

基于 SPCA 的房地产发展指数 的建立及应用

基于 SPCA 的房地产发展指数的建立及应用

摘要

房地产行业近年发展迅猛，已经成为国民经济中的支柱行业之一。随着行业的发展，房价一路走高，房价成为了人们日常谈论的话题之一。政府也在想方设法控制房价的上升以使大部分人都能住上自己的房子。而房地产市场本身非常复杂，这就需要一种方法来分析房地产市场，为政府宏观调控提供依据。

本文采用稀疏主成分分析（SPCA）——一种主成分分析的改进版，对房地产市场发展的综合评价进行分析，得到了北京房地产市场的发展状况。从而我们能够通过几个关键的指标比如房地产投资金额、年房屋竣工面积和年 GDP 等来评价房地产市场的发展好坏。验证了这些指标的合理性之后，我们结合 VaR 方法提出一个房地产市场发展指数，通过这个指数我们预测了未来几年宏观政策改变不大时房地产市场的发展走向。并考虑了这个指数与 CPI 之间的联系，以及 CPI 对房地产市场发展的影响。

关键词：

房产发展 SPCA CPI 评价

Abstract

Nowadays, real estate industry has become an essential part of GDP. As a result, researches in the very field are becoming more and more important. However, data in the real estate industry are complex and then we need a specific approach towards the field and to provide proofs for the government to make some certain decisions.

We use SPCA(sparse Principal Component Analysis) which is an advanced method of PCA to analyze the development of real estate industry of Beijing. During the research, we find some certain indexes to access the development and finally use an index(maybe CPI or something) to make a decision.

Key words:

real estate industry development , SPCA, access, CPI

目录

0. 背景介绍	5
➤ 课题背景.....	5
➤ 方法背景.....	5
1. 方法步骤	7
➤ 数据展示.....	7
➤ 数据处理.....	14
➤ 结果展示.....	17
➤ 房地产发展指数的构造.....	18
➤ 可视化:	21
2. 结果解释及分析:	21
➤ 住宅相关综合指标.....	21
➤ 房地产开发企业相关综合指标.....	21
➤ 政策相关综合指标.....	22
➤ 居民相关综合指标.....	22
➤ 结果分析.....	22
3. 初步预测	23
➤ 趋势图展示及分析.....	23
➤ 预测方法应用.....	24
➤ 初步预测.....	24
➤ 合理性分析.....	24
4. 优势、不足与展望	27
➤ 优势、创新.....	27
➤ 不足之处.....	27
➤ 展望	28
5. 结语	28
6. 参考文献	29

背景介绍

➤ 课题背景

房地产业是我国近十年发展最快、关联度最大、对 GDP 增长贡献最大的行业。并且在未来几年里，房产依然会起到至关重要的地位。然而在房产飞速发展的今天，同样也暴露许多问题。面对现当今房产投资飞速上涨，投资者实际上很难了解投资的房地产项目到底风险几何，针对如此现象，我们决定建立一个房产投资风险指数来预估房产的发展与涨跌。

现当今社会有许多投资者青睐于房产这一项目，然而从某一方面来讲，许多人都只是看重了房产的背后巨大的资金流动力。与股市类似，房产项目也隐藏了巨大的资金波动。针对这一现象，已有许多学者开始了房产方面的研究和统计调查。流行的方法大多是通过模糊统计方法来进行数据的处理与预估，例如：建立模糊矩阵并综合排序来进行数据的整合，使用层次分析法分析指标的权重。然而这些理论方法却只能在模糊程度上对房产风险进行预估，并且模糊矩阵的计算量较大，也增大了统计数据的难度。

针对如此现象，我们通过采用金融统计中的 VaR 模型来对房产风险进行代入评估，VaR 模型是上世纪 90 年代由金融动荡而产生的市场风险测量方法，先被广泛应用于许多金融系统与各大管理系统中以测控各大管理系统的风险值。同时由于房产的多重相关性与高维性，我们通过使用稀疏主成分分析(SPCA)对其降维处理从而方便数据的统计以及模型的代入处理，稀疏主成分分析(SPCA)是 2006 年提出的一个新型降维理论，这一点也是在房产数据的统计中最大的创新点。

通过运用新型的降维方法与金融系统中的权威模型 VaR 模型，可以突破房产风险统计中传统模糊不清的束缚，可以大幅改善原有房产风险系数的精确性。

➤ 方法背景

❖ SPCA 方法

SPCA 是一种较新的降维方法，第一篇文献发表于 2006 年。它处理面对的就是前面搜集的数据。因为数据比较复杂，既是高维（指变量维数较大），又存在多重共线性（指数据之间是相关的，就像线性代数里的线性相关）。主成分分析是对这类数据降维的好办法。稀疏主成分 SPCA 改进了原来的 PCA 主成分方法，可以获得更有效的降维效果，而且使得结

果更容易解释。而房产方面数据正是具有这种高维的特点。

稀疏主成分分析 (Sparse Principal Component Analysis, 简称 SPCA) 的主要思想是: 寻找一个已经做过主成分分析的回归优化框架。而且, 回归框架应该通过应用 Lasso (elastic net) 惩罚而允许直接修正, 这样就可以得到稀疏的得分。这样的稀疏主成分不仅达到了有效降维的目的, 而且很大程度上简化了主成分的解释, 使提取的主成分往往能对应于某些实际的含义, 便于后续应用。由于该方法使用 Lasso, 因此, 所有对于 Lasso 的改进都可以直接运用于稀疏主成分上, 极大推进了稀疏主成分的研究进程, 目前已取得不少理论成果。一个典型算法如下:

第一步, 计算一般主成分的前 k 个主成分对应的向量 α_j ($j=1, 2, \dots, k$)。

第二步, 在给定 $A=(\alpha_1, \dots, \alpha_j, \dots, \alpha_k)$ 的情况下分别解如下的回归问题:

$$\begin{aligned} \hat{\beta}_j &= \arg \min_{\beta} (\alpha_j - \beta)^T X^T X (\alpha_j - \beta) + \lambda \|\beta\|^2 + \lambda_{1,j} \|\beta\|_1 \\ \text{s.t. } \alpha^T \alpha &= I_k \end{aligned}$$

其中 $\|\beta\|^2 = \beta^T \beta$, $\|\beta\|_1$ 表示 β 各分量的绝对值之和。 $\lambda, \lambda_{1,j}$ 是参数, 对它们给定一系列值后, 相应地计算出 $\hat{\beta}_j$, 再回头确定它们的最优值。

