Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Προηγμένα Θέματα Αλγορίθμων

Αλγόριθμοι Δικτύων και Πολυπλοκότητα

Εαρινό εξάμηνο 2019-2020

(ΕΜΠ – ΑΛΜΑ)

Διδάσκοντες: Α. Παγουρτζής – Δ. Φωτάκης

2η Σειρά Ασκήσεων

Προθεσμία παράδοσης: 27/5/2020

Άσκηση 1 (10 μον.). Θεωρούμε μια δημοσκόπηση για 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 $\mathbb{P}\mathrm{r}[|\hat{p} - p| \leq \varepsilon p] > 1 - \delta$, 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-Hoeffding 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 $p \in [0.1, 0.7]$ (και δείτε ότι αυτή η τιμή είναι ανεξάρτητη του πληθυσμού της χώρας!). Να υπολογίσετε ακόμη το ελάχιστο μέγεθος N' δείγματος (ως συνάρτηση των ε και δ) ώστε η εκτίμησή μας $\hat{p}'$ να ικανοποιεί $\mathbb{P}\mathrm{r}[|\hat{p}' - p| \leq \varepsilon] > 1 - \delta$. Ποια είναι η τιμή του N' για για $\varepsilon = 0.02$ και $\delta = 0.05$;

Άσκηση 2 (Sparsification, 12 μον.). (α) Έστω $\boldsymbol{a}, \boldsymbol{x} \in [0, 1]^n$, με $\sum_i x_i = 1$ (θα λέμε ότι το $\boldsymbol{x}$ είναι ένα διάνυσμα πιθανοτήτων στο $[n] \equiv \{1, \dots, n\}$). Έστω ακόμη $k(\varepsilon) = \lceil \ln(2)/(2\varepsilon^2) \rceil + 1$. Να δείξετε ότι για κάθε $\varepsilon > 0$, υπάρχει ένα $k(\varepsilon)$-ομοιόμορφο διάνυσμα πιθανοτήτων $\boldsymbol{y}$ στο $[n]$ τέτοιο ώστε $|\boldsymbol{a} \cdot \boldsymbol{x} - \boldsymbol{a} \cdot \boldsymbol{y}| \leq \varepsilon$. Ένα διάνυσμα πιθανοτήτων $\boldsymbol{y}$ είναι k-*ομοιόμορφο* (k-uniform) αν κάθε y_i είναι ακέραιο πολλαπλάσιο του $1/k$.

(β) Έστω A πίνακας $m \times n$ με όλα τα στοιχεία του στο $[0, 1]$ και έστω $\boldsymbol{x}$ ένα διάνυσμα πιθανοτήτων στο $[n]$. Έστω ακόμη $k(m, \varepsilon) = \lceil \ln(2m)/(2\varepsilon^2) \rceil + 1$. Να δείξετε ότι για κάθε $\varepsilon > 0$, υπάρχει ένα $k(m, \varepsilon)$-ομοιόμορφο διάνυσμα πιθανοτήτων $\boldsymbol{y}$ στο $[n]$ τέτοιο ώστε $\|A\boldsymbol{x} - A\boldsymbol{y}\|_\infty \leq \varepsilon$.

Υπόδειξη: Να χρησιμοποιήσετε το ακόλουθο Hoeffding bound: Έστω $X_1, \dots, X_n$ ανεξάρτητες τυχαίες μεταβλητές στο $[0, 1]$, και έστω $X = \left(\sum_{i=1}^n X_i\right)/n$. Για κάθε $\varepsilon > 0$, $\mathbb{P}\mathrm{r}[|X - \mathbb{E}\mathrm{xp}[X]| > \varepsilon] \leq 2e^{-2n\varepsilon^2}$.

Άσκηση 3 (12 μον.). Crossing number $\mathrm{cr}(G)$ ενός γραφήματος G είναι ο ελάχιστος αριθμός συναντήσεων ακμών σε μια επίπεδη αποτύπωση του G. Το G είναι επίπεδο αν και μόνο αν $\mathrm{cr}(G) = 0$, ενώ π.χ., $\mathrm{cr}(_5) = \mathrm{cr}(_{3,3}) = 1$.

(α) Να δείξετε ότι για κάθε απλό γράφημα G με n κορυφές και m ακμές, $\mathrm{cr}(G) \geq m - 3n + 6$.

