最小２乗法に関する公式の導出

ここでは、直線や平面を求める代表的な方法である線形最小2乗推定問題を考えます。

直線の方程式は、

$$ax+by+c=0 \tag{1}$$

円の方程式は、

$$x^2+y^2+ax+by+c=0 \tag{2}$$

平面の方程式は、

$$ax+by+cz+d=0 \tag{3}$$

で表されます。

これらは各々、

$$\begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=-c \tag{4}$$

$$\begin{pmatrix} a & b & c \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}=-\left(x^2+y^2\right) \tag{5}$$

$$\begin{pmatrix} a & b & c \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=-d \tag{6}$$

と表すことができます。

これらの式において最小2乗法で求めたいのはaやb、それにcなので各々を以下のように変形します。

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=-c \tag{7}$$

$$\begin{pmatrix} x & y & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}=-\left(x^2+y^2\right) \tag{8}$$

$$\begin{pmatrix} x & y & z \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}=-d \tag{9}$$

これらは全て

$$A\cdot\vec{x}=\vec{c} \tag{10}$$

と表すことができます。

よってこれは、$A\cdot\vec{x}=\vec{c}$における固有値が１の時の固有ベクトル$\vec{x}$を求める問題に帰着されます（図２・図３参照）。

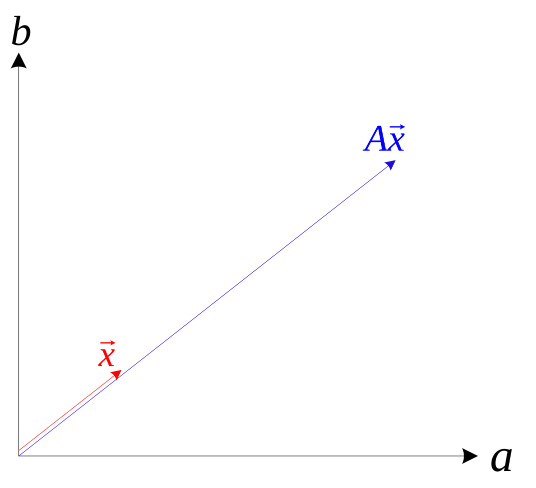

図1　直線の場合

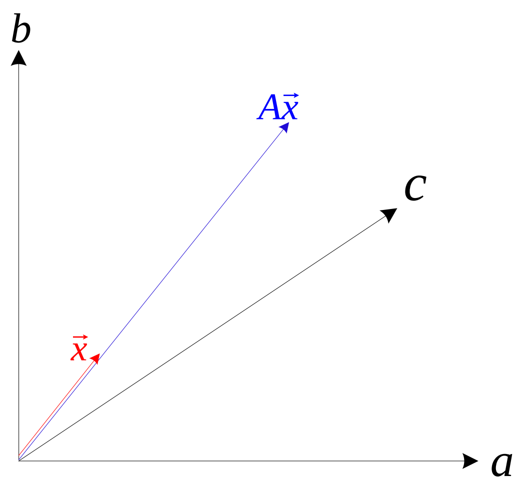

図2　円や平面の場合

以下、$\vec{x}$ を求めましょう。

まず、残差 R を下式で定義します。

$$R = \left\| A \cdot \vec{x} - \vec{c} \right\|^2 \tag{11}$$

この式は、

$$\begin{aligned}
R &= \left(A \cdot \vec{x} - \vec{c}\right)^T \cdot \left(A \cdot \vec{x} - \vec{c}\right) \\
&= \left(\left(A \cdot \vec{x}\right)^T - \vec{c}^{\,T}\right) \cdot \left(A \cdot \vec{x} - \vec{c}\right) \\
&= \left(\vec{x}^T \cdot A^T - \vec{c}^{\,T}\right) \cdot \left(A \cdot \vec{x} - \vec{c}\right) \\
&= \vec{x}^T \cdot A^T \cdot A \cdot \vec{x} - \vec{x}^T \cdot A^T \cdot \vec{c} - \vec{c}^{\,T} \cdot A \cdot \vec{x} + \vec{c}^{\,T} \cdot \vec{c} \\
&= \vec{x}^T \cdot A^T \cdot A \cdot \vec{x} - \vec{x}^T \cdot A^T \cdot \vec{c} - \vec{x}^T \cdot \left(\vec{c}^{\,T} \cdot A\right)^T + \vec{c}^{\,T} \cdot \vec{c} \\
&= \vec{x}^T \cdot A^T \cdot A \cdot \vec{x} - \vec{x}^T \cdot A^T \cdot \vec{c} - \vec{x}^T \cdot A^T \cdot \vec{c} + \vec{c}^{\,T} \cdot \vec{c}
\end{aligned} \tag{12}$$

と展開できますので、$\vec{x}$ で微分すると、

$$\begin{aligned}
\frac{\partial R}{\partial \vec{x}} &= A^T \cdot A \cdot \vec{x} + A^T \cdot A \cdot \vec{x} - A^T \cdot \vec{c} - A^T \cdot \vec{c} \\
&= 2 \cdot \left(A^T \cdot A \cdot \vec{x} - A^T \cdot \vec{c}\right)
\end{aligned} \tag{13}$$

となります。ここで、R を極小化する $\vec{x}$ を求めるため、

$$\frac{\partial R}{\partial \vec{x}} = \vec{0} \tag{14}$$

の性質を考慮すると、

$$2 \cdot \left(A^T \cdot A \cdot \vec{x} - A^T \cdot \vec{c}\right) = 0 \tag{15}$$

ですから、

$$\vec{x} = \left(A^T \cdot A\right)^{-1} \cdot A^T \cdot \vec{c} \tag{16}$$

が得られます。

ここで、

$$B = \left(A^T \cdot A\right)^{-1} \cdot A^T \tag{17}$$

と置きます。

もし A が正方行列ならば $A \cdot \vec{x} = \vec{c}$ から

$$\vec{x} = A^{-1} \cdot \vec{c} \tag{18}$$

と計算できますので、

$$B = A^{-1} \tag{19}$$

となります。

A が非正方行列の場合はこのようにはできませんが、概念的にはその延長線上にあると考えます。つまり、B を非正方行列の逆行列のように考えることで抽象化を図ります。このことから B を一般逆行列と称しています。

以　上