

Χρονικές σειρές

**4^ο μάθημα: Μη στάσιμες χρονοσειρές
Μετασχηματισμός σε στάσιμες**

**Εαρινό εξάμηνο 2018-2019
Τμήμα Μαθηματικών ΑΠΘ**

Διδάσκουσα: Αγγελική Παπάνα

Μεταδιδακτορική Ερευνήτρια

Πολυτεχνική σχολή, Α.Π.Θ. & Οικονομικό Τμήμα, Πανεπιστήμιο Μακεδονίας

<http://users.auth.gr/~agrapana/>

ΜΗ ΣΤΑΣΙΜΕΣ ΧΡΟΝΟΣΕΙΡΕΣ

ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ ΜΗ-ΣΤΑΣΙΜΗΣ ΣΕ ΣΤΑΣΙΜΗ ΧΡΟΝΟΣΕΙΡΑ

Είναι συχνό φαινόμενο στις χρονοσειρές η μέση τιμή τους να παρουσιάζει μια αυξητική, ή φθίνουσα, τάση ή/και να έχει εναλλαγές μεταξύ αυξητικών φάσεων και φθινουσών φάσεων, να παρουσιάζει δηλαδή μια “κυκλικά” επαναλαμβανόμενη δομή σε διαδοχικά χρονικά διαστήματα ή εποχές. Τέτοια συμπεριφορά είναι ιδιαίτερα εμφανής σε μετεωρολογικές χρονοσειρές καθώς και σε εμπορικά και οικονομικά μεγέθη.

Πριν προχωρήσουμε σε οποιαδήποτε προσπάθεια μοντελοποίησης μιας χρονοσειράς είναι απαραίτητο να διερευνήσουμε αν εμφανίζει τέτοιου είδους συμπεριφορά και γενικά αν παραβιάζεται η απαίτηση της, υπό ευρεία έννοια, στασιμότητας.

Επίσης, από τη γραφική παράσταση μιας χρονοσειράς είναι δυνατόν να διαπιστώσουμε αν υπάρχουν ακραίες τιμές. Οι τιμές αυτές ενδέχεται να δημιουργήσουν σοβαρά προβλήματα στην μοντελοποίηση μιας χρονοσειράς και ως εκ τούτου χρειάζονται ειδική μεταχείριση αφού όμως πρώτα προσδιοριστεί το αίτιο το οποίο τις προκάλεσε.

Υπάρχουν συγκεκριμένες τεχνικές καθώς και στατιστικοί έλεγχοι για να διερευνήσουμε τη στασιμότητα σε μια χρονοσειρά αλλά θα τις δούμε σε επόμενη ενότητα.

Εδώ θα δούμε όμως κάποιους βασικούς μετασχηματισμούς που εφαρμόζουμε σε μια μη-στάσιμη χρονοσειρά για να την κάνουμε στάσιμη και να προχωρήσουμε με την ανάλυση.

Γενικά μια χρονοσειρά μπορούμε να τη χωρίσουμε σε τρεις συνιστώσες ως εξής:

$$Y_t = \mu_t + s_t + X_t$$

όπου

- μ_t είναι η συνιστώσα της τάσης
- s_t η συνιστώσα της περιοδικότητας για κάποια περίοδο d και ισχύει $s_{t-d} = s_t$ και
- X_t είναι η χρονοσειρά των υπολοίπων αν αφαιρέσουμε από την παρατηρούμενη χρονοσειρά την τάση και την περιοδικότητα, που συγκεντρώνει την πληροφορία (αν υπάρχει) για τη δυναμική του συστήματος απαλλαγμένο από τάση και περιοδικότητα.

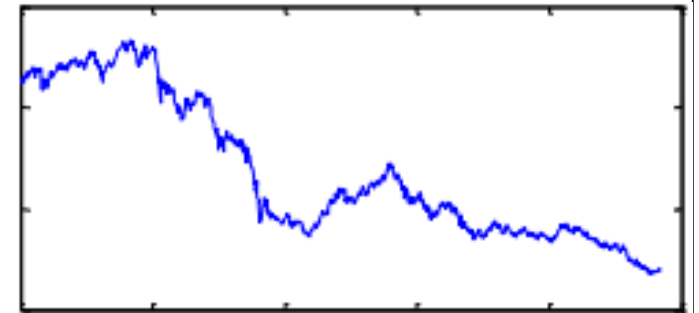
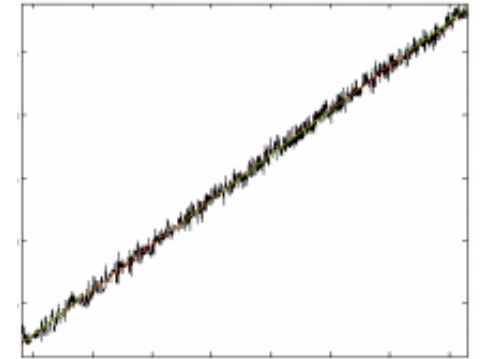
Παρατηρήσεις

- Η συνιστώσα της τάσης μ_t είναι μια χαμηλών μεταβολών συνάρτηση του χρόνου t .
- Η συνιστώσα της περιοδικότητας s_t είναι μια περιοδική συνάρτηση η οποία εκφράζει την εποχική συνιστώσα
- Η συνιστώσα X_t αποτελεί τον θόρυβο και είναι μια στάσιμη στοχαστική ανέλιξη.
- Η τάση και η περιοδικότητα είναι και οι δυο συναρτήσεις του χρόνου και δεν περιέχουν πληροφορία για τη δυναμική του συστήματος, δηλαδή την εξάρτηση της παρατήρησης x_t από τις προηγούμενες παρατηρήσεις.

- Σε πολλά βιβλία θεωρούν την εποχικότητα και περιοδικότητα ως δύο διαφορετικές συνιστώσες, η πρώτη με γνωστή περίοδο (μέρα, μήνας, εποχή) ενώ για τη δεύτερη η περίοδος πρέπει να εκτιμηθεί.
- Σε συγκεκριμένα προβλήματα, όλη η πληροφορία που μας ενδιαφέρει μπορεί να εντοπίζεται στη συνιστώσα της τάσης και της περιοδικότητας, οπότε το πρόβλημα περιορίζεται στην εκτίμηση των μ_t και s_t .

ΤΑΣΗ

- Όταν η **τάση** σε μια χρονοσειρά μπορεί να περιγραφεί από κάποια γνωστή ή εκτιμώμενη συνάρτηση του χρόνου, $\mu_t = f(t)$, τότε ονομάζεται **καθοριστική τάση** (deterministic trend).
- Αν η τάση σε μια χρονοσειρά δεν είναι δυνατόν να περιγραφεί από μια γνωστή (παραμετρική) συνάρτηση του χρόνου (π.χ. αργές μεταβολές με το χρόνο αλλά όχι με κάποιο καθοριστικό τρόπο), τότε ονομάζεται **στοχαστική τάση** (stochastic trend).



Ας υποθέσουμε πως η μη-στασιμότητα της χρονοσειράς οφείλεται αποκλειστικά στην **ύπαρξη τάσης**, και έστω ότι Y_t δίνεται ως

$$Y_t = \mu_t + X_t$$

□ Αν η τάση στη χρονοσειρά είναι **καθοριστική** μπορούμε να την εκτιμήσουμε και στη συνέχεια να την απαλείψουμε (detrend) με κάποια παραμετρική συνάρτηση $f(t)$ του χρόνου.

□ Συνήθως η τάση είναι **γραμμική** $\mu_t = \alpha + \beta t$

ή **εκθετική** $\mu_t = \alpha \exp\{\beta t\}$

ή ένα **πολυώνυμο** του t βαθμού p : $\mu_t = a_0 + a_1 t + \dots + a_p t^p$

Παρατήρηση

Αν η τάση δε φαίνεται να είναι κάποια γνωστή συνάρτηση του χρόνου, τότε μπορούμε να **προσαρμόσουμε τη συνάρτηση $f(t)$ τοπικά**, σε συνεχόμενα τμήματα της χρονοσειράς.

ΕΚΤΙΜΗΣΗ ΤΑΣΗΣ

Υπάρχουν τρεις βασικές μέθοδοι προσδιορισμού της τάσης:

- **Μέθοδος ελαχίστων τετραγώνων**
- **Μέθοδος των διαφορών**
- **Μέθοδος εξομάλυνσης κινητού μέσου**

Μέθοδος ελαχίστων τετραγώνων

Έχοντας παρατηρήσει τη χρονοσειρά $\{Y_t\}$ για χρόνο T , έχοντας δηλαδή τις τιμές $\{y_t\}$, $t = 1, \dots, T$ και διαπιστώνοντας από τη γραφική παράστασή της ότι η τάση είναι ένα χαμηλού βαθμού πολυώνυμο του χρόνου t , μπορούμε με εφαρμογή της μεθόδου ελαχίστων τετραγώνων να προσδιορίσουμε την τάση αυτή.

Έστω ότι έχουμε γραμμική τάση, δηλαδή $\mu_t = \alpha + \beta t$. Η μέθοδος ελαχίστων τετραγώνων ελαχιστοποιεί το άθροισμα:

$$\sum_{t=1}^T (y_t - \alpha - \beta t)^2$$

και μας δίνει μια εκτίμηση της τάσης

$$\hat{\mu}_t = \hat{a} + \hat{\beta}t$$

με

$$\hat{a} = \bar{y} - \hat{\beta}\bar{t}$$

$$\hat{\beta} = C_{ty}/C_{tt}$$

όπου

$$\bar{y} = \frac{1}{T} \sum y_t$$

$$\bar{t} = \frac{1}{T} \sum t = \frac{T+1}{2}$$

$$C_{ty} = \sum (t - \bar{t})(y_t - \bar{y})$$

$$C_{tt} = \sum (t - \bar{t})^2 = \frac{T(T^2 - 1)}{12}$$

Παρατηρήσεις

- Οι ίδιοι τύποι ισχύουν για το λογάριθμο της εκθετικής τάσης.
- Επίσης, ανάλογα αποτελέσματα προκύπτουν για πολυωνυμικές τάσεις.

