

2024 年度 修士論文

機械学習を用いた イジング模型の臨界現象に関する研究

2024 年 2 月 00 日

指導教員 藤原高德

茨城大学大学院
理工学研究科 量子線科学専攻

学籍番号 22NM021S

氏名 須賀 勇貴

要旨

研究の要旨を書く.

目次

第 1 章	はじめに	3
1.1	研究背景	3
1.2	研究目的	3
第 2 章	イジング模型における相転移	4
2.1	相転移, 臨界現象とは	4
2.2	強磁性イジング模型	5
2.3	一次元イジング模型	9
2.4	二次元イジング模型	10
2.5	二次元イジング模型の高温展開	10
第 3 章	機械学習と深層学習	12
3.1	機械学習の定義	12
3.2	統計入門	14
3.3	深層学習とは	15
3.4	ボルツマンマシンの基礎	15
第 4 章	イジング模型の相転移検出に関する研究	16
4.1	相の分類器による相転移検出	16
4.2	温度測定器による相転移検出	16
4.3	相転移検出がなぜ可能なのか?	16
第 5 章	機械学習と繰り込み群	17
5.1	繰り込み群と特徴抽出	17
第 6 章	まとめ	18
	謝辞	19
	参考文献	20

付録 A	aaa	21
A.1	bbb	21

第 1 章

はじめに

1.1 研究背景

1.2 研究目的

第 2 章

イジング模型における相転移

2.1 相転移，臨界現象とは

2.1.1 物質の三態と相転移

私たちにとって身近な物質である水 (H_2O) は，固体 (氷)，液体 (水)，気体 (水蒸気) の 3 つの状態をとる．これら状態は温度の違いによって明確な規則性があり，1 気圧の環境であれば，氷を徐々に温めていくと，ちょうど 0°C で氷から水へ相転移がおき，さらにその水を温めておけば，ちょうど 100°C で水から水蒸気への相転移が起きる．特に 1 気圧の環境では，1 気圧の環境では，水と氷が共存するのはちょうど 0°C に限られ，水と水蒸気が共存するのはちょうど 100°C に限られる．これが摂氏という温度目盛りの基準になっていることは言うまでもない．

液体の水の温度が 0°C から 100°C に変化する間，密度や粘性などの水の物性は少しずつ変化する．しかし，これはあくまでも定量的な変化であり，水から氷，水から水蒸気への変化のような定性的な (あるいは質的な) 変化とは違う．氷と水蒸気のそれぞれの範囲内での変化も，やはり質的な変化を伴わない，定量的な変化である．このように，定性的な変化を伴わずに移り変われるような一連の状態をひとまとめにして，相 (Phase) と呼ぶ．氷，水，水蒸気はそれぞれ固相，液相，気相という三つの相に対応する．温度などのパラメータを変化させたときに，物質が異なった相の間を移り変わる現象を相転移 (phase transition) である．

なぜ同じ H_2O という物質が，温度を変えただけで，固相，液相，気相という全く性質の異なる状態をとるのか？．素朴に考えれば，このような変化は「部品」である． H_2O 分子の性質の変化からくると考えたい．つまり，何らかの意味での「ミクロなルール」が変化することを想定するということである．

しかし，相転移が起こるのは「ミクロなルール」が変化するからではない．「ミクロなルール」が不変であっても，刑を構成する要素 (今回の場合なら分子) の数がきわめて大きければ，それら相互の関連が変化することで，マクロな性質の不連続な変化が生じうる．つまり，相転移は，無数の要素が複雑に絡み合ったときに全体として生じる協力現象の一種なのである．

2.2 強磁性イジング模型

この説では、強磁性体の相転移を調べるために、強磁性イジング模型を定義する。イジング模型は、実在の強磁性体のモデルとしては全く忠実ではないが、教師全体での相転移の本質をつかむためには、きわめてすぐれたモデルである。つまり、相転移や臨界現象を引き起こすために必要最低限の要素だけを持ったモデルと言える。

「イジング模型」という呼び名は、このモデルの一次元でのふるまいを 1924 年の学位論文で調べたイジングにちなんだものである。ただし、モデルの発案者は当時イジングの指導教員であったレンツである。

