Πανεπιστήμιο Αιγαίου

Χωρική Ανάλυση

Ενότητα 4: Τοπική ένταση χωρικού σημειακού προτύπου

Κυριακίδης Φαίδων

Τμήμα Γεωγραφίας

Χρηματοδότηση

Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.

Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.

Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Τοπική Ένταση Χωρικού Σημειακού Προτύπου

Φαίδων Κυριακίδης

Καθηγητής

phkyriakidis@geo.aegean.gr

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
Λόφος Πανεπιστημίου, 81100 Μυτιλήνη

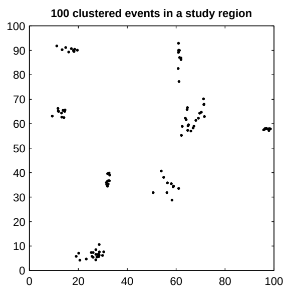
Χωρική Ανάλυση

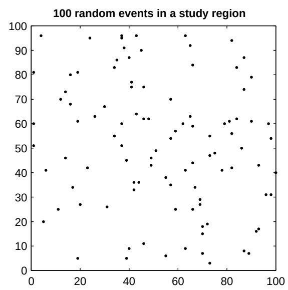
ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΓΕΩΓΡΑΦΙΑ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΗ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ

Χωρικό Σημειακό Πρότυπο

Σύνολο σημειακών θέσεων με καταγεγραμμένα συμβάντα, π.χ., θέσεις δέντρων, θέσεις εκδήλωσης νόσου ή εγκληματικής ενέργειας





Για το μάθημα αυτό:

- βασική παραδοχή: όλες οι θέσεις συμβάντων έχουν καταγραφεί
- αντικείμενο ανάλυσης: θέση συμβάντων κι όχι άλλες ιδιότητες που έχουν τυχόν καταμετρηθεί στις θέσεις εκδήλωσης

Στόχος μαθήματος

Επισκόπηση εργαλείων της περιγραφικής στατιστικής για την ποσοτικοποίηση της χωρικής κατανομής της έντασης (αριθμού ανά επιφάνεια) συμβάντων

Γενικά για την Ανάλυση Έντασης Συμβάντων

Στόχοι της ανάλυσης

- μελέτη της χωρικής κατανομής των συμβάντων και προσδιορισμός πιθανότητας εμφάνισης συμβάντων σε οποιοδήποτε σημείο στο χώρο
- διερεύνηση πιθανών συσχετίσεων μεταξύ έντασης ή πιθανότητας εμφάνισης συμβάντων και άλλων χωρικών μεταβλητών, όπως απόσταση από μια πηγή μόλυνσης ή συγκέντρωση ρύπων

Σημειώσεις

- τα παραπάνω έρχονται σε αντίθεση με την περίπτωση χωρικά συνεχών μεταβλητών, όπου ο σκοπός της ανάλυσης συχνά έγγειται στον χαρακτηρισμό της χωρικής μεταβολής χαρακτηριστικών που μετρήθηκαν σε συγκεκριμένες θέσεις παρατήρησης
- η ανάλυση χωρικής έντασης δεν υπεισέρχεται στη μελέτη της σχετικής θέσης συμβάντων, δηλαδή στην ανάλυση της χωρικής ελκτικότητας ή απώθησης μεταξύ συμβάντων

Βασικές Έννοιες

Σημειακά συμβάντα

Σύνολο N θέσεων συμβάντων που καταγράφηκαν σε μια περιοχή μελέτης:

$$\{\mathbf{u}_i, i = 1, \ldots, N\}, \quad \mathbf{u}_i \in D$$

$\mathbf{u}_i$ = διάνυσμα συντεταγμένων της θέσης του συμβάντος i, π.χ., σε δύο διαστάσεις $\mathbf{u}_i = (x_i \ y_i)$, $\in$ = ανήκει, D = περιοχή μελέτης

Γενικευμένη ένταση λ

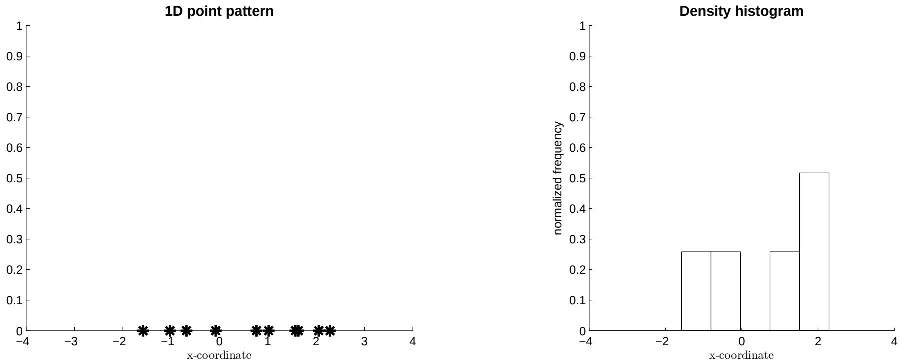
Μέσος (αναμενόμενος) αριθμός συμβάντων στο μοναδιαίο εμβαδό, που εκτιμάται ως: $\hat{\lambda} = N/|D|$, όπου $|D|$ = είναι το μέτρο (εμβαδόν) της υπο μελέτης περιοχής D. Αν λ είναι η ένταση στο μοναδιαίο εμβαδόν, τότε για μια περιοχή s με εμβαδό $|s|$ ισχύει: $\lambda(s) = |s|\lambda$

Τοπική ένταση $\lambda(s)$

Αριθμός συμβάντων $n(s)$ σε μια οποιοδήποτε υποδιαίρεση του χώρου, π.χ., περιοχή, ζώνη, ή φατνίο, s, δια το εμβαδόν $|s|$ του υπο-χώρου αυτού: $\lambda(s) = n(s)/|s|$

Εκτίμηση Έντασης Συμβάντων σε Μία Διάσταση

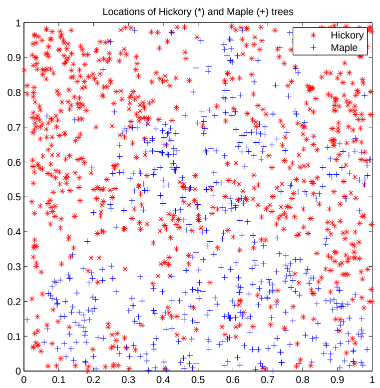
Έστω ένα υποθετικό μονοδιάστατο χωρικό πρότυπο, το οποίο αποτελείται από $N = 10$ συμβάντα (αριστερά) και μια εκτίμηση της τοπικής πυκνότητάς τους, δηλαδή μια μονοδιάστατη κατανομή της σχετικής συχνότητας συμβάντων στο μοναδιαίο διάστημα

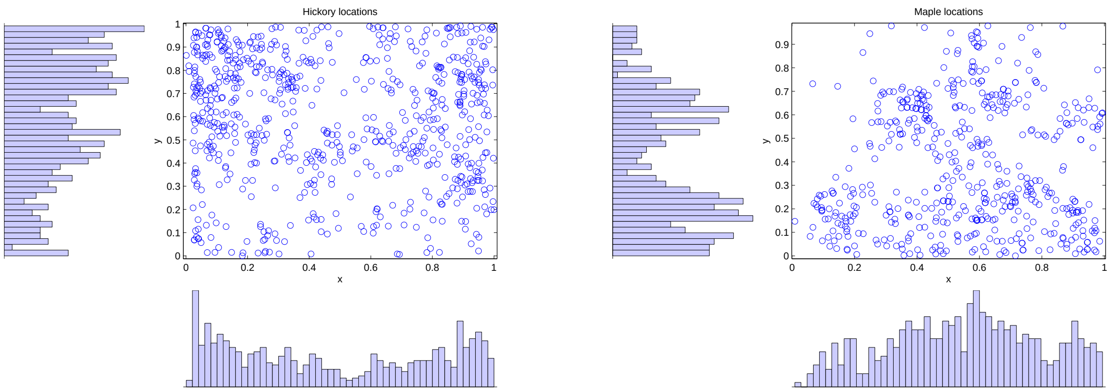


Στατιστική αναλογία

Ο στόχος εδώ είναι η περιγραφή της έντασης των x-συντεταγμένων. Μια αρχική εκτίμηση της κατανομής της έντασης δίνεται από το ιστόγραμμα πυκνότητας (δεξιά). Με άλλα λόγια: οι N συντεταγμένες x των N σημείων ενός μονοδιάστατου χωρικού προτύπου μπορούν να θεωρηθούν και ως N τιμές μιας μεταβλητής, π.χ., θερμοκρασίας...

