Adam Berenzweig
EE 4830
HW-3

1. a) We're looking for a matrix of the form $A_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ to transform $\mathbf{u}$ into $\mathbf{v} = A_\theta \mathbf{u}$ such that v_0 and v_1 are uncorrelated. In other words, the covariance matrix of $\mathbf{v}$ is diagonal, $R_v = \left(\begin{smallmatrix} \lambda_0 & 0 \\ 0 & \lambda_1 \end{smallmatrix}\right)$. The transformation that satisfies this property is the KLT transform, $\mathbf{v} = \Phi^{*T}\mathbf{u}$, where $R_u = \Phi\Lambda\Phi^{*T}$ (the diagonalization of R_u). The KLT has the desired property because

$$R_v = E[\mathbf{v}\mathbf{v}^{*T}] = E[\Phi^{*T}\mathbf{u}\mathbf{u}^{*T}\Phi] = \Phi^{*T}R_u\Phi = \Lambda$$

So we determine Φ by finding the eigenvalues and eigenvectors of $R_u = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ as follows:

$$\begin{vmatrix} 1-\lambda & \rho \\ \rho & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - \rho^2 = \lambda^2 - 2\lambda + (1-\rho^2)$$

Solving the characteristic equation gives the eigenvalues $\lambda = 1 \pm \rho$. The eigenvectors are found by solving $(R_u - \lambda I)x = 0$ for each eigenvalue. The solutions can by multiplied by any scalar, so we choose to make them of length one so that Φ is orthonormal. So we have

$$\Phi = \begin{bmatrix} x_0 & x_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Now if $A_\theta = \Phi^{*T}$, then we have

$$\cos\theta = \frac{1}{\sqrt{2}}, \text{ and } \sin\theta = \frac{1}{\sqrt{2}}$$

So $\theta = \frac{\pi}{4}$. In fact, any multiple $(8n+1)\frac{\pi}{4}$ will work. Also, I believe that the KLT isn't the only way to decorrelate $\mathbf{v}$, but it is the way that also leads to the maximum average energy per dimension (see part b).

b) Another property of the KLT is that it maximizes the average energy packed into the first N dimensions. All unitary transforms preserve energy (they can be thought of as rotations), but since the KLT orders the eigenvalues, as much energy as possible is put into the first N coefficients. Therefore, the KLT also maximizes the average energy in v_0. This all makes sense if you notice that the matrix A_θ is a rotation of negative θ radians. The eigenvectors of R_u were pointing in the direction of $\begin{bmatrix} 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 1 \end{bmatrix}$, in other words at angles of $\frac{\pi}{4}$ and $-\frac{\pi}{4}$. So rotating by $-\frac{\pi}{4}$ radians brings these axes to the coordinate axes v_0 and v_1, which lets most of the energy be along the principle axes v_0. Also, as seen in part a, it decorrelates v_0 and v_1.