International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2020; 5(5): 01-06

www.mathsjournal.com
Received: 01-07-2020
Accepted: 04-08-2020

Fred Nyamitago Monari
Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Kenya

Dr. Joseph Kyalo Mungatu
Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Kenya

George Otieno Orwa
Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Kenya

Romanus Odhiambo Otieno
Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Kenya

Corresponding Author:
Fred Nyamitago Monari
Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Kenya

Comparison between weighted adjustment, diagonal adjustment, quasi optimization and Gibbs sampler methods in deriving the generator matrix of an 8×8 credit transition matrix

Fred Nyamitago Monari, Dr. Joseph Kyalo Mungatu, George Otieno Orwa and Romanus Odhiambo Otieno

Abstract
A transition matrix P is said to be embeddable if it has a generator matrix Q such that P =exp (Q) If the approximated transition matrix $\hat{P}$ is embeddable, then the estimator $\hat{Q}$ can be got for the generator matrix Q. What of cases when $\hat{P}$ is not embeddable? This paper will show how to evaluate $\hat{P}$ and find estimatorˆin cases where Pˆ is embeddable (Kingman, 1962). This problem can be solved by using four methods namely

1. Diagonal and Weighted adjustments Method
2. The Generator Quasi-Optimization method
3. The EM logarithm Method
4. The Gibbs sampler (Markov Chain Monte Carlo Method)

Later after sampling each of the above methods, an L norm will be performed on the results so as to deduce which method is the best for estimating the Generator Matrix for an 8×8 Credit transition Matrix.

Keywords: Series, converges, generator, credit, transition, matrix, parameter, embeddable

1. Introduction

1.1 The Maximum Likelihood Estimator

Given any observations set, the maximum likelihood estimator is one of the methods for estimating statistical model parameters. It does this by getting the values of the parameter which maximize the likelihood function. The estimates are referred to as Maximum Likehoood function (MLE). The parameter $Q = [q_{ij}]_{i,j} \in S$ is the generator matrix. The likelihood function of Q is given by;

$$L_c(Q, Y) = \prod_{i=1}^{l} \prod_{jneq1} q_{ij}^{N_{(ij)(T)}R_i(T)} \qquad (1)$$

Where

- (C) indicates the continuous time observations.
- $N_{(ij)}\ (T)$ is the translation number from state i to state j in $[0, T]$ time interval.
- $R_i(T) = \int_0^t 1\{Y(s) = i\}ds$

Taking the logarithm of equation 1 above and then its partial derivative with respect to q_{ij} and equating it to 0 gives the MLE of Q as;

$$\hat{q}_{ij} = \frac{N_{ij}(T)}{R_i(t)} \qquad (2)$$

The MLE of parameter $P^{\hat{}}$ is given as $\hat{P} = \left(\hat{p}_{ij}\right)_{ij\in S}$ Such that;

$$\hat{P}_{ij} = \frac{K_{ij}(n)}{K_i(n)} \quad (3)$$

Where

$$K_i(n) = \sum_{j=1}^{l} K_{ij}(n)$$

If the $\widehat{P}$ shown by equation 3 is embeddable, then the MLE $\widehat{Q}$ of the generator matrix exists and is obtained by $\widehat{Q} = \log\widehat{P}$. If $\widehat{P}$ is not embeddable, then there is either existence of Q and other methods of obtaining $\widehat{Q}$ using the transition estimator $\widehat{P}$ or else there is no existence of MLE $\widehat{Q}$.

2. Materials and Methods

2.1 The Diagonal and Weighted adjustment

Assuming $l \times l$ transition matrix estimator $\hat{P}$ given by equation (3) above is not embeddable. The next step is to obtain a matrix $\hat{Q}$ such that $\hat{P} = \exp(\hat{Q})$. The first thing to do is to be sure that the log function exists for $\hat{Q}$ using the theorem below.

2.1.1 Theorem

Let P be $l \times l$ TM and let $F = max\{(a-1)^2 + b^2, a + b\}$ is an Eigen value of $P, a, b \in R$. Assuming that $F < 1$, then series;

$$\hat{Q} = \log P = (\hat{P} - 1) - \frac{(\hat{P}-1)^2}{2} - \frac{(\hat{P}-1)^3}{3} - \cdots \quad (4)$$

will converge geometrically giving matrix $\hat{Q}$ having row sums equal to 0 and $\exp(\hat{Q}) = \hat{P}$. The condition $F < 1$ will not be needed if the series absolutely converges. Due to embedding problem, it is not definite that the condition of generator $matrix\ \hat{Q}$ to have positive off-diagonal entries though it satisfies $\exp(\hat{Q}) = \hat{P}$.

