

The use of the Chat GPT to solve mathematical programming tasks: a didactic experiment with the participation of Warsaw University of Life Sciences students

Włodzimierz Wojas¹ and Piotr Stachura¹

Department of Applied Mathematics, Warsaw University of Life Sciences (SGGW),
Nowoursynowska 159, 02-776 Warsaw, Poland
wlodzimierz_wojas@sggw.edu.pl, piotr_stachura1@sggw.edu.pl

Abstract. In this article, the authors present three didactic examples of using Chat GPT in mathematical programming optimization tasks: linear programming, nonlinear programming and convex programming. These examples are analyzed in terms of the correctness of the methods used and the solutions obtained. The article also describes a didactic experiment with the participation of Informatics and Econometrics students of Warsaw University of Life Sciences, consisting in solving optimization tasks on their own using Chat GPT. The final conclusions of the article also present a comparison of the approach based on the CAS methodology and the approach using Chat GPT.

Keywords: AI · Chat GPT · CAS · mathematical programming · optimization · mathematical didactic.

1 Introduction

Chat GPT (Chat Generative Pre-trained Transformed) was developed by research organization OpenAI and launched in November 2022. It is available in online versions: a free version based on GPT-3.5 and more advanced commercial versions based on GPT-4 and GPT-4o [11, 9]. Based on the LLM (Large Language Model), Chat GPT works by learning statistical language patterns from huge databases of online texts, which also contain various types of false information and outdated knowledge. It allows users to conduct and refine conversations, and subsequent answers, questions and suggestions are treated as context at every stage of the conversation. It can both download and create images, as well as use math and coding software [11, 13, 16, 7]. Chatbots, in general, can generate fabricated information called hallucinations. The topic of using artificial intelligence, in particular Chat GPT, in teaching and learning various subjects - including mathematics and computer science - has been undertaken and widely discussed in recent years [5, 4, 14, 17, 3, 6, 15, 12, 1, 2, 10]. Assessing the usefulness of Chat GPT in learning and teaching mathematics requires solving a number of computational tasks from different areas, with different levels of difficulty, and

analyzing obtained solutions. For example, in the article [15] a scan of solving a quadratic equation in a math lesson using GPT Chat is presented. The paper [17] contains scans from GPT Chat sessions presenting solving algebraic equations, calculating the limit of a function of one variable and elementary geometry tasks using the Pythagorean theorem.

Related articles highlight the educational utility of Chat GPT as a comprehensive tool that provides users with access to basic knowledge of mathematics and other fields, including access to basic definitions, methods, instructions and algorithms. On the other hand, attention was also drawn to Chat GPT's lack of deep understanding of some issues, as well as the lack of effective opportunities to correct its "false beliefs". It is also worth emphasizing that the creators of Chat GPT themselves recognize the limitations of current versions of Chat GPT in solving mathematical problems, stating that Chat GPT is good at writing texts but is less effective in mathematical computations where there is a specific final answer [8].

Chat GPT has become a popular tool in the last two years, also among students e.g. at the Faculty of Applied Informatics and Mathematics of Warsaw University of Life Sciences, a student research club dedicated to Chat GPT has been operating for some time. This paper takes up the subject of usefulness of Chat GPT as a didactic tool in teaching and learning advanced topics and optimization algorithms in mathematical programming. Several quite typical calculation tasks are analyzed, falling within the scope of course subjects such as Mathematical Analysis and Mathematical Programming.

2 The analyzed examples

Three examples of calculation tasks will be analyzed: linear programming (LP) task, nonlinear programming (NLP) task and convex programming (CP) task. For the LP task, we will analyze Chat GPT's use of two methods: the graphical method and the simplex algorithm. For the NLP task, we will analyze Chat GPT's use of the Karush-Kuhn-Tucker conditions method (KKT). And for the CP task, we will analyze Chat GPT's use of the method based on the fact that a local minimum of a strictly convex function in a convex region is also the global minimum of this function.

Presented scans are fragments of solutions proposed by Chat GPT. They come from a Chat GPT sessions using version based on GPT-4. The sessions were held in April 2024.

3 Selected Chat GPT scans for the LP task

Scan 1 shows the initial form of the LP task, the task formulation in standard form by Chat GPT, and the correct final answer.

