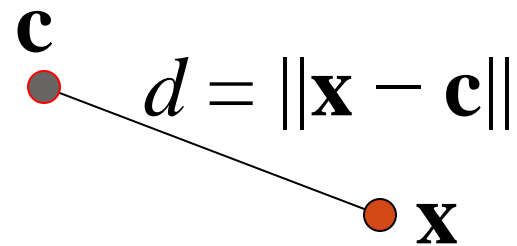


Δίκτυα Συναρτήσεων Βάσης Ακτινικού Τύπου

Κώστας Διαμαντάρας
Τμήμα Πληροφορικής
ΤΕΙ Θεσσαλονίκης

Συναρτήσεις Ακτινικού τύπου

- Radial Basis Functions (RBF)
- \mathbf{x} = διάνυσμα εισόδου
- \mathbf{c} = διάνυσμα κέντρου



Έξοδος = συνάρτηση της απόστασης του \mathbf{x} από το κέντρο \mathbf{c}

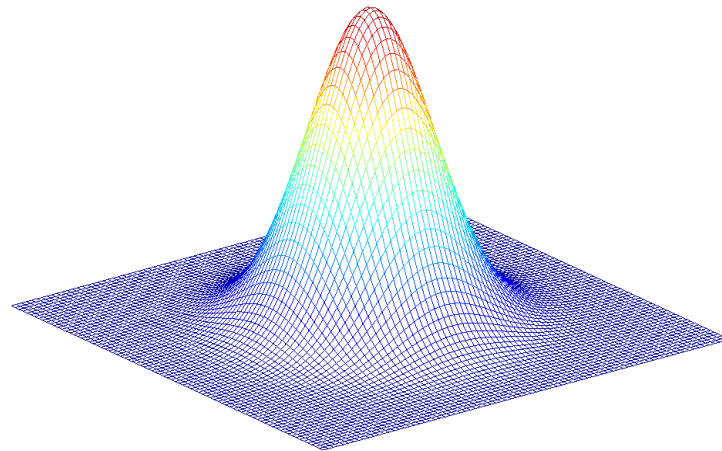
$$f(\mathbf{x}) = f(\|\mathbf{x} - \mathbf{c}\|)$$

Συναρτήσεις Ακτινικού τύπου

Παραδείγματα (1)

- *Συνάρτηση Gauss*

$$f(x) = e^{-\frac{\|x-c\|^2}{\sigma^2}}$$

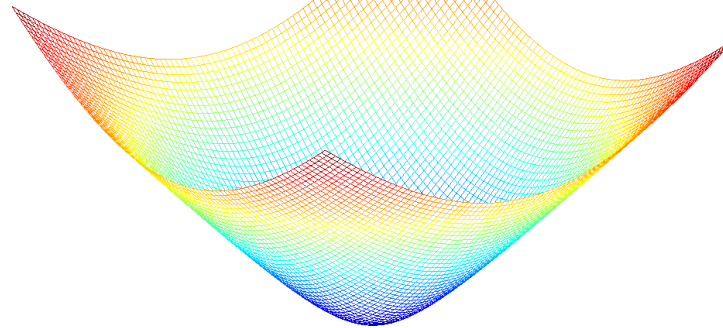


Συναρτήσεις Ακτινικού τύπου

Παραδείγματα (2)

