Μάθημα ΕΦΑΡΜΟΣΜΕΝΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗ
Καθ. Γ. Κοντογιάννης
Τρίτη 10 Δεκεμβρίου 2013

Σετ ασκήσεων #4

ΕΝΤΟΛΕΣ: Για τις ασκήσεις αυτού του σετ, για τη δημιουργία τυχαίων αριθμών να χρησιμοποιηθεί **μόνο** η απλή εντολή `rand()`.

1. Έστω $x=(b_1,b_2,\ldots,b_m)$ μια ακολουθία από *m* bits (δηλαδή το κάθε b_i είναι 0 ή 1) και έστω $k(x)$ το πλήθος από «1» που περιέχει η *x*. Κάθε τέτοια ακολουθία έχει πιθανότητα ανάλογη προς το $1+k(x)^5$, δηλαδή η πιθανότητα της *x* εξαρτάται μόνο από το πόσα «1» περιέχει, και η κατανομή της $\pi(x)$ ορίζεται από τον τύπο

$$\pi(x)=\frac{1}{Z}\left[1+k(x)^5\right],$$

όπου,

$$Z=\sum_{y\in\{0,1\}^m}[1+k(y)^5].$$

Για *m* = 30 bits, θέλουμε να ξέρουμε πόσα «1» κατά μέσο όρο θα περιέχει μία τυχαία τέτοια ακολουθία.

a. Χρησιμοποιώντας το `Scilab`, γράψτε ένα πρόγραμμα που να προσομοιώνει μια Α.Μ. βασισμένη στον αλγόριθμο Metropolis και που να έχει στάσιμη κατανομή την $\pi(x)$, ως εξής.

   i. Επιλέγουμε το αρχικό σημείο $X_0=(1,0,1,0\ldots,1,0)$.

   ii. Αν το τωρινό σημείο του Metropolis είναι το $X_n=x=(b_1,b_2,\ldots,b_m)$, για το προτεινόμενο επόμενο σημείο $y=(b'_1,b'_2,\ldots,b'_m)$ επιλέγουμε μια τυχαία θέση i από το 1 ως το *m*, και με πιθανότητα ½ θέτουμε $b'_i=1-b_i$, αλλιώς θέτουμε $b'_i=b_i$. Για τα υπόλοιπα bits, θέτουμε $b'_j=b_j\ (j\neq i)$.

   iii. Θέτουμε το X_{n+1} να είναι ή το x ή το y, ανάλογα με τη διαδικασία επιλογής του αλγορίθμου Metropolis. Ποια είναι η τιμή του *α* εδώ, δηλαδή της πιθανότητας μετάβασης στην υποψήφια κατάσταση; [**ΠΡΟΣΟΧΗ**: Οι καταστάσεις που αντιστοιχούν στα *i* και *j* των διαλέξεων, εδώ είναι δυαδικές ακολουθίες *x* μήκους *m*. **ΥΠΟΔΕΙΞΗ**: Η αντίστοιχη κατανομή της *π* των διαλέξεων την οποία θέλουμε να προσομοιώσουμε εδώ, είναι μια (από κοινού) κατανομή $\pi(i)$ όπου τα *i* είναι όλες οι δυαδικές ακολουθίες μήκους *m*.]

   Τρέξτε το πρόγραμμά σας για 1000 βήματα και δώστε 3 διαφορετικά αποτελέσματα για την *x*.

   Χρησιμοποιώντας την εντολή `plot(x,'*')` κάντε ένα γράφημα του αποτελέσματός σας μετά από κάθε επανάληψη των πιο πάνω βημάτων. [**ΥΠΟΔΕΙΞΗ**: Για να μη γίνονται τα διαδοχικά βήματα τόσο γρήγορα που να είναι αδύνατο να τα παρακολουθήσετε, μπορείτε να χρησιμοποιήσετε τις εντολές

`f1=scf(1);, clf, plot(X,'*'), xpause(50000)` σε κάθε βήμα του `forloop`.]

b. Χρησιμοποιώντας το πιο πάνω πρόγραμμα βρείτε μια εκτίμηση για το μέσο πλήθος από «1» που περιέχει η *x*. Πόσα βήματα του Metropolis απαιτούνται για να συγκλίνει η εκτίμησή σας με ακρίβεια $\pm 0{,}5$;

2. Έστω ότι $x = (x_u)$ είναι μια εικόνα με την κατανομή $\pi_\beta(x)$ του μοντέλου Ising, και με μέγεθος 25x25 pixels. Δηλαδή, στη θέση του pixel $u = (i, j)$ έχουμε την τιμή x_u η οποία είναι +1 (άσπρο) ή -1 (μαύρο). Επιλέγουμε την τιμή της παραμέτρου β να ισούται με 0,5. Θυμηθείτε πως η (δεσμευμένη) κατανομή των τιμών κάποιου pixel $u = (i, j)$ δεδομένων των τιμών στα υπόλοιπα pixels της εικόνας είναι απλά:

$$\pi_\beta(x_{(i,j)} = +1 \mid x_u : u \neq (i, j)) = \left[1 + e^{-2\beta \sum_{u \sim (i,j)} x_u} \right]^{-1}.$$

a. Χρησιμοποιήστε τον συστηματικό δειγματολήπτη Gibbs για να παράγετε ένα δείγμα εικόνας x με κατανομή $\pi_\beta(x)$. Ο συστηματικός δειγματολήπτης διαφέρει από αυτόν που είδαμε στο μάθημα στο ότι τα pixels που αλλάζουν τιμή δεν επιλέγονται τυχαία, αλλά σαρώνονται όλα από αριστερά προς δεξιά και από την πρώτη γραμμή μέχρι την τελευταία. Ξεκινήστε από μια τυχαία εικόνα, και κάντε ένα γράφημα του αποτελέσματός σας μετά από κάθε πλήρη σάρωση. Πόσα βήματα θεωρείτε πως απαιτούνται για να συγκλίνει ο αλγόριθμος;

b. Επαναλάβετε το πιο πάνω βήμα με β = 10. Σχολιάστε τη διαφορά στα αποτελέσματά σας.

[**ΥΠΟΔΕΙΞΗ**: Για ευκολία , μπορείτε να επιλέξτε μόνο θέσεις $u = (i, j)$ που έχουν ακριβώς 4 γείτονες, δηλαδή που δεν βρίσκονται σε κάποιο από τα άκρα της εικόνας. Για να αναπαραστήσετε μια εικόνα `x`, χρησιμοποιήστε την εντολή `Matplot(10.*x)`.]