Άσκηση 4

Χρήστος Αλέξανδρος Τσιγγιρόπουλος

29 December 2021

1 Ερώτημα 1

Αρχικά καλείται η συνάρτηση initialize_a() που αρχικοποιει ένα πίνακα A με τις τιμές της άσκησης. Στην συνέχεια, καλείται η συνάρτηση create_g(A) που δέχεται σαν όρισμα έναν πίνακα γειτνίασης A και δημιουργεί ένα πίνακα G όπου και επιστρέφει. Όπου ο πίνακας G έχει για στοιχεία του :

$$G[i][j] = \frac{q}{n} + \frac{A[j][i] \cdot (1-q)}{n_j}$$

με q την πιθανότητα ο χρήστης να μετακινηθεί σε μία τυχαία σελίδα, n τα στοιχεία του πίνακα A $(n \times n)$, A[j][i] το στοιχείο στην j-γραμμή και i-στήλη και n_j είναι το άθροισμα της j-οστής γραμμής του A.

Για να αποδείξουμε ότι ο πίνακας G είναι στοχαστικός και ειδικότερα αριστερά στοχαστικός, βρίσκουμε το άθροισμα κάθε στήλης και το τυπώνουμε στην οθόνη. Σαν έξοδο θα πάρουμε τα n αθροίσματα που βγαίνούν όλα 1.

2 Ερώτημα 2

Έχουμε ήδη δημιουργήσει τον πίνακα G απο το ερώτημα 1. Οπότε καλούμε την συνάρτηση methodos_dynamewn(G) που δέχεται σαν όρισμα τον πίνακα G και βρίσκει το ιδιοδιάνυσμα της μέγιστης ιδιοτιμής με την μέθοδοδ της δυνάμεως. Ειδικότερα, αρχικοποιεί δύο πίνακες τους b0,b1 (n στοιχεία) στον b0 την πρώτη στήλη του G και στον b1 το 0. Δημιουργεί τον $bn = A \cdot b0$, μετά διαιρεί κάθε στοιχείο του bn, με το πρώτο στοιχείο του, το bn[0], δλδ $bn[i] = \frac{bn[i]}{bn[0]}$ και στην συνέχεια αποθηκεύει τον bn στον b0. Η διαδικασία αυτή επαναλαμβάνεται για $2 \cdot n$ φορές. Τότε ο bn τείνει να ταυτιστεί με τον φορέα του ιδιοδιανύσματος που αντιστοιχεί στην μεγαλύτερη κατα απόλυτη τιμή ιδιοτιμή. Τέλος, κανονικοποιούμε το ιδιοδιάνυσμα, (δλδ διαιρούμε κάθε στοιχείο του πίνακα bn με το άθροισμα όλων στοιχείων του πίνακα bn, $bn[i] = \frac{bn[i]}{sum(bn)}$), ώστε το άθροισμα να είναι ίσο με 1 και επιστρέφουμε το κανινικοποιημένο ιδιοδιάνυσμα.

3 Ερώτημα 3

Εδώ αλλάζουμε τον πίνακα A ώστε να βελτιωθεί ο βαθμός σημαντικότας της 1ης σελίδας. Ειδικότερα προσθέτουμε τις συνδέσεις a[9][0],a[10][0],a[12][0],a[14][0]=1 και αφαιρούμε την σύνδεση a[0][7] = 0.

Για τον νέο πίνακα A′ τώρα βρίσκουμε τον νέο πίνακα G′ μέσα απο την συνάρτηση create_g(A′). Έπειτα βρίσκουμε το νέο κανονικοποιημένο ιδιοδυάνυσμα μέσο της συνάρτησης methodos_dynamewn(G′) και τον αποθηκεύουμε στον πίνακα p. Τέλος, τυπώνουμε τον πίνακα p και παρατηρούμε ότι πλέον η τάξη σελίδας είναι υψηλότερη για την σελίδα 1 με πιθανότητα 0.15518 ενώ πρίν ήταν οι σελίδες 13,15 με πιθανότητα 0.12509. Αξίζει να σημειωθεί οτί με τις 5 προσθαφαιρέσεις ο βαθμός σημαντικότητας της σελίδας 1 αυξήθηκε απο 0.02682 σε 0.15518 .