2.2.1 モデルの定義

一辺が L の d 次元立方格子を考える。全格子点の数を $N = L^d$ とし、それぞれの格子点に $i = 1, 2, \dots, N$ と番号をつけておく。各格子点には、上向きと下向きの二つの状態をとるスピンののっている。格子点 i のスピンを表すスピン変数を $\sigma_i = \pm 1$ とする。+1 が上向きスピン、-1 が下向きスピンのに対応する。また、スピン変数 σ_i に対応する物理量を $\hat{\sigma}_i$ と書く。

系のエネルギー固有状態は、すべての格子点のスピン変数 $\{\sigma\} = (\sigma_1, \sigma_2, \dots, \sigma_N)$ と列挙することで指定できる、このようなスピンの並びのことをスピン配位と呼ぶ。それぞれのスピンの二通りの状態をとるため、系の全状態数、あるいは、スピン配位の総数は 2^N である。スピン配位 $\{\sigma\}$ に対応するエネルギー固有値を

$$E(\{\sigma\}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - \mu_0 H \sum_{i=1}^N \sigma_i \quad (2.1)$$

とする。第一項はスピン間の相互作用を表す項で、第二項は外部磁場 H とスピン磁気モーメント μ_0 の相互作用を表す項である。第一項の和 $\sum_{\langle i,j \rangle}$ は互いに隣り合う格子点 i, j すべてについての和という意味である。

ここでは相互作用定数 J の値は正とする。したがって、隣り合う格子点の組 i, j に関わる相互作用を取り出すと、

$$-J\sigma_i\sigma_j = \begin{cases} -J, & \sigma_i = \sigma_j \text{ のとき} \\ J, & \sigma_i \neq \sigma_j \text{ のとき} \end{cases} \quad (2.2)$$

となる。つまり、隣り合う格子点のスピンの揃う方がエネルギーが小さくなる。このように互いにスピンを揃えようとする相互作用を、強磁性的相互作用という。

2.2.2 平衡状態での物理量

ここで、イジング模型の逆温度 β での平衡状態を調べる。この場合、カノニカル分布で平衡状態を記述するのが自然である。

分配関数は

$$Z_L(\beta, H) = \sum_{(\sigma_1, \dots, \sigma_N)} \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \quad (2.3)$$

である。和は 2^N 通りのすべてのスピン配位についてとる。後の便利のために格子サイズを L とした。物理量 \hat{g} の期待値は

$$\langle \hat{g} \rangle_{\beta, H}^{\text{can}} := \frac{1}{Z_L(\beta, H)} \sum_{(\sigma_1, \dots, \sigma_N)} g_{(\sigma_1, \dots, \sigma_N)} \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \quad (2.4)$$

である。ここで、 $g_{(\sigma_1, \dots, \sigma_N)}$ は状態 $(\sigma_1, \dots, \sigma_N)$ における \hat{g} の値である。スピン一つあたりの自由エネルギーを

$$f_L(\beta, H) := -\frac{1}{\beta H} \ln Z_L(\beta, H) \quad (2.5)$$

，物理量としての磁化を

$$\hat{m} := \frac{1}{N} \sum_{j=1}^N \mu_0 \hat{\sigma}_j \quad (2.6)$$

と定義する。このとき、磁化の期待値は

$$\begin{aligned} m_L(\beta, H) &:= \langle \hat{m} \rangle_{\beta, H}^{\text{can}} \\ &= \frac{1}{Z_L(\beta, H)} \sum_{(\sigma_1, \dots, \sigma_N)} m_{(\sigma_1, \dots, \sigma_N)} \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \\ &= \frac{1}{Z_L(\beta, H)} \sum_{(\sigma_1, \dots, \sigma_N)} \frac{1}{N} \sum_{j=1}^N \mu_0 \sigma_j \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \\ &= \frac{1}{Z_L(\beta, H)} \sum_{(\sigma_1, \dots, \sigma_N)} \frac{1}{\beta N} \frac{\partial}{\partial H} \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \\ &= \frac{1}{\beta N} \frac{1}{Z_L(\beta, H)} \frac{\partial}{\partial H} Z_L(\beta, H) \\ &= \frac{1}{\beta N} \frac{\partial}{\partial H} \ln Z_L(\beta, H) \\ &= -\frac{\partial}{\partial H} f_L(\beta, H) \end{aligned} \quad (2.7)$$

