Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Multivariate Analysis 97 (2006) 179 – 197

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Bayesian inference in spherical linear models: robustness and conjugate analysis

R.B. Arellano-Valle, G. del Pino, P. Iglesias

Facultad de Matematicas, Pontificia Universidad Católica de Chile, Casilla 306 Santiago, Chile

Received 10 July 2003
Available online 29 January 2005

Abstract

The early work of Zellner on the multivariate Student-t linear model has been extended to Bayesian inference for linear models with dependent non-normal error terms, particularly through various papers by Osiewalski, Steel and coworkers. This article provides a full Bayesian analysis for a spherical linear model. The density generator of the spherical distribution is here allowed to depend both on the precision parameter ϕ and on the regression coefficients $\boldsymbol{\beta}$. Another distinctive aspect of this paper is that proper priors for the precision parameter are discussed.

The normal-chi-squared family of prior distributions is extended to a new class, which allows the posterior analysis to be carried out analytically. On the other hand, a direct joint modelling of the data vector and of the parameters leads to conjugate distributions for the regression and the precision parameters, both individually and jointly. It is shown that some model specifications lead to Bayes estimators that do not depend on the choice of the density generator, in agreement with previous results obtained in the literature under different assumptions. Finally, the distribution theory developed to tackle the main problem is useful on its own right.

Keywords: Linear regression models; Elliptical and squared-radial distributions; Elliptical density generator; Dispersion and dispersion-location elliptical models; Bayes estimator; Conjugate families

1. Introduction and outline

The Bayesian analysis of regression models with multivariate spherical or elliptical error terms has received considerable attention, especially since the seminal paper of Zellner

E-mail address: reivalle@mat.puc.cl.

0047-259X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2004.12.002

[14], who considers the multivariate spherical Student-t linear model with v degrees of freedom, say $\boldsymbol{y}|\boldsymbol{\beta}, \phi \sim t_n(\boldsymbol{X\beta}, \phi; v)$. It is shown there that when this model is combined with an improper reference prior on the (nuisance) precision parameter ϕ, i.e. $\pi(\phi) \propto \phi^{-1}$, posterior inferences on the regression vector $\boldsymbol{\beta}$ and predictive analysis agree exactly with those obtained under the normality assumptions. These results were extended to the entire family of multivariate spherical linear models in Osiewalski and Steel [10,11]. Under similar assumptions, Arellano-Valle et al. [3] show how the posterior of ϕ can be used for detecting the presence of outliers or influential observations, in the context of elliptical linear regression models. They also provide a convenient way to derive the posterior distribution of ϕ and its moments (see also [13]). Fernández et al. [7] revisit these robustness results and showed that they are induced by the improper reference prior on ϕ, in the much wider context of all the continuous multivariate location-scale distributions with density of the form $\phi^n g(\phi(\boldsymbol{y}-\boldsymbol{\mu}))$, $(\boldsymbol{y},\ \boldsymbol{\mu} \in \mathbb{R}^n,\ \phi > 0)$, where the density g is a general function which must not depend upon $(\boldsymbol{\mu}, \phi)$.

The present paper examines further extensions of the spherical linear models in two directions. Firstly, the density generator may be related with both the precision ϕ and the regression (or location) parameters $\boldsymbol{\beta}$. Secondly, proper priors for the precision parameter are examined. Specifically, this work starts with the conventional spherical linear model

$$\begin{aligned} &\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon},\ \text{ where } \boldsymbol{\epsilon}|\phi \sim S_n(\phi; h) \\ &\boldsymbol{\beta} \perp\!\!\!\perp \boldsymbol{\epsilon}|\phi \text{ and } h \text{ does not depend on } (\boldsymbol{\beta}, \phi), \end{aligned} \tag{1.1}$$

where $S_n(\phi; h)$ denotes the n-dimensional spherical distribution with precision parameter ϕ and density generator h (see, for example, [6]). The parameters ϕ and h determine the density function through $\phi^{\frac{n}{2}} h(\phi\|\boldsymbol{\epsilon}\|^2)$, where the symbols $\|\ \|, \perp\!\!\!\perp, \perp\!\!\!\perp|$, and $\stackrel{d}{=}$ have their usual meaning (length, independence, conditional independence, and equality in distribution respectively).

Most of the Bayesian analysis of (1.1) is developed adopting the conventional product noninformative prior

$$\pi(\boldsymbol{\beta}, \phi) \propto \phi^{-1}\pi(\boldsymbol{\beta}), \tag{1.2}$$

under which Osiewalski and Steel [10] show, e.g., that the posterior of $\boldsymbol{\beta}$ does not depend on the density generator. Under (1.2), similar robustness results are discussed in more general settings by Osiewalski and Steel [12], Fernández et al. [8] and Ng [9].

The present work considers a full Bayesian analysis of an extension of (1.1), where the generator h is allowed to depend on the model parameters $(\boldsymbol{\beta}, \phi)$. The simpler case where h depends only on the precision parameter ϕ is first discussed, and this makes it possible to preserve the prior independence assumption between $(\boldsymbol{\epsilon}, \phi)$ and $\boldsymbol{\beta}$. Here a joint modelling of $(\boldsymbol{y}, \phi)$ within the spherical class is necessary, while an arbitrary prior distribution for $\boldsymbol{\beta}$ may be adopted. Next, the situation where h depends both on $\boldsymbol{\beta}$ and ϕ is tackled. It is then necessary to model jointly the distribution of $(\boldsymbol{y}, \boldsymbol{\beta}, \phi)$ within the elliptical family, without any prior independence assumptions. Moreover, in both cases an informative prior distribution for ϕ is specified using the squared-radial versions of spherical distributions (see [1]).

With the idea of extending the usual conjugate analysis for the normal linear model, the class of *dispersion elliptical squared-radial distribution* is first defined. In this setting, we obtain two types of special results: (a) robust posterior inference for $\boldsymbol{\beta}$, in the sense defined by Osiewalski and Steel (1993a), together with a conjugate posterior distribution for ϕ; and (b) joint conjugate posterior inferences for $(\boldsymbol{\beta}, \phi)$. A positive aspect is that in both situations the whole posterior can be obtained analytically, and it is straightforward to find explicit solutions for the posterior means and variances. The fact that in both of the situations considered the posterior mean $E(\boldsymbol{\beta}|\boldsymbol{y}, h)$ does not depend on the generator h, indicates that it is a *robust* Bayes estimator under quadratic loss, with respect to changes in the specification of the sampling mechanism within the wide class of elliptical sampling process.

Although the primary aim of this paper is to provide a general Bayesian analysis, this needs the development of a substantial amount of distribution theory, which has some interest on its own right. To sharpen the focus of this paper we have collected in Section 2 all the required results on elliptical and spherical distributions, as well as most of the notation.

The outline of the rest of the paper is as follows. Sections 3 and 4 provide some basic distribution theory that is needed in subsequent sections. Section 3 describes the main properties of a multivariate version of the univariate *squared-radial* (SR) *distributions* (described in Section 2), which will later be needed to prove some of the main results in other sections. Section 4 defines a *dispersion elliptical squared-radial* (DESR) *distribution*, which essentially consists of a scaled family of elliptical conditional distributions given $T = t$, where T is the squared length of a spherical random vector (which therefore has a squared-radial distribution). The scale factor takes the form $\frac{1}{bt^r}$. The main results of this section are a formula for the p.d.f. and Theorem 1, which provides an important link with the bivariate squared radial distributions, and have also a key role in the specification of the Bayesian models discussed in the following sections.

Sections 5 and 6 perform a full Bayesian analysis for two extensions of (1.1). Section 5 considers a spherical linear model, whose generator h may be depend on ϕ. Theorems 2 and 3 in Section 5.1 provide Bayesian version of the classical pivotal quantities for $(\boldsymbol{\beta}, \phi)$. In Section 5.2 the prior independence of $\boldsymbol{\beta}$ and ϕ, as well as a flat improper prior for $\boldsymbol{\beta}$ are assumed. It is shown there that a squared-radial prior for ϕ produces a posterior for ϕ in this same class. For $r = 1$, Theorem 4 shows that the posterior of $\boldsymbol{\beta}$ is a (generalized) Student-t distribution, depending neither on h nor on ϕ. For $r = 0$, Theorem 5 shows that $\boldsymbol{\beta}$ has an elliptical posterior. Section 6 proposes the direct joint modelling of $(\boldsymbol{y}, \boldsymbol{\beta}, \phi)$ as a way to achieve a conjugate distributions for $\boldsymbol{\beta}$ and ϕ, both individually and jointly. The *dispersion-location spherical squared-radial linear model* defined there, is one alternative to achieve this (the normal-gamma priors do not lead to an analytically tractable posterior, unless the spherical distribution is normal). The starting point is a DESR prior for $(\boldsymbol{\beta}, \phi)$, with the conditional distribution of the error vector $\boldsymbol{\epsilon}$ given $\boldsymbol{\beta}$ and ϕ having a special form. The main results are Theorem 6, which states the validity of yet another SR characterization, and Theorem 7, which details the posterior analysis. The proofs of all theorems are deferred to Appendix A. Appendix B contains two summary tables. Table B.1 shows the squared-radial distributions associated to various spherical distributions, while Table B.2 provides their means and variances.

2. Elliptical and spherical distributions

To establish the notation, facilitate referencing and provide some basic results employed in the proofs, this section reviews material on spherical and elliptical distributions, which can be found in standard sources, like Cambanis et al. [4], Dickey and Chen [5] or Fang et al. [6]. For notational simplicity we often write elliptical (resp. spherical) random vector for one having an elliptical (resp. spherical) distribution.

