

2η Υποχρεωτική Εργασία Στο Μάθημα της Αριθμητικής Ανάλυσης

Οδηγίες: Η εργασία είναι ατομική. Θα πρέπει να γραφεί παραδοτέο σε L^AT_EX ή σε οποιονδήποτε επεξεργαστή κειμένου στο οποίο να περιγράφονται οι λύσεις των ασκήσεων, να δίνονται και να σχολιάζονται τα αποτελέσματα, να περιγράφεται η εκτέλεση του κώδικα. Θα πρέπει το αρχείο του παραδοτέου (σε .pdf) μαζί με τα αρχεία του κώδικα (σε οποιαδήποτε γλώσσα προγραμματισμού επιθυμείτε) να συμπιεστούν σε ένα αρχείο το οποίο θα υποβληθεί μέσω του elearning μέχρι και την ημερομηνία παράδοσης:

22 Ιανουαρίου 2017, ώρα 24:00

Για κάθε ημέρα αργοπορημένης υποβολής της εργασίας και για 5 ημέρες μειώνεται η βαθμολογία κατά 10%. Μετά από την παράδοση όλων των εργασιών θα ακολουθήσει προφορική εξέταση ή παρουσίαση πάνω στις εργασίες, στην οποία θα περιλαμβάνεται και προφορική εξέταση του κώδικα για τις περιπτώσεις πιθανής αντιγραφής.

Για οποιαδήποτε απορία σχετικά με τις εργασίες μπορείτε να επικοινωνείτε μέσω email με τους βοηθούς μαθήματος στην διεύθυνση : natutor@aiia.csd.auth.gr.

