
Analysis of credit-rating migrations with genetic algorithms

Yuri Kaniovski*

Faculty of Economics and Management,
Free University of Bozen-Bolzano,
Bolzano, BZ, Italy
Fax: +39-0471-013-009
Email: YKaniovskiyi@unibz.it
*Corresponding author

Yuriy Kaniovskiyi

Research Group Scientific Computing,
Faculty of Computer Science,
University of Vienna,
Vienna, Austria
Fax: +43-1-4277-878447
Email: yuriy.kaniovskiyi@univie.ac.at

Georg Pflug

Department of Statistics and Operations Research,
University of Vienna,
Vienna, Austria
Fax: +43-1-4277-38639
Email: georg.pflug@univie.ac.at

Abstract: Modelling dependent credit-rating migrations of assets classified into M credit classes and S industries, $M \times S + 2^{M \times S}$ parameters have to be estimated. For a realistic choice of M and S , this number is huge and it greatly exceeds the number of available observations. To avoid brute-force calculations, we suggest sequential and parallel genetic algorithms. Considering a practically important combination of $M = 7$ and $S = 6$, the approach is tested on Standard and Poor's data.

Keywords: heuristics; encoding; nonlinear programming; mutation; parallel; sequential; maximum likelihood; selection; threshold; random search.

Reference to this paper should be made as follows: Kaniovski, Y., Kaniovskiyi, Y. and Pflug, G. (2020) 'Analysis of credit-rating migrations with genetic algorithms', *Int. J. Bio-Inspired Computation*, Vol. 16, No. 4, pp.264–274.

Biographical notes: Yuri Kaniovski holds a PhD in Operations Research from the V.M. Glushkov Institute of Cybernetics in Kyiv. He has published on stochastic optimisation, credit risk analysis, diffusion of innovations and evolutionary economics.

Yuriy Kaniovskiyi received his Master's in Computer Science at the University of Vienna. His research interests include parallel algorithms, cloud computing, adaptive execution strategies for data- and compute-intensive processing systems in bioinformatics.

Georg Pflug holds a PhD in Mathematics from the University of Vienna. He is an editor of eight books and author of five books and more than 100 publications in refereed journals on optimisation of stochastic systems and computationally intensive methods in simulation and optimisation.

This paper is a revised and expanded version of a paper entitled 'Dependent credit-rating migrations: a heuristics for estimating unknown parameters' presented at The 21st IFORS (International Federation of Operational Research Societies) Triennial Conference, Quebec City, Canada, 17–21 June 2017; and a paper 'Model identification and parameter estimation for huge latent distributions: macroeconomic scenarios and credit-rating migrations' presented at The 30th European Conference on Operational Research (EURO 2019), Dublin, Ireland, 23–26 June 2019.

1 Introduction

A credit-rating agency classifies economic entities according to their creditworthiness. Among other features, such a rating determines the interest rate paid by the entity when borrowing money. For example, the Standard and Poor (S&P) consider nine non-default credit classes. In descending order of creditworthiness they are referred to as: *AAA*, *AA*, *A*, *BBB*, *BB*, *B*, *CCC*, *CC* and *C*. Defaulted entities are labelled by *D* (a more subtle classification involves more classes: *AA+*, *R*, *NR*, etc. Also it may indicate an ‘outlook’ for a credit class. We do not discuss such refinements.). Merging *CCC*, *CC* and *C* into a single credit class *C*, a reduction with seven levels is often studied in research papers and textbooks.

We consider $M > 1$ non-default credit classes. They are numbered in descending order of creditworthiness by 1, 2, ..., M . Defaulted entities are indexed by $M + 1$. In particular, there are investment-grade and non-investment-grade debtors or assets if $M = 2$. The former comprises the credit classes *AAA*, *AA*, *A* and *BBB* of the standard S&P’s classification, while the latter is formed by the debtors rated at *BB*, *B*, *CCC*, *CC* and *C*.

The agency reconsiders periodically its ratings. Then some of the entities may receive a higher mark of creditworthiness while the others may be downgraded. That is, the classification evolves in time. A balance of the migrations over a period of time is given by transition counts.

Within the CreditMetrics approach, a transition matrix P is estimated from the counts (see Gupton et al., 1997). The entry P_{ij} of P is interpreted as the probability that a debtor belonging to the credit class i at time t will move to the credit class j at time $t + 1$. A defaulted debtor never returns to business. Therefore, $P_{M+1,j} = 0$ for $j = 1, 2, \dots, M$ and $P_{M+1,M+1} = 1$. In other words, $M + 1$ is an absorbing state of this time-homogeneous, discrete-time, finite Markov chain. Typically, only first M rows of P are quoted. We refer to this $M \times (M + 1)$ sub-matrix as a *migration matrix*.

Since the economic environment is common for all debtors, their credit-rating migrations are affected by the same market forces. Consequently, the migrations are dependent. Considering only two macroeconomic scenarios, contraction and expansion, the models employing regime-switching are used. See Bangia et al. (2002) for a description of them. In a similar vein, Fei et al. (2012) envisage three macroeconomic scenarios: expansion, ‘mild’ recession and ‘severe’ recession. A still more detailed analysis distinguishes among the market conditions affecting different credit classes. Then, assuming for each credit class two possibilities, favourable and adverse economic conditions, there are 2^M scenarios to analyse. Assigning 1 to a favourable outcome and 0 to an adverse outcome, these economic scenarios can be represented as binary strings v with M positions. Denote by $\{0, 1\}^M$ the set of all such vectors. Let the i^{th} coordinate v_i of v characterise the economic conditions affecting the credit class i . We call v_i the i^{th} *tendency variable*.

To introduce dependence among credit-rating migrations, a coupling scheme is used. Then every migration is modelled as a mixture of an idiosyncratic component ζ and a common component η ,

$$\delta\zeta + (1 - \delta)\eta \tag{1}$$

The Bernoulli random variable δ indicates which of the components actually drives the migration. To simplify the notation, equation (1) contains neither time nor an identifier of the asset. Time is discrete $t = 1, 2, \dots$. For every asset, the random variables δ ; ζ and η defining the credit-rating migration at a time instant are independent. These triples of random variables are independent in time as well. Finally, the random variables δ and ζ are independent across debtors, while three models of dependence across debtors are known for the random variables η [see Kaniowski and Pflug (2007), Wozabal and Hochreiter (2012) and Boreiko et al. (2017) for details].

If debtors are classified into S industry sectors, the distribution of δ depends on a combination of a credit class i and an industry s . Therefore, the probabilities of success $q_{i,s}$ form a matrix q with M rows and S columns.

