



ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
& ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης  
Πολυτεχνική Σχολή  
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
Τομέας Ηλεκτρονικής & Υπολογιστών

## Αναφορά Εξαμηνιαίας Εργασίας στις Χρονοσειρές

**Μάθημα:** Χρονοσειρές

**Συντάκτης:** Θανάσης Χαρισούδης  
ΑΕΜ: 9026

**Διδάσκων:** κ. Δημήτρης Κουγιουμτζής  
Καθηγητής Α.Π.Θ.

8 Φεβρουαρίου 2021

# Περιεχόμενα

<b>1</b>	<b>Διερεύνηση Χρονοσειράς A</b>	<b>1</b>
1.1	Αρχική Ανάλυση . . . . .	1
1.2	Απαλοιφή Στοχαστικής Τάσης . . . . .	3
1.3	Απαλοιφή Στοχαστικής Τάσης & Σταθεροποίηση Διασποράς . . . . .	5
<b>2</b>	<b>Προσαρμογή Γραμμικού Μοντέλου</b>	<b>7</b>
2.1	Προσαρμογή ARIMA(p,1,q) . . . . .	7
2.1.1	Προσαρμογή ARMA(p,q) στη χρονοσειρά πρώτων διαφορών . . . . .	7
2.2	Μοντέλο και Σφάλματα Προσαρμογής . . . . .	8

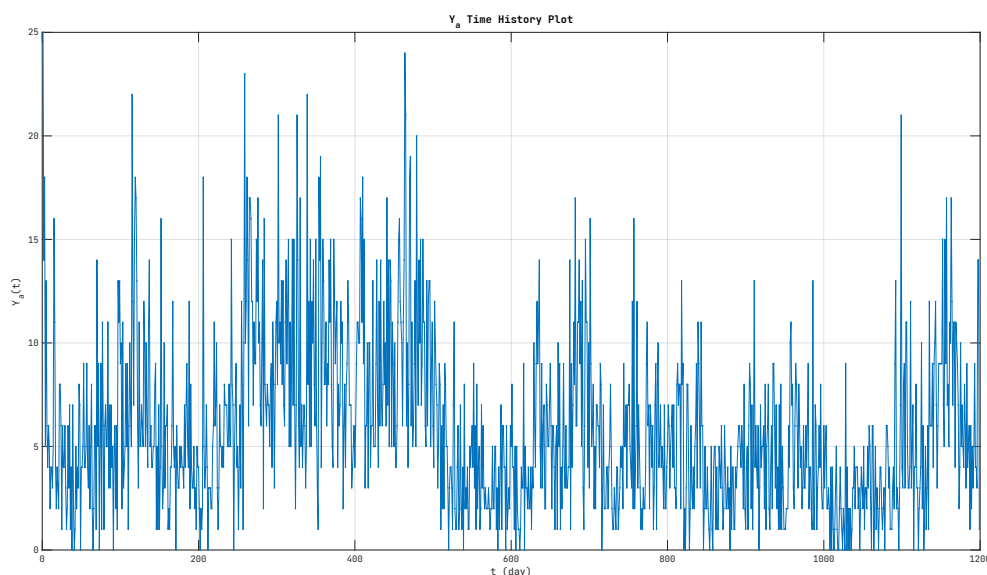
# Βήμα 1

## Διερεύνηση Χρονοσειράς A

Σκοπός της συγκεκριμένης εργασίας του μαθήματος “Χρονοσειρές” είναι να γίνει ανάλυση κάποιων χρονοσειρών από προβολές βίντεο με σκοπό τη δημιουργία αντιπροσωπευτικών μοντέλων αυτών και κατ’ επέκταση, την πρόβλεψή τους. Η ομάδα μου είναι η ομάδα Νο. 3 και επομένως οι χρονοσειρές που θα αναλυθούν στα πλαίσια αυτής της εργασίας είναι αυτές με δείκτες 3 (η “A”) και 13 (η “B”). Ακολουθεί η αρχική ανάλυση της κάθε μίας χρονοσειράς ξεχωριστά, όπου δίνονται τα διαγράμματα ιστορίας, αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Στη συνέχεια γίνεται αφαίρεση τάσης και σταθεροποίηση διασποράς προκειμένου να γίνει στάσιμη η χρονοσειρά, ενώ παρατίθενται και τα αντίστοιχα διαγράμματα.

