Noname manuscript No.
(will be inserted by the editor)

Small area estimation under a measurement error bivariate Fay-Herriot model

Supplementary file with appendices

Jan Pablo Burgard · María Dolores Esteban · Domingo Morales · Agustín Pérez

the date of receipt and acceptance should be inserted later

Appendix A: Auxiliary results

The following lemma gives a computationally efficient formula for Ψ_d that requires inverting a 2×2 matrix, instead of inverting a $p \times p$ matrix.

Lemma A.1. It holds that

$$\Psi_d = \Sigma_d - \Sigma_d B_d'(V_{\lambda d} + V_{ud} + V_{ed})^{-1} B_d \Sigma_d.$$

Proof. We apply the inverting formula

$$(A + TCD)^{-1} = A^{-1} - A^{-1}T(C^{-1} + DA^{-1}T)^{-1}DA^{-1},$$

with $A = \Sigma_d^{-1}$, $T = B_d'$, $C = (V_{ud} + V_{ed})^{-1}$ and $D = B_d$. We obtain

$$\Psi_d = \Big(B_d'\big(V_{ud} + V_{ed}\big)^{-1} B_d + \Sigma_d^{-1}\Big)^{-1} = \Sigma_d - \Sigma_d B_d'\big(V_{ud} + V_{ed} + B_d \Sigma_d B_d'\big)^{-1} B_d \Sigma_d.$$

As $V_{\lambda d} = B_d \Sigma_d B_d'$, the result follows. □

Propositions A.1-A.3 derive some expectation of cross products. They are later needed to calculate the mean squared error of the BP of μ_d.

Supported by the Spanish grant PGC2018-096840-B-I00 and by the grant "Algorithmic Optimization (ALOP) - graduate school 2126" funded by the German Research Foundation.

Jan Pablo Burgard
Research Institute for Official and Survey Statistics (RIFOSS), Trier University, Germany.
E-mail: burgardj@uni-trier.de

María Dolores Esteban and Domingo Morales
Operations Research Center, University Miguel Hernández de Elche, Spain.

Agustín Pérez
Departments of Economic and Financial Studies, University Miguel Hernández de Elche, Spain.

Proposition A.1. Under model (4), the covariance matrices between u_d, v_d, $\hat{u}_d^{bp}$ and $\hat{v}_d^{bp}$ are

$$
\begin{aligned}
E\big[u_d \hat{u}_d^{bp\,\prime}|x_d\big] &= V_{ud}(V_{\lambda d}+V_{ed})^{-1}\Phi_d, \quad E\big[v_d \hat{u}_d^{bp\,\prime}|x_d\big] = \Sigma_d B_d'(V_{\lambda d}+V_{ed})^{-1}\Phi_d,\\
E\big[\hat{v}_d^{bp} \hat{u}_d^{bp\,\prime}|x_d\big] &= \Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d,\\
E\big[u_d \hat{v}_d^{bp\,\prime}|x_d\big] &= V_{ud}(V_{ud}+V_{ed})^{-1}B_d\Psi_d, \quad E\big[v_d \hat{v}_d^{bp\,\prime}|x_d\big] = \Sigma_d B_d'(V_{ud}+V_{ed})^{-1}B_d\Psi_d.
\end{aligned}
$$

Proof. We define $\xi_d = y_d - Z_d\eta - X_d\lambda = B_d v_d + u_d + e_d$. It holds that

$$
E\big[u_d\xi_d'|x_d\big] = V_{ud},\ E\big[v_d\xi_d'|x_d\big] = \Sigma_d B_d',\ E\big[\xi_d\xi_d'|x_d\big] = V_{\lambda d}+V_{ud}+V_{ed}.
$$

The covariance matrices are

$$
\begin{aligned}
E\big[u_d \hat{u}_d^{bp\,\prime}|x_d\big] &= E\big[u_d\xi_d'|x_d\big](V_{\lambda d}+V_{ed})^{-1}\Phi_d = V_{ud}(V_{\lambda d}+V_{ed})^{-1}\Phi_d,\\
E\big[v_d \hat{u}_d^{bp\,\prime}|x_d\big] &= E\big[v_d\xi_d'|x_d\big](V_{\lambda d}+V_{ed})^{-1}\Phi_d = \Sigma_d B_d'(V_{\lambda d}+V_{ed})^{-1}\Phi_d,\\
E\big[\hat{v}_d^{bp} \hat{u}_d^{bp\,\prime}|x_d\big] &= \Psi_d B_d'(V_{ud}+V_{ed})^{-1}E\big[\xi_d\xi_d'|x_d\big](V_{\lambda d}+V_{ed})^{-1}\Phi_d\\
&= \Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d,\\
E\big[u_d \hat{v}_d^{bp\,\prime}|x_d\big] &= E\big[u_d\xi_d'|x_d\big](V_{ud}+V_{ed})^{-1}B_d\Psi_d = V_{ud}(V_{ud}+V_{ed})^{-1}B_d\Psi_d,\\
E\big[v_d \hat{v}_d^{bp\,\prime}|x_d\big] &= E\big[v_d\xi_d'|x_d\big](V_{ud}+V_{ed})^{-1}B_d\Psi_d = \Sigma_d B_d'(V_{ud}+V_{ed})^{-1}B_d\Psi_d.
\end{aligned}
$$

