

平日夜の数学勉強会 2 回目

桑畑英資 (@mather314)

2017/10/24

前回のおさらい

- 機械学習は抽象的には「判定関数を漸近的に改善する処理」
- 「判定をする処理」と「改善しているか評価する関数」が重要
- データの表現はベクトル、評価のモデルに統計・確率
- ベクトルと和・スカラー倍の演算、写像（関数）の定義
- 行列の定義、和・積の演算
- ベクトルの線形写像と行列とベクトルの積の関係

前回のおさらい - ベクトル

K : 体 (field, 和・積・商の定義された集合)、 V : K のベクトル空間 (vector space)

ベクトル (vector)

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \boldsymbol{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in V$$

ベクトルの和

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \in V$$

スカラー倍

前回のおさらい - 写像

集合 A, B の写像 (mapping, function, 関数) $f : A \rightarrow B$

写像

任意の $a \in A$ に対して $f(a) \in B$ は一つに定まる

線形写像 (linear mapping) $f : V \rightarrow V$

線形写像

任意の $a, b \in K$, $v, w \in V$ について、

$$f(av + bw) = af(v) + bf(w)$$

が成り立つ

前回のおさらい - 行列

要素 K の $m \times n$ -行列の集合 M_{mn}

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{mn}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{pmatrix} \in M_{nk}$$

行列の積

$$AB = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1k} + \cdots + a_{1n}b_{nk} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1k} + \cdots + a_{mn}b_{nk} \end{pmatrix} \in M_{mk}$$

行列とベクトルの積

$$A\mathbf{v} = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix}$$

一般に $AB \neq BA$ となる

前回の反省

色々飛ばしすぎました…

今回は数学の基礎に少し戻って考えます

集合

集合とは「要素」の集まりで、各要素を等しいか等しくないかで判定可能なもの。

$$\{1, 2, 3, \dots\}, \emptyset(\text{空集合}), \mathbb{Z}(\text{整数})$$

重複するものは一つの要素とみなされるので、例えば

$$\{1, 2, 2, 4\} = \{1, 2, 4\}$$

内包的記法

要素を列挙する外延的記法に対して、要素の条件を書き集合を定義する記法を内包的記法という

$$\{2n \mid n \in \mathbb{Z}\}(\text{偶数全体})$$

数学で「集合」と言うとき、特定の条件を満たすものや集合の要素全体が法則性を持つものを考えることが多い

集合の等価性

いろいろな集合

- $A \cap B$: A と B の共通部分、どちらにも含まれる要素の集合
- $A \cup B$: A と B の和集合、どちらかに含まれる要素の集合
- \bar{A}, A^c : A の補集合、普遍集合内の A に含まれない要素の集合, complement
- $\wp(A), 2^A$: A のべき集合、 A の部分集合の全体 ($\{B \mid B \subset A\}$)

演算のある集合

- 集合の演算とは、集合の幾つかの要素の組み合わせで新しい要素を作る写像のこと
- 和 ($+$) や積 (\cdot) は二項演算 (2 つの要素の組み合わせによる演算)
- 整数にも有理数にも実数にも、加えてベクトルや行列にも演算があるが、それらの共通点や法則は何か？
- 数学ではこのように抽象化した概念を考える
- 逆に抽象概念を具体的な概念に落とし込むこともある (演繹とかアナロジー)

群

群の定義

ある集合 G が群 (group) であるとは、演算 (\cdot) が定義されており、以下の条件を満たしていることを指す (以下 $a \cdot b$ を ab と省略する)

- $\forall x, y \in G, xy \in G$ (積に関して閉じている)
- $\exists e \in G, \forall x \in G, ex = xe = x$ (単位元の存在)
- $\forall x, y, z \in G, (xy)z = x(yz)$ (結合法則)
- $\forall x \in G, \exists x^{-1} \in G, xx^{-1} = x^{-1}x = e$ (逆元の存在)

ここで「演算」は一般的な積とは異なります。例えば、整数は「和 (+)」に関して群といえますが、積に関しては群ではありません。

課題

整数が和の演算について単位元と逆元を定義し、群になっていることを確認せよ

環の定義

集合 R が環 (ring) であるとは、和と積の集合が定義されており以下の条件が成り立つ時にいう

- $\forall x, y \in R, xy \in R, x + y \in R$ (演算に関して閉じている)
- $\exists 0_R \in R, \forall x \in R, 0_R + x = x + 0_R = x$ (和の単位元の存在)
- $\forall x, y \in R, x + y = y + x$ (和の交換法則)
- $\forall x, y, z \in R, (x + y) + z = x + (y + z)$ (和の結合法則)
- $\forall x \in R, \exists -x \in R, x + (-x) = (-x) + x = 0_R$ (和の逆元の存在)
- $\exists 1_R \in R, \forall x \in R, 1_R x = x 1_R = x$ (積の単位元の存在)
- $\forall x, y, z \in R, (xy)z = x(yz)$ (積の結合法則)
- $\forall x, y, z \in R, (x + y)z = xz + yz, x(y + z) = xy + xz$ (分配法則)

整数 (\mathbb{Z}) はその和と積の演算に関して環である

体

体の定義

集合 K が体 (field^a) であるとは、集合 K が環であり、かつ積に関して以下の条件を満たすときにいう

- $\forall x \in K \setminus \{0\}, \exists x^{-1} \in K, xx^{-1} = x^{-1}x = 1_K$ (積の逆元の存在)

^aドイツ語の Körper から

体はいわゆる割り算が可能な集合の性質である
整数は体ではないが、有理数の集合 (\mathbb{Q}) は体となっている。

$$\mathbb{Q} \stackrel{\text{def}}{=} \left\{ \frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\} \right\}$$

実数 (\mathbb{R}) も体である

課題

複素数 (\mathbb{C}) は体か？

写像

写像

$f: A \rightarrow B$ が写像であるとは、以下の条件を満たすときに言う

- 任意の $a \in A$ に対して、 $f(a) \in B$ が必ず存在する
- $a \in A$ に対して $f(a) \in B$ はただひとつに定まる

$f(a)$ を a の像という。全ての $a \in A$ に対して像が存在することを強調して全域写像ということもある。一般には写像というときは全域写像を指す。

逆に、 A の部分集合のみにしか $f(a)$ が存在しない（定義されていない）写像を部分写像という。

逆写像

写像 $f: A \rightarrow B$ に対し、写像 $g: B \rightarrow A$ が存在して

$$\forall a \in A, g(f(a)) = a$$

統計の目的はいくつかの観測データやサンプルから実際にデータ全体はどうなっているのか、どのような構造がデータの中にあるのかを推定することにある

確率変数

$X : \mathfrak{P}(\Omega) \rightarrow \mathbb{R}$ が確率変数であるとは

- $\forall E \subset \Omega, X(E)$