応用数学

第一章　線形代数

（1）基本事項

・スカラー…いわゆる普通の数。四則演算が可能。

・ベクトル…「大きさ」と「向き」を持つ。スカラーのセットで表示される。

・行列…スカラーを表にしたもの。ベクトルを並べたもの。

　　　　ベクトルの変換に使用する。

（2）連立方程式

連立方程式は、行列の演算で解くことができる。

講義資料より、下記のとおり対応させることができる。

＜連立方程式＞　　　　　　　　＜行列＞

（行基本変形）　　⇔　　（行列の変形）

$$\begin{cases} x_1 + 4x_2 = 7 \\ 2x_1 + 6x_2 = 10 \end{cases} \qquad \begin{pmatrix} 1 & 4 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 10 \end{pmatrix}$$

2行目を1/2倍する　$\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \times$

$$\begin{cases} x_1 + 4x_2 = 7 \\ x_1 + 3x_2 = 5 \end{cases} \qquad \begin{pmatrix} 1 & 4 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

1行目に2行目の-1倍を加える　$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \times$

$$\begin{cases} x_2 = 2 \\ x_1 + 3x_2 = 5 \end{cases} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

2行目に1行目の-3倍を加える　$\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \times$

$$\begin{cases} x_2 = 2 \\ x_1 = -1 \end{cases} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

1行目と2行目を入れ替える　$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \times$

$$\begin{cases} x_1 = -1 \\ x_2 = 2 \end{cases} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

・単位行列…かけてもかけられても相手が変化しない行列。

$$I = \begin{pmatrix} 1 & & \\ & 1 & \\ & & \ddots \end{pmatrix}$$

※空白部分は 0

・逆行列…逆数のような働きをする行列。

$$AA^{-1} = A^{-1}A = I$$ （A^{-1} が逆行列）

・掃き出し法…逆行列の求め方。
　元となる行列と逆行列をかけた結果が単位行列となることを利用する。

（３）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、この横ベクトルの始点を合わせたときに作られる平行四辺形の面積を表す。

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \vec{v}_1 \\ \vec{v}_2 \end{pmatrix}$ のとき、 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} \vec{v}_1 \\ \vec{v}_2 \end{vmatrix}$ と表す。

　ある 1 つの正方行列の行列式には、ある 1 つの数値が対応する。正方行列の大きさのようなもの。

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

　また、3 つ以上のベクトルからできる行列式は下記の通り展開できる。

$$\begin{vmatrix} \vec{v}_1 \\ \vec{v}_2 \\ \vec{v}_3 \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & c \\ d & e & f \\ 0 & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & c \\ 0 & e & f \\ g & h & i \end{vmatrix}$$
$$= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - d\begin{vmatrix} b & c \\ h & i \end{vmatrix} + g\begin{vmatrix} b & c \\ e & f \end{vmatrix}$$

（４）固有値、固有ベクトル
　行列 A とその特殊なベクトル$\vec{x}$の積が、ただのスカラーの数λとその特殊なベクトル$\vec{x}$との積と同じ値になるとき、この特殊なベクトル$\vec{x}$とその係数λを，行列 A に対する，固有ベクトル、固有値という。

$A\vec{x}=\lambda\vec{x}$

・固有値、固有ベクトルの求め方
　$A\vec{x}=\lambda\vec{x}$　から、左辺にまとめると　$(A-\lambda I)\vec{x}=\vec{0}$　となる。　$x\neq 0$ より、
　$|A-\lambda I| = 0$　となる値を求める。

・固有値分解
　ある実数を正方形にならべて作られた行列 A が､固有値λ_1、λ_2､…と固有ベクトル$\vec{v}_1, \vec{v}_2,$ …を持ったとする。この固有値を対角線上に並べた行列（それ以外の成分は 0）