(β) Να δείξετε ότι για κάθε απλό γράφημα G με n κορυφές και $m \geq 4n$ ακμές, $\mathrm{cr}(G) \geq m^3/(64n^2)$.

Υπόδειξη: Για το (α), να χρησιμοποιήσετε ότι για κάθε απλό επίπεδο γράφημα με n κορυφές και m ακμές, $m \leq 3n - 6$. Για το (β), θεωρήστε ένα τυχαίο επαγόμενο υπογράφημα G' του G και υπολογίστε το αναμενόμενο πλήθος κορυφών και ακμών στο G', καθώς και το αναμενόμενο $\mathrm{cr}(G')$. Στη συνέχεια, χρησιμοποιήστε το (α).

Άσκηση 4 (Linear Integer Programming Formulations, 12 μον.). Να διατυπώσετε τα παρακάτω προβλήματα βελτιστοποίησης ως προβλήματα ακέραιου γραμμικού προγραμματισμού:

(α) Δίνονται m παράλληλες υπολογιστικές μηχανές και n υπολογιστικές διεργασίες. Κάθε διεργασία i έχει χρόνο εκτέλεσης $p_{ij} \in \mathbb{N}^*$ στην μηχανή j. Ζητείται να αναθέσουμε κάθε διεργασία σε μία υπολογιστική μηχανή ώστε ο συνολικός χρόνος εκτέλεσης των εργασιών στη μηχανή που έχει δεχθεί το μεγαλύτερο υπολογιστικό φορτίο να ελαχιστοποιηθεί.

(β) Δίνονται ένα σύνολο C με τις θέσεις n πελατών και ένα σύνολο F με m πιθανές θέσεις καταστημάτων. Δίνονται ακόμη το κόστος $f_j \in \mathbb{N}$ για να ανοίξουμε κατάστημα σε κάθε θέση $j \in F$ και οι αποστάσεις $d : C \times F \to \mathbb{N}$ για κάθε ζευγάρι θέσεων $(i, j) \in C \times F$. Ζητείται ένα υποσύνολο $F' \subseteq F$ θέσεων όπου θα ανοίξουμε καταστήματα, ώστε να ελαχιστοποιήσουμε το συνολικό κόστος για να ανοίξουμε καταστήματα στις θέσεις του F' και τη συνολική απόσταση κάθε πελάτη $i \in C$ από το κοντινότερό του ανοικτό κατάστημα.

Άσκηση 5 (Partial Vertex Cover, 14 μον.). Δίνονται ένα γράφημα $G(V, E)$ και μια παράμετρος $\beta > 0$. Ζητείται ένα υποσύνολο κορυφών $C \subseteq V$ που ελαχιστοποιεί το $\beta|C| + |U(C)|$, όπου $U(C) = \{\{u, v\} \in E : u, v \notin C\}$ είναι οι ακμές που δεν καλύπτονται από το C.

(α) Να διατυπώσετε το παραπάνω πρόβλημα ως πρόβλημα (γραμμικού) ακέραιου προγραμματισμού, να δώσετε το αντίστοιχο LP relaxation, και να βρείτε το αντίστοιχο δυϊκό γραμμικό πρόγραμμα. Να διατυπώσετε το πρόβλημα βελτιστοποίησης που περιγράφεται από το δυϊκό πρόγραμμα σε φυσική γλώσσα.

(β) Με βάση το LP relaxation του (α), να διατυπώσετε έναν προσεγγιστικό αλγόριθμο που βασίζεται σε deterministic rounding και να αναλύσετε τον λόγο προσέγγισης που επιτυγχάνει ο αλγόριθμός σας.

Άσκηση 6 (12 μον.). Να λύσετε την [1, Άσκηση **5.3**] και την [1, Άσκηση **5.6**].

