Journal of Multivariate Analysis 126 (2014) 14–24

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On standard conjugate families for natural exponential families with bounded natural parameter space ☆

Kurt Hornik [a], Bettina Grün [b,*]

[a] Institute for Statistics and Mathematics, WU Wirtschaftsuniversität Wien, Welthandelsplatz 1, 1020 Vienna, Austria

[b] Department of Applied Statistics, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria

ARTICLE INFO

Article history:
Received 19 September 2012
Available online 13 January 2014

AMS subject classifications:
62E10
62F15
62H11

Keywords:
Bayesian analysis
Conjugate prior
Elliptical distribution
Exponential family
Linear posterior expectation
Spherical distribution

ABSTRACT

Diaconis and Ylvisaker (1979) give necessary conditions for conjugate priors for distributions from the natural exponential family to be proper as well as to have the property of linear posterior expectation of the mean parameter of the family. Their conditions for propriety and linear posterior expectation are also sufficient if the natural parameter space is equal to the set of all d-dimensional real numbers. In this paper their results are extended to characterize when conjugate priors are proper if the natural parameter space is bounded. For the special case where the natural exponential family is through a spherical probability distribution η, we show that the proper conjugate priors can be characterized by the behavior of the moment generating function of η at the boundary of the natural parameter space, or the second-order tail behavior of η. In addition, we show that if these families are non-regular, then linear posterior expectation never holds. The results for this special case are also extended to natural exponential families through elliptical probability distributions.

1. Introduction

Let η be a σ-finite measure on $\mathbb{R}^d$, and consider the natural exponential family (NEF) $\mathcal{F}$ through η, with densities

$$f(x|\theta) = e^{\theta' x - M(\theta)}$$

with respect to η, where the cumulant generating function $M(\theta)$ is defined by

$$e^{M(\theta)} = \int_{\mathbb{R}^d} e^{\theta' x} \, d\eta(x)$$

(e.g., [1]). Let $\Theta = \{\theta : M(\theta) < \infty\}$ be the natural parameter space of $\mathcal{F}$. The family of standard conjugate distributions for $\mathcal{F}$ (relative to the natural parameter) has densities

$$\pi(\theta|s, \nu) \propto e^{s'\theta - \nu M(\theta)}$$

with respect to the Lebesgue measure on Θ (e.g., [6]).

* Corresponding author.
E-mail addresses: Kurt.Hornik@wu.ac.at (K. Hornik), Bettina.Gruen@jku.at (B. Grün).

0047-259X/$ – see front matter © 2014 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmva.2014.01.003

Let $J(s, \nu) = \int_{\Theta} e^{s'\theta - \nu M(\theta)} d\theta$. Then the hyperparameters s and ν giving proper standard conjugate distributions are the ones for which $J(s, \nu) < \infty$. For Bayesian inference on θ it is natural to employ priors from the standard conjugate family, and it is important to know when these are proper, or yield proper posteriors.

For regular NEFs (i.e., Θ is open) where Θ is non-empty, Diaconis and Ylvisaker [4] show in Theorem 1 that if $\mathcal{X}$, the interior of the convex hull of the support of η, is a non-empty open set in $\mathbb{R}^d$, then $J(s, \nu) < \infty$ if $\nu > 0$ and $s/\nu \in \mathcal{X}$, and conversely, if $\Theta = \mathbb{R}^d$, then $J(s, \nu) < \infty$ implies that $\nu > 0$ and $s/\nu \in \mathcal{X}$. (Note that the reference uses νs where we use s.) This gives a complete characterization of all proper conjugate distributions for the case $\Theta = \mathbb{R}^d$, leaving open the cases where $\Theta \subset \mathbb{R}^d$.

In this paper, we prove that if Θ is *bounded*, there exists $-\infty \le \nu_0 \le 0$ such that for arbitrary s, the conjugate priors with hyperparameters ν and s are proper for $\nu > \nu_0$ and improper for $\nu < \nu_0$. We provide examples showing that all values for ν_0 in the range $-\infty \le \nu_0 \le 0$ are possible.

More specific results are obtained when η is a (non-degenerate) spherical probability distribution on $\mathbb{R}^d$, i.e., a distribution invariant to orthogonal transformations. In this case, $\mathcal{X}$ is of the form $\{x : \|x\| < \sigma\}$, where σ is finite if and only if η has bounded support, and Θ is an open or closed ball with radius ρ for some $0 < \rho \le \infty$. For $\rho = \infty$, $\Theta = \mathbb{R}^d$, and the result of Theorem 1 in Diaconis and Ylvisaker [4] yields that the hyperparameters giving proper conjugate priors are those for which $\nu > 0$ and $\|s\| < \sigma\nu$. For $\rho < \infty$, our characterization applies, and we show that lower (and/or upper) bounds for ν_0 can be derived if the behavior of the moment generating function of η at the boundary of the natural parameter space can be characterized via asymptotic lower (and/or upper) bound functions. In addition we establish that ν_0 can be related to the "second order tail behavior" of η.

