近似計算（ロバスト推定）のアルゴリズム

（Ver.1.4）

2016 年 4 月

株式会社 アイディール

目次

1.　記法と公式

以下の説明にあたり，次の記法を用います．

- $\langle \mathbf{u}, \mathbf{v} \rangle$：列ベクトル $\mathbf{u}$ と列ベクトル $\mathbf{v}$ の内積，すなわち $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^{\mathsf{T}}\mathbf{v}$ （ T は転置）
- $diag(w_1, \cdots, w_n)$： $w_1, \cdots, w_n$ を対角要素とする対角行列

また，次の微分公式を用います．

- $\mathbf{u}, \mathbf{v}$ を n 次元列ベクトル， $f(\mathbf{u})$ をスカラー値関数， $\mathbf{A}$ を $n \times n$ 行列とするとき，

$$\frac{\partial f}{\partial \mathbf{u}} = \left(\frac{\partial f}{\partial u_1}, \cdots, \frac{\partial f}{\partial u_n} \right)^{\mathsf{T}}, \quad \frac{\partial \langle \mathbf{u}, \mathbf{v} \rangle}{\partial \mathbf{u}} = \mathbf{v}, \quad \frac{\partial \langle \mathbf{u}, \mathbf{u} \rangle}{\partial \mathbf{u}} = 2\mathbf{u},$$

$$\frac{\partial \mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{v}}{\partial \mathbf{u}} = \mathbf{A}\mathbf{v}, \quad \frac{\partial \mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{v}}{\partial \mathbf{v}} = \mathbf{A}^{\mathsf{T}}\mathbf{u}, \quad \frac{\partial \mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{u}}{\partial \mathbf{u}} = \mathbf{A}\mathbf{u} + \mathbf{A}^{\mathsf{T}}\mathbf{u} = (\mathbf{A} + \mathbf{A}^{\mathsf{T}})\mathbf{u}.$$

2.　ロバスト推定[1]

M 推定法によるロバスト推定では，第 i 番データの残差を ε_i とするとき，推定関数 $\rho(\varepsilon_i)$ の和の最小化を行います．すなわち，求めるべきパラメータをベクトルで $\mathbf{p}$ と表し，

$$f(\mathbf{p}) = \sum_i \rho(\varepsilon_i(\mathbf{p})) \tag{2-1}$$

とおくとき， $f(\mathbf{p})$ を最小化するパラメータ $\mathbf{p}$ を推定することになります．

ここで，推定関数 $\rho(x)$ は $x = 0$ で唯一の最小値を持つ正定値偶関数です．また，影響力関数および重み関数が次のように定義されます．

$$\text{影響力関数：}\ \psi(x) = \frac{\partial \rho(x)}{\partial x}, \qquad \text{重み関数：}\ w(x) = \frac{\psi(x)}{x}. \tag{2-2}$$

さらに，次の定義をしておきます．

$$w_i \equiv w(\varepsilon_i), \quad W \equiv \sum_i w(\varepsilon_i) = \sum_i w_i. \tag{2-3}$$

さて，式(2-1)を最小化する解は $f(\mathbf{p})$ を $\mathbf{p}$ で微分しそれを $\mathbf{0}$ とおくことで得られます．すなわち，

$$\frac{\partial f}{\partial \mathbf{p}} = \sum_i \frac{\partial \rho(\varepsilon_i)}{\partial \varepsilon_i} \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \sum_i \psi(\varepsilon_i) \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \sum_i w(\varepsilon_i)\varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \sum_i w_i\, \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{p}} = \mathbf{0} \tag{2-4}$$

を解くことにより解となるパラメータ $\mathbf{p}$ を求めます．式(2-4)は反復計算にて求めます．つまり，(k) を反復回数として，

$$\sum_i w(\varepsilon_i(\mathbf{p}^{(k-1)}))\varepsilon_i(\mathbf{p}^{(k)}) \frac{\partial \varepsilon_i(\mathbf{p}^{(k)})}{\partial \mathbf{p}} = \mathbf{0} \tag{2-5}$$

から $\mathbf{p}^{(k)}$ を求めて，次の重み $w(\varepsilon_i(\mathbf{p}^{(k)}))$ を決めていきます．重みの初期値は最小二乗解を用います．

ところで，式(2-4)を解くことは，次の重み付き最小二乗問題を解くことと等価になっています．

$$\text{minimize} \quad \sum_i w_i \varepsilon_i^2(\mathbf{p}). \tag{2-6}$$

なお， $\rho(x) = x^2/2$ のとき式(1-1)の最小化は最小二乗法となります．このとき， n をデータ数とすれば $w_i = 1\ (i = 1, \cdots, n)$ ， $W = n$ となります．

3. 超平面のロバスト推定

直線（２次元空間内）および平面（３次元空間内），あるいは一般に超平面の方程式は，その上にある１点 $\mathbf{a}$ と法線ベクトル $\mathbf{u}$ を用いて，

$$\langle \mathbf{r}-\mathbf{a}, \mathbf{u} \rangle = 0 \tag{3-1}$$

と表されます．この方程式の左辺は，法線ベクトル $\mathbf{u}$ が単位ベクトルであれば, 点 $\mathbf{r}$ と超平面との (符号付き) 距離を表しています. そこで,

