New paradigm! On-demand publishing

著者：平鍋 健児
Kenji Hiranabe

図解 線形代数
ストラング流直感的理解

補遺

初版: 2026-4-1 更新: 2026-04-15

書籍『図解線形代数：ストラング流直感的理解』の付録として、本文中に（▶ 補遺）マークを付けたトピックについて、詳しく補足しました。本書をコンパクトにするために、数学的に詳細な証明を多くこの補遺に移動しています。
特に、ベクトル空間の基礎的な定理（基底の存在など）の証明部分で、抽象代数への有用な入り口になるものが含まれます。射影についても部分空間との双対性への橋渡しとなる部分、CR 分解による行空間と零空間の $\mathbb{R}^n$ の直交・直和分解の詳説、またシュアの定理の証明やケイリー・ハミルトンの定理の証明など、正方行列による固有値分解の理解に重要な内容も含まれます。さらに、特異値分解による行列の近似に関するエカート・ヤングの定理の証明なども含んでいます。特に定理の証明では、知られているものの中から、なるべく直感的に理解しやすいものを選んで丁寧に説明しています。

X: @hiranabe, hiranabe@mail.com, https://anagileway.com

目 次

1　数ベクトルと行列

1.1　シュワルツの不等式の証明

性質 1.3(3) シュワルツ（Schwarz）の不等式（p.21）　ベクトル $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ について、次の性質が成り立つ。

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\| \quad \text{等号成立は } \boldsymbol{u} \text{ と } \boldsymbol{v} \text{ が平行のときのみ}$$

この不等式の証明は有名なものが複数あって技巧的ですが、私が気に入っている簡単なものを一つ紹介しましょう[1]。ベクトル $\boldsymbol{v}, \boldsymbol{u}$ と実数 α に対して $\|\boldsymbol{v} - \alpha\boldsymbol{u}\|^2$ という式を考え、これが任意の α に対して常に非負の実数になることを使います。与えられた $\boldsymbol{v}, \boldsymbol{u}$ からうまく α を調整することで、不等式を導きます。

Proof. まず、$\boldsymbol{u}, \boldsymbol{v}$ どちらかが $\boldsymbol{0}$ の時は両辺とも 0 になるので等号成立。次に $\boldsymbol{u}, \boldsymbol{v} \neq 0$ のとき、任意の α について、

$$0 \leq \|\boldsymbol{u} - \alpha\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - \alpha\boldsymbol{u} \cdot \boldsymbol{v} - \alpha\{\boldsymbol{u} \cdot \boldsymbol{v} - \alpha\|\boldsymbol{v}\|^2\}$$

が成立。α を上手に選んで、$\{\cdots\} = 0$ にする。すなわち、$\alpha = \boldsymbol{u} \cdot \boldsymbol{v}/\|\boldsymbol{v}\|^2$ とすれば、上式は、

$$0 \leq \|\boldsymbol{u}\|^2 - \alpha\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|^2 - \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|^2}\right)\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|^2 - \frac{|\boldsymbol{u} \cdot \boldsymbol{v}|^2}{\|\boldsymbol{v}\|^2}$$

となる。最後の不等式の両辺に $\|\boldsymbol{v}\|^2 (\neq 0)$ を掛けることによって、$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$ が証明された。また、等号は $\|\boldsymbol{u} - \alpha\boldsymbol{v}\|^2 = 0 \iff \boldsymbol{u} = \alpha\boldsymbol{v}$ すなわち、2 つのベクトルが平行のとき。なお、平行は定義 1.3（p.20）参照。どちらかが $\boldsymbol{0}$ の場合も含めて平行であるとき等号が成立する。 □

2　図で見る行列計算

この章には、対称行列と交代行列の特徴を追記しました。

2.1　対称行列と交代行列

行列同士には積が定義されていますが、一旦置いておき、$m \times n$ 行列の和とスカラー倍の演算のみに注目すると、行列もまたベクトル空間をなします。このことは、性質 1.5（p.30）で述べたように、行列の和とスカラー倍が行列の集合に閉じていることと、行列が従う演算規則から分かります。

特に、$n \times n$ の正方行列に限ってベクトル空間を考えることができます。さらに、その中でも対称行列（2.2、p.59）や交代行列（2.2、p.59）の集合もまた、それぞれ和とスカラー倍に関して閉じており、ベクトル空間をなします。すなわち、それらの集合は正方行列の集合がなすベクトル空間の部分空間（3.4、p.69）となります。いくつかの大切な性質を問として紹介します。

問　対称行列であり、交代行列でもあるような行列は存在するか？

解説　そのような行列を A とすると、$A = A^\mathsf{T}$ かつ $A = -A^\mathsf{T}$ が成り立つ。これにより、$A = -A$ となるので $2A = O$。したがって A は零行列 O である。確かに、$O^\mathsf{T} = O = -O$ なので、零行列は対称行列と交代行列の両方の定義に合致する。よって、答えは YES であり、その行列は零行列のみである。

[1] 参考文献 [15]、参考文献 [12] より。他にも 2 次方程式の判別式を用いる方法などが有名です。

問　どんな正方行列 A についても、$A + A^{\mathsf{T}}$ は対称行列になることを示せ。また、A と A^{T} を使って交代行列を作れ。

解説　$A + A^{\mathsf{T}}$ が対称行列であることは、転置をとってみれば分かる。

$$(A + A^{\mathsf{T}})^{\mathsf{T}} = A^{\mathsf{T}} + (A^{\mathsf{T}})^{\mathsf{T}} = A^{\mathsf{T}} + A = A + A^{\mathsf{T}}$$

となり、転置の前後が等しく、対称行列になる。計算には、転置の性質 1.6（p.32）を使っている。交代行列は、$A - A^{\mathsf{T}}$ で作れる。計算は同様。

問　どんな正方行列 A も、対称行列 S と交代行列 T の和に分解できること、さらにその分解は一意であることを示せ。これは、正方行列の集合が対称行列の集合と交代行列の集合に直和分解（定義 3.7、p.70）できることを意味する。

解説　前問での A から作られる対称行列と交代行列の和を考える。適当にスカラー倍で調整して、対称行列（S）と交代行列（T）の和が A に戻るようにする。

Proof. 勘を働かせて、

$$S = \frac{A + A^{\mathsf{T}}}{2}, \quad T = \frac{A - A^{\mathsf{T}}}{2}$$

とすれば、

$$A = S + T, \quad S^{\mathsf{T}} = S \text{（対称）}, \quad T^{\mathsf{T}} = -T \text{（交代）}$$

となる。この分解の一意性は、仮に $A = S' + T'$ と別の対称行列と交代行列に分解できたとして、$S' = S, T' = T$ となることを示せばよい。

$$S' + T' = S + T \quad \text{より} \quad S' - S = T - T'$$

ここで、左辺の $S' - S$ は対称行列同士の差なので対称、右辺の $T - T'$ は交代行列同士の差なので交代である。前問より、対称行列でもあり交代行列でもある行列は零行列しかないので、両辺は O に等しく、$S' = S, T' = T$ となる。よって、この分解は一意である。 □

この証明の最後も、技巧的に感じると思われます。線形結合を「同じ特徴のもの同士になるよう等号の左右に割り振る」という証明戦略は他の証明でもでよく使われます。こうすることで、もしその特徴が和とスカラー倍によって保存されるのであれば、両辺はその特徴を持つ集合の共通部分に属することになります。この問題では、共通部分が零行列しかないことを示し、分解の一意性が示されます。

まとめ 2.1　正方行列全体の集合は、ベクトル空間をなす。さらに、

(1) 対称行列の集合は、正方行列全体のなすベクトル空間の部分空間である
(2) 交代行列の集合も、正方行列全体のなすベクトル空間の部分空間である
(3) 正方行列全体のベクトル空間は、対称行列全体の部分空間と交代行列全体の部分空間の直和である。すなわち、どんな正方行列も、対称行列と交代行列の和に一意に分解できる。

$$A = \frac{A + A^{\mathsf{T}}}{2} + \frac{A - A^{\mathsf{T}}}{2}$$

どんな行列も、対称行列と交代行列の和に一意に分解できる、という事実はとても興味深いですね。さらに、対称行列と交代行列がちょうど複素数の中の実数と純虚数のような関係にあるという議論につながっていますので、興味のある方は調べてみて下さい。例えば、対称行列の固有値はすべて実数であり、交代行列の固有値はすべて純虚数です（問 7.9、p.211）。

3　部分空間・線形変換

この章には、ベクトル空間についての多くの証明を加えました。特に、連立方程式の議論を使わずにこの章を説明しているので、抽象的な議論が多く、応用に向かう方には少々難しいかもしれません。ただし、その分妙味を感じる面白い証明も含まれています。特に、線形空間の次元が一意に決まることを示すのに使われる補題、「Steinitz の交換定理」の証明は、一読の価値があります。

3.1　直和の言い換え証明

定理 3.2 直和の言い替え（p.71）　W が U, V の直和であることと、W が U, V の和空間であって、さらに U, V の共通部分が $\{\mathbf{0}\}$ であることは同値。すなわち、

$$W = U \oplus V \iff \begin{array}{l} (1)\ W = U + V\ \text{（充満条件）、　かつ} \\ (2)\ U \cap V = \{\mathbf{0}\}\ \text{（排他条件）} \end{array}$$

直和（和空間であり一意の表現を持つ、定義 3.7、p.70）が、空間の充満条件かつ排他条件と同値であることを示します。

Proof. （$\Longrightarrow$）：$W = U \oplus V$ ならば、充満条件 $W = U + V$ は定義より。次に排他条件を示す。任意のベクトル $\boldsymbol{x} \in U \cap V$ を取ると、

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x} + \mathbf{0}, \quad \boldsymbol{x} \in U,\ \mathbf{0} \in V \\ \boldsymbol{x} &= \mathbf{0} + \boldsymbol{x}, \quad \mathbf{0} \in U,\ \boldsymbol{x} \in V \end{aligned}$$

と 2 通りに表現できるので、直和表現の一意性から、必ず $\boldsymbol{x} = \mathbf{0}$ となり、$U \cap V = \{\mathbf{0}\}$。
（$\Longleftarrow$）：あるベクトル $\boldsymbol{x} \in W = U + V$ が

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{u}' + \boldsymbol{v}', \quad \boldsymbol{u}, \boldsymbol{u}' \in U, \quad \boldsymbol{v}, \boldsymbol{v}' \in V$$

と 2 通りに表現できるならば、

$$\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{v}' - \boldsymbol{v}$$

が成り立つ。左辺は U に、右辺は V に属するベクトルなので、排他条件 $U \cap V = \{\mathbf{0}\}$ より $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{v}' - \boldsymbol{v} = \mathbf{0}$ が成り立ち、$\boldsymbol{u} = \boldsymbol{u}', \boldsymbol{v} = \boldsymbol{v}'$ となり、表現は一意である。すなわち、$W = U \oplus V$。　□

3.2　基底の延長定理の証明

定理 3.3 基底の延長（p.76）　$\mathbb{R}^n$ の部分空間 V に対して、V の線形独立なベクトルの集合 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ があるとき、これにいくつかの（0 も含む個数の）ベクトルを追加した V の基底 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}, \ldots, \boldsymbol{v}_q\}$, $(p \leq q)$ が存在する。

これを証明するために、まず、簡単な補題を示します。

補題 3.1 線形独立なベクトルの追加　$\mathbb{R}^n$ の線形独立なベクトルの集合 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ がある。この集合に線形独立な $\boldsymbol{v}_{k+1}$ を追加しても、その新たなベクトル集合は線形独立のままである。逆に、あるベクトル $\boldsymbol{v}_{k+1}$ を追加した新たなベクトル集合が線形独立のままであれば、$\boldsymbol{v}_{k+1}$ はもとのベクトル集合に線形独立である。

$$\boldsymbol{v}_{k+1} \notin \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle \iff \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}\ \text{は線形独立}$$

対偶を取って言い換えると、

$$\boldsymbol{v}_{k+1} \in \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle \iff \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\} \text{ は線形従属}$$

Proof. （ $\Longrightarrow$ ）線形関係式、

$$\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k + \alpha_{k+1} \boldsymbol{v}_{k+1} = \boldsymbol{0} \tag{*1}$$

を作る。$\boldsymbol{v}_{k+1}$ が集合に線形独立（定義 3.3、p.67）すなわち、

$$\boldsymbol{v}_{k+1} \notin \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$$

を仮定して、(*1) が自明な線形関係式（すべての係数が 0）であることを導く。

もし、新しく追加したベクトルの係数 $\alpha_{k+1} \neq 0$ なら、

$$\boldsymbol{v}_{k+1} = -\frac{\alpha_1}{\alpha_{k+1}} \boldsymbol{v}_1 - \cdots - \frac{\alpha_k}{\alpha_{k+1}} \boldsymbol{v}_k$$

となり仮定に矛盾。よって $\alpha_{k+1} = 0$。すると線形関係式 (*1) は

$$\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k = \boldsymbol{0}$$

となり、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ の線形独立性から

$$\alpha_1 = \cdots = \alpha_k = 0$$

合わせて全ての係数について $\alpha_i = 0 \ (i = 1, \ldots, k+1)$ となり、新たな $k+1$ 本のベクトル集合も線形独立である。

（ $\Longleftarrow$ ）新たな集合 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ が線形独立であるとする。もし、$\boldsymbol{v}_{k+1} \in \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$ ならば、$\boldsymbol{v}_{k+1}$ は $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ の線形結合で表現できるので、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ は線形従属であり矛盾。よって、$\boldsymbol{v}_{k+1} \notin \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$ でなければならない。 □

では、基底の延長定理を証明しましょう。ただし、部分空間 V は有限個のベクトルの集合によって生成されるとします[2]。

Proof. 出発点となる V の線形独立なベクトルの集合 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}, (1 \leq p)$ があるとする。もし、これが V を張れば、線形独立な生成系となり、すなわち基底。証明は終わり。そうでなければ、この集合で生成されないベクトルが V 内に存在する。そのベクトルを $\boldsymbol{v}_{p+1}$ とする。$\boldsymbol{v}_{p+1}$ は、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ では生成できないので、先の補題より、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}\}$ は線形独立なベクトルの集合になる。これが V を張れば、V の基底となる。このように、V を張るまで、線形独立なベクトルを追加し続けることができ、最終的に V の基底（線形独立な生成系）に到達する。 □

