第１章:線形代数

行列

- スカラーを表にしたもの
- ベクトルを並べたもの ⇒なんに使うのか
  - ベクトルの変換
  - 連立方程式を解く

連立方程式を行列を使って記述

$$\begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + 5x_2 = 5 \end{cases}$$

$$\downarrow$$

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

連立１次方程式の解き方

- 行基本変形で解ける
  - 一般に次の３つの操作になる
    - i行目をC倍する
    - s行目にt行目のc倍を加える
    - p行目とq行目を入れ替える

単位行列

- 数字の１の行列版
- 元の行列を変化させない

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cdots & \\ & & & 1 \end{pmatrix}$$

逆行列

- 行列の「逆数」のようなもの
  - 行基本変形を行った行列を掛け合わせると求めることができる

逆行列が存在しない条件

- 連立方程式の解がない
- 解が１組に定まらないタイプ
  - 形式的には

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

a:b = c:d の時逆行列を持たない
⇒　ad-bc = 0

行列式

- ある行列が２つの横ベクトルの組み合わせだと考えたとき

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \overrightarrow{v_1} \\ \overrightarrow{v_2} \end{pmatrix}$$

で作られる平行四辺形の面積が、逆行列の有無を判別する。 この面積を

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} \overrightarrow{v_1} \\ \overrightarrow{v_2} \end{vmatrix}$$

と表し、行列式と呼ぶ

行列式の展開

- ３つ以上のベクトルからできている行列式は展開できる

行列式の求め方

- 2行２列
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$
- 2行２列以上　⇒行列式を展開して2行２列にして計算する。

固有値と固有ベクトル

- 次の関係が成り立つときλを固有値、Aを固有ベクトルという
$$A\vec{x} = \lambda\vec{x}$$

固有値分解

- ある種の行列は分解することが可能
- 固有値を対角線上に並べた行列
$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \dots \end{pmatrix}$$
それに対応する固有ベクトルを並べた行列
$$V = \begin{pmatrix} \vec{v_1} & \vec{v_1} & \dots \end{pmatrix}$$
を用意すると　AV = VΛ となり
$$A = V\Lambda V^{-1}$$
となる
    - 計算が楽になる

特異値分解

- 正方行列以外の行列を固有値分解する方法が特異値分解
$$M = USV^T$$

演習問題（線形代数）

線形代数 演習問題

1.1. $\vec{a} = \begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix}$ $\vec{b} \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix}$

1.1.1. $\vec{a} + \vec{b} = \begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix}$

1.1.2 $\vec{a} - \vec{b} = \begin{pmatrix} -4 \\ 4 \\ -1 \end{pmatrix}$

1.1.3

$7\vec{a} = \begin{pmatrix} 7 \\ 42 \\ 21 \end{pmatrix}$

1.1.4

$8(\vec{a} + \vec{b}) = 8\begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix} = \begin{pmatrix} 48 \\ 64 \\ 54 \end{pmatrix}$

1.2.

$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 4 \\ 1 & 5 \end{pmatrix}$

1.2.1

$A + B = \begin{pmatrix} 3 & 5 \\ 6 & 8 \end{pmatrix}$

1.2.2

$A - 3B = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} - \begin{pmatrix} 3 & 12 \\ 3 & 15 \end{pmatrix} = \begin{pmatrix} -1 & -11 \\ 2 & -12 \end{pmatrix}$

2.1.

$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ $A = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 9 & 0 \\ 3 & 1 & 2 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 5 \end{pmatrix}$

2.1.1

$A\vec{v} = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 9 & 0 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 13 \\ 5 \\ 9 \end{pmatrix}$

2.1.2

$$B\vec{v} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 15 \end{pmatrix}$$

2.1.3

$$BA = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 3 & 4 \\ 5 & 9 & 0 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 6 & 10 \\ 25 & 23 & 10 \end{pmatrix}$$

2.1.4

$$B^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 5 \end{pmatrix}$$

2.2.