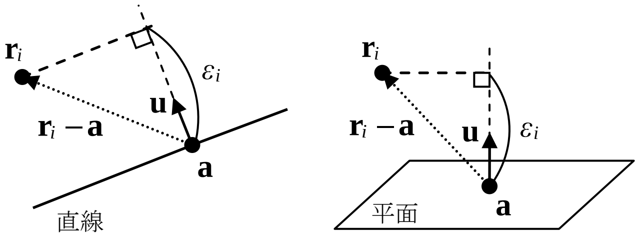

図１．データ点 $\mathbf{r}_i$ の残差 ε_i

データ数 n の点群 $\mathbf{r}_i$ $(i=1,\cdots,n)$ の各点 $\mathbf{r}_i$ に対して，この距離を誤差評価値（残差） ε_i とすれば,

$$\varepsilon_i(\mathbf{a},\mathbf{u}) = \left| \langle \mathbf{r}_i - \mathbf{a}, \mathbf{u} \rangle \right| = \sqrt{\langle \mathbf{r}_i - \mathbf{a}, \mathbf{u} \rangle^2}\ , \tag{3-2}$$

$$\text{ただし，}\ \langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^\mathsf{T}\mathbf{u} = 1\ \text{（単位ベクトル）} \tag{3-3}$$

と書けます．すると，ロバスト推定すなわち式(2-1)の最小化は，式(3-3)の制約条件のもとに行うためラグランジュ乗数 λ' を用いて， $\mathbf{a}$ と $\mathbf{u}$ を未知パラメータとする次の最小化問題に帰着されます.

$$\text{minimize} \quad J(\mathbf{a},\mathbf{u}) = \sum_i \rho(\varepsilon_i(\mathbf{a},\mathbf{u})) - \lambda'(\mathbf{u}^\mathsf{T}\mathbf{u} - 1)\,. \tag{3-4}$$

この最小化問題を解くために， J をパラメータ $\mathbf{a}$ および $\mathbf{u}$ で微分してゼロと置きます．まず， $\mathbf{a}$ で微分します．このとき，式(3-4)の右辺第１項の微分は式(2-4)を用います.

$$\frac{\partial J}{\partial \mathbf{a}} = \sum_i w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{a}} = \sum_i w_i \varepsilon_i \frac{2\langle \mathbf{r}_i - \mathbf{a}, \mathbf{u} \rangle (-\mathbf{u})}{2\varepsilon_i} = \left\langle \sum_i w_i (\mathbf{r}_i - \mathbf{a}), \mathbf{u} \right\rangle (-\mathbf{u}) = \mathbf{0}\,. \tag{3-5}$$

よって，$\mathbf{u} \neq \mathbf{0}$ より $\left\langle \sum w_i (\mathbf{r}_i - \mathbf{a}), \mathbf{u} \right\rangle = 0$ ，すなわち $\sum w_i(\mathbf{r}_i - \mathbf{a})$ は $\mathbf{u}$ と直交するベクトルとなります．そこで， $\sum w_i(\mathbf{r}_i - \mathbf{a}) = \mathbf{0}$ を採用すれば,

$$\mathbf{a} = \frac{1}{W}\sum_i w_i \mathbf{r}_i \tag{3-6}$$

が得られます．つまり， $\mathbf{a}$ はデータ点 $\mathbf{r}_i$ の重み付き重心となります.

次に， $\mathbf{u}$ で微分します.

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{u}} &= \sum_i w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{u}} - 2\lambda'\mathbf{u} = \sum_i w_i \varepsilon_i \frac{2\langle \mathbf{r}_i - \mathbf{a}, \mathbf{u} \rangle (\mathbf{r}_i - \mathbf{a})}{2\varepsilon_i} - 2\lambda'\mathbf{u} \\ &= \left(\sum_i w_i (\mathbf{r}_i - \mathbf{a})(\mathbf{r}_i - \mathbf{a})^\mathsf{T} \right)\mathbf{u} - 2\lambda'\mathbf{u} \\ &= \left(\mathbf{D}^\mathsf{T}\mathbf{W}\mathbf{D} \right)\mathbf{u} - 2\lambda'\mathbf{u} = \mathbf{0}\,. \end{aligned} \tag{3-7}$$

ただし，式(3-6)の $\mathbf{a}$ を用いて,

$$\mathbf{D} = \begin{pmatrix} \vdots \\ (\mathbf{r}_i - \mathbf{a})^\mathsf{T} \\ \vdots \end{pmatrix}, \qquad \mathbf{W} = diag(w_1, \cdots, w_n) \tag{3-8}$$

とおきました．さらに，$\lambda = 2\lambda'$ および $\mathbf{C} = \mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{D}$ とおけば式(3-7)は，

$$\mathbf{C}\mathbf{u} = \lambda\mathbf{u} \tag{3-9}$$

となり，これは行列 $\mathbf{C}$ の固有値問題です．よって，求めるべき法線ベクトル $\mathbf{u}$ は行列 $\mathbf{C}$ の固有値 λ に対する固有ベクトルとなります．

