山下研の常識

線形代数の巻

山下幸彦
東京工業大学・大学院理工学研究科

概要

- 「山下研の常識」は，山下研でやっていく上での基本的な常識をまとめたものです．線形代数，パターン認識，画像符号化，計算機アーキテクチャ，などについて，基本的な部分を独断と偏見と愚痴と長嶋監督的に擬態音(「ヒュー」とか「ビュン」とか)を混ぜながら述べていく予定です．馬鹿馬鹿しいと思っても，あまり怒らないように．
- 「線形代数の巻」では，主に線形代数やその応用に関して述べていく予定です．

1　はじめに

1.1　線形代数はお嫌いですか？

- 線形空間です．なぜか嫌いな学生が多いです．
- 社会人にも少なからずいらしゃいますので，心配は無用です．
  - 特に無限次元の線形空間は嫌いな方がいらっしゃいます．
  - 論文をヒルベルト空間を使って書くと,「(有限次元の) ユークリッド空間で書け」という方がいらっしゃいます．
  - (有限次元の) ユークリッド空間はヒルベルト空間の特別な場合です．従って，ヒルベルト空間で書けるならば，それで書いておいた方が問題の性質を正確に切り出していることになります．「『分数でなく小数で書け』と言っているのと同じだろ」とか「学会は学問をするところだろう．そうでないなら学会なんて名前を降ろして，お役立ちグッズ発表会とかにしたら」と言いたくなります．
  - だいたい，カーネル法では関数空間へ写像したりするので，無限次元なんだからそれを有限ベクトル空間で書けとは変です．不要な条件を付加して適用範囲を狭めることはないといと思います．

(話題閑休)

1.2　空間を頭に描けるようになろう

- 普通の人間には3次元空間までしか頭に描くとことができません.
- 従って，頭に描くときにユークリッド空間に落として考えることは必要です.
- ３次元までのユークリッド空間については，頭で図を描くことができるようにして下さい.
- このように頭に描けるということが，**線形代数の最大の利点**なのですから.
- 無限次元の線形空間は，それに収束に関する議論が加わるだけです.
- そこで，「ヒュー」とか「ピュー」とか言う感覚が必要になります.
  点列が「ヒュー」と近付いていってくれるか,「ビュン」と飛んでいってしまうか,「グッパ，グッパ」と近付きそうに見えて近付かないかです.

1.3 線形代数とは

- 線形代数は，線形空間の点であるベクトルの演算と，その性質をまとめた学問です.
- 線形代数の応用としては，ベクトルを，いま我々が存在している物理的な空間の点や電界とするもの，画像や信号を変換した数の並びとするもの，関数とするものなどがあります.
- 何で関数を**わざわざ**ベクトルなんかで表すんだと思うかもしれません.
  - ベクトルで関数の全ての情報を表すことはできません.
  - 例えば，空間がある性質(再生核ヒルベルト空間など)を持っていない限り，ある一点での関数の値を論じることはできません.
  - 関数をベクトルとして扱うことは，関数の情報の一部を切り出したもの，または，ある角度から見たものがベクトルであるということです.
  - 例えば，2つの関数の近さを，関数値の差の絶対値の2乗を積分したもので評価するものとします.
  - このとき，関数を近似することを論じるなどということが，ベクト

ル間の距離で論じることができるようになり，そのまま扱うよりも非常にわかりやすくなります．

– わかりやすくなるというのは，記述が短くなる他に，空間を頭に描くことによって近似の様子がビジュアル的に描けるようになります．
– **関数の線形空間としての性質だけが必要なときに，それ以上の複雑な所を見ない方が良いわけです．**

(人間関係も相手のイヤなところは見ない方が精神的に安定します．それが，ダチョウの安全保障になっても困るわけですが).

- このように「粗視化」して議論を容易にすることによって，理論や応用の発展が容易になるわけです．
- 数学は，**「ある性質が成立すれば別の性質が成立する」**というものの集まりです．即ち，ある粗視化したものと別の粗視化したものの間に成立する厳密な関係を論じているのです．
- 例えば,「連続関数の一様収束関数は連続関数である」という定理を考えるとき，実際には,「連続関数」というよりも，もう少し具体的な関数の列を考えています．ただ，その「連続」という性質だけが必要な場合は，このような具体的な関数をその連続性だけを見て粗視化した

扱っていきます．もちろん，連続関数といっても，ほとんど不連続のような関数も作ることができますので，そのような粗視化が十分か不充分か考える必要があります．

- ある結果を導き出すためには，どう見るかが重要になります．線形代数はそのものの見方の重要な一つです．
- 確かに現在はコンピュータの計算速度が上がっいるため，コンピュータ万能で，問題を線形代数を使って解析的に考えるよりも，プログラムを買ってきて，具体的に値を代入して計算させた方が容易に解けてしまうということがあります．
- 線形代数など知らなくてもほとんどのことができるでしょう．
- しかし，
  - そのプログラムにはやはり線形代数的な考えかたが使われている場合が多いです．
  - 物理的なものは無限次元の関数として扱いますから，それが有限で次元で扱って良い近似になるかとか，どのようなものが誤差で残るかなどを議論するときに必要になります．
  - また，パターン認識におけるカーネル法は，最終的には有限次元に

落としてきますが，その過程の議論には無限次元のヒルベルト空間が必要です．

- また，線形空間は高度な数学のすべての基礎となっています．多様体にしても，もっと高度な代数にしても，少なくとも有限次元の線形空間が基礎となっています．
- 例えば，定規とコンパスでは角の3等分が作図できないことを知っていると思いますが，もとの座標値がなす体に対して，作図された座標値で拡張された体を，元の体上の線形空間と見なすことができます．コンパスと定規だけでは，この次元が2のベキ乗にしかならないことが証明されますが，角が三等分できるとするとその次元の因数に3を含むことが証明されます．従って，角が三等分が作図できないこととが証明されます．
- これも，ある性質を次元というところまで粗く見る(空間の次元だけ)ことによって，証明しているわけです．

• せっかく山下研に入ったのですから，不本意ながら山下研に入らざるを得なかったとしても，線形空間に関して勉強すると良いと思います．

1.4　ベクトル

線形代数の点または要素となるものがベクトルです.

　高校までで，2次元のベクトル

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{1}$$

みたいなものは扱ったことがあると思います.

　このベクトルには，ベクトルとベクトルの和

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \tag{2}$$

とベクトルと数(一般には実数か複素数，以後，ベクトルでなく単なる数の方をスカラーと呼ぶことにします)のスカラー積

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

が定義されていました.

一般にはベクトルとベクトルの積は定義されていなかったはずです.

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \end{pmatrix} = ? \tag{4}$$

2次元ベクトルは，一般的に，

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{5}$$

と書くことができます．こう書くとx_1, x_2は単なる数です．これは，スカラー積のときの数と同じ集合に属します．x_1やx_2は，ベクトルxの成分または要素と呼ばれます．また，αをスカラーとすれば，

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \tag{6}$$

$$\alpha \boldsymbol{x} = \alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix} \tag{7}$$

となります．「何を当り前のことを，ぐだぐだ書くなと.」と思うかもしれませんが，これから，これを少しずつ一般化していきます.

2次元が分かれば，N次元ベクトルというものも考えられるでしょう．

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \tag{8}$$

と，成分がN個並ぶことになります．和とスカラー積も，

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{pmatrix} \tag{9}$$

$$\alpha x = \alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_N \end{pmatrix} \tag{10}$$

として定義できます．

さて，これらの演算にはどのような性質が成り立つでしょうか．その性質をまとめてみましょう．ここでは，α, βをスカラーとします．

1.和の交換則：「任意のN次元ベクトルx, yに対して，

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

が成立する.」証明は，ベクトルの要素が数で交換則が成り立つので簡単です.

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{pmatrix} = \begin{pmatrix} y_1 + x_1 \\ y_2 + x_2 \\ \vdots \\ y_N + x_N \end{pmatrix} = \boldsymbol{y} + \boldsymbol{x}$$

2.和の結合則：「任意のN次元ベクトル$\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$に対して，

$$(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$$

が成立する.」この証明も，ベクトルの要素が数で交換則が成り立つので簡単でしょう．是非，トライしてみて下さい.

3.和の単位元$\boldsymbol{\theta}$(N次元ベクトル)の存在：「任意のベクトルN次元ベクトル$\boldsymbol{x}$に対して,

$$\boldsymbol{x}+\boldsymbol{\theta}=\boldsymbol{x}$$

が成立する.」$\boldsymbol{\theta}$を,

$$\boldsymbol{\theta}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$$

とすれば条件を満たすことは明らかです.

4.和のスカラー積に対する分配則:「任意のN次元ベクトル$\boldsymbol{x}$, $\boldsymbol{y}$と任意のスカラαに対して,

$$\alpha(\boldsymbol{x}+\boldsymbol{y})=\alpha\boldsymbol{x}+\alpha\boldsymbol{y}$$

が成立する.」これは,「$\boldsymbol{x}+\boldsymbol{y}$という和の結果(ベクトル)に対する$\alpha$のスカラ積の結果(ベクトル)は，ベクトル$\boldsymbol{x}$に対する$\alpha$のスカラ積の結果(ベクトル)と，ベクトル$\boldsymbol{y}$に対する$\alpha$のスカラ積の結果(ベクトル)の和の結果(ベクトル)に等しい」ということです．証明は明らかでしょう.

また,「任意のN次元ベクトル$\boldsymbol{x}$と任意のスカラα, βに対して,

$$(\alpha+\beta)\boldsymbol{x} = (\alpha+\beta)\boldsymbol{x}$$

が成立する.」これは,「ベクトル$\boldsymbol{x}$に対する$\alpha+\beta$というスカラーの和のスカラ積の結果(ベクトル)は,ベクトル$\boldsymbol{x}$に対するαのスカラ積の結果(ベクトル)と,ベクトル$\boldsymbol{y}$に対するβのスカラ積の結果(ベクトル)の和の結果(ベクトル)に等しい」ということです.この証明も明らかでしょう.

5. スカラー積の結合配則:任意のN次元ベクトル$\boldsymbol{x}$と任意のスカラ$\alpha\beta$に対して,

$$(\alpha\beta)\boldsymbol{x} = \alpha(\beta\boldsymbol{x})$$

が成立する.これは,「数としてのαとβの積(数)のベクトル$\boldsymbol{x}$に対するスカラ積の結果(ベクトル)は,βの$\boldsymbol{x}$に対するスカラー積の結果(ベクトル)に対する,αのスカラ積の結果(ベクトル)と等しい」ということです.これも,証明は明らかです.

6.その他にスカラー積の性質として，

$$
\begin{aligned}
0\boldsymbol{x} &= \boldsymbol{\theta} \\
1\boldsymbol{x} &= \boldsymbol{x}
\end{aligned}
$$

が成立します．これらの式は，それぞれ，数0と$\boldsymbol{x}$のスカラー積，数1と$\boldsymbol{x}$のスカラー積を意味しています．１番目，２番目の式は，

$$
0\boldsymbol{x} = 0\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} 0 \times x_1 \\ 0 \times x_2 \\ \vdots \\ 0 \times x_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \boldsymbol{\theta} \qquad 1\boldsymbol{x} = 1\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} 1 \times x_1 \\ 1 \times x_2 \\ \vdots \\ 1 \times x_N \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \boldsymbol{x}
$$

より明らかです．上式の×は，数としての積です．

これを，抽象化したものがベクトルです．$\boldsymbol{x}, \boldsymbol{y}$をベクトルとし，$\alpha$をスカラー(一般には実数か複素数)とします．もうここでは，ベクトル$\boldsymbol{x}$には$x_1, x_2, \ldots, x_N$のような成分から構成されているとは考えません．もっと**抽象的**に，集合の全ての元に和とスカラー積が定義されているものとします．

Kを体(実数や複素数など)とします．Kの元をスカラーと呼ぶことにします．また，$\mathcal{X}$をある集合とします．**このX$\mathcal{X}$の要素に対して和とスカラー積が定義され，上の性質1-6を満たすとき，$\mathcal{X}$をK上の線形空間あるいはベクトル空間と呼びます．**

和とスカラー積が定義されているとは,

$$\boldsymbol{x} + \boldsymbol{y} \tag{11}$$

$$\alpha \boldsymbol{x} \tag{12}$$

という演算が一意に定まり，$\boldsymbol{x} + \boldsymbol{y} \in \mathcal{X}$かつ$\alpha x \in \mathcal{X}$となることです．

すなわち，$\mathcal{X}$の全ての2組の元に対して和が求まり，$\mathcal{X}$の全ての元と全てのスカラーに対して，スカラー積が求まるということです．そして，線形空間であるためには，**この2種類の演算が，上の性質1〜6を満たさなくてはいけないのです．**

この1〜6の性質さえ成り立てば，その抽象的なベクトルxに対しても，N次元ベクトル空間のようにして，議論を進めて行くことができます．

1.4.1　例 : $C[0,1]$

例えば，$\mathcal{X} = C[0,1]$ とします．$C[0,1]$ は区間 $[0,1]$ で定義された連続関数の全体です．$\boldsymbol{x} \in C[0,1]$ とすれば，任意の t $(0 \le t \le 1)$ に対して $\boldsymbol{x}(t)$ は実数値を持ち，$\boldsymbol{x}(t)$ は関数として連続です．

まず，ベクトルの和を定義します．$\boldsymbol{x}, \boldsymbol{y} \in C[0,1]$ に対して，$\boldsymbol{x}(t) + \boldsymbol{y}(t)$ を考えると，これも区間 $[0,1]$ で定義された連続関数となります．従って，この関数は $C[0,1]$ に含まれ，この関数をベクトルとして和 $\boldsymbol{x}+\boldsymbol{y}$ として定義します．すなわち，任意の t $(0 \le t \le 1)$ に対して，

$$(\boldsymbol{x} + \boldsymbol{y})(t) = \boldsymbol{x}(t) + \boldsymbol{y}(t) \tag{13}$$

です．

(この式の意味が分かりますでしょうか．右辺の $\boldsymbol{x}+\boldsymbol{y}$ は，$C[0,1]$ の元で連続関数ですので，t $(0 \le t \le 1)$ に対して値を持たなくてはいけません．それが，任意の t に対して，$\boldsymbol{x}(t) + \boldsymbol{y}(t)$ に等しいという意味です．)

次に，スカラー積を定義します．$\boldsymbol{x} \in C[0,1]$ に対して，$\alpha\boldsymbol{x}(t)$ を考えると，これも区間 $[0,1]$ で定義された連続関数となります．この関数を $\alpha\boldsymbol{x}$ として定義します．すなわち，任意の t $(0 \leq t \leq 1)$ に対して，

$$(\alpha\boldsymbol{x})(t) = \alpha\boldsymbol{x}(t) \tag{14}$$

です.