と書ける。これ以降、期待値 $\langle \hat{m} \rangle_{\beta, H}^{\text{can}}$ のことも、単に磁化と呼ぶ。磁化とは「系がどの程度磁石になっているか」の目安である。それに対して、「系がどの程度磁石になりやすいか」の目安になるのがゼロ磁場での磁化率は

$$\chi_L(\beta) := \left. \frac{\partial m_L(\beta, H)}{\partial H} \right|_{H=0} \quad (2.8)$$

である。

2.2.3 絶対零度

まずは、絶対零度での系のふるまいを見ていく。つまり、基底状態を求めるということである。

式 (2.2) から、一つのスピンに着目すれば、 $\sigma_i = \sigma_j$ のときエネルギーが最小になる。したがって、相互作用項 $-J\sigma_i\sigma_j$ を最小化する状態は、すべてのスピンの値が等しくなっているときであることがわかる。つまり、すべての i について $\sigma_i = +1$ もしくは $\sigma_i = -1$ のいずれかの状態である。

一方、外部磁場とスピンの相互作用項 $-\mu_0 H \sigma_i$ を最小化する状態は、磁場の符号によって変わってくる。式 (2.1) の第二項を最小化するのは、 $H > 0$ なら、すべての i について $\sigma_i = 1$ とした状態であり、 $H < 0$ なら、すべての i について $\sigma_i = -1$ とした状態である。 $H = 0$ のときは、値は常に 0 になるため、すべてのスピン配位で最小値を与える。

以上をまとめると、 $H \geq 0$ のときはすべての i について、 $\sigma_i = +1$ としたものが基底状態となり、 $H \leq 0$ のときはすべての i について、 $\sigma_i = -1$ としたものが基底状態となる (図 2.1)。



図 2.1: 絶対零度でのイジングモデルの基底状態。 $H > 0$ のときはスピンはすべて上向きになり、 $H < 0$ のときはスピンがすべて下向きになる $H = 0$ のときは、これら両方の状態が基底状態になる。

これに基づいて、絶対零度での磁化のふるまいを見てみる。すべてのスピンの値が $+1$ の状態での磁化は μ_0 、すべてのスピンの値が -1 の状態での磁化は $-\mu_0$ なので、

$$m_L(\infty, H) = \begin{cases} -\mu_0, & H \leq 0 \text{ のとき} \\ \mu_0, & H \geq 0 \text{ のとき} \end{cases} \quad (2.9)$$

とわかる。(図 2.2) つまり、磁化は H の関数として不連続である。もちろん、これはエネルギーを最小化する状態が入れ替わったことの単純な反映にすぎない。これに対して、同じような不連続性が 0 でない温度で見られるかどうかは、本質的に難しい問題である。

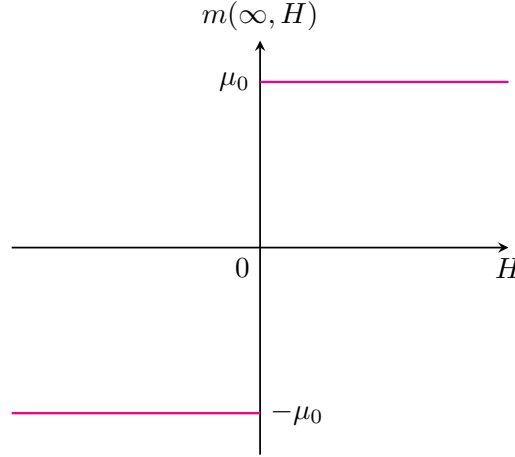


図 2.2: 絶対零度でのイジングモデルの磁化のふるまい. $H = 0$ を境に, $-\mu_0$ から μ_0 に不連続に変化する.

