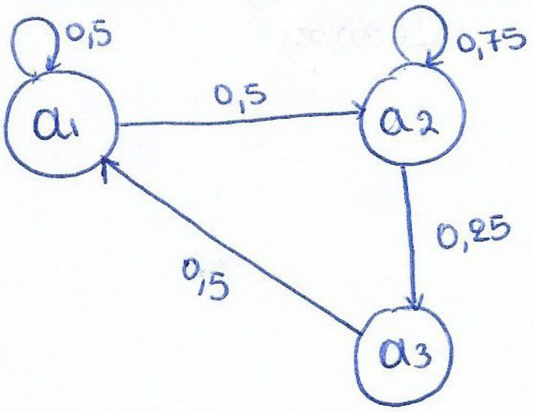
Μάθημα 5.

*** 22/3/18 ***

Κανονικό Σύστημα Markov : $(\Pi^T)^k$ δεν περιέχει μηδενικά για κάποια τιμή του k.



$$\Pi = \begin{bmatrix} 0,5 & 0,5 & 0 \\ 0 & 0,75 & 0,25 \\ 0,5 & 0 & 0,5 \end{bmatrix} \Rightarrow \Pi^2 = \begin{bmatrix} 0,25 & 0,62 & 0,13 \\ 0,13 & 0,56 & 0,31 \\ 0,5 & 0,25 & 0,25 \end{bmatrix}$$

- Εργοδικό
- Κανονικό;

Ιδιότητες κανονικών συστημάτων

- $\lambda = 1$ είναι ιδιοτιμή του $\Pi^T \Rightarrow \boxed{\Pi^T \underline{e_1} = e_1}$ (στάσιμη ή οριακή κατανομή)
- $|\lambda_i| \le 1$ για όλες τις ιδιοτιμές του Π^T.
- $\lim_{k\to+\infty} \Pi^k = \lim_{k\to+\infty} \left[(\Pi^T)^k_{\Pi_0} \right] = c \cdot e_1$

σταθερά από το Π^0 → εξαρτάται

$$A v = \lambda v$$

A: $N \times N$, v → ιδιοδιάνυσμα, λ → ιδιοτιμή του A

$$A v = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \dots & \vec{a_N} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} =$$

$$v_1 \vec{a_1} + v_2 \vec{a_2} + \dots + v_N \vec{a_N} =$$

$$v_1 \begin{bmatrix} \vec{a_1} \\ \downarrow \end{bmatrix} + v_2 \begin{bmatrix} \vec{a_2} \\ \downarrow \end{bmatrix} + \dots + v_N \begin{bmatrix} \vec{a_N} \\ \end{bmatrix}$$

- $$\lim_{k\to+\infty} (\Pi^T)^k = \begin{bmatrix} \pi_{11}(k) & \pi_{21}(k) & \dots & \pi_{N1}(k) \\ \pi_{12}(k) & \pi_{22}(k) & & \pi_{N2}(k) \\ \vdots & \vdots & & \vdots \\ \pi_{1N}(k) & \pi_{2N}(k) & & \pi_{NN}(k) \end{bmatrix}$$

$$= \begin{bmatrix} v_1 & v_1 & \dots & v_1 \\ v_2 & v_2 & \dots & v_2 \\ \vdots & \vdots & & \vdots \\ v_N & v_N & \dots & v_N \end{bmatrix}$$
$$\quad \begin{matrix} \downarrow & \downarrow & & \downarrow \\ e_1 & e_1 & & e_1 \end{matrix}$$

→ Η πιθανότητα να πάει το σύστημα στην κατάσταση i είναι ανεξάρτητο από την κατάσταση εκκίνησης.

με $e_1 = [v_1 \; v_2 \cdots v_N]^T$

Το ιδιοδιάνυσμα του Π^T είναι η οριακή κατανομή.

