幾何学的ＭＤＬとそのメディア応用

金谷 健一[*]　松永 力[†]

[*] 群馬大学工学部情報工学科　[†](株) 朋栄 放送システム開発部

「統計的推測」のデータ数 ∞ の極限での「漸近的」性質が「幾何学的当てはめ」では誤差 0 の「無限小摂動」の極限で成立するという「双対性」に基き、「Rissanen の MDL」の双対として「幾何学的 MDL」が得られることを述べ、その情報理論的解釈を与える。そして平面上の点列への楕円と円と直線の当てはめ、および空間中の点列への直線と平面の当てはめのシミュレーションを行い、幾何学的 AIC と幾何学的 MDL の特性がどのように異なるかを観察する。最後にシミュレーション実験および実画像による仮想スタジオへの応用実験を行い、幾何学的 AIC のほうがデータにより忠実であり、幾何学的 MDL のほうが退化したモデルを選びやすいことを示す。

キーワード: 統計的推測、幾何学的当てはめ、幾何学的 AIC、幾何学的 MDL、モデル選択、仮想スタジオ

Geometric MDL and Its Media Applications

Kenichi Kanatani[*] and Chikara Matsunaga[†]

[*]Department of Computer Science, Gunma University, Kiryu, Gunma 376-8515 Japan
[†]Broadcast Division, FOR-A Co. Ltd., 2-3-3 Ohsaku, Sakura, Chiba 285-0802 Japan

Based on the duality that the *asymptotic* properties of *statistical inference* in the limit of an infinite number of observations hold for *geometric fitting* in the limit of *infinitesimal perturbations*, we define the *geometric MDL* as the dual to *Rissanen's MDL* and give a rigorous information-theoretic interpretation. We then do simulations for fitting an ellipse, a circle, and aline to points in two dimensions and a line and a plane to points in three dimensions examples and observe that the geometric AIC and the geometric MDL exhibit different performance characteristics. Finally, we do virtual studio experiments using real images and observe that the geometric AIC is more faithful to the data than the geometric MDL, which is more likely to choose degenerate models.

Key words: statistical inference, geometric fitting, geometric AIC, geometric MDL, model selection, virtual studio

謝辞: 本研究の一部は文部省科学研究費基盤研究Ｃ (2) (No. 11680377) によった。本論文の内容に関して有益な議論を頂いたオーストラリアのモナシュ大学の David Suter 教授、英国レディング大学の Steve Maybank 教授、理化学研究所の甘利俊一博士、電気通信大学の韓太舜教授、群馬大学の関庸一助教授に感謝します。

[*]376-8515 桐生市天神町 1–5–1　群馬大学工学部情報工学科
Tel: (0277)30-1841, Fax: (0277)30-1801, E-mail: kanatani@cs.gunma-u.ac.jp
[†]285-0802 千葉県佐倉市大作 2–3–3 (株) 朋栄　放送システム開発部
Tel: (043)498-1239, Fax: (043)498-2223, E-mail: matsunaga@for-a.co.jp

1. 序論

前報 [15] では「統計的推測」のデータ数 ∞ の極限での「漸近的」性質が「幾何学的当てはめ」では誤差 0 の「無限小摂動」の極限で成立するという「双対性」を示した。そして、確率現象の発生機構を説明する「確率モデル」とデータ空間中の多様体としての「幾何学的モデル」を対比することにより (赤池の)「AIC」と「幾何学的 AIC」の導出や誤差の推定方式がどのように異なるかを示した。

統計的推測において赤池の AIC とともによく知られているものに (Rissanen の)「MDL」[28, 29] がある。そして、コンピュータビジョンやロボット工学の分野でも従来から MDL を用いると称する研究もあった。しかし、統計的推測の形の問題に Rissanen の MDL をそのまま適用したものを別とすれば [23]、統計的推測と幾何学的当てはめを区別しなかったり [31]、単に記述の短い解を得るというものが多かった [4, 19, 25]。

本論文では統計的推測と幾何学的当てはめの双対性を利用して Rissanen の MDL の双対として「幾何学的 MDL」を定義し、その情報理論的解釈を与える。そしてシミュレーション実験により幾何学的 AIC と幾何学的 MDL の特性を比較するとともに、仮想スタジオへの応用実験を行い、幾何学的 AIC のほうがデータにより忠実であり、幾何学的 MDL のほうが退化したモデルを選びやすいことを示す。

2. 幾何学的当てはめ

前報 [15] で幾何学的当てはめを次のように定式化した。 N 個の m 次元データ $\{\boldsymbol{x}_\alpha\}$ が与えられ、各 $\boldsymbol{x}_\alpha$ は真の位置 $\bar{\boldsymbol{x}}_\alpha$ から期待値 $\boldsymbol{0}$、共分散行列 $V[\boldsymbol{x}_\alpha]$ の独立な正規分布に従う誤差だけずれているとする。そして、それらの真の位置 $\bar{\boldsymbol{x}}_\alpha$ は p 次元ベクトル $\boldsymbol{u}$ でパラメータ化された次の形の r 個の拘束条件を満たすとする。

$$F^{(k)}(\bar{\boldsymbol{x}}_\alpha, \boldsymbol{u}) = 0, \qquad k = 1, ..., r \tag{1}$$

データの共分散行列 $V[\boldsymbol{x}_\alpha]$ は定数倍を除いて既知とし、次のように書けるとする。

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha] \tag{2}$$

パラメータの $\boldsymbol{u}$ の最尤推定は $V_0[\boldsymbol{x}_\alpha]$ に関するマハラノビス距離の 2 乗和

$$J = \sum_{\alpha=1}^{N} (\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha, V_0[\boldsymbol{x}_\alpha]^{-1}(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha)) \tag{3}$$

を拘束条件 (1) のもとで最小化すればよい。ただし、ベクトル $\boldsymbol{a}, \boldsymbol{b}$ の内積を $(\boldsymbol{a}, \boldsymbol{b})$ と記す。

