



ΕΚΤΙΜΗΣΗ ΑΠΟΘΕΜΑΤΩΝ ΜΕ ΧΡΗΣΗ ΜΕΘΟΔΩΝ

KALMAN FILTER

Ι. Μπαντούνας, Γ. Πιτσέλης

Πανεπιστήμιο Πειραιώς

ibadounas@yahoo.gr & pitselis@unipi.gr

ΠΕΡΙΛΗΨΗ

Η εκτίμηση των αποθεμάτων μίας ασφαλιστικής εταιρίας στον κλάδο ζημιών αποτελεί μία άκρως σημαντική διαδικασία καθώς έχει ως σκοπό την οποιαδήποτε κάλυψη των μελλοντικών αξιώσεων από ασφαλιστικούς κινδύνους οι οποίοι έχουν ήδη επέλθει. Η διαδικασία εκτίμησης των αποθεμάτων περιλαμβάνει αρκετά περίπλοκες μεθόδους, οι οποίες πρέπει να συνδυάζονται με την κρίση και την εμπειρία του εκάστοτε αναλογιστή. Οι περισσότερες μέθοδοι αποθεματοποίησης απαιτούν την ύπαρξη παρελθοντικών στοιχείων της εταιρείας, πάνω στα οποία θα βασιστεί η εκτίμηση των μελλοντικών αποθεμάτων. Στην παρούσα εργασία παρουσιάζεται μία μέθοδος εκτίμησης των αποθεμάτων με χρήση του φίλτρου Kalman (Kalman Filter), έναν αλγόριθμο που εφαρμόζεται σε συστήματα που δέχονται εξωτερικές φυσικές διαταραχές (θορύβους) και έχει ως σκοπό τη δημιουργία μιας νέας εκτίμησης της κατάστασης του συστήματος, απαλλαγμένη από διαταραχές. Επιπλέον, επεκτείνουμε τον αλγόριθμο κάνοντάς τον ανθεκτικό (robust) σε ακραίες τιμές, οι οποίες αν αγνοηθούν θα υπερεκτιμήσουν το τελικό απόθεμα που απαιτείται για την ασφαλιστική υποχρέωση της εταιρίας.

Λέξεις Κλειδιά: Αποθεματοποίηση, Φίλτρο Kalman, Ανθεκτικός (Robust).

1. Εισαγωγή

Η αποθεματοποίηση ζημιών είναι μια σημαντική διαδικασία των ασφαλιστικών εταιριών με σκοπό την πρόβλεψη των απαιτούμενων κεφαλαίων για την κάλυψη μελλοντικών απαιτήσεων. Στην αναλογιστική βιβλιογραφία υπάρχουν πολλές εργασίες που περιγράφουν αναλυτικά στοχαστικές μεθόδους αποθεματοποίησης απαιτήσεων, μεταξύ των οποίων είναι τα άρθρα των Kremer (1982), Verrall (1994) καθώς και τα βιβλία των Taylor (2000), Wüthrich & Merz (2008) και των Radtke et al. (2012). Ένα από τα βασικά προβλήματα των βασικών τεχνικών εκτίμησης των αποθεμάτων μίας ασφαλιστικής εταιρίας (όπως είναι αυτή της μεθόδου Chain Ladder) είναι ότι δημιουργούν μεγάλες «σχέσεις - συσχετίσεις» μεταξύ των ετών που συνέβη το ατύχημα (ή των ετών που καταγράφηκε μία ζημιά). Αυτό συμβαίνει κυρίως διότι τα μοντέλα αυτά βασίζονται στην υπόθεση ότι οι παράγοντες εξέλιξης έχουν την ίδια μορφή για κάθε έτος ατυχήματος. Χρησιμοποιώντας μοντέλα χώρου καταστάσεων (state space models) δίνεται η δυνατότητα στους παράγοντες εξέλιξης να έχουν διαφορετική μορφή ανάλογα με το έτος ατυχήματος. Συγκεκριμένα, τα μοντέλα χώρου καταστάσεων είναι πολύ σημαντικά γιατί κυρίως μπορούν να δώσουν μία ποικιλία από διαφορετικά μοντέλα τα οποία είναι εύκολα ερμηνεύσιμα, αλλά και ταυτόχρονα οδηγούν σε συνεπείς εκτιμήσεις μέσω της μεθόδου του φίλτρου Kalman (Kalman Filter). Στις ασφαλίσεις ζημιών μπορούν να χρησιμοποιηθούν ως ένα εργαλείο ώστε να αυτοματοποιηθεί η διαδικασία αποθεματοποίησης.

