Επεξεργασία Φωνής & Φυσικής Γλώσσας - ΡΟΗ Σ $7^{ο}$ Εξάμηνο

$1^{η}$ Σειρά Αναλυτικών Ασκήσεων - Ακ. Έτος: 2020-2021

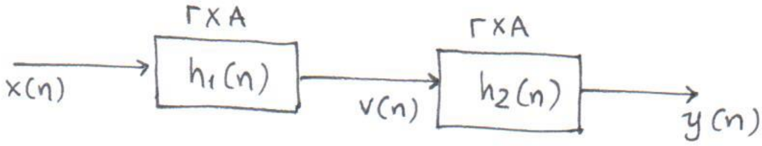
Ονοματεπώνυμο: Χρήστος Δημόπουλος

Αρ. Μητρώου: 031 17 037

E-mail: chrisdim1999@gmail.com

Άσκηση 1:

Θεωρούμε δύο χρονικά αμετάβλητα γραμμικά συστήματα, όπως φαίνονται στο παρακάτω σχήμα, δηλαδή η έξοδος του πρώτου συστήματος είναι η είσοδος του δεύτερου:



1) Έστω $h(n) = h_1(n) * h_2(n) = \sum_{k=-\infty}^{+\infty} h_1[k] \cdot h_2(n-k)$

Έχουμε:

$$v(n) = x(n) * h_1(n) = \sum_{\ell=-\infty}^{+\infty} x(\ell) \cdot h_1(n-\ell) \quad (1)$$

επίσης:

$$x(n) * h(n) = \sum_{\ell=-\infty}^{+\infty} x(\ell) \cdot h(n-\ell) = \sum_{\ell=-\infty}^{+\infty} x(\ell) \cdot \left[\sum_{k=-\infty}^{+\infty} h_1(k) \cdot h_2(n-\ell-k) \right]$$

Ακολούθως, ισχύει:

$$y(n) = v(n) * h_2(n) = \sum_{k=-\infty}^{+\infty} v(k) \cdot h_2(n-k) \overset{(1)}{=} \sum_{k=-\infty}^{+\infty} \left(\sum_{\ell=-\infty}^{+\infty} x(\ell) h_1(k-\ell) \right) h_2(n-k)$$

$$= \sum_{\ell=-\infty}^{+\infty} x(\ell) \cdot \left[\sum_{k=-\infty}^{+\infty} h_1(k-\ell) \cdot h_2(n-k) \right]$$

Θέτοντας $k' = k - \ell \Rightarrow k = k' + \ell$, λαμβάνουμε το εξής:

$$y(n) = \sum_{l=-\infty}^{+\infty} x(l) \left[\sum_{k'=-\infty}^{+\infty} h_1(k') \cdot h_2(n-k'-l) \right]$$

$$\Rightarrow \boxed{y(n) = x(n) * h(n)}, \text{ όπου } h(n) = h_1(n) * h_2(n)$$

Συνεπώς, η κρουστική απόκριση του συνολικού συστήματος είναι η $h(n) = h_1(n) * h_2(n)$.

2) Εξ ορισμού, έχουμε :

$$h_1(n) * h_2(n) = \sum_{k=-\infty}^{+\infty} h_1(k) \cdot h_2(n-k)$$

$$h_2(n) * h_1(n) = \sum_{k=-\infty}^{+\infty} h_2(k) \cdot h_1(n-k)$$

Θέτοντας $k = n-k'$ στο πρώτο άθροισμα, λαμβάνουμε :

$$h_1(n) * h_2(n) = \sum_{k'=-\infty}^{+\infty} h_1(n-k') \cdot h_2(k') = \sum_{k'=-\infty}^{+\infty} h_2(k') \cdot h_1(n-k') = h_2(n) * h_1(n)$$

Επομένως, ισχύει ότι :

$$\boxed{h_1(n) * h_2(n) = h_2(n) * h_1(n)}$$

άρα η συνολική κρουστική απόκριση $h(n)$ δεν εξαρτάται από τη σειρά με την οποία εμφανίζονται τα συστήματα.

3) Θεωρούμε τη συνάρτηση :

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

δηλαδή σαν σειρά δύο συστημάτων.

Εξ ορισμού, για τη συνάρτηση μεταφοράς έχουμε : $H(z) \triangleq \frac{Y(z)}{X(z)}$.

Επομένως :

$$\left(1 - \sum_{k=1}^{N} a_k z^{-k}\right) \cdot Y(z) = \left(\sum_{r=0}^{M} b_r z^{-r}\right) X(z) \qquad (I)$$

Time - Shift Property : $\mathcal{Z}\{x[n-n_0]\} = z^{-n_0} \cdot X(z)$, $ROC = R_x$.

Εφαρμόζοντας Αντίστροφο Μετασχηματισμό $\mathcal{Z}$ στην (I) λαμβάνουμε :

$$\boxed{y(n) - \sum_{k=1}^{N} a_k \cdot y(n-k) = \sum_{r=0}^{M} b_r \cdot x(n-r)}$$

Εξίσωση Διαφορών Ολικού Συστήματος

4) Τώρα θεωρούμε τα δύο συστήματα του ερωτήματος (3) με την ανάποδη σειρά, δηλαδή : $H(z) = H_2(z) \cdot H_1(z) = \left(\dfrac{1}{1 - \sum_{k=1}^{N} a_k z^{-k}}\right) \cdot \left(\sum_{r=0}^{M} b_r z^{-r}\right)$

Ομοίως με πριν :

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{r=0}^{M} b_r z^{-r}}{\left(1 - \sum_{k=1}^{N} a_k z^{-k}\right)} \Rightarrow \left(1 - \sum_{k=1}^{N} a_k z^{-k}\right) Y(z) = \left(\sum_{r=0}^{M} b_r z^{-r}\right) X(z)$$

Παίρνοντας τον Αντίστροφο $\mathcal{Z}$ Μετασχηματισμό, έχουμε :

$$y(n) - \sum_{k=1}^{N} a_k \cdot y(n-k) = \sum_{r=0}^{M} b_r \, x(n-r)$$

Δηλαδή, η ίδια εξίσωση διαφορών με αυτήν του προηγούμενου ερωτήματος .

Άσκηση 2:

Δίνεται η συνάρτηση αυτοσυσχέτισης:

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

1) Έχουμε ότι:

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

Στο παραπάνω άθροισμα, θέτω $m' = m-k \Rightarrow m = m'+k$ και λαμβάνω:

$$R_n(-k) = \sum_{m'=-\infty}^{+\infty} x(m'+k)w(n-m'-k)\,x(m'+\cancel{k}-\cancel{k})w(n+\cancel{k}-m'-\cancel{k})$$

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

Έχουμε, λοιπόν, δείξει ότι: $\boxed{R_n(k) = R_n(-k)}$

δηλαδή η συνάρτηση αυτοσυσχέτισης είναι άρτια συνάρτηση του k.

2) Θέτουμε:

$$h_k(n) = w(n)w(n+k)$$

οπότε: $h_k(n-m) = w(n-m)\cdot w(n-m+k)$ (1)

Επίσης, έχουμε:

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

$$= \sum_{m=-\infty}^{+\infty} x(m)\cdot x(m-k)\cdot\left[w(n-m)\cdot w(n+k-m)\right]$$

$$\overset{(1)}{=} \sum_{m=-\infty}^{+\infty} x(m)\cdot x(m-k)\cdot h_k(n-m)$$

Όμως, από προηγούμενο ερώτημα, δείξαμε ότι $R_n(k) = R_n(-k)$.

Συνεπώς, θα ισχύει:

$$\boxed{R_n(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot h_k(n-m)}$$

όπου:

$$h_k(n) = w(n)\cdot w(n+k)$$

3) Θεωρούμε ότι:

$$w(n) = \begin{cases} a^n, & \text{if } n \geq 0 \\ 0, & \text{if } n < 0 \end{cases} \Rightarrow w(n) = a^n \cdot u(n)$$

, όπου $u(n)$ μοναδιαία βηματική συνάρτηση.

