レポート_応用数学

第一章：線形代数

要点

行列とは、連立方程式をよりシンプルに表現したもの。行列どうしの加減乗除が可能だが、次元数を合わせる必要がある。例えば、行列の加算・減算はすべての行列の次元が $m \times n$ である必要があり、乗算・除算は $m \times n$ と $n \times o$ の次元である必要がある。
行列は、スカラー値の定数倍をかけることが可能。ただし、行列どうしの交換法則は成立しない。つまり、 $AB = BA$ は必ずしも成り立たない。
$A\vec{x} = B$ の $\vec{x}$ を求めるには、 A の **逆行列** （ A^{-1} ）を左からかける。

$$\begin{aligned} A\vec{x} =& \quad B \\ A^{-1}A\vec{x} =& \quad A^{-1}B \\ \vec{x} =& \quad A^{-1}B \end{aligned}$$

特に、対角成分がすべて1で他が0の行列を **単位行列** と呼び、 I であらわされる。 $AA^{-1} = I$ が成り立つ。

逆行列は、 **掃き出し法** を使った導出があるが、特に2x2行列ならば以下の公式で求めることができる。

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

ここで、 $ad - bc$ は **行列式** と呼び、以下のように表される。

$$det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - dc$$

固有値と固有ベクトル

$A\vec{x} = \lambda\vec{x}$ が成立するとき、 λ を固有値、 $\vec{x}$ を固有ベクトルと呼ぶ。

固有値分解

固有値は $(A - \lambda I)\vec{x} = \vec{0}$ と変形したとき、行列式 $|A - \lambda I| = 0$ から求められる。算出された固有値から固有ベクトル $\vec{x}$ を求める。pythonでは `numpy.linalg.eig()` 関数を使って求めることができる。

特異値分解

固有値分解では、正方行列を扱ったが正方行列以外の行列でも、固有値分解のようなことを行いたい。その手続きを特異値分解と呼ぶ。 $M = USV^{-1}$ で表されるとき、 S の対角成分が特異値となる。pyhtonでは `numpy.linalg.svd()` 関数で求めることができる。

演習問題

1.1

次のベクトル

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

に関して，以下の計算をせよ。

問 1.1.1 $\vec{a}+\vec{b}$

$$\begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix} + \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix}$$

問 1.1.2 $\vec{a}-\vec{b}$

$$\begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} -4 \\ 4 \\ -1 \end{pmatrix}$$

問 1.1.3 $7\vec{a}$

$$7 \cdot \begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 42 \\ 21 \end{pmatrix}$$

問 1.1.4 $8(\vec{a}+\vec{b})$

$$8 \cdot \begin{pmatrix} 6 \\ 8 \\ 7 \end{pmatrix} = \begin{pmatrix} 48 \\ 64 \\ 56 \end{pmatrix}$$

1.2

次の行列 $A=\begin{pmatrix}2&1\\5&3\end{pmatrix}$ $B=\begin{pmatrix}1&4\\1&5\end{pmatrix}$ に関して、以下の計算をせよ.

問1.2.1 $A+B$

$$\begin{pmatrix}2&1\\5&3\end{pmatrix}+\begin{pmatrix}1&4\\1&5\end{pmatrix}=\begin{pmatrix}3&5\\6&8\end{pmatrix}$$

問1.2.2 $A-3B$

$$\begin{pmatrix}2&1\\5&3\end{pmatrix}-3\begin{pmatrix}1&4\\1&5\end{pmatrix}=\begin{pmatrix}2&1\\5&3\end{pmatrix}-\begin{pmatrix}3&12\\3&15\end{pmatrix}=\begin{pmatrix}-1&-11\\2&-12\end{pmatrix}$$

2.1 次のベクトルと行列

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

に関して、以下の計算をせよ.

問2.1.1 $A\vec{v}$

$$\begin{pmatrix}1&3&4\\5&9&0\\3&1&2\end{pmatrix}\begin{pmatrix}1\\0\\3\end{pmatrix}=\begin{pmatrix}1+12\\5\\3+6\end{pmatrix}=\begin{pmatrix}13\\5\\9\end{pmatrix}$$

問2.1.2 $B\vec{v}$

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

問2.1.3 BA

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

問2.1.4 B^T

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

2.2 次の行列 $A=\begin{pmatrix}2&1\\4&1\end{pmatrix}$, $B=\begin{pmatrix}1&3\\3&1\end{pmatrix}$ に関して、以下の計算をせよ.

