Linear Unmixing of Multivariate Observations: A Structural Model – Remarks

Werner Stahel
Seminar für Statistik,
Swiss Federal Institute of Technology (ETH),
CH-8092 Zurich, Switzerland

October 19, 2005

This is a collection of notes and remarks referring to the paper in JASA, December 2005. They mainly result from my efforts of going through Marcel Wolber's work in some more depth. They are subject to misunderstandings from my part.

Contents

1 Notation

We collect the notation for ease of reference:

		Model
$\mathbf{X}_i$	m	data vector
$\mathbf{C}$	$m \times p$	matrix of source profiles (= columns)
$\mathbf{S}_i$	p	scores
$\mathbf{E}_i$	m	error term
$\boldsymbol{\mu}$	p	expectation of log scores
$\boldsymbol{\Psi}$	$p \times p$	covariance matrix of log scores
$\boldsymbol{\Sigma}$	$m \times m$	covariance matrix of log errors, diagonal
σ_j^2	1	variance log error for variable j
		Extended Model
$\mathbf{u}_i$	q	covariates for observation i
$\mathbf{B}$	$p \times p$	coefficients for the covariates
Additional Quantities for Theoretical Considerations		
$\boldsymbol{\theta}$	q	collection of all parameters, $= (\mathbf{C}, \boldsymbol{\mu}, \boldsymbol{\Psi}, \boldsymbol{\Sigma})$. $q = mp + p + p(p+1)/2 + m$
$\mathbf{I}$	$q \times q$	Fisher Information matrix
$\mathbf{Z}_i$	m	$= \log(\mathbf{X}_i)$ (elementwise) log data vector
$\mathbf{V}_i$	p	$= \log(\mathbf{S}_i)$ (elementwise) log scores vector
$\mathbf{U}$	p	$= \mathbf{R}^{-t}(\mathbf{V} - \boldsymbol{\mu})$ standardized log scores
$\mathbf{R}$	$p \times p$	(Cholesky) root of $\boldsymbol{\Psi}$

2 Proof of Theorem 3

Structure of the Proof.

The structure of the proof is given by the following relations between Lemmas (L) and conditions.

Theorem 3 $\Longleftarrow$ (A1), (A2), (B), (C), (D)

(A1) $\Longleftarrow$ L A.2 $\Longleftarrow$ L A.1

(A2, B) $\Longleftarrow$ (A.3) in L A.2 and (L A.5 or L A.6) and L A.3

(C) $\Longleftarrow$ Th 2

(D) $\Longleftarrow$ L A.7

L A.3 $\Longleftarrow$ L A.1

L A.5 $\Longleftarrow$ L A.1

L A.6 $\Longleftarrow$ L A.1

L A.7 $\Longleftarrow$ L A.2 $\Longleftarrow$ L A.1

The lemmas state the following:

Lemma A.1: Bounding $f_{\mathbf{Z}|\mathbf{V}=\mathbf{v}}(\mathbf{z};\boldsymbol{\theta})$ from below and above by an exponential with the same terms in $\max\left(v^{(k)}\right)$ and $\|z\|$.

Lemma A.2: Continuous differentiability of $\boldsymbol{\theta} \mapsto \sqrt{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}$.

Lemma A.3: Bounding $f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})$.

Lemma A.4: Bounding the cumulative standard normal distribution function.

Lemma A.5 and A.6: Bounding $\mathcal{E}(||\mathbf{V}||^2|\mathbf{Z}=\mathbf{z};\boldsymbol{\theta})$.

Lemma A.7: Fisher Information is non-singular.

Lemmas and Proofs.

For ease of readability, we reproduce the whole text of Appendix A.2 of the paper, starting from Lemma A.1. The few modifications will be printed in *slanted* font.

Lemma A.1 *There exist constants* $\kappa_l^{(1)}, \kappa_l^{(2)}, \kappa_l^{(3)}, \kappa_u^{(1)}, \kappa_u^{(2)}, \kappa_u^{(3)}$ *such that for all* $\boldsymbol{\theta} \in \mathcal{U}_{\boldsymbol{\theta}^0}$ *the following inequalities hold*

$$\begin{aligned}
&\frac{(2\pi)^{-m/2}}{\sigma_1 \cdots \sigma_m} \cdot \exp\Big(-\frac{1}{2} \sum_{j=1}^{m} \frac{(z^{(j)} - v^*)^2}{\sigma_j^2} + \kappa_l^{(1)} + \kappa_l^{(2)} |v^*| + \kappa_l^{(3)} ||\mathbf{z}|| \Big) \\
\leq\ & f_{\mathbf{Z}|\mathbf{V}=\mathbf{v}}(\mathbf{z}; \boldsymbol{\theta}) \quad \leq \\
&\frac{(2\pi)^{-m/2}}{\sigma_1 \cdots \sigma_m} \cdot \exp\Big(-\frac{1}{2} \sum_{j=1}^{m} \frac{(z^{(j)} - v^*)^2}{\sigma_j^2} + \kappa_u^{(1)} + \kappa_u^{(2)} |v^*| + \kappa_u^{(3)} ||\mathbf{z}|| \Big)
\end{aligned}$$

where $v^* := \max_{1 \leq k \leq p} \{v^{(k)}\}$ *and* $\mathcal{U}_{\boldsymbol{\theta}^0}$ *is as defined at the beginning of the proof of Theorem 3.*

Proof: It is easy to see that the following inequalities hold: For $k_\ell := \min_{1 \leq k \leq p} \big(\log(c_{jk})\big) \geq \log(\triangle_l^c)$ and $k_u := \log(p) + \max_{1 \leq k \leq p} \big(\log(c_{jk})\big) < \log(p)$ we have

$$k_\ell + v^* \leq \max_{1 \leq k \leq p} (\log(c_{jk}) + v^{(k)}) \leq Q^{(j)} \leq k_u + v^*, \tag{A.2}$$

where $Q^{(j)} := \log(\sum_{k=1}^{p} c_{jk} \exp(v^{(k)}))$, *since* $c_{jk} \leq 1$. *Writing* $(z^{(j)} - Q^{(j)})^2 = z^{(j)2} + Q^{(j)}(Q^{(j)} - 2z^{(j)})$, *using the bounds for* $Q^{(j)}$ *and summing over* j *leads to the desired result.*