(右辺の $\alpha\boldsymbol{x}$ は，$C[0,1]$ の元で連続関数ですので，t $(0 \leq t \leq 1)$ に対して値を持たなくてはいけません．それが $\alpha\boldsymbol{x}(t)$ に等しいという意味です．)

また, 2つのベクトルが等しいということも定義が必要です．$\boldsymbol{x}, \boldsymbol{y} \in C[0,1]$ が等しいとは，任意の t $(0 \leq t \leq 1)$ に対して,

$$\boldsymbol{x}(t) = \boldsymbol{y}(t) \tag{15}$$

となることです．すなわち，関数としてまったく等しいということです.

注：「等しい」という定義は，ベクトル空間を考える前の集合の時点で定義されています.

これが性質1から6を満たし，線型空間となることを見ていきましょう.

1.

$$(\boldsymbol{x}+\boldsymbol{y})(t)=\boldsymbol{x}(t)+\boldsymbol{y}(t)=\boldsymbol{y}(t)+\boldsymbol{x}(t)=(\boldsymbol{y}+\boldsymbol{x})(t)$$

となりますので，$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$となります．これは，関数値における交換則がそのまま反映した形になっています.

2.

$$\begin{aligned}((\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z})(t) &= (\boldsymbol{x}+\boldsymbol{y})(t)+\boldsymbol{z}(t)=\boldsymbol{x}(t)+\boldsymbol{y}(t)+\boldsymbol{z}(t)\\ &= \boldsymbol{x}(t)+(\boldsymbol{y}+\boldsymbol{z})(t)=(\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z}))(t)\end{aligned}$$

となりますので，$(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$となります．これも，関数値における結合則がそのまま反映した形になっています.

3.$\boldsymbol{\theta}\in C$は，任意のt $(0\leq t\leq 1)$に対して，いつでも関数値が0となる連続関数$\boldsymbol{\theta}$ $(\boldsymbol{\theta}(t)=0)$を導入すれば,

$$(\boldsymbol{x}+\boldsymbol{\theta})(t)=\boldsymbol{x}(t)+\boldsymbol{\theta}(t)=\boldsymbol{x}(t)+0=\boldsymbol{x}(t)$$

となりますので，$\boldsymbol{x}+\boldsymbol{\theta}=\boldsymbol{x}$となります.

4.

$$
\begin{aligned}
(\alpha(\boldsymbol{x}+\boldsymbol{y}))(t) &= \alpha((\boldsymbol{x}+\boldsymbol{y})(t)) = \alpha(\boldsymbol{x}(t)+\boldsymbol{y}(t)) = \alpha\boldsymbol{x}(t)+\alpha\boldsymbol{y}(t) \\
&= (\alpha\boldsymbol{x})(t)+(\alpha\boldsymbol{y})(t) = (\alpha\boldsymbol{x}+\alpha\boldsymbol{y})(t) \\
((\alpha+\beta)\boldsymbol{x})(t) &= \alpha\boldsymbol{x}(t)+\beta\boldsymbol{x}(t) = \alpha\boldsymbol{x}(t)+\beta\boldsymbol{x}(t) = (\alpha\boldsymbol{x})(t)+(\beta\boldsymbol{x})(t) \\
&= (\alpha\boldsymbol{x}+\beta\boldsymbol{x})(t)
\end{aligned}
$$

となりますので，$\alpha(\boldsymbol{x}+\boldsymbol{y}) = \alpha\boldsymbol{x}+\alpha\boldsymbol{y}$ と $(\alpha\beta)\boldsymbol{x}) = \alpha\boldsymbol{x}+\beta\boldsymbol{x}$ が成立します.

5.

$$((\alpha\beta)\boldsymbol{x})(t) = (\alpha\beta)(\boldsymbol{x}(t)) = \alpha(\beta\boldsymbol{x}(t)) = \alpha((\beta\boldsymbol{x})(t)) = (\alpha(\beta\boldsymbol{x}))(t)$$

となりますので，$(\alpha\beta)\boldsymbol{x} = \alpha(\beta\boldsymbol{x})$ となります.

6.

$$(0\boldsymbol{x})(t) = 0(\boldsymbol{x}(t)) = 0$$

と，

$$(1\boldsymbol{x})(t) = 1(\boldsymbol{x}(t)) = \boldsymbol{x}(t)$$

より明らかです.

問題 1. 定義に従って，次の関係を示しなさい．

1. $-\boldsymbol{x}$を$(-1)\boldsymbol{x}$で定義すれば，和に対する$\boldsymbol{x}$の逆元になる．即ち，$\boldsymbol{x}+(-\boldsymbol{x})=\boldsymbol{\theta}$が成立する．
   $\boldsymbol{a}+(-\boldsymbol{b})$を$\boldsymbol{a}-\boldsymbol{b}$と記すことにする．

2. $\boldsymbol{\theta}_1$と$\boldsymbol{\theta}_2$が，公理*3*を満たすならば，$\boldsymbol{\theta}_1=\boldsymbol{\theta}_2$となる．

3. $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{z}$ならば$\boldsymbol{y}=\boldsymbol{z}$になる．

4. $\alpha\boldsymbol{x}=\alpha\boldsymbol{y}$かつ$\alpha\neq 0$ならば$\boldsymbol{x}=\boldsymbol{y}$となる．

5. $\alpha\boldsymbol{x}=\beta\boldsymbol{x}$かつ$\boldsymbol{x}\neq\boldsymbol{\theta}$ならば$\alpha=\beta$となる．

6. $(\alpha-\beta)\boldsymbol{x}=\alpha\boldsymbol{x}-\beta\boldsymbol{x}$が成立する．

7. $\alpha(\boldsymbol{x}-\boldsymbol{y})=\alpha\boldsymbol{x}-\alpha\boldsymbol{y}$が成立する．

8. $\alpha\boldsymbol{\theta}=\boldsymbol{\theta}$が成立する．

注: 等号の公理は使って良い．
一般に$\boldsymbol{x}=\boldsymbol{y}$ならば，式の$\boldsymbol{x}$の部分を$\boldsymbol{y}$に置き換えても値は変らない．一般に$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{w}+\boldsymbol{z}$ならば，式の$\boldsymbol{x}+\boldsymbol{y}$の部分を$\boldsymbol{w}+\boldsymbol{z}$に置き換えても値は変らない．

1.5　部分空間

部分空間とは，3次元空間の中の**原点を通る**平面や直線みたいなものです．定義は次の通りです．線形空間Xの部分集合Sが部分空間であるとは，(Sが線形空間の定義を満たすことですが) 任意の$\boldsymbol{x}, \boldsymbol{y} \in S$と任意のスカラー$\alpha, \beta$に対して，

$$\alpha\boldsymbol{x} + \beta\boldsymbol{y} \tag{16}$$

がSに含まれるということです．これは，

$$\boldsymbol{x} + \boldsymbol{y}, \qquad \alpha\boldsymbol{x} \tag{17}$$

の2つのベクトルがSに含まれるという条件と等価です．これらの式は，和とスカラ積に関して計算が閉じている(Sのベクトルに関する演算の結果がSに含まれるということです) ことを意味しています．Sの元はXの元ですので，計算が閉じていれば，結合則や分配則は満たされることになります．また,「任意のスカラーα, βに対して」とありますので，$\alpha = \beta = 0$に対しても成立することになります．従って，

$$0\boldsymbol{x} + 0\boldsymbol{y} = \boldsymbol{\theta} + \boldsymbol{\theta} = \boldsymbol{\theta} \tag{18}$$

となりますので，$\boldsymbol{\theta} \in S$ となります．すなわち，線形部分空間は$\boldsymbol{\theta}$(原点)を含んでいなくてはいけません.

線形部分空間を平行移動したものを**線形多様体**と呼びます．Vを部分空間Sをある固定したベクトル$\boldsymbol{x}$で並行移動した線形多様体とすれば,

$$V = \{\boldsymbol{y} + \boldsymbol{x} \mid \boldsymbol{y} \in S\} \tag{19}$$

で定義されます.

1.6　線形独立と空間の次元

n個のベクトル$\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ (これはあるベクトルの要素でなくて，ベクトルがn個あるということです．太文字ですので，それぞれがベクトルになります) とn個のスカラの組$\alpha_1, \alpha_2, \dots, \alpha_n$に対して,

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots + \alpha_n \boldsymbol{x}_n \tag{20}$$

を，$\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$の1次結合または線形結合と呼びます．(1次結合でのベクトルの和は**有限個**です．)

　n個のベクトル$\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$が「**線形独立**」であるとは，$n$個のスカラの組$\alpha_1, \alpha_2, \dots, \alpha_n$に対して,

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots + \alpha_n \boldsymbol{x}_n = \boldsymbol{\theta} \tag{21}$$

となるならば，必ず,

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \tag{22}$$

となることである．

　逆に，少なくとも1つが0でない$\alpha_1, \alpha_2, \dots, \alpha_n$が存在して，式(21)が成立するとき,「**線形従属**」と呼ぶ.

1.7　基底と空間の次元

n 個のベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ が $\mathcal{X}$ の「**基底**」であるとは，それらが線形独立であり，任意の $\mathcal{X}$ のベクトルがその線形結合で表わせることです．

$\mathcal{X}$ に，その数が有限の基底ベクトルが存在するとき，$\mathcal{X}$ は n 次元であるといいます．逆に，有限の基底ベクトルが存在しないとき，$\mathcal{X}$ は無限次元であるといいます．

有限次元 (n 次元) 空間の基底ベクトルの数は，かならず一定 (n) です

問題 2. 上の文を証明して下さい．

n 次元 R^n の基底として，次の，

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ \ldots, \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

を取ることができます．これを，R^n の「**自然基底**」と呼びます．

これが基底になること(基底の公理を満たすこと)は明らかです.

問題 3. 上の文を証明して下さい.

部分空間は線形空間ですから，有限次元の線形空間の部分空間は次元を持ちます．N 次元の線形空間の部分空間の次元は，N 以下になります

問題 4. 上の文を証明して下さい.

線形多様体は部分空間を平行移動したものですから，その部分空間で次元を定義することができます.

超平面とは，N 次元の線形空間においては，$N-1$ 次元の線形多様体のことを指します(無限次元の場合は，商空間(詳しくは後ほど)の次元が 1 として定義されます).

1.8　線形作用素と行列

線形作用素は，行列の概念を一般の線形空間に拡張したものです．微分やフーリエ変換などの線形作用素です.

2つの線形空間$\mathcal{X}$と$\mathcal{Y}$を考えます．Aを$\mathcal{X}$から$\mathcal{Y}$の写像とします．すなわち，$\boldsymbol{x} \in \mathcal{X}$に対して，$\mathcal{Y}$の元$\boldsymbol{y}$が1つ定まります．$A$による写像は，一般的には,

$$A(\boldsymbol{x}) \tag{23}$$

と表します．Aが線形作用素であるとは，2つのベクトル$\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$と2つのスカラ$\alpha_1, \alpha_2$に対して,

$$A(\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2) = \alpha_1 A(\boldsymbol{x}_1) + \alpha_2 A(\boldsymbol{x}_2) \tag{24}$$

を満たすことです．これは,

$$A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A(\boldsymbol{x}_1) + (\boldsymbol{x}_2) \tag{25}$$

$$A(\alpha_1\boldsymbol{x}_1) = \alpha_1 A(\boldsymbol{x}_1) \tag{26}$$

という2つの条件をまとめて書いたものです．和やスカラー積を写像したものは，写像したものの和やスカラー積に等しい，すなわち，写像と和およぼスカラー積が可換であるというものです.

