Unmixing of Hyperspectral Images using Bayesian Nonnegative Matrix Factorization with Volume Prior

TECHNICAL NOTE v1.0

Morten Arngren
Technical University of Denmark, DTU Informatics, Bldg. 321, Richard Petersens Plads, DK-2800 Lyngby
and FOSS Analytical A/S, Slangerupgade 69, DK-3400 Hillerød
ma@imm.dtu.dk,moa@foss.dk, info@arngren.com

Mikkel N. Schmidt
Technical University of Denmark, DTU Informatics, Bldg. 321, Richard Petersens Plads, DK-2800 Lyngby
mns@imm.dtu.dk

Jan Larsen
Technical University of Denmark, DTU Informatics, Bldg. 321, Richard Petersens Plads, DK-2800 Lyngby
jl@imm.dtu.dk

Version History

Version	Date	Change Revision
1.0	May 2010	Initial issue.

Contents

1 Introduction

This paper acts as an appendix and supplement to the Journal paper entitled 'Unmixing of Hyperspectral Images using Bayesian Nonnegative Matrix Factorization with Volume Prior' [1]. It provides detailed derivation as a supplement to the Bayesian unmixing model.

If you should have any comments or corrections, please forward any inquiry directly to Morten Arngren (ma@imm.dtu.dk or info@arngren.com).

2 Bayesian NMF with volume prior

This section present the detailed derivations of the different probability distributions used in the Gibbs sampling scheme [1].

2.1 Posterior

The joint posterior distribution of the NMF parameters, $\boldsymbol{W}$ and $\boldsymbol{H}$, can be expressed using Bayes' rule

$$p(\boldsymbol{W}, \boldsymbol{H}|\boldsymbol{X}, \mathcal{H}) = \frac{p(\boldsymbol{X}|\boldsymbol{W}, \boldsymbol{H}, \sigma^2) p(\boldsymbol{H}) p(\boldsymbol{W}|\gamma) p(\sigma^2|\alpha, \beta)}{p(\boldsymbol{X})}, \tag{1}$$

where $\mathcal{H} = \{\alpha, \beta, \gamma\}$ are hyperparameters and $\mathcal{P} = \{\boldsymbol{W}, \boldsymbol{H}, \sigma^2\}$ denote the parameters of the model.

2.2 Sampling the noise variance

The noise is assumed independent and identically distributed white Gaussian noise, giving rise to the following likelihood function,

$$p(\boldsymbol{X}|\boldsymbol{W}, \boldsymbol{H}, \sigma^2) = \prod_{n=1}^{N} \prod_{m=1}^{M} \mathcal{N}\left(x_{mn}|\boldsymbol{w}_{m:}\boldsymbol{h}_{:n}, \sigma^2\right) = \prod_{n=1}^{N} \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2}{-2\sigma^2}\right). \tag{2}$$

The noise variance σ^2 is modeled by the inverse gamma probability distribution given by

$$p(\sigma^2|\alpha, \beta) = \mathcal{IG}(\sigma^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\sigma^2}\right), \tag{3}$$

where α and β are the parameters defining the distribution. The conditional probability of the noise variance can be expressed as

$$p(\sigma^2|\boldsymbol{X}, \mathcal{P}_{\setminus\sigma^2}, \mathcal{H}) = \frac{p(\sigma^2, \boldsymbol{X}|\mathcal{P}_{\setminus\sigma^2}, \mathcal{H})}{p(\boldsymbol{X}|\mathcal{P}_{\setminus\sigma^2})} = \frac{p(\boldsymbol{X}|\mathcal{P}_{\setminus\sigma^2})p(\sigma^2|\mathcal{H})}{\int p(\sigma^2, \boldsymbol{X}|\mathcal{P}_{\setminus\sigma^2}, \mathcal{H})\, \mathsf{d}\sigma^2} \tag{4}$$

$$\propto p(\boldsymbol{X}|\mathcal{P}_{\setminus\sigma^2})p(\sigma^2|\mathcal{H}) \tag{5}$$

$$= \prod_{n=1}^{N}\prod_{m=1}^{M}\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2}{-2\sigma^2}\right)\right]\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left(-\frac{\beta}{\sigma^2}\right). \tag{6}$$

The conjugate prior in Eq. (3) gives rise to a new set of parameters $\bar{\alpha}$ and $\bar{\beta}$ for the inverse gamma probability distribution. In estimating the new parameter set, constant terms are not considered leading to a reduction of Eq. (6).

$$p(\sigma^2|\boldsymbol{X}, \mathcal{P}_{\setminus\sigma^2}, \mathcal{H}) \propto \prod_{n=1}^{N}\prod_{m=1}^{M}\left[\left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}}\right]\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2}{-2\sigma^2}\right)\exp\left(-\frac{\beta}{\sigma^2}\right) \tag{7}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}NM}\left(\frac{1}{\sigma^2}\right)^{\alpha+1}\exp\left(\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2}{-2\sigma^2} - \frac{\beta}{\sigma^2}\right) \tag{8}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\alpha+1+\frac{1}{2}NM}\exp\left(-\frac{1}{\sigma^2}\left[\beta + \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{M}(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2\right]\right) \tag{9}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma^2}\right)^{\bar{\alpha}+1}\exp\left(-\frac{\bar{\beta}}{\sigma^2}\right). \tag{10}$$

