Μάθημα 4

15/3/18

Πηγή $A = \{a_1, a_2, \ldots, a_N\}$
$\Pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$

$$\boxed{H(A) = -\sum_{i=1}^{N} \pi_i \log \pi_i}$$ ($\log \rightarrow$ δυαδικός: bits/σύμβολο)

$$\boxed{H(A) = -\sum_{i=1}^{N}\sum_{j=1}^{M} \pi(a_i, b_j) \log \pi(a_i, b_j)}$$

$H(A,B) \leq H(A) + H(B)$

Υπό συνθήκη εντροπία (δεσμευμένη εντροπία)

Πηγή $A = \{a_1, a_2, \ldots, a_N\}$, $\Pi_A = \{\pi(a_1), \pi(a_2), \ldots, \pi(a_N)\}$
Πηγή $B = \{b_1, b_2, \ldots, b_N\}$, $\Pi_B = \{\pi(b_1), \pi(b_2), \ldots, \pi(b_N)\}$

Δίνεται ότι η B εμφανίζει το σύμβολο b_j

Ποια είναι η μέση αβεβαιότητα (εντροπία) για την A;

$$H(A|b_j) = -\sum_{i=1}^{N}\sum_{j=1}^{M} \pi(a_i, b_j) \cdot \log \pi(a_i|b_j)$$

$H(A|b_j) \rightarrow$ η αβεβαιότητα στη πηγή A δοθέντος ότι η B έχει εκπέμψει τη b_j

Αλλά: $\pi(a_i, b_j) = \pi(a_i|b_j) \cdot \cancelto{1}{\pi(b_j)} = \pi(a_i|b_j)$

Άρα $$\boxed{H(A|b_j) = -\sum_{i=1}^{N} \pi(a_i|b_j) \cdot \log \pi(a_i|b_j)}$$

$$H(A|B) = \sum_{j=1}^{M} \pi(b_j) \cdot H(A|b_j) = -\sum_{i=1}^{N}\sum_{j=1}^{M} \pi(b_j) \cdot \pi(a_i|b_j) \cdot \log \pi(a_i|b_j)$$

$$\Leftrightarrow \boxed{H(A|B) = -\sum_{i=1}^{N}\sum_{j=1}^{M} \pi(a_i, b_j) \log \pi(a_i|b_j)}$$

Ιδιότητες

$H(A|B) = H(A,B) - H(B)$

→ από κοινού αβεβαιότητα

→ αβεβαιότητα πηγής B

→ αβεβαιότητα πηγής A δοθέντος ότι η B έχει εκπέμψει κάτι

$H(B|A) = H(A,B) - H(A)$

$H(A|B) \leq H(A)$

Εκπέμπει a_i ο πομπός και λαμβάνει b_j ο δέκτης

Δοθέντος ότι στο δέκτη έχουμε λάβει κάτι αυτό αμέσως μειώνει την αβεβαιότητα με την οποία εξέπεμψε η πηγή.

Έστω ότι έχουμε μια πηγή που στέλνει σύμβολα 0 και 1 κι ο δέκτης εμφανίζει κάποια σύμβολα (ιδανικά θα ήταν 0 και 1, αλλά υπάρχει και ο θόρυβος)

Καθόμαστε εμείς στο δέκτη και περιμένουμε να μας στείλει κάτι ο πομπός. Αλλά δεν έχει στείλει κάτι. Ποια είναι η αβεβαιότητα για μας που καθόμαστε δέκτη για το τι σύμβολο θα εκπέμψει ο πομπός; Απ. $H(A)$

Έστω ότι η A εκπέμπει κάτι, μεταδίδεται από το κανάλι κι έρχεται στο B. Ποια η αβεβαιότητα στο δέκτη; Απ. Η δεσμευμένη πιθανότητα. Όταν όμως ο δέκτης εμφανίζει σύμβολα ανεξάρτητα από το τι του έστειλε ο πομπός ισχύει η ισότητα. (Ανεξάρτητα δηλαδή A και B)

Η μέση αβεβαιότητα μειώνεται αν εμφανιστεί ένα σύμβολο στο B.

