Power Spectrum Analysis of Large Baseline Redshift Surveys♣

Hume A. Feldman

Physics Department
University of Michigan
Ann Arbor, MI 48109

Given a set of redshifts of galaxies with known angular and luminousity criteria, our goal is to construct a descriptive statistic that measures the power spectrum of the underlying density fluctuation field. We assume that the fluctuation field is some homogeneous and statistically isotropic random process[1]. The present study follows the program laid down by Peebles in his pioneering series of papers[2] for statistical analysis of galaxy catalogues via low order correlation functions.

We make the usual assumption that the galaxies form a poisson sample[3] of the density field $1 + f(\mathbf{r}) = \rho(\mathbf{r})/\bar{\rho}$:

$$P(\text{vol element } \delta V \text{ contains a galaxy}) = \delta V \bar{n}(\mathbf{r})(1 + f(\mathbf{r})) \tag{1}$$

where $\bar{n}(\mathbf{r})$ is the expected mean space density of galaxies given the angular and luminosity selection criteria, and we wish to estimate the power spectrum

$$P(k) = P(\mathbf{k}) \equiv \int d^3r \xi(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2}$$

where $\xi(\mathbf{r}) = \xi(r) = \langle f(\mathbf{r}')f(\mathbf{r}' + \mathbf{r})\rangle$ is the 2–point correlation function.

Our approach is to take the fourier transform of the real galaxies minus the transform of a synthetic catalogue with the same angular and radial selection function as the real galaxies but otherwise without structure. We also incorporate a weight function $w(\mathbf{r})$ which will be adjusted to optimize the performance. We define the weighted galaxy fluctuation field, to be

$$F(\mathbf{r}) \equiv \frac{w(\mathbf{r})(n_g(\mathbf{r}) - \alpha n_s(\mathbf{r}))}{\left(\int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r})\right)^{1/2}} \tag{3}$$

where $n_g(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ with $\mathbf{r}_i$ being the location of the ith galaxy and similarly for the synthetic catalogue which has number density $1/\alpha$ times that of the real catalogue.

Taking the fourier transform of $F(\mathbf{r})$, squaring it and taking the expectation value we find:

$$\langle |F(\mathbf{k})|^2 \rangle = \frac{\int d^3r \int d^3r' w(\mathbf{r}) w(\mathbf{r}') \langle [n_g(\mathbf{r}) - \alpha n_s(\mathbf{r})][n_g(\mathbf{r}') - \alpha n_s(\mathbf{r}')] \rangle e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}}{\int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r})} \tag{4}$$

♣ This work was supported in part by the National Science Foundation Grant NSF-PHY-92-96020

With the model of equation 1, the two point functions of n_g, n_s are

$$\begin{aligned}\langle n_g(\mathbf{r})n_g(\mathbf{r}')\rangle &= \overline{n}(\mathbf{r})\overline{n}(\mathbf{r}')(1+\xi(\mathbf{r}-\mathbf{r}')) + \overline{n}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \\ \langle n_s(\mathbf{r})n_s(\mathbf{r}')\rangle &= \alpha^{-2}\overline{n}(\mathbf{r})\overline{n}(\mathbf{r}') + \alpha^{-1}\overline{n}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \\ \langle n_g(\mathbf{r})n_s(\mathbf{r}')\rangle &= \alpha^{-1}\overline{n}(\mathbf{r})\overline{n}(\mathbf{r}')\end{aligned} \tag{5}$$

so

$$\langle |F(\mathbf{k})|^2\rangle = \int \frac{d^3k'}{(2\pi)^3} P(\mathbf{k}')|G(\mathbf{k}-\mathbf{k}')|^2 + \mathrm{P_{shot}} \tag{6}$$

where

$$G(\mathbf{k}) \equiv \frac{\int d^3r\overline{n}(\mathbf{r})w(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}}{\left(\int d^3r\overline{n}^2(\mathbf{r})w^2(\mathbf{r})\right)^{1/2}} \tag{7}$$

and

$$\mathrm{P_{shot}} \equiv (1+\alpha)\frac{\int d^3r\overline{n}(\mathbf{r})w^2(\mathbf{r})}{\int d^3r\overline{n}^2(\mathbf{r})w^2(\mathbf{r})}\;. \tag{8}$$

For a large baseline survey $G(\mathbf{k})$ is a rather compact function with width $\sim 1/D$, where D characterizes the depth of the survey. Provided we restrict attention to $|\mathbf{k}| \gg 1/D$, which is really just the requirement that we have a 'fair sample', and provided $P(\mathbf{k})$ is locally smooth on the same scale, then

$$\langle |F(\mathbf{k})|^2\rangle \simeq P(\mathbf{k}) + P_{\mathrm{shot}}, \tag{9}$$

so the raw power spectrum $|F(\mathbf{k})|^2$ is the true power spectrum plus the constant shot noise component and our estimator is

$$\hat{P}(\mathbf{k}) = |F(\mathbf{k})|^2 - P_{\mathrm{shot}}, \tag{10}$$

finally we average over a shell in k-space:

$$\hat{P}(k) \equiv \frac{1}{V_k}\int_{V_k} d^3k'\hat{P}(\mathbf{k}'), \tag{11}$$

where V_k is the volume of the shell.