問2.2.1 AB

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

問2.2.2 A^{-1}

$$\begin{pmatrix}2&1\\4&1\end{pmatrix}^{-1}=\frac{1}{2\cdot1-1\cdot4}\begin{pmatrix}1&-1\\-4&2\end{pmatrix}=-\frac{1}{2}\begin{pmatrix}1&-1\\-4&2\end{pmatrix}$$

問2.2.3 B^{-1}

$$\begin{pmatrix}1&3\\3&1\end{pmatrix}^{-1}=\frac{1}{1\cdot1-3\cdot3}\begin{pmatrix}1&-3\\-3&1\end{pmatrix}=-\frac{1}{8}\begin{pmatrix}1&-3\\-3&1\end{pmatrix}$$

第二章：確率・統計

要点

集合には、以下のようなものがある。

- 和集合（ $A \cup B$: A 又は B ）
- 積集合（ $A \cap B$: A かつ B ）
- 絶対補（ $U \backslash A = \bar{A}$: A 以外のすべて）
- 相対補（ $B \backslash A$: B から $A \cap B$ を除いた領域 ）

確率には、 **頻度確率** , **ベイズ確率** がある。頻度確率は、事象が発生する頻度のことで、一般に言われる確率はこちらを指す。ベイズ確率は、ある事象が発生したときにその原因となる確率のこと。

事象付き確率 とは、ある事象が発生していることを前提としたときに、別の事象が発生する確率のこと。以下の式で表される。

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

条件付き確率は、言葉だけだと概念が分かりにくいため、以下のリンク先にあるような練習問題を解くことで理解が深まる。
https://bellcurve.jp/statistics/course/6438.html
https://www.geisya.or.jp/~mwm48961/kou2/mobile/joken_p_m.html

期待値 は、ある事象 X の発生する確率を $P(X)$ としたとき、以下の式で表される。

- 確率変数 $f(X)$ が **離散型** のとき
  - $E(f) = \Sigma_{k=1}^{n} P(X = x_k) f(X = x_k)$
- 確率変数 $f(X)$ が **連続型** のとき
  - $E(f) = \int P(X = x) f(X = x) dx$

分散 $Var(f)$ は、データの散らばりを表し、以下の式で表される。
$E(\{f(X) - E(f(X))\}^2)$
また、分散の平方根をとった値を **標準偏差** と呼ぶ。

共分散 $Cov(f, g)$ は、2つのデータ系列の関係の強さを表す。正に大きい値だと、正の相関が強く、負に大きい値だと負の相関が強い。0に近いと、2データ間の相関は乏しいと言える。

様々な確率分布として代表的なものに **ベルヌーイ分布** がある。これは、結果が2通りしか存在しない試行を1回行ったときの確率分布。
ベルヌーイ試行を2回以上行ったときの確率分布を **二項分布** と呼ぶ。
また、 **正規分布** という釣鐘型の分布があり、これは平均の部分が最も確率が高く、平均から離れていくにつれて確率が低くなる。

母集団から標本を抽出したとき、標本から母集団の特徴を予測することを **推定** と呼ぶ。
標本平均は、標本数が多いほど母平均に近づくという性質がある。これを **一致性** と呼ぶ。また、標本の期待値は母数と一致する。これを **不偏性** と呼ぶ。

標本分散は、一致性を満たすが不偏性は満たさない。一般に、データ数が少ないほどデータのばらつきは小さくなることが知られている。そのため、標本分散は母分散よりも小さくなってしまう。

演習問題

問3.1
次の選択肢のうち，確率変数として適当なものはどれか。すべて選べ。
a. さいころを振ったときに出た目の数。
b. ３枚のコインを同時に投げて裏表どちらが出るか試行したときのコインの枚数。
c. 赤，青，黄の3色のボールが壷の中に入っている。ここから１個を取り出したときの色。
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。

解答
a, d
bは、コインの枚数は不変のため不適切。
cは、色の名前は数値ではないため不適切。

問3.2
コインを同時に投げる試行を1200回行ったとして，下の表を作った。空欄を埋めよ。

事象	裏0,表4	裏1,表3	裏2,表2	裏3,表1	裏4,表0
確率変数	4	3	2	1	0
事象の発生回数	75	300	450	300	75
事象の発生確率	1/16	4/16	6/16	4/16	1/16

裏1,表3の確率：${}_4C_3(1/2)^3(1/2)^1 = 4/16$
裏2,表2の確率：${}_4C_2(1/2)^2(1/2)^2 = 6/16$
裏3,表1の確率：${}_4C_1(1/2)^1(1/2)^3 = 4/16$
裏4,表0の確率：${}_4C_0(1/2)^0(1/2)^4 = 1/16$

問5.1
1年のうち洗濯物を干していた日数を60日，洗濯物を干していてかつ雨が降ってきた日数を12日として，ある年における，洗濯物を干していた日という条件下で，雨が降ってきた日の発生する確率と，洗濯物を干していてかつ雨が降ってきた日の発生する確率を求めよ。

解答
雨の降ってきた日の確率を A ,洗濯物を干した日の確率を B とする。
$P(A) = 60/365$, $P(A \cap B) = 12/365$

洗濯物を干していた日という条件下で，雨が降ってきた日の発生する確率は $P(A|B)$ なので、
$P(A|B) = P(A \cap B)/P(B) = \frac{12/365}{60/365} = \frac{12}{60} = \frac{1}{5}$

