ΘΕΩΡΙΑ ΠΑΙΓΝΙΩΝ – ΤΕΛΙΚΗ ΕΞΕΤΑΣΗ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΣΕΕ / ΑΛΜΑ (2023–2024)

ΦΕΒΡΟΥΑΡΙΟΣ 2024

Προθεσμία: Δευτέρα 12 Φεβρουαρίου, 23:00.

Οδηγίες: Λύστε όλα τα θέματα, και **καθαρογράψτε τις λύσεις σας.** Ο τρόπος και η παρουσιάση της σκέψης σας μετράνε πιο πολύ από το τελικό αποτέλεσμα. Αν καταφέρετε να λύσετε όλα τα θέματα, συγχαρητήρια· αν δεν τα καταφέρετε, μη νιώσετε άσχημα (δε χρειάζεται να λύσετε όλα τα θέματα για να πάρετε το μέγιστο βαθμό). Το βαθμολογικό σχήμα έχει μείνει εσκεμμένα εκτός εξέτασης: ο σκοπός της εξέτασης είναι να σας κάνει να σκεφτείτε και να δουλέψετε στις έννοιες και τις τεχνικές που είδαμε στο μάθημα, όχι να ενθαρρύνει στρατηγικές μεγιστοποίησης βαθμολογίας.

Τυπογραφικά: Σχεδόν σίγουρα, όλο και κάτι θα έχει ξεφύγει. Αν εντοπίσετε κάτι, ενημερώστε με.

Θέμα 1 (*Παίγνια συμφόρησης*). Έστω ένα δίκτυο που αποτελείται από 2 σημεία, A και B, που ενώνονται από n παράλληλους δρόμους (όπως το δίκτυο Pigou που είδαμε στο μάθημα αλλά με n δρόμους αντί για 2). Η συνάρτηση κόστους του k-στού δρόμου είναι $c_k(x) = a_k x$, με $a_k > 0$ για κάθε $k = 1, \ldots, n$.

(1) Υπολογίστε το τίμημα της αναρχίας (price of anarchy) του δικτύου όταν η συνολική εισροή κίνησης είναι $M = 1$.

(2) Επαναλάβετε τον παραπάνω υπολογισμό για γενικό $M > 0$, και βρείτε το όριο του τιμήματος της αναρχίας όταν $M \to 0^+$ ή $M \to \infty$. ⁂

Θέμα 2 (*Σημεία ισορροπίας*). Υπολογίστε όλα τα σημεία ισορροπίας των διπινακοπαιγνίων

$$\begin{matrix} (1,1) & (0,0) & (0,0) \\ (0,0) & (2,2) & (0,0) \\ (0,0) & (0,0) & (3,3) \end{matrix} \qquad \text{και} \qquad \begin{matrix} (1,-1) & (0,0) & (0,0) \\ (0,0) & (-2,2) & (0,0) \\ (0,0) & (0,0) & (3,-3) \end{matrix}$$

Σημείωση: Αυτό που ενδιαφέρει στο εν λόγω θέμα είναι η σκέψη και η ανάλυσή σας για την εύρεση των σημείων ισορροπίας, όχι το τελικό αποτέλεσμα. ⁂

Θέμα 3 (*Supinf στρατηγικές*). Έστω δύο αριθμήσιμα σύνολα A, B και μία αντικειμενική συνάρτησης ελαχιστομεγιστοποιήσης $f\colon A \times B \to \mathbb{R}$. Να δείξετε ότι

$$\sup_{a\in A} \inf_{b\in B} f(a,b) = \inf_{\beta:A\to B} \sup_{a\in A} f(a,\beta(a)) \tag{1}$$

Σημείωση: Ο περιορισμός σε αριθμήσιμα σύνολα δεν έχει σημασία, αλλά αν σας διευκολύνει, μπορείτε να το υποθέσετε. ⁂