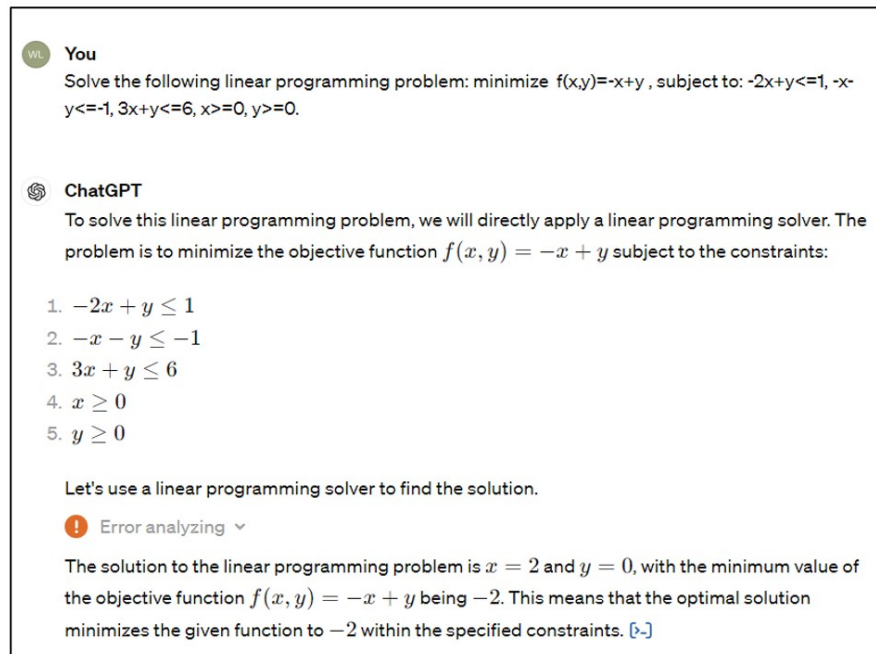


Fig. 1. The initial form of the LP task

Scan 2 contains a question about the method used to solve the LP task and the answer that the "linprog" function from the SciPy library was used.

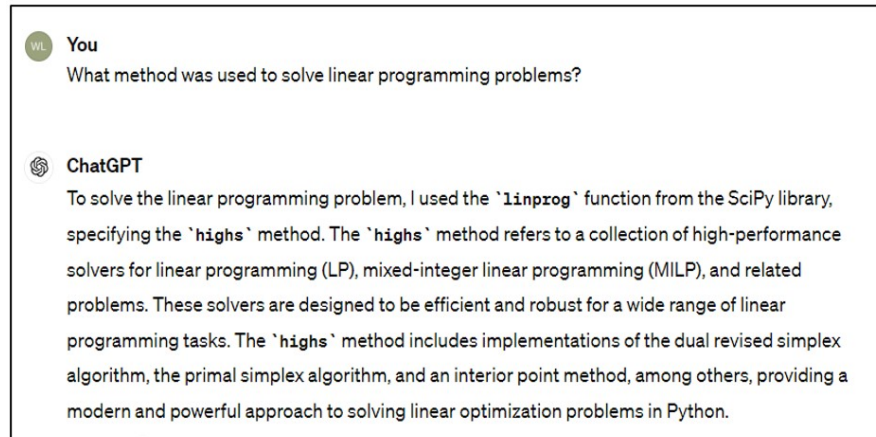


Fig. 2. Method used to solve the LP task

Scan 3 contains a command to draw the feasible region for the LP task. The region drawn by Chat GPT is not correct. One vertex of the feasible region is not included in the figure.

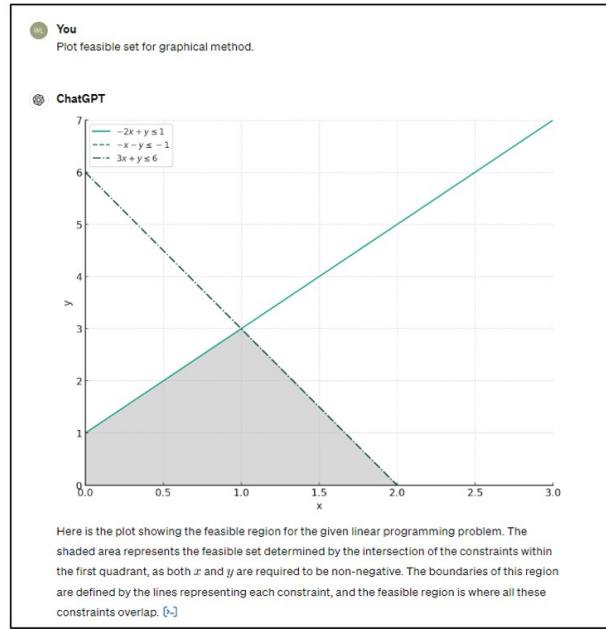


Fig. 3. Chat GPT feasible region for the LP task

The correct feasible region for the LP problem drawn in Mathematica is shown in Fig. 4. The graphical method of solving LP tasks is based on a correct drawing of the feasible region taking into account all the vertices of this region.

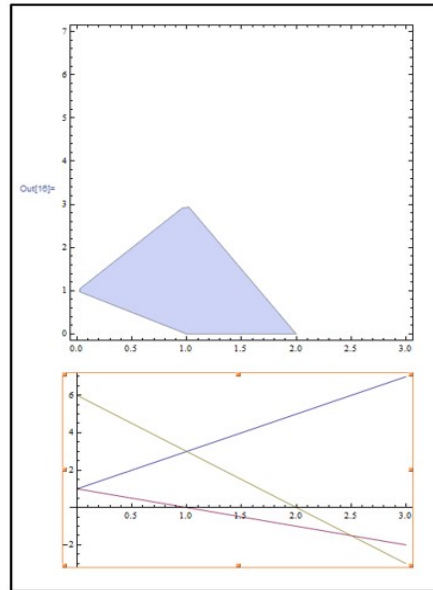


Fig. 4. Feasible region for the LP task drawn in Mathematica

Scan 5 presents the so-called canonical form of the LP task. It is required by the simplex algorithm and created from the standard form. Then, subsequent simplex tables are created until the optimal table is obtained in which the optimal solution and the optimal value of the objective function are determined.

