立教大学大学院人工知能科学研究科

# 数理科学概論

第2回

2023/4/22

担当：石川真之介

本日の内容

- 高校数学のベクトル (+数列、三角比) の復習 (+$\alpha$)

# 授業の目標

機械学習の仕組みを数学用語を使って
説明できるようになる

# 今週の目標

ベクトル、三角比を思い出し、
ベクトルがデータ分析にどのように使われるか
イメージできるようになる

# 本日の内容

1. 数理科学と機械学習との関係の概要
2. 高校数学の復習： ベクトル
3. 高校数学の復習： 微分、積分
4. 高校数学の復習： 場合の数、確率
5. 行列の計算
6. 関数と微分・偏微分
7. 多変数関数の勾配
8. 確率、統計の基礎
9. 情報量の基礎
10. 最適化問題(1) 最適化問題の概要
11. 最適化問題(2) 制約なし最適化と解法アルゴリズム
12. 最適化問題(3) 線形計画問題、双対定理
13. 最適化問題(4) ラグランジュ未定乗数法
14. 最適化問題(5) サポートベクターマシンの数理

# ベクトルの復習 (+数列、三角比)

# 数列とは数の列

いくつかの数値を並べたものを数列という。

1, 2, 3, 4, …

0, 1, 0, 1, 1, 1, 1, 2, 2, 1, 1, 0, 2

数列の各数値を項、最初の項を初項と呼ぶ。項の数は項数と呼び、
有限の場合、無限の場合がある。i 番目の項を $a_i$ のように右下に添字を
つけて表す。つまり単に「i 番目の a」ということ。かんたん。

高校数学では主に法則性のある数列を扱っていた。

例：3, 5, 7, 9, 11, ... (等差数列)、
1, 3, 9, 27, 81, ... (等比数列)

これらの場合、各項は i を使った式で表すことができる
(i の値が決まると、i 番目の項の値 $a_i$ が決まる)

データサイエンスではこのような例はあまり出てこない。
(各項が全く関係ないか、関係はあるが一般項を式で書けない場合が多い)

# 数列の和はΣで表す

法則性がある場合でもない場合でも、数列の和を扱う必要が出てくる場合がしばしばある。数列の和は、Σ (シグマ) 記号で表す。

例えば、1, 2, 3 という数列の和は、i 番目の値が i であるため
一般項は $a_i = i$ と書けるため、以下のように記述する。

$$\sum_{i=1}^{3} i$$

Σ の下に示した式の左辺のパラメータ i について、初期値が右辺 1 となり、Σ の上に示した値まで変化する。

項数が n の数列 $a_n$ の和は、一般には以下の式の左辺のように書ける。
その内容は右辺の通り。

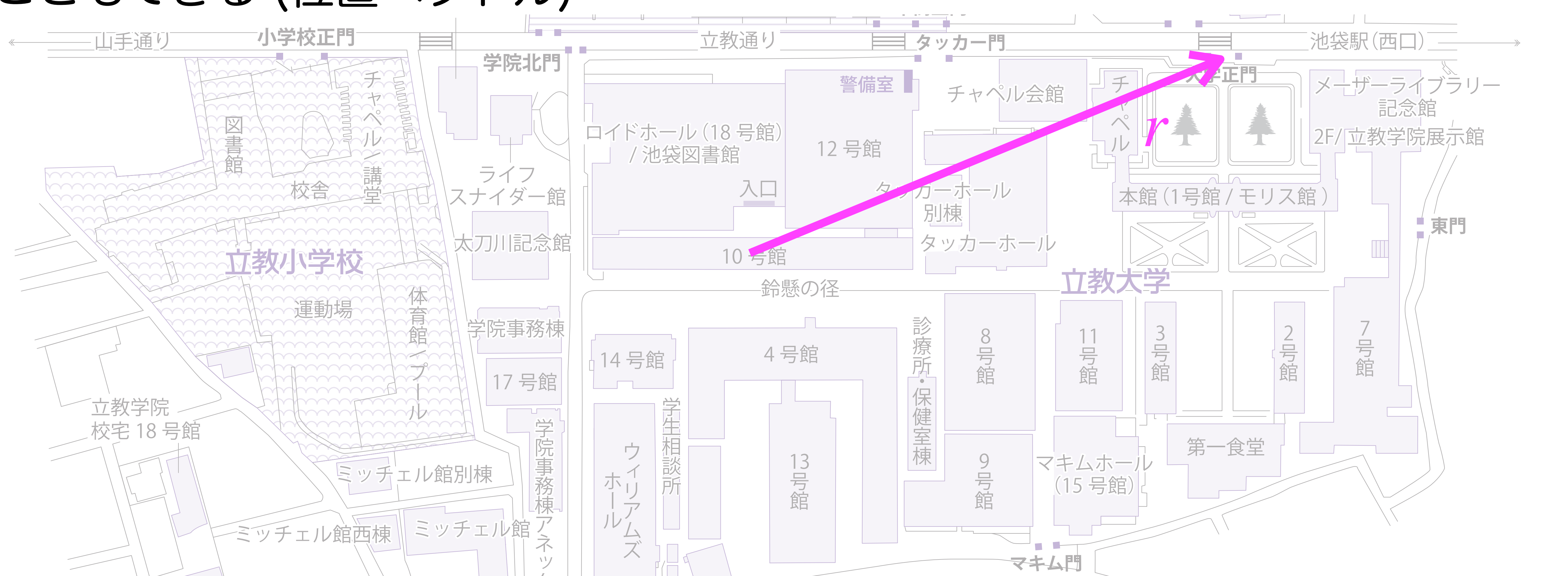
$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n$$

