
Investigating the impact of ChatGPT on teaching at LS3.

Maximilian Esser

03.07.2023

Faculty of Mathematics, LS 3

TU Dortmund

Overview

Introduction

Solving a Numeric I exam

Seminar presentation

Summary

Introduction

ChatGPT

- GPT stands for Generative Pre-Trained Transformer.
- Multiple Versions developed by OpenAI, for us relevant GPT3.5 and GPT4. [Vas+17]
- ChatGPT is a webchat-interface to a fine-tuned version of GPT3.5/4 .[23]
- Generates (sensical) answers to almost any question.
- There are other generative large language models... which we ignore for this presentation.

Introduction

ChatGPT

- GPT stands for Generative Pre-Trained Transformer.
- Multiple Versions developed by OpenAI, for us relevant GPT3.5 and GPT4. [Vas+17]
- ChatGPT is a webchat-interface to a fine-tuned version of GPT3.5/4 .[23]
- Generates (sensible) answers to almost any question.
- There are other generative large language models... which we ignore for this presentation.
- Let's see a small example...

Introduction

Why should we care?

- LLM's can generate a large quantity of text in a very short time.
- Based on the quality it could introduce major shortcuts to certain tasks.
- This could help to work more efficient...
- But could undermine effort and accomplishments of students.
- About 50% of students already use chatGPT, MINT students having the biggest proportion[GMM23].

Introduction

A few notes

- GPT works with a wide number of languages, there is no real difference between German and English.
- The generated text is inherently random.
- While the overall direction is always very similar, it deviates in details.
- The shown examples are the "best" answers.

Solving a Numeric I exam

A short question

- Argumentieren Sie, für welches $\omega \in \mathbb{R}$ der Exaktheitsgrad der Quadraturformel

$$Q(f) = \frac{1+\omega}{8} f\left(\frac{1}{6}\right) + \frac{1}{4} f\left(\frac{3}{6}\right) + \frac{3}{8} f\left(\frac{7-\omega}{6}\right).$$

auf dem Intervall $I = [0, 1]$ am größten ist.

- Due to unicode characters in PDF's, simple copy&paste is not the best idea.
- So the input and output are adjusted at a few points without changing the meaning.

Solving a Numeric I exam

With GPT3.5

Prompt

Argumentieren Sie, für welches $w \in \mathbb{R}$ der Exaktheitsgrad der Quadraturformel $Q(f) = (1+w)/8 f(1/6) + 1/4 f(3/6) + 3/8 f((7-w)/6)$ auf dem Intervall $I = [0, 1]$ am größten ist

Answer

Um den Exaktheitsgrad der gegebenen Quadraturformel auf dem Intervall $I = [0, 1]$ zu maximieren, müssen wir die Gewichtungen optimieren. Die Quadraturformel hat die Form:

$$Q(f) = (1+w)/8 * f(1/6) + 1/4 * f(3/6) + 3/8 * f((7-w)/6)$$

Der Exaktheitsgrad einer Quadraturformel hängt von der Anzahl der Stützstellen ab, die sie verwendet. In diesem Fall haben wir drei Stützstellen: $1/6$, $3/6$ und $(7-w)/6$.

Um den Exaktheitsgrad zu maximieren, müssen wir die Quadraturformel so konstruieren, dass sie mindestens alle Polynome vom Grad 2 exakt integriert. Das bedeutet, dass die Quadraturformel für lineare Funktionen exakt sein muss und auch für quadratische Funktionen [...]

Answer

Wir können die Quadraturformel verwenden, um die Integrale der Funktionen $f(x) = 1$, $f(x) = x$ und $f(x) = x^2$ zu berechnen und die Ergebnisse mit den exakten Werten vergleichen.

Für $f(x) = 1$: $Q(1) = (1+w)/8 * 1 + 1/4 * 1 + 3/8 * 1 = (1+w)/8 + 1/4 + 3/8 = (1+w+2+3)/8 = (6+w)/8$

Das Integral von $f(x) = 1$ auf dem Intervall $[0, 1]$ ist genau 1. Um den Exaktheitsgrad zu maximieren, sollte also $(6+w)/8 = 1$ sein, was zu $w = 2$ führt [...]

- That's correct.

Solving a Numeric I exam

Answer

$$\begin{aligned}\text{Für } f(x) = x: Q(x) &= (1+w)/8 * (1/6) + 1/4 * (3/6) + 3/8 * ((7-w)/6) \\ &= (1+w)/48 + 1/8 + 3(7-w)/48 = (1+w+8+21-3w)/48 = (30-2w)/48 = \\ &= (15-w)/24\end{aligned}$$

Das Integral von $f(x) = x$ auf dem Intervall $[0, 1]$ ist genau $1/2$. Um den Exaktheitsgrad zu maximieren, sollte also $(15-w)/24 = 1/2$ sein, was zu $w = 6$ führt [...]

Solving a Numeric I exam

Answer

$$\begin{aligned} \text{Für } f(x) = x: Q(x) &= (1+w)/8 * (1/6) + 1/4 * (3/6) + 3/8 * ((7-w)/6) \\ &= (1+w)/48 + 1/8 + 3(7-w)/48 = (1+w+8+21-3w)/48 = (30-2w)/48 = \\ &= (15-w)/24 \end{aligned}$$

Das Integral von $f(x) = x$ auf dem Intervall $[0, 1]$ ist genau $1/2$. Um den Exaktheitsgrad zu maximieren, sollte also $(15-w)/24 = 1/2$ sein, was zu $w = 6$ führt [...]

- A small calculation error...

Solving a Numeric I exam

Answer

$$\begin{aligned} \text{Für } f(x) = x: Q(x) &= (1+w)/8 * (1/6) + 1/4 * (3/6) + 3/8 * ((7-w)/6) \\ &= (1+w)/48 + 1/8 + 3(7-w)/48 = (1+w+8+21-3w)/48 = (30-2w)/48 = \\ &= (15-w)/24 \end{aligned}$$

Das Integral von $f(x) = x$ auf dem Intervall $[0, 1]$ ist genau $1/2$. Um den Exaktheitsgrad zu maximieren, sollte also $(15-w)/24 = 1/2$ sein, was zu $w = 6$ führt [...]

- A small calculation error...
- A bigger one...