- Πολυτετραγωνική συνάρτηση

$$f(x) = (\|\mathbf{x} - \mathbf{c}\|^2 + \sigma^2)^{1/2}$$

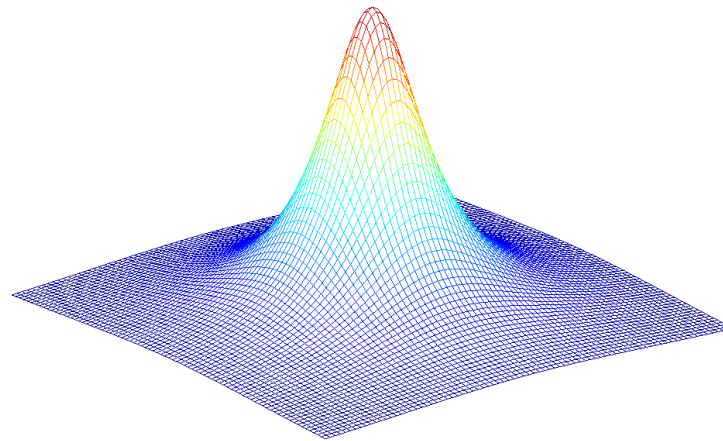


Συναρτήσεις Ακτινικού τύπου

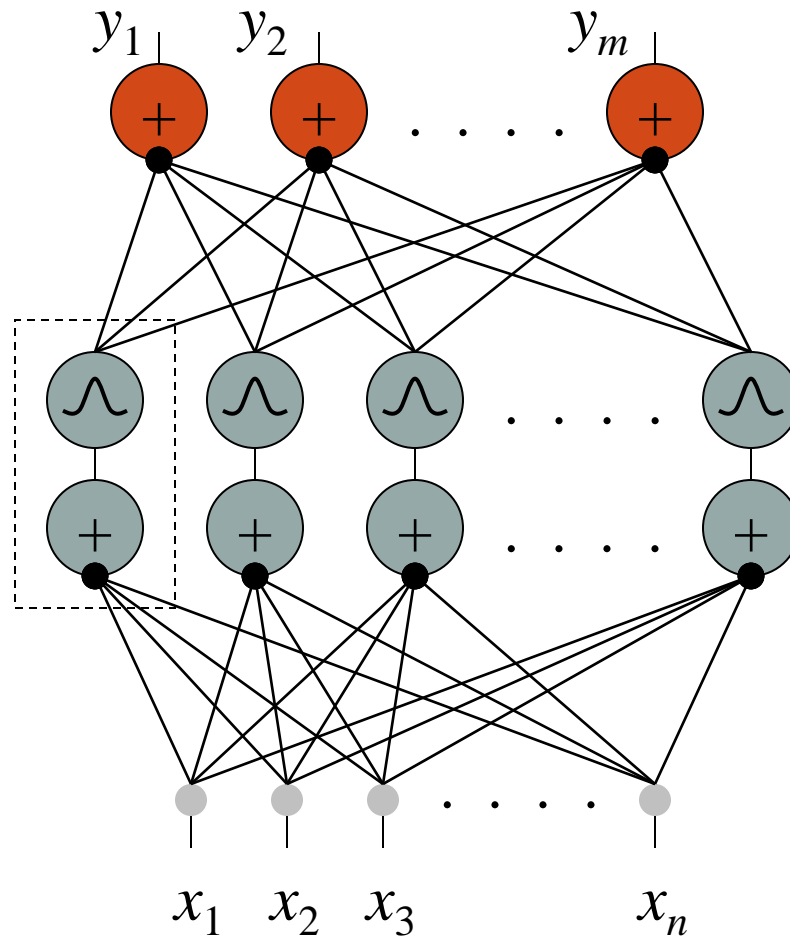
Παραδείγματα (3)

- *Συνάρτηση Cauchy*

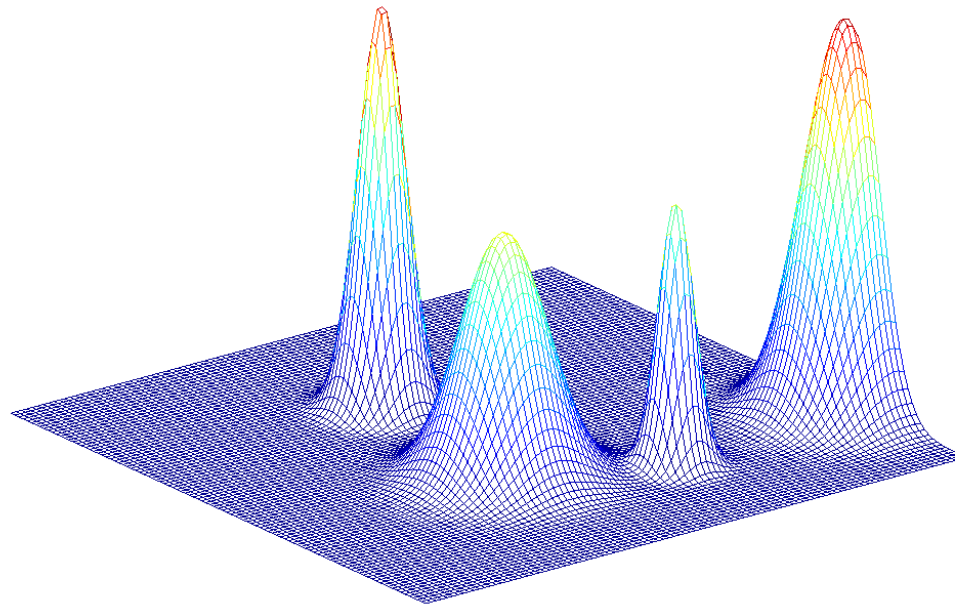
$$f(x) = \frac{(\|\mathbf{x} - \mathbf{c}\|^2 + \sigma^2)^{-1}}{\sigma}$$



