

高校数学の知識から、

人工知能・機械学習・データ解析へ
つなげる、

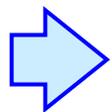
必要最低限の教科書

明治大学 理工学部 応用化学科
データ化学工学研究室 金子 弘昌

どうして人工知能について学ぶ必要があるのか？¹

✓人工知能・数学・統計・・・を学ぶ理由

- 世の中にある怪しい人工知能・統計関係の話にだまされなくなる
- 自分で人工知能をつくれる
 - 暗黙知を形式知化できる
 - = 誰かの頭の中に知識・知見・経験としてあるけど言葉・文字にして他の人に伝えられないこと (暗黙知) を、伝えられる形として表現 (形式知化) できる
- 自分でデータ解析ができるようになる



人工知能・機械学習・データ解析を武器にして、自分で、新しい研究を開拓できる！

どうやって人工知能について学ぶのか？

✓人工知能・数学・統計・・・を学ぶ方法

- この資料を順番に、
 - 自分の頭で考えながら、
 - 実際に手を動かしながら、
 - 分からないところはインターネットなどで調べながら
(高校数学で対応できるように作りましたが、念のため)、
- 理解してってください

内容

✓行列・ベクトルの表現

✓いろいろな行列

- 正方行列・単位行列・逆行列・線形変換

✓固有値・固有ベクトル

✓偏微分・全微分・Lagrangeの未定乗数法

✓確率

- 同時確率・条件付き確率・周辺化・
確率の加法定理・確率の乗法定理・ベイズの定理

連立方程式 (二元一次方程式)

✓二元一次方程式

- 二元：2つの変数がある (たとえば、 x と y)
- 一次：すべての変数は、1乗 (x^2 とか y^3 とかはない)

- 例：
$$2x + 3y = 7$$
$$x + 2y = 4$$

- 解き方：

- 2つの式を使って1つ文字を減らして、もう1つの値を求める
- その値と1つの式を使って、残りの文字の値を求める

- 答え： $x = 2, y = 1$

二元一次方程式を別の形で表現する

$$2x + 3y = 7$$

$$x + 2y = 4$$

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 7 \\ 4 \end{pmatrix} \quad \text{とおき、}$$

$$2x + 3y = 7$$

$$x + 2y = 4$$

を、

$$\mathbf{Ab} = \mathbf{c}$$

と書くことにする (後で詳しく
説明します)

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$$

左のように、縦に2つ以上、横に2つ以上
数字が並んだものを**行列**とよぶ

列

行 $\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$

… 2行2列の行列、 2×2 の行列

$$\begin{pmatrix} 4 & -2 & 1.2 & 6 \\ 4.5 & 0 & -4 & -1.2 \\ 7 & 9.5 & 8.1 & -3 \end{pmatrix}$$

… 3行4列の行列、 3×4 の行列

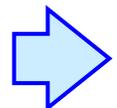
行列の表し方

行列をローマ字であらわすときは、

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 4 & -2 & 1.2 & 6 \\ 4.5 & 0 & -4 & -1.2 \\ 7 & 9.5 & 8.1 & -3 \end{pmatrix}$$

の \mathbf{A} , \mathbf{B} のように、大文字の太字 (ボールド体, bold) を必ず使う
(a , a , \mathbf{a} , A , A とかではダメ)



ひと目で、この文字は行列をあらわしている！ とわかるメリット

転置行列

- ✓行列の縦と横を入れ替えたもの
- ✓ \mathbf{A} の転置行列を \mathbf{A}^T とあらわす

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \quad \mathbf{A}^T = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 4 & -2 & 1.2 & 6 \\ 4.5 & 0 & -4 & -1.2 \\ 7 & 9.5 & 8.1 & -3 \end{pmatrix}^T = \begin{pmatrix} 4 & 4.5 & 7 \\ -2 & 0 & -1.2 \\ 1.2 & -4 & 8.1 \\ 6 & -1.2 & -3 \end{pmatrix}$$