Παράδειγμα 1

Δίνεται η
χρονοσειρά του
πλήθους των
νέων οδηγών
που πήραν το
δίπλωμά τους
σε μία περιοχή
για ένα πλήθος
συνεχόμενων
ετών.

t	y_t	$\hat{\mu}_t$	t	y_t	$\hat{\mu}_t$	t	y_t	$\hat{\mu}_t$
1	6577	8056.8	13	10409	8881.1	25	10118	9705.5
2	5855	8125.5	14	11351	8949.8	26	10889	9774.2
3	6939	8194.2	15	8701	9018.5	27	11140	9842.9
4	7557	8262.9	16	8168	9087.2	28	10183	9911.6
5	8065	8331.6	17	9752	9155.9	29	10398	9980.3
6	9314	8400.3	18	10826	9224.6	30	9833	10049.0
7	9009	8469.0	19	10946	9293.3	31	9160	10117.7
8	8357	8537.6	20	10357	9362.0	32	9234	10186.3
9	9404	8606.3	21	8761	9430.7	33	8054	10255.0
10	9447	8675.0	22	8444	9499.4			
11	8388	8743.7	23	7754	9568.1			
12	9831	8812.4	24	8924	9636.8			

Είναι:

$$\bar{y} = \frac{1}{T} \sum_t y_t = \frac{1}{33} (6577 + 5855 + \dots + 8054) = 9155,9$$

$$\bar{t} = \frac{1}{T} \sum_t t = \frac{T + 1}{2} = \frac{33 + 1}{2} = 17$$

$$C_{ty} = \sum_t (t - \bar{t})(y_t - \bar{y}) = (1 - 17)(6577 - 9155,9) + \dots = 205538$$

$$C_{tt} = \sum_t (t - \bar{t})^2 = \frac{T(T^2 - 1)}{12} = 2992$$

οπότε

$$\hat{\beta} = \frac{C_{ty}}{C_{tt}} = 68.69$$

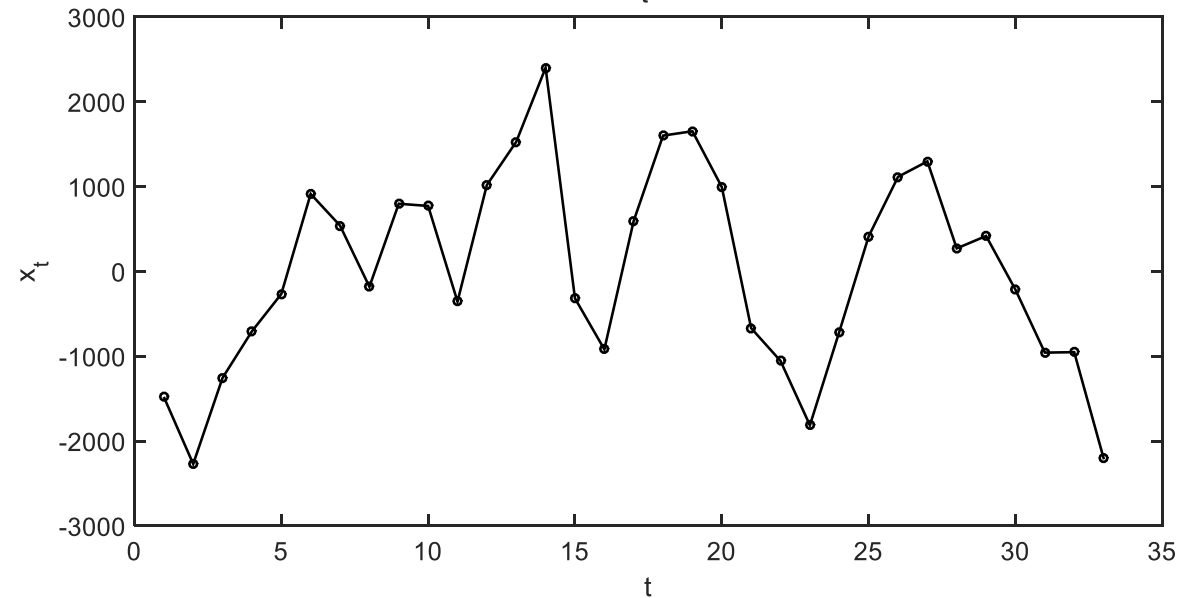
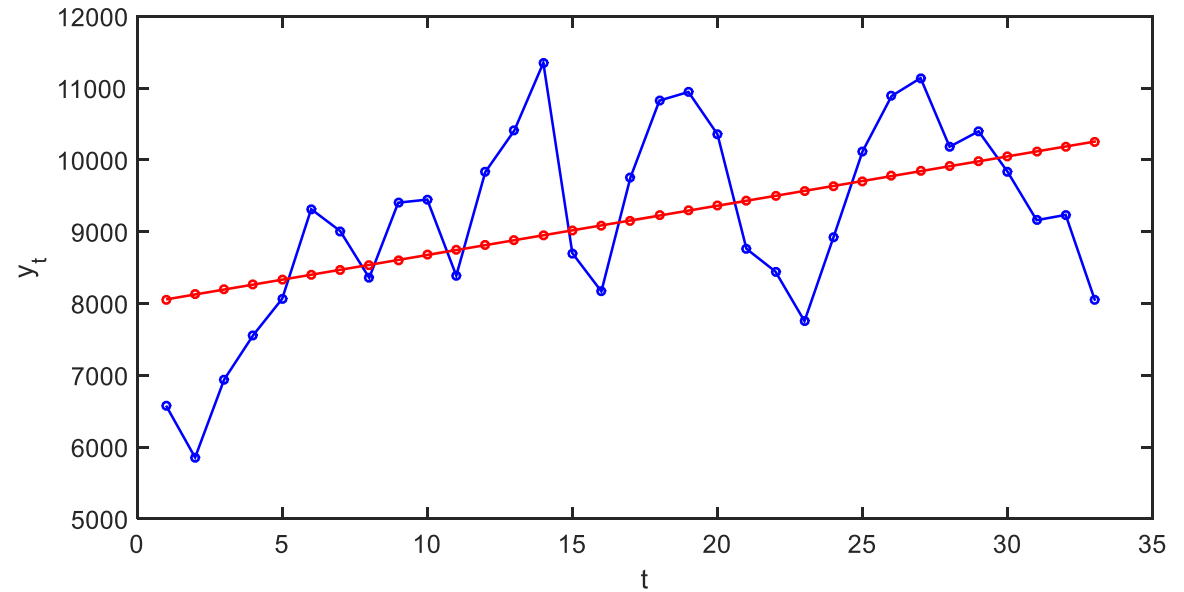
$$\hat{a} = \bar{y} - \hat{\beta}\bar{t} = 7988.1$$

Οπότε η εκτιμώμενη
ευθεία της τάσης είναι

$$\hat{\mu}_t = \hat{a} + \hat{\beta}t$$

$$\hat{\mu}_t = 7988.1 + 68.69 t$$

Η εκτίμηση της κλίσης για την
τάση δηλώνει ότι υπάρχει κατά
μέσο όρο ετήσια αύξηση κατά
68.7 νέων οδηγών.



ΕΠΙΛΥΣΗ ΜΕ ΤΟ ΜΑΤΛΑΒ

Εισαγωγή δεδομένων

```
>> yV2 = [6577 5855 6939 7557 8065 9314 9009 8357 9404 ...  
9447 8388 9831 10409 11351 8701 8168 9752 10826 ...  
10946 10357 8761 8444 7754 8924 10118 10889 11140 ...  
10183 10398 9833 9160 9234 8054]';
```

ή

```
>> cd c:\bin\  
>> load yV2.dat  
>> T = length(yV2);
```

Γραφική παράσταση

```
>> figure(1)  
>> plot(yV2,'.-')  
>> xlabel('t')  
>> ylabel('yV2(t)')
```

```
>> my = mean(yV2)
>> mt = (T+1)/2
>> for t=1:T
    CtyV(t) = (t - mt)*(yV(t)-my);
end
>> Cty = sum(CtyV)

>> Ctt = T*(T^2-1)/12

>> b = Cty / Ctt
>> a = my - b*mt
```