Γενικά, ισχύει: $u(n-n_0) = \begin{cases} 1, & \text{if } n \geq n_0 \\ 0, & \text{if } n < n_0 \end{cases}$ οπότε έχουμε $u(n+k) = \begin{cases} 1, & \text{if } n \geq -k \\ 0, & \text{if } n < -k \end{cases}$

Επομένως:

$$h_k(n) = w(n)\cdot w(n+k) = a^n \cdot u(n) \cdot a^{n+k} u(n+k) = a^{2n+k} u(n)\cdot u(n+k)$$

$$\Rightarrow h_k(n) = \begin{cases} a^{2n+k}\cdot u(n), & \text{if } k \geq 0 \\ a^{2n+k}\cdot u(n+k), & \text{if } k < 0 \end{cases}$$

4) Γνωρίζουμε ότι:

$$a^n u(n) \overset{\mathcal{Z}}{\rightleftharpoons} \frac{z}{z-a}, \quad |z| > |a|$$

· <u>Για $k \geq 0$</u>:

$$h_k(n) = a^{2n+k} u(n) = a^{+k} (a^2)^n u(n)$$

Επομένως: $$\mathcal{Z}[h_k(n)] = a^k \cdot \frac{z}{z-a^2}, \quad |z|>|a^2|$$

· Για $k<0$:

$$h_k(n) = a^{2n+k} u(n+k) = a^{-k} \cdot a^{2(n+k)} u(n+k) = a^{-k} \cdot (a^2)^{n+k} u(n+k).$$

Επομένως: $$\mathcal{Z}[h_k(n)] = a^{-k} \frac{z}{z-a^2} \cdot z^k, \quad |z|>|a^2|$$

Συνεπώς, έχουμε:

$$H_k(z) = \mathcal{Z}[h_k(n)] = \begin{cases} a^k \dfrac{z}{z-a^2}, & |z|>|a^2| \text{ if } k \geq 0. \\ a^{-k} \dfrac{z}{z-a^2} \cdot z^k, & |z|>|a^2| \text{ if } k<0. \end{cases}$$

Στη συνέχεια, θα εκφράσουμε το $R_n(k)$ με αναδρομική συνάρτηση βάσει του $h_k(n)$. Ειδικότερα:

· Για $k \geq 0$: $h_k(n) = a^{2n+k} u(n)$

$$h_k(n+1) = a^{2(n+1)+k} u(n+1) = a^2 \cdot a^{2n+k} \cdot u(n+1) =$$

$$= a^k \cdot \delta(n+1) + a^2 \cdot h_k(n)$$

· Για $k<0$: $h_k(n) = a^{2n+k} u(n+k)$

$$h_k(n+1) = a^{2(n+1)+k} u(n+k+1) = a^2 a^{2n+k} u(n+k+1)$$

$$= a^{-k} \cdot \delta(n+k+1) + a^2 \cdot h_k(n)$$

Επομένως: $$R_n(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot h_k(n-m)$$

$$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\,h_k(n-m+1)$$

Για $k \geq 0$: $$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot\left[a^k\,\delta(n-m+1) + a^2\cdot h_k(n-m)\right]$$

$$\Rightarrow R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot a^k\,\delta(n-m+1) + a^2\cdot R_n(k)$$

$$\Rightarrow \boxed{R_{n+1}(k) = x(n+1)\,x(n+1-k)\cdot a^k + a^2\,R_n(k)}$$

Για $k < 0$: $$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot\left[a^{-k}\,\delta(n-m+k+1) + a^2\,h_k(n-m)\right]$$

$$\Rightarrow R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\,x(m-k)\cdot a^{-k}\cdot\delta(n-m+k+1) + a^2\cdot R_n(k)$$

$$\Rightarrow \boxed{R_{n+1}(k) = x(n+k+1)\,x(n+1)\cdot a^{-k} + a^2\cdot R_n(k)}$$

Παρατηρούμε ότι ισχύει η άρτια συμμετρία ως προς k, δηλ. $R_n(k) = R_n(-k)$.

Συνεπώς:

$$\boxed{R_{n+1}(k) = x(n+1)\cdot x(n+1-k)\cdot a^k + a^2\cdot R_n(k)}$$

Αναδρομική Σχέση Αυτοσυσχέτισης

5) Έστω τώρα ότι:

$$w'(n) = \begin{cases} n \cdot a^n, & \text{if } n \geq 0 \\ 0, & \text{if } n < 0 \end{cases} \Rightarrow w'(n) = n \cdot a^n \cdot u(n)$$

Επομένως:

$$h'_k(n) = w(n) \cdot w(n+k) = n \cdot a^n \cdot u(n) \cdot \left[(n+k) \cdot a^{n+k} u(n+k)\right] =$$

$$= n^2 \cdot a^{2n+k} u(n) \cdot u(n+k) + k \cdot n \cdot a^{2n+k} u(n) \cdot u(n+k)$$

$$\Rightarrow h'_k(n) = \begin{cases} n^2 \cdot a^{2n+k} \cdot u(n) + k \cdot n\, a^{2n+k} u(n), & \text{if } k \geq 0 \\ n^2 \cdot a^{2n+k} \cdot u(n+k) + k \cdot n \cdot a^{2n+k} \cdot u(n+k), & \text{if } k < 0 \end{cases}$$

$$\Rightarrow h'_k(n) = \begin{cases} (n+k) \cdot n \cdot a^{2n+k} \cdot u(n), & \text{if } k \geq 0 \\ (n+k) \cdot n \cdot a^{2n+k} \cdot u(n+k), & \text{if } k < 0 \end{cases}$$

Γνωρίζουμε ότι ισχύει: $n \cdot x(n) \overset{\mathcal{Z}}{\rightleftarrows} -z \dfrac{dX(z)}{dz}$

Παρατηρούμε, λοιπόν, ότι ισχύει: $h'_k(n) = n \cdot h_k(n)$. Επομένως, είχαμε βρει:

$$H_k(z) = \begin{cases} a^k \dfrac{z}{z-a^2}, & |z| > |a^2| \quad \text{if } k \geq 0 \\ a^{-k} \dfrac{z^{k+1}}{z-a^2}, & |z| > |a^2| \quad \text{if } k < 0 \end{cases}$$

· <u>Για $k \geq 0$</u>: $\dfrac{d}{dz}\left(\dfrac{z}{z-a^2}\right) = \dfrac{\cancel{z} - a^2 - \cancel{z}}{(z-a^2)^2} = \dfrac{-a^2}{(z-a^2)^2}$

οπότε: $H'_k(z) = a^k \cdot a^2 \dfrac{z}{(z-a^2)^2}, \quad |z| > |a^2|$

Για $k<0$: $\frac{d}{dz}\left(\frac{z^{k+1}}{z-a^2}\right) = \frac{(k+1)z^k\cdot(z-a^2) - z^{k+1}}{(z-a^2)^2}$

οπότε: $H'_k(z) = \frac{z^{k+2} - (k+1)z^{k+1}(z-a^2)}{(z-a^2)^2}$

Συνολικά:

$$H'_k(z) = \begin{cases} a^2\cdot a^k \frac{z}{(z-a^2)^2}, & |z|>|a^2|, \text{ if } k\geq 0 \\ a^{-k}\,\frac{z^{k+2}-(k+1)z^{k+1}(z-a^2)}{(z-a^2)^2}, & |z|>|a^2|, \text{ if } k<0 \end{cases}$$

Υπολογισμός αναδρομικής σχέσης:

Για $k\geq 0$: $h'_k(n) = (n+k)\cdot n\cdot a^{2n+k}\cdot u(n)$

$$h'_k(n+1) = (n+1+k)\cdot(n+1)\cdot a^{2n+2+k}\cdot u(n+1)$$

$$= (n^2+n+n\cdot k+n+1+k)\cdot a^{2n+k}\cdot a^2\cdot u(n+1)$$

$$= (n^2+2n+n\cdot k+k+1)\cdot a^{2n+k}\, a^2\cdot u(n+1)$$

$$= (2n+k+1)\cdot a^{2n+k}\cdot a^2\cdot u(n+1) + a^2\cdot(n+k)\cdot n\cdot a^{2n+k}\cdot u(n+1)$$

$$= \cancel{(k-1)\cdot a^{k-2}\cdot a^2}\cdot\delta(n+1) + \cancel{a^2\cdot(k-1)\cdot(-1)\cdot a^{k-2}}\cdot\delta(n+1) +$$

$$+ (2n+k+1)\, a^{2n+k}\cdot a^2\cdot u(n) + a^2(n+k)\cdot n\cdot a^{2n+k}\cdot u(n)$$

$$\Rightarrow h'_k(n+1) = (2n+k+1)\cdot a^{2(n+1)+k}\cdot u(n) + a^2\cdot h'_k(n)$$

