Ασκήσεις στη Θεωρία Πληροφορίας

Κωστής Ξενούλης

1. Η χωρητικότητα ενός καναλιού είναι καλά ορισμένη γιατί η αμοιβαία πληροφορία

$$I(X;Y) = -D_{\mathrm{KL}}(p_Y || p_{X,Y})$$

είναι κυρτή συνάρτηση της πιθανότητας εισόδου.

Λύση: Η απόσταση Kullback-Leibler μεταξύ δύο συναρτήσεων πιθανοτήτων p, q μιας τυχαίας μεταβλητής X με πεδίο τιμών $\mathcal{X}$ ορίζεται ως

$$D_{\mathrm{KL}}(p_X || q_X) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}.$$

Βάσει του ορισμού της αμοιβαίας πληροφορίας μεταξύ δύο τυχαίων μεταβλητών , με κατανομή $p_{X,Y}(x,y)$

$$I(X;Y) = \sum_{x,y} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}.$$

Συνεπώς ισχύει $I(X;Y) = D_{\mathrm{KL}}(p_{X,Y} || p_X p_Y)$. Επιπλέον, η αμοιβαία πληροφορία είναι κοίλη συνάρτηση της συνάρτησης πιθανότητας εισόδου.

2. Αν $X, Y, Z = X + Y$ διακριτές τυχαίες μεταβλητές, τότε $H(Z|X) = H(Y|X)$.

Λύση: Η συνάρτηση πιθανότητας $p_{Z|X}(z|x)$ ικανοποιεί

$$p_{Z|X}(Z = z|X = x) = p_{Y|X}(Y = z - X|X = x) = p_{Y|X}(Y = z - x|X = x). \tag{1}$$

Αντικαδιστώντας την παραπάνω σχέση στον τύπο της εντροπίας $H(Z|X)$, προκύπτει η ισότητα.

3. Η διαφορική εντροπία ικανοποιεί $h(X) \leq \frac{1}{2}\log(2\pi e\sigma^2)$. Πότε ισχύει η ισότητα?

Λύση: Έστω οι ακόλουθες περιπτώσεις συναρτήσεων πιθανότητας της τυχαίας μεταβλητής X

- $q_X(x)$ με μέση τιμή 0 και διακύμανση σ^2, και
- κανονική κατανομή $\mathcal{N}(0, \sigma^2)$ με συνάρτηση πιθανότητας

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}.$$

Η απόσταση Kullback-Leibler μεταξύ των συναρτήσεων πιθανότητας q και f_X ικανοποιεί

$$\begin{aligned} D_{\mathrm{KL}}(q||f) &= \int q_X(x) \log \frac{q_X(x)}{f_X(x)} dx = -\int q_X(x) \log \frac{1}{q_X(x)} dx + \int q_X(x) \log f_X(x) dx \\ &= h(X) + \log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{1}{2\sigma^2} \int x^2 q(x) dx. \end{aligned}$$

Από την υπόθεση για τη συνάρτηση πιθανότητας q_X ισχύει $\int x^2 q(x)dx = \sigma^2$. Συνεπώς

$$D_{\mathrm{KL}}(q||f) = h(X) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2} = \frac{1}{2}\log(2\pi e\sigma^2).$$

Επειδή η απόσταση Kullback-Leibler είναι μη-αρνητική ποσότητα και $D_{\mathrm{KL}}(p||q) = 0$ αν και μόνο αν $p = q$, ισχύει $h(X) \leq \frac{1}{2}\log(2\pi e\sigma^2)$. Επιπλέον η ισότητα ισχύει αν η X ακολουθεί την κανονική κατανομή $\mathcal{N}(0, \sigma^2)$.

4. Η χωρητικότητα ενός BSC καναλιού δίνεται από τη σχέση $C = 1 - H_2(p)$. Τι είναι το p και το $H_2(p)$;

Λύση: Σε ένα δυαδικό συμμετρικό κανάλι χωρίς μνήμη, p είναι η πιθανότητα το κανάλι να αλλάξει το μεταδιδόμενο bit είτε από 0 σε 1, είτε από 1 σε 0. Επίσης $H_2(p)$ είναι η δυαδική εντροπία, η οποία ορίζεται ως $H_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$.