Εκτίμηση τάσης

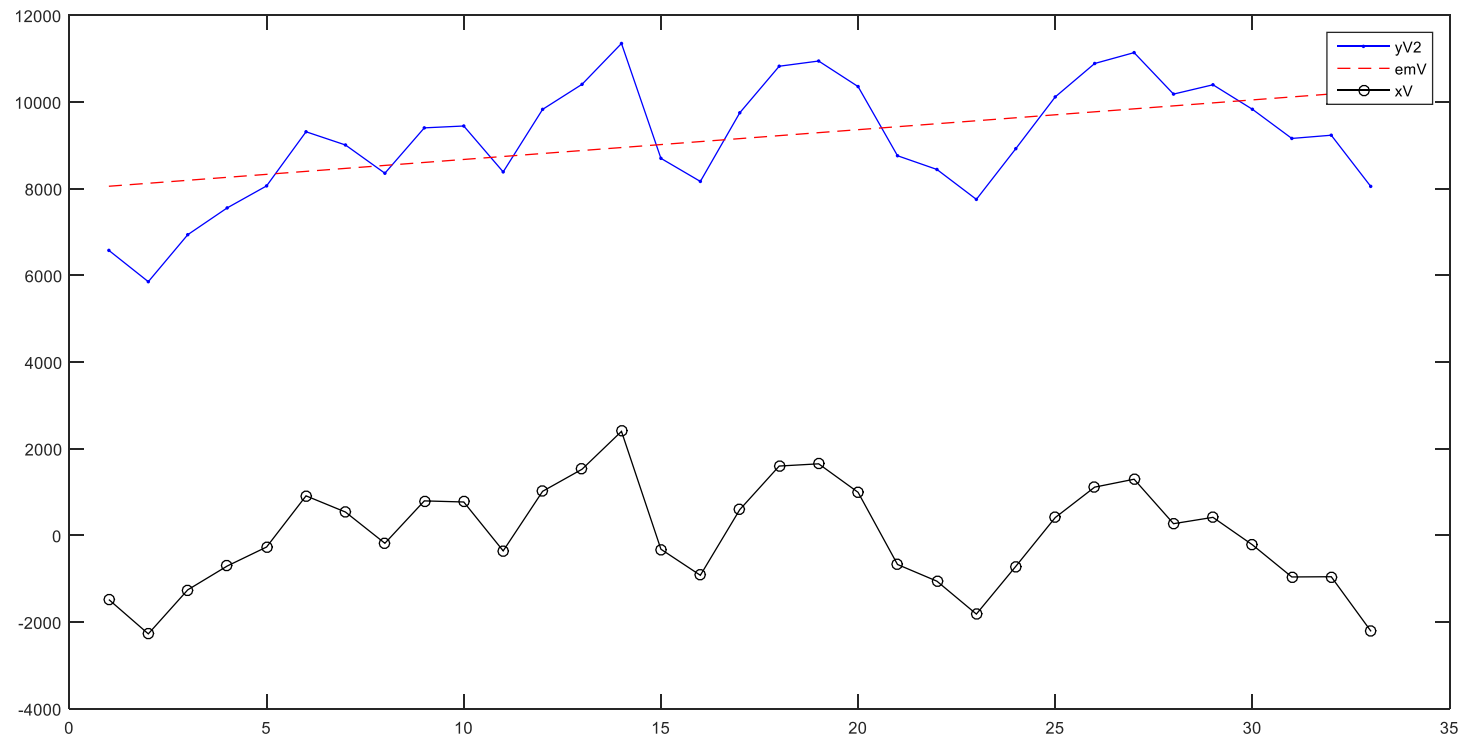
```
>> emV = NaN*ones(T,1);
>> for t=1:T
    emV(t,1) = a + b*t;
end
```

Γραφική παράσταση τάσης

```
>> figure(1)
>> plot(emV,'.-')
```

Γραφική παράσταση χρονοσειρών

- Μη στάσιμη χρονοσειρά: $yV2$
- Γραμμική τάση: emV
- Στάσιμη χρονοσειρά: xV



```
>> xV = yV2 - emV;
```

```
>> figure(1)
```

```
>> plot(yV2,'b.-')
```

```
>> hold on
```

```
>> plot(emV,'r--')
```

```
>> hold on
```

```
>> plot(xV,'ko-')
```

```
>>
```

```
legend('yV2','emV','xV')
```

Προσαρμογή πολυωνύμου μεγαλύτερου βαθμού

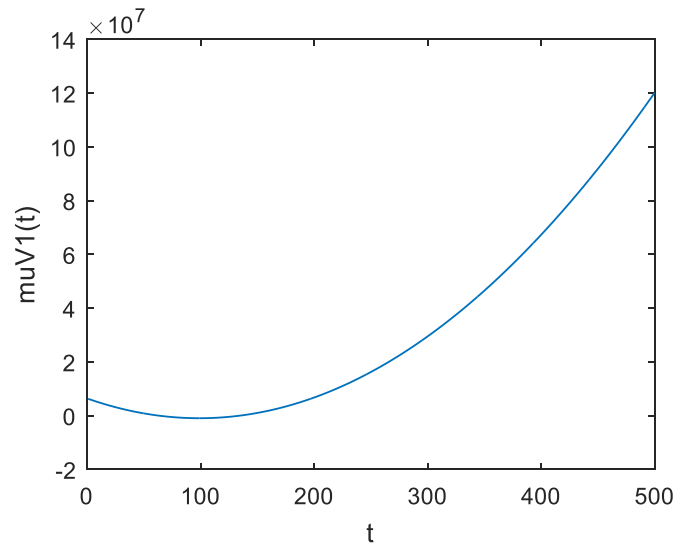
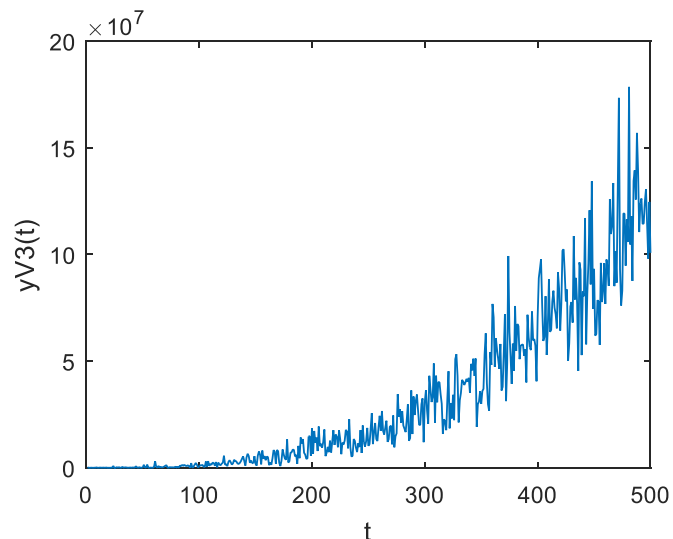
Συνάρτηση [polynomialfit.m](#) : προσαρμογή τάσης με πολώνυμο δεδομένου βαθμού. Π.χ.

```
>> load yV3.dat
```

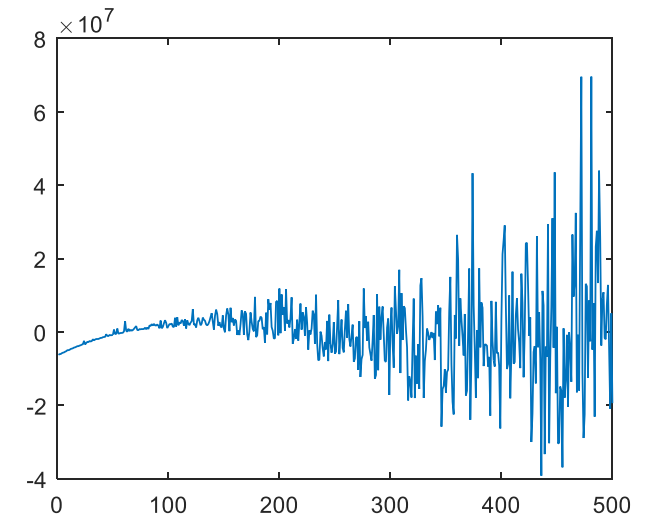
```
>> muV1 = polynomialfit(yV3,3);
```

```
>> plot(muV1)
```

```
>> xV3 = yV3 - muV1;
```



Στάσιμη χρονοσειρά



Μέθοδος των διαφορών

Με τη μέθοδο των αυτή επιδιώκουμε να εξαλείψουμε την τάση μ_t που υπάρχει σε μια χρονοσειρά y_t σχηματίζοντας μια νέα χρονοσειρά από τις διαφορές μεταξύ των διαδοχικών όρων $y_t - y_{t-1}$ για $t = 2, 3, \dots, T$.

Έτσι, όταν η τάση μ_t είναι γραμμική, η χρονοσειρά $x_t = y_t - y_{t-1}$ που παράγεται έχει μηδενική τάση.

Όταν η τάση είναι πολυωνυμική, η διαδικασία των διαφορών μεταξύ διαδοχικών όρων επαναλαμβάνεται μέχρις ότου εξαλειφθεί η τάση πλήρως.

Με την διαδικασία αυτή μπορούμε να απαλείψουμε αλλά όχι να εκτιμήσουμε) την τάση.

Συγκεκριμένα:

- Ο μετασχηματισμός **πρώτων διαφορών** είναι:

$$X_t = \nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t$$

όπου ∇ δηλώνει τον τελεστή της διαφοράς πρώτης τάξης και B είναι ο τελεστής υστέρησης (lag operator), δηλ.

$$BY_t = Y_{t-1}$$

$$\nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t$$

- Αν η χρονοσειρά δε γίνει στάσιμη μπορούμε να πάρουμε πάλι πρώτες διαφορές και αυτός είναι ο μετασχηματισμός **διαφορών δεύτερης τάξης** (second differences):

$$\begin{aligned} X_t &= \nabla^2 Y_t = \nabla(\nabla Y_t) = (1 - B)(1 - B)Y_t \\ &= (1 - 2B + B^2)Y_t = Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned}$$

Γενικά στην πράξη, συνήθως χρησιμοποιούμε **πρώτες διαφορές** γιατί στις περισσότερες περιπτώσεις οι τοπικές τάσεις μπορεί να προσεγγιστούν ικανοποιητικά από γραμμικά πολυώνυμα.

Πράγματι αν η τάση είναι **τοπικά γραμμική**, δηλαδή ισχύει $\mu_t = \alpha_0 + \alpha_1 t$ τότε απαλείφεται με τις πρώτες διαφορές. Ο μετασχηματισμός των πρώτων διαφορών δίνει:

$$\nabla Y_t = Y_t - Y_{t-1} = \mu_t + X_t - (X_{t-1} + \mu_{t-1}) = (\mu_t - \mu_{t-1}) + X_t - X_{t-1}$$

όπου η τάση στην νέα χρονοσειρά είναι

$$\mu_t - \mu_{t-1} = (\alpha_0 + \alpha_1 t) - (\alpha_0 + \alpha_1 (t - 1)) = \alpha_1$$

δηλαδή σταθερή ως προς το χρόνο και άρα έχει απαλειφθεί.

