Astronomical Data Analysis Software and Systems III
ASP Conference Series, Vol. 61, 1994
D. R. Crabtree, R. J. Hanisch, and J. Barnes, eds.

An Alternative Algorithm for CMBR Full Sky Harmonic Analysis

C. A. Wuensche[1] and P. Lubin

Department of Physics, University of California, Santa Barbara CA, 93106 USA

T. Villela

Divisão de Astrofísica, INPE São José dos Campos, SP, 12201 Brazil

Abstract. We discuss an alternative algorithm to deal with many-point data sets and show that, for the kind of problem we want to solve, it is more robust, although a little slower, than the currently used algorithm and can always present a numerical solution, based on the maximum likelihood test. We describe the algorithm and the kind of data it is suitable for. The basic idea is to avoid the inversion of large matrices, usually the core of the algorithms currently chosen, by using the Singular Value Decomposition method. Results of Monte Carlo simulations of full sky maps and a comparison of both methods are discussed.

1. Introduction

The recent advances in space astronomy and astrophysics have created a demand for more powerful computers and more "robust" data analysis softwares. As an example, astronomical observations which map large portions of the sky, and eventually the whole celestial sphere, will become natural, following the development of the so-called satellite astronomy, complementing ground-based observations. These missions generate very large data sets that demand large amounts of memory to be handled, because the finer the information per sky patch, the larger the storage matrices required by the data reduction/analysis software.

Satellites have been used in various ranges of the electromagnetic spectrum, covering from microwaves to γ-rays, with great success. For instance, one of the most exciting questions in cosmology is concerned with the angular distribution of the Cosmic Microwave Background Radiation (hereafter CMBR). By studying the CMBR angular distribution and its properties, one can learn about the density fluctuations that may have generated the large scale structures seen today. Since the discovery of the CMBR in 1965 (Penzias and Wilson, 1965), a large number of experiments have been contributing to the studies of CMBR spectrum and angular distribution. Particularly, the COBE satellite FIRAS,

[1] On leave of absence from Instituto Nacional de Pesquisas Espaciais — INPE São José dos Campos, SP, 12201 Brazil

DIRBE and DMR's experiments have mapped the entire celestial sphere, studying the spectrum and the angular distribution of CMBR (Mather et al. 1990, Smoot et al. 1992). The software developed to analyze the data from the DMR's has to deal with very large data sets, when binning the raw data and also doing a least square fit to a spherical harmonics series (Torres et al. 1989, Jansky and Gulkis 1990). We present here an alternative algorithm using the Singular Value Decomposition method, hereafter referred to as SVD, to deal with these many-point data sets, show that it is more robust than the currently used algorithm (known as the Normal Equations method, hereafter NE), and can always present a numerical solution, based on the standard maximum likelihood test.

2. The Algorithm

One way to do CMBR anisotropy experiments is to collect data that are difference temperatures of two sky regions, as the DMR's do. In order to convert the measurements of temperature differences in a map of the sky it is necessary for some manipulation of the data. The general idea is to find a temperature distribution T_i which minimizes the χ^2 of the map compared to the actual difference data. One can use, for instance, a linear least squares algorithm used to fit the data set $(pixel_i, T_i)$ to a spherical harmonic series. The data set we deal with is a set of sky temperatures in 6144 pixels distributed over the 6 faces of a cube — the quadrilateralized sphere (Chan and O'Neil 1986), and these pixels are easily converted to the true spherical angle pair (θ,ϕ) (or right ascension and declination; galactic latitude and longitude, for example). The figure of merit defined to analyse the goodness-of-fit is the usual χ^2 test:

$$\chi^2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left[T_i - \sum_{k=1}^{M} a_k Y_k \right]^2 , \qquad (1)$$

where T_i is the sky temperature at a pixels (obtained as a combination of various DMR measurements around a given position in the sky), $\sum_{k=1}^{M} a_k Y_k$ is the general spherical harmonics expansion and the pixel k is associated to θ and ϕ, the true spherical angles. What one usually sees in the literature is the use of NE when trying to fit a temperature distribution into its eigenmodes (Lubin and Villela 1986, Smoot et al. 1991). However, the reason that led us to search for an alternative algorithm was the instability in the matrix inversion behavior in our least square fits. Specially after removing part of the celestial sphere, the basis functions become non-orthogonal and there is a mode mixing among different multipole components. Nevertheless, SVD handles this problem quite elegantly, while keeping the χ^2 low. We also discovered that, for low-order fits, both algorithms run at about the same speed. Actually, for $l \leq 5$, SVD is faster than NE (for l=2, $t_{SVD} = 0.089 min$ and $t_N E = 0.178 min$). Both algorithms can run fits up to order l=18 (limited only by the machine's memory), which, naively, represents the division of the celestial sphere into 10° patches. Following Press et al. (1992) and Strang (1982), we define the weighted basis functions and the weighted data points as

$$A_{ij} = \frac{Y_j(x_i)}{\sigma_i}; b_i = \frac{y_i}{\sigma_i}. \tag{2}$$

(with A generally $M \times N, M \geq N$). The basis functions chosen are the spherical harmonics and, among various possible ways to generate the P_{lm}, we chose one described in Abramowitz and Stegun (1972), and coded in Press et al. (1992). The chosen recurrence relation is:

$$(l-m)P_l^m = x\,(2l-1)\,P_{l-1}^m - (l+m-1)\,P_{l-2}^m, \tag{3}$$

for $l \geq m$, $m \geq 0$ and $-1 \leq x \leq 1$ ($x \equiv cos\theta$, as usual). Please see Press et al. (1992) for a detailed discussion of the stability test for recurrence relations.

