ΕΡΓΑΣΙΑ ΣΤΟ ΜΑΘΗΜΑ

ΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ

Ομάδα: 4 **Σειρά: 8**

Μοντελοποίηση χαοτικής χρονοσειράς με χρήση μοντέλων TSK

Στόχος της εργασίας αυτής είναι να διερευνηθεί η ικανότητα των μοντέλων TSK στην μοντελοποίηση χρονοσειρών. Συγκεκριμένα, επιλέγεται η χαοτική χρονοσειρά Mackey-Glass, η οποία έχει χρησιμοποιηθεί ευρέως στην βιβλιογραφία για την αξιολόγηση της ακρίβειας μοντελοποίησης διαφόρων μοντέλων.

1) **Ορισμός του προβλήματος**: Η χρονοσειρά Mackey-Glass προσομοιώνει την δυναμική του αίματος και χρησιμοποιήθηκε για την μελέτη του φαινομένου της αρρυθμίας. Η χρονοσειρά περιγράφεται από την παρακάτω διαφορική εξίσωση με καθυστέρηση:

$$\dot{x}(t) = \frac{0.2\,x(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1\,x(t) \qquad (1)$$

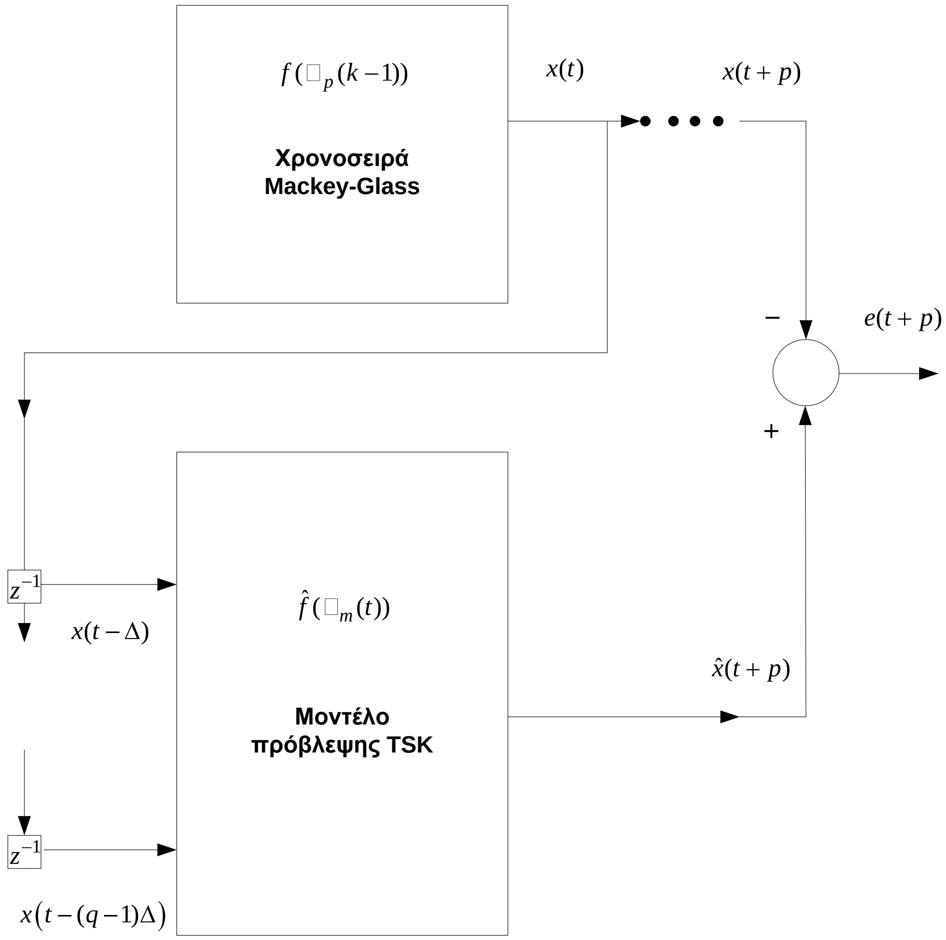
Η χρονοσειρά αυτή για $\tau > 17$ παρουσιάζει μια απεριοδική εξέλιξη στον χρόνο. Αν θεωρήσουμε ότι t είναι η τρέχουσα χρονική στιγμή, στόχος στης εργασίας είναι η ανάπτυξη ενός μοντέλου TSK με σκοπό την πρόβλεψη (prediction) της χρονοσειράς $x(t+p)$ σε κάποιο χρονικό σημείο στο μέλλον $(t+p)$ (future time step), χρησιμοποιώντας παλαιότερη πληροφορία της χρονοσειράς μέχρι την χρονική στιγμή t. Συγκεκριμένα, θεωρούμε q προηγούμενες τιμές $\{x(t-(q-1)\Delta),....,x(t-1)\Delta, x(t)\}$ για τον υπολογισμό της πρόβλεψης $\hat{x}(t+p)$ της $x(t+p)$:

$$\{x(t-(q-1)\Delta),....,x(t-\Delta), x(t)\} \rightarrow \hat{x}(t+p) \qquad (2)$$

όπου Δ είναι το διάστημα διακριτοποίησης στο παρελθόν, και επιλέγεται συνήθως ως $\Delta = 6$.

2) **Δημιουργία μοντέλου**: Δημιουργούμε ένα μοντέλο TSK όπως φαίνεται στο Σχ.1, θεωρώντας ότι $q = 3$. Οι παρελθοντικές τιμές $\{x(t-(q-1)\Delta),....,x(t-1)\Delta, x(t)\}$ αποτελούν τις εισόδους του μοντέλου, $Y_m(t) = \left[x(t), x(t-\Delta),...., x(t-(q-1)\Delta)\right]^T$, ενώ η εκτίμηση $\hat{x}(t+p)$ είναι η έξοδος του μοντέλου. Υποθέτουμε ότι η πραγματική διαδικασία περιγράφεται από μια άγνωστη

συνάρτηση $f(\boldsymbol{Y}_p(k-1))$, όπου η τρέχουσα τιμή $x(t)$ παράγεται με βάση ένα αριθμό παρελθόντων τιμών $Y_p(t) = \left[x(t-1), x(t-2), \ldots, x(t-s)\right]^T$.



Σχ.1

3) **Δημιουργία των συνόλων εκπαίδευσης**: Επιλέγουμε $\tau = 22$ και επιλύουμε την (1) αριθμητικά για $0 \le t \le 3000$, χρησιμοποιώντας την μέθοδο Runge-Kutta τέταρτης τάξεως, θεωρώντας σαν αρχική τιμή $x(0) = 1.2$. Υποθέτουμε ότι $x(t) = 0$ για $t < 0$. Από τα αποτελέσματα της αριθμητικής επίλυσης αμελούμε τις $N_0 = 500$ πρώτες τιμές, έτσι ώστε να απαλειφθεί η επίδραση της αρχικής τιμής στην διακύμανση της χρονοσειράς. Στην συνέχεια δημιουργούμε τα σύνολα D_{trn}, D_{val} και D_{chk}, τα οποία περιλαμβάνουν $N_{trn} = 500$, $N_{val} = 300$ και $N_{chk} = 500$ πρότυπα, αντίστοιχα. Τα πρότυπα εκπαίδευσης για κάθε σύνολο δημιουργούνται θεωρώντας διαδοχικά και συνεχή τμήματα της χρονοσειράς. Συγκεκριμένα, χρησιμοποιούμε το χρονικό διάστημα $N_0 \le t < N_{trn}$ για την δημιουργία του D_{trn}, το διάστημα $N_{trn} \le t < N_{val}$ για την δημιουργία του D_{val}, και το χρονικό διάστημα $N_{val} \le t < N_{chk}$ για την δημιουργία του D_{chk}. Τα σύνολα εκπαίδευσης περιλαμβάνουν πρότυπα εισόδου-εξόδου της μορφής:

$\left\{x(t-12), x(t-6), x(t)\right\} \to \hat{x}(t+p)$.