Ιστογράμματα Συχνότητας Επιμέρους Συντεταγμένων





Επιλογή Εύρους και Αριθμού Διαστημάτων

Εύρος διαστημάτων

- $h = 3.5\, s_X / N^{1/3}$, όπου s_X είναι η τυπική απόκλιση των N μετρήσεων (κανόνας του Scott)
- $h = 2\, IQR_X / N^{1/3}$, όπου $IQR_X = |x_{0.75} - x_{0.25}|$ είναι το ενδο-τεταρτημοριακό εύρος των N μετρήσεων (κανόνας των Freedman-Diaconis)

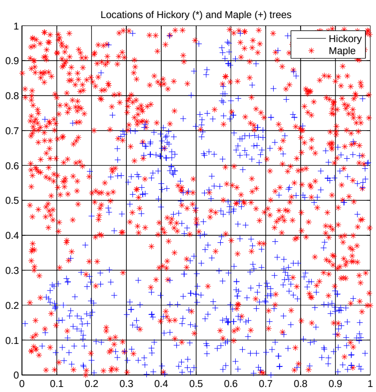
Αριθμός διαστημάτων

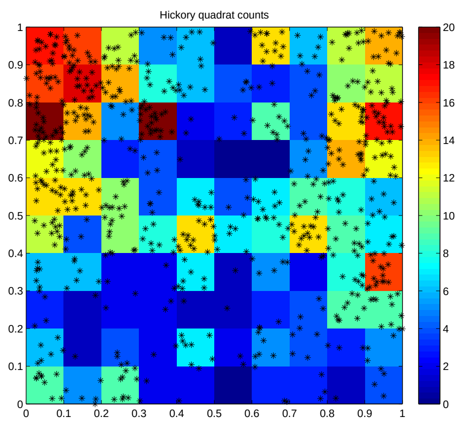
- $K = |x_{max} - x_{min}|/h$, και στρογγυλοποίηση στον πλησιέστερο μεγαλύτερο ακέραιο
- $K = \log_2(N) + 1$, και στρογγυλοποίηση στον πλησιέστερο μεγαλύτερο ακέραιο (φόρμουλα του Sturges)
- $K = \sqrt{N}$, με πολύ διαδεδομένη χρήση (Excel)

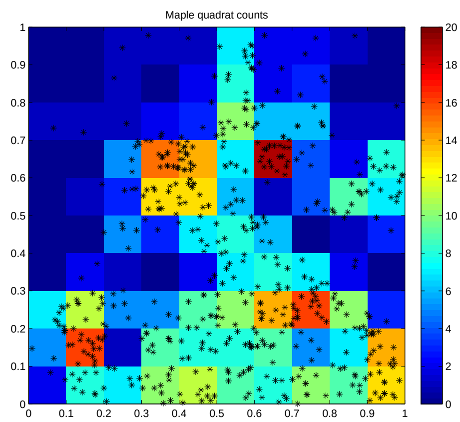
Σημειώσεις

Οι παραπάνω κανόνες βασίζονται σε παραδοχές σχετικά με την υποκείμενη συνεχή κατανομή που θέλουμε να προσεγγίσουμε μέσω του ιστογράμματος. Π.χ., πολλές φορές θεωρείται ότι η άγνωστη κατανομή έχει μία μόνο επικρατούσα τιμή και είναι συμμετρική. Με λίγα λόγια, συνίσταται η κατασκευή ιστογράμματος με πολλαπλούς συνδιασμούς εύρους και αριθμού διαστημάτων...