なお、**基底の削減定理**（定理 3.4、p.77）の方は明示的に定理としている教科書は少ないですが、対照的な関係なので本文中に言及しました。証明は、V を張る基底から初めて、線形従属なベクトルを取り除いていくことで、張る空間を維持したまま基底に到達できることが示せます。

3.3 基底は最大の独立系、最小の生成系の証明

[2]有限生成といいます。V が有限次元（次元についてはこの後で定義される）であっても、部分空間に対してこの仮定を外して基底の存在を示すことはさらに難しい証明になります（参考文献 [13]）。

定理 3.6 最大の独立系、最小の生成系（基底の特徴づけ）（p.78） $\mathbb{R}^n$ の部分空間 V に対して、ベクトルの集合が V の基底であることは、次の 2 つの条件どちらとも同値。

(1) V の最大の独立系（線形独立なベクトルの集合）である。この集合に任意のベクトルを 1 本足すと、線形従属になる

(2) V の最小の生成系（V を張るベクトルの集合）である。この集合から任意のベクトルを 1 本の取り除くと、V を張ることができなくなる

基底が 1 組存在するとして、その基底に 1 本のベクトルを追加すると線形従属になり、1 本のベクトルを削除すると生成系にならなくなります。ここで、最小・最大という言葉は「今手元にあるベクトルの集合」を対象にして、それに追加できるか、削除できるか、という意味です（極大・極小という言葉が正確でしょう）。

Proof. 基底 $\Longrightarrow$ (1) 最大の独立系： 基底にもう 1 本のベクトルを追加することを考える。そのベクトルは基底の生成性より必ず基底と線形従属なので、線形独立性が失われてしまう。よって基底は最大の独立系。

(1) 最大の独立系 $\Longrightarrow$ 基底： 最大の独立系が生成系でもあることを示す。最大の独立系を $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ とする。最大ということは、その集合に属さないどんな $\boldsymbol{x} \in V$ をこれに追加しても、独立系が壊れて線形従属になるということである。このことから、補題 3.1（p.6）の最下行の（ $\Longleftarrow$ ）より、$\boldsymbol{x} \in \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$ が導かれる。よって、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ は V を張る生成系である。独立系であり生成系であることが示されたので、基底。

(2) 最小の生成系 $\Longrightarrow$ 基底：最小の生成系が独立系であることを示す。もし、最小の生成系が線形従属であるとすると、少なくとも 1 つのベクトルは他のベクトルの線形結合なので、そのベクトルを取り除いても生成性を保つことができる（それが担っていた部分は、残りのベクトルの線形結合で担える）。これは最小性に反するので、最小の生成系は線形独立である。生成系であり独立系であることが示されたので、基底。

基底 $\Longrightarrow$ 最小の生成系 (2)：基底であれば生成系である。さらに、独立系でもあるので、基底にもう一本のベクトル（生成系なので必ず基底の線形結合）を加えると線形独立でなくなる。よって基底は最小の生成系である。 □

3.4 次元の一意性の証明

定理 3.7 次元（p.78） ベクトル空間 $\mathbb{R}^n$ の部分空間 V に対して、基底の要素数は一定である。この数を次元 $\dim V$ と呼ぶ。

証明するために準備として次の補題を示します。この定理は後述の連立方程式の議論からも導けますが、ここでは、1 つ面白い別の証明を紹介したいと思います。

補題 3.2 生成系は独立系より大きい ベクトル空間 $\mathbb{R}^n$ の部分空間 V の任意の生成系のベクトルの数は、任意の独立系のベクトル数よりも大きい。

この証明は、「Steinitz の交換定理」と呼ばれます[3]。

Proof. V の任意の生成系 $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p\}$ との任意の独立系 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_q\}$ をそれぞれ考え、$p \geq q$ を示す。 2 つのベクトル集合を、$|$ で区切って並べます（文字 $\boldsymbol{s}$ は span の意味）。やりたいことは、

[3] 参考文献 [12]。

左辺を生成系、右辺を独立系に保ったままでこのリストを変形すること。

$$\text{生成系} \mid \text{独立系}$$

$$\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p \mid \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_q$$

縦棒 | の左を左辺、右を右辺と呼ぶことにする。右辺には $\mathbf{0}$ が含まれないことに注意して、まず、ベクトル $\boldsymbol{v}_1 \neq \mathbf{0}$ を左辺の先頭に移動する。

$$\boldsymbol{v}_1, \boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p \mid \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_q$$

$\boldsymbol{s}_1, \ldots, \boldsymbol{s}_p$ は V を張るので、$\boldsymbol{v}_1$ はこれらの線形結合で表現できるはずである。その線形結合の係数の中には、0 でないものが必ず含まれるので、対応する i 番目の $\boldsymbol{s}_i$ を左辺から消し、さらに、$\boldsymbol{s}$ の添字を振り替えて、$\boldsymbol{s}_2, \ldots, \boldsymbol{s}_p$ とする。ここで、左辺に $\boldsymbol{v}$ を 1 本追加して $\boldsymbol{s}$ を一本削除したことになる。

$$\boldsymbol{v}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p \mid \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_q$$

左辺は生成系を維持し、右辺は線形独立性を維持していることに再度注意。次に $\boldsymbol{v}_2$ を左半分に移動する。

$$\boldsymbol{v}_2, \boldsymbol{v}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p \mid \boldsymbol{v}_3, \ldots, \boldsymbol{v}_q$$

$\boldsymbol{v}_2$ は元の左辺の線形結合で表現できるので、$\boldsymbol{v}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_p$ の線形結合で表現できる。その係数の中には、0 でないものが必ず含まれるが、「$\boldsymbol{s}_2, \ldots, \boldsymbol{s}_p$ の係数がすべて 0」にはならない。そんなことが起こると、$\boldsymbol{v}_1$ と $\boldsymbol{v}_2$ に非自明な線形関係式が成り立ってしまい、もともと線形独立であった $\boldsymbol{v}_1, \boldsymbol{v}_2$ は線形従属になってしまう。よって、今回左辺から消せるベクトル（線形結合の係数が非零になるベクトル）は、残りの $\boldsymbol{s}$ の中にあります。それを取り除き、$\boldsymbol{s}$ の添字を再度 3 から振りなおします。

$$\boldsymbol{v}_2, \boldsymbol{v}_1, \boldsymbol{s}_3, \ldots, \boldsymbol{s}_p \mid \boldsymbol{v}_3, \ldots, \boldsymbol{v}_q$$

ここまでで、左辺は p 本の生成系を維持し、右辺は 1 本ずつ減らして線形独立性を維持していることに注意。このように、$\boldsymbol{v}$ の 1 本を左辺に移動し、$\boldsymbol{s}$ を 1 本削除し、それらの係数を振り替えて行く。

このプロセスを続けていくと、左辺は p 本で、右辺の本数は $q, q-1, \ldots$ と減っていきます。このとき、右辺の $\boldsymbol{v}$ が尽きる前に左辺の $\boldsymbol{s}$ がすべて $\boldsymbol{v}$ で置き換わってしまったとすると、すなわち、左辺から $\boldsymbol{s}$ が消えてしまい、$\boldsymbol{v}$ ばかりになった形で両辺が存在すると不都合が生じる。左辺が生成系であることから、右辺の残りの $\boldsymbol{v}$ が、左辺の線形結合で表されてしまう。もともと、$\boldsymbol{v}$ は線形独立であったのですから矛盾が起こっている。よって、こんな事態は起こらず、必ず左辺に $\boldsymbol{s}$ を残した状態で**右辺の $\boldsymbol{v}$ が先に尽きる**。これで、$p \geq q$ が示された。 □

ここで、ようやく次元の一意性定理 3.7 の証明に入ります。基底の延長定理 3.3（p.76）によって、基底となるベクトルの集合が存在することが分かりました。後は、どのような基底の取り方をしても、その要素数が等しくなることを示します。

Proof. 2 つの基底を $(\boldsymbol{u}) = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_p\}$ と $(\boldsymbol{v}) = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_q\}$ が取れたとして、必ず $p = q$ となることを示せばよい。

$(\boldsymbol{u})$ も $(\boldsymbol{v})$ も基底なので、両方が同じ V の生成系かつ独立系。$(\boldsymbol{u})$ が生成系であり $(\boldsymbol{v})$ が独立系であることから、準備した補題 3.2（p.8）より、$p \geq q$ が成り立つ。また、役割を取り替えて、$(\boldsymbol{v})$ が独立系で $(\boldsymbol{u})$ が生成系であることから、$q \geq p$ が成り立つ。よって、$p = q$。 □

3.5 次元数本のベクトルと基底の証明

定理 3.8 次元数本のベクトルと基底（p.79）　$\mathbb{R}^n$ の部分空間 V の次元 $r = \dim V$ が分かっているとき、次元数 r 本の V に属するベクトルの集合について、以下の 3 条件は同値。
(1) V の基底である
(2) 独立系である（線形独立である）
(3) 生成系である（V を張る）

Proof. (1) $\Longrightarrow$ (2)(3)：基底であれば、独立系かつ生成系である。
(2) $\Longrightarrow$ (1)：V の次元が r なので、定理 3.6（p.78）より最大の独立系のベクトルは r 本。(2) 自身が r 本の独立系であることから基底である。
(3) $\Longrightarrow$ (1)：V 次元が r なので、同定理より最小の生成系のベクトルは r 本。(3) 自身が r 本の生成系であることから基底である。 □

3.6　基底が空間を直和分解することの証明

定理 3.9 基底と直和分解（p.79）　$\mathbb{R}^n$ の部分空間 V のベクトルの集合 $\{v_1, v_2, \ldots, v_r\}$ について、以下は同値。
(1) $\{v_1, v_2, \ldots, v_r\}$ は V の基底である。
(2) V は $\langle v_1 \rangle, \langle v_2 \rangle, \ldots, \langle v_r \rangle$ の直和である。すなわち、
$V = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus \cdots \oplus \langle v_r \rangle$

Proof. (1) $\Longrightarrow$ (2)：基底であれば、V を張るので、

$$\begin{aligned} V &= \langle v_1, v_2, \ldots, v_r \rangle \\ &= \langle v_1 \rangle + \langle v_2 \rangle + \cdots + \langle v_r \rangle \end{aligned}$$

となって、それぞれの部分空間の和集合となる。次に表現の一意性を示す。$\boldsymbol{x} \in V$ が 2 つの表現、

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_r\boldsymbol{v}_r = x'_1\boldsymbol{v}_1 + x'_2\boldsymbol{v}_2 + \cdots + x'_r\boldsymbol{v}_r$$

を持てば、基底は線形独立なので、

$$\begin{aligned} &(x_1 - x'_1)\boldsymbol{v}_1 + (x_2 - x'_2)\boldsymbol{v}_2 + \cdots + (x_r - x'_r)\boldsymbol{v}_r = \boldsymbol{0} \\ &\Longrightarrow\ x_1 = x'_1, x_2 = x'_2, \ldots, x_r = x'_r \end{aligned}$$

となる。よって、V は $\langle \boldsymbol{v}_1 \rangle, \langle \boldsymbol{v}_2 \rangle, \ldots, \langle \boldsymbol{v}_r \rangle$ の直和である。
(2) $\Longrightarrow$ (1)：V が $\langle \boldsymbol{v}_1 \rangle, \langle \boldsymbol{v}_2 \rangle, \ldots, \langle \boldsymbol{v}_r \rangle$ の直和であるとき、線形関係式 $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_r\boldsymbol{v}_r = \boldsymbol{0}$ の表現は一意であるので、各係数について、$x_1 = x_2 = \cdots = x_r = 0$。したがって、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ は線形独立である。さらに、$V = \langle \boldsymbol{v}_1 \rangle + \langle \boldsymbol{v}_2 \rangle + \cdots + \langle \boldsymbol{v}_r \rangle$ であるので、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ は V を張る。よって、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ は線型独立な生成系、すなわち V の基底である。 □

3.7　部分空間の包含関係と次元 (4) の証明

定理 3.10 部分空間の包含関係と次元 (4)（p.81）　$\mathbb{R}^n$ の部分空間 U, V の次元について、次が成り立つ。

$$\dim U + \dim V = \dim(U + V) + \dim(U \cap V)$$

特に、

$$\dim U + \dim V \geq \dim(U + V)$$

等号成立は、$\dim(U \cap V) = 0$ すなわち $U \cap V = \{\mathbf{0}\}$ のとき

参考文献 [15] を参考にした証明です。

Proof. $\dim U = m,\ \dim V = n,\ \dim(U \cap V) = k$ とし、

$$\begin{aligned} U \cap V \text{ の基底 } k \text{ 本：} \quad & \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\} \text{ を延長して、} \\ U \text{ の基底 } m \text{ 本：} \quad & \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_m\} \\ V \text{ の基底 } n \text{ 本：} \quad & \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\} \end{aligned}$$

とする。この基底を重複を省いて並べた $(m+n-k)$ 本のベクトルの集合を $U+V$ の基底の候補とする。

$$\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\}$$

これが $U+V$ を張ることは、任意の $\boldsymbol{x} \in U+V$ が、

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v}, \quad \boldsymbol{u} \in U, \boldsymbol{v} \in V$$

と書け、$\boldsymbol{u}, \boldsymbol{v}$ ともに基底候補の線形結合で書けることから明らかである。あとは、基底候補が線形独立であることを示せばよい。線形関係式を作る。

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{w}_i + \sum_{i=k+1}^{m} \alpha_i \boldsymbol{u}_i + \sum_{i=m+1}^{m+n-k} \alpha_i \boldsymbol{v}_{i-m+k} = \mathbf{0} \tag{*1}$$

最後の $\sum$ を右辺に移項して、

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{w}_i + \sum_{i=k+1}^{m} \alpha_i \boldsymbol{u}_i = -\sum_{i=m+1}^{m+n-k} \alpha_i \boldsymbol{v}_{i-m+k}$$