· Για $k<0$:

$$h'_k(n) = (n+k)\cdot n \cdot a^{2n+k} \cdot u(n+k)$$

$$h'_k(n+1) = (2n+k+1)\, a^{2(n+1)+k} \cdot u(n+k) + a^2 \cdot h'_k(n)$$

Επομένως: $R_n(k) = \sum_{m=-\infty}^{+\infty} x(m)\, x(m-k) \cdot h'_k(n-m)$

$$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\, x(m-k) \cdot h'_k(n-m+1)$$

Για $k \geq 0$:

$$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\, x(m-k) \cdot \left[(2n-2m+k+1)\, a^{2(n-m+1)+k} \cdot u(n-m) + a^2 \cdot h'_k(n-m) \right]$$

$$\Rightarrow R_{n+1}(k) = \sum_{m=n}^{+\infty} x(m)\, x(m-k) \cdot (2(n-m)+k+1) \cdot a^{2(n-m+1)+k} + a^2 \cdot R_n(k)$$

Για $k<0$:

$$R_{n+1}(k) = \sum_{m=-\infty}^{+\infty} x(m)\, x(m-k) \cdot \left[(2(n-m)+k+1) \cdot a^{2(n-m+1)+k} \cdot u(n-m+k) + a^2 \cdot h'_k(n-m) \right]$$

$$\Rightarrow R_{n+1}(k) = \sum_{m=n+k}^{+\infty} x(m)\, x(m-k) \cdot (2n-2m+k+1) \cdot a^{2(n-m+1)+k} + a^2 \cdot R_n(k)$$

Θέτουμε στο πρώτο άθροισμα $m' = m-k \Rightarrow m = m'+k$ και λαμβάνουμε:

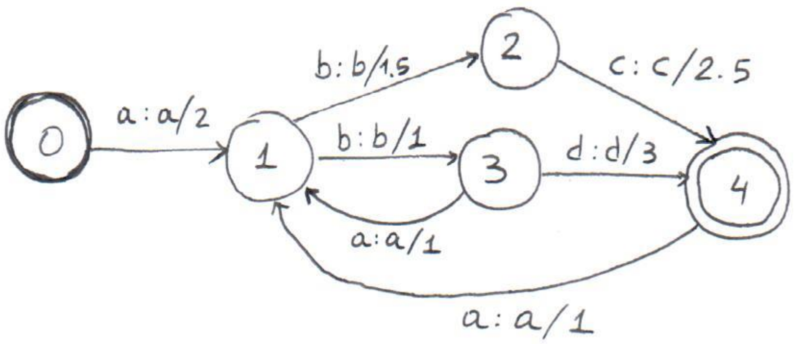
$$R_{n+1}(k) = \sum_{m'=n}^{+\infty} x(m'+k)\, x(m') \cdot (2n-2m'-k+1) \cdot a^{2(n-m+1)-k} + a^2 \cdot R_n(k)$$

Παρατηρούμε ότι και πάλι διατηρείται η άρτια συμμετρία ως προς k, δηλ.

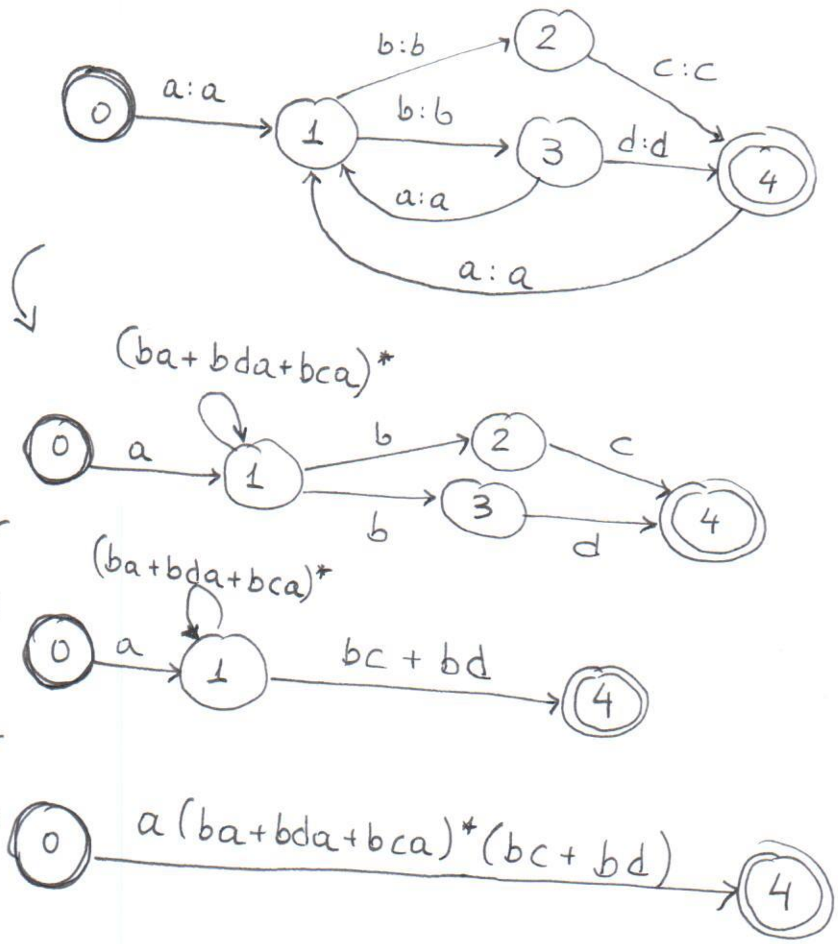
$$R_n(k) = R_n(-k).$$

Άσκηση 3:

Δίνεται η παρακάτω μηχανή πεπερασμένης κατάστασης:



1) Για την εύρεση της κανονικής έκφρασης που αντιστοιχεί στη μηχανή, απαλοίφουμε διαδοχικά καταστάσεις:



Σημείωση:
Ως + θεωρούμε την πράξη της ένωσης $\cup$.

Συνεπώς, η κανονική έκφραση που προκύπτει είναι:

$$\boxed{L = a(ba+bda+bca)^*(bc+bd)}$$

2) Εφόσον χρησιμοποιούμε τον τροπικό ημιδακτύλιο, η πιο πιθανή γραμματοσειρά που αποδέχεται η μηχανή είναι εκείνη που αθροιστικά δίνει το μικρότερο κόστος κατά το μονοπάτι μετάβασης από την αρχική στην τελική κατάσταση (shortest path).

Στη συγκεκριμένη μηχανή, παρατηρούμε ότι οι γραμματοσειρές abd και abc είναι αυτές που αθροιστικά δίνουν το μικρότερο κόστος, καθώς:

(i) $0 \xrightarrow{a:a/2} 1 \xrightarrow{b:b/1} 3 \xrightarrow{d:d/3} 4$

Συνολικό κόστος $C_1 = 2+1+3 = 6$

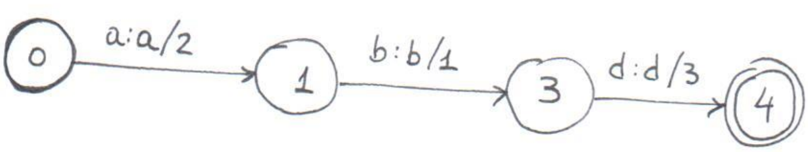
(ii) $0 \xrightarrow{a:a/2} 1 \xrightarrow{b:b/1.5} 2 \xrightarrow{c:c/2.5} 4$

Συνολικό κόστος $C_2 = 2+1.5+2.5 = 6$

Οι συμβολοσειρές "abd", "abc" είναι ισοπίθανες, αφού $C_1 = C_2$? ΟΧΙ!

Ωστόσο, για να δούμε ποιά θα είναι η επόμενη κατάσταση μετά την κατάσταση 1, χρησιμοποιούμε την πράξη $\oplus$ του τροπικού ημιδακτυλίου. Ειδικότερα: $\min\{1, 1.5\} = 1$.