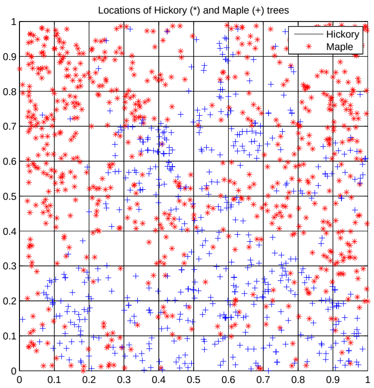
Καταμέτρηση Συμβάντων σε Φατνία

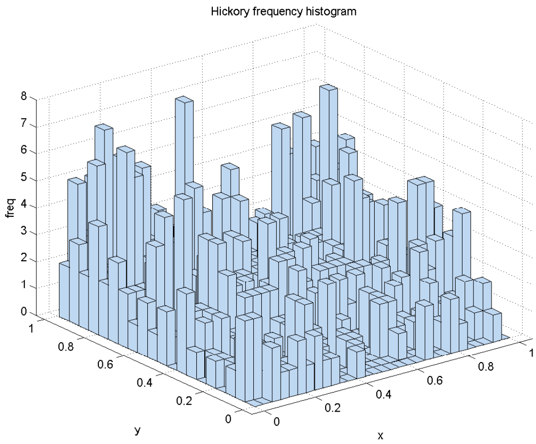


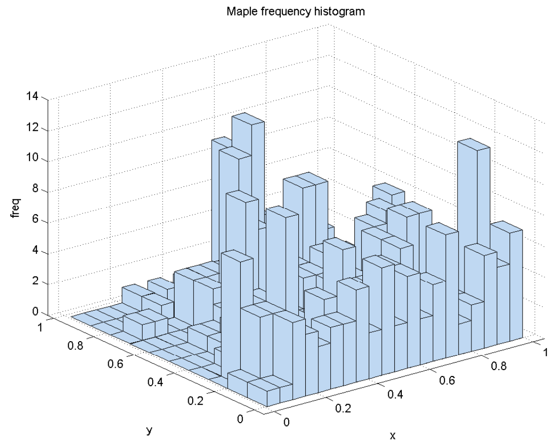




Διδιάστατο Ιστόγραμμα Συχνότητας σε 2.5Δ Προβολή



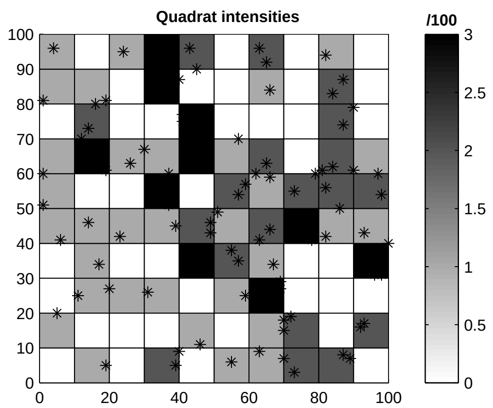




Τοπική Εκτίμηση Έντασης σε Φατνία I

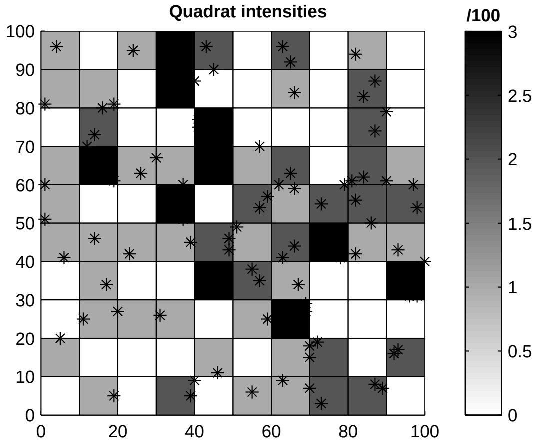
Διαδικασία

1. διαχωρισμός της περιοχής μελέτης D σε $L = K_x K_y$ υπο-περιοχές ή φατνία $\{s_l, l = 1, \dots, L\}$ με *ίσο* μέτρο $|s_l|$. Αυτό εμπεριέχει τη διακριτοποίηση του εύρους των τιμών των συντεταγμένων x_i και y_i των N συμβάντων, όπως και στη μία διάσταση
2. καταμέτρηση του αριθμού (συχνότητας) συμβάντων $n_l = n(s_l)$ σε κάθε φατνίο s_l
3. μετατροπή του αριθμού σε ένταση ως: $\hat{\lambda}_l = \hat{\lambda}(s_l) = n_l/|s_l|$



Διδιάστατο ή διμεταβλητό ιστόγραμμα έντασης, όπου τα φατνία παίζουν το ρόλο παραλληλογράμμων (διαστήματα σε 2 διαστάσεις) και οι ράβδοι είναι παραλληλεπίπεδα

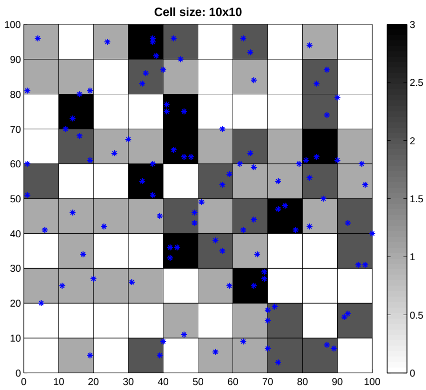
Τοπική Εκτίμηση Χωρικής Έντασης σε Φατνία II

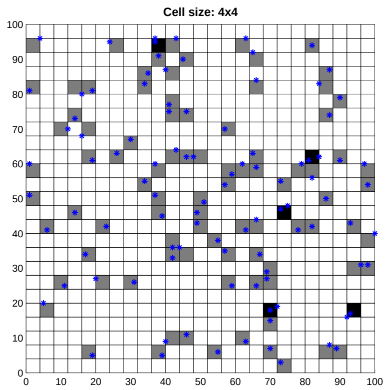


Χαρακτηριστικά

- εκτίμηση έντασης $\hat{\lambda}(s_l)$, αριθμού ανά επιφάνεια, συμβάντων σε ένα σύνολο φατνίων
- αποσκοπεί στην απεικόνιση ευρείας κλίμακας προτύπων της χωρικής κατανομής της έντασης στην περιοχή μελέτης D
- είναι επίσης δυνατή η χρήση τυχαία τοποθετημένων φατνίων

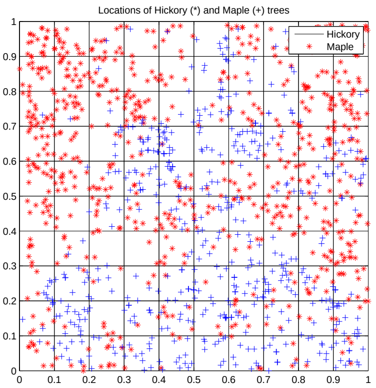
Τοπική Εκτίμηση Χωρικής Έντασης σε Φατνία III

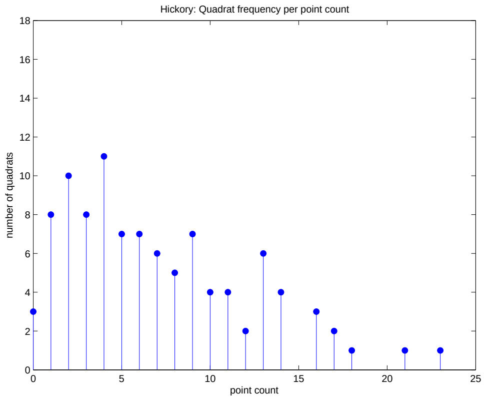


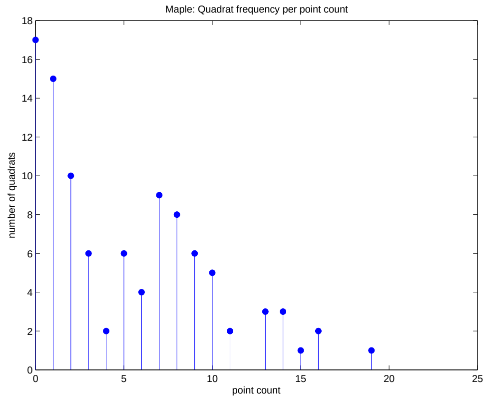


- το μέγεθος και το σχήμα των φατνίων, καθώς επίσης και η αρχή της καννάβου επηρεάζουν τη μορφή της χωρικής κατανομής της εκτιμόμενης έντασης. Π.χ., η χρήση μεγαλύτερων φατνίων οδηγεί σε πιο ομαλή χωρική κατανομή έντασης, ενώ η χρήση μικρότερων φατνίων οδηγεί σε πιο τραχειά κατανομή με πολλές "ψηλόλιγνες" ράβδους και άδεια φατνία
- συχνά, ο στόχος είναι η εκτίμηση μιας επιφάνειας έντασης, η οποία μπορεί (θεωρητικά) να προκύψει όταν το μέγεθος των φατνίων τείνει στο 0. Για να αποφευχθεί το πρόβλημα των μικρών φατνίων με μηδενική ένταση, χρησιμοποιούνται επικαλυπτόμενα φατνία με πολύ κοντινά κέντρα

Ιστόγραμμα Συχνότητας Φατνίων ανά Αριθμό Συμβάντων

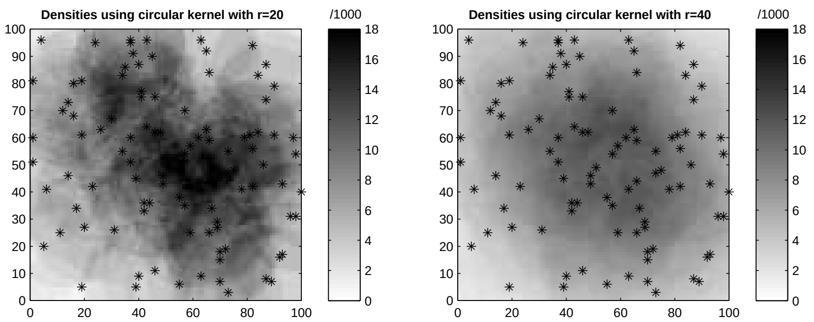






Τοπική Εκτίμηση Έντασης με Ολισθαίνοντα Κυκλικά Φατνία

1. ορισμός ενός κύκλου $C(\mathbf{u}, b)$ με ακτίνα b, και άρα εμβαδό $|C(\mathbf{u}, b)| = \pi b^2$, σε ένα οποιοδήποτε σημείο $\mathbf{u} \in D$
2. υπολογισμός της τοπικής έντασης στο σημείο $\mathbf{u}$ ως: $\hat{\lambda}(\mathbf{u}) = n(\mathbf{u}, b)/|C(\mathbf{u}, b)|$, όπου $n(\mathbf{u}, b)$ είναι η συχνότητα συμβάντων στον κύκλο $C(\mathbf{u}, b)$. <u>Σημείωση</u>: Τα βήματα 1 και 2 αποτελούν και μια συνάρτηση βαρών που απεικονίζεται ως ένας κύλινδρος με ακτίνα βάσης b και ύψος $1/(\pi b^2)$
3. επανάληψη του παραπάνω υπολογισμού για ένα σύνολο σημείων, συνήθως διατεταγμένων σε μια κάνναβο, για την κατασκευή μιας "επιφάνειας" έντασης συμβάντων στην περιοχή

