Report 1: Vector and Matrix Algebra, Multivariate Normal Distribution

Paweł Grabiński

2024-04-22

Exercises

Exercise 1

Consider the matrix

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

1. Is A symmetric?

For a matrix to be symmetric, it must be equal to its transpose.

$$A^T = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

Thus, $A = A^T$ and A is symmetric.

2. and 3. Perform the spectral decomposition of A.

To obtain the spectral decomposition we have to find the eigenvalues and eigenvectors of A. To do so, we have to solve the following equation:

$$\det(A - \lambda I) = 0$$

In this case:

$$\det \begin{bmatrix} 3 - \lambda & -1 \\ -1 & 3 - \lambda \end{bmatrix} = 0$$

That is:

$$\det \begin{bmatrix} 3 - \lambda & -1 \\ -1 & 3 - \lambda \end{bmatrix} = (3 - \lambda)^2 - 1 = 8 - 6\lambda + \lambda^2 = (\lambda - 4)(\lambda - 2) = 0$$

The eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = 2$. To find the eigenvectors, we have to solve the following equations:

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

For $\lambda_1 = 4$:

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 3v_{11} - v_{12} \\ -v_{11} + 3v_{12} \end{bmatrix} = 4 \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

We get the following equations:

$$3v_{11} - v_{12} = 4v_{11} - v_{11} + 3v_{12} = 4v_{12}$$

That gets us:

$$v_{11} = -v_{12} - v_{11} = v_{12}$$

As these equations are equivalent, we can choose $v_{11} = 2^{-\frac{1}{2}}$ and $v_{12} = -2^{-\frac{1}{2}}$ for normalization purpose. Thus, the eigenvector corresponding to $\lambda_1 = 4$ is $\mathbf{v}_1 = \begin{bmatrix} 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} \end{bmatrix}$.

For $\lambda_2 = 2$:

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 3v_{21} - v_{22} \\ -v_{21} + 3v_{22} \end{bmatrix} = 2 \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}$$

We get the following equations:

$$3v_{21} - v_{22} = 2v_{21} - v_{21} + 3v_{22} = 2v_{22}$$

Simplifying, we get:

$$v_{21} = v_{22} v_{21} = v_{22}$$

Analogically, for normalization we pick $v_{21} = 2^{-\frac{1}{2}}$ and $v_{22} = 2^{-\frac{1}{2}}$. Thus, the eigenvector corresponding to $\lambda_2 = 2$ is $\mathbf{v}_2 = \begin{bmatrix} 2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} \end{bmatrix}$.

The spectral decomposition of A is:

$$\lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T$$

$$= 4 \begin{bmatrix} 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \end{bmatrix} + 2 \begin{bmatrix} 2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$$

$$= 4 \begin{bmatrix} 2^{-1} & -2^{-1} \\ -2^{-1} & 2^{-1} \end{bmatrix} + 2 \begin{bmatrix} 2^{-1} & 2^{-1} \\ 2^{-1} & 2^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = A$$

4. Use the spectral decomposition of A given above and find $\sqrt{A}$. Check that the matrix you found satisfies: $\sqrt{A}\sqrt{A} = A$.

To find $\sqrt{A}$, we can use the spectral decomposition of A:

$$A = [v_1, v_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} [v_1, v_2]^T$$

The square root of A will be:

$$\sqrt{A} = [v_1, v_2] \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} [v_1, v_2]^T$$

In our case it is:

$$\sqrt{A} = \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}^T$$

$$= \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2^{\frac{1}{2}} & -2^{\frac{1}{2}} \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 2^{-\frac{1}{2}} & -1 + 2^{-\frac{1}{2}} \\ -1 + 2^{-\frac{1}{2}} & 1 + 2^{-\frac{1}{2}} \end{bmatrix}$$

We can check if $\sqrt{A}\sqrt{A} = A$:

$$\sqrt{A}\sqrt{A} = \begin{bmatrix} 1 + 2^{-\frac{1}{2}} & -1 + 2^{-\frac{1}{2}} \\ -1 + 2^{-\frac{1}{2}} & 1 + 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 1 + 2^{-\frac{1}{2}} & -1 + 2^{-\frac{1}{2}} \\ -1 + 2^{-\frac{1}{2}} & 1 + 2^{-\frac{1}{2}} \end{bmatrix}$$

2

$$= \begin{bmatrix} (1 + 2^{-\frac{1}{2}})^2 + (-1 + 2^{-\frac{1}{2}})^2 & 2(-1 + 2^{-\frac{1}{2}})(1 + 2^{-\frac{1}{2}}) \\ 2(-1 + 2^{-\frac{1}{2}})(1 + 2^{-\frac{1}{2}}) & (-1 + 2^{-\frac{1}{2}})^2 + (1 + 2^{-\frac{1}{2}})^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 2 \cdot 2^{-\frac{1}{2}} + 2^{-1} + 1 - 2 \cdot 2^{-\frac{1}{2}} + 2^{-1} & 2(2^{-1} - 1) \\ 2(2^{-1} - 1) & 1 + 2 \cdot 2^{-\frac{1}{2}} + 2^{-1} + 1 - 2 \cdot 2^{-\frac{1}{2}} + 2^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = A$$

Exercsise 2

Consider the spectral decomposition of a positive definite matrix as given in Lecture 1:

$$A = P\Lambda P^T.$$

The columns of P are made of eigenvectors e_i, $i = 1, ..., n$ and they are orthonormalized, i.e. their lengths are one and they are orthogonal (peripendicular) one to another. The diagonal matrix Λ has the corresponding (positive) eigenvalues on the diagonal. Provide argument for the following: ### 1. $P^T = P^{-1}$

Now, by the definition of the inverse matrix, we have:

$$P^{-1}P = I,$$

where I is the identity matrix.

Similarly, we can consider:

$$(P^T P)_{km} = \sum_{i=1}^{n} (e_k^T)_i (e_m)_i = e_k^T e_m.$$

By the orthonormality we have that:

$$e_i^T e_j = \delta_{ij}.$$

That gives us:

$$(P^T P)_{km} = e_k^T e_m = \delta_{km}.$$

And in the matrix form:

$$P^T P = I.$$

Then we have:

$$P^T P - P^{-1} P = 0 (P^T - P^{-1})P = 0$$

As P is non-singular, we can multiply both sides by P^{-1}:

$$P^T - P^{-1} = 0 P^T = P^{-1}$$

2. Determinant of Λ is the product of the terms on the diagonal.

We can show that by induction. Let's consider diagonal square matrices $A^{(k)}$ where k is the number of dimensions and their entries are positive.

For $n = 1$ according to the lecture notes:

$$\det(A) = a_{11} = \prod_{i=1}^{1} a_{ii}$$

For $n > 1$, we assume that $\det(A^{(n)}) = \prod_{i=1}^{n} a_{ii}$. We can consider the matrix $A^{(n+1)}$ with the formula from the lecture:

$$\det(A^{(n+1)}) = \sum_{j}^{k} a_{1j}(-1)^{1+j}\det(A_{1j}),$$

where A_{1j} is the matrix obtained if row 1 and column j are deleted.

As the matrix $A^{(n+1)}$ is diagonal, the only non-zero element is a_{11}. Thus, the determinant is:

$$\det(A^{(n+1)}) = a_{11}(-1)^{1+1}\det(A_{11}) = a_{11}\det(A_{11})$$

Now, we removed the row and column with indices 1, so the matrix A_{11} is diagonal and belongs the our family of matrices. It can be denoted as $A^{(n)}$ (with its entries being entries of the original $A^{(n+1)}$ matrix with indices starting from 2 going to $n + 1$) Thus, we can write:

$$\det(A^{(n+1)}) = a_{11}\det(A_{11}) = a_{11}\prod_{i=2}^{n+1} a_{ii} = \prod_{i=1}^{n+1} a_{ii}$$

As we have shown that the formula holds for $n = 1$ and that if it holds for n it holds for $n + 1$, we have shown that the formula holds for all n.

3. Determinat of A is the same as that of Λ.

We can consider the determinants through the spectral decomposition of A:

$$\det(A) = \det(P\Lambda P^T) = \det(P)\det(\Lambda)\det(P^T) = \det(\Lambda)\det(P)\det(P^T) = \det(\Lambda)\det(PP^T).$$

Where we used the property of the determinant that $\det(AB) = \det(A)\det(B)$ from the lecture notes and the fact that determinants are scalars so their multiplication is commutative.

As we have shown in the previous point that $P^T = P^{-1}$, we have:

$$\det(A) = \det(\Lambda)\det(PP^T) = \det(\Lambda)\det(I) = \det(\Lambda).$$

4. Find the inverse of Λ, i.e. Λ^{-1}.

Since the eigenvalues λ_i are positive, we can define the matrix B as:

$$B = \mathrm{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}\right)$$

Then, we can consider the product of Λ and B:

$$\Lambda B = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_n\right)\mathrm{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}\right) = \mathrm{diag}\left(\frac{\lambda_1}{\lambda_1}, \ldots, \frac{\lambda_n}{\lambda_n}\right) = \mathrm{diag}\left(1, \ldots, 1\right) = I$$

By the definition of the inverse matrix, we have: that $B = \Lambda^{-1}$.

5. A simple way to determine the inverse of a matrix A from its spectral decomposition is through

$$A^{-1} = P\Lambda^{-1}P^T.$$

Verify that the right hand side of the above indeed define the inverse of A.