Υπολογισμός της οριακής κατανομής

$$\Pi^T e_1 = e_1 \iff \underbrace{\begin{cases} \pi_{11}v_1+\pi_{21}v_2+\dots+\pi_{N1}v_N=v_1 \\ \pi_{12}v_1+\pi_{22}v_2+\dots+\pi_{N2}v_N=v_2 \\ \vdots \\ \pi_{1N}v_1+\pi_{2N}v_2+\dots+\pi_{NN}v_N=v_N \end{cases}}_{N\times N \text{ σύστημα}}$$

Οι συνιστώσες αυτές είναι γραμμικά εξαρτημένες γιατί αν τις προσθέσω παίρνω *

*

$$v_1(\cancelto{1}{\pi_{11}+\pi_{12}+\dots+\pi_{1N}})+v_2(\cancelto{1}{\pi_{21}+\pi_{22}+\dots+\pi_{2N}})+\dots+v_N(\cancelto{1}{\pi_{N1}+\pi_{N2}+\dots+\pi_{NN}})=\sum_{i=1}^{N} v_i$$

$$\iff v_1+v_2+\dots+v_N=1 \iff 1=1$$

Οπότε προσθέτω στο σύστημα τη σχέση

$$v_1+v_2+\dots+v_N=1$$

και τα κάνω γραμμικώς εξαρτημένα

Γιατί ο Π^T έχει ιδιοτιμή $\lambda=1$;

$$\det(\Pi^T-\lambda I)=0 \xLeftrightarrow{\text{έστω } \lambda=1} \det(\Pi^T-I)=0$$

$$\Rightarrow \lambda=1 \text{ ιδιοτιμή του } \Pi^T$$

Η ορίζουσα ενός πίνακα είναι 0 όταν οι γραμμές και οι στήλες του είναι γραμμικώς εξαρτημένα.

είναι κάθετα μεταξύ τους
και ορθογώνια

ιδιοδιανύσματα του Π^T

ιδιοτιμές του Π^T

Οριακή κατανομή με βάση τις αρχικές συνθήκες

Διαγωνοποίηση του Π^T: $\Pi^T = S\Lambda S^{-1} = \begin{bmatrix} e_1 & e_2 & e_N \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & & \varnothing \\ & \lambda_2 & \\ \varnothing & & \ddots \\ & & & \lambda_N \end{bmatrix} S^{-1}$

$$(\Pi^T)^k = \underbrace{(S\Lambda S^{-1})(S\Lambda S^{-1})\dots(S\Lambda S^{-1})}_{\text{k-όροι}} = S\Lambda^k S^{-1}$$

εξ. από ΑΣ.

$$\Pi^k = (\Pi^T)^k \pi^0 = S\Lambda^k \underbrace{S^{-1}\pi^0}_{P} = S\Lambda^k P = \begin{bmatrix} e_1 & e_2 & \dots & e_N \\ \downarrow & \downarrow & \vdots & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_N^k \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_N \end{bmatrix} =$$

$P_1\lambda_1^k e_1 + P_2\lambda_2^k e_2 + \dots + P_N\lambda_N^k e_N$

$\downarrow$
αριθμός

Ισχύει $|\lambda_i| \leq 1$
και μόνο 1 ιδιοτιμή
είναι 1 ($\lambda_1 = 1$)

$$\lim_{k\to+\infty} \Pi^k = \lim_{k\to+\infty} (P_1\lambda_1^k e_1 + P_2\lambda_2^k e_2 + \dots + P_N\lambda_N^k e_N) = P_1 \cdot 1 \cdot e_1 = P_1 e_1$$

