Discrimination by Decision tree: A new Splitting Criterion

CHAMLAL HASNA[1], BAZZI MEHDI[2]

[1] Laboratory of Computer Sciences & Decision Aiding Faculty of Science Ain chock, Casablanca ,Morocco.
 (E-mail: chamlal@yahoo.com)
[2] Laboratory of Computer Sciences & Decision Aiding Faculty of Science Ain chock, Casablanca,Morocco.
 (E-mail: bazzimehdi@gmail.com)

Abstract. This paper presents new method to build a decision tree with new Splitting criterion. Splitting criterion specifies the best splitting variable and its threshold for further splitting in a tree. This new criterion is based on concept of Preordonnance [1] whose advantage is its validity for an explanatory quantitative, categorical or ordinal variable. In the case of a categorical variable, the Splitting criterion is equivalent to d on the chi-2 contingency criterion. In the case of quantitative explanatory variable Chamlal and Chah proved that the new criterion is more accuracy than that one based on the minimization of the intra class variance and the maximization (Fischer Snedecor test) [1]. The final model of the decision tree is applied with real dataset which cover the area of Marketing in the company of insurance. The model results are compared with three other classification techniques; CHAID decision tree, logistic regression and Discriminant analysis.

Keywords—Decision Tree; Splitting Criterion,Preordonnace,Pshycor.

1. Introduction

The decision tree is one of the most important knowledge representation approaches which attempts to resolve problems of regression or classification. The tree is built by recursive process from top to bottom using splitting rules. The splitting criterion in a tree is specified by choosing the best splitting variable and its threshold for the further split. When applied to classification problems, terminal nodes represent classification groups. The following figure shows an example for the DT.
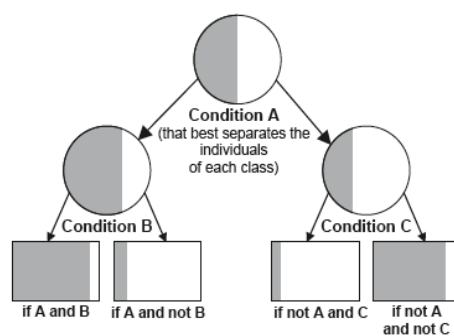

Fig.1.Example DT

Despite strong competitors like logit regression and linear regression, DTs have several advantages. Indeed, the selected model is a set of rules that can be easily implemented in the computer systems of companies, the DTs copes well with heterogeneous data, missing value and non linear effects Gepp and al. [3].

Different building algorithms can be used to generate DTs that have a large variation in classification and prediction accuracy. CART algorithm Breiman and al. [4] selects split using the Twoing Criterion, tolerate all kind of variables, can be used for classification and regression problems and prune the tree by cost-complexity pruning. Another DT algorithm (C5.0) uses

information gain as splitting criterion, the C5.0 Quinlan [5] is an evolution of C4.5 Quinlan [6] and ID3 Quinlan [7], C5.0 works by aiming to maximize the information gain achieved by assigning each individual to a branch of the tree. It shares with CART its suitability for the investigation of all kind of variables, its exhaustive search for all the possible splits and devise for optimizing the tree by the building of maximal tree followed by its pruning. However, its pruning procedure is different from that of CART.C5.0 uses error-based pruning. Due the wide applicability of DT algorithm for data exploration, classification and regression, numerous researchers focus on improving this principal algorithm, or even propose a new splitting criterion with exclusive characteristics.

There are several old and new algorithms that work totally different from CART, C5.0.CHAID algorithm Kass [8] as an old one uses the χ^2 test to define the most significant variable for each node, so it can only be used with discrete or qualitative independent variables. An evolution of the CHAID algorithm Loh and Shih [9] (QUEST) supports univariate and linear combination splits. For each split, the measure of association between each input attribute and the target attribute use the ANOVA F-test or Levene's test (for nominal attributes). Quadratique Discriminant Analysis (QDA) is applied to find the optimal splitting point for input attribute. Another algorithm based on the area under the Receiver Operating Characteristics (ROC) curve, named Roc-tree Maruf and al. [10], and has better simulation results on classifying sample databases then very common used algorithm like C4.5 and C5.0.

Most of researchers try optimizing the DTs algorithms by improving classic DTs characteristics, like changing splitting criterion, minimizing of DT or improving the DT pruning mechanism by using the Automatic Programming, the researchers Hansen and Olsson [11]could improve the DT pruning mechanism (error-based pruning) of C5.0 decision tree. In Sieling [12] the author talkes about the minimization of DT and its importance in decision making speed, the author has demonstrated that, the minimizing of decision trees is hand to approximate.

The classification-based DTs are used in various fields, such as hospitals Yoshikazu and al.[13], for calculating the probability of cardiac arrest in the emergency department, in banks Wisaeng [14] and insurance Bhowmik [15], for calculating the probability of customer defaults and claims. The authors in Kazunor and al.[16] suggested a concept of adding fuzzy to C4.5 algorithm. The addition of fuzzy to C4.5algorithm resulted in better results in terms of accuracy and interpretability. These examples are just covering some of many cases where classifications are used, and decision trees have been used successfully.

We introduce in this paper a new method to build a decision tree with the criterion based on concept of Preordonnance [1] whose advantage is its validity for an explanatory quantitative, categorical or ordinal variable.

In the section 2 we introduce the splitting criterion. In the section 3, we describe the algorithm of Decision tree with the new splitting criterion. In the section 4 we present an empirical study in order to challenge the new method and before the conclusion, we present some elements of discussion.

2. Background

As proposed DT uses a new criterion of division based on the concept of preordonance to find the best splitting variable and to stop the growth of the tree. Here we introduce this new Splitting criterion.

2.1 Notation and Preliminary calculations

We suppose that we have a qualitative random variable to explain Y and p heterogeneous explanatory variables $(X_1 \dots X_p)$.

E: a sample of n possible values $\{1, 2, 3 \dots n\}$

H : $\{(i, j) \in E^2 \: / \: i < j\}$

$M = |H| = \dfrac{n(n-1)}{2}$

A\B: a complementary set B in A

Definition: we call preordonnance defined on E, a relation *of weak order* denoted P defines on H

Definition: we call order relation, a reflexive and transitive binary relation

A préordonnance P can be encoded in binary or ternary.

Codage ternaire

$T_P(i, j, k, l) = \begin{cases} 1 \text{ si } (i, j) >p \text{ (k, l)} \\ 0 \text{ si } (i, j) =p \text{ (k, l)} \\ -1 \text{ si } (i, j) <p \text{ (k, l)} \end{cases}$

Codage binaire

$B_p(i, j, k, l) = \begin{cases} 1 \text{ si } (i, j) >p \text{ (k, l)} \\ \\ 0 \text{ sinon} \end{cases}$

$$T_P(i, j, k, l) = B_P(i, j, k, l) - B_P(k, l, i, j)$$

T_P is centred $T_P(i, j, k, l) = -T_P(k, l, i, j)$

The >p symbol denotes that the pair (i, j) above the pair (k, l) in the order induced by P on H, =p denotes the fact the pairs are ties, <p denotes the fact that the pair (k, l) preceding the pair (i, j)

A preordonnance P may be induced by a random variable X of any kind

if X is qualitative :	(i, j) P_X(k, l)	$\Leftrightarrow$ (X(i)=X(j)) et (X(k) $\neq$ X(l))
if X is ordinale :	(i, j) P_X(k, l)	$\Leftrightarrow$ |r(i)-r(j)| < |r(k)-r(l)|
if X is quantitative :	(i, j) P_X(k, l)	$\Leftrightarrow$ |X(i)-X(j)| < |X(k)-X(l)|

Given two preordonnances P and Q, and, T_P and T_Q associated coding, the covariance coefficient (resp. correlation) induced on all preordonnance, a measure of association denoted ψ_{cov} (resp. (ψ_{cor}))

$$\psi_{cov} = \text{cov}\left(T_P, T_Q\right) = \frac{\sum T_P(i, j, k, l) T_Q(i, j, k, l)}{M(M-1)}$$

$$\psi_{cor} = \text{cor}\left(T_P, T_Q\right) = \frac{\sum T_P(i, j, k, l) T_Q(i, j, k, l)}{\sqrt{\sum T_P(i, j, k, l)^2 \sum T_Q(i, j, k, l)^2}}$$

The summation covers all $\{((i, j), (j, k)) \in H * (H - \{(i, j)\})\}$.

In the absence of ties

$$\sum T_P(i,j,k)^2 = M(M-1) => \psi_{cov} = \psi_{cor}$$

2.2 Splitting criterion

Given a set of (ordinal, quantitative, nominal) explanatory variables and a qualitative dependent variable Y. The purpose of this test is to find the explanatory variable X that maximizes the association within the meaning of ψ_{cov} (ou ψ_{cor}), with the dependent variable for each node:

$$\max_{i=1...p} \psi_{cor}(P_{X_i}, P_Y)$$

$$\psi_{cov}(P_X, P_Y) = \frac{2\sum_{(i,j)\in H \; et \; (k,l)\in H-(i,j)}(B_X(i,j,k,l) - B_X(k,l,i,j))B_Y(i,j,k,l)}{M(M-1)}$$

$$T_X(i,j,k,l) = B_X(i,j,k,l) - B_X(k,l,i,j)$$

$$T_Y(i,j,k,l) = B_Y(i,j,k,l) - B_Y(k,l,i,j)$$

An explanatory variable which maximize this criterion must bring together two individuals unified by the dependent variable, and make away two individuals separated by this variable; in other words, the explanatory variable must satisfy this condition:

$$\max_{Y(i)=Y(j)}|X(i) - X(j)| \; < \; \min_{Y(k)\neq Y(l)}|X(k) - X(l)|$$

This ensures compactness and separability of groups described by the explanatory variable.
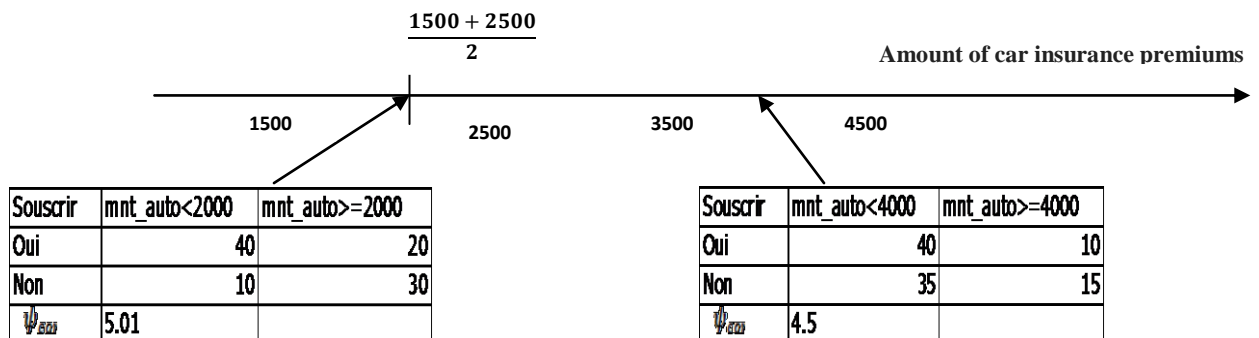
The criterion based on ψ_{cor} is finer because it measures the degree of agreement (ψ_{cor} or ψ_{cov} positive) or disagreement (ψ_{cor} or ψ_{cov} negative) between preordonnances..

In the following sections, we consider the ψ_{cor} as a splitting criterion. It specifies the best splitting variable and its thre threshold for further splitting in a tree.

2.3 Selection of Split point

Given a node T, it is assumed that the random variable X was chosen to split the node. In this case we must seek an optimal cutting X.

➤ If X is continuous, we sort at the first time the variable, then we test each possible threshold between two values of the variable by calculating the ψ_{cor} ou (ψ_{cov}) of the binary variable created.

$$\frac{1500 + 2500}{2}$$

Amount of car insurance premiums

1500   2500   3500   4500

Souscrir	mnt_auto<2000	mnt_auto>=2000
Oui	40	20
Non	10	30
ψ_{cor}	5.01	

Souscrir	mnt_auto<4000	mnt_auto>=4000
Oui	40	10
Non	35	15
ψ_{cor}	4.5	

Discrétization example

314

> ➤ If the variable is qualitative, we look for a regrouping of the variable(merging categories) that give the best measure ψ_{cor} (ou ψ_{cov}) and divide the node t in as many nodes as modality son after regrouping of the variable

2.4 <u>Stopping Criterion</u>

<u>Proposal [1]</u>: the expressions τ_1 et τ_2 for Rate Kendall [18] induces measures of association between two variables of any kind (ψ_{cov} et ψ_{cor}) :

$$\psi_{cov}\left(P_{X_1}, P_{X_2}\right) = \tau_1(r_1, r_2)$$

$$\psi_{cor}\left(P_{X_1}, P_{X_2}\right) = \tau_2(r_1, r_2)$$

The demonstration of the equations above come from the proposition s (1) and (2) in the Annex and the fact that:

$$\begin{cases} r_m(i,j) < r_m(k,l) <=> (i,j) >_{p_{X_m}} (k,l) <=> T_{X_m}(i,j,k,l) = 1 \\[2mm] r_m(i,j) = r_m(k,l) <=> (i,j) =_{p_{X_m}} (k,l) <=> T_{X_m}(i,j,k,l) = 0 \qquad m = 1 \text{ or } 2 \\[2mm] r_m(i,j) > r_m(k,l) <=> (i,j) >_{p_{X_m}} (k,l) <=> T_{X_m}(i,j,k,l) = -1 \end{cases}$$

$\forall \, (i,j) \in H, \, \forall \, (k,l) \in H$

This result allows us to have a criterion for stopping the deepening of the tree coherent with the select criterion.

Indeed, we accept the division if the ψ_{cor} (ouψ_{cov}) calculated on a peak is significantly higher than a threshold that we fixate. The Formalization involves a statistical hypothesis test: the null hypothesis (H0) is the independence of the explanatory variable X with the dependent variable Y. if ψ_{cor} (ouψ_{cov}) calculated is above the threshold correspondent theoretical risk of error that can be fixed (often a risk level of 5%), we accept the Splitting[1].

2.5 <u>Handlin Missing values</u>

There are cases where the explanatory variable X has a rate of missing (and not the whole variable), so in this case we use «surrrogates variables» that can leave individuals (pretty near) in the same way as the original variable, these variables are called equally reducingNakache and Confais [19].

3. Presentation of the Decision tree algorithm

In this section we recapitulate the most steps followed in order to construct the decision tree with the new splitting criterion.

Algorithm: Decision tree with new splitting criterion

Input:
- Training dataset described by categorical features $X_1 \dots X_d$ and continuous features $X_{d+1} \dots X_f$
- L: Minimum of individual in the leaf before splitting it
- Minimum of individual in each leaf.
- Confidence threshold α

The algorithm:

Create a new tree T with single root node

IF *individual in the node> L* then

    Mark T as leaf with the most common value of Y in dataset as a label

ELSE

1. IF d > 0 (there are continuous features)
   a. FOR i= 1,……,d
      ψ_i ⟵ the ψ_{cor} statistic for feature X_i
   b. $best_1$ ⟵ arg $\max_{i=1\dots d} \psi_i$
   c. α_1 ⟵ p-value of adequate ψ distribution for feature $best_1$
      END FOR
2. IF f > d (there are categorical features)
   a. FOR i= d+1,……,f
      p_i ⟵ p-value of the χ^2 test of independence between feature X_i and class labels
   b. $best_2$ ⟵ arg $\min_{i=d+1\dots f} p_i$
   c. α_1 ⟵ p_{best_2}
      END FOR
3. α_{12} ⟵ min(α_1, α_2)
4. IF $\alpha_{12} = \alpha_1$ THEN $best_{12}$ ⟵ $best_1$ ELSE $best_{12}$ ⟵ $best_2$
5. Split $best_{12}$
6. IF d > 0 (there are continuous features)
   a. Sort $best_{12}$, then test each possible threshold between two values of the $best_{12}$ by calculating the ψ_{cor} of the binary variable created.
   b. Select threshold which gives the best ψ_{cor}
7. IF f > d (there are categorical features)
   a. look for a regrouping of the $best_{12}$ (merging categories) that give the best measure ψ_{cor} and divide the node t in as many nodes as modality son after regrouping of the $best_{12}$
   b. Select variable after regrouping which gives the best ψ_{cor}.
8. Split node

END IF

9. FOR each child nodes
       IF *individual in the node< L* THEN Stop the growing of the tree
       ELSE Repeat steps 1-8
   END FOR

---

4. Results Marketing Studies:
4.1 Datasets Description:

We describe in Table 1 the variables used in the empirical studies (Marketing) in order to compare the ψ_{cor} and the measure of association. There are three kinds of variables:

- ➢ Socio demographic variables
- ➢ behavior variables
- ➢ Dependent variable or variable to predict.

Table 1: The variables used in the computational techniques to identify the buyer Customers

Property	value	
Source	Sentent Machine Research, Amsterdam, 2000	
Sample Size	5822 total:5474 Buyer and 348 Not Buyer	
Dependent variable	Binary variable which describe whether the customers purchases the insurance caravan or not.	
Buyer (1)	Purchase an insurance caravan contract	
Not Buyer (0)	Don't Purchase an insurance caravan contract	
Explanatory Variables (18)		**Type**
mt_auto	Amount of automobile insurance premium	continues
mt_incendie	Amount of fire insurance premium	continues
mt_RC	Preium Civil Liability insurance	continues
revenu_moyen	Average income	continues
mt_securite_soc	Premium social security insurance	continues
mt_moto	Premium motorcycle insurance	continues
mt_RC_entreprise	Premium civil liability insurance company	continues
mt_bateau	Amount of the premium ship insurance	continues
mt_tracteur	Amount of the insurance premium tractor	continues
mt_cyclomoteur	Amount of the insurance premium motorcycle	continues
mt_invalidite	Premium disability insurance	continues
mt_velo	Premium cyclomoteur insurance	continues
mt_accident_famil	Premium family accident insurance	continues
nbpers_au_foyer	number of homemake	continues
age_moyen	Average age	continues
locataire	tenant	discrete; for example, if the tenant variable is 7, it means that he lives an area composed of 76-88% Catholic
proprietaire	Owner	discrete.like tenant
marie	Marié	discrete.like tenant
niv_etud_bas	Low level study	discrete.like tenant
PCSagri	Socioprofessionale categoriy	discrete.like tenant
concubin	concubine	discrete.like tenant
autre_relation	other Relation	discrete.like tenant
celibataire	Unmarried	discrete.like tenant
niv_etude_haut	High level study	discrete.like tenant
sans_religion	Without Religion	discrete.like tenant
auto0	Not owning car	discrete.like tenant
catholique	Catholic	discrete.like tenant
autre_religion	other religion	discrete.like tenant
niv_etud_moy	Average level study	discrete.like tenant
auto1	owning car	discrete.like tenant
PCSinter	Socio Professional category intermediate	discrete.like tenant
mt_assur_vie	Premium life insurance	discrete.like tenant
PCSouvr	Socio professionnal category worker	discrete.like tenant
protestant	Protestant	discrete.like tenant
PCStop	Socio-professional category operator	discrete.like tenant
avec_enfant	with children	discrete.like tenant
PCScadre	Socio professionnal category executive	discrete.like tenant
sans_enfant	childless	discrete.like tenant
PCSouvr_quali	Socio professionnal category worker	discrete.like tenant
auto2	Two car possession	discrete.like tenant

4.2 Final Model Development
- ➢ The decision tree model with the new Spitting Criterion has ultimately selected six variables:

- o The amount of car insurance premiums
- o The amount of fire insurance premiums
- o The high level of study
- o The amount of RC insurance premiums
- o Purchasing power
- o Owner

➢ The CHAID decision tree has ultimately selected six variables:
- o The amount of car insurance premiums
- o The amount of fire insurance premiums
- o Socio professional category intermediate
- o The low level of study
- o The high level of study
- o The amount of RC insurance premiums

➢ The resulting of logistic Regression ,which is significant at the 5% level, is :

Logit Y=-2.9411 + (mt_auto>=6)*1.4911 - (mt_incendie<=2)*1.1303 - (mt_incendie>=5)* 1.1303 -(mt_incendie= 3)* 0.2967 - (revenu_moyen<=3)*0.6029 + (niv_etud_bas<=2)*0.5460

Here, probability that the ith customer is buyer is obtained as follows:

$$P \text{ (customer }_i \text{ is buyer)} = e^{Logit\ i} / (1 + e^{Logit\ i})$$

➢ The Discriminant Analysis is also significant at the 5% level with the following equation:

Discriminant Score=-15.50 + 0.242 mt_auto + 1.384 pouv_achat + 0.428 mt_bateau + 0.95mt_RC + 1.546niv_etud_bas + 1.386marie

Test and Training validation model :

	mt_auto	mt_incendie	mt_bateau	mt_RC	revenu_moyen	pouv_achat	niv_etud_bas	niv_etud_haut	proprietere	marie	Training AUC	Test AUC
DT with new Splitting criterion	X	X		X		X		X	X		76%	71%
CHAID decision tree	X	X		X			X	X			78%	73%
Logistic Regression	X	X			X		X				78%	74%
Discriminante Analysis	X		X	X		X				X	75%	74%

Analysis of result:

➢ All the LA, DA, CHAID DT and DT with new splitting criterion models included the variable *Amount of automobile insurance premium*, which indicates that it was the most important in discriminating between potential buyer and others. This can be explained by the fact that this amount indirectly integrates the power of the vehicle, which determines much of its ability to tow a caravan.

➢ To Conclude with the Discriminant Analysis, we see that the AUC (74%) in validation is nearly as elevated as the AUC (75%) in learning sample, which is the sign of an excellent robustness of the model

➢ The best logistic model was obtained by an appropriate selection of the explanatory variables, based on a combination of statistical tests. We sought to optimize validation AUC (not learning AUC), but we did not only relied on the latter. Last model obtained

was therefore selected as offering the best compromise between simplicity and performance with its 4 variables, 6 attributes and the validation AUC is equal to 74%. By contrast we have practically the same level of accuracy with the other methods; Discriminant Analysis (75%), CHAID DT (73%) and DT with new splitting criterion (71%) but with 6variables.

➢ For the DT with the new Splitting criterion, the model finally selected contains six variables. The AUC in training sample is 76% equal to the AUC of the Discriminant Analysis and very close to that of the logistic regression and CHAID decision tree. In the other hand, validation AUC(71%) has decreased by 5%. The significant difference between the area under the ROC curve in learning and validation reports an overfitting and lack of robustness of the tree, which can be explained by the small leaves (we have only 348 customers which have the target variable equal to 1)**.**

5. Discussion

➢ The DT with the new splitting criterion yields comparable to that of logistic regression which according to practitioners remains by far the most used method performance.

➢ The new selection criterion based on the notion of preordonnance only interested in positive associations. A negative association is a disagreement between the variable to predict and the explanatory variable; therefore the latter cannot be selected. Thus, we show that the proposed method allows to identify the potentially interesting variables and saves us from wasting time on variables that are not interesting from a statistical discrimination

➢ The variable **amount of the contribution of the tractor** insurance (as described in the analysis of Marketing study) is concentrated on a single value, this variable is irrelevant for discrimination; we cannot keep it initially. However, our objective is to show that the ψ_{cor} detects this kind of variables (variable with unbalanced distribution) and penalize them.

➢ For continuous variables, we chose a binary discrimination. The drawback of this binary structure is that it produces trees which are 'narrow' but may be very deep, making the trees rather complex and difficult to read in some case.

➢ Using the criterion ψ_{cor} (or ψ_{cov}) requires the measure of association between each pair of individual to be classified and consequently a storage problem occurs especially if the sample is large. However, with the evolution of Big Data technology, strategies have been proposed to address the problem of storage and improve the speed of the system to large databases, without degrading performance Granville [20].

6. Conclusion

In this paper, a new splitting criterion to build decision tree is proposed, the same criterion (ψ_{cor}) is used for stopping the growing of the tree. The number of possible separation condition allowed by the best splitting variable depends on its type. A continuous variable allows a single separation condition. For a qualitative variable, the node is divided in as many sub-nodes as modality son after regrouping this variable. The empirical studies show that the decision tree with the new splitting criterion gives as accurate results as those given by the CHAID DT, logistic regression or Discriminant Analysis.

Annex[18]

Kendall tau τ

Given $r_1, and\ r_2$ the rank variables induced by two random variables X and Y.

Case of absence of ties

Given P the number of pair (i, j) Є H which verifies:

$(r_1(i) - r_1(j))(r_1(i) - r_1(j)) > 0$ And Q the number of of pair (i, j) Є H which verifies: $r_1(i) - r_1(j))(r_1(i) - r_1(j)) < 0$.Set S = P – Q, the maximal number which can reach P is $\frac{n(n-1)}{2}$ and τ(X,Y) is defined as:

$$\tau(X, Y) = \frac{S}{M}$$

Proposition 1. Given A and B the coding comparison pair of the rank variables r_1 and r_2:

$$A_{ij}=\begin{cases} 1 \text{ if } r_1(i) < r_1(j) \\ \\ -1 \text{ if } r_1(i) > r_1(j) \end{cases} \qquad B_{ij}=\begin{cases} 1 \text{ if } r_2(i) < r_2(j) \\ \\ -1 \text{ if } r_2(i) > r_2(j) \end{cases} \qquad A_{iI}=B_{ii}=0$$

Alors τ(X, Y)=cor (A, B)=cov (A, B)

Indeed: $\sum_{i\text{H}\neq j} A_{ij}B_{ij} = 2\ S = 2(P - Q)$ and $\sum_{i\text{H}\neq j} A_{ij}^2 = \sum_{i\text{H}\neq j} B_{ij}^2 = n(n - 1)$

Distribution under the independence assumption:

The distribution of **τ** is tabulated for small values of n, but it can be approximated by a Laplace Gauss with mean equal to 0 and variance equal to $\frac{2(2n+5)}{9n(n-1)}$. The approximation is valid as soon as n> = 10. In the calculation of the distribution of τ under the null assumption, we are interested in the quantity S which is approximated by a discrete and continuous variable, a correction continuity is done which consist on subtracting 1 from the observed value of S if it is positive, adding 1 if it is negative

Case of Existence of ties

In this case the coding for rank variables A and B :

$$A_{ij}=\begin{cases} 1 \text{ if } r_1(i) < r_1(j) \\ 0 \text{ if } r_1(i) = r_1(j) \\ -1 \text{ if } r_1(i) > r_1(j) \end{cases} \qquad B_{ij}=\begin{cases} 1 \text{ if } r_2(i) < r_2(j) \\ 1 \text{ if } r_2(i) = r_2(j) \\ -1 \text{ if } r_2(i) > r_2(j) \end{cases}$$

Beforehand, are assigned to pairs tied the arithmetic average of the ranks they would have if it does there was no tie.

Precisely, suppose there is n_1 (resp. n_2) Ties for groups r_1 (resp.r_2) and denote by u_i ($1 <= i <= n_1$) (resp.v_1, $1 <= j <= n_2$) the number of individuals tied for the i^{eme} (resp. j^{eme}) group.

Two expressions τ_1 and τ_2 are proposed for the coefficient of Kendall, the first 1 is written in the form:

$$\tau_2(X,Y) = \frac{P-Q}{\sqrt{((\frac{n(n-1)}{2}-V)(\frac{n(n-1)}{2}-U))}}$$

$$V = \frac{\sum_j v_j(v_j-1)}{2} \; ; U = \frac{\sum_i u_j(u_j-1)}{2}$$

Proposition 2. The other expression is written in the form:

$$\tau_2 = \frac{P-Q}{\frac{n(n-1)}{2}}$$

In the case of absence of ties: $\qquad \tau_1(X,Y) = \tau_2(X,Y)$

<u>Distribution under the null assumption</u>:

In the case of the presence of a tie (the most common case) and if we denote S = P - Q then The distribution of S under the assumption of independence of X and Y follows a normal distribution with mean 0 and variance:

$$\frac{1}{18}\left[n(n-1)(2n+1) - \sum_j\left(v_j(v_j-1)(2v_j+5)\right) - \sum_i u_i(u_i-1)(2u_i+5)\right]+$$

$$+\frac{1}{9n(n-1)(n-2)}[\sum_j\left(v_j(v_j-1)(v_j-2)\right)\sum_i u_i(u_i-1)(u_i-2)] + \frac{1}{2n(n-1)}[\sum_j\left(v_j(v_j-1)\right)\sum_i u_i(u_i-1)]$$

Reference

[1]S.CHAH, H.CHAMLAL, " Nouvelle approche pour la sélection de variable discriminantes ", Revue de la statistique Appliquée, tome 48, n°4 (2000), p.59-82.

[2]NAKACHE J.P, CELEUX G,'' Analyse discriminante sur variables qualitatives, (1994),Polytechnica, Chap.5.pp.134-135.

[3] Adrian Gepp,J.Holton Wilson,Kuldeep Kuman and Sukanto Bhatacharya,"A comparative Analysis of decision Tree Vis-à-vis Other computational Datamining Techniques in Atomotive Insurance Fraud Detection",journal of Data Science 10 (2012),537-561.

[4]Breiman L., Friedman J., Olshen R., and Stone C." Classification and Regression Trees", CA: Wadsworth and Brooks (1984).

[5]J.R.Quinlan "C5.0:An informal Tutorial,(1998), http://www.rulequest.com/ see5-unix.html.

[6]J.R.Quinlan "C4.5 :Programs for machine Learning"Morgan Kaufmann,San Mateo,California,(1993)

[7] J.R.Quinlan " Induction of Decision Trees ",Machine Learning 1 (1986),pp 81-106.

[8] G.V.Kass," An exploratory technique for investigating large quantities of categorical data ",Applied Statistics, vol.29,(1980),pp.119-127.

[9]W.-.Loh and Y-S. Shih,"Split selection methods for classification tree",statistica Sinica vol. 7,(1997),pp 815-840.

[10]Maruf Hossain, M.Rafiul Hassan and B.J.,"ROC-tree": A novel Decision Trees Induction Algorithm Based on Receive Operating Characteristics to classify Gene Expression data", In the 8th SIAM International Conference on Data Mining (SDM08),(2008) April.

[11] S. E. Hansen and R.Olsson, "Improving decision tree pruning through automatic programming "in

Proceedings of the Norwegian Conference on Informatics (NIK-2007), pp.31-40, (2007).

[12] D.Sieling,"Minimization of decision trees is hard to approximate",journal of Computer and System Sciences,vol. 74,(2008),pp. 394-403.

[13]G.Yoshikazu,M.Testuo,G.Yumiko,"Decision-tree model for predicting outcomes after out-of-hospital cardiac arrest in the emergency department",PMC journal , Crit Care. 2013; 17(4): R133. Jul 2013.

[14]K.Wisaeng," A Comparison of Different Classification Techniques for Bank Direct Marketing», International Journal Of Soft Computing and Engineering (IJSCE) ISSN:2231-2307,Volume-3,Issue-4,September 2013.

[15]Bhowmik R(2011),"Detection auto insurance Fraud by datamining technique.Journal of Emerging Trends in Computing and Information Sciences 2,156-162.

[16]Kazunor, H., U. motohide, S. Hiroshi and U. Yuushi, 1999. "Fuzzy C4.5 for generating fuzzy decision trees and its improvement". Faji Shisutemu Shinpojiumu Koen Ronbunshu, 15: 515-518.

[17] LC Thomas, Edelman DB, JN Crook "Credit Scoring And its Applications",SIAM Philadelphia

 [18]KENDALL M.G , "Rank correlation methods ", (1962), Griffin,Londres

[19] Jean-Pierre Nakache, Josiane Confais « Statistique explicative appliquée » Edition Technip 2003

 [20] Vincent Granville," Fast Combinatorial Feature Selection with New Definition of Predictive Power"

On the existence and uniqueness of solution of MRE and applications

Yunhui Hou[1], Nikolaos Limnios[2], and Walter Schön[3]

[1] Sorbonne Universités, Université de Technologie de Compiègne, CNRS, UMR
7253 Heudiasyc - CS 60 319 - 60 203 Compiègne cedex, France
(E-mail: `yunhui.hou@hds.utc.fr`)
[2] Sorbonne Universités, Université de Technologie de Compiègne, LMAC
Laboratory of Applied Mathematics of Compiègne - CS 60 319 - 60 203
Compiègne cedex, France
(E-mail: `nlimnios@utc.fr`)
[3] Sorbonne Universités, Université de Technologie de Compiègne, CNRS, UMR
7253 Heudiasyc - CS 60 319 - 60 203 Compiègne cedex, France (E-mail:
`wschon@utc.fr`)

Abstract. In this paper, we study the existence and uniqueness of the solution for
Markov renewal equation (MRE) of a semi-Markov process with countable state space.
This method and its proof are based on an iterative scheme. A numerical solution is
also given as well as a case study on system reliability assessment.
Keywords: semi-Markov process, Markov renewal equation, semi-Markov chain,
semi-Markov transition function keyword.

1 Introduction

Markov renewal or semi-Markov process gives a general and more flexible presentation for processes having Markov properties for modeling real life problems in finance (Janssen and Manca[5]), engineering (Grabski[4], Limnios[9],Gikhman and Skorokhod[2]), biology(Barbu and Limnios[1]), etc. However, an important drawback in these applications is the difficulty of obtaining tractable solution in application with a large enough number of states. Several methods were proposed in the past using algebraic and complementary variables approaches (Cox[3], Limnois[7–9]) to obtain transition functions of the semi-Markov process.

In this article, we study the existence and uniqueness solution of a Markov renewal equation(MRE). We also propose a method solving Markov renewal equations based on an iterative scheme inspired by n-fold convolution. Theoretical and numerical applications on system reliability and availability calculation as well as a case study on system availability and reliability assessment are also provided in the following sections.

2 Preliminaries and assumptions

In this section, we present the semi-Markov process framework which is used in the following sections.

2.1 Semi-Markov process framework

Consider a regular semi-Markov process Z_t, $t \geq 0$ on a countable set E state space (Limnios[9]). Let $(J_n, S_n), n \in \mathbb{N}$, be the (embedded) Markov renewal process (MRP) of Z_t where $S_0 \leq S_1 \leq ...$ are the time of jumps and J_n are the nth visited states. This process satisfies the following Markov property $\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_0, ..., J_n, S_0, ..., S_n) = \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n)$, almost surely (a.s.), for any $n \in \mathbb{N}$, $t \in \mathbb{R}_+$ and $j \in E$.

The semi-Markov kernel Q, and the initial probability α are defined as follows:

$$Q_{ij}(t) := \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = i)$$
$$\alpha(i) := \mathbb{P}(J_0 = i)$$

with $i, j \in E$.

Then the distribution function of the holding (sojourn) time in state $i \in E$, $H_i(t)$, is given by

$$H_i(t) := \sum_{j \in E} Q_{ij}(t)$$

Another important function is the semi-Markov transition function defined by

$$P_{ij}(t) := \mathbb{P}(Z_t = j \mid Z_0 = i), \quad i, j \in E, t \geq 0,$$

which is the conditional marginal law of the process.

Notice that the chain $(J_n, n \in \mathbb{N})$ also satisfies Markov property so that its transition probabilities can be defined as follow

$$P(i, j) := Q_{ij}(\infty) = \lim_{t \to \infty} Q_{ij}(t)$$

The counting process of jumps $N(t)$, $t \geq 0$, defined by $N(t) = \sup\{n \geq 0 : S_n \leq t\}$, gives the number of jumps of the Markov renewal process in the time interval $(0, t]$. The semi-Markov process Z_t can also be given by the relation

$$Z_t = J_{N(t)}, \quad t \geq 0. \tag{1}$$

2.2 Markov renewal function and n-fold convolution

Since Z_t is regular, which means that the number of jumps in any finite time interval is finite almost surely, we suppose that Z_t is continuous on the right having left limits in any point of time $t > 0$.

Let $\phi(i, t)$, $i \in E$, $t \geq 0$, be a real-valued measurable function, then the convolution of ϕ by Q is defined as follows

$$Q * \phi(i, t) := \sum_{k \in E} \int_0^t Q_{ik}(ds)\phi(k, t - s). \tag{2}$$

It is easy to prove the following fundamental equality

$$Q_{ij}^{(n)}(t) = P_i(J_n = j, S_n \leq t). \tag{3}$$

where $P_i(\cdot)$ means $\mathbb{P}(\cdot \mid J_0 = i)$ and $Q_{ij}^{(n)}(t)$, $i, j \in E$, is the n-fold convolution of Q by itself, i.e.

$$Q_{ij}^{(n)}(t) = \begin{cases} \sum_{k \in E} \int_0^t Q_{ik}(ds)Q_{kj}^{(n-1)}(t - s)ds & n \geq 2 \\ Q_{ij}(t) & n = 1 \\ \delta_{ij}\mathbf{1}_{\{t \geq 0\}} & n = 0, \end{cases}$$

with $\delta_{ij} = 1$, if $i = j$ and $\delta_{ij} = 0$, if $i \neq j$.

Then the Markov renewal function $\psi_{ij}(t)$, $i, j \in E, t \geq 0$, can be defined by

$$\psi_{ij}(t) := \sum_{n=0}^{\infty} Q_{ij}^{(n)}(t). \tag{4}$$

Let us write the Markov renewal function (4) in matrix form

$$\psi(t) = (I(t) - Q(t))^{(-1)} = \sum_{n=0}^{\infty} Q^{(n)}(t). \tag{5}$$

This can also be written as

$$\psi(t) = I(t) + Q * \psi(t), \tag{6}$$

where $I(t) = I$ (the identity matrix), if $t \geq 0$ and $I(t) = 0$, if $t < 0$.

2.3   Markov renewal equation

Equation (6) is a special case of what is called Markov renewal equation (MRE) which is generally defined as

$$U(t) = V(t) + Q * U(t), \tag{7}$$

where $U(t) = (U_{ij}(t))_{i,j \in E}$, $V(t) = (V_{ij}(t))_{i,j \in E}$ are matrix-valued measurable functions, with $U_{ij}(t) = V_{ij}(t) = 0$ for $t < 0$. The function $V(t)$ is a given matrix-valued function and $U(t)$ is an unknown matrix-valued function.

3 MRE solution given by iterative method

Assumption 1 *Let $\mathbb{B}$ be the space of all locally bounded, on $\mathbb{R}^+$, matrix functions $U(t)$, i.e., $\|U(t)\| = \sup_{i,j} |U_{i,j}(t)|$ is bounded on sets $[0, \xi]$, for every $\xi \in \mathbb{R}^+$.*

The following result is of our concern in the present paper.

Proposition 1 *(Pyke[10], Limnios[9]). The transition function*

$$P(t) = (P_{ij}(t), i, j \in E)$$

satisfies the following MRE

$$P(t) = \bar{H}(t) + Q * P(t), \tag{8}$$

which, under Assumption A, has the unique solution

$$P(t) = \psi * \bar{H}(t), \tag{9}$$

Here $\bar{H}(t) = I(t) - H(t)$ with $H(t) = diag(H_i(t))$ is a diagonal matrix.

Let $(P_{ij}^n(t), n \in \mathbb{N})$ a sequence of functions given by

$$P_{ij}^n(t) = \begin{cases} \delta_{ij}\bar{H}_i(t) & \text{if } n = 0 \\ \delta_{ij}\bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds)P_{kj}^{n-1}(t-s) & \text{if } n \geq 1. \end{cases} \tag{10}$$

where $\bar{H}_i(t) := 1 - H_i(t)$.

Proposition 2 *For any fixed $t \geq 0$ and $i, j \in E$, the limit*

$$\lim_{n \to \infty} P_{ij}^n(t)$$

exists and it is the smallest solution of the MRE (8) that we denote by $\tilde{P}_{ij}(t)$.

Proof. From equation (10) we get

$$\sum_{j \in E} P_{ij}^n(t) = \bar{H}_i(t) + \sum_{j \in E} \sum_{k \in E} \int_0^t Q_{ik}(ds)P_{kj}^{n-1}(t-s)$$

$$\leq \bar{H}_i(t) + \int_0^t H_i(ds)$$

$$= 1.$$

On the other hand, we have $P_{ij}^n(t) \geq 0$ for any $n \geq 0$, $i, j \in E$ and $t \geq 0$.

Define

$$D_{ij}^n(t) := \begin{cases} P_{ij}^0(t) & \text{if } n = 0 \\ P_{ij}^n(t) - P_{ij}^{n-1}(t) & \text{if } n \geq 1. \end{cases} \tag{11}$$

Then we get

$$D_{ij}^n(t) = \sum_{k \in E} \int_0^t Q_{ik}(ds) D_{kj}^{n-1}(t-s), \quad n \geq 1. \tag{12}$$

Since $D_{ij}^0(t) \geq 0$ we have $D_{ij}^n(t) \geq 0$ for all $n \geq 1$ and then $P_{ij}^n(t) \geq P_{ij}^{n-1}(t)$.

Finally, from the above inequality and the inequality $0 \leq P_{ij}^n(t) \leq 1$ we get the desired result that the limit $\tilde{P}_{ij}(t)$ exists.

Let us now prove that $\tilde{P}_{ij}(t)$ is the smallest solution of the MRE (8). The fact that $\tilde{P}_{ij}(t)$ is a solution of the MRE is obtained directly by considering limits in both sides of equation (10).

Now let us consider another solution of the MRE, say $P_{ij}^\sharp(t)$. Then we have $P_{ij}^\sharp(t) \geq \delta_{ij} \bar{H}_i(t) =: P_{ij}^0(t)$ and, suppose that $P_{ij}^\sharp(t) \geq P_{ij}^{n-1}(t)$, we get

$$P_{ij}^n(t) = \delta_{ij} \bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds) P_{kj}^{n-1}(t-s)$$

$$\leq \delta_{ij} \bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds) P_{kj}^\sharp(t-s)$$

$$= P_{ij}^\sharp(t),$$

and passing to limit when $n \to \infty$, the proof is achieved.

Moreover, the probabilistic meaning of $P_{ij}^n(t)$ can be given by

$$P_{ij}^n(t) = P_i(Z_t = j, S_{n+1} > t) \tag{13}$$

It indicates that for a fixed instant t, the convergence speed of the sequence $P_{ij}^n(t)$ depends on distribution of the jump number $N(t)$ or $H_i(t)$, the sejourn distribution of each states i

As the time t increases, the minimal n needed to obtain a precise value of $P_{ij}(t)$ using our method increases too.

Proof. Let us start with $n = 0$

$$P_{ij}^0(t) = \delta_{ij} \bar{H}_i(t) = P_i(Z_t = j, S_1 > t)$$

If for $n \in \mathbb{N}$ the proposition stands, then at $n+1$, we have

$$P_{ij}^{n+1}(t) = \delta_{ij} \bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds) P_{kj}^n(t-s), \qquad n \geq 1$$

$$= P_i(Z_t = j, S_1 > t) + \sum_{k \in E} \int_0^t \mathbb{P}(J_1 = k, S_1 \in ds | J_0 = i) \cdot$$

$$\mathbb{P}(Z_{t-s} = j, S_{n+2} - S_1 > t - s | J_1 = k, S_1 = s)$$

$$= P_i(Z_t = j, S_1 > t) + \sum_{k \in E} \int_0^t P_i(Z_t = j, S_1 \in ds, S_{n+2} > t, J_1 = k)$$

$$= P_i(Z_t = j, S_1 > t) + P_i(Z_t = j, S_1 \leq t < S_{n+2})$$

$$= P_i(Z_t = j, S_{n+2} > t)$$

The hypothesis also stands. Finally, we prove equality (13) for all $n \in \mathbb{N}$, and we also have

$$D_{ij}^n(t) = P_i(Z_t = j, N(t) = n)$$

4 Application on system reliability assessment

Besides transition functions, Proposition 2 is also valuable for other probability functions under the same framework. In this section, the method proposed here is applied on system survival or reliability functions and availability assessment.

Let us consider the survival or reliability function in the case where a subset of down states is given, i.e., say $D \subset E$, and the lifetime of the system T is defined by $T := \inf\{t \geq 0 : Z_t \in D\}$. The survival function is then $R_i(t) := P_i(T > t)$, $i \in E \setminus D$. This function satisfy the following MRE (see, e.g., Limnios[9])

$$R_i(t) = \bar{H}_i(t) + \sum_{k \in E \setminus D} \int_0^t Q_{ik}(ds) R_k(t - s).$$

In that case we can use also the above iterative scheme, i.e.,

$$R_i^n(t) = \begin{cases} \bar{H}_i(t) & \text{if } n = 0 \\ \bar{H}_i(t) + \sum_{k \in E \setminus D} \int_0^t Q_{ik}(ds) R_k^{n-1}(t - s) & \text{if } n \geq 1. \end{cases}$$

Similarly, the availability is defined by $A_i(t) := P_i(Z_t \in E \setminus D)$ and satisfy the following MRE

$$A_i(t) = \mathbf{1}_{\{E \setminus D\}}(i) \bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds) A_k(t - s).$$

with the corresponding iterative scheme

$$A_i^n(t) = \begin{cases} \mathbf{1}_{\{E \setminus D\}}(i) \bar{H}_i(t) & \text{if } n = 0 \\ \mathbf{1}_{\{E \setminus D\}}(i) \bar{H}_i(t) + \sum_{k \in E} \int_0^t Q_{ik}(ds) A_k^{n-1}(t - s) & \text{if } n \geq 1. \end{cases}$$

where $\mathbf{1}_{\{E \setminus D\}}(i)$ is the indicator function gives 1 if $i \in E \setminus D$, 0 otherwise.

In both cases, the limits $\tilde{R}_i(t) := \lim_{n \to \infty} R_i^n(t)$ and $\tilde{A}_i(t) := \lim_{n \to \infty} A_i^n(t)$ give the smallest solutions of the above corresponding MREs. It can be also applied directly on Markov renewal equation of semi-Markov chains (Barbu and Limnios[1]).

5 Numerical application: A case study

In this section, we apply our method discussed in the previous section on a specific case study: a multi-task machine.

5.1 System description

Consider a multi-task machine which performs three different tasks T1, T2 and T3 whose arrival times follow exponential distributions with parameter $\lambda = 2$, density function $f_{exp}(x; \lambda)$ and cumulative distribution function $F_{exp}(x; \lambda)$. The arrival proportion is 0.45, 0.35 and 0.20 for Task T1, T2 and T3 respectively. The time to complete of each task follows lognormal distribution with different parameters:

- Task T1: $\mu_1 = 1.3550$, $\sigma_1 = 0.25$;
- Task T2: $\mu_2 = 1.4728$, $\sigma_2 = 0.25$;
- Task T3: $\mu_3 = 1.5782$, $\sigma_3 = 0.25$.

whose distribution function $F_{logn}(x; \mu_i, \sigma_i)$ and density function $f_{logn}(x; \mu_i, \sigma_i)$. We assume that at the end of each operation, the machine is maintained and failures happen only during the operations with a failure time following Weibull distribution with scale parameter $a = 56.4190$ and shape parameter $b = 2$: for $x \geq 0$ the cumulative distribution function and the density function are respectively given by $F_{wbl}(x; a, b)$ and $f_{wbl}(x; a, b)$. Once the machine fails, it takes 20 hours to repair it which indicates the repair time follows a Dirac distribution with parameter $c = 20$ whose cumulative distribution function is a unit step function.

