Automatic Inclusion of Diagonal Effects in Reserve Models

Gary Venter

Abstract Including diagonal effects in reserve models runs into problems of overlap with other parameters and over-parameterization. We address these by using parameter shrinkage in MCMC to control over-parameterization, modeling loss ratios to remove known level differences, and parameter constraints, especially forcing no ongoing trend across accident years, to control for overlap. We also look at modeling changing settlement rates along the lines proposed by Glenn Meyers.

1 Basic Model

Start with data in a rectangle y with rows indexed by w = 1, ..., 10 and columns d = 0, ..., 10–1. Here the symbol ″10″ can be considered a number in any base, but since we are using Schedule P data, we will take it to be in the usual base. The columns (lags) start at zero so the $y_{w,d}$ cell will be on diagonal (calendar year) $w+u$. The basic model has parameters p_w, q_d, r_{w+d} and a constant c so that

$$m_{w,d} = Ey_{w,d} = cp_wq_dr_{w+d} \tag{1}$$

For the sake of identifiability, $p_1 = r_1 = 1 = \Sigma_d q_u$.

We use this model for incremental paid loss ratios. Incremental losses have some chance of being independent, but cumulative cannot be. Also at any point the prediction problem is the next incremental loss. The next cumulative might be already largely known. Using loss ratios helps with the constraints.

2 Constraints

Sometimes very good fits can be found by letting the accident years and calendar years have sharp but largely offsetting trends. This creates problems for projections, however. To prevent this, we assume there is no overall trend across the accident years. They do not have to start and end at the same point, but a regression line through them has to have zero slope.

Gary Venter
University of New South Wales, e-mail: gary.venter@gmail.com

This is accomplished by first estimating the p parameters as discussed below, then fitting a regression line through them and making the revised parameters the residuals around the line. For a regression of $y_1, ..., y_n$ on the integers 1,...,n, the slope and intercept are:

$$slope = \frac{6\sum_j (2j-n-1)y_j}{n^3-n} \tag{2}$$

$$intercept = average(y) - (slope)(n+1)/2 \tag{3}$$

Initially $p_1 = 1$, but its residual will not be 1. We increase each residual by 1 – the lowest, to make the minimum of the resulting p parameters = 1 after the constraints.

All the p, q and r parameters are forced to be positive.

3 Estimation

To keep from having too many parameters, we put the p, q, and r parameters on line segments (linear splines). The estimation problem then becomes one of estimating the slope changes between line segments – where slope changes should go and how large they should be. We use a Bayesian approach for this, with shrinkage priors, starting with parameters for every slope change (10 of these for q and 9 each for p and r) but give them mean-zero priors that push them to be small unless justified by the data.

The slope at any point is the previous slope plus the slope change at the point. Thus the slopes are the cumulative sum of the slope changes. Similarly the final p, q, and r level parameters are the previous parameter plus the current slope, and so are cumulative summations of the slopes. Thus a double sum of the slope changes gives the level parameters.

We use the popular double exponential shrinkage prior here. It is the exponential for positive values, and its mirror image for negative values, with density:

$$x > 0 : f(x) = e^{-x/b}/2b \tag{4}$$

$$x < 0 : f(x) = e^{x/b}/2b \tag{5}$$

The constant was given a positive prior proportional to $1/x$. With the variance assumptions discussed below, the posterior mean parameters given by the Stan MCMC application then become the estimates.

4 Variance

There is usually an issue with heteroscedasticity in reserve models – the variance can be different in different cells. One way to deal with this is to postulate that cell i that has mean = m_i has variance sm_i^o, with parameters s and o that are constants across the dataset.

A typical loss triangle has aggregate losses in the cells, so is a combination of number of claims and claim size. In the case of a paid loss triangle that would be the number of claims with payments in the cell and the size of the payments. It is pretty common for larger claims to pay later. So if there are not too many partial payments, the later cells in the triangle would have much lower frequency but higher severity than the earlier cells.

