Statistical Unmixing of SAR Images

Stian Normann Anfinsen,
University of Tromsø – The Arctic University of Norway
Department of Physics and Technology
Email: stian.normann.anfinsen(at)uit.no. Tel: +47 776 45173.

Abstract

A method is presented which uses logarithmic statistics to detect and characterise class mixtures and targets in background clutter in synthetic aperture radar (SAR) images. Mixtures of ground cover types show up as extreme radar texture in statistical analysis of SAR images. Instead of modelling this as a spatially nonstationary radar cross section, this paper demonstrates how a mixture model analysis can be used to characterise the separate components and estimate their mixing proportions.

1 Theory

1.1 Mixture Model

Let X be a real and positive random variate which represents a measurement obtained within a certain region of a SAR image. It is assumed that the region of interest is heterogeneous, and that X can be modelled with a two-component mixture model. This means that the observation X will be drawn from a distribution with probability density function (pdf) $p_{X_1}(x)$ with probability π_1, or from a distribution with pdf $p_{X_2}(x)$ with probability π_2. The pdfs are distinct, meaning that $p_{X_1}(x) \neq p_{X_2}(x)$, and the mixing proportions obey

$$\pi_1 + \pi_2 = 1\,. \tag{1}$$

The overall pdf of X thus becomes

$$p_X(x) = \pi_1 \cdot p_{X_1}(x) + \pi_2 \cdot p_{X_2}(x)\,. \tag{2}$$

1.2 General Mixture Moments

We shall now express the moments of X in terms of the moments of the mixture components $\{X_i\}_{i=1}^n$, for a general n. Denote the mean, the variance, and the mixing proportion of X_i as μ_i, σ_i^2 and π_i, respectively. The general jth-order moment of an n-component mixture can be written in terms of a binomial expansion as [1, Ch. 1.2.4]

$$\begin{aligned}\mathrm{E}\{(X-\mu)^j\} &= \sum_{i=1}^n \pi_i\,\mathrm{E}\left\{(X_i-\mu_i+\mu_i-\mu)^j\right\} \\ &= \sum_{i=1}^n \sum_{k=0}^{j} \pi_i \binom{j}{k} \delta_i^{j-k}\,\mathrm{E}\left\{(X_i-\mu_i)^k\right\} \\ &= \sum_{i=1}^n \pi_i\,\mathrm{E}\left\{(X_i-\mu_i)^j\right\} \\ &\quad + \sum_{i=1}^n \sum_{k=0}^{j-1} \pi_i \binom{j}{k} \delta_i^{j-k}\,\mathrm{E}\left\{(X_i-\mu_i)^k\right\}\end{aligned} \tag{3}$$

where $\mathrm{E}\{\cdot\}$ is the expectation value operator, $\binom{j}{k}$ is the binomial coefficient, and $\delta_i = \mu_i - \mu$, with

$$\mu = \mathrm{E}\{X\} = \sum_{i=1}^n \pi_i \mu_i\,. \tag{4}$$

Observe that the final expression in (3) consists of two parts: a sum and a double-sum. The former is a proportional mixture of the jth-order moment for the individual components. The latter contains cross-terms between the components, since all terms in the double-sum depend on the common mean, μ, through δ_i. Hence, the general jth-order moment is rewritten as

$$\mathrm{E}\{(X-\mu)^j\} = W_j + B_j \tag{5}$$

with the within-class contribution defined as

$$W_j\{X\} = \sum_{i=1}^n \pi_i\,\mathrm{E}\left\{(X_i-\mu_i)^j\right\} \tag{6}$$

and the between-class contribution as

$$B_j\{X\} = \sum_{i=1}^n \sum_{k=0}^{j-1} \pi_i \binom{j}{k} \delta_i^{j-k}\,\mathrm{E}\left\{(X_i-\mu_i)^k\right\}\,, \tag{7}$$

both indexed by the moment order, j.
Another interpretation of B_j is that it quantifies the excess portion of the central moments elicited by the mixing of the distributions $\{p_{X_i}(x)\}_{i=1}^n$. W_j is merely a weighted mean of the moments produced by random variables $\{X_i\}_{i=1}^n$ that are not mixed. The idea is to use the B_j to detect the presence of a mixture and to characterise the kind of mixture by resolving the mixing proportions and the parameters of the mixing distributions. It is seen from (7) that $B_1 \triangleq 0$, so the mixture causes no excess in the mean, but all higher-order B_j are generally nonzero.

