In PCA, you can write the following:

$$\begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix} = \underbrace{\begin{bmatrix} X_{1,1}-\overline{X_1} & X_{2,1}-\overline{X_2} & ... & X_{n,1}-\overline{X_n} \\ X_{1,2}-\overline{X_1} & X_{2,2}-\overline{X_2} & ... & X_{n,2}-\overline{X_n} \\ ... & ... & ... & ... \\ X_{1,t}-\overline{X_1} & X_{2,t}-\overline{X_2} & ... & X_{n,t}-\overline{X_n} \end{bmatrix}}_{\text{Call it matrix } \overline{X}} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} \quad (1)$$

The above equation is a consequence of the singular value decomposition theorem. The left-hand side of the equation above is the collection of n principal components time-series (as column vectors), in descending order of importance (from the right to the left). In the right-hand side, we have the product of two matrices: the first one, $\overline{X}$, is the raw data of candidate factors demeaned (i.e. each observation less the mean value of the factor $\overline{X_i}$, $i = 1, 2, ..., n$); the second matrix (n rows by n columns) encompasses the coefficients to transform the factor candidates into the principal components. For example, the first column has the coefficients to (linearly) transform the n factor candidates in each point t in time into the *first* principal component observation at time t. This latter matrix is exactly the matrix of eigenvectors of the variance-covariance matrix of the raw matrix of data (in descending order relative to their associated eigenvalues). The eigenvectors are not correlated and they are normalized to a unitary length! Both these conditions[1] assure the following relationship (provided that we do not have linear dependence on the raw data and hence the matrix of coefficients is invertible):

$$\begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix}^{-1} = \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix}^{T} = \begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} \quad (2)$$

where T means the transpose matrix.

Using this last identity and the definition of eigenvectors and eigenvalues, we can write the following development:

$$\begin{bmatrix} Cov_{1,1} & Cov_{2,1} & ... & Cov_{n,1} \\ Cov_{1,2} & Cov_{2,2} & ... & Cov_{n,2} \\ ... & ... & ... & ... \\ Cov_{1,n} & Cov_{2,n} & ... & Cov_{n,n} \end{bmatrix} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} = \begin{bmatrix} \lambda_1 E_{1,1} & \lambda_2 E_{2,1} & ... & \lambda_n E_{n,1} \\ \lambda_1 E_{1,2} & \lambda_2 E_{2,2} & ... & \lambda_n E_{n,2} \\ ... & ... & ... & ... \\ \lambda_1 E_{1,n} & \lambda_2 E_{2,n} & ... & \lambda_n E_{n,n} \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} Cov_{1,1} & Cov_{2,1} & ... & Cov_{n,1} \\ Cov_{1,2} & Cov_{2,2} & ... & Cov_{n,2} \\ ... & ... & ... & ... \\ Cov_{1,n} & Cov_{2,n} & ... & Cov_{n,n} \end{bmatrix} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} = \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} * \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \quad (4)$$

[1] Try to figure out why... It's not difficult!

$$\begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} * \begin{bmatrix} Cov_{1,1} & Cov_{2,1} & ... & Cov_{n,1} \\ Cov_{1,2} & Cov_{2,2} & ... & Cov_{n,2} \\ ... & ... & ... & ... \\ Cov_{1,n} & Cov_{2,n} & ... & Cov_{n,n} \end{bmatrix} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \tag{5}$$

The variance-covariance matrix can be written as $\frac{1}{t} * \overline{X}^T * \overline{X}$ (note that this is only possible because $\overline{X}$ is demeaned: this is precisely the reason we MUST have the demeaned version of the raw matrix of data X). We can then write:

$$\begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} * \frac{1}{t} * \overline{X}^T * \overline{X} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \tag{6}$$

But, using equation 1, we can easily see that we have the following equalities:

$$\begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} * \frac{1}{t} * \overline{X}^T = \frac{1}{t} * \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix}^T \tag{7}$$

$$\overline{X} * \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} = \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix} \tag{8}$$

Substituting both above expressions into 6 yields:

$$\frac{1}{t} * \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix}^T * \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \tag{9}$$

But hold on! What do we have on the left-hand side of this last equation? It's exactly the variance-covariance matrix of the principal components!!! So we are just saying that the principal components are not correlated (because the right-hand matrix has zeros outside the diagonal) and that the eigenvalues λ_i are in fact the variances of the principal components (remember that the diagonal in a variance-covariance matrix has the variances!). So, we have just demonstrated that the eigenvalues of the variance-covariance matrix of the raw data matrix are the variances of the principal components.