Similarly, we obtain the corresponding transpose matrices. □

Proposition A.2. Under model (4), it holds that

$$
\begin{aligned}
E[\hat{\mu}_d^{bp}|x_d] &= Z_d\eta + X_d\lambda,\\
E\big[\hat{\mu}_d^{bp}\hat{\mu}_d^{bp\,\prime}|x_d\big] &= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)'\\
&+ B_d\Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'\\
&+ \Phi_d(V_{\lambda d}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d\\
&+ B_d\Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d\\
&+ \Phi_d(V_{\lambda d}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'.
\end{aligned}
$$

Proof. As $\hat{\mu}_d^{bp} = Z_d\eta + X_d\lambda + B_d\hat{v}_d^{bp} + \hat{u}_d^{bp}$ and $E[\hat{u}_d^{bp}|x_d] = E[\hat{v}_d^{bp}|x_d] = 0$, by applying Proposition A.1 we get

$$
\begin{aligned}
E\big[\hat{\mu}_d^{bp}\hat{\mu}_d^{bp\,\prime}|x_d\big] &= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)'\\
&+ B_d E\big[\hat{v}_d^{bp}\hat{v}_d^{bp\,\prime}|x_d\big]B_d' + E\big[\hat{u}_d^{bp}\hat{u}_d^{bp\,\prime}|x_d\big] + B_d E\big[\hat{v}_d^{bp}\hat{u}_d^{bp\,\prime}|x_d\big] + E\big[\hat{u}_d^{bp}\hat{v}_d^{bp\,\prime}|x_d\big]B_d'\\
&= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)'\\
&+ B_d\Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'\\
&+ \Phi_d(V_{\lambda d}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d\\
&+ B_d\Psi_d B_d'(V_{ud}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{\lambda d}+V_{ed})^{-1}\Phi_d\\
&+ \Phi_d(V_{\lambda d}+V_{ed})^{-1}(V_{\lambda d}+V_{ud}+V_{ed})(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'. \quad \square
\end{aligned}
$$

Proposition A.3. Under model (4), the cross-product matrix between $\hat{\mu}_d^{bp}$ and μ_d is

$$
\begin{aligned}
E\big[\mu_d \hat{\mu}_d^{bp\,\prime}|x_d\big] &= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)' \\
&+ B_d\Sigma_d B_d'(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d' + V_{ud}(V_{\lambda d}+V_{ed})^{-1}\Phi_d \\
&+ B_d\Sigma_d B_d'(V_{\lambda d}+V_{ed})^{-1}\Phi_d + V_{ud}(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'.
\end{aligned}
$$

Proof. As $\mu_d = Z_d\eta + X_d\lambda + B_d v_d + u_d$, $\hat{\mu}_d^{bp} = Z_d\eta + X_d\lambda + B_d\hat{v}_d^{bp} + \hat{u}_d^{bp}$, $E[\hat{u}_d^{bp}|x_d] = E[\hat{v}_d^{bp}|x_d] = 0$ and $E[u_d|x_d] = E[v_d|x_d] = 0$, by applying Proposition A.1 we get

$$
\begin{aligned}
E\big[\mu_d \hat{\mu}_d^{bp\,\prime}|x_d\big] &= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)' \\
&+ B_d E\big[v_d\hat{v}_d^{bp\,\prime}|x_d\big]B_d' + E\big[u_d\hat{u}_d^{bp\,\prime}|x_d\big] + B_d E\big[v_d\hat{u}_d^{bp\,\prime}|x_d\big] + E\big[u_d\hat{v}_d^{bp\,\prime}|x_d\big]B_d' \\
&= (Z_d\eta + X_d\lambda)(Z_d\eta + X_d\lambda)' \\
&+ B_d\Sigma_d B_d'(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d' + V_{ud}(V_{\lambda d}+V_{ed})^{-1}\Phi_d \\
&+ B_d\Sigma_d B_d'(V_{\lambda d}+V_{ed})^{-1}\Phi_d + V_{ud}(V_{ud}+V_{ed})^{-1}B_d\Psi_d B_d'. \qquad \square
\end{aligned}
$$