4) **Αξιολόγηση μοντέλων:** Η ακρίβεια πρόβλεψης αξιολογείται με βάση το σφάλμα MSE, μεταξύ των προβλέψεων που παρέχει το μοντέλο και των πραγματικών τιμών της χρονοσειράς:

$$MSE(\boldsymbol{\theta}^*) = \sigma_e^2 = \frac{1}{N}\sum_{t=1}^{N} e^2(t+p) = \frac{1}{N}\sum_{t=1}^{N}\left\{\hat{x}(t+p) - x(t+p)\right\}^2 \tag{3}$$

Επίσης, μπορεί να χρησιμοποιηθεί ο δείκτης $RMSE$:

$$RMSE = \sqrt{MSE(\boldsymbol{\theta}^*)} = \sqrt{\frac{1}{N}\sum_{t=1}^{N} e^2(t+p)} \tag{4}$$

Τέλος, για πιο εύρωστη αξιολόγηση, θα μπορούσαν να χρησιμοποιηθούν οι δείκτες

$$NMSE = \frac{\sigma_e^2}{\sigma_x^2} = \frac{\frac{1}{N}\sum_{k=1}^{N}\left\{\hat{x}(t+p) - x(t+p)\right\}^2}{\frac{1}{N}\sum_{k=1}^{N}\left\{x(t+p) - m_x\right\}^2} \tag{5}$$

και

$$NDEI = \sqrt{NMSE} = \frac{\sigma_e}{\sigma_x} \tag{6}$$

Ζητούμενα του προβλήματος

1) **Μοντέλα singleton**: Κάθε είσοδος του συστήματος περιγράφεται από 2 ασαφή σύνολα, με bell-shaped συναρτήσεις συμμετοχής. Η αρχική τοποθέτηση θα πρέπει να είναι τέτοια ώστε τα διαδοχικά ασαφή σύνολα σε κάθε είσοδο να παρουσιάζουν επικάλυψη σε βαθμό περίπου 0.5. Σε αυτή την φάση, να διαμορφωθεί ένα μοντέλο με κανόνες singleton (zero-order TSK). Οι παράμετροι του τμήματος συμπεράσματος να αρχικοποιηθούν τυχαία. Να γίνει εκπαίδευση του δικτύου ANFIS χρησιμοποιώντας δύο εκδοχές του αλγορίθμου εκπαίδευσης: αναπροσαρμογή παραμέτρων με χρήση μόνο του BP και χρήση του υβριδικού αλγορίθμου. Η εκπαίδευση του δικτύου θα πρέπει να περατώνεται όταν το MSE_{val} παρουσιάζει την ελάχιστη τιμή.

   Το μοντέλο που θα αναπτυχθεί εδώ, θα πρέπει να παράγει προβλέψεις $\hat{x}(t+p)$ για $p=6$:

   $\left\{x(t-12), x(t-6), x(t)\right\} \to \hat{x}(t+p)$, δηλαδή προβλέψεις p χρονικές στιγμές μπροστά.

   Για τις δύο περιπτώσεις εκπαίδευσης:

   - Να δώσετε τα αντίστοιχα διαγράμματα όπου να αποτυπώνονται οι προβλέψεις του μοντέλου και οι πραγματικές τιμές της χρονοσειράς, για όλα τα πρότυπα του συνόλου εκπαίδευσης, αξιολόγησης και ελέγχου.
   - Επίσης δώστε τα διαγράμματα όπου να φαίνονται τα σφάλματα πρόβλεψης $e(t+p), t=1,\ldots,N$.

- Να δοθούν τα διαγράμματα εκμάθησης (learning curves), όπου να απεικονίζεται το σφάλμα συναρτήσει των επαναλήψεων εκπαίδευσης (iterations).
- Να δοθούν οι τελικές μορφές των ασαφών συνόλων που προέκυψαν μετά την εκπαίδευση και τα τελικά μοντέλα με την μορφή ασαφών κανόνων.
- Τέλος, να δοθούν σε πίνακα οι τιμές του δείκτη $RMSE$, $NMSE$ και $NDEI$ που προέκυψαν.

