On estimation problems for multivariate skew-symmetric distributions

Adelchi Azzalini

University of Padua, Italy, email: azzalini@stat.unipd.it

Keywords: anomalies of maximum likelihood estimation, skew-symmetric distributions

A currently active stream of literature deals with continuous multivariate distributions whose density function is the form

$$f(x) = 2\, f_0(x)\, G\{w(x)\}, \qquad\qquad x \in \mathbb{R}^d,$$

where f_0 is a density function such that $f_0(x) = f_0(-x)$, G is a distribution function on the real line such that G' exists and is an even function, and w is odd in the sense that $w(-x) = -w(x) \in \mathbb{R}$. The term 'skew-symmetric' is often used to refer to a density $f(x)$ of this type, although the effect of perturbation of the symmetric density $f_0(x)$ by the factor $G\{w(x)\}$ can be more complex than turning it into an asymmetric distribution.

Two important special cases of this construction are the so-called skew-normal and the skew-t distribution, which are obtained by choosing the ingredients as follows:

f_0	$G(w)$	$w(x)$
$N_d(0,\Omega)$ density	$\Phi(w)$	$\alpha^\top \omega^{-1} x$
$t_d(\nu,\Omega)$ density	$T(w,\nu+d)$	$\alpha^\top \omega^{-1} x \left(\dfrac{\nu+d}{\nu + x^\top \Omega^{-1} x}\right)^{1/2}$

where a standard type of notation is adopted, and ω is a diagonal matrix whose non-null terms are the standard deviations associated to the variance matrix Ω.

While the probability side of this formulation leads to a smooth mathematical development, and several nice properties follow with relatively little effort, its statistics side has shown to be more challenging. More specifically, maximum likelihood estimation (MLE) of the above parameters Ω, α and ν, when this is present, complemented by a location parameter, can exhibit two sort of anomalies:

(i) the observed and the expected information matrices are singular at $\alpha = 0$ for certain families, in particular for the skew-normal family indicated above,

(ii) for finite sample size, the MLE of α may happen to diverge with non-null probability.

We shall first review the state of the art for this estimation problem, and then focus on its case (ii) which so far has not yet been given a satisfactory general solution. A proposal based on a form of penalized likelihood function will be put forward.