**Εθνικό Μετσόβιο Πολυτεχνείο**

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

**Κυρτή Βελτιστοποίηση με Εφαρμογές στη Μηχανική Μάθηση**

Διδάσκοντες: Κ. Χρυσαφίνος, Δ. Φωτάκης

**2η Σειρά Ασκήσεων, Ημ/νια Παράδοσης: 10/7/2024**

---

**Άσκηση 1.** Θεωρούμε τον Halving Learner που επιδιώκει να ελαχιστοποιήσει το πλήθος των λαθών του classifier σε περιβάλλον online learning. Να δείξετε ότι αν τα δείγματα έρχονται από μια οποιαδήποτε (άγνωστη, αλλά συγκεκριμένη) κατανομή $\mathcal{D}$ και το σύνολο έγκυρων υποθέσεων $V_t$ δεν μεταβάλλεται για $\Omega(\log(1/\delta)/\varepsilon)$ συνεχόμενα δείγματα, τότε με πιθανότητα τουλάχιστον $1-\delta$, κάθε έγκυρη υπόθεση $h \in V_t$ επιτυγχάνει loss $L_{(\mathcal{D},h^*)}(h) \leq \varepsilon$ (δηλαδή έχουμε επιτύχει το guarantee του PAC Learning – εδώ υποθέτουμε realizability και $h^*$ είναι μια υπόθεση με μηδενικό λάθος). Ποια είναι η δειγματική πολυπλοκότητα που επιτυγχάνουν οι δύο αλγόριθμοι για μια πεπερασμένη κλάση υποθέσεων $\mathcal{H}$;

**Άσκηση 2.** Να διατυπώσετε τον αλγόριθμο Follow the Leader (FTL) για την "πρόβλεψη" μιας ακολουθίας σημείων $\vec{x}_1, \dots, \vec{x}_T \in B$, όπου $B = \{\vec{x} \in \mathbb{R}^n : \|\vec{x}\|_2 \leq 1\} \subseteq \mathbb{R}^n$ η μοναδιαία μπάλα στις $n$ διαστάσεις, όταν η συνάρτηση απώλειας (loss function) είναι τετραγωνική. Δηλαδή, σε κάθε γύρο $t = 1, 2, \dots, T$, ο FTL επιλέγει σημείο $\vec{p}_t \in B$ και εμφανίζει απώλεια $\ell(\vec{p}_t, \vec{x}_t) = \|\vec{p}_t - \vec{x}_t\|_2^2$.

1. Να εκφράσετε κάθε σημείο $\vec{p}_t$ που επιλέγεται από τον FTL ως συνάρτηση των σημείων $\vec{x}_1, \dots, \vec{x}_{t-1}$ της ακολουθίας (μπορείτε να θεωρήσετε ότι $\vec{p}_1 = 0$).

2. Να δείξετε ότι ο FTL εμφανίζει regret $O(\log T)$ για μια τέτοια ακολουθία. Θα χρειαστεί να χρησιμοποιήσετε ότι τα $\vec{x}_1, \dots, \vec{x}_T$ ανήκουν στη μοναδιαία μπάλα, και να προσπαθήσετε να μιμηθείτε την απόδειξη για το regret του Online Gradient Descent για την αντίστοιχη κατηγορία συναρτήσεων.

3. Να δώσετε μια ακολουθία σημείων $\vec{x}_1, \dots, \vec{x}_T \in B$ για την οποία το regret του FTL είναι $\Omega(\log T)$ (και πάλι μπορείτε να θεωρήσετε ότι $\vec{p}_1 = 0$).

**Άσκηση 3.** Στην μέθοδο Support Vector Machines (SVM), χρησιμοποιούμε ως συνάρτηση κόστους (loss function) την $\ell((\vec{x}, y), \vec{w}) = \max\{0, 1 - y\, \vec{w} \cdot \vec{x}\}$ και υπολογίζουμε τον classifier $\vec{w}$ που ελαχιστοποιεί την παρακάτω συνάρτηση regularized loss:

$$f(\vec{w}) = \lambda\, \mathbb{E}\mathrm{xp}_{(\vec{x},y)\sim\mathcal{D}}\big[\ell((\vec{x}, y), \vec{w})\big] + \frac{\|\vec{w}\|^2}{2}\,,$$

όπου $\lambda$ είναι το regularization factor και $(\vec{x}, y)$ τυχαίο δείγμα από άγνωστη κατανομή $\mathcal{D}$. Να εξειδικεύσετε τον αλγόριθμο Stochastic Gradient Descent για την μέθοδο Support Vector Machines. Τι βήμα $\eta$ (ή $\eta_t$) θα χρησιμοποιήσετε και τι regret θα επιτύχετε με αυτό;

**Άσκηση 4.** Θεωρούμε τον παρακάτω αλγόριθμο $A$ για Online Linear Optimization, για τον οποίο θα αναλύσουμε το regret που επιτυγχάνει.