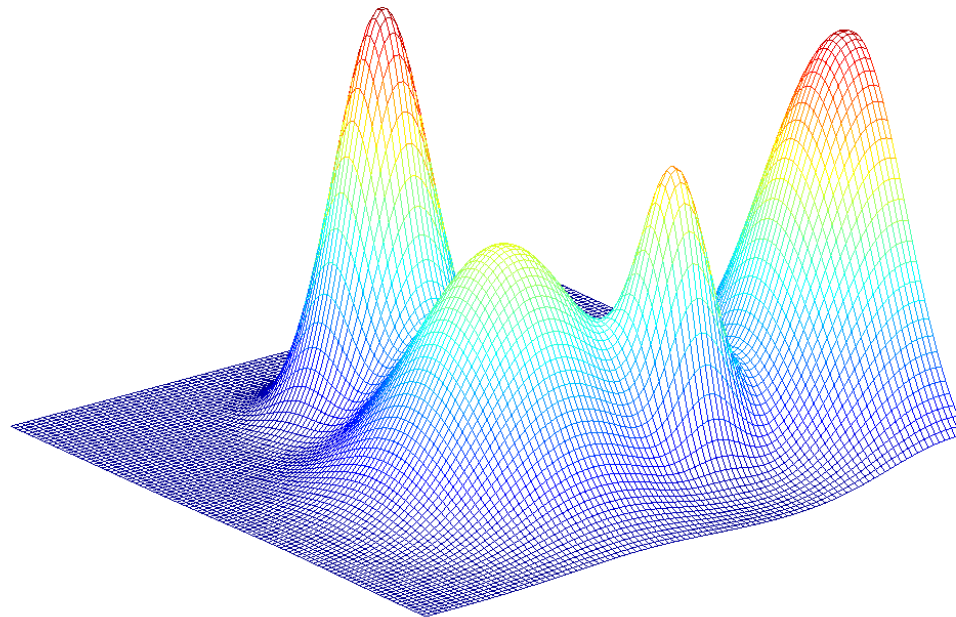
Τοπολογία Δικτύου RBF



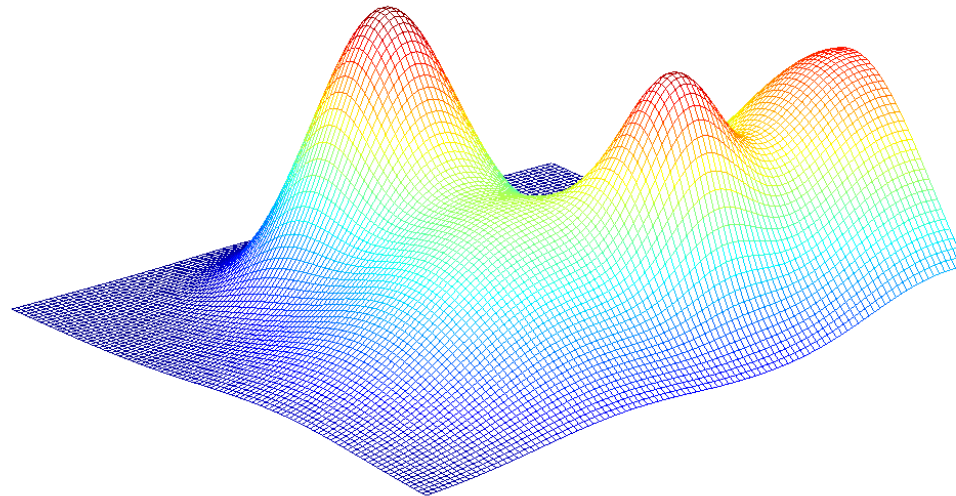
Έξοδος RBF (1)



Έξοδος RBF (2)



Έξοδος RBF (3)



Δυνατότητες Δικτύων RBF

- Θεώρημα Παγκόσμιου προσεγγιστή, Park & Sandberg 1991

οι συναρτήσεις ακτινικού τύπου είναι βάση του χώρου των συνεχών συναρτήσεων



προσεγγίζουν οποιαδήποτε συνεχή συνάρτηση
ικανό πλήθος κρυφών νευρώνων

Ομοιότητες/Διαφορές με MLP

- RBF: μόνο 2 στρώματα
 - δεν υπάρχει ικανοποιητικός κανόνας εκπαίδευσης για 3, 4, κλπ στρώματα
- MLP: μοιάζει περισσότερο με πραγματικούς νευρώνες
- RBF, MLP: εκπαίδευση με επίβλεψη
- RBF: αυτο-οργανούμενο κρυφό στρώμα

Σύγκριση RBF – MLP

- Και τα δύο παγκόσμιοι προσεγγιστές



- Ισοδυναμία μεταξύ δικτύων RBF και μηχανών ασαφούς λογικής (*fuzzy logic*).

