Lecture 5

1

Outline:

1) Simplified data model
2) bandpowers & $\hat{C}_\ell^{TT}$
3) $\mathrm{var}(\hat{C}_\ell^{TT})$, noise bias, cosmic variance
4) Posterior samples of C_ℓ^{TT}

Start by studying estimation in a simplified case: $T(\hat{n})$ observed everywhere with white noise & a beam

Assume the observed field $d(\hat{n})$ is measured on all $\hat{n} \in S^2$ & has the form

$$(*) \quad d(\hat{n}) = \varphi * T(\hat{n}) + \varepsilon(\hat{n})$$

φ models a beam point spread function.

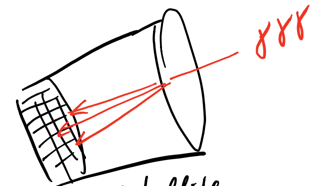
ε: white noise s.t. $E(\varepsilon_{\ell m} \bar{\varepsilon}_{\ell' m'}) = \sigma^2 \delta_{\ell\ell'} \delta_{mm'}$

Note: the fact that $\varepsilon(\hat{n})$ is white noise we expect statistical properties of Model $(*)$ to be similar to the case of finite pixel observations

The beam $\varphi(\hat{n})$

2

As the CMB pass through the detector lense they get defracted and spread out



satellite.

The effect on the observed $T(\hat{n})$ is modeled as a convolution with a point spread function $\varphi(\hat{n})$.

on S^2

- Signal $T(\hat{n}), \hat{n} \in S^2$
- $\varphi * T(\hat{n}) = \int_{g \in SO^3} T(g\,0)\, \varphi(g^{-1}\hat{n})\, dg$ (0: North pole)
- $(\varphi * T)_{\ell m}$
$$= 2\pi \sqrt{\frac{4\pi}{2\ell+1}}\, T_{\ell m}\, \varphi_{\ell,0}$$
- How do we model $\varphi(\hat{n})$ (equiv $\varphi_{\ell m}$) on S^2

on $\mathbb{R}^2$

- Signal $T(x)$, $x \in \mathbb{R}^2$
- $\varphi * T(x) = \int_{\mathbb{R}^2} T(y)\, \varphi(x-y)\, dy$
- $(\varphi * T)_k = \iint_{\mathbb{R}^2 \mathbb{R}^2} e^{-ix\cdot k}\, T(y)\, \varphi(x-y) \frac{dx\,dy}{(2\pi)^{d/2}}$

$z = x-y$

$$\hookrightarrow = \iint e^{-iz\cdot k} e^{-iy\cdot k}\, T(y)\, \varphi(z) \frac{dz\,dy}{(2\pi)^{d/2}}$$
$$= (2\pi)^{d/2}\, T_k\, \varphi_k$$

- $\varphi(x)$ modeled as
$$\frac{1}{2\pi^{d/2}\sigma^n} \exp\left(-\frac{1}{2}\frac{|x|^2}{\sigma^2}\right)$$

$$\therefore \varphi_k = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{\sigma^2}{2}|k|^2\right)$$
$$\cong \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{b^2}{8\log(2)}|k|^2\right)$$

where b = full width half max

$$\therefore (\varphi * T_k) = \exp\left(-\frac{b^2}{8\log 2}|k|^2\right) T_k$$

3

For $x \in \mathbb{R}^2$, $\varphi(x)$ has the property that the solution to

$$\Delta u^t(x) = \frac{d}{dt} u^t(x)$$

is expressed

$$u_k^t = \exp\left(-t\,|k|^2\right) u_k^0$$

$$= (\varphi * u^0)_k, \quad \text{for } t = \frac{b^2}{8\log(2)}.$$

On the sphere Δ is defined and has the property that the solution to

$$\Delta u^t(\hat{n}) = \frac{d}{dt} u^t(\hat{n})$$

is expressed as

doesn't depend on m

$$u_{\ell m}^t = \exp(-t\,\ell(\ell+1))\, u_{\ell m}^0$$

$$= (\varphi * u^0)_{\ell m}$$

i.e. $\varphi_{\ell m} = \frac{1}{2\pi}\sqrt{\frac{2\ell+1}{4\pi}}\exp(-t(\ell+1)\ell)$ is how we model the beam in the simplified case on S^2.

