Supporting document for "Bayesian Cramér-Rao Bounds for factorized model based low-rank matrix reconstruction"

Martin Sundin

August 8, 2016

1 Introduction

I this document we show the calculations and derivations for the paper "Bayesian Cramér-Rao Bounds for factorized model based low-rank matrix reconstruction" by Martin Sundin, Magnus Jansson and Saikat Chatterjee. The paper was presented at EUSIPCO 2016 in Budapest, Hungary.

When citing, please cite the main article.

Supplementary Material

1.1 Proof of Propostion 1

Proposition 1 gives the BCRB bounds (7) and (8). The derivation of the BCRB is similar to the derivation of the deterministic CRB in [1, 2] and was earlier given in [3, 4, 5]. For completeness we here repeat the derivation of the bounds.

Let $\hat{\boldsymbol{\eta}}$ be an unbiased estimator of $\boldsymbol{\eta} = \mathbf{g}(\mathbf{z}) \in \mathbb{R}^K$ from data $\mathbf{y} \in \mathbb{R}^m$ and assume that the probability distribution function $p(\mathbf{y}, \mathbf{z})$ is defined for $\mathbf{z} \in \Omega \subset \mathbb{R}^n$ with $p(\mathbf{y}, \mathbf{z}) = 0$ for points $\mathbf{z}$ on the boundary, $\mathbf{z} \in \partial\Omega$. Let also $\mathbf{a} \in \mathbb{R}^K$ be a constant vector and $\mathbf{b} = \mathbf{b}(\mathbf{z}) \in \mathbb{R}^n$ be a vector which depends on $\mathbf{z}$. We find that

$$
\begin{aligned}
&\mathcal{E}_{\mathbf{y},\mathbf{z}}\left[\mathbf{a}^\top(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})\mathbf{b}^\top \frac{\partial \log p(\mathbf{y}, \mathbf{z})}{\partial \mathbf{z}}\right] \\
&= \int_{\mathbb{R}^m}\int_{\Omega} \mathbf{a}^\top(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})\mathbf{b}^\top \frac{\partial p(\mathbf{y}, \mathbf{z})}{\partial \mathbf{z}} d\mathbf{y} d\mathbf{z} \\
&= -\int_{\mathbb{R}^m}\int_{\Omega} \mathrm{tr}\left(\frac{\partial}{\partial \mathbf{z}}(\mathbf{a}^\top(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})\mathbf{b})\right) p(\mathbf{y}, \mathbf{z}) d\mathbf{y} d\mathbf{z} \\
&= \mathcal{E}_{\mathbf{z}}\left[\mathbf{a}^\top \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \mathbf{b}\right] - \mathcal{E}_{\mathbf{z}}\left[\mathbf{a}^\top \mathcal{E}_{\mathbf{y}}\left[(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})\right] \mathrm{tr}\left(\frac{\partial \mathbf{b}}{\partial \mathbf{z}}\right)\right] \\
&= \mathcal{E}_{\mathbf{z}}\left[\mathbf{a}^\top \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \mathbf{b}\right]
\end{aligned}
$$

where we used that $\hat{\boldsymbol{\eta}}$ only depends on $\mathbf{y}$ and that $\mathcal{E}_{\mathbf{y}}\left[\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\right] = \mathbf{0}$.

The Cauchy-Schwartz inequality gives us that

$$
\begin{aligned}
&\left(\mathcal{E}_{\mathbf{z}}\left[\mathbf{a}^\top \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \mathbf{b}\right]\right)^2 \leq \\
&\mathcal{E}_{\mathbf{y},\mathbf{z}}\left[\left((\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^\top \mathbf{a}\right)^2\right] \mathcal{E}_{\mathbf{y},\mathbf{z}}\left[\left(\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathbf{z}}^\top \mathbf{b}\right)^2\right] \\
&= \mathbf{a}^\top \mathbf{C}_{\boldsymbol{\epsilon}} \mathbf{a} \cdot \mathcal{E}_{\mathbf{z}}\left[\mathbf{b}^\top \mathbf{F} \mathbf{b}\right]
\end{aligned}
\tag{1}
$$

where we set

$$
\begin{aligned}
\mathbf{C}_{\boldsymbol{\epsilon}} &= \mathcal{E}_{\mathbf{y},\mathbf{z}}\left[(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^\top\right], \\
\mathbf{F} &= \mathcal{E}_{\mathbf{y}}\left[\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathbf{z}} \frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathbf{z}}^\top\right].
\end{aligned}
$$

