Optimal and Additive Loss Reserving for Dependent Lines of Business

By Klaus D. Schmidt

Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden

Abstract

In the present paper we review and extend two stochastic models for loss reserving and study their impact on extensions of the additive method and of the chain–ladder method. The first of these models is a particular linear model while the second one is a sequential model which is composed of a finite number of conditional linear models. These models lead to multivariate extensions of the additive method and of the chain–ladder method, respectively, which turn out to resolve the problem of additivity.

Contents

1 Introduction

For a portfolio consisting of several lines of business, it is well–known that the chain–ladder predictors for the aggregate portfolio usually differ from the sums of the chain–ladder predictors for the different lines of business; see Ajne [1994] and Klemmt [2004]. It is one of the purposes of the present paper to point out that the non–coincidence between a chain–ladder predictor for the aggregate portfolio and the sum of the chain–ladder predictors for the different lines of business has its origin in the univariate character of the chain–ladder method which neglects the dependence structure existing between the different lines of business.

The problem of dependence between different lines of business has already been addressed in a paper by Holmberg [1994]. His paper is remarkable since it adopts a general point of view and considers
- correlation within accident years,
- correlation between accident years, and
- correlation between different lines of business.

Nevertheless, the major part of Holmberg's paper is devoted to correlation within and between accident years and the author expresses the opinion that, in practical applications, the great majority of the effects causing correlation between different lines of business are already captured in the correlation within and between accident years. It is another purpose of the present paper to show that correlation between different lines of business can be modelled and that the resulting models, combined with general optimization principles, lead to multivariate predictors which are superior to the univariate ones. Here and in the sequel, the term *univariate* refers to prediction for a single line of business and the term *multivariate* refers to simultaneous prediction for several lines of business or for different types of losses (like paid and incurred losses) of the same line of business.

The papers by Ajne [1994] and Holmberg [1994] were slightly preceded in time by a paper by Mack [1993] which, similar to the paper by Hachemeister and Stanard [1975], turned out to be path–breaking in the discussion of stochastic models for the chain–ladder method. In the model of Mack, dependence within accident years is expressed by conditioning, but it is also assumed that the accident years are independent. The assumption of independent accident years was subsequently relaxed in the model of Schnaus presented by Schmidt and Schnaus [1996]. Both of these models are univariate and hence do not reflect dependence between lines of business.

After the publication of the paper of Mack [1993], about a decade had to pass before the emergence of the first bivariate models related to the chain–ladder method. One of these models, due to Quarg and Mack [2004], expresses dependence between the paid and incurred losses of a single line of business (a topic which had already been studied before by Halliwell [1997] within the theory of linear models) and has been used as a foundation for the construction of certain bivariate predictors which are now known as Munich chain–ladder predictors. The other of these models, due to

Braun [2004], expresses dependence between two lines of business and has been used to construct new estimators for the prediction errors of the univariate chain–ladder predictors, but it has not been used to construct bivariate predictors. Each of these models extends the model of Mack.

Quite recently, Pröhl and Schmidt [2005] as well as Hess, Schmidt and Zocher [2006] proposed multivariate models which reflect dependence between an arbitrary number of lines of business. The model of Pröhl and Schmidt extends the model of Braun in essentially the same way as the model of Schnaus extends the model of Mack, while the model of Hess, Schmidt and Zocher extends in a rather straightforward way the particular linear model which may be used to justify the additive method; see Radtke and Schmidt [2004]. These models, combined with a general optimality criterion, lead to multivariate versions of the chain–ladder method and of the additive method, respectively, which turn out to resolve the problem of additivity.

In the present paper we review these recent multivariate models and methods of loss reserving. In order to avoid the accumulation of technicalities, we start with a systematic review of the univariate case (Section 2) and of prediction in conditional linear models (Section 3). We then pass to the multivariate case (Section 4) and show that, due to the multivariate approach, the predictors for the single lines of business sum up to the corresponding predictors for the aggregate portfolio (Section 5). We also show how the unbiased estimators of variances and covariances proposed by Braun [2004] can be adapted to the multivariate models considered here (Section 6). We conclude with some complementary remarks (Section 7) and a numerical example illustrating the multivariate chain–ladder method (Section 8).

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be a probability space on which all random variables, random vectors and random matrices are defined. We assume that all random variables are square integrable and that all random vectors and random matrices have square integrable coordinates. Moreover, all equalities and inequalities involving random variables are understood to hold almost surely with respect to the probability measure P.

2 Univariate Loss Prediction

In the present section we review two univariate stochastic models which are closely related to two current methods of loss reserving.

We consider a single line of business which is described by a family $\{Z_{i,k}\}_{i,k\in\{0,1,\dots,n\}}$ of random variables. We interpret $Z_{i,k}$ as the loss of *accident year* i which is reported or settled in *development year* k, and hence in *calendar year* $i+k$, and we refer to $Z_{i,k}$ as the *incremental loss* of accident year i and development year k.

We assume that the incremental losses $Z_{i,k}$ are *observable* for calendar years $i+k \leq n$ and that they are *non–observable* for calendar years $i+k \geq n+1$. The observable incremental losses are represented by the following *run–off triangle*:

Accident				Development Year					
Year	0	1	$\dots$	k	$\dots$	$n-i$	$\dots$	$n-1$	n
0	$Z_{0,0}$	$Z_{0,1}$	$\dots$	$Z_{0,k}$	$\dots$	$Z_{0,n-i}$	$\dots$	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	$\dots$	$Z_{1,k}$	$\dots$	$Z_{1,n-i}$	$\dots$	$Z_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$Z_{i,0}$	$Z_{i,1}$	$\dots$	$Z_{i,k}$	$\dots$	$Z_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	$\dots$	$Z_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
n	$Z_{n,0}$								

Besides looking at the incremental losses, we also consider the *cumulative losses* $S_{i,k}$ which are defined by

$$S_{i,k} \;:=\; \sum_{l=0}^{k} Z_{i,l}$$

Then the cumulative losses $S_{i,k}$ are observable for calendar years $i+k \leq n$ and they are non–observable for calendar years $i+k \geq n+1$. Just like the observable incremental losses, the observable cumulative losses are represented by a run–off triangle:

Accident				Development Year					
Year	0	1	$\dots$	k	$\dots$	$n-i$	$\dots$	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$	$\dots$	$S_{0,k}$	$\dots$	$S_{0,n-i}$	$\dots$	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$	$\dots$	$S_{1,k}$	$\dots$	$S_{1,n-i}$	$\dots$	$S_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$S_{i,0}$	$S_{i,1}$	$\dots$	$S_{i,k}$	$\dots$	$S_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$	$\dots$	$S_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
n	$S_{n,0}$								

Of course, the incremental losses can be recovered from the cumulative losses.

2.1 Univariate Additive Model

Let us first consider the univariate additive model:

Univariate Additive Model: *There exist real numbers* $\nu_0, \nu_1, \ldots, \nu_n > 0$ *and* $\sigma_0, \sigma_1, \ldots, \sigma_n > 0$ *as well as real parameters* $\zeta_0, \zeta_1, \ldots, \zeta_n$ *such that*

$$E[Z_{i,k}] = \nu_i \, \zeta_k$$

and

$$\mathrm{cov}[Z_{i,k}, Z_{j,l}] = \begin{cases} \nu_i \, \sigma_k & \text{if } i = j \text{ and } k = l \\ 0 & \text{else} \end{cases}$$

holds for all $i, j, k, l \in \{0, 1, \ldots, n\}$.

For $i, k \in \{1, \ldots, n\}$ such that $i + k \geq n + 1$, the estimators and predictors

$$\begin{aligned} \widehat{\zeta}_k^{\mathrm{AD}} &:= \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \nu_j} \\ \widehat{Z}_{i,k}^{\mathrm{AD}} &:= \nu_i \, \widehat{\zeta}_k^{\mathrm{AD}} \\ \widehat{S}_{i,k}^{\mathrm{AD}} &:= S_{i,n-i} + \nu_i \sum_{l=n-i+1}^{k} \widehat{\zeta}_l^{\mathrm{AD}} \end{aligned}$$

are said to be the estimators and the predictors of the univariate additive method. Under the assumptions of the univariate additive model, these estimators and predictors are indeed reasonable, as will be shown in Section 4 below.

2.2 Univariate Chain–Ladder Model

Let us now consider the univariate chain–ladder model due to Schnaus which was proposed by Schmidt and Schnaus [1996] and is a slight but convenient extension of the model of Mack [1993].

The chain–ladder model is a sequential model since it involves successive conditioning with respect to the σ–algebras $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ where, for each $k \in \{1, \ldots, n\}$, the σ–algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $S_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$, which is at the same time the information provided by the incremental losses $Z_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$.

We assume that $S_{i,k} > 0$ holds for all $i, k \in \{0, 1, \ldots, n\}$.