Επομένως, το συντομότερο μονοπάτι θα είναι:



και η πιο πιθανή γραμματοσειρά η "abd".

3) Δίνεται η γραμματοσειρά "abdababc", η οποία επιτυγχάνεται με τις εξής μεταβάσεις:

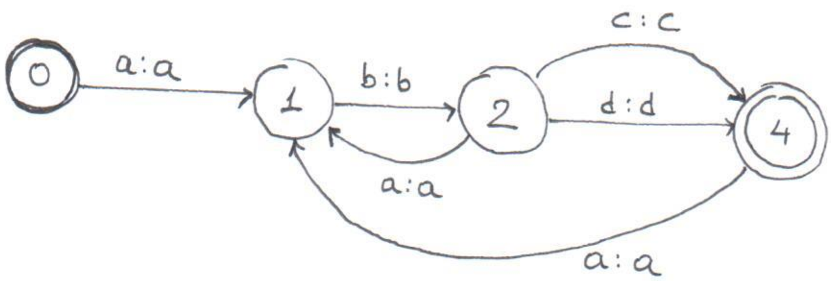
$$0 \xrightarrow{a:a/2} 1 \xrightarrow{b:b/1} 3 \xrightarrow{d:d/3} 4 \xrightarrow{a:a/1} 1 \xrightarrow{b:b/1} 3 \xrightarrow{a:a/1} 1 \longrightarrow$$

$$\xrightarrow{b:b/1.5} 2 \xrightarrow{c:c/2.5} 4$$

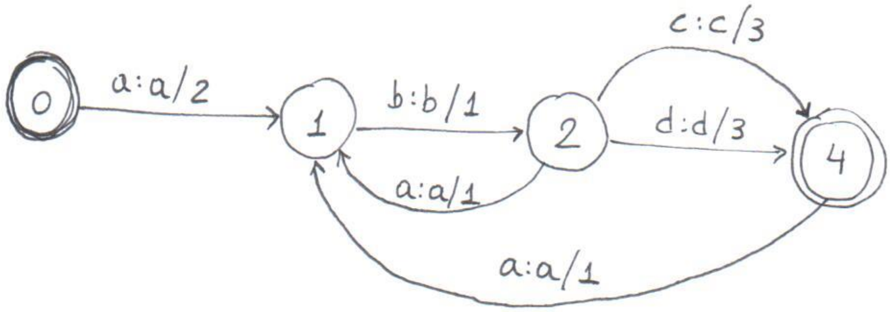
Επομένως, το συνολικό κόστος είναι:

$$C = 2 + 1 + 3 + 1 + 1 + 1 + 1.5 + 2.5 \implies \boxed{C = 13}$$

4) Η ισοδύναμη ντετερμινιστική μηχανή χωρίς κόστος προκύπτει με συγχώνευση των καταστάσεων 2 και 3. Ειδικότερα:



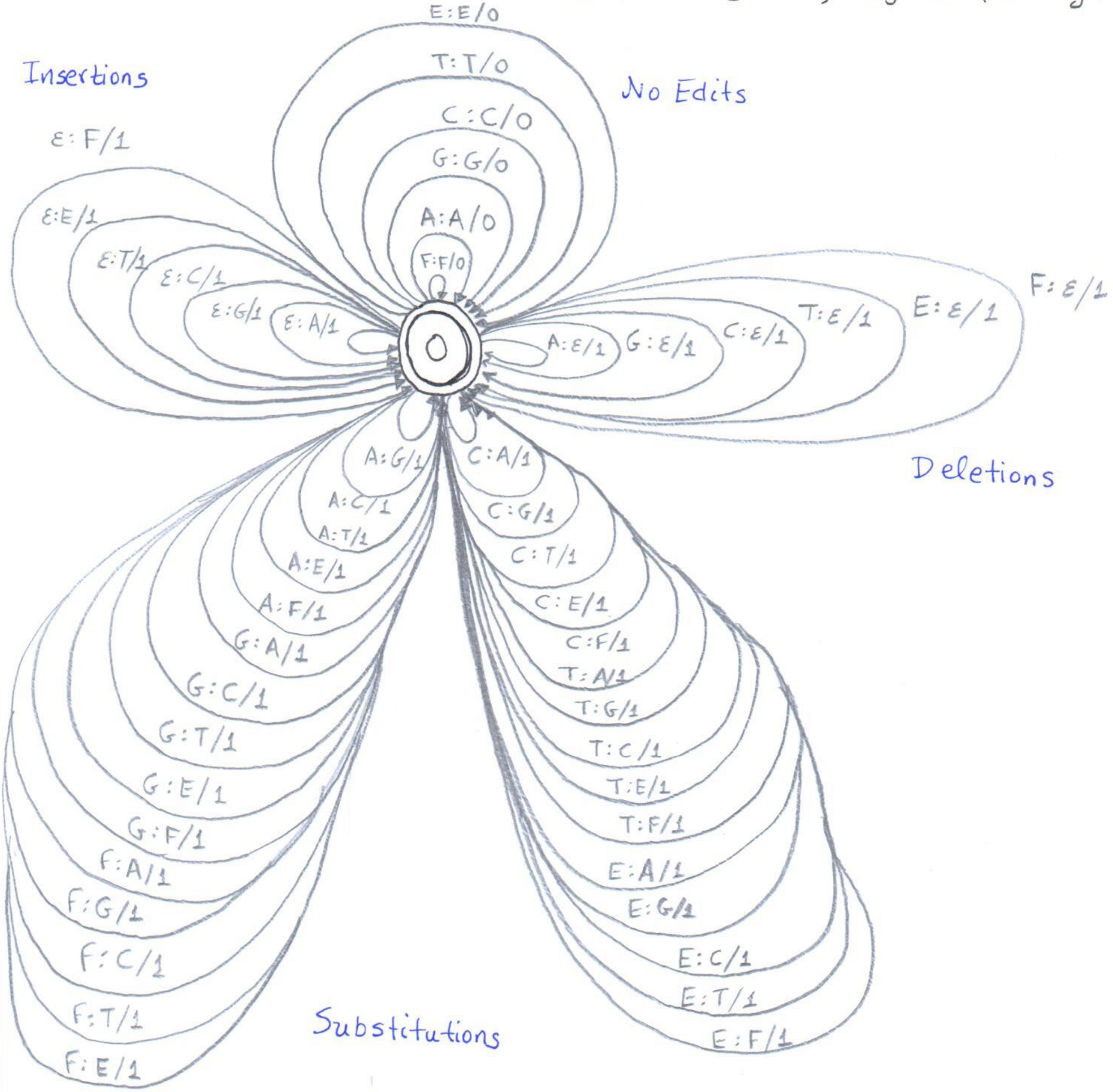
5) Παρακάτω φαίνεται η ισοδύναμη ντετερμινιστική μηχανή με κόστος:



Άσκηση 4:

Δίνεται το αλφάβητο $\Sigma = \{A, G, C, T, E, F\}$

1) Παρακάτω φαίνεται ο μετατροπέας (transducer) που υλοποιεί την απόσταση Levenshtein, δηλαδή $d(x,x) = 0$ και $d(x,\varepsilon) = d(\varepsilon,x) = d(x,y) = 1$, $x,y \in \Sigma$ με $x \neq y$.



