4.3 Optimal Linear Combinations of Bands

Several useful algorithmms for reducing dimensionality involve finding optimal linear combinations of bands. The meaning of *optimal* is encoded in an *objective function*. The process of minimizing or maximizing the objective function leads to an algorithm for reducing dimensionality.

4.3.1 Principal Components Analysis (PCA) and Transform (PCT)

Reducing dimensionality of spectra via PCA produces vector representatives of spectra that have two important properties. The first is that the reduced dimensionality vectors have uncorrelated components. The second is that the reduced dimensionality vectors of any dimension $D < B$ have the highest percentage of the total variance of any linear transformation from dimension B to dimension D. These concepts will be made precise below.

Principal Components Analysis refers to the act of analyzing data using a linear change-of-basis transform, referred to as a Principal Component Transform. The spectral data $\mathbf{X}$ are considered to be samples of a random vector $\boldsymbol{x}$. There is a *true* PCT associated with $\boldsymbol{x}$ and a sample PCT associated with the sample $\mathbf{X}$, which can be referred to as $\text{PCT}_{\boldsymbol{x}}$ and $\text{PCT}_{\mathbf{X}}$, respectively. However, in practice this distinction is rarely made and $\text{PCT}_{\mathbf{X}}$ is what is normally used and so we will only discuss it. It is important to keep in mind that $\text{PCT}_{\mathbf{X}}$ depends on the sample $\mathbf{X}$.

A PCT transforms the standard basis to a basis of eigenvectors of a covariance matrix. More precisely, let $\bar{\mu}_X$ and $\bar{\mathbf{C}}_X$ denote the sample mean and sample covariance of the sample spectra in $\mathbf{X}$. Recall that $\bar{\mathbf{C}}_X = V^t \Lambda V$ where $V^t V = V V^t = I$ and Λ is a diagonal matrix with non-negative values. The PCT of a spectrum $\mathbf{x}$ with respect to the sample $\mathbf{X}$ of the random vector is defined to be

$$\mathbf{y} = V^t \left(\mathbf{x} - \bar{\mu}_X \right).$$

PCA is based on using the entities involved in the PCT to analyze spectral data. It can be used to mitigate the effects of the *Curse of Dimensionality*.

Example. Two-dimensional data. Suppose $\mathbf{x}$ is a 2D Gaussian random variable with mean, $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ given by

$$\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 0.7562 & 0.4222 \\ 0.4222 & 0.2687 \end{bmatrix}$$

The eigenvalues of $\boldsymbol{\Sigma}$ are $\lambda_1 = 1$ and $\lambda_2 = 0.025$. A diagonalizing matrix is

$$V = \begin{bmatrix} 0.8660 & -0.5000 \\ 0.5000 & 0.8660 \end{bmatrix}$$

since $V^t \boldsymbol{\Sigma} V diag\,(1, 0.025)$. A data set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ was created by generating 1000 pseudo-random vectors from the multivariate Gaussian above and had sample mean and covariance:

$$\boldsymbol{\mu}_{data} = \begin{bmatrix} 3.00 \\ 2.01 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_{data} = \begin{bmatrix} 0.7376 & 0.4092 \\ 0.4092 & 0.2592 \end{bmatrix}$$

.

Diagonalizing $\mathbf{\Sigma}_{data}$ yielded eigenvalues $\lambda_1 = 0.972$ and $\lambda_2 = 0.025$ and eigenvectors

$$V_{data} = \begin{bmatrix} -0.8674 & -0.4977 \\ -0.4977 & 0.8674 \end{bmatrix}$$

.

Note that the eigenvectors corresponding to λ_1 and λ_2, respectively, define the direction of the major and minor axes. The major and minor axes are depicted in Figure 4.1.