線形作用素の場合は，写像の結果$A(\boldsymbol{x})$を一般に$A\boldsymbol{x}$と表します．

$\mathcal{X}$および$\mathcal{Y}$をそれぞれ，mおよびn次元の空間とします．それぞれの基底を，$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$および$b \leq'_1, \boldsymbol{e}'_2, \ldots, \boldsymbol{e}'_n$で表します．任意の$j$ $(1 \leq j \leq m)$に対して，$A\boldsymbol{e}_j$は$\mathcal{Y}$のベクトルですので，$\boldsymbol{e}'_1, \boldsymbol{e}'_2, \ldots, \boldsymbol{e}'_m$の一次結合で表すことができます．これを，$a_{ij}$をスカラーとして，

$$A\boldsymbol{e}_j = a_{1j}\boldsymbol{e}'_1 + a_{2j}\boldsymbol{e}'_2 + \cdots + a_{nj}\boldsymbol{e}'_n \tag{27}$$

と書くことができます．$\mathcal{X}$の任意のベクトル$\boldsymbol{x}$は$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$の一次結合で表すことができます．従って，

$$\boldsymbol{x} = \xi_1\boldsymbol{e}_1 + \xi_2\boldsymbol{e}_2 + \cdots + \xi_m\boldsymbol{e}_m \tag{28}$$

となります．また，同様に，$\boldsymbol{y} = A\boldsymbol{x}$を，

$$\boldsymbol{y} = \eta_1\boldsymbol{e}'_1 + \eta_2\boldsymbol{e}'_2 + \cdots + \eta_m\boldsymbol{e}'_n \tag{29}$$

と表わすことができます．このとき，

$$
\begin{aligned}
A\boldsymbol{x} &= A(\xi_1\boldsymbol{e}_1 + \xi_2\boldsymbol{e}_2 + \cdots + \xi_m\boldsymbol{e}_m) \\
&= \xi_1 A\boldsymbol{e}_1 + \xi_2 A\boldsymbol{e}_2 + \cdots + \xi_m A\boldsymbol{e}_m \\
&= \xi_1(a_{11}\boldsymbol{e}'_1 + a_{21}\boldsymbol{e}'_2 + \cdots + a_{n1}\boldsymbol{e}'_n) \\
&\quad +\xi_2(a_{12}\boldsymbol{e}'_1 + a_{22}\boldsymbol{e}'_2 + \cdots + a_{n2}\boldsymbol{e}'_n) \\
&\quad +\cdots \\
&\quad +\xi_m(a_{1m}\boldsymbol{e}'_1 + a_{2m}\boldsymbol{e}'_2 + \cdots + a_{nm}\boldsymbol{e}'_n) \\
&= (a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_m)\boldsymbol{e}'_1 \\
&\quad +(a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_m)\boldsymbol{e}'_2 \\
&\quad +\cdots \\
&\quad +(a_{n1}\xi_1 + a_{n2}\xi_2 + \cdots + a_{nm}\xi_m)\boldsymbol{e}'_n \\
&= \sum_{i=1}^{n}\left(\sum_{j=1} a_{ij}\xi_j\right)\boldsymbol{e}'_i
\end{aligned}
\tag{30}
$$

となります．すなわち，行列

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \tag{31}$$

と，列ベクトル

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix}$$

に対して，行列に列ベクトルをかける演算

$$\boldsymbol{\eta} = \boldsymbol{A}\boldsymbol{\xi}$$

を行ない，

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

を計算しておけば,

$$\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} \eta_i \boldsymbol{e}'_i \tag{32}$$

が成立します．従って，有限次元の線形作用素は行列を使って表現することができることになります.

特に，空間が$\mathcal{X} = R^m$, $\mathcal{Y} = R^n$でe_i, e'_jが自然基底ならば,

$$\boldsymbol{x} = \boldsymbol{\xi} \tag{33}$$

$$\boldsymbol{y} = \boldsymbol{\eta} \tag{34}$$

となっています.

1.9　距離

線形空間の定義だけでは，なかなか有意義な話を展開できません．それはベクトル間にに近いとか遠いという議論ができないからです

まず，一般的な集合(線形空間とは限らない)の距離を定義します．

定義 1. 集合Sの距離とは，写像$d : S \times S \to R$で，任意の$a, b, c \in S$に対して，以下の条件を満たすものです．

1.$d(a,b) \geq 0$であり，
$d(a,b) = 0$ならば，$a = b$となる．

2.$d(a,b) = d(b,a)$

3.$d(a,b) \leq d(a,c) + d(c,b)$

3番目の性質が三角不等式と呼ばれるものです．a地点からb地点までまで行く距離は，c地点に寄っていくよりも，遠くなることはないというものです．何故だと考えて夜が眠れなくなるようなことはないようにして下さい(若い人にはわからなそう)．ただ，そのようなものは距離と呼ぶにはふさわしくないことは，なんとなくは理解できると思います．

性質1と2を満たす写像で三角不等式(性質3)が成り立たない写像が存在します．例えば，2次元の例で，

$$d(\boldsymbol{a},\boldsymbol{b}) = \left(\sqrt{|a_1 - b_1|} + \sqrt{|a_2 - b_2|}\right)^2$$

とします．性質1と2が成立することは明らかです．

$\boldsymbol{a} = (1,0)$，$\boldsymbol{b} = (0,1)$，$\boldsymbol{c} = (1,1)$とすると，

$$\begin{aligned} d(\boldsymbol{a},\boldsymbol{b}) &= \left(\sqrt{1} + \sqrt{1}\right)^2 = 4, \\ d(\boldsymbol{a},\boldsymbol{c}) &= \left(\sqrt{0} + \sqrt{1}\right)^2 = 1, \\ d(\boldsymbol{c},\boldsymbol{b}) &= \left(\sqrt{1} + \sqrt{0}\right)^2 = 1 \end{aligned}$$

となり，三角不等式が成立しないことがわかります．

1.10 ノルム

ノルムはベクトルの大きさを定めるものです．ノルムの定義は次の通りです.

定義 2. 体 K は任意の元 α に対してその絶対値 $|\alpha|$ が存在するものとします．体 K 上の線形空間 $\mathcal{X}$ のノルムとは，写像 $\|\cdot\| : \mathcal{X} \to K$ で，任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ と $\alpha \in K$ に対して，以下の性質を満たすものです.

1. $\|\boldsymbol{x}\| \geq 0$ であり，
$\|\boldsymbol{x}\| = 0$ ならば $\boldsymbol{x} = \boldsymbol{\theta}$ となる.

2. $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$.

3. $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

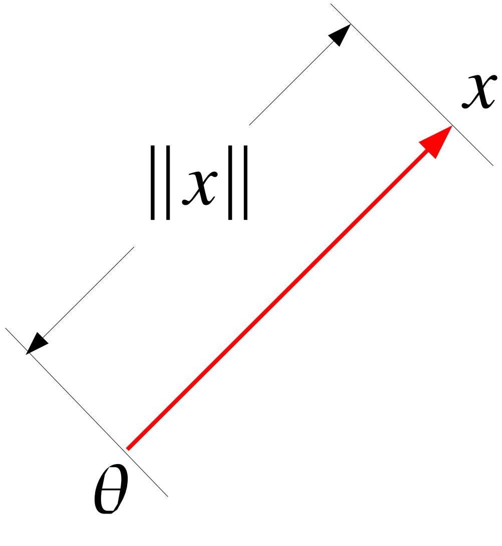
ノルムを使えば，ベクトルの差のノルムで，そのベクトル間の距離を定義することにができます.

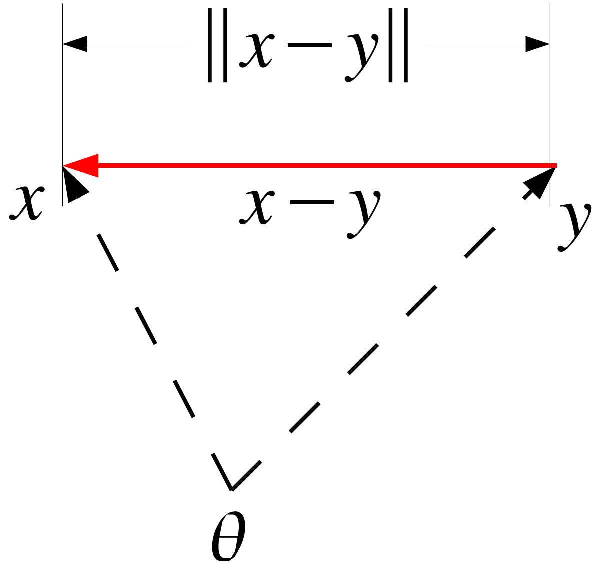
$$d(x, y) = \|x - y\|$$

ベクトルのノルムというときのベクトルは矢印をイメージしていて，ノルムはその長さになります．ベクトル間の距離というときのベクトルは,

ベクトルは矢印の先の点を表しています．その差によって決まるベクトルの大きさ(ノルム)が，その2点間の距離と言っているわけです．

問題 5. ノルムから定義された距離が，三角不等式を満たすことを証明して下さい．





1.11 内積

ノルムを定義できれば，大きさとか距離は定義できるのですが，ベクトル間の成す角のようなものが表せません．そのようなものを表するため，内積を定義します．R^2では，2つの元$\boldsymbol{x} = (x_1, x_2)$と$\boldsymbol{y} = (y_1, y_2)$に対して，内積$\langle \boldsymbol{x}, \boldsymbol{y} \rangle$は，

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + x_2 y_2$$

で定義することができます．もう少し一般に，R^Nでは，2つの元
$\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$と$\boldsymbol{y} = (y_1, y_2, \ldots, y_N)$に対して，

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{N} x_i y_i$$

で定義することができます．R^Nでこの内積を持っている空間をユークリッド空間E^Nと呼びます．

N次元の複素空間C^Nでは，複素数αの複素共役(虚数部をマイナスにす

る演算)を$\overline{\alpha}$で表せば,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{N} x_i \overline{y_i}$$

で定義されます．C^Nで，この内積を持っている空間を複素ユークリッド空間と呼ぶこともあります.

線形空間における内積の公理は以下の通りです.

定義 3. 体Kは任意の元αに対してその共役な要素$\overline{\alpha}$が存在するものとします．体K上の線形空間$\mathcal{X}$の内積とは，写像$\langle \cdot, \cdot \rangle \mathcal{X} \times \mathcal{X} \to K$で，任意の$\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$と$\alpha \in K$に対して，以下の性質を満たすものです.

1. $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$であり,
$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ならば$\boldsymbol{x} = \boldsymbol{\theta}$となる.

2. $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.

3. $\langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

4. $\langle y, x \rangle = \overline{\langle x, y \rangle}$.

線形空間の公理と内積の公理から，次の関係が成立します.

- $\langle \boldsymbol{\theta}, \boldsymbol{y} \rangle = 0$.
- $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$.
- $\langle \boldsymbol{x}, \alpha \boldsymbol{y} \rangle = \overline{\alpha} \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.
- $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} + \boldsymbol{w} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{x}, \boldsymbol{w} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{w} \rangle$
- $\sqrt{\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle} \leq \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} + \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle}$.

問題 6. 内積の公理から上の関係を証明してみて下さい．*(*線形空間の公理から導かれる命題は使ってよいです．*)*

この公理を満たせば内積なので，R^2 では，2つの元 $\boldsymbol{x} = (x_1, x_2)$ と $\boldsymbol{y} = (y_1, y_2)$ に対して，

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + 2 x_2 y_2$$

や

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + x_1 y_2 + x_2 y_2 + 2 x_2 y_2$$

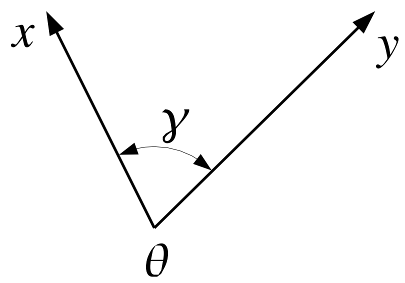
で定義することもできます．もちろん，これらの内積を持つ空間はユークリッド空間とは呼びません.

問題 7. 上の2つの内積が本当に内積の公理を満たすことを証明してみて下さい.

内積を使って,
$$\|\boldsymbol{x}\| = \sqrt{\langle x, x \rangle}$$
と定義すれば，これはノルムの公理を満たします.

問題 8. 上で定義したノルムが，本当にノルムの公理を満たすことを証明してみて下さい.

内積が定義できれば実数空間ではベクトル$\boldsymbol{x}$と$\boldsymbol{y}$がなす角の角度γが,
$$\cos\gamma = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|x\|\|y\|}$$
で定義できます(この場合はベクトルを，2つの矢印として考えているイメージ).