2.2.4 無限体積の極限

有限温度での相転移や臨界現象を調べるには, 自由エネルギー $f_L(\beta, H)$ や磁化 $m_L(\beta, H)$ といった物理量を計算し, それらの振る舞いを見ればよいと思える. しかし, このとき格子サイズ L はどの程度の大きさにとればよいのか?. 実は格子サイズが有限である場合, 相転移を起こさないことが示せる. ここではこれを証明する.

まず, 格子サイズ L が任意の有限な整数とする. このとき, 式 (2.5) の形から明らかなように自由エネルギー $f_L(\beta, H)$ は β, H について何度でも微分できる. よって, 磁化 $m_L(\sigma, H)$ は H の連続関数であることがわかる. そして, 系の対称性を考える. あるスピン配位 $(\sigma_1, \dots, \sigma_N)$ が与えられたとき, これらすべてのスピンを反転させた新しいスピン配位を $(\tilde{\sigma}_1, \dots, \tilde{\sigma}_N)$ で定義する. つまり, すべての i に対して $\tilde{\sigma}_i := -\sigma_i$ とするということである. このとき, エネルギーの表式 (2.1) をスピン反転した変数を使って書き直せば, $\tilde{\sigma}_i \tilde{\sigma}_j = \sigma_i \sigma_j$ が成り立つことを使って

$$E_{(\sigma_1, \dots, \sigma_N)} = -J \sum_{\langle i, j \rangle} \tilde{\sigma}_i \tilde{\sigma}_j - H \sum_{i=1}^N \tilde{\sigma}_i \quad (2.10)$$

のように, 磁場を反転させた, 元のエネルギーと同じ表式が得られる. つまり, すべてのスピンを反転させることは, 磁場を反転させることと等価ということである. すべての $(\sigma_1, \dots, \sigma_N)$ について足し上げることは, すべての $(\tilde{\sigma}_1, \dots, \tilde{\sigma}_N)$ について足し上げることと同じなので, 分配関数の

表式 (2.3) を,

$$\begin{aligned}
 Z_L(\beta, H) &= \sum_{(\tilde{\sigma}_1, \dots, \tilde{\sigma}_N)} \exp[-\beta E_{(\sigma_1, \dots, \sigma_N)}] \\
 &= \sum_{(\tilde{\sigma}_1, \dots, \tilde{\sigma}_N)} \exp \left[\beta \left(J \sum_{\langle i, j \rangle} \tilde{\sigma}_i \tilde{\sigma}_j - H \sum_{i=1}^N \tilde{\sigma}_i \right) \right] \\
 &= \sum_{(\sigma_1, \dots, \sigma_N)} \exp \left[\beta \left(J \sum_{\langle i, j \rangle} \sigma_i \sigma_j - H \sum_{i=1}^N \sigma_i \right) \right] \\
 &= Z_L(\beta, -H)
 \end{aligned} \tag{2.11}$$

となり、磁場を反転させても分配関数は変わらないという結果になる。自由エネルギーの定義である式 (2.5) より

$$f_L(\beta, H) = f_L(\beta, -H) \tag{2.12}$$

という自由エネルギーの対称性が示せる。さらに磁化は式 (2.7) より、 $f_L(\beta, H)$ の H 微分で書けるので、

$$m_L(\beta, H) = -m_L(\beta, -H) \tag{2.13}$$

のように書け、磁化は磁場の反転について反対称になることがわかる。

$f_L(\beta, H)$ が微分可能なので、 $m_L(\beta, H)$ はすべての β, H において定義されている。よって式 (2.13) で $H = 0$ とすれば、 $m_L(\beta, 0) = -m_L(\beta, 0)$ となり、 $m_L(\beta, 0) = 0$ が得られる。したがって、有限系では $0 \leq \beta < \infty$ の任意の逆温度 β で磁化がゼロになり、相転移を示さないことがわかる。

2.3 一次元イジング模型

一次元のイジング模型を考える。ハミルトニアンは

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - h \sum_{i=1}^L \sigma_i \tag{2.14}$$

