Připomenutí:

Víme: Pro f symetrickou bilineární formu na VP V nad T ("$\neq \mathbb{Z}_2$") existuje báze $\mathcal{B}$ tž. matice f vůči $\mathcal{B}$ je diagonální.

tzv. polární báze

Důsledek (Sylvestrův zákon setrvačnosti kvadratických forem)
Pro každou kvadratickou formu g na VP konečné dim. nad $\mathbb{R}$ existuje báze, vůči níž má g diagonální matici pouze s $0, 1, -1$. Navíc počet 1 a počet -1 je pro každou takovou bázi stejný. tzv. signatura

Příklad: nechť $g(x) = x^T A x$, pro $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 3 & -3 \end{pmatrix}$

Najdeme rozklad (diagonalizace)

$$A = \begin{pmatrix} 1 & & \\ 2 & -1 & \\ & 1 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & & \\ & -3 & \\ & & 0 \end{pmatrix}}_{D} \begin{pmatrix} 1 & 2 & \\ & -1 & 1 \\ & & 1 \end{pmatrix}$$

Pak pro

$$S = \begin{pmatrix} 1 & & \\ & \sqrt{3} & \\ & & 1 \end{pmatrix}, \quad D' = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} \text{ máme } D = S^T D' S.$$

Důsledek: v $\mathbb{R}^2$ existuje 6 typů kvadratických forem:

- klasifikace podle jejich diagonální matice s $0, 1, -1$:

"parabola3D"	pringles	otočená "parabola3D"	parabola 2D		rovina
$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & \\ & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & \\ & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$
$x_1^2 + x_2^2$	$x_1^2 - x_2^2$	$-x_1^2 - x_2^2$	x_1^2	$-x_1^2$	0

$\underbrace{x_1^2 + x_2^2 \quad x_1^2 - x_2^2 \quad -x_1^2 - x_2^2}_{\text{paraboloid}}$

Např. ↑
- satelitní anténa
- zrcadlo světla
- olympijský oheň

v $\mathbb{R}^1$ jen 3 typy:

parabola		
(1)	(-1)	(0)
x_1^2	$-x_1^2$	0

OHLÉDNUTÍ

Zimní semestr:

1. Soustavy lin. rovnic a Gauss. eliminace
2. Vektorové prostory a lineární zobrazení
3. Matice a operace s nimi

   klíčový pojem: lineární kombinace

Letní semestr:

4. Determinanty (1, 3)
5. Skalární součin ve VP (1, 2, 3) ortogonalita $R(A)$ a $\mathrm{Ker}(A)$ ... QR rozklad
6. Vlastní čísla a vlastní vektory (1, 2, 3, 4)
7. Pozitivně definitní matice (3, 4, 5, 6)
8. Bilineární a kvadratické formy (2, 5, 6, 7)

Rozklady matic: QR, Choleského, Sylvestr

Důležité: práce s maticemi – násobení – ERÚ – Gauss.

K učení na zkoušku:

- nechte si dost času! (týden)
- připomeňte si LA 1 – navazujeme!
- nepřeskakujte důkazy – pomáhají pochopit, o co jde → snáze si zapamatujete znění vět
- přemýšlejte, ptejte se proč?
- hodně pište (nejen číst a poslouchat (koukat))
- bavte se spolu
- hýbejte se!

Anketa!

GPT a LINEÁRNÍ ALGEBRA

Generative Pre-trained Transformer

hra: „přidej slovo", vytvoř příběh. Např:

Honza šel do sklepa a našel tam mrtvou myš ....
1 2 3 4 5 6 7 8 9

V ... množina všech slov (tokenů), $n = |V|$ GPT3 ~ 50 000

tzv. slovník

auto|dráha, moto|technika, ...

(10 000 slov - plynulá mluva)

Úkol (hráč hry „přidej smysluplné slovo"):

vstup: posloupnost slov $w_1, w_2, \dots, w_k \in V$

výstup: navazující slovo $w_{k+1} \in V$

Jak reprezentovat slova?

- posloupnost písmen (pro řadu jiných aplikací OK)

  nevýhody: samotná písmeno nenese skoro žádnou informaci o významu

- vektor délky n, $e_i = (0, \dots, 1, \dots 0)$ reprezentuje i-té slovo ze slovníku V

  i-tá složka ~ očíslování slov

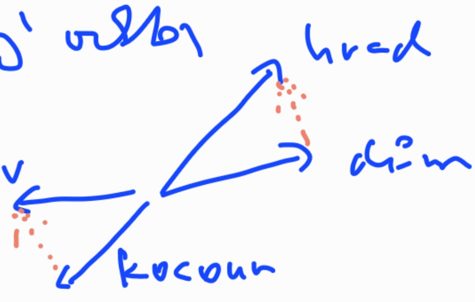
  nevýhody: není úsporné, opět skoro žádná informace o významu slov