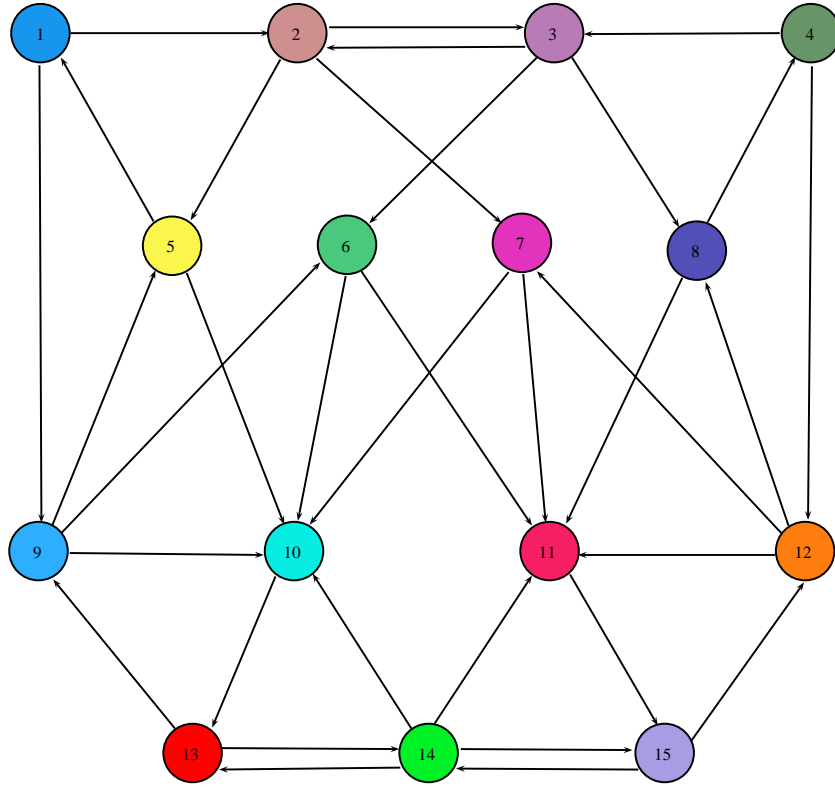
Άσκηση 1

Οι μηχανές αναζήτησης, όπως το <http://www.google.com/> χαρακτηρίζονται από τα εξαιρετικά αποτελέσματα που επιστρέφουν σε ερωτήματα. Θα δώσουμε μία χονδρική εξήγηση του πως γίνεται αυτή η ταξινόμηση χρησιμοποιώντας την πληροφορία σύνδεσης μεταξύ ιστοσελίδων στο WWW. Το google.com δίνει έναν θετικό πραγματικό αριθμό, την τάξη σελίδας (page rank), σε κάθε σελίδα που επιστρέφει. Η τάξη σελίδας υπολογίζεται από τη google.com σε έναν από τους μεγαλύτερους όπως λέγεται υπολογισμούς ιδιοδιανυσμάτων με τη μέθοδο των δυνάμεων στον κόσμο. Στο γράφημα του Σχήματος 1 κάθε κόμβος n αναπαριστά μία ιστοσελίδα και μία ακμή από το i στο j αναπαριστά ένα σύνδεσμο από την ιστοσελίδα του i στην ιστοσελίδα του j .

Έστω $\mathbf{A} \in \mathbb{R}^{n \times n}$, $n = 15$, ο πίνακας γειτνίασης του γραφήματος που απεικονίζεται στο Σχήμα 1.

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Για τον \mathbf{A} ισχύει $\mathbf{A}_{(i,j)} = 1$ όταν υπάρχει σύνδεσμος από τον κόμβο i στον κόμβο j και 0 διαφορετικά.



Σχήμα 1

Οι κατασκευαστές του google.com φαντάστηκαν έναν χρήστη σε αυτό το δίκτυο που είναι στην σελίδα i με πιθανότητα p_i . Έπειτα, ο χρήστης είτε μετακινείται σε μία τυχαία σελίδα (με συγκεκριμένη πιθανότητα q που είναι ίση συνήθως με 0.15) ή με πιθανότητα $1 - q$ επιλέγει τυχαία έναν από τους συνδέσμους που βρίσκονται στη σελίδα i . Η πιθανότητα ότι ο χρήστης μετακινείται από την σελίδα i στη σελίδα j δίνεται από τη σχέση:

$$p_{\{i \rightarrow j\}} = \frac{q}{n} + (1 - q) * \frac{\mathbf{A}_{(i,j)}}{n_i}$$

όπου n_i είναι το άθροισμα της i -οστής γραμμής του \mathbf{A} (στην ουσία το πλήθος των συνδέσεων της σελίδα i). Η πιθανότητα ο χρήστης σε οποιαδήποτε χρονική στιγμή να είναι στην σελίδα j είναι το άθροισμα της $p_{\{i \rightarrow j\}}$ στην i -οστή γραμμή του πίνακα \mathbf{A} .

$$p_j = \sum_{i=1}^n \left(\frac{qp_i}{n} + (1 - q) \frac{p_i}{n_i} \mathbf{A}_{(i,j)} \right) \quad (1)$$

Η Σχέση (1) μπορεί να γραφεί και με μορφή πινάκων ως εξής:

$$\mathbf{p} = \mathbf{Gp} \quad (2)$$

όπου \mathbf{p} είναι το διάνυσμα με τις n πιθανότητες του να βρίσκεται ο χρήστης στις n ιστοσελίδες και \mathbf{G} είναι ο πίνακας του οποίου το κελί $\mathbf{G}_{(i,j)} = \frac{q}{n} + \frac{\mathbf{A}_{(j,i)}(1-q)}{n_j}$. Θα ονομάσουμε τον πίνακα \mathbf{G} , πίνακα Google. Κάθε στήλη του πίνακα Google δίνει άθροισμα 1 και επομένως είναι στοχαστικός πίνακας που σημαίνει ότι η μέγιστη ιδιοτιμή του είναι 1. Το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή 1 είναι το σύνολο των μη μεταβλητών πιθανοτήτων, που είναι επί της ουσίας η τάξη της σελίδας.