The difference between NE and SVD comes mainly from the singularity handling branch in the algorithm. Both algorithms solve a set of simultaneous linear equations that map the data vector b into the solution space, through the A matrix performing the linear transformation from one space to the other. However, if A is singular, the vector x will be partially mapped onto b, and the "misbehaved" part of it will be mapped in a *null* subspace. SVD constructs orthonormal basis for both the *null* space and the mapping space, and the set of solutions we obtain tells us if b lies in the range of A or not. If it does, then the problem will have a solution coming from the A mapping, plus a linear combination of vectors in the *null* space (more than one solution). If we want to single out a specific one, the best approach is to throw away the linear combination of vectors in the *null* space. Following this approach we minimize the residuals of the fitting $r = |A \cdot x - b|$, and in throwing away the solutions included in the *null* space, we discard either round-off errors or non-suitable sets of basis functions.

3. Monte Carlo Simulations

We have done a number of Monte Carlo simulations to verify the stability of the algorithm. We simulate the sky as it is seen by the DMR experiment, including higher order multipole terms, run the program and check the χ^2 of each fit. The residuals from the multipole recovery are very small ($\simeq 1.5\%$ in the worst case), and the χ^2/DOF $\simeq 1$, for a sky fitted to standard amplitude spherical harmonics (as defined in Jackson 1975), using different noise levels. A noise level of 1 means the noise has roughly the same amplitude as the dipole and quadrupole. We also observed the behavior of both algorithms when fitting a simulated DMR map to quadrupole and higher order multipoles. The behavior of both algorithms for different l's was tested both with uniform ($\sigma = 1$) and non-uniform, simulated sky coverage. No galaxy cuts were performed on these tests, and we found that the χ^2/DOF for both lies between 0.9 and 1.1. We also studied the stability with the galactic cut, and found that the SVD fit is reasonably more stable to the removal of portions of the celestial sphere, as opposed to NE. For a 30° cut out off the equator plane, the NE χ^2/DOF is 3.6, more than 3 times the SVD value (1.16). On the other hand, considered as a major disadvantage, the time spent in high order fittings ($l \geq 10$) can be as much as 5 times longer for SVD than for NE. A more detailed study of the method, including the covariance matrix and mode mixing analysis will be published elsewhere (Wuensche et al. 1994).

4. Conclusions

One of the interesting features of SVD is the automatic computation of an orthonormal set of basis functions. Trying to do a straight Gram-Schmidt orthonormalization procedure is a slow and error-prone process, as pointed out by Wright (1993). Using the SVD method, it is possible to decompose the input matrix A into three "orthogonal" matrixes $U(NxM, N \geq M$, returns the desired orthogonal functions), $W(MxM$, diagonal, contains the singular values), and V^T (MxM, used to compute the covariance matrix and the fitting coefficients), and the new orthonormalized functions are generated automatically during the execution of the SVD process.

SVD was proven to be efficient and robust, and is a good alternative to the generally used NE. To compensating the time disadvantage, it offers the stability of the algorithm, and the automatic computation of a new set of orthonormalized basis functions (analogous to a Gram-Schmidt method) in a single pass. Further studies are necessary to better understand the mode mixing problem, and will represent a major step towards doing harmonic analysis on sky maps.

References

Abramowitz & Stegun. 1972, "Handbook of Mathematical Funtions"

Chan, F.K. & O'Neill, E.M. 1976, EPRF Technical Report 2-75 (CSC)

Gulkis, S. & Janssen, M. 1981 NASA COBE report, 4004

Jackson, J. D. 1975, "Classical Electrodynamics", 2nd. Edition, John Wiley & Sons

Janssen, M. & Gulkis, S. 1992, in Infrared and Submillimetre Sky after COBE, M. Signore & C. Dupraz, Dordrecht: Boston, 1992

Lubin, P. & Villela, T. 1986, in Galaxy Distances and Deviations from Universal Expansion, B. Tully, New York: Plenum, 169

Mather et al., 1990, ApJ, 354, L37

Penzias A. & Wilson, R. 1965 ApJ, 771

Press et al., 1992, "Numerical Recipes - The Art of Scientific Computing", 2nd. Edition, Cambridge University Press

Smoot et al., 1991 ApJ, 371, L1

Smoot et al., 1992 ApJ, 396, L1

Strang, G. 1980, "Linear Algebra and its Applications", Academic Press

Torres, S. et al., 1989, in Data Analysis in Astronomy III, V. Di Gesu et al., New York: Plenum, 319

Wright, E. 1993 preprint

Wuensche et al., 1994, Experimental Astronomy, submitted