More Details:

$$\begin{aligned}
(z^{(j)} - Q^{(j)})^2 &= z^{(j)2} + Q^{(j)}(Q^{(j)} - 2z^{(j)}) \\
&\leq z^{(j)2} + (k_u + v^*)(k_u + v^* - 2z^{(j)}) \\
&= z^{(j)2} + v^*(v^* - 2z^{(j)}) + k_u^2 + 2k_u v^* - 2k_u z^{(j)} \\
&= (z^{(j)} - v^*)^2 + k_u^2 + 2k_u v^* - 2k_u z^{(j)}
\end{aligned}$$

Summation over j *with "weights"* $1/\sigma_j^2$ *shows the right hand inequality of the lemma for* $\kappa_u^{(1)} = mk_u^2$, $\kappa_u^{(2)} = \kappa_u^{(3)} = 2mk_u$.

□

Lemma A.2 *The map* $\boldsymbol{\theta} \mapsto \sqrt{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}$ *is continuously differentiable on* $\mathcal{U}_{\boldsymbol{\theta}^0}$ *for every* $\mathbf{z} \in \mathbb{R}^p$ *where* $\mathcal{U}_{\boldsymbol{\theta}^0}$ *is as defined at the beginning of the proof of Theorem 3. Moreover,* $\frac{\partial}{\partial\boldsymbol{\theta}} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) = \int \frac{\partial}{\partial\boldsymbol{\theta}} f_{\mathbf{Z},\mathbf{U}}(\mathbf{z},\mathbf{u};\boldsymbol{\theta}) d\mathbf{u}$.

Proof: We prove continuous differentiability of $\boldsymbol{\theta} \mapsto f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})$. Since $f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) > 0 \; \forall \boldsymbol{\theta} \in \mathcal{U}_{\boldsymbol{\theta}^0}, \mathbf{z} \in \mathbb{R}^p$ the result then follows by the chain rule.

$\boldsymbol{\theta} \mapsto f_{\mathbf{Z},\mathbf{U}}(\mathbf{z},\mathbf{u};\boldsymbol{\theta})$ is continuously differentiable on $\mathcal{U}_{\boldsymbol{\theta}^0}$ for every $\mathbf{z} \in \mathbb{R}^p$. Thus, it is only necessary to justify the interchange of derivative and integral. Writing explicitly out the derivatives of $f_{\mathbf{Z},\mathbf{U}}(\mathbf{z},\mathbf{u};\boldsymbol{\theta})$ and bounding

$$\begin{aligned} \frac{\exp(\boldsymbol{\mu} + \mathbf{R}^t\mathbf{u})^{(k)}}{h_j(\mathbf{u})} &\leq \frac{1}{\triangle_l^c} \quad \text{and} \\ |\log(h_j(\mathbf{u}))| &\leq \log(p) + |\log(\triangle_l^c)| + \triangle_u^\mu + \triangle_u^R||\mathbf{u}||\,, \end{aligned}$$

where $h_j(\mathbf{u}) := \sum_{k=1}^p c_{jk} \exp(\boldsymbol{\mu} + \mathbf{R}^t\mathbf{u})^{(k)}$ and we used (A.2) to establish the second inequality, it is easy to see that

$$\begin{aligned} ||\frac{\partial}{\partial\boldsymbol{\theta}} f_{\mathbf{Z},\mathbf{U}}(\mathbf{z},\mathbf{u};\boldsymbol{\theta})|| &\leq (k_1 + k_2||\mathbf{z}||^2 + k_3||\mathbf{u}||^2) f_{\mathbf{Z},\mathbf{U}}(\mathbf{z},\mathbf{u};\boldsymbol{\theta}) \qquad \text{(A.3)} \\ &\leq \frac{(2\pi)^{m/2}}{\triangle^m}(k_1 + k_2||\mathbf{z}||^2 + k_3||\mathbf{u}||^2) f_{\mathbf{U}}(\mathbf{u}) \end{aligned}$$

holds uniformly over $\mathcal{U}_{\boldsymbol{\theta}^0}$ for suitably chosen constants k_1, k_2, k_3. Since this upper bound is integrable with respect to $\mathbf{u}$, the claim follows from standard results, see, for example Bauer (1992, Kor. 16.3, p. 103). □

Lemma A.3 *There exist constants* κ_1 *and* κ_2 *with* $\kappa_2 > 0$ *such that for all parameter sets in* $\mathcal{U}_{\boldsymbol{\theta}^0}$ *the following inequality holds:*

$$f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) \leq \kappa_1 \exp(-\kappa_2||\mathbf{z}||^2) \qquad \text{(A.4)}$$

where $\mathcal{U}_{\boldsymbol{\theta}^0}$ *is as defined at the beginning of the proof of Theorem 3.*

Proof: We prove (A.4) below for a specific $\boldsymbol{\theta} \in \mathcal{U}_{\boldsymbol{\theta}^0}$. From this proof, the constants κ_1, κ_2 can be seen to depend continuously on $\boldsymbol{\theta}$. Thus, to prove the global result one can set $\kappa_1 = \sup_{\boldsymbol{\theta}\in\overline{\mathcal{U}}_{\boldsymbol{\theta}0}} \kappa_1(\boldsymbol{\theta})$ and $\kappa_2 = \inf_{\boldsymbol{\theta}\in\overline{\mathcal{U}}_{\boldsymbol{\theta}0}} \kappa_2(\boldsymbol{\theta})$. Since $\kappa_1(\boldsymbol{\theta}) < \infty$ and $\kappa_2(\boldsymbol{\theta}) > 0$ for all $\boldsymbol{\theta}$ in the (compact) closure of $\mathcal{U}_{\boldsymbol{\theta}^0}$, the same holds for κ_1 and κ_2.