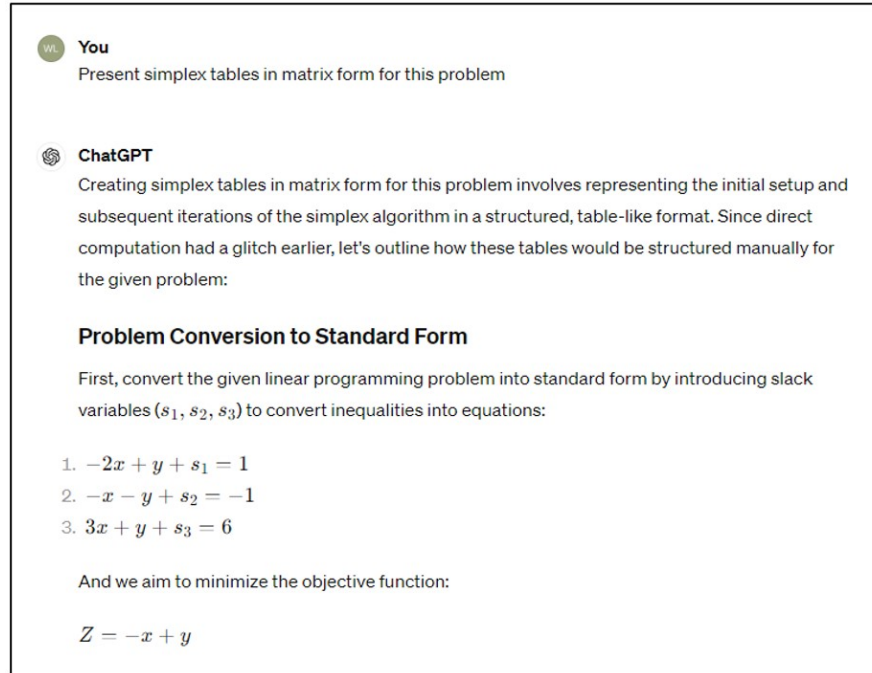


Fig. 5. Canonical form of the LP task

Scan 6 shows the initial simplex table presented by Chat GPT with commentary on the method of creating subsequent simplex tables. Despite additional commands, we were unable to "persuade" Chat GPT to create further simplex tables beyond the initial one.

To sum up: Chat GPT provided the correct solution of the LP task at the beginning. The drawing of the feasible region, which is the basis for the graphical method, was not correct. We were unable to create any simplex tables other than the first one using Chat GPT. Chat GPT provided rules for creating further simplex tables but did not create them.

