

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

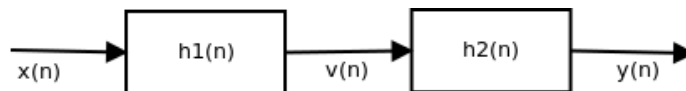
Επεξεργασία Φωνής και Φυσικής Γλώσσας

Εαρινό εξάμηνο 2021-2022

1^η Σειρά Αναλυτικών Ασκήσεων

Άσκηση 1

Θεωρήστε 2 χρονικά αμετάβλητα γραμμικά συστήματα, όπως φαίνονται στο παρακάτω σχήμα, δηλαδή η έξοδος του πρώτου συστήματος είναι η είσοδος του δεύτερου.



1. Δείξτε ότι η χρονική απόκριση του συνολικού συστήματος είναι

$$h(n) = h_1(n) * h_2(n) \quad (1)$$

2. Δείξτε ότι

$$h_1(n) * h_2(n) = h_2(n) * h_1(n) \quad (2)$$

άρα η συνολική χρονική απόκριση δεν εξαρτάται από την σειρά με την οποία εμφανίζονται τα συστήματα.

3. Θεωρείστε τη συνάρτηση

$$H(z) = \left(\sum_{r=0}^M b_r z^{-r} \right) \left(\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right) = H_1(z) H_2(z) \quad (3)$$

δηλαδή σαν σειρά δυο συστημάτων. Γράψτε τις εξισώσεις διαφορών του ολικού συστήματος από αυτήν την οπτική.

4. Τώρα θεωρείστε τα δυο συστήματα του ερωτήματος (3) με την ανάποδη σειρά, δηλαδή:

$$H(z) = H_2(z) H_1(z) \quad (4)$$

Άσκηση 2

Όπως φαίνεται στην παρακάτω εξίσωση, η συνάρτηση αυτοσυσχέτισης είναι

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m) w(n-m) x(m+k) w(n-k-m) \quad (5)$$

1. Ορίστε ότι

$$R_n(k) = R_n(-k) \quad (6)$$

δηλαδή ότι η $R_n(k)$ είναι μια άρτια συνάρτηση του k .

2. Δείξτε ότι η $R_n(k)$ μπορεί να εκφραστεί ως

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m)x(m-k)h_k(n-m) \quad (7)$$

όπου

$$h_k(n) = w(n)w(n+k) \quad (8)$$

3. Υποθέστε ότι

$$w(n) = \begin{cases} \alpha^n & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \quad (9)$$

Βρείτε την κρουστική απόκριση $h_k(n)$.

4. Βρείτε το μετασχηματισμό z του $h_k(n)$ από το προηγούμενο ερώτημα και εκφράστε το $R_n(k)$ με αναδρομική συνάρτηση βάσει αυτού.

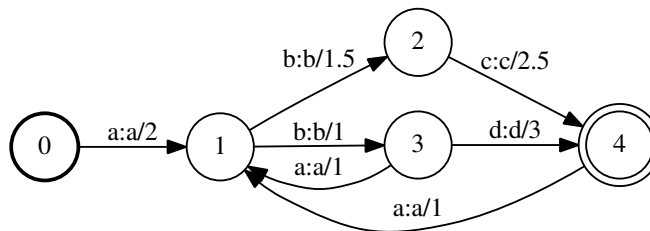
5. Επαναλάβετε τα βήματα (3) και (4) για

$$w(n) = \begin{cases} n\alpha^n & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \quad (10)$$