Με βάση το παράδειγμα που δώσαμε προηγουμένως, οι τάξεις κάθε σελίδας για $q = 0.15$ δίνονται παρακάτω:

$\mathbf{p} = [0.0268238, 0.0298611, 0.0298612, 0.0268241, 0.0395877, 0.0395878, \dots, 0.039588, 0.0395881, 0.074565, 0.106319, 0.10632, 0.0745657, 0.125089, \dots, \dots, 0.11633, 0.12509]^T$

Το συγκεκριμένο ιδιοδιάνυσμα έχει κανονικοποιηθεί ώστε το άθροισμα να είναι ίσο με 1 (όπως θα πρέπει να ισχύει για πιθανότητες). Η τάξη σελίδας είναι υψηλότερη για τις σελίδες 13 και 15 ακολουθούμενες από τις σελίδες 14, 10 και 11. Προσέξτε ότι η τάξη σελίδας δεν εξαρτάται μόνο από το βαθμό εισόδου (πλήθος σελίδων που δείχνουν σε μία σελίδα) αλλά προκύπτει με πιο πολύπλοκο τρόπο. Παρόλο που οι κόμβοι 10 και 11 έχουν το μεγαλύτερο βαθμό εισόδου το γεγονός ότι δείχνουν στους κόμβους 13 και 15 μεταφέρει τη σημαντικότητά τους σε αυτούς. Αυτή είναι η ιδέα πίσω από το φαινόμενο του google-bombing, δηλαδή της τακτικής του να αυξάνεται η σημαντικότητα μίας ιστοσελίδας πείθοντας άλλες ιστοσελίδες υψηλής κίνησης να δείχνουν σε αυτή. Προσέξτε την σύνδεση μεταξύ της τάξης σελίδας και της λέξης σημαντικότητας. Η τάξη σελίδας είναι αυτή τη στιγμή ο καλύτερος τρόπος να αναπαραστήσουμε τη σημαντικότητα μία σελίδας αν και κανείς ακόμα δεν ξέρει τι είναι σημαντικότητα. Επομένως, κάποιος μπορεί να βρει μία καλύτερη μέθοδο για να μοντελοποιεί τη σημαντικότητα.

Ζητούμενα:

1. Αποδείξτε αναλυτικά ότι ο πίνακας \mathbf{G} είναι στοχαστικός.
2. Κατασκευάστε τον πίνακα \mathbf{G} για το συγκεκριμένο παράδειγμα και ελέγξτε ότι το ιδιοδιάνυσμα της μέγιστης ιδιοτιμής είναι αυτό που δίνεται παραπάνω (με τη μέθοδο της δυνάμεως).
3. Προσθέστε 3 συνδέσεις που εσείς θα επιλέξετε στο γράφημα και κατασκευάστε τον νέο πίνακα \mathbf{G} έτσι ώστε να βελτιώσετε τον βαθμό σημαντικότητας μιας σελίδας που θα επιλέξετε.
4. Αλλάξτε την πιθανότητα μεταπήδησης q σε (α) $q = 0.01$ και (β) $q = 0.7$ στον νέο γράφο. Περιγράψτε τις αλλαγές στην τάξη σελίδας. Ποιος είναι ο σκοπός της πιθανότητας μεταπήδησης;
5. Έστω ότι στο αρχικό γράφημα του Σχήματος 1 η ιστοσελίδα 11 ήθελε να βελτιώσει την τάξη της σε σχέση με τον ανταγωνιστή της, την σελίδα 10. Έστω ότι αυτό θέλει να το πετύχει πείθοντας τις σελίδες 8 και 12 να δείχνουν τη σύνδεση στην 11 με καλύτερο τρόπο. Έστω ότι μοντελοποιούμε αυτή την κατάσταση αντικαθιστώντας το $\mathbf{A}_{(8,11)}$ και το $\mathbf{A}_{(12,11)}$ με 2 στον πίνακα γειτνίασης. Αυτή η στρατηγική δουλεύει;
6. Μελετήστε την επίδραση της διαγραφής της σελίδας 11 από το γράφημα. Ποιες τάξεις σελίδων αυξάνονται και ποιες μειώνονται;

Άσκηση 2

Πρέπει να προγραμματίσετε μια συνάρτηση που να υπολογίζει το ημίτονο μιας οποιασδήποτε γωνίας. Δηλαδή να φτιάξετε μια προσομοίωση του τρόπου με τον οποίο υπολογίζεται το ημίτονο στο κομπιουτεράκι σας. Για να δημιουργήσετε την συνάρτηση μπορείτε να χρησιμοποιήσετε 12 τιμές του ημιτόνου με όποια ακρίβεια θέλετε που εσείς θα επιλέξετε και ακολουθώντας να προσεγγίσετε την συνάρτηση του ημιτόνου σε αυτές τις τιμές με τις παρακάτω μεθόδους που αναπτύχθηκαν στις διαλέξεις:

- Με πολυωνυμική προσέγγιση.
- Με Splines.
- Με την μέθοδο ελαχίστων τετραγώνων.