Να σχολιάσετε τα αποτελέσματα των μοντέλων singleton σε σχέση με την ακρίβεια μοντελοποίησης. Επίσης, να σχολιάσετε την επίδοση των μοντέλων που προέκυψαν μετά από εκπαίδευση μέσω BP και υβριδικού αλγορίθμου, αντίστοιχα, και να ερμηνεύσετε που οφείλεται η διαφορά επίδοσης.

2) **Μοντέλα TSK:** Να ακολουθήσετε την ίδια διαδικασία όπως περιγράφεται στη περίπτωση 1, με την διαφορά ότι τώρα θα πρέπει να αναπτυχθούν μοντέλα TSK με πολυωνυμικές συναρτήσεις στο τμήμα συμπεράσματος. Οι παράμετροι $g_{i,j}$ να αρχικοποιηθούν τυχαία.

- Να σχολιάσετε και πάλι τα αποτελέσματα των μοντέλων TSK σε σχέση με τον τύπο του αλγορίθμου εκπαίδευσης.
- Αν θεωρήσουμε την εκπαίδευση με χρήση του υβριδικού αλγορίθμου, να συγκρίνετε τις αποδόσεις του μοντέλου TSK και singleton. Να ερμηνεύσετε που οφείλετε η ενδεχόμενη διαφορά επίδοσης μεταξύ των δύο μοντέλων.

3) **Υπερ-εκπαίδευση:** Θεωρούμε τώρα ένα μοντέλο TSK με αυξημένη πολυπλοκότητα, όπου κάθε είσοδος διαμερίζεται σε τέσσερα ασαφή σύνολα. Η αρχική τοποθέτηση θα πρέπει να είναι τέτοια ώστε τα διαδοχικά ασαφή σύνολα σε κάθε είσοδο να παρουσιάζουν επικάλυψη σε βαθμό περίπου 0.5. Δημιουργούμε ένα μοντέλο πρόβλεψης της $\hat{x}(t+p)$ με την τιμή του p όπως καθορίστηκε στην περίπτωση 1, ενώ η εκπαίδευση επιτυγχάνεται με χρήση του υβριδικού αλγορίθμου.

- Να δώσετε διαγράμματα όπου να αποτυπώνονται οι προβλέψεις $\hat{x}(t+p)$ του μοντέλου καθώς και οι πραγματικές τιμές $x(t+p)$ της χρονοσειράς, για όλα τα πρότυπα του συνόλου εκπαίδευσης, αξιολόγησης και ελέγχου.
- Επίσης, δώστε τα διαγράμματα όπου να φαίνονται τα σφάλματα πρόβλεψης $e(t+p), t=1,\dots,N$.
- Να δοθούν τα διαγράμματα εκμάθησης (learning curves), όπου να απεικονίζεται το σφάλμα συναρτήσει των επαναλήψεων εκπαίδευσης (iterations).
- Να δοθούν οι τελικές μορφές των ασαφών συνόλων που προέκυψαν μετά την εκπαίδευση.
- Να δοθούν σε πίνακα οι τιμές του δείκτη $RMSE$, $NMSE$ και $NDEI$ που προέκυψαν.

- Να σχολιάσετε τα αποτελέσματα σε σχέση με αυτά του μοντέλου TSK με δύο ασαφή σύνολα σε κάθε είσοδο, και να ανιχνεύσετε αν παρατηρείται το φαινόμενο της υπερ-εκπαίδευσης.