2.1. Basic definitions and notations

Denote by $\mathrm{El}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; h)$ the N-dimensional elliptical distribution with location vector $\boldsymbol{\mu}$, dispersion matrix $\boldsymbol{\Sigma}$ and generator h, whose density is

$$f(\boldsymbol{y}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} h(q(\boldsymbol{y})), \quad q(\boldsymbol{y}) = (\boldsymbol{y}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}), \tag{2.1}$$

with h satisfying

$$\int_0^\infty u^{\frac{N}{2}-1} h(cu)\, du = \frac{\Gamma(\frac{N}{2})}{(c\pi)^{\frac{N}{2}}}, \quad \forall c > 0. \tag{2.2}$$

Letting $\boldsymbol{I}_N$ be the $N \times N$ identity matrix and letting ϕ be a positive (precision) parameter, $\mathrm{El}_N(\boldsymbol{0}, \frac{1}{\phi}\boldsymbol{I}_N; h)$ and $\mathrm{El}_N(\boldsymbol{\mu}, \frac{1}{\phi}\boldsymbol{I}_N; h)$ become a spherical distribution $S_N(\phi; h)$ and translate spherical distribution $S_N(\boldsymbol{\mu}, \phi; h)$ respectively. Clearly, if $\boldsymbol{z} \sim S_N(\phi; h)$, then $\boldsymbol{\mu} + \boldsymbol{z} \sim S_N(\boldsymbol{\mu}, \phi; h)$. In the spherical case $q(\boldsymbol{y})$ in (2.1) becomes $\phi\|\boldsymbol{y}-\boldsymbol{\mu}\|^2$. Spherical distributions may be viewed as standardized versions of elliptical distributions:

$$\boldsymbol{y} \sim \mathrm{El}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; h) \Leftrightarrow \boldsymbol{z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{y}-\boldsymbol{\mu}) \sim S_N(1; h). \tag{2.3}$$

To make explicitly the dimension N, h may be denoted by h^N. On the other hand, when h is unspecified or $\phi = 1$, these arguments may optionally be omitted.

When $\boldsymbol{z} \sim S_N(1; h)$, the components Z_j have a common symmetric distribution, whose variance and kurtosis are denoted by α_h and κ_h, respectively. By the additivity of expected values $\alpha_h = E(\frac{T}{N})$ and $\kappa_h = \frac{1}{\alpha_h^2} E(\frac{T^2}{N(N+2)}) - 1$, where T has the same distribution as $\|\boldsymbol{z}\|^2$. Thus, by the spherical symmetry $E(\boldsymbol{z}) = \boldsymbol{0}$ and $\mathrm{Var}(\boldsymbol{z}) = \alpha_h \boldsymbol{I}_N$ and, from (2.3) $E(\boldsymbol{y}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{y}) = \alpha_h \boldsymbol{\Sigma}$, where $\boldsymbol{y} \sim \mathrm{El}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; h)$. To save space we have adopted the convention that a formula for an expected value is implicitly assumed to hold only when this expected value exists.

Under sphericity, the distribution of $\boldsymbol{z}$ is fully determined by that of its squared length. Let $\boldsymbol{z} \sim S_N(1; h)$, $T = \|\boldsymbol{z}\|^2$, and $U = T^{-1}$. As in Arellano-Valle [1], the distributions of T and U will be called (univariate) *squared-radial* and *inverse squared radial* and will be denoted by $\mathrm{SR}(N; h)$ and $\mathrm{ISR}(N; h)$, respectively. These distributions are quite simple conceptually and have appeared in the literature under various names (see [6]). For further reference we write

$$\boldsymbol{z} \sim S_N(1; h) \Leftrightarrow T = \|\boldsymbol{z}\|^2 \sim \mathrm{SR}(N; h). \tag{2.4}$$

Under (2.1), it is easily shown that $q(\mathbf{y}) \sim \mathrm{SR}(N; h)$. Table B.1 in Appendix B shows the squared-radial distributions associated with various spherical distributions, while Table B.2 provides their means and variances. In general these results have appeared in previous papers, but are collected here for convenience (see, e.g., [5,13]).

2.2. *Marginal and conditional generators*

Let $z = (Z_j, j = 1, \ldots, N) \sim S_N(1; h)$ and let T be the squared length of z. Define also the random vectors $z_k = (Z_j, j = 1, \ldots, k), z_{(k)} = (Z_j, j = k+1, \ldots, N)$, and let T_k and $T_{(k)}$ be their squared lengths. Clearly, $z_k, z_{(k)}$ are spherical, $T_k, T_{(k)}$ are squared radial, and the distributions of z_k and $z_{(k)}$ are fully determined by those of T_k and $T_{(k)}$, respectively.

Now, let us denote by h^N the generator for z. The generator for z_k is $h^{k|N}$, defined by

$$h^{k|N}(u) = \int_0^\infty \frac{\pi^{\frac{N-k}{2}}}{\Gamma(\frac{N-k}{2})} v^{\frac{N-k}{2}-1} h^N(u+v)\, dv. \tag{2.5}$$

The conditional distribution of $z_k | z_{(k)} = \boldsymbol{\omega}$ is also spherical and $z_k \perp\!\!\!\perp z_{(k)} | T_{(k)}$. Thus $z_k | z_{(k)} = \boldsymbol{\omega} \stackrel{d}{=} z_k | T_{(k)} = t$, with $t = \|\boldsymbol{\omega}\|^2$. Furthermore, the generator for $z_k | T_{(k)} = t$ is $h_t^{k|N}$, defined by

$$h_t^{k|N}(u) = \frac{h^N(u+t)}{h^{N-k|N}(t)}, \quad u \geqslant 0. \tag{2.6}$$

From the formulae for the marginal and conditional generators we get

$$\int_0^\infty v^{\frac{N-k}{2}-1} h^N(u+v) dv \propto h^{k|N}(u),$$

and

$$h^N(u+t) = h_t^{k|N}(u) h^{N-k|N}(t), \quad u \geqslant 0. \tag{2.7}$$

Notational convention: *When the superscript N is clear from the context, it will usually be omitted, that is,* $h^N = h$, $h^{k|N} = h^k$ *and* $h_t^{k|N} = h_t^k$. *Furthermore, the corresponding marginal and conditional spherical distributions will be denoted by* $z_k \sim S_k(1; h)$ *and* $z_k | T_{(k)} = t \sim S_k(1; h_t)$, *and their respective squared-radial distributions by* $T_k \sim \mathrm{SR}(k; h)$ *and* $T_k | T_{(k)} = t \sim \mathrm{SR}(k; h_t)$.

The moments of z_k and $z_k | T_{(k)} = t$ are also determined by those of T_k and $T_k | T_{(k)} = t$, respectively. In particular,

$$E\left(\frac{T_k}{k}\right) = \alpha_h, \quad \mathrm{Var}\left(\frac{T_k}{k}\right) = \left\{\frac{k+2}{k}(\kappa_h + 1) - 1\right\} \alpha_h^2$$

and

$$E\left(\frac{T_k}{k} | T_{(k)} = t\right) = \alpha_h(t),$$

$$\mathrm{Var}\left(\frac{T_k}{k} | T_{(k)} = t\right) = \left\{\frac{k+2}{k}(\kappa_h(t) + 1) - 1\right\} \alpha_h^2(t), \tag{2.8}$$

where $\alpha_h(t) = \alpha_{h_t}$ and $\kappa_h(t) = \kappa_{h_t}$ are the variance and kurtosis parameters corresponding to the conditional generator h_t, respectively.

2.3. The generalized Student-t distribution

Throughout the paper we will illustrate our main results using the generalized elliptical N-dimensional Student-t distribution with v degrees of freedom, which is discussed in Arellano-Valle and Bolfarine [2]. It will be denoted by $t_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \lambda, v)$. We will concentrate first on the spherical case $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_N$, and write instead $t_N(\lambda, v)$. Its density generator is

$$h^N(u) = c(N, v)\lambda^{\frac{v}{2}}\{\lambda + u\}^{-\frac{N+v}{2}}, \text{ where } c(N, v) = \frac{\Gamma(\frac{N+v}{2})}{\Gamma(\frac{v}{2})\pi^{\frac{N}{2}}}. \tag{2.9}$$

For a spherical random vector z, (2.4) implies

$$z \sim t_N(\lambda, v) \Leftrightarrow T = \|z\|^2 \sim \frac{N\lambda}{v} F_{N,v}, \tag{2.10}$$

and $E(z) = \mathbf{0}$, for $v > 1$, and Var $(z) = \lambda(v-2)^{-1}\boldsymbol{I}_N$, for $v > 2$. Moreover, the marginal generator (2.5) yields $h^{k|N}(u) = h^k(u) = c(k, v)\lambda^{\frac{v}{2}}\{\lambda+u\}^{-\frac{k+v}{2}}$, i.e., $z_k \sim t_k(\lambda, v)$, while the conditional generator (2.6) yields $h_t^{k|N}(u) = c(k, v_{(k)})\lambda_t^{\frac{v_{(k)}}{2}}\{\lambda_t + u\}^{-\frac{k+v_{(k)}}{2}}$, i.e., $z_k | T_{(k)} = t \sim t_k(\lambda_t, v_{(k)})$, where $\lambda_t = \lambda + t$ and $v_{(k)} = v + N - k$. Thus,

$$\alpha_h(t) = \frac{\lambda_t}{v_{(k)} - 2} \quad \text{and} \quad \kappa_h(t) = \frac{v_{(k)} - 2}{v_{(k)} - 4} - 1. \tag{2.11}$$

Notice finally that the generalized Student-t distribution can be represented as scale-mixture of the normal distribution by assuming that $z|V = \upsilon \sim N_N(\mathbf{0}, \upsilon\boldsymbol{I}_N)$ and $V \sim \mathcal{IG}(\frac{v}{2}, \frac{\lambda}{2})$, the *inverse-gamma* distribution with parameter v and λ. Thus, $t_N(v, v) = t_N(v)$ is the usual *multivariate spherical Student-t distribution* with v degrees of freedom, and $t_N(\lambda, 2\upsilon - N) = \text{PVII}_N(\lambda, \upsilon)$, $\upsilon > \frac{N}{2}$, is the *multivariate spherical Pearson VII distribution* discussed in Fang et al. [6]. Thus, for instance most of the results derived from the generalized Student-t distribution can be adapted for the normal scale-mixture class, as well for the entire spherical class, in a rather simple way.

