Θεόδωρος Αλεξόπουλος, Αναπληρωτής Καθηγητής
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ - ΤΟΜΕΑΣ ΦΥΣΙΚΗΣ
ΗΡΩΩΝ ΠΟΛΥΤΕΧΝΕΙΟΥ 9
ΑΘΗΝΑ 157 80
Τηλ: 210 772-3019, Fax: 210 772-3025
e-mail: theoalex@central.ntua.gr

Theodoros Alexopoulos, Associate Professor
NATIONAL TECHNICAL UNIVERSITY
DEPARTMENT OF PHYSICS
ZOGRAFOU CAMPUS
157 80 ATHENS - GREECE
Phone : +30 210 772-3019, Fax: +30 210 772-3025
e-mail: Theodoros.Alexopoulos@cern.ch
html://www.physics.ntua.gr/Faculty/theoalex

Αναγνώριση Προτύπων & Νευρωνικά Δίκτυα

Προβλήματα 3

(Επιστροφή: 26 Ιανουαρίου 2004)

0.1

Ορίζουμε ως θ την πιθανότητα να φέρουμε "κορώνα" όταν ρίχνουμε ένα νόμισμα. Υποθέστε ότι το νόμισμα έχει ριχτεί n φορές. Έστω ότι n_H είναι ο αριθμός που έχουμε φέρει "κορώνα". Να δείξετε ότι η εκτίμηση της μέγιστης πιθανοφάνειας (MLE) της μεταβλητής θ είναι $\hat{\theta} = n_H/n$.

0.2

Έστω ότι μια τυχαία μεταβλητή x υπακούει στην εκθετική πυκνότητα πιθανότητας:

$$p(x/\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0, \\ 0 & \textbf{αλλού}. \end{cases}$$

(α) Να δώσετε ένα γράφημα της $p(x/\theta)$ ως προς x για $\theta = 1$. Επίσης να δώσετε ένα γράφημα της $p(x/\theta)$ ως προς θ για $x = 2$.
(β) Έστω ότι έχουμε n δείγματα $x_1, x_2, \ldots, x_n$ που προέρχονται από την $p(x/\theta)$. Να δείξετε ότι η εκτίμηση της μέγιστης πιθανοφάνειας (MLE) της μεταβλητής θ είναι:

$$\hat{\theta} = \frac{1}{(1/n)\sum_{k=1}^{n} x_k}.$$

Στο γράφημα του ερωτήματος (α) να δείξετε τη MLE $\hat{\theta}$ για μεγάλα n.

0.3

Τα διανύσματα εκπαίδευσης για δυο κλάσεις ω_1, ω_2 είναι:

$$S_1 = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\} \in \omega_1,$$

και

$$S_2 = \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \in \omega_2,$$

Να ταξινομήσετε το διάνυσμα $\vec{x} = (1,5 \quad 0)^t$ σύμφωνα με τον ταξινομητή $k = 3$ πλησιεστέρων γειτόνων.

0.4

Να δώσετε τη πυκνότητα πιθανότητας $p(x)$ μιας διάστασης με τη μέθοδο της παραθυροποίησης του Parzen χρησιμοποιώντας τα ακόλουθα δείγματα:

$$S = \{2,5 \quad 2,8 \quad 3,4 \quad 4,2 \quad 4,5 \quad 4,7 \quad 5,2 \quad 5,6 \quad 7,5\}\,.$$

Να θεωρήσετε το εύρος της συνάρτησης παραθύρου $h_n = 1$.

0.5

Έστω ότι οι συναρτήσεις πιθανοφάνειας για δυο κλάσεις ω_1, ω_2, είναι$p(\vec{x}/\omega_i)$, με μέσες τιμές και πίνακες διασποράς $\vec{\mu}_i$, $\overline{\overline{\Sigma}}_i$ αντίστοιχα (το i είναι $1,\ 2$). Δεν υποθέτουμε γκαουσιανές συναρτήσεις στο πρόβλημα αυτό.
Θεωρούμε μια μονοδιάστατη μεταβλητή $y = \vec{w}^t\vec{x}$. Επομένως μπορούμε να θεωρήσουμε νέες συναρτήσεις πιθανοφάνειας για τις δυο κλάσεις ω_1, ω_2, που είναι $p(y/\omega_i)$.
Να δείξετε ότι η συνάρτηση κόστους:

$$J(\vec{w}) = \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2},$$

γίνεται μέγιστη για:

$$\vec{w} = (\overline{\overline{\Sigma}}_1 + \overline{\overline{\Sigma}}_2)^{-1}(\vec{\mu}_1 - \vec{\mu}_2),$$

όπου m_i, σ_i^2 είναι οι μέσες τιμές και οι διασπορές των συναρτήσεων $p(y/\omega_i)$.