Answer

$$\begin{aligned} \text{Für } f(x) = x^2: Q(x^2) &= (1+w)/8*(1/6)^2 + 1/4*(3/6)^2 + 3/8*((7-w)/6)^2 \\ &= (1+w)/288 + 1/48 + 3(7-w)^2/288 = (1+w+6+3(49-14w+w^2))/288 \\ &= (56 - 13w + w^2)/288 \end{aligned}$$

Das Integral von $f(x) = x^2$ auf dem Intervall $[0, 1]$ ist genau $1/3$. Um den Exaktheitsgrad zu maximieren, sollte also $(56 - 13w + w^2)/288 = 1/3$ sein, was zu keiner ganzzahligen Lösung für w führt.

Daher ist der größte Exaktheitsgrad der Quadraturformel gegeben, wenn $w = 6$ ist.

- And it finishes with an incorrect deduction.

Solving a Numeric I exam

Another small exercise

- Begründen Sie, warum die Lagrange-Polynome $L_0(x), \dots, L_n(x)$ zu den Knoten x_0, \dots, x_n die Eigenschaft

$$\sum_{i=0}^n L_i(x) \equiv 1$$

erfüllen.

- This time more of a prove... or at least the combination of two key properties of Lagrange-Polynomials.

Solving a Numeric I exam

With GPT3.5

Prompt

Begründen Sie, warum die Lagrange-Polynome $L_0(x), \dots, L_n(x)$ zu den Knoten x_0, \dots, x_n die Eigenschaft $\sum_{i=0}^n L_i(x) = 1$ erfüllen.

Answer

[...] Die Lagrange-Polynome sind definiert als:

$$L_i(x) = \prod_{j=0, j \neq i}^{n-1} (x - x_j) / (x_i - x_j)$$

Hier sind x_0, x_1, \dots, x_n die gegebenen Knotenpunkte, und i ist der Index, der den aktuellen Lagrange-Polynomterm angibt. Die Idee hinter den Lagrange-Polynomen besteht darin, dass jedes Polynom $L_i(x)$ genau an einem Knotenpunkt x_i den Wert 1 annimmt und an den übrigen Knotenpunkten den Wert 0. [...]

- Until here, it is fine.

Answer

Da für jeden Lagrange-Polynomterm nur ein einziger Knotenpunkt den Wert 1 hat und alle anderen den Wert 0 haben, ergibt sich die Eigenschaft:

$$\sum_{i=0}^n L_i(x) = L_0(x) + L_1(x) + \dots + L_n(x) = 1$$

Dies bedeutet, dass die Lagrange-Polynome eine sogenannte "Partitions der Eins" bilden, was eine wichtige Eigenschaft für die Interpolation ist. Sie gewährleistet, dass die interpolierende Funktion an den Knotenpunkten den entsprechenden Funktionswerten entspricht und somit eine genaue Approximation der gegebenen Funktion darstellt.

- This is not an explanation.
- Let's see if a better model can help here...

New contender: GPT 4

Answer

[...]

Es ist bekannt, dass jedes $L_i(x)$ die Eigenschaft hat, dass $L_i(x_j) = 0$ für $i \neq j$ und $L_i(x_i) = 1$.

Wenn wir die Summe aller $L_i(x)$ von $i = 0$ bis n berechnen, dann ist diese Summe gleich zu:

$$\text{Summe}_{\{i = 0\} \wedge n} L_i(x) = L_0(x) + L_1(x) + \dots + L_n(x).$$

Da jedes $L_i(x)$ den Wert 1 an der Stelle x_i und den Wert 0 an allen anderen Knoten x_j ($j \neq i$) hat, kann jedes $L_i(x)$ als eine Art "Indikatorfunktion" für den Knoten x_i verstanden werden. Also, wenn $x = x_i$ ist, dann ist $L_i(x) = 1$ und alle anderen $L_j(x)$ ($j \neq i$) sind 0, und die Summe der Lagrange-Polynome ist dann 1 [...]

Answer

Auf diese Weise gilt für jedes x , das gleich einem der Knoten x_i ist, dass die Summe der Lagrange-Polynome 1 ergibt. Außerdem ist das Polynom $\sum_{i=0}^n L_i(x)$ ein Polynom vom Grad höchstens n , und zwei verschiedene Polynome vom Grad höchstens n können nur dann an $n+1$ Punkten übereinstimmen, wenn sie identisch sind. Daher gilt, dass $\sum_{i=0}^n L_i(x) = 1$ für alle x , und nicht nur für $x = x_i$.

- It uses the interpolation property, which is indeed the key for the "proof".

Solving a Numeric I exam

A multistep exercise

In dieser Aufgabe soll iterativ die Stärke $d \in (0, R)$ einer Kugelschale mit Volumen $V > 0$ und gegebenem Außenradius $R > 0$ bestimmt werden, wobei angenommen werden kann, dass V durch das Volumen einer Vollkugel mit Radius R beschränkt ist, also

$$V < \frac{4\pi}{3} R^3.$$

1. Bestimmen Sie eine Funktion f , welche die Stärke d als Nullstelle besitzt und deren Auswertung ohne die Berechnung von Wurzeln auskommt. Formulieren Sie die Funktion so, dass diese im Fall kleiner Stärken d gut konditioniert ist und keine Auslöschung auftreten kann.
2. Argumentieren Sie, dass f auf dem Intervall $[0, R]$ eine Nullstelle besitzt und begründen Sie mit Hilfe von Abschätzungen aus der Vorlesung, warum die Startwerte immer besser gewählt werden müssen, wenn das gewünschte Volumen der Kugelschale V (für festes R) gegen das Volumen einer Vollkugel konvergiert?
3. Gesucht sei nun die Stärke d zu gegebenem Radius $R = 4$ und Volumen $V = 84\pi$. Führen Sie hierzu 4 Iteration mit dem Newton-Verfahren zum Startwert $d_0 = 2$ und mindestens 4 Nachkommastellen durch. Gegen welchen Wert konvergiert die Folge?
4. Zeigen Sie, dass das iterative Verfahren

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k) + p_k f(x_k)}$$

für eine allgemeine, hinreichend glatte Funktionen f mit Ordnung 2 gegen eine einfache Nullstelle konvergiert, wenn $p_k \in \{-1, 1\}$ so gewählt wird, dass $f'(x_k)$ und $p_k f(x_k)$ das gleiche Vorzeichen haben. Warum ist diese Wahl von p_k wichtig?

Back to GPT3.5

Prompt

In dieser Aufgabe soll iterativ die Stärke $d \in (0, R)$ einer Kugelschale mit Volumen $V > 0$ und gegebenem Außenradius $R > 0$ bestimmt werden, wobei angenommen werden kann, dass V durch das Volumen einer Vollkugel mit Radius R beschränkt ist, also $V < 4\pi/3 R^3$. a) Bestimmen Sie eine Funktion f , welche die Stärke d als Nullstelle besitzt und deren Auswertung ohne die Berechnung von Wurzeln auskommt. Formulieren Sie die Funktion so, dass diese im Fall kleiner Stärken d gut konditioniert ist und keine Auslöschung auftreten kann.

