RESERVE VARIABILITY CALCULATIONS

Moderator: Daniel Murphy, President, Trinostics

Panelists:

Markus Gesmann, Manager, Lloyd's Analysis, Lloyd's of London

Wayne Yanwei Zhang, Data Modeling Analyst, Statistical Research, CNA

Jimmy Curcio Jr., Senior Actuarial Analyst, Specialty Markets Actuarial, Munich Reinsurance America

San Diego May 2010

LDF CURVE-FITTING AND STOCHASTIC RESERVING

Jimmy Curcio

Munich RE

Agenda

Munich RE

1. The Method: A Brief Overview
2. Advantages/Disadvantages of the Model
3. Example (Excel File)
4. Appendix

Agenda

1. The Method: A Brief Overview
2. Advantages/Disadvantages of the Model
3. Example (Excel File)
4. Appendix

The Method: A Brief Overview

Munich RE

The **goals** of the model are as follows:

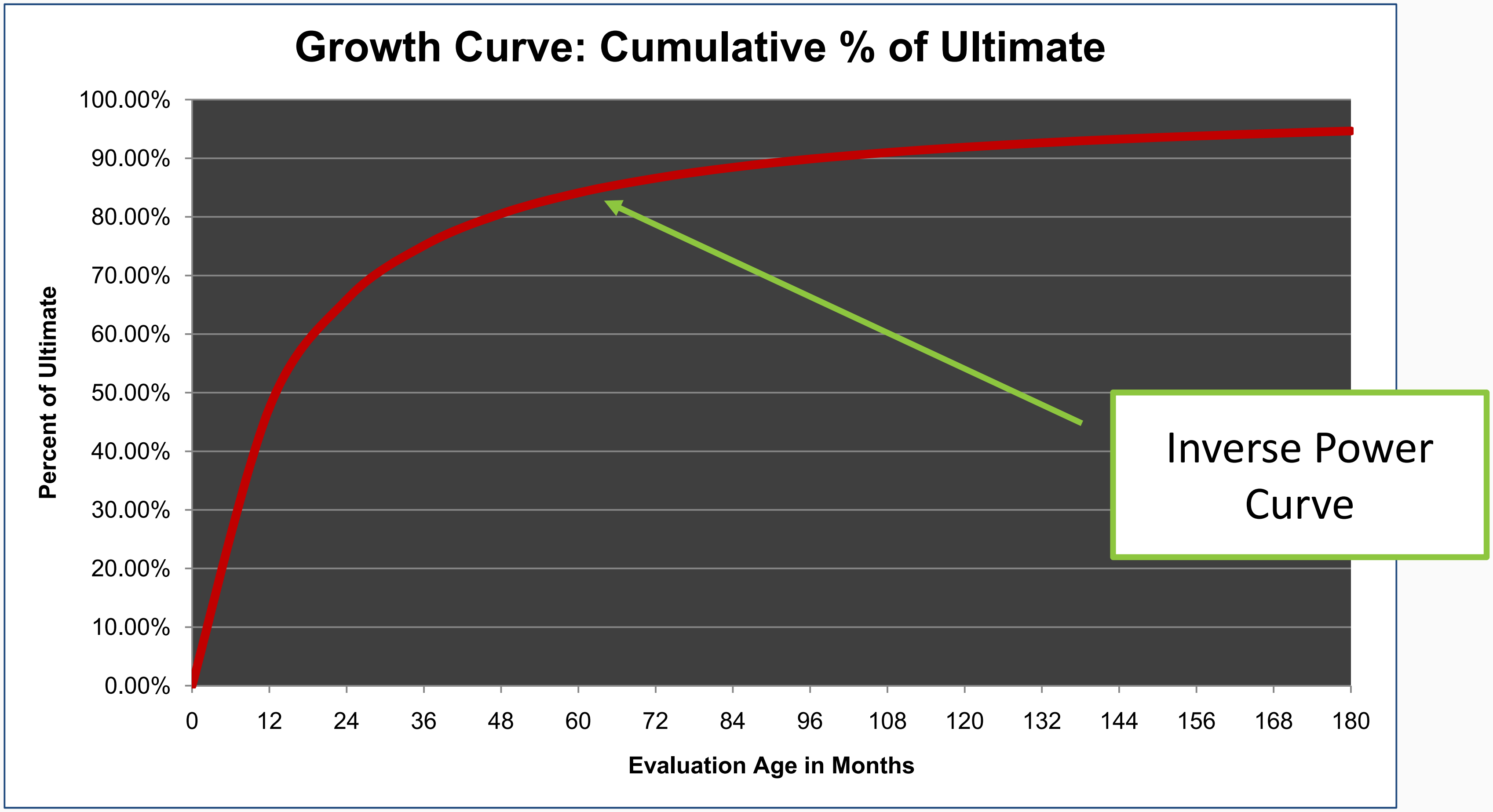
1. Describe loss emergence in a mathematical model to assist in estimating needed reserves

2. Calculate the variability around the estimated reserves

3. Estimate tail factor

The Method: A Brief Overview

The model utilizes the **Inverse Power Curve**.