1.11.1 l^2

無限次元の空間の代表例として，l^2 に関して述べます．l^2 はユークリッド空間 E^N の無限次元版という感じです．l^2 の各要素は，複素数の無限列で絶体2乗和が有界になるものです．

即ち，$\boldsymbol{x} \in l^2$ は，$\boldsymbol{x} = \{x_i\}_{i=1}^{\infty} = (x_1, x_2, \ldots, x_n, \ldots)$ で，

$$\sum_{i=1}^{\infty} |x_i|^2 < \infty$$

となるものです．この $< \infty$ で有界であることを表してます．有界の定義は，ある実数 α が存在して，任意の M に対して，

$$\sum_{i=1}^{M} |x_i|^2 < \alpha$$

が成立することです．つまり，ある天井 α が存在して，和をとる範囲をいくら増やしていってもそれを越えないということです．

ちなみに，論理学の話ですが，α と M を逆にすると意味がありません．

任意のMに対して，ある実数αが存在して，

$$\sum_{i=1}^{M} |x_i|^2 < \alpha$$

はいつでも成立します．1つの要素は複素数ですので，それが無限ということはありません．数の集合には無限大は入っていません．ですので，$\alpha = \sum_{i=1}^{M} |x_i|^2 + 1$とすれば良いです．従って，いつでも成立してしまい意味はありません．

l^2は線形空間としては，次の演算で閉じる(その空間の元の演算の結果がその空間に必ず入る)ことになります．l^2の元$\boldsymbol{x} = \{x_i\}_{i=1}^{\infty} = (x_1, x_2, x_3, \ldots)$と，$\boldsymbol{y} = \{y_i\}_{i=1}^{\infty} = (y_1, y_2, y_3, \ldots)$と，スカラー$\alpha$に対して，和とスカラー積を，

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= (x_1 + y_1,\ x_2 + y_2,\ x_3 + y_3, \ldots) \\ \alpha \boldsymbol{x} &= (\alpha x_1, \alpha x_2, \alpha x_3, \ldots) \end{aligned}$$

と定義します(ほとんど，自明ですが)．この演算で閉じていること示さなくてはいけません．すなわち，この和やスカラー積が絶体2乗和が有

界であることを示さなくてはいけません．$\boldsymbol{x}$ と $\boldsymbol{y}$ の2乗和が α_x と α_y で抑えられているとすれば，和は，

$$\sum_{i=1}^{M}|x_i+y_i|^2 \leq \sum_{i=1}^{M}(|x_i|^2+|y_i|^2+2|x_i||y_i|) \leq \sum_{i=1}^{M}(|x_i|^2+|y_i|^2+|x_i|^2+|y_i|^2) < 2(\alpha_x+\alpha_y)$$

となり(2番目から3番目は相加相乗平均の不等式)，有界になります．スカラー積も，

$$\sum_{i=1}^{M}|\alpha x_i|^2 \leq |\alpha|^2 \sum_{i=1}^{M}|x_i|^2 < |\alpha|^2\alpha_x$$

となり有界になります．また，これが線形空間になること(6つの公理を満たすこと)は簡単に証明できます.

　このとき，$\boldsymbol{x}$ と $\boldsymbol{y}$ の内積を

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{\infty} x_i y_i$$

で定義します．右辺の無限和の絶対和は,

$$\sum_{i=1}^{M} |x_i y_i| \leq \frac{1}{2} \sum_{i=1}^{\infty} (|x_i|^2 + |y_i|^2) < \frac{1}{2}(\alpha_x + \alpha_y).$$

となり，有界になります．単調増加する有界列は必ず収束しますから，和は絶対収束します．絶対収束する級数は収束しますので，l^2 の元に対して内積が定義できます．これが内積の公理を満たすことは明らかであると思います.

問題 9.

- l^2 が線型空間になることを示して下さい.
- 上の定義が内積の公理を満たすことを示して下さい.

1.11.2　直交

内積が定義されている実空間では角度が定義できますが，この角度が$\pi/2$ (90度) になっている場合を「直交」していると呼びます．$\cos\pi/2 = 0$ですから，2つのベクトル$\boldsymbol{x}, \boldsymbol{y}$が直交しているならば,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \tag{35}$$

となります．複素空間でも，このように内積が0の場合を**直交**していると呼びます．このとき，次のピタゴラスの定理が成立します.

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \tag{36}$$

問題 10. 上の式を証明して下さい.

1.11.3　正規直交基底

正規直交基底(orthonormal basis, ONB)$\{\boldsymbol{\varphi}_i\}$とは，ノルムが1で互いに直交している基底のことです.

$$\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \tag{37}$$

とすれば,

$$\langle \boldsymbol{\varphi}_i, \boldsymbol{\varphi}_j \rangle = \delta_{ij} \tag{38}$$

となます.

次に示す**Schmidtの直交化法**によって，任意の基底$\{\boldsymbol{\psi}_i\}$はONBにすることができます.

1. $n = 1$
2. $\boldsymbol{a}_n = \boldsymbol{\psi}_n - \sum_{i=1}^{n-1} \langle \boldsymbol{\psi}_n, \boldsymbol{\varphi}_i \rangle \boldsymbol{\varphi}_i$
3. $\boldsymbol{\varphi}_n = \boldsymbol{a}_n / \|\boldsymbol{a}_n\|$
4. go to 2

問題 11. 数学的帰納法を使って，$\{\boldsymbol{\psi}_i\}$が*ONB*であることを示して下さい.

1.11.4　平行四辺形公式

ノルムが内積から定義されたものであるとき ($\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$)，次の式が成立します.

$$2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\| = \|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 \tag{39}$$

これは，平行四辺形の2つの対角線の2乗の和が，平行でない2辺($\boldsymbol{x}$と$\boldsymbol{y}$)の2乗の和の2倍であることを表わしています.

問題 12. 上の式を証明して下さい.

この逆が成立する．すなわち，式(39)が成立すれば，ある内積が存在して，そのノルムは内積から定義されたものとすることができます.

ノルムが存在しても，それに対応する内積が存在するとは限りません(そのノルムが内積から導かれたものになるとは限らないということです). しかしながら，式(39)が成立すれば，存在することになります.

この証明は結構面倒です

2　共役作用素

定義 4. 作用素Aに対して，その共役作用素A^*は，任意のベクトル$\boldsymbol{x}, \boldsymbol{y}$に対して，

$$\langle \boldsymbol{x}, A\boldsymbol{y} \rangle = \langle A^*\boldsymbol{x}, \boldsymbol{y} \rangle$$

を満たすものとして定義されます.

問題 13. 任意のベクトル$\boldsymbol{x}, \boldsymbol{y}$に対して，

$$\langle A^*\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, A\boldsymbol{y} \rangle$$

が成立することを示して下さい.

$A^{**} = (A^*)^* = A$が成立します.

問題 14. 上の式を示して下さい.

問題 15. 上の問題のために，あるベクトル$\boldsymbol{x}$が任意のベクトル$\boldsymbol{y}$に対して，

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$$

を満たすならば，$\boldsymbol{x} = \boldsymbol{\theta}$を示して下さい.

実ベクトル空間では，行列Aの共役作用素は転置行列A^Tになります．

$$\langle \boldsymbol{x}, A\boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{m} A_{ij} y_j \right) \qquad \langle A^T \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} A_{ij} x_i \right) y_j$$

より，明らかです．

$$(A^T)_{ij} = A_{ji}$$

定義 5. 作用素Aが**自己共役**であるとは，Aが

$$A^* = A$$

を満たすことです．

実ベクトル空間で作用素を行列で表わした場合，**自己共役作用素とは対称行列**のことです．

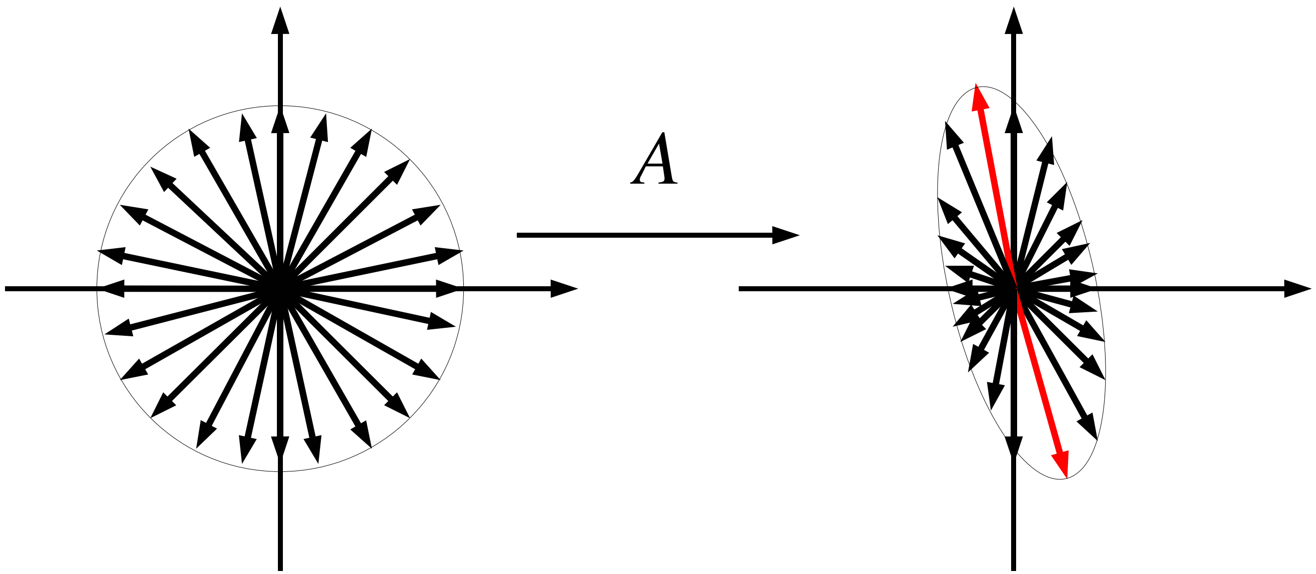
$$A_{ij} = A_{ji}$$

3　作用素のノルム

ノルムが定義された空間からノルムが定義された空間への作用素 A にはノルム $\|A\|$ を定義することができます．一般に，作用素のノルムと言えば，

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \tag{40}$$

で定義されます．これは，写像元のノルムが1のベクトルを様々に変えるとき，そのベクトルの写像先でのノルムの最大値として定義されています．どれだけベクトルを拡大するか示したものです．



4　作用素のトレース

Xを内積が定義された空間とします．$\{\boldsymbol{\varphi}_i\}$を$X$のONBとします．$X$から$X$への作用素$A$に対して，トレース$\mathrm{tr}(A)$を,

$$\mathrm{tr}(A) = \sum_i \langle \boldsymbol{\varphi}_i, A\boldsymbol{\varphi}_i \rangle \tag{41}$$

で定義します．この定義にONBが使われていて，ONBの選び方によってはトレースの値が変化しそうですが，変化せず一定です．$\{\boldsymbol{\varphi}_i\}$とは別のONBを使っても，上式の和は同じ値になります．

問題 16. トレースの値が，*ONB*の選び方に依らないことを示して下さい．

問題 17. XからYへの任意の作用素AとYからXへの任意の作用素Bに対して,

$$\begin{aligned} \mathrm{tr}(AB) &= \mathrm{tr}(BA) \\ \mathrm{tr}(A^*) &= \overline{\mathrm{tr}(A)} \end{aligned}$$

を示して下さい．

ユークリッド空間では，ONBとして自然基底を使えば,

$$\mathrm{tr}(A) = \sum_i A_{ii} \tag{42}$$

という，行列の対角和になることがすぐにわかります．(自然基底でなくても，トレースはONBの選び方に依らないので，対角和になる．)

4.1 作用素のHilbert-Schmidt ノルム

作用素Aに対して，$\mathrm{tr}(AA^*)$の値が一定のとき，AのHilbert-Schmidt ノルムを,

$$\|A\|_2 = \sqrt{\mathrm{tr}(AA^*)} \tag{43}$$

によって定義します．($\|A\|_2^2$は，$\|A\|^2$の2乗という意味です．)

一般に,

$$\|A\|_2 \geq \|A\| \tag{44}$$

が成立します.

5　射影作用素

部分空間への正射影という概念は図形にわかると思います．図のような場合です.

ここでは正射影作用素の前に，まず射影作用素を定義します．これは，図のように斜めに射影することを含みます.

定義 6. 作用素Pが射影作用素であるとは，

$$PP = P \tag{45}$$

を満たすことです.

ベクトル的に考えると，$\boldsymbol{f}$を一度Pで変換した$P\boldsymbol{f}$をもう一度Pで変換しても，$PP\boldsymbol{f} = P\boldsymbol{f}$となり変らないということを意味しています.

