



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Στοχαστική Ανάλυση Χρονοσειρών

Περιεχόμενο:

Τοπικό Μοντέλο (Local Level Model)

Μοντέλα Καταστάσεως – Χώρου (SSM)



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Τοπικό Μοντέλο (Local Level Model)



ARIMA και Μοντέλα Χώρου Καταστάσεων (State Space)

Τα μοντέλα ARIMA περιγράφουν την τρέχουσα τιμή της χρονοσειράς ως συνδυασμό:

- προηγούμενων παρατηρήσεων (AR),
- προηγούμενων σφαλμάτων πρόβλεψης (MA),
- και διαφορών, οι οποίες χρησιμοποιούνται για την αντιμετώπιση της μη στασιμότητας.

Μία διαφορετική προσέγγιση είναι τα μοντέλα χώρου καταστάσεων, τα οποία εισάγουν λανθάνουσες (μη παρατηρούμενες) καταστάσεις που περιγράφουν την υποκείμενη δυναμική της διαδικασίας.

Το μοντέλο διατυπώνεται μέσω δύο εξισώσεων:

- Εξίσωση κατάστασης: περιγράφει την εξέλιξη των λανθανουσών καταστάσεων στον χρόνο.
- Εξίσωση παρατήρησης: συνδέει τις λανθάνουσες καταστάσεις με τα παρατηρούμενα δεδομένα.

Η δομή αυτή επιτρέπει ευέλικτη μοντελοποίηση συνιστωσών όπως τάση, εποχικότητα και κύκλοι.



Γένεση ενός Local Level μοντέλου

Για να μετρήσουμε την εξέλιξη της στάθμης ενός ποταμού, πηγαίνουμε κάθε χρόνο σε ένα συγκεκριμένο σημείο και κάνουμε τη μέτρηση. Όμως, η μέτρηση επηρεάζεται από τυχαίους παράγοντες που δεν γνωρίζουμε, όπως για παράδειγμα μία πιθανή βροχή που προηγήθηκε ή ένα σφάλμα του εργαλείου μέτρησης. Οπότε, μετρούμε τη στάθμη του νερού με μια πρόσθετη, μη ελεγχόμενη και τυχαία ανακρίβεια. Η κατάσταση αυτή αντιπροσωπεύεται στην εξίσωση:

$$\text{Μέτρηση Στάθμης}_t = \text{Πραγματική Στάθμη}_t + \varepsilon_t$$

όπου ε_t το σφάλμα της μέτρησης.

Είναι επίσης λογικό να υποθέσουμε ότι η πραγματική στάθμη του νερού του ποταμού αλλάζει με την πάροδο του χρόνου όπως για παράδειγμα στην περίπτωση όπου χτιστεί ένα φράγμα. Δηλαδή, η τρέχουσα πραγματική στάθμη εξαρτάται από την στάθμη του προηγούμενου έτους συν την στοχαστική μεταβολή της τελευταίας περιόδου.

Συνεπώς, είναι λογικό να ενσωματώσουμε και την ακόλουθη εξίσωση:

$$\text{Πραγματική Στάθμη}_t = \text{Πραγματική Στάθμη}_{t-1} + \eta_t$$

όπου η_t το σφάλμα της τιμής.



Γένεση ενός Local Level μοντέλου

Οι δύο εξισώσεις

$$\text{Μέτρηση Στάθμης}_t = \text{Πραγματική Στάθμη}_t + \varepsilon_t$$

$$\text{Πραγματική Στάθμη}_t = \text{Πραγματική Στάθμη}_{t-1} + \eta_t$$

αποτελούν ένα Local Level μοντέλο.

(το οποίο αργότερα θα το αναγνωρίσουμε ως μία περίπτωση State Space μοντέλου).

Το "local" δηλώνει ότι το επίπεδο (level) δεν είναι καθολικό χαρακτηριστικό της σειράς, αλλά ένα χρονικά μεταβαλλόμενο, σημείο αναφοράς.



Βασική Ιδέα Local Level μοντέλου

- Υποθέτουμε ότι η χρονοσειρά κινείται γύρω από ένα λανθάνον επίπεδο (level).
- Το επίπεδο αυτό δεν είναι σταθερό, αλλά μεταβάλλεται τυχαία στον χρόνο.
- Δεν επιβάλλεται ντετερμινιστική τάση ή προκαθορισμένη μορφή εξέλιξης.

Το local level model γράφεται ως:

$$X_t = \mu_t + \varepsilon_t,$$

$$\mu_t = \mu_{t-1} + \eta_t.$$

X_t : παρατηρούμενη χρονοσειρά

μ_t : λανθάνουσα στάθμη (level)

ε_t : σφάλμα παρατήρησης

η_t : σφάλμα εξέλιξης του επιπέδου



Σκοπός Local Level μοντέλου

Ο σκοπός αυτού του μοντέλου είναι:

- α) Να εκτιμήσει την τάση της χρονοσειράς μ_t , η οποία δεν παρατηρείται άμεσα, επομένως το μοντέλο βοηθά στην εξαγωγή της από τα παρατηρούμενα δεδομένα.
- β) Να ποσοτικοποιήσει την αβεβαιότητα που σχετίζεται με το θόρυβο μέτρησης ε_t όσο και με την εξέλιξη της υποκείμενης τάσης η_t .

Οι άγνωστες παράμετροι του μοντέλου είναι:

- α) Η διασπορά του σφάλματος της μέτρησης σ_ε^2 .
- β) Η διασπορά του σφάλματος της πραγματικής τιμής σ_η^2 .
- γ) Η αρχική τιμή μ_0 .



