Appendix A: The Sherman-Morrison-Woodbury formula

The Sherman-Morrison-Woodbury formula is

$$(\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} = \mathbf{P}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1} \tag{3.57}$$

There are many ways to prove this, all involving matrix multiplications. Here is one possible proof. Expand the left hand side:

$$\begin{aligned}
(\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} &= (\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}[\mathbf{R}(\mathbf{H}^{\mathrm{T}})^{-1}]^{-1} \\
&= [(\mathbf{R}(\mathbf{H}^{\mathrm{T}})^{-1})(\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})]^{-1} \\
&= [(\mathbf{R}(\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{P}^{-1} + \mathbf{H}]^{-1} \\
&= [(\mathbf{R}(\mathbf{H}^{\mathrm{T}})^{-1} + \mathbf{H}\mathbf{P})\mathbf{P}^{-1}]^{-1} \\
&= [(\mathbf{R} + \mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}})(\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{P}^{-1}]^{-1} \\
&= \mathbf{P}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}})^{-1}
\end{aligned}$$

Appendix B: Proof of some derivative formula

Verify that

$$\frac{d\mathrm{Tr}(\mathbf{AB})}{d\mathbf{A}} = \mathbf{B}^{\mathrm{T}}$$

Let $\mathbf{A}$ be $n \times r$ and $\mathbf{B}$ be $r \times n$ since $\mathbf{AB}$ is symmetric. We can write the individual matrix elements as

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}.$$

Thus we can write the trace of this matrix as

$$\mathrm{Tr}(\mathbf{AB}) = \sum_{i=1}^{n}(\mathbf{AB})_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{r} a_{ik}b_{ki}.$$

Now we can write that

$$\frac{d\mathrm{Tr}(\mathbf{AB})}{d\mathbf{A}} = \frac{d}{da_{lm}}\left[\sum_{i=1}^{n}\sum_{k=1}^{r} a_{ik}b_{ki}\right] = b_{ml} = \mathbf{B}^{\mathrm{T}}.$$

Appendix C: Eigenvalues of covariance matrices

A covariance matrix is real, symmetric and positive definite. Since it is real, symmetric, its eigenvalue decomposition may be written as

$$\mathbf{A} = \mathbf{E}\mathbf{D}\mathbf{E}^{\mathrm{T}} \tag{3.58}$$

where $\mathbf{D}$ is a diagonal matrix of eigenvalues and $\mathbf{E}$ is a unitary matrix of eigenvectors. That is, $\mathbf{E}^{\mathrm{T}} = \mathbf{E}^{-1}$. Now if $\mathbf{A}$ is positive definite, then for all vectors, $\boldsymbol{x}$, we have that

$$\boldsymbol{x}^{\mathrm{T}}\mathbf{A}\boldsymbol{x} > 0.$$

We can substitute for $\mathbf{A}$ using (3.58):

$$\boldsymbol{x}^{\mathrm{T}}\mathbf{E}\mathbf{D}\mathbf{E}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\mathbf{D}\boldsymbol{y} > 0$$

where $\boldsymbol{y} = \mathbf{E}^{\mathrm{T}}\boldsymbol{x}$. Expanding this we have that

$$\sum_{i=1}^{n}(y_i)^2\lambda_i > 0. \tag{3.59}$$

But this must be true for all $\boldsymbol{x}$ and therefore for all $\boldsymbol{y}$. The only way to ensure this is when all eigenvalues, λ_i are positive. Another way to see this is to choose a particular $\boldsymbol{y}$ since (3.59) must hold for all $\boldsymbol{y}$. Choose $\boldsymbol{y} = (0,0,\ldots,0,1,0,\ldots,0)$, i.e. the vector with all 0 elements except for the ith element (which is a 1). For this choice of $\boldsymbol{y}$, (3.59) becomes

$$\lambda_i > 0.$$

This can be repeated for all i, $1 \leq i \leq n$.

The eigenvalues of a real, symmetric, positive definite matrix are real and positive. Thus covariance matrices have real and positive eigenvalues.

An excellent reference on eigenvalue problems is Wilkinson (1965).