2) Έστω S ο transducer που υλοποιεί την απόσταση Levenshtein. Ορίζουμε τις μηχανές $\mathcal{L}$ και V, οι οποίες αποδέχονται τις γραμματοσειρές "AECAGEF" και "TETCGAG" αντίστοιχα με μηδενικό κόστος. Ειδικότερα:

$$(\mathcal{L}):\quad \circledcirc\!\!0 \xrightarrow{A:A} 1 \xrightarrow{E:E} 2 \xrightarrow{C:C} 3 \xrightarrow{A:A} 4 \xrightarrow{G:G} 5 \xrightarrow{E:E} 6 \xrightarrow{F:F} \circledcirc\!\!7$$

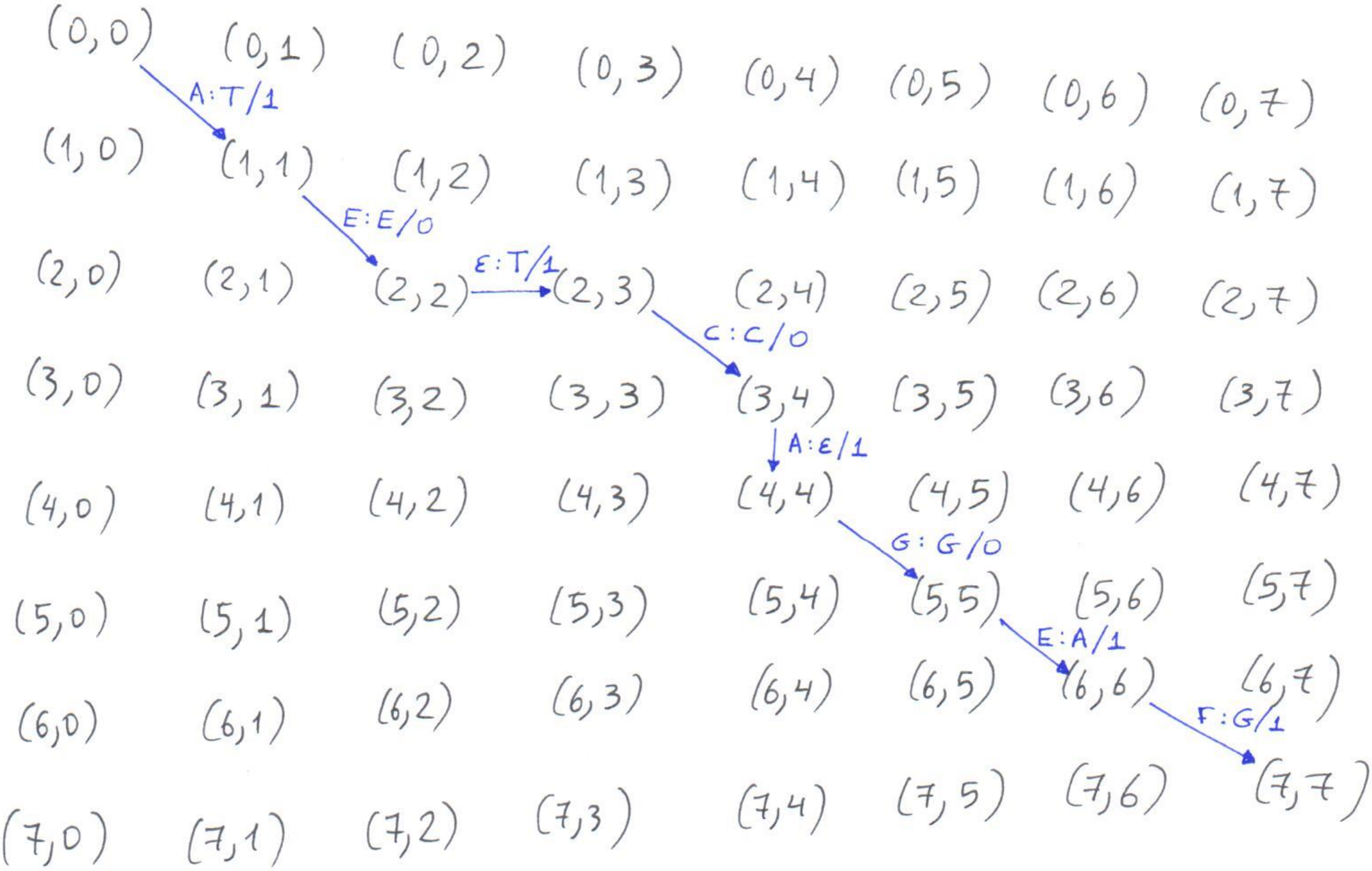
$$(V):\quad \circledcirc\!\!0 \xrightarrow{T:T} 1 \xrightarrow{E:E} 2 \xrightarrow{T:T} 3 \xrightarrow{C:C} 4 \xrightarrow{G:G} 5 \xrightarrow{A:A} 6 \xrightarrow{G:G} \circledcirc\!\!7$$

Τότε, η φτηνότερη αντιστοίχιση της πρώτης γραμματοσειράς στη δεύτερη, δίνεται από το συντομότερο μονοπάτι της μηχανής που προκύπτει από τη σύνθεση $\mathcal{L} \circ S \circ V$. Αναλυτικότερα:

$$
\begin{tikzcd}
(0,0) \arrow[dr, "A:T/1"] & (0,1) & (0,2) & (0,3) & (0,4) & (0,5) & (0,6) & (0,7) \\
(1,0) & (1,1) \arrow[dr, "E:E/0"] & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) & (1,7) \\
(2,0) & (2,1) & (2,2) \arrow[dr, "C:T/1"] & (2,3) & (2,4) & (2,5) & (2,6) & (2,7) \\
(3,0) & (3,1) & (3,2) & (3,3) \arrow[dr, "A:C/1"] & (3,4) & (3,5) & (3,6) & (3,7) \\
(4,0) & (4,1) & (4,2) & (4,3) & (4,4) \arrow[dr, "G:G/0"] & (4,5) & (4,6) & (4,7) \\
(5,0) & (5,1) & (5,2) & (5,3) & (5,4) & (5,5) \arrow[dr, "E:A/1"] & (5,6) & (5,7) \\
(6,0) & (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \arrow[dr, "F:G/1"] & (6,7) \\
(7,0) & (7,1) & (7,2) & (7,3) & (7,4) & (7,5) & (7,6) & (7,7)
\end{tikzcd}
$$

Παρατηρούμε, λοιπόν, ότι η Levenshtein απόσταση μεταξύ των δύο συμβολοσειρών "AECAGEF" και "TETCGAG" είναι $d = 5$ και μπορεί να προκύψει χρησιμοποιώντας ως μοναδικά edits πέντε αντικαταστάσεις συμβόλων, κόστους 1 η καθεμία.

3) Η δεύτερη καλύτερη αντιστοίχιση ανάμεσα στις δύο γραμματοσειρές μπορεί να φανεί παρακάτω:



Επιλέγουμε, δηλαδή, αντί να κάνουμε 2 substitutions $C \to T$ και $A \to C$ συνολικού κόστους 2, να κάνουμε ένα insertion "T", μετά no edit και τελικώς ένα deletion του "A", συνολικού κόστους επίσης 2.

Ως εκ τούτου, η συνολική απόσταση είναι η βέλτιστη, δηλαδή μήκους 5.

Άσκηση 5:

Δίνονται τα ακόλουθα τμήματα προτάσεων που χρησιμοποιούν το λεξικό

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

... lovely mother grand grandmother lovely grandmother grand mother... (4 φορές)
... lovely mother lovely grandmother lovely... (9 φορές)
... mother grandmother mother... (3 φορές)
... lovely grand lovely... (1 φορά)

1) Υπολογίζουμε τις αρχικές πιθανότητες:

$N(lovely) = 2 \cdot 4 + 3 \cdot 9 + 2 \cdot 1 = 37$ φορές

$N(grand) = 2 \cdot 4 + 1 \cdot 1 = 9$ φορές

$N(mother) = 2 \cdot 4 + 1 \cdot 9 + 2 \cdot 3 = 23$ φορές   δηλ. $N(\underline{o}) = 89$

$N(grandmother) = 2 \cdot 4 + 1 \cdot 9 + 1 \cdot 3 = 20$ φορές

Στη συνέχεια, βρίσκουμε:

$N(lovely\ lovely) = 9 - 1 = 8$ φορές

$N(lovely\ grand) = 1$ φορά

$N(lovely\ mother) = 4 + 9 = 13$ φορές

$N(lovely\ grandmother) = 4 + 9 = 13$ φορές

$N(grand\ grand) = 0$ φορές

$N(grand\ lovely) = 1$ φορά

$N(grand\ mother) = 4$ φορές

$N(grand\ grandmother) = 4$ φορές

$N(mother\ mother) = 2$ φορές

$N(mother\ lovely) = 12$ φορές

$N(mother\ grand) = 4$ φορές

$N(mother\ grandmother) = 3$ φορές

$N(\text{grandmother grandmother}) = 0$ φορές

$N(\text{grandmother lovely}) = 4+9 = 13$ φορές

$N(\text{grandmother grand}) = 4$ φορές

$N(\text{grandmother mother}) = 3$ φορές.