The Method: A Brief Overview

The **Model Game Plan** is as follows:

1. Convert loss development triangle to an incremental basis

The Method: A Brief Overview

Munich RE

The **Model Game Plan** is as follows:

2. For each "cell" of the triangle, we have

$c_{i,t}$ = actual loss for AY *i*, between ages *t* and *t-1*

$\mu_{i,t}$ = expected loss for AY *i*, between ages *t* and *t-1*

The Method: A Brief Overview

Munich RE

The **Model Game Plan** is as follows:

2. For each "cell" of the triangle, we have

Actual				
AY	12	24	36	48
2006	$c_{2006,12}$	$c_{2006,24}$	$c_{2006,36}$	$c_{2006,48}$
2007	$c_{2007,12}$	$c_{2007,24}$	$c_{2007,36}$	
2008	$c_{2008,12}$	$c_{2008,24}$		
2009	$c_{2009,12}$			

Expected				
AY	12	24	36	48
2006	$\mu_{2006,12}$	$\mu_{2006,24}$	$\mu_{2006,36}$	$\mu_{2006,48}$
2007	$\mu_{2007,12}$	$\mu_{2007,24}$	$\mu_{2007,36}$	
2008	$\mu_{2008,12}$	$\mu_{2008,24}$		
2009	$\mu_{2009,12}$			

The Method: A Brief Overview

The **Model Game Plan** is as follows:

3. These $\mu_{i,t}$ are treated as the mean of a distribution.

The Method: A Brief Overview

Munich RE

The **Model Game Plan** is as follows:

3. These $\mu_{i,t}$ are treated as the mean of a distribution.

   How do we estimate these $\mu_{i,t}$?

   - Cape Cod Method

     Requires exposure base (more information → more accurate)

     Number of Parameters: **3**

The Method: A Brief Overview

The **Model Game Plan** is as follows:

3. These $\mu_{i,t}$ are treated as the mean of a distribution.

   How do we estimate these $\mu_{i,t}$?

   - LDF Method

     Each AY reserve is estimated independently

     Number of Parameters: **n + 2**

The Method: A Brief Overview

Munich RE

The **Model Game Plan** is as follows:

4. Parameters are estimated via **M**aximum **L**ikelihood **E**stimation (MLE)

- *The distribution for **each cell** uses an Overdispersed-Poisson*
- *The Cape Cod and LDF method are exact MLE results*

Agenda

Munich RE

1. The Method: A Brief Overview
2. Advantages/Disadvantages of the Model
3. Example (Excel File)
4. Appendix

Advantages/Disadvantages of the Model

Munich RE

The use of a continuous curve gives this model some **advantages**:

1. Smoothing of Development Pattern
2. Interpolation & Extrapolation
3. Handle irregular evaluation dates
   (e.g., latest diagonal less than 12 months from penultimate diagonal)
4. Avoid OverParameterization

Advantages/Disadvantages of the Model

The use of a continuous curve gives this model some **disadvantages**:

1. Need curve-fitting engine
   (answers not in “real time”)
2. Less precise estimate of parameter variance
3. May not fit well unless the “right” curve form is used

Agenda

1. The Method: A Brief Overview
2. Advantages/Disadvantages of the Model
3. Example (Excel File)
4. Appendix

Example (Excel File)

Let us look at an example in Excel.

LDF_Curve-Fit_and_Stoch_Res.xls

Agenda

1. The Method: A Brief Overview
2. Advantages/Disadvantages of the Model
3. Example (Excel File)
4. Appendix

Munich RE

Parameterization of the **Inverse Power Curve**

$$G(t|\theta,\omega) = \frac{1}{1+\left(\frac{\theta}{t}\right)^{\omega}}$$

Munich RE

Estimating the expected loss in each cell ($\mu_{i,t}$) for **Cape Cod**

Formulation:

$$\mu_{i,t} = Premium_i \times ELR \times [G(t|\theta,\omega) - G(t-1|\theta,\omega)]$$

Parameters:

ELR	expected loss ratio for all years
θ	"scale" parameter of *G(t)*
ω	"shape" parameter of *G(t)*

Munich RE

Estimating the expected loss in each cell ($\mu_{i,t}$) for **LDF Method.**

Formulation:

$$\mu_{i,t} = Ultimate_i \times [G(t|\theta,\omega) - G(t-1|\theta,\omega)]$$

Parameters:

$Ultimate_i$ expected ultimate loss for accident year *i*

θ “scale” parameter of *G(t)*

ω “shape” parameter of *G(t)*

Munich RE

Multivariate reserving models and the ChainLadder package

Wayne (Yanwei) Zhang

Data modeling analyst
Statistical research
CNA insurance company
Yanwei.Zhang@cna.com

Motivation

Why joint development of triangles?

- Estimation is more efficient: use more information
- Uncertainty calculation is more accurate: positive correlation
- Structural relation exists among triangles: paid and incurred triangles

Need to recognize two types of dependency:

- Contemporaneous correlation
- Structural relation

Modeling methodology

How to handle these dependencies?

- We consider proposed models in the chain ladder framework: sequential development
- The multivariate chain ladder (MCL) model in Pröhl and Schmidt (2005), Merz and Wüthrich (2008) considers how to incorporate contemporaneous correlation
- The General Multivariate Chain Ladder (GMCL) in Zhang (2010) handles both types
- We will focus on the GMCL model which includes various existing multivariate models

Notation

We introduce some notation and formulate the model.

