第 36 章

最小二乗解と最小ノルム解

解けない線形方程式

$\boldsymbol{x}$ についての線形方程式を考える。

$$A\boldsymbol{x} = \boldsymbol{b}$$

ここで、A は $m \times n$ 型行列、$\boldsymbol{x}$ は n 次元ベクトル、$\boldsymbol{b}$ は m 次元ベクトルである。

$m = n$ で、A が正則であれば、その逆行列を用いて、一意な解を求めることができる。

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

しかし、この線型方程式が解けない場合を扱うこともある。

i. $m > n$（A が縦長）の場合、条件式の数が未知数の数より多く、方程式は不能（解なし）
ii. $m < n$（A が横長）の場合、条件式の数が未知数の数より少なく、方程式は不定（解が無数にある）

このような場合には、次のようなアプローチが考えられる。

i. 解が存在しないなら、なるべく近い解を探す
ii. 解が無数にあるなら、その中から適した解を探す

なにをもって「近い」「適した」とすべきかは状況によって異なるが、近い解は最小二乗解、適した解は最小ノルム解とすることが多い。

最小二乗解

$A\boldsymbol{x} = \boldsymbol{b}$ の解がないなら、「せめて $A\boldsymbol{x}$ が $\boldsymbol{b}$ にできるだけ近くなるような $\boldsymbol{x}$ を求めよう」というのが最小二乗解のアプローチである。

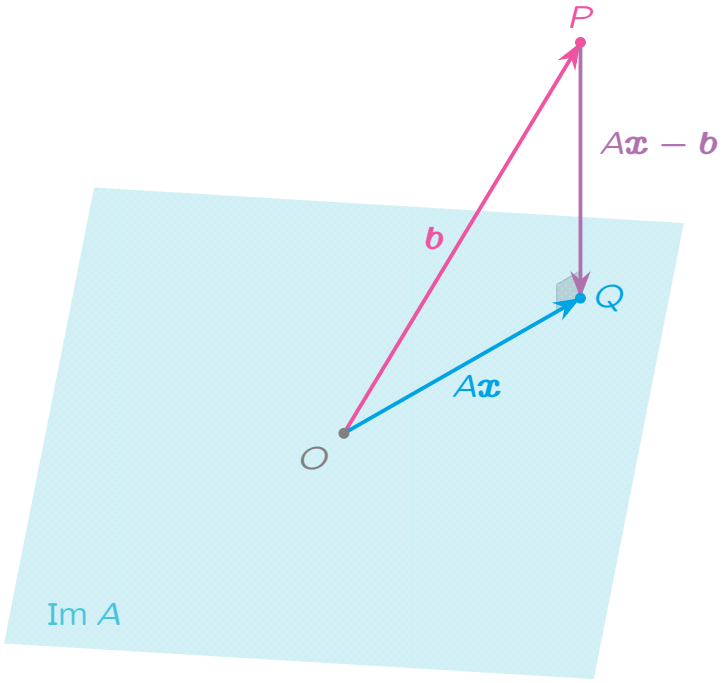
最小二乗解とは、次のような残差 J を最小化するような $\boldsymbol{x}$ のことである。

$$J = \|A\boldsymbol{x} - \boldsymbol{b}\|^2$$

残差の最小化と直交射影

$\boldsymbol{x}$ を動かすと、$A\boldsymbol{x}$ は Im A 上のさまざまなベクトルをとる。

$m > n$ の場合、$\boldsymbol{b}$ は $\boldsymbol{x}$ よりも高次元のベクトルなので、Im A からはみ出してしまう。



ここで、直交射影と最短距離［第 23 章］より、点 P から最短となる Im A 上の点は、P を Im A へ直交射影した点 Q である。

このことから、$\|A\boldsymbol{x} - \boldsymbol{b}\|$ が最小となるのは、

$$\boldsymbol{b} \text{ を列空間 } \mathrm{Im}\, A \text{ へ直交射影したベクトルが } A\boldsymbol{x}$$

になっているときだとわかる。

$\mathrm{Im}\, A$ は A の列空間（A の列ベクトルが張る空間）であるから、これまで通り $\mathcal{U}$ と表すことにしよう。
すると、最小二乗解 $\boldsymbol{x}$ が満たすべき条件は、次のように書ける。

$$A\boldsymbol{x} = P_{\mathcal{U}}\boldsymbol{b}$$

⚓ **theorem 36.1 - 最小二乗解と列空間への直交射影の関係**

A の列空間を $\mathcal{U}$ とし、$\mathcal{U}$ への直交射影行列を $P_{\mathcal{U}}$ とすると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小二乗解は次の関係を満たす。