Ένα μοντέλο χώρου καταστάσεων προσδιορίζεται από δύο εξισώσεις οι οποίες δηλώνουν την εξάρτηση που υπάρχει μεταξύ των παρατηρήσεων και της κατάστασης

- Η εξίσωση παρατήρησης (observation equation) ορίζεται ως

$$\mathbf{y}_t = \mathbf{F}_t \boldsymbol{\theta}_t + \mathbf{w}_t, \quad t \geq 0,$$

όπου \mathbf{y}_t είναι ένα διάνυσμα με τις παρατηρήσεις, $\boldsymbol{\theta}_t$ είναι ένα άγνωστο διάνυσμα κατάστασης, \mathbf{F}_t είναι ένας γνωστός πίνακας, ο οποίος συνδέει το διάνυσμα κατάστασης $\boldsymbol{\theta}_t$ με τις παρατηρήσεις και τέλος \mathbf{w}_t είναι τα πιθανά σφάλματα.

- Η εξίσωση κατάστασης (state equation) ορίζεται ως

$$\boldsymbol{\theta}_{t+1} = \mathbf{G}_{t+1}\boldsymbol{\theta}_t + \mathbf{u}_{t+1}, \quad t \geq 0,$$

η οποία χαρακτηρίζει το διάνυσμα $\boldsymbol{\theta}_t$ ως Μαρκοβιανή μορφή χρησιμοποιώντας τον πίνακα καταστάσεων μετάβασης \mathbf{G}_{t+1} . Όπως και η εξίσωση παρατήρησης έτσι και η εξίσωση κατάστασης συμπεριλαμβάνει ένα τυχαίο διάνυσμα \mathbf{u}_t σφαλμάτων.

Το βασικό σημείο αυτής της προσέγγισης είναι η εξαγωγή συμπερασμάτων σχετικά με το διάνυσμα κατάστασης $\boldsymbol{\theta}_t$ ενώ έχει παρατηρηθεί ένα σύνολο δεδομένων \mathbf{y}_t . Γίνεται η υπόθεση ότι \mathbf{y}_t είναι ένα διάνυσμα $n \times 1$, $\boldsymbol{\theta}_t$ είναι ένα διάνυσμα καταστάσεων $r \times 1$, \mathbf{F}_t είναι πίνακας $n \times r$ και \mathbf{G}_t είναι πίνακας $r \times r$. Τα διανύσματα με τα σφάλματα \mathbf{w}_t και \mathbf{u}_t είναι διαστάσεων $n \times 1$ και $r \times 1$ με πίνακες διακύμανσης-συνδιακύμανσης

$$E[\mathbf{u}_t \mathbf{u}'_{t+k}] = \begin{cases} \mathbf{Q}_t, & k = 0, \\ 0, & \text{αλλού} \end{cases} \quad \text{και} \quad E[\mathbf{w}_t \mathbf{w}'_{t+k}] = \begin{cases} \mathbf{H}_t, & k = 0, \\ 0, & \text{αλλού,} \end{cases}$$

όπου \mathbf{Q}_t και \mathbf{H}_t είναι $r \times r$ και $n \times n$ πίνακες οι οποίοι εξαρτώνται από το χρόνο t . Επιπλέον γίνεται η υπόθεση $E[\mathbf{u}_t \mathbf{w}'_{t+k}] = 0, \forall t, k$.

Υποθέτοντας ότι τα σφάλματα ακολουθούν την κανονική κατανομή, για την εκτίμηση του διανύσματος $\boldsymbol{\theta}_t$ χρησιμοποιείται ο αλγόριθμος του φίλτρου Kalman.

2. Ο Αλγόριθμος του Φίλτρου Kalman (Kalman Filter)

Η εκτίμηση του διανύσματος $\boldsymbol{\theta}_t$ βασίζεται σε μία σειρά δεδομένων που έχουν παρατηρηθεί $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t$. Ο αλγόριθμος του φίλτρου Kalman είναι ένα σύνολο από αναδρομικές σχέσεις στις οποίες γίνεται η εκτίμηση του διανύσματος $\boldsymbol{\theta}_t$ σύμφωνα με το διάνυσμα $\boldsymbol{\theta}_{t-1}$ και το νέο σετ δεδομένων \mathbf{y}_t . Τα βήματα του αλγορίθμου περιγράφονται παρακάτω:

$$\text{Βήμα 1: } \hat{\mathbf{y}}_t = \mathbf{F}_t \hat{\boldsymbol{\theta}}_{t|t-1},$$