Για να πετύχετε καλύτερη ακρίβεια μπορείτε να χρησιμοποιήσετε μη ομοιόμορφα κατανομημένα τα 12 σημεία που θα επιλέξετε. Να συγκρίνετε τις παραπάνω προσεγγίσεις όσον αφορά στην ακρίβεια προσέγγισης στο διάστημα $[-\pi, \pi]$. Να προβάλετε σε διάγραμμα το σφάλμα προσέγγισης για 200 σημεία στο $[-\pi, \pi]$. Πόσα ψηφία ακρίβειας πετυχαίνετε;

Άσκηση 3

Να υπολογίσετε το ολοκλήρωμα της συνάρτησης του ημιτόνου στο διάστημα $[0, \pi/2]$ χρησιμοποιώντας μερικά ή όλα τα σημεία που επιλέξατε στην προηγούμενη άσκηση και τις μεθόδους Simpson και τραπεζίου. Ποιο είναι το σφάλμα προσέγγισης θεωρητικά και αριθμητικά.

Άσκηση 4

Να εκτιμήσετε την τιμή κλεισίματος ημέρας, τριών διαφορετικών μετοχών εταιρειών του Χρηματιστηρίου Αθηνών που εσείς θα επιλέξετε για την ημέρα που είναι η κοντινότερη στα γενέθλιά σας στο 2016 με βάση τις προηγούμενες 20 τιμές κλεισίματος ημέρας που θα πάρετε από την σελίδα <http://www.capital.gr/finance/historycloses> ή οποιαδήποτε άλλη σελίδα που δίνει τέτοια στοιχεία χρησιμοποιώντας τα σύμβολα των διαφόρων μετοχών που μπορείτε να βρείτε στο <http://www.capital.gr/finance/el/allstocks>. Για παράδειγμα αν είχα γενέθλια στις 9/11/2016 θα επιλέξω 3 διαφορετικές μετοχές και θα χρησιμοποιήσω τις τιμές κλεισίματος των μετοχών αυτών για τις 20 συνεδριάσεις που έγιναν πριν τις 9/11/2016 και θα προβλέψω την τιμή των μετοχών για την επόμενη συνεδρίαση (10/11/2016) καθώς και για 5 συνεδριάσεις μετά (14/11/2016). Αν το χρηματιστήριο ήταν κλειστό χρησιμοποιήστε τιμές από το έτος 2015.

Θα προσεγγίσετε την συνάρτηση τιμής κλεισίματος με πολυώνυμο 2ου, 3ου και 4ου βαθμού με την μέθοδο ελαχίστων τετραγώνων. Ακολούθως θα βρείτε την τιμή της συνάρτησης στην θέση που σας ενδιαφέρει. Συγκρίνετε τις προσεγγίσεις σας ποιοτικά για τις τιμές που έχετε διαθέσιμες αλλά και για την ημέρα πρόβλεψης. Να αναφέρετε σε ποιά ημερομηνία κάνατε πρόβλεψη της τιμής της κάθε μετοχής. Κάντε πρόβλεψη τιμής για 5 συνεδριάσεις μετά τις τιμές που έχετε διαθέσιμες και σχολιάστε τα αποτελέσματα. Για τις λύσεις των ασκήσεων δεν μπορείτε να χρησιμοποιήσετε έτοιμες συναρτήσεις προσέγγισης που παρέχονται από τις διάφορες γλώσσες προγραμματισμού. Θα πρέπει να κάνετε τις δικές σας συναρτήσεις.