The new set of hyperparameters are now given by

$$\bar{\alpha} = \alpha + \tfrac{1}{2}NM, \tag{11}$$

$$\bar{\beta} = \beta + \tfrac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{N}(x_{mn} - \boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2. \tag{12}$$

The inverse gamma distribution then becomes

$$p(\sigma^2|\boldsymbol{X}, \mathcal{P}_{\setminus\sigma^2}) = \mathcal{IG}(\sigma^2|\bar{\alpha}, \bar{\beta}), \tag{13}$$

2.3 Sampling the fractional abundances

The prior distribution of the fractional abundances is modeled as a uniform prior on the unit simplex expressed as

$$p(\boldsymbol{H}) \propto \prod_{n=1}^{N}\mathbb{I}\left[\sum_{k=1}^{K}h_{kn} = 1\right]\prod_{k=1}^{K}\mathbb{I}\left[h_{kn} \geq 0\right], \tag{14}$$

$$p(\boldsymbol{h}_{:n}) \propto \mathbb{I}\left[\sum_{k=1}^{K} h_{kn} = 1\right] \prod_{k=1}^{K} \mathbb{I}\left[h_{kn} \geq 0\right], \tag{15}$$

The conditional probability distribution can then be calculated from the general expression as

$$\begin{aligned}
p(\boldsymbol{h}_{:n}|\boldsymbol{X}, \mathcal{P}_{\backslash \boldsymbol{h}_{:n}}) &= \frac{p(\boldsymbol{W}, \boldsymbol{h}_{:n}|\boldsymbol{X}, \sigma^2)}{p(\boldsymbol{W}|\boldsymbol{X}, \sigma^2)} = \frac{p(\boldsymbol{W}, \boldsymbol{h}_{:n}|\boldsymbol{X}, \sigma^2)}{\int p(\boldsymbol{W}, \boldsymbol{h}_{:n}|\boldsymbol{X}, \sigma^2)\ \mathsf{d}\boldsymbol{h}_{:n}} && (16)\\
&= p(\boldsymbol{W}, \boldsymbol{h}_{:n}|\boldsymbol{X}, \sigma^2) \cdot d && (17)\\
&= \frac{p(\boldsymbol{X}|\boldsymbol{W}, \boldsymbol{h}_{:n}, \sigma^2)p(\boldsymbol{h}_{:n})p(\boldsymbol{W}|\gamma)p(\sigma^2|\alpha, \beta)}{p(\boldsymbol{X})} \cdot d && (18)\\
&\propto p(\boldsymbol{X}|\boldsymbol{W}, \boldsymbol{h}_{:n}, \sigma^2)p(\boldsymbol{h}_{:n}) && (19)\\
&= \prod_{n=1}^{N} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(\boldsymbol{x}_{:n} - \boldsymbol{W}\boldsymbol{h}_{:n})^2}{-2\sigma^2}\right)\right] p(\boldsymbol{h}_{:n}), && (20)
\end{aligned}$$

where d denote any irrelevant constant terms. The argument of the exponential function can be rewritten in quadratic form if we denote the covariance matrix $\boldsymbol{\Sigma} = \sigma^2\mathcal{I}$, where $\mathcal{I}$ is the identity matrix.

$$\begin{aligned}
&(\boldsymbol{x}_{:n} - \boldsymbol{W}\boldsymbol{h}_{:n})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_{:n} - \boldsymbol{W}\boldsymbol{h}_{:n}) = \\
&\qquad \boldsymbol{x}_{:n}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n} - \boldsymbol{x}_{:n}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{h}_{:n} - \boldsymbol{h}_{:n}^\top\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n} + \boldsymbol{h}_{:n}^\top\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{h}_{:n}. && (21)
\end{aligned}$$

As Eq. (20) is also a distribution over the parameter $\boldsymbol{h}_{:n}$, it can be rewritten in quadratic form as

$$p(\boldsymbol{h}_{:n}|\boldsymbol{X}, \boldsymbol{W}, \sigma^2) = d \cdot \prod_{n=1}^{N} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left((\boldsymbol{h}_{:n} - \bar{\boldsymbol{\mu}}_n)^\top \bar{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{h}_{:n} - \bar{\boldsymbol{\mu}}_n)\right)\right] p(\boldsymbol{h}_{:n}), \tag{22}$$

where d denote the new normalization constant. The mean offset is denoted by $\bar{\boldsymbol{\mu}}_n$ and $\bar{\boldsymbol{\Sigma}}$ is the covariance matrix for the fractional abundances $\boldsymbol{h}_{:n}$. These two new parameters are calculated by expanding the quadratic form in Eq. (22)

$$(\boldsymbol{h}_{:n} - \bar{\boldsymbol{\mu}}_n)^\top \bar{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{h}_{:n} - \bar{\boldsymbol{\mu}}_n) = \boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1}\boldsymbol{h}_{:n} - \boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\mu}}_n - \bar{\boldsymbol{\mu}}_n^\top\bar{\boldsymbol{\Sigma}}^{-1}\boldsymbol{h}_{:n} + \bar{\boldsymbol{\mu}}_n^\top\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\mu}}_n \tag{23}$$

