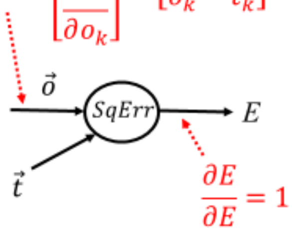
Οικονομικό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής
Μάθημα: Αλληλεπίδραση Ανθρώπου – Υπολογιστή, 2023–24
Διδάσκων: Ι. Ανδρουτσόπουλος

Ασκήσεις μελέτης της ενότητας B5 (επεξεργασία φυσικής γλώσσας και εικόνας με πολυεπίπεδα Perceptron)

1. Επιβεβαιώστε τον υπολογισμό του $\frac{\partial E}{\partial \vec{o}^{(2)}}$ στο γράφο υπολογισμού της διαφάνειας 28. Μια που μας ενδιαφέρει η πύλη *SqErr* να μπορεί να χρησιμοποιηθεί και σε άλλους γράφους υπολογισμού, θέστε χάριν γενικότητος $\vec{o}^{(2)} = \vec{o}$ και $k_2 = k$.

$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \cdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix} = \begin{bmatrix} o_1 - t_1 \\ \cdots \\ o_k - t_k \end{bmatrix} = \vec{o} - \vec{t}$$



Απάντηση: Το διάνυσμα κλίσης (gradient) που χρειάζεται να υπολογίσουμε είναι:

$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \cdots \\ \frac{\partial E}{\partial o_i} \\ \cdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix}$$

Ας εξετάσουμε ξεχωριστά κάθε μία παράγωγο $\frac{\partial E}{\partial o_i}$ (κάθε στοιχείο του διανύσματος κλίσης):

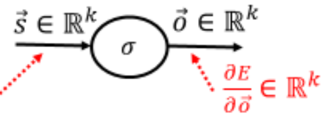
$$\frac{\partial E}{\partial o_i} = \frac{\partial}{\partial o_i} \sum_{j=1}^{k} \frac{1}{2} \left(t_j - o_j\right)^2 = \frac{\partial}{\partial o_i} \frac{1}{2} (t_i - o_i)^2 = \frac{1}{2} \cdot 2 \cdot (t_i - o_i) \cdot \frac{\partial}{\partial o_i} (t_i - o_i)$$
$$= (t_i - o_i) \cdot (-1) = (o_i - t_i)$$

Επομένως:

$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \cdots \\ \frac{\partial E}{\partial o_i} \\ \cdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix} = \begin{bmatrix} o_1 - t_1 \\ \cdots \\ o_i - t_i \\ \cdots \\ o_k - t_k \end{bmatrix} = \vec{o} - \vec{t}$$

Σημείωση: Δεν χρειάζεται να υπολογίσουμε το $\frac{\partial E}{\partial \vec{t}}$, γιατί δεν ενημερώνουμε το $\vec{t}$ (το σωστό διάνυσμα εξόδου για το συγκεκριμένο παράδειγμα εισόδου).

2. Δείξτε ότι για μια σιγμοειδή πύλη $\sigma(\vec{s}) = \vec{o}$, το $\frac{\partial E}{\partial \vec{s}}$ μπορεί να υπολογιστεί όπως παρακάτω, όπου J ο Ιακωβιανός πίνακας.[1]



$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial \sigma(s_1)}{\partial s_1} & \frac{\partial \sigma(s_2)}{\partial s_1} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \sigma(s_1)}{\partial s_i} & \frac{\partial \sigma(s_2)}{\partial s_i} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_i} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \sigma(s_1)}{\partial s_k} & \frac{\partial \sigma(s_2)}{\partial s_k} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_k} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \vdots \\ \frac{\partial E}{\partial o_i} \\ \vdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix} = J^T \frac{\partial E}{\partial \vec{o}}$$

$$= \begin{bmatrix} \sigma(s_1)(1-\sigma(s_1)) & 0 & \cdots & 0 \\ 0 & \sigma(s_2)(1-\sigma(s_2)) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma(s_k)(1-\sigma(s_k)) \end{bmatrix} \frac{\partial E}{\partial \vec{o}}$$

