Εξόρυξη Δεδομένων

Τρίτη Σειρά Ασκήσεων

Ευάγγελος Τζώρτζης

ΑΜ: 3088

Εξόρυξη Δεδομένων / Τρίτη Σειρά Ασκήσεων / Ευάγγελος Τζώρτζης / ΑΜ:3088

Ερώτηση 2

1. $P_v^T = (1-a)P_v^T \cdot P + a v^T \Rightarrow P_v^T - (1-a)\cdot P_v^T \cdot P = a v^T \Rightarrow P_v^T (I-(1-a)\cdot P) = a \cdot v^T \Rightarrow$

$$P_v^T \cdot \overbrace{(I-(1-a)P)\cdot(I-(1-a)P)^{-1}}^{I} = a\cdot v^T \cdot (I-(1-a)\cdot P)^{-1} \Rightarrow P_v^T = a\cdot v^T \cdot (I-(1-a)\cdot P)^{-1} \Rightarrow$$

Ισχύει: $P_v^T = v^T \cdot Q$

$$\Rightarrow v^T \cdot Q = a \cdot v^T \cdot (I-(1-a)\cdot P)^{-1} \Rightarrow \underbrace{(v^T)^{-1}\cdot v^T}_{I} \cdot Q = a \underbrace{(v^T)^{-1}\cdot v^T}_{I} \cdot (I-(1-a)\cdot P)^{-1} \Rightarrow$$

$$\Rightarrow \boxed{Q = a \cdot (I-(1-a)\cdot P)^{-1}}$$

2. $P_v^T = v^T \cdot Q = v^T \cdot a \cdot (I-(1-a)\cdot P)^{-1}$

- Οι γραμμές του Q μετατρέπουν το διάνυσμα v^T με βάση τον πίνακα P, καθώς και την πιθανότητα άλματος επανεκκίνησης a.

3. $P_u = v_u^T \cdot Q = [1/n, 1/n, \ldots, 1/n] \cdot Q = \frac{1}{n}[1,1,1,1\ldots,1]\cdot Q =$

$$= \frac{1}{n}\Big[\underbrace{[1,0,0\ldots,0]}_{v_1^T} + \underbrace{[0,1,\ldots,0]}_{v_2^T} + \ldots + \underbrace{[0,0,\ldots,1]}_{v_n^T}\Big]\cdot Q =$$

$$= \frac{1}{n}\sum_{i=1}^{n} v_i^T \cdot Q \quad \boxed{= \frac{1}{n}\sum_{i=1}^{n} P_i}$$

Q ανεξάρτητο του v^T

4. Θα χρησιμοποιηθεί ένας πίνακας για τη μετατροπή του διανύσματος κάποιας κατανομής σε ομοιόμορφη κατανομή.

- Έστω διάνυσμα κάποιας κατανομής γενικά v^T και διάνυσμα ομοιόμορφης κατανομής v_u^T.

Χρήση πίνακα A για μετατροπή του v^T σε v_u^T.

$$\boxed{v^T = A \cdot v_u^T} \ (1) \Rightarrow v^T \cdot (v_u^T)^{-1} = A \cdot v_u^T \cdot (v_u^T)^{-1} \Rightarrow \boxed{A = v^T \cdot (v_u^T)^{-1}} \ (2)$$

$$\hookrightarrow A^{-1} \cdot v^T = A^{-1} \cdot A \cdot v_u^T \Rightarrow \boxed{v_u^T = A^{-1} \cdot v^T} \ (3)$$

- Πρώτα μέσω της (2) υπολογίζεται ο A (και συνεπώς και ο A^{-1}).
- Ύστερα μέσω του τύπου $P_u = v_u^T \cdot Q \overset{(3)}{=} \underbrace{A^{-1} \cdot v^T \cdot Q}_{P_v} = \frac{1}{n}\sum_{i=1}^{n} P_i$ (από ερώτημα 3)

- Άρα το P_v εκφράζεται ως $\frac{1}{n}\sum_{i=1}^{n} P_i$, με τη διαφορά ότι ζητείται η μετατροπή του σε v_u^T μέσω του πίνακα A^{-1}.