For debtors belonging to the credit class i , ζ takes the value j with probability $P_{i,j}$, while all three known dependence models for η employ the following conditional distribution $P_{i,j}(v_i)$ of η :

$$P_{i,j}(1) = \begin{cases} \frac{P_{i,j}}{P_i} & \text{if } j \leq i \\ 0 & \text{if } j > i \end{cases} \text{ and } P_{i,j}(0) = \begin{cases} \frac{P_{i,j}}{1 - P_i} & \text{if } j > i \\ 0 & \text{if } j \leq i \end{cases}$$

Here, P_i stands for $P_{i,1} + P_{i,2} + \dots + P_{i,i}$, $i = 1, 2, \dots, M$. With such assumptions, the credit rating at $t + 1$ of a debtor belonging to the industry s whose rating at t is i will be j with probability $q_{i,s}P_{i,j} + (1 - q_{i,s})P_{i,j}(v_i)$ conditional on v_i . This parametrisation reflects the empirical fact that migrations towards more (less) secure credit ratings are more frequent under favourable (adverse) economic conditions.

To specify completely the dynamics given by equation (1), we need a probability distribution d over binary strings. Then $d(v) \geq 0$ for every $v \in \{0, 1\}^M$ and

$$\sum_{v \in \{0, 1\}^M} d(v) = 1 \tag{2}$$

The conditional migration probabilities $q_{i,s}P_{i,j} + (1 - q_{i,s})P_{i,j}(v_i)$ can deviate from the corresponding $P_{i,j}$. In the long run, these deviations have to compensate each other. In order to keep for every i the unconditional distribution in equation (1) equal to i^{th} row of P , the following M linear constraints have to be satisfied:

$$\sum_{v \in \{0, 1\}^M} v_i d(v) = P_i \tag{3}$$

Conceptually, relations (3) mean that the i^{th} tendency variable equals 1 with probability P_i .

By the maximum likelihood principle, maximising the likelihood function $l(d, q)$ implied by one of the three dependence models for random variables η , the true distribution d^* and Bernoulli probabilities q^* can be estimated. The corresponding nonlinear programming problem involves $\ln l(d, q)$, the linear constraints (2) and (3) and, since all unknowns are probabilities, the box-constraints, $0 \leq d(v) \leq 1$ and $q_{i,s} \in [0, 1]$. We call this optimisation problem a *basic setting*. Its solution is used below to ‘assemble’ a heuristic solution of a more realistic and involved computationally *complete setting*.

The basic setting is not a computationally hard problem. Wozabal and Hochreiter (2012) found a solution considering $M = 4$ and $S = 6$. For this choice of M and S , the distribution d is nested in $2^4 = 16$ binary vectors. They used a particle swarm optimisation method. A desktop computer requires several minutes to solve such a problem. The interior point (IP) method and the sequential quadratic programming method were used in Boreiko et al. (2017) who considered $M = 7$ and $S = 6$. In this case, the distribution of d is nested in $2^7 = 128$ binary strings. A desktop computer requires a couple of minutes to find a solution. Boreiko et al. (2017) reports estimates for all three known dependence models. If d^* and q^* are available, simulating credit-rating migrations according to the formula (1), the loss generated by a portfolio can be estimated with the Monte-Carlo method. Wozabal and Hochreiter (2012) compare the loss estimated according to their coupling scheme with the estimates obtained when alternative models of dependent credit-rating migrations are used.

The models reviewed so far assume either economic scenarios common to the whole economy or the economic scenarios common to all debtors belonging to a credit class. A sectoral classification of economic scenarios is the next fine-tuning in modelling of dependent credit-rating migrations. Remaining with two possibilities, a favourable and an adverse outcome, an economic scenario has to distinguish among the economic conditions affecting assets characterised by every combination of a credit class and an industry. Considering M credit classes and S industries, a binary string V corresponding to such a scenario must have $M \times S$ positions. Assigning 1 to a favourable outcome and 0 to an adverse outcome, we consider the set of binary strings $\{0, 1\}^{M \times S}$. It contains $2^{M \times S}$ elements. We need a rule for allocation coordinates of a string V to industries and credit classes. To this end, let the coordinate $V_{M(s-1)+i}$ of a binary $(M \times S)$ -vector V characterise the economic conditions affecting the credit class i of the industry s . It is called the *tendency variable* for this combination of a credit class and an industry. That is, the industries occupy blocks of M coordinates each. The blocks are numbered in ascending order of s . Within a block, the credit classes are listed in ascending order of i .

Computational complexity of this setting is enormous. In fact, if $M = 7$ and $S = 6$, the number of binary strings $2^{7 \times 6} = 2^{42}$ is huge. Some of these macroeconomic scenarios are plausible while the others do not make conceptual sense.

In real life, the plausible scenarios are not equally frequent. To model this phenomenon, consider a distribution D such that $D(V)$ is the probability of observing the economic scenario encoded by the binary string V . We have to identify the distribution D^* that fits the observed credit-rating migrations the best. Analysing all binary strings V simultaneously is a task for a supercomputer. Using a desktop computer, we can test just some of them. Consequently, a strategy for choosing the subsets of strings is necessary as well as a strategy for testing them.

This paper suggests heuristics for estimating parameters of models involving industry and credit class specific economic scenarios. As a first step, we consider a reduction of the search space. Second, a decomposition of the reduced search space is introduced. Third, a couple of random search methods are suggested to avoid analysing all of the outcomes in the reduced search space. Finally, a risk threshold typical for the recent regulations in the financial sector is used to streamline the search.

Genetic algorithms (GAs), especially those implying a decomposition of the space of decision variables, is an appropriate instrument to tackle this problem. See Ma et al. (2018, 2019, 2020) for the state of the art in the evolutionary algorithms of this kind.

Two important aspects of the estimation problem have to be mentioned. On the one hand, $2^{M \times S}$ typically greatly exceeds the number of available transition counts $I_{m_1, m_2}^{t,s}$.

We denote by $I_{m_1, m_2}^{t,s}$ the number of debtors in the industry s that migrated from the credit class m_1 to the credit class m_2 in the period t . Therefore, the conventional wisdom suggests that no statistical inference should be possible under such circumstances. On the other hand, estimating parameters of historical migrations is a practical problem. As in any statistical application, there is a threshold for distinguishing between the important and the insignificant outcomes. The Basel accords (BAs) – the banking supervision accords (recommendations on banking regulations) – Basel I, Basel II and Basel III-issued by the Basel Committee on Banking Supervision – set an upper bound of 0.1% for the admissible in the financial sector risk. Then at most 1,000 binary strings have to be considered even in the extreme case when all of them were equally probable. In sum, evoking the conceptual interpretation of binary strings as economic scenarios it appears that the problem is not as hopeless as it looks at first glance.

Analysis of the credit risk associated with a portfolio and evaluation of the systemic risk implied by the banking system of a country are the most important applications for the estimates suggested here. See Wang et al. (2017) who analyse the recent developments in modelling of dependent credit-rating migrations and Glasserman and Young (2016) who present the state of the art in modelling of systemic risk.