# ベクトルは「大きさと方向を持つ量」

x, y 等の文字で表される単なる大きさのみの数値 (スカラー) に対し、「大きさ」と「方向」を持つ量はベクトルと呼ばれる。

例えば、11号館からみた正門の位置は「北北東」に「100 m」のベクトルと捉えることができる。

ベクトルは<u>座標ではない</u>。例えば、正門の位置を表す量ではない。ただし、出発点をあらかじめ決めておけば、ベクトルで位置を表すこともできる (位置ベクトル)

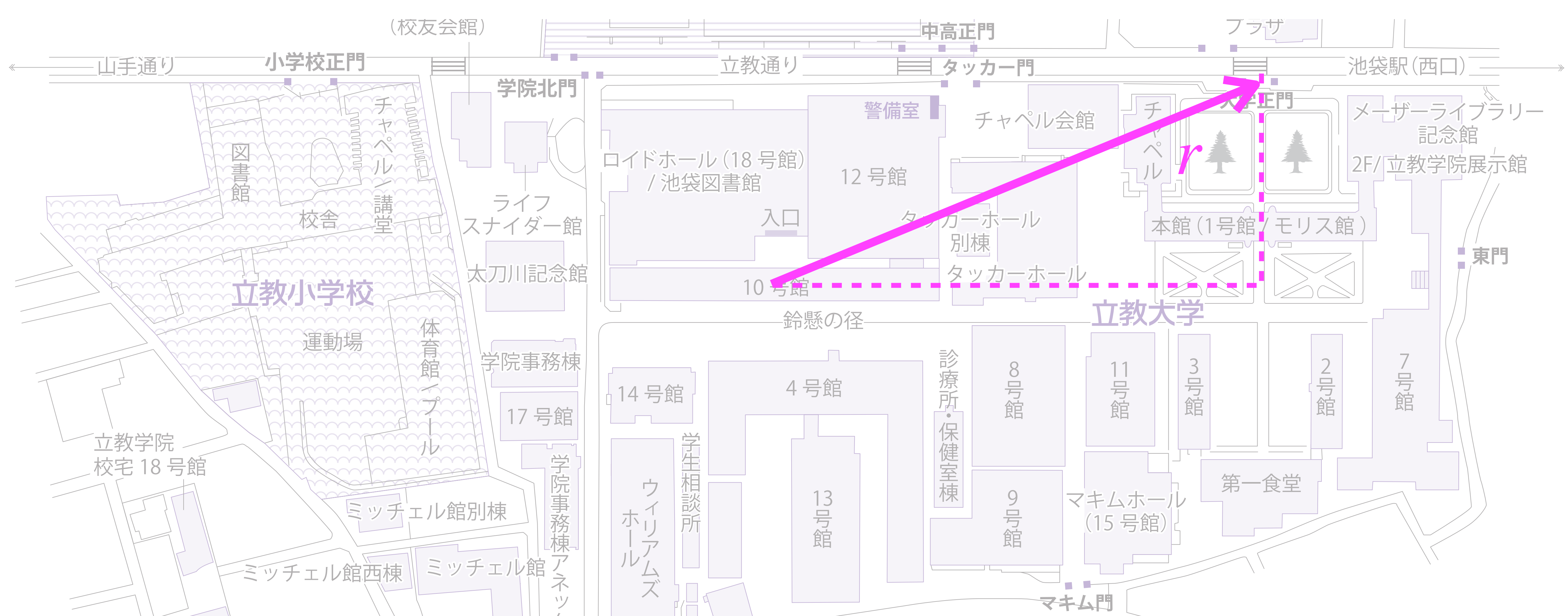



# ベクトルの成分は「基準方向で分解した表記」

あるベクトルを表現するには、大きさと方向を表現する方法以外に、
「東に○○ m」「北に○○ m」という、東西方向と南北方向に
それぞれどれだけ進めばよいか、という形で表現することもできる。

