

# Re72: Gradients and Partial Derivatives Part 3

## (AIMA4e pp. 119–122)

retraice.com

*Be in the math.*

Air date: Tuesday, 6th Dec. 2022, 11:00 PM Eastern/US.

We're focusing on the math and code of AIMA4e<sup>1</sup> 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's a student's attempt, not an expert's.**

We're working on **the airport problem**<sup>2</sup>: Where to locate three airports amongst several cities in Romania such that the sum of the squares of the straight-line distances between each city and its nearest airport is minimized? If we start with a pile of numbers (three pairs of coordinates for airports, and several pairs of coordinates of cities), our path through mathematics will involve algebraic symbols instead of those numbers, but in the end, we want to arrive back at a new pile of numbers for our solution.

The **solution vector**, three pairs of coordinates, one for each airport. Our guesses and our final answer will take this form.

$$\mathbf{x} = (x_1, y_1, x_2, y_2, x_3, y_3)$$

The **objective function**, the sum of the squares of the straight-line distances between each city and its nearest airport, a way of deciding which locations are better and worse. We want the value of this function to be as low as possible.

$$f(\mathbf{x}) = f(x_1, y_1, x_2, y_2, x_3, y_3)$$

The **gradient vector** of the objective function. This will give us, given an initial 'guess' (hypothesis) value of  $\mathbf{x}$ , the 'direction' of the steepest 'elevation' change in our function, which will lead us to a solution  $\mathbf{x}_m$  such that  $f(\mathbf{x}_m)$  is a local maximum (or minimum, if we use the negative) value. By itself, it cannot help us find a *global* maximum (or minimum) if our guess happens to be in the vicinity of a *local* maximum (or minimum)—the gradient can only tell us about the 'hill' we're on, not all the other hills in the solutions space.

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

The **partial derivative** (a 'rate' of change\*) of  $f$  with respect to the independent variable  $x_1$  at any given 'point'  $\mathbf{x}$ . To find the gradient vector  $\nabla f$  (i.e. the *total* derivative of  $f$  at point  $\mathbf{x}$ ) we'll need to calculate one of these partial derivatives for each of the six independent variables  $x_1, y_1, x_2, y_2, x_3, y_3$  of a proposed solution.†

$$\frac{\partial f}{\partial x_1} = \lim_{h \rightarrow 0} \frac{f(x_1 + h, y_1, x_2, y_2, x_3, y_3) - f(\mathbf{x})}{h}$$

\*On the subtleties of 'rates' and 'ratios' and the like, see [Margin \(2020/10/28\)](#), "A philosophical and arithmetical digression on satisfaction, ratios and counting."

† Adapted from the definition of a partial derivative given by [Deisenroth et al. \(2020\)](#), p. 126.

□

<sup>1</sup>Russell & Norvig (2020).

<sup>2</sup>Russell & Norvig (2020) p. 120.

## 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>

Margin (2020/10/28). Ma8: Revolution Before Evolution. *retraice.com*.  
<https://www.retraice.com/segments/ma8> Retrieved 30th Oct. 2020.

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>