式(45)が成立すると，

$$\begin{aligned} R(P) &= N(I - P) && (46) \\ N(P) &= R(I - P) && (47) \\ R(P) \cap N(P) &= \{\boldsymbol{\theta}\} && (48) \\ R(P) + N(P) &= R^N && (49) \end{aligned}$$

が成立します．この式から，P は $N(P)$ に沿った $R(P)$ への射影と解釈することができます．

証明

$\boldsymbol{f} \in R(A)$ とすれば，$\boldsymbol{f} = P\boldsymbol{g}$ となる $\boldsymbol{g}$ が存在し，

$$(I-P)\boldsymbol{f} = (I-P)P\boldsymbol{g} = P\boldsymbol{g} - PP\boldsymbol{g} = P\boldsymbol{g} - P\boldsymbol{g} = 0$$

となります．従って，

$$R(P) \subset N(I-P) \tag{50}$$

となります．

逆に，$\boldsymbol{f} \in N(I-P)$ とすれば，$(I-P)\boldsymbol{f} = \boldsymbol{\theta}$ となり，$\boldsymbol{f} = P\boldsymbol{f}$ となるので，$\boldsymbol{f} \in R(A)$ となります．従って，

$$R(P) \supset N(I-P) \tag{51}$$

となります．式(50)，(51) より，式(46)が成立します．

$$(I-P)(I-P) = I - P - P - P^2 = I - P \tag{52}$$

より，$I-P$ も射影作用素になりますので，$I-P$ を P と考えれば，式(46) より式(47)が成立します．

$\boldsymbol{f} \in R(P)$ならば$\boldsymbol{f} = P\boldsymbol{g}$となる$\boldsymbol{g}$が存在するので，$\boldsymbol{f} = P\boldsymbol{g} = PP\boldsymbol{g} = P\boldsymbol{f}$となります．従って，$\boldsymbol{f} \in R(P) \cap N(P) = R(P) \cap R(I-P)$とすれば，$\boldsymbol{f} = P\boldsymbol{f}$と$\boldsymbol{f} = (I-P)\boldsymbol{f}$となり，$\boldsymbol{f} = \boldsymbol{\theta}$より式(48)が成立します．一般に，$R(A+B) \subset R(A) + R(B)$であるので，式(49)が成立します．

定義 7. 作用素Pが正射影作用素であるとは，式*(45)*の他に，

$$P^* = P \tag{53}$$

を満たすことです．

正射影ですから，沿う空間$N(A)$と射影先の空間$R(P)$が直交していることになります．任意の$x \in N(A)$と任意の$y \in R(P)$に対して，

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, P\boldsymbol{y} \rangle = \langle P^*\boldsymbol{x}, \boldsymbol{y} \rangle = \langle P\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{\theta}, \boldsymbol{y} \rangle = 0$$

となり，直交していることが分ります．

式(45)，(53)の2つの式は，

$$P^*P = P$$

とまとめることができます．

問題 18. 上式を示して下さい．

Pを部分空間Sへの正射影作用素とします($R(P) = S$)．そのベクトル空間の元$\boldsymbol{x}$からSにもっとも近い元は，Pxによって与えられます．Sの元yに対して，$Py = y$であるので，

$$\langle y, x\rangle = \langle Py, x\rangle = \langle P^2 y, x\rangle = \langle Py, P^* x\rangle = \langle Py, Px\rangle$$

が成立します．従って，

$$\begin{aligned}
\|y - x\|^2 &= \|Py - x\|^2 \\
&= \|Py\|^2 - \langle Py, x\rangle - \langle x, Py\rangle + \|x\|^2 \\
&= \|Py\|^2 - \langle Py, Px\rangle - \langle Px, Py\rangle + \|Px\|^2 - \|Px\|^2 + \|x\|^2 \\
&= \|Py - Px\|^2 - \|Px\|^2 + \|x\|^2 \\
&= \|y - Px\|^2 - \|Px\|^2 + \|x\|^2
\end{aligned}$$

となります．上式より，$y = Px$のとき，xとの距離が最小になります．

6　固有値・固有ベクトル

A : n次対称行列

固有方程式

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \tag{54}$$

λ : 固有値

$\boldsymbol{x}$: 固有ベクトル ($\boldsymbol{x} \neq \boldsymbol{\theta}$)

縮退を含めてn個の解が存在します.

固有値 : $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

固有ベクトル : $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

このとき，$\{u_i\}_{i=1}^{n}$ を正規直交基底となるように取ることができます.

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{i,j} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \tag{55}$$

(証明)

式(54) は,

$$(A - \lambda I)\boldsymbol{x} = 0, \tag{56}$$

となります．ここで，I は単位行列です．$\boldsymbol{x} \neq 0$ より $A\boldsymbol{x} - \lambda I$ は正則であってはいけません．従って,

$$\det(A - \lambda I) = 0 \tag{57}$$

が成立します．これは，λ に関する n 次方程式です．従って,「代数学の基本定理」より，複素数まで含めれば n 個の解が存在します．その1つの解を λ と記します．この λ に対して，成分が複素数の n 次ベクトル $\boldsymbol{x} = (x_i)$ が存在して,

$$(A - \lambda I)\boldsymbol{x} = 0 \tag{58}$$

が成立します．Aの成分A_{ij}は実数ですので，

$$\lambda\sum_{i=1}^{n}|x_i|^2 = \sum_{i=1}^{n}\lambda x_i\overline{x_i} = \sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_j\overline{x_i} = \sum_{j=1}^{n}\sum_{i=1}^{n}x_j\overline{A_{ij}x_i} \tag{59}$$

$$= \sum_{j=1}^{n}x_j\overline{\sum_{i=1}^{n}A_{ji}x_i} = \sum_{j=1}^{n}x_j\overline{\lambda x_j} = \bar{\lambda}\sum_{j=1}^{n}|x_j|^2 \tag{60}$$

となり，λが実数であることがわかります．また，

$$\sum_{j=1}^{n}A_{ij}(x_j+\overline{x_j}) = \lambda(x_i+\overline{x_i}), \tag{61}$$

$$\sum_{j=1}^{n}A_{ij}\Re e(x_i) = \lambda\Re e(x_i) \tag{62}$$

となるので，一般性を失うことなく，全てのx_iを実数の中から取ることができます．

従って，少くとも1組の実数の固有値と固有ベクトルが存在します．$\lambda_n = \lambda$，$\boldsymbol{u}_n = \boldsymbol{x}/\|\boldsymbol{x}\|$とおき，$\{\boldsymbol{v}_i\}_{i=1}^{n-1}$を$u_n$の直交補空間($n-1$次元の部分空間) の正規直交系とします．従って，$n$以下の任意の自然数$i, j$に対して，$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij}$と$\langle \boldsymbol{v}_i, \boldsymbol{u}_n \rangle = 0$が成立します．$(n, n-1)$行列$V$を，

$$V = (\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \cdots \ \boldsymbol{v}_{n-1}).$$

と定義します．

$$V^T V = I_{n-1}.$$

は明らかです．また，

$$(V^T AV)^T = V^T A^T V^{TT} = V^T AV,$$

となりますので，$V^T AV$は，$n-1$次対称行列になります．

あとは数学的帰納法を使って証明します．$n = 1$のときに，固有ベクトルと固有値が存在することは明らかです．$V^T AV$は，$n-1$次対称行列ですので，数学的帰納法の仮定より，固有値$\lambda_1, \cdots, \lambda_{n-1}$と，$n-1$次元実ベクトル空間の正規直交基底になる対応する固有ベクトル$\{u'_1, \cdots, u'_{n-1}\}$が存在します．定義より，次の式が成立します．

$$V^T AV\boldsymbol{u}'_i = \lambda_i \boldsymbol{u}'_i$$

従って,

$$\langle \boldsymbol{u}_n, AV\boldsymbol{u}'_i\rangle = \langle A\boldsymbol{u}_n, V\boldsymbol{u}'_i\rangle = \lambda_n\langle V^T\boldsymbol{u}_n, \boldsymbol{u}'_i\rangle = \lambda_n\langle \boldsymbol{\theta}, \boldsymbol{u}'_i\rangle = 0$$

となります．$\boldsymbol{u}_n \perp AV\boldsymbol{u}'_i$であり,

$$VV^TAV\boldsymbol{u}'_i = \sum_{j=1}^{n-1}\langle AV\boldsymbol{u}'_i, \boldsymbol{v}_j\rangle\boldsymbol{v}_j = AVu'_i$$

ですので,

$$AV\boldsymbol{u}'_i = \lambda_i V\boldsymbol{u}'_i \quad (i = 1, 2, \cdots, n)$$

と

$$\langle V\boldsymbol{u}'_i, V\boldsymbol{u}'_j\rangle = \langle V^TV\boldsymbol{u}'_i, \boldsymbol{u}'_j\rangle = \langle \boldsymbol{u}'_i, \boldsymbol{u}'_j\rangle = \delta_{ij}, \quad \langle V\boldsymbol{u}'_i, \boldsymbol{u}_n\rangle = 0.$$

$\lambda_1, \cdots \lambda_n, V\boldsymbol{u}'_1, \cdots V\boldsymbol{u}'_{n-1}, \boldsymbol{u}_n$が，求める固有値と固有ベクトルになります.

注意 $\lambda_i \neq \lambda_j$のとき，$\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle = 0$となります．実際,

$$\lambda_i\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle = \langle A\boldsymbol{u}_i, \boldsymbol{u}_j\rangle = \langle \boldsymbol{u}_i, A\boldsymbol{u}_j\rangle = \lambda_j\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle$$

従って，$(\lambda_i - \lambda_j)\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle = 0$となり，題意が成立します.

Remark $\lambda_i = \lambda_j$のときは，$\boldsymbol{u}_i$と$\boldsymbol{u}_j$の任意の線型結合が固有ベクトルになります．

$A\boldsymbol{u}_i = \lambda\boldsymbol{u}_i$と$A\boldsymbol{u}_j = \lambda\boldsymbol{u}_j$が成立するので，実数$\alpha$と$\beta$に対して，

$$A(\alpha\boldsymbol{u}_i + \beta\boldsymbol{u}_j) = \lambda(\alpha\boldsymbol{u}_i + \beta\boldsymbol{u}_j)$$

となり，$\alpha\boldsymbol{u}_i + \beta\boldsymbol{u}_j$は固有ベクトルになっています．

$U = (\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots\ \boldsymbol{u}_n)$と置けば，次の式が成立します．

$$U^T U = U U^T = I,$$

$$U^T A U = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad A = U \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} U^T,$$

$$A = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T$$

(後の式の証明)

$\{u_i\}$はONBなので，任意のxに対して，$x = \sum_i \langle x, u_i \rangle u_i$と書くことができます．従って，

$$A\boldsymbol{x} = A\sum_i \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i = \sum_i \langle x, \boldsymbol{u}_i \rangle A\boldsymbol{u}_i = \sum_i \lambda_i \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i$$

となり，ベクトル$\boldsymbol{a}$と$\boldsymbol{b}$に対して，$\boldsymbol{a}\boldsymbol{b}^T\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{b} \rangle \boldsymbol{a}$となりますので，

$$\left(\sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T \right) x = \sum_i \lambda_i \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i$$

となります．

問題 19. 正射影行列Pの固有値は，*0*または*1*であることを示して下さい．従って，Pは下のように書くことができます．

$$P = \sum_{i=1}^{d} \boldsymbol{u}_i \boldsymbol{u}_i^T$$

ここで，d = (*1*の固有値の数) = ($R(P)$の次元) = (Pの*rank*)です．

問題 20. m個のn次元データ$\{\boldsymbol{f}_i\}_{i=1}^m$に対して，$\mathrm{rank}(X) \le l$の$n$次正方行列の中で,

$$\sum_{i=1}^m \|X\boldsymbol{f}_i - \boldsymbol{f}_i\|^2$$

を最小にするものを求めて下さい(十分なものでで構いません).

(ヒント)

- Pを$R(X)$への正射影行列とすれば，$\|P\boldsymbol{f} - \boldsymbol{f}\| \le \|X\boldsymbol{f} - \boldsymbol{f}\|$.
- $\|P\boldsymbol{f} - \boldsymbol{f}\|^2 = \|\boldsymbol{f}\|^2 - \|P\boldsymbol{f}\|^2$.
- $\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \mathrm{tr}(\boldsymbol{f}\boldsymbol{g}^T)$.
- $R = \sum_{i=1}^m \boldsymbol{f}_i\boldsymbol{f}_i^T$とおけば，$\sum_{i=1}^m \|P\boldsymbol{f}_i\|^2 = \mathrm{tr}(PRP)$.
- Rの固有値を$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$，それに対応するONBをなす固有ベクトルを$\{\boldsymbol{\varphi}_i\}_{i=1}^n$とおけば，$\mathrm{tr}(PRP) = \sum_{i=1}^n \lambda_i\|P\boldsymbol{\varphi}_i\|^2$.
- $\sum_{i=1}^n \|P\boldsymbol{\varphi}_i\|^2 = \mathrm{tr}(P) \le l$.
- 一般に，α_i $(0 \le \alpha_i \le 1)$かつ$\sum_{i=1}^n \alpha_i \le l$の条件の元，$\sum_{i=1}^n \lambda_i\alpha_i$を最大にする$\alpha_i$は，$\alpha_1 = \alpha_2 = \cdots \alpha_l = 1$，$\alpha_{l+1} = \alpha_{l+2} = \cdots \alpha_n = 0$.