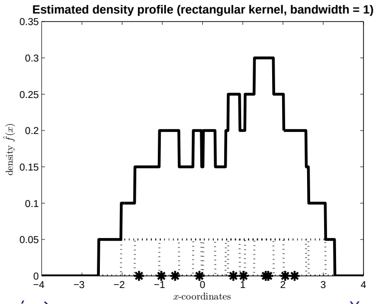


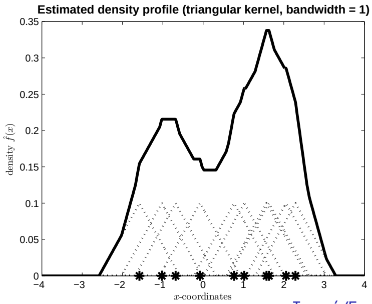
Αντί να προσεγγίσουμε την επιφάνεια έντασης από την πλευρά των σημείου εκτίμησης $\mathbf{u}$*, μπορούμε να την προσεγγίσουμε ισοδύναμα από την πλευρά των θέσεων συμβάντων* $\mathbf{u}_i$

Εκτίμηση Κατανομής Πυκνότητας: Προεπισκόπηση

Βασικές έννοιες

- η εκτίμηση πυκνότητας μέσω ενός ιστογράμματος πυκνότητας απαιτεί την επιλογή: (1) του αριθμού των διαστημάτων στα οποία διακριτοποιείται το εύρος τιμών μιας μεταβλητής, και (2) των αντίστοιχων μεσοδιαστημάτων
- αντί να επιλέξουμε ένα μικρό αριθμό διαστημάτων, μπορούμε να επιλέξουμε: (1) τόσα διαστήματα όσα και τα συμβάντα, και (2) μεσοδιαστήματα ταυτόσημα με τις θέσεις συμβάντων
- θα κατασκευάσουμε ιστογράμματα πυκνότητας αθροίζοντας το ύψος ράβδων με παραλληλόγραμμο ή άλλο σχήμα ...
- κάθε τέτοια ράβδος θα αντιστοιχεί σε ένα συμβάν, και κατά συνέπεια στα διαστήματα του άξονα των X όπου υπάρχουν πολλά συμβάντα η τελική κατανομή πυκνότητας θα έχει μεγάλο ύψος

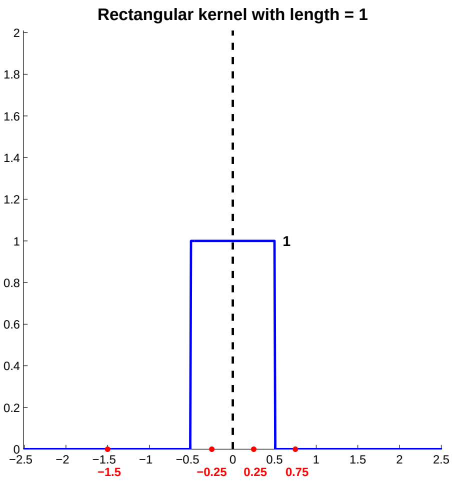




Εκτίμηση Κατανομής Συχνότητας: Επεξήγηση (I)

Παράδειγμα

Εκτίμηση συχνότητας (αριθμού) $f(x)$ συμβάντων στο μοναδιαίο διάστημα με μέσον στο σημείο $x = 0$, από ένα σύνολο $N = 4$ συμβάντων με συντεταγμένες $\{-1.5, -0.25, 0.25, 0.75\}$



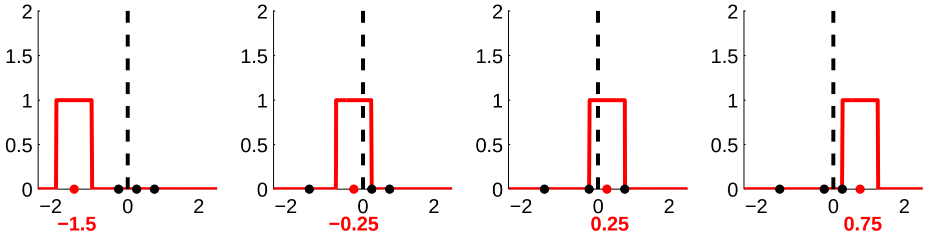
Διαδικασία εκτίμησης

(1) υπολογισμός N απόστασεων μεταξύ των N θέσεων συμβάντων x_i και του σημείου x, (2) καταμέτρηση συμβάντος στη θέση x_i αν $x_i - x < 0.5$, και (3) υπολογισμός (άθροισμα) του αριθμού των συμβάντων που κατακρατήθηκαν στο βήμα (2)

Εκτίμηση Κατανομής Συχνότητας: Επεξήγηση (II)

Ισοδύναμη διαδικασία

(1) "μετατόπιση" του μοναδιαίου διαστήματος ώστε να έχει μέσον ένα συμβάν x_i,
(2) ορισμός ενός (0/1) δείκτη (βάρους) για το αν το μετατοπισμένο διάστημα εμπεριέχει ($= 1$) ή όχι ($= 0$) το σημείο x, και (3) επανάληψη για όλα τα N συμβάντα και άθροισμα



Εκτίμηση πυρήνα

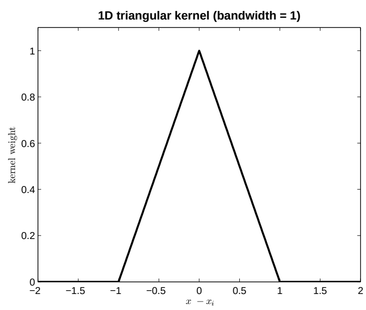
$$f(x) = \sum_{i=1}^{N} k(x - x_i) = k(0 + 1.5) + k(0 + 0.25) + k(0 - 0.25) + k(0 - 0.75) = 0 + 1 + 1 + 0$$

όπου $k(x, x_i)$ είναι μια συνάρτηση απόστασης, που ορίζεται ως: $k(x - x_i) = 1$ αν $\|x - x_i\| \leq 1/2$, και $k(x - x_i) = 0$ σε κάθε άλλη περίπτωση

Η συνάρτηση πυρήνα $k(x, x_i)$ αντιστοιχεί σε κάθε απόσταση $h_i = \|x - x_i\|$ μεταξύ ενός σημείου x και του κέντρου του πυρήνα x_i ένα **βάρος** $k_i(x)$ με το οποίο η θέση συμβάντος x_i επηρεάζει τον αριθμό συμβάντων στη θέση x