Let's consider the product of A and A^{-1}:

$$A^{-1}A = P\Lambda^{-1}P^T P\Lambda P^T = P\Lambda^{-1}I\Lambda P^T = P\Lambda^{-1}\Lambda P^T = PIP^T = I$$

6. Check all these statements on the little example of Exercise 1.

1. $P^T = \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}^T = \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $P^{-1} = \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}^{-1} = ((2^{-\frac{1}{2}})^2 - (-2^{-\frac{1}{2}} \cdot 2^{-\frac{1}{2}}))^{-1} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $P^T = P^{-1}$

2. $\det(\Lambda) = \det(\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}) = 4\det(2) = 4 \cdot 2 = 8$

3. $\det(A) = \det(\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}) = 3 \cdot 3 - (-1) \cdot (-1) = 9 - 1 = 8 = \det(\Lambda)$

4. $\Lambda^{-1} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$

5. $A^{-1}A = \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $= \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}^{-1} \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $= \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $= \begin{bmatrix} 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \\ -2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & 2^{-\frac{1}{2}} \end{bmatrix}$

   $= I$

Exercise 3

In a medical study, length L and weight W of newborn children is considered. It was assumed that (W, L) will be modeled through a bivariate normal distribution. The following information has been known: the mean weight is $3343[g]$, with the standard deviation of $528[g]$, while the mean length is $49.8[cm]$, with the standard deviation of $2.5[cm]$. Additionally the correlation between the length and the weight has been established and equal to 0.75. The joint distribution of (W, L) is bivariate normal, i.e. $(W, L) \sim N(\mu, \Sigma)$. Perform the following tasks and answer the questions:

1. **Write explicitly the parameters μ and Σ.**

$$\mu = \begin{bmatrix} 3343\ g \\ 49.8\ cm \end{bmatrix} \quad \Sigma = \begin{bmatrix} 528^2 & 0.75 \cdot 528 \cdot 2.5 \\ 0.75 \cdot 528 \cdot 2.5 & 2.5^2 \end{bmatrix} = \begin{bmatrix} 278784 & 990 \\ 990 & 6.25 \end{bmatrix}$$

2. **Write explicitly the density of the joint distribution.**

$$f_{W,L}(w, l) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} w \\ l \end{bmatrix} - \mu \right)^T \Sigma^{-1} \left(\begin{bmatrix} w \\ l \end{bmatrix} - \mu \right) \right)$$

$$\sqrt{|\Sigma|} = \sqrt{278784 \cdot 6.25 - 990^2} = \sqrt{762300} = 330\sqrt{7}$$

$$f_{W,L}(w, l) = \frac{1}{660\pi\sqrt{7}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} w \\ l \end{bmatrix} - \begin{bmatrix} 3343 \\ 49.8 \end{bmatrix} \right)^T \begin{bmatrix} 278784 & 990 \\ 990 & 6.25 \end{bmatrix}^{-1} \left(\begin{bmatrix} w \\ l \end{bmatrix} - \begin{bmatrix} 3343 \\ 49.8 \end{bmatrix} \right) \right)$$

$$= \frac{1}{660\pi\sqrt{7}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} w \\ l \end{bmatrix} - \begin{bmatrix} 3343 \\ 49.8 \end{bmatrix} \right)^T \begin{bmatrix} 6.25 & -990 \\ -990 & 278784 \end{bmatrix} \left(\begin{bmatrix} w \\ l \end{bmatrix} - \begin{bmatrix} 3343 \\ 49.8 \end{bmatrix} \right) \right)$$

3. Find eigenvalues and eigenvectors of the covariance matrix Σ. Sketch few elipses corresponding to the constant density contours of the joint distributions. Mark on the plot the eigenvectors scaled by the square roots of the corresponding eigenvalues and comment.

To find the eigenvalues, we have to solve the following equation:

$$\det(\Sigma - \lambda I) = 0$$

$$\det \begin{bmatrix} 278784 - \lambda & 990 \\ 990 & 6.25 - \lambda \end{bmatrix} = (278784 - \lambda)(6.25 - \lambda) - 990^2 = 1742400 - 278777.75\lambda + \lambda^2$$

$$1742400 - 278777.75\lambda + \lambda^2 = 0$$

$$\lambda_1 = \frac{1115161}{8} + \frac{\sqrt{1243535268721}}{8} \approx 278787.52$$

$$\lambda_2 = \frac{1115161}{8} - \frac{\sqrt{1243535268721}}{8} \approx 2.7343405$$

To find the eigenvectors, we have to solve the following equations:

$$\begin{bmatrix} 278784 & 990 \\ 990 & 6.25 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 278787.52 \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

$$\begin{bmatrix} 278784 & 990 \\ 990 & 6.25 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = 2.7343405 \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}$$

By solving with a numerical solver, we get the eigenvectors:

$$v_1 = \begin{bmatrix} 0.999993695 \\ 0.003551149 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} -0.003551149 \\ 0.999993695 \end{bmatrix}$$

Let's approximate them further:

$$v_1 = \begin{bmatrix} 1 \\ 0.00355 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} -0.00355 \\ 1 \end{bmatrix}$$

As noted during the lecture, the eigenvectors scaled by the square root of the eigenvalues give the directions of the axis of the ellipses. The ellipses are the constant density contours of the joint distribution. The vectors are orthogonal, but due to the radically different orders of magnitude of the eigenvalues and scales on the plot we can't see that.

4. How many parameters characterize a bivariate normal distribution? How many parameters characterize a p-dimensional normal distribution?

A bivariate normal distribution is characterized by 5 parameters: 2 means, 2 variances and 1 correlation coefficient - or equivalently by 3 elements of the covariance matrix as it is symmetric.

A p-dimensional normal distribution is characterized by p means, p variances and $\frac{p(p-1)}{2}$ correlation coefficients - as again the covariance matrix is symmetric, so we count only the diagonal and upper half of the matrix that is:

$$p + p - 1 + \cdots + 2 + 1 = \frac{p-1}{2} \cdot p = \frac{p(p+1)}{2} = p + \frac{p(p-1)}{2}$$

Together it is:

$$p + \frac{p(p+1)}{2} = \frac{p(p+3)}{2}$$

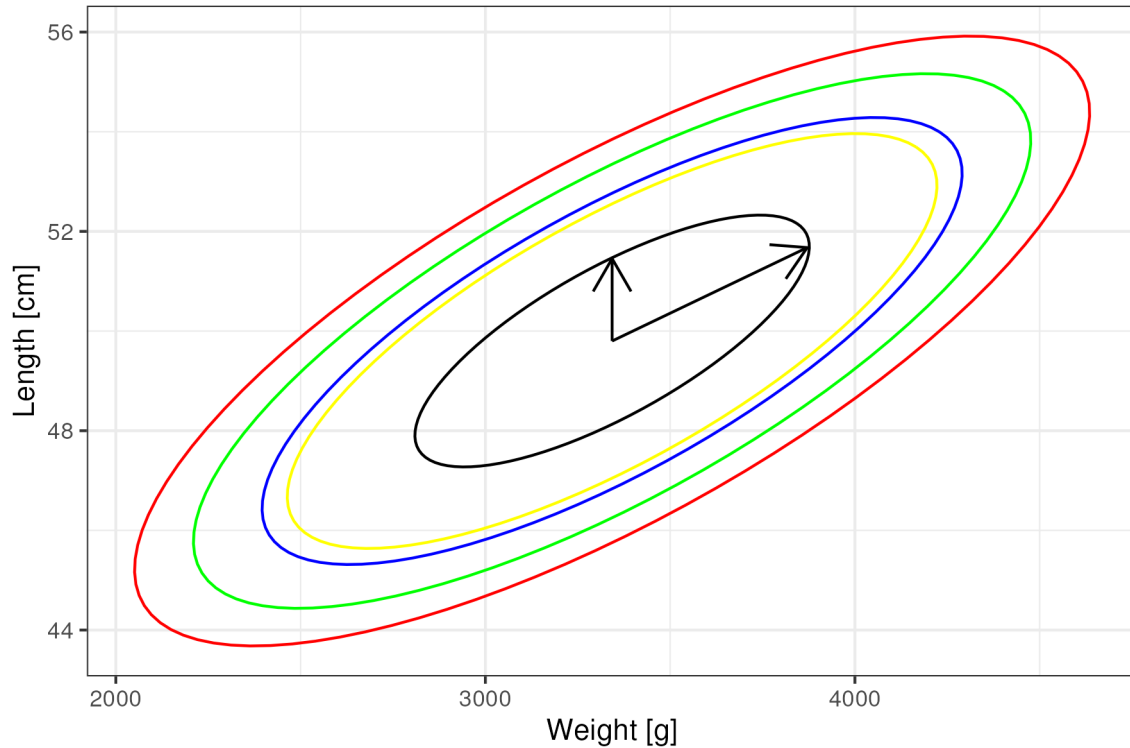We need $\frac{p(p+3)}{2}$ parameters.

Figure 1: Constant density elipsoids and the covariance matrix's eigenvectors

5. What is the distribution of L? Give its name and parameters.

As stated during the lecture, distribution of a subset of random variables from a multivariate normal distribution are also normal. Hence, the distribution of L is a normal distribution with the mean $\mu_L = 49.8$ and the variance $\sigma_L^2 = 6.25$.

6. Suppose that the hospital records of a new-born child was lost. Give a best guess for the value of their length. Provide with accuracy bounds of your 'educated' guess based on the $3 - \sigma$ rule.

The best guess for the length of the child is the mean of the distribution as the probability density has the highest value there, i.e. $49.8 \; cm$.

The accuracy bounds are given by the $3 - \sigma$ rule, that is:

$$49.8 - 3 \cdot 2.5 \text{ cm} \leq L \leq 49.8 + 3 \cdot 2.5 \text{ cm}$$

$$L \in [42.3 \text{ cm}, 57.3 \text{ cm}]$$

Exercise 4

In the setup of the previous problem, assume that it was reported by the mother of the child that weight was $4025g$.

1. What is the distribution of L given this additional information? Give its name and parameters.

2. Improve your previous guess and provide with accuracy limits.

3. Compare the answers from this and previous problems and comment how additional information affected the prediction value and accuracy.

Exercise 5

Let X_1, X_2, and X_3 be independent $N(\mu, \Sigma)$ random vectors of a dimension p.

1. Find the distribution of each of the following vectors:

$$V_1 = \frac{1}{4}X_1 - \frac{1}{2}X_2 + \frac{1}{4}X_3$$

$$V_2 = \frac{1}{4}X_1 - \frac{1}{2}X_2 - \frac{1}{4}X_3$$

First, let's construct a new X vector that is $3p$-dimensional and is a concatenation of X_1, X_2 and X_3:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

It's mean is:

$$\mu_X = \begin{bmatrix} \mu \\ \mu \\ \mu \end{bmatrix}$$

And its covariance matrix is:

$$\Sigma_X = \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix}$$

Then, we can construct the matrix A_1 that is:

$$A_1 = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix}$$

And $I_{p\times p}$ is $p \times p-$dimensional identity matrix. It's dimension is $p \times 3p$.