Solving a Numeric I exam

Answer

Um eine Funktion zu formulieren, die die Stärke d als Nullstelle besitzt, ohne Wurzeln zu verwenden, können wir das Volumen der Kugelschale als Funktion in Bezug auf die Stärke d ausdrücken. Wir wissen, dass das Volumen einer Kugelschale gegeben ist durch:

$$V_{\text{shell}} = 4\pi/3 (R^3 - (R - d)^3)$$

Die Funktion $f(d)$ soll nun die Stärke d als Nullstelle haben. Um dies zu erreichen, können wir die Funktion $g(d) = V_{\text{shell}} - V$ definieren, wobei V das gegebene Volumen ist. Die Nullstelle von $g(d)$ entspricht dann der gesuchten Stärke d .

$$g(d) = 4\pi/3 (R^3 - (R - d)^3) - V$$

Um sicherzustellen, dass die Funktion gut konditioniert ist und keine Auslöschung auftritt, können wir die Funktion $g(d)$ umschreiben, um die Subtraktion von ähnlichen Werten zu vermeiden. Dazu verwenden wir die Identität $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$:

$$g(d) = 4\pi/3 [(R^3 - (R^3 - 3R^2d + 3Rd^2 - d^3))] - V = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$$

Diese Funktion $f(d)$ erfüllt die Bedingungen und ermöglicht die Bestimmung der Stärke d als Nullstelle, ohne die Berechnung von

Wurzeln und mit guter Konditionierung.

Solving a Numeric I exam

Answer

Um eine Funktion zu formulieren, die die Stärke d als Nullstelle besitzt, ohne Wurzeln zu verwenden, können wir das Volumen der Kugelschale als Funktion in Bezug auf die Stärke d ausdrücken. Wir wissen, dass das Volumen einer Kugelschale gegeben ist durch:

$$V_{\text{shell}} = 4\pi/3 (R^3 - (R - d)^3)$$

Die Funktion $f(d)$ soll nun die Stärke d als Nullstelle haben. Um dies zu erreichen, können wir die Funktion $g(d) = V_{\text{shell}} - V$ definieren, wobei V das gegebene Volumen ist. Die Nullstelle von $g(d)$ entspricht dann der gesuchten Stärke d .

$$g(d) = 4\pi/3 (R^3 - (R - d)^3) - V$$

Um sicherzustellen, dass die Funktion gut konditioniert ist und keine Auslöschung auftritt, können wir die Funktion $g(d)$ umschreiben, um die Subtraktion von ähnlichen Werten zu vermeiden. Dazu verwenden wir die Identität $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$:

$$g(d) = 4\pi/3 [(R^3 - (R^3 - 3R^2d + 3Rd^2 - d^3))] - V = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$$

Diese Funktion $f(d)$ erfüllt die Bedingungen und ermöglicht die Bestimmung der Stärke d als Nullstelle, ohne die Berechnung von Wurzeln und mit guter Konditionierung.

- Not sure where it uses the identity, but it is a valid iteration function.

Solving a Numeric I exam

Answer

Um eine Funktion zu formulieren, die die Stärke d als Nullstelle besitzt, ohne Wurzeln zu verwenden, können wir das Volumen der Kugelschale als Funktion in Bezug auf die Stärke d ausdrücken. Wir wissen, dass das Volumen einer Kugelschale gegeben ist durch:

$$V_{\text{shell}} = 4\pi/3 (R^3 - (R - d)^3)$$

Die Funktion $f(d)$ soll nun die Stärke d als Nullstelle haben. Um dies zu erreichen, können wir die Funktion $g(d) = V_{\text{shell}} - V$ definieren, wobei V das gegebene Volumen ist. Die Nullstelle von $g(d)$ entspricht dann der gesuchten Stärke d .

$$g(d) = 4\pi/3 (R^3 - (R - d)^3) - V$$

Um sicherzustellen, dass die Funktion gut konditioniert ist und keine Auslöschung auftritt, können wir die Funktion $g(d)$ umschreiben, um die Subtraktion von ähnlichen Werten zu vermeiden. Dazu verwenden wir die Identität $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$:

$$g(d) = 4\pi/3 [(R^3 - (R^3 - 3R^2d + 3Rd^2 - d^3))] - V = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$$

Diese Funktion $f(d)$ erfüllt die Bedingungen und ermöglicht die Bestimmung der Stärke d als Nullstelle, ohne die Berechnung von Wurzeln und mit guter Konditionierung.

- Not sure where it uses the identity, but it is a valid iteration function.
- But it renamed f to g .

Solving a Numeric I exam

Answer

Um eine Funktion zu formulieren, die die Stärke d als Nullstelle besitzt, ohne Wurzeln zu verwenden, können wir das Volumen der Kugelschale als Funktion in Bezug auf die Stärke d ausdrücken. Wir wissen, dass das Volumen einer Kugelschale gegeben ist durch:

$$V_{\text{shell}} = 4\pi/3 (R^3 - (R - d)^3)$$

Die Funktion $f(d)$ soll nun die Stärke d als Nullstelle haben. Um dies zu erreichen, können wir die Funktion $g(d) = V_{\text{shell}} - V$ definieren, wobei V das gegebene Volumen ist. Die Nullstelle von $g(d)$ entspricht dann der gesuchten Stärke d .

$$g(d) = 4\pi/3 (R^3 - (R - d)^3) - V$$

Um sicherzustellen, dass die Funktion gut konditioniert ist und keine Auslöschung auftritt, können wir die Funktion $g(d)$ umschreiben, um die Subtraktion von ähnlichen Werten zu vermeiden. Dazu verwenden wir die Identität $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$:

$$g(d) = 4\pi/3 [(R^3 - (R^3 - 3R^2d + 3Rd^2 - d^3))] - V = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$$

Diese Funktion $f(d)$ erfüllt die Bedingungen und ermöglicht die Bestimmung der Stärke d als Nullstelle, ohne die Berechnung von Wurzeln und mit guter Konditionierung.

- Not sure where it uses the identity, but it is a valid iteration function.
- But it renamed f to g .
- Except in the answer... Probably it saw a very similar exercise somewhere else.

Prompt

Argumentieren Sie, dass f auf dem Intervall $[0, R]$ eine Nullstelle besitzt und begründen Sie mit Hilfe von Abschätzungen aus der Vorlesung, warum die Startwerte immer besser gewählt werden müssen, wenn das gewünschte Volumen der Kugelschale V (für festes R) gegen das Volumen einer Vollkugel konvergiert?

