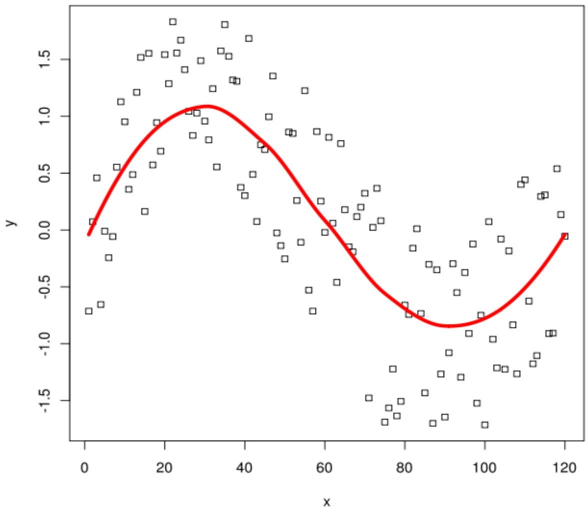
Οικονομικό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής
Μάθημα: Τεχνητή Νοημοσύνη, 2025–26
Διδάσκων: Ι. Ανδρουτσόπουλος

Ασκήσεις μελέτης της 10ης διάλεξης

10.1. Οι τελείες του σχήματος στα δεξιά παριστάνουν έναν δείγμα που προέρχεται από πληθυσμό ο οποίος ακολουθεί στην πραγματικότητα την άγνωστη συνάρτηση *y = f(x)*, της οποίας η γραφική παράσταση είναι η συνεχής καμπύλη.[1] Λόγω θορύβου κατά τις μετρήσεις της δειγματοληψίας, όμως, τα σημεία του δείγματος δεν βρίσκονται ακριβώς πάνω στη συνεχή καμπύλη. Θέλουμε να μάθουμε από το δείγμα μια συνάρτηση *y = h(x)*, που να προσεγγίζει κατά το δυνατόν περισσότερο την *f*(*x*).



(α) Εξηγήστε γιατί η χρήση γραμμικής παλινδρόμησης ελαχίστων τετραγώνων δεν θα οδηγούσε σε ικανοποιητική *h(x)*.

Απάντηση: Η γραμμική παλινδρόμηση ελαχίστων τετραγώνων μαθαίνει ευθείες γραμμές (ή γενικότερα επίπεδα ή υπερ-επίπεδα). Στη συγκεκριμένη περίπτωση η καμπύλη της y= f(x) δεν μπορεί να προσεγγιστεί καλά από μια μόνο ευθεία γραμμή.

(β) Πώς θα μπορούσαμε να μάθουμε μια πιο ικανοποιητική *h(x)* χρησιμοποιώντας τον αλγόριθμο των *k* κοντινότερων γειτόνων (ή μια παραλλαγή του); Τι θα κάναμε κατά το στάδιο της εκπαίδευσης και τι όποτε (κατόπιν) μας δίνουν ένα *x* για το οποίο πρέπει να επιστρέψουμε το *y* = *h(x)*;

Απάντηση: Κατά το στάδιο της εκπαίδευσης απλά αποθηκεύουμε όλες τις συντεταγμένες (x, y) των σημείων του δείγματος. Κατόπιν, όποτε μας δίνουν ένα νέο x' και μας ζητούν το y' = h(x'), ανακτούμε τα k σημεία του δείγματος (γείτονες) των οποίων οι τιμές x βρίσκονται πιο κοντά στο x' και επιστρέφουμε το μέσο όρο των τιμών y αυτών των k σημείων (των γειτόνων). Μπορούμε επίσης να ζυγίζουμε κατά τον υπολογισμό του μέσου όρου τις y τιμές των γειτόνων, δίνοντας σε κάθε μία y τιμή γείτονα βάρος π.χ. αντιστρόφως ανάλογο της απόστασης της x τιμής του γείτονα από το x'.

10.2. Γράψτε τους υπολογισμούς με τους οποίους προκύπτει ο κανόνας ενημέρωσης βαρών της γραμμικής παλινδρόμησης ελαχίστων τετραγώνων, όταν χρησιμοποιείται στοχαστική κατάβαση κλίσης.