1.3 Two-class Mixture Moments

The scope is from here on limited to the two-component model as we enter a study of the second, third and fourth-order central moment of a two-class mixture. These are referred to as variance, skewness and kurtosis, while noting that the final two may alternatively be defined as standardised central moments of their respective order.

The central moments up to fourth-order have already been given by Kim and White [2] in a form which is easily obtained from (3). We elaborate on their result by deriving simplified expressions for the between-class contribution, B_j, for $j = \{2, 3, 4\}$. The following presentation uses the notation: $\delta = \mu_1 - \mu_2$.

For the variance of a two-class mixture, the between-class contribution becomes

$$\begin{aligned} B_2\{X\} &= \sum_{i=1}^{2} \sum_{k=0}^{1} \pi_i \binom{2}{k} \delta_i^{2-k} \operatorname{E}\left\{(X_i - \mu_i)^k\right\} \\ &= \pi_1 \delta_1^2 + \pi_2 \delta_2^2 = \pi_1 \pi_2 \delta^2 \end{aligned} \tag{8}$$

The derivation is straight-forward algebra using (1) and (4). The factor δ^2 in B_2 is the square of the difference in component means and has an obvious interpretation as between-class dispersion. We also note that $\pi_1\pi_2 = \pi_1 - \pi_1^2$ is maximum when $\pi_1 = \pi_2 = 0.5$.

The between-class contribution to the skewness of a two-class mixture is

$$\begin{aligned} B_3\{X\} &= \sum_{i=1}^{2} \sum_{k=0}^{2} \pi_i \binom{3}{k} \delta_i^{3-k} \operatorname{E}\left\{(X_i - \mu_i)^k\right\} \\ &= \pi_1 \delta_1^3 + \pi_2 \delta_2^3 + 3\pi_1 \delta_1 \sigma_1^2 + 3\pi_2 \delta_2 \sigma_2^2 \\ &= \pi_1 \pi_2 \delta \left[(\pi_2^2 - \pi_1^2)\delta^2 + 3(\sigma_1^2 - \sigma_2^2)\right] . \end{aligned} \tag{9}$$

The sign of B_3 is determined by the difference in means, $\delta = \mu_1 - \mu_2$, in combination with the relative size of the difference in squared mixing proportions, $\pi_2^2 - \pi_1^2$, and the difference in variances, $\sigma_1^2 - \sigma_2^2$.

The kurtosis of a two-class mixture has the following between-class contribution:

$$\begin{aligned} B_4\{X\} &= \sum_{i=1}^{2} \sum_{k=0}^{3} \pi_i \binom{4}{k} \delta_i^{4-k} \operatorname{E}\left\{(X_i - \mu_i)^k\right\} \\ &= \sum_{i=1}^{2} \pi_i \left(\delta_i^4 + 6\delta_i^2 \sigma_i^2 + 4\delta_i \gamma_i\right) \\ &= \pi_1 \pi_2 \delta^2 \left[(\pi_1^3 + \pi_2^3)\delta^2 + 6(\pi_1 \sigma_2^2 + \pi_2 \sigma_1^2) \right. \\ &\qquad \left. + 4(\gamma_1 - \gamma_2)/\delta\right] \end{aligned} \tag{10}$$

where $\gamma_i = \operatorname{E}\{(X_i - \mu_i)^3\}$ is the skewness of component i. The sign of B_4 depends on the relative size of the mixing proportions, the variances, the difference in means, δ, and the difference in skewnesses, $\gamma_1 - \gamma_2$.

1.4 Gamma Mixture Moments

We now insert two gamma distributions with equal shape parameter, $L > 0$, but unequal means, $\mu_1 \neq \mu_2$, into the two-class mixture model assumed in Section 1.1. The common shape parameter is reasonable for SAR data, because it corresponds to the equivalent number of looks [3], which is an image constant determined by the level of multilook averaging [4].

The pdf of X is thus given by (2), defined as a mixture of the gamma distributions

$$p_{X_i}(x; \mu_i, L) = \left(\frac{L}{\mu_i}\right)^L \frac{x^{L-1}}{\Gamma(L)} \exp\left(-\frac{L}{\mu_i} x\right) \tag{11}$$

for $i = \{1, 2\}$, where $\Gamma(\cdot)$ is the gamma function, $\mu_i > 0$ and $L > 0$. This is denoted $X_i \sim \gamma(\mu_i, L)$.

