宇宙論における統計的手法 (個人用ノート、随時更新)

2018 年 5 月 16 日, 並河 俊弥

概要

このノートでは、宇宙論的観測から得られたデータを利用することで、宇宙論的情報を引き出す手法についてまとめる。

目次

1 準備

数理統計学における推定の基本事項は、宇宙論観測で得られるデータから宇宙論的情報を引き出す手法を説明する上で不可欠となる。宇宙論観測を例に挙げながら、確率や推定法に関する基礎事項を挙げていく。

1.1 確率分布関数

1.1.1 揺らぎの確率分布関数

標準とされている宇宙論では、我々が観測する CMB の温度揺らぎの統計的性質はほぼガウス分布で記述できる。より正確には、CMB の温度揺らぎを球面調和関数展開して得た展開係数

$$a_{\ell m} = \int \mathrm{d}2\hat{\boldsymbol{n}}\, \Theta(\hat{\boldsymbol{n}}) Y^*_{\ell m} \tag{1}$$

は、平均ゼロ、分散 C_ℓ のガウス分布に従う：

$$P(a_{\ell m}|C_\ell) = \frac{1}{\sqrt{2\pi C_\ell}} \exp\left(-\frac{|a_{\ell m}|^2}{2C_\ell}\right) \tag{2}$$

ここで C_ℓ は揺らぎ $a_{\ell m}$ の分散を表す。等方的な揺らぎの場合、分散は m に依存しない。確率変数である $a_{\ell m}$ は各 m について互いに独立であるので、各 m すべてをまとめてデータベクトル $\boldsymbol{d}_\ell = \{a_{\ell,-\ell}, \cdots, a_{\ell m}\}$ で表すと、このデータを得る確率は多次元ガウス分布

$$P(\boldsymbol{d}_\ell|C_\ell) = \prod_{m=-\ell}^{\ell} P(a_{\ell m}|C_\ell) = \frac{1}{(2\pi C_\ell)^{(2\ell+1)/2}} \exp\left(-\sum_{m=-\ell}^{\ell} \frac{|a_{\ell m}|^2}{2C_\ell}\right) \tag{3}$$

となる。

1.1.2 揺らぎの和の確率分布関数

一般に確率変数 x, y の和 $z = x + y$ の分布関数は

$$P(z) = \int \mathrm{d}x \int \mathrm{d}y\, \delta(x+y-z) P(x)P(y) = \int \mathrm{d}x\, P(x)P(z-x) \tag{4}$$

で与えられる。x, y がそれぞれ平均 μ_x, μ_y、分散 σ_x, σ_y のガウス分布に従うとすれば

$$\begin{aligned}
P(z) &= \int_{-\infty}^{\infty} \mathrm{d}x\, \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(z-x-\mu_y)^2}{2\sigma_y^2}\right) \\
&= \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} \mathrm{d}x\, \exp\left[\frac{-(\sigma_x^2+\sigma_y^2)x^2 + 2(\mu_x\sigma_2^y + (\mu_y - z)\sigma_x^2)x - (\mu_x^2\sigma_y^2 + (z-\mu_y)^2\sigma_x^2)}{2\sigma_x^2\sigma_y^2}\right] \\
&= \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} \mathrm{d}x\, \exp\left[-\frac{(\sigma_x^2+\sigma_y^2)(x - \frac{\mu_x\sigma_y^2 + (\mu_y - z)\sigma_x^2}{\sigma_x^2+\sigma_y^2})^2 - (z-\mu_x-\mu_y)^2}{2(\sigma_x^2+\sigma_y^2)}\right] \\
&= \frac{1}{2\pi\sigma_x\sigma_y} \sqrt{2\sigma^2\pi}\, \exp\left[-\frac{(z-\mu_x-\mu_y)^2}{2(\sigma_x^2+\sigma_y^2)}\right] \\
&= \frac{1}{\sqrt{2\pi(\sigma_x^2+\sigma_y^2)}} \exp\left[-\frac{(z-\mu_x-\mu_y)^2}{2(\sigma_x^2+\sigma_y^2)}\right]
\end{aligned} \tag{5}$$

これより、和の分布は、平均 $\mu_x + \mu_y$、分散 $\sigma_x^2 + \sigma_y^2$ のガウス分布となる。

ここで、球面上のマップとして与えられた CMB の揺らぎの統計的性質についてみてみる。球面上の揺らぎ $X(\hat{\boldsymbol{n}}) = \sum_{\ell m} X_{\ell m} Y_{\ell m}(\hat{\boldsymbol{n}})$ は、平均ゼロ、分散 C_ℓ のガウス分布に従う確率変数 $X_{\ell m}$ の和で与えられる。したがって、$X(\hat{\boldsymbol{n}})$ は平均ゼロであり、分散

$$\sigma_0^2 \equiv \langle X^2(\hat{\boldsymbol{n}})\rangle = \sum_{\ell m \ell' m'} \langle X_{\ell m} X_{\ell' m'}\rangle Y_{\ell m}(\hat{\boldsymbol{n}}) Y_{\ell' m'}(\hat{\boldsymbol{n}}) = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell \tag{6}$$

のガウス分布に従う。注意として、異なる二点間の相関も存在するため、共分散行列は対角にならない。

1.1.3　揺らぎの分散の確率分布関数

観測から宇宙論的情報を引き出すさい、揺らぎの分散である角度パワースペクトル C_ℓ を測定し、理論で計算したものと比較することがよく行われる。理論における角度パワースペクトルは、その定義（$a_{\ell m}$ の分散）から

$$C_\ell \delta_{mm'} = \langle a_{\ell m} a^*_{\ell' m'} \rangle \tag{7}$$

で関係付けられる。ただし $\langle \cdots \rangle$ は、$P(\boldsymbol{d}_\ell | C_\ell)$ の統計平均

$$\langle X(a_{\ell m}) \rangle = \int \mathrm{d}(2\ell+1) a_{\ell m}\, X(a_{\ell m}) P(\boldsymbol{d}_\ell | C_\ell) \tag{8}$$

を表す。ここで、C_ℓ を観測から測定するさい、理論側の定義から

$$\widehat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \tag{9}$$

が推定量の候補である。

ここで、$\widehat{C}_\ell$ が従う確率分布について見てみる。揺らぎの確率分布関数は

$$P(\boldsymbol{d}_\ell | C_\ell) \mathrm{d}\boldsymbol{d}_\ell = \frac{1}{(2\pi C_\ell)^{(2\ell+1)/2}} \exp\left[-(2\ell+1) \frac{\widehat{C}_\ell}{2C_\ell} \right] \mathrm{d}\boldsymbol{d}_\ell \tag{10}$$

これを変数変換すると

$$\begin{aligned} P(\widehat{C}_\ell | C_\ell) \mathrm{d}\widehat{C}_\ell &= \frac{1}{(2\pi C_\ell)^{(2\ell+1)/2}} \exp\left[-(2\ell+1) \frac{\widehat{C}_\ell}{2C_\ell} \right] \frac{\mathrm{d}\boldsymbol{d}_\ell}{\mathrm{d}\widehat{C}_\ell} \mathrm{d}\widehat{C}_\ell \\ &= A_\ell \frac{(\widehat{C}_\ell)^{\frac{2\ell-1}{2}}}{(C_\ell)^{\frac{2\ell+1}{2}}} \exp\left[-(2\ell+1) \frac{\widehat{C}_\ell}{2C_\ell} \right] \mathrm{d}\widehat{C}_\ell \end{aligned} \tag{11}$$

ただし規格化因子 A_ℓ は

$$1 = \int_0^\infty \mathrm{d}\widehat{C}_\ell\, P(\widehat{C}_\ell | C_\ell) = A_\ell \int_0^\infty \mathrm{d}x\, x^{\frac{2\ell-1}{2}}\, \mathrm{e}^{-\frac{2\ell+1}{2}x} = A_\ell \left(\frac{2}{2\ell+1} \right)^{\frac{2\ell+1}{2}} \Gamma\left[\frac{2\ell+1}{2} \right] \tag{12}$$

を満たす。一般に多重ガウス型確率分布関数に確率変数に対し、その分散が従う確率分布関数は Wishart 分布と呼ばれる。平均は

$$\begin{aligned} \langle \widehat{C}_\ell \rangle &= \int_0^\infty \mathrm{d}\widehat{C}_\ell\, \widehat{C}_\ell P(\widehat{C}_\ell | C_\ell) \\ &= C_\ell A_\ell \int_0^\infty \mathrm{d}x\, x^{\frac{2\ell+1}{2}}\, \mathrm{e}^{-\frac{2\ell+1}{2}x} \\ &= C_\ell A_\ell \int_0^\infty \mathrm{d}x\, x^{\frac{2\ell-1}{2}}\, \mathrm{e}^{-\frac{2\ell+1}{2}x} = C_\ell \end{aligned} \tag{13}$$

となる。二乗期待値は

$$\begin{aligned} \langle \widehat{C}_\ell^2 \rangle &= \int_0^\infty \mathrm{d}\widehat{C}_\ell\, \widehat{C}_\ell^2 P(\widehat{C}_\ell | C_\ell) \\ &= C_\ell^2 A_\ell \int_0^\infty \mathrm{d}x\, x^{\frac{2\ell+3}{2}}\, \mathrm{e}^{-\frac{2\ell+1}{2}x} \\ &= C_\ell^2 A_\ell \frac{2\ell+3}{2\ell+1} \int_0^\infty \mathrm{d}x\, x^{\frac{2\ell+1}{2}}\, \mathrm{e}^{-\frac{2\ell+1}{2}x} = C_\ell^2 \frac{2\ell+3}{2\ell+1} \end{aligned} \tag{14}$$

1.1.4 Cumulant

確率分布がガウス分布の場合は、平均と分散で確率分布が決定される。しかし上記の分散のように、確率分布がガウス分布と異なる場合には、さらにその確率分布を特徴づける量を導入する必要がある。

確率変数 X のに対し、$\phi(t) \equiv \langle \mathrm{e}^{\mathrm{i}tX} \rangle$ として定義される関数を特性関数 (characteristic function) という。特性関数の対数 $\ln\phi$ を t で展開を考える：

$$\ln\phi(t) = \sum_{n=1}^{\infty} \frac{c_n}{n!} (\mathrm{i}t)^n \tag{15}$$

ここに現れる c_n を n 次の cumulant とよぶ。ガウス分布の場合には $n \geq 3$ の cumulant は存在しないため、高次の cumulantc_n $(n \geq 3)$ はガウス分布からのずれを特徴づける。X の n 乗期待値を n 次の moment と呼ぶが、n 次の cumulant は n 次までの moment で表される。これを見るために、特性関数と n 次の momentm_n の関係が以下で与えられることに着目する：

$$m_n = \frac{1}{\mathrm{i}^n} \frac{\partial^n \phi(t)}{\partial t^n}\bigg|_{t=0} \tag{16}$$

これから $\phi(t)$ の t に関する Taylor 展開は

$$\phi(t) = 1 + \sum_{n=1}^{\infty} \frac{\mathrm{i}^n m_n}{n!} t^n \tag{17}$$

両辺の対数をとることで

$$\ln\phi(t) = \ln\left(1 + \sum_{n=1}^{\infty} \frac{\mathrm{i}^n m_n}{n!} t^n\right) = \sum_{n=1}^{\infty} \frac{\mathrm{i}^n m_n}{n!} t^n - \frac{1}{2}\left(\sum_{n=1}^{\infty} \frac{\mathrm{i}^n m_n}{n!} t^n\right)^2 + \ldots \tag{18}$$

これと、$\ln\phi(t)$ の t に関する展開を比較することで、

$$c_1 = m_1 \tag{19}$$

$$c_2 = m_2^2 - m_1^2 \tag{20}$$

$$c_3 = m_3 - 3m_2 m_1 + 2m_1^3 \tag{21}$$

$$c_4 = m_4 - 4m_3 m_1 - 3m_2^2 + 12 m_2 m_1^2 - 6m_1^4 \tag{22}$$

などが得られる。

例として、種々の確率分布での特性関数を求めておく。Wishart 分布の特性関数は

$$\begin{aligned}
\langle \mathrm{e}^{\mathrm{i}t\widehat{C}_\ell} \rangle &= \int_0^\infty \mathrm{d}\widehat{C}_\ell \ \mathrm{e}^{\mathrm{i}t\widehat{C}_\ell} P(\widehat{C}_\ell | C_\ell) \\
&= A_\ell \int_0^\infty \mathrm{d}x \, x^{\frac{2\ell-1}{2}} \ \mathrm{e}^{-(\frac{2\ell+1}{2} - \mathrm{i}tC_\ell)x} \\
&= A_\ell \left(\frac{2}{2\ell + 1 - 2\mathrm{i}tC_\ell}\right)^{\frac{2\ell+1}{2}} \Gamma\left[\frac{2\ell+1}{2}\right] \\
&= \left(\frac{2\ell+1}{2\ell + 1 - 2\mathrm{i}tC_\ell}\right)^{\frac{2\ell+1}{2}} \\
&= \left(\frac{1}{1 - \mathrm{i}tC_\ell/\beta_\ell}\right)^{\beta_\ell}
\end{aligned} \tag{23}$$

で与えられる。ただし $\beta_\ell = (2\ell+1)/2$ である。

1.1.5 漸近展開

ガウス分布から少しずれた確率分布を表現するさいに、確率分布関数をガウス分布周りで展開しておくと便利である。ここでは、その漸近展開方法である Edgeworth 展開について述べる。

いま n 個の独立変数 X_n があり、それらは平均 μ と分散 σ のガウス ian とは限らない確率分布に従うとする。このとき、標準化された確率変数

$$S_n \equiv \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma} \tag{24}$$

が従う確率分布を求めたいとする。中心極限定理より、$n\to\infty$ でガウス分布に従う確率変数となるので、有限 n であってもある程度大きい n であれば、S_n の確率分布はガウス分布に近い形となる。

標準化された確率変数 $Y_i=(X_i-\mu)/\sigma$ を用いると、S_n の特性関数 $\phi_n(t)$ と Y_i の特性関数 $\phi(t)$ は

$$\phi_n(t)=\langle \mathrm{e}^{\mathrm{i}tS_n}\rangle=\prod_{i=1}^{n}\langle \mathrm{e}^{\mathrm{i}tY_i/\sqrt{n}}\rangle=[\phi(t/\sqrt{n})]^n \tag{25}$$

となる。ここで $\phi(t)$ は

$$\ln\phi(t)=-\frac{1}{2}t^2+\sum_{i=3}^{\infty}\frac{c_i}{i!}(\mathrm{i}t)^i \tag{26}$$

と展開できるので、

$$\ln\phi_n(t)=-\frac{1}{2}t^2+\sum_{i=3}^{\infty}\frac{c_i}{n^{i/2-1}i!}(\mathrm{i}t)^i \tag{27}$$

となる。これから、

$$\phi_n(t)=\exp\left[-\frac{t^2}{2}+\sum_{i=3}^{\infty}\frac{c_i}{n^{i/2-1}i!}(\mathrm{i}t)^i\right]=\exp\left[\sum_{i=3}^{\infty}\frac{c_i}{n^{i/2-1}i!}(\mathrm{i}t)^i\right]e^{-t^2/2} \tag{28}$$

これを逆フーリエ変換することで、求めたい確率分布関数が得られる：

$$\begin{aligned}f_n(x)&=\exp\left[\sum_{i=3}^{\infty}\frac{c_i}{n^{i/2-1}i!}\left(\frac{\partial}{\partial x}\right)^i\right]\mathrm{e}^{-x^2/2}\\&=\left[1-\frac{1}{\sqrt{n}}\frac{c_3}{3!}\left(\frac{\partial}{\partial x}\right)^3+\frac{1}{n}\left(\frac{c_4}{4!}\left(\frac{\partial}{\partial x}\right)^4+\frac{1}{2}\frac{c_3^2}{3!}\left(\frac{\partial}{\partial x}\right)^6\right)+\dots\right]\mathrm{e}^{-x^2/2}\end{aligned} \tag{29}$$

上記の展開を Edgeworth 展開と呼ぶ[*1]。$n\to\infty$ でガウス分布に近づくため、$\sqrt{n}$ に関してまとめられた項ごとに収束の速さが異なる。Hermite 多項式 $h_n(x)$ を

$$\mathrm{e}^{-x^2/2}h_n(x)=(-1)^n\left(\frac{\partial}{\partial x}\right)^n\mathrm{e}^{-x^2/2} \tag{30}$$

で定義すると、

$$f_n(x)=\left[1-\frac{1}{\sqrt{n}}\frac{c_3}{3!}h_3(x)+\frac{1}{n}\left(\frac{c_4}{4!}h_4(x)+\frac{1}{2}\frac{c_3^2}{3!}h_6(x)\right)+\dots\right]\mathrm{e}^{-x^2/2} \tag{31}$$

