PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1108 — 2006
Statystyka aktuarialna – stan i perspektywy rozwoju w Polsce

Wojciech Otto
University of Warsaw

INSURER'S SURPLUS MODEL WITH VARYING RISK PARAMETER AND DELAYED REPORTING

1. Introduction

The simple standard discrete-time model of the insurer's surplus process S_t assumes:

$$S_t = S_{t-1} + W_t,\ t = 1, 2, \ldots,$$

where $W_1, W_2, \ldots$ are i.i.d. random variables, claimed to represent yearly premium less yearly aggregate claims, and the initial surplus S_0 is fixed. Typically it is assumed that the premium component of W_t is constant, and the distribution of W_t is known.

In real life however, premium is written in advance to cover claims over the coming exposure period that are often reported and paid a number of years later. The inadequacy is even more obvious in the case of the continuous-time model, where the time elapsed between receiving premium and paying (eventually) compensations is totally neglected. In order to restore correspondence of the model to real life processes, we could change the interpretation of variables involved. The variable W_t could be interpreted as corresponding to accounting concepts of premium earned and claims occurred (claims paid plus increment of the outstanding claims reserve). This leads to interpreting the surplus as the amount of free assets, and consequently the ruin as insolvency. Under this interpretation the surplus model is meaningful for practice, as in fact focuses on phenomena of crucial importance for all involved parties: shareholders, tax authority, policyholders, and insurance supervision. However, the problem arises when we take into account that:

- outstanding claims amount is a random variable, and the corresponding reserve is in fact its point predictor, based on information available at the accounting date. Additional problem that makes predictions complex is that:
- in real life the risk parameter characterising the claim process changes as time goes on, and predictions are needed as well for premium setting as for reserving.

The paper concerns on incorporating the two above mentioned complications into the model. The incorporation could be read as such reinterpretation of the surplus S_t itself and the increment W_t, that leaves classical probabilistic assumptions unaffected. So, in a way the paper is focused on "calibrating" the basic model to the empirical evidence. However, some more important consequences arise when the basic surplus model is generalized by allowing for intervention. Typically, interventions are interpreted as:

- ceding a certain part of risk (and premium) to the reinsurer when the current surplus drops down the lower threshold,
- starting dividend payments when the surplus goes beyond the upper threshold.

In terms of the model, interventions are just modifications of the distribution of the increment W_t, undertaken immediately in response to the current level of the surplus S_{t-1}. This leads to the illusion of perfect controllability of the process. The illusion is especially apparent in some versions of the model (continuous time, diffusion approximation) when such results as zero ruin probability are obtained. Such result is in contradiction to the empirical evidence, where bankruptcies of insurers do happen, and could not be explained just by stupidity or fraudulent management. However, such results are only partially due to simplifications of the probabilistic structure of the process. They are rather due to neglecting the time elapsing between the information on the current state of the business and the effects of interventions undertaken.

There are no explicit considerations on the stochastic control issues in the paper. However, the basic surplus model is proposed, that allows next for incorporating the intervention mechanisms in a way that is at hand in real life.

Techniques used in the paper resembles in general those used by Scheike [3], who introduced the notion of fair premium for the claim process with dependent increments, applying to this purpose the Doob-Meyer decomposition for submartingales. Model assumptions have been chosen so as to enable casting the model into the state-space form, which allows for explicit expressions of premium and reserves as predictors of respective claim payments. Although using Kalman Filter for reserving has been proposed as early as in 1983 [1], its application for restoring the correspondence of the simple surplus model to real life processes is, to my best knowledge, new.

2. Notations, assumptions, and immediate results

Two sets of assumptions are taken: first set concerns the nature of the claim process, whereas the second concerns information on quantities of interest needed to define the increment of the insurer surplus process (which quantities are known, which of them are observed, and which have to be predicted).

In order to describe the (stylized) real life process, we will use following notations:

- $X_{t,j}$ – the amount of claims occurred in year t and paid in year $t+j$,
- $OX_t = \sum_{j=0}^{J} X_{t,j}$ – the aggregate amount of claims occurred in year t,
- $PX_t = \sum_{j=0}^{J} X_{t-j,j}$ – the aggregate amount of claims paid in year t,
- $LX_t = \sum_{k=1}^{J}\sum_{j=k}^{J} X_{t+1-k,j}$ – claims outstanding (liabilities) at the end of year t,
- J – is the maximum delay in claims settlement.

Figures $X_{t,j}$ ordered by years of claims occurrence (rows) and years of delay (columns) are depicted in Table 1. Diagonals represent claims that are paid in the same year. The sum of all elements shown on the shadow background can be expressed by two equivalent sums of aggregates:

$$LX_{t-1} + OX_t = PX_t + LX_t.$$

On the LHS claims are classified by time of occurrence with LX_{t-1} and OX_t representing claims occurred before and after the beginning of year t. On the RHS the same claims are classified by time of payment with PX_t and LX_t representing claims paid before and after the end of year t.

The aggregate claim process is assumed to depend on the underlying process of risk parameter μ_t. Given the value of this parameter, all components $X_{t,j}$ of the variable OX_t are normally distributed, with expectations and covariances equal to:

- $\mathrm{E}\left(X_{t,j} \middle| \mu_t\right) = \mu_t \cdot r_j$,
- $\mathrm{Cov}\left(X_{t,j}, X_{\tau,i} \middle| \mu_t, \mu_\tau\right) = \begin{cases} \sigma^2 \cdot r_j & \text{when} \quad t=\tau \text{ and } j=i, \\ 0 & \text{when} \quad t \neq \tau \text{ or } j \neq i, \end{cases}$
- $r_0, r_1, \ldots, r_J$ are non-negative delay coefficients that sum up to unity: $\sum_{j=0}^{J} r_j = 1$.

Table 1. Aggregate claims by calendar years of claim occurrence and by number of years of delay. Shadow area corresponds to amounts paid after the beginning of year t

Delay in years / Year of occurrence	0	1	...	...	...	$J-1$	J
...	...	...	...	...	...	...	...
$t-J-1$	$X_{t-J-1,0}$	$X_{t-J-1,1}$	$X_{t-J-1,2}$	...	...	$X_{t-J-1,J-1}$	$X_{t-J-1,J}$
$t-J$	$X_{t-J,0}$	$X_{t-J,1}$	$X_{t-J,2}$	...	...	$X_{t-J,J-1}$	$X_{t-J,J}$
$t-J+1$	$X_{t-J+1,0}$	$X_{t-J+1,1}$	$X_{t-J+1,2}$	...	...	$X_{t-J+1,J-1}$	$X_{t-J+1,J}$
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
$t-2$	$X_{t-2,0}$	$X_{t-2,1}$	$X_{t-2,2}$	...	...	$X_{t-2,J-1}$	$X_{t-2,J}$
$t-1$	$X_{t-1,0}$	$X_{t-1,1}$	$X_{t-1,2}$	...	...	$X_{t-1,J-1}$	$X_{t-1,J}$
t	$X_{t,0}$	$X_{t,1}$	$X_{t,2}$	...	...	$X_{t,J-1}$	$X_{t,J}$

As a result the aggregate variables OX_t are also normal (given μ_t) with parameters:

$$\operatorname{Cov}\left(OX_t, OX_\tau \middle| \mu_t, \mu_\tau\right) = \begin{cases} \sigma^2 & \text{when} \quad t = \tau, \\ 0 & \text{when} \quad t \neq \tau. \end{cases}$$

The risk parameter μ_t is assumed to follow an autoregressive process of order n:

$$\mu_t = a_1\mu_{t-1} + a_2\mu_{t-2} + \ldots + a_n\mu_{t-n} + (1 - a_1 - a_2 - a_n)\mu + \gamma_t,$$

with fixed coefficients $a_1, a_2, \ldots, a_n, \mu$ and independent normally distributed disturbances $\gamma_t = \mu_t - \mathrm{E}\left(\mu_t \middle| \mu_{t-1}, \mu_{t-2}, \ldots, \mu_{t-n}\right)$ characterized by:

$$\mathrm{E}(\gamma_t) = 0,$$

$$\operatorname{Cov}(\gamma_t, \gamma_\tau) = \begin{cases} \phi^2 & \text{when} \quad t = \tau, \\ 0 & \text{when} \quad t \neq \tau. \end{cases}$$

It is also assumed that disturbance term γ_t is independent of all past, current and future conditional deviations of the claim process. Denoting these deviations by:

$$\varepsilon_{t,j} \doteq X_{t,j} - \mathrm{E}\left(X_{t,j} \middle| \mu_t\right),$$

we can thus write:

$$\mathrm{Cov}\left(\gamma_\tau, \varepsilon_{t,j}\right) = 0 \text{ for arbitrary } \tau,\ t,\ j.$$

Unless otherwise stated, it is assumed also that both variances σ^2 and ϕ^2 are positive.

