Άσκηση 2

Στην άσκηση αυτή επιθυμούμε να βρούμε γραμμικό διαχωριστή για να κατηγοριοποιήσουμε τα σημεία ενός δισδιάστατου χώρου. Καθότι ο εντοπισμός μιας γραμμικής σχέσης για τον διαχωρισμό των δεδομένων στον αρχικό χώρο μπορεί να είναι αδύνατος, σε περίπτωση που αυτά διέπονται από κάποιον μη γραμμικό (π.χ. πολυωνυμικό, εκθετικό κλπ κλπ) κανόνα, συχνά τα δεδομένα αναπαρίστανται σε έναν υψηλότερης (ενδεχομένως άπειρης) διάστασης χώρο, όπου είναι πιο πιθανό να υπάρχει ένα γραμμικό διαχωριστικό όριο.

Για το σκοπό της άσκησης, θα πρέπει να δημιουργηθεί ένας classifier που θα διακρίνει τα σύνολα "stars" και "circles" χρησιμοποιώντας τη μέθοδο με τα kernel. Συγκεκριμένα, δίνονται δύο μήτρες "stars" και "circles" περιέχουν τα δεδομένα ως δισδιάστατα διανύσματα. Για να ορίσουμε ένα διαχωριστικό σύνορο, αντιστοιχίζουμε την αριθμητική ετικέτα "1" στο σύνολο "stars" και την "-1" στο σύνολο "circles". Η μέθοδος με τα kernel μας επιτρέπει να υπολογίσουμε την απόσταση μεταξύ των σημείων στον υψηλότερης διάστασης χώρο, χρησιμοποιώντας μια συνάρτηση πυρήνα. Ένας δημοφιλής πυρήνας που χρησιμοποιείται συχνά είναι ο γκαουσιανός πυρήνας (Gaussian kernel), ο οποίος δίνεται από τον τύπο $K(X,Y) = e^{-\frac{\|X-Y\|^2}{h}}$.

Όπως και στα περισσότερα προβλήματα μηχανικής μάθησης, ενδιαφερόμαστε ουσιαστικά για την ελαχιστοποίηση μιας συνάρτησης κόστους. Καθότι μάλιστα έχουμε μεταβεί στον υψηλότερης διάστασης χώρο, η ελαχιστοποίηση θα γίνει ως προς φ.

$$\min_{\phi\in\mathcal{V}}\left\{\sum_{X_i\in\text{stars}}\left(1-\phi(X_i)\right)^2+\sum_{X_j\in\text{circles}}\left(1+\phi(X_j)\right)^2+\lambda\|\phi(X)\|^2\right\}$$

Όπως όμως προαναφέρθηκε, ο χώρος V που περιέχει την συνάρτηση φ μπορεί να έχει πολύ μεγάλη, ακόμα και άπειρη διάσταση, οπότε η συμβατική ελαχιστοποίηση είναι αδύνατη ή τουλάχιστον υπολογιστικά δαπανηρή. Σύμφωνα όμως με το Reprsenter Theorem, το οποίο αποδεικνύεται στις διαφάνειες του μαθήματος, κοντά στα σημεία ενδιαφέροντος μπορούμε να αντικαταστήσουμε τις $\phi(X_i)$, $\phi(X_j)$ με τις κάθετες προβολές τους στον δισδιάστατο χώρο Ω:

$$\hat{\phi}(X)=\sum_{X_i\in\text{stars}}\alpha_i\mathsf{K}(X,X_i)+\sum_{X_j\in\text{circles}}\beta_j\mathsf{K}(X,X_j).$$

Επίσης εύκολα αποδεικνύεται ότι φ(X) >= $\hat{\varphi}(X)$, οπότε και στον τελευταίο όρο (regularization term) μπορούμε να αντικαταστήσουμε την φ(X) με μια τυχαία μικρότερη ποσότητα χωρίς να επηρεαστεί το πρόβλημα βελτισοποίησης.