3. Multivariate squared-radial distributions

A multivariate extension of the squared-radial distribution is given next:

Definition 1. Let $z = (Z_j, j = 1, \ldots, N) \sim S_N(1; h)$, let $(M_1, \ldots, M_r)$ be an orthogonal decomposition of $\mathbb{R}^N$, with $\dim(M_j) = N_j$, and let $P_j z$ be the orthogonal projection of z onto M_j. We call the distribution of

$$\boldsymbol{w} = (W_j, j = 1, \ldots, r) = (\|P_j z\|^2, j = 1, \ldots, r) \tag{3.1}$$

multivariate squared-radial distribution and write $\boldsymbol{w} \sim \text{SR}(N_1, \ldots, N_r; h)$.

This is a valid definition because the spherical symmetry of $\boldsymbol{z}$ implies that the distribution of $\boldsymbol{w}$ depends on the subspaces only through their dimensions. A canonical form is obtained choosing M_j as the subspace spanned by a set of unit vectors. Let $\boldsymbol{v}_i$ be the ith unit vector and define $A_1 = \{1, \ldots, N_1\}$, $A_j = \{N_1 + \cdots + N_{j-1} + 1, \ldots, N_1 + \cdots + N_j\}$, $j = 2, \ldots, r$. Then the ith component of $P_j\boldsymbol{z}$ equals Z_i for $i \in A_j$ and it is equal to 0 otherwise. Thus

$$\boldsymbol{w} \stackrel{d}{=} \left(\sum_{i \in A_j} Z_i^2, \; j = 1, \ldots, r \right). \tag{3.2}$$

Some important properties of the family $\mathrm{SR}(N_1, \ldots, N_r; h)$ are:

SR1 *Closure under aggregation:* When the components are aggregated, the new distribution is still squared-radial, with the corresponding degrees of freedom being added. For example, $(W_1, W_2, W_3, W_4, W_5) \sim \mathrm{SR}(N_1, N_2, N_3, N_4, N_5; h) \Rightarrow (W_1 + W_2, W_3 + W_4, W_5) \sim \mathrm{SR}(N_1 + N_2, N_3 + N_4, N_5; h)$.

SR2 *Invariance for homogeneous functions:* If $g(a\boldsymbol{w}) = g(\boldsymbol{w})$ for any positive a, then $g(\boldsymbol{w})$ has the same distribution, say G, for all h and G may be obtained assuming $\boldsymbol{z}$ to be a normal random vector. For instance, the distribution of the ratio of two components does not depend on the generator.

SR3 *Independence of total and ratios:* $\frac{\boldsymbol{w}}{\sum_{j=1}^r W_j} \perp\!\!\!\perp \sum_{j=1}^r W_j$.

SR4 *Distribution of ratios:* Since the distribution does not depend on h, it may be computed under normality. In particular, $\frac{\boldsymbol{w}}{\sum_{j=1}^r W_j} \sim$ Dirichlet $(\frac{N_1}{2}, \ldots, \frac{N_r}{2})$.

SR5 *Invariance under conditioning:* $(W_1, W_2, \ldots, W_k | \sum_{j=k+1}^r W_j = s) \sim \mathrm{SR}(N_1, \ldots, N_r, h_s)$.

From SR1–SR5 several general properties of spherical distributions may be proved. For instance, taking $W_1 = T_k$ and $W_2 = T_{(k)}$, SR3 and SSR4 become $\frac{T_k}{T} \perp\!\!\!\perp T$ and $V = \frac{T_k}{T} \sim \mathrm{Beta}\left(\frac{k}{2}, \frac{N-k}{2}\right)$ respectively; while SR2 becomes $\frac{N-k}{k} \frac{T_k}{T_{(k)}} \stackrel{d}{=} \frac{N-k}{k} \frac{V}{1-V} \sim F_{k,N-k}$, regardless of the generator h. Under sphericity, the distribution of $(\boldsymbol{z}_k, \boldsymbol{z}_{(k)})$ is determined by that of $(T_k, T_{(k)})$ and the same holds for $(\boldsymbol{z}_k, T_{(k)})$. This translates into the following key relationship:

$$(T_k, T_{(k)}) \sim \mathrm{SR}(k, N-k; h) \Leftrightarrow \boldsymbol{z}_k | T_{(k)} = t \sim S_k(1; h_t) \quad \text{and} \\ T_{(k)} \sim \mathrm{SR}(N-k; h). \tag{3.3}$$

4. Dispersion elliptical squared-radial distributions

This section introduces a general class of distributions which will have a key place in the statistical analysis of the spherical linear model to be developed in the next sections. These distributions are called *dispersion elliptical squared-radial* (DESR) and they can be viewed as extension of the well-known normal-gamma family, which leads to a conjugate Bayesian analysis for the normal linear model.

Definition 2. Let $\boldsymbol{x}$ be a $m \times 1$ random vector and let T be a positive random variable. Let $h = h^{m+d}$ be a given generator. If

$$\boldsymbol{x}|T = t \sim \mathrm{El}_m\left(\boldsymbol{\mu}, \frac{1}{bt^r}\boldsymbol{\Sigma}; h_{at}\right) \quad \text{and} \quad T \sim \frac{1}{a}\mathrm{SR}(d; h), \tag{4.1}$$

then $(\boldsymbol{x}, T)$ is said to have a *dispersion elliptical squared-radial distribution* and it is denoted by $\mathrm{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}; h)$.

By ellipticity, the p.d.f. of $(\boldsymbol{x}, T) \sim \mathrm{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}; h)$ depends on $\boldsymbol{x}$ only through $q(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu})^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$. Straightforward computations lead to

$$f(\boldsymbol{x}, t) = \frac{(a\pi)^{\frac{d}{2}} b^{\frac{m}{2}}}{\Gamma\left(\frac{d}{2}\right)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} t^{\frac{rm+d}{2}-1} h(bt^r q(\boldsymbol{x}) + at), \quad (\boldsymbol{x} \in \mathbb{R}^m, \ t > 0). \tag{4.2}$$

4.1. Parameters and special cases

(a) $\frac{1}{b}$ is a scale parameter. In fact, $\mathrm{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}; h) = \mathrm{DESR}(m, d; r, a, 1; \boldsymbol{\mu}, \frac{1}{b}\boldsymbol{\Sigma}; h)$.

(b) The parameter r allows the conditional variances of $\boldsymbol{x}|T = t$ to depend on t. For $(r = 0, b = 1)$, the dispersion is fully determined by $\boldsymbol{\Sigma}$.

(c) The *dispersion spherical squared-radial* distribution $\mathrm{DSSR}(m, d; r, a, b; h)$ is just the special case $\mathrm{DESR}(m, d; r, a, b; \boldsymbol{0}, \boldsymbol{I}_m; h)$. On the other hand, it can be used as an alternative starting point by defining $\boldsymbol{z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu})$ (see (2.3)). Then

$$(\boldsymbol{x}, T) \sim \mathrm{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}; h) \Leftrightarrow (\boldsymbol{z}, T) \sim \mathrm{DSSR}(m, d; r, a, b; h).$$

(d) The *spherical squared-radial* distribution, denoted by $\mathrm{SSR}(m, d; h)$ is the particular case $\mathrm{DSSR}(m, d; 0, 1, 1; h)$. Directly from the definitions we get

$$(\boldsymbol{z}, T) \sim \mathrm{SSR}(m, d; h) \Leftrightarrow \boldsymbol{z}|T = t \sim S_m(1; h_t) \quad \text{and} \quad T \sim \mathrm{SR}(d; h).$$

When the conditional distributions of $\boldsymbol{z}|T = t$ are all spherical, they are determined by those of $\|\boldsymbol{z}\|^2|T = t$, which are in turn determine by the distribution of $(\boldsymbol{z}, T)$ (see also (3.3)). Therefore, the SSR and SR families are linked by

$$(\boldsymbol{z}, T) \sim \mathrm{SSR}(m, d; h) \Leftrightarrow (\|\boldsymbol{z}\|^2, T) \sim \mathrm{SR}(m, d; h).$$

We are now ready to state the following key result for the general DESR family:

Theorem 1. *If* $(\boldsymbol{x}, T) \sim \mathrm{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}); h)$, *then*

$$(bT^r q(\boldsymbol{x}), aT) \sim \mathrm{SR}(m, d; h).$$

Applying the properties SR1–SR5, as well as (2.10), (2.4) and (3.3) we get:

Corollary 1. *Let* $(\boldsymbol{x}, T) \sim \text{DESR}(m, d; r, a, b; \boldsymbol{\mu}, \boldsymbol{\Sigma}; h)$. *Then*

(a) $T^r q(\boldsymbol{x}) \sim \frac{1}{b}\text{SR}(m; h)$; $\quad T \sim \frac{1}{a}\text{SR}(d; h)$; $\quad q(\boldsymbol{x})|T = t \sim \frac{1}{bt^r}\,\text{SR}(m; h_{at})$.
(b) $T^{r-1} q(\boldsymbol{x}) \sim \frac{am}{bd}\, F_{m,d}$.
(c) *For* $r = 1$,

$$\boldsymbol{x} \sim t_m\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \frac{a}{b}, d\right) = t_m\left(\boldsymbol{\mu}, \frac{1}{b}\boldsymbol{\Sigma}; a, d\right), \quad \textit{which does not depend on } h. \tag{4.3}$$

(d) *For* $r = 0$, $\boldsymbol{x} \sim \text{El}_m(\boldsymbol{\mu}, \frac{1}{b}\boldsymbol{\Sigma}; h)$ *and* $T|\boldsymbol{x} \sim \frac{1}{a}\,\text{SR}(d; h_{bq(x)})$.