第三步, 对于给定的 $B=(\hat{\beta}_1, \dots, \hat{\beta}_j, \dots, \hat{\beta}_k)$, 对矩阵进行奇异值分解:

$$X^T X B = U D V^T$$

其中 $U U^T = I, V V^T = I$, 并且令 $A = U V^T$ 。

第四步, 重复步骤 2、3 至收敛。

$$\hat{V}_j = \frac{\beta_j}{\|\beta_j\|} \quad (j=1, 2, \dots, k)$$

第五步,

从以上算法可以看出, 稀疏主成分是在主成分的基础上, 牺牲主成分的部分方差贡献率, 更进一步的对载荷进行稀疏, 从而更清晰地揭示出数据的内在结构, 显示出比主成分分析更强大的优势。

❖ VaR 模型

我们在过程中, 通过现有的数据, 我们将对已有数据进行整合, 同时通过房产的各个影响与调控因素 (如综合考虑国家相关政策等因素) 来设定一个房产风险指数值来设定风险度, 并通过建立金融系统中的权威模型 VaR 模型来进行数据的处理。VaR 模型中也融合了金融

方面的国家政策等因素，我们将会把他们用于房产投资风险分析。我们将采用方差-协方差法来进行数据的计算与编排，并将风险指数代入与实际数据进行统计，从而得出一个风险指数值。同时，由于房产信息的多重相关性和多维的特点，我们将在 VaR 模型中的方差-协方差法的基础上运用稀疏主成分分析(SPCA)对房产信息进行降维处理，从而方便了数据的计算，同时也可以更适应 VaR 模型的建立与应用。

0. 方法步骤

➤ 数据展示

在确立了与房地产发展的相关因素后，我组成员积极查找数据。最终在中国北京统计年鉴网上查询到所需数据 (http://www.bjstats.gov.cn/nj/main/****-tjnj/index.htm)

为方便统计，利于最终解释，我们将数据分为如下五个方面：

❖ 北京住宅相关数据：

住宅相关							
	住宅开发 投资额(亿 元)	住宅销售 额(亿元)	住宅施工面 积(万平方 米)	住宅竣工 面积(万平 方米)	中央竣工 面积(万平 方米)	地方竣工 面积(万平 方米)	住宅销售 面积(万平 方米)
1990	12.3	18.5	1561.9	573.0	637.4	443.8	152.0
1991	14.0	21.6	1612.0	601.8	396.2	640.2	152.5
1992	20.0	24.7	1784.7	681.2	399.3	712.1	153.0
1993	38.1	41.0	1866.4	654.8	320.5	837.5	182.0
1994	50.3	40.8	2315.1	832.1	366.8	1003.9	149.0
1995	142.4	63.1	2897.6	937.3	386.4	1143.8	180.0
1996	124.9	70.8	2696.9	870.4	452.0	1065.5	183.1
1997	132.9	140.4	2881.3	996.8	492.2	1133.5	256.2
1998	168.0	179.8	3473.7	1093.1	508.4	1313.1	377.0
1999	236.6	232.0	3754.8	1519.9	655.3	1666.1	484.7

2000	288.3	409.3	4083.3	1499.7	618.8	1739.4	898.2
2001	464.2	531.7	5226.4	1804.9	490.6	2064.0	1127.5
2002	586.7	716.7	6193.3	2191.4	441.8	2680.0	1604.4
2003	633.0	789.2	7011.3	2322.3	242.1	2980.7	1771.1
2004	776.0	1085.1	7513.1	2649.5	301.9	3901.3	2285.8
2005	779.5	1501.8	8043.2	3024.1	256.6	4422.6	2566.0
2006	863.6	1626.3	7113.0	2391.6	388.6	3802.4	2205.0
2007	991.7	1846.0	6788.8	2098.0	399.6	3466.8	1731.5
2008	940.6	1201.4	6656.3	1871.1	496.8	3343.9	1031.4
2009	906.6	2486.8	7058.4	2369.6	388.4	3864.2	1880.5
2010	1509.0	2060.5	7932.9	2263.5	329.2	3579.1	1201.4
2011	1778.3	1606.0	8817.1	2121.8	425.5	3607.4	1035.0

表 1

❖ 房地产开发企业相关数据：

房地产开发企业相关					
	房地产开发企业个数（个）	房地产开发企业(单位)资金来源（亿元）			
		国内贷款	利用外资	自筹资金	其他资金来源
1998	585	120.63	6.36	77.24	141.56
1999	716	127.32	4.51	88.99	161.20
2000	893	158.65	2.97	106.84	220.28
2001	1142	193.83	2.39	144.54	286.79
2002	1508	254.32	2.76	181.24	360.96
2003	1546	359.46	2.99	249.56	477.08
2004	2704	361.77	4.01	344.66	669.02

2005	3123	448.79	4.53	463.32	798.64
2006	2882	613.60	7.04	568.99	998.64
2007	2688	803.59	11.27	779.16	1410.20
2008	3433	871.17	12.81	1013.42	1248.04
2009	3171	1301.72	8.43	1187.95	2188.18
2010	3190	1439.08	13.90	1762.97	2574.66
2011	3069	1576.43	15.00	2866.13	2961.14

房地产开发企业相关（续）					
房地产开发企业(单位)资产（亿元）					
	资产总计	本年折旧	负债总计	所有者权益	资产负债率 (%)
1998	2280.94	4.96	1842.33	447.19	80.80
1999	2189.67	5.07	1768.75	429.18	80.80
2000	2942.09	7.34	2360.02	589.42	80.27
2001	3337.02	8.58	2658.07	683.02	79.63
2002	3859.92	10.40	3070.86	792.92	79.52
2003	4729.41	12.32	3806.68	937.49	80.48
2004	7217.88	17.71	5677.26	1533.02	78.68
2005	8433.27	20.02	6512.67	1884.28	77.24
2006	10326.17	24.38	8119.24	2195.41	78.64
2007	12975.55	29.47	10252.52	2719.97	79.03
2008	16918.67	43.28	12993.22	3836.12	76.76
2009	19880.00	40.77	15505.55	4323.67	78.04
2010	26221.03	48.33	20745.22	5475.77	79.10
2011	29395.96	53.37			