- Assume N triangles are available and each has I rows and I columns.
- Let $\mathbf{Y}_{i,k} = (Y_{i,k}^{(1)}, Y_{i,k}^{(2)}, \cdots, Y_{i,k}^{(N)})'$ be the $N \times 1$ vector of losses at accident year i and development year k, where (n) refers to the n_{th} triangle
- Let $\mathcal{D}_{i,k}$ denote the losses up to and including development year k for accident year i
- Let D be the diagonal operator that transforms a vector to a diagonal matrix

Model structure

GMCL is dynamic, where the development from year k to $k+1$ follows

$$\mathbf{Y}_{i,k+1} = \mathbf{A}_k + \mathbf{B}_k \cdot \mathbf{Y}_{i,k} + \epsilon_{i,k}, \tag{1}$$

with

$$\mathbf{A}_k = \begin{pmatrix} \beta_{10} \\ \cdots \\ \beta_{N0} \end{pmatrix}, \mathbf{B}_k = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1N} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \cdots & \beta_{NN} \end{pmatrix}$$

Discussion

- Overparameterized model: always specify parameter restriction
- Reduce to MCL if $\mathbf{A}_k = \mathbf{0}$ and $\mathbf{B}_k$ is diagonal
- Intercepts: improve model adequacy
- Nondiagonal development matrix $\mathbf{B}_k$: allows structural connection
- Tail years with few data: use separate chain ladder to stabilize estimation

Model assumptions

The GMCL model has the following assumptions:

$$E(\epsilon_{i,k}|\mathcal{D}_{i,k}) = \mathbf{0} \tag{2}$$

$$cov(\epsilon_{i,k}|\mathcal{D}_{i,k}) = D(\mathbf{Y}_{i,k})^{\delta/2} \cdot \mathbf{\Sigma}_k \cdot D(\mathbf{Y}_{i,k})^{\delta/2} \tag{3}$$

losses from different accident years are independent (4)

$$\mathbf{\Sigma}_k = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{pmatrix}$$

Estimation of parameters

- Estimation is using seemingly unrelated regressions
- Assume $\delta = 0$ for easy presentation
- Each triangle is a regression equation as follows

$$\begin{pmatrix} Y_{1,k+1}^{(n)} \\ Y_{2,k+1}^{(n)} \\ Y_{3,k+1}^{(n)} \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & Y_{1,k}^{(1)} & \cdots & Y_{1,k}^{(N)} \\ 1 & Y_{2,k}^{(1)} & \cdots & Y_{2,k}^{(N)} \\ 1 & Y_{3,k}^{(1)} & \cdots & Y_{3,k}^{(N)} \\ \vdots & & \ddots & \vdots \end{pmatrix} \begin{pmatrix} \beta_{n0} \\ \beta_{n1} \\ \beta_{n2} \\ \vdots \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \end{pmatrix} \quad (5)$$

- Write (5) as $\mathbf{Y}^{(n)} = \mathbf{X}^{(n)} \cdot \boldsymbol{\beta}^{(n)} + \boldsymbol{\epsilon}^{(n)}$
- Error covariance matrix for this equation is $var(\boldsymbol{\epsilon}^{(n)}) = \sigma_{nn} \cdot \mathbf{I}$

Estimation of parameters

- Combine all N regression equations

$$\begin{pmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \\ \vdots \\ \mathbf{Y}^{(N)} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}^{(2)} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}^{(N)} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^{(1)} \\ \boldsymbol{\beta}^{(2)} \\ \vdots \\ \boldsymbol{\beta}^{(N)} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}^{(1)} \\ \boldsymbol{\epsilon}^{(2)} \\ \vdots \\ \boldsymbol{\epsilon}^{(N)} \end{pmatrix}$$

- Write as $\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\epsilon}$, and error terms are correlated under (3) with $var(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}_k \otimes \mathbf{I}$
- Apply generalized least squares to get $\hat{\boldsymbol{\beta}} = (\mathbf{X}'(\boldsymbol{\Sigma}_k \otimes \mathbf{I})^{-1}\mathbf{X})^{-1}\mathbf{X}'(\boldsymbol{\Sigma}_k \otimes \mathbf{I})^{-1}\mathbf{Y}$
- Replace $\boldsymbol{\Sigma}_k$ by its estimator $\hat{\boldsymbol{\Sigma}}_k$ (usually from OLS) to make estimation feasible

Estimation of ultimate loss

- The augmented approach: want to keep the multiplicative recursive estimation
- We add a constant equality to model (1) as follows:

$$\begin{pmatrix} 1 \\ \mathbf{Y}_{i,k+1} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{A}_k & \mathbf{B}_k \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \mathbf{Y}_{i,k} \end{pmatrix} + \begin{pmatrix} 0 \\ \epsilon_{i,k} \end{pmatrix} \quad (6)$$

$$\Rightarrow \mathbf{Z}_{i,k+1} = \mathbf{E}_k \cdot \mathbf{Z}_{i,k} + \mathbf{e}_{i,k} \quad (7)$$

- Now we can use the recursive estimation as

$$\hat{\mathbf{Z}}_{i,I} = \left(\prod_{k=I-1}^{I+1-i} \hat{\mathbf{E}}_k\right) \cdot \mathbf{Z}_{i,I+1-i} \quad (8)$$

- Removing the constant 1 in $\hat{\mathbf{Z}}_{i,I}$ will result in $\hat{\mathbf{Y}}_{i,I}$.

Triangles and problems

Three triangles from NAIC Schedule P are analyzed. They are composed of personal auto paid, personal auto incurred and commercial auto paid.

