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, $\mathcal{B}$:

$$\mathbf{u} \underset{\mathbf{P}_{\mathcal{B}}^{-1}}{\overset{\mathbf{P}_{\mathcal{B}}}{\rightleftarrows}} [\mathbf{u}]_{\mathcal{B}} \ .$$

Change of Basis (contd).

Let $\mathcal{B}$ consist of N basis vectors, $\mathbf{b}_1 \ldots \mathbf{b}_N$. Since $[\mathbf{u}]_{\mathcal{B}}$ is the representation of $\mathbf{u}$ in $\mathcal{B}$, it follows that

$$\mathbf{u} = ([\mathbf{u}]_{\mathcal{B}})_1 \mathbf{b}_1 + ([\mathbf{u}]_{\mathcal{B}})_2 \mathbf{b}_2 + \ldots ([\mathbf{u}]_{\mathcal{B}})_N \mathbf{b}_N.$$

But this is just the matrix vector product

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

where

$$\mathbf{B} = \left[\, \mathbf{b}_1 \middle| \mathbf{b}_2 \middle| \ldots \middle| \mathbf{b}_N \,\right].$$

We see that $\mathbf{P}_{\mathcal{B}} = \mathbf{B}^{-1}$ and $\mathbf{P}_{\mathcal{B}}^{-1} = \mathbf{B}$.

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 $\mathcal{B}$:

$$\begin{array}{ccc} \mathbf{u} & \xrightarrow{\mathbf{A}} & \mathbf{Au} \\ \uparrow \mathbf{B} & & \downarrow \mathbf{B}^{-1} \\ [\mathbf{u}]_{\mathcal{B}} & \xrightarrow{[\mathbf{A}]_{\mathcal{B}}} & [\mathbf{Au}]_{\mathcal{B}} \end{array}$$

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 $\mathbf{B}$, $\mathbf{A}$, and $\mathbf{B}^{-1}$:

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

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, x . Let $\mathbf{A}$ be its representation in the standard basis and let the columns of $\mathbf{X}$ be the eigenvectors of $\mathbf{A}$. Then $\mathbf{X}$ and $\mathbf{X}^{\mathrm{T}} = \mathbf{X}^{-1}$ take us back and forth between the standard basis and x :

$$\mathbf{u} \underset{\mathbf{X}}{\overset{\mathbf{X}^{\mathrm{T}}}{\rightleftarrows}} [\mathbf{u}]_{x} \ .$$

Diag. of Symmetric Matrices (contd.)

The matrix we seek maps $[\mathbf{u}]_x$ into $[\mathbf{Au}]_x$:

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

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

$$\Lambda = \mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}.$$

We observe that Λ is diagonal with $\Lambda_{ii} = \lambda_i$, the eigenvalue of $\mathbf{A}$ associated with eigenvector, $\mathbf{x}_i$.

Spectral Thm. for Sym. Matrices

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

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{\mathrm{T}}$$

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

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

with eigenvalue $\Lambda_{ii} = \lambda_i$. Note that $\mathbf{x}_i$ is orthogonal to $\mathbf{x}_j$ when $i \neq j$:

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

In other words, $\mathbf{X}\mathbf{X}^{\mathrm{T}} = \mathbf{I}$. Consequently,

$$\mathbf{X}^{\mathrm{T}} = \mathbf{X}^{-1}.$$

Spectral Thm. for Sym. Matrices (contd).

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

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

Since $\mathbf{X} = \left[\, \mathbf{x}_1 \middle| \mathbf{x}_2 \middle| \ldots \middle| \mathbf{x}_N \,\right]$ and $\Lambda_{kk} = \lambda_k$

$$\begin{aligned}(\mathbf{A})_{ij} &= \sum_{k=1}^{N} (\mathbf{x}_k)_i \lambda_k (\mathbf{x}_k)_j \\ &= \sum_{k=1}^{N} \left(\lambda_k \mathbf{x}_k \mathbf{x}_k^{\mathrm{T}}\right)_{ij} \\ \mathbf{A} &= \sum_{k=1}^{N} \lambda_k \mathbf{x}_k \mathbf{x}_k^{\mathrm{T}}\end{aligned}$$

where $\lambda_k \mathbf{x}_k = \mathbf{A}\mathbf{x}_k$.

Spectral Thm. for Sym. Matrices (contd).

The *spectral factorization* of $\mathbf{A}$ is:

$$\mathbf{A} = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^{\mathrm{T}} + \lambda_2 \mathbf{x}_2 \mathbf{x}_2^{\mathrm{T}} + \cdots + \lambda_N \mathbf{x}_N \mathbf{x}_N^{\mathrm{T}}.$$

Note that each $\lambda_n \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}}$ is a rank one matrix. Let $\mathbf{A}_i = \lambda_i \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}}$. Now, because $\mathbf{x}_i^{\mathrm{T}} \mathbf{x}_i = 1$:

$$\begin{aligned} \lambda_i \mathbf{x}_i &= \left(\lambda_i \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}}\right) \mathbf{x}_i \\ &= \mathbf{A}_i \mathbf{x}_i \end{aligned}$$

i.e., $\mathbf{x}_i$ is the only eigenvector of $\mathbf{A}_i$ and its only eigenvalue is λ_i.

