Supplemental Online Material for "A heteroscedastic measurement error model based on skew and heavy-tailed distributions with known error variances"

ARTICLE HISTORY
Compiled October 8, 2017

ABSTRACT
Supplemental Online Material contains the proofs of Propositions 2.1 and 2.2 as well as the details of the E step of the ECM algorithm in Section 2.3. Finally, adaptations of the ECM algorithm to fit the particular models used in Section 3 are briefly presented.

Supplemental Online Material A. Proofs

Proof of Proposition 2.1 The result in (a) follows from (5) and (4). For (b), the result follows directly from (a) upon applying the law of iterated expectations and using well-known results about moments involving normal and gamma random variables. $\square$

It is worth mentioning that expressions such as $\mathbf{\Pi}_{v_i}$ and λ_{v_i} defined in (6) were adapted from Corollary 2 of Theorem 1 in [1]. Moreover, other results found in the literature were also applied for the measurement error model framework in (2) in order to derive the marginal pdf of $\boldsymbol{Z}_i$. These results will be cited whenever needed.

Proof of Proposition 2.2 First, consider the expressions in (6). For the result in (a), we use (5) to write $\boldsymbol{Z}_i$ as

$$\boldsymbol{Z}_i|x_i, W_{e_i} \sim \mathcal{N}_2(\boldsymbol{a} + \boldsymbol{b}x_i, \boldsymbol{\Sigma}/W_{e_i}) \quad \text{and} \quad x_i|W_{x_i} \sim \mathcal{SN}(\xi, \omega^2/W_{x_i}, \lambda),$$

where $W_{x_i} \sim \mathcal{G}\,(1/2\eta_x, 1/2\eta_x)$ and $W_{e_i} \sim \mathcal{G}(1/2\eta_e\,, 1/2c(\eta_e))$ with $c(\eta_e) = \eta_e/(1-2\eta_e)$. The result now holds from an application of Corollary 2 in [1] and after transforming $(W_{x_i}, W_{e_i})^\top$ to $(U_i = W_{x_i}, V_i = W_{x_i}/W_{e_i})^\top$. For (b), we use (a) to write the joint pdf of $(\boldsymbol{Z}_i^\top, U_i, V_i)^\top$ as $f(\boldsymbol{z}_i, u_i, v_i; \boldsymbol{\theta}) = f(\boldsymbol{z}_i|u_i, v_i; \boldsymbol{\theta}) f(u_i, v_i; \boldsymbol{\theta})$ and after simplification, the result follows. For (c), the result is obtained by integrating $f(\boldsymbol{z}_i, u_i, v_i; \boldsymbol{\theta})$ with respect to u_i by using Lemma 1 in [2]. Finally, for (d), we integrate out v_i from $f(\boldsymbol{z}_i, v_i|\boldsymbol{\theta})$ given in (c). $\square$

Supplemental Online Material B. Computing expectations in the E step

In this section, we describe how to compute all the expectations involved in (14). We omit the subscript of each random variable to simplify the notation. Next, our strategy is to obtain the joint distribution of the missing vector $(x, D, U, V)^\top$ conditional on the observed $\boldsymbol{Z}$ and then, use it to derive the desired expectations in the E step of the ECM algorithm. Define, for $v > 0$,

$$\boldsymbol{\Omega}_v = \psi \boldsymbol{b}\boldsymbol{b}^\top + v\boldsymbol{\Sigma},\ \rho_v = (\psi^{-1} + \boldsymbol{b}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{b}/v)^{-1} \text{ and } \zeta_v^2 = (1 + \gamma^2 \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}\boldsymbol{b})^{-1}. \quad \text{(S1)}$$

Using well-known matrix results [3, p. 467], one can verify that $\boldsymbol{\Omega}_v$ and ρ_v are related by $\rho_v = \psi - \psi^2 \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}\boldsymbol{b}$ and $\rho_v \boldsymbol{b}^\top \boldsymbol{\Sigma}^{-1}/v = \psi \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}$. In addition, ξ_v given in (6) can be written in terms of (S1) as $\xi_v = \zeta_v \gamma \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}\boldsymbol{\Delta}$.