5. Να δείξετε με τη βοήθεια της ανισότητας επεξεργασίας δεδομένων ότι $I(Y;Z) \geq I(X;Z)$ για την αλυσίδα Markov $X \to Y \to Z$.

Λύση: Αν οι τυχαίες μεταβλητές X, Y, Z ικανοποιούν την αλυσίδα Markov $X \to Y \to Z$, τότε οι X, Y είναι ανεξάρτητες μεταξύ τους δοθέντος της Y

$$p_{Z,Y|X}(Z = z, Y = y|X = x) = p_{Z|X}(Z = z|X = x)p_{Y|X}(Y = y|X = x).$$

Βάσει του κανόνα αλυσίδας για την αμοιβαία πληροφορία $I(X, Y; Z)$

$$\begin{aligned} I(X,Y;Z) &= I(X;Z) + I(Y;Z|X) \\ &= I(Y;Z) + I(X;Z|Y) \end{aligned}$$

Λόγω της αλυσίδας Markov και της παραπάνω σχέσης μεταξύ των συναρτήσεων πιθανότητας, ισχύει $I(X;Z|Y) = 0$. Συνεπώς, επειδή η αμοιβαία πληροφορία είναι μη-αρνητική ποσότητα έχουμε

$$I(X;Z) + I(Y;Z|X) = I(Y;Z) \to I(Y;Z) \geq I(X;Z).$$

6. Για μία στατική στοχαστική διαδικασία $X_1, X_2, \ldots, X_n$ να δείξετε ότι

$$\frac{H(X_1, X_2, \ldots, X_n)}{n} \leq \frac{H(X_1, X_2, \ldots, X_{n-1})}{n-1}.$$

Λύση: Για την απλούστευση της παρουσίασης της λύσης, εξετάζουμε την περίπτωση όπου $n = 4$. Θέλουμε να δείξουμε

$$3H(X_1, X_2, X_3, X_4) \leq 4H(X_1, X_2, X_3).$$

Βάσει του κανόνα αλυσίδας για την εντροπία αρκεί να δείξουμε

$$\begin{aligned} &3H(X_4|X_1, X_2, X_3) + 3H(X_1, X_2, X_3) \leq 4H(X_1, X_2, X_3) \leftrightarrow \\ &3H(X_4|X_1, X_2, X_3) \leq H(X_1, X_2, X_3) \leftrightarrow \\ &3H(X_4|X_1, X_2, X_3) \leq H(X_3|X_1, X_2) + H(X_2|X_1) + H(X_1) \end{aligned}$$

Γνωρίζοντας ότι οι συνθήκες μειώνουν την εντροπία ($H(X|Y) \leq H(X)$), ισχύει

$$\begin{aligned} H(X_4|X_1, X_2, X_3) &\leq H(X_4|X_2, X_3) \\ H(X_4|X_1, X_2, X_3) &\leq H(X_4|X_3) \\ H(X_4|X_1, X_2, X_3) &\leq H(X_4). \end{aligned}$$

Επιπλέον, λόγω της υποθέσης ότι η διαδικασία είναι στατική ισχύει

$$\begin{aligned} H(X_4|X_2, X_3) &= H(X_3|X_2, X_1) \\ H(X_4|X_3) &= H(X_3|X_2) \\ H(X_4) &= H(X_1). \end{aligned}$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις καταλήγουμε στην αρχική ζητούμενη ανισότητα.

7. Έστω κανάλι πολλαπλής πρόσβασης με δύο ανεξάρτητους μεταδότες X_1, X_2 και δέκτη Y. Να δείξετε ότι $I(X_1; X_2|Y) \leq I(X_1; Y|X_2)$.

Λύση: Στόχος του δέκτη είναι να αποκωδικοποιήσει και τις δύο ακολουθίες X_1, X_2 των πομπών βάσει της ακολουθίας Y. Χρησιμοποιώντας τον κανόνα αλυσίδας για την αμοιβαία πληροφορία έχουμε

$$\begin{aligned} I(X_1; X_2|Y) &= I(X_1; X_2, Y) - I(X_1; Y) \\ I(X_1; Y|X_2) &= I(X_1; X_2, Y) - I(X_1; X_2). \end{aligned}$$

Επειδή οι X_1, X_2 είναι μεταξύ τους ανεξάρτητοι έχουμε $I(X_1; X_2) = 0$. Συνεπώς από τις παραπάνω ισότητες

$$I(X_1; X_2|Y) = I(X_1; Y|X_2) - I(X_1; Y).$$

Η ζητούμενη ανισότητα προκύπτει καθώς $I(X_1; Y) \geq 0$.