洗濯物を干していてかつ雨が降ってきた日の発生する確率は $P(A \cap B)$ なので
$P(A \cap B) = \frac{12}{365}$

5.2 袋の中に赤い玉3個と白い玉2個が入っている。赤い玉はA，B，Cの文字が，白い玉にはA，Bの文字が，それぞれ1個に対して1文字ずつ記されている。以下の問いに答えよ。

問5.2.1
出てきた玉が赤色であったとき，それに記されている文字がBである確率

解答
赤かつBの確率は $1/5$
赤の確率は $3/5$
よって、 $\frac{1/5}{3/5} = \frac{1}{3}$

問5.2.2
出てきた玉に記されている文字がAであったとき，その玉の色が白色である確率。

解答
文字がAの確率は $2/5$
文字がAかつ白の確率は $1/5$
よって、 $\frac{1/5}{2/5} = \frac{1}{2}$

第三章：情報理論

要点

確率 $p(X)$ について、$-log_2(p(X))$ で表す値を **自己情報量** と呼び、 I の記号を使い単位をbitと定義する。
例えば、1から2への変化と9から10への変化は、同じく1増えているが、1から2のほうが大きく増えているように見え、9から10は1から2に比べるとそこまで増えていないように見える。そういう所からも、対数をとるというのは納得できると言える。

自己情報量の期待値を **シャノンエントロピー** と呼び、以下の式で表される。
$H(X) = -\Sigma(P(x)log(P(x)))$
シャノンエントロピーが一番大きくなるのは、情報に偏りがないとき。逆に、最も偏りが大きい、つまり、一方の事象ばかり発生し他方の事象が全く発生しないときに、シャノンエントロピーは一番小さくなる。

同じ事象・確率変数において、確率分布Qに対し実際に計測した確率分布PがQとどれくらい異なっているかを表したものを **カルバック・ライブラー(KL)　ダイバージェンス** と呼ぶ。

また、KLダイバージェンスの一部分を取り出したとき、Qについての自己情報量をPの分布で平均化しているものを **交差エントロピー** と呼ぶ。こちらもKLダイバージェンスのように、PとQの確率分布がどれくらい離れているかを表したものとなる。

演習問題

問4.1.1 1枚のコインを1回投げて表が出た事象の情報量は何bitか？
情報量 $I = -log_2(1/2) = 1$ (bit)

問4.1.2 2枚のコインを1回投げてすべて表が出た事象の情報量は何bitか？
確率 $p(X) = 1/2 * 1/2 = 1/4$
情報量 $I = -log_2(1/4) = -log_2 2^{-2} = 2$ (bit)

問4.1.3 n枚のコインを1回投げて1枚の表が出た事象の情報量は何bitか？
確率 $p(X) = {}_nC_1(1/2)^n = n(1/2)^n$
情報量 $I = -log_2(n(1/2)^n) = -log_2(n) - log_2 2^{-n} = -log_2(n) + n$ (bit)

問6.1

$X = AB$ が成り立つとき

$log(X) = log(AB) = log(A) + log(B)$

問6.2

$X = \frac{A}{B}$ が成り立つとき

$log(X) = log(\frac{A}{B}) = log(A) - log(B)$

問6.3

$X = x_1 x_2 x_3 x_4$ が成り立つとき

$log(X) = log(x_1 x_2 x_3 x_4) = log(x_1) + log(x_2) + log(x_3) + log(x_4) = \Sigma_{k=1}^{4} log(x_k)$

確認テスト

問7.1 次のベクトル和を求めよ.

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

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

問7.2 $\begin{pmatrix}1&4\\2&3\end{pmatrix}$ の固有ベクトルは $\begin{pmatrix}1\\1\end{pmatrix}$ の定数倍だった。$\begin{pmatrix}1\\1\end{pmatrix}$ に対する固有値を求めよ。

$\begin{pmatrix}1&4\\2&3\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\lambda\begin{pmatrix}1\\1\end{pmatrix}$ の λ が固有値となる.

$\begin{pmatrix}5\\5\end{pmatrix}=\lambda\begin{pmatrix}1\\1\end{pmatrix}$ より, $\underline{\lambda=5}$

問7.3 離散的な確率分布の下では,ある確率変数 $f(x)$ の期待値は $\underline{\sum P(x)f(x)}$ で表される.

問7.4 ある確率変数 $f(x)$ の分散 $Var(f)$ は

$$\begin{aligned}Var(f)&=E\left((f(x)-E(f(x)))^2\right)\\&=E\left(f(x)^2-2f(x)E(f(x))+E(f(x))^2\right)\\&=E(f(x)^2)-E(2f(x)E(f(x)))+E(E(f(x))^2)\\&=E(f(x)^2)-2E(f(x))E(f(x))+E(f(x))^2\\&=E(f(x)^2)-2E(f(x))^2+E(f(x))^2\\&=E(f(x)^2)-E(f(x))^2\end{aligned}$$

(a が定数のとき, $E(ax)=aE(x)$
また $E(f(x))$ は定数)

問7.5 シャノンエントロピーとは,平均情報量とも呼ばれ、以下の式で表される.

$$-\sum P(x)\log(P(x))$$