3. Rissanen の MDL

Rissanen の MDL は情報理論的な符号長をモデルのよさの尺度とするものであり、データ長 ∞ の確率密度の漸近解析から導出される。したがって、漸近解析が意味を持たず、確率構造とは無関係の幾何学的モデルには転用できない。赤池の AIC が本質的に期待値評価であり、したがって幾何学的モデル選択にも容易に転用できるのと著しい対照である。

このことは Rissanen の MDL が対数尤度の項と $k \log n$ の項から成っていることに現れている。 k はモデルのパラメータ数、 n は「データ長」とされているが、データ長とは何かについて解説書 (例えば [2, 33]) には何も書かれていない。例えば平面上に 10 点 (x_1, y_1), ..., (x_{10}, y_{10}) が与えられたとき、 $n = 10$ なのか、それとも実数データとして $n = 20$ なのか。あるいは 20 次元直積空間に発生した 1 個のデータとみなして $n = 1$ であろうか。

解説書の導出をよく検討すると、 n とは「全データのフィッシャー情報行列の $1/n$ が $n \to \infty$ で定数行列に収束するような数」であることがわかる。そのような数なら何でもよい。平面上の点の例では $n = 10$ でも $n = 5$ でもよい。なぜなら Rissanen の MDL は漸近評価式であり、 $n \to \infty$ のとき $O(\log n)$ 以下で増加する量は無視されるからである [2, 33]。例えば n と $2n$ とでは $\log 2n = \log n + \log 2 \approx \log n$ である。

すなわち Rissanen の MDL を定義する「データ長」がそもそも漸近的な概念である。独立な観測の場合は n を観測数とも解釈できるが、観測数の任意の (固定した) 定数倍でもよい。そして Rissanen の MDL は定数 c に対して $\log cn = \log n + \log c \approx \log n$ であるような n に対して適用される。

よく考えるとこれは当然である。実数を扱う問題では実数を完全に符号化するのに無限大の符号長を要する。しかし、実数は完全には推定できないと考えれば適当な量子化によって有限長で記述できる。その量子化幅は推定精度が上がるにつれて小さくすればよい。統計的推定ではデータが多いほど推定精度が上がるので、 “データ長”n に応じた最適な量子化幅を用いればよい。データ長 n を無限に増やすと記述長が ∞ になるので、漸近評価が必要となる。要するに MDL は最短 “記述長” というより最短 “記述長増大速度” である。

4. 幾何学的 MDL

前報 [15] に述べたように、幾何学的当てはめでは観測数は常に 1 であり、したがって観測数 ∞ でのみ意味をもつ Rissanen の MDL は適合しない。しかし、 $n \to \infty$ の統計的推測と $\epsilon \to 0$ の幾何学的当てはめとの双対性に着目すると、 “時間” ϵ の $\epsilon \to 0$ に対する “記述長増大速度” を評価すればよい。

前報 [15] で考えた仮想的な漸近解析によると、仮

想的観測数 n が増えればノイズレベル ϵ が実質的に $O(1/\sqrt{n})$ になる。したがって $n = O(1/\epsilon^2)$ である。Rissanen の MDL は負の対数尤度に $(k/2)\log n$ が加わったものであり、誤差が正規分布の場合の負の対数尤度は $J/2\epsilon^2$ と書けるから、その "双対" が次のように予想される。

$$\text{G-MDL} = \hat{J} - (Nd + p)\epsilon^2 \log \epsilon^2 \tag{4}$$

これを幾何学的 **MDL** と呼ぼう[1]。もちろん幾何学的当てはめにおいては ϵ^2 はモデルパラメータではないから、一般モデルから次のように推定する [15]。

$$\hat{\epsilon}^2 = \frac{\hat{J}}{rN - p} \tag{5}$$

5. 情報理論的解釈

式 (4) の幾何学的 MDL が実際に ϵ に関する最短記述長増大速度と解釈できることを示そう。幾何学的当てはめ問題は (i) データ $\{\boldsymbol{x}_\alpha\}$、(ii) それらの真の位置 $\{\bar{\boldsymbol{x}}_\alpha\}$、(iii) パラメータ $\boldsymbol{u}$ とで指定されるので、これらを符号化すればよい。

(i) データ $\{\boldsymbol{x}_\alpha\}$ の符号化。データの密度関数は次のように書ける。

$$P = \prod_{\alpha=1}^{N} \frac{e^{-(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha, V_0[\boldsymbol{x}_\alpha]^{-1}(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha))/2\epsilon^2}}{(2\pi\epsilon^2)^m \det V_0[\boldsymbol{x}_\alpha]} = \frac{e^{-J/2\epsilon^2}}{(2\pi\epsilon^{2m})^N \prod_{\alpha=1}^{N} \det V_0[\boldsymbol{x}_\alpha]} \tag{6}$$

式 (6) の密度関数によってデータ $\{\boldsymbol{x}_\alpha\}$ を最短語頭符号化するにはほぼ次の長さが必要となる[2] [2, 33]。

$$-\log P = \frac{J}{2\epsilon^2} + N\log(2\pi\epsilon^{2m}) + \sum_{\alpha=1}^{N} \log\det V_0[\boldsymbol{x}_\alpha] \tag{7}$$

右辺の最後の 2 項は (幾何学的) モデルに依存しない定数である。

(ii) 真の位置 $\{\bar{\boldsymbol{x}}_\alpha\}$ の符号化。真の位置はデータに誤差があるかぎり厳密には求まらないのでその最尤推定量 $\hat{\boldsymbol{x}}_\alpha$ を符号化する [2, 33]。$\hat{\boldsymbol{x}}_\alpha$ はデータ $\boldsymbol{x}_\alpha$ をその共分散行列 $V_0[\boldsymbol{x}_\alpha]$ の定義するマハラノビス距離が最小になるように、拘束条件の定義する d 次元モデル多様体 $\mathcal{S}$ へ射影した点である [11]。その点の共分散行列 $V[\hat{\boldsymbol{x}}_\alpha]$ は次のようになる[3] [9, 11]。

$$V[\hat{\boldsymbol{x}}_\alpha] = \epsilon^2 V_0[\hat{\boldsymbol{x}}_\alpha] + O(\epsilon^4) \tag{8}$$