4

Since we assume observations of $(*)$ everywhere on S^2 we can take the spherical transform to obtain:

$$d_{\ell m} = \exp\left(-\frac{b^2}{8\log(2)}\ell(\ell+1)\right) T_{\ell m} + \varepsilon_{\ell m}$$

Dividing out the beam (& renaming the $d_{\ell m}, \varepsilon_{\ell m}$) we can assume w.l.g the data is in the form of an infinite sequence of spherical coefficients:

$$d_{\ell m} = T_{\ell m} + \varepsilon_{\ell m}$$

for $\ell = 0, 1, 2, \dots$, $m = -\ell, -\ell+1, \dots, \ell$ s.t.

$$E(T_{\ell m} T^*_{\ell' m'}) = \delta_{\ell\ell'}\delta_{mm'} C_\ell^{TT}$$

$$E(\varepsilon_{\ell m}\varepsilon^*_{\ell' m'}) = \delta_{\ell\ell'}\delta_{mm'}\,\sigma^2 \exp\left(\frac{b^2}{8\log(2)}\ell(\ell+1)\right)$$

additive noise level

beam.

Notice

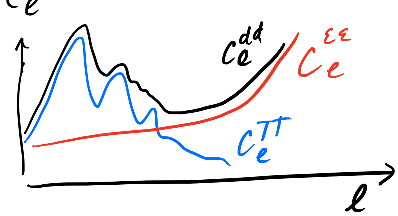
$$C_\ell^{dd} = E|d_{\ell m}|^2$$

$$= E|T_{\ell m}|^2 + \underbrace{E(T_{\ell m}\bar{\varepsilon}_{\ell m})}_{=0} + \underbrace{E(\bar{T}_{\ell m}\varepsilon_{\ell m})}_{=0} + E|\varepsilon_{\ell m}|^2$$

$$= C_\ell^{TT} + \sigma^2 \exp\left(\frac{b^2}{8\log(2)}\ell(\ell+1)\right)$$

$$= C_\ell^{dd}$$

=



Nonparametric estimates of C_ℓ^{TT}

Let $\sigma_\ell = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} |d_{\ell m}|^2 =$ "band power"

Now $E(\sigma_\ell) = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} E|d_{\ell m}|^2$

$$= C_\ell^{dd}$$

$$= C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon}$$

Noise bias... but known and can be subtracted.

Now

$$var(\sigma_\ell) = E(|\sigma_\ell|^2) - (C_\ell^{dd})^2$$
$$= \frac{1}{(2\ell+1)^2}\sum_{m=-\ell}^{\ell}\sum_{m'=-\ell}^{\ell} E(d_{\ell m} d_{\ell m}^* d_{\ell m'} d_{\ell m'}^*) - (C_\ell^{dd})^2$$

where

$$E(d_{\ell m} d_{\ell m}^* d_{\ell m'} d_{\ell m'}^*) = E(d_{\ell m} d_{\ell m}^* d_{\ell m'} d_{\ell m'}^*) + E(d_{\ell m} d_{\ell m}^* d_{\ell m'} d_{\ell m'}^*) + E(d_{\ell m} d_{\ell m}^* d_{\ell m'} d_{\ell m'}^*)$$

Wicks thm by Gaussianity

$$= E|d_{\ell m}|^2 E|d_{\ell m'}|^2 + E(d_{\ell m} d_{\ell m'}) E(d_{\ell m}^* d_{\ell m'}^*) + E(d_{\ell m} d_{\ell m'}^*) E(d_{\ell m}^* d_{\ell m'})$$

$$= (C_\ell^{dd})^2 + (-1)^{m'}\delta_{m,-m'} C_\ell^{dd} (-1)^m \delta_{-m,m'} C_\ell^{dd} + \delta_{mm'} C_\ell^{dd} \delta_{m,m'} C_\ell^{dd}$$

$$\therefore var(\sigma_\ell) = (C_\ell^{dd})^2 + \frac{2}{(2\ell+1)^2}\sum_{m=-\ell}^{\ell}(C_\ell^{dd})^2 - (C_\ell^{dd})^2$$

$$= \frac{2}{2\ell+1}(C_\ell^{dd})^2$$

$$\therefore var(\sigma_\ell) = \frac{2}{2\ell+1}(C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon})^2$$