γ

Σύμφωνα πάλι με τις διαφάνειες του μαθήματος, για να προσδιορίσουμε τους συντελεστές α_i και β_j αρκεί να επιλύσουμε το γραμμικό σύστημα με πίνακα συντελεστών A και B το διάνυσμα με τους σταθερούς όρους.

$$A=\begin{bmatrix}\langle z_1,z_1\rangle & \dots & \langle z_1,z_n\rangle\\ \vdots & \ddots & \vdots\\ \langle z_1,z_n\rangle & \dots & \langle z_n,z_n\rangle\end{bmatrix},\quad B=\begin{bmatrix}\langle x,z_1\rangle\\ \vdots\\ \langle x,z_n\rangle\end{bmatrix}$$

Στους παραπάνω τύπους ορίζουμε ως $z_i = K(X,X_i)$, οπότε τελικά το εσωτερικό γινόμενο $\langle z_i,z_j\rangle$ τελικά μετατρέπεται σε $K\left(X_i,X_j\right)$. Θεωρούμε επίσης ότι το διάνυσμα-στήλη X περιέχει όλα τα αστέρια και έπειτα όλους τους κύκλους. Από αυτό συνεπάγεται ότι ο πίνακας των συντελεστών έχει μέγεθος 42x42. Η πρώτη του στήλη για παράδειγμα θα είναι $[K(X_1^{stars},\ X_1^{stars}) \quad \dots \quad K(X_1^{stars},\ X_{21}^{stars}) \quad K(X_1^{stars},\ X_1^{circles}) \quad \dots \quad K(X_1^{stars},\ X_{21}^{circles})]^T$

Σημειώνεται ότι ο πίνακας των συντελεστών είναι συμμετρικός, καθώς $K\left(X_i,X_j\right) = K(X_j,X_i)$. Έτσι, οι άγνωστοι του συστήματος θα είναι μια στήλη $[a_1 \quad \dots \quad a_{21} \quad b_1 \quad \dots \quad b_{21}]^T$ και το διάνυσμα των σταθερών όρων περιλαμβάνει τις τιμές y για κάθε στοιχείο, δηλαδή θα είναι 21 ‘1’ και ακολούθως 21 ‘-1’. Αποδεικνύεται επίσης ότι αντί να λύσουμε το εν λόγω σύστημα ορθότερο είναι να λυθεί το σύστημα $(A+\lambda I_{42})\begin{bmatrix}\alpha_i\\ b_j\end{bmatrix} = B$, όπου I_{42} ο μοναδιαίος πίνακας διάστασης 42, ούτως ώστε να λάβουμε υπόψιν και τον συντελεστή του όρου εξομάλυνσης (regularization) λ. Το σύστημα επιλύεται εύκολα με την βοήθεια της βιβλιοθήκης numpy.