7　特異値分解

A: (n, m)-行列

目的： 対称とは限らないAを，Aが対称行列のときの固有値よる方法のように分解する．

以下の式より明らかに，AA^T は対称な(n, n)次行列です．

$$(AA^T)^T = (A^T)^T A^T = AA^T.$$

従って，AA^Tは，固有値$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0$とそれに対応する固有ベクトル$u_1 \geq u_2 \geq \cdots u_n \geq 0$を持ち，$\{u_i\}_{i=1}^n$をONBとすることができます．従って，

$$AA^T u_i = \lambda_i u_i.$$

が成立します．また，

$$\lambda_i \|u_i\|^2 = \langle \lambda_i u_i, u_i \rangle = \langle AA^T u_i, u_i \rangle = \langle A^T u_i, A^T u_i \rangle = \|A^T u_i\|^2.$$

より，$\lambda_i \geq 0 \ (i = 1, 2, \cdots, n)$となります．

同様にして，$A^T A$は固有値$\mu_1 \geq \mu_2 \geq \cdots \mu_n \geq 0$を持つ．このとき，

$$A^T A(A^T u_i) = A^T (AA^T u_i) = \lambda_i A u_i$$

が成立するので，λ_i と $A^T u_i$ は，それぞれ，$A^T A$ の固有値とそれに対応する固有ベクトルです．従って，$i = 1, 2, \cdots, \min(m, n)$ に対して，

$$\mu_i = \lambda_i$$

が成立します．

いま，$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$，$\lambda_{k+1} = \cdots = \lambda_n = 0$ とします．$\boldsymbol{v}_i$ を $i \leq k$ に対して，

$$v_i = A^T u_i / \sqrt{\lambda_i}$$

とおき，$\{\boldsymbol{v}_i\}_{i=k+1}^m$ を，$\{\boldsymbol{x} \mid A^T A \boldsymbol{x} = 0\}$ の任意の正規直交系とおきます．

$\{\boldsymbol{v}_i\}_{i=1}^m$ がONBになることを示す．任意の $i \leq k$ と任意の $j \leq k$ に対して，

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \left\langle \frac{A^T u_i}{\sqrt{\lambda_i}}, \frac{A^T u_j}{\sqrt{\lambda_j}} \right\rangle = \frac{\langle A A^T u_i, u_j \rangle}{\sqrt{\lambda_i}\sqrt{\lambda_j}} = \frac{\lambda_i}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij}$$

が成立します．さらに，任意の $i (\leq k)$ と任意の $j (> k)$ に対して，

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \langle A^T \boldsymbol{u}_i / \sqrt{\lambda_i}, \boldsymbol{v}_j \rangle = \langle A^T A A^T \boldsymbol{u}_i / (\lambda_i \sqrt{\lambda_i}), \boldsymbol{v}_j \rangle = \langle \boldsymbol{v}_i / \lambda_i, A^T A \boldsymbol{v}_j \rangle = 0$$

とります．従って，任意の $i\ (\leq m)$ と任意の $j\ (\leq m)$ に対して，

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta_{ij}.$$

が成立します.

$i \leq k$ のとき，$\boldsymbol{v}_i$ の定義より，

$$A^T \boldsymbol{u}_i = \sqrt{\lambda_i} \boldsymbol{v}_i,$$

と

$$A\boldsymbol{v}_i = AA^T \boldsymbol{u}_i / \sqrt{\lambda_i} = \sqrt{\lambda_i} \boldsymbol{u}_i.$$

が成立します．また，$\{\boldsymbol{x} \,|\, A\boldsymbol{x} = 0\} = \{\boldsymbol{x} \,|\, A^T A\boldsymbol{x} = 0\}$ が成立するので，$i > k$ に対して $A\boldsymbol{v}_i = 0$ となります．従って，

$$A^T \boldsymbol{u}_i = \begin{cases} \sqrt{\lambda_i} \boldsymbol{v}_i & (i \leq m) \\ \boldsymbol{\theta} & (\text{else}) \end{cases}, \tag{63}$$

$$A\boldsymbol{v}_i = \begin{cases} \sqrt{\lambda_i} \boldsymbol{u}_i & (i \leq n) \\ \boldsymbol{\theta} & (\text{else}) \end{cases}. \tag{64}$$

となります．(n,n)-行列 U と (m,m)-行列 V を

$$U = (\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots\ \boldsymbol{u}_n) \tag{65}$$

$$V = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_m) \tag{66}$$

と定義します．$U^TU = I$と$V^TV = I$より，

$$UU^T = U^TU = I, \quad VV^T = V^TV = I,$$

となります．また，

$$\begin{aligned} U^TAV &= U^T(A\boldsymbol{v}_1\ A\boldsymbol{v}_2\ \cdots\ A\boldsymbol{v}_m) \\ &= \begin{pmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T \end{pmatrix} \left(\sqrt{\lambda_1}\boldsymbol{u}_1\ \sqrt{\lambda_2}\boldsymbol{u}_2\ \cdots\ \sqrt{\lambda_{\min(m,n)}}\boldsymbol{u}_{\min(m,n)}\ \boldsymbol{\theta}\cdots\boldsymbol{\theta} \right) \end{aligned}$$

となります．

この右辺を計算すると,

$$U^T AV = \begin{cases} \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & & & \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & \boldsymbol{\theta} & \cdots & \boldsymbol{\theta} \\ \vdots & \ddots & \ddots & \vdots & & & \\ 0 & 0 & \cdots & \sqrt{\lambda_{\min(m,n)}} & & & \end{pmatrix} & (m \geq n) \\ \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_{\min(m,n)}} \\ & & \boldsymbol{\theta}^T & \\ & & \vdots & \\ & & \boldsymbol{\theta}^T & \end{pmatrix} & (m < n) \end{cases}$$

となり，2つの直交行列によって対角化されることがわかります.

また，Uを左から，V^Tを右からかけることによって，次の式を得ます．

$$A = U \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & & & \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & [\boldsymbol{\theta}] & \cdots & [\boldsymbol{\theta}] \\ \vdots & \ddots & \ddots & \vdots & & & \\ 0 & 0 & \cdots & \sqrt{\lambda_{\min(m,n)}} & & & \\ & & [\boldsymbol{\theta}^T] & & & & \\ & & \vdots & & & & \\ & & [\boldsymbol{\theta}^T] & & & & \end{pmatrix} V^T \tag{67}$$

これは次のように書くこともできます．

$$A = \sum_{i=1}^{k} \sqrt{\lambda_1} \boldsymbol{u}_i \boldsymbol{v}_i^T = \sqrt{\lambda_i} \boldsymbol{u}_1 \boldsymbol{v}_1^T + \sqrt{\lambda_1} \boldsymbol{u}_2 \boldsymbol{v}_2^T + \cdots + \sqrt{\lambda_k} \boldsymbol{u}_k \boldsymbol{v}_k^T$$

(Proof.)

$\{\boldsymbol{v}_i\}_{i=0}^m$はONBであり，$A$と上式の左辺を$\boldsymbol{v}_i$に作用させると，両方から$\sqrt{\lambda_i} \boldsymbol{u}_i$が得られます．従って，両者は等しくなります．

(n,k)-行列 U' と (m,k)-行列 V' を,

$$
\begin{aligned}
U' &= (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k), \\
V' &= (\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k)
\end{aligned}
$$

とおきます. $((U')^T U' = I_k$, $(V')^T V' = I_k$ は成立しています. $)$

$$
{U'}^T A V = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_k} \end{pmatrix}
$$

および

$$
A = U' \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_k} \end{pmatrix} {V'}^T
$$

が成立します.

問題 21.

$$A = U' \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_k} \end{pmatrix} V'^T$$

問題 22. $\{\boldsymbol{x} \mid A\boldsymbol{x} = 0\} = \{\boldsymbol{x} \mid A^T A\boldsymbol{x} = 0\}$ を示して下さい.

問題 23. 任意の (n, m)-行列 A に対して，(n, m)-行列 X が，$\mathrm{rank}(X) \leq l$ の条件の元で

$$\|X - A\|_2^2 = \mathrm{tr}((X - A)^T (X - A))$$

を最小にする必要十分条件は，固有値の縮退による自由度を除いて,

$$A = \sum_{i=1}^{l} \sqrt{\lambda_i} \boldsymbol{u}_i \boldsymbol{v}_i^T$$

が成立することであることを示して下さい.

8　一般逆行列(作用素)

A : (n, m) 行列(正則でなくてもよい)

一般逆行列の定義：

1-逆 $AXA = A,$ 　　 $X = A^{(1)}$ or A^{-}

2-逆 $XAX = X,$ 　　 $X = A^{(2)}$

3-逆 $(AX)^T = AX,$ 　　 $X = A^{(3)}$

4-逆 $(XA)^T = XA,$ 　　 $X = A^{(4)}$

一般逆の表記の例：

$A^{(1)}, A^{-}$: A の 1-逆行列

$A^{(1,2)}$: A の 1, 2-逆行列

$A^{(1,2,3,4)}, A^{\dagger}$: 1, 2, 3, 4-行列, Moore-Penrose 一般逆行列

$A\{1\}$: A の 1-逆行列全体の集合

$A\{1, 2\}$: A の 1, 2-逆行列全体の集合

補題 1. 任意の行列 A に対して，次式が成立します.

$$R(A)^{\perp} = N(A^T), \quad R(A) = N(A^T)^{\perp} \tag{68}$$
$$R(A^T)^{\perp} = N(A), \quad R(A^T) = N(A)^{\perp} \tag{69}$$

(証明)

任意の $\boldsymbol{x} \in R(A)^{\perp}$ と任意の $\boldsymbol{y} \in R^m$ に対して，$A\boldsymbol{y} \in R(A)$ ですので，

$$\langle x, Ay \rangle = 0$$

が成立します.

$$\Rightarrow \langle A^T x, y \rangle = 0 \Rightarrow A^T x = 0 \Rightarrow x \in N(A^T) \Rightarrow R(A)^{\perp} \subset N(A^T).$$

逆に，任意の $\boldsymbol{x} \in N(A)$ と任意の $\boldsymbol{y} \in R^m$ に対して，$A\boldsymbol{y} \in \boldsymbol{\theta}$ ですので，

$$0 = \langle A^T x, y \rangle = \langle x, Ay \rangle \Rightarrow x \in R(A)^{\perp} \Rightarrow N(A^T) \subset R(A)^{\perp}$$

となり，$R(A)^{\perp} = N(A^T)$ が成立します.

残りの式は，両辺の直交補空間を考える，あるいは，A に A^T を代入することによって得られます.

8.1　1-逆行列

次の連立1次方程式を考えます.

$$Ax = y. \tag{70}$$

もし，式(70)が解を持てば($\Leftrightarrow y \in R(A)$)，その解の1つは，次の式で与えられます.

$$A^{-}y$$

(証明)

$x = A^{-}y$とおけば，下式より明らかです.

$$Ax = A(A^{-}y) = AA^{-}Ax = Ax = y$$

$A^{-}A$とAA^{-}は，射影行列になります.

$$(A^{-}A)^2 = A^{-}AA^{-}A = A^{-}A$$

$$(AA^{-})^2 = AA^{-}AA^{-} = AA^{-}$$

一般に，1-逆行列は一意に決まりません．次の式は，A の一般形です．

$$X = A^- + W - A^-AWAA^-. \tag{71}$$

すなわち，任意の W に対して X は A の 1-逆行列で，A の任意の 1-逆行列はこの形で書くことができます．

(証明)

$$\begin{aligned} AXA &= A(A^- + W - A^-AWAA^-)A \\ &= AA^-A + AWA - AA^-AWAA^-A \\ &= A + AWA - AWA = A. \end{aligned}$$

また，B を A の任意の 1-逆行列とすれば，$W = B - A^-$ とおいて，

$$X = A^- + (B - A^-) - A^-A(B - A^-)AA^- = B + A^-(A - A)A^- = B.$$

が成立します．従って，式(71)は一般形です．

次の式が成立します.