Απάντηση: Ο κανόνας ενημέρωσης βαρών της γραμμικής παλινδρόμησης ελαχίστων τετραγώνων, όταν χρησιμοποιείται στοχαστική κατάβαση κλίσης, δίνεται από τον τύπο:

$$\vec{w} \leftarrow \vec{w} - \eta \cdot \nabla_{\vec{w}} E_i(\vec{w}), \text{ όπου } E_i(\vec{w}) = \frac{1}{2}\left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right]^2$$

Επίσης:

[1] Το σχήμα προέρχεται από την ιστοσελίδα http://en.wikipedia.org/wiki/Local_regression.

$$\nabla_{\vec{w}} E_i(\vec{w}) = \langle \frac{\partial E_i(\vec{w})}{\partial w_0}, \frac{\partial E_i(\vec{w})}{\partial w_1}, \cdots, \frac{\partial E_i(\vec{w})}{\partial w_l}, \cdots, \frac{\partial E_i(\vec{w})}{\partial w_n} \rangle$$

Για $l \in \{0, \dots, n\}$:

$$\begin{aligned}
\frac{\partial E_i(\vec{w})}{\partial w_l} &= 2 \cdot \frac{1}{2} \cdot \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot \frac{\partial\left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right]}{\partial w_l} \\
&= \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot \frac{\partial\left[w_0 x_0 + \cdots + w_l x_l + \cdots + w_n x_n\right]}{\partial w_l} \\
&= \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot x_l^{(i)}
\end{aligned}$$

Άρα,

$$\nabla_{\vec{w}} E_i(\vec{w}) = \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot \langle x_0^{(i)}, \cdots, x_l^{(i)}, \cdots x_n^{(i)} \rangle =$$

$$= \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot \vec{x}^{(i)}$$

και ο τύπος ενημέρωσης βαρών γίνεται:

$$\vec{w} \leftarrow \vec{w} - \eta \cdot \left[f_{\vec{w}}\left(\vec{x}^{(i)}\right) - y^{(i)}\right] \cdot \vec{x}^{(i)}$$

10.3. Καταγράψαμε μεγάλο αριθμό παρτίδων (συνολικά M παρτίδες) μεταξύ παικτών Othello. Σε κάθε παρτίδα καταγράψαμε όλα τα στιγμιότυπα της σκακιέρας (τις θέσεις όλων των πιονιών), ένα στιγμιότυπο αμέσως μετά από κάθε κίνηση παίκτη. Τα στιγμιότυπα όλων των παρτίδων συνολικά ήταν K, ενώ τα *διαφορετικά* στιγμιότυπα (μετρώντας τα ίδια στιγμιότυπα μόνο μία φορά το καθένα) όλων των παρτίδων συνολικά ήταν L. Για κάθε στιγμιότυπο, καταγράψαμε τις τιμές των ιδιοτήτων $f_1(n)$, $f_2(n)$, $f_3(n)$, όπου τώρα n είναι το στιγμιότυπο· εμφανίστηκαν συνολικά N *διαφορετικοί* συνδυασμοί τιμών των $f_1(n)$, $f_2(n)$, $f_3(n)$ στις M παρτίδες. Υποθέστε ότι οι ιδιότητες είναι τρεις χειρωνακτικά κατασκευασμένες ευρετικές συναρτήσεις του συγκεκριμένου παιχνιδιού. (Σε επόμενες ενότητες θα δούμε πώς θα μπορούσαν να παράγονται και αυτές αυτόματα, μέσω μοντέλων βαθιάς μάθησης, και να είναι περισσότερες.) Για κάθε στιγμιότυπο, καταγράψαμε ακόμη αν η παρτίδα στην οποία εμφανίστηκε το στιγμιότυπο έληξε υπέρ του μαύρου παίκτη, υπέρ του άσπρου ή ισόπαλη.