Proof. (a) follows from SR1 and SR5, while (b) follows from SR2 applied to the ratio of the two components. On the other hand, (b) implies $q(\boldsymbol{x}) \sim \frac{am}{bd}\,F_{m,d}$, for $r = 1$, so that (c) is obtained from (2.10). Part (d) is a consequence of (3.3) and the equivalence of $q(\boldsymbol{x}) \sim \frac{1}{b}\,\text{SR}(m; h)$ and $\boldsymbol{x} \sim \text{El}_m(\boldsymbol{\mu}, \frac{1}{b}\,\boldsymbol{\Sigma}; h)$.

4.2. Some important examples

Example 1. The normal-gamma family: It is a well-known family in the Bayesian literature, and now it may be obtained as special case of (4.1) by letting $b = r = 1$ and taking h to be the normal generator. In fact, taking $h(u) = (2\pi)^{-\frac{m+d}{2}} e^{-\frac{u}{2}}$, then $\boldsymbol{x}|T = t \sim N_m(\boldsymbol{\mu}, \frac{1}{t}\boldsymbol{\Sigma})$ and $T \sim \mathcal{G}(\frac{d}{2}, \frac{a}{2})$. Notice that the same case, but with $r = 0$, yields independent normal and gamma models for $\boldsymbol{x}$ and T, respectively. This is the only situation within the DESR class where we have independence for any value of a.

Example 2. The DSSR-family induced by the generalized Student-t distribution (see Section 2.3): It is defined by

$$\begin{aligned} \boldsymbol{x}|T = t &\sim t_m\left(\boldsymbol{\mu}, \frac{1}{bt^r}\,\boldsymbol{\Sigma}; \lambda_{at}; v + d\right) = t_m\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \frac{\lambda_{at}}{bt^r}; v + d\right) \quad \text{and} \\ T &\sim \frac{d\lambda}{av}\,F_{m,d}, \end{aligned} \tag{4.4}$$

where $\lambda_c = \lambda + c$. It p.d.f is easy obtained by applying (2.9), with $N = m + d$, in (4.2). This family will be used next in order to illustrate some results.

5. Inference in dispersion spherical squared-radial linear models

This section deals with the Bayesian analysis of a spherical linear model, with a generator depending on the precision parameter ϕ which is given a squared-radial prior. It is shown here that the prior independence of $\boldsymbol{\beta}$ and ϕ leads to DSSR distribution for $(\boldsymbol{\epsilon}, \phi)$. Moreover, under a flat improper prior for $\boldsymbol{\beta}$ the posterior for $(\boldsymbol{\beta}, \phi)$ belongs to the DSSR class, and so ϕ has a squared-radial posterior. Finally, the Bayes estimator of $\boldsymbol{\beta}$ is shown to be invariant with respect to changes in the generator. This robustness holds also under the much more general condition that the prior $\pi(\boldsymbol{\beta})$ does not depend on that generator, which generalizes analogous results in Zellner [14] and Osiewalski and Steel [10].

5.1. The linear model

Start with a $(n + d_0)$-dimensional spherical distribution with generator $h = h^{n+d_0}$ and assume the linear model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, with the assumptions that

$$\boldsymbol{\epsilon}|\phi \sim S_n(b_0\phi^r; h_{a_0\phi}) \quad \text{and} \quad \boldsymbol{\beta} \perp\!\!\!\perp \boldsymbol{\epsilon}|\phi, \tag{5.1}$$

which includes (1.1) as the particular case $r = 1, a_0 = 0$. The error components ε_i are still uncorrelated, with a common variance $\sigma_h^2(\phi)$ which depends both on h and ϕ. The general formula is $\sigma_h^2(\phi) = \frac{\alpha_h(a_0\phi)}{b_0\phi^r}$, where $\alpha_h(t) = \alpha_{h_t}$. For instance, the generalized Student-t model satisfies $\sigma_h^2(\phi) = \frac{\lambda}{(v+d_0-2)b_0\phi^r} + \frac{\lambda_0}{(v+d_0-2)\phi^{r-1}}$, where $\lambda_0 = \frac{a_0}{b_0}$ and $v + d_0 > 2$.

The likelihood of (5.1) is

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \phi) \propto \phi^{\frac{rn}{2}} h^n_{a_0\phi}(b_0\phi^r \parallel \boldsymbol{y} - \boldsymbol{X\beta} \parallel^2), \tag{5.2}$$

where a formula for the conditional generator $h^n_{a_0\phi}$ is given in (2.6), with the substitutions $t = a_0\phi; k = n; N = n + d_0$.

Consider the proper prior specification

$$\pi(\boldsymbol{\beta}|\phi) \text{ arbitrary}$$

and

$$\pi(\phi) \propto \phi^{\frac{d_0}{2}-1}\, h^{d_0}(a_0\phi), \quad \text{i.e.} \quad \phi \sim \frac{1}{a_0}\,\mathrm{SR}(d_0; h). \tag{5.3}$$

Note that the usual non-informative prior $\pi(\phi) \propto \phi^{-1}$ is formally obtained taking $d_0 = a_0 = 0$. At some points we will need the prior independence assumption

$$\boldsymbol{\beta} \perp\!\!\!\perp \phi, \tag{5.4}$$

the flat prior

$$\pi(\boldsymbol{\beta}|\phi) = \text{constant} \tag{5.5}$$

or both.

From (5.2), (5.3) and (2.7) the following joint density becomes

$$f(\boldsymbol{y}, \boldsymbol{\beta}, \phi) \propto \phi^{\frac{rn+d_0}{2}-1}\, h^{n+d_0}(b_0\phi^r \|\boldsymbol{y} - \boldsymbol{X\beta}\|^2 + a_0\phi)\, \pi(\boldsymbol{\beta}|\phi). \tag{5.6}$$

Under (5.4), $(\boldsymbol{\epsilon}, \phi) \perp\!\!\!\perp \boldsymbol{\beta}$ and

$$(\boldsymbol{\epsilon}, \phi) \sim \mathrm{DSSR}(n, d_0; r, a_0, b_0; h), \tag{5.7}$$

which includes the joint assumptions (1.1–1.2) as the particular case $r = b_0 = 1, a_0 = d_0 = 0$. From (5.7) one immediately gets the following theorem:

Theorem 2. *Under* (5.1–5.4)

$$(b_0\phi^r \|\boldsymbol{\epsilon}\|^2,\ a_0\phi) \sim \mathrm{SR}(n, d_0; h) \quad \textit{and} \quad \phi^{r-1}\|\boldsymbol{y} - \boldsymbol{X\beta}\|^2 \sim \frac{a_0 n}{b_0 d_0}\, F_{n,d_0} \tag{5.8}$$

define pivotal quantities for $(\boldsymbol{\beta}, \phi)$.

The posterior $\pi(\boldsymbol{\beta}|\boldsymbol{y})$ does not admit a general closed expression, unless the geometric structure is preserved by the choice of the improper prior (5.5). In this case, (5.6) may be used to get the DSSR posterior density:

$$\pi(\boldsymbol{\beta}, \phi|\boldsymbol{y}) \propto \phi^{\frac{rn+d_0}{2}-1}\, h^{n+d_0}(b_0\phi^r \parallel \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta} \parallel^2 + a_0\phi). \tag{5.9}$$

We may also get a Bayesian version of the classical pivotal quantities for $\boldsymbol{\beta}$ and ϕ. In fact, (5.7), (5.9) and a constant $\pi(\boldsymbol{\beta})$ imply $(b_0\phi^r \parallel \boldsymbol{\epsilon} \parallel^2, a_0\phi)|\boldsymbol{y} \stackrel{d}{=} (b_0\phi^r \parallel \boldsymbol{\epsilon} \parallel^2, a_0\phi)|\boldsymbol{\beta}$, and their common distribution is $\mathrm{SR}(n, d_0; h)$. Thus, conditioning on either $\boldsymbol{\beta}$ or $\boldsymbol{y}$ produces the same squared-radial distribution, which may be applied to get the posterior of $\boldsymbol{\beta}$ and ϕ.

