Topic: Decorrelating and then Whitening data

Extra notes for MAS622J/1.126J by Rosalind W. Picard

1. Let \( \mathbf{x} \) be a vector of zero-mean data. Form its covariance matrix,

\[
\Sigma = E(\mathbf{x}\mathbf{x}^T)
\]

If the data points in \( \mathbf{x} \) are correlated, then their covariance, \( \Sigma \), will NOT be a diagonal matrix.

2. In order to decorrelate the data, we need to transform it so that the transformed data will have a diagonal covariance matrix. This transform can be found by solving the eigenvalue problem. We find the eigenvectors and associated eigenvalues of \( \Sigma \) by solving

\[
\Sigma \Phi = \Phi \Lambda
\]

\( \Lambda \) is a diagonal matrix having the eigenvalues as its diagonal elements. The matrix \( \Phi \) thus diagonalizes the covariance matrix of \( \mathbf{x} \). The columns of \( \Phi \) are the eigenvectors of the covariance matrix.

We can also write the diagonalized covariance as:

\[
\Phi^T \Sigma \Phi = \Lambda
\]

(1)

If we wish to apply this diagonalizing transform to a single vector of data we just form:

\[
\mathbf{y} = \Phi^T \mathbf{x}
\]

(2)

Thus, the data \( \mathbf{y} \) has been decorrelated: its covariance, \( E[\mathbf{y}\mathbf{y}^T] \) is now a diagonal matrix, \( \Lambda \).

3. The diagonal elements (eigenvalues) in \( \Lambda \) may be the same or different. If we make them all the same, then this is called whitening the data. Since each eigenvalue determines the length of its associated eigenvector, the covariance will correspond to an ellipse when the data is not whitened, and to a sphere (having all dimensions the same length, or uniform) when the data is whitened. Whitening is easy:

\[
\Lambda^{-1/2} \Lambda \Lambda^{-1/2} = I
\]

Equivalently, substituting in (1), we write:

\[
\Lambda^{-1/2} \Phi^T \Sigma \Phi \Lambda^{-1/2} = I
\]
Thus, to apply this whitening transform to $y$ we simply multiply it by this scale factor, obtaining the whitened data $w$:

$$w = \Lambda^{-1/2} y = \Lambda^{-1/2} \Phi^T x$$

(3)

Now the covariance of $w$ is not only diagonal, but also uniform (white), since the covariance of $w$, $E(ww^T) = I$:

$$E(\Lambda^{-1/2} \Phi^T xx^T \Phi \Lambda^{-1/2}) = I$$

In DHS, the diagonalizing transform applied to $x$ is denoted $A$ and the whitening transform is represented by $A_w$. These map onto the notation above as follows:

$$y = A^T x, \quad A = \Phi$$

$$w = A_w^T x, \quad A_w = \Phi \Lambda^{-1/2}$$