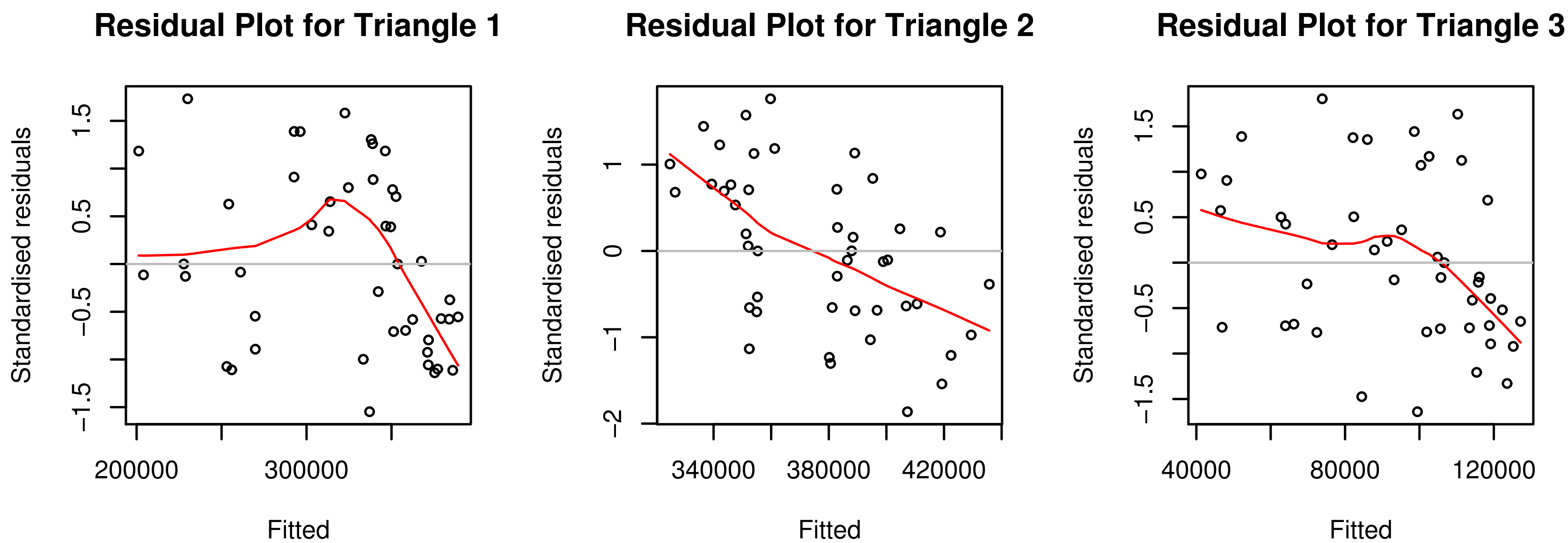
Problems with separate chain ladders:

#1 It ignores contemporaneous correlations among triangles, and the corresponding mean square error estimation will be underestimated if positive correlation exists.

#2 The plots of standardized residuals against the fitted values for each triangle show clear downward trend, which indicates that the assumption under the model is not appropriate.

#3 The predicted ultimate paid-to-incurred loss ratios go far beyond 100% for accident years 8-10.

Residual plots for separate chain ladders



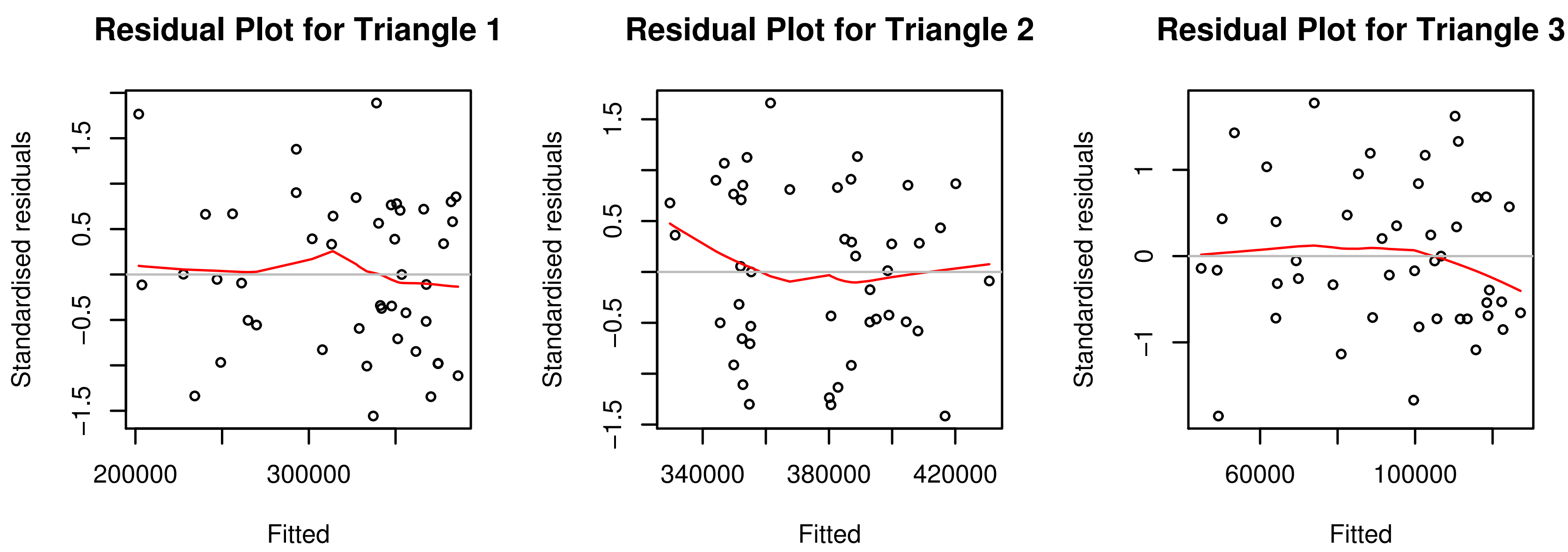
Fit GMCL

- Begin with a model that has the following structure

$$(\mathbf{A}_k, \mathbf{B}_k) = \begin{pmatrix} \beta_{10} & \beta_{11} & \beta_{12} & 0 \\ \beta_{20} & \beta_{21} & \beta_{22} & 0 \\ \beta_{30} & 0 & 0 & \beta_{33} \end{pmatrix} \tag{9}$$