Μονοδιάστατος ή Μονομεταβλητός Πυρήνας

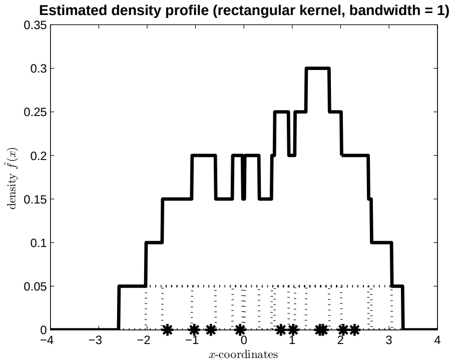
Εξίσωση που δίνει την πιθανότητα εμφάνισης μιας τιμής της συντεταγμένης x, δηλαδή την πιθανότητα εμφάνισης ενός συμβάντος σε οποιαδήποτε θέση x, <u>δεδομένης</u> της παρουσίας ενός συμβάντος στη θέση x_i

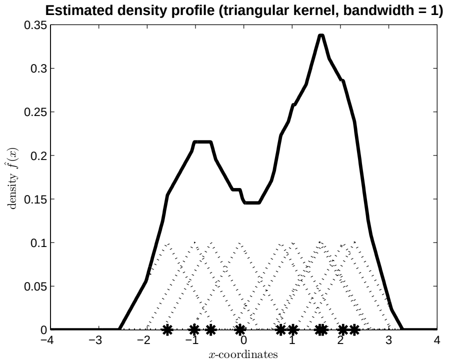


Χαρακτηριστικά πυρήνα

- συνάρτηση της απόσταση $h_i = \|x - x_i\|$ μεταξύ μιας οποιασδήποτε θέσης x και της θέσης x_i ενός συμβάντος: $\boxed{k(x, x_i) = k(\|x - x_i\|) = k_i(h)}$, όπου h είναι η απόσταση μεταξύ της θέσης x και του κέντρου του πυρήνα, εδώ της θέσης συμβάντος x_i (στο γράφημα $x_i = 0$)
- ως πυρήνες λαμβάνονται (συμμετρικές) συναρτήσεις πυκνότητας πιθανότητας (ΣΠΠ), οι οποίες είναι μη αρνητικές και έχουν ολοκλήρωμα 1, δηλαδή: $k(h) \geq 0$, και $\int k(h)dh = 1$
- ως ΣΠΠ, οι πυρήνες έχουν μια αναμενόμενη (μέση) τιμή, 0 γιατί ο άξονας των X απεικονίζει απόσταση από το κέντρο, και μια διακύμανση: $\int hk(h)dh = 0$ και $0 < \int h^2 k(h)dh < \infty$

Πυρήνες και Εκτίμηση Πυκνότητας





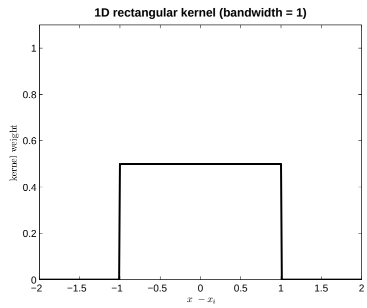
Σχέση με εκτίμηση πυκνότητας

- Αντί να κρατήσουμε σταθερό τον αριθμό και το κέντρο των διαστημάτων ενός ιστογράμματος πυκνότητας, μπορούμε να εκτιμήσουμε την πυκνότητα $p(x)$ σε μια οποιαδήποτε θέση x ως ένα σταθμευμένο άθροισμα N τιμών $k(x - x_i)$, όπου κάθε τέτοια τιμή προέρχεται από ένα διαφορετικό πυρήνα $k_i(h)$ "τοποθετημένο" σε μια διαφορετική θέση συμβάντος x_i
- συνήθως, όλοι οι N πυρήνες που αντιστοιχούν στις N θέσεις συμβάντων λαμβάνονται ίδιοι: $k_i(h) = k(h), \forall i$

Παραδείγματα Συναρτήσεων Μονοδιάστατων Πυρήνων I

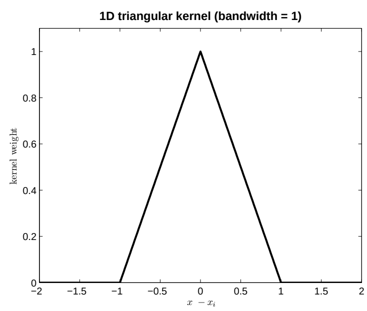
- Ομοιόμορφος ή Parzen:

$$k(h) = \begin{cases} 1/2 & \text{αν } h \in [-1\ 1] \\ 0 & \text{αν όχι} \end{cases}$$



- Τριγωνικός:

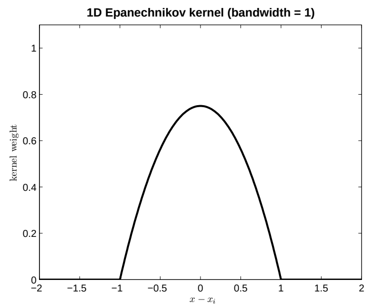
$$k(h) = \begin{cases} 1 - |h| & \text{αν } h \in [-1\ 1] \\ 0 & \text{αν όχι} \end{cases}$$



Παραδείγματα Συναρτήσεων Μονοδιάστατων Πυρήνων II

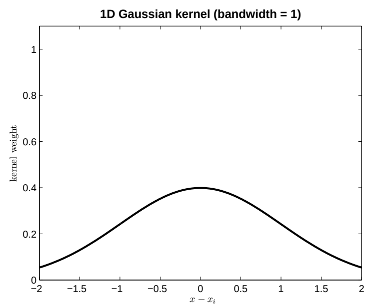
- Τετραγωνικός ή Epanechnikov:

$$k(h) = \begin{cases} \frac{3}{4}(1 - h^2) & \text{αν } h \in [-1 \; 1] \\ 0 & \text{αν όχι} \end{cases}$$



- Κανονικός ή Gaussian:

$$k(h) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}h^2)$$



Πυρήνες που τείνουν ασυμπτωτικά στο 0, π.χ., αυτός του Gauss, λέγονται μη-συμπαγείς πυρήνες

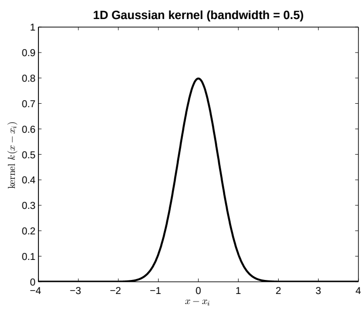
Κλιμακωτός Μονοδιάστατος Πυρήνας

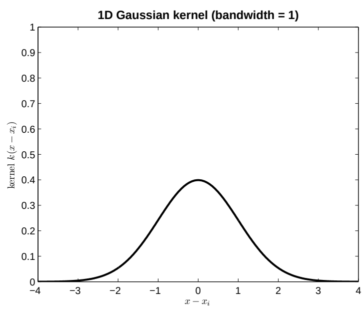
Εναλλακτική θεώρηση πυρήνα

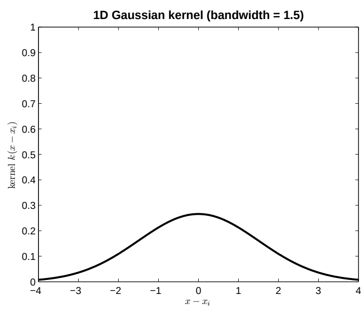
Μια συνάρτηση πυρήνα $k(x, x_i)$ ποσοτικοποιεί την "επίδραση" ενός συμβάντος στη θέση x_i στο "περιβάλλον" του – σε οποιαδήποτε θέση x – με βάση την απόσταση $h_i = \|x - x_i\|$ μεταξύ του συμβάντος x_i και του σημείου x, δηλαδή: $k(x, x_i) = k(\|x - x_i\|) = k(h_i)$

