European Journal of Operational Research 212 (2011) 123–130

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Modelling the profitability of credit cards by Markov decision processes

Meko M.C. So *, Lyn C. Thomas

School of Management, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

ARTICLE INFO

Article history:
Received 20 May 2008
Accepted 14 January 2011
Available online 21 January 2011

Keywords:
OR in banking
Markov decision process
Credit card
Behavioural score
Profitability
Probability of default

ABSTRACT

This paper derives a Markov decision process model for the profitability of credit cards, which allows lenders to find an optimal dynamic credit limit policy. The states of the system are based on the borrower's behavioural score and the decisions are what credit limit to give the borrower each period. In determining which Markov chain best describes the borrower's performance, second order as well as first order Markov chains are considered and estimation procedures developed that deal with the low default levels that may exist in the data. A case study is given in which the optimal credit limit is derived and the results compared with the actual outcomes.

1. Introduction

Since the advent of credit cards in the 1960s, lenders have used credit scoring, both application and behavioural scoring, to monitor and control default risk. However in the last decade the lenders' objectives have changed from minimising default rates to maximising profit. Lenders have recognized that operating decisions are crucial in determining how much profit is achieved from a card and this paper focuses on the most important operating policy: The management of the credit limit. Soman and Cheema (2002) conducted a study on the use of credit limit policies in encouraging spending and found that the availability of additional credit does promote card usage in some consumers.

Currently lenders set credit limits by subjectively determining an appropriate value for each cell in a risk/return matrix. They decide on a credit limit for each combination of risk band and average balance, where the average balance is considered to be a surrogate for the return to the lender from that customer. This approach is static in that it does no consider how the customer's default risk and the customer's profitability will change over time. Nor is there any optimization involved in deciding what credit limits to choose.

We propose using a Markov decision processes (MDP) to improve the credit limit decision. A MDP model determines the optimal sequence of credit limit decisions by considering the evolution of a customer's behaviour over time. It calculates the profitability of a credit card customer under the optimal dynamic credit limit policy.

Lenders keep a wealth of historical credit card data, including customers' behavioural scores each month. Behavioural scores are a way of assessing customers' default risk in the next year. Building the Markov decision process model on behavioural scores has the advantage that most lenders have been keeping this data on customers for a number of years. With the advent of the Basel Accord in 2008, lenders are required to keep such data for five years and are encouraged to keep it through a whole economic cycle.

MDPs have been used in a number of different contexts (Heyman and Sobel, 1982; Ross, 1983; White, 1985, 1988, 1993; Kijma, 1997). The first application of MDPs in consumer credit was by Bierman and Hausman (1970) who looked at the repayment of a loan where no further borrowing was allowed. The model assumed the repayment of the customer followed a prior probability distribution. Using a Bayesian approach, the model revises the probability of repayment in the light of the collection history. Modifications of the basic model were made both in the accounting rules (Dirickx and Wakeman, 1976) and in the form of the Markov chain (Frydman et al., 1985). A MDP model to optimise consumer lifetime value was developed by Trench et al. (2003). Here the actions in the MDP model were the consumer's credit card limit or the interest rate charged. So the actions in that paper are similar to the ones in this paper. However their state space did not involve behavioural scores nor were they concerned with the problems of estimating the transition probabilities if there are low default rates. Instead they used a six dimensional state space each dimension having only two or three categories describing the recency and frequency of purchases and payments. The authors developed mechanisms for reducing the size of the transition matrix through

* Corresponding author. Tel.: +44 0 7796035843.
E-mail addresses: M.So@soton.ac.uk (M.M.C. So), L.Thomas@soton.ac.uk (L.C. Thomas).

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.01.023

merging states. Ching et al. (2004) used MDPs to manage the customer lifetime value generated from telecommunication customers, but again the state space used in that study involved marketing factors not risk factors and the decision was whether to implement promotions.

This paper is the first to use behavioural score bands as the basis for MDP models. The advantage of basing the model on behavioural scores is considerable. Almost all lenders calculate such scores every month for every borrower as a basis for their Basel Accord probability of default calculations and as a way of segmenting the population by risk.

When modelling real problems using Markov decision processes, the curse of dimensionality (Puterman, 1994) can mean the state space is very large and that one would need a large amount of data to obtain robust estimates of the transition probabilities. Using behavioural scores helps to overcome this first difficulty because it itself is a "sufficient statistic" of the risk of the account and already contains information from a number of different characteristics. Also by aggregating states one can obtain a simple but meaningful state space. In our case we make part of each state an interval of behavioural scores and similarly combine possible credit limits into bands, which make up the other part of the state. The actions are then which of these credit limit bands should be applied to the borrower in the next month.

