Υπολογιστική Στατιστική

Κατερίνα Ορφανογιαννάκη

Τμήμα Μαθηματικών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
korfanog@math.uoa.gr

2020-2021

Εισαγωγή στον EM και άλλους αριθμητικούς αλγορίθομους

1. Αριθμητικοί αλγόριθμοι
2. *Newton-Raphson* και άλλα σχήματα
3. Ανασκόπηση της θεωρίας της Μέγιστης Πιθανοφάνειας
4. Ο αλγόριθμος EM
5. Εφαρμογές

Η ανάγκη για αριθμητικές μεθόδους

Κάτω από συγκεκριμένες συνθήκες δεν υπάρχουν υπάρχουν κλειστού τύπου εξισώσεις (για την εξεύρεση εκτιμητών/ριών ΜΠ). Χρειάζεται λοιπόν να λύσουμε ένα σύστημα πολλών μη-γραμμικών εξισώσεων ή ισοδύναμα να μεγιστοποιήσουμε τη συνάρτηση πιαθανοφάνειας με αριθμητικές μεθόδους. Υπάρχουν διάφορες διαδικασίες και για τους 2 τρόπους, οι οποίοι σε πολλές περιπτώσεις είναι πολύ παρόμοιοι. Επιπλέον υπάρχουν επι τούτου διαδικασίες για εύρεση λύσης.

Η μέθοδος *Newton-Raphson* (σε μία διάσταση)

Έστω ότι θέλουμε να λύσουμε την εξίσωση $F(\theta) = 0$. Ξεκίνα από μία αρχική τιμή $\theta^{(0)}$ και ενημέρωσε με την καινούρια τιμή:

$$\theta^{(r+1)} = \theta^{(r)} - \frac{F(\theta^{(r)})}{F'(\theta^{(r)})}$$

μέχρι η μεταβολή ανάμεσα σε 2 διαδοχικά βήματα να είναι ασήμαντη.

Πλεονεκτήματα και Μειονεκτήματα

Πλεονεκτήματα

- Εύκολο υπολογιστικά
- Τετραγωνική σύγκλιση

Μειονεκτήματα

- Μπορεί να προκύψει λύση έξω από το αποδεκτό εύρος
- Μπορεί να συγκλίνει σε τοπικό και όχι ολικό μέγιστο
- Εξαρτάται από καλή επιλογή της αρχικής τιμής $\theta^{(0)}$.

Εκτίμηση με τη μέθοδο της Μέγιστης πιθανοφάνειας

Έστω ένα μοντέλο $f(x \mid \theta)$, όπου x είναι ένα διάνυσμα με παρατηρήσεις και θ το διάνυσμα με τις παραμέτρους που θέλουμε να εκτιμήσουμε. Παρατηρούμε ένα τυχαίο δείγμα μεγέθους n το οποίο το συμβολίζουμε ως: $(x_1, x_2, \ldots, x_n)$.

Συνάρτηση Πιθανοφάνειας $L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$

Λογαριθμοποιημένη συνάρτηση πιθανοφάνειας:

$$\ell(\theta) = log(L(\theta)) = \sum_{i=1}^{n} log f(x_i \mid \theta)$$

Θεωρία

Έστω p παράμετροι $\theta_1, \ldots, \theta_p$.

Διανύσματα *Score*: $U(\theta) = \left(\frac{\partial \ell(\theta)}{\partial \theta_1}, \ldots, \frac{\partial \ell(\theta)}{\partial \theta_p}\right)'$

Η εκτίμηση ΜΠ $\hat{\theta}$ προϋποθέτει ότι $U(\hat{\theta}) = 0$. Ασυμπτωτικά (και κάτω από κάποιες συνθήκες κανονικοποίησης) ισχύει:

$$U(\theta) \sim N_p(0, J(\theta))$$

όπου $J(\theta)$ είναι ένας πίνακας του οποίου τα στοιχεία J_{ij} υπολογίζονται ως:

$$J_{ij}(\theta) = -E\left[\frac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j}\right]$$

$J(\hat{\theta})$ είναι ο αναμενόμενος πίνακας πληροφορίας του *Fisher*. Επίσης ισχύει ότι:

$$\hat{\theta} \sim N_p(\theta, J^{-1}(\theta))$$

Θεωρία

Συνήθως ο αναμενόμενος πίνακας πληροφορίας του *Fisher* προσεγγίζεται από τον παρατηρούμενο πίνακα πληροφορίας του *Fisher* ο οποίος για το $\hat{\theta}$ συμβολίζεται ως $I(\hat{\theta})$ με στοιχεία $I_{ij}(\theta)$:

$$I_{ij}(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial\theta_i \partial\theta_j}$$

ο οποίος αξιολογείται για $\hat{\theta}$, και άρα κάνουμε χρήση ότι:

$$\hat{\theta} \sim N_p(\theta, I^{-1}(\theta))$$

Επίσης αποδεικνύεται ότι:

$$E\left[-\frac{\partial^2 \ell(\theta)}{\partial\theta_i \partial\theta_j}\right] = E\left[\frac{\partial \ell(\theta)}{\partial\theta_i}\frac{\partial \ell(\theta)}{\partial\theta_j}\right]$$

Αυτό μπορεί να απλοποιήσει σημαντικά τους υπολογισμούς.

Παράδειγμα: Περικομένη κατανομή *Poisson*

Η Περικομένη κατανομή *Poisson* προκείπτει από την κατανομή *Poisson* όταν η τιμή μηδέν δεν μπορεί να παρατηρηθεί. Είναι χρήσιμη σε συγκεκριμένες εφαρμογές όπως: επιδημίες, βιβλιομετρία, ζωολογία κλπ. Τρυνςατεδ Ποισσον διστριβυτιον αρισες φρομ τηε Ποισσον διστριβυτιον ωηεν τηε ζερο-τη αλυε ςαννοτ βε οβσερεδ. Υσεφυλ φορ ςερταιν αππλιςατιονς ιν επιδεμιςς, βιβλιομετρψ, ζοολογψ ετς. Η συνάρτηση πιθανότητας δίνεται από τη σχέση:

$$P(X=x)=\frac{exp(-\lambda)\lambda^x}{x!(1-exp(-\lambda))}=\frac{\lambda^x}{x!(e^\lambda-1)},\quad \lambda>0, x=1,2,\ldots$$

Δοθέντος ενός δείγματος $(x_1, x_2, \ldots, x_n)$ η λογαριθμοπιθανοφάνεια είναι:

$$\ell(\lambda)=\log\lambda\sum_{i=1}^n x_i - n\log\left(e^\lambda-1\right)-\sum_{i=1}^n \log(x_i!)$$

Παίρνοντας τη μερική παράγωγο ως προς λ έχουμε ότι:

$$\frac{d\ell(\lambda)}{d\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - \frac{ne^{\lambda}}{e^{\lambda} - 1} = 0$$

η οποία είναι ξεκάθαρα μη γραμμική και πρέπει να λυθεί αριθμητικά. Η εξίσωση που πρέπει να λύσουμε γράφεται:

$$F(\lambda) = \frac{\lambda}{\bar{x}} - 1 + e^{-\lambda} = 0$$

Η *Newton-Raphson* επανάληψη είναι: (όπου με λ συμβολίζουμε την τρέχουσα τιμή)

$$\lambda^{(new)} = \lambda - \frac{\frac{\lambda}{\bar{x}} - 1 + e^{-\lambda}}{\frac{1}{\bar{x}} - e^{-\lambda}}$$