Internal Linear Combination Method for Foreground Subtraction

Justin Lazear

Jan 7, 2014

1 Internal Linear Combination method of foreground subtraction

Consider a measurement of the full CMB sky in k different frequency bands. The product of such a measurement is k temperature maps

$$T_i(p) = \text{Temperature map of frequency band } \nu_i \text{ with pixel index } p \tag{1}$$

where i is the frequency band index with $i = 1, \dots, k$ and p is the pixel index with $p = 1, \dots, N$. We will choose to represent the maps in thermodynamic temperature units[1].

Such maps are usually encoded in the HEALpix[4] format. Care must be taken to ensure that all maps that are used have the same number of pixels (typically $N_{\text{side}} = 512$ for WMAP) and that all maps have been smoothed to the same resolution (typically 1°).

The temperature maps can (theoretically) be decomposed into a CMB component and a residual component

$$T_i(p) = T_{\text{CMB}}(p) + R_i(p), \tag{4}$$

where the $T_{\text{CMB}}(p)$ component does not dependent on frequency (since we have chosen thermodynamic temperature units), and the $R_i(p)$ component encodes *all* sources of signal that are not from the CMB.

[1] Thermodynamic temperature units are defined relative to the Planck distribution, and are a measure of power per unit area. The key point to keep in mind that the thermodynamic temperature units are referencing a blackbody (Planck) spectrum, so any additional information required to get to the desired units must be pulled from the properties of the map (e.g. area).

Sky maps are usually given in terms of spectral intensity, i.e. energy per unit area per unit solid angle per unit frequency, with SI unit $\frac{\text{W}}{\text{m}^2\,\text{sr}\,\text{Hz}}$. Commonly, intensity is given in terms of mega-Janskies per steradian, where $1\,\text{Jy} = 10^{-26}\,\frac{\text{W}}{\text{m}^2\,\text{Hz}}$, so $1\,\frac{\text{w}}{\text{m}^2\,\text{sr}\,\text{Hz}} = 10^{20}\,\frac{\text{MJy}}{\text{sr}}$.

Spectral intensity I_ν is related to thermodynamic temperature units T by the Planck distribution,

$$I_\nu = \frac{2h\nu^3}{c^2}\frac{1}{\exp\left(h\nu/k_BT\right) - 1} = B_\nu(T). \tag{2}$$

From inspection, we can see that T has units of Kelvin, but encodes the same information as spectral intensity.

The last common unit is antenna temperature T_A, which is the low-frequency limit of the above relation,

$$I_\nu = \frac{2k_B\nu^2}{c^2}T_A, \tag{3}$$

where we note that all we have done is expand the inverse exponential term in the $h\nu/k_BT \ll 1$ limit. The antenna temperature also has units of Kelvin.

We note that the conversions between spectral intensity and the temperatures are frequency-dependent, and do not simply have a numerical value. Additionally, since the CMB spectrum is a blackbody (ignoring spectral distortions), its thermodynamic temperature is constant for all frequencies.

Since the CMB component will be constant in frequency, we expect that we can combine the various maps in different frequencies to construct an estimator of the CMB map $T_{\mathrm{CMB}}(p)$. We construct an estimator using a linear combination of the maps in different frequency bands with to-be-determined weights $w_i(p)$[1, 3, 5],

$$\begin{aligned}\hat{T}(p) &= \sum_{i=1}^{k} w_i(p) T_i(p) &(5)\\ &= \sum_i w_i(p) \left[T_{\mathrm{CMB}}(p) + R_i(p)\right]\\ \hat{T}(p) &= T_{\mathrm{CMB}}(p) \sum_i w_i(p) + \sum_i w_i(p) R_i(p). &(6)\end{aligned}$$

We note that there are $\mathrm{num}\,(w_i(p)) = Nk$ unknown parameters in our estimator. In order for the estimator $\hat{T}(p)$ to have unity gain in $T_{\mathrm{CMB}}(p)$, we must have

$$\sum_i w_i(p) = 1 \qquad \forall\ p \tag{7}$$

which provides N constraint equations. This results in

$$\hat{T}(p) = T_{\mathrm{CMB}}(p) + \sum_{i=1}^{k} w_i(p) R_i(p). \tag{8}$$

Further properties of the estimator $\hat{T}(p)$ will be determined by how we choose to constrain the remaining $(N-1)k$ degrees of freedom.

Let us now introduce a matrix notation for this system.

Define vectors as column vectors of the frequency bands, so

$$\mathbf{T}^T = (T_1(p), \ldots, T_k(p)) \quad \text{and} \quad \mathbf{w}^T = (w_1(p), \ldots, w_k(p))\,,$$

then we may write our estimator as

$$\hat{T} = \mathbf{w}^T \mathbf{T} \tag{9}$$

and the unity gain constraint as

$$\mathbf{1}^T \mathbf{w} = 1 \tag{10}$$

2 Constant weighting factors

Suppose the weight factors $\mathbf{w}(p)$ were uniform across the entire map, so $\mathbf{w}(p) = \mathbf{w}$. We note in this case that our estimator $\hat{T}$ is a linear function with regards to $\mathbf{T}$, so then the variance of $\hat{T}$ has a particularly simple form

$$\mathrm{var}\,\hat{T} = \mathbf{w}^T \,\mathrm{cov}\,(\mathbf{T}, \mathbf{T}) \mathbf{w} \tag{11}$$

where $(\mathrm{cov}\,(\mathbf{T},\mathbf{T}))_{ij} = \mathrm{cov}\,(T_i, T_j)$[2] is the covariance matrix of $\mathbf{T}$.