Remark 1

No more assumptions are needed to conclude that conditional moments $\mathrm{E}\left(\mu_t \middle| \mu_{t-k}, \ldots, \mu_{t-k-n}\right)$ and $\mathrm{Var}\left(\mu_t \middle| \mu_{t-k}, \ldots, \mu_{t-k-n}\right)$ are well defined for $k = 1, 2, \ldots$. If in addition the process were stationary, then unconditional moments $\lim_{k\to\infty} \mathrm{E}\left(\mu_t \middle| \mu_{t-k}, \ldots, \mu_{t-k-n}\right)$ and $\lim_{k\to\infty} \mathrm{Var}\left(\mu_t \middle| \mu_{t-k}, \ldots, \mu_{t-k-n}\right)$ are well defined, too. In the case of AR(1) process conditional moments are given simply as: $\mathrm{E}\left(\mu_t \middle| \mu_{t-k}\right) = a_1^k \mu_{t-k} + \left(1 - a_1^k\right)\mu$, and $\mathrm{Var}\left(\mu_t \middle| \mu_{t-k}\right) = \left(1 + a_1^2 + \ldots + a_1^{2(k-1)}\right)\phi^2$, stationarity is ensured for $\left|a_1\right| < 1$, and then unconditional moments equal μ and $\phi^2 / \left(1 - a_1^2\right)$, respectively.

Throughout the rest of the paper we will assume that the following quantities are known:

- parameters $a_1, a_2, \ldots, a_n$, μ and ϕ^2 of the process of the risk parameter μ_t,
- parameters $r_0, r_1, \ldots, r_J$ of the delay distribution,
- conditional variance parameter σ^2.

Moreover, we will assume that at the end of year t we observe all figures on claims paid $X_{t,0}, X_{t-1,1}, \ldots, X_{t-J,J}$ during this year. Let us denote the set of observations made at the end of year t and all past years by $\Im_t$. As a result we obtain the ascending sequence of sets of information $\ldots \subset \Im_{t-1} \subset \Im_t \subset \Im_{t+1} \subset \ldots$. Of course, some starting conditions for the underlying risk process μ_t and for observation process $X_{t,0}, X_{t-1,1}, \ldots, X_{t-J,J}$ have to be assumed, but this will be considered later.

Now the ultimate surplus at the end of year t can be defined naturally as:

$$US_t \doteq A_t - LX_t,$$

where A_t denotes the value of assets at the end of year t. The notion of ultimate surplus makes sense provided writing new business has been stopped and we ask about the final balance after settlement of all payments. However, the value of the ultimate surplus US_t will be known J years later. When the surplus at the moment t is to be assessed just at this moment then its definition must be based on information available then. So we will use the notion "surplus" for such an assessment:

$$S_t \doteq \mathrm{E}\left(US_t \middle| \Im_t\right) = A_t - \mathrm{E}\left(LX_t \middle| \Im_t\right),$$

where the expectation operator E serves here as the best unbiased prediction rule. Once the surplus is defined, the final balance can be defined too:

$$FB_t \doteq US_t - S_t .$$

Let us assume that the premium c_t written at the beginning of year t consists of a predictor of the amount of claims that will occur during this year and a fixed loading c:

$$c_t = c + \mathrm{E}\left(OX_t \,|\, \Im_{t-1}\right).$$

Under this assumption (supplemented by the assumption that assets are not invested) we can define the increment of the surplus as:

$$\Delta S_t = c + \mathrm{E}\left(OX_t \,|\, \Im_{t-1}\right) - PX_t + \mathrm{E}\left(LX_{t-1} \,|\, \Im_{t-1}\right) - \mathrm{E}\left(LX_t \,|\, \Im_t\right).$$

First three components of the RHS represent current cash flows (premium less claim expenses), whereas last two represent reduction of the surplus by the increment of the outstanding claims reserve. Let us notice that obviously $\mathrm{E}\left(PX_t \,|\, \Im_t\right) \equiv PX_t$. This allows for another rearrangement:

$$\Delta S_t = c + \mathrm{E}\left(LX_{t-1} + OX_t \,|\, \Im_{t-1}\right) - \mathrm{E}\left(PX_t + LX_t \,|\, \Im_t\right)$$

showing that the increment of the surplus is just a sum of loading c and the difference of expectations of the same random variable (represented either by $LX_{t-1} + OX_t$ or by $PX_t + LX_t$) conditional on $\Im_{t-1}$ and $\Im_t$, respectively.

Some general results could be obtained immediately. First of all, it could be shown easily that the surplus process S_t is a submartingale with drift c. This is because:

$$\mathrm{E}\left(\Delta S_t \,|\, S_{t-1}\right) = c + \mathrm{E}\left\{\mathrm{E}\left(LX_{t-1} + OX_t \,|\, \Im_{t-1}\right) \middle| S_{t-1}\right\} - \mathrm{E}\left\{\mathrm{E}\left(PX_t + LX_t \,|\, \Im_t\right) \middle| S_{t-1}\right\},$$

but information contained in S_{t-1} is included in the set $\Im_{t-1}$, and of course also in the set $\Im_t$, so making use of the iterative expectation rule we obtain:

$$\mathrm{E}\left(\Delta S_t \,|\, S_{t-1}\right) = c + \mathrm{E}\left(LX_{t-1} + OX_t - PX_t - LX_t \,|\, S_{t-1}\right) = c,$$

that finally leads to the conclusion:

$$\mathrm{E}\left(\Delta S_t \,|\, S_{t-1}\right) = c .$$

The above argument shows that the premium formula is a sum of the constant c and a "fair premium" as defined by Scheike [3]. Following general ideas of Scheike, the Doob-Meyer decomposition for submartingales can be used also to show that the final balance FB_t and increments $\Delta S_t, \Delta S_{t-1}, \Delta S_{t-2}, \ldots$ are all mutually uncorrelated random variables. The derivation presented below is based on more elementary rules of probability calculus. At first let us consider the covariance of the final balance with one of the preceding increments of the surplus process. Using well known formula for decomposition of covariance we obtain:

$$\operatorname{Cov}\left(FB_t, \Delta S_{t-k}\right) = \mathrm{E}\left\{\operatorname{Cov}\left(FB_t, \Delta S_{t-k} \mid \Im_t\right)\right\} + \operatorname{Cov}\left\{\mathrm{E}\left(FB_t \mid \Im_t\right), \mathrm{E}\left(\Delta S_{t-k} \mid \Im_t\right)\right\}.$$

However, for any $k = 0, 1, 2, \ldots$ both components equal zero. The first component equals zero because ΔS_{t-k} given $\Im_t$ is fixed, whereas the second component equals zero because $\mathrm{E}\left(FB_t \mid \Im_t\right)$ equals zero. Hence we obtain:

$$\operatorname{Cov}\left(FB_t, \Delta S_{t-k}\right) = 0.$$

In order to show that increments of the surplus process are not serially correlated, similar decomposition of covariance could be considered:

$$\operatorname{Cov}\left(\Delta S_t, \Delta S_{t-k}\right) = \mathrm{E}\left\{\operatorname{Cov}\left(\Delta S_t, \Delta S_{t-k} \mid \Im_{t-1}\right)\right\} + \operatorname{Cov}\left\{\mathrm{E}\left(\Delta S_t \mid \Im_{t-1}\right), \mathrm{E}\left(\Delta S_{t-k} \mid \Im_{t-1}\right)\right\}$$

In this case for any $k = 1, 2, 3, \ldots$ both components equal zero. The first component equals zero because ΔS_{t-k} given $\Im_{t-1}$ is fixed, whereas the second component equals zero because $\mathrm{E}\left(\Delta S_t \mid \Im_{t-1}\right)$ is fixed. Hence we obtained also the result:

$$\operatorname{Cov}\left(\Delta S_t, \Delta S_{t-k}\right) = 0.$$

To the contrary, the attempt to repeat the same argument for the increments of the ultimate surplus process US_t generally fails.