If $\theta \in \text{int}(\Theta)$, $\mu(\theta) = \nabla M(\theta) = \int_{\mathbb{R}^d} x f(x|\theta) d\eta(x)$ is the mean parameter of the NEF. Diaconis and Ylvisaker [4] show in Theorem 2 that if Θ is open and θ has a distribution which corresponds to a proper conjugate prior with hyperparameters s and ν satisfying $s/\nu \in \mathcal{X}$ and $\nu > 0$, then $\mathbb{E}(\nabla M(\theta)) = s/\nu$. Clearly, in this case the posterior from an observation x is a conjugate distribution with parameters $s + x$ and $\nu + 1$, so that $\mathbb{E}(\nabla M(\theta)|x) = (s + x)/(\nu + 1)$ is linear in x. For NEFs through a spherical probability distribution with bounded Θ, we show that $\mathbb{E}(\nabla M(\theta))$ does not exist for $\nu \le 0$ if Θ is open, and exists for all s and ν if Θ is closed, where in this case $\mathbb{E}(\nabla M(\theta)) \ne s/\nu$ unless $s = 0$ and $\nu \ne 0$. Finally, we show that if Θ is closed, linear posterior expectation never holds when using canonical priors.

These results for η a (non-degenerate) spherical probability distribution on $\mathbb{R}^d$ are extended to the case of elliptical distributions as given in Fang et al. [5, p. 31f]. We show that propriety of conjugate priors is only possible if the matrix in the linear transformation is a square matrix of full rank and that if the natural parameter space is bounded the value of ν_0 and the characterization of propriety for $\nu = \nu_0$ are the same as for the corresponding spherical probability distribution. Similarly linear posterior expectation only holds in the regular case for $\nu > 0$ and never holds in the non-regular case when canonical priors are used.

2. General NEFs with bounded natural parameter space

We first establish a general result on the propriety of conjugate priors for NEFs with bounded natural parameter space.

Theorem 1. *Let η be a σ-finite measure on $\mathbb{R}^d$. Suppose the natural parameter space Θ of the NEF through η is bounded and non-empty. Then $\nu \ge 0$ and arbitrary s give proper conjugate distributions, and there exists $\nu_0 = \nu_0(\eta)$ with $-\infty \le \nu_0 \le 0$ such that for arbitrary s, the conjugate distributions with parameters s and ν are proper for $\nu > \nu_0$, and improper for $\nu < \nu_0$.*

Proof. If Θ is bounded, then clearly $J(s, \nu) < \infty$ if and only if $K(\nu) = \int_{\Theta} e^{-\nu M(\theta)} d\theta < \infty$, and $K(0) = \int_{\Theta} d\theta < \infty$. To establish the theorem, it suffices to show that if $K(\nu_1) < \infty$, then $K(\nu) < \infty$ for all $\nu > \nu_1$. Now M is convex; the assumptions on Θ are readily seen to imply that M is *proper* in the sense of Rockafellar [9]. By Corollary 12.1.2 of Rockafellar [9], there are $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}$ such that $M(\theta) \ge x'\theta + \alpha$ for all θ (i.e., M can be bounded below by a hyperplane). Hence, writing $\gamma = |\alpha| + \|x\| \sup_{\theta \in \Theta} \|\theta\| < \infty$, $-M(\theta) \le \gamma$ for all $\theta \in \Theta$.

Now suppose $K(\nu_1)$ is finite and $\nu > \nu_1$. Clearly, for all $\theta \in \Theta$, $-\nu M(\theta) = -\nu_1 M(\theta) + (\nu - \nu_1)(-M(\theta)) \le -\nu_1 M(\theta) + (\nu - \nu_1)\gamma$ so that

$$K(\nu) \le \int_{\Theta} e^{-\nu_1 M(\theta)} e^{(\nu-\nu_1)\gamma} d\theta = e^{(\nu-\nu_1)\gamma} K(\nu_1)$$

and hence $K(\nu)$ is finite as well. Taking $\nu_0 = \inf\{\nu : K(\nu) < \infty\}$, the proof is complete. □

Remark. In contrast to the case where the natural parameter space is equal to $\mathbb{R}^d$, negative values of ν also give proper prior distributions. In this case the parameter ν cannot be interpreted as a prior sample size. Furthermore, the mean for the prior distribution does not necessarily exist as indicated in the example given by Diaconis and Ylvisaker [4, p. 275].

Remark. If Θ is not bounded, $J(s, \nu_1) < \infty$ does not necessarily imply that $J(s, \nu) < \infty$ for all $\nu \ge \nu_1$. This can straightforwardly be seen for $\Theta = \mathbb{R}^d$, taking, e.g., η to have the density with respect to the Lebesgue measure given by $f(x) \propto e^{-\|x\|^2}$ for $\min(x) := \min(x_1, \dots, x_d) > 1$, and zero otherwise. Then clearly $\Theta = \mathbb{R}^d$ and $\mathcal{X} = \{x : \min(x) > 1\}$. By Theorem 1 of Diaconis and Ylvisaker [4], $J(s, \nu) < \infty$ if and only if $\nu > 0$ and $s/\nu \in \mathcal{X}$, or equivalently, if and only if $0 < \nu < \min(s)$.