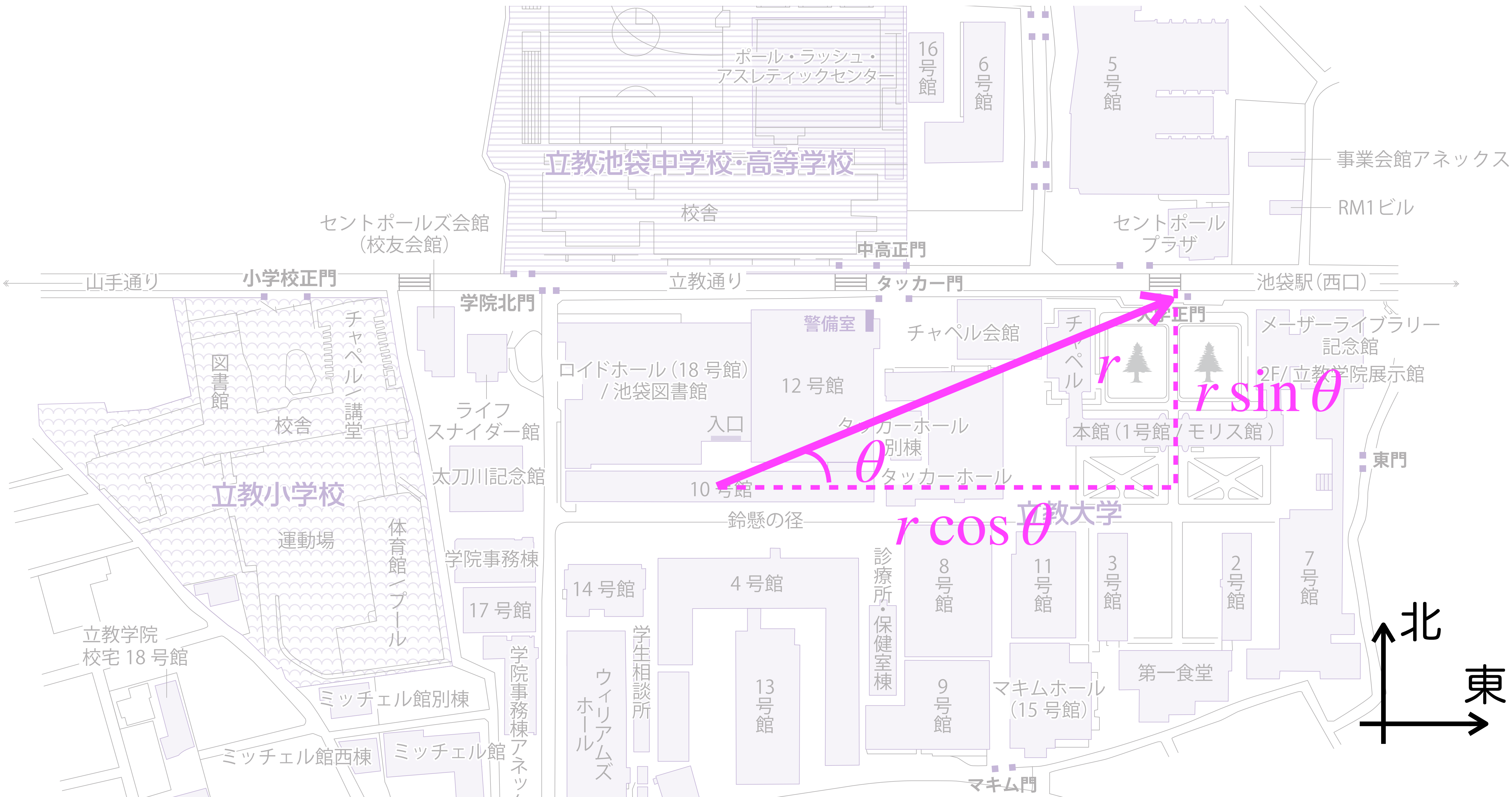
ベクトルのうち東西方向等、特定の方向の要素のみを
取り出した量を成分と呼ぶ。

例えば、あるベクトルに対して、x方向の要素をx成分と呼ぶ、等。



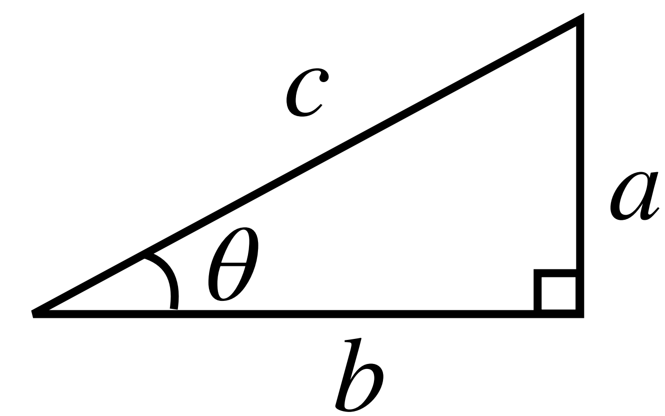
# ベクトルの成分を方角で示す方法

ベクトルの大きさを r として、東西方向とベクトルとのなす角を $\theta$ とすると、東西成分は $r\cos\theta$、南北成分は $r\sin\theta$ と表すことができる。



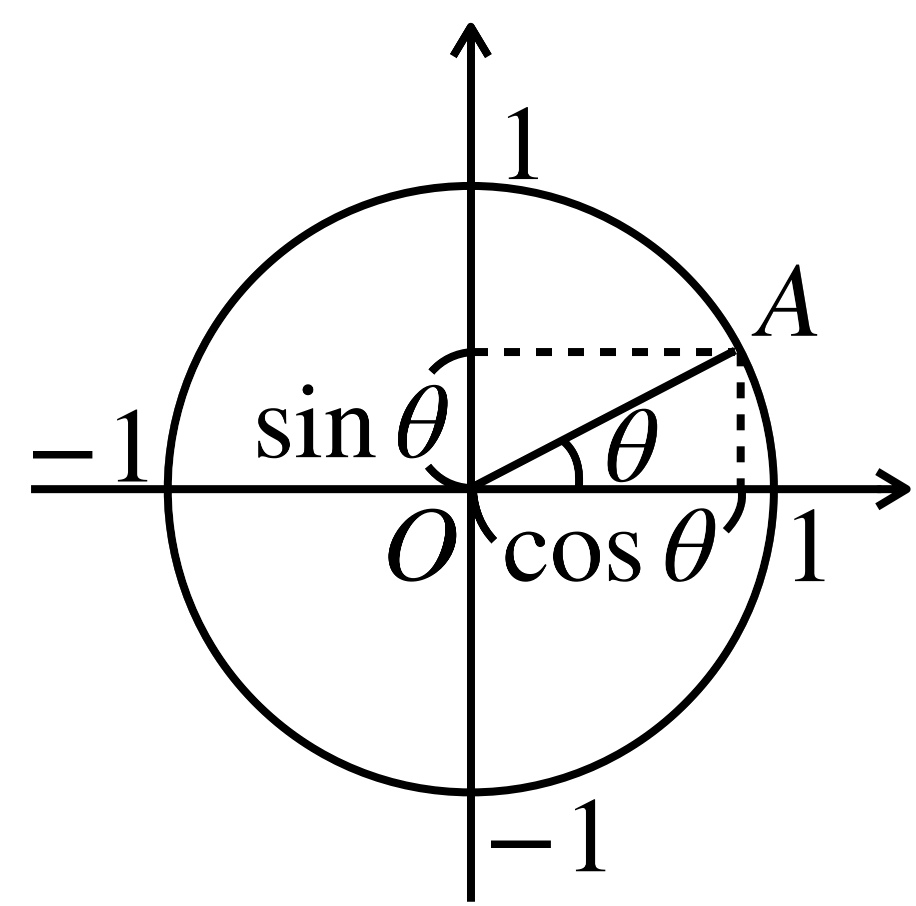
# …… sin, cos ってなんだっけ。。。

$\theta$ が0°から90° の間の場合、
内角のひとつが $\theta$ である右のような直角三角形を
考えると、sin, cos, tan は以下のようになる。




$$\sin\theta = \frac{a}{c}, \quad \cos\theta = \frac{b}{c}, \quad \tan\theta = \frac{a}{b}$$