Remark 2

We have just derived, that reserving and premium setting based on expected values (premium supplemented by constant loading) stands for a proper "calibration" of the classical insurer's surplus model. Several detailed assumptions taken in this section have not been necessary to come to the above conclusion (nor linearity of the time series process μ_t neither normality and constant variances of both deviations γ_t and $\varepsilon_{t,j}$ have been explored so far).

3. Casting the surplus process into the state-space form

In this section the matrix notation is introduced in order to cast the surplus process into the state-space form. Kalman filtering technique is used then to derive explicit expressions for the premium formula, outstanding claims reserve, variance of the increment of the process S_t and variance of the final balance FB_t.

Let us take following notations:

$X_t = \left[X_{t,0}, X_{t-1,1}, ..., X_{t-J,J}\right]'$, a column vector of claims paid in the year t,

$m_t = \left[\mu_t, \mu_{t-1}, ..., \mu_{t-J}\right]'$, a column vector of current and past values of risk parameter μ_t,

$E_t = \left[\varepsilon_{t,0}, \varepsilon_{t-1,1}, ..., \varepsilon_{t-J,J}\right]'$, a vector of deviations of paid-claims figures around their conditional expectations,

$m_{t/\tau} = \mathrm{E}\left(m_t \mid \Im_\tau\right)$, the best unbiased predictor (BUP) of the vector m_t given information collected until the end of year τ;

$X_{t/\tau} = \mathrm{E}\left(X_t \mid \Im_\tau\right)$, the BUP of the vector X_t given information collected until the end of year τ (of course for $\tau \geq t$ X_t is known, so for this case $X_{t/\tau} \equiv X_t$),

$V_{t/\tau} = \mathrm{E}\left(m_{t/\tau} - m_t\right)\left(m_{t/\tau} - m_t\right)'$, a $(J+1 \times J+1)$ covariance matrix of errors of predictor $m_{t/\tau}$ of the vector m_t,

$r = \left[r_0, r_1, ..., r_J\right]'$, a column vector of lag coefficients,

$R = diag(r)$, a diagonal $(J+1 \times J+1)$ matrix with elements of the vector r on the main diagonal.

Now the autoregressive process generating the risk parameter μ_t could be cast in the form of the transition equation:

$$m_t = \begin{bmatrix} a_1 & a_2 & \dots & a_J & a_{J+1} \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & & 1 & 0 \end{bmatrix} m_{t-1} + \begin{bmatrix} 1-\left(a_1 + \dots + a_{J+1}\right) \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \mu + \begin{bmatrix} \gamma_t \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix},$$

which can be rewritten in the compact form as:

$$m_t = Am_{t-1} + b\mu + \Gamma_t.$$

The *observation equation* in turn can be written also in the matrix form:

$$X_t = Rm_t + E_t.$$

Let us notice that in practice the maximal lag J is usually fairly greater than the order n sufficient to express the dynamics of the underlying process generating the risk parameter μ_t. Then the lasting elements of the first row of the matrix A equal zero. However, if it is not the case, the dimension of the vector r can be enlarged enough to satisfy the requirement $J+1 \geq n$ by adding a respective number of zero elements.

The assumptions taken in section 2 can be now expressed in the matrix form:

- both Γ_t and E_t are normally distributed and have zero expectations,
- for each t and τ $\mathrm{E}(E_t\Gamma_\tau') = 0$, and for $t \neq \tau$ also $\mathrm{E}(E_tE_\tau') = \mathrm{E}(\Gamma_t\Gamma_\tau') = 0$,
- $\mathrm{E}(E_tE_t') = \sigma^2 R$,
- $\mathrm{E}(\Gamma_t\Gamma_t') = \Phi$, a $(J+1 \times J+1)$ matrix containing almost only zeroes except the upper-left element that equals ϕ^2.

Best unbiased predictors and their covariance matrices in the state-space model are defined recursively. Hence there is a need to supplement the above assumptions by starting conditions. For simplicity we assume that the claim process has satisfied the above assumptions at least since claims that occurred in the period $(-J)$, and that earliest observables are elements of the vector X_1. Thus the vector m_1 and m_0 are well defined and satisfy the transition equation, as well as the vector X_1 satisfies the observation equation. Our knowledge about the underlying risk parameter m_1 prior to any statistical observation is represented by a guess $m_{1/0}$ and some non-singular covariance matrix $V_{1/0}$.

Now one-year-ahead BUP of the vector m_t is defined by the prediction equation:

$$m_{t/t-1} = Am_{t-1/t-1} + b\mu, \tag{1}$$

whereas this year BUP of this vector by the updating (or filtering) equation:

$$m_{t/t} = V_{t/t}\left(\sigma^{-2}X_t + V_{t/t-1}^{-1}m_{t/t-1}\right). \tag{2}$$

Covariance matrices of one-year ahead prediction errors and filtering errors are given by:

$$V_{t/t-1} = AV_{t-1/t-1}A' + \Phi, \tag{3}$$

$$V_{t/t} = \left(\sigma^{-2}R + V_{t/t-1}^{-1}\right)^{-1}. \tag{4}$$

Additionally, error $F_t = X_t - X_{t/t-1}$ of one-year-ahead predictor $X_{t/t-1}$ of X_t is also considered. The point predictor itself simply equals $X_{t/t-1} = Rm_{t/t-1}$. The error can be decomposed into two uncorrelated components $F_t = E_t - R(m_{t/t-1} - m_t)$, so the covariance matrix equals:

$$\mathrm{Cov}(F_t) = \sigma^2 R + RV_{t/t-1}R. \tag{5}$$

Especially relevant result is that auto-covariances of vectors F_t are zero:

$$\operatorname{Cov}\left(F_t, F_{t-k}'\right)=0 \text{ for } k=1, 2, \ldots. \tag{6}$$

However, one has to be aware that lack of serial correlation of prediction errors F_t cannot be extended to prediction errors $m_t - m_{t/t-1}$, as for $k = 1, 2, \ldots$ the following equality holds:

$$\begin{aligned}&\operatorname{Cov}\left(\left(m_t - m_{t/t-1}\right),\left(m_{t-k} - m_{t-k/t-k-1}\right)'\right)=\\ &=\left(AV_{t-1/t-1}V_{t-1/t-2}^{-1}\right)\left(AV_{t-2/t-2}V_{t-2/t-3}^{-1}\right)\ldots\left(AV_{t-k/t-k}V_{t-k/t-k-1}^{-1}\right)V_{t-k/t-k-1}.\end{aligned} \tag{7}$$

The above results come from Kalman and Bucy [2]. Despite the literature is easily accessible, an outline of the proof is given in the next section. A reader having no doubts concerning the results can skip section 4 and go straight to section 5.

4. Basic properties of the state-space model

The prediction equation (1) comes from replacing in the transition equation the term γ_t by zero and m_{t-1} by its predictor $m_{t-1/t-1}$, which is nothing else than replacing the RHS of the transition equation by its expectation given $\Im_{t-1}$. This replacement preserves unbiasedness and minimizes (in the class of unbiased predictors) the covariance matrix. The formula (3) for the resulting covariance matrix $V_{t/t-1}$ is obvious.