$V_0[\hat{\boldsymbol{x}}_\alpha]$ は次のランク d の $m \times m$ 特異半正値対称行列である。

$$V_0[\hat{\boldsymbol{x}}_\alpha] = V_0[\boldsymbol{x}_\alpha] - \sum_{k,l=1}^{r} W_\alpha^{(kl)} (V[\boldsymbol{x}_\alpha]\nabla_{\mathbf{x}} F_\alpha^{(k)})(V[\boldsymbol{x}_\alpha]\nabla_{\mathbf{x}} F_\alpha^{(k)})^\top \tag{9}$$

ただし $W_\alpha^{(kl)}$ は $(\nabla_{\mathbf{x}} F^{(k)}, V[\boldsymbol{x}_\alpha]\nabla_{\mathbf{x}} F^{(l)})$ を (kl) 要素とする $r \times r$ 行列の逆行列の (kl) 要素である [15]。$\hat{\boldsymbol{x}}_\alpha$ の精度はほぼ ϵ に比例するので、$\bar{\boldsymbol{x}}_\alpha$ の定義域の d 次元モデル多様体 $\mathcal{S}$ を幅 $c\epsilon$ で刻み、その格子点を指定すればよい。その符号長は次のようになる。

$$(\text{定数}) - \log(c\epsilon)^d = (\text{定数}) - d\log\epsilon \tag{10}$$

N 個データがあるから、合計は次のようになる。

$$(\text{定数}) - Nd\log\epsilon \tag{11}$$

(iii) パラメータ $\boldsymbol{u}$ の符号化。これも真の値はデータに誤差があるかぎり厳密には求まらないので、その最尤推定量 $\hat{\boldsymbol{u}}$ を符号化する。その共分散行列 $V[\hat{\boldsymbol{u}}]$ は次のようになる[4] [8, 11]。

$$V[\hat{\boldsymbol{u}}] = \epsilon^2 V_0[\hat{\boldsymbol{u}}] + O(\epsilon^4) \tag{12}$$

$$V_0[\hat{\boldsymbol{u}}] = \Big(\sum_{\alpha=1}^{N} \sum_{k,l=1}^{r} W_\alpha^{(kl)} (\nabla_{\mathbf{u}} F_\alpha^{(k)})(\nabla_{\mathbf{u}} F_\alpha^{(l)})^\top\Big)^{-1} \tag{13}$$

$\bar{\boldsymbol{x}}_\alpha$ の場合と同様に、$\boldsymbol{u}$ の p 次元パラメータ空間を幅 $C\epsilon$ で刻めば、格子点を指定する符号長は次のようになる。

$$(\text{定数}) - \log(C\epsilon)^p = (\text{定数}) - p\log\epsilon \tag{14}$$

式 (7) でモデルに依存しない定数を無視し、関数 J を $\hat{J}$ に置き換えても差は ϵ の高次の項であることを考慮すると、全体の符号長は次のようにみなせる。

$$\frac{\hat{J}}{2\epsilon^2} - (Nd + p)\log\epsilon \tag{15}$$

これを $2\epsilon^2$ 倍すると式 (4) が得られる[5]。

6.. 最適量子化

前節の解析はおおざっぱな評価であるが、Rissanen [2, 29, 33] にならって、$\hat{\boldsymbol{x}}_\alpha$ の分割幅 $c\epsilon$ の係数 c

[1] AIC の場合と同様に、幾何学的当てはめに MDL と G-MDL の 2 種類が定義されるわけではないので、単に MDL と書いてもよい。

[2] 情報理論的には log の底は 2 であるが、符号長を測るのに $\log_2 e$ ビットを 1 とする単位を用いれば自然対数とみなしてよい。

[3] 最後の $O(\epsilon^4)$ を除いたものが統計的推測におけるクラメル・ラオの下界に相当する共分散行列の下界であり、誤差が正規分布に従う最尤推定では $O(\epsilon^4)$ を除いてこれが達成される [11]。

[4] 最後の $O(\epsilon^4)$ を除いたものが統計的推測におけるクラメル・ラオの下界に相当する共分散行列の下界であり、誤差が正規分布に従う最尤推定では $O(\epsilon^4)$ を除いてこれが達成される [11, 13]。

[5] 赤池の AIC や Rissanen の MDL にデータ数 n の式を掛けたり足したりしても、n は各モデルに共通であるから相対的な比較に影響しない。同様に、幾何学的 AIC や幾何学的 MDL にノイズレベル ϵ の式を掛けたり足したりしても、ϵ は各モデルに共通であるから相対的な比較に影響しない。

をデータごとおよび方向ごとに変え、$\hat{\boldsymbol{u}}$ の分割幅 $C\epsilon$ の係数 C も座標ごとに変え、それらを全符号長が最小になるように最適に量子化することを考える。

基礎となるのは、式 (8), (12) から式 (3) が次のように書けるという事実である [8, 11]。

$$J = \hat{J} + \sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha - \hat{\boldsymbol{x}}_\alpha, V_0[\hat{\boldsymbol{x}}_\alpha]^-(\boldsymbol{x}_\alpha - \hat{\boldsymbol{x}}_\alpha)) + (\boldsymbol{u} - \hat{\boldsymbol{u}}, V_0[\hat{\boldsymbol{u}}]^{-1}(\boldsymbol{u} - \hat{\boldsymbol{u}})) + O(\epsilon^3) \tag{16}$$

ただし $\hat{J}$ は J に $\boldsymbol{x}_\alpha = \hat{\boldsymbol{x}}_\alpha$, $\boldsymbol{u} = \hat{\boldsymbol{u}}$ を代入した値であり、$V_0[\hat{\boldsymbol{x}}_\alpha]^-$ は (ムーア・ペンローズの) 一般逆行列である。