RBF ή MLP?

- Κριτήριο = ποιότητα προσέγγισης
 - πεπερασμένο πλήθος νευρώνων
 - συγκεκριμένα δεδομένα

Ανάκληση

Είσοδοι: Διάνυσμα εισόδου $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$
κρυφό στρώμα = p νευρώνες

Έξοδοι: y_1, y_2, \dots, y_m

Για κάθε κρυφό νευρώνα $i = 1, \dots, p$

$$\{ a_i = f(\|\mathbf{x} - \mathbf{c}_i\|, \sigma_i) \}$$

Για κάθε νευρώνα εξόδου $i = 1, \dots, m$

$$\{ y_i = \sum_{j=1}^p w_{ij} a_j + w_{i0} \}$$

Εκπαίδευση Δικτύων RBF

Κανόνας εκπαίδευσης κρυφού στρώματος

≠

Κανόνας εκπαίδευσης εξωτερικού στρώματος

- Το κρυφό στρώμα αποτελεί το βασικό στρώμα αναπαράστασης της συνάρτησης. Εκπαιδεύεται ξεχωριστά και πιο αργά.
- Το στρώμα εξόδου εκπαιδεύεται κατόπιν, εύκολα και γρήγορα

Εκπαίδευση κρυφού στρώματος (1)

- Εκπαίδευση χωρίς επίβλεψη (unsupervised)
- Επιλογή κέντρων $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_p$
- Επιλογή εύρους $\sigma_1, \sigma_2, \dots, \sigma_p$

Εκπαίδευση κρυφού στρώματος: Μέθοδοι (1)

Κάθε πρότυπο εισόδου = ένα κέντρο

Ίσα πλάτη

$$\mathbf{x}_i = \mathbf{c}_i$$

$$\sigma_i = \dots$$

$$\frac{d}{\sqrt{2p}}$$

Εκπαίδευση κρυφού στρώματος: Μέθοδοι (2)

Επιλογή στην τύχη κάποιων προτύπων εισόδου για να
είναι κέντρα

$$\mathbf{c}_i = \mathbf{x}_j$$

Πλάτη ίσα μεταξύ τους

$$\sigma_i = \dots$$

$$\frac{d}{\sqrt{2p}}$$

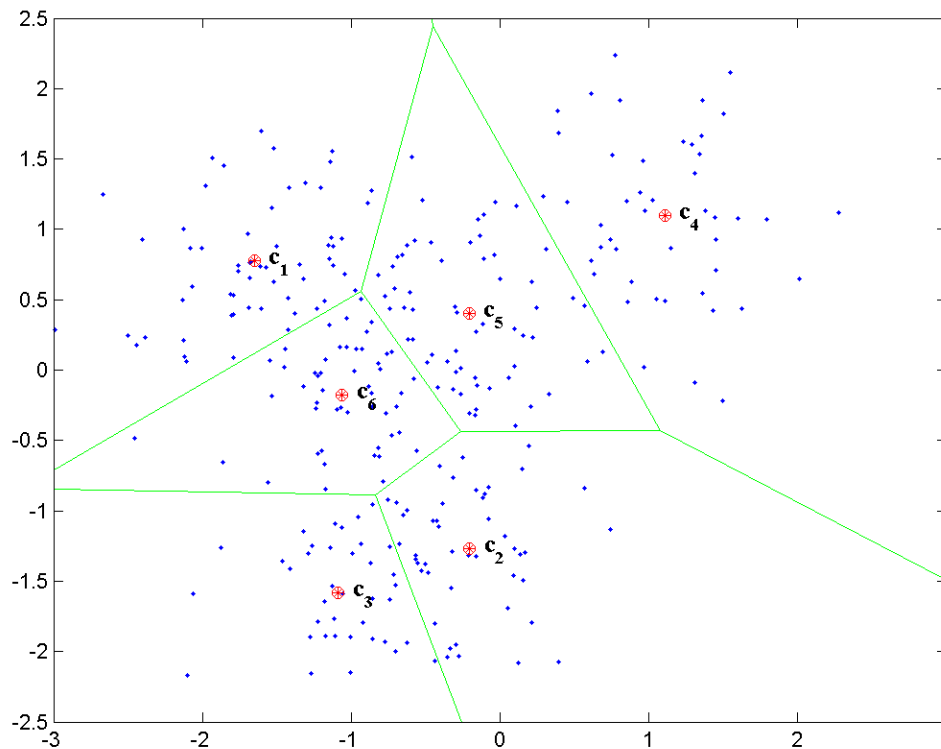
Εκπαίδευση κρυφού στρώματος: Μέθοδοι (3)