と書けば、左辺は U の要素であり右辺は V の元であるので、両辺とも $U \cap V$ の要素。よって $U \cap V$ の基底の線形結合で書ける。右辺をそのように記述すると、新しい係数 α' を使って、

$$-\sum_{i=m+1}^{m+n-k} \alpha_i \boldsymbol{v}_{i-m+k} = \sum_{i=1}^{k} \alpha'_k \boldsymbol{w}_i$$

と書ける。$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n$ は V の基底で線形独立なので、この式の係数はすべて 0 でなければならない。特に左辺の係数について、

$$\alpha_{m+1} = \cdots = \alpha_{m+n-k} = 0$$

がいえる。これをもとの線形関係式 (*1) に代入すると、

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{w}_i + \sum_{i=k+1}^{m} \alpha_i \boldsymbol{u}_i = \mathbf{0}$$

を得る。左辺は U の基底の線形結合なので、係数はすべて 0 でなければならない。以上より、

$$\alpha_1 = \cdots = \alpha_m = \cdots = \alpha_{m+n-k} = 0$$

がいえて、基底候補 $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\}$ は線形独立である。証明の先頭に示した生成系であること、さらに、線形独立であることがいえたので、正式に $U+V$ の基底であることが示された。次元は基底の要素数なので、

$$\dim(U+V) = m+n-k$$

である。すなわち、 $\dim U = m, \quad \dim V = n, \quad \dim(U \cap V) = k$ であったので、

$$\dim U + \dim V = \dim(U+V) + \dim(U \cap V)$$

□

3.8 基底の取り替え行列は可逆

定理 3.16 基底の取り替え行列は可逆（p.93） 基底の取り替え行列は可逆であり、その逆行列は逆方向の基底の取り替え行列である。

を証明します。

Proof. $(\boldsymbol{v}) = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ が基底なので、$(\boldsymbol{v}') = \{\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\}$ の各ベクトルはすべて $(\boldsymbol{v})$ の線形結合であり、逆に $(\boldsymbol{v}')$ が基底なので、$(\boldsymbol{v})$ の各ベクトルはすべて $(\boldsymbol{v}')$ の線形結合である。よって行列 P が存在して、

$$\begin{bmatrix} \boldsymbol{v}'_1 & \boldsymbol{v}'_2 & \cdots & \boldsymbol{v}'_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} P \tag{*1}$$

と書けると同時に、別の行列 Q が存在して、

$$\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}'_1 & \boldsymbol{v}'_2 & \cdots & \boldsymbol{v}'_n \end{bmatrix} Q \tag{*2}$$

とも書ける。両式が成り立つためには (*1) を (*2) に代入して、

$$\begin{aligned} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} &= \left(\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} P \right) Q \\ &= \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} (PQ) \\ &= \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} R \end{aligned} \tag{*3}$$

が成り立つ必要がある。ここで、$PQ = R = (r_{ij})$ と書いた。式 (*3) の両辺を 1 列ずつ見ていくと、例えば 1 列目は

$$\boldsymbol{v}_1 = r_{11}\boldsymbol{v}_1 + r_{21}\boldsymbol{v}_2 + \cdots + r_{n1}\boldsymbol{v}_n$$

となる。$(\boldsymbol{v}) = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ は線形独立なので、上記の線形結合の係数が $r_{11} = 1$、その他 $r_{21} = r_{31} = \cdots = r_{n1} = 0$ でなければならない（基底による表示の一意性、定理 3.9、p.79）。同様に、2 列目以降も調べると、R の各列が標準基底ベクトル（単位ベクトルの各列）になる。すなわち、$R = PQ = I$ となる。よって、P は可逆であり、さらに逆の取り替え行列は $Q = P^{-1}$ であることが分かる。 □

4 LU 分解と連立一次方程式

この章には補遺はありません。

5 CR 分解と 4 つの部分空間

この章には、CR 分解と 4 つの部分空間、特に行空間と零空間の直和関係に関する、例題を含む詳細を追記しています。

5.1 線形独立な列ベクトルの見分け方

5.1.4 節（p.123）で、CR 分解は理論上の分解法であり、CR 分解を実際に行うには、A から線形独立な列ベクトルを一つずつ見分けなければならないことを述べました。

この列ベクトルの独立性の見分け作業を、複雑な行列について手計算で行うには、結局ガウスの消去法や LU 分解を使うことになるでしょう。A を行階段形（REF）に変形し、そのピボット列番号

を見つけることによって行うことができます。ピボット列は、A の線形独立な列ベクトルに対応しています。ピボット列番号が $(1, 3, 4)$ であれば、C に入る A の列ベクトルは A の 1 列目、3 列目、4 列目になります。

この理由をあらためて書いておきます。行基本変形（定義 4.1、p.101）では、行の入れ替えやスカラー倍、行の加減を行いますが、この操作によって連立方程式の解は変わりません（これが「行基本変形」であり、「列基本変形」にはこの性質はありません)。行変形では A の零空間は変わらない、と言ってもいいでしょう。

理解するには、連立方程式を加減法で解く過程（すなわち行基本変形）を思い浮かべます。行基本変形は、連立方程式を変形して見た目を変えるだけで、解そのものは維持しています。$(x_1, x_2, \dots, x_n)$ のそれぞれが満たすべき関係は変わらないのです。

ガウス・ジョルダン法によって A が行簡約階段形（RREF）A_{rref} に到達したとき、A_{rref} の各列は「ピボット列」と「自由列」に分かれます（式 (4.5)、p.115)。C に集められる線形独立な列は、RREF のピボット列に対応しており、この列番号（x_i の変数番号 i）はオリジナルの A の線形独立な列番号に対応していることから、C に入る列ベクトルを A_{rref} のピボット列から遡って対応するオリジナルの A の列ベクトルを特定できるのです。

そして、R は A_{rref} から零行を取り除いたものに一致します。この一致は、RREF のピボットが必ず 1 であるということから、対応する R の成分が 1 となり、同じ列の残りの行成分は 0 になるためです。

大きな行列になると手作業はどちらにしても難しいので、計算機を使った LU 分解等で判定を行うことになるでしょう。

5.2 行空間と零空間による直和分解の証明

定理 5.8 行空間と零空間による直和分解（p.133） $m \times n$ 行列 A によって、$\mathbb{R}^n$ は 2 つの直交する部分空間、行空間 $\mathbf{C}(A^\mathsf{T})$ と零空間 $\mathbf{N}(A)$ によって直和分解される。

$$\mathbb{R}^n = \mathbf{C}(A^\mathsf{T}) \oplus \mathbf{N}(A), \quad \mathbf{C}(A^\mathsf{T}) \perp \mathbf{N}(A)$$

すなわち、任意のベクトル $\boldsymbol{x} \in \mathbb{R}^n$ は

$$\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n, \quad \boldsymbol{x}_r \in \mathbf{C}(A^\mathsf{T}),\ \boldsymbol{x}_n \in \mathbf{N}(A),\ \boldsymbol{x}_r \perp \boldsymbol{x}_n$$

と一意に分解できる。

Proof. （排他条件）$\mathbf{N}(A), \mathbf{C}(A^\mathsf{T})$ は直交するので（定理 5.7、p.131）共通部分は $\{\mathbf{0}\}$ である。

（充満条件）$\mathbf{N}(A), \mathbf{C}(A^\mathsf{T})$ はともに $\mathbb{R}^n$ の部分空間なので、和空間 $W = \mathbf{C}(A^\mathsf{T}) + \mathbf{N}(A)$ も $\mathbb{R}^n$ の部分空間（$W \in \mathbb{R}^n$)。この和空間が $\mathbb{R}^n$ を充満すること（$W = \mathbb{R}^n$）を示せばよい。定理 3.10 の (4)（p.81）より、排他条件を使って $\mathbf{C}(A^\mathsf{T}) \cap \mathbf{N}(A) = \{\mathbf{0}\}$、すなわち共通部分の次元が 0 であること（定理 5.7、p.131）から、

$$\begin{aligned}\dim W &= \dim(\mathbf{C}(A^\mathsf{T}) + \mathbf{N}(A)) \\ &= \dim \mathbf{C}(A^\mathsf{T}) + \dim \mathbf{N}(A) \\ &= n\end{aligned}$$

最終行は、本文中に示した議論で、行空間の次元を n から引いた自由度が零空間の次元となることから。以上より、W は $\mathbb{R}^n$ の部分空間であり、$\dim W = n$。定理 3.10 の (2)（p.81）より、$W = \mathbb{R}^n$ となる。 □

5.3　零空間・行空間と CR 分解

本文中に書ききれなかった、零空間と行空間に関する補足説明をしておきます。5.2 の零空間と行空間の直和分解について、CR 分解の例を使って説明します。この補遺例題を使って進めます。

補遺例題 5.1 零空間と CR 分解　次の行列 A の零空間（$Ax = \mathbf{0}$ の解空間）を CR 分解を用いて求めよ。

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 4 & 1 & 5 & 6 \end{bmatrix}$$

今回の問題では、少し横長の行列 A を使った問題にしています。横長の方が零空間が大きくなる傾向にあるので、本文で行き届かなかった説明を補いたいと思います。

5.3.1　CR 分解と行空間・零空間

では、CR 分解をしてみます。左から列ベクトルを取り出して、線形独立なものを選びます。1 列目と 2 列目を選ぶと、次のような行列 C ができます。R の 1 列目と 2 列目に単位行列ができます。A の 3 列目と 4 列目は、1 列目と 2 列目の線形結合で表せるので、その線形結合の係数を R の 3 列目と 4 列目に書き入れます（グレー部分）。

$$A = CR$$
$$\begin{bmatrix} 2 & 1 & 3 & 4 \\ 4 & 1 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

復習すると、3,4 列目は、

$$\begin{bmatrix} 3 \\ 5 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 6 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

でした。この結果から、A のランクは 2 であり、行空間と列空間の次元はともに 2 であることが分かります（節 5.1、p.118 からの議論です）。

ここでは、R に注目します。C の列が線形独立なので、

$$A\boldsymbol{x} = \mathbf{0} \iff R\boldsymbol{x} = \mathbf{0}$$

であることは、定理 3.14（p.82）で示しました。零空間は、以下の連立方程式の解空間です。

$$A\boldsymbol{x} = \mathbf{0} \iff \begin{cases} 2x_1 + x_2 + 3x_3 + 4x_4 = 0 \\ 4x_1 + x_2 + 5x_3 + 6x_4 = 0 \end{cases}$$

この行列 A は次のように CR 分解できたので、$R\boldsymbol{x} = \mathbf{0}$ を解きます。

$$A = CR$$
$$\begin{bmatrix} 2 & 1 & 3 & 4 \\ 4 & 1 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

これは、前述の連立方程式を加減法で変形したものと同じ結果です。

$$R\boldsymbol{x} = \mathbf{0} \iff \begin{cases} x_1 \qquad + x_3 + \ \ x_4 = 0 \\ \qquad x_2 + x_3 + 2x_4 = 0 \end{cases}$$

$$R = \begin{bmatrix} 1 & 0 & & 1 & 1 \\ 0 & 1 & & 1 & 2 \end{bmatrix}$$
$$\underbrace{\qquad\qquad}_{\substack{\text{ピボット列} \\ (x_1, x_2)}} \quad \underbrace{\qquad\qquad}_{\substack{\text{自由列} \\ (x_3, x_4)}}$$

R の左半分には、単位行列 I があります。よって、この行列のランクは 2 です。そして、この式には自由変数が 2 つあります。次の式を見ると分かるでしょう。

$$\begin{cases} x_1 \qquad = -x_3 - x_4 \\ \qquad x_2 = -x_3 - 2x_4 \end{cases}$$

この連立方程式は、変数 x_3 と x_4 を自由変数として自由な値を入れ、ピボット変数 x_1 と x_2 について必ず解くことができます。$(x_3, x_4) = ($ 1,0 $)$, $(x_3, x_4) = ($ 0,1 $)$ として求めた 2 つの零空間特殊解を、$\boldsymbol{s}_1 = (-1, -1, 1, 0)$, $\boldsymbol{s}_2 = (-1, -2, 0, 1)$ とすると、零空間の一般解が、$\boldsymbol{x} = x_3\boldsymbol{s}_1 + x_4\boldsymbol{s}_2$ と求まります。

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{0} \longrightarrow \boldsymbol{s}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{s}_2 = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{2 つのピボット変数} \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{2 つの自由変数} \end{matrix}$$

$$\longrightarrow \text{一般解} \quad \boldsymbol{x} = x_3\boldsymbol{s}_1 + x_4\boldsymbol{s}_2$$

x_1, x_2 の値は、それぞれ R の第 3 列と第 4 列を符号反転したものになっていることに注意してください。零空間の次元は、自由変数（x_3, x_4）の数で 2 になります。

より一般的には、R 行列のランクを r とすると、R 左端に r 次の単位行列 I_r があるので、この連立方程式は $n-r$ 個の自由変数にどんな値を入れても必ず解け、零空間の次元は $n-r$ になるのです。

5.3.2 零空間の一般化解法

この求解を一般化して、零空間特殊解 $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots$ が直接求められないか考えてみましょう。零空間の基底を作りたいのです。

問題を一般的に書くと、$m \times n$ 行列 A のランクを r として、$r \times n$ 行列 $R = [I_r \ F]$ と、左側に A の線形独立な列を表す r 次単位行列 I_r、右側に線形従属な列からなる $r \times (n-r)$ 係数行列 F を並べた形になります[4]。C の列が線形独立なので、定理 3.14（p.82）から、

$$A\boldsymbol{x} = \mathbf{0} \iff R\boldsymbol{x} = \mathbf{0} \iff \begin{bmatrix} I_r & F \end{bmatrix} \boldsymbol{x} = \mathbf{0}$$

が得られ、これを $\boldsymbol{x}$ について解いて零空間特殊解 $\boldsymbol{x} = \boldsymbol{s}_1, \boldsymbol{s}_2, \ldots$ を求めます。自由変数のうち 1 つを 1 に、他を 0 にして全ピボット変数について解けばよいのでした。