まず $\boldsymbol{u}$ について考える。$\boldsymbol{u}$ の p 次元パラメータ空間 $\mathcal{R}^p$ を各座標軸方向に幅 δu_i, $i = 1, ..., p$ で分割すれば、その頂点を指定する符号長は定数項を除いて $-\sum_{i=1}^{p} \log \delta u_i$ となる [2, 33]。したがって式 (16) から式 (7) の符号長の $\boldsymbol{u}$ に関係する部分が次のように見積もれる。

$$\frac{(\delta\boldsymbol{u}, V_0[\hat{\boldsymbol{u}}]^{-1}\delta\boldsymbol{u})}{2\epsilon^2} - \sum_{i=1}^{p}\log \delta u_i \tag{17}$$

ただし $\delta\boldsymbol{u} = (\delta u_i)$ である。δu_i で微分して 0 と置くと次式を得る。

$$\frac{1}{\epsilon^2}\Big(V_0[\hat{\boldsymbol{u}}]^{-1}\delta\boldsymbol{u}\Big)_i = \frac{1}{\delta u_i} \tag{18}$$

パラメータ空間 $\mathcal{R}^p$ の座標系を $V_0[\hat{\boldsymbol{u}}]$ が対角行列になるようにとれば解 $\delta u_i = \epsilon\sqrt{\lambda_i}$ を得る。ただし λ_i は $V_0[\hat{\boldsymbol{u}}]$ の第 i 固有値である。これを代入し、$\prod_{i=1}^{p}\lambda_i = \det V_0[\hat{\boldsymbol{u}}]$ に注意すると式 (17) は次のようになる。

$$\frac{p}{2} - \frac{p}{2}\log\epsilon^2 - \frac{1}{2}\log\det V_0[\hat{\boldsymbol{u}}] \tag{19}$$

ϵ が小さいとき $-\log\epsilon \gg 1$ であり、主要項のみ残して $2\epsilon^2$ 倍すると式 (4) 中の項 $-p\epsilon^2\log\epsilon^2$ が得られる。

同様に各 $\boldsymbol{x}_\alpha$ の定義域の d 次元モデル多様体 $\mathcal{S}$ を離散化し、式 (16) の第 2 項と合わせて最適符号化を行って $2\epsilon^2$ 倍すれば式 (4) 中の項 $-Nd\epsilon^2\log\epsilon^2$ が得られる。

結果として、このような厳密な解析[6]をしても幾何学的 MDL には反映されず、実質的に用いたのは “推定量の誤差がほぼ ϵ に比例する” という事実のみである[7]。そして、幾何学的 MDL は定数 c に対して $-\log c\epsilon = -\log\epsilon - \log c \approx -\log\epsilon$ であるような小さいノイズレベル ϵ に対して適用されることもわかる。

[6] 統計的推定ではこのような解析は確率的複雑度と呼ばれ、種々の学習問題の評価として用いられているそうである [32]。本節の解析は幾何学的当てはめに対する “双対” 確率的複雑度とも解釈できる。

[7] 結果的には赤池 [1] の指摘するように、Rissanen の最適量子化は log の項を出すための言い訳であると思えなくもない。

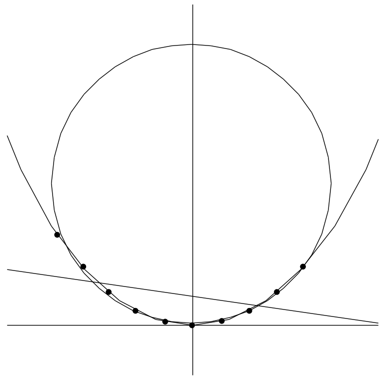


図 1: 楕円と円と直線の当てはめ。

7. 平面上の直線と円の検出

x 軸に原点で接して y 軸方向の半径が 50、離心率が $1/\beta$ の楕円上に x 座標が等間隔の 11 個の点をとり、x, y 座標に期待値 0、標準偏差 σ の正規乱数誤差を加えて、直線、円、楕円を最適に当てはめる。この当てはめは「くりこみ法」と呼ぶ手法により効率的に計算できる [6, 11, 16, 17]。図 1は $\beta = 2.5$、$\sigma = 0.1$ の場合の一例である。

直線、円、楕円はどれも 1 次元であり、自由度はそれぞれ 2, 3, 5 であるから、N 点に対する幾何学的 AIC と幾何学的 MDL はそれぞれ次のようになる。

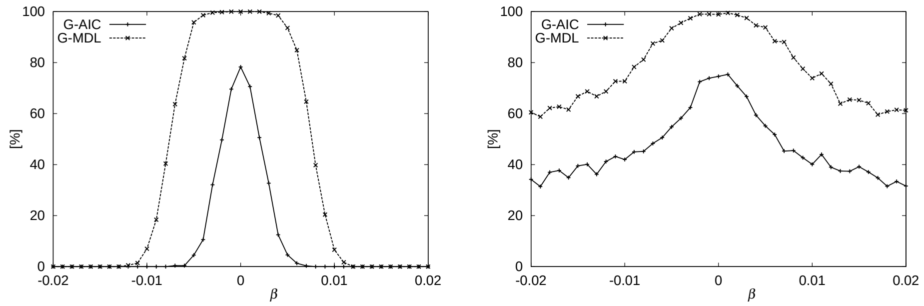
$$\begin{aligned}
\text{G-AIC}_l &= \hat{J}_l + 2(N+2)\sigma^2 \\
\text{G-AIC}_c &= \hat{J}_c + 2(N+3)\sigma^2 \\
\text{G-AIC}_e &= \hat{J}_e + 2(N+5)\sigma^2 \\
\text{G-MDL}_l &= \hat{J}_l - (N+2)\sigma^2\log\sigma^2 \\
\text{G-MDL}_c &= \hat{J}_c - (N+3)\sigma^2\log\sigma^2 \\
\text{G-MDL}_e &= \hat{J}_e - (N+5)\sigma^2\log\sigma^2
\end{aligned} \tag{20}$$