### 1.1 Αρχική Ανάλυση

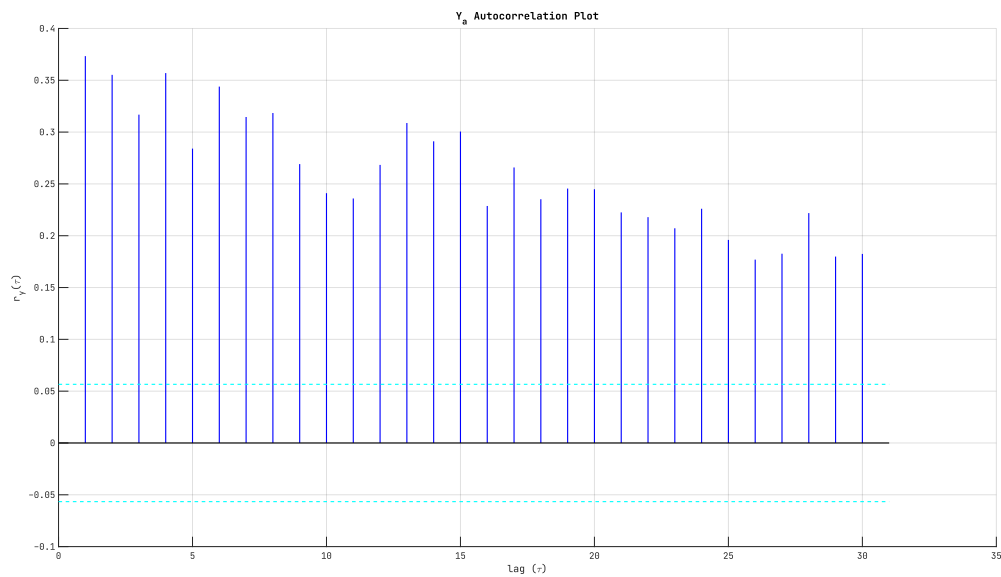
Αρχικά, παρατίθεται το διάγραμμα ιστορίας της χρονοσειράς A,  $Y_a(t)$ , ακολούθως.



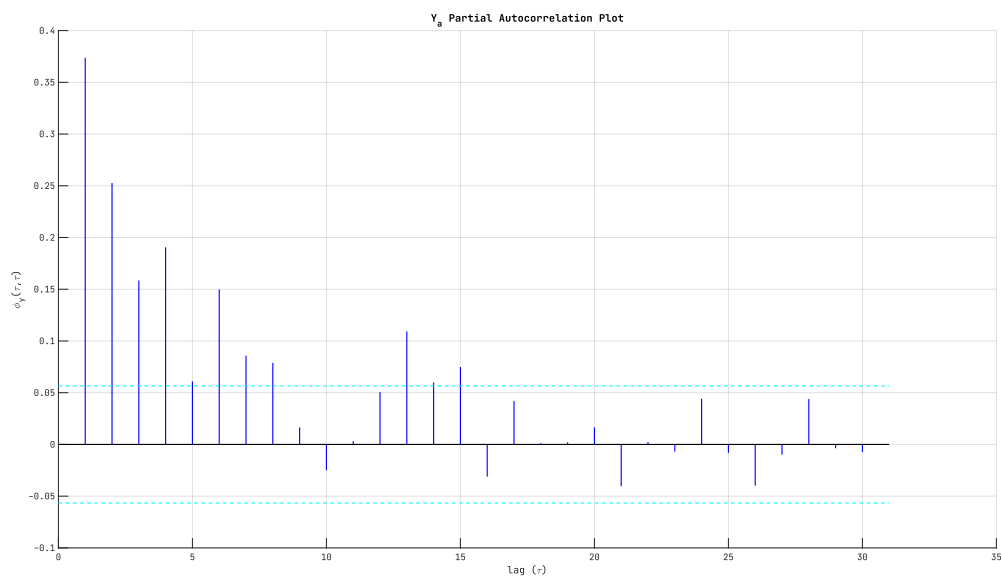
Σχήμα 1.1: Διάγραμμα ιστορίας της χρονοσειράς A,  $Y_a(t)$

Από το διάγραμμα ιστορίας της  $Y_a(t)$  φαίνεται πως η χρονοσειρά δεν είναι στάσιμη καθώς παρατηρείται κάποια τάση. Αυτό επιβεβαιώνεται και από τα διαγράμματα της (δειγματικής) αυτοσυσχέτισης

