Derivaiton of Maximum Orthogonal Complement Analysis (MOCA)

Let the N pixels in a hyperspectral (HS) image be represented as a set of vectors in $\{\mathbf{r}_i\}_{i=1}^{N}$ where $\mathbf{r}_i \in \mathscr{R}^{Lx1}$. The goal of MOCA and MOSP [1, 2, 3] is to find a signal subspace $\mathbf{S}_l$ of rank l so that the vectors in $\mathbf{S}_l$ span the signal subspace and exclude the noise subspace. When applied to a HS image the algorithms iteratively finds a vector which is part of the signal subspace one at a time. This is done via SVD for MOCA and can be done with Automatic Target Generation Process (ATGP) developed by Chang [3]. This increases the signal subspace by rank 1 each time a target vector is chosen. In order to find the next target vector, the algorithms divide the projected pixels into two, mutually exclusive, sets $I_B(l)$ and $I_T(l)$ in the projected subspace in $\mathbf{S}_l$ where the orthogonal projector is $\mathbf{P}_{\mathbf{S}_l}^{\perp} = I - \mathbf{S}_l\left(\mathbf{S}_l^T\mathbf{S}_l\right)^{-1}\mathbf{S}_l^T$ for $1 \le l \le L$. The measure of the signal in the projected pixels is defined in equations 1, 2 and 3.

$$\nu_l = \max_{i \in I_B(l)} \|\mathbf{P}_{\mathbf{S}_l}^{\perp}\mathbf{r}_i\|^2 \tag{1}$$

$$\xi_l = \max_{i \in I_T(l)} \|\mathbf{P}_{\mathbf{S}_l}^{\perp}\mathbf{r}_i\|^2 \tag{2}$$

$$\eta_l = \max\{\xi_l, \nu_l\} \tag{3}$$

This algorithm implies that the subspace is monotonically increasing in the sense that $\mathbf{S}_1 \subset \mathbf{S}_2 \subset ... \subset \mathbf{S}_l$ and that each measure η_l, which is the max residual remaining in the pixels, is monotonically decreasing $\eta_0 \ge \eta_1 \ge ... \ge \eta_l$. Therefore, the question arises of how to stop the search of the subspace at l^* which contains the target space and not the background noise. MOCA and MOSP both derive a hypothesis test which can be resolved using a maximum likelihood test or a Neyman-Pearson test.

$$H_0 : \eta_l \approx p(\eta_l|H_0) = p_0(\eta_l) \tag{4}$$

versus

$$H_1 : \eta_l \approx p(\eta_l|H_1) = p_1(\eta_l) \tag{5}$$

Here the null hypothesis H_0 represents the maximum residual of the subspace remaining in the image is from the background pixels. This approach requires the distribution function under H_0 and H_1. We use a Gumbel distribution for $p_0(\eta_l)$ and uniform for $p_1(\eta_l)$. Here we derive the Gumbel distribution. We assume that the background is Gaussian with zero mean because the signal has been removed via projection. Therefore, we start with equation 6 where $\mathbf{r}_i$ are background pixels and $\mathbf{P}_{\mathbf{S}_l}^{\perp}\mathbf{r}_i = (n_{i1}, n_{i2}, ..., n_{iL-l})^T$ where $n_{ij} \sim N(0, \sigma^2)$. We assume the background projections are IID and stationary. Interestingly, the random variable of 6 is a composite random variable where n_{ij} is a centered Gaussian, n_{ij}^2 is Chi-squared, $\sum_{j=1}^{L-l} n_{ij}^2$ is Gaussian and the $\max_{i \in I_B(l)}\left(\sum_{j=1}^{L-l} n_{ij}^2\right)$ is Gumbel.

$$\eta_l = \max_{i \in I_B(l)} \|\mathbf{P}_{\mathbf{S}_l}^{\perp}\mathbf{r}_i\|^2 = \max_{i \in I_B(l)}\left(\sum_{j=1}^{L-l} n_{ij}^2\right) \tag{6}$$

We then proceed with the approach of Chang, et. al. in [1] in appendix A. The random variable $m_{ij} = n_{ij}^2$ has a Chi-Squared distribution of degree 1 and the mean and variance of this new random variable is defined in equations 8 and 9.