From (1) we can derive bounds by choosing $\mathbf{b}$ appropriately. Setting

$$
\mathbf{b} = \left(\mathcal{E}_{\mathbf{z}}[\mathbf{F}]\right)^{-1} \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right]^\top \mathbf{a},
$$

gives us that

$$
\mathbf{a}^\top \mathbf{C}_{\boldsymbol{\epsilon}} \mathbf{a} \geq \mathbf{a}^\top \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right] \left(\mathcal{E}_{\mathbf{z}}[\mathbf{F}]\right)^{-1} \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right]^\top \mathbf{a},
$$

for all $\mathbf{a} \in \mathbb{R}^n$. It follows that

$$
\mathbf{C}_{\boldsymbol{\epsilon}} \succeq \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right] \left(\mathcal{E}_{\mathbf{z}}[\mathbf{F}]\right)^{-1} \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}}\right]^\top.
$$

This is the bound (7). Another choice is to set

$$
\mathbf{b} = \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top \left(\mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}} \mathbf{F} \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top\right]\right)^{-1} \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}} \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top\right] \mathbf{a}
$$

which gives us that

$$
\mathbf{C}_{\boldsymbol{\epsilon}} \succeq \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}} \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top\right] \left(\mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}} \mathbf{F} \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top\right]\right)^{-1} \mathcal{E}_{\mathbf{z}}\left[\frac{\partial \mathbf{g}}{\partial \mathbf{z}} \frac{\partial \mathbf{g}}{\partial \mathbf{z}}^\top\right].
$$

This gives us the bound (8).

1.2 Proof of Proposition 2

Proposition 2 gives the Fisher information matrix of the factorized model. In the factorized model, we have that

$$\begin{aligned}
\log p(\mathbf{y},\mathbf{z}) = & -\frac{\beta}{2}||\mathbf{y}-\mathbf{A}\mathrm{vec}(\mathbf{L}\mathbf{R}^\top)||_2^2 + \frac{m+2(c-1)}{2}\log\beta \\
& - d\beta + \frac{p+q+2(a-1)}{2}\log|\boldsymbol{\Gamma}| - \frac{1}{2}\mathrm{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top) - \frac{1}{2}\mathrm{tr}(\mathbf{R}\boldsymbol{\Gamma}\mathbf{R}^\top) \\
& - b\mathrm{tr}(\boldsymbol{\Gamma}).
\end{aligned}$$

We find that

$$\begin{aligned}
\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathrm{vec}(\mathbf{L})} = & \beta(\mathbf{R}\otimes\mathbf{I}_p)^\top\mathbf{A}^\top(\mathbf{y}-\mathbf{A}\mathrm{vec}(\mathbf{L}\mathbf{R}^\top)) \\
& - \mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}), \\
\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathrm{vec}(\mathbf{R})} = & \beta\mathbf{K}_{q,r}(\mathbf{I}_q\otimes\mathbf{L})^\top\mathbf{A}^\top(\mathbf{y}-\mathbf{A}\mathrm{vec}(\mathbf{L}\mathbf{R}^\top)) \\
& - \mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma}), \\
\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \gamma_i} = & \frac{p+q+2(a-1)}{2\gamma_i} - \frac{||\mathbf{l}_i||_2^2 + ||\mathbf{r}_i||_2^2}{2} - b = h_i, \\
\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \beta} = & -\frac{1}{2}||\mathbf{y}-\mathbf{A}\mathrm{vec}(\mathbf{L}\mathbf{R}^\top)||_2^2 + \frac{m+2(c-1)}{2\beta} - d.
\end{aligned}$$