ところで，式(2-6)の評価値を用いて誤差の総和を評価してみると，

$$\begin{aligned}\sum_i w_i \varepsilon_i^{\,2} &= \sum_i w_i \left(\mathbf{u}^{\mathsf{T}}(\mathbf{r}_i - \mathbf{a})(\mathbf{r}_i - \mathbf{a})^{\mathsf{T}}\mathbf{u}\right) = \mathbf{u}^{\mathsf{T}}\left(\sum_i w_i(\mathbf{r}_i - \mathbf{a})(\mathbf{r}_i - \mathbf{a})^{\mathsf{T}}\right)\mathbf{u} \\ &= \mathbf{u}^{\mathsf{T}}\left(\mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{D}\right)\mathbf{u} = \mathbf{u}^{\mathsf{T}}\mathbf{C}\mathbf{u} = \lambda\mathbf{u}^{\mathsf{T}}\mathbf{u} = \lambda\end{aligned} \tag{3-10}$$

となります．これは，行列 $\mathbf{C}$ の最小固有値にて最小化が達成されることを示しています．すなわち，法線ベクトル $\mathbf{u}$ は行列 $\mathbf{C}$ の最小固有値に対する固有ベクトルとなります．

なお，式(3-6)および式(3-9)において $w_i = 1\ (i = 1, \cdots, n)$，$W = n$ および $\mathbf{W} = \mathbf{I}$（単位行列）とすれば，最小二乗解が得られます．

4.　２次曲線のロバスト推定

ここでは，２次曲線のロバスト推定として「楕円」および「円」の近似を考えます．

係数パラメータ $\mathbf{u} = (a, b, c, d, e, f)^{\mathsf{T}}$ と点 $\mathbf{r} = (x, y)^{\mathsf{T}}$ （$\mathbf{r}^* = (x^2, xy, y^2, x, y, 1)^{\mathsf{T}}$）に対して，２次曲線を

$$F(\mathbf{u};\mathbf{r}) = \mathbf{u}^{\mathsf{T}}\mathbf{r}^* = ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{4-1}$$

と表すことにし，データ数 n の点群 $\mathbf{r}_i = (x_i, y_i)^{\mathsf{T}}\ (i = 1, \cdots, n)$ を近似するパラメータ $\mathbf{u}$ のロバスト推定を考えます．このときのデータ点 $\mathbf{r}_i$ と 2 次曲線 $F(\mathbf{u};\mathbf{r}) = 0$ の距離（残差）ε_i は「代数的距離」，すなわち

$$\varepsilon_i(\mathbf{u}) = F(\mathbf{u};\mathbf{r}_i) = \mathbf{u}^{\mathsf{T}}\mathbf{r}_i^* \tag{4-2}$$

とします．

4.1.　楕円のロバスト推定

２次曲線 式(4-1)が楕円となる条件は $4ac - b^2 > 0$ ですが，不等式を制約条件として解くのは困難であるため，文献[2]にしたがい等式条件である次の制約条件を設けます．

$$\mathbf{u}^{\mathsf{T}}\mathbf{C}\mathbf{u} = \mathbf{u}^{\mathsf{T}}\begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}\mathbf{u} = 1. \tag{4-3}$$

この条件は，$4ac - b^2 = 1$ を表しています．すると，ロバスト推定すなわち式(2-1)の最小化は，式(4-3)の制約条件のもとに行うためラグランジュ乗数 λ' を用いて，$\mathbf{u}$ を未知パラメータとする次の最小化問題に帰着されます．

$$\text{minimize} \quad J(\mathbf{u}) = \sum_i \rho(\varepsilon_i(\mathbf{u})) - \lambda'(\mathbf{u}^\top \mathbf{C}\mathbf{u} - 1)\,. \tag{4-4}$$

この最小化問題を解くために，J をパラメータ $\mathbf{u}$ で微分してゼロと置きます．式(2-4)を用いると，

$$\frac{\partial J}{\partial \mathbf{u}} = \sum_i w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{u}} - 2\lambda' \mathbf{C}\mathbf{u} = \frac{1}{2}\frac{\partial}{\partial \mathbf{u}} \sum_i (w_i^{1/2}\varepsilon_i)^2 - 2\lambda' \mathbf{C}\mathbf{u} = \mathbf{0} \tag{4-5}$$

となります．ここで，

$$\mathbf{D} = \begin{pmatrix} \mathbf{r}_1^{*\top} \\ \vdots \\ \mathbf{r}_i^{*\top} \\ \vdots \\ \mathbf{r}_n^{*\top} \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ & & \vdots & & & \\ x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \\ & & \vdots & & & \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{pmatrix}, \quad \mathbf{W} = diag(w_1, \cdots, w_n)\,, \quad \lambda = 2\lambda' \tag{4-6}$$