Acquiring enough data to calculate robust estimators of the transition probabilities is not a problem in the consumer credit context because of the size of the data sets available to lenders. The only problem is that with some portfolios of loans, the number of movements directly into default from some states is so low (quite possibly zero) that the resultant zero transition probability estimate may affect the structure of the Markov chain. This problem of estimating default probabilities in low default portfolios was also highlighted in the Basel Accord. We therefore use an approach suggested in that context (Pluto and Tasche, 2006) and extended by the UK regulators (Benjamin et al., 2006) which ensures the resulting Markov chain model is robust and conservative. The conservativeness is reasonable as one would prefer the model to underestimate rather than over estimate the profitability of a credit card account.

The main contribution of this paper is to show how one can use Markov decision process models based on states consisting of behavioural score bands – scores which most lenders calculate on a monthly basis – to determine optimal credit limit policies. Such policies maximise the expected profitability of each borrower and allow for policy rules such as never decreasing the credit limit which lenders may impose. The paper deals with some practical issues of modelling these transitions such as avoiding zero probability transition estimates which could affect the structure of the Markov chain. The approach is applied to a case study of a Hong Kong credit card portfolio of over 1.4 million accounts. The results of the modelling are compared with what actually happened in the portfolio.

The rest of the paper is organized as follows: Section 2 describes the MDP model formulation. Section 3 discusses the estimation of the transition probabilities including the probabilities of defaulting immediately. Section 4 presents the practical issues in applying the MDP model to the real credit card data while the results of the case study are described in Section 5. The final section draws some conclusion on the model and the resultant case study.

2. The MDP model

Consider a discrete state, discrete time discounted Markov decision process with decision epochs T (indexed by $t = 1,2,\ldots,T$) based on a state space S. Each state in the state space consists of two parts-which behavioural score band the borrower is in and what is the borrower's current credit limit band. The state space thus consists of the current credit limit band represent by L (indexed by $l = 0,1,\ldots,L$) and the current behavioural score band I (indexed by $i = 0,1,\ldots,I$). The latter is extended to allow absorbing states corresponding to various types of default or closure of the account. In our model the actions are limited to keeping the credit limit at its current level or raising it to a higher limit band. This policy of not decreasing credit limits is used by many lenders but the methodology we describe will not change if this restriction is dropped. With this limitation the action set is defined as $A_l = \{l' : l \leqslant l'\}$.

Two further elements need to be defined to complete the Markov decision process model. Let $p(i'|l,i)$ be the probability that if l is the current customer's credit limit band and the customer is in behavioural score band i, then the next period the customer will be in behavioural score band i'. Secondly let $r(l,i)$ be the profit obtained in the current period from a customer with credit limit l and in behavioural score band i.

The objective is to maximise the discounted profit obtained from the customer over the next t periods where the discount factor λ describes the time value of money. This leads to the following optimality equation for $V_t(l,i)$, the maximum expected profit over the next t periods that can be obtained from an account which is currently in behaviour score band i, and with a credit limit of l:

$$V_t(l,i) = \max_{l' \in A_l} \left\{ r(l,i) + \sum_{i'} p(i' \mid l,i) \lambda V_{t-1}(l',i') \right\}. \tag{1}$$

The right-hand-side of (1) corresponds to the profit over the next t periods if we change the credit limit to l' from l at the end of the current period for an account with behavioural score state i. We assume that a change in the credit limit takes effect in the next time period since such a decision is usually included in the next monthly balance statement sent to the customer. Removing this delay makes no difference to the methodology though of course the optimality equation will be slightly different. The profit to the lender from the credit card during the current period is $r(l,i)$. The $p(i'|l,i)$ is the probability that the behavioural score changes to band i'. In that case, the profit on the remaining $t-1$ period is $V_{t-1}(l',i')$. The discount factor λ is introduced because the profits in the remaining $t-1$ periods actually occur one period after those used in calculating $V_{t-1}(l',i')$ since that assumes the $t-1$ periods start now.

The optimality principle says that the decision l', which maximises the right hand side of (1) is the one to use when there are t more periods to go if one wants to maximise the sum of the future profits, when credit limits can only remain the same or be increased.

3. Estimating the PDs of low default portfolios

Maximum likelihood estimators are used to estimate the transition probabilities of a Markov chain. In the Markov chain described in Section 2, let $n_t(l,i)$ be the number of accounts in state (l,i) at time t and let $n_t^j(l,i)$ of them move to behavioural score state j at time $t+1$. Assuming the Markov chain is stationary means the maximum likelihood estimate $\tilde{p}(j \mid l,i)$ for the probability $p(j|l,i)$ is

$$\frac{\sum_{t=1}^{T-1} n_t^j(l,i)}{\sum_{t=1}^{T-1} n_t(l,i)}.$$

In reality, moving directly to the default state is a rare event, particularly for high value (low risk) behavioural scores. There may be no examples in the data of transaction from certain states (l,i) to the default state D. Thus it is possible that $p(D|l,i)$ may be very small or even equal to zero. Putting such estimates into the MDP model