και (δειγματικής) μερικής αυτοσυσχέτισης που παραθέτονται ακολούθως:



Σχήμα 1.2: Διάγραμμα (δειγματικής) αυτοσυσχέτισης της χρονοσειράς A,  $r_y(\tau)$



Σχήμα 1.3: Διάγραμμα (δειγματικής) μερικής αυτοσυσχέτισης της χρονοσειράς A,  $\phi_y(\tau)$

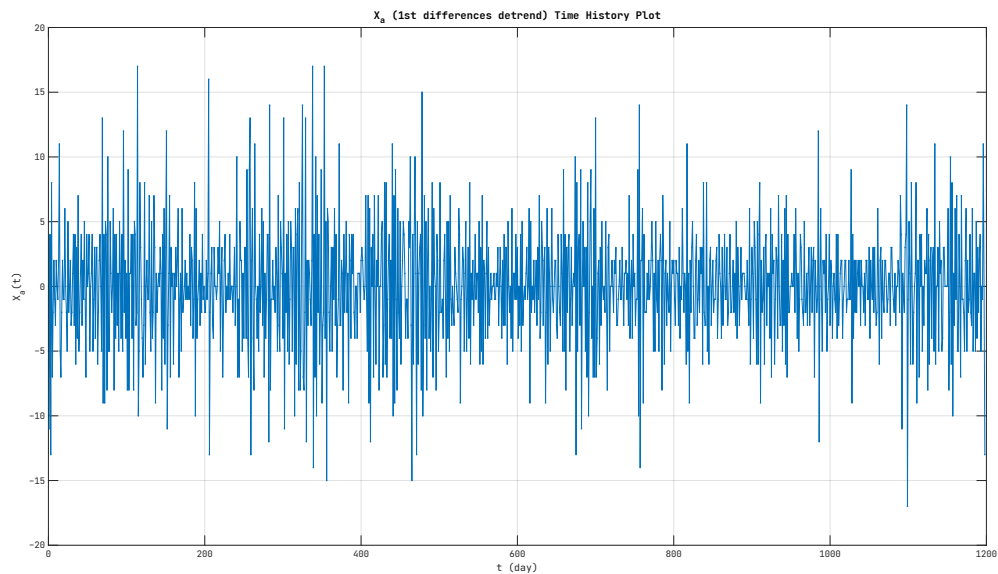
Στα παραπάνω διαγράμματα έχουν σημειωθεί και τα όρια σημαντικότητας για επίπεδο εμπιστοσύνης 95%. Ειδικά στο διάγραμμα αυτοσυσχέτισης φαίνεται έντονα η ύπαρξη τάσης καθώς η αυτοσυσχέτιση έχει υψηλές τιμές και φθίνει πολύ αργά. Η τάση όπως φαίνεται από το διάγραμμα ιστορίας (σχήμα 1.1) είναι στοχαστική και επομένως για να γίνει στάσιμη η χρονοσειρά θα πρέπει να εφαρμοστεί κάποια μέθοδος απαλοιφής της στοχαστικής τάσης (όπως η μέθοδος των πρώτων διαφορών).

## 1.2 Απαλοιφή Στοχαστικής Τάσης

Για την απαλοιφή της τάσης αρχικά δοκιμάστηκαν οι πρώτες διαφορές. Η χρονοσειρά που προκύπτει λοιπόν θα είναι:

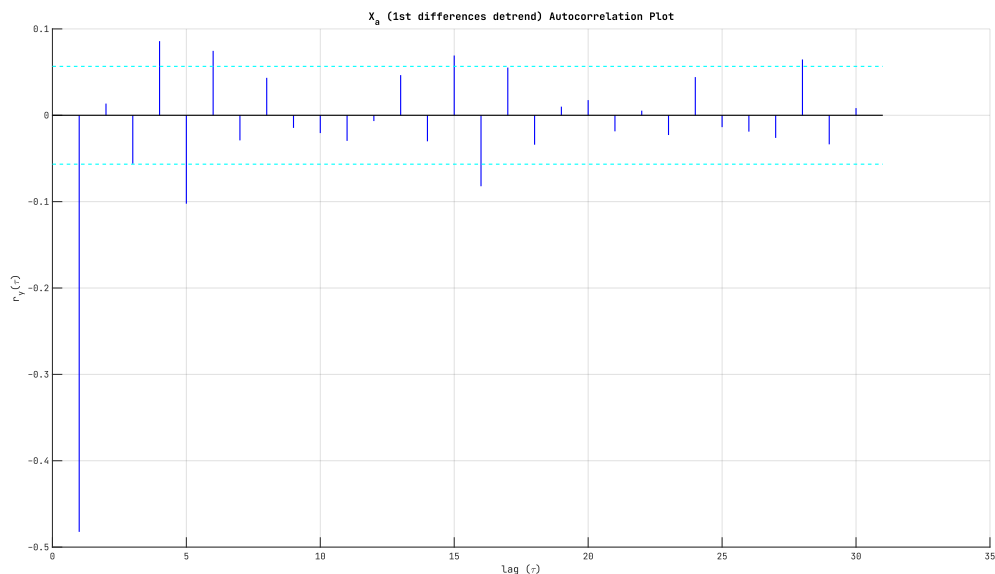
$$X_a(t) = Y_a(t) - Y_a(t - 1)$$

Παρακάτω παρατίθεται το διάγραμμα ιστορίας της  $X_a(t)$  όπου φαίνεται ότι ο μετασχηματισμός των πρώτων διαφορών έχει καταφέρει να απαλείψει τη στοχαστική τάση και επομένως δεν υπάρχει η ανάγκη για επανάληψη της διαδικασίας (δηλαδή να πάρω διαφορές δεύτερης τάξης).

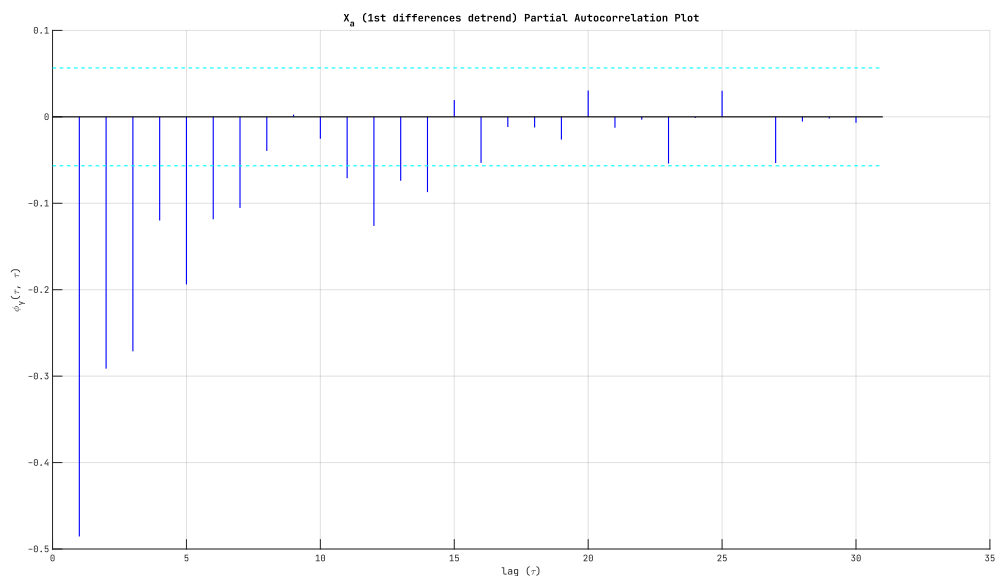


Σχήμα 1.4: Διάγραμμα ιστορίας στη χρονοσειρά των πρώτων διαφορών,  $X_a(t)$

Ο παραπάνω ισχυρισμός περί απαλοιφής τάσης ενισχύεται και από τα ακόλουθα διαγράμματα δειγματικής αυτοσυσχέτισης της χρονοσειράς των πρώτων διαφορών:



Σχήμα 1.5: Διάγραμμα (δειγματικής) αυτοσυσχέτισης της χρονοσειράς πρώτων διαφορών,  $r_x(\tau)$



Σχήμα 1.6: Διάγραμμα (δειγματικής) μερικής αυτοσυσχέτισης της χρονοσειράς πρώτων διαφορών,  $\phi_x(\tau)$

όπου αν και δεν φαίνεται η χρονοσειρά να παρουσιάζει αυτοσυσχέτιση λευκού θορύβου επίσης όμως δεν φαίνεται να παρουσιάζει και (στατιστικά) σημαντικές αυτοσυσχετίσεις (σχήμα 1.5) ή αργές μεταβολές.