$$m_{ij} = n_{ij}^2 \tag{7}$$

$$\mu_{m_{ij}} = E\left[m_{ij}\right] = E\left[n_{ij}^2\right] = \sigma^2 E\left[\chi_1^2\right] = \sigma^2 \tag{8}$$

$$\sigma_{m_{ij}}^2 = E\left[\left(m_{ij} - \mu_{m_{ij}}\right)^2\right] = E\left[\left(n_{ij}^2 - \sigma^2\right)^2\right] = E\left[n_{ij}^4\right] - 2\sigma^2 E\left[n_{ij}^2\right] + \sigma^4 = 3\sigma^4 - \sigma^4 = 2\sigma^4 \tag{9}$$

Now we can define a new random variable ζ_i for the ith pixel as the sum of the Chi-Squared distributed m_{ij} random variables across the bands $\zeta_i = \sum_{j=1}^{L-l} m_{ij}$. The total number of bands in an HSI image is large and the signal subspace is small such as $l << L$. Therefore, due to the central limit theorem, the sum in equation 6 approaches a Gaussian distribution with mean and variance of the new random variable ζ_i of equations 11 and 12.

$$\zeta_i = \sum_{j=1}^{L-l} m_{ij} = \sum_{j=1}^{L-l} n_{ij}^2 \tag{10}$$

$$\mu_{\zeta_i} = E[\zeta_i] = E\left[\sum_{j=1}^{L-l} m_{ij}\right] = E\left[\sum_{j=1}^{L-l} n_{ij}^2\right] = \sum_{j=1}^{L-l} E\left[n_{ij}^2\right] = \sum_{j=1}^{L-l} \mu_{m_{ij}} = (L-l)\sigma^2 \tag{11}$$

$$\begin{aligned}
\sigma_\zeta^2 &= E\left[\left(\zeta_i - \mu_{\zeta_i}\right)^2\right] = E\left[\left(\sum_{j=1}^{L-l} m_{ij} - (L-l)\sigma^2\right)^2\right] && (12)\\
&= E\left[\left(\sum_{j=1}^{L-l} m_{ij}\right)^2\right] - 2(L-l)\sigma^2 E\left[\sum_{j=1}^{L-l} m_{ij}\right] + (L-l)^2\sigma^4 && (13)\\
&= E\left[\left(\sum_{j=1}^{L-l} n_{ij}^2\right)^2\right] - (L-l)^2\sigma^4 = VAR\left[\sum_{j=1}^{L-l} m_{ij}\right] + E\left[\sum_{j=1}^{L-l} m_{ij}\right]^2 - (L-l)^2\sigma^4 && (14)\\
&= \sum_{j=1}^{L-l} VAR\left[m_{ij}\right] + \mu_{\zeta_i}^2 - (L-l)^2\sigma^4 = \sum_{j=1}^{L-l} 2\sigma^4 + (L-l)^2\sigma^4 - (L-l)^2\sigma^4 = 2(L-l)\sigma^4 && (15)
\end{aligned}$$

Therefore, we can define a normalized version of the random variables ζ_i for each pixel by subtracting the mean and dividing by the standard deviation as in equation 16.

$$X_i = \frac{\zeta_i - \mu_{\zeta_i}}{\sigma_{\zeta_i}} = \frac{\zeta_i - \sigma^2(L-l)}{\sigma^2\sqrt{2(L-l)}} \sim N(0,1) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{x^2/2} \tag{16}$$

So for each pixel in the set $\mathbf{r}_i \in I_B(l)$ we can define a new random variable $\eta_l = M_N$ as the max of all the L_2 norms of the projections of the background pixels where the N is the number of pixels in the image.

$$\eta_l = M_N = \max_{I_B(l)}\{X_1, X_2, ..., X_N\} \tag{17}$$

Now we can find the probability distribution for M_N which turns out to be Gumbel which we will derive. We start with the cumulative distribution function (CDF) of M_N as in equation 18. Note that $\Phi(x)$ is the normal distribution function.

$$P(M_N \le x) = P(X_1 \le x, X_2 \le x, ..., X_N \le x) = \prod_{i=1}^{N} P(X_i \le x) = \Phi^N(x) = \left[\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz\right]^N \tag{18}$$

Note that as $N \to \infty$ the distribution of M_N is degenerate because the Gaussian is continuous and you will always exceed any value of x with enough samples so the PDF is degenerate to 1. The Extreme Value Theorem (EVT) finds normalizing values for scale $a_N \ge 0$ and location b_N so that the distribution converges to non-degenerate $G(x)$ as in equation19.