2.1.2 The Diagonal Adjustment

Let generator matrix Q^{DA} elements given as q_{ij}^{WA} which will be obtained using the Diagonal Adjustment (DA). The negative value of $\hat{Q}$ will be replaced by 0 since they are usually very small and then the difference added to the diagonal entries so as to preserve the property of row sums totalling to 0. The DA method is given by;

$$q_{ij}^{DA} = \max(\widehat{q_{ij}}, 0), i \neq j$$
$$q_{ii}^{DA} = -\sum_{j} q_{ij}^{DA} \quad (5)$$

2.1.3 The Weighted Adjustment

Let the generator matrix Q^{WA} elements be given as q_{ij}^{WA} which will be obtained using the Weighted Adjustment method. Let;

$$G_i = |\widehat{q_{ii}}| + \sum_{j\neq i} \max(\widehat{q_{ij}}, 0), B_i = \sum_{j\neq i} \max(\widehat{q_{ij}}, 0)$$

$$q_{ij}^{WA} = \begin{cases} 0, if\ i \neq j\ and\ \widehat{q_{ij}} \leq 0 \\ \widehat{q_{ij}} - B_i|\widehat{q_{ij}}|/G_i \\ \widehat{q_{ij}}, otherwise\ if\ G_i = 0 \end{cases} \quad (6)$$

Since $\sum_{j=1}^{l} \widehat{q\_ij} = 0, then\ G_i \geq B_i, hence\ that\ q_i^{WA} \geq 0\ \forall i \neq j$

2.2 Generator Quasi-Optimization method

Using transition matrix P given in figure 4.2 above, $Q = |\hat{q}_{ij}|_{i,j\in S}$ is set to be solution of $\hat{Q} = \log\hat{P}$. This Generator Quasi-Optimization method is used to get the generator matrix approximation similarly to getting a maximization problem solution given as;

$$\min_{Q\in Q} ||Q - \hat{Q}|| \quad (7)$$

such that the set of all generator Matrices is given by Q and it is also the Euclidean norm ||.||. This problem is solved row by row since the Q conditions () are closed on each row. Next solve () which is equal to solving independent minimization problems of l given as;

$$\min_{q\in(l)} \sum_i^l (p_i - q_i)^2 \quad (8)$$

such that $\hat{Q}$ has row vector $p = (p_1, \ldots \ldots, p_l)$ also permuted as $q_i q_{i+1}$ meaning that the diagonal element in this row is q_i. Also;

$$p(l) = \{q \in \mathbb{R} | \sum q_i = 0, q_i \leq 0, q_i \geq 0, for\ i \geq 2\} \quad (9)$$

2.3 The EM logarithm

Let a data set which is complete be denoted by $Y = (X, Z)$. Expectation Maximization also known as EM is a method of obtaining the MLE when only X observation exists (Pfeuffer, Reis, *et al*., 2018). It comprises of two steps;

Step 1: The E-step
Involves evaluating the conditional expectation of $\log(L(\theta, Y))$ when given X and MLE θ_0 i.e computation of $\mathbb{E}[\log L(\theta, Y)|X, \theta_0]$.

Step 2: The M-step
Involves obtaining the new MLE θ by getting the maximization of $E[\log L(\theta, Y)|X, \theta_0]$. Then setting $\theta_0 = \theta$ and repeating the two steps till the sequence converges. The maximum likelihood function of the data set Y which is continuous is given by equation 1.