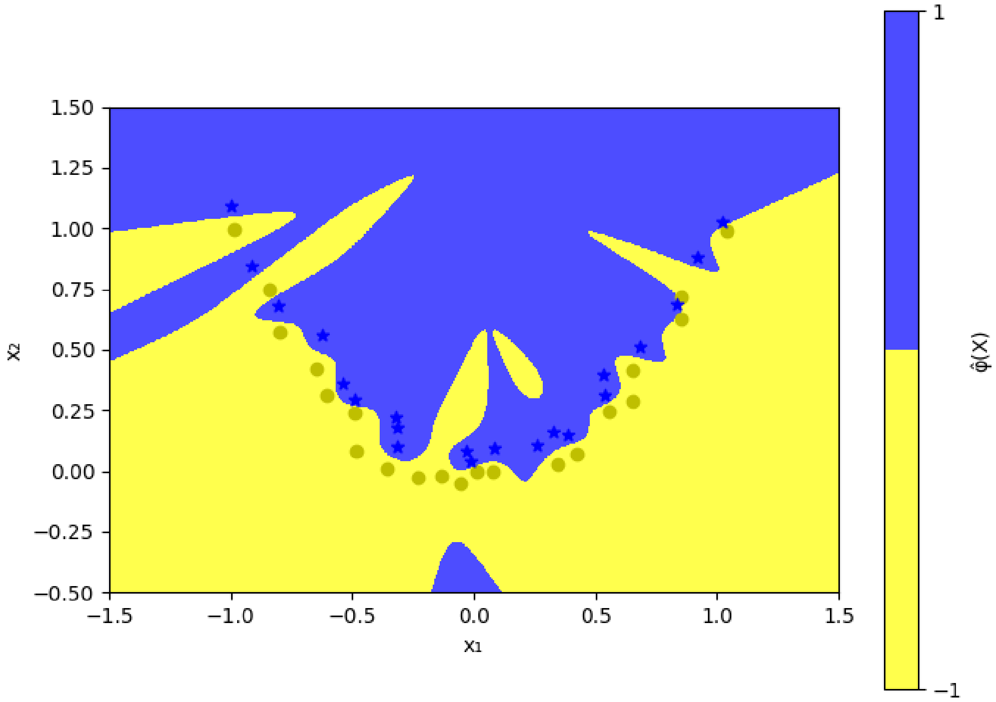
δ

Επιλύοντας το παραπάνω γραμμικό σύστημα παίρνουμε τους βέλτιστους συντελεστές a_i και b_j, ώστε η συνάρτηση $\hat{\varphi}(X)$ να έχει την τιμή 1 πάνω σε όλα τα σημεία του σετ με τα αστέρια και -1 πάνω σε όλου τους κύκλους. Έστω ότι επιθυμούμε τώρα να κατηγοριοποιήσουμε ένα άλλο, καινούριο σημείο του χώρου, έστω X_{new} σε μία από τις δύο κατηγορίες. Δεδομένου ότι η τιμή $\hat{\varphi}(X_{new})$ δεν θα έχει ούτε την τιμή 1 ούτε -1, μπορούμε να θέσουμε ως κριτήριο την σύγκριση με την μέση τιμή των κατηγοριών, δηλαδή 0. Αν $\hat{\varphi}(X_{new}) > 0$ τότε ισχυριζόμαστε ότι ανήκει στην ομάδα 1 (αστέρια), διαφορετικά στο σύνολο με τους κύκλους.

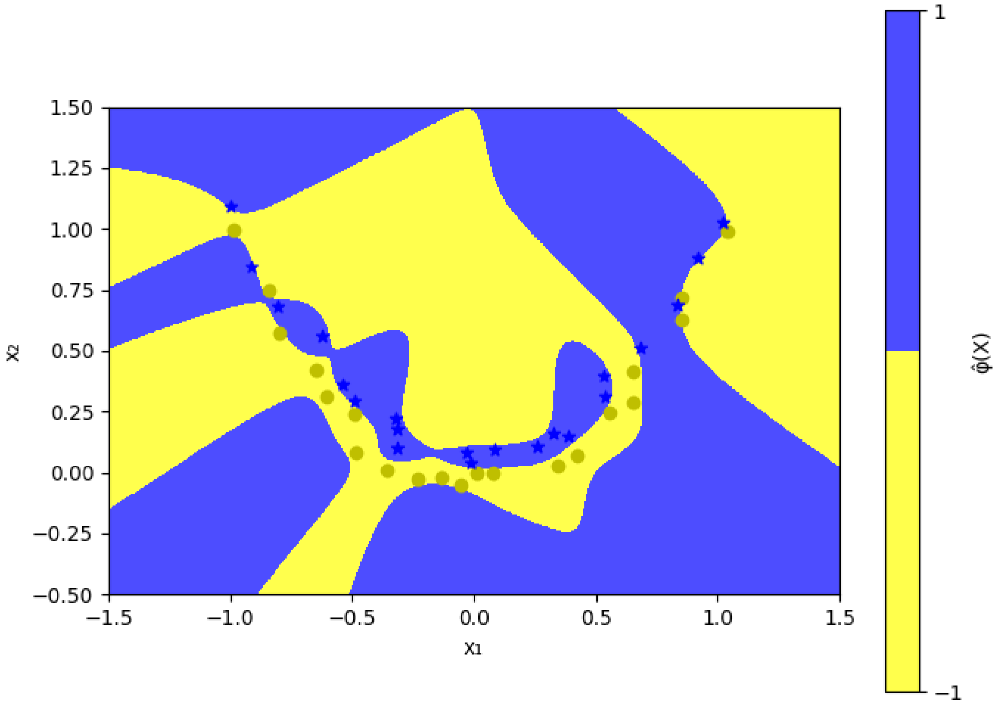
ε

Βάσει όλων των παραπάνω, ορίζουμε τα όρια του δισδιάστατου κόσμου x ∈ [-1.5, 1.5] και y ∈ [-0.5, 0.5]. Τον χώρο αυτό τον διαμερίζουμε σε 500x500 στοιχεία. Για καθένα στοιχείο του χώρου υπολογίζουμε την τιμή $\hat{\varphi}(X)$ και το χρωματίζουμε ανάλογα με την τιμή του ($\gtrless 0$). Στις αρνητικές τιμές δίνουμε κίτρινο χρώμα και στις θετικές μπλε. Τα σημεία ζωγραφίζονται αντίστοιχα. Ακολουθούν διάφορες υλοποιήσεις για διάφορες τιμές των παραμέτρων h και λ.