The filtering equation (2) can be justified on the basis of the linear regression model, where last observables and predictor based on past information are cast together:

$$\begin{bmatrix} X_t \\ m_{t/t-1} \end{bmatrix}=\begin{bmatrix} R \\ I \end{bmatrix} m_t+\begin{bmatrix} E_t \\ \left(m_{t/t-1}-m_t\right) \end{bmatrix},$$

where I denotes the $(J+1\times J+1)$ identity matrix. LHS of the equation represents observed dependent variables, first component of RHS represents their expectations given by the known matrix multiplied by the unknown vector m_t, treated here as a vector of regression coefficients to be estimated, and the last component represents random deviations around expectations. Their covariance matrix Ω equals:

$$\Omega=\begin{bmatrix} \sigma^2 R & 0 \\ 0 & V_{t/t-1} \end{bmatrix}.$$

Estimation of the vector of unknown regression coefficients by GLS produces its best linear unbiased predictor (BLUP) based on information $\Im_t$, with clearly separated role of older information $\Im_{t-1}$ and newest data X_t. The filtering equation (2) is just the GLS estimator $m_{t/t}$ of the vector of regression coefficients m_t, and the covariance matrix of the estimator is given by formula (4). Under linear regression model with normally distributed disturbances BLUP is of course equivalent to BUP.

It should be noted however, that the GLS estimator used here is not the most standard one. It is because it may happen that the vector r contains some zero elements. Then the matrix R is singular, as well as the whole matrix Ω. Despite that, the formula for the GLS estimator still works. However, its derivation requires the generalized Penrose-Moore inversion technique to be applied to the matrix Ω, yielding as a result:

$$\Omega^{+} = \begin{bmatrix} \sigma^{2}R^{+} & 0 \\ 0 & V_{t/t-1}^{-1} \end{bmatrix},$$

where only the left-upper block of this matrix differs from the standard inversion (see [4]). The generalized Penrose-Moore inverse R^{+} of the matrix R is a diagonal matrix with elements r_j^{-1} corresponding to non-zero coefficients r_j, and zeros for $r_j = 0$. Thus the product of matrices R and R^{+} is an idempotent diagonal matrix with ones corresponding to $r_j \neq 0$, and zeros corresponding to $r_j = 0$. Let us denote this matrix by I_R. Now the formally proper filtering equation should be written as:

$$m_{t/t} = \left(\sigma^{-2}R + V_{t/t-1}^{-1}\right)^{-1}\left(\sigma^{-2}I_R X_t + V_{t/t-1}^{-1}m_{t/t-1}\right).$$

However, each element of the vector X_t corresponding to zero of the diagonal of the matrix I_R has – according to assumptions taken – both expectation and variance equal zero, which ensures that $\Pr\left(X_t = I_R X_t\right) = 1$.

Some comment is needed also on non-singularity of covariance matrices $V_{t/t-1}$ and $V_{t/t}$, obviously presumed in formulas (2), (4) and in the argumentation based on the GLS estimation technique. In fact non-singularity of these matrices is ensured by the assumption that the starting matrix $V_{1/0}$ is nonsingular. This is a quite natural assumption, as $V_{1/0}$ represents our knowledge on m_1 prior to any statistical observation. Non-singularity of all next matrices $V_{t/t}$ and $V_{t/t-1}$ can be proven then by induction. The implication:

$$V_{t/t-1} \text{ nonsingular} \rightarrow V_{t/t} \text{ nonsingular}$$

is ensured, because non-singularity of covariance matrix $V_{t/t-1}$ means that it is a positive definite matrix, so its inversion is also positive definite, which means that for any non-zero vector $y = \left[y = y_{\bar{0}} \quad y_{\bar{1}} \quad \cdots \quad y_{\bar{J}} \right]$ the quadratic form $yV_{t/t-1}^{-1}y' > 0$ is strictly positive. That implies matrix $\sigma^{-2}R + V_{t/t-1}^{-1}$ being also positive definite, because:

$$y\left(\sigma^{-2}R + V_{t/t-1}^{-1}\right)y' = \sigma^{-2}yRy' + yV_{t/t-1}^{-1}y' \geq yV_{t/t-1}^{-1}y'.$$

So the matrix $\sigma^{-2}R + V_{t/t-1}^{-1}$ is nonsingular, and thus can be inverted to yield the matrix $V_{t/t}$ as a result, which implies non-singularity of $V_{t/t}$ itself. The proof of the implication:

$$\text{nonsingular} \rightarrow V_{t+1/t} \text{ nonsingular}$$

is more tedious. Let denote by C the difference $V_{t+1/t} - \Phi$ (for a given t). Of course $C = AV_{t/t}A'$. This ensures that C is positive semidefinite and so has a non-negative determinant. Let us denote elements of the matrix C by $c_{i,j}$, where $i, j = 0, 1, ..., J$. Let us also denote by $C_{i,j}$ (and similarly, by $V_{i,j}$) the matrix $J \times J$ obtained by deleting row i and column j out of C (and respectively out of $V_{t/t}$). Consider now the determinant of the matrix $V_{t+1/t}$, calculated by expansion of cofactors in respect of the first row:

$$\det\left(V_{t+1/t}\right) = \left(c_{0,0} + \phi^2\right)\det\left(C_{0,0}\right) + \sum_{j=1}^{J}(-1)^j c_{0,j}\det\left(C_{0,j}\right).$$

Hence it is clear that $\det\left(V_{t+1/t}\right) = \det\left(C\right) + \phi^2\det\left(C_{0,0}\right)$. But inspection of the special structure of matrix A shows that $C_{0,0} = V_{J,J}$. Matrix $V_{J,J}$ in turn, being a covariance matrix of the sub-vector of the vector $m_{t/t}$ (obtained by deleting its last element) is non-singular as well as the covariance matrix $V_{t/t}$ of the entire vector $m_{t/t}$. Thus we can conclude that:

$$\det\left(V_{t+1/t}\right) = \det\left(C\right) + \phi^2\det\left(V_{J,J}\right) \geq \phi^2\det\left(V_{J,J}\right) > 0,$$

which completes the proof of non-singularity of matrices $V_{t/t}$ and $V_{t+1/t}$ for $t = 1, 2, ...$ [1].

[1] In fact non-singularity of $V_{1/0}$ has been assumed only for convenience. Without the assumption, it still can be shown (after tedious derivation) that $V_{t/t}$ and $V_{t+1/t}$ are both non-singular at least for $t > J$.

As the errors F_t of one-year-ahead prediction of X_t are considered, the result concerning zero auto-covariances (equation 6) needs justification. In order to do that, let us first consider the case, when all $(J+1)$ elements of the vector r are non-zero. In this case the covariance matrix of the error vector F_t (as well as the vector F_{t-k} for an arbitrary positive integer k) is nonsingular. Now the auto-covariance matrix $\mathrm{Cov}(F_t, F'_{t-k})$ can be expressed as:

$$\mathrm{E}(F_t F'_{t-k}) = B\mathrm{Cov}(F_{t-k}),$$

where matrix B exists and is unique by the virtue of non-singularity of $\mathrm{Cov}(F_{t-k})$. This virtue ensures also that B is a zero matrix when and only when the auto-covariance matrix $\mathrm{E}(F_t F'_{t-k})$ equals zero. Consider now a predictor of the vector X_t of the form:

$$X_t^* = X_{t/t-1} + BF_{t-k}.$$

Covariance matrix of that predictor's errors equals:

$$\mathrm{Cov}(X_t - X_t^*) = \mathrm{Cov}(F_t - BF_{t-k}) = \mathrm{Cov}(F_t) + B\mathrm{Cov}(F_{t-k})B' -$$
$$-B\mathrm{E}(F_{t-k}F_t') - \mathrm{E}(F_t F'_{t-k})B',$$

which equals (by definition of the matrix B):

$$\mathrm{Cov}(X_t - X_t^*) = \mathrm{Cov}(F_t) - B\mathrm{Cov}(F_{t-k})B'.$$

However, the matrix $B\mathrm{Cov}(F_{t-k})B'$ is obviously positive semidefinite, which is in contradiction with optimal properties of the predictor $X_{t/t-1}$, unless $B\mathrm{Cov}(F_{t-k})B' = 0$. This implies $B = 0$, and so $\mathrm{Cov}(F_t, F_{t-k}) = 0$.

The intuitive explanation of the above result is based on the remark, that the vector F_{t-k} contains no more information than $\Im_{t-1}$, and this information has been efficiently used already in the predictor $X_{t/t-1}$, thus no improvement of the predictor $X_{t/t-1}$ is possible.