ベクトル

$$\mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}$$

左のように、縦に2つ以上、横に1つ
数字が並んだものを縦ベクトルとよぶ

$$(3 \quad -1.1 \quad -7)$$

左のように、縦に1つ、横に2つ以上
数字が並んだものを横ベクトルとよぶ

縦ベクトルと横ベクトルとを合わせて、ベクトルとよぶ

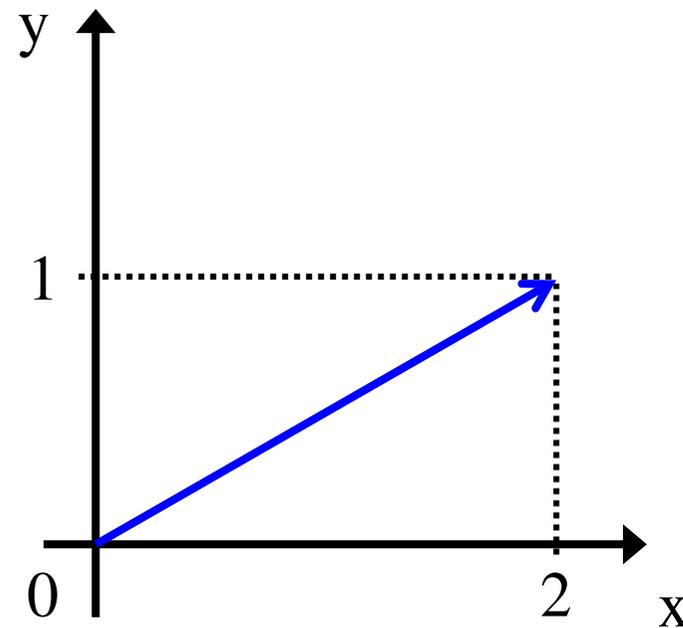
座標系でのベクトル

たとえば、 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ のとき (これは、 $x = 2, y = 1$ をあらわす)、

右のような x, y 座標 (二次元座標) においてベクトルは原点からの矢印で表わされる

(2 1) も右と同じ矢印になる

[縦ベクトルも横ベクトルも同じこと]



(-2 3.4 5.1) は x, y, z 座標 (三次元座標) での矢印、

(1.3 0 -3 -2.1) は四次元座標での矢印、...

ベクトルの大きさ

ベクトルの大きさとは、ベクトルの矢印の長さのこと

各要素の二乗を足し合わせたものの平方根で計算できる

例1) ベクトル (2, 1) の大きさは、 $\sqrt{2^2 + 1^2} = \sqrt{5}$

例2) ベクトル (-2, 3.4, 5.1) の大きさは、

$$\sqrt{(-2)^2 + 3.4^2 + 5.1^2} = 6.44\dots$$

ベクトルの表し方

ベクトルをローマ字であらわすときは、の **a**, **b** のように、
小文字の太字 (ボールド体, bold) を**必ず**使う
(*a*, *a*, **A**, **A**, **A** とかでは**ダメ**)



ひと目で、この文字はベクトルをあらわしている！ とわかるメリット

ベクトルの大きさは、 $\|\mathbf{a}\|$ や $\|\mathbf{b}\|$ のように表される

$$\mathbf{B} = \begin{pmatrix} 4 & -2 & 1.2 & 6 \\ 4.5 & 0 & -4 & -1.2 \\ 7 & 9.5 & 8.1 & -3 \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} 4 \\ 4.5 \\ 7 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} -2 \\ 0 \\ 9.5 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 1.2 \\ -4 \\ 8.1 \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} 6 \\ -1.2 \\ -3 \end{pmatrix} \quad \text{のとき、}$$

$$\mathbf{B} = [\mathbf{a} \quad \mathbf{b} \quad \mathbf{c} \quad \mathbf{d}] = \begin{bmatrix} \mathbf{a}^T \\ \mathbf{b}^T \\ \mathbf{c}^T \\ \mathbf{d}^T \end{bmatrix}^T$$

正方行列

✓縦の長さと同じ横の長さが等しい行列

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1.1 & 0.6 & 11.1 \\ -3.5 & 1.7 & -9.5 \\ 2.6 & -9.4 & 3.2 \end{pmatrix} \quad \begin{pmatrix} 0.5 & 5.4 & -3 & 4 \\ 2.5 & 2 & 2.8 & -3.2 \\ 3 & 6.8 & 5.9 & -0.4 \\ -8.1 & 9 & -2.3 & 10 \end{pmatrix}$$

単位行列

✓ 対角成分が 1 で、他が 0 の正方行列

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \dots$$

✓ \mathbf{I} や \mathbf{E} で表されることが多い

$$\mathbf{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{とか}$$

行列・ベクトルの足し算・引き算

行列・ベクトルの足し算・引き算は、それぞれの**中身**を足し算・引き算する

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

$$(a \ b \ c) + (d \ e \ f) = (a+d \ b+e \ c+f)$$

$$\begin{pmatrix} 2.1 & 5.6 \\ -1.6 & 3.2 \end{pmatrix} + \begin{pmatrix} -0.6 & 9.6 \\ 1.5 & -2.2 \end{pmatrix} = \begin{pmatrix} 1.5 & 15.2 \\ -0.1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} - \begin{pmatrix} g & h \\ i & j \\ k & l \end{pmatrix} = \begin{pmatrix} a-g & b-h \\ c-i & d-j \\ e-k & f-l \end{pmatrix}$$

注意！ 行列やベクトルの**大きさが同じ**でないと足し算・引き算できない

行列同士の掛け算

足し算・引き算と比べると、ちょっと複雑

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 6 & 1 \times 7 + 2 \times 8 \\ 3 \times 5 + 4 \times 6 & 3 \times 7 + 4 \times 8 \end{pmatrix} \\ = \begin{pmatrix} 17 & 23 \\ 39 & 53 \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} = \begin{pmatrix} ag + bj & ah + bk & ai + bl \\ cg + dj & ch + dk & ci + dl \\ eg + fj & eh + fk & ei + fl \end{pmatrix}$$

$$\begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} ga + hc + ie & gb + hd + if \\ ja + kc + le & jb + kd + lf \end{pmatrix}$$