$\theta$ が90° 以上の場合も含め、より一般的には：
半径1の円周上の点Aに対して、
x軸と直線 OA のなす角が $\theta$ だった場合に、
A のx座標が $\cos\theta$, y座標が $\sin\theta$




直線 OA の傾きに相当する、
[A の垂直座標]/[A の水平座標] が $\tan\theta$
そのため、以下の関係が成り立つ

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

# ベクトルの成分と大きさの記法

点 A から点 B に向かうベクトル：$\overrightarrow{\mathrm{AB}}$

あるベクトル x： $\overrightarrow{x}$ もしくは $\mathbf{x}$ (太字)

機械学習系の文献には太字表記のものも多いので戸惑わないように注意。

参考：かっこいい太字ベクトルの手書き方法 (ファインマン物理学 Volume 2, Chapter 2 より)

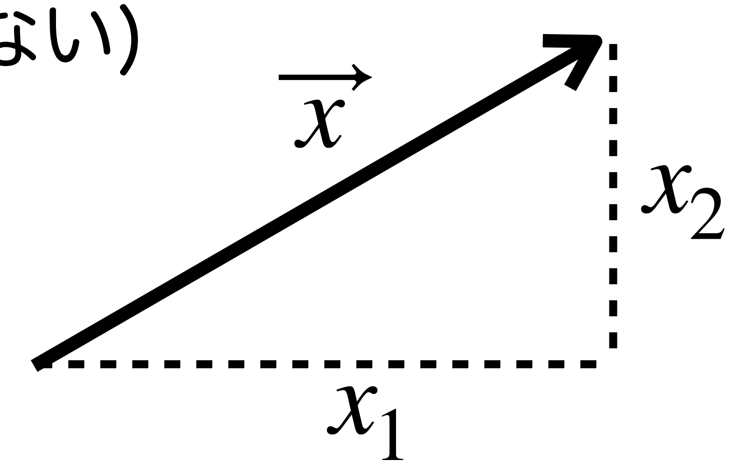
https://www.feynmanlectures.caltech.edu/II_02.html

$\overrightarrow{x}$ の x 成分が $x_1$、y 成分が $x_2$ の場合の表記

$$\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

大学入学共通テスト範囲の高校数学では横に並べて表記するが、
今後の応用もふまえここでは縦並び表記を使うことが多い

ベクトルの大きさ：$|\overrightarrow{x}|$、$\|\overrightarrow{x}\|$ もしくは $x$ (太字でない)
ノルムとも呼ばれる (正確にはユークリッドノルム)



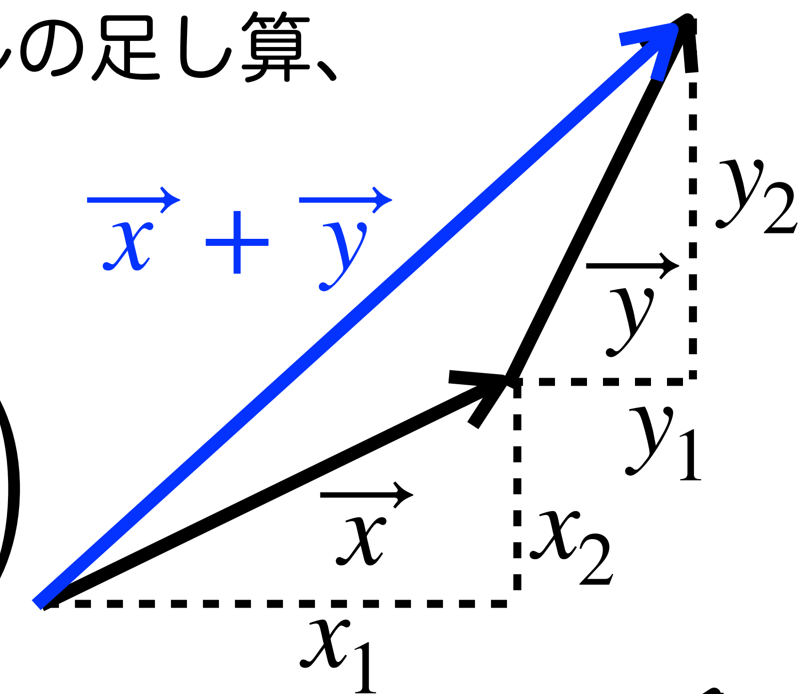

三平方の定理から、$|\overrightarrow{x}|^2 = x_1^2 + x_2^2 = \sum_{i=1}^{2} x_i^2$

# ベクトルの足し算/引き算は成分を足す/引く

2つのベクトル $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \overrightarrow{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ に対して、