なり、周期的境界条件 $\sigma_{N+1} = \sigma_1$ を課す。分配関数は

$$Z = \sum_{(\sigma_1, \dots, \sigma_N)} \exp \left[\beta J \sum_{i=1}^L \sigma_i \sigma_{i+1} + \beta h \sum_{i=1}^N \sigma_i \right] \tag{2.15}$$

となる。ここで、直交行列 T を定義する。

$$(T)_{\sigma_i, \sigma_{i+1}} = \exp \left[\beta J \sigma_i \sigma_{i+1} + \frac{\beta h}{2} (\sigma_i + \sigma_{i+1}) \right], \tag{2.16}$$

$$T = \begin{pmatrix} (T)_{1,1} & (T)_{1,-1} \\ (T)_{-1,1} & (T)_{-1,-1} \end{pmatrix} = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta h} \end{pmatrix} \tag{2.17}$$

このようにすることで、分配関数 (2.15) を

$$\begin{aligned} Z &= \sum_{(\sigma_1, \dots, \sigma_L)} (T)_{\sigma_1, \sigma_2} (T)_{\sigma_2, \sigma_3} \cdots (T)_{\sigma_N, \sigma_1} \\ &= \sum_{(\sigma_1, \dots, \sigma_L)} \prod_{i=1}^N (T)_{\sigma_i, \sigma_{i+1}} \end{aligned} \quad (2.18)$$

を書くことができる。スピン変数はそれぞれ ± 1 の値をとるため、 (σ_i, σ_{i+1}) の組のとりうる値は 4 通りである。これを次のように行列で表す。

$$\hat{T} = \begin{pmatrix} T(1, 1) & T(1, 2) \\ T(2, 1) & T(2, 2) \end{pmatrix} = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta h} \end{pmatrix} \quad (2.19)$$

式 (2.14) を

$$E_{(\sigma_1, \dots, \sigma_L)} = - \sum_{i=1}^L \left\{ J \sigma_i \sigma_{i+1} + h \frac{\sigma_i + \sigma_{i+1}}{2} \right\} \quad (2.20)$$

と書き換える。これより、分配関数 (2.15) は、

$$\begin{aligned} z_L(\beta, H) &= \sum_{(\sigma_1, \dots, \sigma_L)} \prod_{i=1}^L \exp \left[\beta \left(J \sigma_i \sigma_{i+1} + H \frac{\sigma_i + \sigma_{i+1}}{2} \right) \right] \\ &= \sum_{\sigma_1 = \pm 1} \sum_{\sigma_2 = \pm 1} \cdots \sum_{\sigma_L = \pm 1} \prod_{i=1}^L (M)_{\sigma_i, \sigma_{i+1}} \end{aligned} \quad (2.21)$$

と書ける。ここで、 $\sigma, \sigma' = \pm 1$ について、

$$(M)_{\sigma, \sigma'} := \exp \left[\beta \left(J \sigma \sigma' + H \frac{\sigma + \sigma'}{2} \right) \right] \quad (2.22)$$

という量を定義した。

σ と σ' はそれぞれ ± 1 の二つの値をとるため、四つの量が定義されている。これらを成分とした実対称行列

$$M = \begin{pmatrix} e^{\beta(J+H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-H)} \end{pmatrix} \quad (2.23)$$

を

2.4 二次元イジング模型

2.5 二次元イジング模型の高温展開

二次元イジング模型の厳密解は Onsager によって求められた。Onsager は、転送行列を対角化することで磁場のないときの自由エネルギーを求め、比熱をある温度で発散することを示した。いくつかの解法があるが、ここでは、高温展開を用いた解法を説明する。高温展開は、温度が高いと