Appendix B: Tables

Table B.1. Men (top) and women (bottom) poverty proportions in 2008.

province	n_d	N_d	dir	eblup	rmse.dir	rmse.ebp
Soria	14	19607	0.3600	0.1236	0.1440	0.0329
Teruel	82	62529	0.0699	0.1071	0.0332	0.0256
Lérida	126	212149	0.2119	0.1772	0.0407	0.0275
Salamanca	158	145415	0.3026	0.2392	0.0432	0.0292
León	189	187253	0.1841	0.1803	0.0328	0.0263
Ceuta	212	36775	0.3955	0.3108	0.0394	0.0285
Toledo	288	325132	0.2502	0.2295	0.0319	0.0250
Pontevedra	440	474953	0.1558	0.1648	0.0194	0.0184
La Coruña	507	528266	0.1613	0.1607	0.0174	0.0160
Baleares	584	528518	0.1132	0.1182	0.0147	0.0136
Madrid	1235	3018699	0.1209	0.1160	0.0112	0.0111
Soria	9	11760	0.5303	0.1416	0.1846	0.0354
Teruel	81	63996	0.1291	0.1452	0.0502	0.0295
Lérida	123	201188	0.2130	0.1720	0.0382	0.0289
Salamanca	155	145453	0.3484	0.2719	0.0458	0.0315
León	227	256784	0.1825	0.1795	0.0293	0.0231
Ceuta	238	39496	0.3976	0.3305	0.0375	0.0260
Toledo	270	307611	0.2713	0.2468	0.0336	0.0253
Pontevedra	460	457734	0.1931	0.1955	0.0220	0.0183
La Coruña	605	632648	0.1894	0.1876	0.0175	0.0157
Baleares	613	520372	0.1445	0.1427	0.0191	0.0173
Madrid	1371	3161588	0.1474	0.1409	0.0116	0.0111

Table B.2. Men (top) and women (bottom) poverty gaps in 2008.

province	n_d	N_d	dir	eblup	rmse.dir	rmse.eblup
Soria	14	19607	0.0589	0.0206	0.0233	0.0055
Teruel	82	62529	0.0159	0.0244	0.0084	0.0070
Lérida	126	212149	0.0797	0.0535	0.0251	0.0139
Salamanca	158	145415	0.0767	0.0634	0.0136	0.0110
León	189	187253	0.0505	0.0490	0.0108	0.0091
Ceuta	212	36775	0.1720	0.1140	0.0208	0.0135
Toledo	288	325132	0.1255	0.0769	0.0307	0.0133
Pontevedra	440	474953	0.0486	0.0505	0.0073	0.0067
La Coruña	507	528266	0.0480	0.0477	0.0064	0.0060
Baleares	584	528518	0.0437	0.0441	0.0077	0.0070
Madrid	1235	3018699	0.0392	0.0393	0.0059	0.0055
Soria	9	11760	0.0868	0.0239	0.0300	0.0060
Teruel	81	63996	0.0232	0.0286	0.0106	0.0083
Lérida	123	201188	0.0482	0.0423	0.0108	0.0100
Salamanca	155	145453	0.0916	0.0827	0.0163	0.0121
León	227	256784	0.0440	0.0458	0.0085	0.0079
Ceuta	238	39496	0.1973	0.1330	0.0217	0.0130
Toledo	270	307611	0.1204	0.0809	0.0317	0.0141
Pontevedra	460	457734	0.0625	0.0616	0.0088	0.0077
La Coruña	605	632648	0.0531	0.0528	0.0062	0.0054
Baleares	613	520372	0.0481	0.0473	0.0079	0.0070
Madrid	1371	3161588	0.0396	0.0411	0.0064	0.0061

Appendix C: Maximum likelihood method

This appendix derives the Fisher-scoring algorithm for calculating the ML estimators of the MEBFH model parameters. Section C.1 presents some matrix derivatives and Section C.2 gives the fitting algorithm.

C.1. Matrix derivatives

For $k = 1, 2$, $j = 1, \dots, p_k$, let us define $e_{kj} = \underset{1\le i\le p_k}{\text{col}}(\delta_{ij}) = (0,\dots,0,1^{(j)},0,\dots,0)'_{p_k\times 1}$.
The first derivatives of $V_{\lambda d}$ are

$$V_{\lambda d1j} = \frac{\partial V_{\lambda d}}{\partial \lambda_{1j}} = \begin{pmatrix} 2e'_{1j}\Sigma_{d11}\lambda_1 & e'_{1j}\Sigma_{d12}\lambda_2 \\ \lambda'_2\Sigma_{d21}e_{1j} & 0 \end{pmatrix}, \quad V_{\lambda d2j} = \frac{\partial V_{\lambda d}}{\partial \lambda_{2j}} = \begin{pmatrix} 0 & \lambda'_1\Sigma_{d12}e_{2j} \\ e'_{2j}\Sigma_{d21}\lambda_1 & 2e'_{2j}\Sigma_{d22}\lambda_2 \end{pmatrix},$$