$F = [\boldsymbol{f}_1 \ \boldsymbol{f}_2 \ \cdots \ \boldsymbol{f}_{n-r}]$ とします。まず最初の自由変数の値を 1 として他の自由変数の値を 0 にし、ピボット変数について解いたものを $\boldsymbol{s}_1$ とすると、$\boldsymbol{s}_1$ の最初の r 成分には、$-\boldsymbol{f}_1$ が現れ、その後に

[4] A の列のうち、線形独立なものが左側に集まっていて残りが線形従属に並んでいる例です。より一般的な議論は後述。

$(1,0,\ldots)$ が続きます。また、$\boldsymbol{s}_2$ の最初の r 成分には $-\boldsymbol{f}_2$ が現れ、その後に $(0,1,0,\ldots)$ が続きます。このようにして、$n-r$ 本の零空間特殊解が得られます。

$$\boldsymbol{s}_1=\begin{bmatrix}-\boldsymbol{f}_1\\ 1\\ 0\\ \vdots\\ \\ 0\end{bmatrix},\quad \boldsymbol{s}_2=\begin{bmatrix}-\boldsymbol{f}_2\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix},\cdots,\quad \boldsymbol{s}_{n-r}=\begin{bmatrix}-\boldsymbol{f}_{n-r}\\ 0\\ \vdots\\ \\ 0\\ 1\end{bmatrix}$$

これらが線形独立であることは、$\boldsymbol{s}$ を並べた行列の下半分に単位行列 I_{n-r} が現れることから、定理 3.11（p.81）によって示されます。

また、これらの列ベクトルは、すべての $R\boldsymbol{s}=\boldsymbol{0}$ を満たすこと、すなわち零空間に属することも確認できます。そして、どんな自由変数 $(x_{r+1},\ldots,x_n)$ の値の組み合わせでも、零空間のベクトル $\boldsymbol{x}$ は、$n-r$ 個の自由変数 $x_{r+1},\ldots,x_n$ をパラメータとして、

$$\boldsymbol{x}=x_{r+1}\boldsymbol{s}_1+x_{r+2}\boldsymbol{s}_2+\cdots+x_n\boldsymbol{s}_{n-r}$$

と書くことができるのです。

零空間特殊解である $(n-r)$ 本の列ベクトル $\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_{n-r}$ は、零空間の線形独立な生成系、すなわち基底になり、零空間の次元は $n-r$ となります。すなわち、零空間と行空間の次元の和は、入力側 $\mathbb{R}^n$ 全体の次元 n すなわち行列 A の列数となるのです（定理 5.7、p.131 の証明の再説明）。

$$\dim\mathbf{N}(A)+\dim\mathbf{C}(A^{\mathsf{T}})=\dim\mathbb{R}^n=n$$

上記の $\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_{n-r}$ を並べた $n\times(n-r)$ 行列 X を考えます。すると上記の計算を、行列同士の演算として一気に書くことができます。

$$X=\begin{bmatrix}\boldsymbol{s}_1 & \boldsymbol{s}_2 & \cdots & \boldsymbol{s}_{n-r}\end{bmatrix}=\begin{bmatrix}-F\\ I_{n-r}\end{bmatrix}$$

とすると、F は $r\times(n-r)$ 行列であることを思い出して、

$$RX=\begin{bmatrix}I_r & F\end{bmatrix}\begin{bmatrix}-F\\ I_{n-r}\end{bmatrix}=-I_rF+FI_{n-r}=F-F=O$$

によって、$RX=O$、すなわち X の各列ベクトルは R の各行と直交します。今回の例では、

$$RX=\begin{bmatrix}I_r & F\end{bmatrix}\begin{bmatrix}-F\\ I_{n-r}\end{bmatrix}=\begin{bmatrix}1&0&1&1\\0&1&1&2\end{bmatrix}\begin{bmatrix}-1&-1\\-1&-2\\ \\ 1&0\\0&1\end{bmatrix}=O$$

と確認できます。つまり、X の列空間が A の零空間の基底です。

$$\begin{aligned}\mathbf{N}(A)=\langle\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_{n-r}\rangle=\mathbf{C}(X)\\ \dim\mathbf{N}(A)=\dim\mathbf{C}(X)=n-r\end{aligned}\tag{1}$$

この例で、行空間と零空間の基底をまとめておきます。

$$\begin{array}{c}\text{行空間の基底}\\ (R^{\mathsf{T}}\text{ の列ベクトル})\end{array}\begin{bmatrix}1\\0\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\\2\end{bmatrix}\qquad\begin{array}{c}\text{零空間の基底}\\ (X\text{ の列ベクトル})\end{array}\begin{bmatrix}-1\\-1\\1\\0\end{bmatrix},\begin{bmatrix}-1\\-2\\0\\1\end{bmatrix}$$

5.3.3　行空間と零空間による直和分解

行空間の基底（R の行ベクトルすなわち R^{T} の列ベクトル）と零空間の基底（X の列ベクトル）とが得られています。

$$R^{\mathsf{T}} = \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \cdots & \boldsymbol{r}_r \end{bmatrix} \in \mathbb{R}^{n \times r}$$
$$X = \begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \cdots & \boldsymbol{s}_{n-r} \end{bmatrix} \in \mathbb{R}^{n \times (n-r)}$$

各列ベクトルは n 次元であり、これらを並べて $\mathbb{R}^n$ 全体の基底が作れれば、それぞれの基底ベクトルの直和に全体が分解できるので、この直和分解が完成します。

$$\begin{aligned} \mathbb{R}^n &= \langle \boldsymbol{r}_1 \rangle \oplus \cdots \oplus \langle \boldsymbol{r}_r \rangle \oplus \langle \boldsymbol{s}_1 \rangle \oplus \cdots \oplus \langle \boldsymbol{s}_{n-r} \rangle \\ &= \underbrace{\langle \boldsymbol{r}_1, \ldots, \boldsymbol{r}_r \rangle}_{\text{行空間}} \oplus \underbrace{\langle \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{n-r} \rangle}_{\text{零空間}} \end{aligned}$$

定理 5.7（p.131）で示したように、行空間と零空間の共通部分は $\{\mathbf{0}\}$ のみであり、両者の次元の和は n です。

$$\mathbf{C}(A^{\mathsf{T}}) \cap \mathbf{N}(A) = \{\mathbf{0}\}$$
$$\dim \mathbf{C}(A^{\mathsf{T}}) + \dim \mathbf{N}(A) = \dim \mathbb{R}^n = n$$

この 2 つから、直和に必要な排他条件と充満条件が満たされていることが分かります（本補遺の 5.2 節、p.13 に書かれた証明です）。

5.4　検査と生成

本書にも書きましたが、ベクトル空間の部分空間を定義するには、検査的な方法と生成的な方法があります。

- 検査的定義：ベクトルが部分空間に属する**検査法**を用いて部分空間を示す。
  例：零空間　$\mathbf{N}(A) = \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \mathbf{0}\}$
- 生成的定義：部分空間のベクトルの**生成法**を用いて部分空間を示す。
  例：列空間　$\mathbf{C}(A) = \{A\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^n\}$

これらは、集合の定義の 2 つの仕方、すなわち、要素の条件を示す内包的定義と、要素を示す外延的定義に対応します。検査的定義の例が零空間や左零空間の定義であり、生成的定義の例が行空間や列空間の定義です。

列空間や行空間にあるベクトルのすべてを記述するには、生成的定義が有用です。すなわち、その行列の列や行の線形結合としての表現です。例題にある行列 A の列空間に属するベクトルは、すべて、$A\boldsymbol{x}$ もしくは $C\boldsymbol{x}$、行空間に属するベクトルは、$A^{\mathsf{T}}\boldsymbol{x}$ もしくは $R^{\mathsf{T}}\boldsymbol{x}$ と記述できます。

一方、あるベクトルが行空間に属するかどうかを調べるには、検査的定義が有用です。ある $\boldsymbol{x}$ を一つ持ってきて、それが A の行空間に属するかどうかを調べるには、どうすればよいでしょうか。本文のように、零空間の基底ベクトルが X（の列ベクトル）として求まっていれば、$X^{\mathsf{T}}\boldsymbol{x} = \mathbf{0}$ を計算して調べるだけでよいのです。逆に、あるベクトルが零空間に属するかどうかを調べるには、$A\boldsymbol{x}$ を計算して $\mathbf{0}$ になるかを調べるのです（これは、連立方程式 $A\boldsymbol{x} = \mathbf{0}$ を検査することになり零空間の定義そのものですね）。

問　$\boldsymbol{v} = (1,1,2,3)$ は補遺例題 5.1 （補遺の p.14）の行列 A の行空間に含まれるか。対称行列であり、交代行列でもあるような行列は存在するか？

解説　行空間の検査表示は、補遺の式 (1)（p.16）より、$\mathbf{C}(A^{\mathsf{T}}) = \mathbf{N}(X^{\mathsf{T}}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid X^{\mathsf{T}}\boldsymbol{x} = \mathbf{0}\}$ なので、$X^{\mathsf{T}}\boldsymbol{v} = \mathbf{0}$ を確かめればよい。

$$X^{\mathsf{T}}\boldsymbol{v} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1-1+2-0 \\ -1-2+0+3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

この例の $\boldsymbol{v}$ は、R の 1 行目と 2 行目の和なので含まれて当然ですね。

これを知っていると、面白い応用があります。列数の等しい行列 A（$m \times n$）、B（$m' \times n$）の零空間の共通部分は、A, B を縦にならべた新しい $(m+m') \times n$ 行列の零空間になります。

$$\begin{aligned} \mathbf{N}(A) \cap \mathbf{N}(B) &= \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \mathbf{0} \text{ かつ } B\boldsymbol{x} = \mathbf{0}\} \\ &= \{\boldsymbol{x} \in \mathbb{R}^n \mid \begin{bmatrix} A \\ B \end{bmatrix} \boldsymbol{x} = \mathbf{0}\} = \mathbf{N}(\begin{bmatrix} A \\ B \end{bmatrix}) \end{aligned}$$

これは、検査的定義の性質を利用したものです。$A\boldsymbol{x} = \mathbf{0}$ が表現する連立方程式 m 本と $B\boldsymbol{x} = \mathbf{0}$ が表現する連立方程式 m' 本を合わせたものを解くと考えることもできます。検査的に定義された 2 つの部分空間の共通部分を検査的に表示するのに便利です。

また、行数の等しい行列 A（$m \times n$）と B（$m \times n'$）の列空間の和空間は、A, B を横にならべた新しい $m \times (n+n')$ 行列の列空間になります。

$$\begin{aligned} \mathbf{C}(A) + \mathbf{C}(B) &= \{A\boldsymbol{x} + B\boldsymbol{y} \mid \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n\} \\ &= \{\begin{bmatrix} A & B \end{bmatrix} \boldsymbol{z} \mid \boldsymbol{z} \in \mathbb{R}^{n+n'}\} = \mathbf{C}(\begin{bmatrix} A & B \end{bmatrix}) \end{aligned}$$

これは、生成的定義の性質を利用したものです。列ベクトルを並列にならべてその線形結合を一度に作っています。生成的定義は、生成的に定義された 2 つの部分空間の和空間を生成的に表示するのに便利です。

問　次の A, B の両方の零空間に含まれる、$\boldsymbol{x}$ を求めよ。

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

解説　$A\boldsymbol{x} = \mathbf{0}, B\boldsymbol{x} = \mathbf{0}$ を満たす $\boldsymbol{x}$ を求めることになる。この式を変形すると、

$$\begin{bmatrix} A \\ B \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \boldsymbol{x} = \mathbf{0}$$

自由列に対応する x_4 が自由変数なので、1 とおいて解くと、零空間特殊解 $(1, -1, -1, 1)$ を得る。一般解は、x_4 をパラメータとして、$\boldsymbol{x} = x_4(1, -1, -1, 1)$ となる。

5.4.1　左零空間

あるベクトルが列空間に入っていることを調べるにはどうすればよいでしょうか。列空間 $C(A)$ は生成的に定義されましたが、検査的定義を考えることもできます。

ここで、これまで目立たなかった 4 つ目の部分空間、左零空間（定義 5.5、p.133）は、列空間の検査的定義として機能します。

ただ今回の例題 5.1（p.118）での左零空間は、$A^\mathsf{T}\boldsymbol{y}$ を求めると $\boldsymbol{y}=\boldsymbol{0}$ となり、

$$\mathbf{N}(A^\mathsf{T})=\{\boldsymbol{0}\},\quad \dim\mathbf{N}(A^\mathsf{T})=0$$

となります。すなわち、左零空間は、零ベクトルのみからなる空間となり、出力側 $\mathbb{R}^m$ は列空間のみによって充満されます（全ての検査はパスし、必ず方程式は解 $\boldsymbol{x}$ を持ちます)。この左零空間の役割については、本文 5.3 節にて、詳しく解説します。

5.5　列の順序と置換

本節の CR 分解の中では、議論を簡易にするため、A の線形独立な列ベクトルは先頭に、線形従属な列ベクトルは後ろに配置しています。実際はこのような順序で列ベクトルが並んでいるとは限らず、中間に線形従属なベクトルが挟まることもあるでしょう。すると、R の先頭にあった単位行列 I は、C の先頭ではなく、列ごとにばらばらに散らばってしまいます。

次のような A を考えます。2 列目に線形従属な列が現れ、C は A の 1 列目と 3 列目で構成されます。

$$\begin{aligned}A&=CR\\&=\begin{bmatrix}1&2&1&3\\2&4&0&4\end{bmatrix}=\begin{bmatrix}1&1\\2&0\end{bmatrix}\begin{bmatrix}1&2&0&2\\0&0&1&1\end{bmatrix}\end{aligned}$$

このような場合では、R の列交換を行うことで先頭に単位行列を作ることができます。列交換の操作は、行列の列を入れ替えるもので、置換行列 P_{ij}（定義 4.2、p.113）を右から掛けることで実現できます。

$$\begin{aligned}A&=\begin{bmatrix}1&1\\2&0\end{bmatrix}\begin{bmatrix}1&0&2&2\\0&1&0&1\end{bmatrix}P_{23}\\&=C\begin{bmatrix}I&F\end{bmatrix}P_{23}\end{aligned}$$