表 2

❖ 买房居民相关数据：

买房居民相关						
	收入（元）		消费（元）			
	家庭总收入	可支配收入	居民消费水平（元）	消费性支出	居民消费价格指数（上年=100）	家庭恩格尔系数（%）
1991	2359.9	2040.4	2071	1860.2	111.9	54.7
1992	2813.1	2363.7	2375	2134.7	109.9	52.8
1993	3935.4	3296.0	3206	2939.6	119.0	47.8
1994	5585.9	4731.2	4508	4134.1	124.9	46.4
1995	6748.7	5868.4	5663	5019.8	117.3	48.5
1996	7945.8	6885.5	6497	5729.5	111.6	46.6
1997	8741.7	7813.1	7418	6531.8	105.3	43.7
1998	10098.2	8472.0	7954	6970.8	102.4	41.1
1999	10654.8	9182.8	8598	7498.5	100.6	39.5
2000	12560.3	10349.7	8820	8493.5	103.5	36.3
2001	13768.8	11577.8	9057	8922.7	103.1	36.2
2002	13253.3	12463.9	10882	10285.8	98.2	33.8
2003	14959.3	13882.6	12014	11123.8	100.2	31.7
2004	17116.5	15637.8	13425	12200.4	101.0	32.2
2005	19533.3	17653.0	14662	13244.2	101.5	31.8
2006	22417.0	19978.0	16487	14825.0	100.9	30.8
2007	24576.0	21989.0	18553	15330.0	102.4	32.2
2008	27678.0	24725.0	20113	16460.0	105.1	33.8
2009	30674.0	26738.0	22023	17893.0	98.5	33.2

2010	33360.0	29073.0	24982	19934.0	102.4	32.1
2011	37124.0	32903.0	27760	21984.0	105.6	31.4

买房居民相关（续）				
	人均住房		常住人口（万人）	就业者负担人数 （人）
	人均住房建筑面积 （平方米）	人均住房消费支出 （元）		
1991	5.3	60.498	1094.0	1.47
1992	6.5	73.253	1102.0	1.43
1993	7.8	95.422	1112.0	1.42
1994	9.0	188.08	1125.0	1.41
1995	10.2	264.55	1251.1	1.41
1996	11.5	375.46	1259.4	1.41
1997	12.7	492.88	1240.0	1.43
1998	13.9	563.62	1245.6	1.40
1999	15.2	633.76	1257.2	1.41
2000	16.4	808.53	1363.6	1.41
2001	17.6	913.95	1385.1	1.39
2002	19.2	1049.21	1423.2	1.41
2003	19.7	1209.27	1456.4	1.39
2004	21.5	1354.23	1492.7	1.44
2005	22.0	2248.02	1538.0	1.39
2006	23.7	2615.26	1601.0	1.40
2007	24.8	3039.03	1676.0	1.40
2008	26.9	3385.6	1771.0	1.40
2009	27.7	3821.5	1860.0	1.40

2010	28.9	4648.2	1961.9	1.40
2011	29.4	5474.9	2018.6	1.50

表 3

❖ 政府政策相关数据：

政府政策相关							
	地区生产 总值比上年 增长 (%)	全社会固 定资产投 资比上年 增长 (%)	全社会固 定资产投 资 (亿元)	房地产开 发投资 (亿 元)	本年完成 的土地开 发面积 (万 平方米)	土地开 发投资 (亿 元)	土地购置 费 (亿元)
1991	9.9	7.1	192.0	24.0	305.9	4.6	2.5
1992	11.3	38.5	266.0	33.7	590.0	4.6	2.5
1993	12.3	54.3	410.4	58.4	182.5	4.6	2.5
1994	13.7	58.1	648.8	99.5	382.5	11.6	4.0
1995	12.0	29.7	841.5	352.8	709.0	58.7	52.4
1996	9.0	4.2	876.9	328.2	239.8	33.0	15.0
1997	10.1	9.6	961.2	330.3	419.8	22.2	23.9
1998	9.5	20.2	1155.6	377.4	195.0	22.7	28.4
1999	10.9	1.3	1170.6	421.5	187.2	24.6	36.6
2000	11.8	10.8	1297.4	522.1	460.1	29.3	57.6
2001	11.7	18.0	1530.5	783.8	1162.7	38.1	115.6
2002	11.5	18.5	1814.3	989.4	925.0	44.0	149.2
2003	11.1	18.9	2157.1	1202.5	1088.2	39.5	213.2
2004	14.1	17.2	2528.3	1473.3	634.2	36.5	275.8
2005	12.1	11.8	2827.2	1525.0	314.2	19.4	239.8
2006	13.0	19.3	3371.5	1719.9	840.5	35.7	477.9
2007	14.5	17.6	3966.6	1995.8	248.7	25.6	644.7
2008	9.1	-3.0	3848.5	1908.7	351.5	24.0	639.0
2009	10.2	26.2	4858.4	2337.7	364.0	28.7	587.7

2010	10.3	13.1	5493.5	2901.1	376.5	11.5	1292.7
2011	8.1	13.3	5910.6	3036.3	389	4.9	1301.2

表 4

❖ 北京土地利用情况：

土地面积及利用状况(2008 年)					
单位：平方公里					
地 区	土地面积	农用地		建设用地	未利用地
			#耕 地		
全 市	16410.54	10959.81	2316.88	3377.15	2073.58
东 城 区	25.34	0.00	0.00	25.34	0.00
西 城 区	31.62	0.00	0.00	31.62	0.00
崇 文 区	16.52	0.00	0.00	16.52	0.00
宣 武 区	18.91	0.00	0.00	18.91	0.00
朝 阳 区	455.08	137.25	47.26	308.93	8.89
丰 台 区	305.80	78.75	31.61	203.26	23.80
石 景 山 区	84.32	31.79	2.14	49.44	3.10
海 淀 区	430.73	194.89	26.90	229.24	6.60
门 头 沟 区	1450.70	1091.28	18.17	95.35	264.07
房 山 区	1989.54	1159.06	282.78	350.55	479.94
通 州 区	906.28	561.87	350.35	309.02	35.39
顺 义 区	1019.89	603.97	310.31	328.24	87.69
昌 平 区	1343.54	921.67	117.76	365.20	56.68
大 兴 区	1036.32	667.75	381.17	311.94	56.63
怀 柔 区	2122.62	1568.02	97.51	133.37	421.23
平 谷 区	950.13	713.84	123.67	126.98	109.31
密 云 县	2229.45	1540.66	229.33	329.52	359.27
延 庆 县	1993.75	1689.03	297.92	143.73	160.99