掛け算の順番を
変えたら、答えが
異なった！

行列同士の掛け算のイメージ 1/3

$$\begin{array}{c} 1 \\ \left(\begin{array}{cc} a & b \\ c & d \\ e & f \end{array} \right) \end{array} \begin{array}{c} 1 \\ \left(\begin{array}{ccc} g & h & i \\ j & k & l \end{array} \right) \end{array} = \begin{array}{c} 1 \\ \left(\begin{array}{ccc} ag + bj & ah + bk & ai + bl \\ cg + dj & ch + dk & ci + dl \\ eg + fj & eh + fk & ei + fl \end{array} \right) \end{array}$$

$$\begin{array}{c} 1 \\ \left(\begin{array}{cc} a & b \\ c & d \\ e & f \end{array} \right) \end{array} \begin{array}{c} 2 \\ \left(\begin{array}{ccc} g & h & i \\ j & k & l \end{array} \right) \end{array} = \begin{array}{c} 1 \\ \left(\begin{array}{ccc} ag + bj & ah + bk & ai + bl \\ cg + dj & ch + dk & ci + dl \\ eg + fj & eh + fk & ei + fl \end{array} \right) \end{array}$$

$$\begin{array}{c} 1 \\ \left(\begin{array}{cc} a & b \\ c & d \\ e & f \end{array} \right) \end{array} \begin{array}{c} 3 \\ \left(\begin{array}{ccc} g & h & i \\ j & k & l \end{array} \right) \end{array} = \begin{array}{c} 3 \\ \left(\begin{array}{ccc} ag + bj & ah + bk & ai + bl \\ cg + dj & ch + dk & ci + dl \\ eg + fj & eh + fk & ei + fl \end{array} \right) \end{array}$$

行列同士の掛け算のイメージ 2/3

$$2 \begin{pmatrix} a & b \\ e & d \\ e & f \end{pmatrix} \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} \Rightarrow 2 \begin{pmatrix} ag + bj & ah + bk & ai + bl \\ \boxed{cg + dj} & ch + dk & ci + dl \\ eg + fj & eh + fk & ei + fl \end{pmatrix}$$

Diagram description: A blue arrow points from the second row of the first matrix to the second row of the second matrix. A blue arrow points from the second column of the second matrix to the second row of the first matrix. The element $cg + dj$ in the resulting matrix is boxed in blue.

$$2 \begin{pmatrix} a & b \\ e & d \\ e & f \end{pmatrix} \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} \Rightarrow 2 \begin{pmatrix} ag + bj & ah + bk & ai + bl \\ cg + dj & \boxed{ch + dk} & ci + dl \\ eg + fj & eh + fk & ei + fl \end{pmatrix}$$

Diagram description: A red arrow points from the second row of the first matrix to the second row of the second matrix. A red arrow points from the second column of the second matrix to the second column of the first matrix. The element $ch + dk$ in the resulting matrix is boxed in red.

$$2 \begin{pmatrix} a & b \\ e & d \\ e & f \end{pmatrix} \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} \Rightarrow 2 \begin{pmatrix} ag + bj & ah + bk & ai + bl \\ cg + dj & ch + dk & \boxed{ci + dl} \\ eg + fj & eh + fk & ei + fl \end{pmatrix}$$

Diagram description: A green arrow points from the second row of the first matrix to the second row of the second matrix. A green arrow points from the third column of the second matrix to the second row of the first matrix. The element $ci + dl$ in the resulting matrix is boxed in green.

行列同士の掛け算の注意点

✓ $m \times n$ の行列と、 $p \times q$ の行列との掛け算のとき、

$n = p$ でなければならない (m と q は何でもよい)

- 例: 2×3 の行列と 3×6 の行列、 5×4 の行列と 4×10 の行列

✓ 行列の掛け算の順番を変えたとき、答えが同じになるとは限らない

- 例:
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 17 & 23 \\ 39 & 53 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 26 & 38 \\ 30 & 44 \end{pmatrix}$$

行列とベクトルとの掛け算

ベクトルは行列の一部、として考えると、行列同士の掛け算とやり方は同じ

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \times \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+3y \\ x+2y \end{pmatrix} \quad (\text{p.5の連立方程式})$$

$$(a \quad b \quad c) \begin{pmatrix} d & e & f \\ g & h & i \\ j & k & l \end{pmatrix} = (ad+bg+cj \quad ae+bh+ck \quad af+bi+cl)$$

ベクトル同士の掛け算

ベクトルは行列の一部、として考えると、行列・ベクトルの掛け算と同じ

$$(2 \ 3) \times \begin{pmatrix} 4 \\ 5 \end{pmatrix} = (2 \ 3) \begin{pmatrix} 4 \\ 5 \end{pmatrix} = 2 \times 4 + 3 \times 5 = 23$$

$$(a \ b \ c) \times \begin{pmatrix} d \\ e \\ f \end{pmatrix} = (a \ b \ c) \begin{pmatrix} d \\ e \\ f \end{pmatrix} = ad + be + cf$$

$$\begin{pmatrix} 4 \\ 5 \end{pmatrix} \times (2 \ 3) = \begin{pmatrix} 4 \\ 5 \end{pmatrix} (2 \ 3) = \begin{pmatrix} 4 \times 2 & 4 \times 3 \\ 5 \times 2 & 5 \times 3 \end{pmatrix} = \begin{pmatrix} 8 & 12 \\ 10 & 15 \end{pmatrix}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times (d \ e \ f) = \begin{pmatrix} a \\ b \\ c \end{pmatrix} (d \ e \ f) = \begin{pmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{pmatrix}$$