このように拡張することで、R の先頭に単位行列を移動します。$A\boldsymbol{x}=CR\boldsymbol{x}=\boldsymbol{0}\iff R\boldsymbol{x}=\boldsymbol{0}$ であり、$R=[I_r\ F]$ であるときの $R\boldsymbol{x}=\boldsymbol{0}$ の解を求める式に修正が入ります。列の入れ替えが必要な場合は、P を置換行列として $R=[I_r\ F]P$ となり、

$$\begin{aligned}R\boldsymbol{x}&=\begin{bmatrix}I_r&F\end{bmatrix}P\boldsymbol{x}=O\quad \text{の零空間特殊解は}\\X&=P^\mathsf{T}\begin{bmatrix}-F\\I_{n-r}\end{bmatrix}\quad \text{の列ベクトル}\end{aligned}$$

となります。P は置換行列なので $PP^\mathsf{T}=I$ を使えば、

$$\begin{aligned}RX&=(\begin{bmatrix}I_r&F\end{bmatrix}P)(P^\mathsf{T}\begin{bmatrix}-F\\I_{n-r}\end{bmatrix})\\&=\begin{bmatrix}I_r&F\end{bmatrix}(PP^\mathsf{T})(\begin{bmatrix}-F\\I_{n-r}\end{bmatrix}=\begin{bmatrix}I_r&F\end{bmatrix}\begin{bmatrix}-F\\I_{n-r}\end{bmatrix}=O\end{aligned}$$

が成り立ち、置換が必要な場合でも零空間特殊解 X の列ベクトルすべてが得られ、同様に議論を進めることができます。

5.6 写像における単射・全射とランク

A による線形変換を、写像 f として

$$\begin{array}{rcccll} f\colon & U & \longrightarrow & V & \ldots & U=\mathbb{R}^n,\ V=\mathbb{R}^m \\ & \cup & & \cup & & \\ & \boldsymbol{x} & \longmapsto & f(\boldsymbol{x}) & \ldots & f(\boldsymbol{x})=A\boldsymbol{x} \end{array}$$

と定義したとき、**単射**と**全射**は次のように定義されます。

- 単射（injection）：$f(\boldsymbol{x}_1)=f(\boldsymbol{x}_2) \implies \boldsymbol{x}_1=\boldsymbol{x}_2$
- 全射（surjection）：すべての $\boldsymbol{y}\in V$ について、$f(\boldsymbol{x})=\boldsymbol{y}$ となる $\boldsymbol{x}\in U$ が存在する
- 全単射（bijection）：単射かつ全射（逆写像が存在する）

一般の写像として $f(\boldsymbol{x})=\boldsymbol{b}$ を考えたとき、f が単射であれば、解の一意性が保証され、全射であれば、解の存在が保証されます。これは行列による**線形変換に限らない写像の性質**です。

$m\times n$ 行列 A のランクが r のとき、$\boldsymbol{x}\mapsto A\boldsymbol{x}$ を上記の写像 $f(\boldsymbol{x})$ として扱った線形代数では、5.2（p.147）で示したように、解の存在と一意性について、以下が同値になります。

表 1: 単射・全射と行列のランク・部分空間の関係

写像	**行列のランクと行数・列数**	**4 つの部分空間**
単射	$r=n$（列フルランク）	$\mathbf{N}(A)=\{\mathbf{0}\}$
全射	$r=m$（行フルランク）	$\mathbf{C}(A)=\mathbb{R}^m$
全単射	$r=m=n$（フルランク）	$\mathbf{N}(A)=\{\mathbf{0}\}$ かつ $\mathbf{C}(A)=\mathbb{R}^m$
		このとき $\mathbf{C}(A^\mathsf{T})=\mathbb{R}^n, \mathbf{N}(A^\mathsf{T})=\{\mathbf{0}\}$

5.7 行列の可逆性とランク

単射・全射と 4 つの部分空間の考え方を使うと、なぜ $m\times n$ 行列 A が正方行列でかつフルランク（$\operatorname{rank}A=n=m$）であるとき、そのときに限り A が可逆となるのかも明快に説明できます（式 (2.1)、p.62）。可逆（逆行列（A^{-1}）が存在する）とは、「$A\boldsymbol{x}=\boldsymbol{b}$ の**任意の**ベクトル $\boldsymbol{b}$ に対して、**一意な**解 $\boldsymbol{x}$ が存在する」ことと同値です。$r=\operatorname{rank}A$（列ランク）とすると、これまでの議論から、

- 解が必ず存在（全射）： 任意の $\boldsymbol{b}$ に対して解が存在するためには、A の列空間が出力側の空間全体を覆っている必要があります（$\mathbf{C}(A)=\mathbb{R}^m$）。これは、$A$ の列ランク $r=m$ であることを意味します。逆に、列空間が出力側の空間全体であれば、どんな $\boldsymbol{b}$ にも適当な $\boldsymbol{x}$ を選んで $A\boldsymbol{x}=\boldsymbol{b}$ とすることができます。
- 解が一意（単射）：解が一意であることと、零空間の次元が 0（$n-r=0$）、すなわち $r=n$ であることは同値です。

これら 2 つの条件を満たすことは、$r=m$ かつ $r=n$、すなわち $m=n=r$ であることと同値、すなわち、行列 A が正方行列（$m=n$）であり、かつ列フルランク（$r=n$）であることを意味します。また、列フルランクな正方行列は、行フルランクでもあることから、行列 A が行フルランクであることも同時に意味します。

定理 行列の可逆性とランク　行列 A について、以下は同値。

(1) A が可逆（逆行列 A^{-1} が存在）

(2) 任意の $\boldsymbol{b}$ に対して $A\boldsymbol{x} = \boldsymbol{b}$ が解を持ち、解は一意

(3) A が列および行フルランクな正方行列（列数 = 行数 = $\operatorname{rank} A$）

6　QR 分解と射影

射影について言い残したことを追記します。

6.1　A と $A^\mathsf{T}A$ のランクが等しいことの証明

定理 6.3 A と $A^\mathsf{T}A$ のランク（p.158）　$m \times n$ 行列 A の列ベクトルが線形独立であることと、$A^\mathsf{T}A$ の列ベクトルが線形独立であることは同値。このとき $A^\mathsf{T}A$ は n 次正方行列なので可逆であり、$(A^\mathsf{T}A)^{-1}$ が存在する。

さらに、列ベクトルの線形独立性を条件としない一般の長方行列 A についても、$\operatorname{rank} A = \operatorname{rank}(A^\mathsf{T}A)$ が成立する。

Proof. $A\boldsymbol{x} = \boldsymbol{0} \implies A^\mathsf{T}A\boldsymbol{x} = \boldsymbol{0}$ は明らか。次に、$A^\mathsf{T}A\boldsymbol{x} = \boldsymbol{0}$ の左から $\boldsymbol{x}^\mathsf{T}$ を掛けると、

$$\boldsymbol{x}^\mathsf{T}A^\mathsf{T}A\boldsymbol{x} = 0 \implies \|A\boldsymbol{x}\|^2 = 0 \implies A\boldsymbol{x} = \boldsymbol{0}$$

最後の式で内積の正値性（性質 1.2、p.20）を使っている。上記より、$A\boldsymbol{x} = \boldsymbol{0}$ と $A^\mathsf{T}A\boldsymbol{x} = \boldsymbol{0}$ は同値である。よって、A の列ベクトルと $A^\mathsf{T}A$ の列ベクトルの線形独立性が同値。

次に、A の列ランクの定義 5.1（p.118）より、A を CR 分解していく要領で A の列ベクトルの中で何本線形独立なものがあるか、を考える。何本取ってきても、その線形独立性の判定は、対応する $A^\mathsf{T}A$ を使っても同じであることが前の議論で分かったので、線形独立な列ベクトルの最大数は A と $A^\mathsf{T}A$ で同じである。すなわち、$\operatorname{rank} A = \operatorname{rank}(A^\mathsf{T}A)$ である。 □

$\boldsymbol{x}^\mathsf{T}$ を左から掛けるのは 2 乗を作るための少しトリッキーな操作です。
$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ が線形独立であれば、$A^\mathsf{T}A$ は可逆で、射影行列 $P = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}$ が安心して定義されます（定義 6.2, p.158）。

6.1.1　射影行列はほとんど不可逆

また、P はほとんどの場合不可逆であることに注意してください。実際、P は単位行列の場合以外は不可逆であり、単位行列の場合は P は $\boldsymbol{b}$ を動かしません。「何もしない」という特殊な射影を意味します（これも $P^2 = P, P^\mathsf{T} = P$ という射影の性質を満たします）。では、「この条件を満たす可逆な射影は I のみ」であることを示しておきます。

Proof. $\operatorname{rank} P = \operatorname{rank} QQ^\mathsf{T} = \operatorname{rank} Q = r \le m$（定理 6.3、p.158）なので、$P$ は m 次正方行列でありながらランク r の行列であるから、可逆であるには $m = r$ が必要条件。その場合、Q は正方行列かつ $Q^\mathsf{T}Q = I$ であるから、$Q^{-1} = Q^\mathsf{T}$ が導かれる。よって、$QQ^\mathsf{T} = I$ である。結果、$P = QQ^\mathsf{T} = I$ となり単位行列となる。 □

これは、一回射影してしまうと低い次もとの部分空間に押し込められてしまってもとに戻せないことから、可逆にならないことが直感的にも理解できます。

6.2 正規直交系は線形独立の証明

定理 6.4 正規直交系は線形独立（p.163） 正規直交系は線形独立である。

Proof. 正規直交系を $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_r\}$ として、線形関係式を作って両辺を $\boldsymbol{q}_j$ との内積を取る。$\boldsymbol{q}_j$ の係数が 0 になることが分かり、j を動かせばすべての係数が 0 であることが分かる。

$$\begin{aligned}
&\alpha_1\boldsymbol{q}_1 + \alpha_2\boldsymbol{q}_2 + \ldots + \alpha_r\boldsymbol{q}_r = \boldsymbol{0} \\
\Longrightarrow\ & (\alpha_1\boldsymbol{q}_1 + \alpha_2\boldsymbol{q}_2 + \cdots + \alpha_r\boldsymbol{q}_r) \cdot \boldsymbol{q}_j = 0 \\
\Longrightarrow\ & \alpha_j = 0 \quad (j = 1, 2, \ldots, r)
\end{aligned}$$

よって、$\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_r$ は線形独立。 □

「直交 $\Longrightarrow$ 独立」は覚えておくとよいでしょう（当然逆は成り立ちません）。

6.3 冪等な対称行列は射影行列

美しい射影の性質として、

$$\begin{aligned}
P^2 &= P \quad \text{冪等（idempotent）} \\
P^\mathsf{T} &= P \quad \text{対称（symmetric）}
\end{aligned}$$

の 2 つが射影を特徴づけることを示します。本書では射影といったら直交射影を指す（両方の性質を満たす）ことにしていますが、より一般には、冪等性のみで定義される斜交も含めた一般の射影があり、そこからさらに対称性を加えることで、直交射影が特徴づけられることになります。

では、冪等性（定理 6.4、p.176）と対称性を持つ P は、必ず P の列空間への直交射影になっていること、すなわち、定理 6.2（p.158）にある $P = QQ^\mathsf{T}$ となる列正規直交行列 Q を P から逆に構成して、その存在を示します。

Proof. $\mathbf{C}(P) = \mathbf{C}(Q)$ となるように、Gram-Schmidt 法を用いて $P = QR$ と正規直交化する。正方行列 P は一般的に不可逆行列なので、列正規直交行列 Q は r 本（$r = \operatorname{rank} P$）の列ベクトルからなる $m \times r$ の縦長行列になる。式 (6.10) で A を P と見て、

$$P = QR \tag{*1}$$

列正規直交行列 Q を転置して左から掛ければ、

$$Q^\mathsf{T}P = R \quad (\because Q^\mathsf{T}Q = I_r)$$

この R を式 (*1) に代入し、

$$P = Q(Q^\mathsf{T}P) \tag{*2}$$

となる。後は、上式で $Q^\mathsf{T}P = Q^\mathsf{T}$ となることを示せばよい。Q の列ベクトル $\boldsymbol{q}_i$ は $\mathbf{C}(P)$ 内にあるので、定理 6.4 から、$P\boldsymbol{q}_i = \boldsymbol{q}_i$ すなわち $PQ = Q$ なので、転置して、

$$\begin{aligned}
Q^\mathsf{T}P^\mathsf{T} &= Q^\mathsf{T} \\
Q^\mathsf{T}P &= Q^\mathsf{T} \quad (\because P^\mathsf{T} = P)
\end{aligned}$$

が得られた。これを前の式 (*2) に代入すると、

$$P = Q(Q^\mathsf{T}P) = QQ^\mathsf{T}$$

となり、P が定理 6.2（p.158）にある列正規直交行列 $Q = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r]$ の各列の張る空間への射影行列であることが示された。 □

これで、P の射影先の正規直交基底を求めるには、P を QR 分解して Q を得ればよいことが分かりました。ただし、ここで求めた Q は一意ではありません。Gram-Schmidt 法ではベクトルの並び順によって結果が異なることを思い出すと、P の列ベクトルの順番を変えれば、Q は順番だけでなく内容も変わります。それでも $P = QQ^\mathsf{T}$ という関係は変わりません。

実は、この 2 つの特徴から冪等性のみ（直交性を省く）で定義される一般の射影行列（冪等行列）からも、多くの面白い帰結が得られます。特に 1 の分解（後述）によって、P と $I - P$ が対になり、空間を 2 つの（直交とは限らない）部分空間の直和に分割することができます。

本書では解説しなかった、もう一つの線形代数のクライマックスとして、Jordan 分解と呼ばれる行列分解があります。そこで現れる $\mathbb{R}^n$ の「A の広義固有空間」への直和分解は、総和が I でどの 2 つも積が O になる冪等行列の集合を構成することによって、それぞれの射影先の空間として得られます。興味のある方は調べてみて下さい。本書で詳解した対称行列のスペクトル分解は、その分解が直交射影によって行われる特別なケースといえます。