The E-step: From equation 1 and given the initial generator matrix Q_0;

$$\mathbb{E}Q_0[LogL_{(C)}(Q, Y)|X, Q_0] = \sum_{i=1}^{l}\sum_{j\neq i} \log(q_{ij})\, \mathbb{E}Q_0[N_{ij}(T)|X] - \sum_{i=1}^{l}\sum_{j\neq i} q_{ij}\mathbb{E}_{Q_0}[R_i(T)|X] \quad (10)$$

Hence there is need to evaluate $EQ_0[N_{ij}(T)|X]$ and $EQ_0[R_i(T)|X]$ and since there is;

$$\mathbb{E}_{Q_0}[N_{ij}(T)|X] = \sum_{k=0}^{n-1} \tilde{F}_{x_k x_{k+1}}^{ij}(t_{k+1} - t_k) \quad (11)$$

Such that

$$\tilde{F}_{kl}^{ij}(t) = \mathbb{E}[N_{ij}(t)|Y(t) = l, Y(0) = k] \quad (12)$$

And

$$\mathbb{E}_{Q_0}[R_i(T)|X] = \sum_{k=0}^{n-1} \widetilde{M}_{X_k X_{k+1}} (t_{K+1} - t_k) \quad (13)$$

Such that

$$\widetilde{M}_{kl}^{i} = \mathbb{E}[|R_i(t)|Y(t) = l, Y(0) = k] \quad (14)$$

is sufficient to evaluate $\widetilde{M}_{kl}^{i}$ and $F_{kl}^{ij}(t)$.
$\lambda = \max_{i\in S}(-q_{ii}, 0)$ is chosen and $B = I + \lambda^{-1}Q_0$ defined. Letting $e_j, e_j^{!}$to denote the unit vector with j^{th} coordinate equalling to 1 and getting its transpose, it is given that,

$$\widetilde{M}_{kl}^{i}(t) = \frac{\widetilde{M}_{kl}^{i}(t)}{e_k^{t} e^{Q_0 t} e_l} \quad (15)$$

Such that;

$$M^{i}(t) = \left[M_{kl}^{i}(l)\right]_{k,l\in S} = e^{-\lambda t}\lambda^{-1}\sum_{n=1}^{\infty}\frac{(\lambda t)^{n+1}}{(n+1)!}\sum_{s=0}^{n} B^{s}(e_i, e_i^{!})B^{n-s} \quad (16)$$

and that,

$$\widetilde{F}_{kl}^{ij}(t) = \frac{F_{kl}^{ij}(t)}{e_k^{l} e^{Q_0 t} el} \quad (17)$$

Such that;

$$F^{ij}(t) = \left[F_{kl}^{ij}(t)\right]_{k,l\in S} q_0 ij e^{-\lambda t}\lambda^{-1}\sum_{n=1}^{\infty}\frac{(\lambda t)^{n+1}}{(n+1)!}\sum_{s=0}^{n} B^{s}(e_i, e_i^{!})B^{n-s} \quad (18)$$

The M-Step
It is seen that $\tilde{Q} = [\tilde{q}_{ij}]_{i,j\in S}$ where,

$$\tilde{q}_{ij} = \frac{\mathbb{E}_{Q_0}[N_{ij}(T)|X]}{\mathbb{E}_{Q_0}[R_i(T)|X]} \quad (19)$$

Maximizes $\mathbb{E}_{Q_0}[\log L_{(C)}(Q,Y)|X, Q_0]$ hence the new MLE is $\tilde{Q}$. Putting $\tilde{Q} = \tilde{Q}_0$ and repeating this till there is convergence of sequence. The series of generator matrices $\{Q_k\}_{k=1}^{K}$, obtained will depend on the initial matrix generator Q_0 choice hence it is prudent to select it in a way that the det (e^{Q_k}) is greater than 0.

2.4 The Gibbs sampler (Markov Chain Monte Carlo Method)
This section will demonstrate how to make use of Monte Carlo Chain estimation methods. There are various Monte Carlo Markov Chain methods and the Gibbs Sampler is one of them which will be employed here. The Gibbs Sampler method avoids the non-existence problem of MLE and also makes it very easy to compute. The Gibbs sampler is a method that makes use of samples from joint distribution of some conditional distribution. To understand how this algorithm works its steps are outlined as; Suppose K samples are obtained of $X = (x_1, \dots, x_n)$ from the joint distribution $p(x_1, \dots, x_n)$. Let the ith sample be given as $X^{i} = (x_1^{(0)}, \dots, x_n^{(0)})$, then the algorithm is:

1. Initialize $X^{0} = (x_1^{(0)}, \dots, x_n^{(0)})$ for $t = 0$.
2. For $i = 0, \dots, k$ sample $x^{(i+1)}$ by sampling each component of $x_j^{(i+1)}$, $j = 1, \dots, n$ and evaluating it using the probability distribution,

$$p(x_j^{(i+1)}|x_1^{(i+1)}, \dots, x_{j-1}^{(i+1)}, x_{j+1}^{(i+1)}, \dots, x_n^{(i)}$$