Local Level Model

$$X_t = \mu_t + \varepsilon_t,$$
$$\mu_t = \mu_{t-1} + \eta_t.$$

Παρατηρήσεις

- Σε ένα Local Level μοντέλο, οι χρονοσειρές X_t και μ_t είναι εν γένει μη στάσιμες, καθιστώντας τα μοντέλα αυτού του είδους κατάλληλα για μη στάσιμες σειρές.
- Παρατηρούμε ότι το Local Level μοντέλο είναι ένας τυχαίος περίπατος με πρόσθετο Measurement Error. Ειδικότερα, αν $\sigma_\varepsilon = 0$, τότε το μοντέλο γίνεται ο τυχαίος περίπατος $X_t = X_{t-1} + \eta_t$.
- Αν $\sigma_\eta = 0$, τότε το μοντέλο γίνεται ο λευκός θόρυβος με σταθερή μέση τιμή $X_t = \mu + \varepsilon_t$.



Local Level Model

$$X_t = \mu_t + \varepsilon_t,$$
$$\mu_t = \mu_{t-1} + \eta_t.$$

Παρατηρήσεις

- Μεγάλο σ_η^2 : μεγάλη μεταβλητότητα στο πραγματικό επίπεδο μ_t .
- Μικρό σ_η^2 : σχεδόν σταθερή μέση τιμή.
- Μεγάλο σ_ε^2 : μεγάλο σφάλμα μέτρησης.
- Λόγος $\sigma_\varepsilon^2/\sigma_\eta^2$: περιγράφει την κατανομή της μεταβλητότητας μεταξύ του
 - θορύβου παρατήρησης (ε_t) και της
 - πραγματικής μεταβολής του επιπέδου (η_t).
- Αν $\sigma_\varepsilon^2/\sigma_\eta^2$ μεγάλο, τότε οι αποκλίσεις της χρονοσειράς ερμηνεύονται ως θόρυβος.
- Αν $\sigma_\varepsilon^2/\sigma_\eta^2$ μικρό, τότε οι αποκλίσεις της χρονοσειράς ερμηνεύονται ως πραγματικές μεταβολές του επιπέδου.



Παραδείγματα Local Level σειρών

$$X_t = \mu_t + \varepsilon_t,$$
$$\mu_t = \mu_{t-1} + \eta_t.$$

```
ll.sample.ts = function(sd.eta = 1, sd.epsilon = 1, n = 100){
  mu = numeric(n)
  xt = numeric(n)
  eta.values = rnorm(n, 0, sd.eta)
  epsilon.values = rnorm(n, 0, sd.epsilon)

  mu[1] = eta.values[1]
  for(i in 2:n){
    mu[i] = mu[i-1] + eta.values[i]
  }
  for(i in 2:n){
    xt[i] = mu[i] + epsilon.values[i]
  }
  return(xt)
}

library(latex2exp)
par(mfrow=c(3,2))
for(i in 1:6){
  sd.eta = 30
  sd.epsilon = 0.5
  plot(ll.sample.ts(sd.eta = sd.eta, sd.epsilon = sd.epsilon), type = 'l',
       ylab = TeX('$x_t$'), xlab = 't',
       main = TeX(paste0('$Local \\\, Level \\\, with \\\, \eta_t \\\sim N(0, ', sd.eta, '^2), \varepsilon_t \\\sim N(0, ', sd.epsilon, '^2)$'))))
}
```



Υπολογισμός σ_η , σ_ε στο

$$\begin{aligned}X_t &= \mu_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \eta_t.\end{aligned}$$

Έστω η Local Level σειρά $X_1, X_2, \dots, X_n$.

Από την $X_t = \mu_t + \varepsilon_t$, έχουμε $X_{t-1} = \mu_{t-1} + \varepsilon_{t-1}$, και

$$Y_t = \Delta X_t$$

$$= X_t - X_{t-1}$$

$$= \mu_t - \mu_{t-1} + \varepsilon_t - \varepsilon_{t-1}$$

$$= \eta_t + \varepsilon_t - \varepsilon_{t-1}$$



Υπολογισμός σ_η , σ_ε στο

$$\begin{aligned} X_t &= \mu_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \eta_t. \end{aligned}$$

Βρήκαμε $Y_t = \Delta X_t = \eta_t + \varepsilon_t - \varepsilon_{t-1}$.

Είναι $\varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2)$, $\eta_t \sim \text{iid } N(0, \sigma_\eta^2)$ και ε_t , η_t , ανεξάρτητα μεταξύ τους, άρα:

- $E(Y_t) = 0$.
- $\text{Var}(Y_t) = \sigma_\eta^2 + 2\sigma_\varepsilon^2$.
- $E(Y_t Y_{t-1}) = E[(\eta_t + \varepsilon_t - \varepsilon_{t-1})(\eta_{t-1} + \varepsilon_{t-1} - \varepsilon_{t-2})] = -\sigma_\varepsilon^2$.
- $E(Y_t Y_{t-k}) = E[(\eta_t + \varepsilon_t - \varepsilon_{t-1})(\eta_{t-k} + \varepsilon_{t-k} - \varepsilon_{t-k-1})] = 0$, $k > 1$.
- Αυτοδιασπορά $\gamma(1) = E(Y_t Y_{t-1}) - E(Y_t)E(Y_{t-1}) = -\sigma_\varepsilon^2$ και $\gamma(k) = 0$, $k > 1$.
- Αυτοσυσχέτιση $\rho(1) = -\frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_\varepsilon^2} = -\frac{1}{2+c}$, $c = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$ και $\rho(k) = 0$, $k > 1$.



Υπολογισμός σ_η , σ_ε στο $X_t = \mu_t + \varepsilon_t$,
 $\mu_t = \mu_{t-1} + \eta_t$.