$$\text{Βήμα 2: } \boldsymbol{\varepsilon}_t = \mathbf{y}_t - \hat{\mathbf{y}}_t,$$

$$\text{Βήμα 3: } \hat{\boldsymbol{\theta}}_{t|t} = \hat{\boldsymbol{\theta}}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t' (\mathbf{F}_t' \mathbf{P}_{t|t-1} \mathbf{F}_t + \mathbf{H}_t)^{-1} \boldsymbol{\varepsilon}_t,$$

$$\text{Βήμα 4: } \hat{\boldsymbol{\theta}}_{t+1|t} = \mathbf{G}_{t+1} \hat{\boldsymbol{\theta}}_{t|t},$$

$$\text{Βήμα 5: } \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t' (\mathbf{F}_t' \mathbf{P}_{t|t-1} \mathbf{F}_t + \mathbf{H}_t)^{-1} \mathbf{F}_t' \mathbf{P}_{t|t-1},$$

$$\text{Βήμα 6: } \mathbf{P}_{t+1|t} = \mathbf{G}_{t+1} \mathbf{P}_{t|t} \mathbf{G}_{t+1}' + \mathbf{Q}_{t+1}.$$

Στον παραπάνω αλγόριθμο, $\hat{\boldsymbol{\theta}}_{t|t}$ είναι η εκτίμηση του $\boldsymbol{\theta}_t$ όταν είναι γνωστό το σύνολο των δεδομένων $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t$ στο χρόνο t , ενώ $\mathbf{P}_{t|t}$ είναι ο πίνακας διακύμανσης - συνδιακύμανσης αυτής της εκτίμησης. Αντίστοιχα, το διάνυσμα $\hat{\boldsymbol{\theta}}_{t+1|t}$ και ο πίνακας $\mathbf{P}_{t+1|t}$ είναι η εκτίμηση του $\boldsymbol{\theta}_{t+1}$ και της διακύμανσης σύμφωνα με το ίδιο σύνολο δεδομένων σε χρόνο t . Στην πραγματικότητα ο αλγόριθμος δεν επανεκτιμά μόνο το διάνυσμα $\boldsymbol{\theta}_{t+1}$ αλλά και τον πίνακα διακύμανσης-συνδιακύμανσης αυτής της εκτίμησης. Στην περίπτωση που τα σφάλματα κατανέμονται κανονικά, είναι εφικτό να κατασκευαστούν $(1-\alpha)100\%$ διαστήματα εμπιστοσύνης για τις παραμέτρους σύμφωνα με τους τύπους $\hat{\boldsymbol{\theta}}_{k,t+1|t} \pm t_a \sqrt{\mathbf{p}_{k,k,t+1|t}}$, όπου $\hat{\boldsymbol{\theta}}_{k,t+1|t}$ είναι το k στοιχείο του $\hat{\boldsymbol{\theta}}_{t+1|t}$, $\mathbf{p}_{k,k,t+1|t}$ είναι το (k, k) διαγώνιο στοιχείο του πίνακα $\mathbf{P}_{t+1|t}$ και t_a είναι το αντίστοιχο ποσοστημόριο.

3. Ο Ανθεκτικός (Robust) Αλγόριθμος Φίλτρου Kalman

Μία από τις υποθέσεις του αλγορίθμου του φίλτρου Kalman είναι η κανονικότητα των σφαλμάτων και τα περισσότερα μοντέλα βασίζονται σε αυτή την υπόθεση. Σε περίπτωση όμως ασυμμετρίας της κατανομής των σφαλμάτων είτε έκτοπων παρατηρήσεων (outliers), κάποιες υποθέσεις όπως η προηγούμενη παραβιάζονται, το οποίο σημαίνει ότι οι εκτιμήσεις των παραμέτρων, των διαστημάτων εμπιστοσύνης και άλλων στατιστικών ποσοτήτων είναι αναξιόπιστες (Cipra, & Romera, 1991). Στο Βήμα 3 του αλγορίθμου, η έλλειψη κανονικότητας των σφαλμάτων ή η ύπαρξη έκτοπων τιμών θα οδηγήσει σε λανθασμένες εκτιμήσεις για το $\boldsymbol{\theta}_{t|t}$ αλλά και για όλο το διάνυσμα $\boldsymbol{\theta}_{t+i|t+i}, \forall i > 0$. Για το λόγο αυτό, μία απλή ανθεκτική (robust) στρατηγική είναι να απαλειφθούν τέτοιες επιρροές αντικαθιστώντας τα σφάλματα $\boldsymbol{\varepsilon}_t$ με μία φραγμένη συνάρτηση $\psi(\boldsymbol{\varepsilon}_t)$. Αντί επομένως να χρησιμοποιηθούν τα σφάλματα $\boldsymbol{\varepsilon}_t$, εφαρμόζεται η συνάρτηση του Huber (Huber, 1981)