2.4.1 Bayes Theorem
Given equation;

$$p(\phi|Y) = \frac{p(Y|\phi)p(\phi)}{p(Y)} \quad (20)$$

ignoring the constant p(Y) and utilizing (Israel *et al*., 2001);

$$p(\phi|Y)\alpha p(Y|\phi)p(\phi) \quad (21)$$

Where α denotes the proportion, $p(\phi)$ the prior distribution and $p(\phi|Y)$ the posterior distribution which is the conditional distribution of parameter ϕ when given data Y. Consider taking a complete data as having discreet time observations $X = \{Y(t_1), \dots, Y(t_n)\}$ observed at time $(t_1 = 0, \dots, t_n = T$. An application of Q and J Gibbs sample will be done by drawing J from (Q, X) and Q from (J, X) hence $\{Q^{(k)}, J^{(k)}\}_{k=1}^{k}$ will be obtained where Q_s denote the generator matrices and J_s denote the simulated Markov Chain samples. Given the prior distribution of Q, the gamma distribution proposed is;

$$p(Q) = \prod_{i=1}^{l}\prod_{j\neq i} q_{ij}^{\alpha_{ij}-1} e^{q_{ij}-\beta_i} \quad (22)$$

Where $\alpha_{ij} > 0, i, j \in S$ and $\beta_i > 0, i \in S$. Using the likelihood function, given the complete data in equation 1, the posterior of Q is given as;

$$p(Q|J, X) = p(Q|J) \propto p(Q)L_{(c)}(Q, Y) = \prod_{i=1}^{l}\prod_{j\neq i} q_{ij}^{N_{ij}(T)+\alpha_{ij}-1} e^{-q_{ij}(R_i(T)+\beta_i)} \quad (23)$$

The Gibbs sampler algorithm is summarized as below

1. Begin by constructing the initial generator matrix Q_0 by getting $q_{ji,0}$ from the prior distribution having $j \neq i$ (Hanks, 2016).
2. Perform the steps below K times;
   a. Perform a simulation of continuous Markov Chain sample J having generator matrix Q so that a realization of all observations is achieved.
   b. Evaluate $N_{ij}(T)$ and $R_i(T)$ from the Markov chain simulated.
   c. Evaluate a new Q by getting q_{ij} from the posterior distribution $\Gamma(+\alpha_{ij}, 1/(+\beta_i))$.
   d. This new Q is saved and is used in the next simulation.

3. Let $Q_1, \dots, Q_k$ be generator matrices got using Gibbs sampler algorithm and take the mean of N proportions. Hence the estimator $\tilde{Q}$ of the generator matrix Q will be given by;

$$\tilde{Q} = \frac{1}{K-N}\sum_{i=N+1}^{K} Q_i \quad (24)$$

3. Results
Given the data below got from Standard and Poors website https://cerep.esma.europa.eu/cerep-web/statistics/transitionMatrice.xhtml.

