応用数学

2021 年 11 月 19 日

1 線形代数

1.1 スカラー、ベクトル、行列

線形代数ではいくつかの数学的な対象を扱う。

- スカラー
  スカラーは単純な数字
- ベクトル
  ベクトルは数字の配列で次のように表現される:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

- 行列
  行列は数字の 2 次元の配列で次のように表現される:

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

  行列の操作は以下の通り:

  – 和

  $$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} \rightarrow C_{i,j} = A_{i,j} + B_{i,j}$$

  – 積

  $$\boldsymbol{C} = \boldsymbol{AB} \rightarrow C_{i,j} =_k A_{i,k} \times B_{k,j})_]$$

  – 要素積 (アダマール積)
  同じインデックスの要素同士を掛けわせせる。$\boldsymbol{A} \otimes \boldsymbol{B}$ と表記する:

  $$\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} \rightarrow C_{i,j} = A_{i,j} \times B_{i,j}$$

1.2 逆行列

1.2.1 行列基本変形

- ある行を定数倍する

- ある行に別の行の定数倍を加える
- ある行と別の行を入れ替える

行列 $\begin{pmatrix} 1 & 4 \\ 2 & 6 \end{pmatrix}$ の逆行列は行基本変形の組み合わせで求められる：

$\left(\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 2 & 6 & 0 & 1 \end{array}\right) \to$ 2 行目に 1/2 を掛ける $\to \left(\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ 1 & 3 & 0 & \frac{1}{2} \end{array}\right)$

$\to$ 1 行目に 2 行目 $\times -1$ を加える $\to \left(\begin{array}{cc|cc} 0 & 1 & 1 & -\frac{1}{2} \\ 1 & 3 & 0 & \frac{1}{2} \end{array}\right)$

$\to$ 2 行目に 1 行目 $\times -3$ を加える $\to \left(\begin{array}{cc|cc} 0 & 1 & 1 & -\frac{1}{2} \\ 1 & 0 & -3 & 2 \end{array}\right)$

$\to$ 1 行目と 2 行目を入れ替える $\to \left(\begin{array}{cc|cc} 1 & 0 & -3 & 2 \\ 0 & 1 & 1 & -\frac{1}{2} \end{array}\right)$

逆行列は $\begin{pmatrix} -3 & 2 \\ 1 & \frac{1}{2} \end{pmatrix}$ となる。この逆行列の求め方をガウスの掃き出し法という。

1.3　行列式

ある行列が 2 つのベクトルの組み合わせだと考えてみる

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

$\overrightarrow{v_1}$ と $\overrightarrow{v_2}$ からなる平行四辺形の面積 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} \overrightarrow{v_1} \\ \overrightarrow{v_2} \end{vmatrix}$ を**行列式**という。行列式が 0 のとき、逆行列は存在しない。

1.3.1　行列式の性質

- 同じベクトルを複数持つとき、行列式は 0
  $\begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{w} \\ \overrightarrow{w} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix} = 0$
- 1 つのベクトルが定数倍されると、行列式も定数倍
  $\begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \lambda\overrightarrow{v_l} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix} = \lambda \begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{v_l} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix}$
- 1 つのベクトルだけ違うとき、行列式は足し合わせ

$$\begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{v_l}+\overrightarrow{w} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix} = \begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{v_l} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix} + \begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{w} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix}$$

- 行を入れ替えると符号が変わる

$$\begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{v_s} \\ \overrightarrow{v_t} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix} = - \begin{vmatrix} \overrightarrow{v_1} \\ \vdots \\ \overrightarrow{v_t} \\ \overrightarrow{v_s} \\ \vdots \\ \overrightarrow{v_2} \end{vmatrix}$$

1.4 固有値と固有ベクトル

$A\overrightarrow{v} = \lambda\overrightarrow{v}$ となる λを固有値、$\overrightarrow{v}$を固有ベクトルという。固有値は $|A - \lambda I)| = 0$ の解。

1.5 固有値分解

ある行列を、固有値を対角成分とする行列と、固有ベクトルからなる行列で表現することを固有値分解という。

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ & 0 & & \end{pmatrix}, V = \begin{pmatrix} \overrightarrow{v_1} \cdots \end{pmatrix} \Rightarrow AV = V\Lambda \Rightarrow A = V\Lambda V^{-1}]$$