Comparing Eq. (21) and (23), the new mean offset and covariance matrix can easily be identified as

$$\boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1}\boldsymbol{h}_{:n} = \boldsymbol{h}_{:n}^\top \underbrace{\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W}}_{\bar{\boldsymbol{\Sigma}}^{-1}} \boldsymbol{h}_{:n} \qquad \Rightarrow \qquad \bar{\boldsymbol{\Sigma}} = (\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W})^{-1} = \sigma^2(\boldsymbol{W}^\top\boldsymbol{W})^{-1} \tag{24}$$

$$\begin{aligned}
\boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\mu}}_n &= \boldsymbol{h}_{:n}^\top\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n} = \boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1}\bar{\boldsymbol{\Sigma}}\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n} && (25)\\
&= \boldsymbol{h}_{:n}^\top\bar{\boldsymbol{\Sigma}}^{-1} \underbrace{(\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n}}_{\bar{\boldsymbol{\mu}}_n} && (26)
\end{aligned}$$

$$\Rightarrow \qquad \bar{\boldsymbol{\mu}}_n = (\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_{:n} = (\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{x}_{:n}, \tag{27}$$

where the last reduction step utilizes $\boldsymbol{\Sigma}^{-1}$ as a diagonal covariance matrix resulting in the pseudoinverse of $\boldsymbol{W}$. The new posterior conditional parameters are hence expressed by Eq. (24) and (27) and becomes

$$p(\boldsymbol{h}_{:n}|\boldsymbol{X},\mathcal{P}\backslash\boldsymbol{h}_{:n}) \propto \mathcal{N}(\boldsymbol{h}_{:n}|\bar{\boldsymbol{\mu}}_n,\bar{\boldsymbol{\Sigma}}_n)\prod_{n=1}^{N}\mathbb{I}\left[\sum_{k=1}^{K}h_{kn}=1\right]\prod_{k=1}^{K}\mathbb{I}\left[h_{kn}\geq 0\right], \tag{28}$$

2.4 Sampling endmembers

The prior probability of the endmembers are modeled as a constrained exponential distribution to encourage small spanning volumes of the endmembers and is expressed as

$$p(\boldsymbol{W}|\gamma) \propto \exp\big(-\gamma J_w(\boldsymbol{W})\big)\prod_{m=1}^{M}\prod_{k=1}^{K}\mathbb{I}\left[w_{mk}\geq 0\right], \tag{29}$$

where 3 different volume measures are considered given by

$$J_w^{\mathrm{pp}}(\boldsymbol{W}) = \det(\boldsymbol{W}^{\top}\boldsymbol{W}), \tag{30}$$

$$J_w^{\mathrm{sv}}(\boldsymbol{W}) = \det(\tilde{\boldsymbol{W}}^{\top}\tilde{\boldsymbol{W}}), \text{ and} \tag{31}$$

$$J_w^{\mathrm{dist}}(\boldsymbol{W}) = \sum_{k=1}^{K}\left|\left|\boldsymbol{w}_{:k}-\frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{:k'}\right|\right|_2^2. \tag{32}$$

The conditional prior distribution of a single endmember element then becomes a truncated Gaussian distribution due to the quadratic terms in any of the $J_w(\boldsymbol{W})$, i.e.

$$\begin{aligned} p(w_{mk}|\mathcal{P}_{\backslash w_{mk}}) &\propto \mathcal{N}(w_{mk}|c_{mk},s_{mk}^2)\mathbb{I}\left[w_{mk}\geq 0\right] \\ &= \frac{1}{\sqrt{2\pi}s_{mk}}\exp\left(\frac{(w_{mk}-c_{mk})^2}{-2s_{mk}^2}\right)\mathbb{I}\left[w_{mk}\geq 0\right], \end{aligned} \tag{33}$$

where c_{mk} and s_{mk}^2 denote the corresponding set of parameters for the Gaussian distribution, mean and covariance respectively. Using the likelihood written in Eq. (2) the conditional posterior distribution then becomes

$$p(w_{mk}|\boldsymbol{X},\mathcal{P}_{\backslash w_{mk}}) = \frac{p(w_{mk},\boldsymbol{H}|\boldsymbol{X},\boldsymbol{W}_{\backslash w_{mk}},\sigma^2)}{p(\boldsymbol{H}|\boldsymbol{X},\boldsymbol{W}_{\backslash w_{mk}}\sigma^2)} = \frac{p(w_{mk},\boldsymbol{H}|\boldsymbol{X},\boldsymbol{W}_{\backslash w_{mk}}\sigma^2)}{\int p(w_{mk},\boldsymbol{H}|\boldsymbol{X},\boldsymbol{W}_{\backslash w_{mk}}\sigma^2)\,\mathrm{d}w_{mk}} \tag{34}$$