Εξηγήστε πώς θα μπορούσαμε να εκμεταλλευτούμε τα καταγεγραμμένα στοιχεία των M παρτίδων, για να μάθουμε με γραμμική παλινδρόμηση ελαχίστων τετραγώνων τις καλύτερες δυνατές τιμές των w_0, w_1, w_2, w_3, ώστε για κάθε κόμβο n του δέντρου αναζήτησης του αλγορίθμου MiniMax, η $h(n) = w_1 \cdot f_1(n) + w_2 \cdot f_2(n) + w_3 \cdot f_3(n) + w_0$ να επιστρέφει μια κατά το δυνατόν ακριβέστερη εκτίμηση του *αναμενόμενου* οφέλους με το οποίο θα τελειώσει το παιχνίδι, αν βρεθεί στην κατάσταση του κόμβου n. Θεωρήστε ότι το όφελος ενός παιχνιδιού είναι 1 όταν κερδίζει ο Max (μαύρος), –1 όταν κερδίζει ο Min (άσπρος) και 0 όταν το παιχνίδι λήγει ισόπαλο.

(α) Πόσα παραδείγματα (διανύσματα ιδιοτήτων) εκπαίδευσης θα είχαμε στη γραμμική παλινδρόμηση και ποιες θα ήταν οι μεταβλητές της συνάρτησης που θα μάθαινε η γραμμική παλινδρόμηση; Φροντίστε να μην υπάρχουν ασυνεπή παραδείγματα εκπαίδευσης.

Απάντηση: Θα είχαμε N παραδείγματα εκπαίδευσης, ένα για κάθε διαφορετικό συνδυασμό τιμών των $f_1(n)$, $f_2(n)$, $f_3(n)$ που παρατηρήθηκε στις M παρτίδες. Οι μεταβλητές θα ήταν οι ιδιότητες $f_1(n)$, $f_2(n)$, $f_3(n)$.

(β) Πώς ακριβώς (δώστε μαθηματικό τύπο) θα υπολογίζαμε για κάθε παράδειγμα εκπαίδευσης την «ορθή» (επιθυμητή) απόκριση της συνάρτησης που θα θέλαμε να μάθει η γραμμική παλινδρόμηση;

Απάντηση: Κάθε παράδειγμα εκπαίδευσης παριστάνει έναν από τους N διαφορετικούς συνδυασμούς τιμών των τριών ιδιοτήτων που παρατηρήθηκε στις M παρτίδες. Έστω σ ένας από αυτούς τους διαφορετικούς συνδυασμούς και σ_1, σ_2, ..., σ_k τα καταγεγραμμένα στιγμιότυπα (όχι αναγκαστικά διαφορετικά μεταξύ τους) που είχαν το συγκεκριμένο συνδυασμό τιμών του σ στις $f_1(n)$, $f_2(n)$, $f_3(n)$. Η επιθυμητή απόκριση για το σ θα ήταν:

$$\frac{1}{k}\cdot\sum_{i=1}^{k}u(\sigma_i)$$

όπου $u(\sigma_i)$ είναι το όφελος με το οποίο τελείωσε η παρτίδα στην οποία καταγράφηκε το στιγμιότυπο σ_i.

(γ) Ποια θα ήταν η σχέση της συνάρτησης $f_{\vec{w}}(\vec{x}) = \vec{w}\cdot\vec{x}$ που θα μάθαινε η γραμμική παλινδρόμηση με την *h(n)* που θα χρησιμοποιούσαμε τελικά;

Απάντηση: Θα χρησιμοποιούσαμε ως h(n) την $f_{\vec{w}}(\vec{x})$, με $\vec{x} = \langle f_1(n), f_2(n), f_3(n)\rangle$:

$$f_{\vec{w}}(\vec{x}) = f_{w_0,w_1,w_2,w_3}(\langle f_1(n), f_2(n), f_3(n)\rangle) = w_1\cdot f_1(n) + w_2\cdot f_2(n) + w_3\cdot f_3(n) + w_0 = h(n)$$

(δ) Πώς θα μπορούσαμε να συλλέξουμε στιγμιότυπα του παιχνιδιού από έναν μεγάλο αριθμό (M) παρτίδων, χωρίς να βάλουμε ανθρώπους να παίξουν μεταξύ τους;