Setting $\mathbf{h} = [h_1, h_2, \ldots, h_r]^\top$, we get that

$$\begin{aligned}
& \mathbf{F}_{\mathbf{ll}} = \mathcal{E}_{\mathbf{y}}\left[\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathrm{vec}(\mathbf{L})}\frac{\partial \log p(\mathbf{y},\mathbf{z})}{\partial \mathrm{vec}(\mathbf{L})}^\top\right] \\
& = \beta(\mathbf{R}\otimes\mathbf{I}_p)^\top\mathbf{A}^\top\mathbf{A}(\mathbf{R}\otimes\mathbf{I}_p) + \mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma})\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma})^\top, \\
& \mathbf{F}_{\mathbf{lr}} = \beta(\mathbf{R}\otimes\mathbf{I}_p)^\top\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q\otimes\mathbf{L})\mathbf{K}_{r,q} \\
& + \mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma})\mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma})^\top, \\
& \mathbf{F}_{\mathbf{rr}} = \beta\mathbf{K}_{q,r}(\mathbf{I}_q\otimes\mathbf{L})^\top\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q\otimes\mathbf{L})\mathbf{K}_{r,q} \\
& + \mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma})\mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma})^\top, \\
& \mathbf{F}_{\mathbf{l}\boldsymbol{\gamma}} = -\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma})\mathbf{h}^\top,\ \mathbf{F}_{\mathbf{r}\boldsymbol{\gamma}} = -\mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma})\mathbf{h}^\top, \\
& \mathbf{F}_{\boldsymbol{\gamma}} = \mathbf{h}\mathbf{h}^\top,\ \mathbf{F}_{L\beta} = \left(d - \frac{c-1}{\beta}\right)\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}), \\
& \mathbf{F}_{\mathbf{r}\beta} = \left(d - \frac{c-1}{\beta}\right)\mathrm{vec}(\mathbf{R}\boldsymbol{\Gamma}),\ \mathbf{F}_{\boldsymbol{\gamma}\beta} = \left(d - \frac{c-1}{\beta}\right)\mathbf{h}, \\
& F_{\beta\beta} = \frac{m}{2\beta^2} + \left(d - \frac{c-1}{\beta}\right)^2.
\end{aligned}$$

We find that $\mathbf{F}_{\mathbf{l}\boldsymbol{\gamma}}$, $\mathbf{F}_{\mathbf{r}\boldsymbol{\gamma}}$, $\mathbf{F}_{\mathbf{l}\beta}$, $\mathbf{F}_{\mathbf{r}\beta}$ and $\mathbf{F}_{\boldsymbol{\gamma}\beta}$ are zero when $\boldsymbol{\gamma}$ and β are deterministic ($a = c = 1$ and $b = d = 0$) and zero-mean when $\boldsymbol{\gamma}$ and β are random.

Using the expression for $\mathbf{F}$ and (11) we get that

$$\begin{aligned}
\mathbf{G}_{\mathbf{ww}} = \quad & \beta(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p) \\
& +\beta(\mathbf{I}_q \otimes \mathbf{L}\mathbf{L}^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q \otimes \mathbf{L}\mathbf{L}^\top) \\
& +\beta(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q \otimes \mathbf{L}\mathbf{L}^\top) \\
& +\beta(\mathbf{I}_q \otimes \mathbf{L}\mathbf{L}^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p), \\
& +4\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top)\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top)^\top, \\
\mathbf{G}_{\mathbf{w}\boldsymbol{\gamma}} = \quad & -2((\nabla_{\boldsymbol{\gamma}} s)^\top \mathbf{h})\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top), \\
\mathbf{G}_{\mathbf{w}\beta} = \quad & 2\left(d - \tfrac{c-1}{\beta}\right)\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top), \\
G_{\boldsymbol{\gamma\gamma}} = \quad & (\nabla_{\boldsymbol{\gamma}} s)^\top \mathbf{F}_{\boldsymbol{\gamma}}(\nabla_{\boldsymbol{\gamma}} s) = ((\nabla_{\boldsymbol{\gamma}} s)^\top \mathbf{h})^2, \\
G_{\boldsymbol{\gamma}\beta} = \quad & \left(d - \tfrac{c-1}{\beta}\right)((\nabla_{\boldsymbol{\gamma}} s)^\top \mathbf{h}), \\
G_{\beta\beta} = \quad & F_{\beta\beta}.
\end{aligned} \tag{2}$$

We find that the parameters $\mathbf{G}_{\mathbf{w}\boldsymbol{\gamma}}$, $\mathbf{G}_{\mathbf{w}\beta}$ and $G_{\boldsymbol{\gamma}\beta}$ are zero when $\boldsymbol{\gamma}$ and β are deterministic and zero mean when they are random.