- **Input**: $n$ actions $\{1, \dots, n\}$, time horizon $T$, $\vec{w}_1 = (1, \dots, 1)$, $\vec{x}_1 = (1/n, \dots, 1/n)$
- for $t = 1$ to $T$ do:
  - Select action $i_t \in \{1, \dots, n\}$ with probability $\vec{x}_t(i_t)$

- Get loss $\vec{\ell}_t \in [0,1]^n$ for all actions and incur loss $\vec{\ell}_t(i_t)$
- Update weights $\vec{w}_{t+1}(i) = \vec{w}_t(i) e^{-\eta \vec{\ell}_t(i)}$, for all $i \in \{1, \dots, n\}$
- Update probabilities $\vec{x}_{t+1}(i) = \frac{\vec{w}_{t+1}(i)}{\sum_{i=1}^n w_{t+1}(i)}$, 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 $\Phi(T) \geq e^{-\eta \sum_{t=1}^T \vec{\ell}_t(i^*)}$, όπου $i^* = \arg\min_i \sum_{t=1}^T \vec{\ell}_t(i)$ η βέλτιστη επιλογή.

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

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

όπου $\vec{\ell}_t^2(i) = (\vec{\ell}_t(i))^2$, 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 $\vec{\ell}_t \in [0,1]^n$, για κάθε $t \in \{1, \dots, T\}$, να δείξετε ότι:

$$\text{Regret}_A(T) \leq \eta T + \ln(n)/\eta$$

Ποια τιμή του $\eta$ θα επιλέγατε και ποιο είναι το $\text{Regret}_A(T)$ για αυτή;

(δ) Να εντοπίσετε (και να αναφέρετε συνοπτικά) τις ομοιότητες και τις διαφορές της παραπάνω ανάλυσης σε σχέση με την ανάλυση του EXP3 αλγόριθμου για adversarial bandits που είδαμε στις διαλέξεις.

**Άσκηση 5.** Θεωρούμε ένα stochastic bandit σενάριο με 2 βραχίονες (arms), όπου η ακολουθία κερδών για κάθε βραχίονα $i \in \{1, 2\}$ προκύπτει ως ακολουθία ανεξάρτητων δειγμάτων από μια κατανομή Bernoulli. Οι μέσες τιμές των δύο κατανομών Bernoulli είναι $a$ και $b$ (τα οποία είναι γνωστά), με $0 < a < b < 1$ και $\Delta = b - a$. Εκείνο που δεν γνωρίζουμε είναι ποιος βραχίονας αντιστοιχεί στην κατανομή με υψηλή μέση τιμή $b$. Γνωρίζουμε ότι ο $UCB$ εξασφαλίζει regret $O((\log T)/\Delta)$ σε αυτό το σενάριο, αλλά ο αλγόριθμος (και η ανάλυση) δεν εκμεταλλεύονται την πλήρη γνώση των κατανομών κέρδους για τους δύο βραχίονες.
Θεωρούμε τον ακόλουθο αλγόριθμο: Για $t = i$, $i \in \{1, 2\}$, επιλέγουμε τον $i$-οστό βραχίονα. Για κάθε $t \geq 3$, αν υπάρχει βραχίονας με μέσο εμπειρικό κέρδος μεγαλύτερο του $(a+b)/2$, επιλέγουμε τον βραχίονα με το υψηλότερο μέσο εμπειρικό κέρδος. Διαφορετικά, επιλέγουμε την 1ο βραχίονα για τη στιγμή $t$ και τον 2ο βραχίονα για τη στιγμή $t+1$. Θέλουμε να δείξουμε ότι το regret αυτού του αλγόριθμου εξαρτάται μόνο από το $\Delta$ (και όχι από τον χρονικό ορίζοντα $T$).

1. Έστω $t \geq 3$. Χρησιμοποιώντας την ανισότητα Hoeffding[1], να δείξετε ότι η πιθανότητα ο βραχίονας με μέση τιμή $a$ να εμφανίζει μέσο εμπειρικό κέρδος μεγαλύτερο του $(a+b)/2$ (και άρα να είναι επιλέξιμος) τη χρονική στιγμή $t$ δεν ξεπερνά το $\exp(-k\Delta^2/2)$, όπου $k \leq t$ το πλήθος των χρονικών στιγμών που έχουμε επιλέξει αυτόν τον βραχίονα στο παρελθόν.

2. Έστω $t \geq 3$. Χρησιμοποιώντας την ανισότητα Hoeffding, να δείξετε ότι η πιθανότητα ο βραχίονας με μέση τιμή $b$ να εμφανίζει μέσο εμπειρικό κέρδος μικρότερο του $(a+b)/2$ τη χρονική στιγμή $t$ (γεγονός που καθιστά επιλέξιμο τον βραχίονα με μέση τιμή $a$) δεν ξεπερνά το $\exp(-k\Delta^2/2)$, όπου $k \leq t$ το πλήθος των χρονικών στιγμών που έχουμε επιλέξει αυτόν τον βραχίονα στο παρελθόν.