The variance of aggregate losses is generally proportional to the mean of frequency and the square of the mean of severity. Thus the later cells would have both aggregate mean and variance going down by the frequency mean, and mean going up by the increase in severity, and variance going up by the square of the mean of severity. Thus the variance would not be going down as fast as the mean would be. This would lead to the parameter o being less than 1, but we are not requiring that. While s is required to be positive, o can be any real number.

For a normal distribution, this can be implemented by setting $\mu_i = m_i$ and $\sigma_i^2 = sm_i^o$. For other distributions it gets a little more complicated. For instance consider a gamma distribution parameterized to have mean = $\alpha\beta$ and variance = $\alpha\beta^2$. Setting $\alpha = m_i^{2-o}/s$ and $\beta = sm_i^{o-1}$ gives mean m_i and variance sm_i^o.

With mean and variance specified, the choice of distribution would be based on other statistical properties. Often the smallest observations in a triangle are quite volatile, leading to a skewed distribution overall. If $o < 1$, the CV of the smaller elements is relatively large, and its skewness would be as well. A fairly skewed distribution that has been found to work satisfactorily in some cases is the lognormal. We adopt that here.

To get the μ and σ parameters of the lognormal that match the mean and variance from our model, note that the lognormal mean for cell i is $m_i = e^{\mu_i + \frac{1}{2}\sigma_i^2}$. Thus set $\mu_i = log(m_i) - \frac{1}{2}\sigma_i^2$. The variance is $m_i^2\left(e^{\sigma_i^2} - 1\right) = sm_i^o$. From this it follows that $\sigma_i^2 = log(1 + sm_i^{o-2})$.

5 Application Details

The model was fit to triangles of incremental net loss ratios in whole percents – that is, a 60% loss ratio is 60. Some experimentation found a double exponential b parameter of 0.05 reasonable. Fits were not particularly sensitive to this choice. A lower value leads to smaller changes in the level parameters over the 10 years. However this parameter was scaled up for the AY parameters, since making those residuals around the trend provides a degree of constraint anyway. Also b was scaled down a bit for the lag parameters. These do not usually change quickly across the triangle anyway, and allowing more change sometimes results in very strange, poor fitting level parameters. That suggests a specification issue, but it can be handled in an ad hoc way for now by omitting the very occasional chain where it occurs.

6 Changing Settlement Rate

Glenn Meyers introduces a model of changing settlement rates that provides for the payout pattern changing over time. He is modeling cumulative payouts, not incremental, and expresses the model for the log of the payments. To convert the notation, let x_d be the cumulative sum across the lags of q_d, so $x_9 = 1$. Meyers then adds a payment shift factor power z near 1, and replaces x_d by $x_d^{z^{w-1}}$ in the mean for $y_{w,d}$.

If $z < 1$, the new lag factor increases towards 1, so more of the eventual total loss is paid earlier, and increasingly so for the most recent accident years. You compensate by increasing the row parameters for the later years to match the data, but less is projected for future payments.

It would be possible to implement this in the incremental model, by differencing the revised lag factors. However an alternative is to model shifts in the incremental lags. It is difficult to discern in the data changes in the payout pattern beyond $d = 2$, so we will model shifts among the first 3 lags, $d = 0, 1, 2$.

Postulate then a payment shift factor z near 1, and for $m_{w,0}$ replace q_0 by $z^w q_0$. This increases or decreases the payment at lag 0 increasingly for later accident years. We will keep the sum of the qs at 1.0 before the application of trend, so will allocate $(1 - z^w)q_0$ between lags 1 and 2 by the use of a single parameter g, with $g(1 - z^w)q_0$ added to q_1 and $(1 - g)(1 - z^w)q_0$ added to q_2 for $m_{w,1}$ and $m_{w,2}$. This is done also for rows 9 and 10 for projections.

Some experimentation finds that a double exponential prior with $b = 0.01$ improves the fit for triangles that appear to have payout shifts, without hurting it too much for other triangles. A small value of z tends to show up for the latter triangles, and seems to improve the fit slightly, but hurt the loo measure, since more parameters are used.

For g we tried a uniform prior on [–2,2]. This allows lag 1 to have more of a change than lag 0.