民经济其他行业关系、⑤房地产行业态势分析)，在对房地产行业态势分析时，没有将其影响因子的反馈关系考虑在内，这在一定程度上会对指标的精确性造成影响，下一步目标是引入反馈机制，改进房地产发展的稳定性指标。

针对问题三 (⑥房地产行业可持续发展)，我们选取了天津市为例，使用层次分析法分析该市房地产行业可持续发展水平。层次分析法一般使用 Delphi 法来衡量经济、人口、环境、资源等因素的权重，一般 Delphi 法需要依赖房地产专家的专业知识，而我们仅仅使用互联网的搜索结果以及主观猜测，确定了参数的权重，可能会对可持续发展的指标准确性造成一定的影响，下一步工作是寻找一个更加准确的权值关系。

由于时间仓促，本文在模型建立上还存在一些问题，比如，对于问题三，我们仅仅对天津市可持续发展水平进行了研究，而没对其他城市数据进行下一步分析，这为我们未来的房地产问题研究提供了一个方向。

8. 参考文献

- [1] 李启明等, 房地产投资风险与决策[M], 南京: 东南大学出版社, 1998。
- [2] 曹云, 徐卫亚, 系统工程评估方法的进展[J], 中国工程科学, 7(6): 88-94, 2005。
- [3] 齐雪琴, 姜照华, 姜朝妮, 胡晓玮, 我国房地产价格影响因素分析[J], 商场现代化, 10:121-123, 2010。
- [4] 王志强, 北京市房地产价格影响因素浅析从实证中看房地产政策[J], 商场现代化, 1: 88-90, 2010。
- [5] 李熙娟, 李斌, 房地产业与国民经济增长的实证研究[J], 商业研究, 4: 201-204, 2006。
- [6] 夏丹, 我国房地产业发展与经济增长关系的实证分析[J], 重庆工学院学报社会科学版, 10: 41-43, 2007。
- [7] 郑慧开, 论灰色预测在房地产开发管理中的应用[J], 经济师, 206(4): 67-72, 2002。
- [8] 廖英敏, 房地产投资供给需求回落价格继续上涨[J], 中国房地信息, 10: 27-32, 2004。
- [9] 姚翠友, 基于 GM(1, 1)模型的北京市房地产投资分析[J], 工业技术经济, 7: 69-71, 2007。
- [10] 杨贵中, 邓学芬, 成都市房价影响因素的回归分析与事后模拟[J], 价值工程, 4: 46-50, 2007。
- [11] 张吉军, 区间数多指标决策问题的模糊层次分析法[J], 工业工程与管理, 6: 20-23, 2003。
- [12] 闫天兵, 沈丽, 基于因子分析法的中外商业银行盈利能力比较研究[J], 金融发展研究 4: 36-40, 2009。
- [13] 周京奎, 房地产价格波动与投机行为[J], 当代经济科学, 7: 19-24, 2005。