Univariate Chain–Ladder Model: *For each $k \in \{1,\ldots,n\}$, there exists a random variable φ_k and a strictly positive random variable σ_k such that*

$$E^{\mathcal{G}_{k-1}}(S_{i,k}) \;=\; S_{i,k-1}\,\varphi_k$$

and

$$\mathrm{cov}^{\mathcal{G}_{k-1}}(S_{i,k}, S_{j,k}) \;=\; \begin{cases} S_{i,k-1}\,\sigma_k & \text{if } i=j \\ 0 & \text{else} \end{cases}$$

holds for all $i,j \in \{0,1,\ldots,n-k+1\}$.

For $i,k \in \{1,\ldots,n\}$ such that $i+k \geq n+1$, the estimators and predictors

$$\widehat{\varphi}_k^{\mathrm{CL}} \;:=\; \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

$$\widehat{S}_{i,k}^{\mathrm{CL}} \;:=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \widehat{\varphi}_l^{\mathrm{CL}}$$

(such that $\widehat{\varphi}_{n-i}^{\mathrm{CL}} = S_{i,n-i}$) are said to be the estimators and the predictors of the univariate chain–ladder method. Under the assumptions of the univariate chain–ladder model, these estimators and predictors are indeed reasonable, as will be shown in Section 4.

3 Estimation and Prediction in the Conditional Linear Model

In the present section we consider a random vector $\mathbf{X}$ and a sub–σ–algebra $\mathcal{G}$ of $\mathcal{F}$. The σ–algebra $\mathcal{G}$ represents information which is provided by some other random quantities.

The Conditional Linear Model: *There exists a $\mathcal{G}$–measurable random matrix $\mathbf{A}$ and a $\mathcal{G}$–measurable random vector $\boldsymbol{\beta}$ such that*

$$E^{\mathcal{G}}[\mathbf{X}] \;=\; \mathbf{A}\boldsymbol{\beta}$$

The random matrix $\mathbf{A}$ is assumed to be observable and is said to be the *design matrix* and the random vector $\boldsymbol{\beta}$ is assumed to be non–observable and is said to be the *parameter vector* or the *parameter* for short.

In the subsequent discussion, we assume that the assumption of the conditional linear model is fulfilled.

We assume further that some of the coordinates of $\mathbf{X}$ are *observable* whereas some other coordinates are *non–observable*. Then the random vector $\mathbf{X}_1$ consisting of the observable coordinates of $\mathbf{X}$ and the random vector $\mathbf{X}_2$ consisting of the non–observable coordinates of $\mathbf{X}$ satisfy

$$\begin{aligned} E^{\mathcal{G}}[\mathbf{X}_1] &= \mathbf{A}_1\boldsymbol{\beta} \\ E^{\mathcal{G}}[\mathbf{X}_2] &= \mathbf{A}_2\boldsymbol{\beta} \end{aligned}$$

for some submatrices $\mathbf{A}_1$ and $\mathbf{A}_2$ of $\mathbf{A}$.

We also assume that the matrix $\mathbf{A}_1$ has full column rank, that the random matrices

$$\begin{aligned} \boldsymbol{\Sigma}_{11} &:= \mathrm{Var}^{\mathcal{G}}[\mathbf{X}_1] \\ \boldsymbol{\Sigma}_{21} &:= \mathrm{Cov}^{\mathcal{G}}[\mathbf{X}_2, \mathbf{X}_1] \end{aligned}$$

are known, and that $\boldsymbol{\Sigma}_{11}$ is (almost surely) invertible.

Since the random vector $\mathbf{X}_2$ is non–observable, only the random vector $\mathbf{X}_1$ can be used for the estimation of the parameter $\boldsymbol{\beta}$.

3.1 Gauss–Markov Estimation

Let us first consider the estimation problem for a random vector of the form $\mathbf{C}\boldsymbol{\beta}$, where $\mathbf{C}$ is a $\mathcal{G}$–measurable random matrix of suitable dimension.

A random variable $\widehat{\mathbf{Y}}$ is said to be an *estimator* of $\mathbf{C}\boldsymbol{\beta}$ if it is a measurable transformation of the observable random vector $\mathbf{X}_1$. For an estimator $\widehat{\mathbf{Y}}$ of $\mathbf{C}\boldsymbol{\beta}$, the random variable

$$E^{\mathcal{G}}\Big[\Big(\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\Big)'\Big(\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\Big)\Big]$$

is said to be the *$\mathcal{G}$–conditional expected squared estimation error* of $\widehat{\mathbf{Y}}$. Since

$$\begin{aligned} & E^{\mathcal{G}}\Big[\Big(\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\Big)'\Big(\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\Big)\Big] \\ &= \mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}}[\widehat{\mathbf{Y}}]\Big) + E^{\mathcal{G}}[\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}]' E^{\mathcal{G}}[\widehat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}] \end{aligned}$$

the $\mathcal{G}$–conditional expected squared estimation error is determined by the $\mathcal{G}$–conditional variance of the estimator and the $\mathcal{G}$–conditional expectation of the estimation error. An observable random vector $\widehat{\mathbf{Y}}$ is said to be

- a *linear estimator* of $\mathbf{C}\boldsymbol{\beta}$ if there exists a $\mathcal{G}$–measurable random matrix $\mathbf{Q}$ such that $\widehat{\mathbf{Y}} = \mathbf{Q}\mathbf{X}_1$.
- a *$\mathcal{G}$–conditionally unbiased estimator* of $\mathbf{C}\boldsymbol{\beta}$ if $E^{\mathcal{G}}[\widehat{\mathbf{Y}}] = E^{\mathcal{G}}[\mathbf{C}\boldsymbol{\beta}]$.
- a *Gauss–Markov predictor* of $\mathbf{C}\boldsymbol{\beta}$ if it is a $\mathcal{G}$–conditionally unbiased linear estimator of $\mathbf{C}\boldsymbol{\beta}$ and minimizes the $\mathcal{G}$–conditional expected squared estimation error over all $\mathcal{G}$–conditionally unbiased linear estimators of $\mathbf{C}\boldsymbol{\beta}$.

We have the following result:

3.1 Proposition (Gauss–Markov Theorem for Estimators). *There exists a unique Gauss–Markov estimator* $\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{C}\boldsymbol{\beta})$ *of* $\mathbf{C}\boldsymbol{\beta}$ *and it satisfies*

$$\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{C}\boldsymbol{\beta}) \;=\; \mathbf{C}(\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_1$$

In particular, $\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{C}\boldsymbol{\beta}) = \mathbf{C}\widehat{\mathbf{Y}}^{\mathrm{GM}}(\boldsymbol{\beta})$.

Proposition 3.1 implies that the coordinates of the Gauss–Markov estimator

$$\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} \;:=\; (\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_1$$

of the parameter $\boldsymbol{\beta}$ coincide with the Gauss–Markov estimators of its coordinates.

3.2 Gauss–Markov Prediction

Let us now consider the prediction problem for a non–observable random vector of the form $\mathbf{D}\mathbf{X}_2$, where $\mathbf{D}$ is a matrix of suitable dimension.

A random variable $\widehat{\mathbf{Y}}$ is said to be a *predictor* of $\mathbf{D}\mathbf{X}_2$ if it is a measurable transformation of the observable random vector $\mathbf{X}_1$. For a predictor $\widehat{\mathbf{Y}}$ of $\mathbf{D}\mathbf{X}_2$, the random variable

$$E^{\mathcal{G}}\Big[\Big(\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\Big)'\Big(\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\Big)\Big]$$

is said to be the *$\mathcal{G}$–conditional expected squared prediction error* of $\widehat{\mathbf{Y}}$. Since

$$\begin{aligned} &E^{\mathcal{G}}\Big[\Big(\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\Big)'\Big(\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\Big)\Big] \\ &\quad= \;\mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}}\big[\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\big]\Big) + E^{\mathcal{G}}\big[\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\big]'E^{\mathcal{G}}\big[\widehat{\mathbf{Y}}-\mathbf{D}\mathbf{X}_2\big] \end{aligned}$$

the $\mathcal{G}$–conditional expected squared prediction error is determined by the $\mathcal{G}$–conditional variance and the $\mathcal{G}$–conditional expectation of the prediction error. An observable random vector $\widehat{\mathbf{Y}}$ is said to be

- a *linear predictor* of $\mathbf{D}\mathbf{X}_2$ if there exists a $\mathcal{G}$–measurable random matrix $\mathbf{Q}$ such that $\widehat{\mathbf{Y}} = \mathbf{Q}\mathbf{X}_1$.
- a *$\mathcal{G}$–conditionally unbiased predictor* of $\mathbf{D}\mathbf{X}_2$ if $E^{\mathcal{G}}[\widehat{\mathbf{Y}}] = E^{\mathcal{G}}[\mathbf{D}\mathbf{X}_2]$.
- a *Gauss–Markov predictor* of $\mathbf{D}\mathbf{X}_2$ if it is a $\mathcal{G}$–conditionally unbiased linear predictor of $\mathbf{D}\mathbf{X}_2$ and minimizes the $\mathcal{G}$–conditional expected squared prediction error over all $\mathcal{G}$–conditionally unbiased linear predictors of $\mathbf{D}\mathbf{X}_2$.