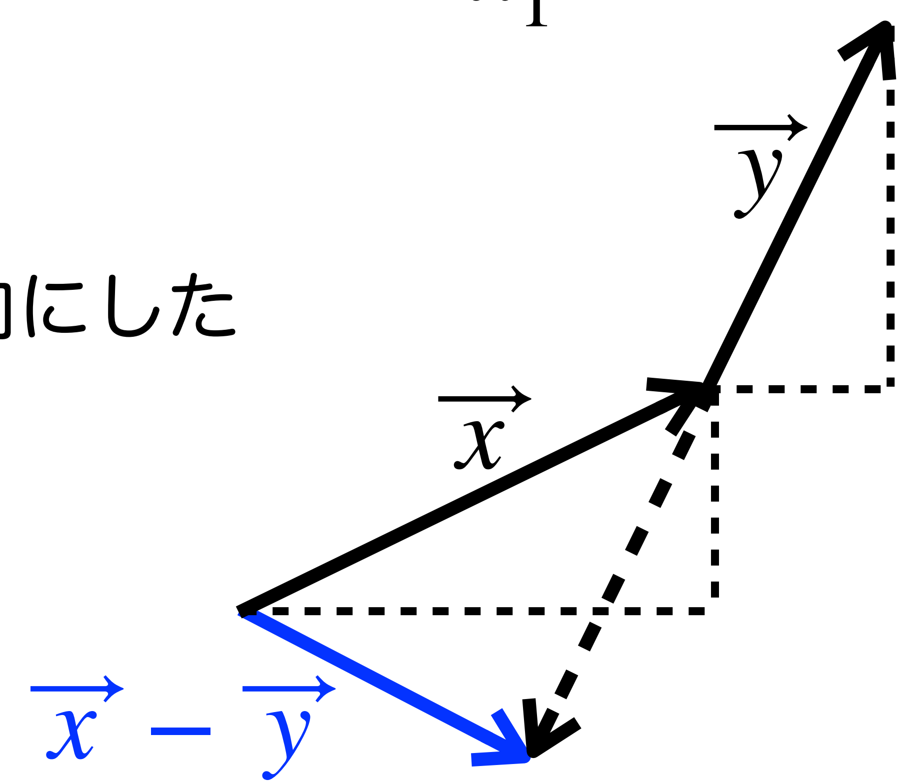
成分ごとに足し算してベクトルを得ることでベクトルの足し算、
成分ごとにを引き算してベクトルを得ることで
ベクトルの引き算を行うことができる。

$$\overrightarrow{x} + \overrightarrow{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}, \ \overrightarrow{x} - \overrightarrow{y} = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix}$$




足し算の結果は、$\overrightarrow{x}$ と $\overrightarrow{y}$ をつないだ
始点から終点に向かうベクトルに相当する。

引き算の結果は、$\overrightarrow{x}$ に対して $\overrightarrow{y}$ の向きを逆方向にした
ベクトルを足した結果に相当する。

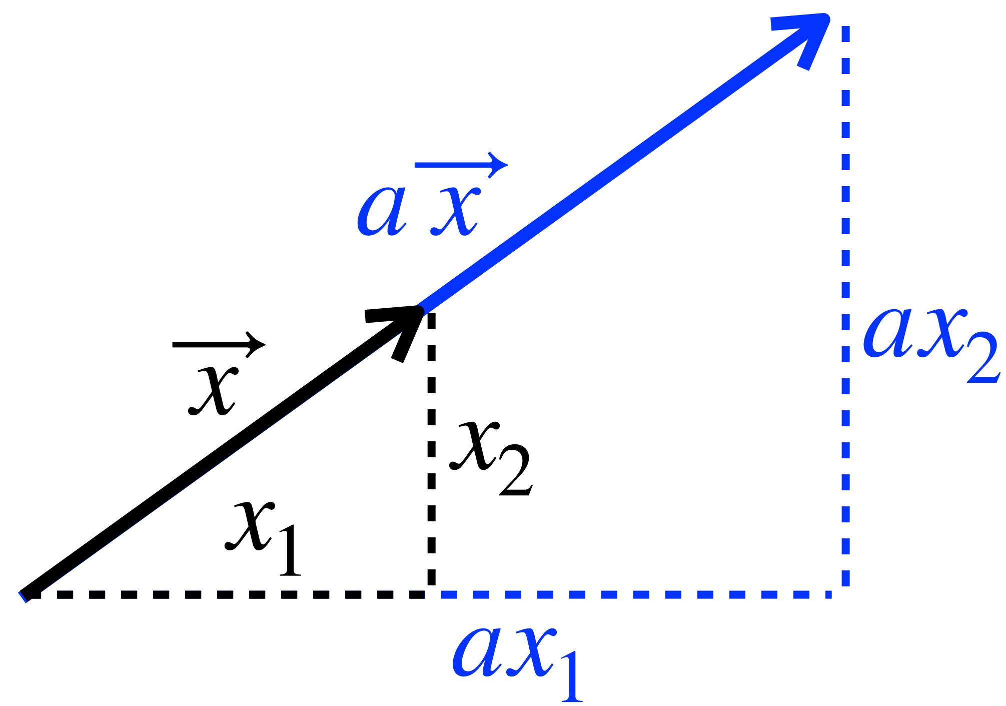


# ベクトルのスカラー倍は向きは同じで大きさが変わる

ベクトルに対してスカラー倍するという演算を考えることができ、
結果は方向が同じで大きさがスカラー倍されたベクトルとなる。

例えば、$\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ を a 倍すると、

$$a\overrightarrow{x} = a\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \end{pmatrix}$$




と、各成分が a 倍されることとなる。

a が負の場合には、スカラー倍の計算結果は向きが180° 変わる。
例えば、あるベクトルを-3倍する場合、大きさは3倍になり向きは逆方向になる。

ベクトルの引き算は、-1倍したベクトルを足すと捉えることもできる。

# ベクトル計算のルール (足し算、スカラー倍)

ベクトルには以下のような計算法則がある。
これらは、成分を計算すると簡単に示すことができる。

足し算関連

$$\overrightarrow{x} + \overrightarrow{y} = \overrightarrow{y} + \overrightarrow{x}$$

$$(\overrightarrow{x} + \overrightarrow{y}) + \vec{z} = \overrightarrow{x} + (\overrightarrow{y} + \vec{z})$$

スカラー倍関連

$$(a + b)\overrightarrow{x} = a\overrightarrow{x} + b\overrightarrow{x}$$

$$ab\overrightarrow{x} = a(b\overrightarrow{x})$$

$$a(\overrightarrow{x} + \overrightarrow{y}) = a\overrightarrow{x} + a\overrightarrow{y}$$

# ベクトルの内積: 同じ向きを取り出しかけ算する

2つのベクトル $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\overrightarrow{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ とそれらのなす角 $\theta$ に対し、