Απάντηση: Μπορούμε να αναπτύξουμε αρχικά ένα πρόγραμμα Othello το οποίο να χρησιμοποιεί τον αλγόριθμο MiniMax (ιδανικά με πριόνισμα α-β) και έναν γραμμικό συνδυασμό των ιδιοτήτων (ευρετικών) $w_1 \cdot f_1(n) + w_2 \cdot f_2(n) + w_3 \cdot f_3(n) + w_0$ με χειρωνακτικά επιλεγμένα βάρη (π.χ. ίσα βάρη για όλες τις ευρετικές και $w_0 = 0$). Μπορούμε να βάλουμε το πρόγραμμα να παίξει με αντίπαλο τον εαυτό του πολλές παρτίδες και να συλλέξουμε τα στιγμιότυπα των παρτίδων. Αλλά έτσι θα προκύπτει πάντα η ίδια παρτίδα, γιατί οι δύο παίκτες θα κάνουν πάντα τις ίδιες κινήσεις.

Για να αντιμετωπίσουμε αυτό το πρόβλημα και να έχουμε μεγαλύτερη ποικιλία παρτίδων μπορούμε να κάνουμε το εξής. Όποτε είναι η σειρά του Max να παίξει, αν το δέντρο MiniMax που κατασκευάζει δείχνει ότι οι διαθέσιμες κινήσεις του $a_1, a_2, \ldots, a_r$ έχουν τιμές MiniMax $v_1, v_2, \ldots, v_r$ αντίστοιχα, μπορούμε να μετατρέπουμε τις τιμές αυτές σε κατανομή πιθανότητας $p_1, p_2, \ldots, p_r$. Αυτό μπορούμε να το πετύχουμε περνώντας τις $v_1, v_2, \ldots, v_r$ από softmax (βλ. διάλεξη 11), δηλαδή $p_i = \frac{exp\,(v_i/T)}{\sum_{j=1}^{r} exp\,(v_r/T)}$, όπου το T (θερμοκρασία) ρυθμίζει πόσο ομοιόμορφη (για ψηλό T) είναι η κατανομή πιθανοτήτων. Έτσι ο Max μπορεί να επιλέγει κίνηση με πιθανοτικό τρόπο, σύμφωνα με την κατανομή $p_1, p_2, \ldots, p_r$. Αντίστοιχα, όποτε είναι η σειρά του Min να παίξει, αν το δέντρο MiniMax που κατασκευάζει για να επιλέξει την κίνησή του δείχνει ότι οι διαθέσιμες κινήσεις του $a_1, a_2, \ldots, a_r$ έχουν τιμές MiniMax $v_1, v_2, \ldots, v_r$, μπορούμε να μετατρέπουμε τις τιμές $-v_1, -v_2, \ldots, -v_r$ σε κατανομή πιθανότητας $p_1, p_2, \ldots, p_r$, περνώντας τις από softmax, ώστε και ο Min να επιλέγει κίνηση με πιθανοτικό τρόπο.

Αφού συλλέξουμε πολλά στιγμιότυπα παρτίδων με τον παραπάνω τρόπο, μπορούμε να μάθουμε τα w_0, w_1, w_2, w_3 με γραμμική παλινδρόμηση, όπως στο σκέλος (β), και να χρησιμοποιήσουμε ως ευρετική την $f_{\vec{w}}(\vec{x})$, ώστε το πρόγραμμα να παίζει πλέον καλύτερα. Θα μπορούσαμε κατόπιν να ξαναβάλουμε το πρόγραμμα (με τη νέα ευρετική) να παίζει πολλές παρτίδες με αντίπαλο τον εαυτό του, να επιλέξουμε πρόσθετα στιγμιότυπα παρτίδων (ελπίζοντας ότι αυτά είναι πιο διδακτικά, αφού τώρα το πρόγραμμα παίζει καλύτερα) και να μάθουμε από αυτά μια (ελπίζουμε) ακόμα καλύτερη $f_{\vec{w}}(\vec{x})$.

Μπορούμε επίσης να χρησιμοποιήσουμε Αναζήτηση Δέντρου Monte Carlo (Monte Carlo Tree Search, MCTS) και Ενισχυτική Μάθηση (Reinforcement Learning, RL). Όσοι ενδιαφέρεστε δείτε προαιρετικά την ενότητα 5.4 (MCTS) και το κεφάλαιο 22 (RL) των Russel & Norvig ($4^{η}$ έκδοση), καθώς και την ενότητα 18.14 των Βλαχάβα κ.ά.