We have the following result:

3.2 Proposition (Gauss–Markov Theorem for Predictors). *There exists a unique Gauss–Markov predictor* $\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{D}\mathbf{X}_2)$ *of* $\mathbf{D}\mathbf{X}_2$ *and it satisfies*

$$\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{D}\mathbf{X}_2) \;=\; \mathbf{D}\Big(\mathbf{A}_2\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\Big(\mathbf{X}_1-\mathbf{A}_1\widehat{\boldsymbol{\beta}}^{\mathrm{GM}}\Big)\Big)$$

In particular, $\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{D}\mathbf{X}_2) = \mathbf{D}\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{X}_2)$.

Proposition 3.2 shows that the Gauss–Markov predictor

$$\widehat{\mathbf{X}}_2^{\mathrm{GM}} \; := \; \mathbf{A}_2\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\Big(\mathbf{X}_1 - \mathbf{A}_1\widehat{\boldsymbol{\beta}}^{\mathrm{GM}}\Big)$$

of the non–observable random vector $\mathbf{X}_2$ depends not only on the Gauss–Markov estimator $\widehat{\boldsymbol{\beta}}^{\mathrm{GM}}$ of the parameter $\boldsymbol{\beta}$ but also on the $\mathcal{G}$–conditional covariance $\boldsymbol{\Sigma}_{21}$ between the non–observable random vector $\mathbf{X}_2$ and the observable random vector $\mathbf{X}_1$. Moreover, the final assertion of Proposition 3.2 implies that the coordinates of the Gauss–Markov predictor of the non–observable random vector coincide with the Gauss–Markov predictors of its coordinates.

For a single non–observable random variable, the Gauss–Markov predictor has been determined by Goldberger [1962]; see also Rao and Toutenburg [1995]. We also refer to the paper by Halliwell [1996] and to the discussion of his paper by Schmidt [1999a] and Hamer [1999] and the author's response by Halliwell [1999]. Related results can also be found in Radtke and Schmidt [2004] and in Schmidt [1998, 2004].

The proof of Propositions 3.1 and 3.2 can be achieved in exactly the same way as in the unconditional case (which corresponds to the case $\mathcal{G} = \{\emptyset, \Omega\}$, where the $\mathcal{G}$–conditional expectations, variances and covariances are nothing else than the ordinary expectations, variances and covariances).

It is sometimes also of interest to predict a random vector of the form

$$\mathbf{DX} \; = \; \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}$$

An obvious candidate is the predictor

$$\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{DX}) \; := \; \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \widehat{\mathbf{X}}_2^{\mathrm{GM}} \end{pmatrix}$$

Extending the definitions and repeating the discussion with $\mathbf{X}$ in the place of $\mathbf{X}_2$, it is easily seen that the predictor $\widehat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{DX})$ is indeed the Gauss–Markov predictor of $\mathbf{DX}$; see also Hamer [1999] for the even more general case of Gauss–Markov estimation/prediction of $\mathbf{D}_0\boldsymbol{\beta} + \mathbf{D}_1\mathbf{X}_1 + \mathbf{D}_2\mathbf{X}_2$.

4 Multivariate Loss Prediction

We are now prepared to consider multivariate loss prediction.

We consider m lines of business all having the same number of development years. The m lines of business may be interpreted as subportfolios of an aggregate portfolio.

For the line of business $p \in \{1, \ldots, m\}$, we denote by

$$Z_{i,k}^{(p)}$$

and

$$S_{i,k}^{(p)}$$

the incremental loss and the cumulative loss, respectively, of accident year $i \in \{0, 1, \dots, n\}$ and development year $k \in \{0, 1, \dots, n\}$.

For $i, k \in \{0, 1, \dots, n\}$, we thus obtain the m–dimensional random vectors

$$\mathbf{Z}_{i,k} := \left(Z_{i,k}^{(p)} \right)_{p \in \{1,\dots,m\}}$$

and

$$\mathbf{S}_{i,k} := \left(S_{i,k}^{(p)} \right)_{p \in \{1,\dots,m\}}$$

of incremental losses and cumulative losses of the combined subportfolios. The observable incremental losses and the observable cumulative losses are represented by the run–off triangles

Accident	**Development Year**								
Year	0	1	...	k	...	$n-i$	...	$n-1$	n
0	$\mathbf{Z}_{0,0}$	$\mathbf{Z}_{0,1}$	...	$\mathbf{Z}_{0,k}$	...	$\mathbf{Z}_{0,n-i}$	...	$\mathbf{Z}_{0,n-1}$	$\mathbf{Z}_{0,n}$
1	$\mathbf{Z}_{1,0}$	$\mathbf{Z}_{1,1}$	...	$\mathbf{Z}_{1,k}$	...	$\mathbf{Z}_{1,n-i}$	...	$\mathbf{Z}_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$\mathbf{Z}_{i,0}$	$\mathbf{Z}_{i,1}$	...	$\mathbf{Z}_{i,k}$	...	$\mathbf{Z}_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$\mathbf{Z}_{n-k,0}$	$\mathbf{Z}_{n-k,1}$	...	$\mathbf{Z}_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$\mathbf{Z}_{n-1,0}$	$\mathbf{Z}_{n-1,1}$							
n	$\mathbf{Z}_{n,0}$								

and

Accident	**Development Year**								
Year	0	1	...	k	...	$n-i$	...	$n-1$	n
0	$\mathbf{S}_{0,0}$	$\mathbf{S}_{0,1}$	...	$\mathbf{S}_{0,k}$	...	$\mathbf{S}_{0,n-i}$	...	$\mathbf{S}_{0,n-1}$	$\mathbf{S}_{0,n}$
1	$\mathbf{S}_{1,0}$	$\mathbf{S}_{1,1}$	...	$\mathbf{S}_{1,k}$	...	$\mathbf{S}_{1,n-i}$	...	$\mathbf{S}_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$\mathbf{S}_{i,0}$	$\mathbf{S}_{i,1}$	...	$\mathbf{S}_{i,k}$	...	$\mathbf{S}_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$\mathbf{S}_{n-k,0}$	$\mathbf{S}_{n-k,1}$	...	$\mathbf{S}_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$\mathbf{S}_{n-1,0}$	$\mathbf{S}_{n-1,1}$							
n	$\mathbf{S}_{n,0}$								

We can now present multivariate extensions of the models considered in Section 2:

4.1 Multivariate Additive Model

Let us first consider a multivariate extension of the univariate additive model which applies to the combined subportfolios and was proposed by Hess, Schmidt and Zocher [2006].

> **Multivariate Additive Model:** *There exist positive definite diagonal matrices* $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_n$ *and positive definite symmetric matrices* $\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_n$ *as well as parameter vectors* $\boldsymbol{\zeta}_0, \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_n$ *such that*
>
> $$E[\mathbf{Z}_{i,k}] \;=\; \mathbf{V}_i \boldsymbol{\zeta}_k$$
>
> *and*
>
> $$\mathrm{Cov}[\mathbf{Z}_{i,k}, \mathbf{Z}_{j,l}] \;=\; \begin{cases} \mathbf{V}_i^{1/2} \boldsymbol{\Sigma}_k \mathbf{V}_i^{1/2} & \textit{if } i = j \textit{ and } k = l \\ \mathbf{O} & \textit{else} \end{cases}$$
>
> *holds for all* $i, j, k, l \in \{0, 1, \ldots, n\}$*.*

In the subsequent discussion, we assume that the assumption of the multivariate additive model is fulfilled and that the matrices $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_n$ are known.

Because of the assumption on the expectations of the incremental losses, the multivariate additive model is a linear model. This can be seen as follows: Define

$$\boldsymbol{\beta} \;:=\; \begin{pmatrix} \boldsymbol{\zeta}_0 \\ \boldsymbol{\zeta}_1 \\ \vdots \\ \boldsymbol{\zeta}_{k-1} \\ \boldsymbol{\zeta}_k \\ \boldsymbol{\zeta}_{k+1} \\ \vdots \\ \boldsymbol{\zeta}_n \end{pmatrix}$$

and, for all $i, k \in \{0, 1, \ldots, n\}$, define

$$\mathbf{A}_{i,k} \;:=\; \begin{pmatrix} \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{V}_i & \mathbf{O} & \ldots & \mathbf{O} \end{pmatrix}$$

where the matrix $\mathbf{V}_i$ occurs in position $1+k$. Then we have

$$E[\mathbf{Z}_{i,k}] \;=\; \mathbf{A}_{i,k} \boldsymbol{\beta}$$

for all $i, k \in \{0, 1, \ldots, n\}$. Let $\mathbf{Z}_1$ and $\mathbf{A}_1$ denote a block vector and a block matrix consisting of the vectors $\mathbf{Z}_{i,k}$ and the matrices $\mathbf{A}_{i,k}$ with $i + k \leq n$ (arranged in the same order) and let $\mathbf{Z}_2$ and $\mathbf{A}_2$ denote a block vector and a block matrix consisting of the vectors $\mathbf{Z}_{i,k}$ and the matrices $\mathbf{A}_{i,k}$ with $i + k \geq n + 1$. Then we have

$$E[\mathbf{Z}_1] \;=\; \mathbf{A}_1 \boldsymbol{\beta}$$

$$E[\mathbf{Z}_2] \;=\; \mathbf{A}_2\boldsymbol{\beta}$$

Therefore, the multivariate additive model is indeed a linear model.