Θέμα 4 (*Σημεία αυστηρής ισορροπίας*). Έστω ένα πεπερασμένο παίγνιο $\Gamma \equiv \Gamma(\mathcal{N}, \mathcal{A}, u)$. Υπενθυμίζεται ότι ένα προφίλ μεικτών στρατηγικών $x^* = (x_1^*, \dots, x_N^*)$ είναι *σημείο αυστηρής ισορροπίας* όταν

$$u_i(x_i^*; x_{-i}^*) > u_i(x_i; x_{-i}^*) \quad \text{για κάθε } x_i \in \mathcal{X}_i \setminus \{x_i^*\} \text{ και για κάθε } i \in \mathcal{N}.$$

Αν $v_i(x) = (u_i(\alpha_i; x_{-i}))_{\alpha_i \in \mathcal{A}_i}$ το διάνυσμα μεικτών πληρωμών του παίκτη $i \in \mathcal{N}$ στο προφίλ μεικτών στρατηγικών $x = (x_1, \dots, x_N)$, να δείξετε ότι τα παρακάτω είναι ισοδύναμα:

(1) Το x^* είναι σημείο αυστηρής ισορροπίας.

(2) Κάθε x_i^*, $i \in \mathcal{N}$, στηρίζεται σε μία μοναδική αμιγή στρατηγική α_i^* (δηλαδή $\operatorname{supp}(x_i^*) = \{\alpha_i^*\}$ για κάθε $i \in \mathcal{N}$) και, επιπλέον, $u_i(\alpha^*) > u_i(\alpha_i; \alpha_{-i}^*)$ για κάθε $\alpha_i \in \mathcal{A}_i \setminus \{\alpha_i^*\}$ και για κάθε $i \in \mathcal{N}$.

(3) $\sum_{i \in \mathcal{N}} \langle v_i(x^*), x_i - x_i^* \rangle < 0$ για κάθε $x \in \mathcal{X} \setminus \{x^*\}$.

(4) $\sum_{i \in \mathcal{N}} \langle v_i(x), x_i - x_i^* \rangle < 0$ για κάθε $x \neq x^*$ σε κάποια (πιθανώς μικρή) περιοχή του x^*.

Χρησιμοποιώντας την τελευταία ιδιότητα δείξτε ότι τα σημεία αυστηρής ισορροπίας είναι ασυμπτωτικά ευσταθή ως προς τη δυναμική των αντιγραφέων. [Υπόδειξη: θεωρήστε τη συνάρτηση $E(x) = -\sum_{i \in \mathcal{N}} \log x_{i\alpha_i^*}$ και μελετήστε την εξέλιξή της ως προς τη δυναμική των αντιγραφέων.] ⁂

Θέμα 5 (*Παίγνια δυναμικού*). Ένα πεπερασμένο παίγνιο $\Gamma \equiv \Gamma(\mathcal{N}, \mathcal{A}, u)$ καλείται παίγνιο δυναμικού όταν επιδέχεται μία συνάρτηση $\phi\colon \mathcal{A} \to \mathbb{R}$ τέτοια ώστε

$$u_i(\alpha_i; \alpha_{-i}) - u_i(\beta_i; \alpha_{-i}) = \phi(\alpha_i; \alpha_{-i}) - \phi(\beta_i; \alpha_{-i}) \quad \text{για κάθε προφίλ } \alpha, \beta \in \mathcal{A} \text{ και κάθε } i \in \mathcal{N}.$$

(1) Να δείξετε ότι ένα παίγνιο είναι παίγνιο δυναμικού αν και μόνο αν επιδέχεται μία συνάρτηση $f\colon \mathcal{X} \to \mathbb{R}$ ώστε

$$\langle v(x), x' - x \rangle = \left.\frac{d}{dt}\right|_{t=0} f(x + t(x' - x)) \quad \text{για κάθε } x, x' \in \mathcal{X}. \tag{2}$$

(2) Εξετάστε αν η παραπάνω συνθήκη είναι ισοδύναμη με την ύπαρξη μίας συνάρτησης $p\colon \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$ ώστε $v(x) = \nabla p(x)$ για κάθε $x \in \mathcal{X}$.

