An Analytic Model of Relation between Companies' Recruitment Activities and Number of Students' Application Based on Mixture Regression Model

Seiya Nagamori †
Graduate School of Creative Science and Engineering,
Waseda University, Tokyo, Japan
Tel: (+81) 3-5286-3290, Email: s.s.7-1t.n.24@akane.waseda.jp

Haruka Yamashita ††
School of Creative Science and Engineering,
Waseda University, Tokyo, Japan
Tel: (+81) 3-5286-3290, Email: h.yamashita@aoni.waseda.jp

Masayuki Goto ††
School of Creative Science and Engineering,
Waseda University, Tokyo, Japan
Tel: (+81) 3-5286-3290, Email: masagoto@waseda.jp

Tairiku Ogihara †††
Customer Action Group, Media Planning Division
Field Planning Unit, Graduate Recruitment Business Development
Recruit Career Co.,Ltd., Tokyo, Japan
Tel: (+81) 3-6835-0281, Email: tairiku_ogihara@r.recruit.co.jp

Abstract. In recent years, most Japanese university students use Internet portal sites for job-hunting activities and many companies also use them for efficient recruitment. From the companies' viewpoints, there are mainly two interests for using an Internet portal site for job hunting: to know the effect of using portal sites for the number of students' applications and to know the effect of recruitment activities on the sites.

In this study, we apply the mixture model that represents the multi-linear structure between the number of students' applications and recruitment activities. Here, it is difficult to build only an accurate regression model, because not only the recruit activities but also external factors of companies such as the brand image have large influences on the number of applications. Therefore, we apply a regression model taking account of potential factors. Moreover, we discuss the improvement of the algorithm for the mixture regression model in terms of the initial setting of the algorithm.

In order to verify the effectiveness of the method, we demonstrate an analysis by using actual data stored in an Internet portal site of job-hunting and show the relationship between the number of applications and the company activities for their recruitment.

Keywords: Mixture Model, Regression Model, Analytics, Data Mining, Job-Hunting

1. INTRODUCTION

In recent years, most students in Japan have been using Internet portal site for their job-hunting activities. Companies also use Internet portal site for their efficient recruitment activities. Companies can provide information about recruitment for many students thorough the site. Using the site, companies expect to increase the number of the applications from students. When companies use Internet portal site for recruitment activities, there are mainly two interests for using an Internet portal site for job hunting: one is the estimated value of the number of applications from students and another is the impact for the number of applications from students by performing some kind of recruitment activities.

In this study, we apply a regression model to verify the relation between the number of applications from students and recruitment activities (i.e., the objective variable is the number of application, and the explanatory variables are recruitment activities). Here, it is difficult to build an accurate regression model only by applying a general regression model. The difficulty of estimating the number of the applications that a company acquires is caused by the existence of various external factors that can affect the number of application from students, (e.g., students′ recognition for each company, business conditions, reputation, recent news about companies, etc.). In addition, basic attributes of company such as industry fields or the scale in terms of the number of employees also affect the relation between recruitment activities of company and the number of the applications from students.

Therefore, we apply a mixture model taking account of potential factors (De Veaux, 1989). For expressing difference of the linear structure between the recruitment activities and the number of applications, we build plural regression models by assuming latent classes and properly mix their models. This model improves estimate accuracy of the number of applications by fitting the analytical data. Moreover, this model discloses the differences of the effectiveness of recruitment activities among companies. It leads to a more accurate analysis in terms of explanation capability.

Moreover, we consider the initial value problem in the algorithm for the estimation of this model. The parameter of the mixture regression model can be estimated by an iterative procedure based on the expectation maximization (EM) algorithm which needs the initialization. Although the randomized initialization of parameter is usually used in the EM algorithm, it is well known the goodness of the acquired parameter depends on the initial value in the first step. Therefore, an effective way of initial value setting is discussed. In order to fit this model to recruitments data, we consider using basic attributes of company such as the industry field for initial value. In this study, the aspect model (AM) is used for a way to reflect basic attributes to initial value in the parameter estimation of the mixture regression model. By using basic attributes, a stable and high accurate model can be acquired, and we discuss the effectiveness of such initial value settings.

In order to verify the effectiveness of models, we demonstrate the analysis and processing by using actual data from on an Internet site for job-hunting.