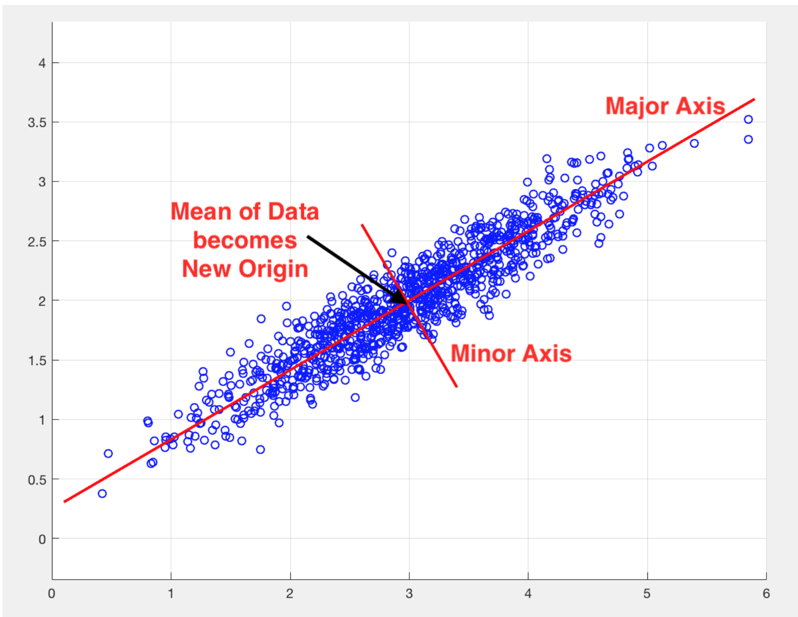

Figure 4.1 Principal Axes of data set generated by Gaussian distribution. The direction of the major axis is given by the eigenvector associated with the largest eigenvalue. The direction of the minor axis is given by the eigenvector associated with the smallest eigenvalue.

The result of transforming the data set from the original domain to the PCA domain is shown in Figure 4.2

Spectral Example. The image in Figure 4.3 is an RGB version of an AVIRIS image with 178-dimensional spectral pixels (the water bands have been removed). The Principal Component Dimensionality Reduction is shown in Figure 4.4.

Mathematical Derivation. some of the mathematical properties of the entities used in PCA. Notice that the mean of $\mathbf{y}$ is $\mathbf{0}$ since
$\mathbb{E}_X[\mathbf{y}] = \mathbb{E}_X[V^t(\mathbf{x} - \bar{\mu}_X)] = V^t(\mathbb{E}_X[\mathbf{x}] - \bar{\mu}_X) = V^t\mathbf{0} = \mathbf{0}$.

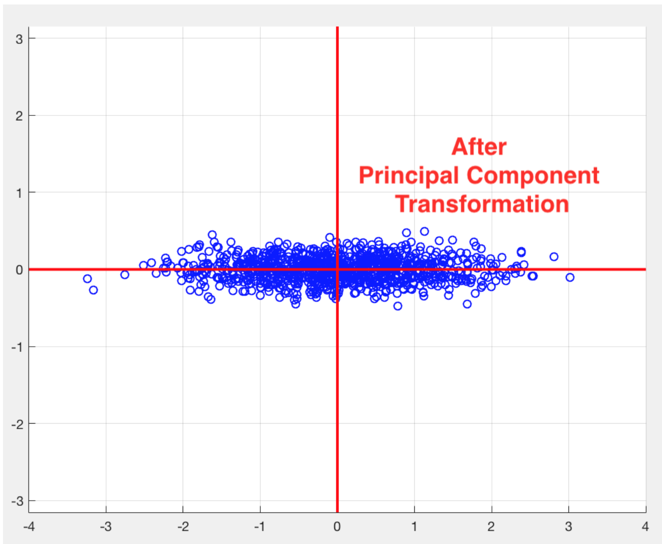

Figure 4.2 PCT of data shown in Figure 4.1

Figure 4.3 An RGB version of a 178D AVIRIS image.

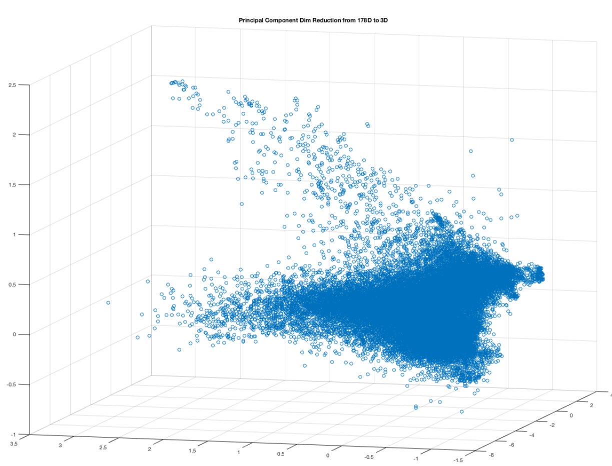

Figure 4.4 A scatterplot of the first 3 Principal Components of an 178D AVIRIS image.