

多元回归 第一周作业

蒋翌坤 20307100013

《实用多元统计分析》P28, 1.4

(a) 为变量 x_1 和 x_2 作散布图和边缘点图，并加以解释

变量 x_1 和 x_2 的散布图及边缘点图

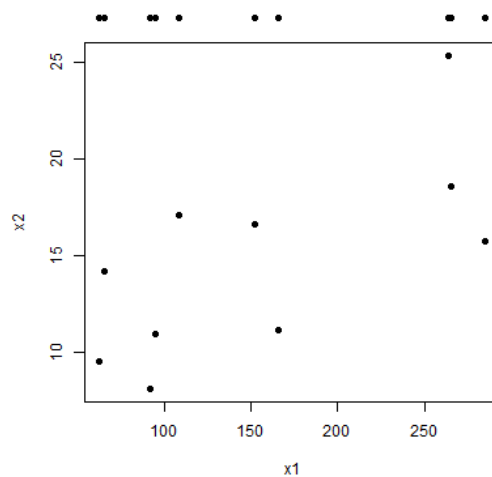


图 1

变量 x_1 和 x_2 的散布图和边缘点图如上图（图 1）所示。图中在方形边框内部的黑色圆点坐标为 (x_1, x_2) ，组成变量 x_1 和 x_2 的散布图。在方形边框上侧和右侧的黑色圆点分别代表变量 x_1 和 x_2 对应的数据点，组成变量 x_1 和 x_2 的边缘点图。

(b) 计算 \bar{x}_1 , \bar{x}_2 , s_{11} , s_{22} , s_{12} 和 r_{12} , 解释 r_{12}

由计算可得（精确到两位小数）：

$$\bar{x}_1 = 155.60, \bar{x}_2 = 14.70, s_{11} = 7476.45, s_{22} = 26.19, s_{12} = 303.62, r_{12} = 0.69$$

r_{12} 表示变量 x_1 和 x_2 之间的相关系数， $r_{12} = 0.69$ 可以表示变量 x_1 和 x_2 之间有较强的正相关性。

《实用多元统计分析》P35, 1.20

参看第 11 章表 11.4 的破产数据和网页 www.prenhall.com/statistics 上的数据，利用适当的计算机软件：

(a) 观察 x_1, x_2, x_3 空间中的全部数据集，向不同的方向旋转坐标轴，检查异常观测值

下图（图 2）展示了从不同方向观测 (x_1, x_2, x_3) 的情况，从散点图中则观察不出有明显的异常观测值，而从三个箱线图中可以发现少量异常观测值， x_1 的异常观测值有 (-0.56) ， x_2 的异常观测值有 $(-0.41, -0.31, -0.30, -0.27)$ ， x_3 的异常观测值有 $(5.06, 4.45, 4.29, 4.24)$ 。

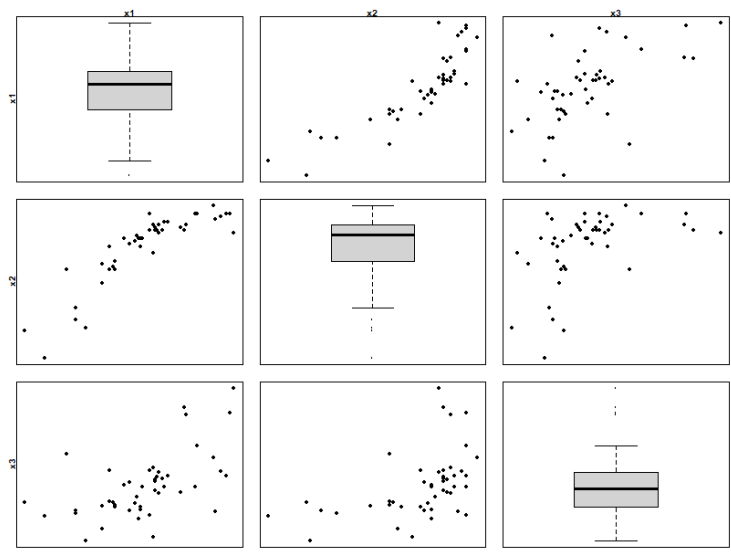


图 2

(b) 高亮显示破产公司对应的点集，检查各种三维透视图，存在能辨别破产公司和非破产公司的三维空间的某些方向性吗？根据数据计算出样本均值、样本方差和样本协方差，以得出划分破产公司与非破产公司的规律，在这两类公司中是否有观测值会对该规律有重大影响（参见练习 11.24）？