Solving a Numeric I exam

Answer

Um zu zeigen, dass die Funktion f auf dem Intervall $[0, R]$ eine Nullstelle besitzt, müssen wir nachweisen, dass $f(0) < 0$ und $f(R) > 0$ gilt.

Für $f(0)$ setzen wir $d = 0$ ein:

$$f(0) = 4\pi/3 [3R^2(0) - 3R(0)^2 + (0)^3] - V = 4\pi/3 [0 - 0 + 0] - V = -V < 0$$

Da $V > 0$ gegeben ist, ist $f(0)$ negativ.

Für $f(R)$ setzen wir $d = R$ ein:

$$f(R) = 4\pi/3 [3R^2(R) - 3R(R)^2 + (R)^3] - V = 4\pi/3 [3R^3 - 3R^3 + R^3] - V = 4\pi/3 R^3 - V > 0$$

Da $V < 4\pi/3 R^3$ gemäß der gegebenen Einschränkung ist, ist $f(R)$ positiv.
[...]

- First part is fine...
- But a short reason why this is enough would be nice.

Solving a Numeric I exam

Answer

Nun zur Wahl der Startwerte: Wenn das gewünschte Volumen V der Kugelschale gegen das Volumen einer Vollkugel konvergiert ($V \rightarrow 4\pi/3 R^3$), müssen die Startwerte immer besser gewählt werden, um eine genaue Nullstelle zu finden.

Dies liegt daran, dass in diesem Fall die Differenz zwischen dem gewünschten Volumen V und dem tatsächlichen Volumen der Kugelschale sehr klein wird. Die Funktion $f(d)$ nähert sich der Nullstelle, und kleine Änderungen in den Startwerten können zu erheblichen Unterschieden in der berechneten Stärke d führen.

Aus der Vorlesung wissen wir, dass die Kondition einer Funktion [...]

- Sounds a bit like a student who tries to at least get a few pity points...

Solving a Numeric I exam

Prompt

Gesucht sei nun die Stärke d zu gegebenem Radius $R = 4$ und Volumen $V = 84\pi$. Führen Sie hierzu 4 Iteration mit dem Newton-Verfahren zum Startwert $d_0 = 2$ und mindestens 4 Nachkommastellen durch. Gegen welchen Wert konvergiert die Folge?

- Since this involves multiple calculations, we can have a good guess how this will be going...

Solving a Numeric I exam

Answer

Um das Newton-Verfahren anzuwenden, benötigen wir die Ableitung der Funktion $f(d)$ in Bezug auf d . Die Ableitung von $f(d) = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$ ergibt:

$$f'(d) = 4\pi/3 [3R^2 - 6Rd + 3d^2]$$

Wir starten mit dem Startwert $d_0 = 2$ und führen insgesamt 4 Iterationen durch. Die Iterationsformel des Newton-Verfahrens lautet:

$$d_{\{n+1\}} = d_n - f(d_n) / f'(d_n)$$

Hier sind die Berechnungen für jede Iteration:

$$\begin{aligned} \text{Iteration 1: } d_1 &= d_0 - f(d_0) / f'(d_0) = 2 - (4\pi/3 [3(4)^2(2) - 6(4)(2)^2 + (2)^3] - 84\pi) / (4\pi/3 [3(4)^2 - 6(4)(2) + (2)^2]) \\ &\approx 2 - (32\pi - 84\pi) / (96\pi - 48\pi + 12) \approx 2 - (-52\pi) / (60\pi - 48\pi + 12) \approx 2 + (52\pi) / (12\pi) \approx 2 + 13/3 \approx 2 + 4.3333 \approx 6.3333 \end{aligned}$$

$$\begin{aligned} \text{Iteration 2: } d_2 &= d_1 - f(d_1) / f'(d_1) \approx 6.3333 - (4\pi/3 [3(4)^2(6.3333) - 6(4)(6.3333)^2 + (6.3333)^3] - 84\pi) / (4\pi/3 \\ &[3(4)^2 - 6(4)(6.3333) + (6.3333)^2]) [\dots] \end{aligned}$$

Solving a Numeric I exam

Answer

Um das Newton-Verfahren anzuwenden, benötigen wir die Ableitung der Funktion $f(d)$ in Bezug auf d . Die Ableitung von $f(d) = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$ ergibt:

$$f'(d) = 4\pi/3 [3R^2 - 6Rd + 3d^2]$$

Wir starten mit dem Startwert $d_0 = 2$ und führen insgesamt 4 Iterationen durch. Die Iterationsformel des Newton-Verfahrens lautet:

$$d_{\{n+1\}} = d_n - f(d_n) / f'(d_n)$$

Hier sind die Berechnungen für jede Iteration:

$$\begin{aligned} \text{Iteration 1: } d_1 &= d_0 - f(d_0) / f'(d_0) = 2 - (4\pi/3 [3(4)^2(2) - 6(4)(2)^2 + (2)^3] - 84\pi) / (4\pi/3 [3(4)^2 - 6(4)(2) + (2)^2]) \\ &\approx 2 - (32\pi - 84\pi) / (96\pi - 48\pi + 12) \approx 2 - (-52\pi) / (60\pi - 48\pi + 12) \approx 2 + (52\pi) / (12\pi) \approx 2 + 13/3 \approx 2 + 4.3333 \approx 6.3333 \end{aligned}$$

$$\begin{aligned} \text{Iteration 2: } d_2 &= d_1 - f(d_1) / f'(d_1) \approx 6.3333 - (4\pi/3 [3(4)^2(6.3333) - 6(4)(6.3333)^2 + (6.3333)^3] - 84\pi) / (4\pi/3 \\ &[3(4)^2 - 6(4)(6.3333) + (6.3333)^2]) [\dots] \end{aligned}$$