表 5

➤ 数据处理

鉴于影响因素多且复杂，为了更明确地找到房地产发展的相关指数，我们采取 SPCA 方法对数据进行处理，得到稀疏主成分分析的载荷：

❖ 北京住宅相关指标：

表 6

住宅相关指标				
变量名		PC1 _a	PC1 _b	PC1 _c
住宅开发投资额 (亿元)	X ₁₁	0.083	0	1
住宅销售额(亿元)	X ₁₂	0.288	0	0
住宅施工面积 (万平方米)	X ₁₃	0.347	0	0
住宅竣工面积(万 平方米)	X ₁₄	0.515	0	0
中央竣工面积(万 平方米)	X ₁₅	0.005	1	0
地方竣工面积(万 平方米)	X ₁₆	0.424	0	0
住宅销售面积(万 平方米)	X ₁₇	0.587	0	0
方差贡献率(%)		66.2	11.6	4.2
总计方差贡献率(%)		82.1		

❖ 房地产开发企业相关指标：

房地产开发企业相关指标					
变量名			PC2 _a	PC2 _b	PC2 _c
房地产开发企业个数（个）		X ₂₁	0.223	0	0
房地产开发 企业(单位) 资金来源 (亿元)	国内贷款	X ₂₂	0.412	0	0
	利用外资	X ₂₃	0.019	0	1
	自筹资金	X ₂₄	0.365	0	0
	其他资金来源	X ₂₅	0.435	0	0
房地产开发 企业(单位) 资产(亿元)	资产总计 (亿元)	X ₂₆	0.360	0	0
	本年折旧 (亿元)	X ₂₇	0.278	0	0
	负债总计 (亿元)	X ₂₈	0.366	0	0
	所有者权益 (亿元)	X ₂₉	0.342	0	0
	资产负债率（%）	X ₂₀	-0.027	1	0
方差贡献率（%）			76.1	6.0	2.6
总结方差贡献率（%）			84.8		

表 6

❖ 北京政策相关指标

北京政策相关指标				
变量名		PC3 _a	PC3 _b	PC3 _c
地区生产总值比上年 年增长（%）	X31	0.035	0.000	0.684
全社会固定资产投资比上年 增长（%）	X32	-0.048	0.000	0.729

全社会固定资产投资（亿元）	X33	0.580	0.000	0.000
房地产开发投资（亿元）	X34	0.583	0.000	0.000
本年完成的土地开发面积（万平方米）	X35	0.037	0.612	0.000
土地开发投资（亿元）	X36	-0.036	0.791	0.000
土地购置费（亿元）	X37	0.563	0.000	0.000
方差贡献率（%）		42.0	22.3	19.2
总计方差贡献率（%）		83.6		

表 7

❖ 买房居民相关指标：

居民相关指标				
变量名		PC4 _a	PC4 _b	PC4 _c
家庭总收入	X ₄₁	0.371	0	0
可支配收入	X ₄₂	0.370	0	0
居民消费水平（元）	X ₄₃	0.369	0	0
消费性支出	X ₄₄	0.359	0	0
居民消费价格指数 （上年=100）	X ₄₅	-0.025	0	1
家庭恩格尔系数 （%）	X ₄₆	-0.254	0	0
人均住房建筑面积 （平方米）	X ₄₇	0.337	0	0
人均住房消费支出	X ₄₈	0.380	0	0

(元)				
常住人口(万人)	X_{49}	0.370	0	0
就业者负担人数 (人)	X_{40}	-0.012	1	0
方差贡献率(%)		76.6	10.0	6.0
总计方差贡献率		92.6		

表 8

➤ 结果展示

为得到房地产发展指数，我们综合分析上述因素，将4个方面的主成分做整合得到综合指数： PC_1, PC_2, PC_3, PC_4 分别代表北京住宅综合指数、北京房地产发展企业综合指数、北京居民综合指数、北京相关政策综合指数

4个方面的综合指数计算公式如下：

$$PC_k = \sum_{i=a}^c \alpha_i PCk_i \quad (k=1, 2, 3, 4)$$

其中，权重 α_i 等于方差贡献率除以三个主成分的总计方差贡献率得到的比值。

进一步利用VaR模型构建各个评价价值考察北京房地产市场发展的历史变化趋势。若评价值为负值，说明该指数的水平处于1998—2010年期间平均水平之下；若为正值，则说明该指数水平处于平均水平之上。

评价价值				
年份 指数	住宅相关综合指数 (PC_1)	房地产开发企业相 关综合指数(PC_2)	政策相关综合指数 (PC_3)	居民相关综合指数 (PC_4)
1998	-1.794	-2.706	0.956	-1.338
1999	-1.549	-2.783	0.947	-1.027
2000	-1.091	-2.566	0.558	-0.591
2001	-0.096	-2.293	0.880	-0.334

2002	0.967	-1.965	0.712	0.237
2003	1.786	-1.795	0.561	0.581
2004	2.726	-0.582	0.677	0.933
2005	3.606	0.177	0.405	1.505
2006	2.626	0.614	1.061	2.156
2007	2.268	1.627	1.254	2.559
2008	1.216	3.209	1.314	3.043
2009	2.931	3.689	1.045	3.866
2010	2.953	5.373	2.242	4.577

表 9

➤ 房地产发展指数的构造

下面，我们仅以这四个综合指标作为影响因素，再次采取 SPCA 方法分析各年的房地产发展状况，得到最后的房地产发展指标，并利用 VaR 模型构造房地产发展指数——综合评价价值：

我们给出了两个方案：

❖ 方案一

房地产发展指标 Z ：

房地产发展指标 (Z)	
指标	载荷
PC_1	0.4396992
PC_2	0.5609281
PC_3	0.4171588
PC_4	0.5639174
方差贡献率	0.7688207

表 10

综合评价 A:

综合评价 (A)	
1998	-2.0631974
1999	-1.9360310
2000	-1.9980078
2001	-1.3438708
2002	-1.0040783
2003	-0.8061174
2004	-0.1266558
2005	0.1665167
2006	0.7987280
2007	1.2127943
2008	1.4843750
2009	1.9956066
2010	3.6199379

表 11

❖ 方案二

房地产发展指标 S:

房地产发展指标 S		
指标	发展指标 1 载荷	发展指标 2 载荷
PC ₁	0.6609374	0
PC ₂	0.4604285	0
PC ₃	0.0000000	1
PC ₄	0.5925937	0
方差贡献率	64.4	17.9
方差贡献率总计	0.8232373	

表 12

类似综指数的构造，我们可以构造综合评价值 B:

综合评价值 (B)	
1998	-1.837615797
1999	-1.698849426
2000	-1.255601858
2001	-1.028506867
2002	-0.470325878
2003	-0.065418921
2004	0.389183467
2005	1.000603269
2006	0.640126139
2007	0.682027379
2008	0.686958934
2009	1.554907621
2010	1.402512014

表 13

➤ 可视化：

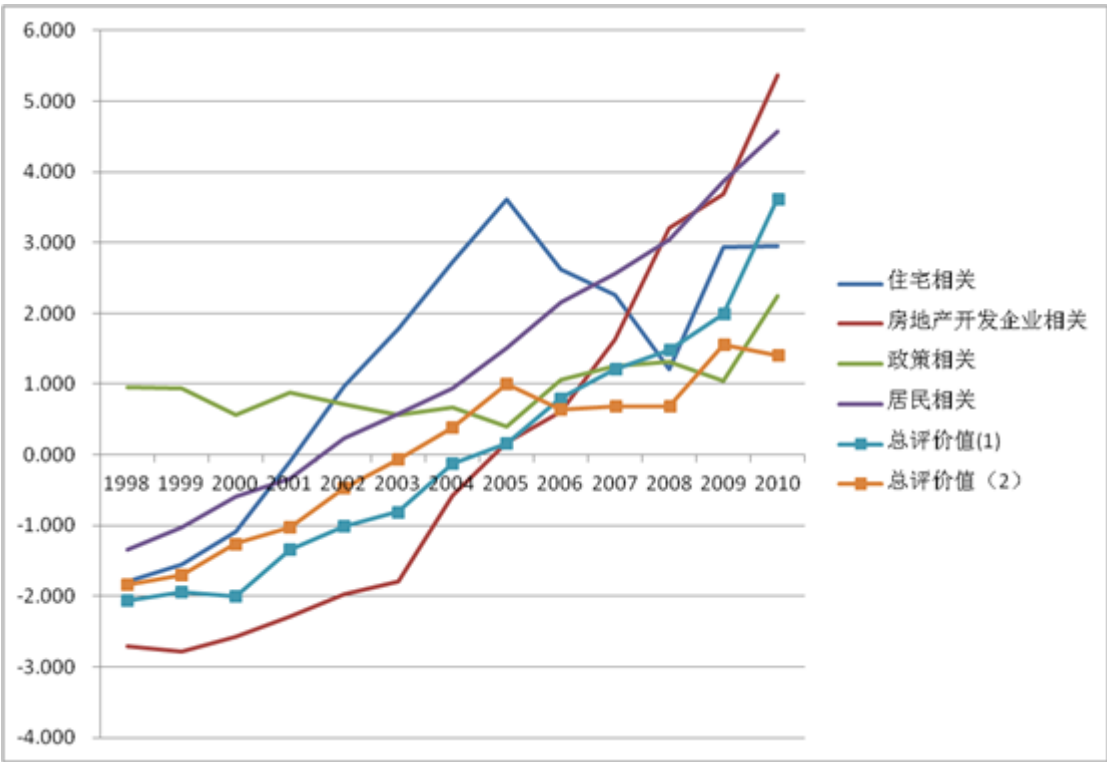


图 1

1. 结果解释及分析：

➤ 住宅相关综合指标

此主成分与几个主要指标比如住宅开发投资额，住宅竣工面积以及销售面积等相关系数相近，这是一个与我们通常认识相符的结果。即房地产市场的发展与住宅方面的投资和销售紧密相连，在住宅方面投入越多的资源，房地产市场发展越迅速。而另一方面，我们的结果表明，中央竣工面积对主成分的贡献为负值，这是一个有趣的结果。中央竣工面积，即由政府相关单位投资竣工的住宅面积，目的是用于其职工的住宿。此指标对房地产市场的发展有负影响，可以一定程度说明政府对住宅市场的介入对房地产市场发展有一定的坏处。

➤ 房地产开发企业相关综合指标

此主成分与房地产开发企业个数以及资金来源的指标都成正相关，仅与房地产开发企业

(单位)资产中的资产负债率成负相关。这说明房地产企业的资金来源对于房地产市场的发展起重要作用,其中又以自筹和其他方式的正面作用最大。房地产开发企业(单位)的总资产以及所有者权益对房地产市场的发展也成正相关的,这些企业的投资开发能力越强,房地产市场也自然会壮大。此外,这也反映了房地产市场对资本的强烈依赖性。由结果我们可以看到,房地产开发企业的资产负债率与房地产的发展是负相关的,企业的负债率决定了它的流动资金,利润等因素,因此在一定程度上限制了企业的项目开发,所以呈现负相关。

➤ 政策相关综合指标

我们的结果表明,房地产政策相关的主成分需要三个才能够达到较高的方差贡献率。这三个主成分分别于全社会固定资产投资,房地产开发投资,土地购置费、土地开发投资以及全社会固定资产投资年增长率呈现较大正相关,且负相关的指标少而且相关程度低。这说明在地价以及全社会的财富增长率与房地产市场的发展是正相关的,地价的增长以及 GDP 的增长对房地产市场的发展是有利的,即 GDP 的增长促进了人们对工作和居住场所的更高需求,也带动了地价的上涨。

➤ 居民相关综合指标

此主成分与家庭可支配收入、消费水平、人均住宅面积、以及常住人口呈现相关系数非常平均的正相关,并与居民消费价格指数以及家庭恩格尔系数成负相关。这些总体说明家庭的总体消费水平越高,对房地产市场的发展也就更加有利。用于家庭的居住房屋,也就是媒体经常提到的“刚性需求”对房地产市场的发展是有促进作用的,而且此需求会长期存在。

➤ 结果分析

我们将各年的四个方面的综合指标的评价值计算出来并进行单位化后得到 1998-2010 年每年评价值。将各年四个主成分得分与整体得分的曲线图画出,对比可以发现综合评价值 A 与房地产开发企业主成分以及居民相关主成分趋势一致,都在不断上升,其中 09-10 年的涨幅最大,这与房地产企业在 09-10 年实力的强力增长以及政策的对房地产销售的促进有密切关系。而房地产企业在 1998-2010 年间的资金来源增长受益于金融业对它的影响,因此房地产企业在这段时间的成长与金融业的发展有密切关系,中国金融业的起伏在房地产市场发

