1 Proof of the Cramer-Rao bound

Normalization of the probability density:

$$\int d^p \underline{\mathbf{x}}\, \underbrace{P_{\left(\{\underline{\mathbf{x}}^{(\alpha)}\};\underline{\mathbf{w}}\right)}}_{\equiv P} = 1 \tag{1}$$

$$\begin{aligned} 0 &= \tfrac{\partial}{\partial \underline{\mathbf{w}}} \int d^p \underline{\mathbf{x}} P \\ &= \int d^p \underline{\mathbf{x}} \tfrac{\partial P}{\partial \underline{\mathbf{w}}} \\ &= \int d^p P \tfrac{\partial \ln P}{\partial \underline{\mathbf{w}}} \\ &= \left\langle \tfrac{\partial \ln P}{\partial \underline{\mathbf{w}}} \right\rangle_p \end{aligned} \tag{2}$$

we then obtain

$$\begin{aligned} \left\langle \left(\hat{\mathrm{w}}_i - \mathrm{w}_i^*\right) \tfrac{\partial \ln P}{\partial \mathrm{w}_j}\Big|_{\underline{\mathbf{w}}^*} \right\rangle_p &= \left\langle \hat{\mathrm{w}}_i \overbrace{\frac{\partial \ln P}{\partial \mathrm{w}_j}}^{=\frac{1}{P}\frac{\partial P}{\partial \mathrm{w}_j}}\Big|_{\underline{\mathbf{w}}^*} \right\rangle_p - \mathrm{w}_i^* \underbrace{\left\langle \frac{\partial \ln P}{\partial \mathrm{w}_j}\Big|_{\underline{\mathbf{w}}^*} \right\rangle_p}_{=0} \\ &= \tfrac{\partial <\hat{\mathrm{w}}_i>_p}{\partial \mathrm{w}_j}\Big|_{\underline{\mathbf{w}}^*} \\ &= \tfrac{\partial \mathrm{w}_i}{\partial \mathrm{w}_j}\Big|_{\underline{\mathbf{w}}^*} \\ &\uparrow \text{estimator without bias} \\ &= \delta_{ij} \end{aligned} \tag{3}$$

let $\underline{\mathbf{a}}$ and $\underline{\mathbf{b}}$ be arbitrary vectors, then:

$$\left\langle \underline{\mathbf{a}}^T \left(\hat{\underline{\mathbf{w}}} - \underline{\mathbf{w}}^*\right) \left(\frac{\partial \ln P}{\partial \underline{\mathbf{w}}}\right)^T \underline{\mathbf{b}} \right\rangle\Big|_{\underline{\mathbf{w}}^*} = \underline{\mathbf{a}}^T \underline{\mathbf{b}} \tag{4}$$

Application of the Cauchy-Schwarz inequality:

$$\left\{ \int f_{(\underline{\mathbf{D}})} g_{(\underline{\mathbf{D}})} h_{(\underline{\mathbf{D}})} d\underline{\mathbf{D}} \right\}^2 \leq \left\{ \int f_{(\underline{\mathbf{D}})} g^2_{(\underline{\mathbf{D}})} d\underline{\mathbf{D}} \right\} \left\{ \int f_{(\underline{\mathbf{D}})} h^2_{(\underline{\mathbf{D}})} d\underline{\mathbf{D}} \right\} \tag{5}$$

with:

$$\begin{aligned} \underline{\mathbf{D}} &= \left\{\underline{\mathbf{x}}^{(\alpha)}\right\} \\ f_{(\underline{\mathbf{D}})} &= P_{\left(\{\underline{\mathbf{x}}^{(\alpha)}\},\underline{\mathbf{w}}\right)} \\ g_{(\underline{\mathbf{D}})} &= \underline{\mathbf{a}}^T \left(\hat{\underline{\mathbf{w}}} - \underline{\mathbf{w}}^*\right) \\ h_{(\underline{\mathbf{D}})} &= \left(\frac{\partial \ln P_{\left(\{\underline{\mathbf{x}}^{(\alpha)}\}\right)}}{\partial \underline{\mathbf{w}}} \right)^T \underline{\mathbf{b}} \end{aligned}$$