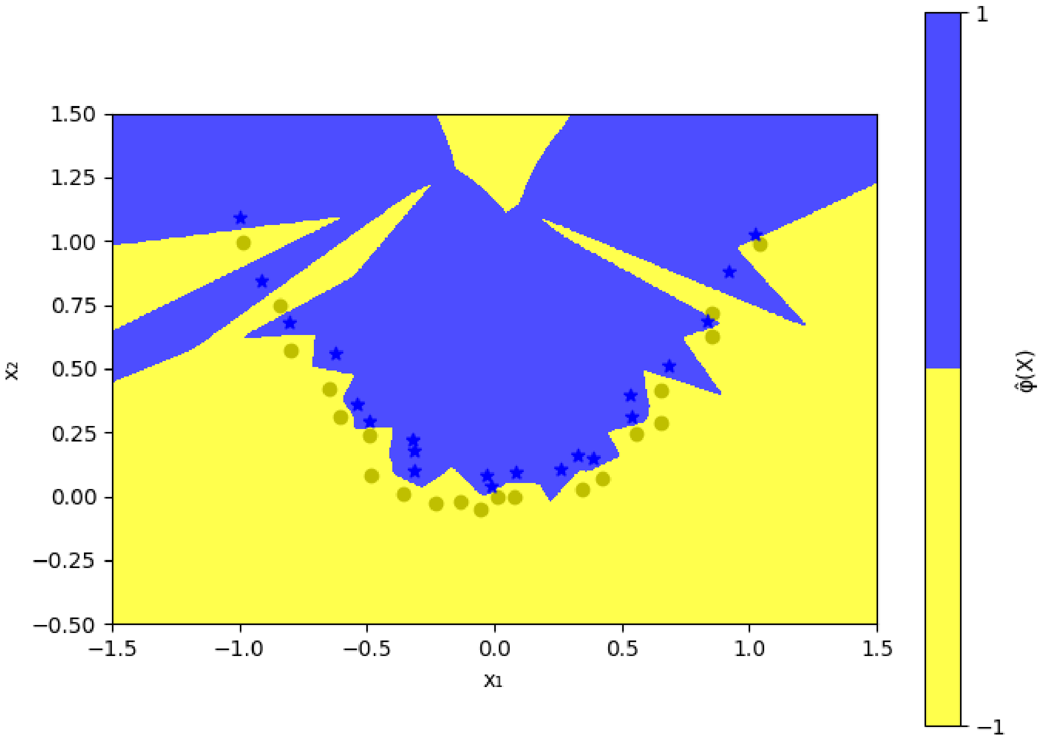
λ=0, h=0.01



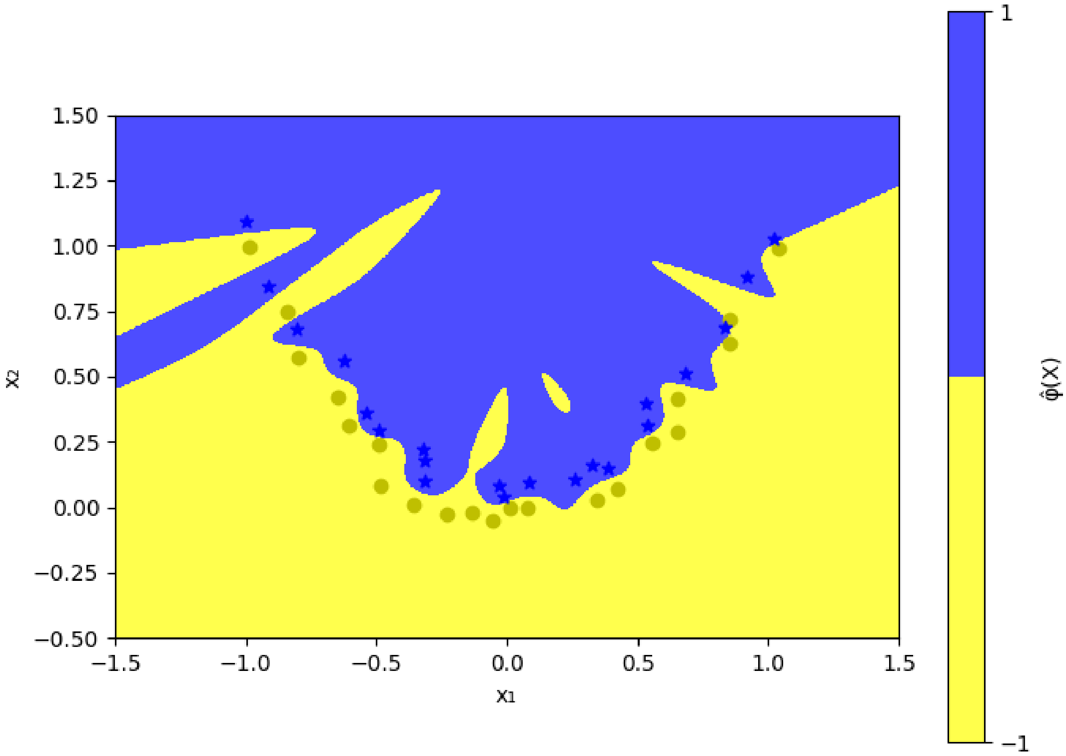
λ=0, h=0.1



