

静大周辺の賃貸アパート/マンションの家賃の重回帰分析を用いた分析

安藤 和敏
静岡大学工学部

平成 17 年 11 月 9 日

1. はじめに

私が浜松に引っ越してきて、かれこれ 3 年半過ぎた。現在のアパートは、浜松という土地柄をあまり知らずに決めてしまったのであるが、3 年以上も住んでいると現在のアパート周辺の環境やアパートの部屋自体にも不満がでてきた。そろそろ引っ越しをしたい。

静大周辺は大学に近くて、かつ、駅にもある程度近くて便利そうだ。学生たちに私生活を暴かれるというリスクはあるが、静大の近くに引っ越そうと思う。広沢、城北、布橋、住吉、和地山、文丘の 6 町にある賃貸アパート/マンションの相場というのを調査し、引っ越しの際の参考にしたいというのが本分析の目的である。

2. データ

アパマンショップの Web ページ [1] で、2005 年 10 月 26 日の夜の時点で、城北、布橋、住吉、和地山、文丘、広沢の賃貸アパート/マンションを検索すると 191 件の物件がヒットした。これを元にして、Excel で分析するためのデータとする。

とりあげる変数として、住所 (広沢、布橋、城北、住吉、文丘、和地山のいずれか)、家賃 (単位万円)、面積 (単位 m^2)、タイプ (アパート、マンション、コーポのいずれか)、築年数 (単位年) をとりあげる。もちろん、家賃が目的変数で、それ以外の変数は説明変数である。

住所とタイプは質的変数であるから住所に対応するダミー変数: { 布橋, 城北, 文丘, 住吉, 和地山 }, 及び、タイプに対応するダミー変数: { マンション, コーポ } を考えた。したがって、説明変数の候補は、住所 ({ 布橋, 城北, 文丘, 住吉, 和地山 }), 面積, 築年数, タイプ ({ マンション, コーポ }) である。このなかから、変数をいくつか選んで価格を説明する回帰モデルを構築しよう。

3. モデル 1 (単回帰)

まず最初に、家賃を一つの変数で説明することを考える。このとき、一つの説明変数として家賃との相関係数が最も高い変数を選ぶのが自然であろう。

以下に家賃と各説明変数 (ダミー変数を含む) との相関係数を示す。

布橋	城北	文丘	住吉	和地山	築年数	面積	コーポ	マンション
-0.31	0.13	-0.07	0.07	0.06	-0.19	0.92	-0.11	0.11

家賃との相関係数が一番高いのは面積である。しかも他の変数に比べて際立って高い。また、築年数というのが、家賃に対して大きな影響を及ぼすはずだと予想していたのであるが、実際には、-0.19 という非常に低い数字となったのが意外であった。ちょっと信じられなかったので、相関図 (散布図) を作ってみた。図 3.1 と図 3.2 に、それぞれ、面積と家賃の相関図と築年数と家賃の相関図を示した。確かに、面積と家賃とは線型の相関があり、築年数と家賃との間には (線型の) 相関は認められない。