Συνεπώς:

- $P(\text{lovely}) = 37/89 \simeq 0.416$
- $P(\text{mother}) = 23/89 \simeq 0.2585$
- $P(\text{grand}) = 9/89 \simeq 0.101$
- $P(\text{grandmother}) = 20/89 \simeq 0.2247$

- $P(\text{lovely} \mid \text{lovely}) = \dfrac{N(\text{lovely lovely})}{N(\text{lovely})} = \dfrac{8}{37}$
- $P(\text{lovely} \mid \text{mother}) = \dfrac{N(\text{lovely mother})}{N(\text{mother})} = \dfrac{13}{23}$
- $P(\text{lovely} \mid \text{grand}) = \dfrac{N(\text{grand lovely})}{N(\text{grand})} = 1/9$
- $P(\text{lovely} \mid \text{grandmother}) = \dfrac{N(\text{grandm. lovely})}{N(\text{grandmother})} = \dfrac{13}{20}$

Αντιστοίχως βρίσκουμε:

$$\boxed{\begin{array}{l}
\cdot\ P(\text{grand} \mid \text{grand}) = 0, \quad P(\text{grand} \mid \text{mother}) = \frac{4}{23}, \quad P(\text{grand} \mid \text{lovely}) = \frac{1}{37} \\
\quad P(\text{grand} \mid \text{grandmother}) = 4/20 \\
\cdot\ P(\text{mother} \mid \text{grand}) = 4/9, \quad P(\text{mother} \mid \text{lovely}) = \frac{13}{37}, \quad P(\text{mother} \mid \text{mother}) = \frac{2}{23} \\
\quad P(\text{mother} \mid \text{grandmother}) = 3/20 \\
\cdot\ P(\text{grandmother} \mid \text{lovely}) = 13/37, \quad P(\text{grandmother} \mid \text{grand}) = 4/9 \\
\quad P(\text{grandmother} \mid \text{mother}) = 3/23, \quad P(\text{grandmother} \mid \text{grandmother}) = 0.
\end{array}}$$

Σημειώνεται ότι τα bigrams "grand grand" και "grandmother grandmother" δεν εμφανίζονται στις προτάσεις, γι' αυτό θα χρησιμοποιήσουμε back-off.

▹ Για την περίπτωση "grand grand":

$$\alpha = 1 - \mathbb{P}(\text{grand}|\text{lovely}) - \mathbb{P}(\text{grand}|\text{mother}) - \mathbb{P}(\text{grand}|\text{grandmother})$$

$$= 1 - 1/37 - 4/23 - 4/20 = \frac{2549}{4255} \simeq 0.599$$

Συνεπώς: $\mathbb{P}(\text{grand}|\text{grand}) = \alpha \cdot \mathbb{P}(\text{grand}) = 0.599 \cdot 0{,}101$

$$\Rightarrow \boxed{\mathbb{P}(\text{grand}|\text{grand}) = 0.06}$$

▹ Για την περίπτωση "grandmother grandmother":

$$\alpha' = 1 - \mathbb{P}(\text{grandmother}|\text{lovely}) - \mathbb{P}(\text{grandmother}|\text{grand}) - \mathbb{P}(\text{grandmother}|\text{mother})$$

$$= 1 - 13/37 - 4/9 - 3/23 = \frac{565}{7659} = 0.074$$

Συνεπώς: $\mathbb{P}(\text{grandmother}|\text{grandmother}) = \alpha' \cdot \mathbb{P}(\text{grandmother}) = 0.074 \cdot 0.2247$

$$\Rightarrow \boxed{\mathbb{P}(\text{grandmother}|\text{grandmother}) = 0.0166}$$

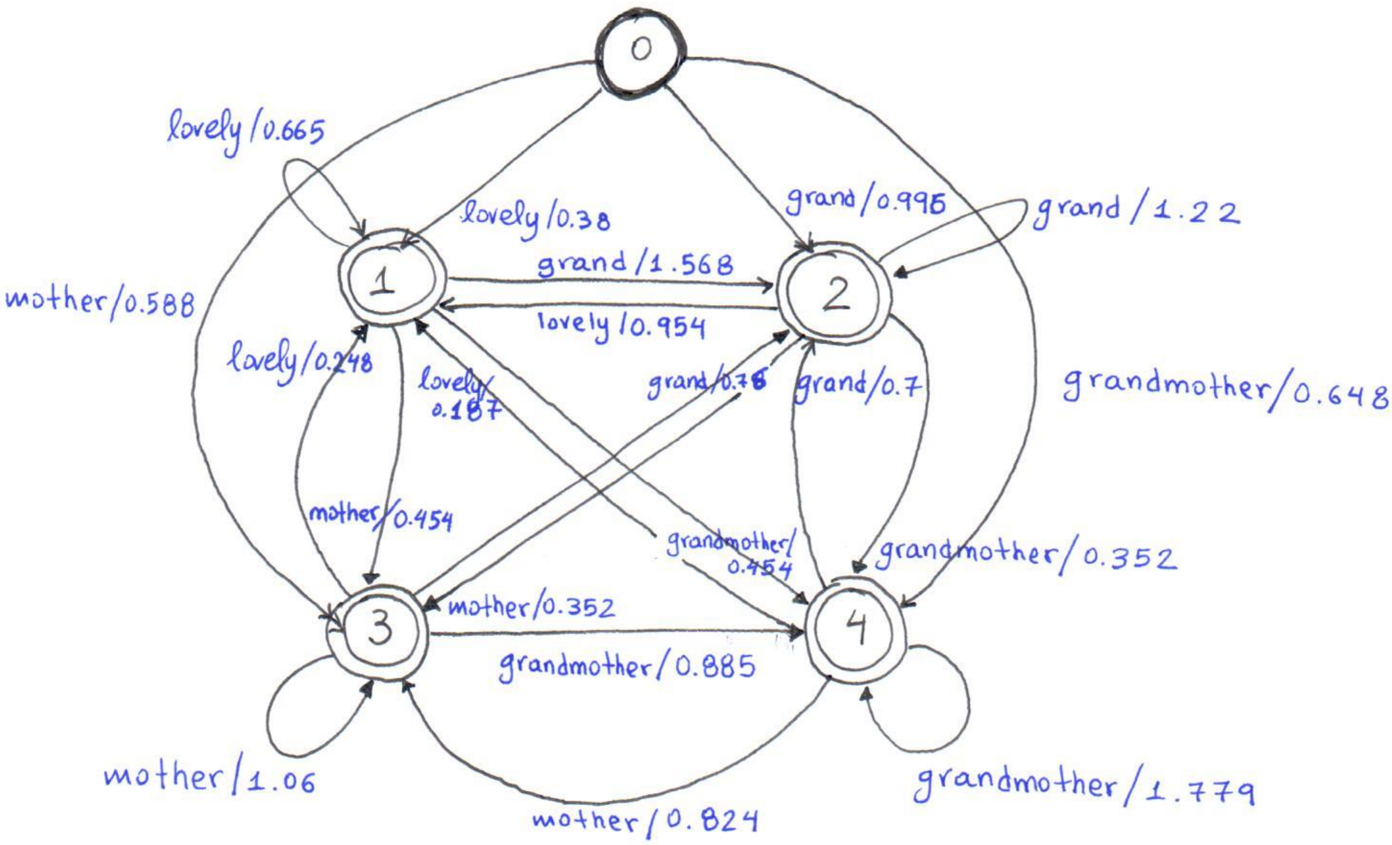
Παρακάτω φαίνεται η αντίστοιχη μηχανή πεπερασμένης κατάστασης με κόστος $-\log \mathbb{P}$ όπου:

(i) Η κατάσταση 1 σηματοδοτεί τελευταία λέξη "lovely"

(ii) Η κατάσταση 2 σηματοδοτεί τελευταία λέξη "grand"

(iii) Η κατάσταση 3 σηματοδοτεί τελευταία λέξη "mother"

(iv) Η κατάσταση 4 σηματοδοτεί τελευταία λέξη "grandmother".



2) Έστω ότι δίνεται η πρόταση χωρίς κενά "lovelygrandmothergrand".