添字の l, c, e はそれぞれ直線、円、楕円を意味する。これを β をいろいろ変えながら計算し、幾何学的 AIC および幾何学 MDL が小さいモデルを選ぶ。

図 2(a) は $\sigma = 0.01$ の場合に $\beta = 0$ の近傍で独立に誤差を変えて 1000 回試行し、直線と判定された割合 (%) である。誤差がなければ $\beta \neq 0$ で 0%、$\beta = 0$ で 100% となる。誤差があるとき、幾何学的 AIC は直線の場合にピークを持ち、直線を楕円と区別する能力が高いが、直線の場合でも一定の割合で楕円と認識する。それに対して幾何学的 MDL は直線を直線と判定する割合が誤差が小さいと 100% に近いが、直線でない場合も広く直線と認識する。

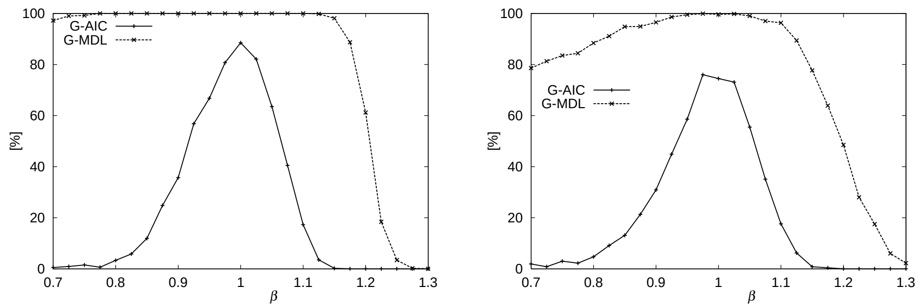
図 2(a) では誤差の標準偏差 σ に真の値を用いたが、これを一般モデルである楕円の当てはめの残差から推定したものが図 2(b) である。ばらつきが増大しているが、やはり同じ傾向が見られる。

図 3は同様に $\sigma = 0.01$ の場合に $\beta = 1$ の近傍で円



(a) (b)

図 2: 幾何学的 AIC(実線) と幾何学的 MDL(破線) による直線の検出率 (%)。(a) 既知のノイズレベル。(b) 推定したノイズレベル。



(a) (b)

図 3: 幾何学的 AIC(実線) と幾何学的 MDL(破線) による円の検出率 (%)。(a) 既知のノイズレベル。(b) 推定したノイズレベル。

と判定された割合 (%) である。誤差がなければ $\beta \neq 1$ で 0%、$\beta = 1$ で 100% となる。誤差があると $\beta < 1$ の短い弧は楕円弧か円弧かの区別が難しいが、幾何学的 AIC は明確に円を楕円と区別する。ただし円でも一定の割合で楕円と認識する。これに対して幾何学的 MDL は $\beta < 1.1$ ではどんな楕円もほとんど常に円と認識する。

8. 空間直線の検出

幾何学的当てはめにおける「幾何学的モデル」は統計的推測における確率密度としての「確率モデル」とは異なり、拘束条件で与えられるデータ空間内の多様体である。このためモデルの自由度 (パラメータ数) だけでなく、その "次元" が重要な役割を果たす [10, 11, 14]。前節の例では楕円も円も直線も 1 次元モデルで次元に差はなく、差は自由度のみに現れた。そこで次元の差があるモデルとして、空間に分布した点に直線および平面を当てはめる。

xyz 空間の xy 面内の長方形領域 $[0, 10] \times [-1, 1]$ に 11 点をランダムにとり、y 軸方向に A 倍に拡大する。そして、各点の x, y, z 座標に期待値 0、標準偏差 σ の正規乱数誤差を加えて、平面および直線を最適に当てはめる (計算手順は [7, 11] 参照)。これを A をいろいろに変えながら各 A に対して幾何学的 AIC と幾何学的 MDL を計算し、平面が直線に退化しているとみなせるかどうかを判定する。

空間直線は 1 次元、平面は 2 次元であり、自由度はそれぞれ 4, 3 であるから N 点に対する幾何学的 AIC と幾何学的 MDL はそれぞれ次のようになる。

$$\begin{aligned} \text{G-AIC}_l &= \hat{J}_l + 2(N+4)\sigma^2 \\ \text{G-AIC}_p &= \hat{J}_p + 2(2N+3)\sigma^2 \\ \text{G-MDL}_l &= \hat{J}_l - (N+4)\sigma^2 \log \sigma^2 \\ \text{G-MDL}_p &= \hat{J}_p - (2N+3)\sigma^2 \log \sigma^2 \end{aligned} \qquad (21)$$

添字の l, p はそれぞれ空間直線および平面を意味する。

図 4(a) は $\sigma = 0.01$ の場合に $A = 0$ の近傍で独立に誤差を変えて 1000 回試行し、直線であると判定された割合 (%) である。誤差がなければ $A \neq 0$ で 0%、$A = 0$ で 100% となる。誤差があるとき、幾何学的 AIC は直線を平面と区別する能力が高いが、直線の場合でも一定の割合で平面と認識する。それに対して幾何学的 MDL は直線を直線と判定する割合が誤差が小さいと 100% に近いが、直線でない場合も広く直線と認識する。図 4(a) では誤差の標準偏差 σ に真の値を用いているが、これを一般モデルである平面の当てはめの残差から推定したものが図 4(b) である。ばらつきが増大しているが、やはり同じような傾向が見られる。