1.3 Proof of Proposition 3

Proposition 3 gives the bounds BCRB-I and BCRB-II of the factorized model. To derive the BCRB-II for the factorized model, we need to compute expectation values with respect to $\mathbf{w}$. To compute $\mathcal{E}_{\mathbf{w}}[\mathbf{G}_{\mathbf{ww}}]$ we use that

$$\begin{aligned}
& \mathcal{E}_{\mathbf{w}}\left[\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top)\mathrm{vec}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{R}^\top)^\top\right] \\
& = \mathcal{E}_{\mathbf{w}}\left[\sum_{i,j=1}^{r} \mathrm{vec}(\gamma_i \mathbf{l}_i \mathbf{r}_i^\top)\mathrm{vec}(\gamma_j \mathbf{l}_j \mathbf{r}_j^\top)^\top\right] \\
& = \mathcal{E}_{\mathbf{w}}\left[\sum_{i,j=1}^{r} \gamma_i\gamma_j(\mathbf{r}_i \otimes \mathbf{l}_i)(\mathbf{r}_j \otimes \mathbf{l}_j)^\top\right] \\
& = \sum_{i,j=1}^{r} \gamma_i\gamma_j(\mathcal{E}_{\mathbf{w}}\left[\mathbf{r}_i\mathbf{r}_j^\top\right] \otimes \mathcal{E}_{\mathbf{w}}\left[\mathbf{l}_i\mathbf{l}_j^\top\right]) \\
& = \sum_{i,j=1}^{r} \delta_{ij}\gamma_i\gamma_j(\gamma_i^{-1}\mathbf{I}_q \otimes \gamma_j^{-1}\mathbf{I}_p) = r\mathbf{I}_{pq}, \\
& \mathcal{E}_{\mathbf{w}}\left[\beta(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q \otimes \mathbf{L}\mathbf{L}^\top)\right] = \\
& \beta(\mathcal{E}_{\mathbf{w}}[\mathbf{R}\mathbf{R}^\top] \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{I}_q \otimes \mathcal{E}_{\mathbf{w}}[\mathbf{L}\mathbf{L}^\top]) = \\
& \beta\left(\sum_{i=1}^{r}\gamma_i^{-1}\right)^2 (\mathbf{I}_q \otimes \mathbf{I}_p)\,\mathbf{A}^\top\mathbf{A}\,(\mathbf{I}_q \otimes \mathbf{I}_p)\,.
\end{aligned}$$

Expectations such as $\mathcal{E}_{\mathbf{w}}\left[\beta(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\right]$ are more challenging to calculate. To compute the expectation, we use that

$$
\begin{aligned}
&\left[(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\right]_{\substack{i+(k-1)p,\\ j+(l-1)p}} \\
&= (\mathbf{e}_k \otimes \mathbf{e}_i)^\top (\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)(\mathbf{e}_l \otimes \mathbf{e}_j) \\
&= (\mathbf{e}_k^\top \mathbf{R}\mathbf{R}^\top \otimes \mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top\mathbf{e}_l \otimes \mathbf{e}_j) \\
&= \operatorname{tr}\left((\mathbf{R}\mathbf{R}^\top\mathbf{e}_l\mathbf{e}_k^\top\mathbf{R}\mathbf{R}^\top \otimes \mathbf{e}_j\mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}\right).
\end{aligned}
$$