Let's calculate the total variance of the original data. Still from 1, we can develop the following:

$$\begin{bmatrix} X_{1,1}-\overline{X_1} & X_{2,1}-\overline{X_2} & ... & X_{n,1}-\overline{X_n} \\ X_{1,2}-\overline{X_1} & X_{2,2}-\overline{X_2} & ... & X_{n,2}-\overline{X_n} \\ ... & ... & ... & ... \\ X_{1,t}-\overline{X_1} & X_{2,t}-\overline{X_2} & ... & X_{n,t}-\overline{X_n} \end{bmatrix} = \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix} \begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} \tag{10}$$

But if:

$$Var\text{-}Covar(X) = \frac{1}{t}\begin{bmatrix} X_{1,1}-\overline{X_1} & X_{2,1}-\overline{X_2} & ... & X_{n,1}-\overline{X_n} \\ X_{1,2}-\overline{X_1} & X_{2,2}-\overline{X_2} & ... & X_{n,2}-\overline{X_n} \\ ... & ... & ... & ... \\ X_{1,t}-\overline{X_1} & X_{2,t}-\overline{X_2} & ... & X_{n,t}-\overline{X_n} \end{bmatrix} \begin{bmatrix} X_{1,1}-\overline{X_1} & X_{2,1}-\overline{X_2} & ... & X_{n,1}-\overline{X_n} \\ X_{1,2}-\overline{X_1} & X_{2,2}-\overline{X_2} & ... & X_{n,2}-\overline{X_n} \\ ... & ... & ... & ... \\ X_{1,t}-\overline{X_1} & X_{2,t}-\overline{X_2} & ... & X_{n,t}-\overline{X_n} \end{bmatrix} \tag{11}$$

We have:

$$Var\text{-}Covar(X) = \frac{1}{t}\begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix}^T \begin{bmatrix} PC_{1,1} & PC_{2,1} & ... & PC_{n,1} \\ PC_{1,2} & PC_{2,2} & ... & PC_{n,2} \\ ... & ... & ... & ... \\ PC_{1,t} & PC_{2,t} & ... & PC_{n,t} \end{bmatrix} \begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} \tag{12}$$

Now we use equation 9:

$$\begin{aligned} Var\text{-}Covar(X) &= \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \begin{bmatrix} E_{1,1} & E_{1,2} & ... & E_{1,n} \\ E_{2,1} & E_{2,2} & ... & E_{2,n} \\ ... & ... & ... & ... \\ E_{n,1} & E_{n,2} & ... & E_{n,n} \end{bmatrix} \\ &= \begin{bmatrix} E_{1,1} & E_{2,1} & ... & E_{n,1} \\ E_{1,2} & E_{2,2} & ... & E_{n,2} \\ ... & ... & ... & ... \\ E_{1,n} & E_{2,n} & ... & E_{n,n} \end{bmatrix} \begin{bmatrix} \lambda_1 E_{1,1} & \lambda_1 E_{1,2} & ... & \lambda_1 E_{1,n} \\ \lambda_2 E_{2,1} & \lambda_2 E_{2,2} & ... & \lambda_2 E_{2,n} \\ ... & ... & ... & ... \\ \lambda_n E_{n,1} & \lambda_n E_{n,2} & ... & \lambda_n E_{n,n} \end{bmatrix} \end{aligned} \tag{13}$$

such that we have the variance of each candidate factor X_i given by the diagonal terms:

$$Var(X_i) = \sum_{j=1}^{n} \lambda_j E_{j,i}^2 \tag{14}$$

We now define the total variance of the original data as the sum of all variances, such that we can develop the following (using the last equation):

$$TotalVar(X) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_j E_{j,i}^2 = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} E_{j,i}^2 \tag{15}$$

However, notice that $\sum_{i=1}^{n} E_{j,i}^2$ is the norm of the j^{th} eigenvector, which is 1!!! Then, the total variation of the original data is given by:

$$TotalVar(X) = \sum_{i=1}^{n} Var(X_i) = \sum_{j=1}^{n} \lambda_j \tag{16}$$

So, we have just proved that the sum of eigenvalues is exactly equal to the sum of variances of all original factors. This means that all principal components (which are non-correlated) have together the same total variance as the original factors (which are correlated). Thus, they 100% explain the variation on X. However, if we choose to work with the first k principal components, they will be able to explain just part of the total variance of X, given by:

$$\frac{ExplainedVariance(X)}{TotalVariance(X)} = \frac{\sum_{j=1}^{k} \lambda_j}{\sum_{j=1}^{n} \lambda_j} \tag{17}$$