From the lecture we know that the distribution, i.e. the mean vector and the covariance matrix, of X changes as:

$$V_1 = A_1 X \sim N(A_1\mu, A_1\Sigma A_1^T)$$

Hence,

$$\mu_{V_1} = A_1\mu_X = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \mu \\ \mu \\ \mu \end{bmatrix}$$

$$= \frac{1}{4}I_{p\times p}\mu - \frac{1}{2}I_{p\times p}\mu + \frac{1}{4}I_{p\times p}\mu = (\frac{1}{4} - \frac{1}{2} + \frac{1}{4})\mu = 0$$

And:

$$\Sigma_{V_1} = A_1\Sigma_X A_1^T$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix}^T$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} \\ -\frac{1}{2}I_{p\times p} \\ \frac{1}{4}I_{p\times p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \frac{1}{4}\Sigma I_{p\times p} \\ -\frac{1}{2}\Sigma I_{p\times p} \\ \frac{1}{4}\Sigma I_{p\times p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \frac{1}{4}\Sigma \\ -\frac{1}{2}\Sigma \\ \frac{1}{4}\Sigma \end{bmatrix} = \left(\frac{1}{4}\right)^2 I_{p\times p}\Sigma + \left(\frac{1}{2}\right)^2 I_{p\times p}\Sigma + \left(\frac{1}{4}\right)^2 I_{p\times p}\Sigma$$

$$= \frac{3}{8}\Sigma$$

Finally, we have:

$$V_1 = A_1 X \sim N(0_{p\times 1}, \frac{3}{8}\Sigma)$$

Analogically, the A_2 matrix is:

$$A_2 = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix}$$

Then, the mean becomes:

$$\mu_{V_2} = A_2\mu_X =$$

$$\begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \mu \\ \mu \\ \mu \end{bmatrix}$$

$$= \frac{1}{4}I_{p\times p}\mu - \frac{1}{2}I_{p\times p}\mu - \frac{1}{4}I_{p\times p}\mu = (\frac{1}{4} - \frac{1}{2} - \frac{1}{4})\mu = -\frac{1}{2}\mu$$

And the covariance matrix:

$$\Sigma_{V_2} = A_2\Sigma_X A_2^T = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix}^T$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} \\ -\frac{1}{2}I_{p\times p} \\ -\frac{1}{4}I_{p\times p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \frac{1}{4}\Sigma I_{p\times p} \\ -\frac{1}{2}\Sigma I_{p\times p} \\ -\frac{1}{4}\Sigma I_{p\times p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \frac{1}{4}\Sigma \\ -\frac{1}{2}\Sigma \\ -\frac{1}{4}\Sigma \end{bmatrix}$$

$$= \left(\frac{1}{4}\right)^2 I_{p\times p}\Sigma + \left(\frac{1}{2}\right)^2 I_{p\times p}\Sigma + \left(\frac{1}{4}\right)^2 I_{p\times p}\Sigma = \frac{3}{8}\Sigma$$

The distribution of V_2 is:

$$V_2 = A_2 X \sim N(-\frac{1}{2}\mu, \frac{3}{8}\Sigma)$$

2. Find the joint distribution of the above vectors.

In this case, the matrix A is:

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix}$$

Then, the mean vector is:

$$\mu_V = A\mu_X = \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \mu \\ \mu \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2}\mu \end{bmatrix}$$

And the covariance matrix is:

$$\Sigma_V = A\Sigma_X A^T$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix}^T$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ -\frac{1}{2}I_{p\times p} & -\frac{1}{2}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & \frac{1}{4}I_{p\times p} \\ \frac{1}{4}I_{p\times p} & -\frac{1}{2}I_{p\times p} & -\frac{1}{4}I_{p\times p} \end{bmatrix} \begin{bmatrix} \frac{1}{4}\Sigma & \frac{1}{4}\Sigma \\ -\frac{1}{2}\Sigma & -\frac{1}{2}\Sigma \\ \frac{1}{4}\Sigma & -\frac{1}{4}\Sigma \end{bmatrix}$$

$$= \begin{bmatrix} \left(\frac{1}{4}\right)^2\Sigma + \left(\frac{1}{2}\right)^2\Sigma + \left(\frac{1}{4}\right)^2\Sigma & \left(\frac{1}{4}\right)^2\Sigma + \left(\frac{1}{2}\right)^2\Sigma - \left(\frac{1}{4}\right)^2\Sigma \\ \left(\frac{1}{4}\right)^2\Sigma + \left(\frac{1}{2}\right)^2\Sigma - \left(\frac{1}{4}\right)^2\Sigma & \left(\frac{1}{4}\right)^2\Sigma + \left(\frac{1}{2}\right)^2\Sigma + \left(\frac{1}{4}\right)^2\Sigma \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{8}\Sigma & \frac{1}{4}\Sigma \\ \frac{1}{4}\Sigma & \frac{3}{8}\Sigma \end{bmatrix}$$

Finally, we have:

$$V = AX \sim N\left(\begin{bmatrix} 0 \\ -\frac{1}{2}\mu \end{bmatrix}, \begin{bmatrix} \frac{3}{8}\Sigma & \frac{1}{4}\Sigma \\ \frac{1}{4}\Sigma & \frac{3}{8}\Sigma \end{bmatrix}\right)$$

Project 1: Weight and length of newborn children

Health services and health insurance companies are interested in determining what kind of medical examinations and diagnostic procedures should be administered to a newborn child. In one approach, there is a score system based on which it is determined when a child is healthy and does not require any special attention or when it is not in which case a series additional medical tests are performed. Weight and length of a newborn child are most standard indicators of the health of a child. In order to decide on the score the following procedure is considered. If the weight and length fall outside 95% of the typical values for the population, the score of zero is given. If the measurements are falling in the category between 75% and 95% the score is one. In all other cases the score of two is assigned. A random sample of records for 736 recently born children (singleton and not prematurely born) has been considered from hospital across a certain region. The records contain a large variety of information but extraction of weight and height data are given in the file WeightHeight.txt.

Part 1

1. Using the data estimate the mean and the covariance for the length and the weight of children.

As given in the notes from the lecture, the sample mean is:

$$\bar{\mu} = \frac{1}{n_{sample}} \sum_i^{n_{sample}} \begin{bmatrix} W_i \\ L_i \end{bmatrix}$$

In our case:

$$\bar{\mu} = \begin{bmatrix} 3233.55 \ g \\ 49.24 \ cm \end{bmatrix}$$

Similarly, the sample covariance matrix is:

$$\bar{\Sigma} = \frac{1}{n_{sample} - 1} \sum_i^{n_{sample}} \left(\begin{bmatrix} W_i \\ L_i \end{bmatrix} - \bar{\mu} \right) \left(\begin{bmatrix} W_i \\ L_i \end{bmatrix} - \bar{\mu} \right)^T$$

For this data it is:

$$\bar{\Sigma} = \begin{bmatrix} 220276.7 & 915.3 \\ 915.3 & 4.44 \end{bmatrix}$$

2. Verify graphically normal distribution of the data. Use a scatterplot and qq-plots for the marginal distributions.

The scatter plot of the data is:

We can see that the data is highly correlated.

The qq-plot for the length marginal distribution:

The qq-plot for the weight marginal distribution:

We can see that the data is normally distributed.

3. Find the ellipsoids that would serve classification regions for scores as described above.

4. How many children would score zero, one, and two, respectively? Illustrate this classification on the graphs.

Score	Number of children
0	38
1	157
2	541

5. Find the spectral decomposition of the estimated covariance matrix.

The spectral decomposition of the estimated covariance matrix is:

$$\bar{\Sigma} = 2.2 \cdot 10^5 \begin{bmatrix} 1 \\ 4.15 \cdot 10^{-3} \end{bmatrix} \begin{bmatrix} 1 \\ 4.15 \cdot 10^{-3} \end{bmatrix}^T + 0.64 \begin{bmatrix} -4.15 \cdot 10^{-3} \\ 1 \end{bmatrix} \begin{bmatrix} -4.15 \cdot 10^{-3} \\ 1 \end{bmatrix}^T$$

$$= \begin{bmatrix} 220276.6577 & 915.298170 \\ 915.2982 & 3.803266 \end{bmatrix} + \begin{bmatrix} 1.105078 \cdot 10^{-5} & -2.659494 \cdot 10^{-3} \\ -2.659494 \cdot 10^{-3} & 0.640036621 \end{bmatrix}$$
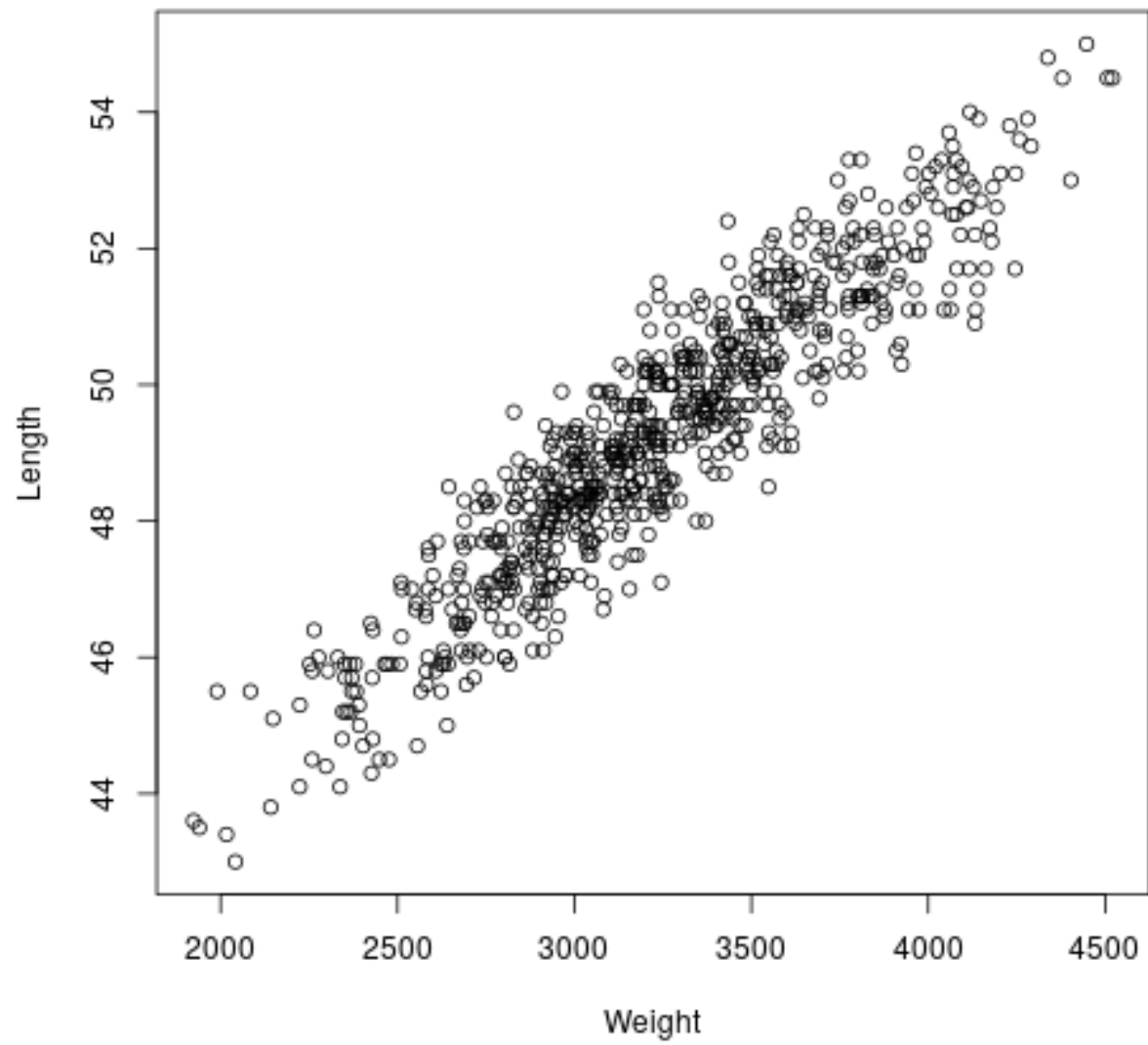
Figure 2: Children body weight and length