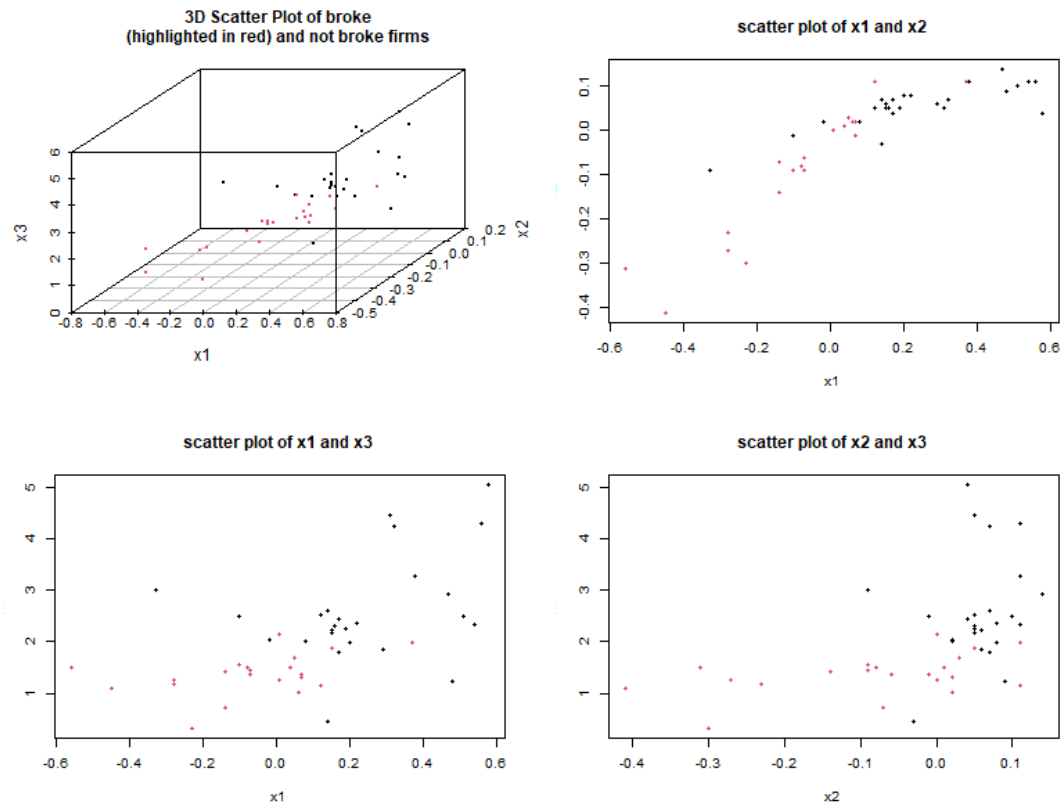


图 3

上图（图 3）为变量 x_1, x_2, x_3 的三维空间图及三维透视图，破产公司对应的点集使用红色高亮表示。从图中可以看出，大部分破产公司的 x_1, x_2, x_3 小于非破产公司。因此，可以

通过 x_1, x_2, x_3 的大小来辨别破产公司和非破产公司, x_1, x_2, x_3 较大的为非破产公司。

令破产公司组成的样本由上标 b 表示, 非破产公司组成的样本由上标 c 表示, 由计算可得(精确到三位小数):

$$\overline{X^b} = (-0.069, -0.081, 1.367)'$$

$$S_3^b = \begin{pmatrix} 0.044 & 0.028 & 0.034 \\ 0.028 & 0.021 & 0.026 \\ 0.034 & 0.026 & 0.164 \end{pmatrix}$$

$$\overline{X^c} = (0.235, 0.056, 2.594)'$$

$$S_3^c = \begin{pmatrix} 0.047 & 0.009 & 0.075 \\ 0.009 & 0.002 & 0.009 \\ 0.075 & 0.009 & 1.047 \end{pmatrix}$$

可以发现, 破产公司的 x_1, x_2, x_3 平均值均比非破产公司小, 因此可以通过比较 x_1, x_2, x_3 的大小来划分破产公司和非破产公司, 例如: 小于 $(\overline{X^b} + \overline{X^c})/2$ 的为破产公司, 大于该值的为非破产公司; 通过样本均值和样本方差也可以得出变量的置信区间, 用于划分破产公司

和非破产公司, 例如, 当 $X \in (\overline{X^b} - (s_{11}^b + s_{22}^b + s_{33}^b)', \overline{X^b} - (s_{11}^b + s_{22}^b + s_{33}^b)')$ 时, 该样本为破产公司。

两类公司中的大多数观测值只会对该规律有一些数值上的变化, 不会产生重大影响。然而, 考虑多个观测值, 则有可能使破产公司的 x_1, x_2, x_3 平均值不都比非破产公司小, 因此不能通过比较 x_2 的大小来划分的规律, 产生了重大影响。