$$H_b(X) = X \cdot \min \left\{ 1, \frac{b}{\|X\|} \right\},$$

όπου $\| \cdot \|$ είναι η ευκλείδεια νόρμα και $0 < b < \infty$. Σε αυτή την περίπτωση, το Βήμα 3 του αλγορίθμου μετασχηματίζεται ως εξής:

$$\hat{\boldsymbol{\theta}}_{t|t} = \hat{\boldsymbol{\theta}}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t' (\mathbf{F}_t' \mathbf{P}_{t|t-1} \mathbf{F}_t + \mathbf{H}_t)^{-1} H_b(\boldsymbol{\varepsilon}_t),$$

όπου η σταθερά b είναι μία παράμετρος προς εκτίμηση. Ο αλγόριθμος μπορεί να γίνει ανθεκτικός στην περίπτωση που μπορούν να χρησιμοποιηθούν διαφορετικές συναρτήσεις αλλά παρόμοιες με αυτή του Huber (Pitselis et al., 2015, Cipra & Romera, 1991). Μία προσέγγιση για να επιλεγεί η σταθερά είναι χρησιμοποιώντας την αποδοτικότητα (efficiency) σε σχέση με το απλό μοντέλο του φίλτρου Kalman. Αν οι τιμές της σταθεράς b είναι μικρές, τότε η σχετική αποδοτικότητα παίρνει μεγάλες τιμές (Moore et. al, 1980).

4. Δημιουργία Μοντέλων Φίλτρου Kalman σε Τρίγωνα Ζημιών

Οι ζημιές μίας ασφαλιστικής εταιρίας στον κλάδο ζημιών μελετώνται κυρίως χρησιμοποιώντας τρίγωνα εξέλιξης ζημιών. Τη μορφή των δεδομένων δηλώνει ο Πίνακας 4. Συγκεκριμένα, οι γραμμές δηλώνουν το έτος που συνέβη το ατύχημα, ενώ οι στήλες δηλώνουν το χρόνο εξέλιξης της ζημιάς, δηλαδή το χρονικό διάστημα που χρειάστηκε ώστε να καταβληθεί αποζημίωση (Verrall, 1994).

Πίνακας 4: Τρίγωνο εξέλιξης αποθεμάτων

Έτος ατυχήματος i	Έτος εξέλιξης j				
	1	2		$n-1$	n
1	$C_{1,1}$	$C_{1,2}$...	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$		$C_{2,n-1}$	
⋮	⋮		⋮		
	⋮	$C_{n-1,2}$			
n	$C_{n,1}$				

Επίσης, η διαγώνιος δηλώνει το ημερολογιακό έτος. Η τιμή $C_{i,j}$ δηλώνει μία ζημιά που έγινε το έτος i αλλά πληρώθηκε το έτος j . Φυσικά, τα δεδομένα παρουσιάζονται ως τρίγωνο διότι οι γνωστές παρατηρήσεις είναι έως και την τελευταία διαγώνιο με στοιχεία $C_{n-i+1,i}, \forall i=1,2,\dots,n$, τα οποία δηλώνουν το ποσό των ζημιών που έγιναν το έτος $n-i+1$ και πληρώθηκαν το έτος n , όπου n είναι το σύνολο των ετών που είναι γνωστά τα μεγέθη των ζημιών. Όλες οι υπόλοιπες τιμές, δηλαδή το τρίγωνο δεξιά κάτω, πρέπει να εκτιμηθούν διότι αποτελούν την εξέλιξη των ζημιών που έγιναν και πρόκειται να πληρωθούν σε μελλοντικό χρόνο. Συγκεκριμένα, θα πρέπει

να υπολογιστεί το απόθεμα $R_i = \sum_{n-i+2}^n C_{i,j}$ που αναμένεται να έχει μία ασφαλιστική

εταιρία για κάθε έτος ατυχήματος $i = 2, \dots, n$, με $R_1 = 0$.