Normal Q-Q Plot
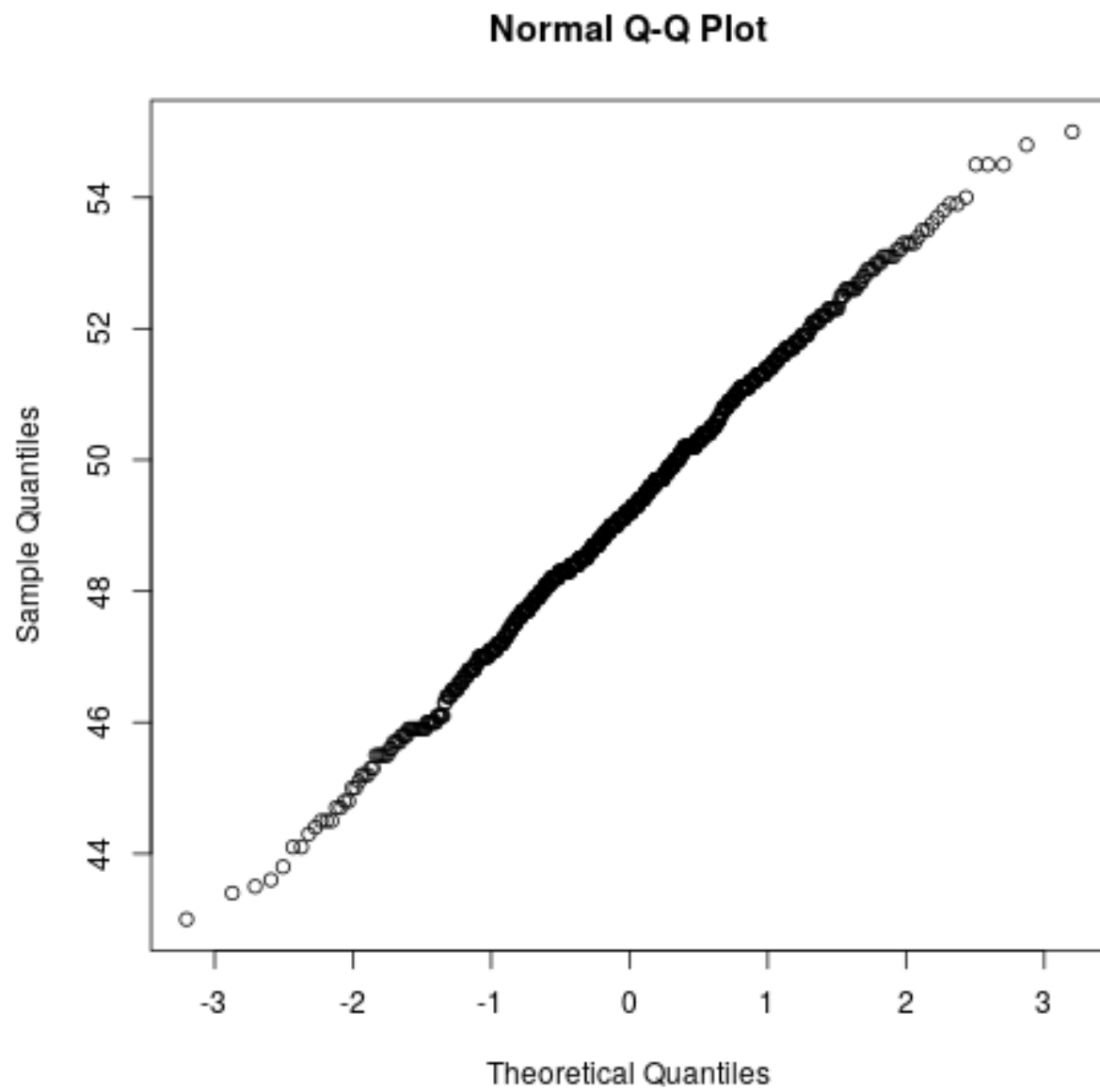

Figure 3: Children body length quantile plot
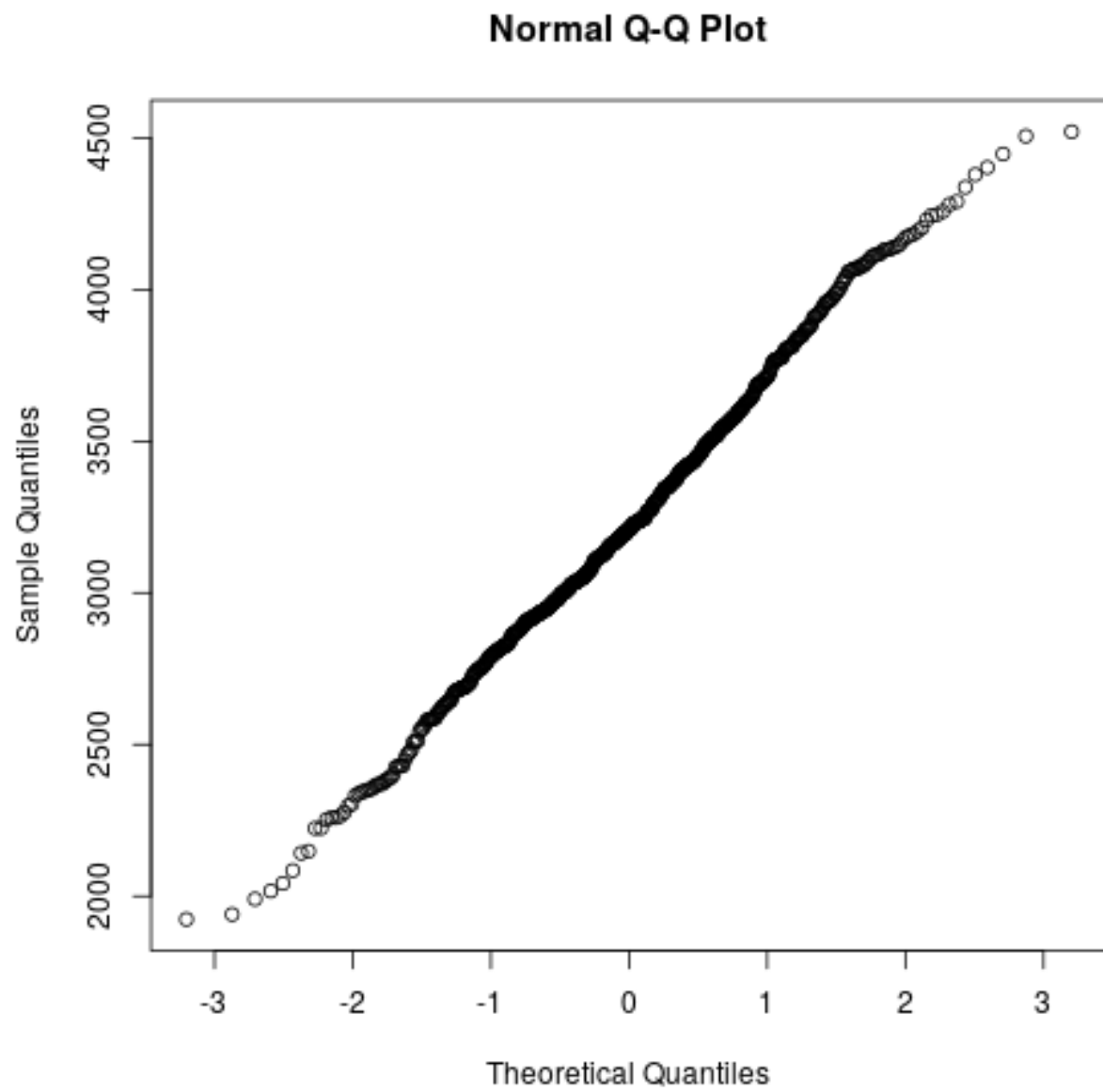
Normal Q-Q Plot
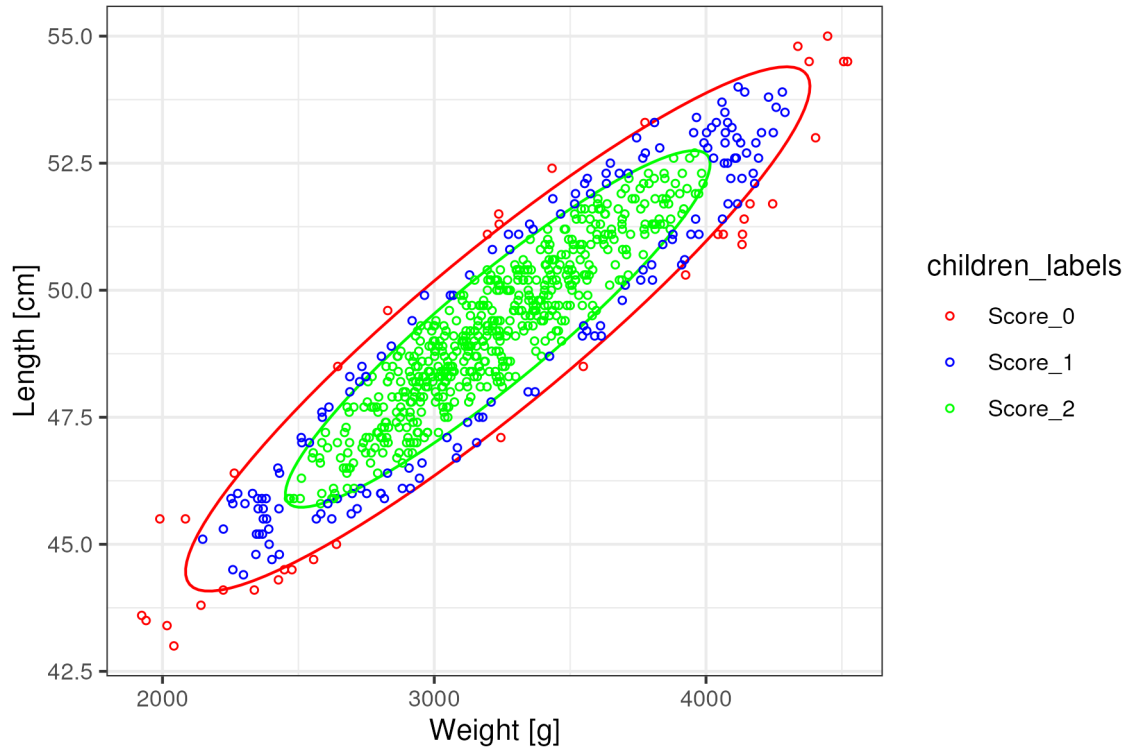

Figure 4: Children body weight quantile plot

Figure 5: Classification constant density ellipsoids

6. Plot the data transformed according to $P^T X$, where P is the matrix made of the eigenvectors standing as the columns. Interpret the transformed data.

