Change of Basis

Consider a linear transform, $\mathbf{P}_\mathcal{B}$, and its inverse, $\mathbf{P}_\mathcal{B}^{-1}$, which map a vector back and forth between its representation in the standard basis and its representation in the basis, $\mathcal{B}$:

$$\begin{align*}
\mathbf{P}_\mathcal{B} & \quad \mathbf{u} \quad \mathbf{P}_\mathcal{B}^{-1} \\
\mathbf{u} & \quad \leftrightarrow \\
\mathbf{P}_\mathcal{B}^{-1} &
\end{align*}$$

\[ [\mathbf{u}]_\mathcal{B} \]
Change of Basis (contd).

Let $\mathcal{B}$ consist of $N$ basis vectors, $b_1 \ldots b_N$. Since $[u]_{\mathcal{B}}$ is the representation of $u$ in $\mathcal{B}$, it follows that

$$u = ([u]_{\mathcal{B}})_1 b_1 + ([u]_{\mathcal{B}})_2 b_2 + \ldots ([u]_{\mathcal{B}})_N b_N.$$ 

But this is just the matrix vector product

$$u = B [u]_{\mathcal{B}}$$

where

$$B = \begin{bmatrix} b_1 & b_2 & \ldots & b_N \end{bmatrix}.$$ 

We see that $P_B = B^{-1}$ and $P_B^{-1} = B$. 
Similarity Transforms

Now consider a linear transformation represented in the standard basis by the matrix \( \mathbf{A} \). We seek \([\mathbf{A}]_{\mathcal{B}}\), i.e., the representation of the corresponding linear transformation in the basis \( \mathcal{B} \):

\[
\begin{align*}
\mathbf{u} & \xrightarrow{\mathbf{A}} \mathbf{Au} \\
\uparrow \mathcal{B} & \quad \downarrow \mathcal{B}^{-1} \\
[u]_{\mathcal{B}} & \xrightarrow{[\mathbf{A}]_{\mathcal{B}}} [\mathbf{Au}]_{\mathcal{B}}
\end{align*}
\]

The matrix we seek maps \([\mathbf{u}]_{\mathcal{B}}\) into \([\mathbf{Au}]_{\mathcal{B}}\). From the above diagram, we see that this matrix is the composition of \(\mathcal{B}\), \(\mathbf{A}\), and \(\mathcal{B}^{-1}\):

\[
[\mathbf{A}]_{\mathcal{B}} = \mathcal{B}^{-1}\mathbf{A}\mathcal{B}.
\]

We say that \(\mathbf{A}\) and \([\mathbf{A}]_{\mathcal{B}}\) are related by a similarity transform.
Diag. of Symmetric Matrices

Because of linearity, one might expect that a transformation will have an especially simple representation in the basis of its eigenvectors, $\lambda$. Let $A$ be its representation in the standard basis and let the columns of $X$ be the eigenvectors of $A$. Then $X$ and $X^T = X^{-1}$ take us back and forth between the standard basis and $\lambda$:

$$\begin{align*}
    X^T \\
    \mathbf{u} \leftrightarrow [\mathbf{u}]_{\lambda} \\
    X
\end{align*}$$
Diag. of Symmetric Matrices (contd.)

The matrix we seek maps $[u]_x$ into $[Au]_x$:

$$
\begin{array}{c}
\begin{array}{ccc}
  u & \xrightarrow{A} & Au \\
  \uparrow X & & \downarrow X^T \\
 [u]_x & \xrightarrow{[A]_x} & [Au]_x \\
\end{array}
\end{array}
$$

From the above diagram, we see that this matrix is the composition of $X$, $A$, and $X^T$:

$$
\Lambda = X^TAX.
$$

We observe that $\Lambda$ is diagonal with $\Lambda_{ii} = \lambda_i$, the eigenvalue of $A$ associated with eigenvector, $x_i$. 
Spectral Thm. for Sym. Matrices

Any symmetric $N \times N$ matrix, $A$, with $N$ distinct eigenvalues, can be factored as follows:

$$A = X\Lambda X^T$$

where $\Lambda$ is $N \times N$ and diagonal, $X$ and $X^T$ are $N \times N$ matrices, and the $i$-th column of $X$ (equal to the $i$-th row of $X^T$) is an eigenvector of $A$:

$$\lambda_i x_i = Ax_i$$

with eigenvalue $\Lambda_{ii} = \lambda_i$. Note that $x_i$ is orthogonal to $x_j$ when $i \neq j$:

$$(XX^T)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $XX^T = I$. Consequently, $X^T = X^{-1}$. 
Spectral Thm. for Sym. Matrices (contd).