- Το αυτόματο αναγνωρίζει τη λέξη "lovely" και μεταπηδά από την αρχική κατάσταση 0 στην κατάσταση 1 με κόστος $c_1 = 0.38$
- Στη συνέχεια, έχει την επιλογή να αναγνωρίσει τη λέξη "grand" και να μεταπηδήσει στην κατάσταση 2 με κόστος $c_2 = 1.568$, ή να αναγνωρίσει τη λέξη "grandmother" και να μεταπηδήσει στην κατάσταση 4 με κόστος $c_3 = 0.454$. Επιλέγεται το σενάριο μικρότερου κόστους, δηλαδή η δεύτερη περίπτωση ($c_3 < c_2$).
- Τέλος, αναγνωρίζεται η λέξη "grand" και καταλήγουμε στην τελική

κατάσταση 2 με κόστος $c_4 = 0.7$.

Συνολικά, το εν λόγω μονοπάτι έχει κόστος $c_{ολ} = c_1 + c_3 + c_4 = 1.534$.

Αν επιλεγόταν το μονοπάτι: "lovely" → "grand" → "mother" → "grand", τότε το συνολικό κόστος θα ήταν: $0.38 + 1.568 + 0.352 + 0.76 = 3.06$

δηλαδή: $c'_{ολ} = 3.06 > 1.534 = c_{ολ}$.

Συνεπώς, αφού προτιμάται το μονοπάτι μικρότερου κόστους, η πιό πιθανή σειρά από λέξεις για την πρόταση χωρίς κενά θα είναι:

lovely grandmother grand.

Άσκηση 6 :

Θεωρούμε ένα all pole μοντέλο με συνάρτηση μεταφοράς της μορφής :

$$V(z) = \frac{1}{\prod_{k=1}^{q}(1-c_k z^{-1})(1-c_k^* z^{-1})} \quad , \text{όπου } c_k = r_k e^{j\theta_k}$$

Τότε :

$$\hat{V}(z) = \log V(z) = \cancel{\log 1}^{\,0} - \sum_{k=1}^{q} \log(1-c_k z^{-1}) - \sum_{k=1}^{q} \log(1-c_k^* z^{-1})$$

$$\Rightarrow \hat{V}(z) = -\sum_{k=1}^{q}\left[\log(1-c_k z^{-1}) + \log(1-c_k^* z^{-1})\right]$$

Θεωρώντας ότι $|c_k|<1$, λαμβάνουμε υπόψη πως $\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots$ και λαμβάνουμε :

$$\hat{v}(n) = \begin{cases} 0 & , \text{για } n=0 \\ \sum_{k=1}^{q}\left[\frac{c_k^n}{n} + \frac{(c_k^*)^n}{n}\right] & , \text{για } n>0 \\ 0 & , \text{για } n<0 \end{cases}$$

Όμως, για $n>0$ έχουμε : $\hat{v}(n) = \sum_{k=1}^{q}\left[\frac{(r_k e^{j\theta_k})^n}{n} + \frac{(r_k e^{-j\theta_k})^n}{n}\right]$

$$\Rightarrow \hat{v}(n) = \sum_{k=1}^{q} \frac{r_k^n \cdot \left[e^{j\theta_k \cdot n} + e^{-j\theta_k n}\right]}{n}$$

. Από Τύπο Euler : $\cos x = \frac{e^{jx} + e^{-jx}}{2}$

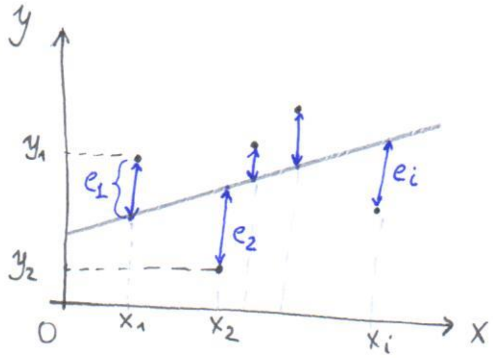
Συνεπώς :

$$\hat{v}(n) = \begin{cases} 2\sum_{k=1}^{q}\left(\frac{r_k^n}{n}\right)\cdot\cos(\theta_k \cdot n) & , \text{για } n>0 \\ 0 & , \text{για } n\le 0. \end{cases}$$

Άσκηση 7:

Θεωρούμε το απλό πρόβλημα linear regression $y = Wx + b$.

1) Ως συνάρτηση σφάλματος του linear regression θεωρούμε το άθροισμα των τετραγώνων των αποκλίσεων μεταξύ προβλεπόμενων και πραγματικών τιμών (Residual Sum of Squares).



Ειδικότερα, αν ορίσουμε ως y_i την i-οστή τιμή της μεταβλητής που πρόκειται να προβλεφθεί και $\hat{y}_i$ την αντίστοιχη προβλεπόμενη τιμή, τότε η συνάρτηση σφάλματος μπορεί να γραφεί:

$$\boxed{E = \frac{1}{2N}\sum_{i=1}^{N}\left\{y_i - \hat{y}_i\right\}^2 = \frac{1}{2N}\sum_{i=1}^{N}\left\{y_i - W^T \cdot x_i\right\}^2}$$

Μέσο Τετραγωνικό Σφάλμα

όπου:

$$W^T = \begin{bmatrix} b & w_1 & w_2 & \dots & w_n \end{bmatrix}$$

$$x_i = \begin{bmatrix} 1 \\ x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_n} \end{bmatrix}$$

2) Καλούμαστε να βρούμε τη λύση κλειστού τύπου για τις βέλτιστες παραμέτρους W, b. Επιδιώκουμε, δηλαδή, να ελαχιστοποιήσουμε τη συνάρτηση σφάλματος:

$$E = \frac{1}{2N}\sum_{i=1}^{N}\left\{ y_i - w^T x_i \right\}^2$$

Παραγωγίζοντας ως προς w, λαμβάνουμε: (αναλυτικό άθροισμα)

$$\nabla E = \frac{\cancel{2}}{\cancel{2}N}\sum_{i=1}^{N}\left\{ y_i - w^T x_i \right\}\cdot x_i^T = \frac{1}{N}\sum_{i=1}^{N}\left\{ y_i - w^T x_i \right\} x_i^T$$

Εξισώνοντας το αποτέλεσμα αυτό με μηδέν:

$$0 = \sum_{i=1}^{N} y_i \cdot x_i^T - w^T \sum_{i=1}^{N} x_i \cdot x_i^T$$

Λύνοντας ως προς w, έχουμε τη λύση κλειστού τύπου:

$$\boxed{w = (X^T X)^{-1} X^T y}$$ [3]

3) Στην πράξη παρόλο που υπάρχει η λύση κλειστού τύπου, χρησιμοποιούμε επαναληπτικές μεθόδους (π.χ. Gradient Descent) για τη λύση του προβλήματος linear regression. Αυτό συμβαίνει διότι, batch techniques, όπως η παραπάνω, οι οποίες επεξεργάζονται το σύνολο των δεδομένων μας μονομιάς, έχουν τεράστιο υπολογιστικό κόστος όταν εφαρμόζονται σε σύνολα δεδομένων μεγάλου μεγέθους. Στις περιπτώσεις αυτές, λοιπόν, είναι προτιμότερο να χρησιμοποιούμε σειριακούς αλγορίθμους που επεξεργάζονται κάθε δεδομένο μόνο του κάθε φορά και το παραμετρικό μοντέλο ενημερώνεται επαναληπτικά. Επιπλέον, ένας επαναληπτικός αλγόριθμος εφαρμόζεται αποτελεσματικότερα σε συστήματα, όπου η πληροφορία έρχεται σειριακά και πρέπει να ληφθούν προβλέψεις, προτού δούμε ολόκληρο το σύνολο δεδομένων μας.

4) Ο stochasting gradient descent επαναληπτικός αλγόριθμος, ενημερώνει τις παραμέτρους του διανύσματος w, χρησιμοποιώντας τον τύπο:

$$w^{(i+1)} = w^{(i)} - \eta \nabla E$$

όπου i δηλώνει τον αριθμό της επανάληψης και η είναι η παράμετρος του ρυθμού μάθησης. Το διάνυσμα w αρχικοποιείται σε μια τιμή $w^{(0)}$.