The second derivatives of $V_{\lambda d}$ are

$$V_{\lambda d11j_1j_2} = \frac{\partial^2 V_{\lambda d}}{\partial \lambda_{1j_1}\partial \lambda_{1j_2}} = \begin{pmatrix} 2e'_{1j_1}\Sigma_{d11}e_{1j_2} & 0 \\ 0 & 0 \end{pmatrix}, V_{\lambda d22j_1j_2} = \frac{\partial^2 V_{\lambda d}}{\partial \lambda_{2j_1}\partial \lambda_{2j_2}} = \begin{pmatrix} 0 & 0 \\ 0 & 2e'_{2j_1}\Sigma_{d22}e_{2j_2} \end{pmatrix},$$

$$V_{\lambda d12j_1j_2} = V_{\lambda d21j_1j_2} = \frac{\partial^2 V_{\lambda d}}{\partial \lambda_{1j_1}\partial \lambda_{2j_2}} = \begin{pmatrix} 0 & e'_{1j_1}\Sigma_{d12}e_{2j_2} \\ e'_{1j_1}\Sigma_{d12}e_{2j_2} & 0 \end{pmatrix},$$

The first derivatives of V_{ud} with respect to σ_{u1}^2, σ_{u2}^2 and ρ are

$$V_{ud1}=\frac{\partial V_{ud}}{\partial \sigma_{u1}^2}=\begin{pmatrix}1 & \frac{\rho\sigma_{u2}}{2\sigma_{u1}}\\ \frac{\rho\sigma_{u2}}{2\sigma_{u1}} & 0\end{pmatrix},\quad V_{ud2}=\frac{\partial V_{ud}}{\partial \sigma_{u2}^2}=\begin{pmatrix}0 & \frac{\rho\sigma_{u1}}{2\sigma_{u2}}\\ \frac{\rho\sigma_{u1}}{2\sigma_{u2}} & 1\end{pmatrix},$$
$$V_{ud3}=\frac{\partial V_{ud}}{\partial \rho}=\begin{pmatrix}0 & \sigma_{u1}\sigma_{u2}\\ \sigma_{u1}\sigma_{u2} & 0\end{pmatrix}.$$

The second derivatives of V_{ud} are

$$V_{ud11}=\frac{\partial^2 V_{ud}}{\partial \sigma_{u1}^2\partial \sigma_{u1}^2}=\begin{pmatrix}0 & \frac{-\rho\sigma_{u2}}{4\sigma_{u1}^3}\\ \frac{-\rho\sigma_{u2}}{4\sigma_{u1}^3} & 0\end{pmatrix},\ V_{ud22}=\frac{\partial^2 V_{ud}}{\partial \sigma_{u2}^2\partial \sigma_{u2}^2}=\begin{pmatrix}\frac{-\rho\sigma_{u1}}{4\sigma_{u2}^3} & 0\\ 0 & \frac{-\rho\sigma_{u1}}{4\sigma_{u2}^3}\end{pmatrix},$$
$$V_{ud33}=\frac{\partial^2 V_{ud}}{\partial \rho\partial \rho}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\ V_{ud12}=\frac{\partial^2 V_{ud}}{\partial \sigma_{u1}^2\partial \sigma_{u2}^2}=\begin{pmatrix}0 & \frac{\rho}{4\sigma_{u1}\sigma_{u2}}\\ \frac{\rho}{4\sigma_{u1}\sigma_{u2}} & 0\end{pmatrix},$$
$$V_{ud13}=\frac{\partial^2 V_{ud}}{\partial \sigma_{u1}^2\partial \rho}=\begin{pmatrix}0 & \frac{\sigma_{u2}}{2\sigma_{u1}}\\ \frac{\sigma_{u2}}{2\sigma_{u1}} & 0\end{pmatrix},\ V_{ud23}=\frac{\partial^2 V_{ud}}{\partial \sigma_{u2}^2\partial \rho}=\begin{pmatrix}0 & \frac{\sigma_{u1}}{2\sigma_{u2}}\\ \frac{\sigma_{u1}}{2\sigma_{u2}} & 0\end{pmatrix}.$$

