

Multiplying a Gaussian Matrix by a Gaussian Vector

Pierre-Alexandre Mattei

► **To cite this version:**

Pierre-Alexandre Mattei. Multiplying a Gaussian Matrix by a Gaussian Vector. *Statistics and Probability Letters*, Elsevier, 2017, <10.1016/j.spl.2017.04.004>. <hal-01462941v2>

HAL Id: hal-01462941

<https://hal.archives-ouvertes.fr/hal-01462941v2>

Submitted on 5 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MULTIPLYING A GAUSSIAN MATRIX BY A GAUSSIAN VECTOR

Pierre-Alexandre Mattei

pierre-alexandre.mattei@parisdescartes.fr

Laboratoire MAP5, UMR CNRS 8145

Université Paris Descartes & Sorbonne Paris Cité

45 rue des Saints-Pères, 75006 Paris, France

Abstract

We provide a new and simple characterization of the multivariate generalized Laplace distribution. In particular, this result implies that the product of a Gaussian matrix with independent and identically distributed columns by an independent isotropic Gaussian vector follows a *symmetric* multivariate generalized Laplace distribution.

Keywords: Gamma distribution; Laplace distribution; Product distribution; Random matrix; Scale mixture; Variance-gamma distribution

1. Introduction

Wishart and Bartlett (1932) proved that the inner product of two independent bidimensional standard gaussian vectors follows a Laplace distribution. This result is deeply linked to the fact that the Laplace distribution can be represented as an infinite scale mixture of Gaussians with gamma mixing distribution. Specifically, if σ^2 follows a $\text{Gamma}(1, 1/2)$ distribution and $x|\sigma \sim \mathcal{N}(0, \sigma^2)$, then x follows a standard Laplace distribution¹. This representation – which was recently named the *Gauss-Laplace representation* by Ding and Blitzstein (2017) following a blog post by Christian P. Robert² – is particularly useful if one wants to simulate a Laplace random variable: its use constitutes for example the cornerstone of the Gibbs sampling scheme for the Bayesian lasso of Park and Casella (2008).

In this short paper, we are interested in studying links between multivariate counterparts of these two characterizations. More specifically, we give a new simple characterization of the *multivariate generalized Laplace distribution* of Kotz, Kozubowski, and Podgórski (2001). As a corollary, we show that the product of a zero-mean Gaussian matrix with independent and identically distributed (i.i.d.) columns and a zero-mean isotropic Gaussian vector follows a symmetric multivariate generalized Laplace distribution, a result that has useful applications for Bayesian principal component analysis (Bouveyron, Latouche, and Mattei, 2016, 2017).

In the remainder of this paper, p and d are two positive integers and \mathcal{S}_p^+ denotes the cone of positive semidefinite matrices of size $p \times p$.

2. The multivariate generalized Laplace distribution

While the definition of the univariate Laplace distribution is widely undisputed, there exist several different generalizations of this distribution to higher dimensions – a comprehensive review of such generalizations can be found in the monograph of Kotz, Kozubowski, and Podgórski (2001). In particular, McGraw and Wagner (1968) introduced a zero-mean el-

1. The shape-rate parametrization of the gamma distribution is used through this paper. Note also that a standard Laplace distribution is centered with variance two.

2. <https://xianblog.wordpress.com/2015/10/14/gauss-to-laplace-transmutation/>

liptically contoured bidimensional Laplace distribution with univariate Laplace marginals. This distribution was later generalized to the p -dimensional setting by Anderson (1992), considering characteristic functions of the form

$$\forall \mathbf{u} \in \mathbb{R}^p, \phi(\mathbf{u}) = \frac{1}{1 + \frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u}},$$

where $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$. This distribution was notably promoted by Eltoft, Kim, and Lee (2006) and is arguably the most popular multivariate generalization of the Laplace distribution (Kotz, Kozubowski, and Podgórski, 2001, p. 229). Among its advantages, this distribution can be slightly generalized to model skewness, by building on characteristic functions of the form

$$\forall \mathbf{u} \in \mathbb{R}^p, \phi(\mathbf{u}) = \frac{1}{1 + \frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u} - i\boldsymbol{\mu}^T \mathbf{u}},$$

where $\boldsymbol{\mu} \in \mathbb{R}^p$ accounts for asymmetry. Similarly to the univariate Laplace distribution, this asymmetric multivariate generalization is infinitely divisible (Kotz, Kozubowski, and Podgórski, 2001, p. 256). Therefore, it can be associated with a specific Lévy process (Kyprianou, 2014, p. 5), whose increments will follow yet another generalization of the Laplace distribution, the *multivariate generalized asymmetric Laplace distribution*. This distribution, introduced by Kotz, Kozubowski, and Podgórski (2001, p. 257) and further studied by Kozubowski, Podgórski, and Rychlik (2013), will be the cornerstone of our analysis of multivariate characterizations of Laplace and Gaussian distributions.

