Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Προηγμένα Θέματα Αλγορίθμων

Αλγόριθμοι Δικτύων και Πολυπλοκότητα

Εαρινό εξάμηνο 2023-2024

(ΕΜΠ – ΑΛΜΑ)

Διδάσκοντες: Α. Παγουρτζής – Δ. Φωτάκης – Θ. Λιανέας – Ο. Πλευράκης

1η Σειρά Ασκήσεων

Προθεσμία παράδοσης: 10/5/2024

Άσκηση 1 (Min-Cut Algorithm, 1.5 μον.). Να λύσετε την [1, Άσκηση **1.24**].

Άσκηση 2 (Random Sampling, 1 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 $\delta = 0.05$; *Σημείωση:* Πρόκειται για παραλλαγή της [1, Άσκησης **4.5**].

Άσκηση 3 (Sparsification, 2 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 $[n] \equiv \{1, \ldots, n\}$). 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 $X_1, \ldots, X_n$ 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 4 (Capacited Max k-Cut, 1.5 μον.). Στο Max k-Cut πρόβλημα δίνεται απλό μη κατευθυνόμενο γράφημα $G(V, E, w)$, με θετικά ακέραια βάρη $w : E \to \mathbb{N}^*$ στις ακμές, και ζητείται μια διαμέριση των κορυφών σε $k \geq 2$ σύνολα $(S_1, \ldots, S_k)$, με $\bigcup_{i=1}^k S_i = V$, ώστε να μεγιστοποιηθεί το συνολικό βάρος των ακμών με άκρα σε διαφορετικά σύνολα. Ειδικότερα, αν συμβολίσουμε με

$$\delta(S_1, \ldots, S_k) = \Big\{ e = \{u, v\} \in E : v \in S_i, u \in S_j \text{ και } i \neq j \Big\}$$

το σύνολο των ακμών στο k-cut $(S_1, \ldots, S_k)$, ζητείται να μεγιστοποιήσουμε το

$$w(S_1, \ldots, S_k) = \sum_{e \in \delta(S_1, \ldots, S_k)} w(e)\,.$$

Θεωρούμε παραλλαγή του προβλήματος Max k-Cut όπου τα επιθυμητά μεγέθη $(c_1, \ldots, c_k)$, με $c_1 + \cdots + c_k = |V|$ των συνόλων που ορίζουν το k-cut είναι δεδομένα. Εφαρμόζουμε τον πιθανοτικό αλγόριθμο που τοποθετεί κάθε κορυφή v στο σύνολο S_i τυχαία, με πιθανότητα $c_i/|V|$, και ανεξάρτητα. Να υπολογίσετε την αναμενόμενη τιμή του $w(S_1, \ldots, S_k)$ (και να διερευνήσετε πότε μεγιστοποιείται). Να υπολογίσετε ακόμη άνω και κάτω φράγματα (ώστε να ισχύουν με μεγάλη πιθανότητα) για τα μεγέθη των συνόλων που προκύπτουν.

Άσκηση 5 (1 μον.). Θεωρούμε τον Αλγόριθμο Υποδιπλασιασμού (Halving Algorithm, όπως παρουσιάζεται στις διαφάνειες τις αντίστοιχης διάλεξης), που επιδιώκει να ελαχιστοποιήσει το πλήθος των λαθών σε περιβάλλον online learning. Εξετάζουμε την περίπτωση που η κλάση υποθέσεων $\mathcal{H}$ είναι πεπερασμένη και τα δείγματα κατηγοριοποιούνται με βάση υπόθεση $f \in \mathcal{H}$ (realizability).