Solving a Numeric I exam

Answer

Um das Newton-Verfahren anzuwenden, benötigen wir die Ableitung der Funktion $f(d)$ in Bezug auf d . Die Ableitung von $f(d) = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$ ergibt:

$$f'(d) = 4\pi/3 [3R^2 - 6Rd + 3d^2]$$

Wir starten mit dem Startwert $d_0 = 2$ und führen insgesamt 4 Iterationen durch. Die Iterationsformel des Newton-Verfahrens lautet:

$$d_{\{n+1\}} = d_n - f(d_n) / f'(d_n)$$

Hier sind die Berechnungen für jede Iteration:

$$\begin{aligned} \text{Iteration 1: } d_1 &= d_0 - f(d_0) / f'(d_0) = 2 - (4\pi/3 [3(4)^2(2) - 6(4)(2)^2 + (2)^3] - 84\pi) / (4\pi/3 [3(4)^2 - 6(4)(2) + (2)^2]) \\ &\approx 2 - (32\pi - 84\pi) / (96\pi - 48\pi + 12) \approx 2 - (-52\pi) / (60\pi - 48\pi + 12) \approx 2 + (52\pi) / (12\pi) \approx 2 + 13/3 \approx 2 + 4.3333 \approx 6.3333 \end{aligned}$$

$$\begin{aligned} \text{Iteration 2: } d_2 &= d_1 - f(d_1) / f'(d_1) \approx 6.3333 - (4\pi/3 [3(4)^2(6.3333) - 6(4)(6.3333)^2 + (6.3333)^3] - 84\pi) / (4\pi/3 \\ &[3(4)^2 - 6(4)(6.3333) + (6.3333)^2]) [...] \end{aligned}$$

Solving a Numeric I exam

Answer

Um das Newton-Verfahren anzuwenden, benötigen wir die Ableitung der Funktion $f(d)$ in Bezug auf d . Die Ableitung von $f(d) = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$ ergibt:

$$f'(d) = 4\pi/3 [3R^2 - 6Rd + 3d^2]$$

Wir starten mit dem Startwert $d_0 = 2$ und führen insgesamt 4 Iterationen durch. Die Iterationsformel des Newton-Verfahrens lautet:

$$d_{\{n+1\}} = d_n - f(d_n) / f'(d_n)$$

Hier sind die Berechnungen für jede Iteration:

$$\begin{aligned} \text{Iteration 1: } d_1 &= d_0 - f(d_0) / f'(d_0) = 2 - (4\pi/3 [3(4)^2(2) - 6(4)(2)^2 + (2)^3] - 84\pi) / (4\pi/3 [3(4)^2 - 6(4)(2) + (2)^2]) \\ &\approx 2 - (32\pi - 84\pi) / (96\pi - 48\pi + 12) \approx 2 - (-52\pi) / (60\pi - 48\pi + 12) \approx 2 + (52\pi) / (12\pi) \approx 2 + 13/3 \approx 2 + 4.3333 \approx 6.3333 \end{aligned}$$

$$\begin{aligned} \text{Iteration 2: } d_2 &= d_1 - f(d_1) / f'(d_1) \approx 6.3333 - (4\pi/3 [3(4)^2(6.3333) - 6(4)(6.3333)^2 + (6.3333)^3] - 84\pi) / (4\pi/3 \\ &[3(4)^2 - 6(4)(6.3333) + (6.3333)^2]) [\dots] \end{aligned}$$

Solving a Numeric I exam

Answer

Um das Newton-Verfahren anzuwenden, benötigen wir die Ableitung der Funktion $f(d)$ in Bezug auf d . Die Ableitung von $f(d) = 4\pi/3 [3R^2d - 3Rd^2 + d^3] - V$ ergibt:

$$f'(d) = 4\pi/3 [3R^2 - 6Rd + 3d^2]$$

Wir starten mit dem Startwert $d_0 = 2$ und führen insgesamt 4 Iterationen durch. Die Iterationsformel des Newton-Verfahrens lautet:

$$d_{\{n+1\}} = d_n - f(d_n) / f'(d_n)$$

Hier sind die Berechnungen für jede Iteration:

$$\begin{aligned} \text{Iteration 1: } d_1 &= d_0 - f(d_0) / f'(d_0) = 2 - (4\pi/3 [3(4)^2(2) - 6(4)(2)^2 + (2)^3] - 84\pi) / (4\pi/3 [3(4)^2 - 6(4)(2) + (2)^2]) \\ &\approx 2 - (32\pi - 84\pi) / (96\pi - 48\pi + 12) \approx 2 - (-52\pi) / (60\pi - 48\pi + 12) \approx 2 + (52\pi) / (12\pi) \approx 2 + 13/3 \approx 2 + 4.3333 \approx 6.3333 \end{aligned}$$

$$\begin{aligned} \text{Iteration 2: } d_2 &= d_1 - f(d_1) / f'(d_1) \approx 6.3333 - (4\pi/3 [3(4)^2(6.3333) - 6(4)(6.3333)^2 + (6.3333)^3] - 84\pi) / (4\pi/3 \\ &[3(4)^2 - 6(4)(6.3333) + (6.3333)^2]) [...] \end{aligned}$$

- At least it did this error consistently...

Answer

Wiederholen Sie die Berechnung für Iteration 2, Iteration 3 und Iteration 4, um den Konvergenzwert der Folge zu bestimmen.

Iteration 3: $d_3 = \dots$ Iteration 4: $d_4 = \dots$

Bitte beachten Sie, dass die Berechnungen aufgrund der Komplexität der Ausdrücke und der Rundungen während des Rechenprozesses nicht in ihrer vollständigen Genauigkeit dargestellt werden. Die endgültigen Werte für d_3 und d_4 sollten jedoch eine Konvergenz gegen einen bestimmten Wert zeigen.

- It even "surrenders" for the last two iterations.

Prompt

Zeigen Sie, dass das iterative Verfahren

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k) + p_k f(x_k)}$$

für eine allgemeine, hinreichend glatte Funktionen f mit Ordnung 2 gegen eine einfache Nullstelle konvergiert, wenn $p_k \in \{-1, 1\}$ so gewählt wird, dass $f'(x_k)$ und $p_k f(x_k)$ das gleiche Vorzeichen haben. Warum ist diese Wahl von p_k wichtig?

- As a last example, a genuine proof.

Solving a Numeric I exam

Answer

Um zu zeigen, dass das iterative Verfahren gegen eine einfache Nullstelle konvergiert, betrachten wir die Taylor-Entwicklung von f um die Nullstelle x :

$$f(x_k) = f(x) + f'(x)(x_k - x) + \frac{1}{2}f''(\xi_k)(x_k - x)^2$$

Hierbei ist ξ_k ein Wert zwischen x und x_k , der von der Wahl des Zwischenwertsatzes abhängt.