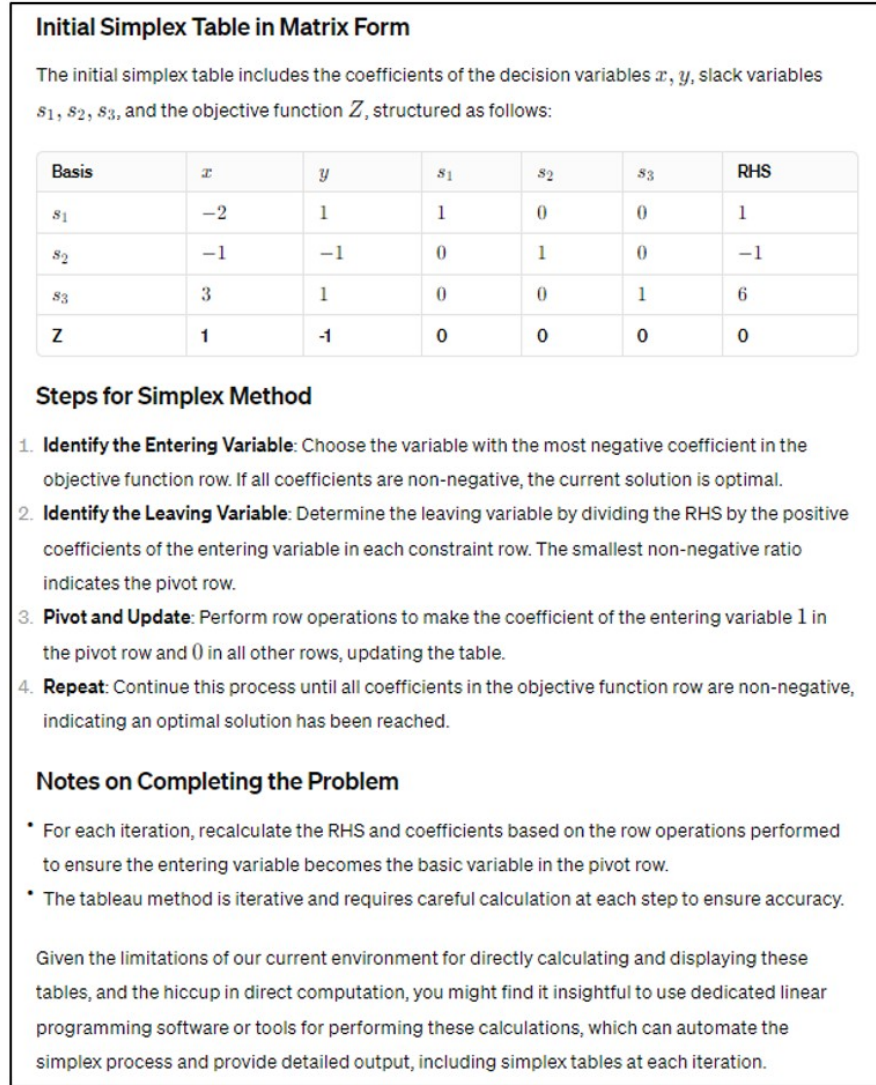


Fig. 6. Initial simplex table and commentary

4 Selected Chat GPT scans for the NLP task

Scan 7 presents the initial form of the NLP task for KKT method.

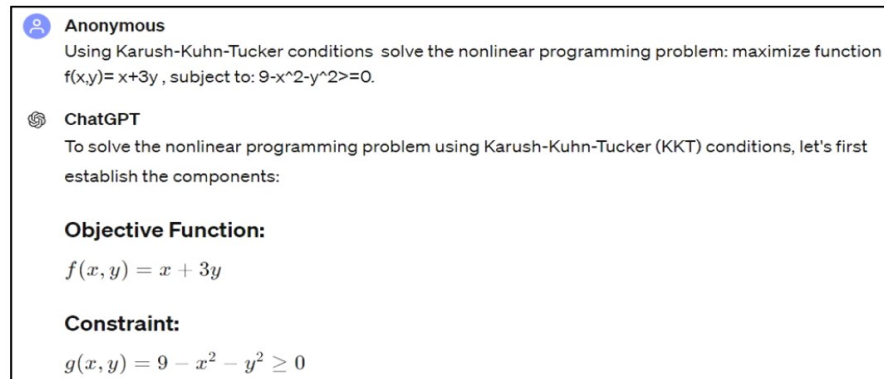


Fig. 7. The initial form of the NLP task

Scan 8 demonstrates the KKT conditions presented by Chat GPT. This is a kind of its own didactic commentary. There was no command on our part to present these conditions.

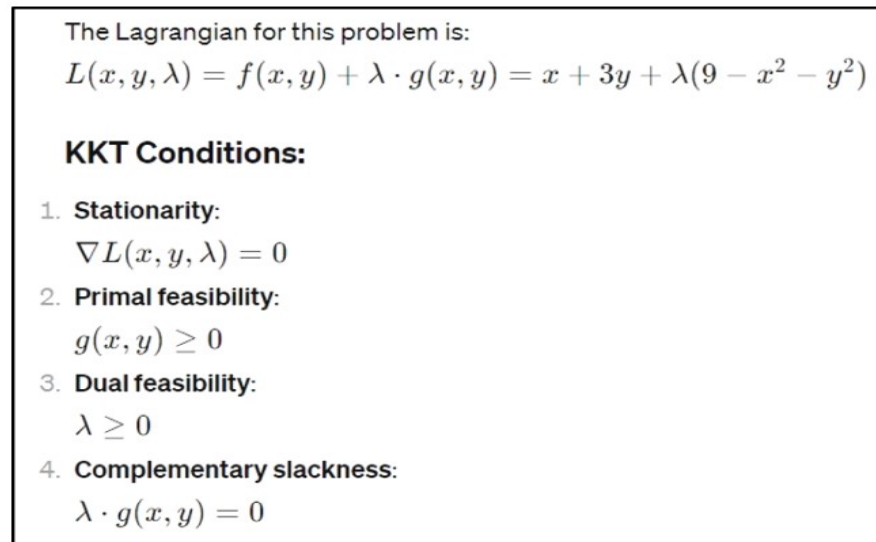


Fig. 8. KKT conditions presented by Chat GPT

Scan 9 presents two solutions of the system of equations from condition 1, only one of which meets condition 3. Details of solving the system of equations were presented by Chat GPT (We omitted this scan).

Compute Stationarity:

Solve the system of equations derived from the partial derivatives of L with respect to x, y , and λ .

Let's proceed to compute these derivatives and solve the resulting system of equations.