In the sequel we employ the standard notation: $\hat{\boldsymbol{\beta}}$ is the ordinary least-squares estimate of $\boldsymbol{\beta}$, i.e., $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}$; $\boldsymbol{e}$ is the residual vector $\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$, $(n-p)s^2$ is the residual sum of squares $\parallel \boldsymbol{e} \parallel^2$, $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$, and $\hat{\boldsymbol{\mu}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$. We will need the simple identities

$$\begin{aligned} &\parallel \boldsymbol{\mu} - \hat{\boldsymbol{\mu}} \parallel^2 \text{ coincides with the quadratic form } (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^t\boldsymbol{X}^t\boldsymbol{X}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}),\\ &\parallel \boldsymbol{\epsilon} \parallel^2 = \parallel \boldsymbol{e} \parallel^2 + \parallel \boldsymbol{\mu} - \hat{\boldsymbol{\mu}} \parallel^2 . \end{aligned} \tag{5.10}$$

Thus, we end this section by stating another theorem, which may be used in conjunction with Corollary 1 to give an alternative proof of Theorems 4 and 5.

Theorem 3. *Under* (5.1–5.5),

$$(b_0\phi^r \parallel \boldsymbol{e} \parallel^2, b_0\phi^r \parallel \boldsymbol{\mu} - \hat{\boldsymbol{\mu}} \parallel^2, a_0\phi) \sim \mathrm{SR}(n-p, p, d_0; h).$$

5.2. *Two special cases*

Throughout this section we assume that (5.1), (5.3)–(5.5) hold (which is equivalent to considering the posterior distribution $(\boldsymbol{\beta}, \phi)|\boldsymbol{y} \sim \mathrm{DSSR}(n, d_0; r, a_0, b_0; h)$ defined by (5.9)) and analyzing separately the case of an identity ($r = 1$) and a constant ($r = 0$) dispersion function.

5.2.1. *Identity dispersion function: $r = 1$*

In this case ϕ disappears from (5.8) and we get a pivotal quantity for $\boldsymbol{\beta}$:

$$\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 \sim \frac{a_0 n}{b_0 d_0} F_{n,d_0}.$$

From Corollary 1 and (4.3) $\boldsymbol{\epsilon}|\boldsymbol{\beta} \sim t_n(\frac{a_0}{b_0}, d_0)$, regardless of $\boldsymbol{\beta}$ or h.

Theorem 4. *Let* $r = 1$ *and assume that* (5.1), (5.3)–(5.5) *hold. Letting* $\sigma^2 = \phi^{-1}$, $a_\omega = a_0 + \omega$ *and* $d = n - p + d_0$ *we have*

$$\boldsymbol{\beta}|\boldsymbol{y} \sim t_p\left(\hat{\boldsymbol{\beta}}, \frac{1}{b_0}(\boldsymbol{X}^t\boldsymbol{X})^{-1}; a_{b_0 q(y)}, d\right), \tag{5.11}$$

$$\phi|\boldsymbol{y} \sim \frac{1}{a_{b_0 q(y)}}\,\mathrm{SR}(d; h), \quad \sigma^2|\boldsymbol{y} \sim a_{b_0 q(y)}\,\mathrm{ISR}(d; h), \tag{5.12}$$

where $q(\boldsymbol{y}) = (n-p)s^2$.

Theorem 4 shows that the class of SR distributions provides a conjugate analysis for ϕ. The invariance of $\pi(\boldsymbol{\beta}|\boldsymbol{y})$ under a change of generator allows the computations to be performed under the simpler normal model. This invariance holds under the much more general condition that $\pi(\boldsymbol{\beta})$ does not depend on h, generalizing thus the results obtained by Osiewalski and Steel [10] under the conventional spherical linear model (1.1) and the improper prior (1.2). From the properties of the generalized Student-t model and (2.8)

$$E(\boldsymbol{\beta}|\boldsymbol{y}) = \hat{\boldsymbol{\beta}} \ \ (d > 1), \quad \text{Var}\,(\boldsymbol{\beta}|\boldsymbol{y}) = \tfrac{a_{b_0 q(y)}}{b_0(d-2)}\,(\boldsymbol{X}^t\boldsymbol{X})^{-1} \ \ (d > 2),$$
$$E(\phi|\boldsymbol{y}) = \tfrac{d\,\alpha_h}{a_{b_0 q(y)}}, \quad \text{Var}\,(\phi|\boldsymbol{y}) = \left\{\tfrac{d+2}{d}\,(\kappa_h + 1) - 1\right\}\{E(\phi|\boldsymbol{y})\}^2,$$

where $q(\boldsymbol{y}) = (n-p)s^2$. For $b_0 = 1, d_0 = a_0 = 0$, $\boldsymbol{\beta}|\boldsymbol{y} \sim t_p(\hat{\boldsymbol{\beta}}, s^2(\boldsymbol{X}^t\boldsymbol{X})^{-1}; n-p)$ and $\phi|\boldsymbol{y} \sim \frac{1}{(n-p)s^2}\ \text{SR}(n-p; h)$, a result due to Zellner [14], which here becomes a special case of (5.11–5.12). Table B.2 makes it easy to get means and variances for a variety of elliptical models.

Example 3 (*Example 2 continued*). Consider the DSSR family induced by generalized spherical Student-t distribution with dimension $m = n + d_0$, which is defined in (4.4). Thus, the model defined by (5.3–5.2) satisfies

$$\boldsymbol{y}|\boldsymbol{\beta}, \phi \sim t_n\left(\boldsymbol{X\beta}; \frac{\lambda_{a_0\phi}}{b_0\phi^r}, v + d_0\right) \quad \text{and} \quad \phi \sim \frac{\lambda\, d_0}{v\, a_0}\, F_{d_0,v}.$$

Substituting $r = 1$ in (5.6), $f(\boldsymbol{y}, \phi|\boldsymbol{\beta}) \propto \phi^{\frac{n+d_0}{2}-1}\,\{\lambda + b_0\phi \parallel \boldsymbol{y} - \boldsymbol{X\beta} \parallel^2 + a_0\phi\}^{-\frac{n+d_0+v}{2}}$. Furthermore $\pi(\phi|\boldsymbol{y}) \propto \phi^{\frac{d}{2}-1}\{\lambda + a_{b_0q(y)}\phi\}^{-\frac{d+v}{2}}$, where $q(\boldsymbol{y}) = (n-p)s^2$, implying $\phi|\boldsymbol{y} \sim \frac{\lambda\, d}{v\, a_{b_0q(y)}}\ F_{d,v}$ and $\sigma^2|\boldsymbol{y} \sim \frac{v\, a_{b_0q(y)}}{\lambda\, d}\ F_{v,d}$. The Bayes estimator and the Bayes risk of ϕ and $\sigma^2 = \phi^{-1}$ are

$$E(\phi|\boldsymbol{y}) = \frac{\lambda}{v-2}\,\frac{d}{a_{b_0q(y)}}, \quad (v > 2),$$

$$\text{Var}\,(\phi|\boldsymbol{y}) = \left\{\frac{d+2}{d}\left(\frac{v-2}{v-4}\right) - 1\right\}\left\{\frac{\lambda}{v-2}\,\frac{d}{a_{b_0q(y)}}\right\}^2, \quad (v > 4),$$

$$E(\sigma^2|\boldsymbol{y}) = \frac{v}{\lambda}\,\frac{a_{b_0q(y)}}{d-2}, \quad (n > p+2),$$

$$\text{Var}\,(\sigma^2|\boldsymbol{y}) = \left\{\frac{v+2}{v}\,\frac{d-2}{d-4} - 1\right\}\left\{\frac{v}{\lambda}\,\frac{a_{b_0q(y)}}{d-2}\right\}^2, \quad (n > p+4),$$

respectively.

5.2.2. *Constant dispersion function:* $r = 0$

Unlike the case $r = 1$, the posterior of $\boldsymbol{\beta}$ depends on the generator h. The next theorem states that under an improper constant prior for $\boldsymbol{\beta}$, the posterior for $\boldsymbol{\beta}$ is elliptical and the posterior of ϕ is squared-radial.

Theorem 5. *Let* $r = 0$ *and assume that* (5.1), (5.3)–(5.5) *hold. Letting* $\sigma^2 = \phi^{-1}$, $a_\omega = a_0 + \omega$ *and* $d = n - p + d_0$ *we have*

$$\boldsymbol{\beta}|\boldsymbol{y} \sim \mathrm{El}_p\left(\hat{\boldsymbol{\beta}}, \frac{1}{b_0}(\boldsymbol{X}^t\boldsymbol{X})^{-1}; h_{b_0 q(y)}\right), \tag{5.13}$$

$$\phi|\boldsymbol{y} \sim \frac{1}{a_0}\,\mathrm{SR}(d_0; h_{b_0 q(y)}), \quad \sigma^2|\boldsymbol{y} \sim a_0\mathrm{ISR}(d_0; h_{b_0 q(y)}), \tag{5.14}$$

where $q(\boldsymbol{y}) = (n-p)s^2$.

The means and variance are easily obtained from the general results given in Section 2. Thus, letting $\alpha_h(t) = \alpha_{h_t}$ and $\kappa_h(t) = \kappa_{h_t}$,

$$E(\boldsymbol{\beta}|\boldsymbol{y}) = \hat{\boldsymbol{\beta}}, \quad (n > p+2), \quad \mathrm{Var}\,(\boldsymbol{\beta}|\boldsymbol{y}) = \frac{\alpha_h(b_0 q(y))}{b_0}(\boldsymbol{X}^t\boldsymbol{X})^{-1},$$
$$E(\phi|\boldsymbol{y}) = \frac{d_0\ \alpha_h(b_0 q(y))}{a_0}, \quad \mathrm{Var}\,(\phi|\boldsymbol{y}) = \left\{\frac{d_0+2}{d_0}(\kappa_h(b_0 q(y))+1)-1\right\}\{E(\phi|\boldsymbol{y})\}^2,$$

where, as was defined above, $q(\boldsymbol{y}) = (n-p)s^2$.