- Eliminate redundant parameters in a step-wise fashion
- Keep the diagonal components of the development matrix due to practical intuition
- Separate chain ladders for development years 7-10

Residual plots for selected GMCL



Predicted ultimate paid-to-incurred loss ratios

Table: Predicted Ultimate Paid-to-Incurred Loss Ratios (%)

AY	SCL	MCL	MuCL	GMCL
1	99.50	99.50	99.50	99.50
2	99.49	99.49	99.55	99.49
3	99.29	99.29	100.23	99.29
4	99.20	99.20	100.23	99.20
5	99.83	99.82	100.04	99.38
6	100.43	100.43	100.03	99.51
7	103.53	103.53	99.95	99.44
8	111.23	111.23	99.81	99.41
9	122.10	122.15	99.67	99.39
10	126.22	126.15	99.69	99.38
Total	105.57	105.56	99.88	99.40

Overview

- The R package *ChainLadder* includes several functions to fit the GMCL model starting from version 0.1.3-3
- The multivariate functions are written in the formal S4 language, where a lot of classes and methods are defined
- Main function *MultiChainLadder* to fit model and calculate mean square errors
- Main class *MultiChainLadder*, where various methods are defined, such as *coef, resid, summary, plot*
- Important class *triangles* for input data, where various methods are defined, such as *'[', '[<-', cbind2*

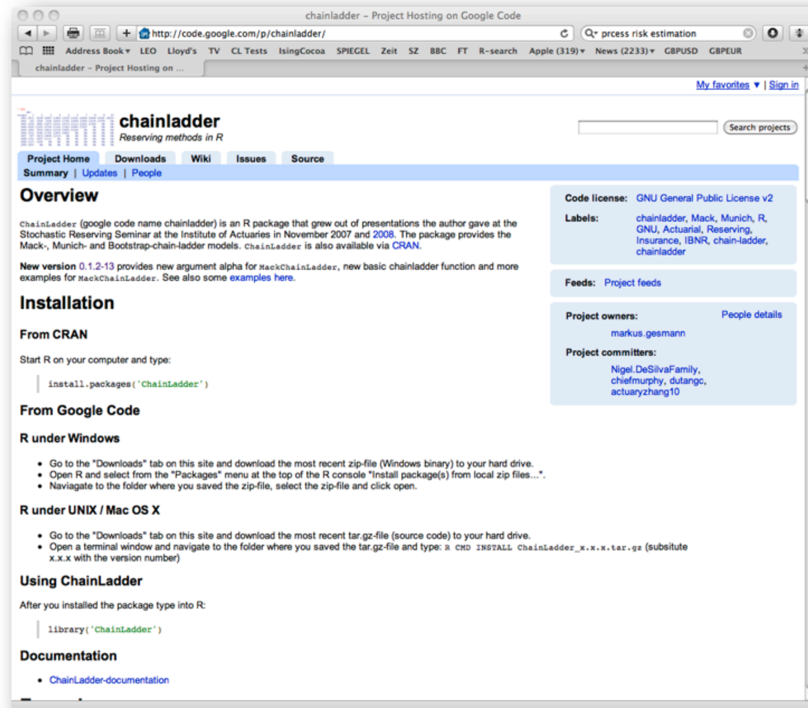
How to integrate ChainLadder into MS Office

Markus Gesmann

CAS San Diego 2010

The *ChainLadder* package

- Reserving functions
  - MackChainLadder
  - MunichChainLadder
  - BootChainLadder
  - MultiChainLadder
- Utility functions for triangles
- Visualisation
- Demos



Project web page: http://code.google.com/p/chainladder/
Current version: 0.1.3-3

Getting started

Start R and type for

- Installation:

```
install.packages("ChainLadder")
```

- Loading the package:

```
library(ChainLadder)
```

- Help:

```
?ChainLadder
```

- Examples:

```
example(ChainLadder)
```

- Demos:

```
demo(package="ChainLadder")
```

Integrating R with MS Office

Agenda

- Using R graphics in MS Office apps
- Exchanging data with Excel and DBs
- R and MS Office interfaces via statconnDCOM server
  - Embed R into VBA
  - Embed R into Excel with RExcel
  - Embed R into Word with SWord
  - Embed MS Office into R: rcom

Using R graphics in MS Office

- Windows meta-files (WMF, or EMF (Enhanced meta-file) are the native vector graphic format for MS Windows
- High quality, and editable format for MS Office
- Create WMF-files in R with *win.metafile()*

```
data(Insurance, package="MASS")
win.metafile(file="C:/Temp/Testplot.wmf")
plot(Claims/Holders ~ Age, data=Insurance)
dev.off()
```