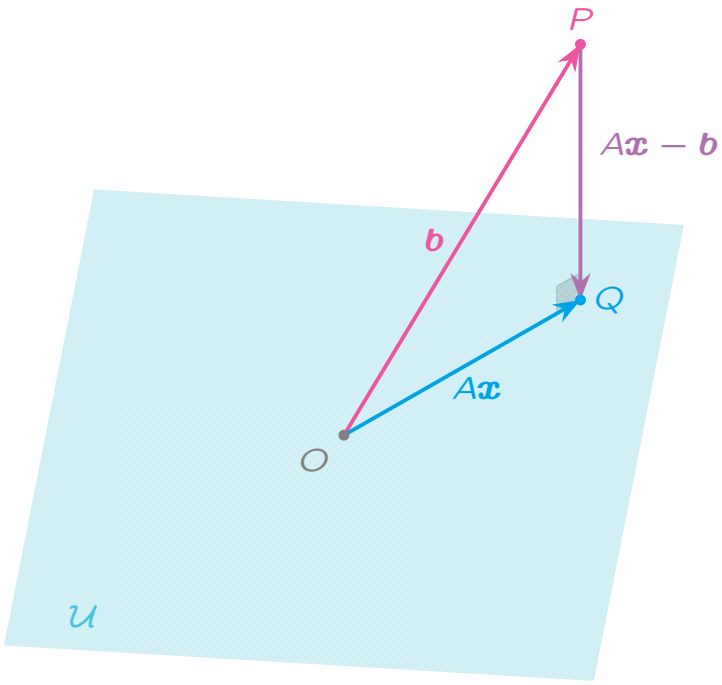
$$A\boldsymbol{x} = P_{\mathcal{U}}\boldsymbol{b}$$

列フルランクの場合の最小二乗解

$n = \mathrm{rank}\, A$ だと仮定する。**theorem 9.4**「階数と線型独立な列ベクトルの最大個数」より、階数は線型独立な列の最大個数なので、A の n 個の列ベクトルは線型独立である。

A の列ベクトルを $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ とすると、これらは $\mathcal{U}$ を張るベクトルであり、かつ線型独立であるので、$\mathcal{U}$ の基底を成す。
よって、$\mathcal{U}$ 上の任意のベクトルは、これらの線型結合で表される。

直交射影の定義より、$A\boldsymbol{x} - \boldsymbol{b}$ は $\mathcal{U}$ 上のすべてのベクトルに直交する。



$A\boldsymbol{x}-\boldsymbol{b}$ が $\mathcal{U}$ 上の任意のベクトルに直交するには、

$$\boldsymbol{a}_i^\top(A\boldsymbol{x}-\boldsymbol{b})=0 \quad (i=1,\ldots,n)$$

が成り立てばよい（$\mathcal{U}$ 上の任意のベクトルは $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ の線型結合で表されるのだから）。この条件を行列の形に書き直すと、

$$A^\top(A\boldsymbol{x}-\boldsymbol{b})=\boldsymbol{o}$$

となる。

補足

$A^\top(A\boldsymbol{x}-\boldsymbol{b})=\boldsymbol{o}$ は、$\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ すべてとの内積を並べた形である。

$$\begin{pmatrix}\boldsymbol{a}_1^\top\\ \vdots\\ \boldsymbol{a}_n^\top\end{pmatrix}(A\boldsymbol{x}-\boldsymbol{b})=\begin{pmatrix}\boldsymbol{a}_1^\top(A\boldsymbol{x}-\boldsymbol{b})\\ \vdots\\ \boldsymbol{a}_n^\top(A\boldsymbol{x}-\boldsymbol{b})\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

これを展開すると、

$$A^\top A\boldsymbol{x}-A^\top\boldsymbol{b}=\boldsymbol{o}$$
$$A^\top A\boldsymbol{x}=A^\top\boldsymbol{b}$$

となるので、$A^\top A$ が正則であれば、$\boldsymbol{x}=\cdots$ の形に整理できる。

$n = \operatorname{rank} A$ という仮定のもとでは、次の定理が成り立つ。

theorem – 線型独立な列ベクトルと自己共役積の正則性

$m \times n$ 型行列 A において、$m \geq n = \operatorname{rank} A$（列ベクトルが線型独立である）とき、$A^\top A$ は正則である。

証明

theorem 11.6「核空間の次元による正則判定」より、$A^\top A$ が正則であることと同値な条件として、$A^\top A$ の核空間の次元が 0 であることを示す。