Άσκηση 3

Για την μηχανή πεπερασμένης κατάστασης που ακολουθεί

1. Ποια είναι η κανονική έκφραση που αντιστοιχεί στην μηχανή;
2. Ποια είναι η πιο πιθανή γραμματοσειρά που αποδέχεται η μηχανή εφόσον χρησιμοποιούμε τον τροπικό ημιδακτύλιο; (tropical semiring collect operation is min, extend operation is +.) Σημείωση: το κόστος των (πιθανών) τελικών καταστάσεων συνυπολογίζεται μόνο εφόσον αυτή η κατάσταση είναι όντως τελική. Το κόστος μίας μη τελικής κατάστασης συνυπολογίζεται κάθε φορά που περνάμε από αυτήν. Το κόστος της κατάστασης 3 είναι 5 και της κατάστασης 4 είναι 0.
3. Ποιο είναι το κόστος της γραμματοσειράς abcababd;
4. Ποια είναι η ισοδύναμη ντετερμινιστική μηχανή χωρίς κόστος;
5. Ποια είναι η ισοδύναμη ντετερμινιστική μηχανή με κόστος;



Άσκηση 4

Δίδεται το εξής αλφάβητο $\Sigma = \{A, B, C, D, E, F\}$.

1. Σχεδιάστε τον μετατροπέα (transducer) που υλοποιεί την απόσταση Levenshtein, δηλαδή $d(x, x) = 0$ και $d(x, \epsilon) = d(\epsilon, x) = d(x, y) = 1$ όπου x και y είναι διαφορετικά γράμματα του αλφάβητου Σ .
2. Ποια είναι η καλύτερη (ποιο φτηνή) αντιστοίχιση ανάμεσα στις γραμματοσειρές EDBAEDC και CDFABEA; Πώς χρησιμοποιήσατε το μετατροπέα από την ερώτηση (1);
3. Ποια είναι η δεύτερη καλύτερη αντιστοιχία ανάμεσα στις γραμματοσειρές της ερώτησης (2);

Άσκηση 5

Δίνονται τα ακόλουθα τμήματα προτάσεων που χρησιμοποιούν το λεξικό

$L = \{\text{lovely, grand, mother, grandmother}\}$:

... lovely mother grand grandmother lovely grandmother grand mother ... (5 φορές)

... lovely mother lovely grandmother lovely ... (7 φορές)

... mother grandmother mother... (2 φορές)

... lovely grand lovely... (1 φορά)

1. Υπολογίστε το γλωσσικό μοντέλο bigram και την αντίστοιχη μηχανή πεπερασμένης κατάστασης με κόστος $-\log P$. Χρησιμοποιήστε back-off για τα bigrams που δεν παρατηρούνται στις παραπάνω προτάσεις.
2. Ποια είναι η πιο πιθανή σειρά από λέξεις για την πρόταση χωρίς κενά: lovelygrandmothergrandmother;

Άσκηση 6

Θεωρήστε ένα all pole μοντέλο με συνάρτηση μεταφοράς της μορφής

$$V(z) = \frac{1}{\prod_{k=1}^q (1 - c_k z^{-1})(1 - c_k^* z^{-1})} \quad (11)$$

όπου

$$c_k = r_k e^{j\theta_k} \quad (12)$$

Δείξτε ότι το αντίστοιχο cepstrum είναι

$$\hat{v}(n) = 2 \sum_{k=1}^q \frac{(r_k)^n}{n} \cos(\theta_k n) \quad (13)$$

Άσκηση 7

Σύμφωνα με την υπόθεση της κατανεμημένης έννοιας των λέξεων, οι λέξεις που απαντώνται σε παρόμοια περιεχόμενα (contexts) θα πρέπει να έχουν παρόμοιο νόημα. Αυτή η υπόθεση αποτελεί την βάση για το μοντέλο Skipgram των Mikolov et al., το οποίο συνιστά αποδοτικό τρόπο μάθησης των σημασιών των λέξεων ως πυκνών διανυσματικών αναπαραστάσεων από μη δομημένο κείμενο. Στόχος του skipgram είναι να μάθει την κατανομή πιθανότητας $P(C | T)$ όπου δοθείσης μιας λέξης στόχου w_t , εκτιμούμε την πιθανότητα να βρίσκεται μια λέξη περικειμένου w_c στο παράθυρο περικειμένου της w_t . Η κατανομή του πιθανοτικού μοντέλου παραμετροποιείται ως εξής:

$$P(C = w_c | T = w_t) = \frac{\exp(\mathbf{u}_{w_c}^\top \cdot \mathbf{v}_{w_t})}{\sum_{w' \in \mathcal{V}} \exp(\mathbf{u}_{w'}^\top \cdot \mathbf{v}_{w_t})} \quad (14)$$