$$= p(w_{mk},\boldsymbol{H}|\boldsymbol{X},\boldsymbol{W}_{\backslash w_{mk}}\sigma^2)\cdot d \tag{35}$$

$$= \frac{p(\boldsymbol{X}|\mathcal{P})p(\boldsymbol{H})p(w_{mk}|\gamma)p(\sigma^2|\alpha,\beta)}{p(\boldsymbol{X})}\cdot d \tag{36}$$

$$\propto p(\boldsymbol{X}|\mathcal{P})p(w_{mk}|\gamma) \tag{37}$$

$$= \prod_{n=1}^{N}\prod_{m=1}^{M}\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{(x_{mn}-w_{mk}h_{kn})^2}{-2\sigma^2}\right)\right] \tag{38}$$

$$\times\exp\big(-\gamma J_w(w_{mk})\big)\mathbb{I}\left[w_{mk}\geq 0\right], \tag{39}$$

where d denote any constant terms. As argued the conditional prior is a Gaussian distribution in Eq. (33) due to the quadratic terms and leads further to

$$p(w_{mk}|\boldsymbol{X},\mathcal{P}_{\backslash w_{mk}}) \propto p(\boldsymbol{X}|\mathcal{P})p(w_{mk}|\mathcal{P}_{\backslash w_{mk}}) \tag{40}$$

$$= \prod_{n=1}^{N}\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{(x_{mn}-\boldsymbol{w}_{m:}\boldsymbol{h}_{:n})^2}{-2\sigma^2}\right)\right]\frac{1}{\sqrt{2\pi}s_{mk}}\exp\left(\frac{(w_{mk}-c_{mk})^2}{-2s_{mk}^2}\right)\mathbb{I}\left[w_{mk}\geq 0\right] \tag{41}$$

$$= \frac{N}{\sqrt{2\pi}\sigma}\frac{1}{\sqrt{2\pi}s_{mk}}\exp\left[-\tfrac{1}{2}\sigma^{-2}\sum_{n=1}^{N}\big(x_{mn}-\boldsymbol{w}_{m:}\boldsymbol{h}_{:n}\big)^2-\tfrac{1}{2}s_{mk}^{-2}\big(w_{mk}-c_{mk}\big)^2\right]\mathbb{I}\left[w_{mk}\geq 0\right] \tag{42}$$

Using the argument of the exponential function in quadratic form, we get

$$-\tfrac{1}{2}\sigma^{-2}\sum_{n=1}^{N}\big(x_{mn}-\boldsymbol{w}_{m:}\boldsymbol{h}_{:n}\big)^2-\tfrac{1}{2}s_{mk}^{-2}\big(w_{mk}-c_{mk}\big)^2 \tag{43}$$

$$= -\tfrac{1}{2}\sigma^{-2}\big(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m:}\boldsymbol{H}\big)\big(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m:}\boldsymbol{H}\big)^\top-\tfrac{1}{2}s_{mk}^{-2}\big(w_{mk}-c_{mk}\big)^2 \tag{44}$$

$$= -\tfrac{1}{2}\sigma^{-2}\big(\boldsymbol{x}_{m:}-w_{mk}\boldsymbol{h}_{k:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:}\big)\big(\boldsymbol{x}_{m:}-w_{mk}\boldsymbol{h}_{k:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:}\big)^\top-\tfrac{1}{2}s_{mk}^{-2}\big(w_{mk}-c_{mk}\big)^2 \tag{45}$$

$$= -\tfrac{1}{2}\sigma^{-2}\big(w_{mk}^2\boldsymbol{h}_{k:}\boldsymbol{h}_{k:}^\top-2w_{mk}\boldsymbol{h}_{k:}(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:})^\top+d\big)-\tfrac{1}{2}s_{mk}^{-2}\big(w_{mk}-c_{mk}\big)^2 \tag{46}$$

$$\propto -\sigma^{-2}\boldsymbol{h}_{k:}\boldsymbol{h}_{k:}^\top w_{mk}^2+2\sigma^{-2}w_{mk}\boldsymbol{h}_{k:}(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:})^\top-s_{mk}^{-2}w_{mk}^2+2s_{mk}^{-2}c_{mk}w_{mk}, \tag{47}$$

where d represent the constant terms. The conditional posterior is also a Gaussian distribution and can hence be expressed in quadratic form similarly to Eq. (23)

$$-\bar{\sigma}_{mk}^{-2}(w_{mk}-\bar{\mu}_{mk})^2 = -\bar{\sigma}_{mk}^{-2}w_{mk}^2+2\bar{\sigma}_{mk}^{-2}w_{mk}\bar{\mu}_{mk}-\bar{\sigma}_{mk}^{-2}\bar{\mu}_{mk}^2. \tag{48}$$