Βρήκαμε ότι $\rho(1) = -\frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_\varepsilon^2} = -\frac{1}{2+c}$, $c = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$ και $\rho(k) = 0$, $k > 1$.

Παρατηρούμε ότι η $Y_t = \Delta X_t$ είναι ισοδύναμη με μία MA(1) χρονοσειρά

$$Y_t = w_t + \theta w_{t-1}, w_t \sim N(0, \sigma_w^2).$$

Γνωρίζουμε ότι $\rho_Y(1) = \theta/(1 + \theta^2)$.

Εξισώνοντας, βρίσκουμε ότι:

$$\theta_{1,2} = \frac{-(2+c) \pm \sqrt{(2+c)^2 - 4}}{2} = -1 - \frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 + 4c}, \quad c = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}.$$

Από τις 2 λύσεις η μία με $|\theta| < 1$, καθιστά το MA(1) μοντέλο αντιστρέψιμο (άρα έχει μοναδική Wold αναπαράσταση):

$$\theta = -1 - \frac{c}{2} + \frac{1}{2}\sqrt{c^2 + 4c}, \quad c = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2}.$$



Υπολογισμός σ_η , σ_ε στο $X_t = \mu_t + \varepsilon_t$, $\mu_t = \mu_{t-1} + \eta_t$.

Με την αρχική υπόθεση $w_0 = 0$, βρίσκουμε επακριβώς την τιμή του θ , ελαχιστοποιώντας με MLE την ποσότητα:

$$\sum_{t=1}^{n-1} w_t^2 = \sum_{t=1}^{n-1} (Y_t - \theta w_{t-1})^2 = Y_1^2 \cdot (Y_2 - \theta w_1)^2 \cdot \dots \cdot (Y_{n-1} - \theta w_{n-2})^2$$

Τώρα, αν $c = \sigma_\eta^2 / \sigma_\varepsilon^2$, είναι

$$\theta = -1 - \frac{c}{2} + \frac{1}{2} \sqrt{c^2 + 4c} \Leftrightarrow c = -\frac{(1+\theta)^2}{\theta}$$

Η επίλυση του συστήματος

$$\frac{\sigma_\eta^2}{\sigma_\varepsilon^2} = -\frac{(1+\theta)^2}{\theta}$$

$$\sigma_\eta^2 + 2\sigma_\varepsilon^2 = \text{Var}(Y_t)$$

οδηγεί στον υπολογισμό των διακυμάνσεων σ_η^2 , σ_ε^2 , του αρχικού local level μοντέλου.



Πρόβλεψη στο $X_t = \mu_t + \varepsilon_t$, $\mu_t = \mu_{t-1} + \eta_t$.

Γνωρίζουμε ότι για οποιαδήποτε τ.μ. Y και οποιοδήποτε χώρο πληροφοριών $F = \sigma(X_1, \dots, X_k)$, είναι:

$$\hat{Y}_{\text{MSE}} = \min_{g \in L^2(F)} E[(Y - g(X_1, X_2, \dots, X_k))^2] = E(Y | X_1, X_2, \dots, X_t)$$

Στο local level model $X_t = \mu_t + \varepsilon_t$, $\mu_t = \mu_{t-1} + \eta_t$, είναι

$$\hat{X}_{t+1} = E(X_{t+1} | X_1, X_2, \dots, X_t) = E(\mu_t + \varepsilon_t | X_1, X_2, \dots, X_t) = \mu_t + E(\varepsilon_t) = \mu_t$$

$$\hat{X}_{t+h} = E(X_{t+h} | X_1, X_2, \dots, X_t) = E(\mu_{t+h} + \varepsilon_{t+h} | X_1, X_2, \dots, X_t) = E(\mu_{t+h}) = \mu_t$$

- Σφάλμα πρόβλεψης για 1 βήμα:

$$\text{Error}(1) = X_{t+1} - \hat{X}_{t+1} = \mu_{t+1} + \varepsilon_{t+1} - \mu_t = \mu_t + \eta_{t+1} + \varepsilon_{t+1} - \mu_t = \eta_{t+1} + \varepsilon_{t+1}$$

$$\text{Var}(\text{Error}(1)) = \sigma_{\eta}^2 + \sigma_{\varepsilon}^2$$

- Σφάλμα πρόβλεψης για h βήματα:

$$\text{Error}(h) = X_{t+h} - \hat{X}_{t+h} = \sum_{j=1}^h \eta_{t+j} + \varepsilon_{t+h}, \quad \text{Var}(\text{Error}(h)) = h\sigma_{\eta}^2 + \sigma_{\varepsilon}^2$$



Πρόβλεψη στο $X_t = \mu_t + \varepsilon_t,$
 $\mu_t = \mu_{t-1} + \eta_t.$

Υπολογισμός μ_t

Γράφουμε $\mu_{t|t-1} = E(\mu_t | X_1, \dots, X_{t-1})$. Αν γνωρίζουμε τα $X_1, \dots, X_{t-1}$, τότε πριν παρατηρήσουμε το X_t :

$$\mu_t | X_1, \dots, X_{t-1} \sim N(\mu_{t-1|t-1}, P_{t|t-1}),$$

όπου $P_{t|t-1} = \text{Var}(\mu_t | X_1, \dots, X_{t-1}) = \text{Var}(\mu_{t-1} + \eta_t | X_1, \dots, X_{t-1}) = P_{t-1|t-1} + \sigma_\eta^2$.