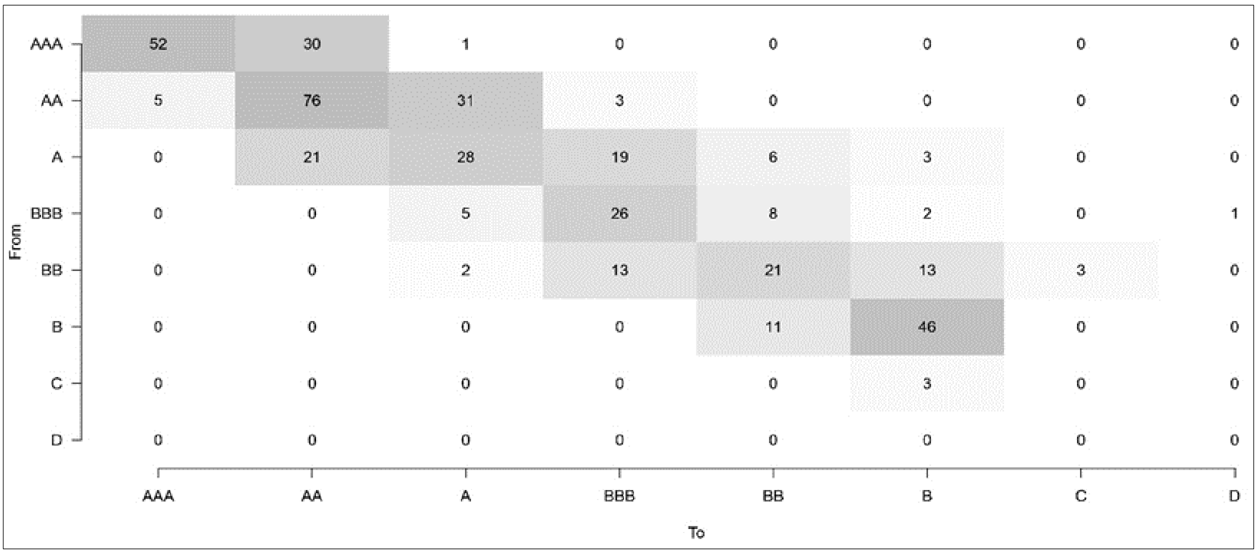

Fig 1: Data from Standard and poors website

The transition matrix P for data from figure 1 is given as;

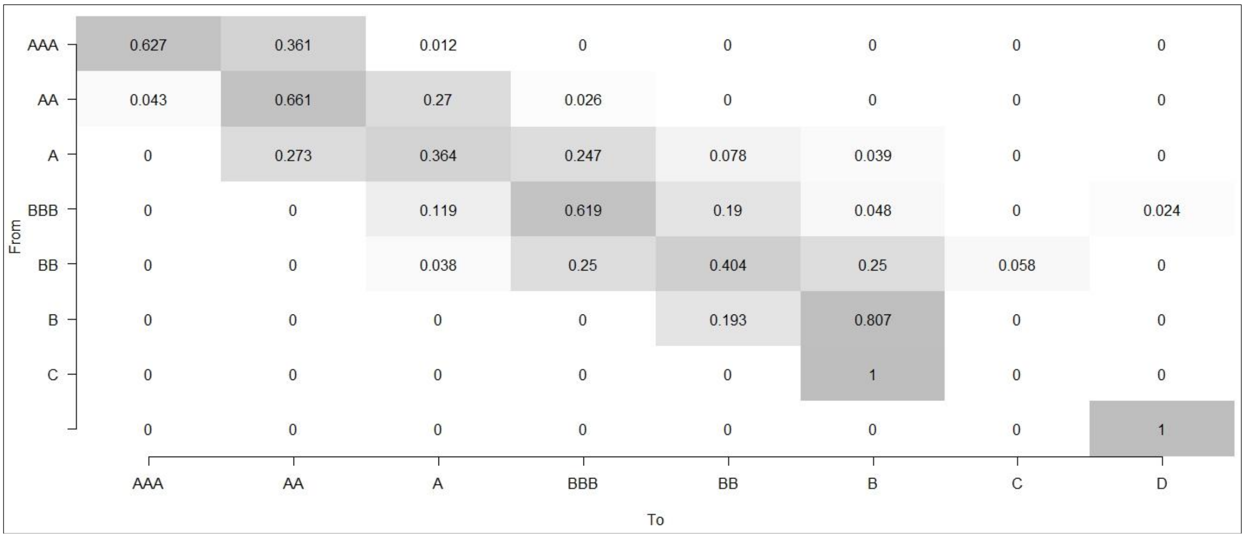

Fig 2: Transition matrix P plot

Using equation 5, the generator matrix qDA using Diagonal Adjusment method (DA) wil be given as;

