Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Αλγόριθμοι και Πολυπλοκότητα

Μεταπτυχιακό ΣΕΜΦΕ, ΜΠΛΑ, Πληροφορική-ΕΚΠΑ

Διδάσκοντες: Σ. Ζάχος, Δ. Φωτάκης

Προτεινόμενες Ασκήσεις

Άσκηση 1. Σε μια φανταστική χώρα, διενεργείται δημοσκόπηση σχετικά με τη στάση των πολιτών ως προς μια σημαντική πολιτικο-οικονομική μεταβολή. Οι πολίτες απαντούν στη δημοσκόπηση με ''ναι'' ή ''όχι'' (υπέρ ή εναντίον της μεταβολής). Αν το (πραγματικό) ποσοστό των πολιτών που τάσσονται υπέρ της μεταβολής είναι p, θέλουμε να υπολογίσουμε μια εκτίμηση $\hat{p}$ του p ώστε $\Pr[|\hat{p}-p| \leq \varepsilon p] > 1-\delta$, για δεδομένα $\varepsilon, \delta \in (0,1)$. Για τη δημοσκόπηση, θα ρωτήσουμε N πολίτες, που επιλέγονται ισοπίθανα και ανεξάρτητα από το σύνολο των πολιτών. Η εκτίμησή μας $\hat{p}$ θα είναι το ποσοστό των N πολιτών που τάσσονται υπέρ της μεταβολής. Χρησιμοποιώντας Chernoff bounds, να υπολογίσετε (ως συνάρτηση των ε, δ, και p) το ελάχιστο μέγεθος N του δείγματος που χρειαζόμαστε. Βρείτε την τιμή του N για $\varepsilon = 0.02$ και $\delta = 0.05$, αν γνωρίζουμε ότι $p \in [0.2, 0.8]$ (και δείτε ότι αυτή η τιμή είναι ανεξάρτητη του πληθυσμού της χώρας!).

Άσκηση 2. Θεωρούμε ένα ασύρματο δίκτυο όπου ένας δέκτης ''ακούει'' μια πηγή μετάδοσης σε μια συγκεκριμένη συχνότητα, και το bit $b(t)$, που μεταδίδεται τη χρονική στιγμή t, αναπαρίσταται είτε ως 1 είτε ως -1. Εκτός όμως από την πηγή μετάδοσης με την οποία είναι συντονισμένος ο δέκτης μας, υπάρχουν και άλλες n πηγές στον περιβάλλοντα χώρο, οι οποίες δημιουργούν παρεμβολές. Θεωρούμε ότι κάθε τέτοια πηγή i έχει ισχύ $p_i \in (0,1)$, και μεταδίδει το bit $b_i(t) \in \{1,-1\}$ τη χρονική στιγμή t. Ας δεχθούμε ότι ο δέκτης μας λαμβάνει το σήμα:

$$s(t) = b(t) + \sum_{i=1}^{n} p_i b_i(t)$$

και συμπεραίνει ότι η πηγή μετέδωσε 1, αν $s(t) \geq 0$, και -1, διαφορετικά. Αν θεωρήσουμε ότι τα bits $b_i(t)$ επιλέγονται ομοιόμορφα και ανεξάρτητα, να εκτιμήσετε (με χρήση Chernoff bounds) την πιθανότητα ο δέκτης μας να υπολογίσει λάθος το $b(t)$.

Άσκηση 3. Έστω το Γραμμικό Πρόγραμμα (Π1):

$$\begin{aligned} \min \quad & 3x_1 + x_2 + x_3 + x_4 \\ \text{s.t.} \quad & 3x_1 - x_2 - 7x_3 + x_4 = -3 \\ & 3x_1 - 4x_3 + x_4 = -1 \\ & x_3 \leq 3 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

1. Να σχεδιάσετε το σύνολο των εφικτών λύσεων του (Π1) και να βρείτε τις κορυφές του (ίσως χρειαστεί να το τροποποιήσετε κατάλληλα). Να σχεδιάσετε την κατεύθυνση βελτιστοποίησης, και να βρείτε μια βέλτιστη λύση.
2. Να διατυπώσετε το δυϊκό πρόγραμμα (ΔΠ1) του (Π1). Να διατυπώσετε τις complementary slackness συνθήκες για τα (Π1) και (ΔΠ1), και να τις χρησιμοποιήσετε για τον υπολογισμό της βέλτιστης λύσης του (ΔΠ1).