Κλιμάκωση πυρήνα

Η επίδραση του συμβάντος στη θέση x_i σε οποιαδήποτε θέση x μπορεί να μεταβληθεί διαιρώντας την απόσταση $h_i = \|x - x_i\|$ με ένα βαθμωτό b που ονομάζεται φάσμα (bandwidth) του πυρήνα: $k(x, x_i; b) = \frac{1}{b} k\left(\frac{\|x - x_i\|}{b}\right)$. Διαιρώντας τα εξερχόμενα της συνάρτησης πυρήνα με $1/b$, ο κλιμακωτός πυρήνας έχει ολοκλήρωμα 1







Μη-τυπική κανονική συνάρτηση πυκνότητας πιθανότητας (ΣΠΠ): $k(h; b) = \frac{1}{b\sqrt{2\pi}} \exp[-\frac{1}{2}(\frac{h}{b})^2]$ με τυπική απόκλιση b

Εκτίμηση Πυκνότητας με Πυρήνα σε Μία Διάσταση I

Διαδικασία και επιλογές

1. επιλογή μιας συνάρτησης πυρήνα $k(x - x_i)$, δηλαδή μιας ΣΠΠ, και μιας τιμής φάσματος b που καθορίζει το εύρος του πυρήνα. Με άλλα λόγια επιλογή μιας συνάρτησης κλιμακωτού πυρήνα $k(x - x_i; b)$
2. διακριτοποίηση του εύρους συντεταγμένων, δηλαδή επιλογή ενός συνόλου διακριτών τιμών x-συντεταγμένων όπου η πυκνότητα $p(x)$ θα εκτιμηθεί
3. για κάθε συμβάν x_i, υπολογισμός της επίδρασης, μέσω του κλιμακωτού πυρήνα $k(x - x_i; b)$, του συμβάντος αυτού σε όλες τις διακριτές τιμές του x. Αυτό δίνει N κλιμακωτούς πυρήνες όπου κάθε πυρήνας αντιστοιχεί σε ένα συγκεκριμένο συμβάν x_i

Εκτίμηση Πυκνότητας με Πυρήνα σε Μία Διάσταση II

Στάδιο εκτίμησης πυκνότητας

Για κάθε διακριτή τιμή x, η εκτιμόμενη πυκνότητα $\hat{p}(x)$ υπολογίζεται ως ο αριθμητικός μέσος όρος N τιμών πυρήνα $k(x - x_i; b)$, όπου κάθε τιμή συμβάλει με βάρος $1/N$:

$$\hat{p}(x) = \sum_{i=1}^{N} \frac{1}{N}\frac{1}{b} k\left(\frac{x - x_i}{b}\right) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{b} k\left(\frac{x - x_i}{b}\right)$$

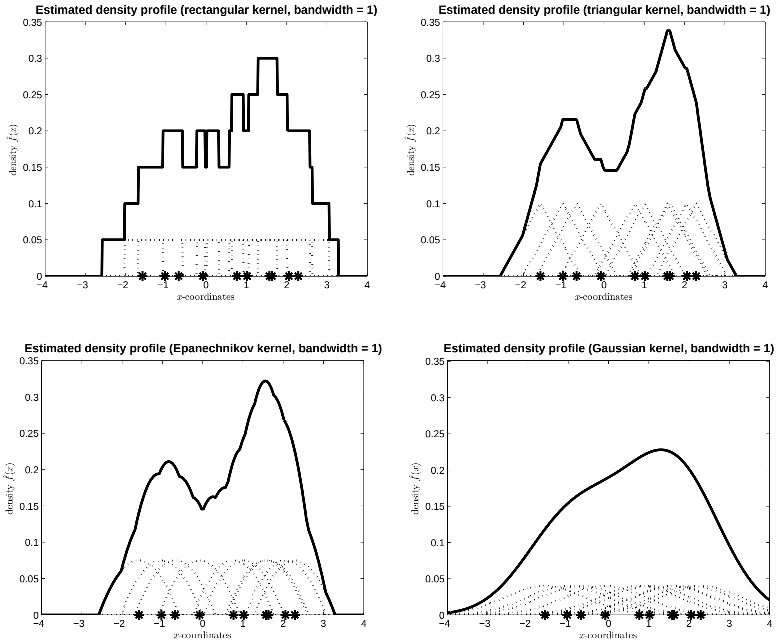
Τελικό αποτέλεσμα

Εκτιμόμενες τιμές πυκνότητας $\hat{p}(x)$ στις διακριτές τιμές του x. Οι N κλιμακωτοί και σταθμισμένοι πυρήνες στο βήμα (3) της προηγούμενης σελίδας (ή στην παρένθεση παραπάνω) μπορούν να θεωρηθούν ως N αρχέτυπες κατανομές πιθανότητας, των οποίων η επίθεση ή επαλληλία χτίζει την τελική κατανομή της εκτιμόμενης πυκνότητας

Σχέση με ένταση συμβάντων

Για να μετατρέψουμε μια τιμή πυκνότητας συμβάντων $p(x)$ στην αντίστοιχη τιμή έντασης συμβάντων $\lambda(x)$ χρειάζεται απλά ο πολλαπλασιασμός της πυκνότητας $p(x)$ με το εύρος $|s|$ του διαστήματος s στο οποίο αυτή εκτιμήθηκε, δηλαδή: $\lambda(x) = p(x) \cdot |s|$

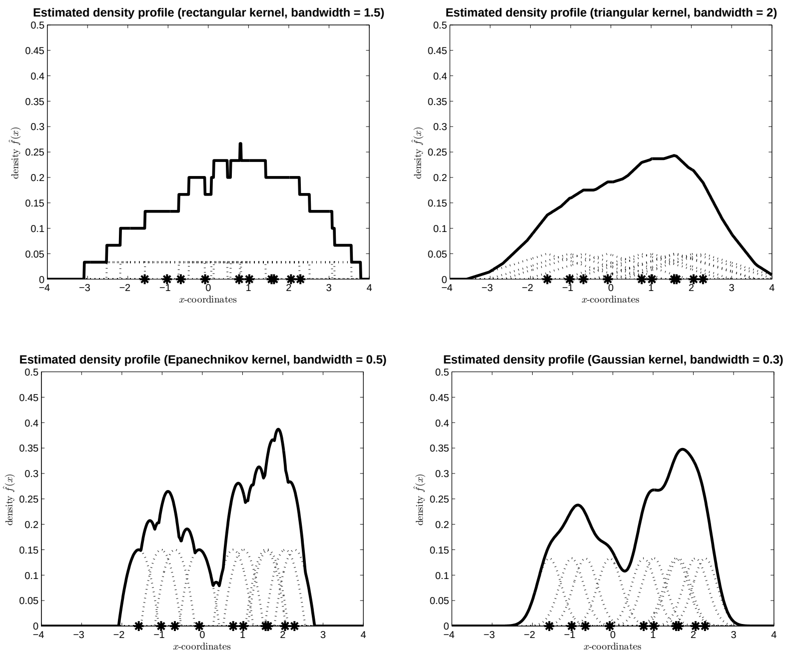
Παραδείγματα Εκτίμησης Πυκνότητας με Πυρήνα I



Οι διακεκομμένες καμπύλες απεικονίζουν σταθμισμένα (με βάρος $1/N$) και κλιμακωτούς πυρήνες

Υπάρχουν *κανόνες* για την επιλογή ενός "βέλτιστου" φάσματος, οι οποίοι συνήθως βασίζονται στην παραδοχή ότι τα N δεδομένα ακολουθούν κάποια θεωρητική κατανομή, π.χ., την κανονική. Η εκτιμώμενη κατανομή επηρεάζεται περισσότερο από το φάσμα παρά από τον τύπο της συνάρτησης πυρήνα. . .

