最小二乗法の基本

least squares method

1 導入

この講義で重要なのは、最小二乗法は「解が存在しない方程式を無理に解く」方法ではなく、右辺を列空間へ射影する方法だということである。

$Ax = b$ が解を持つのは、b が A の列空間に属する場合である。測定値を含む問題では、b が列空間から少し外れることが多い。そのときは、Ax を b に最も近い列空間の点に設定する。

2 用語と定義

最小二乗法とは、行列 A とベクトル b に対して

$$\| Ax - b \|^2$$

を最小化する x を求める方法である。

残差とは、

$$r = b - Ax$$

で定義される誤差ベクトルである。

正規方程式とは、実数行列で

$$A^T Ax = A^T b$$

と表される方程式である。

3 方針

Ax は A の列空間に属する。したがって $\| Ax - b \|$ を最小化するには、b を列空間へ直交射影すればよい。

最良近似 Ax と誤差 $b - Ax$ は直交するため、そこから正規方程式が導かれる。

→ 講義 直交補空間と射影 lecture math linear-algebra
https://study.bem130.com/lecture/math/linear-algebra/直交補空間と射影-講義/

→ 講義 複素内積とユニタリ行列 lecture math linear-algebra
https://study.bem130.com/lecture/math/linear-algebra/複素内積とユニタリ行列-講義/

4 直感的な説明

列空間を 1 つの平面とみなす。b がその平面の外にあるなら、$Ax = b$ は成立しない。しかし平面の中で b に最も近い点は存在する。その点が Ax であり、垂直成分として分離される誤差が残差である。

この垂直という条件が計算へ変換されると、$A^T(b - Ax) = 0$ となる。

なぜ列空間へ射影するのかは、Ax の形から決定される。x をどのように選択しても、候補となる左辺 Ax は A の列の線型結合である。したがって到達可能な右辺の集合は $\mathrm{Col}(A)$ に限定される。b がそこに属しない場合、問題は「b に最も近い $\mathrm{Col}(A)$ の点を選択する」ことへ変換される。

最近点 $\hat{b}$ では、残差 $r = b - \hat{b}$ が列空間のどの方向にも成分を持たない。もし列空間の方向に残差の成分が残存していれば、その方向へ $\hat{b}$ を微小に移動して距離を短縮できるためである。したがって

$$b = \hat{b} + r, \qquad \hat{b} \in \mathrm{Col}(A), \quad r \in \mathrm{Col}(A)^{\perp}$$

という直交分解が最小二乗法の幾何学的内容である。

5 厳密な説明

5.1 1. 列空間への射影

A の列空間を $\mathrm{Col}(A)$ とする。Ax は常に $\mathrm{Col}(A)$ に属する。最小二乗解では、Ax が b の $\mathrm{Col}(A)$ への直交射影になる。

したがって残差

$$r = b - Ax$$

は列空間に直交する。

5.2 2. 直交条件から正規方程式へ

A の列を $a_1, \ldots, a_n$ とする。r が列空間に直交するとは、

$$\langle r, a_j \rangle = 0 \qquad (j = 1, \ldots, n)$$

である。これは行列で

$$A^T r = 0$$

と表現できる。$r = b - Ax$ を代入すると、

$$A^T(b - Ax) = 0$$

すなわち

$$A^T A x = A^T b$$

を得る。これが正規方程式である。

複素の場合は、列との直交条件は同じく $\langle r, a_j \rangle = 0$ であるが、行列では $A^* r = 0$ と書く。したがって

$$A^*(b - Ax) = 0$$

すなわち

$$A^* A x = A^* b$$

を得る。ここで A^* は共役転置であり、複素内積に対する随伴として現れる。

5.3 3. 一意性

A の列が一次独立なら、$A^T A$ は可逆である。この場合、最小二乗解は

$$x = (A^T A)^{-1} A^T b$$

で一意に決定される。列が一次従属なら、最小化する x は複数存在しうる。その場合、最小ノルム解を選択する道具が擬似逆行列である。

5.4 4. 射影行列

A が列満階数なら、列空間への直交射影は

$$P = A(A^T A)^{-1} A^T$$

で表現される。このとき

$$\hat{b} = Pb, \qquad r = b - \hat{b} = (I - P)b$$

であり、$\hat{b}$ が b の列空間への射影、r が直交残差である。さらに

$$P^2 = P, \qquad P^T = P$$

が成立するため、P は「一度射影すれば再度射影しても変化しない」対称な射影行列である。複素の場合は A^T を A^* に置換し、

$$P = A(A^* A)^{-1} A^*$$

を用いる。

6 具体例

$$A = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

とする。Ax は全成分が x のベクトルであり、b を定数で近似する問題である。正規方程式は

$$A^T A x = A^T b$$

すなわち

$$3x = 7$$

である。したがって $x = \frac{7}{3}$ となる。これは $1, 2, 4$ の平均であり、定数で最良に近似する値になる。

7 別の観点

幾何的には、最小二乗法は点を部分空間へ射影する方法である。統計的には、観測値と予測値の差の 2 乗和を最小化する方法である。

8 数値計算上の注意

理論上は正規方程式や射影行列で最小二乗解を記述できる。一方で数値計算では、$A^T A$ を直接構成すると条件数が悪化する場合がある。そのため、実装では QR 分解や特異値分解を用いることが標準的である。

9 判定基準

- $Ax = b$ が過剰な条件を持ち、厳密解が存在しない場合は最小二乗法を検討する。
- 列空間への最近点を求める問題では、残差の直交条件を立式する。
- 列が一次独立なら $A^T A$ の可逆性を利用する。

10 どこまで成り立つか

実数の標準内積では A^T を用いる。複素の場合は共役転置 A^* を用い、正規方程式は $A^*Ax = A^*b$ となる。

列が一次従属の場合、射影 $\hat{b}$ は一意であるが、係数 x は一意とは限らない。

11 最終形

$$\boxed{\min_x \| Ax - b \|^2}$$

$$\boxed{A^T(b - Ax) = 0}$$

$$\boxed{A^T Ax = A^T b}$$

$$\boxed{P = A(A^T A)^{-1} A^T \quad \text{(full column rank)}}$$

$$\boxed{A^* Ax = A^* b \quad \text{(complex case)}}$$

12 一言でいうと

- 最小二乗法は、b を列空間へ直交射影し、残差を列空間に直交させる方法である。

13 演習リンク

→ 基本演習 **内積・直交・射影** exercise math linear-algebra
https://study.bem130.com/exercise/math/linear-algebra/内積・直交・射影-基本演習/

→ 基本演習 **複素内積とユニタリ行列** exercise math linear-algebra
https://study.bem130.com/exercise/math/linear-algebra/複素内積とユニタリ行列-基本演習/

14 関連リンク

→ 講義 **直交補空間と射影** lecture math linear-algebra
https://study.bem130.com/lecture/math/linear-algebra/直交補空間と射影-講義/

→ 講義 **複素内積とユニタリ行列** lecture math linear-algebra
https://study.bem130.com/lecture/math/linear-algebra/複素内積とユニタリ行列-講義/

→ 講義 **擬似逆行列の基本** lecture math linear-algebra
https://study.bem130.com/lecture/math/linear-algebra/擬似逆行列の基本-講義/