The variance and skewness of a gamma distributed variable are $\sigma_i^2 = \mu_i^2/L$ and $\gamma_i = 2\mu_i^3/L^2$. Hence, the between-class contributions to the central mixture moments become

$$B_2\{X\} = \pi_1 \pi_2 \delta^2 , \tag{12}$$

$$B_3\{X\} = \pi_1 \pi_2 \delta[(\pi_2^2 - \pi_1^2)\delta^2 + \frac{3}{L}(\mu_1^2 - \mu_2^2)] , \tag{13}$$

$$\begin{aligned} B_4\{X\} &= \pi_1 \pi_2 \delta^2 \left[(\pi_1^3 + \pi_2^3)\delta^2 \right. \\ &\quad \left. + \frac{6}{L}(\pi_1 \mu_2^2 + \pi_2 \mu_1^2) + \frac{4}{L^2}(\mu_1^2 + \mu_1 \mu_2 + \mu_2^2)\right] . \end{aligned} \tag{14}$$

Since $\delta^2 \geq 0$ and $\pi_i, \mu_i, L > 0$, it is easy to verify that $B_2 \geq 0$ and $B_4 \geq 0$, with equality if and only if $\mu_1 = \mu_2$, in which case $p_{X_1}(x) = p_{X_2}(x)$.

1.5 Logarithmic Gamma Mixture Moments

We still assume $X_i \sim \gamma(\mu_i, L)$, but now consider the moments of $Y_i = \ln X_i$, or equivalently, the logarithmic moments of X_i. The mean of Y_i becomes $\tilde{\mu}_i = \psi^{(0)}(L) + \ln(\mu_i/L)$, the variance is $\tilde{\sigma}_i^2 = \psi^{(1)}(L)$, and the skewness is $\tilde{\gamma}_i = \psi^{(2)}(L)$ [5,6]. Here $\psi^{(r)}(\cdot)$ denotes the polygamma function of order r.

Note that only the first-order moment of Y_i depends on the mean. Due to the logarithmic transformation, the higher-order moments only depend on the common shape parameter L. We thus have

$$\tilde{\mu}_1 - \tilde{\mu}_2 = \ln\left(\frac{\mu_1}{\mu_2}\right) , \tag{15}$$

$$\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2 = \psi^{(1)}(L) - \psi^{(1)}(L) = 0 , \tag{16}$$

$$\tilde{\gamma}_1 - \tilde{\gamma}_2 = \psi^{(2)}(L) - \psi^{(2)}(L) = 0 . \tag{17}$$

When these are inserted into Eqs. (8)-(10), we obtain the between-class contribution to the logarithmic mixture moments, whose expressions are seen to be simpler than for the linear case. They become

$$B_2\{Y\} = \pi_1 \pi_2 \tilde{\delta}^2 , \tag{18}$$

$$B_3\{Y\} = \pi_1 \pi_2 (\pi_2^2 - \pi_1^2) \tilde{\delta}^3 , \tag{19}$$

$$B_4\{Y\} = \pi_1 \pi_2 \tilde{\delta}^2 \left[(\pi_1^3 + \pi_2^3)\tilde{\delta}^2 + 6\psi^{(1)}(L)\right] , \tag{20}$$

where we define

$$\tilde{\delta} = \tilde{\mu}_1 - \tilde{\mu}_2 = \ln\left(\frac{\mu_1}{\mu_2}\right) . \tag{21}$$

1.6 Logarithmic Wishart Mixture Moments

The theory can be extended to multilook polarimetric SAR data, where a pixel is represented by the polarimetric covariance or coherency matrix, denoted $\mathbf{C}$. Let $\mathbf{C} \in \mathbb{C}^{d\times d}$ be a random matrix defined on the cone of complex, Hermitian and positive semidefinite matrices with dimension d, denoted Ω_+. Then assume a two-class mixture model, such the pdf of $\mathbf{C}$ is

$$p_{\mathbf{C}}(\mathbf{C}) = \pi_1 \cdot p_{\mathbf{C}_1}(\mathbf{C}) + \pi_2 \cdot p_{\mathbf{C}_2}(\mathbf{C}) , \qquad (22)$$

with mixing proportions π_1 and π_2, and pdfs $p_{\mathbf{C}_1}(\mathbf{C})$ and $p_{\mathbf{C}_2}(\mathbf{C})$ for the two class components. Further assume that the components follow the scaled complex Wishart distribution, such that