Although (5.13) implies $\pi(\boldsymbol{\beta}|\boldsymbol{y})$ depends on h, the Bayes estimator of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}}$ regardless of the generator.

Table B.2 in Appendix is useful to get the above results for some particular elliptical distributions.

Example 4 (*Example 2 continued*). For $r = 0$, (5.13, 5.14) imply

$$\boldsymbol{\beta}|\boldsymbol{y} \sim t_p\left(\hat{\boldsymbol{\beta}}, \frac{1}{b_0}(\boldsymbol{X}^t\boldsymbol{X})^{-1}; \lambda_{b_0 q(y)}, v+n-p\right) \quad \text{and}$$
$$\phi|\boldsymbol{y} \sim \frac{\lambda_{b_0 q(y)}\ d_0}{(v+n-p)\ a_0}\ F_{d_0, v+n-p},$$

where $\lambda_t = \lambda + t$, $q(\boldsymbol{y}) = (n-p)s^2$. Thus, the posterior of $\boldsymbol{\beta}$ does depend on the generator. The estimator and risk of ϕ are easily obtained from (2.11).

6. Inference in dispersion-location spherical squared-radial linear models

In this section we build a new model that produces conjugate distributions for $\boldsymbol{\beta}$ and ϕ, both individually and jointly. This is achieved by modelling $(\boldsymbol{y}, \boldsymbol{\beta}, \phi)$ jointly.

6.1. Prior distributions

The normal-gamma priors for the linear model (1.1) do not lead to an analytically tractable posterior, unless the spherical distribution is normal. To develop an extension we start with the prior

$$(\boldsymbol{\beta}, \phi) \sim \mathrm{DESR}(p, d_0; r, a_0, b_0; \boldsymbol{m}_0, \boldsymbol{V}_0; h), \tag{6.1}$$

which by (4.1) is equivalent to

$$\boldsymbol{\beta}|\phi \sim \mathrm{El}_m\left(\boldsymbol{m}_0, \frac{1}{b_0\phi^r}\boldsymbol{V}_0; h_{a_0\phi}\right), \quad \phi \sim \frac{1}{a_0}\,\mathrm{SR}(d_0; h). \tag{6.2}$$

Under normality (6.2), with $r = 1$, reduces to the usual normal-gamma prior distribution. Besides, for $r = 1$ $\boldsymbol{\beta} \sim t_p\left(\boldsymbol{m}_0, \frac{1}{b_0}\boldsymbol{V}_0; a_0, d_0\right)$, regardless of the generator chosen.

6.2. Joint distributions

For a fixed generator $h = h^{n+p+d_0}$, define $q_0(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \boldsymbol{m}_0)^t\boldsymbol{V}_0^{-1}(\boldsymbol{\beta} - \boldsymbol{m}_0)$ and $w_0(\boldsymbol{\beta}, \phi) = b_0\phi^r q_0(\boldsymbol{\beta}) + a_0\phi$. The likelihood is then determined by the distribution of $\boldsymbol{\epsilon}|\boldsymbol{\beta}, \phi$. To get conjugacy results we assume a rather special form for this conditional distribution:

$$\boldsymbol{\epsilon}|\boldsymbol{\beta}, \phi \sim S_n(b_0\phi^r; h_{w_0(\beta,\phi)}). \tag{6.3}$$

The likelihood for the linear model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ is

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \phi) \propto \phi^{\frac{rn}{2}} h^n_{w_0(\beta,\phi)}(b_0\phi^r \parallel \boldsymbol{y} - \boldsymbol{X\beta} \parallel^2). \tag{6.4}$$

Let $\boldsymbol{\omega} = (\boldsymbol{\epsilon}, \boldsymbol{\beta})$, $\boldsymbol{t}_0 = (\boldsymbol{0}, \boldsymbol{m}_0)$ and $\boldsymbol{W}_0 = \mathrm{diag}\,\{\boldsymbol{I}_n, \boldsymbol{V}_0\}$. Considering Definition 1 we get that (6.1) and (6.3) imply

$$(\boldsymbol{\omega}, \phi) \sim \mathrm{DESR}(n + p, d_0; r, a_0, b_0; \boldsymbol{t}_0, \boldsymbol{W}_0; h).$$

An explicit formula for the joint density is

$$f(\boldsymbol{y}, \boldsymbol{\beta}, \phi) \propto \phi^{\frac{rn+rp+d_0}{2}-1}\, h^{n+p+d_0}(b_0\phi^r \parallel \boldsymbol{\epsilon} \parallel^2 + b_0\phi^r q(\boldsymbol{\beta}) + a_0\phi), \tag{6.5}$$

which implies that $(\boldsymbol{\beta}, \phi)|\boldsymbol{y}$ follows a DESR model.

From (6.1) and (6.3) we get the following fundamental property:

Theorem 6. *Under* (6.1) *and* (6.3),

$$(b_0\phi^r\|\boldsymbol{\epsilon}\|^2, b_0\phi^r q_0(\boldsymbol{\beta}), a_0\phi) \sim \mathrm{SR}(n, p, d_0; h),$$

and in analogy with Theorem 2 *we get*

$$(b_0\phi^r\|\boldsymbol{e}\|^2, b_0\phi^r\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2, b_0\phi^r q_0(\boldsymbol{\beta}), a_0\phi) \sim \mathrm{SR}(n - p, p, p, d_0; h).$$

The next theorem characterizes the posteriors of $\boldsymbol{\beta}$ and ϕ.

Theorem 7. *Assume that* (6.1) *and* (6.3) *hold. Let* $\boldsymbol{V} = (\boldsymbol{V}_0^{-1} + \boldsymbol{X}^t\boldsymbol{X})^{-1}$ *and* $\boldsymbol{m} = \boldsymbol{V}\boldsymbol{V}_0^{-1}\boldsymbol{m}_0 + (\boldsymbol{I}_p - \boldsymbol{V}\boldsymbol{V}_0^{-1})\hat{\boldsymbol{\beta}}$. *Let*

$$q(\boldsymbol{y}) = (n - p)s^2 + (\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{X}\boldsymbol{m}_0)^t(\boldsymbol{I}_n - \boldsymbol{X}\boldsymbol{V}\boldsymbol{X}^t)(\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{X}\boldsymbol{m}_0) \tag{6.6}$$

and $w(\boldsymbol{y}, \phi) = b_0\phi^r q(\boldsymbol{y}) + a_0\phi$. *Then*

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{\beta} \end{pmatrix} |\phi \sim \mathrm{El}_{n+p}\left(\begin{pmatrix} \boldsymbol{X}\boldsymbol{m}_0 \\ \boldsymbol{m}_0 \end{pmatrix}, \frac{1}{b_0\phi^r}\begin{pmatrix} \boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{V}_0\boldsymbol{X}^t & \boldsymbol{X}\boldsymbol{V}_0 \\ \boldsymbol{V}_0\boldsymbol{X}^t & \boldsymbol{V}_0 \end{pmatrix}; h_{a_0\phi}\right), \tag{6.7}$$

$$\boldsymbol{\beta}|\boldsymbol{y}, \phi \sim \mathrm{El}_p\left(\boldsymbol{m}, \frac{1}{b_0\phi^r}\boldsymbol{V}; h_{w(y,\phi)}\right), \tag{6.8}$$

$$\phi|\boldsymbol{y} \sim \frac{1}{a_{b_0q(y)}}\mathrm{SR}(n+d_0; h). \tag{6.9}$$

From (6.8) $E(\boldsymbol{\beta}|\boldsymbol{y}, \phi) = E(\boldsymbol{\beta}|\boldsymbol{y}) = \boldsymbol{m}$, which depends neither on ϕ nor on h. Therefore the posterior mean coincides with that obtained assuming a normal linear model with a normal-gamma prior distribution. Note however that the choice of h may have an effect on the posterior covariance matrix of $\boldsymbol{\beta}$. The posterior distributions of $\boldsymbol{\beta}$ and ϕ are easily obtained by applying Theorem 7.

Corollary 2. *For* $r = 1$

$$\boldsymbol{\beta}|\boldsymbol{y} \sim t_p\left(\boldsymbol{m}, \frac{1}{b_0}\boldsymbol{V}; a_{b_0q(y)}, n+d_0\right) \quad \text{and} \quad \phi|\boldsymbol{y} \sim \frac{1}{a_{b_0q(y)}}\mathrm{SR}(n+d_0; h); \tag{6.10}$$

while for $r = 0$

$$\boldsymbol{\beta}|\boldsymbol{y} \sim \mathrm{El}_p\left(\boldsymbol{m}, \frac{1}{b_0}\boldsymbol{V}; h_{b_0q(y)}\right) \quad \text{and} \quad \phi|\boldsymbol{y} \sim \frac{1}{a_0}\mathrm{SR}(d_0; h_{b_0q(y)}), \tag{6.11}$$

where $a_{b_0q(y)} = a_0 + b_0q(\boldsymbol{y})$ *and* $q(\boldsymbol{y})$ *is defined in* (6.6).

