Masses des lois multinomiales négatives. Application au traitement d'images polarimétriques.

Philippe Bernardoff, Florent Chatelain, Jean-Yves Tourneret

Masses des lois multinomiales negatives. Application au traitement d images polarimetriques.

Philippe Bernardoff, Florent Chatelain et Jean-Yves Tourneret.

P. Bernardoff : Université de Pau et des Pays de l'Adour, avenue de l'Université, 64000 Pau, France. LMAP UMR 5142. E-mail : philippe.bernardoff @univ-pau.fr

F. Chatelain : Gipsa-lab, Département Signal et Images, 961 rue de la Houille Blanche, BP 46, 38402 Saint Martin d'Heres, France. E-mail : florent.chatelain@gipsa-lab.inpg.fr

J.-Y. Tourneret : Université de Toulouse, IRIT/ENSEEIHT/T esa, 2 rue Charles Camichel, BP 7122, 31071. Toulouse cedex 7, France

Résumé : Cet article est dérivé d'une nouvelle expression des masses des lois multinomiales négatives multivariées. Cette expression des masses peut être utilisée pour déterminer les estimateurs du maximum de vraisemblance de ses paramètres inconnus. Une application au traitement d'images polarimétriques est étudiée. Plus précisément, les estimateurs du degré de polarisation utilisant la méthode du maximum de vraisemblance avec différentes combinaisons d'images sont comparés.

Mots clés : Lois multinomiales négatives, maximum de vraisemblance,images polarimétriques, degrés de polarisation.

Abstract : This paper derives new closed-form expressions of the masses of multivariate negative multinomiale distributions. These masses can be maximized to determine the maximum likelihood estimator of its unknown parameters. An application to polarimetric image processing is investigated. More precisely, estimators of the polarization degree using maximum likelihood methods with different combinations of images are compared.

Keywords : Negative multinomial distributions, maximum likelihood, polarimetric images, degree of polarization

1 Introduction

Bar Lev *et al.* [1] introduced multivariate NMDs whose PGFs are defined as the inverse λth power of any affine polynomial as follows. Let us denote $[n] = \{1, \ldots, n\}$ and $\boldsymbol{z}^T = \prod_{t\in T} z_t$, where $\boldsymbol{z} = (z_1, \ldots, z_n) \in \mathbb{R}^n$ and $T \subset [n]$. Let $P_n(\boldsymbol{z}) = \sum_{T\subset[n],T\neq\varnothing} p_T \boldsymbol{z}^T$ be an affine polynomial, with respect to the n variables $(z_1, \ldots, z_n)$ such as $1-P_n(\mathbf{1}) \neq 0$. The NMD $\mathcal{NM}(n, P_n)$ associated to (n, P_n) is defined by its PGF equal to $\mathcal{G}_{\mathcal{NM}(n,P_n)}(\boldsymbol{z}) = (1 - P_n(\boldsymbol{z}))^{-\lambda} (1 - P_n(\mathbf{1}))^{\lambda}$. These very general multivariate NMDs were recently used for image processing applications in [3].

For a valid NMD, the corresponding expression of the coefficient of $\boldsymbol{z}^{\boldsymbol{\alpha}}$ in the Taylor expansion of $(1 - P_n(\boldsymbol{z}))^{-\lambda}$ is given by the formula (see [2])

$$c_\alpha(\lambda, P_n) = \sum_{k\in K_\alpha} (\lambda)_{|k|} \frac{\boldsymbol{p}^k}{k!}, \tag{1}$$

where $K_\alpha = \{k : \mathcal{P}_n \to \mathbb{N}\}$, and $\mathcal{P}_n$ is the set of all subsets of $[n]$.

The first part of this paper derives a way of computing the masses of multivariate NMDs $\mathcal{NM}(n, P_n)$ defined above. The second part of the paper is devoted to the application of NMDs to image processing, more specifically to polarimetric image processing.