0.6

Έστω μια τυχαία ομοιόμορφη μεταβλητή X **που ορίζεται στο διάστημα** $[0, \theta]$ **με συνάρτηση πιθανοφάνειας:**

$$p(x/\theta) = \begin{cases} 1/\theta & 0 \leq x \leq \theta, \\ 0 & \textbf{αλλού}. \end{cases}$$

Έστω ότι λαμβάνουμε ένα σύνολο n **δειγμάτων** $D = \{x_1, x_2, \ldots, x_n\}$ **από την πυκνότητα πιθανότητας** $p(x/\theta)$. **Να δείξετε ότι η εκτίμηση της μέγιστης πιθανοφάνειας (MLE) της μεταβλητής** θ **είναι:**

$$\hat{\theta} = \max_k x_k.$$

0.7

Έστω ότι $p(x) \sim N(\mu, \sigma^2)$ **και θεωρούμε ως συνάρτηση παραθύρου την** $\phi(x) \sim N(0, 1)$.
Να δείξετε ότι η εκτίμηση Parzen της συνάρτησης $p(x)$:

$$\hat{p}(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \phi\left(\frac{x - x_i}{h_n}\right),$$

έχει την ακόλουθη μέση τιμή:

$$E\left[\hat{p}_n(x)\right] \sim N(\mu, \sigma^2 + h_n^2).$$

0.8

Έστω ότι $p(x)$ **είναι μια ομοιόμορφη κατανομή στο διάστημα** $[0, a]$, **και έστω ότι η συνάρτηση παραθύρου έχει τη μορφή:**

$$\phi(x) = \begin{cases} e^{-x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Να δείξετε ότι η μέση τιμή της εκτίμησης Parzen για τη συνάρτηση $p(x)$ **δίνεται από τη σχέση:**

$$E\left[\hat{p}_n(x)\right] = \begin{cases} 0 & x < 0, \\ (1/a)\left(1 - e^{-x/h_n}\right) & 0 \leq x \leq a, \\ (1/a)\left(e^{a/h_n} - 1\right)e^{-x/h_n} & x > a. \end{cases}$$

0.9

Έστω ότι $\hat{p}_n(\vec{x})$ είναι η εκτίμηση Parzen για μια άγνωστη πυκνότητα πιθανότητας $p(\vec{x})$ και έστω $\phi(\vec{x})$ είναι μια συνάρτηση παραθύρου από την οποία σχηματίζεται η $\hat{p}_n(x)$ με μέση τιμή μηδέν:

$$\int \vec{x}\phi(\vec{x}) = 0.$$

Να βρείτε μια απλή έκφραση της μέσης τιμής της εκτίμησης Parzen:

$$\hat{\vec{\mu}} = \int \vec{x}\hat{p}_n(\vec{x})d\vec{x}.$$

0.10

Έστω ότι $\overrightarrow{x}$ είναι ένα δυαδικό διάνυσμα, και $p(\overrightarrow{x}/\theta)$ είναι ένα μείγμα C πολυδιάστατων κατανομών Bernoulli:

$$p(\overrightarrow{x}/\theta) = \sum_{i=1}^{C} p(\overrightarrow{x}/\omega_i, \overrightarrow{\theta}_i)P(\omega_i),$$

όπου

$$p(\overrightarrow{x}/\omega_i, \overrightarrow{\theta}_i) = \prod_{j=1}^{d} \theta_{ij}^{x_j}(1-\theta_{ij})^{1-x_j}.$$

(α) Να δείξετε ότι:

$$\frac{\partial \ln p(\overrightarrow{x}/\omega_i, \theta_i)}{\partial \theta_{ij}} = \frac{x_j - \theta_{ij}}{\theta_{ij}(1-\theta_{ij})}.$$

(β) Με τη βοήθεια της γενικής εξίσωσης της εκτίμησης της μέγιστης πιθανοφάνειας, να δείξετε ότι η MLE της θ_i ($\hat{\overrightarrow{\theta_i}}$) θα είναι:

$$\hat{\overrightarrow{\theta_i}} = \frac{\sum_{k=1}^{n} P(\omega_i/\overrightarrow{x}_k, \hat{\overrightarrow{\theta_i}})\overrightarrow{x}_k}{\sum_{k=1}^{n} P(\omega_i/\overrightarrow{x}_k, \hat{\overrightarrow{\theta_i}})}.$$