逆行列

- ✓ある正方行列について、掛けると単位行列になる正方行列
- ✓ \mathbf{A} の逆行列は、 \mathbf{A}^{-1} で表される

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \quad \mathbf{A}^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

$$\mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{E}$$

$$\mathbf{A}^{-1}\mathbf{A} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{E}$$

逆行列と連立方程式 1/2

✓ 逆行列は連立方程式を解くことに対応

$$2x + 3y = 7$$

$$x + 2y = 4$$

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 7 \\ 4 \end{pmatrix} \quad \text{とおくと、(p.21より)}$$

$\mathbf{Ab} = \mathbf{c}$ の両辺に左から \mathbf{A}^{-1} をかけると、
(p.20 の行列の掛け算の注意点にあるように、
右から掛けるか左から掛けるかも大事)

$$\mathbf{A}^{-1} \mathbf{Ab} = \mathbf{A}^{-1} \mathbf{c}$$

$$\mathbf{Eb} = \mathbf{A}^{-1} \mathbf{c}$$

逆行列と連立方程式 2/2

$$\mathbf{E}\mathbf{b} = \mathbf{A}^{-1}\mathbf{c}$$

$$\mathbf{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}, \mathbf{A}^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \quad \text{より、}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

逆行列の計算

✓逆行列の計算方法として、

- 掃き出し法
- 余因子法

が存在する

✓いろいろなプログラミング言語で逆行列を計算する関数が用意されているため、利用するとよい

✓例

- Python: `numpy.linalg.inv`
- MATLAB: `inv`
- R: `solve`

逆行列の応用先

- ✓ 最小二乗法による線形重回帰分析

<https://datachemeng.com/ordinaryleastsquares/>

- ✓ 部分的最小二乗回帰(Partial Least Squares Regression, PLS)

<https://datachemeng.com/partialleastsquares/>

など

逆行列を計算できない場合

✓線形従属 (一次従属) のベクトルが存在するとき

$$\begin{pmatrix} 3 & 1 & -2 \\ 9 & 3 & 1 \\ 6 & 2 & -3 \end{pmatrix}$$

において、一列目と二列目は

$$\begin{pmatrix} 3 \\ 9 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

のように、一列目の定数倍で二列目が表される

このようなベクトル間の関係を、線形従属 (一次従属) とよぶ

行列の階数 (ランク)

✓線形独立 (一次独立) なベクトルの数のことを、
行列の階数 (ランク) と呼ぶ

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \quad \text{のランクは 2}$$

$$\begin{pmatrix} 1.1 & 0.6 & 11.1 \\ -3.5 & 1.7 & -9.5 \\ 2.6 & -9.4 & 3.2 \end{pmatrix} \quad \text{のランクは 3}$$

$$\begin{pmatrix} 3 & 1 & -2 \\ 9 & 3 & 1 \\ 6 & 2 & -3 \end{pmatrix} \quad \text{のランクは 2}$$

- ✓ 正方行列に対して与えられる
- ✓ \mathbf{A} の行列式を $\det(\mathbf{A})$ もしくは $|\mathbf{A}|$ で表す

\mathbf{A} が逆行列をもたない \Leftrightarrow \mathbf{A} に線形従属なベクトルがある $\Leftrightarrow \det(\mathbf{A}) = 0$