The transformed data seems to be decorellated - the constant-density ellipsoids would be much more round than the original ones. The data is transformed into the eigenvector space, where the covariance matrix is diagonal.

Part 2

Additionally to weight and length of a child, also the height of parents is included in the records. In order to tune the procedure of scoring the height of parents can be also used. The ParentsWeightLength.txt file contains this information.

1. Using the data estimate the mean and the covariance for all four variables.

The estimated means: |Variable | Mean| |:————|——:| |Weight | 3233.545| |Length | 49.238| |FatherHeight | 177.416| |MotherHeight | 166.920|

The estimated covariance matrix: | | Weight| Length| FatherHeight| MotherHeight| |:————|————-:|————-:|————-:|————-:| |Weight | 220276.658| 915.296| 931.859| 827.288| |Length | 915.296| 4.443| 3.290| 2.852| |FatherHeight | 931.859| 3.290| 12.612| 0.631| |MotherHeight | 827.288| 2.852| 0.631| 9.772|

2. Verify graphically normal distribution of the data. Use a scatterplot and qq-plots for the marginal distributions.

In the plot below, we can see that the data seems to be normally distributed. The correlation between the length and weight of the child and the parents' heights seems to be lower than between the length and weight of the child.
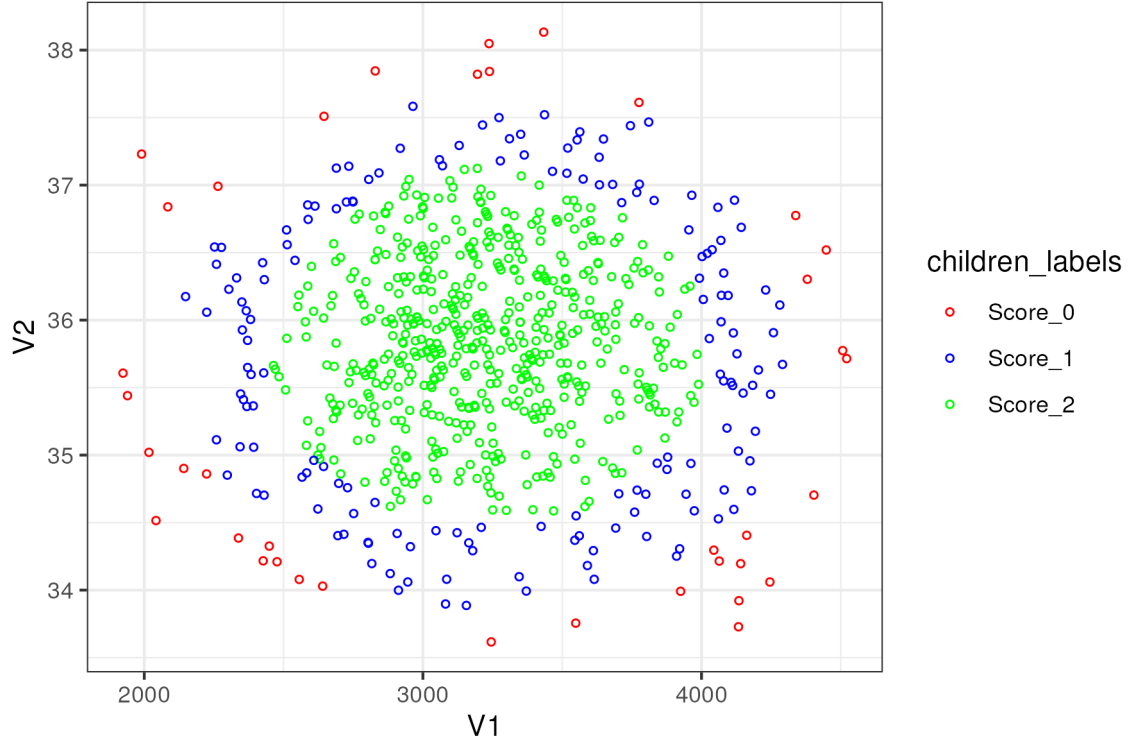
15

Figure 6: Children data transformed into eigenvector space

The quantile plot for the length marginal distribution:

The quantile plot for the weight marginal distribution:

The quantile plot for the father height marginal distribution:

The quantile plot for the mother height marginal distribution:

Judging by the quantile plots, the data seems to be normally distributed.

3. Identify the conditional distribution of the weight and length of a child given the heights of parents. Find an estimate of the covariance matrix of the conditional distribution and compare it with the original unconditional covariance.

From the lecture we know that the conditional distribution of a multivariate normal distribution is given as:

$$\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{x}_2 \sim \mathcal{N}_q \left(\mu_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \left(\mathbf{x}_2 - \boldsymbol{\mu}_2 \right), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} \right)$$

Where $X \sim N_p(\mu, \Sigma)$, X_1 is the $q-$dimensional subvector of X and X_2 is the $(p-q)-$dimensional subvector of X:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Smiliarly the mean vector and the covariance matrix are:

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

In our case, the weight and length of the child are X_1 and the heights of the parents are X_2.
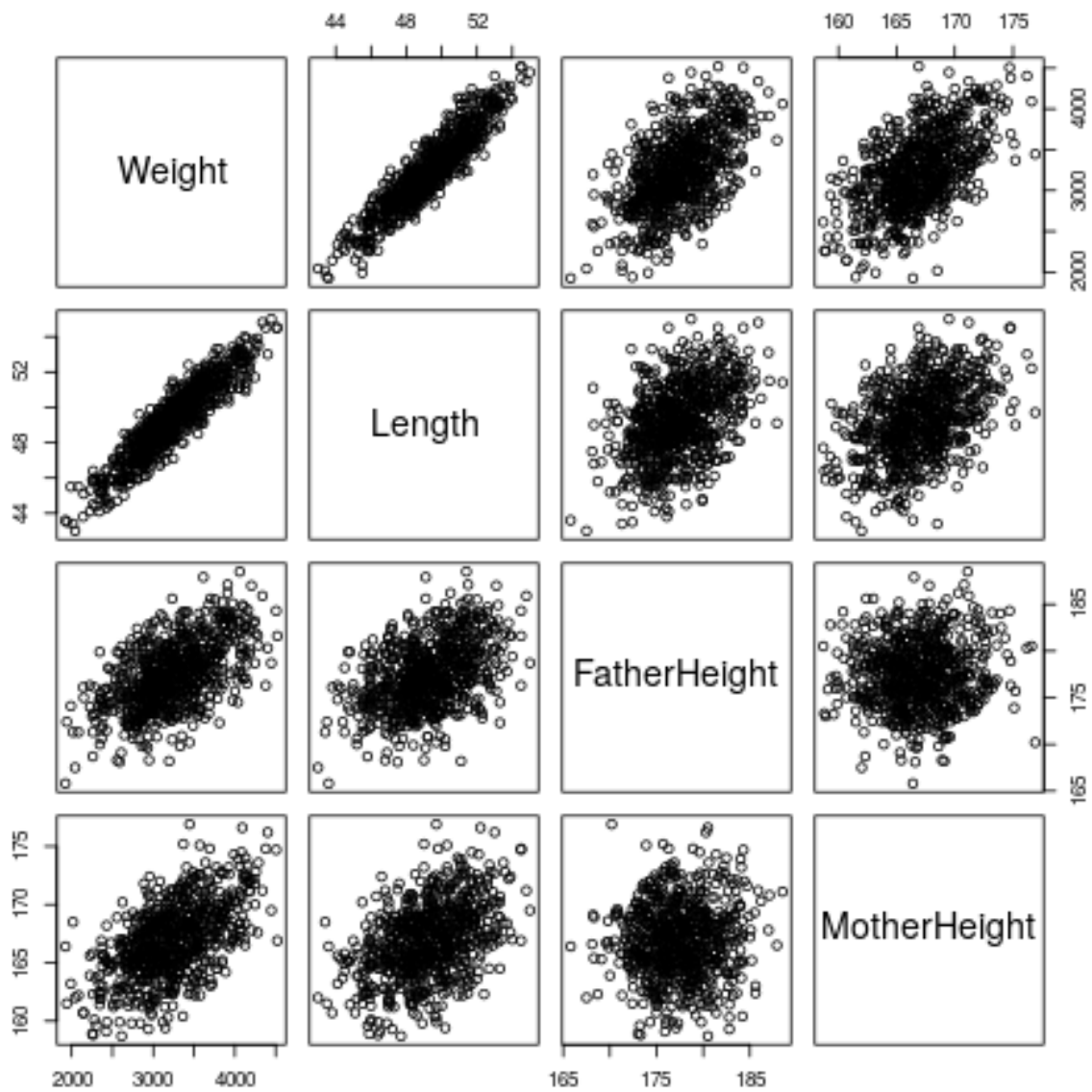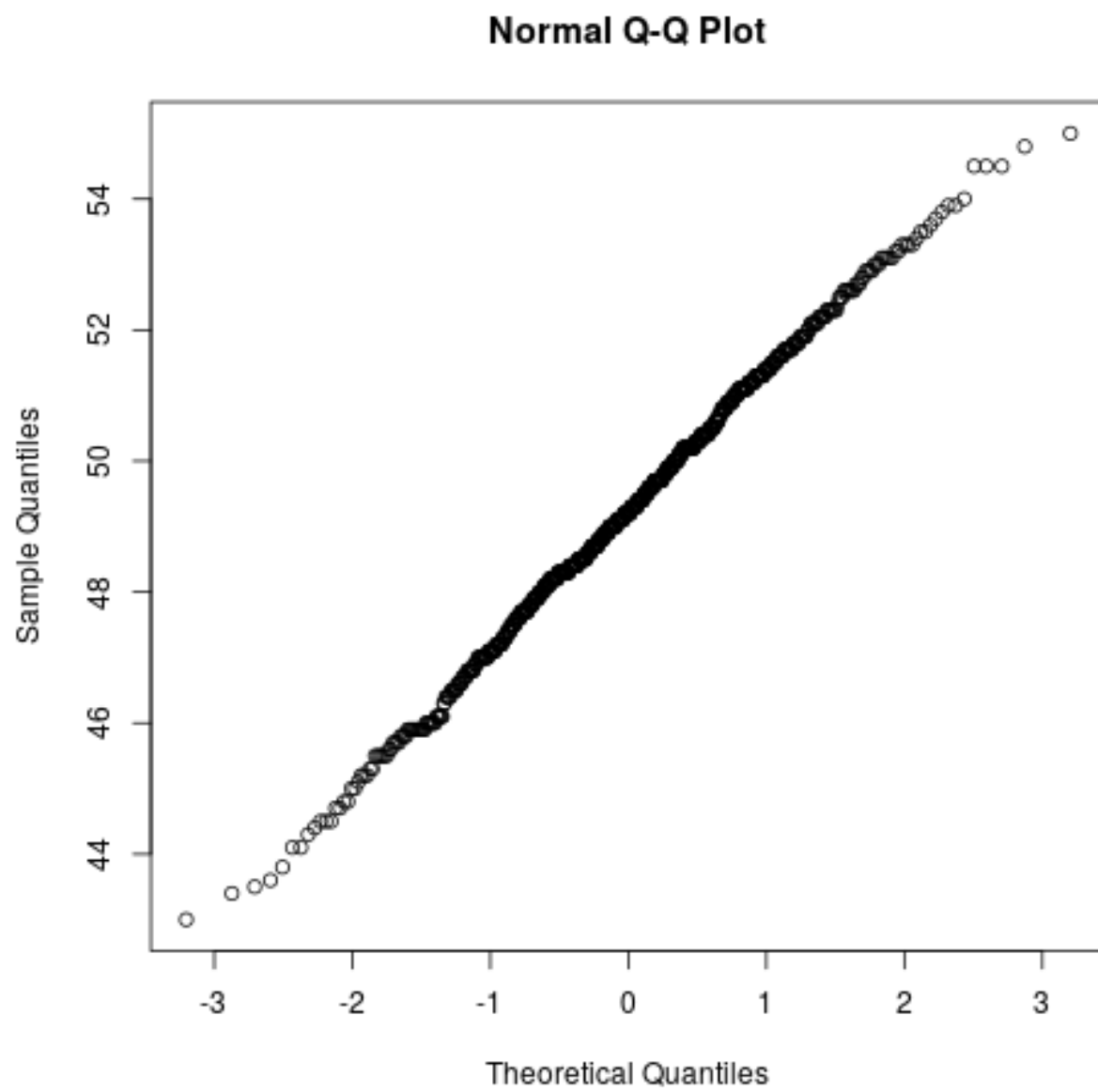
16