As a benchmark, we consider the common for all industries migration matrix P . However, modelling industry-specific transition probabilities should imply a more precise estimate of the credit risk associated with a

portfolio composed by interdependent assets as compared with the models employing the same migration matrix for all debtors. The transitions within an industry allow to estimate the corresponding industry-specific migration matrix. Such matrices typically differ across industries. Denote by $P^{(s)}$ the migration $M \times (M + 1)$ matrix estimated for the industry sector s . Let $P_{i,j}^{(s)}$ be entries of $P^{(s)}$.

The migrations are defined by equation (1) with δ and ζ independent in time and across debtors. For a debtor belonging to the credit class i and to the industry s :

- distribution of ζ coincides with the i^{th} row of $P^{(s)}$
- probability of success of δ equals $Q_{i,s} \in [0, 1]$ an entry of an $M \times S$ matrix Q
- the conditional distribution of η for a given $V_{M(s-1)+i}$ is:

$$P_{i,j}^{(s)}(1) = \begin{cases} \frac{P_{i,j}^{(s)}}{P_i^{(s)}} & \text{if } j \leq i \text{ and} \\ 0 & \text{if } j > i \end{cases}$$

$$P_{i,j}^{(s)}(0) = \begin{cases} \frac{P_{i,j}^{(s)}}{1 - P_i^{(s)}} & \text{if } j > i \\ 0 & \text{if } j \leq i \end{cases}$$

For every time instant and every debtor the corresponding random variables ζ , δ and η are independent. Conditional on $V_{M(s-1)+i}$, the credit rating at $t + 1$ of a debtor belonging to the industry s whose rating at t is i will be j with probability $Q_{i,s}P_{i,j}^{(s)} + (1 - Q_{i,s})P_{i,j}^{(s)}(V_{M(s-1)+i})$.

At a fixed time instant, Boreiko et al. (2017) specify two models of dependence among random variables η :

- 1 η are conditionally independent across debtors
- 2 η remains the same for all debtors characterised by a combination of a credit class i and an industry s , while for different couples, the corresponding random variables η are conditionally independent.

Let $L(D, Q)$ be the likelihood function corresponding to one of these dependence models. To estimate the true values D^* and Q^* , $\ln L(D, Q)$ has to be maximised subject to the $M \times S + 1$ linear constraints,

$$\sum_{V \in \{0,1\}^{M \times S}} D(V) = 1 \tag{4}$$

$$\sum_{V \in \{0,1\}^{M \times S}} V_{M(s-1)+i} D(V) = P_i^{(s)} \tag{5}$$

and the box constraints, $D(V) \in [0, 1]$, $Q_{i,s} \in [0, 1]$. Here, $P_i^{(s)}$ stands for $P_{i,1}^{(s)} + P_{i,2}^{(s)} + \dots + P_{i,i}^{(s)}$. Constraints (5) guarantee that the unconditional distribution in (1) equals to the i^{th} row of $P^{(s)}$ for every combination of i and s . Conceptually, relations (5) mean that a macroeconomic scenario favourable for the industry s and to the credit class i occurs with probability $P_i^{(s)}$. Such an optimisation problem is termed as a *complete setting*.

Considering $M = 2$ and $S = 6$, Boreiko et al. (2017) solved with the IP method the complete setting for both dependence models. There are $2^{2 \times 6} = 4,096$ economic scenarios. With 103,723 transition counts available, this complete setting can be treated within the traditional framework. A desktop computer requires a couple of hours to solve such a problem. Along with more precise and specific credit risk estimates, certain aspects of systemic risk can be analysed using the true distribution D^* . For example, Boreiko et al. (2017), referring to the period of time between 1991 and 2015, concluded that “the investment-grade financial institutions evolve independently of the rest of the economy represented by the data. This might be an evidence of implicit too-big-to-fail bailout guarantee policies of the regulatory authorities.”

Because complexity of the sample space is combinatorial for a realistic choice of M and S , estimation of D^* is a big-parameter problem. Therefore, a clear understanding of what is a solution in such a case is necessary. The next section suggests an interpretation in the context of risk analysis. It presents the framework employed by all the GAs considered here. Numerical tests are described in Section 3. The biological phenomena mimicked by the sequential and the parallel algorithms are discussed. Versions with equally probable mutations as well as versions where mutations are sampled path-dependently are suggested. The numerical results presented in Section 4 allow to indicate the most efficient GAs as well as the fastest of them. Section 5 summarises the main results.

2 The framework

We call $\mathbf{V} \subseteq \{0, 1\}^{M \times S}$ a *suitable set*, if there exists a probability distribution $D_{\mathbf{V}}$ such that:

- $D_{\mathbf{V}}$ satisfies constraints (5)
- the support of $D_{\mathbf{V}}$ belongs to \mathbf{V} .

For a suitable set \mathbf{V} , the set of non-negative measures D defined by the $M \times S + 1$ linear constraints,

$$\sum_{V \in \mathbf{V}} D(V) = 1 \tag{6}$$

$$\sum_{V \in \mathbf{V}} V_{M(s-1)+i} D(V) = P_i^{(s)} \tag{7}$$

is not empty. In fact, $D_{\mathbf{V}}$ satisfies these constraints. If the cardinality of \mathbf{V} is low enough, the corresponding optimal distribution $D_{\mathbf{V}}^*$ can be found by a conventional nonlinear programming method. The estimates obtain by maximising $\ln L(D, Q)$ subject to the linear constraints (6) and (7). The box constraints on $Q_{i,s}$ and $D(V)$ must hold true as well.

For the practical numerical procedure, look for the ϵ -support \mathbf{V}_{ϵ} of $D_{\mathbf{V}}^*$:

$$\mathbf{V}_{\epsilon} = \{V \in \mathbf{V} : D_{\mathbf{V}}^*(V) > \epsilon\}$$

Since $\sum_{V \in V_\epsilon} D_V^*(V) \uparrow 1$ as ϵ decreases and D_V^* satisfies the constraints (6) and (7), V_ϵ is a suitable set for all sufficiently small ϵ . In fact, the linear constraints (6) and (7) with V_ϵ and D_V^* instead of V and D , can hold true with any desired precision.

Here is the structure of our heuristic search:

- 1 At the beginning, a suitable set V with a low cardinality has to be found and a threshold value ϵ has to be chosen.
- 2 Solve the optimisation problem for specifying V_ϵ . Decrease ϵ if V_ϵ is not suitable.
- 3 Better solutions (in terms of the likelihood value) can be found by extending V_ϵ to a larger set. Since any set containing a suitable set is also suitable, the extension preserves suitability. Let $V' = V_\epsilon \cup \bar{V}$, where \bar{V} is a subset of $\{0, 1\}^{M \times S} / V$. This is a suitable set, an extension of V_ϵ . \bar{V} should not be too large so that the optimisation problem with V replaced by V' could be solved.
- 4 We may repeat steps 1–3 with V' replacing V and continue them. The sequence of maximum likelihood values corresponding to the extensions does not decrease.