とおけば式(4-5)は，

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{u}} &= \frac{1}{2}\frac{\partial}{\partial \mathbf{u}} \sum_i w_i (\mathbf{u}^\top \mathbf{r}_i^*)^2 - 2\lambda' \mathbf{C}\mathbf{u} = \frac{1}{2}\frac{\partial}{\partial \mathbf{u}}\Big((\mathbf{D}\mathbf{u})^\top \mathbf{W}(\mathbf{D}\mathbf{u})\Big) - 2\lambda' \mathbf{C}\mathbf{u} \\ &= \frac{1}{2}\frac{\partial}{\partial \mathbf{u}}\Big(\mathbf{u}^\top(\mathbf{D}^\top \mathbf{W}\mathbf{D})\mathbf{u}\Big) - \lambda \mathbf{C}\mathbf{u} = \frac{1}{2}\cdot 2(\mathbf{D}^\top \mathbf{W}\mathbf{D})\mathbf{u} - \lambda \mathbf{C}\mathbf{u} \\ &= (\mathbf{D}^\top \mathbf{W}\mathbf{D})\mathbf{u} - \lambda \mathbf{C}\mathbf{u} = \mathbf{0} \end{aligned} \tag{4-7}$$

と計算でき，結局つぎの一般固有値問題が得られました．

$$(\mathbf{D}^\top \mathbf{W}\mathbf{D})\mathbf{u} = \lambda \mathbf{C}\mathbf{u}\,. \tag{4-8}$$

この一般固有値問題は次のように解くことにします．$\tilde{\mathbf{D}} = \mathbf{D}^\top \mathbf{W}\mathbf{D}$ とおくと，$\tilde{\mathbf{D}}$ は実対称半正定値行列であり，すべての固有値は非負です．そこで，固有値すべてが正値である場合と，ゼロを含む場合に分けて考察します．

まず，$\tilde{\mathbf{D}}$ の固有値すべてが正値である場合です．このとき，$\tilde{\mathbf{D}} = \mathbf{U}^\top \mathbf{U}$ $(\mathbf{U}:$上三角行列$)$ とコレスキー分解（正値対称行列の LU 分解）ができます．すると，式(4-8)は

$$(\text{左辺}) = \tilde{\mathbf{D}}\mathbf{u} = \mathbf{U}^\top \mathbf{U}\mathbf{u} = \mathbf{U}^\top(\mathbf{U}\mathbf{u}), \quad (\text{右辺}) = \lambda \mathbf{C}\mathbf{u} = \lambda \mathbf{C}\mathbf{U}^{-1}(\mathbf{U}\mathbf{u}) \tag{4-9}$$

と書けます．これより，$\tilde{\mathbf{u}} = \mathbf{U}\mathbf{u}$ および $\tilde{\lambda} = 1/\lambda$ とおいて，

$$\mathbf{U}^\top \tilde{\mathbf{u}} = \lambda \mathbf{C}\mathbf{U}^{-1}\tilde{\mathbf{u}} \qquad \therefore\ ((\mathbf{U}^{-1})^\top \mathbf{C}\mathbf{U}^{-1})\tilde{\mathbf{u}} = \tilde{\lambda}\,\tilde{\mathbf{u}} \tag{4-10}$$

と書き直せば，通常の固有値問題となります．文献[2]の定理 1 によれば，$\tilde{\mathbf{D}}$ が正定値であれば式(4-8)の固有値 λ は一つだけ正値となるものがあり，それに対する固有ベクトルが解となることが示されています．よって，式(4-10)の固有値 $\tilde{\lambda}$ も一つだけ正値となり，それに対する固有ベクトル $\tilde{\mathbf{u}}$ から，

$$\mathbf{u} = \mathbf{U}^{-1}\tilde{\mathbf{u}} \tag{4-11}$$

として解が得られます．

つぎに，$\tilde{\mathbf{D}}$ の固有値にゼロが含まれる場合です．固有値ゼロに対する固有ベクトルを $\mathbf{u}_0$ とすれば，$\tilde{\mathbf{D}}\mathbf{u}_0 = \mathbf{0}$ です．このとき，式(2-6)の評価値を調べてみると，

$$\sum_i w_i \varepsilon_i^2 = \sum_i w_i (\mathbf{u}_0^\top \mathbf{r}_i^*)^2 = \mathbf{u}_0^\top \tilde{\mathbf{D}}\mathbf{u}_0 = 0 \tag{4-12}$$

となります．つまり，係数 $\mathbf{u}_0$ の表す２次曲線は誤差なしで重み付きのデータ点群に当てはまっていることになります．よって，$\mathbf{u}_0 = (a_0, b_0, c_0, d_0, e_0, f_0)^{\mathsf{T}}$ に対して $s = 4a_0c_0 - b_0^2$ が $s > 0$ であれば，

$$\mathbf{u} = \frac{1}{\sqrt{s}}\mathbf{u}_0 \tag{4-13}$$

として解が得られます．

なお，式(4-8)で $\mathbf{W} = \mathbf{I}$（単位行列）とすれば，最小二乗解が得られます．

さて，以上でパラメータ $\mathbf{u} = (a, b, c, d, e, f)^{\mathsf{T}}$ が求まったので，これを使って楕円を表す別のパラメータ すなわち「中心 $\mathbf{c}$，長軸半径 r_a，短軸半径 r_b，長軸傾斜角 θ」を求めることにします[3]．

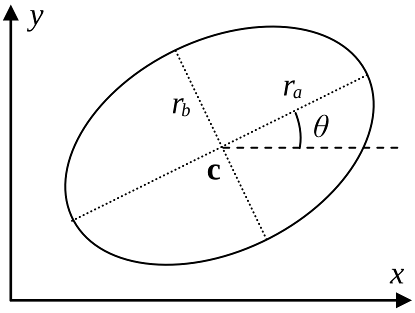

図２．楕円のパラメータ

はじめに，

$$\mathbf{A} = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} d \\ e \end{pmatrix} \tag{4-14}$$