Άσκηση 4. Έστω το Γραμμικό Πρόγραμμα (Π2):

$$\begin{array}{rr} \min & -10x_1 + 57x_2 + 9x_3 + 24x_4 \\ \text{s.t.} & 0.5x_1 - 5.5x_2 - 2.5x_3 + 9x_4 + x_5 = 0 \\ & 0.5x_2 - 1.5x_2 - 0.5x_3 + x_4 + x_6 = 0 \\ & x_1 + x_7 = 1 \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0 \end{array}$$

Να λύσετε (δύο φορές) το (Π2) με τον αλγόριθμο Simplex. Να ξεκινήσετε με βασικές μεταβλητές τις x_5, x_6, x_7. Να ακολουθήσετε τους παρακάτω κανόνες για την εναλλαγή στηλών (pivoting) στη βάση (την πρώτη φορά τον ένα, την δεύτερη τον άλλο):

1. (i) ως νέα βασική μεταβλητή επιλέγεται εκείνη με το ελάχιστο ανηγμένο κόστος, και (ii) αν υπάρχουν δύο ή περισσότερες βασικές μεταβλητές υποψήφιες για έξοδο από τη βάση, επιλέγεται εκείνη με τον ελάχιστο δείκτη.
2. (i) ως νέα βασική μεταβλητή επιλέγεται εκείνη (από τις μεταβλητές με αρνητικό ανηγμένο κόστος) με τον ελάχιστο δείκτη, και (ii) αν υπάρχουν δύο ή περισσότερες βασικές μεταβλητές υποψήφιες για έξοδο από τη βάση, επιλέγεται εκείνη με τον ελάχιστο δείκτη.

Άσκηση 5. Ένας τετραγωνικός ακέραιος πίνακας B καλείται *unimodular* αν η απόλυτη τιμή της ορίζουσάς του είναι ίση με 1. Ένας (όχι κατ' ανάγκη τετραγωνικός) ακέραιος πίνακας A καλείται *totally unimodular* αν κάθε τετραγωνικός μη ιδιάζων (nonsingular) υποπίνακάς του A είναι unimodular.

Θεωρούμε το πολύτοπο $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, όπου A ένας ακέραιος πίνακας $m \times n$ και b ένα ακέραιο m-διάνυσμα. Να δείξετε ότι αν ο A είναι totally unimodular, τότε όλες οι κορυφές του P είναι ακέραια n-διανύσματα.

Άσκηση 6. Έστω A ένας πίνακας $m \times n$, x ένα n-διάνυσμα μεταβλητών, και b ένα m-διάνυσμα. Το γραμμικό σύστημα $Ax \leq b$ καλείται *μη-συμβιβαστό* αν υπάρχει m-διάνυσμα y τέτοιο ώστε $A^T y = 0$, $b^T y < 0$, και $y \geq 0$. Να αποδείξετε ότι το σύστημα $Ax \leq b$ είναι *μη-επιλύσιμο* αν και μόνο αν είναι *μη-συμβιβαστό*.

Άσκηση 7. Στο πρόβλημα του MAX k-CUT, δίνεται ένα απλό μη κατευθυνόμενο γράφημα $G(V, E)$ με θετικά βάρη στις ακμές. Το ζητούμενο είναι μια διαμέριση του V σε k σύνολα $V_1, \dots, V_k$ που μεγιστοποιεί το συνολικό βάρος των ακμών στο αντίστοιχο k-cut. Θεωρούμε τον απλό πιθανοτικό αλγόριθμο που ισοπίθανα τοποθετεί κάθε κορυφή σε ένα από τα σύνολα $V_1, \dots, V_k$. Να δείξετε ότι η μέση τιμή του συνολικού βάρους των ακμών στο k-cut που προκύπτει είναι τουλάχιστον $(1 - \frac{1}{k})\,\mathrm{OPT}$.

Άσκηση 8. Να δείξετε ότι ο άπληστος αλγόριθμος για το (διακριτό) Knapsack που χρησιμοποιεί ως κριτήριο επιλογής το λόγο της αξίας προς το μέγεθος κάθε αντικειμένου δεν επιτυγχάνει σταθερό λόγο προσέγγισης. Να τροποποιήσετε τον αλγόριθμο ώστε (η τροποποιημένη εκδοχή) να επιτυγχάνει λόγο προσέγγισης 2 (ο χρόνος εκτέλεσης πρέπει να παραμείνει $O(n \log n)$, όπου n ο αριθμός των αντικειμένων).