と書き直せる。

1.1.6　Minkowski 汎関数

平均ゼロ、分散 C_ℓ のガウス分布に従う揺らぎ $X_{\ell m}$ が与えられたとする。球面上の揺らぎを X、その分散を σ_0 としたとき、$u=X/\sigma_0$ は規格化されたガウス分布に従う。

一般に、d 次元 Euclid 空間の集合 Q に対し、Minkowski 汎関数は

$$V_0^{(d)}\equiv\int_Q \mathrm{d}v\ , \tag{32}$$

$$V_k^{(d)}\equiv\frac{1}{\omega_{k-1}\binom{d}{j}}\int_{\partial Q}\mathrm{d}s\,S_k(R_1,\dots,R_{d-1}) \tag{33}$$

[*1] $n=1$ としたとき、同様の展開として Gram-Charlier A 展開がある。しかしこの展開方法では微分の次数にしたがって項がまとめられていたため、漸近展開にはならない場合、すなわち級数が発散する状況が多い。

で与えられる。ここで、$\omega_k = 2\pi^{(k+1)/2}/\Gamma[(k+1)/2]$、$S_k$ は k 次の対称式、R_i は ∂Q 上の主曲率である。ここでは、特に $d=2$ 次元の場合を考える：

$$V_0 = \int_Q \mathrm{d}S \ , \tag{34}$$

$$V_1 = \frac{1}{4}\int_{\partial Q} \mathrm{d}\ell \ , \tag{35}$$

$$V_2 = \frac{1}{2\pi}\int_{\partial Q} \mathrm{d}\ell \, \frac{1}{R} \, . \tag{36}$$

集合 $Q_\nu = \{u \,|\, u \geq \nu\}$ に対する Minkowski 汎関数は

$$V_k(\nu) = \int \mathrm{d}\hat{\boldsymbol{n}} \, F_k(\nu; u) \tag{37}$$

で与えられる。ここで、積分範囲は全天であり、また

$$F_k(\nu; u) = \begin{cases} \theta(u-\nu) & (k=0) \\ \frac{1}{4}\delta(u-\nu)\sqrt{u_{;\theta}^2 + u_{;\varpi}^2} & (k=1) \\ \frac{1}{2\pi}\delta(u-\nu)\frac{2u_{;\theta}u_{;\varphi}u_{;\theta\varphi} - u_{;\theta}^2 u_{;\varphi\varphi} - u_{;\varphi}^2 u_{;\theta\theta}}{u_{;\theta}^2 + u_{;\varphi}^2} & (k=2) \end{cases} \tag{38}$$

Minkowski 汎関数の期待値を求めるには、$\langle F_k(\nu; u)\rangle$ を求める必要がある。まず $k=0$ の場合、求めるべきは

$$\begin{aligned} \langle F_0(\nu; u)\rangle &= \int_{-\infty}^{\infty} \mathrm{d}u \, \theta(u-\nu) \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-u^2/2} \\ &= \frac{1}{\sqrt{2\pi}} \int_{\nu}^{\infty} \mathrm{d}u \, \mathrm{e}^{-u^2/2} = \frac{1}{\sqrt{2\pi}} e^{-\nu^2/2} h_{-1}(\nu) \end{aligned} \tag{39}$$

が得られる。

次に $k=1$ について考える。確率変数 u と $u_{;i}$ に対し [1]

$$\langle u_{;i}\rangle = \langle u\rangle_{;i} = 0 \tag{40}$$

$$\langle u u_{;i}\rangle = \frac{1}{2}\langle u^2\rangle_{;i} = 0 \tag{41}$$

$$\langle u_{;i} u_{;j}\rangle = \langle u u_{;i}\rangle_{;j} - \langle u u_{;ij}\rangle = -\langle u u_{;ij}\rangle = \delta_{ij} \sum_{\ell} \ell(\ell+1) \frac{C_\ell}{\sigma_0^2} \frac{2\ell+1}{8\pi} \equiv \frac{\sigma_1^2}{2\sigma_0^2} \tag{42}$$

ただし

$$(Y_{\ell m})_{;\theta} = \frac{\sqrt{\ell(\ell+1)}}{2}(Y_{\ell m}^{1} + Y_{\ell m}^{-1}) \tag{43}$$

$$(Y_{\ell m})_{;\varphi} = \frac{\sqrt{\ell(\ell+1)}}{2}(Y_{\ell m}^{1} - Y_{\ell m}^{-1}) \tag{44}$$

$$\sum_m (-1)^{m+s_1} Y_{\ell,-m}^{s_1} Y_{\ell m}^{s_2} = \sqrt{\frac{2\ell+1}{4\pi}} Y_{\ell s_1}^{s_2}(0,0) \tag{45}$$

$$Y_{\ell 1}^{-1}(0,0) = Y_{\ell 1}^{1}(0,0) = -\sqrt{\frac{2\ell+1}{4\pi}} \tag{46}$$

を用いた。これより u と $u_{;i}$ はそれぞれ分散 σ_0、σ_1 の独立なガウス場である。したがって、

$$\langle F_1(\nu; u)\rangle = \frac{1}{4}\langle \delta(u-\nu)\sqrt{u_{;\theta}^2 + u_{;\varpi}^2}\rangle = \frac{1}{4}\langle \delta(u-\nu)\rangle \langle \sqrt{u_{;\theta}^2 + u_{;\varpi}^2}\rangle \tag{47}$$

と分解できる。ここで

$$\langle \delta(u-\nu)\rangle = \int_{-\infty}^{\infty} \mathrm{d}u \, \delta(u-\nu) \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-u^2/2} = \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\nu^2/2} \tag{48}$$

である。また、$u_{;\theta}$ と $u_{;\varphi}$ は同じ分散 $\sigma \equiv \sigma_1/\sigma_0/\sqrt{2}$ に従うガウス場なので

$$\begin{aligned}\langle\sqrt{u_{;\theta}^2+u_{;\varpi}^2}\rangle &= \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\,\sqrt{x^2+y^2}\,\frac{1}{2\pi\sigma^2}\,\mathrm{e}^{-(x^2+y^2)/2\sigma^2} \\ &= \frac{1}{\sigma^2}\int_0^{\infty}\mathrm{d}x\,r^2\,\mathrm{e}^{-r^2/2\sigma^2} = \sqrt{\pi/2}\sigma \end{aligned} \tag{49}$$

最終的に

$$\langle F_1(\nu;u)\rangle = \frac{1}{8}\frac{\sigma_1}{\sqrt{2}\sigma_0}\,\mathrm{e}^{-\nu^2/2} \tag{50}$$

が得られる。

最後に $k=2$ の場合について計算する。u と $u_{;i}$、$u_{;i}$ と $u_{;ij}$ は独立、u と $u_{;ii}$ は独立ではないことに注意すると、

$$\langle F_2(\nu;u)\rangle = -\frac{1}{2\pi}\left[\langle\delta(u-\nu)u_{;\varphi\varphi}\rangle\left\langle\frac{u_{;\theta}^2}{u_{;\theta}^2+u_{;\varphi}^2}\right\rangle + \langle\delta(u-\nu)u_{;\theta\theta}\rangle\left\langle\frac{u_{;\varphi}^2}{u_{;\theta}^2+u_{;\varphi}^2}\right\rangle\right] \tag{51}$$

ここで

$$\langle\delta(u-\nu)u_{;ii}\rangle = -\langle\delta(u-\nu)u\rangle\langle u_{;i}^2\rangle = -\frac{\nu\,\mathrm{e}^{-\nu^2/2}}{\sqrt{2\pi}}\frac{\sigma_1^2}{2\sigma_0^2} \tag{52}$$

また

$$\left\langle\frac{u_{;j}^2}{u_{;i}u^{;i}}\right\rangle = \int_0^{\infty}\mathrm{d}r\int_0^{2\pi}\mathrm{d}\phi\,r\,\mathrm{e}^{-r^2/2\sigma^2}\cos^2\phi\frac{1}{2\pi\sigma^2} = \frac{1}{2} \tag{53}$$

したがって

$$\langle F_2(\nu;u)\rangle = \frac{1}{(2\pi)^{3/2}}\frac{\sigma_1^2}{2\sigma_0^2}\nu\,\mathrm{e}^{-\nu^2/2} \tag{54}$$

$x = X_{\ell,m}$ として、確率変数の絶対値に対し

$$\begin{aligned}\langle|x|\rangle = \langle x(2\theta(x)-1)\rangle = 2\langle x\theta(x)\rangle &= \int_0^{\infty}\mathrm{d}x\,\frac{x}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-x^2/(2\sigma^2)} \\ = \frac{-\sigma}{\sqrt{2\pi}}[\mathrm{e}^{-x^2/(2\sigma^2)}]_0^{\infty} = \frac{\sigma}{\sqrt{2\pi}} &\end{aligned} \tag{55}$$

1.2 統計的一様性および等方性について

ある揺らぎ $X(\hat{\boldsymbol{n}})$ のフーリエ変換を考え、その二点相関

$$\langle X_{\boldsymbol{\ell}}X_{\boldsymbol{\ell}'}\rangle = \int\mathrm{d}\hat{\boldsymbol{n}}\int\mathrm{d}\hat{\boldsymbol{n}}'\,\xi(\hat{\boldsymbol{n}},\hat{\boldsymbol{n}}')\,\mathrm{e}^{-\mathrm{i}\hat{\boldsymbol{n}}\cdot\boldsymbol{\ell}}\,\mathrm{e}^{-\mathrm{i}\hat{\boldsymbol{n}}'\cdot\boldsymbol{\ell}'} \tag{56}$$

の性質を考える。

$\xi(\hat{\boldsymbol{n}},\hat{\boldsymbol{n}}')$ が二点の差 $\Delta\hat{\boldsymbol{n}} \equiv \hat{\boldsymbol{n}}' - \hat{\boldsymbol{n}}$ にのみ依存することを統計的に一様であるとすると、上述のフーリエ空間の統計量は

$$\begin{aligned}\langle X_{\boldsymbol{\ell}}X_{\boldsymbol{\ell}'}\rangle &= \int\mathrm{d}\hat{\boldsymbol{n}}\int\mathrm{d}\Delta\hat{\boldsymbol{n}}\,\xi(\Delta\hat{\boldsymbol{n}})\,\mathrm{e}^{-\mathrm{i}\hat{\boldsymbol{n}}\cdot(\boldsymbol{\ell}+\boldsymbol{\ell}')}\,\mathrm{e}^{-\mathrm{i}\Delta\hat{\boldsymbol{n}}\cdot\boldsymbol{\ell}'} \\ &= \delta(\boldsymbol{\ell}+\boldsymbol{\ell}')\int\mathrm{d}\hat{\boldsymbol{n}}\,\xi(\Delta\hat{\boldsymbol{n}})\,\mathrm{e}^{-\mathrm{i}\Delta\hat{\boldsymbol{n}}\cdot\boldsymbol{\ell}'} \\ &\equiv \delta(\boldsymbol{\ell}+\boldsymbol{\ell}')P(\boldsymbol{\ell}')\end{aligned} \tag{57}$$

となる。さらに $\xi(\Delta\hat{\boldsymbol{n}})$ が $\Delta\hat{\boldsymbol{n}}$ の方向にも依存しなければ（統計的に等方であれば）

$$\langle X_{\boldsymbol{\ell}}X_{\boldsymbol{\ell}'}\rangle = \delta(\boldsymbol{\ell}+\boldsymbol{\ell}')P(|\boldsymbol{\ell}'|) \tag{58}$$

となる。等方であれば一様である。

さらに一般に非一様非等方の場合をみてみる。例えば実際の観測で起こりうる非一様なポアソンノイズ

$$\xi(\hat{\boldsymbol{n}}, \hat{\boldsymbol{n}}') = S(\hat{\boldsymbol{n}})\delta(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}') \tag{59}$$

を考える。これは非一様かつ非等方である。このとき、

$$\langle X_{\boldsymbol{\ell}} X_{\boldsymbol{\ell}'} \rangle = \int \mathrm{d}\hat{\boldsymbol{n}}\, \xi(\hat{\boldsymbol{n}}) \mathrm{e}^{-\mathrm{i}\hat{\boldsymbol{n}}\cdot(\boldsymbol{\ell}+\boldsymbol{\ell}')} \tag{60}$$

であり、$\boldsymbol{\ell} \neq \boldsymbol{\ell}'$ の値も存在する。

1.3 推定

CMB 観測から、各天球面上での温度揺らぎ $\Theta(\hat{\boldsymbol{n}})$ が与えられたとする。宇宙論では、この温度揺らぎから構成される角度パワースペクトルなどの統計量を測り、さらに宇宙論的に興味のあるパラメータを推定する。

温度揺らぎ Θ はある確率分布に従い生成される。これに伴い、Θ から構成される統計量の推定量 (estimator)$\widehat{X}$ も、ある確率分布に従って値がばらつくことになる。宇宙をいくつも観測した場合、Θ は真の値に近づく。このとき、$\widehat{X}$ も真の値 X に近づくべきである。これは

- 不偏性 (unbiased)：$\langle \widehat{X} \rangle_\Theta = X$

を意味する。ただし、$\langle\cdots\rangle_\Theta$ は Θ に対する統計平均である。現実の観測では、一度の観測でできるだけ真に近い値を推定したい。したがって、

- 最小分散 (optimal)：$\langle |\widehat{X} - X|^2 \rangle_\Theta$ ができるだけ小さい

ものが、推定量として好まれる。どのような不偏推定量を用いても、ある分散以下に分散を抑えることはできないことが知られている（Cramer-Rao の不等式)。

これらの条件を満たす推定量が、特定の確率分布関数の形状を仮定したときにどのように導かれるか、以下で見ていくことにする。

1.3.1 事前・事後確率と尤度関数

例えば、温度揺らぎ $\widehat{\boldsymbol{\Theta}}$ を観測し、それから角度パワースペクトル $\boldsymbol{C}^{\Theta\Theta}$ を得る場合、その確率は

$$P(\boldsymbol{C}^{\Theta\Theta}|\widehat{\boldsymbol{\Theta}})P(\widehat{\boldsymbol{\Theta}})\,, \tag{61}$$

で表される。一方、これは

$$P(\widehat{\boldsymbol{\Theta}}|\boldsymbol{C}^{\Theta\Theta})P(\boldsymbol{C}^{\Theta\Theta})\,, \tag{62}$$

とも表せる。これから、Bayes の定理と呼ばれる等式が導かれる：

$$P(\boldsymbol{C}^{\Theta\Theta}|\widehat{\boldsymbol{\Theta}}) = \frac{P(\widehat{\boldsymbol{\Theta}}|\boldsymbol{C}^{\Theta\Theta})P(\boldsymbol{C}^{\Theta\Theta})}{P(\widehat{\boldsymbol{\Theta}})}\,. \tag{63}$$

左辺は事後確率 (Posterior probability)、左辺の $P(\boldsymbol{C}^{\Theta\Theta})$ と $P(\widehat{\boldsymbol{\Theta}})$ は事前確率分布 (Prior probability) と呼ばれる。また、$P(\widehat{\boldsymbol{\Theta}}|\boldsymbol{C}^{\Theta\Theta})$ は尤度関数 (Likelihood function) という。

1.3.2 Bayes 推定と最尤推定

Bayes 推定では、事後確率分布 (この場合は角度パワースペクトルの分布) を求める。このさい、平均値などをもとにして点推定が行われる。もう一つの推定量構築の指針は最尤原理である。これは、我々が推定すべき量は、尤度関数を最大とするものである。最尤推定では、尤度関数を最大にする角度パワースペクトルを推定する。従って事前確率が一様である場合には、両者の推定法は同じ結果を与える。また Bayes 推定では、事前確率分布を何かしらの方法で与える必要がある。

1.3.3 ガウス型尤度関数

n 個のデータ s_i が、平均 μ、分散 σ^2 のガウス型尤度関数にしたがって得られる場合を考える：

$$P(\boldsymbol{s}|\mu,\sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\sum_{i=1}^{n} \frac{(s_i-\mu)^2}{2\sigma^2}\right\} \tag{64}$$

データから平均 μ と分散 σ^2 を推定したいとする。最尤原理に従うと、最尤推定量は

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} s_i\,, \tag{65}$$

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(s_i-\widehat{\mu})^2 \tag{66}$$

が得られる。

注意として、$\widehat{\mu}$ は不偏推定量であるが、$\widehat{\sigma^2}$ は不偏推定量ではなく

$$\langle\widehat{\sigma^2}\rangle = \frac{n-1}{n}\sigma^2 \tag{67}$$

となる。もし平均がゼロだと分かっている場合は、尤度関数において $\mu=0$ とすることで、分散の最尤推定は

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} s_i^2 \tag{68}$$

となる。これは不偏推定量である。

2 Fisher 解析

ここでは、Fisher 情報行列を利用した、パラメータの決定精度を理論的に見積もる統計的解析手法について述べる。

2.1 準備

観測から得られる n 個のデータ $d_i(i=1,2,\ldots,n)$ に対して、データベクトル $\boldsymbol{d}$ を

$$\boldsymbol{d} \equiv \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \tag{69}$$