Απάντηση: Το διάνυσμα κλίσης που χρειάζεται να υπολογίσουμε είναι:

$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix}$$

Ας εξετάσουμε ξεχωριστά κάθε μία μερική παράγωγο $\frac{\partial E}{\partial s_i}$ του διανύσματος της κλίσης. Σύμφωνα με τον κανόνα αλυσίδας των παραγώγων:

$$\frac{\partial E}{\partial s_i} = \sum_{j=1}^{k} \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial s_i}$$

Κάθε s_i, όμως, επηρεάζει μόνο το $o_i = \sigma(s_i)$. Δεν επηρεάζει κανένα άλλο $o_j = \sigma(s_j)$, για $j \neq i$. Επομένως, $\frac{\partial o_j}{\partial s_i} = 0$ για $j \neq i$ και άρα:

$$\frac{\partial E}{\partial s_i} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial s_i} = \frac{\partial E}{\partial o_i} \frac{\partial \sigma(s_i)}{\partial s_i} = \frac{\partial E}{\partial o_i} \sigma(s_i)(1-\sigma(s_i))$$

όπου χρησιμοποιήσαμε την ιδιότητα της σιγμοειδούς πως $\frac{d\sigma(x)}{dx} = \sigma(x)(1-\sigma(x))$.

[1] See https://en.wikipedia.org/wiki/Jacobian_matrix_and_determinant.

Επομένως:

$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \frac{\partial \sigma(s_1)}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial o_i} \frac{\partial \sigma(s_i)}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial o_k} \frac{\partial \sigma(s_k)}{\partial s_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \sigma(s_1)\big(1-\sigma(s_1)\big) \\ \vdots \\ \frac{\partial E}{\partial o_i} \sigma(s_i)\big(1-\sigma(s_i)\big) \\ \vdots \\ \frac{\partial E}{\partial o_k} \sigma(s_k)\big(1-\sigma(s_k)\big) \end{bmatrix}$$

Η τελευταία εξίσωση μπορεί να γραφτεί και ως:

$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial \sigma(s_1)}{\partial s_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial \sigma(s_2)}{\partial s_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{\partial \sigma(s_k)}{\partial s_k} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \frac{\partial E}{\partial o_2} \\ \vdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix} =$$

$$= \begin{bmatrix} \sigma(s_1)\big(1-\sigma(s_1)\big) & 0 & \cdots & 0 \\ 0 & \sigma(s_2)\big(1-\sigma(s_2)\big) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma(s_k)\big(1-\sigma(s_k)\big) \end{bmatrix} \frac{\partial E}{\partial \vec{o}}$$

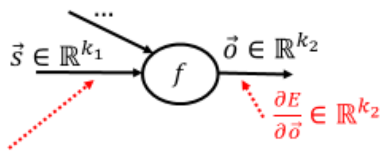
Πιο γενικά, μπορεί να γραφτεί ως:

$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial \sigma(s_1)}{\partial s_1} & \frac{\partial \sigma(s_2)}{\partial s_1} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_1} \\ \frac{\partial \sigma(s_1)}{\partial s_2} & \frac{\partial \sigma(s_2)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \sigma(s_1)}{\partial s_k} & \frac{\partial \sigma(s_2)}{\partial s_k} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_k} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \frac{\partial E}{\partial o_2} \\ \vdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix} = J^T \frac{\partial E}{\partial \vec{o}}$$

όπου J ο Ιακωβιανός πίνακας:

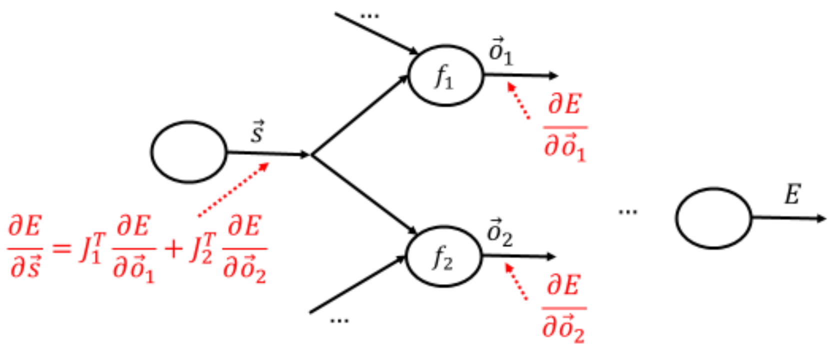
$$J = \begin{bmatrix} \frac{\partial \sigma(s_1)}{\partial s_1} & \frac{\partial \sigma(s_1)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_1)}{\partial s_k} \\ \frac{\partial \sigma(s_2)}{\partial s_1} & \frac{\partial \sigma(s_2)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_2)}{\partial s_k} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \sigma(s_k)}{\partial s_1} & \frac{\partial \sigma(s_k)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_k)}{\partial s_k} \end{bmatrix}$$

Αυτός είναι ένας γενικότερος κανόνας. Για έναν κόμβο που υπολογίζει το $f(\vec{s}, \dots) = \vec{o}$, μπορούμε να υπολογίσουμε το $\frac{\partial E}{\partial \vec{s}}$ ως εξής, υπό την προϋπόθεση ότι το $\vec{s}$ δίνεται ως είσοδος μόνο στον κόμβο f:

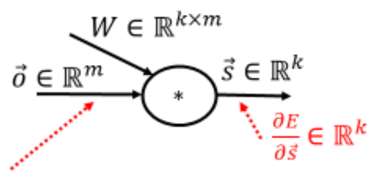


$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial s_{k_1}} \end{bmatrix} = \begin{bmatrix} \frac{\partial o_1}{\partial s_1} & \frac{\partial o_2}{\partial s_1} & \cdots & \frac{\partial o_{k_2}}{\partial s_1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial o_1}{\partial s_i} & \frac{\partial o_2}{\partial s_i} & \cdots & \frac{\partial o_{k_2}}{\partial s_i} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial o_1}{\partial s_{k_1}} & \frac{\partial o_2}{\partial s_{k_1}} & \cdots & \frac{\partial o_{k_2}}{\partial s_{k_1}} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \vdots \\ \frac{\partial E}{\partial o_i} \\ \vdots \\ \frac{\partial E}{\partial o_{k_2}} \end{bmatrix} = J^T \frac{\partial E}{\partial \vec{o}}$$

Αν το $\vec{s}$ δίνεται ως είσοδος σε δύο (ή περισσότερους) κόμβους f_1, f_2, πρέπει να αθροίσουμε τα διανύσματα κλίσης $\frac{\partial E}{\partial \vec{s}}$ που λαμβάνουμε από τους f_1, f_2:



3. Δείξτε ότι σε έναν κόμβο πολλαπλασιασμού πίνακα-διανύσματος $W\vec{o} = \vec{s}$, το $\frac{\partial E}{\partial \vec{o}}$ μπορεί να υπολογιστεί ως εξής:



$$\frac{\partial E}{\partial \vec{o}} = J_{\vec{o}}^T \frac{\partial E}{\partial \vec{s}} = W^T \frac{\partial E}{\partial \vec{s}}$$

Απάντηση:

$$\vec{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \cdots \\ s_k \end{bmatrix} = W\vec{o} = \begin{bmatrix} w_{1,1} & w_{2,1} & \cdots & w_{m,1} \\ w_{1,2} & w_{2,2} & \cdots & w_{m,2} \\ w_{1,3} & w_{2,3} & \cdots & w_{m,3} \\ \cdots & \cdots & \cdots & \cdots \\ w_{1,k} & w_{2,k} & \cdots & w_{m,k} \end{bmatrix} \begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \cdots \\ o_m \end{bmatrix} = \begin{bmatrix} w_{1,1}o_1 + w_{2,1}o_2 + \cdots + w_{m,1}o_m \\ w_{1,2}o_1 + w_{2,2}o_2 + \cdots + w_{m,2}o_m \\ w_{1,3}o_1 + w_{2,3}o_2 + \cdots + w_{m,3}o_m \\ \cdots \\ w_{1,k}o_1 + w_{2,k}o_2 + \cdots + w_{m,k}o_m \end{bmatrix}$$

Το διάνυσμα κλίσης που χρειάζεται να υπολογίσουμε είναι το:

$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \vdots \\ \frac{\partial E}{\partial o_i} \\ \vdots \\ \frac{\partial E}{\partial o_m} \end{bmatrix}$$

Ας εξετάσουμε ξεχωριστά κάθε μία μερική παράγωγο $\frac{\partial E}{\partial o_i}$ *(κάθε στοιχείο του διανύσματος κλίσης). Σύμφωνα με τον κανόνα αλυσίδας των παραγώγων:*

$$\frac{\partial E}{\partial o_i} = \sum_{j=1}^{k} \frac{\partial E}{\partial s_j} \frac{\partial s_j}{\partial o_i}$$

Σύμφωνα με τις παραπάνω εξισώσεις για το $\vec{s} = W\vec{o}$*:*

$$s_j = w_{1,j} o_1 + w_{2,j} o_2 + \cdots + w_{i,j} o_i + \cdots + w_{m,j} o_m$$

Επομένως:

$$\frac{\partial s_j}{\partial o_i} = w_{i,j}$$

και:

$$\frac{\partial E}{\partial o_i} = \sum_{j=1}^{k} \frac{\partial E}{\partial s_j} \frac{\partial s_j}{\partial o_i} = \sum_{j=1}^{k} \frac{\partial E}{\partial s_j} w_{i,j}$$

το οποίο μπορεί να γραφτεί και ως:

$$\frac{\partial E}{\partial o_i} = \begin{bmatrix} \frac{\partial s_1}{\partial o_i} & \frac{\partial s_2}{\partial o_i} & \cdots & \frac{\partial s_k}{\partial o_i} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \frac{\partial E}{\partial s_2} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix} = \begin{bmatrix} w_{i,1} & w_{i,2} & \cdots & w_{i,k} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \frac{\partial E}{\partial s_2} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix}$$

Επομένως, για το συνολικό διάνυσμα κλίσης:

$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \frac{\partial E}{\partial o_2} \\ \vdots \\ \frac{\partial E}{\partial o_i} \\ \vdots \\ \frac{\partial E}{\partial o_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial s_1}{\partial o_1} & \frac{\partial s_2}{\partial o_1} & \cdots & \frac{\partial s_k}{\partial o_1} \\ \frac{\partial s_1}{\partial o_2} & \frac{\partial s_2}{\partial o_2} & \cdots & \frac{\partial s_k}{\partial o_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial s_1}{\partial o_i} & \frac{\partial s_2}{\partial o_i} & \cdots & \frac{\partial s_k}{\partial o_i} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial s_1}{\partial o_m} & \frac{\partial s_2}{\partial o_m} & \cdots & \frac{\partial s_k}{\partial o_m} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \frac{\partial E}{\partial s_2} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,k} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ w_{i,1} & w_{i,2} & \cdots & w_{i,k} \\ \vdots & \vdots & \vdots & \vdots \\ w_{m,1} & w_{m,2} & \cdots & w_{m,k} \end{bmatrix} \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \frac{\partial E}{\partial s_2} \\ \vdots \\ \frac{\partial E}{\partial s_k} \end{bmatrix} =$$

$$J_{\vec{o}}^{T}\frac{\partial E}{\partial \vec{s}} = W^{T}\frac{\partial E}{\partial \vec{s}}$$

Σημείωση: Προτιμούμε να χρησιμοποιούμε πράξεις πινάκων, που μπορούν να υπολογιστούν αποδοτικά με βελτιστοποιημένους αλγορίθμους πράξεων πινάκων και GPUs, αντί να χρησιμοποιούμε δικούς μας βρόχους (π.χ. με for της Python) που υπολογίζουν (συνήθως πολύ πιο αργά) σε κάθε επανάληψη ένα μεμονωμένο στοιχείο των πινάκων.

Όσοι ενδιαφέρεστε ιδιαίτερα για την Τεχνητή Νοημοσύνη και ιδιαίτερα τη Μηχανική Μάθηση, καλό είναι να μελετήσετε (προαιρετικά) και τις υπόλοιπες ασκήσεις μελέτης των διαλέξεων 19 και 20 του μαθήματος «Τεχνητή Νοημοσύνη».