We choose to optimize $\mathbf{w}$ so that $\mathrm{var}\,\hat{T}$ is minimized, i.e.

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \left(\mathrm{var}\,\hat{T}(\mathbf{w};\mathbf{T})\right) \tag{12}$$

We will do this in two spaces: the raw temperature map space $\{\mathbf{T}\}$, and the foreground-background space $\{T_{\mathrm{CMB}}, \mathbf{R}\}$. The first solution will be used to actually perform the foreground cleaning, while the second will give us a sense of what the algorithm is doing. Let us proceed with the first solution.

2.1 Raw temperature map space

Let $\mathbf{C} = \mathrm{cov}\,(\mathbf{T},\mathbf{T})$, i.e. $\mathbf{C}_{ij} = \mathrm{cov}\,(T_i, T_j)$. Then we may write our minimization problem as

$$\mathrm{var}\,\hat{T} = \mathbf{w}^T\mathbf{C}\mathbf{w} \quad \text{subject to} \quad \mathbf{1}^T\mathbf{w} = 1 \tag{13}$$

which may be solved with Lagrange multipliers. Define the function

$$L = \mathrm{var}\,\hat{T} + \lambda\left(\mathbf{1}^T\mathbf{w} - 1\right). \tag{14}$$

Then the solution to our optimization problem is given by the set of equations

$$\frac{\partial L}{\partial \mathbf{w}^T} = 2\mathbf{C}\mathbf{w} + \lambda\mathbf{1} = 0 \tag{15}$$

$$\mathbf{1}^T\mathbf{w} = 1 \tag{16}$$

Solving Eq. (2.2) for $\mathbf{w}$

$$\mathbf{w} = -\frac{\lambda}{2}\mathbf{C}^{-1}\mathbf{1}$$

and plugging it into Eq. (2.2) allows us to solve for λ,

$$\mathbf{1}^T\mathbf{w} = -\frac{\lambda}{2}\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1} = 1$$

$$\lambda = -2\left(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}\right)^{-1} \tag{17}$$

[2]Expectation values are defined over the pixels, so

$$\langle T_i \rangle = \sum_{p=1}^{N} T_i(p)$$

then

$$\mathrm{cov}\,(T_i, T_j) = \langle T_i T_j \rangle - \langle T_i \rangle \langle T_j \rangle\,.$$

It is worth noting that $\mathrm{cov}\,(\mathbf{T},\mathbf{T}) = \mathrm{cov}\,(\mathbf{T},\mathbf{T})^T$ and $\mathrm{cov}\,(\mathbf{T},\mathbf{T}) \succeq 0$, i.e. the covariance matrix is symmetric and positive semi-definite.

and then we may substitute into $\mathbf{w}$,

$$\mathbf{w}^* = \frac{\mathbf{C}^{-1}\mathbf{1}}{(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1})} \tag{18}$$

We note that

$$w_i^* = \frac{\sum_j \mathrm{C}_{ij}^{-1}}{\sum_{jk} \mathrm{C}_{jk}^{-1}} \tag{19}$$

which matches Eriksen et al[3].

$\mathbf{C}$ is a $k \times k$ matrix, where k is the number of frequency bands, which is typically order unity. Thus, computing $\mathbf{w}^*$ from $\mathbf{C}$ requires the inversion of a small matrix and the sum of its components and may be done cheaply. Far more expensive is the computation of each element of the matrix $\mathrm{C}_{ij} = \sum_{p=1}^{N} T_i(p)T_j(p) - \sum_{p=1}^{N} T_i(p) \sum_{p=1}^{N} T_j(p)$, which is $\mathcal{O}(4N)$. Given that matrix inversion of $\mathbf{C}$ is pessimistically $\mathcal{O}(k^3)$ and $N = 12N_{\text{side}}^2$, our total complexity is $\mathcal{O}(48N_{\text{side}}^2 k^3)$.

2.2 Foreground-background space

Let us now perform the same optimization supposing we already know the decomposition between residual (foreground) and CMB (background). The optimization procedure will produce the same set of equations to solve, Eqs. (2.2) and (2.2).