! Error analyzing ▾

The system of equations yields two sets of solutions for x, y , and λ :

1. $x = -\frac{3\sqrt{10}}{10}, y = -\frac{9\sqrt{10}}{10}, \lambda = -\frac{\sqrt{10}}{6}$
2. $x = \frac{3\sqrt{10}}{10}, y = \frac{9\sqrt{10}}{10}, \lambda = \frac{\sqrt{10}}{6}$

Again, the dual feasibility condition ($\lambda \geq 0$) allows only the second solution, as λ must be non-negative.

Thus, the optimal solution that maximizes the function $f(x, y) = x + 3y$ under the constraint $9 - x^2 - y^2 \geq 0$ is $x = \frac{3\sqrt{10}}{10}, y = \frac{9\sqrt{10}}{10}$.

Fig. 9. The solutions of the system of equations

Scan 10 shows compliance of the obtained solution with the remaining conditions: 4 and 2. The solution of the NLP task obtained by Chat GPT meets all four KKT conditions.

Verify with Complementary Slackness and Primal Feasibility:

- **Complementary Slackness:** $\lambda \cdot g(x, y) = 0$
- **Primal Feasibility:** $g(x, y) \geq 0$

We can compute the constraint $g(x, y)$ with the chosen x and y to confirm these conditions. [?]

The constraint function $g(x, y) = 9 - x^2 - y^2$ evaluates to 0 for the valid solution $x = \frac{3\sqrt{10}}{10}, y = \frac{9\sqrt{10}}{10}$, and the product of λ and $g(x, y)$ is also 0. This satisfies both the complementary slackness condition and the primal feasibility condition. Therefore, this solution indeed maximizes the function $f(x, y) = x + 3y$ under the given constraint, and all KKT conditions are satisfied. [?]

Fig. 10. Verify with complementary slackness and primal feasibility conditions

Scan 11 presents the GPT Chat's response to the command to present feasible region for the NLP task with the optimal solution.

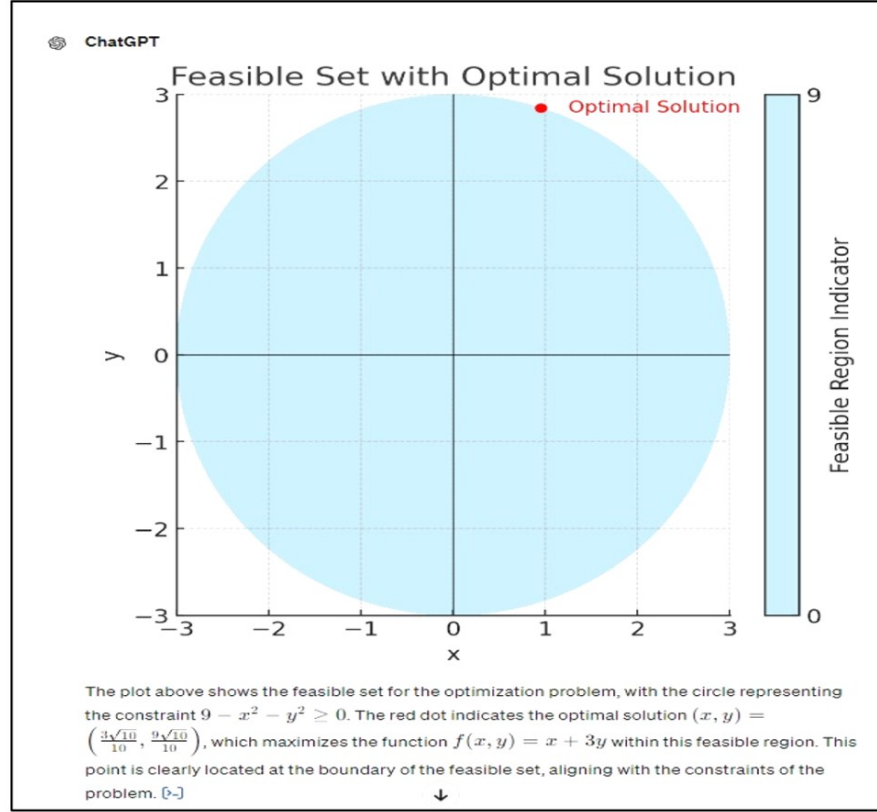


Fig. 11. Feasible region for the NLP task with the optimal solution

The drawing of the feasible region is correct. The obtained solution of the NLP task should be considered fully correct. It additionally contains the GPT Chat's own comments of a didactic nature.

The obtained solution is given in symbolic form. When we asked Chat GPT: what computational program was used by the Chat to obtain a response in symbolic form, the answer was: the Python library for symbolic computations called SymPy was used. Plots were created by Chat GPT using Matplotlib, a widely used library in Python for generating static interactive and animated visualizations.

5 Selected Chat GPT scans for the CP task

Scan 12 demonstrates the initial form of the CP task. We will use the method based on the fact that a local minimum of a strictly convex function in a convex region is also the global minimum of this function.