Definition 1 *A random variable $\mathbf{z} \in \mathbb{R}^p$ is said to have a **multivariate generalized asymmetric Laplace distribution** with parameters $s > 0, \boldsymbol{\mu} \in \mathbb{R}^p$ and $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$ if its characteristic function is*

$$\forall \mathbf{u} \in \mathbb{R}^p, \phi_{\text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s)}(\mathbf{u}) = \left(\frac{1}{1 + \frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u} - i\boldsymbol{\mu}^T \mathbf{u}} \right)^s.$$

It is denoted by $\mathbf{z} \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s)$.

General properties of the generalized asymmetric Laplace distribution are discussed by Kozubowski, Podgórski, and Rychlik (2013). We list here a few useful ones.

Proposition 2 *Let $s > 0, \boldsymbol{\mu} \in \mathbb{R}^p$ and $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$. If $\mathbf{z} \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s)$, we have $\mathbb{E}(\mathbf{z}) = s\boldsymbol{\mu}$ and $\text{Cov}(\mathbf{z}) = s(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T)$. Moreover, if $\boldsymbol{\Sigma}$ is positive definite, the density of \mathbf{z} is given by*

$$\forall \mathbf{x} \in \mathbb{R}^p, f_{\mathbf{z}}(\mathbf{x}) = \frac{2e^{\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}}}{(2\pi)^{p/2} \Gamma(s) \sqrt{\det \boldsymbol{\Sigma}}} \left(\frac{Q_{\boldsymbol{\Sigma}}(\mathbf{x})}{C(\boldsymbol{\Sigma}, \boldsymbol{\mu})} \right)^{s-p/2} K_{s-p/2}(Q_{\boldsymbol{\Sigma}}(\mathbf{x}) C(\boldsymbol{\Sigma}, \boldsymbol{\mu})),$$

where $Q_{\boldsymbol{\Sigma}}(\mathbf{x}) = \sqrt{\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}}$, $C(\boldsymbol{\Sigma}, \boldsymbol{\mu}) = \sqrt{2 + \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}$ and $K_{s-p/2}$ is the modified Bessel function of the second kind of order $s - p/2$.

Note that the $\text{GAL}_1(2b^2, 0, 1)$ case corresponds to a centered univariate Laplace distribution with scale parameter $b > 0$. In the symmetric case ($\boldsymbol{\mu} = 0$) and when $s = 1$, we recover the multivariate generalization of the Laplace distribution of Anderson (1992).

An appealing property of the multivariate generalized Laplace distribution is that it is also endowed with a multivariate counterpart of the Gauss-Laplace representation.

Theorem 3 (Generalized Gauss-Laplace representation) Let $s > 0$, $\boldsymbol{\mu} \in \mathbb{R}^p$ and $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$. If $u \sim \text{Gamma}(s, 1)$ and $\mathbf{x} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ is independent of u , we have

$$u\boldsymbol{\mu} + \sqrt{u}\mathbf{x} \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s). \quad (1)$$

A proof of this result can be found in Kotz, Kozubowski, and Podgórski (2001, chap. 6). This representation explains why the multivariate generalized Laplace distribution can also be seen as a multivariate generalization of the *variance-gamma distribution* which is widely used in the field of quantitative finance (Madan, Carr, and Chang, 1998). Infinite mixtures similar to (1) are called *variance-mean mixtures* (Barndorff-Nielsen, Kent, and Sørensen, 1982) and are discussed for example by Yu (2017).

Another useful property of the multivariate generalized Laplace distribution is that, under some conditions, it is closed under convolution.

Proposition 4 Let $s_1, s_2 > 0$, $\boldsymbol{\mu} \in \mathbb{R}^p$ and $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$. If $\mathbf{z}_1 \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_1)$ and $\mathbf{z}_2 \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_2)$ are independent random variables, then

$$\mathbf{z}_1 + \mathbf{z}_2 \sim \text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_1 + s_2). \quad (2)$$

Proof Since \mathbf{z}_1 and \mathbf{z}_2 are independent, we have for all $\mathbf{u} \in \mathbb{R}^p$,

$$\phi_{\mathbf{z}_1 + \mathbf{z}_2}(\mathbf{u}) = \phi_{\text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_1)}(\mathbf{u})\phi_{\text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_2)}(\mathbf{u}) = \left(\frac{1}{1 + \frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u} - i\boldsymbol{\mu}^T \mathbf{u}} \right)^{s_1 + s_2}$$

which is the characteristic function of the $\text{GAL}_p(\boldsymbol{\Sigma}, \boldsymbol{\mu}, s_1 + s_2)$ distribution. ■

3. A new characterization involving a product between a Gaussian matrix and a Gaussian vector

We now state our main theorem, which gives a new characterization of multivariate generalized Laplace distributions with half-integer shape parameters.