Data exchange: Clipboard

Use clipboard for small data sets

- **Write R data into clipboard and paste into Excel**

```
data(Insurance, package="MASS")
write.table(Insurance, file="clipboard",
          sep="\t", na="",rownames=FALSE)
```

- **Copy in Excel and read into R**

```
read.table(file = "clipboard", sep="\t")
```

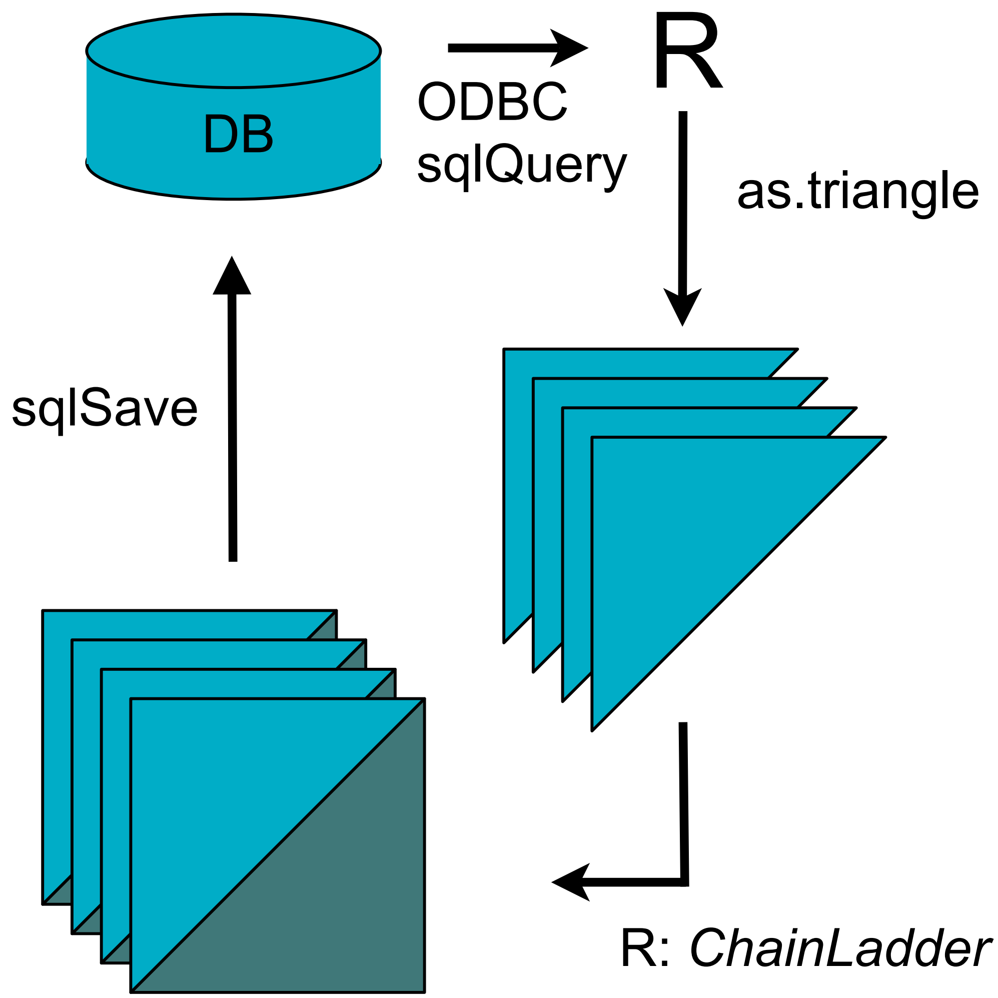

Data exchange: Excel

The **xlsReadWrite** package provides functions to read and write Excel files natively

- **read.xls**(file, colNames = TRUE, sheet = 1, type = "data.frame", from = 1, rowNames = NA, colClasses = NA, checkNames = TRUE, dateTimeAs = "numeric", stringsAsFactors = default.stringsAsFactors())
- **write.xls**(x, file, colNames = TRUE, sheet = 1, from = 1, rowNames = NA)

Data Exchange: ODBC / Databases

- Use SQL to interact with databases via RODBC
- Use R to transform tables into triangles
- Apply *ChainLadder* function across many triangles in **one statement**
- Write results back into database




```
library(ChainLadder)
demo(DatabaseExamples)
```

R and MS Office Interfaces

We need:

- **statconnDCOM Server:** server which allows to integrate R into other applications.
- **rscproxy**: R package which is required for rcom and for the statconn (D)COM Server.
- **rcom**: R package allowing R and other apps to communicate
- **RExcel:** embeds R into MS Excel.
- **SWord:** embeds R into MS Word