9. 幾何学的 AIC vs. 幾何学的 MDL

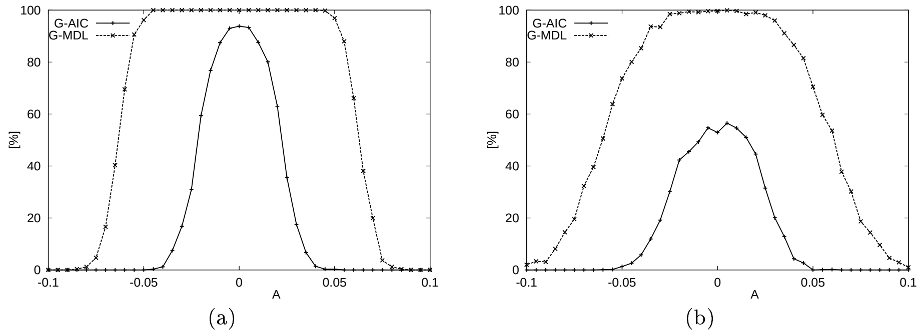


図 4: 幾何学的 AIC(実線) と幾何学的 MDL(破線) による空間直線の検出率 (%)。 (a) 既知のノイズレベル。 (b) 推定したノイズレベル。

以上の実験からわかるように、幾何学的 AIC は退化モデルを非退化モデルと区別する能力が高いが、退化の場合も一定の割合で非退化モデルを選ぶ。それに対して幾何学的 MDL は退化を退化と判定する割合が誤差 0 の無限小摂動の極限で 100% に近づく。これが統計的推測において **MDL** の一致性、すなわち Rissanen の MDL が真のモデルを選ぶ割合がデータ数 ∞ の極限で 100% に収束することの幾何学的当てはめにおける双対命題である。

一致性の代償として、幾何学的 MDL は退化でない場合も幅広く退化と判定し、幾何学的 AIC と比べてより安全側に偏っている。これは幾何学的 MDL のモデル自由度に対するペナルティが幾何学的 AIC より大きいことから当然である。このため幾何学的 AIC のほうがデータにより忠実であり、幾何学的 MDL のほうが退化モデルを選びやすい。

10. 仮想スタジオへの応用

仮想スタジオとは人物を撮影したビデオ画像から人物像を切り出し、グラフィクスで生成した背景と合成する技術である [3, 24, 27, 30]。ビデオカメラを任意に移動しながらズームを任意に変えて撮影するので、適合した背景を生成するにはフレームごとにカメラの向きや位置やズームを知らなければならない。

これを知る一つの方法は濃紺と淡青に塗り分けた格子パタンを背景に置き、クロマキー手法によって人物像の背景からパタンを抽出し、それからカメラの向きや位置やズームを計算することである。原パタンの形状や位置は既知であるから、原理的には 4 個以上の格子点が原パタンと対応づけられればこれが可能となる [18, 20, 21]。しかし、次の二つの問題が生じる。

一つはカメラがパタンに正対する位置に来ると、カメラの前進とズームとが区別できず、計算が破綻することである。もう一つは、人物が移動すると背後に見える部分が異なり、例えカメラが静止していてもフレームごとに解が多少異なり、計算上カメラが振動しているという解釈が得られることである。

これに対して従来は発見的な対処がなされてきたが、これは幾何学的モデル選択によって解決できる。具体的には、通常の撮影で生じやすいカメラの移動やズームの仕方をモデル化しておき、各フレームごとに最も妥当なモデルを選択すればよい [22]。

図 5(a) は 3×3 格子パタンと垂直な平面上でカメラを、格子の中心を注視させながら焦点距離を徐々に変えて移動させたシミュレーションである。位置 13 でカメラがパタンに正対し、位置 20 $\sim$ 24 で停止し、位置 25 $\sim$ 30 で後退する。格子画像の頂点に標準偏差 1 画素の正規分布に従う誤差を毎回独立に加えて焦点距離とカメラの経路を計算した。ここでは次のようなモデルを用いた (計算の詳細は [22] 参照)。

- カメラは静止し、ズームは固定。
- ズーム固定でカメラが回転。
- ズーム固定でカメラは直進運動。
- ズーム固定でカメラは任意運動。
- ズームは線形に変化し、カメラは任意運動。
- すべてが変化。

図 5(b), (c) は焦点距離およびカメラの経路を示したものである。位置 13, 14 で退化が生じ、モデル選択を用いない方法では計算が破綻している。破綻したことを見やすくするために図 5(b) では f を ∞、図 5(c) ではカメラ位置をパタンの中心として表示している。

図 5(d) は位置 20 $\sim$ 24 付近の拡大図である。このようにフレーム毎に独立に計算すると、実際にはカメラが静止していても計算値は不規則にゆらいでいるが、モデル選択によるとゆらぎが除去されている。

図 6は実画像列である (代表的な 5 画像のみを表示)。図 7(a), (b) はこれから計算した焦点距離とカメラの軌跡である。フレームごとの独立な推定は 15 フレーム目で破綻している。そして、幾何学的 AIC の解はデータにより忠実であるという意味の “柔軟さ” があり、幾何学的 MDL の解は多少データを無視しても単純なモデルを選ぶという “剛直さ” があることがわかる。