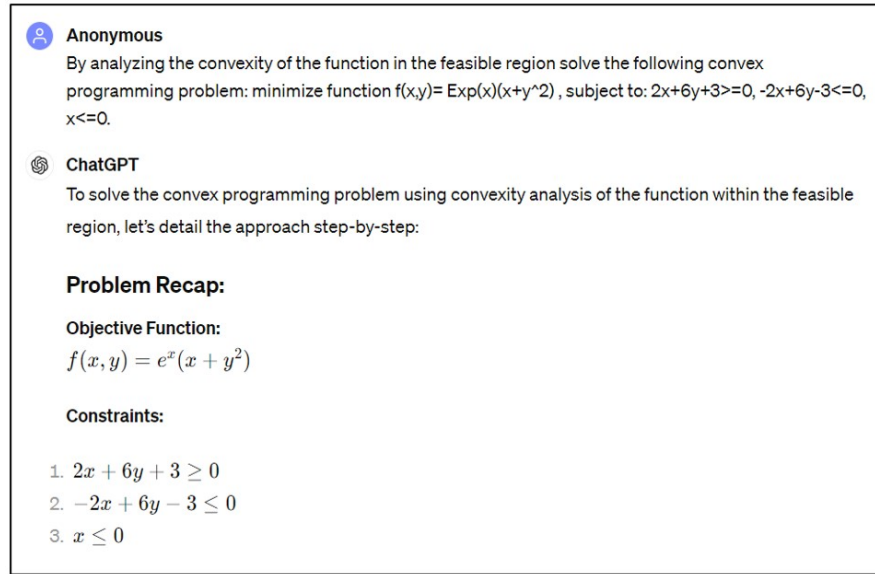


Fig. 12. The initial form of the CP task

Scan 13 contains the convexity analysis performed by Chat GPT with respect to the function $f(x, y)$ as well as feasible region of the CP task. The feasible region is convex and the function $f(x, y)$ is strictly convex in the feasible region. It additionally contains Chat's didactic commentary on the method of solving the task.

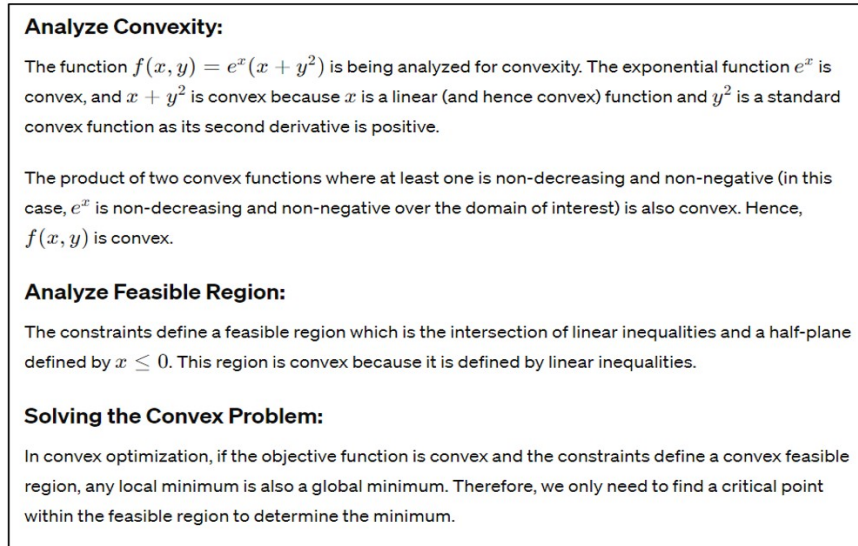


Fig. 13. Convexity analysis

Scan 14: due to a problem that occurred when comparing symbolically calculated objective function values, Chat GPT decided to choose a numerical method of calculating these values to determine which of them is the minimum.

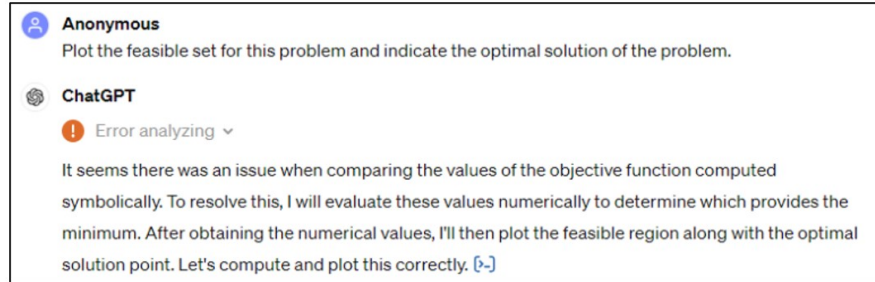


Fig. 14. Chat GPT decided to choose a numerical method

Scan 15 presents feasible region for the CP task with the optimal solution $(-1.03, -0.16)$ determined numerically. Symbolically determined coordinates of the minimum point are $(-1, 0)$.