2 Negative Multinomial Distributions

An n-variate NMD is the distribution of a random vector $\boldsymbol{N} = (N_1, ..., N_n)$ taking its values in $\mathbb{N}_0^n$ whose PGF is

$$G_{\boldsymbol{N}}(\boldsymbol{z}) = \mathbb{E}\left(\prod_{k=1}^{n} z_k^{N_k}\right) = [P_n(\boldsymbol{z})]^{-\lambda} \tag{2}$$

where $\mathbb{E}$ denotes the mathematical expectation, $\boldsymbol{z} = (z_1, ..., z_n)$, $\lambda > 0$ and $P_n(\boldsymbol{z})$ is an affine polynomial of order n[1]. The affine polynomial P_n has to satisfy appropriate conditions to ensure that $G_{\boldsymbol{N}}(\boldsymbol{z})$ is a PGF (see [2]). These conditions include the equality $P_n(1, ..., 1) = 1$. As explained in [2], the affine polynomial $P_n(\boldsymbol{z})$ can be rewritten $P_n(\boldsymbol{z}) = A_n(a_1 z_1, ..., a_n z_n)/A_n(a_1, ..., a_n)$ where $a_1, ..., a_n$ are positive numbers and A_n is an affine polynomial such that $A_n(0, ..., 0) = 1$. The Taylor expansions of $[A_n(\boldsymbol{z})]^{-\lambda}$ and $[P_n(\boldsymbol{z})]^{-\lambda}$ in the neighborhood of $(0, ..., 0)$ will be denoted as follows

$$[A_n(\boldsymbol{z})]^{-\lambda} = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^n} c_{\boldsymbol{\alpha}}(\lambda, A_n) \boldsymbol{z}^{\boldsymbol{\alpha}}, \quad [P_n(\boldsymbol{z})]^{-\lambda} = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^n} c_{\boldsymbol{\alpha}}(\lambda, P_n) \boldsymbol{z}^{\boldsymbol{\alpha}} \tag{3}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$ and $\boldsymbol{z}^{\boldsymbol{\alpha}} = \prod_{i=1}^n z_i^{\alpha_i}$. The masses of multivariate NMDs are $c_{\boldsymbol{\alpha}}(\lambda, P_n) = c_{\boldsymbol{\alpha}}(\lambda, A_n) A_n(a_1, ..., a_n)^\lambda \prod_{i=1}^n a_i^{\alpha_i}$.

3 Masses of Negative Multinomial Distributions

Before computing the $c_{\boldsymbol{\alpha}}(\lambda, A_n)$, we need the following result

Theorem 1 *Denote as $\mathfrak{P}_n^*$ the set of non empty subsets of $[n] = \{1, ..., n\}$. Any affine polynomial A_n such that $A_n(\boldsymbol{0}) = 1$ denoted as $A_n(\boldsymbol{z}) = 1 - \sum_{T \in \mathfrak{P}_n^*} a_T \boldsymbol{z}^T$. Moreover*

$$A_n(\boldsymbol{z}) = \left[\prod_{i \in [n]} (1 - a_i z_i)\right] \left(1 - Q_n\left(\frac{z_1}{1 - a_1 z_1}, \ldots, \frac{z_n}{1 - a_n z_n}\right)\right) \tag{4}$$

where $|T|$ is the cardinal of the set T, and Q_n is the polynomial defined by $Q_n(\boldsymbol{z}) = \sum_{T \in \mathfrak{P}_n^*, |T| \geqslant 2} d_T^n \boldsymbol{z}^T$ and d_T^n is related to the $2^{|T|} - 1$ variables a_S, $S \in \mathfrak{P}_T^*$ as follows $d_T^n = \sum_{T \in \mathfrak{P}_n, |T| > 1}^{|T|} a_T a^{[n] \setminus T} + (|T| - 1) \prod_{i \in T} a_i$.