Παρατήρηση

Γενικά αν η τάση εκφράζεται τοπικά με πολυώνυμο βαθμού p , απαλείφεται με χρήση **διαφορών τάξης p** ($\nabla^p \gamma_t$).

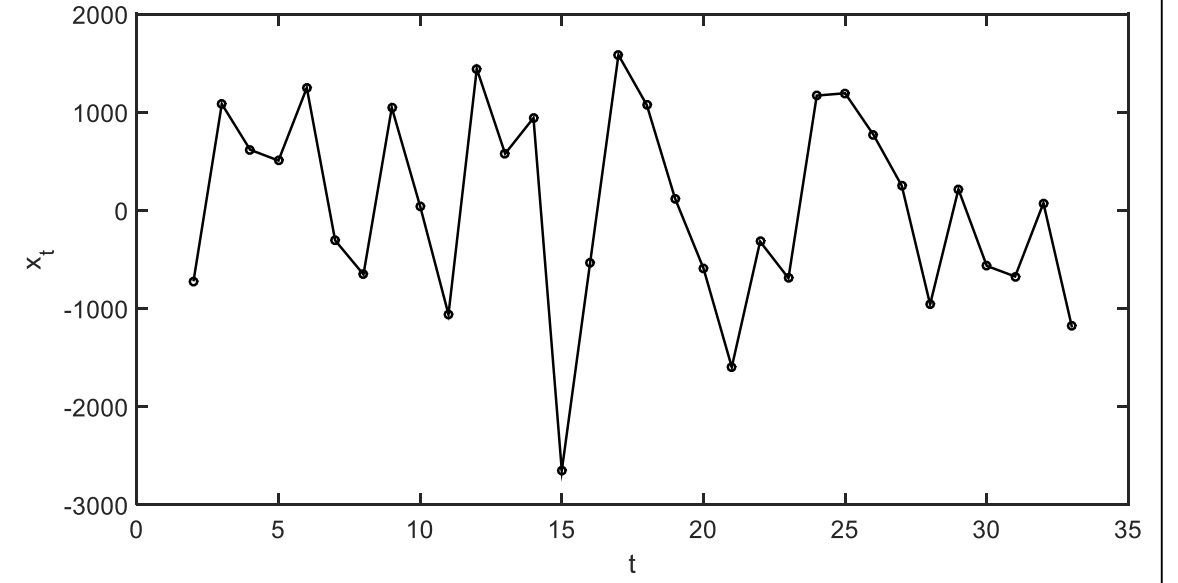
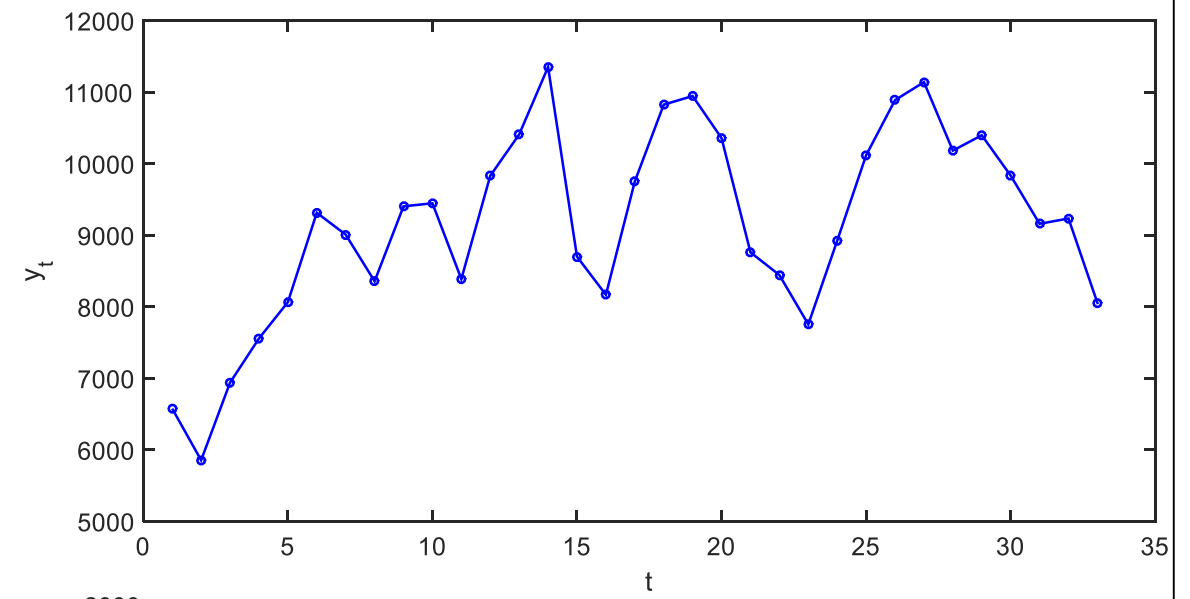
Παράδειγμα 2

Να βρεθεί η στάσιμη χρονοσειρά με την μέθοδο πρώτων διαφορών.

t	y_t	t	y_t	t	y_t
1	6577	13	10409	25	10118
2	5855	14	11351	26	10889
3	6939	15	8701	27	11140
4	7557	16	8168	28	10183
5	8065	17	9752	29	10398
6	9314	18	10826	30	9833
7	9009	19	10946	31	9160
8	8357	20	10357	32	9234
9	9404	21	8761	33	8054
10	9447	22	8444		
11	8388	23	7754		
12	9831	24	8924		

Για $t = 2, \dots, T$:

$$x_t = y_t - y_{t-1}$$



ΕΠΙΛΥΣΗ ΜΕ ΤΟ MATLAB

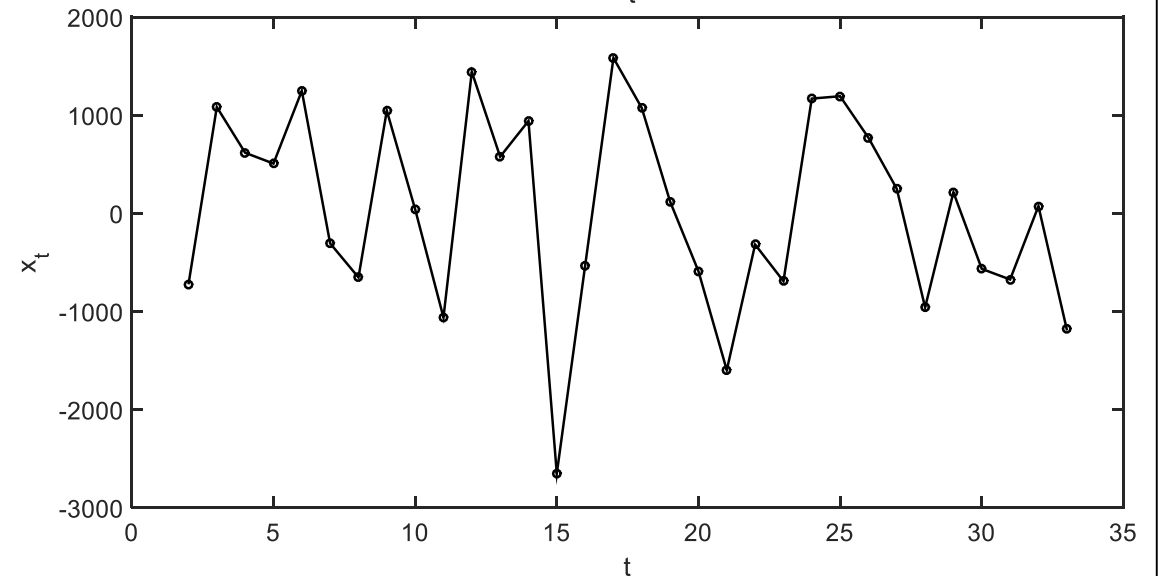
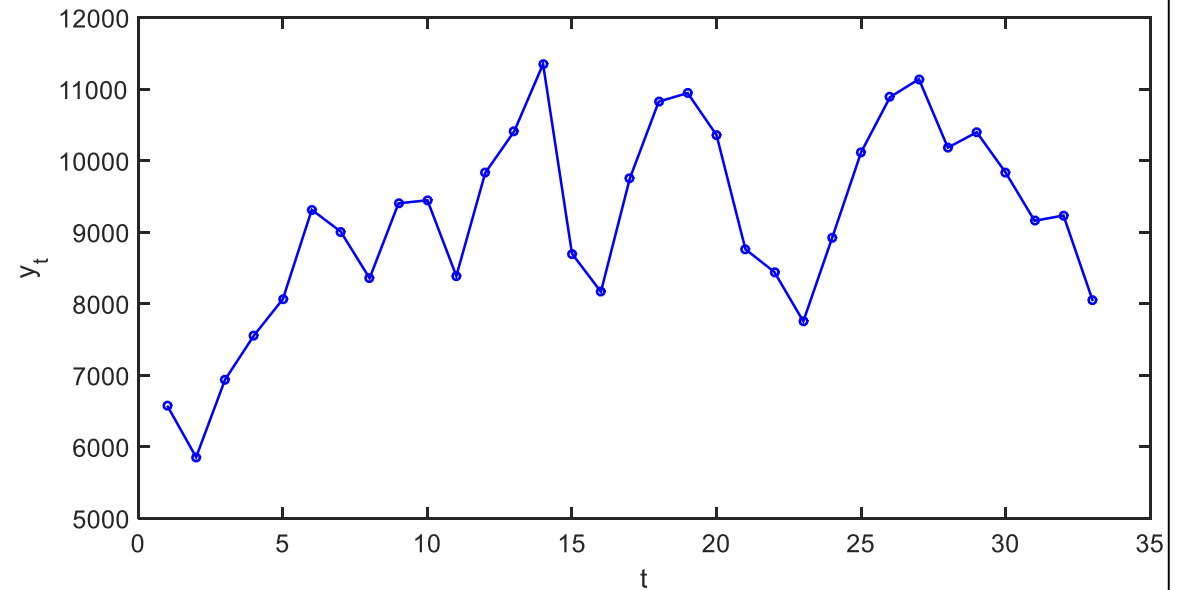
Κατεβάστε το αρχείο με τα δεδομένα της χρονοσειράς (yV2.dat) στο directory:

C:\bin

Έπειτα φορτώστε το αρχείο των δεδομένων στο MATLAB με την εντολή

```
>> load c:\bin\yV2.dat
```

```
>> xV = diff(yV2) ;
```



Μέθοδος εξομάλυνσης κινητού μέσου

Ένας άλλος τρόπος απαλοιφής της τάσης είναι με τη χρήση φίλτρου **κινούμενου μέσου** (moving average (MA) filter).