5.2 Semi-Markov modeling

According to the description, the system can be presented by a five-state semi-Markov process (Fig. 1) where each state corresponds a specific system state (Table 1) and the Markov kernel $(Q_{ij}(t))$ are shown in Table 2 using method mentioned in Korolyuk and Tomusyak[6]. (The conditional reliability and availability are shown in Fig. 5 and Figure 4).

State	Description
1	Machine fails
2	Machine waits for new task
3	Operation task T1
4	Operation task T2
5	Operation task T3

Table 1. States and their description

Using our proposed method, the obtained conditional reliability $R_i^n(t_c)$ and availability $A_i^n(t_c)$ at instant $t_c = 100$ on function of iteration number n are shown in Fig. 3 and Fig. 2.

Except the constant sojourn time of state 1, we observe that the obtained system reliability and availability estimates have the same order as their expected sojourn times. The convergence rates of all curves are almost the same as they share the same jumping count process $N(t)$.

6   Conclusion

In this article, we studied the existent and unique solution for Markov renewal equation and proposed an iterative method for solving MRE. We proved the existence and uniqueness of the obtained solution. The advantage of this method is that there is little limitation on state space which can be extended to infinite and its simplicity. The convergence speed of our method depends on the jump count process which is easier to be estimated. We also applied it on calculating system reliability and availability. The future work will focus on simplifying for long term study and application on large scale state space.

References

1. V. S.Barbu and N. Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications*. Vol. 191 of Lecture Notes in Statistics. Springer New York, New York, NY, 2008.

2. I.I. Gikhman and A.V. Skorokhod. *Theory of Stochastic Processes*, Vol 2, Springer-Verlag, Berlin,1974.

3. D. R.Cox, . The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Mathematical Proceedings of the Cambridge Philosophical Society* 51 (03), 433, 2008.

4. F. Grabski. Applications of semi-Markov processes in reliability. *Reliability Theory and Applications*, 3-4 (December - Special Issue), 60–75, 2007.

5. J.Janssen and R.Manca. *Semi-Markov Risk Models for Finance, Insurance and Reliability*, Springer US, Boston, MA, 2007.

6. V. S. Korolyuk, and A. A. Tomusyak. Description of the functioning of systems with reserves by means of semi-Markov processes, *Cybernetics* 1 (5), 58–63, 1966.

7. N. Limnios. Reliability Measures of Semi-Markov Systems with General State Space, *Methodology and Computing in Applied Probability*, 14 (4), 895–917, Jan. 2011.

8. N. Limnios. Interval Reliability, Corrections and Developments of "Reliability Measures of Semi-Markov Systems with General State Space", *Methodology and Computing in Applied Probability* 16, 765–770, 2014.

9. N. Limnios and G. Oprian. *Semi-Markov Processes and Reliability*, Birkhäuser Boston, Boston, MA, 2001.

10. R. Pyke. Markov Renewal Processes: Definitions and Preliminary Properties, *The Annals of Mathematical Statistics* 32 (4), pp. 1231–1242, 1961.
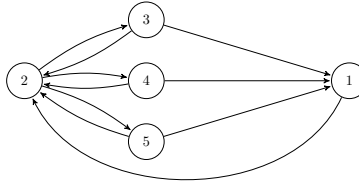
Fig. 1. The semi-Markov process of the case study

$Q_{12}(t)$	$H(t;20)$
$Q_{23}(t)$	$0.45F_{exp}(t;\lambda)$
$Q_{24}(t)$	$0.35F_{exp}(t;\lambda)$
$Q_{25}(t)$	$0.20F_{exp}(t;\lambda)$
$Q_{31}(t)$	$\int_0^t (1-F_{logn}(s;\mu_1,\sigma_1))f_{wbl}(s;a,b)ds$
$Q_{32}(t)$	$\int_0^t f_{logn}(s;\mu_1,\sigma_1)(1-F_{wbl}(s;a,b))ds$
$Q_{41}(t)$	$\int_0^t (1-F_{logn}(s;\mu_2,\sigma_2))f_{wbl}(s;a,b)ds$
$Q_{42}(t)$	$\int_0^t f_{logn}(s;\mu_2,\sigma_2)(1-F_{wbl}(s;a,b))ds$
$Q_{51}(t)$	$\int_0^t (1-F_{logn}(s;\mu_3,\sigma_3))f_{wbl}(s;a,b)ds$
$Q_{52}(t)$	$\int_0^t f_{logn}(s;\mu_3,\sigma_3)(1-F_{wbl}(s;a,b))ds$

Table 2. $Q_{ij}(t)$ for passage between states(for those mot mentioned $Q_{ij}(t)=0$)
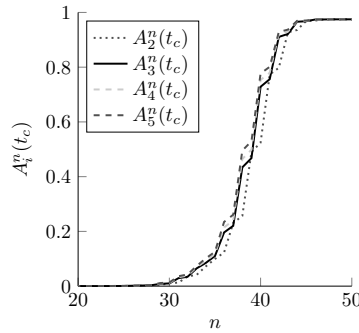

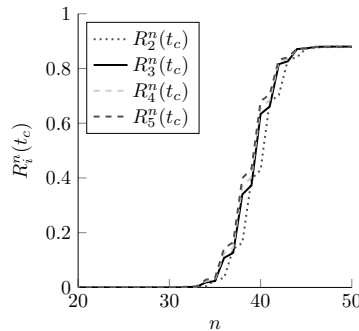**Fig. 2.** Plot of $A_i^n(t_c)$, $t_c = 100$ on function of n

Fig. 3. Plot of $R_i^n(t_c)$, $t_c = 100$ on function of n
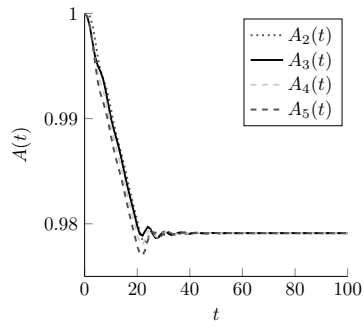
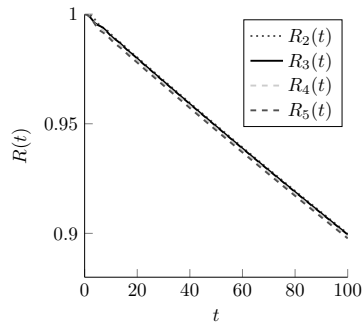**Fig. 4.** Plot of $A_i(t)$, $t = 0 : 100$ on function of t

Fig. 5. Plot of $R_i(t)$, $t = 0 : 100$ on function of t

Joint mixture modeling of survival and longitudinal data with multiple features

Yangxin Huang[1,2] and Huahai Qiu[1]

[1] Department of Mathematics, School of Mathematics and Computer, Wuhan Textile University, Hubei 430073, P.R. China
   (E-mail: `yhuang@health.usf.edu`; `qiuhuahai2006@163.com`)
[2] Department of Epidemiology and Biostatistics, University of South Florida, Tampa, FL 33612, U.S.A.

Abstract. It often happens in longitudinal studies that repeated measurements of markers are observed with various data features of a heterogeneous population comprising of several subclasses, left-censoring due to a limit of detection (LOD) and covariates measured with error. Moreover, repeatedly measured markers in time may be associated with a time-to-event of interest. Inferential procedures may become very complicated when one analyzes data with these features together. This article explores a finite mixture of hierarchical joint models of event times and longitudinal measures with an attempt to alleviate departures from homogeneous characteristics, tailor observations below LOD as missing values, mediate accuracy from measurement error in covariate and overcome shortages of confidence in specifying a parametric time-to-event model with a nonparametric distribution. The Bayesian joint modeling is employed to not only estimate all parameters in mixture of joint models, but also evaluate probabilities of class membership. A real data example is analyzed to demonstrate the methodology by jointly modeling the viral dynamics and the time to decrease in CD4/CD8 ratio in the presence of CD4 cell counts with measurement error and the analytic results are reported by comparing potential models for various scenarios.
Keywords: Bayesian joint modeling ; Dirichlet process; Limit of detection; Longitudinal data analysis; Mixture of joint models; Time-to-event.

1 Introduction

Many studies aim at exploring the relationship between a time-to-event outcome and a longitudinal marker. For example, the relationship of HIV viral suppression and immune restoration after a treatment has received great attention in HIV/AIDS research [1]. One is often interested in simultaneously studying the HIV dynamics and immune restoration, which may be characterized by the time to CD4/CD8 ratio decline [2]. Relatively little work has been published about statistical analysis for the particular association of HIV viral dynamics and time trend in CD4/CD8 ratio in the presence of CD4 covariate process. The research was motivated by an AIDS Clinical Trials Group study 388 (ACTG388) [1] to understand within-subject patterns of change in HIV-1 RNA copies (also referred to as viral load) or CD4 cell count, and to study the relationship of features of viral load and CD4 profiles with time to decrease in CD4/CD8 ratio.

Joint analysis of event times and longitudinal measures is an active area of biostatistics and statistics research. There have been a considerable number of statistical approaches in the literature. Researchers may often confront the task of developing inference where longitudinal outcomes of interest may follow heterogeneous (not homogeneous) characteristics, suffer from left-censoring due to a limit of detection (LOD), and measure with substantial error. Modeling such data has many challenges due to the following issues of inherent data features. First, in the literature, most studies of longitudinal modeling assume that all subjects come from a homogeneous population where large between- and within-individual variations were accommodated by random-effects and/or time-varying covariates in the models. These typical between- and within-individual variations are shown in Figure 1(a), viral load trajectory profiles of three representative patients in ACTG388 (see Section 2 for details of this study and data description); these viral load trajectories can be roughly classified into three classes, which are biologically and clinically interpretable. We, therefore, can reasonably assume that these patients are from a population which consists of three relatively distinct classes and, thus, we consider a finite mixture of nonlinear mixed-effects (NLME) models for such data set. Second, the outcome of a longitudinal study may be subject to LOD because of the low sensitivity of current standard assays [3]. For example, for the ACTG388 study, which was designed to collect data on every individual at each assessment, the response (viral load) measurements may be subject to left-censoring due to an LOD. It can be seen from Figure 1(a) that for some patients their viral loads

are below LOD. The proportion of data censored may not be trivial, so failure to account for censoring in the analysis may result in significant biases in the estimates of the fixed-effects and variance components [4,5]. Third, measurement error in covariates is another typical feature of longitudinal data and ignoring this phenomenon may result in unreliable statistical inference. This is the usual case in the longitudinal studies; for instance, CD4 cell counts are often measured with substantial measurement error [6–8]. Finally, it is a commonplace in clinical and public health research that longitudinal studies consist of repeated measurements on continuous variables, an observation on a possibly censored time-to-event and additional covariate information collected for each subject. Interest often focuses on interrelationships among these variables [9]. A joint model that links the event time to these longitudinal measures, which can also incorporate information about left-censoring due to LOD and measurement error in covariates, is becoming increasingly important in many applications.
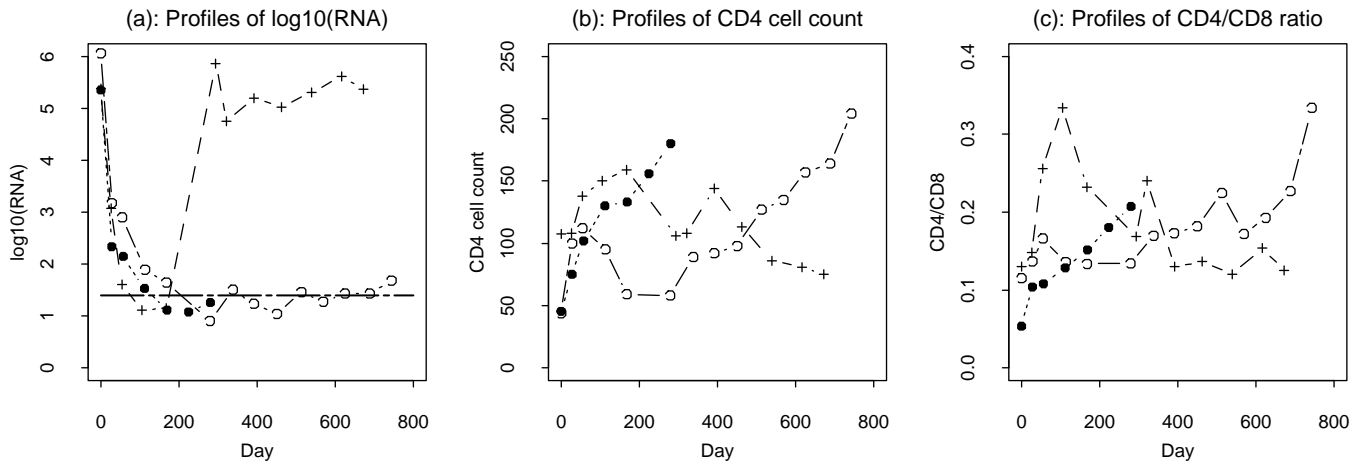

Fig. 1. Profile of viral load in $\log_{10}$ scale, CD4 cell count and CD4/CD8 ratio trajectories for three representative patients. Trajectory class 1: decrease rapidly and constantly in a short-term period (dotted line with solid dot sign); class 2: decrease at the beginning and then maintain stable at a low level (solid line with circle sign) and class 3: decrease at the beginning, but rebound later (dashed line with plus sign).

Although joint modeling of both longitudinal and event time data has received a great deal of attention in the statistical literature in recent years, the majority of the statistical literature for the joint modeling of longitudinal and/or event time data has focused on the development of models that aim at capturing only specific aspects of the motivating studies. These specific data features considered in the literature include longitudinal data with, but are not limited to, censored response [4,5,8,10], mixture response [11–14], covariate measurement error [6–8,15], and longitudinal-survival (or event time) data [9,16–19]. However, there is relatively few studies on simultaneous inference for longitudinal data with features of heterogeneity and left-censoring due to LOD in response and measurement error in covariate as well as event time data incorporated. It is not clear how heterogeneity, left-censoring, covariate measurement error and event time of data may interact and simultaneously influence inferential procedures. Statistical inference and analysis complicate dramatically when all of these issues are present. In this article, we develop a mixture of mixed-effects joint model to investigate the effects on inference when all of these typical data features exist.

We employ a Bayesian framework for a mixture of nonlinear mixed-effects joint (MNLMEJ) models for longitudinal response and time-to-event in the presence of covariates with measurement error to accommodate a large class of data structures with various features. In particular, we consider the MNLMEJ model with the following components to quantify these data features: (i) a mixture of NLME models for viral load response process in connection with a Tobit model [20] to tailor observations below LOD; (ii) nonparametric mixed-effects model for CD4 covariate process to modulate accuracy from measurement error; and (iii) an accelerated failure time (AFT) model for time to decline of CD4/CD8 ratio using the nonparametric Dirichlet process (DP) prior as a distribution [21,22], which is linked to longitudinal models through random-effects, providing further bridge between time-to-event and longitudinal observations. To the best of our knowledge, there has been relatively

little work done to address the simultaneous impact of longitudinal data with heterogeneity, left-censoring due to LOD, measurement error in covariate and event time by joint modeling the three processes above based on Bayesian MNLMEJ model. It is noted that some models published in the literature [6,12,19] can be treated as special cases of MNLMEJ model proposed in this article when our model is tailored in this way. Approaching the mixture joint modeling problem from the Bayesian perspective is more natural and straightforward. It avoids many complicated approximations required by the frequentist approach. Yet, with noninformative priors, we can still get maximum likelihood-type estimates.

The remainder of this article evolves as follows. Section 2 describes the data set that motivated this research and discusses the specific MNLMEJ models, where mixture models are formulated by three nonlinear mean functions of different mixture components for viral load response with left-censoring, CD4 covariate process with measurement errors and AFT model for time to decline of CD4/CD8 ratio. Simultaneous Bayesian inferential procedure for estimating both parameters and probabilities of class membership are presented in Section 3. In Section 4, we demonstrate to apply the proposed methodologies to an AIDS data set described in Section 2 and report the analytic results. Finally, we conclude the article with discussion in Section 5.

2   Motivating data set and a mixture of joint models for HIV dynamics

2.1   Motivating data set

The data set that motivated this research is from ACTG388 study [1]. ACTG388 study was a randomized, open-label study comparing 2 different 4-drug regimens with a standard 3-drug regimen for 517 subjects with no or limited previous experience with antiretroviral treatment (ART) who had a CD4 cell count $\leq$200 cells/mm^3 or a plasma HIV-1 RNA level $\geq$ 80,000 copies/mL at screening. The planned study duration was 72 weeks, which was subsequently increased to 96 weeks beyond the enrollment of the last subject. The plasma HIV-1 RNA (viral load) is repeatedly quantified at weeks 0, 4, 8, 16, and every 8 weeks until the last patient on study. The number of viral load measurements for each individual varies from 2 to 18. Out of total 517 patients, sixty-five subjects had no viral load measurements because all RNA assay results are not available for these subjects and twenty-five subjects had less than three viral load measurements (which is not suitable for classification) because of permanently discontinued study prior to study completion with various reasons, which were excluded from this study. The remaining 427 subjects whose viral load measurements vary from 3 to 18 were included in data analysis. Thirty four percent of viral load observations were measured below the LOD of 25 copies/mL in our data. The CD4 and CD8 cell counts were also measured throughout study on a similar scheme. The detailed data description can be found in Fischl et al.[1]. The $\log_{10}$ transformation of viral load and square-root of CD4 were used in the analysis in order to make data more symmetric, to stabilize the variation of the measurement errors and to speed up estimation algorithm. In addition, in order to avoid too small or large estimates which may be unstable, we rescaled the original time (in days) so that the time scale is between 0 and 1.

Since viral load is an important marker for assessing virologic response of an ART regimen, our objective is to develop a mixture of joint modeling for characterizing population-level and individual-level viral load trajectories. As discussed in Section 1, viral load trajectory profiles of three representative patients displayed in Figure 1(a) can be roughly classified into three classes in terms of the effect of treatment on viral load responses: For class 1, patient's viral loads decrease rapidly and constantly in a short-term period (dotted line with solid circle sign) due to the fact that some patients withdrew too early to be clustered into either class 2 (with viral load suppression or without viral load rebound) or class 3 (with viral load rebound); for class 2, patient's viral loads decrease at the beginning and then stay stable at a low level (solid line with circle sign). The classes 1 and 2 with suppression of plasma HIV-1 RNA levels indicate that the treatment can be thought successful without serious clinical problems arose, suggesting a confirmed virologic response. While for class 3 patient's viral loads decrease at the beginning, but they experience viral load increase, which results in viral load rebound eventually (dashed line with plus sign), implying a virologic failure. Along with these observations, we can reasonably assume that patients from a population consist of three relatively homogeneous classes.

Figure 1 presents the trajectories of HIV viral load, CD4 cell count and CD4/CD8 ratio of three representative subjects from ACTG388 study [1]. While the data indicate a close association of the longitudinal viral load, CD4 cell count and CD4/CD8 ratio, the data show a large variation in the association across subjects and over time within each subject. This observation together with the published researches in the area [7,12,16,19,23] led us to

propose the following MNLMEJ models for HIV dynamics in connection with three outcome variables of concern. Due to the fact that our objective in this study is HIV viral dynamics, we choose viral load as a response variable and consider CD4 cell count as a covariate [5,7,8,16] to develop MNLMEJ models below.

2.2 Measurement error models

Denote the number of subjects by n and the number of measurements on the ith subject by n_i. Let $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{in_i})^T$, where z_{ij} is time-varying CD4 covariate for the ith subject at time t_{ij} ($i = 1, 2, \ldots, n; j = 1, 2, \ldots, n_i$). As is evident from Figure 1(a), the inter-patient variations in viral load appear to be large and these variations change over time. Previous studies suggest that the inter-patient variation in viral load may be partially quantified by time-varying CD4 cell count [6–8]. The covariate (CD4 cell counts) often have substantial measurement errors, and ignoring these errors may lead to misleading results [24]. With CD4 measures collected over time, we may model the CD4 process to partially address the measurement errors [8]. However, the CD4 trajectories are often complicated, and there is no well established model for the CD4 process. We, thus, model the CD4 process empirically using a nonparametric mixed-effects model, which is flexible and works well for complex longitudinal CD4 data as follows.

$$z_{ij} = w(t_{ij}) + h_i(t_{ij}) + \epsilon_{ij} \ (\equiv z_{ij}^* + \epsilon_{ij}) \ \ \boldsymbol{\epsilon}_i \overset{\text{iid}}{\sim} N_{n_i}(\boldsymbol{0}, \sigma_1^2 \boldsymbol{I}_{n_i}), \tag{1}$$

where $w(t_{ij})$ and $h_i(t_{ij})$ are unknown nonparametric smooth fixed-effects and random-effects functions, respectively, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^T$ follows a multivariate normal distribution. $z_{ij}^* = w(t_{ij}) + h_i(t_{ij})$ are the true (but unobservable) covariate values at time t_{ij}. The fixed smooth function $w(t)$ represents population average of the covariate process, while the random smooth function $h_i(t)$ is introduced to incorporate the large inter-individual variation in the covariate process. We assume that $h_i(t)$ is the realization of a zero-mean stochastic process.

Model (1) is more flexible than parametric mixed-effects models. To fit model (1), we apply a regression spline method to $w(t)$ and $h_i(t)$. The working principle is briefly described as follows and more details can be found in literature [25,26]. The main idea of regression spline is to approximate $w(t)$ and $h_i(t)$ by using a linear combination of spline basis functions. For instance, $w(t)$ and $h_i(t)$ can be approximated by a linear combination of basis functions $\boldsymbol{\Psi}_p(t) = \{\psi_0(t), \psi_1(t), \ldots, \psi_{p-1}(t)\}^T$ and $\boldsymbol{\Phi}_q(t) = \{\phi_0(t), \phi_1(t), \ldots, \phi_{q-1}(t)\}^T$, respectively. That is,

$$w(t) \approx w_p(t) = \sum_{l=0}^{p-1} \alpha_l \psi_l(t) = \boldsymbol{\Psi}_p(t)^T \boldsymbol{\alpha}, \quad h_i(t) \approx h_{iq}(t) = \sum_{l=0}^{q-1} a_{il} \phi_l(t) = \boldsymbol{\Phi}_q(t)^T \boldsymbol{a}_i, \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{p-1})^T$ is a $p \times 1$ vector of fixed-effects and $\boldsymbol{a}_i = (a_{i0}, \ldots, a_{i,q-1})^T$ ($q \leq p$) is a $q \times 1$ vector of random-effects with $\boldsymbol{a}_i \overset{\text{iid}}{\sim} N_q(\boldsymbol{0}, \boldsymbol{\Sigma}_a)$. Based on the assumption of $h_i(t)$, we can regard $\boldsymbol{a}_i$ as iid realizations of a zero-mean random vector. Substituting $w(t)$ and $h_i(t)$ by their approximations $w_p(t)$ and $h_{iq}(t)$, we can approximate model (1) by the following linear mixed-effects (LME) model.

$$z_{ij} \approx \boldsymbol{\Psi}_p(t_{ij})^T \boldsymbol{\alpha} + \boldsymbol{\Phi}_q(t_{ij})^T \boldsymbol{a}_i + \epsilon_{ij} \approx z_{ij}^* + \epsilon_{ij}, \quad \boldsymbol{\epsilon}_i \overset{\text{iid}}{\sim} N_{n_i}(\boldsymbol{0}, \sigma_1^2 \boldsymbol{I}_{n_i}), \tag{3}$$

For this model, we consider natural cubic spline bases with the percentile-based knots. To select an optimal degree of regression spline and numbers of knots, i.e., optimal sizes of p and q, the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) is often applied [25]. Following the study in Liu and Wu [7], we set $\psi_0(t) = \phi_0(t) = 1$ and take the same natural cubic splines in the approximation (2) with $q \leq p$ (in order to limit the dimension of random-effects). The numbers of knots (see equation (2) in detail), p and q, are determined based on the AIC/BIC criteria. For the ACTG388 data, the AIC/BIC values are evaluated for various models with $(p, q) = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 2), (3, 3)\}$ which was found that the model with $(p, q) = (3, 3)$ has the smallest AIC/BIC values being 5733.5/5749.8. We, thus, adopted the following CD4 covariate model.

$$z_{ij} \approx (\alpha_0 + a_{i0})\psi_0(t_{ij}) + (\alpha_1 + a_{i1})\psi_1(t_{ij}) + (\alpha_2 + a_{i2})\psi_2(t_{ij}) + \epsilon_{ij}(\equiv z_{ij}^* + \epsilon_{ij}) \tag{4}$$

where z_{ij} is the observed $\sqrt{\text{CD4}}$ value at time t_{ij}, $\psi_1(\cdot)$ and $\psi_2(\cdot)$ are two basis functions given in Section 2.1, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)^T$ is a vector of population (fixed-effects) parameters, $\boldsymbol{a}_i = (a_{i0}, a_{i1}, a_{i2})^T$ is a vector of random-effects.

2.3 A mixture of NLME Tobit models for HIV dynamics

Let y_{ij} be the value of viral load response for the individual i at time t_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, n_i)$. In order to introduce Tobit model to deal with observations below LOD in our mixture joint modeling framework, denote the observed value y_{ij} by (q_{ij}, d_{ij}), where d_{ij} is the censoring indicator and q_{ij} is the latent response variable. The latent q_{ij} is observed, as y_{ij}, if and only if $y_{ij} > \rho$ (a known constant LOD). When q_{ij} is observed we have $d_{ij} = 0$. Otherwise we have $d_{ij} = 1$. Let $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^T$, $\boldsymbol{q}_i = (q_{i1}, \ldots, q_{in_i})^T$ and $\boldsymbol{d}_i = (d_{i1}, \ldots, d_{in_i})^T$. Finite mixture models for longitudinal studies were considered in the literature [13,14], where the latent classes corresponding to the mixture components and cluster individuals may provide a better inference. However, most finite mixture models for longitudinal data are currently based on linear (polynomial) [13] or piecewise linear [14] mean functions. The partial reason is that the computation can be conveniently carried out because the likelihood function of a model based on these 'linear' mean functions has a closed form [13]. However, in practice, most longitudinal trajectories appear to be nonlinear patterns. When a mixture model is extended to incorporate nonlinear mean functions which will be conducted in this article, inferential procedures complicate dramatically because a closed form of likelihood function no longer exists.

We assume that there are K plausible nonlinear trajectory classes with mean functions $g_k(\cdot)$ $(k = 1, \ldots, K)$, which are known to be specified (in our case conducted K=3). The true trajectory mean function of the ith subject might be $g_k(\cdot)$ with unknown probability $\pi_k = P(c_i = k)$ which satisfies $\sum_{k=1}^K \pi_k = 1$, where c_i is a latent indicator. For the response process with left-censoring, an NLME model for ith subject, given $c_i = k$, can be formulated by

$$(\boldsymbol{y}_i | c_i = k) = \boldsymbol{g}_k(\boldsymbol{t}_i, \boldsymbol{A}_k \boldsymbol{\beta}_{ij}) + \boldsymbol{e}_i, \quad \boldsymbol{e}_i \overset{\text{iid}}{\sim} N_{n_i}(\boldsymbol{0}, \sigma_2^2 \boldsymbol{I}_{n_i}), \tag{5}$$

where $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})^T$, the vector of random errors $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})^T$ follows a multivariate normal distribution with mean zero and unknown variance parameter σ_2^2, and $\boldsymbol{\beta}_{ij} = (\beta_{1ij}, \ldots, \beta_{sij})^T$ is a vector of individual parameters for the ith subject, which is expressed by

$$\boldsymbol{\beta}_{ij} = \boldsymbol{Z}_{ij} \boldsymbol{\beta} + \boldsymbol{X} \boldsymbol{b}_i, \tag{6}$$

where $\boldsymbol{Z}_{ij}$ $(s \times r)$ is a design matrix including time-independent and/or time-varying covariates such as CD4 cell count, $\boldsymbol{X}$ $(s \times m)$ is random-effects related indicator matrix; $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)^T$ is a vector of population parameters, and the vector of random-effects $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{im})^T \overset{\text{iid}}{\sim} N_m(\boldsymbol{0}, \boldsymbol{\Sigma_b})$ with an unknown variance-covariance matrix $\boldsymbol{\Sigma_b}$ $(m \times m)$. We assume one of the elements in $\boldsymbol{Z}_{ij}$, z_{ij}^* in model (3), is a summary of true (but unobservable) time-varying covariate value at time t_{ij}. It is noted that $r \geq s \geq m$. $\boldsymbol{g}_k(\boldsymbol{t}_i, \boldsymbol{A}_k \boldsymbol{\beta}_{ij}) = (g_k(t_{i1}, \boldsymbol{A}_k \boldsymbol{\beta}_{i1}), \ldots, g_k(t_{in_i}, \boldsymbol{A}_k \boldsymbol{\beta}_{in_i}))^T$, and $\boldsymbol{A}_k$ $(k = 1, \ldots, K)$ is known $s \times s$ square-matrix indicator, of which diagonal elements are either 0 or 1 and off-diagonal elements are all 0. $\boldsymbol{A}_k$ is introduced because the mean functions of y_{ij} $(i = 1, \ldots n; j = 1, \ldots, n_i)$, specified by the nonlinear functions $g_1(\cdot), \ldots, g_K(\cdot)$, may only involve different subsets of $\boldsymbol{\beta}_{ij}$. By introducing $\boldsymbol{A}_k$, $\boldsymbol{A}_k \boldsymbol{\beta}_{ij}$ will set unrelated elements of $\boldsymbol{\beta}_{ij}$ to 0 in the kth trajectory class, respectively. We will illustrate the choice of $\boldsymbol{A}_k$ and specify nonlinear mean functions $g_k(\cdot)$ in the HIV dynamic application below.

Similar to discussion in Pauler and Laird [14] and Lu and Huang [12], model (5) can be specified conditionally and marginally, respectively, by

$$(\boldsymbol{y}_i | c_i = k) \sim N_{n_i}(\boldsymbol{g}_k(\boldsymbol{t}_i, \boldsymbol{A}_k \boldsymbol{\beta}_{ij}), \sigma_2^2 \boldsymbol{I}_{n_i}), \tag{7}$$

$$\boldsymbol{y}_i \sim \sum_{k=1}^K \pi_k N_{n_i}(\boldsymbol{g}_k(\boldsymbol{t}_i, \boldsymbol{A}_k \boldsymbol{\beta}_{ij}), \sigma_2^2 \boldsymbol{I}_{n_i}). \tag{8}$$

Thus, equation (8) along with the Tobit formulation forms the finite mixture of NLME Tobit models for response variable with missing values due to LOD in the presence of measurement error model (3). In (8) the vector of mixture probabilities $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)^T$ can be also viewed as the mixture weights of all plausible components within the finite mixture model framework. Model (8) is identifiable, as long as each of the component models is identifiable and distinguishable from each other; when the component models are identifiable but not distinguishable from each other, some constraints may be required to make model (8) identifiable[27].

For the viral load response in HIV dynamics, we consider one- and two-compartment models (see Appendix A in detail) with constant decay rate(s) for trajectory classes 1 and 2, respectively, and a two-compartment model with a time-varying decay rate in the second compartment for trajectory class 3 described above. Toward this end, the mean functions of $K = 3$ components in the mixture model are specified by

1. One-compartment model with a constant decay rate for class 1 trajectory

$$g_1(t_{ij}, \boldsymbol{A}_1\boldsymbol{\beta}_{ij}) = \log_{10}(e^{p_{1i}-\lambda_{1i}t_{ij}}), \tag{9}$$

2. Two-compartment model with constant decay rates for class 2 trajectory

$$g_2(t_{ij}, \boldsymbol{A}_2\boldsymbol{\beta}_{ij}) = \log_{10}(e^{p_{1i}-\lambda_{1i}t_{ij}} + e^{p_{2i}-\lambda_{2i}t_{ij}}), \tag{10}$$

3. Two-compartment model with constant and time-varying decay rates for class 3 trajectory

$$g_3(t_{ij}, \boldsymbol{A}_3\boldsymbol{\beta}_{ij}) = \log_{10}(e^{p_{1i}-\lambda_{1i}t_{ij}} + e^{p_{2i}-\lambda_{2ij}t_{ij}}). \tag{11}$$

In (9)–(11),

$$\beta_{1i} = p_{1i} = \beta_1 + b_{i1}, \ \beta_{2i} = \lambda_{1i} = \beta_2 + \beta_3 z_{i0} + b_{i2}, \ \beta_{3i} = p_{2i} = \beta_4 + b_{i3},$$
$$\beta_{4i} = \lambda_{2i} = \beta_5 + b_{i4}, \ \beta_{5ij} = \lambda_{2ij} = \beta_5 + \beta_6 E(z_{ij}) + b_{i4},$$
$$\boldsymbol{\beta}_{ij} = (\beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_{4i}, \beta_{5ij})^T, \ \boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^T, \boldsymbol{b}_i = (b_{i1}, \dots, b_{i4})^T,$$
$$\boldsymbol{A}_1 = \mathrm{diag}(1,1,0,0,0), \boldsymbol{A}_2 = \mathrm{diag}(1,1,1,1,0), \boldsymbol{A}_3 = \mathrm{diag}(1,1,1,0,1),$$

$$\boldsymbol{Z}_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & z_{i0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & z_{ij}^* \end{pmatrix} \text{ and } \boldsymbol{X} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{12}$$

where z_{i0} is the baseline $\sqrt{\mathrm{CD4}}$ and $E(z_{ij}) = z_{ij}^*$ is true (but unobservable) value of $\sqrt{\mathrm{CD4}}$ at time t_{ij} defined in (4). The decay rate of the second compartment in (11), β_{5ij}, is time-varying due to z_{ij}^*, but other parameters in $\boldsymbol{\beta}_{ij}$ are time independent. As mentioned previously, the mean functions in different components may involve different subsets of $\boldsymbol{\beta}_{ij}$; for example, $g_1(\cdot)$ only involves parameters β_{1i} and β_{2i}, and $g_2(\cdot)$ and $g_3(\cdot)$ share the same parameters β_{1i}, β_{2i}, and β_{3i} but have different second-phase decay rate, β_{4i} and β_{5ij}, respectively. The diagonal indicator matrices, $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, and $\boldsymbol{A}_3$, determine which elements of $\boldsymbol{\beta}_{ij}$ are involved and set other unrelated parameters to be 0 in the mean functions, $g_1(\cdot)$, $g_2(\cdot)$, and $g_3(\cdot)$, respectively. It is noted that (11) is a natural extension of (10) to consider a time-varying decay rate for capturing viral rebound in class 3 trajectories. With this mixture clustering, our mixture modeling can be used to estimate probabilities of class membership which is either viral rebound eventually, suggesting virologic failure (class 3) or viral decrease continuously, indicating a confirmed shorter-term virologic suppression (class 1) and longer-term virologic suppression (class 2).

2.4 Time-to-event models with an unspecified distribution

Various accelerated failure time (AFT) models for time-to-event were investigated in the literature [9,16]. However, the common assumption of distributions for model errors is log-normal and this assumption may lack the robustness and/or may violate the agreement with observed data. Thus, statistical inference and analysis with log-normal assumption may lead to misleading results. We consider AFT models with error term to have a nonparametric prior distribution, which follows a Dirichlet process (DP) prior [21,22]. The event time $\Im_i$ is likely related to the longitudinal response and covariate processes. We specify the association by assuming that, conditional on the random-effects $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$, the event time is related to the longitudinal processes through the random-effects that characterize the individual-specific effects [16,19].

Assume that $\Im_i$ is time to the first decline in the CD4/CD8 ratio for the ith subject ($i = 1, 2, \dots, n$). We are interested in the association of time to immune suppression of the individual-specific initial viral decay rates and the true CD4 trajectory, which are characterized by the random-effects in the viral load response and CD4 covariate models. We may view the (unobservable) random-effects as error-free covariates in time-to-event models. The associated inference can be computationally challenging when joint models consist of (nonlinear) longitudinal models and the semiparametric Cox proportional hazards model for time-to-event, a commonly used failure time

model [9], especially when the event times are interval censored. To focus on the primary interests discussed in the article, we consider the following AFT model for time to first decline of the CD4/CD8 ratio.

$$
\begin{aligned}
\ln(\Im_i) &= \boldsymbol{\gamma}^T \boldsymbol{\tau}_i + \varepsilon_i, \ \ \varepsilon_i \sim G(\cdot), \\
G(\cdot) &\sim DP(\eta G_0), \ G_0 \sim N(\zeta_0, \sigma_0^2), \ \eta \sim \Gamma(\eta_{10}, \eta_{20}).
\end{aligned}
\tag{13}
$$

where $\boldsymbol{\tau}_i = (1, a_{i0}, a_{i1}, a_{i2}, b_{i2}, b_{i4})^T$ with unknown coefficients $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_5)^T$ and we assume ε_i follows an unspecified distribution $G(\cdot)$ that is the DP prior [21,22] with the concentration parameter η and the base measure G_0 following a Gamma distribution and a specified normal distribution, respectively. Note that the parameter G_0 is the prior mean of $G(\cdot)$ and represents a guess; the other parameter of the DP prior, η, reflects the degree of closeness of $G(\cdot)$ to the prior mean G_0. Large values of η make $G(\cdot)$ very close to G_0, while small values of η allow $G(\cdot)$ to deviate from G_0. This nonparametric analysis of specifying the distribution of ε_i is robust to misspecification of the distributional assumptions about the model error. To implement the AFT model with the Dirichlet Process (DP) prior distribution, a popular way for specifying the DP prior $DP(\eta G_0)$ is the stick-breaking prior representation [22]. That is, the distribution of ε_i can be expressed as

$$
\varepsilon_i \sim \sum_{g=1}^{\infty} w_g N(\mu_g, \kappa_g) \text{ with } \mu_g = \mu_g^* - \sum_{g=1}^{\infty} w_g \mu_g^*, \ (\mu_g^*, \kappa_g) \sim G_0,
$$

where w_g is a random probability weight chosen to be independent of (μ_g^*, κ_g) such that $0 \leq w_g \leq 1$ and $\sum_{g=1}^{\infty} w_g = 1$. For the sake of computational convenience, we consider the following mixture formulation of the truncated approximation DP.

$$
\varepsilon_i \sim \sum_{g=1}^{L} w_g N(\mu_g, \kappa_g), \text{ with } \mu_g = \mu_g^* - \sum_{g=1}^{L} w_g \mu_g^*, \ (\mu_g^*, \kappa_g) \sim G_0,
\tag{14}
$$

where $0 \leq G < \infty$ and w_g is defined by the following stick-breaking procedure.

$$
w_1 = \nu_1 \text{ and } w_g = \nu_g \prod_{l=1}^{g-1} (1 - \nu_l) \text{ for } g = 2, \ldots, G,
\tag{15}
$$

where $\nu_g \sim \text{Beta}(1, \eta)$ for $g = 1, 2, \ldots, L-1$ and $\nu_L = 1$ so that $\sum_{g=1}^{L} w_g = 1$. The prior distribution for the unknown parameter η is given by $\Gamma(\eta_{10}, \eta_{20})$ with prespecified hyperparameters η_{10} and η_{20}, and the details for selecting L can refer to the publication [22].

In model (13), the random-effects b_{i2} and b_{i4} represent individual variations in the first- and second-phase viral decay rates, respectively, so they may be predictive for event times. While b_{i1} and b_{i3} represent variations in the baseline viral loads, they do not appear to be highly predictive of event time, so they are excluded from the model to reduce the number of parameters. Model (13) offers the following advantages: (i) the random-effects in the covariate model summarize the history of the covariate process and the summary quantities are likely better predictors than the covariates at several particular times; (ii) the random-effects in the response model summarize individual variations in the first- and second-phase viral decay rates to predict time to immune suppression; (iii) the link among the three models is made clear by the shared random-effects; (iv) the model error is assumed to be more flexible and unspecified distribution that has DP prior; (v) it is easy to implement.

Model (13) may be a good choice when the event times are thought to depend on individual-specific longitudinal trajectories, such as initial intercepts and slopes, or summaries of the longitudinal trajectories, and it is closely related to so-called shared parameter models [28]. As pointed out in the studies by Tsiatis and Davidian [9] and Wu et al.[16], in some practical situations, the event times cannot be observed but are only known as being contained in some time intervals, i.e., being interval-censored. In the AIDS study mentioned previously, for example, given that CD4 and CD8 were collected at a finite number of times, we can only know that the time to the first occurrence of CD4/CD8 decrease in a subject is between two data collection time points. The observed event time data are then $(u_i, v_i], i = 1, \ldots, n$, where $(u_i, v_i]$ is the smallest observed interval containing $\Im_i$. We take u_i as the subject's latest time in the study and $v_i = \infty$ if the ith subject did not experience the event of interest during the whole study period.

3 Simultaneous Bayesian inferential approach

In a longitudinal study, such as the AIDS study described previously, the longitudinal response, the covariate processes, and the time-to-event are usually connected physically or biologically. Models (3), (6) and (8) in connection with the time-to-event (13) define the MNLMEJ models considered in this article. They can be jointly modeled through the shared random-effects. Statistical inference on all of the model parameters can then be made simultaneously. Although a simultaneous inferential method based on a joint likelihood for the response, covariate and event time data may be favorable, the computation associated with the simultaneous likelihood inference in the MNLMEJ models can be extremely complex and may lead to convergence problems; in some cases it can even be computationally infeasible [7–10]. Here, we propose a simultaneous Bayesian inferential method via Markov chain Monte Carlo (MCMC) procedure to estimate both probabilities of class membership and all model parameters. The MNLMEJ models-based Bayesian inferential approach may pave a way to balance the computational burdens and convergence problems for such complex model setting.

Let $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_1^2, \sigma_2^2, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b\}$ be the collection of unknown parameters in the MNLMEJ models except for the mixture weight $\boldsymbol{\pi}$ in (8). Under the Bayesian framework, we next need to specify prior distributions for all the unknown parameters in the MNLMEJ models as follows.

$$\begin{aligned} \boldsymbol{\alpha} &\sim N_3(\boldsymbol{\tau}_1, \boldsymbol{\Lambda}_1), \ \boldsymbol{\beta} \sim N_6(\boldsymbol{\tau}_2, \boldsymbol{\Lambda}_2), \ \boldsymbol{\gamma} \sim N_6(\boldsymbol{\tau}_3, \boldsymbol{\Lambda}_3), \\ \sigma_1^2 &\sim IG(\omega_1, \omega_2), \ \sigma_2^2 \sim IG(\omega_3, \omega_4), \ \boldsymbol{\Sigma}_a \sim IW(\boldsymbol{\Omega}_1, \nu_1), \ \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{\Omega}_2, \nu_2), \end{aligned} \tag{16}$$

where the mutually independent Normal (N), Inverse Gamma (IG) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations. The super-parameter matrices $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$, $\boldsymbol{\Lambda}_3$, $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ can be assumed to be diagonal for convenient implementation. By its definition, the latent indicating variables c_i $(i = 1, ..., n)$ follow a Categorical distribution(Cat)

$$c_i \overset{iid}{\sim} Cat((1, 2, 3), (\pi_1, \pi_2, \pi_3)), \tag{17}$$

in which $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3)^T$ follows a Dirichlet distribution(Dir)[11],

$$\boldsymbol{\pi} \sim Dir(\phi_1, \phi_2, \phi_3). \tag{18}$$

Under the umbrella of the MNLMEJ models (3), (6), (8) and (13), the MCMC procedure consists of the following two iterative steps:

(i). Sampling class membership indicators c_i $(i = 1, ..., n)$, conditional on population parameters, $\boldsymbol{\theta}$, and individual random-effects, $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$.

Generate c_i $(i = 1, ..., n)$ from

$$P(c_i = k | \boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{\theta}, \boldsymbol{y}_i) = \frac{\pi_k f(\boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, c_i = k, \boldsymbol{\theta})}{\sum_{m=1}^3 \pi_m f(\boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, c_i = m, \boldsymbol{\theta})}, \tag{19}$$

where $f(\boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, c_i = k, \boldsymbol{\theta})$ $(k = 1, 2, 3)$ is a conditional density function of $\boldsymbol{y}_i$ based on (7). Then, the probability $\boldsymbol{\pi}$ can be updated from the following distribution for next iteration

$$(\boldsymbol{\pi} | num_1, num_2, num_3) \sim Dir(\phi_1 + num_1, \phi_2 + num_2, \phi_3 + num_3), \tag{20}$$

where $num_k = \sum_{i=1}^n I(c_i = k), (k = 1, 2, 3)$, in which $I(\cdot)$ is an indicator function.

(ii). Sampling parameters $\boldsymbol{\theta}$, and individual random-effects $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$, conditional on class membership indicator $\boldsymbol{c} = (c_1, \ldots, c_n)^T$.

It can be shown, conditional on c_i determined in step (i), that $\boldsymbol{z}_i$ and $\boldsymbol{y}_i$ (in the presence of left-censoring) with respective random-effects $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ in conjunction with the time-to-event model (13) can be hierarchically formulated as

$$\begin{aligned} \boldsymbol{z}_i | \boldsymbol{a}_i &\sim N_{n_i}(\boldsymbol{z}_i^*, \sigma_1^2 \boldsymbol{I}_{n_i}), \ \boldsymbol{a}_i \sim N_3(\boldsymbol{0}, \boldsymbol{\Sigma}_a), \\ \boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, c_i &\sim N_{n_i}(\boldsymbol{g}_{c_i}(\boldsymbol{t}_i, \boldsymbol{A}_{c_i}\boldsymbol{\beta}_i), \sigma_2^2 \boldsymbol{I}_{n_i}), \ \boldsymbol{b}_i \sim N_4(\boldsymbol{0}, \boldsymbol{\Sigma}_b), \\ \Im_i | \boldsymbol{\tau}_i &\sim G\left(\ln(t_i) - \boldsymbol{\gamma}^T \boldsymbol{\tau}_i\right), \end{aligned} \tag{21}$$

where $\boldsymbol{z}_i^* = (z_{i1}^*, ..., z_{in_i}^*)^T$.

