応用数学

第一章：線形代数

要点

- ベクトルはスカラーのセット、行列はスカラーの表でありかつベクトルのベクトル、ベクトルの変換に使う
- ベクトルの計算は連立方程式の研究から生まれた、行列を用いた連立方程式の解法は手順自体も行列の掛け算として表現することができる
- 逆行列は行列にかけると単位行列になる、逆行列を求めるためには掃き出し法
- 逆行列の有無を判別するのが行列式
- $Ax=\lambda x$ を満たすようなベクトル x を固有ベクトル、λを固有値という
- 固有値を対角線上に並べた行列Λと固有ベクトルを並べた行列Vに対して$AV=V\Lambda$となるため、$A=V\Lambda V$-1と変形できることを固有値分解という
- 正方行列以外に対して類似した操作は特異値分解という
- 特異値分解はデータ量削減に利用可能

演習問題や参考記事、修了課題など関連記事

応用数学演習問題（Practice_Mathematics Q）

第１章　ベクトルと行列の演算Ⅰ

1.1

次のベクトル

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

に関して，以下の計算をせよ。

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

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

問 1.1.3　$7\vec{a}=\begin{pmatrix}7\\42\\21\end{pmatrix}$

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

問 1.2.2　$A-3B=\begin{pmatrix}2&1\\5&3\end{pmatrix}-\begin{pmatrix}3&12\\3&15\end{pmatrix}=\begin{pmatrix}-1&-11\\2&-12\end{pmatrix}$

第２章　ベクトルと行列の演算Ⅱ

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　$A\vec{v}=\begin{pmatrix}13\\5\\9\end{pmatrix}$

問 2.1.2　$B\vec{v}=\begin{pmatrix}10\\15\end{pmatrix}$

問 2.1.3　$BA=\begin{pmatrix}10&6&10\\25&23&10\end{pmatrix}$

問 2.1.4　$B^{\top}=\begin{pmatrix}1&0\\0&2\\3&5\end{pmatrix}$

第二章：統計・確率

要点

- 頻度確率は客観確率ともいい発生する頻度であるのに対してベイズ確率は主観確率ともいい信念の度合い
- 条件付き確率はある事象X=xが与えられたときにY=yとなる確率
- 独立な事象の同時確率は単純な積
- ベイズ則でP(X=x|Y=y)P(Y=y)=P(Y=y|X=x)P(X=x)が成立、これを利用して条件付き確率を求めることができる
- 分散は「確率変数と期待値との差の2乗」の期待値、で解釈としては「データの散らばり具合」
- 共分散は２つの確率変数について「確率変数と期待値の差同士の積」の期待値、正を取れば似た傾向
- 分散の平方根は標準偏差
- ベルヌーイ分布は「コイントス」のイメージで２カテゴリー、マルチヌーイは３つ以上のカテゴリカルに拡張されている
- 二項分布はベルヌーイ分布の多試行版

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 $AB=\begin{pmatrix}5&7\\7&13\end{pmatrix}$

問 2.2.2 $A^{-1}=\frac{1}{-2}\begin{pmatrix}1&-1\\-4&2\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}&\frac{1}{2}\\2&-1\end{pmatrix}$

問 2.2.3 $B^{-1}=\frac{1}{-8}\begin{pmatrix}1&-3\\-3&1\end{pmatrix}=\begin{pmatrix}-\frac{1}{8}&\frac{3}{8}\\\frac{3}{8}&-\frac{1}{8}\end{pmatrix}$

問 2.2.4 $BAB^{-1}=\begin{pmatrix}1&3\\3&1\end{pmatrix}\begin{pmatrix}2&1\\4&1\end{pmatrix}\begin{pmatrix}-\frac{1}{8}&\frac{3}{8}\\\frac{3}{8}&-\frac{1}{8}\end{pmatrix}$

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

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

第 3 章　確率変数と確率分布

問 3.1

「異なる絵柄の描かれた 5 枚のカードから 1 枚を引く」というような，同じ条件下で繰り返し行うことのできる実験や観測などのことを「試行」といい，それによって起こる「★印の描かれたカードを引いた」というような，結果を「事象」という。