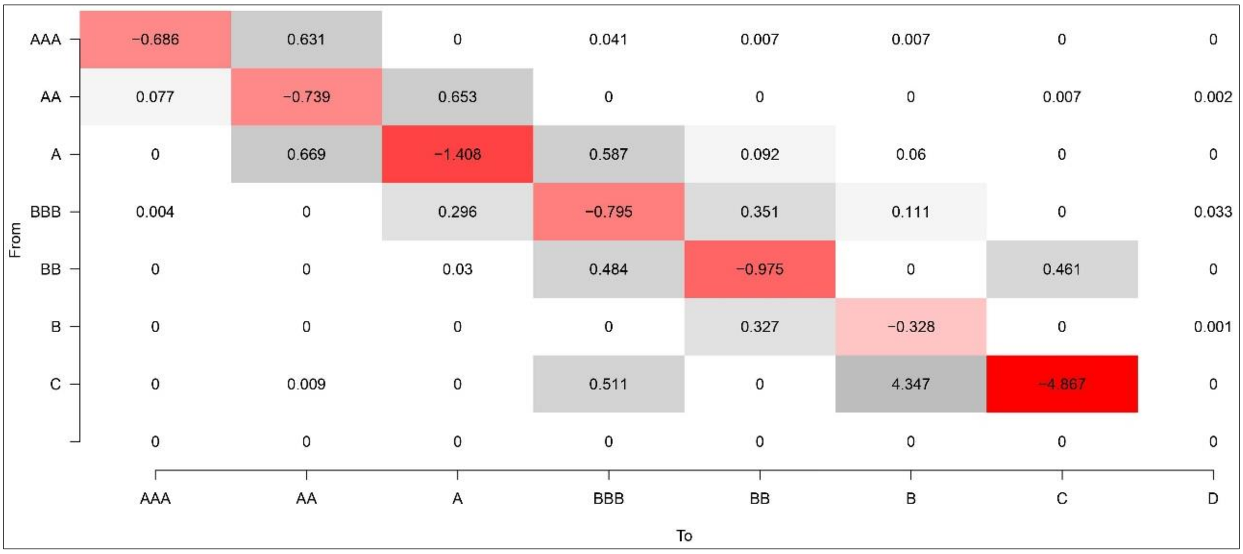

Fig 3: Generator Matrix obtained using Diagonal Adjustment Method

Running equation 6 for Weighted Adjustment (WA) method using data from figure 1 yields;

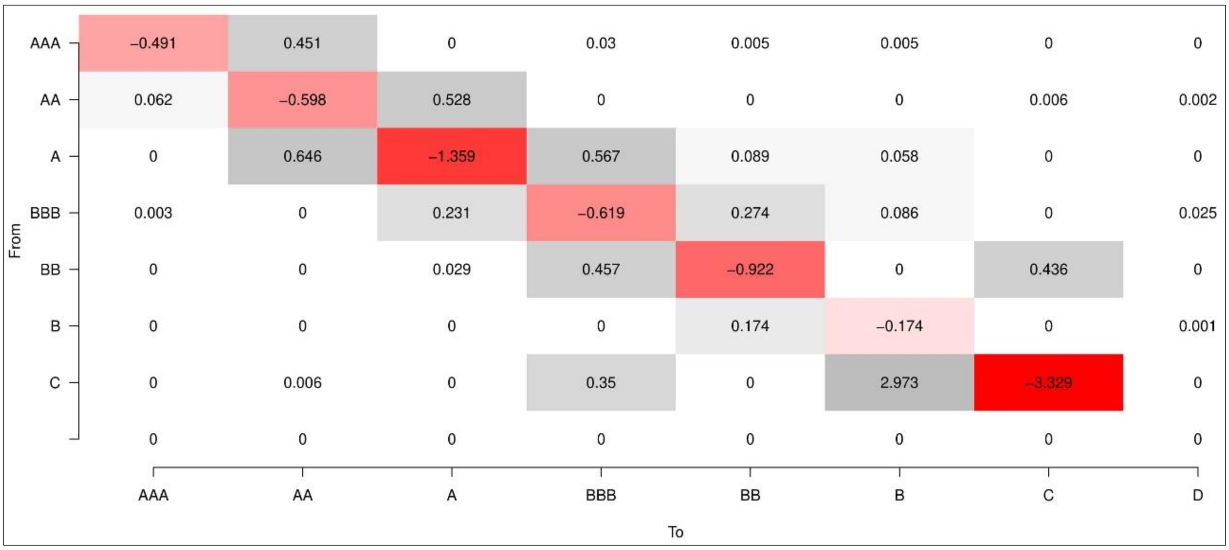

Fig 4: Generator Matrix obtained using Weighted Adjustment Method

Evaluating equation 8 using R programming yield the data below;

