hw5

Jiahao Tian

2023-03-07

Question 1

- From the book we knew that:

$$\Omega = \sum_{k=1}^{r} \sigma_k^2 \Gamma_k$$

- where the variance components σ_k^2 are nonnegative and the matrices Γ_k are known covariance matrices.

$$L(\gamma) = \frac{-m}{2} ln 2\pi - \frac{1}{2} ln det\Omega - \frac{1}{2}(y - A\mu)^t \Omega^{-1}(y - A\mu)$$

- In the multivariate normal loglikelihood, $det\Omega$ denotes the determinant of Ω, and $\gamma = (\mu_1, ..., \mu_p, \sigma_1^2, ..., \sigma_r^2)^t$ denotes the parameters collected into a column vector. Because we assume $\Gamma_r = I$, Ω is nonsingular whenever $\sigma_r^2 > 0$.
- Show: $\sum_{i=1}^{s} m_i = \sum_{i=1}^{s} (Y^i - A_i\hat{\mu})^t \cdot \hat{\Omega}_i^{-1} \cdot (Y^i - A_i\hat{\mu})$.
- Proof:
- From the hint we knew that:

$$\sum_{k=1}^{r} \hat{\sigma}_k^2 \frac{d}{d\sigma_k^2} L(\hat{\gamma}) = 0$$

- plug in the loglikelihood.

$$\frac{d}{d\sigma_k^2}L(\gamma) = -\frac{1}{2}\frac{d}{d\sigma_k^2}lndet\Omega - \frac{1}{2}(y-A\mu)^t\frac{d}{d\sigma_k^2}\Omega^{-1}(y-A\mu)$$

$$From\ the\ book,\ one\ of\ the\ property\ is\ that:$$

$$\begin{aligned}\frac{d}{d\theta}lndetB &= \sum_{ij}(\frac{d}{db_{ij}}lndetB)\frac{d}{d\theta}b_{ij} \\ &= \sum_{ij}\frac{B_{ij}}{detB}\frac{d}{d\theta}b_{ij} \\ &= tr(B^{-1}\frac{d}{d\theta}B)\end{aligned}$$

$$= -\frac{1}{2}tr(\Omega^{-1}\Gamma_k) + \frac{1}{2}(y-A\mu)^t\Omega^{-1}\Gamma_k\Omega^{-1}(y-A\mu)$$

- Now consider the whole k dataset:

$$\begin{aligned}\sum_{k=1}^{r}\hat{\sigma}_k^2\frac{d}{d\sigma_k^2}L(\gamma) &= -\frac{1}{2}\sum_{k=1}^{r}\hat{\sigma}_k^2\cdot tr(\Omega^{-1}\Gamma_k) + \frac{1}{2}\sum_{k=1}^{r}\hat{\sigma}_k^2\cdot(y-A\mu)^t\Omega^{-1}\Gamma_k\Omega^{-1}(y-A\mu) \\ &= -\frac{1}{2}tr\sum_{k=1}^{r}\left(\hat{\sigma}_k^2\cdot\Omega^{-1}\Gamma_k\right) + \frac{1}{2}\sum_{k=1}^{r}\hat{\sigma}_k^2\cdot\Omega^{-1}\Gamma_k\cdot\left[(y-A\mu)^t\Omega^{-1}(y-A\mu)\right] \\ &= -\frac{1}{2}tr(I) + \frac{1}{2}\left[(y-A\mu)^t\Omega^{-1}(y-A\mu)\right] \\ &= -\frac{1}{2}m + \frac{1}{2}\left[(y-A\mu)^t\Omega^{-1}(y-A\mu)\right] \\ &= 0\end{aligned}$$

$$So\ we\ get:$$

$$m = (y-A\mu)^t\Omega^{-1}(y-A\mu)$$

- At the end, the ith of s pedigrees evaluated at the maximum likelihood estimates is:

$$\sum_{i=1}^{s}m_i = \sum_{i=1}^{s}(Y^i - A_i\hat{\mu})^t\cdot\hat{\Omega}_i^{-1}\cdot(Y^i - A_i\hat{\mu})$$