試行の結果として起こる事象に整数や実数の数値が結びつけられているときに，その数値を「確率変数」という。

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。

問 3.2

試行の結果に生じる様々な事象は，すべてが同じ割合で生じるとは限らない。たとえば，「4 枚のコインを同時に投げて裏表どちらが出るか」試行した場合「すべて表であったり，すべて裏である場合は少なく，2 枚は裏で 2 枚が表といった裏表が混合している場合の方がより多く発生するだろう」ということは計算しなくとも直観的に理解できるかと思う。事象は様々な割合で（すなわち様々な確率で）発生するのだから，事象と結びつけられた数値（たとえば表という事象に 1 という数値を，裏という事象に 0 という数値を対応させる etc）である確率変数もまた様々な確率で発生することになる。この確率変数と対応する確率との関係を表した関数を「確率分布」という。確率変数が離散値（とびとびの値）ならば対応

する確率もまた離散値となる。つまり確率変数と対応する確率との関係（すなわち確率分布）を表にまとめることができるということである。
上述の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	裏が 0 枚，表が 4 枚	裏が 1 枚，表が 3 枚	裏が 2 枚，表が 2 枚	裏が 3 枚，表が 1 枚	裏が 4 枚，表が 0 枚
確率変数（裏を 0，表を 1 と対応させ和をとった）	4	3	2	1	0
事象が発生した回数	75	300	450	300	75
事象と対応する確率	1/16	1/4	3/8	1/4	1/16

第 4 章　情報量

4.1

滅多に起こらない事象が発生したという情報（例：伝説の生物ツチノコを発見したという情報）の重要度と，しばしば起こる事象が発生したという情報（例：残念なことに風呂場で G から始まる名前を持つ昆虫を発見してしまったという情報）の重要度とでは，どちらの情報により重要さがあるだろうか。これは直観的に前者であるとわかると思う。情報理論の世界の考え方もこの直観には従う。たとえば，10 本に 1 本当たりが含まれているクジに当選したという情報より，1000 本に 1 本当たりが含まれるクジに当選したという情報の方が，（つまり発生する確率の小さい事象の発生する情報の方が）より情報の量が多いと考えるのだ。具体的には 2 回に 1 回は起こるような事象（確率 1/2 の事象）が発生したという情報の大きさを 1bit の「(自己) 情報量」として定義している。この情報量は足し合わせ可能な量だと考え，2 回に 1 回は起こるような事象 A と，また別の 2 回に 1 回は起こるような事象 B が同時に発生するという事象（確率 1/4 の事象）の情報量を 2bit，2 回に 1 回は起こるような事象 A と，また別の 2 回に 1 回は起こるような事象 B と，さらに別の 2 回に 1 回は起こるような事象 C が同時に発生するという事象（確率 1/8 の事象）の情報量を 3bit のように考えていく。この関係を数式で表すと以下のようになる。

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

※I は情報量，p(X)は事象 X の発生する確率。

この定義にしたがって，次の問いに答えよ。

問 4.1.1　1 枚のコインを 1 回投げて表が出たという事象の情報量は何 bit か。

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という事象の情報量は何 bit か。

2 bit

問 4.1.3　n 枚のコインを 1 回投げて 1 枚の表が出たという事象の情報量は何 bit か。

~~n bit~~　$I = -\log_2\left({}_nC_1\left(\frac{1}{2}\right)^n\right) = -\log_2 n \cdot 2^{-n}$

$= n - \log_2 n$ (bit)

第 5 章　条件付き確率

問 5.1

条件付き確率とは，事象 A が起こった条件の下で事象 B が起こる確率のことであり，これを$p(B|A)$と表す（B が先に書かれるのは，現在の数学の記述法の発祥の地であるヨーロッパで使われている言語における語順が反映しているため。例：the probability of B under the condition A)。たとえば，ある年における，洗濯物を干していたときに，雨が降ってきた確率（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）を考える。前述の記述法に従うならば，

$$\mathrm{p}\left(雨が降ってきた日|洗濯物を干していた日\right)$$

のように記述できる。これは洗濯物を干していてかつ雨が降ってきた確率（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）と同様ではない。事象 A と事象 B が同時に起こる確率を同時確率といい，$p(A,B)$あるいは$p(B,A)$で表され（どちらも同じ値となる），条件付き確率とは区別されなければならない。なぜなら，条件付き確率の場合

$$\mathrm{p}\left(雨が降ってきた日|洗濯物を干していた日\right) = \frac{洗濯物を干していてかつ雨が降ってきた日数}{すべての洗濯物を干していた日数}$$

を表していることになるが，同時確率の場合

$$\mathrm{p}\left(雨が降ってきた日,洗濯物を干していた日\right) = \frac{洗濯物を干していてかつ雨が降ってきた日数}{すべての日数（つまり365\ 日）}$$