\mathbf{A} が 2×2 の行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ のとき、

$$\det(\mathbf{A}) = ad - bc$$

線形変換

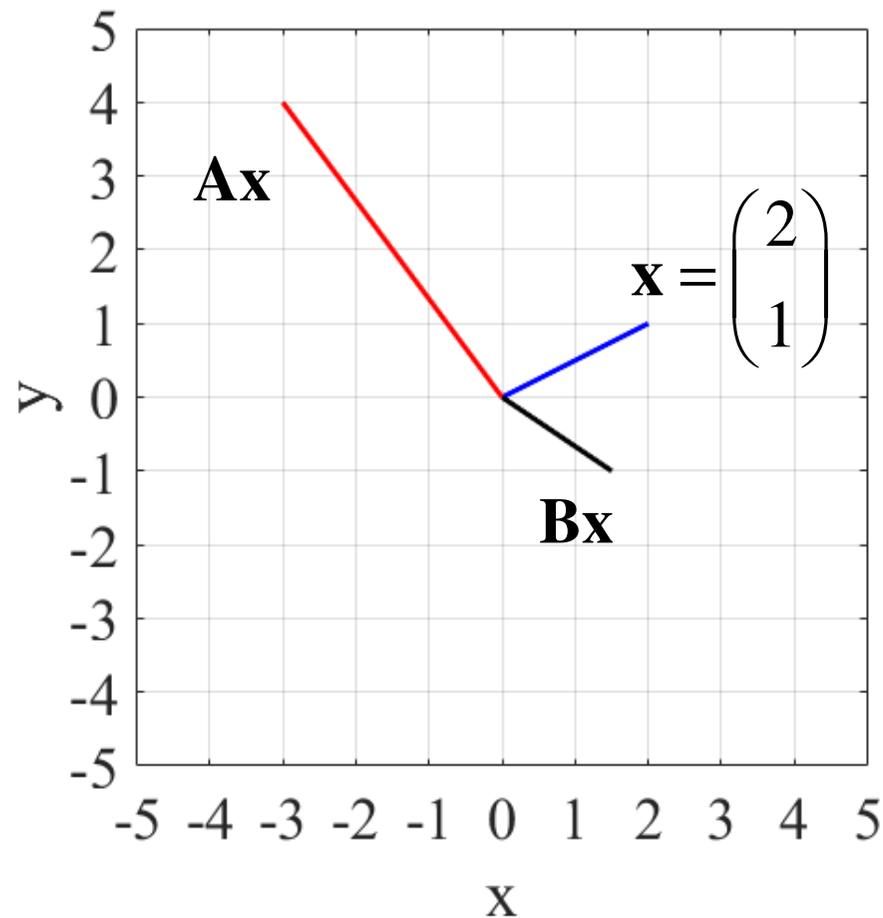
✓ある縦ベクトル \mathbf{x} に対して、左から正方行列をかけることを、線形変換とよぶ

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \text{ のとき、 } \mathbf{Ax} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2x + y \\ x + 2y \end{pmatrix}$$

線形変換 意味合い

- ✓線形変換は、ベクトルの回転 & 伸縮
- ✓どんな回転・伸縮になるかは、正方行列・ベクトルによって異なる

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$$



$$\mathbf{B} = \begin{pmatrix} 0.5 & 0.5 \\ -1 & 1 \end{pmatrix}$$

✓ 正方行列 A に対して、

$$Ax = \lambda x$$

となる λ を固有値、 x を固有ベクトルとよぶ

✓ 意味合い： あるベクトルを線形変換 (ベクトルの回転 & 伸縮) したときに、**向きが同じで長さが定数倍**になった

- そのベクトルが固有ベクトル
- 定数が固有値

✓ 固有値問題： 固有値・固有ベクトルを見つける問題

固有値・固有ベクトルの計算

✓固有値問題を解く方法として、 A の固有方程式

$$\det(\mathbf{A} - \lambda\mathbf{E}) = 0$$

を解く方法が存在する

✓いろいろなプログラミング言語で固有値・固有ベクトルを計算する関数が用意されているため、利用するとよい

✓例

- Python: `numpy.linalg.eig`
- MATLAB: `eig`
- R: `eigen`

固有値・固有ベクトルの応用先

- ✓主成分分析(Principal Component Analysis, PCA)
<https://datachemeng.com/principalcomponentanalysis/>
- ✓部分的最小二乗回帰(Partial Least Squares Regression, PLS)
<https://datachemeng.com/partialleastsquares/>

など

偏微分

- ✓ 複数の変数をもつ関数に対して、一つの変数に着目して、他の変数は定数とみなして、微分すること
- ✓ 関数 $f(x, y, z)$ を x で偏微分することを

$$\frac{\partial f(x, y, z)}{\partial x}$$

であらわす

例) $f(x, y, z) = xyz + x^2y + x^{-1}z^3 + y^{-2}z + 3$ のとき、

$$\frac{\partial f(x, y, z)}{\partial x} = yz + 2xy - x^{-2}z^3$$

全微分

- ✓ 複数の変数をもつ関数に対して、すべての変数が微小変化したときの関数の変化を表現したもの

関数 $z = f(x, y)$ とすると、

$x \rightarrow x + dx$

$y \rightarrow y + dy$

だけ微小変化したとき、 z の微小変化 $z \rightarrow z + dz$ は

$$dz = \frac{\partial f(x, y)}{\partial x} dx + \frac{\partial f(x, y)}{\partial y} dy$$

とあらわされる

Lagrangeの未定乗数法

- ✓ 複数の変数をもつ関数を、制約条件があるなかで最大化 (もしくは最小化) する方法

たとえば、2変数 x, y として、最大化したい関数を $f(x, y)$, 制約条件を $g(x, y) = 0$ とする

Lagrangeの未定乗数法では、ラグランジュ定数を λ として、

$$F(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

とするとき、

$$\frac{\partial F(x, y, \lambda)}{\partial x} = 0 \quad \frac{\partial F(x, y, \lambda)}{\partial y} = 0 \quad \frac{\partial F(x, y, \lambda)}{\partial \lambda} = 0$$

をすべて満たす点が、 $f(x, y)$ を最大にする点となる

Lagrangeの未定乗数法の雑な証明

$$\frac{\partial F(x, y, \lambda)}{\partial \lambda} = 0 \quad \text{は制約条件 } g(x, y) = 0 \text{ と同じ}$$

$$\frac{\partial F(x, y, \lambda)}{\partial x} = 0 \quad \frac{\partial F(x, y, \lambda)}{\partial y} = 0 \quad \Rightarrow \quad \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} = \lambda \begin{pmatrix} \frac{\partial g(x, y)}{\partial x} \\ \frac{\partial g(x, y)}{\partial y} \end{pmatrix}$$