$H(B|A) \leq H(B)$

Αποδείξεις ιδιοτήτων

$H(A|B) = H(A,B) - H(B)$

Απόδειξη

$$H(A|B) = -\sum_i \sum_j \pi(a_i,b_j)\log\pi(a_i/b_j) \xrightarrow{\text{κανόνας Bayes}} -\sum_i \sum_j \pi(a_i,b_j)\cdot\log\left[\frac{\pi(a_i,b_j)}{\pi(b_j)}\right]$$

$$= -\sum_i \sum_j \pi(a_i,b_j)\cdot\log\pi(a_i,b_j) + \sum_i \sum_j \pi(a_i,b_j)\log\pi(b_j)$$

$$= H(A,B) + \overbrace{\sum_i \sum_j \pi(a_i,b_j)}^{\pi(b_j)}\log\pi(b_j)$$

προσπαθώ να διώξω το a_i

$$= H(A,B) + \sum_j \left[\underbrace{\sum_i \pi(a_i,b_j)}_{\to\,\pi(b_j)}\right]\log\pi(b_j)$$

$$= H(A,B) + \sum_j \pi(b_j)\log\pi(b_j) = H(A,B) - H(A)$$

$H(A|B) \le H(A)$

Απόδειξη

Ισχύουν

Αντικαθιστώ την $H(A,B)$ με το άνω φράγμα της

$$\left.\begin{array}{l} H(A,B) \le H(A) + H(B) \\ H(A|B) = H(A,B) - H(B) \end{array}\right\} H(A|B) \le \overbrace{H(A)+H(B)} - H(B) \Leftrightarrow H(A|B) \le H(A)$$

Επέκταση πηγής πληροφορίας

Πηγές A και B $\longrightarrow H(A,B) \leq H(A)+H(B)$

- Η πηγή B ταυτίζεται με την πηγή A (π.χ. εκπέμπω 2 φορές σύμβολο από την A)

$\Pi_A^2 = \{\pi(a_i, a_j)\}$

- Η πηγή A εκπέμπει ανεξάρτητα σύμβολα

$\Pi_A^2 = \{\pi(a_i)\cdot\pi(a_j)\} : H(A,A) = H(A)+H(A) = 2H(A)$

Γενικά για ν-οστή επέκταση της πηγής : $H(A^\nu) = \nu \cdot H(A)$

Παράδειγμα

$A = \{a_1, a_2, a_3\}$ με $\Pi_A = \{1/2, 1/4, 1/4\}$

$B = \{b_1, b_2, b_3\}$

→ Δοθέντος ότι κάτι στέλνει ο A κάτι θα εμφανίσει ο B

$\left.\begin{array}{l} \pi(b_1|a_1) = 1/3 \\ \pi(b_2|a_1) = 1/3 \\ \pi(b_3|a_1) = 1/3 \end{array}\right\} \Sigma() = 1$ $\qquad \Pi_{B|A} = \{\pi(b_j|a_i)\}$

Να υπολογιστούν $H(A), H(B), H(A|B)$

$\pi(b_1|a_2) = 1/2$

$\pi(b_2|a_2) = 1/4$

$\pi(b_3|a_2) = 1/4$

$\pi(b_1|a_3) = 1/4$

$\pi(b_2|a_3) = 1/4$

$\pi(b_3|a_3) = 1/2$

Λύση

▸ $H(A) = -\sum_{i=1}^{3} \pi(a_i) \cdot \log \pi(a_i) = \ldots = 1{,}5$ bits/σύμβολο

▸ $H(B) = -\sum_{i=1}^{3} \pi(b_i) \log \pi(b_i)$

Βάσει τού κανόνα της άθροισης

$$\pi(b_j) = \sum_{i} \pi(a_i, b_j) \xrightarrow[\text{Bayes}]{\text{κανόνας}} \sum_{i=1} \pi(b_j | a_i) \cdot \pi(a_i)$$

$\pi(b_1) = \pi(a_i)\pi(b_1|a_1) + \pi(a_2) \cdot \pi(b_1|a_2) + \pi(a_3) \cdot \pi(b_1|a_3)$

$\pi(b_1) = 17/48$

$\pi(b_2) = 14/48$

$\pi(b_3) = 17/48$

$$H(A|B) = -\sum_{i=1}^{3} \sum_{j=1}^{3} \pi(a_i, b_j) \cdot \log \pi(a_i|b_j) = -\sum_{i=1}^{3} \sum_{j=1}^{3} \pi(b_j|a_i)\pi(a_i) \log \left[\frac{\pi(b_j|a_i) \cdot \pi(a_i)}{\pi(b_j)}\right]$$