Question 8

a)

$$(cA) \otimes B = \left(c \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \right) \otimes B = \begin{pmatrix} ca_{11} & \dots & ca_{1n} \\ \vdots & & \vdots \\ ca_{m1} & \dots & ca_{mn} \end{pmatrix} \otimes B$$

$$= \begin{pmatrix} ca_{11}B & \dots & ca_{1n}B \\ \vdots & & \vdots \\ ca_{m1}B & \dots & ca_{mn}B \end{pmatrix} = c \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} = c(A \otimes B)$$

$$= \begin{pmatrix} ca_{11}B & \dots & ca_{1n}B \\ \vdots & & \vdots \\ ca_{m1}B & \dots & ca_{mn}B \end{pmatrix} = \begin{pmatrix} a_{11}cB & \dots & a_{1n}cB \\ \vdots & & \vdots \\ a_{m1}cB & \dots & a_{mn}cB \end{pmatrix} = A \otimes (cB)$$

b)

$$(A \otimes B)^T = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}^T = \begin{pmatrix} a_{11}B^T & \dots & a_{1n}B^T \\ \vdots & & \vdots \\ a_{m1}B^T & \dots & a_{mn}B^T \end{pmatrix} = A^T \otimes B^T$$

c)

$$(A+B)\otimes C = \left(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \right) \otimes C$$

$$= \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix} \otimes C$$

$$= \begin{pmatrix} (a_{11}+b_{11})C & \dots & (a_{1n}+b_{1n})C \\ \vdots & & \vdots \\ (a_{m1}+b_{m1})C & \dots & (a_{mn}+b_{mn})C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}C+b_{11}C & \dots & a_{1n}C+b_{1n}C \\ \vdots & & \vdots \\ a_{m1}C+b_{m1}C & \dots & a_{mn}C+b_{mn}C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}C & \dots & a_{1n}C \\ \vdots & & \vdots \\ a_{m1}C & \dots & a_{mn}C \end{pmatrix} + \begin{pmatrix} b_{11}C & \dots & b_{1n}C \\ \vdots & & \vdots \\ b_{m1}C & \dots & b_{mn}C \end{pmatrix}$$

$$= (A\otimes C)(B\otimes C)$$

d)

$$A\otimes(B+C) = \begin{pmatrix} a_{11}(B+C) & \dots & a_{1n}(B+C) \\ \vdots & & \vdots \\ a_{m1}(B+C) & \dots & a_{mn}(B+C) \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}B+a_{11}C & \dots & a_{1n}B+a_{1n}C \\ \vdots & & \vdots \\ a_{m1B}+a_{m1}C & \dots & a_{mn}B+a_{mn}C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} + \begin{pmatrix} a_{11}C & \dots & a_{1n}C \\ \vdots & & \vdots \\ a_{m1}C & \dots & a_{mn}C \end{pmatrix}$$

$$= (A\otimes B)(A\otimes C)$$

e)

$$(A \otimes B) \otimes C = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} \otimes C$$

$$= \begin{pmatrix} (a_{11}B) \otimes C & \dots & (a_{1n}B) \otimes C \\ \vdots & & \vdots \\ (a_{m1}B) \otimes C & \dots & (a_{mn}B) \otimes C \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}(B \otimes C) & \dots & a_{1n}(B \otimes C) \\ \vdots & & \vdots \\ a_{m1}(B \otimes C) & \dots & a_{mn}(B \otimes C) \end{pmatrix}$$

$$= A \otimes (B \otimes C)$$

f)

- Let $A \in M_{m,n}$ $B \in M_{p,q}$ $C \in M_{n,m}$ *and* $D \in M_{q,r}$ Then,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

- Proof:

$$(A \otimes B)(C \otimes D) = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} \begin{pmatrix} c_{11}D & \dots & c_{1m}D \\ \vdots & & \vdots \\ c_{n1}D & \dots & c_{nm}D \end{pmatrix}$$

$$= \begin{pmatrix} (\sum_{n=1}^{N} a_{1n}c_{n1})BD & \dots & (\sum_{n=1}^{N} a_{1n}c_{nm})BD \\ \vdots & & \vdots \\ (\sum_{n=1}^{N} a_{mn}c_{n1})BD & \dots & (\sum_{n=1}^{N} a_{mn}c_{nm})BD \end{pmatrix}$$

$$= \begin{pmatrix} (ac)_{11}BD & \dots & (ac)_{1m}BD \\ \vdots & & \vdots \\ (ac)_{m1}BD & \dots & (ac)_{mm}BD \end{pmatrix}$$

$$= (AC) \otimes (BD)$$

- we have used the fact that the multiplication of two block matrices can be carried out as if their blocks were scalars; and we also have used the definition of matrix multiplication to deduce that $(ac)_{mm} = \sum_{n=1}^{N} a_{mn}c_{nm}$).

g)

- If $A \in M_m$ $B \in M_n$ are nonsingular, then $A \otimes B$ is also nonsingular with:

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

- Proof: The following results follows part f.