Let $f(\cdot|\cdot)$, $F(\cdot|\cdot)$ and $h(\cdot)$ denote a probability density function (pdf), cumulative density function (cdf) and prior density function, respectively. Conditional on the random variables and some unknown parameters, with the specification of the Tobit model, a detectable measurement y_{ij} contributes $f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{a}_i, c_i)$, whereas a non-detectable measurement contributes $F(\rho|\boldsymbol{b}_i, \boldsymbol{a}_i, c_i) \equiv P(y_{ij} < \rho|\boldsymbol{b}_i, \boldsymbol{a}_i, c_i)$ in the likelihood. We assume that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_1^2, \sigma_2^2, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b$ are independent of each other, i.e., $h(\boldsymbol{\theta}) = h(\boldsymbol{\alpha})h(\boldsymbol{\beta})h(\boldsymbol{\gamma})h(\sigma_1^2)h(\sigma_2^2)h(\boldsymbol{\Sigma}_a)h(\boldsymbol{\Sigma}_b)$. After we specify the models for the observed data denoted by $\Re = \{(\boldsymbol{q}_i, \boldsymbol{d}_i, \boldsymbol{z}_i, z_{i0}, u_i, v_i), i = 1, \ldots, n\}$ and the prior distributions for the unknown model parameters, we can make statistical inference for the parameters based on their posterior distributions under Bayesian framework. Thus, the joint posterior density of $\boldsymbol{\theta}$ based on the observed data $\Re$ and classification indicator $\boldsymbol{c}$ can be given by

$$f(\boldsymbol{\theta}|\Re, \boldsymbol{c}) \propto \{ \prod_{i=1}^{n} \int \int L_{\boldsymbol{y}_i} f(\boldsymbol{z}_i|\boldsymbol{a}_i) F^*(u_i, v_i|\boldsymbol{\tau}_i) f(\boldsymbol{a}_i) f(\boldsymbol{b}_i) d\boldsymbol{a}_i d\boldsymbol{b}_i \} h(\boldsymbol{\theta}), \qquad (22)$$

where $F^*(u_i, v_i|\boldsymbol{\tau}_i) = G\left(\ln(v_i) - \boldsymbol{\gamma}^T\boldsymbol{\tau}_i\right) - G\left(\ln(u_i) - \boldsymbol{\gamma}^T\boldsymbol{\tau}_i\right)$; $L_{\boldsymbol{y}_i} = \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{a}_i, c_i)^{1-d_{ij}} F(\rho|\boldsymbol{b}_i, \boldsymbol{a}_i, c_i)^{d_{ij}}$ is the likelihood for the observed response data in which d_{ij} is the censoring indicator such that $y_{ij} = q_{ij}$ is observed if $d_{ij} = 0$ and y_{ij} is left-censored (i.e., treated as missing) if $d_{ij} = 1$.

In general, the integrals in (22) are of high dimension and do not have a closed form. Analytic approximations to the integrals may not be sufficiently accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data and class membership. As an alternative, the MCMC procedure can be used to sample population parameters, $\boldsymbol{\theta}$, and random-effects, $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ as well as latent class indicator c_i ($i = 1, ..., n$), from conditional posterior distributions, based on (22), using the Gibbs sampler along with the Metropolis-Hastings (M-H) algorithm. Steps (i) and (ii) considered in the MCMC procedure above are repeated alternatively in iterations of MCMC procedure until convergence is reached. An important advantage of the above representations based on the hierarchical models is that they are easily implemented using the freely available WinBUGS software [29] interacted with a function called bugs in a package R2WinBUGS of R. Note that when WinBUGS software is used to implement our modeling approach, it is not necessary to explicitly specify the full conditional posterior distributions or proportional functions of the density functions of full conditional posterior distributions for parameters to be estimated. Although their derivations are straightforward by working the complete joint posterior, some cumbersome algebra will be involved and are not present here to save space.

4 Analysis of AIDS clinical data

4.1 Model implementation

We conducted the following scenarios of modeling comparison. First, we proposed the joint modeling (JM) approach based on the MNLMEJ model (denoted as Model I) in Section 3. Here, we estimated the model parameters by using the "naive" method (denoted by NM), which ignores measurement error in CD4 covariate. That is, the "naive" method uses only the observed CD4 value z_{ij} rather than expected (unobservable) CD4 value $E(z_{ij}) = z_{ij}^*$ in the mixture model for viral load response (denoted by Model II). We used it as a comparison to the JM approach. This comparison attempted to investigate how the measurement errors in CD4 contribute to modeling results. Second, we further investigated a commonly used NLME joint model (Model III, ignoring data feature of heterogeneous population) where the mean function is specified by (11) alone. This NLME model has been widely applied to analyze viral load data [7,8,26] in significantly advancing the understanding of pathogenesis of HIV infection and the assessment of effectiveness of ART. We compared Model III with Model I proposed in this article to explore how heterogeneous feature influences modeling results.

To carry out the Bayesian inference, we took weakly-informative prior distributions for the parameters. In particular, (i) fixed-effects were taken to be independent normal distribution $N(0, 100)$ for each element of the population parameter vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$; (ii) we assume a noninformative inverse Gamma prior distribution $IG(0.01, 0.01)$, which has mean 1 and variance 100, for variance parameters σ_1^2 and σ_2^2; (iii) the priors for the variance-covariance matrices of the random-effects $\boldsymbol{\Sigma}_a$ and $\boldsymbol{\Sigma}_b$ were taken to be inverse Wishart distributions $IW(\boldsymbol{\Omega}_1, \nu_1)$ and $IW(\boldsymbol{\Omega}_2, \nu_2)$, where the diagonal elements for diagonal variance matrix $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ were 0.01,

and $\nu_1 = \nu_2 = 4$; and (iv) we set hyper-parameters of Dirichlet distribution in (18), $\phi_1 = \phi_2 = \phi_3 = 1$, assuming individuals have equal probabilities of coming from any one of three classes initially. (v) The hyper-parameters for the DP prior are assigned by $\zeta_0 = 0, \eta_{10} = 0.5, \eta_{20} = 0.01$ and $\sigma_0^2 = 100$. The MCMC sampler was implemented using WinBUGS software [29] interacted with R2WinBUGS of R, and the program codes for mixture joint model are available from author upon request. When the MCMC procedure was applied to the actual clinical data, convergence of the generated samples was assessed using standard tools within WinBUGS software such as trace plots and Gelman-Rubin (GR) diagnostics [30]. We observed from the plots of trace, autocorrelation and GR diagnostics that the algorithm has approached convergence. Along with the diagnostic convergence, we proposed that, after an initial number of 50,000 burn-in iterations of three chains of length 100,000, every 50th MCMC sample (i.e., thin set equal to 50) was retained from the next 50,000 for each chain. Thus, we obtained a total of 3,000 samples of targeted posterior distributions of the unknown parameters for statistical inference.

4.2 Results of data analysis

The Bayesian joint modeling approach in conjunction with the three models for the viral load response, the CD4 covariate and the time to first decline of CD4/CD8 ratio with different scenarios was used to fit the data. From the model fitting results, we have seen that, in general, all the models provided a reasonably good fit to the observed viral load data above LOD for most patients in our study, although the fitting for a few patients was not completely satisfactory due to unusual data fluctuation patterns for these patients, particularly for Models II and III. As an example, the three representative individual estimates of viral load trajectories based on Models I (solid lines), II (dash lines) and III (dotted lines) are displayed in Figure 2. The following findings are observed from joint modeling . The estimated individual trajectories for Model I fit the originally observed values above LOD more closely than those for Models II and III which are comparable.
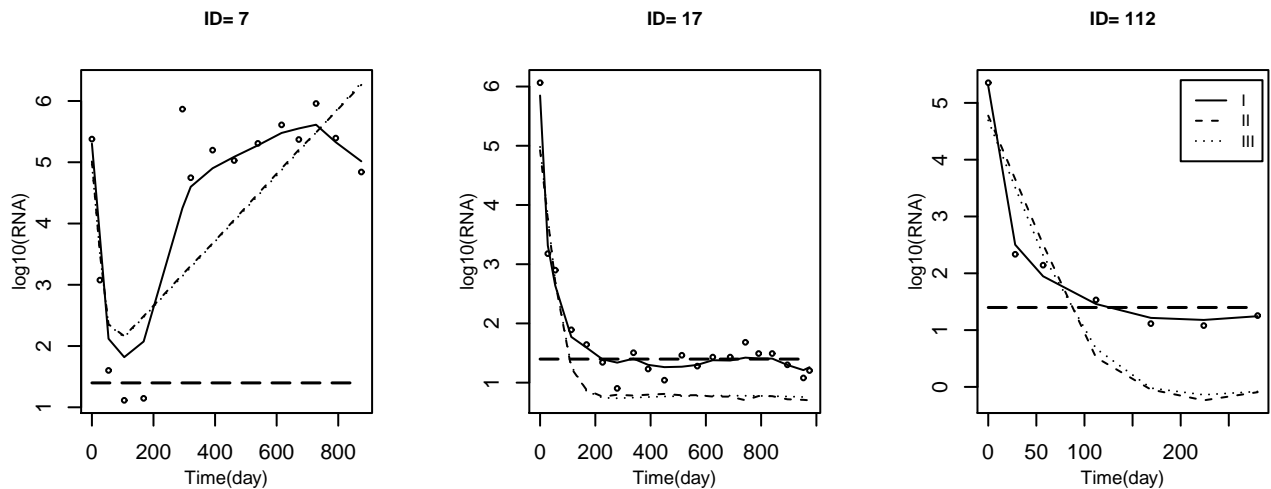
Fig. 2. Individual fitted curves of viral load for three representative patients based on Models I (solid lines), II (dash lines) and Model III (dotted lines). The observed values are indicated by circle "○". The horizontal lines represent LOD at $1.398 = \log_{10}(25)$.

The population posterior mean (PM), corresponding standard deviation (SD) and 95% credible interval (CI) for fixed-effects parameters based on Models I, II and III with two methods are presented in Tables 1 and 2. The following findings are observed for estimated results of parameters.

In the mixture of NLME response model with the three components (9)–(11), the findings, particularly for the fixed-effects (β_5, β_6), which are parameters related to second-phase viral decay rate, show that these estimates are different from zero for the three Models since the 95% credible intervals do not contain zero. Nevertheless, for the estimate of the coefficient of CD4 covariate β_6, although its estimate based on Model I (0.098) is much

Table 1. Summary of estimated posterior mean (PM) of population (fixed-effects) parameters, the corresponding standard deviation (SD), and lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI).

Method	Model		β_1	β_2	β_3	β_4	β_5	β_6	α_1	α_2	α_3	σ_1^2	σ_2^2
JM	I	PM	11.36	103.8	1.854	2.604	-3.157	0.098	13.53	8.712	68.27	4.871	0.940
		L_{CI}	11.04	92.09	0.307	1.740	-4.836	0.027	13.00	7.997	49.93	4.677	0.872
		U_{CI}	11.62	115.7	3.303	3.527	-0.892	0.184	13.98	9.536	86.99	5.058	1.023
		SD	0.145	6.015	0.720	0.510	1.059	0.052	0.271	0.426	9.405	0.097	0.043
NM	II	PM	11.34	95.88	2.364	2.106	-5.693	0.245	13.56	8.648	67.11	4.877	0.912
		L_{CI}	11.04	78.06	0.462	0.853	-7.665	0.168	12.98	7.707	48.58	4.679	0.853
		U_{CI}	11.58	111.8	4.105	3.274	-3.181	0.312	14.15	9.557	85.83	5.081	0.997
		SD	0.134	10.37	0.999	0.711	1.269	0.036	0.290	0.471	9.507	0.103	0.040
JM	III	PM	11.39	104.9	2.097	2.038	-9.012	0.157	13.50	8.854	65.56	4.863	0.915
		L_{CI}	11.16	95.01	1.047	1.360	-7.966	0.079	12.96	8.057	47.71	4.671	0.872
		U_{CI}	11.62	115.8	3.213	2.823	-3.883	0.237	13.99	9.732	84.87	5.061	0.967
		SD	0.115	5.485	0.540	0.368	0.039	0.039	0.264	0.427	9.443	0.101	0.024

Table 2. Summary of estimated posterior mean (PM) of population (fixed-effects) parameters, the corresponding standard deviation (SD), and lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI) as well as DIC and EPD values.

Method	Model		γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	DIC	EPD
JM	I	PM	-2.783	-0.155	-0.265	1.251	0.054	0.026	10368.9	1.675
		L_{CI}	-2.870	-0.493	-0.864	-0.502	-0.478	-0.036		
		U_{CI}	-2.706	0.750	0.219	2.458	1.127	0.230		
		SD	0.043	0.335	0.375	0.735	0.326	0.065		
NM	II	PM	-2.802	-0.073	-0.110	-0.050	-0.020	-0.005	12572.8	1.932
		L_{CI}	-2.890	-0.401	-1.389	-2.068	-0.246	-0.034		
		U_{CI}	-2.715	0.270	0.766	1.085	0.026	0.026		
		SD	0.045	0.158	0.599	0.778	0.064	0.015		
JM	III	PM	-2.785	-0.112	-0.177	0.485	0.001	-0.003	14099.1	2.637
		L_{CI}	-2.871	-0.581	-0.943	-1.611	-0.004	-0.027		
		U_{CI}	-2.701	0.307	0.271	1.446	0.004	0.019		
		SD	0.044	0.199	0.322	0.763	0.003	0.012		

smaller than the counterpart based on Models II and III (0.245 and 0.157), its estimate is always significantly positive. This means that CD4 has a significantly positive effect on the second-phase viral decay rate, suggesting that the CD4 covariate may be an important predictor of the second-phase viral decay rate during the treatment. The fixed-effects (β_2, β_3), which are parameters of the first-phase viral decay rate, show that the estimate of β_3, the coefficient of baseline CD4 count, is significantly positive, indicating that the baseline CD4 has positive effect on the first-phase viral decay rate.

For parameter estimates of the CD4 covariate model (4), the estimates of the coefficients based on the three models show slight differences. The estimated linear coefficient α_2 is significantly different from zero with a positive value, suggesting that there is a positive linear relation between CD4 cell count and measurement time. This result may be explained by the fact that CD4 cell count may increase during treatment, and in turn, indicates an overall improvement of AIDS progression in this AIDS trial study.

The estimates of parameters in the AFT model (13) do not directly show that the time to CD4/CD8 decrease is highly associated with either the two viral decay rates (due to insignificant estimates of γ_4 and γ_5) or the CD4

changes (due to insignificant estimates of γ_2 and γ_3) over time, which is different from what was anticipated. This finding is consistent with the study by Wu et al.[16] in which further explanations are provided.

For comparison, we employed the NM to estimate the model parameters presented in Tables 1 and 2 using the observed CD4 and ignoring the CD4 measurement error in the mixture of NLME response model with the three components (9)–(11). It is seen from estimates of (β_5, β_6), which depend on whether or not to ignore potential CD4 measurement error, that the differences between the naive estimates (-3.157, 0.098) of (β_5, β_6) and the joint modeling estimates (-5.693, 0.245), indicate that CD4 measurement error can not be ignored in the analysis.

We further investigated a commonly used NLME joint model (Model III, where data feature of heterogeneous population is ignored). We compared Model III with Model I to explore how heterogeneous feature influences modeling results. We found the important differences in the estimates of parameters (β_5, β_6), which are associated with the second-phase viral decay rate. It is noted that one advantage of Model I over Model III is its flexibility to provide not only estimates of all model parameters, but also evaluate class membership probabilities at both population and individual levels, which is helpful for clinicians to develop individualized treatment.

We now focus on the lower end of the distribution of the viral load due to LOD where the Tobit model is offered to mediate the observations below LOD as missing values. As it was mentioned in the Section 1, the current assay techniques for quantifying HIV-RNA viral load may not give accurate readings below LOD. In our analysis, we treated those inaccurate observed viral loads as missing values and predict them using the proposed (mixture) joint models in connection with the Tobit model. Note that the main advantage of our proposed Tobit model is its ability to predict the true viral load values ($\log_{10}$ scale) below LOD based on a latent variable approach. The fitted results of these models for values below LOD indicate that most observed values are piled up at the range (0.6, 1.4), whereas for Models I, II and III the predicted values of the unobserved viral load below LOD are spread out as expected (data not shown here). However, Models II and III produce the predicted values exceeded the LOD much more than Model I does. In addition, when we compare the three models in terms of their distributions in predicting viral loads below LOD, we can see that Model I gives more plausible values ranged within (-0.2, 1.5) than Models II and III do in the sense that the distribution is closely fits the lower part of the whole distribution of the predicted viral load values based on Model I as expected, implying that Model I is the better model. This finding is also confirmed by the results (see Table 2 for details) from other criteria such as deviance information criterion (DIC) [31] and expected predictive deviance (EPD) [32]. In summary, our results may suggest that it is important to consider heterogeneous and/or covariate measurement error data features-based the mixture of joint models in order to achieve reliable results. Based on these findings, we will further report our results in detail only for the best Model I below.

As mentioned in Section 1, one of the primary objectives in this analysis was to cluster all individuals' membership into 3 classes based on viral load trajectories. Based on the mixture joint modeling, we are able to obtain a summary of class membership at both the population and individual levels. At population level, the MCMC procedure yields samples from the posterior distribution of (π_1, π_2, π_3) in (20), the population proportion of individuals in each class. The estimates of population proportion and associated 95% CI of (π_1, π_2, π_3) for three classes are 5.29% (3.15%, 7.47%), 57.13% (54.34%, 59.79%) and 37.58% (34.83%, 40.43%), respectively. It can be seen that class 2 (decrease and maintain stable) has the largest proportion 57.13%, followed by class 3 (decrease and rebound later) with proportion 37.58%, and then class 1 (decrease all the time) with the lowest proportion 5.29%. Thus, out of 427 patients, the patterns of changing viral load of 23, 244 and 160 patients followed classes 1, 2 and 3, respectively. This indicates that a confirmed shorter-term (class 1 due to early dropout) and longer-term (class 2) virologic responses were observed in a total of 62.42% of the patients in classes 1 and 2.

At individual level, the posterior probability of individual i belonging to the kth ($k = 1, 2, 3$) class, $p_{ik} = E[I(c_i = k)]$, can be approximated by $\frac{1}{M} \sum_{m=1}^{M} I(c_i^{(m)} = k)$, where $c_i^{(m)}$ is class membership of individual i drawn from the posterior distribution (19) in the mth MCMC iteration ($m = 1, ..., M$) and M is a total number of posterior sample iterations ($M = 3000$ in this study). Barplot shown in Figure 3 displays the probabilities for the selected 20 individuals (from the 1st to the 20th subjects). The probability corresponding to individual patient who is classified as either viral load rebound or not may help physicians to refine treatment strategy and to identify the reason of viral load rebound for such individual patient. As examples, for the three patients shown in Figures 1 and 2, the patient 112 belongs to class 1 because the viral load decreases constantly in a early short-term period with probability 87.5%; the viral load of the patient 17 decreases and then maintain stable, and thus, this patient belongs to class 2 with probability 91.1%; and finally, the patient 7 is in class 3 (indicating viral load rebound) with probability 83.7%.
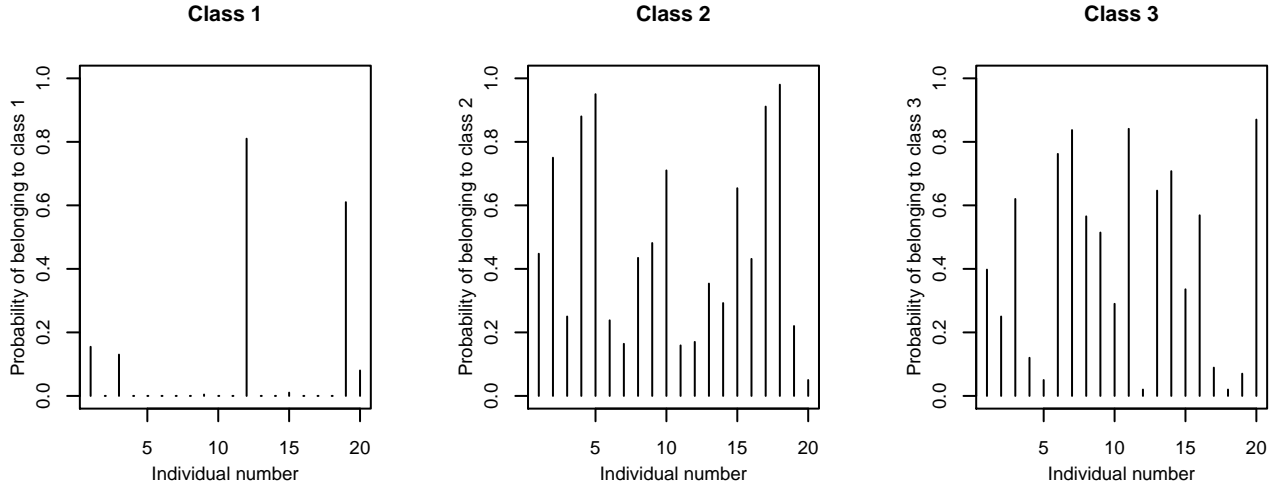
Fig. 3. Posterior probabilities of belonging to 3 trajectory classes for the first 20 individuals based on Model I.

5 Concluding discussion

We present a Bayesian joint modeling for three (response, covariate and time-to-event) processes linked through the random-effects that characterize the underlying individual-specific longitudinal processes. We consider the MNLMEJ models for survival-longitudinal data with multiple features. The approach is applied to jointly model HIV dynamics and time to decrease in CD4/CD8 ratio in the presence of the CD4 covariate process with measurement error to provide a tool to assess ART and to monitor disease progression. Among the models considered in the application, Model I for longitudinal data with features of heterogeneity, missing due to LOD in response and measurement error in covariate was found to be favorable. One advantage of mixture joint modeling approach-based models proposed in this article is its flexibility to study simultaneous impact of various data characteristics (heterogeneity, nonlinearity, missingness due to LOD and incompleteness). Another advantage of our mixture joint modeling approach is to provide not only all model parameter estimates, but also model-based probabilistic clustering to obtain class membership probabilities at both population and individual levels. In addition, the proposed mixture joint modeling approach can be easily implemented using the publicly available WINBUGS interacted with R. This makes our models and methods quite powerful and accessible to practicing statisticians in the field.

The estimated results of fixed-effects presented in Table 1 based on Model I indicate that the first-phase decay rate, and the second-phase decay rate without and with time-varying CD4 covariate may be approximated by $\hat{\lambda}_1 = 103.8 + 1.854z_0$, $\hat{\lambda}_2 = -3.157$ and $\hat{\lambda}_2(t) = -3.157 + 0.098(-13.53 + 8.712\psi_1(t) + 68.27\psi_1(t))$, respectively, where z_0 is the baseline $\sqrt{CD4}$ value, $\psi_1(t)$ and $\psi_2(t)$ are two basis functions given in Section 2.1. Thus, the population viral load processes of 3 classes may be approximated by $\hat{V}_1(t) = \exp\left\{11.36 - \hat{\lambda}_1 t\right\}$, $\hat{V}_2(t) = \exp\left\{11.36 - \hat{\lambda}_1 t\right\} + \exp\left\{2.604 - \hat{\lambda}_2 t\right\}$ and $\hat{V}_3(t) = \exp\left\{11.36 - \hat{\lambda}_1 t\right\} + \exp\left\{2.604 - \hat{\lambda}_2(t)t\right\}$. Since the first-phase viral decay rate, λ_1, is significantly associated with the baseline CD4 (due to significant estimate of β_3) and the second-phase viral decay rate $\lambda_2^*(t)$ in the component three is significantly associated with the time-varying CD4 values (due to significant estimate of β_6), this suggests that the viral load change $V(t)$ may be significantly associated with both the baseline CD4 and time-varying CD4 values. This simple approximation considered here may provide a rough guidance and point to further research even though the true association described above may be more complicated. The analysis results suggest that for patients with viral load rebound, the CD4 process at time t has a significantly positive effect on the second-phase viral decay rate; this finding confirms the fact that the CD4 covariate is a more significant predictor on viral decay rate during late stage and more rapid increase in CD4 cell count may be associated with faster viral decay in late stage, which, in turn, may result in a viral load suppression. In addition, the estimated results of the parameters in the time-to-event model (13) based on (best)

345

Model I suggest that the time to first decline of CD4/CD8 ratio is not highly associated with either the two viral decay rates or the CD4 changes over time. This finding is consistent with that in the study by Wu *et al.* [16] in which further explanations are provided.

In order to examine the sensitivity of parameter estimates to the prior distributions, we conducted a limited sensitivity analysis using the uniform prior distribution $U(0, 100)$ for standard deviation scales σ_1 and σ_2 [33] instead of inverse gamma prior distribution (used in this paper) for variance parameters σ_1^2 and σ_2^2; we refitted data based on Model I. We found from the results (not shown here) of sensitivity analysis that the conclusions with different prior distributions are in agreement. Thus, the findings from our analysis remain unchanged.

For inference of mixture modeling, parameter (or model) identifiability can be an important but difficult problem when a large number of model parameters must be estimated simultaneously. We must ensure each component model to be identifiable to make whole mixture model to be identifiable. To make (10) and (11) to be identifiable, we assume $\lambda_{1i} > \lambda_{2i}$ in (10) and $\lambda_{1i} > \lambda_{2ij}^*$ in (11). In practice, if the models are not identifiable, the MCMC algorithm would diverge quickly. In the application considered in this article, the MCMC algorithm was converged without problems and we did not observe potential identifiability problems. It is noted that the mixture components, $g_k(\cdot)$ in the finite mixture model may have the same family of densities but differ only in specific values of parameters such as in their means, or have completely different functional forms with parameters of different dimensions and meanings across the sub-models [14], which is the case adopted in this article. Because of complexity of longitudinal studies, an advantage of choosing components correspond to different densities with distinguished mean functions $g_k(\cdot)$, which are known and pre-specified, rather than with the same general family in the finite mixture models is that identifiability problem can be avoided [14].

It is noted that the models in our previous publications [6,12,19] can be treated as special cases of MNLMEJ model proposed in this article when the MNLMEJ model is tailored as follows. In comparison with the proposed MNLMEJ model which can be offered to not only estimate all parameters simultaneously, but also evaluate probabilities of class membership, Huang and Dagne [6] considered only commonly used NLME model with mean function specified by (11) for viral load response (not a mixture) in the presence of CD4 measurement error specified by a standard linear mixed-effects model without time-to-event model involved; Huang et al.[19] investigated longitudinal-survival data analysis based on commonly used NLME model with mean function specified by (11) for viral load response, but mixture of model structure is ignored in this sense that classification of viral load trajectories was not evaluated; although Lu and Huang [12] studied a finite mixture of NLME models, both measurement error model for CD4 covariate and AFT model for time to decline of CD4/CD8 ratio were not considered in the sense that model in Lu and Huang [12] is not conducted under framework of joint models.

The final three issues related to this application are noted as follows. (i) The original motivation of this mixture joint modeling was to cluster all patients into two classes with or without viral load rebound, which is of main interest from a clinical prospective. But another class was needed (class 1) to capture some patients who withdrew too early to be clustered into either class 2 with viral load suppression or class 3 with viral load failure. Thus, the number of components in this analysis was determined empirically based on the viral load trajectory patterns and clinical interpretability. Note that class 1 is referred as a confirmed 'short-term virologic response' due to early dropout which may be informative dropout and may not indicate virologic suppression in long-term treatment. Toward this end, we may consider that dropouts are likely to be informative or non-ignorable missing in the sense that the probability of dropouts (missing data) may depend on the missing values. Thus, we can modify the mixture joint models conducted in this article to cluster all patients into two classes with virologic suppression and failure, respectively. (ii) Although the normal distribution assumption in both response and covariate models may be appropriate (Histograms of viral load (in log scale) and CD4 (in square-root) not shown here) for this specific data analysis, the mixture joint models in this article can be extended to assume model errors with skew distributions as considered in the literature [5,6,19,34] if data exhibit non-normality; however, the computational burden and other issues may be involved. (iii) The purpose of this article is to demonstrate the proposed MNLMEJ models using a real data analysis, although a limited simulation study may be conducted to evaluate the performance of the proposed models and method. However, since this article investigated different scenarios based on relatively complex model specifications focusing on real data analysis, the complex natures considered in this article, especially the three model components involved, will pose some challenges for such simulation studies with intensive computational burden which requires additional significant efforts. We are actively investigating these important research problems and hope that the relative results could be reported in the near future. Although this article is motivated by AIDS clinical study, the basic concepts of the developed

MNLMEJ models and methods have generally broader applications whenever the relevant technical specifications are met and longitudinal measurements are assumed to arise from two or more identifiable subclasses within a population.

Appendix A. HIV dynamic models

Viral dynamic models can be formulated through a system of ordinary differential equations (ODE) [3,18,23,35]. Under some reasonable assumptions and simplifications, two useful approximations of ODE solution, which can be used to capture viral load responses, have been proposed as follows.

$$V(t) = \log_{10}(e^{p_1 - \lambda_1 t}) \tag{A.1}$$

$$V(t) = \log_{10}(e^{p_1 - \lambda_1 t} + e^{p_2 - \lambda_2 t}) \tag{A.2}$$

where $V(t)$ is the $\log_{10}$ scaled plasma HIV-1 RNA levels at time t. λ_1 and λ_2 are called the first- and second-phase viral decay rates, which may represent the minimum turnover rate of productively infected cells and that of latently or long-lived infected cells, respectively [3]. The parameters p_1 and p_2 are macro-parameters; e^{p_1} and $e^{p_1} + e^{p_2}$ are the baseline viral load at time $t = 0$ in one- and two-compartment models, respectively. It is generally assumed that $\lambda_1 > \lambda_2$, which assures that the model is identifiable and is appropriate for empirical studies [23]. Models (A.1) and (A.2) offer almost equal performance to capture the early fast-decaying segment of viral load trajectory [23], but model (A.2) performs better for long term of viral load trajectory [6]. It was noted that both models can be only applied to the early segment or longer term of the viral load response with decreasing trajectory patterns as discussed in Section 1 and shown in Figure 1(a)(two solid and two dashed lines), but the viral load trajectory may have a different shape or rebound in later stages (two dotted lines in Figure 1(a)). It was also noted that for some patients the second viral decay rate λ_2, may vary over time because they depend on some phenomenological parameters that hide with considerable microscopic complexity and change over time. Negative values of the decay rates may correspond to viral increase and lead to viral rebound [23], suggesting that variation in the dynamic parameters, particularly λ_2, may be partially associated with time-varying covariates such as repeated CD4 cell counts to capture the viral load change including viral rebound. Thus, we the following extended function are introduced [26].

$$V(t) = \log_{10}(e^{p_1 - \lambda_1 t} + e^{p_2 - \lambda_2(t) t}) \tag{A.3}$$

where the second-phase decay rate $\lambda_2(t)$ is either a function of time-varying CD4 cell count (covariate) and/or an unknown smooth function. Intuitively, this model is more flexible because it assumes that the second-phase viral decay rate can vary with time as a result of drug resistance, medication adherence and other relevant clinical factors likely to affect changes in the viral load during the treatment. Therefore, all data obtained during whole study period can be used by fitting this model. The following are a couple of issues related to the proposed mixture of joint models which should be addressed.

References

1. Fischl, M.A., Ribaudo, H.J., Collier, A.C., et al.. (2003). A randomized trial of 2 different 4-drug antiretroviral regimens versus a 3-drug regimen, in advanced human immunodeficiency virus disease. *Journal of Infectious Disease* **188**, 625–634.
2. Pawitan, Y. and Self, S. (1993). Modeling disease marker processes in AIDS. *Journal of the American Statistical Association* **88**, 719–726.
3. Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M. and Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* **387**, 188–191.
4. Hughes, J.P. (1999). Mixed effects models with censored data with applications to HIV RNA levels. *Biometrics* **55**, 625–629.
5. Lachos, V.H., Bandyopadhyay, D. and Dey, D.K. (2011). Linear and nonlinear mixed-effects models for censored HIV viral loads using normal/independent distributions. *Biometrics* **67**, 1594–1604.

6. Huang, Y. and Dagne, G.(2011). A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates. *Biometrics* **67**, 260-269.

7. Liu, W. and Wu, L. (2007). Simultaneous inference for semiparametric nonlinear mixed-effects models with covariate measurement errors and missing responses. *Biometrics* **63**, 342–350.

8. Wu, L. (2002). A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of the American Statistical Association* **97**, 955–964.

9. Tsiatis, A.A. and Davidian, M. (2004). An overview of joint modeling of longitudinal and time-to-event data. *Statistica Sinica* **14**, 809–834.

10. Thiebaut, R., Jacqmin-Gadda, H., Babiker, A. and Commenges, D. (2005). Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of CD4+ cell count and HIV RNA viral load in response to treatment of HIV infection. *Statistics in Medicine* **24**(1), 65–82.

11. Diebolt, J. and Robert, C. (1994). Estimation of finite mixture distributions by Bayesian sampling. *J. Royal Statist. Soc. Series B* **56**, 363–375.

12. Lu, X. and Huang, Y. (1024). Bayesian analysis of nonlinear mixed-effects mixture models for longitudinal data with heterogeneity and skewness. *Statistics in Medicine* **33**, 2830–2849.

13. Muthén, B. and Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* **55**, 463–469.

14. Pauler, D. and Laird, N.M. (2000). A mixture model for longitudinal data with application to assessment of noncompliance. *Biometrics* **56**, 464–472.

15. Yi, G.Y., Liu, W. and Wu, L. (2011). Simultaneous inference and bias analysis for longitudinal data with covariate measurement error and missing responses. *Biometrics* **67**, 67–75.

16. Wu, L., Liu, W. and Hu, X.J. (2010). Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics* **66**, 327-335.

17. Brown, E.R. and Ibrahim, J.G. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics* **59**, 221-228.

18. Guedj, J., Rodolphe Thiebaut, R. and Commenges, D. (2010). Joint modeling of the clinical progression and of the biomarkers dynamics using a mechanistic model. *Biometrics* **67**, 59–66.

19. Huang, Y., Dagne, G.A. and Wu, L. (2011). Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors. *Statistics in Medicine* **30**, 2930–2946.

20. Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica* **26**, 24–36.

21. Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Annals of Statistics* **1**, 209–230.

22. Ishwaran, H. and Zarepour, M. (2000). Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika* **87**, 371–390.

23. Wu, H. and Ding, A.A. (1999). Population HIV-1 dynamics in vivo: Applicable models and inferential tools for virological A data from AIDS clinical trials. *Biometrics* **55**, 410–418.

24. Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models: Modern Perspective*, 2nd edition. London, Chapman and Hall.

25. Davidian, M. and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. London, Chapman and Hall.

26. Wu, H, and Zhangm J,T. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. Hoboken, New Jersey, Wiley.

27. Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). *Statistical Analysis of Finite Mixture Distributions*. Chichester, U.K., John Wiley and Sons.

28. DeGruttola, V. and Tu, X.M. (1994). Modeling progression of CD4-lymphocyte count and its relationship to survival time. *Biometrics* **50**, 1003–1014.

29. Lunn, D.J., Thomas, A., Best, N. and Spiegelhalter, D. (2000). WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.

30. Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–511.

31. Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statiscal Socity, Series B* **64**, 583–639.

32. Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2003). *Bayesian Data Analysis*. London, Chapman and Hall.

33. Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* **1**, 515-533

34. Bandyopadhyay et al.(2012). Skew-normal/independent linear mixed models for censored responses with applications to HIV viral loads. *Biometrical Journal* **54**(3), 405–425.

35. Huang, Y., Liu, D. and Wu, H. (2006). Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics* **62**, 413–423.

On understanding company perceptions of risk – a text mining application

Zhuopeng Jiang[1] & Jim Freeman[2]

Manchester Business School, The University of Manchester, Booth Street East, Manchester, M15

6PB, UK

[1] Email: Paul.Jiang@cn.ey.com

[2] Email: jim.freeman@mbs.ac.uk

Abstract. Two methodologies – thematic analysis and laddering theory – were used to analyse qualitative data on risk management. The data – in the form of interview transcripts – were obtained from an in-depth survey of how risk is perceived by UK business. Using the NVivo software, four main themes were identified: concepts of risk, risk management, roles in risk management and data encryption. From these, inferential codes on attributes, consequences and values were used in conjunction with means–end theory to help explain the motivations behind the competing core values of the participating organisations.

Keywords: Hierarchy value map, Implication matrix, Laddering theory, Risk management.

1 Introduction

Risk is the potential of losing something of value – be it economic, financial or physical - and arises from the uncertainty associated with a particular course of action [1]. IRM (2015) defines risk as "the combination of the probability of an event and its consequence" [2] – consequences potentially being considered positively as well as negatively.

Risk management is concerned with helping organisations perform more effectively in a world of uncertainty [3] and is fundamental to the need for making better decisions, protecting business continuity, learning from the past, predicting the future and managing out interruptions. Dealt with constructively, risks can often be turned into opportunities.

Managing risk is not necessarily the same as removing it. According to the Joint Information Systems Committee [4] it is helpful to see risk management - the systematic process of identifying, analysing and responding to various business risks – in terms of five stages (see Fig. 1):

1. Risk identification;
2. Qualitative risk analysis;
3. Quantitative risk assessment;
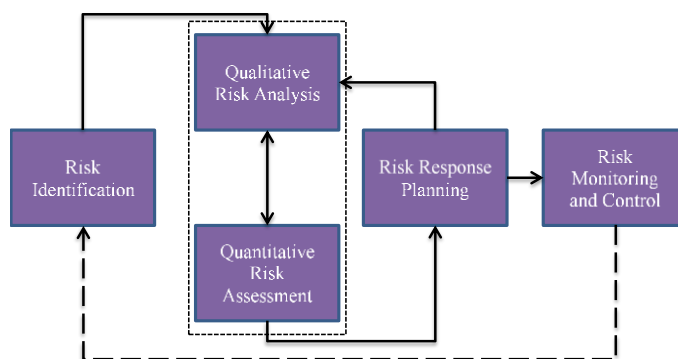4. Risk response planning;
5. Risk monitoring and control.

Fig 1. The Process of Risk Management

2 Qualitative Data Analysis

Focussing on the second stage here, most risk analysis is qualitative in nature - qualitative data being widely considered 'richer' and 'more valid' than quantitative data [5] - as well as more concrete, vivid and meaningful, especially when presented as an event or a story [6].

A variety of different approaches to qualitative analysis have evolved in recent years [7], [8] including *Grounded Theory*, *Content Analysis*, *Thematic Analysis, Framework Analysis, Discourse Analysis*, *Conversation Analysis*, *Case Study*, *Narrative Inquiry*, *Phenomenology*, *Ethnography* and *Laddering Theory.*

And where once qualitative data analysis used to be an arduous, exhausting and time-consuming process, the development of computer assisted qualitative data analysis software (CAQDAS) - has greatly enhanced both the speed and quality of analytical provisions in the area [9].

Table 1 lists some of the better-known CAQDAS packages in recent use [10]. Of these, Atlas.ti and NVivo are particularly well rated [11]. However, Barry [12] and Bazeley [13] most favour NVivo, the software adopted for the study.

CAQDAS Package	Developer	Current Version (Year of Release)
NVivo	QSR International	NVivo 10 (2012)
Atlas.ti	ATLAS.ti Scientific Software Development GmbH	Atlas.ti 7.0 (2012)
MAXQDA	VERBI GmbH	MAXQDA 11 (2012)
Aquad	Günter Huber	Aquad 7 (2012)
QDA Miner	Provalis Research	QDA Miner 4 (2011)

Table 1. CAQDAS Packages

For the latter, data was supplied by IIRSM in the form of interview transcripts. These were generated as part of a project by the *International Institute of Risk and Safety Management* (*IIRSM*) and *Manchester Business School* (*MBS*), *The University of Manchester* aimed at improving the reliability of management information for risk practitioners [14]. Relevant background on the three survey respondents concerned are provided in Table 2 below.

Using NVivo, the data were first exposed to a thematic analysis. This formed the basis of a subsequent laddering formulation based on means end theory.

Participating Organisation	Participant	About the Organisation
Organisation 1: A Leading Logistics Organisation	Head of Safety	Organisation 1 is a leading logistics and postal service of the UK responsible for universal mail collection and delivery.
Organisation 2: A Leading Chain of British Supermarkets	Strategic Safety Manager	Organisation 2 is an upmarket chain of British supermarkets, forming the food retail division of Britain's largest employee-owned retailer.
Organisation 3: A Publicly-Funded Research Organisation	Head of Safety, Security and Resilience	Organisation 3 supports research across the entire spectrum of medical sciences, in universities and hospitals, in its own units, centres and institutes in the UK, and in its units in Africa.

Table 2. Participating Organisations by Type of Business

Thematic analysis is the qualitative data analysis approach most widely used for pinpointing, analysing and reporting patterns (or "themes") within data. There are six distinct phases to the analysis: (1) becoming familiar with the data (2) generating initial codes (3) searching for themes (4) reviewing themes (5) defining and naming themes, and (6) producing the report [15]. The analysis typically progresses

iteratively with constant moving back and forth through the phases as necessary, rather than simply moving linearly, from one phase to the next.

Using NVivo four main themes were uncovered from the IIRSM transcript data:

- Concepts of Risk
- Risk Management
- Roles in Risk Management
- Data Encryption

and these formed the basis of the subsequent laddering analysis.

Laddering Theory

Laddering methodology provides a simple, systematic method for establishing the core constructs by which an individual perceives the world [16]. Relying on Means-End Theory from marketing [17], it provides a model for assessing how consumers choose products or services with specific attributes (A), the consequences (C) of which link to desired goals and individual values (V).

For example, Fig. 2 shows a respondent's ladder from a luxury car study.

(A) sleek look
↓
(A) quality
↓
(C) willing to pay more
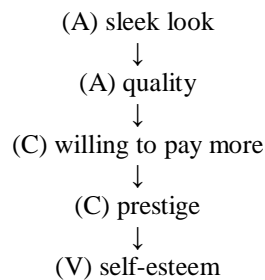↓
(C) prestige
↓
(V) self-esteem

Fig. 2. An Example of a Ladder

Once ladders for a given application have been aggregated across multiple interviewees, they can be summarised using standard content analysis procedures. The resultant table ('implication matrix') can then be used to form a so-called Hierarchical Value Map (HVM) to graphically highlight the A-C-V connections that principally determine individuals' core values [18]. For the data in hand, the steps were as follows:

1. Ladder Construction. To maintain data quality and the practicality of subsequent findings, the interviews – all conducted by IIRSM - closely followed the guidelines recommended by [19]. Resultant ladders of Organisations 1, 2 and 3 are shown respectively in Fig. 3, 4, and 5.

Ladder 1:
A: a fantastic database - A: a systematic approach - C: manage cost (full control over cost) -
C: get a payback for it - V: have a competitive advantage

Ladder 2:
A: an effective safety management system - C: manage risk effectively - V: stable and
corporately responsible organisation

Ladder 3:
A: wider awareness of risks - C: make individuals behave differently - C: help decision
makers to make proper decisions - V: protect the value of the overall organisation

Ladder 4:
A: a whole systematic suite of things - C: manage the risk - C: protect standards, projects
consistency across the business - V: protect the value of the overall organisation
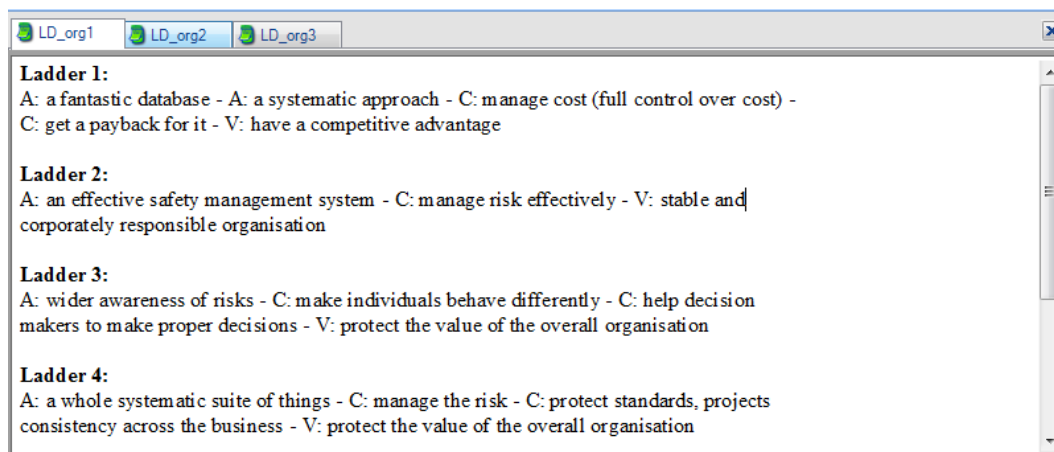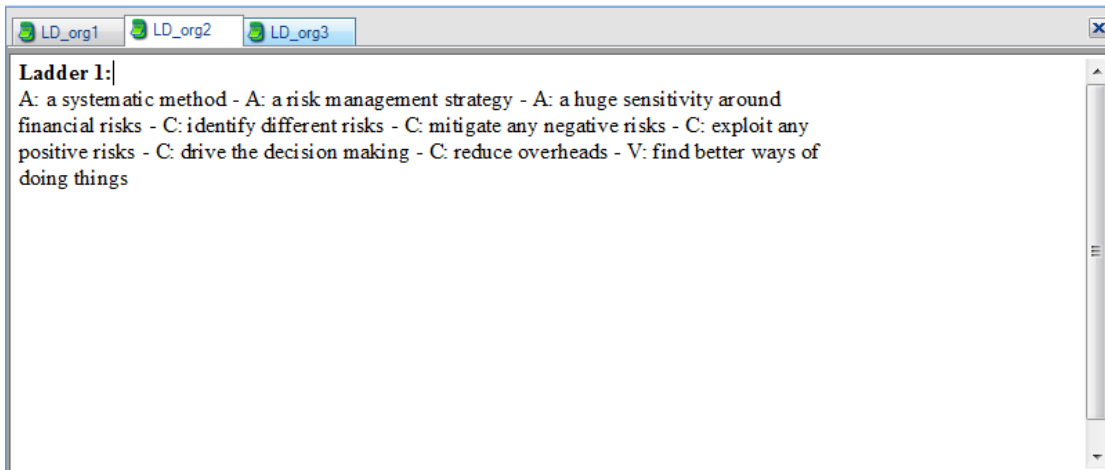
Fig. 3. Ladders for Organisation 1
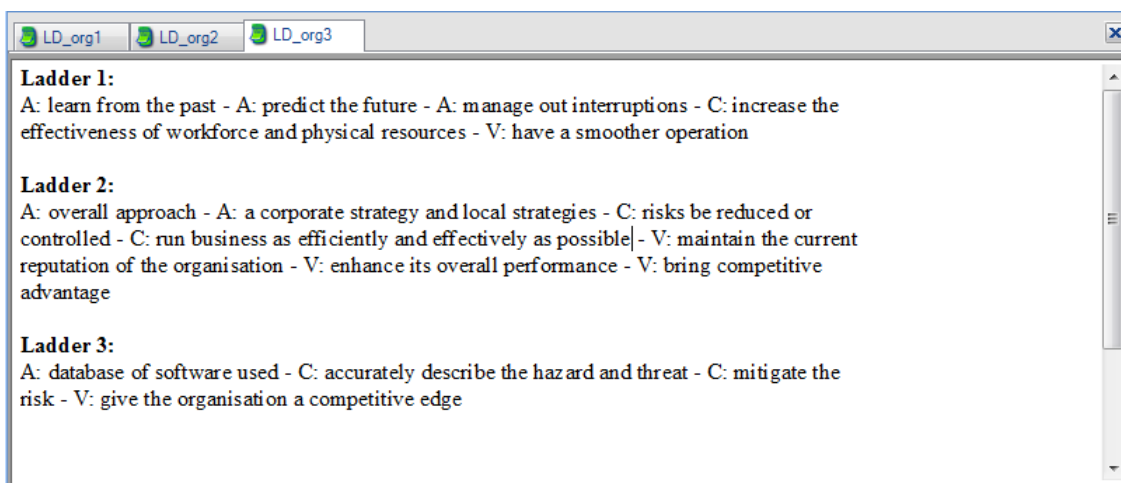
Fig. 4. Ladders for Organisation 2

Fig. 5. Ladders for Organisation 3

2. Content Analysis starts with recording the entire set of ladders across interviewees on a separate coding form, and then develops a set of summary codes for the (A-C-Vs) elements elicited. The focus of interest in content analysis is the relationships <u>between</u> the elements instead of the elements themselves. Therefore, the elements are classified in groups according to the closest in meaning. For example, "a whole systematic suite of things" is similar in meaning with "a systematic approach", so they are classified in the same group. Moreover, "manage risk effectively" is a summary of several more specific attributes such as "identify different risks", "mitigate any negative risks" and "exploit any positive risks". The summary content codes of attributes, consequences and values are shown respectively in Tables 3, 4 and 5:

No.	Attributes	Count
1	A systematic approach	10
2	A Risk Management strategy	8
3	Database	7
4	Awareness of risks	5
5	Predict the future	4
6	Learn from the past	2
7	Manage out interruptions	1

Table 3. Summary Content Codes of Attributes

352

No.	Consequences	Count
8	Manage risk effectively	20
9	Run business effectively	9
10	Manage cost	5
11	Drive decision making	5
12	Make individuals behave differently	1

Table 4. Summary Content Codes of Consequences

No.	Values	Count
13	Have a competitive advantage	4
14	Maintain the current reputation	2
15	Have a smoother operation	2
16	Find better ways of doing things	1

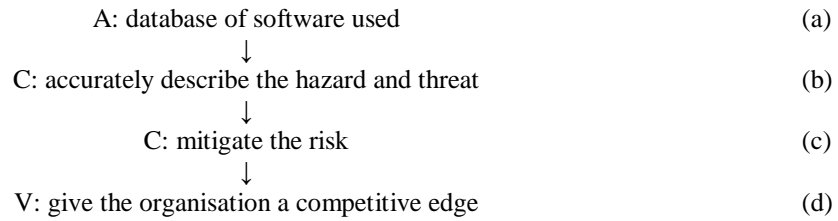Table 5. Summary Content Codes of Values

Numbers are assigned to the codes and used to label each element in each ladder, producing a matrix in Table 6. Where the original ladders have more than one attribute or value they have been separated into multiple sub-ladders – no direct relations existing between attributes and values. Thus 16 sub-ladders have been derived from the initial eight ladders. Ladders in the table have been expressed in fractional form showing for example, that sub-ladder 13 is derived from Ladder 2 of Organisation 3. "Content Codes" on the other hand represent the elements comprising each sub-ladder.