Για την περίπτωση του linear regression, όπου η συνάρτηση σφάλματος είναι αυτή που δείξαμε στο ερώτημα (a), παίρνουμε:

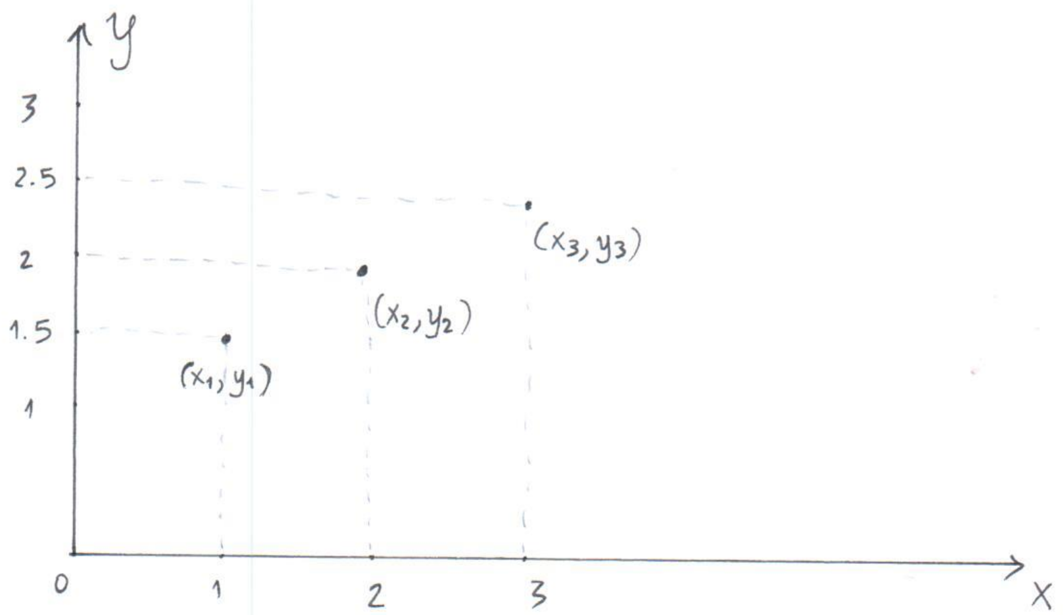
$$\boxed{w^{(i+1)} = w^{(i)} + \frac{\eta}{N}\left(y_n - w^{(i)T} x_n\right) x_n}$$

[3] Βιβλιογραφίας

$$\boxed{w^{(i+1)} = w^{(i)} - \frac{\eta}{N} X^T \left(X \cdot w^{(i)} - y\right)}$$

Εναλλακτική Μορφή

5) Δίνονται τρία δείγματα $x = \{1, 2, 3\}$ και οι αντίστοιχες επισημειώσεις $y = \{1.5, 2, 2.5\}$. Σχεδιάζουμε στο επίπεδο τα ζεύγη (x_i, y_i):



6) Δίνεται ως αρχική συνθήκη $(w^{(0)}, b^{(0)}) = (0, 0)$. Ως εκ τούτου, η ευθεία του linear regression $y = wx + b$ ταυτίζεται με τον οριζόντιο άξονα x.

Επομένως, η τιμή της συνάρτησης σφάλματος είναι:

$$E = \sum_{i=1}^{3} \{y_i - \hat{y}_i\}^2 = (1.5-0)^2 + (2-0)^2 + (2.5-0)^2 = 2.25 + 4 + 6.25$$

$\Rightarrow \boxed{E = 12.5}$ Επομένως: $E_n = \dfrac{\sum_{i=1}^{3} \{y_i - \hat{y}_i\}^2}{2 \cdot 3} \Rightarrow \boxed{E_n = 2.0833}$

7) Λύση Κλειστού Τύπου:

Θεωρούμε το διάνυσμα εισόδου $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ και το διάνυσμα επισημειώσεων $y = \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix}$.

Προκειμένου να λάβουμε το bias και τα βάρη στο διάνυσμα W, εισάγουμε μια μοναδιαία στήλη στο διάνυσμα x, ώστε να λάβουμε ορθά αποτελέσματα.

Δηλαδή: $X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$

Τότε:

$$W = (X^T X)^{-1} X^T y = \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} =$$

$$= \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} = \begin{bmatrix} 7/3 & -1 \\ -1 & 1/2 \end{bmatrix} \begin{bmatrix} 6 \\ 13 \end{bmatrix}$$

$\Rightarrow W = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$. Δηλαδή $\boxed{(w, b) = (0.5, 1)}$

ώστε $y = 0.5x + 1$

8) Μέθοδος Gradient Descent

Θεωρούμε τις αρχικές συνθήκες $(w^{(0)}, b^{(0)}) = (0,0)$, δηλαδή $w^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ και ρυθμό μάθησης $\eta = 1$.

1η Επανάληψη:

$$w^{(1)} = w^{(0)} - \frac{\eta}{N} X^T (X \cdot w^{(0)} - y)$$

$$\Rightarrow w^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} \right)$$

$$\Rightarrow w^{(1)} = \frac{1}{3} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} \Rightarrow w^{(1)} = \begin{bmatrix} 2 \\ 13/3 \end{bmatrix}$$

Δηλαδή $\boxed{(w^{(1)}, b^{(1)}) = \left(\frac{13}{3}, 2 \right) \simeq (4{,}33, 2)}$

Υπολογίζουμε το αντίστοιχο σφάλμα:

$$E^{(1)} = \frac{1}{2N} \sum_{i=1}^{N} \{ y_i - w x_i - b \}^2 = \frac{1}{6} \left[(1.5 - 4.33 \cdot 1 - 2)^2 + (2 - 4.33 \cdot 2 - 2)^2 \right.$$

$$\left. + (2.5 - 4.33 \cdot 3 - 2)^2 \right] = \frac{1}{6} \cdot \left[28.409 + 75{,}099 + 156{,}25 \right]$$

$$\Rightarrow \boxed{E^{(1)} = 43{,}292}$$

2η Επανάληψη:

$$w^{(2)} = w^{(1)} - \frac{\eta}{N} X^T (X w^{(1)} - y)$$

$$\Rightarrow w^{(2)} = \begin{bmatrix} 2 \\ 13/3 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 13/3 \end{bmatrix} - \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} \right)$$

$$\Rightarrow w^{(2)} = \begin{bmatrix} 2 \\ 13/3 \end{bmatrix} - \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 2/3 & 1 \end{bmatrix} \left(\begin{bmatrix} 19/3 \\ 32/3 \\ 45/3 \end{bmatrix} - \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix} \right)$$

$$\Rightarrow w^{(2)} = \begin{bmatrix} 2 \\ 13/3 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 14.5 \\ 26 \\ 37.5 \end{bmatrix} \cdot \frac{1}{3}$$

$$\Rightarrow w^{(2)} = \begin{bmatrix} 2 \\ 13/3 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 78 \\ 179 \end{bmatrix} \Rightarrow w^{(2)} = \frac{1}{9} \cdot \begin{bmatrix} -60 \\ -140 \end{bmatrix}$$

$$\Rightarrow w^{(2)} = \begin{bmatrix} -6.667 \\ -15.556 \end{bmatrix}$$

Δηλαδή : $\boxed{(w^{(2)}, b^{(2)}) = (-15.556, -6.667)}$

Υπολογίζουμε το αντίστοιχο σφάλμα :

$$E^{(2)} = \frac{1}{2N} \sum_{i=1}^{N} \{y_i - w^{(2)}x - b^{(2)}\}^2 \Rightarrow \boxed{E^{(2)} = 877,063}$$

Παρατηρούμε ότι καθώς εκτελούμε παραπάνω επαναλήψεις της μεθόδου Gradient Descent, το μέσο τετραγωνικό σφάλμα, αντί να μειώνεται, αυξάνεται!

Αυτό συμβαίνει διότι η επιλογή μας της παραμέτρου ρυθμού μάθησης $\underline{\eta = 1}$ είναι ανεπιτυχής, αποτρέποντας τον αλγόριθμο να συγκλίνει στη σωστή απάντηση $(w, b) = (0.5, 1)$.

ΒΙΒΛΙΟΓΡΑΦΙΑ:

[1] "Ψηφιακή Επεξεργασία Φωνής: Θεωρία & Εφαρμογές"
Lawrence R. Rabiner, Ronald W. Schafer

[2] "Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition and Computational Linguistics"
D. Jurafski, J. H. Martin

[3] "Pattern Recognition and Machine Learning"
C. M. Bishop