

with the same eigenvalue, so we can select one linear combination arbitrarily, and then choose the second to be orthogonal to the first (it can be shown that the degenerate eigenvectors are never linearly dependent). Hence, the eigenvectors can be chosen to be orthogonal, and by normalizing can be set to unit length. Because there are M eigenvalues, the corresponding M orthogonal eigenvectors form a complete set and so any M -dimensional vector can be expressed as a linear combination of the eigenvectors.

We can take the eigenvectors \mathbf{u}_i to be the columns of an $M \times M$ matrix \mathbf{U} , which from orthonormality satisfies

$$\mathbf{U}^T \mathbf{U} = \mathbf{I}. \quad (\text{A.37})$$

Such a matrix is said to be *orthogonal*. Interestingly, the rows of this matrix are also orthogonal, so that $\mathbf{U}\mathbf{U}^T = \mathbf{I}$. To show this, note that (A.37) implies $\mathbf{U}^T \mathbf{U} \mathbf{U}^{-1} = \mathbf{U}^{-1} = \mathbf{U}^T$ and so $\mathbf{U}\mathbf{U}^{-1} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$. Using (A.12), it also follows that $|\mathbf{U}| = 1$.

The eigenvector equation (A.29) can be expressed in terms of \mathbf{U} in the form

$$\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{\Lambda} \quad (\text{A.38})$$

where $\mathbf{\Lambda}$ is an $M \times M$ diagonal matrix whose diagonal elements are given by the eigenvalues λ_i .

If we consider a column vector \mathbf{x} that is transformed by an orthogonal matrix \mathbf{U} to give a new vector

$$\tilde{\mathbf{x}} = \mathbf{U}\mathbf{x} \quad (\text{A.39})$$

then the length of the vector is preserved because

$$\tilde{\mathbf{x}}^T \tilde{\mathbf{x}} = \mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{x} = \mathbf{x}^T \mathbf{x} \quad (\text{A.40})$$

and similarly the angle between any two such vectors is preserved because

$$\tilde{\mathbf{x}}^T \tilde{\mathbf{y}} = \mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{y} = \mathbf{x}^T \mathbf{y}. \quad (\text{A.41})$$

Thus, multiplication by \mathbf{U} can be interpreted as a rigid rotation of the coordinate system.

From (A.38), it follows that

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{\Lambda} \quad (\text{A.42})$$

and because $\mathbf{\Lambda}$ is a diagonal matrix, we say that the matrix \mathbf{A} is *diagonalized* by the matrix \mathbf{U} . If we left-multiply by \mathbf{U} and right-multiply by \mathbf{U}^T , we obtain

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T. \quad (\text{A.43})$$

Taking the inverse of this equation and using (A.3) together with $\mathbf{U}^{-1} = \mathbf{U}^T$, we have

$$\mathbf{A}^{-1} = \mathbf{U} \mathbf{\Lambda}^{-1} \mathbf{U}^T. \quad (\text{A.44})$$