Σε ότι αφορά τη διασπορά της χρονοσειράς των πρώτων διαφορών, επιστρέφοντας στο διάγραμμα ιστορίας της (σχήμα 1.4) και συγκρίνοντάς το με το διάγραμμα ιστορίας της αρχικής χρονοσειράς (σχήμα 1.1) παρατηρούμε ότι η διασπορά της χρονοσειράς των πρώτων διαφορών φαίνεται να μεταβάλλεται κάπως "σχετικά" με τις "καμπύλες" της τάσης της αρχικής χρονοσειράς. Επομένως, ίσως να ήταν σκόπιμο να προσπαθούσαμε να σταθεροποιήσουμε τη διασπορά της προκύπτουσας χρονοσειράς ως προς τη τάση της αρχικής, κάτι που αναλύεται ακολούθως.

### 1.3 Απαλοιφή Στοχαστικής Τάσης & Σταθεροποίηση Διασποράς

Η απαλοιφή της τάσης (detrending) θεωρούμε πως έχει επιτευχθεί ικανοποιητικά παίρνοντας τις πρώτες διαφορές στην αρχική χρονοσειρά. Ωστόσο, πριν πάρουμε τις πρώτες διαφορές θα μας ενδιέφερε να απαλείψουμε την εξάρτηση της διασποράς της (αρχικής) χρονοσειράς από τη τάση. Για αυτό θα χρησιμοποιηθεί ο μετασχηματισμός των Box & Cox με  $\lambda = 0.5$  αφενός διότι φαίνεται η διασπορά να εξάρταται με ανάλογο (γραμμικό) τρόπο από τη τάση και αφετέρου επειδή η χρονοσειρά περιέχει μηδενικά δείγματα (καμία προβολή του αντίστοιχου βίντεο εκείνη την ημέρα) και άρα οι μετασχηματισμοί λογαρίθμου ή σχετικών μεταβολών δεν αποτελούν πρακτική λύση.

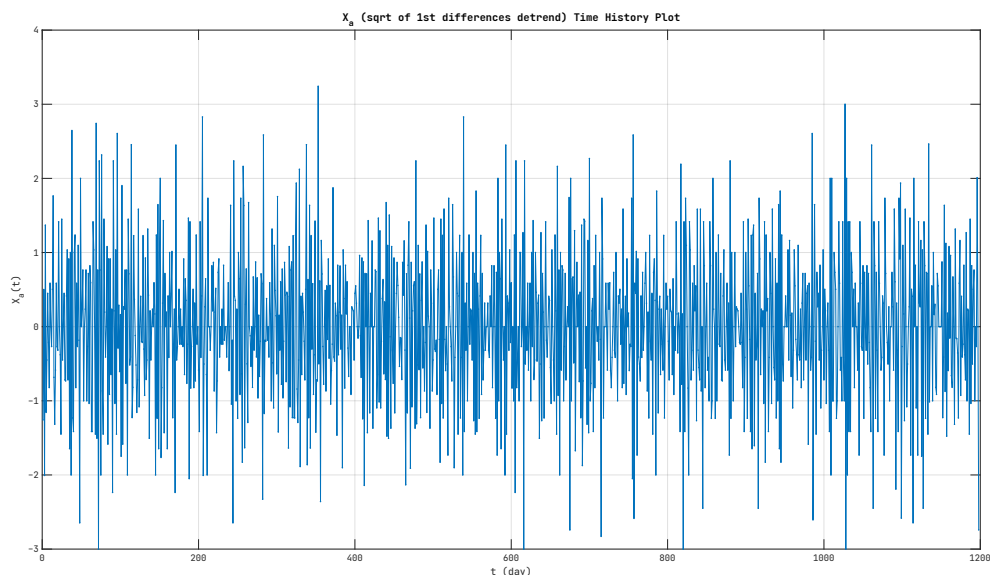
Ο μετασχηματισμός, λοιπόν, που χρησιμοποιήθηκε στην αρχική χρονοσειρά είναι αυτός της τετραγωνικής ρίζας. Δηλαδή, αρχικά η  $Y_a(t)$  μετασχηματίστηκε σε:

$$Y'_a(t) = \sqrt{Y_a(t)}$$

ενώ στη συνέχεια χρησιμοποιήθηκαν οι πρώτες διαφορές για απαλοιφή της τάσης. Συνολικά, λοιπόν, ο μετασχηματισμός που υλοποιήθηκε για μετασχηματισμό της αρχικής χρονοσειράς σε στάσιμη είναι ο ακόλουθος:

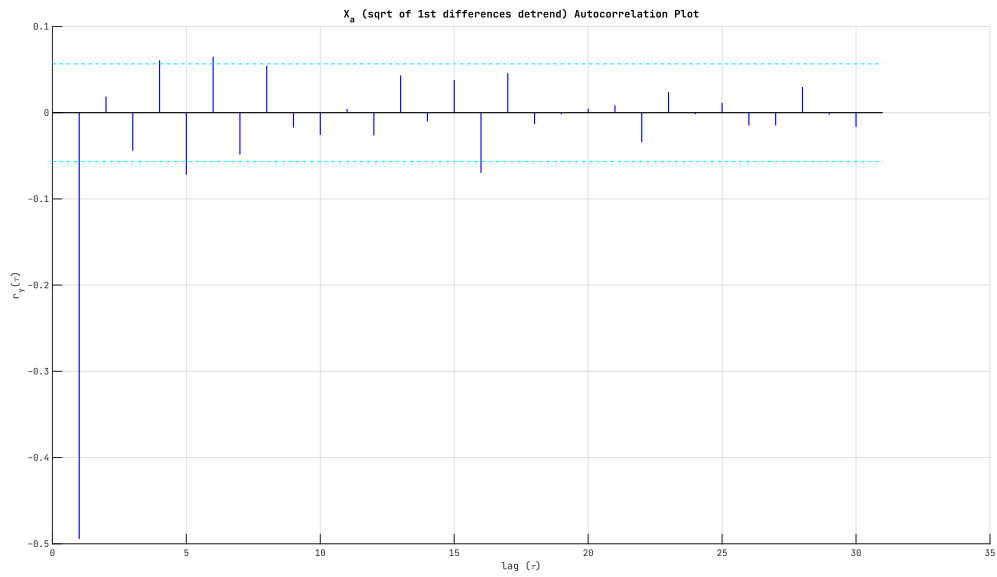
$$X_a(t) = \sqrt{Y_a(t)} - \sqrt{Y_a(t-1)}$$

με την  $X_a(t)$  να είναι η στάσιμη εκδοχή της  $Y_a(t)$ . Παρακάτω, φαίνεται το διάγραμμα ιστορίας της (νέας)  $X_a(t)$  όπου επιβεβαιώνεται η υπόθεσή μας για συσχέτισης της διασποράς με την τάση (δεν φαίνεται αυτή η “κυματοειδής” μεταβολή της διασποράς).

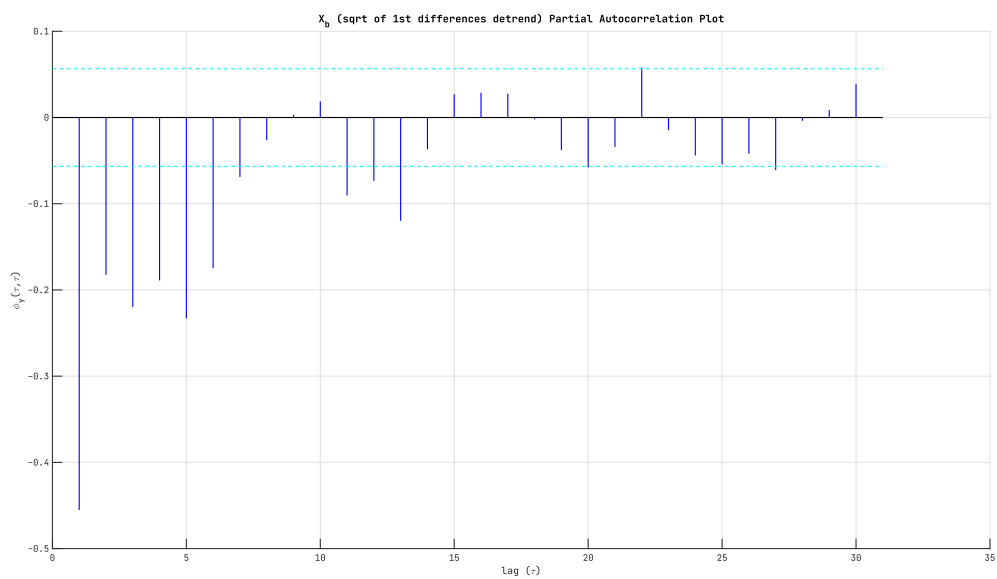


Σχήμα 1.7: Διάγραμμα ιστορίας στη χρονοσειρά των πρώτων διαφορών των τετρ. ριζών,  $X_a(t)$