とする。また、理論の m 個のパラメータ $p_i(i=1,2,\ldots,m)$ に対して、ベクトル

$$\boldsymbol{p} \equiv \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix} \tag{70}$$

を定義する。

理論のパラメータ $\boldsymbol{p}$ が与えられたときに観測データ $\boldsymbol{d}$ が得られる確率をを尤度関数という。これを $\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})$ と表すことにする。尤度関数は $\boldsymbol{d}$ と $\boldsymbol{p}$ の二変数関数である。ただし

$$\int \mathrm{d}\boldsymbol{d}\,\mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) = 1 \qquad (\forall\boldsymbol{p}) \tag{71}$$

が成り立つように規格化する。以下では、簡単のために尤度関数による平均操作を

$$\langle X\rangle \equiv \int \mathrm{d}\boldsymbol{d}\,X(\boldsymbol{d})\mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) \tag{72}$$

と表す。$\langle X\rangle$ は $\boldsymbol{p}$ の関数である。

2.1.1　共分散行列

観測データから計算される角度パワースペクトル、あるいは一般にデータの関数である統計量の誤差について考える。データ $\boldsymbol{d}$ から計算される i 次元の量 $\boldsymbol{X}(\boldsymbol{d}) = (X_1(\boldsymbol{d}), X_2(\boldsymbol{d}), \ldots, X_i(\boldsymbol{d}))$ に対して

$$\mathbf{C}(\boldsymbol{X}) \equiv \langle {}^t[\boldsymbol{X} - \langle \boldsymbol{X} \rangle][\boldsymbol{X} - \langle \boldsymbol{X} \rangle] \rangle \tag{73}$$

で定義される行列 $\mathbf{C}$ を共分散行列（covariance matrix）とよぶ。これは、一次元における分散を i 次元に拡張したものである。

2.1.2　不偏推定量

あるデータから理論のパラメータ p_i を推定する状況を考える。データから求められた推定量を $\hat{\boldsymbol{p}}$ としたとき、$\hat{\boldsymbol{p}}$ の期待値が理論のパラメータ $\boldsymbol{p}$ に一致する場合、$\hat{\boldsymbol{p}}$ は不偏推定量とよばれる。これを式で表すと

$$\langle \hat{\boldsymbol{p}} \rangle \equiv \int \mathrm{d}\boldsymbol{d}\, \hat{\boldsymbol{p}}(\boldsymbol{d}) \mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) = \boldsymbol{p} \tag{74}$$

となる。

2.2　パラメータ決定精度

宇宙論ではモンテカルロ・マルコフ連鎖法によって推定したいパラメータの事後確率分布を求めることがよく行われている。ここでは、Bayes 推定に基づくパラメータ決定精度について考察する。

まず、特定の realization での事後確率分布を得たとし、それから平均 $\boldsymbol{p}_0$ を求めたとする（事後確率分布を最大にするものとしても結果は同じとなる）。平均値まわりで対数事後確率分布を展開しておくと

$$\begin{aligned}\ln P(\boldsymbol{p}|\boldsymbol{d}) = \ln P(\boldsymbol{p}_0|\boldsymbol{d}) &+ \left.\frac{\partial \ln P(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i}\right|_{\boldsymbol{p}=\boldsymbol{p}_0} (\boldsymbol{p} - \boldsymbol{p}_0) \\ &+ \frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}_0)^t \left.\frac{\partial^2 \ln P(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i \partial p_j}\right|_{\boldsymbol{p}=\boldsymbol{p}_0} (\boldsymbol{p} - \boldsymbol{p}_0) + \mathcal{O}(|\boldsymbol{p} - \boldsymbol{p}_0|^3)\end{aligned} \tag{75}$$

ここで $\boldsymbol{p} = \boldsymbol{p}_0$ における対数事後確率分布の Hessian を $\mathbf{H}(\boldsymbol{p}_0)$ とおいた。上式において $|\boldsymbol{p} - \boldsymbol{p}_0|$ の三次以上の寄与を無視する、すなわち事後確率分布がパラメータ $\boldsymbol{p}$ に対する Gauss 分布であるとする。事後確率分布は

$$P(\boldsymbol{p}|\boldsymbol{d}) = P(\boldsymbol{p}_0|\boldsymbol{d}) \exp\left[-\frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}_0)^t \mathbf{H}(\boldsymbol{p}_0, \boldsymbol{d})(\boldsymbol{p} - \boldsymbol{p}_0)\right] \tag{76}$$

ただし、$\boldsymbol{p} - \boldsymbol{p}_0$ に比例する項は

$$\begin{aligned}0 &= \int \mathrm{d}\boldsymbol{p}\, (\boldsymbol{p} - \boldsymbol{p}_0) P(\boldsymbol{p}|\boldsymbol{d}) \\ &= \int \mathrm{d}\boldsymbol{p}\, (\boldsymbol{p} - \boldsymbol{p}_0) P(\boldsymbol{p}_0|\boldsymbol{d}) \exp\left[\alpha(\boldsymbol{p} - \boldsymbol{p}_0) - \frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}_0)^t \mathbf{H}(\boldsymbol{p}_0, \boldsymbol{d})(\boldsymbol{p} - \boldsymbol{p}_0)\right]\end{aligned} \tag{77}$$

から消える。

パラメータ p_1 に対する事後確率分布関数は、p_1 への射影

$$P(p_1|\boldsymbol{d}) = \int \mathrm{d}^{m-1} p_2 p_3 \cdots p_m\, P(\boldsymbol{p}|\boldsymbol{d}) \tag{78}$$

で与えられる。この確率分布から、パラメータ p_1 への決定精度が求められる。

ここで、一般にパラメータ p_i に対する決定精度 σ_i が $\{\{\mathbf{H}^{-1}\}_{ii}\}^{1/2}$ で与えられることを示す。このために、数理統計でよく用いられる特性関数を利用する。これは、確率分布関数の Fourier 変換で与えられる：

$$\Phi(\boldsymbol{k}, \boldsymbol{d}) = \int \mathrm{d}\boldsymbol{p}\, P(\boldsymbol{p}|\boldsymbol{d}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{p} - \boldsymbol{p}_0)\cdot \boldsymbol{k}} \tag{79}$$

特性関数を用いる利点としては、例えば両辺を k_i で二回微分して $\boldsymbol{k}=0$ とおくことで、

$$\left.\frac{\partial^2\Phi(\boldsymbol{k},\boldsymbol{d})}{\partial k_i^2}\right|_{\boldsymbol{k}=\boldsymbol{0}} = \int \mathrm{d}\boldsymbol{p}\,(p_i-p_{i,0})^2 P(\boldsymbol{p}|\boldsymbol{d}) = \sigma_i^2 \tag{80}$$

が得られる。異なる k_i と k_j で微分した場合は二つのパラメータの相関を表す。今の場合、特性関数は

$$\Phi(\boldsymbol{k},\boldsymbol{d}) = P(\boldsymbol{p}_0|\boldsymbol{d}) \int \mathrm{d}\boldsymbol{p}\,\exp\left[-\frac{1}{2}\boldsymbol{p}^t\mathbf{H}(\boldsymbol{p}_0,\boldsymbol{d})\boldsymbol{p} - \mathrm{i}\boldsymbol{p}\cdot\boldsymbol{k}\right] \tag{81}$$

である。右辺の積分は Gauss 積分であることは、あるいは単に Gauss 積分の Fourier 変換であることから

$$\Phi(\boldsymbol{k},\boldsymbol{d}) = P(\boldsymbol{p}_0|\boldsymbol{d})\exp\left[-\frac{1}{2}\boldsymbol{k}^t\mathbf{H}^{-1}(\boldsymbol{p}_0,\boldsymbol{d})\boldsymbol{k}\right] \tag{82}$$

が導かれる。したがって、$\sigma_i^2 = \{\mathbf{H}^{-1}\}_{ii}$ である。あるいは一般に、パラメータ間の共分散行列は $\mathbf{H}^{-1}$ である。

2.3 Fisher 行列

2.3.1 決定精度の予測

パラメータの決定精度は、さらに prior に仮定を置くことで Fisher 行列を用いて表すことができる。Bayse の関係式から、対数事後確率分布のパラメータ微分は対数尤度関数のパラメータ微分に加えて対数 prior のパラメータ微分で書ける：

$$\frac{\partial \ln P(\boldsymbol{p}|\boldsymbol{d})}{\partial p_i} = \frac{\partial \ln \mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i} + \frac{\partial \ln P(\boldsymbol{p})}{\partial p_i}\,. \tag{83}$$

宇宙論ではパラメータの prior として区間一様分布がよく仮定される。$\boldsymbol{p}_0$ がこの区間に入っていれば、$\boldsymbol{p}=\boldsymbol{p}_0$ においては、対数事後確率分布のパラメータ微分と対数尤度関数のパラメータ微分は等しくなる。したがって、$\mathbf{H}$ の対数事後確率分布は対数尤度関数に置き換えられる。$\mathbf{H}$ の期待値、すなわち $\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})$ に従う確率分布で $\boldsymbol{d}$ に対し平均操作を行って得られる行列

$$F(\boldsymbol{p}_0) \equiv \langle H(\boldsymbol{p}_0,\boldsymbol{d})\rangle = -\left\langle \left.\frac{\partial^2 \ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i\partial p_j}\right|_{\boldsymbol{p}=\boldsymbol{p}_0}\right\rangle = -\left.\left\langle \frac{\partial^2 \ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i\partial p_j}\right\rangle\right|_{\boldsymbol{p}=\boldsymbol{p}_0} \tag{84}$$

は Fisher 情報行列（Fisher information matrix)、あるいは単に Fisher 行列と呼ばれる。上記の条件を想定すれば、煩雑な事後確率分布を求めることなく、Fisher 行列を決定精度の予測に使える。

上記の議論と同様のことを繰り返すことで、理論予測におけるパラメータ間の共分散行列が $\mathbf{F}^{-1}$ で与えられることが導かれる。そのさい、期待されるパラメータの確率分布関数は

$$\ln P(\boldsymbol{p}) \equiv \langle \ln P(\boldsymbol{p}|\boldsymbol{d})\rangle = \mathrm{const.} - \frac{1}{2}(\boldsymbol{p}-\boldsymbol{p}_0)^t\mathbf{F}(\boldsymbol{p}_0)(\boldsymbol{p}-\boldsymbol{p}_0) \tag{85}$$

2.3.2 二次元楕円

二次元平面上での 1σ の制限の等高線を書くには、確率分布関数の積分を行うことで、確率が 0.68 となる中心 $\boldsymbol{p}_0$ の二次元面を求めればよい。具体的には、$\boldsymbol{p}_0$ が原点となるように平行移動させ、Fisher 行列を対角化する。このとき、

$$P = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right] \tag{86}$$

の確率分布関数に対する積分を実行すればよい。ただし $\sigma_i^2 = 1/\tilde{F}_{ii}$、また対角化した Fisher 行列を $\tilde{F}$ と表す。さらに変数変換

$$x = \sigma_x t\sin\theta\,,\quad y = \sigma_y t\cos\theta \tag{87}$$

を行うと、

$$\mathrm{d}x = \sigma_x(\sin\theta\mathrm{d}t + t\cos\theta\mathrm{d}\theta \tag{88}$$

$$\mathrm{d}y = \sigma_y(\cos\theta\mathrm{d}t - t\sin\theta\mathrm{d}\theta \tag{89}$$

$$|\mathrm{d}x\wedge\mathrm{d}y| = \sigma_x\sigma_y t\mathrm{d}t\wedge\mathrm{d}\theta \tag{90}$$

を用いて

$$0.68 = \frac{1}{2\pi\sigma_x\sigma_y}\int_0^s\mathrm{d}t\int_0^{2\pi}\mathrm{d}\theta\,\mathrm{e}^{-t^2/2}\sigma_x\sigma_y t = \int_0^s\mathrm{d}t\,t\,\mathrm{e}^{-t^2/2} = 1-\mathrm{e}^{-s^2/2}. \tag{91}$$

これから $s\simeq 1.51$ となる。また二次元での 1σ に相当する楕円は、

$$(\boldsymbol{p}-\boldsymbol{p}_0)^t\mathbf{F}(\boldsymbol{p}_0)(\boldsymbol{p}-\boldsymbol{p}_0) = (1.51)^2 \tag{92}$$

で表現される。

2.3.3 情報の和

二つの独立な観測量を組み合わせて決定精度を求める場合を考える。尤度関数は、各々の尤度関数の積として表される。このため、情報を組み合わせた場合の Fisher 行列は、各 Fisher 行列の和として表される。

二つの Fisher 行列を足して描いた楕円は、もとの楕円の内側に必ず入る。これを確認するため、$\boldsymbol{p}_0 = \mathbf{0}$ とする。$\mathbf{F} = \mathbf{F}^{(1)} + \mathbf{F}^{(2)}$ の描く楕円は

$$\boldsymbol{p}^t(\mathbf{F}^{(1)} + \mathbf{F}^{(2)})\boldsymbol{p} = s^2 \tag{93}$$

である。これと $\mathbf{F}^{(1)}$ が描く楕円の交点は

$$\boldsymbol{p}^t\mathbf{F}^{(2)}\boldsymbol{p} = 0 \tag{94}$$

を満たす必要があるが、左辺はゼロとならない。共通の中心をもち、もとの楕円より面積が小さいことから、$\mathbf{F}$ が描く楕円は $\mathbf{F}^{(1)}$ が描く楕円より内側に存在する。同様に、$\mathbf{F}^{(2)}$ が描く楕円より内側に存在するから、$\mathbf{F}^{(1)}$ と $\mathbf{F}^{(2)}$ の楕円の共通領域より内側になければならない。これは n 次元でも成立する。

2.3.4 Fisher 行列とスコア関数

まず、確率分布関数 $\mathcal{L}$ の規格化条件（71）において両辺を p_i で微分すると

$$0 = \int\mathrm{d}\boldsymbol{d}\,\frac{\partial\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i} = \int\mathrm{d}\boldsymbol{d}\,\frac{\partial\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i}\mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) = \left\langle\frac{\partial\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i}\right\rangle \tag{95}$$

となる。右辺の

$$V_i \equiv \frac{\partial\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i} \tag{96}$$

はスコア関数と呼ばれる。さらに p_j で微分を行うと

$$\left\langle\frac{\partial^2\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i\partial p_j}\right\rangle + \left\langle\frac{\partial\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_i}\frac{\partial\ln\mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_j}\right\rangle = 0 \tag{97}$$

が得られる。上記の式は、Fisher 行列とスコア関数を用いて

$$F_{ij}(\boldsymbol{p}) = \langle V_iV_j\rangle \tag{98}$$

と書ける。スコア関数の平均はゼロなので、Fisher 行列はスコア関数の共分散行列で表される。

2.3.5 Fisher 行列と Cramér-Rao の不等式

不偏推定量を用いて推定される理論パラメータの分散には最小値が存在する（Cramér-Rao の不等式）。この分散を最小分散（minimum variance）という。ここでは、その不偏推定量に対する最小分散を求める。

式（74）の両辺を理論のパラメータ p_j で微分すると

$$\int \mathrm{d}\boldsymbol{d}\,\hat{p}_i \frac{\partial \ln \mathcal{L}(\boldsymbol{d}|\boldsymbol{p})}{\partial p_j} \mathcal{L}(\boldsymbol{d}|\boldsymbol{p}) = \langle \hat{p}_i V_j \rangle = \delta_{ij} \tag{99}$$

が得られる。スコア関数の平均がゼロであるから、さらに $\langle (\hat{p}_i - p_i) V_j \rangle = \delta_{ij}$ が成り立つ。ここで、ベクトル $\boldsymbol{A}^t \equiv (\hat{\boldsymbol{p}}, \boldsymbol{V})$ を考える。ただし $\boldsymbol{V}$ はスコア関数を一まとめにしたベクトルである。$\hat{\boldsymbol{p}}$ の共分散行列を $\mathbf{P} = \langle (\hat{\boldsymbol{p}} - \boldsymbol{p})^\dagger (\hat{\boldsymbol{p}} - \boldsymbol{p}) \rangle$ と表すと、このベクトルの共分散行列 $\boldsymbol{C}$ は、

$$\mathbf{C} = \begin{pmatrix} \mathbf{P} & \mathbf{I} \\ \mathbf{I} & \mathbf{F} \end{pmatrix} \tag{100}$$