Both parameters can now be derived

$$-\bar{\sigma}_{mk}^{-2}w_{mk}^2 = -\sigma^{-2}\boldsymbol{h}_{k:}\boldsymbol{h}_{k:}^\top w_{mk}^2-s_{mk}^{-2}w_{mk}^2 \qquad \Leftrightarrow \tag{49}$$

$$\bar{\sigma}_{mk}^{-2} = s_{mk}^{-2}+\sigma^{-2}\boldsymbol{h}_{k:}\boldsymbol{h}_{k:}^\top. \tag{50}$$

$$2\bar{\sigma}_{mk}^{-2}w_{mk}\bar{\mu}_{mk} = 2\sigma^{-2}w_{mk}\boldsymbol{h}_{k:}(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:})^\top+2s_{mk}^{-2}c_{mk}w_{mk} \qquad \Leftrightarrow \tag{51}$$

$$\bar{\sigma}_{mk}^{-2}\bar{\mu}_{mk} = \sigma^{-2}\boldsymbol{h}_{k:}(\boldsymbol{x}_{m:}-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:})^\top+s_{mk}^{-2}c_{mk} \qquad \Leftrightarrow \tag{52}$$

$$\bar{\mu}_{mk} = \bar{\sigma}_{mk}^2\Big(c_{mk}s_{mk}^{-2}+\big(\boldsymbol{x}_{m:}\boldsymbol{h}_{k:}^\top-\boldsymbol{w}_{m\tilde{k}}\boldsymbol{h}_{\tilde{k}:}\boldsymbol{h}_{k:}^\top\big)\sigma^{-2}\Big). \tag{53}$$

The parameters of the conditional Gaussian distribution is hence given by Eq. (50) and (53). The mean c_{mk} and covariance s_{mk}^{-2} of the conditional prior can now be determined for each of the volume measures in Eq. (30)-(32). These parameter are calculated from the conditional prior distribution rewritten as

$$p(w_{mk}|\mathcal{P}_{\backslash w_{mk}}) \propto \frac{1}{\sqrt{2\pi}s_{mk}}\exp\Big(-\gamma J_w(\boldsymbol{W})\Big)\mathbb{I}\left[w_{mk}\geq 0\right]. \tag{54}$$

In order to find the equivalent mean and covariance for the determinant based volume measures the argument of the determinant must be decomposed into scalar values using the following expression for the determinant of a symmetric matrix,

$$\det\begin{bmatrix} a & \boldsymbol{b}^\top \\ \boldsymbol{b} & \boldsymbol{C}\end{bmatrix} = a\det\left(\boldsymbol{C} - \tfrac{1}{a}\boldsymbol{b}\boldsymbol{b}^\top\right) \tag{55}$$

$$= a\det(\boldsymbol{C})\left(1 - \tfrac{1}{a}\boldsymbol{b}^\top\boldsymbol{C}^{-1}\boldsymbol{b}\right) \tag{56}$$

$$= a\det(\boldsymbol{C}) - \boldsymbol{b}^\top\mathrm{adj}(\boldsymbol{C})\boldsymbol{b}, \tag{57}$$

where Eq. (55) is the expression for the determinant of a block matrix, Eq. (56) follows from the matrix determinant lemma, and in Eq. (57) we have used the definition of the matrix adjugate.

2.4.1 Parellelepiped Volume Measure

The parallelepipped volume given in Eq. (30) is decomposed using Eq. (57) to isolate w_{mk} for the conditional prior distribution.

$$\det(\boldsymbol{W}^\top\boldsymbol{W}) = \det\left(\begin{bmatrix} w_{mk} & \boldsymbol{w}_{m\tilde{k}} \\ \boldsymbol{w}_{\tilde{m}k} & \boldsymbol{W}_{\tilde{m}\tilde{k}}\end{bmatrix}^\top \begin{bmatrix} w_{mk} & \boldsymbol{w}_{m\tilde{k}} \\ \boldsymbol{w}_{\tilde{m}k} & \boldsymbol{W}_{\tilde{m}\tilde{k}}\end{bmatrix}\right) \tag{58}$$

$$= \det\left(\begin{bmatrix} w_{mk} & \boldsymbol{w}_{\tilde{m}k}^\top \\ \boldsymbol{w}_{m\tilde{k}}^\top & \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top\end{bmatrix} \begin{bmatrix} w_{mk} & \boldsymbol{w}_{m\tilde{k}} \\ \boldsymbol{w}_{\tilde{m}k} & \boldsymbol{W}_{\tilde{m}\tilde{k}}\end{bmatrix}\right) \tag{59}$$

$$= \det\left(\begin{bmatrix} w_{mk}^2 + \boldsymbol{w}_{\tilde{m}k}^\top\boldsymbol{w}_{\tilde{m}k} & w_{mk}\boldsymbol{w}_{m\tilde{k}} + \boldsymbol{w}_{\tilde{m}k}^\top\boldsymbol{W}_{\tilde{m}\tilde{k}} \\ \boldsymbol{w}_{m\tilde{k}}^\top w_{mk} + \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top\boldsymbol{w}_{\tilde{m}k} & \boldsymbol{w}_{m\tilde{k}}^\top\boldsymbol{w}_{m\tilde{k}} + \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top\boldsymbol{W}_{\tilde{m}\tilde{k}}\end{bmatrix}\right) \tag{60}$$

$$= \det\left(\begin{bmatrix} w_{mk}^2 + \boldsymbol{w}_{\tilde{m}k}^\top\boldsymbol{w}_{\tilde{m}k} & w_{mk}\boldsymbol{w}_{m\tilde{k}} + \boldsymbol{w}_{\tilde{m}k}^\top\boldsymbol{W}_{\tilde{m}\tilde{k}} \\ \boldsymbol{w}_{m\tilde{k}}^\top w_{mk} + \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top\boldsymbol{w}_{\tilde{m}k} & \boldsymbol{W}_{m\tilde{k}}^\top\boldsymbol{W}_{m\tilde{k}}\end{bmatrix}\right) \tag{61}$$