The following result provides formulas for the Gauss–Markov estimators of the parameters of the multivariate additive model:

4.1 Theorem. *For each $k \in \{0,1,\ldots,n\}$, the Gauss–Markov estimator $\widehat{\boldsymbol{\zeta}}_k^{\mathrm{GM}}$ of $\boldsymbol{\zeta}_k$ satisfies*

$$\widehat{\boldsymbol{\zeta}}_k^{\mathrm{GM}} \;=\; \left(\sum_{j=0}^{n-k}\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)^{-1}\sum_{j=0}^{n-k}(\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k}$$

Proof. Because of the diagonal block structure of $\boldsymbol{\Sigma}_{11} = \mathrm{Var}[\mathbf{Z}_1]$ and the block structure of $\mathbf{A}_1$ we obtain

$$\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1 \;=\; \mathrm{diag}\left(\sum_{j=0}^{n-k}\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)_{k\in\{0,1,\ldots,n\}}$$

and

$$\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{Z}_1 \;=\; \left(\sum_{j=0}^{n-k}(\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k}\right)_{k\in\{0,1,\ldots,n\}}$$

Now the Gauss–Markov Theorem for estimators yields

$$\begin{aligned}\widehat{\boldsymbol{\beta}}^{\mathrm{GM}} &= \left(\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{Z}_1\\ &= \left(\left(\sum_{j=0}^{n-k}\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)^{-1}\sum_{j=0}^{n-k}(\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k}\right)_{k\in\{0,1,\ldots,n\}}\end{aligned}$$

and hence

$$\widehat{\boldsymbol{\zeta}}_k^{\mathrm{GM}} \;=\; \left(\sum_{j=0}^{n-k}\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)^{-1}\sum_{j=0}^{n-k}(\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k}$$

for all $k \in \{0,1,\ldots,n\}$. □

The following result provides formulas for the Gauss–Markov predictors of the non–observable incremental losses and for the Gauss–Markov predictors of the non–observable cumulative losses:

4.2 Theorem. *For all $i,k \in \{1,\ldots,n\}$ such that $i+k \geq n+1$, the Gauss–Markov predictor $\widehat{\mathbf{Z}}_{i,k}^{\mathrm{GM}}$ of $\mathbf{Z}_{i,k}$ satisfies*

$$\widehat{\mathbf{Z}}_{i,k}^{\mathrm{GM}} \;=\; \mathbf{V}_i\,\widehat{\boldsymbol{\zeta}}_k^{\mathrm{GM}}$$

and the Gauss–Markov predictor $\widehat{\mathbf{S}}_{i,k}^{\mathrm{GM}}$ *of* $\mathbf{S}_{i,k}$ *satisfies*

$$\widehat{\mathbf{S}}_{i,k}^{\mathrm{GM}} \;=\; \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \widehat{\boldsymbol{\zeta}}_l^{\mathrm{GM}}$$

Proof. Since $\boldsymbol{\Sigma}_{21} = \mathrm{Cov}[\mathbf{Z}_2, \mathbf{Z}_1] = \mathbf{O}$, the first assertion is immediate from the Gauss–Markov Theorem for predictors and the second assertion follows from the final remark of Section 3. □

The Gauss–Markov Theorem for predictors implies that

– the Gauss–Markov predictors of the sum of the non–observable incremental losses of a given accident year,
– the Gauss–Markov predictors of the sum of the non–observable incremental losses of a given calendar year, and
– the Gauss–Markov predictors of the sum of all non–observable incremental losses

are obtained by summation over the Gauss–Markov predictors of the corresponding single non–observable incremental losses.

For $i,k \in \{1,\dots,n\}$ such that $i+k \geq n+1$, the estimators and predictors

$$\begin{aligned}
\widehat{\boldsymbol{\zeta}}_k^{\mathrm{AD}} &:= \left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)^{-1} \sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k} \\
\widehat{\mathbf{Z}}_{i,k}^{\mathrm{AD}} &:= \mathbf{V}_i\,\widehat{\boldsymbol{\zeta}}_k^{\mathrm{AD}} \\
\widehat{\mathbf{S}}_{i,k}^{\mathrm{AD}} &:= \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \widehat{\boldsymbol{\zeta}}_l^{\mathrm{AD}}
\end{aligned}$$

are said to be the estimators and predictors of the *multivariate additive method.* Except for $m=1$ or $k=n$ they usually differ from the estimators and predictors

$$\begin{aligned}
\widetilde{\boldsymbol{\zeta}}_k &:= \left(\sum_{j=0}^{n-k} \mathbf{V}_j\right)^{-1} \sum_{j=0}^{n-k} \mathbf{Z}_{j,k} \\
\widetilde{\mathbf{Z}}_{i,k} &:= \mathbf{V}_i\,\widetilde{\boldsymbol{\zeta}}_k \\
\widetilde{\mathbf{S}}_{i,k} &:= \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \widetilde{\boldsymbol{\zeta}}_l
\end{aligned}$$

whose coordinates coincide with those of the univariate additive method.

4.2 Multivariate Chain–Ladder Model

Let us now consider a multivariate extension of the univariate chain–ladder model which applies to the combined subportfolios and was proposed by Pröhl and Schmidt [2005]. This model is a slight but convenient extension of the model of Braun [2004]; see also Kremer [2005].

The multivariate chain–ladder model involves successive conditioning with respect to the σ–algebras $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ where, for each $k \in \{1, \ldots, n\}$, the σ–algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $\mathbf{S}_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$, which is at the same time the information provided by the incremental losses $\mathbf{Z}_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$.

For all $i, k \in \{0, 1, \ldots, n\}$ we denote by

$$\boldsymbol{\Delta}_{i,k} \ := \ \mathrm{diag}(\mathbf{S}_{i,k})$$

the diagonal random matrix whose diagonal elements are the coordinates of the random vector $\mathbf{S}_{i,k}$.

We assume that all coordinates of $\mathbf{S}_{i,k}$ are strictly positive. Then each $\boldsymbol{\Delta}_{i,k}$ is invertible and the identity

$$\mathbf{S}_{i,k} \ = \ \boldsymbol{\Delta}_{i,k-1}\big(\boldsymbol{\Delta}_{i,k-1}^{-1}\mathbf{S}_{i.k}\big)$$

holds for all $i \in \{0, 1, \ldots, n\}$ and $k \in \{1, \ldots, n\}$.

> **Multivariate Chain–Ladder Model:** *For each $k \in \{1, \ldots, n\}$, there exists a random parameter vector $\boldsymbol{\Phi}_k$ and a positive definite symmetric random matrix $\boldsymbol{\Sigma}_k$ such that*
>
> $$E^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}] \ = \ \boldsymbol{\Delta}_{i,k-1}\,\boldsymbol{\Phi}_k$$
>
> *and*
>
> $$\mathrm{Cov}^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}, \mathbf{S}_{j,k}] \ = \ \begin{cases} \boldsymbol{\Delta}_{i,k-1}^{1/2}\,\boldsymbol{\Sigma}_k \boldsymbol{\Delta}_{i,k-1}^{1/2} & \textit{if } i = j \\ \mathbf{O} & \textit{else} \end{cases}$$
>
> *holds for all $i, j \in \{0, 1, \ldots, n-k+1\}$.*

In the subsequent discussion, we assume that the assumption of the multivariate chain–ladder model is fulfilled.