statcom home page

http://rcom.univie.ac.at/

Embed R into your VBA code

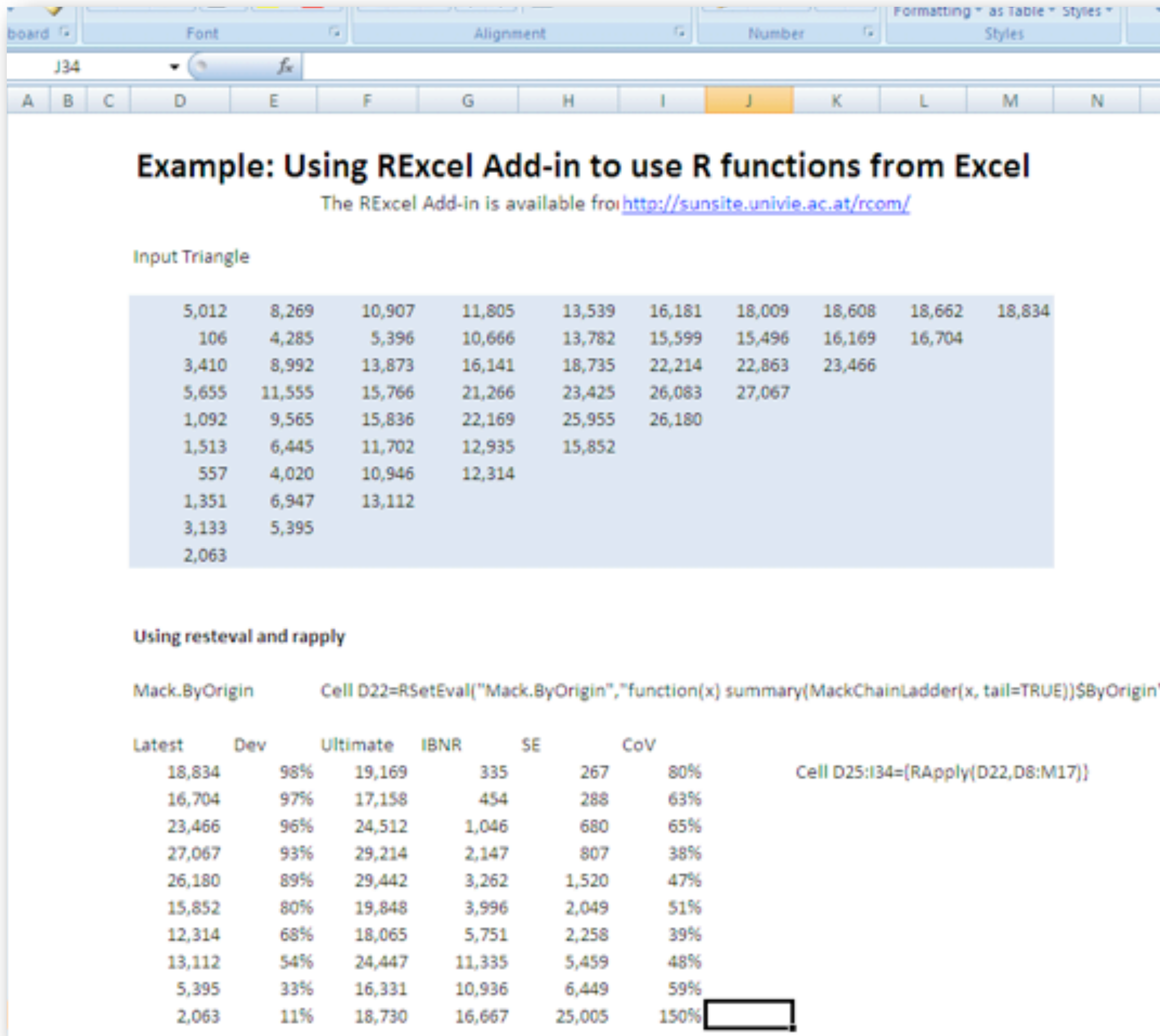
StatConnector allows to use R within MS Office VBA

Add reference to **StatConnectorSrv** 1.1 Type Library

```
Sub FirstR()
 Dim nrandom As Integer, x As Double
 nrandom = 100
 Set StaR = New StatConnector
 StaR.Init ("R")
  With StaR
    .SetSymbol "n", nrandom
   .EvaluateNoReturn ("x <- rnorm(n)")
   .EvaluateNoReturn ("pdf(file='c:/Temp/Testplot.pdf')")
   .EvaluateNoReturn ("hist(x)")
   .EvaluateNoReturn ("dev.off()")
  x = .Evaluate("mean(x)")
 End With
 Debug.Print x
End Sub
```

RExcel: Embed R into Excel

- RExcel allows to use R functions within Excel
- R function can be embedded and are interactive
- RExcel comes with example spreadsheets
- See example in ChainLadder package



Example: Using RExcel Add-in to use R functions from Excel

The RExcel Add-in is available fro http://sunsite.univie.ac.at/rcom/

Input Triangle

5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1,092	9,565	15,836	22,169	25,955	26,180				
1,513	6,445	11,702	12,935	15,852					
557	4,020	10,946	12,314						
1,351	6,947	13,112							
3,133	5,395								
2,063									