Define, *for* $q \in \{-1, 1\}$,

$$g_q(\mathbf{z}, \mathbf{v}, \nu) \;:=\; \exp\Big(-\frac{1}{2}\sum_{j=1}^{m} \frac{(z^{(j)} - \nu)^2}{\sigma_j^2} + \kappa_u^{(1)} + q \cdot \kappa_u^{(2)} \nu + \kappa_u^{(3)} ||\mathbf{z}|| - \frac{1}{2}(\mathbf{v} - \boldsymbol{\mu})^t \boldsymbol{\Psi}^{-1} (\mathbf{v} - \boldsymbol{\mu}) \Big) ,$$

where $\kappa_u^{(1)}, \kappa_u^{(2)}, \kappa_u^{(3)}$ are the constants from Lemma A.1. Because $|v^*|$ is $v^{(k)}$ or $-v^{(k)}$ for some k, and g_q is positive, it follows that *the bound obtained from Lemma A.1 can in turn be bounded by a sum over all terms* $g_q(..., v^{(k)})$*, which leads to*

$$f_{\mathbf{Z},\mathbf{V}}(\mathbf{z}, \mathbf{v}; \boldsymbol{\theta}) \;\le\; \frac{(2\pi)^{-(m+p)/2}}{\sigma_1 \cdots \sigma_m} (\det \boldsymbol{\Psi})^{-1/2} \times \Big(\sum_{k=1}^{p} g_1(\mathbf{z}, \mathbf{v}, v^{(k)}) + \sum_{k=1}^{p} g_{-1}(\mathbf{z}, \mathbf{v}, v^{(k)}) \Big) \tag{A.5}$$

We will show specifically for the first term in the first sum of (A.5) that it can be bounded as required for Lemma A.3. All other summands can be treated analogously. Therefore, the lemma is proved if we can show that there exist constants κ_1^* and κ_2^* with $\kappa_2^* > 0$ such that

$$I \;:=\; \int_{\mathbb{R}^p} \exp\Big(-\frac{1}{2}\sum_{j=1}^{m} \frac{(z^{(j)} - v^{(1)})^2}{\sigma_j^2} - \frac{1}{2}\mathbf{v}^t \boldsymbol{\Psi}^{-1} \mathbf{v} \Big) d\mathbf{v} \;\le\; \kappa_1^* \exp(-\kappa_2^* ||\mathbf{z}||^2) ,$$

where we have dropped terms of lower order as a convenience. (The calculations below would be the same, though more cumbersome, if these terms were not neglected.) *Quadratic completion will allow for explicit integration of* I*.*

Set $\mathbf{w} := \mathbf{R}^{-t}\mathbf{v}$ where $\mathbf{R}$ is the Choleski decomposition of $\boldsymbol{\Psi}$. Then I transforms to

$$\begin{aligned}
I &= (\det \mathbf{R}) \int_{\mathbb{R}^p} \exp\Big(-\frac{1}{2}\sum_{j=1}^{m} \frac{(z^{(j)} - r_{11} w^{(1)})^2}{\sigma_j^2} - \frac{1}{2}\sum_{k=1}^{p} (w^{(k)})^2 \Big) d\mathbf{w} \\
&= (\det \mathbf{R})(2\pi)^{(p-1)/2} \int_{\mathbb{R}} \exp\Big(-\frac{1}{2}\Big[\sum_{j=1}^{m} \frac{(z^{(j)} - r_{11} w)^2}{\sigma_j^2} + w^2 \Big] \Big) dw \\
&= (\det \mathbf{R})(2\pi)^{(p-1)/2} \int_{\mathbb{R}} \exp\Big(-\frac{1}{2}\Big[\Big(aw - \frac{1}{a} r_{11} \sum_{j=1}^{m} \frac{z^{(j)}}{\sigma_j^2} \Big)^2 \\
&\qquad + \sum_{j=1}^{m} \frac{(z^{(j)})^2}{\sigma_j^2} - \frac{1}{a^2} r_{11}^2 \Big(\sum_{j=1}^{m} \frac{(z^{(j)})}{\sigma_j^2} \Big)^2 \Big] \Big) dw \\
&= (\det \mathbf{R})(2\pi)^{p/2} \frac{1}{a} \exp\Big(-\frac{1}{2}\Big[\sum_{j=1}^{m} \frac{(z^{(j)})^2}{\sigma_j^2} - \frac{1}{a^2} r_{11}^2 \Big(\sum_{j=1}^{m} \frac{z^{(j)}}{\sigma_j^2} \Big)^2 \Big] \Big)
\end{aligned}$$

where we have set $a := \sqrt{1 + r_{11}^2 \sum_{j=1}^{m} 1/\sigma_j^2}$.

Because $Q_1(\mathbf{z}, w) := \sum_{j=1}^m (z^{(j)} - r_{11}w)^2/\sigma_j^2 + w^2$ is a positive definite quadratic form in $\mathbf{z}$ and w, the same is true for the quadratic form

$$Q_2(\mathbf{z}) := \sum_{j=1}^{m} \frac{(z^{(j)})^2}{\sigma_j^2} - \frac{1}{a^2} r_{11}^2 \Big(\sum_{j=1}^{m} \frac{z^{(j)}}{\sigma_j^2} \Big)^2 ,$$

as can be seen from the third equality in the above calculations to determine I. Setting κ_2^* to be half of the smallest eigenvalue of Q_2 leads to the desired bound,

$$I \leq (\det \mathbf{R})(2\pi)^{p/2} \frac{1}{a} \exp(-\kappa_2^* ||\mathbf{z}||^2).$$

$\square$

Lemma A.4 *Let* $\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ *and* $\Phi(x) = \int_{-\infty}^{x} \phi(t)dt$ *be the standard normal density and distribution function, respectively. Then for any* $\delta > 0$ *there exists a constant* $K = K(\delta)$ *such that* $K\phi((1+\delta)x) \leq \Phi(x) \leq \sqrt{\frac{\pi}{2}}\phi(x)$ *for all* $x < 0$.