と書ける。$\mathbf{F}$ は Fisher 行列、$\mathbf{I}$ を N 次元の単位行列である。共分散行列は半正定値行列であるから、任意のベクトル $\boldsymbol{U}$ に対して

$${}^t\boldsymbol{U}\mathbf{C}\boldsymbol{U} \geq 0 \tag{101}$$

が成り立つ。ここで N 次元ベクトル $\boldsymbol{u}, \boldsymbol{v}$ を用いて

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \tag{102}$$

と分解すると、式（101）は

$${}^t\boldsymbol{u}\mathbf{P}\boldsymbol{u} + {}^t\boldsymbol{u}\boldsymbol{v} + {}^t\boldsymbol{v}\boldsymbol{u} + {}^t\boldsymbol{v}\mathbf{F}\boldsymbol{v} \geq 0 \tag{103}$$

と書き直せる。この式の左辺は

$${}^t\boldsymbol{u}(\mathbf{P} - \mathbf{F}^{-1})\boldsymbol{u} + (\boldsymbol{u} + \mathbf{F}\boldsymbol{v})^t \mathbf{F}^{-1} (\boldsymbol{u} + \mathbf{F}\boldsymbol{v}) \geq 0 \tag{104}$$

任意の $\boldsymbol{u}$,$\boldsymbol{v}$ で成り立つので、

$$\mathbf{P}_{ii} \geq \{\mathbf{F}^{-1}\}_{ii} \tag{105}$$

となる。この不等式は Cramer-Rao の不等式と呼ばれる。

2.4 Fisher 行列の計算例

2.4.1 Gauss 分布

ここで、尤度関数を Gauss 分布と仮定した場合の Fisher 行列を求める。n 次元 Gauss 分布に対する尤度関数 $\mathcal{L}$ は

$$\mathcal{L} \equiv \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left[-\frac{1}{2} (\boldsymbol{d} - \langle \boldsymbol{d} \rangle)^\dagger \mathbf{C}^{-1} (\boldsymbol{d} - \langle \boldsymbol{d} \rangle) \right] \tag{106}$$

で与えられる。ここで、$\boldsymbol{d}$ は n 次元データベクトル、$\mathbf{C}$ は共分散行列である。対数尤度は

$$\ln \mathcal{L} = -\frac{1}{2} \mathrm{Tr}(\ln \mathbf{C} + \mathbf{C}^{-1} \widehat{\mathbf{C}}) \tag{107}$$

と表せる[*2]。ただし $\widehat{\mathbf{C}} = (\boldsymbol{d} - \langle \boldsymbol{d} \rangle)^\dagger (\boldsymbol{d} - \langle \boldsymbol{d} \rangle)$ である。スコア関数は

$$V_i = \frac{\partial \ln \mathcal{L}}{\partial p_i} = -\frac{1}{2} \mathrm{Tr} \left(\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial p_i} - \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial p_i} \mathbf{C}^{-1} \widehat{\mathbf{C}} + \mathbf{C}^{-1} \frac{\partial \widehat{\mathbf{C}}}{\partial p_i} \right) \tag{108}$$

[*2] カイ二乗は対数尤度を用いて $\chi^2 \equiv -2 \ln \mathcal{L}$ と定義される。

であり、さらに微分したものは

$$-2\frac{\partial^2 \ln\mathcal{L}}{\partial p_i \partial p_j} = \mathrm{Tr}\Bigg(\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j} + \mathbf{C}^{-1}\frac{\partial^2\mathbf{C}}{\partial p_i\partial p_j} + \mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\widehat{\mathbf{C}}$$
$$- \mathbf{C}^{-1}\frac{\partial^2\mathbf{C}}{\partial p_i\partial p_j}\mathbf{C}^{-1}\widehat{\mathbf{C}} + \mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j}\mathbf{C}^{-1}\widehat{\boldsymbol{C}} - \mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\frac{\partial\widehat{\mathbf{C}}}{\partial p_j}$$
$$- \mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j}\mathbf{C}^{-1}\frac{\partial\widehat{\mathbf{C}}}{\partial p_i} + \mathbf{C}^{-1}\frac{\partial^2\widehat{\mathbf{C}}}{\partial p_i\partial p_j}\Bigg) \tag{109}$$

となる。Fisher 行列を得るため、式（(109)）の両辺を $\boldsymbol{d}$ に関し、$\mathcal{L}(\boldsymbol{d},\boldsymbol{p})$ で平均をとる。ここで、$\widehat{\mathbf{C}}$ の平均は、その定義から

$$\langle\widehat{\mathbf{C}}\rangle = \mathbf{C} \tag{110}$$

となる。一方、一回微分は

$$\left\langle\frac{\partial\widehat{\mathbf{C}}}{\partial p_i}\right\rangle = -2\left\langle(\mathbf{d}-\langle\mathbf{d}\rangle)\frac{\partial\langle\mathbf{d}^\dagger\rangle}{\partial p_i}\right\rangle = -2\langle\mathbf{d}-\langle\mathbf{d}\rangle\rangle\frac{\partial\langle\mathbf{d}^\dagger\rangle}{\partial p_i} = 0 \tag{111}$$

であり、さらに二回微分は

$$\left\langle\frac{\partial^2\widehat{\mathbf{C}}}{\partial p_i\partial p_j}\right\rangle = 2\left\langle\frac{\partial\langle\mathbf{d}\rangle}{\partial p_i}\frac{\partial\langle\mathbf{d}^\dagger\rangle}{\partial p_j}\right\rangle = 2\frac{\partial\langle\mathbf{d}\rangle}{\partial p_i}\frac{\partial\langle\mathbf{d}^\dagger\rangle}{\partial p_j} \tag{112}$$

これを用いると、式（109）を Fisher 行列の定義式に代入することで、尤度関数が Gauss 分布の場合の Fisher 行列

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{Tr}\left(\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_i}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial p_j}\right) + \frac{\partial\langle\mathbf{d}^\dagger\rangle}{\partial p_j}\mathbf{C}^{-1}\frac{\partial\langle\mathbf{d}\rangle}{\partial p_i} \tag{113}$$

が得られる。

CMB 観測の場合、データベクトルを $a_{\ell m}$ として Gauss 型尤度関数をもとに Fisher 行列を導くと、

$$\mathbf{F}_{ij} = \sum_{\ell=2}\frac{2\ell+1}{2}\mathrm{Tr}\left(\mathbf{C}_\ell^{-1}\frac{\partial\mathbf{C}_\ell}{\partial p_i}\mathbf{C}_\ell^{-1}\frac{\partial\mathbf{C}_\ell}{\partial p_j}\right) \tag{114}$$

である。ここで

$$\mathbf{C}_\ell \equiv \begin{pmatrix} C_\ell^{\Theta\Theta} & C_\ell^{\Theta E} & 0 \\ C_\ell^{\Theta E} & C_\ell^{EE} & 0 \\ 0 & 0 & C_\ell^{BB} \end{pmatrix} \tag{115}$$

であり、p_i は理論のパラメータである。また C_ℓ^{XY} はパワースペクトルである。

2.4.2 Wishart 分布の場合での Fisher 行列

Wishart 分布の場合（例えば角度パワースペクトル）、対数尤度関数は

$$-\ln\mathcal{L}(\widehat{C}_\ell|C_\ell) = \frac{2\ell+1}{2}\left[\frac{\widehat{C}_\ell}{C_\ell} + \ln C_\ell - \frac{2\ell-1}{2\ell+1}\ln\widehat{C}_\ell\right] \tag{116}$$

で与えられる。これからスコア関数は

$$-\frac{\partial\ln\mathcal{L}(\widehat{C}_\ell|C_\ell)}{\partial p_i} = \frac{2\ell+1}{2}\frac{\partial C_\ell}{\partial p_i}\left[-\frac{\widehat{C}_\ell}{C_\ell^2} + \frac{1}{C_\ell}\right] \tag{117}$$

Fisher 行列は

$$\begin{aligned}
F_{ij} &= \left\langle \frac{\partial \ln \mathcal{L}(\widehat{C}_\ell|C_\ell)}{\partial p_i} \frac{\partial \ln \mathcal{L}(\widehat{C}_\ell|C_\ell)}{\partial p_j} \right\rangle \\
&= \left(\frac{2\ell+1}{2}\right)^2 \frac{\partial C_\ell}{\partial p_i}\frac{\partial C_\ell}{\partial p_j}\left[\frac{\langle \widehat{C}_\ell^2\rangle}{C_\ell^4} + \frac{1}{C_\ell^2} - \frac{\langle \widehat{C}_\ell\rangle}{C_\ell^3}\right] \\
&= \left(\frac{2\ell+1}{2}\right)^2 \frac{1}{C_\ell^2}\frac{\partial C_\ell}{\partial p_i}\frac{\partial C_\ell}{\partial p_j}\left[\frac{2\ell+3}{2\ell+1} + 1 - 2\right] = \frac{2\ell+1}{2}\frac{1}{C_\ell^2}\frac{\partial C_\ell}{\partial p_i}\frac{\partial C_\ell}{\partial p_j}
\end{aligned} \tag{118}$$

となり、Gauss 型の場合と等しい。これは、Wishart 分布が Gauss 分布から確率変数を変換して得られるものであり、確率変数の変換で Fisher 行列は不変であることによる。すなわち、データの変換を行っても、母数に対する情報量は変わらないことを意味する。

2.4.3　弱非 Gauss 近似：一次元

まずは一次元において尤度関数の Edgeworth 展開を考える：

$$\mathcal{L}(x) \equiv A[1+f(x)]\mathcal{L}^{\mathrm{G}}(x) \tag{119}$$

ここで $\mathcal{L}^{\mathrm{G}}$ は平均ゼロ、分散 σ の Gauss 分布とする。f は高次の cumulant で表される補正項である：

$$f(x) \equiv -\frac{c_3}{3!\sigma^3}h_3(x/\sigma) + \left(\frac{c_4}{4!\sigma^4}h_4(x/\sigma) + \frac{1}{2}\frac{c_3^2}{3!\sigma^6}h_6(x/\sigma)\right) + \ldots \tag{120}$$

また A は規格化因子である：

$$A = \left\{\int \mathrm{d}x\,(1+f(x))\mathcal{L}^{\mathrm{G}}(x)\right\}^{-1} \tag{121}$$

ただし、すべての cumulant が消える極限で $A \to 1$ とする。スコア関数は

$$V_i = \frac{\partial \ln(1+f)}{\partial p_i} + \frac{\partial \ln \mathcal{L}^{\mathrm{G}}}{\partial p_i} + \frac{\partial \ln A}{\partial p_i} \simeq \frac{\partial f}{\partial p_i} + \frac{\partial \ln \mathcal{L}^{\mathrm{G}}}{\partial p_i} \tag{122}$$

である。ただし補正項は $f \ll 1$ を満たすとし、A には p_i の依存性はないとした。以降では、$f \ll 1$ と考え、統計平均は $\mathcal{L}^{\mathrm{G}}$ で行うとする。

ここで Edgeworth 展開の一次までを考える。このとき、補正項の微分は

$$\frac{\partial f}{\partial p_i} = -\frac{1}{6\sigma^3}h_3(x/\sigma)\frac{\partial c_3}{\partial p_i} + \frac{c_3}{6}\frac{\partial}{\partial p_i}\frac{h_3(x/\sigma)}{\sigma^3} \tag{123}$$

となる。これらは x の奇数次しか含まないので、スコア関数の第一項と第二項の相関はゼロとなる。したがって Fisher 行列は

$$F_{ij} = \left\langle \frac{\partial f}{\partial p_i}\frac{\partial f}{\partial p_j}\right\rangle + F_{ij}^{\mathrm{G}} \tag{124}$$

と表される。ただし第二項は $\mathcal{L}^{\mathrm{G}}$ の Fisher 行列であり、

$$F_{ij}^{\mathrm{G}} = \frac{1}{2}\frac{\partial \ln \sigma^2}{\partial p_i} \tag{125}$$

となる。Fisher 行列の第一項は

$$\begin{aligned}
\left\langle \frac{\partial f}{\partial p_i}\frac{\partial f}{\partial p_j}\right\rangle &= \frac{1}{36\sigma^6}\frac{\partial c_3}{\partial p_i}\frac{\partial c_3}{\partial p_j}\langle [h_3(x/\sigma)]^2\rangle + \frac{c_3^2}{36}\left\langle \left[\frac{\partial}{\partial p_i}\frac{h_3(x/\sigma)}{\sigma^3}\right]\left[\frac{\partial}{\partial p_j}\frac{h_3(x/\sigma)}{\sigma^3}\right]\right\rangle \\
&\quad - \frac{c_3}{36\sigma^3}\frac{\partial c_3}{\partial p_i}\left\langle h_3(x/\sigma)\frac{\partial}{\partial p_j}\frac{h_3(x/\sigma)}{\sigma^3}\right\rangle - (i \leftrightarrow j)
\end{aligned} \tag{126}$$

ここで分散にはパラメータの情報は含まれないとして、その微分をゼロとする。このとき第一項のみが残る。$t = x/\sigma$ として、

$$\langle [h_3(t)]^2 \rangle = \langle (t^3 - 3t)^2 \rangle = \langle t^6 - 6t^4 + 9t^2 \rangle \tag{127}$$

である。変数変換 $x \to t$ をすることで

$$\langle \cdots \rangle = \int_{-\infty}^{\infty} \mathrm{d}t \, \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \tag{128}$$

に注意すると、

$$\langle [h_3(t)]^2 \rangle = 15 - 18 + 9 = 6 \tag{129}$$

となる。したがって、

$$F_{ij} = \frac{1}{6\sigma^6} \frac{\partial c_3}{\partial p_i} \frac{\partial c_3}{\partial p_j} + F^{\mathrm{G}}_{ij} \tag{130}$$

が得られる。

2.4.4　弱非 Gauss 近似：多次元

ここで、実際の観測ではデータベクトルは多次元なので一次元の計算方法を多次元に拡張する。Edgeworth 展開した尤度関数は、一次までを拾うと [2]

$$\mathcal{L} = (1 + f)\mathcal{L}^{\mathrm{G}} \tag{131}$$

ただし

$$\mathcal{L}^{\mathrm{G}} \equiv \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} \prod_{\ell m \ell' m'} \exp\left(-\frac{1}{2} a^*_{\ell m} \mathbf{C}^{-1} a_{\ell' m'} \right) \tag{132}$$

および

$$\begin{aligned} f &\equiv -\frac{1}{3!\mathcal{L}^{\mathrm{G}}} \sum_{\ell_i m_i} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \frac{\partial}{\partial a_{\ell_1 m_1}} \frac{\partial}{\partial a_{\ell_2 m_2}} \frac{\partial}{\partial a_{\ell_3 m_3}} \mathcal{L}^{\mathrm{G}} \\ &= \frac{1}{6} \sum_{\ell_i m_i} b_{\ell_1 \ell_2 \ell_3} \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \left[\overline{a}_{\ell_1 m_1} \overline{a}_{\ell_2 m_2} \overline{a}_{\ell_3 m_3} - (-1)^{m_3} \left(\frac{\overline{a}_{\ell_1 m_1}}{C_{\ell_3}} \delta_{\ell_2 \ell_3} \delta_{m_2, -m_3} + (\text{cyc.}) \right) \right] \end{aligned} \tag{133}$$

上式では、共分散行列は対角であるとし、$\overline{a}_{\ell m} = a_{\ell m} C^{-1}_{\ell m}$、また $b_{\ell_1 \ell_2 \ell_3}$ は Gaunt 積分

$$\mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \equiv \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{16\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{134}$$

を用いて

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} b_{\ell_1 \ell_2 \ell_3} \tag{135}$$

で定義される。Wigner-3j の性質から揺らぎの一次の項は消え、

$$f = \frac{1}{6} \sum_{\ell_i m_i} b_{\ell_1 \ell_2 \ell_3} \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \overline{a}_{\ell_1 m_1} \overline{a}_{\ell_2 m_2} \overline{a}_{\ell_3 m_3} \tag{136}$$

が導かれる。

いま、b にのみパラメータの依存性があるとすると、非 Gauss 性によって生じる Fisher 行列への補正項は

$$F^f_{ij} \equiv \frac{1}{36} \sum_{\ell_i m_i \ell'_i m'_i} \frac{\partial b_{\ell_1 \ell_2 \ell_3}}{\partial p_i} \frac{\partial b_{\ell'_1 \ell'_2 \ell'_3}}{\partial p_i} \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \mathcal{G}^{\ell'_1 \ell'_2 \ell'_3}_{m'_1 m'_2 m'_3} \langle \overline{a}_{\ell_1 m_1} \cdots \overline{a}_{\ell'_3 m'_3} \rangle \tag{137}$$