➡ c を定数として、 $f(x, y) = c$ と $g(x, y) = 0$ のそれぞれの勾配ベクトル (法線ベクトル、曲線に垂直なベクトルのこと) が平行

➡ $f(x, y) = c$ と $g(x, y) = 0$ が接する

接しない、つまり $f(x, y) = c$ と $g(x, y) = 0$ が交わるとき、 $g(x, y) = 0$ で $f(x, y) > c$ となる点が存在する

➡ $f(x, y) = c$ と $g(x, y) = 0$ が接する点において、 c が最大

勾配ベクトル、法線ベクトル 1/2

曲線 $f(x, y) = c$ において、ある点 (x, y) から曲線上に $(\Delta x, \Delta y)$ だけ微小変化させる

$(\Delta x, \Delta y)$ は接線ベクトル

$$\begin{aligned} f(x + \Delta x, y + \Delta y) &= f(x, y) + \frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y \\ &= c + \frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y \end{aligned}$$

曲線上の変化なので、 $f(x + \Delta x, y + \Delta y) = c$

よって、
$$\frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y = 0$$

勾配ベクトル、法線ベクトル 2/2

$$\frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y = 0 \quad \Rightarrow \quad \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0$$

$$\begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix}$$

は、 $f(x, y) = c$ の接線ベクトル $(\Delta x, \Delta y)$ に垂直なベクトル
(法線ベクトル)

このベクトルを、勾配ベクトルとよぶ

Lagrangeの未定乗数法の応用先

✓部分的最小二乗回帰(Partial Least Squares Regression, PLS)

<https://datachemeng.com/partialleastsquares/>

✓サポートベクターマシン(Support Vector Machine, SVM)

<https://datachemeng.com/supportvectormachine/>

など

✓ある事象 A が起こる確率は $p(A)$ とあらわされる

- 例) $p(\text{サイコロを振って } 1 \text{ が出る}) = 1/6$

✓確率変数 X の値が x_i となる確率は $p(X = x_i)$ とあらわされる

- 上の例のとき、 $p(X = 1) = 1/6$

✓確率変数 X が任意の値をもつとき、“ $= x_i$ ” を省略して $p(X)$ とあらわす

同時確率・条件付き確率

✓ 2つの確率変数 X, Y が、それぞれ $X = x_i, Y = y_j$ となる確率を **同時確率** とよび、 $p(X = x_i, Y = y_j)$ とあらわす

- 例) X : サイコロPを振る、 Y : サイコロQを振る、のとき、
 $p(X = 2, Y = 3) = 1/36$

✓ $X = x_i$ の場合だけを考えたとき、 $Y = y_j$ となる確率を、 $X = x_i$ が与えられた下での $Y = y_j$ の **条件付き確率** とよび、 $p(Y = y_j | X = x_i)$ とあらわす

- 例) X : サイコロPを振る、 Y : サイコロQを振る、のとき、
 $p(Y = 3 | X = 2) = 1/6$

X:喫煙・Y:パチンコ 人口

- ✓日本の全人口(2016年) : およそ 12,000 万人 [1]

- ✓ $X=1$ ・・・喫煙者(2016年) : およそ 2,000 万人 [2]
 - $X=0$ ・・・非喫煙者(2016年) : およそ 10,000 万人 [2]

- ✓ $Y=1$ ・・・パチンコ参加者(2016年) : およそ 1,000 万人 [3]
 - $Y=0$ ・・・パチンコ非参加者(2016年) : およそ 11,000 万人 [3]

- ✓ $X=1$ かつ $Y=1$ (2015年)・・・ およそ 500 万人 [3]
 - $X=1$ かつ $Y=0$ (2015年)・・・ およそ 1,500 万人 [2][3]
 - $X=0$ かつ $Y=1$ (2015年)・・・ およそ 500 万人 [2][3]
 - $X=0$ かつ $Y=0$ (2015年)・・・ およそ 9,500 万人 [2][3]