とおいて，２次曲線 式(4-1)を次のように書き直しておきます．

$$\mathbf{r}^{\mathsf{T}}\mathbf{A}\mathbf{r} + \mathbf{b}^{\mathsf{T}}\mathbf{r} + f = 0. \tag{4-15}$$

まず，原点を楕円の中心 $\mathbf{c}$ へもっていくために，$\mathbf{r}' = \mathbf{r} - \mathbf{c}$ として式(4-15)へ代入すると，

$$\mathbf{r}'^{\mathsf{T}}\mathbf{A}\mathbf{r}' + \mathbf{r}'^{\mathsf{T}}(2\mathbf{A}\mathbf{c} + \mathbf{b}) + \mathbf{c}^{\mathsf{T}}(\mathbf{A}\mathbf{c} + \mathbf{b}) + f = 0 \tag{4-16}$$

となります．楕円は中心に関して点対称であるので，$-\mathbf{r}'$ も式(4-16)をみたす必要があります．つまり，１次の項が $2\mathbf{A}\mathbf{c} + \mathbf{b} = \mathbf{0}$ でなくてはなりません．よって，中心 $\mathbf{c}$ が

$$\mathbf{c} = -\frac{1}{2}\mathbf{A}^{-1}\mathbf{b} \tag{4-17}$$

として得られます．このとき，式(4-16)は，

$$\mathbf{r}'^{\mathsf{T}}\mathbf{A}\mathbf{r}' = -\frac{1}{2}\mathbf{c}^{\mathsf{T}}\mathbf{b} - f \tag{4-18}$$

と書き直せるので，両辺を $s = -(\mathbf{c}^{\mathsf{T}}\mathbf{b})/2 - f$ で割って $\tilde{\mathbf{A}} = 1/s \cdot \mathbf{A}$ とおけば，式(4-18)は

$$\mathbf{r}'^{\mathsf{T}}\tilde{\mathbf{A}}\mathbf{r}' = 1 \tag{4-19}$$

となります．

つぎに，$\tilde{\mathbf{A}}$ の固有値 λ_1, λ_2（$\lambda_1 \le \lambda_2$ としておきます）を求めます．楕円では，これら二つの固有値はいずれも正値となります．また，主軸方向は固有値 λ_1, λ_2 に対するそれぞれの固有ベクトル $\mathbf{v}_1, \mathbf{v}_2$ の方向となります．このとき，$\tilde{\mathbf{A}}$ を対角化する直交行列を $\mathbf{T}$ すなわち $\mathbf{T}^{\mathsf{T}}\tilde{\mathbf{A}}\mathbf{T} = diag(\lambda_1, \lambda_2)$ として $\tilde{\mathbf{r}} = \mathbf{T}^{\mathsf{T}}\mathbf{r}' = (\tilde{x}, \tilde{y})^{\mathsf{T}}$ とおけば，式(4-19)はつぎの標準形で書けます．

$$\frac{\tilde{x}^2}{1/\lambda_1} + \frac{\tilde{y}^2}{1/\lambda_2} = 1. \tag{4-20}$$

したがって，長軸半径 r_a および短軸半径 r_b は，

$$r_a = \frac{1}{\sqrt{\lambda_1}}, \quad r_b = \frac{1}{\sqrt{\lambda_2}} \tag{4-21}$$

となります．また，長軸傾斜角 θ は固有ベクトル $\mathbf{v}_1 = (u, v)^{\mathsf{T}}$ を方向ベクトルとする直線が x 軸とな

す角となるので，もし $u<0$ ならば $-\mathbf{v}_1$ を改めて $\mathbf{v}_1$ と置き直すことにより，$u\neq 0$ のとき

$$\theta=\tan^{-1}\frac{v}{u}\quad(-\pi/2<\theta<\pi/2) \tag{4-22}$$

として得られます．$u=0$ のときは $\theta=\pi/2$ です．

4.2. 円のロバスト推定

中心 (x_0, y_0)，半径 r の円における，データ点 $\mathbf{r}_i=(x_i, y_i)^\top$ の代数的距離（残差）ε_i は，

$$\begin{aligned}\varepsilon_i&=(x_i-x_0)^2+(y_i-y_0)^2-r^2\\&=(x_i^2+y_i^2)-(2x_0x_i+2y_0y_i-x_0^2-y_0^2+r^2)\end{aligned} \tag{4-23}$$

となります．ここで，$\mathbf{x}=(2x_0, 2y_0, r^2-x_0^2-y_0^2)^\top$ および $\tilde{\mathbf{r}}_i=(x_i, y_i, 1)^\top$ とおくと，未知パラメータを $\mathbf{x}$ として式(4-23)を，

$$\varepsilon_i(\mathbf{x})=\mathbf{r}_i^\top\mathbf{r}_i-\tilde{\mathbf{r}}_i^\top\mathbf{x} \tag{4-24}$$