6 点相関を 2 点に分解するさい、プライムのつくものとつかないものでペアをとらない場合はすべて Wigner-3j の性質によって消える。これを利用すると、6 点相関は 6 個の項に分解でき、$b_{\ell_1\ell_2\ell_3} = (-1)^{\ell_1+\ell_2+\ell_3} b_{\ell_1\ell_3\ell_2}$ や Wigner-3j の対称性を使うと、補正項は

$$\begin{aligned}
F^f_{ij} &= \frac{1}{6}\sum_{\ell_1\ell_2\ell_3 m_1 m_2 m_3} \frac{\partial b_{\ell_1\ell_2\ell_3}}{\partial p_i}\frac{\partial b_{\ell'_1\ell'_2\ell'_3}}{\partial p_i}\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3}\mathcal{G}^{\ell_1\ell_2\ell_3}_{-m_1,-m_2,-m_3}\frac{1}{C_{\ell_1}C_{\ell_2}C_{\ell_3}} \\
&= \frac{1}{6}\sum_{\ell_1\ell_2\ell_3}\frac{1}{C_{\ell_1}C_{\ell_2}C_{\ell_3}}\frac{\partial b_{\ell_1\ell_2\ell_3}}{\partial p_i}\frac{\partial b_{\ell'_1\ell'_2\ell'_3}}{\partial p_i}\sum_{m_1m_2m_3}\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3}\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3} \\
&= \frac{1}{6}\sum_{\ell_1\ell_2\ell_3}\frac{1}{C_{\ell_1}C_{\ell_2}C_{\ell_3}}\frac{\partial B_{\ell_1\ell_2\ell_3}}{\partial p_i}\frac{\partial B_{\ell'_1\ell'_2\ell'_3}}{\partial p_i}
\end{aligned} \tag{138}$$

ただし

$$B_{\ell_1\ell_2\ell_3} \equiv \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{16\pi}}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0\end{pmatrix} b_{\ell_1\ell_2\ell_3} \tag{139}$$

および

$$\sum_{m_1m_2m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3\end{pmatrix}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3\end{pmatrix} = 1 \tag{140}$$

を用いた。

2.4.5 有限観測領域での Fisher 行列

実際の観測ではピクセル数は有限であり情報は有限しかない。そのため、$\ell = \ell_{\max}$ より小さいスケールの情報は得られないと考えられる。また観測領域が有限の場合は、全天に対する観測領域の割合 $f_{\rm sky}$ を用いて

$$F_{ij} = \sum_{\ell=2}^{\ell_{\max}}\frac{(2\ell+1)f_{\rm sky}}{2}\mathrm{Tr}\left(\boldsymbol{C}_\ell^{-1}\frac{\partial \boldsymbol{C}_\ell}{\partial p_i}\boldsymbol{C}_\ell^{-1}\frac{\partial \boldsymbol{C}_\ell}{\partial p_j}\right) \tag{141}$$

2.5 系統誤差の評価

パワースペクトルの理論計算に誤りがある場合、それが伝搬してパラメータ推定にバイアスを生じる。ここでは、誤った理論パワースペクトル $\widetilde{C}$ を用いて推定した宇宙論パラメータと真の宇宙論パラメータとの差 $\delta\boldsymbol{p}$ を定量的に評価する方法について述べる。

尤度関数は平均 0、共分散行列 $\mathbf{C}$ の Gauss 分布に従うとし、共分散行列に理論のパワースペクトルが含まれている。バイアスが小さいと仮定し、誤りを含む理論パワースペクトルをもとに計算された対数尤度 $\ln\mathcal{L}'$ を真のパラメータ $\overline{\boldsymbol{p}}$ 周りで展開して 2 次までの寄与をとると

$$\ln\mathcal{L}'(\boldsymbol{p}) = \ln\mathcal{L}'(\overline{\boldsymbol{p}}) + \sum_i\frac{\partial\ln\mathcal{L}'(\overline{\boldsymbol{p}})}{\partial p_i}(p_i - \overline{p}_i) + \frac{1}{2}\sum_{ij}(p_i-\overline{p}_i)\frac{\partial^2\ln\mathcal{L}'(\overline{p})}{\partial p_i\partial p_j}(p_j - \overline{p}_j) \tag{142}$$

となる。バイアスされた推定量 $\overline{\boldsymbol{p}} + \delta\boldsymbol{p}$ は、パラメータ空間において対数尤度を最小にする。そこで、上式をさらに p^k で微分したものを考える：

$$\frac{\partial\ln\mathcal{L}'(\boldsymbol{p})}{\partial p_k} = \frac{\partial\ln\mathcal{L}'(\overline{\boldsymbol{p}})}{\partial p_k} + \sum_j\frac{\partial^2\ln\mathcal{L}'(\overline{\boldsymbol{p}})}{\partial p_k\partial p_j}(p_j - \overline{p}_j) \tag{143}$$

上式に $\boldsymbol{p} = \overline{\boldsymbol{p}} + \delta\boldsymbol{p}$ を代入すると左辺は消える。さらにデータに関して統計平均をとることで

$$\langle\delta V_i(\overline{\boldsymbol{p}})\rangle - \sum_j\mathbf{F}'_{ij}(\overline{\boldsymbol{p}})\delta p_j = 0 \tag{144}$$

となる。ここで δV_i はバイアスされたスコア関数と真のスコア関数の差、$\mathbf{F}'_{ij}$ はバイアスされた Fisher 行列である：

$$\delta V_i(\boldsymbol{d},\boldsymbol{p}) \equiv \frac{\partial \ln \mathcal{L}'(\boldsymbol{d},\boldsymbol{p})}{\partial p_i} - \frac{\partial \ln \mathcal{L}(\boldsymbol{d},\boldsymbol{p})}{\partial p_i} \tag{145}$$

$$\mathbf{F}'_{ij}(\boldsymbol{p}) \equiv -\left\langle \frac{\partial^2 \ln \mathcal{L}'(\boldsymbol{p})}{\partial p_i \partial p_j} \right\rangle \tag{146}$$

これより、

$$\delta p_i = \sum_{ij} \{\mathbf{F'}_{ij}\}^{-1} \langle \delta V_i \rangle \bigg|_{\boldsymbol{p}=\overline{\boldsymbol{p}}} \tag{147}$$

が得られる。右辺の δV_i は、バイアスされた共分散行列 $\mathbf{C}'$ を用いると

$$-2\delta V_i = \mathrm{Tr}\left(\mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i} - \mathbf{C'}^{-1}\frac{\partial \mathbf{C}'}{\partial p_i}\mathbf{C'}^{-1}\boldsymbol{\Delta} + \mathbf{C'}^{-1}\frac{\partial \boldsymbol{\Delta}}{\partial p_i} \right) \tag{148}$$

と書けるので、

$$\begin{aligned} -2\langle \delta V_i \rangle &= \mathrm{Tr}\left(\mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i} - \mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i}\mathbf{C'}^{-1}\langle\boldsymbol{\Delta}\rangle + \mathbf{C'}^{-1}\left\langle\frac{\partial \boldsymbol{\Delta}}{\partial p_i}\right\rangle \right) \\ &= \mathrm{Tr}\left(\mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i} - \mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i}\mathbf{C'}^{-1}\mathbf{C} \right) \\ &= \mathrm{Tr}\left(-\mathbf{C'}^{-1}\frac{\partial \mathbf{C'}}{\partial p_i}\mathbf{C'}^{-1}(\mathbf{C}-\mathbf{C}') \right) \end{aligned} \tag{149}$$

となる。ただし、$\partial\langle\boldsymbol{\Delta}\rangle/\partial p_i = 0$ を用いた。これを式（147）に代入すると、共分散行列の差 $\delta\mathbf{C} \equiv \mathbf{C}' - \mathbf{C}$ を用いて

$$\delta p_i = -\sum_j \{F'_{ij}\}^{-1}\frac{1}{2}\mathrm{Tr}\left(\mathbf{C'}^{-1}\frac{\partial \mathbf{C}'}{\partial p_j}\mathbf{C'}^{-1}\delta\mathbf{C} \right) \tag{150}$$

が得られる。

3　宇宙論における推定 I:揺らぎとその角度パワースペクトル

3.1　CMB 観測における尤度関数

3.1.1　ビームパターン

ある方向（ここではピクセル）において CMB を検出したさい、その信号は有限の広がりをもつ。i 番目のピクセルにおいて検出されるシグナル s_i は、i 番目のピクセルにおける光の揺らぎ X_i、i 番目のピクセルにおけるビームパターン B_i を用いて

$$s_i = \int \mathrm{d}\hat{\boldsymbol{n}}\, X_i(\hat{\boldsymbol{n}}) B_i(\hat{\boldsymbol{n}}) \tag{151}$$

と表される。ここで、ビームパターンは各ピクセルごとの特性を表す。ここでは簡単のため、$\hat{\boldsymbol{n}}_i$ を i 番目のピクセルの方向とし、ビームパターン B_i を

$$B_i(\hat{\boldsymbol{n}}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}_i)}{2\sigma^2} \right) \tag{152}$$

のように（Gauss 型を）仮定する。このとき、ビームパターンを Fourier 変換したものは

$$B_i(\boldsymbol{\ell}) = e^{-\ell^2\sigma^2} \tag{153}$$

となる。σ は角度分解能を意味する。以降では FWHM に直した角度分解能 $\theta \equiv \sqrt{8\ln 2}\sigma$ [arcmin] を用いる。

3.1.2　ホワイトノイズ

ビームパターンを角度分解能 θ の Gauss 型と仮定する。本来のシグナル X に一様なノイズ n^X が加わると、測定されるパワースペクトル $\hat{C}_\ell^{XX}$ は

$$\hat{C}_\ell^{XX} = e^{-\ell^2\theta^2/8\ln 2}(C_\ell^{XX} + N_\ell^{XX}) \tag{154}$$

となる。ただし

$$N_\ell^{XX}\delta_{\ell\ell'}\delta_{mm'} \equiv \langle n_{\ell m}^{X,*} n_{\ell m}^{X}\rangle e^{\ell^2\theta^2/8\ln 2} \tag{155}$$

であり、ノイズとシグナルの相関は 0 とした。

$\langle n_{\ell m}^{X,*} n_{\ell m}^{X}\rangle$ は、揺らぎ X を測定するさいのノイズの統計的性質を表す。ここでは、ノイズは一様であるとする。CMB の温度 $T_{\rm CMB}$ を用い、感度 σ_X を

$$\frac{\sigma_X}{T_{\rm CMB}} \equiv n^X \tag{156}$$

と定義すると、ノイズのスペクトル N_ℓ^{XX} は

$$N_\ell^{XX}\delta_{\ell\ell'} \equiv \left(\frac{\sigma_X}{T_{\rm CMB}}\right)^2 e^{\ell^2\theta^2/8\ln 2} \tag{157}$$

と表せる。

複数の振動数で観測を行っている場合、複数の CMB のマップを得ることができる。それぞれから得られたデータを組み合わせることで、ノイズの寄与を小さくすることができる。i を振動数 ν_i のチャンネルとして

$$\frac{1}{N} \equiv \sum_i \frac{1}{N_i} = \sqrt{\sum_{i\geq j}\frac{2}{N_iN_j(1+\delta_{ij})}} \tag{158}$$

と一般化されたノイズ N を用いると、データを組み合わせた場合に生じるノイズの寄与が最小になる。

3.1.3　尤度関数

得られるデータ $\boldsymbol{D}$ を、温度揺らぎ・偏光の展開係数 $X_{\ell m}, Y_{\ell m}$ とする。このとき、部分データベクトル

$$\boldsymbol{D}_\ell = (X_{\ell,-\ell}, X_{\ell,-\ell+1}, \ldots, X_{\ell,\ell}, Y_{\ell,-\ell}, \ldots, Y_{\ell,\ell}) \tag{159}$$

を用いて、データベクトルを

$$\boldsymbol{D} = \{\boldsymbol{D}_\ell\}_{\ell=2,3,\ldots} \tag{160}$$

と表すことにする。式（73）で定義される共分散行列は

$$\boldsymbol{C} = \delta_{\ell\ell'}\langle \boldsymbol{D}_\ell^\dagger \boldsymbol{D}_\ell\rangle = \begin{pmatrix} \langle \boldsymbol{D}_2^\dagger \boldsymbol{D}_2\rangle & & \\ & \langle \boldsymbol{D}_3^\dagger \boldsymbol{D}_3\rangle & \\ & & \ddots \end{pmatrix} \tag{161}$$

のように、部分行列 $\boldsymbol{c}_\ell \equiv \langle \boldsymbol{D}_\ell^\dagger \boldsymbol{D}_\ell\rangle (\ell = 2, 3, \ldots)$ を用いて表される。$\boldsymbol{c}_\ell$ は

$$\boldsymbol{c}_\ell = \begin{pmatrix} \boldsymbol{C}_\ell^{\Theta\Theta} & \boldsymbol{C}_\ell^{\Theta E} \\ \boldsymbol{C}_\ell^{\Theta E} & \boldsymbol{C}_\ell^{EE} \end{pmatrix} \tag{162}$$

と表される。ただし $\boldsymbol{C}_\ell^{XY}$ は

$$\boldsymbol{C}_\ell^{XY} \equiv \delta_{mm'}C_\ell^{XY} = \begin{pmatrix} C_\ell^{XY} & & \\ & \ddots & \\ & & C_\ell^{XY} \end{pmatrix} \tag{163}$$

のように、パワースペクトル C_ℓ^{XY} を対角成分にもつ $2\ell+1$ 次元の対角行列である。

尤度関数 $\mathcal{L}$

3.2　角度パワースペクトルの推定

宇宙論解析において興味がある量は、揺らぎから構成される角度パワースペクトルなどの統計量である。

簡単な具体例として、いま、観測した温度揺らぎ $\widehat{\Theta}(\hat{\boldsymbol{n}})$ から、その角度パワースペクトルの推定量を構築することを考える。以下では、観測量、あるいは推定量にはすべて ^ をつけて区別する。$\Theta(\hat{\boldsymbol{n}})$ は、宇宙論パラメータを決めた段階で得られる 1 realization の温度ゆらぎとする。観測器ノイズ (Instrumental noise)、前景輻射 (Foreground) などが存在せず、$\widehat{\Theta}(\hat{\boldsymbol{n}}) = \Theta(\hat{\boldsymbol{n}})$ と仮定する。また、異なる Fourier モードが互いに独立で、各モードは平均 0、分散 $C_\ell^{\Theta\Theta}$ の Gauss 分布に従うとする。

3.2.1　温度揺らぎの角度パワースペクトル

理論的には、温度揺らぎの角度パワースペクトルは

$$\delta_{\ell\ell'}\delta_{mm'}C_\ell^{\Theta\Theta} = \langle\Theta_{\ell m}\Theta_{\ell' m'}\rangle_\Theta\,, \tag{164}$$

のように定義される。今は一様等方性を仮定しているため、異なる m は独立である。

3.2.2　不偏性

いま、角度パワースペクトルの推定量 $\widehat{C}_\ell^{\Theta\Theta}$ を構築するため、条件 1 から考える。すなわち、

$$\langle\widehat{C}_\ell^{\Theta\Theta}\rangle_\Theta = C_\ell^{\Theta\Theta}\,. \tag{165}$$

この条件 1 を満たす推定量はいろいろとれる。例えば、単純に角度パワースペクトルの定義式 (164) から、$-\ell \le m \le \ell$ として

$$\widehat{C}_{\ell,(m)}^{\Theta\Theta} = |\widehat{\Theta}_{\ell m}|^2\,, \tag{166}$$

とする。このとき、各 m に対し、$\widehat{C}_{\ell,(m)}^{\Theta\Theta}$ は $C_\ell^{\Theta\Theta}$ の不偏推定量である。

3.2.3　最小分散

次に最小分散について考える。(166) の分散より小さい推定量を構成することは可能だろうか。

■Optimal weighting　一般に、不偏性を満たす独立な n 個の推定量 $\widehat{X}_i$ $(i = 1, \ldots, n)$ から、その線形結合を用いて新たに

$$\widehat{X} = \sum_{i=1,\ldots,n} w_i \widehat{X}_i \tag{167}$$

として表される推定量を構築することを考える。このとき、不偏性、最小分散を満たす w_i は、Lagrange の未定乗数法から

$$w_i = \frac{1}{W\sigma^2(\widehat{X}_i)}\,, \tag{168}$$

$$W = \sum_{i=1}^{n}\frac{1}{\sigma^2(\widehat{X}_i)}\,. \tag{169}$$

また、$1/W$ は estimator の分散となる。

■最小分散をもつ温度揺らぎの角度パワースペクトルの推定量　いま、独立に $2\ell+1$ 個の推定量 $\widehat{C}_{\ell,(m)}^{\Theta\Theta}$ とることができる。これらの線形結合を考えると

$$\widehat{C}_\ell^{\Theta\Theta} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}|\widehat{\Theta}_{\ell m}|^2\,. \tag{170}$$