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

2.2.1

$$AB = \begin{pmatrix} 2 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 7 \\ 7 & 13 \end{pmatrix}$$

2.2.2

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 1 \\ 4 & -2 \end{pmatrix}$$

2.2.3

$$B^{-1} = -\frac{1}{8} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

2.2.4

$$BAB^{-1} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{8} & \frac{3}{8} \\ \frac{3}{8} & -\frac{1}{8} \end{pmatrix} = \begin{pmatrix} 14 & 4 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} -\frac{1}{8} & \frac{3}{8} \\ \frac{3}{8} & -\frac{1}{8} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{2}{8} & \frac{38}{8} \\ \frac{2}{8} & \frac{26}{8} \end{pmatrix}$$

$$= \frac{1}{4} \begin{pmatrix} -1 & 19 \\ 1 & 13 \end{pmatrix}$$

7.1

$$\vec{a} = \begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix} \quad \vec{b} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

$$\vec{a} + \vec{b} = \begin{pmatrix} 3 \\ 7 \\ 7 \end{pmatrix}.$$

7.2

$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ の固有ベクトル.

$$\begin{vmatrix} \lambda - 1 & -4 \\ -2 & \lambda - 3 \end{vmatrix} = 0 . \quad (\lambda - 1)(\lambda - 3) - 8 = 0$$

$$\lambda^2 - 4\lambda + 3 - 8 = 0$$

$$\lambda^2 - 4\lambda - 5 = 0$$

$$(\lambda - 5)(\lambda + 1) = 0$$

$\lambda = 5, -1.$ となる。

(ヒントから求める). $\lambda = 5$ のとき.

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \lambda = 5 \text{ が固有値.}$$

第２章:確率・統計

- 集合とは
  物の集まり
  数学的には
  $$S = \{a, b, c, d, e, f, g\}$$
  $$a \in S$$
  などと記載する
- 和集合
  $$A \cup B$$
- 共通部分
  $$A \cap B$$
- 絶対補
  $$U \setminus A = \overline{A}$$
- 相対補
  $$B \setminus A$$
- 確率
  - 頻度確率（客観確率）
    - 発生する頻度
    - 例：「１０本の内１本だけ当たりのくじを引いて当選する確率が１０%」
  - ベイズ確率（主観確率）
    - 信念の度合い
    - 例：「あなたは４０%の確率でインフルエンザです」という診断
- 確率の定義
  $$P(A) = \frac{n(A)}{n(U)} = \frac{\text{事象}A\text{が起こる数}}{\text{すべての事象の数}}$$
- 条件付確率
  - ある事象Bが与えられたもとで、Aとなる確率
  $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{n(A \cap B)}{n(B)}$$
- 独立な事象の同時確率
  - お互いの発生には因果関係のない事象Aと事象Bが同時に発生する確率
  $$P(A \cap B) = P(A)P(B|A) = P(A)P(B)$$
- ベイズ則
  $$P(A)P(B|A) = P(B)P(A|B)$$
- 記述統計と推測統計
  - 母集団の要約
    ⇒「記述統計：集団の性質を要抑止記述する」
  - 母集団から抽出 ⇒　標本　⇒　推測
    「推測統計：集団から一部を取り出し元の集団を推測する」
- 確率変数
  - 事象と結びつけられた数値
  - 事象そのものを指すと解釈する
- 確率分布
  - 事象の発生する確率の分布
  - 離散値であれば表にできる
- 期待値
  - 分布における確率変数の　平均の値　又は　「ありえそうな」値

$$E(f) = \sum_{k=1}^{n} P(X = x_k) f(X = x_k)$$

* すべての値を足して、確率をかける
連続値に対しては同じ考えで期待値を求められる（積分）

- 分散と共分散
  - データの散らばり具合を示す
  - 分散

$$Var(f) = E((f_{(X=x)} - E_{(f)})^2) = E(f^2_{(X=x)}) - (E_{(f)})^2$$

  * 2乗は、ずれが正でも負でも対応するためについている
  - 共分散
    - 2つのデータに対する傾向
      - 正の値：似た傾向
      - 負の値：逆の傾向
      - 0:関係がない

$$Cov(f, g) = E((f_{(X=x)} - E(f))(g_{(Y=y)} - E(g)))$$

- 分散と標準偏差
  - 分散は２乗しているので単位が違う
    ⇒逆演算（平方根をとる）
    単位が戻る（標準偏差という）
- ベルヌーイ分布
  - コイントスのイメージ
  - 表と裏が出る確率が違っていてもよい
    (例えばひしゃげているコインなど)
  - マルチヌーイ分布
    - さいころを転がすイメージ
    - 出る目が等しくなくてもいい (例えばおもりが入っているサイコロ)
- ２項分布
  - ベルヌーイ分布の多試行版

$$P(x|\lambda, n) = \frac{n!}{x!(n-x)!} \lambda^x (1-\lambda)^{n-x}$$

- ガウス分布
  - 真の分布がわからなくても、サンプルが多ければ正規分布になる
    ⇒　とりあえず分布がわからないときは、ガウス分布にすることが多い
- 推定
  - 母集団を特徴づけるパラメータを統計学的に推測すること
  - 点推定：
    平均値などを１つの値に推定すること。
- 推定量
  - パラメータを推定するために利用する数値の計算方法や計算式
    （中に関数が入っている）
- 推定値
  - 実際に施行をした結果の値（式ではなく値）
    - 慣例として$\hat{\Theta}$(ハットをつける）
- 標本分散