この場合は幾何学的 MDL のほうがより実際のカメラの運動に近い解を与えているが、これはズームを固

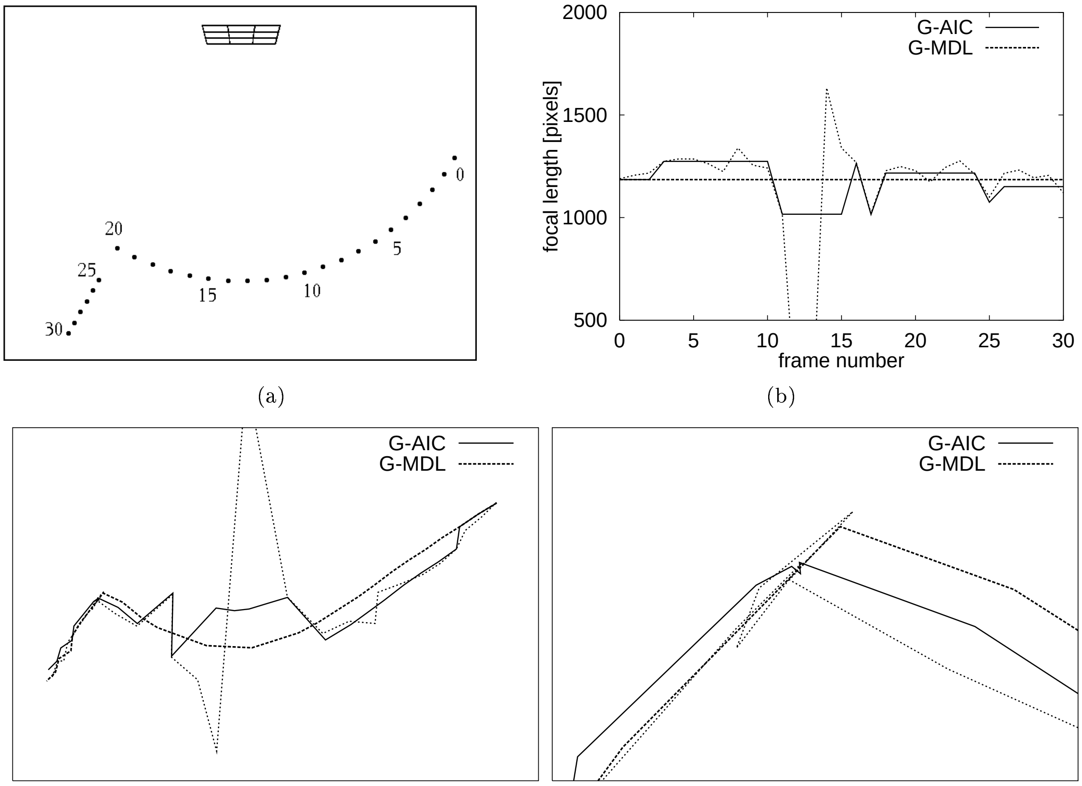


図 5: (a) カメラの経路。 (b) 推定した焦点距離。 (c) 推定したカメラの経路。 (d) 径路の拡大図。 (b)–(d) で実線は幾何学的 AIC、破線は幾何学的 MDL、点線はモデル選択を用いないもの。

定してカメラを滑らかに動かしたためである。ズームやカメラ位置を変動させながら撮影しても幾何学的 MDL はやはり滑らかな解を選ぶから、その場合には実際のカメラ運動とは異なる解が得られる。したがって、幾何学的 AIC と幾何学的 MDL のどちらが正解に近いかを論じることはあまり意味がなく、判断の基準は “その応用ではどのような解が望ましいと期待するか” による。

11. まとめ

本論文では統計的推測と幾何学的当てはめの双対性を利用して Rissanen の MDL の双対として「幾何学的 MDL」を定義し、その情報理論的解釈を与えた。そしてシミュレーション実験を行って幾何学的 AIC と幾何学的 MDL の特性を比較するとともに、仮想スタジオへの応用実験を行い、幾何学的 AIC のほうがデータにより忠実であり、幾何学的 MDL のほうが退化したモデルを選びやすいことを示した。結論として、幾何学的 AIC と幾何学的 MDL のどちらがよいかはその応用目的によることがわかった。

最近、長尾ら [26] も「幾何学的 AIC」、「幾何学的 MDL」と呼ぶ基準を定義しているが、これらは隠れ変数の導入により $N \to \infty$ の統計的推測として定式化したものである。したがって彼らの MDL が $\log N$ の項を含んでいる (本論文の MDL は $\log \epsilon$ の項を含んでいる)。また、期待値や分散のような誤差の特性もモデルパラメータとして含んでいるために、それらをモデルごとに独立に推定しなければならない。

本論文では例として直線、円、楕円の当てはめを実験したが、長尾らの基準を適用しようとすれば、曲線上に何らかの分布を仮定してそのパラメータまで最適に計算しなければならない。円や楕円のような有限の曲線上には正規分布は定義できず、どう定義するか難しい。さらにそのパラメータの推定は直線上の正規分布の場合ですら複雑で、曲線や一般の場合に計算アルゴリズムを与えることは極めて困難である。

それに対して本論文の基準は誤差の特性をモデルパラメータとしていないため、推定は純粋な幾何学的問題であり、「くりこみ法」[6] 等、さまざまなツールが利用できる。例えば動画像解析では「エピ極線拘束条件」の定義する 4 次元空間中の 3 次元代数多様体を最適に当てはめなければならないが [5, 7, 11]、その場合にも簡単に適用できる [12]。このような応用ではデータ点数は多くないので長尾らの基準の前提は満たされず、具体的な計算法を与えることも困難である。

参考文献

[1] 赤池弘次, AIC と MDL と BIC, オペレーションズ・リサーチ, **41**-7 (1996), 375–378.

[2] 韓太舜, 小林欣吾, 「情報と符号化の数理」, 培風館, 1999.

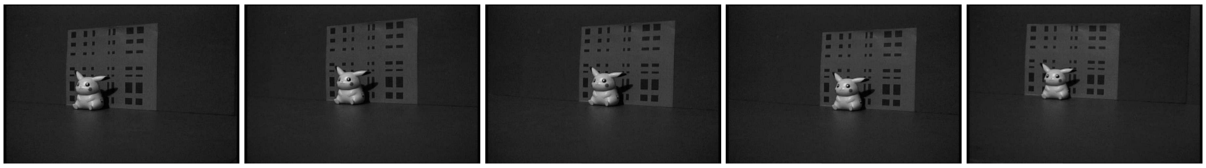
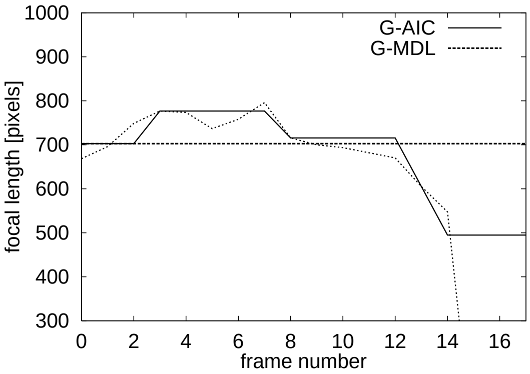
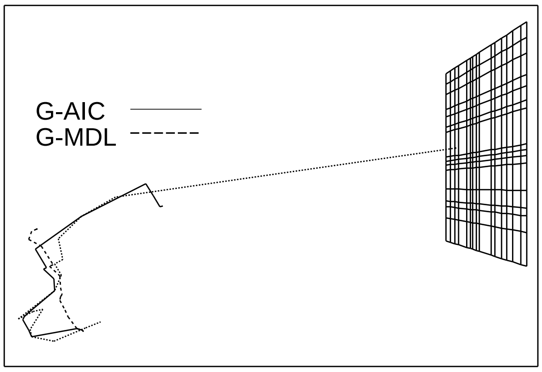