This leads to the specific definition of the parameters a, $\boldsymbol{b}$ and $\boldsymbol{C}$ to be

$$a = w_{mk}^2 + \boldsymbol{w}_{\tilde{m}k}^\top\boldsymbol{w}_{\tilde{m}k} \tag{62}$$

$$\boldsymbol{b} = \boldsymbol{w}_{m\tilde{k}}^\top w_{mk} + \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top\boldsymbol{w}_{\tilde{m}k} \tag{63}$$

$$\boldsymbol{C} = \boldsymbol{W}_{m\tilde{k}}^\top\boldsymbol{W}_{m\tilde{k}}. \tag{64}$$

Using Eq. (57) we get

$$\gamma \det(\boldsymbol{W}^\top \boldsymbol{W}) = \gamma \left(w_{mk}^2 + \boldsymbol{w}_{\tilde{m}k}^\top \boldsymbol{w}_{\tilde{m}k} \right) \det(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) - \gamma \left(w_{mk} \boldsymbol{w}_{m\tilde{k}} + \boldsymbol{w}_{\tilde{m}k}^\top \boldsymbol{W}_{\tilde{m}\tilde{k}} \right) \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \left(\boldsymbol{w}_{m\tilde{k}}^\top w_{mk} + \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k} \right) \tag{65}$$

$$= w_{mk}^2 \gamma \det(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) - \gamma w_{mk} \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{w}_{m\tilde{k}}^\top w_{mk} - 2\gamma w_{mk} \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k} + d \tag{66}$$

$$= w_{mk}^2 \left(\gamma \det(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) - \gamma \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{w}_{m\tilde{k}}^\top \right) - 2\gamma w_{mk} \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k} + d, \tag{67}$$

where d denote the irrelevant constant terms not including the variable w_{mk}. The quadratic form of the argument of the exponential function in Eq. (33) is expanded to find the new parameters

$$s_{mk}^{-2} (w_{mk} - c_{mk})^2 = s_{mk}^{-2} w_{mk}^2 - 2 s_{mk}^{-2} w_{mk} c_{mk} + s_{mk}^{-2} c_{mk}^2 \tag{68}$$

The new mean and covariance can now easily be identified

$$s_{mk}^{-2} w_{mk}^2 = w_{mk}^2 \left(\gamma \det(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) - \gamma \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{w}_{m\tilde{k}}^\top \right) \quad \Leftrightarrow \tag{69}$$

$$s_{mk}^{-2} = \gamma \left(\det(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) - \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{w}_{m\tilde{k}}^\top \right). \tag{70}$$

$$-2 s_{mk}^{-2} w_{mk} c_{mk} = -2\gamma w_{mk} \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k} \quad \Leftrightarrow \tag{71}$$

$$c_{mk} = \gamma s_{mk}^2 \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k}. \tag{72}$$

2.4.2 Simplex Volume Measure

The simplex volume expressed in Eq. (31) is also decomposed using Eq. (57) to isolate w_{mk} for the conditional prior distribution.

$$\det(\tilde{\boldsymbol{W}}^\top \tilde{\boldsymbol{W}}) = \det \left(\begin{bmatrix} w_{mk} - w_{m\rho} & \boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1} \\ \boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho} & \boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1} \end{bmatrix}^\top \begin{bmatrix} w_{mk} - w_{m\rho} & \boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1} \\ \boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho} & \boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1} \end{bmatrix} \right) \tag{73}$$

$$= \det \left(\begin{bmatrix} w_{mk} - w_{m\rho} & (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho})^\top \\ (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1})^\top & (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1})^\top \end{bmatrix} \begin{bmatrix} w_{mk} - w_{m\rho} & \boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1} \\ \boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho} & \boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1} \end{bmatrix} \right) \tag{74}$$

$$= \det \left(\begin{bmatrix} \begin{matrix} (w_{mk} - w_{m\rho})^2 \\ +(\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho})^\top (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho}) \end{matrix} & \begin{matrix} (w_{mk} - w_{m\rho})(\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1}) \\ +(\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho})^\top (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1}) \end{matrix} \\ \begin{matrix} (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1})^\top (w_{mk} - w_{m\rho}) \\ +(\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1})^\top (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho}) \end{matrix} & \begin{matrix} (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1})^\top (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1}) \\ +(\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1})^\top (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1}) \end{matrix} \end{bmatrix} \right). \tag{75}$$