と表すことができます．そこで，評価式 式(2-1)を最小化するために，パラメータ $\mathbf{x}$ で微分してゼロとおきます．このとき，

$$\mathbf{v}=\begin{pmatrix}\mathbf{r}_1^\top\mathbf{r}_1\\ \vdots\\ \mathbf{r}_i^\top\mathbf{r}_i\\ \vdots\\ \mathbf{r}_n^\top\mathbf{r}_n\end{pmatrix}=\begin{pmatrix}x_1^2+y_1^2\\ \vdots\\ x_i^2+y_i^2\\ \vdots\\ x_n^2+y_n^2\end{pmatrix},\quad \mathbf{A}=\begin{pmatrix}\tilde{\mathbf{r}}_1^\top\\ \vdots\\ \tilde{\mathbf{r}}_i^\top\\ \vdots\\ \tilde{\mathbf{r}}_n^\top\end{pmatrix}=\begin{pmatrix}x_1 & y_1 & 1\\ & \vdots & \\ x_i & y_i & 1\\ & \vdots & \\ x_n & y_n & 1\end{pmatrix} \tag{4-25}$$

とおくと，式(2-4)は次のように計算できます．

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}}&=\frac{1}{2}\frac{\partial}{\partial \mathbf{x}}\sum_i w_i\varepsilon_i^2=\frac{1}{2}\frac{\partial}{\partial \mathbf{x}}\sum_i w_i(\mathbf{r}_i^\top\mathbf{r}_i-\tilde{\mathbf{r}}_i^\top\mathbf{x})^2\\&=\frac{1}{2}\frac{\partial}{\partial \mathbf{x}}\Big((\mathbf{v}-\mathbf{A}\mathbf{x})^\top\mathbf{W}(\mathbf{v}-\mathbf{A}\mathbf{x})\Big)\\&=\frac{1}{2}\frac{\partial}{\partial \mathbf{x}}\Big(\mathbf{v}^\top\mathbf{W}\mathbf{v}-\mathbf{v}^\top\mathbf{W}\mathbf{A}\mathbf{x}-\mathbf{x}^\top\mathbf{A}^\top\mathbf{W}\mathbf{v}+\mathbf{x}^\top\mathbf{A}^\top\mathbf{W}\mathbf{A}\mathbf{x}\Big)\\&=\frac{1}{2}(-2\mathbf{A}^\top\mathbf{W}\mathbf{v}+2\mathbf{A}^\top\mathbf{W}\mathbf{A}\mathbf{x})=\mathbf{A}^\top\mathbf{W}\mathbf{A}\mathbf{x}-\mathbf{A}^\top\mathbf{W}\mathbf{v}=\mathbf{0}.\end{aligned} \tag{4-26}$$

$$\therefore\quad \mathbf{A}^\top\mathbf{W}\mathbf{A}\mathbf{x}=\mathbf{A}^\top\mathbf{W}\mathbf{v}. \tag{4-27}$$

したがって，解 $\mathbf{x}$ は，

$$\mathbf{x}=(\mathbf{A}^\top\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{W}\mathbf{v} \tag{4-28}$$

として得られます．

なお，式(4-28)で $\mathbf{W}=\mathbf{I}$（単位行列）とすれば，最小二乗解が得られます．

ところで，式(4-27)は最小二乗問題における正規方程式の形をしています．すなわち，重み $w_i=0$ であるデータ点を省き，

$$\sqrt{\mathbf{W}} = diag(\sqrt{w_1}, \cdots, \sqrt{w_n}), \quad \hat{\mathbf{v}} = \sqrt{\mathbf{W}}\,\mathbf{v}, \quad \hat{\mathbf{A}} = \sqrt{\mathbf{W}}\,\mathbf{A} \tag{4-29}$$

とおけば，式(4-28)はつぎの最小二乗問題

$$\hat{\mathbf{v}} = \hat{\mathbf{A}}\mathbf{x} \tag{4-30}$$

の解となっています[4]．最小二乗問題の数値計算法として一般的なのは，式(4-30)を QR 分解で解く方法です. 行列 $\mathbf{Q}$ を $n\times 3$ の長方行列で列ベクトルが直交（ $\mathbf{Q}^\mathsf{T}\mathbf{Q}=\mathbf{I}$，$\mathbf{Q}\mathbf{Q}^\mathsf{T}\neq\mathbf{I}$ ），また, 行列 $\mathbf{R}$ を 3×3 の上三角行列として， $\hat{\mathbf{A}}=\mathbf{QR}$ と分解します．すると，式(4-27)へ代入したとき，左辺において，

$$\mathbf{A}^\mathsf{T}\mathbf{W}\mathbf{A} = \hat{\mathbf{A}}^\mathsf{T}\hat{\mathbf{A}} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{Q}\mathbf{R} = \mathbf{R}^\mathsf{T}\mathbf{R} \tag{4-31}$$

であることより，式(4-27)は,

$$\mathbf{R}^\mathsf{T}\mathbf{R}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\hat{\mathbf{v}}, \qquad \therefore\ \mathbf{R}\mathbf{x} = \mathbf{Q}^\mathsf{T}\hat{\mathbf{v}} \tag{4-32}$$

となります．後は，行列 $\mathbf{R}$ が上三角行列であることより，後退代入で解 $\mathbf{x}$ を求めることができます．以上は概要ですが，詳しくは文献[4]を参照してください.