Finding an initial suitable set V is a quite difficult task. Note that existence of a suitable set is equivalent to existence of an optimal distribution D^* . In fact, a set containing a suitable set is suitable. Therefore, $\{0, 1\}^{M \times S}$ is a suitable set, if a suitable V exists. In this case we set $D^* = D_{\{0,1\}^{M \times S}}^*$.

On the other hand, if an optimal distribution D^* exists, the whole $\{0, 1\}^{M \times S}$ with $D_{\{0,1\}^{M \times S}} = D^*$ is a suitable set.

The choice of ϵ is affected by two contradictory factors. On the one hand, a larger threshold can imply a smaller support V_ϵ . On the other hand, a larger ϵ can cause non-suitability of the support. Note that V_ϵ contains at most ϵ^{-1} elements.

Even if the search space for a suitable V consists of binary strings, a mutation or a recombination based on the conventional GA encoding can be, given the notation adopted here, an implausible conceptual transformation. See Holland (1992) for a general introduction into GAs. Some recent applications in engineering can be found in Cai et al. (2020a, 2020b) and Ma et al. (2014a, 2014b, 2014c, 2016a, 2016b).

Ceteris paribus, a string (... , 1, 1, ...) can appear intuitively more likely than its mutation (... , 0, 1, ...). For example, let the former economic scenario be favourable for debtors belonging to two subsequent credit classes, say i and $i + 1$, in an industry. Then the latter scenario is favourable for the more risky credit class $i + 1$ while it is adverse for the less risky credit class i . This is a situation having little empirical justification, at least without a deeper insight into the economic scenarios represented by the strings. Consequently, the conventional GA encoding of the

problem cannot be used. In fact, an allele in this case has to consist of several symbols 0 and 1. A natural candidate for the smallest unit of evolutionary selection is a binary M -vector characterising the economic conditions affecting credit classes of an industry. This approach allows to avoid brute-force evaluation of all possible $2^{M \times S}$ macroeconomic scenarios. The particular extent of the reduction of the search space and thus the complexity of the estimation problem is characterised below in several steps.

Let us identify a set of alleles for running a selection process. Choose a dependence models for the basic setting and a dependence model for the complete setting. Having three possibilities for the first choice and two possibilities for the second choice, six couples can be considered. Write down the corresponding likelihood functions l and L . Fix a threshold $\epsilon > 0$. Solve the basic setting.

Let v^k , $k = 1, 2, \dots, K$, be the support of d^* defined according to the threshold. That is, $d^*(v^k) > \epsilon$ for $k = 1, 2, \dots, K$. We refer to these binary vectors as blocks. Combining them, we create economic scenarios for an economy, where economic conditions affecting different industries can vary. As the simplest economic scenario of this kind, let V^k consist of S identical blocks v^k , $V^k = (v^k, v^k, \dots, v^k)$. It is referred to as the k^{th} block-structure. Let us denote by V the set of all such block-structures. It contains K elements. Conceptually, v^k is an economic scenario for an economy, where economic conditions are not classified according to industry sectors, while V^k is the same scenario for the same economy, where economic conditions are additionally classified into S groups to account for the industry-specific phenomena. V^k represents the particular case when identical economic conditions affect all industries. The conditions are represented by v^k .

Consider first the benchmark case of the common for all industries migration matrix P . Let us show that the defined above set V of block-structures is a suitable set if ϵ is sufficiently small. To this end, define a probability distribution D_V on V by setting $D_V(V^k) = d^*(v^k)$ for $k = 1, 2, \dots, K$. Observe that equation (7) with the common for all industries transition matrix P are just S times repeated equation (3). In formal terms, the estimation problem for (d^*, q^*) is nested, if all $P^{(s)} = P$ for all s , in the estimation problem for (D_V^*, Q^*) . Since $\sum_{i=1}^K d^*(v^i) \uparrow 1$ as ϵ decreases and d^* satisfies constraints (2) and (3), we conclude that D_V will satisfy constraints (6) and (7) with any desired precision. It is enough to choose a sufficiently small ϵ . In sum, V is a suitable set.

Solve the optimisation problem for identifying V_ϵ . Recall that the likelihood function is L in this case. Denote by \tilde{K} the number of elements in V_ϵ . Then $\tilde{K} \leq K$, because some of the binary vectors from V may receive a probability smaller than ϵ .

As it has been explained, a chromosome has S loci. They correspond to S industry sectors. If $V^{ij} = (v^{ij}, v^{ij}, \dots, v^{ij}) \in V_\epsilon$, then v^{ij} is an allele. In sum, the GA encoding involves S loci and \tilde{K} alleles in our case.

There are \tilde{K}^S possibilities for allocating \tilde{K} building blocks among S positions. If $\tilde{K} < 2^M$, this number is smaller than $2^{M \times S}$. Consequently, the suggested GA encoding in fact implies $\left(\frac{2^M}{\tilde{K}}\right)^S$ times reduction of the search space. On the other hand, there is a price of the reduction: if even all possible \tilde{K}^S chromosomes were considered, the corresponding solution could still not be optimal. Let us order the block-structures so that $D_V^*(V^{i_1}) \geq D_V^*(V^{i_2}) \geq \dots \geq D_V^*(V^{i_{\tilde{K}}})$.

Living communities evolve from simpler to more complicated forms. For the case in hand, this observation translates into a search process employing extensions that progressively deviate from the block-structures, as the simplest chromosomes. Sets \mathbf{V}^n ; $n \geq 1$; of mutants with (exactly) n mutations of the block-structures, on the one hand, increasingly deviate from \mathbf{V}_ϵ and, on the other hand, complexity of the set $\cup_{k=0}^n \mathbf{V}^k$ of mutants with at most n mutations increases in n . If indicating the exact threshold value is not necessary, \mathbf{V}_ϵ will be denoted also as \mathbf{V}^0 for notation consistency. In fact, block-structures are mutants with zero mutations.

Even if the notion of a mutation is fundamental for GAs, let us describe formally, keeping in mind the conceptual problem in hand, the mutations of a block-structure. This is necessary for specifying unambiguously the extension schemes introduced below. Substituting one block v^{ik} in V^{ik} by a block v^{il} $l \neq k$, we get a *mutant* with a *single mutation* of the block-structure V^{ik} . Conceptually, this is an economic scenario, where all industries, except for one, are affected by the economic conditions summarised in v^{ik} , while the remaining industry is affected by the economic conditions represented by v^{il} . Trying all possible V^{ik} , $k = 1, \dots, \tilde{K}$, industries $s = 1, \dots, S$ or, equivalently, positions of the mutated block, and its type v^{il} , $l = 1, \dots, \tilde{K}$, $l \neq k$, the set \mathbf{V}^1 of all mutants with a single mutation will be obtained. Similarly \mathbf{V}^n , can be defined for $n > 1$. Note that a block v^{il} can appear more than one time in a mutant stemming from the block-structure V^{ik} .