This leads to the specific definition of the parameters a, $\boldsymbol{b}$ and $\boldsymbol{C}$ to be

$$
\begin{aligned}
a &= (w_{mk} - w_{m\rho})^2 + (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho})^\top (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho}) && (76)\\
&= (w_{mk} - w_{m\rho})^2 + \tilde{\boldsymbol{w}}_{\tilde{m}k}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} && (77)\\
\boldsymbol{b} &= (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1})^\top (w_{mk} - w_{m\rho}) + (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1})^\top (\boldsymbol{w}_{\tilde{m}k} - \boldsymbol{w}_{\tilde{m}\rho}) && (78)\\
&= \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top (w_{mk} - w_{m\rho}) + \tilde{\boldsymbol{W}}_{\tilde{m}\tilde{k}}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} && (79)\\
\boldsymbol{b}^\top &= (w_{mk} - w_{m\rho})\tilde{\boldsymbol{w}}_{m\tilde{k}} + \tilde{\boldsymbol{w}}_{\tilde{m}k}^\top \tilde{\boldsymbol{W}}_{\tilde{m}\tilde{k}} && (80)\\
\boldsymbol{C} &= (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1})^\top (\boldsymbol{w}_{m\tilde{k}} - w_{m\rho}\mathbf{1}) + (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1})^\top (\boldsymbol{W}_{\tilde{m}\tilde{k}} - \boldsymbol{w}_{\tilde{m}\rho}\mathbf{1}) && (81)\\
&= (\boldsymbol{W}_{m\tilde{k}} - \boldsymbol{w}_{:\rho}\mathbf{1})^\top (\boldsymbol{W}_{m\tilde{k}} - \boldsymbol{w}_{:\rho}\mathbf{1}) && (82)\\
&= \tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}} && (83)
\end{aligned}
$$

Applying Eq. (57) we get

$$
\begin{aligned}
\gamma \det(\tilde{\boldsymbol{W}}^\top \tilde{\boldsymbol{W}}) &= \gamma \left((w_{mk} - w_{m\rho})^2 + \tilde{\boldsymbol{w}}_{\tilde{m}k}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} \right) \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}})\\
&- \gamma \left((w_{mk} - w_{m\rho})\tilde{\boldsymbol{w}}_{m\tilde{k}} + \tilde{\boldsymbol{w}}_{\tilde{m}k}^\top \tilde{\boldsymbol{W}}_{\tilde{m}\tilde{k}} \right) \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \left(\tilde{\boldsymbol{w}}_{m\tilde{k}}^\top (w_{mk} - w_{m\rho}) + \tilde{\boldsymbol{W}}_{\tilde{m}\tilde{k}}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} \right) && (84)\\
&= (w_{mk} - w_{m\rho})^2 \gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma (w_{mk} - w_{m\rho}) \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top (w_{mk} - w_{m\rho})\\
&- 2\gamma (w_{mk} - w_{m\rho}) \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} + d_1 && (85)\\
&= w_{mk}^2 \gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - 2 w_{mk} w_{m\rho} \gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}})\\
&- \gamma w_{mk}^2 \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top + 2\gamma w_{mk} w_{m\rho} \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top\\
&- 2\gamma w_{mk} \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \tilde{\boldsymbol{w}}_{\tilde{m}k} + d_2 && (86)\\
&= w_{mk}^2 \left(\gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top \right)\\
&- w_{mk} \Big(2 w_{m\rho} \gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - 2\gamma w_{m\rho} \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top\\
&+ 2\gamma \boldsymbol{w}_{m\tilde{k}} \operatorname{adj}(\boldsymbol{W}_{m\tilde{k}}^\top \boldsymbol{W}_{m\tilde{k}}) \boldsymbol{W}_{\tilde{m}\tilde{k}}^\top \boldsymbol{w}_{\tilde{m}k} \Big) + d_2, && (87)
\end{aligned}
$$

where $\{d_i\}_{i=1}^2$ denote the irrelevant constant terms not including the variable w_{mk}. The new mean and covariance can now easily be derived using the quadratic form in Eq. (68).

$$
\begin{aligned}
s_{mk}^{-2} w_{mk}^2 &= w_{mk}^2 \left(\gamma \det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top \right) \qquad \Leftrightarrow\\
s_{mk}^{-2} &= \gamma \left(\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) - \tilde{\boldsymbol{w}}_{m\tilde{k}} \operatorname{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^\top \tilde{\boldsymbol{W}}_{m\tilde{k}}) \tilde{\boldsymbol{w}}_{m\tilde{k}}^\top \right), && (88)
\end{aligned}
$$

$$-2s_{mk}^{-2}w_{mk}c_{mk} = -w_{mk}\Big(2w_{m\rho}\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - 2\gamma w_{m\rho}\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top} + 2\gamma\boldsymbol{w}_{m\tilde{k}}\mathrm{adj}(\boldsymbol{W}_{m\tilde{k}}^{\top}\boldsymbol{W}_{m\tilde{k}})\boldsymbol{W}_{\tilde{m}\tilde{k}}^{\top}\boldsymbol{w}_{\tilde{m}k}\Big) \quad \Leftrightarrow \tag{89}$$

