ディープラーニングの仕組みを知ろう！

第 2 回　人工知能勉強会

Shion MORISHITA

June 24, 2024

目次

はじめに

目的

-

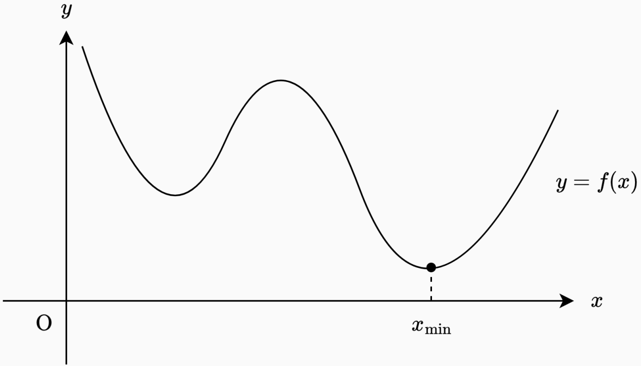
勾配降下法

勾配降下法

勾配降下法の基本概念

勾配降下法とその目的

- 機械学習や最適化の分野で広く用いられる最適化アルゴリズム
- 目的：最小化（または最大化）したい関数の最適なパラメータを見つけること



勾配降下法のアイデア

どのように関数が最小となるパラメータを見つけるか？

- 【重要】多変数関数の最小条件 を利用（第 1 回）
- 斜面を転がるボールのイメージ

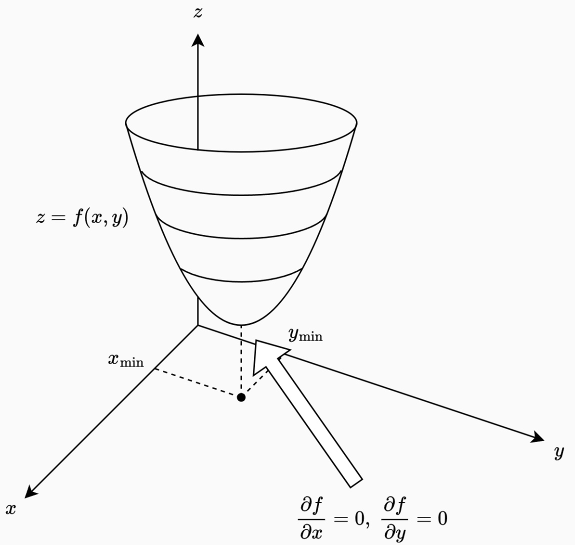
【重要】多変数関数の最小条件（第 1 回）

関数 $z = f(x, y)$ が最小になる必要条件は、$\dfrac{\partial f}{\partial x} = 0$ かつ $\dfrac{\partial f}{\partial y} = 0$

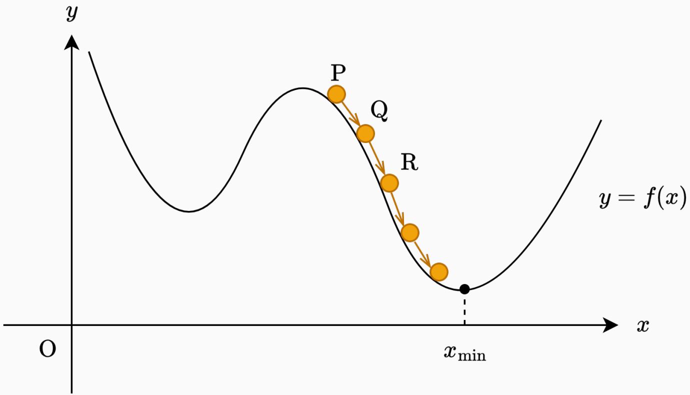
ポイント

どの成分から見ても傾きが 0 なら、最小値の可能性あり！

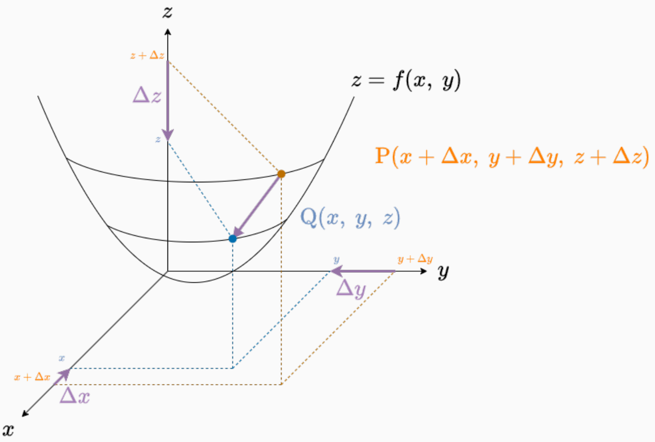
多変数関数の最小条件のイメージ



斜面を転がるボールのイメージ



斜面を転がるボール（多変数関数 ver.）



- 最速で転がる（Δz が最小になる）には $\Delta x,\ \Delta y$ をどう決める？

勾配降下法の基本式

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = -\eta \begin{bmatrix} \dfrac{\partial z}{\partial x} \\ \dfrac{\partial z}{\partial y} \end{bmatrix}$$ （η：正の小さな定数）

勾配降下法の基本式の導出 i

「関数の近似公式　簡潔 ver.」（第 1 回）より、

$$\Delta z \simeq \left\langle \begin{bmatrix} \dfrac{\partial z}{\partial x} \\ \\ \dfrac{\partial z}{\partial y} \end{bmatrix},\ \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right\rangle .$$

「ベクトルの基本公式」（第１回）の内積の式

$$\langle \boldsymbol{a},\ \boldsymbol{b} \rangle \triangleq \|\boldsymbol{a}\| \|\boldsymbol{b}\| \cos\theta$$

および、コーシー・シュワルツの不等式

$$-\|\boldsymbol{a}\| \|\boldsymbol{b}\| \leq \langle \boldsymbol{a}, \boldsymbol{b} \rangle \leq \|\boldsymbol{a}\| \|\boldsymbol{b}\|$$

勾配降下法の基本式の導出 ii

より、内積が最小となるのは $\cos\theta = -1$ のとき、すなわち、ベクトルの向きが反対のとき（$\theta = 180°$）。

ベクトルの向きが反対というのは、ベクトルの符号が異なるという意味なので、

$$\boldsymbol{a} = -k\boldsymbol{b} \quad （k：正の定数）$$

と表せる。今回の表記に合わせれば、

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = -\eta \begin{bmatrix} \dfrac{\partial z}{\partial x} \\ \dfrac{\partial z}{\partial y} \end{bmatrix} \quad （\eta：正の定数）$$

と導かれる。 □

結局どういうこと??

（勾配降下法の基本式の図的なイメージをここに載せる）

【参考】ギリシャ文字一覧

文字	名称	文字	名称
α	アルファ	ν	ニュー
β	ベータ	ξ	グザイ
γ	ガンマ	o	オミクロン
δ	デルタ	π	パイ
ϵ	イプシロン	ρ	ロー
ζ	ゼータ	σ	シグマ
η	イータ	τ	タウ
θ	シータ	υ	ウプシロン
ι	イオタ	ϕ	ファイ
κ	カッパ	χ	カイ
λ	ラムダ	ψ	プサイ
μ	ミュー	ω	オメガ