C.2. ML Fisher-scoring algorithm

It holds that $y|X \sim N(Z\eta+X\lambda, V)$, with covariance matrix $V=\underset{1\le d\le D}{\text{diag}}(V_d)$, $V_d=V_{\lambda d}+V_{ud}+V_{ed}$. The vector of model parameters is $\psi=(\eta',\lambda',\theta')'$, where $\theta=(\theta_1,\theta_2,\theta_3)'=(\sigma_{u1}^2,\sigma_{u2}^2,\rho)'$. The log-likelihood is $l=\sum_{d=1}^D l_d$, where

$$l_d=-\log 2\pi-\frac{1}{2}\log|V_d|-\frac{1}{2}(y_d-Z_d\eta-X_d\lambda)'V_d^{-1}(y_d-Z_d\eta-X_d\lambda)$$

For $i=1,\dots,q_k$, $j=1,\dots,p_k$, $k=1,2$, $\ell=1,2,3$, the first partial derivatives of ℓ_d are

$$\begin{aligned}
\frac{\partial l_d}{\partial \eta_{1i}}&=(z_{d1i},0)V_d^{-1}(y_d-Z_d\eta-X_d\lambda),\quad \frac{\partial l_d}{\partial \eta_{2i}}=(0,z_{d2i})V_d^{-1}(y_d-Z_d\eta-X_d\lambda),\\
\frac{\partial l_d}{\partial \lambda_{1j}}&=-\frac{1}{2}\text{tr}\big(V_d^{-1}V_{\lambda d1j}\big)+(x_{d1j},0)V_d^{-1}(y_d-Z_d\eta-X_d\lambda)\\
&\quad+\frac{1}{2}(y_d-Z_d\eta-X_d\lambda)'V_d^{-1}V_{\lambda d1j}V_d^{-1}(y_d-Z_d\eta-X_d\lambda),\\
\frac{\partial l_d}{\partial \lambda_{2j}}&=-\frac{1}{2}\text{tr}\big(V_d^{-1}V_{\lambda d2j}\big)+(0,x_{d2j})V_d^{-1}(y_d-Z_d\eta-X_d\lambda)\\
&\quad+\frac{1}{2}(y_d-Z_d\eta-X_d\lambda)'V_d^{-1}V_{\lambda d2j}V_d^{-1}(y_d-Z_d\eta-X_d\lambda),\\
\frac{\partial l_d}{\partial \theta_\ell}&=-\frac{1}{2}\text{tr}\big(V_d^{-1}V_{ud\ell}\big)+\frac{1}{2}(y_d-Z_d\eta-X_d\lambda)'V_d^{-1}V_{ud\ell}V_d^{-1}(y_d-Z_d\eta-X_d\lambda).
\end{aligned}$$

The second partial derivatives of ℓ_d are

$$
\begin{aligned}
\frac{\partial^2 l_d}{\partial \eta_{1i_1}\partial \eta_{1i_2}} &= -(z_{d1i_1},0)V_d^{-1}(z_{d1i_2},0)', \quad \frac{\partial^2 l_d}{\partial \eta_{2i_1}\partial \eta_{2i_2}} = -(0,z_{d2i_1})V_d^{-1}(0,z_{d2i_2})', \\
\frac{\partial^2 l_d}{\partial \eta_{1i_1}\partial \eta_{2i_2}} &= -(z_{d1i_1},0)V_d^{-1}(0,z_{d2i_2})', \quad \frac{\partial^2 l_d}{\partial \eta_{1i_1}\partial \lambda_{1i_2}} = -(z_{d1i_1},0)V_d^{-1}(x_{d1i_2},0)', \\
\frac{\partial^2 l_d}{\partial \eta_{2i_1}\partial \lambda_{2i_2}} &= -(0,z_{d2i_1})V_d^{-1}(0,x_{d2i_2})', \quad \frac{\partial^2 l_d}{\partial \eta_{1i_1}\partial \lambda_{2i_2}} = -(z_{d1i_1},0)V_d^{-1}(0,x_{d2i_2})', \\
\frac{\partial^2 l_d}{\partial \lambda_{1j_1}\partial \lambda_{1j_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d1j_2}V_d^{-1}V_{\lambda d1j_1}\big) - \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d11j_1j_2}\big) - (x_{d1j_1},0)V_d^{-1}(x_{d1j_2},0)' \\
&- (x_{d1j_1},0)V_d^{-1}V_{\lambda d1j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) - (x_{d1j_2},0)V_d^{-1}V_{\lambda d1j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d1j_2}V_d^{-1}V_{\lambda d1j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&+ \frac{1}{2}(y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d11j_1j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \lambda_{2j_1}\partial \lambda_{2j_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d2j_1}\big) - \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d22j_1j_2}\big) - (0,x_{d2j_1})V_d^{-1}(0,x_{d2j_2})' \\
&- (0,x_{d2j_1})V_d^{-1}V_{\lambda d2j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) - (0,x_{d2j_2})V_d^{-1}V_{\lambda d2j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d2j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&+ \frac{1}{2}(y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d22j_1j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \lambda_{1j_1}\partial \lambda_{2j_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d1j_1}\big) - \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d12j_1j_2}\big) - (x_{d1j_1},0)V_d^{-1}(0,x_{d2j_2})' \\
&- (x_{d1j_1},0)V_d^{-1}V_{\lambda d2j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) - (0,x_{d2j_2})V_d^{-1}V_{\lambda d1j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d1j_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&+ \frac{1}{2}(y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{\lambda d12j_1j_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \eta_{1i}\partial \theta_\ell} &= -(z_{d1i},0)V_d^{-1}V_{ud\ell}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \eta_{2i}\partial \theta_\ell} &= -(0,z_{d2i})V_d^{-1}V_{ud\ell}V_d^{-1}(y_d - Z_d\eta - X_d\lambda),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_d}{\partial \lambda_{1j}\partial \theta_\ell} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d1j}\big) - (x_{d1j},0)V_d^{-1}V_{ud\ell}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d1j}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \lambda_{2j}\partial \theta_\ell} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d2j}\big) - (0,x_{d2j})V_d^{-1}V_{ud\ell}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d2j}V_d^{-1}(y_d - Z_d\eta - X_d\lambda), \\
\frac{\partial^2 l_d}{\partial \theta_{\ell_1}\partial \theta_{\ell_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell_2}V_d^{-1}V_{ud\ell_1}\big) - \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell_1\ell_2}\big) \\
&- (y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{ud\ell_2}V_d^{-1}V_{ud\ell_1}V_d^{-1}(y_d - Z_d\eta - X_d\lambda) \\
&+ \frac{1}{2}(y_d - Z_d\eta - X_d\lambda)'V_d^{-1}V_{ud\ell_1\ell_2}V_d^{-1}(y_d - Z_d\eta - X_d\lambda),
\end{aligned}
$$