The i'th row vector of $\mathbf{R}$ is $\mathbf{R}^\top\mathbf{e}_i$, this gives us that

$$
\begin{aligned}
&\mathcal{E}_{\mathbf{w}}\left[\mathbf{e}_m^\top\mathbf{R}\mathbf{R}^\top\mathbf{e}_l\mathbf{e}_k^\top\mathbf{R}\mathbf{R}^\top\mathbf{e}_n\right] \\
&= \mathcal{E}_{\mathbf{w}}\left[(\mathbf{R}^\top\mathbf{e}_m)^\top(\mathbf{R}^\top\mathbf{e}_l)(\mathbf{R}^\top\mathbf{e}_k)^\top(\mathbf{R}^\top\mathbf{e}_n)\right] \\
&= \left(\sum_{i=1}^r \gamma_i^{-1}\right)^2 \delta_{ml}\delta_{kn} + \left(\sum_{i=1}^r \gamma_i^{-2}\right)(\delta_{lk}\delta_{mn} + \delta_{mk}\delta_{ln}) \\
&= \mathbf{e}_m^\top\left[\left(\sum_{i=1}^r \gamma_i^{-1}\right)^2 \mathbf{e}_l\mathbf{e}_k^\top + \left(\sum_{i=1}^r \gamma_i^{-2}\right)(\delta_{lk}\mathbf{I}_q + \mathbf{e}_k\mathbf{e}_l^\top)\right]\mathbf{e}_n.
\end{aligned}
$$

This gives us that

$$
\begin{aligned}
&\mathcal{E}_{\mathbf{w}}\left[(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\mathbf{A}^\top\mathbf{A}(\mathbf{R}\mathbf{R}^\top \otimes \mathbf{I}_p)\right]_{\substack{i+(k-1)p,\\ j+(l-1)p}} \\
&= \left(\sum_{i=1}^r \gamma_i^{-1}\right)^2 (\mathbf{e}_k^\top \otimes \mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{e}_l \otimes \mathbf{e}_j) \\
&+ \left(\sum_{i=1}^r \gamma_i^{-2}\right)(\mathbf{e}_l^\top \otimes \mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{e}_k \otimes \mathbf{e}_j) \\
&+ \left(\sum_{i=1}^r \gamma_i^{-2}\right)\delta_{lk}\operatorname{tr}\left[\left(\mathbf{I}_q \otimes \mathbf{e}_i^\top\right)\mathbf{A}^\top\mathbf{A}\left(\mathbf{I}_q \otimes \mathbf{e}_j\right)\right].
\end{aligned}
$$

Let $\mathcal{T}_1$ be the operator defined in section 1.1, we find that

$$
\begin{aligned}
(\mathbf{e}_k^\top \otimes \mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{e}_l \otimes \mathbf{e}_j) &= [\mathbf{A}^\top\mathbf{A}]_{i+(k-1)p,j+(l-1)p}, \\
(\mathbf{e}_l^\top \otimes \mathbf{e}_i^\top)\mathbf{A}^\top\mathbf{A}(\mathbf{e}_k \otimes \mathbf{e}_j) &= [\mathcal{T}_1(\mathbf{A}^\top\mathbf{A})]_{i+(k-1)p,j+(l-1)p},
\end{aligned}
$$

Using that $(\mathbf{e}_n \otimes \mathbf{I}_p)\mathbf{e}_j = (\mathbf{e}_n \otimes \mathbf{I}_p)(1 \otimes \mathbf{e}_j) = (\mathbf{e}_n \otimes \mathbf{e}_j)$, we get that

$$
\begin{aligned}
&\delta_{lk}\operatorname{tr}\left[\left(\mathbf{I}_q \otimes \mathbf{e}_i^\top\right)\mathbf{A}^\top\mathbf{A}\left(\mathbf{I}_q \otimes \mathbf{e}_j\right)\right] \\
&= \mathbf{e}_l^\top\mathbf{e}_k\operatorname{tr}\left[\left(\mathbf{I}_q \otimes \mathbf{e}_j\mathbf{e}_i^\top\right)\mathbf{A}^\top\mathbf{A}\right] \\
&= \mathbf{e}_l^\top\mathbf{e}_k\sum_{n=1}^r\left(\mathbf{e}_n^\top \otimes \mathbf{e}_i^\top\right)\mathbf{A}^\top\mathbf{A}\left(\mathbf{e}_n \otimes \mathbf{e}_j\right) \\
&= \left[\left(\mathbf{I}_q \otimes \left(\sum_{n=1}^r\left(\mathbf{e}_n^\top \otimes \mathbf{I}_p\right)\mathbf{A}^\top\mathbf{A}\left(\mathbf{e}_n \otimes \mathbf{I}_p\right)\right)\right)\right]_{\substack{i+(k-1)p,\\ j+(l-1)p}}.
\end{aligned}
$$

A similar computation can be made for $\mathbf{L}$.