Για τη μοντελοποίηση των ζημιών ενός τριγώνου χρησιμοποιείται το μοντέλο των λογαρίθμων των δεδομένων

$$\log(C_{i,j}) = \mu + \alpha_i + \beta_j + e_{i,j},$$

με $\mu > 0$, α_i μία παράμετρος που περιγράφει το έτος ατυχήματος και β_j μία παράμετρος που περιγράφει το έτος εξέλιξης της ζημιάς ενώ $e_{i,j}$ είναι τα σφάλματα που ακολουθούν την κανονική κατανομή με μέση τιμή 0 και κάποια σταθερή διακύμανση (De Jong and Zehnwirth, 1983). Για την εκτίμηση των αποθεμάτων, χρησιμοποιώντας το φίλτρο Kalman τα δεδομένα θα έχουν την ακόλουθη μορφή σύμφωνα με το ημερολογιακό έτος (βλέπε Verrall, 1994):

$$C_{1,1}, \begin{bmatrix} C_{1,2} \\ C_{2,1} \end{bmatrix}, \begin{bmatrix} C_{1,3} \\ C_{2,2} \\ C_{3,2} \end{bmatrix}, \dots$$

Η αναπαράσταση του λογαριθμικού μοντέλου σε ένα Μοντέλο Χώρου Καταστάσεων είναι

$$\underbrace{\begin{pmatrix} \log(C_{1,t}) \\ \log(C_{2,t-1}) \\ \vdots \\ \log(C_{t,1}) \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 1 & 0 \end{pmatrix}}_{F_t} \underbrace{\begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \beta_t \end{pmatrix}}_{\theta_t} + \underbrace{\begin{pmatrix} w_{1,t} \\ w_{2,t-1} \\ \vdots \\ w_{t,1} \end{pmatrix}}_{w_t},$$

με το διάνυσμα καταστάσεων $\theta_t = (\mu, \alpha_1, \dots, \beta_t)'$ να ικανοποιεί την εξίσωση

$$\boldsymbol{\theta}_{t+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{G}_{t+1}} \boldsymbol{\theta}_t + \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{u}_t,$$

όπου \mathbf{u}_t είναι ένα διάνυσμα τυχαίων σφαλμάτων. Να τονιστεί ότι οι διαστάσεις όλων των πινάκων μεταβάλλονται δυναμικά καθώς νέες πληροφορίες εισάγονται στο μοντέλο. Εφαρμόζοντας τον αλγόριθμο, προκύπτουν οι εκτιμήσεις των αγνώστων παραμέτρων ώστε να γίνει η τελική εκτίμηση του αποθέματος.

5. Αριθμητική εφαρμογή

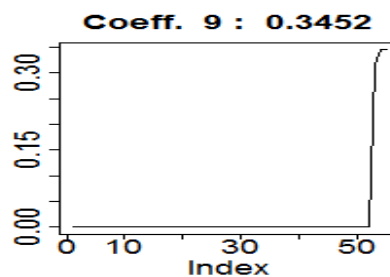
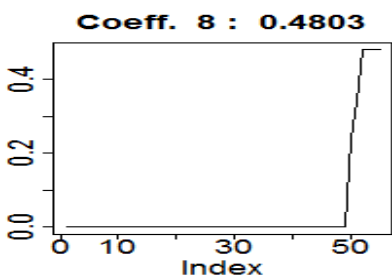
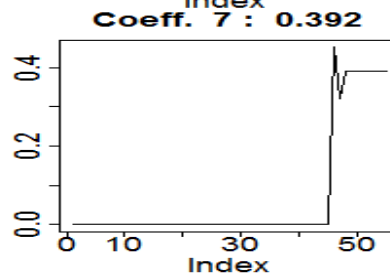
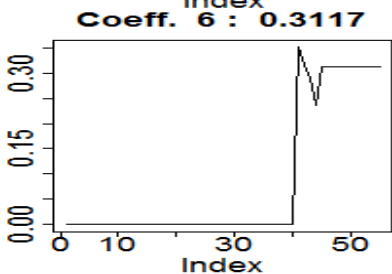
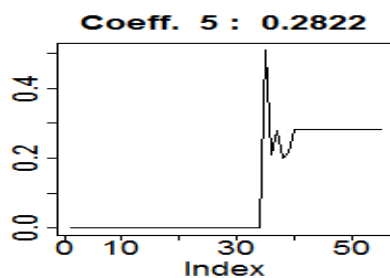
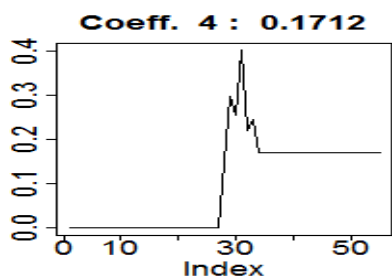
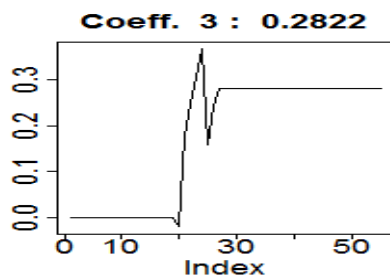
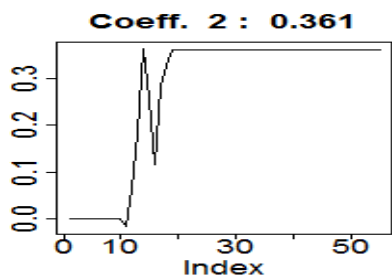
Για την εφαρμογή της μεθόδου θα χρησιμοποιήθηκε το πακέτο R και ένα ευρέως γνωστό σύνολο δεδομένων το οποίο δείχνει ο Πίνακας 5 (Taylor & Ashe, 1983). Στο συγκεκριμένο τρίγωνο εξέλιξης αποθεμάτων υπάρχουν δεδομένα για 10 έτη ατυχήματος καθώς και για 10 έτη εξέλιξης. Στόχος είναι η εκτίμηση του κάτω τριγώνου το οποίο θα οδηγήσει και στην εκτίμηση των αναγκαίων αποθεμάτων που καλείται να διαθέτει η εκάστοτε ασφαλιστική εταιρία ώστε να είναι φερέγγυα για τις μελλοντικές της υποχρεώσεις.