話が逸れたが、以下の単回帰モデル

$$\text{家賃} = a + b \times \text{面積} \quad (3.1)$$

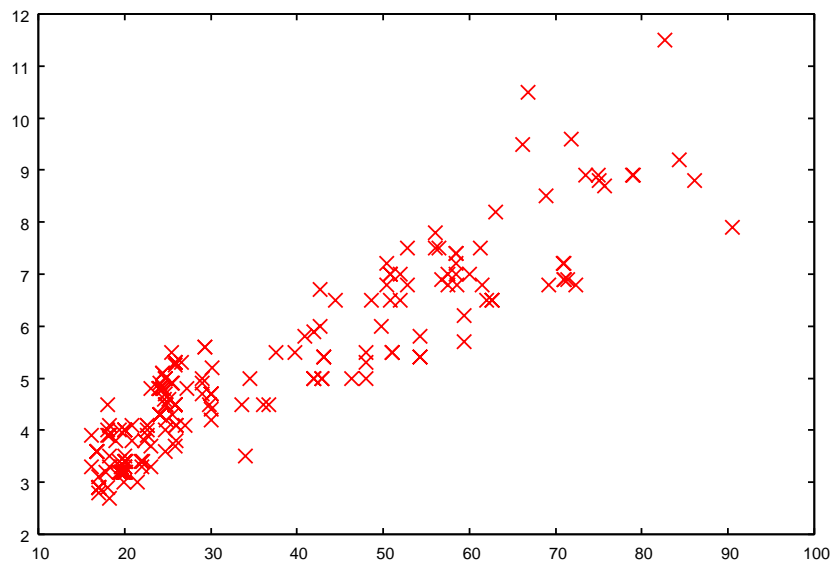


図 3.1: 面積 (横軸) と家賃 (縦軸) の相関図

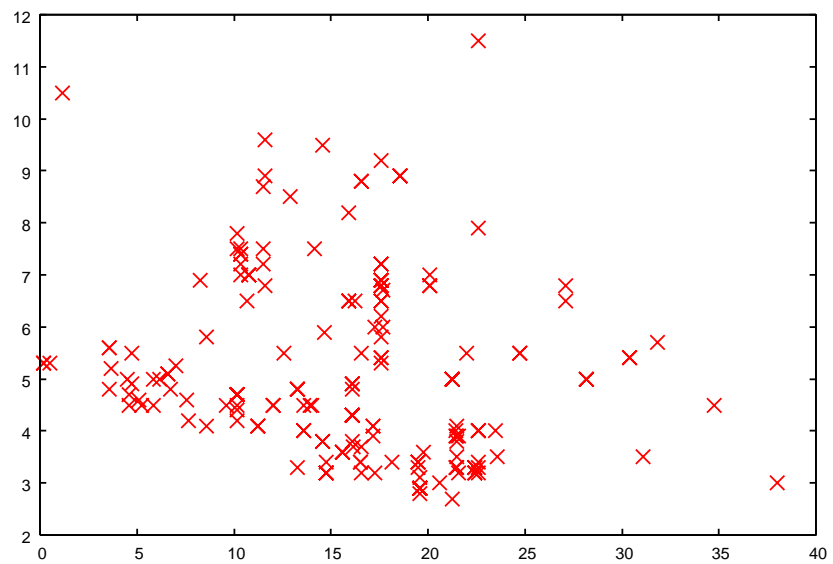


図 3.2: 築年数 (横軸) と家賃 (縦軸) の相関図

について単回帰分析を行った結果,

$$\text{家賃} = 2.1 + 0.08 \times \text{面積} \quad (3.2)$$

という回帰方程式を得た. R, R^2, R^{*2} は以下の通りである.

重相関 R	0.92
重決定 R2	0.84
補正 R2	0.84

悪くない結果である. だが, もっと欲ばっていいモデルを追求していこう.

4. モデル 2

次に, 説明変数を面積に加えてもうひとつ追加して 2 つにすることを考える. もっとも良い補正決定係数 (補正 R2) を与える説明変数を選ぶことにする.

	面積+築年数	面積+タイプ	面積+住所
重相関 R	0.9466	0.9218	0.9321
重決定 R2	0.8961	0.8498	0.8688
補正 R2	0.8950	0.8474	0.8645

面積と築年数を説明変数とする組合せが最も良さそうなので,

$$\text{家賃} = a + b \times \text{面積} + c \times \text{築年数} \quad (4.1)$$

という重回帰モデルを仮定して, 回帰方程式を求めると,

$$\text{家賃} = 3.04 + 0.08 \times \text{面積} - 0.06 \times \text{築年数} \quad (4.2)$$

が得られた.

5. モデル 3

モデル 2 を得たのと同じ方法で, 面積と築年数に追加する第 3 の説明変数を選ぼう.

	面積+築年数+住所	面積+築年数+タイプ
重相関 R	0.9545	0.9470
重決定 R2	0.9112	0.8969
補正 R2	0.9078	0.8947

追加する説明変数は住所ということになって,

$$\text{家賃} = a + b \times \text{面積} + c \times \text{築年数} + d \times \text{布橋} + e \times \text{城北} + f \times \text{文丘} + g \times \text{住吉} + h \times \text{和地山} \quad (5.1)$$

という重回帰モデルを仮定して, 回帰方程式を求めると,

$$\text{家賃} = 3.33 + 0.09 \times \text{面積} - 0.06 \times \text{築年数} - 0.56 \times \text{布橋} - 0.72 \times \text{城北} - 0.6 \times \text{文丘} - 0.3 \times \text{住吉} - 0.27 \times \text{和地山} \quad (5.2)$$

が得られた.

住所を示すダミー変数の係数が全て負になっているのが気になるかも知れないが, これは広沢を 0 としたときの係数である. つまり, 広沢のアパートの家賃が最も高いということである. 例えば, 城北の係数は-0.72 であるから, 他の条件が同じであれば, 城北のアパートは広沢のアパートよりも 7200 円安いということである. (なんか変かも...)

6. おわりに

全ての説明変数を使った重回帰分析もやってみたが、ほとんど補正決定係数は増加しなかったので、本分析の結論としては、モデル 3 で得られた回帰方程式 (5.2) を、家賃を説明するモデルとして結論付ける。分析の結果、家賃に最も影響を及ぼす要素は面積であり、築年数や住所はほとんど影響を及ぼさないということが分かった。築年数の影響が小さいことについては、築年数が経っていてもリフォームをして家賃を下げないのが理由であろうと考えられる。また実際は、その部屋が何階にあるかとか、日当たり、鉄筋か木造かなど、今回の分析には用いなかった変数が家賃に影響を及ぼしていると考えられる。こうした変数を用いた重回帰分析については今後の課題としたい。

講義では説明しなかったし、ここでも適当にやっているのであるが、どのように説明変数を選ぶかは難しい問題である。この話をしただしたら、あと 3 回くらい講義をやらなければならないかも知れない。この講義は、あくまで「入門」ということなので、この話には踏み込まないことにするが、重回帰分析に興味をもった学生はさらに勉強したら良い。

また、例えば SPSS [2] のような統計解析ソフトを使うと説明変数の選択を自動的にやってくれたりもする¹。SPSS のユーザになるというのも一つの考えではあるが、システム工学科の学生ならば SPSS のようなソフトを作ってしまうくらいの志が欲しい。

参考文献

[1] <http://www.apamanshop.com/>.

[2] <http://www.spss.co.jp/>.

¹価格は 10 万円以上もする!