Da x eine Nullstelle von f ist, gilt $f(x) = 0$. Substituieren wir dies in die Taylor-Entwicklung:

$$f(x_k) = f'(x)(x_k - x) + \frac{1}{2}f''(\xi_k)(x_k - x)^2$$

Das Ziel des iterativen Verfahrens ist es, einen neuen Punkt x_{k+1} zu finden, der näher an der Nullstelle liegt. Wir setzen $x_{k+1} = x_k + h_k$, wobei h_k den Fehler zwischen x_{k+1} und der Nullstelle x darstellt. Dann können wir $x_k - x = -h_k$ schreiben. Wir können die Gleichung umstellen, um h_k auszudrücken:

$$f'(x)h_k = \frac{1}{2}f''(\xi_k)h_k^2 - f(x_k)$$

Dividieren wir beide Seiten durch $f'(x)$:

$$h_k = \frac{\frac{1}{2}f''(\xi_k)h_k^2 - f(x_k)}{f'(x)}$$

Um die Konvergenz zu analysieren, nehmen wir an, dass h_k nahe genug an 0 liegt, so dass h_k^2 vernachlässigt werden kann. Dann vereinfacht sich der Ausdruck zu:

$$h_k \approx -\frac{2f(x_k)}{f'(x)} [\dots]$$

- While Taylor-Expansion and error terms are somewhat the right direction...
- This leads nowhere.

Solving a Numeric I exam

Results

- Compared to students: (Far) below average...
- But the main problems are calculations and longer transformations.
- Clearly the base ideas are there.

Seminar presentation

Premise

- Maybe a more realistic application: Help on presentation and write-up for some topic.
- Topic we focus on here: QR Algorithm.
- Broad topic, interesting implementation aspects.
- Base question: What can the generative AI do by itself? Where can it help?
- Not as straight forward as: "Write a draft on topic x".

Seminar presentation

How to proceed?

1. Get a broad overview. Maybe a few references?
2. Structure the presentation/draft.
3. Generate text, codes and examples.
 - From now on, we only use GPT4.

Seminar presentation

How to proceed?

1. Get a broad overview. Maybe a few references?
2. Structure the presentation/draft.
3. Generate text, codes and examples.
 - From now on, we only use GPT4.

What is the QR algorithm? Do you know of any good references?

With GPT4

The QR algorithm is an eigenvalue algorithm, that is, a procedure to calculate the eigenvalues and eigenvectors of a matrix. It was developed in the [...]. The algorithm gets its name from the QR decomposition which forms the basis of the procedure.

The basic idea of the algorithm is to perform a QR decomposition, write the original matrix as a product of an orthogonal and a triangular matrix (QR), then form a new matrix from the reverse product of these (RQ). This new matrix is similar to the original (they represent the same linear transformation in different bases), and after a few iterations, the procedure converges to a form where the original matrix's eigenvalues can be read off the diagonal. Here is a very simple pseudocode of the QR algorithm for a square matrix A:

```
while A is not sufficiently close to upper triangular
  decompose A into A = QR
  set A = RQ
end
```

This QR algorithm is efficient and numerically stable, but requires the input matrix to be full rank (no linearly dependent rows or columns). There are variants that work for other cases. [...] "Matrix Computations" by Gene H. Golub and Charles F. Van Loan (4th Edition, Johns Hopkins University Press, 2013). This is [...] "Numerical Recipes: The Art of Scientific Computing" by William H. Press et. al. (3rd Edition, Cambridge University Press, 2007). This is [...] The Wikipedia page on the QR algorithm is also a good starting point [...] Remember, understanding the QR algorithm requires a good foundation in linear algebra, especially the concepts of matrix decompositions and eigenvalues. If you're not already familiar with these topics, you may want to review them first.

- Start is very close to Wikipedia... but diverges after the first sentence.
- References exist, and are cited correctly. This is not to be expected!

Seminar presentation

Small Intermission

Prompt

Generate a bibtex file with this references and a small explanation on how to use it.

Answer

```
@book{golub2013matrix,  
title={Matrix Computations},  
author={Golub, Gene H and Van Loan, Charles F},  
year={2013},  
publisher={Johns Hopkins University Press},  
edition={4th}  
}
```

```
@book{press2007numerical,  
title={Numerical Recipes: The Art of Scientific Computing},  
author={Press, William H and others},  
year={2007},  
publisher={Cambridge University Press},  
edition={3rd}  
}
```

[...] Save this as a .bib file. Suppose you name it as references.bib.

To use these citations in a LaTeX document, use the `\cite{}` command where [...]

- This is a neat example for one of the small automation-tools implicitly introduced with LLM's...

Seminar presentation

Prompt

I need to create a presentation about the QR-algorithm and its application with a focus on mathematics. Could you give me steps on how to approach this?

- Now we want to generate a structure for the presentation.

Seminar presentation

Sure, I'd be happy to help you outline a presentation on the QR algorithm. This is a fundamental topic in numerical linear algebra, with many applications.

I. Introduction

1. Brief Overview

- Define the QR algorithm
- Importance of the QR algorithm in mathematics and computer science
- What will be covered in the presentation

II. Historical Background

1. Development of the QR algorithm

- Francis and Kublanovskaya's contribution
- How it has evolved over time

III. Understanding Matrices

1. Quick review of matrices

- What is a matrix?
- Different types of matrices (e.g., square, rectangular, diagonal, etc.)
- Eigenvalues and eigenvectors

2. Matrix Decomposition

- Explain what it is and why it is important
- Briefly introduce other types of decompositions (LU, Cholesky, etc.) [...]

Seminar presentation

IV. QR Decomposition

1. What is QR Decomposition

- Explain Q (orthogonal matrix) and R (upper triangular matrix)

2. Mathematical explanation of QR Decomposition

- Explain how to perform QR Decomposition with an example
- Discuss properties of QR Decomposition

3. Algorithms for QR Decomposition

- Gram-Schmidt process
- Householder reflections
- Givens rotations

V. QR Algorithm

1. How QR Decomposition leads to the QR Algorithm

2. Detailed explanation of the QR Algorithm

- How it's used for finding eigenvalues
- Discuss shift strategies for accelerating convergence

3. Work through an example

- Apply the QR Algorithm on a simple matrix and iterate to convergence

4. Discuss the convergence of the QR Algorithm [...]

Seminar presentation

VI. Applications

1. Real-world examples where the QR Algorithm is useful

- Solve linear least squares for overdetermined systems
- Solve eigenvalue/eigenvector problems in physics and engineering
- Use in the Google PageRank algorithm
- QR decomposition in QR codes

2. Mathematical software that uses the QR Algorithm

VII. Limitations and Advantages

1. Discuss the limitations of the QR Algorithm

- Large matrices and computational cost
- Precision errors

2. Discuss the advantages of the QR Algorithm

- Stability and accuracy
- Applicability to various problems

VIII. Conclusion [...]

Seminar presentation

VI. Applications

1. Real-world examples where the QR Algorithm is useful

- Solve linear least squares for overdetermined systems
- Solve eigenvalue/eigenvector problems in physics and engineering
- Use in the Google PageRank algorithm
- QR decomposition in QR codes

2. Mathematical software that uses the QR Algorithm

VII. Limitations and Advantages

1. Discuss the limitations of the QR Algorithm

- Large matrices and computational cost
- Precision errors

2. Discuss the advantages of the QR Algorithm

- Stability and accuracy
- Applicability to various problems

VIII. Conclusion [...]

Seminar presentation

- While a bit too long, the essential steps are there, maybe an explicit mention of Hessenberg-Matrices missing.
- It somewhat mixes the QR algorithm with the QR decomposition.