The number of binary vectors in \mathbf{V}^n equals,

$$\frac{S!}{n!(S-n)!} \tilde{K}(\tilde{K}-1)^n$$

if $n < S/2$. For an even S , $\mathbf{V}^{S/2}$ contains,

$$\frac{S!}{\left(\frac{S}{2}\right)^2} \tilde{K}(\tilde{K}-1)^{\frac{S}{2}-1} (\tilde{K}-2)$$

chromosomes. Note that $\mathbf{V}^n \cap \mathbf{V}^{n'} = \emptyset$, if $n, n' < S/2$; $n \neq n'$. For $n \geq S/2$, the mutants originating from different block-structures should not be listed several times. Already

for $n = S/2$, if S is an even number, a mutant containing $S/2$ blocks v^{i_1} and $S/2$ blocks v^{i_2} can be regarded as a mutation of the block-structure V^{i_1} as well as a mutation of the block-structure V^{i_2} . Such mutants have to be listed in $\mathbf{V}^{S/2}$ only one time. If there are too many binary vectors in a set \mathbf{V}^n , it can be split into subsets and each of them can be used for extensions.

Instead of dealing with all binary vectors from \mathbf{V}^n , a random sample can be drawn. Using such a random search strategy the space of decision variables and, consequently, the complexity of the estimation problem can be further reduced. The extent of this reduction depends upon the sample size.

Sampling a binary vector from \mathbf{V}^n , for the first mutation there are S equally probable positions, for the second – the remaining $S-1$, etc. For the probability of choosing a particular block as a mutation there are two possibilities. First, the probability of choosing the block v^{i_1} can be proportional to $D_V^*(V^{i_1})$. In particular, generating mutations in the block-structure V^{i_1} , v^{i_1} is sampled with probability

$$D_V^*(V^{i_1}) / \sum_{j=1, j \neq k}^{\tilde{K}} D_V^*(V^{i_j}).$$

We refer to this distribution as *natural*. This way of increasing the variety of genotypes under consideration is an example of a *path-dependent search*. Arguing about species in a living population, the word ‘frequency’ is more appropriate than ‘probability’. Then, instead of more probable binary strings, we can argue about more frequent alleles or chromosomes. Intuitively, alleles, corresponding to more probable block-structures, should be used more frequently as mutations. Choosing mutations with equal probabilities is an alternative way for sampling mutants. This distribution is termed as *uniform*. Since each of the possible alleles is always sampled with probability $\frac{1}{\tilde{K}-1}$, this technique can imply a richer variety of mutants.

Crossing over v^k and v^l can produce a binary M -vector that does not belong to the support of d^* . In other words, this economic scenario has been sorted out by the basic setting as an unlikely outcome. In fact, the fitness function l assigned to this scenario a probability that falls below the threshold ϵ . In view of this observation, we do not use recombination of block-structures in our GAs.

If the transition matrices are industry-specific, the relations (7) are not a consequence of the relations (3): for a combination of an i and an s , the mismatch depends upon $P_i^{(s)} - P_i$. Consequently, the defined above set \mathbf{V} may be regarded only as an *approximation* to a suitable set. However, the alleles v^k , $k = 1, 2, \dots, K$, can be used in the evolutionary search discussed above. Since every time a larger set of binary vectors is involved, passing from $\cup_{k=0}^n \mathbf{V}^k$ to $\cup_{k=0}^{n+1} \mathbf{V}^k$, the corresponding maximal likelihood value cannot decrease and the corresponding mismatch of the constraints cannot increase. Therefore, a suitable set can be eventually found.

3 Inputs and numerical experiments

Solutions of all nonlinear optimisation problems reported below were found with the IP method. The likelihood functions l and L tested correspond to the model with conditionally independent across debtors random variables η :

$$l(d, q) = \prod_{t=1}^T \sum_{v \in \{0,1\}^M} d(v) \prod_{s=1}^S \prod_{m_1=1}^M \prod_{m_2=1}^{M+1} f_{m_1, m_2}(s, v, q)^{I_{m_1, m_2}^{t, s}}$$

$$L(D, Q) = \prod_{t=1}^T \sum_{V \in \{0,1\}^{M \times S}} D(V) \prod_{s=1}^S \prod_{m_1=1}^M \prod_{m_2=1}^{M+1} F_{m_1, m_2}(s, V, Q)^{I_{m_1, m_2}^{t, s}}$$

where

$$f_{m_1, m_2}(s, v, q) = \begin{cases} \frac{1 - q_{m_1, s}(1 - P_{m_1})}{P_{m_1}} & \text{if } m_1 \geq m_2, v_{m_1} = 1 \\ \frac{1 - q_{m_1, s}P_{m_1}}{1 - P_{m_1}} & \text{if } m_1 < m_2, v_{m_1} = 0 \\ q_{m_1, s} & \text{otherwise} \end{cases}$$

$$F_{m_1, m_2}(s, V, Q) = \begin{cases} \frac{1 - Q_{m_1, s}(1 - P_{m_1}^{(s)})}{P_{m_1}^{(s)}} & \text{if } m_1 \geq m_2, V_{M(s-1)+m_1} = 1 \\ \frac{1 - Q_{m_1, s}P_{m_1}^{(s)}}{1 - P_{m_1}^{(s)}} & \text{if } m_1 < m_2, V_{M(s-1)+m_1} = 0 \\ Q_{m_1, s} & \text{otherwise} \end{cases}$$

In the annual transition counts used for numerical experiments $t = 1$ corresponds to 1991 and $T = 25$ corresponds to 2015.

The analysis is restricted to the sample points whose probabilities exceed 10^{-3} , the threshold fixed by the BAs.

3.1 Alleles, sample and evaluation criterion

Considering $M = 7$ and $S = 6$, we use the $K = 13$ blocks v^k identified in Boreiko et al. (2017) for the threshold 10^{-3} :

$$\begin{aligned} v^1 &= (1, 1, 1, 1, 1, 1, 1), v^2 = (1, 1, 1, 1, 1, 0) \\ v^3 &= (0, 1, 1, 1, 1, 1), v^4 = (1, 1, 1, 1, 0, 1, 1) \\ v^5 &= (1, 1, 1, 0, 1, 1, 1), v^6 = (1, 0, 1, 1, 1, 1, 1) \\ v^7 &= (1, 1, 1, 1, 1, 0, 0), v^8 = (1, 0, 0, 1, 0, 1, 1) \\ v^9 &= (1, 0, 1, 1, 1, 0, 1), v^{10} = (1, 1, 0, 1, 0, 0, 0) \\ v^{11} &= (0, 0, 1, 1, 1, 0, 0), v^{12} = (1, 1, 0, 0, 1, 1, 1) \\ v^{13} &= (0, 0, 0, 1, 0, 0, 0) \end{aligned}$$

Probability of the remaining $2^7 - 13 = 115$ binary strings is $8.6 \cdot 10^{-50}\%$.