すなわち、次を示せばよい。

$$A^\top A\boldsymbol{x} = \boldsymbol{o} \Longrightarrow \boldsymbol{x} = \boldsymbol{o}$$

$A^\top A\boldsymbol{x} = \boldsymbol{o}$ の両辺に $\boldsymbol{x}^\top$ をかけると、

$$\begin{aligned} \boldsymbol{x}^\top A^\top A\boldsymbol{x} &= \boldsymbol{o} \\ (A\boldsymbol{x})^\top (A\boldsymbol{x}) &= 0 \\ \|A\boldsymbol{x}\|^2 &= 0 \end{aligned}$$

ノルムが 0 であるベクトルは零ベクトルのみであるから、

$$A\boldsymbol{x} = \boldsymbol{o}$$

ここで、$A\boldsymbol{x} = \boldsymbol{o}$ ということは、A の列ベクトルの線形結合として $\boldsymbol{o}$ になる $\boldsymbol{x}$ が存在することを意味する。

しかし、仮定より A の列ベクトルは線型独立なので、$\boldsymbol{x} = \boldsymbol{o}$ しか解がない。

よって、$A^\top A\boldsymbol{x} = \boldsymbol{o}$ の解空間である核空間 $\operatorname{Ker}(A^\top A)$ は、零ベクトルしか含まない。

すなわち、$\dim \operatorname{Ker}(A^\top A) = 0$ であり、$A^\top A$ は正則である。 ■

$A^\top A$ が正則であることがわかったので、$\boldsymbol{x}$ は次のように求められる。

$$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$$

これが、$m \geq n = \mathrm{rank}\, A$ という条件のもとでの最小二乗解である。

⚓ **theorem - 線型独立な列ベクトルに対する最小二乗解**

$m \times n$ 型行列 A において、$m \geq n = \mathrm{rank}\, A$（列ベクトルが線型独立である）とき、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小二乗解 $\boldsymbol{x}$ は次のように表される。

$$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$$

列空間への直交射影行列

$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$ の両辺に左から A をかけると、

$$A\boldsymbol{x} = A(A^\top A)^{-1} A^\top \boldsymbol{b}$$

この式を、**theorem 36.1**「最小二乗解と列空間への直交射影の関係」と見比べると、

$$P_{\mathcal{U}} = A(A^\top A)^{-1} A^\top$$

であることがわかる。

⚓ **theorem 36.2 - 列空間への直交射影の行列表現**

$m \times n$ 型行列 A において、$m \geq n = \mathrm{rank}\, A$（列ベクトルが線型独立）のとき、$A$ の列空間への直交射影行列は次のように表される。

$$P_{\mathcal{U}} = A(A^\top A)^{-1} A^\top$$

最小ノルム解

$A\boldsymbol{x} = \boldsymbol{b}$ の解が無数にある場合は、「その中で有用な解を選ぼう」というアプローチをとる。このような場合、$\|\boldsymbol{x}\|$ が最小となる解を選ぶことが多い。

$\|\boldsymbol{x}\|$ が最小となる解を選ぶメリットはさまざまある。

- 解空間の中で「原点に最も近い」解

- 「変位が最小」など、物理学でも頻出する条件を満たす解

しかし、$\|\boldsymbol{x}\|$ は平方根を含むため、計算が複雑になりがちである。

そのため、2 乗して平方根を外したもの、すなわち $\|\boldsymbol{x}\|^2$ が最小となる解を考えることが多い。
$\|\boldsymbol{x}\|^2$ が最小となる解は、最小ノルム解と呼ばれる。

核空間と直交する空間への射影

解のパラメータ表示の再解釈［第 8 章］でも考察したように、$A\boldsymbol{x} = \boldsymbol{b}$ の解は次のようにして構成することができる。

1. 1 つの解 $\boldsymbol{x}_0$ を見つける
2. Ker A の任意のベクトル $\boldsymbol{u}$ を持ってくれば、$\boldsymbol{x}_0 + \boldsymbol{u}$ も解である

補足：なんで $\boldsymbol{x}_0 + \boldsymbol{u}$ も解だといえるのか？

A によって $\boldsymbol{o}$ に写ってしまうような、つまり $A\boldsymbol{u} = \boldsymbol{o}$ となるような $\boldsymbol{u}$ すべての集合が Ker A である。