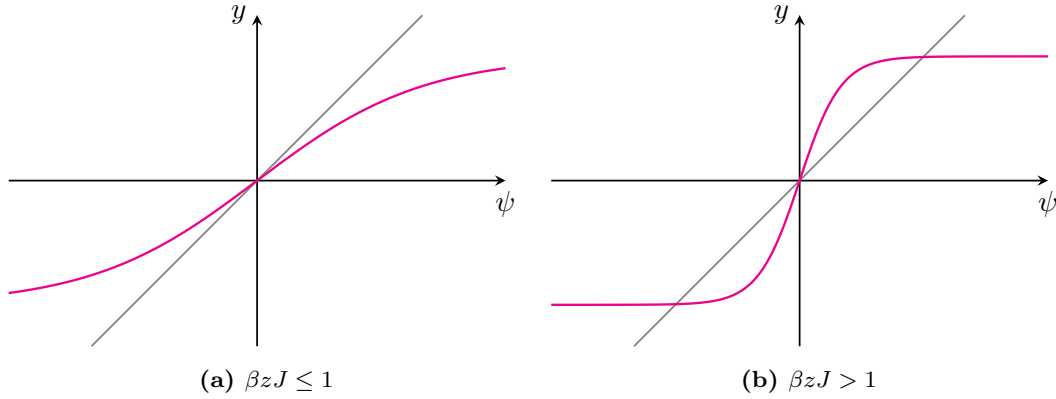


図 2.3: $h = 0$ での自己整合方程式 $\psi = \tanh(\beta zJ\psi)$

して自由エネルギーを βJ のべきで展開する方法である。十分高温であれば、その展開を数項で打ち切って自由エネルギーの近似値とする。計算が比較的簡単なため、汎用的な手法として用いられているが、近似の正当性に十分に注意する必要がある。この近似のみを用いて相転移を調べることはできない。有限項の和から特異性が生じることはないからである。本節では、無限和を計算することにより厳密解を求め、特異性が生じることを示す。

2.5.1 高温展開

$h = 0$ の場合に高温展開を行う。二次元イジング模型の分配関数は

$$Z = \text{Tr} \prod_{\langle i,j \rangle} \exp(\beta J \sigma_i \sigma_j) \quad (2.24)$$

と書ける。積はスピンの最近近接についてとる。スピン変数に関する恒等式

$$\begin{aligned} e^{x\sigma_i\sigma_j} &= \frac{e^x + e^{-x}}{2} + \sigma_i\sigma_j \frac{e^x - e^{-x}}{2} \\ &= \cosh x + \sigma_i\sigma_j \sinh x \end{aligned} \quad (2.25)$$

を用いれば,

$$\begin{aligned} Z &= \text{Tr} \prod_{\langle i,j \rangle} (\cosh \beta J + \sigma_i\sigma_j \sinh \beta J) \\ &= (\cosh \beta J)^{N_B} \text{Tr} \prod_{\langle i,j \rangle} (1 + \sigma_i\sigma_j \tanh \beta J) \end{aligned} \quad (2.26)$$

と書ける。 N_B は最近接の数を表す。 $v = \tanh \beta J$ は有限温度で 1 より有限温度で 1 より小さい非負の量なので、(2.26) を v について次のように展開する。

$$\prod_{\langle i,j \rangle} (1 + v\sigma_i\sigma_j) = 1 + v \sum_{\langle i,j \rangle} \sigma_i\sigma_j + v^2 \sum_{\langle i,j \rangle, \langle k,l \rangle} \sigma_i\sigma_j\sigma_k\sigma_l + \cdots \quad (2.27)$$

ここで、 $\langle i,j \rangle \neq \langle k,l \rangle$ である。

第3章

機械学習と深層学習

ここでは、イジングモデルに対して機械学習、深層学習による手法を応用する上で必要となる機械学習と深層学習の基礎知識について説明する。

3.1 機械学習の定義

機械学習 (machine learning) とは、人間がこなすような学習や知的作業を計算機に実行させるためのアプローチの研究、あるいはその手法そのもののことを意味する。機械学習では、知識を人間が直接アルゴリズムに具体的に書き込んだり教え込んだりせず、データという具体例の集まりから計算機に自動的に学ばせるという方法をとる。

機械学習の定義として T. M. ミッチェル (Tom Michael Mitchell) の書籍 [1] で書かれている定義が有名である。それは

”コンピュータプログラムが、ある種のタスク T とパフォーマンス評価尺度 P において、経験 E から学習するとは、タスク T におけるその性能を P によって評価した際に、経験 E によってそれが改善されている場合である。”