$\pi(a_i, b_j) \xrightarrow[\text{Bayes}]{\text{κανόνας}} \pi(b_j|a_i) \cdot \pi(a_i)$ $\qquad = \ldots = 1{,}463$ bits/σύμβολο

$$\pi(a_i/b_j) = \frac{\pi(a_i, b_j)}{\pi(a_i)}$$

$H(A|B) < H(A)$

Πηγή πληροφορίας με μνήμη

π.χ. ποια η πιθανότητα μια πηγή να εκπέμψει κάτι αν προηγουμένως έχει εκπέμψει κάτι άλλο

Πηγή Α,ΠΑ εκπέμπει διαδοχικά τα σύμβολα
$a_i(1), a_i(2), \dots, a_i(k) \dots$ με $a_i \in A$

→ χρόνος

Αν η εμφάνιση κάθε συμβόλου εξαρτάται από τα μ προηγούμενα σύμβολα → η πηγή έχει τάξη μ:

$$\pi(a_i(k)) = \pi(a_i(k) / a_i(k-1), a_i(k-2), \dots, a_i(k-\mu))$$

Το μαθηματικό μοντέλο που περιγράφει πηγές με μνήμη:

Στοχαστική διαδικασία Markov:

Ορισμός: Στοχαστική διαδικασία $X(t)$ είναι ένα σύνολο τυχαίων μεταβλητών:

$\{X_i = X(t_i), i = 1, 2, \dots\}$ που αντιστοιχεί σε ένα σύνολο τιμών της παραμέτρου t_i με καθορισμένες από κοινού κατανομές: $\Pi_{x_1 x_2 \dots}((x_1, t_1), (x_2, t_2), \dots)$

Ορισμός: Η στοχαστική διαδικασία Markov $X(t)$, $t \in \{t_1 \le t_2 \le \dots \le t_N\}$ είναι εκείνη η στοχαστική διαδικασία όπου:

$$\Pi_{x_1 x_2 x_3 \dots x_N}((x_1, t_1), (x_2, t_2), \dots, (x_N, t_N)) = \underbrace{\Pi_{x_N x_{N-1}}((x_N, t_N) / (x_{N-1}, t_{N-1}))}$$

Η τ.μ. τη χρονική στιγμή t_N εξαρτάται μόνο από την τιμή της τ.μ. τη στιγμή t_{N-1}

Πίνακας (μητρώο) μεταβάσεων απλού βήματος για πηγή με ρ σύμβολα

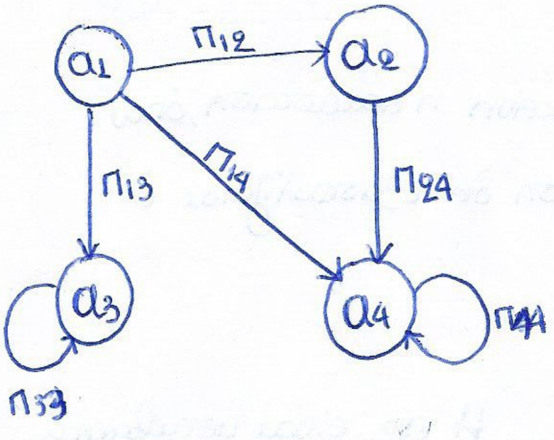
$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1\rho} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2\rho} \\ \vdots & \vdots & & \vdots \\ \pi_{\rho 1} & \pi_{\rho 2} & & \pi_{\rho\rho} \end{bmatrix}, \quad \sum_{i=1}^{\rho} \pi_{ij} = 1$$: Άθροισμα κάθε γραμμής 1

$$\pi_{ij} = \pi\left(a_j(k) / a_i(k-1)\right)$$

από προς

χρόνος

Διάγραμμα καταστάσεων



Χρονική στιγμή k=0: $\Pi^0 = [\pi_1 \; \pi_2 \ldots \pi_\rho]^T$ → πιθανότητες "εμφάνισης" των συμβόλων της πηγής

$\pi(a_i(0))$

Χρονική στιγμή k=1: Πιθανότητες εμφάνισης των συμβόλων : $\Pi^1 = \Pi^T \Pi^0 =$

χρόνος χρόνος

$$\begin{bmatrix} \pi_{11} & \pi_{21} & \cdots & \pi_{\rho 1} \\ \pi_{12} & \pi_{22} & \cdots & \pi_{\rho 2} \\ \vdots & \vdots & & \vdots \\ \pi_{1\rho} & \pi_{2\rho} & & \pi_{\rho\rho} \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_\rho \end{bmatrix}$$