Πίνακας 5: Taylor and Ashe data (1983)

	1	2	3	4	5	6	7	8	9	10
1	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
2	352118	884021	933894	1183289	445745	320996	527804	266172	425046	
3	290507	1001799	926219	1016654	750816	146923	495992	280405		
4	310608	1108250	776189	1562400	272482	352053	206286			
5	443160	693190	991983	769488	504851	470639				
6	396132	937085	847498	805037	705960					
7	440832	847631	1131398	1063269						
8	359480	1061648	1443370							
9	376686	986608								
10	344014									

Εφαρμόζοντας τον αλγόριθμο του φίλτρου Kalman, λαμβάνουμε τις εκτιμήσεις του διανύσματος κατάστασης του μοντέλου (Πίνακας 6). Η Εικόνα 1 παρουσιάζει τη διαχρονική εκτίμηση κάποιων στοιχείων του διανύσματος κατάστασης (β_1 έως β_8) για τα οποία οι μεταβολές στις εκτιμώμενες τιμές είναι εμφανείς, καθώς εισάγονται νέες πληροφορίες στο τρίγωνο. Συγκεκριμένα, διαχρονικά εισάγονται νέες ζημιές στο τρίγωνο και επομένως επανεκτιμούνται οι παράμετροι του μοντέλου. Υπάρχουν περιπτώσεις όπου αρχικά κάποια από τα στοιχεία του διανύσματος κατάστασης δεν μπορούν να εκτιμηθούν, καθώς όμως εισέρχονται πληροφορίες στο μοντέλο τότε είναι εφικτό να εκτιμηθούν, αλλά και να επανεκτιμηθούν στα επόμενα βήματα του αλγορίθμου.

Εικόνα 1: Διαχρονική εκτίμηση στοιχείων του διανύσματος κατάστασης



Να τονιστεί ότι πρόκειται για ένα διάνυσμα, το οποίο περιέχει τις εκτιμήσεις των παραμέτρων και για τα έτη ατυχήματος (παράμετροι α_i), αλλά και για τα έτη εξέλιξης (παράμετροι β_i).

Πίνακας 6: Εκτιμήσεις διανύσματος κατάστασης

μ	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	
12.51	0.3610	0.2822	0.1712	0.2822	0.3118	0.392	0.4802	0.3452	
β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9

0.229	0.9111	0.9387	0.9649	0.3832	-0.005	-0.118	-0.439	-0.053	-1.393
-------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Χρησιμοποιώντας τις εκτιμήσεις του διανύσματος κατάστασης εκτιμάται το απόθεμα για τα υπόλοιπα έτη. Ο Πίνακας 7 παρουσιάζει τις εκτιμήσεις με τα μελλοντικά αναμενόμενα αποθέματα. Να τονιστεί ότι οι εκτιμήσεις αυτές αποτελούν το αναμενόμενο απόθεμα που χρειάζεται να έχει η ασφαλιστική εταιρία ώστε να είναι συνεπής στις μελλοντικές της υποχρεώσεις, οι οποίες προέκυψαν από την εξέλιξη των ζημιών που έχουν ήδη γίνει.