Συγκεκριμένα, με $1 \leq q \ll T$, εξομαλύνουμε την $y_t = \mu_t + x_t, t = 1, \dots, T$ από την σχέση

$$w_t = \frac{1}{2q+1} \sum_{j=-q}^q y_{t+j} = \frac{1}{2q+1} \sum_{j=-q}^q \mu_{t+j} + \frac{1}{2q+1} \sum_{j=-q}^q x_{t+j}$$

όπου $q+1 \leq t \leq T-q$.

Επειδή $E(X_t) = 0$, συμπεραίνουμε ότι το δεύτερο άθροισμα στα δεξιά της παραπάνω σχέσης είναι πολύ κοντά στο μηδέν και συνεπώς μπορούμε να δεχθούμε ότι τα μ_t ικανοποιούν το σύστημα εξισώσεων:

$$\frac{1}{2q+1} \sum_{j=-q}^q \mu_{t+j} = w_t = \frac{1}{2q+1} \sum_{j=-q}^q y_{t+j}$$

όπου $q + 1 \leq t \leq T - q$.

Για κάθε χρονική στιγμή t , τα μ_{t+j} , $-q \leq j \leq q$, συνδέονται γραμμικά μεταξύ τους, τότε έχουμε τις παρακάτω εκτιμήτριες κινητού μέσου:

$$\hat{\mu}_t = \frac{1}{2q+1} \sum_{j=-q}^q y_{t+j}, \text{ όπου } q + 1 \leq t \leq T - q.$$

Οι εκτιμώμενες τιμές της τάσης $\hat{\mu}_j$ αφαιρούνται από τις αντίστοιχες αρχικές παρατηρήσεις και η χρονοσειρά $\{x_j\}$ που προκύπτει είναι απαλλαγμένη από τάσεις. Οι πρώτες και οι τελευταίες παρατηρήσεις παραλείπονται ή θέτονται ίσες με τις αρχικές.

Παρατηρήσεις

- Η εκλογή του q παίζει σημαντικό ρόλο στην εξομάλυνση της σειράς και στη συνέχεια στην απαλοιφή της τάσης.
- Αν θέλουμε να απαλείψουμε μόνο πολύ αργές μεταβολές (τάσεις) θα πρέπει να χρησιμοποιήσουμε μεγάλη τάξη.
- Για την απαλοιφή μεταβολών σε μικρότερη χρονική κλίμακα η τάξη πρέπει να είναι αντίστοιχα μικρή.
- Αν δεν έχουμε κάποια πληροφορία για τη μορφή της τάσης που θέλουμε να απαλείψουμε, τότε δεν είναι πάντα ξεκάθαρο ποια μέθοδος ή παράμετρος μεθόδου απαλοιφής τάσης είναι η πιο κατάλληλη.

ΕΠΙΛΥΣΗ ΜΕ ΤΟ ΜΑΤΛΑΒ

Έστω ότι θέλουμε να χρησιμοποιήσουμε την χρονοσειρά $yV2$

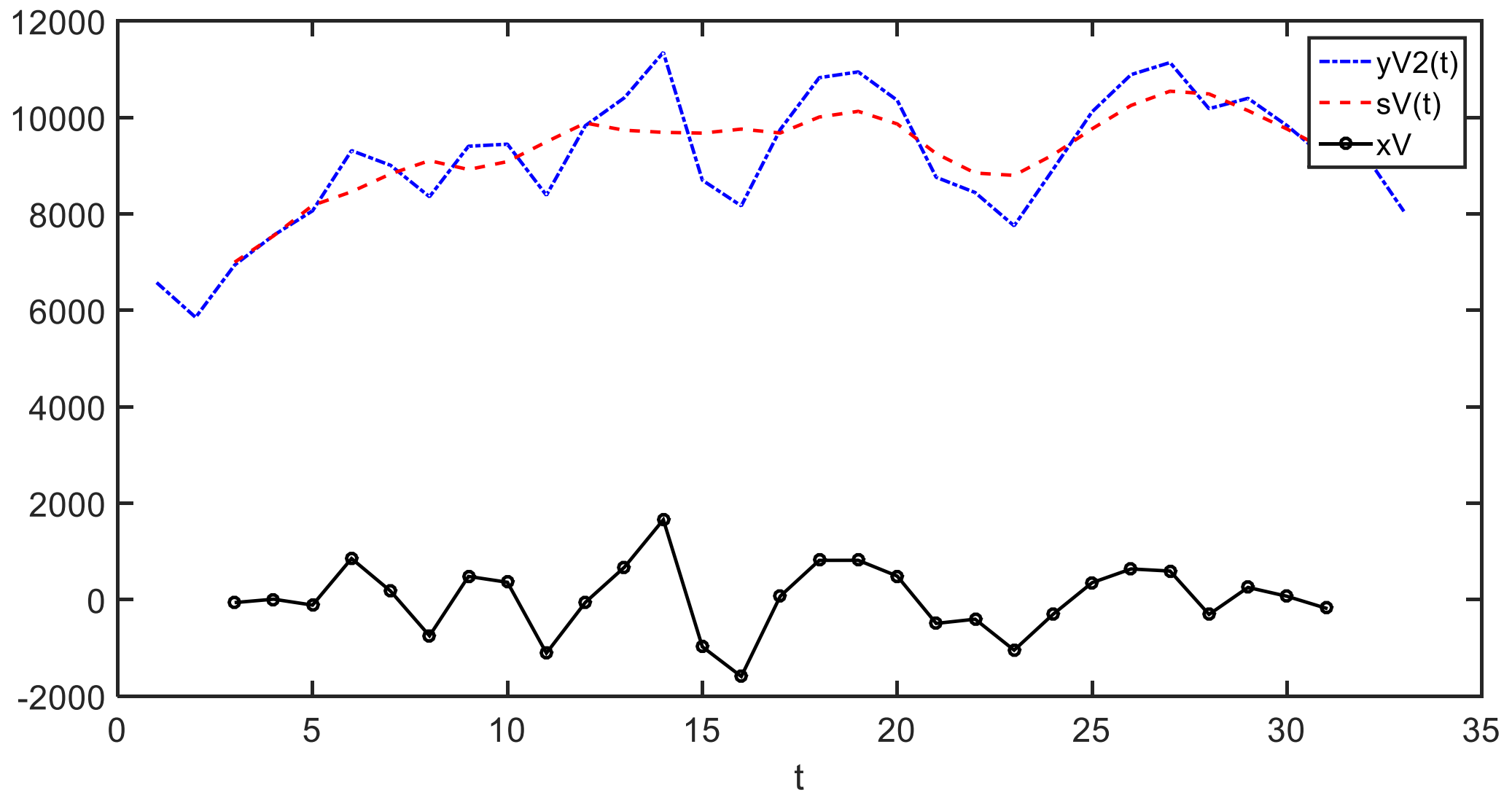
```
>> load c:\bin\yV2.dat
```

Συνάρτηση [movingaveragesmooth.m](#) : προσαρμογή τάσης με φίλτρο κινούμενου μέσου τάξης $2q+1$ (ή $2q$) (κυλιόμενος τοπικός μέσος όρος, τα πρώτα και τελευταία q στοιχεία είναι κενά).

```
>> maorder =5;      % τάξη φίλτρου κινούμενου μέσου
```

```
>> sV = movingaveragesmooth(yV2, maorder);
```

```
>> xV = yV2 - sV;
```

ΠΕΡΙΟΔΙΚΟΤΗΤΑΣ Η΄ ΕΠΟΧΙΚΟΤΗΤΑΣ

Η εποχική συνιστώσα s_t είναι της μορφής

$$s_t = a \sin 2\pi\omega t = a \sin \frac{2\pi t}{d}$$

όπου ω είναι η συχνότητα και $d = \frac{1}{\omega}$ η περίοδος, ή γενικότερα είναι της μορφής:

$$s_t = \sum_k \alpha_k \sin(2\pi\omega_k t + \theta_k)$$

ΑΠΑΛΟΙΦΗ ΠΕΡΙΟΔΙΚΟΤΗΤΑΣ Η΄ ΕΠΟΧΙΚΟΤΗΤΑΣ

Έστω ότι εξετάζουμε μια χρονοσειρά που έχει μόνο περιοδικότητα ή εποχικότητα, η οποία δίνεται ως

$$Y_t = s_t + X_t$$

όπου η **περίοδος** της περιοδικής συνάρτησης s_t είναι γνωστή και ίση με d .