Using the definition of matrix product and the fact that $\Lambda$ is diagonal, we can write $A = X\Lambda X^T$ as

$$(A)_{ij} = \sum_{k=1}^{N} (X)_{ik} \Lambda_{kk} (X^T)_{kj}.$$ 

Since $X = [x_1 | x_2 | \ldots | x_N]$ and $\Lambda_{kk} = \lambda_k$

$$(A)_{ij} = \sum_{k=1}^{N} (x_k)_i \lambda_k (x_k)_j$$

$$= \sum_{k=1}^{N} (\lambda_k x_k x_k^T)_{ij}$$

$$A = \sum_{k=1}^{N} \lambda_k x_k x_k^T$$

where $\lambda_k x_k = Ax_k$. 
Spectral Thm. for Sym. Matrices (contd).

The *spectral factorization* of $A$ is:

$$A = \lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T + \cdots + \lambda_N x_N x_N^T.$$  

Note that each $\lambda_n x_n x_n^T$ is a rank one matrix. Let $A_i = \lambda_i x_i x_i^T$. Now, because $x_i^T x_i = 1$:

$$\lambda_i x_i = (\lambda_i x_i x_i^T) x_i = A_i x_i$$

*i.e.*, $x_i$ is the only eigenvector of $A_i$ and its only eigenvalue is $\lambda_i$. 
Diag. of Non-symmetric Matrices

The situation is more complex when the transformation is represented by a non-symmetric matrix, $P$. Let the columns of $X$ be $P$’s right eigenvectors and the rows of $Y^T$ be its left eigenvectors. Then $X$ and $Y^T = X^{-1}$ take us back and forth between the standard basis and $x$:

$$Y^T$$

$$u \leftrightarrow [u]_x.$$  

$X$
The matrix we seek maps $[u]_x$ into $[Pu]_x$:

$$
\begin{align*}
    u & \xrightarrow{P} Pu \\
    \uparrow X & \quad \downarrow Y^T \\
    [u]_x & \xrightarrow{\Lambda} [Pu]_x
\end{align*}
$$

From the above diagram, we see that this matrix is the composition of $X$, $P$, and $Y^T$:

$$
\Lambda = Y^T PX
$$

We observe that $\Lambda$ is diagonal with $\Lambda_{ii} = \lambda_i$, the eigenvalue of $P$ associated with right eigenvector, $x_i$, and left eigenvector, $y_i$. 

Spectral Theorem

Any $N \times N$ matrix, $\mathbf{P}$, with $N$ distinct eigenvalues, can be factored as follows:

$$\mathbf{P} = \mathbf{X}\Lambda\mathbf{Y}^T$$

where $\Lambda$ is $N \times N$ and diagonal, $\mathbf{X}$ and $\mathbf{Y}^T$ are $N \times N$ matrices, and the $i$-th column of $\mathbf{X}$ is a right eigenvector of $\mathbf{P}$:

$$\lambda_i \mathbf{x}_i = \mathbf{P}\mathbf{x}_i$$

with eigenvalue $\Lambda_{ii} = \lambda_i$ and the $i$-th row of $\mathbf{Y}^T$ is a left eigenvector of $\mathbf{P}$:

$$\lambda_i \mathbf{y}_i^T = \mathbf{y}_i^T\mathbf{P}$$

with the same eigenvalue.
Spectral Theorem (contd.)

Note that $x_i$ is orthogonal to $y_j$ when $i \neq j$:

$$(XY^T)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $XY^T = I$. Consequently,

$$Y^T = X^{-1}.$$
Spectral Theorem (contd).

The *spectral factorization* of $P$ is:

$$P = \lambda_1 x_1 y_1^T + \lambda_2 x_2 y_2^T + \cdots + \lambda_N x_N y_N^T.$$  

Note that each $\lambda_n x_n y_n^T$ is a rank one matrix. Let $P_i = \lambda_i x_i y_i^T$. Now, because $y_i^T x_i = 1$:

$$\lambda_i x_i = (\lambda_i x_i y_i^T) x_i = P_i x_i$$

and

$$\lambda_i y_i^T = y_i^T (\lambda_i x_i y_i^T) = y_i^T P_i$$

*i.e.*, $x_i$ and $y_i$ are the sole right and left eigenvectors of $P_i$. The only eigenvalue is $\lambda_i$. 
Stochastic Matrices

If $P$ is stochastic, then

$$1 = \sum_{i} P_{ij}.$$  

Let $y^{T} = [1 \ 1 \ \ldots \ \ 1]$. It follows that

$$(y^{T})_{j} = \sum_{i} (y^{T})_{i} P_{ij}.$$  

Consequently, $y^{T}$ is a left eigenvector with unit eigenvalue of every stochastic matrix:

$$y^{T} = y^{T}P.$$
Stochastic Matrices (contd.)

What is the representation of an arbitrary distribution, \( z \), in the basis of eigenvectors of an arbitrary stochastic matrix, \( P \)?

\[
z = c_1 x_1 + c_2 x_1 + \cdots + c_N x_N = Xc.
\]

Solving for \( c \):

\[
c = X^{-1} z = Y^T z.
\]

Since \( y_1^T = [1 \ 1 \ \ldots \ 1] \), it follows that:

\[
c_1 = \sum_j Y_{1j}^T z_j = \sum_j z_j = 1
\]

which is independent of the specific \( z \) and \( P \)!