Πίνακας 7: Τρίγωνο εξέλιξης ζημιών με εκτιμήσεις

	1	2	3	4	5	6	7	8	9	10
1	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
2	352118	884021	933894	1183289	445745	320996	527804	266172	425046	10927
3	290507	1001799	926219	1016654	750816	146923	495992	280405	379508	102650
4	310608	1108250	776189	1562400	272482	352053	206286	228731	340091	91988
5	443160	693190	991983	769488	504851	470639	351067	256033	380684	102968
6	396132	937085	847498	805037	705960	404479	362429	264319	393004	106300
7	440832	847631	1131398	1063269	646825	439787	394066	287392	427310	115580
8	359480	1061648	1443370	1268870	710402	483014	432799	315640	469311	126940
9	376686	986608	1090101	1120657	627422	426594	382246	278771	414492	112113
10	344014	973601	1001989	1030076	576709	392113	351349	256238	380990	103051

Σύμφωνα με τις εκτιμήσεις του διανύσματος κατάστασης και το μελλοντικό πίνακα σχεδιασμού προκύπτει ο Πίνακας 8 όπου παρουσιάζονται οι εκτιμήσεις για το τελικό απόθεμα για κάθε έτος ατυχήματος.

Πίνακας 8: Αθροιστικές ζημιές και εκτίμηση αποθεμάτων και τελικών πληρωθεισών ζημιών ανά έτος ατυχήματος

	1	2	3	4	5	6	7	8	9	10	Συνολικές Τελικές Πληρωθείσες Ζημιές	Τελικό Εκτιμώμενο Απόθεμα
1	357848	1124788	1735330	2218270	2745596	3319994	3466336	3606286	3833515	3901463	3901463	0
2	352118	1236139	2170033	3353322	3799067	4120063	4647867	4914039	5339085	5450013	5350012	110928
3	290507	1292306	2218525	3235179	3985995	4132918	4628910	4909315	5288823	5391473	5391473	482158
4	310608	1418858	2195047	3757447	4029929	4381982	4588268	4816999	5157090	5249079	5249078	660811

5	443160	1136350	2128333	2897821	3402672	3873311	4224379	4480411	4861096	4964064	4964063	1090753	
6	396132	1333217	2180715	2985752	3691712	4096191	4458621	4722940	5115945	5222245	5222243	1530533	
7	440832	1288463	2419861	3483130	4129955	4569742	4963809	5251200	5678511	5794091	5794090	2310961	
8	359480	1421128	2864498	4133368	4843771	5326785	5759584	6075224	6544536	6671476	6671474	3806978	
9	376686	1363294	2453395	3574052	4201475	4628070	5010315	5289086	5703579	5815692	5815690	4452398	
10	344014	1317616	2319606	3349683	3926393	4318507	4669856	4926095	5307085	5410136	5410130	5066122	
											Σύνολο	53769716	19511642

Το ολικό απόθεμα εκτιμάται 19511642 ενώ το τυπικό σφάλμα εκτιμάται 6266024 (32.11%). Επίσης, η διακύμανση των σφαλμάτων του μοντέλου εκτιμάται 0.116217. Προφανώς, καθώς αυξάνεται διαχρονικά το έτος ατυχήματος, το συνολικό αναμενόμενο απόθεμα αυξάνεται. Να τονιστεί επίσης ότι γίνεται η υπόθεση ότι μετά τα 10 έτη δεν υπάρχει εξέλιξη των ζημιών, με αποτέλεσμα το πρώτο έτος ατυχήματος η εκτίμηση του αποθέματος να είναι μηδενική, όπως δείχνει ο Πίνακας 8.

Τέλος, εφαρμόζεται το μοντέλο με το ανθεκτικό φίλτρο Kalman για διάφορες τιμές της παραμέτρου b . Συγκεκριμένα, για διάφορες τιμές της παραμέτρου b εκτιμάται το απόθεμα, η διακύμανση των σφαλμάτων, η σχετική αποδοτικότητα σε σχέση με το απλό φίλτρο Kalman (δηλαδή ο λόγος των μέσων τετραγωνικών σφαλμάτων) καθώς επίσης και το ποσοστό του αποθέματος ως προς τη συνολική διακύμανση των αποθεμάτων. Όπως φαίνεται από τα αποτελέσματα, καθώς αυξάνεται η τιμή της σταθεράς b , μειώνεται το απόθεμα που απαιτείται και επίσης η αποδοτικότητα τείνει στο 1, καθώς σε αυτή την περίπτωση δεν εμφανίζεται κάποια ακραία τιμή και οι δύο αλγόριθμοι βγάζουν ίδια αποτελέσματα. Επίσης, μειώνεται και το ποσοστό του αποθέματος ως προς τη συνολική διακύμανση. Ο Πίνακας 9 παρουσιάζει τα αποτελέσματα της εφαρμογής του Ανθεκτικού φίλτρου Kalman στα δεδομένα.