Using resteval and rapply

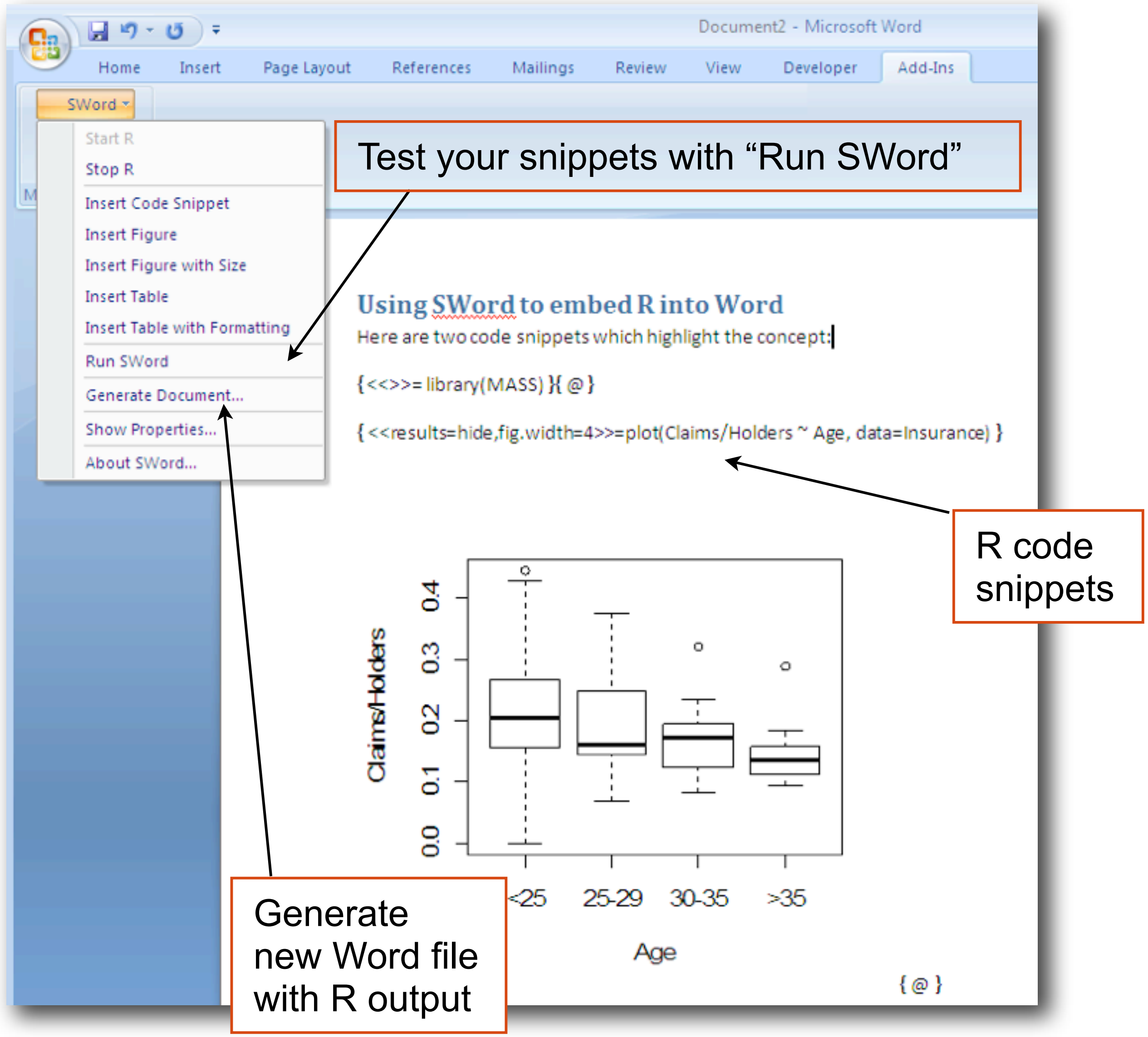
Mack.ByOrigin Cell D22=RSetEval("Mack.ByOrigin","function(x) summary(MackChainLadder(x, tail=TRUE))$ByOrigin")

Latest	Dev	Ultimate	IBNR	SE	CoV
18,834	98%	19,169	335	267	80%
16,704	97%	17,158	454	288	63%
23,466	96%	24,512	1,046	680	65%
27,067	93%	29,214	2,147	807	38%
26,180	89%	29,442	3,262	1,520	47%
15,852	80%	19,848	3,996	2,049	51%
12,314	68%	18,065	5,751	2,258	39%
13,112	54%	24,447	11,335	5,459	48%
5,395	33%	16,331	10,936	6,449	59%
2,063	11%	18,730	16,667	25,005	150%

Cell D25:I34={RApply(D22,D8:M17)}

SWord: Embed R into Word

- SWord is similar to Sweave
- R code snippets are embed into the Word file
- R graphics work as well
- Not GNU! Read licence!
- See example in ChainLadder package



rcom: Control MS Office from R

- The *rcom* R-package allows to communicate with MS Office. Here we create a PowerPoint slide with a chart.

```
data(Insurance, package="MASS")
myf <- tempfile()
win.metafile(file=myf)
 plot(Claims/Holders ~ Age, data=Insurance)
dev.off()
library(rcom)
ppt <- comCreateObject("Powerpoint.Application")
comSetProperty(ppt,"Visible",TRUE)
myPresColl <- comGetProperty(ppt,"Presentations")
myPres <- comInvoke(myPresColl,"Add")
mySlides <- comGetProperty(myPres,"Slides")
mySlide <- comInvoke(mySlides,"Add",1,12)
myShapes <- comGetProperty(mySlide,"Shapes")
myPicture <- comInvoke(myShapes,"AddPicture",myf,0,1,100,10)
```

Summary

- R graphics can easily be integrated and edited in MS Office applications
- Data exchange between R and Excel is fairly flexible
- R can be used as a server for MS Office
- Recommended:
  - Heiberger R. and Neuwirth E. R Through Excel, Springer Verlag 2009
  - RODBC to connect to data bases
- R and VBA code of this presentation attached:

R_and_MS_Office.R

Conclusion

- R is ideal for reserving
  - Built-in functions for statistical modelling
  - Powerful language for data manipulations and visualisation
  - Easy to set-up connections to databases (ODBC)
  - RExcel add-in allows to share R functions with colleagues without R knowledge
  - *rcom* allows to control MS Office from R
  - Effective knowledge transfer - plain text files
  - See also Formatted Actuarial Vignettes in R: http://www.favir.net