Θέματα 06/28/2021:

Ερώτηση 5:

SVD: Έστω $X \in \mathbb{R}^{m \times n}$ τότε οι $X^T \cdot X \in \mathbb{R}^{n \times n}$ και $X \cdot X^T \in \mathbb{R}^{m \times m}$ είναι συμμετρικοί, $(X^T X)^T = X^T (X^T)^T = X^T X$

$(AB)^T = B^T A^T$

όμοιως και $X \cdot X^T$ είναι και αυτός συμμετρικός.

$A = A^T \Rightarrow A$ συμμετρικός.

· Κάθε συμμετρικός πίνακας έχει μη αρνητικές ιδιοτιμές και τα ιδιοδιανύσματα είναι ανά δύο κάθετα (ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές).

Παραγοντοποίηση SVD: $\boxed{(X = U \cdot \Sigma \cdot V^T)}$

1ο Βήμα: Βρίσκουμε ιδιοτιμές και ιδιοδιανύσματα του $X^T X$

2ο Βήμα: Οι στήλες του V είναι τα ιδιοδιανύσματα $X^T \cdot X$.

Ο Σ έχει την μορφή $\Sigma = \begin{pmatrix} \sigma_1 & & & \emptyset \\ & \ddots & & \\ & & \sigma_k & \\ \emptyset & & & 0 \ddots 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$, όπου $X \in \mathbb{R}^{m \times n}$

και $\sigma_1 \geq \ldots \geq \sigma_k$ με σ_i να αντιστοιχούν στις μη μηδενικές ρίζες των ιδιοτιμών ιδιοτιμές του $X^T \cdot X$.

3ο Βήμα: Βρίσκω ιδιοτιμές και ιδιοδιανύσματα του $X \cdot X^T$ και οι στήλες του U είναι ιδιοδιανύσματα του $X \cdot X^T$.

1) Από την $X = U \cdot \Sigma V^T$ και οι U, V είναι αντιστρέψιμοι $\Rightarrow$ τότε το $\text{rank}(\Sigma)$ είναι το ίδιο με το $\text{rank}(X)$.

2) Οι ιδιάζουσες τιμές είναι μοναδικές το SVD όχι (οι U, V δεν είναι μοναδικοί)

~~4~~) Οι ιδιάζουσες τιμές και οι ιδιοτιμές του X είναι ίδιες $\rightarrow$ Λάθος!

Άλλο οι ιδιοτιμές του X και άλλο οι ρίζες των ιδιοτιμών του $X^T X$

~~4~~3) Δεν έχει νόημα να μιλάμε για ιδιοδιανύσματα/ιδιοτιμές για μη τετραγωνικούς πίνακες.

Ερώτηση 6: Έστω $v \in \mathbb{R}^n$ και $\{v_1, \ldots, v_n\}$ γραμμικά ανεξάρτητα με

$\boxed{\langle v, v_i \rangle = 0 \quad \forall i = 1, \ldots n}$

Αφού τα $\{v_1, \dots, v_n\}$ είναι n το πλήθος γραμμικά ανεξάρτητα στον $\mathbb{R}^n$ θα είναι βάση του $\mathbb{R}^n$. Άρα το v θα γράφεται ως γραμμικός συνδυασμός στοιχείων της βάσης, δηλαδή

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n \quad , \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Τότε $\forall i = 1, \dots, n$ έχουμε

$$\langle v, v \rangle = \langle v, \lambda_1 v_1 + \dots + \lambda_n v_n \rangle$$

$$= \lambda_1 \cancel{\langle v, v_1 \rangle}^{=0} + \dots + \lambda_n \cancel{\langle v, v_n \rangle}^{=0} = 0$$

Όμως $\langle v, v \rangle = \|v\|_2^2 = 0 \Rightarrow \boxed{v = 0}$.

Ερώτηση 7:

$$h_n = f(h_{n-1}; \theta) \quad , \quad n \geq 1.$$

με συνάρτηση μετάβασης $f(x;\theta) = \tanh(\theta x)$ με $|\theta| < 1$.

