Ασκήσεις Κεφαλαίου 4

1. Δημιουργείστε μια συνάρτηση που θα υπολογίζει τη μερική αυτοσυσχέτιση σε μια χρονοσειρά. Ως είσοδο η συνάρτηση θα παίρνει το διάνυσμα της χρονοσειράς και την τιμή της μέγιστης τάξης (υστέρησης). Ως έξοδο θα δίνει ένα διάνυσμα με τις τιμές της μερικής αυτοσυσχέτισης για κάθε τάξη ως τη μέγιστη τάξη. Δοκιμάστε τη συνάρτηση αυτή σε χρονοσειρές από AR διαδικασία για διαφορετικές τάξεις.

   Βοήθεια matlab: Η μερική αυτοσυσχέτιση για κάθε τάξη p είναι ο συντελεστής του τελευταίου όρου του AR(p). Για την εκτίμηση των παραμέτρων του AR(p) μπορείς να χρησιμοποιήσεις τις συναρτήσεις ar για τον υπολογισμό του μοντέλου και polydata για την εξαγωγή του πολυωνύμου (διανύσματος) από τη δομή του μοντέλου που δίνει η συνάρτηση ar. Για το σχηματισμό χρονοσειρών άπό AR διαδικασία, χρησιμοποιείστε τη συνάρτηση AR.m που δίνεται στην ιστοσελίδα του μαθήματος.

2. Ένας έλεγχος για την ανεξαρτησία χρονοσειράς (τυχαία σειρά παρατηρήσεων) είναι από την συνάρτηση αυτοσυσχέτισης, δηλαδή ελέγχοντας αν η αυτοσυσχέτιση $r(\tau)$ για κάθε υστέρηση τ, από ένα ως κάποια μέγιστη υστέρηση $\tau_{\max}$, είναι στατιστικά ασήμαντη, δηλαδή είναι μέσα στα όρια $\pm 1.96/\sqrt{N}$ (για επίπεδο εμπιστοσύνης 95%), όπου N είναι το μήκος της χρονοσειράς. Έχει βρεθεί από τους Ljung και Box ότι το παρακάτω άθροισμα των τετραγώνων των αυτοσυσχετίσεων

$$Q = N(N+2)\sum_{\tau=1}^{\tau_{\max}} \frac{r(\tau)^2}{N-\tau}$$

   ακολουθεί κατανομή $\mathcal{X}^2$ με $\tau_{\max}$ βαθμούς ελευθερίας. Άρα μπορούμε να κάνουμε έλεγχο υπόθεσης με μηδενική υπόθεση ότι η χρονοσειρά είναι τυχαία (έχει μηδενική αυτοσυσχέτιση για όλες τις υστερήσεις). Υπολογίζουμε το Q από τη χρονοσειρά και απορρίπτουμε αυτή τη μηδενική υπόθεση σε επίπεδο σημαντικότητας α (δηλαδή σε 1-α επίπεδο εμπιστοσύνης) αν $Q > \mathcal{X}^2_{1-a,\tau_{\max}}$, όπου $\mathcal{X}^2_{1-a,\tau_{\max}}$ είναι η αντίστοιχη κρίσιμη τιμή (το όριο της ουράς της κατανομής) για πιθανότητα 1-α και $\tau_{\max}$ βαθμούς ελευθερίας. Αντίστοιχα μπορεί κάποιος να υπολογίσει την p-τιμή του ελέγχου ως το α που δίνει ως κρίσιμη τιμή $\mathcal{X}^2_{1-a,\tau_{\max}}$ την τιμή της Q.

   Δημιουργείστε μια συνάρτηση που κάνει αυτόν τον έλεγχο. Η συνάρτηση θα πρέπει να παίρνει ως είσοδο τη χρονοσειρά και τη μέγιστη υστέρηση και να δίνει ως έξοδο την p-τιμή του ελέγχου Ljung-Box. Δοκιμάστε τη συνάρτηση αυτή σε χρονοσειρές από AR διαδικασία κάποιας τάξης p εφαρμόζοντας μοντέλα AR μικρότερης τάξης και τάξης p.