There are 103,723 annual transition counts for estimating $42 + 2^{42}$ parameters. That is, the number of unknowns exceeds $4.2 \cdot 10^7$ times the number of observations. See <https://doi.org/10.1371/journal.pone.0175911.s001> for the transition counts.

The fitness function of the selection process in hand is L . The quality of a solution whose likelihood equals L^* is characterised by the percentage of increase, $\frac{\ln L^* - \ln L^0}{\ln L^0} \cdot 100$ against the maximum likelihood value L^0 attained on the set of block-structures \mathbf{V}^0 . Recall that the block-structures are the simplest chromosomes in this ‘vivarium’.

3.2 Suitable sets and extensions

If the migration matrix P is the same for all industries, the set $\mathbf{V}_{10^{-3}}$ consists of 11 block-structures. They were assigned the following probabilities:

$$\begin{aligned} D_V^*(V^1) &= 0.5294, D_V^*(V^2) = 0.1986, D_V^*(V^3) = 0.0714, \\ D_V^*(V^5) &= 0.0531, D_V^*(V^{10}) = 0.0320, D_V^*(V^4) = 0.0266, \\ D_V^*(V^6) &= 0.0239, D_V^*(V^{11}) = 0.0210, D_V^*(V^9) = 0.0166, \\ D_V^*(V^8) &= 0.0142, D_V^*(V^{13}) = 0.0133 \end{aligned}$$

The remaining two block-structures received together the probability $7.9 \cdot 10^{-60}\%$. All 13 block-structures belong to $\mathbf{V}_{10^{-3}}$, if the industry-specific transition matrices $P^{(s)}$ are considered. The corresponding probabilities $D_V^*(V^{ik})$ are:

$$\begin{aligned} D_V^*(V^1) &= 0.4440, D_V^*(V^2) = 0.2280, D_V^*(V^3) = 0.0688, \\ D_V^*(V^{10}) &= 0.0421, D_V^*(V^5) = 0.0397, D_V^*(V^7) = 0.0363, \\ D_V^*(V^8) &= 0.0342, D_V^*(V^6) = 0.0329, D_V^*(V^4) = 0.0325, \\ D_V^*(V^{13}) &= 0.0255, D_V^*(V^{11}) = 0.0217, D_V^*(V^{12}) = 0.0212, \\ D_V^*(V^9) &= 0.0167 \end{aligned}$$

These block-structures do not form a suitable set, but just an approximation to it.

The total number of chromosomes generated by these alleles is $11^6 = 1,771,561$ and $13^6 = 4,826,809$, correspondingly. In relative terms, these economic scenarios cover $4.0 \cdot 10^{-50}\%$ and, correspondingly, $1.1 \cdot 10^{-40}\%$, of the total number 2^{42} of them. In other words, by restricting the search space to all possible mutations of the block-structures, computational complexity of the estimation problem reduces $\left(\frac{2^7}{11}\right)^6 \approx 2.5 \cdot 10^6$ and,

correspondingly, $\left(\frac{2^7}{13}\right)^6 \approx 9.1 \cdot 10^5$ times. As Table 1

demonstrates, dealing already with two mutations represents a hard computational problem. Consequently, an extension strategy is needed. To streamline the discussion, first we will quote the numerical characteristics of the GAs employing the common for all industry migration matrix P and then, in parentheses, the corresponding characteristics of the GAs involving industry-specific migration matrices.

Table 1 Number of elements \tilde{K} in V^n

\tilde{K} / n	1	2	3
11	660	16,500	198,000
13	936	28,080	411,840

We split the set of all mutants with two mutations into 5 (10) subsets. Each of them contains 3,300 (2,808) binary strings. In Figures 1 and 2, they correspond to the sets SS_k^2 . Since we are not able to analyse all mutants with $i = 3, 4$ and 5 mutations, we sample randomly 5 (10) subsets SS_k^i of V^i . Each of the block-structures is represented in these SS_k^i by 300 (250) mutants. Therefore, every such SS_k^i contains 3,300 (3,250) binary strings. None of the mutants can be listed more than one time. In particular, a newly generated mutant is added into the pool if it differs from all chromosomes already present there. To this end, none of the randomly generated mutants with four identical mutations, may be included in the list. In fact, all such mutants are already there because they belong to one of the subsets SS_k^2 . In the same way, randomly sampled mutants with five identical mutations will be excluded from the consideration because they are already included in V^1 .

The values presented in Table 2 are evaluated in the following way. Consider $n = 3$ for example. Dealing with the common for all industries transition matrix P , the total number of tested chromosomes with at most $n = 3$ mutations equals $11 + 660 + 16,500 + 5 \cdot 3,300 = 33,671$. The addends in the sum are: the number of elements in $V_{10^{-3}}$, the number of elements in V^1 , the number elements in V^2 , the total number of binary strings sampled from V^3 . The percentage $\frac{33,671}{11^6} \cdot 100$ characterises reduction of the search space due to, one the one hand, this particular extension scheme and, on the other hand, due to limiting by 3 the maximum number of mutations and using a random sample rather than all of them. We can argue in this case about $\frac{11^6}{33,671} \approx 53$ times reduction of complexity of the estimation problem. While the first three terms are fixed, the number of mutants sampled from V^3 can be changed. Then the percentage will be adjusted accordingly. Recall that 11^6 stands for the total number of chromosomes formed by eleven alleles that occupy six loci.

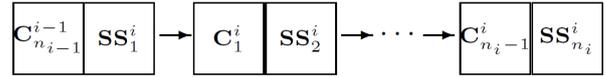
Table 2 Percentage of chromosomes from $\cup_{i=0}^n V^i$ tested

	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
P	0.06	1.46	2.87	4.28	5.69
$P^{(s)}$	0.02	0.60	1.27	1.95	2.62

Since the threshold value is fixed, we can denote everywhere $V_{10^{-3}}$ by V^0 . First, we estimate parameters allowing at most one mutation. That is, economic scenarios belonging to $V' = V^0 \cup V^1$ are considered. Note, that

this turns out to be already a suitable set in the case of industry-specific migration matrices. Then, mutants with at most two, three, four and five mutations are analysed. Dealing with at most two mutations, first, $V'_{10^{-3}}$ is extended using SS_1^2 . Remark that $V'_{10^{-3}}$ corresponds to C_1^1 in Figure 1. Then, the result C_2^2 is extended with the second subset SS_2^2 , etc. In sum, within a set V^n , the subsets of mutants SS_k^n are used for extension *sequentially* one after another.