The above argumentation extends to the case when some elements of the vector r equal zero. In this case all corresponding elements of vectors X_τ, $X_{\tau/\tau-1}$ and F_τ equal zero, as well for $\tau = t$, as for $\tau = t-k$. Hence all elements of corresponding rows and columns of covariance and auto-covariance matrices equal zero. So it suffices to prove that all other elements of the auto-covariance matrix $\mathrm{Cov}(F_t, F'_{t-k})$ are also zero. It can be done by repeating the argumentation given above to vectors and matrices of reduced dimensions, obtained by the operation of deleting all elements from vectors (and all rows and columns from matrices) that correspond to zero elements of the vector r.

In order to derive the formula (7) for auto-covariance matrices of one-year-ahead prediction errors of vectors m_t it is convenient to express first the prediction error of the vector m_{t+1} as a function of the prediction error of the vector m_t. Making use of the assumed transition equation and derived prediction and filtering equations (1) and (2) we obtain:

$$m_{t+1} - m_{t+1/t} = A\left(m_t - V_{t/t}\left(\sigma^{-2}X_t + V_{t/t-1}^{-1}m_{t/t-1}\right)\right) + \Gamma_{t+1}.$$

Replacing now X_t by $E_t + Rm_t$ we obtain:

$$m_{t+1} - m_{t+1/t} = A\left(\left(I - \sigma^{-2}V_{t/t}R\right)m_t - V_{t/t}V_{t/t-1}^{-1}m_{t/t-1}\right) - \sigma^{-2}AV_{t/t}E_t + \Gamma_{t+1}.$$

However, inverting both sides of formula (4) and multiplying both sides from the left by $V_{t/t}$ we obtain $I = \sigma^{-2}V_{t/t}R + V_{t/t}V_{t/t-1}^{-1}$, that justifies the following recursive formula:

$$m_{t+1} - m_{t+1/t} = AV_{t/t}V_{t/t-1}^{-1}\left(m_t - m_{t/t-1}\right) - \sigma^{-2}AV_{t/t}E_t + \Gamma_{t+1}.$$

Noticing now, that neither m_t nor $m_{t/t-1}$ depend on $\{E_t, \Gamma_{t+1}\}$, we come to the conclusion that $\text{Cov}\left(\left(m_{t+1} - m_{t+1/t}\right), \left(m_t - m_{t/t-1}\right)'\right) = AV_{t/t}V_{t/t-1}^{-1}\text{Cov}\left(m_t - m_{t/t-1}\right)$. This leads to the formula (7) for $k = 1$. Iterating the recursion many times and noticing the general lack of dependency of $\{m_t, m_{t/t-1}\}$ on $\{E_t, \Gamma_{t+1}, E_{t+1}, \Gamma_{t+2}, E_{t+2}, \Gamma_{t+3}, \ldots\}$ we come to the conclusion that formula (7) works also for $k > 1$. Needles to say, it works as well for $k = 0$, as the RHS reduces then to $V_{t/t-1}$.

Remark 3

All the results derived so far hold without normality assumption, provided we replace the operator E by the operator BLUP denoting Best Linear Unbiased Prediction, and stop to interpret lack of correlation as independency.

5. Premium, outstanding claims reserve, and characteristics of increments of the surplus and of the final balance

Some additional notations help express the reserves conveniently. Let us denote by:

- l the $(J+1)$-element column vector of ones, and by:

- p the $(J+1)$-element column vector of cumulated lag coefficients, so as $p_j = \sum_{i=0}^{j} r_i$ for $j = 0, 1, ..., J$.

Now the outstanding claims reserve at the end of year t could be expressed as:

$$\mathrm{E}(LX_t | \Im_t) = (\iota - p)' m_{t/t} \tag{8}$$

and the expectation of the sum $(LX_t + PX_t)$ made at the beginning of the year t as:

$$\mathrm{E}(LX_t + PX_t | \Im_{t-1}) = (\iota - p + r)' m_{t/t-1}. \tag{9}$$

Conditional variance of the ultimate surplus US_t given information $\Im_t$ is the same as variance of outstanding claims LX_t around the reserve $\mathrm{E}(LX_t | \Im_t)$. Conditioning LX_t by vector of risk parameters m_t we obtain:

$$\begin{aligned}\mathrm{Var}(LX_t | \Im_t) &= \mathrm{Var}(\mathrm{E}(LX_t | m_t) | \Im_t) + \mathrm{E}(\mathrm{Var}(LX_t | m_t) | \Im_t) = \\ &= \mathrm{Var}((\iota - p)' m_t | \Im_{t-1}) + \mathrm{E}(\sigma^2 (\iota - p)' \iota).\end{aligned}$$

The last result allows to express the variance as a sum of two quadratic forms:

$$\mathrm{Var}(FB_t) = \mathrm{Var}(US_t | \Im_t) = \mathrm{Var}(LX_t | \Im_t) = (\iota - p)' V_{t/t} (\iota - p) + \sigma^2 (\iota - p)' \iota. \tag{10}$$

Using the same conditioning for the variance of US_t given $\Im_{t-1}$ we obtain:

$$\begin{aligned}\mathrm{Var}(US_t | \Im_{t-1}) &= \mathrm{Var}(\mathrm{E}(PX_t + LX_t | m_t) | \Im_{t-1}) + \mathrm{E}(\mathrm{Var}(PX_t + LX_t | m_t) | \Im_{t-1}) = \\ &= \mathrm{Var}((\iota - p + r)' m_t | \Im_{t-1}) + \mathrm{E}(\sigma^2 (\iota - p + r) \iota).\end{aligned}$$

The variance thus reads:

$$\mathrm{Var}(US_t | \Im_{t-1}) = (\iota - p + r)' V_{t/t-1} (\iota - p + r) + \sigma^2 (\iota - p + r)' \iota. \tag{11}$$

Variance given by formula (11) is just the variance of the sum of the final balance FB_t and last year increment ΔS_t. As both variables are uncorrelated (see section 2), the variance of ΔS_t is equal to the difference between variances given by (11) and (10):

$$\mathrm{Var}(\Delta S_t) = (\iota - p + r)' V_{t/t-1} (\iota - p + r) - (\iota - p)' V_{t/t} (\iota - p) + \sigma^2. \tag{12}$$

Summarizing results of section 2 and the current section, we could say that the variance of the ultimate surplus US_t given $\Im_{t-1}$ (11) can be decomposed into two parts:

- part given by formula (12) that is attributable to the technical result of year t,
- part given by formula (10), representing either the final technical result over lasting activity period (provided writing new business is stopped), or contribution to next years technical results (provided the process is continued).

Last but not least, lack of correlation between $FB_t, \Delta S_t, \Delta S_{t-1}, \Delta S_{t-2}, \ldots$ means mutual independence, as these variables are linear functions of disturbance terms γ_t and fluctuations E_t that by assumption are normally distributed.

Obtained results help express conveniently also the premium formula. According to assumptions taken in section 2, premium for year t equals $c_t =$ $= c + \mathrm{E}\left(OX_t \middle| \Im_{t-1}\right)$, that can be rewritten as:

$$c_t = c + \mu_{t/t-1}, \tag{13}$$

where $\mu_{t/t-1}$ is the first element of the vector $m_{t/t-1}$.