Proposition S1 Consider $\boldsymbol{Z} \in \mathbb{R}^2$ as in (2), (S1) and the quantities given in Proposition 2.2.

(a) The conditional pdf of $V|\boldsymbol{Z}$ is $f(v|\boldsymbol{z};\boldsymbol{\theta}) = f(\boldsymbol{z}, v;\boldsymbol{\theta})/f(\boldsymbol{z};\boldsymbol{\theta})$, $v > 0$.
(b) The conditional pdf of $U|\boldsymbol{Z}, V = v$ is $f(u|\boldsymbol{z}, v;\boldsymbol{\theta}) = f(\boldsymbol{z}, u, v;\boldsymbol{\theta})/f(\boldsymbol{z}, v;\ \boldsymbol{\theta})$, $u, v > 0$.
(c) $D|\boldsymbol{Z}, U = u, V = v \sim \mathcal{TN}(\xi_v\zeta_v, \zeta_v^2/u; (0,\infty))$.
(d) $x|\boldsymbol{Z}, D = d, U = u, V = v \sim \mathcal{N}(\xi + \psi \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}(\boldsymbol{Z} - \boldsymbol{a} - \boldsymbol{b}\xi) + \rho_v \psi^{-1}\gamma d,\ \rho_v/u)$.

Proof. For (a) and (b), we apply the definition of conditional pdf and using the density functions in (a), (b) and (c) given in Proposition 2.2, the results hold. For (c), from (7) we obtain $\boldsymbol{Z}|D = d, U = u, V = v \sim \mathcal{N}_2(\boldsymbol{a} + \boldsymbol{b}\xi + \boldsymbol{b}\gamma d, \boldsymbol{\Omega}_v/u)$ and $D|U = u, V = v \sim \mathcal{TN}(0, 1/u; (0, +\infty)$. Moreover, by applying the definition of conditional pdf to $D|\boldsymbol{Z}, U, V$, it follows that $f(d|\boldsymbol{z}, u, v;\boldsymbol{\theta}) \propto f(\boldsymbol{z}|d, u, v;\boldsymbol{\theta}) f(d|u, v;\boldsymbol{\theta})$ and by applying Lemma 2 in [1] to the product on the right side, we obtain $f(d|\boldsymbol{z}, u, v;\boldsymbol{\theta}) \propto \phi_2(\boldsymbol{z};\boldsymbol{a} + \boldsymbol{b}\xi, (\Omega_v + \gamma^2 \boldsymbol{b}\boldsymbol{b}^\top)/u) 2\phi(d;\xi_v\zeta_v, \zeta_v^2/u) I(d > 0)$. Then, we can see that the pdf of $D|\boldsymbol{Z}, U, V$ is proportional to the pdf of the $\mathcal{TN}(\xi_v\zeta_v, \zeta_v^2/u; (0,\infty))$ distribution evaluated at d. Therefore, this establishes the desired conditional distribution. For (d), from a similar manner to (c), we use (7) to obtain $\boldsymbol{Z}|x, D = d, U = u, V = v \sim \mathcal{N}_2(\boldsymbol{a} + \boldsymbol{b}x, v\boldsymbol{\Sigma}/u)$ and $x|D = d, U = u \sim \mathcal{N}(\xi + \gamma d, \psi/u)$. By applying the definition of conditional pdf to $x|\boldsymbol{Z}, D, U, V$, it follows that $f(x|\boldsymbol{z}, d, u, v;\boldsymbol{\theta}) \propto f(\boldsymbol{z}|x, d, u, v;\boldsymbol{\theta}) f(x|d, u;\boldsymbol{\theta})$. Then, using again Lemma 2 in [1], we obtain that $f(x|\boldsymbol{z}, d, u, v;\boldsymbol{\theta}) \propto \phi_2(\boldsymbol{Z};\boldsymbol{a} + \boldsymbol{b}\xi + \boldsymbol{b}\gamma d, \Omega_v/u)$ $\phi(x;\xi + \gamma d + \rho_v \boldsymbol{b}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{Z} - \boldsymbol{a} - \boldsymbol{b}\xi - \boldsymbol{b}\gamma d)/v, \rho_v/u)$. Using (S1), we simplify the mean in the second pdf. Thus, we notice that the pdf of $x|\boldsymbol{Z}, D, U, V$ is proportional to the pdf of the $\mathcal{N}(\xi + \psi \boldsymbol{b}^\top \boldsymbol{\Omega}_v^{-1}(\boldsymbol{Z} - \boldsymbol{a} - \boldsymbol{b}\xi) + \rho_v\psi^{-1}\gamma d,\ \rho_v/u)$ distribution evaluated at x. Therefore, it follows that this must also be the conditional distribution of $x|\boldsymbol{Z}, D, U, V$ and the result holds. □