Example 5 (*Example 2 continued*). Choose $m = n + p + d_0$ and apply (6.5) to get

$$f(\boldsymbol{y}, \boldsymbol{\beta}, \phi) \propto \phi^{\frac{nr+pr+d_0}{2}-1}\{\lambda + b_0\phi^r \parallel \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta} \parallel^2 + b_0\phi^r q(\boldsymbol{\beta}) + a_0\phi\}^{-\frac{n+p+d_0+v}{2}}.$$

Following (6.7) the model may be specified as

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{\beta} \end{pmatrix} |\boldsymbol{X}, \phi \sim t_{n+p}\left(\begin{pmatrix} \boldsymbol{X}\boldsymbol{m}_0 \\ \boldsymbol{m}_0 \end{pmatrix}, \frac{1}{b_0\phi^r}\begin{pmatrix} \boldsymbol{I}_n + \boldsymbol{X}\boldsymbol{V}_0\boldsymbol{X}^t & \boldsymbol{X}\boldsymbol{V}_0 \\ \boldsymbol{V}_0\boldsymbol{X}^t & \boldsymbol{V}_0 \end{pmatrix}; \lambda_{a_0\phi}, v+d_0\right)$$

and

$$\phi \sim \frac{\lambda\, d_0}{v\, a_0} F_{d_0,v}.$$

From (6.3) and (6.2) the likelihood and the prior may written as

$$\boldsymbol{y}|\boldsymbol{\beta}, \phi \sim t_n\left(\boldsymbol{X}\boldsymbol{\beta}, \frac{\lambda_{w_0(\beta,\phi)}}{b_0\phi^r}, v+d_0\right), \quad \text{with} \quad w_0(\boldsymbol{\beta}, \phi) = b_0\phi^r q_0(\boldsymbol{\beta}) + a_0\phi,$$

and

$$\boldsymbol{\beta}|\phi \sim t_p\left(\boldsymbol{m}_0, \frac{1}{b_0\phi^r}\boldsymbol{V}_0; \lambda_{a_0\phi}, v+d_0\right), \quad \phi \sim \frac{\lambda\, d_0}{v\, a_0} F_{d_0,v},$$

respectively. Hence, when $r = 1$, in (6.10), we get

$$\phi|\mathbf{y} \sim \frac{\lambda\ (n+d_0)}{v\ a_{b_0 q(y)}}\ F_{n+d_0,v},$$

where $a_{b_0 q(y)} = a_0 + b_0 q(\mathbf{y})$. For $r = 0$, (6.11) yields

$$\boldsymbol{\beta}|\mathbf{y} \sim t_p\left(\boldsymbol{m}, \frac{1}{b_0}\boldsymbol{V}; \lambda_{b_0 q(y)}, v + d_0\right) \quad \text{and} \quad \phi|\mathbf{y} \sim \frac{\lambda_{b_0 q(y)}\ d_0}{(v+n)\ a_0}\ F_{d_0,v+n}.$$

The results for the usual Student-t model follow by setting $\lambda = v$.

Table B.2 is easily adapted to cover this case.

7. Discussion

One main concern of this paper is to provide basic distribution theory for the elliptical linear model, with special emphasis on the analytic study of Bayesian analysis for these models. From these general results we wish to remark some points. First, for a normal generator and $r = 0$ the distribution in (5.1) and (6.3) reduces to $N_n(\boldsymbol{X\beta}, \frac{1}{b_0}\boldsymbol{I}_n)$ and the parameter ϕ fades away. This surprising result means that ϕ has no impact at all on inferences about $\boldsymbol{\beta}$, which is more likely to reflect a deficiency of this model than a physical property. A similar phenomenon takes place with the likelihood in (5.2) and (6.4). For $r = 1$, together with particular values of the constants a_0, b_0, and the matrix $\boldsymbol{V}_0$, (1.1) is obtained. Although, by definition, d_0 is a positive integer, a convenient meaning may be given to expressions involving the choice $d_0 = 0$. For instance, the prior $\pi(\phi) \propto \phi^{\frac{d_0}{2}-1} h^{d_0}(a_0\phi)$ formally becomes $\pi(\phi) \propto \phi^{-1} h^0(0)$, which yields the reference improper prior $\pi(\phi) \propto \phi^{-1}$, interpreting $h^0(0)$ as a positive constant. The same improper prior may be formally obtained from (6.1), taking $a_0 = 0$ and either $\boldsymbol{V}_0^{-1} = \mathbf{0}$ or $b_0 = 0$, provided $rp + d_0 = 0$.

A nice feature of the results obtained here is that posterior means and variances are easily obtained from the general results for elliptical distributions. One just needs to make sure that the suitable scale factor is used. We have purposely tried to minimize the explicit computations of densities, but it is sometimes simpler to prove results in this way. For instance, the identities

$$\frac{h^{n+p+d_0}(u+a+b)}{h^{p+d_0}(a+b)} = \frac{h_b^{n+p}(u+a)}{h_b^{p}(a)} = \frac{h_a^{n+d_0}(u+b)}{h_a^{d_0}(b)},$$

which are easily derived from (2.6, 2.7) may be applied to get (6.5) and other results in Section 5. They admit a probabilistic interpretation in terms of the sequential conditioning.

Finally, a much more general situation that was not considered in this work is when the generator h is indexed by additional "nuisance" parameters, so that it will be unspecified. Such situations can be adapted in our approaches following the related well discussion in Osiewalski and Steel [10].

Acknowledgments

We acknowledge the detailed comments of one referee, which contributed to a substantial improvement of the manuscript. This work was partially supported by Grants 8000004, 1040865 and 1030588, of Fondecyt/Chile.

Appendix A. Proofs of Theorems

Proof of Theorem 2. It follows immediately from $b_0\phi^r\|\boldsymbol{\epsilon}\|^2 \sim \mathrm{SR}(n; h)$, the orthogonal projections associated with (5.10) and the general properties of the multivariate radial distributions. □

Proof of Theorem 4. From $d = n - p + d_0$, (5.4), (5.6) and (5.10) $\pi(\boldsymbol{\beta}, \phi|\boldsymbol{y}) \propto \phi^{\frac{p+d}{2}-1} h(\{a_{b_0q(y)} + b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2\}\phi)\ \pi(\boldsymbol{\beta})$. By integrating with respect to ϕ and applying (2.2) we get $\pi(\boldsymbol{\beta}|\boldsymbol{y}) \propto \{a_{b_0q(y)} + b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2\}^{-\frac{p+d}{2}}\ \pi(\boldsymbol{\beta})$, from which (5.11) follows. By integrating with respect to $\boldsymbol{\beta}$, $\pi(\phi|\boldsymbol{y}) \propto \phi^{\frac{p+d}{2}-1} \int_{\mathbb{R}^p} h^{p+d}(\{a_{b_0q(y)} + b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2\}\phi)\ \pi(\boldsymbol{\beta})d\boldsymbol{\beta}$. Applying now (5.5) and the change of variable $u = b_0\phi\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2$ reduces this to $\pi(\phi|\boldsymbol{y}) \propto \phi^{\frac{p+d}{2}-1} \int_0^\infty u^{\frac{p}{2}-1} h^{p+d}(u + a_{b_0q(y)}\phi)du$, and (5.12) follows from (2.5) by letting $v = a_{b_0q(y)}\phi$. □

Proof of Theorem 5. Substituting $r = 0$ and (5.10) in (5.6) we get $\pi(\boldsymbol{\beta}, \phi|\boldsymbol{y}) \propto \phi^{\frac{d_0}{2}-1} h^{n+d_0}(b_0q(\boldsymbol{y}) + b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2 + a_0\phi)\ \pi(\boldsymbol{\beta})$. Letting $u = b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2 + a_0\phi$, $h^{n+d_0}(u + b_0q(\boldsymbol{y})) = h^{p+d_0+n-p}(u + q(\boldsymbol{y})) = h^{p+d_0}_{b_0q(y)}(u)\ h^{n-p}(b_0q(\boldsymbol{y}))$, the joint posterior density becomes $\pi(\boldsymbol{\beta}, \phi|\boldsymbol{y}) \propto \phi^{\frac{d_0}{2}-1}\ h^{p+d_0}_{b_0q(y)}(b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2 + a_0\phi)\ \pi(\boldsymbol{\beta})$. From $h^{k|p+d_0}_{b_0q(y)}(u) = h^{k|n+d_0}_{b_0q(y)}(u)$ we get $\pi(\boldsymbol{\beta}|\boldsymbol{y}) \propto h^{p}_{b_0q(y)}(b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2)\ \pi(\boldsymbol{\beta})$ and $\pi(\phi|\boldsymbol{y}) \propto \phi^{\frac{d_0}{2}-1} \int_{\mathbb{R}^p} h^{p+d_0}_{b_0q(y)}(b_0\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2 + a_0\phi)\ \pi(\boldsymbol{\beta})d\boldsymbol{\beta}$, where $h^k_{b_0q(y)}(u) = \frac{h^{k+n-p}(u+b_0q(y))}{h^{n-p}(b_0q(y))},\ u \geqslant 0,\ k = p, p + d_0$. The required results follows by choosing a constant prior for $\boldsymbol{\beta}$. □

Proof of Theorem 7. Proof of (6.7). By assumption, the distribution of $(\boldsymbol{\epsilon}, \boldsymbol{\beta})|\phi$ is elliptical. Since linear transformations preserve ellipticity, the distribution of $(\boldsymbol{y}, \boldsymbol{\beta})|\phi = (\boldsymbol{\epsilon} + \boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\beta})|\phi$ is also elliptical. Its mean and dispersion matrix satisfy are identical to those derived under normality. The conditioning on ϕ justifies the subscript $a_0\phi$ in h. □

Proof of (6.8). It follows from (6.7) or by appealing to the results under normality. □