Αν η **περιοδικότητα** αντιστοιχεί σε κάποια γνωστή περίοδο, όπως 24 ώρες ή 12 μήνες, αναφέρεται ως **εποχικότητα**.

Όπως για την τάση, μπορεί η περιοδική συνάρτηση να είναι τέτοια που επιτρέπει την προσέγγιση της με κάποια γνωστή (παραμετρική) περιοδική συνάρτηση, $s_t = f(t)$, όπως π.χ. μια ημιτονοειδής συνάρτηση. Συχνά όμως η μορφή της περιοδικής συνάρτησης δεν είναι συγκεκριμένη.

Όταν σε μια χρονοσειρά υπάρχει εποχικότητα, τότε είναι αναγκαίο κατά την εκτίμηση της τάσης να ληφθεί υπόψη η εποχική συνιστώσα.

Αυτό σημαίνει ότι θα πρέπει να έχουμε ήδη εκτιμήσει την εποχική συνιστώσα πράγμα το οποίο πάλι προϋποθέτει ότι γνωρίζουμε την τάση. Έτσι η εκτίμηση των δύο αυτών συνιστωσών γίνεται με επαναληπτικές διαδικασίες διαδοχικών εκτιμήσεων των δύο αυτών συνιστωσών ή με μία διαδικασία ταυτόχρονου προσδιορισμού τους.

Για απλούστευση του προβλήματος θα θεωρήσουμε ότι η εποχική συνιστώσα παρουσιάζει ένα μόνο κύκλο με συγκεκριμένη και γνωστή χρονική διάρκεια, π.χ. ετήσια. Αυτό συμβαίνει συχνά σε οικονομικές χρονοσειρές, όμως συχνά μπορεί να υπάρχουν κύκλοι διαφορετικών περιόδων, π.χ. σε μετεωρολογικά δεδομένα, όπου εκτός του ετήσιου κύκλου υπάρχουν και άλλοι κύκλοι μεγαλύτερης διάρκειας.

Το πρόβλημα της αναγνώρισης της ύπαρξης ή όχι κύκλων σε μια χρονοσειρά είναι βασικό αντικείμενο της Φασματικής Ανάλυσης και δεν θα ασχοληθούμε περαιτέρω εδώ.

Για την περιγραφή των μεθόδων θα θεωρήσουμε εδώ ότι στη χρονοσειρά υπάρχει ένας μόνο κύκλος με γνωστή περίοδο d ($d > 1$).

Έστω μια χρονοσειρά $\{y_t\}$, $t = 1, \dots, T$ η οποία παρουσιάζει ετήσια περιοδικότητα και ο χρόνος t μετράται σε μήνες. Η περίοδος της χρονοσειράς είναι συνεπώς $d = 12$. Η εποχική συνιστώσα της χρονοσειράς θα πρέπει να ικανοποιεί τις εξισώσεις:

$$s_t = s_{t+d} \text{ και } \sum_{l=1}^d s_{t+l} = 0 \text{ για κάθε } t.$$

Ως εκ τούτου θα πρέπει να προσδιορίσουμε την εποχική συνιστώσα κατά τρόπο που να ικανοποιούνται οι εξισώσεις αυτές.

Για διευκόλυνση στην ανάπτυξη των μεθόδων που εφαρμόζονται σε περιοδικές χρονοσειρές με περίοδο d , συμβολίζουμε με $y_{k,l}$ την παρατήρηση y_t με

$$t = (k - 1)d + l \quad k = 1, \dots, K \text{ και } l = 1, \dots, d.$$

Για απλούστευση θεωρούμε ότι το μήκος T της χρονοσειράς $\{y_t\}$ είναι πολλαπλάσιο της περιόδου d και συγκεκριμένα

$$T = Kd .$$

Μέθοδοι εκτίμησης περιοδικής συνιστώσας

- Μέθοδος χαμηλής τάσης
- Μέθοδος κινητού μέσου
- Μέθοδος διαφορών με υστέρηση d

Μέθοδος χαμηλής τάσης

Όταν η συνιστώσα της τάσης μ_t παρουσιάζει χαμηλό βαθμό μεταβολής τότε μπορούμε να θεωρήσουμε ότι, σε ετήσια βάση, η τάση διατηρεί ένα σταθερό επίπεδο τιμών. Το σταθερό αυτό επίπεδο τιμών μπορεί να εκτιμηθεί από τον μέσο όρο των τιμών y_t του αντίστοιχου έτους.

Έτσι για το έτος k , το επίπεδο τιμών εκτιμάται από το (δειγματικό) μέσο ετήσιο επίπεδο τιμών, δηλαδή

$$\hat{\mu}_k = \frac{1}{d} \sum_{l=1}^d y_{k,l} \quad k = 1, \dots, K$$

όπου K το πλήθος των ετών στα οποία αναφέρεται η εν λόγω χρονοσειρά.

Μένει τώρα να εκτιμήσουμε τις εποχικές συνιστώσες $s_l, l = 1, \dots, d$. Χρησιμοποιούμε τις μέσες αποκλίσεις από το εκάστοτε μέσο ετήσιο επίπεδο τιμών. Έχουμε δηλαδή ως εκτιμήτρια της περιοδικής συνιστώσας για το μήνα l την

$$\hat{s}_l = \frac{1}{K} \sum_{k=1}^K (y_{k,l} - \hat{\mu}_k), \quad l = 1, \dots, d$$

από την οποία προκύπτει ότι οι εκτιμήτριες των εποχικών συνιστωσών ικανοποιούν την συνθήκη

$$\sum_{l=1}^d \hat{s}_l = 0$$

Συνδυάζοντας τα παραπάνω έχουμε

$$y_{k,l} = \hat{\mu}_k + \hat{s}_l + \hat{X}_{k,l}, \quad k = 1, \dots, K, l = 1, \dots, d.$$

από όπου προκύπτει η χωρίς τάση και εποχικότητα χρονοσειρά

$$\hat{X}_{k,l} = y_{k,l} - \hat{\mu}_k - \hat{s}_l, \quad k = 1, \dots, K, l = 1, \dots, d.$$

ΕΠΙΛΥΣΗ με το MATLAB

Συνάρτηση [seasonalcomponents.m](#) : προσαρμογή εποχικότητας περιόδου s με υπολογισμό του μέσου όρου του κάθε στοιχείου της περιόδου.

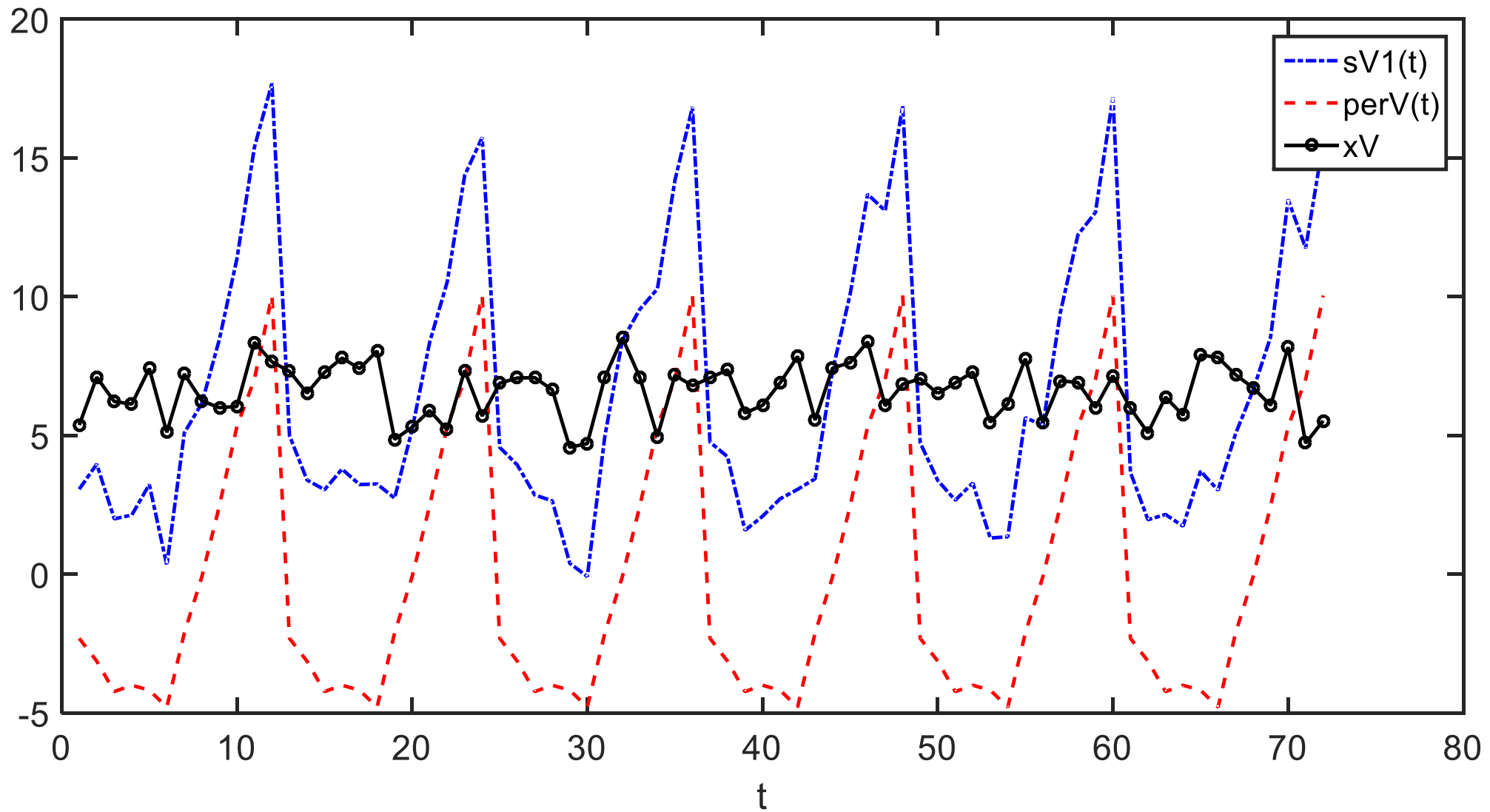
π.χ. μια χρονοσειρά με ετήσια εποχικότητα: sV1.dat

```
>> load cd:\bin\sV1.dat
```

```
>> per = 12; % season or period
```

```
>> perV = seasonalcomponents(sV1, per); % περιοδική συνιστώσα
```

```
>> xV = sV1 - perV; % στάσιμη χρονοσειρά
```



Μέθοδος κινητού μέσου

Η μέθοδος κινητού μέσου εφαρμόζεται σε χρονοσειρές στις οποίες η υπόθεση ότι η συνιστώσα τάσης διατηρεί “σταθερό” επίπεδο τιμών σε όλη τη διάρκεια του έτους δεν ευσταθεί.