yields:

$$\left(\underline{\mathbf{a}}^T\underline{\mathbf{b}}\right)^2 \leq \underline{\mathbf{a}}^T\left\langle\left(\underline{\hat{\mathbf{w}}} - \underline{\mathbf{w}}^*\right)\left(\underline{\hat{\mathbf{w}}} - \underline{\mathbf{w}}^*\right)^T\right\rangle_p \underline{\mathbf{a}}\underline{\mathbf{b}}^T\left\langle\frac{\partial \ln P}{\partial \underline{\mathbf{w}}}\left(\frac{\partial \ln P}{\partial \underline{\mathbf{w}}}\right)^T\right\rangle_p\Bigg|_{\underline{\mathbf{w}}^*}\underline{\mathbf{b}} \tag{6}$$

Using:

$$\begin{aligned}
\left\langle \tfrac{\partial^2 \ln P}{\partial \mathrm{w}_i \partial \mathrm{w}_j}\right\rangle_p &= \left\langle \tfrac{\partial}{\partial \mathrm{w}_i}\left(\tfrac{1}{P}\tfrac{\partial P}{\partial \mathrm{w}_i}\right)\right\rangle_p \\
&= \left\langle\left\{ -\tfrac{1}{P^2}\tfrac{\partial P}{\partial \mathrm{w}_i}\tfrac{\partial P}{\partial \mathrm{w}_j} + \tfrac{1}{P}\tfrac{\partial}{\partial \mathrm{w}_i}\tfrac{\partial P}{\partial \mathrm{w}_j}\right\}\right\rangle_p \\
&= -\left\langle \tfrac{\partial \ln P}{\partial \mathrm{w}_i}\tfrac{\partial \ln P}{\partial \mathrm{w}_j}\right\rangle_p + \left\langle \tfrac{1}{P}\tfrac{\partial}{\partial \mathrm{w}_i}\left(P\tfrac{\partial \ln P}{\partial \mathrm{w}_j}\right)\right\rangle_p \qquad (7)\\
&= -\left\langle \tfrac{\partial \ln P}{\partial \mathrm{w}_i}\tfrac{\partial \ln P}{\partial \mathrm{w}_j}\right\rangle_p + \tfrac{\partial}{\partial \mathrm{w}_i}\underbrace{\left\langle \frac{\partial \ln P}{\partial \mathrm{w}_j}\right\rangle_p}_{=0}
\end{aligned}$$

we obtain:

$$\left(\underline{\mathbf{a}}^T\underline{\mathbf{b}}\right)^2 \leq \left(\underline{\mathbf{a}}^T\underline{\sum}\underline{\mathbf{a}}\right)\left(\underline{\mathbf{b}}^T\underline{\mathbf{M}}\underline{\mathbf{b}}\right) \tag{8}$$

let $\underline{\mathbf{b}} = \underline{\mathbf{M}}^{-1}\underline{\mathbf{a}}$ (ok, because $\underline{\mathbf{b}}$ can be an arbitrary vector), then:

$$\left(\underline{\mathbf{a}}^T\underline{\mathbf{M}}^{-1}\underline{\mathbf{a}}\right)^2 \leq \left(\underline{\mathbf{a}}^T\underline{\sum}\underline{\mathbf{a}}\right)\left(\underline{\mathbf{b}}^T\underline{\mathbf{M}}^{-1}\underline{\mathbf{b}}\right) \tag{9}$$

$$\left(\underline{\mathbf{a}}^T\underline{\mathbf{M}}^{-1}\underline{\mathbf{a}}\right) \leq \left(\underline{\mathbf{a}}^T\underline{\sum}\underline{\mathbf{a}}\right) \text{ for all vectors } \underline{\mathbf{a}} \tag{10}$$

Consequence:

$$\underline{\sum} - \underline{\mathbf{M}}^{-1} \text{ is positive semidefinite} \tag{11}$$