<u>スカラーの値</u> $xy\cos\theta$ を内積と呼び、$\overrightarrow{x} \cdot \overrightarrow{y}$ で表す。
(ただし、$\theta$ は0°以上180°以下の値とする)

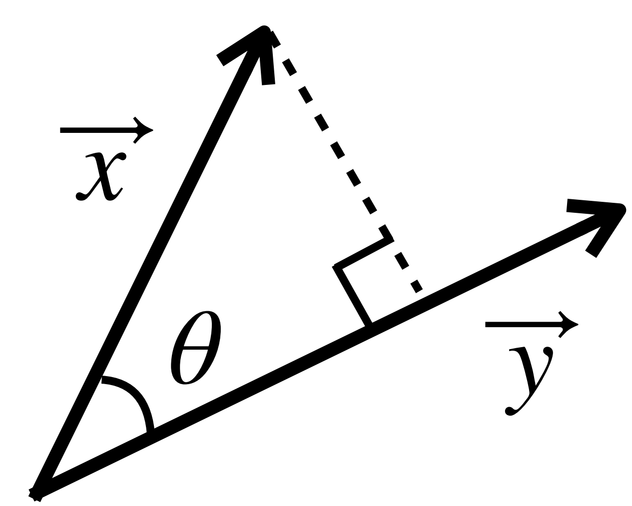
$\overrightarrow{x}$, $\overrightarrow{y}$ は入れ替えても計算結果は同じになるため、

$$\overrightarrow{x} \cdot \overrightarrow{y} = \overrightarrow{y} \cdot \overrightarrow{x}$$

が成り立つ。

$\cos\theta$ が入っているため、2つのベクトルが直交すれば
ゼロとなり、平行であれば大きさの積となる。

※ 内積は必ず「・」で表す。
通常のかけ算のような省略は不可、
×での代用不可。



# ベクトルの内積の計算法則 (ふつうのかけ算と同じ)

ベクトル$\overrightarrow{x}$, $\overrightarrow{y}$, $\vec{z}$ に対し、

$$\overrightarrow{x}\cdot(\overrightarrow{y}+\vec{z}) = \overrightarrow{x}\cdot\overrightarrow{y}+\overrightarrow{x}\cdot\vec{z}$$

が成り立つ。その理由は右下図の通り、

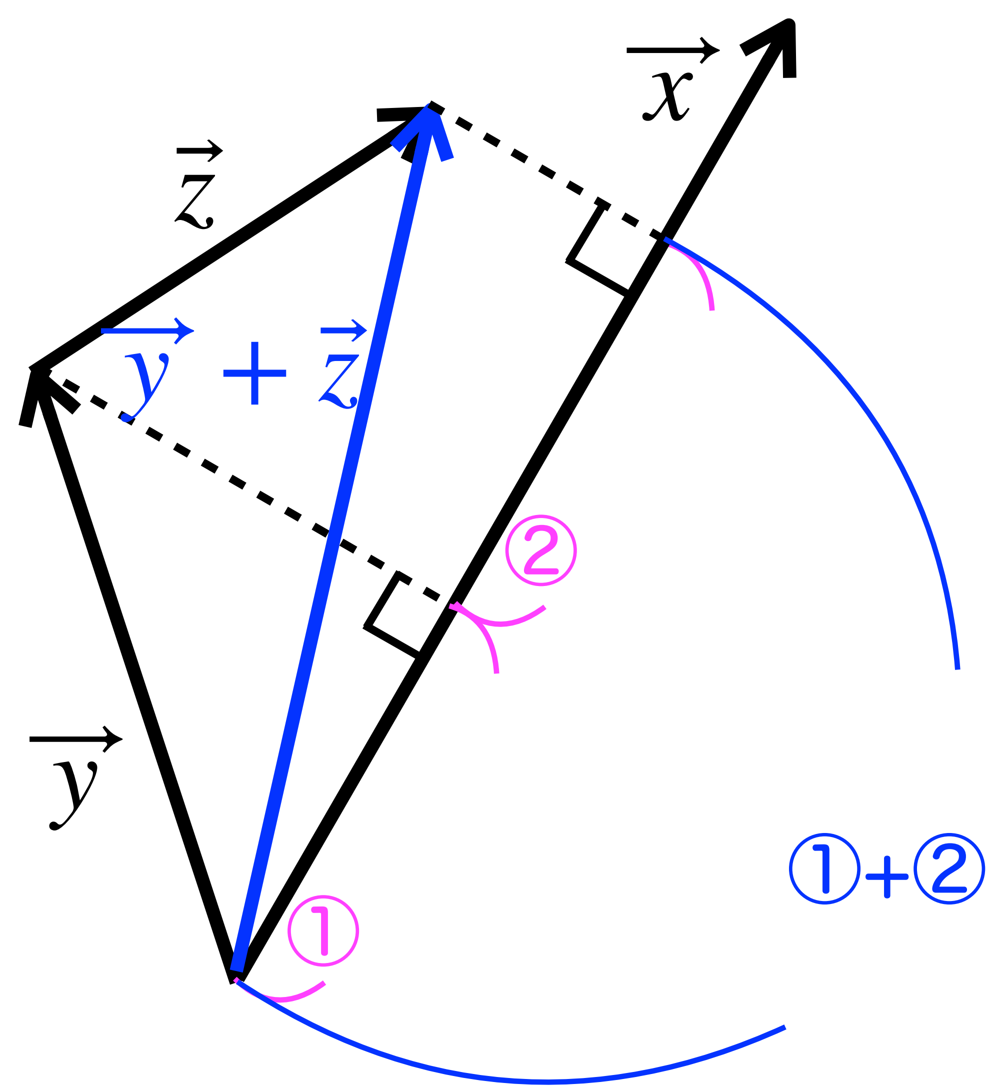
$x \times$ (①+②) $=$ $x \times$ ①$+$$x \times$ ②