The multivariate chain–ladder model consists of n conditional linear models corresponding to the development years $k \in \{1, \ldots, n\}$. This can be seen as follows: Fix

$k \in \{1, \dots, n\}$, let $\mathbf{S}_1$ and $\mathbf{A}_1$ denote a block vector and a block matrix consisting of the random vectors $\mathbf{S}_{i,k}$ and the random matrices $\mathbf{\Delta}_{i,k}$ with $i \le n-k$ (arranged in the same order) and let $\mathbf{S}_2 := \mathbf{S}_{n-k+1,k}$ and $\mathbf{A}_2 = \mathbf{\Delta}_{n-k+1,k}$. Then the random vectors $\mathbf{S}_1$ and $\mathbf{S}_2$ and the random matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ depend on k and we have

$$\begin{aligned} E^{\mathcal{G}_{k-1}}[\mathbf{S}_1] &= \mathbf{A}_1 \mathbf{\Phi}_k \\ E^{\mathcal{G}_{k-1}}[\mathbf{S}_2] &= \mathbf{A}_2 \mathbf{\Phi}_k \end{aligned}$$

Therefore, the multivariate chain–ladder model consists indeed of n conditional linear models.

The following result provides formulas for the Gauss–Markov estimators of the parameters in the multivariate chain–ladder model:

4.3 Theorem. *For each $k \in \{1, \dots, n\}$, the Gauss–Markov estimator $\widehat{\mathbf{\Phi}}_k^{\mathrm{GM}}$ of $\mathbf{\Phi}_k$ satisfies*

$$\widehat{\mathbf{\Phi}}_k^{\mathrm{GM}} = \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{j,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{j,k-1}^{1/2} \right)^{-1} \sum_{j=0}^{n-k} \big(\mathbf{\Delta}_{j,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{j,k-1}^{1/2} \big) \mathbf{\Delta}_{j,k-1}^{-1} \mathbf{S}_{j,k}$$

Theorem 4.3 is immediate from the Gauss–Markov Theorem for estimators.

The following result provides formulas for the Gauss–Markov predictors of the cumulative losses of the first non–observable calendar year:

4.4 Theorem. *For each $i \in \{1, \dots, n\}$, the Gauss–Markov predictor $\widehat{\mathbf{S}}_{i,n-i+1}^{\mathrm{GM}}$ of $\mathbf{S}_{i,n-i+1}$ satisfies*

$$\widehat{\mathbf{S}}_{i,n-i+1}^{\mathrm{GM}} = \mathbf{\Delta}_{i,n-i} \widehat{\mathbf{\Phi}}_{n-i+1}^{\mathrm{GM}}$$

Theorem 4.4 is immediate from the Gauss–Markov Theorem for predictors.

For $i, k \in \{1, \dots, n\}$ such that $i + k \ge n+1$, the estimators and predictors

$$\begin{aligned} \widehat{\mathbf{\Phi}}_k^{\mathrm{CL}} &:= \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{j,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{j,k-1}^{1/2} \right)^{-1} \sum_{j=0}^{n-k} \big(\mathbf{\Delta}_{j,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{j,k-1}^{1/2} \big) \mathbf{\Delta}_{j,k-1}^{-1} \mathbf{S}_{j,k} \\ \widehat{\mathbf{S}}_{i,k}^{\mathrm{CL}} &:= \widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}} \widehat{\mathbf{\Phi}}_k^{\mathrm{CL}} \end{aligned}$$

with

$$\widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}} := \begin{cases} \mathrm{diag}(\mathbf{S}_{i,n-i}) & \text{if } k = n-i+1 \\ \mathrm{diag}(\widehat{\mathbf{S}}_{i,k-1}^{\mathrm{CL}}) & \text{else} \end{cases}$$

are said to be the estimators and predictors of the *multivariate chain–ladder method.* Except for $m = 1$ or $k = n$ they usually differ from the estimators and predictors

$$
\begin{aligned}
\widetilde{\mathbf{\Phi}}_k &:= \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{j,k-1}\right)^{-1} \sum_{j=0}^{n-k} \mathbf{S}_{j,k} \\
\widetilde{\mathbf{S}}_{i,k} &:= \widetilde{\mathbf{\Delta}}_{i,k}\widetilde{\mathbf{\Phi}}_k
\end{aligned}
$$

with

$$
\widetilde{\mathbf{\Delta}}_{i,k-1} := \begin{cases} \mathrm{diag}(\mathbf{S}_{i,n-i}) & \text{if } k = n-i+1 \\ \mathrm{diag}(\widetilde{\mathbf{S}}_{i,k-1}) & \text{else} \end{cases}
$$

whose coordinates coincide with those of the *univariate chain–ladder method.*

In the case $i + k = n + 1$, the multivariate chain–ladder predictors are justified by Theorem 4.4, but another justification is needed in the case $i + k \geq n + 2$; this can be achieved by minimizing the $\mathcal{G}_{k-1}$–conditional expected prediction error over the collection of all predictors $\widehat{\mathbf{S}}_{i,k}$ of $\mathbf{S}_{i,k}$ satisfying

$$
\widehat{\mathbf{S}}_{i,k} = \widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1}\widehat{\mathbf{\Phi}}_k
$$

for some $\mathcal{G}_{k-1}$–conditionally unbiased linear estimator $\widehat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$; see Schmidt [1999b] for the univariate case. We have the following result:

4.5 Theorem. *For all $i, k \in \{1, \ldots, n\}$ such that $i + k \geq n + 1$, the chain–ladder predictor $\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}$ minimizes the $\mathcal{G}_{k-1}$–conditional expected prediction error over all predictors $\widehat{\mathbf{S}}_{i,k}$ of $\mathbf{S}_{i,k}$ satisfying*

$$
\widehat{\mathbf{S}}_{i,k} = \widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1}\widehat{\mathbf{\Phi}}_k
$$

for some $\mathcal{G}_{k-1}$–conditionally unbiased linear estimator $\widehat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$.

A proof of Theorem 4.5 will be given in the Appendix.

The optimality of the multivariate chain–ladder method guaranteed by Theorem 4.5 is sequential and one–step ahead. Of course, one would like to have a condition ensuring some kind of global optimality of the chain–ladder predictors; however, even in the univariate case, no such condition seems to be known.

To illustrate the situation without introducing additional notation, let us recall two results for the univariate case:

– The assumption of the univariate chain–ladder model is fulfilled in the model of Mack [1993] in which it is assumed that the accident years are independent and that the parameters φ_k and σ_k are non–random; see Schmidt and Schnaus [1996]. Under the assumptions of the model of Mack, it can be shown that all chain–ladder predictors are unbiased, but it can also be shown that many other predictors are unbiased as well. Therefore, unbiasedness does not distinguish the chain–ladder predictors among all other predictors.

– One might hope that the chain–ladder predictors minimize the $\mathcal{G}_{n-i}$–conditional expected squared predictor error over all predictors of the form

$$\widehat{S}_{i,k} \;=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \widehat{\varphi}_l$$

where, for each $l \in \{n-i+1,\ldots,k\}$, $\widehat{\varphi}_l$ is a $\mathcal{G}_l$–conditionally unbiased linear estimator of φ_l. Again, under the assumptions of the model of Mack, it has been shown in Schmidt [1997] that this kind of optimality may fail for the chain–ladder predictors.

Thus, even in the univariate case and under the stronger assumptions of the model of Mack, it remains an open question whether there exists a condition which is less restrictive than the sequential optimality criterion of Theorem 4.5 and still ensures some kind of global optimality of the chain–ladder predictors.

5 Additivity

Let $\mathbf{1}$ denote the m–dimensional vector with all coordinates being equal to 1. For $i,k \in \{0,1,\ldots,n\}$ define

$$\begin{aligned} Z_{i,k} &:= \mathbf{1}'\mathbf{Z}_{i,k} \\ S_{i,k} &:= \mathbf{1}'\mathbf{S}_{i,k} \end{aligned}$$

We shall now study prediction of the non–observable incremental losses $Z_{i,k}$ and of the non–observable cumulative losses $S_{i,k}$ of the aggregate portfolio.

5.1 Multivariate Additive Model

In the multivariate additive model it is immediate from the Gauss–Markov Theorem for predictors that, for all $i,k \in \{1,\ldots,n\}$ such that $i+k \geq n+1$, the Gauss–Markov predictor $\widehat{Z}_{i,k}^{\mathrm{GM}}$ of $Z_{i,k}$ and the Gauss–Markov predictor $\widehat{S}_{i,k}^{\mathrm{GM}}$ of $S_{i,k}$ satisfy

$$\begin{aligned} \widehat{Z}_{i,k}^{\mathrm{GM}} &= \mathbf{1}'\widehat{\mathbf{Z}}_{i,k}^{\mathrm{AD}} \\ \widehat{S}_{i,k}^{\mathrm{GM}} &= \mathbf{1}'\widehat{\mathbf{S}}_{i,k}^{\mathrm{AD}} \end{aligned}$$

This means that the Gauss–Markov predictors for the aggregate portfolio are obtained by summation over the Gauss–Markov predictors for the single lines of business. Therefore, the multivariate additive method is consistent in the sense that there is no problem of additivity.

Warning: One might believe that the Gauss–Markov predictors for the aggregate portfolio could also be obtained by applying the univariate additive method to the aggregate portfolio. This, however, is not the case since the multivariate additive model for the combined subportfolios does not lead to a univariate additive model for the aggregate portfolio.