- vektor délky $d < n$ GPT3 ~ 13 tis.

  přání: podobná slova — podobné vektory

Kde vektory vzít?

příklad:



Matice $W_e \in \mathbb{R}^{d \times n}$ ← předučená

- tzv. vnořovací matice

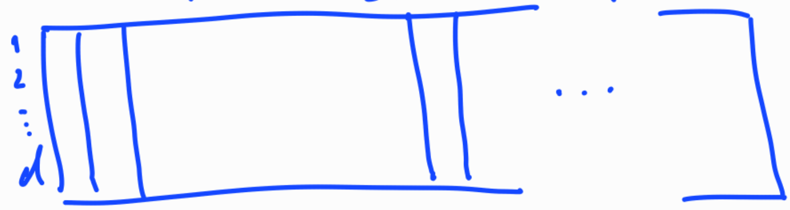
reprezentuje zobrazení $e : V \to \mathbb{R}^d$, $e(w) = W_e \cdot e_i$, (embedding)

kde i je pořadí slova w

W_e:

i-tý sloupec W_e je vektorová reprezentace i-tého slova ze slovníku V

Budeme potřebovat také něco jako inverzní matici/zobrazení, co nám k vektoru $x \in \mathbb{R}^d$ dá pstní distribuci na množině všech slov ze slovníku V

$W_u^T =$ 

$W_u \cdot x \in \mathbb{R}^n$

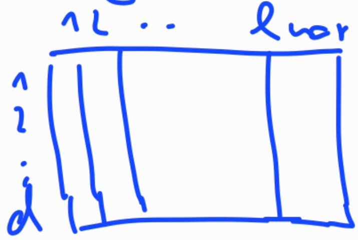
normal$(W_u x)$

též ji předučíme

Jak reprezentovat slova ve větě / delším textu?

- přidání informace o pořadí slova (ve větě/textu)
- pracujeme s omezenou délkou úseku – ℓ_{max} (GPT3 ~ 1–4 tis., GPT4 8 tis., GPT4 Turbo 128 tis.)

reprezentuje zobrazení $p: \{1, \dots, \ell_{max}\} \to \mathbb{R}^d$

$W_p =$ 

také předučíme

Pro vstup – posloupnost slov $w_1, \dots, w_k$ – spočítáme počáteční vektorovou reprezentaci $x_1, \dots, x_k \in \mathbb{R}^d$ zohledňující pořadí slov:

$$x_i = e(w_i) + p(i)$$

~ maticové násobení a sčítání

Jak rozlišit různé významy téhož slova?

Zajíc se chytil do oka.

Jan Žižka přišel o oko.

Idea: vektorovou reprezentaci x_i slova w_i upravit tak, aby zohledňovala kontext – předešlá slova $w_1, \dots, w_{i-1}$ resp. jejich vektorovou reprezentaci $x_1, \dots, x_{i-1}$

Zajišťuje to tzv. mechanismus pozornosti (attention)

Potřebujeme nastavit:

- jak moc, které ze slov $w_1, \dots, w_i$ ovlivňuje x_i?

podle toho, jak moc spolu slova souvisí –
- nesouvisející slova se neovlivňují
- související slova upřesňují význam

Jak poznat, zi spolu slova w a w' souvisí?
podle skalárního součinu x a x'!

~ nesouvisející slova mají kolmé vektory

• jakým způsobem, které ze slov $w_1, \dots, w_i$ ovlivňuje x_i?
 ... směrem

upřesňujeme reprezentaci slova w_k - vektor x_k

$x_1 \quad x_2 \quad \dots \quad x_k$ ← skalární součiny

x_k: $\langle x_k, x_1\rangle \; \langle x_k, x_2\rangle \; \dots \; \langle x_k, x_k\rangle$ ← posloupnost čísel od $-\infty$ k ∞

normalizujeme tak, aby
- všechna čísla byla ≥ 0
- součet $= 1$

} tzv. pravděpodobn. rozdělení na $x_1, \dots, x_k$

$\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_k$

přejdeme od x_k k $x_k' = \sum_{i=1}^{k} \alpha_i \cdot x_i$

K čemu to je? pouzijeme odvozovací matici W_u:

$y = W_u \cdot x_k' \in \mathbb{R}^n$

$z = \mathrm{Normal}(y) \in \mathbb{R}^n$ normalizace: $\sum_{i=1}^{n} z_i = 1$, $\forall i$: $z_i \geq 0$

tj. z je pravděpodobnostní rozdělení na V

$\Rightarrow$ slovo w_{k+1} vybereme podle něj.

Model pracuje se třemi dalšími maticemi:

$W_k, W_q, W_v \in \mathbb{R}^{d \times n}$, které pro každé

slovo ze slovníku uchovávají tři další vektory

⋮

Nejznámější se ručně opakuje.