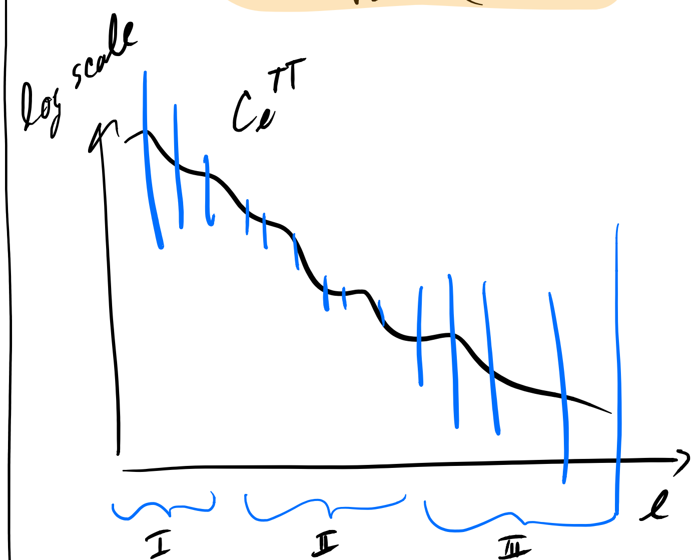
Now a nonparametric estimate of C_ℓ^{TT} can be constructed as:

$$\hat{C}_\ell^{TT} := \sigma_\ell - C_\ell^{\varepsilon\varepsilon}$$

where

$$E(\hat{C}_\ell^{TT}) = C_\ell^{TT} \quad \&$$

$$sd(\hat{C}_\ell^{TT}) = \sqrt{\frac{2}{2\ell+1}}(C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon})$$



Relative s.d.: $\dfrac{\sqrt{\frac{2}{2\ell+1}}(C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon})}{C_\ell^{TT}}$ ← the s.d. of the est of C_ℓ^{TT} ; ← what your estimating

$$= \sqrt{\frac{2}{2\ell+1}}\left(1 + \frac{C_\ell^{\varepsilon\varepsilon}}{C_\ell^{TT}}\right)$$

$$= \begin{cases} \text{big}\left(1+\frac{\text{small}}{\text{big}}\right) & \text{in I} \\ \text{small}(1+1) & \text{in II} \\ \text{small}\left(1+\frac{\text{big}}{\text{small}}\right) & \text{in III} \end{cases}$$

in I ⟶ Cosmic var dominated, beam effect small

in III ⟶ noise dominated by beam

Cosmic variance

Notice that even in the noise free case

$$d(\hat{n}) = T(\hat{n}) \iff d_{\ell m} = T_{\ell m}$$

The estimate $\hat{C}_\ell^{TT} = \hat{\sigma}_\ell$ one still has uncertainty @ each ℓ:

$$sd(\hat{C}_\ell^{TT}) = \sqrt{\frac{2}{2\ell+1}}\, C_\ell^{TT}$$

This is called cosmic variance and comes from the fact we only have one realization of the universe i.e. just one map $T(\hat{n})$.

$\therefore$ if a parameter θ_i only effects C_ℓ^{TT} for $0 \le \ell \le \ell_{max}$ then θ_i can not be consistently estimated from $d(\hat{n})$ as noise $\to 0$.

However since $sd(\hat{C}_\ell^{TT}) \to 0$ as $\ell \to \infty$ (in the noise free case) consistent estimation is possible when observing only one realization of $T(\hat{n})$ on $\hat{n} \in S^2$.

The likelihood $P(\vec{d} \mid C_\ell^{TT}, C_\ell^{\mathcal{E}\mathcal{E}})$

Going back to the Model

$$d_{\ell m} = T_{\ell m} + \mathcal{E}_{\ell m}$$

where $\mathcal{E}_{\ell m}$'s are mean zero complex Gaussian s.t. $\mathcal{E}_{\ell,-m} = (-1)^m \overline{\mathcal{E}_{\ell,m}}$ &

$$E(\mathcal{E}_{\ell m} \overline{\mathcal{E}_{\ell' m'}}) = \delta_{\ell\ell'} \delta_{mm'} \sigma^2 \exp\left(\frac{b^2}{8\log(2)} \ell(\ell+1)\right).$$