By changing the sign and taking expectations, conditionally on X_d, we have

$$\begin{aligned}
F_{d\eta_{1i_1}\eta_{1i_2}} &= (z_{d1i_1},0)V_d^{-1}(z_{d1i_2},0)', \quad F_{d\eta_{2i_1}\eta_{2i_2}} = (0,z_{d2i_1})V_d^{-1}(0,z_{d2i_2})', \\
F_{d\eta_{1i_1}\eta_{2i_2}} &= (z_{d1i_1},0)V_d^{-1}(0,z_{d2i_2})', \quad F_{d\eta_{1i}\lambda_{1j}} = (z_{d1i},0)V_d^{-1}(x_{d1j},0)', \\
F_{d\eta_{2i}\lambda_{2j}} &= (0,z_{d2i})V_d^{-1}(0,x_{d2j})', \quad F_{d\eta_{1i}\lambda_{2j}} = (z_{d1i},0)V_d^{-1}(0,x_{d2j})', \\
F_{d\lambda_{1j_1}\lambda_{1j_2}} &= -\frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d1j_2}V_d^{-1}V_{\lambda d1j_1}\big) + \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d11j_1j_2}\big) \\
&\quad + (x_{d1j_1},0)V_d^{-1}(x_{d1j_2},0)' + \mathrm{tr}\big(V_d^{-1}V_{\lambda d1j_2}V_d^{-1}V_{\lambda d1j_1}\big) - \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d11j_1j_2}\big) \\
&= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d1j_2}V_d^{-1}V_{\lambda d1j_1}\big) + (x_{d1j_1},0)V_d^{-1}(x_{d1j_2},0)', \\
F_{d\lambda_{2j_1}\lambda_{2j_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d2j_1}\big) + (0,x_{d2j_1})V_d^{-1}(0,x_{d2j_2})', \\
F_{d\lambda_{1j_1}\lambda_{2j_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{\lambda d2j_2}V_d^{-1}V_{\lambda d1j_1}\big) + (x_{d1j_1},0)V_d^{-1}(0,x_{d2j_2})',
\end{aligned}$$

$$\begin{aligned}
F_{d\lambda_{1j}\theta_\ell} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d1j}\big), \quad F_{d\lambda_{2j}\theta_\ell} = \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell}V_d^{-1}V_{\lambda d2j}\big), \\
F_{d\theta_{\ell_1}\theta_{\ell_2}} &= \frac{1}{2}\mathrm{tr}\big(V_d^{-1}V_{ud\ell_2}V_d^{-1}V_{ud\ell_1}\big), \quad F_{d\eta_{1i}\theta_\ell} = F_{d\eta_{2i}\theta_\ell} = 0.
\end{aligned}$$