6.4　$(I - P)$ は直交補空間への射影

性質 6.5（p.177）の (5)、P が射影なら $I - P$ も射影となること、そしてそれが直交補空間への射影であることを証明しましょう。

Proof. まず、$P^\perp = I - P$ が射影行列の定義を満たすことを示す。

$$\begin{aligned} (P^\perp)^2 &= (I-P)(I-P) = I - 2P + P = I - P = P^\perp \\ (P^\perp)^\mathsf{T} &= (I-P)^\mathsf{T} = I - P^\mathsf{T} = I - P = P^\perp \end{aligned}$$

冪等性と対称性が示せた。さらに、P と $P^\perp$ この 2 つ積が必ず O になること（直交性）を示す。

$$\begin{aligned} (P^\perp)P &= (I-P)P = P - P^2 = P - P = O \\ P(P^\perp) &= P(I-P) = P - P^2 = P - P = O \end{aligned} \tag{直交性}$$

そして、$P^\perp$ の定義より単位行列を分け合うことができます。これを 1 の分解と呼びます。

$$I = P + P^\perp \tag{1 の分解}$$

直交性から、それぞれの射影先の空間に存在する任意のベクトル $\boldsymbol{x}, \boldsymbol{y}$ は直交し、共通部分は $\{\boldsymbol{0}\}$。$\boldsymbol{x} \in \mathbf{C}(P), \boldsymbol{y} \in \mathbf{C}(P^\perp)$ とすると、適当なベクトル $\boldsymbol{x}', \boldsymbol{y}'$ を使って $\boldsymbol{x} = P\boldsymbol{x}', \boldsymbol{y} = P^\perp\boldsymbol{y}'$ と表せるので、

$$\begin{gathered} \boldsymbol{x} \cdot \boldsymbol{y} = (P\boldsymbol{x}')^\mathsf{T}(P^\perp\boldsymbol{y}') = \boldsymbol{x}'^\mathsf{T}(P^\mathsf{T}P^\perp)\boldsymbol{y}' = \boldsymbol{x}'^\mathsf{T}(PP^\perp)\boldsymbol{y}' = 0 \\ \boldsymbol{x} \in \mathbf{C}(P) \cap \mathbf{C}(P^\perp) \implies \boldsymbol{x} \cdot \boldsymbol{x} = \|\boldsymbol{x}\|^2 = 0 \implies \boldsymbol{x} = \boldsymbol{0} \end{gathered}$$

そして、次の **1 の分解**を使うと、$\mathbb{R}^m$ の任意のベクトル $\boldsymbol{z}$ は、$\boldsymbol{x} \in \mathbf{C}(P)$ と $\boldsymbol{y} \in \mathbf{C}(P^\perp)$ の和で表せる。任意の $\boldsymbol{z} \in \mathbb{R}^m$ について、

$$\begin{aligned} \boldsymbol{z} = I\boldsymbol{z} &= (P + P^\perp)\boldsymbol{z} = P\boldsymbol{z} + P^\perp\boldsymbol{z} \\ &= \boldsymbol{x} + \boldsymbol{y}, \quad \boldsymbol{x} \in \mathbf{C}(P),\ \boldsymbol{y} \in \mathbf{C}(P^\perp) \end{aligned}$$

以上より、$\mathbb{R}^m$ は必ず P の列空間と $P^\perp$ の列空間の和で表せ（充満条件）、かつ、共通部分は $\mathbf{0}$（排他条件）なので、$P^\perp$ の列空間は P の列空間の直交補空間（定義 5.6、p.138）。

$$\mathbb{R}^m = \mathbf{C}(P) \oplus \mathbf{C}(P^\perp), \quad \mathbf{C}(P^\perp) = \mathbf{C}(P)^\perp$$

□

7 固有値分解 $X\Lambda X^{-1}$

行列式、固有値と固有ベクトルについての補足をします。特にシュアの定理はなかなか手強いですが、一度は証明を読んでみることをおすすめします。

7.1 「積の行列式は行列式の積」の証明

性質 7.7(1) 行列の積の行列式（p.195） 行列 A, B の積 AB の行列式について、$\det(AB) = \det(A)\det(B)$ が成り立つ。

行列の特徴づけ（性質 7.1、p.190）を使った証明を示します。

Proof. $AB = A[\boldsymbol{b}_1\ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_n] = [A\boldsymbol{b}_1\ A\boldsymbol{b}_2\ \cdots\ A\boldsymbol{b}_n]$ である。A が不可逆であれば AB も不可逆であり $\det(AB) = 0 = \det(A)\det(B)$ が成り立つので、以下は A を可逆とする。A を定数のように考えて、B の関数 $F(B)$ を、

$$\begin{aligned} F(B) = F(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) &= \det(AB)/\det(A) &(2)\\ &= |A\boldsymbol{b}_1\ A\boldsymbol{b}_2\ \cdots\ A\boldsymbol{b}_n|/|A| &(3)\end{aligned}$$

と定義する。ここから、$F(B)$ が B の列に関して多重線形性、交代性がいえて、さらに $F(I) = 1$ がいえれば、$F(B) = \det(B)$ がいえ、$\det(AB)/\det(A) = \det(B)$ すなわち $\det(AB) = \det(A)\det(B)$ がいえる。

まず、$F(B)$ の j 列目に関する線形性は、次のように示される。

$$\begin{aligned} F(\ldots, \boldsymbol{b}_j + \boldsymbol{b}'_j, \ldots) &= |\cdots A(\boldsymbol{b}_j + \boldsymbol{b}'_j)\cdots|/|A| \\ &= (|\cdots A\boldsymbol{b}_j\cdots| + |\cdots A\boldsymbol{b}'_j\cdots|)/|A| \\ &= |\cdots A\boldsymbol{b}_j\cdots|/|A|\ +\ |\cdots A\boldsymbol{b}'_j\cdots|/|A| \\ &= F(\ldots, \boldsymbol{b}_j, \ldots) + F(\ldots, \boldsymbol{b}'_j, \ldots) &(\text{和})\end{aligned}$$

$$\begin{aligned} F(\ldots, c\boldsymbol{b}_j, \ldots) &= |\cdots c\boldsymbol{b}_j\cdots|/|A| \\ &= c|\cdots \boldsymbol{b}_j\cdots|/|A| \\ &= cF(\ldots, \boldsymbol{b}_j, \ldots) &(\text{スカラー倍})\end{aligned}$$

次に、$F(B)$ の交代性は j, k 列を入れ替えて、次のように示される。

$$\begin{aligned} F(\ldots, \boldsymbol{b}_k, \ldots, \boldsymbol{b}_j, \ldots) &= |\cdots A\boldsymbol{b}_k\cdots A\boldsymbol{b}_j\cdots|/|A| \\ &= -|\cdots A\boldsymbol{b}_j\cdots A\boldsymbol{b}_k\cdots|/|A| \\ &= -F(\ldots, \boldsymbol{b}_j, \ldots, \boldsymbol{b}_k, \ldots) &(\text{交代性})\end{aligned}$$

最後に、単位行列での値が 1 であることを確かめる。

$$\begin{aligned} F(I) &= F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) \\ &= |A\boldsymbol{e}_1 \cdots A\boldsymbol{e}_n|/|A| \\ &= |\boldsymbol{a}_1 \cdots \boldsymbol{a}_n|/|A| = |A|/|A| = 1 && \text{(単位行列の行列式)} \end{aligned}$$

これで、$F(B)$ が B の行列式としての性質 7.1（p.190）をすべて満たすことが示されたので、$F(B) = \det(B)$ がいえ、$\det(AB)/\det(A) = \det(B)$ すなわち $\det(AB) = \det(A)\det(B)$ が示された。 □

この証明は、少し不思議な感じがするかもしれませんが、行列式の特徴づけをうまく使っています。もう一つ、お気に入りの証明を紹介しておきます。こちらは、行基本変形と基本変形行列を使った証明で、直感的に分かりやすいものになっています。

Proof. (別証明) 行列 A を可逆とし（不可逆の場合は上と同じ証明)、行基本変形（ガウス・ジョルダン法）で簡約行階段形に変形することを考える。行基本変形は 3 種類あり、それぞれに対応する行基本変形行列 P_{ij}, $P_i(\alpha)(\alpha \neq 0)$, $P_{ij}(\alpha)$ がある（定義 4.1、p.101)。それぞれ行列は可逆であり、その逆行列も基本行列である。

$$\begin{aligned} P_{ij}: &\quad P_{ij}^{-1} = P_{ji} \\ P_i(\alpha): &\quad P_i(\alpha)^{-1} = P_i(1/\alpha)) \\ P_{ij}(\alpha): &\quad P_{ij}(\alpha)^{-1} = P_{ij}(-\alpha) \end{aligned}$$

A が可逆であれば、簡約行階段形は単位行列なので、A は行基本変形行列の積で表される (*1)。すなわち、行基本変形を k 回適用して A を単位行列に変形したとき、

$$P_k \cdots P_2 P_1 A = I \implies A = P_1^{-1} P_2^{-1} \cdots P_k^{-1} = \text{(基本変形行列の積)}$$

さらに、それぞれの基本変形行列 P はどんな B に対しても $\det(PB) = \det(P)\det(B)$ を満たす (*2)。

$$\begin{aligned} P_{ij}: &\quad \det(P_{ij}B) = -\det(B) = \det(P_{ij})\det(B) \\ P_i(\alpha): &\quad \det(P_i(\alpha)B) = \alpha\det(B) = \det(P_i(\alpha))\det(B) \\ P_{ij}(\alpha): &\quad \det(P_{ij}(\alpha)B) = \det(B) = \det(P_{ij}(\alpha))\det(B) \end{aligned}$$

(*1)(*2) より、$\det(AB) = \det(A)\det(B)$ が成り立つ。 □

7.2 固有値の和と積の証明

定理 7.4 固有値の和と積（p.204） n 次正方行列 A の固有値の和はトレースに等しく、積は行列式に等しい。すなわち、A の固有値を $\lambda_1, \lambda_2, \ldots, \lambda_n$ とすると、

$$\begin{aligned} \lambda_1 + \lambda_2 + \cdots + \lambda_n &= \mathrm{tr}(A) \\ \lambda_1 \lambda_2 \cdots \lambda_n &= \det(A) \end{aligned}$$

Proof. 固有多項式 $\phi_A(\lambda) = |\lambda I - A|$ を行列式の大公式を使って展開する。

$$|A| = \underbrace{\sum_{\sigma \in S_n}}_{n!\text{個の項の和}} \underbrace{\mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}}_{\text{行、列から 1 つずつ選択した } n \text{ 成分の積}}$$

より、

$$
\begin{aligned}
\phi_A(\lambda) = |\lambda I - A| &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \\
&= \lambda^n - \mathrm{tr}(A)\lambda^{n-1} + \cdots + (-1)^n|A|
\end{aligned}
$$

まず、λ の降ベキに並んだ先頭の 2 項のみを観察する。大公式は項数が $n!$ 個であり、Σ 中の各項（n 個の成分の積）は、各行 i と各列 j からそれぞれ重複ないように選んでいる。この計算ルールから、λ^n の項は $A - \lambda I$ の対角成分の積のみから現れる。λ^{n-1} の係数についても、n 個の成分の積のうち、$n-1$ 個の対角成分と 1 個の非対角成分を選ぶ、ということができないため、対角成分の積のみから現れる。対角成分の積を計算すると、λ^n の係数が 1、λ^{n-1} の係数が $-(a_{11} + a_{22} + \cdots + a_{nn}) = \mathrm{tr}(A)$ であることが分かる。

最後に、定数項は、$|\lambda I - A|$ の中で λ を含まない項であり、これは $\lambda = 0$ とすれば、$|-A| = (-1)^n|A|$ に等しいことも分かる。

一方で、代数学の基本定理（定理 7.2）から、固有多項式を一次式の積で次のように表現できる。

$$
\begin{aligned}
\phi_A(\lambda) &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\
&= \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n(\lambda_1\lambda_2\cdots\lambda_n)
\end{aligned}
$$

2 つの式が恒等的に等しいことから、係数を比較して、
$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(A), \quad \lambda_1\lambda_2\cdots\lambda_n = \det(A)$ が得られる。 □

7.3 シュアの定理の証明

定理 7.8 シュアの定理（三角化）（p.210） すべての n 次正方行列 A は三角化できる。すなわち、ある可逆行列 Q によって、

$$A = QTQ^{-1} \quad (T \text{ は上三角行列})$$

と分解できる。

すこし長いですが、帰納法による証明を示します。

Proof. n に関する帰納法で証明する。$n = 1$ のとき、$A_1 = [1][a][1]^*$ と分解できる。次に A_{n-1} が $n-1$ 次正方行列の場合に成り立つと仮定して、A_n が n 次正方行列の場合を示す。

A_n の固有値を重複を込めて $\lambda_1, \ldots, \lambda_n$ とする（必ず n 個とれるのは、代数学の基本定理 7.2、p.202 による）。λ_1 に対応する A_n の固有ベクトルを長さを調整して $\boldsymbol{q}_1 (\|\boldsymbol{q}_1\| = 1)$ とし、基底を延長して $\boldsymbol{q}_1, \boldsymbol{q}'_2 \ldots, \boldsymbol{q}'_n$ を $\mathbb{C}^n$ の正規直交基底とする（底の延長の定理 3.3（p.76）および QR 分解の定理 6.6（p.165）を使う）。ただし、直交の定義および内積は複素数の場合の定義 7.8（p.208））を使う。すると、$Q'_n = [\boldsymbol{q}_1\ \boldsymbol{q}'_2\ \cdots\ \boldsymbol{q}'_n]$ はユニタリ行列であり、

$$
\begin{aligned}
A_nQ'_n &= \begin{bmatrix} A_n\boldsymbol{q}_1 & A_n\boldsymbol{q}'_2 & \cdots & A_n\boldsymbol{q}'_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{q}_1 & A_n\boldsymbol{q}'_2 & \cdots & A_n\boldsymbol{q}'_n \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}'_2 & \cdots & \boldsymbol{q}'_n \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ \boldsymbol{0} & A_{n-1} \end{bmatrix} = Q'_n \begin{bmatrix} \lambda_1 & * \\ \boldsymbol{0} & A_{n-1} \end{bmatrix} \qquad (*1)
\end{aligned}
$$