Παραδείγματα Εκτίμησης Πυκνότητας με Πυρήνα II



Οι διακεκομμένες καμπύλες απεικονίζουν σταθμισμένα (με βάρος $1/N$) και κλιμακωτούς πυρήνες

Όσο μικρότερο είναι το φάσμα του πυρήνα, τόσο πιο "τραχειά" είναι η εκτιμώμενη κατανομή. Πολύ μεγάλες τιμές φάσματος οδηγούν σε υπερ-εξομαλυνσμένες κατανομές, οι οποίες δεν παρουσιάζουν καμία ενδιαφέρουσα λεπτομέρεια. . .

Διαχωρίσιμοι Πυρήνες

Δύο μονοδιάστατοι πυρήνες Gauss

$$k_x(x - x_i; b_x) = \frac{1}{b_x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_i}{b_x}\right)^2\right]$$

$$k_y(y - y_i; b_y) = \frac{1}{b_y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y - y_i}{b_y}\right)^2\right]$$

$\mathbf{u}_i = (x_i, y_i)$ = συντεταγμένες συμβάντων, $\mathbf{u} = (x, y)$ = συντεταγμένες άλλης θέσης, b_x και b_y = φάσματα πυρήνων

Σύνθετος διδιάστατος πυρήνας

$$k(x - x_i, y - y_i; b_x, b_y) = \frac{1}{2\pi b_x b_y} \exp\left[-\frac{1}{2}\left(\frac{x - x_i}{b_x}\right)^2 - \frac{1}{2}\left(\frac{y - y_i}{b_y}\right)^2\right]$$

διδιάστατη κανονική ΣΠΠ δύο ανεξάρτητων τυχαίων μεταβλητών, γινόμενο δύο μονομεταβλητών κανονικών ΣΠΠ

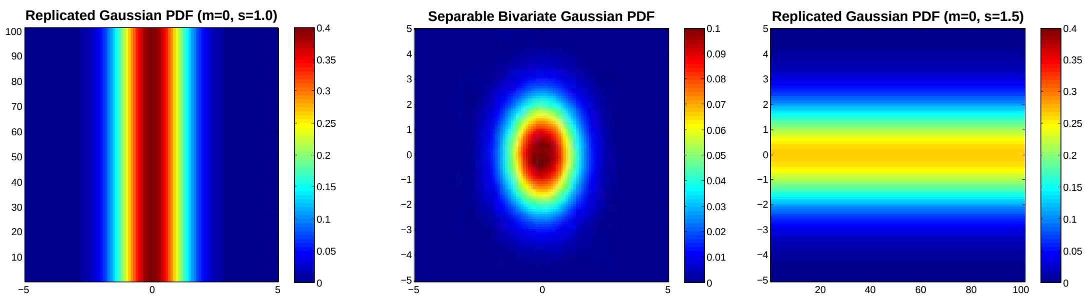
Διαχωρισιμότητα

Ένας διδιάστατος πυρήνας που μπορεί να κατασκευαστεί ως το γινόμενο δύο μονοδιάστατων πυρήνων λέγεται *διαχωρίσιμος* πυρήνας

Κατασκευάζοντας Ένα Διαχωρίσιμο Διδιάστατο Πυρήνα

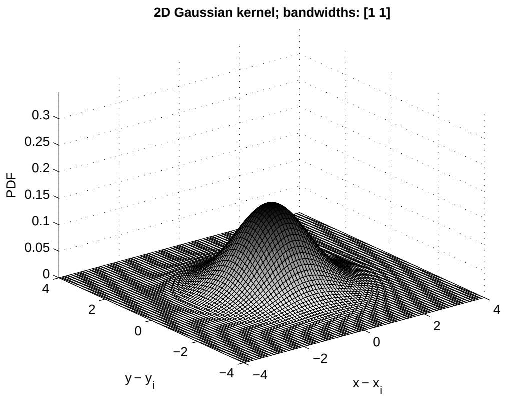
Δύο μονοδιάστατοι πυρήνες Gauss για την x- και y-διάσταση

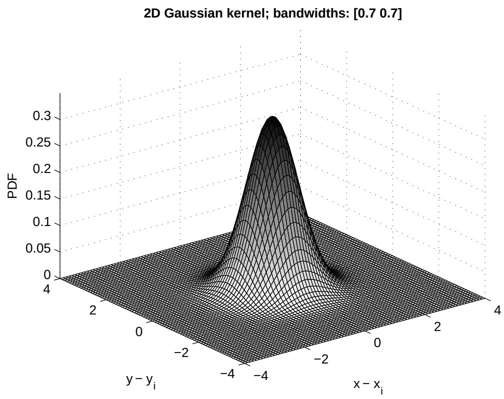
Αντιγραφή μονοδιάστατων πυρήνων και διδιάστατος διαχωρίσιμος πυρήνας

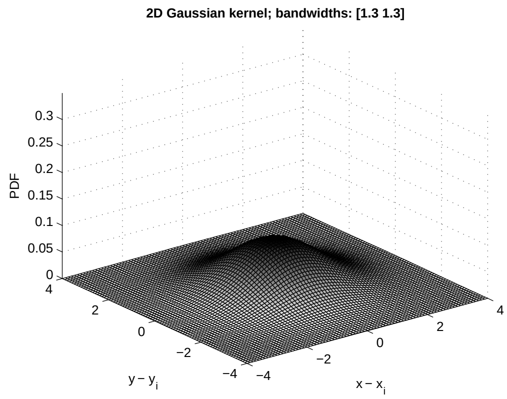


Ανισότροπος πυρήνας = πολυδιάστατος πυρήνας με διαφορετικές τιμές φάσματος στις διάφορες κατευθύνσεις

Παραδείγματα Διδιάστατων Πυρήνων Gauss







<u>ισότροπος πυρήνας</u> = πολυδιάστατος πυρήνας με ίδια τιμή φάσματος σε όλες τις διευθύνσεις

Πολυδιάστατη Εκτίμηση Πυκνότητας με Πυρήνα

Πολυμεταβλητοί διαχωρίσιμοι πυρήνες

Ένας K-διάστατος πυρήνας μπορεί να κατασκευαστεί ως το γινόμενο $K \geq 2$ μονοδιάστατων πυρήνων:

$$f(\mathbf{u} - \mathbf{u}_i; \mathbf{b}) = \prod_{k=1}^{K} \frac{1}{b_k} f\left(\frac{u_k - u_{ik}}{b_k}\right)$$

u_k είναι η συντεταγμένη του σημείου $\mathbf{u}$ στη διάσταση k, u_{ik} είναι η συντεταγμένη της θέσης συμβάντος $\mathbf{u}_i$ στη διάσταση k, b_k είναι η τιμή του φάσματος στη διάσταση k, και $\mathbf{b}$ είναι ένα διάνυσμα με K τιμές φάσματος

Εκτιμόμενη πυκνότητα

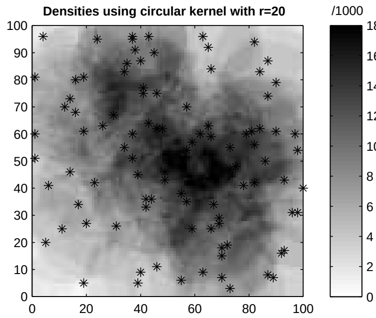
Εφαρμόζωντας την ίδια διαδικασία όπως και στη μονοδιάστατη περίπτωση:

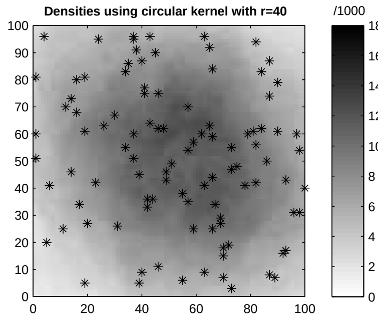
$$\hat{p}(\mathbf{u}) = \sum_{i=1}^{N} \frac{1}{N} f(\mathbf{u} - \mathbf{u}_i; \mathbf{b}) = \frac{1}{N b_1 \cdots b_K} \sum_{i=1}^{N} \left[\prod_{k=1}^{K} f\left(\frac{u_k - u_{ik}}{b_k}\right)\right]$$

Υπάρχουν κανόνες, ανάλογοι με αυτούς για τη μονοδιάστατη περίπτωση, για την επιλογή "βέλτιστων" τιμών φάσματος στις διάφορες διευθύνσεις

Εκτίμηση Κατανομής Έντασης σε Δύο Διαστάσεις I

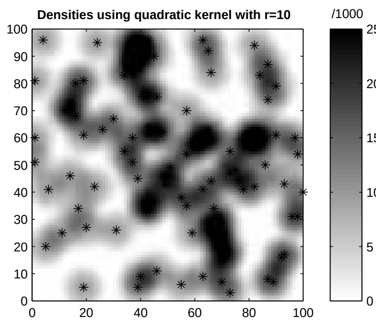
- μετατροπή σημειακής χωρικής κατανομής (θέσεις συμβάντων) σε "συνεχή επιφάνεια" έντασης $\{\hat{\lambda}(\mathbf{u}), \mathbf{u} \in D\}$, που χρησιμοποιείται για την απεικόνιση μεγάλης κλίμακας πρότυπων (large-scale patterns) και για την αναγνώριση "θερμών κηλίδων" (hot spots) στη χωρική κατανομή των συμβάντων
- μεγαλύτερες τιμές φάσματος πυρήνα b οδηγούν σε "ομαλές" επιφάνειες έντασης, ενώ το αντίθετο συμβαίνει για μικρότερες τιμές

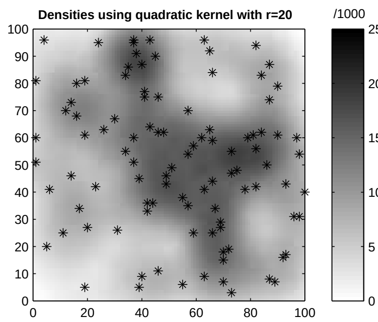




Εκτίμηση Κατανομής Έντασης σε Δύο Διαστάσεις II

- διαφορετικοί τύποι πυρήνων προσδίδουν περισσότερο βάρος στα γειτονικά συμβάντα κατά τον υπολογισμό της έντασης $\hat{\lambda}(\mathbf{u})$
- θεωρητικά, θα πρέπει να υπάρχει ένα θεωρητικό υπόβαθρο για την επιλογή μιας συνάρτησης πυρήνα, π.χ., ο κανονικός (Gauss) πυρήνας είναι συμβατός με διαδικασίες χωρικής διάχυσης
- είναι επίσης δυνατή η χρησιμοποίηση τοπικά μεταβαλόμενου φάσματος $b(\mathbf{u})$, σε συνάρτηση με το πλήθος των συμβάντων στη γειτονιά του σημείου $\mathbf{u}$. Η διαδικασία αυτή λέγεται προσαρμοσμένη (adaptive) εκτίμηση





Ανακεφαλαίωση I

Ένταση συμβάντων

- ορισμός: συχνότητα (αριθμός) συμβάντων σε ένα διάστημα ή μία επιφάνεια
- γενικευμένη ένταση: αριθμός συμβάντων στο μοναδιαίο εμβαδόν: $\hat{\lambda} = N/|D|$
- τοπική ένταση: ένα μέτρο συγκέντρωσης συμβάντων μέσα σε μια προ-επιλεγμένη περιοχή s, π.χ., σε ένα φατνίο, η οποία εκτιμάται ως: $\hat{\lambda}(s) = n(s)/|s|$

Εκτίμηση τοπικής έντασης με χρήση καννάβου

- ορισμός συστήματος φατνίων που αποτελούν μια κάνναβο, δηλαδή ορισμός του: (1) αριθμού φατνίων στις δύο διαστάσεις, (2) του μήκους κάθε φατνίου στις δύο διαστάσεις, και (3) των συνεταγμένων του σημείου αφετηρίας της καννάβου
- υπολογισμός της έντασης συμβάντων σε κάθε φατνίο της καννάβου και απεικόνιση – διδιάστατο ανάλογο ενός ιστογράμματος
- η "επιφάνεια" της χωρικής έντασης που προκύπτει επηρεάζεται από όλες τις παραμέτρους που ορίζουν την κάνναβο των φατνίων, π.χ., μικρά φατνία αφήνουν κενά στην επιφάνεια, ενώ μεγάλα φατνία αποκρύπτουν ενδιαφέρουσες λεπτομέρειες
- η χρήση χωρικά επικαλυπτόμενων φατνίων οδηγεί στην έννοια μιας συνεχούς κατανομής (επιφάνειας) χωρικής έντασης (μη-παραμετρική εκτίμησης πυκνότητας)

Ανακεφαλαίωση II

Εκτίμηση τοπικής έντασης με χρήση πυρήνα

- μετατροπή σημειακής κατανομής συμβάντων σε "συνεχή επιφάνεια" έντασης
- η στατιστική μέθοδος της πολυμεταβλητής εκτίμησης πυκνότητας πιθανότητας χρησιμοποιείται για την κατασκευή της επιφάνειας πυκνότητας σε πολυδιάστατους χώρους. Σημείωση: Το ολοκλήρωμα της εκτιμόμενης επιφάνειας πυκνότητας πιθανότητας είναι 1, και συνεπώς οι τιμές της επιφάνειας αυτής μετατρέπονται (με πολλαπλασιασμό με τον αριθμό των συμβάντων N) σε τιμές έντασης, έτσι ώστε η τελική επιφάνεια έντασης να έχει ολοκλήρωμα N
- η χωρική κατανομή της επιφάνειας έντασης επηρεάζεται από: (1) τον τύπο του πυρήνα (κυρίως), και (2) από το φάσμα του πυρήνα
- δύο εναλλακτικές μέθοδοι μη-παραμετρικής εκτίμησης πυκνότητας πιθανότητας είναι η μέθοδος πλησιέστερου γείτονα (nearest neighbor) και η μέθοδος μίγματος κανονικών κατανομών (mixture of Gaussian densities)
- η εκτιμόμενη επιφάνεια έντασης είναι δυνατό να συσχετισθεί (μέσω ενός μοντέλου παλινδρόμησης) με άλλες χωρικές μεταβλητές, π.χ., συσχέτιση έντασης εμφάνισης νόσου με ποιότητα αέρα. Σημείωση: Οποιαδήποτε σχέση βρεθεί με την παραπάνω διαδικασία ισχύει για το συγκεκριμένο τύπο πυρήνα, το φάσμα του, και τη χωρική ανάλυση (discretization) που επιλέχθηκαν