Υπολογίστε την τιμή της μόνιμης κρυφής κατάστασης h του συστήματος σε έναν μακρινό ορίζοντα.

$$h_2 = f(h_1; \theta) \quad (1)$$

$$h_3 = f(h_2; \theta) \overset{(1)}{=} f(f(h_1; \theta); \theta)$$

άρα $\quad h_{n+1} = f^{(n)}(h_1; \theta)$

$$f^{(n)} = \underbrace{f \circ f \dots \circ f}_{n}$$

Τότε $|h_{n+1}| = |f^{(n)}(h_1; \theta)|$

$$= |\tanh \theta h_{n-1}| \simeq |\theta| |h_{n-1}|$$

$\tanh x \simeq x$ κοντά στο 0.

$$\simeq |\theta|^2 \cdot |h_{n-2}|$$

$$\simeq \dots |\theta|^{n-1} |h_1| \xrightarrow{n \to \infty} 0$$

Ερώτηση 8: $2^3 = 8$.

$(i, j, k) \quad , \quad i, j, k \in \{0, 1\}$.

Ερώτηση 9:

$$\underset{\text{input}}{r \times c} \rightsquigarrow \underset{\text{conv layer}}{r^{(1)} \times c^{(1)} \times f^{(1)}} \rightsquigarrow \underbrace{r^{(2)} \times c^{(2)} \times f^{(2)}}_{\text{conv layer}} \rightsquigarrow \underbrace{N \text{ neurons}}_{\text{FC layer}} \xRightarrow{\text{softmax}}$$

Πού εφαρμόζουμε dropout και γιατί;

<u>Απάντηση</u>: Εφαρμόζουμε dropout στα δύο Conv Layers ως τεχνική regularization (μείωση overfitting) και αντίστοιχα μπορούμε να εφαρμόσουμε και στο FFN Layer. Δεν μπορούμε να εφαρμόσουμε στο Layer της softmax γιατί δεν θέλω να χάσω το output. Στο input layer παίζεται...

<u>Ερώτηση 10</u>: <u>Q-Learning</u>:

<u>Ερώτηση 11</u>: Ποια μέθοδος ομαλοποίησης οδηγεί σε πιο αραιές αναπαραστάσεις;

Αραιές αναπαραστάσεις πετυχαίνουμε μέσω τοποθέτησης όρων ποινής στις **συναρτήσεις ενεργοποίησης** των νευρώνων.

<u>Ερώτηση 12</u>:

$$Q(x) = \frac{1}{2} x^T A x - bx$$

$$\frac{\partial}{\partial x} Q(x) = \frac{1}{2} \cdot (x^T A + Ax) - b$$

$$x^{n+1} = x_n - \eta \frac{\partial}{\partial x} Q(x) \Rightarrow x_{n+1} = x_n - \eta \cdot \left(\frac{1}{2} (x^T A + Ax) - b \right).$$

Η ερώτηση έχει νόημα εάν θεωρήσουμε ότι:

- $Action(t) = State(t+1)$
- Reward $= -1 \Rightarrow$ αδύνατη μετάβαση
- Δεν έχει νόημα το «κατώφλι στην κατάσταση 0»
- υπάρχει κάποιο learning rate $0 < \eta \leq 1$

Τότε:
$$Q(2,3) \leftarrow (1-\eta)\,Q(2,3) + \eta\left(R(2,3) + \gamma \cdot \max_i \{Q(3,i)\}\right)$$
$$\leftarrow (1-\eta)\cdot 64 + \eta\,(0 + 1 \cdot 80)$$
$$\leftarrow (1-\eta)\cdot 64 + 80\eta$$
$$\leftarrow 64 + 16\eta$$

Όμως $0 < \eta \leq 1 \Rightarrow 64 < 64 + 16\eta \leq 80 \Rightarrow 64 < Q(2,3) \leq 80$

Η μόνη δυνατή επιλογή είναι, λοιπόν, η F9