[1] A polynomial $P_n(\boldsymbol{z})$ with respect to $\boldsymbol{z} = (z_1, \ldots, z_n)$ is affine if the one variable polynomial $z_j \mapsto P_n(\boldsymbol{z})$ can be written $A^{(-j)} z_j + B^{(-j)}$ (for any $j = 1, \ldots, d$), where $A^{(-j)}$ and $B^{(-j)}$ are polynomials with respect to the z_i's with $i \neq j$.

Theorem 2 *Let* $A_n(\boldsymbol{z}) = 1 - \sum_{T \in \mathfrak{P}_n^*} a_T z^T$, $\mathbf{a} = (a_1, \ldots, a_n)$ *and* Q_n *the affine polynomial defined in Theorem 1. For any* α *and* γ *in* $\mathbb{N}^n$, *denote as* $c_\gamma(\lambda, A_n)$ *the coefficient of* $\boldsymbol{z}^\gamma$ *in the Taylor expansion of* $[A_n(\boldsymbol{z})]^{-\lambda}$ *and* $c_\alpha(\lambda, 1 - Q_n)$ *the coefficient of* $\boldsymbol{z}^\alpha$ *in the Taylor expansion* $[1 - Q_n(\boldsymbol{z})]^{-\lambda}$. *The following relation can be obtained*

$$c_\gamma(\lambda, A_n) = \sum_{\alpha+\beta=\gamma} c_\alpha(\lambda, 1 - Q_n)(\lambda\mathbf{1} + \alpha)_\beta \frac{\mathbf{a}^\beta}{\beta!} \tag{5}$$

$$= \sum_{0 \leqslant \beta_i \leqslant \gamma_i,\ i=1,\ldots,n} c_{\gamma-\beta}(\lambda, 1 - Q_n) \prod_{i=1}^{n} (\lambda + \gamma_i - \beta_i)_{\beta_i} \frac{a_i^{\beta_i}}{\beta_i!} \tag{6}$$

$$= \sum_{0 \leqslant \alpha_i \leqslant \gamma_i,\ i=1,\ldots,n} c_\alpha(\lambda, 1 - Q_n) \prod_{i=1}^{n} (\lambda + \alpha_i)_{\beta_i} \frac{a_i^{\gamma_i - \alpha_i}}{(\gamma_i - \alpha_i)!} \tag{7}$$

3.1 Bivariate NMDs

Theorem 3 *The coefficient of* $\boldsymbol{z}^\gamma$ *in the Taylor expansion of* $[A_2(\boldsymbol{z})]^{-\lambda}$, *can be computed as follows*

$$c_\gamma(\lambda, P_2) = (\lambda)_{\max(\gamma_1,\gamma_2)} \sum_{\ell=0}^{\min(\gamma_1,\gamma_2)} \frac{(\lambda + \ell)_{\min(\gamma_1,\gamma_2)-\ell}}{(\gamma_1 - \ell)!\,(\gamma_2 - \ell)!\,\ell!} a_1^{\gamma_1-\ell} a_2^{\gamma_2-\ell} b_{1,2}^{\ell}. \tag{8}$$

3.2 Trivariate NMDs

Theorem 4 *The coefficient of* $\boldsymbol{z}^\gamma$ *in the Taylor expansion of* $[A_3(\boldsymbol{z})]^{-\lambda}$ *can be expressed as follows*

$$c_\gamma(\lambda, P_3) = \sum_{\beta_1=0}^{\gamma_1} \sum_{\beta_2=0}^{\gamma_2} \sum_{\beta_3=0}^{\gamma_3} \sum_{v=\|\gamma-\beta\|}^{\left\lfloor \frac{|\gamma-\beta|}{2} \right\rfloor} (\lambda)_v \frac{b_{2,3}^{v-\gamma_1+\beta_1} b_{1,3}^{v-\gamma_2+\beta_2} b_{1,2}^{v-\gamma_3+\beta_3}}{\prod_{i=1}^{3} (v - \gamma_i + \beta_i)!} \frac{b_{1,2,3}^{|\gamma-\beta|-2v}}{(|\gamma-\beta| - 2v)!}$$
$$\times \frac{(\lambda + \gamma_1 - \beta_1)_{\beta_1}}{\beta_1!} \frac{(\lambda + \gamma_2 - \beta_2)_{\beta_2}}{\beta_2!} \frac{(\lambda + \gamma_3 - \beta_3)_{\beta_3}}{\beta_3!} a_1^{\beta_1} a_2^{\beta_2} a_3^{\beta_3}. \tag{9}$$