4 Ερώτημα 4

Για τον αλλαγμένο A του 3ου Ερωτήματος έχουμε:

- (α) q = 0.02

  Βρίσκουμε ξανά τον πίνακα g και έπειτα τον πίνακα p με το κανονικοποιημένο ιδιοδιάνυσμα για την νέα πιθανότητα μεταπήδησης q = 0.02 και τον τυπώνουμε. Βλέπουμε ότι η τάξη σημαντικότητας των μικρών ιστοσελίδων (με προηγούμενη σημαντικότητα < 0.05 για q=0.15) έγινε ακόμα πιο μικρή ενώ των μεγάλων (≥ 0.05) αυξήθηκε. Η μέγιστη τιμή είναι 0.17288 της πρώτης ιστοσελίδας.

- (b) q = 0.6

  Παρόμοια, βρίσκω g και p για την νέα πιθανότητα μεταπήδησης $q = 0.6$ και τυπώνω τον p. Σε αυτή την περίπτωση βλέπουμε το αντίθετο, δλδ ότι όσο μικρή ήταν η σημαντικότητα πρίν(για q=0.15) τόσο μεγαλύτερη η αύξηση για (q=0.6) και όσο μεγάλη ήταν η σημαντικότητα πρίν τόσο μεγαλύτερη η μείωση για (q=0.6). Η μέγιστη τιμή είναι 0.10397 της πρώτης ιστοσελίδας.

Ο σκοπός της πιθανότητας μεταπήδησης είναι για να δείξει πόσο εύκολα μπορώ να μεταβώ απο μία σελίδα σε μία άλλη. Όσο πιο μεγάλη είναι η πιθανότητα αυτή τόσο μικραίνουν οι διαφορές σημαντικότητας των ιστοσελίδων και όσο μικραίνει τόσο αυξάνονται οι διαφορές στην τάξη των ιστοσελίδων.

5 Ερώτημα 5

Τυπώνουμε τον πίνακα p με το κανονικοποιημένο ιδιοδιάνυσμα για τις κανονικές συνδέσεις. Στην συνέχεαι αλλάζουμε τον πίνακα A a[8][11],a[12][11] = 3. Δημιουργούμε τον πίνακα G' και τον p' και τον τυπώνουμε.

Παρατηρούμε ότι αυτή η στρατηγική δουλεύει, καθώς βελτιώθηκε η τάξη σημαντικότητας της ιστοσελίδας 11 (απο 0.10632 σε 0.13211) και μειώθηκε της 10 (απο 0.10632 σε 0.07858). Βλέπουμε όμως, να αυξήθηκε αρκετά και η σημαντικότητα της σελίδας 12 (απο 0.07456 σε 0.14822) κανοντάς την έτσι, πίο σημαντική απο την 11 κάτι που δέν ίσχυε πριν. Άυξηση δέχτηκαν και οι σελίδες 8 και 14, με την 14 να εξακολουθεί να είναι πρώτη, με δεύτερη την 12 και τρίτη την 11 στην τάξη σημαντικότητας.

6 Ερώτημα 6

Διαγράφουμε την ιστοσελίδα 10 απο τον πίνακα A δλδ πλέον A(14×14). Βρίσκουμε τους νέους πίνακες G και p και τυπώνουμε τον p.

Παρατηρούμε ότι αύξηση δεχτηκαν οι περισσότερες σελίδες. Τις μεγαλύτερες αυξήσεις τις δέχτηκαν η 11 (απο 0.10632 σε 0.17096) και η 13 (απο 0,12509 σε 0,18648) που πλέον έχει την υψηλότερη τάξη. Μείωση δέχτηκαν μόνο η 12 (απο 0.07456 σε 0.04822), η 14 (μια μικρή) ενώ την μεγαλύτερη την δέχτηκε η 15 (από 0.12509 σε 0.04116).

Παρατηρούμε ότι οι περισσότερες σελίδες δέχτηκαν αύξηση αφού το n απο 15 έγινε 14 οπότε και ο πίνακας p. Αφού ο p είναι ένα κανονικοποιημένο ιδιοδιάνυσμα θα πρέπει πλέον τα 14 στοιχεία του να βγάζουν άθροισμα 1.