Θεωρούμε την ποσότητα (Kalman gain) $K_t = \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2}$

Η νέα τιμή του επιπέδου ορίζεται ως εξής:

$$\begin{aligned}\mu_{t|t} &= (1 - K_t)\mu_{t-1|t-1} + K_t X_t \\ &= \mu_{t-1|t-1} + K_t(X_t - \mu_{t-1|t-1})\end{aligned}$$

Στη συνέχεια υπολογίζεται η νέα διακύμανση

$$P_{t|t} = (1 - K_t)P_{t|t-1}$$

Η διαδικασία συνεχίζεται επαναληπτικά.

Συνήθης αρχικοποίηση: $P_{0|0}$ πολύ μεγάλο.



Κέρδος Kalman (kalman gain) $K_t = \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2}$

Το κέρδος Kalman K_t είναι η ποσότητα που ελαχιστοποιεί το δεσμευμένο ως προς τα $X_{1:t-1}$ μέσο τετραγωνικό σφάλμα μεταξύ όλων των γραμμικών εκτιμητών της μορφής

$$\hat{\mu}_t = \mu_{t|t-1} + K_t(X_t - \mu_{t|t-1})$$

Είναι:

$$\hat{\mu}_t = \mu_{t|t-1} + K_t(X_t - \mu_{t|t-1}) = \mu_{t|t-1} + K_t(\mu_t + \varepsilon_t - \mu_{t|t-1})$$

$$\mu_t - \hat{\mu}_t = (1 - K_t)(\mu_t - \mu_{t|t-1}) - K_t \varepsilon_t$$

$$\text{Var}(\mu_t - \hat{\mu}_t | X_{1:t-1}) = (1 - K_t)^2 \text{Var}(\mu_t - \mu_{t|t-1} | X_{1:t-1}) + K_t^2 \sigma_\varepsilon^2$$

$$\frac{\partial}{\partial K_t} \text{Var}(\mu_t - \hat{\mu}_t | X_{1:t-1}) = 0$$

$$0 = -2(1 - K_t) \text{Var}(\mu_t - \mu_{t|t-1}) + 2K_t \sigma_\varepsilon^2$$

$$K_t = \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2}, \quad P_{t|t-1} = \text{Var}(\mu_t - \mu_{t|t-1} | X_{1:t-1})$$



Προσαρμογή Local Level μοντέλου

$$X_t = \mu_t + \varepsilon_t,$$
$$\mu_t = \mu_{t-1} + \eta_t.$$

Προσαρμογή μοντέλου και εμφάνιση τυπικών αποκλίσεων σ_ε και σ_η

```
dev.off()
plot.ts(Nile)
fitNile <- StructTS(Nile, type = "level")
print(fitNile$coef)
print(sqrt(fitNile$coef))
```

Έλεγχος υπολοίπων

```
library(forecast)
checkresiduals(fitNile$residuals)
```

Αναπαράσταση προσαρμογής

```
plot(Nile, type = "o")
lines(fitted(fitNile), lty = "dashed", lwd = 2)
```



Προσαρμογή Local Level μοντέλου

$$X_t = \mu_t + \varepsilon_t,$$
$$\mu_t = \mu_{t-1} + \eta_t.$$

Πρόβλεψη h βημάτων μπροστά

```
h <- 20
```

```
fc <- predict(fitNile, n.ahead = h)
```

```
fc$pred # point forecasts
```

```
fc$se # standard errors των forecasts
```

95% διαστήματα εμπιστοσύνης

```
lower <- fc$pred - 1.96 * fc$se
```

```
upper <- fc$pred + 1.96 * fc$se
```

```
cbind(mean = fc$pred, lower = lower, upper = upper)
```

Γραφική αναπαράσταση

```
ts_start <- end(Nile)[1] + (end(Nile)[2] / frequency(Nile))
```

```
pred_ts <- ts(fc$pred, start = end(Nile)[1] + 1, frequency = frequency(Nile))
```

```
low_ts <- ts(lower, start = end(Nile)[1] + 1, frequency = frequency(Nile))
```

```
up_ts <- ts(upper, start = end(Nile)[1] + 1, frequency = frequency(Nile))
```

```
plot(Nile, xlim = c(start(Nile)[1], end(Nile)[1] + h), main = "Nile: Local level  
forecasts")
```

```
lines(pred_ts, lty = 1)
```

```
lines(low_ts, lty = 2)
```

```
lines(up_ts, lty = 2)
```



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Τοπικό Μοντέλο με τάση (Local Level Model with Trend)



Μοντελοποιώντας τη στοχαστική τάση

Έστω, η ευθεία $y = \alpha x$ και τα σημεία $A(x, y_t)$, $B(x+1, y_{t+1})$.

Φανερά, $y_{t+1} = y_t + \alpha$.

Συμπεραίνουμε ότι η εξίσωση $y_{t+1} = y_t + \alpha$ αντιστοιχεί σε χρονοσειρά που ακολουθεί ντετερμινιστική γραμμική τάση με κλίση α .

Αν προστεθεί σφάλμα μέτρησης $\varepsilon_t \sim \text{iid } N(0, \sigma^2)$, τότε προκύπτει το μοντέλο:

$$y_{t+1} = y_t + \alpha + \varepsilon_t$$

$$\Delta y_t = \alpha + \varepsilon_t$$

y_t : ARIMA(0, 1, 0) με drift.



Μοντελοποιώντας τη στοχαστική τάση

Στην περίπτωση όπου η κλίση α μεταβάλλεται, τότε στη θέση του α μπορούμε να θεωρήσουμε μία στοχαστική διαδικασία α_t .