$$p_{\mathbf{C}_i}(\mathbf{C}; \mathbf{\Sigma}_i, L) = \frac{L^{Ld}}{\Gamma_d(L)} \frac{|\mathbf{C}|^{L-d}}{|\mathbf{\Sigma}_i|^L} \operatorname{etr}(-L\mathbf{\Sigma}_i^{-1}\mathbf{C}) \quad (23)$$

for $i \in \{1, 2\}$, where $| \cdot |$ is the determinant, $\operatorname{etr}(\cdot) = \exp(\operatorname{tr}(\cdot))$ is the exponential trace operator, and

$$\Gamma_d(L) = \pi^{\frac{d(d-1)}{2}} \prod_{i=0}^{d-1} \Gamma(L - i) \qquad (24)$$

is the multivariate gamma function of the complex kind [6,7]. The distribution parameters are the component specific scale matrix $\mathbf{\Sigma}_i = \mathrm{E}\{\mathbf{C}\}$ and the common shape parameter L.

A matrix-variate pdf defined on Ω_+ can be characterised by statistics known as matrix log-cumulants (MLCs). For the scaled complex Wishart distribution the low-order MLCs are given as [6, 7]

$$\kappa_1 = \mathrm{E}\{\ln |\mathbf{C}|\} = \psi_d^{(0)}(L) + \ln |\mathbf{\Sigma}| - d \ln L , \qquad (25)$$

$$\kappa_2 = \mathrm{E}\{(\ln |\mathbf{C}| - \kappa_1)^2\} = \psi_d^{(1)}(L) , \qquad (26)$$

$$\kappa_3 = \mathrm{E}\{(\ln |\mathbf{C}| - \kappa_1)^3\} = \psi_d^{(2)}(L) , \qquad (27)$$

with the rth-order multivariate polygamma function of the complex kind defined as

$$\psi_d^{(r)}(L) = \frac{d^{r+1}}{dL^{r+1}} \ln \Gamma_d(L) . \qquad (28)$$

The first three matrix log-cumulants are identical to the first three central moments of $\ln |\mathbf{C}|$ (which is not true for higher-order MLCs). We may therefore apply the theory of mixture moments and their decomposition into within-class and between-class contributions, as presented in Section 1.2 and 1.3. The between-class contributions to the second, third, and fourth-order MLC of a two-class mixture of scaled complex Wishart distributions become

$$B_2\{\mathbf{C}\} = \pi_1\pi_2\Delta^2 , \qquad (29)$$

$$B_3\{\mathbf{C}\} = \pi_1\pi_2(\pi_2^2 - \pi_1^2)\Delta^3 , \qquad (30)$$

$$B_4\{\mathbf{C}\} = \pi_1\pi_2\Delta^2 \left[(\pi_1^3 + \pi_2^3)\Delta^2 + 6\psi_d^{(1)}(L) \right] , \quad (31)$$

where we define

$$\Delta = \ln \left(\frac{|\mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_2|} \right) . \qquad (32)$$

Remark that all expressions in this section reduces to those in Section 1.5 when $d = 1$, in which the scaled complex Wishart matrix becomes a gamma variable.

2 Inference

This section outlines a way of extracting the mixing proportions and the scale parameters of the mixture model from the moments reviewed in previous sections. We assume that the image constant L is known for the given SAR focusing scheme, and are left with estimating π_1, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ in the general polarimetric case. This is done by relating the desired mixture model parameters to sample moments that can be computed from the data subset.

Denote the sample MLCs as $\langle\kappa_i\rangle, i \in \{1, 2, 3\}$. These are finite sample estimates of the population MLCs defined by (25)-(27). The sample MLCs can be computed as sample means or by unbiased k-estimators [8]. They are related to the estimates of the mixing proportions, $\hat{\pi}_1$ and $\hat{\pi}_2$, and of the scale matrices, $\hat{\mathbf{\Sigma}}_1$ and $\hat{\mathbf{\Sigma}}_2$, by the following estimation equations:

$$\langle\kappa_1\rangle - \psi_d^{(0)}(L) = \ln |\hat{\pi}_1\hat{\mathbf{\Sigma}}_1 + \hat{\pi}_2\hat{\mathbf{\Sigma}}_2| , \qquad (33)$$

$$\langle\kappa_2\rangle - \psi_d^{(1)}(L) = \hat{\pi}_1\hat{\pi}_2\hat{\Delta}^2 , \qquad (34)$$

