

How Smart is ChatGPT 4?

Jeffrey Morgan, Associate Provost and Professor of Mathematics, University of Houston

This is a popular question. Some media sources describe ChatGPT 4 as god-like, and we have been told that AI will eventually be more impactful than electricity. It's not hard for me to imagine that AI will someday achieve the hype. But we don't know when. This is like green energy. Eventually, it's going to be amazing. But it's not amazing yet, and we will live in a world mixed with green and nongreen energy for some time. Similarly, ChatGPT and AI will get there, but it could take a minute.

As a faculty member, I decided to determine how ChatGPT 4 handles the questions I ask my students, and learn whether ChatGPT can pass my classes. I've experimented with questions from several courses, with mixed results. I'm currently teaching a sophomore level linear algebra class this semester, so I gave my midterm exam to ChatGPT (one question at a time). The results were mixed. ChatGPT did not pass the test, but it did answer some difficult questions correctly. It also surprised me with some miserable responses on some simple questions. Recently, I gave an assignment to ChatGPT and I'll give the responses to my students next week as an assignment for them to determine which responses are correct, and why. I also gave ChatGPT 4 a recent problem that was purely computational, and I thought it was sure to ace. I was horribly wrong.

The problem has a simple statement. How many positive definite matrices are there of the form

$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ where a , b and c are integers from -10 to 10 ? For readers who are not familiar with basic linear algebra, this question is equivalent to asking how many points (a, b, c) have integer coefficients from -10 to 10 , with $a > 0$ and $ac - b^2 > 0$. This is a simple problem that can be answered by writing a few short lines of code. As a result, I assumed ChatGPT 4 would have no trouble.

Before I discuss the response from ChatGPT, let me comment that ChatGPT 4 understands the LaTeX mathematics typesetting language, and can give its responses in LaTeX. So, mathematics questions can be given to ChatGPT using this language. Of course, these questions can also be given by using the words that would be included in alt text for students who are visually impaired. So, in this case, I used the latter approach.

Me: Can you determine the number of positive definite 2 by 2 real symmetric matrices whose entries are integers from -10 to 10 ?

Overview of ChatGPT's response: ChatGPT correctly restated the problem in mathematical terms, and said a matrix of this form is positive definite if and only if its eigenvalues are positive. It named the set of all possible matrices M , and wrote $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ for one of the possible matrices. It then used the quadratic formula to find the eigenvalues, and concluded that the values a , b and c must be chosen so that $ac - b^2 > 0$. ChatGPT did not notice that $a > 0$ is also required. But I assumed it would discover this at some point. Instead, ChatGPT proceeded to justify some computer code to find the answer.

It began by saying,

“For each possible value of a from -10 to 10 , there are 21 possible values of b , and 21 possible values of c . Therefore, there are $21^2 = 441$ possible choices for the entries of A . To determine how many of these matrices are positive definite, we need to calculate the value of $ac - b^2$ for each matrix.”

Unfortunately, the value 21^2 is much smaller than the actual number of matrices. ChatGPT then said,

“Since $-10 \leq a, b, c \leq 10$, we have $ac - b^2 \leq 10^2 - (-10)^2 = 200$. Therefore, we only need to check the values of $ac - b^2$ for $0 \leq ac - b^2 \leq 200$.”

Forget the logical error and notice that $10^2 - (-10)^2 = 0$. ChatGPT continued by saying,

“There are 441 possible matrices in M . We can calculate the value of $ac - b^2$ for each of these matrices in $O(441)=O(1)$ time. For each value of k from 0 to 200, we can count the number of matrices in M whose value of $ac - b^2$ is exactly k , and whose eigenvalues are positive. This can be done in $O(441)$ time for each value of k .”

This keeps getting worse. ChatGPT continues with some more incorrect logic, writes some pseudo code, and concludes that there are 310 positive definite matrices of this type.

BTW – The correct answer is 986.

So, ChatGPT 4 is not god-like. It seemed to understand the question, but its logic after that point was horribly flawed.

As I said above, ChatGPT can give correct answers to questions, and I encourage faculty to experiment by visiting <https://openai.com/product/gpt-4> and clicking on “Try on ChatGPT Plus.” After my first month, I decided to pay the \$20 per month to (hopefully) get better access. It’s not clear what I’m getting, but I don’t buy a latte each morning, so I can afford it.

You can also test these features through the new Bing, and the new Bing understands LaTeX. Enjoy.