Για τη χρονοσειρά των πρώτων διαφορών των τετραγωνικών ριζών που καταλήξαμε, παραθέτονται επίσης και τα διαγράμματα δειγματικής αυτοσυσχέτισης και δειγματικής μερικής αυτοσυσχέτισης με τα όρια σημαντικότητας (για εμπιστοσύνη 95%):



Σχήμα 1.8: Διάγραμμα (δειγματικής) αυτοσυσχέτισης της χρονοσειράς πρώτων διαφορών των τετραγωνικών ριζών,  $r_x(\tau)$



Σχήμα 1.9: Διάγραμμα (δειγματικής) μερικής αυτοσυσχέτισης της χρονοσειράς πρώτων διαφορών των τετραγωνικών ριζών,  $\phi_x(\tau)$



## Βήμα 2

# Προσαρμογή Γραμμικού Μοντέλου

Με βάση τα συμπεράσματα που εξήχθηκαν στην ανάλυση του πρώτου βήματος το γραμμικό μοντέλο που θα προσαρμόσουμε στη χρονοσειρά  $A$  είναι ARIMA με  $d = 1$ , καθώς θα χρησιμοποιηθούν διαφορές πρώτης τάξης. Η χρονοσειρά που θα προσαρμοστεί το μοντέλο αυτό δεν είναι η αρχική χρονοσειρά,  $Y_a(t)$ , αλλά η χρονοσειρά των τετραγωνικών ριζών,  $\sqrt{Y_a(t)}$  ώστε να έχουμε σταθεροποιήσει τη διασπορά (συγκ. την εξάρτησή της από την τάση) πριν την προσαρμογή του γραμμικού μοντέλου. Η ανάλυση που ακολουθεί, επομένως, αφορά τη χρονοσειρά των τετραγωνικών ριζών και παρατίθεται ακολούθως.

## 2.1 Προσαρμογή ARIMA(p,1,q)

Μετά την εφαρμογή των πρώτων διαφορών στη χρονοσειρά των τετραγωνικών ριζών έχουμε υλοποιήσει το πρώτο στάδιο προσαρμογής ενός μοντέλου ARIMA(p,1,q) που είναι η εφαρμογή διαφορών  $p$ -οστής τάξης. Το δεύτερο στάδιο είναι η εύρεση των παραμέτρων του μοντέλου ARMA(p,q), κάτι που αναλύεται στην επόμενη υπο-ενότητα.

### 2.1.1 Προσαρμογή ARMA(p,q) στη χρονοσειρά πρώτων διαφορών

Οι παράμετροι  $p$  και  $q$  είναι hyperparameters του μοντέλου και επομένως το πρώτο μας μέλημα είναι να βρούμε το βέλτιστο συνδυασμό των τάξεων του μοντέλου. Προς το σκοπό αυτό θα κάνουμε grid search για τιμές των τάξεων από 0 (απουσία του αντίστοιχου όρου) έως και 10.

Για την αξιολόγηση του κάθε συνδυασμού χρησιμοποιήθηκαν τα κριτήρια πληροφορίας Akaike (AIC) και Forward Prediction Error (FPE), τα οποία ορίζονται ως εξής:

$$AIC(p, q) = \ln(s_z^2) + \frac{2 \times (p + q)}{n}$$

και

$$FPE(p, q) = s_z^2 \times \frac{n + (p + q)}{n - (p + q)}$$

όπου  $n$  είναι ο αριθμός των δειγμάτων που χρησιμοποιήθηκαν για την εκτίμηση των ροπών (π.χ για τη δειγματική αυτοσυσχέτιση) και  $s_z^2$  είναι η (δειγματική) διασπορά των σφαλμάτων ή υπολοίπων

που προκύπτουν όταν συγκρίνουμε τις τιμές του προσαρμοσμένου μοντέλου τάξης  $(p, q)$  με τις πραγματικές τιμές της χρονοσειράς (ενν. τη χρονοσειρά  $\{X_a(t)\}$  που προέκυψε ως οι πρώτες διαφορές της χρονοσειράς των τετραγωνικών ριζών της αρχικής χρονοσειράς προβολών του βίντεο  $A, \{Y_a(t)\}$ ).