Figure 1 Sequential algorithm, passage from $i - 1$ to $i, i \geq 2$



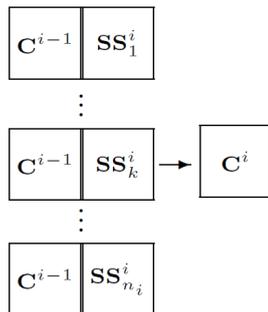
When the maximum likelihood method is applied to multinomial distributions, multiple maxima are a typical complication. Allman et al. (2009) present a comprehensive analysis of this phenomenon. Therefore, the IP method, which is used for the numerical experiments, can lock into a local maximum. As a consequence, extending a suitable set, the maximum value will not necessarily increase. Switching to a smaller threshold may not help in this case. Then, along with trying several initial approximations, as it is typically suggested, the order in which subsets of mutants are tested can be changed. Modifying the shape of the likelihood function, it may imply convergence to a larger likelihood value. In sum, if a local method is used for maximisation, the heuristic search can exhibit path dependence. That is, the result can depend upon the order in which suitable subsets are considered.

Such complications can be partly avoided if the subsets SS_k^n of a set V^n are analysed by a *parallel* algorithm. If 5 (10) parallel processors maximise simultaneously the corresponding likelihood functions, all calculations regarding V^n can be done 5 (10) times faster for every $n > 1$. In particular, dealing with at most $n = 2$ mutations, the k^{th} of 5 (10) parallel processors maximises the likelihood function based on the extension of $V'_{10^{-3}}$ the k^{th} subset SS_k^2 . Recall that $V' = V^0 \cup V^1$ and, consequently, $V'_{10^{-3}}$ corresponds to C^1 in Figure 2. Then, the support of the solution with the largest likelihood value, denoted as C^2 in Figure 2, is extended using 5 (10) subsets SS_k^3 of V^3 . Again, the problems are solved simultaneously by 5 (10) processors. The support C^3 corresponding to the largest likelihood value attained at this stage is extended with subsets of V^4 , etc. Figure 2 depicts the situation when the greatest likelihood value among all suitable sets $C^{i-1} \cup SS_j^i, 1 \leq j \leq n_i$, is attained on $C^{i-1} \cup SS_i^i$.

Intuitively, a solution obtained according to this scheme cannot be better than its counterpart found by the sequential algorithm. In fact, passing from k to $k + 1, k \geq 2$, the parallel method retains only the information contained in one of the 5 (10) subsets SS_i^k . Even if this subset corresponds to the

largest likelihood value, the remaining 4 (9) subsets can contain useful information as well.

Figure 2 Parallel algorithm, passage from $i - 1$ to $i, i \geq 2$



In order to compare the two techniques for sampling mutants, we consider the sequential as well as the parallel algorithms with both the natural and the uniform distribution. These numerical experiments were conducted for the case of the common for all industries migration matrix P and for the case of industry-specific migrations matrices.

4 Numerical results

Tables 3–5 demonstrate that, all other things being equal, the sequential algorithm always achieves a greater likelihood value than its parallel analogue. This is an intuitively plausible result. In the case of industry specific migration matrices, it is surprising that the algorithm employing the uniform distribution performs typically better than its counterpart using the natural distribution. A richer variety of chromosomes in the sample is a plausible explanation of this phenomenon. In fact, generating mutations with equal rather than with different probabilities can imply an evener coverage of the set V^n . There are at least two conceptual explanations for the greater relative increase of the optimal value in the case of the common P . On the one hand, the higher percentages in Table 2 imply that the sets of mutants with at most n mutations are better represented in this case. On the other hand, the blocks are obtained by solving a basic optimisation problem where the migration matrix P is the same for all industries. Therefore, the corresponding block-structures may not completely account for the effect of industry-specific migration matrices $P^{(s)}$.

Table 3 $\frac{\ln L^* - \ln L^0}{\ln L^0} 100$, at most k deterministic mutations

$k = 2$			
	$k = 1$	Sequential algorithm	Parallel algorithm
P	36.63	46.69	43.01
$P^{(s)}$	13.77	21.46	17.69

Table 4 $\frac{\ln L^* - \ln L^0}{\ln L^0} 100$, sequential algorithm, at most k deterministic and random mutations

	Natural distribution			Uniform distribution		
	$k = 3$	$k = 4$	$k = 5$	$k = 3$	$k = 4$	$k = 5$
P	48.52	50.49	51.26	47.41	47.76	48.53
$P^{(s)}$	26.53	28.55	28.86	26.77	29.15	29.20

Table 5 $\frac{\ln L^* - \ln L^0}{\ln L^0} 100$, parallel algorithm, at most k deterministic and random mutations

	Natural distribution			Uniform distribution		
	$k = 3$	$k = 4$	$k = 5$	$k = 3$	$k = 4$	$k = 5$
P	47.38	48.95	49.73	46.66	47.29	48.47
$P^{(s)}$	23.66	26.22	26.52	24.03	25.84	26.63

Table 6 Number of binary vectors in 10^{-3} -support of D^*

Sequential algorithm		Parallel algorithm					
Nat. dist.	Unif. dist.	Nat. dist.	Unif. dist.	Nat. dist.	Unif. dist.		
P	$P^{(s)}$	P	$P^{(s)}$	P	$P^{(s)}$		
51	48	45	55	51	49	48	45

Table 7 Frequencies of identical blocks, P

N	Sequential algorithm		Parallel algorithm	
	Nat. dist.	Unif. dist.	Nat. dist.	Unif. dist.
1	1.96	11.11	0.00	4.17
2	43.14	40.00	29.41	35.42
3	39.22	8.89	47.06	18.75
4	15.69	40.00	23.53	41.67
5	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00

Table 8 Frequencies of identical blocks, $P^{(s)}$

N	Sequential algorithm		Parallel algorithm	
	Nat. dist.	Unif. dist.	Nat. dist.	Unif. dist.
1	2.08	3.64	2.04	6.67
2	29.17	49.09	22.45	31.11
3	54.17	27.27	53.06	11.11
4	14.58	18.18	20.41	44.44
5	0.00	1.82	2.04	6.67
6	0.00	0.00	0.00	0.00

Tables 6–8 characterise different aspects of the heuristic solutions reported here. In all cases, the chromosomes involved in a solution can contain up to five mutations. Table 6 demonstrates that the threshold of 10^{-3} is sufficiently large to bound the 10^{-3} -support of D^* . Frequencies of economic scenarios containing exactly N identical blocks are given in Tables 7 and 8. This is a

different way of characterising the number of mutations involved in a solution. The frequencies are expressed in percent. $N = 1$ corresponds to the economic scenarios, where all blocks are different. In other words, there are five different mutations of a block-structure. If $N = 2$, the row contains frequencies of economic scenarios with exactly two identical blocks. Therefore, there are four mutations in this case. At most two of them can coincide. In other words, economic conditions affecting couples of industries coincide. If $S = 6$, there can be one, two or three such couples. The economic conditions affecting the couples differ and they differ from the economic conditions affecting the remaining industries. Block-structures with a single mutation correspond to $N = 5$. They are present only in the solutions corresponding industry-specific Markovian matrices. Since none of the solutions contains block-structures, all percentages are equal to zero if $N = 6$.