2. PRELIMINARIES

2.1 Job-hunting of University Students in Japan

In this section, the system of job-hunting in Japan is introduced to delimitate the job-hunting problem and the background of our study. Most Japanese students start job-hunting activities when they are 3rd year students and begin working immediately after their graduation from educational institutions such as universities or graduate schools. The schedule for job-hunting is decided by the Federation of Economic Organizations in Japan; therefore, the job-hunting information for almost all companies is published at the same time.

In Japan, most students generally use Internet portal sites for job-hunting. Many Internet portal sites for job-hunting exist, and many students use multiple sites simultaneously. Typically, an Internet portal site for job-hunting provides a comprehensive service from the start to the finish of job-hunting activities; the service include the entry for an internship, the reservation of a briefing session and the reservation of interviews in the employment examination of a company. On the other hand, in order to convey the information to many students, each company publishes information such as job information and their appeal information on the Internet portal site.

2.2 Regression Model

The object of this study is to model the statistical structures between recruitment activities and the number of applications. Here, the recruitment activities are defined by implemented activities which a company can take such as the implementation of internship or information session. Each company is able to select appropriate activities in an Internet portal site. Because the effectiveness of recruitment activities may be different from one another depending on the situation of each company, the detail analysis between recruitment activities and the number of applications is desirable.

The most basic model to achieve the above object is a regression model (Bishop, 2006). Here, for l-th company out of L companies, the explanatory variable vector that represents the recruitment activities is denoted by $\boldsymbol{x}_l = (x_{l0}, x_{l1}, x_{l2}, \dots, \mathrm{x}_{lI})^{\mathrm{T}}$, the objective variable that represents the number of applications from students is defined as y_l. Here, I recruitment activities are considered as explanatory variables, and T is a transpose of a vector or a matrix. In this case, the regression model can be defined as follows:

$$y_l = \sum_{i=0}^{I} \beta_i x_{li} + \varepsilon_l \tag{1}$$

$$\varepsilon_l \sim N(0, \sigma^2) \tag{2}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_I)^{\mathrm{T}}$ represents the parameter vector with $I+1$ dimensions of the regression model and x_{l0} is identical to 1. ε_l represents the error of the l-th company's number of application from students and it follows a normal distribution of $N(0, \sigma^2)$. The parameter $\boldsymbol{\beta}$ is estimated by minimizing the residual sum of squares S^2.

$$S^2 = \sum_{l=1}^{L} (y_l - \sum_{i=0}^{I} \beta_i x_{li})^2 \tag{3}$$

3. MIXTURE REGRESSION MODEL

3.1 Regression Model and Mixture Model

Here, when we directly apply companies' data to a linear regression model shown in the section 2.2, there is a problem in the simplicity of the model. It is not possible to take into account the differences of each company's relation between the number of applications and the recruitment activities. This is because a unique parameter for the explanatory variables is estimated under the assumption of simple linearity. In other words, in a normal regression model, the same linear structure between the number of applications and recruitment activities of each company is assumed regardless of the characteristic difference of companies. However, the relation between the number of applications and recruitment activities of each company are usually different depending on companies' characteristics. For example, even if two companies carry out the same recruitment activities, the impact to the number of applications can be different. The difference is caused by not only basic attributes of company such as the industry field or the scale in terms of the number of employees but also external factors such as recent reputation or business conditions. However, since these factors cannot be observed in an Internet portal site, it is difficult to take account into the model.

In this study, we apply a mixture regression model (Faria and Soromenho, 2010; Jones and McLachlan, 1992; Turner, 2000) which is one of latent class models. The latent class model allocates similar companies to the same latent class and a different linear regression is assumed for each latent class. Then, by mixing these linear regressions, the different structure depending on characteristics of each company is expressed flexibly.

This analysis enables to figure out the relation between the number of applications and recruitment activities with taking account of latent characteristics of each company. Furthermore, we can conduct more accurate analysis about heterogeneity of recruitment activities by introducing latent classes for companies. Appropriate estimation of the latent class model leads to build a model that has high estimate accuracy for the number of students' applications and high interpretability for the relation between the number of applications from students and recruitment activities of companies.

Learning procedure of this model consists of parameter estimation of individual regression models and weights updated by EM algorithm.

3.2 Overview of Mixture Regression Model

The mixture regression model is given by a liner combination of probability density functions $P_k(y_l|\boldsymbol{x}_l)$ defined for K latent classes, when a set of latent classes is defined as $\mathrm{Z} = \{z_k\colon\ 1 \le k \le K\}$.