Theorem 5 Let \mathbf{W} be a $p \times d$ random matrix with i.i.d. columns following a $\mathcal{N}(0, \boldsymbol{\Sigma})$ distribution, $\mathbf{y} \sim \mathcal{N}(0, \mathbf{I}_d)$ be a Gaussian vector independent from \mathbf{W} and let $\boldsymbol{\mu} \in \mathbb{R}^p$. We have

$$\mathbf{W}\mathbf{y} + \|\mathbf{y}\|_2^2 \boldsymbol{\mu} \sim \text{GAL}_p(2\boldsymbol{\Sigma}, 2\boldsymbol{\mu}, d/2). \quad (3)$$

Proof For each $k \in \{1, \dots, d\}$ let \mathbf{w}_k be the k -th column of \mathbf{W} , $u_k = y_k^2$ and $\boldsymbol{\xi}_k = y_k \mathbf{w}_k + y_k^2 \boldsymbol{\mu}$. To prove the theorem, we will prove that $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_d$ follow a GAL distribution and use the decomposition

$$\mathbf{W}\mathbf{y} + \|\mathbf{y}\|_2^2 \boldsymbol{\mu} = \sum_{k=1}^d \boldsymbol{\xi}_k.$$

Let $k \in \{1, \dots, d\}$. Since \mathbf{y} is standard Gaussian, $u_k = y_k^2$ follows a $\chi^2(1)$ distribution, or equivalently a $\text{Gamma}(1/2, 1/2)$ distribution. Therefore, $u_k/2 \sim \text{Gamma}(1/2, 1)$.

Moreover, note that $\sqrt{u_k}\mathbf{w}_k = |y_k|\mathbf{w}_k = y_k\text{sign}(y_k)\mathbf{w}_k \stackrel{d}{=} y_k\mathbf{w}_k$ since $|y_k|$ and $\text{sign}(y_k)$ are independent and $\text{sign}(y_k)\mathbf{w}_k \stackrel{d}{=} \mathbf{w}_k$. Therefore, according to the generalized Gauss-Laplace representation, we have

$$\boldsymbol{\xi}_k \stackrel{d}{=} \sqrt{\frac{u_k}{2}}\sqrt{2}\mathbf{w}_k + \frac{u_k}{2}\mathbf{2}\boldsymbol{\mu} \sim \text{GAL}_p(2\boldsymbol{\Sigma}, 2\boldsymbol{\mu}, 1/2).$$

Since $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_d$ are i.i.d. and following a $\text{GAL}_p(2\boldsymbol{\Sigma}, 2\boldsymbol{\mu}, 1/2)$ distribution, we can use Proposition 4 to conclude that

$$\mathbf{W}\mathbf{y} + \|\mathbf{y}\|_2^2\boldsymbol{\mu} = \sum_{k=1}^d \boldsymbol{\xi}_k \sim \text{GAL}_p(2\boldsymbol{\Sigma}, 2\boldsymbol{\mu}, d/2).$$

■

In the symmetric case ($\boldsymbol{\mu} = 0$), this result gives the distribution of the product between a Gaussian matrix with i.i.d. columns and a isotropic Gaussian vector.

Corollary 6 *Let \mathbf{W} be a $p \times d$ random matrix with i.i.d. columns following a $\mathcal{N}(0, \boldsymbol{\Sigma})$ distribution and let $\mathbf{y} \sim \mathcal{N}(0, \alpha\mathbf{I}_d)$ be a Gaussian vector independent from \mathbf{W} . Then*

$$\mathbf{W}\mathbf{y} \sim \text{GAL}_p(2\alpha\boldsymbol{\Sigma}, 0, d/2). \quad (4)$$

Moreover, if u is a standard Gamma variable with shape parameter $d/2$ and if $\mathbf{x} \sim \mathcal{N}(0, 2\alpha\boldsymbol{\Sigma})$ is a Gaussian vector independent of u , then

$$\mathbf{W}\mathbf{y} \stackrel{d}{=} \sqrt{u}\mathbf{x}. \quad (5)$$

Less general versions of Theorem 5 have been proven in the past, dating back to the derivation of the inner product of two i.i.d. standard Gaussian vectors by Wishart and Bartlett (1932). In particular, the unidimensional case ($p = 1$) was recently proven by Gaunt (2014) in order to obtain bounds for the convergence rate of random sums involving Gaussian products. The multivariate symmetric isotropic case ($\boldsymbol{\mu} = 0$ and $\boldsymbol{\Sigma}$ proportional to \mathbf{I}_p) was proven by Bouveyron, Latouche, and Mattei (2016) in order to derive the marginal likelihood of the noiseless probabilistic principal component analysis model of Roweis (1998). While the proof of Bouveyron, Latouche, and Mattei (2016) relied on characteristic functions and the properties of Bessel functions, the proof that we presented here is closer in spirit to the one of Gaunt (2014), based on representations of distributions.

