These notes may help you follow section 4.2.

The scores on $x_1$, $x_2$, $\ldots$, $x_p$ are standardized scores on the original variables (the original variables are the variables measured in the study). The quantity $p$ is the number of variables. The $n \times p$ matrix $X$ contains the standardized scores on the original variables. The first column of $X$ contains the standardized scores on $x_1$, the second contains the standardized scores on $x_2$, and so forth. The quantity $n$ is the number of participants.

It can be shown that $\frac{1}{(n-1)}X'X$ is the correlation matrix for the original variables. That is

$$R = \frac{1}{(n-1)}X'X.$$ 

You can demonstrate this to yourself, using the fact that the correlation between two variables $j$ and $k$ can be computed by using

$$r_{jk} = \frac{1}{(n-1)}(z_{1j}z_{1k} + \cdots + z_{nj}z_{nk})$$

where

- $z_{1j}$ is the z-score for the first person on the jth variable,
- $z_{1k}$ is the z-score for the first person on the kth variable,
- $z_{nj}$ is the z-score for the nth person on the jth variable,
• \( z_{nk} \) is the z-score for the \( n \)th person on the \( k \)th variable,

• \( \cdots \) means all the products in between \( z_{ij}z_{ik} \) and \( z_{ij}z_{nk} \).

The vectors \( \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p \) are vectors of weights as is the vector \( \mathbf{u} \). The matrix \( \mathbf{U} \) is formed by stacking the vectors \( \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p \) next to one another.

The relationship between \( \mathbf{Z} \) and \( \mathbf{X} \) is

\[
\mathbf{Z} = \mathbf{XU}
\]

indicating that \( \mathbf{Z} \) is a matrix of scores on linear combinations of the standardized original variables. The matrix \( \mathbf{Z} \) is a \( n \times p \) matrix. The first column of \( \mathbf{Z} \) contains the standardized scores on the first linear combination, the second contains the standardized scores on the second linear combination, and so forth. The vector \( \mathbf{z} \) is a \( n \times 1 \) vector of scores on one linear combination of the standardized variables.

You do not have to understand the details of equations 4.3 to 4.6. The author is just showing that if you want to select the weight vector \( \mathbf{u} \), so that the variance of the linear combination formed by using \( \mathbf{u} \) is as large as possible, the solution is to set \( \mathbf{u} \) equal to the first eigenvector of \( \mathbf{R} \).

When \( \mathbf{U} \) is chosen so that its columns contain the eigenvectors of \( \mathbf{R} \), \( \mathbf{Z} \) contains scores on the principal components.

The \( \lambda \)'s are the eigenvalues of \( \mathbf{R} \). The eigenvalues are ordered by size: \( \lambda_1 \) is the largest eigenvalue and \( \lambda_p \) is the smallest eigenvalue. Eigenvalues and eigenvectors come in pairs. The \( j \)th eigenvalue \( \lambda_j \) is the variance of the linear combination formed by using \( \mathbf{u}_j \), the \( j \)th eigenvector. The section "Variance
Accounted for by Principal Components" simply shows that $\lambda_1 + \cdots + \lambda_p = p$; the eigenvalues of $R$ must add to $p$, the number of variables.

Let $\mathbf{x}' = \begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}$ be the vector of standardized original variables. The relationship between $z_j$, the $j$th principal component, and $\mathbf{x}$ is $z_j = x'u_j$.

Since $\mathbf{x}$ contains standardized variables $\mathbf{x}' = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}$. By the relationship $z_j = x'u_j$, $z_j = 0$. This means that each principal component variable has mean zero and since its mean is zero it is a deviation score variable. Since each principal component is a deviation score variable, if we divide the $j$th principal component variable by its standard deviation, we get standardized principal components. We know that $\lambda_j$ is the variance of $z_j$, so $\sqrt{\lambda_j}$ is its standard deviation. As a result $\frac{z_j}{\sqrt{\lambda_j}}$ is a standardized principal component. On pages 100 and 101

$$D = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_p \end{bmatrix}$$

is a diagonal matrix containing the eigenvalues. The matrix $D^{1/2}$ is defined as a matrix such that $D = D^{1/2}D^{1/2}$ (this means $D$ is equal to $D^{1/2}$ multiplied by itself):

$$D^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \ddots \\ & & & \sqrt{\lambda_p} \end{bmatrix}.$$
The matrix $D^{-\frac{1}{2}}$ is the inverse of $D^{\frac{1}{2}}$:

$$D^{-\frac{1}{2}} = \begin{bmatrix}
\frac{1}{\sqrt{\lambda_1}} \\
\frac{1}{\sqrt{\lambda_2}} \\
\vdots \\
\frac{1}{\sqrt{\lambda_p}}
\end{bmatrix}.$$ 

The matrix $Z_s$ is a matrix of standardized principal component scores. The $j$th column contains the standardized scores on the $j$th principal component. To calculate $Z_s$ from $Z$, one needs to divide each score in the $j$th column of $Z$ by $\sqrt{\lambda_j}$. This is accomplished using the relationship

$$Z_s = ZD^{-\frac{1}{2}}.$$ 

Because $Z = XU$, $Z_s = XUD^{-\frac{1}{2}}$. Then postmultiplying both sides of $Z_s = XUD^{-\frac{1}{2}}$ by $D^{\frac{1}{2}}U'$

$$Z_sD^{\frac{1}{2}}U' = XUD^{-\frac{1}{2}}D^{\frac{1}{2}}U'.$$

Since $D^{-\frac{1}{2}}D^{\frac{1}{2}} = I$, $Z_sD^{-\frac{1}{2}}U' = XUU' = XU'$. Since $UU' = I$ ($U$ is orthogonal) $Z_sD^{-\frac{1}{2}}U' = X$. The expressions $Z_s = XUD^{-\frac{1}{2}}$ and $Z_sD^{-\frac{1}{2}}U' = X$ (which can also be expressed as $X = Z_sD^{-\frac{1}{2}}U'$) show that scores on $Z_s$ (the matrix of standardized scores on the principal components) can be computed from scores on $X$ (the matrix of standardized scores on the original variables) and scores on $X$ can be computed from $Z_s$. 