1 Karhunen-Loeve Eigenmode Filter

Let's define $d_i(\nu,t) = s_i(\nu,t) + f_i(\nu,t)$ where d is the data we observe, s is our desired signal, and f is some contaminant (e.g. rfi). In baseband data, assuming we want to measure signal from a signal point source, we can model our expected signal at input i, $s_i(\hat{n})$ as

$$s_i(\hat{n}) = e^{2\pi i\nu \frac{\hat{u}_i \cdot \hat{n}}{c}} = e^{2\pi i \nu \phi_i} \tag{1}$$

If the signal is not correlated with the contaminant, we can express the visibilities as

$$\langle dd^\dagger \rangle = S + F \tag{2}$$

Let's suppose there exists some basis where F is a diagonal matrix and S is a diagonal matrix with entries monotonically increasing along the diagonal (note in this case S is rank one for a point source). Denoting this change of basis via the transform $\tilde{d} = R^{-1}$, we have

$$\langle \tilde{d}\tilde{d}^\dagger \rangle = \underbrace{\begin{bmatrix} 1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 1 & 0 \\ 0 & ... & 0 & 1 \end{bmatrix}}_{\tilde{F}} + \underbrace{\begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & 0 \\ 0 & ... & 0 & 0 \end{bmatrix}}_{\tilde{S}} \tag{3}$$

where $\lambda_1 \leq \lambda_2 \leq ...\lambda_n$. Then we can immediately see that in this basis, λ is a direct measure of the signal to noise ratio. Just as "bad" antenna inputs are excluded completely in baseband data, we could construct a filter by excluding any modes corresponding to $\lambda > 1$ (in practice we can also choose our threshold to be higher).

The procedure then for this filter is to first conduct a change of basis where 3 holds ("K-L" space), remove (zero-out) the modes where λ_n are high, and undo the change of basis transformation:

$$dd^\dagger_{\text{cleaned}} = RXR^{-1}dd^\dagger \tag{4}$$

where X is a diagonal matrix with zeroes corresponding to the modes we wish to remove.

Thus the task at hand is to find the change of basis matrix R that will (1) diagonalize both S and F and (2) turn S into the identity matrix. We can obtain this by noting that

$$\langle \tilde{d}\tilde{d}^\dagger \rangle = R^{-1}SR + R^{-1}FR \tag{5}$$

with the constraint

$$R^{-1}FR = I_n \Rightarrow \tilde{S} = R^{-1}FR\tilde{S} \tag{6}$$

We can massage Equation 6 as follows:

$$R^{-1}SR = R^{-1}FR\tilde{S} \tag{7}$$

$$SR = FR\tilde{S} \tag{8}$$

Recall from (3) that $\tilde{S}$ is a diagonal matrix, which we will conveniently rename Λ. Let us also conveniently rename R to be V with column vectors v_i. Then we have

$$SV = FV\Lambda \tag{9}$$

which is the *generalized* eigenvalue problem, where we have packed the eigenvectors v_i into the matrix V and the eigenvalues λ_i into the matrix Λ. **Thus, to find R, we simply need to solve for the eigenvectors v_i in the generalized eigenvalue problem $Sv = \lambda F v$.**

In the case that our signal is a point source, S is a rank 1 matrix ($\Lambda_{ij} = \lambda\delta(j-N, i-N)$). Since our signal only corresponds to a single mode, we can throw away all modes except the one that maximizes the signal to noise.

The full operation of the KLT filter is then

$$b(t) = (s^*)^m R^k_m \Lambda^j_k (R^{-1})^i_j d_i(t) \tag{10}$$

$$b(t) \propto C(R^{-1})^i_N d_i(t) \tag{11}$$

$$\tag{12}$$

where C is a complex number and $(R^{-1})^i_N$ is a vector.