Claim: If m is restricted to be ≥ 0 then the $d_{\ell m}$'s are independent &

$$m > 0 \Rightarrow \begin{pmatrix} Re(d_{\ell m}) \\ Im(d_{\ell m}) \end{pmatrix} \overset{(1)}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{C_\ell^{TT}+C_\ell^{\mathcal{E}\mathcal{E}}}{2} & 0 \\ 0 & \frac{C_\ell^{TT}+C_\ell^{\mathcal{E}\mathcal{E}}}{2} \end{pmatrix} \right)$$

$$m = 0 \Rightarrow \begin{pmatrix} Re(d_{\ell m}) \\ Im(d_{\ell m}) \end{pmatrix} \overset{(2)}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_\ell^{TT}+C_\ell^{\mathcal{E}\mathcal{E}} & 0 \\ 0 & 0 \end{pmatrix} \right)$$

Proof: First Note $E|d_{\ell m}|^2 = C_\ell^{TT} + C_\ell^{\mathcal{E}\mathcal{E}}$ implies

$$E(Re(d_{\ell m})^2) + E(Im(d_{\ell m})^2) \overset{(3)}{=} C_\ell^{TT} + C_\ell^{\mathcal{E}\mathcal{E}}.$$

Since $d_{\ell 0}$ is real we immediatly get (2).

Also $E(d_{\ell m} d_{\ell m}) = 0$ implies

$$E(Re(d_{\ell m})^2) - E(Im(d_{\ell m})^2) \overset{(4)}{=} 0$$

$$2E(Re(d_{\ell m}) Im(d_{\ell m})) \overset{(5)}{=} 0$$

$$(z_1 + i z_2)(z_1 + i z_2) = z_1^2 - z_2^2 + i 2 z_1 z_2$$

$\therefore$ (3), (4) & (5) implies (1)

To see why the $d_{\ell m}$'s are indep $\forall m \geq 0$, let $\ell \neq \ell'$ & $m \neq m'$ both ≥ 0.

write $d_{\ell m} = R + iI$ &

$$d_{\ell' m'} = R' + iI'.$$

Now $d_{\ell m} \overline{d_{\ell' m'}} = (R+iI)(R'-iI')$ so that

$$E(RR' + II') + iE(IR' - RI')$$

$$= E(d_{\ell m} \overline{d_{\ell' m'}}) \overset{(6)}{=} 0$$

Also $d_{\ell m} d_{\ell' m'} = (R+iI)(R'+iI')$ so

$$E(RR'-II') + i\,E(IR'+RI') = E(d_{\ell m} d_{\ell' m'})$$

$$= E(d_{\ell m} (-1)^{m} \overline{d_{\ell', -m'}})$$

$$\overset{(7)}{=} 0 \quad \text{since } m \neq -m' \text{ because } m, m' \geq 0$$

Adding (6) & (7) gives

$$E(RR') = 0 \quad \& \quad E(IR') = 0$$

Subtracting (6) & (7) gives

$$E(II') = 0 \quad \& \quad E(RI') = 0$$

$\therefore$ R, R', I, I' are indep. $\blacksquare$

If we stack the $d_{\ell m}$'s into a vector $\vec{d}$ for all $\ell = 0, 1, \ldots, \ell_{max}$ & $m = 0, 1, \ldots, \ell$ we get a Gaussian likelihood:

$$P(\vec{d} \mid \underbrace{C_\ell^{dd}}_{= C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon}}) = \prod_{\ell=0}^{\ell_{max}} \prod_{m=0}^{\ell} P(d_{\ell m} \mid C_\ell^{dd})$$

$$= \prod_{\ell=0}^{\ell_{max}} P(d_{\ell 0} \mid C_\ell^{dd}) \prod_{m=1}^{\ell} P(d_{\ell m} \mid C_\ell^{dd})$$

$$\propto \prod_{\ell=0}^{\ell_{max}} \frac{1}{\sqrt{C_\ell^{dd}}} e^{-\frac{1}{2} \frac{d_{\ell 0}^2}{C_\ell^{dd}}} \prod_{m=1}^{\ell} \frac{\sqrt{2}}{C_\ell^{dd}} e^{-\frac{1}{2} \frac{|d_{\ell m}|^2}{C_\ell^{dd}/2}}$$

$$\propto \prod_{\ell=0}^{\ell_{max}} (C_\ell^{dd})^{-\frac{2\ell+1}{2}} \exp\left(-\frac{1}{2} \frac{|d_{\ell 0}|^2 + 2\sum_{m=1}^{\ell} |d_{\ell m}|^2}{C_\ell^{dd}}\right)$$

$$= \prod_{\ell=0}^{\ell_{max}} (C_\ell^{dd})^{-\frac{2\ell+1}{2}} \exp\left(-\frac{2\ell+1}{2} \frac{\sigma_\ell}{C_\ell^{dd}}\right)$$