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}^T} &= 2\mathbf{C}\mathbf{w} + \lambda\mathbf{1} = 0 \\ \mathbf{1}^T\mathbf{w} &= 1 \end{aligned}$$

We note that the covariance matrix may be decomposed into foreground and background components, since $\mathbf{T} = T_{\text{CMB}}\mathbf{1} + \mathbf{R}$,

$$\begin{aligned} \mathbf{C} &= \operatorname{cov}(T_{\text{CMB}}\mathbf{1} + \mathbf{R}, T_{\text{CMB}}\mathbf{1} + \mathbf{R}) \\ &= \operatorname{var}(T_{\text{CMB}})\mathbf{1}\mathbf{1}^T + \mathbf{cov}(T_{\text{CMB}}, \mathbf{R})\mathbf{1}^T + \mathbf{1}\,\mathbf{cov}(T_{\text{CMB}}, \mathbf{R})^T + \operatorname{cov}(\mathbf{R}, \mathbf{R}) \\ \mathbf{C} &= \sigma_B^2\mathbf{1}\mathbf{1}^T + \mathbf{X}\mathbf{1}^T + \mathbf{1}\mathbf{X}^T + \mathbf{F} \end{aligned} \tag{20}$$

where we've defined $\sigma_B^2 \equiv \operatorname{var}(T_{\text{CMB}})$, $\mathbf{X} \equiv \mathbf{cov}(T_{\text{CMB}}, \mathbf{R})$[3], and $\mathbf{F} = \operatorname{cov}(\mathbf{R}, \mathbf{R})$. So our system of equations becomes

$$\begin{aligned} 2\sigma_B^2\mathbf{1} + 2\mathbf{X} + 2\mathbf{1}\mathbf{X}^T\mathbf{w} + 2\mathbf{F}\mathbf{w} + \lambda\mathbf{1} &= 0 \\ \mathbf{1}^T\mathbf{w} &= 1 \end{aligned} \tag{21}$$

[3]Explicitly, $X_i = \operatorname{cov}(T_{\text{CMB}}, R_i)$.

which may be solved in an identical manner,

$$
\begin{aligned}
\mathbf{w} &= -\left(\frac{\lambda}{2} + \sigma_B^2\right)(\mathbf{F} + \mathbf{1}\mathbf{X}^T)^{-1}\mathbf{1} - (\mathbf{F} + \mathbf{1}\mathbf{X}^T)^{-1}\mathbf{X} \\
\mathbf{w} &= -\left(\frac{\lambda}{2} + \sigma_B^2\right)\mathbf{G}^{-1}\mathbf{1} - \mathbf{G}^{-1}\mathbf{X}
\end{aligned} \tag{22}
$$

where we've defined $\mathbf{G} \equiv \mathbf{F} + \mathbf{1}\mathbf{X}^T$. So

$$
\frac{\lambda}{2} + \sigma_B^2 = -\frac{1 + (\mathbf{1}^T\mathbf{G}^{-1}\mathbf{X})}{(\mathbf{1}^T\mathbf{G}^{-1}\mathbf{1})} \tag{23}
$$

and our solution is

$$
\mathbf{w}^* = \frac{1 + (\mathbf{1}^T\mathbf{G}^{-1}\mathbf{X})}{(\mathbf{1}^T\mathbf{G}^{-1}\mathbf{1})}\mathbf{G}^{-1}\mathbf{1} - \mathbf{G}^{-1}\mathbf{X} \tag{24}
$$

In component form, this is

$$
w_i^* = \frac{1 + \sum_{jk}\left(\mathbf{F} + \mathbf{1}\mathbf{X}^T\right)^{-1}_{jk} X_j}{\sum_{jk}\left(\mathbf{F} + \mathbf{1}\mathbf{X}^T\right)^{-1}_{jk}} \sum_j \left(\mathbf{F} + \mathbf{1}\mathbf{X}^T\right)^{-1}_{ij} - \sum_j \left(\mathbf{F} + \mathbf{1}\mathbf{X}^T\right)^{-1}_{ij} X_j \tag{25}
$$

which differs slightly from the result of Efstathiou et al[2] in that $\mathbf{F} \to \mathbf{F} + \mathbf{1}\mathbf{X}^T$.

References

[1] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-year wilkinson microwave anisotropy probe (WMAP) observations: Foreground emission. *The Astrophysical Journal Supplement Series*, 148:97–117, September 2003.

[2] G. Efstathiou, S. Gratton, and F. Paci. Impact of galactic polarized emission on b-mode detection at low multipoles. *MNRAS*, 397(3):1355–1373, August 2009.

[3] H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje. On foreground removal from the wilkinson microwave anisotropy probe data by an internal linear combination method: Limitations and implications. *The Astrophysical Journal*, 612(2):633–646, September 2004. arXiv: astro-ph/0403098.

[4] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman. HEALPix – a framework for high resolution discretization, and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759–771, April 2005. arXiv: astro-ph/0409513.

[5] G. Hinshaw, M. R. Nolta, C. L. Bennett, R. Bean, O. Dor, M. R. Greason, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, N. Odegard, S. S. Meyer, L. Page, H. V. Peiris, D. N. Spergel, G. S. Tucker, L. Verde, J. L. Weiland, E. Wollack, and E. L. Wright. Three-year wilkinson microwave anisotropy probe (WMAP) observations: Temperature analysis. *ApJS*, 170(2):288, June 2007.