Unconditional Optimization: Directional Derivative, Stationary Points, Convexity

Video: Introduction to the Week
1 min

Video: Directional Derivative: Definition
13 min

Video: Directional Derivative: Calculation
10 min

Video: Direction of Maximal Growth
4 min

Interactive Plot: Directional Derivative
15 min

Practice Quiz: Practice Quiz #1
4 questions

Video: Multivariate Extrema
13 min

Practice Quiz: Practice Quiz #2
5 questions

Introduction to Gradient Descent: Motivations, Step Length and Final Project

Video: Gradient Descent
12 min

Video: Introduction to the Final Project
10 min

Reading: Models and Parameters: Clearing the Air

Models and Parameters: Clearing the Air

Data Science always involves modeling your data. And what is a model?

Suppose, we have some data about an object with a certain *structure*. For example, it might be data about an apartment: *vector* $[2, 84.5, 6]$, where 2 is the number of rooms, 84.5 is the total area of the apartment and 6 is the floor number. Actually, we might have the data about N different apartments. Then we will represent them as a set of vectors $X = [x_1, ..., x_N]$, where each x_i is a set of m **features** of an **object**: $x_i = [x_{i1}, ..., x_{im}]$. X can also be considered as a **matrix**:

$$X = \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \dots & \dots & \dots \\ x_{N1} & \dots & x_{Nm} \end{pmatrix}$$

where x_{ij} is the j_{th} feature of the i_{th} object. $\mathbf{x_i}$ **is a vector of numeric features of the** $\mathbf{i_{th}}$ **object.**

Besides, we have some numeric characteristic for each of N objects $y = [y_1, ..., y_N]$ that we want to be able to calculate (or to **model**) having that object's data. $\mathbf{y_i}$ **is the result, the target or "ground truth" for the** $\mathbf{i_{th}}$ **object.**

That is the data from an actual example. That could be the actual price of the apartment, which we consider the true price. Given such data, we make an assumption:

$$\hat{y}_i = a(x_i) = w_0 + w_1 x_{i1} + ... + w_m x_{im} \quad (1)$$