Here, the $I+1$ dimensional parameter of the regression model on k-th latent class out of K latent classes is denoted by $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{Ik})^{\mathrm{T}}$. In this case, the mixture regression model can be represented by the equation (4).

$$P(y_l|\boldsymbol{x}_l) = \sum_{k=1}^{K} w_{lk} P_k(y_l|\boldsymbol{x}_l) \tag{4}$$

w_{lk} denotes the weight of l-th company to k-th latent class. The sum of weights for all latent classes is equal to 1. The probability density function $P_k(y_l|\boldsymbol{x}_l)$ is the normal distribution assumed for each latent class and given by the equation (5).

$$P_k(y_l|\boldsymbol{x}_l) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp(-\frac{(y_l - f_k(\boldsymbol{x}_l))^2}{2\sigma_k^2}) \tag{5}$$

Here, the mean $f_k(\boldsymbol{x}_l)$ is represented by the equation (6).

$$f_k(\boldsymbol{x}_l) = \sum_{i=0}^{I} \beta_{ik} x_{li} \quad (6)$$

3.3 Method of Parameter Estimation

In this section, we represent the way to estimate the parameter of regression models $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{Ik})^{\mathrm{T}}$, weights w_{lk} and variances of each latent class σ_k^2. For the parameter estimation, the EM algorithm can be applied (Dempster et al., 1977; Wedel and DeSarbo, 1995). The EM algorithm consists of Expectation (E) step and Maximization (M) step.

In the E-step, the weight parameters w_{lk} are updated. It is desirable to allocate a high weight w_{lk} to the regression model which has high estimation accuracy for l-th company. In the E-step, each weight w_{lk} is updated by the equation (7).

$$w_{lk} = \frac{P(z_k) P_k(y_l|\boldsymbol{x}_l)}{\sum_{k'=1}^{K} P\,(z_{k'}) P_{k'}(y_l|\boldsymbol{x}_l)} \quad (7)$$

In the M-step, the parameters of regression models $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{Ik})^{\mathrm{T}}$ and variances of each latent class σ_k^2 are updated. In this model, each company is allocated to latent classes probabilistically and each regression model is estimated by using the weights w_{lk}. The regression parameters $\boldsymbol{\beta}_k$ are estimated by minimizing the residual sum of squares considering the weights S_k^2 shown in the following equation (8).

$$S_k^2 = \sum_{l=1}^{L} w_{lk} (y_l - f_k(\boldsymbol{x}_l))^2 \quad (8)$$

When estimating the parameter $\boldsymbol{\beta}_k$, each regression model is learned to adapt to the high weighted data. As the result, it is possible to reflect the characteristics of companies to the parameter estimation of each regression model.

The variances of each latent class σ_k^2 are estimated by the equation (9). In addition, $P(k)$ is calculated by the equation (10).

$$\sigma_k^2 = \frac{\sum_{l=1}^{L} w_{lk} (y_l - f_k(\boldsymbol{x}_l))^2}{\sum_{l=1}^{L} w_{lk}} \quad (9)$$

$$P(k) = \frac{\sum_{l=1}^{L} w_{lk}}{L} \quad (10)$$

3.3 Algorithm of Mixture Regression Model

Here, the algorithm of model construction is shown as follows:

[Steps in mixture regression model]

Step1) Initialize belonging probabilities of each company to latent classes w_{lk} (weight) randomly.

Step2) Estimate the parameter of regression models $\boldsymbol{\beta}_k$ for each latent class by minimalizing S_k^2 shown in the equation (8).

Step3) Calculate parameters, σ_k^2 and $P(k)$ shown in the equations (9)-(10) and update the weight w_{lk} by the equation (7)

Step4) End if convergence condition is satisfied, otherwise return to Step2.

4. DISCUSSION FOR APPLICSTION TO ACTUAL DATA

4.1 Consideration of Initial Value in the Algorithm

The estimation accuracy of the mixture regression model depends on the initial value of w_{lk} which is set in Step1 (Govaert and Nadif, 1996). For efficient setting of initial value, we introduce the aspect model (AM) as a method of company clustering stochastically (Hofmann, 1999; Hofmann 1999; Hofmann and Puzicha, 1999; Hofmann, 2001). AM is a statistical model that assumes the existence of a discrete latent class among factors.