5. 直線のロバスト推定

２次元以上の空間における直線の方程式は，その上にある１点 $\mathbf{a}$ と方向ベクトル $\mathbf{v}$ を用いて,

$$\mathbf{r} = \mathbf{a} + t\mathbf{v} \quad (t \text{ は実数}) \tag{5-1}$$

と表されます．このとき，データ数 n の点群 $\mathbf{r}_i$ $(i=1,\cdots,n)$ の各点 $\mathbf{r}_i$ と直線との距離を誤差評価値（残差） ε_i とすれば,

$$\varepsilon_i(\mathbf{a},\mathbf{v}) = \left\| (\mathbf{r}_i - \mathbf{a}) - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle \mathbf{v} \right\|, \tag{5-2}$$

ただし， $\langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{v}^\mathsf{T}\mathbf{v} = 1$ （単位ベクトル） (5-3)

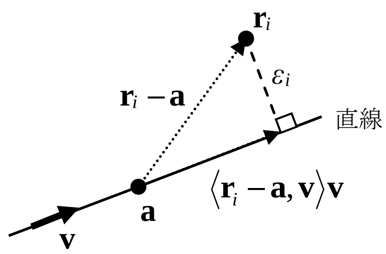

図３．データ点 $\mathbf{r}_i$ の残差 ε_i

と書けます．ここで，$\|\mathbf{x}\|$ はベクトル $\mathbf{x}$ のユークリッドノルム，すなわち $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ です．そこで， $\mathbf{v}$ が単位ベクトルであることに注意して， ε_i を次のように書き換えておきます.

$$\begin{aligned}
\varepsilon_i(\mathbf{a},\mathbf{v}) &= \left\| (\mathbf{r}_i - \mathbf{a}) - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle \mathbf{v} \right\| \\
&= \sqrt{\langle (\mathbf{r}_i - \mathbf{a}) - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle \mathbf{v},\ (\mathbf{r}_i - \mathbf{a}) - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle \mathbf{v} \rangle} \\
&= \sqrt{\langle \mathbf{r}_i - \mathbf{a}, \mathbf{r}_i - \mathbf{a} \rangle - 2\langle \mathbf{r}_i - \mathbf{a}, \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle \mathbf{v} \rangle + \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle^2 \langle \mathbf{v}, \mathbf{v} \rangle^2} \\
&= \sqrt{\langle \mathbf{r}_i - \mathbf{a}, \mathbf{r}_i - \mathbf{a} \rangle - 2\langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle^2 + \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle^2} \\
&= \sqrt{\langle \mathbf{r}_i - \mathbf{a}, \mathbf{r}_i - \mathbf{a} \rangle - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle^2}\ .
\end{aligned} \tag{5-4}$$

すると，ロバスト推定すなわち式(2-1)の最小化は，式(5-3)の制約条件のもとに行うためラグランジュ乗数 λ' を用いて， $\mathbf{a}$ と $\mathbf{v}$ を未知パラメータとする次の最小化問題に帰着されます.

$$\text{minimize} \quad J(\mathbf{a},\mathbf{v}) = \sum_i \rho(\varepsilon_i(\mathbf{a},\mathbf{v})) - \lambda'(\mathbf{v}^\mathsf{T}\mathbf{v} - 1)\,. \tag{5-5}$$

この最小化問題を解くために， J をパラメータ $\mathbf{a}$ および $\mathbf{v}$ で微分してゼロと置きます．まず， $\mathbf{a}$ で微分します．このとき，式(5-5)の右辺第１項の微分は式(2-4)を用います.

$$\frac{\partial J}{\partial \mathbf{a}} = \sum_i w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{a}} = \sum_i w_i \varepsilon_i \frac{-2(\mathbf{r}_i - \mathbf{a}) - 2\langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle (-\mathbf{v})}{2\varepsilon_i}$$
$$= \sum_i w_i (\mathbf{v}\mathbf{v}^\top - \mathbf{I})(\mathbf{r}_i - \mathbf{a}) = (\mathbf{v}\mathbf{v}^\top - \mathbf{I}) \sum_i w_i (\mathbf{r}_i - \mathbf{a}) = \mathbf{0}. \tag{5-6}$$

ただし，$\mathbf{I}$ は単位行列です．ところで，$\mathbf{P} = \mathbf{v}\mathbf{v}^\top$ とおけば，$\mathbf{P}^\top = (\mathbf{v}\mathbf{v}^\top)^\top = (\mathbf{v}^\top)^\top \mathbf{v}^\top = \mathbf{v}\mathbf{v}^\top = \mathbf{P}$ （対称）かつ $\mathbf{P}^2 = (\mathbf{v}\mathbf{v}^\top)(\mathbf{v}\mathbf{v}^\top) = \mathbf{v}(\mathbf{v}^\top\mathbf{v})\mathbf{v}^\top = \mathbf{v}\mathbf{v}^\top = \mathbf{P}$ （べき等）であるから，$\mathbf{P}$ は射影行列です．すると，その固有値は 0 および 1 であり，式(5-6)の最後の等式はベクトル $\sum w_i(\mathbf{r}_i - \mathbf{a})$ が自明な解（零ベクトル）の他に，行列 $\mathbf{v}\mathbf{v}^\top$ の固有値 1 に対する固有ベクトルも解となることを示しています．その固有空間はベクトル $\mathbf{v}$ を基底とする 1 次元空間です．よって，式(5-6)の解は $\sum w_i(\mathbf{r}_i - \mathbf{a}) = t\mathbf{v}$ (t は実数) となります．そこで，$t = 0$ の場合，すなわち $\sum w_i(\mathbf{r}_i - \mathbf{a}) = \mathbf{0}$ を採用すれば，