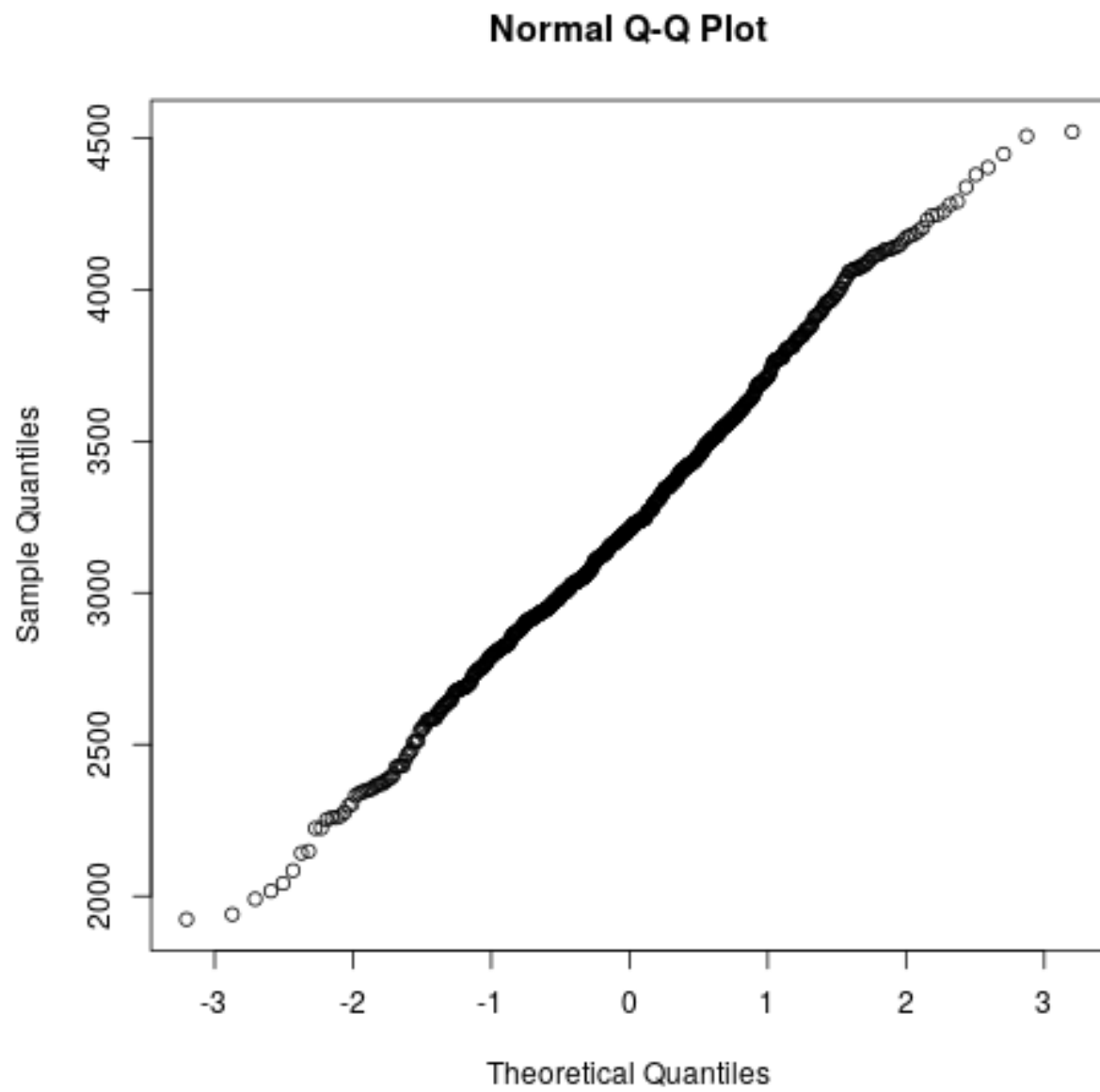Figure 7: Parents data scatter plot

Figure 8: Parents length qq plot

Figure 9: Parents weight qq plot
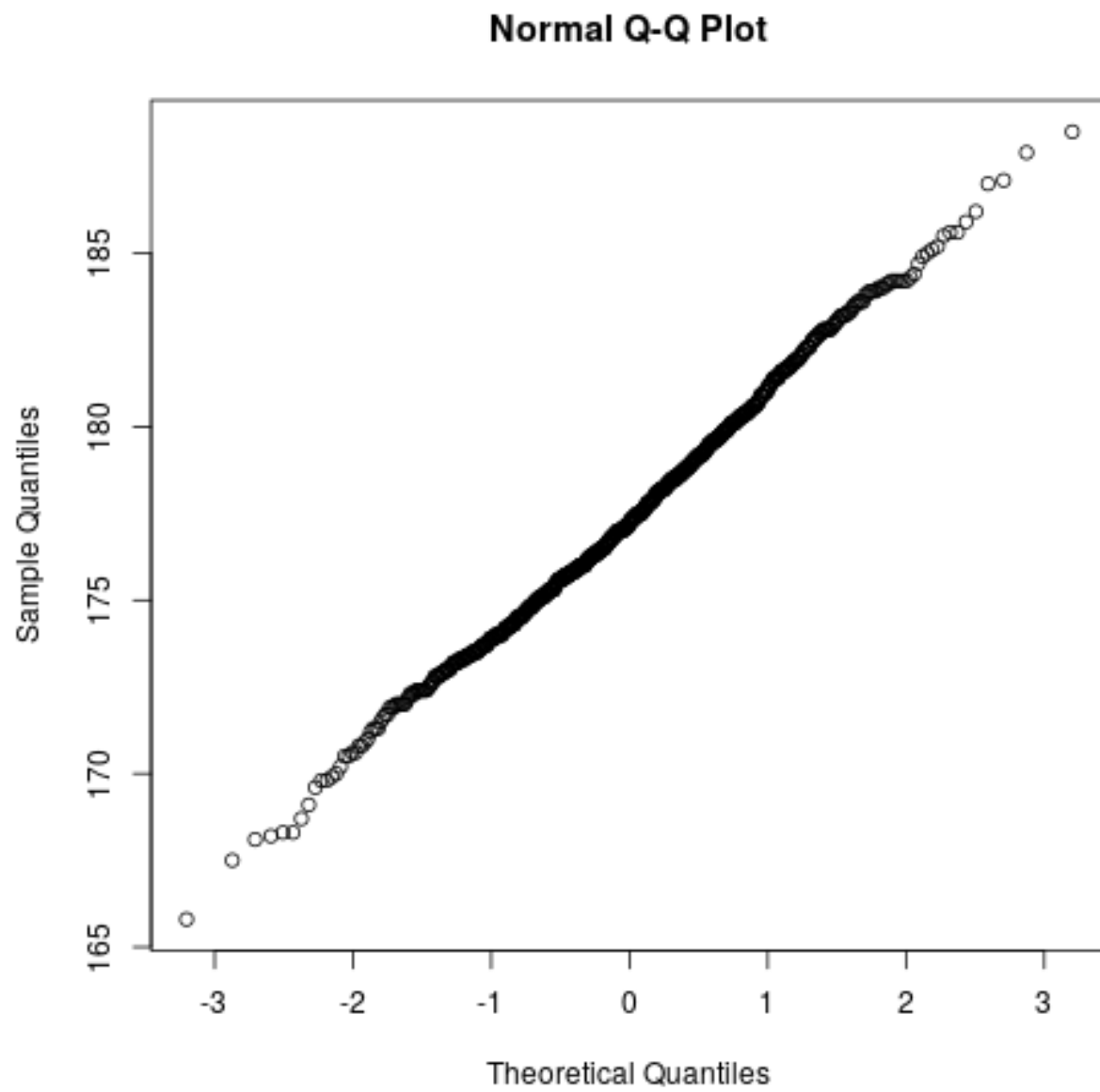
Normal Q-Q Plot