3.3　温度揺らぎの角度パワースペクトル：最尤推定

ここでは、温度揺らぎの角度パワースペクトルを例にとり、宇宙論における最尤推定法について説明する。

3.3.1　例：角度パワースペクトルに対する最尤推定法

$\boldsymbol{C}^{\boldsymbol{\Theta\Theta}} = \{C_\ell^{\Theta\Theta}\}$ が与えられたときに、温度揺らぎ $\boldsymbol{\Theta}$ を実現する確率は

$$P(\boldsymbol{\Theta}|\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}) \propto \frac{1}{\sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}\boldsymbol{\Theta}^\dagger \mathbf{C}^{-1}\boldsymbol{\Theta}\right) . \tag{171}$$

ここで、$\mathbf{C} = \langle \boldsymbol{\Theta}\boldsymbol{\Theta}^\dagger \rangle$ は共分散行列であり、規格化定数は除いた。この尤度関数を最大にする角度パワースペクトルは

$$\left.\frac{\partial P(\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}|\boldsymbol{\Theta})}{\partial C_\ell^{\Theta\Theta}}\right|_{\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}=\widehat{\boldsymbol{C}}^{\boldsymbol{\Theta\Theta}}} = 0 . \tag{172}$$

あるいは、$\mathcal{L} \equiv -2\ln P$ として、

$$\mathcal{L}(\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}|\boldsymbol{\Theta}) = \mathrm{Tr}\ln\mathbf{C} + \boldsymbol{\Theta}^\dagger \mathbf{C}^{-1}\boldsymbol{\Theta} + \mathrm{const.} , \tag{173}$$

を用いれば、推定量は

$$\left.\frac{\partial \mathcal{L}(\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}|\boldsymbol{\Theta})}{\partial C_\ell^{\Theta\Theta}}\right|_{\boldsymbol{C}^{\boldsymbol{\Theta\Theta}}=\widehat{\boldsymbol{C}}^{\boldsymbol{\Theta\Theta}}} = 0 , \tag{174}$$

を満たす。ここで、$\ln\det\mathbf{C} = \mathrm{Tr}\ln\mathbf{C}$ を用いた。実際に微分すると、

$$\mathrm{Tr}\left(\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial C_\ell^{\Theta\Theta}}\right) - \boldsymbol{\Theta}^\dagger \mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial C_\ell^{\Theta\Theta}}\mathbf{C}^{-1}\boldsymbol{\Theta} = 0 . \tag{175}$$

今、任意の行列 $\mathbf{A}$ に対して

$$\mathrm{Tr}(\mathbf{A}) = \langle \boldsymbol{\Theta}^\dagger \mathbf{A}\mathbf{C}^{-1}\boldsymbol{\Theta}\rangle \tag{176}$$

が成り立つことを利用すれば、$\overline{\boldsymbol{\Theta}} = \mathbf{C}^{-1}\boldsymbol{\Theta}$ (inverse variance filtered multipole) を用いて

$$\left\langle \overline{\boldsymbol{\Theta}}^\dagger \frac{\partial \mathbf{C}}{\partial C_\ell^{\Theta\Theta}}\overline{\boldsymbol{\Theta}}\right\rangle - \overline{\boldsymbol{\Theta}}^\dagger \frac{\partial \mathbf{C}}{\partial C_\ell^{\Theta\Theta}}\overline{\boldsymbol{\Theta}} = 0 . \tag{177}$$

この方程式の解が推定量 $\widehat{\Theta}$ となる。

3.3.2　理想的な場合

共分散行列を $\{\mathbf{C}\}_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2} = \langle \Theta_{\boldsymbol{\ell}_1}\Theta^*_{\boldsymbol{\ell}_2}\rangle = \delta_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2} C_{\ell_1}^{\Theta\Theta}$ と仮定すると、

$$\frac{1}{\widehat{C}_\ell^{\Theta\Theta}} - \frac{1}{(\widehat{C}_\ell^{\Theta\Theta})^2}\int_{-\pi}^{\pi}\frac{\mathrm{d}\varphi_{\boldsymbol{\ell}}}{2\pi}|\Theta_{\boldsymbol{\ell}}|^2 = 0 , \tag{178}$$

であり、最尤推定量 $\widehat{C}_\ell^{\Theta\Theta}$ は式 (170) に一致する。

3.3.3　窓関数の影響

観測領域が限定される場合：

$$\widehat{\Theta}(\hat{\boldsymbol{n}}) = W(\hat{\boldsymbol{n}})\Theta(\hat{\boldsymbol{n}}) . \tag{179}$$

$W(\hat{\boldsymbol{n}})$ は窓関数で、観測領域で 1、それ以外で 0 をとる関数。フーリエ空間では

$$\widehat{\Theta}_{\boldsymbol{\ell}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} W_{\boldsymbol{L}-\boldsymbol{\ell}}\Theta_{\boldsymbol{L}} . \tag{180}$$

共分散行列は

$$\langle \widehat{\Theta}_{\boldsymbol{\ell}} \widehat{\Theta}^*_{\boldsymbol{\ell}'} \rangle = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, W_{\boldsymbol{L}-\boldsymbol{\ell}} W^*_{\boldsymbol{L}-\boldsymbol{\ell}'} C_L^{\Theta\Theta} \tag{181}$$

擬パワースペクトルは

$$\widehat{C}_\ell \delta_{\mathbf{0}} \equiv \langle \widehat{\Theta}_{\boldsymbol{\ell}} \widehat{\Theta}^*_{\boldsymbol{\ell}'} \rangle = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, M_{\boldsymbol{L},\boldsymbol{\ell}} C_L^{\Theta\Theta} \simeq \sum_{\boldsymbol{L}} \{\mathbf{M}\}_{\boldsymbol{L},\boldsymbol{\ell}} C_L^{\Theta\Theta} \tag{182}$$

ただし

$$M_{\boldsymbol{L},\boldsymbol{\ell}} \equiv \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, W^2_{\boldsymbol{L}-\boldsymbol{\ell}} \tag{183}$$

であり、$\mathbf{M}$ は $M_{\boldsymbol{L},\boldsymbol{\ell}}$ を成分にもつ行列である。したがって、測定したいパワースペクトルは擬パワースペクトルに $\mathbf{M}^{-1}$ をかけて得られる。ただし、この逆行列の操作は、実験によっては数値計算上困難な場合もある。

3.4 銀河サーベイ

3.4.1 ノイズ

ここでは、銀河の空間分布から数密度揺らぎを推定する場合、および銀河のシアを推定する場合のノイズについて考察する。

銀河数密度の場合をまず考える。ある微小体積を考え、その中での銀河の数密度を $n(\boldsymbol{x})$ とする。いま、この揺らぎの推定量として

$$\widehat{\delta}(\boldsymbol{x}) = \frac{1}{\bar{n}} n(\boldsymbol{x}) - 1 \tag{184}$$

を考えてみる。ただし $\bar{n}$ は全天での銀河数である。この推定量の統計平均は 0 である。このフーリエ変換は

$$\widehat{\delta}_{\boldsymbol{k}} = \frac{1}{\bar{n}} \int \mathrm{d}3\boldsymbol{x} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \, n(\boldsymbol{x}) \tag{185}$$

である。ただし $k = 0$ のモードは無視する。ここで、密度揺らぎがない場合、異なる座標の銀河数密度は各宇宙において独立に、かつ同じ確率分布で決まると考えられる。2 つの領域での数密度の相関は

$$\langle n(\boldsymbol{x}) n(\boldsymbol{x}') \rangle = \sigma^2 \delta_D^3(\boldsymbol{x} - \boldsymbol{x}') \tag{186}$$

となる。分散 σ^2 は、ある特定の領域の銀河数を考えることで

$$\langle [n(\boldsymbol{x}) d^3\boldsymbol{x}][n(\boldsymbol{x}') d^3\boldsymbol{x}'] \rangle = \left\langle \sum_{i\in\boldsymbol{x}} \sum_{j\in\boldsymbol{x}'} \right\rangle = d^3\boldsymbol{x}' \delta_D^3(\boldsymbol{x} - \boldsymbol{x}') \sum_{i\in\boldsymbol{x}} = n(\boldsymbol{x}) d^3\boldsymbol{x} d^3\boldsymbol{x}' \delta_D^3(\boldsymbol{x} - \boldsymbol{x}') \tag{187}$$

となることを使い、両辺を積分することで

$$N = \int \mathrm{d}3\boldsymbol{x} \, n(\boldsymbol{x}) = \sigma^2 \int \mathrm{d}3\boldsymbol{x} \tag{188}$$

すなわち $\sigma^2 = \bar{n}$ である。このとき、異なるフーリエモードの相関は

$$\langle \widehat{\delta}_{\boldsymbol{k}} \widehat{\delta}^*_{\boldsymbol{k}'} \rangle = \frac{1}{\bar{n}} \delta_D^3(\boldsymbol{k} - \boldsymbol{k}') \tag{189}$$

となる。これは銀河分布の推定におけるショットノイズである。

次に、銀河シアのノイズを考える。シアの場合は、もともとの銀河の形状が歪んでいるのでそれをエラーとして計上する。シグナルがないとした場合、推定量は

$$\widehat{\gamma}(\boldsymbol{x}) = \epsilon(\boldsymbol{x}) \tag{190}$$

である。数密度と同様に推定量をフーリエ変換し、その分散を考える。シアの形状ノイズは、各銀河でそれぞれ独立であり、かつ同じ確率分布に従うとする。このとき、

$$\langle \epsilon(\boldsymbol{x}) \epsilon(\boldsymbol{x}') \rangle = \sigma^2 \delta_D^3(\boldsymbol{x} - \boldsymbol{x}') \tag{191}$$

である。ある特定の領域の銀河数を考えることで

$$\langle[\epsilon(\boldsymbol{x})d^3\boldsymbol{x}][\epsilon(\boldsymbol{x}')d^3\boldsymbol{x}']\rangle = \left[\sum_{i\in\boldsymbol{x}}\sum_{j\in\boldsymbol{x}'}\right]^{-1}\left\langle\sum_{i\in\boldsymbol{x}}e_i\sum_{j\in\boldsymbol{x}'}e_j\right\rangle$$
$$= \sigma_{\rm int}^2 d^3\boldsymbol{x}'\delta_D^3(\boldsymbol{x}-\boldsymbol{x}')\left[\sum_{i\in\boldsymbol{x}}\right]^{-1} = \sigma_{\rm int}^2 n^{-1}(\boldsymbol{x})d^3\boldsymbol{x}d^3\boldsymbol{x}'\delta_D^3(\boldsymbol{x}-\boldsymbol{x}') \quad (192)$$

数密度の場合と同様にして

$$\sigma_{\rm int}^2\int \mathrm{d}3\boldsymbol{x} \ = \sigma^2\int \mathrm{d}3\boldsymbol{x}\, n(\boldsymbol{x}) \quad (193)$$

すなわち $\sigma^2 = \sigma_{\rm int}^2/\bar{n}$

3.4.2 重み関数

銀河サーベイで得られた分布に適度な重み $W(z)$ をつけ、S/N の最大化を考える。推定量は

$$\widehat{\delta}(\hat{\boldsymbol{n}}) = \int \mathrm{d}z\, W(z)\delta(z,\hat{\boldsymbol{n}}) \quad (194)$$

である。この推定量の分散を最小にし、かつ

$$\int \mathrm{d}z\, W(z) = 1 \quad (195)$$

を満たすように $W(z)$ を決める。分散としてショットノイズのみを考えると、

$$\langle|\widehat{\delta}(\hat{\boldsymbol{n}})|^2\rangle = \int \mathrm{d}z\,\frac{1}{n(z)}W^2(z) \quad (196)$$

である。ただし $n(z)$ は奥行き方向の銀河分布である。ラグランジュの未定乗数 λ を用いて

$$F(W) = \int \mathrm{d}z\,\frac{1}{n(z)}W^2(z) + \lambda\left(\int \mathrm{d}z\, W(z) - 1\right) \quad (197)$$

を最小にする W は

$$W(z) = \frac{\lambda}{2}n(z) = \frac{n(z)}{N} \quad (198)$$

となる。ただし N は $n(z)$ の積分である。$W(z)$ は、奥行き方向の銀河分布を 1 に規格化したものとなる。

4 宇宙論における推定 II:重力レンズ効果

4.1 レンズ場の推定：Quadratic 近似

レンズ場推定法の基本的アイデアは、Zaldarriaga & Seljak (1998) [3] において見受けられる。現在使われている方法は、 Hu & Okamoto (2002) の二点推定量（quadratic estimator）[4] に基づき、種々の改良を行い利用されている。

4.1.1 重力レンズによる非等方性

重力レンズ効果を受けた揺らぎは

$$\widetilde{\Theta}_{\boldsymbol{\ell}} = \Theta_{\boldsymbol{\ell}} - \sum_{x=\phi,\varpi}\int\frac{\mathrm{d}2\boldsymbol{L}}{2\pi}\,[(\boldsymbol{\ell}-\boldsymbol{L})\odot_x \boldsymbol{L}]x_{\boldsymbol{\ell}-\boldsymbol{L}}\Theta_{\boldsymbol{L}} + \mathcal{O}(x_{\boldsymbol{\ell}}^2)\,. \quad (199)$$

となる。重力レンズにより、温度揺らぎの異なるフーリエモードが混じる。

4.1.2 レンズ場推定量の不偏性

いま推定したい量は、我々の宇宙におけるレンズ場 $x_{\boldsymbol{\ell}}$ である。この推定量 $\widehat{x}_{\boldsymbol{\ell}}$ は、重力レンズを受ける前の温度揺らぎはある確率分布に従って生成されるため、

$$\langle \widehat{x}_{\boldsymbol{\ell}} \rangle_{\rm CMB} = x_{\boldsymbol{\ell}} \,, \tag{200}$$

が条件 1 となる。

4.1.3 不偏性を満たす推定量の候補

1 realization のレンズ場は、光源である温度揺らぎを非等方に歪める。そこで、$\boldsymbol{\ell} \neq \mathbf{0}$ として、異なるモードの相関に着目すると、

$$\langle \widetilde{\Theta}_{\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \rangle_{\rm CMB} = \sum_{x=\phi,\varpi} x_{\boldsymbol{\ell}} f^x_{\boldsymbol{\ell}\boldsymbol{L}} \,. \tag{201}$$

ただし

$$f^x_{\boldsymbol{\ell}\boldsymbol{L}} = [\boldsymbol{\ell} \odot_x \boldsymbol{L}] C^{\Theta\Theta}_L + [\boldsymbol{\ell} \odot_x (\boldsymbol{\ell} - \boldsymbol{L})] C^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|} \,. \tag{202}$$

これから、

$$x_{\boldsymbol{\ell}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} A_{\boldsymbol{\ell}} f^x_{\boldsymbol{\ell}\boldsymbol{L}} \langle \widetilde{\Theta}_{\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \rangle_{\rm CMB} \,; \qquad A_{\boldsymbol{\ell}} = \left\{ \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} (f^x_{\boldsymbol{\ell}\boldsymbol{L}})^2 \right\}^{-1} \,. \tag{203}$$

以上に基づくと、推定量の候補として以下の二点推定量

$$\widehat{x}_{\boldsymbol{\ell}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} A_{\boldsymbol{\ell}} f^x_{\boldsymbol{\ell}\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \,. \tag{204}$$

などが考えられる。しかし実際には、以下のようにして、この推定量に比べてさらにノイズの寄与を小さくすることが可能である。

4.1.4 分散の抑制

より一般化した二点推定量を考えてみる：

$$\widehat{x}_{\boldsymbol{\ell}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} F^x_{\boldsymbol{\ell}\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \,. \tag{205}$$

ここで、重み関数 $F^x_{\boldsymbol{\ell}\boldsymbol{L}}$ は、以下の二条件を満たす必要がある：

1. 不偏推定量であることから、$x_{\boldsymbol{\ell}} = \langle \widehat{x}_{\boldsymbol{\ell}} \rangle_{\rm CMB}$、あるいは

$$x_{\boldsymbol{\ell}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} F^x_{\boldsymbol{\ell}\boldsymbol{L}} \sum_{x'=\phi,\varpi} x'_{\boldsymbol{\ell}} f^{x'}_{\boldsymbol{\ell}\boldsymbol{L}} \,. \tag{206}$$

すなわち、

$$\int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} F^x_{\boldsymbol{\ell}\boldsymbol{L}} f^{x'}_{\boldsymbol{\ell}\boldsymbol{L}} = \delta^{xx'} \,. \tag{207}$$

2. 関数形を適切に選ぶことで、推定量におけるノイズの寄与をできるだけ小さくしたい。つまり

$$\frac{\delta}{\delta F^x_{\boldsymbol{\ell}\boldsymbol{L}}} \langle |\widehat{x}_{\boldsymbol{\ell}}|^2 \rangle_{\rm CMB} = 0 \,. \tag{208}$$