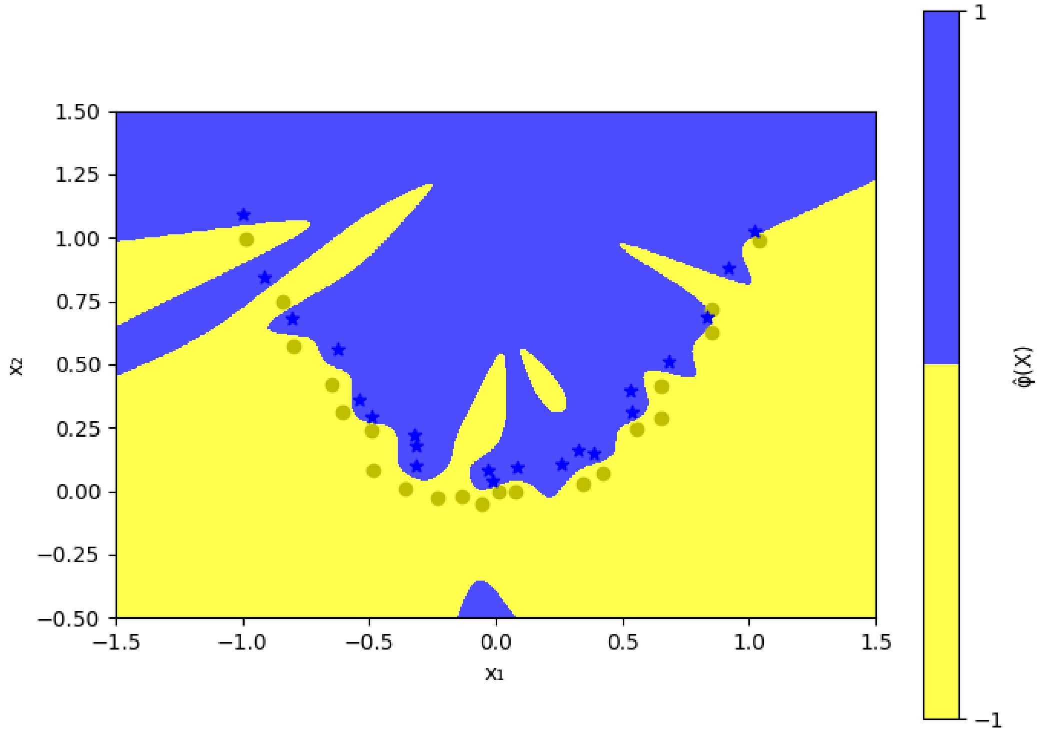
λ=0, h=0.001



λ=0.05, h=0.01



λ=0.01, h=0.01



λ=0.2, h=0.01