Στην ειδική περίπτωση όπου το α_t είναι Random Walk, τότε η σειρά y_t γράφεται:

$$y_{t+1} = y_t + \alpha_t + \varepsilon_t$$

$$\alpha_t = \alpha_{t-1} + \zeta_t$$

Αν, επιπλέον δεχθούμε ότι υπάρχει και μία τοπική μεταβολή στο επίπεδο της σειράς τότε το μοντέλο γράφεται:

$$y_{t+1} = \mu_{t+1} + \varepsilon_{t+1},$$

$$\mu_{t+1} = \mu_t + \alpha_t + \eta_t,$$

$$\alpha_{t+1} = \alpha_t + \zeta_t,$$



Local Linear Trend

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad (\text{measurement equation})$$

$$\mu_{t+1} = \mu_t + \alpha_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \quad (\text{state equation 1})$$

$$\alpha_{t+1} = \alpha_t + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta^2) \quad (\text{state equation 2})$$

όπου:

y_t : παρατηρούμενη χρονοσειρά

μ_t : τοπικό επίπεδο (local level)

α_t : τοπική κλίση / στοχαστική τάση (local slope)

η_t, ζ_t : καινοτομίες του επιπέδου και της τάσης αντίστοιχα



Προσαρμογή LLT στην R

```
alldata_en_csv = "https://utopia.duth.gr/epdiaman/files/dpth/alldata_en.csv"  
stocks.df <- readr::read_csv(alldata_en_csv)
```

```
# Local Level with Trend (Local Linear Trend)
```

```
fit <- StructTS(stocks.df$FTSE, type = "trend")
```

```
fit
```

```
# Εξομαλυμένες εκτιμήσεις (υπολογίζονται με το Kalman Smoother)
```

```
smoothed <- tsSmooth(fit)
```

```
# Οπτικοποίηση
```

```
plot(stocks.df$FTSE, ylab = "FTSE")
```

```
lines(smoothed[,1], col = "red", lty = 2) # level
```

```
lines(smoothed[,2], col = "blue", lty = 2) # slope (trend)
```

```
legend("topleft", legend = c("Παρατηρήσεις", "Level", "Slope"),  
      col = c("black", "red", "blue"), lty = c(1, 2, 2))
```

```
# Προβλέψεις
```

```
predictions <- predict(fit, n.ahead = 10)
```




ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Τοπικό Μοντέλο με εποχικότητα (Local Level Model with Seasonality)



Μοντελοποιώντας τη στοχαστική εποχικότητα

Έστω, μία περιοδική χρονοσειρά x_t με **περίοδο $T = 4$ και επίπεδο 0** . Η εποχική συνιστώσα έχει άθροισμα 0 σε κάθε πλήρη περίοδο.

Για να περιγράψουμε την περιοδικότητα ορίζουμε 4 βοηθητικές καταστάσεις: $y_{1,t}$, $y_{2,t}$, $y_{3,t}$, $y_{4,t}$, όπου η $y_{1,t}$ είναι η εποχική συνιστώσα που παρατηρείται στο χρόνο t , ενώ οι υπόλοιπες αποθηκεύουν τις επόμενες εποχικές φάσεις.

Το άθροισμά τους είναι ίσο με το επίπεδο 0 :

$$y_{1,t} + y_{2,t} + y_{3,t} + y_{4,t} = 0 \text{ ή ισοδύναμα } y_{4,t} = -y_{1,t} - y_{2,t} - y_{3,t}.$$

Επιπλέον, οι καταστάσεις αυτές πρέπει να επαναλαμβάνονται κυκλικά:

$$y_{1,t+1} = y_{4,t} = -y_{1,t} - y_{2,t} - y_{3,t}$$

$$y_{2,t+1} = y_{1,t}$$

$$y_{3,t+1} = y_{2,t}$$



Μοντελοποιώντας τη στοχαστική εποχικότητα

Συγκεντρωτικά, το σύνολο των εξισώσεων που περιγράφει την εξέλιξη της περιοδικής σειράς x_t με περίοδο $T = 4$ είναι:

$$x_t = y_{1,t} + \varepsilon_t$$

$$y_{1,t+1} = -y_{1,t} - y_{2,t} - y_{3,t}$$

$$y_{2,t+1} = y_{1,t}$$

$$y_{3,t+1} = y_{2,t}$$

Για εποχικότητα περιόδου 4 αρκούν τρία state variables, καθώς η τέταρτη προσδιορίζεται από τον περιορισμό μηδενικού αθροίσματος και δεν χρειάζεται ξεχωριστή εξίσωση.



Μοντελοποιώντας τη στοχαστική εποχικότητα

Χρονοσειρά $x_t = 0, 1, 0, -1, 0, 1, 0, -1, 0, \dots$

$$x_t = y_{1,t}$$

$$y_{1,t+1} = -y_{1,t} - y_{2,t} - y_{3,t}$$

$$y_{2,t+1} = y_{1,t}$$

$$y_{3,t+1} = y_{2,t}$$

t	x_t	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	Άθροισμα
1	0	0	-1	0	1	0
2	1	1	0	-1	0	0
3	0	0	1	0	-1	0
4	-1	-1	0	1	0	0
5	0	0	-1	0	1	0
6	1	1	0	-1	0	0
...	...	...	...	...	...	...