On the other hand, because of the factors that affect the linear structures between recruitment activities and the number of applications, basic attributes of companies such as the industry field or the scale in terms of the number of employees should be considered. Therefore, we conduct company clustering by using basic attributes information stochastically. Then, the estimated belonging probabilities to latent classes are set as initial values of w_{lk}. As a result, more stable and accurate estimation of mixture regression model can be expected by the consideration of basic attributes to regression mixture model.

4.2 Aspect Model

Here, let the number of the basic attribute items of company be J. The j-th attribute takes a value on a discrete set $D^j = \{d_{n_j}^j : 1 \le n_j \le N_j\}$. $d_{n_j}^j$ means the n_j-th element of j-th basic attribute. In this case, the AM can be stochastically represented by the equation (11).

Table 1: Comparison results of the absolute average error obtained by each model.

	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$
Piecewise linear regression model	110.206	74.322	55.231	44.801	37.558
Mixture regression model (random)	103.161	59.771	45.457	41.865	37.455
Mixture regression model (AM)	103.156	55.634	50.812	40.268	36.202

$$P(d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J) = \sum_{k=1}^{K} P(z_k) \prod_{j=1}^{J} P(d_{n_j}^j | z_k) \quad (11)$$

In the equation (11), the parameters $P(z_k)$ and $P(d_{n_j}^j|z_k)$ can be estimated by the EM algorithm. By using estimated values, the belonging probabilities of companies to each latent class are calculated by the following equation (12).

$$P(z_k|\boldsymbol{x}'_l) = \frac{\{\prod_{j=1}^{J} P(x'_{lj}|z_k)\}P(z_k)}{\sum_{k=1}^{K}\{\prod_{j=1}^{J} P(x'_{lj}|z_k)\}P(z_k)} \quad (12)$$

The vector that represents the basic attributes of l-th company is denoted by $\boldsymbol{x}'_l = (x_{l1}, x_{l2}, \dots, x_{lJ})$. The conditional probability $P(z_k|\boldsymbol{x}'_l)$ shows the characteristics of basic attributes of company. Therefore, this estimated probability can be utilized to set an initial value of w_{lk}.

5. DATA ANALYSIS FOR ACTUAL DATA

In this section, we show how to apply the mixture regression model to actual data stored in an Internet portal site. For the evaluation of the model, we introduce the absolute average error between the estimated value and actual value of the number of application as estimation accuracy. If the prediction accuracy improves, the estimated model can be judged as a better model that is good representation of the relation between the number of applications and recruitment activities and the analysis based on this model is better from the viewpoint of reliability. In the section 5.3, we show a case of data analysis as an application of the model and discuss the results.

5.1 Data Set and Analysis Condition

In this analysis, we used the data accumulated on an Internet portal site which is used by university students who graduated from universities in 2015. We just use the data of about 6,000 companies whose number of application from students are from 100 to 1,000 in an Internet portal site. We use the number of applications from students as the objective variable and the six options about recruitment activities that can be taken by each company as explanatory variables ($I = 6$). Four variables of basic information of each company are used for setting an initial value ($J = 4$).

The number of latent classes K is set as 2, 3, 4, 5 and 6. As a comparison model, we use the piecewise linear regression model (Malash and El-Khaiary, 2010) with K ($K = 2$, 3, 4, 5 and 6) divided regions. In the piecewise linear regression model, the data space is divided into K regions and the regression models are built for each region separately. For making divided regions, we use the way that the range of the amount of applications from students is equally divided into K regions.

As an evaluation of models, the absolute average error between estimated value $\hat{y}_l$ by using each model and actual value y_l is used. The absolute average error is defined as follows:

$$\text{absolute average error} = \frac{\sum_{l=1}^{L} |y_l - \hat{y}_l|}{L} \quad (13)$$

5.2 Evaluation of Mixture Regression Model

We estimate each model and evaluate the absolute average error between estimated value and actual value. Table 1 and Figure 1 present the results of this experiment. The result shows average of 100 different initial values.