Άσκηση 7 (Public Project, 10 μον.). Η δημοτική αρχή μελετά την κατασκευή ενός νέου πάρκου, το οποίο αναμένεται να κοστίσει C ευρώ. Κάθε δημότης i, $i = 1, \dots, n$, εκτιμά την ωφέλειά του από το νέο πάρκο σε $v_i \geq 0$ ευρώ, τιμή που είναι γνωστή μόνο στον ίδιο. Η δημοτική αρχή ζητάει από τους δημότες να δηλώσουν τις εκτιμήσεις τους $v_1, \dots, v_n$, και θα προχωρήσει στην κατασκευή του πάρκου μόνο αν $\sum_{i=1}^{n} v_i \geq C$. Σε αυτή την περίπτωση, κάθε δημότης θα πληρώσει μια εισφορά $p_i \geq 0$. Η εισφορά έχει στόχο οι δημότες να δηλώσουν τις πραγματικές τους εκτιμήσεις, και όχι να καλύψει το συνολικό κόστος του πάρκου. Να σχεδιάσετε έναν φιλαλήθη (truthful) μηχανισμό για αυτό το πρόβλημα.

Άσκηση 8 (Δημοπρασίες και Multiplicative Price Updates, 18 μον.). Θεωρούμε μια δημοπρασία για m διαφορετικά αντικείμενα (υποθέτουμε, για λόγους απλότητας, ότι κάθε αντικείμενο υπάρχει σε απεριόριστα αντίγραφα και ότι κάθε παίκτης μπορεί να πάρει το πολύ ένα αντίγραφο από κάθε αντικείμενο). Στη δημοπρασία συμμετέχουν n παίκτες. Κάθε παίκτης i έχει μια αύξουσα συνάρτηση αποτίμησης (valuation function) $v_i : 2^{[m]} \to \mathbb{R}_{\geq 0}$ που ορίζει πόσο αξίζει κάθε υποσύνολο αντικειμένων για αυτόν. Ο μηχανισμός ορίζει την αρχική τιμή κάθε αντικειμένου $p_1^j = p_0 = L/(4m)$, για κάποιο κατάλληλα επιλεγμένο $L > 0$. Στη συνέχεια, ζητά με τη σειρά από κάθε παίκτη $i \in [n]$ να επιλέξει το σύνολο S_i των αντικειμένων που μεγιστοποιεί τη συνάρτηση $v_i(S_i) - \sum_{j \in S_i} p_i^j$ και του το δίνει (ο παίκτης i μπορεί να επιλέξει το $\emptyset$, έτσι ισχύει πάντα ότι $v_i(S_i) - \sum_{j \in S_i} p_i^j \geq 0$). Πριν προχωρήσει στον επόμενο παίκτη, ο μηχανισμός διπλασιάζει τις τιμές όλων των αντικειμένων $j \in S_i$ πολλαπλασιαστικά, θέτοντας $p_{i+1}^j = 2p_i^j$ (θέτουμε ακόμη $p_{i+1}^j = p_i^j$, για κάθε $j \notin S_i$).

(α) Να διερευνήσετε αν ο παραπάνω μηχανισμός είναι φιλαλήθης (truthful). Να αιτιολογήσετε συνοπτικά την απάντησή σας.

(β) Υποθέτουμε ότι $L = \max_i\{v_i([m])\}$. Να δείξετε ότι ο μηχανισμός διαθέτει το πολύ $1 + \log_2(4m)$ αντίγραφα από κάθε αντικείμενο.

(γ) Να δείξετε ότι αν $L = \max_i\{v_i([m])\}$, τότε ο μηχανισμός επιτυγχάνει συνολική αξία $\sum_i v_i(S_i) \geq 3V^*/8$, όπου V^* η συνολική αξία της βέλτιστης λύσης που όμως διαθέτει το πολύ ένα αντίγραφο από κάθε αντικείμενο. Για να αποδείξετε αυτόν τον ισχυρισμό, πρέπει πρώτα να δείξετε (i) ότι λόγω της πολλαπλασιαστικής αναπροσαρμογής των τιμών, στο τέλος του μηχανισμού ισχύει ότι

$$\sum_{i \in [n]} v_i(S_i) \geq \sum_{j \in [m]} p_{n+1}^j - mp_0 \Rightarrow \sum_i v_i(S_i) \geq \sum_{j \in [m]} p_{n+1}^j - \frac{L}{4},$$

και (ii) ότι επειδή η βέλτιστη λύση δίνει το πολύ ένα αντίγραφο από κάθε αντικείμενο,

$$\sum_{i \in [n]} v_i(S_i) \geq V^* - \sum_{j \in [m]} p_{n+1}^j$$

Αναφορές

[1] D.P. Williamson and D.B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2010.