Βοηθητικές states μεταβλητές



Μοντελοποιώντας τη στοχαστική εποχικότητα

Εποχικότητα ($T = 4$)

$$y_t = \gamma_{1,t} + \varepsilon_t,$$

$$\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, \quad \omega_t \sim N(0, \sigma^2_\omega)$$

$$\gamma_{2,t+1} = \gamma_{1,t}$$

$$\gamma_{3,t+1} = \gamma_{2,t}$$

Local Level με Εποχικότητα ($T = 4$)

$$y_t = \mu_t + \gamma_{1,t} + \varepsilon_t,$$

$$\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

$$\mu_{t+1} = \mu_t + \xi_t,$$

$$\xi_t \sim N(0, \sigma^2_\xi)$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, \quad \omega_t \sim N(0, \sigma^2_\omega)$$

$$\gamma_{2,t+1} = \gamma_{1,t},$$

$$\gamma_{3,t+1} = \gamma_{2,t},$$

Local Level με trend και εποχικότητα ($T = 4$)

$$X_t = \mu_t + \gamma_{1,t} + \varepsilon_t,$$

$$\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

$$\mu_{t+1} = \mu_t + u_t + \xi_t,$$

$$\xi_t \sim N(0, \sigma^2_\xi)$$

$$u_{t+1} = u_t + \zeta_t,$$

$$\zeta_t \sim N(0, \sigma^2_\zeta)$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, \quad \omega_t \sim N(0, \sigma^2_\omega)$$

$$\gamma_{2,t+1} = \gamma_{1,t}$$

$$\gamma_{3,t+1} = \gamma_{2,t}$$



Παράδειγμα LL Seasonal στην R

```
set.seed(123)
T <- 120
season <- rep(c(0, 1, 0, -1), length.out = T)
mu <- cumsum(rnorm(T, sd = 0.2))
y <- mu + season + rnorm(T, sd = 0.3)
y_ts <- ts(y, frequency = 4)
plot(y_ts, main = "Χρονοσειρά με Local Level + Εποχικότητα (T=4)")
```

$$y_t = \mu_t + \gamma_{1,t} + \varepsilon_t,$$

$$\mu_{t+1} = \mu_t + \xi_t,$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t,$$

$$\gamma_{2,t+1} = \gamma_{1,t},$$

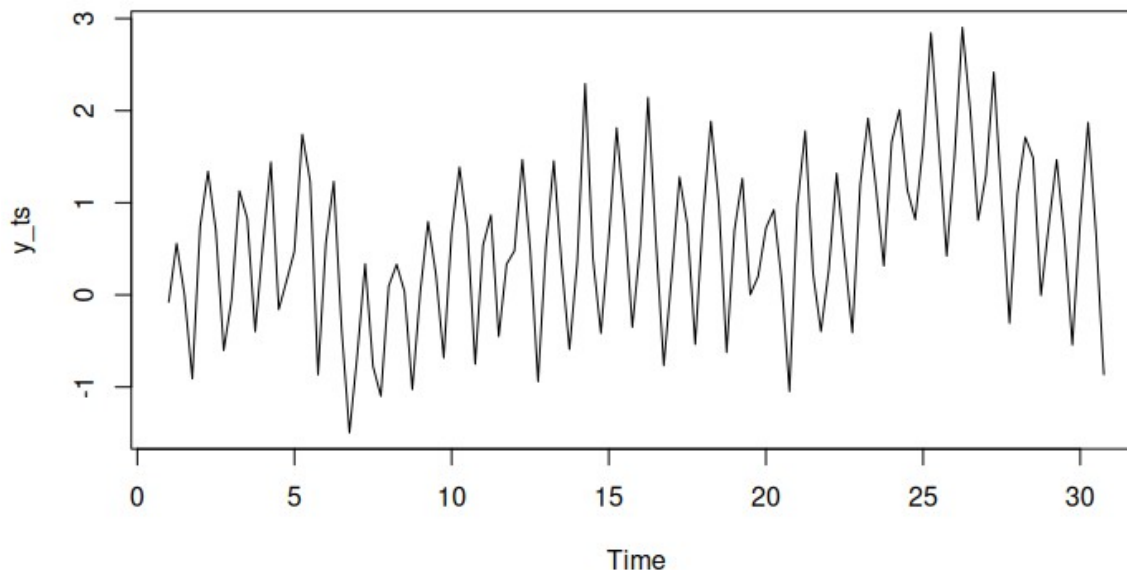
$$\gamma_{3,t+1} = \gamma_{2,t},$$

$$\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

$$\xi_t \sim N(0, \sigma^2_\xi)$$

$$\omega_t \sim N(0, \sigma^2_\omega)$$

Χρονοσειρά με Local Level + Εποχικότητα (T=4)





Παράδειγμα LL Seasonal στην R

```
fit <- StructTS(y_ts, type = "BSM")  
Fit
```

```
> fit <- StructTS(y_ts, type = "BSM")  
> fit  
Call:  
StructTS(x = y_ts, type = "BSM")  
Variances:  
   level    slope    seas  epsilon  
0.02853  0.00000  0.01335  0.10186
```

$$y_t = \mu_t + \gamma_{1,t} + \varepsilon_t,$$

$$\mu_{t+1} = \mu_t + \xi_t,$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t,$$

$$\gamma_{2,t+1} = \gamma_{1,t},$$

$$\gamma_{3,t+1} = \gamma_{2,t},$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad \sigma_\varepsilon^2 = 0.10186.$$

$$\xi_t \sim N(0, \sigma_\xi^2), \quad \sigma_\xi^2 = 0.02853.$$

$$\omega_t \sim N(0, \sigma_\omega^2), \quad \sigma_\omega^2 = 0.01335.$$



Παράδειγμα LL Seasonal στην R

```
fit <- StructTS(y_ts, type = "BSM")  
Fit
```

```
> fit <- StructTS(y_ts, type = "BSM")  
> fit  
Call:  
StructTS(x = y_ts, type = "BSM")  
Variances:  
   level    slope    seas  epsilon  
0.02853 0.00000 0.01335 0.10186
```

$$y_t = \mu_t + \gamma_{1,t} + \varepsilon_t,$$

$$\mu_{t+1} = \mu_t + \xi_t,$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t,$$

$$\gamma_{2,t+1} = \gamma_{1,t},$$

$$\gamma_{3,t+1} = \gamma_{2,t},$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad \sigma_\varepsilon^2 = 0.10186.$$