Sub-Ladder	Ladder	Content Codes				
1	1.1	3	10	13		
2	1.1	1	10	13		
3	1.2	3	8	15		
4	1.3	4	11	12	14	
5	1.4	1	8	9	14	
6	2.1	1	8	10	11	16
7	2.1	2	8	10	11	16
8	2.1	4	8	10	11	16
9	3.1	5	9	15		
10	3.1	6	9	15		
11	3.1	7	9	15		
12	3.2	1	8	9	13	
13	3.2	1	8	9	14	
14	3.2	2	8	9	13	
15	3.2	2	8	9	14	
16	3.3	3	8	13		

Table 6: Content Codes of Ladders

3. Implication Matrix. This displays the number of times one element leads to another. Operationally, it is defined as those elements in a given row precede other elements in the same row. Two kinds of

relations are included in the matrix, direct and indirect. Direct relations refer to those in which one element leads to another without any intervening element. Take Ladder 3 of Organisation 3 as an example:

<div align="center">

A: database of software used   (a)

↓

C: accurately describe the hazard and threat   (b)

↓

C: mitigate the risk   (c)

↓

V: give the organisation a competitive edge   (d)

</div>

In this example the relation a→b (database of software used → accurately describe the hazard and threat) is a direct relation, as is b→c and c→d. In addition, there are also indirect relations, a→c, a→d and b→d, which are helpful for exploring deeper understandings. It is critical to identify both kinds of relations in order to determine which paths are dominant in the hierarchy value map.

Elements		C					V			
		8	9	10	11	12	13	14	15	16
A	1	(4,0)	(0,3)	(1,1)	(0,1)		(0,2)	(0,2)		(0,1)
	2	(3,0)	(0,2)	(0,1)	(0,1)		(0,1)	(0,1)		(0,1)
	3	(2,0)		(1,0)			(0,2)		(0,1)	
	4	(1,0)		(0,1)	(1,1)	(0,1)		(0,1)	(0,1)	(0,1)
	5		(1,0)						(0,1)	
	6		(1,0)						(0,1)	
	7		(1,0)						(0,1)	
C	8		(5,0)	(3,0)	(0,3)		(1,2)	(0,3)	(0,1)	(0,3)
	9						(2,0)	(3,0)	(3,0)	
	10				(3,0)		(2,0)			(0,3)
	11					(1,0)		(0,1)		(3,0)
	12							(1,0)		

<div align="center">Table 7: Implication Matrix</div>

The implication matrix indicates the number of times directly or indirectly a row element leads to a column element. The numbers are expressed in the form (x, y), in which x represents <u>direct</u> relations whilst y represents <u>indirect</u> relations. For example, the first attribute "a systematic approach" (row 1) leads to "manage cost" (row 10) one time directly and one time indirectly. In the same way "a Risk Management strategy" (row 2) leads to "manage risk effectively" (row 8) two times directly and zero indirectly.

4. Hierarchy Value Map (HVM) follows from the implication matrix. The HVM reconstructs "chains" from the aggregate implication matrix to represent the linkages across different levels of abstractions, starting with attributes and ending with values. The term "chains" refers to sequences of elements that emerge from the implication matrix.

The most common approach of constructing an HVM is to set a "cut-off" - a minimum number of relations that must be present before that element is considered. According to Henneberg *et al* [20], the HVM should include all the connections which are composed of at least 4 or more direct relations. In this paper, however, the "cut-off" is only set to 2 because of the relatively small sample size (only 3 interviews and 16 sub-ladders) involved.

The most effective way to build the HVM is to start with the first row of the implication matrix for which there is a number equal to or larger than the chosen cut-off level. Beginning with the cut-off of 2, the first noteworthy value is the relationship of "a systematic approach → manage risk effectively" (Row 1, Column 8), with 4 direct relations and zero indirect relations between these two elements. Since "a

systematic approach" is related to "manage risk effectively", the next row to look at is row 8. "Run business effectively" in column 9 has the first significant value above the cut-off. Moving to row 9 it is easy to find that the final value in the growing chain is "have a competitive advantage" in column 13. Therefore, the chain of 1-8-9-13 is constructed.

Having reached the end of a chain, it is necessary to go back to the beginning to check whether there are other significant links in the same row of the matrix that were not inspected in the completed chain. For example, "a systematic approach" (1) is also linked to "manage cost" (10), with one direct relation and one indirect relation. Moving to row 10, the next value is "drive decision making" (11), and following that, the final relationship of "drive decision making → find better ways of doing things" (16) is found, producing another chain of 1-10-11-16. Repeating these steps, the emergent HVM is shown in Fig. 6.
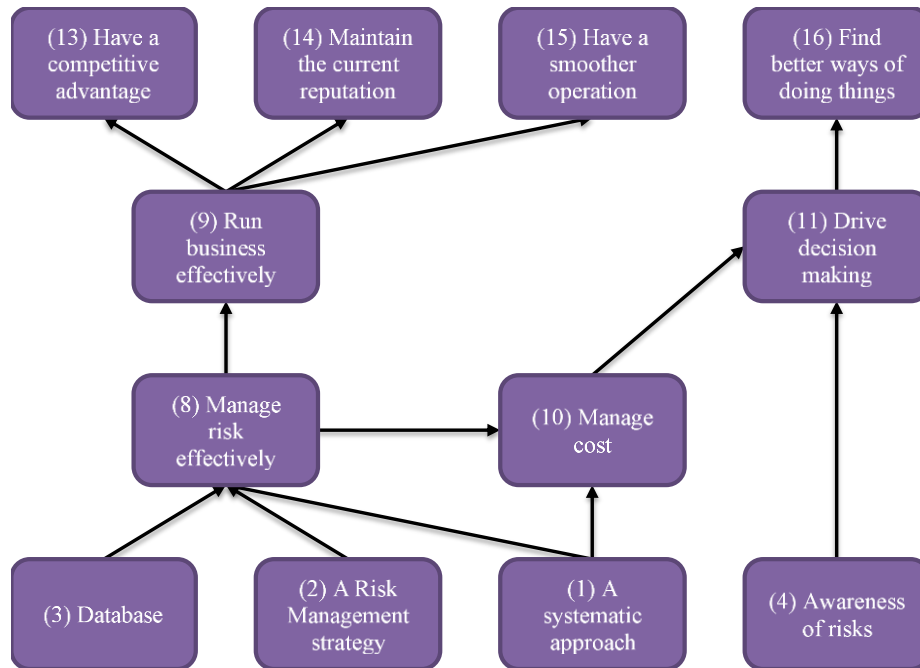

Fig. 6. Hierarchy Value Map of Risk Management

The objective of constructing the HVM is to interconnect all the meaningful chains in a map where relations are plotted without crossing lines. The HVM must include all relevant relations and be easy to read and understand. Clearly attributes 5, 6 and 7 (i.e. "predict the future", "learn from the past" and "manage out interruptions") do not appear in Fig. 6. From the implication matrix these attributes have only one direct relation to "run business effectively" (15) and one indirect relation to "have a smoother operation", so therefore do not merit inclusion. This does not necessarily mean that learning from the past, predicting the future and managing out interruptions are not important for an organisation's risk management. It merely signifies that the particular sample of interviews used for this analysis is not sufficiently supportive of these relations.

Once the HVM is constructed, it is important to consider the number of direct and indirect relations an element has with other elements. Table 8 presents the sums of direct and indirect relations for each element. "A systematic approach" (1) appears to be the most important attribute leading to other elements, whilst "manage risk effectively" (8), at the consequence level, leads to the most values. At the value level, "have a competitive advantage" (13), "maintain the current reputation" (14) and "find better ways of doing things" (16) have the largest sums of direct and indirect relations, which indicates that they would be judged to be the core values for the three interviewing organisations.

Code	To	From
1	(5,10)	(0,0)
2	(3,7)	(0,0)
3	(3,3)	(0,0)
4	(2,4)	(0,0)
5	(1,1)	(0,0)
6	(1,1)	(0,0)
7	(1,1)	(0,0)
8	(9,12)	(10,0)
9	(8,0)	(8,5)
10	(5,3)	(5,3)
11	(4,1)	(4,6)
12	(1,0)	(1,1)
13	(0,0)	(5,7)
14	(0,0)	(4,8)
15	(0,0)	(3,5)
16	(0,0)	(3,9)

Table 8: Summary of Direct and Indirect Relations for Each Element

The element with the highest frequency of elements leading from it is "run business effectively" (9). In fact, the chain of "a systematic approach → manage risk effectively → run business effectively → maintain the current reputation" appears to have the highest number of relations among its respective elements. Therefore, it is considered the dominant perceptive pathway. While this dominant pathway provides a genuine insight into why and how the organisations implement risk management, all the pathways in the HVM deserve our attention because the weaker pathways may highlight where these ties need to be strengthened.

3 Conclusions

A sample of interview transcripts obtained during a risk management survey of UK business was subjected to a thematic analysis using the NVivo software. The interview transcripts - each comprising approximately 2900 words – yielded a total of 40 nodes, from which four main themes were eventually identified viz. Concepts of Risk, Risk Management, Roles in Risk Management and Data Encryption.

Building on this encouraging start, a laddering analysis was undertaken next to investigate the impact of organisational values on risk management. The resulting hierarchical value map (HVM) enabled managers' core values elicited via the laddering technique e.g. "have a competitive advantage", "maintain the current reputation" and "find better ways of doing things", to be systematically traced back to specific Risk Management attributes. One 'perceptive pathway' was found to be especially dominant:
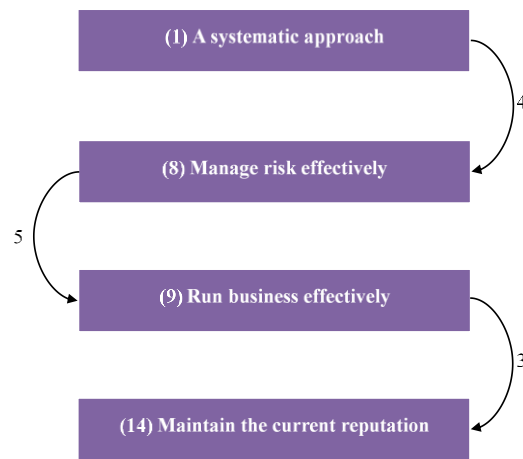
Fig. 7 Dominant Perceptive Pathway

The analysis has provided no less valuable insights for the individual respondents: for Organisation 1, for example, a clear priority is to develop a database-driven risk management strategy or face the prospect of losing its competitive advantage and current reputation.

Organisation 2 on the other hand, already demonstrates good risk management practice: risk assessments are carried out internally in every department, and staff encouraged to take risks with support as necessary from their senior managers. To improve things further the organisation could consider institutionalising a web-based, tailor-made centralised system.

Organisation 3 is similarly performing well: they have a risk register and a risk management committee to manage risk effectively. As the leading organisation in the medical science market, Organisation 3 should perhaps look at further embedding risk management within its culture.

References

1.   D. Cooper, and C. B. Chapman. Risk analysis for large projects: models, methods and cases. New York: Wiley, 1987.

2.   The Institute of Risk Management (IRM). 2013. Available at: http://www.theirm.org/ (Accessed: 16th August 2013).

3.   International Standards Organization (ISO) Risk Management - Principles and Guidelines. ISO Guide 31000, Geneva: International Standards Organization, 2009.

4.   Joint Information Systems Committee (JISC). (JISCinfoNet Risk Management), 2012. Available at: http://www.jiscinfonet.ac.uk/infokits/risk-management/ (Accessed: 19th July 2013).

5.   N. Mays and C. Pope (Eds.). Qualitative research in health care. London: BMJ, 1996.

6.   M. B. Miles, and A. M. Huberman. Qualitative data analysis: An expanded sourcebook. SAGE Publications, Inc., 1994.

7.   D. J. Banner and J. W. Albarrran. Computer-assisted qualitative data analysis software: a review, Canadian Journal of Cardiovascular Nursing, 19, 3, 24{31, 2009.

8.   J. W. Creswell. Qualitative inquiry and research design: Choosing among five approaches. 3rd edn. SAGE Publications Inc., 2012.

9.   J. W. Creswell. Research design: Qualitative, quantitative, and mixed methods approaches. 4th edn. SAGE Publications, Inc., 2013.

10. A. Lewins and C. Silver. Choosing a CAQDAS package 6th edn Qualitative Innovations in CAQDAS, 2009.

11.   E. A. Weitzman and M. B. Miles. Computer programs for qualitative data analysis. SAGE Publications, Inc., 1995.

12. C. A. Barry. Choosing qualitative data analysis software: Atlas/ti and Nudist compared, Sociological Research Online, **3** 3, http://www.socresonline.org.uk/socresonline/3/3/4.html [Accessed: 7th July 2013]

13. P. Bazeley, Qualitative Data Analysis with NVivo. SAGE Publications Inc., 2007.

13. A. Herrmann, F. Huber & C. Braunstein. Market-driven product and service design: Bridging the gap between customer needs, quality management, and customer satisfaction, International Journal of Production Economics, 66, 1, 77{96, 2000.

14. S. A. Malik, B. Holt, J. Freeman. Enterprise Risk Management (ERM): A New Way of Looking at Risk Management at an Organisational Level. 2013. In EUROXXVI 01 July 2013 - 04 July 2013. Rome, Italy.

15. V. Braun and V Clarke. 'Using thematic analysis in psychology', Qualitative research in psychology, **3**, 2, 77{101, 2006.

17. A. Saaka, C. Sidon, and B. F. Blake. Laddering: a "how to do it" manual – with a note of caution, Research Reports in Consumer Behaviour, Cleveland State University, 2004.

18. H. H. Kassarjian. Content analysis in consumer research, Journal of Consumer Research, **4**, 8{18, 1977.

19. T. J. Reynolds, and J. Gutman. 'Laddering theory, method, analysis, and interpretation', Journal of advertising research, **28**, 1, 11{31, 1988.

20. S. C., Henneberg, T., Gruber, A., Reppel, B. Ashnai, and P Naudé. 'Complaint management expectations: An online laddering analysis of small versus large firms', Industrial Marketing Management, **38**, 6, 584{598, 2009.

Capture-recapture Type model for Estimating Elusive Events with Tag Loss

Jibasen, Danjuma

Department of Statistics and Operations Research
Modibbo Adama University of Technology, Yola, Nigeria.
djibasen@mautech.edu.ng djibasen1@yahoo.co.uk

Abstract

In capture-recapture experiments, if individuals lost their tags, the observed recaptures will be smaller than expected. This phenomenon results in overestimation of the population under consideration. Drugs addicts are usually referred for treatment or for rehabilitation, on the process of these, they are likely to change their identity hence losing their tags. Tags loss method was therefore incorporated in the estimation of the size of elusive population. The simulation studies revealed that the proposed coverage probability tag loss model (CPTLM) is consistent with small and large population sizes. The proposed model was applied to addicts' data collected from North east, Nigeria.

Key words

Capture-recapture, tag-loss, overestimation, drug addicts

Introduction

Capture-recapture (C-R) methods were originally applied to animal populations in which sequence of samples were taken from a well-defined population; any animal found without a tag in a particular sample were given a unique tag before returning that sample to the population. In that way, estimate of the population size and other relevant parameters are obtained. These methods have now been applied extensively to epidemiological and public health events with the aim of estimating the incidence and prevalence of such events (Seber *et al*., 2000). The technique has also been adopted for other areas such as; the evaluation of census undercount (Ericksen and Kandame 1985 and Darroch *et al*., 1993), software testing and reliability (Wohlin *et al*., 1995; Ebrahimi 1997; Briand *et al*., 2000, Yip *et al*., 2003), to mention a few. When there are only two samples, the method is called the Petersen method (or the Lincoln index).

In this work, we are concerned with estimating the number of drugs addicts within a given location; a case study of North eastern states of Nigeria, consisting of six states namely; Adamawa, Bauchi, Borno, Gombe, Taraba and Yobe States. The primary sources of data on drugs addicts in these States are from the National Drug Law Enforcement Agency (NDLEA) State command headquarters and the Psychiatric centres of the Specialists Hospital in these States. Patients are taking to the Centres for treatment and are also to the NDLEA for rehabilitation. NDLEA also makes direct arrest of barons and drugs users.

In practice, C-R methods can be applied to any situation in which records of individuals from the same population are kept in at least two different but incomplete lists. Thus "being on list i" can be equated to "being on sample i". The problem is to estimate those missing from both lists. These lists can come from different units or departments of the same agency (e.g. Doctors' and Pharmacists' record), or different agencies (e.g. The Police Force and the Prison Services' records). When applied to list, the Petersen method is known by the nomenclature; Dual System Methods (DSM); Dual System Estimation (DSE) or Dual Record Systems (DRS), IWGDMF (1995); El-khorazaty, *et al*., (1976) and Ericksen and Kadane (1985).

The assumptions required for this estimate to be valid can be spelt out in a number of ways. However the key ingredients are: (1) the population is closed, that is, the population has a

constant size for the entire period of the study, (2) the lists are independent, (3) each member of the population has the same chance of being on a given list, and (4) individuals are matched correctly, that is, individuals will not change their identity, in the terminology of C-R, no tags loss. Assumption (1) holds if the experiment is conducted within a reasonably short period of time. For (2) the listing systems may not be independent, since addicts can be referred across systems for rehabilitation or treatment; the NDLEA usually refer addicts to Psychiatric centres for treatment, likewise, Psychiatric centres can also refer psychoactive patients to the NDLEA for rehabilitation. We assumed that addicts have similar behaviours, hence assumption (3) holds, that is, addicts have the same probability of being on a given list.

Assumption (4) will completely be false; matching will depend on the quality of records, the truthfulness of the information and the uniqueness of the tags used. Addicted individuals are likely to give false information about their identity deliberately to avoid stigmatization or arrest, or even unconsciously under the effect of intoxicant. This leads to tag loss. According to Pollock (1991), the loss or overlooked of marks (tags) can be serious, he suggests that one way to estimate tag loss is to use double marks. Pollock *et al*., (1990) stated that, if tag loss is likely to occur, an attempt should be made to estimate rate of loss and that if individuals lose their tags, N will be overestimated; this situation is referred to as positive biased (IWGFDMF,1995).

The paper is organized as follows: the tag loss method of Seber, (1982) is presented in section 2, followed by the coverage probability model (CPM) for estimating elusive events of Jibasen (2011) incorporating tag loss. In section (3), the incorporated tag loss method called, coverage probability tag loss model (CPTLM) was applied to a set of simulated data and compared to the Petersen method and the CPM. Finally, in section 4, the proposed CPTLM was applied to data set obtained from the NDLEA and Psychiatric centres in the north eastern states of Nigeria.

2. Methodology

The Petersen estimator is well-known, if we assume the proportion on list 1, n_1/N for the whole population is roughly the same as it is on list 2, m_2/n_2, then,

$$n_1/N = m_2/n_2$$

and, solving for N yields the Petersen estimator:

$$\bar{N} = \frac{n_1 n_2}{m_2}$$

(1)

where, m_2, n_1 and n_2 are the numbers of individuals on both lists, on list 1 and on list 2, respectively.

According to IWGDMF (1995), if we exploit the mathematical statistics result, (1) can be written as;

$$
\left.
\begin{aligned}
E(\bar{N}) &= \frac{E(n_1)E(n_2)}{E(m_2)} \\
&= \frac{Np_1 p_2}{p_{12}} \\
&= NR
\end{aligned}
\right\}
$$

(2)

Where, R > 1, if being on list 1 tends to decrease the chance of being on list 2, N will be overestimated, this is negative dependence (or positive biased) which can be credited to the fact that, individuals 'elusified' by losing their tags. It is therefore evident that, with elusive events, tags loss is therefore apparent which leads to negative dependence between lists. In epidemiological studies, tag loss has received little attention (Seber *et al*., 2000).

This paper assumed that addicts are referred across agencies and that on reference, they are likely to change their information identification making it difficult for matching. This will thus imply that being on the first list will decrease the chance of being on the second, thus $\bar{N}$ will lead to over estimation of N.

Addicts on a list are considered as having an identifying string of information, these are; first name, surname, age, religion, address and type of substance abused. These information were group into two, forming two tags. Tag A consists of name, age, and religion; tag B consists of individuals' address and type of substance abuse.

2.1 Tag Loss Method of Seber (1982)

Each individual on a list has a string of identifying information subdivided into tag A and tag B. Tag A consists of name, age, and religion; these are items we assume individuals are

likely to be truthful about. Tag B consist of address and type of substance abused; these are items addicts are likely to lie about. If either substring is correct the individual is identified uniquely. We assume further that these tags are independent. This assumptions and assertion are in line with Seber *et al.,* (2000).

Let,

π_x = the probability that a tag x is lost on the second list (x = A, B)

π_{AB} = the probability that both tags are lost

m_x = number of tagged individuals on the second list, with tag x only (x = A, B)

m_{AB} = number of tagged individuals on the second list with both tags

m_2 = those on both lists.

As earlier stated that, the tags are assumed to be independent, that is, $\pi_{AB} = \pi_A \pi_B$, according to Seber (1982), the joint probability function of m_A, m_B, m_{AB} and m_2 is given by

$$f(m_A, m_B, m_{AB}, m_2 \mid \llbracket n \rrbracket_1, n \rrbracket_2) = f(m_A, m_B, m_{AB} \mid m_2) f(m_2 \mid \llbracket n_1, n \rrbracket_2)$$

(3)

where

$$f(m_A, m_B, m_{AB} \mid m_2) = (m_2!)/(m_A \llbracket!m\rrbracket_B! m_{AB}! m_0!)[(1 - \pi_A) \pi_B]^T(m_A) [\pi_A (1 - \pi_B$$

$$\times \left[(1 - \pi_A)(1 - \pi_B)\right]^{m_{AB}} [\pi_A \pi_A]^{m_0}$$

$m_0 = m_2 - m_A - m_B - m_{AB}$,

and

$$f(m_2 \mid \llbracket n_1 1, n \rrbracket_2) = \frac{\binom{n_1}{m_2}\binom{N - n_1}{n_2 - m_2}}{\frac{N}{n_2}}$$

While, maximum-likelihood estimates of *N*, m_2, π_A and π_B are given by

$$\widehat{N} = \frac{N_1 N_2}{\widehat{m}_2}$$

$$m_A = \widehat{m}_2 \left(1 - \widehat{\pi}_A\right)\widehat{\pi}_B$$

$$m_B = \widehat{m}_2 \left(1 - \widehat{\pi}_B\right)\widehat{\pi}_A$$

$$m_{AB} = \widehat{m}_2 \left(1 - \widehat{\pi}_A\right)\left(1 - \widehat{\pi}_B\right)$$

With the solutions,

$$\widehat{\pi}_A = \frac{m_B}{m_B + m_{AB}}$$

$$\widehat{\pi}_B = \frac{m_A}{m_A + m_{AB}}$$

$$\widehat{m}_2 = \frac{[(m]_A + m_{AB})(m_B + m_{AB})}{m_{AB}} \tag{4}$$

Provided $m_x \neq 0$ for at least one x. If at least one $m_x = 0$, the model collapsed to the classical model.

2.2 M_c: Coverage probability model (CPM)

Huggins (1991) used a form of a Hurwitz- Thompson (H-T) method to model heterogeneity of individual animals, where animals were assigned probabilities, following this idea, Jibasen (2011) introduced a model for estimating elusive events for two lists (called, coverage probability model (CPM)) based on the H-T method; this model was discovered to give better results than the Petersen method, in the presence of low recaptures. The model is presented below incorporating the method of estimating tag loss (4).

The joint probability density function for the coverage probability model is given as;

$$L = f(n_1, n_2, n_{11}) = \binom{r}{n_1}\binom{r - n_1}{n_2 - n_{11}}\binom{n_1}{n_{11}}p^{n_1 + n_2}(1 - p)^{sr - n_1 - n_2} \tag{5}$$

(see Jibasen, 2011 and Jibasen., *et al.,* 2012)

The ML estimator of p is $\widehat{p} = \frac{n}{sr}$

The two sample estimator for estimating elusive events is give in (6) as;

$$\tilde{N}_c = \frac{r}{p} = \frac{sr^2}{n}$$

(6)

where, $r = n_1 + n_2 - \hat{m}_2$

where, n_{11} has been replaced by, $\hat{m}_2$.

3. Simulation Studies

The simulated data was thus based on the hypergeometric settings, the marginal totals (n_1 and n_2) were fixed as well as the assumed population size N, while the recaptures n_{11} were randomly generated. The variates: m_A, m_B and m_{AB} were simulated from n_{11}, from where $\hat{m}_2$ was estimated to replace n_{11}. Series of simulation was carried out for $n_1 = 50$, $n_2 = 10$ and the assumed population size $N = 90$, as showed in Tables 1 to 5. Simulation for other values of n_1, n_2 and N are summarized in Tables 6 to 7. The Petersen estimator $\tilde{N}$, the coverage probability model (CPM) estimator $\tilde{N}_c$, and the coverage probability tag loss model (CPTLM) estimator $\tilde{N}_{m2}$, were compared using the Akaike information criterion (AIC), while mean absolute deviation (MAD) was used to checked for the overall model performance for each set of simulated data. All simulated data were from Jibasen (2011).

Table 1 shows, no tag loss; $n_{11} = m_A + m_B + m_{AB}$ the analyses thus showed that the CPTLM performs well compared with the CPM, but both performs better compared to the Petersen method. This reveals that the CPTLM performs well even when there is no tag loss.

Table1: Simulated data for $n_1 = 50$, $n_2 = 10$ and $N = 90$

S/No.	n_{11}	M_A	M_B	M_{AB}	$\hat{m}_2$	$\tilde{N}$	AIC ($\tilde{N}$)	$\tilde{N}_c$	AIC ($\tilde{N}_c$)	$\tilde{N}_{m2}$	AIC ($\tilde{N}_{m2}$)
1	7	1	3	3	8	63	4.366	94	4.272	90	4.010
2	7	2	1	4	8	67	4.366	94	4.272	92	4.140
3	6	2	1	3	7	75	3.490	97	4.543	95	4.361
4	8	2	1	5	8	60	5.112	90	4.010	89	4.092
5	8	2	3	3	10	50	5.112	90	4.010	83	4.490
6	8	1	2	5	8	60	5.112	90	4.010	89	4.092
7	10	3	3	4	12	41	-	83	4.490	76	5.025
8	7	1	4	2	9	56	4.366	94	4.272	87	4.244

S/No.	n_{11}	M_A	M_B	M_{AB}	$\overline{M}_2$	$\widehat{N}$	AIC $(\widehat{N})$	$\widehat{N}_c$	AIC $(\widehat{N}_c)$	$\widehat{N}_{m2}$	AIC $(\widehat{N}_{m2})$
9	8	2	2	4	9	56	5.112	90	4.010	87	4.244
10	5	0	1	4	5	100	3.759	101	4.824	101	4.824
MAD						29		4		5	

Tables 2 shows a situation where one individual losses one tag. The result from Table 2 shows that if an individual loss both tags the CPTLM performs better than both the Petersen and the CPM. But if up to two individuals lost both tags the CPM performs better.

Table 2: Simulated data for n_1= 50, n_2 =10 and $N =$ 90 with one tag loss

S/No.	n_{11}	M_A	M_B	M_{AB}	$\overline{M}_2$	$\widehat{N}$	AIC $(\widehat{N})$	$\widehat{N}_c$	AIC $(\widehat{N}_c)$	$\widehat{N}_{m2}$	AIC $(\widehat{N}_{m2})$
1	7	1	2	3	7	63	4.366	94	4.272	95	4.361
2	7	1	1	4	6	67	4.366	94	4.272	96	4.474
3	6	2	1	3	7	75	3.490	97	4.543	95	4.361
4	8	2	1	4	8	60	5.112	90	4.010	92	4.140
5	8	1	2	4	8	50	5.112	90	4.010	92	4.140
6	8	1	1	5	7	60	5.112	90	4.010	93	4.219
7	10	1	2	6	9	41	-	83	4.490	86	4.327
8	7	1	2	3	7	56	4.366	94	4.272	95	4.361
9	8	2	1	4	8	56	5.112	90	4.010	92	4.140
10	5	1	1	2	5	100	3.759	101	4.824	103	4.968
MAD						29		4		5	

When individuals loss 2 tags, one from each string, Table 3 shows that the CPM is better model.

Table 3: Simulated data for n_1= 50, n_2 =10 and $N =$ 90 with two tag loss

S/No.	n_{11}	M_A	M_B	M_{AB}	$\overline{M}_2$	$\widehat{N}$	AIC $(\widehat{N})$	$\widehat{N}_c$	AIC $(\widehat{N}_c)$	$\widehat{N}_{m2}$	AIC $(\widehat{N}_{m2})$
1	7	1	2	2	6	63	4.366	94	4.272	97	4.543
2	7	1	1	3	5	67	4.366	94	4.272	100	4.729
3	6	2	0	3	5	75	3.490	97	4.543	101	4.824
4	8	2	1	3	7	60	5.112	90	4.010	95	4.361
5	8	1	2	3	7	50	5.112	90	4.010	95	4.361
6	8	0	1	5	6	60	5.112	90	4.010	97	4.543
7	10	1	2	5	8	41	-	83	4.490	89	4.092
8	7	0	2	3	5	56	4.366	94	4.272	101	4.824
9	8	2	1	3	7	56	5.112	90	4.010	95	4.361

							AIC		AIC		AIC
10	5	1	1	1	4	100	3.759	101	4.824	105	5.115
MAD						29		4		8	

Table 4 shows tag loss by substring A, while Table 5 shows tag loss by substring B. Each shows that the CPTLM performs well but the CPM is better.

Table 4: Simulated data for $n_1 = 50$, $n_2 = 10$ and $N = 90$ with one tag loss from substring A.

S/No.	n_{11}	M_A	M_B	M_{AB}	$\overline{M}_2$	$\widehat{N}$	AIC $(\widehat{N})$	$\widehat{N}_c$	AIC $(\widehat{N}_c)$	$\widehat{N}_{m2}$	AIC $(\widehat{N}_{m2})$
1	7	0	3	3	6	63	4.366	94	4.272	95	4.361
2	7	1	1	4	6	67	4.366	94	4.272	97	4.543
3	6	1	1	3	5	75	3.490	97	4.543	101	4.824
4	8	1	1	5	7	60	5.112	90	4.010	94	4.272
5	8	1	3	3	8	50	5.112	90	4.010	89	4.076
6	8	0	2	5	7	60	5.112	90	4.010	93	4.219
7	10	2	3	4	11	41	-	83	4.490	82	4.611
8	7	0	4	2	6	56	4.366	94	4.272	92	4.140
9	8	1	2	4	8	56	5.112	90	4.010	92	4.140
10	5	0	1	4	5	100	3.759	101	4.824	105	5.115
MAD						29		4		6	

Table 5: Simulated data for $n_1 = 50$, $n_2 = 10$ and $N = 90$ with one tag loss from substring B.

S/No.	n_{11}	M_A	M_B	M_{AB}	$\overline{M}_2$	$\widehat{N}$	AIC $(\widehat{N})$	$\widehat{N}_c$	AIC $(\widehat{N}_c)$	$\widehat{N}_{m2}$	AIC $(\widehat{N}_{m2})$
1	7	1	2	3	7	63	4.366	94	4.272	95	4.361
2	7	2	0	4	6	67	4.366	94	4.272	97	4.543
3	6	2	0	3	5	75	3.490	97	4.543	101	4.824
4	8	2	0	5	7	60	5.112	90	4.010	94	4.272
5	8	2	2	3	8	50	5.112	90	4.010	89	4.076
6	8	1	1	5	7	60	5.112	90	4.010	93	4.219
7	10	3	2	4	11	41	–	83	4.490	82	4.611
8	7	1	3	2	8	56	4.366	94	4.272	92	4.140
9	8	2	1	4	8	56	5.112	90	4.010	92	4.140

10	5	0	0	4	4	100	3.759	101	4.824	105	5.115
MAD						29		4		6	

For lack of space, simulation for various values of n_1, n_2 and N are summarized in Tables 6 and 7. The simulated results in Table 6 revealed that the CPTLM performs well when the elusiveness is high, that is, when more individuals 'elusifies'. The results further show that, even when all the recaptures are lost CPTLM performs better than the Petersen but not the CPM.

Table 6: Simulated Results for Various values of n_1, n_2 and N

N	n_1	n_2	n_{11}	$\widehat{N}_{m2}$ with 3 loss	$\widehat{N}_{m2}$ with 10 loss	$\widehat{N}_{m2}$ with10 loss	$\widehat{N}_c$	$\widehat{N}$
200	100	10	10	193	220	200	182	100
	100	10	8	200	220	208	189	125
	100	10	6	208	220	216	197	167
	MAD			5	20	8	11	69
				$\widehat{N}_{m2}$ with 10 loss	$\widehat{N}_{m2}$ with 20 loss	$\widehat{N}_{m2}$ With 25 loss		
200	100	40	24	227	264	284	192	167
	100	40	25	223	260	280	189	160
	100	40	27	216	253	272	182	148
	MAD			22	59	79	12	42
				$\widehat{N}_{m2}$ with 25 loss	$\widehat{N}_{m2}$ with 30 loss	$\widehat{N}_{m2}$ With 40 loss		
200	90	70	46	242	259	177	162	137
	90	70	43	252	270	186	171	147
	90	70	42	256	274	189	174	150
	MAD			50	68	16	31	56

When the population size is large, Table 7 shows that CPTLM performs better with smaller number of the tag loss compare to the CPM and far better than the classical method, the Petersen.

Table 7: Simulated Results for Various values of n_1, n_2 with large values of N

N	n_1	n_2	n_{11}	$\widehat{N}_{m2}$ with 1 loss	$\widehat{N}_{m2}$ with 2 loss	$\widehat{N}_{m2}$ With 5 loss	$\widehat{N}_c$	$\widehat{N}$
300	150	10	7	296	300	312	293	214
	150	10	6	300	304	316	296	250
	150	10	8	293	296	308	289	188
MAD				4	3	12	7	83
1000	500	80	57	947	950	961	943	702
	500	80	51	969	972	983	965	784
	500	80	49	976	980	991	972	816
MAD				36	33	22	40	233
2000	900	80	49	1773	1777	1788	1769	1469
	900	80	48	1777	1780	1792	1773	1500
	900	80	57	1742	1746	1758	1739	1263
MAD				236	232	221	240	589

N	n_1	n_2	n_{11}	$\widehat{N}_{m2}$ with 100 loss	$\widehat{N}_{m2}$ with 150 loss	$\widehat{N}_{m2}$ With 5 loss	$\widehat{N}_c$	$\widehat{N}$
3000	1500	300	195	3230	3422	2880	2862	2308
	1500	300	207	3185	3376	2837	2820	2174
	1500	300	194	3234	3426	2884	2866	2320
MAD				216	408	133	151	733

4. Estimation of the Number of Addicts from North east Nigeria Using CPTLM

369

The CPTLM was used to estimate the population of addicts in the north east alongside the CPM. A 95% Confidence intervals based on the suggestion of Chao(1989) was constructed for $\widehat{N}_{mz}$ with 5 tag loss, as shown in Table 9, the interval estimates shows that all the estimates of $\widehat{N}_{mz}$ with 1 tag loss, 2 tag loss and for $\widehat{N}_c$ falls within the acceptance region.

Table 9 Estimated populations of addicts by States Using CPTLM

State	Year	$\widehat{N}_{mz}$ with 1 loss	$\widehat{N}_{mz}$ with 2 loss	$\widehat{N}_{mz}$ with 5 loss	LCL $\widehat{N}_{mz}$ with 5 loss	UCL $\widehat{N}_{mz}$ with 5 loss	$\widehat{N}_c$
Adamawa	2006	239	243	244	150.0758	694.0903	235
	2007	172	176	178	99.56733	679.4685	168
	2006	170	174	176	101.8159	622.2887	166
	2007	206	210	212	134.5086	566.4152	203
	2008	283	287	289	188.6322	649.4115	280
Taraba	2006	197	200	202	122.1806	617.8411	193
	2007	179	183	185	111.1938	590.6701	175
	2006	224	228	230	146.0445	603.8456	221
	2007	218	222	224	132.0206	700.2673	214
	2008	233	237	239	149.9038	633.1515	229
Gombe	2006	386	390	392	250.1097	921.735	382
	2007	413	417	419	272.5436	928.3745	409
	2006	357	360	362	230.3457	870.088	353
	2007	320	324	326	216.1508	715.0046	317
	2008	313	317	319	213.2949	685.7079	309
Bauchi	2006	364	367	369	265.3076	688.6959	360
	2006	224	227	229	159.4315	486.0392	220
	2007	207	210	212	152.8454	409.0217	204
	2006	208	211	213	172.1324	323.4389	206
	2007	251	253	255	205.3606	388.0259	248
	2008	231	233	234	203.7887	303.1091	228
Borno	2006	553	557	558	386.1228	1056.389	549
	2007	593	597	599	432.5211	1032.522	590

2006	678	682	684	477.3404	1240.633	674
2007	506	510	512	366.2576	901.735	503
2008	536	540	542	375.8013	1003.603	532
Total	8260	8354	8401	5715.297	18311.57	8167

Source: This data was originally presented by Jibasen (2011)

Discussion and Conclusion.

Tag loss is inevitable when dealing with elusive events. The simulation studies revealed that the CPTLM performs competitively alongside the CPM, and that the CPTLM performs better when the population size is large, whereas CPM performs better with smaller population sizes. It was also discovered that CPTLM performs well when there is no tag loss and as the number of loss tags reduces CPTLM performs better. The proposed tag loss model was applied to addicts' data, where a 95% confidence interval shows that all the estimates fall within the acceptance region.

Remark.

The robustness of CPM was established by Jibasen (2011), this is really, the first attempt at improving (or even disapproving) on the performance of CPM. With this work, it has been established that CPM is a robust method of estimating elusive events from two sources, even when tag loss is inevitable.

Reference

Briand , L. C, El-Emam , K., Freimut ,B.G. and Laitenberger, O. (2000). A Comprehensive Evaluation of Capture-Recapture Models for Estimating Software Defect Content (2001) IEEE Transactions on Software Engineering. **26**, 6, 518-540

Chao, A.(1989). Estimating size of population size for sparse data in capture-recapture experiments. *Biometrics* **45**,427-438.

Darroch J. N., Fienberg S. E., Glonek G. F. V., Junker, B. W. (1993). "A Three-Sample Multiple-Recapture Approach to Census Population Estimation with Heterogeneous Catchability."*Journal of the American Statistical Assoc.*, 88, 1137–1148.

Ebrahimi N. B. (1997). "On the Statistical Analysis of the Number of Errors Remaining in Software Design Document After Inspection." IEEE Transactions on Software Engineering, **23**, 529–532.

El-Khorazaty, M. N., Imrey, P. B., Koch, G. G and Wells, H. B. (1976). A Review Of

Methodological Strategies For Estimating The Total Number Of Events With Data From Multiple-Record. University of North Carolina Institute of Statistics Mimeo Series No. 1095. Chapel Hill, N.C.

Huggins, R. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*. **47**, 725-732.

International Working Group for Disease Monitoring and Forecasting (IWGDMF) (1995). Capture-Recapture and
Multiple-Record Systems Estimation I: History and Theoretical Development. *Am. J. Epidemiol*.
142,1047-58.

Jibasen, D (2011). Capture-recapture Type Models for Estimating the Size of An Elusive Population. A PhD Thesis submitted to the University of Ilorin, Ilorin, Nigeria.

Jibasen, D. Yahya, W. B. and Jolayemi, E. T.(2012). Capture-Recapture Estimation for Elusive events with Two Lists. *Mathematical Theory and Modeling* www.iiste.org **2** (8),1-6

Ericksen, E.P. and Kadame, J. B. (1985). Estimating the population in a census year. *Journal of the American Statistical Assoc.* **80**, 98-109.

Pollock, K.H., Nichols, J.D., Brownie, C. and Hines, J.E. (1990). Statistical inference from capture-recapture experiment. *Wildlife Monographs* **107**, 1-97

Pollock, K. H. (1991). Modelling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present and future. *Journal of the American Statistical Association* **86**, 225-238

Seber, G. A. F. (1982). *Estimating Animal Abundance and related Parameters*, 2nd edition. London: Charles Griffin & Sons.

Seber, G. A. F., Huakau, J. T., and Simmons, D. (2000). Capture-recapture, Epidemiology and Lists Mismatches: Two lists. *Biometrics*, **56** 1227- 1232

Wohlin C, Runeson P, Brantestam B (1995). "An Experimental Evaluation of Capture-Recapture in Software Inspection." *Software Testing, Verification and Reliability*, **5**, 213–232.

Yip, P. S. F., Wang, Y. and Chao, A. (2003). The Application of Capture-recapture Methods in Software Reliability. In H. Pham (ed.) *Handbook of Reliability Engineering*, Springer, U. K 493-510.

Optimal Inspection Policy for Age-dependent Markov Deteriorating System

Lu Jin[1] and Kouta Kashiwai[2]

[1] Graduate School of Informatics and Engineering, University of Electro-Communications, Tokyo, Japan
(E-mail: `jinl@inf.uec.ac.jp`)
[2] Graduate School of Information Systems, University of Electro-Communications, Tokyo, Japan

Abstract. On-condition maintenance, by which a system is inspected periodically, is widely used. The deterioration state of the system can be determined only by a costly inspection. On the basis of the inspection results, one of two actions, keep operating or replace, is selected. If keep operating is selected, the timing of the next inspection is determined, and the system continues to be operated until the next inspection. In most previous research, inspections are carried out at regular intervals irrespective of the system state. This research provides a flexible inspection policy in which the interval between the current and next inspection is determined on the basis of the current information about the system. The focus is on an age-dependent deteriorating system in which the deterioration process can be formulated as a non-stationary Markov process. The optimal decision-making problem is investigated, and the structural properties of the resulting optimal expected cost function are obtained. These properties establish the existence of an optimal decision policy in which the action changes from keep operating to replace at most once. The properties of the inspection interval when keep operating is selected were investigated, and the optimal inspection interval was found to decrease with respect to both the system's deterioration and age under certain assumptions. Since the action keep operating in this research is accompanied by a flexible inspection interval, the keep-operating action is called *inspection*.

Keywords: Monotone policy, Non-stationary Markov decision process, On-condition maintenance, Optimal decision-making, Totally positive of order 2.

1 Introduction

1.1 Background

Systems that deteriorate over time are generally subject to stochastic breakdowns, and breakdowns can impose societal costs. On-condition maintenance is an effective way to avoid such breakdowns. In on-condition maintenance, systems are inspected periodically, and the results of the inspection are compared with pre-specified criteria to determine whether potential failures or problems exist. If necessary, an appropriate countermeasure such as replacement is taken before problems occur.