この 2 条件を満たす関数形 $F^x_{\boldsymbol{\ell}\boldsymbol{L}}$ は、以下のようにして一意に決まる。ノイズの分散は

$$\begin{aligned} N_\ell = \langle |\widehat{x}_{\boldsymbol{\ell}}|^2 \rangle &= \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \int \frac{\mathrm{d}2\boldsymbol{L}'}{2\pi} F^x_{\boldsymbol{\ell}\boldsymbol{L}} F^x_{\boldsymbol{\ell}\boldsymbol{L}'} \langle \widetilde{\Theta}_{\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \widetilde{\Theta}^*_{\boldsymbol{L}'} \widetilde{\Theta}^*_{\boldsymbol{\ell}-\boldsymbol{L}'} \rangle \\ &\simeq \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} F^x_{\boldsymbol{\ell}\boldsymbol{L}} (F^x_{\boldsymbol{\ell}\boldsymbol{L}} + F^x_{\boldsymbol{\ell},\boldsymbol{\ell}-\boldsymbol{L}}) \widetilde{C}^{\Theta\Theta}_L \widetilde{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|} \,. \end{aligned} \tag{209}$$

ただし、レンズを受けた 4 点相関は、Gaussian の部分 (disconnected part) だけを取り出した。普遍性を満たし、かつノイズ N_ℓ の Gaussian 部分 (Gaussian bias) を最小にする $F^x_{\boldsymbol{\ell},\boldsymbol{L}}$ の関数形は、Lagrange の未定乗数法を用いることで

$$F^x_{\boldsymbol{\ell}\boldsymbol{L}} = A^x_\ell \frac{f^x_{\boldsymbol{\ell}\boldsymbol{L}}}{2\widetilde{C}^{\Theta\Theta}_L \widetilde{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|}} \, ; \qquad A^x_\ell = \left\{ \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \frac{(f^x_{\boldsymbol{\ell}\boldsymbol{L}})^2}{2\widetilde{C}^{\Theta\Theta}_L \widetilde{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|}} \right\}^{-1} . \tag{210}$$

となることが分かる。

実際には式 (209) における Gaussian 以外の寄与は大きく、Planck やそれ以上の重力レンズに感度を持つ CMB 観測ではこれらの寄与は無視できなくなる [5, 6]。

4.2 Bias hardened estimator

ここでは、重力レンズ以外のモード相関から生じる系統誤差を取り除く方法について述べる [7]。

4.2.1 Mean-field bias

二点推定量の導出過程では、重力レンズ以外のモード相関の影響がないことを仮定している。そこで、例えば以下のようにマスク型の系統誤差（地球の固有運動、非等方再イオン化など）がある場合を考える：

$$\widehat{\Theta}(\hat{\boldsymbol{n}}) = (1 + \epsilon(\hat{\boldsymbol{n}}))\widetilde{\Theta}(\hat{\boldsymbol{n}}) \, . \tag{211}$$

ただし、ϵ は、その二次以上が無視できる微小量とする。このとき、

$$\widehat{\Theta}_{\boldsymbol{\ell}} = \widetilde{\Theta}_{\boldsymbol{\ell}} + \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, \epsilon_{\boldsymbol{\ell}-\boldsymbol{L}} \widetilde{\Theta}_{\boldsymbol{L}} \, , \tag{212}$$

であり、$\boldsymbol{\ell} \neq 0$ として

$$\langle \widehat{\Theta}_{\boldsymbol{L}} \widehat{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} \rangle = (\widetilde{C}^{\Theta\Theta}_L + \widetilde{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|} \epsilon_{\boldsymbol{\ell}} \equiv f^\epsilon_{\boldsymbol{\ell},\boldsymbol{L}} \epsilon_{\boldsymbol{\ell}} \, , \tag{213}$$

が得られる。この式から、レンズ場二点推定量の統計平均をとると

$$\langle \widehat{x}_{\boldsymbol{\ell}} \rangle_{\rm CMB} = x_{\boldsymbol{\ell}} + R^{x,\epsilon}_\ell \epsilon_{\boldsymbol{\ell}} \, , \qquad R^{x,\epsilon}_\ell = A^{xx}_\ell \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \frac{f^x_{\boldsymbol{\ell},\boldsymbol{L}} f^\epsilon_{\boldsymbol{\ell},\boldsymbol{L}}}{\widehat{C}^{\Theta\Theta}_L \widehat{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|}} \, . \tag{214}$$

となる。すなわち、$\widehat{x}_{\boldsymbol{\ell}}$ は不偏推定量になっていない。第二項は平均場バイアス（mean-field bias）と呼ばれる。このため、平均場バイアスを何かしらの方法で見積もる必要が出てくる。

4.2.2 Bias-hardened estimator: Mask-type distortion

まず、重力レンズと同じようにして ϵ の推定量を構築できる：

$$\langle \widehat{\epsilon}_{\boldsymbol{\ell}} \rangle_{\rm CMB} = \epsilon_{\boldsymbol{\ell}} + R^{\epsilon,x}_\ell x_{\boldsymbol{\ell}} \, . \tag{215}$$

ただし

$$A^{a,b}_\ell = \left\{ \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \frac{f^a_{\boldsymbol{\ell},\boldsymbol{L}} f^b_{\boldsymbol{\ell},\boldsymbol{L}}}{\widehat{C}^{\Theta\Theta}_L \widehat{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|}} \right\}^{-1} ; \qquad R^{a,b}_\ell = \frac{A^{a,a}_\ell}{A^{a,b}_\ell} \, . \tag{216}$$

ϵ に対する推定量も平均場バイアスを含む。そこで、$\widehat{x}$ の平均場バイアスと相殺するように $x_{\boldsymbol{\ell}}$ の推定量を定義しなおす：

$$\widehat{x}^{\rm BHE}_{\boldsymbol{\ell}} = \frac{\widehat{x}_{\boldsymbol{\ell}} - R^{x,\epsilon}_\ell \widehat{\epsilon}_{\boldsymbol{\ell}}}{1 - R^{\epsilon,x}_\ell R^{x,\epsilon}_\ell} \, . \tag{217}$$

4.2.3 Bias-hardened estimator: Point-source type distortion

同様のことは、unresolved point source や非一様ノイズ、非対称ビームがある場合にも行える。unresolved point source の場合は、

$$\widehat{\Theta}(\hat{\boldsymbol{n}}) = \widetilde{\Theta}(\hat{\boldsymbol{n}}) + n(\hat{\boldsymbol{n}}) \,, \tag{218}$$

と書ける。$n(\hat{\boldsymbol{n}})$ は、その点にける点光源の寄与であり、そこに点光源ができる確率は他の点と相関がないとする（すなわち Poisson 的である）。この場合に生じる平均場バイアスは

$$\langle \widehat{x}_{\boldsymbol{\ell}} \rangle_{\rm CMB} = x_{\boldsymbol{\ell}} + R^{x,s}_{\ell} s_{\boldsymbol{\ell}} \,, \qquad R^{x,s}_{\ell} = A^{xx}_{\ell} \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, \frac{f^x_{\boldsymbol{\ell},\boldsymbol{L}} f^s_{\boldsymbol{\ell},\boldsymbol{L}}}{\widehat{C}^{\Theta\Theta}_{L} \widehat{C}^{\Theta\Theta}_{|\boldsymbol{\ell}-\boldsymbol{L}|}} \,. \tag{219}$$

ただし $s(\hat{\boldsymbol{n}}) = \langle n^2(\hat{\boldsymbol{n}}) \rangle$、および $f^s_{\boldsymbol{\ell},\boldsymbol{L}} = 1$ である。これから、式 (217) と同様の手続きを行うことで、平均場バイアスの影響を受けない推定量を構成できる。

4.2.4 マスクによる平均場バイアス

上記の推定法は、known source に対しても利用できる。例えば、点光源マスクや有限観測領域の影響といった、ϵ の形が分かっている場合には、Monte Carlo シミュレーションなどに基づいて、その平均場バイアスが推定される。しかし、Monte Carlo シミュレーションに基づく方法の信頼性は、生成された擬似マップが観測されたマップをどの程度再現できているかに依存する。この信頼性をチェックする方法として、bias-hardened estimator で得られた結果と比較を行い、一致するかどうか調べることが考えられる [7]。Bias-hardened estimator においても、角度パワースペクトルの推定によっては系統誤差が生じるため、異なる手法による cross check を行うのが望ましいと考えられる。

4.3 レンズ場に対する最尤推定法

レンズ場二点推定量は、最尤推定法を簡略化することでも導くことができる [8]. ここでは温度揺らぎの場合に着目してその説明を行う。

4.3.1 確率分布関数

重力レンズ場 $x_{\boldsymbol{\ell}}$ と角度パワースペクトルが与えられたもとで、重力レンズを受けた温度揺らぎ $\widehat{\Theta}(\hat{\boldsymbol{n}})$ を得る確率は、重力レンズを受ける前の揺らぎがガウス統計であることを仮定すると、

$$\mathcal{L}(\widetilde{\boldsymbol{\Theta}}|\boldsymbol{x}) = \widetilde{\boldsymbol{\Theta}}^{\dagger}[\mathbf{C}(\boldsymbol{x})]^{-1}\widetilde{\boldsymbol{\Theta}} - \mathrm{Tr}\ln[\mathbf{C}(\boldsymbol{x})]^{-1} + \mathrm{const.} \,. \tag{220}$$

ただし、共分散行列は、重力レンズを受ける前の温度揺らぎに対する統計平均 $\langle\cdots\rangle_{\rm CMB}$ を用いて、

$$\{\mathbf{C}\}_{\boldsymbol{\ell},\boldsymbol{\ell}'} = \langle \widetilde{\Theta}_{\boldsymbol{\ell}} \widetilde{\Theta}^*_{\boldsymbol{\ell}'} \rangle_{\rm CMB} = \delta_{\boldsymbol{\ell}-\boldsymbol{\ell}'} C^{\Theta\Theta}_{\ell} + f^x_{\boldsymbol{\ell}+\boldsymbol{\ell}',\boldsymbol{\ell}} x_{\boldsymbol{\ell}+\boldsymbol{\ell}'} + \mathcal{O}(x^2) \,. \tag{221}$$

4.3.2 最尤推定

温度揺らぎの角度パワースペクトルの場合と同じ議論を行うことで、$x_{\boldsymbol{\ell}}$ に対する推定量が満たす式が得られる：

$$\mathcal{H}_{\boldsymbol{\ell}}(\widehat{x}_{\boldsymbol{\ell}}) - \langle \mathcal{H}_{\boldsymbol{\ell}}(\widehat{x}_{\boldsymbol{\ell}}) \rangle = 0 \,. \tag{222}$$

ここで

$$\mathcal{H}_{\boldsymbol{\ell}}(\widehat{x}_{\boldsymbol{\ell}}) = \overline{\boldsymbol{\Theta}}^{\dagger} \frac{\partial \mathbf{C}}{\partial x_{\boldsymbol{\ell}}} \overline{\boldsymbol{\Theta}} = \int \frac{\mathrm{d}2\boldsymbol{L}}{2\pi} \, f^x_{\boldsymbol{\ell}\boldsymbol{L}} \overline{\Theta}_{\boldsymbol{L}} \overline{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}} + \mathcal{O}(x) \,. \tag{223}$$

式 (222) を $x_{\boldsymbol{\ell}}$ について解くため、$x_{\boldsymbol{\ell}}$ に対して Taylor 展開すると

$$\mathcal{L} = \mathcal{L}|_{x_{\boldsymbol{\ell}}=0} + \sum_{\boldsymbol{\ell}} \frac{\partial \mathcal{L}}{\partial x_{\boldsymbol{\ell}}}\bigg|_{x_{\boldsymbol{\ell}}=0} x_{\boldsymbol{\ell}} + \sum_{\boldsymbol{\ell},\boldsymbol{\ell}'} \frac{1}{2} \frac{\partial^2 \mathcal{L}}{\partial x_{\boldsymbol{\ell}} \partial x'_{\boldsymbol{\ell}'}}\bigg|_{x_{\boldsymbol{\ell}}=0} x_{\boldsymbol{\ell}} x'_{\boldsymbol{\ell}'} + \mathcal{O}(x^3) \,. \tag{224}$$

これから

$$x_{\boldsymbol{\ell}} = \left[\left.\frac{\partial^2\mathcal{L}}{\partial x_{\boldsymbol{\ell}}\partial x'_{\boldsymbol{\ell}'}}\right|_{x_{\boldsymbol{\ell}}=0}\right]^{-1} \left.\frac{\partial\mathcal{L}}{\partial x_{\boldsymbol{\ell}}}\right|_{x_{\boldsymbol{\ell}}=0} . \tag{225}$$

さらに簡単化するため、規格化を統計平均で置き換えると

$$\left\langle \left.\frac{\partial^2\mathcal{L}}{\partial x_{\boldsymbol{\ell}}\partial x'_{\boldsymbol{\ell}'}}\right|_{x_{\boldsymbol{\ell}}=0} \right\rangle = \left\langle \left.\frac{\partial\mathcal{L}}{\partial x_{\boldsymbol{\ell}}}\frac{\partial\mathcal{L}}{\partial x'_{\boldsymbol{\ell}'}}\right|_{x_{\boldsymbol{\ell}}=0} \right\rangle = [\langle\mathcal{H}_{\boldsymbol{\ell}}\mathcal{H}_{\boldsymbol{\ell}'}\rangle - \langle\mathcal{H}_{\boldsymbol{\ell}}\rangle\langle\mathcal{H}_{\boldsymbol{\ell}'}\rangle]_{x_{\boldsymbol{\ell}}=0} = \mathcal{F}_{\boldsymbol{\ell}\boldsymbol{\ell}'} . \tag{226}$$

これを用いると

$$\widehat{x}_{\boldsymbol{\ell}} = \int\frac{\mathrm{d}2\boldsymbol{\ell}'}{2\pi}\,\mathcal{F}^{-1}_{\boldsymbol{\ell}\boldsymbol{\ell}'}[\mathcal{H}_{\boldsymbol{\ell}'} - \langle\mathcal{H}_{\boldsymbol{\ell}'}\rangle]_{x_{\boldsymbol{\ell}}=0} . \tag{227}$$

書き換えると、

$$\widehat{x}_{\boldsymbol{\ell}} = \widehat{x}^{\mathrm{S}}_{\boldsymbol{\ell}} - \langle\widehat{x}^{\mathrm{S}}_{\boldsymbol{\ell}}\rangle , \tag{228}$$

ただし

$$\widehat{x}^{\mathrm{S}}_{\boldsymbol{\ell}} = \int\frac{\mathrm{d}2\boldsymbol{\ell}'}{2\pi}\,\mathcal{F}^{-1}_{\boldsymbol{\ell}\boldsymbol{\ell}'}\frac{1}{2}\int\frac{\mathrm{d}2\boldsymbol{L}}{2\pi}\,f^{x}_{\boldsymbol{\ell}'\boldsymbol{L}}\overline{\Theta}_{\boldsymbol{L}}\overline{\Theta}_{\boldsymbol{\ell}'-\boldsymbol{L}} . \tag{229}$$

共分散行列が対角であれば、$\mathcal{F}$ は対角となり、その成分は Gaussian noise に一致する。したがって上記推定量は最初に導いたレンズ場二点推定量に帰着する。

4.3.3　二点推定量の導出過程との比較

最尤推定では、重力レンズを受けた温度揺らぎのパワースペクトルが出てこない。これは、「余分な寄与」の定義が二点推定におけるそれと異なることによる。二点推定では、**重力レンズを受けた**後の CMB 揺らぎのばらつきはノイズとして寄与する。これにより、重力レンズの推定は線形の連立方程式を解くことになり、ある程度簡略化される。一方、最尤推定では重力レンズを受けた温度揺らぎに含まれる重力レンズ効果も情報として取り出され、こちらのほうが情報量が多くなる。このとき、ノイズとして寄与するのは**重力レンズを受ける前**の CMB 揺らぎのばらつきである。ただし、計算を行う際には一般に非線形な連立方程式を解く必要があり、反復計算が必要となる。

二点推定量との違いは、偏光を用いた場合に顕著である。小角度スケールにおける B モード偏光の主成分は重力レンズ効果で生じるため、二点推定量と最尤推定におけるノイズが大きく異なる。上で述べたように、二点推定ではレンズを受けた後の CMB 揺らぎが、最尤推定ではレンズを受ける前の CMB 揺らぎがそれぞれノイズとして寄与する。B モード偏光を使う場合、二点推定は重力起源 B モードによるノイズが生じるのに対し、最尤推定ではこのノイズがない。Planck 程度の偏光観測ではほとんど違いが出ないが、小角度スケールに感度のある実験においては、最尤推定を用いることで検出精度が飛躍的に向上する。