πρώτο στοιχείο του $S^{-1}\pi^0$

ιδιοδιάνυσμα για ιδιοτιμή $\lambda=1$

Ισχύει: $\lim \lambda_i^k = 0$, $\forall k = 1,2,3\dots,N$

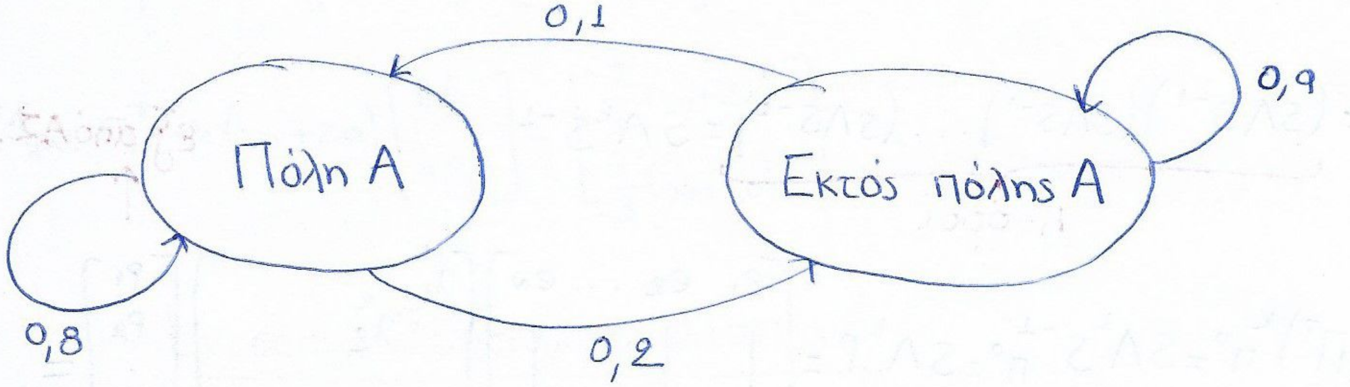
$\lim \lambda_1^k = 1$

Παράδειγμα

Ο πληθυσμός πόλης Α μεταβάλλεται ως εξής:

- Το 1/10 εκτός της πόλης Α εγκαθίσταται στην πόλη Α
- Τα 2/10 της πόλης Α την εγκαταλείπουν

Να υπολογιστεί η οριακή κατανομή τού συστήματος



$$\Pi^T = \begin{bmatrix} 0,9 & 0,2 \\ 0,1 & 0,8 \end{bmatrix} : \det(\Pi^T - \lambda I) = 0 \Leftrightarrow \lambda^2 - 1,7\lambda + 0,7 = 0 \Leftrightarrow \begin{cases} \lambda_1 = 1 \\ \lambda_2 = 0,7 \end{cases}$$

$$\begin{cases} \Pi^T v_1 = v_1 & \text{σύστημα 1 (μαζί με } v_1 + v_2 = 1) \\ \Pi^T v_2 = 0,7 v_2 & \text{σύστημα 2 (μαζί με } v_1 + v_2 = 1) \end{cases} \Rightarrow \begin{cases} v_1 = [\,2/3 \quad 1/3\,]^T \\ v_2 = [\,1/3 \quad -1/3\,]^T \end{cases}$$

→ οριακή κατανομή χωρίς αρχικές συνθήκες

$$\Pi^T = \underbrace{\begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0,7 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}}_{S^{-1}}$$

Αν οι Α.Σ. είναι $\pi^0 = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$, τότε $\lim_{k \to +\infty} \Pi^k = C \cdot v_1$ με C το πρώτο στοιχείο του $S^{-1}\pi^0$

$$S^{-1}\pi^0 = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} y_0 + z_0 \\ y_0 - 2z_0 \end{bmatrix}$$

Άρα $\Pi^k \to (y_0+z_0)\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3}(y_0+z_0) \\ \frac{1}{3}(y_0+z_0) \end{bmatrix}$ → εντός Α, → εκτός Α

Εντροπία Συστήματος Markov

Πηγή Α, π_A με ρ σύμβολα. Κάποια στιγμή βρίσκεται στην κατάσταση a_i.
Η αβεβαιότητα για την εξέλιξη της διαδικασίας: $H(a_i) = -\sum_{j=1}^{\rho} \pi_{ij} \log(\pi_{ij})$

Η μέση αβεβαιότητα που ονομάζεται Εντροπία Απλού Βήματος:

$$H(A) = \sum_{i=1}^{\rho} \pi(a_i) H(a_i) \iff \boxed{H(A) = -\sum_{i=1}\sum_{j=1} \pi_i \pi_{ij} \log(\pi_{ij})}$$