$$\xi_t \sim N(0, \sigma_\xi^2), \quad \sigma_\xi^2 = 0.02853.$$

$$\omega_t \sim N(0, \sigma_\omega^2), \quad \sigma_\omega^2 = 0.01335.$$



Παράδειγμα LL Seasonal στην R

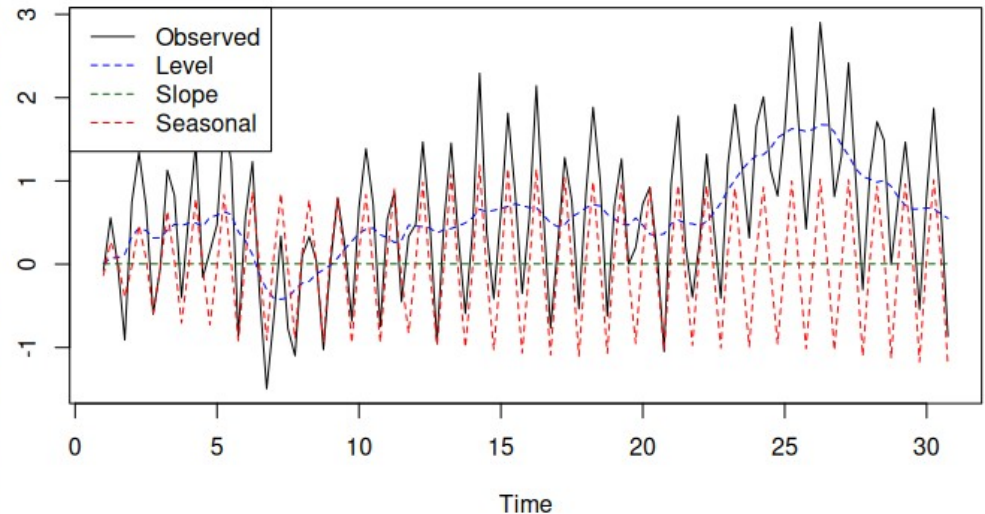
```
comp <- tsSmooth(fit)
plot(cbind(y_ts,
           level = comp[, "level"],
           slope = comp[, "slope"],
           sea    = comp[, "sea"]),
     plot.type = "single",
     col = c("black", "blue", "darkgreen", "red"),
     lty = c(1, 2, 2, 2),
     main = "Observed, Level, Slope, Seasonal (Kalman smoothing)")

legend("topleft",
      legend = c("Observed", "Level", "Slope", "Seasonal"),
      col = c("black", "blue", "darkgreen", "red"),
      lty = c(1, 2, 2, 2))
```

$$\begin{aligned}y_t &= \mu_t + \gamma_{1,t} + \varepsilon_t, \\ \mu_{t+1} &= \mu_t + \xi_t, \\ \gamma_{1,t+1} &= -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, \\ \gamma_{2,t+1} &= \gamma_{1,t}, \\ \gamma_{3,t+1} &= \gamma_{2,t},\end{aligned}$$

$$\begin{aligned}\varepsilon_t &\sim N(0, \sigma_{\varepsilon}^2) & \sigma_{\varepsilon}^2 &= 0.10186. \\ \xi_t &\sim N(0, \sigma_{\xi}^2) & \sigma_{\xi}^2 &= 0.02853. \\ \omega_t &\sim N(0, \sigma_{\omega}^2) & \sigma_{\omega}^2 &= 0.01335.\end{aligned}$$

Observed, Level, Slope, Seasonal (Kalman smoothing)





ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

State Space Models



Μοντέλα Καταστάσεως – Χώρου (SSM) Χρονοσειρά μίας μεταβλητής

Τα μοντέλα καταστάσεως–χώρου (state space models) βασίζονται στην υπόθεση ότι η παρατηρούμενη χρονοσειρά X_t , στη χρονική στιγμή t , μπορεί να περιγραφεί ως γραμμική συνάρτηση ενός διανύσματος λανθανουσών μεταβλητών (states), οι οποίες εξελίσσονται δυναμικά στο χρόνο σύμφωνα με μία γραμμική στοχαστική εξίσωση.

$$X_t = F_t \alpha_t + \varepsilon_t, \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2_\varepsilon),$$

Όπου:

X_t : η παρατηρούμενη τιμή της χρονοσειράς στη χρονική στιγμή t .

α_t : το $m \times 1$ διάνυσμα των λανθανουσών καταστάσεων στη χρονική στιγμή t .

F_t : $1 \times m$ πίνακας συντελεστών των state μεταβλητών α_t

ε_t : τα υπόλοιπα του μοντέλου.



Μοντέλα Καταστάσεως – Χώρου (SSM) Χρονοσειρά μίας μεταβλητής

Η εξίσωση $X_t = F_t \alpha_t + \varepsilon_t$, $\varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2)$, αποτελεί το measurement equation του συστήματος. Επιπλέον, θεωρούμε και την state (ή transition) εξίσωση που περιγράφει τον τρόπο που εξελίσσονται στο χρόνο οι καταστάσεις α_t .

$$\alpha_{t+1} = G_t \alpha_t + R_t \eta_t, \eta_t \sim \text{i.i.d. } N(0, W_\eta),$$

Όπου:

α_{t+1} : οι state μεταβλητές του μοντέλου τη χρονική στιγμή $t + 1$ (πίνακας $m \times 1$).

G_t : $m \times m$ πίνακας μετάβασης των α_t .

η_t : τα g ($\leq m$) υπόλοιπα του μοντέλου ως πίνακας $g \times 1$.

R_t : ο $g \times g$ πίνακας συντελεστών των υπολοίπων η_t .

