Julien Fer
10649441
julien.fer88@gmail.com

1 MAP solution with correlated responses

1.1 Question 1

a) *Write down the likelihood $p(\boldsymbol{D} \mid \boldsymbol{\theta})$ in vector/matrix form, i.e. in terms of $\boldsymbol{t}, \boldsymbol{\Psi}, \boldsymbol{w}$ and $\boldsymbol{\Omega}$. Note that the distribution can not be factored into independent multiplicands in this basis.*

$$\begin{aligned} p(\mathbf{D} \mid \boldsymbol{\theta}) &= p(\mathbf{t} \mid \boldsymbol{\Psi}, \mathbf{w}, \boldsymbol{\Omega}) \\ &= \mathcal{N}(\boldsymbol{\Psi}\mathbf{w}, \boldsymbol{\Omega}) \\ &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Omega})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})^T \boldsymbol{\Omega}^{-1} (\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})\right]} \end{aligned} \tag{1}$$

b) *Write the likelihood in terms of a Gaussian distribution with a diagonal covariance matrix by changing the basis of the space in which the targets are expressed. Specifically, express the covariance matrix in its eigenbasis, i.e. write it as $\boldsymbol{\Omega} = \boldsymbol{A}^T \boldsymbol{D} \boldsymbol{A}$ with $\boldsymbol{D} := diag(d_1, \ldots, d_N)$ being a diagonal matrix containing the eigenvalues of $\boldsymbol{\Omega}$ and $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$ being an orthogonal change of basis. This is possible in general since covariance matrices are symmetric.*

Let $\boldsymbol{\Omega} := \mathbf{A}^T\mathbf{D}\mathbf{A}$. From (1) we see that it is necessary to determine the determinant and the inverse of the 'new' defined matrix $\boldsymbol{\Omega}$:

$$\begin{aligned} \det(\boldsymbol{\Omega}) &= \det(\mathbf{A}^T\mathbf{D}\mathbf{A}) \\ &= \det(\mathbf{A}^{-1}\mathbf{D}\mathbf{A}) \quad \text{(as } \mathbf{A} \text{ is orthogonal)} \\ &= \det(\mathbf{A}^{-1})\det(\mathbf{D})\det(\mathbf{A}) = \det(\mathbf{A})^{-1}\det(\mathbf{D})\det(\mathbf{A}) \\ &= \det(\mathbf{D}) \end{aligned} \tag{2}$$

$$\begin{aligned} \boldsymbol{\Omega}^{-1} &= (\mathbf{A}^T\mathbf{D}\mathbf{A})^{-1} = (\mathbf{A}^{-1})(\mathbf{D}^{-1})(\mathbf{A}^T)^{-1} \\ &= \mathbf{A}^T\mathbf{D}^{-1}\mathbf{A} \quad \text{(as } \mathbf{A} \text{ is orthogonal)} \end{aligned} \tag{3}$$

Substituting (2) and (3) in (1) gives the following likelihood:

$$\begin{aligned} p(\mathbf{D} \mid \boldsymbol{\theta}) &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Omega})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})^T \boldsymbol{\Omega}^{-1} (\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})\right]} \\ &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{D})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})^T \mathbf{A}^T\mathbf{D}^{-1}\mathbf{A} (\mathbf{t}-\boldsymbol{\Psi}\mathbf{w})\right]} \\ &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{D})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\mathbf{t}^T\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\mathbf{t} - \mathbf{t}^T\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\boldsymbol{\Psi}\mathbf{w} - \mathbf{w}^T\boldsymbol{\Psi}^T\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\mathbf{t} + \mathbf{w}^T\boldsymbol{\Psi}^T\mathbf{A}^T\mathbf{D}^{-1}\mathbf{A}\boldsymbol{\Psi}\mathbf{w})\right]} \\ &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{D})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\boldsymbol{\tau}^T\mathbf{D}^{-1}\boldsymbol{\tau} - \boldsymbol{\tau}^T\mathbf{D}^{-1}\boldsymbol{\Phi}\mathbf{w} - \mathbf{w}^T\boldsymbol{\Phi}^T\mathbf{D}^{-1}\boldsymbol{\tau} + \mathbf{w}^T\boldsymbol{\Phi}^T\mathbf{D}^{-1}\boldsymbol{\Phi}\mathbf{w})\right]} \quad \text{(with } \boldsymbol{\tau} := \mathbf{A}\mathbf{t} \text{ and } \boldsymbol{\Phi} := \mathbf{A}\ \boldsymbol{\Psi}\text{)} \\ &= \frac{1}{(2\pi)^{\frac{N}{2}} \det(\mathbf{D})^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}\left[(\boldsymbol{\tau}^T\mathbf{D}^{-1}\boldsymbol{\tau} - 2\boldsymbol{\tau}^T\mathbf{D}^{-1}\boldsymbol{\Phi}\mathbf{w} + \mathbf{w}^T\boldsymbol{\Phi}^T\mathbf{D}^{-1}\boldsymbol{\Phi}\mathbf{w})\right]} \\ &= \end{aligned} \tag{4}$$