Seminar presentation

In the same Conversation

Prompt

You said as an application we can use the QR algorithm to solve linear least squares for overdetermined systems. Could you give me an example for this?

Yes, indeed. The QR factorization can be used to solve the least squares problem for overdetermined systems, which is a common scenario in applications such as data fitting, where you have more equations (data points) than unknowns. [...]

- Indeed, it confuses this, but at least explains what it means here.

Seminar presentation

```
% Create a system Ax = b where A is m x n with m > n
m = 100;
n = 10;
A = rand(m, n);
x_true = rand(n, 1);
b = A * x_true + 0.05 * randn(m, 1); % Add some noise

% Compute the QR decomposition of A
[Q, R] = qr(A);

% Solve the system Rx = Q'b for x using back-substitution
x_est = R \ (Q' * b);

% Compare the true and estimated x
error = norm(x_true - x_est);
disp(error);
```

- As a side-effect, it gives a nice method to generate a test-setup for a least-squares problem and a explanation why it works.
- To be exact, this was after asking a few other things, so the idea to directly implement this in Matlab was probably due to a prompt before.

Seminar presentation

Back to the overview

- If we stop here, chatGPT at least helps for organizing the work.
- To go further, we want to let it subdivide the single points again.

Seminar presentation

Back to the overview

- If we stop here, chatGPT at least helps for organizing the work.
- To go further, we want to let it subdivide the single points again.

Prompt

What are important steps for number IV ? What should I focus on?

Seminar presentation

Section IV focuses on the QR Decomposition, which is the foundation for the QR Algorithm. Here's a breakdown of what each part should contain:

1. What is QR Decomposition?

- Define QR Decomposition: Explain that it is a method of decomposing a matrix into an orthogonal matrix (Q) and an upper triangular matrix (R).
- Detail the properties of orthogonal and upper triangular matrices. This is crucial because the properties of Q and R are what make QR decomposition useful.

[...]

- Results in more fine grained steps.

Seminar presentation

Section IV focuses on the QR Decomposition, which is the foundation for the QR Algorithm. Here's a breakdown of what each part should contain:

1. What is QR Decomposition?

- Define QR Decomposition: Explain that it is a method of decomposing a matrix into an orthogonal matrix (Q) and an upper triangular matrix (R).
- Detail the properties of orthogonal and upper triangular matrices. This is crucial because the properties of Q and R are what make QR decomposition useful.

[...]

- Results in more fine grained steps.
- We can now use a sub-point to "self-prompt" chatGPT.

Seminar presentation

Prompt

Detail the properties of orthogonal and upper triangular matrices, especially in context of the QR algorithm. Give mathematical formulas as needed.

Seminar presentation

Prompt

Detail the properties of orthogonal and upper triangular matrices, especially in context of the QR algorithm. Give mathematical formulas as needed.

Sure, let's dive into the properties of orthogonal and upper triangular matrices, which are crucial in the context of the QR decomposition and algorithm.

Orthogonal Matrix (Q)

An orthogonal matrix is a square matrix whose columns and rows are orthogonal unit vectors (also known as orthonormal vectors). This leads to two important properties:

The transpose of an orthogonal matrix equals its inverse, mathematically represented as:

$$Q^T = Q^{-1}$$

Multiplication of an orthogonal matrix with its transpose results in the identity matrix:

$$QQ^T = I \text{ [...]}$$

Seminar presentation

Upper Triangular Matrix (R)

[...]

$R_{ij} = 0$ for $i > j$

[...]

Given a matrix A , the QR decomposition is $A=QR$. So if we want to solve a system $Ax=b$, we can substitute to get $QRx=b$. Then, we set $y=Rx$ and solve the system $Qy=b$, which is straightforward because $Q^{-1} = Q^T$. Once we have y , we solve $Rx=y$, which is easy because R is an upper triangular matrix. [...]

- For a presentation this is enough level of detail...
- Although it could be more exact, for example in definitions.
- You could write this up yourself... or:

Prompt

Could you summarize this information into two latex beamer slides?

- Prompt leads to working latex code resulting in the following two slides.
- Prompt also includes a few setups and explanation, with which you are able to directly compile the slides...
- While the order of points is a bit confusing and the formatting could lead to consistency problems, the slides are good enough to adapt...

Orthogonal Matrices

Definition

An **orthogonal matrix** Q is a square matrix whose rows and columns are orthogonal unit vectors.

Properties

- The transpose of Q equals its inverse: $Q^T = Q^{-1}$
- Multiplication of Q with its transpose results in the identity matrix: $QQ^T = I$
- These properties stem from the orthonormality condition: $Q^T Q = I$

Importance in QR Decomposition

- QR Decomposition: $A = QR$
- Solving system: $Ax = b$ becomes $QRx = b$, then $Qy = b$ and $Rx = y$
- Orthogonality of Q provides numerical stability

Upper Triangular Matrices

Definition

An **upper triangular matrix** R is a square matrix where all entries below the main diagonal are zero: $R_{ij} = 0$ for $i > j$

Properties

- Systems of linear equations with an upper triangular matrix can be easily solved through back substitution

Importance in QR Decomposition

- Form of R aids in maintaining appropriate matrix form in the QR algorithm
- Systems involving R can be solved easily, improving the efficiency of the QR algorithm

Seminar presentation

Generating Code

Could you write me a matlab implementation, that uses Givens rotation to calculate the QR factorization of a square matrix?

```
function [Q, R] = givensQR(A)
% Check if the input matrix is square
[m, n] = size(A);
if m ~= n
error('The input matrix must be square. ');
end

% Initialization
Q = eye(n); % Identity matrix
R = A; % Copy of A

for j = 1:n
for i = n:-1:(j+1)
G = eye(n);
[c, s] = givensrotation( R(i-1,j),R(i,j) );
G([i-1, i],[i-1, i]) = [c -s; s c];% GivensRot
R = G'*R; % Apply on R
Q = Q*G; % Apply on Q
end
end
end
```

```
function [c, s] = givensrotation(a, b)
if b == 0
c = 1;
s = 0;
else
if abs(b) > abs(a)
r = a / b;
s = 1 / sqrt(1 + r^2);
c = s * r;
else
r = b / a;
c = 1 / sqrt(1 + r^2);
s = c * r;
end
end
end
```


Seminar presentation

- We can go even further with the self prompting.
- This is directly after our request for the Given's Rotation, so it can reference it.