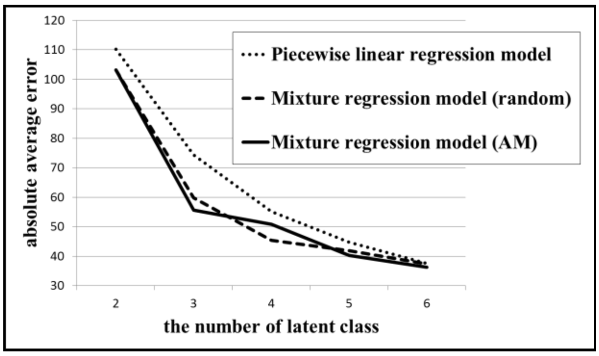

Figure 1: Comparison results of the absolute average error obtained by each model.

The interior of the parenthesis ( ) in Table 1 means the way to set initial values. From Table 1 and Figure 1, it can be seen that the mixture regression models have good results in terms of estimation accuracy. The mixture regression models represent the characteristics of companies well by building plural latent classes stochastically. In the mixture regression model, the weight w_{lk} is learned as it takes high points if l-th company is suitable for k-th regression model in the point of estimation of the number of applications from students. As a result, an accurate estimation model is built. Moreover, the mixture regression models whose initial values are set by AM show the good result in many situations. The effect of weight setting for latent classes which take account of basic attributes of companies is shown. The result suggests that the basic attributes of company affect the linear structure between the number of applications and recruitment activities. However, the mixture regression model whose initial values are set by AM is worse result than the normal mixture regression model when the number of latent classes is set as $K = 4$. This suggests that company clustering using basic attributes by AM is not conducted successfully. There is the number of latent classes that are not suitable for company clustering by basic attributes, but it is found that accurate models can be generally constructed by taking account with basic attributes of companies.

5.3 Application of Mixture Regression Model

Here, we discuss the interpretation about the estimated mixture regression model. By analyzing the result from estimated parameters, an analyst can set the relevant number of the latent classes. The number of latent classes K can be set to make it easier to interpret the results in accordance with the purpose. The estimated parameters of regression models are shown in the table 2. From some condition, the number of the latent classes is set as $K = 5$ and the initial value is set by belonging probabilities of the estimated AM.

Table 2: Parameters of regression models estimated from the mixture regression model (AM).

Latent class k	β_{0k}	β_{1k}	β_{2k}	β_{3k}
1	153.76	-2.16	0.13	-4.05
2	276.80	15.63	0.45	31.22
3	372.38	22.09	-0.10	-9.30
4	584.20	-8.23	1.53	42.99
5	950.07	9.14	0.05	10.03

Latent class k	β_{4k}	β_{5k}	β_{6k}	Mixture ratio $P(z_k)$
1	6.07	-9.42	-7.64	0.16
2	618.87	-26.97	-30.52	0.23
3	-128.19	287.11	-76.11	0.26
4	-102.20	-176.22	-74.85	0.25
5	-234.53	-60.62	-20.75	0.10

From Table 2, the parameters are different between the numbers of latent classes. This result shows that the recruitment activity that affects the number of application from students can be different between each latent class. This suggests that recruit activities for increasing the number of application are different for each company.

For example, a company A has the weights w_{lk} shown in Table3.

Table 3: Estimated weights w_{lk} for a company A

w_{l1}	w_{l2}	w_{l3}	w_{l4}	w_{l5}
0.00	0.00	0.24	0.00	0.76

The company A has high weight values on the latent classes z_3 and z_5. When analyzing the characteristics of the company A, we need to focus on the latent class z_3 and z_5. By using Table 2 together, the company A can realize their recruitment activities′ impact for the application from students.

Here, the effect vector of l-th company′s recruitment activities are defines as follow:

$$\tilde{\boldsymbol{\beta}}_l = (\sum_{k=1}^{K} w_{l1}\beta_{1k}, \sum_{k=1}^{K} w_{lk}\beta_{2k}, \ldots, \sum_{k=1}^{K} w_{lk}\beta_{Ik})^{\mathrm{T}} \tag{14}$$

From equation (14), the effect vector of the company A's recruitment activities is shown by $\widetilde{\boldsymbol{\beta}}_{company\ A} = (12.25, 0.01, 5.39, -209.01, 22.83, -34.04)^{\mathrm{T}}$. From this analysis, the biggest effect of the company A's recruitment activity is fifth recruitment activity. On the other hand, the smallest effect of the company A's recruitment activity is fourth recruitment activity. By considering these results, cost and effort of each recruitment activity, each company can consider more efficient plans.

In summary, an analyst can get effective information for planning recruitment activities by analyzing regression parameters and weights w_{lk} of each company.

