Re71: *Gradients and Partial Derivatives Part 2 (AIMA4e pp. 119–122)*

retraice.com

Put the airport problem first.

Air date: Monday, 5th Dec. 2022, 11:00 PM Eastern/US.

*We're focusing on the math and code of AIMA4e[1] right now, December 2022. This is in service of our plan to deep-dive the book from Jan.–Jun., 2023. **DISCLAIMER: The below mathematics cannot be trusted; it is a student's attempt, not an expert's.***

The airport toy problem

Toy problems, like Legos, are good if and only if your mind is on the real world as you play, not the toy itself. A little imagination will suggest the many ways the airport problem,[2] properly reduced to numbers and math, resembles other more interesting problems in the world.[3] **Figures** for the airport problem, e.g. **3.1 and 3.20**, are available at aima.cs.berkeley.edu or via links.retraice.com. The simple version of the airport problem has two cities and one airport, from Retraice (2022/12/04):

Sums *of squares* reflect 'better' airport locations; sums *of distances* don't necessarily do so.

A pile of numbers

The full airport problem is a **continuous (not discrete) search problem** that can be **formulated (using vectors)** by representing the cities and airports as **points in a Cartesian coordinate system** so that tools from **calculus (partial derivatives)** can tell us how to find a **local-best solution** given an initial solution hypothesis. To represent the problem in this way is to view it as a 'pile of numbers'[4] in order to make it tractable using math and machines. We can imagine a 'landscape' of possible solutions, and trying to get to the highest point (or lowest valley) of that landscape, without the benefit of sight.[5]

The gradient equation again:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial y_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial y_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial y_3} \right)$$

Elements of the problem and a tool to help solve it:

- **The (x_i, y_i) values (the <u>i</u>ndependent variables) are the three pairs of solution coordinates for our airport locations**, represented as a vector $\mathbf{x} = (x_1, y_1, x_2, y_2, x_3, y_3)$.
- **The sum-of-squares objective function $f(\mathbf{x})$ generates a 'score'[6] (the <u>d</u>ependent variable)** that enables comparing different solutions (sets of values of the six coordinates) to the airport location problem;
- **The gradient vector of the objective function, ∇f, a tool to help us improve on guesses:** Given a *hypothesis*, i.e. a set of values for the six independent variables $(x_1, y_1, x_2, y_2, x_3, y_3)$ that represent the coordinates of the three airport locations, it tells us, in the form of a vector whose magnitude is [I don't know yet] and whose direction is [I don't know yet],[7] the 'direction' to go in our six-dimensional space[8] to improve our score to a local maximum (really, we should be saying local minimum, since we're trying minimize our objective function) by using calculus to find the 'slope' of the function $f(\mathbf{x})$ with respect to each of the six dimensions, i.e. the partial derivatives $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial y_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial y_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial y_3}$. □

---

[1] Russell & Norvig (2020).

[2] Russell & Norvig (2020) p. 120.

[3] E.g. How to find the best combination of delivery locations given dozens of refugee camps and only three air-dropped pallets of supplies? How to find the best combination of missile targets given dozens of enemy positions and only three missiles?

[4] Cf. Retraice (2022/03/07).

[5] The 'sight' analogue I mentioned during the livestream: transformer architecture and 'attention'. See aima4e.retraice.com on chpt. 24 of AIMA4e; Vaswani *et al.* 2017, *Attention Is All You Need*, arxiv.org/abs/1706.03762; and Russell & Norvig (2020) pp. 866, 868 ff.

[6] It is also common to think in terms of 'cost' instead of score; in our case, we want to minimize the 'cost' of the airport locations, or 'maximize' the score of our airport locations by considering lower numbers to be better 'scores', confusingly. Our maximum *utility* corresponds to the minimum value of the objective function.

[7] Each partial derivative looks like this:

$$\frac{\partial f}{\partial x_1} = \lim_{h \to 0} \frac{f(x_1 + h, x_2, \ldots, x_n) - f(\mathbf{x})}{h}$$

Deisenroth et al. (2020) p. 126. They also say that the gradient *is* the "generalization of the derivative to functions of several variables". So, we're going to have a row vector in $\mathbb{R}^{1 \times 6}$ composed of six scalars, and that then gets its magnitude from the distance to the origin, and its direction from the line through the origin and its coordinates? And we would continue in that direction by means of applying the gradient ∇f at $f(\mathbf{x})$ using what operation(s)?

[8] On thinking about objects in higher dimensions, see aima4e.retraice.com on Appendix A section A.2. and Retraice (2022/11/24).

© 2022 Retraice, Inc.

Last revised: December 6, 2022

References

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for Machine Learning*. Cambridge University Press. ISBN: 978-1108455145.
https://mml-book.github.io/ Searches:
https://www.amazon.com/s?k=9781108455145
https://www.google.com/search?q=isbn+9781108455145
https://lccn.loc.gov/2019040762

Retraice (2022/03/07). Re18: Plan of Attack. *retraice.com*.
https://www.retraice.com/segments/re18 Retrieved 25th Mar. 2022.

Retraice (2022/11/24). Re60: Complexity, Linear Algebra, Probability (AIMA4e Appendix A). *retraice.com*.
https://www.retraice.com/segments/re60 Retrieved 25th Nov. 2022.

Retraice (2022/12/04). Re70: Gradients and Partial Derivatives Part 1 (AIMA4e pp. 119–122). *retraice.com*.
https://www.retraice.com/segments/re70 Retrieved 5th Dec. 2022.

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. Pearson, 4th ed. ISBN: 978-0134610993. Searches:
https://www.amazon.com/s?k=978-0134610993
https://www.google.com/search?q=isbn+978-0134610993
https://lccn.loc.gov/2019047498

Last revised: December 6, 2022