$$\langle\kappa_3\rangle - \psi_d^{(2)}(L) = \hat{\pi}_1\hat{\pi}_2(\hat{\pi}_2^2 - \hat{\pi}_1^2)\hat{\Delta}^3 , \qquad (35)$$

where $\hat{\Delta} = \ln(|\hat{\mathbf{\Sigma}}_1|/|\hat{\mathbf{\Sigma}}_2|)$. In addition, we may use

$$\langle\mathbf{C}\rangle = \hat{\pi}_1\hat{\mathbf{\Sigma}}_1 + \hat{\pi}_2\hat{\mathbf{\Sigma}}_2 , \qquad (36)$$

where $\langle\mathbf{C}\rangle$ is a sample mean estimate of the scale matrix $\mathbf{\Sigma}$ of the mixture.

There are alternative ways of combining the estimation equations using various optimization techniques. A simple method is shown here. To isolate $\hat{\pi}_1$, we may combine (34) and (35) into

$$\rho = \frac{\left(\langle\kappa_3\rangle - \psi_d^{(2)}(L)\right)^2}{\left(\langle\kappa_2\rangle - \psi_d^{(1)}(L)\right)^3} = \frac{1 - 2\hat{\pi}_1}{\hat{\pi}_1(1 - \hat{\pi}_1)} . \qquad (37)$$

To infer the mixing proportions, we compute the statistic ρ and solve (37) numerically for $\hat{\pi}_1$. Figure 1 shows the nonmonotonic relationship between π_1 and ρ. Due to the ambiguity between the mixing proportions, we can introduce the convention $\pi_1 \geq \pi_2$ and limit the search for π_1 to the interval $[0.5, 1]$ to ensure that the problem has a unique solution.

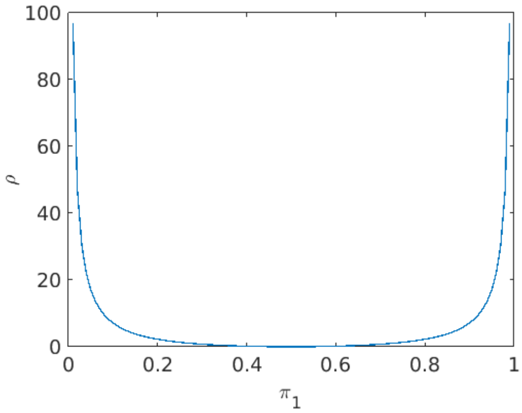

Figure 1: Relation between mixing proportion π_1 and ρ.

3 Future Work

The theory will be verified and illustrated by experiments on simulated data, that show the capabilities of statistical unmixing under both ideal conditions. It will then be applied to relevant applications using real data, such as estimation of melt pond fractions over sea ice and forest density in sparse forest areas with low to medium above-ground biomass levels.

References

[1] S. Frühwirth-Schnatter: *Finite Mixture and Markov Switching Models*. Springer, New York, USA, 2006.

[2] T.-H. Kim: *On more robust estimation of skewness and kurtosis*, Finance Res. Lett., vol. 1, no. 1, pp. 56–73, 2003.

[3] S.N. Anfinsen, A.P. Doulgeris, and T. Eltoft: *Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery*, IEEE Trans. Geosci. Remote Sens., vol. 47, no. 11, pp. 3795–3809, 2009.

[4] C. Oliver and S. Quegan: *Understanding Synthetic Aperture Radar Images*. 2nd ed., SciTech Publishing, Raleigh, USA, 2004.

[5] J.-M. Nicolas: *Introduction aux statistique de deuxième espèce: Application des logs-moments et des logs-cumulnats à l'analyse des lois d'images radar*, Traitement du Signal, vol. 19, no. 3, pp. 139–167, 2002. In French. English translation in [7].

[6] S.N. Anfinsen and T. Eltoft: *Application of the matrix-variate Mellin transform to analysis of polarimetric radar images*, IEEE Trans. Geosci. Remote Sens., vol. 49, no. 6, pp. 2281–2295, 2011.

[7] S.N. Anfinsen: *Statistical analysis of multilook polarimetric radar images with the Mellin transform*, Ph.D. dissertation, University of Tromsø, Dept. of Physics and Technology, Tromsø, Norway, 2010.

[8] M. Kendall: *Kendall's Advanced Theory of Statistics: Distribution Theory*, 5th ed., vol. 1, ch. x, London, UK: Charles Griffin, 1987.