Στην περίπτωση αυτή εφαρμόζεται μια μέθοδος τριών βημάτων:

- Στο πρώτο βήμα γίνεται αμφίπλευρη εξομάλυνση κινητού μέσου.
- Στο δεύτερο βήμα υπολογίζεται η φασματική συνιστώσα.
- Στο τρίτο υπολογίζεται η τάση έχοντας προηγουμένως “από-εποχικοποιήσει” την χρονοσειρά.

1ο βήμα

Με άρτια περίοδο $d = 2q$, λαμβάνουμε τους κινητούς μέσους της y_t , $t = 1, \dots, T$:

$$\hat{\mu}_t = \{0.5y_{t-q} + y_{t-q+1} + \dots + y_{t+q-1} + 0.5y_{t+q}\}/d$$

ενώ με περιττή περίοδο $d = 2q + 1$ λαμβάνουμε τους κινητούς μέσους

$$\hat{\mu}_t = \{y_{t-q} + y_{t-q+1} + \dots + y_{t+q-1} + y_{t+q}\}/d$$

Έχουμε δηλαδή και στις δύο περιπτώσεις εξομάλυνση κινητού μέσου της μορφής

$$\hat{\mu}_t = \sum_{j=-q}^q a_j y_{t+j}$$

με

$$\sum_{j=-q}^q a_j = 1$$

Με την παραπάνω διαδικασία προσδιορισμού των $\hat{\mu}_t$, $t = 1, 2, \dots$ εξαλείφεται η επίδραση της εποχικής συνιστώσας αφού στις δύο παραπάνω περιπτώσεις έχουμε

$$\sum_{j=-q}^q a_j s_{t+j} = 0$$

για κάθε t .

2ο βήμα

Λαμβάνουμε τους μέσους

$$w_l = \frac{1}{K} \sum_{k=1}^K \{x_{(k-1)d+l} - \hat{\mu}_{(k-1)d+l}\}, \quad l = 1, \dots, d.$$

Οι μέσοι αυτοί εκφράζουν την εποχική συνιστώσα όμως δεν ικανοποιούν την συνθήκη

$$\sum_{l=1}^d w_l = 0$$

Για αυτό χρησιμοποιούμε τις διαφορές $w_l - \bar{w}$, όπου \bar{w} είναι ο αριθμητικός μέσος των w_l , $l = 1, \dots, d$.

Συνεπώς η εκτιμήτρια της εποχικής συνιστώσας είναι:

$$\hat{s}_l = w_l - \bar{w} = w_l - \frac{1}{d} \sum_{l=1}^d w_l, \quad l = 1, \dots, d.$$

από την οποία προκύπτει ότι

$$\hat{s}_t = \hat{s}_l \quad \text{με} \quad l = t \bmod d, \quad t = 1, \dots, T.$$

$$\begin{aligned} &(\text{δηλαδή } t = (k - 1)d + l \\ &k = 1, \dots, K, \quad l = 1, \dots, d) \end{aligned}$$

3ο βήμα

Έχοντας εκτιμήσει την εποχική συνιστώσα μπορούμε να επανέλθουμε στη αρχική χρονοσειρά και να την “από-εποχικοποιήσουμε” λαμβάνοντας τις διαφορές

$$d_t = y_t - \hat{s}_t, t = 1, \dots, T.$$

Οι παραπάνω διαφορές μπορούν να χρησιμοποιηθούν για την επανεκτίμηση της τάσης εφαρμόζοντας μια από τις μεθόδους εκτίμησης τάσης της προηγούμενης Ενότητας.

Έχοντας ολοκληρώσει τα παραπάνω βήματα, χρειάζεται να εξετάσουμε πάλι γραφικά αν ο εκτιμώμενος θόρυβος είναι στάσιμος ή όχι:

$$\hat{X}_t = y_t - \hat{\mu}_t - \hat{s}_t, t = 1, \dots, T.$$

ΕΠΙΛΥΣΗ με το MATLAB

Συνάρτηση [movingaverageseasonal.m](#) : προσαρμογή εποχικότητας περιόδου s με φίλτρο κινούμενου μέσου τάξης s (αφαίρεση κυλιόμενου τοπικού μέσου όρου, τα πρώτα και τελευταία $s/2$ στοιχεία είναι κενά).

π.χ. η χρονοσειρά με ετήσια εποχικότητα: `sV1.dat`

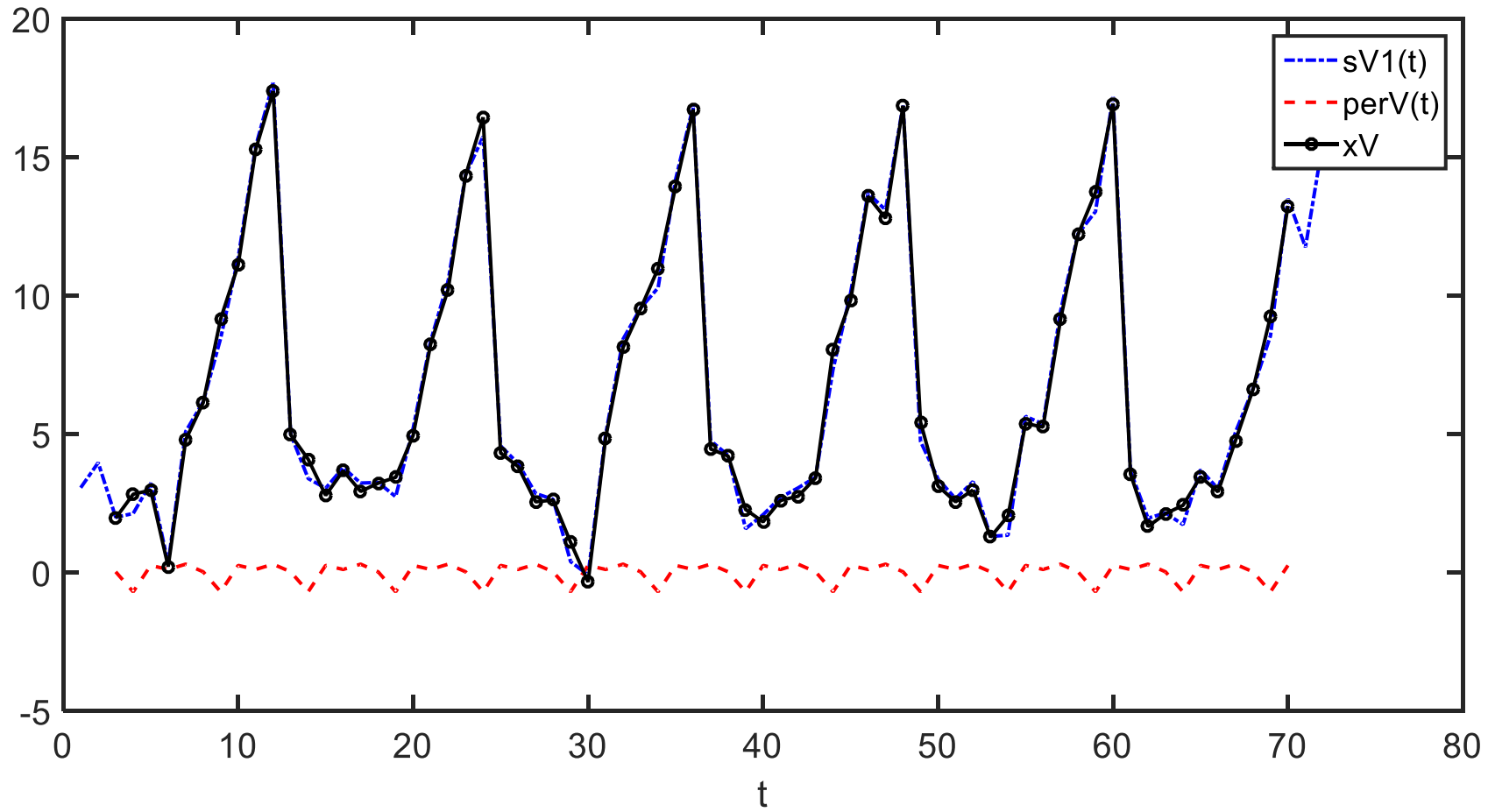
```
>> load cd:\bin\sV1.dat
```

```
>> maorder = 5; % τάξη φίλτρου κινούμενου μέσου
```