On-condition maintenance is used widely in various industries. For example, the wireless devices now being installed in airplanes and ships are subject to an on-

---

condition maintenance policy. The Japanese Ministry of Internal Affairs and Communications requires that such devices be inspected annually [10]. However, a report by a wireless device inspection committee released by the Ministry [11] contained three interesting findings related to Japanese aviation operators: 1) very few problems have been found by periodic inspection, 2) other countries do not require such frequent periodic inspection, and 3) the burden of such periodic inspections is quite heavy. Similar problems related to on-condition maintenance can be found in other industries. Therefore, there is a pressing need to develop more flexible and effective inspection policies and to improve existing ones.

1.2 Previous Research

The optimal maintenance decision-making problem for stochastically deteriorating systems has been investigated, and it is typically formulated as a Markov decision process (MDP). Derman [3] investigated a decision-making problem for a deteriorating system for which the true state can be known by regular inspections, and one of two actions, keep operating for one more period or replace, is selected on the basis of the true state. He proposed a monotone structure for maintenance policies and provided a sufficient condition for the optimality of a monotone policy. Lam and Yeh [7] presented algorithms for deriving optimal maintenance policies that minimize the expected long-run cost for continuous-time Markov deteriorating systems. Unlike those in Dermanfs model, the time durations for inspection and replacement are non-negligible in their model. Moreover, the state of a system in their model can transit to and only to a) the next stage due to deterioration or b) the failure state due to a shock. Structural optimal policies have also been obtained under certain conditions. Elwany, Gebraeel, and Maillart [4] presented an exponentially increasing deteriorating model in which the true state can be fully captured only by inspection every period. The replacement problem is formulated as an MDP, and the optimal maintenance policy is shown to be a monotone policy. Tamura [14] proposed an optimal policy with a dynamic inspection interval that minimizes the expected total discounted cost for a deteriorating system.

In most previous research, the system was assumed to deteriorate in accordance with a stationary MDP. However, the deterioration of a system is sometimes affected by its age. For example, analysis using actual data has shown that the deterioration rate of wireless devices in airplanes increases with the devicefs age. We call such systems *aging systems* in this research. Few researchers have considered the effect of age in their models and simply investigated the decision-making policy on the basis of simulation results. Chen [2] used an aging factor to describe an aging system and presented an algorithm for determining the actions over a finite time period. Abeygunawardane, Jirutitijaroen, and Xu [1] formulated the decision-making problem for aging systems by using an Markov decision process and proposed a solution for adaptive decision-making.

Using a proposed non-stationary MDP model, we formulated a decision-making problem for an aging system and analytically investigated the optimal policies for on-condition maintenance. We obtained several structural properties to establish the existence of an optimal policy in which there is at most one transition from keep operating to replace with respect to both the system's deterioration state and age under

certain conditions. A monotone property of the inspection interval can be identified under the derived conditions. A comparison of the models used in previous research and in this research is shown by Table 1. In this research, the action keep operating is accompanied by a flexible inspection interval. In the following, we call the action keep operating and inspect the system after l periods *inspection l.*

Table 1. Comparison of previous and current research models

	Regular inspection interval	Flexible inspection interval
Stationary deterioration	Derman [3], Lam and Yeh [7] Elwany, Gebraeel, and Maillart [4]	Tamura [14]
Non-stationary deterioration	Chen [2] Abeygunawardane, Jirutitijaroen, and Xu [1]	Current research

The rest of this manuscript is organized as follows. In section 2, we formulate a non-stationary MDP model for the optimal maintenance decision-making problem. Section 3 first provides preliminary assumptions and then investigates the optimal decision-making problem for an aging system. A monotone property of the optimal inspection interval is derived under certain conditions. Finally, in section 4, we briefly summarize the key points and mention future work.

2 Model Description

The deterioration state of the system i can be classified into one of a finite number of states, $i \in \mathbb{S} = \{1, 2, \ldots, n\}$. State numbers are ordered by increasing degree of deterioration. This means that state 1 is the initial, new state and state n is the final, failed state. The true deterioration state can only be measured by using an inspection device (taking a negligible amount of time).

Let $t \in \mathbb{N} = \{0, 1, 2 \cdots\}$ indicate the system age. It can be calendar time, total system operating time, or total system operating amount. The start time is 0. The system becomes one age unit older each period. A two-dimensional information vector (i, t) is defined as the status of the deterioration process in space $\mathbb{S} \times \mathbb{N}$. Once the system is inspected, the true deterioration state is known, and the maintenance manager decides whether the system should be replaced.

In previous research, the system was assumed to be inspected every period irrespective of the system's deterioration state, and the maintenance manager selects one of two actions: keep one more period or replace the system with a new one. However, in an actual situation, the inspection interval may be extended if the system is apparently in a good state, and an inspection may be carried out immediately after one more period of operation if the system is apparently in a deteriorated state.

In the model proposed in this research, if keep operating is selected, the inspection interval is determined simultaneously on the basis of the current deterioration state and age. To distinguish it from the keep with a regular inspection interval that was used in previous research, we call the action keep operating accompanied with a flexible

inspection interval l as *inspection l*. Once *inspection l* is selected, the system is inspected after l periods of operation. We assume that the maximum inspection interval is L. This means the system can be continuously operated for at most L periods.

If *inspection l* is selected for status (i, t), a certain sequence of events is initiated. A real-valued operation cost, $K(i, t)$, is incurred for the current period for status (i, t). We assume that a system failure cannot be detected immediately and that the loss due to a failure is included in $K(n, t)$. After l periods of operation, the system is inspected at the beginning of the $(l + 1)$-th period, and an inspection cost of I is incurred. The time for inspection is neglected since the state of the system does not change during the inspection. The system state transits in accordance with a non-stationary transition law over l periods:

$$\boldsymbol{P}(l|t) = \left(p_{ij}(l|t) \right)_{i,j \in \mathbb{S}}, l \leq L, t \in \mathbb{N}.$$

Here, $p_{ij}^{(l|t)}$ is the probability that the system transits from state i to j after l periods of operation given that it starts operating at age t. Let

$$C_l(i, t) = \begin{cases} K(i, t) & (l = 1) \\ K(i, t) + \displaystyle\sum_{k=1}^{l-1} \beta^k \sum_{j=1}^{n} p(k|t)_{ij} K(j, t + k) & (2 \leq l \leq L) \end{cases} \tag{1}$$

be the total operation cost for a system starting at status (i, t) over l periods, and β $(0 < \beta < 1)$ be the discount factor for one period.

If the maintenance manager decides to replace the system when the system status is (i, t), replacement cost $R(i, t)$ is incurred for the current period. The system is replaced with a new one, and the system's state and age are both reset to their initial values from the next period.

Let $a(i, t)$ be the optimal action, i.e., the action that minimizes the expected long-run cost of the process starting when the system status is (i, t). An optimal policy is a sequence of optimal actions for every possible status $(i, t) \in \mathbb{S} \times \mathbb{N}$. Let M be the number of periods and let

$$V_M(i, t) = \min \begin{cases} \displaystyle\min_{1 \leq l \leq L} V_M^{(l)}(i, t) & \equiv V_M^{(I)}(i, t) \\ R + \beta V_{M-1}(1, 0) & \equiv V_M^{(R)}(i, t) \end{cases} \tag{2}$$

be the optimal expected cost of the process starting when the system status is (i, t) over M periods. Here,

$$V_M^{(l)}(i, t) = C_l(i, t) + \beta^l I + \beta^l \sum_{j=1}^{n} p_{ij}(l|t) V_{M-l}(j, t + l) \tag{3}$$

is the total expected cost if *inspection l* is selected when the system status is (i, t) and the optimal decision-making policy is followed from the $l + 1$-th period. For the current status (i, t), we determine the optimal interval l by minimizing $V_M^{(l)}(i, t)$. We call an inspection with such an interval an *optimal inspection* when the system status is (i, t). Let $V_M^{(I)}(i, t)$ denote the total expected cost for an optimal inspection over M

periods. $V_M^{(R)}(i, t)$ is the total expected cost if the system is replaced at the beginning of the current period and the optimal decision-making policy is followed from the next period. For every possible status (i, t), the maintenance manager compares the expected costs of *optimal inspection* ($V_M^{(I)}(i, t)$) and *replacement* ($V_M^{(R)}(i, t)$) and takes the action with the lower total expected cost as the optimal action for status (i, t). Let $V_0(i, t) = 0$.

From the standard argument of contraction mapping theory [13], $V_M(i, t)$ converges to $V(i, t)$ as M tends to infinity. $V_M^{(I)}(i, t)$ and $V_M^{(R)}(i, t)$ converge to $V^{(I)}(i, t)$ and $V^{(R)}(i, t)$, respectively. Therefore, we have a recursive function for the optimal total expected cost function:

$$V(i, t) = \min \begin{cases} V^{(I)}(i, t) \\ V^{(R)}(i, t) \end{cases}, \tag{4}$$

where

$$V^{(I)}(i, t) = \min_{1 \le l \le L} V^{(l)}(i, t). \tag{5}$$

Optimal action $a(i, t)$ can be found by calculating the recursive functions given by (4) and (5).

3 Optimal Inspection Policy

In this section, we investigate the properties of the optimal inspection policy for aging systems.

3.1 Assumptions

We first define a totally positive of order 2 (Karlin [6]) property, abbreviated as TP$_2$, for the matrix used in this research. For an $(n \times m)$ matrix $\boldsymbol{Q}$, if

$$q_{ij} q_{i'j'} \ge q_{ij'} q_{i'j} \tag{6}$$

holds for any $1 \le i < i' \le n$ and $1 \le j < j' \le m$, it is said that $\boldsymbol{Q}$ has a property of TP$_2$. Here, q_{ij} is the element of $\boldsymbol{Q}$ in the i-th row and j-th column. If we focus on vectors $\{\boldsymbol{Q}\}_i$ and $\{\boldsymbol{Q}\}_{i'}$, which are the i-th line and i'-th line of matrix $\boldsymbol{Q}$, the TP$_2$ property given by (6) between two line vectors can be denoted by $\{\boldsymbol{Q}\}_i \overset{\text{TP}_2}{\prec} \{\boldsymbol{Q}\}_{i'}$.

We made the following assumptions about the deterioration process. Note that we considered functions as either increasing or decreasing in the weak sense throughout this research.

(A-1) $\boldsymbol{P}(l|t)$ has a property of TP$_2$ for any age t and $l \le L$.

(A-2) $\{\boldsymbol{P}(l|t_1)\}_i \overset{\text{TP}_2}{\prec} \{\boldsymbol{P}(l|t_2)\}_i$ for any $t_1 < t_2$ and $l \le L$.

(A-3) $K(i, t)$ is increasing in both deterioration state $i \in \mathbb{S}$ and age $t \in \mathbb{N}$.

Assumption (A-1) means that, as the state of the system deteriorates, the system at any age is more likely to move to an even more deteriorated state after several periods of operation. This assumption is widely used to analyze the structure of the optimal

cost function in terms of reliability and maintenance optimization [12] [8] [9] [5]. Assumption (A-2) means that, as the system ages, it is more likely to move to a more deteriorated state after several periods of operation. In other words, given two systems with the same level of deterioration, the older system is more likely to be in a worse condition after several periods of operation. In most previous studies [3] [12] [8] [9] [5], a one-step deterioration state transition probability was considered that did not depend on the system age. Assumption (A-3) means that the costs for keep operating for one period do not decrease as the system becomes more deteriorated and older.

3.2 Structural Properties

In this section, we investigate the optimal decision-making problem for aging systems under the assumption of on-condition maintenance. The structural properties of the optimal expected cost function are examined, and these properties establish that the optimal action changes from an optimal inspection to replacement at most once under certain conditions. In addition, the optimal inspection interval is found to be decreasing in both the system's deterioration state and age.

To prove our main theorem, we use four lemmas. Lemma 1 was originally proved by Derman [3], and Lemma 2 was originally proved by Karlin [6]. The latter is known as the variation diminishing property.

Lemma 1 (Derman [3]). *If assumption (A-1) holds and $f(j,t)$ is an increasing function in j for any $l \ (\leq L)$ and t,*

$$\sum_{j=1}^{n} p_{ij}(l|t)f(j,t) \leq \sum_{j=1}^{n} p_{i'j}(l|t)f(j,t)$$

for any $1 \leq i < i' \leq n$.

Lemma 2 (Karlin [6]). *Under assumption (A-1), if the sign of $f(i)$ changes from negative to positive at most once in i, the sign of $\sum_{j=1}^{n} p_{ij}(l|t)f(j)$ changes in the same way in i for any $l \ (\leq L)$ and $t \in \mathbb{T}$.*

Lemmas 3 and 4 investigate the properties of the total expect cost function if inspection is selected. Lemma 3 provides sufficient conditions under which $V^{(l)}(i,t)$ is increasing in deterioration state i. This means that, for two systems at the same age, the total expected cost of inspection increases as the system deteriorates if the inspection intervals stay the same. Lemma 4 provides sufficient conditions under which $V^{(l)}(i,t)$ is increasing in age t. This means that, for two systems in the same state of deterioration, the total expected cost of inspection for the newer system is less than that of the older one if their next inspections are done at the same time.

Lemma 3. *Under assumptions (A-1) through (A-3), $V^{(l)}(i,t)$ is increasing in i for any $l \ (\leq L)$ and t.*

Lemma 4. *Under assumptions (A-1) through (A-3), $V^{(l)}(i,t)$ is increasing in t for any $l \ (\leq L)$ and $i \in \mathbb{S}$.*

Theorems 1 and 2 investigate the properties of the difference between the total expected cost if inspection is selected for (i, t) and the optimal policy is followed for the remaining periods and that if replacement is selected. They provide a set of sufficient conditions for the monotonicity of $V^{(I)}(i, t) - V^{(R)}(i, t)$ in both deterioration state and age. Corollary 1 shows that the optimal action changes from an optimal inspection to replacement at most once. This result comes from Theorems 1 and 2.

Theorem 1. *Under assumptions (A-1) through (A-3), the sign of $V^{(I)}(i, t) - V^{(R)}(i, t)$ changes at most once in deterioration state i for any age t.*

Theorem 2. *Under assumptions (A-1) through (A-3), the sign of $V^{(I)}(i, t) - V^{(R)}(i, t)$ changes at most once in age t for any deterioration state i.*

Corollary 1. *Under assumptions (A-1) through (A-3), the optimal action changes from optimal inspection to replacement at most once as the system deteriorates and ages.*

Theorems 3 and 4 investigate the total expected cost functions of inspections with two different intervals by checking the property of function $V^{(l)}(i, t) - V^{(l+1)}(i, t)$. If $V^{(l)}(i, t) - V^{(l+1)}(i, t)$ is increasing in deterioration state i and age t, inspection with operation interval l leads to a greater increase in cost than inspection with operation interval $l + 1$ as the system either deteriorates or ages. This means that the optimal inspection interval transits from l to $l + 1$ at most once if the decision is made on the basis of current status (i, t). From Theorems 3 and 4, we can derive Corollary 2, which shows that the assumptions presented in Section 3.1 are sufficient for the inspection interval to be monotone decreasing in both state i and age t.

Theorem 3. *Under assumptions (A-1) through (A-3), the sign of $V^{(l)}(i, t) - V^{(l+1)}(i, t)$ changes at most once in deterioration state i for any age t.*

Theorem 4. *Under assumptions (A-1) through (A-3), the sign of $V^{(l)}(i, t) - V^{(l+1)}(i, t)$ changes at most once in age t for any deterioration state i.*

Corollary 2. *Under assumptions (A-1) through (A-3), the optimal inspection interval decreases in both state i and age t.*

Figure 1 shows an example of policies having monotone properties given by Corollaries 1 and 2.

4   Summary

In previous research, the system was assumed to be inspected at a regular interval irrespective of the system's deterioration state and age. However, the inspection interval should differ in accordance with the systemfs status. In the proposed inspection policy, the interval is flexible and is determined on the basis of the current deterioration state and age. We investigated the optimal inspection policy and obtained a monotone policy for which the optimal action changes from optimal inspection to replacement at most once with respect to the deterioration state and age under certain conditions. Furthermore, we showed that the optimal inspection interval decreases as deterioration

and age increase. The proposed inspection policy enables the maintenance manager to carry out an effective on-condition maintenance plan. The next inspection can be scheduled on the basis of the current system status dynamically. Therefore, the inspection cost and number of maintenance personnel can be reduced.

The inspection of a deteriorating system was assumed to be perfect, meaning that the true state of the system was known exactly. However, in many cases, inspection provides only incomplete information related to the true state. In future work, we will take these limitations into account and investigate optimal decision-making for a system with imperfect inspection.

$$(i < i' < i'', t < t' < t'', l < l')$$
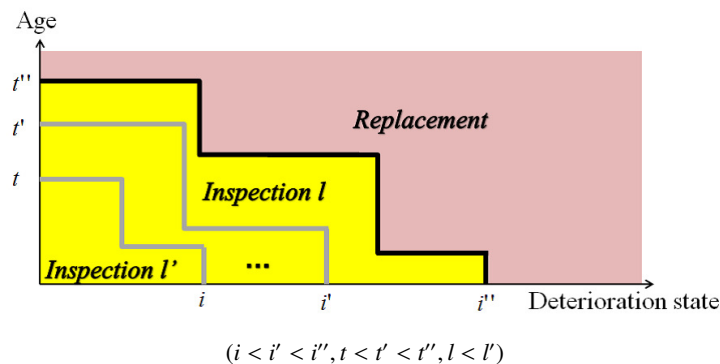
Fig. 1. Example of policies having monotone properties given by Corollaries 1 and 2.

References

1. S. K. Abeygunawardane, P. Jirutitijaroen and H. Xu. Adaptive maintenance policies for aging devices using a Markov decision process. *IEEE Transactions on Power Systems*, 28, 3194-3203, 2013.
2. C-T. Chen. Dynamic preventive maintenance strategy for an aging and deteriorating production system. *Expert Systems with Applications*, 38, 6287-6293, 2011.
3. C. Derman. On optimal replacement rules when changes of state are Markovian. *Mathematical Optimization Techniques* (edited by R. Bellman), University of California Press, 201-210, 1963.
4. A. H. Elwany, N. Z. Gebraeel and L. M. Maillart. Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research*. 59, 684-695, 2011.
5. L. Jin, K. Suzuki and K. Kumagai. Necessary and sufficient condition for control limit policy for partially observable Markov decision process. *Journal of the Japanese Society for Quality Control*, 41, 238-249, 2011.
6. S. Karlin. *Total Positivity*, Vol. I, Stanford University Press, Stanford, California, 1968.
7. C. T. Lam and R. H. Yeh. Optimal maintenance-policies for deteriorating systems under various maintenance strategies. *IEEE Transactions on Reliability*, 43, 423-430, 1994.
8. W. S. Lovejoy. Some monotonicity results for partially observed Markov decision processes. *Operations Research*, 35, 736-743, 1997.

9. H. Maehara and K. Suzuki. An optimal policy for condition monitoring maintenance with uncertain repair. *The Journal of Reliability Engineering Association of Japan*, 27, 219-230 (in Japanese), 2005.

10. Ministry of Internal Affairs and Communications of Japan. *Japanese radio wave laws*, 2014. http://law.e-gov.go.jp/htmldata/S25/S25HO131.html (in Japanese).

11. Ministry of Internal Affairs and Communications of Japan. *Report on inspection for wireless devices*, 2013. http://www.soumu.go.jp/main_content/000211500.pdf (in Japanese).

12. M. Ohnishi, H. Kawai and H. Mine. An optimal inspection and replacement policy under incomplete state information. *European Journal of Operational Research*, 27, 117-128, 1986.

13. M. L. Puterman. *Markov Decision Processes: discrete stochastic dynamic programming*, Wiley, 1994.

14. N. Tamura. An optimal inspection policy for a Markovian deteriorating system with stochastically decreasing lifetime. *International Journal of Reliability Quality and Operations Management* 2, 13-28, 2011.

Expected Shortfall Estimation via Importance Sampling GARCH Predictive Distribution

Kanako Nakajima[1] and Takayuki Shiohama[2]

[1] Graduate School of Engineering, Tokyo University of Science, Kagurazaka 1-3, Shinjuku, Tokyo, Japan
   (E-mail: `kanako@ms.kagu.tus.ac.jp`)
[2] Department of Management Science, Faculty of Engineering, Tokyo University of Science, Kagurazaka 1-3, Shinjuku, Tokyo, Japan
   (E-mail: `shiohama@ms.kagu.tus.ac.jp`)

Abstract. Estimations of Value at Risk (VaR) and Expected Shortfall (ES) are important for banks and other participants in financial markets. These financial risk measures capture the amount of financial risk caused by a rare event. While precise estimations of these risk measures are desirable, standard Monte Carlo simulation-based VaR and ES estimates require significant computing time. We propose an estimator procedure for computing ES in a Bayesian framework using GARCH predictive distributions and importance sampling methods. Approximations for the GARCH predictive distributions are also proposed. The proposed methods are compared to the standard cases, and show less computing time for the same degree of numerical accuracy. Finally, the proposed procedures are used to estimate the ES of Japanese stock market data.
Keywords: Bayesian estimation, GARCH model, Laplace approximation, importance sampling, predictive distribution.

1 Introduction

Market disruptions caused by the worldwide financial crisis required that financial institutions plan, monitor, evaluate, and assess financial risks. The financial crisis in the United States, triggered by the collapse of Lehman Brothers in 2008, resulted in global downturn. Further examples of financial crises include the Great East Japan Earthquake in 2011, and the escalating debt crisis in Europe between 2010 and 2012. Utilizing these data to study and/or forecast the rare events associated with a financial crisis is becoming increasingly important.

The Expected Shortfall (ES) and the Value at Risk (VaR) are popular measures of financial risks for an asset or a portfolio of assets. The VaR measures the maximum loss of a financial asset price due to market movements over a given holding period with a certain level of confidence. Let X be the negative log return of the market values of an asset or a portfolio of assets, and F be

the distribution of X. Then, for a positive value p close to zero, the $100p\%$ VaR is defined as

$$VaR(p) = \min[x : F(x) \geq (1-p)],$$

which is the $(1-p)$th quantile of the loss distribution F. The ES is the expected loss given that the VaR is exceeded and measures the financial risk of very rare events. The ES is defined as follows:

$$ES(p) = E[X|X \geq VaR(p)].$$

The VaR is a financial risk measure of a rare event and is widely used as a standard tool in financial markets. Since the work of Artzner *et al.* [1,2], it is well known that the VaR is not a coherent measure of risk. If the risk measure satisfies the four requirements of homogeneity, monotonicity, translation invariance, and subadditivity, then it is a "coherent" measure of risk. The VaR satisfies the first three requirements, but lacks the subadditivity property. Let X_1 and X_2 denote the random returns of two financial assets, then the subadditivity of a risk measure $\rho(\cdot)$ means

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2).$$

This property is desirable because it is consistent with the diversification principle of modern portfolio theory.

The Basel Committee identified the ES, a theoretically coherent risk measure, as an alternative to the VaR. On May 3, 2012, the Basel Committee explicitly raised the prospect of phasing out the VaR and replacing it with the ES in the Third Basel Accord (Basel Committee [3]).

Various studies have attempted to obtain accurate ES values, including sample averages based on the order statistics of Yamai and Yoshiba [24] and estimators using the extreme-value theories of Cotter and Dowd [7]. For the conditional distribution approach, we need to incorporate changes in market volatility to reflect the most recent risk level. So and Wong [20] proposed statistical methods to estimate the multi-period ES under GARCH models using the conditional kurtosis. Hoogerherde and van Dijk [12] proposed a Bayesian framework for forecasting the VaR and ES using mixture of t distributions.

The classic Monte Carlo methods to calculate the VaR and ES are possible. One major limitation of any Monte Carlo method is that the computational effort can be too heavy to make the method widely applicable. Thus, one objective of this study is to develop a new ES estimator based on the importance sampling distribution of the conditional predictive distribution of GARCH returns. For further details about GARCH models, Francq and Zakoian [9] provides extensive reviews on the variants of GARCH models and their empirical applications.

We propose an estimation procedure for computing the ES in a Bayesian framework using GARCH predictive distributions and the importance sampling methods. To this end, we present new analytical methods to approximate the cumulant generating function of the predictive distribution from the GARCH

384

models. Several approximations are proposed, including the usual fully exponential Laplace approximation, Lindley approximation, and Miyata approximation. We also discuss and compare the precisions of these estimations.

This paper is organized as follows. Section 2 defines the GARCH models and describes the ES estimation procedures. Bayesian approximation is also explained. Section 3 presents the estimation procedures based on the importance sampling methods. To illustrate how our estimation procedures work, we conduct a data analysis using Japanese stock market data in Section 4. Conclusions are presented in Section 5.

2 Models and Methodology

The first step in a Bayesian analysis is to choose a probability model for the data. In this research, we have chosen models widely used in financial analyses to formulate the movement of financial asset prices. The ARCH model of Engle [8] and the GARCH model of Bollerslev [4] aim to more accurately describe the phenomenon of volatility clustering. Volatility clustering refers to the empirical idea that large changes tend to be followed by large changes.

Let p_t be the financial asset price at time t. The daily percent log return is defined as $y_t = 100 \log(p_t/p_{t-1})$. Let the volatility be defined as h_t. Then, the Autoregressive Conditional Heteroskedasticity(ARCH) model is described as follows:

$$y_t = \varepsilon_t h_t, \quad h_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2, \quad \varepsilon_t \sim \text{i.i.d.}(0,1), \tag{1}$$

where $\alpha_0 > 0$, and $0 \leq \alpha_1 \leq 1$ for positive volatility.

The generalized ARCH (GARCH) model is often preferred by financial professionals as it provides a more real-world context than other forms when trying to predict prices. The GARCH(1,1) model can be expressed as follows:

$$y_t = \varepsilon_t h_t, \quad h_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta h_{t-1}^2, \quad \varepsilon_t \sim \text{i.i.d.}(0,1), \tag{2}$$

where $\alpha_0 > 0, 0 \leq \alpha_1 \leq 1, 0 \leq \beta \leq 1$, and $\alpha_1 + \beta < 1$. These restrictions ensure that non-negative volatility and model stationarity. There is a special case of the GARCH model called RiskMetrics (Longerstaey [16]) model whose volatility equation is given by

$$h_t^2 = (1 - \lambda)y_{t-1}^2 + \lambda h_{t-1}^2, \tag{3}$$

where $\lambda \in (0, 1)$ is a smoothing parameter, for which, according to the RiskMetrics model, a reasonable choice is $\lambda = 0.94$ for daily series. This model can be viewed as the IGARCH(1,1) model without intercept.

Computing the ES in a Bayesian framework is one of the main purposes of this research. Suppose we have observed data $\boldsymbol{y}_T = (y_1, \ldots, y_T)$ from the (G)ARCH models defined by equations (1) or (2). Denote $p_T(\boldsymbol{y}_T|\boldsymbol{\theta})$ by the probability density function $\boldsymbol{y}_T$, given $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta)'$. Conditional on $\boldsymbol{\theta}$ and y_{i-1}, y_i are independent and identically distributed random variables with density $p(y_i|y_{i-1}, \boldsymbol{\theta})$. Assume that ε_t follows a standard normal distribution. Then,

$p_T(\boldsymbol{y}_T|\boldsymbol{\theta}) = \prod_{i=1}^{T} p(y_i|y_{i-1}, \boldsymbol{\theta})$ is a quasi-maximum likelihood (QML) even if the true distribution of ε_t is non-normal. It is well known that, provided the error distribution has a finite fourth moment, QML estimators are asymptotically normally distributed in the case of the ARCH by Weiss [23] and the GARCH(1,1) model by Lee and Hansen [14].

If we specify a prior density $\pi(\boldsymbol{\theta})$, then by the usual Bayesian calculus, the posterior distribution of $\boldsymbol{\theta}$ from a sample $\boldsymbol{y}_T$ is

$$p_T(\boldsymbol{\theta}|\boldsymbol{y}_T) = \frac{p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}.$$

Once the posterior distribution for the unknown parameters is obtained, we can predict future data conditional on the data $\boldsymbol{y}_T$. Then, the predictive density for y_{T+1} becomes

$$p(y_{T+1}|\boldsymbol{y}_T) = \frac{\int_\Theta p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{4}$$

This is called the Bayesian predictive distribution. In large deviation theory, the tail probability of the predictive distribution is bounded by in terms of the cumulant generating function of $p(y_{T+1}|\boldsymbol{y}_T)$. To do this, we need to define the moment generating function of the predictive distribution, which is given by

$$M_{y_{T+1}}(s) = E\left[e^{sY_{T+1}}\right] = \int_{-\infty}^{\infty} e^{sy_{T+1}}p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})dy_{T+1}.$$

Then the cumulant generating function is the log of the moment generating function, which is given by

$$\psi(s) = \log E[e^{sY_{T+1}}] = \log \int e^{sy_{T+1}}p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})dy_{T+1}$$

$$= \log \frac{\int_{-\infty}^{\infty} e^{sy_{T+1}} \int_\Theta p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})p_T(\boldsymbol{\theta}|\boldsymbol{y}_T)d\boldsymbol{\theta} \cdot dy_{T+1}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}$$

$$= \log \frac{\int_\Theta \int_{-\infty}^{\infty} e^{sy_{T+1}}p(y_{T+1}|\boldsymbol{y}, \boldsymbol{\theta})dy_{T+1}p_T(\boldsymbol{\theta}|\boldsymbol{y}_T)d\boldsymbol{\theta}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}$$

$$= \log \frac{\int_\Theta E[e^{sy_{T+1}}|\boldsymbol{y}_T, \boldsymbol{\theta}]p_T(\boldsymbol{\theta}|\boldsymbol{y}_T)d\boldsymbol{\theta}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{5}$$

Let g be a smooth, positive function of the parameter space, and set

$$g^{(MGF)}(\boldsymbol{\theta}) = E[e^{sy_{T+1}}|\boldsymbol{y}_T, \boldsymbol{\theta}] = \int e^{sy_{T+1}}p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})dy_{T+1},$$

and $g^{(p)}(\boldsymbol{\theta}) = p(y_{T+1}|\boldsymbol{y}_T, \boldsymbol{\theta})$ as the moment generating and the distribution functions of a vector $\boldsymbol{\theta}$ conditional on $\boldsymbol{y}_T$, respectively. The posterior mean of $g^{(m)}(\boldsymbol{\theta})$ can be written as

$$E(g^{(m)}(\boldsymbol{\theta})|\boldsymbol{y}_T) = \frac{\int_\Theta g^{(m)}(\boldsymbol{\theta})p_T(\boldsymbol{\theta}|\boldsymbol{y}_T)d\boldsymbol{\theta}}{\int_\Theta p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}, \qquad m \in \{MGF, p\}. \tag{6}$$

In general, we cannot calculate the exact posterior mean of function (6). Thus, we provide approximations to (6) using Laplace methods.

The Laplace approximation is widely used in the asymptotic expansion of integrals, which involves expanding the parameters in a Taylor series about the mean values and then using the second-order terms to approximate the integrals. It is well known that a posterior distribution is going to have a similar form to the normal distribution as the number of observations go to infinity. The posterior distribution is approximated using normal distribution in the Laplace approximation. Implementation details of the Laplace method may be found in Tierney and Kadane [21], Tierney et $al.$ [22], and Miyata [17,18], who introduce applied Laplace approximation methods.

We consider the fully exponential form of Laplace method by defining $h_T(\boldsymbol{\theta}) = -T^{-1}\log\{p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\}$. Suppose $\{\log p_T(\boldsymbol{y}_T|\boldsymbol{\theta}) + \log \pi(\boldsymbol{\theta})\}$ has strict local maximum $\hat{\boldsymbol{\theta}}$ and a positive definite matrix $D^2 h_T(\boldsymbol{\theta}) = (\partial^2/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}')h_T(\boldsymbol{\theta})$. Then, the approximation for an integral of the form $I = \int_\Theta b(\boldsymbol{\theta})\exp\{-Th_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}$ is given by a Taylor series expansion of $h_T(\boldsymbol{\theta})$, as follows:

$$I \approx b(\hat{\boldsymbol{\theta}})\exp(-Th(\hat{\boldsymbol{\theta}}))(2\pi/T)^{D/2}|D^2 h_T(\hat{\boldsymbol{\theta}})|^{-1/2}, \qquad (7)$$

where $\hat{\boldsymbol{\theta}}$ is a point in $\mathbb{R}^D$. Here, $b(\boldsymbol{\theta})$ is continuous, differentiable, and nonzero at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta} \in \mathbb{R}^D$. The approximation of I has the error term $O(T^{-1})$.

2.1   Tierney and Kadane Approximation

Tierney and Kadane [21] argued in favor of a specific implementation of the Laplace approximation, called the fully exponential form, that produces results more accurate than those of other approaches. Instead of errors typically being $O(T^{-1})$ for the conventional Laplace approximation, they found that the errors are $O(T^{-2})$ owing to the cancellation of $O(T^{-1})$ error terms using their method. Their proposed approximation is

$$\widehat{g}_{\mathrm{TK}}^{(m)} = \left(\frac{|D^2 h_T(\hat{\boldsymbol{\theta}})|}{|D^2 h_T^*(\hat{\boldsymbol{\theta}}^*)|}\right)^{1/2}\frac{\exp\{-Th_T^*(\hat{\boldsymbol{\theta}}^*)\}}{\exp\{-Th_T(\hat{\boldsymbol{\theta}})\}}, \qquad m \in \{MGF, p\}, \qquad (8)$$

where $h_T^* = -T^{-1}\log\{g(\boldsymbol{\theta})p_T(\boldsymbol{y}_T|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\}$ and $\hat{\boldsymbol{\theta}}^* = \arg\sup_{\theta\in\Theta}\{-h_T^*(\boldsymbol{\theta})\}$.

2.2   Miyata Approximation

Miyata [18] proposed a simple approximation to (6) based on the posterior means of $\boldsymbol{\theta}$. Denote $\mathrm{tr}(\cdot)$ by the trace of a matrix and let $\boldsymbol{\theta}^B$ be the posterior mean. Then, the approximation to the posterior mean of $g^{(m)}(\boldsymbol{\theta})$ in (6) is given by

$$\widehat{g}_{\mathrm{M}}^{(m)} = g^{(m)}(\hat{\boldsymbol{\theta}}^B; \boldsymbol{y}_T) + \frac{1}{2T}\mathrm{tr}\{D^2 g^{(m)}(\hat{\boldsymbol{\theta}}^B)[D^2 h_T(\hat{\boldsymbol{\theta}}^B)]^{-1}\}, \; m \in \{MGF, p\}, \; (9)$$

where $\hat{\boldsymbol{\theta}}^B$ is a second-order approximation to the posterior mean $\boldsymbol{\theta}^B$. Miyata [18] shows that the approximation is valid under regularity conditions,

such that $g^{(m)}(\boldsymbol{\theta}) = \widehat{g}_M^{(m)} + O(T^{-2})$. We only have to calculate this approximation using the second derivative. When T is comparably small, the result $\widehat{g}_M^{(m)} < 0$ cannot be avoided.

2.3 Lindley Approximation

Lindley [15] developed an asymptotic expansion to evaluate the ratio of the integral of the form as

$$\widehat{g}_{\mathrm{L}}^{(m)} = g^{(m)}(\hat{\boldsymbol{\theta}}; \boldsymbol{y}_T) - \frac{1}{2T} \sum_{ijrs} \left\{ \frac{\partial}{\partial \theta_i} g^{(m)}(\hat{\boldsymbol{\theta}}) \right\} h^{ij} h^{rs} h_{ijk}$$

$$+ \frac{1}{2T} \sum_{ij} h^{ij} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} g^{(m)}(\hat{\boldsymbol{\theta}}) \right\}, \qquad m \in \{MGF, p\}, \qquad (10)$$

where h^{ij} are the (i,j)-components of the inverse of the Hessian of $h_T(\hat{\boldsymbol{\theta}})$ and $h_{ijk} = (\partial^3 / \partial \theta_i \partial \theta_j \partial \theta_k) h_T(\boldsymbol{\theta})$.

3 Importance Sampling Estimation Procedures

There is a large body of work on the use of asymptotic deviations to identify an effective change of measure for estimating rate event probabilities by simulation. For example, see Glasserman and Wang [10], Glynn [11], Muller [19], and Heidelberger [13]. For numerous references and discussions on these studies, see Chen *et al.* [6]. In order to use importance sampling for loss probabilities estimations, an importance sampling density has to be chosen. For this, we first define the exponential change of measure as below.

Let Y be a random variable on $\mathbb{R}$ and have distribution function F and cumulant generating function ψ under the original probability measure P. Then,

$$\psi(s) = \log \int \exp(sy) dF(y).$$

Set ψ's domain to $S = \{s : \psi(s) < \infty\}$ and for each $s \in S$, let F_s denote the distribution function obtained by the exponential change of measure. In other words,

$$dF_s(y) = \exp(sy - \psi(s)) dF(y).$$

Then the cumulant generating function ψ contains important information about the distributions F_s. Let ψ' denote the derivative of ψ. It is a standard consequence of exponential twisting that $\psi'(s)$ is the mean and $\psi''(s)$ is the variance of a random variable with distribution function F_s. Each F_s is a probability density function, and $\{F_s; s \in S\}$ form an exponential family of distributions. A transformation from F to F_s is called an exponential change of measure. The density function of F_s is given by

$$f_s(y) = e^{sy - \psi(s)} f(y), \qquad (11)$$

where f is supposed to be the density of F.

The exponential changes of measure stem from the large deviation theory (see Bucklew [5]). According to the large deviation theory, the tail of the distribution of $F(Y)$ can be approximated as

$$P(Y > y) \approx \exp(-ys_y + \psi(s_y)),$$

for $y \gg E(Y)$, where s_y is the root of the equation $\psi(s_y) = y$. The cumulant generating function $\psi(s)$ can be approximated by equations (8), (9), and (10). In particular, let $\tilde{s}_p$ be the root of the equation

$$-\tilde{s}_p \psi'(\tilde{s}_p) + \psi(\tilde{s}_p) = \log(1 - p).$$

Then, for p closer to 1, we observe

$$P(Y \geq \psi'(\tilde{s}_p)) \approx 1 - p,$$

which suggests that $\psi'(\tilde{s}_p)$ is an approximation to the quantile $F^{-1}(p)$.

Let Y have the initial distribution function F_0. If we apply an exponential change of measure such that Y has the new distribution F_s, the corresponding likelihood ratio is given by

$$\frac{dF_0(Y)}{dF_s(Y)} = \exp(-sY + \psi(s)).$$

Now, we propose the following algorithm to compute the VaR and ES based on the importance sampling methods, for $method \in \{TK, M, L\}$.

Step 1. Choose the model and set the prior for each parameter, then construct a likelihood from the model. These are required to calculate the approximate predictive distribution:

$$\hat{g}^{(p)}_{method} = \hat{p}(y_{T+1}|\boldsymbol{y}_t, \boldsymbol{\theta}).$$

Step 2. Let $\widehat{\psi}_{method}(s) = \log \widehat{g}^{(MGF)}_{method}$. Obtain an approximation to the cumulant generating function using (8), (9), and (10).

Step 3. Obtain the value of s_x, which is the root of the nonlinear equation

$$\widehat{\psi}'_{method}(s_x) = x,$$

for fixed x. Note that the cumulant generating function of Y is approximated by $\widehat{\psi}_{method}(s)$. Under the importance sampling distribution, the expected value of Y is equal to x.

Step 4. Simulate a set of draws $Y_1^{(IS)}, \ldots, Y_N^{(IS)}$ from the approximated importance sampling density (11) using the acceptance-rejection method:

$$\hat{p}_{s,method}(y) = \exp(sy - \widehat{\psi}_{method}(s))\hat{p}_{0,method}(y).$$

Step 5. Take samples from the distribution in step 3. Then the importance weight for collecting the bias can be determined as follows:

$$\frac{dF_0}{dF_s} = \exp(-sY^{(IS)} + \widehat{\psi}_{method}(s)). \tag{12}$$

Step 6. Estimate the VaR and ES, as follows:

$$\widehat{VaR}(\alpha) = Y^{(IS)}_{k,N}, \quad \text{where} \quad k = \min\Big\{j : \sum_{i=1}^{j} w_{i,N} \le \alpha\Big\}, \tag{13}$$

and

$$\widehat{ES}(\alpha) = \frac{1}{N}\sum_{i=1}^{N} I\{Y_i > x\} \exp(-sY_i + \widehat{\psi}_{method}(s)), \tag{14}$$

respectively, where $Y^{(IS)}_{1,N} \ge Y^{(IS)}_{2,N} \ge \cdots \ge Y^{(IS)}_{N,N}$ are the ordered samples and $w_{i,N} = \frac{1}{N}\exp(-sY^{(IS)}_{i,N} + \widehat{\psi}_{method}(s))$.

Standard Monte Carlo would generate N identically distributed samples $Y^{(SMC)}_1,\ldots,Y^{(SMC)}_N$ from the approximated posterior predictive distribution $\widehat{g}^{(p)}_{method}$. Denote $Y^{(SMC)}_{1,N} \ge Y^{(SMC)}_{2,N} \ge \cdots \ge Y^{(SMC)}_{N,N}$ be the ordered samples of the random draws of $Y^{(SMC)}_i$, $i = 1,\ldots,N$. Then, the unbiased estimates of the VaR and ES have the form

$$\widehat{VaR}(\alpha) = \hat{F}_Y^{-1}(1-\alpha) = Y^{(SMC)}_{[N\alpha]+1,N},$$

and

$$\begin{aligned}
\widehat{ES}(\alpha) &= \frac{1}{\alpha}\int_{1-\alpha}^{1} \hat{F}_Y^{-1}(p)dp \\
&= \frac{1}{\alpha}\left(\sum_{i=1}^{[N\alpha]} \frac{Y^{(SMC)}_{i,N}}{N} + \left(\alpha - \frac{[N\alpha]}{N}\right)Y^{(SMC)}_{[N\alpha]+1,N}\right),
\end{aligned}$$

respectively. Here $\hat{F}_Y^{-1}(p) = \min\{x \in \mathbb{R} : \hat{F}_Y(x) \ge p\}$ is the empirical quantile function for $p \in (0,1)$.

4 Data Analysis

The daily price of the Nikkei 225 index p_t for March 1, 2010 - March 15, 2015 was corrected. The daily percent log returns are constructed from the price series $\{p_t\}$, such that $y_t = 100\log(p_t/p_{t-1})$. The total number of returns in the sample is 1,236.

How volatility is evaluated depends on the specification of the volatility dynamics models. Here, we consider the ARCH(1) and RiskMetrics models for the process y_t. The ARCH(1) parameter α_0 is assumed such that the sample variance is equal to the model variance as $\alpha_0 = S^2(1-\alpha_1)$, where S^2 is the

sample variance. In Figure 1, we plot the return series $\{y_t\}$ together with the estimated volatilities for the ARCH(1) and RiskMetrics model. As can be seen in this figure, the RiskMetrics model is sensitive to volatility changes.

We compute the VaR and ES for the risk levels $\alpha \in \{0.01, 0.05\}$. We choose a flat prior for both α_1 on the interval $[0, 1)$ in the ARCH model and λ on the interval $[0, 1)$ in the RiskMetrics model. The number of random samples for each simulation is $N = 1,000$. Figure 2 displays the kernel density plot for the ARCH(1) predictive density for $T = 1,236$ using Tierney and Kadane and Miyata approximation, together with its importance sampling distributions for $\alpha = 0.01, 0.05$. These figures show that the approximations give similar densities.

Table 1 shows the estimation results for the VaR and ES for Tierney and Kadane, Lindley, and Miyata approximation. From this table, we can see that the estimated tail risks are quite close, such that the difference of procedure does not affect the risk evaluation. Note that the estimated values acquired using the RiskMetrics model are smaller to those obtained using the ARCH model. This is because the volatility estimate for $\hat{\sigma}_{T+1}$ is quite different between these two models, as can be verified by Figure 1. Table 2 shows the relative computing time for the estimation by each model and method. The Miyata approximation requires less computing time than the other methods. However, Lindley approximation needs more computing time because it requires a third derivative of the kernel functions.

To compare the convergence performance of our proposed method with Standard Monte Carlo, we estimate the VaR and ES for each simulation sample from 100 to 1,000. The results are plotted in Figure 3 , where we use the Miyata approximation. We can see that the importance sampling estimation has small spread even though the smaller sample size, whereas the Standard Monte Carlo have a larger variability.

Table 1. Estimated VaR and ES

Method	$\widehat{VaR}(0.05)$	$\widehat{VaR}(0.01)$	$\widehat{ES}(0.05)$	$\widehat{ES}(0.01)$
ARCH (1)				
TK	2.294	3.189	2.790	3.565
Miyata	2.285	3.180	2.796	3.563
Lindley	2.253	3.181	2.763	3.556
RiskMetrics				
TK	1.378	1.936	1.703	2.188
Miyata	1.339	1.938	1.717	2.172
Lindley	1.395	1.931	1.711	2.161

Table 2. Relative computing time for calculating VaR and ES for each model and method.