$$R(A) = R(AA^-) \tag{72}$$

$$N(A) = N(A^-A) \tag{73}$$

(証明)

$$R(A) = R(AA^-A) \subset R(AA^-) \subset R(A)$$

となり $R(A) = R(AA^-)$ が成立します．同様に,

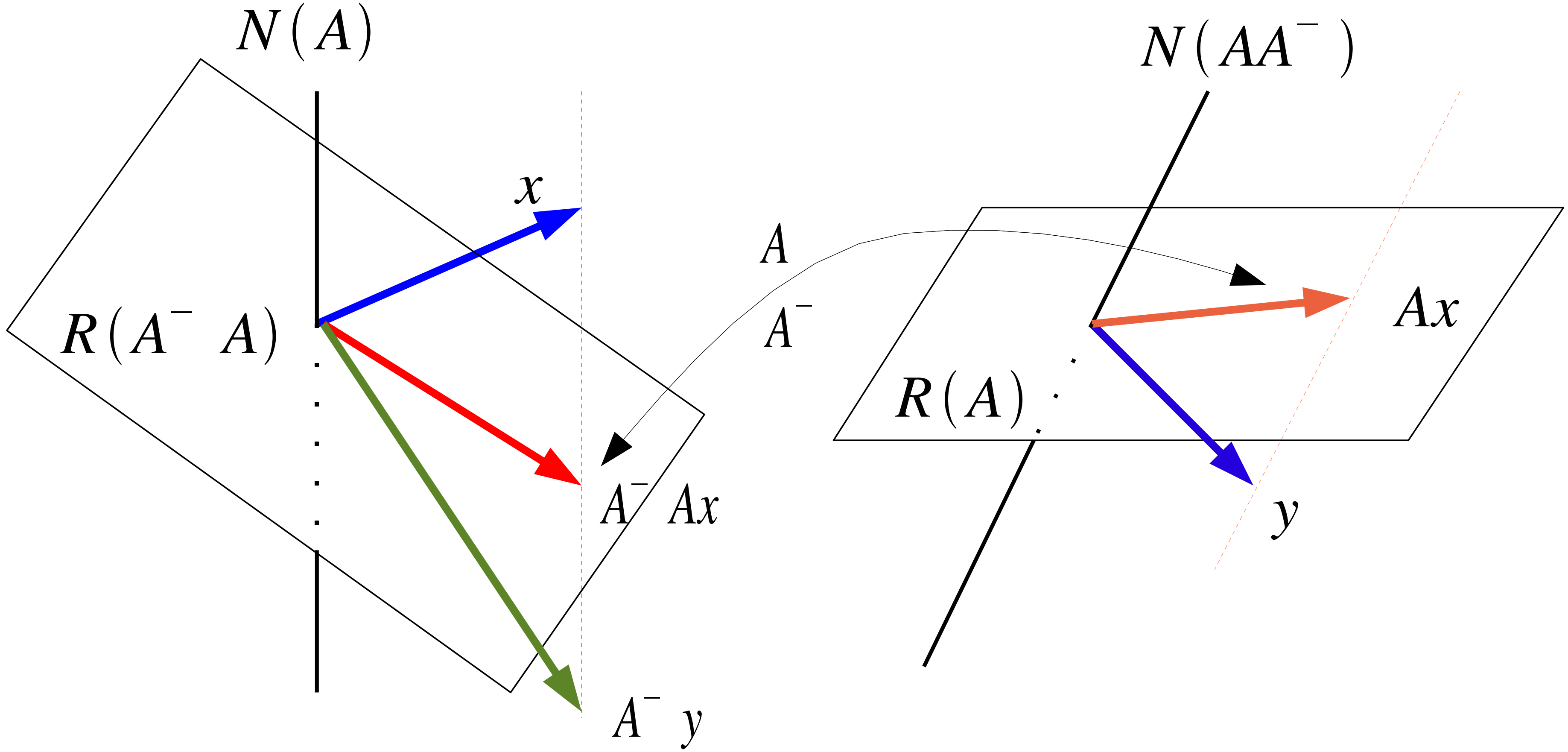
$$N(A) = N(AA^-A) \supset N(A^-A) \subset N(A)$$

が成立し，$N(A) = N(A^-A)$ が成立します.

しかし，一般には次の式は成立しません

$$\times\ R(A^-) = R(A^-A) \tag{74}$$

$$\times\ N(A^-) = N(AA^-) \tag{75}$$



- $R(A^{-}A)$ と $R(A)$ のベクトルは1対1に対応します.
  $A^{-}A\boldsymbol{x}$ と $A\boldsymbol{x}$ が対応します.
- $\boldsymbol{x} + N(A)$ のベクトルはすべて $A\boldsymbol{x}$ に写像されます.
- $A\boldsymbol{x} + N(AA^{-})$ のベクトルはすべて $A^{-}A\boldsymbol{x} + N(A)$ に写像されます.

8.2　1,2-逆

2-逆を加えることによる最大の特長は,

$$\mathrm{rank}(A) = \mathrm{rank}\left(A^{(1,2)}\right)$$

が成立することです．1-逆だけでは，$\mathrm{rank}(A) \leq \mathrm{rank}\left(A^{(1,2)}\right)$ です．ここで，$\mathrm{rank}(A)$ は，部分空間 $R(A)$ の次元で定義されます.

(証明) 一般に，$\mathrm{rank}(AB) \leq \mathrm{rank}(A)$ と $\mathrm{rank}(AB) \leq \mathrm{rank}(B)$ が成立します.
従って,

$$\mathrm{rank}(A) = \mathrm{rank}\left(AA^{(1,2)}A\right) \leq \mathrm{rank}\left(A^{(1,2)}\right)$$

と

$$\mathrm{rank}\left(A^{(1,2)}\right) = \mathrm{rank}\left(A^{(1,2)}AA^{(1,2)}\right) \leq \mathrm{rank}(A)$$

が成立します.

補題 2. *1-逆では成立しなかった,*

$$R\left(A^{(1,2)}\right) = R\left(A^{(1,2)}A\right) \tag{76}$$

$$N\left(A^{(1,2)}\right) = N\left(AA^{(1,2)}\right) \tag{77}$$

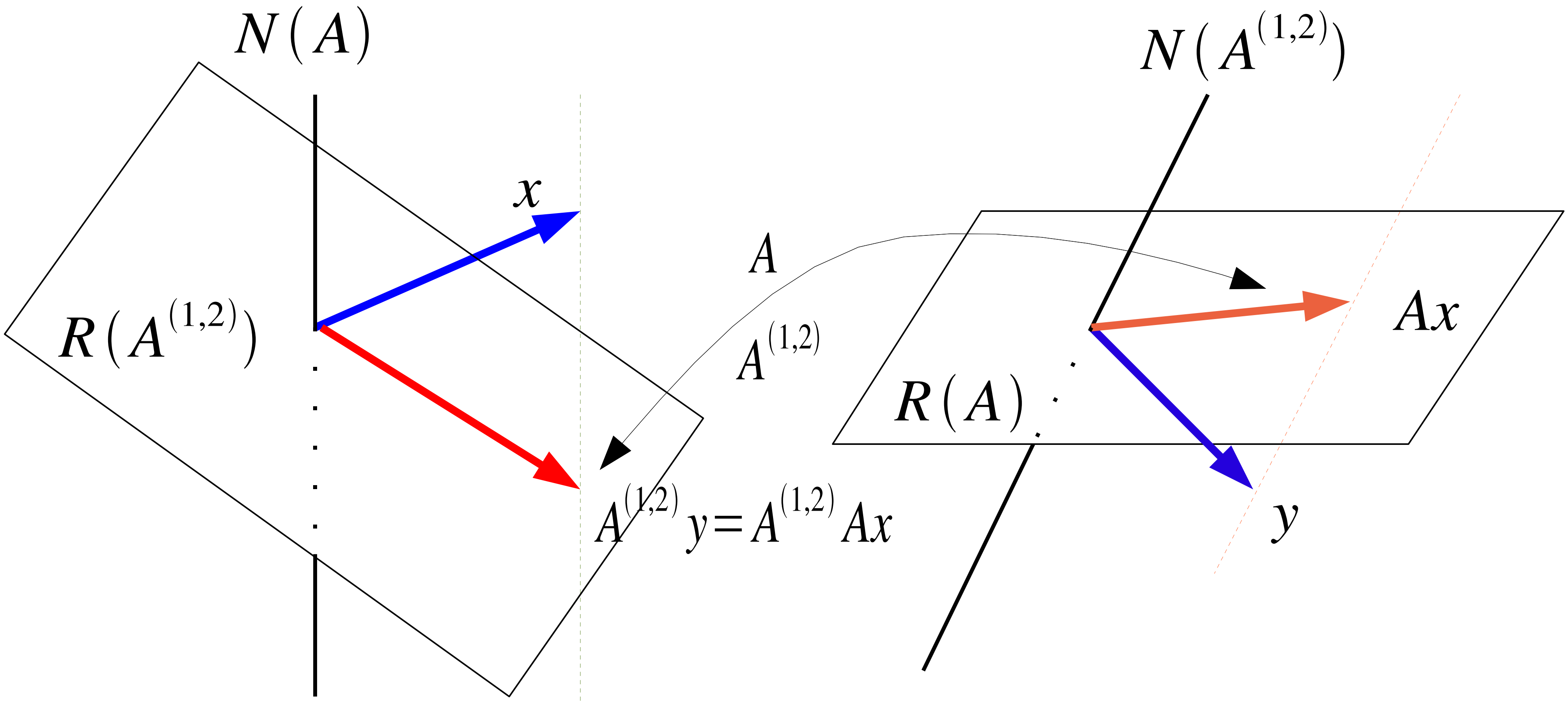
が成立します.

問題 24. 上の式を示してて下さい.

問題 25. $R(A^{(1,2)})$ の元と $R(A)$ の元の間に, *1* 対 *1* 対応が存在します. $A^{(1,2)}\boldsymbol{y}$ に対して $A(A^{(1,2)}\boldsymbol{y})$ が対応し, $A\boldsymbol{x}$ に対して $A^{(1,2)}(A\boldsymbol{x})$ が対応しています. これを示して下さい.

問題 26. *1*-逆における, $R(A)$ と $R(A^{-}A)$ の *1* 対 *1* 対応を示して下さい.

まだ, $R(A^{(1,2)})$ と $N(A^{(1,2)})$ は, それぞれ, $N(A)$ と $R(A)$ の補空間に任意に選ぶことができます.



- $R\left(A^{(1,2)}\right)$ と $R(A)$ のベクトルは1対1に対応します.
  $A^{(1,2)}A\boldsymbol{x}$ と $A\boldsymbol{x}$ が対応します.
- $\boldsymbol{x}+N(A)$ のベクトルはすべて $A\boldsymbol{x}$ に写像されます.
- $A\boldsymbol{x}+N\left(A^(1,2)\right)$ のベクトルはすべて $A^{(1,2)}A\boldsymbol{x}$ に写像されます.

8.3　1,3-逆

1,3-inverse: $N(AA^{(1,3)}) = R(A)^{\perp}$: 最小2乗誤差を与えます.

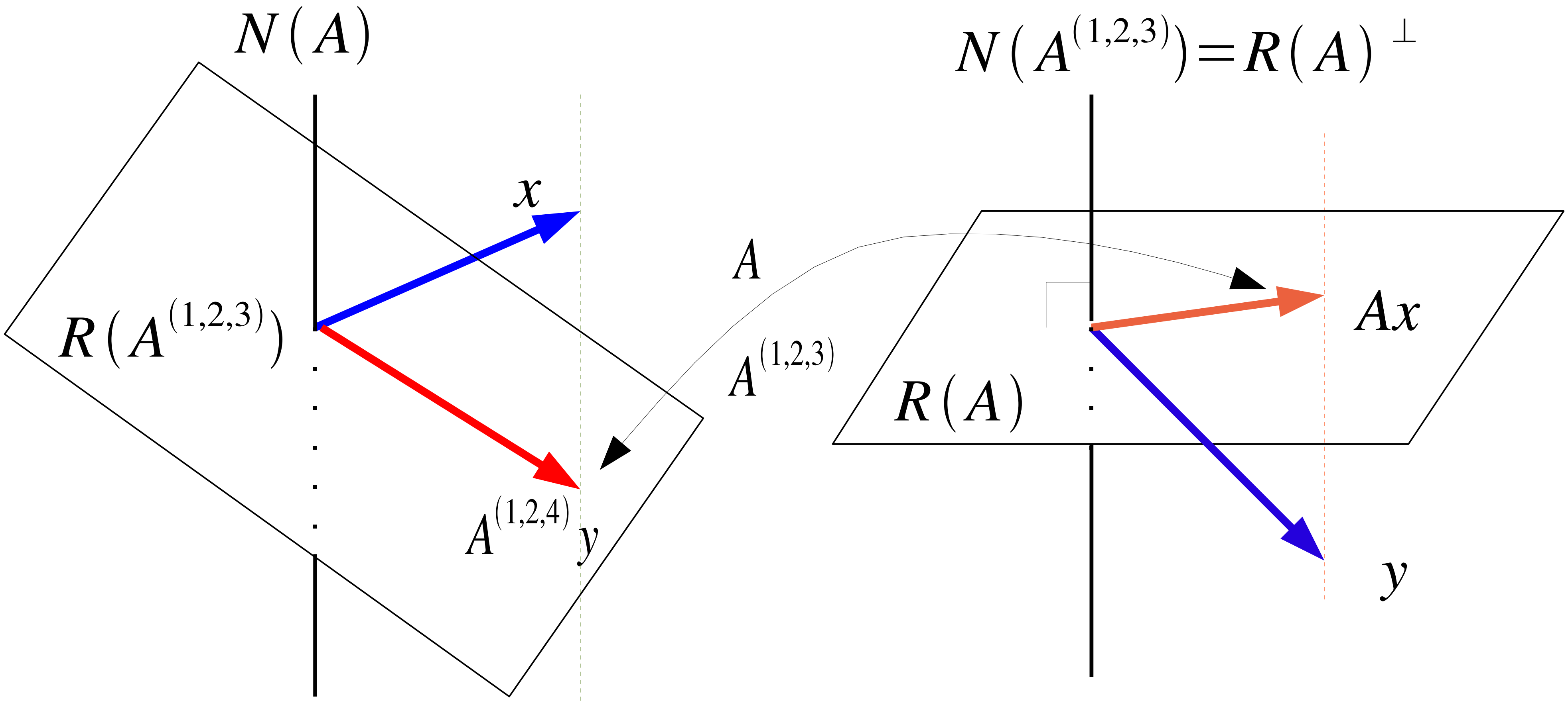
定理 1. 与えられた$\boldsymbol{y}$に対して,

$$\|A\boldsymbol{x} - \boldsymbol{y}\|$$

を最小にする$\boldsymbol{x}$は, $A^{(1,3)}\boldsymbol{y}$によって与えられます.

補題 3. $AA^{(1,3)}$は正射影行列になります.