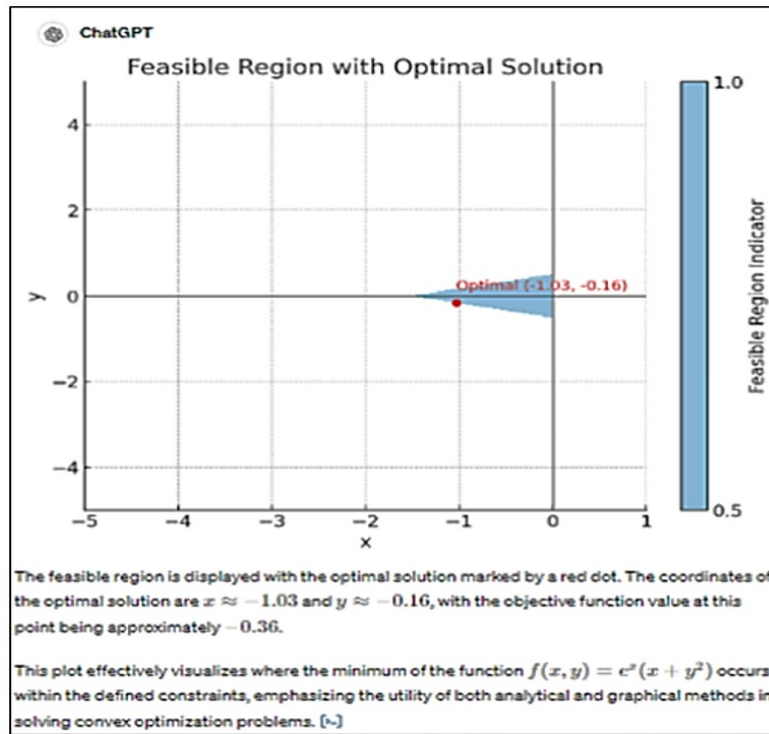


Fig. 15. Feasible region with optimal solution of CP task

The drawing of the feasible region is correct. But in reality, the optimal point lies inside the feasible region, not on its boundary.

Chat GPT examined the convexity of the feasible region and the objective function but determined the coordinates of the optimal point only approximately. Nevertheless, in the commentary it described a way to determine the global minimum as the local minimum of a strictly convex function on a convex set.

6 Didactic experiment

The three examples of using GPT Chat to solve mathematical programming (MP) tasks presented above were prepared by us. As a part of the MP course for students of Informatics and Econometrics of Warsaw University of Life Sciences, we solve a certain number of calculation tasks in the field of linear and nonlinear programming during lectures and exercises. The didactic experiment consisted of using Chat GPT to solve certain selected computational tasks previously solved during MP classes and assessing the usefulness of Chat GPT in solving these tasks. Six students attending MP classes and declaring their ability to use Chat GPT, volunteered for the experiment. We have not used GPT Chat before in MP classes. Each student participating in the experiment had to choose two tasks from the tasks we had previously solved during the exercises: one LP task and one NLP task and solve them using Chat GPT and then assess the usefulness of GPT Chat in solving these tasks. They used version 3.5 or 4.0 of Chat GPT. Students sent us scans of the solutions they received with their own comments. The students' experiences were similar to ours. Virtually all students had problems solving the LP task using the graphical method and/or the simplex algorithm with the help of Chat GPT. They had to formulate additional questions or suggestions in the form of a dialogue with Chat GPT to obtain final solutions. They did much better with the NLP task using the KKT method or the Lagrange multipliers method. In the NLP task, almost all students received the correct solution in the first attempt. In comments sent to us, students emphasized that Chat GPT is an imperfect tool for solving PM tasks and does not guarantee the correctness of the solutions obtained. On the other hand, they emphasized that the descriptive comments and tips formulated by Chat GPT were useful and educational for them.

7 Conclusions

Chat GPT is a new educational tool and in publications of recent years one can come across innovative attempts to use Chat GPT in education – testing Chat GPT in various examples of educational applications attempts to assess its didactic usefulness. This paper addresses the topic of the usefulness of Chat GPT as a didactic tool in learning and teaching optimization algorithms within such course subjects as Mathematical Programming or Mathematical Analysis. Three computational tasks were presented and analyzed: LP, NLP and CP task solved using Chat GPT. In the articles [18, 19], the authors discussed the topic of teaching MP using CAS programs such as Mathematica or wxMaxima. If we compare the approach in MP teaching based on the use of CAS with the approach based

on the use of Chat GPT, some similarities and differences are visible. Both CAS and Chat GPT programs use symbolic calculus and graphics to visualize mathematical objects. CAS programs such as Mathematica, however, are much more specialized when it comes to access to various mathematical tools. On the other hand, when comparing CAS and Chat GPT, the way the user communicates with the program is different. The advantage of using Chat GPT's help when solving mathematical tasks is undoubtedly the ease of formulating commands and questions. They only need to be formulated in a descriptive form using everyday language, unlike specialized CAS programs such as Mathematica, Wx-Maxima or Maple, which require the use of formal, precise notation. It is worth emphasizing that Chat GPT adds its own didactic comments to the solution of the task, describing the chosen solution method. It seems that such a way of presenting the issue could be useful for pupils and students in their own work and for teachers and lecturers as a material to use during lessons, exercises or lectures. Nevertheless, interactive dialogues with GPT Chat are often necessary to guide it to correct solutions.