4 Estimating the polarization degree of low-flux polarimetric images using maximum likelihood methods

4.1 Low-flux polarimetric images

The state of polarization of the light can be described by the random behavior of a complex vector $\boldsymbol{A} = (A_X, A_Y)$, called the Jones vector, whose covariance matrix, called the polarization matrix, is

$$\Gamma = \begin{pmatrix} \mathbb{E}(A_X A_X^*) & \mathbb{E}(A_X A_Y^*) \\ \mathbb{E}(A_Y A_X^*) & \mathbb{E}(A_Y A_Y^*) \end{pmatrix} \triangleq \begin{pmatrix} a_1 & a_3 + ia_4 \\ a_3 - ia_4 & a_2 \end{pmatrix}$$

where * denotes the complex conjugate. The covariance matrix Γ is a non negative hermitic matrix whose diagonal terms are the intensity components in the X and Y directions. The cross terms of Γ are the correlation between the Jones components. If we assume a fully developed speckle, the Jones vector $\boldsymbol{A}$ is distributed according to a complex Gaussian distribution with probability density function (pdf) (see [5]) $p(\boldsymbol{A}) = 1/(\pi^2|\Gamma|)\exp\left(-\boldsymbol{A}^\dagger\Gamma^{-1}\boldsymbol{A}\right)$ where $|\Gamma|$ is the determinant of the matrix Γ and † denotes the conjugate transpose operator. As a consequence, the statistical properties of $\boldsymbol{A}$ are fully characterized by the covariance matrix Γ. The different components of Γ can be classically estimated by using four intensity images that are related to the components of the Jones vector as follows (see [3], for more details)

$$
\begin{aligned}
&I_1 = |A_X|^2\,,\ I_2 = |A_Y|^2\,,\ I_3 = \frac{1}{2}|A_X|^2 + \frac{1}{2}|A_Y|^2 + \mathrm{Re}\left(A_X A_X^*\right),\\
&I_4 = \frac{1}{2}|A_X|^2 + \frac{1}{2}|A_Y|^2 + \mathrm{Im}\left(A_X A_Y^*\right).
\end{aligned}
$$

The state of polarization of the light is classically characterized by the square DoP defined by ([5])

$$
P^2 = 1 - 4\frac{|\Gamma|}{[\mathrm{trace}(\Gamma)]^2} = 1 - \frac{4\left[a_1a_2 - (a_3^2 + a_4^2)\right]}{(a_1+a_2)^2} \tag{10}
$$

where $\mathrm{trace}(\Gamma)$ is the trace of the matrix Γ. The light is totally depolarized for $P = 0$, totally polarized for $P = 1$ and partially polarized when $P \in]0,1[$. Different estimation methods of P^2 using several combinations of intensity images were studied in [3]. Since only one realization of the random vector $\boldsymbol{I} = (I_1, ..., I_4)^T$ was available for a given pixel of a polarimetric image, the image was supposed to be locally stationary and ergodic. These assumptions were used to derive square DoP estimators using several neighbor pixels belonging to a so-called estimation window.