W_η : ο $g \times g$ πίνακας συνδιακύμανσης των υπολοίπων η_t .



Μοντέλα Καταστάσεως – Χώρου (SSM) Χρονοσειρά μίας μεταβλητής

Οι δύο εξισώσεις

- $X_t = F_t \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2),$ (measurement equation)
- $\alpha_{t+1} = G_t \alpha_t + R_t \eta_t, \quad \eta_t \sim \text{i.i.d. } N(0, W_\eta),$ (state equation)

προσδιορίζουν πλήρως την εξέλιξη του μοντέλου.

Τα σφάλματα ε_t, η_t θεωρούνται ανεξάρτητα μεταξύ τους και ανεξάρτητα με την αρχική κατάσταση του συστήματος α_0 .

Η πλήρης περιγραφή του μοντέλου ολοκληρώνεται με την υπόθεση κατανομής της αρχικής κατάστασης $\alpha_0 \sim N(\mu_0, P_0)$, όπου μ_0 και P_0 είναι αντίστοιχα η μέση τιμή και ο πίνακας συνδιακύμανσης της αρχικής κατάστασης.

Στα δομικά μοντέλα χρονοσειρών, η πρόβλεψη βασίζεται αποκλειστικά στην εκτίμηση των διακυμάνσεων που ελέγχουν την εξέλιξη των καταστάσεων.



Μοντέλα Καταστάσεως – Χώρου (SSM)

$$\begin{aligned} X_t &= F_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim \text{i.i.d. } N(0, \sigma_\varepsilon^2), & \text{(measurement equation)} \\ \alpha_{t+1} &= G_t \alpha_t + R_t \eta_t, & \eta_t &\sim \text{i.i.d. } N(0, W_\eta), & \text{(state equation)} \end{aligned}$$

Παράδειγμα I (local level model)

Υποθέτουμε ότι η χρονοσειρά X_t έχει τάση μ_t και δεν έχει εποχικότητα. Στην περίπτωση αυτή μπορούμε να γράψουμε:

$$X_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad \text{(measurement equation)}$$

$$\mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \quad \text{(state equation)}$$

Τότε:

$$\alpha_t = [\mu_t] \text{ (ένα μοναδικό state)}$$

$$F_t = G_t = R_t = [1],$$

$$V = [\sigma_\varepsilon^2],$$

$$W_\eta = [\sigma_\eta^2].$$



Μοντέλα Καταστάσεως – Χώρου (SSM)

$$\begin{aligned} X_t &= F_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim \text{i.i.d. } N(0, \sigma_\varepsilon^2), & \text{(measurement equation)} \\ \alpha_{t+1} &= G_t \alpha_t + R_t \eta_t, & \eta_t &\sim \text{i.i.d. } N(0, W_\eta), & \text{(state equation)} \end{aligned}$$

Παράδειγμα II (local level trend model)

Υποθέτουμε ότι η χρονοσειρά X_t έχει τάση μ_t με μεταβαλλόμενη κλίση.

$$X_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad \text{(measurement equation)}$$

$$\mu_{t+1} = \mu_t + u_t + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2) \quad \text{(state equation 1)}$$

$$u_{t+1} = u_t + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta^2) \quad \text{(state equation 2)}$$

Παρατηρούμε, ότι οι δύο τελευταίες εξισώσεις γράφονται:

$$\begin{pmatrix} \mu_{t+1} \\ u_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mu_t \\ u_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix} \quad \text{ή} \quad \alpha_{t+1} = G_t \alpha_t + R_t \eta_t,$$

$$F_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \alpha_t = \begin{pmatrix} \mu_t \\ u_t \end{pmatrix}, \quad R_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}, \quad W_\eta = \begin{pmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{pmatrix}.$$



Μοντέλα Καταστάσεως – Χώρου (SSM)

$$\begin{aligned} X_t &= F_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim \text{i.i.d. } N(0, \sigma_\varepsilon^2), & \text{(measurement equation)} \\ \alpha_{t+1} &= G_t \alpha_t + R_t \eta_t, & \eta_t &\sim \text{i.i.d. } N(0, W_\eta), & \text{(state equation)} \end{aligned}$$

Παράδειγμα III (local level trend model με εποχικότητα)

Υποθέτουμε ότι η χρονοσειρά y_t έχει τάση μ_t με μεταβαλλόμενη κλίση και εποχικότητα περιόδου 4.

$$X_t = \mu_t + \gamma_{1,t} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad \text{(measurement equation)}$$

$$\mu_{t+1} = \mu_t + u_t + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2) \quad \text{(state equation 1)}$$

$$u_{t+1} = u_t + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta^2) \quad \text{(state equation 2)}$$

$$\gamma_{1,t+1} = -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, \quad \omega_t \sim N(0, \sigma_\omega^2) \quad \text{(state equation 3)}$$

$$\gamma_{2,t+1} = \gamma_{1,t} \quad \text{(state equation 4)}$$

$$\gamma_{3,t+1} = \gamma_{2,t} \quad \text{(state equation 5)}$$

Το μοντέλο είναι της μορφής

ME: $X_t = F_t \alpha_t + \varepsilon_t$, $\varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2)$, SE: $\alpha_{t+1} = G_t \alpha_t + R_t \eta_t$, $\eta_t \sim \text{i.i.d. } N(0, W)$, για

$$F_t = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad G_t = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \alpha_t = \begin{pmatrix} \mu_t \\ u_t \\ \gamma_{1,t} \\ \gamma_{2,t} \\ \gamma_{3,t} \end{pmatrix}, \quad R_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \zeta_t \\ \omega_t \end{pmatrix}, \quad W_\eta = \begin{pmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \sigma_\omega^2 \end{pmatrix}.$$