Να δείξετε ότι αν τα δείγματα $(x_t, f(x_t)$ έρχονται από μια οποιαδήποτε (άγνωστη, αλλά συγκεκριμένη) κατανομή $\mathcal{D}$ και το σύνολο έγκυρων υποθέσεων S_t δεν μεταβάλλεται για $\Omega(\log(1/\delta)/\varepsilon)$ συνεχόμενα δείγματα, τότε με πιθανότητα τουλάχιστον $1-\delta$, κάθε έγκυρη υπόθεση $h \in S_t$ επιτυγχάνει loss $L_{(\mathcal{D},f)}(h) \leq \varepsilon$ (δηλαδή έχουμε επιτύχει την εγγύηση του PAC Learning). Ποια είναι η δειγματική πολυπλοκότητα του αλγορίθμου για πεπερασμένη κλάση υποθέσεων $\mathcal{H}$;

Άσκηση 6 (1 μον.). Να επαναλάβετε την ανάλυση του Weighted Majority Algorithm (WMA, όπως παρουσιάζεται στις διαφάνειες της αντίστοιχης διάλεξης) για την περίπτωση που το βάρος εμπιστοσύνης κάθε ειδικού / υπόθεσης $h \in \mathcal{H}$ πολλαπλασιάζεται με $(1-\varepsilon)$ (αντί για $1/2$) κάθε φορά που ο ειδικός h κάνει λάθος (έχουμε δηλαδή ότι για κάθε h με $h(x_t) \neq y_t$, $w_{t+1}(h) = (1-\varepsilon)w_t(h)$).

Άσκηση 7 (Ανάλυση Regret του Αλγόριθμου Hedge, 2 μον.). Θεωρούμε τον παρακάτω αλγόριθμο για online learning / online linear optimization και αναλύουμε το regret που επιτυγχάνει.

- **Input**: n actions $\{1,\dots,n\}$, time horizon T, $\boldsymbol{w}_1 = (1,\dots,1)$, $\boldsymbol{x}_1 = (1/n,\dots,1/n)$
- for $t = 1$ to T do:
  - Select action $i_t \in \{1,\dots,n\}$ with probability $\boldsymbol{x}_t(i_t)$
  - Get loss $\boldsymbol{\ell}_t \in [0,1]^n$ for all actions and incur loss $\boldsymbol{\ell}_t(i_t)$
  - Update weights $\boldsymbol{w}_{t+1}(i) = \boldsymbol{w}_t(i)e^{-\eta\boldsymbol{\ell}_t(i)}$, for all $i \in \{1,\dots,n\}$
  - Update probabilities $\boldsymbol{x}_{t+1}(i) = \frac{\boldsymbol{w}_{t+1}(i)}{\sum_{i=1}^n w_{t+1}(i)}$, for all $i \in \{1,\dots,n\}$

(α) Θεωρούμε τη συνάρτηση δυναμικού $\Phi(t) = \sum_{i=1}^n w_t(i)$ (συνολικό βάρος εμπιστοσύνης των actions τη χρονική στιγμή t). Αρχικά είναι $\Phi(1) = n$. Να δείξετε ότι $\Phi(T) \geq e^{-\eta\sum_{t=1}^T \boldsymbol{\ell}_t(i^*)}$, όπου $i^* = \arg\min_i \sum_{t=1}^T \boldsymbol{\ell}_t(i)$ η βέλτιστη επιλογή.

(β) Να δείξετε ακόμη ότι για κάθε $t \geq 1$,

$$\Phi(t+1) \leq \Phi(t)e^{-\eta\boldsymbol{x}_t\cdot\boldsymbol{\ell}_t + \eta^2\boldsymbol{x}_t\boldsymbol{\ell}_t^2}\,,$$

όπου $\boldsymbol{\ell}_t^2(i) = (\boldsymbol{\ell}_t(i))^2$, για όλα τα $i \in \{1,\dots,n\}$. Χρησιμοποιώντας αυτή τη σχέση, να βρείτε ένα άνω φράγμα για το $\Phi(T)$ ως συνάρτηση του $\Phi(1) = n$.

(γ) Χρησιμοποιώντας τα άνω και κάτω φράγματα για την $\Phi(T)$ που υπολογίσατε στα (α) και (β) και το γεγονός ότι $\boldsymbol{\ell}_t \in [0,1]^n$, για κάθε $t \in \{1,\dots,T\}$, να δείξετε ότι:

$$\text{Regret}(T) = \sum_{t=1}^T \ell_t(i_t) - \sum_{t=1}^T \ell_t(i^*) \leq \eta T + \ln(n)/\eta$$

Ποια τιμή του η θα επιλέγατε; Ποιο είναι το Regret(T) που επιτυγχάνει ο αλγόριθμος για αυτή την τιμή του η;

Αναφορές

[1] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.