と $n-1$ 次正方行列 A_{n-1} を使って書ける。A_{n-1} の固有値は A_n の固有値の残り $\lambda_2, \ldots, \lambda_n$ であり（余因子展開の式 (7.5)、p.196 より、$\phi_{A_n}(\lambda) = (\lambda - \lambda_1)\phi_{A_{n-1}}(\lambda)$)、帰納法の仮定からあるユニタリ行列 Q_{n-1} によって、

$$A_{n-1} = Q_{n-1} T_{n-1} Q_{n-1}^*, \quad T_{n-1} = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ O & & \lambda_n \end{bmatrix}$$

$$A_{n-1} Q_{n-1} = Q_{n-1} T_{n-1} \tag{*2}$$

とできる。ここから、Q_n' とこの Q_{n-1} を使ってうまく Q_n を構成する。

$$Q_n = Q_n' \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q_{n-1} \end{bmatrix} \tag{*3}$$

と定義すると、Q_n もユニタリ行列である（ユニタリ行列の積はユニタリ行列)。A_n とこの Q_n の積を計算すると、

$$\begin{aligned}
A_n Q_n &= A_n Q_n' \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q_{n-1} \end{bmatrix} && (\text{*3 より}) \\
&= Q_n' \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & A_{n-1} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q_{n-1} \end{bmatrix} && (\text{*1 より}) \\
&= Q_n' \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & A_{n-1} Q_{n-1} \end{bmatrix} \\
&= Q_n' \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & Q_{n-1} T_{n-1} \end{bmatrix} && (\text{*2 より}) \\
&= Q_n' \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q_{n-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & T_{n-1} \end{bmatrix} \\
&= Q_n \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & T_{n-1} \end{bmatrix} && (\text{*3 より})
\end{aligned}$$

となる。両辺に $Q_n^* = Q^{-1}$ をかけると、

$$A_n = Q_n \begin{bmatrix} \lambda_1 & * \\ 0 & T_{n-1} \end{bmatrix} Q_n^* = Q_n T_n Q_n^*, \quad T_n = \begin{bmatrix} \lambda_1 & * \\ 0 & T_{n-1} \end{bmatrix}$$

と上三角行列 T_n に三角化できる。$A = A_n, Q = Q_n, T = T_n$ とすれば、定理の主張が成り立つ。

なお、A の成分がすべて実数で、固有値もすべて実数ならば、この証明における正規直交基底や行列の成分はすべて実数にとれる。したがって、Q として直交行列をとれる。 □

この証明は、参考文献 [16] を参考にしました。

7.4 対称行列の対角化の証明

性質 7.12 対称行列の対角化（p.211） 実対称行列について以下が成り立つ。

(1) 固有値はすべて実数

(2) 異なる固有値に対応する固有ベクトルは直交する

(3) 直交行列によって対角化可能である

Proof. (1) 固有値が実数であることを示す。固有値 λ に対応する固有ベクトルを $\boldsymbol{x} \neq \boldsymbol{0}$ とする。$S\boldsymbol{x}$ と $\boldsymbol{x}$ の複素内積を 2 通り考える。

$$(S\boldsymbol{x}) \cdot (\boldsymbol{x}) = (\lambda \boldsymbol{x}) \cdot (\boldsymbol{x}) = \lambda(\boldsymbol{x} \cdot \boldsymbol{x})$$
$$(\boldsymbol{x}) \cdot (S\boldsymbol{x}) = (\boldsymbol{x}) \cdot (\lambda \boldsymbol{x}) = \overline{\lambda}(\boldsymbol{x} \cdot \boldsymbol{x})$$

この 2 つの式は対称行列の性質 $S^* = S$ と式 (7.8)（p.210）より等しいので、

$$\lambda(\boldsymbol{x} \cdot \boldsymbol{x}) = \overline{\lambda}(\boldsymbol{x} \cdot \boldsymbol{x}) \implies (\lambda - \overline{\lambda})(\boldsymbol{x} \cdot \boldsymbol{x}) = 0$$

$\boldsymbol{x} \neq \boldsymbol{0}$ より、$(\boldsymbol{x} \cdot \boldsymbol{x}) \neq 0$ である。したがって、$\lambda = \overline{\lambda}$ となり、λ は実数である。

(2) 異なる固有値に対応する固有ベクトルは直交することを示す。異なる固有値 λ_1, λ_2 に対応する固有ベクトルを $\boldsymbol{x}_1, \boldsymbol{x}_2 \neq \boldsymbol{0}$ とする。S を $\boldsymbol{x}_1, \boldsymbol{x}_2$ に掛けたの複素内積を 2 通り考える。

$$(S\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = (\lambda_1 \boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = \lambda_1(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2)$$
$$(\boldsymbol{x}_1 \cdot S\boldsymbol{x}_2) = (\boldsymbol{x}_1 \cdot \lambda_2 \boldsymbol{x}_2) = \overline{\lambda_2}(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2)$$

この 2 つの式は $S = S^*$ と式 (7.8) より等しいので、(1) より固有値が実数であることから $\lambda_2 = \overline{\lambda_2}$ である。

$$\lambda_1(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = \lambda_2(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) \implies (\lambda_1 - \lambda_2)(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = 0$$

$\lambda_1 \neq \lambda_2$ なので、$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 0$ が成り立つ。

(3) 対角化可能であることを示す。S を三角化（定理 7.8、p.210）する。直交行列 Q（$Q^{\mathsf{T}} = Q^{-1}$）を使って $T = Q^{\mathsf{T}}SQ$ と上三角行列 T にできる。これを転置すると、$S = S^{\mathsf{T}}$ より、

$$T^{\mathsf{T}} = (Q^{\mathsf{T}}SQ)^{\mathsf{T}} = Q^{\mathsf{T}}S^{\mathsf{T}}Q = Q^{\mathsf{T}}SQ = T$$

となり、T は対称行列でもある。対称な上三角行列は対角行列しかなく、したがって T は対角行列である。すなわち、S を三角化すれば、自動的に対角化される。 □

さらに、蛇足ながら (2) について 4 つの部分空間を使って直感的な面白い証明を加えます。4 つの部分空間の理解の助けになるでしょう。

Proof. (別証明) 2 つの異なる固有値を α, β とし、それぞれに対応する固有ベクトルを $\boldsymbol{x}, \boldsymbol{y}$ とする。仮に $\alpha = 0$ とすると、$\beta \neq 0$ であり、$\boldsymbol{x}$ は S の零空間に、$\boldsymbol{y}$ は列空間に属する。

$$S\boldsymbol{x} = \boldsymbol{0} \iff \boldsymbol{x} \in \mathbf{N}(S)$$
$$S\boldsymbol{y} = \beta \boldsymbol{y} \iff S\boldsymbol{y}/\beta = \boldsymbol{y} \iff \boldsymbol{y} \in \mathbf{C}(S)$$

ところが、$S = S^{\mathsf{T}}$ より行空間と列空間は等しく、行空間と零空間は直交する（定義 5.4、p.130）ので、$\boldsymbol{x}$ と $\boldsymbol{y}$ は直交。

$$\boldsymbol{x} \in \mathbf{N}(S), \quad \boldsymbol{y} \in \mathbf{C}(S) = \mathbf{C}(S^{\mathsf{T}}) \quad (\because S = S^{\mathsf{T}})$$
$$\implies \boldsymbol{x} \perp \boldsymbol{y} \quad (\because \mathbf{N}(S) \perp \mathbf{C}(S^{\mathsf{T}}))$$

次に、$\alpha \neq 0$ の場合、S をシフトして $T = S - \alpha I$ を考えると、T も対称行列であり、T の固有値は $0, (\beta - \alpha)$（性質 7.9、p.205)。固有ベクトルは S と T で同じ。よって、$\alpha = 0$ の場合と同じ議論で、対応する固有ベクトル $\boldsymbol{x}, \boldsymbol{y}$ が直交する。 □

どうです？対称性が列空間と行空間を等しくし、固有値のシフトをうまく使って、一方を零空間に移動することによって、行空間が零空間と直交するという構造を使っています。

7.5 固有値と固有ベクトルの順序の入れ替え

A を対角化したとき、固有値と対応する固有ベクトルの順序入れ替えても相似であることを証明します。

Proof. $A = X\Lambda X^{-1}$ とし、P を置換行列（定義 4.2、p.113）とすると、P は可逆であり、

$$\begin{aligned} A &= X\Lambda X^{-1} \\ &= X(PP^{-1})\Lambda(PP^{-1})X^{-1} \\ &= (XP)(P^{-1}\Lambda P)(XP)^{-1} \\ &= X'\Lambda' X'^{-1} \qquad \text{ここで、} X' = XP,\ \Lambda' = P^{-1}\Lambda P \end{aligned}$$

となる。$X' = XP$ は X の列を並び替えた（固有ベクトルの順序を変えた）行列であり、$\Lambda' = P^{-1}\Lambda P$ は Λ の対角成分を並び替えた行列である。これは、$A\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$ の添字 i を付け替えたことに対応する。 □

7.6 行列のランクと固有値の関係の証明

性質 7.14 行列のランクと固有値（p.216） n 次正方行列 A のランクは、零でない固有値の個数以上である。A が対角化可能の場合、ランクは零でない固有値の個数に等しい。

Proof. 対角化可能の場合、$A = X\Lambda X^{-1}$ と可逆行列 X、対角行列 Λ を使って表せる。Λ の対角成分として並ぶ固有値のうち、零でないものの個数が $\operatorname{rank}\Lambda$ である。また、可逆行列 X を左右から掛けてもランクは変わらないので、$\operatorname{rank} A = \operatorname{rank}\Lambda$ となる。

対角化可能でない場合、三角化して $A = XTX^{-1}$ と可逆行列 X、上三角行列 T を使って表せる。可逆行列 X を左右から掛けてもランクは変わらないので、$\operatorname{rank} A = \operatorname{rank} T$。$T$ の対角成分には A の固有値が並ぶ。T を行階段形として見ると、対角成分が 0 でない列の数は少なくとも列ランクを与えることが分かり、$\operatorname{rank} T$ は零でない固有値の個数以上である。 □

対角化可能な場合は分かりやすい結果ですが、そうでない場合が難問です。例えば、上三角行列で対角成分がすべて 0 の行列を考えると、固有値はすべて 0 ですが、ランクは 0 ではなく、非ゼロの上三角成分によってランクが決まることが分かります。ここで言えるのは、ランクは「零でない固有値の個数以上」ということだけです。

7.7 ケイリー・ハミルトンの定理の証明

定理 7.11 ケイリー・ハミルトン（Cayley-Hamilton）の定理（p.216） 正方行列 A は、その固有多項式 $\phi_A(\lambda)$ に対して、A を代入すると、

$$\phi_A(A) = O \quad \text{(零行列)}$$

を満たす。

Proof. A が対角化可能な場合は、$A = X\Lambda X^{-1}$ とすると、n 個の固有値が Λ の対角成分として並ぶ。

$$\phi_A(A) = \phi_A(X\Lambda X^{-1}) = X\phi_A(\Lambda)X^{-1}$$

より、$\phi_A(\Lambda) = O$ が示されればよい。

$$\begin{aligned}\phi_A(\Lambda) &= (\Lambda - \lambda_1 I)(\Lambda - \lambda_2 I)\cdots(\Lambda - \lambda_n I)\\ &= \begin{bmatrix} 0 & & & \\ & * & & \\ & & \ddots & \\ & & & * \end{bmatrix}\begin{bmatrix} * & & & \\ & 0 & & \\ & & \ddots & \\ & & & * \end{bmatrix}\cdots\begin{bmatrix} * & & & \\ & * & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}\end{aligned}$$

となり、各 $(\Lambda - \lambda_i I)$ は対角行列で i 番目の対角成分が 0 になっている。これらを掛け合わせると、対角成分がすべて 0 になるので、$\phi_A(\Lambda) = O$ が示された。 □

さらに、A が対角化可能でない場合も証明しておきます。

Proof. A が対角化できない場合、A を三角化して $A = XTX^{-1}$（T は上三角行列）とする。A と T は固有値を共有し、T の対角成分として並ぶ。

$$\phi_A(A) = \phi_A(XTX^{-1}) = X\phi_A(T)X^{-1}$$

より、$\phi_A(T) = O$ が示されればよい。

$$\begin{aligned}\phi_A(T) &= (T - \lambda_1 I)(T - \lambda_2 I)\cdots(T - \lambda_n I)\\ &= \begin{bmatrix} 0 & * & \cdots & * \\ & * & \cdots & * \\ & & \ddots & \vdots \\ & & & * \end{bmatrix}\begin{bmatrix} * & * & \cdots & * \\ & 0 & \cdots & * \\ & & \ddots & \vdots \\ & & & * \end{bmatrix}\cdots\begin{bmatrix} * & * & \cdots & * \\ & * & \cdots & * \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}\end{aligned}$$

各 $(T - \lambda_i I)$ は三角行列で i 番目の対角成分が 0 になっている。任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して、$\phi_A(T)\boldsymbol{x} = \boldsymbol{0}$ であることを観察する。上式の右端の $(T - \lambda_n I)$ は最下行がすべて 0 であることから、$(T - \lambda_n I)\boldsymbol{x}$ の最下成分は 0 である。さらに、結果に $(T - \lambda_{n-1} I)$ を掛けると、下から 2 番目の成分も 0 になる。

$$(T - \lambda_n I)\begin{bmatrix} * \\ \vdots \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \\ \vdots \\ * \\ 0 \end{bmatrix}, \quad (T - \lambda_{n-1} I)\begin{bmatrix} * \\ \vdots \\ * \\ 0 \end{bmatrix} = \begin{bmatrix} * \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

このように、$\boldsymbol{x}$ に順々に $(T - \lambda_n I), (T - \lambda_{n-1} I), (T - \lambda_{n-2} I), \ldots$ を掛けていくと下部に 0 が増えて