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1}) \otimes (BB^{-1}) = I \otimes I = I$$
$$(A^{-1} \otimes B^{-1})(A \otimes B) = (A^{-1}A) \otimes (B^{-1}B) = I \otimes I = I$$

- This implies that $A^{-1} \otimes B^{-1}$ is the unique inverse of $A \otimes B$ under conventional matrix multiplication. Therefore, $A \otimes B$ is nonsingular.

h)

- Proof: The following results follows part f.
- consider $Ar = \lambda r \ and \ Bs = \mu s \ for \ r, s \neq 0$.

$$(A \otimes B)(r \otimes s) = (Ar) \otimes (Bs) = (\lambda r) \otimes (\mu s) = \lambda\mu(r \otimes s)$$

- Above shows that λ, μ is an eigenvalue of $A \otimes B$ with corresponding algebraic multiplicity r, s.
- By the triangularization theorem, consider $U \in M_n$ and $V \in M_m$ such that $U^{-1}AU = \Delta_A$ and $V^{-1}BV = \Delta_B$ where Δ_A and Δ_B are upper triangular matrices. By part f:

$$(U \otimes V)^{-1}(A \otimes B)(U \otimes V) = (U^{-1}AU) \otimes (V^{-1}BV) = \Delta_A \otimes \Delta_B$$

- From above, it follows that $\Delta_A \otimes \Delta_B$ is an upper triangular matrix that is similar to $A \otimes B$. The eigenValues of A, B, and $A \otimes B$ are the main diagonal entries of the upper triangular matrices to which they are similar ($\Delta_A, \Delta_B \ and \ \Delta_A \otimes \Delta_B$). Since $\Delta_A \ and \ \Delta_B$ are square matrices, it follows from the definition of the Kronecker product that the entries of the main diagonal of $\Delta_A \otimes \Delta_B$ are the pairwise products of the entries on the main diagonals of $\Delta_A \ and \ \Delta_B$ Therefore, the eigenvalues of $\Delta_A \otimes \Delta_B$ are also the pairwise products of the eigenValues of A and B. Since the eigenvalues of $B \otimes A$ are the pairwise products of the eigenvalues of B and A, they will be the same as the eigenvalues of $\Delta_A \otimes \Delta_B$.

i)

- The trace is the sum of the diagonal entries of a matrix. As a consequence, can also be computed as the sum of the eigenvalues of the matrix. If the eigenvalues of A are λ, and the eigenvalues of B are μ.
- Consider $A \in M_{mm}$

$$tr(A \otimes B) = tr\left(\begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mm}B \end{pmatrix} \right)$$

$$= \sum_{m=1}^{M} tr(a_{mm}B)$$

$$= \sum_{m=1}^{M} a_{mm} \cdot tr(B)$$

$$= tr(A) \cdot tr(B)$$

j)

- Since the detenninant of a matrix is the product of the eigenvalues of the matrix, $det(A \otimes B) = \Pi_{i=1}^{nm} \lambda_i$, where λ_i are the eigenvalues of $A \otimes B$. Let $\lambda_i = \alpha_j \cdot \beta_k$, where α_j is an eigenvalue of A and β_k is an eigenvalue of B.
- Proof:

$$det(A \otimes B) = \Pi_{i=1}^{nm} \lambda_i = \Pi_{j=1}^{m} \Pi_{k=1}^{n} (\alpha_j \beta_k)$$

$$= (\Pi_{j=1}^{m} {\alpha_j}^n)(\Pi_{k=1}^{n} {\beta_k}^m)$$

$$= det(A)^n det(B)^m$$

- Worked with Sherry Zhang.