SUPPLEMENT: Further properties of SNI distributions

From Proposition 1 in [20], the SN distribution defined in (1) has a convenient stochastic representation:

$$\mathbf{Y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}|\mathbf{T}_1| + \boldsymbol{\Sigma}^{1/2}\mathbf{T}_2, \tag{S-1}$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ are two independent $N_p(0, \mathbf{I}_p)$ random vectors. Here, $|.|$ denotes the absolute value. If $\mathbf{Y} \sim SNI_{p,p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}, H)$, the mean vector and the covariance matrix of a SNI random vector are

$$E\{\mathbf{Y}\} = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}, \quad Var\{\mathbf{Y}\} = E\{U^{-1}\}\left(\boldsymbol{\Omega} + \frac{2}{\pi}(\boldsymbol{\lambda}\boldsymbol{\lambda}^\top - \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top)\right) - \frac{2}{\pi}E^2\{U^{-1/2}\}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top.$$

Some members of the SNI class follows.

(i) *Multivariate skew-normal (SN) distribution.* This is the case when $U = 1$ (a degenerate random variable) in (3).

(ii) *Multivariate skew-t (ST) distribution.* It is derived from (3) by taking $U \sim Gamma(\nu/2, \nu/2)$, $\nu > 0$ and is denoted as $St_{p,p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}, \nu)$. It follows from Proposition 1 given in [8] that the pdf of $\mathbf{Y}$ is:

$$f(\mathbf{y}) = 2^p\, t_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Omega}, \nu) T_p\left(\sqrt{\frac{p+\nu}{d+\nu}}\mathbf{A}; \boldsymbol{\Delta}, \nu + p\right), \quad \mathbf{y} \in \mathbb{R}^p, \tag{S-2}$$

where $\mathbf{A} = \boldsymbol{\Lambda}^\top\boldsymbol{\Omega}^{-1}(\mathbf{y} - \boldsymbol{\mu})$ and $d = (\mathbf{Y} - \boldsymbol{\mu})^\top\boldsymbol{\Omega}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$ is the Mahalanobis distance, $t_p(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ denotes the p-dimensional multivariate Student-t distribution with location $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and degrees of freedom (df) ν, and $T_p(\cdot; \boldsymbol{\Sigma}; \nu)$ is the cdf of $t_p(\cdot; \mathbf{0}, \boldsymbol{\Sigma}, \nu)$. A particular case of the skew–t distribution is the skew–Cauchy distribution, when $\nu = 1$. Also, when $\nu \uparrow \infty$, we have the SN distribution as the limiting case. Applications of the ST distribution to robust estimation can be found in [8, 18].

(iii) *Multivariate skew-slash (SSL) distribution.* It is derived from (3), choosing $U \sim Beta(\nu, 1)$, $\nu > 0$. It is denoted by $SSL_{p,p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}, \nu)$ and the p.d.f is given by

$$f(\mathbf{y}) = 2^p\nu \int_0^1 u^{\nu-1}\phi_p(\mathbf{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Omega})\Phi_p(u^{1/2}\mathbf{A}; \boldsymbol{\Delta})du, \quad \mathbf{y} \in \mathbb{R}^p. \tag{S-3}$$

. The SL distribution reduces to the SN distribution when $\nu \uparrow \infty$.

(iv) *Multivariate skew contaminated normal (SCN) distribution.* This arises when the mixing scale factor U is a discrete random variable taking one of two states, i.e. either ν_2 or 1, with $\boldsymbol{\nu} = (\nu_1, \nu_2)^\top$. It is denoted by $SCN_{p,p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}, \nu_1, \nu_2)$. The probability function of U is

$$h(u|\boldsymbol{\nu}) = \nu_1\mathbb{I}_{\{\nu_2\}}(u) + (1 - \nu_1)\mathbb{I}_{\{1\}}(u), \quad 0 < \nu_1 < 1,\ 0 < \nu_2 \leq 1. \tag{S-4}$$

It then follows that

$$f(\mathbf{y}) \quad = \quad 2^p\left\{\nu_1\phi_p(\mathbf{y}; \boldsymbol{\mu}, \nu_2^{-1}\boldsymbol{\Omega})\Phi_p(\nu_2^{1/2}\mathbf{A}; \boldsymbol{\Delta}) + (1 - \nu_1)\phi_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Omega})\Phi_p(\mathbf{A}; \boldsymbol{\Delta})\right\}.$$

Parameter ν_1 can be interpreted as the proportion of outliers while ν_2 may be interpreted as a scale factor. The SCN distribution reduces to the SN distribution when $\nu_2 = 1$.