図 6: 実画像列の一部。



(a)



(b)

図 7: (a) 推定した焦点距離。(b) 推定したカメラの経路。(a), (b) において実線は幾何学的 AIC、破線は幾何学的 MDL、点線はモデル選択を用いないもの。

[3] S. Gibbs, C. Arapis, C. Breiteneder, V. Lalioti, S. Mostafawy and J. Speier, Virtual studios: An overview, *IEEE Multimedia*, **5**-1 (1998), 24–35.

[4] 顧海松, 浅田稔, 白井良明, 動き情報に基づくエッジセグメントの最適分割, 電子情報通信学会論文誌 D-II, **J76-D-II**-8 (1993), 1544–1553.

[5] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, U.K., 1993.

[6] 金谷健一, コンピュータビジョンのためのくりこみ法, 情報処理学会論文誌, **35**-2 (1994), 201–209.

[7] 金谷健一, 「空間データの数理 — 3次元コンピューティングに向けて —」, 朝倉書店, 1995.

[8] 金谷健一, 当てはめ問題の最適推定と精度の理論限界, 情報処理学会論文誌, **36**-8 (1995), 1865–1873.

[9] 金谷健一, 幾何学的補正問題の最適計算と精度の理論限界, 情報処理学会論文誌, **37**-3 (1996), 363–370.

[10] 金谷健一, 情報量基準による幾何学的モデル選択, 情報処理学会論文誌, **37**-6 (1996), 1073–1080.

[11] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, The Netherlands, 1996.

[12] 金谷健一, 自己評価を伴うアクティブビジョン, 日本ロボット学会誌, **15**-2 (1997), 268–274.

[13] K. Kanatani, Cramer-Rao lower bounds for curve fitting, *Graphical Models Image Processing*, **60**-2 (1998), 93–99.

[14] K. Kanatani, Geometric information criterion for model selection, *Int. J. Comput. Vision*, **26**-3 (1998), 171–189.

[15] 金谷健一, 統計的推測と幾何学的当てはめにおけるモデル選択, 情報処理学会コンピュータビジョンとイメージメディア研究会, 2000-CVIM-122 (2000-5).

[16] Y. Kanazawa and K. Kanatani, Optimal line fitting and reliability evaluation, *IEICE Trans. Inf. & Syst.*, **E79-D**-9 (1996), 1317–1322.

[17] Y. Kanazawa and K. Kanatani, Optimal conic fitting and reliability evaluation, *IEICE Trans. Inf. & Syst.*, **E79-D**-9 (1996), 1323–1328.

[18] 金澤靖, 松永力, 金谷健一, 複比による識別のための最良マーカーパタンの設計: 理論解析, 情報処理学会コンピュータビジョンとイメージメディア研究会, 99-CVIM-115-13 (1999-3), 97–104.

[19] Y. G. Leclerc, Constructing simple stable descriptions for image partitioning. *Int. J. Comput. Vision* **3**-1 (1989), 73–102.

[20] 松永力, 新島和孝, 金谷健一, 複比による識別のための最良マーカーパタンの設計: 応用実験, 情報処理学会コンピュータビジョンとイメージメディア研究会, 99-CVIM-115-14 (1999-3), 105–110.

[21] 松永力, 金谷健一, 平面パタンを用いる移動カメラのキャリブレーション, 情報処理学会コンピュータビジョンとイメージメディア研究会, 99-CVIM-116-1 (1999-5), 1–8.

[22] 松永力, 金谷健一, 平面パタンを用いる移動カメラの校正：最適計算，信頼性評価，および幾何学的 AIC による安定化, 電子情報通信学会論文誌 A, **J83-A**-6 (2000), to appear.

[23] B. A. Maxwell, Segmentation and interpretation of multicolored objects with highlights, *Comput. Vision Image Understand.*, **77**-1 (2000), 1–24.

[24] 宮川一乗, 黒川雅人, 広田源太郎, 画像とＣＧの融合, 情報処理, **38**-3 (1997), 189–194.

[25] 宮島耕治, 武川直樹, 岡田守, MDL 原理に基づく正則化 — 不連続性に対する正則化パラメータの推定 —, 電子情報通信学会論文誌 D-II, **J80-D-II**-9 (1997), 2369–2378.

[26] 長尾淳平, 韓太舜, かく乱母数を含む場合の MDL 基準の構築と空間図形モデル推定問題への応用, 電子情報通信学会論文誌 A, **J83-A**-1 (2000), 83–95.

[27] 中嶋正之, ディジタル映像処理の基礎と応用 (II)— 二次元ディジタル映像処理入門 —, 電子情報通信学会誌, **18**-1 (1998), 68–75.

[28] J. Rissanen, Universal coding, information, prediction and estimation, *IEEE Trans. Inform. Theory*, **30**-4 (1984), 629–636.

[29] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, Singapore, 1989.

[30] 坂村健, デジタルミュージアム — コンピュータを駆使した新しい博物館の構築 —, 情報処理, **39**-5 (1998), 385–392.

[31] P. H. S. Torr, An assessment of information criteria for motion model selection, *Proc. IEEE Conf. Comput. Vision Patt. Recogn.*, June 1997, Puerto Rico, pp. 47–53.

[32] 山西健司, 確率的コンプレキシティと学習理論, オペレーションズ・リサーチ, **41**-7 (1996), 379–386.

[33] 山西健司, 韓太舜, MDL 入門: 情報理論の立場から, 人工知能学会誌, **7**-3 (1992), 427–434.