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}$$

と，それに対応する固有ベクトルを並べた行列

$$V = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots \end{pmatrix}$$

を用意したとき、それらは

$AV=V\Lambda$

$A=V\Lambda V$-1

と関係付けられる。このように正方形の行列を上述の様な3つの行列の積に変換することを固有値分解という。この変換によって行列の累乗の計算が容易になる。

・特異値分解

正方行列以外でも、下記のような特殊な単位ベクトルがあるならば特異値分解できる。

$$M\vec{v} = \sigma\vec{u}$$
$$M^{\top}\vec{u} = \sigma\vec{v}$$

特異値分解を用いることで、成分の小さい行列（不必要な要素）を取り除いて、データ量を削減することができる。

第二章　確率・統計

（１）基本事項

・頻度確立（客観確立）…実際に発生した頻度。

・ベイズ確立（主観確立）…信念の度合い。どれくらいその事象が起こりそうかを表す。

・条件付確率…ある事象 X=x　が与えられたもとで、Y=y　となる確率。

$$P(Y=y|X=x)=\frac{P(Y=y,X=x)}{P(X=x)}$$

・独立な事象の同時確立…お互いの発生には因果関係のない事象 X=x　と事象 Y=y　が同時に発生する確率。

$$P(X=x,Y=y) \quad = \quad P(X=x)\,P(Y=y) \quad = \quad P(Y=y,X=x)$$

・ベイズ則…ある事象Yに関連する可能性のある条件Xについて、その事象が原因となった確率。

$$P(X=x\,|\,Y=y)\,P(Y=y) \quad = \quad P(Y=y\,|\,X=x)\,P(X=x)$$

（２）確率変数

・確率変数…事象と結びつけられた数値。

・確率分布…事象の発生する確率の分布。

・期待値…ある分布に対する確率変数の平均値、またはありえそうな値。

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

連続値の場合、　$\mathrm{E}(f) \quad = \int P(X=x)f(X=x)dx$

（$P(X=x)$は確率、$f\,(X=x)$は確率変数を表す）

・分散…データの散らばり具合。

$$\begin{aligned}\mathrm{Var}(f) \quad &= \mathrm{E}\Big(\big(f_{(X=x)}-\mathrm{E}_{(f)}\big)^2\Big)\\ &= \mathrm{E}\big(f^2_{(X=x)}\big)-\big(\mathrm{E}_{(f)}\big)^2\end{aligned}$$

・共分散…2 つのデータ系列の傾向の違い。

正の場合：似た傾向

負の場合：逆の傾向

0 に近い場合：関係性が薄い

・標準偏差…分散の平方根。分散では、2 乗しているため、元の単位と異なってしまっているので、元に戻している。

（3）確率分布

・ベルヌーイ分布…確率変数を 0 と 1 のみで表現できる場合。

$$P(x \mid \mu) = \mu^{x}(1-\mu)^{1-x}$$

・マルヌーイ分布…複数の離散値を確率変数として扱う場合。

・二項分布…ベルヌーイ分布の多試行版。

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

※n が全体、x が表のイメージ。

・ガウス分布…釣り鐘型の連続分布。

$$\mathcal{N}(x; \mu, \sigma^2) = \sqrt{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

第 3 章　情報理論

・自己情報量…増加の比率を表す。
対数の底が 2 のとき、単位はビット(bit)
対数の底がネイピアの e のとき、単位は(nat)

$$I(x) \quad = \quad -\log(P(x)) \quad = \quad \log(W(x))$$

・シャノンエントロピー（微分エントロピー）…自己情報量の期待値。

$$H(x) \quad = \quad E(I(x)) \quad = \quad -E(\log(P(x)) \quad = \quad -\Sigma(P(x)\log(P(x))$$

・カルバック・ライブラー・ダイバージェンス（KL ダイバージェンス）
…同じ事象・確率変数における異なる確率分布 P、Q の違い。
Q からみて、P の分布がどのくらい異なるかを表す。

$$D_{\mathrm{KL}}(P\|Q) = \mathbb{E}_{\mathrm{x}\sim P}\left[\log\frac{P(x)}{Q(x)}\right] = \mathbb{E}_{\mathrm{x}\sim P}\left[\log P(x) - \log Q(x)\right]$$

・交差エントロピー…KL ダイバージェンスの一部を取り出したもの。
Q についての自己情報量を P の分布で平均している。

$$H(P,Q) = H(P) + D_{\mathrm{KL}}(P\|Q)$$

$$H(P,Q) = -\mathbb{E}_{\mathrm{x}\sim P}\log Q(x)$$

【参考サイト】
・Wikipedia（ベイズの定理）(2021/7/18)
https://ja.wikipedia.org/wiki/%E3%83%99%E3%82%A4%E3%82%BA%E3%81%AE%E5%AE%9A%E7%90%86