Proof: For $-\sqrt{2/\pi} \leq x < 0$, the proof is straightforward. For $x < -\sqrt{2/\pi}$, it follows from

$$\left(\frac{1}{x^3} - \frac{1}{x} \right) \phi(x) \leq \Phi(x) \leq -\frac{1}{x} \phi(x) \quad \forall x < 0 \tag{A.6}$$

established in Feller (1968, p. 175).

More Details: *Let* $x_0 = -\sqrt{2/\pi} \leq x < 0$. *We have*

$$\phi((1+\delta)x) < \phi(0) = K_0 \Phi(-\sqrt{2/\pi} \leq K_0 \Phi(x)$$

with $K_0 = \phi(0)/\Phi(-\sqrt{2/\pi})$ *independently of* δ. *Furthermore,* $g(x) := \sqrt{\pi/2}\,\phi(x) - \Phi(x)$ *has a negative derivative* $\phi(x)(-\sqrt{\pi/2}x - 1)$ *and is* ≥ 0 *for* $x = x_0$ *and* $x = 0$, *which proves the right hand inequality. For* $x < x_0$, *... !!! Since* $\exp(-\delta x^2/2)$ *goes to 0 faster than any power of* x,

Lemma A.5 *Assume that* $\boldsymbol{\Psi}$ *is constrained to be diagonal. Then for every* $\varepsilon > 0$ *there exists a constant* $K = K(\varepsilon)$ *such that*

$$\sup_{\boldsymbol{\theta} \in \mathcal{U}_{\boldsymbol{\theta}^0}} \mathcal{E}(||\mathbf{V}||^2 | \mathbf{Z} = \mathbf{z}; \boldsymbol{\theta}) \leq K \exp(\varepsilon ||\mathbf{z}||^2)$$

where $\mathcal{U}_{\boldsymbol{\theta}^0}$ *is as defined at the beginning of the proof of Theorem 3.*

Proof: Intuitively, the lemma is true since the bounds in Lemma A.1 are sharp with respect to the dominating (quadratic) terms. We shall prove the result for a specific $\boldsymbol{\theta} \in \mathcal{U}_{\boldsymbol{\theta}^0}$. Since $K = K(\varepsilon; \boldsymbol{\theta})$ will depend continuously on $\boldsymbol{\theta}$ and $\overline{\mathcal{U}}_{\boldsymbol{\theta}^0}$ is compact, $K(\varepsilon) := \sup_{\boldsymbol{\theta} \in \overline{\mathcal{U}}_{\boldsymbol{\theta}^0}} K(\varepsilon, \boldsymbol{\theta}) < \infty$.

For ease of notation, *we introduce*

$$S(v) := \sum_{j=1}^{m} (z^{(j)} - v)^2 / \sigma_j^2$$

(In the paper, we have set $m = 1$ *to simplify notation, but this is somewhat confusing since it contradicts the whole setup.)*

Set $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1^2, \ldots, \psi_p^2)$. In the sequel, $\lesssim$ means less or equal up to multiplicative terms of order at most $O(\exp(||\mathbf{z}||))$. !!! achtung !!! From Lemma A.1 it is clear that

$$\mathcal{E}(||\mathbf{V}||^2 | \mathbf{Z} = \mathbf{z}; \boldsymbol{\theta}) \lesssim \frac{\int_{\mathbb{R}^p} ||\mathbf{v}||^2 g_u(\mathbf{v}) dv}{\int_{\mathbb{R}^p} g_l(\mathbf{v}) dv} \tag{A.7}$$

where g_l *and* g_u *are defined as*

$$\begin{aligned} g_h(\mathbf{v}) &= \exp\Big(-\frac{1}{2} S(v^*) + \kappa_h^{(2)} |v^*| - \frac{1}{2} \sum_{k=1}^{p} \Big(\frac{v^{(k)} - \mu_k}{\psi_k} \Big)^2 \Big) d\mathbf{v} \\ v^* &= \max_{1 \leq k \leq p} \{ v^{(k)} \} \end{aligned}$$

with $h = l$ and $h = u$, respectively, and the constants $\kappa_u^{(2)}, \kappa_l^{(2)}$ are those of Lemma A.1. We split the integrals over $\mathbb{R}^p$ as follows:

$$\int_{\mathbb{R}^p} = \int_{v^* = v^{(1)}; v^* \leq \mu_*} + \int_{v^* = v^{(1)}; v^* > \mu_*} + \ldots + \int_{v^* = v^{(p)}; v^* > \mu_*}$$

where $\mu_* := \min\{\mu_1, \ldots, \mu_p\}$. Thus, the numerator and the denominator of (A.7) can be rewritten as the sum of $2p$ integrals each:

$$\mathcal{E}(||\mathbf{V}||^2 | \mathbf{Z} = \mathbf{z}; \boldsymbol{\theta}) \lesssim \frac{I_u^{(1,1)} + I_u^{(1,2)} + \ldots + I_u^{(p,1)} + I_u^{(p,2)}}{I_l^{(1,1)} + I_l^{(1,2)} + \ldots + I_l^{(p,1)} + I_l^{(p,2)}}$$

We now show specifically that the bound (A.7) applies to $I_u^{(1,1)} / I_l^{(1,1)}$ (the integrals over $\{v^* = v^{(1)}; v^* \leq \mu_*\}$). For every $\delta > 0$ there exists a constant $K = K(\delta)$ such that

$$||\mathbf{v}||^2 \leq K \exp\Big(\frac{\delta}{2} \sum_{k=1}^{p} \Big(\frac{v^{(k)} - \mu_k}{\psi_k} \Big)^2 \Big). \tag{A.8}$$