The components of the score vector are

$$S_{\eta_1 j} = \sum_{d=1}^{D}\frac{\partial l_d}{\partial \eta_{1j}},\ 1\le j\le p_1; \quad S_{\eta_2 j} = \sum_{d=1}^{D}\frac{\partial l_d}{\partial \eta_{2j}},\ 1\le j\le p_2;$$

$$S_{\lambda_1 j} = \sum_{d=1}^{D}\frac{\partial l_d}{\partial \lambda_{1j}},\ 1\le j\le p_1; \quad S_{\lambda_2 j} = \sum_{d=1}^{D}\frac{\partial l_d}{\partial \lambda_{2j}},\ 1\le j\le p_2; \quad S_{\theta_\ell} = \sum_{d=1}^{D}\frac{\partial l_d}{\partial \theta_\ell},\ \ell=1,2,3.$$

The score vector is

$$S(\psi) = \big(S_{\eta_1 1},\dots,S_{\eta_1 q_1},S_{\eta_2 1},\dots,S_{\eta_2 q_2},S_{\lambda_1 1},\dots,S_{\lambda_1 p_1},S_{\lambda_2 1},\dots,S_{\lambda_2 p_2},S_{\theta_1},S_{\theta_2},S_{\theta_3}\big)'.$$

The components of the Fisher information matrix are

$$F_{\eta_{1j_1}\eta_{1j_2}} = \sum_{d=1}^{D}F_{d\eta_{1j_1}\eta_{1j_2}},\ F_{\eta_{2j_1}\eta_{2j_2}} = \sum_{d=1}^{D}F_{d\eta_{2j_1}\eta_{2j_2}},\ F_{\eta_{1j_1}\lambda_{2j_2}} = \sum_{d=1}^{D}F_{d\eta_{1j_1}\eta_{2j_2}},$$

$$F_{\lambda_{1j_1}\lambda_{1j_2}} = \sum_{d=1}^{D}F_{d\lambda_{1j_1}\lambda_{1j_2}},\ F_{\lambda_{2j_1}\lambda_{2j_2}} = \sum_{d=1}^{D}F_{d\lambda_{2j_1}\lambda_{2j_2}},\ F_{\lambda_{1j_1}\lambda_{2j_2}} = \sum_{d=1}^{D}F_{d\lambda_{1j_1}\lambda_{2j_2}},$$

$$F_{\eta_{1j_1}\lambda_{1j_2}} = \sum_{d=1}^{D}F_{d\eta_{1j_1}\lambda_{1j_2}},\ F_{\eta_{2j_1}\lambda_{2j_2}} = \sum_{d=1}^{D}F_{d\eta_{2j_1}\lambda_{2j_2}},\ F_{\eta_{1j_1}\lambda_{2j_2}} = \sum_{d=1}^{D}F_{d\eta_{1j_1}\lambda_{2j_2}},$$

$$F_{\eta_{1j}\theta_\ell} = \sum_{d=1}^{D}F_{d\eta_{1j}\theta_\ell},\ F_{\eta_{2j}\theta_\ell} = \sum_{d=1}^{D}F_{d\eta_{2j}\theta_\ell},$$

$$F_{\lambda_{1j}\theta_\ell}=\sum_{d=1}^{D}F_{d\lambda_{1j}\theta_\ell},\ F_{\lambda_{2j}\theta_\ell}=\sum_{d=1}^{D}F_{d\lambda_{2j}\theta_\ell},\ F_{\theta_{\ell_1}\theta_{\ell_2}}=\sum_{d=1}^{D}F_{d\theta_{\ell_1}\theta_{\ell_2}}.$$

The blocks of the Fisher information matrix are

$$F_{\eta_1\eta_1}=\Big(F_{\eta_{1j_1}\eta_{1j_2}}\Big)_{i_1,i_2=1,\dots,q_1},F_{\eta_2\eta_2}=\Big(F_{\eta_{2j_1}\eta_{2j_2}}\Big)_{i_1,i_2=1,\dots,q_2},F_{\eta_1\eta_2}=\Big(F_{\eta_{1j_1}\eta_{2j_2}}\Big)_{i_1=1,\dots,q_1;i_2=1,\dots,q_2},$$

$$F_{\lambda_1\lambda_1}=\Big(F_{\lambda_{1j_1}\lambda_{1j_2}}\Big)_{j_1,j_2=1,\dots,p_1},F_{\lambda_2\lambda_2}=\Big(F_{\lambda_{2j_1}\lambda_{2j_2}}\Big)_{j_1,j_2=1,\dots,p_2},F_{\lambda_1\lambda_2}=\Big(F_{\lambda_{1j_1}\lambda_{2j_2}}\Big)_{j_1=1,\dots,p_1;j_2=1,\dots,p_2},$$