The obtained premium formula helps also deriving the variance of the increment ΔUS_t of the ultimate surplus process. This is because we can write:

$$\Delta US_t = c + \mathrm{E}\left(OX_t \middle| \Im_{t-1}\right) + LX_{t-1} - PX_t - LX_t = c + \mathrm{E}(OX_t \mid \Im_{t-1}) - OX_t,$$

and so $\mathrm{Var}(\Delta US_t) = \mathrm{Var}(OX_t \mid \Im_{t-1})$. Conditioning OX_t by risk parameter μ_t we obtain $\mathrm{Var}(\Delta US_t) = \mathrm{Var}\left(\mathrm{E}(OX_t \mid \mu_t) \mid \Im_{t-1}\right) + \mathrm{E}\left(\mathrm{Var}(OX_t \mid \mu_t) \mid \Im_{t-1}\right)$, that leads to the formula:

$$\mathrm{Var}(\Delta US_t) = \begin{bmatrix}1 & 0 & \ldots & 0\end{bmatrix} V_{t/t-1} \begin{bmatrix}1 & 0 & \ldots & 0\end{bmatrix}' + \sigma^2. \tag{14}$$

6. Surplus process of a mature insurer

We need one more step to obtain the classical surplus model as it was declared in the introduction. We have obtained already the surplus process with independent and normally distributed increments having the same expectation. What remains is changing variance. However, when the aim is to model the surplus process of the mature insurer, we can replace current values of covariance matrices $V_{t/t}$ and $V_{t/t-1}$ by their limits:

$$V_0 \doteq \lim_{t\to\infty} V_{t/t} \text{ and } V_1 \doteq \lim_{t\to\infty} V_{t/t-1}.$$

The limits can be found by iterating the recursion composed from equations (3) and (4):

$$V_{t/t-1} = A\left(V_{t-1/t-2}^{-1} + \sigma^{-2}R\right)^{-1} A' + \Phi, \tag{15}$$

and complementing the result V_1 by $V_0 = \left(\sigma^{-2}R + V_1^{-1}\right)^{-1}$. However, the technical question arises whether limiting values V_1 and V_0 do exist and are finite. In simplest cases an analytical solution could be derived.

Example 1

Let us assume $n = 1$ and $J = 0$, which means all involved vectors and matrices are in fact scalars. So $R = 1$, and the process of the risk parameter is of the form $\mu_t = a\mu_{t-1} + (1-a)\mu + \gamma_t$. Denoting for simplicity $V_{t/t-1}$ by v_t we can write the recursion (15) in a simple form:

$$v_{t+1} = a^2\sigma^2 v_t \big/ \left(\sigma^2 + v_t\right) + \phi^2. \tag{16}$$

It is easy to show, that for arbitrary positive σ^2 and ϕ^2, and any real number a the recursion starting at any non-negative value v_0 leads to the following stable point:

$$\lim_{t\to\infty} v_t = \frac{1}{2}\left(\phi^2 + \sigma^2\left(a^2-1\right) + \sqrt{\left(\phi^2 + \sigma^2\left(a^2-1\right)\right)^2 + 4\phi^2\sigma^2}\right),$$

so the solution exists and is unique regardless stationarity of the process μ_t.

The next example concerns a generalization that may be especially useful when the model is based on quarterly data, the risk parameter follows the auto-regressive seasonality, but all claims are reported within one year period.

Example 2

Let us assume that $\mu_t = a_{J+1}\mu_{t-J-1} + (1-a_{J+1})\mu + \gamma_t$ for some $J > 0$, and all claims are paid with delay no greater than J periods (for $J = 3$ the assumption could reflect reality in case of the model based on quarterly data). In this case the stable point of the recursion (15) is a diagonal matrix V_1 of the form:

$$V_1 = \operatorname{diag}\left[a_{J+1}^2 \frac{\sigma^2 v}{\sigma^2 + v} + \phi^2,\ \frac{\sigma^2 v}{\sigma^2 + v p_0},\ \frac{\sigma^2 v}{\sigma^2 + v p_1},\ \cdots\ \frac{\sigma^2 v}{\sigma^2 + v p_{J-1}}\right],$$

with the corresponding matrix V_0 being also diagonal:

$$V_0 = \mathrm{diag}\left[\frac{\sigma^2 v}{\sigma^2 + vp_0}, \ \frac{\sigma^2 v}{\sigma^2 + vp_1}, \ \dots \ \frac{\sigma^2 v}{\sigma^2 + vp_{J-1}}, \ \frac{\sigma^2 v}{\sigma^2 + v}\right],$$

where $v = \frac{1}{2}\left(\phi^2 + \sigma^2\left(a_{J+1}^2 - 1\right) + \sqrt{\left(\phi^2 + \sigma^2\left(a_{J+1}^2 - 1\right)\right)^2 + 4\phi^2\sigma^2}\right)$.

Formally, the solution again does not require stationarity of the underlying process μ_t. In the special sub-case when $a_{J+1} = 0$ (so that the risk parameter μ_t just varies independently around the mean μ), the last formula simplifies to $v = \phi^2$, and subsequent diagonal elements of the matrix V_0 coincide with well known formulas for variances of outstanding claims estimates made for each cohort of claims as the weighted average of the Chain-Ladder and Bornhuetter-Fergusson estimates, with weights based on credibility theory (and optimal in this sense).

Remark 4

The suspicion that the existence of the solution depends on stationarity of the underlying process μ_t might ensue from the simple remark that predictions are functions of this process (disturbed by E_t). However, the suspicion that covariances may expand unlimitedly could be dispelled, as the covariance matrix V_t could be bounded from above. In order to show that, let us remind that $V_{t+1/t}$ is a covariance matrix of BUP of m_{t+1} given information set $\Im_t$. Thus the matrix has to be smaller than covariance matrix of any other unbiased predictor of m_{t+1} based on a subset of information set $\Im_t$. An inferior predictor with covariance matrix easy to be assessed analytically can be obtained by iterating the transition equation J times:

$$m_{t+1} = A^{J+1} m_{t-J} + \sum_{j=0}^{J} A^j \left(b\mu + \Gamma_{t-j+1}\right).$$

Omission of unknown random terms and replacement of each element of m_{t-J} by claims occurred in the respective period leads to the following predictor:

$$m_{t+1}^{*} \doteq A^{J+1}\begin{bmatrix} OX_{t-J} \\ \vdots \\ OX_{t-2J} \end{bmatrix} + \sum_{j=0}^{J} A^j b\mu$$

that is based on information on those cohorts of claims, which have been fully compensated before the end of year t. The prediction error of the above predictor could be decomposed into two components:

$$m_{t+1} - m_{t+1}^{*} \doteq A^{J+1} \begin{bmatrix} \mu_{t-j} - OX_{t-J} \\ \vdots \\ \mu_{t-2J} - OX_{t-2J} \end{bmatrix} + \sum_{j=0}^{J} A^{j} \Gamma_{t-j+1}$$

both having zero expectations and being mutually independent (the first one depends only on vectors $E_t, E_{t-1}, ..., E_{t-2J}$, the second one depends only on $\gamma_{t+1}, \gamma_t, ..., \gamma_{t-J+1}$). Thus the covariance matrix of prediction errors equals:

$$\mathrm{Cov}\left(m_{t+1} - m_{t+1}^{*}\right) = \sigma^2 A^{J+1}\left(A^{J+1}\right)' + \sum_{j=0}^{J} A^{j} \Phi \left(A^{j}\right)'.$$

For arbitrary finite number J the above covariance matrix is finite. The predictor m_{t+1}^{*} ignores all information contained in $X_{t-j,k}$ for $j = 0, 1, ..., J-1$ and $k \leq j$, that belong to the information set $\Im_t$, so cannot be superior to $m_{t+1/t}$. Usually it is strictly inferior. Both predictors are identical only in the case, when $r_0 = ... = r_{J-1} = 0$ and $r_J = 1$, and at the same time $a_1 = ... = a_{J+1} = 0$. So we can conclude that:

$$V_1 \leq \sigma^2 A^{J+1}\left(A^{J+1}\right)' + \sum_{j=0}^{J} A^{j} \Phi \left(A^{j}\right)',$$

where the symbol "$\leq$" means that RHS less LHS equals a positive semidefinite matrix.

Remark 5

A researcher focused on realism of assumptions should take into account that one of them was that the conditional variance σ^2 is fixed. The more realistic assumption is that the conditional variance is proportional to the conditional expectation μ_t. Thus the more dispersion in the possible paths of the process μ_t is allowed, the more doubtful is the fixed variance assumption. Hence the problem of finite variances of the surplus process under non-stationary risk parameter process should be read as interesting technical issue, but not so much relevant for practical applications. However, the random walk case $a = 1$ is still of some practical interest. It is because the true parameter is rarely known, and empirical tests of hypothesis that $a = 1$ against the alternative $a < 1$, are in fact weak. In lights of prudence required in actuarial practice, it is safer to assume $a = 1$ in case when the test is not really conclusive.