$$\lim_{N\to\infty} P\left(a_N(M_N - b_N) \le x\right) = G(x) \tag{19}$$

According to the Extreme Type Theorem (ETT) as explored in Leadbetter (1983) [4], when $G(x)$ exists, it converges to one of three forms with normalization constants a_N and b_N which are dependent on the underlying distribution of X_i. Under the condition that X_i is Gaussian, we can further derive the Gumbel distribution starting in equation 20.

$$\lim_{N\to\infty} P\left(M_N \le a_N^{-1}x + b_N\right) = \lim_{N\to\infty} \prod_{i=1}^{N} P\left(X_i \le a_N^{-1}x + b_N\right) = \lim_{N\to\infty} \Phi^N\left(a_N^{-1}x + b_N\right) = G(x) \tag{20}$$

Taking the log of both sides we eventually arrive at the Gumbel distribution in terms of a_N and b_N.

$$\lim_{N\to\infty} N\log\left(\Phi(a_N^{-1}x + b_N)\right) = \log(G(x)) \tag{21}$$

Note that as $N \to \infty$, then $\log\left(\Phi(a_N^{-1}x + b_N)\right) \approx \Phi(a_N^{-1}x + b_N) - 1$. This is due to the bound of $1 - 1/x \leq \log(x) \leq x - 1$.

$$\lim_{N\to\infty} N(1 - \Phi(a_N^{-1}x + b_N)) = -\log(G(x)) \tag{22}$$

$$\lim_{N\to\infty} N(1 - \Phi(a_N^{-1}x + b_N)) = \tau \tag{23}$$

Where τ is a constant. If we let $\tau = e^{-x}$, as in Leadbetter (1983), theorem 1.5.1, then the following equations result.

$$\lim_{N\to\infty} N(1 - \Phi(a_N^{-1}x + b_N)) = e^{-x} \tag{24}$$

$$\lim_{N\to\infty} \Phi(a_N^{-1}x + b_N) = \lim_{N\to\infty} \left\{1 - e^{-x}/N\right\} \tag{25}$$

Then we raise both sides to the power of N.

$$\lim_{N\to\infty} \Phi^N(a_N^{-1}x + b_N) = \lim_{N\to\infty} \left\{(1 - e^{-x}/N)^N\right\} \tag{26}$$

$$\lim_{N\to\infty} P(M_N \leq a_N^{-1}x + b_N) = exp\left\{-e^{-x}\right\} \tag{27}$$

We've arrived at the standard Gumbel distribution in equation 27. Now the task is to find the normalizing constants a_N and b_N when X_i is a standard normal distribution. We proceed by following the derivation in Leadbetter (1983) [4] starting at theorem 1.5.3. For this derivation, it will simplify things if we let $u_N = a_N^{-1}x + b_N$. We will develop an expression for u_N, then extract the a_N and b_N constants from there. In addition, we will use the equality of $1 - \Phi(u_N) \sim \phi(u_N)/u_N$. This can be shown by using L'Hopital's rule as in equation 28.

$$\lim_{N\to\infty} \frac{1 - \Phi(u_N)}{\phi(u_N)/u_N} = \lim_{N\to\infty} \frac{-\phi(u_N)}{-\phi(u_N)(1 + 1/u_N^2)} = -\lim_{N\to\infty} \frac{1}{1 + 1/u_n^2} = 1 \tag{28}$$

Now we can use equations 24 and the equality in 28 to make an equivalency using $1 - \Phi(u_N) = \phi(u_N)/u_N = 1/Ne^{-x}$. Then we can proceed with the division to create equation 29.

$$\lim_{N\to\infty} 1 - \Phi(u_N) = \lim_{N\to\infty} 1/Ne^{-x} = \lim_{N\to\infty} \phi(u_N)/u_N \implies \lim_{N\to\infty} (1/Ne^{-x})(u_N)/\phi(u_N) = 1 \tag{29}$$

Taking the log of the lim in equation 29, we get the following.

$$\lim_{N\to\infty} \left\{-\log(N) - x + \log(u_N) - \log(\phi(u_N))\right\} = 0 \tag{30}$$

Then we arrive at 31 which matches the equation 1.5.6 in Leadbetter (1983) [4].

$$\lim_{N\to\infty} \left\{-\log(N) - x + \log(u_N) + 1/2\log(2\pi) + u_N^2/2\right\} = 0 \tag{31}$$