Figure 10: Parents father height qq plot
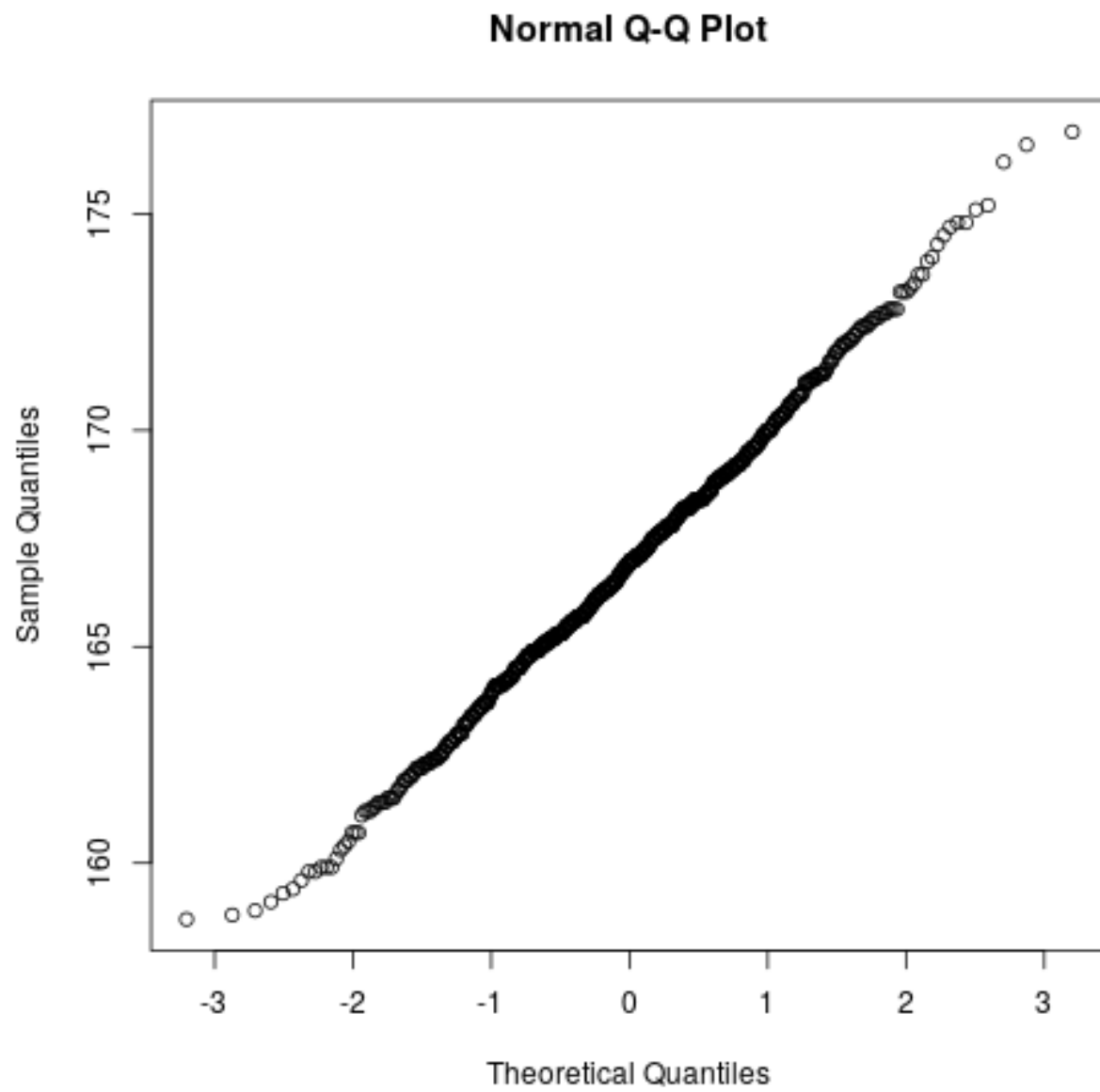
Normal Q-Q Plot

Figure 11: Parents mother height qq plot

The means become:
$$\mu_1 = \begin{bmatrix} 3233.55 \\ 49.24 \end{bmatrix} \quad \mu_2 = \begin{bmatrix} 177.42 \\ 166.92 \end{bmatrix}$$

The covariance matrices become:
$$\Sigma_{11} = \begin{bmatrix} 220276.7 & 915.3 \\ 915.3 & 4.44 \end{bmatrix} \quad \Sigma_{12} = \begin{bmatrix} 931.859 & 827.288 \\ 3.290 & 2.852 \end{bmatrix}$$

$$\Sigma_{21} = \begin{bmatrix} 931.859 & 3.290 \\ 827.288 & 2.852 \end{bmatrix} \quad \Sigma_{22} = \begin{bmatrix} 12.612 & 0.631 \\ 0.631 & 9.772 \end{bmatrix}$$

The new mean is:
$$\tilde{\mu} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$

$$= \begin{bmatrix} 3233.55 \\ 49.24 \end{bmatrix} + \begin{bmatrix} 931.859 & 827.288 \\ 3.290 & 2.852 \end{bmatrix} \begin{bmatrix} 12.612 & 0.631 \\ 0.631 & 9.772 \end{bmatrix}^{-1} \left(x_2 - \begin{bmatrix} 177.42 \\ 166.92 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 3233.55 \\ 49.24 \end{bmatrix} + \begin{bmatrix} 69.876 & 80.146 \\ 0.247 & 0.276 \end{bmatrix} \left(x_2 - \begin{bmatrix} 177.42 \\ 166.92 \end{bmatrix} \right)$$

The new covariance matrix is:
$$\tilde{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \begin{bmatrix} 88857.9955 & 456.846078 \\ 456.8461 & 2.843767 \end{bmatrix}$$

The conditional distribution of the weight and length of a child given the heights of parents is:
$$\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{x}_2 \sim \mathcal{N}_q \left(\tilde{\mu}(x_2), \tilde{\Sigma} \right)$$

Taking the element-wise division of $\tilde{\Sigma}$ by $\bar{\Sigma}$ we get: | | Weight| Length| |:——|——:|——:| |Weight | 0.403| 0.499| |Length | 0.499| 0.640|

That means the conditional distribution is more concentrated around the mean.

4. How the ellipsoids based on the conditional distribution will look like?

We consider the conditional distribution of the weight and length of a child given the heights of parents equal to their means $x_2 = \mu_2$, so the $\tilde{\mu} = \mu_1$.

The red and dark green are 0.95 and 0.75 ellipsoids for the unconditional distribution, while the orange and light green are 0.95 and 0.75 ellipsoids for the conditional distribution.

We can see that first of all the ellipsoids of the conditional distribution are more concentrated around the mean. Secondly, the correlation between the weight and length of the child is higher in the conditional distribution as can be inferred from the change of the covariance matrix as the diagonal elements decreased more than the off-diagonal elements.

5. How many children would score zero, one, and two, respectively? Illustrate this classification on the graph and compare with the one obtained without considering the heights of parents.

Scores	Occurances
0	40
1	137
2	559

The elipsoids on the graph are the 0.95 and 0.75 elipsoids of the conditional distribution with the center at the unconditional sample mean. The classification does not allign with the given elipsoids as the classificaiton depends on the conditional distribution, i.e. the height of the parents.
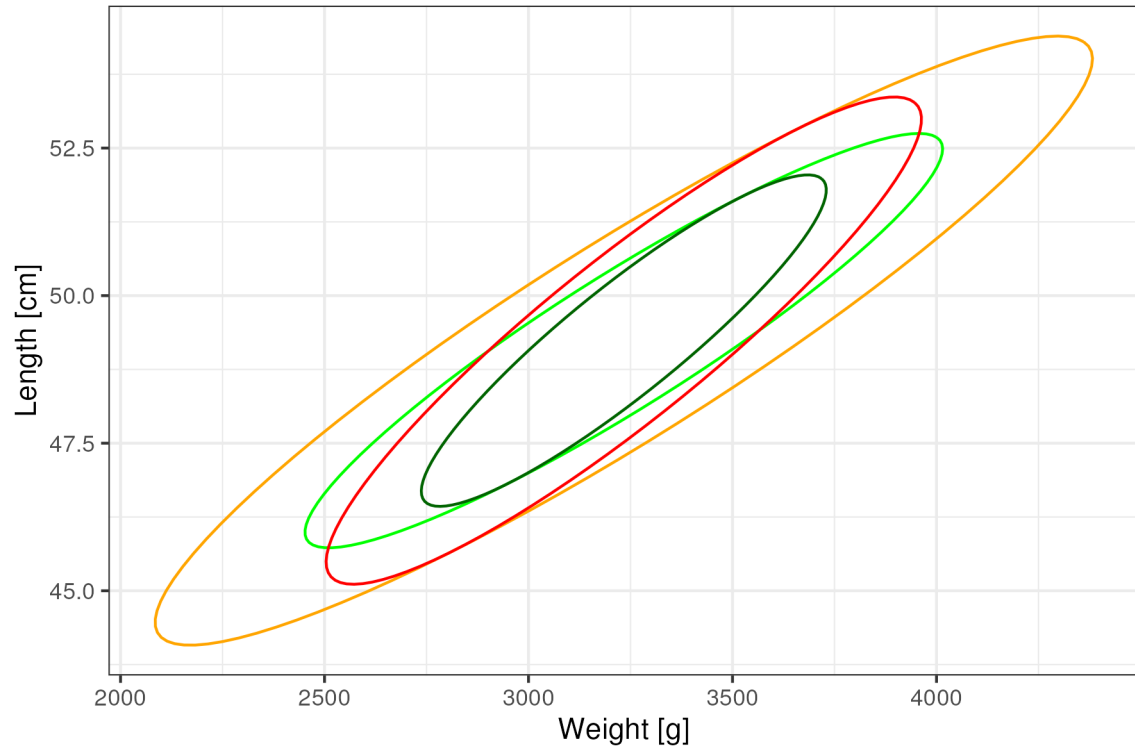
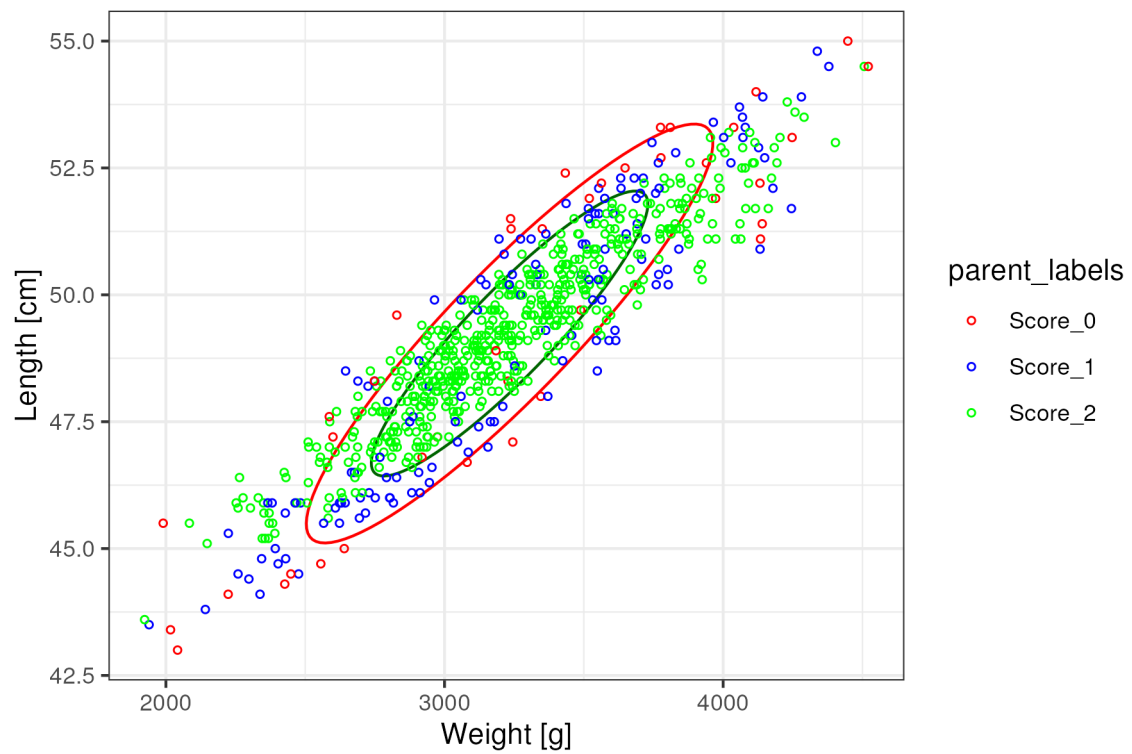Figure 12: Comparison of the classification elipsoids - conditional and independent