Now we can simplify formulas for variances concerning the "mature insurer". So the larger t, the better is the following approximation of the variance of the outstanding claims reserve (final balance):

$$\text{Var}(US_t|\Im_t) = \text{Var}(LX_t|\Im_t) \approx (l-p)'V_0(l-p) + \sigma^2(l-p)'l. \quad (17)$$

It is worthwhile to notice, that the second of the two components appearing on the RHS is just the product of the conditional variance times the average delay. This component is usually recognized and taken into account. The first component is due to prediction errors of the risk parameter, and its recognition is less common.

Before presenting next formulas it is worthwhile to explore first the special structure of matrix A. Let us denote the first row of this matrix (containing autoregression parameters $a_1, a_2, ..., a_{J+1}$) by a. It is easy to verify through element by element calculation, that the following vectors are equal:

$$(l-p+r)'A = (l-p)' + a.$$

Taking into account the recursion (3) we can conclude also that the upper-left element of the matrix $V_{t/t-1}$ equals $aV_{t-1/t-1}a' + \phi^2$.

Now the formula (11) for the variance of the sum of the final balance and last year increment and formulas (12) and (14) for variances of increments of the surplus and the ultimate surplus processes simplify to:

$$\text{Var}(US_t|\Im_{t-1}) \approx (l-p+a')'V_0(l-p+a') + \phi^2 + \sigma^2(l-p+r)'l, \quad (18)$$

$$\text{Var}(\Delta S_t) \approx aV_0a' + 2aV_0(l-p) + \phi^2 + \sigma^2, \quad (19)$$

$$\text{Var}(\Delta US_t) \approx aV_0a' + \phi^2 + \sigma^2. \quad (20)$$

Comparing last two formulas we find that the difference between variances of increments of the surplus and the ultimate surplus processes equals $2aV_0(l-p)$. Moreover, the following implication holds:

$$aV_0(l-p) \neq 0 \Rightarrow \underset{k\in\{1,2,3,...\}}{\exists} \text{Cov}(\Delta US_{t+k}, \Delta US_t) \neq 0. \quad (21)$$

This means that the difference between variances of increments of the process S_t and the process US_t imply existence of serial correlation in the latter.

We start the proof of implication (21) by comparing two different decompositions of the variance of the difference $\text{Var}(US_{t+k} - S_t)$. The first decomposition is based on lack of correlation of increments of the surplus process and final balance:

$$\text{Var}(US_{t+k} - S_t) = \text{Var}(FB_{t+k}) + \sum_{j=1}^{k} \text{Var}(\Delta S_{t+k}).$$

This means that for t large enough the approximate equality holds:

$$\mathrm{Var}(US_{t+k+1}-S_t)-\mathrm{Var}(US_{t+k}-S_t)\approx\mathrm{Var}(\Delta S_{t+k+1}),$$

where the approximation is due to neglecting the difference $\mathrm{Var}(FB_{t+k+1})-$ $-\mathrm{Var}(FB_{t+k})$.

The second decomposition:

$$\mathrm{Var}(US_{t+k+1}-S_t)=\mathrm{Var}(US_{t+k}-S_t)+\mathrm{Var}(\Delta US_{t+k+1})+\\ +2\mathrm{Cov}(\Delta US_{t+k+1},US_{t+k}-S_t),$$

confronted with the first one allows to obtain the result:

$$\mathrm{Cov}(\Delta US_{t+k+1},US_{t+k}-S_t)\approx\frac{1}{2}\left(\mathrm{Var}(\Delta S_{t+k+1})-\mathrm{Var}(\Delta US_{t+k+1})\right).$$

Making use again of the assumption that t is large we can replace the RHS by the constant d. Decomposing the covariance appearing on the LHS we can conclude that for arbitrary $k=0,1,2,\ldots$:

$$\mathrm{Cov}(\Delta US_{t+k+1},FB_t)+\sum_{j=1}^{k}\mathrm{Cov}\left(\Delta US_{t+k+1},\Delta US_{t+j}\right)\approx d.$$

Let us take now the working assumption (contradictory to implication 21) that at the same time $d\neq 0$ and all serial correlations of ΔUS_t are zero. But if this is so, then $\mathrm{Cov}(\Delta US_{t+k},FB_t)\approx d$, and $\mathrm{Var}(US_{t+k}-US_t)\approx k\mathrm{Var}(\Delta US_t)$, and thus we obtain:

$$\mathrm{Var}\left(US_{t+k}-US_t-k\frac{d}{\mathrm{Var}(FB_t)}FB_t\right)\approx k\mathrm{Var}(\Delta US_t)-k^2\frac{d^2}{\mathrm{Var}(FB_t)}.$$

But for any $d\neq 0$ we could find k large enough to assure that the RHS is negative. As variance cannot be negative we have come to the contradiction that dispels the working assumption, so the implication (21) has to be true.

Implication (21) raises the question about cases when variances of ΔUS_t and ΔS_t are equal, and whether serial correlation of ΔUS_t is present then. Both questions are answered by the two following examples.

Example 3

Let us assume $r'=\begin{bmatrix}1 & 0 & \ldots & 0\end{bmatrix}$, which means that there are no delays in reporting claims. So $LX_t\equiv 0$ and $US_t\equiv S_t$, thus no serial correlation exist.

Less trivial is the next example.

Example 4

Let us assume that the vector a equals zero, so that the risk parameter just varies around the mean $\mu_t = \mu + \gamma_t$. Then $aV_0(l-p) = 0$, so variance of ΔUS_t equals that of ΔS_t. In this case mutual independency of deviations γ_t suffice to conclude that $\mathrm{E}(OX_t|\Im_{t-1}) = \mu$, and so $\mathrm{Cov}(\Delta US_{t+k}, \Delta US_t) = \mathrm{Cov}(OX_{t+k}, OX_t) = 0$ for all $k = 1, 2, \ldots$.

Conjecture

Most probably examples 3 and 4 illustrate all possible cases when increments of the ultimate surplus process ΔUS_t shows no serial correlation. So the conjecture reads that the model produces trivial results only in two cases:

- when there are no delays in claims settlement, or:
- when there is no serial correlation in the process driving the risk parameter μ_t.

Generally, under simultaneous appearance of both delay and serial correlations the processes ΔUS_t and ΔS_t are essentially different.

7. Conclusions and their validity under generalizations of the model

Direct conclusions concern explicit results derived in the paper:

- Properties of the surplus process S_t defined as $\mathrm{E}(US_t|\Im_t)$ are (under assumptions of the model) analogous to those of the classical surplus process with i.i.d increments. However, this is true as long as we assume that the process is continued. Once writing new business is stopped at time t, the question arises whether the lasting amount of assets $S_t + \mathrm{E}(LX_t|\Im_t)$ will suffice to cover outstanding claims LX_t.
- Variances of both the final balance and the increment of the surplus, are derived in the paper (for the mature insurer there are given by formulas (17) and (19)).
- The variance of the increment of the process can be a multiple of the conditional variance of the aggregate amount of claims occurred in a year given the risk parameter μ_t. This is just this conditional variance that is usually estimated on the basis of exploration of the cross-section dataset on claims occurred over a one- or two-year period in all lines of business provided by the company.
- The formula for the final balance could be decomposed into the product of the conditional variance and the average delay (that is usually taken into account) and the other part due to prediction errors of the risk parameter (the explicit recognition is less common in that case).

- The Insurance Supervision Authority has to set the lower bound S_{LOW} that triggers the stopping decision, provided the current surplus is lower. The proper level of S_{LOW} has to be set on the basis of the assumption, that otherwise (if the process is not stopped) the sum of next increment of the process and the final balance should cover liabilities with a sufficiently high probability.
- The event of the surplus S_t falling below the imposed lower bound S_{LOW} could be treated by the insurer as the event of ruin, as it usually means serious troubles to managers and shareholders.