5.2 Multivariate Chain–Ladder Model

In the multivariate chain–ladder model it is immediate from the Gauss–Markov Theorem for predictors that, for all $i \in \{1, \ldots, n\}$, the Gauss–Markov predictor $\widehat{S}^{\mathrm{GM}}_{i,n-i+1}$ of $S_{i,n-i+1}$ satisfies

$$\widehat{S}^{\mathrm{GM}}_{i,n-i+1} = \mathbf{1}'\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,n-i+1}$$

This means that the Gauss–Markov predictors for the aggregate portfolio are obtained by summation over the multivariate Gauss–Markov predictors for the different lines of business. Moreover, it is easy to see that, for all $i, k \in \{1, \ldots, n\}$ such that $i + k \geq n + 2$, the predictor

$$\begin{aligned} S^*_{i,k} &:= \mathbf{1}'\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k} \\ &= \mathbf{1}'\widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1}\widehat{\mathbf{\Phi}}^{\mathrm{CL}}_k \\ &= \big(\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k-1}\big)'\widehat{\mathbf{\Phi}}^{\mathrm{CL}}_k \end{aligned}$$

minimizes the $\mathcal{G}_{k-1}$–conditional expected prediction error over all predictors $\widehat{S}_{i,k}$ of $S_{i,k}$ satisfying

$$\begin{aligned} \widehat{S}_{i,k} &= \mathbf{1}'\widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1}\widehat{\mathbf{\Phi}}_k \\ &= \big(\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k-1}\big)'\widehat{\mathbf{\Phi}}_k \end{aligned}$$

for some $\mathcal{G}_{k-1}$–conditionally unbiased linear predictor $\widehat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$. Therefore, the multivariate chain–ladder method is consistent in the sense that there is no problem of additivity.

Warning: As in the case of the multivariate additive model, it would be a serious mistake to predict the non–observable cumulative losses of the aggregate portfolio on the basis of the observable cumulative losses of the aggregate portfolio since such an approach would ignore the correlation structure between the different lines of business; see Pröhl and Schmidt [2005].

6 Estimation of the Variance Parameters

In the case $m = 1$, which is the univariate case, the variance parameters $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_n$ drop out in the formulas for the Gauss–Markov predictors in the multivariate additive model and in the multivariate chain–ladder model.

In the case $m \geq 2$, however, only the variance parameter $\mathbf{\Sigma}_n$ drops out in the formulas for the Gauss–Markov predictors in the multivariate additive model and in the multivariate chain–ladder model; in this case, the variance parameters $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_{n-1}$ must be estimated.

6.1 Multivariate Additive Model

Under the assumptions of the multivariate additive model and for $k \leq n-1$, the random matrix

$$\widehat{\boldsymbol{\Sigma}}_k^{\mathrm{AD}} := \frac{1}{n-k} \sum_{j=0}^{n-k} \mathbf{V}_j^{-1/2} \Big(\mathbf{Z}_{j,k} - \mathbf{V}_j \widetilde{\boldsymbol{\zeta}}_k\Big) \Big(\mathbf{Z}_{j,k} - \mathbf{V}_j \widetilde{\boldsymbol{\zeta}}_k\Big)' \mathbf{V}_j^{-1/2}$$

is a positive semidefinite estimator of the positive definite matrix $\boldsymbol{\Sigma}_k$; moreover, its diagonal elements are unbiased estimators of the diagonal elements of $\boldsymbol{\Sigma}_k$ whereas its non–diagonal elements slightly underestimate the corresponding elements of $\boldsymbol{\Sigma}_k$.

Although unbiasedness of an estimator is usually considered to be desirable, this property would not be helpful in the present situation since any estimator of $\boldsymbol{\Sigma}_k$ has to be inverted and since the inverse of an unbiased estimator of $\boldsymbol{\Sigma}_k$ is very likely to be biased anyway. Moreover, the relative bias of the estimators proposed before can be shown to be very small.

By contrast, for any estimator of $\boldsymbol{\Sigma}_k$, the property of being positive semidefinite is a necessary, although not sufficient, condition for being positive definite and hence invertible. In fact, the estimator of $\boldsymbol{\Sigma}_k$ proposed before is always singular when $k \geq n-m+2$ since in this case the dimension of the linear space generated by any realizations of the random vectors $\mathbf{V}_j^{-1/2}(\mathbf{Z}_{j,k} - \mathbf{V}_j \widetilde{\boldsymbol{\zeta}}_k)$ with $j \in \{0, 1, \ldots, n-k\}$ is at most $m-1$ such that there exists at least one nonzero vector which is orthogonal to each of the realizations of these random vectors; moreover, the realizations of the random vectors $\mathbf{V}_j^{-1/2}(\mathbf{Z}_{j,k} - \mathbf{V}_j \widetilde{\boldsymbol{\zeta}}_k)$ may be linearly dependent also for some $k \leq n-m+1$, which implies that the corresponding realization of the estimator of $\boldsymbol{\Sigma}_k$ proposed before may be singular also for some $k \leq n-m+1$.

In practical applications, it is thus necessary to check whether the estimators proposed before are invertible or not, and to modify those estimators which are not invertible. Such modifications could be obtained by extrapolation or by the use of external information; see below.

6.2 Multivariate Chain–Ladder Model

Under the assumptions of the multivariate chain–ladder model and for $k \leq n-1$, the random matrix

$$\widehat{\boldsymbol{\Sigma}}_k^{\mathrm{CL}} := \frac{1}{n-k} \sum_{j=0}^{n-k} \boldsymbol{\Delta}_{j,k-1}^{-1/2} \Big(\mathbf{S}_{j,k} - \boldsymbol{\Delta}_{j,k-1} \widetilde{\boldsymbol{\Phi}}_k\Big) \Big(\mathbf{S}_{j,k} - \boldsymbol{\Delta}_{j,k-1} \widetilde{\boldsymbol{\Phi}}_k\Big)' \boldsymbol{\Delta}_{j,k-1}^{-1/2}$$

is a positive semidefinite estimator of the positive definite matrix $\boldsymbol{\Sigma}_k$; moreover, its diagonal elements are unbiased estimators of the diagonal elements of $\boldsymbol{\Sigma}_k$ whereas its non–diagonal elements slightly underestimate the corresponding elements of $\mathbf{\Sigma}_k$ and hence differ from the unbiased estimators proposed by Braun [2004].

The comments on the variance estimators proposed for the multivariate additive model apply as well to the variance estimators proposed for the multivariate chain–ladder model.

6.3 Extrapolation

In the case where the proposed estimators of the variances for late development years are singular or almost singular, it could be reasonable to replace these estimators with estimators obtained by extrapolation from the estimators for the first development years which are usually invertible.

6.4 Iteration

In both models, one may try to improve the estimators of the variances and hence the Gauss–Markov estimators of the parameters by iteration, as proposed by Kremer [2005]. However, the iterates of some of the estimators of the variances may again be singular, and it seems to be difficult to prove that the resulting empirical Gauss–Markov estimators of the parameters are indeed improved by iteration.

6.5 External Information

In both models, another possibility for the estimation of the variance parameters $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_{n-1}$ consists in the use of external information, which is not contained in the run–off triangle and could be obtained, e. g., from the run–off triangle of a similar portfolio or from market statistics.

7 Remarks

Another bivariate model of loss reserving is the model of Quarg and Mack [2004]. Under the assumptions of their model, Quarg and Mack propose bivariate chain–ladder predictors for the paid and incurred cumulative losses of a single line of business with the aim of reducing the gap between the univariate chain–ladder predictors for the paid and incurred cumulative losses; see also Verdier and Klinger [2005] for a related model. None of these two models is contained in the multivariate models proposed in the present paper.

Since no conditions at all are imposed on the character of the different lines of business in the multivariate models presented here, the multivariate additive method and the multivariate chain–ladder method could, in principle, also be applied to the paid and incurred cumulative losses of a single line of business.

Let us finally note that the problem of additivity can also be solved in quite different models like credibility models; see Radtke and Schmidt [2004] and Schmidt [2004].

8 A Numerical Example

In this section we present a numerical example for the multivariate chain–ladder method in the case of $m = 2$ subportfolios and $n = 3$ development years.