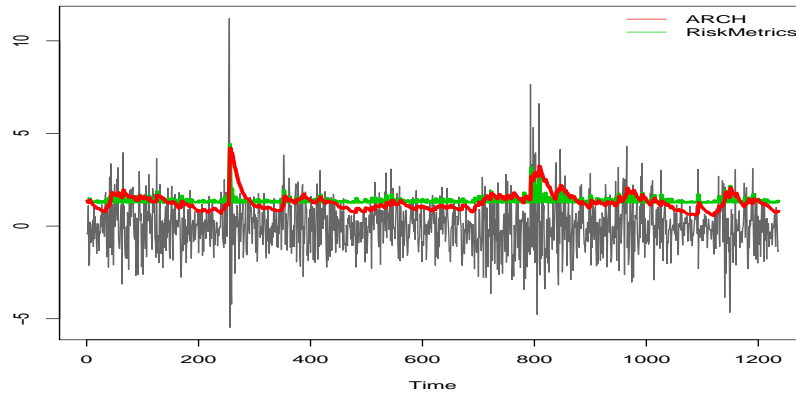| Method	TK	Miyata	Lindley
ARCH(1)	2.08	1.00	8.40
RiskMetrics	2.13	1.00	7.61

Fig. 1. Time series plot of negative log returns of Nikkei 225 stock market index together with estimated volatilities using the ARCH and RiskMetrics models.
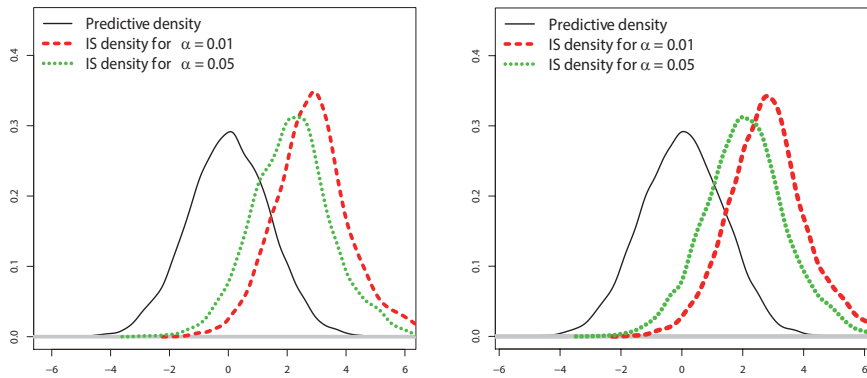

Fig. 2. Approximation of the predictive distribution together with importance sampling densities obtained by Tierney and Kadane approximation (left) and Miyata approximation (right). The dotted and dashed curves represent risk levels $\alpha = 0.01$ and $\alpha = 0.05$, respectively.
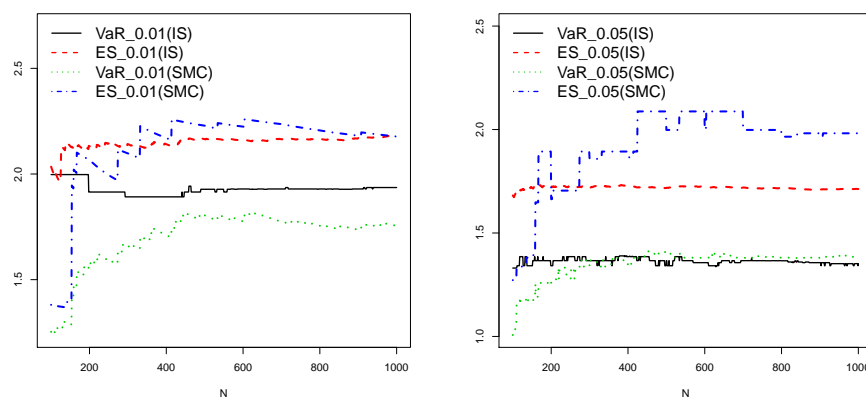
Fig. 3. Performance of convergence of importance sampling method with Miyata approximation together with Standard Monte Carlo method, for risk levels $\alpha = 0.05$ (left) and $\alpha = 0.01$ (right).

5   Conclusion

We performed a tail risk estimation using a Bayesian importance sampling predictive distribution on the Nikkei 225 log returns. Although we did not find much difference between the estimated values, the Miyata approximation needed less computing time than the other two methods. Other GARCH models are also worth considering using the proposed method. In future research, we must evaluate the difference between the estimated values and the true values. Thus, several open questions remain to be investigated.

Acknowledgments

References

1. P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Thinking coherently, *Risk*, **10**, 68–71, 1997.
2. P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk, *Mathematical Finance*, **9**, 203–228, 1999.
3. Basel Committee on Banking Supervision. Fundamental Review of the Trading Book, *Bank for International Settlements*, 2012.
4. T. Bollerslev. Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327, 1986.

5. J. A. Bucklew. *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley, New York, 1990.

6. J.-C. Chen, D. Lu, J. S. Sadowsky, and K. Yao. On importance sampling in digital communications, Part I: Fundamentals IEEE, *IEEE Journal on Selected Areas in Communications*, **11**, 289–299, 1993.

7. J. Cotter and K. Dowd. Extreme spectral risk measures: An application to futures clearing house margin requirements, *Journal of Banking & Finance*, **30**, 3469–3485, 2006.

8. R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007, 1982.

9. J. Francq and J. M. Zakoian. *GARCH Models*, Wiley, New York, 2004.

10. P. Glasserman and Y. Wang. Counterexamples in importance sampling for large deviations probabilities, *The Annals of Applied Probability*, **7**, 731–746, 1997.

11. P. Glynn. Importance sampling for Monte Carlo estimation of quantiles, *Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation*, 180–185, 1996.

12. L. Hoogerherde and H. K. van Dijk. Bayesian forecasting of Value at Risk and Expected Shortfall using adaptive importance sampling, *International Journal of Forecasting*, **26**, 231–247, 2010.

13. P. Heidelberger. Fast simulation of rate events in queueing and reliability models, *ACM Transactions on Modeling and Computer Simulation*, **5**, 43–85, 1995.

14. S.-W. Lee and B. E. Hansen. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory*, **10**, 29–52, 1994.

15. D. V. Lindley. Approximate Bayesian methods, *Bayesian Statistics, eds*, 223–245, 1980.

16. J. Longerstaey. RiskMetrics technical document. Technical Report fourth edition, JP Morgan/Reuters, 2010.

17. Y. Miyata. Fully exponential Laplace approximations using asymptotic modes, *Journal of American Statistical Association*, **99**, 1037–1049, 2004.

18. Y. Miyata. Simple approximations to posterior expectations of nonlinear functions using posterior means, Working paper, 2015.

19. P. Muller. Computation of risk measures using importance sampling, Master Thesis submitted to Swiss Federal Institute of Technology Zurich, 2010.

20. M. K. P. So and C.-M. Wong. Estimation of multiple period expected shortfall and median shortfall for risk management, *Quantitative Finance*, **12**, 739–754, 2012.

21. L. Tierney and J. B. Kadane. Accurate approximations for posterior moments and marginal densities, *Journal of American Statistical Association*, **81**, 82–86, 1986.

22. L. Tierney, R. E. Kass, and J. B. Kadane. Fully exponential Laplace approximations to expectations and variances of nonpositive functions, *Journal of American Statistical Association*, **84**, 710–716, 1989.

23. A. A. Weiss. Asymptotic theory for ARCH models: Estimation and testing, *Econometric Theory*, **2**, 107–131, 1986.

24. Y. Yamai and T. Yoshiba. Value-at-risk versus expected shortfall: A practical perspective, *Journal of Banking & Finance*, **29** , 997–1015, 2005.

Multivariate Birnbaum-Saunders Distribution and its Applications*

Yury S. Khokhlov[1], Ekaterina N. Smirnova[2], and Maxim A. Filippov[3]

[1] Department of Mathematical Statistics, Lomonosov Moscow State University, Leninskie Gory, GSP-1, 119991, Moscow, Russia
(E-mail: `yskhokhlov@yandex.ru`)
[2] Department of Applied Informatics and Probability Theory, Peoples Friendship University of Russia, Ordzhonikidze str. 3, 117 198, Moscow, Russia
(E-mail: `sukmanova-kate@mail.ru`)
[3] Department of Mathematical Statistics and Systems Analysis, Tver State University, Zhelyabova str. 33, 170 100, Tver, Russia
(E-mail: `makson@inbox.ru`)

Abstract. We consider multivariate analog of Birnbaum-Saunders distribution and investigate some its properties. In bivariate case our definition coinside with definition by Kundu et.al (2010). We use this approach for estimation of ruin probability in multivariate collective risk model.
Keywords: Multivariate Birnbaum-Saunders distribution, multivariate collective risk model, ruin probability.

1 Univariate Birnbaum-Saunders distribution

The two-parameter Birnbaum-Saunders (BS) distribution was originally proposed by Birnbaum and Saunders [1] as a failure time distribution for fatigue failure caused under cyclic loading. Later Desmond provided a more general derivation. So in what follows we we use the technique from the paper Desmond[2].

The specimen in the form of a metal bar secured to the ends is subjected to a cyclic strain, which leads to the emergence of a crack. Let X_k be the length of the crack at time $k = 0, 1, \ldots$. Suppose the dynamics of this process can be described by the following relation:

$$X_{k+1} = X_k + \xi_{k+1} \cdot g(X_k) , \; k = 0, 1, \ldots , \tag{1}$$

where $X_0 = 0$, $g(x), x \geq 0$ is a continuous positive function, $g(0) \neq 0$, $\{\xi_k\}$ are independent identically distributed nonnegative random variables with finite second moments.

Due to relation (1) we have:

$$\sum_{k=0}^{r-1} \xi_{k+1} = \sum_{k=0}^{r-1} \frac{X_{k+1} - X_k}{g(X_k)} . \tag{2}$$

Next we take large $t > 0$ and fixed $s > 0$. Let $r = t \cdot s$, $E(\xi_k) = \mu/t$, $D(\xi_k) = \sigma^2/t$. Then we have the following relation:

$$\sum_{k=0}^{r-1} \frac{X_{k+1} - X_k}{g(X_k)} \approx \int_0^{X_r} \frac{dx}{g(x)} \, . \tag{3}$$

Now consider some fixed $h > 0$ and denote by τ the moment of first passage time of the level h by the sequence X_k.

Due to Central Limit Theorem we have

$$P(\tau > r = t \cdot s) = P(X_r < h) =$$

$$P\left(\int_0^{X_r} \frac{dx}{g(x)} < \int_0^h \frac{dx}{g(x)} =: a(h) \right) \approx$$

$$P\left(\sum_{k=0}^{r-1} \xi_{k+1} < a(h) \right) \approx \Phi\left(\frac{a(h) - (\mu/t) \cdot r}{(\sigma/\sqrt{t}) \cdot \sqrt{r}} \right) =$$

$$\Phi\left(\frac{a(h) - \mu \cdot s}{\sigma \cdot \sqrt{s}} \right) \, .$$

Denote $\sqrt{\theta}/\lambda = a(h)/\sigma$, $1/\lambda\sqrt{\theta} = \mu/\sigma$. Then for $t \to \infty$ and fixed $s > 0$ we get

$$P(\tau \leq r = t \cdot s) \to 1 - \Phi\left(\frac{1}{\lambda} \left(\sqrt{\frac{\theta}{s}} - \sqrt{\frac{s}{\theta}} \right) \right) \, . \tag{4}$$

The right hand of (4) is the comulative function of Birnbaum-Saunders distribution (BS).

Standard normal distribution is symmetric with respect zero. Then the right hand of (4) can be rewritten in the form:

$$\Phi\left(\frac{1}{\lambda} \left(\sqrt{\frac{s}{\theta}} - \sqrt{\frac{\theta}{s}} \right) \right) \, . \tag{5}$$

$\theta > 0 \ \lambda > 0$ are the paprameters of the distribution.

If $\theta = 1$ we have standard BS-distribution (with parameter λ!).

Let τ be a random variable, which has BS-distribution with parameters θ and λ, X be a random variable with standard normal distribution.

In many cases it is useful the following relation:

$$X \overset{d}{=} \frac{1}{\lambda} \left[\left(\frac{\tau}{\theta} \right)^{1/2} - \left(\frac{\tau}{\theta} \right)^{-1/2} \right] \, , \tag{6}$$

or

$$\tau \overset{d}{=} \theta \left(1 + \frac{\lambda^2}{2} X^2 + \lambda X \left(1 + \frac{\lambda^2}{4} X^2 \right)^{1/2} \right) \, . \tag{7}$$

In particular we can calculate the moments of random variable τ:

$$E(\tau) = \theta \left(1 + \frac{\lambda^2}{2} \right) \, , \tag{8}$$

$$D(\tau) = (\theta\lambda)^2 \left(1 + \frac{5\lambda^2}{4}\right) \ . \tag{9}$$

2   Multivariate Birnbaum-Saunders distribution

Now we want to define the multivariate analog of Birbaum-Sanders distribution using the analogous approach.

Let $\{\xi_n\}$ be a sequence of independent identically distributed random vectors in R^m with finite second moments and all components of these vectors are nonnegative. Next consider vector-valued function

$$g(x) = (g_1(x_1), \ldots, g_m(x_m))^T, x \in R^m \ ,$$

whose components are continuous positive functions.

Similarly to univariate case (see Desmond[2]) we consider the multivariate process of crack extension in R^m:

$$X_{k+1} = X_k + \xi_{k+1} \circ g(X_k) \ , \ \ X_0 = 0, \tag{10}$$

where "$\circ$" is the operation of termwise multiplication of vectors.

Due to relation (10) we have:

$$\sum_{k=0}^{r-1} \xi_{k+1} = \sum_{k=0}^{r-1} \frac{X_{k+1} - X_k}{g(X_k)} \ . \tag{11}$$

Again we consider all operations with vectors as termwise. Next we take some real $t > 0$ large enough and fixed vector $s \in R^m$ with positive components. Let $r = t \cdot s$ and random vectors ξ_k have vector of means $\mu/t = (\mu_1, \ldots, \mu_m)^T/t$ and covariance matrix $A/t = (a_{ij}/t)$.

Again we have

$$\sum_{k=0}^{r-1} \frac{X_{k+1} - X_k}{g(X_k)} \approx \int_0^{X_r} \frac{dx}{g(x)} \ . \tag{12}$$

Fix some vector

$$h = (h_1, \ldots, h_m)^T$$

with positive components and consider random vector

$$\tau = (\tau_1, \ldots, \tau_m)^T \ ,$$

where τ_j is the moment of first passage time of the level h_j by the sequence $X_{k,j}$.

Let $r = t \cdot s$ and define vector $a(h) = (a_1(h_1), \ldots, a_m(h_m))^T$ with components

$$a_j(h_j) = \int_0^{h_j} \frac{dx_j}{g_j(x_j)} \ .$$

Below we consider all operation for vector as termwise. Then due to multivariate Central Limit Theorem we get:

$$P(\tau > r) = P(X_r < h) =$$

$$P\left(\int_0^{X_r} \frac{dx}{g(x)} < \int_0^h \frac{dx}{g(x)} =: a(h)\right) \approx$$

$$P\left(\sum_{k=0}^{r-1} \xi_k < a(h)\right) \approx \Phi_R\left(\frac{a(h) - \mu/t \circ r}{\sigma/\sqrt{t \cdot r}}\right) =$$

$$\Phi_R\left(\frac{a(h) - \mu \cdot s}{\sigma \cdot \sqrt{s}}\right) \ ,$$

where $\Phi_R(x), x \in R^m$ is a cumulative distribution function of multivariate normal distribution in R^m with matrix of correlation coefficients R.

As in univariate case we consider the following parametrization:

$$\sqrt{\theta_j}/\lambda_j = a_j(h_j)/\sigma_j \ , 1/\lambda_j\sqrt{\theta_j} = \mu_j/\sigma_j \ j = \overline{1, m} \ . \tag{13}$$

Then for large real t and fixed $s \in R^m$ we have

$$P(\tau > r = t \cdot s) \approx \Phi_R\left(\frac{1}{\lambda}\left(\sqrt{\frac{\theta}{s}} - \sqrt{\frac{s}{\theta}}\right)\right)$$

$$= \Phi_R\left(\frac{1}{\lambda_1}\left(\sqrt{\frac{\theta_1}{s_1}} - \sqrt{\frac{s_1}{\theta_1}}\right), \dots, \frac{1}{\lambda_m}\left(\sqrt{\frac{\theta_m}{s_m}} - \sqrt{\frac{s_m}{\theta_m}}\right)\right) \tag{14}$$

By the definition the right hand of relation (14) is the tail of multivariate Birnbaum-Saunders distribution.

It is easy to see that random variable τ_k has univariate BS-distribution with parameters λ_k, θ_k.

Let $X = (X_1, \dots, X_m)$ be random vectors with multivariate normal distribution zero means, unit variances and matrix of correlation coefficients R. Now consider new random vector $T = (T_1, \dots, T_m)$ such that

$$\tau_k = \theta_k\left(1 + \frac{\lambda_k^2}{2}X^2 + \lambda_k X\left(1 + \frac{\lambda_k^2}{4}X^2\right)^{1/2}\right) \ , \tag{15}$$

or

$$X_k = f_k(\tau_k) := \frac{1}{\lambda_k}\left[\left(\frac{\theta_k}{\tau_k}\right)^{1/2} - \left(\frac{\tau_k}{\theta_k}\right)^{-1/2}\right] \ , \tag{16}$$

It is easy to see that the functions $x_k = f_k(t_k)$ are strictly decreasing. Using this remark, relation (14) and the properties of multivariate normal distribution we get that this random vector τ has multivariate BS-distribution with parameters $\lambda = (\lambda_1, \dots, \lambda_m)$, $\theta = (\theta_1, \dots, \theta_m)$ and R.

Some years ago in paper Kundu *et al.*[3] it has been proposed the following variant of bivariate BS-distribution. Random vector $\tau = (\tau_1, \tau_2)$ has bivariate

BS-distribution with parameters λ, θ if its cumulative distribution function has the form

$$P(\tau_1 \leq t_1, \tau_2 \leq t_2) = \Phi_\rho \left[\frac{1}{\lambda_1} \left(\sqrt{\frac{t_1}{\theta_1}} - \sqrt{\frac{\theta_1}{t_1}} \right), \left(\sqrt{\frac{t_2}{\theta_2}} - \sqrt{\frac{\theta_2}{t_2}} \right) \right] . \quad (17)$$

Let some random vector $\tau = (\tau_1, \tau_2)$ has the distribution (14). Then

$$P(\tau_1 \leq t_1, \tau_2 \leq t_2) = P(Z_1 \leq -f_1(t_1), Z_2 \leq -f_2(t_2)) =$$

$$= \Phi_\rho \left[\frac{1}{\lambda_1} \left(\sqrt{\frac{t_1}{\theta_1}} - \sqrt{\frac{\theta_1}{t_1}} \right), \left(\sqrt{\frac{t_2}{\theta_2}} - \sqrt{\frac{\theta_2}{t_2}} \right) \right] .$$

So our definition in bivariate case coinside with definition in the paper Kundu *et al.*[3]. In this interesting paper it has been obtained some important properties this distribution. We remind some of them. We will write $(\tau_1, \tau_2) \sim BVBS(\lambda_1, \theta_1, \lambda_2, \theta_2, \rho)$ if (τ_1, τ_2) has bivariate BS-distribution with corresponding parameters. In the paper Kundu *et al.*[3] it has been proved the following

Theorem 1. . *If* $(\tau_1, \tau_2) \sim BVBS(\lambda_1, \theta_1, \lambda_2, \theta_2, \rho)$, *then*
  1) $(\tau_1^{-1}, \tau_2^{-1}) \sim BVBS(\lambda_1, \theta_1^{-1}, \lambda_2, \theta_2^{-1}, \rho)$,
  2) $(\tau_1^{-1}, \tau_2) \sim BVBS(\lambda_1, \theta_1^{-1}, \lambda_2, \theta_2, -\rho)$,
  3) $(\tau_1, \tau_2^{-1}) \sim BVBS(\lambda_1, \theta_1, \lambda_2, \theta_2^{-1}, -\rho)$.

Moreover in our case it can be proved that any subvector has multivariate BS-distribution.

3 Multivariate collective risk model

In this section we use the analogous approach for estimation of ruin probability in multivariate risk model.

In actuarial mathematics great importance has always been given to estimation of ruin probability of insurer. The multivariate collective risk model allows to consider the dependence between claims of different kinds of insurance, operated by insurance company. Claims happened in different kinds of insurance are mutually dependent very often, that affects the process of changing the value of the insurance reserve. The dynamics of reserve process can be written in the following form:

$$U(t) = u + c \cdot t - S(t) ,$$

where $U(t) = (U_1(t), \ldots U_m(t))$ – the process of insurer reserve; $u = (u_1, \ldots, u_m)$ – the initial insurer capital distributed between m kinds of insurance; $c = (c_1, \ldots, c_m)$ – vector of intensity of premiums incoming for each kind of insurance; $S(t) = (S_1(t), \ldots, S_m(t))$ – multivariate process of total payments for each kind of insurance up to the moment t.

Let us consider the multivariate index $i = (i_1, \ldots, i_m)$, where i_k equals to 0 or 1. $i_k = 1$ when claims of the k-th kind were paid and 0 in other

case. Denote by I the set of all possible values of i and $I_k = \{i : i_k = 1\}$. Let $N^{(i)}(t), t \geq 0$ be the number of claims happened up to the moment t and having the structure corresponding to the index i. We assume that $N^{(i)}$ are Poisson processes with parameters $\lambda^{(i)}$ and they are independent for different i. Define $N(t) = \sum_{i \in I} N^{(i)}(t)$ – total number of claims happened up to the moment t. N is Poisson process with parameter $\lambda = \sum_{i \in I} \lambda^{(i)}$. Let $(\varepsilon_j, j \geq 1)$ be a sequence of i.i.d. random vectors with values in I such that $P(\varepsilon_j = i) = \lambda^{(i)}/\lambda$ which is independent of $\{N^{(i)}, i \in I\}$. In our model

$$S_k(t) = \sum_{j=1}^{N(t)} \sum_{i \in I_k} I(\varepsilon_j = i) \cdot X_{j,k}^{(i)} ,$$

where $X_{j,k}^{(i)}$ is the part of the claim of structure i corresponding to k-th kind of insurance. We assume that processes $N^{(i)}$ and the sequences of claims $(X_j^{(i)})$ are independent. More details about such model see in the paper Ivanova[4].

Let τ_k be the moment of ruin in k-th kind of insurance. It is very important for the insurer that each kind of insurance is not lossmaking. So we define the ruin moment of the insurance company as following $\tau = \min\{\tau_k, k = \overline{1,m}\}$.

It is very difficult to find exact distribution of these random variables (especially in multivariate case). So as usual we try to find asymptotic distribution for large times. More exactly we want to find

$$P(\tau_1 > t_1, \ldots, \tau_m > t_m)$$

for the case $\min(t_1, \ldots, t_m) \to \infty$. In what follows we assume that $0 < t_1 < t_2 < \ldots < t_m$.

First we find the asymptotic distribution payments process $\boldsymbol{S}(\boldsymbol{t}), \boldsymbol{t} = (t_1, \ldots, t_m)$, which can be decomposed in the following form:

$$\boldsymbol{S}(\boldsymbol{t}) := (S_1(t_1), S_2(t_2), \ldots, S_m(t_m)) =$$

$$(S_1(t_1), S_2(t_1), \ldots, S_m(t_1)) + (0, S_2(t_2) - S_2(t_1), \ldots, S_m(t_2) - S_m(t_1)) +$$

$$\ldots + (0, 0, \ldots, 0, S_m(t_m) - S_m(t_{m-1})) .$$

Denote

$$\boldsymbol{S}_{t_1} = (S_1(t_1), S_2(t_1), \ldots, S_m(t_1)) ,$$

$$\boldsymbol{S}_{t_1, t_2} = (0, S_2(t_2) - S_2(t_1), \ldots, S_m(t_2) - S_m(t_1)) ,$$

$$\vdots$$

$$\boldsymbol{S}_{t_{m-1}, t_m} = (0, 0, \ldots, 0, S_m(t_m) - S_m(t_{m-1})) ,$$

$$X_{j,k}^* = \sum_{i \in I_k} I(\varepsilon_j = i) \cdot X_{j,k}^{(i)} .$$

It follows that random vectors

$$S_k(t_1) = \sum_{j=1}^{N(t_1)} X_{j,k}^*, 0 < t_1, k = \overline{1, m}$$

$$S_k(t_2) - S_k(t_1) = \sum_{j=N(t_1)+1}^{N(t_2)} X_{j,k}^*, 0 < t_1 < t_2, k = \overline{2, m}$$

$$\vdots$$

$$S_m(t_m) - S_m(t_{m-1}) = \sum_{j=N(t_{m-1})+1}^{N(t_m)} X_{j,m}^*, 0 < t_{m-1} < t_m,$$

$\boldsymbol{S}_{t_1}, \boldsymbol{S}_{t_1,t_2}, \ldots, \boldsymbol{S}_{t_{m-1},t_m}$ are independent.

It is well known the following variant of Central Limit Theorem for random sums.

Theorem 2. . *Let $X = (X_j, j \geq 1)$ be a sequence of i.i.d. random vectors in R^m with vector of means $a = (a_1, \ldots, a_m)^T$ and matrix of second moments $M = (m_{pq} = E(X_{jp}X_{jq}))$ and N be a random variable which has Poisson distribution with parameter λ and independent of sequence X. If $S = \sum\limits_{j=1}^{N} X_j$ and $\lambda \to \infty$ then random vector*

$$S_\lambda := \frac{S - \lambda \cdot a}{\sqrt{\lambda}}$$

has asymptotically multivariate normal distribution with zero means and covariance matrix M.

Now calculate the moments of random vector $X_j^* = (X_{j,k}^*)$.

$$a = (a_1, \ldots, a_m)^T = E(X_j^*) = \sum_{i \in I} E\left(I(\varepsilon_j = i) \cdot X_j^{(i)}\right) = \sum_{i \in I} \frac{\lambda^{(i)}}{\lambda} \cdot E(X_j^{(i)}) \,,$$

$$\sigma_{p,q} = E\left(X_{j,p}^* \cdot X_{j,q}^*\right) = \sum_{i \in I} \sum_{l \in I} E\left(I(\varepsilon_j = i) \cdot I(\varepsilon_j = l) \cdot X_{j,p}^{(i)} \cdot X_{j,q}^{(l)}\right) =$$

$$\sum_{i \in I} E\left(I(\varepsilon_j = i) \cdot X_{j,p}^{(i)} \cdot X_{j,q}^{(i)}\right) = \sum_{i \in I} \frac{\lambda^{(i)}}{\lambda} \cdot E\left(X_{j,p}^{(i)} \cdot X_{j,q}^{(i)}\right) \,.$$

Denote $\Sigma = (\sigma_{pq})$.

Let $t_j = s \cdot r_j$, so $\boldsymbol{t} = (t_1, t_2, \ldots, t_m)^T = (s \cdot r_1, s \cdot r_2, \ldots, s \cdot r_m)^T$, where $r_1 < r_2 < \ldots < r_m$.

Consider multivariate random process $\boldsymbol{S}_{t_1}$. Due to multivariate Central Limit Theorem for fixed r_1 and $s \to \infty$ random vector

$$\boldsymbol{S}_{t_1}^* = \frac{\boldsymbol{S}_{t_1} - E(X_j^*) \cdot \lambda \cdot s \cdot r_1}{\sqrt{\lambda \cdot s}}$$

has asymptotically multivariate normal distribution with zero means and co-variance matrix $r_1 \cdot \Sigma$.

Analogously it can be shown that for any $k = \overline{2, m}$ random vector $\boldsymbol{S}_{t_{k-1}, t_k}$ (under some normalization) has asymptotically multivariate normal distribution with zero means covariance matrix $(r_k - r_{k-1}) \cdot \Sigma^{(k-1)}$, where $\Sigma^{(k-1)}$ is constructed from Σ by taking zero instead of first $(k-1)$ rows and columns. Using the above representation and independence of summands we get the following

Theorem 3. . *Under above notations for* $\min(t_1, \ldots, t_m) \to \infty$ *random vector*

$$\boldsymbol{S}^*(\boldsymbol{t}) = \frac{\boldsymbol{S}(\boldsymbol{t}) - E(X_j^*) \circ r \cdot \lambda \cdot s}{\sqrt{\lambda \cdot s}} \ ,$$

where $a \circ b = (a_1 \cdot b_1, \ldots, a_m \cdot b_m)^T$, *has asymptotically multivariate normal distribution with zero means and covariance matrix* Σ_0, *whose elements have the form* $\sigma_{p,q} \cdot \min(r_p, r_q)$.

Now we come back to our initial problem. It is easy to see that

$$P(\tau_1 > t_1, \ldots, \tau_m > t_m) =$$

$$P(U_1(v_1) > 0, \ldots, U_m(v_m) > 0 \ , 0 \le v_1 < t_1, \ldots, 0 \le v_m < t_m) =$$

$$P(S_1(v_1) < u_1 + c_1 \cdot v_1, \ldots, S_m(v_m) < u_m + c_m \cdot v_m \ , 0 \le v_1 < t_1, \ldots, 0 \le v_m < t_m) \le$$

$$P(S_1(t_1) < u_1 + c_1 \cdot t_1, \ldots, S_m(t_m) < u_m + c_m \cdot t_m) \ .$$

For large t_k and u_k we come again (under some reparametrization!) to multi-variate BS-distribution and get some lower bound for ruin probability.

References

1. Birnbaum Z.W., Saunders S.C. A new family of life distribution. *J.Appl.Probab.*, 6, 319-327, 1962.
2. Desmond A. Stochastic models of failure in random environments. *Canadian J. of Statist.*, 13, 171-183, 1985.
3. Debasis Kundu, N. Balakrishnan, A. Jamalizadeh. Bivariate Birnbaum-Saunders distribution and associated inference. *Journal of Multivariate Analysis*, 101, 113-125, 2010.
4. N. L. Ivanova and Yu. S. Khokhlov, Multidimensional collective risk model. *Journal of Mathematical Sciences*, 146, 4, 6000-6007, October 2007.

Financial News Classification Based on Topographic Independent Component Analysis: Optimization on the Stiefel Manifold

Akiko Kitao[1], Takayuki Shiohama[2], and Hiroyuki Sato[3]

[1] Graduate School of Engineering, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo, Japan
(Eamil: `akiko527@ms.kagu.tus.ac.jp`)
[2] Department of Management Science, Faculty of Engineering, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo, Japan
(E-mail: `shiohama@ms.kagu.tus.ac.jp`)
[3] Department of Management Science, Faculty of Engineering, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo, Japan
(E-mail: `hsato@ms.kagu.tus.ac.jp`)

Abstract. Owing to the vast number of digitally stored text documents, the computational analysis of natural language text data has become increasingly important. Methods using statistical pattern recognition, such as latent semantic analysis (LSA) or independent component analysis (ICA), are completely unsupervised, and make use of second and higher order statistical data in their analyses. In this paper, we propose an algorithm using a trust-region method for a topographic independent component analysis (TICA). By applying our proposed method, we classify financial news using TICA. Furthermore, we make predictions about the stock market through this application of our method to the classification of financial news.
Keywords: text classification, topographic independent component analysis, financial news, stock price prediction.

1 Introduction

It is widely accepted that financial news provides a key source of investment information. Owing to the vast number of financial news volumes and sources, computational analysis of the relevant news is required by market investors and traders in order to make good and timely decisions. Although news-based trading has long been a part of investment decision making, it has become possible only in the recent years to quantify the impact of news events on financial markets by using natural language processing (NLP) technology. With this technology being available, it is easier to analyze financial news to gain a good insight into the current economic and financial situations.

Stock market predictions based on news content present an attractive research topic. One application of text mining is the discovery and exploitation of relationships between documented text and external sources of information

such as time stamped streams of data; specifically, stock market quotes. Predicting the movements of stock prices based on the contents of news articles is one of the many possible applications of text mining techniques. Literature on the application of computational methods to textual data for making financial market prediction includes Bollen *et al.* [4], Schumaker and Chan [14], Kumar *et al.* [10], and Nikolaos and Markellos [12]. A good review of recent results on the analysis of financial news can be found in the monograph of Mitra and Mitra [11].

Text classification is also an important part of text mining. Supervised learning for text classification begins with a training set of documents that are already labeled with a class. For example, the sentiments of news categories are "positive", "negative", or "neutral". In this paper, we consider a framework for traditional topic-based categorization of financial news using a topographic independent component analysis (TICA) model with unsupervised learning. This clustering process could help to illuminate the structure of the financial news, thus aiding the study of the relationships between the words featuring in news items and financial markets.

Methods using statistical pattern recognition, such as latent semantic analysis (LSA) or independent component analysis (ICA), are completely unsupervised, and make use of second and higher order statistical data in their analysis, such as, tree based learning (Bach and Jordan [2,3]) and independent subspace analysis (Theis [15]). The LSA approach is based on the summarization of a term by document matrix, i.e., a count of how often a given set of terms occur in the set of documents under analysis. As in ICA, the term by document matrix is considered to be a linear mixture of a set of independent sources, each activating its characteristic semantic network.

TICA is another method for multidimensional ICA, which can be seen as a generalization of the ICA model. Instead of using clusters to model dependencies between components, TICA applies a topographic representation to the independent components. The topographic order is also based on the mutual dependencies of the components, such that components that can be close to each other in the topographic model are assumed to be statistically dependent, and vice versa.

Various topologies can be used as structural models for TICA. As a measure for modeling the dependencies between components, Hyvärinen and Köster [9] and Hoyer *et al.* [8] proposed the notion of the correlation of energies, which measures higher-order correlations instead of independence.

The algorithms for (T)ICA can be roughly divided into two categories. The algorithms in the first category rely on batch computations for minimizing or maximizing some relevant criterion functions. The second category contains adaptive algorithms, often based on stochastic gradient methods, which may have implementations in neural networks. In this paper, we present a method for solving the TICA problem using Riemannian optimization, as introduced by Absil *et al.* [1] and Edelman *et al.* [6]. Sato and Iwai [13] recently proposed an optimization algorithm that solves the singular value decomposition problem on two Stiefel manifolds of different sizes.

Our contribution highlights the 30 basis vectors that result from applying TICA to the extraction of features from financial news texts. Furthermore, this TICA is obtained using the Riemannian optimization method.

The rest of this paper is organized as follows. In Section 2, topographic ICA is briefly explained. In Section 3, we propose a new optimization algorithm for TICA, and show that our new algorithm can improve existing methods. In Section 4, the results of financial news classifications based on TICA are presented. In Section 5, we conclude the paper.

2 Methodology

ICA is a statistical model in which the observed data is expressed as a linear combination of underlying latent variables, which are usually called signals, and are denoted by $\boldsymbol{s}$. The signals are assumed to be non-Gaussian and mutually independent. Let the observed mixture signals be denoted by $\boldsymbol{x}$. The ICA model assumes that $\boldsymbol{s}$ and $\boldsymbol{x}$ have zero-mean and finite covariance, and that the mixture signal $\boldsymbol{x}$ can be decomposed into a mixing matrix $\boldsymbol{A}$ with a set of source signals $\boldsymbol{s}$, such that for the sample index $t = 1, 2, \ldots, T$, it holds that

$$\boldsymbol{x}^{(t)} = \boldsymbol{A}\boldsymbol{s}^{(t)},$$

where $\boldsymbol{x}^{(t)} = (x_1^{(t)}, \ldots, x_m^{(t)})^\top$ is the vector of observed random variables, $\boldsymbol{s}^{(t)} = (s_1^{(t)}, \ldots, s_n^{(t)})^\top$ is the vector of source signals, and the $m \times n$ mixing matrix $\boldsymbol{A}$ is an unknown full rank constant matrix. The columns of $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ are often called basis functions or basis vectors. Three forms of ICA exist, based on the relationship between the number of the data m and the hidden number of sources n. These three forms are called square ($n = m$), over-complete ($n > m$), and under-complete ($n < m$). The standard ICA model is the square ($n = m$) case, while in this paper we consider a model in the under-complete ($n < m$) case.

A limitation of ordinary ICA is its strong independence assumption, which is difficult to satisfy with real-world data. In order to capture the dependence between the components, Hyvärinen and Köster [9] proposed the topographic independent component analysis (TICA) model, to find higher-order correlations for the components using the correlation of energies method. We consider the situation where the sources $\boldsymbol{s}$ are generated according to

$$s_i = z_i \sigma \quad \text{for } i = 1, \ldots, n,$$

where z_i is an independent component with zero-mean and unit variance such that $E\{z_i z_j\} = 0$ for $i \neq j$ and σ is a common variance. The correlation of energies becomes

$$\text{cov}\left(s_i^2, s_j^2\right) = E(s_i^2 s_j^2) - E(s_i^2)E(s_j^2) \neq 0,$$

if s_i and s_j are close in the topography. The TICA model relaxes the independence assumption of the classical ICA approach, thus seeking components that

are as independent as possible. To define the joint density $\boldsymbol{s}^{(t)}$, the variances σ_i^2 of the s_i are not constant and modeled by a nonlinearity:

$$\sigma_i = \phi\left(\sum_{j=1}^{n} V_{ij} u_j\right),$$

where the u_j are the higher order independent components used to generate the variances, and ϕ describes some nonlinearity. Then the dependency between two signals originates from the correlation between their variances. In order to define our topography, the following neighborhood function for each component is necessary:

$$V_{ij} = \begin{cases} 1 & \text{if } |i-j| \leq L, \\ 0 & \text{if } |i-j| > L. \end{cases} \tag{1}$$

The standard probabilistic model for ICA takes the form

$$p(\boldsymbol{x}|\boldsymbol{A}) = \int_{\boldsymbol{s}} p(\boldsymbol{x}|\boldsymbol{s}, \boldsymbol{A}) \prod_{j=1}^{n} p_j(s_j) d\boldsymbol{s},$$

where $p(\boldsymbol{x}|\boldsymbol{s}, \boldsymbol{A})$ is the likelihood term. In the square noiseless case $(m = n)$, the sources are estimated using a linear model of the data, such that $\boldsymbol{s} = \boldsymbol{W}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{x}$. Using this assumption, the model above simplifies to

$$p(\boldsymbol{x}|\boldsymbol{W}) = \int_s \delta(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s}) \prod_{j=1}^{n} p_j(s_j) d\boldsymbol{s} = |\boldsymbol{W}| \prod_{j=1}^{n} p_j(\boldsymbol{w}_j^\top \boldsymbol{x}),$$

where $|\cdot|$ is the absolute value of the determinant, and $|\boldsymbol{W}|$ normalizes the change in volume by the transformation. For under-complete models, an approximate maximum likelihood learning rule can be derived by replacing the stochastic relation between the source variables and input variables using the pseudo-inverse of the mixing matrix. That is,

$$\hat{\boldsymbol{s}} = \boldsymbol{W}\boldsymbol{x} = \boldsymbol{A}^{\#}\boldsymbol{x}, \quad \text{with } \boldsymbol{A}^{\#} = \boldsymbol{A}^\top (\boldsymbol{A}\boldsymbol{A}^\top)^{-1}.$$

If the separating matrix $\boldsymbol{W} = (\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n)^\top$ is square $(m = n)$, this expression reduces to $\boldsymbol{A}^{\#} = \boldsymbol{A}^{-1}$. Using this approximation, we obtain the likelihood for (T)ICA in the under-complete case as

$$\log L(\boldsymbol{W}) \approx \sum_{t=1}^{T} \sum_{i=1}^{n} G\left(\sum_{j=1}^{n} V_{ij}(\boldsymbol{w}_j^\top \boldsymbol{x}^{(t)})^2\right) + T \log\left|\boldsymbol{A}^{\#}\right|, \tag{2}$$

where the scalar function G is obtained from the probability density function of $p(u)$. One possible example of the function $G(y)$ is given by $G(y) = -\alpha\sqrt{y} + \beta$ for $\alpha > 0$. The last term $\log\left|\boldsymbol{A}^{\#}\right|$ gives the scaling of the probability mass when the linear transformation given by $\boldsymbol{W}$ is performed. Here, we discard the last term in (2), because the term $\log\left|\boldsymbol{A}^{\#}\right|$ is basically included to make the

$\boldsymbol{w}_i$ more or less orthonormal, with the assumption that the $\boldsymbol{a}_i$ are orthonormal. As in the case of basic ICA, prewhitening of the data allows us to consider the $\boldsymbol{w}_i$ to be orthonormal.

The data, which are assumed to have zero-mean, are first whitened as $\boldsymbol{z} = \boldsymbol{Hx} = \boldsymbol{HAs}$. The original mixing matrix of the unwhitened data can be computed as $\boldsymbol{A} \approx \boldsymbol{H}^{-1}\boldsymbol{B}^{\#} = \boldsymbol{H}^{-1}(\boldsymbol{B}^{\top}\boldsymbol{B})^{-1}\boldsymbol{B}^{\top}$, where $\boldsymbol{B} = (\boldsymbol{HA})^{\#} = \boldsymbol{A}^{\#}\boldsymbol{H}^{-1}$. The rows of $\boldsymbol{A}^{\#}$ provide the filters in the original, not whitened space.

More precisely, given an input pattern $\boldsymbol{x}$, the activation of each second layer unit is $p_i(\boldsymbol{x}^{(t)}; \boldsymbol{W}, \boldsymbol{V}) = \sqrt{\sum_{j=1}^{n} V_{ij}(\sum_{k=1}^{m} W_{jk} x_k^{(t)})^2}$. TICA learns the parameter $\boldsymbol{W}$ through sparse feature representations in the second layer, by solving

$$\text{minimize}_{\boldsymbol{W}} \sum_{t=1}^{T} \sum_{i=1}^{n} p_i(\boldsymbol{x}^{(t)}; \boldsymbol{W}, \boldsymbol{V}), \quad \text{subject to} \quad \boldsymbol{W}\boldsymbol{W}^{\top} = \boldsymbol{I}_n,$$

where the input patterns $\{\boldsymbol{x}^{(t)}\}_{t=1}^{T}$ are whitened.

Classic ICA is obtained as a special case of the topographic model, by taking a neighborhood function V_{ij} that is equal to the Kronecker delta function $V_{ij} = \delta_{ij}$.

3 Riemannian Optimization for TICA

3.1 Trust-region method for TICA

In this section, we define $\boldsymbol{Y} = \boldsymbol{W}^{\top}$ so that $\boldsymbol{Y}$ is vertically long. We also regard the objective function in the previous problem as a function on $\boldsymbol{Y}$, which we denote by $f(\boldsymbol{Y})$. This leads to the following optimization problem.

Problem 1.

$$\text{minimize} \quad f(\boldsymbol{Y}) = \sum_{t=1}^{T} \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{n} V_{ij} \left(\sum_{k=1}^{m} Y_{kj} x_k^{(t)} \right)^2},$$

$$\text{subject to} \quad \boldsymbol{Y}^{\top}\boldsymbol{Y} = \boldsymbol{I}_n.$$

This problem is a constrained optimization problem on the Euclidean space $\mathbb{R}^{m \times n}$. Here, we note that the search space $\{\boldsymbol{Y} \in \mathbb{R}^{m \times n} | \boldsymbol{Y}^{\top}\boldsymbol{Y} = \boldsymbol{I}_n\}$ admits a manifold structure, called the Stiefel manifold, which we denote by $\text{St}(n, m)$. For further details see, for example, Edelman *et al.* [6]. The problem can then be reformulated as an unconstrained problem on the Stiefel manifold $\text{St}(n, m)$. The objective function can be also rewritten as

$$f(\boldsymbol{Y}) = \sum_{t=1}^{T} \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{n} V_{ij}(\boldsymbol{Y}^{\top}\boldsymbol{x}^{(t)})_j^2}$$

$$= \sum_{t=1}^{T} \sum_{i=1}^{n} \sqrt{(\boldsymbol{Y}^{\top}\boldsymbol{x}^{(t)})^{\top} \boldsymbol{V}_i'(\boldsymbol{Y}^{\top}\boldsymbol{x}^{(t)})} =: \sum_{t=1}^{T} \sum_{i=1}^{n} f_i^{(t)}(\boldsymbol{Y}),$$

where the matrix $\boldsymbol{V}_i'$ is defined as $\boldsymbol{V}_i' = \operatorname{diag}(V_{i1}, V_{i2}, \ldots, V_{in}) \in \mathbb{R}^{n \times n}$. Note that we have defined $f_i^{(t)}(\boldsymbol{Y}) := \sqrt{(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i' \boldsymbol{Y}^\top \boldsymbol{x}^{(t)}}$. Thus, we are led to the following Riemannian optimization problem.

Problem 2.

$$\text{minimize} \quad f(\boldsymbol{Y}) = \sum_{t=1}^{T} \sum_{i=1}^{n} \sqrt{(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i' \boldsymbol{Y}^\top \boldsymbol{x}^{(t)}},$$

$$\text{subject to} \quad \boldsymbol{Y} \in \operatorname{St}(n, m).$$

We shall develop the Riemannian trust-region method for this problem. See Absil *et al.* [1] for general theory of the Riemannian trust-region method.

In Euclidean optimization, we use the updating formula

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k, \tag{3}$$

where $\boldsymbol{x}_k$ and $\boldsymbol{d}_k$ are the current point and the search direction vector, at the k-th iteration. However, we cannot use Eq. (3) on Riemannian manifolds, since addition cannot be defined in general. An alternative way of reframing (3) is to retract the search direction vector onto the manifold in question. In particular, we can use the following updating formula on the Stiefel manifold $\operatorname{St}(n, m)$:

$$\boldsymbol{Y}_{k+1} = \operatorname{qf}(\boldsymbol{Y}_k + \boldsymbol{\xi}_k),$$

where $\boldsymbol{Y}_k \in \operatorname{St}(n, m)$ and $\boldsymbol{\xi}_k \in T_{\boldsymbol{Y}_k} \operatorname{St}(n, m)$ are the current point and the search direction tangent vector, respectively, and $\operatorname{qf}(\cdot)$ denotes the Q-factor of the QR decomposition of the matrix in the parentheses. In other words, if a matrix $\boldsymbol{M} \in \mathbb{R}^{m \times n}$ is decomposed into $\boldsymbol{M} = \boldsymbol{Q} \boldsymbol{R}$ with $\boldsymbol{Q} \in \operatorname{St}(n, m)$ and $\boldsymbol{R} \in S_{\mathrm{upp}}^{+}(n)$, then $\operatorname{qf}(\boldsymbol{M}) = \boldsymbol{Q}$, where $S_{\mathrm{upp}}^{+}(n)$ is the set of all upper triangular matrices with strictly positive diagonal elements. For more details about QR decomposition, we refer the reader to Golub and Van Loan [7]. Because the Q-factor of the QR decomposition of an $m \times n$ matrix is an orthonormal matrix, qf can be regarded as a map $\mathbb{R}^{m \times n} \to \operatorname{St}(n, m)$.

In order to develop the Riemannian trust-region method for Problem 2, we must investigate several geometric quantities on the Stiefel manifold $\operatorname{St}(n, m)$. Let $T_{\boldsymbol{Y}} \operatorname{St}(n, m)$ denote the tangent space of $\operatorname{St}(n, m)$ at $\boldsymbol{Y} \in \operatorname{St}(n, m)$. Because $\operatorname{St}(n, m) \subset \mathbb{R}^{m \times n}$, the tangent space $T_{\boldsymbol{Y}} \operatorname{St}(n, m)$ can be regarded as a subspace of $\mathbb{R}^{m \times n}$. Therefore, we can endow $\operatorname{St}(n, m)$ with a Riemannian metric $\langle \cdot, \cdot \rangle$, as

$$\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle_{\boldsymbol{Y}} := \operatorname{tr}(\boldsymbol{\xi}^\top \boldsymbol{\eta}), \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in T_{\boldsymbol{Y}} \operatorname{St}(n, m), \ \boldsymbol{Y} \in \operatorname{St}(n, m), \tag{4}$$

which is induced from the natural inner product $(\cdot, \cdot)$ on $\mathbb{R}^{m \times n}$. That is,

$$(\boldsymbol{X}_1, \boldsymbol{X}_2) := \operatorname{tr}(\boldsymbol{X}_1^\top \boldsymbol{X}_2) = \sum_{i=1}^{m} \sum_{j=1}^{n} (X_1)_{ij} (X_2)_{ij}, \quad \boldsymbol{X}_1, \boldsymbol{X}_2 \in \mathbb{R}^{m \times n}.$$

Under the Riemannian metric (4), the orthogonal projection $P_{\boldsymbol{Y}} : \mathbb{R}^{m \times n} \to T_{\boldsymbol{Y}} \operatorname{St}(n, m)$ is written as

$$P_{\boldsymbol{Y}}(\boldsymbol{X}) = \boldsymbol{X} - \boldsymbol{Y} \operatorname{sym}(\boldsymbol{Y}^\top \boldsymbol{X}), \quad \boldsymbol{X} \in \mathbb{R}^{m \times n}, \tag{5}$$

where sym$(\cdot)$ denotes the symmetric part of the matrix in the parentheses. That is, sym$(\boldsymbol{Z}) = (\boldsymbol{Z} + \boldsymbol{Z}^\top)/2$.