$$\mathbf{a} = \frac{1}{W} \sum_i w_i \mathbf{r}_i \tag{5-7}$$

が得られます．つまり，$\mathbf{a}$ はデータ点 $\mathbf{r}_i$ の重み付き重心となります．

次に，$\mathbf{v}$ で微分してゼロと置きます．

$$\frac{\partial J}{\partial \mathbf{v}} = \sum_i w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial \mathbf{v}} - 2\lambda' \mathbf{v} = \sum_i w_i \frac{-2\langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle (\mathbf{r}_i - \mathbf{a})}{2\varepsilon_i} - 2\lambda' \mathbf{v}$$
$$= -\left(\sum_i w_i (\mathbf{r}_i - \mathbf{a})(\mathbf{r}_i - \mathbf{a})^\top \right) \mathbf{v} - 2\lambda' \mathbf{v} \tag{5-8}$$
$$= -\left(\mathbf{D}^\top \mathbf{W} \mathbf{D} \right) \mathbf{v} - 2\lambda' \mathbf{v} = \mathbf{0}.$$

ただし，式(5-7)の $\mathbf{a}$ を用いて，

$$\mathbf{D} = \begin{pmatrix} \vdots \\ (\mathbf{r}_i - \mathbf{a})^\top \\ \vdots \end{pmatrix}, \quad \mathbf{W} = diag(w_1, \cdots, w_n) \tag{5-9}$$

とおきました．さらに，$\lambda = -2\lambda'$ および $\mathbf{C} = \mathbf{D}^\top \mathbf{W} \mathbf{D}$ とおけば式(5-8)は，

$$\mathbf{C}\mathbf{v} = \lambda \mathbf{v} \tag{5-10}$$

となり，これは行列 $\mathbf{C}$ の固有値問題です．よって，求めるべき方向ベクトル $\mathbf{v}$ は行列 $\mathbf{C}$ の固有値 λ に対する固有ベクトルとなります．

ところで，式(2-6)の評価値を用いて誤差の総和を評価してみると，

$$\sum_i w_i \varepsilon_i^2 = \sum_i w_i \left(\langle \mathbf{r}_i - \mathbf{a}, \mathbf{r}_i - \mathbf{a} \rangle - \langle \mathbf{r}_i - \mathbf{a}, \mathbf{v} \rangle^2 \right)$$
$$= \sum_i w_i (\mathbf{r}_i - \mathbf{a})^\top (\mathbf{r}_i - \mathbf{a}) - \mathbf{v}^\top \left(\sum_i w_i (\mathbf{r}_i - \mathbf{a})(\mathbf{r}_i - \mathbf{a})^\top \right) \mathbf{v} \tag{5-11}$$
$$= \mathrm{tr}\left(\mathbf{D}^\top \mathbf{W} \mathbf{D} \right) - \mathbf{v}^\top \left(\mathbf{D}^\top \mathbf{W} \mathbf{D} \right) \mathbf{v} = \mathrm{tr}\mathbf{C} - \mathbf{v}^\top \mathbf{C} \mathbf{v} = \mathrm{tr}\mathbf{C} - \lambda \mathbf{v}^\top \mathbf{v}$$
$$= \mathrm{tr}\mathbf{C} - \lambda$$

となります．ただし，$\mathrm{tr}\mathbf{C}$ は行列 $\mathbf{C}$ のトレース（対角成分の和）です．すると式(5-11)は，行列 $\mathbf{C}$ の最大固有値にて最小化が達成されることを示しています．すなわち，方向ベクトル $\mathbf{v}$ は行列 $\mathbf{C}$ の最大固有値に対する固有ベクトルとなります．

なお，式(5-7)および式(5-10)において $w_i = 1 \ (i = 1, \cdots, n)$，$W = n$ および $\mathbf{W} = \mathbf{I}$（単位行列）とすれば，最小二乗解が得られます．

〈参考文献〉

[1] 徐剛, 辻三郎, "３次元ビジョン", 共立出版, 1998.
[2] A. Fitzgibbon, M. Pilu and R. B. Fisher, "Direct Least Square Fitting of Ellipses," IEEE Trans. Pattern Anal.Mach. Intell., vol.21, no.5, pp.476–480, 1999.
[3] 金谷健一, "空間データの数理", 朝倉書店, 1995.
[4] 中川徹, 小柳義夫, "最小二乗法による実験データ解析", 東京大学出版会, 1982.

改訂履歴

Version No.	内　　容
1.0	• 新規発行
1.1	• 誤記の修正
1.2	• 表記法の修正
1.3	• 第3章の修正および第5章の新規追加
1.4	• 第4章の誤記および表記法の修正

以上