8.1 The Data

The following run–off triangles contain the observable cumulative losses $S_{i,k}^{(1)}$, $S_{i,k}^{(2)}$, and $S_{i,k}$ of the two subportfolios and of the aggregate portfolio, respectively:

Subportfolio 1

AY	**DY**			
	0	1	2	3
0	2423	3123	3567	3812
1	2841	3422	3952	
2	3700	3977		
3	5231			

Subportfolio 2

AY	**DY**			
	0	1	2	3
0	3546	6578	7650	8123
1	4001	7566	8822	
2	4040	7813		
3	4300			

Aggregate Portfolio

AY	**DY**			
	0	1	2	3
0	5969	9701	11217	11935
1	6842	10988	12774	
2	7740	11790		
3	9531			

8.2 Univariate Chain–Ladder Method

Applying the univariate chain–ladder method to each of these run–off triangles yields the univariate chain–ladder factors (CLF) and the univariate chain–ladder predictors of the non–observable cumulative losses:

Subportfolio 1

AY	DY			
	0	1	2	3
0	2423	3123	3567	3812
1	2841	3422	3952	4223
2	3700	3977	4569	4883
3	5231	6140	7054	7538
CLF		1.1738	1.1488	1.0687

Subportfolio 2

AY	DY			
	0	1	2	3
0	3546	6578	7650	8123
1	4001	7566	8822	9367
2	4040	7813	9099	9662
3	4300	8148	9490	10076
CLF		1.8950	1.1646	1.0618

Aggregate Portfolio

AY	DY			
	0	1	2	3
0	5969	9701	11217	11935
1	6842	10988	12774	13592
2	7740	11790	13672	14547
3	9531	15063	17467	18585
CLF		1.5804	1.1596	1.0640

8.3 Multivariate Chain–Ladder Method

We now combine the run–off triangles of the two subportfolios into a single run–off triangle which contains the vectors $\mathbf{S}_{i,k}$ of cumulative losses:

Combined Subportfolios

AY	DY			
	0	1	2	3
0	$\begin{pmatrix} 2423 \\ 3546 \end{pmatrix}$	$\begin{pmatrix} 3123 \\ 6578 \end{pmatrix}$	$\begin{pmatrix} 3567 \\ 7650 \end{pmatrix}$	$\begin{pmatrix} 3812 \\ 8123 \end{pmatrix}$
1	$\begin{pmatrix} 2841 \\ 4001 \end{pmatrix}$	$\begin{pmatrix} 3422 \\ 7566 \end{pmatrix}$	$\begin{pmatrix} 3952 \\ 8822 \end{pmatrix}$	
2	$\begin{pmatrix} 3700 \\ 4040 \end{pmatrix}$	$\begin{pmatrix} 3977 \\ 7813 \end{pmatrix}$		
3	$\begin{pmatrix} 5231 \\ 4300 \end{pmatrix}$			

Transforming the vectors $\mathbf{S}_{i,k}$ of cumulative losses into diagonal matrices, we obtain the following run–off triangle for the matrices $\boldsymbol{\Delta}_{i,k} = \text{diag}(\mathbf{S}_{i,k})$ which is completed by the vectors $\widetilde{\boldsymbol{\Phi}}_k$ of univariate chain–ladder factors:

Combined Subportfolios

AY	DY			
	0	1	2	3
0	$\begin{pmatrix} 2423 & 0 \\ 0 & 3546 \end{pmatrix}$	$\begin{pmatrix} 3123 & 0 \\ 0 & 6578 \end{pmatrix}$	$\begin{pmatrix} 3567 & 0 \\ 0 & 7650 \end{pmatrix}$	$\begin{pmatrix} 3812 & 0 \\ 0 & 8123 \end{pmatrix}$
1	$\begin{pmatrix} 2841 & 0 \\ 0 & 4001 \end{pmatrix}$	$\begin{pmatrix} 3422 & 0 \\ 0 & 7566 \end{pmatrix}$	$\begin{pmatrix} 3952 & 0 \\ 0 & 8822 \end{pmatrix}$	
2	$\begin{pmatrix} 3700 & 0 \\ 0 & 4040 \end{pmatrix}$	$\begin{pmatrix} 3977 & 0 \\ 0 & 7813 \end{pmatrix}$		
3	$\begin{pmatrix} 5231 & 0 \\ 0 & 4300 \end{pmatrix}$			
$\widetilde{\boldsymbol{\Phi}}_k$		$\begin{pmatrix} 1.1738 \\ 1.8950 \end{pmatrix}$	$\begin{pmatrix} 1.1488 \\ 1.1646 \end{pmatrix}$	$\begin{pmatrix} 1.0687 \\ 1.0618 \end{pmatrix}$

For the estimators of the variances we thus obtain

$$\widehat{\boldsymbol{\Sigma}}_1^{\text{CL}} = \begin{pmatrix} 35.4968 & -14.3861 \\ -14.3861 & 5.9200 \end{pmatrix}$$

$$\widehat{\boldsymbol{\Sigma}}_2^{\text{CL}} = \begin{pmatrix} 0.2637 & 0.0926 \\ 0.0926 & 0.0325 \end{pmatrix}$$

and hence

$$\left(\widehat{\boldsymbol{\Sigma}}_1^{\mathrm{CL}}\right)^{-1} = \begin{pmatrix} 1.8616 & 4.5239 \\ 4.5239 & 11.1624 \end{pmatrix}$$

$$\left(\widehat{\boldsymbol{\Sigma}}_2^{\mathrm{CL}}\right)^{-1} = \begin{pmatrix} 25876.4330 & -73727.6467 \\ -73727.6467 & 210097.0596 \end{pmatrix}$$

Note that estimators of the variances $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_3$ are not needed. Applying the multivariate chain–ladder method to the combined subportfolios yields the multivariate chain–ladder predictors of the non–observable cumulative losses:

Combined Subportfolios

AY	**DY** 0	1	2	3
0	$\begin{pmatrix} 2423 \\ 3546 \end{pmatrix}$	$\begin{pmatrix} 3123 \\ 6578 \end{pmatrix}$	$\begin{pmatrix} 3567 \\ 7650 \end{pmatrix}$	$\begin{pmatrix} 3812 \\ 8123 \end{pmatrix}$
1	$\begin{pmatrix} 2841 \\ 4001 \end{pmatrix}$	$\begin{pmatrix} 3422 \\ 7566 \end{pmatrix}$	$\begin{pmatrix} 3952 \\ 8822 \end{pmatrix}$	$\begin{pmatrix} 4223 \\ 9367 \end{pmatrix}$
2	$\begin{pmatrix} 3700 \\ 4040 \end{pmatrix}$	$\begin{pmatrix} 3977 \\ 7813 \end{pmatrix}$	$\begin{pmatrix} 4569 \\ 9099 \end{pmatrix}$	$\begin{pmatrix} 4883 \\ 9661 \end{pmatrix}$
3	$\begin{pmatrix} 5231 \\ 4300 \end{pmatrix}$	$\begin{pmatrix} 6105 \\ 8167 \end{pmatrix}$	$\begin{pmatrix} 7013 \\ 9512 \end{pmatrix}$	$\begin{pmatrix} 7495 \\ 10100 \end{pmatrix}$
$\widehat{\boldsymbol{\Phi}}_k$		$\begin{pmatrix} 1.1670 \\ 1.8994 \end{pmatrix}$	$\begin{pmatrix} 1.1489 \\ 1.1646 \end{pmatrix}$	$\begin{pmatrix} 1.0687 \\ 1.0618 \end{pmatrix}$

8.4 Comparison

Predictors for non–observable aggregate cumulative losses may be computed by the following three methods:

- **Method A:** Apply the univariate chain–ladder method to the aggregate portfolio.
- **Method B:** Apply the univariate chain–ladder method to each of the subportfolios and take sums of the univariate predictors.
- **Method C:** Apply the multivariate chain–ladder method to the combined subportfolios and take sums of the multivariate predictors.

For example, for the ultimate aggregate cumulative loss of accident year 3,

- Method A yields the value 18585.
- Method B yields the value $7538 + 10076 = 17614$.
- Method C yields the value $7495 + 10100 = 17595$.

The following table presents several reserves obtained by these three methods:

Reserve	Method A	Method B	Method C
Accident Year 1	818	817	817
Accident Year 2	2757	2754	2754
Accident Year 3	9054	8084	8064
Total	12628	11655	11635
Calendar Year 4	8231	7452	7436
Calendar Year 5	3279	3131	3129
Calendar Year 6	1118	1071	1070
Total	12628	11655	11635

Due to round–off errors, some of the total reserves differ slightly from the sums of the reserves over accident years of calendar years.

In the present example, the results obtained by Methods B and C are quite similar, but they differ considerably from those obtained by Method A.

8.5 Preliminary Conclusions

Of course, one should not draw general conclusions from a single numerical example. Nevertheless, the present example and experience with other sets of data justify the following rules of thumb:

- Method C is optimal when the model assumptions and the optimality criteria for the multivariate chain–ladder method can be accepted.
- Method B may in many cases provide a reasonable approximation of Method C.
- Method A may be disastrous since it ignores correlation between the different lines of business.

Experience with other sets of data also indicates that the similarities and differences between the three methods may vary with

- the lines of business under consideration,
- the number of lines of business, and
- the number of development years.

It is therefore indispensable for the actuary to acquire practical experience for every combined portfolio of interest.

Appendix

Here we give a proof of Theorem 4.5.