Next, we discuss the Euclidean gradient and Hessian of the objective function. Strictly speaking, we have to define the extension $\bar{f}$ of f defined on the whole of $\mathbb{R}^{m \times n}$. Specifically, we define

$$\bar{f}(\boldsymbol{Y}) := \sum_{t=1}^{T} \sum_{i=1}^{n} \bar{f}_i^{(t)}(\boldsymbol{Y}), \; \bar{f}_i^{(t)}(\boldsymbol{Y}) := \sqrt{(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i' \boldsymbol{Y}^\top \boldsymbol{x}^{(t)}}, \quad \boldsymbol{Y} \in \mathbb{R}^{m \times n}.$$

The directional derivative of $\bar{f}_i^{(t)}$ at $\boldsymbol{Y}$ in the direction of $\boldsymbol{\xi} \in \mathbb{R}^{m \times n}$ can be easily computed as

$$\mathrm{D}\bar{f}_i^{(t)}(\boldsymbol{Y})[\boldsymbol{\xi}] = \frac{(\boldsymbol{x}^{(t)})^\top \boldsymbol{\xi} \boldsymbol{V}_i' \boldsymbol{Y}^\top \boldsymbol{x}^{(t)}}{\bar{f}_i^{(t)}(\boldsymbol{Y})}.$$

Hence, we have

$$\mathrm{D}\bar{f}(\boldsymbol{Y})[\xi] = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\mathrm{tr}(\xi \boldsymbol{V}_i' \boldsymbol{Y}^\top \boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top)}{\bar{f}_i^{(t)}(\boldsymbol{Y})}$$

$$= \mathrm{tr}\left(\boldsymbol{\xi}^\top \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i'}{\bar{f}_i^{(t)}(\boldsymbol{Y})} \right).$$

It follows that the Euclidean gradient $\nabla \bar{f}(\boldsymbol{Y})$ of f at $\boldsymbol{Y}$ is

$$\nabla \bar{f}(\boldsymbol{Y}) = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i'}{\bar{f}_i^{(t)}(\boldsymbol{Y})}.$$

Furthermore, the Hessian $\nabla^2 \bar{f}(\boldsymbol{Y})$ acts on $\boldsymbol{\xi} \in \mathbb{R}^{m \times n}$ as

$$\nabla^2 \bar{f}(\boldsymbol{Y})[\boldsymbol{\xi}] = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top \boldsymbol{\xi} \boldsymbol{V}_i' \bar{f}_i^{(t)}(\boldsymbol{Y}) - \boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top \boldsymbol{Y} \boldsymbol{V}_i' \mathrm{D}\bar{f}_i^{(t)}(\boldsymbol{Y})[\boldsymbol{\xi}]}{\bar{f}_i^{(t)}(\boldsymbol{Y})^2}$$

$$= \sum_{t=1}^{T} \boldsymbol{x}^{(t)}(\boldsymbol{x}^{(t)})^\top \sum_{i=1}^{n} \frac{\left(\bar{f}_i^{(t)}(\boldsymbol{Y})^2 \boldsymbol{\xi} - \left((\boldsymbol{x}^{(t)})^\top \boldsymbol{\xi} \boldsymbol{V}_i' \boldsymbol{Y}^T \boldsymbol{x}^{(t)} \right) \boldsymbol{Y} \right)}{\bar{f}_i^{(t)}(\boldsymbol{Y})^3} \boldsymbol{V}_i'.$$

Because we regard the Stiefel manifold $\mathrm{St}(n, m)$ as a Riemannian submanifold of $\mathbb{R}^{m \times n}$ with the natural induced metric, the Riemannian gradient $\mathrm{grad}\, f(\boldsymbol{Y})$ and the Riemannian Hessian $\mathrm{Hess}\, f(\boldsymbol{Y})$ of the objective function f with respect to this metric can be computed as in Absil *et al.* [1]:

$$\mathrm{grad}\, f(\boldsymbol{Y}) = P_{\boldsymbol{Y}}(\nabla \bar{f}(\boldsymbol{Y})), \; \mathrm{Hess}\, f(\boldsymbol{Y})[\boldsymbol{\xi}] = P_{\boldsymbol{Y}}(\mathrm{D}(\mathrm{grad}\, f)(\boldsymbol{Y})[\boldsymbol{\xi}]), \quad (6)$$

where $\boldsymbol{\xi}$ is an arbitrary tangent vector at $\boldsymbol{Y} \in \mathrm{St}(n, m)$. Combining (6) with (5) yields that

$$\mathrm{Hess}\, f(\boldsymbol{Y})[\boldsymbol{\xi}] = P_{\boldsymbol{Y}}(\mathrm{D}(P \nabla \bar{f})(\boldsymbol{Y})[\boldsymbol{\xi}])$$

$$= P_{\boldsymbol{Y}} \left(-\boldsymbol{\xi} \,\mathrm{sym}(\boldsymbol{Y}^\top \nabla \bar{f}(\boldsymbol{Y})) - \boldsymbol{Y} \,\mathrm{sym}(\boldsymbol{\xi}^\top \nabla \bar{f}(\boldsymbol{Y})) + P_{\boldsymbol{Y}}(\nabla^2 \bar{f}(\boldsymbol{Y})[\boldsymbol{\xi}]) \right)$$

$$= P_{\boldsymbol{Y}} \left(\nabla^2 \bar{f}(\boldsymbol{Y})[\boldsymbol{\xi}] - \boldsymbol{\xi} \,\mathrm{sym}(\boldsymbol{Y}^\top \nabla \bar{f}(\boldsymbol{Y})) - \boldsymbol{Y} \,\mathrm{sym}(\boldsymbol{\xi}^T \nabla \bar{f}(\boldsymbol{Y})) \right),$$

where we have used the fact that $P_{\boldsymbol{Y}}^2 = P_{\boldsymbol{Y}}$.

In the Riemannian trust-region method, we make full use of geometric information relating to the objective function, such as the gradient and the Hessian. At each iterate $\boldsymbol{Y}_k$, we construct a quadratic model $\hat{m}_{\boldsymbol{Y}_k} : T_{\boldsymbol{Y}_k} \mathrm{St}(n, m) \to \mathbb{R}$ of the objective function f, as

$$\hat{m}_{\boldsymbol{Y}_k}(\boldsymbol{\xi}) = f(\boldsymbol{Y}_k) + \langle \mathrm{grad}\, f(\boldsymbol{Y}_k), \boldsymbol{\xi} \rangle_{\boldsymbol{Y}_k} + \frac{1}{2} \langle \mathrm{Hess}\, f(\boldsymbol{Y}_k)[\boldsymbol{\xi}], \boldsymbol{\xi} \rangle_{\boldsymbol{Y}_k}.$$

We minimize the quadratic model $\hat{m}_{\boldsymbol{Y}_k}$ in a trust-region with a radius $\Delta > 0$, which is defined by $\{\boldsymbol{\xi} \in T_{\boldsymbol{Y}_k} \mathrm{St}(n, m) \,|\, \|\boldsymbol{\xi}\|_{\boldsymbol{Y}_k} \leq \Delta\}$. To this end, we use the truncated conjugate gradient method introduced by Absil *et al.* [1]. The obtained solution $\boldsymbol{\xi}_k$ to the subproblem is then accepted or rejected, depending on whether $\boldsymbol{\xi}_k$ satisfactorily decreases the original function f. The algorithm of the trust-region method for Problem 2 is described by Algorithm 1.

Algorithm 1 Riemannian trust-region method for Problem 2

1: Choose parameters $\bar{\Delta} > 0$, $\Delta_0 \in (0, \bar{\Delta})$, $\rho' \in [0, \frac{1}{4})$, and an initial point $\boldsymbol{Y}_0 \in \mathrm{St}(n, m)$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3:     Solve the following trust-region subproblem to obtain $\boldsymbol{\xi}_k$:

$$\begin{aligned} \text{minimize} \quad & \hat{m}_{\boldsymbol{Y}_k}(\boldsymbol{\xi}), \\ \text{subject to} \quad & \|\boldsymbol{\xi}\|_{\boldsymbol{Y}_k} \leq \Delta_k, \ \boldsymbol{\xi} \in T_{\boldsymbol{Y}_k} \mathrm{St}(n, m). \end{aligned}$$

4:     Evaluate $\rho_k := \dfrac{f(R_{\boldsymbol{Y}_k}(0)) - f(R_{\boldsymbol{Y}_k}(\boldsymbol{\xi}_k))}{\hat{m}_{\boldsymbol{Y}_k}(0) - \hat{m}_{\boldsymbol{Y}_k}(\boldsymbol{\xi}_k)}$.
5:     **if** $\rho_k < \frac{1}{4}$ **then**
6:         $\Delta_{k+1} = \frac{1}{4} \Delta_k$.
7:     **else if** $\rho_k > \frac{3}{4}$ and $\|\boldsymbol{\xi}_k\|_{\boldsymbol{Y}_k} = \Delta_k$ **then**
8:         $\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$.
9:     **else**
10:        $\Delta_{k+1} = \Delta_k$.
11:     **end if**
12:     **if** $\rho_k > \rho'$ **then**
13:        $\boldsymbol{Y}_{k+1} = \mathrm{qf}(\boldsymbol{Y}_k + \xi_k)$.
14:     **else**
15:        $\boldsymbol{Y}_{k+1} = \boldsymbol{Y}_k$.
16:     **end if**
17: **end for**

Recall that we have defined $\boldsymbol{Y} = \boldsymbol{W}^\top$. Once we have obtained an optimal solution $\boldsymbol{Y}$ to Problem 2 using Algorithm 1, the corresponding solution $\boldsymbol{W}$ to the original TICA problem in Section 2 can also be determined as $\boldsymbol{W} = \boldsymbol{Y}^\top$.

3.2   Numerical experiments

In our experiments, we first prepared a 30×6502 term by document matrix $\boldsymbol{x}$, which we will describe and analyze in more detail in the next section. This $\boldsymbol{x}$

corresponds to the case with $m = 30$ and $T = 6502$. We set $n = 5$, and let $\boldsymbol{x}^{(t)}$ denote the t-th column of the matrix $\boldsymbol{x}$. Using the matrix $\boldsymbol{I}_{m,n} := \begin{pmatrix} \boldsymbol{I}_n \\ 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$ as the initial guess $\boldsymbol{Y}_0$, we applied our algorithm to Problem 2 by means of Manopt, which is a MATLAB toolbox for optimization on manifolds by Boumal *et al.* [5]. Let $\boldsymbol{Y}_*$ and $\boldsymbol{Y}_{\mathrm{MAT}}$ denote the estimated solutions to Problem 2 that are obtained by the proposed algorithm and the interior point method provided by MATLAB Optimization Toolbox, respectively. The values of f at $\boldsymbol{Y}_0$, $\boldsymbol{Y}_*$, and $\boldsymbol{Y}_{\mathrm{MAT}}$ are found to be

$$f(\boldsymbol{Y}_0) = 1.265 \times 10^3, \ f(\boldsymbol{Y}_*) = 5.128 \times 10^2, \ f(\boldsymbol{Y}_{\mathrm{MAT}}) = 5.711 \times 10^2,$$

implying that our proposed algorithm is more efficient.

4 Data Analysis

We obtained textual news data from Bloomberg between September 1 2014 and November 30 2014. The total number of textual news items is 6502, taken over 91 days. On average, there were 71.5 news items delivered per day. All of the textual news are in Japanese, and extracting meaningful terms yields a total number of 28,697 terms. In order to reduce the number of terms, we use singular value decomposition, and obtain an input matrix $\boldsymbol{x}$ of size 30×6502. The most common weighting function is the *term frequency-inverse document frequency* (*tf-idf*) weights, which multiplies *term frequency* and *inverse document frequency* weights. In this study, the term by document matrix is based on the *tf-idf* weights.

The latent features can be extracted by TICA by setting $n = 30$ and the neighborhood threshold $L \in \{0, 2, 4\}$, as in (1). Notice that $L = 0$ corresponds to the usual ICA model. For these values of L, we optimize $\boldsymbol{W}$ using the optimization algorithms described in the previous section.

In order to visualize the higher order correlations among latent signals, we plot the correlation matrices of squared estimated signals, as presented in Figure 1. According to this figure, we can see that the extracted feature signals have a strong correlation within the axis labels (11,12), (18,19,20), and (25,26) for TICA with $L = 4$, whereas the usual ICA model with $L = 0$ does not exhibit these higher order correlations.

Table 1 shows the words that have high values in the feature vector for some selected feature axes. As in Table 1 and Figure 1, we pick up five latent topic categories, which indicates that the representations of the feature vectors are similar when s_i and s_j are close.

The clustering of financial news stories refers to the task of the unsupervised grouping of data objects into predetermined categories or topics, based on some similarity measure. The lists of news topics were created by selecting the major news categories, such as foreign exchange rates, international current affairs, domestic stock market, on central bank monetary policy. The (T)ICA model uses the clustering on its independent components, in order to maximize inter-cluster independence and intra-cluster dependence. Therefore, the similarity
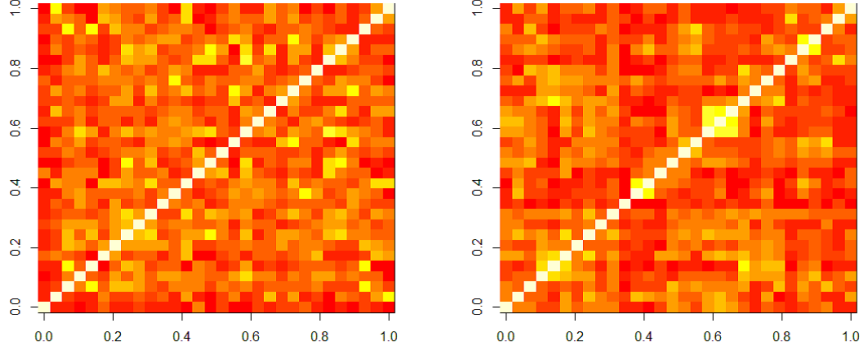
Fig. 1. Plots of the correlation matrices of squared latent signals with $L = 0$ (left) and $L = 4$ (right).

function used in this approach is based on statistical independence. The cosine similarity measure is applied, which is given by

$$C(\boldsymbol{w}_1, \boldsymbol{w}_2) = \frac{|\boldsymbol{w}_1^\top \boldsymbol{w}_2|}{|\boldsymbol{w}_1||\boldsymbol{w}_2|},$$

where $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are the documents' feature vectors.

The k-means cluster algorithm is applied, which divides the T observations $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(T)}\}$ into a set of k-clusters $\{C_1, C_2, \ldots, C_k\}$. For each given news document $\boldsymbol{x}^{(t)}$, we categorize this document with most closest feature topic axis.

Financial news can be split into regular synchronous announcements or expected news, and event-driven asynchronous announcements or unexpected news. Table 2 shows the clustering results for the feature vectors of 6502 financial news items, with $k = 10$ for the TICA with $L = 4$. A principal component analysis (PCA) is applied to detect the main directions of data variance. Figure 2 shows the biplot of the first and second principle components and their k-means clustering results. As seen from this figure, the largest category, whose topic is news relating to the world economy, lies close to the origin with small variability, indicating that the majority of the financial news items could not be classified into accurate topics. The regular and expected news items reporting on financial markets are classified effectively in this study.

Finally, we investigate the relationship between Japanese major market indices, including the Nikkei 225 stock market index, USD/JPY rates, and 10 year Japanese government bond yields, and the obtained feature axes. For this purpose, we perform simple linear regressions, such that

$$y_i^{(\text{Index})} = \alpha + \beta x_i^{(\text{feature axis})} + \varepsilon_i,$$

for the period from September 2014 to November 2014. The total sample size is 60 market opening days. The feature axis score for a given day is computed

as follows. For $j = 1, 2, \ldots, 30$ and day i, the set of the indices of whose news items is T_i, we have

$$x_i^{(\text{feature axis } j)} = \sum_{t \in T_i} C(\boldsymbol{x}^{(t)}, \boldsymbol{w}_j).$$

The scores are standardized to have zero-mean and unit variance. The parameter estimates with R^2 are presented in Table 3. According to this table, the world financial market axes with labels 12 and 25 have a significant impact on both the Japanese stock market and USD/JPY rates, whereas these axes do not show any effect on the Japanese bond market. The feature axes labeled with 11 and 26, whose financial news topics are crude oil and energy markets and European political issues, have a significant impact on the Japanese bond market.

Table 1. Five extracted financial news topics are shown whose topographic ordering between the categories are close, along with the words that occur in them with high absolute value scores.

Feature axis	Related topic and words with high values
Topic 1	Crude oil and energy and resource markets
11	Crude old, OPEC, ironstone, production, barrel
Topic 2	World stock market
12	China, RMB, India, SENSEX, LME, COMEX, stock market
Topic 3	International financial markets
18	FOMC, MOF, central bank, forward price, ruble, ECB
19	JPT, RMB, fund, LME, PIMCO, interest rate
20	USD, EURO, currency, bond price
Topic 4	European political issues
25	Ukraine, Russia, EU, independent, Scotland
Topic 5	Japanese and Asian economy
26	JPY, bond, FOMC, bp, interest rate, inflation, China, Hong Kong

Table 2. Results of k-means clustering with $k = 20$, and the relevant corresponding topics.

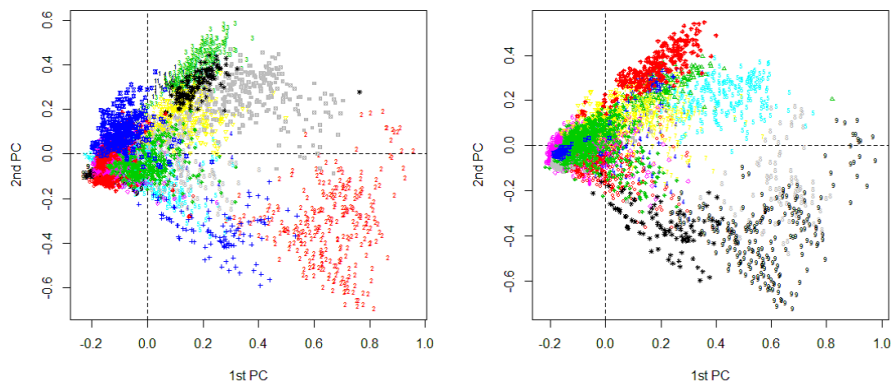| Clusters	Number of news stories
World's monetary policy of the central bank	413
EU and US bond markets	269
Japanese stock markets and currency market	256
World's stock markets	253
Asian stock markets	190
Performance of the world's major companies	170
Crude oil and OPEC	147
China and Hong Kong current news	120
World's residential real estate markets	65

Fig. 2. Biplots of the 1st and 2nd principal components of TICA with $L = 0$ (left) and $L = 4$ (right).

Table 3. Performance of market predictability with extracted financial news information.

Feature axis	11	12	18	19	20	25	26
y: Changes in 10 years JGB							
$\hat{\beta}$	-0.0123^*	0.0017	0.0033	-0.0048	0.0068	0.0010	-0.0105^*
t value	-2.6241	0.3514	0.6656	-0.9654	1.3920	0.2023	-2.1889
R^2	0.1061	0.0021	0.0076	0.0158	0.0323	0.0007	0.0763
y: Log returns for Nikkei 225 index							
$\hat{\beta}$	0.0074	-0.4028^*	-0.0174	-0.1078	-0.0492	0.5313^*	0.0055
t-value	0.0397	-2.2672	-0.0939	-0.5883	-0.2653	3.0937	0.0295
R^2	0.0000	0.0814	0.0002	0.0059	0.0012	0.1416	0.0000
y: Log returns for USD/JPY rate							
$\hat{\beta}$	0.0817	-0.1827^*	0.0845	-0.0494	0.0205	0.2132^*	0.0247
t-value	1.0381	-2.4131	1.0745	-0.6238	0.2587	2.8676	0.3108
R^2	0.0182	0.0912	0.0195	0.0067	0.0012	0.1242	0.0017

$*$ Statistically significant at 5% level.

5 Summary and Conclusions

In this study, we proposed an algorithm to solve (T)ICA optimization problems on the Stiefel manifold. TICA feature extraction was applied to textual data relating to financial news, and the results indicated that our proposed method is useful. By using TICA based feature extraction, we were able to classify financial news stories into some categories that are useful to predict market indices.

In order to apply news data effectively in trading decisions, we need to be able to identify news that is both relevant and current. As a means of doing so, latent sentiment analysis based on (T)ICA should be considered.

Acknowledgments

References

1. P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
2. F.R. Bach and M. I. Jordan. Tree-dependent component analysis, In: *Proc. of UAI*, 36–44, 2002.
3. F.R. Bach and M. I. Jordan. Beyond independent components: Trees and clusters, *Journal of Machine Learning Research*, **4**, 1205–1233, 2003.
4. J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock markets, *Journal of Computational Science*, **2**, 1–8, 2011.
5. N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds, *The Journal of Machine Learning Research*, **15**, 1455–1459, 2014.
6. A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints, *SIAM Journal on Matrix Analysis and Applications*, **20**, 303–353, 1998.
7. G. H. Golub and C. F. Van Loan. *Matrix Computations*, Johns Hopkins University Press, 2012.
8. P. O. Hoyer, A. Hyvärinen, and M. Inki. Topographic independent component analysis, *Neural Computation*, **13**, 1527–1558, 2001.
9. A. Hyvärinen and U. Köster. FastICA: A fast fixed-point algorithm for independent subspace analysis. In: *Proc. of ESANN*, 371–376, 2006.
10. R. B. Kumar, B. S. Kumar, and C. S. S. Prasad. Financial news classification using SVM, *International Journal of Scientific and Research Publications*, **2**, 2012.
11. G. Mitra and L. Mitra (eds). *The Handbook of News Analytics in Finance*, Wiley, New York, 2011.
12. J. Nikolaos and R. N. Markellos. Information demand and stock market volatility, *Journal of Banking & Finance*, **36**, 1808–1821, 2012.
13. H. Sato and T. Iwai. A Riemannian optimization approach to the matrix singular value decomposition, *SIAM Journal on Optimization*, **23**, 188–212, 2013.
14. R. Schumaker and H. Chen. A discrete stock prediction engine based on financial news, *IEEE Computer*, **43**, 51–56, 2010.
15. F. J. Theis. Towards a general independent subspace analysis, In: *Proc. of NIPS*, 2006.

A New Notion of
Bivariate Lack of Memory Property

Nikolai Kolev and Jayme Pinto

Department of Statistics, University of Sao Paulo, Brazil
(E-mail: `kolev.ime@gmail.com` and `jaymeaugusto@gmail.com`)

Abstract. Our main purpose is to establish necessary and sufficient conditions such that the sum of the components of hazard gradient vector of bivariate continuous non-negative distributions is a linear function of both arguments. The class of distributions obtained has as particular cases many classical ones. Joint survival distribution expression, restrictions for the marginals, geometric interpretations, generalizations and multivariate extension are presented.

Keywords: Bivariate lack of memory property, characterization, conditional failure rate, hazard gradient vector..

1   Introduction

The lack of memory property (LMP) of a univariate exponential distribution namely

$$P(X > x + t \mid X > t) = P(X > x) \quad \text{for all} \quad x \geq 0,\ t \geq 0, \tag{1}$$

plays a central role in survival analysis, reliability, insurance, finance and many other fields. Marshall and Olkin [11] extend LMP in a multivariate framework. In the bivariate case the authors look for continuous distributions such that

$$P(X_1 > x_1 + t, X_2 > x_2 + t \mid X_1 > t, X_2 > t) \quad \text{is independent of } t \tag{2}$$

for all $x_1, x_2 \geq 0$. Marshall and Olkin [11] characterize the bivariate exponential distribution (having singularity along the line $x_1 = x_2$) with exponential marginals. Denote by $S_{X_1,X_2}(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$ the joint survival function of (X_1, X_2). The Marshall-Olkin (MO) bivariate exponential distribution is given by

$$S_{X_1,X_2}(x_1, x_2) = \exp\{-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_3 \max(x_1, x_2)\} \quad \text{for all} \quad x_1, x_2 \geq 0, \tag{3}$$

where $\lambda_i > 0,\ i = 1, 2, 3$.

Apart from Marshall and Olkin's bivariate exponential distribution (3), other known solutions of (2) are the bivariate distributions obtained by Freund [4], Block and Basu [3], Proschan and Sullo [13], Friday and Patil [5] and all

distributions shown in Kulkarni [8], see Chapter 10 in Balakrishnan and Lai [1] for related discussion as well.

Relation (2) can be equivalently rewritten as a functional equation

$$S_{X_1,X_2}(x_1 + t, x_2 + t) = S_{X_1,X_2}(x_1, x_2)S_{X_1,X_2}(t, t) \qquad (4)$$

for all $x_1, x_2 \geq 0$ and $t > 0$ and it is known as a *bivariate lack-of-memory property*, to be abbreviated $BLMP_1$.

Following Johnson and Kotz [7], let us define the bivariate failure (hazard) rates as $r_i(x_1, x_2) = -\frac{\partial}{\partial x_i} \ln S_{X_1,X_2}(x_1, x_2)$, $i = 1, 2$. The hazard gradient vector

$$\mathbf{R}(x_1, x_2) = (r_1(x_1, x_2), r_2(x_1, x_2))$$

uniquely determines the joint distribution of (X_1, X_2), see Marshall [10].

Distributions possessing $BLMP_1$ satisfy the equation $r_1(x_1, x_2) + r_2(x_1, x_2) = a_0$, for all (x_1, x_2) such that $r_1(x_1, x_2)$ and $r_2(x_1, x_2)$ exist, where a_0 is a nonnegative constant, see Theorem 2 in Kulkarni [8] for details.

Johnson and Kotz [7] introduce another version of the bivariate LMP under the name *local lack of memory property*. Imposing the condition of "local constancy" of the failure rates $r_i(x_1, x_2)$, $i = 1, 2$, the authors enforce the conditional distributions $\{X_1 \mid X_2 > x_2\}$ and $\{X_2 \mid X_1 > x_1\}$ to keep the univariate LMP defined by (1). Equivalently,

$$P(X_i > x_i + y_i \mid X_j > x_j) = P(X_i > x_i \mid X_j > x_j)P(X_i > y_i \mid X_j > x_j) \quad (5)$$

for $i = 1, 2$ and $j = 3 - i$. Roy [14] rediscovered the local LMP and named it $BLMP_2$.

The only absolutely continuous distribution satisfying (5) is the *Gumbel's [6] type I bivariate exponential distribution*

$$S_{X_1,X_2}(x_1, x_2) = \exp\{-\lambda_1 x_1 - \lambda_2 x_2 - \theta\lambda_1\lambda_2 x_1 x_2\}, \quad x_1, x_2 \geq 0, \qquad (6)$$

where $\lambda_i > 0$, $i = 1, 2$ and $\theta \in [0, 1]$. For Gumbel's distribution (6) one can verify that $r_i(x_1, x_2) = \lambda_i + \theta\lambda_1\lambda_2 x_{3-i}$, $i = 1, 2$. Substituting $a_0 = \lambda_1 + \lambda_2$ and $a_1 = \theta\lambda_1\lambda_2$, the sum of bivariate hazard rates is $r_1(x_1, x_2) + r_2(x_1, x_2) = a_0 + a_1 x_1 + a_1 x_2$.

Our main goal is to link $BLMP_1$ and $BLMP_2$ in a new class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ of bivariate continuous distributions specified by

$$r(x_1, x_2) = r_1(x_1, x_2) + r_2(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 \qquad (7)$$

for all $x_1, x_2 \geq 0$, where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{a} = (a_0, a_1, a_2)$ is the parameter vector with nonnegative elements.

In Section 2 we characterize the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ by a general expression for the joint survival function of its members. Only certain marginal distributions are allowed in this class and the corresponding restrictions in terms of marginal densities are presented. Geometric interpretation of the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$, a possible generalization, multivariate extension and conclusions finalize the article.

2 Model specification, geometrical interpretation and extensions

Note that the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ defined by (7) is composed by nonnegative bivariate continuous distributions such that the sum of the components of hazard gradient is a linear function of both arguments x_1 and x_2. In Theorem 1 we obtain the expression of the joint survival function $S_{X_1,X_2}(x_1, x_2)$ for the elements of the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ and the restrictions for the constants a_0, a_1 and a_2 will be given in Theorem 2. The following Lemma equivalently characterizes the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ via a Cauchy functional equation.

Lemma 1. *The class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ of nonnegative bivariate continuous distributions specified by relation (7) can be equivalently defined by linear (additive) functionals*

$$A_1(x_1) = r(x_1, 0) - a_0 \quad \text{and} \quad A_2(x_2) = r(0, x_2) - a_0,$$

being the only continuous solutions of the functional equation $f(x+y) = f(x) + f(y)$.

Several known bivariate distributions follow property (7). Besides Marshall-Olkin's [11] bivariate exponential, Gumbel's [6] type I bivariate exponential distributions and the distributions presented in Kulkarni [8], other examples are:

(i) One obvious extension of the univariate LMP to the bivariate case is given by $S_{X_1,X_2}(x_1 + y_1, x_2 + y_2) = S_{X_1,X_2}(x_1, x_2)S_{X_1,X_2}(y_1, y_2)$ for all $x_1, x_2, y_1, y_2 \geq 0$. The only solution of this functional equation is given by $S_{X_1,X_2}(x_1, x_2) = \exp\{-b_1 x_1 - b_2 x_2\}$, for $b_1, b_2 > 0$, i.e. X_1 and X_2 are *independent and exponentially distributed* with parameters b_1 and b_2, see Marshall and Olkin [11]. One can get this (absolutely continuous) case from (7) by setting $a_0 = b_1 + b_2$ and $a_1 = a_2 = 0$;

(ii) The *bivariate Schur-constant law* $S_{X_1,X_2}(x_1, x_2) = \exp\{-x_1 - x_2\}$ can be obtained letting $a_0 = 2$ and $a_1 = a_2 = 0$ in (7). Barlow and Mendel [2] have characterized it in terms of *bivariate no-aging property* by relation $P(X_1 > x_1 + t \mid X_1 > x_1, X_2 > x_2) = P(X_2 > x_2 + t \mid X_1 > x_1, X_2 > x_2)$, which has applications in Bayesian reliability analysis;

(iii) Setting $a_0 = \lambda_3 > 0$, $a_1 = 2\lambda_1 > 0$ and $a_2 = 2\lambda_2 > 0$ in (7) corresponds to the joint survival function $S_{X_1,X_2}(x_1, x_2) = \exp\{-\lambda_1 x_1^2 - \lambda_2 x_2^2 - \lambda_3 \max(x_1, x_2)\}$, $x_1, x_2 \geq 0$, which has a *singular component along the line* $x_1 = x_2$. This distribution is *a member of the generalized Marshall-Olkin (GMO) distributions* introduced by Li and Pellerey [9].

When $S_{X_1,X_2}(x_1, x_2)$ is continuous, its vector hazard gradient $\mathbf{R}(x_1, x_2)$ is useful even when it does not exist everywhere in the interior of the set $\left\{(x_1, x_2) \in \mathbb{R}_+^2 \mid S_{X_1,X_2}(x_1, x_2) > 0\right\}$, where $\mathbb{R}_+^2$ is the first quadrant, see Marshall [10].

Theorem 1 holds for bivariate distributions belonging to the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$, whose survival functions possess all first partial derivatives and hence hazard

gradient vector $\mathbf{R}(x_1, x_2)$. A possible anomaly (singularity) that may happen when $x_1 = x_2$ is discussed in Remark 1.

Theorem 1. *If the first partial derivatives of $S_{X_1, X_2}(x_1, x_2)$ exist, relation (7) is fulfilled if and only if the corresponding joint survival function can be represented by*

$$S_{X_1, X_2}(x_1, x_2) = \begin{cases} S_{X_1}(x_1 - x_2) \exp\left\{-a_0 x_2 - a_1 x_1 x_2 - \frac{a_2 - a_1}{2} x_2^2\right\}, & \text{if } x_1 \geq x_2, \\ S_{X_2}(x_2 - x_1) \exp\left\{-a_0 x_1 - a_2 x_1 x_2 - \frac{a_1 - a_2}{2} x_1^2\right\}, & \text{if } x_2 \geq x_1. \end{cases}$$

(8)

Proof. The necessary part is obtained from property of path invariance of line integrals, valid for conservative vector fields. The sufficiency part is obtained from direct calculus of the hazard components $r_i(x_1, x_2) = -\frac{\partial}{\partial x_i} \ln S_{X_1, X_2}(x_1, x_2)$. ∎

Remark 1 (hazard vector elements in singularity case). The bivariate survival functions considered in Theorem 1 are not necessarily absolutely continuous (and therefore not differentiable almost everywhere). We do not exclude the possibility of existence of a singular component along the line $x_1 = x_2 = x \geq 0$ in our model, i.e. it may happen that $P(X_1 = X_2) > 0$. Therefore, in the presence of such a singular component, the function $S_{X_1, X_2}(x_1, x_2)$ is not differentiable in the set $\left\{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1 = x_2 = x\right\}$. In this case the hazard gradient vector $\mathbf{R}(x_1, x_2)$ does not exist along the line $x_1 = x_2$ and one should consider the derivative of $-\ln S_{X_1, X_2}(x, x)$ with respect to x. ∎

Theorem 1 characterizes bivariate distributions belonging to the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$, i.e. having joint survival function specified by relation (8), which imply some more restrictions on the margins of these distributions. In other words, $S_{X_1, X_2}(x_1, x_2)$ in (8) is a valid survival function only for certain marginal distributions of X_1 and X_2. Theorem 2 shows the corresponding constraints in terms of marginal densities when $a_1 + a_2 > 0$. The case $a_1 = a_2 = 0$ is detailed studied by Kulkarni [8].

Theorem 2. *Let X_i be a random variable with absolutely continuous density $f_{X_i}(x_i)$, $i = 1, 2$. Then $S_{X_1, X_2}(x_1, x_2)$ in (8) is a proper bivariate survival function if and only if there exist non-negative constants a_0, a_1 and a_2 with $a_1 + a_2 > 0$, such that*

$$A(x_i, x_j) - a_i x_j + \frac{d}{dx_i} \log f_{X_i}(x_i - x_j) + a_i [x_j A(x_i, x_j) - 1] \frac{S_{X_i}(x_i - x_j)}{f_{X_i}(x_i - x_j)} \geq 0 \quad (9)$$

where $A(x_i, x_j) = a_0 + a_i x_i + (a_j - a_i) x_j$ for all $x_i \geq x_j \geq 0$, $i \neq j$, $i, j = 1, 2$. Moreover,

$$[f_{X_1}(0) + f_{X_2}(0) - a_0] \sqrt{\frac{\pi}{2(a_1 + a_2)}} \exp\left\{\frac{a_0^2}{2(a_1 + a_2)}\right\} \left[1 - Erf\left(\frac{a_0}{\sqrt{2(a_1 + a_2)}}\right)\right] \in [0, 1],$$

(10)

where $Erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\{-t^2\} dt$. In addition, the survival function (8) is absolutely continuous if and only if $f_{X_1}(0) + f_{X_2}(0) = a_0$.

Proof. Necessary and sufficient conditions come from the fact that a bivariate distribution, with a (possible) singular component along the line $x_1 = x_2$, has to satisfy $\frac{\partial^2}{\partial x_1 \partial x_2} S_{X_1,X_2}(x_1, x_2) \geq 0$ for all $x_1, x_2 \geq 0$, where the mixed derivative exists. Also,

$$S_{X_1,X_2}(x_1, x_2) = (1 - \alpha)S_{X_1,X_2}^{ac}(x_1, x_2) + \alpha S_{X_1,X_2}^{si}(\max\{x_1, x_2\})$$

for $\alpha \in [0, 1]$, where $\alpha = P(X_1 = X_2)$ and $S_{X_1,X_2}^{ac}(x_1, x_2)$ and $S_{X_1,X_2}^{si}(\max\{x_1, x_2\})$ denote the absolutely continuous and singular components of $S_{X_1,X_2}(x_1, x_2)$, respectively. ∎

Direct calculus show that relation (10) represents $\alpha = P(X_1 = X_2)$, so that the obtained bivariate distribution is absolutely continuous if and only if $\alpha = f_{X_1}(0) + f_{X_2}(0) = a_U$ and will possess a singular component whenever $\alpha < f_{X_1}(0) + f_{X_2}(0)$. The lower bound for a_0 is obtained from (9) and (10) simultaneously, and is denoted by a_L.

The sum $r(x_1, x_2)$ specified by relation by (7) has a geometrical interpretation and is illustrated in Figure 1 below. The vertical axis $0z$ represents the values of the sum $r(x_1, x_2)$ of the hazard vector elements, i.e. $r_1(x_1, x_2) + r_2(x_1, x_2)$. One can localize the lower and upper bounds a_L and a_U of the constant a_0, obtained from (9) and (10). In general, all members of the
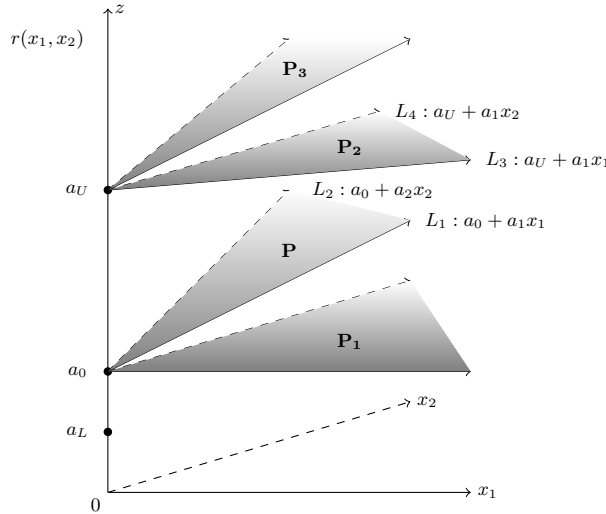

Fig. 1. Geometric representation of the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$

class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ lie in the first octant over the plane $\{P\}$ given analytically by $P : r(x_1, x_2) = a_0 + a_1x_1 + a_2x_2$. The plane $\{P\}$ crosses the axis $0z$ at the point $(0, 0, a_0)$ and crosses the planes $\{x_2 = 0\}$ and $\{x_1 = 0\}$ in lines $L_1 : a_0 + a_1x_1$ and $L_2 : a_0 + a_2x_2$, respectively. The lines $\{L_1\}$ and $\{L_2\}$ have always nonnegative inclination, since $a_1, a_2 \geq 0$.

One can recognize the following basic cases displayed on Figure 1.

(i) The plane $P_1 : r(x_1, x_2) = a_0$ crosses $0z$ in the point $(0, 0, a_0)$ and is *parallel to the plane* $\{z = 0\}$. The plane $\{P_1\}$ *represents distributions possessing* $BLMP_1$ *and exhibiting singularity along the line* $x_1 = x_2$;

(ii) The plane $P_2 : r(x_1, x_2) = a_U + a_1x_1 + a_1x_2$ is *the image of* $BLMP_2$ (characterizing the Gumbel's bivariate exponential distribution from (6)). The plane $\{P_2\}$ crosses the planes $\{x_2 = 0\}$ and $\{x_1 = 0\}$ in the lines $L_3 : a_U + a_1x_1$ and $L_4 : a_U + a_1x_2$, with equal inclination.
Observe that the image of all exchangeable absolutely continuous distributions belonging to the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ have similar characteristics;

(iii) The planes representing the *absolutely continuous bivariate distributions or those with independent marginals* cross the axis $0z$ at the point $(0, 0, a_U)$. The corresponding common lines with planes $\{x_2 = 0\}$ and $\{x_1 = 0\}$ have nonnegative inclination which may be *symmetric* (if $a_1 = a_2$), *asymmetric* (when $a_1 \neq a_2$), or *parallel to the plane* $\{z = 0\}$, whenever $a_1 = a_2 = 0$. The plane $P_3 : r(x_1, x_2) = a_U + a_1x_1 + a_2x_2$ is a typical example.

In the bivariate case, the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ can be generalized considering some expansion of summands in relation (7). One possibility is

$$r_1(x_1, x_2) + r_2(x_1, x_2) = a_0 + B_1(x_1) + B_2(x_2), \quad x_1, x_2 > 0,$$

where $B_i(x_i) = a_i(x_i)^{q_i}$ with nonnegative constants a_i and q_i, $i = 1, 2$.

The multivariate extension of (7) is readily obtained. Considering the $n-$variate nonnegative continuous random vector $\mathbf{X} = (X_1, ..., X_n)$, $n \geq 2$, the components of the hazard vector are given by $r_i(x_1, ...x_n) = -\frac{\partial}{\partial x_i} \ln [S_{X_1,...,X_n}(x_1, ...x_n)]$, $i = 1, 2, ...n$. Thus, the class $\mathcal{L}^{(n)}(\mathbf{x}; \mathbf{a})$ in n-dimensional case can be defined by relation

$$\sum_{i=1}^{n} r_i(x_1, ...x_n) = a_0 + a_1x_1 + \ldots + a_nx_n, \quad a_j \geq 0, \; j = 0, 1, ..., n.$$

3 Conclusion

We define the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ and characterize it by the following equivalent relations

$$\text{Lemma } 1 \Leftrightarrow \mathcal{L}(\mathbf{x}; \mathbf{a}) \Leftrightarrow (7) \Leftrightarrow (8).$$

It was shown that many bivariate distributions, in addition to those possessing $BLMP_1$ and $BLMP_2$, belong to $\mathcal{L}(\mathbf{x}; \mathbf{a})$. We do believe that a parallel methodology will be elaborated in bivariate and multivariate discrete setting as well.

Thus, the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ helps to deep the BLMP-notion giving possibility to model the aging phenomena in the complement to the "non-aging" one, which fixes the world on equations (4) or (5), via $BLMP_1$ and $BLMP_2$ correspondingly. So, our "world" and base is now equation (7), with graphical interpretation given on Figure 1. We are convinced that the class introduced is promising in modeling dynamic aging dependence, being much more realistic than the virtual "nonaging world".

The class $\mathcal{L}(\mathbf{x}; \mathbf{a})$ is very flexible, including symmetric and asymmetric continuous distributions, with possible singularity, and those which are positive or negative quadrant dependent. In Pinto [12] one can find many new bivariate distributions belonging to the class $\mathcal{L}(\mathbf{x}; \mathbf{a})$, see his Chapter 5. Such a huge variety of bivariate distributions would help to choose the "right" model consistent with the physical nature of the observations. The selection of bivariate distribution to be used should depend on considerations involving both the physical scenario at hand, and the properties of chosen distribution.

References

1. N. Balakrishnan and C-D. Lai. *Continuous Bivariate Distributions*, 2nd Edition, Springer, 2009.
2. R. Barlow and M. Mendel. De-Finetti type representations for life distributions. *Journal of the American Statistical Association* **87**, 1116–1122, 1992.
3. H.W. Block and A.P Basu. A continuous bivariate exponential extension. *Journal of the American Statistical Association* **69**, 1031–1037, 1974.
4. E. Freund. A bivariate extension of the exponential distribution. *Journal of American Statistical Association* **56**, 971–977, 1961.
5. D.S. Friday and G.P. Patil. A bivariate exponential model with applications to reliability and computer generation of random variables. *In* The Theory and Applications of Reliability, Volume 1, C.P. Tsokos and I.N. Shimi (eds.), 527–549. Academic Press, New York, 1977.
6. E. Gumbel. Bivariate exponential distributions. *Journal of American Statistical Association* **55**, 698–707, 1960.
7. N.L. Johnson and S. Kotz. A vector multivariate hazard rate. *Journal of Multivariate Analysis* **5**, 53–66, 1975.
8. H.V. Kulkarni. Characterizations and modelling of multivariate lack of memory property. *Metrika* **64**, 167–180, 2006.
9. X. Li and F. Pellerey. Generalized Marshall-Olkin distributions and related bivariate aging properties. *Journal of Multivariate Analysis* **102**, 1399–1409, 2011.
10. A.W. Marshall. Some comments on the hazard gradient. *Stochastic Processes and their Applications* **3**, 293–300, 1975.
11. A.W. Marshall and I. Olkin. A multivariate exponential distribution. *Journal of the American Statistical Association* **62**, 30–41, 1967.
12. J. Pinto. *Deepening the Notions of Dependence and Aging in Bivariate Probability Distributions*. PhD Thesis, Sao Paulo University, 2014.
13. F. Proschan and P. Sullo. Estimating the parameters of a bivariate exponential distribution in several sampling situations. In: Reliability and Biometry: Statistical Analysis of Life Lengths, F. Proschan and R.J. Serfling (eds.), 423–440, 1974.
14. D. Roy. On bivariate lack of memory property and a new definition. *Annals of the Institute of Statistical Mathematics* **54**, 404–410, 2002.

Hypotheses Testing on Covariance Structures: Comparison of Likelihood Ratio Test, Rao's Score Test and Wald's Score Test

Tõnu Kollo[1], Dietrich von Rosen[2], Marju Valge[1]

[1] *University of Tartu, Tartu, Estonia;* [2] *Swedish University of Agricultural Sciences, Uppsala, Sweden*

Abstract

For a normal population likelihood ratio test, Rao's score test and Wald's score test for testing covariance structures are compared in the situation when the number of variables and the sample size are growing. Expressions of all three test statistics are derived under the general null-hypothesis $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, using matrix derivative techniques. The special cases $\boldsymbol{\Sigma} = \gamma \mathbf{I}_p$ and $\boldsymbol{\Sigma} = \mathbf{I}_p$ are also under consideration. The tests are compared in a simulation experiment with sample sizes varying from 100 to 5000 and dimensionalities from 2 to 50. When the number of variables is growing Rao's score test behaves most adequately.

Keywords: Covariance structure, hypotheses testing, likelihood ratio test, Rao's score test, Wald's score test.