Πίνακας 9: Αποτελέσματα Ανθεκτικού Φίλτρου Kalman

a/a	--b--	Απόθεμα	σ^2	Αποδοτικότητα	%SE
1	10.00	26165879	0.17	1.43	0.36
2	10.20	25527872	0.16	1.37	0.36
3	10.40	24919698	0.15	1.32	0.35

4	10.60	24339883	0.15	1.27	0.35
5	10.80	23787045	0.14	1.22	0.35
6	11.00	23259881	0.14	1.18	0.34
7	11.20	22757171	0.13	1.14	0.34
8	11.40	22277765	0.13	1.11	0.33
9	11.60	21820582	0.13	1.08	0.33
10	11.80	21384606	0.12	1.05	0.33
11	12.00	20968880	0.12	1.03	0.33
12	12.20	20572506	0.12	1.02	0.32
13	12.40	20194635	0.12	1.01	0.32
14	12.60	19834471	0.12	1.00	0.32
15	12.80	19511642	0.12	1.00	0.32
16	13.00	19511642	0.12	1.00	0.32

ABSTRACT

The estimate of the total loss reserve of an insurance company is a very important process as it intends to cover future claims from insurance risks that have already occurred. The procedure of loss reserving involves several complicated methods, which must be combined with the judgment and experience of the actuary. Most loss reserving methods require past claims data for the estimation of future reserves. In this work, we present a loss reserving method using Kalman Filter, an algorithm applied to systems that receive external physical disturbances (noises) and aims to create a new estimate of the state of the system without disturbances. In addition, we extend the algorithm by making it robust at extreme values, which, if ignored, will overestimate the final reserve required for the company's liability.

ΑΝΑΦΟΡΕΣ

- Busse, M., Müller, U., & Dacorogna, M., (2010). Robust Estimation of Reserve Risk, *ASTIN Bulletin*, 40(2), 453-489.
- Cipra, T., & Romera, R., (1991). Robust Kalman Filter and its application in time series analysis. *Kybernetika*, 27(6), 481-494.
- Huber, P. J., (1981). *Robust Statistics*, Wiley, New York.

- De Jong, P., & Zehnwirth, B., (1983). Claims reserving, state-space models and the Kalman filter. *Journal of the Institute of Actuaries*, 110, 157-181.
- Kremer, E., (1982). IBNR-Claims and the two-way model of ANOVA. *Scandinavian Actuarial Journal*, 47–55.
- Moore, J. B., & Anderson, B. D. O., (1980). Coping with singular transition matrices in estimation and control stability theory, *Int. J. Contr.*, vol. 3.
- Pitselis, G., (2008). Robust Regression Credibility: The Influence Function Approach, *Insurance Mathematics and Economics*, 288-300.
- Pitselis, G., Grigoriadou, V., & Badounas., I., (2015). Robust loss reserving in a log-linear model. *Insurance: Mathematics and Economics*, 64, 14–27.
- Radtke, M., Schmidt, K. D., & Schnaus, A., (2012). *Handbook on Loss Reserving, EAA Lecture Notes*, European Actuarial Academy, Springer.
- Taylor, G., C. & Ashe, F.R., (1983). Second moments of estimates of outstanding claims. *Journal of Econometrics*, 23, 37–61.
- Taylor, G., C., (2000). *Loss Reserving - An Actuarial Perspective*. Kluwer Academic Publishers, Norwell, Mass.
- Verdonck, T., Van Wouwe, & M., Dhaene, J., (2009). A Robustification of the Chain-Ladder Method. *North American Actuarial Journal*, 13(2), 280-298.
- Verrall, R. J., (1989). A state space representation of the chain-ladder linear model. *Journal of the Institute of Actuaries*, 116, 589-610.
- Verrall, R. J., (1994). *Statistical methods for the chain-ladder technique*. Casualty Actuarial Society Forum, Spring 1994, 393-446.
- Wüthrich, M. V., & Merz, M., (2008). *Stochastic Claims Reserving Methods in Insurance*, John Wiley.