5 Conclusions

Four GAs are suggested for estimating the distribution governing dependent credit-rating migrations of debtors belonging to $M = 7$ credit classes and $S = 6$ industries. Migration counts is the only input required by these algorithms. A S&P's dataset was used. Since the sample space contains 2^{42} points, while some 10^5 counts are available, conventional statistical techniques are not straightforwardly applicable in this case. There are two parallel GAs considered and two sequential GAs. To avoid brute-force calculations, two distributions for generating mutants are tested: a natural and an uniform. The natural distribution implies path-dependence of the selection process. In fact, it generates mutations with probabilities proportional to the frequencies of the corresponding alleles in the initial population. Each of the GAs employs a threshold that limits from above the population size by eliminating the genotypes underrepresented in the population. The tested value 10^{-3} was chosen according to the recommendations of the BAs. All other things being equal, a sequential GA achieves a higher likelihood value than its parallel analogue. Considering at most $2.3 \cdot 10^{-8}\%$ of all 2^{42} possible economic scenarios, the logarithm of the likelihood value increased by up to 51.26% against the level corresponding to the initial population.

Since not all of $2^{M \times S}$ binary strings can be interpreted as economic scenarios, associating a gene with a single locus of the string is not possible in our case. According to the maximum likelihood principle, we identify alleles that comprise M subsequent loci of a string. Conceptually, they represent combinations of favourable and adverse economic factors affecting every debtor in an economy, where M credit classes are considered. If the number of the alleles identified in this way equals K , the search-space reduces

$$\left(\frac{2^M}{K}\right)^S \text{ times.}$$

References

- Allman, E.L., Matias, C. and Rhodes, J.A. (2009) 'Identifiability of parameters in latent structure with many observed variables', *Annals of Statistics*, Vol. 37, No. 6A, pp.3099–3132.
- Bangia, A., Diebold, F.X., Kronimus, A., Schagen, C. and Schuermann, T. (2002) 'Ratings migration and the business cycle, with applications to credit portfolio stress testing', *Journal of Banking and Finance*, Vol. 26, Nos. 2–3, pp.445–474.
- Boreiko, D.V., Kaniovski, S.Y., Kaniovski, Y.M. and Pflug, G.C. (2017) 'Traces of business cycles in credit-rating migrations', *PLOS ONE*, 20 April, pp.1–29 [online] <https://doi.org/10.1371/journal.pone.0175911>.
- Cai, X., Hu, Z. and Chen, J. (2020a) 'A many-objective optimization recommendation algorithm based on knowledge mining', *Information Sciences*, Vol. 537, pp.148–161.
- Cai, X., Hu, Z., Zhao, P., Zhang, W. and Chen, J. (2020b) 'A hybrid recommendation system with many-objective evolutionary algorithm', *Expert Systems with Applications*, Vol. 159, No. 30, in Progress [online] <https://doi.org/10.1016/j.eswa.2020.113648>.
- Fei, F., Fuertes, A-M. and Kalotychou, E. (2012) 'Credit rating migration risk and business cycles', *Journal of Business Finance and Accounting*, Vol. 39, Nos. 1–2, pp.229–263.
- Glasserman, P. and Young, H.P. (2016) 'Contagion in financial networks', *Journal of Economic Literature*, Vol. 54, No. 3, pp.779–831.
- Gupton, G.M., Finger, C.C. and Bhatia, M. (1997) *CreditMetrics – Technical Document*, J.P. Morgan Inc., New York, NY, USA.
- Holland, J.H. (1992) *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, USA.
- Kaniovski, Y.M. and Pflug, G.C. (2007) 'Risk assessment for credit portfolios: a coupled Markov chain model', *Journal of Banking and Finance*, Vol. 31, No. 8, pp.2303–2323.
- Ma, X., Chen, Q., Yu, Y., Sun, Y., Ma, L. and Zhu, Z. (2020) 'A two-level transfer learning algorithm for evolutionary multitasking', *Frontiers in Neuroscience*, 14 January, Vol. 13, p.1408 [online] <https://doi.org/10.3389/fnins.2019.01408>.
- Ma, X., Li, X., Zhang, Q., Tang, K., Liang, Z., Xie, W. and Zhu, Z. (2019) 'A survey on cooperative co-evolutionary algorithms', *IEEE Transactions on Evolutionary Computation*, Vol. 23, No. 3, pp.421–441.
- Ma, X., Liu, F., Qi, Y., Gong, M., Yin, M., Li, L., Jiao, L. and Wu, J. (2014a) 'MOEA/D with opposition-based learning for multiobjective optimization problem', *Neurocomputing*, Vol. 146, pp.48–64.
- Ma, X., Liu, F., Qi, Y., Li, L., Jiao, L., Liu, M. and Wu, J. (2014b) 'MOEA/D with Baldwinian learning inspired by the regularity property of continuous multiobjective problem', *Neurocomputing*, 5 December, Vol. 145, pp.336–352.
- Ma, X., Qi, Y., Li, L., Liu, F., Jiao, L. and Wu, J. (2014c) 'MOEA/D with uniform decomposition measurement for many-objective problems', *Soft Computing*, Vol. 18, pp.2541–2564, <https://doi-org.libproxy.unibz.it/10.1007/s00500-014-1234-8>.
- Ma, X., Liu, F., Qi, Y., Li, L., Jiao, L., Deng, X., Wang, X., Dong, B., Hou, Z., Zhang, Y. and Wu, J. (2016a) 'MOEA/D with biased weight adjustment inspired by user preference and its application on multi-objective reservoir flood control problem', *Soft Computing*, Vol. 20, pp.4999–5023, <https://doi-org.libproxy.unibz.it/10.1007/s00500-015-1789-z>.

- Ma, X., Liu, F., Qi, Y., Wang, X., Li, L., Jiao, L., Yin, M. and Gong, M. (2016b) 'A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables', *IEEE Transactions on Evolutionary Computation*, Vol. 20, No. 2, pp.275–298.
- Ma, X., Zhang, Q., Tian, G., Yang, J. and Zhu, Z. (2018) 'On Techebycheff decomposition approaches for multiobjective evolutionary optimization', *IEEE Transactions on Evolutionary Computation*, Vol. 22, No. 2, pp.226–244.
- Wang, Y., Ding, M., Pan, J.J. and Malone, S. (2017) *Credit Transition Model 2017 Update: Methodology and Performance Review*, Moody's Analytics Inc., Report Number 186801, USA.
- Wozabal, D. and Hochreiter, R. (2012) 'A coupled Markov chain approach to credit risk modeling', *Journal of Economic Dynamics and Control*, Vol. 36, No. 3, pp.403–415.