Proof of (6.9). This involves some long algebraic manipulation and an application of (6.5), which produces the expression $f(\boldsymbol{y}, \phi) \propto \phi^{\frac{rn+rp+d_0}{2}-1} \int_{\mathbb{R}^p} h^{n+p+d_0}(b_0\phi^r(\boldsymbol{\beta} - \boldsymbol{m})^t\boldsymbol{V}^{-1}(\boldsymbol{\beta} - \boldsymbol{m}) + b_0\phi^r q(\boldsymbol{y}) + a_0\phi)d\boldsymbol{\beta}$. Letting now $u = b_0\phi^r(\boldsymbol{\beta} - \boldsymbol{m})^t\boldsymbol{V}^{-1}(\boldsymbol{\beta} - \boldsymbol{m})$, this integral is reduces to $f(\boldsymbol{y}, \phi) \propto \phi^{\frac{rn+d_0}{2}-1} \int_0^\infty u^{\frac{p}{2}-1}\ h^{n+p+d_0}(u + b_0\phi^r q(\boldsymbol{y}) + a_0\phi)du$. An application of (2.5) yields $f(\boldsymbol{y}, \phi) \propto \phi^{\frac{rn+d_0}{2}-1}\ h^{n+d_0}(b_0\phi^r q(\boldsymbol{y}) + a_0\phi)$, from which (6.9) follows. □

Table B.1
Squared-radial distributions corresponding to some subclasses of n-dimensional spherical distributions ($u = \| z \|^2,\ z \in \mathbb{R}^N$)

Spherical distribution $S_N(1;h)$	Density generator h^N	Squared-radial distribution $\mathrm{SR}(N;h)$
Normal	$c_N \exp\{-\frac{1}{2}u\}$	χ^2_N
Contaminated normal	$(1-\varepsilon)\,\phi^{(N)}(u) + \varepsilon\,\lambda^N \phi^{(N)}(\lambda u)$, $0 < \varepsilon < 1,\ \lambda > 0$	$(1-\varepsilon)\,\chi^2_N + \varepsilon\,\lambda\,\chi^2_N$
Cauchy	$c_N\{1+u\}^{-\frac{N+1}{2}}$	$N F_{N,1}$
Student-t	$c_N\{v+u\}^{-\frac{N+v}{2}},\ v > 0$	$N F_{N,v}$
Generalized student-t	$c_N\{\lambda+u\}^{-\frac{N+v}{2}},\ v,\lambda > 0$	$\frac{N\lambda}{v}\,F_{N,v}$
Power exponential	$c_N \exp\{-\frac{1}{2}u^s\},\ s > 0$	$\mathcal{G}^{\frac{1}{s}}(\frac{N}{2s}, \frac{1}{2})$
Kotz type	$c_N u^q \exp\{-\frac{r}{2}u^s\},\ r, s > 0$, $2q + N > 0$	$\mathcal{G}^{\frac{1}{s}}(\frac{2q+N}{2s}, \frac{r}{2})$
Pearson type II	$c_N\{1-u\}^{\frac{\alpha}{2}-1},\ \alpha > 0$	$\mathrm{Beta}(\frac{N}{2}, \frac{\alpha}{2})$
Normal scale mixture	$c_N \int_0^\infty v^{-\frac{N}{2}}\,\phi^{(N)}(\frac{u}{v})\,dF(v)$, F a c.d.f.	$c_N u^{\frac{N}{2}-1} \int_0^\infty v^{-\frac{N}{2}}\,\phi^{(N)}(\frac{u}{v})\,dF(v)$, F a c.d.f.

c_N is just an appropriate constant; $T \sim \mathcal{G}^{\frac{1}{s}}(\alpha,\lambda)$ means that $T^s \sim \mathcal{G}(\alpha,\lambda)$; $\phi^{(N)}(t) = (2\pi)^{-\frac{N}{2}} e^{-\frac{t}{2}}$ is the generator of the n-dimensional normal distribution.

Table B.2
Means and variances of squared radial distributions ($\mathrm{SR}(d;h)$)

Likelihood	$\mathrm{SR}(d;h)$	Mean	Variance
Normal	χ^2_d	d	$2d$
Contaminated normal	$(1-\varepsilon)\chi^2_d + \varepsilon\lambda\chi^2_d$	$1-\varepsilon+\varepsilon\lambda$	$\left\{\frac{d+2}{d}\,\frac{1-\varepsilon+\varepsilon\lambda^2}{(1-\varepsilon+\varepsilon\lambda)^2} - 1\right\}\left\{1-\varepsilon+\varepsilon\lambda\right\}^2$
Cauchy	$d\ F_{d,1}$	—	—
Student-t	$d\ F_{d,v}$	$\frac{v}{v-2},\ v > 2$	$\left\{\frac{d+2}{d}\,\frac{v-2}{v-4} - 1\right\}\left\{\frac{v}{v-2}\right\}^2,\ v > 4$
Generalized student-t	$\frac{d\,\lambda}{v}\ F_{d,v}$	$\frac{\lambda}{v-2}\,\frac{1}{v},\ v > 2$	$\left\{\frac{d+2}{d}\,\frac{v-2}{v-4} - 1\right\}\left\{\frac{\lambda}{v-2}\,\frac{1}{v}\right\}^2,\ v > 2$
Power exponential	$\chi^{\frac{2}{s}}_{\frac{d}{s}}$	$\frac{2^{\frac{1}{s}}\,\Gamma(\frac{d+2}{2s})}{\Gamma(\frac{d}{2s})}$	$\frac{2^{\frac{2}{s}}\left\{\Gamma(\frac{d+4}{2s})\Gamma(\frac{d}{2s}) - \Gamma^2(\frac{d+2}{2s})\right\}}{\Gamma^2(\frac{d}{2s})}$
Kotz type	$r^{\frac{1}{s}}\ \chi^{\frac{2}{s}}_{\frac{2q+d}{s}}$	$\frac{(\frac{2}{r})^{\frac{1}{s}}\,\Gamma(\frac{2q+d+2}{2s})}{\Gamma(\frac{2q+d}{2s})}$	$\frac{(\frac{2}{r})^{\frac{2}{s}}\left\{\Gamma\left(\frac{2q+d+4}{2s}\right)\Gamma\left(\frac{2q+d}{2s}\right) - \Gamma^2\left(\frac{2q+d+2}{2s}\right)\right\}}{\Gamma^2\left(\frac{2q+d}{2s}\right)}$
Pearson type II	$\mathrm{Beta}(\frac{d}{2}, \frac{\alpha+d}{2})$	$\frac{d}{\alpha+2d}$	$\frac{2d(\alpha+d)}{\alpha+2(d+1)}\left\{\frac{1}{\alpha+2d}\right\}^2$

Appendix B. Tabular summary of SSR distributions

This appendix contains two tables related with squared-radial distributions (see Sections 2.4 and 3). Table B.1 shows the squared-radial distributions associated to various spherical distributions, while Table B.2 provides their means and variances (see also [6,13]).

References

[1] R.B. Arellano-Valle, On characterizations of spherical distributions, Statist. Probab. Lett. 54 (2001) 227–232.
[2] R.B. Arellano-Valle, H. Bolfarine, On characterizations of the t distribution, Statist. Probab. Lett. 25 (1995) 79–85.
[3] R.B. Arellano-Valle, M. Galea-Rojas, P. Iglesias, Bayesian sensivity analysis in elliptical linear regression models, J. Statist. Plann. Inference 86 (2000) 175–199.
[4] S. Cambanis, S. Huang, G. Simons, On the theory of elliptical contoured distributions, J. Multivariate Anal. 11 (1981) 368–385.
[5] J.M. Dickey, C.-H. Chen, Direct Subjective-Probability Modelling using Ellipsoidal Distributions, in: J.M. Bernardo, M.H. DeGroot, D.V. Lindley, A.F.M. Smith (Eds.), Bayesian Statistics 2, North-Holland, Amsterdam, 1985, pp. 157–182.
[6] K.T. Fang, S. Kotz, K.W. Ng, Symmetric Multivariate and Related Distributions, Chapman & Hall, London, 1990.
[7] C. Fernández, J. Osiewalski, M. Steel, The continuous multivariate location-scale model revisited: a tale of robustness, Biometrika 81 (1994) 588–594.
[8] C. Fernández, J. Osiewalski, M. Steel, Classical and Bayesian inference robustness in multivariate regression models, J. Amer. Statist. Assoc. 92 (1997) 1434–1444.
[9] V.M. Ng, Robust Bayesian inference for seemingly unrelated regressions with elliptical errors, J. Multivariate Anal. 83 (2002) 409–414.
[10] J. Osiewalski, M. Steel, Robust Bayesian inference in elliptical regression models, J. Econometrics 57 (1993) 345–363.
[11] J. Osiewalski, M. Steel, Bayesian marginal equivalence of elliptical regression models, J. Econometrics 59 (1993) 391–403.
[12] J. Osiewalski, M. Steel, Robust Bayesian inference in l_q-spherical models, Biometrika 80 (1993) 456–460.
[13] J. Osiewalski, M. Steel, Posterior moments of scale parameters in elliptical regression models, in: D. Berry et al. (Eds.), Bayesian Analysis in Statistics and Econometrics. Essays in Honor of Arnold Zellner, Wiley, New York, 1996, pp. 323–335.
[14] A. Zellner, Bayesian and non-Bayesian analysis of the regression model with multivariate student-t error terms, J. Amer. Statist. Assoc. 71 (1976) 400–405.