4) Μοντελοποίηση με χρήση Subtractive Clustering: Στην φάση αυτή επαναλαμβάνουμε την διαδικασία των μοντέλων TSK με πολυωνυμικές εξόδους κανόνων (περίπτωση 2), όπου για την αρχικοποίηση των ασαφών μοντέλων χρησιμοποιείται η μεθοδολογία Subtractive Clustering (SC) (genfis2). Ως αλγόριθμος εκπαίδευσης να χρησιμοποιηθεί ο υβριδικός αλγόριθμος. Επιλέξτε κατάλληλα τις παραμέτρους του SC με σκοπό την δημιουργία μεταβλητού αριθμού ασαφών κανόνων.

- Να μελετήσετε την επίδοση των μοντέλων TSK για μεταβλητό αριθμό κανόνων n = 3,4,…, για πρόβλεψη ενός βήματος p = 6.
- Για κάποιες περιπτώσεις, να δείξετε τα αρχικά μοντέλα (κανόνες και ασαφή σύνολα) όπως προκύπτουν μετά τον SC, καθώς και τα τελικά μοντέλα μετά την εκπαίδευση.
- Να διερευνήσετε και πάλι εάν και πότε εμφανίζεται το φαινόμενο της υπερεκπαίδευσης. Επιλέξτε τον κατάλληλο αριθμό κανόνων για βέλτιστη μοντελοποίηση.
- Να εξετάσετε επίσης το ζήτημα της πολυπλοκότητας των μοντέλων με SC και grid partition, σε σχέση με τον αριθμό των παραμέτρων των δικτύων. Συγκρίνετε και σχολιάστε την απόδοση των μοντέλων με SC και grid partition.

5) Πρόβλεψη πολλαπλών βημάτων (Multiple step ahead prediction): Θεωρούμε το βέλτιστο μοντέλο TSK της περίπτωσης 4 (subtractrive clustering), όπου η εκπαίδευση επιτελείται με χρήση του υβριδικού αλγορίθμου. Να παραχθούν από το μοντέλο προβλέψεις μέχρι την χρονική στιγμή $\hat{x}(t+P)$, $P=96$, μπροστά. Οι προβλέψεις αυτές παράγονται, με ένα παράλληλο τρόπο, με βάση τις προβλέψεις $\hat{x}(t+p)$. Συγκεκριμένα, αν θεωρήσουμε ότι $p=6$, η πρόβλεψη $\hat{x}(t+6)$ αναδράται στην είσοδο του μοντέλου, το οποίο παράγει την πρόβλεψη $\hat{x}(t+12)$: $\{x(t-6), x(t), \hat{x}(t+6)\} \rightarrow \hat{x}(t+12)$. Στην συνέχεια, η πρόβλεψη $\hat{x}(t+12)$ αναδράται επίσης στην είσοδο για την δημιουργία της $\hat{x}(t+18)$: $\{x(t), \hat{x}(t+6), \hat{x}(t+12)\} \rightarrow \hat{x}(t+18)$. Η διαδικασία επαναλαμβάνεται μέχρις ότου να φτάσουμε στο όριο $t+P$. Επίσης, η παραπάνω διαδικασία ανακύκλωσης των προβλέψεων εφαρμόζεται για κάθε $t=1,\ldots,N$, και οι προβλέψεις πολλαπλών βημάτων χρησιμοποιούνται για τον υπολογισμό των αντίστοιχων δεικτών MSE, $RMSE$, $NMSE$ και $NDEI$.

- Να δώσετε διαγράμματα όπου να αποτυπώνονται οι προβλέψεις πολλαπλών βημάτων $\hat{x}(t+P)$ του μοντέλου καθώς και πραγματικές τιμές $x(t+P)$ της χρονοσειράς, για όλα τα πρότυπα του συνόλου εκπαίδευσης, αξιολόγησης και ελέγχου.
- Επίσης, δώστε τα διαγράμματα όπου να φαίνονται τα σφάλματα πρόβλεψης $e(t+p), t=1,\ldots,N$, για μερικά βήματα του p από την αρχική στιγμή ως το όριο $t+P$.
- Να σχολιάσετε τα αποτελέσματα πρόβλεψης σε σχέση με αυτά των προβλέψεων $\hat{x}(t+p)$. Δώστε την ερμηνεία σας για την διαφορά των επιδόσεων.