だから。

同様に、

$$(\overrightarrow{y}+\vec{z})\cdot\overrightarrow{x} = \overrightarrow{y}\cdot\overrightarrow{x}+\vec{z}\cdot\overrightarrow{x}$$

も成り立つ。



# ベクトルの成分による内積の計算

$\overrightarrow{y}$ を成分ごとに分解し、以下のように書いてみる。

$$\boxed{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cdot \boxed{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}} = \boxed{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cdot \boxed{\left[\begin{pmatrix} y_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y_2 \end{pmatrix}\right]}$$

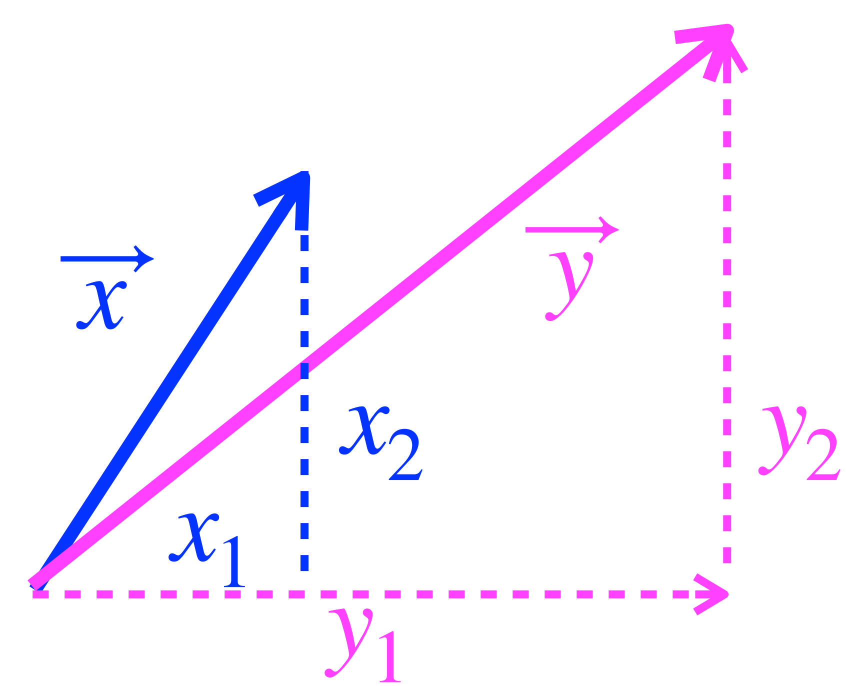
先ほどの法則 $\overrightarrow{x} \cdot (\overrightarrow{y} + \overrightarrow{z}) = \overrightarrow{x} \cdot \overrightarrow{y} + \overrightarrow{x} \cdot \overrightarrow{z}$ より、

$$\boxed{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cdot \boxed{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}} = \boxed{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cdot \boxed{\begin{pmatrix} y_1 \\ 0 \end{pmatrix}} + \boxed{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cdot \boxed{\begin{pmatrix} 0 \\ y_2 \end{pmatrix}}$$

右図より、以下のように書ける。

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 + x_2 y_2$$

$$= \sum_{i=1}^{2} x_i y_i$$



# データはベクトルで表せる

前回紹介した通り、データ分析ではデータをベクトルと捉える。
しかしデータは単なる数値の羅列であってベクトルではないのでは？
→ データに関するさまざまな計算でベクトルの計算法則を使うと便利。
直感的にわかりやすい平面 (や空間) のベクトルのイメージから拡張し
ベクトルの概念を一般化すればデータもベクトルと捉えられる

なお、「1つの対象に対する複数の変数」をベクトルと考える場合、
「複数の対象に対する1種類の変数」をベクトルと考える場合、両方ある

例:
右表のようなデータ
があったとき…

| | りんご (個) | みかん (個) |
|---|---|---|
| バスケットA | 3 | 5 |
| バスケットB | 4 | 2 |

バスケットごとのベクトル $\begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

果物ごとのベクトル $\begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \end{pmatrix}$

# ベクトルの計算はそのまま高次元に拡張できる

ここまでは、平面上のベクトルについて扱ってきた。
平面のベクトルは2つの成分で表すことができた。
ベクトルの成分数のことを次元と呼ぶ。平面のベクトルは2次元。

これをそのまま空間 (3次元) に拡張することができる。

- 各ベクトルは3つの成分で表すことができる
- 足し算、引き算、スカラー倍、内積の法則はそのまま
- ベクトルの大きさおよび内積を成分で表した式は、Σの上の数が2 → 3

さらに、n 次元空間にまで拡張することができる (もはやイメージは不可能)

- 各ベクトルは n 個の成分で表すことができる
- 足し算、引き算、スカラー倍、内積の法則はそのまま
- ベクトルの大きさおよび内積を成分で表した式は、Σの上の数が2 → n

ただし、足し算、引き算、内積計算は同じ次元のベクトル同士でしか
できないことに注意。

# 演習

3店のラーメン店について、ラーメンの価格、ギョウザの価格、1日の来客人数を調査したところ、左下の表のようになった。ラーメンの価格、ギョウザの価格、来客数の平均は700円、200円、85人であった。
各項目から平均を引いた値 (偏差) をまとめると右下表のようになる。
ここで、各項目をベクトルとして以下のように表す。

$$\vec{r} = \begin{pmatrix} 0 \\ -100 \\ 100 \end{pmatrix}, \quad \vec{g} = \begin{pmatrix} -50 \\ -100 \\ 150 \end{pmatrix}, \quad \vec{n} = \begin{pmatrix} -5 \\ 15 \\ -10 \end{pmatrix}$$

このとき、内積 $\vec{r} \cdot \vec{g}$, $\vec{g} \cdot \vec{n}$, $\vec{n} \cdot \vec{r}$ をそれぞれ求めよ。