Setting $q_k := \psi_k/(1-\delta)^{1/2}$ (where $\delta < 1$ was assumed) and using (A.8), the integral $I_u^{(1,1)}$ can be simplified by directly integrating over $v^{(2)}, \dots, v^{(p)}$ from $-\infty$ to $v^{(1)}$,

$$I_u^{(1,1)} \lesssim \int_{-\infty}^{\mu_*} \exp\Big(-\frac{1}{2}S(v) + \kappa_u^{(2)}|v| - \frac{1}{2}\Big(\frac{v-\mu_1}{q_1}\Big)^2\Big) \times \prod_{k=2}^{p} \Phi\Big(\frac{v-\mu_k}{q_k}\Big) dv$$

For $v \in]-\infty, \mu_*]$,

$$(v-\mu_k)/q_k < 0 \quad \text{and} \quad \kappa_u^{(2)}|v| < -\kappa_u^{(2)} v + 2\kappa_u^{(2)}|\mu_*|$$

are satisfied, where we assumed, without loss of generality, that $\kappa_u^{(2)} > 0$. Using $\Phi(x) \leq \sqrt{\frac{\pi}{2}}\phi(x)$ for $x < 0$ (see Lemma A.4), we get

$$I_u^{(1,1)} \lesssim \int_{-\infty}^{\mu_*} \exp\Big(-\frac{1}{2}S(v) - \kappa_u^{(2)} v - \frac{1}{2}\sum_{k=1}^{p}\Big(\frac{v-\mu_k}{q_k}\Big)^2\Big) dv \tag{A.9}$$

The integral can be evaluated explicitly after quadratic completion. The result is of the form

$$I_u^{(1,1)} \lesssim \exp(\alpha_u^{(1)} + \alpha_u^{(2)} z + \alpha_u^{(3)} z^2)\Phi(\beta_u^{(1)} + \beta_u^{(2)} z). \tag{A.10}$$

In a very similar way, we can obtain an upper bound for $I_l^{(1)}$ which is of the same form as (A.10):

$$I_l^{(1,1)} \gtrsim \exp(\alpha_l^{(1)} + \alpha_l^{(2)} z + \alpha_l^{(3)} z^2)\Phi(\beta_l^{(1)} + \beta_l^{(2)} z). \tag{A.11}$$

The key fact about inequalities (A.10) and (A.11) is that the dominating terms $\alpha_u^{(3)}$ and $\alpha_l^{(3)}$ as well as $\beta_u^{(2)}$ and $\beta_l^{(2)}$ differ by an arbitrarily small number which goes to zero as $\delta \downarrow 0$, since all applied inequalities are (almost) sharp with respect to the dominating terms. Thus, (A.10) and (A.11) can be used to bound $I_u^{(1,1)}/I_l^{(1,1)}$ from above and applying Lemma A.4 once again to the terms of this ratio yields the desired bound *for* $I_u^{(1,1)}/I_l^{(1,1)}$.

The case of $I_u^{(1,2)}/I_l^{(1,2)}$ *runs along the same lines. Finally, since all the ratios* $I_u^{(k,l)}/I_l^{(k,l)}$ *are bounded by the right hand side in (A.7), so is* $\sum_{k,l} I_u^{(k,l)}/\sum_{k,l} I_l^{(k,l)}$. □

Lemma A.6 *The assertion of Lemma A.5 holds for general* $\mathbf{\Psi}$ *if* $p = 2$.

Proof: The proof is very similar to that of Lemma A.5 . Therefore we will be somewhat sketchy. For ease of notation, we set again $m = 1$, so that $\sum_{j=1}^{m}(z^{(j)} - v^*)^2/\sigma_j^2$ in Lemma A.1 reduces to $(z-v^*)^2/\sigma^2$. Using Lemma A.1, $\mathcal{E}(||\mathbf{V}||^2|\mathbf{Z} = \mathbf{z}; \boldsymbol{\theta})$ can be bounded by a ratio as in (A.7). Let's examine the numerator of this ratio:

For every $\delta > 0$ there exists a constant $K = K(\delta)$ such that

$$||\mathbf{v}||^2 \leq K \exp\Big(\frac{\delta}{2}(\mathbf{v} - \boldsymbol{\mu})^t \mathbf{\Psi}^{-1}(\mathbf{v} - \boldsymbol{\mu})\Big).$$

Then the numerator of the above mentioned ratio satisfies

$$\begin{aligned}\int_{\mathbb{R}^2} ||\mathbf{v}||^2 f_{\mathbf{Z},\mathbf{V}}(\mathbf{z}, \mathbf{v}; \boldsymbol{\theta}) \;\; &\lesssim \;\; \int_{\mathbb{R}^2} \exp\Big(-\frac{1}{2}(z - v^*)^2/\sigma^2 + \kappa_u^{(2)}|v^*| \\ &\quad -\frac{1}{2}(1-\delta)(\mathbf{v} - \boldsymbol{\mu})^t \mathbf{\Psi}^{-1}(\mathbf{v} - \boldsymbol{\mu})\Big) d\mathbf{v}\ . \end{aligned} \tag{A.12}$$

Now, split the integral (A.12): $\int_{\mathbb{R}^2} = \int_{v^*=v^{(1)}} + \int_{v^*=v^{(2)}}$. Consider specifically the first integral which is equal to

$$\begin{aligned}&\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{v^{(1)}} \exp\Big(-\frac{1}{2}(1-\delta)(\mathbf{v} - \boldsymbol{\mu})^t \mathbf{\Psi}^{-1}(\mathbf{v} - \boldsymbol{\mu})\Big) dv^{(2)}\Big) \cdot \\ &\qquad \exp\Big(-\frac{1}{2}(z - v^{(1)})^2/\sigma^2 + \kappa_u^{(2)}|v^{(1)}|\Big) dv^{(1)}. \end{aligned} \tag{A.13}$$