1.6 特異値分解

正方行列以外は固有値分解と似たようなことができる

$$M\overrightarrow{v} = \sigma\overrightarrow{v}$$

$$M^T\overrightarrow{u} = \sigma\overrightarrow{u}$$

このような特殊な単位ベクトルが存在すれば、次のように特異値分解できる

$$M = USV^{-1}$$

1.7 特異値の求め方

$MV = US \Rightarrow \mathrm{M} = \mathrm{USV}^{-1}$

$M^TV = VS^T \Rightarrow \mathrm{M}^T = VSU_{-1}$

これらの関係から以下が成り立つ

$$MM^T = USV^{-1}VS^Tu^{-1} = USS^TU^{-1}$$

この MM^T を固有値分解すればよい。

2　確率・統計

2.1　確率

頻度確率 (客観確率)　発生する頻度
ベイズ確率 (主観確率)　信念の度合い

2.2　条件付き確率

$X = x$ が与えられたとき、$Y = y$ となる確率

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

2.3　ベイズ則

$$P(X = x|Y = y)P(Y = y) = P(Y = y|X = x)P(X = x)$$

2.4　確率変数と確率分布

確率変数　事象と結び付けられた数値
確率分布　事象の発生する確率の分布

2.5　期待値

ある分布における確率変数の平均値、「あり得そうな」値

$$E(f) = \begin{cases} \sum_{k=1}^{n} P(X = x_k)f(X = x_k) & (\text{離散確率変数}) \\ \int_{-\infty}^{\infty} P(X = x)f(X = x)dx & (\text{連続確率変数}) \end{cases}$$

2.6　分散と共分散

分散　データの散らばり具合

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

共分散　2 つのデータ系列の傾向の違い

$$\begin{aligned} Cov(f, g) &= E((f(X = x) - E(f))(g(Y = y) - E(g))) \\ &= E(fg) - E(f)E(g) \end{aligned}$$

標準偏差$\sigma = \sqrt{Var(f)}$

2.7　確率分布

ベルヌーイ分布　コイントス

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

マルチヌーイ (カテゴリカル) 分布　サイコロ

$P(x_1, x_2, \cdots, x_k|\mu_1, \mu_2, \cdots, \mu_k) = \frac{n!}{x_1!x_2!\cdots x_k!}\mu_1^{x_1}\mu_2^{x_2}\cdots\mu_k^{x_k}$

二項分布　ベルヌーイ試行の多試行版

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

ガウス分布　釣鐘型の連続分布

$N(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

2.8　記述統計と推測統計

推測統計　集団から一部を取り出し、元の集団 (母集団という) の性質を推測する

記述統計　集団の性質を要約し記述する

2.9　推定

母集団を特徴づける母数 (パラメータ) を統計学的に推測すること。

- 推定量
  パラメータ推定に使う計算方法や計算式のこと。推定関数ともいう
- 推定値
  計算した値

2.10　標本平均

標本の平均値

一致性　サンプル数が大きいほど母集団の値に近づく

不偏性　サンプル数がいくらであっても、その期待値は母集団と同様。 $E(\hat{\theta}) = \theta$

2.11　標本分散

一致性は満たすが、不偏性は満たさない

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

2.12 不偏分散

サンプル数が少ないときに不偏性を満たすように標本分散を修正

$$s^2 = \frac{n}{n-1}\hat{\sigma}^2$$

3 情報理論

3.1 自己情報量

$$I(x) = -\log(P(x)) = \log(W(x)) \quad (\text{底が 2 のとき単位は } bit\text{、} e \text{ のとき } nat)$$

3.2 シャノンエントロピー

自己情報量の期待値

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

3.3 カルバック・ライブラーダイバージェンス

同じ事象・確率変数における、異なる確率分布 P、Q の違いを表す

$$\begin{aligned} D_{KL}(P||Q) &= E_{x\sim P}\left[\log\left(\frac{P(x)}{Q(x)}\right)\right] \\ &= P(x)\frac{P(x)}{Q(x)} \\ &= P(x)\log(P(x)) - P(x)\log(Q(x)) \end{aligned}$$

3.4 交差エントロピー

KL ダイバージェンスの一部分を取り出したもの

$$\begin{aligned} H(P,Q) &= H(P) + D_{KL}(P||Q) \\ &= -E_{x\sim P}\left[\log Q(x)\right] \end{aligned}$$