Prompt

Could you give an example for solving eigenvalue/eigenvector problems in physics and engineering and solving it with the QR algorithm?

Seminar presentation

Consider a vibrating system, which can be represented by the following mass-spring system with two masses and three springs: [...]

$$m_1 * x_1'' = -k_1 * x_1 + k_2 * (x_2 - x_1)$$

$$m_2 * x_2'' = -k_3 * x_2 - k_2 * (x_2 - x_1)$$

[...]

$$X'' + M^{-1} * K * X = 0$$

Here, $M^{-1} * K$ is a matrix, which we can represent as A . So, we have:

$$A * X = \omega^2 * X$$

- Until here, nice short explanation on how to construct the problem.
- It also directly gives us the code for the problem:

Seminar presentation

```
% Define the constants
m1 = 1; m2 = 1; % masses
k1 = 1; k2 = 1; k3 = 1; % spring constants

% Define the M and K matrices
M = [m1 0; 0 m2];
K = [k1+k2 -k2; -k2 k2+k3];

% Calculate A
A = inv(M) * K;

% QR algorithm
tol = 1e-6; % tolerance for convergence
maxiter = 1000; % maximum number of iterations
Q = eye(size(A));
R = A;
iter = 0;
while max(max(abs(triu(R,1)))) > tol && iter < maxiter
    [Q, R] = givensQR(Q*R); % use the function givensQR defined in the previous response
    iter = iter + 1;
end
eigenvalues = diag(R); % eigenvalues are the diagonal elements of R
eigenvectors = Q; % eigenvectors are the columns of Q

% The natural frequencies are the square roots of the eigenvalues
frequencies = sqrt(eigenvalues)
```

Seminar presentation

```
% Define the constants
m1 = 1; m2 = 1; % masses
k1 = 1; k2 = 1; k3 = 1; % spring constants

% Define the M and K matrices
M = [m1 0; 0 m2];
K = [k1+k2 -k2; -k2 k2+k3];

% Calculate A
A = inv(M) * K;

% QR algorithm
tol = 1e-6; % tolerance for convergence
maxiter = 1000; % maximum number of iterations
Q = eye(size(A));
R = A;
iter = 0;
while max(max(abs(triu(R,1)))) > tol && iter < maxiter
    [Q, R] = givensQR(Q*R); % use the function givensQR defined in the previous response
    iter = iter + 1;
end
eigenvalues = diag(R); % eigenvalues are the diagonal elements of R
eigenvectors = Q; % eigenvectors are the columns of Q

% The natural frequencies are the square roots of the eigenvalues
frequencies = sqrt(eigenvalues)
```

Seminar presentation

- Two pretty nasty errors, presented with absolute confidence and the convergence condition is also wrong.
- If you do not know what you are doing, you will not spot this, since the program will not abort.
- In the end you have pretty wrong results.
- chatGPT can somewhat fix that... but only if you directly call it out.
- For a nice overview what chatGPT can programm an what not, see for example [[KM23](#)].

Summary

What did we observe?

- GPT can generate a coherent presentation and elaboration.
- It can even generate and work on standard examples and exercises...
- At least on a broad/beginner level.
- If you are lucky it can even generate more detailed parts.
- But it **will** do hard to spot errors in any detailed part.
- This can change rapidly due to development, since the gap is not that large.
- Still one concern: Using AI to modify/translate sources and using these.
- At least for mathematical texts this is harder due to unicode inconsistencies.

Summary

A few guidelines on how to make AI harder to cheat with

- Avoid broad topics and summaries of fields.
- Give specific tasks.
- Weight in-person efforts more.
- Modify known results/examples, ingenuity is (for now) the biggest hurdle of LLM's.
- Raise the level to pass.

Summary



OpenAI. <https://openai.com/chatgpt>. Zugriff am 2. Juli 2023. 2023.



Jörg von Garrel, Jana Mayer, and Markus Mühlfeld. “Künstliche Intelligenz im Studium Eine quantitative Befragung von Studierenden zur Nutzung von ChatGPT & Co.”. 2023.



Ali Kashefi and Tapan Mukerji. *ChatGPT for Programming Numerical Methods*. 2023. arXiv: 2303.12093 [cs.LG].



Ashish Vaswani et al. *Attention Is All You Need*. 2017. arXiv: 1706.03762 [cs.CL].

Summary

Curiosity: GPT3.5 vs GPT4 for math exercises...

- A bit strange: Overall GPT3.5 answers this exam questions more consistently than GPT4.

Summary

Curiosity: GPT3.5 vs GPT4 for math exercises...

- A bit strange: Overall GPT3.5 answers this exam questions more consistently than GPT4.
- Probably since GPT3.5 strays less away from "known" examples.

Summary

Curiosity: GPT3.5 vs GPT4 for math exercises...

- A bit strange: Overall GPT3.5 answers this exam questions more consistently than GPT4.
- Probably since GPT3.5 strays less away from "known" examples.
- Only an assumption: GPT4 generates content more self-reliant, and is more reluctant to give a straight forward answer. A language model is simply not intended to do arithmetic.

Summary

Curiosity: GPT3.5 vs GPT4 for math exercises...

- A bit strange: Overall GPT3.5 answers this exam questions more consistently than GPT4.
- Probably since GPT3.5 strays less away from "known" examples.
- Only an assumption: GPT4 generates content more self-reliant, and is more reluctant to give a straight forward answer. A language model is simply not intended to do arithmetic.
- In the case the exercise is very similar to some training data, it is a more "successful" strategy to simply copy the other solution.

Summary

Curiosity: GPT3.5 vs GPT4 for math exercises...

- A bit strange: Overall GPT3.5 answers this exam questions more consistently than GPT4.
- Probably since GPT3.5 strays less away from "known" examples.
- Only an assumption: GPT4 generates content more self-reliant, and is more reluctant to give a straight forward answer. A language model is simply not intended to do arithmetic.
- In the case the exercise is very similar to some training data, it is a more "successful" strategy to simply copy the other solution.
- This will lead to strange consistency errors for GPT3.5, but higher chances to give a useful answer.