8. Αν X, Y, Z διακριτές τυχαίες μεταβλητές, να αποδείξετε

$$I(X; Y, Z) \geq I(X, Y; Z) - I(Y; Z|X).$$

Λύση: Βάσει του κανόνα αλυσίδας για την αμοιβαία πληροφορία, το δεξιό μέλος της ανισότητας που θέλουμε να αποδείξουμε ικανοποιεί

$$I(X, Y; Z) - I(Y; Z|X) = I(X; Z) + I(Y; Z|X) - I(Y; Z|X) = I(X; Z).$$

Επίσης

$$I(X; Y, Z) = I(X; Z) + I(Y; Z|X) \geq I(X; Z).$$

Από τις δύο παραπάνω σχέσεις, καταλήγουμε στην ζητούμενη ανισότητα.

9. Έστω οι διακριτές τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ σχηματίζουν Markov αλυσίδα της μορφής $X_1 \to X_2 \to \cdots \to X_n$. Να δείξετε ότι $I(X_1; X_2, \ldots, X_n) = I(X_1; X_2)$.

Λύση: Για λόγους απλούστευσης της λύσης θεωρούμε την περίπτωση όπου $n = 4$. Θέλουμε να δείξουμε $I(X_1; X_2, X_3, X_4) = I(X_1; X_2)$. Βάσει του κανόνα αλυσίδας για την αμοιβαία πληροφορία ισχύει

$$I(X_1; X_2, X_3, X_4) = I(X_1; X_2, X_3) + I(X_1; X_4|X_2, X_3).$$

Λόγω της αλυσίδας Markov, οι τυχαίες μεταβλητές X_1, X_4, είναι ανεξάρτητες μεταξύ τους δοθέντος των X_2 και X_3. Συνεπώς

$$I(X_1; X_4|X_2, X_3) = 0 \to I(X_1; X_2, X_3, X_4) = I(X_1; X_2, X_3).$$

Με παρόμοιο τρόπο $I(X_1; X_2, X_3) = I(X_1; X_2) + I(X_1; X_3|X_2) = I(X_1; X_2)$.

10. Έστω διακριτές τυχαίες μεταβλητές X, Y_1, Y_2 τέτοιες ώστε Y_1, Y_2 είναι στατιστικά ανεξάρτητες δοθέντος της X, $P(Y_1, Y_2|X) = P(Y_1|X)P(Y_2|X)$. Επίσης ισχύει $H(Y_1) = H(Y_2)$ και $H(Y_1|X) = H(Y_2|X)$. Να δείξετε ότι $I(X; Y_1, Y_2) = 2I(X; Y_1) - I(Y_1; Y_2)$.

Λύση: Βάσει των υποθέσεων $H(Y_1) = H(Y_2)$ και $H(Y_1|X) = H(Y_2|X)$ προκύπτει

$$I(X;Y_1) = I(X;Y_2).$$

Επίσης βάσει του κανόνα αλυσίδας για την αμοιβαία πληροφορία ισχύει

$$I(X;Y_1,Y_2) = I(Y_1,Y_2;X) = I(Y_1;X) + I(Y_2;X|Y_1)$$

και

$$I(Y_2;X|Y_1) = I(X,Y_1;Y_2) - I(Y_1;Y_2).$$

Συνδυάζοντας τις δύο παραπάνω ισότητες προκύπτει

$$I(X;Y_1,Y_2) = I(X;Y_1) - I(Y_1;Y_2) + I(X,Y_1;Y_2).$$

Για την ολοκλήρωση της απόδειξης της ζητούμενης σχέσης αρκεί να δείξουμε

$$I(X,Y_1;Y_2) = I(X;Y_1).$$

Χρησιμοποιώντας πάλι τον κανόνα αλυσίδας

$$I(X,Y_1;Y_2) = I(X;Y_2) + I(Y_1;Y_2|X).$$

Επειδή οι Y_1, Y_2 είναι στατιστικά ανεξάρτητες δοθέντος της X, έχουμε $I(Y_1;Y_2|X) = 0$. Τέλος, σύμφωνα με την αρχική σχέση καταλήγουμε στο ζητούμενο.