Equations (3), (9-11) provide our operational definition of $\hat{P}(k)$. To use these we must specify the weight function $w(\mathbf{r})$ which so far has been arbitrary, and we must choose some sampling grid in k-space. In order to set these wisely — and also to put error bars on our estimate of the power — we need to understand the statistical fluctuations in $\hat{P}(\mathbf{k})$.

From equation (11) the mean square fluctuation in $\hat{P}(k)$ is

$$\sigma_P^2 \equiv \langle\left(\hat{P}(k) - P(k)\right)^2\rangle = \frac{1}{V_k^2}\int_{V_k} d^3k \int_{V_k} d^3k' \langle\delta\hat{P}(\mathbf{k})\delta\hat{P}(\mathbf{k}')\rangle. \tag{12}$$

An interesting model for the two point function of $\delta\hat{P}(\mathbf{k})$ is to assume that the coefficients $F(\mathbf{k})$ are gaussian distributed, in which case $\langle\delta\hat{P}(\mathbf{k})\delta\hat{P}(\mathbf{k}')\rangle = |\langle F(\mathbf{k})F^*(\mathbf{k}')\rangle|^2$. A generalisation of the steps leading to (6) gives us

$$\langle F(\mathbf{k})F^*(\mathbf{k}')\rangle = \int \frac{d^3k''}{(2\pi)^3} P(\mathbf{k}'')G(\mathbf{k}-\mathbf{k}'')G^*(\mathbf{k}'-\mathbf{k}'') + S(\mathbf{k}'-\mathbf{k}), \quad (13)$$

where we have defined

$$S(\mathbf{k}) \equiv \frac{(1+\alpha)\int d^3r\bar{n}(\mathbf{r})w^2(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}}{\int d^3r\bar{n}^2(\mathbf{r})w^2(\mathbf{r}),} \quad (14)$$

and, in the same approximation that led to equation (9) we obtain

$$\langle F(\mathbf{k})F^*(\mathbf{k}+\delta\mathbf{k})\rangle \simeq P(\mathbf{k})Q(\delta\mathbf{k}) + S(\delta\mathbf{k}) \quad (15)$$

where

$$Q(\mathbf{k}) \equiv \frac{\int d^3r\bar{n}^2(\mathbf{r})w^2(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}}{\int d^3r\bar{n}^2(\mathbf{r})w^2(\mathbf{r})}, \quad (16)$$

and therefore

$$\langle\delta\hat{P}(\mathbf{k})\delta\hat{P}(\mathbf{k}')\rangle = |P(\mathbf{k})Q(\delta\mathbf{k}) + S(\delta\mathbf{k})|^2. \quad (17)$$

If the shell we average over in equation (11) has a width which is large compared to the coherence length then the double integral in (12) reduces to

$$\sigma_P^2(k) \simeq \frac{1}{V_k}\int d^3k'|P(k)Q(\mathbf{k}') + S(\mathbf{k}')|^2, \quad (18)$$

so, with the definition of $Q(\mathbf{k})$ and $S(\mathbf{k})$ and using Parseval's theorem, the fractional variance in the power is

$$\sigma_P^2(k)/P(k)^2 = (2\pi)^3\int d^3r\bar{n}^4w^4(1+1/\bar{n}P(k))^2/V_k\left[\int d^3r\bar{n}^2w^2\right]^2. \quad (19)$$

We seek $w(\mathbf{r})$ which minimises this. Writing $w(\mathbf{r}) = w_0(\mathbf{r}) + \delta w(\mathbf{r})$ and requiring that $\sigma_P^2(k)$ be stationary with respect to arbitrary variations $\delta w(\mathbf{r})$ we obtain

$$\frac{\int d^3r\bar{n}^4w_0^3\left(\frac{1+\bar{n}P}{\bar{n}P}\right)^2\delta w(\mathbf{r})}{\int d^3r\bar{n}^4w_0^4\left(\frac{1+\bar{n}P}{\bar{n}P}\right)^2} = \frac{\int d^3r\bar{n}^2w_0\delta w(\mathbf{r})}{\int d^3r\bar{n}^2w_0^2} \quad (20)$$

and it is easy to see by direct substitution that this is satisfied if we take

$$w_0(\mathbf{r}) = [1+\bar{n}(\mathbf{r})P(k)]^{-1}\ . \quad (21)$$

This is the optimal weighting (under the assumption that the fluctuations are gaussian).

References

[1] Feldman, H.A., Kaiser, N. & Peacock, J., 1993 in preparation.

[2] Peebles, P.J.E. & Hauser, G.M., 1974 *ApJ Suppl.* **28** 19; Yu, J.T. & Peebles, P.J.E., 1969 *ApJ* **158** 103.

[3] Peebles, P.J.E. 1980, *The Large Scale Structure of the Universe*, Princeton U. Press.