An Estiamtion of Moody's Model

Xiaotong Ding

October 19, 2014

1 Moody's Model

Concering the rating of a technology company, Moody's has a staight forward idea: to combine all the factors in linear model. In Moody's Model, ther are 10 factors, each factor makes effect to the model.
In Moody's Model,

$$Aaa = 1, Aa = 3, A = 6, Baa = 9, Ba = 12, B = 15, Caa = 18, Ca = 21$$

. This is just the a period. For example, if one company have rating score of 2, its rating will be Aa.
After the combination of all the factors that concerns, one can find that the Moody's rating is still different from just the linear combination.
Since Moody's has 10 factors to estimate the rating, we can denot:

$$\hat{Y_j} = \sum_{i=1}^{10} a_{ij} X_{ij} \tag{1}$$

Hereby, $\sum_{i=1}^{10} a_i = 1$, which is know.

The Moody's rating is

$$Y_j = \hat{Y_j} + \epsilon \tag{2}$$

Where ϵ might be a function of X_i maybe not.

2 Problems

Just as what we have discussed, this is in fact not a **statisitcs**. It is just a linear combination. However, if one wants to do the prediction or analysis, we are faced with several problems.

2.1 ϵ

The ϵ is unknown, which needs us to drive it out.

2.2 Unknow factors

Several factors are unknown. They may be familiar with Moody's, or maybe Moody's does not know as well.

3 Model

Before all, we need to try our best to get all the data we need. However, maybe we still can not find the data we want. And I found that **most of the data we known are the inside status of a company**.Therefore, my idea is to use the historical data to estimate the unknown factors. First of all, one can change the order of the variables:

$$Y_j = \sum_{i=1}^{7} X_{(ij)} + \sum_{i=8}^{10} X_{(ij)} + \epsilon_j \tag{3}$$

Hereby, in order to make the question clear, we can rearrage the order of the variables, such that

$$Y = \sum_{i=1}^{7} X_i + \sum_{i=8}^{10} X_i + \epsilon_j \tag{4}$$

Hereby, I suppose $X_i, i = 1, 2, ..., 7$ are known and the rest are unknown. Therefore, we emphasize on $X_i, i \geq 7$.
Consider X_8. X_8 might be affected by the internal ones, and it might also be affected by the outside one. Suppose

$$X_8 = f_8(X_1, X_2, X_3, ..., X_7) + \eta_8 \tag{5}$$

Hereby,$f()$ means the part of X_8 only affected by the internal factors, and η_8 are the ouside effect, which is not affected by the company itself.
In order to approxiamte $f()$, suppose

$$f_8(X_1, .., X_7) = f_8(0, ..., 0) + \sum_{i=1}^{7} \frac{\partial f}{\partial X_i} X_i + \gamma_8 \tag{6}$$

γ_8 might be related to $X_1, ..., X_7$. However, we can still supppose:

$$f_8(X_1, ..., X_7) = \sum_{i=1}^{7} b_{8i} X_i + \Gamma_8 \tag{7}$$

Hereby $\Gamma_8 = \gamma_8 + f_8(0, ..., 0)$. Denote $\theta_8 = \Gamma_8 + \eta_8$. Therefore,

$$Y_j = \sum_{i=1}^{7} (a_i + b_{8i} + b_{9i} + b_{10i}) X_{ij} + \zeta_j \tag{8}$$

Where

$$\zeta_j = \epsilon_j + \theta_{8j} + \theta_{9j} + \theta_{10j}$$

Therefore, our aim is to get b_8, b_9 and b_{10}.

4 Method

4.1 Linear Regression

4.2 Logistics Regression

4.3 Tree

5 Potential Problems

5.1 Non-normal

5.2 Correlation

5.3 What's more