を表していることになるからである。

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。

(前者) $= \frac{12}{60} = \frac{1}{5}$　(後者) $= \frac{12}{365}$

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。

$$P(B|赤) = \frac{P(B \cap 赤)}{P(赤)} = \frac{(3/5)(1/3)}{3/5} = \frac{1}{3}$$

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。

$$P(白|A) = \frac{P(白 \cap A)}{P(A)} = \frac{2/5 \cdot 1/2}{3/5 \cdot 1/3 + 2/5 \cdot 1/2} = \frac{1/5}{1/5 + 1/5} = \frac{1}{2}$$

第 6 章　対数と乗算除算の関係

問 6.1

実数 X, A, B に対して $X = AB$ という関係が成り立っている。次の式の空欄に入れることのできる適切な選択肢を選べ。

$$\log(X) = \boxed{}$$

ア　$\log\left(\frac{A}{B}\right)$　イ　$\log(A) - \log(B)$　ウ　$\log(A) + \log(B)$　エ　$A\log(B)$

問 6.2

実数 X, A, B に対して　$X = \frac{A}{B}$　という関係が成り立っている。次の式の空欄に入れることのできる適切な選択肢を選べ。

$$\log(X) = \boxed{}$$

ア　$\log(AB)$　イ　$\log(A) - \log(B)$　ウ　$\log(A) + \log(B)$　エ　$A\log(B)$

問 6.3

実数 X, x_1, x_2, x_3, x_4 に対して　$X = x_1x_2x_3x_4$　という関係が成り立っている。次の式の空欄に入れることのできる適切な選択肢を選べ。

$$\log(X) = \boxed{}$$

ア　$\log(x_1)\log(x_2)\log(x_3)\log(x_4)$　イ　$\log(x_1x_2) - \log(x_3x_4)$

ウ　$\sum_{k=1}^{4}\log(x_k)$　エ　$\prod_{k=1}^{4}\log(x_k)$

▾ 第三章：情報理論

要点

- 自己情報量Ｉ（ｘ）は$I(x)=-\log P(x)=\log W(x)$、２を底とするとき単位はbit
- シャノンエントロピーH(x)は自己情報量の期待値で$H(x)=E(I(x))=E(-\log P(x))=-\Sigma(P(x)\log P(x))$
- カルバック・ライブラーダイバージェンス$D_{KL}=E(x\sim P)[\log(P(x)-\log Q(x)]$は同じ事象・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$H(P, Q)=H(P)+D_{KL}(P||Q)$はKLダイバージェンスの一部分を取り出したもので、Qについての自己情報量をPの分布で平均している

演習問題や参考記事、修了課題など関連記事

第７章　確認テスト

問 7.1　次のベクトルの和を求めよ。

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

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

問 7.2　　行列$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$の固有ベクトルを求めたところ，$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$の定数倍であることがわかった。固有ベクトル$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$に対応する固有値を求めよ。

ヒント：$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ 1 \end{pmatrix}$が成り立つとき，$\lambda$は，固有ベクトル$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$に対応する固有値であるといえる。

固有値をλとすると $\begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^2 - 4\lambda - 5 = 0$

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

これらのうち，$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ を満たすのは

$\lambda = 5$ (答).

問 7.3　　離散的な確率分布$P(x)$の下では，ある確率変数$f(x)$の期待値は

$$\mathrm{E}(f) = \boxed{}$$

と表される。空欄にふさわしい選択肢はどれか。

ア　$\sum P(x)f(x)$　　　~~イ　$\int P(x)f(x)$~~　　　ウ　$\sum \frac{f(x)}{P(x)}$　　　エ　$\sum \frac{P(x)}{f(x)}$

問 7.4　ある確率変数$f(x)$の分散は

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

$$= \boxed{}$$

と表される。空欄にふさわしい選択肢はどれか。

ア　$\mathrm{E}(f(x)) - \mathrm{E}(f(x))$　　　イ　$\mathrm{E}(f(x)^2)\mathrm{E}(f(x))^2$　　　ウ　$\mathrm{E}(f(x)^2) - \mathrm{E}(f(x))^2$

問 7.5　　シャノンエントロピーは自己情報量の平均である。ある離散的な事象の確率分布を$P(x)$としたとき，シャノンエントロピーとしてふさわしいものは，次のうちどれか。

ア　$\sum \frac{1}{P(x)} \log(P(x))$　　　イ　$-\sum P(x) \log(P(x))$　　　ウ　$\sum P(x) + \log(P(x))$