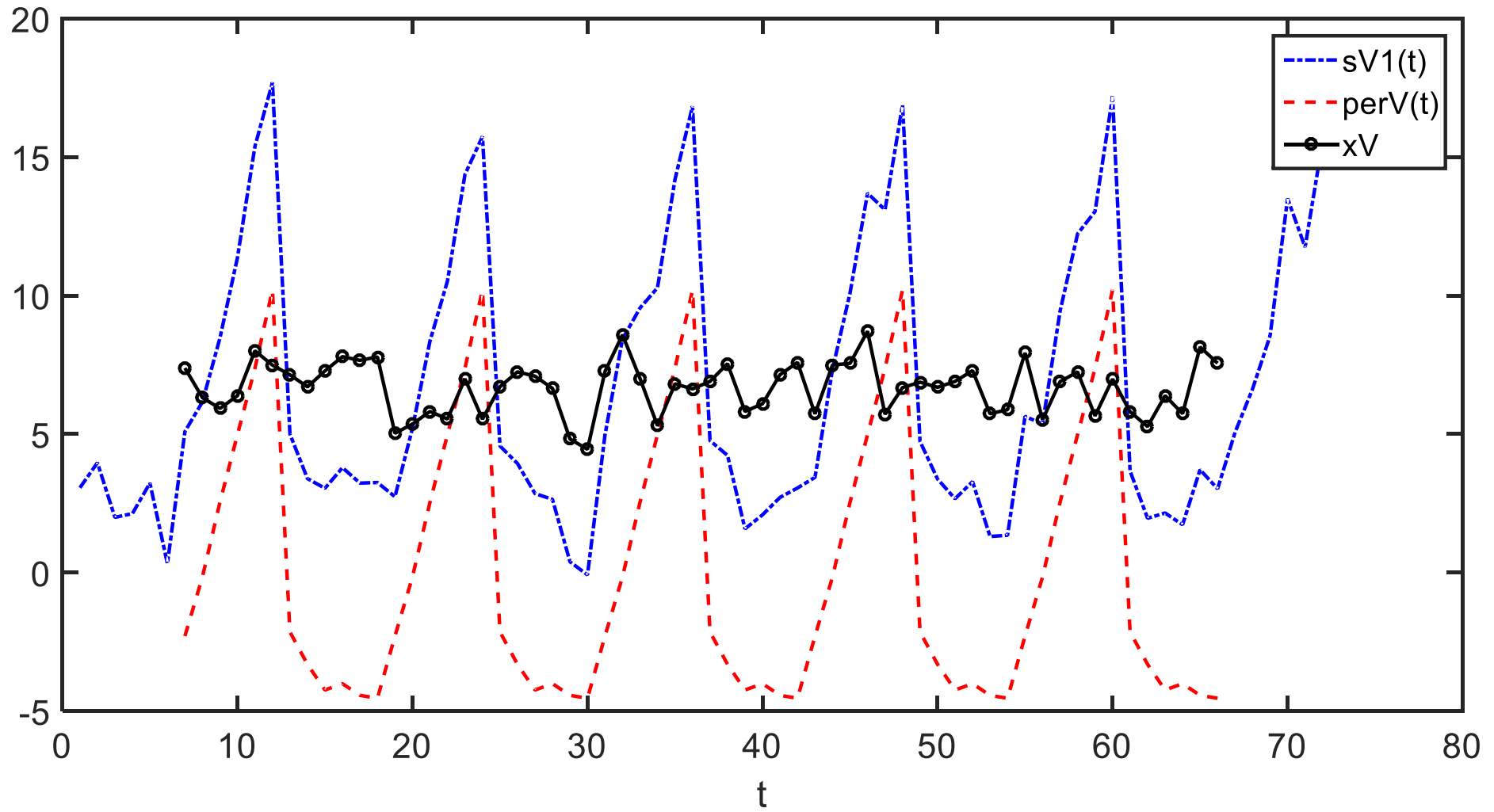
```
>> perV = movingaverageseasonal(sV1, maorder); % περιοδική συνιστώσα
```

```
>> xV = sV1 - perV; % στάσιμη χρονοσειρά
```

For $maorder = 5$



For $maorder = 12$



Μέθοδος διαφορών με υστέρηση d

Επειδή στις περιοδικές, χρονοσειρές ισχύει η σχέση

$$s_t = s_{t-d} \quad t = 1, 2, \dots$$

συμπεραίνουμε ότι αν λάβουμε τις διαφορές

$$y_t - y_{t-d} \quad t = d, d + 1, \dots, T$$

η εποχική συνιστώσα s_t εξαλείφεται.

Εφαρμόζοντας συνεπώς τον τελεστή διαφοράς με υστέρηση d

$$\begin{aligned} \nabla_d y_t &= y_t - y_{t-d} = \nabla_d (\mu_t + s_t + x_t) = \\ &= \mu_t - \mu_{t-d} + x_t - x_{t-d} \end{aligned} \quad t = d + 1, \dots$$

Έχοντας τώρα την απεριοδική χρονοσειρά $\{\nabla_d Y_t\}, t = d + 1, \dots$ οι διαφορές

$$\delta_t = \nabla_d \mu_t = \mu_t - \mu_{t-d}, t = d + 1$$

εκτιμώνται εφαρμόζοντας μια από τις μεθόδους της προηγούμενης ενότητας, όπως τη μέθοδο των διαφορών τάξης k , δηλαδή εφαρμόζοντας τον τελεστή $\nabla^k = (I - B)^k$ πάνω στην δ_t , με κατάλληλο k .

ΕΠΙΛΥΣΗ με το MATLAB

```
>> load cd:\bin\sV1.dat % Χρονοσειρά με εποχικότητα
```

```
>> T = length(sV1); (T=72)
```

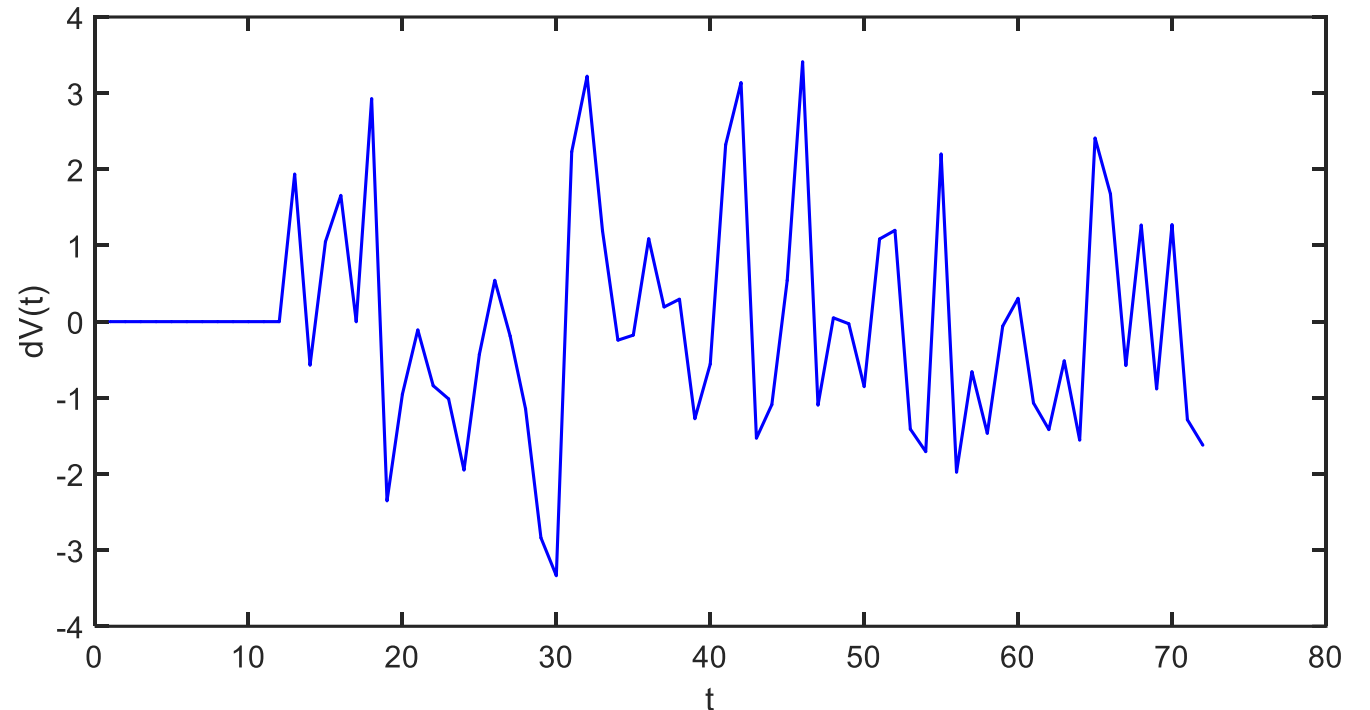
```
>> d = 12;
```

```
>> for t=d+1:T
```

```
    dV(t,1) = sV1(t)-sV1(t-d);
```

```
end
```

```
>> plot(dV)
```



Βιβλιογραφία

1. Ε. Μπόρα – Σέντα, Χ. Μωυσιάδης. Εφαρμοσμένη στατιστική, Β' έκδοση, Εκδόσεις Ζήτη, 1995.
2. Γ. Κ. Χρήστου. Εισαγωγή στην Οικονομετρία, Β τόμος (Γ' έκδοση), Εκδόσεις Gutenberg, 2007.
3. Δ. Κουγιουμτζής. Σημειώσεις μαθήματος Χρονοσειρών. Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΑΠΘ.
4. Γ.Ε. Κοκολάκης. Σημειώσεις ανάλυσης Χρονοσειρών. Σχολή Εφαρμοσμένων Μαθηματικών & Φυσικών Επιστημών, Αθήνα.