Diag. of Non-symmetric Matrices

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

$$\mathbf{u} \underset{\mathbf{X}}{\overset{\mathbf{Y}^{\mathrm{T}}}{\rightleftarrows}} [\mathbf{u}]_x \ .$$

Diag. of Non-symmetric Matrices (contd.)

The matrix we seek maps $[\mathbf{u}]_x$ into $[\mathbf{Pu}]_x$:

$$\begin{array}{ccc} \mathbf{u} & \xrightarrow{\mathbf{P}} & \mathbf{Pu} \\ \uparrow \mathbf{X} & & \downarrow \mathbf{Y}^{\mathrm{T}} \\ [\mathbf{u}]_x & \xrightarrow{\Lambda} & [\mathbf{Pu}]_x \end{array}$$

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

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

We observe that Λ is diagonal with $\Lambda_{ii} = \lambda_i$, the eigenvalue of $\mathbf{P}$ associated with right eigenvector, $\mathbf{x}_i$, and left eigenvector, $\mathbf{y}_i$.

Spectral Theorem

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

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

where Λ is $N \times N$ and diagonal, $\mathbf{X}$ and $\mathbf{Y}^{\mathrm{T}}$ 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 $\mathbf{Y}^{\mathrm{T}}$ is a *left eigenvector* of $\mathbf{P}$:

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

with the same eigenvalue.

Spectral Theorem (contd.)

Note that $\mathbf{x}_i$ is orthogonal to $\mathbf{y}_j$ when $i \neq j$:

$$\left(\mathbf{X}\mathbf{Y}^{\mathrm{T}}\right)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $\mathbf{X}\mathbf{Y}^{\mathrm{T}} = \mathbf{I}$. Consequently,

$$\mathbf{Y}^{\mathrm{T}} = \mathbf{X}^{-1}.$$

<u>Spectral Theorem (contd).</u>

The *spectral factorization* of $\mathbf{P}$ is:

$$\mathbf{P} = \lambda_1 \mathbf{x}_1 \mathbf{y}_1^{\mathrm{T}} + \lambda_2 \mathbf{x}_2 \mathbf{y}_2^{\mathrm{T}} + \cdots + \lambda_N \mathbf{x}_N \mathbf{y}_N^{\mathrm{T}}.$$

Note that each $\lambda_n \mathbf{x}_n \mathbf{y}_n^{\mathrm{T}}$ is a rank one matrix. Let $\mathbf{P}_i = \lambda_i \mathbf{x}_i \mathbf{y}_i^{\mathrm{T}}$. Now, because $\mathbf{y}_i^{\mathrm{T}} \mathbf{x}_i = 1$:

$$\begin{aligned} \lambda_i \mathbf{x}_i &= \left(\lambda_i \mathbf{x}_i \mathbf{y}_i^{\mathrm{T}}\right) \mathbf{x}_i \\ &= \mathbf{P}_i \mathbf{x}_i \end{aligned}$$

and

$$\begin{aligned} \lambda_i \mathbf{y}_i^{\mathrm{T}} &= \mathbf{y}_i^{\mathrm{T}} \left(\lambda_i \mathbf{x}_i \mathbf{y}_i^{\mathrm{T}}\right) \\ &= \mathbf{y}_i^{\mathrm{T}} \mathbf{P}_i \end{aligned}$$

i.e., $\mathbf{x}_i$ and $\mathbf{y}_i$ are the sole right and left eigenvectors of $\mathbf{P}_i$. The only eigenvalue is λ_i.

Stochastic Matrices

If **P** is stochastic, then

$$1 = \sum_i P_{ij}.$$

Let $\mathbf{y}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$. It follows that

$$\left(\mathbf{y}^{\mathrm{T}}\right)_j = \sum_i \left(\mathbf{y}^{\mathrm{T}}\right)_i P_{ij}.$$

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

$$\mathbf{y}^{\mathrm{T}} = \mathbf{y}^{\mathrm{T}}\mathbf{P}.$$

Stochastic Matrices (contd.)

What is the representation of an arbitrary distribution, $\mathbf{z}$, in the basis of eigenvectors of an arbitrary stochastic matrix, $\mathbf{P}$?

$$\mathbf{z} = c_1\mathbf{x}_1 + c_2\mathbf{x}_1 + \cdots + c_N\mathbf{x}_N = \mathbf{X}\mathbf{c}.$$

Solving for $\mathbf{c}$:

$$\mathbf{c} = \mathbf{X}^{-1}\mathbf{z} = \mathbf{Y}^{\mathrm{T}}\mathbf{z}.$$

Since $\mathbf{y}_1^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$, it follows that:

$$c_1 = \sum_j Y_{1j}^{\mathrm{T}} z_j = \sum_j z_j = 1$$

which is independent of the specific $\mathbf{z}$ and $\mathbf{P}$!