縦長ヘッセンベルグ行列で定められる
最小二乗問題の数値解法

1　一般の優決定系における最小二乗問題

縦長フルランク行列 $A \in \mathbb{R}^{n\times m}$ と行列 $b \in \mathbb{R}^n$ について、以下の最小二乗問題を解く。

$$\underset{x}{\text{minimize}} \qquad \|Ax - b\|^2 \tag{1}$$

この最適解は A の擬似逆行列 A^+ を用いて

$$x = A^+ b \tag{2}$$

である。

2　QR 分解に基づく擬似逆行列の表現

縦長列フルランク行列 $A \in \mathbb{R}^{n\times m}$ が $A = QR$ と QR 分解されるとき、A の擬似逆行列 A^+ は

$$A^+ = R^{-1}Q^T \tag{3}$$

となる。このことを証明する。

まず、一般に、行列 A に対して以下の 4 式を満たす行列 A^+ はただ一つ存在し、この A^+ がムーアペンローズの擬似逆行列であることが知られている。

$$AA^+A = A \tag{4}$$

$$A^+AA^+ = A^+ \tag{5}$$

$$(AA^+)^T = AA^+ \tag{6}$$

$$(A^+A)^T = A^+A \tag{7}$$

行列 $B = R^{-1}Q^T$ がこの式を満たすことを確認する。Q^TQ は単位行列 I に等しいので、$BA = (R^{-1}Q^T)(QR) = I$ であることがわかる。ゆえに、

$$ABA = A(BA) = A \tag{8}$$

$$BAB = (BA)B = B \tag{9}$$

$$(BA)^T = I = BA \tag{10}$$

となる。さらに $AB = QRR^{-1}Q^T = QQ^T$ は明らかに対称行列なので、

$$(AB)^T = AB \tag{11}$$

以上より、B は A の擬似逆行列に等しい。

3　ヘッセンベルグ行列の QR 分解

まず A が上三角行列であるときの QR 分解は、直交行列 Q として単位行列 I、上三角行列 R として A を取り次式のように QR 分解される：

$$A = IA \tag{12}$$

次に、今回の主題である A がヘッセンベルグである場合を考える。ヘッセンベルグ行列とは上三角行列から一つだけ左下にはみ出た場合の行列であり、以下のような形で表される。ただし今回は縦長を仮定するので、$n = m + 1$ である。

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & & \vdots \\ & a_{3,2} & \ddots & \vdots \\ & & \ddots & \vdots \\ O & & & a_{n,m} \end{bmatrix} \tag{13}$$

この QR 分解は、はみ出た余分な部分を右上に押し込む直交行列が取れれば実現できそうである。この”余分な部分を押し込む”ための直交変換として、次式で定義されるギブンス変換を用いることで高速に計算を行うことができる。

$$G(a,b,\theta)_{i,j} = \begin{cases} \cos\theta & (i,j) = (a,a),(b,b) \\ \sin\theta & (i,j) = (a,b) \\ -\sin\theta & (i,j) = (b,a) \\ 1 & i = j \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

この行列を具体的に書き下すと以下のようになる。るので、$n = m + 1$ である。

$$A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos\theta & \cdots & \sin\theta & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -\sin\theta & \cdots & \cos\theta & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \tag{15}$$

この行列は各列ベクトルの大きさが 1 であり、かつ列ベクトル同士は直交しているため直交行列である。

ギブンス変換を用いた上三角化戦略の戦略を述べる。2×2 行列と回転行列の積

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{16}$$

の左下成分は $-a\sin\theta + c\cos\theta$ であり、

$$\cos\theta = \frac{a}{\sqrt{a^2 + c^2}} \tag{17}$$

$$\sin\theta = \frac{c}{\sqrt{a^2 + c^2}} \tag{18}$$

となる θ を取れば上三角化できることがわかる。[1] このギブンス変換を $(a, b) = (1, 2), (2, 3), ..., (m, n)$ までについて順番に繰り返せば良いのである。

4 一回冷静になる

最小二乗解は A^+b であり、A^+ は QR 分解を用いて具体的に $R^{-1}Q^T$ と与えられる。したがって、求める解 x は

$$x = R^{-1}Q^T b \tag{19}$$

であるらしい。ここで、右辺に R^{-1} となっていることに疑問を持ちたい。R は正則だから、

$$Rx = Q^T b \tag{20}$$

を解いてもいいはずである。この連立一次方程式は上三角行列 R を係数行列とするから後退代入で簡単に解を求めることができる。

一般に、アルゴリズムの設計において逆行列を用いるかというのはオーダーが大きく変わる深刻な問題であるので、実装前に一回冷静になることが大事である。

5 まとめ

1. QR 分解
    - ヘッセンベルグ行列 A にギブンス変換 (特殊な直交変換)G を作用させる
    - この変換を同時に b にも作用させる
2. 上三角行列を係数行列とする連立一次方程式 $Rx = Q^T b$ を解く
3. x を解として出力

[1] 言うまでもないが、実際に θ を求める必要などない