όπου τα διανύσματα \mathbf{u}_{w_c} και \mathbf{v}_{w_t} αντιπροσωπεύουν την λέξη περικειμένου w_c και την λέξη στόχου w_t αντίστοιχα. Σημειώστε τη χρήση της συνάρτησης softmax και πώς το πρόβλημα μάθησης embeddings σε αυτό το μοντέλο έχει μετατραπεί σε πρόβλημα ταξινόμησης. Τα διανύσματα για όλες τις λέξεις του λεξιλογίου \mathcal{V} μπορούν να αναπαρασταθούν σε συμπαγή μορφή

με δύο πίνακες \mathbf{U} and \mathbf{V} , όπου το διάνυσμα στην j -στη στήλη στον \mathbf{U} και στον \mathbf{V} αντιστοιχούν στα διανύσματα περικειμένου και στόχου για την j -οστή λέξη στο \mathcal{V} . Σημειώστε ότι οι \mathbf{U} και \mathbf{V} είναι οι παράμετροι του μοντέλου. Να απαντήσετε στις ακόλουθες ερωτήσεις για το μοντέλο Skipgram.

1. Το σφάλμα διασταυρούμενης εντροπίας μεταξύ δύο πιθανοτικών κατανομών p και q , δίνεται από:

$$L_{CE}(p, q) = - \sum_m p_m \log(q_m). \quad (15)$$

Για δοθείσα λέξη στόχου w_t , μπορούμε να θεωρήσουμε ότι η αληθινή κατανομή \mathbf{y} είναι το διάνυσμα one-hot μεγέθους $|\mathcal{V}|$ με τιμή 1 μόνο στην θέση της πραγματικής λέξης περικειμένου w_c , και τιμή 0 αλλού. Η προβλεπόμενη κατανομή $\hat{\mathbf{y}}$ (ίδιο μήκος με την \mathbf{y}) είναι η κατανομή πιθανότητας $P(C | T = w_t)$. Η j -οστή συνιστώσα σε αυτά τα διανύσματα είναι η πιθανότητα της j -οστής λέξης του \mathcal{V} να είναι λέξη περικειμένου. Γράψτε μια απλοποιημένη μορφή του σφάλματος της διασταυρούμενης εντροπίας, $L_{CE}(\mathbf{y}, \hat{\mathbf{y}})$, για το μοντέλο Skipgram σε ένα μόνο ζεύγος λέξεων w_c και w_t . Η απάντησή σας θα πρέπει να είναι ως προς $P(C = w_c | T = w_t)$.

2. Βρείτε την κλίση (gradient) του σφάλματος διασταυρούμενης εντροπίας που υπολογίστηκε στο βήμα 1 ως προς το διάνυσμα της λέξης στόχου \mathbf{v}_{w_t} . Η απάντησή σας θα πρέπει να είναι ως προς $\mathbf{y}, \hat{\mathbf{y}}$ και \mathbf{U} .
3. Βρείτε την κλίση (gradient) του σφάλματος διασταυρούμενης εντροπίας που υπολογίστηκε στο βήμα 1 ως προς το κάθε διάνυσμα περικειμένου \mathbf{u}_{w_c} . Κάντε το για τις δύο περιπτώσεις $C = w_c$ (πραγματική λέξη περικειμένου) και $C \neq w_c$ (όλες οι υπόλοιπες λέξεις). Η απάντησή σας θα πρέπει να είναι ως προς $\mathbf{y}, \hat{\mathbf{y}}$ και \mathbf{v}_{w_t} .