Proof. Consider any $\mathcal{G}_{k-1}$–conditionally unbiased linear estimator $\widehat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$. Then there exist $\mathcal{G}_{k-1}$–measurable matrices $\mathbf{Q}_{0,k-1}, \mathbf{Q}_{1,n-1}, \dots, \mathbf{Q}_{n-k,k-1}$ satisfying

$$\widehat{\mathbf{\Phi}}_k \quad = \quad \sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}\mathbf{S}_{j,k}$$

and $\sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}\mathbf{\Delta}_{j,k-1} = \mathbf{I}$. Also, letting

$$\mathbf{Q}_{j,k-1}^{\mathrm{CL}} := \left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{s,k-1}^{1/2}\right)^{-1} \left(\mathbf{\Delta}_{j,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{j,k-1}^{1/2}\right)\mathbf{\Delta}_{j,k-1}^{-1}$$

we obtain

$$\widehat{\mathbf{\Phi}}_k^{\mathrm{CL}} = \sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\mathbf{S}_{j,k}$$

and $\sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\mathbf{\Delta}_{j,k-1} = \mathbf{I}$. We thus obtain

$$\sum_{j=0}^{n-k}\left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathbf{\Delta}_{j,k-1} = \mathbf{O}$$

Since

$$\mathbf{Q}_{j,k-1}^{\mathrm{CL}} = \left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{s,k-1}^{1/2}\right)^{-1} \mathbf{\Delta}_{j,k-1}\left(\mathrm{Var}^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}]\right)^{-1}$$

this yields

$$\begin{aligned}
\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\widehat{\mathbf{\Phi}}_k - \widehat{\mathbf{\Phi}}_k^{\mathrm{CL}}, \widehat{\mathbf{\Phi}}_k^{\mathrm{CL}}\right] &= \sum_{j=0}^{n-k}\sum_{l=0}^{n-k}\left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}, \mathbf{S}_{l,k}\right]\left(\mathbf{Q}_{l,k-1}^{\mathrm{CL}}\right)' \\
&= \sum_{j=0}^{n-k}\left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathrm{Var}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}\right]\left(\mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)' \\
&= \sum_{j=0}^{n-k}\left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathbf{\Delta}_{j,k-1}\left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{s,k-1}^{1/2}\right)^{-1} \\
&= \mathbf{O}
\end{aligned}$$

Since $i + k \geq n + 1$, we also have $\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}, \mathbf{S}_{i,k}\right] = \mathbf{O}$ and thus

$$\begin{aligned}
\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\widehat{\mathbf{S}}_{i,k} - \widehat{\mathbf{S}}_{i,k}^{\mathrm{CL}}, \mathbf{S}_{i,k}\right] &= \mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}}\widehat{\mathbf{\Phi}}_k - \widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}}\widehat{\mathbf{\Phi}}_k^{\mathrm{CL}}, \mathbf{S}_{i,k}\right] \\
&= \widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}}\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\widehat{\mathbf{\Phi}}_k - \widehat{\mathbf{\Phi}}_k^{\mathrm{CL}}, \mathbf{S}_{i,k}\right] \\
&= \widehat{\mathbf{\Delta}}_{i,k-1}^{\mathrm{CL}}\sum_{j=0}^{n-k}\left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathrm{Cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}, \mathbf{S}_{i,k}\right] \\
&= \mathbf{O}
\end{aligned}$$

Using the two identities established before, we thus obtain

$$\begin{aligned}
\mathrm{Cov}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k},\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big] &= \mathrm{Cov}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k},\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}\Big] \\
&= \widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1}\mathrm{Cov}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{\Phi}}_{k}-\widehat{\mathbf{\Phi}}^{\mathrm{CL}}_{k},\widehat{\mathbf{\Phi}}^{\mathrm{CL}}_{k}\Big]\,\widehat{\mathbf{\Delta}}^{\mathrm{CL}}_{i,k-1} \\
&= \mathbf{O}
\end{aligned}$$

and hence

$$\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\Big] = \mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}\Big]+\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big]$$

We thus obtain

$$\begin{aligned}
& E^{\mathcal{G}_{k-1}}\Big[\Big(\widehat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\Big)'\Big(\widehat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\Big)\Big] \\
&= \mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\Big]\Big) \\
&= \mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}_{i,k}-\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}\Big]\Big)+\mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big]\Big) \\
&\geq \mathrm{trace}\Big(\mathrm{Var}^{\mathcal{G}_{k-1}}\Big[\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big]\Big) \\
&= E^{\mathcal{G}_{k-1}}\Big[\Big(\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big)'\Big(\widehat{\mathbf{S}}^{\mathrm{CL}}_{i,k}-\mathbf{S}_{i,k}\Big)\Big]
\end{aligned}$$

which proves the theorem. $\square$

9 Acknowledgement

I am grateful to Leigh Halliwell for his comments on earlier drafts of this paper and to Mathias Zocher who provided the numerical example.

10 References

Ajne, B. [1994]: *Additivity of chain–ladder projections.* ASTIN Bull. **24**, 313–318.

Braun, C. [2004]: *The prediction error of the chain–ladder method applied to correlated run–off triangles.* ASTIN Bull. **34**, 399–423.

Goldberger, A. S. [1962]: *Best linear unbiased prediction in the generalized linear regression model.* J. Amer. Statist. Assoc. **57**, 369–375.

Hachemeister, C. A., and Stanard, J. N. [1975]: *IBNR claims count estimation with static lag functions.* Unpublished.

Halliwell, L. J. [1996]: *Loss prediction by generalized least squares.* Proc. CAS **83**, 436–489.

Halliwell, L. J. [1997]: *Conjoint prediction of paid and incurred losses.* CAS Forum Summer 1997, vol. I, pp. 241–379.

Halliwell, L. J. [1999]: *Loss prediction by generalized least squares – Author's response.* Proc. CAS **86**, 764–769.

Hamer, M. D. [1999]: *Loss prediction by generalized least squares – Discussion.* Proc. CAS **86**, 748–763.

Hess, K. T., and Schmidt, K. D. [2002]: *A comparison of models for the chain–ladder method.* Insurance Mathematics and Economics **31**, 351–364.

Hess, K. T., Schmidt, K. D., and Zocher, M. [2006]: *Multivariate loss reserving in the multivariate additive model.* Insurance Mathematics and Economics (in press).

Holmberg, R. D. [1994]: *Correlation and the measurement of loss reserve variability.* CAS Forum Spring 1994, pp. 247–278.

Klemmt, H. J. [2004]: *Trennung von Schadenarten und Additivität bei Chain–Ladder Prognosen.* Paper presented at the 2004 Fall Meeting of the German ASTIN Group in München.

Kremer, E. [2005]: *The correlated chain–ladder method for reserving in case of correlated claims developments.* Blätter DGVFM **27**, 315–322.

Mack, T. [1993]: *Distribution–free calculation of the standard error of chain–ladder reserve estimates.* ASTIN Bull. **23**, 213–225.

Pröhl, C., and Schmidt, K. D. [2005]: *Multivariate chain–ladder.* Dresdner Schriften zur Versicherungsmathematik 3/2005. Paper presented at the International ASTIN Colloquium 2005 in Zürich.

Quarg, G., and Mack, T. [2004]: *Munich chain–ladder – A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses.* Blätter DGVFM **26**, 597–630.

Radtke, M., and Schmidt, K. D. (eds.) [2004]: *Handbuch zur Schadenreservierung.* Karlsruhe: Verlag Versicherungswirtschaft.

Rao, C. R., and Toutenburg, H. [1995]: *Linear Models – Least Squares and Alternatives.* Berlin – Heidelberg – New York: Springer.

Schmidt, K. D. [1997]: *Non–optimal prediction by the chain–ladder method.* Insurance, Mathematics and Economics **21**, 17–24.

Schmidt, K. D. [1998]: *Prediction in the linear model – A direct approach.* Metrika **48**, 141–147.

Schmidt, K. D. [1999a]: *Loss prediction by generalized least squares – Discussion.* Proc. CAS **86**, 736–747.

Schmidt, K. D. [1999b]: *Chain ladder prediction and asset liability management.* Blätter DGVM **24**, 1–9.

Schmidt, K. D. [2004]: *Prediction.* In: *Encyclopedia of Actuarial Science*, vol. 3, pp. 1317–1321. Chichester: Wiley.

Schmidt, K. D., and Schnaus, A. [1996]: *An extension of Mack's model for the chain–ladder method.* ASTIN Bull. **26**, 247–262.

Verdier, B., and Klinger, A. [2005]: *JAB chain – A model based calculation of paid and incurred loss development factors.* Paper presented at the International ASTIN Colloquium 2005 in Zürich.

Klaus D. Schmidt
Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden
D–01062 Dresden

e–mail: klaus.d.schmidt@tu-dresden.de

July 15, 2006