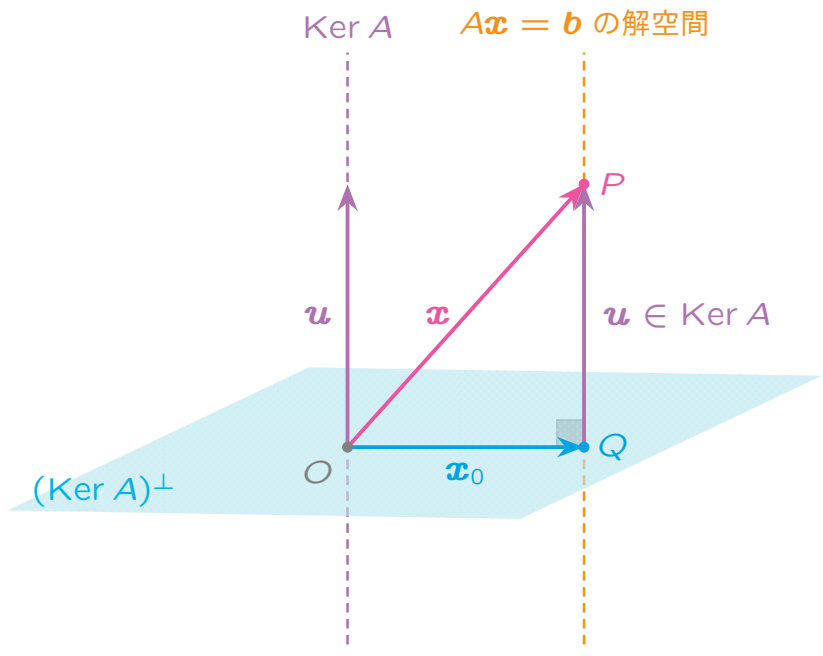
$A\boldsymbol{x}_0 = \boldsymbol{b}$ が成り立つことと合わせると、

$$A(\boldsymbol{x}_0 + \boldsymbol{u}) = A\boldsymbol{x}_0 + A\boldsymbol{u} = A\boldsymbol{x}_0 + \boldsymbol{o} = A\boldsymbol{x}_0 = \boldsymbol{b}$$

となり、たしかに $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{u}$ も $A\boldsymbol{x} = \boldsymbol{b}$ を満たすことがわかる。

そのため、Ker A を $\boldsymbol{x}_0$ だけ平行移動したものが、$A\boldsymbol{x} = \boldsymbol{b}$ の解空間となる。

$$\text{解空間} = \{\boldsymbol{x}_0 + \boldsymbol{u} \mid \boldsymbol{u} \in \mathrm{Ker}\, A\}$$



さて、$A\boldsymbol{x} = \boldsymbol{b}$ の解

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{u} \quad (\boldsymbol{u} \in \mathrm{Ker}\, A)$$

において、$\boldsymbol{x}$ のノルム（長さ）が最小となるのは、明らかに $\boldsymbol{u} = \boldsymbol{o}$ のときである。

つまり、$\boldsymbol{x} = \boldsymbol{x}_0$ が最小ノルム解となるが、$\boldsymbol{x}_0$ を使わずに最小ノルム解を表現したい。

図を見ると、$\boldsymbol{x}_0$ は、$\boldsymbol{x}$ を平面 $(\mathrm{Ker}\, A)^{\perp}$ へ直交射影したものであることがわかるので、

解 $\boldsymbol{x}$ のノルムを最小化するには、

$\boldsymbol{x}$ を $\mathrm{Ker}\, A$ と直交する空間に直交射影すればよい

ということがいえる。

ここで、**theorem 28.1**「核空間と行空間の直交関係」より、「$\mathrm{Ker}\, A$ と直交する空間」は、A の行空間 $\mathrm{Im}\, A^{\top}$ である。

まとめると、$\|\boldsymbol{x}\|$ が最小となるのは、

$\boldsymbol{x}_0$ を行空間 $\mathrm{Im}\, A^{\top}$ へ直交射影したベクトルが $\boldsymbol{x}$

となっているときだとわかる。

行空間 $\mathrm{Im}\, A^\top$ をこれまで通り $\mathcal{V}$ と表すことにすると、最小ノルム解は次のように表現できる。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0 \quad (A\boldsymbol{x}_0 = \boldsymbol{b})$$

theorem 36.3 - 最小ノルム解と行空間への直交射影の関係

A の行空間を $\mathcal{V}$ とし、$\mathcal{V}$ への直交射影行列を $P_{\mathcal{V}}$ とすると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小ノルム解は次の関係を満たす。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0 \quad (A\boldsymbol{x}_0 = \boldsymbol{b})$$