Indirect conclusions concern possible interventions into the process and related strategies:

- Having the surplus S_{t-1} known at the end of the year $t-1$ we can consider possible interventions in respect of components of the increment over the coming year: $\Delta S_t = c + \mathrm{E}(OX_t|\Im_{t-1}) - PX_t + \mathrm{E}(LX_{t-1}|\Im_{t-1}) - \mathrm{E}(LX_t|\Im_t)$.
- Switching on and off dividend payments poses no problem. However, usually reinsurance contracts are written in respect of new business, so that the difference between the amount of outstanding claims LX_{t-1} and the corresponding reserve $\mathrm{E}(LX_{t-1}|\Im_{t-1})$ is left uncovered. Of course, only a part of this risk will be revealed during the current year, and the other part will be passed to next years as a component of the difference $LX_t - \mathrm{E}(LX_t|\Im_t)$.
- Reinsurance contracts concerning the run-off of claims already occurred are known, and sometimes happen in practice. However, the insurer whose surplus has fallen down recently to a dangerously low level will have serious problems to obtain (or extend) such reinsurance cover at a reasonable price.

8. Advantages and limitations of the model

The model is intended to give answer to the question, how the ever-changing economic environment together with delay in information influences the ruin probability and other solvency criteria.

- That is why assumptions about changes in time of the risk parameter μ_t are quite general. They can reflect as well quite stable, as fairly volatile economies or lines of business.
- Also assumptions on distribution of the delay in claims settlement are quite general, enable to cover as well portfolios with majority of "short tailed" as those with majority of "long tailed" lines of business. They reflect more or less the practice of reserve calculation, and so the mechanisms of creation of profit and loss account figures.

- The significant simplification concerns in this respect confounding the process of reporting and settlement of claims. It is done to clarify the discrimination between "already known" and "still unknown" figures, and to avoid considering the third category of "partly known" information. A realistic model should include the separated process of gradual improvement of prediction of amounts of individual claims between the date of first report and the date of final settlement. The simplification does not distort basic qualitative results, rather it means that the model cannot be directly applied to solve practical problems (as setting solvency requirements for instance).

A comment is needed also on the discrete-time character of the model. On the one hand, it simplifies the reality evidently. On the other hand, it reflects in a simple manner some fairly complex phenomena:

- It reflects in a simplified manner the fact, that premium is usually written for one-year risk exposure period, during which rates cannot be updated even if the newest empirical evidence shows that their level is no more adequate.
- The above-mentioned one-year cycle justifies (at least to some extent) neglecting the difference between the premium written and premium earned.
- One-year cycle coincides also with the frequency of publishing financial statements verified by independent auditor, which has substantial impact on decisions made by various agents. In our context the most important are external bodies entitled to suppress the company to write new business, and internal bodies entitled to set premium rates (provided writing business is continued). Also decisions on capital (subdivision of profits into dividends and retained part for instance) are made cyclically.

Some comment should be given also on assumption on normality of conditional (given μ_t) distribution of claim figures $X_{t,j}$. This assumption contradicts the tradition of treating the ruin theory as a tool for analysis of the danger of large individual claims (or catastrophes causing enormous number of claims). In this paper it is assumed to the contrary that the model describes the situation of an insurer prudent enough to be always secured by well suited XL and CAT covers. However, even such prudent insurer is still sensitive to the risk of changes of basic parameters, enlarged additionally because the information concerning these changes comes delayed.

Another simplification is embodied in the assumption that the risk parameter changes are described realistically enough by the linear time series model with normally distributed innovations. In fact, linearity is just a simplification. On the other hand, normality (as well as normality of conditional distribution of claim amounts) is not crucial for conclusions drawn from the model. In fact it is needed when we insist on obtaining the process with independent increments. This assumption is required in standard ruin theory models, where the long term behavior of the surplus process stands for the basis of business decisions. However, once we restrict our interest to decisions undertaken on the basis of a short horizon, the independency of increments is no more required. Assuming then that both premium setting and calculation of reserves is based on Best Linear Unbiased Predictors, we come to the same results concerning point predictors and covariances of prediction errors.

Literature

[1] De Jong P., Zehnwirth B., *Claims Reserving, State-Space Models and the Kalman Filter*, „Journal of the Institute of Actuaries" 1983, vol. 110, 157-181.

[2] Kalman R.E., Bucy R.S., *New Results in Linear Filtering and Prediction Theory*, „ASME Journal of Basic Engineering" 1961, 83, 95-108.

[3] Scheike T.H., *A General Risk Process and its Properties*, „Journal of Applied Probability" 1992, 29, 73-81.

[4] Theil H., *Principles of Econometrics*, Wiley and Sons, New York 1971.

MODEL NADWYŻKI UBEZPIECZYCIELA ZE ZMIENNYM W CZASIE PAREMETREM RYZYKA I OPÓŹNIONYMI REALIZACJAMI

Streszczenie

Typowy model nadwyżki ubezpieczyciela S_t z czasem dyskretnym zakłada, że:

$$S_t - S_{t-1} + W_t, \quad t = 1, 2, ...,$$

gdzie $W_1, W_2, ...$ to zmienne losowe i.i.d. reprezentujące saldo składki i wartości szkód za okres roku. Zakłada się też, że nadwyżka początkowa S_0 jest ustalona, zaś rozkład W_t jest znany. Model służy uzyskaniu odpowiedzi na pytania o ruinę – jej prawdopodobieństwo, czas zajścia, deficyt w momencie ruiny itd.

W rzeczywistości jednak procesy zgłaszania i likwidacji szkód zachodzących w ciągu roku rozciągają się z reguły na lata następne. Związek modelu z rzeczywistością można przywrócić, przyjmując, że zmienna W_t odpowiada pojęciowo temu, co w języku sprawozdań finansowych określamy jako wynik techniczny, a więc różnicy pomiędzy wartością składki zarobionej a wartością szkód wypłaconych powiększoną o przyrost zobowiązań (rezerw) z tytułu szkód zaszłych i niewypłaconych. Nadwyżkę S_t możemy wtedy interpretować jako stan środków własnych, zaś ruinę jako utratę wypłacalności. Taka interpretacja prowadzi jednak do komplikacji, ponieważ wartość szkód zaszłych, ale niewypłaconych jest zmienną losową, a odpowiednia rezerwa to w istocie predyktor punktowy tej zmiennej oparty na informacji dostępnej w dniu bilansowym.

W rzeczywistości proces predykcji dodatkowo utrudnia fakt, iż parametry rozkładu łącznej wartości szkód zachodzących w ciągu roku nie są stałe w czasie, toteż kalkulacja tak rezerw, jak i składki wymaga predykcji tych parametrów.

W artykule zaprezentowany jest model z wbudowanymi komplikacjami obu rodzajów. Kosztem pewnych upraszczających założeń obie komplikacje wprowadzić można w taki sposób, aby zachować niezależność i taki sam rozkład zmiennych $W_1, W_2, \ldots$ i otrzymać realistyczne oceny jego parametrów.

Przy okazji okazało się, że:

- uwzględnienie obu komplikacji jest niezbędne do poprawnej oceny wariancji rocznego przyrostu nadwyżki oraz wariancji błędu predykcji zobowiązań z tytułu szkód zaszłych i niewypłaconych,
- model pozwala uzyskać bardziej realistyczny obraz gry toczonej pomiędzy nadzorem, określającym reguły stopowania działalności a ubezpieczycielem, pragnącym prowadzić działalność niezakłóconą interwencjami nadzoru,
- model pozwala bardziej realistycznie zakreślić granice sterowalności procesem nadwyżki poprzez takie standardowe techniki, jak stopowanie/uruchamianie wypłat dywidend czy zwiększanie/zmniejszanie stopnia, w jakim ryzyko jest reasekurowane.

Zastosowane w pracy techniki można określić językiem matematyki jako wykorzystanie twierdzenia Dooba-Meyera o dekompozycji submartyngałów, zaś językiem ekonomii jako wykorzystanie teorii racjonalnych oczekiwań. Przyjęto na tyle prostą wersję założeń, aby móc oprzeć predykcję (kalkulację rezerw i kalkulację składki) na filtrze Kalmana i aby w rezultacie otrzymać analityczne wyniki dotyczące wariancji błędów predykcji. Praca prezentuje więc oryginalnie postawione pytanie, jednak techniki użyte w celu uzyskania odpowiedzi na nie są od dawna znane (np. kalkulacja rezerw na podstawie filtru Kalmana została zaproponowana w latach osiemdziesiątych).