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

  - 和があまりとれない場合は値が小さくなってしまう
    ⇒不偏分散（元の値に近づけるため）
- 不偏分散

$$s^2 = \frac{n}{n-1} * \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

演習問題（確率統計）

確率・統計　演習問題

問3.1. 確率変数として適切なもの.

a. さいころを振ったときに出た目の数.

d. 8本のうち1本が当たりであるクジを、当たりが出るまで抽選し続けたときの回数.

(の2つ).

3.2.

事象	裏：0枚 表：4枚	裏：1枚 表：3枚	裏：2枚 表：2枚	裏：3枚 表：1枚	裏：4枚 表：0枚
確率変数	4	3	2	1	0
事象の発生回数	75	300	450	300	75
事象と対応する確率	1/16	4/16	6/16	4/16	1/16

5.1.

$$\frac{\text{洗たく物を干していてかつ雨の日数}}{\text{すべての洗たく物を干していた日数}} = \frac{12}{60} = \frac{1}{5}$$

$$\frac{\text{洗たく物を干していてかつ雨の日数}}{\text{すべての日数}} = \frac{12}{365}$$

5.2.1

$$P(B \mid \text{赤}) = \frac{1}{3}$$

5.2.2

$$P(\text{白} \mid A) = \frac{1}{2}$$

7.3. 確率変数f(x)の期待値、(離散値).

$$E(x) = \Sigma\, P(x) f(x) \quad \cdots (\beta)$$

7.4. 確率変数 $f(x)$ の分散

$$\mathrm{Var}(f) = E((f(x) - E(f(x))^2)$$

$$= E(f(x^2)) - E(f(x))^2 \quad \cdots ウ.$$

第３章：情報理論

- 人間の目では
  - １．１０この点が入った箱　：　１１個の点が入った箱
    ⇒　人間が見て一瞬で見分けるのは難しい
  - ２．1この点が入った箱　：　2個の点が入った箱
    ⇒　人間が見て一瞬で見分けられる
  - １．２．を比べるとき
    - 変化の数：１個
    - 変化の割合：
      1. $$\frac{1}{10}$$
      2. $$\frac{1}{1}$$
    - 情報の増え方を数式にしたい

- 自己情報量
  - 何か定義できる式が欲しいので作った
$$I(x) = -log(P(x)) = log(W(x))$$
- シャノンエントロピー
  - 情報の珍しさ
  - 自己情報量の期待値
$$H(x) = E(I(x)) = -E(log(P(x))) = -\sum(P(x)log(P(x)))$$
- カルバック・ライブラー　ダイバージェンス
  - コインを投げるとき、理想上のコインで試したが、
    実際に使われたのはいかさまのコインだった場合の違いを知りたいなど
  - 同じ事象・確率変数における異なる確率分布Ｐ，Ｑの違いを表す
- 交差エントロピー
  - Qについての自己情報量をPの分布で平均している
  - カルバック・ライブラー　ダイバージェンスの一部を取り出したもの

演習問題（情報理論）

情報理論 演習問題

4.1.

$$I = \log_2\left(\frac{1}{P(X)}\right) = -\log_2(P(X))$$

4.1.1.

$$I = -\log_2\left(\frac{1}{2}\right) = 1 \qquad \underline{1\,\mathrm{bit}} //$$

4.1.2.

$$I = -\log_2\left(\frac{1}{2}\times\frac{1}{2}\right) = 2 \qquad \underline{2\,\mathrm{bit}} //$$

4.1.3.

$$I = -\log_2\left(n\cdot\left(\frac{1}{2}\right)^n\right) = -\log_2(n) - \log_2\left(\frac{1}{2}\right)^n$$

$$= -\log_2(n) + n$$

$$\underline{-\log_2(n)+n} \ \underline{\mathrm{bit}} //$$

7.5. 離散的な事象の確率分布 $P(x)$ の
シャノンエントロピー

$$H(x) = -\sum \left(P(x)\log\left(P(x)\right)\right) \quad \text{---} \ \underline{(ア)} //$$