π_{ij}: πιθανότητα μετάβασης από Π^T

π_i: υπάρχει γιατί το σύστημά μας δεν ξεκινάει από την ηρεμία

Εντροπία πολλαπλού βήματος

$$\boxed{H^k(A) = -\sum_{i=1}^{\rho}\sum_{j=1}^{\rho} \pi_i \pi_{ij}^{(k)} \log[\pi_{ij}^{(k)}]}$$

Αν το σύστημα είναι κανονικό κι έχει φτάσει σε ισορροπία:

Εντροπία Απλού Βήματος: $\boxed{H(A) = -\sum_{i=1}^{\rho}\sum_{j=1}^{\rho} v_i \pi_{ij} \log(\pi_{ij})}$, v_i → οριακή πιθανότητα για την κατάσταση a_i (αντίστοιχη "συνιστώσα" του "πρώτου" ιδιοδιανύσματος)

Εντροπία Πολλαπλού Βήματος: $H^k(A) = -\sum_{i=1}^{\rho}\sum_{j=1}^{\rho} v_i \pi_{ij}^{(k)} \log[\pi_{ij}^{(k)}]$

Ιδιότητα: $H^{k+\lambda}(A) = H^k(A) + H^\lambda(A) = (k+\lambda)H(A)$

Οριακή Εντροπία: $H^{\infty}(A) = -\sum_{i=1}^{p} v_i \log v_i$

Διαφορά Απλού Βήματος και Οριακής Εντροπίας είναι ότι η Απλού Βήματος έχει μνήμη.

Ιδιότητα: $\boxed{H(A) \leq H^{\infty}(A) \leq \log(p)}$

Η μνήμη μειώνει την αβεβαιότητα

Απόδειξη

$H(A) \leq H^{\infty}(A) \leq \log(p)$

$-\sum_i \sum_j v_i \pi_{ij} \log(\pi_{ij}) \leq -\sum_{i=1}^{p} v_i \log v_i$

Λύση

Απόσταση kullback-Leibler μεταξύ κατανομών π και τ

$D(\pi, \tau) = \sum_{i=1}^{p} \pi_i \log\left(\frac{\pi_i}{\tau_i}\right) \geq 0$

Εδώ: $\pi_i = \pi(a_i, a_j) \rightarrow$ από κοινού πιθανότητα $= \pi(a_j / a_i)\pi(a_i) = \pi_{ij} v_i$

$\tau_i = \pi(a_i)\pi(a_j) \rightarrow$ περιθώριες πιθανότητες $= v_i v_j$

Κανονικό Σύστημα Markov / Οριακή κατάσταση

Άρα: $D[\pi(a_i, a_j), \pi(a_i)\pi(a_j)] = \sum_{i=1}^{p}\sum_{j=1}^{p} \pi(a_i, a_j) \log\left[\frac{\pi(a_i, a_j)}{\pi(a_i)\pi(a_j)}\right]$

$= \sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log\left(\frac{v_i \pi_{ij}}{v_i v_j}\right) = \sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log\left(\frac{\pi_{ij}}{v_j}\right) \geq 0$

$\Leftrightarrow \sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log(\pi_{ij}) - \sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log(v_i) \geq 0$

$\Leftrightarrow -\sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log(v_i) \geq -\sum_{i=1}^{p}\sum_{j=1}^{p} v_i \pi_{ij} \log(\pi_{ij})$

$\Leftrightarrow -\sum_{j=1}^{p} \Big(\underbrace{\sum_{i=1}^{p} v_i \pi_{ij}}_{v_j}\Big) \log(v_i) \geq -\sum_{i=1}^{p} v_i \sum_{j=1}^{p} \pi_{ij} \log(\pi_{ij})$

$$\Longleftrightarrow -\sum_{j=1}^{p} v_j \log(v_j) \geq -\sum_{i=1}^{p} v_i H(a_i)$$

$$\Longleftrightarrow \boxed{H^\infty(A) \geq H(A)}$$