The inner integral can be evaluated explicitly after quadratic completion. The result has the form

$$\text{const} \cdot \exp(\alpha + \beta v^{(1)} + \gamma (v^{(1)})^2) \cdot \Phi(a + b v^{(1)}) \tag{A.14}$$

for constants $\alpha, \beta, \gamma, a, b$. Now the outer integral in (A.13) can be split further into integrals over $]-\infty, a/b]$ and $[a/b, \infty[$ so that $a + bv^{(1)} \leq 0$ on one region and ≥ 0 on the other. It is clear that Lemma A.4 can be applied to (A.14) if $a + bv^{(1)} \leq 0$, and that $1/2 \leq \Phi(a + bv^{(1)}) \leq 1$ if $a + bv^{(1)} > 0$. Thus, we can proceed analogously to the proof of Lemma A.5. □

Lemma A.7 *If $\boldsymbol{\theta}$ is an inner point of the parameter space the Fisher information matrix $\mathbf{I}_{\boldsymbol{\theta}}$ is nonsingular.*

Proof: It suffices to show that

$$\int_{\mathbb{R}^p} (\boldsymbol{\alpha}^t(\dot{f}_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})/f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})))^2 f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) = 0$$

for a vector $\boldsymbol{\alpha} \in \mathbb{R}^{(m-1)p+p(p+3)/2+m}$ implies $\boldsymbol{\alpha} = \mathbf{0}$.

$f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})$ is continuous and positive on its entire domain. By Lemma A.2 and by continuity of $(\mathbf{z},\mathbf{u}) \mapsto \frac{\partial}{\partial \boldsymbol{\theta}} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})$ the lemma is proved if it can be shown that $\boldsymbol{\alpha}^t \mathbf{d} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) \equiv 0$ implies $\boldsymbol{\alpha} = \mathbf{0}$, where $\mathbf{d}$ is defined by $\frac{\partial}{\cdot}\partial\boldsymbol{\theta} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) = \mathbf{d} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})$. (Remember that $\mathbf{C}$ is parametrized by $m(p-1)$ parameters, i.e. its first $m-1$ rows.) Define

$$\begin{aligned} g(\mathbf{z},\mathbf{u}) = \boldsymbol{\alpha}^t \mathbf{d} &= \sum_{j=1}^{m-1}\sum_{k=1}^{p} \alpha_{jk} d^{(c_{jk})} + \sum_{k=1}^{p} \beta_k d^{(\mu_k)} \\ &\quad + \sum_{\ell=1}^{p}\sum_{k=\ell}^{p} \gamma_{\ell k} d^{(r_{\ell k})} + \sum_{j=1}^{m} \delta_j d^{(\sigma_j)} , \end{aligned}$$

where we have set $\boldsymbol{\alpha} = ((\alpha_{jk}), (\beta_k), (\gamma_{lk}), (\delta_j))$.

Quadratic terms of $g(\mathbf{z},\mathbf{u})$ in $z^{(j)}$ appear only in $d^{(\sigma_j)}$. Thus $g(\mathbf{z},\mathbf{u}) \equiv 0$ implies $\delta_j = 0$. Now consider the coefficients of the terms that are linear in $z^{(j)}$ $(j = 1, \ldots, m-1)$,

$$\ell_j(\mathbf{u}) = \frac{1}{\sigma_j^2 h_j(\mathbf{u})} \sum_{k=1}^{p} \Big(\alpha_{jk} + c_{jk}\beta_k + c_{jk} \sum_{\ell=1}^{k} \gamma_{\ell k} u^{(\ell)} \Big) \exp(\boldsymbol{\mu} + \mathbf{R}^t \mathbf{u})^{(k)} ,$$

where $h_j(\mathbf{u})$ is as before (A.3). By setting $\mathbf{v} = \boldsymbol{\mu} + \mathbf{R}^t\mathbf{u}$, expanding $\ell_j(\mathbf{u})$ (written in terms of $\mathbf{v}$) in a Taylor series and equating coefficients of the polynomial terms to 0 it can be seen that $\ell(\mathbf{u};j) \equiv 0$ implies $\alpha_{jk} + c_{jk}\beta_k = 0$ and $\gamma_{\ell k} = 0$. Taking into account linear terms of $g(\mathbf{z},\mathbf{u})$ in $z^{(m)}$ in the same way then establishes that $\alpha_{jk} = 0$ and $\beta_k = 0$. □

3 Asymptotic Covariance Matrix

The use of the asymptotic normal distribution for inference is sketched in Section 4.5.3 of Wolbers (2002), which is reproduced here as Subsection 1 for convenience. Subsection 2 expands the formulas. Note that Marcel Wolbers found the asymptotic results to be unreliable. They are therefore provided only for researchers who want to collect more experience, and maybe for getting a crude idea about precision.

3.1 Estimation of the Fisher information and asymptotic confidence regions

We would be very surprised if asymptotic normality of the MLE would only hold for the special cases covered by Theorem 2. Rather, it seems plausible that asymptotic normality holds generally (as long as the parameters are interior points of the parameter space) but it should be kept in mind that the following discussion implicitely requires asymptotic normality.

One major reason for asymptotic theory is that it allows to construct **approximate confidence regions** for the estimated parameters. One of the simplest ways to do so is to use that if $\widehat{\mathbf{I}}_n$ is a consistent estimate of the Fisher information matrix $\mathbf{I}_{\boldsymbol{\theta}^0}$, then

$$\{\boldsymbol{\theta} : n(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)^t \widehat{\mathbf{I}}_n (\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n) < \chi^2_{0.95, df=m(p-1)+p(p+3)/2+m}\}$$

is an asymptotic 95%-confidence ellipsoid for $\boldsymbol{\theta}^0$. (Here, as in Section 4 of the paper, $\mathbf{C}$ is parametrized by its first $m-1$ rows. Thus, the total number of parameters is $m(p-1) + p(p+3)/2 + m$.) Likewise, one can construct univariate **95%-confidence intervals** for specific parameters θ: $\{\widehat{\theta}_n \pm 1.96(\widehat{\mathbf{I}}_n^{-1}(\theta)/n)^{1/2}\}$.

