Slides created by Sofia Eleftheriou

Ασκήσεις μελέτης Β5
Lab 7

Human-Computer Interaction, AUEB
Εαρινό εξάμηνο 2023-2024

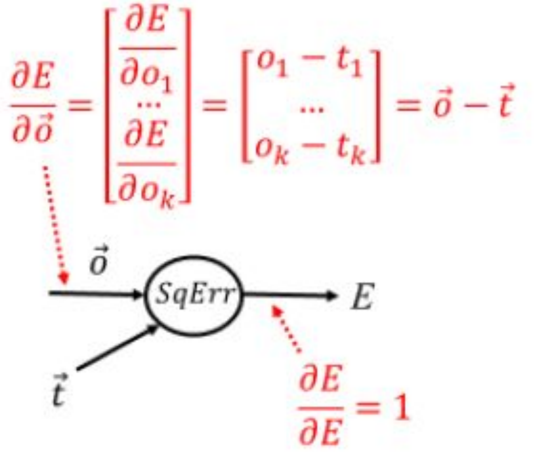
Lab Assistant: Sofia Eleftheriou

Άσκηση B5.1.

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

Απάντηση:

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



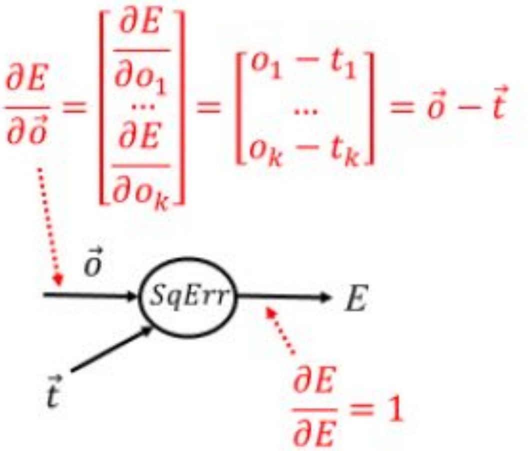
$$\frac{\partial E}{\partial \vec{o}} = \begin{bmatrix} \frac{\partial E}{\partial o_1} \\ \cdots \\ \frac{\partial E}{\partial o_i} \\ \cdots \\ \frac{\partial E}{\partial o_k} \end{bmatrix}$$

Ας εξετάσουμε ξεχωριστά κάθε μία παράγωγο ∂E/∂o_i (κάθε στοιχείο του διανύσματος κλίσης):

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

Απάντηση -συνέχεια:
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 ∂E/∂$\vec{t}$, γιατί δεν ενημερώνουμε το $\vec{t}$ (το σωστό διάνυσμα εξόδου για το συγκεκριμένο παράδειγμα εισόδου).

Στη Διανυσματική Ανάλυση, ο Πίνακας Τζακόμπι ή Ιακωβιανή Μήτρα είναι ο πίνακας όλων των παραγώγων 1ης τάξης ενός διανύσματος ή μιας βαθμωτής συνάρτησης σε σχέση με ένα άλλο διάνυσμα.

Έστω ότι η f: $\mathbb{R}^n \rightarrow \mathbb{R}^m$ είναι μία συνάρτηση από τον ευκλείδειο χώρο n προς τον ευκλείδειο χώρο m. Μια τέτοια συνάρτηση δίνεται από πραγματικές στοιχεία - συναρτήσεις, $y_1(x_1, \ldots, x_n), \ldots, y_m(x_1, \ldots, x_n)$. Οι μερικές παράγωγοι όλων αυτών των συναρτήσεων (εάν υπάρχουν) μπορούν να αναπαρασταθούν σε έναν m επί n πίνακα, τον πίνακα Τζακόμπι J της f, ως εξής:

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Άσκηση B5.2.

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

Απάντηση:

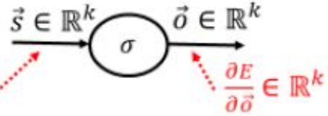
$(o_1, o_2, \ldots, o_k) = (\sigma(s_1), \sigma(s_2), \ldots, \sigma(s_k))$

Δηλαδή έχουμε μια συνάρτηση f: $\mathbb{R}^k \rightarrow \mathbb{R}^k$ με

$$f(\vec{s}) = \begin{bmatrix} \sigma(s_1) \\ \ldots \\ \sigma(s_k) \end{bmatrix} = \begin{bmatrix} o_1 \\ \ldots \\ o_k \end{bmatrix}$$

Το διάνυσμα κλίσης που χρειάζεται να υπολογίσουμε είναι:

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



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

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

Απάντηση – συνέχεια:

Ας εξετάσουμε ξεχωριστά κάθε μία μερική παράγωγο $\partial E/\partial s_i$ του διανύσματος της κλίσης.