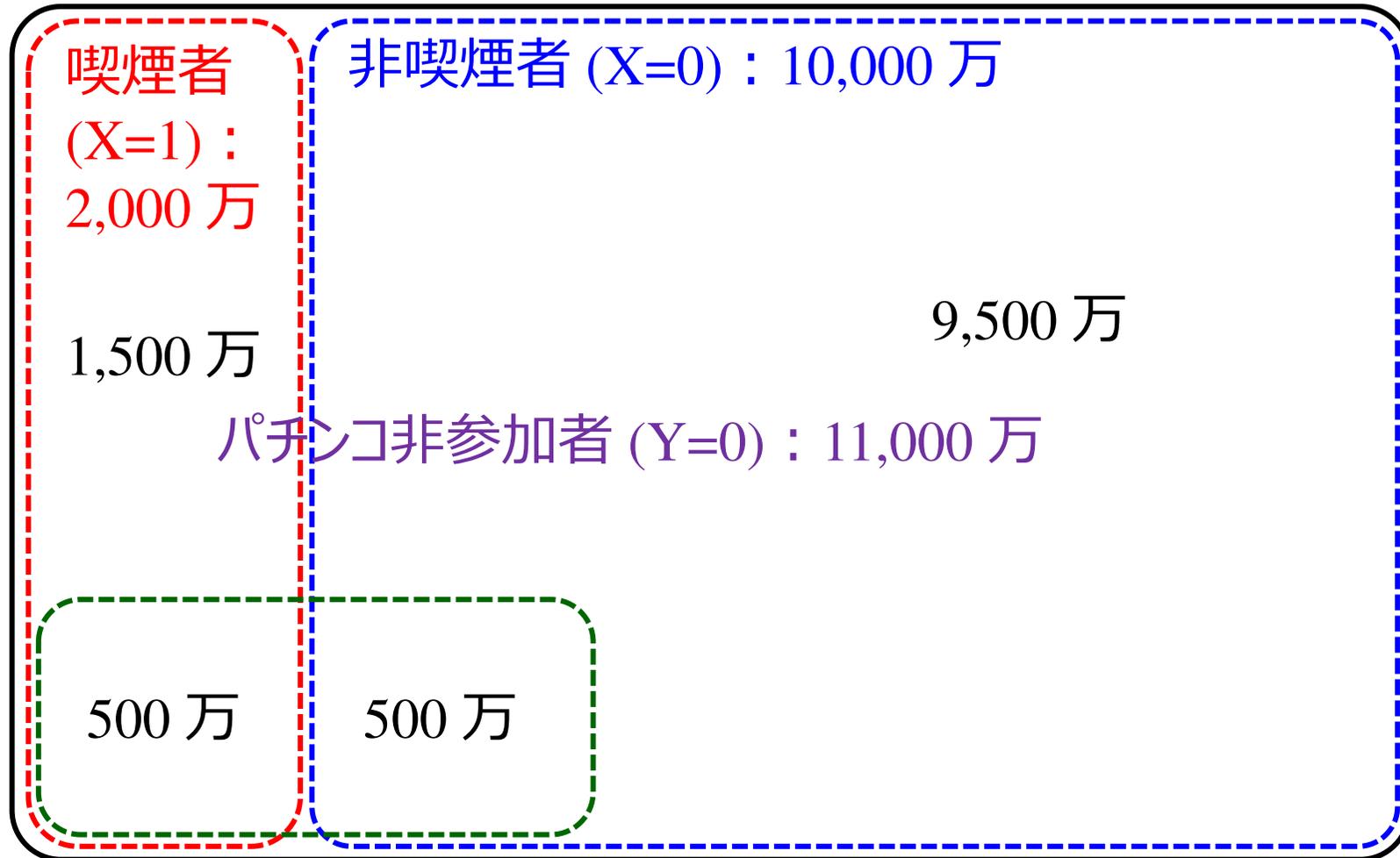
[1] <https://ja.wikipedia.org/wiki/%E6%97%A5%E6%9C%AC%E3%81%AE%E4%BA%BA%E5%8F%A3%E7%B5%B1%E8%A8%88>

[2] https://www.jti.co.jp/investors/library/press_releases/2016/0728_01.html

[3] http://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000110201_3.pdf

X:喫煙・Y:パチンコ ベン図

全人口：12,000 万



パチンコ参加者 (Y=1)：1,000 万

X:喫煙・Y:パチンコ 同時確率

$$\checkmark p(X = 1, Y = 1) = 500 / 12,000 = 0.04$$

$$\checkmark p(X = 1, Y = 0) = 1,500 / 12,000 = 0.13$$

$$\checkmark p(X = 0, Y = 1) = 500 / 12,000 = 0.04$$

$$\checkmark p(X = 0, Y = 0) = 9,500 / 12,000 = 0.79$$

- すべて足すと 1 になる

X:喫煙・Y:パチンコ 条件付き確率

$$\checkmark p(X = 1 | Y = 1) = 500 / 1,000 = 0.50$$

- $p(X = 0 | Y = 1) = 500 / 1,000 = 0.50 (= 1 - 0.50)$

$$\checkmark p(X = 1 | Y = 0) = 1,500 / 11,000 = 0.14$$

- $p(X = 0 | Y = 0) = 9,500 / 11,000 = 0.86 (= 1 - 0.14)$

$$\checkmark p(Y = 1 | X = 1) = 500 / 2,000 = 0.25$$

- $p(Y = 0 | X = 1) = 1,500 / 2,000 = 0.75 (= 1 - 0.25)$

$$\checkmark p(Y = 1 | X = 0) = 500 / 10,000 = 0.05$$

- $p(Y = 0 | X = 0) = 9,500 / 10,000 = 0.95 (= 1 - 0.05)$

確率の加法定理

✓加法定理

$$p(Y) = \sum_X p(X, Y)$$

例) 前ページのサイコロ

$$\begin{aligned} p(Y=3) &= p(X=1, Y=3) + p(X=2, Y=3) + p(X=3, Y=3) + \\ &\quad p(X=4, Y=3) + p(X=5, Y=3) + p(X=6, Y=3) \\ &= 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 \\ &= 1/6 \end{aligned}$$