(3) Χρησιμοποιώντας τα παραπάνω, να δείξετε ότι σε ένα παίγνιο δυναμικού, ή δυναμική των αντιγραφέων συγκλίνει σε ισορροπία Nash από οποιαδήποτε πλήρως μεικτή αρχική συνθήκη. ⁂

Θέμα 6 (*Βρείτε το παιχνίδι*). Έστω ένα παίγνιο 2×2 όπου ο πρώτος παίκτης επιλέγει μεταξύ των στρατηγικών "T" και "B" (top/bottom), ενώ ο δεύτερος παίκτης επιλέγει μεταξύ των στρατηγικών "L" και "R" (left/right). Συμβολίζουμε με x την πιθανότητα με την οποία ο πρώτος παίκτης παίζει T, και y την πιθανότητα με την οποία ο δεύτερος παίκτης παίζει L.

Δίνονται οι παρακάτω δυναμικές

(1) $\dot{x} = x(1-x)(4y-1)$ και $\dot{y} = y(1-y)(4x-3)$

(2) $\dot{x} = x(1-x)(2y+1)$ και $\dot{y} = y(1-y)(2x+1)$

(3) $\dot{x} = -3x(1-x)y$ και $\dot{y} = -3y(1-y)x$

Για κάθε μία από τις παραπάνω περιπτώσεις, να βρείτε έναν πίνακα πληρωμών τέτοιο ώστε η δυναμική των αντιγραφέων για το παίγνιο να παίρνει την εκάστοτε μορφή. ⁂

Θέμα 7 (*Bandidos*). Αυτό το θέμα αφορά μία διαφορετική εκδοχή του αλγορίθμου EXP3 που είδαμε στο μάθημα. Το βασικό πρόβλημα αφορά έναν ανταγωνιστικό ληστή (adversarial multi-armed bandit): κάθε χρονική στιγμή $t = 1, 2, \dots$, ο παίκτης επιλέγει μία δράση α_t από ένα πεπερασμένο σύνολο δράσεων $\mathcal{A} = \{1, \dots, A\}$. Ταυτόχρονα, το "περιβάλλον" επιλέγει ένα διάνυσμα πληρωμών $v_t = (v_{1,t}, \dots, v_{A,t}) \in [0,1]^A$, και ο παίκτης λαμβάνει ως πληρωμή τη συνιστώσα $v_{\alpha_t,t}$ του v_t που αντιστοιχεί στη δράση α_t που επέλεξε.

Όπως συζητήσαμε στο μάθημα, αν ο παίκτης επιλέξει τη δράση $\alpha_t \in \mathcal{A}$ τη χρονική στιγμή t με βάση τη μεικτή στρατηγική $x_t \in \mathcal{X} \coloneqq \Delta(\mathcal{A})$, η μεταμέλεια του παίκτη ως προς μία σταθερή μεικτή στρατηγική $p \in \mathcal{X}$ ορίζεται ως

$$\mathrm{Reg}_p(T) = \sum_{t=1}^{T} \langle v_t, p - x_t \rangle$$

Στο πρόβλημα αυτό υποθέτουμε ότι η μόνη πληροφορία που έχει ο παίκτης τη χρονική στιγμή $t = 1, 2, \dots$ είναι η πληρωμή $u_t = v_{\alpha_t,t}$ που έλαβε. Με βάση αυτήν την παρατήρηση, ο παίκτης μπορεί να υπολογίσει τον εκτιμητή σταθμισμένης σημαντικότητας (importance weighted estimator)

$$\hat{v}_{\alpha,t} = \frac{\mathbb{1}\{\alpha = \alpha_t\}}{x_{\alpha,t}} u_t \quad \text{για κάθε } \alpha \in \mathcal{A},\ t = 1, 2, \dots \tag{IWE}$$