Άσκηση 9. Θεωρούμε τον παρακάτω αλγόριθμο προσέγγισης για το πρόβλημα του weighted Vertex Cover σε ένα γράφημα $G(V, E)$, όπου κάθε κορυφή $v \in V$ έχει θετικό βάρος $w(v)$.

Να δείξετε ότι ο αλγόριθμος επιτυγχάνει λόγο προσέγγισης 2. Ειδικότερα, να δείξετε ότι για τις τιμές των C και $c(e)$, $e \in E$, στο τέλος του αλγορίθμου, ισχύουν τα εξής:

1. Το συνολικό βάρος του C είναι μικρότερο ή ίσο του $2\sum_{e\in E} c(e)$.
2. Το βάρος του βέλτιστου συνόλου κάλυψης είναι μεγαλύτερο ή ίσο του $\sum_{e\in E} c(e)$. *Υπόδειξη:* Να δείξετε ότι στο τέλος του αλγόριθμου, οι τιμές των $c(e)$ αποτελούν feasible dual λύση.

Ακόμη να βρείτε ένα γράφημα όπου ο αλγόριθμος υπολογίζει ένα σύνολο κάλυψης με βάρος διπλάσιο από αυτό του βέλτιστου συνόλου κάλυψης.

WeightedVertexCover($G(V, E)$, w)
- $C \leftarrow \emptyset$;
- $\forall v \in V$, $t(v) \leftarrow w(v)$;
- $\forall e \in E$, $c(e) \leftarrow 0$;
- **while** C δεν είναι vertex cover **do**
  - $e = \{u, v\}$ μια ακάλυπτη ακμή;
  - $\delta \leftarrow \min\{t(u), t(v)\}$;
  - $t(u) \leftarrow t(u) - \delta$;
  - $t(v) \leftarrow t(v) - \delta$;
  - $c(e) \leftarrow \delta$;
  - $C \leftarrow C \cup \{z \in \{u, v\} : t(z) = 0\}$;
- **return**(C);

Άσκηση 10. Στο πρόβλημα του Maximum Coverage, δίνονται ένα σύνολο U με n στοιχεία, μια συλλογή $S_1, \dots, S_m$ από υποσύνολα του U, και ένας θετικός ακέραιος k. Το ζητούμενο είναι η επιλογή k υποσυνόλων $S_{i_1}, \dots, S_{i_k}$ ώστε να μεγιστοποιηθεί ο αριθμός των στοιχείων που καλύπτονται από αυτά (δηλ. να μεγιστοποιηθεί το $|S_{i_1} \cup \cdots \cup S_{i_k}|$).

1. Να δείξετε ότι ο άπληστος αλγόριθμος που σε κάθε επανάληψη επιλέγει το σύνολο που περιέχει τα περισσότερα ακάλυπτα στοιχεία έχει λόγο προσέγγισης $1 - (1 - \frac{1}{k})^k > 1 - \frac{1}{e}$.
2. Να δείξετε ότι για κάθε $\alpha \in (0, 1)$, η ύπαρξη ενός $(1 - \alpha)$-προσεγγιστικού αλγόριθμου για το Maximum Coverage οδηγεί σε έναν $\lceil \log_{1/\alpha} n \rceil$-προσεγγιστικό αλγόριθμο για το πρόβλημα Set Cover.

Άσκηση 11. Δίνεται ένα πλήρες μη κατευθυνόμενο γράφημα $G(V, E, \ell)$, κάθε ακμή e του οποίου έχει μήκος $\ell(e) \geq 0$. Υποθέτουμε ότι τα μήκη των ακμών ικανοποιούν την τριγωνική ανισότητα. Το ζητούμενο είναι ένας κύκλος Hamilton που το μέγιστο βάρος των ακμών του είναι το ελάχιστο δυνατόν. Να δείξετε ότι δεν υπάρχει προσεγγιστικός αλγόριθμος (πολυωνυμικού χρόνου) με λόγο προσέγγισης μικρότερο του 2 για αυτό το πρόβλημα (εκτός αν $\mathbf{P} = \mathbf{NP}$). Τι συμβαίνει αν τα μήκη των ακμών είναι αυθαίρετα;