[1] Έστω $X_1, \dots, X_n$ ανεξάρτητες Bernoulli τυχαίες μεταβλητές με μέση τιμή $\mu$. Η ανισότητα Hoeffding δίνει $\text{Prob}[X_1 + \cdots + X_n \geq n(\mu + \varepsilon)] \leq \exp(-2n\varepsilon^2)$, για κάθε $n \in \mathbb{N}_+$ και κάθε $\varepsilon > 0$. Παρατηρήστε ακόμη ότι $\frac{1}{e^x - 1} \leq \frac{1}{x}$, για κάθε $x > 0$.

3. Να δώσετε ένα άνω φράγμα στο αναμενόμενο regret του αλγορίθμου, χρησιμοποιώντας το $\Delta$ και τις πιθανότητες που υπολογίσατε στα προηγούμενα δύο βήματα. Χρησιμοποιώντας το γεγονός ότι οι πιθανότητες αυτές μειώνονται εκθετικά με το $k$, να καταλήξετε στο συμπέρασμα ότι το regret του αλγορίθμου εξαρτάται μόνο από το $\Delta$ (και όχι από τον χρονικό ορίζοντα $T$).

**Άσκηση 6.** Να γενικεύσετε την ανάλυση και να δώσετε ένα άνω φράγμα στο regret του Upper Confidence Bounds (UCB) Elimination αλγόριθμου για stochastic bandits για την περίπτωση όπου το πλήθος των arms είναι $K \geq 3$ (στις διαλέξεις αναλύσαμε το regret του UCB Elimination για $K = 2$ arms).

**Άσκηση 7.** Έστω $\mathcal{D} = \{\mathcal{P}_1, \ldots, \mathcal{P}_n\}$ κλάση αποτελούμενη από $n$ κατανομές. Για την κλάση κατανομών $\mathcal{D}$, γνωρίζουμε hypothesis testing αλγόριθμο $A(s_1, \ldots, s_k, \mathcal{P}_1, \mathcal{P}_2, \varepsilon, \delta)$, ο οποίος με είσοδο κατανομές $\mathcal{P}_1, \mathcal{P}_2 \in \mathcal{D}$ σε total variation distance μεγαλύτερη του $\varepsilon$ και $k = k(\varepsilon, \delta)$ ανεξάρτητα δείγματα που έχουν ληφθεί είτε από την $\mathcal{P}_1$ είτε από την $\mathcal{P}_2$, εντοπίζει σωστά την κατανομή από την οποία προέρχονται τα δείγματα (μεταξύ των $\mathcal{P}_1$ και $\mathcal{P}_2$) με πιθανότητα τουλάχιστον $1 - \delta$. Στην περίπτωση που τα δείγματα εισόδου έχουν ληφθεί από κατανομή $\mathcal{P} \in \mathcal{D}$ διαφορετική των $\mathcal{P}_1, \mathcal{P}_2$ ή οι $\mathcal{P}_1, \mathcal{P}_2$ βρίσκονται σε total variation distance το πολύ $\varepsilon$, η απάντηση του $A$ μπορεί να είναι οποιαδήποτε από τις $\mathcal{P}_1, \mathcal{P}_2$ (και δεν έχουμε καμία εγγύηση για αυτή την περίπτωση).
Με βάση τον αλγόριθμο $A$, να διατυπώσετε αλγόριθμο μη επιβλεπόμενης μάθησης για την κλάση κατανομών $\mathcal{D}$ που επιτυγχάνει όσο το δυνατόν μικρότερη δειγματική πολυπλοκότητα. Ειδικότερα, ο αλγόριθμός σας για κάθε $\varepsilon, \delta \in [0, 1]$, θα πρέπει να καθορίζει ένα πλήθος δειγμάτων $m = m(\varepsilon, \delta)$. Στη συνέχεια, με είσοδο $m$ ανεξάρτητα δείγματα από άγνωστη κατανομή $\mathcal{P} \in \mathcal{D}$, ο αλγόριθμος θα πρέπει να υπολογίζει και να επιστρέφει κατανομή $\mathcal{P}' \in \mathcal{D}$ σε total variation distance το πολύ $\varepsilon$ από την $\mathcal{P}$ με πιθανότητα τουλάχιστον $1 - \delta$. Ποια είναι η δειγματική πολυπλοκότητα $m(\varepsilon, \delta)$ του αλγορίθμου σας;

**Υποβολή.** Οι εργασίες πρέπει να αναρτηθούν στο `https://helios.ntua.gr/mod/assign/view.php?id=24409` μέχρι τα μεσάνυχτα της Τετάρτης 10 Ιουλίου.