行空間への直交射影行列

転置すると行と列が入れ替わるため、A の行空間は、$A^\top$ の列空間と言い換えられる。

そのため、A の **theorem 36.2**「列空間への直交射影の行列表現」の式

$$P_{\mathcal{U}} = A(A^\top A)^{-1}A^\top$$

において、A を $A^\top$ に置き換えたものが、A の行空間への直交射影行列となる。

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

theorem - 行空間への直交射影の行列表現

$m \times n$ 型行列 A において、$n \geq m = \mathrm{rank}\, A$（行ベクトルが線型独立）のとき、$A$ の行空間への直交射影行列は次のように表される。

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

行フルランクの場合の最小ノルム解

$\mathrm{rank}\, A = m$ という条件のもとでは、$P_{\mathcal{V}}$ を

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

という式で表現できるので、これを $\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0$ という関係に代入すると、

$$\begin{aligned} \boldsymbol{x} &= A^\top(AA^\top)^{-1}A\boldsymbol{x}_0 \\ &= A^\top(AA^\top)^{-1}\boldsymbol{b} \end{aligned} \quad A\boldsymbol{x}_0 = \boldsymbol{b}$$

という形で、最小ノルム解が求まる。

⚓ **theorem – 線型独立な行ベクトルに対する最小ノルム解**

$m \times n$ 型行列 A において、$n \geq m = \operatorname{rank} A$（行ベクトルが線型独立である）とき、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小ノルム解 $\boldsymbol{x}$ は次のように表される。

$$\boldsymbol{x} = A^\top (AA^\top)^{-1} \boldsymbol{b}$$

擬似逆行列による解

擬似逆行列と行空間・列空間への射影［第 35 章］より、ムーア・ペンローズの擬似逆行列と直交射影行列は、次のような関係で結ばれていた。

$$A^+ A = P_{\mathcal{V}}, \quad AA^+ = P_{\mathcal{U}}$$

これらの関係を用いると、最小二乗解や最小ノルム解などのすべての場合を包括した解の表現が得られる。

最小二乗解

theorem 36.1「最小二乗解と列空間への直交射影の関係」より、最小二乗解は次の条件を満たすものだった。

$$A\boldsymbol{x} = P_{\mathcal{U}} \boldsymbol{b}$$

擬似逆行列と直交射影行列の関係 $AA^+ = P_{\mathcal{U}}$ を用いると、

$$A\boldsymbol{x} = AA^+ \boldsymbol{b}$$

この式から、$\boldsymbol{x} = A^+ \boldsymbol{b}$ が、$A\boldsymbol{x}$ を $\boldsymbol{b}$ に最も近づける近似解であることがわかる。

> A が正則でない場合を考えているので、両辺に A^{-1} をかけて $\boldsymbol{x}$ を得るという同値変形は成り立たない。それゆえ、$\boldsymbol{x} = A^+ \boldsymbol{b}$ は「近似解」でしかない。

よって、最小二乗解は、次のように表すことができる。

$$\boldsymbol{x} = A^+ \boldsymbol{b}$$

最小ノルム解

theorem 36.3「最小ノルム解と行空間への直交射影の関係」より、最小ノルム解は次の条件を満たすものだった。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0$$

擬似逆行列と直交射影行列の関係 $A^+A = P_{\mathcal{V}}$ を用いると、

$$\boldsymbol{x} = A^+A\boldsymbol{x}_0 = A^+\boldsymbol{b}$$

よって、最小ノルム解は、次のように表すことができる。

$$\boldsymbol{x} = A^+\boldsymbol{b}$$

擬似逆行列はあらゆる場合の解を表せる

このように、最小二乗解も最小ノルム解も、ムーア・ペンローズの擬似逆行列を用いると同じ式で表現できる。
そして、ここでの議論は、射影による条件のみを前提としており、A の行や列の線型独立性（フルランクかどうか）に依存していない。

つまり、フルランクでない場合も、擬似逆行列 A^+ が求められれば、

$$\boldsymbol{x} = A^+\boldsymbol{b}$$

という形で、最小二乗解や最小ノルム解を求めることができる。

また、theorem 35.1「正則行列に対するムーア・ペンローズ逆」より、A が正則な場合は $A^+ = A^{-1}$ となるので、$\boldsymbol{x} = A^{-1}\boldsymbol{b}$ という逆行列を用いた解の表現も包括していることがわかる。

⚓ **theorem – 擬似逆行列による線形方程式の解**

A のムーア・ペンローズ擬似逆行列 A^+ を用いると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解は次のように表される。

$$\boldsymbol{x} = A^+\boldsymbol{b}$$