Το σφάλμα Ε είναι μια συνάρτηση κ μεταβλητών αφού

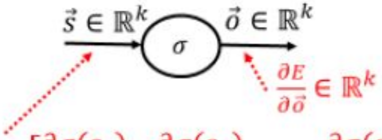
$$E = g(o_1, o_2, \dots, o_k)$$

όπου τα o_i είναι ανεξάρτητα μεταξύ τους και κάθε o_i είναι μια συνάρτηση των s_i που είναι επίσης ανεξάρτητα.

Σύμφωνα με τον κανόνα αλυσίδας των παραγώγων:

$$\frac{\partial E}{\partial s_i} = \frac{\partial E}{\partial o_1}\frac{\partial o_1}{\partial s_i} + \frac{\partial E}{\partial o_2}\frac{\partial o_2}{\partial s_i} + \dots + \frac{\partial E}{\partial o_k}\frac{\partial o_k}{\partial s_i} = \sum_{j=1}^{k}\frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial s_i}$$

$$\frac{\partial o_j}{\partial s_i} = 0$$



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

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

Απάντηση – συνέχεια:

Επομένως θα είναι

$$\frac{\partial E}{\partial s_i} = \frac{\partial E}{\partial o_i}\frac{\partial o_i}{\partial s_i}$$

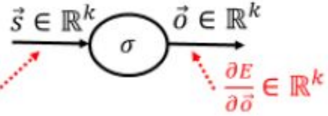
Το $\partial E/\partial o_i$ έχει υπολογιστεί στο προηγούμενο βήμα οπότε εστιάζουμε στο $\partial o_i/\partial s_i$

Για τη σιγμοειδή συνάρτηση $\sigma(x) = \frac{1}{1+e^{-x}}$ ισχύει ότι

$$\frac{d\sigma(x)}{dx} = \sigma(x)\big(1 - \sigma(x)\big)$$

Άρα

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



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

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

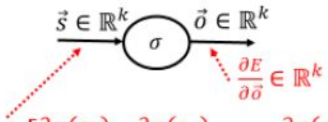
Απάντηση – συνέχεια:
Επομένως

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

Η τελευταία εξίσωση μπορεί να γραφτεί και ως:

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

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



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

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

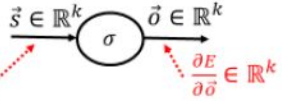
Απάντηση – συνέχεια:

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:

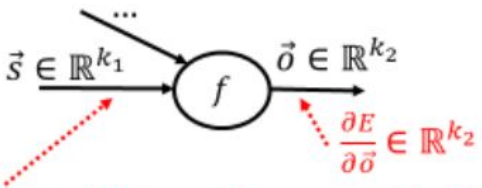
$$J = \begin{bmatrix} \frac{\partial \sigma(s_1)}{\partial s_1} & \frac{\partial \sigma(s_1)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_1)}{\partial s_k} \\ \frac{\partial \sigma(s_2)}{\partial s_1} & \frac{\partial \sigma(s_2)}{\partial s_2} & \cdots & \frac{\partial \sigma(s_2)}{\partial s_k}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



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

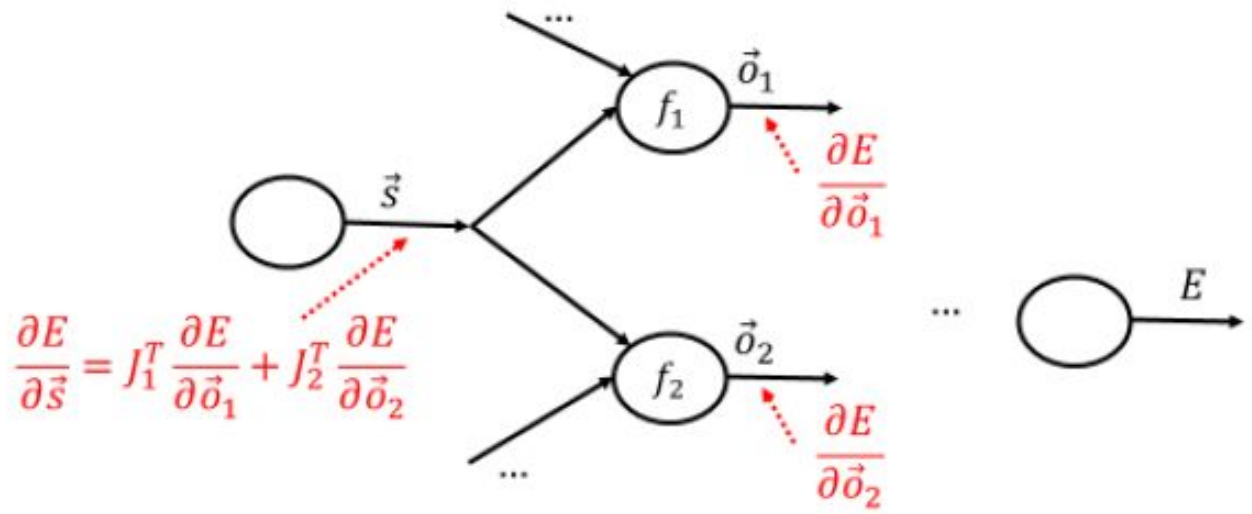
$$= \begin{bmatrix} \sigma(s_1)(1-\sigma(s_1)) & 0 & \cdots & 0 \\ 0 & \sigma(s_2)(1-\sigma(s_2)) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma(s_k)(1-\sigma(s_k)) \end{bmatrix} \frac{\partial E}{\partial \vec{o}}$$