$$F_{\eta_1\lambda_1}=\Big(F_{\eta_{1i}\lambda_{1j}}\Big)_{i=1,\dots,q_1;j=1,\dots,p_1},F_{\eta_2\lambda_2}=\Big(F_{\eta_{2i}\lambda_{2j}}\Big)_{i=1,\dots,q_2;j=1,\dots,p_2},$$

$$F_{\eta_1\lambda_2}=\Big(F_{\eta_{1i}\lambda_{2j}}\Big)_{i=1,\dots,q_1;j=1,\dots,p_2},\ F_{\eta_1\theta}=\Big(F_{\eta_{1i}\theta_\ell}\Big)_{i=1,\dots,q_1;\ell=1,2,3},\ F_{\eta_2\theta}=\Big(F_{\eta_{2i}\theta_\ell}\Big)_{i=1,\dots,q_2;\ell=1,2,3}.$$

$$F_{\lambda_1\theta}=\Big(F_{\lambda_{1j}\theta_\ell}\Big)_{j=1,\dots,p_1;\ell=1,2,3},\ F_{\lambda_2\theta}=\Big(F_{\lambda_{2j}\theta_\ell}\Big)_{j=1,\dots,p_2;\ell=1,2,3},\ F_{\theta\theta}=\Big(F_{\theta_{\ell_1}\theta_{\ell_2}}\Big)_{\ell_1=1,2,3;\ell_2=1,2,3}.$$

The Fisher information matrix is

$$F(\psi)=\begin{pmatrix} F_{\eta_1\eta_1} & F_{\eta_1\eta_2} & F_{\eta_1\lambda_1} & F_{\eta_1\lambda_2} & F_{\eta_1\theta}\\ F_{\eta_2\eta_1} & F_{\eta_2\eta_2} & F_{\eta_2\lambda_1} & F_{\eta_2\lambda_2} & F_{\eta_2\theta}\\ F_{\lambda_1\eta_1} & F_{\lambda_1\eta_2} & F_{\lambda_1\lambda_1} & F_{\lambda_1\lambda_2} & F_{\lambda_1\theta}\\ F_{\lambda_2\eta_1} & F_{\lambda_2\eta_2} & F_{\lambda_2\lambda_1} & F_{\lambda_2\lambda_2} & F_{\lambda_2\theta}\\ F_{\theta\eta_1} & F_{\theta\eta_2} & F_{\theta\lambda_1} & F_{\theta\lambda_2} & F_{\theta\theta} \end{pmatrix}_{(p+q+3)\times(p+q+3)}.$$

The ML Fisher-scoring algorithm is

1. Set the initial values $\psi^{(0)}=(\eta_1^{(0)},\eta_2^{(0)},\lambda_1^{(0)},\lambda_2^{(0)},\theta^{(0)})$ and $\varepsilon_j>0$, $j=1,\dots,p+3$.
2. Repeat the following steps until the tolerance or the boundary conditions are fulfilled.
   (a) Updating equation: Do $\psi^{(i+1)}=\psi^{(i)}+F^{-1}(\psi^{(i)})S(\psi^{(i)})$.
   (b) Boundary condition: If $\theta_1^{(i+1)}>0$, $\theta_2^{(i+1)}>0$ and $\big|\theta_3^{(i+1)}\big|<1$, continue. Otherwise, do $\hat{\psi}=\psi^{(i)}$ and stop.
   (c) Tolerance condition: If $\big|\psi_j^{(i+1)}-\psi_j^{(i)}\big|<\varepsilon$, $j=1,\dots,p+3$, do $\hat{\psi}=\psi^{(i+1)}$ and stop. Otherwise, continue.
3. Output: $\hat{\psi}$, $F^{-1}(\hat{\psi})$.

Remark. As starting values, we take $\eta_1^{(0)}=\hat{\eta}_1^{(0)}$, $\eta_2^{(0)}=\hat{\eta}_2^{(0)}$, $\lambda_1^{(0)}=\hat{\lambda}_1^{(0)}$, $\lambda_2^{(0)}=\hat{\lambda}_2^{(0)}$, $\hat{\theta}_{3,0}=0$, $\hat{\theta}_{k,0}=\hat{\sigma}_{uk,0}^2$, $k=1,2$, where $\hat{\eta}_1^{(0)}$, $\hat{\eta}_2^{(0)}$, $\hat{\lambda}_1^{(0)}$, $\hat{\lambda}_2^{(0)}$ and $\hat{\sigma}_{uk,0}^2$ are the REML or the ML estimators of the corresponding bivariate Fay-Herriot model without measurement errors.