いく。

$$
\begin{aligned}
\phi_A(T)\boldsymbol{x} &= (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)\begin{bmatrix} * \\ \vdots \\ * \\ * \end{bmatrix} \\
&= (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_{n-1} I)\begin{bmatrix} * \\ \vdots \\ * \\ 0 \end{bmatrix} \\
&= (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_{n-2} I)\begin{bmatrix} * \\ \vdots \\ 0 \\ 0 \end{bmatrix} \\
&\vdots \\
&= (T-\lambda_1 I)\begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \boldsymbol{0}
\end{aligned}
$$

これで、任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して $\phi_A(T)\boldsymbol{x} = \boldsymbol{0}$ であること、すなわち、$\phi_A(T) = O$ が示された[5]（定理 2.1 の (2)、p.57）。 □

7.8 数列と微分方程式の書き換え

問題の先頭で変数変換を行って、数列と微分方程式の両方を固有ベクトル基底で表現します。

例題 7.3 数列と微分方程式の解（p.221） 次の数列 $x^{(n)}$ および微分方程式 $x(t)$ の一般解を求めよ。

$$
\begin{aligned}
x^{(n+2)} &= 3x^{(n+1)} - 2x^{(n)}, &\quad \text{初期値 } x^{(0)} = 1,\ x^{(1)} = 0 \\
\ddot{x}(t) &= 3\dot{x}(t) - 2x(t), &\quad \text{初期値 } x(0) = 1,\ \dot{x}(0) = 0
\end{aligned}
$$

図 3.2（p.92）より、固有ベクトル基底 $X = [\boldsymbol{x}_1\ \boldsymbol{x}_2]$ による座標を $\boldsymbol{y}^{(n)}, \boldsymbol{y}(t)$ とします。

$$
\begin{aligned}
\boldsymbol{y}^{(n)} &= X^{-1}\boldsymbol{x}^{(n)} \\
\boldsymbol{y}(t) &= X^{-1}\boldsymbol{x}(t)
\end{aligned}
$$

すると、数列と微分方程式の問題は、あらためて $\boldsymbol{c}$ を $\boldsymbol{y}^{(0)}$ および $\boldsymbol{y}(0)$ と書くと、

$$
\begin{aligned}
\boldsymbol{y}^n &= \Lambda^n \boldsymbol{y}^{(0)} \\
\boldsymbol{y}(t) &= e^{\Lambda t}\boldsymbol{y}(0)
\end{aligned}
$$

[5] この証明法は、参考文献 [18] から借りました。

となります。Λ が対角行列であることから、このベクトル計算を「1 行ずつ分解できる」こと、つまり $\boldsymbol{y}$ の各成分が独立に変換されることが、はっきり分かるでしょう。

$$\begin{array}{cc} \text{数列の場合} & \text{微分方程式の場合} \\ y_1^{(n)} = \lambda_1^n y_1^{(0)} & y_1(t) = e^{\lambda_1 t} y_1(0) \\ y_2^{(n)} = \lambda_2^n y_2^{(0)} & y_2(t) = e^{\lambda_2 t} y_2(0) \\ \vdots & \vdots \end{array}$$

7.8.1　初等的導出との対応

高校数学においてよく知られている解法と対応づけてみます。この例題では、以下の式変形によって x を y に置き換えて解いたことと同等になります。特性方程式と呼ばれる、$x^{(n+1)}, \dot{x}(t)$ を λ で、$x^{(n+2)}, \ddot{x}(t)$ を λ^2 で置き換えた式は、

$$\lambda^2 = 3\lambda - 2 \text{ より、} \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2) = 0$$

となり、これまで議論してきた行列の固有方程式と同じです。固有値 $\lambda_1 = 1, \lambda_2 = 2$ が得られます。数列の場合は、

$$\begin{aligned} y^{(n)} &= 2x^{(n)} - x^{(n+1)} \text{ と置くと、} \\ y^{(n+1)} &= \lambda_1 y^{(n)} = y^{(n)} \text{ となって項比 1 の等比数列なので、} \\ y^{(n)} &= y^{(0)} = 2x^{(0)} - x^{(1)} = 2 \end{aligned} \tag{*1}$$

$$\begin{aligned} y^{(n)} &= -x^{(n)} + x^{(n+1)} \text{ と置くと、} \\ y^{(n+1)} &= \lambda_2 y^{(n)} = 2y^{(n)} \text{ 項比 2 の等比数列なので、} \\ y^{(n)} &= 2^n y^{(0)} = 2^n(-x^{(0)} + x^{(1)}) = -2^n \end{aligned} \tag{*2}$$

(*1), (*2) を $x^{(n)}$ について解く。(*1)+ (*2) より、

$$x^{(n)} = 2 - 2^n$$

が得られます。$1^n = 1$ と 2^n の線形結合になっています。微分方程式の場合も同様に、

$$\begin{aligned} y(t) &= 2x(t) - \dot{x}(t) \text{ と置くと、} \\ \dot{y}(t) &= 1y(t) \text{ を解いて、} \\ y(t) &= y(0)e^t = (2x(0) - \dot{x}(0))e^t = 2e^t \end{aligned} \tag{**1}$$

$$\begin{aligned} y(t) &= -x(t) + \dot{x}(t) \text{ と置くと、} \\ \dot{y}(t) &= 2y(t) \text{ を解いて、} \\ y(t) &= y(0)e^{2t} = (-x(0) + \dot{x}(0))e^{2t} = -e^{2t} \end{aligned} \tag{**2}$$

この 2 式を $x(t)$ について解くと、(**1)+ (**2) より、

$$x(t) = 2e^t - e^{2t}$$

が得られます。e^t と e^{2t} の線形結合になっています。

8　特異値分解 $U\Sigma V^T$

特異値分解に関して、本文で省いた証明をいくつか示します。また、重要な「エカート・ヤングの定理」については、ぜひ読んでおいて下さい。直感を導く美しい定理です。

8.1　ランクと正の特異値の個数

定理 非零特異値の個数とランク（定理 8.1、p.232　行列 A の非零（正）の特異値の個数は、$\mathrm{rank}\,A$ に等しい。

の証明を示します。

Proof. 一般に、対角化可能な行列のランクが r であるとき、非零の固有値は重複を込めてちょうど r 個（性質 7.14、p.216）。$A^\top A$ は半正定値（対称行列）であるため対角化可能なので、$\mathrm{rank}\,A^\top A = \mathrm{rank}\,A$（定理 6.3、p.158）より、$A^\top A$ の非零の固有値は重複を込めてちょうど $r = \mathrm{rank}\,A$ 個で残りは 0 である。A の特異値は、$A^\top A$ の固有値の非負の平方根であるから、$\mathrm{rank}\,A$ 個の非零特異値（正）が存在し、残りの固有値はすべて 0 である。 □

8.2　AB と BA の非零固有値は等しいの証明

定理 AB と BA の非零固有値は等しい（定理 8.2、p.236）　$m \times n$ 行列 A と $n \times m$ 行列 B に対して、AB の非零固有値は、BA の非零固有値である。

の証明を示します。

Proof. AB の固有値 $\lambda \neq 0$ に対して、$AB\boldsymbol{v} = \lambda\boldsymbol{v}$ を満たす $\boldsymbol{v}(\neq \mathbf{0})$ が存在する。両辺に B を掛けると、$BA(B\boldsymbol{v}) = \lambda(B\boldsymbol{v})$ となり、$B\boldsymbol{v}$ は BA の固有値 λ に対応する固有ベクトルである。ただし、$B\boldsymbol{v} \neq \mathbf{0}$ であることを示す必要がある。もし $B\boldsymbol{v} = \mathbf{0}$ ならば、$\lambda\boldsymbol{v} = AB\boldsymbol{v} = A(B\boldsymbol{v}) = A\mathbf{0} = \mathbf{0}$ となり、$\lambda \neq 0$ かつ $\boldsymbol{v} \neq \mathbf{0}$ に矛盾する。以上より、λ は BA の固有値である。同様にして、BA の固有値は AB の固有値であることが示される。 □

8.3　エカート・ヤング（Eckart-Young）の定理の証明

定理 エカート・ヤングの定理（8.5、p.245）　$m \times n$ 行列 A（ランク r）の特異値分解を $A = U\Sigma V^\top$ とし、Σ の対角成分を降順に $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots$ とする。A と同じ型の任意のランク $k(\leq r)$ の行列 B について、

$$\|A - \hat{A}_k\|_F \leq \|A - B\|_F$$

が成り立つ。ここで、$\hat{A}_k$ は以下で定義されるランク k で打ち切った近似行列（式 (8.2)、p.243）である。

$$\hat{A}_k = U\hat{\Sigma}_k V^\top$$

ただし、$\hat{\Sigma}_k$ は Σ の最初の k 個の対角成分以外を 0 にした行列。

証明の概略は、$A = U\Sigma V^\top$ と同じ U, V を使って $B = UCV^\top$ と変形し、A, B の比較を Σ, C の比較にすり替えることです。すなわち、$\|A - B\|_F = \|\Sigma - C\|_F$ に帰着させ、Σ に最も近い C を探します。

Proof. まず、直交行列の掛け算によって成分ノルムが不変であることを確認する。A に任意の直交行列 P を左から掛けて、

$$\|PA\|_F^2 = \operatorname{tr}((PA)^\top(PA)) = \operatorname{tr}(A^\top P^\top PA) = \operatorname{tr}(A^\top A) = \|A\|_F^2$$

となる。右から掛けても同様に不変（トレースの性質 1.7（p.34）を使った)。A の特異値分解（完全版）$A = U\Sigma V^\top$ を用いて、評価するノルムを l を変形する。

$$l = \|A - B\|_F = \|U^\top(A - B)V\|_F = \|\Sigma - U^\top BV\|_F$$

変形できる。ここで、$C = U^\top BV$ とおいて成分ノルムの 2 乗（全成分の 2 乗和）を計算する。対角成分と非対角成分に分けて考えると、

$$l^2 = \|\Sigma - C\|_F^2 = \sum_{i=1}^{r}(\sigma_{ii} - c_{ii})^2 + \sum_{i\neq j}(c_{ij})^2 \tag{*1}$$

この値が最小になる C を求める。ここで、B のランクが k であるという条件から、$C = U^\top BV$ のランクも k である。なぜなら、可逆行列の積でランクは不変（定理 5.2、p.137）である。l^2 を最小にするには、非対角成分 $c_{ij}(i \neq j)$ をすべて 0 にし、対角成分 c_{ii} を σ_{ii} に近づけるのがよい。この観察から C は対角行列であり、$\operatorname{rank} C = k$ の制約から、

$$c_{11} = \sigma_1,\ c_{22} = \sigma_2,\ \ldots, c_{kk} = \sigma_k, \quad \text{残りの成分は 0}$$

すなわち、$C = \hat{\Sigma}_k$ とするのが l を最小化する (*2)。以上より、l を最小化する B は、

$$B_{\min} = U\hat{\Sigma}_k V^\top = \hat{A}_k$$

である。 □

以上がエカート・ヤングの定理の短かい直感的証明です。実は (*2) の部分は端折った説明になっています。(*1) の l^2 は完全には 0 にならない（$\sum_{i=k+1}^{r}\sigma_i^2$ が残る）ので、本当にランク k という制約で $C = \hat{\Sigma}_k$ が最小かどうかを確認する必要があります。以下に、より正確に証明を追加しておきます。

Proof. 追加証明: (*2) を示すために、(*1) の式をより正確に評価する。$\|A\|_F^2 = \operatorname{tr}(A^\top A)$（定義 8.4、p.244 参照）は、$A$ の全成分の 2 乗和であり全成分を縦一列に並べた mn 成分のベクトル $\boldsymbol{v}_A$ のノルムの 2 乗 $\|\boldsymbol{v}_A\|^2$ と一致すること、同様に、行列 A, B の内積 $\boldsymbol{v}_A \cdot \boldsymbol{v}_B$ は $\operatorname{tr}(A^\top B)$ で定義できることを思い出すと、

$$l^2 = \|\Sigma - C\|_F^2 = \|\Sigma\|_F^2 + \|C\|_F^2 - 2\operatorname{tr}(\Sigma^\top C) \tag{*3}$$

となる。 C のランクは B と同じ k であるから、非零特異値を降順に $\tau_1 \geq \cdots \geq \tau_k > 0$ とすると、von Neumann のトレース不等式 (**) より、

$$\operatorname{tr}(\Sigma^\top C) \leq \sigma_1\tau_1 + \cdots + \sigma_k\tau_k$$

よって、(*3) は $k \leq r$ であることに注意して、

$$\begin{aligned}
l^2 &\geq \|\Sigma\|_F^2 + \|C\|_F^2 - 2(\sigma_1\tau_1 + \cdots + \sigma_k\tau_k) \\
&= \sum_{i=1}^{r}\sigma_i^2 + \sum_{i=1}^{k}\tau_i^2 - 2\sum_{i=1}^{k}\sigma_i\tau_i \\
&= \sum_{i=1}^{k}(\sigma_i - \tau_i)^2 + \sum_{i=k+1}^{r}\sigma_i^2 \\
&\geq \sum_{i=k+1}^{r}\sigma_i^2 \quad (i \text{ の範囲に注意})
\end{aligned}$$

となる。一方、$\hat{\Sigma}_k$ を C として選ぶと、(*1) の式は、

$$l^2 = \|\Sigma - \hat{\Sigma}_k\|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2$$

となり、この下限を達成できる。よって、$C = \hat{\Sigma}_k$ は最適解である。そのとき、$B_{\min} = U\hat{\Sigma}_k V^\mathsf{T} = \hat{A}_k$ である。 □

(**) で使った「フォン・ノイマン（von Neumann）のトレース不等式」を挙げておきます。

補題 8.1 フォン・ノイマンのトレース不等式 A, B を同じ型の行列とし、$\operatorname{rank} A = r_A, \operatorname{rank} B = r_B$ とする。A, B の非零特異値をそれぞれ $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{r_A} > 0,\ \tau_1 \geq \tau_2 \geq \cdots \geq \tau_{r_B} > 0$ とする。すると、$r = \min(r_A, r_B)$ として、以下の不当式が成り立つ。

$$\operatorname{tr}(A^\mathsf{T} B) \leq \sigma_1\tau_1 + \sigma_2\tau_2 + \cdots + \sigma_r\tau_r$$