$$c_{mk} = s_{mk}^{2}\Big(w_{m\rho}\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma w_{m\rho}\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top} + \gamma\boldsymbol{w}_{m\tilde{k}}\mathrm{adj}(\boldsymbol{W}_{m\tilde{k}}^{\top}\boldsymbol{W}_{m\tilde{k}})\boldsymbol{W}_{\tilde{m}\tilde{k}}^{\top}\boldsymbol{w}_{\tilde{m}k}\Big) \quad \Leftrightarrow \tag{90}$$

$$c_{mk} = \frac{w_{m\rho}\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma w_{m\rho}\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top}}{\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top}} \cdot s_{mk}^{2}\gamma\boldsymbol{w}_{m\tilde{k}}\mathrm{adj}(\boldsymbol{W}_{m\tilde{k}}^{\top}\boldsymbol{W}_{m\tilde{k}})\boldsymbol{W}_{\tilde{m}\tilde{k}}^{\top}\boldsymbol{w}_{\tilde{m}k} \quad \Leftrightarrow \tag{91}$$

$$c_{mk} = \frac{w_{m\rho}\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma w_{m\rho}\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top}}{\gamma\det(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}}) - \gamma\tilde{\boldsymbol{w}}_{m\tilde{k}}\mathrm{adj}(\tilde{\boldsymbol{W}}_{m\tilde{k}}^{\top}\tilde{\boldsymbol{W}}_{m\tilde{k}})\tilde{\boldsymbol{w}}_{m\tilde{k}}^{\top}} \cdot s_{mk}^{2}\gamma\boldsymbol{w}_{m\tilde{k}}\mathrm{adj}(\boldsymbol{W}_{m\tilde{k}}^{\top}\boldsymbol{W}_{m\tilde{k}})\boldsymbol{W}_{\tilde{m}\tilde{k}}^{\top}\boldsymbol{w}_{\tilde{m}k} \quad \Leftrightarrow \tag{92}$$

$$c_{mk} = w_{m\rho} + s_{mk}^{2}\gamma\boldsymbol{w}_{m\tilde{k}}\mathrm{adj}(\boldsymbol{W}_{m\tilde{k}}^{\top}\boldsymbol{W}_{m\tilde{k}})\boldsymbol{W}_{\tilde{m}\tilde{k}}^{\top}\boldsymbol{w}_{\tilde{m}k} \tag{93}$$

The derived mean and covariance shows how using the relative simplex vectors $\tilde{\boldsymbol{W}}$ leads to the intuitive result of only adding the offset $m_{m\rho}$ to the mean c_{mk}.

2.4.3 Euclidian Distance Measure

The volume measure based on euclidian distance metric can be derived by isolating w_{mk} in Eq. (32).

$$\gamma\sum_{k=1}^{K}\left|\left|\boldsymbol{w}_{:k} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{:k'}\right|\right|_2^2 = \sum_{\bar{k}=1}^{K}\left[\left(w_{m\bar{k}} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{mk'}\right)^2 + \sum_{\tilde{m}\neq m}\left(\boldsymbol{w}_{\tilde{m}\bar{k}} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{\tilde{m}k'}\right)^2\right] \tag{94}$$

$$= \sum_{\bar{k}=1}^{K}\left(w_{m\bar{k}} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{mk'}\right)^2 + d_1 \tag{95}$$

$$= \left(w_{mk} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{mk'}\right)^2 + \sum_{\bar{k}\neq k}\left(w_{m\bar{k}} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{mk'}\right)^2 + d_1 \tag{96}$$

$$= \left(w_{mk} - \frac{1}{K}\sum_{k'=1}^{K}\boldsymbol{w}_{mk'}\right)^2 + d_2 \tag{97}$$

$$= \left(w_{mk} - \frac{1}{K}w_{mk} - \frac{1}{K}\sum_{k'\neq k}\boldsymbol{w}_{mk'}\right)^2 + d_2 \tag{98}$$

$$= \left(\frac{K-1}{K}w_{mk} - \frac{1}{K}\sum_{k'\neq k}\boldsymbol{w}_{mk'}\right)^2 + d_2 \tag{99}$$

$$= \left(\frac{K-1}{K}\right)^2\left(w_{mk} - \frac{1}{K-1}\sum_{k'\neq k}\boldsymbol{w}_{mk'}\right)^2 + d_2 \tag{100}$$

where $\{d_i\}_{i=1}^2$ denote the irrelevant constant terms not including the variable w_{mk}. The mean and covariance can now easily be derived as

$$s_{mk}^{-2} = \gamma \left(\frac{K-1}{K} \right)^2, \tag{101}$$

$$c_{mk} = \frac{1}{K-1} \sum_{k' \neq k} w_{mk'}. \tag{102}$$

3 Comments

If you should have any comments or corrections, please forward any inquiry directly to Morten Arngren (ma@imm.dtu.dk or info@arngren.com).

References

1. Arngren, M., Schmidt, M.N., Larsen, J.: Unmixing of hyperspectral images using bayesian nonnegative matrix factorization with volume prior. Journal of Signal Processing Systems (2010)