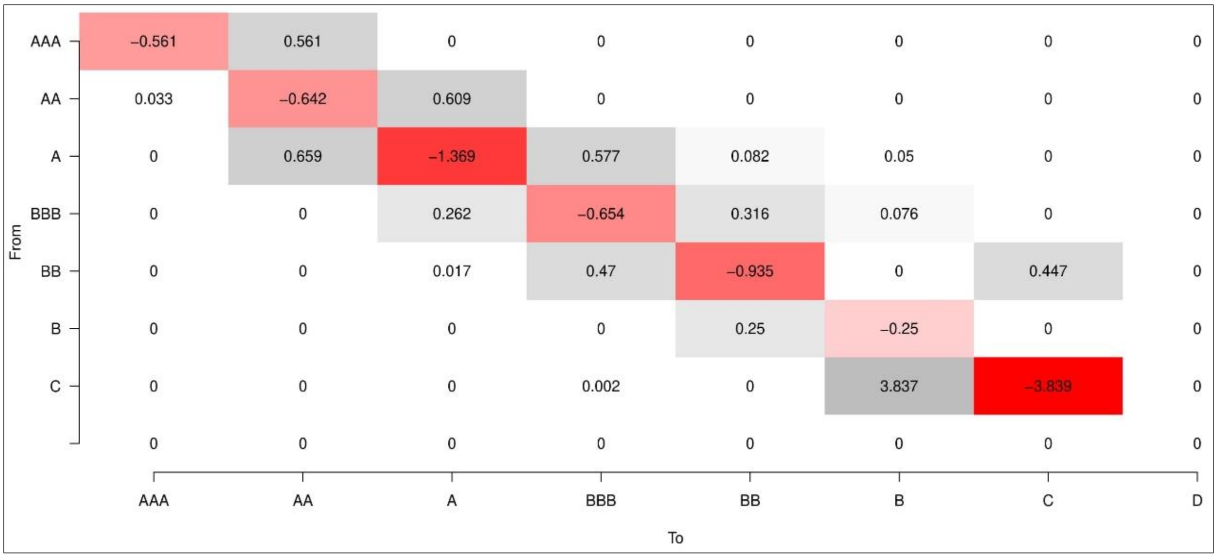

Fig 5: Generator Matrix obtained using Quasi Optimization Method

Using data from figure 1 above and running equation 19 using R yields the generator matrix below;

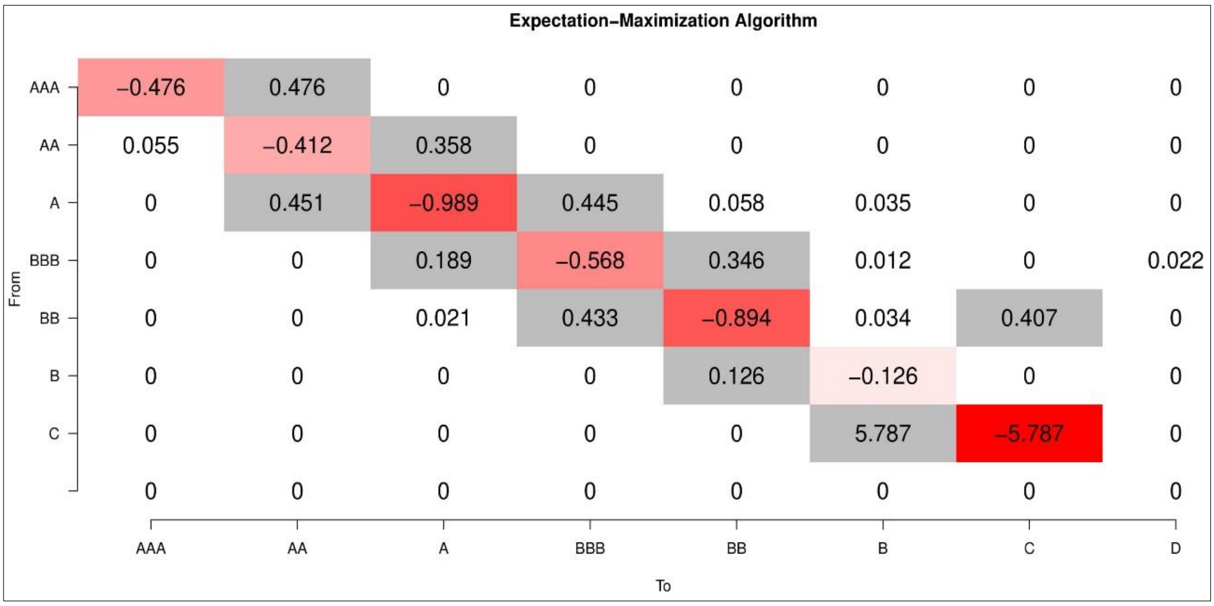

Fig 6: Generator Matrix obtained using Expectation-Maximization Algorithm Method

Using using data from figure 1 above and running equation 23 using R yields the generator matrix using Gibbs Sampler method is given below;