Two possible **consistent estimates of the Fisher information** are given by

$$\begin{aligned} \widehat{\mathbf{I}}_n^{(1)} &:= \frac{1}{n}\sum_{i=1}^{n}(\frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\mathbf{Z}}(\mathbf{z}_i; \widehat{\boldsymbol{\theta}}_n))(\frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\mathbf{Z}}(\mathbf{z}_i; \widehat{\boldsymbol{\theta}}_n))^t \\ \widehat{\mathbf{I}}_n^{(2)} &:= -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \log f_{\mathbf{Z}}(\mathbf{z}_i; \widehat{\boldsymbol{\theta}}_n) \end{aligned}$$

Since the log-likelihood is not available in closed form, it is not trivial to actually calculate these estimates. However, it is discussed in Section 5 of the paperhow to approximate the log-likelihood and its derivatives by Monte Carlo importance sampling and exactly the same method can be used to obtain approximations of $\widehat{\mathbf{I}}_n^{(1)}$ and in principle also of $\widehat{\mathbf{I}}_n^{(2)}$. It is much easier to efficiently implement the approximation of $\widehat{\mathbf{I}}_n^{(1)}$ since only first derivatives of $f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})$ need to be coded (see next Subsection) to obtain the importance sampling approximation.

Preliminary experiments for the MLEs calculated in Section 7 show that for the case $n = 100$ the coverage of the confidence ellipsoids (based on $\widehat{\mathbf{I}}_n^{(1)}$) is rather too high (especially for source profile 2). For the case $n = 250$, the coverage for source profiles 2 and 3 is acceptable (98 resp. 94 out of 100 confidence ellipsoids covered the true profiles), but too low for source profile 1 (only 78 out of 100; however, all univariate confidence intervals for individual components of source 1 had a coverage of at least 88%). Further research is needed to clarify in which cases asymptotic confidence intervals are valid for the lognormal structural mixing model.

If the structural lognormal mixing model is **misspecified** in that the true underlying distribution does not belong to the model neither $\big(\widehat{\mathbf{I}}_n^{(1)}\big)^{-1}$ nor $\big(\widehat{\mathbf{I}}_n^{(2)}\big)^{-1}$ is a consistent estimator of the asymptotic covariance matrix. An alternative estimator that will be consistent is $\big(\widehat{\mathbf{I}}_n^{(2)}\big)^{-1}\widehat{\mathbf{I}}_n^{(1)}\big(\widehat{\mathbf{I}}_n^{(2)}\big)^{-1}$, see van der Vaart (1998, Ex. 5.25, p. 55).

An alternative to asymptotic confidence intervals is to determine confidence intervals by resampling techniques such as the **bootstrap**, see e.g. Davison & Hinkley (1997). Since computation of the MLE is computer intensive, as described in Section 5 of the paper,the bootstrap takes considerable computing time. Nevertheless, we chose to use the blockwise bootstrap to assess uncertainty in an application to account for dependent observations, see Section 8 of the paper.

3.2 Formulas for the Fisher Information Matrix

In order to calculate $\widehat{\mathbf{I}}_n^{(1)}$, we need the derivatives of the log likelihood with respect to the parameter components. The respective formulas are collected from Sections 4.4 and Appendix A.1 of Wolbers (2002) and partly rewritten to ease understanding and programming. The likelihood can be written as

$$f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) = \int f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}$$

where $f_{\mathbf{U}}(\mathbf{u};\boldsymbol{\theta}) = (2\pi)^{-p/2}\exp\Big(-\frac{1}{2}\mathbf{u}^t\mathbf{u}\Big)$ does not depend on the parameters, and

$$\begin{aligned} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) &= (2\pi)^{-m/2}\cdot\frac{1}{\sigma_1\cdots\sigma_m}\cdot\exp\Big(-\frac{1}{2}\sum\nolimits_{j=1}^{m} g_j(\mathbf{u};\boldsymbol{\theta})^2\Big) \\ g_j(\mathbf{u};\boldsymbol{\theta}) &= \big(z^{(j)} - \log(h_j(\mathbf{u};\boldsymbol{\theta}))\big)/\sigma_j \\ h_j(\mathbf{u};\boldsymbol{\theta}) &= \sum_{k=1}^{p} c_{jk}\exp(\boldsymbol{\mu} + \mathbf{R}^t\mathbf{u})^{(k)} = \sum_{k=1}^{p} c_{jk} S^{(k)} \end{aligned}$$

Note that h_j is the fit of the model for the variable $X^{(j)}$, g_j is the residual on the log scale, and $S^{(k)}$ is the score of source k.

The derivatives of $\log(f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}))$ with respect to the parameter, the Fisher scores, are as follows. The standard deviation σ_j appears in the normalizing constant and in g_j, and we get

$$\psi(z;\boldsymbol{\theta})^{(\sigma_j)} = -\sigma_j^{-1} + \frac{1}{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}\int \frac{\partial}{\partial\sigma_j} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}$$

whereas all the other parameter components only show up in g_j, and we obtain

$$\psi(z;\boldsymbol{\theta})^{(t)} = \frac{1}{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}\int \frac{\partial}{\partial\theta_t} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}$$

The derivatives of $f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})$ are of the form

$$\frac{\partial}{\partial\boldsymbol{\theta}} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) = -f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) \sum_{j=1}^{m} g_j(\mathbf{u};\boldsymbol{\theta})\ \frac{\partial}{\partial\boldsymbol{\theta}} g_j(\mathbf{u};\boldsymbol{\theta})$$

The derivative of g_j with respect to σ_j is

$$\frac{\partial}{\partial\sigma_j} g_j(\mathbf{u};\boldsymbol{\theta}) = -g_j(\mathbf{u};\boldsymbol{\theta})/\sigma_j$$

and the other derivatives are of the form

$$\frac{\partial}{\partial\theta_\ell} g_j(\mathbf{u};\boldsymbol{\theta}) = -\sigma_j^{-1}\frac{\partial}{\partial\theta_\ell} h_j(\mathbf{u};\boldsymbol{\theta}) \Big/ h_j(\mathbf{u};\boldsymbol{\theta})\ .$$