- X についての周辺化とも呼ばれる
- $p(Y)$: 周辺確率

X:喫煙・Y:パチンコ 確率の加法定理

$$\checkmark p(X = 1) = p(X = 1, Y = 1) + p(X = 1, Y = 0) = 0.04 + 0.13 = 0.17 \\ (= 2,000/12,000 = 0.17)$$

$$\checkmark p(X = 0) = p(X = 0, Y = 1) + p(X = 0, Y = 0) = 0.04 + 0.79 = 0.83 \\ (= 10,000/12,000 = 0.83)$$

$$\checkmark p(Y = 1) = p(X = 0, Y = 1) + p(X = 1, Y = 1) = 0.04 + 0.04 = 0.08 \\ (= 1,000/12,000 = 0.08)$$

$$\checkmark p(Y = 0) = p(X = 1, Y = 0) + p(X = 0, Y = 0) = 0.13 + 0.79 = 0.92 \\ (= 11,000/12,000 = 0.92)$$

確率の乗法定理

✓乗法定理

$$p(X, Y) = p(Y | X) p(X)$$

- 意味合い： X の確率に、 X が与えられたときの Y の確率をかけると X と Y が同時に起こる確率

X:喫煙・Y:パチンコ 確率の乗法定理

$$\begin{aligned}\checkmark p(X=1, Y=1) &= p(X=1 | Y=1) p(Y=1) = 0.50 \times 0.08 = 0.04 \\ &= p(Y=1 | X=1) p(X=1) = 0.25 \times 0.17 = 0.04\end{aligned}$$

$$\begin{aligned}\checkmark p(X=1, Y=0) &= p(X=1 | Y=0) p(Y=0) = 0.14 \times 0.92 = 0.13 \\ &= p(Y=0 | X=1) p(X=1) = 0.75 \times 0.17 = 0.13\end{aligned}$$

$$\begin{aligned}\checkmark p(X=0, Y=1) &= p(X=0 | Y=1) p(Y=1) = 0.50 \times 0.08 = 0.04 \\ &= p(Y=1 | X=0) p(X=0) = 0.05 \times 0.83 = 0.04\end{aligned}$$

$$\begin{aligned}\checkmark p(X=0, Y=0) &= p(X=0 | Y=0) p(Y=0) = 0.86 \times 0.92 = 0.79 \\ &= p(Y=0 | X=0) p(X=0) = 0.95 \times 0.83 = 0.79\end{aligned}$$

ベイズの定理

確率の乗法定理より、

$$\begin{aligned} p(X, Y) &= p(Y | X) p(X) \\ &= p(X | Y) p(Y) \end{aligned}$$

よって、

$$p(X | Y) p(Y) = p(Y | X) p(X)$$

$$p(X | Y) = \frac{p(Y | X) p(X)}{p(Y)}$$

確率の加法定理より、

$$p(Y) = \sum_X p(X, Y) = \sum_X p(Y | X) p(X)$$

ベイズの定理：

$$p(X | Y) = \frac{p(Y | X) p(X)}{\sum_X p(Y | X) p(X)}$$

ベイズの定理 メリット

$$\text{ベイズの定理: } p(X|Y) = \frac{p(Y|X)p(X)}{\sum_Y p(Y|X)p(X)}$$

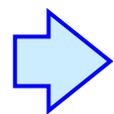
X が与えられたときの Y の条件付き確率と X の周辺確率のみから、
 Y が与えられたときの X の条件付き確率 を計算できる

$p(X)$: Y が与えられる前の X の確率 (事前確率)

$p(Y)$: X が与えられる前の Y の確率 (事前確率)

$p(Y|X)$: X が与えられた後の Y の確率 (事後確率)

$p(X|Y)$: Y が与えられた後の X の確率 (事後確率)



X の事前確率と X が与えられた後の Y の事後確率のみから、
 Y が与えられた後の X の事後確率 を計算できる

X:喫煙・Y:パチンコ ベイズの定理

$$\begin{aligned} p(X=1|Y=1) &= \frac{p(Y=1|X=1)p(X=1)}{p(Y=1|X=1)p(X=1)+p(Y=1|X=0)p(X=0)} \\ &= \frac{0.25 \times 0.17}{0.25 \times 0.17 + 0.05 \times 0.83} \\ &= 0.50 \end{aligned}$$

- 喫煙者の確率と
喫煙者におけるパチンコ利用者の確率のみから、
パチンコ利用者における喫煙者の確率を計算できた
- パチンコで出口調査をしなくても、パチンコ利用者における
喫煙者の確率がわかる

確率・ベイズの定理の応用先

✓ Generative Topographic Mapping (GTM)

<https://datachemeng.com/generativetopographicmapping/>

✓ ガウス過程による回帰 (Gaussian Process Regression, GPR)

<https://datachemeng.com/gaussianprocessregression/>

など