Παρακάτω παρατίθενται οι τιμές του AIC για τους συνδυασμούς των παραμέτρων  $(p, q) = (0 \dots 10, 0 \dots 10)$ , φυσικά με εξαίρεση το συνδυασμό  $(p, q) = (0, 0)$ :

		$p$										
$q$		0	1	2	3	4	5	6	7	8	9	10
	0		-0.465	-0.463	-0.462	-0.461	-0.460	-0.459	-0.459	-0.457	-0.458	-0.457
	1	-0.211	-0.463	-0.463	-0.462	-0.462	-0.461	-0.460	-0.458	-0.456	-0.457	-0.456
	2	-0.306	-0.462	-0.462	-0.465	-0.463	-0.461	-0.460	-0.457	-0.456	-0.459	-0.458
	3	-0.376	-0.461	-0.462	-0.463	-0.462		-0.458	-0.459	-0.457	-0.458	-0.457
	4	-0.395	-0.46	-0.461	-0.462			-0.463	-0.460	-0.462	-0.457	-0.456
	5	-0.427	-0.460	-0.459	-0.460	-0.459	-0.461	-0.459		-0.455	-0.455	-0.457
	6	-0.436	-0.459	-0.458	-0.459	-0.459	-0.460		-0.463	-0.457	-0.458	-0.455
	7	-0.448	-0.458	-0.456	-0.457	-0.457	-0.457	-0.456	-0.457	-0.462	-0.454	-0.453
	8	-0.449	-0.458	-0.456	-0.459	-0.458	-0.456	-0.456	-0.457	-0.454	-0.462	-0.455
	9	-0.447	-0.456	-0.456	-0.455	-0.456	-0.456	-0.455	-0.455	-0.454	-0.452	-0.456
	10	-0.446	-0.456	-0.454	-0.454	-0.455	-0.454	-0.453	-0.459	-0.456		-0.456

Πίνακας 2.1: Αναζήτηση Πλέγματος με βάση τη μετρική AIC για διάφορες τιμές των τάξεων  $(p, q)$

από όπου φαίνεται πως η χαμηλότερη τιμή του AIC επιτυγχάνεται όταν προσαρμόζεται μοντέλο AR-MA(0,1) ή, ισοδύναμα, μοντέλο MA(1). Ακολούθως δίνονται οι τιμές του FPE για τους αντίστοιχους συνδυασμούς τιμών των  $p$  και  $q$ :

		$p$										
$q$		0	1	2	3	4	5	6	7	8	9	10
	0		0.628	0.629	0.630	0.631	0.632	0.632	0.632	0.633	0.632	0.633
	1	0.810	0.629	0.629	0.630	0.630	0.631	0.631	0.632	0.634	0.633	0.634
	2	0.737	0.630	0.630	0.628	0.629	0.630	0.631	0.633	0.634	0.632	0.632
	3	0.687	0.631	0.630	0.629	0.630		0.633	0.632	0.633	0.632	0.633
	4	0.673	0.631	0.631	0.630			0.629	0.631	0.630	0.633	0.634
	5	0.652	0.632	0.632	0.631	0.632	0.631	0.632		0.634	0.634	0.633
	6	0.647	0.632	0.633	0.632	0.632	0.631		0.630	0.633	0.633	0.634
	7	0.639	0.633	0.634	0.633	0.633	0.633	0.634	0.633	0.630	0.635	0.635
	8	0.639	0.633	0.634	0.632	0.632	0.634	0.634	0.633	0.635	0.630	0.635
	9	0.640	0.634	0.634	0.635	0.634	0.634	0.634	0.634	0.635	0.636	0.634
	10	0.640	0.634	0.635	0.635	0.635	0.635	0.636	0.632	0.634		0.634

Πίνακας 2.2: Αναζήτηση Πλέγματος με βάση τη μετρική FPE για διάφορες τιμές των τάξεων  $(p, q)$

Όπως επιβεβαιώνεται και από τους δύο πίνακες παραπάνω, φαίνεται πως από τα γραμμικά μοντέλα καλύτερα προσαρμόζεται το μοντέλο MA(1).

## 2.2 Μοντέλο και Σφάλματα Προσαρμογής

μπλα μπλα μπλα