To compute

$$G_{\boldsymbol{\gamma}\boldsymbol{\gamma}} = \mathcal{E}_{\mathbf{w}}[((\nabla_{\boldsymbol{\gamma}} s)^\top \mathbf{h})^2] = (\nabla_{\boldsymbol{\gamma}} s)^\top \mathcal{E}_{\mathbf{w}}[\mathbf{h}\mathbf{h}^\top](\nabla_{\boldsymbol{\gamma}} s)$$

we use that

$$\mathcal{E}_{\mathbf{w}}[\mathbf{h}\mathbf{h}^\top] = \mathcal{E}_{\mathbf{w}}[\mathbf{h}]\mathcal{E}_{\mathbf{w}}[\mathbf{h}]^\top + \mathrm{Cov}(\mathbf{h}),$$

where the covariance is diagonal since the precisions are independent. We find (after a somewhat lengthy calculation) that

$$\begin{aligned} \mathcal{E}_{\mathbf{w}}[h_i] &= \frac{a-1}{\gamma_i} - b, \\ \mathcal{E}_{\mathbf{w}}\left[(h_i - \mathcal{E}_{\mathbf{z}}[h_i])^2\right] &= \frac{p+q}{2\gamma_i^2}. \end{aligned}$$

We see that $\mathcal{E}_{\mathbf{w}}[h_i] = 0$ when $\boldsymbol{\gamma}$ is deterministic. So

$$\begin{aligned} G_{\boldsymbol{\gamma}\boldsymbol{\gamma}} &= ((\nabla_{\boldsymbol{\gamma}} s)^\top((a-1)\boldsymbol{\gamma}^{-1} - b\mathbf{1}_r))^2 + \frac{p+q+2(a-1)}{2}(\nabla_{\boldsymbol{\gamma}} s)^\top \boldsymbol{\Gamma}^{-2}(\nabla_{\boldsymbol{\gamma}} s) \\ &= \frac{p+q}{2}(\nabla_{\boldsymbol{\gamma}} s)^\top \boldsymbol{\Gamma}^{-2}(\nabla_{\boldsymbol{\gamma}} s). \end{aligned}$$

1.4 Proof of Proposition 4

The bound BCRB-III can be computed from BCRB-II by taking the appropriate expectation values with respect to $\boldsymbol{\gamma}$ and β. Using that

$$\begin{aligned} \mathcal{E}_\beta[\beta] = \tfrac{c}{d}, \quad \mathcal{E}_\beta[\beta^{-2}] &= \tfrac{d^2}{(c-1)(c-2)}, \\ \mathcal{E}_{\boldsymbol{\gamma}}[\gamma_i^{-k}] = b^k \frac{\Gamma(a-k)}{\Gamma(a)} &= b^k \prod_{i=0}^{k-1}(a-1-i)^{-1}, \end{aligned}$$

we are able to compute the respective expectation values.

Full the full reference list, see the conference paper.

References

[1] S.M. Kay, "Fundamentals of statistical signal processing, volume I: estimation theory", Pearson Education (1993).

[2] H. Cramér, "Mathematical methods of statistics", Princeton university press (1999).

[3] H.L. Van Trees, "Detection, estimation, and modulation theory", John Wiley & Sons (2004).

[4] R.D. Gill and B.Y. Levit, "Applications of the van Trees inequality: a Bayesian Cramér-Rao bound", Bernoulli, pp. 59 - 79 (1995).

[5] A.A. Borovkov and A.U. Sakhanienko, "On estimates of the expected quadratic risk", Probab. Math. Statist, vol. 1, pp. 185 - 195 (1980).