T. M. ミッチェル

である。ここで、タスク T とは解きたい問題のこと、パフォーマンス評価尺度 P は精度、誤差率などの評価指標のこと、経験 E はデータセットのことを指す。

機械学習では、何かしらのモデルを構築する必要がある。このモデルというのは、なにかしらの入力を与えられたときに出力を返すものであり、その正体は多数のパラメータを持った非線形関数 $f_{\theta}(x)$ 、 $\theta = \{\theta_1, \theta_2, \theta_3, \dots\}$ である。このパラメータの値は最初はランダムに初期化されており、

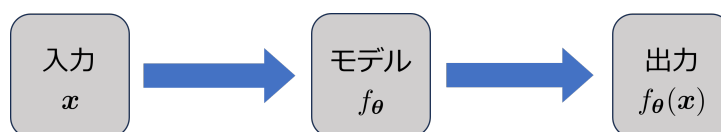


図 3.1: モデルの概略図

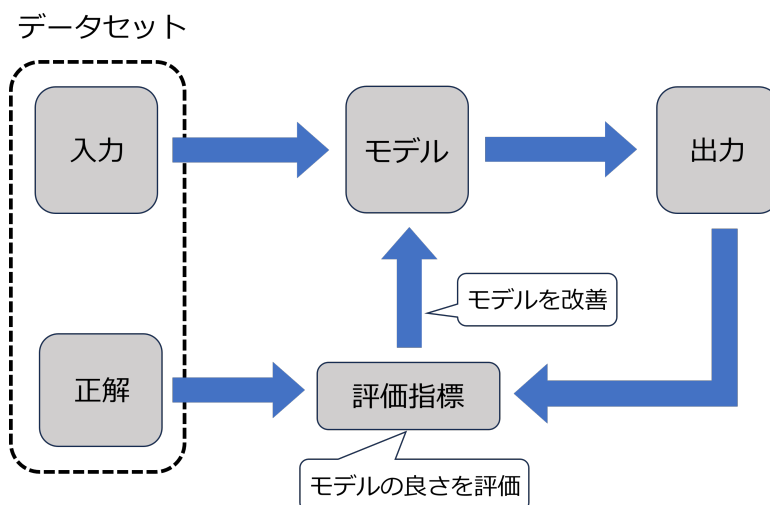


図 3.2: 機械学習の大まかな流れ

入力を与えると出力が返ってくるが、このときの出力結果は本来出力してほしい正解の値からは外れた値が帰ってくる。そこで、出力と正解の誤差を測るような評価基準を決めておき、算出された評価結果を良くするようにモデルのパラメータを更新してあげることによってモデルを改善していく。この一連の流れを繰り返し行うことで、モデルを改善して、正解に近い出力結果を得られるモデルを作成すること機械学習とよぶ。ここで注目してほしいのは、機械学習ではモデルの改善を機械に行わせているという点である。

メトロポリス法とメトロポリス・ヘイスティング法は同じですか

3.1.1 代表的なタスク

ここでは、機械学習で解きたい問題であるタスクの代表的な例として「クラス分類」と「回帰」を紹介する。

クラス分類

分類 (classification) とは、データをいくつかのクラスに仕分ける作業のことである。例えば送られてきた電子メールをみて、それがスパムメールか通常メールを判別するような作業は、2 クラス分類と呼べる。ここでのクラスというのは「スパムメール or 通常メール」といった分類先のことを指す。ここで、スパムメールを C_0 、通常メールを C_1 とラベル付けしよう。また、入力データを \mathbf{x} とする。このとき、クラス分類というのは、与えられた \mathbf{x} が C_0 と C_1 のどちらに属するかを決定する作業である。数値的な変数 $y = 0, 1$ を導入すると、 \mathbf{x} をクラス C_y へ分類するということは、 \mathbf{x} の所属クラスを表す離散ベクトル $y(\mathbf{x})$ の値を決めることであると言い換えられる。

$$\mathbf{x} \longrightarrow y(\mathbf{x}) \in \{0, 1\} \quad (3.1)$$

分類先のクラスが多数にわたる場合、多クラス分類とよび、分類先が C_1, C_2, \dots, C_K となり、ラベル $y(\mathbf{x})$ も 1 から K の整数値をとる。これは先ほどの例でいえば、電子メールを「仕事」「家族」「友人」「スパム」と複数のフォルダに分けることに対応する。