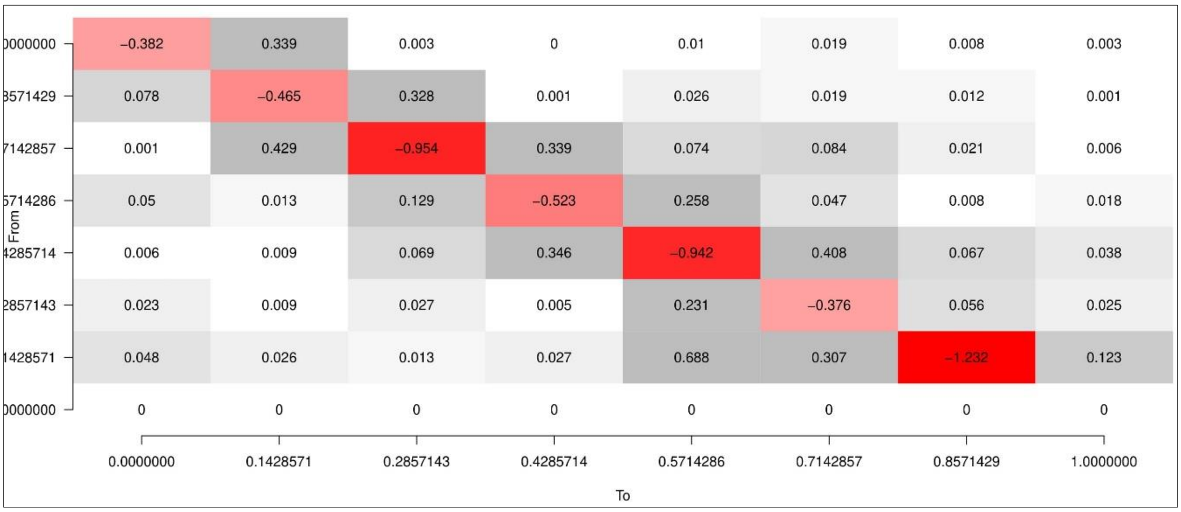

Fig 7: Generator Matrix obtained using Gibbs Sampler Method

4. Discussion

L Norm comparison between Diagonal Adjustment, Weighted Adjustment, Quasi Optimization and Gibbs Sampler Monte Carlo Methods

In order for one to know which Generator approximation is more suitable or accurate for an 8×8 Credit Generator Matrix, the L norm is computed for each Generator matrix obtained by the four methods above.

Using R software, it is computed as follows;

$$\begin{aligned} norm[P - \exp(Q_{WA})] &= 24.40658 \\ norm[P - \exp(Q_{DA})] &= 82.1496 \\ norm[P - \exp(Q_{QO})] &= 51.21381 \\ norm[P - exp(Q_{GS})] &= 8.207688 \end{aligned} \quad (26)$$

Hence from norm results 26 above, the norm of Gibbs Sampler method given by 8.207688 is the least hence the most suitable for computing the generator for embeddable 8×8 Credit transition Matrix and this paper proposes the Gibbs sampler method as the most accurate for generator matrix approximation for a Credit Transition Matrix.

5. References

1. Carette P. Characterizations of embeddable 3 3 stochastic matrices with a negative eigenvalue. New York J Math. 1995; 1(120):129.
2. Hanks E. Ctmcmove: Modeling animal movement with continuous-time discrete-spacemarkov chains. R package version, 2016, 1(3).
3. Israel RB, Rosenthal JS, Wei JZ. Finding generators for markov chains via empirical transition matrices, with applications to credit ratings. Mathematical finance. 2001; 11(2):245–265.
4. Kingman JFC. The imbedding problem for finite markov chains. Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete. 1962; 1(1):14–24.
5. Pfeuffer M, Reis GD *et al*. Capturing model risk and rating momentum in the estimation of probabilities of default and credit rating migrations. arXiv preprint ar. 2018; Xiv:1809.09889.