展上得到了显著反应。

综合评价值 B 的上升与下降与住宅相关主成分的大致保持同步，而与其他三个方面的主成分趋势有较大差异，这从另外一个方面评价了北京市的房地产市场发展情况。不同于总得分 1 的一直上升趋势，总得分 2 在 05-07 年持续低低迷，直到 08 年才有回升的趋势。这可能与 05-07 年房市波动造成的住宅购买信心不足有关，08 年后人们对房地产市场信心回复，使得房地产市场回暖。另外由于总得分 2 的发展趋势与住宅投资金额以及竣工面积和销售面积的发展趋势较为一致，我们认为总得分 2 比较合理的反应了房地产市场的总发展情况。

2. 初步预测

➤ 趋势图展示及分析

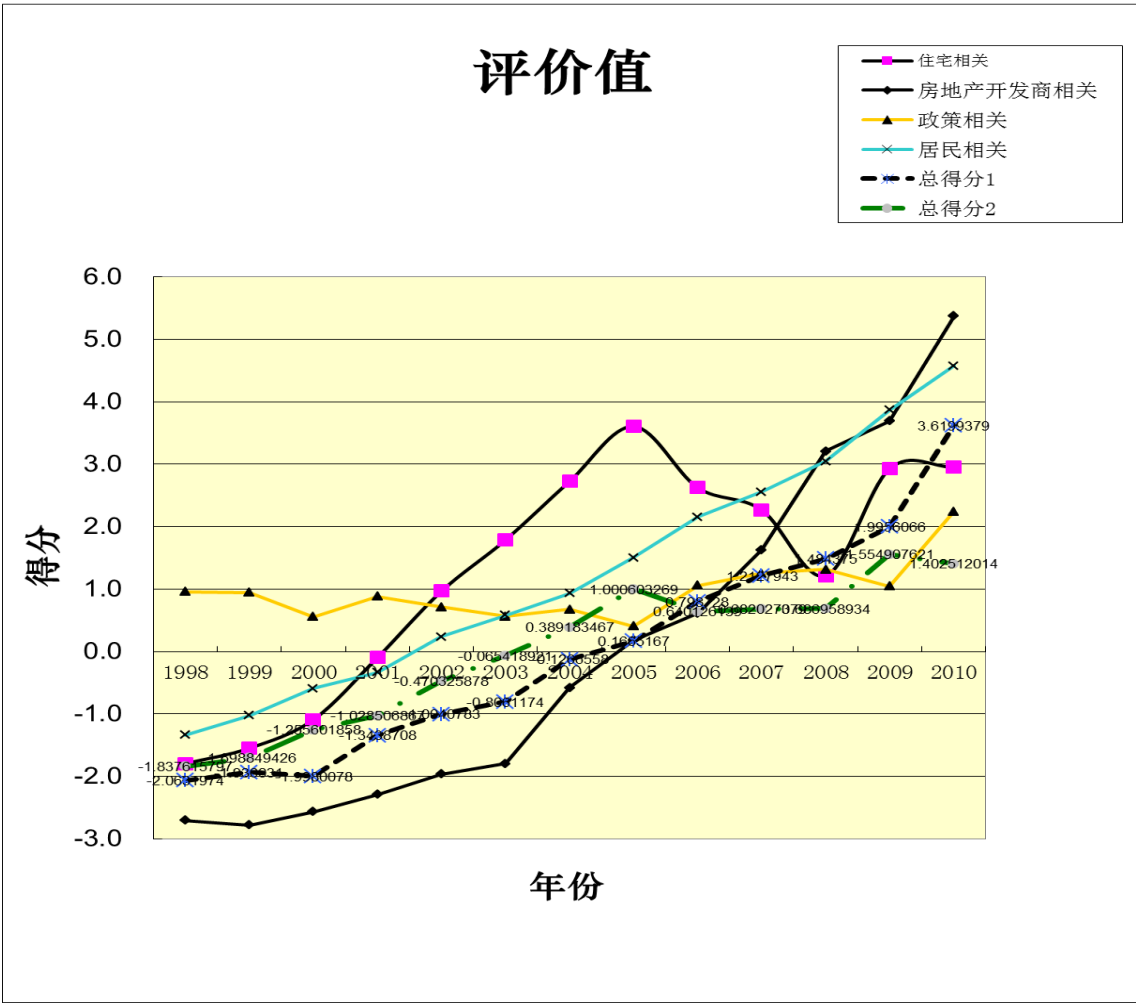


图 2

上图是由 SPCA 方法中的得分一项的最终各项趋势图，可以看出，住宅相关组指标得分走势与其他数据组有较显著的差距，主要在于 2003~2006 的高得分波动。其实细致寻找房地产相关组指标走向也可以发现，在 2003~2006 年期间，中央竣工面积、地方竣工面积、住宅销售面积以及商品房销售面积都有较高的值，对该指标组也做出了不小的贡献。另外，政策相关指标组得分一维拟合斜率最小，可以知道政策变化幅度较小，也符合国家稳定发展的基本政策方向。而开发商相关指标组数据的一维拟合斜率最大，可以知道开发商相关指标组在 1998~2010 年之间增长幅度最大，显示出开发商一方在近几年我国房地产发展过程中的些许潜在引导功能。另外，居民相关的指标组数据趋势居于中间地位，并且走势较为平稳，受住宅与开发商相关指标影响较大。

➤ 预测方法应用

那么我们下面将展示北京市 CPI 以及 GDP 数据变化图，具体方法就是找出下述指数变化趋势与上面一个小节中的那个折线走向最接近，并以此来确定我们最终的房产发展指数模型。

乍一看，相应时间的北京市 CPI 指数走势与我们的 SPCA 算法中的政策相关得分较接近（见上图中黑色线条括住的部分）。然而，仔细分析其中的具体走势，发现数据波动较大，这一点政策相关不能很好地反应，综合考虑我们认为还是住房相关指标组得分更加接近 CPI 变动情况。

➤ 初步预测

因此，我们决定在初步预测中运用住房相关指标组得分来预测房产发展。这里要先对 CPI 与房价的关系进行一下评价：

➤ 合理性分析

CPI 与房价之间的关系

实际上，即便在资产价格不直接计入 CPI 的情况下，二者之间也存在着广泛的联系。拿房屋价格作为例子，房价的波动性很强，一般而言，当房价上涨，房租也会随之上涨，但房租不会剧烈波动，因为即便房地产价格急速上涨，也很难马上反映在房租上。所以当房价