簡単のため1,2,3-逆に関して図で説明します.



- $R(A^{(1,2,3)})$ と $R(A)$ のベクトルは1対1に対応します.
  $A^{(1,2,3)}A\boldsymbol{x}$ と $A\boldsymbol{x}$ が対応します.
- $\boldsymbol{x} + N(A)$ のベクトルはすべて $A\boldsymbol{x}$ に写像されます.
- $A\boldsymbol{x}+R(A)^{\perp}$ のベクトルはすべて $A^{(1,2,3)}A\boldsymbol{x}$ に写像されます.

8.4　1,4-逆

1,4-inverse: $R(A^{(1,4)}A) = N(A)^{\perp}$: 最小ノルム解

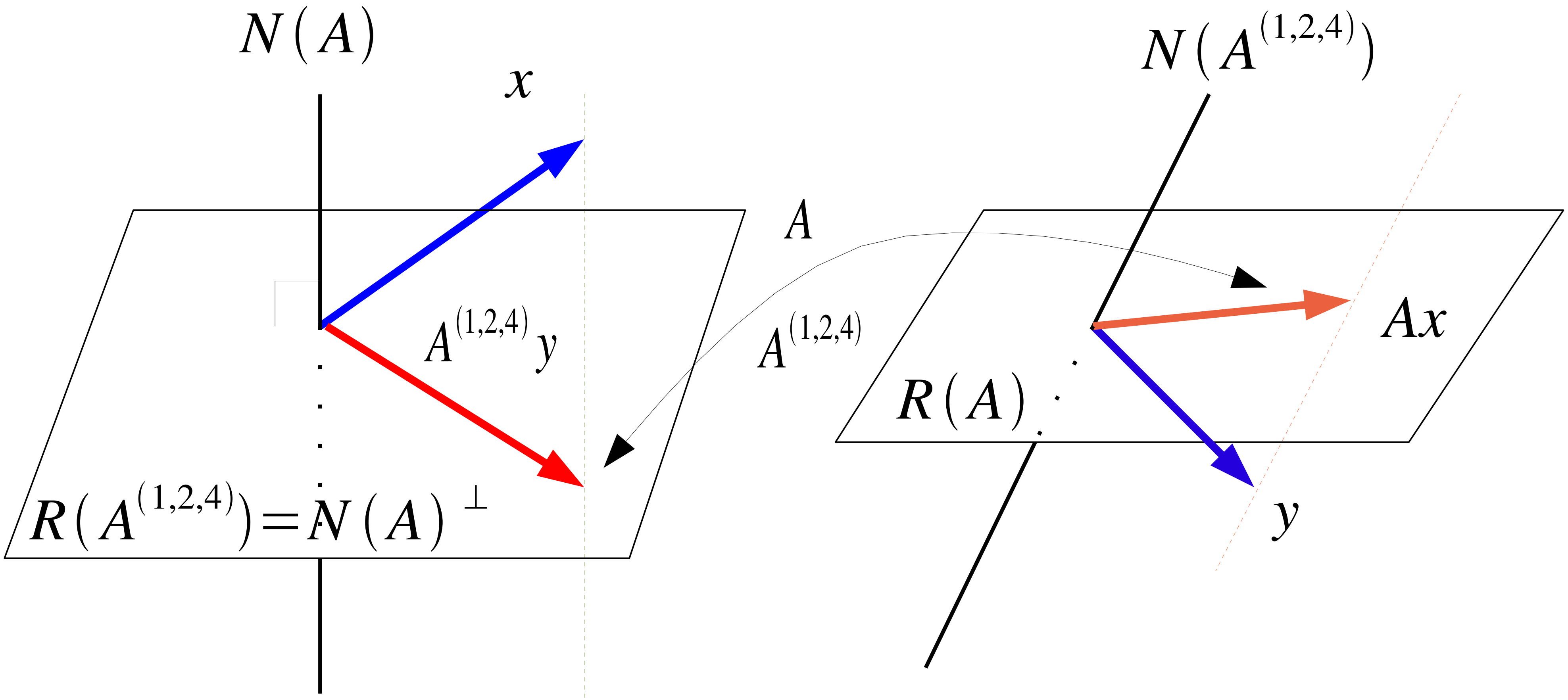
定理 2. $R(A)$ に含まれる $\boldsymbol{y}$ が与えられたとき，$\boldsymbol{y} = A\boldsymbol{z}$ を満たすものの中で,

$$\|\boldsymbol{x}\|$$

を最小にする $\boldsymbol{x}$ は，$A^{(1,4)}\boldsymbol{y}$ によって与えられます.

補題 4. $A^{(1,4)}A$ は正射影行列になります.

簡単のため 1,2,4-逆に関して図で説明します.



- $N(A)^{\perp}$ と $R(A)$ のベクトルは1対1に対応します.
  $A^{(1,2,4)}A\boldsymbol{x}$ と Ax が対応します.
- $\boldsymbol{x}+N(A)$ のベクトルはすべて $A\boldsymbol{x}$ に写像されます.
- $A\boldsymbol{x}+N(A^{(1,2,4)})$ のベクトルはすべて $\boldsymbol{x}$ の $N(A)^{\perp}$ への正射影へ写像されます.

8.5　1,2,3,4-逆

1,2,3,4-逆は，最小2乗解かつ最小ノルム解を与えます.

定理 3. 与えられた $\boldsymbol{y}$ に対して,

$$\|A\boldsymbol{x} - \boldsymbol{y}\|$$

を最小にする $\boldsymbol{x}$ の中で,

$$\|\boldsymbol{x}\|$$

を最小にする $\boldsymbol{x}$ は，$A^{\dagger}\boldsymbol{y}$ によって与えられます.

補題 5. $AA^{\dagger}$ と $A^{\dagger}A$ は，は正射影行列になります.

定理 4. $N(A^{\dagger}) = R(A)^{\perp}$, $R(A^{\dagger}) = N(A)^{\perp}$ が成立します.

問題 27. 上の定理を証明して下さい.
($N(A^{\dagger}) = N(A\ T)$, $R(A^{\dagger}) = R(A^{T})$ を示します.　)

定理 5. *1,2,3,4*-逆は，存在すれば一意です.

(証明)

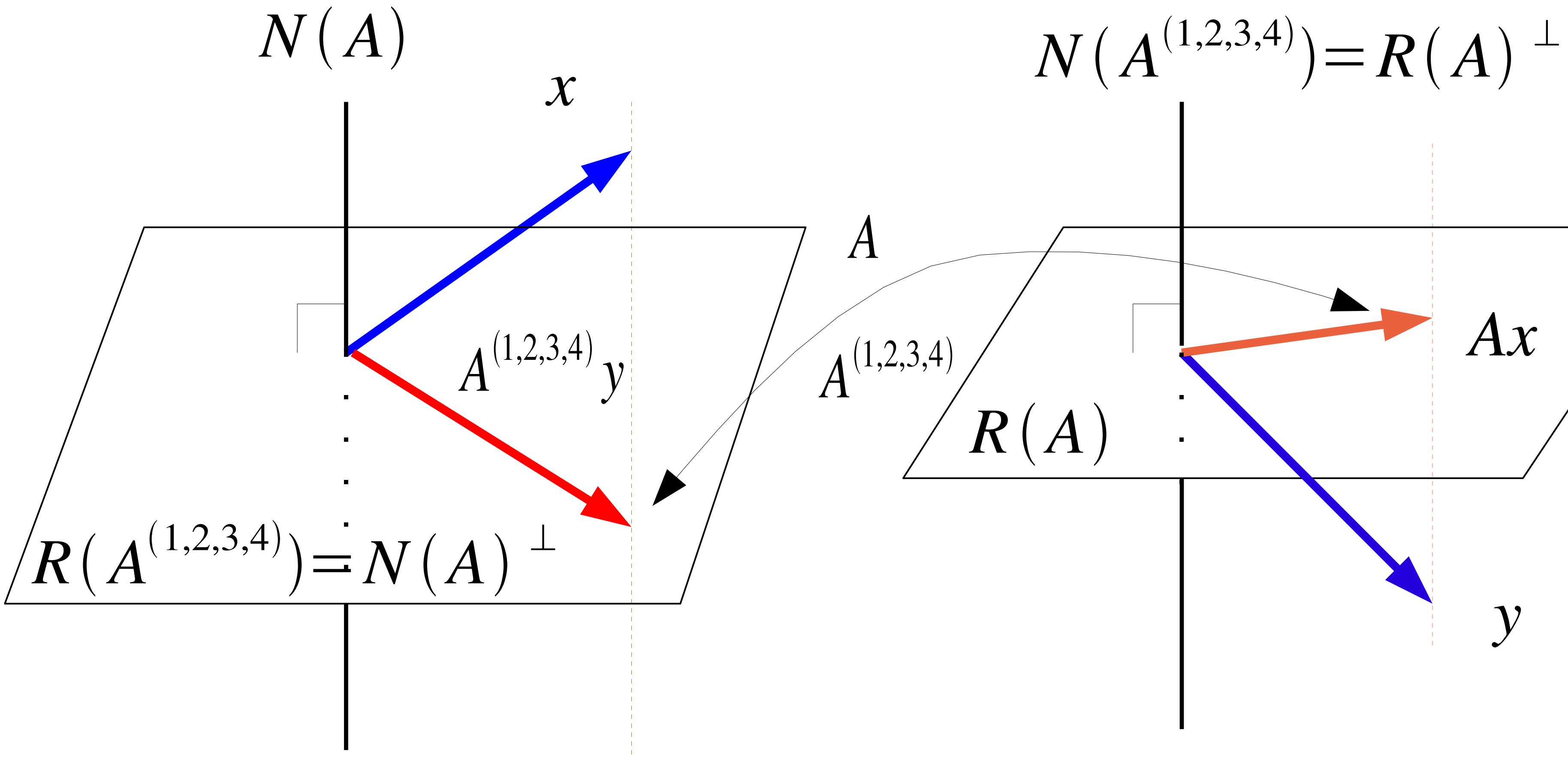
Bも，$ABA = A$，$BAB = B$，$(AB)^T = AB$，$(BA)^T = BA$を満たすものとします.

$$
\begin{aligned}
B &= BAB = BAA^\dagger AB = B(A^\dagger)^T A^T B^T A^T = B(A^\dagger)^T (ABA)^T \\
&= B(A^\dagger)^T A^T = B(AA^\dagger)^T = BAA^\dagger = BAA^\dagger AA^\dagger \\
&= A^T B^T A^T (A^\dagger)^T A^\dagger = A^T (A^\dagger)^T A^\dagger = (A^\dagger A)^T A^\dagger = A^\dagger AA^\dagger = A^\dagger.
\end{aligned}
$$

より，一意であることがわかります.

定理 6. 作用素の場合も含めて，一般逆は$R(A)$が閉集合ならば存在します.
有限部分空間は閉集合ですので，一般逆行列は必ず存在します.

証明には，Banachの可逆定理(Harn-Banachの定理から得られます)を使い，複雑ですので省略します.



- $N(A)^{\perp}$と$R(A)$のベクトルは1対1に対応します.
  $A^{\dagger}A\boldsymbol{x}$と$Ax$が対応します.
- $\boldsymbol{x}+N(A)$のベクトルはすべて$A\boldsymbol{x}$に写像されます.
- $A\boldsymbol{x}+R(A)^{\perp}$のベクトルはすべて$\boldsymbol{x}$の$N(A)^{\perp}$への正射影へ写像されます.

8.6　一般逆行列の計算法

A^- : 掃き出し法，LU-分解

$A^\dagger$:

- 1-逆の利用 : $A^\dagger = A^T(AA^T)^- A(A^TA)^- A^T$
- SVD:

$$A = \sum_{i=1}^{k} \mu_i u_i v_i^T$$

と分解し，$\mu_i \neq 0$ とすれば，以下の式が成立します.

$$A^\dagger = \sum_{i=1}^{k} \frac{1}{\mu_i} v_i u_i^T$$

- 正則化で近似的に計算できます.

$$A^\dagger = \lim_{\varepsilon \to 0} A^T(AA^T + \varepsilon I)^{-1} = \lim_{\varepsilon \to 0} (A^TA + \varepsilon I)^{-1} A^T$$

ε を小さくしていくと，$A^\dagger$ に収束する.

8.7 $N(A) = N(A^TA)$

補題 6. 次式が成立します.

$$N(A^TA) = N(A), \tag{78}$$
$$R(AA^T) = R(A). \tag{79}$$

(証明)

$N(A^TA) \supset N(A)$ は明らかです．$\boldsymbol{x} \in N(A^TA)$ とすれば,

$$\|A\boldsymbol{x}\|^2 = \langle A\boldsymbol{x}, A\boldsymbol{x}\rangle = \langle A^TA\boldsymbol{x}, \boldsymbol{x}\rangle = 0,$$

となり，$A\boldsymbol{x} = 0$ となります．従って，$N(A^TA) \subset N(A)$ となり，$N(A^TA) \supset N(A)$ は明らかですから式(78)が成立します.

A に A^T を代入すれば，$N(AA^T) = N(A^T)$ となりますので，補題1より,

$$R(A) = N(A^T)^\perp = N(AA^T)^\perp = R(AA^T).$$

が成立します.