Proposition S2 Consider $\boldsymbol{Z} \in \mathbb{R}^2$ as defined in (2), (6) and the quantities in Proposition S1. Then, we have the following conditional expectations on $\boldsymbol{Z}$ and V.

(a) For an integer k such that $\eta^* + 2k > 0$,

$$E[U^k|\boldsymbol{Z}, V] = 2^k\ \frac{\Gamma((\eta^* + 2k)/2)}{\Gamma(\eta^*/2)} \frac{F_t(\xi_V[(\eta^* + 2k)/\varsigma_V]^{1/2};\eta^* + 2k)}{\varsigma_V^k F_t(\xi_V[\eta^*/\varsigma_V]^{1/2};\eta^*)}.$$

(b) For an integer k such that $\eta^* + k > 0$,

$$E\left[U^{k/2}\frac{\phi(\xi_V U^{1/2})}{\Phi(\xi_V U^{1/2})}|\boldsymbol{Z}, V\right] = \frac{2^{(k-1)/2}}{\pi^{1/2}}\frac{\Gamma((\eta^*+k)/2)}{\Gamma(\eta^*/2)} \times \frac{\varsigma_V^{\eta^*/2}}{(\varsigma_V + \xi_V^2)^{(\eta^*+k)/2}\, F_t(\xi_V[\eta^*/\varsigma_V]^{1/2}; \eta^*)}.$$

(c) For an integer k such that $\eta^* + 2k > 1$,

$$E[DU^k|\boldsymbol{Z}, V] = \zeta_V\left\{\xi_V E[U^k|\boldsymbol{Z}, V] + E\left[U^{(2k-1)/2}\frac{\phi(\xi_V U^{1/2})}{\Phi(\xi_V U^{1/2})}|\boldsymbol{Z}, V\right]\right\}.$$

(d) For an integer k such that $\eta^* + 2k > 2$,

$$E[D^2U^k|\boldsymbol{Z}, V] = \zeta_V^2 E[U^{k-1}|\boldsymbol{Z}, V] + \xi_V\zeta_V E[DU^k|\boldsymbol{Z}, V].$$

(e) $E[xDU|\boldsymbol{Z}, V] = \big(\xi + \psi\boldsymbol{b}^\top\boldsymbol{\Omega}_V^{-1}\boldsymbol{\Delta}\big)E[DU|\boldsymbol{Z}, V] + \rho_V\psi^{-1}\gamma E[D^2U|\boldsymbol{Z}, V]$.

(f) $E[xU|\boldsymbol{Z}, V] = \big(\xi + \psi\boldsymbol{b}^\top\boldsymbol{\Omega}_V^{-1}\boldsymbol{\Delta}\big)E[U|\boldsymbol{Z}, V] + \rho_V\psi^{-1}\gamma E[DU|\boldsymbol{Z}, V]$.