4. Perspectives

The new characterization presented in this paper may notably prove useful in two contexts.

First, it indicates a new way of handling situations involving the product of a Gaussian matrix and a Gaussian vector. An important instance is the Bayesian factor analysis model (Lopes and West, 2004), of which principal component analysis is a particular case.

In this framework, the marginal distribution of the data, which is essential for model selection purposes, can be derived using representation (5) together with the Gauss-Laplace representation (Bouveyron, Latouche, and Mattei, 2016, 2017).

Moreover, our characterization offers a means to get around problems encountered when dealing with distributions related to the GAL distribution. For example, representation (3) might lead to alternative estimation strategies for some problems related to portfolio allocation (Mencía and Sentana, 2009; Breyman and Lüthi, 2013) or cluster analysis (McNicholas, McNicholas, and Browne, 2013; Franczak, Browne, and McNicholas, 2014).

Acknowledgements

I thank Charles Bouveyron, Pierre Latouche, Brendan Murphy and Christian Robert for fruitful advices and discussions. Part of this work was made during a visit to University College Dublin, funded by the Fondation Sciences Mathématiques de Paris (FSMP).

References

- D. N. Anderson. A multivariate Linnik distribution. *Statistics & Probability Letters*, 14(4): 333–336, 1992.
- O. Barndorff-Nielsen, J. Kent, and M. Sørensen. Normal variance-mean mixtures and z distributions. *International Statistical Review/Revue Internationale de Statistique*, 50(2): 145–159, 1982.
- C. Bouveyron, P. Latouche, and P.-A. Mattei. Bayesian variable selection for globally sparse probabilistic PCA. *Technical report, HAL-01310409, Université Paris Descartes*, 2016.
- C. Bouveyron, P. Latouche, and P.-A. Mattei. Exact dimensionality selection for Bayesian PCA. *Technical report, HAL-01484099, Université Paris Descartes*, 2017.
- W. Breyman and D. Lüthi. `ghyp`: A package on generalized hyperbolic distributions. *Manual for R Package ghyp*, 2013.
- P. Ding and J. K. Blitzstein. On the Gaussian mixture representation of the Laplace distribution. *The American Statistician*, in press, 2017.
- T. Eltoft, T. Kim, and T.-W. Lee. On the multivariate Laplace distribution. *IEEE Signal Processing Letters*, 13(5):300–303, 2006.
- B. C. Franczak, R. P. Browne, and P. D. McNicholas. Mixtures of shifted asymmetric Laplace distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(6):1149–1157, 2014.
- R. E. Gaunt. Variance-Gamma approximation via Stein’s method. *Electronic Journal of Probability*, 19(38):1–33, 2014.
- S. Kotz, T. Kozubowski, and K. Podgórski. *The Laplace distribution and generalizations: a revisit with applications to communications, exconomics, engineering, and finance*. Springer Science & Business Media, 2001.

- T. Kozubowski, K. Podgórski, and I. Rychlik. Multivariate generalized Laplace distribution and related random fields. *Journal of Multivariate Analysis*, 113:59–72, 2013.
- A. Kyprianou. *Fluctuations of Lévy processes with applications: Introductory Lectures*. Springer Science & Business Media, 2014.
- H. F. Lopes and M. West. Bayesian model assessment in factor analysis. *Statistica Sinica*, pages 41–67, 2004.
- D. B. Madan, P. P. Carr, and E. C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.
- D. McGraw and J. Wagner. Elliptically symmetric distributions. *IEEE Transactions on Information Theory*, 14(1):110–120, 1968.
- S. M. McNicholas, P. D. McNicholas, and R. P. Browne. Mixtures of variance-gamma distributions. *arXiv preprint arXiv:1309.2695*, 2013.
- J. Mencía and E. Sentana. Multivariate location–scale mixtures of normals and mean–variance–skewness portfolio allocation. *Journal of Econometrics*, 153(2):105–121, 2009.
- T. Park and G. Casella. The Bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.
- S. Roweis. EM algorithms for PCA and SPCA. *Advances in Neural Information Processing Systems*, 10:626–632, 1998.
- J. Wishart and M. S. Bartlett. The distribution of second order moment statistics in a normal system. *Mathematical Proceedings of the Cambridge Philosophical Society*, 28(4):455–459, 1932.
- Y. Yu. On normal variance–mean mixtures. *Statistics & Probability Letters*, 121:45–50, 2017.