1.1 Equivalence to maximal signal to noise beamformer

Masui 2019 derived the optimal beamforming weight in visibility space. We will show that this result in "baseband data" space is equivalent to the KLT filter with a very similar derivation. Adopting the same formalism, the electric field measured at antenna a at some time t will be

$$d_a(t) = s_a(t) + n_a(t) \tag{13}$$

(note the polarization hand is also contained in index a). Our beamformed baseband data then is

$$b(t) = w^a d_a(t). \tag{14}$$

Our signal $\langle b_s \rangle$ is simply

$$b_s(t) = w^a s_a(t) \tag{15}$$

and our noise is

$$\langle \Delta b^2 \rangle \propto w^a N_{a,b} (w^b)^* \tag{16}$$

where $N_{a,b} \equiv n_a n_b^*$. The optimal beamform weight w_{opt} can be found by finding the value at which $\langle b_s \rangle / \Delta b^2$ is maximized, which is

$$w_{opt,a} \propto N_{ab}^{-1} s_a. \tag{17}$$

In the KLT filter,

$$w_{KLT,a}(t) \propto (R^{-1})^i_N \tag{18}$$

where

$$SR = NR\Lambda \tag{19}$$

or for nonzero λ

$$s_i \underbrace{s_b^* R_b^N}_{scalar} = N_{ib} R_b^N \lambda \tag{20}$$

$$s_i N_{ib}^{-1} \propto R_b^N \tag{21}$$

$$\tag{22}$$

1.2 Polarized Data

In the case that the polarization information of our signal is known, we can repeat the procedure in section 1 except with the stokes information included in S. This generally requires that two modes be saved instead of jut one, so our beamformed data is

$$b^\alpha(t) = N^{\alpha\beta} V^i_\beta d_i(t) \tag{23}$$

where α, β index from 0 to 1 and $V^i_\beta = \Lambda^j_\beta (R^{-1})^i_j$. In general, the operation $V^i_\beta d^\alpha_i(t)$ will result mix the polarization information in our data, so to obtain the beam b^α in the same basis in which our data is measured, an "unmixing" matrix $N^{\alpha\beta}$ is required. The constraint on $N^{\alpha\beta}$ can be obtained by first expressing our signal as

$$s^\alpha_i(t) = c^\alpha_i(t) s_i \tag{24}$$

(note $c^\alpha_i(t) = 0$ for antenna indices i which do not have polarization α), and assuming s is normalized ($s^*_i s^i = 1$)

$$\langle c_\alpha c^*_\beta \rangle_t = I_{\alpha\beta} \tag{25}$$

where I is the intensity matrix. We want our filter to have the property that it preserves the signal, i.e.

$$b^\alpha_{\text{signal}}(t) = N^{\alpha\beta} V^i_\beta s_i(t) = s^{*i} s^\alpha_i(t) = c^\alpha(t) \tag{26}$$

and

$$b^\alpha_{\text{signal}}(t) = N^{\alpha\beta} V^i_\beta s_i(t) = s^{*i} s^\alpha_i(t) = c^\alpha(t). \tag{27}$$

Since this must hold for all times t, we can recast this constraint as

$$N^{\alpha\beta} V^i_\beta s^\gamma_i = \delta_{\alpha,\gamma} \tag{28}$$

or

$$N^{\alpha\beta} = \left(V^i_\beta s^\gamma_i \right)^{-1} \tag{29}$$

1.2.1 Simulating Polarized Data

We want to simulate a polarized source with the covariance matrix

$$\begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_x E_y^* \rangle \end{bmatrix} \tag{30}$$

We can do this by finding the change of basis matrix that diagonalizes the covariance matrix. This is straightforward with an eigendecomposition

$$\begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_x E_y^* \rangle \end{bmatrix} = Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^{-1} \tag{31}$$

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = Q \begin{bmatrix} \mathcal{N}(0, \lambda_1) \\ \mathcal{N}(0, \lambda_2) \end{bmatrix} \tag{32}$$

where $\mathcal{N}(0, \lambda)$ is a gaussian random variable with variance λ and Q is the eigenvector matrix of the original covariance matrix.