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



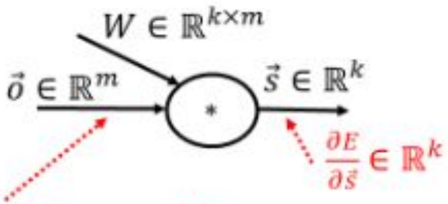
$$\frac{\partial E}{\partial \vec{s}} = \begin{bmatrix} \frac{\partial E}{\partial s_1} \\ \vdots \\ \frac{\partial E}{\partial s_i} \\ \vdots \\ \frac{\partial E}{\partial s_{k_1}} \end{bmatrix} = \begin{bmatrix} \frac{\partial o_1}{\partial s_1} & \frac{\partial o_2}{\partial s_1} & \cdots & \frac{\partial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

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:



Άσκηση B5.3

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



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

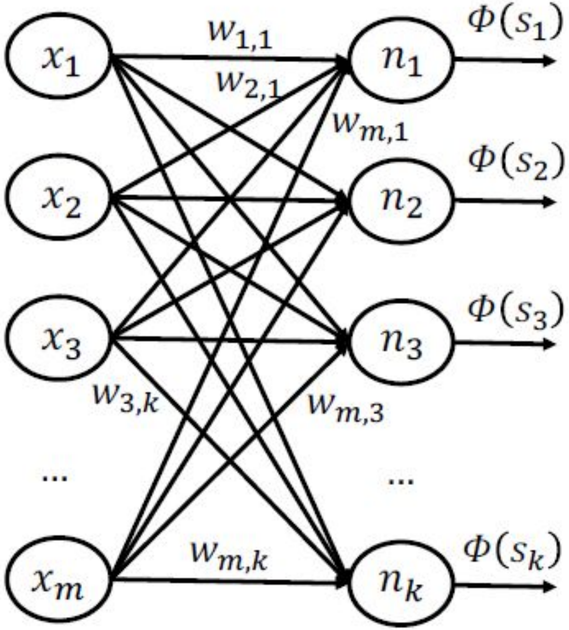
Απάντηση:

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

Το διάνυσμα κλίσης που χρειάζεται να υπολογίσουμε είναι το:

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

Πιο συμπαγής συμβολισμός



$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \dots \\ s_k \end{bmatrix} = \begin{bmatrix} w_{1,1}x_1 + w_{2,1}x_2 + \cdots + w_{m,1}x_m \\ w_{1,2}x_1 + w_{2,2}x_2 + \cdots + w_{m,2}x_m \\ w_{1,3}x_1 + w_{2,3}x_2 + \cdots + w_{m,3}x_m \\ \dots \\ w_{1,k}x_1 + w_{2,k}x_2 + \cdots + w_{m,k}x_m \end{bmatrix}$$

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \dots \\ s_k \end{bmatrix} = \begin{bmatrix} w_{1,1} & w_{2,1} & \dots & w_{m,1} \\ w_{1,2} & w_{2,2} & \dots & w_{m,2} \\ w_{1,3} & w_{2,3} & \dots & w_{m,3} \\ \dots & \dots & \dots & \dots \\ w_{1,k} & w_{2,k} & \dots & w_{m,k} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_m \end{bmatrix}$$

$$\vec{s} = W\vec{x}$$

Μαθαίνουμε τον ***W*** με **ανάστροφη μετάδοση**.

$$\vec{o} = \begin{bmatrix} o_1 \\ o_2 \\ o_3 \\ \dots \\ o_k \end{bmatrix} = \begin{bmatrix} \Phi(s_1) \\ \Phi(s_2) \\ \Phi(s_3) \\ \dots \\ \Phi(s_\kappa) \end{bmatrix} = \Phi(\vec{s}) = \Phi(W\vec{x})$$

Ας εξετάσουμε ξεχωριστά κάθε μία μερική παράγωγο $\partial E/\partial o_i$ (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.