持续上涨时，房租则缓速上涨，CPI 也会受其影响而有所提高。此外，在将贷款利率计入 CPI 中的情况下，当政府为控制房地产过快增长势头而提高利率时，则不可避免地也会引起抵押贷款利率升高，这也会推高 CPI。

房屋价格与其他资产的一个重要区别在于，房屋本身不仅是一项投资品，同时也能够为人们提供消费服务。因此有的国家也将房屋价格计入 CPI（如澳大利亚和新西兰），但大多数国家都将消费房屋的日常成本，如房租、自住房屋的隐含租金、水电、燃气等费用纳入 CPI 中。这些支出项目都比房价要稳定的多，比如我们看到房价在飞速上涨，但房租却比较稳定。因为一般房租是一年签订一次或是数年签订一次，在签订后便会在一段时间内固定下来，而不会发生波动，即便在长期持续的房地产价格上涨的时候，它也只是会稳步上涨。当然也有例外，比如在北京奥运会期间，奥运场馆周边的房屋月租金就达到了平常的 4~6 倍，有的甚至超过 10 万元。而 2000 年悉尼奥运会期间，一些地处“奥运通道”周围的悉尼居民将自己的住房出租给前来悉尼的观光者，房租提高到原先水平的 4 倍。2004 年雅典奥运会期间租金平均上涨幅度也在 300%~400%之间。但这些都是受到特殊需求突然增加的推动，它只是一种局部地区的短期现象，随着奥运会的结束，这些地方的房租也都会恢复如初。

为了满足居住需求，我们知道除了支付房租或者月供款之外，还需要购买一些配套商品或服务，如装修材料、家具、家电、家政服务等等。当我们购买住房时，这些相关的消费品或服务同样也需要购买。也就是说，对住房配套设施的需求是对住房服务需求的衍生需求，当对住房的需求提高时，那么对这些配套设施或服务的需求也会提高。因此当房价因需求提高而上升时，其他配套设施或服务的价格也会提高，这也会带动 CPI 提高。当然，这些物品对 CPI 的影响程度除了取决于其价格的上升幅度外，还取决于它们在 CPI 商品篮子中的权重，权重越大，其对 CPI 的影响也越大。

此外，如果房屋的购买是采取抵押贷款的方式，即将房屋抵押给银行后获得购房贷款，再支付购房总价，若房价在购房后上涨，那么也就表明抵押品的价值也提高了。一些西方发达国家，在这种情况下，银行允许抵押房贷款者可在房价提高的基础上增加贷款融资。这就会鼓励房屋所有者进一步向银行借款，来为更多的消费融资。这被人们称为房价的信贷市场效应。当利率降低时，它就会使未来住房服务的贴现值（也就是房价）提高，于是就会产生这种信贷市场效应，它会刺激人们的消费，从而从总体上带动对其他商品的需求。

人们的消费需求提高，是相关商品价格提高的一个前提条件，但实际上价格是否会提高还取决于市场上对这些产品的供给情况。如果人们的需求提高不能为现有的供给所满足，那么价格将会作出上涨的反应，但如果人们的需求提高可为现有的生产能力所满足，那么就不

会给这些商品的价格带来上涨的压力。因此房价上涨通过影响消费而对物价所产生的影响也要取决于经济中的供给情况。

房价上涨在一定程度上会带动CPI的提高,相反地,CPI的持续膨胀也会带动房价的上涨。尤其是当通货膨胀率高于名义利率时,这意味着银行存款的实际利率是负数,人们存钱不仅不会获得好处,反而会被通货膨胀把获取的一点点利息侵蚀掉。于是人们就会把钱投资于其他地方,而在这个时候房贷(实际)利率也会较低,这便会刺激人们把银行中的积蓄拿出来购买房产,进而推动房价上涨。

虽然在高通胀的情况下,房贷金额在开始时占资产价值(房屋总价)较高,但随着房价的上涨,二者的比值会越来越低。而未付完的房贷金额则相对稳定,或者减少。同时人们偿还的月供可能在一开始较高,但人们的收入也会随着通货膨胀的影响而增加,工资中用来支付贷款利息的比例也会不断下降(这实际上也反映了货币在不断贬值),于是很快房贷金额与资产价值之间的比值就会下降,这会降低借款人的风险。在一些高通胀的国家,购买房产还有税负的优势,即自有住宅的资本利得是免税的,而其他资产(如股票)却无减税的优势,这就使得购买房产具有更大的优势。这样,因为人们的追捧而导致房价上涨,而房价的上涨又会进一步刺激人们去购买,由此导致房价攀升。

虽然在通货膨胀时期,可能所购房产的实际价值(指扣除了物价上涨因素后的价值)并不会如人们所看到的那样巨幅上涨,但由于人们贷款的货币价值会下跌,这实际上就补偿了借款人的损失。随着房价的不断上涨和人们收入因通货膨胀而提高,人们会发现自己的债务负担不断减轻,虽然在购房的前几年负担会比较沉重,但随着时间推移,人们就会感到越来越安全。

此外,一般人都很容易犯“货币幻觉”的毛病,就是只关注物品的名义价格,而不注意其实际价值。在高通胀的情况下,虽然房屋的实际价值可能并未变化,房屋根本就没有增值,但购房者却只会看到不断攀升的房价,把增加的房价看成是自己的所得,这会促使他们去借更多的钱,买更大的房子,这就出现了人们追赶房价的现象。因此在高通胀的时期,房价会比较容易上涨。

当通货膨胀加剧,超出了政府认为合理的范围,政府就会出台紧缩性政策,紧缩银根使房地产商贷款变得日益艰难,而同时购房者在房价大幅度上涨后也越来越会选择驻足观望,于是房价就会在高位动摇,直至下跌。如先期房价涨幅过大,形成泡沫,这时可能就会导致泡沫破裂。而这可能又会反作用于物价,因为受房价下跌、资产缩水的影响人们会缩减消费,推动物价下跌。