(g) $E[x^2U|\boldsymbol{Z}, V] = \tau^* E[xU|\boldsymbol{Z}, V] + \big\{1 + \psi^{-1}\gamma E[xDU|\boldsymbol{Z}, V]\big\}\rho_V$, where $\tau^* = \xi + \boldsymbol{\Delta}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{b}\rho_V/V$.

Proof. For (a), from the definition of expectation, we can write $E[U^k|\boldsymbol{z}, v] = \int_0^\infty u^k f(u|\boldsymbol{z}, v; \boldsymbol{\theta})du$. Then, from Proposition S1 and by applying Lemma 1 in [2], we have after lengthy algebra that the result follows. For (b), as in (a), we use again the definition of conditional expectation and Proposition S1. After some algebraic manipulations, the result follows by making use of the gamma integral. For (c), using the law of iterated expectations we can write

$$E[DU^k|\boldsymbol{z}, v] = E[U^k\ E[D|\ \boldsymbol{z}, v, U]|\ \boldsymbol{z}, v]. \tag{S2}$$

Using part (c) of Proposition S1 and properties about moments of a truncated normal random variable [4, Section 10.1], we can write the inner expectation as $E[D|\boldsymbol{z}, v, u] = \zeta_v[\xi_v + u^{-1/2}\ \phi(\xi_v u^{1/2})/\Phi(\xi_v u^{1/2})]$. The result now follows after substituting this expression into (S2). For (d)–(g), we proceed in a similar manner as in (c). Using Proposition S1, the result in [4, Section 10.1] and the law of the iterated expectations along with $\boldsymbol{\Omega}_v^{-1}\boldsymbol{b}\psi = \boldsymbol{\Sigma}^{-1}\boldsymbol{b}\rho_v/v$, the results follow after simplifications. □

Finally, to compute expectations required in the E step of the ECM algorithm, first we compute the expectation conditional on $(\boldsymbol{Z}^\top, V)^\top$ using Proposition S2 and then average it over the conditional distribution of $(V|\boldsymbol{Z})$. This computation requires a unidimensional numerical integration and here it was computed using the `statmod` package [5] implemented in the R language [6]. All the expectations are evaluated at $\boldsymbol{\theta}^* = \boldsymbol{\theta}^{*(r)}$.

Supplemental Online Material C. Adapting the ECM algorithm for some particular models

In this section, we briefly point out how the ECM algorithm was modified to fit each one of the particular models used in this work. According to Section 2.2, we have the following models: (a) STN MEM, (b) TcT MEM, (c) TN MEM and (d) NN MEM. Recall that the values of η_x and η_e were fixed according to each specific model. In case (a), the gamma distribution for $W_e = U/V$ is not needed, so that we can apply the algorithm after setting $U/V \equiv 1$. In case (b), the algorithm works by setting $\gamma = 0$ in the expected log-likelihood given in (13) and omitting the CM step 3. In case (c), the modifications we need are as in cases (a) and (b) and can be applied after modifying the expected log-likelihood in (13). Finally, in case (d), the gamma distribution for $W_x = U$ is not required, so that we can apply the ECM algorithm after setting $U \equiv 1$ and by using the modifications as in cases (a) and (b).

References

[1] Arellano-Valle RB, Bolfarine H, Lachos VH. Skew-normal linear mixed models. Journal of Data Science. 2005;3:415–438.

[2] Azzalini A, Capitanio A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. Journal of the Royal Statistical Society, Series B. 2003;65:367–389.

[3] Seber GAF, Lee AJ. Linear regression analysis. New York: Wiley; 2003.

[4] Johnson NL, Kotz S, Balakrishnan N. Continuous univariate distributions. 2nd ed.; Vol. 1. New York: Wiley; 1994.

[5] Giner G, Smyth GK. statmod: Probability calculations for the inverse gaussian distribution. The R Journal. 2016;8:339–351.

[6] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2017.