As seen in the presented examples, Chat GPT can make substantive errors. It did not correctly determine the feasible region in the graphical method of the LP task, nor did it create subsequent simplex tables (only the initial one) for the simplex method in this task, limiting itself only to describing the rules for creating these tables. It was also visible that the solution to the task presented by Chat may depend on the current knowledge of Chat GPT on this subject. Students informed us that in the case of the LP task solved graphically, they had to formulate additional questions or commands to guide Chat to the right solution path. In the opinion of students taking part in the experiment, using the Chat GPT to solve MP tasks did not provide certainty as to the correctness of the solutions obtained. Students declared problems with obtaining the correct solution to the LP task using the graphical method as well as the simplex algorithm with the help of Chat GPT. On the other hand they found the comments and tips presented by the Chat useful. Taking all this into account, due to the possibility of Chat GPT giving incorrect answers (creating hallucinations), it would be difficult to consider Chat GPT in the currently available versions as a fully reliable tool that can be used to learn and teach advanced optimization issues in the field of mathematical programming or mathematical analysis. It may be no less interesting in this context to follow the development, changes and improvement of mathematical possibilities in subsequent versions of Chat GPT.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. von Hippel, P.T.: Chatgpt is not ready to teach geometry (yet). Education Next (2023), <https://www.educationnext.org/chatgpt-is-not-ready-to-teach-geometry-yet/>

2. Imran, M., Almusharraf, N.: Analyzing the role of chatgpt as a writing assistant at higher education level: a systematic review of the literature. *Contemp. Educ. Technol.* **15**(4), ep464 (2023)
3. Kleinman, Z.: Upgraded chatgpt teaches maths and flirts - but still glitches (2024), <https://www.bbc.com/news/articles/cv2xx1xe2evo>, accessed 28.01.2025
4. Long, P.P.V., Vu, D.A., Hoang, N.M., Do, X.L., TuanLuu, A.: Chatgpt as a math questioner? evaluating chatgpt on generating pre-university math questions. In: *Proceedings of ACMSAC Conference (SAC'24)*. ACM, New York, NY, USA (2024)
5. Memarian, B., Doleck, T.: Chatgpt in education: Methods, potentials, and limitations. *Computer i Human Behavior: Artificial Humans* **1**(2) (August-December 2023)
6. Milne, S.: Ai researcher discusses the new version of chatgpt's advances in math and reasoning. *UW News* (2024), <https://www.washington.edu/news/2024/09/17/ai-chatgpt-openai-math-reasoning-o1/>, accessed 28.01.2025
7. OpenAI: Data analysis with chat gpt (2024), <https://help.openai.com/en/articles/8437071-data-analysis-with-chatgpt>, accessed 28.01.2025
8. OpenAI: Doing math with openai models (2024), <https://help.openai.com/en/articles/6681258-doing-math-with-openai-models>, accessed 28.01.2025
9. OpenAI: Hello gpt-4o (2024), <https://openai.com/index/hello-gpt-4o/>, accessed 28.01.2025
10. Salih, S., Husain, O., Hamdan, M., Abdelsalam, S., Motwakel, H.E.A.: Transforming education with ai: A systematic review of chatgpt's role in learning, academic practices, and institutional adoption. *Results in Engineering* **25**, 103837 (2025)
11. Sanderson, K.: Gpt-4 is here: what scientists think. *Nature* **615**, 773 (2023)
12. Spreitzer, C. and Straser, O., Zehetmeier, S., Maab, K.: Mathematical modelling abilities of artificial intelligence tools: The case of chatgpt. *Educ. Sci.* **14**(7) (2024)
13. Stokel-Walker, C., Van Noorden, R.: What chatgpt and generative ai mean for science. *Nature* **614**(7947), 214–216 (2023)
14. Supriyadi, E., Kuncoro, K.: Exploring the future of mathematics teaching: Insight with chatgpt. *Union: Jurnal Ilmiah Pendidikan Matematika* **11**(2), 305–316 (2023)
15. Taani, O., Alabidi, S.: Chatgpt in education: benefits and challenges of chatgpt for mathematics and science teaching practices. *IJMEST* (2024)
16. Van Noorden, R., Webb, R.: Chatgpt and science: the ai system was a force in 2023 – for good and bad. *Nature* **624**, 509 (2023)
17. Wardat, Y., Tashtoush, M.A., AlAli, R., Jarrah, A.M.: Chatgpt: A revolutionary tool for teaching and learning mathematics. *EURASIA Journal of Mathematics, Science and Technology Education* **19**(7) (2023)
18. Wojas, W., Krupa, J.: Teaching students nonlinear programming with computer algebra system. *Math.Comput.Sci.* **13**, 297–309 (2019)
19. Wojas, W., Krupa, J.: Supporting education in algorithms of computational mathematics by dynamic visualizations using computer algebra system. In: *ICCS*. vol. 7, pp. 634–647 (2020)