因此，CPI 与房价还是有一定关联，另外，CPI 又由我们的分析可知与 SPCA 分项得分关联度较大，因此我们运用稀疏主成分分析得分来表示房产发展是合理的。

3. 优势、不足与展望

➤ 优势、创新

1. 应用范围大型化，预控细化。房地产风险评价有很多方法，各有优劣。本项目在综述当前已有方法的基础上，获得更具普适应的应用模型。该模型不仅获得一个综合指数评价投资风险，而且有助于我们了解影响投资风险的因素以及影响效果，从而为决策提供依据。
2. 采用最新的降维方法进行数据处理。对于计算过程中出现的高维数据和多重变量相关性，采用较新的稀疏主成分分析(SPCA)对其降维处理，并采用统计软件 R 编程，对各种稀疏算法的稀疏效果进行比较，并作相应的理论分析。同时，我们采用适用范围更广的稀疏主成分综合评价方法，更好地对房地产投资风险进行了评价，从而为决策提供依据。
3. 算法实例例证，理论建模与实证数据分析相结合。通过模拟出的数据我们将与实际进行预对比并进行整合从而整理出最后一个合理的风险指数与波动值。这正是模糊风险指数分析中所缺少的。
4. 充分考虑了政策对房地产发展的重要影响。

➤ 不足之处

1. 由于数据不充分，笔者仅就所收集到的房地产市场的基本指标进行了分析，但这并不影响采用房地产市场发展指数的模型构造思想分析和预测未来的波动状态。
2. 直接用住房指标组得分情况来反映房产发展指数情况稍稍有失偏颇，我们并没有综合几个指标组的得分情况。还有，关于 CPI 来反映房产发展，虽然有较多证据，但是似乎还有更好反应这个的，比方说全社会固定资产投资等指标。不过这只是我们的初步预测，我们会在接下来的研究中做到更加完善的。
3. 我国房地产市场统计历史不长，存在基础指标类别和序列不仅齐全的问题。
4. 笔者仅就北京地区总体上进行了分析和预测，并未给出北京市各区的具体情况，也没有

对地域方面的问题作进一步分析和阐述。

➤ 展望

如果条件允许,我国应该建立起各大类房地产市场统计指标体系,以满足社会对日益增长的房地产市场信息量的需求,这样,我们可以再加入更多的房地产发展影响因素,进行全面分析。

房地产的固定性和差异性以及由此而决定的房地产准市场性和地区性市场特性,导致了房地产市场运行规律不同于一般商品市场的运行规律,其表现为各区域性房地产市场具有一定的特殊性,故此,要求对其进行适度的细分,以达到体现房地产市场的区域差异性,又要便于时间序列数据的统计和管理。接下来笔者打算进一步以北京市各区为划分,进一步探讨房地产发展。

笔者只涉及到了房地产发展指数的构造和对实际情况的分析描述以及预测,对如何运用该指数对房地产市场波动的重要转折点、波动的危害性、波动的质量以及波动的警度分析还有待进一步探讨。

4. 结语

由于房地产企业具有投资规模大、资金回收期长、负债比率高等特点,并受当前的宏观经济形势的影响,所面临的财务风险较大,因此,房地产发展经常出现较大波动。近几年,房地产市场在我国得到快速的发展,往往成为地区经济发展的重要支柱产业,这在北京尤其显著。然而,房地产市场的发展受到了众多因素的影响,而且都具有很强的相关性,因此如何更好地对房地产市场在进行实证研究就需要谨慎对待。

本文先围绕房地产相关因素,将影响指标按“卖家、买家、房地产业、宏观政策”分为四大部分,并采用稀疏主成分分析房地产市场的发展状况进行综合评价。该方法是在主成分的基础上,牺牲主成分的部分方差贡献率,更进一步的对载荷进行稀疏,从而更清晰的揭示出数据的内在结构,显示出比主成分分析更强大的优势。接着,我们通过风险评估的 VaR 模型,确立了发展评价值——即房地产发展指数,并根据该指数对当前房地产事业进行了评价和分析,给出了该指数与 CPI 等的关系,并最终确立了可以衡量房地产发展的便利指数。

5. 参考文献

- 【1】. 丁烈云, 李 斌. 房地产市场预警调控系统的构筑技术要点及流程设计[J]. 系统工程理论与实践, 2002, 22(4): 58-62.
- 【2】. Quigley John M. Real estate prices and economic cycles [J]. International Real Estate Review, 1999, 2(1): 1-20.
- 【3】. Dipasquale D, Wheaton WC. Housing market dynamics and the future of housing prices [J]. Journal of Urban Economics, 1994, 35: 1-27.
- 【4】. 张尧庭, 方开泰. 多元统计分析引论[M]. 北京: 科学出版社, 1997.
- 【5】. 王洛林. 经济周期研究[M]. 北京: 经济科学出版社, 1998.
- 【6】. Niemira MP, Klein PA. Forecasting Financial and Economic Cycles[M]. New York: John Wiley & Sons, Inc, 1994.
- 【7】. Zou, H., Hastie, T., and Tibshirani, R. (2006), Sparse principal component analysis, *Journal of Computational and Graphical Statistics*, 15, 265–286.
- 【8】. Y. Wang and Qiang Wu, Sparse PCA by iterative elimination algorithm, *Advances in Computational Mathematics*, 36 (2012), 137–151.
- 【9】. 钱争鸣, 陈伟彦, 我国工业经济效益指标评价与主成分分析的实证研究, 统计研究, 1999, (7): 49-52
- 【10】. 北京市统计局编, 北京统计年鉴: 1991-2011, 中国统计出版社, 1991-2011。
- 【11】. 刘超, 吴丹丹, 齐嘉悦. 基于稀疏主成分的房地产市场发展分析: 以北京为例[J]. 经济研究导刊, 2012(30).
- 【12】. 赵雪红. 城市房地产投资风险管理研究——以济南景泉花园为例[D]. 山东大学, 2006.
- 【13】. Zou H, Tibshirani T H A R. Sparse Principal Component Analysis[J]. *Journal of Computational and Graphical Statistics*, 2006(15:2).
- 【14】. 杨志勇, 申映华, 霍腾飞, 等. 基于 PCA-SVM 的中国省域建筑产业竞争力预测[J]. 统计观察, 2012(23).
- 【15】. 刘宇飞. VaR 模型及其在金融监管中的应用[J]. 经济科学, 1999(1).
- 【16】. 学习 . CPI 与 房 价 之 间 的 关 系 [EB/OL]. [2013-4-27].
http://szbbs.soufun.com/salon~1~55/77949794_77949794.htm.