Ομαδοποίηση (clustering)

Μέθοδος k-μέσων

1. Κάθε πρότυπο ανήκει στην ομάδα (cluster) του κέντρου που βρίσκεται πιο κοντά σ' αυτό
2. Κάθε κέντρο είναι ο μέσος όρων των προτύπων που ανήκουν στην ομάδα του

Εκπαίδευση κρυφού στρώματος: Μέθοδος k-μέσων



Ο κανόνας k-μέσων

Είσοδος: $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$, πλήθος κέντρων K

Έξοδος: K κέντρα $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_K$

Αρχικοποίησε τα $\mathbf{c}_1, \dots, \mathbf{c}_K$ σε τυχαίες τιμές.

Επανάλαβε {

Για κάθε πρότυπο $i = 1, \dots, P$

{ Βρες το κοντινότερο κέντρο \mathbf{c}_j στο \mathbf{x}_i
 $label(i) \leftarrow j$ }

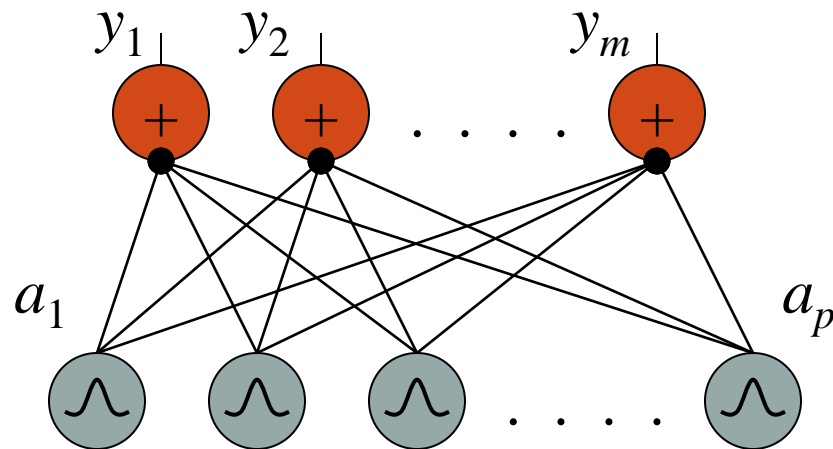
Για κάθε κέντρο $k = 1, \dots, K$

{ $\mathbf{c}_k \leftarrow$ μέσος όρος των \mathbf{x}_i τα οποία ανήκουν στην κλάση k , δηλαδή $label(i) = k$ }

} Μέχρι να μην υπάρξει καμία αλλαγή στα κέντρα $\mathbf{c}_1, \dots, \mathbf{c}_K$

Εκπαίδευση εξωτερικού στρώματος

- Χρήση του κανόνα δέλτα (delta-rule) όπως στο ADALINE



Εκπαίδευση εξωτερικού στρώματος (2)

Κανόνας Δέλτα (ADALINE)

- Στόχοι $d(k)$
- Διάνυσμα ενεργοποίησης κρυφού στρώματος $\mathbf{a}(k)$
- Βήμα εκπαίδευσης $\beta(k)$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \beta(k) (d(k) - y(k)) \mathbf{a}(k)$$

Εφαρμογές (1)

- Μοντελοποίηση χρονοσειρών

$$y(t) = F(\mathbf{x}(t))$$

- Πρόβλεψη χρονοσειράς
- Εκτίμηση παραμέτρων
- Moody & Darken εκτίμηση παραμέτρων χαοτικής χρονοσειράς Mackey-Glass
- RBF: 27 φορές περισσότερα δεδομένα από το MLP για ίδιο επίπεδο απόδοσης
- RBF: 16 φορές πιο γρήγορο από το BP

Εφαρμογές (2): Ταξινόμηση φωνημάτων

Πρόβλημα: Ταξινόμηση δέκα ήχων φωνηέντων

- 338 φωνήματα για εκπαίδευση
- 333 φωνήματα για έλεγχο
- Μέθοδος k-μέσων
- 100 κρυφοί Γκαουσιανοί νευρώνες
 - Σφάλμα = 18% στα πρότυπα ελέγχου

Εφαρμογές (3)

- Τιμολόγηση Χρηματιστηριακών παραγώγων (option pricing)
- Εξισορρόπηση τηλεπικοινωνιακού καναλιού (channel equalization)