回帰

データに対するクラス y は離散的とは限らない。例えば、過去数日の気象データ \mathbf{x} から明日の気温を予測したとき、 y は温度の数値に相当し、連続的な実数値をもつ変数になる。このようにデータから、それに対応する実数値 (を並べたベクトル) \mathbf{y} を予測作業のことを回帰 (regression) という。つまり、回帰とは、与えられた \mathbf{x} を、対応する \mathbf{y} に変換するために関数 $\mathbf{y}(\mathbf{x})$ を決定する作業である。

$$\mathbf{x} \longrightarrow \mathbf{y}(\mathbf{x}) \in \mathbb{R} \quad (3.2)$$

回帰による応用的なタスクとして、言語を翻訳する機械翻訳や音声文字起こし音声認識、状況が通常と異なることを自動的に検知させる異常検知、データのサイズを圧縮させるデータ次元削減などがあげられる。

3.2 統計入門

機械学習とは、データ (経験) をもとにしてプログラムがいろいろなタスクをこなせるようにさせることであった。ではどのようにしたらプログラムはデータからタスクをこなすための知識を学びとれるか？データを科学的に分析する数理的手法といえば、統計学である。機械学習の手法もまた統計を基礎として構築される。そのような視点から特に統計的機械学習と呼ばれている。ここではまず、統計の基礎を確認し、どのようにして学習アルゴリズムや評価指標を設計できるかについて説明する。

3.2.1 標本と推定

まず、データ (集合)(data(set)) やサンプル、標本 (sample) という用語についてきちんと定めておく。これらはデータ点 (data point) の集まりからなる。手書き文字による画像認識の例でいえば、データは統計分析に用いるために用意した画像の集合のことで、1 枚 1 枚の画像がそれぞれデータ点とよばれる。ただしデータ点も略してデータとよばれる。さらにはサンプルをサンプルの要素であるデータ点の意味でも用いられる。これらは用語の乱用であり、全く正しくない用法であるが、しばしば用いられているのが現状であるため、本論文でも文脈に応じてこの使い方に準じることにする。

このようなデータ点 (サンプル) の集まりからなるデータを分析するのが統計である。ここでは推定について考える。まず統計解析に用いるデータは、母集団から抽出 (sampling) されたものとみなす。データの要素を 1 つ取り出すことは、正確には抜き取り (draw) という。そしてデータの分析から、データ自身ではなくその背後にある母集団についての知識を獲得することを目標とす

る。これは例えば疫学者が飲酒量と健康の関係を調べたいときに、地球上のすべての(=母集団)を調査する必要はないということと同じである。その代わりにランダムに選び取った(抽出した)少人数に対する調査結果(データ)を分析することで、人類すべてに通用する飲酒量と健康の相関関係を読み取ろうとするということである。

母集団の性質はデータ生成分布 (generative distribution) $P_{\text{data}}(x)$ により特徴づけられているものとする。つまり、不確実性を伴う現象を確率的にモデル化するということである。

3.2.2 教師あり学習と教師なし学習，強化学習

機械学習には大きく 3 つに分類することができる。それは教師あり学習と教師なし学習，そして強化学習である。そのうち，ここでは本研究に関係する教師あり学習と教師なし学習について説明する。

教師あり学習

教師なし学習

3.3 深層学習とは

3.3.1 ニューラルネットワーク

3.3.2 畳み込みニューラルネットワーク

3.4 ボルツマンマシンの基礎

第 4 章

イジング模型の相転移検出に関する研究

- 4.1 相の分類器による相転移検出
- 4.2 温度測定器による相転移検出
- 4.3 相転移検出がなぜ可能なのか？

第 5 章

機械学習と繰り込み群

5.1 繰り込み群と特徴抽出

第 6 章

まとめ

研究のまとめを書く．

謝辞

謝辞を書く.

参考文献

- [1] Tom Michael Mitchell. Machine Learning. Mc Graw-Hill, 1997.

付録 A

aaa

aaa

A.1 bbb

bbb