We have

$$\begin{aligned}
\frac{\partial}{\partial c_{jk}} h_j(\mathbf{u};\boldsymbol{\theta}) &= S^{(k)} \\
\frac{\partial}{\partial \mu_k} h_j(\mathbf{u};\boldsymbol{\theta}) &= \sum_{j=1}^{m} c_{jk} S^{(k)} \\
\frac{\partial}{\partial r_{\ell k}} h_j(\mathbf{u};\boldsymbol{\theta}) &= \sum_{j=1}^{m} c_{jk} u^{(\ell)} S^{(k)}
\end{aligned}$$

Since the columns of $\mathbf{C}$ must sum up to 1, the Fisher Information for the full parameter is singular. If we drop the last row of $\mathbf{C}$ as suggested in Section 4 of the paper, the derivatives of $f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})$ with respect to the c_{jk}, $j \le m-1$ are

$$\frac{\partial}{\partial c_{jk}} f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) = f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta}) \left(\frac{g_j(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_j(\mathbf{u};\boldsymbol{\theta})} - \frac{g_m(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_m(\mathbf{u};\boldsymbol{\theta})} \right) \exp(\boldsymbol{\mu} + \mathbf{R}^t\mathbf{u})^{(k)}\ .$$

Collecting the results, we have

$$\psi(z;\boldsymbol{\theta}) = \frac{1}{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}\int \mathbf{q}(\mathbf{u};\boldsymbol{\theta})\, f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}(\mathbf{z};\boldsymbol{\theta})\, f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}$$

with

$$
\begin{aligned}
\mathbf{q}(\mathbf{u};\boldsymbol{\theta})^{(\sigma_j)} &= (-1+g_j(\mathbf{u};\boldsymbol{\theta})^2)/\sigma_j \\
\mathbf{q}(\mathbf{u};\boldsymbol{\theta})^{(c_{jk})} &= \left(\frac{g_j(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_j(\mathbf{u};\boldsymbol{\theta})} - \frac{g_m(\mathbf{u};\boldsymbol{\theta})}{\sigma_m h_m(\mathbf{u};\boldsymbol{\theta})}\right) S^{(k)} \\
\mathbf{q}(\mathbf{u};\boldsymbol{\theta})^{(\mu_k)} &= \sum_{j=1}^{m} \frac{g_j(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_j(\mathbf{u};\boldsymbol{\theta})}\, c_{jk} S^{(k)} \\
\mathbf{q}(\mathbf{u};\boldsymbol{\theta})^{(r_{\ell k})} &= \sum_{j=1}^{m} \frac{g_j(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_j(\mathbf{u};\boldsymbol{\theta})}\, c_{jk} u^{(\ell)} S^{(k)}
\end{aligned}
$$

If the scores are modeled by regression with respect to covariates, these results are easily extended. Since U has been used above, we call the covariates a_ℓ. The model says that the expected value for the log scores for observation i depends on the covariates through $\boldsymbol{\mu}_i = \mathbf{a}_i^t \mathbf{B}$, where the element $\beta_{k\ell}$ is the coefficient of covariate ℓ for score k. The parameter $\boldsymbol{\mu}$ is therefore replaced by $\mathbf{B}$, and the derivatives are

$$
\begin{aligned}
\frac{\partial}{\partial \beta_{\ell k}} h_j(\mathbf{u};\boldsymbol{\theta}) &= c_{jk} S^{(k)} a_{i\ell} \\
\mathbf{q}(\mathbf{u};\boldsymbol{\theta})^{(\beta_{\ell k})} &= \sum_{j=1}^{m} \frac{g_j(\mathbf{u};\boldsymbol{\theta})}{\sigma_j h_j(\mathbf{u};\boldsymbol{\theta})}\, c_{jk} S_i^{(k)} a_{i\ell}\,.
\end{aligned}
$$

Calculations are done by importance sampling as for the E-step in the algorithms. Note that normalizing constants of $f_{\mathbf{Z}}$ and $f_{\mathbf{Z}|\mathbf{U}=\mathbf{u}}$ cancel in the formulas, and therefore need not be computed.

Finally, the estimate of the Fisher Information ist

$$
\widehat{\mathbf{I}}_n^{(1)} = \frac{1}{n} \sum_{i=1}^{n} \psi(z_i, a_i; \boldsymbol{\theta})\, \psi(z_i, a_i; \boldsymbol{\theta})^t
$$

The estimate of the covariance matrix of the estimated parameter vector is $(\widehat{\mathbf{I}}_n^{(1)})^{-1}/n$. Standard errors of the estimated parameters are, of course, the diagonal elements. The standard error of $\widehat{c}_{mk}$ is obtained as the sum over all elements of the covariance matrix of $\widehat{c}_{1k}, ..., \widehat{c}_{m-1,k}$. The estimated covariance matrix $\widehat{\Psi}$ of the scores is less important. Its covariance matrix would need to be calculated from a linearization of $\widehat{\Psi} = \widehat{R}^t \widehat{R}$.

4 Regression of Scores

The relationship between the scores and potential explanatory variables may be introduced into the model as described in Section 6.2 of the paper. This should not only give insight into effects of such variables as traffic counts or weather conditions, but should help to identify the source profiles at the same time.

Some preliminary experience and thoughts indicate that identification is facilitated only, or mainly, if (partly) different explanatory variables are provided for the different sources, or, in other words, some elements of the coefficient matrix are fixed (at 0).

References

Bauer, H. (1992), *Maß- und Integrationstheorie (Second Edition)*, Walter de Gruyter & Co., Berlin.

Davison, A. C. & Hinkley, D. V. (1997), *Bootstrap methods and their application*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Feller, W. (1968), *An Introduction to Probability Theroy and Its Applications (Third Edition)*, Vol. I, Wiley, New York.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press.

Wolbers, M. (2002), Linear unmixing of multivariate observations, PhD thesis, ETH Zürich (Diss. ETH 14723).