Now the limit in equation 31 contains two expressions for u_N. One is in the term of $\log(u_N)$ and the other is in the term $u_N^2/2$. We would like to solve this equation for u_N. If we divide 31 by $\log(N)$, then take the limit at $n \to \infty$, then we can eliminate all terms but the $u_N^2/2$ term. This is possible because $\log(N) \to \infty$ as $N \to \infty$ and the other terms grow slowly compared to the squared term. The other terms are collected in a constant denotes $o(1)$.

$$\lim_{N\to\infty} \left\{\frac{1}{\log(N)}\left(-\log(N) - x + \log(u_N) + 1/2\log(2\pi) + u_N^2/2\right)\right\} = 0 \tag{32}$$

$$\lim_{N\to\infty} \left\{\frac{u_N^2}{2\log(N)}\right\} + o(1) = 1 \tag{33}$$

Then we take the log of this limit in equation 33 and then solve for $\log(u_N)$.

$$\lim_{N\to\infty} \left\{2\log(u_N) - \log(2) - \log(\log(N)) + \log(o(1))\right\} = 0 \tag{34}$$

$$\lim_{N\to\infty} \log(u_N) = \lim_{n\to\infty} \left\{1/2\log(2) + 1/2\log(\log(N)) - 1/2\log(o(1))\right\} \tag{35}$$

If we follow Leadbetter (1983) on page 15 [4] we take 35 and substitute it back into equation 31, then we get the following.

$$\lim_{N\to\infty}\left\{-\log(N) - x + 1/2\log(2) + 1/2\log(\log(N)) - 1/2\log(o(1)) + 1/2\log(2\pi) + u_N^2/2\right\} = 0 \tag{36}$$

Then solving for u_N.

$$\lim_{N\to\infty} u_N^2/2 = \lim_{N\to\infty}\left\{x + \log(N) - 1/2\log(4\pi) - 1/2\log(\log(N)) + 1/2\log(o(1))\right\} \tag{37}$$

Then we multiple by 2 and then factor out a $\log(N)$ term.

$$\lim_{N\to\infty} u_N^2 = \lim_{N\to\infty}\left\{2\log(N)\left\{1 + \frac{x - 1/2\log(4\pi) - 1/2\log(\log(N))}{\log(N)} + o\left(\frac{1}{\log(N)}\right)\right\}\right\} \tag{38}$$

Then we take a square root as was done in Leadbetter (1983) on page 15 [4]. Notice that we take advantage of the fact that $\sqrt{1+x} \approx 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{5x^4}{128}\ldots$ and then we only use the first order of $1 + \frac{x}{2}$.

$$\lim_{N\to\infty} u_N = \lim_{N\to\infty}\left\{(2\log(N))^{1/2}\left\{1 + \frac{x - 1/2\log(4\pi) - 1/2\log(\log(N))}{2\log(N)} + o\left(\frac{1}{\log(N)}\right)\right\}\right\} \tag{39}$$

Now this is in the form of $u_N = a_N x + b_N$

$$a_N = (2\log(N)^{-1/2} \tag{40}$$

$$b_N = (2\log(N))^{1/2} - \frac{1}{2}(2\log(N))^{-1/2}(\log(\log(N)) + \log(4\pi)) \tag{41}$$

Combining these constants with the definition of X_i in equation 16 and the standard Gumbel 27, we arrive at the final Gumbel distribution for a background pixel under MOCA and MOSP for H_0 in equation 42.

$$F_{\eta_l}(x) = F_{M_N}(x) = exp\left\{-e^{-a_N(x+b_N)}\right\} \tag{42}$$

$$= exp\left\{-e^{-(2\log(N))^{1/2}\left[\frac{x-\sigma^2(L-l)}{\sigma^2\sqrt{2(L-l)}} - (2\log(N))^{1/2} + \frac{1}{2}(2\log(N))^{-1/2}(\log(\log(N))+\log(4\pi))\right]}\right\} \tag{43}$$

REFERENCES

[1] C. Chang, W. Xiong, H. Chen, and J. Chai, "Maximum orthogonal subspace projection approach to estimating the number of spectral signal sources in hyperspectral imagery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 504–520, June 2011.

[2] O. Kuybeda, D. Malah, and M. Barzohar, "Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5579–5592, Dec 2007.

[3] C.-I. Chang, *Hyperspectral Data Processing*. John Wiley & Sons, Inc., 2013.

[4] M. R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, ser. Springer Series in Statistics. Springer Verlag, 1983.

[5] C. Chang, "A review of virtual dimensionality for hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1285–1305, April 2018.