και ο αλγόριθμος που θα μελετήσουμε είναι ο εξής:

$$\begin{aligned} \text{Μεικτή στρατηγική:} &\quad x_t = (1-\delta)\,\Lambda(y_t) + \delta\,\mathrm{unif} \\ \text{Επιλογή δράσης:} &\quad \alpha_t \sim x_t \\ \text{Βήμα επανάληψης:} &\quad y_{t+1} = y_t + \eta \hat{v}_t \end{aligned}$$

όπου, όπως στην περίπτωση του EXP3:

- Το διάνυσμα $y_t \in \mathbb{R}^{\mathcal{A}}$, $t = 1, 2, \dots$, είναι ένα διάνυσμα "βαθμολογίας" των δράσεων του παίκτη, που θεωρούμε ότι αρχικοποιείται ως $y_1 \leftarrow 0$.
- Η απεικόνιση $\Lambda\colon \mathbb{R}^{\mathcal{A}} \to \mathcal{X}$ δίνεται από τη σχέση

$$\Lambda(y) = \frac{(e^{y_1}, \dots, e^{y_A})}{e^{y_1} + \cdots + e^{y_A}}$$

- Ο συμβολισμός $\mathrm{unif} \coloneqq (1/A, \dots, 1/A)$ αναφέρεται στην ομοιόμορφη κατανομή στο $\mathcal{A}$.
- Τα $\delta, \eta > 0$ είναι παράμετροι του αλγορίθμου.

Απαντήστε στα παρακάτω ερωτήματα:

(1) Υπολογίστε τη μέση τιμή $\mathbb{E}[\hat{v}_{\alpha,t}]$ του εκτιμητή (IWE) και δείξτε ότι ικανοποιεί τη σχέση

$$\max_{\alpha \in \mathcal{A}} |\hat{v}_{\alpha,t}| \leq \frac{A}{\delta} \qquad \text{για κάθε } \alpha = 1, \dots, A.$$

(2) Θεωρήστε τη συνάρτηση ενέργειας (Fenchel coupling)

$$F(p, y) = \sum_{\alpha \in \mathcal{A}} p_\alpha \log p_\alpha + \log \sum_{\alpha \in \mathcal{A}} \exp(y_\alpha) - \langle y, p \rangle$$

που χρησιμοποιήσαμε στο μάθημα για την ανάλυση των αλγορίθμων εκθετικών/πολλαπλασιαστικών βαρών (exponential/multiplicative weights). Αν $F_t = F(p, y_t)$, δείξτε ότι

$$F_{t+1} \leq F_t + \frac{\eta}{1-\delta} \langle \hat{v}_t, x_t - p \rangle + \frac{\delta\eta}{1-\delta} \langle \hat{v}_t, p - \mathrm{unif} \rangle + \frac{\eta^2 A^2}{2\delta^2}$$

και χρησιμοποιήστε αυτήν τη σχέση για να δείξετε ότι

$$\frac{\eta}{1-\delta} \mathbb{E}[\mathrm{Reg}_p(T)] \leq F_1 + \frac{\delta\eta}{1-\delta} \sum_{t=1}^{T} \langle v_t, p - \mathrm{unif} \rangle + \frac{\eta^2 A^2}{2\delta^2} T$$

(3) Με βάση τα παραπάνω, υπολογίστε ένα άνω φράγμα – δηλαδή μία *εγγύηση* του αλγορίθμου – για τη μέση μεταμέλεια $\mathbb{E}[\mathrm{Reg}_p(T)]$ του παίκτη ως προς τις ποσότητες A, T, η και δ. Ποιά είναι η καλύτερη εγγύηση που μπορείτε να επιτύχετε αν χρησιμοποιήσετε παραμέτρους της μορφής $\eta \propto 1/T^p$ και $\delta \propto 1/T^q$ για $p, q \geq 0$?

(4) Πώς συγκρίνονται οι εγγυήσεις αυτές με τις αντίστοιχες εγγυήσεις του EXP3? Μπορεί να χρησιμοποιηθεί η παραπάνω ανάλυση για τον EXP3? Αν ναι, πώς? Αν όχι, γιατί?

Καλή επιτυχία!