   Βοήθεια matlab: Για τον υπολογισμό της κρίσιμης τιμής της $\mathcal{X}^2$ με $\tau_{\max}$ βαθμούς ελευθερίας χρησιμοποίησε τη συνάρτηση chi2inv και για την p-τιμή τη συνάρτηση chi2cdf.

3. Δίνεται η χρονοσειρά της τιμής κλεισίματος του δείκτη S&P500 για την περίοδο 20/10/1982 – 13/3/2008. Δημιουργείστε τη χρονοσειρά των αποδόσεων. Ελέγξετε αν η χρονοσειρά αυτή είναι ανεξάρτητη χρησιμοποιώντας τον έλεγχο Ljung-Box (δες Άσκηση 2). Αν δεν προκύπτει ανεξαρτησία, κάνετε προσαρμογή κατάλληλου AR μοντέλου. Επιλέξτε την τάξη του AR μοντέλου από τη μερική αυτοσυσχέτιση (δες Άσκηση 1) και το κριτήριο AIC.

   Βοήθεια matlab: Για τον υπολογισμό του κριτηρίου AIC μπορείς να χρησιμοποιήσεις τη συνάρτηση aic που παίρνει ως όρισμα τη δομή του μοντέλου που υπολογίζει η συνάρτηση ar και δίνει την τιμή της συνάρτησης AIC για το συγκεκριμένο μοντέλο. Η χρονοσειρά του SP500 δίνεται στην ιστοσελίδα του μαθήματος.

4. Κάνετε την ίδια διαδικασία όπως στην Άσκηση 3 για τη χρονοσειρά των πρώτων διαφορών που θα δημιουργήσετε από τη χρονοσειρά του δείκτη της μηνιαίας τιμής συναλλάγματος βρετανικής λίρας σε δολλάριο (GBR/USD) για την περίοδο από Ιανουάριο 1971 ως Αύγουστο 2000. Η χρονοσειρά αυτού του δείκτη δίνεται στην ιστοσελίδα του μαθήματος (gbpusdmonthly.dat).

5. Για τη χρονοσειρά των αποδόσεων της τιμής κλεισίματος του δείκτη S&P500 στην άσκηση 3, κάνετε προβλέψεις για μια ημέρα μπροστά για τις τελευταίες 100 μέρες της χρονοσειράς και υπολογίστε το NRMSE. Κάνετε το ίδιο για τους τελευταίους 36 μήνες της χρονοσειράς των πρώτων διαφορών της μηνιαίας τιμής συναλλάγματος βρετανικής λίρας σε δολλάριο της άσκησης 4.

6. Για να αξιολογίσουμε την πρόβλεψη για μια μη-στάσιμη χρονοσειρά συγκρίνουμε την πρόβλεψη από το επιλεγμένο μοντέλο (όπως ARIMA) με την πρόβλεψη που δίνεται απλά από την παρούσα τιμή, δηλαδή $y_n(1) = y_n$ και αναφέρεται ως πρόβλεψη διατήρησης (persistence). Διαιρώντας το rmse της πρόβλεψης με το επιλεγμένο μοντέλο με το rmse της πρόβλεψης διατήρησης προκύπτει το στατιστικό σφάλματος πρόβλεψης που λέγεται και normalized (by) persistence root mean square error (nprmse). Γενικά για πρόβλεψη T βημάτων μπροστά το nprmse δίνεται ως

$$\text{nprmse}(T) = \frac{\sqrt{\dfrac{1}{N-N_1-T+1}\displaystyle\sum_{j=N_1}^{N-T}\left(y_{j+T}-y_j(T)\right)^2}}{\sqrt{\dfrac{1}{N-N_1-T+1}\displaystyle\sum_{j=N_1}^{N-T}\left(y_{j+T}-y_j\right)^2}}$$

   Υπολογίστε το nprmse για μια ημέρα μπροστά για τις τελευταίες 100 μέρες της χρονοσειράς του κλεισίματος του δείκτη S&P500 (δες άσκηση 3 και 5) και για τους τελευταίους 36 μήνες της χρονοσειράς της μηνιαίας τιμής συναλλάγματος βρετανικής λίρας σε δολλάριο (δες άσκηση 4 και 5).