π.χ. το π_1^1 (πρώτο στοιχείο του Π^1)

$$\Pi^1 = \begin{bmatrix} \pi_1^1 \\ \pi_2^1 \\ \vdots \\ \pi_\rho^1 \end{bmatrix} \quad \pi_1^1 = \pi_{11}\cdot\pi_1 + \pi_{21}\cdot\pi_2 + \cdots + \pi_{\rho 1}\pi_\rho$$

πιθανότητα να εμφανιστεί η κατάσταση a_1

$(AB)^T = B^T A^T$
$(AB\Gamma)^T = \Gamma^T B^T A^T$

Χρονική στιγμή $\kappa=2$: $\pi^2 = \Pi^T \cdot \pi^1 = \Pi^T \Pi^T \pi^0$

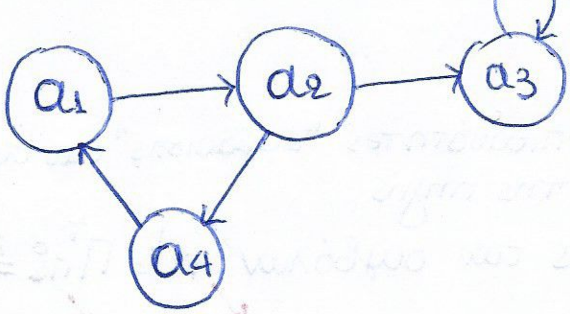
Χρονική στιγμή κ (γενικό) $\pi^k = \Pi^T \pi^{k-1} = \underbrace{\left(\Pi^T \Pi^T \dots \Pi^T\right)}_{\kappa \text{ φορές}} \pi^0 = \left(\Pi^T\right)^k \pi^0 = \left(\Pi^k\right)^T \pi^0$

Έννοιες σχετικές με τη στοχαστική διαδικασία Markov:

Αλυσίδα Markov: Ακολουθία διαδοχικών καταστάσεων Markov (το τώρα εξαρτάται από το πριν)

Στάσιμη Αλυσίδα: Στοιχεία δεν αλλάζουν στον χρόνο

Μεταβατική κατάσταση διαδικασίας Markov: Εκείνη η κατάσταση, όπου από μία τουλάχιστον διέξοδο σε γειτονική κατάσταση δεν εξασφαλίζεται ότι θα επανέλθουμε στην εν λόγω κατάσταση



Η a_2 είναι μεταβατική. Δηλαδή αν πάμε στην a_3 δεν μπορούμε να επανέλθουμε.

Επανερχόμενη κατάσταση: Το αντίθετο της μεταβατικής

Περιοδική κατάσταση: Είναι η επανερχόμενη κατάσταση που επιτρέπει την επιστροφή σε αυτήν μετά από ακέραιο αριθμό βημάτων πολλαπλάσιων κάποιων $\lambda \in N$.

$\pi_{ii}(k) = 0$ για $k \neq \lambda, 2\lambda, \dots$

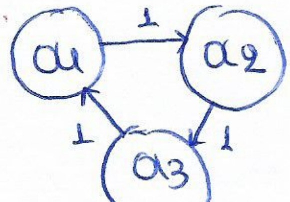
Εργοδική αλυσίδα Markov: Η αλυσίδα που επιτρέπει τη μετάβαση από οποιαδήποτε κατάσταση σε οποιαδήποτε άλλη. (όχι απαραίτητα σ' ένα βήμα)

Κανονική αλυσίδα Markov: Αν υπάρχει $\lambda \in \mathbb{N}$: Π^k για $k \geq \lambda$ να μην έχει μηδενικά. Επιτρέπει τη μετάβαση από οποιαδήποτε σε οποιαδήποτε κατάσταση σε ένα βήμα.

↳ Ειδική περίπτωση Εργοδικής

π.χ.

Οι ακμές είναι οι πιθανότητες μετάδοσης



$$\Pi = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Είναι εργοδική, αλλά όχι κανονική

Αν δεν μας δινόταν το σχήμα, αλλά ένας πιο πολύπλοκος πίνακας (π.χ. 100x100)
Κάθε στήλη βγάζει ένα σύμβολο.
Πώς θα τσεκάραμε αν είναι κανονικό;
Απ: Από τον ορισμό.
Αν υπάρχει κάποιο λ τ.ω. ο Π^k να μην έχει 0.
Αν δεν έχει 0 σημαίνει ότι από οποιαδήποτε κι αν ξεκινήσω μπορώ να πάω σε όλες. Άρα θα παίρναμε τα $\Pi^2, \Pi^3, \Pi^4, \Pi^5 \ldots$ μέχρι να καταλήξουμε σε πίνακα χωρίς 0, είναι κανονική.