6. CONCLUSION AND FUTURE WORKS

In this study, we apply the mixture regression model for the analysis about the relation between the number of applications from students and recruitment activities of companies. In the evaluation of the applied models and data analysis, we show the effectiveness of estimated suitable latent classes for improving the estimation accuracy. It is a useful way to express characteristics of companies more flexibly by applying the regression mixture model. In addition, the way to set initial value is discussed and its effectiveness is also clarified.

Future works are improving estimation accuracy and extending this model for a method of support planning company's recruitment activities. Moreover, it is desirable to predict the number of applications from students for the companies that publish information about recruitment newly by extension of this model.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to all members of Goto laboratory, Waseda University, and members of Recruit Career Co.,Ltd. who support us for their helpful comments in this research. A part of this study was supported by JSPS KAKENHI Grant Numbers 26282090 and 26560167.

APPENDIX A: EM algorithm of AM

Here, we present a method to estimate the parameters of the AM by using the EM algorithm.

The EM algorithm is a method for parameter estimation that locally maximizes the likelihood function for incomplete data. The EM algorithm consists of two steps: the Expectation (E) step and the Maximization (M) step. In order to estimate the parameters, these two steps are repeated until convergence of the log-likelihood is achieved.

Log-likelihood function

$$LL = \sum_{\boldsymbol{n}} \delta \log P(d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J) \tag{15}$$

Expectation Step (E-step)

$$P(z_k | d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J) = \frac{P(z_k) \prod_{j=1}^{J} P(d_{n_j}^j | z_k)}{\sum_k P(z_k) \prod_{j=1}^{J} P(d_{n_j}^j | z_k)} \tag{16}$$

Maximization Step (M-step)

$$P(z_k) = \frac{\sum_k \delta P(z_k | d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J)}{\sum_{\boldsymbol{n}} \sum_k \delta P\left(z_k \middle| d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J\right)} \tag{17}$$

$$P(d_{n_j}^j | z_k) = \frac{\sum_{\boldsymbol{n} \neq n_j} \delta P(z_k | d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J)}{\sum_{\boldsymbol{n}} \delta P(z_k | d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J)} \tag{18}$$

$$\delta = \begin{cases} 1 & \text{(event of } \left(d_{n_1}^1, d_{n_2}^2, \dots, d_{n_J}^J\right) \text{is occur} \\ 0 & \text{(otherwise)} \end{cases} \tag{19}$$

Here, $\sum_{\boldsymbol{n}}$ represents $\sum_{n_1} \sum_{n_2} \cdots \sum_{n_J}$ for the sake of simplicity.

REFERENCES

Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*, Springer.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Society, Series B*, **39**, 1-38.

De Veaux, R. D. (1989). Mixtures of linear regressions. *Computational Statistics & Data Analysis*, **8**, 227-245.

Faria, S., and Soromenho, G. (2010) Fitting mixtures of linear regressions. *Journal of Statistical Computation and Simulation*, **80**, 201-225.

Govaert, G., and Nadif, M. (1996). Comparison of the mixture and the classification maximum likelihood in cluster analysis with binary data. *Computational Statistics & Data Analysis,* **23**, 65-81.

Hofmann, T. (1999) Probabilistic Latent Semantic Analysis. *Proc. the Fifteenth Conference on Uncertainty in Artificial Intelligence*, 289-296.

Hofmann, T. (1999) Probabilistic Latent Semantic Indexing. *Proc. the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 50-57.

Hofmann, T., and Puzicha, J. (1999) Latent Class Models for Collaborative Filtering. *Proc. 16th International Joint Conference on Artificial Intelligence*, 688-693.

Hofmann, T. (2001) Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Machine Learning Journal*, **42**, 177–196.

Jones, P. N., and McLachlan, G. J. (1992). Fitting finite mixture models in a regression context. *Australian Journal of Statistics*, **34**, 233-240.

Malash, G. F., and El-Khaiary, M. I. (2010). Piecewise linear regression: A statistical method for the analysis of experimental adsorption data by the intraparticle-diffusion models. *Chemical Engineering Journal*, **163**, 256-263.

Turner, T. R. (2000). Estimating the propagation rate of a viral infection of potato plants via mixtures of regressions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **49**, 371-384.

Wedel, M., & DeSarbo, W. S. (1995). A mixture likelihood approach for generalized linear models. *Journal of Classification*, **12**, 21-55.