This section considers practical applications where the intensity level of the reflected light is very low (low-flux assumption), which leads to an additional source of fluctuations on the detected signal. Under the low-flux assumption, the quantum nature of the light leads to a Poisson-distributed noise which can become very important relatively to the mean value of the signal at low photon level. As a consequence, the observed pixels of the low-flux polarimetric image are discrete random variables contained in the vector $\boldsymbol{N} = (N_1, ..., N_4)$ such that the conditional distributions of the random variables $N_l|I_l$, for $l = 1, ..., 4$ are independent and distributed according to Poisson distributions with means I_l, for $l = 1, ..., 4$. The resulting joint distribution of $\boldsymbol{N}$ is a multivariate mixed Poisson distribution (see [4]) $P\left(N = \boldsymbol{k}\right) = \int\cdots\int_{(\mathbb{R}^+)^4}\prod_{l=1}^4 \frac{I_l^{k_l}}{k_l!}\exp\left(-I_l\right) f\left(\boldsymbol{I}\right) d\boldsymbol{I}$, where $\boldsymbol{k} = (k_1, ..., k_4)$, $k_i \in \mathbb{N}$ and $f(\boldsymbol{I})$ is the joint pdf of the intensity vector. This section studies estimators of the square DoP P^2 defined in (10) based on several vectors $\boldsymbol{N}^1, \ldots, \boldsymbol{N}^n$ belonging to the estimation window.

The joint distribution of the intensity vector $\boldsymbol{I}$ is known to be a multivariate gamma distribution whose Laplace transform is (see [3]) $\mathbb{E}\left[\exp\left(\sum_{k=1}^4 z_k I_k\right)\right] = 1/P_4(\boldsymbol{z})$ where the affine polynomial P_4 is $P_4(\boldsymbol{z}) = 1 + \boldsymbol{z}\boldsymbol{\mu} + k_a[2z_1z_2 + z_3z_4 + (z_1+z_2)(z_3+z_4)]$ with $\boldsymbol{z} = (z_1, ..., z_4)$ and $k_a = \frac{1}{2}\left(a_1a_2 - a_3^2 - a_4^2\right)$, $\boldsymbol{\mu} = \left(a_1, a_2, a_3 + (a_1+a_2)/2, a_4 + (a_1+a_2)/2\right)^T$. As a consequence, the distribution of $\boldsymbol{N}$ is an NMD whose PGF can be written

$$
G_{\boldsymbol{N}}(\boldsymbol{z}) = \left[P_4(z_1-1, z_2-1, z_3-1, z_4-1)\right]^{-1}. \tag{11}
$$

The results of Section 3 allow us to compute the masses of $\boldsymbol{N}$ that will be useful for studying the maximum likelihood estimator (MLE) of the square DoP.

4.2 MLE using three polarimetric images

The PGF of $\widetilde{\boldsymbol{N}} = (N_1, N_2, N_3)$ can be computed from (11) by setting $z_4 = 1$. The following result can be obtained

$$G_{\widetilde{\boldsymbol{N}}}(\boldsymbol{z}) = [P_3(\boldsymbol{z})]^{-1} \tag{12}$$

with $P_3(\boldsymbol{z}) = P_3(\boldsymbol{0}) + z_1(\mu_1 - 3k_a) + z_2(\mu_2 - 3k_a) + z_3(\mu_3 - 2k_a) + k_a(2z_1z_2 + z_1z_3 + z_2z_3)$, $\boldsymbol{z} = (z_1, z_2, z_3)$ and $P_3(\boldsymbol{0}) = 1 - \sum_{i=1}^{3} \mu_i + 4k_a$. The results of Section 3.2 can then be used to express the masses of $\widetilde{\boldsymbol{N}}$ as a function of $\boldsymbol{\theta} = (a_1, a_2, a_3, a_4^2)^T$.

The ML estimator of $\boldsymbol{\theta}$ based on several vectors $\widetilde{\boldsymbol{N}}^k$ belonging to the estimation window (where $k = 1, ..., K$ and K is the number of pixels of the observation window) is obtained by maximizing the log-likelihood $l_3\left(\widetilde{\boldsymbol{N}}^1, ..., \widetilde{\boldsymbol{N}}^K | \boldsymbol{\theta}\right) = \sum_{k=1}^{K} \log\left[P\left(\widetilde{\boldsymbol{N}}^k\right)\right]$ with respect to $\boldsymbol{\theta}$. The practical determination of the ML estimator of $\boldsymbol{\theta}$ is achieved by using a Newton-Raphson procedure. The ML estimators of the vector $\boldsymbol{\theta}$ elements, denoted as $\widetilde{\boldsymbol{\theta}} = \left(\widetilde{a}_1, \widetilde{a}_2, \widetilde{a}_3, \widetilde{a_4^2}\right)^T$, are then plugged into (10) to provide the ML estimator of the square DoP based on three polarimetric images

