corresponding variances are related to the *Eigenvalues* of the *covariance matrix*. Therefore, the problem of *PCA* reduces to the following *Eigenvalue problem*,

$$\boldsymbol{\Sigma}\mathbf{v} = \lambda\mathbf{v} \tag{12.2}$$

where $\mathbf{v} : \mathscr{R}^1 \mapsto \mathscr{R}^D$ is an *Eigenvector* associated with the feature vectors of interest, $\mathbf{x} : \mathscr{R}^1 \mapsto \mathscr{R}^D$, and $\boldsymbol{\Sigma} : \mathscr{R}^D \mapsto \mathscr{R}^D$ is the associated variance-covariance (*covariance*) matrix (see Section 6.10). λ is known as an *Eigenvalue* associated with *Eigenvector*, $\mathbf{v}$ and the matrix, $\boldsymbol{\Sigma}$.

We may rearrange Equation 12.2 in the following form,

$$(\lambda\mathbf{I} - \boldsymbol{\Sigma})\mathbf{v} = \mathbf{0} \;\forall\; \mathbf{v} \tag{12.3}$$

For any general matrix, $\boldsymbol{\Sigma}$, Equation 12.3 may only be true if the determinant of $\lambda\mathbf{I} - \boldsymbol{\Sigma}$ is zero. There will generally be D values of λ (the *Eigenvalues*) for which the *determinant* can become zero, given any general matrix. $\lambda_i, i \in \{1, 2, \cdots, D\}$ are said to be the solutions to the characteristic equation,

$$\begin{aligned} |\lambda\mathbf{I} - \boldsymbol{\Sigma}| &= \prod_{i=1}^{D}(\lambda - \lambda_i) \\ &= \mathbf{0} \end{aligned} \tag{12.4}$$

where $|\lambda\mathbf{I} - \boldsymbol{\Sigma}|$ denotes the determinant of $\lambda\mathbf{I} - \boldsymbol{\Sigma}$. Note that in general $\lambda_i \in \mathbb{C}$.

for every *Eigenvalue*, λ_i, Equation 12.2 must be true. There is a single *Eigenvector*, $\mathbf{v}_i$, associated with every λ_i which makes Equation 12.2 valid, namely,

$$\boldsymbol{\Sigma}\mathbf{v}_i = \lambda_i\mathbf{v}_i \;\forall\; i \in \{1, 2, \cdots, D\} \tag{12.5}$$

Solving Equation 12.2 for all i will produce a set of D *Eigenvectors* associated with the D Eigenvalues, λ_i. Let us construct a matrix, $\mathbf{V} : \mathscr{R}^D \mapsto \mathscr{R}^D$ whose columns are the *Eigenvectors*, $\mathbf{v}_i$, such that the first *Eigenvector* is associated with the largest *Eigenvalue* and the last one is associated with the smallest *Eigenvalue*. Also, note that Equation 12.5 may be multiplied by any constant from both sides. Therefore, the magnitude of the *Eigenvectors*, $\mathbf{v}_i$ is arbitrary. However, in here, we assume that all $\mathbf{v}_i$ have been normalized to have unit magnitude,

$$\|\mathbf{v}_i\|_{\mathscr{E}} = 1 \;\forall\; i \in \{1, 2, \cdots, D\} \tag{12.6}$$

This may be simply achieved by dividing the computed *Eigenvectors* by their corresponding Euclidean norms. After applying the normalization, we will have the following relation, based on Equation 12.2,

$$\boldsymbol{\Sigma}\mathbf{V} = \mathbf{V}\boldsymbol{\Lambda} \tag{12.7}$$