Figure 13: Classification based on the conditional distribution

6. Suppose that the father of a child is 185[cm] tall and mother is 178[cm] tall. Plot the classification ellipsoids for their child.
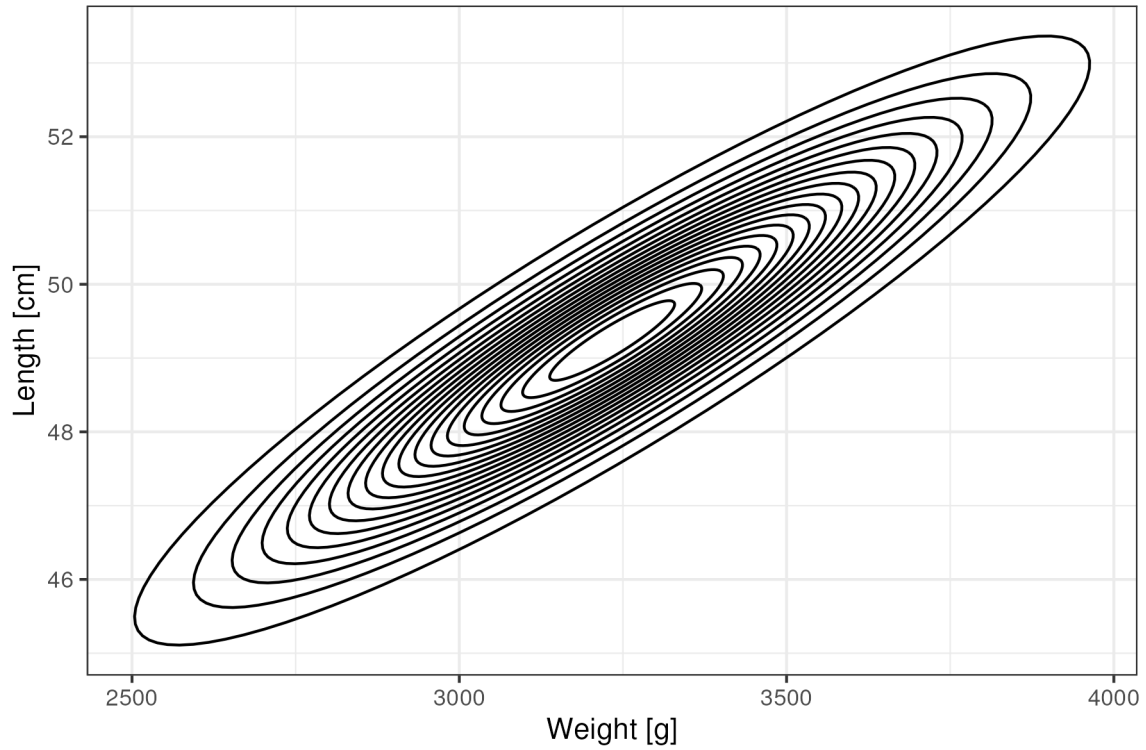

Figure 14: Classification based on the conditional distribution for the given parents' heights

7. Find spectral decomposition of the estimated covariance matrix for the complete set of the data.

The eigenvalues are:

λ_i
2.202875e+05
1.070822e+01
4.788382e+00
4.783659e-01

The eigenvectors are:

v_1	v_2	v_3	v_4
0.9999754	0.0012288	0.0046024	0.0051550
0.0041551	0.0136742	0.1925554	-0.9811820
0.0042304	-0.8164120	-0.5643934	-0.1221213
0.0037556	0.5773066	-0.8027212	-0.1494712

24

The decomposition is:

$$\bar{\Sigma} = \sum_i \lambda_i v_i v_i^T = 2.2 \cdot 10^5 \begin{bmatrix} 1 \\ 4.15 \cdot 10^{-3} \\ 4.23 \cdot 10^{-3} \\ 3.76 \cdot 10^{-3} \end{bmatrix} \begin{bmatrix} 1 \\ 4.15 \cdot 10^{-3} \\ 4.23 \cdot 10^{-3} \\ 3.76 \cdot 10^{-3} \end{bmatrix}^T + 10.71 \begin{bmatrix} 1.12 \cdot 10^{-3} \\ 1.37 \cdot 10^{-2} \\ -0.816 \\ 0.577 \end{bmatrix} \begin{bmatrix} 1.12 \cdot 10^{-3} \\ 1.37 \cdot 10^{-2} \\ -0.816 \\ 0.577 \end{bmatrix}^T$$

$$+4.79 \begin{bmatrix} 4.60 \cdot 10^{-3} \\ 0.192 \\ -0.564 \\ -0.803 \end{bmatrix} \begin{bmatrix} 4.60 \cdot 10^{-3} \\ 0.192 \\ -0.564 \\ -0.803 \end{bmatrix}^T + 0.48 \begin{bmatrix} 5.16 \cdot 10^{-3} \\ -0.981 \\ -0.122 \\ -0.149 \end{bmatrix} \begin{bmatrix} 5.16 \cdot 10^{-3} \\ -0.981 \\ -0.122 \\ -0.149 \end{bmatrix}^T$$

$$= \begin{bmatrix} 220276.6575 & 915.293507 & 931.882445 & 827.298245 \\ 915.2935 & 3.803227 & 3.872158 & 3.437589 \\ 931.8824 & 3.872158 & 3.942337 & 3.499893 \\ 827.2982 & 3.437589 & 3.499893 & 3.107104 \end{bmatrix}$$

$$+ \begin{bmatrix} 0.0000162 & 0.0001799 & -0.0107427 & 0.0075964 \\ 0.0001799 & 0.0020023 & -0.1195446 & 0.0845332 \\ -0.0107427 & -0.1195446 & 7.1373337 & -5.0469986 \\ 0.0075964 & 0.0845332 & -5.0469986 & 3.5688671 \end{bmatrix}$$

$$+ \begin{bmatrix} 0.0001014 & 0.0042435 & -0.012438 & -0.0176903 \\ 0.0042435 & 0.1775416 & -0.520387 & -0.7401320 \\ -0.0124380 & -0.5203870 & 1.525291 & 2.1693794 \\ -0.0176903 & -0.7401320 & 2.169379 & 3.0854484 \end{bmatrix}$$

$$+ \begin{bmatrix} 0.0000127 & -0.0024196 & -0.0003011 & -0.0003686 \\ -0.0024196 & 0.4605315 & 0.0573193 & 0.0701564 \\ -0.0003011 & 0.0573193 & 0.0071342 & 0.0087319 \\ -0.0003686 & 0.0701564 & 0.0087319 & 0.0106875 \end{bmatrix}$$

$$= \begin{bmatrix} 220276.6577 & 915.295511 & 931.8589629 & 827.2877829 \\ 915.2955 & 4.443303 & 3.2895454 & 2.8521470 \\ 931.8590 & 3.289545 & 12.6120963 & 0.6310056 \\ 827.2878 & 2.852147 & 0.6310056 & 9.7721065 \end{bmatrix}$$
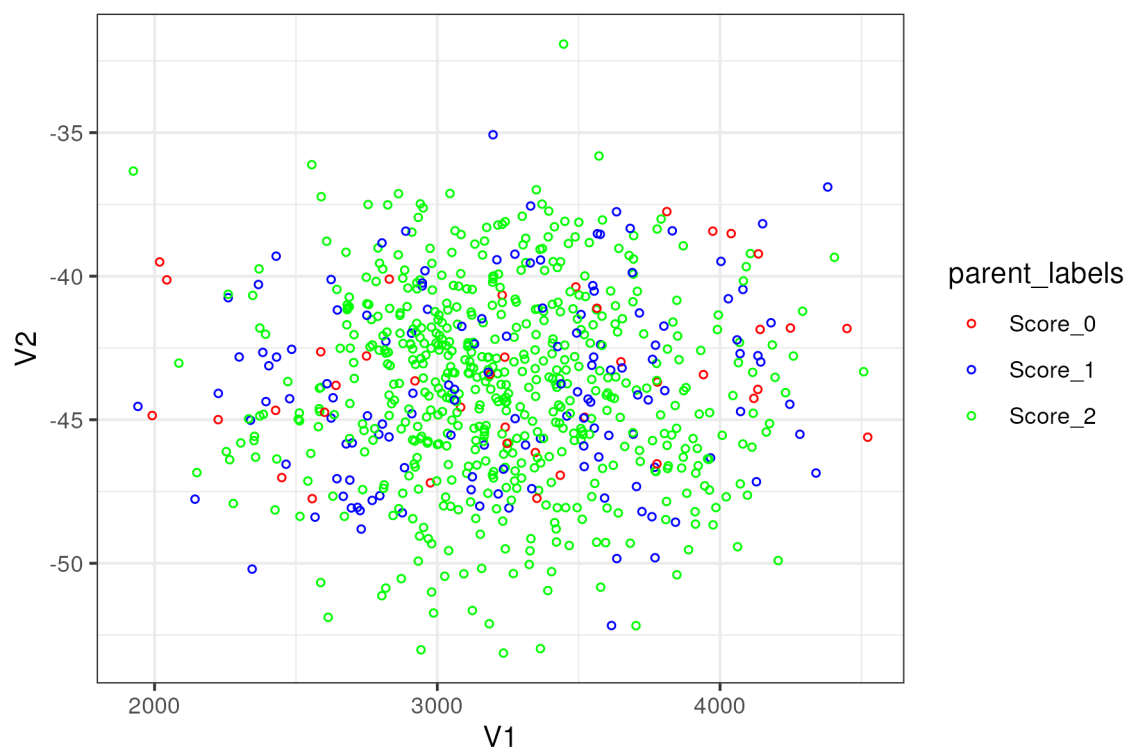
8. Transforme the data the according to $P^T X$. Plot scatter plots of the transformed data.

Figure 15: Transformed data