0.11

Οι μέσες τιμές δύο κλάσεων ω_1, ω_2 **είναι** $\overrightarrow{m}_1 = (-2\ \ -2)^t$ **και** $\overrightarrow{m}_2 = (2\ 2)^t$ **και οι πίνακες διασποράς είναι:**

$$\overline{\overline{\Sigma}}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \overline{\overline{\Sigma}}_2 = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \quad a > 0 \in \mathbb{R}.$$

(α) Να βρείτε την κλίση της συνάρτησης διάκρισης του Fisher από τη σχέση:

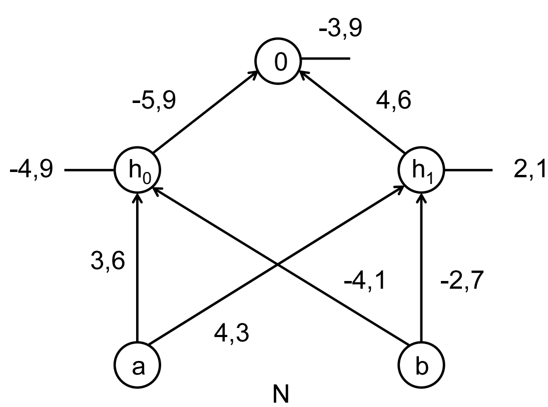
$$\hat{\overrightarrow{w}} = \overline{\overline{S}}_w^{-1}(\overrightarrow{m}_1 - \overrightarrow{m}_2).$$

(β) Να βρείτε την κλίσης της συνάρτησης διάκρισης του Fisher υπολογίζοντας τις ιδιοτιμές του πίνακα:

$$\overline{\overline{S}}_w^{-1}\overline{\overline{S}}_B.$$

0.12

Το δίκτυο N **(που φαίνεται στο σχήμα 1) εκπαιδεύτηκε χρησιμοποιώντας τον αλγόριθμο οπισθοδρομικής διάδοσης,** $(a, b) \in \{-1, 1\}$ **και η συνάρτηση μεταφοράς είναι** $\tanh(x)$**.**



Σχήμα 1:

(α) Να ελέγξετε αν το δίκτυο N **υπολογίζει τη συνάρτηση XOR.**
(β) Περιγράψτε πώς γίνεται η εκμάθηση του δικτύου σύμφωνα με τον αλγόριθμο οπισθοδρομικής διάδοσης, υπολογίζοντας την πρώτη αλλαγή του βάρους Δw**, για την κρυφή μονάδα** h_1**. Για την**

εκπαίδευση θεωρείστε $p = (-1, 1),\ t = 1,\ n = 0, 1$, καθώς και τα βάρη και τις μεροληψίες που φαίνονται στο σχήμα 1.

0.13

Το δίκτυο N έχει συμμετρικά βάρη $(w_{ab} = w_{ba})$ και δεν έχει μεροληψία $(w_{0a} = 0)$, και έχει 3 νευρώνες $(N = \{1, 2, 3\})$. Ο στόχος μας είναι να αποθηκεύσουμε το πρότυπο: $\xi = (-1, 1, 1)$. Να θέσετε τα βάρη ώστε να επιτευχθεί ο στόχος μας ακολουθώντας τον κανόνα του Hebb $w_{ab} = (1/|N|) \sum_i \xi_a^i \xi_b^i$, να σχεδιάσετε το δίκτυο και να τοποθετήσετε στο σχήμα τα αντίστοιχα βάρη. Να αποδείξετε ότι το ξ είναι ένα σταθερό σημείο έλξης (σημείο ισορροπίας) για το δίκτυο αυτό. Σταθερές καταστάσεις είναι τυπικά minima της ενέργειας:

$$E(w, s) = -\frac{1}{2} \sum_{a,b} w_{ab} s_a s_b - \sum_a w_{0a} s_a.$$

Υπόδειξη: Συγκρίνετε την ενέργεια για το σημείο ξ με την ενέργεια στις γειτονικές καταστάσεις.

0.14

Έστω μια τυχαία ομοιόμορφη μεταβλητή X που ορίζεται στο διάστημα $[\theta, 1]$ με συνάρτηση πιθανοφάνειας:

$$p(x/\theta) = \begin{cases} 1/(1-\theta) & \theta \leq x \leq 1, \\ 0 & \text{αλλού}. \end{cases}$$

Έστω ότι λαμβάνουμε ένα σύνολο n δειγμάτων $D = \{x_1, x_2, \ldots, x_n\}$ από την πυκνότητα πιθανότητας $p(x/\theta)$. Να βρείτε την εκτίμηση της μέγιστης πιθανοφάνειας (MLE) της μεταβλητής θ.
