Maximum likelihood analysis for an ideal survey

November 22, 2010

Consider the type of ideal survey that you would get from an N-body simulation, with constant mean galaxy number density $\bar{n}$ over a periodic volume $V = L^3$. Let $n(\boldsymbol{r})$ be the observed galaxy number density in real space (i.e. a sum of delta functions). Following the maximum likelihood method of Tegmark et al., we collapse this number density into pixelized values

$$x_i = \int_V d^3r \left(\frac{n(\boldsymbol{r})}{\bar{n}} - 1 \right) \psi_i(\boldsymbol{r}) \tag{1}$$

for some set of mode functions $\psi_i(\boldsymbol{r})$. The expected covariance of these pixelized values is

$$\begin{aligned} C_{ij} &= \int_V d^3r_1 \int_V d^3r_2 \; \psi_i^*(\boldsymbol{r}_1)\psi_j(\boldsymbol{r}_2)\xi(\boldsymbol{r}_2 - \boldsymbol{r}_1) + \int_V d^3r \; \frac{\psi_i^*(\boldsymbol{r})\psi_j(\boldsymbol{r})}{\bar{n}} && (2) \\ &= V \sum_{\boldsymbol{k}} \tilde{\psi}_i^*(\boldsymbol{k})\tilde{\psi}_j(\boldsymbol{k}) P(\boldsymbol{k}) + \int_V d^3r \; \frac{\psi_i^*(\boldsymbol{r})\psi_j(\boldsymbol{r})}{\bar{n}} && (3) \\ &\equiv S_{ij} + N_{ij}, && (4) \end{aligned}$$

where the sum is over all wavevectors $\boldsymbol{k} = \frac{2\pi}{L}(n_x, n_y, n_z)$.

For the mode functions $\psi_i(\boldsymbol{r})$ we choose simple plane waves,

$$\psi_i(\boldsymbol{r}) = A_i e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}}, \tag{5}$$

for some set of wavevectors $\{\boldsymbol{k}_i\}$. The Fourier transform of this function is a Kronecker delta ($\tilde{\psi}_i(\boldsymbol{k}) = A_i \delta_{\boldsymbol{k}_i,\boldsymbol{k}}$) so we find

$$C_{ij} = \delta_{ij} V A_i^2 \left(P(\boldsymbol{k}_i) + \frac{1}{\bar{n}} \right). \tag{6}$$

In this case the pixelized values x_i are just the Fourier coefficients of the density contrast $\delta(\boldsymbol{r}) = n(\boldsymbol{r})/\bar{n} - 1$,

$$x_i = V A_i \tilde{\delta}_{\boldsymbol{k}_i}. \tag{7}$$

We take our model power spectrum to be a piecewise-constant function

$$P(k) = \sum_m p_m \theta(K_m \leq k < K_{m+1}), \tag{8}$$

where p_m represents the average band power over the interval $[K_m, K_{m+1}]$. We then define the quadratic estimator

$$\hat{p}_m = \mathsf{x}^T \mathsf{Q}_m \mathsf{x} - \mathrm{Tr}[\mathsf{Q}_m \mathsf{N}] \tag{9}$$

where x is a column vector with components x_i, N is the noise matrix with components $N_{ij} = \delta_{ij} V A_i^2 \frac{1}{\bar{n}}$, and Q_m is given by

$$\mathsf{Q}_m = \frac{1}{2} \sum_n M_{mn} \mathsf{C}^{-1} \mathsf{C}_{,n} \mathsf{C}^{-1} \tag{10}$$

In terms of components we have

$$C_{m,ij} = \delta_{ij} V A_i^2 \theta(K_n \leq |\boldsymbol{k}_i| < K_{n+1}). \tag{11}$$

The matrix M_{mn} is arbitrary, except that it should be normalized so that $\hat{p}_m$ gives a sensible estimate of the actual band power p_m. This translates to the requirement that the rows of the matrix MF sum to 1, i.e.

$$\sum_n (MF)_{mn} = 1, \tag{12}$$

where F_{mn} is the Fisher matrix,

$$F_{mn} = \frac{1}{2} \mathrm{Tr}[\mathsf{C}^{-1} \mathsf{C}_{,m} \mathsf{C}^{-1} \mathsf{C}_{,n}] \tag{13}$$

$$= \frac{1}{2} \mathrm{Tr} \left[\delta_{ij} \theta(K_m \leq |\boldsymbol{k}_i| < K_{m+1}) \theta(K_n \leq |\boldsymbol{k}_i| < K_{n+1}) \left(P(\boldsymbol{k}_i) + 1/\bar{n} \right)^{-2} \right] \tag{14}$$

$$= \delta_{mn} \, \frac{\mathcal{N}_m}{2} \left(p_m + 1/\bar{n} \right)^{-2}, \tag{15}$$

and $\mathcal{N}_m$ is the number of wavevectors $\boldsymbol{k}_i$ lying in the interval $K_m \leq |\boldsymbol{k}_i| < K_{m+1}$. Since F_{mn} is diagonal, we can (but don't necessarily have to) take M_{mn} to be diagonal as well, in which case the normalization requirement forces

$$M_{mn} = (F^{-1})_{mn} = \delta_{mn} \, \frac{2}{\mathcal{N}_m} \left(p_m + 1/\bar{n} \right)^2. \tag{16}$$

Putting everything together we find

$$
\begin{aligned}
\hat{p}_m &= \sum_{ij} VA_i \tilde{\delta}^*_{\boldsymbol{k}_i} \cdot \delta_{ij} \frac{1}{VA_i^2} \frac{1}{\mathcal{N}_m} \theta(K_m \leq |\boldsymbol{k}_i| < K_{m+1}) \cdot VA_j \tilde{\delta}_{\boldsymbol{k}_j} - \frac{1}{\bar{n}} && (17) \\
&= V \, \frac{1}{\mathcal{N}_m} \sum_{\boldsymbol{k} \in S_m} |\tilde{\delta}_{\boldsymbol{k}}|^2 - \frac{1}{\bar{n}}, && (18) \\
\langle \hat{p}_m \rangle &= p_m, && (19) \\
\mathrm{Cov}[\hat{p}_m, \hat{p}_n] &= 2\, \mathrm{Tr}[\mathsf{Q}_m C \mathsf{Q}_n C] && (20) \\
&= \delta_{mn} \, \frac{2}{\mathcal{N}_m} (p_n + 1/\bar{n})^2 && (21)
\end{aligned}
$$

where $S_m = \{\boldsymbol{k}_i : K_m \leq |\boldsymbol{k}_i| < K_{m+1}\}$ includes all wavevectors lying in the mth shell. This is exactly the estimator and error properties we expect for such an ideal survey.