$$\widetilde{P}^2 = 1 - \frac{4\left[\widetilde{a}_1\widetilde{a}_2 - (\widetilde{a}_3^2 + \widetilde{a_4^2})\right]}{(\widetilde{a}_1 + \widetilde{a}_2)^2}. \tag{13}$$

4.3 MLE using two polarimetric images

The PGF of $\underline{\boldsymbol{N}} = (N_1, N_2)$ can be computed from (12) by setting $z_3 = 1$. The following result can be obtained

$$G_{\underline{\boldsymbol{N}}}(\boldsymbol{z}) = [P_2(\boldsymbol{z})]^{-1} \tag{14}$$

with $P_2(\boldsymbol{z}) = P_2(\boldsymbol{0}) + z_1(\mu_1 - 2k_a) + z_2(\mu_2 - 2k_a) + 2k_a z_1 z_2$, $\boldsymbol{z} = (z_1, z_2)$ and $P_2(\boldsymbol{0}) = 1 - \sum_{i=1}^{2} \mu_i + 2k_a$. The results of Section 3.1 can then be used to express the masses of $\underline{\boldsymbol{N}}$ as a function of $\boldsymbol{\theta} = (a_1, a_2, k_a)^T$.

The ML estimator of $\boldsymbol{\theta}$ based on several vectors $\underline{\boldsymbol{N}}^k$ belonging to the estimation window is obtained by maximizing the log-likelihood $l_2\left(\underline{\boldsymbol{N}}^1, ..., \underline{\boldsymbol{N}}^P | \boldsymbol{\theta}\right) = \sum_{k=1}^{K} \log\left[P\left(\underline{\boldsymbol{N}}^k\right)\right]$ with respect to $\boldsymbol{\theta}$. The practical determination of the ML estimator of $\boldsymbol{\theta}$ is achieved by using a Newton-Raphson procedure. The ML estimators of the vector $\boldsymbol{\theta}$ elements, denoted as $\underline{\boldsymbol{\theta}} = \left(\underline{a}_1, \underline{a}_2, \underline{k_a}\right)^T$, are then plugged into (10) to provide the ML estimator of the square DoP based on two polarimetric images

$$\underline{P}^2 = 1 - \frac{8\underline{k_a}}{(\underline{a}_1 + \underline{a}_2)^2}. \tag{15}$$

5 Simulations results

Simulations results showing the estimation performance and illustrating the application to polarimetric imagery will be presented during the conference.

Acknowledgements
The authors would like to thank Gérard Letac for fruitful discussions regarding multivariate gamma distributions and negative multinomial distributions.

Références

[1] S. K. Bar-Lev, D. Bshouty, P. Enis, G. Letac, I. Li-Lu, and D. Richards. The diagonal multivariate natural exponential families and their classification. *Journal of Theoretical Probability*, 7(4) :883–929, October 1994.

[2] Philippe Bernardoff. Which negative multinomial distributions are infinitely divisible ? *Bernoulli*, 9(5) :877–893, 2003.

[3] F. Chatelain, J.-Y. Tourneret, M. Roche, and M. Alouini. Estimating the polarization degree of polarimetric images in coherent illumination using maximum likelihood methods. *JOSA A*, 26(6) :1348–1359, June 2009.

[4] A. Ferrari, G. Letac, and J.-Y. Tourneret. Multivariate mixed Poisson distributions. In F. Hlawatsch, G. Matz, M. Rupp, and B. Wistawel, editors, *EUSIPCO-04*, pages 1067–1070, Vienna, Austria, September 2004.

[5] J. Goodman. *Statistical Optics*. Wiley, New York, 1985.