それでは、理想的な極限（観測機器由来のノイズなし、全天観測など）では、最尤推定におけるノイズはゼロになるかというとそうではない。単純なマッピングで記述される重力レンズ起源 B モード以外に、多重散乱による寄与、重力ポテンシャルの非線形成長によるカールモードの存在によって限界がある。これらも同時推定すればよいのでは、となるが、自由度の観点から、すくなくとも有限ピクセル数では不可能だと考えられる。注意として、実はここらへんの議論は未だ収束していない（例えば、勾配、カールが両方ある場合の原理的限界はどうなっているのか、など)。

4.3.4　Bias-hardened estimator **との関連**

最尤推定法の観点では、bias-hardened estimator は平均場バイアスを生じる場に対し同時推定することを意味し、その簡略化した推定法が、すでに導いたものとなっている。もし、ϵ や n の高次の寄与が効いてくる場合には、簡略化した方法は使えず、反復計算などを用いる必要がある。

4.4　レンズ場の角度パワースペクトルの推定

二点推定で求めたレンズ場から、レンズ場の角度パワースペクトルを推定する方法を考える。

4.4.1　レンズ場高次のバイアス

レンズ場の推定量の自乗期待値は

$$\langle|\widehat{x}_{\boldsymbol{\ell}}|^2\rangle = \int\frac{\mathrm{d}2\boldsymbol{L}}{2\pi}\int\frac{\mathrm{d}2\boldsymbol{L}'}{2\pi}\,\langle\widetilde{\Theta}_{\boldsymbol{L}}\widetilde{\Theta}_{\boldsymbol{\ell}-\boldsymbol{L}}\widetilde{\Theta}^*_{\boldsymbol{L}'}\widetilde{\Theta}^*_{\boldsymbol{\ell}-\boldsymbol{L}'}\rangle\,. \tag{230}$$

これを

$$\langle|\widehat{x}_{\boldsymbol{\ell}}|^2\rangle = N^{x,(0)}_\ell + C^{xx}_\ell + N^{(1),x}_\ell + N^{(2),x}_\ell + \cdots\,, \tag{231}$$

と形式的に分解する。ここで $N^{x,(0)}_\ell = A^{xx}_\ell$ であり、$N^{(j),x}_\ell, (j\geq 1)$ は、左辺から $N^{x,(0)}_\ell + C^{xx}_\ell$ を引いた残りの寄与のうち、C^{xx}_ℓ の j 次の寄与を含むものである。

定義より、$N^{(1)}_\ell$ は、4 点相関のうち

$$\begin{aligned}&\langle\langle\Theta_{\boldsymbol{L}}\Theta^*_{\boldsymbol{L}'}\rangle_{\mathrm{CMB}}\langle\Theta_{\boldsymbol{\ell}-\boldsymbol{L}}\Theta^*_{\boldsymbol{\ell}-\boldsymbol{L}'}\rangle_{\mathrm{CMB}}\rangle + \langle\langle\Theta_{\boldsymbol{L}}\Theta^*_{\boldsymbol{\ell}-\boldsymbol{L}'}\rangle_{\mathrm{CMB}}\langle\Theta_{\boldsymbol{\ell}-\boldsymbol{L}}\Theta^*_{\boldsymbol{L}'}\rangle_{\mathrm{CMB}}\rangle \\ &= C^{xx}_{|\boldsymbol{L}-\boldsymbol{L}'|}f^x_{\boldsymbol{L}-\boldsymbol{L}',\boldsymbol{L}}f^x_{\boldsymbol{L}-\boldsymbol{L}',\boldsymbol{\ell}-\boldsymbol{L}} + C^{xx}_{|\boldsymbol{L}+\boldsymbol{L}'-\boldsymbol{\ell}|}f^x_{\boldsymbol{L}+\boldsymbol{L}'-\boldsymbol{\ell},\boldsymbol{L}}f^x_{\boldsymbol{L}+\boldsymbol{L}'-\boldsymbol{\ell},\boldsymbol{\ell}-\boldsymbol{L}}\,,\end{aligned} \tag{232}$$

から出てくる。

4.4.2　尤度推定を用いたレンズ場パワースペクトルの推定

重力レンズ場の角度パワースペクトルは、最尤推定法でも求めることができる [7]。この推定法は、上述の二点推定による方法に比べてより信頼性が高く、かつより検出精度の良い方法となっている。

重力レンズを受けた揺らぎの確率分布は、ガウス統計から少しずれたものになる。一般論として、いまガウス統計からずれた確率分布に従うある確率変数 a を考える。その 4 点相関の寄与を考えると、確率分布関数は

$$P(a) = \left[1 + \sum_{\boldsymbol{\ell}_i}\kappa^{(4)}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2\boldsymbol{\ell}_3\boldsymbol{\ell}_4}\frac{\partial}{\partial a_{\boldsymbol{\ell}_1}}\frac{\partial}{\partial a_{\boldsymbol{\ell}_2}}\frac{\partial}{\partial a_{\boldsymbol{\ell}_3}}\frac{\partial}{\partial a_{\boldsymbol{\ell}_4}}\right]P_{\mathrm{g}}(a)\,, \tag{233}$$

これから、共分散行列を $\mathbf{C}$ として

$$\begin{aligned}\mathcal{R}(a) &= \frac{P(a)}{P_{\mathrm{g}}(a)} \\ &= 1 + \sum_{\boldsymbol{\ell}_i}\kappa^{(4)}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2\boldsymbol{\ell}_3\boldsymbol{\ell}_4}\{\bar{a}_{\boldsymbol{\ell}_1}\bar{a}_{\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} - [\mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} + (5\text{ perm.})] + [\mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\mathbf{C}^{-1}_{\boldsymbol{\ell}_3\boldsymbol{\ell}_4} + (2\text{ perm.})]\} \\ &= 1 + \sum_{\boldsymbol{\ell}_i}\kappa^{(4)}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2\boldsymbol{\ell}_3\boldsymbol{\ell}_4}(\bar{a}_{\boldsymbol{\ell}_1}\bar{a}_{\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} - 6\mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} + 3\mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\mathbf{C}^{-1}_{\boldsymbol{\ell}_3\boldsymbol{\ell}_4})\,.\end{aligned} \tag{234}$$

ただし $\bar{a}_{\boldsymbol{\ell}} = \sum_{\boldsymbol{\ell}'}\mathbf{C}^{-1}_{\boldsymbol{\ell}\boldsymbol{\ell}'}a_{\boldsymbol{\ell}'}$ である。$a = \widetilde{\Theta}$ とすると、4 次のキュムラントは

$$\begin{aligned}\kappa^{(4)}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2\boldsymbol{\ell}_3\boldsymbol{\ell}_4} &= f_{\boldsymbol{\ell}_{12}\boldsymbol{\ell}_1}f_{-\boldsymbol{\ell}_{12},\boldsymbol{\ell}_3}\mathbf{C}^{\phi\phi}_{|\boldsymbol{\ell}_{12}|}\delta_{\boldsymbol{\ell}_{12},-\boldsymbol{\ell}_{34}} + f_{\boldsymbol{\ell}_{13}\boldsymbol{\ell}_1}f_{-\boldsymbol{\ell}_{13},\boldsymbol{\ell}_2}\mathbf{C}^{\phi\phi}_{|\boldsymbol{\ell}_{13}|}\delta_{\boldsymbol{\ell}_{13},-\boldsymbol{\ell}_{24}} \\ &\quad + f_{\boldsymbol{\ell}_{14}\boldsymbol{\ell}_1}f_{-\boldsymbol{\ell}_{14},\boldsymbol{\ell}_2}\mathbf{C}^{\phi\phi}_{|\boldsymbol{\ell}_{14}|}\delta_{\boldsymbol{\ell}_{14},-\boldsymbol{\ell}_{23}} + \mathcal{O}(|\phi|^4)\,,\end{aligned} \tag{235}$$

と書ける。ここで $\boldsymbol{\ell}_{ij} = \boldsymbol{\ell}_i + \boldsymbol{\ell}_j$ である。これを用いると

$$\begin{aligned}\mathcal{R}(a) = 1 + 3\sum_{\boldsymbol{\ell}_i}\{&f_{\boldsymbol{\ell}_{12}\boldsymbol{\ell}_1}f_{-\boldsymbol{\ell}_{34},\boldsymbol{\ell}_3}\mathbf{C}^{\phi\phi}_{|\boldsymbol{\ell}_{12}|}\delta_{\boldsymbol{\ell}_{12}\boldsymbol{\ell}_{34}}\bar{a}_{\boldsymbol{\ell}_1}\bar{a}_{\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} \\ &+ \kappa^{(4)}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2\boldsymbol{\ell}_3\boldsymbol{\ell}_4}[-2\mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\bar{a}_{\boldsymbol{\ell}_3}\bar{a}_{\boldsymbol{\ell}_4} + \mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_2}\mathbf{C}^{-1}_{\boldsymbol{\ell}_3\boldsymbol{\ell}_4}]\}\,,\end{aligned} \tag{236}$$

が得られる。

最尤推定法に基づくレンズ場角度パワースペクトルの推定量は、

$$\frac{\partial\mathcal{R}(a)}{\partial C^{xx}_\ell} = 0\,, \tag{237}$$

の解である。$\mathcal{R}(a)$ を C_ℓ^{xx} で展開すると

$$\mathcal{R}(a) = 1 + \sum_{\boldsymbol{\ell}} \mathcal{L}_\ell^{(1)} C_\ell^{xx} + \frac{1}{2} \sum_{\boldsymbol{\ell},\boldsymbol{\ell}'} \mathcal{L}_{\ell\ell'}^{(2)} C_\ell^{xx} C_{\ell'}^{xx} + \mathcal{O}([C_\ell^{xx}]^3)\,, \tag{238}$$

ここで

$$\mathcal{L}_\ell^{(1)} = \left.\frac{\partial \mathcal{L}}{\partial C_\ell^{xx}}\right|_{C_\ell^{xx}=0}, \qquad \mathcal{L}_{\ell\ell'}^{(2)} = \left.\frac{\partial^2 \mathcal{L}}{\partial C_\ell^{xx} \partial C_{\ell'}^{xx}}\right|_{C_\ell^{xx}=0}. \tag{239}$$

これより

$$\hat{C}_\ell^{xx} = \sum_{\ell'} \{\boldsymbol{\mathcal{L}}^{(2)}\}^{-1}_{\ell\ell'} \mathcal{L}_{\ell'}^{(1)}\,. \tag{240}$$

まずは $\mathcal{L}^{(1)}$ を評価する。4 次のキュムラントの式を用いると、

$$\begin{aligned}\mathcal{L}_\ell^{(1)} = 3 \sum_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_3} \{ & f_{\boldsymbol{\ell},\boldsymbol{\ell}_1} f_{-\boldsymbol{\ell},\boldsymbol{\ell}_3} \bar{a}_{\boldsymbol{\ell}_1} \bar{a}_{\boldsymbol{\ell}-\boldsymbol{\ell}_1} \bar{a}_{\boldsymbol{\ell}_3} \bar{a}_{-\boldsymbol{\ell}-\boldsymbol{\ell}_3} \\ & - 2[f_{\boldsymbol{\ell}\boldsymbol{\ell}_1} f_{-\boldsymbol{\ell},\boldsymbol{\ell}_3} \mathbf{C}^{-1}_{\boldsymbol{\ell}_1,\boldsymbol{\ell}-\boldsymbol{\ell}_1} \bar{a}_{\boldsymbol{\ell}_3} \bar{a}_{\boldsymbol{\ell}-\boldsymbol{\ell}_3} + 2 f_{\boldsymbol{\ell}\boldsymbol{\ell}_1} f_{-\boldsymbol{\ell},\boldsymbol{\ell}_3} \mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_3} \bar{a}_{\boldsymbol{\ell}-\boldsymbol{\ell}_1} \bar{a}_{-\boldsymbol{\ell}-\boldsymbol{\ell}_3}] \\ & + f_{\boldsymbol{\ell}\boldsymbol{\ell}_1} f_{-\boldsymbol{\ell},\boldsymbol{\ell}_3} \mathbf{C}^{-1}_{\boldsymbol{\ell}_1,\boldsymbol{\ell}-\boldsymbol{\ell}_1} \mathbf{C}^{-1}_{\boldsymbol{\ell}_3,\boldsymbol{\ell}-\boldsymbol{\ell}_3} + 2 f_{\boldsymbol{\ell}\boldsymbol{\ell}_1} f_{-\boldsymbol{\ell},\boldsymbol{\ell}_3} \mathbf{C}^{-1}_{\boldsymbol{\ell}_1\boldsymbol{\ell}_3} \mathbf{C}^{-1}_{\boldsymbol{\ell}-\boldsymbol{\ell}_1,-\boldsymbol{\ell}-\boldsymbol{\ell}_3} \}\,. \end{aligned} \tag{241}$$

が得られる。規格化されていないレンズ場二点推定量

$$\overline{x}_{\boldsymbol{\ell}}^{aa'} = \sum_{\boldsymbol{\ell}'} f^x_{\boldsymbol{\ell}\boldsymbol{\ell}'} \bar{a}_{\boldsymbol{\ell}'} \bar{a}'_{\boldsymbol{\ell}-\boldsymbol{\ell}'} \tag{242}$$

で書き直すと

$$\mathcal{L}_\ell^{(1)} \propto |\overline{x}_{\boldsymbol{\ell}}^{aa}|^2 - 2(\langle 2|\overline{x}_{\boldsymbol{\ell}}^{aa_1}|^2 - |\overline{x}^{a_1a_2}|^2\rangle_{1,2})\,. \tag{243}$$

ただし、平均場バイアスはゼロとした。また a_1, a_2 は $\mathbf{C}$ を共分散にもつランダムな変数である。$\langle\cdots\rangle_{1,2}$ はこれらに対する統計平均である。第一項は単に二点推定量のパワースペクトルである。それにノイズ除去を意味する第二項が加わっている。ノイズ項は観測量に依存している。この方法を用いることで、共分散行列に含まれる不定性の影響を受けづらくなる [7]。

規格化 $\mathcal{L}^{(2)}$ は、二階微分を評価するのは困難であるため、以下のように統計平均で置き換える：

$$\mathcal{L}_{\ell\ell'}^{(2)} \to \langle \mathcal{L}_{\ell\ell'}^{(2)} \rangle = \langle \mathcal{L}_\ell^{(1)} \mathcal{L}_{\ell'}^{(1)} \rangle \Big|_{C_\ell^{xx}=0}. \tag{244}$$

規格化はフィッシャー情報行列である。厳密な規格化は複雑であるが、計算可能とするために $\mathcal{L}_{\ell\ell'}^{(2)}$ は対角行列とする。このとき、上式右辺は $\overline{x}_{\boldsymbol{\ell}}^{aa}$ でパワースペクトルを計算した場合のノイズパワースペクトルの二乗で表される。このノイズパワースペクトルは、レンズ場二点推定量の規格化と一致するので、最終的に得られるパワースペクトルの推定量は

$$\hat{C}^{xx} = |\widehat{x}_{\boldsymbol{\ell}}^{aa}|^2 - 2(\langle 2|\widehat{x}_{\boldsymbol{\ell}}^{aa_1}|^2 - |\widehat{x}^{a_1a_2}|^2\rangle_{1,2})\,. \tag{245}$$

となる。

参考文献

[1] T. Matsubara, *Statistical perturbation theory of cosmic fields. 1. basic formalism and second order theory*, *Astrophys. J.* (2003) [astro-ph/0006269].

[2] D. Babich, *Optimal estimation of non-gaussianity*, *Phys. Rev. D* **72** (2005) 043003, [astro-ph/0503375].

[3] M. Zaldarriaga and U. Seljak, *Reconstructing projected matter density from cosmic microwave background*, *Phys. Rev. D* **59** (1999) 123507, [astro-ph/9810257].

[4] W. Hu and T. Okamoto, *Mass reconstruction with cmb polarization*, *Astrophys. J.* **574** (2002) 566–574.

[5] M. H. Kesden, A. Cooray, and M. Kamionkowski, *Lensing reconstruction with cmb temperature and polarization*, *Phys. Rev. D* **67** (2003) 123507, [astro-ph/0302536].

[6] D. Hanson, G. Rocha, and K. Gorski, *Lensing reconstruction from planck sky maps: inhomogeneous noise*, *Mon. Not. R. Astron. Soc.* **400** (2009) 2169–2173, [arXiv:0907.1927].

[7] T. Namikawa, D. Hanson, and R. Takahashi, *Bias-hardened cmb lensing*, *Mon. Not. R. Astron. Soc.* **431** (2013) 609–620, [arXiv:1209.0091].

[8] C. M. Hirata and U. Seljak, *Reconstruction of lensing from the cosmic microwave background polarization*, *Phys. Rev. D* **68** (2003) 083002, [astro-ph/0306354].