So

$$P(\vec{d} \mid C_\ell^{dd}) \propto \prod_{\ell=0}^{\ell_{max}} (C_\ell^{dd})^{-\frac{2\ell+1}{2}} \exp\left(-\frac{2\ell+1}{2} \frac{\sigma_\ell}{C_\ell^{dd}}\right)$$

where $C_\ell^{dd} = C_\ell^{TT} + C_\ell^{\varepsilon\varepsilon}$.

* This also shows that the bandpower σ_ℓ is a sufficient statistic for esitimating C_ℓ^{TT} at any fixed ℓ. (Note: this derivation was done under the assumption that $\varepsilon(\hat{n})$ is isotropic generalized noise.)

Bayesian Posterior $P(C_\ell^{TT} \mid \sigma_\ell)$

Since the likelihood factors over ℓ one can construct Bayesian posteriors for C_ℓ^{TT} given σ_ℓ at each ℓ. In particular one can model the uncertainty in C_ℓ^{TT} with a prior probability distribution $\pi(C_\ell^{TT})$. Following Bayes Rule for swapping the events in a conditional probability calculation we have

$$\underbrace{P}_{\text{denotes density}}(\underbrace{C_\ell^{TT}}_{\text{treating } C_\ell^{TT} \text{ as random}} \mid \sigma_\ell) = \frac{P(C_\ell^{TT}, \sigma_\ell)}{P(\sigma_\ell)}$$

$$= \frac{\overbrace{P(\sigma_\ell \mid C_\ell^{TT})}^{\text{The likelihood}} \overbrace{\pi(C_\ell^{TT})}^{\text{Prior}}}{\underbrace{\int P(\sigma_\ell \mid C_\ell^{TT}) \pi(C_\ell^{TT})\, dC_\ell^{TT}}_{\text{a normalizing constant w.r.t. } C_\ell^{TT}}}$$

"Non-informative"

Jeffreys prior for C_ℓ^{TT} is given by (11)

$$\pi(C_\ell^{TT}) \propto \sqrt{E\left(-\frac{d^2}{dC_\ell^{TT\,2}} \underbrace{\log P(\sigma_\ell|C_\ell^{TT})}\right)}$$

$$-\frac{2\ell+1}{2}\log\left(C_\ell^{TT}+C_\ell^{EE}\right)-\frac{2\ell+1}{2}\frac{\sigma_\ell}{C_\ell^{TT}+C_\ell^{EE}}$$

$$\frac{d}{dC_\ell^{TT}}(\cdots) = -\frac{2\ell+1}{2}\frac{1}{C_\ell^{TT}+C_\ell^{EE}} + (2\ell+1)\frac{\sigma_\ell}{(C_\ell^{TT}+C_\ell^{EE})^2}$$

$$\frac{d^2}{dC_\ell^{TT\,2}}(\cdots) = \frac{(2\ell+1)}{(C_\ell^{TT}+C_\ell^{EE})^2} - 2(2\ell+1)\frac{\sigma_\ell}{(C_\ell^{TT}+C_\ell^{EE})^3}$$

$$\therefore \pi(C_\ell^{TT}) \propto \sqrt{-\frac{(2\ell+1)}{(C_\ell^{TT}+C^{EE})^2} + 2(2\ell+1)\frac{C_\ell^{TT}+C_\ell^{EE}}{(C_\ell^{TT}+C_\ell^{EE})^3}}$$

$$\propto \frac{1}{C_\ell^{TT}+C_\ell^{EE}}$$

Resulting a posterior

$$P(C_\ell^{TT}|\sigma_\ell) \propto \left(C_\ell^{TT}+C_\ell^{EE}\right)^{-\frac{2\ell+1}{2}-1} \exp\left(-\frac{2\ell+1}{2}\frac{\sigma_\ell}{C_\ell^{TT}+C_\ell^{EE}}\right)$$