調査結果

| | ラーメンの価格 (円) | ギョウザの価格 (円) | 1日の来客 (人) |
|---|---|---|---|
| A店 | 700 | 150 | 80 |
| B店 | 600 | 100 | 100 |
| C店 | 800 | 350 | 75 |

偏差

| | ラーメンの価格 (円) | ギョウザの価格 (円) | 1日の来客 (人) |
|---|---|---|---|
| A店 | 0 | -50 | -5 |
| B店 | -100 | -100 | 15 |
| C店 | 100 | 150 | -10 |

# 演習解答

Python によるベクトル計算紹介とともに別ファイルで示す。

# 内積とコサイン類似度

この演習問題にいったいどんな意味があったのか？
そもそも内積とは結局何なのか？

内積という量自体を明確なイメージで捉えるのは難しい。
ただ、$\overrightarrow{x} \cdot \overrightarrow{y} = xy\cos\theta$ なので、[大きさ]×[大きさ]×[向きの関係]
と捉えることができる (これは何次元でも同じ)

向きの関係 $\cos\theta$ は、2つのベクトルの向きが同じとき1 (最大)、
直交するとき0、向きが180°逆のとき-1 (最小) となる。
これは「<u>向きがどのぐらい近いか</u>」を表す量と捉えられ、
「コサイン類似度」とも呼ばれる。

多次元になるほど角度 $\theta$ 自体は計算が難しくなるが、コサイン類似度は
内積から以下の通り比較的簡単に求められる。

$$\cos\theta = \frac{\overrightarrow{x} \cdot \overrightarrow{y}}{xy}$$

# (参考) ベクトル演算と統計量の間には関係がある

データ分析においては、演習問題のように変数の種類ごとの偏差を
ベクトルと考えると (ラーメンベクトル、餃子ベクトル、来客ベクトル)、
ベクトルの演算と統計量が以下のような関係を持つ。

[ベクトルの大きさ]の二乗/N：分散
ベクトルの内積/N：共分散
コサイン類似度: 相関係数
(N はデータ数、つまりベクトルの次元数とする)

※ 統計量については第8回で改めて説明
(データサイエンス概論も参照)

# ベクトルの線型結合 ～ベクトルを組み合わせて新たなベクトルを作る

複数のベクトルがあったとき、あるベクトルが他のベクトルを計算して(足し算、引き算、スカラー倍で) 得られる場合には、それらは線形従属であるという。

そうではない場合、それらのベクトルは線型独立であるという。
(※方向という概念がなくなる大きさゼロのベクトルは除外)

線型独立なベクトルは互いに並行でない。
(並行であればスカラー倍で表せてしまう)

$\overrightarrow{x}, \overrightarrow{y}$ という2つのベクトルとあるスカラー a, bに対して
$a\overrightarrow{x} + b\overrightarrow{y}$ という形の式を線形結合と呼ぶ。
線形独立とは、ゼロベクトルを除き、互いに線形結合で表すことのできないベクトルの関係。

# 任意のベクトルの線型結合による表現

平面のベクトルの場合、2つの線型独立なベクトルの線型結合で
平面上のあらゆるベクトルを表現することができる。

ベクトル $\vec{r} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$ を $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\overrightarrow{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ で表現するには、

$$\vec{r} = a\overrightarrow{x} + b\overrightarrow{y} = a\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + b\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

の式を a, b について解いて、

$$a = \frac{r_1 y_2 - r_2 y_1}{x_1 y_2 - x_2 y_1}, \quad b = \frac{r_1 x_2 - r_2 x_1}{x_2 y_1 - x_1 y_2}$$

と書ける。分母が0でない限り、このように a, b を設定すれば
任意のベクトルを2つのベクトルの線型結合で表現できることになる。

$\overrightarrow{x}, \overrightarrow{y}$ が線型独立であれば分母はゼロにはならない。

# 線形結合の例

先ほどの例を使って、線形結合の例を考えてみよう。

| | りんご (個) | みかん (個) |
|---|---|---|
| バスケットA | 3 | 5 |
| バスケットB | 4 | 2 |

バスケットAベクトル $\overrightarrow{b_A} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ バスケットBベクトル $\overrightarrow{b_B} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

りんご10個、みかん12個欲しかった場合、バスケットAを2つ、バスケットBを1つ買えばよい。ベクトルでは以下のような線形結合で表せる。係数の2と1は前ページの式でも求められる。

$$\begin{pmatrix} 10 \\ 12 \end{pmatrix} = 2\overrightarrow{b_A} + \overrightarrow{b_B}$$

この例では係数 a, b は正の整数だが、線形結合一般には小数を含む数や負の値となることもある (整数でなくとも問題ない)

# 単位ベクトルの線型結合

以下のような、ある成分のみ1で、他の成分は0であるようなベクトルを単位ベクトルと呼ぶ。

$$\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \ \overrightarrow{e_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

この2つのベクトルは、一方をスカラー倍しても他方にはできないので、線型独立である。
単位ベクトルは互いに内積がゼロ、つまり垂直である。

線型結合 $a\overrightarrow{e_1} + b\overrightarrow{e_2}$ で任意のベクトルを表現することができることは簡単に示すことができる。

# 線形結合の高次元のベクトルへの拡張

線形結合の考え方についても、平面のベクトルをそのまま空間 (3次元) に拡張できる。

- 空間上のあらゆるベクトルを線型独立な3つのベクトルで表せる
- 単位ベクトルの数は3つ

n 次元空間にも拡張できる。

- 空間上のあらゆるベクトルを線型独立な n 個のベクトルで表せる
- 単位ベクトルの数は n 個

# 本日のまとめ

- ベクトルは大きさと向きを持つ量で、成分を用いて表すことができる
- ベクトルの内積は大きさ同士のかけ算と、向きの関係を反映した量で、成分からも算出できる
- 内積のうち向きの関係のみを取り出したコサイン類似度はベクトル同士の向きの近さを表す量
- n 個の要素を持つデータは n 次元空間上のベクトルとして表現することができる