MSC Classification: 62H10, 62H15, 62F03, 62F05

1. Introduction
In data analysis and modelling one is often interested in hypotheses testing: hypotheses about mean or testing presence of a specific covariance structure. In a simple case the hypothized covariance matrix is an identity matrix, whereas in more complex situations e.g. when analysing spacial-temporal data the Kronecker product structure is often present. Probably the most commonly used test is the likelihood ratio test (LRT). However, it is known that when the number of parameters is greater than the sample size the likelihood ratio test will almost always reject the null hypothesis. In order to overcome this problem, corrections to the test have been made so that it could be used in a high-dimensional setup as well, see Bai, Jiang, Yao and Zheng (2009), for example. In practice one still carries on using the uncorrected test. In this paper a simulation experiment shows that instead of the likelihood ratio test or Wald's score test (WST) more consistent Rao's score test (RST) should be used to test a particular covariance structure.

In Section 2 notation is described. In Section 3 we will introduce the likelihood ratio test, Rao's score test and Wald's score test and derive formulas for

hypotheses testing under an assumption of multivariate normality. Both the general formula for testing $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ and specific formulae for the sphericity test and for testing if the covariance matrix is an identity matrix are derived. In Section 4 using simulations we compare the behaviour of all three tests in a growing dimensional situation, both under H_0 and the alternative. The effect of the sample size on the behaviour of the test statistics is also investigated. Finally, in Section 5 we summarize the results and draw some conclusions.

2. Notation

Later on we shall use matrix derivatives repeatedly, and use the definition of Kollo and von Rosen (2005), p. 127.

Definition 1. *Let the elements of $\mathbf{Y} \in \mathbb{R}^{r \times s}$ be functions of $\mathbf{X} \in \mathbb{R}^{p \times q}$. The matrix $\frac{d\mathbf{Y}}{d\mathbf{X}} \in \mathbb{R}^{pq \times rs}$ is called matrix derivative of $\mathbf{Y}$ by $\mathbf{X}$ in a set A, if the partial derivatives $\frac{dy_{kl}}{dx_{ij}}$ exist, are continuous in A, and*

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{\partial}{\partial \text{vec}\,(\mathbf{X})} \text{vec}\,'\mathbf{Y}$$

where

$$\frac{\partial}{\partial \text{vec}\,\mathbf{X}} = (\frac{\partial}{\partial x_{11}}, ..., \frac{\partial}{\partial x_{p1}}, \frac{\partial}{\partial x_{12}}, ..., \frac{\partial}{\partial x_{p2}}, ..., \frac{\partial}{\partial x_{1q}}, ..., \frac{\partial}{\partial x_{pq}})'$$

and $\text{vec}\,(\cdot)$ is the usual vectorization operator.

Further, the following properties of the matrix derivative are used (Kollo, von Rosen, 2005, p. 149):

1. $\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{I}_{pq}$;
2. $\frac{d\mathbf{Y}+\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Y}}{d\mathbf{X}} + \frac{d\mathbf{Z}}{d\mathbf{X}}$;
3. $\frac{d\mathbf{AXB}}{d\mathbf{X}} = \mathbf{B} \otimes \mathbf{A}'$;
4. When $\mathbf{Z} = \mathbf{Z}(\mathbf{Y})$, $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$ then $\frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Y}}{d\mathbf{X}} \frac{d\mathbf{Z}}{d\mathbf{Y}}$;
5. When $\mathbf{W} = \mathbf{YZ}$, $\mathbf{Z} \in \mathbb{R}^{s \times t}$ then $\frac{d\mathbf{W}}{d\mathbf{X}} = \frac{d\mathbf{Y}}{d\mathbf{X}}(\mathbf{Z} \otimes \mathbf{I}_r) + \frac{d\mathbf{Z}}{d\mathbf{X}}(\mathbf{I}_t \otimes \mathbf{Y}')$;
6. When $\mathbf{X} \in \mathbb{R}^{p \times p}$ then $\frac{d\mathbf{X}^{-1}}{d\mathbf{X}} = -\mathbf{X}^{-1} \otimes (\mathbf{X}')^{-1}$;
7. When $\mathbf{X} \in \mathbb{R}^{p \times p}$ then $\frac{d|\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}|\text{vec}\,(\mathbf{X}^{-1})'$, where $|\cdot|$ denotes the determinant;
8. $\frac{d\text{tr}(\mathbf{A}'\mathbf{X})}{d\mathbf{X}} = \text{vec}\,(\mathbf{A})$, where $\text{tr}(\cdot)$ denotes the trace function.

Let X be a continuous random variable with distribution $P_X(\boldsymbol{\theta})$ and density function $f_X(x, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of unknown parameters.

Definition 2. *The score vector of random variable X is given by the matrix derivative*

$$U(X, \boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}} \ln f(X, \boldsymbol{\theta}).$$

Definition 3. *The information matrix of random variable X is the covariance matrix of the score vector $U(X, \boldsymbol{\theta})$:*

$$\mathrm{I}(X, \boldsymbol{\theta}) = D(U(X, \boldsymbol{\theta})) = E(U(X, \boldsymbol{\theta})U'(X, \boldsymbol{\theta})).$$

Definition 4. *The Hessian matrix of random variable X is given by the second order matrix derivative*

$$\mathrm{H}(X, \boldsymbol{\theta}) = \frac{d^2}{d\boldsymbol{\theta}^2} \ln f(X, \boldsymbol{\theta}) = \frac{dU(X, \boldsymbol{\theta})}{d\boldsymbol{\theta}}.$$

Note that when the distribution of the random variable X is regular, then the information matrix of X can be calculated as an expectation:

$$\mathrm{I}(X, \boldsymbol{\theta}) = -\mathrm{E}(\mathrm{H}(X, \boldsymbol{\theta})).$$

Let $\mathcal{X} = (X_1, ..., X_n)$ denote a theoretical sample from the distribution $P_X(\boldsymbol{\theta})$. Then the log-likelihood of the sample is given by $l(\mathcal{X}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f(X_i, \boldsymbol{\theta})$ and the score function of the sample, $U(\mathcal{X}, \boldsymbol{\theta})$, is given by $U(\mathcal{X}, \boldsymbol{\theta}) = \sum_{i=1}^{n} U(X_i, \boldsymbol{\theta})$. The information matrix of the sample is given by $\mathrm{I}(\mathcal{X}, \boldsymbol{\theta}) = n \cdot \mathrm{I}(X, \boldsymbol{\theta})$.

3. Test statistics
3.1 Likelihood Ratio Test
For testing the hypothesis

$$\begin{cases} \mathrm{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0, \\ \mathrm{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0 \end{cases} \tag{1}$$

we use the likelihood ratio test (LRT) in logarithmic form:

$$LRT(\mathcal{X}, \boldsymbol{\theta}_0) = -2\ln\left(\frac{L(\mathcal{X}, \boldsymbol{\theta}_0)}{\max_{\boldsymbol{\theta}} L(\mathcal{X}, \boldsymbol{\theta})}\right) = -2[\ln L(\mathcal{X}, \boldsymbol{\theta}_0) - \ln L(\mathcal{X}, T_{\boldsymbol{\theta}}(\mathcal{X}))]$$

where $L(\cdot)$ is the likelihood function of X, $\mathcal{X} = (X_1, ..., X_n)$ is a random sample from $P_X(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_r)'$ and $T_{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$. When sample size $n \to \infty$ and H_0 holds then $LRT(\mathcal{X}, \boldsymbol{\theta}) \sim \chi_r^2$.

When under the null hypothesis there are k parameters $\boldsymbol{\theta}_1$ not fixed then the LRT in logarithmic form has the representation

$$LRT(\mathcal{X}, \boldsymbol{\theta}_0) = -2\ln\left(\frac{\max_{\boldsymbol{\theta}_1} L(\mathcal{X}, \boldsymbol{\theta}_0)}{\max_{\boldsymbol{\theta}} L(\mathcal{X}, \boldsymbol{\theta})}\right)$$

$$= -2[\ln L(\mathcal{X}, T_{\boldsymbol{\theta}_1}(\mathcal{X}))) - \ln L(\mathcal{X}, T_{\boldsymbol{\theta}}(\mathcal{X}))]$$

and $LRT(\mathcal{X}, \boldsymbol{\theta}) \sim \chi^2_{r-k}$ when sample size $n \to \infty$ and H_0 holds (see Rao (1973, §6e), for instance).

Let us present the likelihood ratio test statistic for testing hypothesis (1) in the multivariate normal case. Let $X \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)'$ and $\boldsymbol{\Sigma} > 0 : p \times p$ and let $\boldsymbol{\theta} = (\boldsymbol{\mu}', \text{vec}\,'(\boldsymbol{\Sigma}))'$. Then

$$f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

The likelihood ratio test statistic for the theoretical sample $\mathcal{X} = (X_1, ..., X_n)$ when testing hypothesis (1) has the following form

$$LRT(\mathcal{X}, \boldsymbol{\theta}_0) = -2\ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}|^{\frac{n}{2}}}{|\boldsymbol{\Sigma}_0|^{\frac{n}{2}}} \cdot \frac{\exp(\sum\limits_{i=1}^{n}(X_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1}(X_i - \boldsymbol{\mu}_0))}{\exp(\sum\limits_{i=1}^{n}(X_i - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1}(X_i - \widehat{\boldsymbol{\mu}}))}\right) \quad (2)$$

where $\widehat{\boldsymbol{\mu}} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ is the maximum likelihood estimator of $\boldsymbol{\mu}$ and $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \widehat{\boldsymbol{\mu}})(X_i - \widehat{\boldsymbol{\mu}})'$ is the maximum likelihood estimator of $\boldsymbol{\Sigma}$.

Proposition 1. *The likelihood ratio test statistic for testing* $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ *when no constraints are imposed on* $\boldsymbol{\mu}$ *has the following form:*

$$LRT(\mathcal{X}, \boldsymbol{\Sigma}_0) = n(\text{tr}(\boldsymbol{\Sigma}_0^{-1}\widehat{\boldsymbol{\Sigma}}) - p) - n\ln|\boldsymbol{\Sigma}_0^{-1}\widehat{\boldsymbol{\Sigma}}|, \quad (3)$$

where $\mathcal{X} = (X_1, ..., X_n)$ *is a sample from* $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *When* $n \to \infty$ *and* H_0 *holds then* $LRT(\mathcal{X}, \boldsymbol{\Sigma}_0) \sim \chi^2$ *with degrees of freedom* $df = \frac{p(p+1)}{2}$.
The proof can be found in Anderson (2003, §10.8).

Corollary 1. *The likelihood ratio test statistic for testing* $H_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ *when no constraints are imposed on* $\boldsymbol{\mu}$ *has the following form:*

$$LRT(\mathcal{X}, \mathbf{I}_p) = n(\text{tr}(\widehat{\boldsymbol{\Sigma}}) - p) - n\ln|\widehat{\boldsymbol{\Sigma}}|, \quad (4)$$

where $\mathcal{X} = (X_1, ..., X_n)$ *is a sample from* $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *When* $n \to \infty$ *and* H_0 *holds then the* $LRT(\mathcal{X}, \mathbf{I}_p) \sim \chi^2$ *with degrees of freedom* $df = \frac{p(p+1)}{2}$.
Proof. It follows directly from (3) by replacing $\boldsymbol{\Sigma}_0$ with $\mathbf{I}_p$.

Next, the LRT statistic for testing sphericity is given in Proposition 2. The derivation of this statistic can be found from Anderson (2003, ch. 10) or Bilodeau, Brenner (1999, §8.6), for example.
Proposition 2. *The likelihood ratio test statistic for testing* $H_0 : \boldsymbol{\Sigma} = \gamma\mathbf{I}_p$ *when no constraints are imposed on* $\boldsymbol{\mu}$ *has the following form:*

$$LRT(\mathcal{X}, \gamma\mathbf{I}_p) = np\ln\left(\frac{1}{p}\text{tr}(\widehat{\boldsymbol{\Sigma}})\right) - n\ln|\widehat{\boldsymbol{\Sigma}}|, \quad (5)$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\gamma > 0$ is unknown parameter. When $n \to \infty$ and H_0 holds then $LRT(\mathcal{X}, \gamma \mathbf{I}_p) \sim \chi^2$ with degrees of freedom $df = \frac{p(p+1)}{2} - 1$.

3.2 Rao's Score Test

Definition 5. *Rao's score test statistic (RST) for testing hypothesis (1) is given by*

$$RST(\mathcal{X}, \boldsymbol{\theta}_0) = U'(\mathcal{X}, \boldsymbol{\theta}_0) \cdot I(\mathcal{X}, \boldsymbol{\theta}_0)^{-1} \cdot U(\mathcal{X}, \boldsymbol{\theta}_0)$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a random sample from $P_X(\boldsymbol{\theta})$ and $\boldsymbol{\theta} = (\theta_1, ..., \theta_r)'$. When the sample size $n \to \infty$ and H_0 holds then $RST(\mathcal{X}, \boldsymbol{\theta}) \sim \chi_r^2$ (Rao, 1948).

Let us derive Rao's score statistic for testing (1) in the multivariate normal case. The logarithm of the density of X is given by

$$l(\mathbf{x}, \boldsymbol{\theta}) = c - \frac{1}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})'),$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix and c is a constant that does not depend on the parameters $\boldsymbol{\theta}$.

According to Definition 2 the score vector can be derived as

$$U(X, \boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}} l(X, \boldsymbol{\theta})$$

where $\frac{d}{d\boldsymbol{\theta}} = \begin{pmatrix} \frac{d}{d\boldsymbol{\mu}} \\ \frac{d}{d\text{vec}(\boldsymbol{\Sigma})} \end{pmatrix} : (p + p^2) \times 1$.

Using the differentiation rules given in Section 2 the derivative $\frac{dl}{d\boldsymbol{\mu}}$ equals:

$$\frac{d}{d\boldsymbol{\mu}} l(X, \boldsymbol{\theta}) = -\frac{1}{2} \frac{d}{d\boldsymbol{\mu}} (X - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (X - \boldsymbol{\mu}) = -\frac{1}{2} \frac{d(X - \boldsymbol{\mu})}{d\boldsymbol{\mu}} \frac{d(X - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (X - \boldsymbol{\mu})}{d(X - \boldsymbol{\mu})}$$

$$= -\frac{1}{2} (-\mathbf{I}_p) 2 \boldsymbol{\Sigma}^{-1} (X - \boldsymbol{\mu}) = \boldsymbol{\Sigma}^{-1} (X - \boldsymbol{\mu}).$$

The derivative of $l(\cdot)$ by $\boldsymbol{\Sigma}$ gives the following expression:

$$\frac{d}{d\text{vec}(\boldsymbol{\Sigma})} l(\mathbf{x}, \boldsymbol{\theta}) = -\frac{1}{2} \frac{d|\boldsymbol{\Sigma}|}{d\boldsymbol{\Sigma}} \frac{d \ln |\boldsymbol{\Sigma}|}{d|\boldsymbol{\Sigma}|} - \frac{1}{2} \frac{d\boldsymbol{\Sigma}^{-1}}{d\boldsymbol{\Sigma}} \frac{d\text{tr}(\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu})(X - \boldsymbol{\mu})')}{d\boldsymbol{\Sigma}^{-1}}$$

$$= -\frac{1}{2} |\boldsymbol{\Sigma}| \text{vec}(\boldsymbol{\Sigma}^{-1}) \cdot \frac{1}{|\boldsymbol{\Sigma}|} - \frac{1}{2} (-\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \cdot \text{vec}((X - \boldsymbol{\mu})(X - \boldsymbol{\mu})')$$

$$= -\frac{1}{2} \text{vec}(\boldsymbol{\Sigma}^{-1}) + \frac{1}{2} (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \text{vec}(\mathbf{S}_1) = \frac{1}{2} \text{vec}(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_1 \boldsymbol{\Sigma}^{-1} - \mathbf{I}_p))$$

where $\mathbf{S}_1 = (X - \boldsymbol{\mu})(X - \boldsymbol{\mu})'$.

Hence the score vector can be written as a partitioned matrix:

$$U(X, \boldsymbol{\theta}) = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \\ \frac{1}{2}\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_1\boldsymbol{\Sigma}^{-1} - \mathbf{I}_p)) \end{pmatrix}.$$

The score vector for the sample equals

$$U(\mathcal{X}, \boldsymbol{\theta}) = \begin{pmatrix} n\boldsymbol{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \\ \frac{n}{2}\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_n\boldsymbol{\Sigma}^{-1} - \mathbf{I_p})) \end{pmatrix}$$

where $\mathbf{S}_n = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \boldsymbol{\mu})(X_i - \boldsymbol{\mu})'$.

Next, the Hessian matrix $\mathrm{H}(X, \boldsymbol{\theta})$ is derived. According to Definition 4 $\mathrm{H}(X, \boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}}U(X, \boldsymbol{\theta})$:

$$\frac{d}{d\boldsymbol{\theta}}U(X, \boldsymbol{\theta}) = \begin{pmatrix} \frac{d}{d\boldsymbol{\mu}}\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu}) & \frac{d}{d\boldsymbol{\mu}}\frac{1}{2}\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_1\boldsymbol{\Sigma}^{-1} - \mathbf{I}_p)) \\ \frac{d}{d\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu}) & \frac{d}{d\boldsymbol{\Sigma}}\frac{1}{2}\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_1\boldsymbol{\Sigma}^{-1} - \mathbf{I}_p)) \end{pmatrix}$$

$$=: \begin{pmatrix} \mathrm{H}_{1,1} & \mathrm{H}_{1,2} \\ \mathrm{H}_{2,1} & \mathrm{H}_{2,2} \end{pmatrix} : (p + p^2) \times (p + p^2) \ .$$

First, let us derive $\mathrm{H}_{1,1}$:

$$\mathrm{H}_{1,1} = \frac{d}{d\boldsymbol{\mu}}\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu}) = \frac{d(X - \boldsymbol{\mu})}{d\boldsymbol{\mu}}\frac{d(\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu}))}{d(X - \boldsymbol{\mu})} = -\boldsymbol{\Sigma}^{-1}.$$

Next, let us derive $\mathrm{H}_{2,1}$:

$$\mathrm{H}_{2,1} = \frac{d}{d\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu}) = \frac{d\boldsymbol{\Sigma}^{-1}}{d\boldsymbol{\Sigma}}\frac{d\boldsymbol{\Sigma}^{-1}(X - \boldsymbol{\mu})}{d\boldsymbol{\Sigma}^{-1}}$$

$$= (-\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})((X - \boldsymbol{\mu}) \otimes \mathbf{I}_p) = (\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - X)) \otimes \boldsymbol{\Sigma}^{-1}.$$

Since $\mathrm{H}_{1,2} = \mathrm{H}'_{2,1}$ then $\mathrm{H}_{1,2} = ((\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - X)) \otimes \boldsymbol{\Sigma}^{-1})'$.
Finally, we derive $\mathrm{H}_{2,2}$:

$$\mathrm{H}_{2,2} = \frac{1}{2}\frac{d\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}(\mathbf{S}_1\boldsymbol{\Sigma}^{-1} - \mathbf{I}_p))}{d\boldsymbol{\Sigma}} = \frac{1}{2}\frac{d\boldsymbol{\Sigma}^{-1}}{d\boldsymbol{\Sigma}}\frac{d\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}\mathbf{S}_1\boldsymbol{\Sigma}^{-1}) - \mathrm{vec}\,(\boldsymbol{\Sigma}^{-1})}{d\boldsymbol{\Sigma}^{-1}}.$$

Since

$$\frac{d(\mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}\mathbf{S}_1\boldsymbol{\Sigma}^{-1}) - \mathrm{vec}\,(\boldsymbol{\Sigma}^{-1}))}{d\boldsymbol{\Sigma}^{-1}} = \frac{d\boldsymbol{\Sigma}^{-1}}{d\boldsymbol{\Sigma}^{-1}}(\mathbf{S}_1\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}_p) + \frac{d\mathbf{S}_1\boldsymbol{\Sigma}^{-1}}{d\boldsymbol{\Sigma}^{-1}}(\mathbf{I}_p\otimes\boldsymbol{\Sigma}^{-1}) - \mathbf{I}_{p^2}$$

$$= (\mathbf{S}_1\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_p) + (\mathbf{I}_p \otimes \mathbf{S}_1)(\mathbf{I}_p \otimes \boldsymbol{\Sigma}^{-1}) = (\mathbf{S}_1\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_p) + (\mathbf{I}_p \otimes \mathbf{S}_1\boldsymbol{\Sigma}^{-1}) - \mathbf{I}_{p^2},$$

then

$$\mathrm{H}_{2,2} = -\frac{1}{2}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})((\mathbf{S}_1\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_p) + (\mathbf{I}_p \otimes \mathbf{S}_1\boldsymbol{\Sigma}^{-1}) - \mathbf{I}_{p^2})$$

Hence, the Hessian matrix of X is given by

$$\mathrm{H}(X, \boldsymbol{\theta}) =$$

$$\begin{pmatrix} -\boldsymbol{\Sigma}^{-1} & ((\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - X)) \otimes \boldsymbol{\Sigma}^{-1})' \\ (\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - X)) \otimes \boldsymbol{\Sigma}^{-1} & -\frac{1}{2}(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})^{-1}[(\mathbf{S}_1 \boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_p) + (\mathbf{I}_p \otimes \mathbf{S}_1 \boldsymbol{\Sigma}^{-1}) - \mathbf{I}_{p^2}] \end{pmatrix}.$$

Finally, the information matrix $\mathrm{I}(X, \boldsymbol{\theta})$ is of the form

$$\mathrm{I}(X, \boldsymbol{\theta}) = -\mathrm{E}(\mathrm{H}(X, \boldsymbol{\theta})) = \begin{pmatrix} \mathbf{I}_{11} & \mathbf{I}_{12} \\ \mathbf{I}_{21} & \mathbf{I}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \end{pmatrix}$$

as

$$\mathrm{E}(\mathbf{S}_1) = \mathrm{E}(X - \boldsymbol{\mu})(X - \boldsymbol{\mu})' = D(X) = \boldsymbol{\Sigma}.$$

For a sample of size n

$$\mathrm{I}(\mathcal{X}, \boldsymbol{\theta}) = n\mathrm{I}(X, \boldsymbol{\theta}).$$

Now, Rao's score test (RST) statistic for testing $\mathrm{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ can be calculated using the expressions of $U(\mathcal{X}, \boldsymbol{\theta})$ and $\mathrm{I}(\mathcal{X}, \boldsymbol{\theta})$, substituting $\boldsymbol{\mu}$ with $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma}_0$, and noting that the inverse of a block-diagonal matrix is a block-diagonal matrix with the inverses of the diagonal blocks.

Proposition 3. *Rao's score test statistic for testing* $\mathrm{H}_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ *when no constraints are imposed on* $\boldsymbol{\mu}$ *has the following form:*

$$RST(\mathcal{X}, \boldsymbol{\Sigma}_0) = \frac{n}{2}\mathrm{tr}((\boldsymbol{\Sigma}_0^{-1}\widehat{\boldsymbol{\Sigma}} - \mathbf{I}_p)^2), \tag{6}$$

where $\mathcal{X} = (X_1, ..., X_n)$ *is a sample from* $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *When* $n \to \infty$ *and* H_0 *holds then* $RST(\mathcal{X}, \boldsymbol{\Sigma}_0) \sim \chi^2$ *with degrees of freedom* $df = \frac{p(p+1)}{2}$.

Proof. When there are no restrictions set on $\boldsymbol{\mu}$ then maximum likelihood estimate of $\boldsymbol{\mu}$ is inserted into score vector to replace $\boldsymbol{\mu}$ (see Gombay, 2002). Then the part of the score vector corresponding to $\boldsymbol{\mu}$ turns to zero and hence only the second part of the score vector and element $\mathrm{I}_{2,2}$ of the information matrix will contribute to the calculation of the test statistic. Also, substituting $\boldsymbol{\mu}$ with $\widehat{\boldsymbol{\mu}}$ turns $\mathbf{S}_n$ into $\widehat{\boldsymbol{\Sigma}}$. Taking this into account, one gets

$$RST(\mathcal{X}, \boldsymbol{\Sigma}_0) = \frac{n}{2}\mathrm{vec}\,'(\boldsymbol{\Sigma}_0^{-1}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}_0^{-1} - I_p))(\boldsymbol{\Sigma}_0^{-1} \otimes \boldsymbol{\Sigma}_0^{-1})^{-1}\mathrm{vec}\,(\boldsymbol{\Sigma}_0^{-1}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}_0^{-1} - \mathbf{I}_p))$$

$$= \frac{n}{2}\mathrm{vec}\,'(\boldsymbol{\Sigma}_0^{-1}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}_0^{-1} - \mathbf{I}_p))\mathrm{vec}\,(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0) = \frac{n}{2}\mathrm{tr}((\boldsymbol{\Sigma}_0^{-1}\widehat{\boldsymbol{\Sigma}} - \mathbf{I}_p)^2).$$

Note that while deriving Rao's score test statistic we have taken derivatives by vec $\boldsymbol{\Sigma}$, not by vech $\boldsymbol{\Sigma}$ which consists of the elements of the lower triangle of $\boldsymbol{\Sigma}$. A transformation via the duplication matrix which is used in the liturature (see Yuan, Bentler (1999), for example) complicates derivation and is not necessary. It turns out that both approaches will lead to the test statistic given in (6) which follows from the next lemma.

Let $\mathbf{D} : \frac{p(p+1)}{2} \times p^2$ denote the elimination matrix, i.e.,

$$\mathbf{D}\mathrm{vec}\,\boldsymbol{\Sigma} = \mathrm{vec}\,\boldsymbol{\Sigma}_{\triangle},$$

where $\boldsymbol{\Sigma}_\triangle$ represents the elements of the lower triangle of $\boldsymbol{\Sigma}$.

Lemma 1. *Let U_2 and $\mathbf{I}_{22}$ denote the part of the score vector and the part of the information matrix corresponding to $\boldsymbol{\Sigma}$. Let $U_\triangle = \mathbf{D}U_2$ and $\mathbf{I}_\triangle = \mathbf{D}\mathbf{I}_{22}\mathbf{D}'$ denote the part of score vector and the part of information matrix corresponding to the unique elements of $\boldsymbol{\Sigma}$, respectively. Then $RST_\triangle(\mathcal{X}, \boldsymbol{\Sigma}_0) = U_\triangle' \mathbf{I}_\triangle^{-1} U_\triangle$ is identical to $RST(\mathcal{X}, \boldsymbol{\Sigma}_0)$.*
Proof.

$$RST_\triangle(\mathcal{X}, \boldsymbol{\Sigma}_0) = (\mathbf{D}U_2)'(\mathbf{D}\mathbf{I}_{22}\mathbf{D}')^{-1}(\mathbf{D}U_2) = U_2'\mathbf{D}'(\mathbf{D}\mathbf{I}_{22}\mathbf{D}')^+ \mathbf{D}U_2$$

$$= U_2'\mathbf{D}'(\mathbf{D}')^+\mathbf{I}_{22}^{-1}\mathbf{D}^+\mathbf{D}U_2 = U_2'\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})\mathbf{I}_{22}^{-1}\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})U_2.$$

On the last line the equality $\mathbf{D}^+\mathbf{D} = \frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})$ was used (Magnus, Neudecker, 1999, p.49).
As $\mathbf{I}_{22}^{-1} = 2(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})$, by a property of the commutation matrix $\mathbf{K}_{p,p}$ (see Kollo, von Rosen (2005), p. 82, for example)

$$(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})\mathbf{I}_{22}^{-1} = \mathbf{I}_{22}^{-1}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p}).$$

From here

$$RST_\triangle(\mathcal{X}, \boldsymbol{\Sigma}_0) = U_2'\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})\mathbf{I}_{22}^{-1}U_2 = U_2'\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})\text{vec}\,(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)$$

$$= U_2'\text{vec}\,(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0) = RST(\mathcal{X}, \boldsymbol{\Sigma}_0).$$

Remark. The statement of the lemma above is a consequence of a more general

relation in Kollo (1994) for patterned matrices.

Corollary 2. *Rao's score test statistic for testing $\mathrm{H}_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ when no constraints are imposed on $\boldsymbol{\mu}$ has the following form:*

$$RST(\mathcal{X}, \mathbf{I}_p) = \frac{n}{2}\text{tr}(\widehat{\boldsymbol{\Sigma}} - \mathbf{I}_p)^2, \tag{7}$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $n \to \infty$ and H_0 holds then $RST(\mathcal{X}, \mathbf{I}_p) \sim \chi^2$ with degrees of freedom $df = \frac{p(p+1)}{2}$.
Proof. It follows directly from (6).

Corollary 3. *Rao's score test statistic for testing $\mathrm{H}_0 : \boldsymbol{\Sigma} = \gamma\mathbf{I}_p$ when no constraints are imposed on $\boldsymbol{\mu}$ has the following form:*

$$RST(\mathcal{X}, \gamma\mathbf{I}_p) = \frac{n}{2}\text{tr}\Big(\Big(\frac{p}{\text{tr}(\widehat{\boldsymbol{\Sigma}})}\widehat{\boldsymbol{\Sigma}} - \mathbf{I}_p\Big)^2\Big), \tag{8}$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\gamma > 0$ is an unknown parameter. When $n \to \infty$ and H_0 holds then $RST(\mathcal{X}, \gamma\mathbf{I}_p) \sim \chi^2$ with degrees of freedom $df = \frac{p(p+1)}{2} - 1$.
Proof. It follows from (6) when substituting $\boldsymbol{\Sigma}_0$ by $\widehat{\gamma}\mathbf{I}_p$, where $\widehat{\gamma} = \frac{1}{p}\text{tr}(\widehat{\boldsymbol{\Sigma}})$ is the maximum likelihood estimator of γ.

3.3 Wald's score test

Definition 6. *Wald's score test statistic for testing hypothesis (1) is given by*

$$WST(\mathcal{X}, \boldsymbol{\theta}_0) = (T_{\boldsymbol{\theta}}(\mathcal{X}) - \boldsymbol{\theta}_0)' \cdot I(\mathcal{X}, T_{\boldsymbol{\theta}}(\mathcal{X})) \cdot (T_{\boldsymbol{\theta}}(\mathcal{X}) - \boldsymbol{\theta}_0), \qquad (9)$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_r)'$, $T_{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$ and $\mathcal{X}$ is a random sample. When sample size $n \to \infty$ and H_0 holds then $WST(\mathcal{X}, \boldsymbol{\theta_0})$ converges to χ_r^2-distribution. (Wald, 1943)

Using the results obtained earlier one can calculate Wald's score test statistic for the multivariate normally distributed observations and $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ by substituting $\boldsymbol{\mu}$ with maximum likelihood estimate $\widehat{\boldsymbol{\mu}}$ and $\boldsymbol{\Sigma}$ with maximum likelihood estimate $\widehat{\boldsymbol{\Sigma}}$ in the expression of $I(\mathcal{X}, T_{\boldsymbol{\theta}}(\mathcal{X}))$.

Proposition 4. *Wald's score test statistic for testing $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ when no constraints are imposed on $\boldsymbol{\mu}$ has the following form:*

$$WST(\mathcal{X}, \boldsymbol{\Sigma}_0) = \frac{n}{2}\text{tr}((\mathbf{I}_p - \boldsymbol{\Sigma}_0 \widehat{\boldsymbol{\Sigma}}^{-1})^2), \qquad (10)$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $n \to \infty$ and H_0 holds then $WST(\mathcal{X}, \boldsymbol{\Sigma}_0) \sim \chi^2$ with degrees of freedom $df = \frac{p(p+1)}{2}$.

Proof. When there are no restrictions set on $\boldsymbol{\mu}$ then $T_{\boldsymbol{\theta}}(\mathcal{X})$ is replaced by $\widehat{\boldsymbol{\Sigma}}$, $\boldsymbol{\theta}_0$ by $\boldsymbol{\Sigma}_0$ and $I(\mathcal{X}, T_{\boldsymbol{\theta}}(\mathcal{X}))$ by $I_{2,2}(\mathcal{X}, \widehat{\boldsymbol{\Sigma}})$ in expression (9). This gives

$$\begin{aligned} WST(\mathcal{X}, \boldsymbol{\Sigma}_0) &= \frac{n}{2}\text{vec}\,'(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)(\widehat{\boldsymbol{\Sigma}}^{-1} \otimes \widehat{\boldsymbol{\Sigma}}^{-1})\text{vec}\,(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0) \\ &= \frac{n}{2}\text{vec}\,'(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)\text{vec}\,(\widehat{\boldsymbol{\Sigma}}^{-1}(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)\widehat{\boldsymbol{\Sigma}}^{-1}) \\ &= \frac{n}{2}\text{tr}((\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)\widehat{\boldsymbol{\Sigma}}^{-1}(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0)\widehat{\boldsymbol{\Sigma}}^{-1}) = \frac{n}{2}\text{tr}((\mathbf{I}_p - \boldsymbol{\Sigma}_0\widehat{\boldsymbol{\Sigma}}^{-1})^2). \end{aligned}$$

Corollary 4. *Wald's score test statistic for testing $H_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ when no constraints are imposed on $\boldsymbol{\mu}$ has the following form:*

$$WST(\mathcal{X}, \mathbf{I}_p) = \frac{n}{2}\text{tr}((\mathbf{I}_p - \widehat{\boldsymbol{\Sigma}}^{-1})^2), \qquad (11)$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $n \to \infty$ and H_0 holds then $WST(\mathcal{X}, \mathbf{I}_p) \sim \chi^2$ with degrees of freedom $df = \frac{p(p+1)}{2}$.
Proof. It follows directly from (10).

Corollary 5. *Wald's score test statistic for testing $H_0 : \boldsymbol{\Sigma} = \gamma\mathbf{I}_p$ when no constraints are imposed on $\boldsymbol{\mu}$ has the following form:*

$$WST(\mathcal{X}, \gamma\mathbf{I}_p) = \frac{n}{2}\text{tr}\left(\left(\mathbf{I}_p - \frac{\text{tr}(\widehat{\boldsymbol{\Sigma}})}{p}\widehat{\boldsymbol{\Sigma}}^{-1}\right)^2\right), \qquad (12)$$

where $\mathcal{X} = (X_1, ..., X_n)$ is a sample from $N_p(\boldsymbol{\mu}, b\Sigma)$ and $\gamma > 0$ is an unknown parameter. When $n \to \infty$ and H_0 holds then $WST(\mathcal{X}, \gamma\mathbf{I}_p) \sim \chi^2$ with degrees

of freedom $df = \frac{p(p+1)}{2} - 1$.

Proof. It follows from (10) when substituting $\boldsymbol{\Sigma}_0$ by $\widehat{\gamma}\mathbf{I}_p$, where $\widehat{\gamma} = \frac{1}{p}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}})$ is the maximum likelihood estimator of γ.

4. Simulation experiment

In the simulation experiment we have compared the behaviour of the likelihood ratio test, Rao's score test and Wald's score test in different situations: both under H_0 and the alternative, with dimension of the covariance matrix going from 2 to 50 and with the sample size growing from 100 to 5000. The simulation experiment was carried out using the R software. Note that we have used relatively small sample sizes and critical values from asymptotic distributions for testing the behaviour of the test statistics. This was done in order to mirror realistic analysis situations, where the amount of data is limited and statistical softwares produce p-values from asymptotic distributions.

Testing $\boldsymbol{\Sigma} = \mathbf{I}_p$ under H_0

The following procedure was used for the simulation:
1. Generate a random sample $\mathbf{x}_1, ..., \mathbf{x}_n$ from multivariate N_p distribution with mean vector $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma} = \mathbf{I}_p$.
2. Use the LRT, RST and WST to test $\boldsymbol{\Sigma} = \mathbf{I}_p$.
3. Repeat $N = 300$ times for each combination of $p = 2, 4, 6, 8, 10, 12, 14,$ 16, 18, 20, 25, 30, 35, 40, 45, 50 and $n = 100, 250, 500, 1000, 2500, 5000$.
4. Count the number of H_0 rejected by LRT, RST and WST for each combination of p and n.

The results of the simulation experiment are presented in Figure 1.

Testing $\boldsymbol{\Sigma} = \mathbf{I}_p$ when $\boldsymbol{\Sigma} = \gamma\mathbf{I}_p$

The following procedure was used for the simulation:
1. Generate random sample $\mathbf{x}_1, ..., \mathbf{x}_{300}$ from multivariate N_p distribution with mean vector $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma} = \gamma\mathbf{I}_p$, where $\gamma = 1.1, 1.2,$ 1.3, 1.4, 1.5, 2.0.
2. Use LRT, RST and WST to test $\boldsymbol{\Sigma} = \mathbf{I}_p$.
3. Repeat $N = 300$ times for each combination of $p = 2, 4, 6, 8, 10, 12, 14,$ 16, 18, 20, 25, 30, 35, 40, 45, 50 and γ.
4. Count the number of H_0 rejected by LRT, RST and WST for each combination of p and γ.

The results of the simulation experiment are presented in Figure 2.

Figure 1. Simulation results for testing H$_0$: $\boldsymbol{\Sigma} = \mathbf{I}_p$ under H$_0$
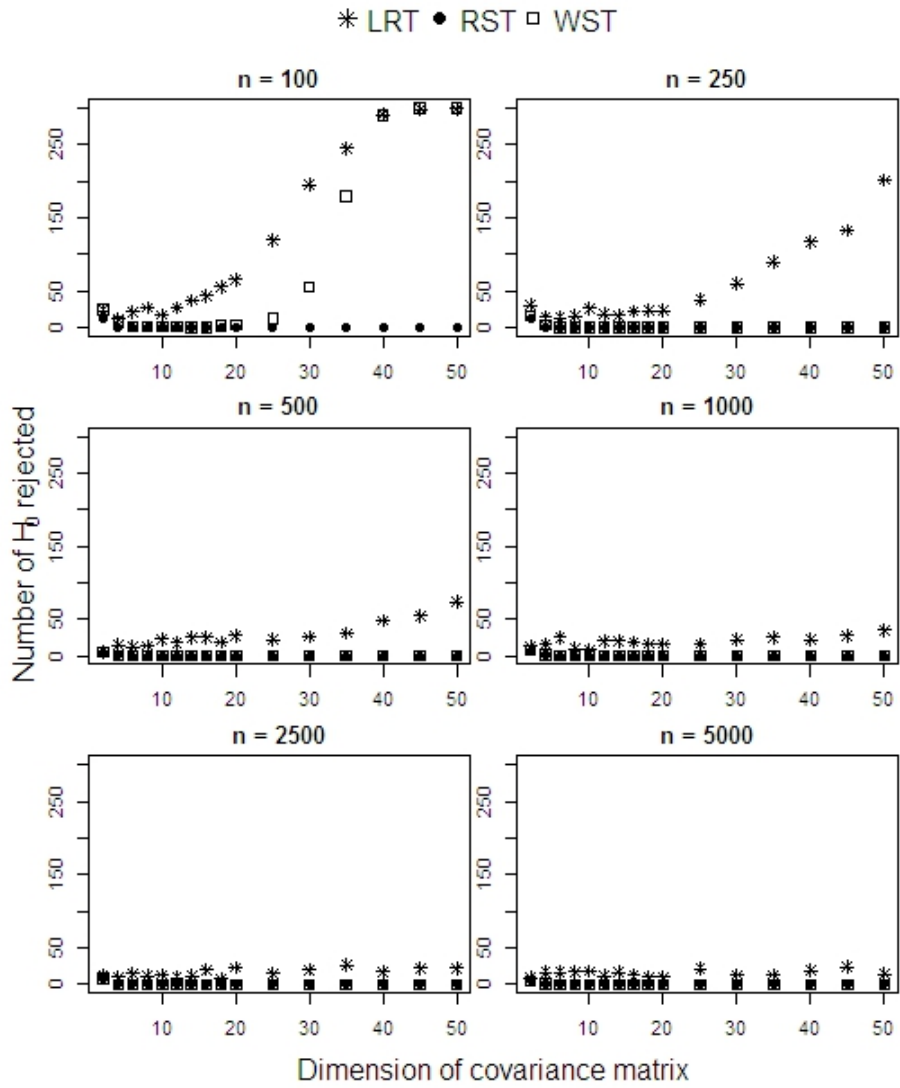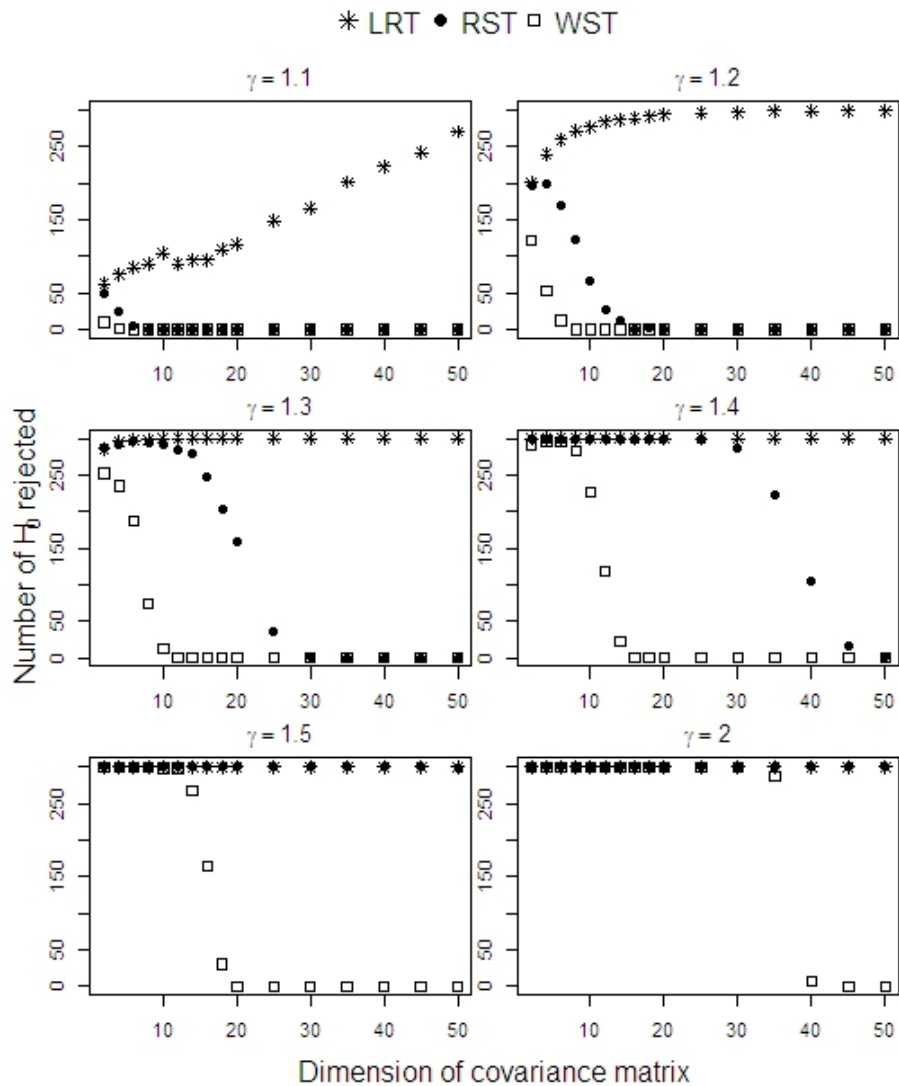

435

Figure 2. Simulation results for testing $H_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ under the alternative

Simulation results

Figure 1 shows that when H_0 is true, then RST and WST almost never reject the null-hypothesis. At the same time the LRT rejected the correct null-hypothesis too often, especially in the case when the sample size was smaller than the number of different parameters in the covariance matrix.

From Figure 2 one can see that the LRT rejected the null-hypothesis mostly when it should have been rejected. Here we tested $\boldsymbol{\Sigma} = \mathbf{I}_p$ but simulated data from a multivariate normal distribution with $\boldsymbol{\Sigma} = \gamma\mathbf{I}_p$, and $\gamma = 1.1$, 1.2, 1.3, 1.4, 1.5, 2.0. LRT rejected the null hypothesis almost all the time even if $\gamma\mathbf{I}_p$ was quite close to the hypothetical H_0. Rao's score test was not as sensitive and a larger difference from the null hypothesis was needed in order to reject it. However, with $\gamma = 1.3$, 1.4 and dimension of the covariance matrix realtively small, RST gave correct results. When $\gamma{=}1.5$, then RST rejected the null hypothesis all the time, as did the LRT. WST tended to stay with the null hypothesis and even with $\gamma{=}2.0$, it still stayed with the null-hypothesis when the dimension of the covariance matrix was large.

5. Summary
In chapter 3 we derived the likelihood ratio test statistic, Rao's score test statistic and Wald's score test statistic for testing $\mathrm{H}_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, $\mathrm{H}_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ and $\mathrm{H}_0 : \boldsymbol{\Sigma} = \gamma\mathbf{I}_p$.

In chapter 4 we carried out simulation experiment and reached the following conclusion. Under realistic conditions and while using critical values from the asymptotic distribution (as implemented in most statistical software), Rao's score test statistic should be preferred to the likelihood ratio and Wald's score test statistics. RST stays with H_0 when it holds and rejects it realtively easily when it does not hold - hence reaching the correct conclusion both under null and the alternative. Meanwhile, the LRT tends to reject the null hypothesis even if it holds (particularly when sample size is less than the number of unique parameters in the covariance matrix) and WST tends to stay with the null hypothesis even if it should be rejected.

References:

Anderson, T. W. (2003). An Introduction to Multivariate Statistical Analysis, 3rd Edition. Wiley, New York.

Bai, Z., Jiang, D., Yao, J., Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics* 37, 3822-3840.

Bilodeau, M., Brenner, D. (1999). Theory of Multivariate Statistics. Springer, New York.

Gombay, E. (2002). Parametric sequential tests in the presence of nuisance parameters. *Theory Stochastic Processes* 8, 106-118.

Kollo, T. (1994). A note on patterned matrices with applications in multivariate analysis. *Acta et Commentationes Universitatis Tartuensis de Mathematica* 968, 17-27.

Kollo, T., von Rosen, D. (2005). Advanced Multivariate Statistics with Matrices. Springer, Dordrecht.

Magnus, J., Neudecker, H. (1999). Matrix Differential Calculus with Applications in Statistics and Econometrics. Revised edition. Wiley, New York.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society* 44, 50-57.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. 2nd edition. Wiley, New York.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters, when the number of observations is large. *Trans. Amer. Math. Soc.* 54, 426-482.

Yuan, K.-H., Bentler, P. M. (1999). On normal theory and associated test statistics in covariance structure analysis under two classes of nonnormal distributions. *Statistica Sinica* 9, 831-853.